

**SEMINARIO DE INVESTIGACIÓN EN METODOS DE DIAGNOSTICO DE
FALLAS DE SISTEMAS COMPLEJOS**

MANUEL ORLANDO CARREÑO MARTINEZ

LUZ AMPARO ALVAREZ TORRES

SERGIO ARMANDO MEJIA FAJARDO

RUBEN DARIO DUARTE ANAYA

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICOMECÁNICAS
ESCUELA DE INGENIERÍA MECÁNICA
BUCARAMANGA**

2011

**SEMINARIO DE INVESTIGACIÓN EN METODOS DE DIAGNOSTICO DE
FALLAS DE SISTEMAS COMPLEJOS**

MANUEL ORLANDO CARREÑO MARTINEZ

LUZ AMPARO ALVAREZ TORRES

SERGIO ARMANDO MEJIA FAJARDO

RUBEN DARIO DUARTE ANAYA

Trabajo de grado para optar al título de Ingeniero(a) Mecánico(a)

Director

Ing. Rómulo Niño Delgado

UNIVERSIDAD INDUSTRIAL DE SANTANDER

FACULTAD DE INGENIERÍAS FÍSICOMECÁNICAS

ESCUELA DE INGENIERÍA MECÁNICA

BUCARAMANGA

2011

AGRADECIMIENTOS

Le dedico este logro a mi Padre celestial, por sus bendiciones sin fin y su compañía; A mi madre Lilian Zoraida Anaya Bohórquez por su amor y apoyo incondicional; A mi hermano Gilberto Andrés Duarte Anaya por su ayuda, compañía y aliento; y a mi padre Gilberto Duarte Torres que se encuentra con el SEÑOR.

Rubén Darío Duarte Anaya

Agradezco a Dios por haberme permitido alcanzar este logro, a mi madre y a todos mis familiares y amigos que de alguna manera especial, contribuyeron con su apoyo a la culminación exitosa de este largo camino.

Sergio Armando Mejía Fajardo.

El primer pensador fue sin duda alguna el primer maniático del

Por qué.

Pues nacemos en la consternación.

A mis padres y Hermanas queridas.

Manuel O. Carreño Martínez.

CONTENIDO

INTRODUCCIÓN	34
1. GENERALIDADES DEL SEMINARIO DE INVESTIGACIÓN	36
1.1 QUE ES EL SEMINARIO DE INVESTIGACIÓN	36
1.3 VENTAJAS DEL SEMINARIO DE INVESTIGACIÓN	38
1.4. CARACTERÍSTICAS	40
1.5. ORGANIZACIÓN DEL SEMINARIO DE INVESTIGACIÓN	41
1.6. TEMA DEL SEMINARIO DE INVESTIGACIÓN	42
1.7. DIRECCIÓN DEL SEMINARIO DE INVESTIGACIÓN EN DIAGNOSTICO DE FALLAS EN SISTEMAS COMPLEJOS.....	43
2. PLANEACIÓN	44
2.1. OBJETIVOS DEL SEMINARIO.....	45
2.2. ESTUDIO Y SELECCIÓN DEL MATERIAL BIBLIOGRÁFICO.....	45
2.3. TEMARIO	46
2.4 PLANECIÓN DE LAS SESIONES	51
3. EJECUCIÓN	53
3.1.EL ACTA	53
4. PRESENTACIÓN, DOCUMENTOS Y APLICACIONES	58
4.1. PRESENTACIONES	58

4.1.1 CARACTERÍSTICAS.....	58
4.2 DOCUMENTO	61
4.3 APLICACIONES EN MATLAB	62
4.3.1 DISTOOL.....	62
4.3.2 DISTRIBUCIÓN UNIFORME.....	63
4.3.3 DISTRIBUCIÓN NORMAL (GAUSIANA)	65
4.3.4 PRUEBAS DE HIPÓTESIS	68
4.3.5 FUNDAMENTOS EN ALGEBRA LINEAL.....	70
5. CONCLUSIONES Y RECOMENDACIONES.....	83
6. BIBLIOGRAFÍA	85
ANEXOS.....	88
A.7. PERSPECTIVA HISTORICA DE LOS METODOS DE DETECCION Y DIAGNÓSTICO DE FALLAS.....	89
A.7.1.1 HISTORIA DE LOS METODOS DE DETECCION Y DIAGNÓSTICO DE FALLAS	91
A.7.1.2 HISTORIA DE LOS METODOS DE DIAGNÓSTICO Y DETECCION DE FALLAS EN EL SECTOR DE LA AVIACION.....	94
A.7.1.3 ESTADO Y TENDENCIA ACTUAL	99
A.7.2.1 ESQUEMAS DE DIAGNÓSTICO Y MONITOREO DE FALLAS.....	113
A.7.3 GENERALIDADES	117
A.7.3.1 PROCEDIMIENTOS DE MONITOREO DEL PROCESO.....	117
A.7.3.2 MEDICIONES EN EL MONITOREO DEL PROCESO	118
A.7.3.3 METODOS PARA EL MONITOREO DE PROCESOS	121
A.7.3.4 DIAGNÓSTICO Y DETECCION DE FALLAS BASADO EN MODELOS Y DATA DRIVEN - GENERALIDADES	123
A.7.3.5 ENFOQUES DE DIAGNÓSTICO DE FALLAS BASADO EN MODELOS	126
A.7.3.5.1 Métodos cuantitativos de diagnóstico basado en modelos	126
A.7.3.5.2 Métodos basados en modelos cualitativos	129

A.7.3.6 ENFOQUE DE DIAGNÓSTICO DE FALLAS DATA DRIVEN	133
A.7.3.6.1 Métodos cuantitativos de data driven	138
A. 7.3.6.2 Métodos data driven cualitativos	142
A.7.4 EJEMPLOS DE APLICACIÓN DE LOS MÉTODOS DE FDD MEDIANTE EL ENFOQUE DE DATA-DRIVEN..	144
A.7.4.1 DATA DRIVEN DIAGNÓSTICO DE FALLAS Y CONTROL TOLERANTE DE FALLAS:ALGUNOS AVANCES Y POSIBLES NUEVAS DIRECCIONES	144
A.7.4.2 DATA-DRIVEN METODOLOGÍA DE RED NEURONAL EN LA PREDICCIÓN DE TIEMPO DE VIDA RESTANTE PARA LOS COMPONENTES DEL ACTUADOR EN UN AVIÓN	144
A.7.4.3 DETECCIÓN Y DIAGNÓSTICO DE FALLAS UTILIZANDO TÉCNICAS ESTADÍSTICAS MULTIVARIABLES A PARTIR DE DATA HISTÓRICA.....	145
A.7.5 SENSORES Y ESTRATÉGIAS DE MEDICIÓN	146
A.7.5.1 SENSORES	148
A.7.5.1.1 Transductores y principios de transducción	150
A.7.5.2 SISTEMAS MECÁNICOS Y ESTRUCTURALES DE SENSORES	152
A.7.5.2.1 Acelerómetros: Medición de vibraciones.....	153
A.7.5.2.2 Medidores de deformación.....	155
A.7.5.2.3 Sensor ultrasónico.....	160
A.7.5.3 SENSORES DE RUTINA Y OPERACIONALES	167
A.7.5.3.1 Sensores térmicos / Termografía.....	169
A.7.5.4 MEDICIONES ELÉCTRICAS	176
A.7.5.4.1 Sondas de proximidad por corrientes de Foucault.....	177
A.7.5.4.2 Sistemas de sensores Micro-electromecánicos (MEMS).....	178
A.7.5.5 SENSORES DE FIBRA ÓPTICA.....	179
A.7.5.6 UBIACIÓN DEL SENSOR	180
A.7.5.7 REDES DE SENSORES INALÁMBRICOS	183
A.7.5.7.1 Topología de la Red	185
A.7.5.7.2 Sistemas de Sensores Disponibles Comercialmente.....	186
A.7.6 SISTEMAS GERENCIALES DE PROCESAMIENTO DE SEÑALES Y BASE DE DATOS.....	187
A.7.6.1 INTRODUCCIÓN	187
A.7.6.2 PROCESAMIENTO DE SEÑALES EN CBM/PHM	188
A.7.6.3 PREPROCESAMIENTO DE LA SEÑAL.....	190
A.7.6.3.1 Acondicionamiento de la señal	190

A.7.6.4 ELIMINACIÓN DE RUIDO Y EXTRACCION DE SEÑALES DEBILES	193
A.7.6.5 COMPRESION DE LA SEÑAL DE VIBRACION.....	194
A.7.6.6 PROMEDIO DE TIEMPO SINCRONO (TSA)	195
A.7.6.7 PROCESAMIENTO DE LA SEÑAL.....	201
A.7.6.7.1 Extracción y selección de características.....	201
A.7.6.8 PROCESAMIENTO DE LA SEÑAL EN EL DOMINIO DEL TIEMPO.....	210
A.7.6.9 METODOS DE GERENCIAMIENTO DE BASES DE DATOS	215
A.8. ESTADÍSTICA MULTIVARIANTE.....	218
A.8.1 INTRODUCCIÓN.....	218
A.8.1.2 PRE-TRATAMIENTO DE DATOS.....	219
A.8.1.3 MONITOREO ESTADÍSTICO UNIVARIANTE	222
A.8.2 ESTADÍSTICA MULTIVARIABLE.....	237
A.8.2.1 ESTADÍSTICO DE HOTELLING (T^2)	238
A.9. MÉTODOS USADOS CON DATA DRIVEN.....	249
A.9.1 ANÁLISIS DE COMPONENTES PRINCIPALES	249
A.9.1.1 INTRODUCCIÓN.....	249
A.9.1.2 ANÁLISIS DE COMPONENTES PRINCIPALES.....	251
A.9.1.3 REDUCCIÓN DE ORDEN	259
A.9.1.4 DETECCIÓN DE FALLAS	261
A.9.1.5 IDENTIFICACIÓN DE FALLAS	263
A.9.1.6 DIAGNÓSTICO DE FALLA.....	269
A.9.2 ANÁLISIS DISCRIMINANTE DE FISHER.	270
A.9.2.1 INTRODUCCION.....	270
A.9.2.2 ANALISIS DISCRIMINANTE DE FISHER.....	271
A.9.2.3 REDUCCION DE ORDEN	274
A.9.2.4 DETECCION Y DIAGNOSTICO DE FALLAS	278
A.9.2.5 FDA DINÁMICO.....	289
A.9.3 MÍNIMOS CUADRADOS PARCIALES	291

A.9.3.1 INTRODUCCIÓN	291
A.9.3.2 ALGORITMOS PLS	293
A.9.3.3 REDUCCIÓN DE ORDEN	300
A.9.3.4 DETECCIÓN DE FALLAS, Y DIAGNÓSTICO.....	302
A.9.3.5 COMPARACIÓN ENTRE PCA Y PLS	304
A.9.4 ANÁLISIS DE VARIACIÓN CANÓNICA (CVA)	304
A.9.4.1 ALGORITMO CVA.....	309
A.9.5 APLICACIONES.....	311
A.9.5.1 INTRODUCCIÓN	311
A.9.5.2 CONJUNTO DE DATOS	315
A.9.5.3 INTERVALO DE MUESTREO.....	320
A.9.5.4 TAMAÑO DE LA MUESTRA	323
A.9.5 SELECCIÓN DE RETRASO Y ORDEN	327
A.10. METODOS BASADOS EN CONOCIMIENTO.....	328
A.10.1 INTRODUCCIÓN.....	329
A.10.2 ANALISIS CUALITATIVO DE TENDENCIAS (QTA)	330
A.10.2.1 BASE DE CONOCIMIENTO PARA QTA.	336
A.10.2.2 DIAGNOSTICO DE FALLAS ON-LINE CON QTA	338
A.10.2.3 CONSTRUYENDO LA BASE DE CONOCIMIENTO QTA.....	343
A.10.2 SISTEMAS EXPERTOS	348
A.10.2.1 SISTEMA EXPERTO DE CONOCIMIENTO SUPERFICIAL.....	349
A.10.2.2 SISTEMAS EXPERTOS DE CONOCIMIENTO PROFUNDO.....	350
A.10.2.3 COMBINACIÓN DE SISTEMAS EXPERTOS DE CONOCIMIENTO-SUPERFICIAL Y DE CONOCIMIENTO-PROFUNDO	351
A.10.2.4 TÉCNICAS DE APRENDIZAJE AUTOMÁTICO	352
A.10.2.5 REPRESENTACIÓN DEL CONOCIMIENTO	353
A.10.2.6 MOTOR DE INFERENCIA	354
A.10.2.7 SELECCIONADO PROBLEMAS PARA DESARROLLAR SISTEMAS EXPERTOS	363
A.10.2.8 EL PROCESO DE CONOCIMIENTO INGENIERIL.....	364

A.10.2.9 MODELOS CONCEPTUALES Y SU ROL EN LA ADQUISICIÓN DE CONOCIMIENTO	366
A.10.2.10 EL SISTEMA DE PRODUCCIÓN Y SOLUCIÓN DE PROBLEMAS MEDIANTE GOAL-DRIVEN	368
A.10.3 RECONOCIMIENTO DE PATRONES.	373
A.10.3.1 REDES NEURONALES.	376
A.10.3.1.1 Definiciones de una red neuronal.	376
A.10.3.1.3 ARQUITECTURAS DE REDES NEURONALES.....	382
A.10.3.1.3.1 Organización de las redes.....	383
A.10.3.1.3.2 Funciones neuronales	386
A.10.3.1.4 ALGORITMOS DE APRENDIZAJE	389
A.10.3.1.4.1 Aprendizaje supervisado.	390
A.10.3.1.4.2 Aprendizaje no supervisado.	392
A.10.3.2 MAPAS AUTO-ORGANIZATIVOS (SOM)	393
A.10.4 COMBINACIÓN DE VARIAS TÉCNICAS	395
A.10.4.1 REDES NEURONALES Y SISTEMAS EXPERTOS	395
A.10.4.2 LÓGICA FUZZY.....	396
A.10.4.3 SISTEMAS EXPERTOS FUZZY	400
A.10.4.4 REDES NEURONALES FUZZY	402

LISTA DE FIGURAS

FIGURA 1. FORMATO ACTA	54
FIGURA 2. FORMATO DE LA PRESENTACIÓN	61
FIGURA 3. VISTA DE LA INTERFAZ EN MATLAB DE LA HERRAMIENTA DISTOOL	63
FIGURA 4. GRÁFICA DE LA FUNCIÓN UNIFORME DEL EJEMPLO ANTERIOR.	65
FIGURA 5. VISTA DE LA INTERFAZ EN MATLAB DE LA HERRAMIENTA DISTOOL DE LA FUNCIÓN DE DENSIDAD DE PROBABILIDAD.....	66
FIGURA 6.DISTRIBUCIÓN NORMAL GAUSSIANA PARA EL EJEMPLO ANTERIOR .	68
FIGURA 7. BOSQUEJO DEL DESARROLLO HISTÓRICO DE LA TEORÍA EN EL DIAGNÓSTICO DE FALLAS.....	93
FIGURA 8. EVENTOS PRINCIPALES EN LA EVOLUCIÓN DE LOS MÉTODOS DE DIAGNÓSTICO Y DETECCIÓN DE FALLAS EN LA AVIACIÓN	98
FIGURA 9. CLASIFICACIÓN DE LOS MÉTODOS DE DIAGNÓSTICO DE FALLAS. ...	100
FIGURA.10. DESCRIPCIÓN ESQUEMÁTICA DEL ESQUEMA DE HARDWARE REDUNDANTE	101
FIGURA 11. REDUNDANCIA FÍSICA VERSUS A REDUNDANCIA ANALÍTICA.....	102
FIGURA.12. ESQUEMA GENERAL DE LA REDUNDANCIA ANALÍTICA.....	103

FIGURA.13. DESCRIPCIÓN ESQUEMÁTICA DEL MÉTODO BASADO EN PROCESAMIENTO DE SEÑALES	110
FIGURA. 14.DESCRIPCIÓN ESQUEMÁTICA DE LA PRUEBA DE PLAUSIBILIDAD.	111
FIGURA 15. DESCRIPCIÓN ESQUEMÁTICA DE FALLOS EN LOS SISTEMAS.....	112
FIGURA. 16.MARCO GENERAL DE DIAGNÓSTICO	112
FIGURA. 17. ESTRUCTURA GENERAL FDI.....	113
FIGURA.18. TÉCNICAS BASADAS EN MODELOS Y DATA- DRIVEN	114
FIGURA.19. MONITOREO Y DIAGNOSIS DE FALLAS ONLINE	115
FIGURA 20. TRANSFORMACIONES EN EL SISTEMA DE DIAGNÓSTICO.....	116
FIGURA 21. CLASIFICACIÓN DE LOS MÉTODOS DE DIAGNÓSTICO DE FALLAS..	124
FIGURA 22. DESCRIPCIÓN ESQUEMÁTICA DEL DIAGNÓSTICO DE FALLAS BASADO EN MODELOS	126
FIGURA. 23. GENERACIÓN Y PRUEBA DE HIPÓTESIS SECUENCIAL	127
FIGURA 24.FORMAS DE CONOCIMIENTO CUALITATIVO.....	130
FIGURA 25. EJEMPLO DE DÍGRAFO PARA UN TANQUE SIMPLE.....	131
FIGURA 26.CLASIFICACIÓN DE LOS MÉTODOS BASADOS EN LA HISTORIA DEL PROCESO.....	133

FIGURA 27. CLASIFICACIÓN DE LOS MÉTODOS DE EXTRACCIÓN EN DATA-DRIVEN.....	136
FIGURA 28. EJEMPLO DE UNA RED NEURONAL.....	141
FIGURA 29. CARACTERÍSTICAS DE LOS NODOS USADOS EN REDES NEURONALES.	141
FIGURA 30. MARCO DE INTEGRACIÓN.	144
FIGURA 31. (A) REPRESENTACIÓN ESQUEMÁTICA DE UN DISPOSITIVO DE MEDIDA. (B) OPERACIÓN DE UN ACELERÓMETRO PIEZOELÉCTRICO.....	151
FIGURA 32. MODELO SIMPLIFICADO DE UN ACELERÓMETRO.	155
FIGURA 33. UN ACELERÓMETRO PIEZOELÉCTRICO DE BRUEL AN KJAER (TIPO 7704 A-50).....	156
FIGURA 34. ACELERÓMETRO TRIAXIAL MINIATURA DELTA TRON (TIPO 4524)	156
FIGURA 35.. POSICIONAMIENTO DEL EXTENSÓMETRO.....	157
FIGURA 36. PUENTE WHEATSTONE.....	158
FIGURA 37 CABLE COAXIAL RG58 CON DOS CONECTORES MACHOS BNC.	159
FIGURA 38. TRANSDUCTORES ULTRASÓNICOS ADHERIDOS A PARTES METÁLICAS. (A) DISCOS PIEZOELÉCTRICOS ENCAPSULADOS EN EPOXI Y UNIDA A LA PIEZA. (B) TRANSDUCTORES DE ANGULARES DISPONIBLES	

COMERCIALMENTE PROTEGIDOS CON BLOQUE DE ALUMINIO EN U Y UNIDOS A LA PIEZA.	162
FIGURA 39. (A) SEÑALES DE ONDAS ULTRASÓNICAS LAMB DISPERSAS DE UN AGUJERO PERFORADO EN UNA PLACA DE ALUMINIO CON LAS REGIONES MARCADAS, INDICÁNDOSE LA DISPERSIÓN DEL AGUJERO. (B) LA IMAGEN DE UNA MATRIZ DE FASE RESULTANTE CONSTRUIDA A PARTIR DE LAS SEIS SEÑALES. LOS CÍRCULOS SÓLIDOS SON TRAZADOS EN LOS LUGARES DE LOS TRANSDUCTORES ADJUNTOS, EL CÍRCULO ABIERTO SE CENTRA EN LA UBICACIÓN DEL AGUJERO REAL, Y LA INSERCIÓN ES UNA VISTA AMPLIADA DEL AGUJERO.	164
FIGURA 40. MÉTODO DE TRANSMISIÓN A TRAVÉS DE UN HAZ ANGULAR PARA MONITOREAR IN SITU DE LAS GRIETAS DE FATIGA DE ORIGINADAS POR AGUJEROS DE FIJACIÓN.	166
FIGURA 41. GRÁFICA DEL ÁREA DE GRIETA VERSUS LOS CICLOS DE FATIGA PARA MONITOREO DE AGUJEROS DE FIJACIÓN.	166
FIGURA 42. COHERENCIA LOCAL TEMPORAL DE ONDAS DIFUSAS DE UN COMPONENTE DAÑADO COMPARADO CON UNA SEÑAL DE REFERENCIA DE UN COMPONENTE EN BUEN ESTADO. LAS GRÁFICAS (A) A (D) MUESTRAN LA CAÍDA PROGRESIVA DE LA COHERENCIA COMO UN DEFECTO ARTIFICIAL ES AGRANDADO.	168
FIGURA 43 ARQUITECTURA DE UN SISTEMA DE MONITOREO DE SALUD POR ULTRASONIDO DE UNA ESTRUCTURA.	169
FIGURA 44. RTD DISPONIBLES EN SENSORTEC INC.	171

FIGURA 45. RTD INDUSTRIA DE PLATINO SERIE 21 A.	171
FIGURA 46. ISO-TERMISTORES DE Sonda Y ESFERAS DE VIDRIO	173
FIGURA 47. TERMISTORES DE PLOMO ESTÁNDAR RECUBIERTOS DE EPOXY...174	
FIGURA 48. IMÁGENES INFRARROJAS DEL PROGRESO DEL DAÑO POR FATIGA ALREDEDOR DEL AGUJERO DE REMACHE EN UN ESPÉCIMEN DE PRUEBA 7050- T7451	176
FIGURA 49. UN TÍPICO SENSOR DE TEMPERATURA DE FIBRA ÓPTICA.	181
FIGURA 50. ARQUITECTURA DE SELECCIÓN/ CLASIFICACIÓN	183
FIGURA 51. RED DE SENSORES INALÁMBRICOS	184
FIGURA 52. TOPOLOGÍAS DE RED BÁSICAS.....	186
FIGURA 53. FILTRO ANÁLOGO PASA-BAJA	191
FIGURA 54. DFT DE UN CONJUNTO DE DATOS TSA.....	199
FIGURA 55. REVOLUCIÓN MÚLTIPLE DE DATOS SIN INTERPOLACIÓN EN EL DOMINIO DEL TIEMPO Y LA FRECUENCIA	199
FIGURA 56. TSA EN EL DOMINIO DEL ÁNGULO Y EL CORRESPONDIENTE DOMINIO DE LA FRECUENCIA.	200
FIGURA 57. DATOS SIN INTERPOLACIÓN EN EL DOMINIO DEL ÁNGULO Y DE LA FRECUENCIA.....	201

FIGURA 58. ESQUEMA GENERAL DE EXTRACCIÓN DE CARACTERÍSTICAS Y CLASIFICACIÓN DE MODO DE FALLA.....	202
FIGURA 59. MEDIDA DE LA DISTINCIÓN.....	205
FIGURA 60 SEMEJANZA MÉTRICA.....	207
FIGURA 61. EL PROCEDIMIENTO DE PERCEPCIÓN ACTIVA PARA PATRONES DE RECONOCIMIENTO.....	210
FIGURA 62. OBLICUIDAD, UNA MEDIDA DE LA SIMETRÍA DE LA DISTRIBUCIÓN.....	213
FIGURA 63. KURTOSIS, COMO UNA MEDIDA DEL TAMAÑO DE LOS LÓBULOS DE UNA DISTRIBUCIÓN.....	213
FIGURA 64. SERVIDOR DE DATOS COMO COMPONENTE DE LA ARQUITECTURA CBM/PHM.....	216
FIGURA 65. UNA MUESTRA DE UN ESQUEMA NÚCLEO DE BASE DE DATOS.....	217
FIGURA 66. ESTADÍSTICA EN EL MONITOREO DE PROCESOS.....	219
FIGURA 67. PROCEDIMIENTO PARA EL PRETRATAMIENTO DE DATOS.....	220
FIGURA 68. SIMETRÍA DE LOS DATOS SEGÚN EL SESGO.....	226
FIGURA 69. MUESTRAS CON MEDIDAS DE TENDENCIA CENTRAL IDÉNTICAS, PERO CANTIDADES DISTINTAS DE VARIABILIDAD.....	227
FIGURA 70. ERRORES TIPO I (α) Y TIPO II (β).....	232

FIGURA 71. DIAGRAMA PUNTO DE OBSERVACIONES INDIVIDUALES DE UNA VARIABLE.....	235
FIG 72. ILUSTRACIÓN DE LA GRÁFICA DE CONTROL DE SHEWHART. LOS PUNTOS NEGROS SON OBSERVACIONES.....	236
FIGURA.73. UNA FUNCIÓN DE DOS VARIABLES DE PROBABILIDAD NORMAL $\sigma_1=1, \sigma_2=1, \rho=0,8$.....	241
FIGURA. 74. UN BIVARIADO CONTORNO NORMAL CONTIENE UN 75%Y EL 95%DE LA PROBABILIDAD.....	242
FIGURA 75. ILUSTRACIÓN GRAFICA DE LA CONVERSIÓN DE LA COVARIANZA PARA LA ESTADÍSTICA T^2.....	244
FIGURA 76. UNA COMPARACIÓN DE LA REGIÓN EN ESTADO DE CONTROL USANDO ESTADÍSTICA T^2 Y LA ESTADÍSTICA UNIVARIANTE PARA DOS VARIABLES DE PROCESO.....	245
FIGURA 77. LA PROYECCIÓN DEL VECTOR DE OBSERVACIÓN X DENTRO DEL ESPACIO SCORE Y RESIDUAL, Y LA COMPUTACIÓN DE LA OBSERVACIÓN FILTRADA \hat{x}.....	254
FIGURA 79. LAS PROYECCIONES DE LOS DATOS FISHER PARA LAS TRES CLASES SOBRE LOS DOS PRIMEROS VECTORES DE CARGA PCA.....	258
FIGURA 80. UNA COMPARACIÓN DE LA TRANSFORMACIÓN LINEAL PCA Y FDA DE LOS DATOS EN CLASES "X" Y "O" SOBRE EL PRIMER VECTOR DE CARGA FDA Y PCA.....	283

FIGURA 81. TRANSFORMACIÓN LINEAL DE FISHER DE LOS DATOS PARA TRES CLASES SOBRE LOS PRIMEROS DOS VECTORES DE CARGA FDA Y PCA RESPECTIVAMENTE.....	284
FIGURA 82 DIAGRAMA DE FLUJO DEL PROCESO TEP CON ESTRUCTURA DE CONTROL SECUNDARIA.....	313
FIGURA 83. SIMULACIÓN EN LAZO CERRADO PARA ESTADO ESTABLE SIN FALLAS. LAS LÍNEAS SÓLIDAS Y PUNTEADAS EN LA GRÁFICA INFERIOR DERECHA REPRESENTAN LAS COMPOSICIONES DE G Y H RESPECTIVAMENTE.	322
FIGURA 84.. SIMULACIÓN EN LAZO CERRADO PARA CAMBIO ESCALONADO DE LA COMPOSICIÓN B (IDV(2) EN LA TABLA 9.7).....	325
FIGURA 85 LENGUAJE FUNDAMENTAL: PRIMITIVAS.....	331
FIGURA. 86. SELECCIÓN ADAPTATIVA DE LA VENTANA.....	335
FIGURA. 87 DIAGRAMA DE FLUJO PARA LA EXTRACCIÓN DE TENDENCIAS ON-LINE	336
FIGURA. 88. INSTANTÁNEA DURANTE DOS VENTANAS CONSECUTIVAS	336
FIGURA 89. IMPORTANCIA DEL TIEMPO DE OCURRENCIA DE LA FALLA	339
FIGURA 90. ESTIMACIÓN DEL TIEMPO DE OCURRENCIA DE LA FALLA	347
FIGURA 91. RED SEMÁNTICA REPRESENTADO EL CONOCIMIENTO “FT ES UN TRANSMISOR DE FLUJO CON UNA RATA DE 10m³/hr”	354
FIGURA 92. UN AGENTE SOFTWARE CON INTERACCIÓN DE USUARIO.....	356

FIGURA 93. UN AGENTE HARDWARE	356
FIGURA 94. ARQUITECTURA GENERAL DE UN SISTEMA BASADO EN CONOCIMIENTO.	358
FIGURA 95. ARQUITECTURA DE UN SISTEMA EXPERTO TÍPICO	361
FIGURA 96. DIAGRAMA DE FLUJO DEL CICLO DE DESARROLLO DEL PROGRAMA DE EXPLORACIÓN.....	365
FIGURA 97. REPRESENTACIÓN DE UN MODELO SIMPLIFICADO DEL PROCESO DE ADQUISICIÓN DE CONOCIMIENTO.....	366
FIGURA 98. EL ROL DE MODELOS MENTALES O CONCEPTUALES EN SOLUCIONADORES DE PROBLEMAS.....	368
FIGURA 99. EL SISTEMA DE PRODUCCIÓN AL COMIENZO DE LA CONSULTA E EL EJEMPLO DE DIAGNÓSTICO DE UN AUTOMÓVIL	370
FIGURA 100. EL SISTEMA DE PRODUCCIÓN DESPUÉS QUE LA REGLA 1 HA SIDO LANZADA.....	371
FIGURA 101. SISTEMA DE PRODUCCIÓN DESPUÉS QUE LA REGLA 4 HA SIDO LANZADA.....	372
FIGURA 102. EJEMPLO DE ESQUEMA GENERAL PARA EL RECONOCIMIENTO DE PATRONES.....	374
FIGURA 103. ESQUEMA PARA EL MECANISMO DE GENERACIÓN DE PATRONES.	375

FIGURA 104. SISTEMA GLOBAL DE PROCESO DE UNA RED NEURONAL.....	381
FIGURA 105. EJEMPLO DE UNA RED NEURONAL TOTALMENTE CONECTADA...381	381
FIGURA 106. EJEMPLO DE UNA NEURONA CON 2 ENTRADAS Y UNA SALIDA.....	387
FIGURA 107. ESQUEMA DE APRENDIZAJE SUPERVISADO.....	390
FIGURA 108. ILUSTRACION DE MODO DE OPERAR DEL MAPA AUTO- ORGANIZATIVO	394
FIGURA 109. FUNCIÓN DE PERTENENCIA REPRESENTANDO 3 CONJUNTOS FUZZY PARA LA VARIABLE LINGÜÍSTICA TEMPERATURA CORPORAL.	399

LISTA DE TABLAS

TABLA 1. ROL ASIGNADO A LOS PARTICIPANTES DEL PROYECTO EN LAS DIFERENTES SESIONES.....	51
TABLA 2. ARREGLOS DE PUENTES Y SUS COMPENSACIONES.....	159
TABLA 3. NOMENCLATURA (ANSI), MATERIALES PARA TERMOCUPLAS Y SUS RESPECTIVOS RANGOS DE OPERACIÓN.....	174
TABLA 4. REQUERIMIENTOS DE DATOS PARA LA ECUACIÓN 8.22	247
TABLA 5 ESTADÍSTICA DE LOS DATOS FISHER.....	257
TABLA 6. TASA DE ERROR DE CLASIFICACIÓN PARA LOS DATOS DE ENTRENAMIENTO PARA FDA	285
TABLA 7. TASA DE ERROR DE CLASIFICACIÓN PARA LOS DATOS DE ENTRENAMIENTO, UTILIZANDO VARIOS MÉTODOS PARA EL DIAGNOSTICO DE FALLAS DATA-DRIVEN (60 OBSERVACIONES EN EL CONJUNTO DE ENTRENAMIENTO).....	288
TABLA 8. TASA DE ERROR DE CLASIFICACIÓN GLOBAL DE LOS DATOS DE PRUEBA, UTILIZANDO VARIOS MÉTODOS DE DIAGNOSTICO DE FALLAS DATA-DRIVEN (60 OBSERVACIONES EN EL CONJUNTO DE ENTRENAMIENTO)	288
TABLA 9. TASA DE ERROR DE CLASIFICACIÓN PARA LOS DATOS DE ENTRENAMIENTO UTILIZANDO VARIOS MÉTODOS DE DIAGNOSTICO DATA-DRIVEN (120 OBSERVACIONES EN EL CONJUNTO DE ENTRENAMIENTO)	290

TABLA 10. TASA DE ERROR EN LA CLASIFICACIÓN PARA LOS DATOS DE PRUEBA UTILIZANDO VARIOS MÉTODOS DE DIAGNÓSTICO DE FALLAS DATA-DRIVEN(120 OBSERVACIONES EN EL CONJUNTO DE ENTRENAMIENTO)	291
TABLA 11. FALLAS DEL PROCESO	314
TABLA 12 VARIABLES MANIPULADAS	316
TABLA 13. MEDICIONES DEL PROCESO (INTERVALOS DE MUESTRA DE 3 MINUTOS).....	317
TABLA 14 MEDICIONES DE COMPOSICIONES (UNIDADES: MOL %)	318
TABLA 15 EL NÚMERO DE PARÁMETROS INDEPENDIENTES ESTIMADO PARA VARIOS MODELOS Y ORDENES.....	326

RESUMEN

TITULO: SEMINARIO DE INVESTIGACIÓN EN METODOS DE DIAGNOSTICO DE FALLAS DE SISTEMAS COMPLEJOS¹

AUTORES²: MANUEL ORLANDO CARREÑO MARTINEZ
LUZ AMPARO ALVAREZ TORRES
SERGIO ARMANDO MEJIA FAJARDO
RUBEN DARIO DUARTE ANAYA

PALABRAS CLAVES: Detección y diagnóstico de fallas (FDD), PCA (análisis de componentes principales), PLS (mínimos cuadrados parciales), FDD basado en Data-Driven, FDD basado en modelos.

Este documento presenta las técnicas utilizadas y los actuales avances, en la detección y diagnóstico de fallas para procesos industriales, que cuentan con un alto volumen de datos disponibles, grado de instrumentación y complejidad; al inicio se realiza una clasificación global de los métodos en dos grandes enfoques: FDD basado en modelos y FDD basado en data-driven, haciendo énfasis en los métodos de estadística multivariante de reducción dimensional de datos (data-driven), junto a un análisis comparativo de estos dos enfoques. Dentro de las técnicas expuestas sobresalen PCA, PLS, redes neuronales, lógica fuzzy y clasificadores estadísticos.

Mediante este documento se pretende dar una introducción a técnicas alternativas para la industria en el campo del diagnóstico y detección de fallas, mediante direccionamiento de datos (data-driven), que aún están en fase de desarrollo e implementación. En estas últimas décadas, los métodos de detección y diagnóstico de fallas se han convertido en un tópico de considerable interés dentro de la comunidad de investigadores pertenecientes al área de control, esto es debido a los crecientes requerimientos de operaciones confiables en los sistemas de control, las cuales se encuentran, la mayor parte de los casos, sujetos a un determinado número de fallas internas en lazos cerrados (closed loops) o fallas generadas por factores ambientales. Es por ello que la detección y diagnóstico de fallas (FDD de sus siglas en inglés Fault Detection and Diagnosis) es de vital importancia para la operación segura de las plantas industriales.

¹ Trabajo de grado.

² Universidad Industrial de Santander, Facultad de Ingenierías Físico-mecánicas, Escuela de Ingeniería Mecánica, Director: Ing. Mecánico, Rómulo Niño Delgado.

ABSTRACT

TITLE: SEMINARIO DE INVESTIGACIÓN EN METODOS DE DIAGNOSTICO DE FALLAS DE SISTEMAS COMPLEJOS³

AUTHORS⁴: MANUEL ORLANDO CARREÑO MARTINEZ
LUZ AMPARO ALVAREZ TORRES
SERGIO ARMANDO MEJIA FAJARDO
RUBEN DARIO DUARTE ANAYA

KEY WORDS: Fault Detection and Diagnosis (FDD), PCA (principal components analysis), PLS (least partial squared), FDD based on models, FDD based on Data-Driven.

DESCRIPTION: This document presents the approaches and current advances, in the fault Detection and Diagnosis (FDD), in industrial processes that have a big volume of data available, level of instrumentation and complexity; Beginning with a global classification of the methods, divide in two general approaches: FDD based in models and FDD based in Data-Driven, making emphasis in multivariate statistical methods of dimensional reduction of data (Data-Driven), which is, the most recent method, and a comparative analysis of this two approaches, in the final part. Within the exposed techniques stand, PCA, PLS, Neural Networks, fuzzy logic, and statistical classifiers.

The intend with this document is, give an introduction to alternatives approaches for the industry (Data-Driven) in the field of Fault Detection and Diagnosis, that still are in a phase of development and implementation. In the last decades, the fault detection and diagnosis methods, has become on a main topic within the industry and academy, due to higher requirements of reliable operations in the control systems, which are attached to a number of inner failures in the closed loops or failures due to ambiental issues. That's why the Fault Detection and Diagnosis are very important to the safe and reliable operations in industrial plants.

³ Trabajo de Grado.

⁴ Universidad Industrial de Santander, Facultad de Ingenierías Físico-mecánicas, Escuela de Ingeniería Mecánica, Director: Ing. Mecánico, Rómulo Niño Delgado.

INTRODUCCIÓN

El continuo incremento en los estándares de desempeño, y calidad de producción por un lado y por el otro la optimización de costos, son las variantes que influyen en el grado de complejidad y automatización de los sistemas industriales, los cuales deben garantizar la seguridad y confiabilidad del proceso; por ello actualmente uno de los problemas más críticos que rodean el diseño de sistemas automáticos, es precisamente su confiabilidad y nivel de dependencia.

Una manera común de mejorar la confiabilidad del sistema y nivel de dependencia es aumentar el desempeño y la robustez de los componentes como sensores, servomecanismos, controladores u ordenadores. De manera tal que inclusive en un sistema libre de fallas, su operación no puede ser totalmente garantizada. Por ello el monitoreo de procesos y detección de fallas se ha convertido ya en parte fundamental de los sistemas automáticos de control moderno. Todos los sistemas de detección de fallas ya sea basado en modelos o basado en conocimiento, tienen en común el uso de un modelo del proceso, dependiendo del tipo de algoritmo que sea implementado para el procesamiento de datos, si son recolectados en línea, o grabados durante la operación del sistema.

Las instalaciones de la manufactura moderna son a grande a escala, de alta complejidad y operan con un gran número de variables sometidas al control de lazos cerrados (closed-loop). Un diagnóstico y detección de falla temprana y precisa para estas plantas puede minimizar los tiempos de inactividad,

incrementando la seguridad de las operaciones en la planta, y reduciendo costos de manufactura.

1. GENERALIDADES DEL SEMINARIO DE INVESTIGACIÓN

1.1 QUE ES EL SEMINARIO DE INVESTIGACIÓN⁵

El seminario de Investigación, también conocido como Seminario Alemán, es una actividad académica cuyo origen se dio a finales del siglo XVIII en la Universidad de GOTTINGEN de Alemania, para renovar las estrategias de estudio y formación de los investigadores, además para demostrar que la docencia y la investigación se pueden unir y complementar para obtener mejores resultados.

Consiste en estudiar, discutir e intercambiar experiencias acerca de un tema en particular, en un grupo, en el cual sus participantes se intercomunican exponiendo dicho tema (la Relatoría) complementándolo, evaluándolo (Correlatoria), aportando entre todos (la Discusión), sacando conclusiones y planteando nuevos interrogantes lo que permite que todo ello quede en la memoria escrita (el Protocolo).

El seminario de investigación se programa por subtemas, estos son seleccionados con la orientación del director del seminario, quien con su experiencia y conocimiento del tema central, guía la selección con la debida pertinencia,

⁵Universidad Industrial de Santander. Lineamientos para el seminario de investigación como modalidad para el desarrollo del trabajo de grado. Vicerrectoría académica. Septiembre de 2007

actualidad y ubicación en el contexto. Los temas son desarrollados en sesiones planificadas, en las cuales los miembros del grupo deben asumir diferentes roles de acuerdo con la descripción anterior y mantener una relación de interés y compromiso con el conocimiento sin jerarquías en un clima de colaboración y participación activa.

1.2 OBJETIVO DEL SEMINARIO DE INVESTIGACIÓN

Formar a los participantes para la investigación científica mediante el desarrollo de habilidades aplicadas al asumir los diferentes roles dentro del seminario. Dichas habilidades están orientadas a desarrollar la capacidad del lector crítico de resultados de investigación en cualquiera de las áreas del conocimiento, a fortalecer la capacidad de observar e identificar los problemas presentes en temas bajo análisis, a buscar respuestas a preguntas claves y sustentarlas teórica y metodológicamente en forma verbal y por escrito, y a identificar las relaciones del problema objeto de estudio con el contexto económico, político o social, a fin de enriquecer con una mirada de integralidad, el conocimiento, el conocimiento para el grupo de estudiantes. Para ello se programan y ejecutan ejercicios estructurados que permiten a los estudiantes desarrollar competencias iniciales de investigador, avanzar en el conocimiento y aportar buenas revisiones y análisis sobre tópicos que pueden facilitar el desarrollo de la investigación.

Para alcanzar dicho objetivo es preciso que haya una formación desde el trabajo personal hacia el trabajo en equipo; para esto, cada participante debe reconocer sus intereses, estilos de aprendizaje, su capacidad para aprender en interacción

con pares; debe apropiarse de la metodología e instrumentos con los cuales trabajará, con el fin de lograr, al interactuar con los demás miembros del grupo en las secciones del seminario, compartir, criticar y corregir las ideas que surjan de él, en un ambiente de colaboración mutua.

Los seminarios de investigación, no se enfocan hacia la repetición de trabajos ya realizados, sino hacia la búsqueda de respuestas con nuevos argumentos; por tal razón los trabajos que se deriven del cumplimiento del objetivo del Seminario, deben caracterizarse por su originalidad y estar acordes al nivel científico de formación de sus participantes.

1.3 VENTAJAS DEL SEMINARIO DE INVESTIGACIÓN

El seminario de investigación, como herramienta para el desarrollo integral, presenta las siguientes ventajas:

- Permite a los participantes contar con un director (profesor) durante el seminario, el cual les guía hacia la consecución de los propósitos establecidos, y además, resuelve las dudas e inquietudes o en su defecto, orienta sobre las fuentes de consulta y ayuda a los miembros del grupo en la búsqueda para suplir las necesidades de información.
- Fortalece el hábito de documentarse acerca del tema bajo estudio. Para esto los participantes recurren a fuentes bibliográficas, bases de datos, textos de referencia obligada; este ejercicio refuerza el desarrollo de las competencias interpretativas, argumentativas y propositivas; les permite apropiarse y

aprender de los métodos que emplearon los autores de los artículos y textos, así como también, reconocer su valor y aporte a la investigación.

- Permite que los participantes desempeñen diferentes roles dentro del grupo, desarrollando habilidades comunicativas y de relaciones interpersonales, complementarias para la formación tanto personal como profesional.
- Fomenta el aprendizaje como una experiencia grupal, permite experimentar la eficiencia del trabajo en equipo y si el grupo está conformado con estudiantes de diferentes áreas del conocimiento, la riqueza de la interdisciplinariedad, características todas aplicables y necesarias en el desempeño laboral del mundo de hoy.
- Permite el uso de distintas herramientas didácticas de apoyo al desarrollo de las sesiones, así como un control sobre la planificación establecida de éstas.
- Es una metodología integradora centrada en el estudiante, con amplio potencial para fortalecer la habilidad de aprender a aprender, fundamentalmente para tomar el perfil de ciudadano del siglo XXI, quién deberá asumir el compromiso de aprender a lo largo de la vida, como lo plantea J. Delors⁶. Sus ventajas la hacen aplicable a todas las áreas del

⁶ Jaques Delors et al. La educación Encierra un Tesoro. Informe a la UNESCO de la Comisión internacional sobre La educación para el siglo XXI. Paris, 1996.

conocimiento humano y no solo a algunas como equivocadamente pudiera pensarse⁷.

1.4. CARACTERÍSTICAS

El seminario de investigación posee las siguientes características:

- Participación activa de todos los miembros del seminario, puesto que no sólo el director (profesor) interviene, sino también todos los integrantes del grupo realizan sus aportes desde el rol que estén desempeñando. En este proceso los participantes siendo discípulos empiezan a recorrer el camino hacia maestros.
- El seminario de investigación está conformado por un grupo reducido, de aprendizaje activo y cooperativo, inducido a investigar, reflexionar, descubrir y concluir.
- Empleo del dialogo permanente para compartir los conocimientos adquiridos.
- Ambiente amable y cooperativo fomentando la mayor participación de los integrantes del grupo.
- Sesiones desarrolladas utilizando medios didácticos de apoyo al aprendizaje.

⁷Carlos M. Vélez S, Resumen El Seminario Investigativo, basado en el simposio permanente sobre la Universidad 1990-1992 ASCUN.

- La estructura del seminario y todas las actividades y parámetros para desarrollarlas, son planificados en la primera sesión.
- El seminario de investigación exige a los participantes una alta responsabilidad para lograr la preparación adecuada, que les permita tener bases para llevarlo a cabo.

1.5. ORGANIZACIÓN DEL SEMINARIO DE INVESTIGACIÓN

El seminario de investigación se compone de las siguientes actividades: la Relatoría, la Correlatoría, la Discusión y el Protocolo, las cuales deben girar en torno a un tema, del que se desprenden los subtemas que se trataran durante las sesiones.

Dichas actividades son responsabilidad de los integrantes del grupo. Por lo cual a cada uno de ellos es asignado un rol de carácter rotativo; es así como una persona que en una sesión asume el rol de correlator podrá ser el protocolante en la siguiente, es decir, el responsable del protocolo.

La organización del seminario también implica establecer el lugar donde se llevará a cabo, el número de sesiones y las fechas para realizarlo; así como la duración de cada una de las actividades (Relatoría, Correlatoría, Discusión, Protocolo).

El seminario de investigación en Diagnostico de Fallas en Sistemas Complejos está compuesto por:

Director: Ing. Rómulo Niño Delgado

Participantes: Rubén Darío Duarte Anaya

Manuel Orlando Carreño Martínez

Sergio Armando Mejía Fajardo

Luz Amparo ÁlvarezTorres

A los participantes se les asigna un rol para cada sesión. Las sesiones se llevan a cabo en la sala CAD de la escuela de Ingeniería Mecánica. El número de sesiones depende del tema y los subtemas a los que hace referencia el temario (sección 2.3). Las sesiones están programadas para una duración entre una hora y dos.

1.6. TEMA DEL SEMINARIO DE INVESTIGACIÓN

El seminario de investigación tratará las metodologías de diagnóstico de fallas, las cuales han alcanzado un grado de exigencia cada vez mayor, tanto en la detección como en el tiempo de reacción y las tasas de falsas alarmas o detecciones perdidas, todo con el propósito de disminuir el impacto de éstas sobre los distintos sistemas, procesos, y en global el funcionamiento de las plantas industriales. Es por este motivo que los esfuerzos en desarrollar técnicas que faciliten y aumenten el grado de confiabilidad y control de los sistemas y/o

equipos, es una de las tareas primordiales del control, mantenimiento y automatización industrial. Además dada la inferencia entre las mediciones procedentes del monitoreo de variables críticas de los procesos, se hace necesaria su caracterización y completa identificación, con esta premisa se pretende presentar los conocimientos previos que se deben tener acerca de la implementación de éstas.

El estudio de estas metodologías para el diagnóstico de fallas, se enfocan en el tratamiento de los datos; mediante reducción dimensional y la extracción de las características relevantes que estén relacionadas con indicios de posibles fallas teniendo en cuenta la variabilidad del proceso, el contexto operacional, la capacidad de procesamiento, el control y monitoreo.

1.7. DIRECCIÓN DEL SEMINARIO DE INVESTIGACIÓN EN DIAGNOSTICO DE FALLAS EN SISTEMAS COMPLEJOS

En la dirección del seminario de investigación en Diagnostico de Fallas en Sistemas Complejos estuvo a cargo del profesor: Ing Rómulo Delgado Niño

Titulo obtenido: Ingeniero Mecánico	Universidad Industrial de Santander U.I.S.
Tesis: “Diseño, Construcción y Operación de una planta de Biogás”	Bucaramanga, 1994

<p>Especialista en Ingeniería Mecatronica</p> <p>Monografía:</p> <p>“Laboratorio de Sistemas Dinámicos.</p> <p>Propuesta de Implementación</p>	<p>Universidad Industrial de Santander U.I.S.</p> <p>Bucaramanga, 2003</p>
--	--

Para llevar a cabo el seminario de investigación en Diagnóstico de Fallas en Sistemas Complejos satisfactoriamente y garantizar la fluidez de las sesiones y el cumplimiento de los objetivos planteados, el proceso del seminario se realizó en tres etapas: planeación, ejecución y finalización, los cuales dependen uno del otro para poder avanzar en el proceso; los cuales se explicaran en los siguientes capítulos.

2. PLANEACIÓN

En esta etapa se establecen los lineamientos bajo los cuales se desarrolla el seminario, se definen los alcances y resultados que se esperan obtener mediante

la selección del tema, el estudio bibliográfico, los subtemas; y la planificación de las sesiones.

2.1. OBJETIVOS DEL SEMINARIO

- Elaborar un documento síntesis original donde se recopile los resultados de la investigación realizada en cada una de las sesiones en base a los resúmenes, relatorías y protocolos establecidos en el seminario, de acuerdo con los subtemas especificados.
- Cumplir con los roles asignados a cada integrante del seminario de investigación.
- Elaborar presentaciones en PowerPoint y documentos síntesis de cada uno de los temas investigados, obtenidos de los las relatorías y los protocolos recopilados de cada sesión.
- Presentar el estado del arte con respecto a las metodologías usadas con Data Driven.
- Elaborar un documento base para iniciar una línea de investigación en la escuela.

2.2. ESTUDIO Y SELECCIÓN DEL MATERIAL BIBLIOGRÁFICO

Para llevar a cabo la selección de los temas tratados dentro del seminario de investigación, se desarrolló en base al interés común de los estudiantes y profesor encargado. Empezando por la búsqueda de aplicaciones innovadoras y exitosas en la industria, las cuales se encontraban relacionadas con la formación académica de los integrantes y cumplían para ser tratadas en un seminario de investigación, con el aval del director; él cual cuenta con capacidad de orientación y apoyo en su perfil para la dirección.

Como aspecto fundamental del seminario se definió la bibliografía en la cual se soportarían los temas a tratar, tales como: libros, revistas, páginas web, monografías y aplicaciones específicas de la industria; aparte se reforzaron conceptos básicos matemáticos- computacionales como tareas propias, recurriendo a otras fuentes que aportaran claridad en la investigación de los temas. Para dicho propósito se hizo consulta de los libros y recursos especificados en la bibliografía relacionada en el capítulo 6 de este libro.

2.3. TEMARIO

1. Introducción.
1.1 Perspectiva histórica de los métodos de detección y diagnóstico de fallas.
1.2 Definiciones: falla, detección, diagnóstico, confiabilidad.

1.3 Generalidades

1.3.1 Diagnostico de fallas basado en modelos, (enfoque cualitativo y cuantitativo).

1.3.2 Diagnostico de fallas basado en Data-Driven, (enfoque cualitativo y cuantitativo).

1.4 Casos de aplicación de Data-Driven en el diagnostico de fallas en sistemas y aplicaciones industriales (papers).

1.5 Sensores y estrategias de medición.

1.6 Sistemas procesamiento de señales y direccionamiento de bases de datos.

2. Estadística multivariable y patrones de clasificación

2.1 Introducción

2.2 Pre tratamiento de datos.

2.3 Monitoreo estadístico univariante.

2.4 Estadística T^2 .

2.5 Requerimientos de datos.

2.6 Análisis discriminante.

2.7 Extracción de características.

3. Métodos usados con Data-Driven.

3.1 Análisis de componentes principales.

3.1.1 Introducción

3.1.2 Análisis de componentes principales (PCA).

3.1.3 Reducción de orden.

3.1.4 Detección de fallas.

3.1.5 Identificación de fallas.

3.1.6 Diagnostico de fallas.

3.1.7 PCA dinámico.

3.1.8 Otros métodos basados en PCA.

3.2 Análisis discriminante de Fisher.

3.2.1 Introducción

3.2.2 Análisis discriminante de Fisher (FDA).

3.2.3 Reducción de orden.

3.2.4 Diagnostico y detección de fallas.

3.2.5 Comparación de PCA y FDA.

3.2.6 FDA dinámico.

3.3 Mínimos cuadrados parciales (PLS).

3.3.1 Introducción

3.3.2 Algoritmos PLS.

3.3.3 Reducción de orden y predicción PLS.

3.3.4 Detección, identificación y diagnóstico de fallas.

3.3.5 Comparación de PCA y PLS.

3.3.6 Otros métodos con PLS.

3.4 Análisis de variación canónica (CVA).

3.4.1 Introducción

3.4.2 Teorema del (CVA).

3.4.3 Algoritmos CVA.

3.4.4 Modelo del estado espacial e identificación del sistema

3.4.5 Computación y selección del orden retardado (lagorder).

3.4.6 Selección del orden de estado y criterio de información de Akaike.

3.4.7 Algoritmos de interpretación su espaciales.

3.4.8 Monitoreo estadístico del proceso.

3.5 Aplicaciones

3.5.1 Introducción.

3.5.2 Conjunto de datos.

3.5.3 Intervalo del muestreo.

3.5.4 Tamaño de la muestra.

3.5.5. Selección de orden y retraso.

3.5.6 Detección de fallas.

3.5.7 Identificación de fallas.

3.5.8 Diagnostico de fallas.

3.6 Métodos basados en conocimiento.

3.6.1 Introducción.

3.6.2 Análisis casual.

3.6.2.1 Signedirectedgraph.

3.6.2.2 Modelo de árbol de síntomas.

3.6.3 Sistemas expertos.

3.6.3.1 Sistema experto de conocimiento superficial.

3.6.3.2 Sistema experto de conocimiento profundo.

3.6.3.3 Combinación de los sistemas expertos de conocimiento superficial y profundo.

3.6.3.4 Técnicas de aprendizaje de la maquina.

3.6.3.5. Representación del conocimiento.

3.6.3.6 Máquina de inferencia.

3.6.4 Reconocimiento de patrones

3.6.4.1 Redes neuronales artificiales.

3.6.4.2 Mapa de auto organización.

3.6.5 Combinación de varias técnicas.

3.6.5.1 Sistemas expertos y redes neuronales.

3.6.5.2 Lógica Fuzzy.

3.6.5.3 Sistemas expertos Fuzzy.

3.6.5.4 Redes neuronales Fuzzy.

3.6.5.5 Fuzzy directed graph.

3.6.5.6 Lógica Fuzzy y enfoque analítico.

2.4 PLANEACIÓN DE LAS SESIONES

Para realizar el número de sesiones se repartió de acuerdo a los subtemas relacionados, asignándose a cada participante de manera rotativa un subtema para su exposición y análisis, con el objetivo de que los temas se trataran de manera solvente, también se definieron los roles de cada participante (Relator, Correlator, Protocolante, Asistente), dichos roles se rotan y especifican como se muestra en la tabla 1.

Tabla 1. Rol asignado a los participantes del proyecto en las diferentes sesiones

Expositor	ROMULO NIÑO	MANUEL CARREÑO				LUZ ALVAREZ				RUBEN DUARTE				SERGIO MEJIA			
TEMA	A	C	P	R	D	C	P	R	D	C	P	R	D	C	P	R	D

1ª sesión	X	X				X				X				X	
2ª sesión	X			X				X	X				X		
3ª sesión	X				X	X				X				X	
4ª sesión	X		X				X				X	X			
5ª sesión	X		x			x				X				X	
6ª sesión	X			x				x		x			x		
7ª sesión	X	x					x					x		X	
8ª sesión	X	x						X				X		x	
9ª sesión	X				x		X				X		x		
10ª sesión	X			X		x					x				x
11ª sesión	X		x						x	x				X	
12ª sesión	X				x			X			x			x	
13ª sesión	X	X					X					X			X
14ª sesión	X			X					X	X				X	
15ª sesión	X				X	X					X			X	
16ª sesión	X		X					X					X	X	X

Fuente: Autores

También se fijaron las fechas de cada sesión con el fin que hubiera un lapso de tiempo prudente para la preparación de los subtemas. Las fechas de las ponencias se seleccionaron según la disponibilidad de los miembros con previa confirmación y se apartó la sala con anterioridad.

3. EJECUCIÓN

Para la realización de lo que se planeó con anterioridad en el seminario; consistió en la preparación de un documento por parte del relator de cada subtema adjunto al desarrollo de cada sesión.

3.1.EL ACTA

El acta es un documento escrito donde se hace constar la relación de los roles de los participantes de cada sesión del seminario de investigación. El documento modelo se que se utilizó a todas las sesiones realizadas como se muestra en la figura 3.1.

Estructura del acta:

- **Encabezado:** es donde va el título del seminario, acompañado del logo de la universidad al lado derecho y al lado izquierdo el logo de mecánica.
- **Numeración:** Es el número indicativo de la sesión en curso, en total se realizaron un total de 16 actas, que corresponden a los 16 temas desarrollados a lo largo del seminario.
- **Fecha:** Es la fecha en que se realizó cada sesión.
- **Lugar:** Es el sitio donde se llevó a cabo cada sesión, el cual fue la sala Cad de la escuela de ingeniería mecánica.



 INGENIERÍA MECÁNICA	Seminario de Investigación en Métodos de Diagnóstico de <i>Fallas de Sistemas Complejos</i>	
ACTA No.		
Fecha: Hora de inicio: Hora de finalización: Lugar:		
Relator: Co-relator: Discusión: Protocolo:		
TEMA A DESARROLLAR:		
SÍNTESIS DE LA SESIÓN:		
OBSERVACIONES:		
CONCLUSIONES:		
<hr/> Ing. Rómulo Niño		

Figura 1. Formato Acta

Fuente: Autores

- **Hora:** Se especifica la hora de iniciación de cada sesión.
- **Relator:** Es el encargado de exponer el tema correspondiente en la sesión del seminario. El relator tiene como misión principal enriquecer, como

resultado de su investigación y estudio, el saber de los demás, buscando por medio de una argumentación rigurosa aportar algo nuevo que permita avanzar en el conocimiento sobre el objeto de estudio.

El relator durante su exposición debe estimular e inducir el grupo a la discusión, mantener su atención y despertar interés de participación, a través de la argumentación rigurosa, explicaciones claras y el planteamiento de su punto de vista.

El relator tiene toda la libertad para organizar su material de trabajo, puede utilizar, en términos didácticos, comparaciones, relaciones, ejemplos y pruebas lógicas o experimentales; además debe abordar su exposición desde diferentes posiciones, es decir lo puede hacer desde un enfoque informativo, en el que expresa su opinión y comparte con el grupo sus propias conclusiones, con el fin de llegar a nuevos resultados. El rumbo de la relatoría depende de lo que el grupo espera de la sesión y de la naturaleza de los temas⁸.

- **Correlator:** Es la persona encargada de complementar y enriquecer la relatoría, el cual debe estar muy atento a la exposición del relator, pues no debe repetir lo expuesto, sino evaluar el contenido, la claridad y el manejo de la exposición, con el fin de plantear su reflexión personal, y su posición frente a la relatoría, de una manera crítica, destacando los aportes más relevantes de la relatoría motivando a los participantes al debate.
- **Discusión:** Es el medio a través del cual los participantes con base en lo expuesto en la relatoría y la argumentación del correlator, confrontan sus

⁸Nestor H. Bravo Salinas. El seminario investigativo. El seminario como práctica pedagógica para la formación integral.

concepciones y apreciaciones sobre el tema. Esta dinámica se desarrolla a manera de discusión, la cual se caracteriza por ser de carácter positivo, argumentativo, y enfocado al enriquecimiento del trabajo en grupo. La discusión requiere participación oportuna y respeto por las opiniones de los demás, criticándolas constructivamente. Este ejercicio permite a los participantes fortalecer procesos cognitivos a partir de la escucha atenta, la reflexión, la toma de notas sobre lo que se desea expresar para luego hablar.

- **Protocolo:** El protocolo es la evidencia escrita de cada sesión del seminario y un insumo para la elaboración del informe final. Éste debe ser leído antes de comenzar cada sesión, ya que es un acta de registro de lo ocurrido en la anterior y de los compromisos planteados. Su autor es el protocolante, el cual plasma en un documento las actividades tal y como fueron desarrolladas, utilizando una redacción clara y concisa.

El protocolo, como documento testimonial, debe recoger el tema, día, duración, asistencia, funciones y tareas cumplidas (síntesis de la relatoría y correlatoría), desarrollo de la discusión, aportes y tareas cumplidas (síntesis de la relatoría y correlatoría), desarrollo de la discusión, aportes de los participantes, interrogantes nuevos sobre el tema, validaciones y acuerdos entre los miembros del seminario.

El protocolo además de documentar la relación exacta de lo ejecutado durante la sesión, también juzga y evalúa su desarrollo integral registrando sus aportes en el protocolo, el cual debe ser revisado por el Director de seminario y el Relator antes de iniciar la sesión, teniendo en cuenta que cada integrante del grupo debe tener una copia de este.

- **Observaciones y conclusiones:** Se anotaron en el acta para cada una de las sesiones, las cuales eran revisadas en la sesión posterior.

4. PRESENTACIÓN, DOCUMENTOS Y APLICACIONES

4.1. PRESENTACIONES

Las presentaciones es el documento que se elabora en formato Power Point por la persona con el rol de relator para cada sesión, como apoyo y síntesis del tema a socializar ante los demás participantes. Las presentaciones constituyen el resultado del trabajo de investigativo proferido por el relator, expuesto en un ambiente gráfico y siguiendo ciertos parámetros con el objeto de que la información contenida en ésta sea entendible por el auditorio. Por dicha razón se diseñó un formato guía para las presentaciones de este mismo seminario como se describe a continuación.

4.1.1 CARACTERÍSTICAS

Para las exposiciones, las diapositivas le posibilitan al relator un manejo interactivo y dinámico con el fin de mantener la atención de las personas a las que se dirige. Los aspectos fundamentales de esta herramienta son los siguientes:

1. Servir como guía al expositor
2. Evitar o minimizar el uso de apuntes
3. Dar un orden lógico a la presentación
4. Favorecer la captación del mensaje

A continuación se exponen las características que ayudan a cumplir con estos objetivos:

- La comunicación humana es verbal y corporal. Mientras las palabras envían un mensaje, el cuerpo emite mensajes adicionales que pueden ser complementarios, afirmativos o contradictorios. El balance entre ambas formas de comunicación acompañadas con la exposición de imágenes, gráficas y textos favorecen que el mensaje llegue, por alguna de las vías de percepción del ser humano. Es por ello que el formato resulta relevante. El color blanco ayuda a que las personas que estén presentes durante la presentación centre su atención y no cansen la vista.
- Su identificación está siempre presente por lo que incluye el logo del seminario, el cual fue diseñado por los integrantes, también el logo de la universidad y de la escuela, lo que da sentido de pertenencia.
- El color negro de la letra resalta sobre el fondo claro haciendo atractivo el mensaje y evita la fatiga al espectador, pues las sesiones son de larga duración y por el cansancio es fácil distraerse.
- Se establecieron márgenes en las diapositivas que restringen el uso de textos recargados y quién elabora la diapositiva se ve en la tarea de preparar bien dicha presentación al incluir ideas claras y concretas.
- Se utilizó un tipo de letra y tamaño estándar y adecuado para facilitar la lectura.

- Las imágenes cumplen una función nemotécnica, son más atractivas que las palabras y los números, por lo que son más fácil recordadas por quienes las observan.
- El título del tema está presente en cada diapositiva para que el auditorio siempre sepa que tema se está desarrollando y no pierda el hilo de la exposición.
- La información se organiza de forma lógica, dividida en ideas sencillas que facilitan su exposición.

La totalidad de las diapositivas resultantes de las sesiones en el transcurso del seminario están recopiladas en forma digital, debido al gran número y anexadas al informe.

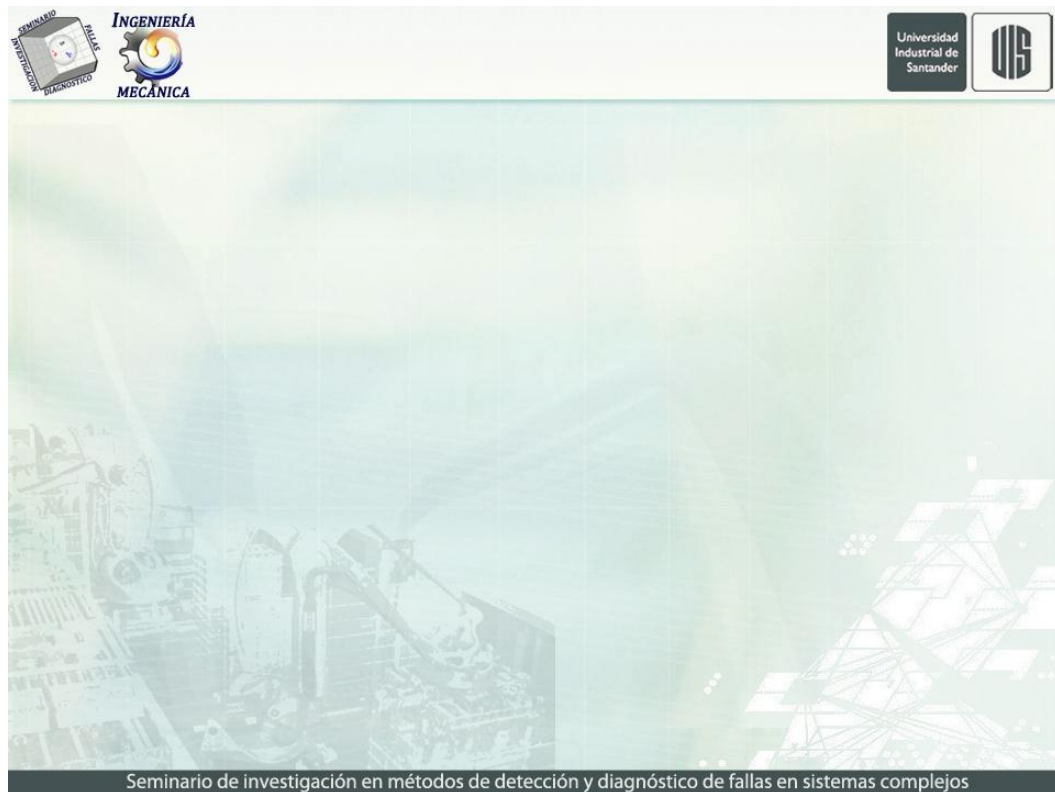


Figura 2. Formato de la presentación

Fuente: autores

4.2 DOCUMENTO

La elaboración del documento se realizó bajo las normas técnicas y metodologías científicas para la presentación de trabajos investigativos, el cual se fue revisando bajo la premisa de la discusión argumentativa, y de fundamentos que se fueron desarrollando a medida que se avanzaba en los temas, por lo que se resalta la asesoría del director en este tipo de proceso comunicativo y cognoscitivo.

El documento está dividido en diez (10) capítulos los cuales están agrupados en tres partes y en donde se recopilan todos los subtemas expuestos durante las sesiones según la tabla x siguiendo dicho orden para que haya continuidad y conexión entre ellos.

En el desarrollo de la investigación y documentación se evaluaron los métodos de detección y diagnóstico de fallas con el objeto de reconocer la aplicabilidad de cada uno sus enfoques y los nuevos caminos de investigación que han surgido hasta la actualidad, con sus posibles campos de acción en la ingeniería.

El documento se elaboró en Microsoft office Word 2010 el cual presenta un formato estándar de facto⁹ con el que pueden transferirse textos con formato o sin formato, o imágenes siendo preferido por muchos usuarios antes que otras

⁵ Norma generalmente usada y aceptada por iniciativa propia en un gran número de usuarios.

opciones, es por esto que se decidió realizar el documento en este formato de texto. Para el desarrollo de diagramas, gráficas y esquemas se utilizaron programas como Microsoft Visio el cual es un software de dibujo vectorial, que en facilita realizar diagramas de flujo de programas y UML¹⁰.

4.3 APLICACIONES EN MATLAB

4.3.1 DISTOOL

Es una herramienta en MATLAB que permite visualizar de forma gráfica las características de cada distribución con la posibilidad de variar sus parámetros. Las funciones que muestra son PDF¹¹ y CDF¹².

Al activar este comando.

```
>>distool
```

¹⁰ Lenguaje unificado de modelado. Es un lenguaje gráfico para visualizar, especificar, construir y documentar un sistema.

¹¹ (función de densidad de probabilidad) Al consignar valores de una variable aleatoria es posible desarrollar una función matemática que asigne una probabilidad a cada realización x de la variable aleatoria X .

¹² (función de densidad acumulativa) es la probabilidad X sea menor o igual a un valor específico de x .

Se despliega una ventana donde podemos elegir el tipo de distribución además de sus parámetros y ver su PDF o su CDF.

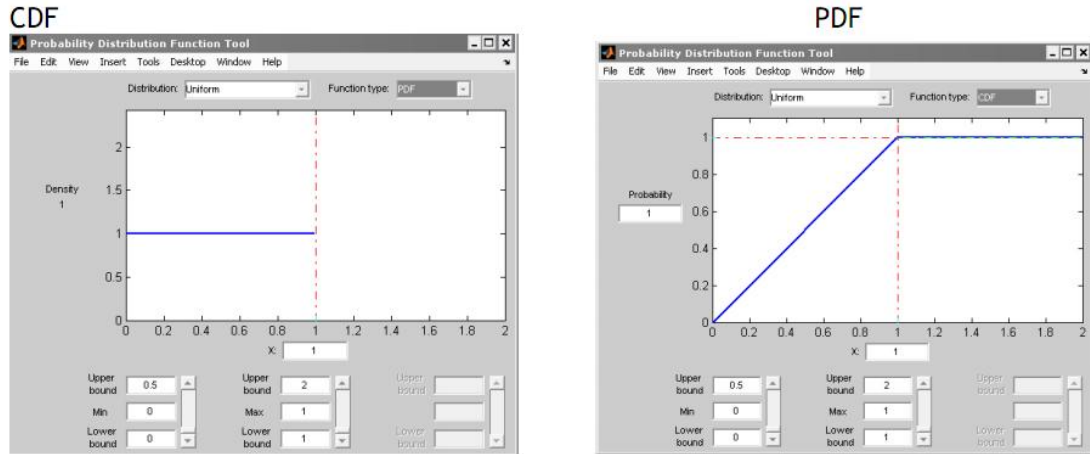


Figura 3. Vista de la Interfaz en Matlab de la herramienta Distool

Fuente: Matlab

4.3.2 DISTRIBUCIÓN UNIFORME

Ocurre en un evento donde la variable aleatoria toma valores de un intervalo finito, de manera que éstos se encuentran distribuidos igualmente sobre el intervalo. Esto es, la probabilidad de que la variable aleatoria tome un valor en cada sub-intervalo de igual longitud es la misma.

Computa la función uniforme continua para el valor X y los parámetros A y B . El resultado Y es la probabilidad de que ocurra X en el intervalo (A,B) .

- **Ejemplo 1.** Considerar la cantidad de agua consumida al mes en una vivienda, medida en m^3 , sigue una distribución uniforme continua en el intervalo (20,40). Calcular la probabilidad de que se consuman: (a) $28 m^3$, (b) $35 m^3$, (c) entre 25 y $35 m^3$, (d) más de $28 m^3$.

```

clc ;
clear all;
close all;
%-----Ejemplo distribución uniforme-----
%Del enunciado obtenemos los extremos del intervalo
%es decir A=20 Y B=40
%(a) se requiere encontrar la probabilidad P(X=28)
Y1 = unifpdf(28,20,40) %---comando distribución uniforme
%(b) menos de 35m3
Y2 = unifcdf(35,20,40) %---se evalua la para CDF menos de 35m3
%(c) entre 25 y 35 m3
Y3 = unifcdf(35,20,40)-unifcdf(25,20,40) %se evalua para la CDF
                                         %el límite superior 35
                                         %menos el límite inferior 25

%(d) mas de 28m3
Y4 = 1-unifcdf(28,20,40)

%se despliegue la grafica de la función de distribución acumulada

X= 20:40;
Y= unifcdf(X,20,40);
plot(X,Y);
grid

```

Y1 =	Probabilidad de que ocurra 28	Y1=5%
0.0500		
Y2 =	Menos de 35	Y2= 75%
0.7500		
Y3 =	Probabilidad de que esté entre 25 y 35	Y3=75%
0.5000		
Y4 =	Más de 28	Y4=60%
0.6000		

Y se despliega la gráfica de la función de distribución acumulada.

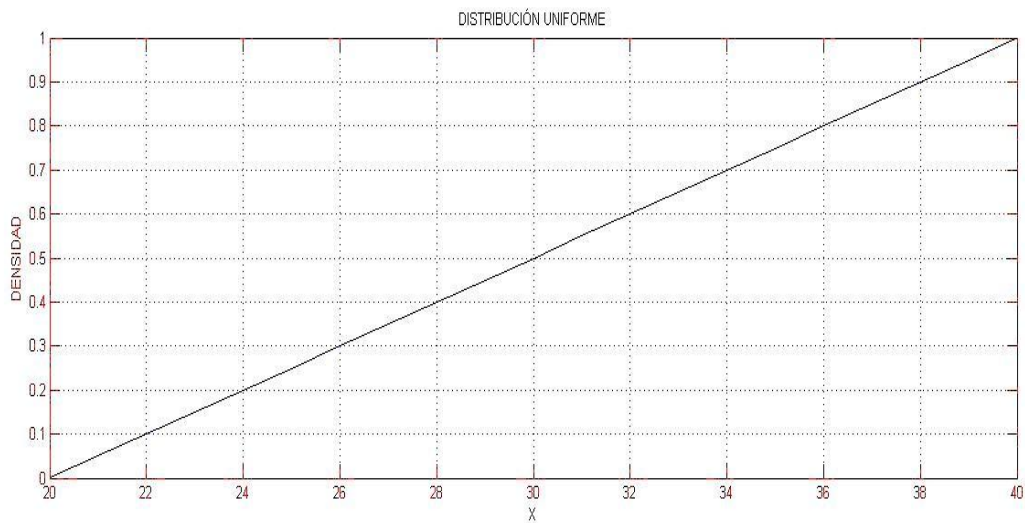


Figura 4. gráfica de la función uniforme del ejemplo anterior.

Fuente: Matlab

4.3.3 DISTRIBUCIÓN NORMAL (GAUSIANA)

Computa la función exponencial negativa para el valor X y los parámetros $(\mu)MU$ y $(\sigma)SIGMA$.¹³ X , MU y $SIGMA$ pueden ser vectores, matrices y arreglos que deben ser del mismo tamaño. El parámetro $SIGMA$ debe ser positivo.

¹³ μ (media poblacional), σ (desviación estándar poblacional) ver capítulo 8.

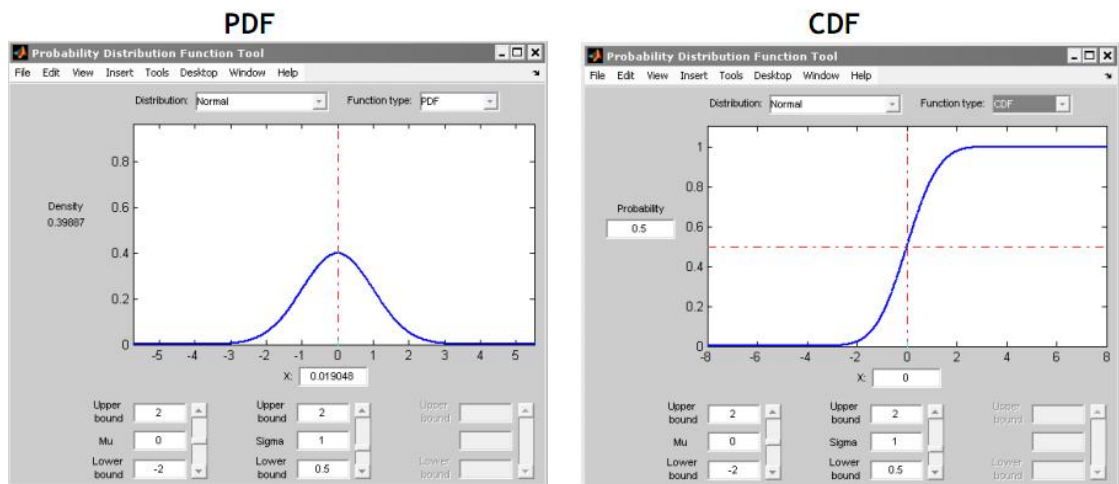


Figura 5. Vista de la Interfaz en Matlab de la herramienta Distool de la función de densidad de probabilidad.

Fuente: Matlab

Sintaxis: $Y = \text{normpdf}(X, MU, SIGMA)$

- **Ejemplo 2.** Una compañía de suministro de electricidad ha determinado que el consumo, medido en Kw/h, de una vivienda familiar durante un mes, sigue una distribución normal de media $\mu=300$ y desviación típica $\sigma=50$. Calcular la probabilidad de que se consuman:
 - 245 Kw/h,
 - entre 200 y 300 Kw/h,
 - más de 300 Kw/h,
 - ¿Qué porcentaje de viviendas familiares consumirán más de 250 Kw/h? ¿Y menos de 350?

```

clc ;
clear all;
close all;
%-----Ejemplo distribución Normal Gaussiana-----
%del enunciado tenemos que mu=200 y sigma=50 se requiere encontrar
%(a) se requiere encontrar la probabilidad P(X=245)
Y1=normpdf(245,200,50)
%(b) P entre 200 y 300
Y2= normcdf(300,200,50)
%(c) se requiere encontrar P mayor a 200
Y3= 1-normcdf(200,200,50)
%(d) se requiere encontrar P con mas de 250 y menos de 350
Y4=1-normpdf(250,200,50)
Y5=normcdf(350,200,50)

%se despliegue la grafica de la función de distribución acumulada
%para un arreglo entre 0 y 400 con una diferencia de 5
X = (1:5:400);
Y = normpdf(X,200,50); % con MU = 165 y SIGMA = 5
plot(X,Y);
grid

```

```

Y1 =
    0.0053

Y2 =
    0.4772

Y3 =
    0.5000

Y4 =
    0.1587

Y5 =
    0.9987

fx >>

```

Probabilidad de que ocurra 245Kw/h 0.53%

Entre 200 y 300Kw/h 97.7%

Más de 300 Kw/h 50%

Más de 250 Kw/h 99.5%

Menos de 350 Kw/h 99.87%

Y se despliega la gráfica de la función de distribución normal gaussiana

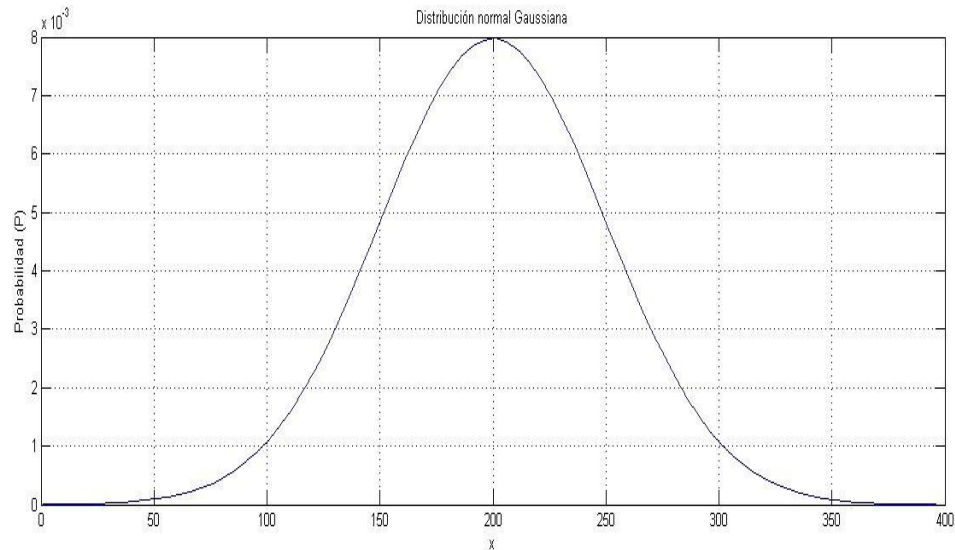


Figura 6. Distribución normal gaussiana para el ejemplo anterior

Fuente: Matlab

4.3.4 PRUEBAS DE HIPÓTESIS

Una hipótesis Estadística es una proposición sobre los parámetros de una población o sobre la distribución de probabilidad de una variable aleatoria (ver capítulo 8).

- **Ejemplo 3.** Se tiene interés en la rapidez de combustión de un agente propulsor para los sistemas de salida de emergencia en aeronaves. (esta rapidez es una variable aleatoria con alguna distribución de probabilidad). Especialmente interesa la rapidez de combustión promedio (que es un parámetro (μ) de dicha distribución). De manera más específica, interesa

decidir si esta rapidez promedio es o no 50 cm/seg. Calcular α para el ejemplo de la rapidez de combustión para una muestra de $N=10$ datos, suponiendo que la desviación estándar de la rapidez de combustión es $s=2.5$ cm/seg.

Hipótesis Nula: $H_0: m = 50$ cm/seg

Hipótesis Alternativa: $H_1: m \neq 50$ cm/seg

A la probabilidad de cometer un error de Tipo I se denota por α , y se le llama el nivel o tamaño de significancia de la prueba es decir

$$\alpha = P(\text{error Tipo I}) = P(\text{rechazar } H_0 \mid H_0 \text{ es verdadera})$$

Solución: es necesario introducir un criterio para decidir qué tanto es muy cercano, para el ejemplo este criterio pudiera ser: $48.5 \leq x \leq 51.5$, si esto ocurre se acepta H_0 en este caso $\alpha = P(x \text{ caiga en la región crítica} \mid m=50)$, es decir:

$$\alpha = P(x < 48.5) + P(x > 51.5)$$

Recordando que La distribución de x es Normal con media $m=50$ y desviación estándar $s/\sqrt{N} = 0.79$, por lo tanto, usando Matlab:

```
clc ;
clear all;
close all;
%-----Ejemplo de prueba de hipótesis-----
%la probabilidad de cometer un error de Tipo I se halla de la siguiente
%manera con media mu=50 y desviación estandar 0.79

a= normcdf(48.5,50,0.79) + (1-normcdf(51.5,50,0.79))

a =

    0.0576

>>
```

- ❖ Esto significa que el 5.76% de las muestras de tamaño 10 conducirán al rechazo de la Hipótesis H_0 : $m=50$ cm/seg, cuando ésta es verdadera.

4.3.5 FUNDAMENTOS EN ALGEBRA LINEAL

En esta sección se repasa conceptos básicos de algebra lineal. No se puede verdaderamente entender la mayoría de las técnicas de FDD.

- Escalares, Vectores y Matrices

Un *escalar* es una cantidad matemática que es completamente descrita por una magnitud es decir un número solo. La variables escalar son generalmente denotadas por una letra minúscula, ej. a . Ejemplos de variables escalares incluyen temperatura, densidad, presión y flujo. En Matlab:

```
>> a=5
a =
    5
fx >>
```

Un *vector* es una cantidad matemática que es completamente descrita por su magnitud y dirección un ejemplo de un vector columna tridimensional puede ser

$$b = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix} \text{ob } = [4; 3; 5];$$

En Matlab este vector puede ser ingresado de la siguiente manera:

```
>> b = [4
3
5]

b =

     4
     3
     5

fx >>
```

La traspuesta de un vector columna es un vector fila y viceversa. La traspuesta es generalmente indicada por el superíndice T es decir T , mientras en otras instancias, incluyendo Matlab, se usa un apóstrofe. Por ejemplo

$$\mathbf{b}^T = [4 \ 3 \ 5]$$

En Matlab, se puede asignar fácilmente \mathbf{b}^T a otra variable \mathbf{c} , así:

```
>> c = b'

c =

     4     3     5

fx >>
```

Una *matriz* es un arreglo escalar, o en algunas instancias, expresiones algebraicas que evalúan escalares.

$$A = \begin{bmatrix} 2 & 5 & 3 & 6 \\ 7 & 3 & 2 & 1 \\ 5 & 2 & 0 & 3 \end{bmatrix}$$

Esta matriz puede ser añadida en Matlab así:

```
»A = [2 5 3 6; 7 3 2 1; 5 2 0 3];
```

- **Multiplicación y Adición de vectores y matrices**

Las matrices pueden ser sumadas, siempre que las dimensiones sean consistentes, es decir, ambas matrices deben tener el mismo número de filas y columnas. Para sumar dos matrices hay que añadir los elementos correspondientes

$$\begin{bmatrix} 1 & 4 & 3 \\ 5 & 4 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 4 & 1 \\ 2 & 6 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 8 & 4 \\ 7 & 10 & 3 \end{bmatrix}$$

En Matlab se hace así:

```
>> x = [1 4 3; 5 4 0]; y = [2 4 1; 2 6 3];
x + y

ans =

     3     8     4
     7    10     3

fx >>
```

La suma de vectores y matrices es asociativa, siempre que **A**, **B** y **C** son matrices con el mismo número de filas y columnas, entonces: **A + B = B + A** como **(A + B) + C = A + (B + C)**.

Multiplicación de una matriz por una escalar en Matlab es hecha usando el operador *

```
>> A= [1 4 3; 5 4 0;2 1 4];
>> c = 2;
c*A

ans =

     2     8     6
    10     8     0
     4     2     8

fx >>
```

Vectores pueden ser multiplicados juntos de dos diferentes maneras. El producto interno de dos vectores deberán tener el mismo número de elementos o dimensión. El producto interno es la suma sobre los productos de los correspondientes elementos, y es así:

$$\mathbf{a}^T \mathbf{b} = [2 \ 5 \ 1] \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix} = [2 * 4 + 5 * 3 + 1 * 5] = 28$$

```
>> a = [2; 5; 1]; b = [4; 3; 5];
a'*b

ans =

    28

fx >> |
```

El producto interno es también usado cuando se calcula la longitud de un vector. La longitud de un vector es la raíz cuadrada de la suma de cuadrados de los coeficientes del vector. La longitud de un vector \mathbf{a} , denotada por $\|\mathbf{a}\|$ es el producto cuadrado del vector por si mismo $\|\mathbf{a}\| = \sqrt{\mathbf{a}^T \mathbf{a}}$. En matlab.

```
>> a = [2; 5; 1];
>> sqrt(a'*a)

ans =

    5.4772

fx >>
```

- **Vectores ortonormales y ortogonales**

Un vector es ortogonal si su producto interno es cero. La interpretación geométrica de esto es que forman un ángulo recto entre sí o perpendicular. Los vectores son ortonormales si son ortogonales y de la unidad de longitud, es decir si su producto

interno entre si es la unidad. Para un conjunto ortonormal de vectores $\mathbf{v}_i^T \mathbf{v}_j =$
 $\begin{cases} 0 \text{ para } i \neq j \\ 1 \text{ para } i = j \end{cases}$

- **Eliminación Gaussiana: Resolviendo sistemas de ecuaciones**

La eliminación gaussiana es un método consistentemente usado para resolver sistemas de ecuaciones. Considerar el siguiente sistema de tres ecuaciones

$$\begin{bmatrix} 2b_1 & b_2 & b_3 \\ 4b_1 & b_2 & \\ -2b_1 & 2b_2 & b_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 7 \end{bmatrix}$$

El sistema puede ser también escrito usando multiplicación de matrices como sigue:

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -7 \end{bmatrix}$$

Y se simplifica como $\mathbf{Xb=y}$ donde

$$\mathbf{X} = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \text{ y } \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ -7 \end{bmatrix}$$

El problema es encontrar los valores de b_1 , b_2 y b_3 que hace el sistema de ecuaciones mantener. Para hacerlo podemos aplicar eliminación gaussiana. En Matlab la eliminación gaussiana es hecha con el operador divisor izquierdo:

```

>> X = [2 1 1; 4 1 0; -2 2 1]; y = [1; -2; 7];
b = X\y

b =

   -1.0000
    2.0000
    1.0000

fx >>

```

Cuando uno de los cero, en este caso, los coeficientes bajo el pivote no pueden ser eliminados por la sustracción de un múltiplo de la ecuación pivote. En este caso la matriz se llama *singular* y no hay esperanza para una solución única solución al problema ej:

```

>> X = [1 3 2; 2 6 9; 3 9 8]; y = [1; -4; -4];
b = X\y
Warning: Matrix is singular to working precision.

b =

   -Inf
    Inf
   -3.5000

fx >>

```

Para esto se halla el rango de la matriz. Nótese que este procedimiento puede ser hecho en cualquier matriz; no se necesita que sea cuadrada. Se muestra que el rango de una matriz m por n debe ser menos que, o igual que, el menor entre m o n , es decir:

$$\text{rank}(x) \leq \min(m, n) \quad (4.2)$$

Una matriz cuyo rango es igual al menor de m o n se dice que es de rango completo. Si el rango es menor que $\min(m, n)$, la matriz es de rango deficiente o colineal; en Matlab se hace así.

```

1     3     2
2     6     9
3     9     8
|
>> rank(X)

ans =

     2

fx >>

```

- **Dependencia e independencia lineal**

Se requiere en ocasiones determinar si un conjunto de vectores de un espacio vectorial es linealmente independiente o dependiente, se halle el determinante de la matriz conformada por las componentes de los vectores

- **Ejemplo 4.** Estudiar si los vectores $\{(2,3,-1),(0,0,1),(2,1,0)\}$ forman una base de espacio vectorial \mathcal{R}^3 .

```

2 -   clear all;
3 -   close all;
4     %-----Ejemplo Dependencia e independencia lineal-----
5     %Estudiar si los vectores {(2,3,-1),(0,0,1),(2,1,0)} forman una base de espacio vector
6     %solucion
7     %Calcularemos el valor del determinante de la matriz cuyas filas (o
8     %columnas) son las coordenadas de los vectores dados.
9     X= [2,3,-1; 0,0,1 ; 2,1,0];
10    det(X)
11
ans =

     4

fx >>

```

El determinante es no nulo, luego el conjunto es linealmente independiente y por tanto base, ya que sabemos que la dimensión del espacio es 3.

- **Matriz de covarianza**

Si x es un vector $\text{cov}(x)$, retorna la varianza. Para matrices, donde cada fila es una observación, y cada columna es una variable

```

>> X = [1 1 2 3;7 8 9 7;1 3 2 1;10 9 11 9;2 2 1 1; 3 4 1 2] ;
>> cov(X)

ans =

    13.6000    11.6000    15.6000    11.8000
    11.6000    10.7000    13.6000     9.9000
    15.6000    13.6000    19.8667    14.6667
    11.8000     9.9000    14.6667    11.3667
fx >>

```

- **Eigenvalores**

El comando **eig** calcula los valores propios de una matriz

```

>> X = [1 1 2 3;7 8 9 7;1 3 2 1;10 9 11 9;2 2 1 1; 3 4 1 2] ;
>> cov(X)

ans =

    13.6000    11.6000    15.6000    11.8000
    11.6000    10.7000    13.6000     9.9000
    15.6000    13.6000    19.8667    14.6667
    11.8000     9.9000    14.6667    11.3667

>> eig(cov(X))

ans =

    0.1590
    0.5731
    1.4018
   53.3994
fx >>

```

- **Ortogonalidad de Subespacios:**

Dos vectores cualquiera v y w son ortogonales si su producto interno es cero. Dos subespacios V y W pueden ser también ortogonales, siempre que cada vector v en V es ortogonal a cada vector w en W

$$v^T w = 0 \quad (4.3)$$

- **Proyecciones sobre líneas:**

Dado un vector y un punto definido por el vector y , a lo largo de la dirección definida por x que sea más cercana a y

$$p = \hat{b}x \quad (4.4)$$

$$x^T(y - \hat{b}x) = 0 \rightarrow x^T y = \hat{b}x^T x \rightarrow \hat{b} = \frac{x^T y}{x^T x} \quad (4.5)$$

$$p = \hat{b}x = \frac{x^T y}{x^T x} x \quad (4.6)$$

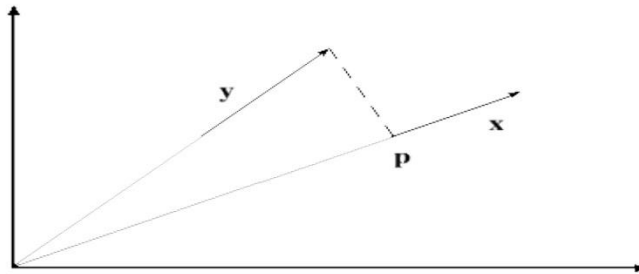


Figura 4. Proyección vector y sobre el vector x

Fuente: PLS_Toolbox 3.0 for use with MATLAB™, Barry M. Wise

Cuando $m > n$, se tienen más muestras filas que variables columnas, el sistema puede ser inconsistente... la única manera de resolver el sistema es haciendo que el error promedio en las ecuaciones sea minimizado.

El error no es más que la longitud del vector

$$xb - y \rightarrow E = \|xb - y\| \quad (4.7)$$

$$E^2 = (xb - y)^T (xb - y) = x^T x b^2 - 2x^T y b + y^T y \quad (4.8)$$

$$\frac{dE^2}{db} = 2x^T x b - 2x^T y = 0 \quad (4.9)$$

La solución con mínimos cuadrados es

$$\hat{b} = \frac{x^T y}{x^T x} \quad (4.10)$$

Lo que se busca es determinar a $\hat{\mathbf{b}}$ tal que se pueda utilizar en nuevas muestras, X_{new} , para estimar sus propiedades $\hat{\mathbf{y}}_{new}$:

$$\hat{\mathbf{y}}_{new} = X_{new} \hat{\mathbf{b}} \quad (4.11)$$

En Matlab se hace de la siguiente manera:

```
>> X = [1 1; 1 2; 2 1; 2 2]; y = [6 6 7 11]';
b = inv(X'*X)*X'*y

b =

     3
     2

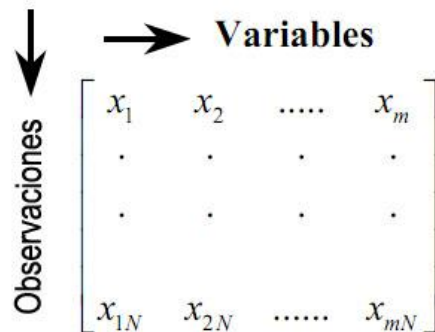
fx >>
```

Entonces nuestro modelo de estimación es:

$$y = 3x_1 + 2x_2 \quad (4.12)$$

- **Ejemplo de procedimiento básico para PCA con Matlab**

El conjunto de datos X consiste en N observaciones y m variables. En nuestro ejemplo $m < N$, y el conjunto de datos está formado de la siguiente manera.



Suponer que el conjunto de datos es $X = \begin{bmatrix} 118 & 30 \\ 119 & 32 \\ 120 & 33 \\ 120 & 31 \\ 121 & 32 \\ 122 & 34 \end{bmatrix}$ con dos variables y 6

observaciones

```

1 -   clc ;
2 -   clear all;
3 -   close all;
4 -   %-----Ejemplo de procedimiento básico para PCA con Matlab-----
5 -   N=6;
6 -   X=[118 30; 119 32; 120 33; 120 31; 121 32; 122 34];
7 -   Xmean=mean(X);%da la media de X
8 -   ZX=X-ones(N,1)*Xmean;% Para obtener la varianza
9 -   Xstd=std(X);% La desviación estandar para cada variable en X
10 -  SX=ZX./ (ones(N,1)*Xstd);% Escala_ losdatos sinprocesar por desviaciónEstandar
11
12   %Construccion de modelo
13 -  C=SX'*SX/(N-1); %matriz de covarianza X
14 -  [eigvec,lamda]=eigs(C, 2) % Donde lamda es una matriz en orden decreciente
15 -  Cusm=cumsum(diag(lamda))% la suma acumulada de los eigenvalues
16 -  Percentage=[];
17 -  for i=1:length(Cusm)
18 -      Percentage=[Percentage,Cusm(i)/Cusm(end)*100];% porcentaje
19 -  end
20 -  npc = 1; % En Este caso, un componente principal es seleccionado del 90%
21 -  Vk = eigvec(:,1:npc);%la matriz transformada Vk
22 -  sigma = lamda(1:npc,1:npc); % sigma
23 -  %b calculo del coeficiente limite para el hoteling T2
24 -  T2lim=np*(N-1)/(N-npc)*finv(0.95,npc,N-npc); % alfa=0.95 como nivel de confianza
25 -  T2lim_1 = sqrt(sigma(1,1))*tinv(0.975,N-1); % para el componente principal inicial
26 -                                     %Para el indice estadístico SPE se calcula
27 -  zeta1=lamda(2,2); % el segundo eigen valor
28 -  zeta2=lamda(2,2)^2; % el cuadrado del segundo eigen valor
29 -  zeta3=lamda(2,2)^3; % el poder del tercer eigen valor
30 -  h0=1-2*zeta1*zeta3/3/zeta2^2;
31 -  ca=norminv(0.95, 0, 1); % ca es un valor para la distribución normal estandar 95% de
32 -                                     %como nivel de confianza

```

```

33 - SPElim=zeta1*(ca*h0*sqrt(2*zeta2)/zeta1+1+zeta2*h0*(h0-1)/zeta1^2)^(1/h0);
34
35 %test para los nuevos datos
36 - Xnew1=[119; 31];
37 - Xnew2=[123; 35];
38 - Xnew3=[118; 43];
39 %se escalan los nuevos datos
40 - Sxnew1=(Xnew1-Xmean')./Xstd'; % porque Xmean es un vector columna
41 - Sxnew2=(Xnew2-Xmean')./Xstd';
42 - Sxnew3=(Xnew3-Xmean')./Xstd';
43 %calculo del score usando la matriz transformada
44 - t1=Vk'* Sxnew1;
45 - t2=Vk'* Sxnew2;
46 - t3=Vk'* Sxnew3;
47
48 %c) calculo del indice de hoteling
49 - T21=Sxnew1'*Vk* inv(sigma)*Vk'*Sxnew1;
50 - T22=Sxnew2'*Vk* inv(sigma)*Vk'*Sxnew2;
51 - T23=Sxnew3'*Vk* inv(sigma)*Vk'*Sxnew3;
52

```

Name ^	Value	Min	Max
C	[1.0000,0.8000;0.8000,...	0.8000	1.0000
Cusm	[1.8000;2.0000]	1.8000	2.0000
N	6	6	6
Percentage	[90,100]	90	100
SPElim	0.7494	0.7494	0.7494
SX	<6x2 double>	-1.4142	1.4142
Sxnew1	[-0.7071;-0.7071]	-0.7071	-0.7071
Sxnew2	[2.1213;2.1213]	2.1213	2.1213
Sxnew3	[-1.4142;7.7782]	-1.4142	7.7782
T21	0.5556	0.5556	0.5556
T22	5	5	5
T23	11.2500	11.2500	11.2500
T2lim	6.6079	6.6079	6.6079
T2lim_1	3.4488	3.4488	3.4488
Vk	[0.7071;0.7071]	0.7071	0.7071
X	<6x2 double>	30	122
Xmean	[120,32]	32	120
Xnew1	[119;31]	31	119
Xnew2	[123;35]	35	123
Xnew3	[118;43]	43	118
Xstd	[1.4142,1.4142]	1.4142	1.4142
ZX	<6x2 double>	-2	2
ca	1.6449	1.6449	1.6449
eigvec	[0.7071,-0.7071;0.707...	-0.7071	0.7071
h0	0.3333	0.3333	0.3333
i	2	2	2
lamda	[1.8000,0;0,0.2000]	0	1.8000
npc	1	1	1
sigma	1.8000	1.8000	1.8000
t1	-1.0000	-1.0000	-1.0000
t2	3.0000	3.0000	3.0000
t3	4.5000	4.5000	4.5000
zeta1	0.2000	0.2000	0.2000
zeta2	0.0400	0.0400	0.0400
zeta3	0.0080	0.0080	0.0080

Comparando los score calculados o índices y sus límites correspondiente. El límite de hotelling T2 es 6.6 para nuestro caso y el índice T2 para los tres nuevos datos son $T_{21}= 0.55$, $T_{22}=5$ Y $T_{23}=11,25$

Podemos decir que el tercer conjunto de datos se colectan cuando el proceso está fallando mientras los otros dos conjunto de datos son recogidos cuando el proceso está en buenas condiciones.

5. CONCLUSIONES Y RECOMENDACIONES

El seminario de investigación proporciona un ambiente propicio para el desarrollo de las capacidades de interpretación, argumentación, deducción análisis, y búsqueda de nuevas perspectivas o puntos de vista que tengan repercusión en el desarrollo de nuestra sociedad; además fomenta el trabajo en equipo e interdisciplinar.

En este trabajo de investigación cumple con la necesidad de recopilar las metodologías que se proponen en la actualidad, como respuesta a una problemática existente en los sistemas complejos; el volumen de datos, y su imposibilidad o extrema dificultad de concebir modelos que prevean fallas o anomalías en los mismos.

Los tópicos que componen el seminario, a pesar de ser de naturaleza interdisciplinar se han enfocado en el área de mantenimiento, control de procesos, y los modos de falla; debido al manejo de datos se vuelve indispensable el manejo de estadística multivariante, probabilidad, haciendo uso de los recursos computacionales, y dando una visión de los nuevos enfoques como como redes neuronales e inteligencia artificial.

El área de investigación en métodos de detección y diagnóstico de fallas es un tema de creciente interés, por su aplicación en las grandes industrias

colombianas; por lo cual el seminario generó bases conceptuales que permiten la integración entre la academia y el desarrollo tecnológico de la región.

6. BIBLIOGRAFÍA

- BLÁZQUEZ L. Felipe. DEMIGUEI Luis J. I. Diagnostico automático de fallos para sistemas dinámicos no lineales.
- CAMACHO Oscar. PADILLA Delfina . PADILLA José L. Fault diagnosis based on multivariate statistical techniques.
- CHIANG, L.H. RUSSELL, E.L. and BRAATZ, R.D. Fault Detection and Diagnosis in Industrial systems. Springer. 2001.
- CINAR Ali, PALAZOGLU Ahmet. Chemical Process Performance Evaluation. CRC Press 2007.
- DASH, Sourabh and VENKATASUBRAMANIAN, Venkat. Challenges In the Industrial Applications of Fault Diagnostic Systems.
- DING, Steven X. Model-Based Fault Diagnosis Techniques, Design Schemes, Algorithms, and Tools. Springer. 2008.
- GERTLER . Janos J. Fault Detection and Diagnosis in Engineering Systems. Marcel Dekker Inc. 1998.
- MAURYA, M. R. RENGASWAMY, R. and VENKATASUBRAMANIAN, V. Fault Diagnosis by Qualitative Trend Analysis of the Principal Components. Ichem. 2005.

- MAURYA, Ram Mano. RENGASWAMY, Raghunathan. VENKATASUBRAMANIAN, Venkat. A framework for On-line Trend Extraction and Fault Diagnosis.2003.
- PLA, Laura E. Análisis multivariado. Método de componentes principales. Universidad Nacional Experimental Francisco de Miranda 1986.
- RODRIGUEZ LLANOS, David Alejandro. Time Misalignments in Fault Detection and Diagnosis. Universitat de Girona. 2008.
- VACHTSEVANOS, George. LEWIS, L. Frank. ROEMER, Michael. HESS, Andrew. Wu, Biqing. Intelligent Fault Diagnosis and Prognosis for Engineering Systems. John Wiley & Sons, INC. 2006.
- VENKATASUBRAMANIAN, Venkat. RENGASWAMY, Raghunathan. YIN, Kewen. N. Surya. A ,Kavuri. review of process fault detection and diagnosis Part 1: Quantitative model-based methods. 2002. Elsevier.
- VENKATASUBRAMANIAN, Venkat. RENGASWAMY, Raghunathan. YIN, Kewen. N. Surya. A ,Kavuri. review of process fault detection and diagnosis Part 2: Qualitative models and search strategies.. 2002. Elsevier.
- VENKATASUBRAMANIAN, Venkat. RENGASWAMY, Raghunathan. YIN, Kewen. N. Surya. A ,Kavuri. review of process fault detection and diagnosis Part 3: Process history based methods.. 2002. Elsevier.

- WANG, Hong. CHAI, Tian-YU. DING, Jin-Liang. BROWN, Martin. Data Driven Fault Diagnosis and Fault Tolerant Control: Some Advances and Possible New Directions.
- XU Jiang-XIn, HOU Zhong-sheng. Notes on Data-Driven System Approaches.
- YANG, Qingsong. Model-based and Data Driven Fault diagnosis methods with applications to process monitoring. Candidate for the Ph.D. degree. Case Western reserve University, May 2004.

ANEXOS

A.7. PERSPECTIVA HISTORICA DE LOS METODOS DE DETECCION Y DIAGNÓSTICO DE FALLAS

A.7.1 INTRODUCCION

En estas últimas décadas, los métodos de detección y diagnóstico de fallas se han convertido en un tópico de considerable interés dentro de la comunidad de investigadores pertenecientes al área de control, esto es debido a los crecientes requerimientos de operaciones confiables en los sistemas de control, las cuales se encuentran, la mayor parte de los casos, sujetos a un determinado número de fallas internas en lazos cerrados (closed loops) o fallas generadas por factores ambientales. Una vez la falla ha ocurrido, puede causar pérdidas irreparables y resultar en contaminación inaceptable; Por ejemplo, se ha observado que la presencia de fallas en sistemas petroleros han causado de un 3% a 8% de reducción en la producción de petróleo en los estados unidos, llevando a pérdidas de 20 billones de dólares en la economía del país por año. También en 1997, las fallas en la planta química de Beijing causaron grandes pérdidas directas, al igual que la explosión en la mina petroquímica de Kuwait Al Ahmedi en junio del 2000, que resulto en pérdidas de alrededor de 100 millones de dólares. Por otro lado, estadísticas industriales señalan que las fallas grandes de carácter catastrófico pueden ser infrecuentes, sin embargo las fallas menores son muy comunes, ocurriendo diariamente y generando muchas lesiones, enfermedades ocupacionales con un costo de billones de dólares cada año.

Es por ello que la detección y diagnóstico de fallas (FDD de sus siglas en inglés Fault Detection and Diagnosis) es de vital importancia para la operación segura

de las plantas industriales. Desde que las industrias petroquímicas pierden un estimado de 20 billones de dólares cada año, han categorizado la detección y diagnóstico de fallas como su problema número uno que necesita resolverse. Así que existe un considerable interés en este campo, desde los practicantes industriales hasta los investigadores académicos, contrario a lo que ocurría varias décadas atrás.

Los sistemas concernientes al FDD abarcan un amplio rango del espectro de la maquinaria creada por el hombre, incluyendo instalaciones de producción industrial (plantas de potencia, plantas químicas, refinerías de petróleo, fábricas de acero, fábricas de papel...), vehículos de transporte (barcos, aviones, automóviles) y aplicaciones domésticas (equipos de aire acondicionado y calentadores, refrigeradores, máquinas de lavado...).

Además del amplio campo de aplicación de los métodos de FDD en la industria cabe resaltar que la detección temprana de la ocurrencia de fallas es crítica para evitar el desperdicio del producto, la degradación de la eficiencia, los daños mayores en la misma máquina y deterioro en la salud humana o incluso pérdidas de vidas. El diagnóstico rápido y correcto del componente defectuoso facilita la toma de decisiones apropiadas y óptimas para acciones correctivas, de emergencia o reparación, y además puede ayudar a evadir una progresión de eventos anormales y reducir la pérdida de productividad. En conclusión la detección y aislamiento de fallas (diagnóstico) en sistemas de ingeniería es de un gran significado práctico.

A.7.1.1 HISTORIA DE LOS METODOS DE DETECCION Y DIAGNÓSTICO DE FALLAS

Desde que los humanos empezaron a construir maquinas, han estado naturalmente preocupados sobre su condición. Por siglos, la única forma de aprender acerca del mal funcionamiento y su ubicación, era mediante los sentidos biológicos (un enfoque todavía ampliamente utilizado); buscando cambios en la forma o el color, escuchando sonidos inusuales en intensidad o tono, uso del tacto para sentir el calor o la vibración, y el olor para detectar fugas o sobrecalentamiento.

Después, se introdujeron los aparatos de medición, los cuales proveían información más exacta sobre variables físicas importantes. Sin embargo estos aparatos (sensores) también mostraron ser propensos al mal funcionamiento, generando el dilema de las falsas alarmas. El potencial de ocurrir fallas en los sensores se convirtió en un tema aún mas critico cuando fueron aplicados en el control automático de máquinas, donde los efectos de tales fallas podían ser más directos y devastadores y donde el operador humano es frecuentemente removido del proceso.

Un desarrollo dramático tomo lugar con la llegada del computador (alrededor de 1939) y la proliferación de sus aplicaciones en tiempo real. El computador podía realizar una supervisión automática de la maquinaria. Su inteligencia nos permitía integrar la información originada de diversas fuentes, abriendo la posibilidad de localizar el componente defectuoso incluso en un sistema complejo, incluyendo los sensores responsables de las falsas alarmas. La velocidad de los computadores

también hizo posible captar las fallas mientras se estaban desarrollando, antes de que condujeran a interrupciones en la operación de la máquina.

Inicialmente los computadores eran complejos y costosos y necesitaban un ambiente especial de operación. Solo se podían utilizar en maquinaria que fuese lo suficientemente costosa y compleja para justificarlos. Grandes instalaciones de producción industrial, tales como plantas de potencia, refinerías de petróleo, plantas químicas, plantas de acero y máquinas para hacer papel fueron las primeras en ser equipadas con un sistema computarizado de monitoreo. A estos le siguieron aplicaciones similares en la mayoría de equipos de transporte, como los barcos y aviones. En tales sistemas una computadora central leería y supervisaría cientos o miles de mediciones de lazos cerrados (control loops). Con la llegada del microprocesador (alrededor de 1971), la situación cambió substancialmente. Los procesadores se convirtieron en meros componentes, los cuales eran lo suficientemente económicos para implementarlos en varios aparatos y operaban bajo un gran rango de condiciones ambientales. Este desarrollo se acopló con la llegada de la comunicación avanzada, resultando en la descentralización de los grandes sistemas de control y monitoreo industrial. Más importante aún, los microcomputadores aparecieron como productos generados en producción masiva al igual que los automóviles, los sistemas de aire acondicionado y calefacción e incluso electrodomésticos. Así el monitoreo computarizado, aunque invisible para muchos, se ha convertido en parte de la vida diaria de la sociedad de consumo.

Actualmente el enfoque tradicional hacia la detección y diagnóstico de fallas consiste en el chequeo de ciertas variables que no sobrepasen determinado límite o la aplicación de sensores redundantes (redundancia física). Métodos más avanzados descansan en el análisis espectral de las señales emanadas por la

maquinaria o la comparación del comportamiento actual de la planta sobre la base de un modelo matemático (redundancia analítica). Los últimos enfoques incluyen métodos los cuales se encuentran en un marco más determinístico (tales como relaciones de paridad y observadores) y otros formulados sobre bases estadísticas (filtros de kalman y estimación de parámetros). Las fronteras entre los varios enfoques es más bien borroso, varios métodos han mostrado tener estrechas relaciones e incluso producir resultados idénticos bajo un amplio número de condiciones.

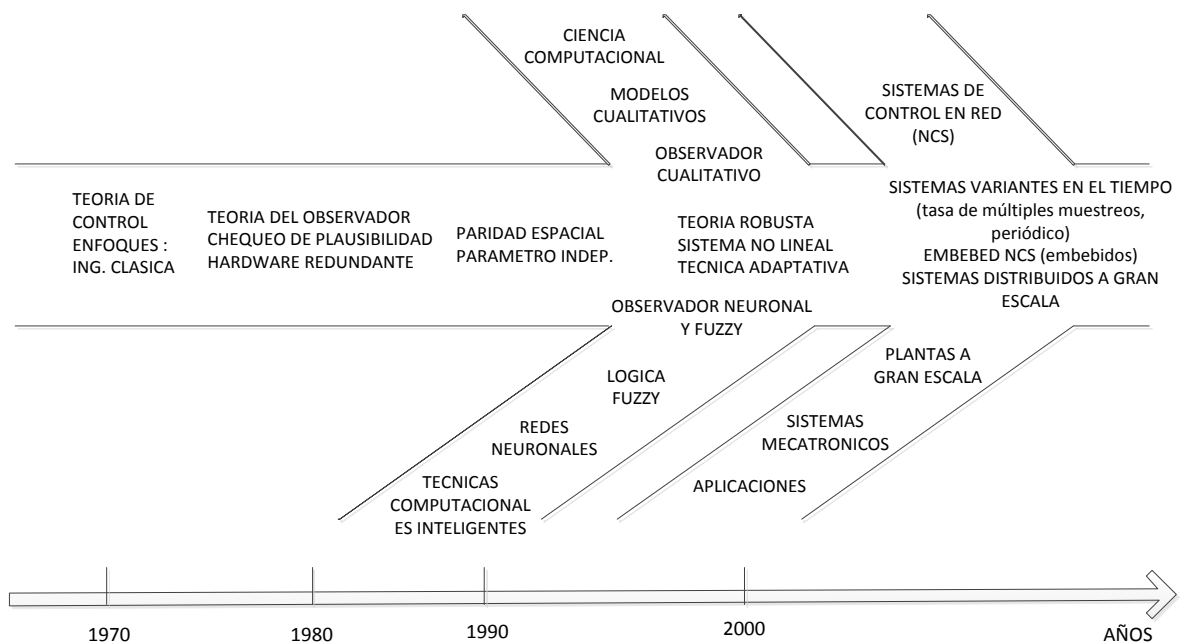


Figura 7. Bosquejo del desarrollo histórico de la teoría en el diagnóstico de fallas.

Fuente: Model-Based fault Diagnosis techniques, design schemes, algorithms, and tools. Steven X. Ding

El estudio sobre el diagnóstico de fallas basado en modelos tuvo su origen a principios de los años 70. Como se muestra en la Fig.7, en los primeros veinte años, la comunidad de control fue la que realizó la contribución decisiva para la

teoría de FDI basado en modelos, mientras que en las últimas décadas, las tendencias en la teoría FDI están siendo marcadas por contribuciones en:

- La comunidad de ciencia computacional con los métodos basados en conocimiento y cualitativos al igual que las técnicas de inteligencia computacional.
- Las aplicaciones, principalmente guiadas por las demandas urgentes de alta confiabilidad y seguridad en los sistemas de control en la industria automotriz, en el sector aeroespacial, en robótica al igual que en plantas y procesos interconectados y distribuidos a gran escala.

A.7.1.2 HISTORIA DE LOS METODOS DE DIAGNÓSTICO Y DETECCIÓN DE FALLAS EN EL SECTOR DE LA AVIACION

En las últimas décadas, han emergido un amplio rango de enfoques y estrategias para la realización manual, semiautomática, o totalmente automática del diagnóstico y pronóstico (gerenciamiento del estado o salud) de fallas en sistemas críticos, de manera especial en los sectores mercantiles de defensa y comercio. De la mano de estos diseños de sistemas de diagnóstico y pronóstico, existen igualmente un diverso número de filosofías y arquitecturas asociadas, utilizadas para su implementación en aplicaciones particulares. Por ejemplo, La primera generación de aeronaves dependía de la detección y aislamiento manual de problemas en tierra para conocer sus necesidades para el gerenciamiento de la salud del sistema. Estos sistemas típicamente eran análogos e independientes unos de otros, con tan solo lecturas esquemáticas y de voltímetro disponibles para solucionar los problemas. A medida que los sistemas de aviación se volvieron más complicados e integrados, se implementó el built-in test equipment (BITE),

alarmas sencillas, y análisis de tendencias para advertir a los operadores de situaciones críticas de seguridad. Sin embargo los encargados del mantenimiento del sistema todavía no utilizaban completamente esta herramienta y a menudo continuaban fiándose del voltímetro, los esquemas, y reportes del operador para resolver problemas o mejorar los diseños.

Continuando con la evolución del sistema de diagnóstico en la aviación como ejemplo (aunque otros equipos poseen paralelos similares), los fabricantes de los equipos comenzaron a darse cuenta de que las salidas de las alarmas para detectar fallas o los equipos de BIT podrían generar un sistema de soporte para solucionar problemas. Con este concepto, los indicadores de fallas fueron implementados en una unidad especial line-replaceable unit (LRU), los cuales indicaban que una falla había sido detectada, originalmente eran mecánicos pero después fueron reemplazados por pequeños diodos (light-emitting diodes LEDs). Estos tipos de indicadores progresaron en la década de los 50s y 60s. En muchos casos el panel frontal del LRU contenía un interruptor que ordenaba al LRU auto examinarse, de forma similar a como el equipo de soporte en tierra podría inspeccionar el LRU. Esta capacidad se denominó *built-in test equipment* (BITE). Esta capacidad comenzó a disminuir la necesidad de ciertos equipos de soporte terrestre que previamente se utilizaban para inspeccionar el equipo crítico. Dependiendo del sistema, el indicador de fallas del LRU mostraba efectivamente el punto mecánico de falla en la dirección correcta, pero las lecturas de esquemas y voltímetro aun se utilizaban en muchas condiciones. Sin embargo el BITE de esa época era a menudo confuso, poco confiable, y difícil de usar. Los mecánicos a menudo desconfiaban del BITE. Muchos sistemas en aviones como el Boeing 707, 727, el temprano 737/747 y el McDonnell Douglas DC-8, DC-9, y DC-10 empleaban este tipo de diagnóstico y practica de mantenimiento.

En la década de los 70s, a medida que los computadores comenzaron a aparecer, algunos de los sistemas de gran complejidad de la época comenzaron a utilizar los computadores para realizar sus cálculos. Esto se llamó el digital BITE. Con estos computadores surgió la habilidad de mostrar la información de la detección y aislamiento de la falla en forma digital, normalmente utilizando códigos numéricos, en el panel frontal del LRU. La lógica digital permitía que se pudiera aislar mejor la causa de la falla. La nueva forma digital ofreció la capacidad de mostrar diferentes códigos para identificar cada tipo de falla que era detectado. Estos códigos a menudo indicaban cierta descripción en un manual que podía ser utilizado para aislar y corregir la falla. Muchos sistemas en el Boeing 757/767, Airbus A300/310, McDonnell Douglas DC-10, y Lockheed L-1011 emplearon este enfoque.

A medida que el número de sistemas crecía, el uso de paneles frontales separados para mantener los sistemas se volvió cada vez menos efectivo, particularmente desde que cada LRU a menudo utilizaba una técnica diferente para mostrar los datos de las fallas. Además, algunos de los sistemas habían incrementado su interconexión. Los buses de datos digitales, tales como ARINC 429, comenzaron a utilizarse en este periodo. Los sistemas autopilotos, por ejemplo, fueron los primeros en usar estos buses de datos digitales y dependían de los datos de los sensores proveídos por muchos otros sistemas. El éxito de estos sistemas llegó a ser una fuerza que empujó a la definición de sistemas de mantenimiento más sofisticados. Un monitoreo más sofisticado era necesario para conocer los requerimientos de integridad y certificación de su sistema automático de aterrizaje. Por ejemplo el control de mantenimiento 767 y el panel de control unieron las funciones de mantenimiento de muchos sistemas relacionados. Como un siguiente paso, ARINC 604, definido en 1986, proveyó un sistema central de fallas (central fault display system CFDS), que mostró todos los indicadores de mantenimiento de los sistemas del avión en un solo lugar. Este enfoque permitió un acceso más consistente a los datos de mantenimiento en

todos los sistemas del avión y una observación mayor de los sistemas de lo que individualmente podría contener, y ahorro los costos de implementar un panel frontal de muchos de los sistemas asociados LRU. En este enfoque, los CFDS fueron utilizados para seleccionar el sistema del cual se deseaban los datos y luego encaminar el texto de mantenimiento de ese sistema a la pantalla. Este enfoque fue aplicado en muchos de los sistemas del Boeing 737s y la mayoría de los sistemas en el Airbus A320/330/340 y el McDonnell Douglas MD11.

Los sistemas continuaron aumentando su complejidad e integración. Una sola falla en el avión podría causar indicaciones de fallas en muchos sistemas, incluso con el uso de sistemas como el CFDS. En muchos casos, la mecánica tenía muy poca ayuda para determinar cuál indicación identificaba la fuente de la falla y cual era simplemente un efecto. Para resolver este y otros problemas se desarrolló en los tempranos 90s el ARINC 624. Con esto se define un sistema de mantenimiento más integrado que puede consolidar las indicaciones de fallas de múltiples sistemas y proveer una funcionalidad adicional para dar soporte al mantenimiento basado en condiciones. Con tal arquitectura, se requiere de un equipo de soporte terrestre mínimo para examinar los sistemas del avión porque la mayoría de estas operaciones están incluidas en el sistema de mantenimiento. Por ejemplo, la mayoría de las pruebas funcionales en los sistemas de los aviones del Boeing 747-400 y 777 consisten de no más que la ejecución de pruebas seleccionadas, monitoreo de las fallas mostradas en pantalla, y monitoreo de ciertos datos de bus utilizando el sistema integrado de mantenimiento. La evolución del diagnóstico y pronóstico de los sistemas se ha esforzado en identificar mejor el LRU o componente que es la fuente de la falla.

Esto permite a quienes mantienen el sistema, remover el componente fallido y corregir la condición de falla confiadamente. Aunque en muchos casos esto es

posible, hay muchas situaciones donde no es posible sin la adición de sensores o cableado.

Además estos sensores incrementan el número de componentes que pueden fallar y a veces puede empeorar el desarrollo del mantenimiento, sin mencionar que añaden costo y peso al avión. Hay casos claros donde la adición de tal hardware puede ser de beneficio, pero el mejoramiento en el aislamiento de la falla debe ser contrastado con la potencial reducción en la confiabilidad y el aumento en el costo y peso adicional para el avión. (Ver línea de tiempo Fig.8.).

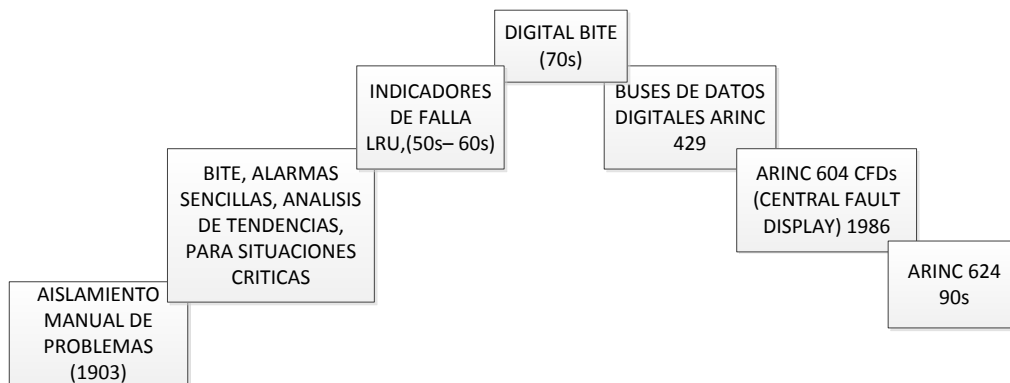


Figura 8. Eventos principales en la evolución de los métodos de diagnóstico y detección de fallas en la aviación

Fuente: autores

A.7.1.3 ESTADO Y TENDENCIA ACTUAL

Es un hecho que los modernos sistemas de control se vuelven cada vez más complejos y que los algoritmos de control que se implementan son cada vez más sofisticados. En consecuencia, las características de fiabilidad, disponibilidad, seguridad y protección medioambiental adquieren cada vez mayor importancia. Dichas características son importantes no solo en aquellos sistemas cuya seguridad es crítica, tales como centrales nucleares, plantas químicas y aeronaves, sino en cualquier tipo de proceso de fabricación automatizado como los implicados por ejemplo en el sector automotriz.

Para los sistemas en los que la seguridad es crítica, las consecuencias de los fallos pueden ser extremadamente serias en términos de vidas humanas, impacto medioambiental y pérdidas económicas; por lo que existe una necesidad creciente en la supervisión en línea y en el diagnóstico de fallos con el objetivo de incrementar la fiabilidad, teniendo en cuenta que los síntomas relacionados a fallos que se están desarrollando pueden ayudar a evitar daños irreversibles como caídas del sistema y catástrofes. Para aquellos sistemas donde la seguridad no es crítica, las técnicas de diagnóstico de fallos en línea se pueden utilizar para mejorar la eficiencia, mantenibilidad, disponibilidad y fiabilidad de la planta. Los métodos modernos de diagnóstico de fallos pueden aportar información del estado del sistema que permita implementar un mejor plan de mantenimiento.

Los métodos tradicionales de detección y diagnóstico de fallos se basan en la comparación de variables medidas del proceso con valores límites constantes y preestablecidos (chequeo de umbrales) o la aplicación de sensores redundantes (redundancia física). Otros métodos más avanzados se basan en la aplicación de

test (univariantes o multivariantes) de hipótesis a propiedades estadísticas de las variables del proceso. Dentro de los multivariantes, se encuentran aquellos métodos basados en análisis estadísticos de los datos (data-driven), especialmente indicados para grandes sistemas que producen una gran cantidad de datos, ya que reducen esa gran cantidad de información quedándose con la parte más significativa. Entre estos métodos están el análisis de la componente principal (PCA), análisis del discriminante de Fisher (FDA), mínimos cuadrados parciales (PLS) y análisis de variables canónicas (CVA).

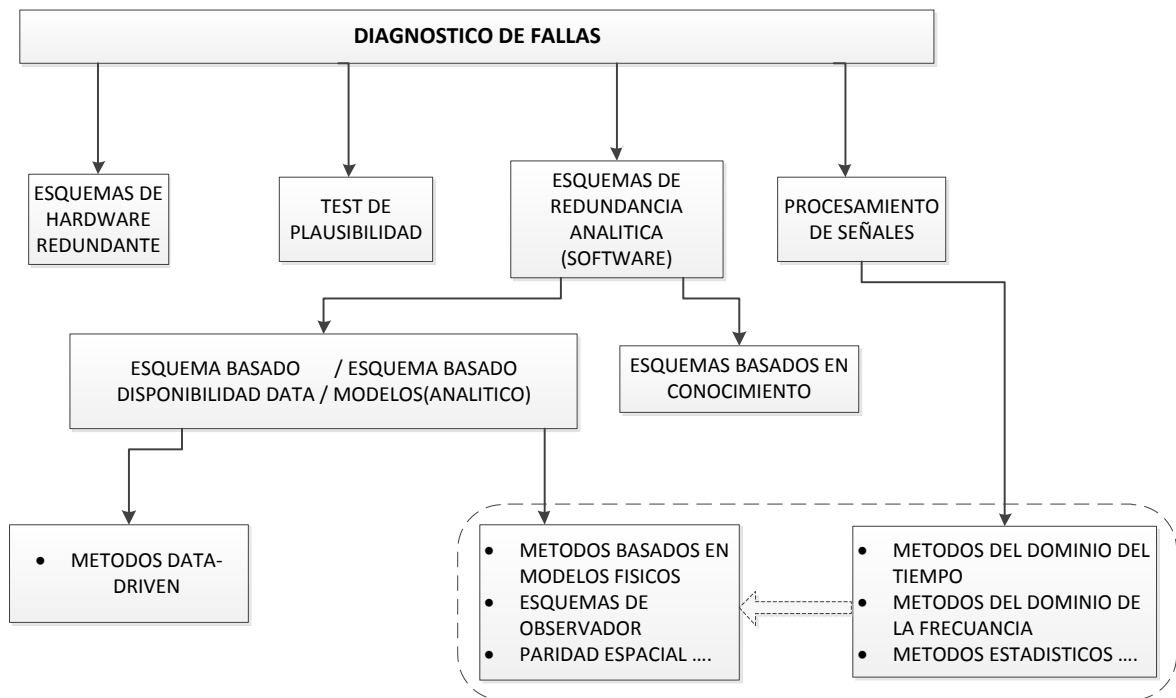


Figura 9. Clasificación de los métodos de diagnóstico de fallas.

Fuente: Model-Based fault Diagnosis techniques, design schemes, algorithms, and tools. Steven X. Ding

La técnica de diagnóstico de fallas basada en modelos es un campo de investigación relativamente joven en el dominio del diagnóstico de fallas técnicas,

dentro de la ingeniería clásica, su desarrollo ha sido rápido y actualmente recibe una considerable atención. En la Fig 9., se encuentra una clasificación general de las técnicas de diagnósticos de fallas y su relación con la técnica basada en modelos.

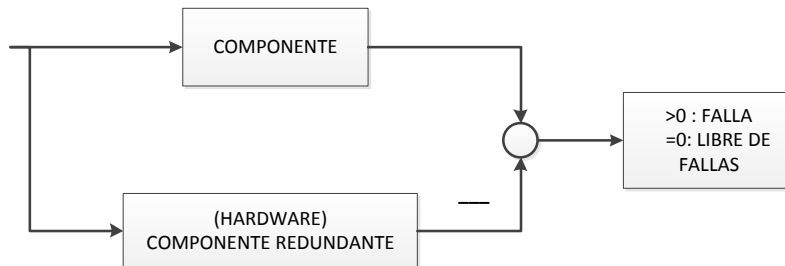


Figura.10. Descripción esquemática del esquema de hardware redundante

Fuente: Model-Based fault Diagnosis techniques, design schemes, algorithms, and tools. Steven X. Ding.

La esencia de este esquema, como se muestra en la figura 10., consiste en la reconstrucción de los componentes del proceso utilizando los componentes de hardware idénticos (redundantes). Una falla en el componente del proceso es detectada si la salida del componente del proceso es diferente de la de su redundancia. La principal ventaja de este esquema es su alta confiabilidad y aislamiento de la falla directo. El uso de hardware redundante por otro lado, resulta en altos costos, por ello la aplicación de este esquema se restringe a un número de componentes claves.

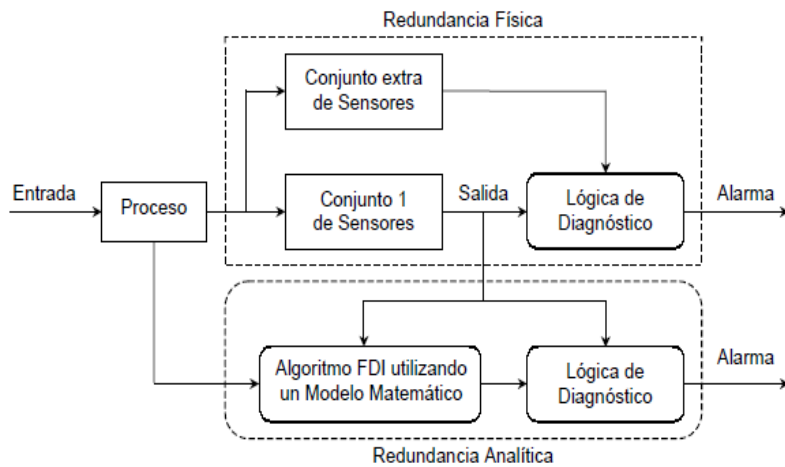


Figura 11. Redundancia física versus a redundancia Analítica

Fuente: *Model-Based fault Diagnosis techniques, design schemes, algorithms, and tools.* Steven X. Ding.

Otros métodos se basan en la redundancia analítica, Fig. 11., es decir la comparación del comportamiento actual de la planta con el esperado obtenido mediante un modelo matemático de la misma. Algunos de estos métodos tienen un marco determinista, como las ecuaciones de paridad a partir del modelo de espacio de estado, ecuaciones de paridad a partir del modelo entrada salida y observadores; mientras que otros métodos se formulan en un contexto estocástico, como los filtros de Kalman y la estimación de parámetros. La teoría lineal de estos métodos está bien desarrollada y sus relaciones bien establecidas.

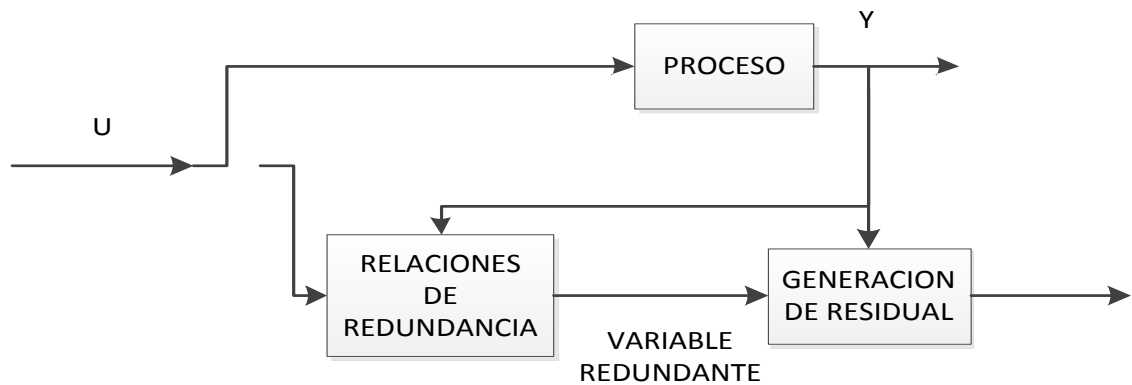


Figura.12. Esquema general de la redundancia analítica

Fuente: A review of process fault detection and diagnosis part I: quantitative model-based methods. Elsevier

Aunque desde los primeros años de la década de los 70 se vienen produciendo avances significativos en el campo del diagnóstico de fallos tanto en la teoría como en la práctica, ha sido en los últimos años cuando las investigaciones en este campo más se han dirigido hacia los sistemas de control no lineales.

Tradicionalmente, el problema del diagnóstico de fallos en sistemas dinámicos no lineales se ha manejado en dos etapas. En la primera se linealiza el modelo alrededor de un punto de operación, para en una segunda etapa aplicar técnicas robustas para generar los residuos.

En general, la mayoría de los sistemas físicos dan lugar a modelos prácticos y reales de alto orden, variables en el tiempo y no lineales. En sistemas de control, sin embargo, es importante mantener el sistema en un punto de equilibrio y se puede considerar la respuesta del sistema como señales pequeñas en torno a dicho equilibrio. Este método solo funciona bien cuando la linealización no causa mucha diferencia entre los modelos lineal y no lineal y el sistema opera cerca del

punto de operación especificado. Al tratar con sistemas con no linealidades altas y amplios rangos de operación, el problema del diagnóstico de fallos debe manejarse utilizando directamente técnicas no lineales.

Existen desarrollos y aplicaciones de diagnóstico de fallos con ecuaciones de paridad no lineales. Cuando los modelos analíticos no son lo suficientemente fiables, una red neuronal correctamente entrenada se puede usar como modelo dinámico no lineal de un sistema.

Algunas veces se requiere información cualitativa para que el modelo represente adecuadamente al sistema, y es aquí donde la lógica borrosa y las redes neuroborrosas tienen su papel en aplicaciones de diagnóstico de fallos. Recientemente en la literatura de diagnóstico de fallos, a estos métodos basados en modelos analíticos que incorporan técnicas cualitativas se les llama métodos de software computacionales.

El trabajo sobre diagnóstico de fallos en la comunidad de la Inteligencia Artificial se enfocó inicialmente sobre métodos basados en el conocimiento o sistemas expertos, donde se aplica la heurística para asociar explícitamente los síntomas con las hipótesis de fallos. Las deficiencias de estos métodos llevaron al desarrollo de técnicas de diagnóstico de fallos basadas en modelos cualitativos en la forma de ecuaciones diferenciales cualitativas, gráficos dirigidos con signo, modelos funcionales cualitativos, modelos estructurales, etc.

Por otro lado el continuo incremento en los estándares de desempeño, calidad de producción y la optimización de costos, han sido las variantes que influyen en

el grado de complejidad y automatización de los sistemas industriales, los cuales deben garantizar la seguridad y confiabilidad del proceso; por ello actualmente uno de los problemas más críticos que rodean el diseño de sistemas automáticos, es precisamente su confiabilidad y nivel de dependencia.

Una manera común de mejorar la confiabilidad y nivel de dependencia del sistema es aumentar el desempeño y la robustez de los componentes como sensores, servomecanismos, controladores u ordenadores. De manera tal que inclusive en un sistema libre de fallas, su operación no puede ser totalmente garantizada. Por ello el monitoreo de procesos y detección de fallas se ha convertido ya en parte fundamental de los sistemas automáticos de control moderno. Todos los sistemas de detección de fallas ya sea basado en modelos o basado en conocimiento, tienen en común el uso de un modelo del proceso, dependiendo del tipo de algoritmo que sea implementado para el procesamiento de datos, si son recolectados en línea, o grabados durante la operación del sistema.

En conclusión la detección y aislamiento de fallas (diagnóstico) en los sistemas ingenieriles son de gran significado práctico. Los sistemas interesados abarcan un gran espectro de maquinaria, incluyendo fábricas de producción industrial (plantas de potencia, plantas químicas, refinerías de petróleo, fábricas de acero, fábricas de papel etc.), vehículos de transporte (barcos, aviones, automóviles) y aplicaciones domesticas (equipos de calefacción y aire acondicionado, máquinas de lavar...). La detección temprana de la ocurrencia de la falla, es crítica para evitar el deterioro del producto, degradación en la eficiencia, mayores daños en la misma máquina y daños en la salud humana o incluso pérdidas de vidas. El diagnóstico rápido y correcto del componente que ha fallado facilitan las decisiones optimas y apropiadas para acciones de emergencia y correctivas. Los enfoques tradicionales en el diagnóstico y detección de fallas involucran el

chequeo límite de ciertas variables o la aplicación de sensores redundantes (redundancia física). Métodos más avanzados recaen en el análisis espectral de las señales emanadas por la maquinaria o en la comparación del comportamiento actual de la planta con el esperado por una base de un modelo matemático (redundancia analítica). Los últimos enfoques incluyen métodos los cuales se encuentran en un marco determinístico (tales como las relaciones de paridad y los observadores) y otros formulados en una base estadística (los filtros kalman y la estimación de parámetros). Las fronteras entre varios enfoques son borrosas, y más recientemente, varios métodos han mostrado estar relacionados de cerca e incluso producir resultados idénticos bajo amplias condiciones.

A.7.2 DEFINICIONES Y CONCEPTOS BASICOS

- **CLASES DE FALLAS:** En general, las fallas son desviaciones del comportamiento normal en la planta o en su instrumentación. Las fallas de interés pertenecen a una de las siguientes categorías:
 - Fallas aditivas del proceso: estas son entradas desconocidas que actúan sobre la planta, que normalmente son cero y las cuales, cuando se presentan, causan un cambio en las salidas de las plantas independientemente de las entradas conocidas. Ej. Fugas en la planta, cargas.
 - Fallas multiplicativas del proceso: Estas son cambio (abrupto o gradual) en algunos parámetros de la planta. Ellos causan cambios en las salidas de la planta las cuales dependen también de la magnitud de las entradas conocidas. Ej. Deterioro del equipo de la planta, contaminación en la superficie, obstrucciones o pérdida parcial o total de potencia.

- Fallas en el sensor: estas son discrepancias entre los valores medidos y los reales de las variables individuales de la planta. Estas fallas se consideran aditivas (independiente de la magnitud medida) aunque algunas fallas (tales como el quedarse pegado o falla total) se describen mejor como multiplicativas.
- Fallas en el actuador: son discrepancias entre el comando de entrada de un actuador y su salida real. Estas fallas se consideran aditivas aunque algunas (tales como el quedarse pegado o falla total) se describen mejor como multiplicativas.

La detección y diagnóstico de la falla implementa las siguientes tareas:

- DETECCIÓN DE LA FALLA: que es, la indicación de que algo está mal en el sistema monitoreado
- AISLAMIENTO DE LA FALLA: que es, la determinación de la ubicación exacta de la falla (el componente averiado)
- IDENTIFICACIÓN DE LA FALLA: que es, la determinación de la magnitud de la falla.

El aislamiento y la identificación se les hacen referencia como diagnóstico de fallas.

Mientras que la detección es un absoluto deber en cualquier sistema práctico y el aislamiento de la falla es casi igual de importante, la identificación de la falla (aunque útil) puede no justificar el esfuerzo extra que requiere. Por ello, los sistemas más prácticos contienen solo las etapas de detección y aislamiento de la falla (y son referenciadas como sistemas FDI). Además, en muchos casos “diagnóstico” se utiliza como sinónimo de “aislamiento”.

Generalmente, la actividad de detección y diagnóstico toma lugar on-line, en tiempo real. Las dos tareas, detección y diagnóstico, se pueden ejecutar en paralelo o secuencialmente. En algunos sistemas de diagnóstico, una sola decisión no solo indica el hecho de que la falla se presentó sino además su ubicación. En otros sistemas, la tarea de detectar se realiza permanentemente mientras que la tarea de diagnóstico es disparada solo cuando se hace presente la detección de la falla.

Particularmente en la detección y diagnóstico de fallas basada en modelos, las siguientes convenciones son generalmente adoptadas:

Se asume que las fallas no se presentan inicialmente en el sistema sino que aparecen en un tiempo posterior. Las fallas se describen usualmente por medio de funciones de tiempo determinísticas las cuales son desconocidas. Casos especiales de importancia son el jump-fault (salto de falla) [función step], y drift-fault (función rampa), en las cuales su magnitud y tiempo de aparición son desconocidos.

También se puede hablar de perturbaciones aditivas, las cuales también son entradas desconocidas y determinísticas para el sistema. La diferencia entre fallas aditivas y perturbaciones aditivas es subjetiva: las fallas son esas entradas desconocidas las cuales deseamos detectar y aislar mientras las perturbaciones son una molestia que deseamos ignorar.

Cualquier ruido, originado por la planta o por los sensores y actuadores, es considerado aleatorio con ningún (cero) significado (cualquier ruido con significado se maneja como una falla o una perturbación).

Los errores de modelado son discrepancias entre el modelo (parámetros del modelo) y el verdadero sistema. Se pueden presentar desde los orígenes del sistema o pueden aparecer debido a cambios en el punto de operación. Los errores de modelado son elementos molestos que obstaculizan el efecto que se quiere suprimir. Se pueden considerar como perturbaciones multiplicativas, en contraste a las fallas multiplicativas las cuales son discrepancias entre el modelo y el sistema verdadero, pero que deseamos detectar.

La etapa de detección dentro de la técnica de diagnóstico se caracteriza por un número de importantes y cuantificables puntos de referencia, denominados:

- Sensibilidad de falla, que es, la habilidad de la técnica para detectar fallas de un razonable pequeño tamaño.
- Velocidad de reacción, que es, la habilidad de la técnica para detectar fallas con una pequeña demora después de su aparición.
- Robustness (robustez), que es, la habilidad de la técnica para operar en presencia de ruido, perturbaciones y errores de modelado, con pocas falsas alarmas.

Sensibilidad de falla, velocidad de reacción y robustez surgen de la interacción entre fallas, ruido y errores de modelado, y son afectados por el diseño del algoritmo de detección.

La etapa de aislamiento, que es, la habilidad del sistema de diagnóstico para distinguir fallas dependiendo de las propiedades físicas de la planta, del tamaño de las fallas, ruido, perturbaciones y errores de modelado, y en el diseño del algoritmo.

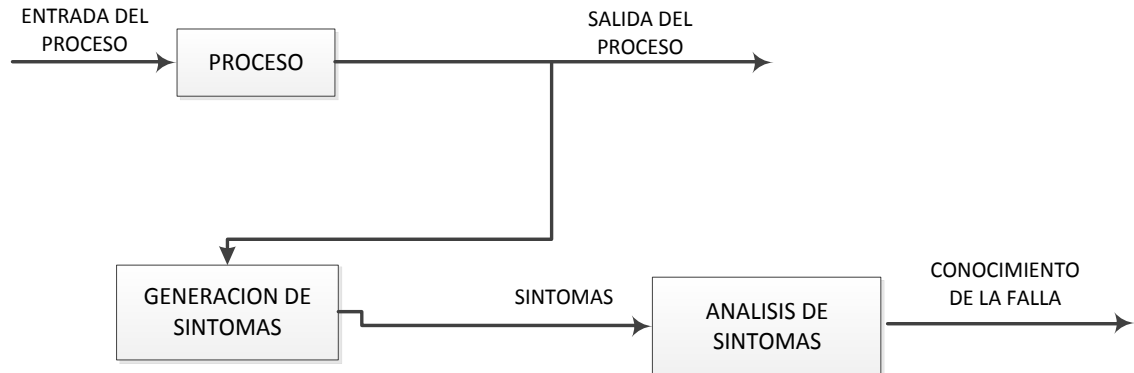


Figura.13. Descripción esquemática del método basado en procesamiento de señales

Fuente: Model-Based fault Diagnosis techniques, design schemes, algorithms, and tools. Steven X. Ding

Asumiendo que ciertas señales del proceso contienen información acerca de las fallas de interés y esta información es presentada en forma de síntomas, un diagnóstico de fallas se puede lograr mediante un adecuado procesamiento de señales. Los síntomas típicos son funciones en el dominio del tiempo como magnitudes, valores de media aritméticos o cuadráticos, valores límite, tendencias, momentos estadísticos de la distribución de amplitud o en voluta, o funciones en el dominio de la frecuencia como densidades de potencia espectral, líneas de frecuencia espectral o espectros, etc. El esquema basado en el procesamiento de señales principalmente se utiliza en procesos en estado estable, y su eficiencia para la detección de fallas en sistemas dinámicos, los cuales poseen un amplio rango de operación debido a la posible variación de las señales de entrada, es

considerablemente limitada. En la fig. 13.se ilustra la idea básica de los esquemas de procesamiento de señales.

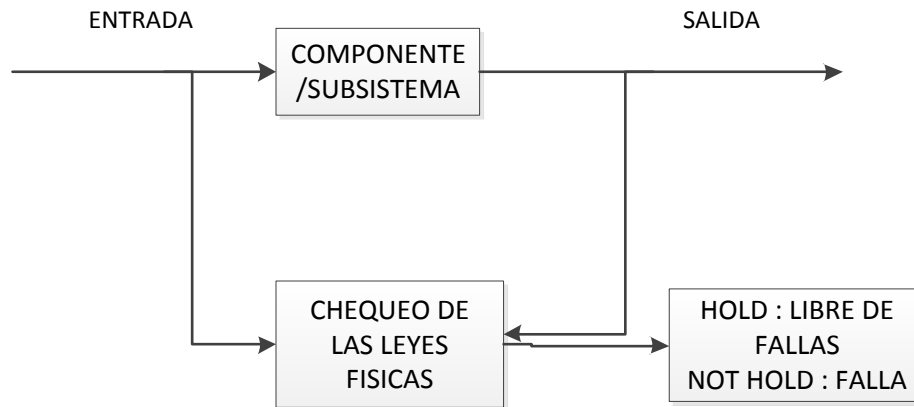


Figura. 14.Descripción esquemática de la prueba de plausibilidad.

Fuente: Model-Based fault Diagnosis techniques, design schemes, algorithms, and tools. Steven X. Ding

Como se muestra en la fig. 14., la prueba de plausibilidad se basa en el chequeo de algunas leyes físicas bajo las cuales el componente del proceso trabaja. Asumiendo que la falla guiara a la pérdida de plausibilidad, chequeando la plausibilidad nos proveerá con información acerca de la falla. La prueba de plausibilidad es limitada en su eficiencia para detectar fallas en un proceso complejo o para aislar las fallas. Observar en la figura 15. el esquema de fallos en sistemas.

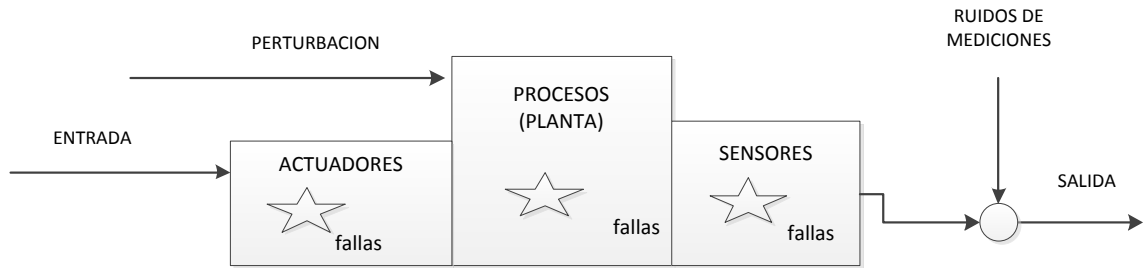


Figura 15. Descripción esquemática de fallos en los sistemas

Fuente: *Model-Based fault Diagnosis techniques, design schemes, algorithms, and tools.* Steven X. Ding

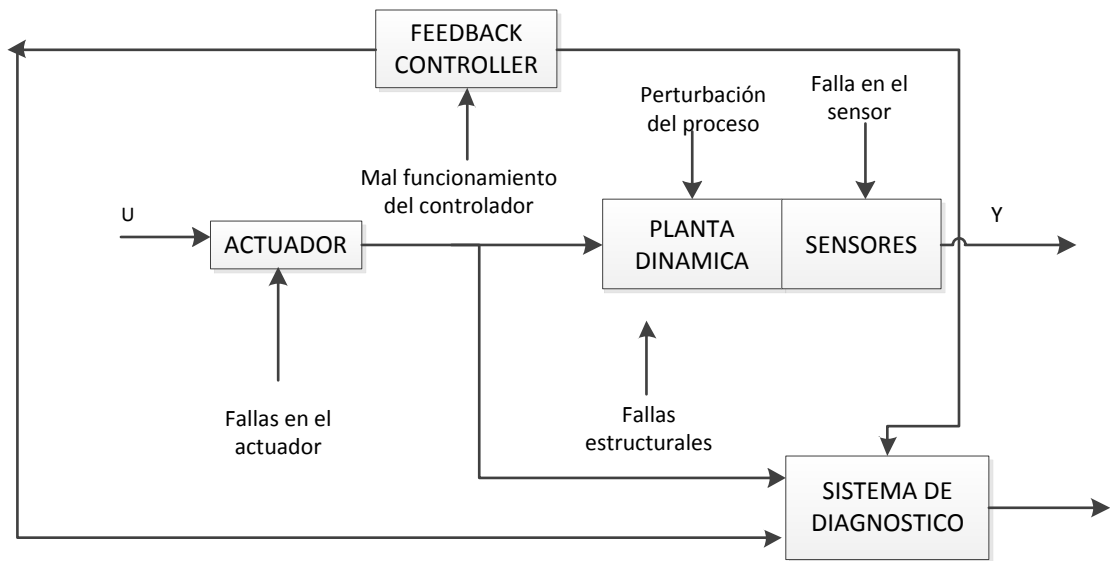


Figura. 16. Marco general de diagnóstico

Fuentes: *A review of process fault detection and diagnosis part I: quantitative model-based methods.* Elsevier

En la Fig. 16 se observan los componentes de un marco general de diagnóstico de fallos. La figura muestra un sistema de un proceso controlado e indica las diferentes fuentes de fallos en el.

A.7.2.1 ESQUEMAS DE DIAGNÓSTICO Y MONITOREO DE FALLAS

Estrategias de detección e identificación de fallas (FDI) han sido desarrolladas en recientes años y han encontrado una extensiva utilidad en un sinnúmero de aplicaciones. Se usa el término FDI aquí para referirse a la detección de falla preliminar y la determinación de los componentes/subsistemas específicos de falla. Se puede visualizar en la figura 16, que la señal residual representa una desviación de las condiciones de operación estándar, que es generada por comparación, por ejemplo un modelo de salida cualquiera con el sistema de la figura de salida. Basados en esta señal, se toman decisiones acerca de las condiciones de operación de la maquinaria. Los métodos para generar lo restante pueden estar basados en estadística de señal de tiempo, identificación de sistemas, y técnicas de dominio de la frecuencia, etc.



Figura. 17. Estructura General FDI

Fuentes: A review of process fault detection and diagnosis part I: quantitative model-based methods. Elsevier

La tecnología disponible generalmente están agrupadas en dos categorías: **Model-based** (basado en modelos) y **Data-Driven** (manejo de datos), como se ilustra en la figura 17. En esta figura las técnicas de *Basado en Modelos* se soportan de un modelo dinámico exacto del sistema y son capaces de detectar inclusive fallas inesperadas. Estos toman ventaja de las salidas del sistema y del

modelo para generar “discrepancia” o sobrante (innovador), entre las dos salidas lo cual es indicio de una condición de falla potencial. En el método, un banco de filtros es usado para precisar la identidad de la falla del componente correspondiente. Por otro lado se tienen las técnicas de *Data-Driven* que ofrecen únicamente la anticipación de las condiciones de falla, donde un “modelo” de falla ahora es la construcción o colección de construcciones como una red neuronal, sistemas expertos, etc. Los cuales primero deben ser entrenados o instruidos, con prototipos conocidos de patrones de falla (data) y después implementados en línea para que puedan detectar la identidad de los componentes defectuosos.

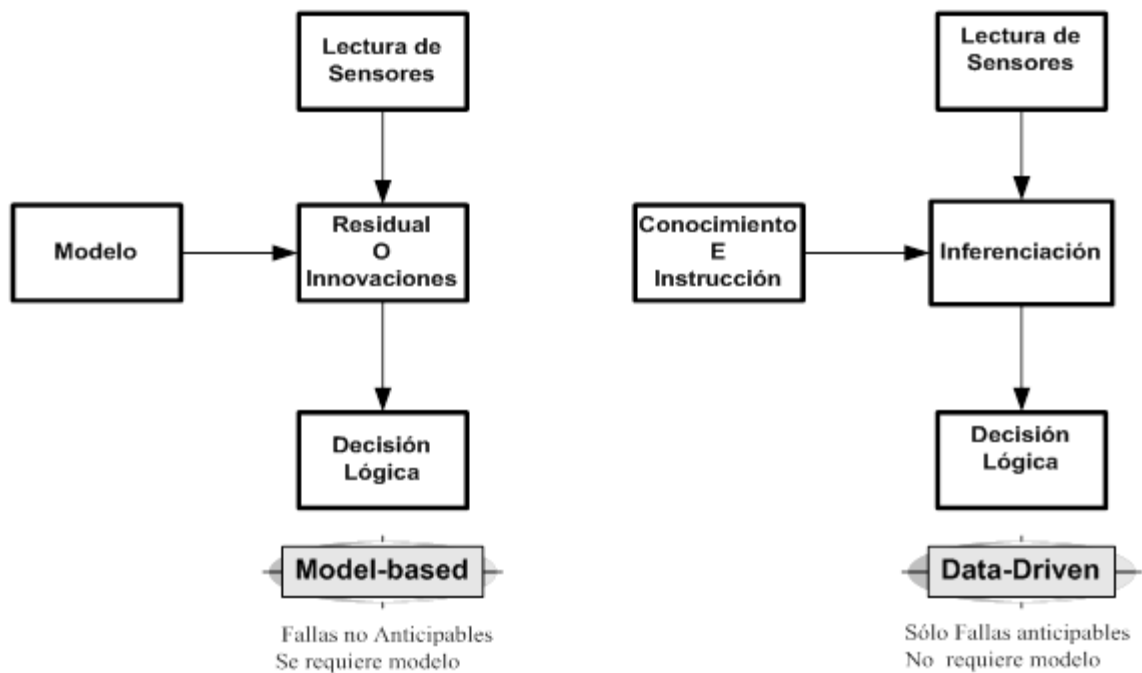


Figura. 18. Técnicas Basadas en modelos y Data- Driven

Fuentes: A review of process fault detection and diagnosis part I: quantitative model-based methods. Elsevier

Un esquema general de monitoreo y diagnóstico de fallas en línea es mostrado en la figura 19. que representa el procedimiento completo en línea requerido para monitoreo y diagnóstico. Los componentes con mayor relevancia de dicho procedimiento son la colección de datos por medio de la salida de los sensores, y la construcción de un *vector característica*¹⁴ que contenga la información suficiente acerca del estado actual de operación de la maquinaria, que permitan la clasificación e identificación de fallas. La información contenida en el vector de forma puede ser obtenida mediante tres formas:

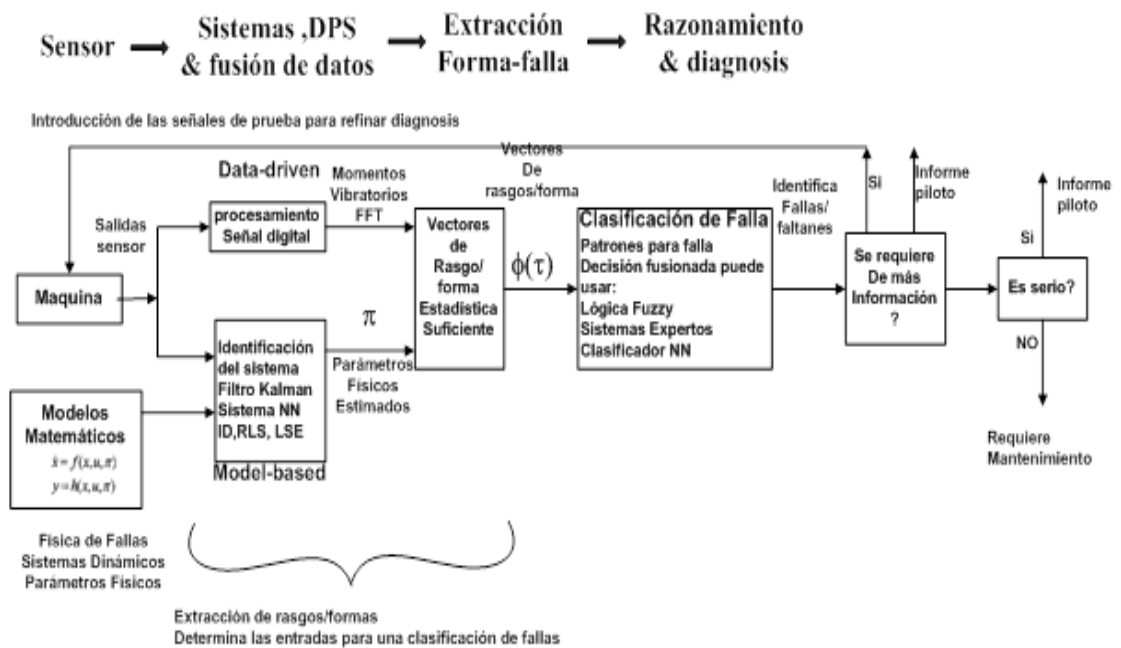


Figura.19. Monitoreo y diagnosis de fallas online

Fuente: Intelligent Fault Diagnosis and Prognosis for Engineering Systems. George Vachtsevanos et al.

¹⁴ Los vectores y valores propios (el vocablo original de raíz anglosajona es “eigen” por lo que se traduce también “eigenvectores” y “eigenvalores”).

- Método basado en modelos (ej. Identificar parámetros físicos usando filtrado de Kalman o recursivos mínimos cuadrados).
- Método Data-driven (ej. Momentos estadísticos de señales vibratorias, información del espectral de la frecuencia vibratoria).
- Regresión estadística y técnicas agrupadas sobre datos históricos existentes.

Una vez el vector característico ha sido obtenido, es usado como entrada de una serie de bloques que contienen varios tipos de algoritmos toma-decisión.

En la fig. 20. se muestra las transformaciones que los datos del proceso a traviesan durante el diagnóstico.

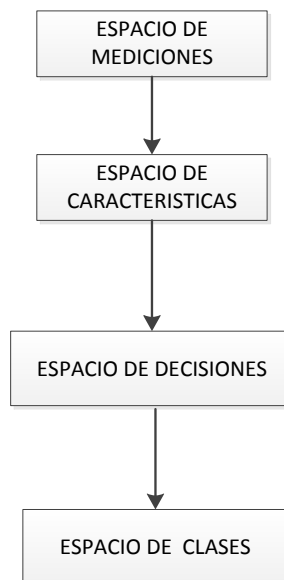


Figura 20. Transformaciones en el sistema de diagnóstico.

Fuente: A review of process fault detection and diagnosis part I : quantitative model-based methods. Elsevier

A.7.3 GENERALIDADES

A.7.3.1 PROCEDIMIENTOS DE MONITOREO DEL PROCESO

Los cuatro procedimientos asociados con el monitoreo del proceso son: detección de falla, identificación de falla, diagnóstico de la falla, y recuperación del proceso. La detección de la falla consiste en determinar si una falla ha ocurrido. Una detección temprana puede proveer de advertencias invaluableles sobre problemas emergentes, que con acciones apropiadas se puede evitar serias alteraciones en el proceso. La identificación de fallas consiste en identificar las variables observadas más relevantes en el diagnóstico de la falla. El propósito de este procedimiento es enfocar la atención del operador de la planta y el ingeniero en los subsistemas más pertinentes para el diagnóstico de la falla, así el efecto de la falla puede ser eliminado de una manera más eficiente. El diagnóstico de falla consiste en determinar cuál falla ocurrió, en otras palabras, determinar la causa del estado observado fuera de control. En conclusión el diagnóstico de fallas se puede definir como la determinación del tipo, ubicación, magnitud y tiempo de la falla.

El procedimiento de diagnóstico de la falla es esencial para la neutralización o eliminación de la falla. La recuperación del proceso, también conocido como intervención, consiste en remover el efecto de la falla, y es el procedimiento en el que finaliza el ciclo del monitoreo del proceso. En el momento en que una falla es detectada, la identificación, diagnóstico y recuperación del proceso se emplean en la respectiva secuencia; de otra forma, solo el proceso de detección de la falla se repite. Mientras que los cuatro procedimientos puedan implementarse en un esquema de monitoreo del proceso, este o siempre es necesario.

Por ejemplo, una falla puede ser diagnosticada (diagnóstico de fallas) sin identificar las variables inmediatamente afectadas por la falla (identificación de la falla). Además, no es necesario automatizar los cuatro procedimientos. Por ejemplo, un procedimiento de identificación de fallas automatizado puede utilizarse para asistir a los operadores e ingenieros de la planta en el diagnóstico de la falla y retornar a la operación normal. A menudo el objetivo del monitoreo de los procesos es el de incorporar a los operadores e ingenieros de la planta en el ciclo eficiente de monitoreo de la operación en vez de automatizar completamente el esquema de monitoreo. Después que una falla ocurre, las deficiencias en control pueden ser recuperadas mediante la reconfiguración del proceso, reparaciones del proceso o reajustando los controladores. Una vez la falla ha sido diagnosticada apropiadamente, el enfoque óptimo para neutralizar la falla puede que no sea obvio. Un enfoque factible puede ser ajustar los controladores estándar del proceso. Muchos métodos se han desarrollado para evaluar el rendimiento de los controladores, y esto se puede utilizar para determinar cual controlador del proceso necesita ser ajustado para restaurar el rendimiento satisfactorio.

En el caso de un problema en el sensor, una técnica de reconstrucción del sensor puede ser aplicada al proceso para restaurar las operaciones en control.

A.7.3.2 MEDICIONES EN EL MONITOREO DEL PROCESO

Un esquema de monitoreo del proceso típico contiene una o más mediciones, basadas en desarrollos de la teoría estadística, teoría de la clasificación de

patrones, la teoría de la información, y /o teoría de sistemas. Estas mediciones en cierta forma representan el estado o comportamiento de los procesos. La idea es convertir datos on-line recolectados del proceso en unas pocas mediciones significativas, y de este modo asistir a los operadores en la determinación del estatus de las operaciones y si es necesario un diagnóstico de las fallas. Para la detección de las fallas, se pueden ubicar ciertos límites en ciertas mediciones, y una falla es detectada cada vez que una de las mediciones esta fuera de los límites. De esta forma las mediciones son capaces de definir el comportamiento del proceso en control y de acuerdo al estado fuera de control.

Desarrollando mediciones que caractericen en forma precisa el comportamiento de cada variable en observación, el valor de una variable puede ser comparada con los valores de otras variables para determinar la variable más afectada por la falla. Las fallas pueden ser también diagnosticadas mediante el desarrollo y la comparación de medidas exactas que representen las diferentes fallas del proceso.

La meta del monitoreo del proceso es desarrollar mediciones que sean maximizadas en su sensibilidad y robustez para todas las posibles fallas. Las fallas se manifiestan en diversas formas; sin embargo, es muy improbable que todas las fallas que hayan ocurrido en un proceso puedan ser detectadas efectivamente y diagnosticada con solo unas pocas mediciones, ya que cada medición caracteriza la falla en una forma diferente, una medida será más sensible a ciertas fallas y menos a otras relativas a otras mediciones. Esto hace que se usen múltiples mediciones en el monitoreo del proceso, con las medidas de un proceso particular y las posibles fallas a mano.

Las mediciones en el monitoreo del proceso se pueden clasificar siendo asociada con uno o más de tres enfoques; llamados data-driven, analítico, y basado en conocimiento. Las medidas con data-driven se derivan directamente de los datos del proceso. Los sistemas industriales modernos, ya sea una planta entera o una máquina de papel, son sistemas a grande escala. Con la instrumentación pesada típica de los procesos modernos, los sistemas a gran escala producen una excepcional cantidad de datos. Aunque mucha información de estos procesos es disponible, está más allá de las capacidades de un operador o ingeniero para evaluar efectivamente las operaciones del proceso, simplemente de la observación de los datos. La fortaleza de las técnicas de data-driven reside en su habilidad de transformar datos de alta dimensionalidad en menores dimensiones, en las cuales la información importante es capturada. Mediante el computo de algunas estadísticas significativas para el proceso, el esquema de monitoreo del proceso para un sistema a gran escala puede ser mejorado significativamente. La principal desventaja de las mediciones de data-driven es que su efectividad es altamente dependiente de la cantidad y calidad de los datos del proceso.

A diferencia del enfoque data-driven, el enfoque analítico utiliza modelos matemáticos a menudo construidos a partir de principios. El enfoque analítico es aplicable a sistemas de información ricos, donde modelos satisfactorios y suficientes sensores están disponibles. La mayoría de las mediciones analíticas están basadas en la estimación de parámetros, diseño basado en observación, y/o relaciones de paridad. Las mayorías de aplicaciones del enfoque analítico han sido en sistemas con un número relativamente pequeño de entradas, salidas y estados. Es difícil aplicar el enfoque analítico a sistemas a gran escala (sistemas que poseen una gran cantidad de entradas, salidas y/o estados) debido a que requieren de modelos detallados para poder ser efectivos. Modelos detallados para sistemas a gran escala son costosos debido a todas las asociaciones del sistema multivariable. La principal ventaja del enfoque analítico es la habilidad de

incorporar el entendimiento físico del proceso en el esquema de monitoreo del proceso. En otras palabras, cuando los modelos analíticos detallados son disponibles, las mediciones analíticas pueden superar las medidas con data – driven.

El enfoque basado en conocimiento utiliza modelos cualitativos para desarrollar las mediciones en el monitoreo del proceso. El enfoque basado en conocimiento es bastante adecuado para sistemas en los cuales modelos matemáticos detallados no están disponibles. La mayoría de las mediciones basadas en modelos están basadas en análisis causal, sistemas expertos, y/o reconocimiento de patrones. Al igual que en el enfoque analítico, la mayoría de aplicaciones para el enfoque basado en conocimiento han sido a sistemas con un número relativamente pequeño de entradas, salidas y estados. Construir los modelos de fallas para un sistema a gran escala puede requerir de una enorme cantidad de esfuerzo. Paquetes de software están siendo desarrollados para permitir que el enfoque basado en conocimiento sea más fácilmente aplicado a sistemas complejos.

A.7.3.3 METODOS PARA EL MONITOREO DE PROCESOS

La capacidad o eficiencia de los enfoques data-driven, analítico, y basado en conocimientos depende de la calidad y el tipo de modelos disponibles, y de la cantidad y calidad de los datos disponibles. Los métodos de monitoreo tradicional constan de la detección límite y detección de discrepancias. La detección límite dispara una alarma cuando las observaciones cruzan un umbral predefinido, y ha sido aplicado tradicionalmente porque es fácil de implementar y entender.

La detección límite, sin embargo pierde sensibilidad en algunas alteraciones en procesos debido a que ignora las interacciones entre las variables del proceso para los diversos sensores. La detección de discrepancia depende altamente de la precisión del modelo, y las inexactitudes del modelo no están disponibles en la práctica. Debido a que es difícil distinguir entre fallas genuinas y errores en el modelo, la detección de discrepancia puede perder robustez. La detección límite determina el umbral para cada variable sin usar ninguna información de las otras variables, y en este sentido es idéntica a las técnicas estadísticas univariantes. Estos métodos ignoran las correlaciones entre las capturadas en solo dos o tres dimensiones, lo cual puede funcionar para algunos procesos, la variabilidad del proceso dominante puede visualizarse con una sola gráfica.

Independiente del número de dimensiones requeridas en espacios de baja dimensionalidad, otras gráficas (T^2 , y las cartas Q) pueden ser utilizadas de una forma similar a las cartas de control univariante pero están basadas en estadística multivariable. Estas cartas de control pueden ayudar a los operadores e ingenieros a interpretar las tendencias significativas en los datos del proceso.

Todas las mediciones basadas en los enfoques de data-driven, analítico, y basado en conocimiento tienen sus ventajas y desventajas, de forma que un solo enfoque no es el mejor para todas las aplicaciones. Generalmente el mejor esquema de monitoreo del proceso emplea estadística múltiple, o métodos para la detección, identificación, y diagnóstico de fallas. Se han realizado esfuerzos para incorporar varias técnicas para el monitoreo del proceso. Esto puede ser beneficioso en muchas aplicaciones.

A.7.3.4 DIAGNÓSTICO Y DETECCION DE FALLAS BASADO EN MODELOS Y DATA DRIVEN - GENERALIDADES

El diagnóstico de fallas está cobrando cada vez más y más importancia para el monitoreo del proceso debido al incremento en la demanda de un mayor rendimiento al igual que por un incremento en la seguridad y confiabilidad de los sistemas dinámicos.

Un diagnóstico temprano de las fallas del proceso mientras el sistema continua operando en una región controlable, puede ayudar a evitar una progresión de eventos anormales y una reducción en la productividad, lo cual ayuda a evadir averías mayores en el sistema y catástrofes. Por lo tanto, el diagnóstico de fallas es un tópico de investigación de gran importancia que genera un considerable interés para los practicantes industriales al igual que para los investigadores académicos.

Existe una abundante fuente literaria acerca del diagnóstico de fallas en un proceso, desde métodos analíticos a la inteligencia artificial y enfoques estadísticos. Desde una perspectiva de modelado, hay métodos que requieren modelos exactos del proceso, modelos semi-cuantitativos, o modelos cualitativos. Al otro lado del espectro, hay métodos que no asumen ninguna forma de la información del modelo y descansan solamente en los datos históricos del proceso. Además, dado el conocimiento del proceso, hay diferentes técnicas de búsqueda que pueden ser aplicadas para realizar el diagnóstico. Los métodos de diagnóstico pueden clasificarse en dos categorías generales, basado en modelos y data driven, teniendo como base el conocimiento del proceso que es un

requerimiento a priori. La jerarquía de los enfoques para diagnóstico de fallas se muestra en la figura 21.

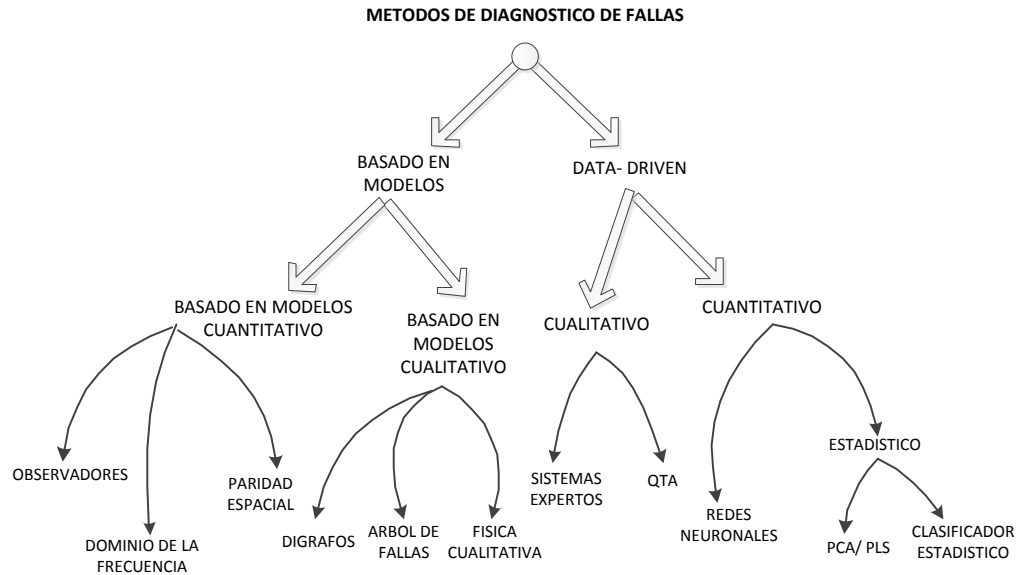


Figura 21. Clasificación de los métodos de Diagnóstico de fallas

Fuente: Model based and data driven fault diagnosis methods with applications to process monitoring. Qinsong yang

El tipo de conocimiento del proceso a priori utilizado es la más importante y distinguida característica para la clasificación de los sistemas de diagnóstico de fallas. El conocimiento a priori básico que se necesita para el diagnóstico de fallas es el conjunto de fallas y la relación entre las observaciones (síntomas) y las fallas. Un diagnóstico del sistema las puede tener explícitamente, o pueden ser inferidas de alguna fuente del dominio del el conocimiento. El dominio del conocimiento a priori puede ser desarrollado de un entendimiento fundamental del proceso usando los primeros principios del conocimiento.

Tal conocimiento se refiere como profundo, causal o conocimiento basado en modelos. Por otro lado, el conocimiento puede ser obtenido de experiencias pasadas en el proceso. A este conocimiento se hace referencia como poco profundo, evidencia o conocimiento basado en la historia. Para métodos basados en modelos, el conocimiento a priori basado en modelos puede clasificarse en forma general como cualitativa o cuantitativa. El modelo generalmente desarrollado basándose en algunos conceptos fundamentales de la física el proceso.

En los modelos cuantitativos este entendimiento es expresado en términos de relaciones funcionales matemáticas entre las entradas y salidas del sistema. En modelos de ecuaciones cualitativas estas relaciones se expresan mediante funciones cualitativas. Para los enfoques data-driven, solo la disponibilidad de una gran cantidad de datos históricos del proceso es asumida. Existen diferentes formas en las cuales estos datos pueden ser transformados y presentados como conocimiento a priori para el diagnóstico del sistema. Esto se conoce como proceso de extracción de características del historial de datos del proceso, y se hace para facilitar el diagnóstico que se realizara posteriormente. Este proceso de extracción puede principalmente proceder como una extracción de características cuantitativas o cualitativas. En la extracción de características cuantitativa se puede realizar ya sea por métodos estadísticos o no estadísticos. Puede haber puntos coincidentes entre el método basado en modelos y el data driven, esto es solo una clasificación basada en si el conocimiento de las características del proceso son requeridas o no.

A.7.3.5 ENFOQUES DE DIAGNÓSTICO DE FALLAS BASADO EN MODELOS

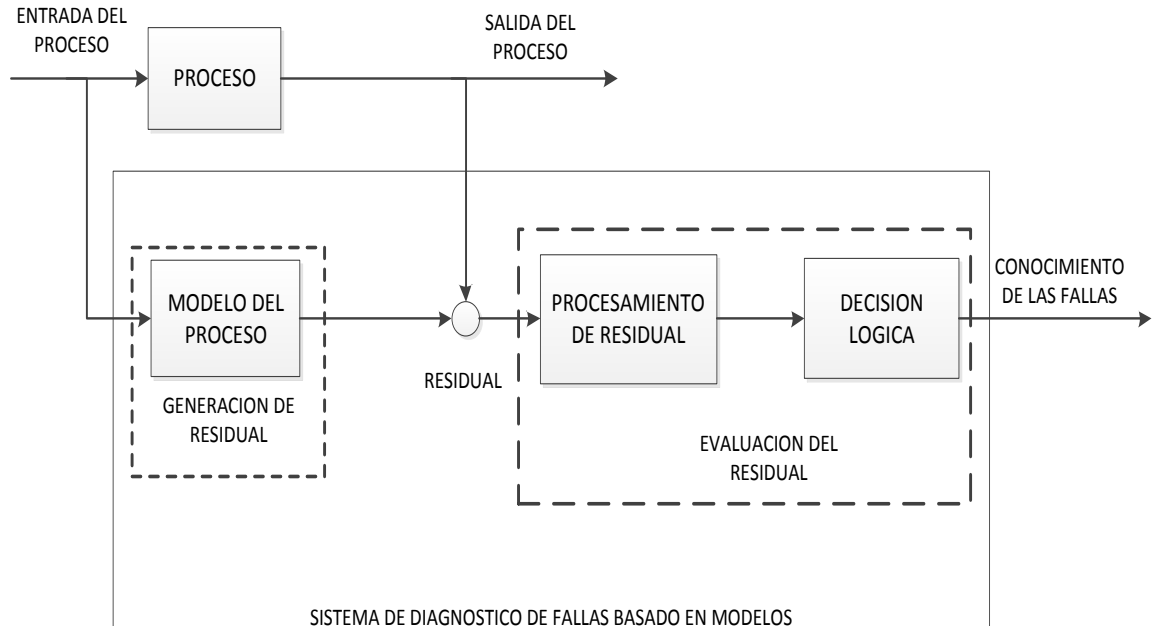


Figura 22. Descripción esquemática del diagnóstico de fallas basado en modelos

Fuente: *Model-Based fault Diagnosis techniques, design schemes, algorithms, and tools.* Steven X. Ding

A.7.3.5.1 Métodos cuantitativos de diagnóstico basado en modelos

La mayoría del trabajo en los métodos cuantitativos basados en modelos tiene su fundamento en utilizar modelos generales de entrada-salida y estados espaciales para generar residuales. Este enfoque se puede clasificar en métodos de observador, paridad espacial y dominio de la frecuencia. El enfoque matemático del método basado en modelos en esta tesis recae sobre la categoría de observador.

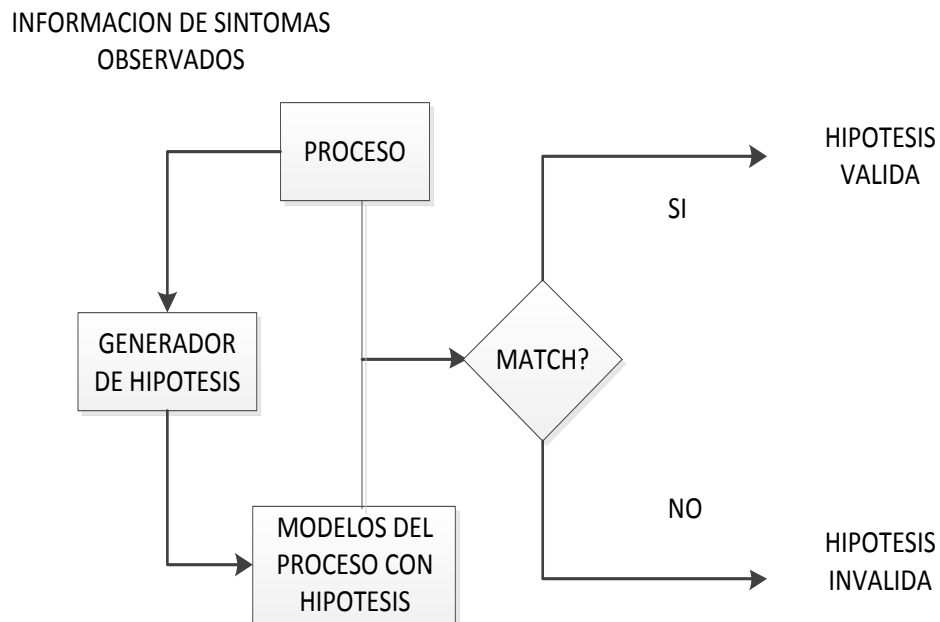


Figura. 23. Generación y prueba de hipótesis secuencial

Fuente: A review of process fault detection and diagnosis part I: quantitative model-based methods. Elsevier

- **Observador o enfoques basados en filtros:**

La idea básica de el observador o enfoque basado en filtros es calcular las salidas de el sistema de las mediciones (o subconjunto de mediciones) utilizando observaciones en un conjunto determinístico o filtros estadísticos (el filtro Kalman) en un conjunto estocástico. Luego, la (pesada) estimación de errores de salida (o innovaciones en el caso estadístico) son usados como residuales. Dependiendo de las circunstancias, se puede utilizar observadores lineales o no lineales, completos o de orden reducido, fijos o adaptativos.

- **Enfoques paridad espacial**

Las ecuaciones de paridad son re arregladas y generalmente variantes transformadas de la entrada-salida o los modelos de estado espacial de la planta. La idea básica es chequear la paridad (consistencia) de la los modelos de la planta con las salidas de los sensores (mediciones) y las entradas conocidas del proceso. La idea de este enfoque es reacomodar la estructura del modelo de tal forma que consiga la mejor aislación de la falla. La redundancia provee libertad en el diseño de la generación de ecuaciones residuales para que la aislación posterior de la falla se lleve a cabo. La aislación de la falla requiere la habilidad de generar vectores residuales los cuales son ortogonales entre sí por las diferentes fallas.

- **Enfoque de el dominio de la frecuencia (factorización)**

Los residuales también se pueden generar en el dominio de la frecuencia vía factorización de la función de transferencia del sistema monitoreado. Un estudio exhaustivo fue realizado por Dingo y Frank, en el cual el problema FDI fue sistemáticamente formulado y resuelto mediante las técnicas de factorización.

- **La estimación de parámetros**

Para la estimación de parámetros, los residuales son la diferencia entre el parámetro del modelo nominal y los parámetros estimados por el modelo.

Desviaciones en los parámetros del modelo sirven de base para detectar y aislar fallas.

- **Observadores**

El método basado en el observador reconstruye la salida del sistema de las mediciones o subconjuntos de mediciones con la ayuda de observadores. La diferencia entre las mediciones de salida y las salidas estimadas se utiliza como el vector de residuales.

- **Relaciones de paridad**

Este método chequea la consistencia de las ecuaciones matemáticas del sistema con las mediciones.

A.7.3.5.2 Métodos basados en modelos cualitativos

Basado en varias formas del conocimiento cualitativo utilizado en el diagnóstico de fallas, el enfoque basado en modelos cualitativo puede ser clasificado en dígrafos, arboles de falla y métodos físicos cualitativos.

El enfoque analítico requiere de modelos matemáticos cuantitativos precisos para llegar a ser efectivos. Para sistemas a gran escala, tal información puede estar no

disponible o puede ser muy costosa y consumir demasiado tiempo para obtenerla. Un método alternativo para el monitoreo del proceso es utilizar métodos basados en conocimiento tales como el análisis causal, sistemas expertos y reconocimiento de patrones. Estas técnicas están basadas en modelos cualitativos, los cuales se pueden obtener a través del modelamiento causal, conocimiento experto, una detallada descripción del sistema o ejemplos de falla- síntoma. Las técnicas de análisis causal se basan en el modelamiento causal de las relaciones de falla- síntoma. Las relaciones cualitativas y semi-cuantitativas en estos modelos causales pueden ser obtenidas sin usar principios primarios. Las técnicas de análisis causal incluyendo signed directed graph y el árbol de síntomas son primordialmente utilizados para el diagnóstico de fallas.

La Figura. 24 muestra la taxonomía del dominio basado en conocimiento sobre estas dos categorías.

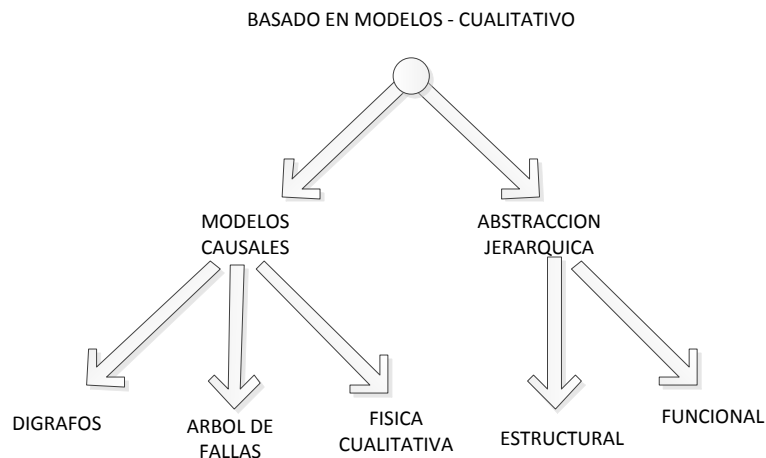


Figura 24. Formas de conocimiento cualitativo

Fuente: A review of process fault detection and diagnosis part II: qualitative models and search strategies Elsevier

- **Modelo causal usando dígrafos**

Las relaciones de causa-efecto o modelos pueden ser representados en la forma de signed digraphs(SDG). Un dígrafo es un grafo con arcos dirigidos entre los nodos y SDG es un grafo en el cual los arcos dirigidos tienen un signo positivo o negativo amarrado a ellos. Los arcos dirigidos van de los nodos “causa” a los nodos “efectos”. Los SDGs proveen una forma muy eficiente de representar modelos cualitativos gráficamente y han sido los más ampliamente usados del conocimiento causal para el diagnóstico de fallas en un proceso. Iri et al fueron los primeros en usar SDG para el diagnóstico de fallas. Del SDG derivaron lo que se llama grafico de causa-efecto. (CE Gráfico). Umeda et al mostro como SDG puede ser obtenido de ecuaciones algebraicas diferenciales para el proceso. En años recientes, Wilcox y Himmelblau se han aproximado a el problema del diagnóstico de fallas utilizando lo que se denomina posibles modelos gráficos causa – efecto (PCEG). Li y Wang han presentado como los di gráficos fuzzy pueden ser utilizados para simulaciones cuantitativas y cualitativas del comportamiento temporal del proceso.

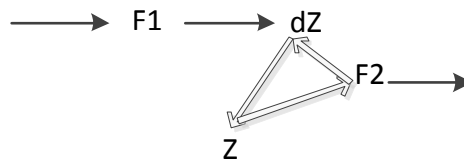


Figura 25. Ejemplo de dígrafo para un tanque simple

Fuente: A review of process fault detection and diagnosis part I: quantitative model-based methods. Elsevier.

- **Árbol de fallas**

Los árboles de fallas son utilizados en el análisis de la confiabilidad y seguridad del sistema. El análisis de árbol de fallas fue originalmente desarrollado en los laboratorios telefónicos Bell en 1961. El árbol de fallas es un árbol lógico que propaga eventos primarios o fallas hacia el evento o daño principal. El árbol generalmente tiene capas de nodos. En cada nodo diferentes operaciones lógicas como AND y OR se realizan para la propagación.

- **Enfoque físico cualitativo**

El conocimiento físico cualitativo en el diagnóstico de fallas ha sido representado primordialmente en dos maneras. El primer enfoque es derivar ecuaciones cualitativas de las ecuaciones diferenciales denominadas como ecuaciones de confluencia. El otro enfoque en la física cualitativa es la derivación de comportamientos cualitativos de las ecuaciones diferenciales ordinarias (ODEs). Estos comportamientos cualitativos para diferentes fallas pueden utilizarse como fuente de conocimiento. En términos de aplicaciones de modelos cualitativos en el diagnóstico de fallas, la simulación cualitativa (QSIM) y la teoría de proceso cualitativa (QPT) han sido los enfoques populares.

A.7.3.6 ENFOQUE DE DIAGNÓSTICO DE FALLAS DATA DRIVEN

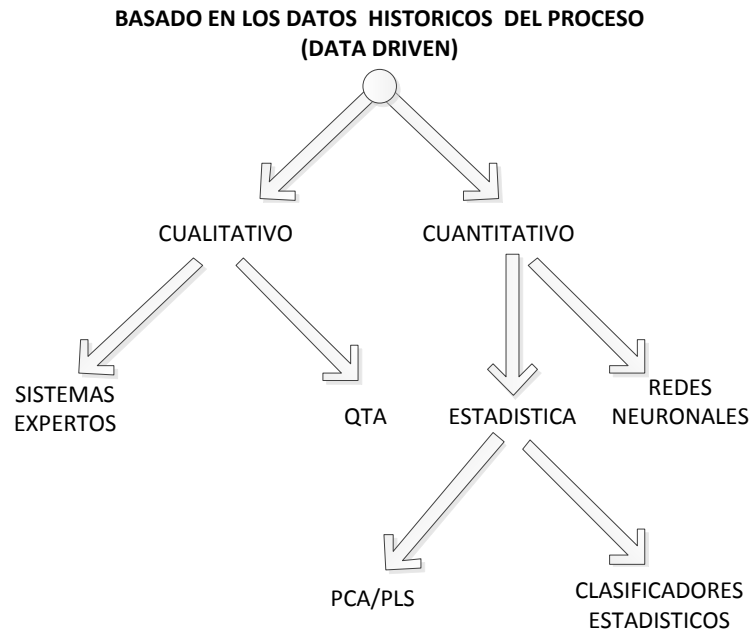


Figura 26. Clasificación de los métodos basados en la historia del proceso.

Fuentes: A review of process fault detection and diagnosis part III: process history based methods Elsevier

Las instalaciones de manufactura moderna son a grande a escala, de alta complejidad y operan con un gran número de variables sometidas al control de lazos cerrados (closed-loop). Un diagnóstico y detección de falla temprana y precisa para estas plantas puede minimizar los tiempos de inactividad, incrementando la seguridad de las operaciones en la planta, y reduciendo costos de manufactura. Las plantas están aumentando rápidamente su grado de instrumentación, resultando en más datos que quedan disponibles para utilizar en el diagnóstico y detección de fallas. Las cartas de control univariante (cartas Shewart) tienen una habilidad limitada para detectar y diagnosticar fallas en tales procesos que albergan múltiples variables. Esto llevó al surgimiento de un esfuerzo académico e industrial concentrado en desarrollar métodos más

efectivos en el monitoreo de los procesos. Un gran número de estos métodos ha sido aplicado regularmente a sistemas industriales reales, lo que hace a estas técnicas convenientes para cubrirlas en cursos de pregrado y postgrado. Además en las industrias de procesamiento y manufactura, ha existido un gran impulso hacia la generación de productos de alta calidad, la reducción de la rata de productos rechazados y a satisfacer las regulaciones cada vez más rigurosas de seguridad y medio ambiente. Operaciones en procesos que en una ocasión se consideraban aceptables ya no son adecuadas. Para alcanzar los altos estándares, los procesos industriales modernos contienen un gran número de variables operando bajo el control de lazo cerrado (under closed-loop control). Los controladores del proceso estándar (controladores PID, controladores modelo predictivo etc.) son diseñados para mantener operaciones satisfactorias compensando los efectos de perturbaciones y cambios ocurridos en el proceso. Aunque estos controladores pueden compensar muchos tipos de perturbación, hay cambios en los procesos los cuales no pueden manejar correctamente. Estos cambios son llamados fallas. De forma más precisa una falla se define como una desviación no permitida de al menos una propiedad característica o variable del sistema. Los tipos de fallas que ocurren en sistemas industriales incluyen cambios en los parámetros del proceso, cambios en los parámetros de perturbación, problemas en el actuador y problemas en el sensor. Catalizador de envenenamiento y obstrucciones en el intercambiador de calor son ejemplos de cambios en los parámetros del proceso. Un cambio en el parámetro de perturbación puede ser un cambio extremo en la concentración de un flujo de alimentos o en la temperatura ambiente. Un ejemplo de un problema en el actuador es una válvula pegada, y un sensor enviando mediciones parciales es un ejemplo de problemas en el sensor. Para asegurar que las operaciones del proceso satisfacen las especificaciones de rendimiento, los fallos en el proceso necesitan ser detectados, diagnosticados y removidos. Estas tareas están asociadas con el monitoreo del proceso. La meta del monitoreo del proceso es asegurar el éxito de las operaciones planeadas mediante el reconocimiento de

anomalías en el comportamiento. La información no solo mantiene al operador de la planta y al personal de mantenimiento mejor informados del estado del proceso. Como resultado de un adecuado monitoreo del proceso, los tiempos muertos se minimizan, la seguridad de las operaciones de la planta se ven mejoradas y los costos de manufactura se reducen. Como los sistemas industriales se han vuelto cada vez más integrados y complejos, la ocurrencia de fallas en procesos modernos presentan retos en el monitoreo que no son fácilmente abordados mediante las cartas de control univariante. La debilidad de las cartas de control univariante para detectar fallas en procesos multivariantes ha llevado a el surgimiento de investigaciones concentradas en desarrollar mejores métodos para el monitoreo del proceso. Este crecimiento en la actividad investigativa puede ser también explicado por el hecho de que los sistemas industriales se están volviendo más instrumentados, dando como resultado grandes cantidades de datos disponibles para utilizar en el monitoreo del proceso y que los computadores modernos cada vez son más potentes. La disponibilidad de datos recolectados durante varias operaciones y las condiciones de falla son esenciales para el monitoreo del proceso. La capacidad de almacenamiento y crecimiento de la velocidad computacional de los ordenadores modernos permiten algoritmos en el monitoreo del proceso para ser computados cuando se aplican a grandes cantidades de datos.

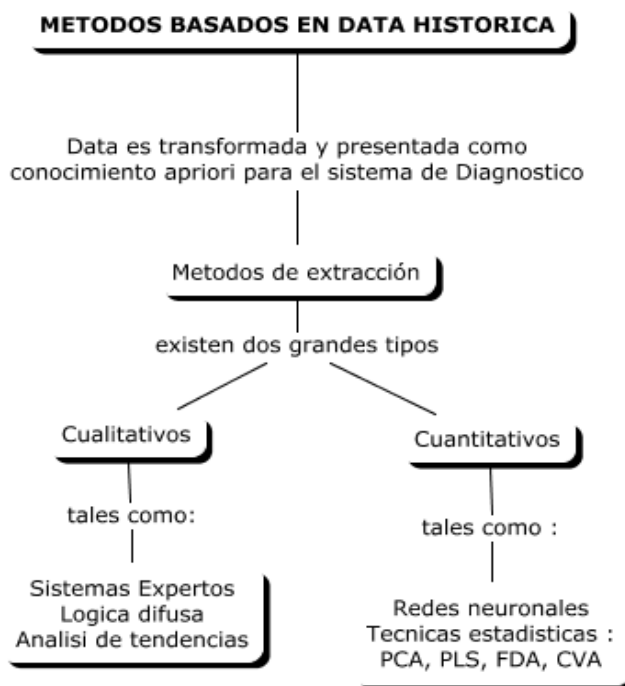


Figura 27. Clasificación de los métodos de extracción en Data- Driven

Fuente: autores

Los métodos de detección y diagnóstico de fallas basados en data histórica son los más aplicados en las industrias de procesos, debido a que son fáciles de implementar, y no necesitan un modelo matemático del proceso y requieren poco conocimiento a priori del proceso y las fallas.

El desarrollo de la instrumentación digital, las redes industriales y los sistemas de control SCADA y Control Distribuido, permiten almacenar un gran volumen de data de los procesos industriales. Esta data puede ser aprovechada para desarrollar sistemas de detección y diagnóstico de fallas; generando, a partir de esta información, patrones que permitan identificar condiciones similares de falla.

En los métodos basados en data histórica, la detección y diagnóstico de fallas se realiza mediante el procesamiento de un gran volumen de ella. Debido a esto, existen varias maneras en que la data pueda ser transformada y presentada como un conocimiento a priori para el sistema de diagnóstico. La primera etapa para realizar esta tarea es conocida como *extracción de características*, la cual puede ser cualitativa o cuantitativa. Los sistemas expertos, los métodos basados en lógica difusa y los análisis de tendencia cualitativos son ejemplos de métodos de extracción cualitativa. Para el caso de métodos de extracción cuantitativa, en la literatura aparecen ampliamente usadas las redes neuronales y las técnicas estadísticas tales como el Análisis de Componentes Principales (PCA), Mínimos Cuadrados Parciales (PLS), el Análisis Discriminante de Fisher (FDA), el Análisis de Variables Canónicas (CVA) y el Análisis de Componentes Independientes (ICA) han sido aplicadas en la detección y diagnóstico de fallas con anterioridad. FDA es un Método lineal para clasificar múltiples clases y permite reducir la dimensionalidad de la data. Ha sido aplicado con éxito en el diagnóstico de fallas, mostrando superioridad sobre PCA, PLS y CVA. La reducción de dimensionalidad es útil cuando el número de variables medidas en el proceso es grande mientras que el número de observaciones sea pequeño, hecho que es muy común en las industrias de procesos actuales, donde se mide y almacena gran cantidad de variables. Los resultados con FDA muestran superioridad respecto a PCA.

El uso de la transformada Wavelet ha sido investigado en los últimos años, debido a su habilidad de correlacionar data auto correlacionada. Los coeficientes Wavelet proveen información Compacta acerca de una señal en diferentes localizaciones de tiempo y frecuencia, lo que permite también reducir la dimensionalidad de la data, pero en este caso en la dirección del tiempo, a diferencia del FDA que lo hace en la dirección de las variables. En este trabajo no se realizará una explicación teórica de la misma, por lo que se recomienda a los lectores interesados consultar.

Para el caso de que la data requiere de un clasificador no lineal se propone el uso del Análisis Discriminante Generalizado (GDA), también llamado Análisis Discriminante Linear Kernel (KLDA), el cual ha sido aplicado con éxito en la clasificación de patrones, obteniéndose buenos resultados en la clasificación de semillas y las flores Iris de Fisher. También ha mostrado superioridad ante otras técnicas como el Análisis de Componentes Principales Kernel (KPCA) y el Análisis de Componentes Independientes Kernel (KICA) en el reconocimiento de voz.

A.7.3.6.1 Métodos cuantitativos de data driven

Los métodos que extraen información cuantitativa en general se pueden clasificar en métodos no estadísticos y métodos estadísticos. Las redes neuronales son una clase importante de clasificadores no estadísticos. El análisis de componente principal (PCA)/ mínimos cuadrados parciales (PLS) y clasificadores de patrones estadísticos forman el mayor componente de los métodos de extracción de características estadístico.

- **Análisis discriminante de Fisher (FDA)**

FDA es una técnica de reducción de dimensionalidad desarrollada y estudiada en la comunidad de clasificación de patrones. El FDA determina la porción del espacio observado que es más efectivo en la discriminación de varios tipos de datos. El análisis discriminante se aplica a esta porción de espacio observado para el diagnóstico de fallas. La técnica de reducción de dimensionalidad es

aplicada a los datos en todas las clases simultáneamente. Así, toda la información de las clases de fallas es utilizada cuando la función discriminante es evaluada para cada clase y se espera un mejor rendimiento en el diagnóstico de las fallas. Los desarrollos teóricos para el FDA sugieren que debería ser más efectivo que el diagnóstico de fallas usando PCA.

- **Mínimos cuadrados parciales (PLS)**

Son métodos para descomponer los datos para maximizar la covarianza entre el bloque predictor (independiente) y el bloque predictivo (dependiente) para cada componente. El PLS pretende encontrar los vectores de carga y de puntuación que están interrelacionados con el bloque predictivo X mientras describe una gran cantidad de la variación en el bloque predictor Y. una aplicación popular del PLS es la de seleccionar X para contener los datos del sensor y Y para contener solo los datos de la calidad del producto.

En forma similar al PCA, tales modelos inferenciales (también conocidos como soft sensors) pueden ser utilizados para detectar, identificar, y diagnosticar fallas. Otra aplicación del PLS en el diagnóstico de fallas consiste en definir Y como perteneciente a una clase. Este método PLS es conocido como transformaciones discriminantes de mínimos parciales de los residuales será relativamente grande cuando existan fallas presentes, y pequeño en la presencia de alteraciones, ruido y/o errores del modelamiento. Esto permite la definición de umbrales que permiten detectar la presencia de fallas. Las tres formas primordiales para generar residuales son: la estimación de parámetros, observadores y relaciones de paridad.

- **Enfoque estadístico multivariante.**

Las aplicaciones exitosas de métodos estadísticos multivariantes al diagnóstico de fallas tales como el análisis del componente principal (PCA) y mínimos cuadrados parciales (PLS) han sido reportadas en forma extensa en la literatura. Una limitación mayor del monitoreo convencional con PCA es que el modelo PCA es invariante con el tiempo, mientras que la mayoría de los procesos reales son variantes en el tiempo. Por lo tanto el modelo PCA debería además ser recursivamente actualizado.

- **Clasificadores estadísticos**

El diagnóstico de fallas es esencialmente un problema de clasificación y por ello puede juntarse con la clásica estructura de reconocimiento de patrones estadísticos. El diagnóstico de fallas puede ser considerado como un problema que combina, con el tiempo, la estimación instantánea del clasificador usando conocimiento acerca de las propiedades estadísticas de los modos de falla del sistema.

- **Redes neuronales**

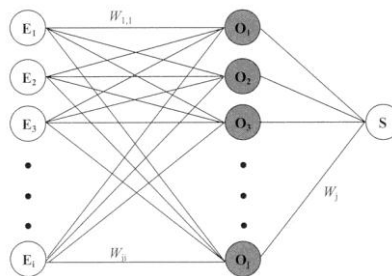


Figura 28. Ejemplo de una red neuronal

Fuente: A review of process fault detection and diagnosis part III : process history based methods

Se ha notado un interés considerable en la aplicación de redes neuronales para el diagnóstico de fallas. En general las redes neuronales que han sido utilizadas en el diagnóstico de fallas pueden clasificarse en dos dimensiones: 1) la arquitectura de la red tal como la sigmoidea, base radial y demás; y 2) la estrategia de aprendizaje como el aprendizaje supervisado y no supervisado. La estrategia de aprendizaje supervisado más popular en redes neuronales ha sido el algoritmo de propagación inversa. Hay numerosas publicaciones que direccionan el problema de diagnóstico de fallas hacia las redes neuronales usando propagación inversa.

	ORTOGONAL	NO ORTOGONAL
LOCAL	WAVELETS	RBFS
GLOBAL	POLINOMIAL TRIGONOMETRICO	SIGMOIDES

Figura 29. Características de los nodos usados en redes neuronales.

Fuente: A review of process fault detection and diagnosis part III : process history based methods

A. 7.3.6.2 Métodos data driven cualitativos

Dos grandes métodos que extraen información histórica cualitativa son sistemas expertos y el análisis de tendencia cualitativo (QTA).

- **Sistemas expertos**

La extracción de características basada en reglas ha sido ampliamente utilizada en sistemas expertos para muchas aplicaciones. Un sistema experto es generalmente un sistema muy especializado que resuelve problemas en un dominio angosto de experticia.

Los sistemas expertos se utilizan para imitar el razonamiento de un humano experto cuando diagnostica las fallas. La experiencia de un dominio experto puede ser formulado en términos de reglas, las cuales se pueden combinar con el conocimiento de los primeros principios o la descripción estructural del sistema para el diagnóstico de las fallas. Los sistemas expertos son capaces de capturar asociaciones a diagnósticos humanos que no son fácilmente traducidas a modelos matemáticos o causales. Las técnicas de reconocimiento de patrones usan la asociación entre patrones de datos y clases de fallas sin un modelamiento específico de los estados internos del proceso o su estructura. Ejemplos incluyen a las redes neuronales y mapas auto-organizados. Estas técnicas están relacionadas con las técnicas aplicadas en data-driven (PCA, PLS, FDA, y CVA) en términos del modelado la relación entre patrones de datos y clases de fallas. Las técnicas de data-driven son técnicas reductoras de dimensionalidad basadas en la estadística multivariable. Por otro lado, las redes neuronales y los

mapas auto-organizados son métodos de caja negra que aprenden el patrón basado enteramente de las sesiones de entrenamiento.

- **Tendencia de el análisis cualitativo**

El análisis de tendencia y predicción son componentes importantes del proceso de monitoreo y control de supervisión. La tendencia de modelado puede utilizarse para explicar los eventos importantes que suceden en un proceso, para diagnosticar el mal funcionamiento y predecir futuros estados.

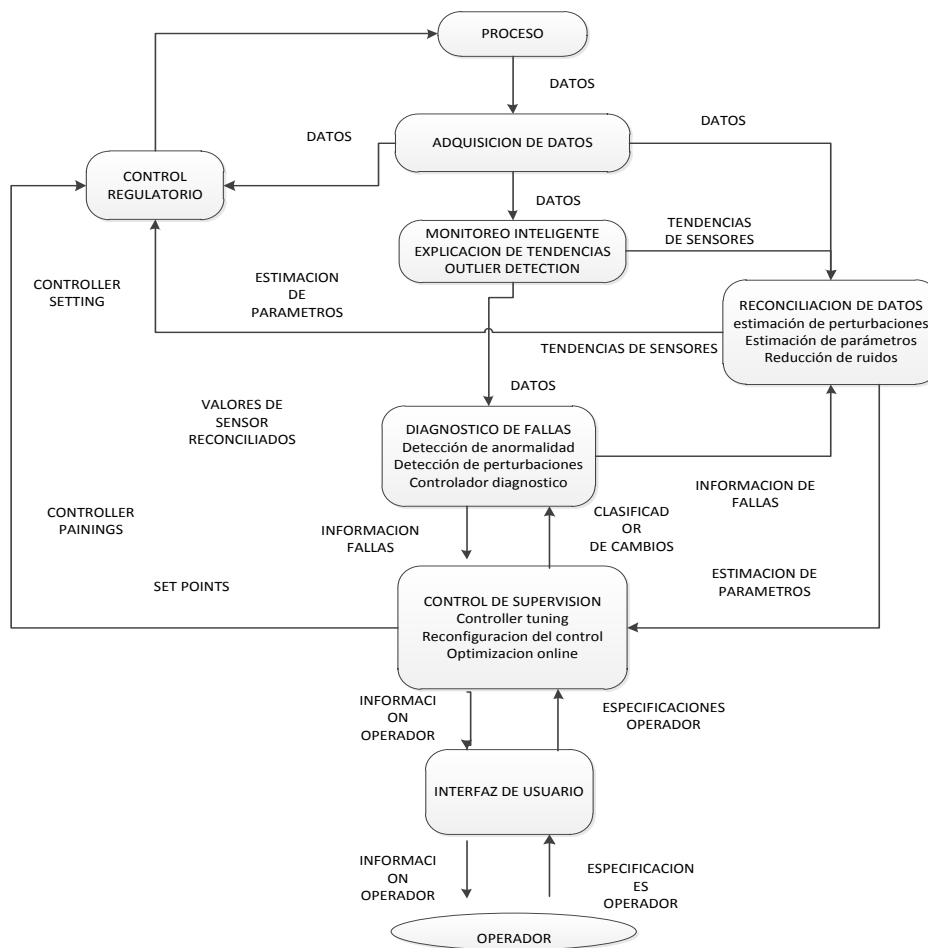


Figura 30. Marco de integración.

Fuente: A review of process fault detection and diagnosis part III: process history based methods Elsevier

A.7.4 EJEMPLOS DE APLICACIÓN DE LOS MÉTODOS DE FDD MEDIANTE EL ENFOQUE DE DATA-DRIVEN

A.7.4.1 DATA DRIVEN DIAGNÓSTICO DE FALLAS Y CONTROL TOLERANTE DE FALLAS:ALGUNOS AVANCES Y POSIBLES NUEVAS DIRECCIONES

En este documento presenta una visión selecta que cubre los avances en diagnóstico de fallas y control tolerante de fallas mediante las técnicas de Data Driven. Un ejemplo de aplicación exitoso es también mostrado, cuando se tiene que enfrentar con fallas causadas por la desubicación los puntos fijados del lazo de control y varias áreas de futuras direcciones potenciales son incluidas en este documento.

A.7.4.2 DATA-DRIVEN METODOLOGÍA DE RED NEURONAL EN LA PREDICCIÓN DE TIEMPO DE VIDA RESTANTE PARA LOS COMPONENTES DEL ACTUADOR EN UN AVIÓN

Los actuadores son mecanismos complejos electro-hidráulicos o mecánicos utilizados en aeronaves para manejar controles superficiales de vuelo, llantas de aterrizaje, puertas de carga, y sistemas de armamentos. Impact ha desarrollado

una metodología de pronóstico y manejo sin errores (PHM, pronostic and health) para éstos sistemas críticos que incluyen procesamiento de señales y técnicas de seguimiento a redes neuronales, a lo largo con razonamiento automatizado, clasificación, fusión de conocimiento, y algoritmos de progresión en modos de falla probabilísticos. El procesamiento utiliza la señal comando/respuesta y datos de la presión hidráulica de los actuadores y provee de una valoración en tiempo real del estado actual/futuro de la condición del actuador. Ésta metodología fue aplicada a las servo válvulas del estabilizador electro-hidráulico (EHSVs) usando el banco de datos provisto por los trabajos Boeing Phantom. El modulo automatizado demostró excelentes resultados clasificados sin errores. El pronóstico fue también exitosamente realizado sin embargo no se tuvo datos para validar la predicción. Los algoritmos fueron desarrollados teniendo en cuenta las limitaciones por sensor/procesamiento para la implementación a bordo. Varios de los elementos PHM presentados pueden ser adaptados para otros tipos de actuadores y aplicaciones.

A.7.4.3 DETECCIÓN Y DIAGNÓSTICO DE FALLAS UTILIZANDO TÉCNICAS ESTADÍSTICAS MULTIVARIABLES A PARTIR DE DATA HISTÓRICA.

En este trabajo son aplicadas algunas técnicas estadísticas multivariantes para realizar la detección, identificación y diagnóstico de fallas en un proceso industrial. El proceso de un tanque de reacción no isotérmico agitado continuamente (CSTR) fue simulado en presencia de fallas típicas, para generar la data histórica de donde se obtiene un patrón que luego permite identificar una falla similar en el futuro.

PCA, FDA y GDA son algunas de las técnicas estadísticas que son estudiadas en este trabajo.

Para probar la eficiencia de estas técnicas estadísticas es necesario contar con una data histórica de un proceso real en presencia de fallas, debido a que este tipo de data no se encuentra a la mano públicamente por motivo de seguridad de las empresas, muchos investigadores han optado por probar los métodos en procesos industriales simulados en computador. El proceso de un tanque de reacción no isotérmico agitado continuamente (CSTR) fue simulado para generar a partir del, la data histórica.

Este proceso se ha utilizado previamente en un buen número de publicaciones.

A.7.5 SENSORES Y ESTRATÉGIAS DE MEDICIÓN

Los sensores y estrategias de detección constituyen los fundamentos básicos para el diagnóstico y pronóstico de fallas. Los hechos estratégicos que se generen con el uso de sensores para recoger datos que eventualmente darán lugar a la realización en línea de algoritmos de diagnóstico y pronóstico están relacionados con el tipo, número, y localización de los sensores; su tamaño, peso, costo, rango dinámico y otras propiedades características; si son de la variedad por cable o inalámbrica, etc. Los datos recogidos por dispositivos de transducción rara vez son útiles en su forma primaria. Estos deben ser procesados apropiadamente y así la información útil puede ser extraída en una versión reducida de los datos originales pero conservando los posibles rasgos característicos o indicadores de los eventos de falla que se están tratando de detectar, aislar y predecir en el tiempo. Así, estos datos deberán ser procesados, lo que significa, filtrado, comprimido, correlacionado, etc., para remover artificios y reducir los niveles de ruido y el volumen de datos que deben ser procesados

posteriormente. Por otra parte, el sensor provee datos que deben ser validados, es decir, los sensores por sí mismos no identifican condiciones de falla. Una vez el módulo de pre-proceso confirma que el dato del sensor está limpio y apropiadamente formateado, los rasgos y señales de condiciones de normalidad o falla deberán ser extraídos. Esto es el paso más significativo en la estructura de diagnóstico oportuno y preciso de los modos de falla.

Los sensores son específicos para el dominio de la aplicación, y están destinados para controlar variables típicas del estado sensible como la temperatura, presión, velocidad, vibraciones, etc. Algunos sensores son insertados específicamente para medir cantidades que están directamente relacionadas con modos de falla como candidatos para el diagnóstico. Entre ellos se encuentran los calibradores de tensión, los sensores ultrasónicos, dispositivos de proximidad, sensores de emisión acústica, los sensores electroquímicos de fatiga, interferómetros, etc., mientras que otros son de la variedad de usos múltiples, tales como la temperatura, velocidad, caudal, etc., y están diseñados para controlarlas variables del proceso para el control y/o de evaluación del desempeño, además de su uso en el diagnóstico, en este campo cabe resaltar que en los últimos tiempos hemos sido testigos de la introducción de dispositivos inalámbricos. Las principales preguntas siguen en el dominio del sensor y de la detección. Algunas técnicas analíticas son capaces de ayudar en la optimización de la colocación y determinación del tipo y número de sensores para una configuración dada del sistema. El procesamiento de datos para caracterizar es principalmente un arte que confía en la intuición y la experiencia del diseñador en cuanto en donde está la información de falla y cómo hacer para extraerla. Las características de selección y extracción pueden seguir varios caminos. Pueden ser extraídos en el dominio del tiempo, de la frecuencia, como onda o incluso dominios caóticos, dependiendo el evento de falla. El vector característico resultante debe ser óptimo en términos de criterios relativos a su capacidad de discriminación para detectar y aislar una falla en particular, mientras que cumple con los requisitos de cómputo y

otras limitaciones. En este capítulo se abordan algunas de las preocupaciones fundamentales en el importante ámbito de los sensores y las estrategias de detección.

A.7.5.1 SENSORES

En esta sección trataremos los principios de transducción y el enfoque de sistemas de sensores que encuentran una extensa utilidad en el campo del diagnóstico y detección de fallas.

Una clase de sensor significativo dentro del dominio de la aplicación incluye acelerómetros (para medidas de vibraciones), sensores de corrientes inducidas (corrientes de Foucault) de emisiones ultrasónicas o acústicos (detección de grietas y otras anomalías), y dispositivos de transducción que aprovechan los recientes avances sistemas electro-mecánicos (MEMS, por sus siglas en inglés) y fibra óptica. Nos enfocaremos en esta sección solamente en unos cuantos típicos e innovadores sensores que tienen un uso generalizado o prometen expandir su utilidad en aplicaciones relacionadas con diagnóstico. Introduciremos además una breve discusión sobre las redes de sensores. Los protocolos y estándares para sensores, interfaces y redes que han sido establecidas y son actualizados continuamente porque la detección y el campo de redes está afectando cada vez más cada aspecto del monitoreo y control de los complejos procesos industriales, sistemas aeroespaciales, etc.

Acerca del hardware de sensores para las aplicaciones de diagnóstico se refieren a la utilidad de transductores, MEMS, redes e interfaces, multitareas basado en DSP, y plataformas multiproceso para la eficiente, confiable y económica implementación de las rutinas de software. La configuración de hardware debe ser capaz de medir múltiples cantidades y eventualmente, combinar la información en conjunto, haciendo uso de las relaciones estadísticas, herramientas de análisis de componentes principales (PCA), filtrar, etc. para reducir los efectos del ruido y las perturbaciones. Los sistemas de metrología dinámica deben ser no intrusivos, operando en línea con una gran banda ancha, confiable y fácil de operar. Una estructura metrológica integrada hardware/ software tiene como objetivo la estabilización de los fundamentos y principios de operación de sensores y metodologías de procesamiento de señales para uso de diagnóstico en línea para una variedad de sistemas mecánicos, electromecánicos y otras ingenierías.

Los sensores y las estrategias de detección constituyen la mayoría de la tecnología con un amplio dominio desde la ingeniería y las ciencias hasta la medicina y otras disciplinas. Por lo general, se determina que dos clases de sensores están logrando avances significativos en los sistemas de monitoreo para el diagnóstico y pronóstico de fallas. El primero de ellos se refiere al clásico o tradicional transductor usado para monitorear máquinas, estructuras, desarrollo y operación, y propiedades eléctricas/electrónicas que se relacionan con mecanismos de falla de máquinas, estructuras y sistemas eléctricos. En esta categoría, incluimos dispositivos que miden propiedades de flujo, termodinámicas, térmica y mecánicas de una variedad de sistemas o procesos – turbinas de gas, vehículos terrestres, bombas, sistemas aeroespaciales, etc. La segunda categoría importante se refiere a los sistemas de sensores que se posicionan casi exclusivamente para interrogar y trazar las propiedades del sistema que están relacionadas con sus mecanismos de falla. Resumiremos brevemente en este capítulo algunos sensores que pertenecen a la primera categoría y nos

extenderemos en un nuevo e innovador sensor que esta específicamente diseñado para aplicaciones de diagnóstico.

A.7.5.1.1 Transductores y principios de transducción

Un transductor es definido como un dispositivo que recibe energía de un sistema y la retransmite, a menudo de una forma diferente a otro sistema. Por otro lado un sensor es definido como un dispositivo que es sensible a la luz, temperatura, impedancia eléctrica, o nivel de radiación y transmite una señal a un medidor o dispositivo de control. Un dispositivo de medida pasa por dos etapas mientras mide una señal. Primero, la magnitud (una cantidad física como aceleración, presión, tensión, temperatura, etc.) es detectada. Luego, la señal de medida es transformada en una forma particular adecuada para ser transmitida, acondicionada y procesada. Por esta razón, la salida del transductor es a menudo una señal eléctrica que es luego digitalizada. El sensor y la etapa de transducción de un dispositivo típico de medida son representados esquemáticamente en la Figura 31.

Usualmente un elemento de transducción se basa en un principio físico que con un cambio en el valor medido se altera la salida del transductor (señal eléctrica). Hay sólo un puñado de principios físicos detrás de muchos miles de transductores disponibles en el mercado. Alloca y Stuart (1984) dieron una completa lista de transductores clasificados de acuerdo a diferentes principios físicos.

Mencionaremos unos cuantos ejemplos que tienen gran relevancia para el diagnóstico y pronóstico en sistemas mecánicos.

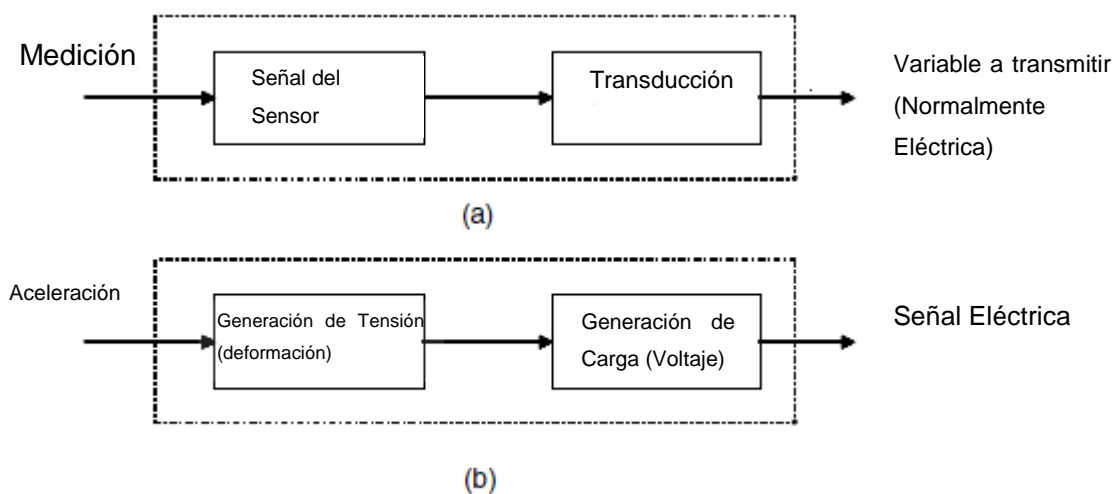


Figura 31. (a) Representación esquemática de un dispositivo de medida. (b) Operación de un acelerómetro piezoeléctrico.

Fuente: Intelligent Fault Diagnosis and Prognosis for Engineering Systems. George Vachtsevanos et al.

Un rasgo deseable para la transducción es tener una estática (no dinámica) relación entrada/salida tal que la salida instantáneamente alcance el valor de entrada (o la variación de la medida). En este caso, la función transferida es ganancia pura. Esto sucede cuando las constantes de tiempo del transductor son pequeñas (la banda ancha del transductor es grande). Sin embargo, en la práctica, los efectos dinámicos del transductor algunas veces no son despreciables debido

a las limitaciones en el tamaño de los componentes y los tipos de materiales usados en la construcción. Usualmente la constante eléctrica de tiempo de un transductor es de magnitud más pequeña que la constante de tiempo mecánico (de Silva, 1989). Cuando escogemos un sistema transductor, es importante atender a las especificaciones tales como, tiempo de subida, banda ancha, linealidad, saturación, etc.

A.7.5.2 SISTEMAS MECÁNICOS Y ESTRUCTURALES DE SENSORES.

Los sistemas mecánicos y estructurales de sensores han sido estudiados extensamente, y un gran número de estos dispositivos están frecuentemente en uso en plataformas de sistema de monitoreo para etapa operativa evaluando y siguiendo indicadores de falla. Un número de cantidades mecánicas – posición, velocidad, aceleración, torque, tensión, etc.- son comúnmente usadas en los sistemas dinámicos. Los dispositivos convencionales para medir estas cantidades están disponible comercialmente (www.sensorland.com/TorqPage001.html, www.oceansensor.com), y su operación ha sido ampliamente descrita en libros y publicaciones técnicas (Allocca y Stuart, 1984; de Silva, 1989). Usaremos solamente dos sistemas mecánicos de sensores- acelerómetros para monitoreo de vibraciones y medidores de deformación- para ilustrar sus características de funcionamiento y resaltar su utilidad en los sistemas de aplicaciones de diagnóstico y detección de fallas. El lector es referido a literatura voluminosa de otros sensores mecánicos de interés. Además debido a que el monitoreo de componentes estructurales y la investigación de causas ha atraído el interés en los últimos años, se estudiará brevemente sobre la detección ultrasónica, que es un medio típico de medir cantidades relevantes de interés.

A.7.5.2.1 Acelerómetros: Medición de vibraciones.

Recientemente que se ha requerido incrementar el interés por un mayor entendimiento de las causas de la vibración y de la respuesta dinámica de las fallas de estructuras y máquinas por fuerzas de vibración. Un exacto y confiable transductor de vibraciones robustas por consiguiente es requerido para monitorear en línea los componentes críticos y las estructuras. Los acelerómetros piezoeléctrico ofrecen un amplio rango dinámico y se encuentran entre los aparatos óptimos para monitoreo de vibraciones. Estos exhiben propiedades deseables como:

- Aplicabilidad para amplios rangos de frecuencia.
- Excelente linealidad para rangos dinámicos amplios.
- Integración electrónica de la señal de aceleración para dar datos de velocidad y desplazamiento.
- Medición de vibraciones para amplios rangos de condiciones ambientales mientras mantienen una excelente exactitud.
- Auto generación de la energía de funcionamiento.
- No tienen partes móviles y por lo tanto tienen una durabilidad extrema
- Extremadamente compactos sumado a una alta sensibilidad por unidad de masa.

Los acelerómetros piezoeléctricos son usados para medir todo tipo de vibraciones sin importar su naturaleza o comportamiento en el tiempo o frecuencia, siempre que el acelerómetro tenga la frecuencia y el rango dinámico correcto. Un modelo básico simplificado de un acelerómetro piezoeléctrico es mostrado en la Figura 32.. (Seridge yLicht, 1987).

Los elementos piezoeléctricos mostrados tienen una rigidez k equivalente y constituyen el elemento activo del transductor. Actúa como un resorte que conecta la masa de la base m_b con la masa sísmica m_s . Una fuerza igual al producto de esta fuerza de vibración actúa sobre el elemento piezoeléctrico, produciendo una carga proporcional a la fuerza aplicada. Como la masa sísmica acelera con la misma magnitud y fase que la base del acelerómetro en un rango de frecuencia amplio, la salida del acelerómetro es proporcional a la aceleración de la base y por lo tanto de la aceleración de la estructura dentro de la que el dispositivo es montado. De la ecuación del movimiento del modelo mostrado en la figura 32., fácilmente podemos llegar a la siguiente relación para la relación R/R_0 donde R_0 es el desplazamiento de las masas a la frecuencia por debajo de la frecuencia natural de resonancia del acelerómetro ($\omega \ll \omega_n$), asumiendo que el desplazamiento varía sinusoidalmente, y R es el desplazamiento a la frecuencia más alta. Así el desplazamiento entre la base y la masa sísmica aumenta cuando la frecuencia forzada es comparable con la frecuencia natural del acelerómetro. En consecuencia, la fuerza sobre el elemento piezoeléctrico y la salida eléctrica del acelerómetro también aumenta. El rango de frecuencia en el cual el acelerómetro puede usarse es determinado por la frecuencia de resonancia montada en el dispositivo. Hay un equilibrio entre el rendimiento de alta frecuencia y la sensibilidad del transductor. Los materiales piezoeléctricos comúnmente usados incluyen titanato de plomo, circonato, titanato de plomo, cristal de cuarzo, etc. La figura 33. muestra un acelerómetro piezoeléctrico Bruel y Kjaer típico, mientras que la figura 34. Representa un acelerómetro triaxial fabricado por Bruel y Kjaer.

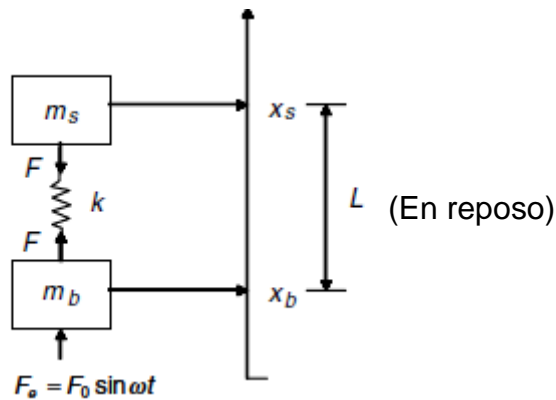


Figura 32. Modelo simplificado de un acelerómetro.

Fuente: Intelligent Fault Diagnosis and Prognosis for Engineering Systems. George Vachtsevanos et al.

A.7.5.2.2 Medidores de deformación.

Un sensor de deformación está basado en un simple principio de la electrónica básica, la resistencia de un conductor es directamente proporcional a su longitud y resistividad, e inversamente proporcional al área de sección transversal. Aplicando tensión o esfuerzo al elemento metálico del transductor ocasiona una variación en la longitud y en la sección transversal, causando así un cambio en la resistencia que puede ser medida como una señal eléctrica. Ciertas sustancias, como los semiconductores, muestran efectos piezo-resistivos, en los que la aplicación de la tensión afecta en gran medida su resistencia. Medidores de deformación de este tipo tienen una sensibilidad de aproximadamente dos órdenes mayor que el primer tipo.

El transductor usualmente es usado con un arreglo del puentede Wheatstone, con uno, dos o los cuatro brazos midiendo una tensión individual, tal que el cambio en la salida de voltaje es una indicación de que la medida (la tensión) cambia.



Figura 33. Un Acelerómetro piezoeléctrico de Bruel an Kjaer (tipo 7704 A-50)

Fuente: Intelligent Fault Diagnosis and Prognosis for Engineering Systems. George Vachtsevanos et al.



Figura 34. Acelerómetro triaxial miniatura Delta Tron (tipo 4524)

Fuente: Intelligent Fault Diagnosis and Prognosis for Engineering Systems. George Vachtsevanos et al.

- *Posicionamiento del Extensómetro.* Para el caso particular de un álabe de turbina, cuatro medidores de tensión son posicionados y conectados para detectar modos de flexión y vibración (Lackner, 2004). Están dispuestos de tal manera que estén paralelos con los bordes anterior y posterior del álabe. Dos extensómetros están en un lado del álabe, y los otros dos indicadores son imágenes de espejo en el otro lado. Los extensómetros se colocan todos aproximadamente a un tercio del espacio del álabe. Cada medidor corresponde a un cuarto del arreglo. La posición de cada uno de los medidores es mostrada en la figura 35. Los extensómetros están resaltados en negro, y la línea punteada indica el cambio de lado del álabe.
- *Extensómetro - Esquema de cableado.* Los calibradores de tensión se conectan como un puente Wheatstone. Un diagrama de esto muestra en la figura 36. Cada resistencia representa uno de los medidores de deformación o extensómetros.

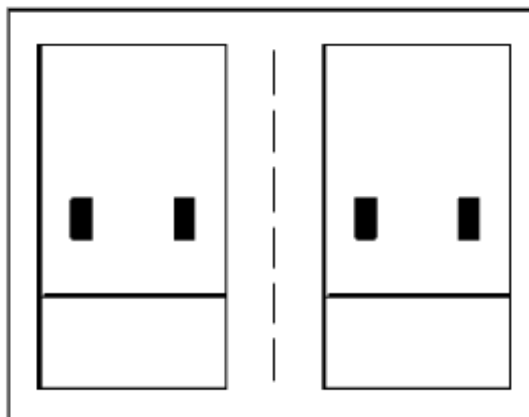


Figura 35.. Posicionamiento del Extensómetro

Fuente: Intelligent Fault Diagnosis and Prognosis for Engineering Systems. George Vachtsevanos et al.

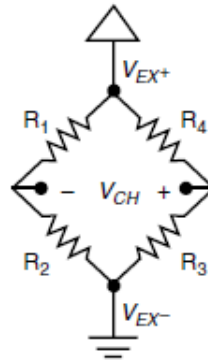


Figura 36. Puente Wheatstone

Fuente: Intelligent Fault Diagnosis and Prognosis for Engineering Systems. George Vachtsevanos et al.

Cuando un medidor se somete a tensión, la resistencia se incrementa. Del mismo modo, cuando se somete a la compresión, la resistencia disminuye. Los cuatro medidores de deformación se pueden conectar entre sí de tal forma que tengan un puente completo. Considere el cada lado del álabe. Cuando el álabe se curva, dos de los medidores de un lado están a fuerza de tensión mientras que los opuestos están a compresión. V_{EX} es el voltaje de excitación, el cual es proporcionado por el amplificador para alimentar el puente. Si los indicadores están conectados de tal manera que los dos que la experimentan tensión son R_1 y R_3 en el puente de Wheatstone, y los dos indicadores que están a compresión son R_2 y R_4 , entonces aparecerá una tensión V_{CH} . Este voltaje, es la señal de salida.

Un puente Wheatstone es una configuración avanzada para extensómetros por varias razones. Primero, un puente elimina los efectos de los esfuerzos centrífugos. Cuando el álabe gira, la fuerza centrífuga produce una tensión a lo largo del álabe. Sin embargo, los cuatro están en el mismo lugar a lo largo de la

envergadura, por lo que experimentan la misma fuerza tensión. Dado esto, la resistencia de cada indicador aumenta en la misma cantidad. El resultado es que el puente no se ve afectado por la tensión centrífuga, y VCH es cero. De la misma manera los efectos térmicos se anulan. Como el álabe que girase calienta debido a la fricción del rotor y la aerodinámica, el álabe también tiende a alargarse un poco. Una vez más, los cuatro indicadores de álabe experimentan la misma fuerza, por lo que la salida del puente es cero. La salida del amplificador de tensión del medidor es una señal de tensión simple que puede estar conectado a un cable BNC(del inglés Bayonet Neill-Concelman 7.32). El cable BNC puede ser conectado a un osciloscopio para ver la salida de la tensión o al sistema de adquisición de datos para extraer las respuestas del medidor.


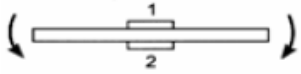
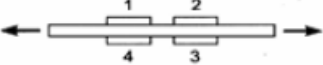
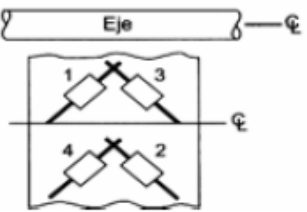
En la tabla 7.1 se observan otros arreglos que se pueden usar para la conexión de extensómetro y se relacionan las respectivas compensaciones.



Figura 37 Cable coaxial RG58 con dos conectores machos BNC.

Fuente: Belkin Components

Tabla 2. Arreglos de puentes y sus compensaciones.

	ARREGLOS	COMPENSACIÓN
	Un solo medidor bajo esfuerzo uniaxial	Ninguna
	Dos medidores detectan deformaciones iguales pero opuestas: arreglo típico para flexión	Temperatura
	Cuatro medidores con pares que detectan deformaciones iguales pero opuestas	Temperatura y flexión
	Cuatro medidores con pares que detectan deformaciones iguales pero opuestas: sensible a la torsión, arreglo típico de eje.	Temperatura y axial

Fuente: *Publicación de Scientia et Technica Año XIII. Universidad Tecnología de Pereira*

A.7.5.2.3 Sensor ultrasónico.

Sistemas de sistemas de sensores de ultrasonidos se consideran para monitorear la salud de estructuras críticas como aviones, puentes y edificios. Los métodos de ultrasonido son particularmente útiles para esto tanto porque las ondas ultrasónicas viajan a largas distancias y así tiene el potencial de monitorear un gran volumen de material y porque los métodos de ultrasonido han demostrado su utilidad para inspecciones no destructivas (NDI, por sus siglas en inglés) de estas estructuras durante el mantenimiento. Hay tres tipos de ondas de ultrasonido que son útiles para el monitoreo de estructuras: ondas guiadas, en masa y difusas. Independientemente del tipo de onda, la estrategia es monitorear los cambios y

luego detectar, localizar y caracterizar el daño basado en la naturaleza del mismo. Esta estrategia de búsqueda permite una sensibilidad de detección similar al NDI a pesar de la limitación de los sensores fijos.

La figura 38. muestra ejemplos de transductores ultrasónicos montados para componentes metálicos. En la figura 38a, un simple disco piezoeléctrico es encapsulado en una masa epóxica y unido a la superficie de un componente de aluminio. La frecuencia nominal es 2.25 MHz, pero los discos son transmisores y receptores de banda ancha, efectivamente operando por debajo de 50Khz hasta por encima de los 3Mhz. El mismo tipo de discos también se pueden embalar en una caja con un conector. La figura 38b muestra transductores de haz angular disponibles comercialmente adjunto a un espécimen; la forma de U es por protección mecánica. Estos transductores tienen una frecuencia central de 10 Mhz y pueden ser usados para monitorear pequeñas grietas.

Algunos métodos de ultrasonido son descritos a continuación:

- *Ondas Guiadas.* Las ondas ultrasónicas guiadas requieren una o más interfaces para direccionar la onda de propagación y poder propagarla a largas distancias mientras se mantiene la suficiente energía para una recepción útil. Las ondas Rayleigh, u ondas de superficie, son adecuadas para el seguimiento de superficie o defectos cerca de la superficie, su penetración es aproximadamente la longitud de onda.

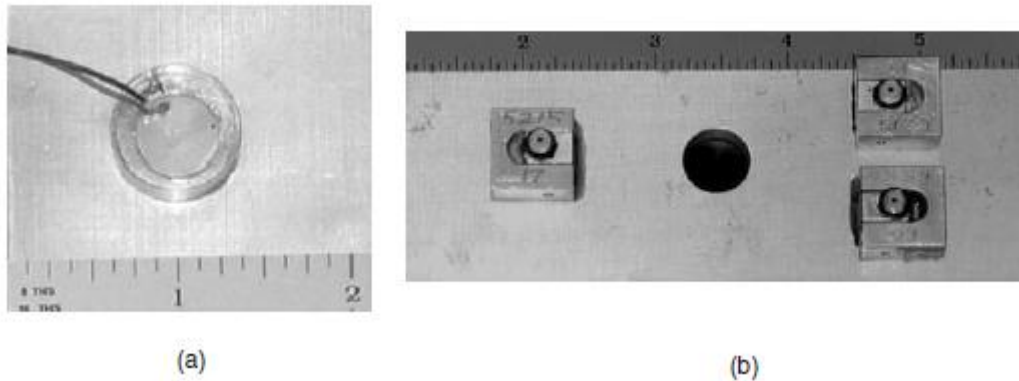


Figura 38. Transductores ultrasónicos adheridos a partes metálicas. (a) Discos piezoeléctricos encapsulados en epoxi y unida a la pieza. (b) Transductores de angulares disponibles comercialmente protegidos con bloque de aluminio en U y unidos a la pieza.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

Las ondas Lamb, o ondas planas, son capaces de monitorear estructuras planas como estructuras como recubrimiento metálicos y compuestos de aeronaves, sin embargo, éstas ondas son dispersivas y pueden ser difíciles de analizar. Las ondas guiadas suelen ser generados *in situ* mediante transductores de montaje de bajo perfil sobre superficies accesibles. Cada transductor puede ser un simple disco piezoeléctrico o parche o también puede ser un dispositivo diseñado para propagaren forma de modos específicos de ondas de placa.

Los defectos pueden ser detectados y localizados *in situ* por ondas guiadas usando variedad de métodos. Cook y Bertholet (2001) usaron ondas Rayleigh para monitorear crecimiento de grietas y aplicaron simples diferenciales en el dominio del tiempo para estimar el tamaño de las mismas. Leonard et al. (2000) y Gao et al (2005) han aplicado métodos de tomografía de reconstrucción para ondas planas y han obtenido imágenes del daño. Giurgiutiu and Bao (2004) y Wilcox (2003) han aplicado técnicas de arreglos de fase para matrices compactas de ondas guiadas. Recientemente, Michaels y Michaels (2006) del Tecnológico de Georgia han

generado imágenes de daños en placas con muy escasa matrices de discos piezoeléctricos. La figura 39a muestra formas de onda diferenciadas debido al campo de sonido disperso por un agujero perforado en una placa de aluminio, y la figura. 39b muestra la imagen resultante del agujero con una matriz dispersa de solo cuatro transductores.

- *Ondas de volumen (Bulk waves)*. Se refieren a las ondas longitudinales y de corte elásticas que se propagan en masa en materiales. Estas son las ondas de ultrasonidos convencionales normalmente utilizados para los métodos del NDI, por el que el transductor se mueve de forma manual o de forma automática para barrerla trayectoria del haz de ultrasonidos a través del volumen de material a que va ser inspeccionado. Los métodos de ondas Bulk son capaces de monitorear “puntos calientes” en estructuras donde la localización de los alrededores del daño ha sido identificada a través de cualquier análisis o experiencia. Por ejemplo, Michaels et al. (2005) y Mi et al (2005) han desarrollado método de haz angular de ondas de corte para monitorear grietas creciente originadas por agujeros de ajuste. Como se muestra en la Fig.40, la onda de corte incidente se convierte en una onda progresiva que viaja por todo el agujero de fijación, y cualquier grieta bloquea esta onda y evitar que se propagación al receptor. A medida que la grieta se abre bajo la carga, este efecto se acentúa y permite estimar la gráfica del área de la grieta versus los ciclos de fatiga, Fig.40.

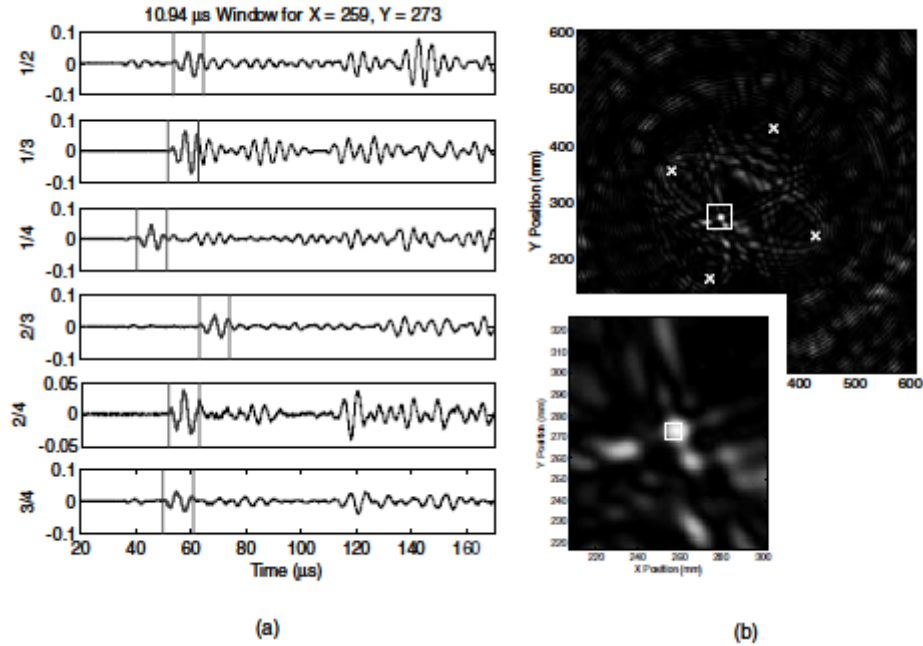


Figura 39. (a) Señales de ondas ultrasónicas Lamb dispersas de un agujero perforado en una placa de aluminio con las regiones marcadas, indicándose la dispersión del agujero. (b) La imagen de una matriz de fase resultante construida a partir de las seis señales. Los círculos sólidos son trazados en los lugares de los transductores adjuntos, el círculo abierto se centra en la ubicación del agujero real, y la inserción es una vista ampliada del agujero.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

- **Ondas Difusas.** El concepto de un campo de ondas difusas se refiere a un campo en el cual todos los modos son excitados igualmente. En palabras prácticas, difuso o como ondas difusa son aquellos en los que los ecos específicos no pueden ser identificados, y la energía esencialmente ha “llenado” la estructura en la cual se están propagando las ondas. El método de ultrasonido acústico (AU) para ensayos no destructivos se refiere a la generación y análisis de ondas difusas que son registradas por un largo periodo de tiempo y eventualmente decaen a niveles ilegibles. La ventaja del

uso de ondas difusas es que son muy fáciles de generar, y son muy sensibles a cualquier cambio estructural. La desventaja de las ondas difusas es que es muy difícil discriminar a los cambios debido a los daños relacionados a los efectos benignos como los cambios del medio ambiente y el acoplamiento del transductor. Sin embargo, la situación es algo más manejable para el monitoreo de salud estructural en comparación con el NDI, porque el carácter fijo de los transductores minimiza la variabilidad debido al acoplamiento del transductor y el posicionamiento. Trabajos recientes en el Tecnológico de Georgia (Lu y Michaels, 2005; Michaels y Michaels, 2005; Michaels et al, 2005) ha tenido éxito en discriminar los daños artificiales de especímenes de aluminio debido a los cambios medio ambiente. Uno de los discriminadores más eficaz es la coherencia local temporal, que es una medida de la similitud en la forma de dos señales como una función del tiempo. La Figura 41. muestra una gráfica de la coherencia local temporal ilustrando el progreso del daño en un espécimen de aluminio. El daño causa una pérdida distintiva de la coherencia en comparación con una forma de onda de referencia, mientras que la coherencia es mantenida a lo largo de la decadencia de la reverberación difusa como señales de cambios modestos del ambiente.

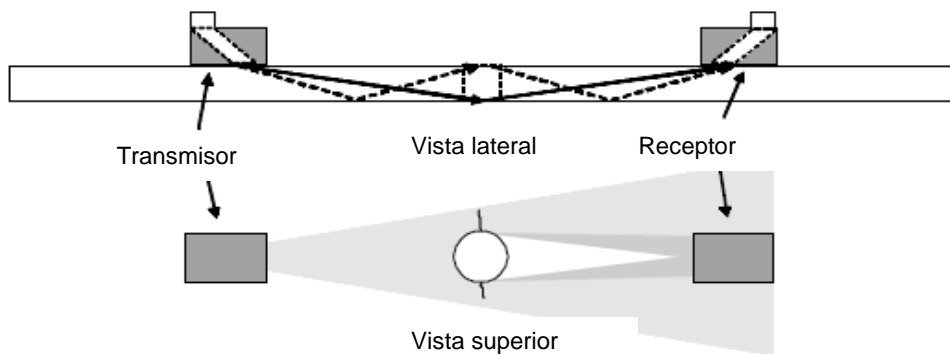


Figura 40. Método de transmisión a través de un haz angular para monitorear in situ de las grietas de fatiga de originadas por agujeros de fijación.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

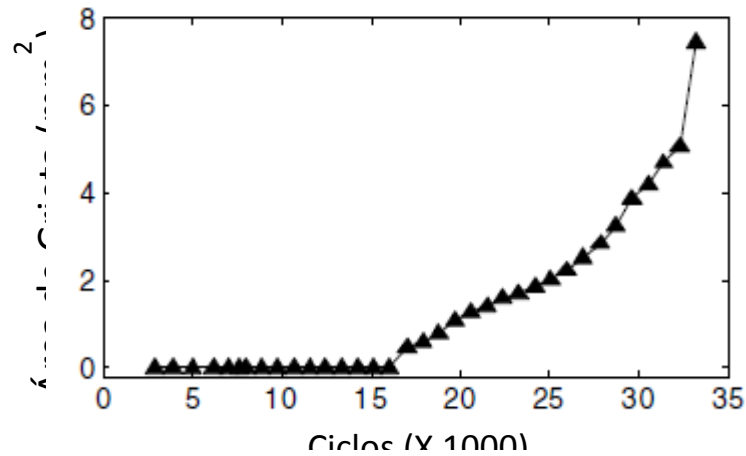


Figura 41. Gráfica del área de grieta versus los ciclos de fatiga para monitoreo de agujeros de fijación.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

- *Otros métodos de Ultrasonido.* Se están realizando investigaciones en la aplicación de otros métodos de inspección no destructiva con ultrasonidos para vigilancia de la salud estructural. La velocidad y la atenuación de ultrasonido se han utilizado en NDI para caracterizar las propiedades del material, una aplicación de control sanitario correspondiente es muy posible y se espera que sea eficaz en la detección del deterioro gradual de las propiedades del material del volumen analizado. Métodos no lineales de ultrasonidos han demostrado ser prometedores para la detección de micro-daños, sin embargo, la técnica es muy difícil de aplicar, incluso en el laboratorio, y en la aplicación in situ sería aún más difícil.

Varios métodos de detección por ultrasonido pueden ser implementados de manera sinérgica para detectar gran variedad de daños tanto local como globalmente. La figura 43 ilustra la arquitectura de un sistema completo de monitoreo de salud (PHM). Los datos actuales e históricos son analizados y se fusionan utilizando métodos basados en datos como aquellos basados en la física para obtener el mejor diagnóstico del estado estructural actual. Esta información del estado es usada luego para el pronóstico de rendimiento futuro de la estructura.

A.7.5.3 SENSORES DE RUTINA Y OPERACIONALES

Los datos operacionales y de rendimiento del sistema son monitoreados rutinariamente en todas las fábricas, operadores de servicios públicos, sistemas de transporte, etc. para procesos de control, evaluación de rendimiento, garantía de calidad y con el propósito de diagnosticar fallas. Un gran número de sistemas de detección han sido desarrollados y usados en los últimos años. La lista incluye dispositivos que tienen por objeto medir propiedades críticas como temperatura; presión; flujo, y propiedades termodinámicas, ópticas; y elementos bioquímicos, entre muchos otros. Los sensores basados en elementos clásicos de medida- inductivo, capacitivo, ultrasonido- han encontrado muchas aplicaciones. Los materiales semiconductores usualmente constituyen las bases de construcción de muchos de los sensores. Los dispositivos de campo magnético (sensores Hall) son acoplados con sensores eléctricos para monitorear propiedades magnéticas de los materiales. Los sensores de fibra óptica son tratados en una sección aparte de este capítulo. Recientemente, los sensores biomecánicos han empezado a llamar la atención, y sus principios de detección y requerimientos son descritos en la literatura técnica (Guardia, 1995). Las propiedades características de los sensores químicos con aplicación potencial en diagnósticos de falla de estructuras

incluye sensores electrolíticos sólidos y líquidos, sensores foto químicos, sensores de humedad, del efecto de campo y de dispositivos sensibles a la masa. Como ejemplo de sus principios de funcionamiento, tenga en cuenta sensores de conductividad. En estos dispositivos, la interacción del gas con el sólido (semiconductor de óxido metálico o semiconductor orgánico) causa un cambio en la conductividad. Un cambio en la resistencia también puede causar un cambio en la temperatura del material del sensor. Un amplio número de textos de referencia y publicaciones técnicas cubren los principios de operación, métodos de diseño, y aplicaciones de estos sistemas de detección. Restringiremos nuestra atención solo en una clase de sensores – sensores térmicos- para propósitos ilustrativos y como principales candidatos para el diagnóstico de falla.

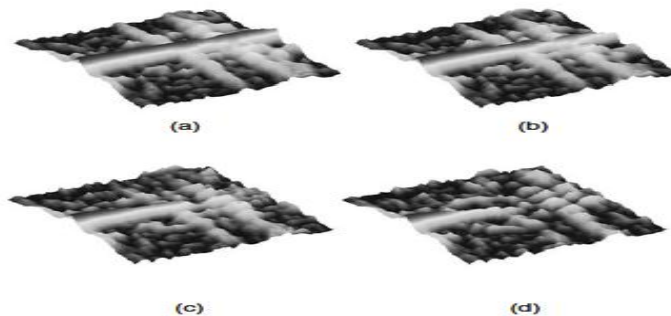


Figura 42. Coherencia local temporal de ondas difusas de un componente dañado comparado con una señal de referencia de un componente en buen estado. Las Gráficas (a) a (d) muestran la caída progresiva de la coherencia como un defecto artificial es agrandado.

Fuente: Intelligent Fault Diagnosis and Prognosis for Engineering Systems. George Vachtsevanos et al.

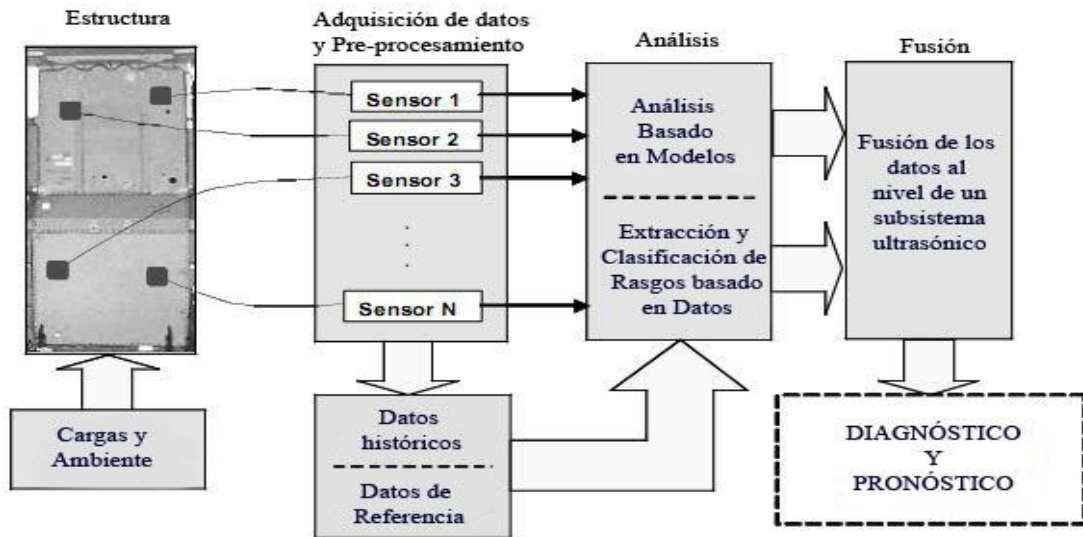


Figura 43 Arquitectura de un sistema de monitoreo de salud por ultrasonido de una estructura.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

A.7.5.3.1 Sensores térmicos / Termografía.

Las variaciones de temperatura en muchos sistemas mecánicos, eléctricos y electrónicos son un excelente indicador de condiciones de falla inminente. Las temperaturas que exceden los límites de control deben ser monitoreadas y usadas en conjunto con otras medidas para detectar y aislar fallas. Los sensores de temperatura han encontrado numerosas aplicaciones en los últimos años en áreas como la ingeniería, la medicina, monitoreo medio ambiental, etc.

Exceptuando el tiempo, la temperatura es el parámetro físico más observado en la ciencia y la ingeniería. Los sensores térmicos son usados ampliamente para medir y rastrear los niveles y variaciones de la temperatura. Las propiedades térmicas de los materiales de los organismos vivos de sólidos, líquidos y gases son

indicadores importantes del estado de estos objetos. Para el diagnóstico de fallas en sistemas, las variaciones de temperatura- puntos calientes, grandes gradientes de temperatura, y temperaturas que exceden niveles específicos- son considerados indicadores claves de condiciones inminentes de falla. Varios de los dispositivos de detección se basan en principios de transferencia de calor. Entre los más comunes está el detector de temperatura por resistencia (RTDs), cuyo principio de operación es la variación de la resistencia de un alambre o película de platino como función de la temperatura. El platino es usualmente empleado porque es estable con la temperatura y por el hecho que su variación de resistencia tiende a ser casi lineal con la temperatura. Un termómetro industrial con resistencia de platino normalmente trabaja en un rango de temperaturas entre 200 y 650°C. RTDs especiales son capaces de trabajar por encima de 850°C (www.watlow.com/products/sensors/se_rtd.cfm). Sensor Tec (www.sensortecinc.com, Figura 44) ofrece una completa línea de RTDs con rangos de temperatura desde -200 hasta 650°C. Los termómetros de resistencia están disponibles en una serie de formas (www.peaksensors.co.uk). Los más comunes son (1) alambre en una cerámica aislante (altas temperaturas a 850 ° C) y (2) alambre encapsulado en vidrio (resistente a altas vibración y ofrece mayor protección a la fina película de platino, con películas de platino en un sustrato de cerámica, producción en masa bajo costo). La figura 7.40 muestra una serie RTD industrial de platino tipo 21A de montaje por prueba directa) con un rango de operación de 195 a 482°C. (www.rdfcorp.com/products/industrial/r-21_1.shtml). En uso, un RTD normalmente está conectado para formar un brazo del arreglo puente Wheatstone, lo que permite alta precisión en la medida, particularmente con una conexión de cuatro conductores.



Figura 44. RTD disponibles en Sensortec Inc.

Fuente: Sensortec Inc.

En los transductores de muy baja resistencia ($\sim 100\Omega$) puede ser necesario en ciertas circunstancias un elemento de configuración de cuatro hilos para compensar la resistencia de los cables de conexión. Los termómetros de resistencia requieren que una pequeña corriente pase para determinar la resistencia. Esto puede causar un calentamiento espontáneo, y los límites de manufactura se deben seguir siempre junto a consideraciones de la ruta de calor en el diseño. También se deberían tomar en cuenta para evitar haya deformación en el termómetro de resistencia en su aplicación.



Figura 45. RTD industria de platino serie 21 A.

Fuente: RDF Corporation

Las principales ventajas de los RTD son la estabilidad, repetitividad y precisión. Ellos también tienen algunas desventajas, tales como el costo (el platino es caro), rango de temperatura (generalmente -200 a 650 C), y la lenta respuesta a los cambios de temperatura. Además, el dispositivo puede resultar dañado por las vibraciones y golpes.

Un termistor es una resistencia térmica sensible cuya resistencia eléctrica varía directamente con la temperatura. Los termistores son semiconductores de cerámica que presentan una tendencia del coeficiente de temperatura negativa (NTC) o positiva (PTC), y por lo general se utilizan dentro de la gama de temperaturas desde -200 hasta 1000 °C para NTC y de 60 a 180 C para PTC. Los NTC deben ser escogidos cuando se requiere un cambio continuo de resistencia en un amplio rango de temperatura, y los PTC cuando se requiere un cambio drástico en la resistencia en una temperatura específica o nivel de corriente. Los termistores sellados herméticamente tiene una estabilidad a largo plazo excelente con una amplia gama de valores de resistencia relativamente alta y un rápido tiempo de respuesta, están disponibles en pequeños tamaños (0.005 pulgadas de diámetro). La Figura 46. muestra termistores ISO-Honeywell de esfera y sonda de vidrio diseñados como sensores de precisión, intercambiables y son necesarios para un control y medición de temperatura preciso. La alta confiabilidad de los productos consiste en que se une un par de termistores herméticamente sellados y se pueden seleccionar las tolerancias limitándose solo por la capacidad del sistema que se esté probando, (<http://content.honeywell.com/sensing/prodinfo/thermistors/isocurves.asp>).

Coeficientes de temperatura positivas de 0,7 por ciento por grado centígrado son típicos, con constantes de disipación del orden de 2,5 mW/°C. La figura 27 muestra termistores de US Sensor (www.ussensor.com/pgthermistors.html). A

nivel militar los montajes de termistores en superficies son manufacturados por Quality Thermistor, Inc. (www.thermistor.com/ptctherms.cfm).

Las termocuplas son por mucho los sensores más usados en la industria. Son muy resistentes y pueden ser usadas para temperaturas bajo cero hasta por encima de los 4000°F (2000°C). La medición de temperatura con termocuplas es basado en el efecto Seebeck, o efecto termoeléctrico en el cual dos materiales distintos son unidos, y si las dos uniones formadas están a diferentes temperaturas, una corriente fluye por los materiales. Cuando este circuito es abierto una fuerza electromotriz aparece (fem), la cual es proporcional a la corriente de Seebeck observada. Las termocuplas responden a diferencias de temperatura en vez de temperaturas absolutas; su salida es una función de la temperatura en los dos puntos, también llamados unión fría y caliente. Las termocuplas más comunes de pares de alambres son listadas en tablas estándar especificando la salida en milivoltios versus temperatura con una temperatura de referencia de unión de 0°C. Los rangos sensibles están desde los 5 a los 80µV/°C. Watlow suministra termocuplas con un amplio rango de temperaturas desde -200 a 2200°C (-328 a 4000°F) (www.watlow.com/products/sensors/se_therm.cfm).



Figura 46. ISO-termistores de Sonda y esferas de vidrio

Fuente: Honeywell



Figura 47. Termistores de plomo estándar recubiertos de epoxy.

Fuente: US Sensor

Algunos materiales típicos para termocuplas son se ilustran en la Tabla 3.

Tabla 3. Nomenclatura (ANSI), materiales para termocuplas y sus respectivos rangos de operación.

TERMOCUPLAS			
Símbolo	Tipo de Metales	Temp. de utilizacion °C	F.e.m. a 100° C

	+	-		
T	Cobre	Cobre-Níquel T	-200°C +350°C	4.277
J	Hierro	Cobre-Níquel J	-40°C +750°C	5.268
E	Níquel-Cromo	Cobre-Níquel E	-150°C +800°C	6.317
K	Níquel-Cromo	Níquel-Aleado	-150°C +1100°C	4.095
N	Níquel-Silicio-Cromo	Níquel-Silicio	-150°C +1100°C	2.774
R	Platino 13% Rodio	Platino	0°C +1600°C	0.647
S	Platino 10% Rodio	Platino	0°C +1550	0.645
B	Pt 30% Rh	Pt 6% Rh	-600°C +1550°C	0.033

Fuente: Tecnoindusil S.A

Es bueno saber que los defectos y grietas en estructuras metálicas generan calor cuando se someten a vibraciones mecánicas. El aumento de la histéresis del defecto genera más actividad térmica que la región circundante en buen estado. Las ondas térmicas lanzadas desde el defecto se propagan a la superficie, y el patrón térmico resultante puede revelar la presencia del defecto cuando se sincronizan con una señal de excitación adecuada. Esta técnica de evaluación no destructiva (END) es usada para propósitos de inspección y usa un método llamado análisis de tensión termoelástico (ATT) y termografía . La técnica ha sido satisfactoriamente demostrada en grietas del subsuelo en una estructura de una sola capa de hasta 20 mm de espesor por El-Hajjar y Haj-Ali (2004) en el Instituto

de Tecnología de Georgia. La Figura 48. Proporciona imágenes infrarrojas de la progresión del daño de la fatiga alrededor de un agujero del remache. Es evidente a partir de las imágenes que el remache impide al calor que se transmite directamente a través de él.

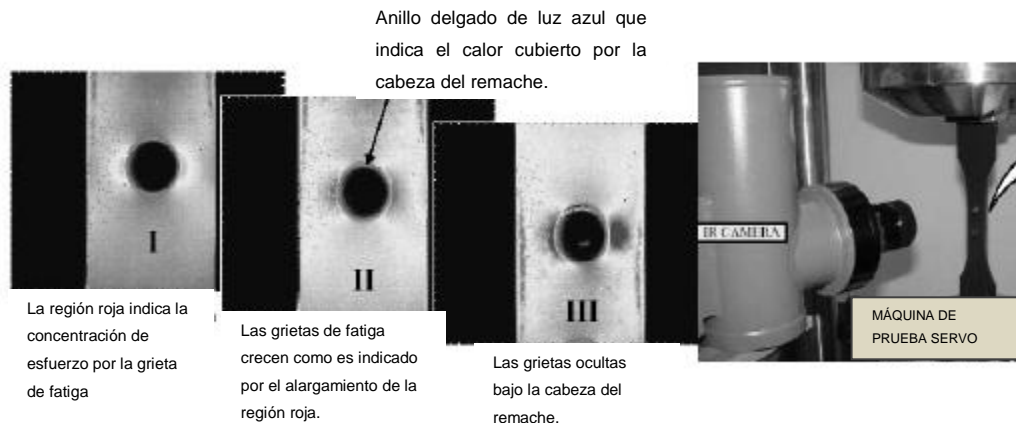


Figura 48. Imágenes infrarrojas del progreso del daño por fatiga alrededor del agujero de remache en un espécimen de prueba 7050-T7451

Fuente: Intelligent Fault Diagnosis and Prognosis for Engineering Systems. George Vachtsevanos et al.

Sin embargo, el calor generado por la región del daño (concentración de esfuerzos) es capturado vívidamente por el método (ATT).

A.7.5.4 MEDICIONES ELÉCTRICAS

Los sistemas electromecánicos, eléctricos y electrónicos constituyen la mayoría de los componentes de nuestras máquinas industriales. Estos son los elementos dominantes en cualquier área como sistemas de transporte, herramientas, instrumentación biomédica, comunicaciones, computación, etc. Los controladores

eléctricos y los sistemas y componentes eléctrico/electrónicos están atrayendo el interés de la comunidad de diagnóstico, y los sensores de vigilancia de su desempeño se están convirtiendo en la actividad central del diseñador. Libros universitarios y cursos de medición presentan un excelente tratamiento de los sistemas de sensores clásicos destinado a seguir cantidades como voltaje, corriente, potencia, par motor y sus derivados –contenido en armónicos, por ejemplo- como indicadores claves de la disminución del rendimiento. Por esta razón, en esta sección trataremos las mediciones eléctricas.

Un número de sistemas de sensores ha sido desarrollado y aplicado en los últimos años en un intento de averiguar *in situ* sistemas y componentes críticos para el diagnóstico y pronóstico de fallas. Principios de transducción basados en características de la respuesta de las corrientes de Foucault, seguimiento de señales ópticas y de infrarrojos, microondas, y otros han sido investigados. Vamos a definir los principios básicos y ámbitos posibles de aplicación de algunos de esos dispositivos.

A.7.5.4.1 Sondas de proximidad por corrientes de Foucault.

Las respuestas características de las corrientes de Foucault en un medio conductor son monitoreadas para detectar los cambios en su comportamiento debido a las anomalías de material, grietas, juntas o desplazamientos, etc. Las sondas de proximidad de Foucault es una tecnología madura que se ha utilizado para la protección y el manejo de la maquinaria de rotación. Son usadas comúnmente en turbo maquinaria de alta velocidad para observar el movimiento del eje relativo directamente, esto es, dentro de las separaciones de los

rodamientos, en la interfaz de la película de fluido. Se puede encontrar información de sensores de proximidad por corrientes de Foucault en http://proximity-sensors.globalspec.com/SpecSearch/Suppliers/Sensors_Transducers_Detectors/Proximity_Presence_Sensing/Eddy_Current_Proximity_Sensors. Entre los fabricantes de estos están: Panasonic Electric Works (antes Aromat; <http://pewa.panasonic.com/acsd/>), Lion Precision (www.lionprecision.com/), and Metrix Instrument Co. (www.metrix1.com/).

Estos sensores son usados para detectar movimientos laterales del eje por desalineación absoluta o relativa, defectos de rodamientos, grietas y otras anomalías de los materiales mediante la generación de gráficos de la respuesta de tres dispositivos ubicados estratégicamente en la estructura bajo prueba.

A.7.5.4.2 Sistemas de sensores Micro-electromecánicos (MEMS).

De interés son también los sensores que se pueden producir económicamente, por separado o en una matriz, mientras que mantienen un alto nivel de confiabilidad operacional. Los MEMS y los sensores basados en tecnologías de fibra óptica son populares debido a su tamaño, costo y capacidad para integrar múltiples transductores en un solo dispositivo. Los MEMS en silicona u otros materiales son fabricados en un proceso por lotes con el potencial de integración con la electrónica, lo que facilita el procesamiento de señales a bordo y otras funciones “inteligentes”. Varios transductores MEMS y sistemas de sensores han sido manufacturados en el laboratorio o están disponibles comercialmente para monitoreo de parámetros críticos como temperatura, presión, aceleración, etc.

Analog Devices, Inc. (ADI), produce un MEMS de tres ejes integrados (iMEMS) de tecnología de proceso de señales de movimiento que permite una desviación muy pequeña. Honeywell fabrica sensores MEMS basados en giroscopio inercial como los HG1900, los cuales usan giroscopio MEMS y acelerómetro RBA500. Esos sensores requieren una fuente de entrada de bajo voltaje y dan una variación de ángulo y de velocidad como salida. (<http://content.honeywell.com/dses/products/tactical/hg1900.htm>).

A.7.5.5 SENSORES DE FIBRA ÓPTICA.

Los sensores de fibra óptica han incursionado en las telecomunicaciones y otros sectores de las tecnologías de punta en los últimos años. Han encontrado utilidad en el campo de los sensores debido a su geometría compacta y flexible, potencial de fabricación dentro de matrices de dispositivos, fabricación por lotes, etc. Los sensores de fibra óptica han sido diseñados para medir deformación, temperatura, desplazamiento, concentración química y aceleración, además de otras propiedades de los materiales y del medio ambiente. Su principal ventaja incluye que son de tamaños pequeños, livianos, inmunes al electromagnetismo y la interferencia de radio frecuencia (EMI/RFI), resistencia a baja y alta temperatura, rápida respuesta, alta sensibilidad, y bajo costo. La tecnología de fibra óptica son basadas en interferometría extrínsecas de Fabry-Perot (EFPI), cambio químico en el revestimiento de la fibra, cambios en la señal óptica debido a tensión y deformación de la fibra, etc. Algunos de los fabricantes de sensores de fibra óptica son: Luna Innovations (www.lunainnovations.com/r&d/fiber_sensing.asp), KEYENCE (http://world.keyence.com/productos/fiber_optic_sensors/fs.html), and Techno-Sciences (www.technosci.com/capabilities/fiberoptic.php).

Luna Innovations ha desarrollado sensores de temperatura de fibra óptica basados en EFPI. La física de base que regulan el funcionamiento de estos sensores los hace relativamente tolerantes o inmunes a los efectos de la alta temperatura y alta-EMI, y ambientes altamente corrosivos. La física de base también lleva a un sensor de temperatura muy estable, preciso y lineal en un amplio rango de temperaturas. Estos sensores también son muy pequeños y por lo tanto ideal para aplicaciones donde la interferencia de medición del espacio restringido o mínimo es una consideración. El tamaño también conduce a un tiempo de respuesta muy pequeño en comparación con la medición de la temperatura por otras técnicas (Fig. 49).

A.7.5.6 UBICACIÓN DEL SENSOR

La ubicación del sensor para monitorear indicadores de falla o firmas con precisión y firmeza en un sistema crítico militar o industrial constituye una función esencial del proceso de diseño de una estructura de diagnóstico y pronóstico de falla. Los sensores típicamente son ubicados con base a la experiencia tomado en consideración el número estimado de sensores necesarios, tipo de medida que se va a tomar, acceso físico para la ubicación del sensor, etc. Las directrices de localización de sensores están disponibles para algunos sistemas militares y aeroespaciales (Kammer, 1991). Unas cuantas técnicas analíticas han sugerido en los últimos años como determinar el número, tipo y localización de los sensores. La mayoría de los frutos de estos trabajos son de la industria química, donde la necesidad de minimizar el número de sensores usados para un proceso en particular es obvio debido al tamaño, complejidad y número de parámetros que son seguidos en los complejos procesos químicos (Alonso et al., 2004; Chmielewski et al., 2002).

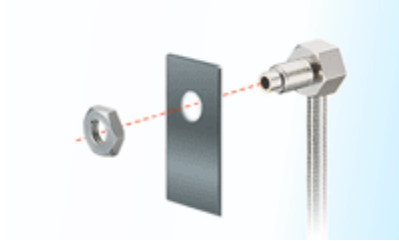


Figura 49. Un típico sensor de temperatura de fibra óptica.

Fuente: KEYENCE

El conjunto de sensores instalados para el monitoreo de fallas debe, necesariamente, poseer atributos de durabilidad, robustez, exactitud, y sensibilidad mientras responde a un rango dinámico de las variables o parámetros monitoreados. La vida útil del sensor debe ser mucho más larga que la de los componentes o sub-sistemas que tiene por objeto medir. Optimizando la tarea de seleccionar un conjunto de sensor apropiado e identificar la mejor localización para el dispositivo físico se contribuye al éxito técnico-económico y a la viabilidad del sistema monitoreo. Tradicionalmente, los sensores se colocan para cumplir con los objetivos de control y supervisión del rendimiento (Anthony y Belinda, 2000; Al Shehabi y Newman, 2001; Giraud y Jouvencel, 1995; Che y Li. 2002). Es útil tomar ventaja de estos sensores en un esquema de seguimiento diagnóstico de fallo, ya que pueden proporcionar información útil relativa a las conductas fallo de las variables críticas del sistema. Las firmas de falla, sin embargo, pueden requerir sensores específicos adicionales; dispositivos típicos de esta clase son los acelerómetros para monitoreo de vibraciones, sensores de proximidad, sensores de emisiones acústicas, transductores ultrasónicos, etc. Más recientemente, la investigación sobre la ubicación de los sensores se ha enfocado en dos niveles diferentes: El nivel componente y el nivel sistema. En el nivel

componente, se ha intentado reportar respecto a la colocación en el rango del componente, por ejemplo, un rodamiento o un objeto visto en tres dimensiones (Anthony y Belinda, 2000; Naimimohasses et al., 1995). Para sistemas complejos y a gran escala constituidos de múltiples componentes o sub-sistemas, una falla puede propagarse a través de varias de sus partes. Con un gran número de posibles localizaciones de sensores, seleccionar la mejor, así como el número y tipos de sensores, plantea un desafío que debe ser abordado al nivel del sistema.

Un posible enfoque de la teoría de los sistemas ve el problema de una ubicación óptima del sensor desde diversas perspectivas, pero el marco teórico subyacente se basa en técnicas de estimación, análisis de sensibilidad, los conceptos de observación, y herramientas de optimización. Básicamente estamos tratando de estimar a través de mediciones el valor de una variable crítica, maximizar la sensibilidad del sensor a la señal de falla mientras se minimizan las falsas alarmas, verificar que las variables de son observables y finalmente, optimizar el número, tipo y localización de los sensores seleccionados. Como una ilustración de estos métodos pueden abordar el problema de ubicación óptima, resumiremos a continuación un enfoque de sistema desarrollado por el Instituto de Tecnología de Georgia (Zhang, 2005). Otros enfoques pueden explotar diferentes técnicas, buscando respuestas a las mismas cuestiones planteadas en esta sección.

El concepto fundamental de este enfoque es una métrica de detección de falla que expresa la capacidad del sensor para detectar un fallo. Una técnica basada en gráficos, llamada *Modelo Cuantificado Directo (QMD, por sus siglas en inglés)*, es usado para modelar la propagación de la falla de un componente o subsistema a el siguiente. Una figura de mérito está orientada finalmente a maximizar la capacidad de detección del fallo al mismo tiempo que reduce al mínimo el número requerido de sensores y logra la ubicación óptima de los mismos. La arquitectura

de selección / localización es mostrada en la Fig.50. Implica varios módulos funcionales: análisis de requerimientos, modelación del sistema, una figura de mérito (FOM), optimización y evaluación del desempeño.

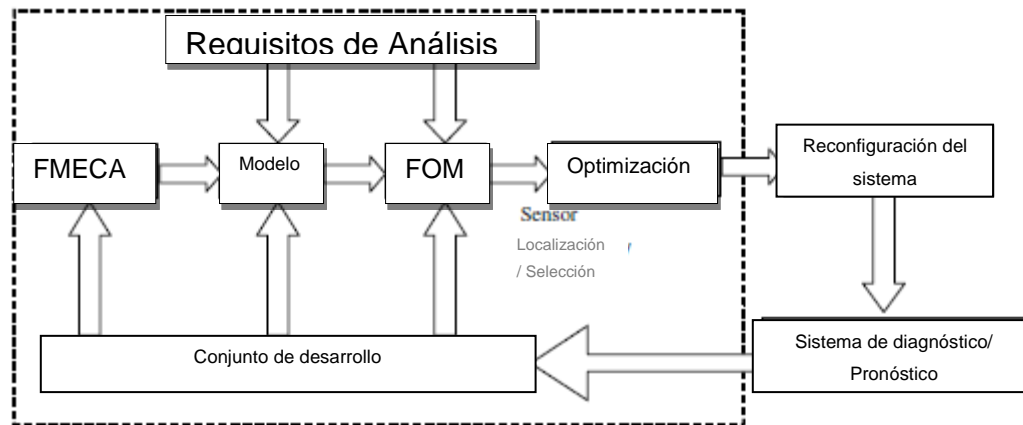


Figura 50. Arquitectura de selección/ Clasificación

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

A.7.5.7 REDES DE SENSORES INALÁMBRICOS

Los sensores inalámbricos (WSNs, por sus siglas en inglés) son importantes en aplicaciones donde no puedan usarse cables debido a su costo, precio y accesibilidad. Un adecuado diseño de estos sensores puede ser instalado y calibrado rápidamente y puede estar en funcionamiento en muy poco tiempo. Los sistemas diseñados en la Universidad de Texas en Arlington (UTA) Automatización y Robótica del Instituto de Investigación (ARRI) pueden ser instalados en maquinaria espacial en pocas horas.

La figura 51. Muestra la complejidad de los WSNs, los cuales generalmente consisten en una red de adquisición de datos y una red de distribución de datos monitoreado y controlado por un centro de control. El gran número de tecnologías disponibles hace incluso que la selección de componentes sea difícil, y mucho más el diseño en general de un constante, confiable y robusto sistema. El estudio de los WSNs es un reto en el que se requiere un enorme cantidad de conocimiento de una gran variedad de disciplinas.

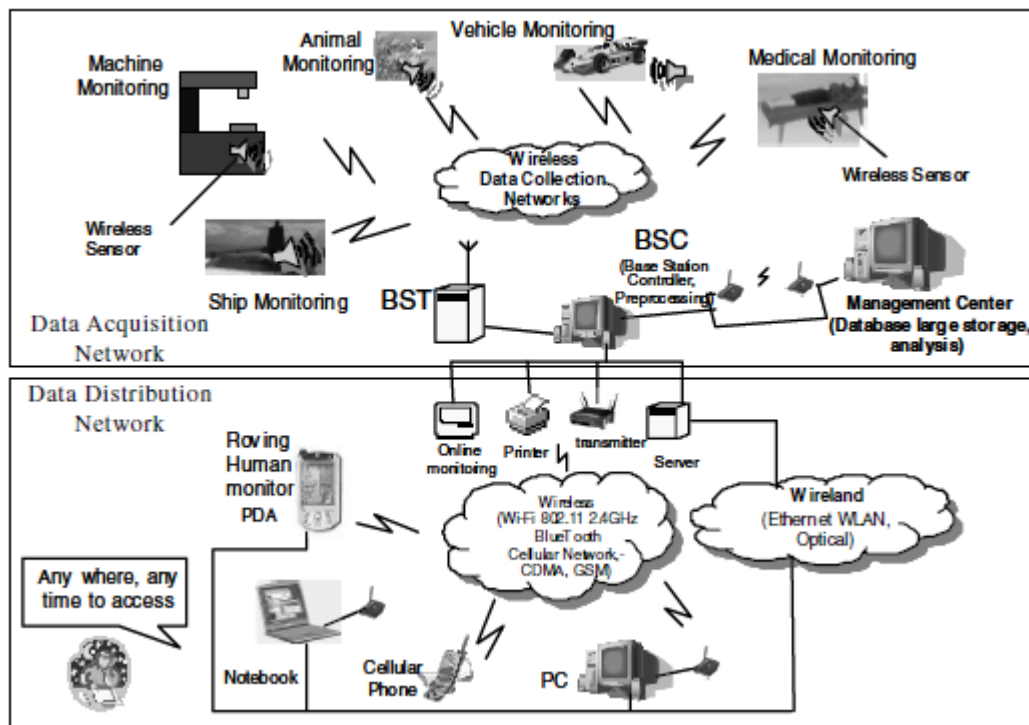


Figura 51. Red de sensores inalámbricos

Fuente: Intelligent Fault Diagnosis and Prognosis for Engineering Systems. George Vachtsevanos et al.

A.7.5.7.1 Topología de la Red

El hecho básico en las redes de comunicación es la transmisión de mensajes para lograr un rendimiento de los mensajes prescritos (cantidad de servicios) y la calidad del servicio (QoS). QoS puede especificarse en términos del retraso del mensaje, fechas de vencimiento, tasas de error de bit, pérdida de paquetes, costo de la transmisión, poder de transmisión, etc. Dependiendo de la calidad del servicio, el entorno de instalación, las consideraciones económicas y la aplicación pueden utilizarse varias topologías de redes básicas.

Una red de comunicación está compuesta por nodos, cada uno de los cuales tienen un poder de computación y puede transmitir y recibir mensajes a través de enlaces de comunicación cableados o inalámbricos. Las topologías básicas se muestran en la Fig. 52. entre los que vemos conexión total, malla, estrella, anillo, árbol y bus. Una sola red puede estar compuesta de varias subredes interconectadas de diferentes topologías. Las redes además se clasifican como redes de área local (LANs, por sus siglas en inglés), por ejemplo, dentro de un edificio, o redes de área grande (WANs, por sus siglas en inglés) por ejemplo entre edificios.

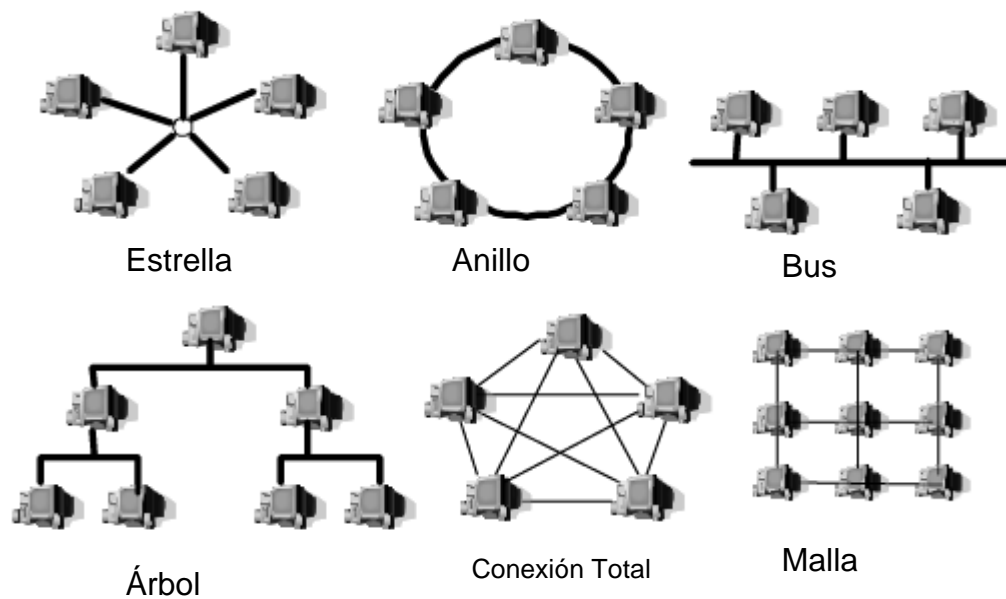


Figura 52. Topologías de red básicas.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

A.7.5.7.2 Sistemas de Sensores Disponibles Comercialmente.

Muchos proveedores ahora producen muchos tipos de sensores que son comercialmente adecuados para las aplicaciones de red inalámbrica. Véase, por ejemplo, los sitios Web de SUNX Sensores, Schaevitz, Keyence, Turck, Pepperl& Fuchs, National Instruments, UE Systems (ultrasonidos), Leake(IR), y CSI(vibración). La tabla 3.1 muestra cuales principios físicos pueden usarse para varias cantidades medibles. Los sensores MEMS están disponibles para la mayoría de estas medidas.

Muchos nodos de comunicación inalámbrica están disponibles comercialmente, incluyendo Lynx Technologies y varios kits bluetooth, como los dispositivos Casira de Cambridge Silicon Radio (CSR).

A.7.6 SISTEMAS GERENCIALES DE PROCESAMIENTO DE SEÑALES Y BASE DE DATOS

A.7.6.1 INTRODUCCIÓN

Durante los últimos años un gran reto ha confrontado a practicantes e investigadores de la gerencia tanto del mantenimiento basado en la condición, CBM¹⁵, como el pronóstico y estado de salud del sistema, PHM¹⁶: ¿cómo son monitoreados o medidos los cambios funcionales (síntomas) del sistema en términos de los estados medibles del sistema (salidas)? ¿Cómo relacionamos los mecanismos de falla con la física fundamental de un complejo sistema dinámico? Es bien sabido que los modos de falla inducen cambios en la energía (potencia) del sistema, su entropía, espectro de potencia, magnitud de la señal, comportamiento caótico, etc. ¿Dónde está la información? Se encuentra posiblemente en el dominio del tiempo, el dominio de la frecuencia o en el dominio del wavelet. ¿Cómo extraemos información útil de un base de datos masiva sin procesar?

Los datos son raramente útiles o manipulables en su forma primaria. Considere, por ejemplo, datos de vibración muestreados a 100 kHz. Tal cantidad de datos son

¹⁵De las siglas en inglés Condition-Based Maintenance

¹⁶De las siglas en inglés Prognostics and Health Management

inmanejables a menos que sean procesados y reducidos a una forma en que puedan ser manipulados fácilmente por algoritmos de diagnóstico y pronóstico de fallas. Nuestro objetivo del procesamiento primario de los datos de sensores es que queremos que los datos entreguen o “confiesen” solo la información verdadera o correcta.

Nos enfocaremos en tecnologías que nos asistan en la selección y extracción de las mejores “características” o indicadores de la condición que contengan tanta información como sea posible. La meta eventual es emplear estas características de una manera óptima para poder maximizar la detección y aislamiento de fallas de forma precisa mientras se minimizan las falsas alarmas. Así, los datos recolectados por aparatos de detección deben ser procesados apropiadamente para que la información útil pueda ser extraída de ellos lo que constituye una versión reducida de los datos originales pero preserva en la mayor forma posible las características o indicadores de falla que muestran los eventos de fallas que buscamos detectar, aislar, y predecir. Tales datos deben ser pre-procesados, es decir, filtrados, comprimidos, correlacionados, etc., para remover artefactos y reducir los niveles de ruido y el volumen de los datos a ser procesados subsecuentemente. Además, el sensor que provee los datos debe ser validado; es decir, los mismos sensores se encuentran sujetos a condiciones de falla. Una vez el módulo de pre-proceso confirma que los datos del sensor son “limpios” y formateados apropiadamente, las características o rasgos de las condiciones normales o de falla deben ser extraídas. Este es el paso más significativo en la arquitectura CBM/PHM, cuya salida ajustará el estado para un diagnóstico preciso y a tiempo de los modos de falla. El vector de extracción de características servirá como una de las entradas esenciales para los algoritmos de diagnóstico de fallas.

A.7.6.2 PROCESAMIENTO DE SEÑALES EN CBM/PHM

Las tecnologías de CBM/PHM han sido desarrolladas y aplicadas a una gran variedad de máquinas, sistemas, y procesos en los sectores del transporte, la industria y la manufactura. El equipo rotatorio ha recibido especial atención debido a su régimen de operación crítico y los modos frecuentes de fallas y la disponibilidad de mediciones (vibración, temperatura, etc.....) las cuales permiten la detección y aislamiento de fallas incipientes. Otra categoría de importancia son los componentes estructurales sujetos a esfuerzos, superficies corroídas, puntos calientes en maquinaria y bajo estrés térmico, etc.

Las mediciones tomadas por el sensor que afectan el diagnóstico de fallas para ciertos sistemas críticos, se pueden ver desde dos diferentes perspectivas: mediciones estáticas o relacionadas al proceso tales como temperatura, velocidad, posición, presión y rata de flujo y aquellas que son caracterizadas por su gran ancho de banda (que contienen componentes de alta frecuencia). En la última categoría listamos las mediciones ultrasónicas, de medición, vibraciones, grabaciones acústicas y corriente o voltaje alternante.

El procesamiento de los datos del sensor con la intención de extraer características de fallas o indicadores de condición en cada una de estas dos categorías puede requerir diferentes herramientas de procesamiento. Por ejemplo, las mediciones de temperatura pueden ser procesadas en el dominio del tiempo, y características tales como picos de amplitud, energía, y rata de cambio pueden ser indicadores apropiados de la condición o salud del equipo bajo prueba. Análisis en la frecuencia o el dominio del wavelet pueden ser más ajustables para vibración o señales ultrasónicas.

El procesamiento de señales generalmente envuelve dos pasos. El primero es generalmente llamado pre-procesamiento de la señal y su objetivo es mejorar la señal característica que eventualmente puede facilitar la extracción eficiente de información útil, que es, las indicaciones de la condición de un componente o subsistema fallido. Las herramientas en esta categoría incluyen filtrado, amplificación, como presión de datos, validación de datos y eliminación de ruido y generalmente son dirigidos para mejorar la relación señal-ruido. El segundo paso pretende extraer las características o indicadores de condición de los datos pre-procesados que son característicos de una avería o falla incipiente.

A.7.6.3 PREPROCESAMIENTO DE LA SEÑAL

A.7.6.3.1 Acondicionamiento de la señal

El acondicionamiento de señal para mejorar la tasa de señal-a-ruido y extraer información útil de un primario data puede ser realizado con hardware y/o software.

Consideraremos ejemplos sencillos de ambos métodos. Las señales que provienen de los sensores pueden ser muy ruidosas, de baja amplitud, parciales, y dependientes de parámetros secundarios tales como la temperatura. Además, no siempre es posible medir la cantidad de interés sino solo una cantidad relacionada. Por ello, el acondicionamiento generalmente es requerido. El acondicionamiento es realizado utilizando circuitos electrónicos, la cual puede ser construida utilizando técnicas de fabricación como la integración a gran escala estándar, VLSI¹⁷, *in situ* con sensores de sistemas micro-electromecánicos (MEMS).

Un problema real con los sensores de los MEMS es la indeseable sensibilidad a cantidades secundarias como la temperatura. La compensación de temperatura a menudo puede ser construida directamente en los circuitos de los sensores de los MEMS o añadidos durante la etapa de acondicionamiento de señal.

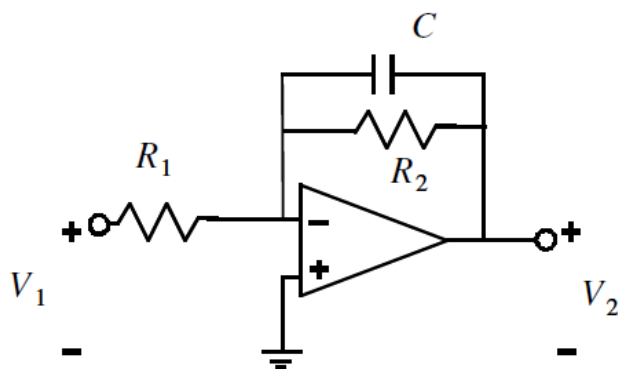


Figura 53. Filtro análogo pasa-baja

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

¹⁷ Por sus siglas en ingles, standard very large scale integration

Una técnica básica para mejorar relación señal-ruido (SNR) es filtrado pasa-baja, ya que el ruido generalmente domina las señales deseables a altas frecuencias.

En la Fig. 53. Se muestra un filtro análogo pasa-baja (LPF) que también amplifica, construido de un amplificador operacional. Tales aparatos son fabricados fácilmente utilizando técnicas semiconductor VLSI. El tiempo constante de este circuito es $\tau = R_2 C$. La función de transferencia de este filtro es $H(s) = ka/(s + a)$, con 3 dB frecuencia aislada dada por $a = 1/\tau$ rad, y ganancia dada por $k = R_2/R_1$. Aquí, s es la variable transformada de Laplace. La frecuencia aislada debería ser escogida para que sea mayor que la frecuencia de señal más alta útil del sensor.

De forma alternativa, se puede utilizar un LPF digital implementado en un computador después del muestreo. Una función de transferencia digital LPF y la ecuación de diferencia asociada para la implementación está dada por:

$$\hat{S}_k = K \frac{z+1}{z-\alpha} S_k \quad (7.1)$$

Y

$$\hat{S}_{k+1} = \alpha S_k + K(S_{k+1} + S_k) \quad (7.2)$$

Respectivamente. Aquí, z es la z variable transformada tratada como una unidad retrasada en el dominio del tiempo, S_k es la señal medida, y \hat{S}_k es la variable filtrada con menor ruido. Los parámetros del filtro son seleccionados en términos de la frecuencia aislada deseada y el periodo de muestreo. A menudo ocurre que se puede medir la variable S_k (por ejemplo, posición) pero se necesita conocer su tasa de cambio v_k (por ejemplo, velocidad). Debido a la presencia de ruido, no se puede simplemente tomar la diferencia entre valores sucesivos de S_k como de velocidad.

Una estimación de velocidad filtrada dada por $v_{k+1} = \alpha v_k + K(S_{k+1} - S_k)$ ambos filtros fuera de ruidos y da una velocidad estimada baja.

Circuitos de amplificadores operacionales diseñados especialmente son útiles para acondicionamiento de señales en general. Los amplificadores de instrumentación proveen entrada diferencial y modo común de rechazo, la impedancia entre sensores y aparatos de procesamiento, calibración, etc. Los amplificadores SLEEPMODE consumen mínima potencia mientras están inactivos y se activan automáticamente cuando la señal del sensor excede el umbral prescrito.

A.7.6.4 ELIMINACIÓN DE RUIDO Y EXTRACCION DE SEÑALES DEBILES

El pre-procesamiento de señales es un paso importante hacia el mejoramiento de la confiabilidad de los datos y así mejorar la precisión de los métodos de análisis de señales mejorando la relación señal-ruido (SNR). Los métodos tradicionales de eliminación de ruido requieren cierta información y poseen varias asunciones con respecto a las señales que deben ser separadas del ruido. La teoría wavelet ha sido aplicada exitosamente como una herramienta de eliminación de ruido, y las implementaciones en hardware y software basados en wavelet están disponibles.

Cuando se aplica wavelets para mejorar el SNR, es suficiente conocer solamente a qué clase de familia pertenece la señal de la tarjeta en lugar de la clase exacta. Una transformada wavelet ortogonal puede comprimir la “energía” de una señal en muy pocos grandes coeficientes, dejando la energía del ruido blanco dispersa a través de la transformada como coeficientes pequeños.

En los métodos tradicionales de eliminación de ruidos basados en filtros, la frecuencia de los componentes fuera de cierto rango a menudo son ajustados a cero, resultando en pérdida de información debida a impulsos en la señal que abarca un amplio rango de frecuencia. Los métodos de eliminación de ruido basados en wavelet solo remueven pequeños coeficientes, y así la información impulso es retenida porque generalmente es expresada como grandes coeficientes en la representación transformada.

A.7.6.5 COMPRESION DE LA SEÑAL DE VIBRACION

Con el creciente número de sensores de diagnóstico y altas tasas de muestreo, la cantidad de datos adquirida de sistemas industriales tiende a ser voluminosa y, en la mayoría de los casos, difícil de manejar. Así, para sistemas de diagnósticos de fallas, especialmente aquellos implementados en línea o técnicas basadas en internet remoto, la compresión de datos es una de las aplicaciones más exitosas de wavelets, incluyendo tanto la señal unidimensional como la compresión bidimensional. Las tareas de compresión y descompresión de datos

generalmente posee cinco grandes pasos: transformada, establecimiento de umbrales o límites, cuantificación/ codificación, decodificación y reconstrucción. La eficiencia de compresión es determinada por dos medidas: la tasa de compresión, definida como el ratio del número de bits de los datos originales con los números de bits de los datos comprimidos y el error cuadrado normalizado, dado por:

$$MSE(x) = \frac{100}{N\sigma_x^2} \sum_{i=1}^N (x_i - \hat{x}_i)^2 \quad (7.3)$$

Donde σ_x es la desviación estándar de $x(t)$, y N es el numero de simple points en la señal. La compresión de señales de vibración mecánica para obtener relaciones de compresión hasta de 10 o 20, compresión de las señales de vibración rodamientos y compresión del espectro de vibración de una caja de engranajes, para almacenamiento, transmisión, y selección de características de falla para el monitoreo de la condición son algunos ejemplos de las técnicas de compresión basadas en wavelet.

La compresión wavelet ha mostrado ser más efectiva en señales no estacionarias.

Un wavelet más compacto, y por ello una función wavelet que es menos suave, es más apropiada para señales no estacionarias e irregulares. De forma similar a los wavelets utilizados para eliminación de ruido, consiguen un mejor rendimiento en la compresión porque no descartan las singularidades en la señal, como otras técnicas basadas en la transformación de la frecuencia.

A.7.6.6 PROMEDIO DE TIEMPO SINCRONO (TSA)

El acondicionamiento de una señal de vibración proveniente de un equipo rotatorio tal como una caja de engranajes, rango de corrección de señal, basado en la

unidad de adquisición de datos, y los amplificadores usados, y la remoción del valor principal mediante promediado síncrono de tiempo y filtrado. Se asume que un pulso de la señal sincronizada a la rotación del engranaje indicando el inicio de una revolución individual se encuentra disponible. El objetivo de la técnica TSA es mejorar las frecuencias de vibración, las cuales son múltiples de la frecuencia del eje, el cual en muchos casos es principalmente relacionado con el enmallado del diente del engranaje después de promediar otros componentes tales como vibraciones aleatorias y perturbaciones externas. Los datos promediados resultantes muestran las vibraciones características de un engranaje en el dominio del tiempo en una revolución completa y las diferencias producidas por el diente individual del engranaje. Ya que la velocidad rotacional de una transmisión típicamente varía ligeramente durante la operación normal, los números de muestras por revolución son diferentes para una frecuencia de muestreo dada. La interpolación es requerida para hacer que el número de muestras por revolución el mismo antes de poder ensamblar el promedio. La interpolación transforma la señal de vibración del dominio del tiempo al dominio angular y redefine la frecuencia de muestreos para ser una función de la posición angular en vez de el tiempo.

Si tenemos que el segmento de datos $x_m(n)$, para $n = 0, 1, \dots, N - 1$, se define como los datos de vibración de la revolución m de un total de M revoluciones después de la interpolación, y $\bar{x}(n)$, para $n = 0, 1, \dots, N - 1$, como los datos del tiempo promediado. La transformada discreta de Fourier (DFT) de los datos de tiempo promedio está dada por

$$X(k) = \sum_{n=0}^{N-1} \bar{x}(n) e^{-j2\pi(kn/N)} \quad k = 0, 1, \dots, N - 1 \quad (7.4)$$

$$= \frac{1}{M} \sum_{n=0}^{N-1} \sum_{m=1}^M x_m(n) e^{-j2\pi(kn/N)} \quad (7.5)$$

Una nueva técnica de pre-proceso incluye tomar la DFT de cada revolución de los datos de vibración (no requiere interpolación) y luego ensamblar promediando los

números complejos (incluyendo la información de amplitud y fase) de los correspondientes índices. El razonamiento detrás de esta técnica es la siguiente:

La resolución de frecuencia del DFT está dado por $\Delta f = \frac{1}{T_{eje}} = f_{eje}$, donde T_{eje} es el tiempo de duración de cada revolución individual y puede variar ligeramente de una revolución a otra. Si el periodo de rotación del eje del engranaje es más largo, el volumen de datos muestreados en una revolución es mayor. La resolución de la frecuencia luego es más fina, donde el muestreo de la frecuencia es el mismo, resultando en un número de índices de frecuencias mayor. Sin embargo, el índice de las frecuencias de enmallado y sus anchos de bandas no cambian. Por ejemplo, las frecuencias de enmallado de un planetario con 228 dientes del engranaje son siempre $(228) \cdot (f_{eje})$ y sus múltiplos enteros, y los índices en el DFT son siempre 228 y sus múltiplos enteros. Podemos promediar las frecuencias con índices comunes y correspondientes y descartar las frecuencias extras de mayores índices. El resultado obtenido utilizando esta técnica de pre-proceso es equivalente a el DFT del promedio del tiempo síncrono (TSA), donde la interpolación y re-muestreo como se requería en el promediado del tiempo no es necesario. Este procedimiento reduce la complejidad computacional porque, en la mayoría de los casos, el DFT (FFT) es más sencillo y rápido que re-muestreo de los datos de vibración, especialmente cuando el muestreo de frecuencias es elevado.

Ya que el DFT es lineal, el DFT del ensamble promediado de los datos interpolados es equivalente al ensamble promediado del DFT de los datos interpolados, como se muestra a continuación:

$$X(k) = \sum_{n=0}^{N-1} \bar{x}(n) e^{-j2\pi(kn/N)} \quad k = 0, 1, \dots, N - 1 \quad (7.6)$$

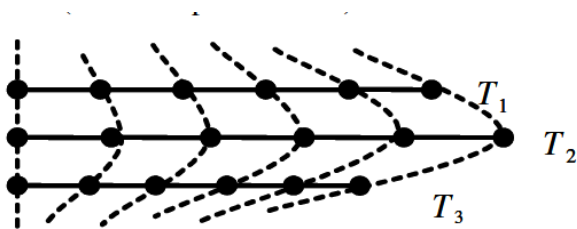
$$= \frac{1}{M} \sum_{n=0}^{N-1} \sum_{m=1}^M x_m(n) e^{-j2\pi(kn/N)} \quad (7.7)$$

$$= \frac{1}{M} \sum_{m=1}^M \sum_{n=0}^{N-1} x_m(n) e^{-j2\pi(kn/N)} \quad (7.8)$$

La figura ilustra el DFT de un conjunto de datos TSA, los cuales pueden ser calculados mediante el promediado de DFT, como se muestra en la parte derecha de la figura. En la figura los puntos de datos conectados por las líneas punteadas van a ser promediados. Ya que la longitud del tiempo varía en cada revolución individual, la resolución de la frecuencia del DFT de cada revolución varía también.

La figura 54 muestra el promediado del DFT para múltiples revoluciones sin interpolación. De las ilustraciones en el dominio de la frecuencia, se puede observar que los resultados son equivalentes a los índices comunes del DFT de los datos TSA mostrados en la figura 55. Ya que la variación en la velocidad rotacional del rotor es muy pequeño (usualmente la diferencia en el número de muestras entre dos revoluciones de datos primarios es alrededor de 0.05 por ciento), la diferencia entre los datos del DFT y TSA y el promediado del DFT del dato primario (sin interpolación) está al final del eje del espectro y es despreciable cuando la tasa de muestreo es muy alta (ej. 100 kHz).

TSA (Con datos interpolados)



Dominio de la frecuencia equivalente

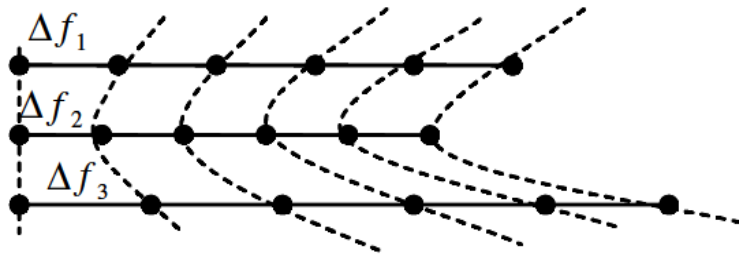


Figura 54. DFT de un conjunto de datos TSA

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

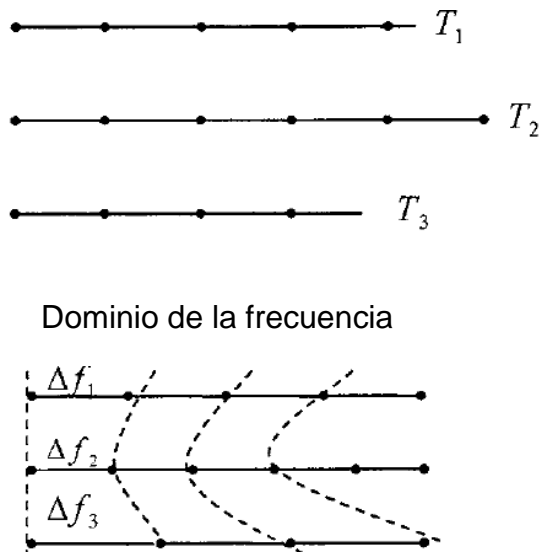


Figura 55. Revolución múltiple de datos sin interpolación en el dominio del tiempo y la frecuencia

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

Otra perspectiva consiste en visualizar el muestreo de las revoluciones individuales en el dominio del ángulo. La frecuencia en este caso es definida en el

dominio del ángulo en lugar del dominio del tiempo. Tiene una unidad de frecuencia del eje rotor en lugar de Hertz; esto es, que las frecuencias son múltiplos de la frecuencia del eje. Es una representación muy útil porque la frecuencia del diente es siempre un múltiplo específico de la frecuencia del eje sin importar como varíe la velocidad rotacional del eje. Los verdaderos valores de la frecuencia pueden cambiar con la velocidad rotacional del eje; sin embargo, los valores de las frecuencias como múltiplos de la frecuencia del eje puede que no cambien. Esto permite características invariantes de los datos de vibración a lo largo de las múltiples revoluciones para ser extraídas y mejoradas. La figura 56. permite observar las muestras en el dominio angular y el correspondiente dominio de frecuencia de los datos interpolados, como en el TSA. La figura 57. Ilustra las muestras de los datos en el dominio del ángulo y el correspondiente dominio de la frecuencia de los datos sin interpolación. De nuevo, podemos ver que en el dominio de la frecuencia las dos técnicas son equivalentes en frecuencias no triviales.

TSA (con datos interpolados) En el dominio del ángulo

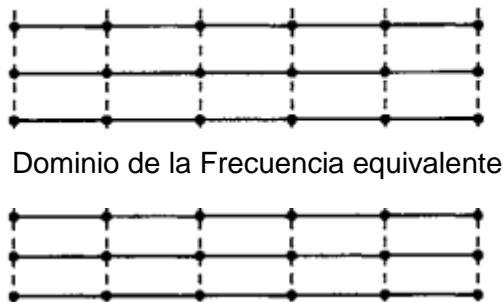


Figura 56. TSA en el dominio del ángulo y el correspondiente dominio de la frecuencia.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

Datos del dominio del ángulo (sin interpolación)



Dominio de la frecuencia



Figura 57. Datos sin interpolación en el dominio del ángulo y de la frecuencia.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

A.7.6.7 PROCESAMIENTO DE LA SEÑAL

A.7.6.7.1 Extracción y selección de características

La selección y extracción de características o estado/indicador de condición constituye la piedra angular de un diagnóstico de fallas preciso y confiable. El paradigma clásico reconocimiento- imagen y procesamiento- señal de:

Datos → información → conocimiento se vuelve más relevante y ocupa una etapa central en el caso particular del diagnóstico de fallas porque tales operaciones deben ser realizadas on-line en un marco de tiempo real. La figura 58. Presenta un esquema conceptual para la extracción de características y la clasificación del modo de falla.

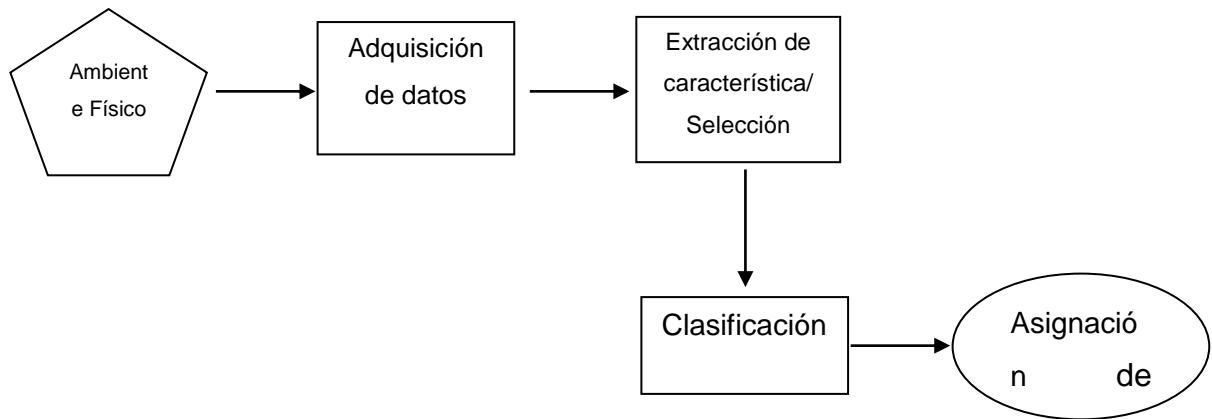


Figura 58. Esquema general de extracción de características y clasificación de modo de falla.

Fuente: Intelligent Fault Diagnosis and Prognosis for Engineering Systems. George Vachtsevanos et al.

El diagnóstico de fallas depende principalmente de la extracción un conjunto de características de los datos del sensor que puedan distinguir entre clases de fallas de interés y detectar e aislar una falla particular en su etapa inicial. Estas características deben ser bastante insensibles al ruido y entre la variación de clases de fallas. El último puede surgir por la ubicación de la falla, tamaño, etc. En el marco de un sensor. “buenas” características deben poseer los siguientes atributos:

Computacionalmente de bajo costo la medición, definible matemáticamente, explicable en términos físicos, caracterizada por una gran distancia media entre clases y pequeña varianza entre clases, insensible a variables externas, y no correlacionadas con otras características.

En años recientes se ha enfocado mucha atención en el problema de la extracción de características, donde la selección de características ha descansado primordialmente sobre la experiencia, observaciones, evidencia histórica, y nuestro entendimiento de signos características de fallas. En la selección de un conjunto óptimo de características, estamos tratando de resolver preguntas como: ¿Dónde está la información? Y ¿cómo los mecanismos de falla se relacionan con la física fundamental de un complejo sistema dinámico? Los modos de falla pueden inducir cambios en: la energía (potencia) del sistema, su entropía, el espectro de potencia, la magnitud de la señal, el comportamiento caótico entre otros. La selección de las características depende de la aplicación. Buscamos aquellas características, para una clase particular de modos de falla, de un gran conjunto de candidatos que poseen propiedades de distinción y detección de la falla mientras se logra una clasificación de fallas confiable en la menor cantidad de tiempo. La extracción de características, por otro lado, es vista como un proceso algorítmico donde, dado los datos del sensor, las características son extraídas en una manera computacional eficiente mientras preserva el máximo contenido de información. Así el proceso de extracción de características convierte los datos de falla en un espacio de características N-dimensional tal que una clase de fallas es agrupada y puede ser distinguida de otras clases. Sin embargo, en general, no todas las fallas de una clase necesitan N características para formar un grupo compacto. Solo las fallas que se encuentran en la región de traslape de dos o más clases que gobiernan el número de características requeridas para realizar la clasificación. Aunque prevalece la selección de características basada en la experiencia y/o evidencia heurística, es posible enfocar este gran problema desde una perspectiva analítica colocándolo como un problema de optimización: maximizar la precisión en detección/aislación mientras se minimizan falsas alarmas. Una función multi-objetivo puede definirse con criterios tales como distinción/detección y aislamiento que obliga mínimos requerimientos de procesado en aplicaciones de tiempo real. Además, las características deben ser ordenadas en una forma estándar, donde para los propósitos de clasificación de

fallas, un grado de certeza métrico puede ser llamado como el número límite de falsos positivos y falsos negativos.

La medida de distinción cuantifica la habilidad característica para diferenciar entre varias clases encontrando el área de la región en la cual las dos clases se traslapan. Entre más pequeña el área, más alta es la habilidad de la característica para distinguir entre las clases. Suponga que hay dos clases w_1 y w_2 que necesitan ser clasificadas basadas en la característica A. Teniendo un conjunto de datos de entrenamiento que incluye N_1 muestras pertenecientes a w_1 y N_2 muestras pertenecientes a w_2 . Asumiendo una distribución gaussiana para los datos de entrenamiento perteneciente a cada clase, la medida de la distinción de la característica A para las clases w_1 y w_2 es:

$$d^{w_1 w_2}_A = \int_{-\infty}^T \frac{1}{\sqrt{2\pi v_2}} \exp\left[-\frac{(z-\mu_2)^2}{2\sigma_2^2}\right] dz + \int_T^{\infty} \frac{1}{\sqrt{2\pi v_1}} \exp\left[-\frac{(z-\mu_1)^2}{2\sigma_1^2}\right] dz \quad (7.9)$$

Donde μ_1 es la media del conjunto de entrenamiento de datos w_1 , μ_2 es la media del conjunto de entrenamiento de datos w_2 , σ_1 es la desviación estándar about μ_1 , y σ_2 es la desviación estándar about μ_2 . Figura 59. Ilustra la medida de la distinguibilidad de las características gráficamente. La región sombreada en la figura es la región para la característica A en la cual las dos clases w_1 y w_2 se traslapan. Para poder calcular la medida de distinción, se necesita determinar T. esto se hace minimizando esta expresión, es decir, diferenciándola con respecto a T e igualando el resultado a cero.

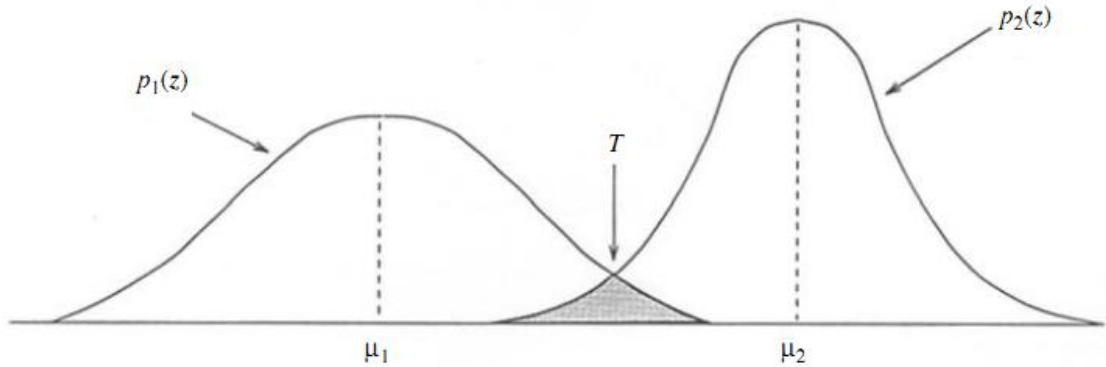


Figura 59. Medida de la distinción.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

La capacidad de detección es el grado para el cual el esquema de diagnóstico puede detectar la presencia de una falla particular; se relaciona con la más pequeña señal de falla que puede ser detectada y el porcentaje de falsas alarmas.

Una expresión que captura esta noción:

$$d_k(a_i) = \max_{a_1 \dots a_n} \left[\frac{F_k(a_i)}{N(a_i)s} (1 - \lambda |L_{\psi_{a_i}} - L_k|) \right] \quad i = 1, \dots, n \quad (7.10)$$

$$Y \quad D_k(a_i) = 1 - e^{-d_k(a_i)} \quad (7.11)$$

Donde

$F_k(a_i)$ = magnitud de la característica en w_i para la escala a_i para la k -ésima falla.

$N(a_i)s$ = valor medio cuadrado de el ruido aleatorio en w_i para la escala a_i

$L_{\psi_{a_i}}$ = longitud de la función wavelet para la escala a_i

s = factor de decimation factor escala en el escaneo fractal.

L_k = longitud de la característica para la k-esima falla

λ = constante menor que 1

La medida de identificación es la capacidad de seguir semejanzas de características como identificar un modo de falla k, pero que también los distingue de las otras clases de fallas, $p \neq k$. se puede expresar como

$$I_k(ai) = \frac{1}{n} \left[\sum_{i=1}^n sim(w_{i,k}, \beta_{ai,k}) - \sum_{i=1}^n \sum_{p=1, p \neq k}^M sim(w_{i,p}, \beta_{ai,k}) \right]$$

$i = 1, \dots, n$ (7.11)

Donde *sim* es una semejanza métrica definida como sigue: dejando que $X, Y \in [0,1]$ sean dos conjuntos con elementos x_i y y_i ; luego una medida de la semejanza se define por

$$Sim(X, Y) = \frac{1}{K} \sum_{i=1}^k \frac{1 - |x_i - y_i|}{1 + \alpha |x_i - y_i|}$$

(7.12)

Y $w_{i,k}$ son características de fallas o señales, y $\beta_{ai,k}$ son los modos de falla correspondientes. La figura 60. ilustra la semejanza métrica.

La capacidad de identificación va un paso más allá en distinguir entre varios modos de falla una vez la presencia de una falla ha sido establecida. Esto implica definiciones tales como la fuente, ubicación, tipo, y consecuencia de una falla y la distinción entre fallas del sensor, actuador y fallas de los componentes del sistema.

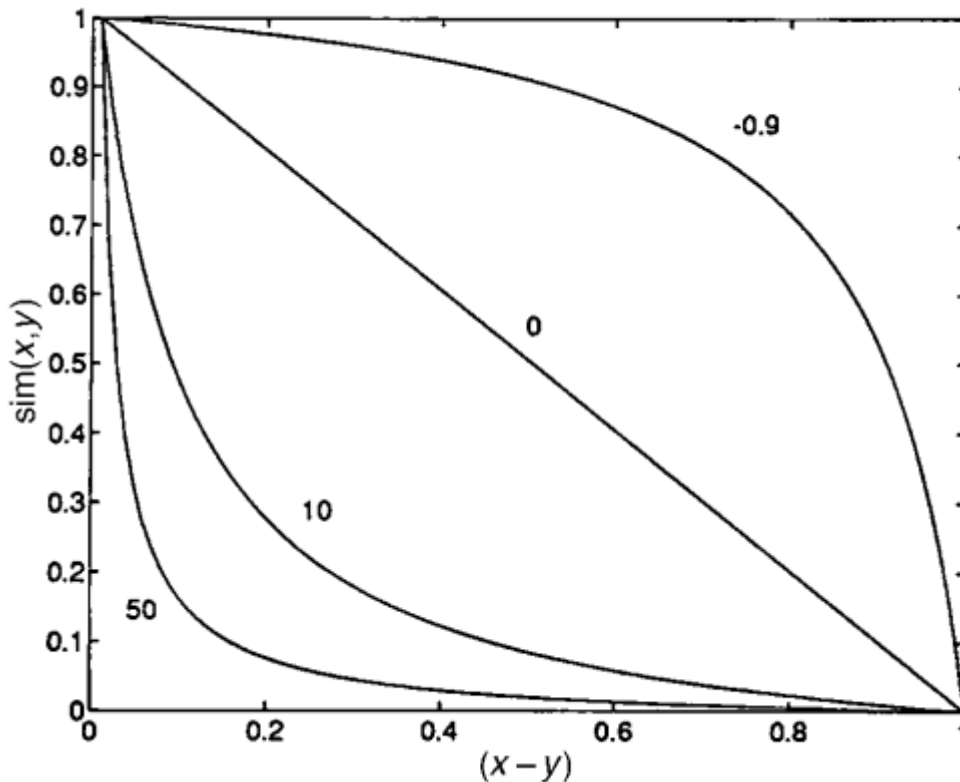


Figura 60 Semejanza métrica

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

Para los enfoques de clasificación de fallas supervisada, la información principal acerca de las clases de fallas está disponible en la forma de conjunto de entrenamiento de datos. Utilizando el conjunto de entrenamiento, la capacidad de distinción de cada característica puede ser determinada. Así, es una asunción razonable que la base de conocimiento contiene información acerca de la distinción de características de cada falla, al igual que el tiempo necesario para medir las características. Es generalmente cierto que el tiempo requerido para derivar un conjunto de funciones difiere ampliamente de una clase de característica a otra tanto como de un sensor a otro aparato similar. Para alcanzar

una clasificación confiable en la mínima cantidad de tiempo, las características deberían ser alineadas en un orden óptimo tal que las características que requieren menor tiempo de procesamiento y proveen una mayor distinción entre las clases, deberían ser los primeros en el esquema de ordenamiento de las características. Asumiendo que un conjunto de características ordenadas se ha obtenido, la clasificación fuzzy es realizada utilizando una característica a la vez del conjunto ordenado de características. Otros esquemas de clasificación también son posibles. Cada resultado de clasificación, basado en una característica, se convierte en un cuerpo de evidencia. Así, para cada característica, el clasificador provee un pedazo de evidencia. Para determinar el grado de confianza (DOC) alcanzado, basado en la evidencia más reciente y toda la evidencia previa, el DOC asociado con el cuerpo combinado de evidencia se calcula. La evidencia total es agregada hasta que el DOC requerido es alcanzado. El proceso de clasificación fuzzy es terminado tan pronto como suficiente evidencia es obtenida para alcanzar el pre-especificado DOC. Así, el esquema sugerido controla el proceso de recolección de información y utiliza una cantidad de información que es suficiente para la tarea en mano. La figura 61. muestra el enfoque de un ambiente multi-sensor. La base de conocimiento, como se muestra en la figura 62, representa la base de datos incluyendo el ordenamiento de características óptimo y el conjunto de datos de entrenamiento para el clasificador fuzzy. El procedimiento es genérico en el sentido de que puede ser aplicado en un ambiente multi-sensor para realizar la fusión de sensor al nivel de característica o solo características extraídas de un solo sensor. Para obtener el orden óptimo de características disponibles, el tiempo de procesamiento para la clasificación y la distinción para cada característica necesita ser evaluada. Estas evaluaciones se pueden efectuar como un proceso de aprendizaje offline utilizando el conjunto de datos de entrenamiento. El ordenamiento considera características de todos los sensores disponibles, así fusionando información en el nivel de características. Un ordenamiento óptimo es efectuado para cada par de clases. Si tenemos que existen c clases de fallas que necesitan clasificarse. El número total de conjuntos de características ordenadas

requerido es $c(c - 1)/2$. Un conjunto de características ordenado F_{ij} representa el ordenamiento de las características que maximiza la distinción entre la i -ésima y j -ésima clase mientras se mantiene el tiempo de procesamiento en un rango bajo. Los DOCs para todas las clases posibles son calculados basados en la evidencia provista para las características actuales y todas las previas. Si el valor más alto del DOC es menor que el umbral pre-especificado, la siguiente característica es seleccionada para así maximizar la distinción entre las clases con mayor DOC y aquella con el segundo más alto DOC. Así, el ordenamiento de características para una falla desconocida es modificado dinámicamente de forma on-line durante el proceso de clasificación. El rango de características cercanas a la determinación una falla específica, típicamente está basado en el cociente del discriminante de Fisher (FDR):

$$FDR = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \quad (7.13)$$

Donde μ_1, σ_1 = desviación estándar media de la característica de falla 1

μ_2, σ_2 = la media y la desviación estándar de la falla.

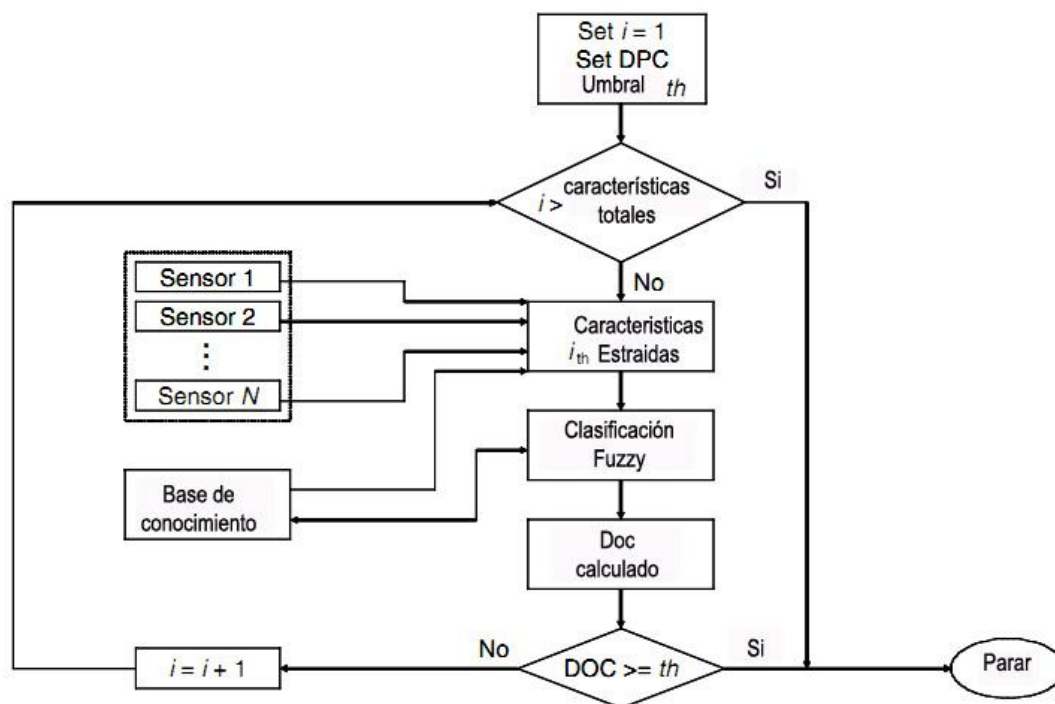


Figura 61. El procedimiento de percepción activa para patrones de reconocimiento.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

A.7.6.8 PROCESAMIENTO DE LA SEÑAL EN EL DOMINIO DEL TIEMPO

Para utilizar información de los sensores CBM/PHM para diagnóstico, se debe saber cómo computar los momentos estadísticos, que se describen a continuación. El cuarto momento, kurtosis, se ha encontrado especialmente útil en el diagnóstico de fallas porque captura el tipo de movimiento abrupto (bang-bang) propio de componentes dañados o desajustados. Dada una variable aleatoria, su p-ésimo momento es computado utilizando el ensamble promedio

$$E(x^p) = \int x^p f(x) dx \quad (7.14)$$

Con $f(x)$ como la función de densidad de probabilidad. Si una variable aleatoria es ergódica¹⁸, luego su ensamble de promedios puede ser aproximado por su promedio de tiempo.

- **Momentos alrededor de cero y energía**

Dada una serie de tiempo discreta x_k , el momento p-esimo sobre el intervalo $[1, N]$ esta dado por

$$\frac{1}{N} \sum_{k=1}^N x_k^p \quad (7.15)$$

El primer momento es el valor medio (muestra)

$$\bar{x} = \frac{1}{N} \sum_{k=1}^N x_k \quad (7.16)$$

Y el segundo momento es el momento de inercia

$$\frac{1}{N} \sum_{k=1}^N x_k^2 \quad (7.17)$$

A medida que N se vuelve más grande, más se aproxima al ensamble de promedios correspondiente.

La energía de la serie de tiempo sobre el intervalo $[1, N]$ esta dado por

$$\sum_{k=1}^N x_k^2 \quad (7.18)$$

Y el valor de la raíz media cuadrada (RMS) es

¹⁸ Una variable ergódica se define como aquella donde los promedios estadísticos son iguales a los temporales.

$$\sqrt{\frac{1}{N} \sum_{k=1}^N x_k^2} \quad (7.19)$$

- **Momentos alrededor de la media**

Dada una serie de tiempo x_k , el momento p-esimo alrededor de la media sobre el intervalo $[1, N]$ esta dado por

$$\frac{1}{N} \sum_{k=1}^N (x_k - \bar{x})^p \quad (7.20)$$

El segundo momento alrededor de la media es la varianza (muestra)

$$\sigma^2 = \frac{1}{2} \sum_{k=1}^N (x_k - \bar{x})^2 \quad (7.21)$$

Donde σ es la desviación estándar. Desafortunadamente, la varianza de la muestra es una estimación sesgada de la varianza del verdadero ensamble, por lo cual en estadística a menudo se utiliza en lugar de la varianza sesgada:

$$S^2 = \frac{1}{N-1} \sum_{k=1}^N (x_k - \bar{x})^2 \quad (7.22)$$

El tercer momento alrededor de la media contiene información acerca de la simetría de la distribución. Una medida común de la simetría es la oblicuidad:

$$\frac{1}{N\sigma^3} \sum_{k=1}^N (x_k - \bar{x})^3 \quad (7.23)$$

La oblicuidad desaparece para distribuciones simétricas y es positiva (negativa) si la distribución desarrolla una larga cola a la derecha (izquierda) de la media $E(x)$. Mide la cantidad de dispersión (spread) de la distribución in either dirección de la media. Ambas funciones de densidad de probabilidad en la figura 62. tienen la misma media y desviación estándar. Aquella que se encuentra a la izquierda es positivamente oblicua. La que está a la derecha es negativamente oblicua.



Figura 62. Oblicuidad, una medida de la simetría de la distribución.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

También es útil la kurtosis, definida como

$$\frac{1}{N\sigma^4} \sum_{k=1}^N (x_k - \bar{x})^4 - 3 \quad (7.24)$$

La kurtosis mide la contribución de las colas de la distribución, como se ilustra en la figura 63. La PDF a la derecha tiene mayor kurtosis que la de la izquierda. Es posible que una distribución tenga la misma media, varianza y oblicuidad y no tener la misma medición de kurtosis.

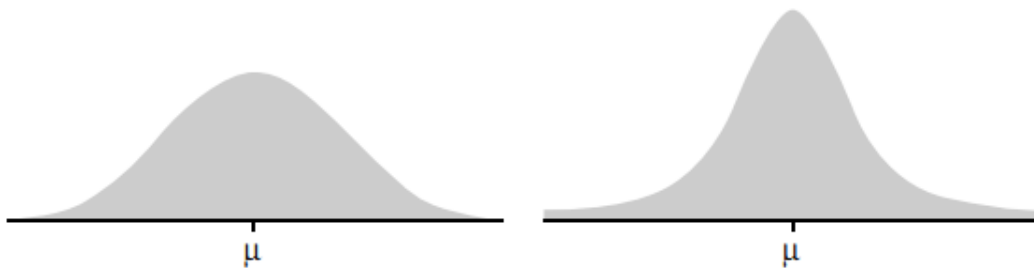


Figura 63. Kurtosis, como una medida del tamaño de los lóbulos de una distribución.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

- **Correlación, covarianza, convolución**

La (auto) correlación está dada por

$$R_x(n) = \frac{1}{N} \sum_{k=1}^N x_k x_{k+n} \quad (7.25)$$

Note que $NR_x(0)$ es la energía. La (auto)covarianza está dada como

$$P_x(n) = \frac{1}{N} \sum_{k=1}^N (x_k - \bar{x})(x_{k+n} - \bar{x}) \quad (7.26)$$

Note que $P_x(0) = \sigma^2$

Dadas dos series de tiempo x_k , y_k la correlacion cruzada es

$$R_{xy}(n) = \frac{1}{N} \sum_{k=1}^N x_k y_{k+n} \quad (7.27)$$

Y la covarianza cruzada es

$$P_{xy}(n) = \frac{1}{N} \sum_{k=1}^N (x_k - \bar{x})(y_{k+n} - \bar{y}) \quad (7.28)$$

A menudo utilizada en estadística en la correlación de Pearson, la cual se define como:

$$r = \frac{\sum x_k y_k - \frac{\sum x_k \sum y_k}{N}}{\sqrt{\left[\sum x_k^2 - \frac{(\sum x_k)^2}{N} \right] \left[\sum y_k^2 - \frac{(\sum y_k)^2}{N} \right]}} \quad (7.29)$$

El tiempo-discreto de convolución para N puntos de secuencias esta dado por

$$x * y(n) = \sum_{k=0}^{N-1} x_k y_{n-k} \quad (7.30)$$

Lo cual no es más que una multiplicación polinomial. Es interesante notar que la correlación es expresada en términos del operador de convolución por

$$R_{xy}(n) = \frac{1}{N} \sum_{k=1}^N x_k y_{k+n} = \frac{1}{N} x(k) * y(-k) \quad (7.31)$$

- **Transformada de Fourier discreta (DFT)**

Dada un N-punto de una serie de tiempo $x(n)$, la DFT está dada por

$$X(k) = \sum_{n=1}^N x(n)e^{-j2\pi(k-1)(n-1)/N} \quad k = 1, 2, \dots, N \quad (7.32)$$

La inversa de DFT está dada por

$$x(n) = \frac{1}{N} \sum_{k=1}^N X(k)e^{j2\pi(k-1)(n-1)/N} \quad n = 1, 2, \dots, N \quad (7.33)$$

La DFT es periódica con un periodo N. La inversa de la eficiencia de DFT es una función de tiempo que es periódica con periodo N. Existe un algoritmo eficiente para computar el DFT conocido como transformada rápida de Fourier (FFT) que se encuentra disponible en paquetes de software como Matlab.

A.7.6.9 METODOS DE GERENCIAMIENTO DE BASES DE DATOS

Lo principal para una arquitectura de pronóstico y diagnóstico confiable, eficiente, y robusta es una base de datos bien diseñada y con un esquema de gerenciamiento flexible. La base de datos debe ser capaz de almacenar una variedad de datos del sensor preferiblemente en una configuración relacional de cliente-servidor. El sistema de gerenciamiento relacional facilita la representación, conversión y la transferencia de datos para utilizar con los algoritmos de extracción de características y de diagnóstico y pronóstico. La figura 64. representa un esquema conceptual de la función del servidor de datos en la arquitectura CBM/ PHM, mientras que la figura 65. Provee una muestra de la esencia del esquema de la base de datos. El servidor de datos debe proveer almacenamiento para la configuración de la información, colección dinámica de los datos (dominio del tiempo, fase, y espectro), datos estáticos (proceso), y análisis

de resultados. Una base de datos secundaria puede extender la funcionalidad de la base de datos principal incluyendo tablas que definan programación de componentes, el mapeado de entrada del multiplexor, la información en la calibración del sensor, el mapeo de la base de datos principal y el mapeo para el reporte de los resultados a la interfaz gráfica de usuario (GUI).

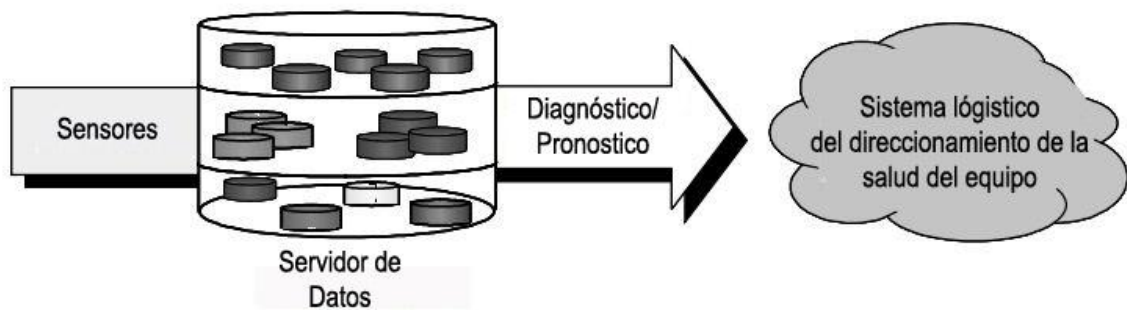


Figura 64. Servidor de datos como componente de la arquitectura CBM/PHM

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

El sistema de gerenciamiento de la base de datos relacional (RDMS) provee las siguientes funciones:

- Organizar una gran cantidad de datos en tablas interconectadas para hacer el diseño más fácil de entender.
- Proveer un completo lenguaje relacional para la definición de datos, recuperación y actualización.
- Provee reglas de integración que definen una consistente base de datos que mejoral la confiabilidad de los mismos.
- Ejemplos de RDBMS incluyen Oracle, Microsoft SQL server, y aplicaciones de base datos de oficina tales como Microsoft Access.

SQL es un lenguaje estándar bien definido, utilizado para cuestionar, actualizar, y administrar las bases de datos relacionales. Mediante el uso de SQL, una aplicación le puede pedir a la base de datos realizar tareas en lugar de requerir el código de aplicación para efectuarlos. Un efectivo uso del SQL puede minimizar la cantidad de datos enviada a través de la red. Finalmente, las directrices SQL SELECT devuelven información de la base de datos como un conjunto de filas seleccionadas.

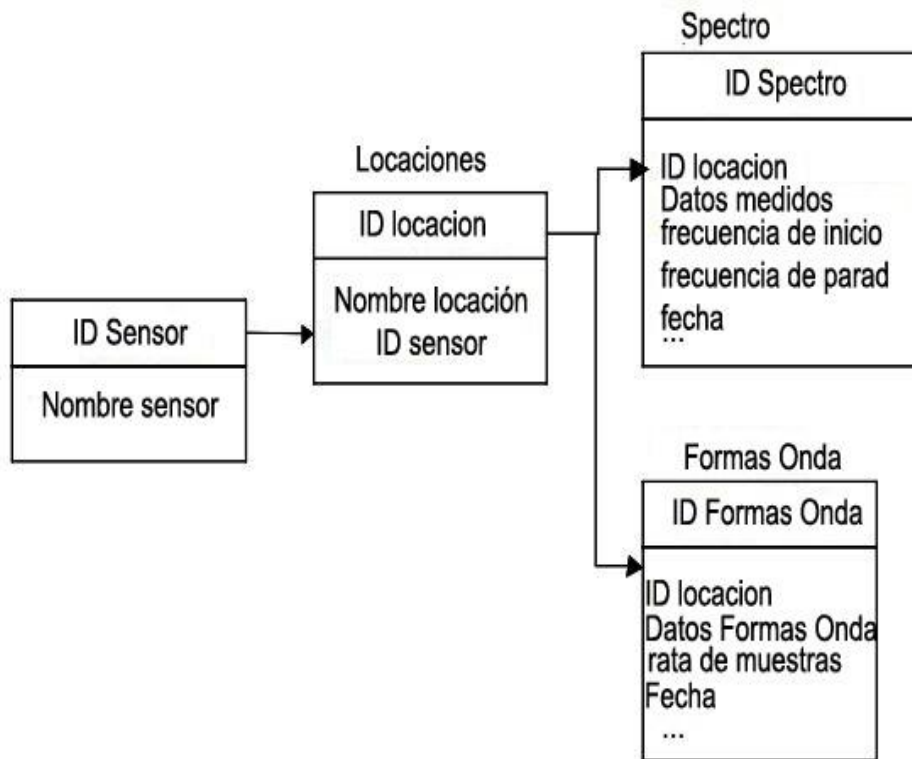


Figura 65. Una muestra de un esquema núcleo de base de datos.

Fuente: *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. George Vachtsevanos et al.

A.8. ESTADÍSTICA MULTIVARIANTE

A.8.1 INTRODUCCIÓN

La efectividad de las mediciones data-driven dependen de la caracterización de las variaciones de los datos del proceso. Existen dos tipos de variaciones en los datos de procesos: **causas comunes y causas especiales**. Las variaciones por causas comunes son las debidas enteramente a ruidos aleatorios (por ejemplo, el asociado a lecturas de sensores), mientras que las variaciones por causas especiales se cuentan como variaciones no atribuidas a causas comunes. Estrategias de control de procesos estándar pueden ser capaces de remover la mayoría de las variaciones por causas especiales, pero dichas estrategias no pueden remover las variaciones por causas comunes, que son inherentes a los datos del proceso. Ya que las variaciones en los datos del proceso son inevitables, teoría estadística juega un papel importante en la mayoría de esquemas de monitoreo de procesos.

La aplicación de teoría estadística al monitoreo de procesos se basa en la suposición que las características de las variaciones en los datos son relativamente invariables a menos que una falla ocurra en el sistema. Mediante la definición de una falla como una condición anormal del proceso, es una suposición razonable. Esto implica que las propiedades de las variaciones de los datos, tales como la media y la varianza, son repetibles para las mismas condiciones de operación, aunque los valores actuales de los datos puedan no ser predecibles. La repetitividad de las propiedades estadísticas permite límites para ciertas

mediciones, definiendo eficazmente el estatus de lo fuera-de-control, ser determinado automáticamente. Éste es un importante paso para la automatizar el esquema de monitoreo de un proceso, como lo ilustra la figura 66.



Figura 66. Estadística en el monitoreo de procesos

Fuente: Autores

El propósito es el de ilustrar como usar los métodos para monitoreo de procesos, en particular métodos usando la estadística multivariante T^2 .

A.8.1.2 PRE-TRATAMIENTO DE DATOS

Para extraer la información relevante de los datos para efectivamente monitorear el proceso, es comúnmente necesario pre-tratar los datos en el conjunto de entrenamiento. El conjunto de entrenamiento contiene datos off-line disponibles para análisis previo a la implementación en línea del plan de monitoreo del proceso y es usado para desarrollar las medidas que representan las operaciones en-control y las distintas fallas. El procedimiento de pre-tratamiento consiste de

tres tareas: **remoción de variables**, **auto-escalamiento**, y **remoción de datos fuera de rango**.

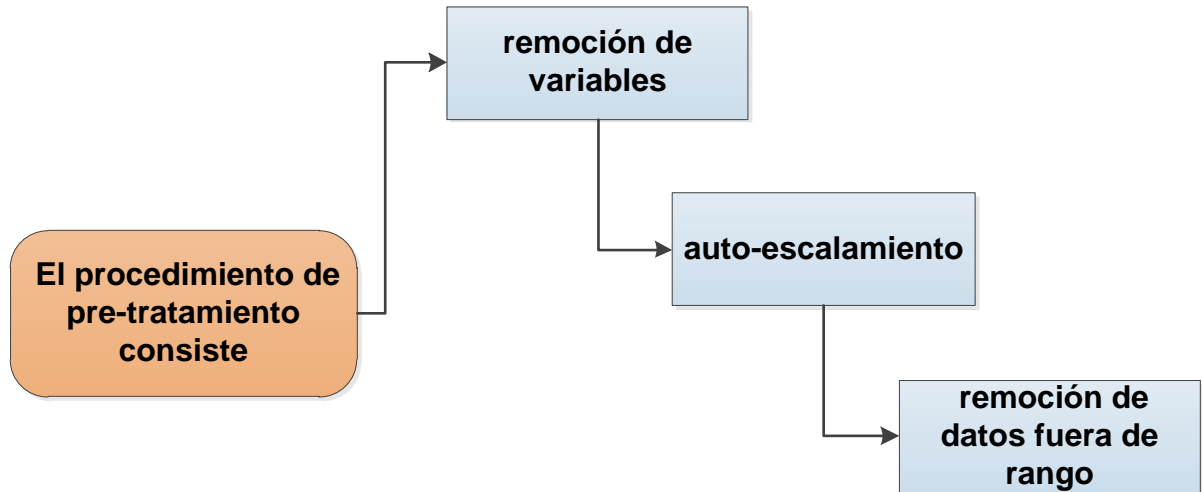


Figura 67. Procedimiento para el pretratamiento de datos.

Fuente: Autores

Los datos en el conjunto de entrenamiento pueden contener variables que no tienen información relevante para el monitoreo del proceso, y dichas variables deben ser removidas antes de más análisis. Por ejemplo, se puede saber *a priori* que ciertas variables exhiben errores de medición extremadamente largos, tal como aquellos de debido a calibraciones impropias de sensor, o algunas de las variables pueden ser físicamente separadas de la porción del proceso que está siendo monitoreado. En dichas instancias, la capacidad del método de monitoreo del proceso puede ser mejorada por la remoción de las variables inapropiadas.

Datos de procesos comúnmente necesitan ser escalados para evitar variables particularmente dominantes del método de monitoreo del proceso, especialmente métodos basados en técnicas de reducción dimensional, tales como PCA y FDA.

Por ejemplo, cuando desempeñando una reducción dimensional sin escala del proceso de mediciones de temperatura variando entre 300K y 320 K y variaciones en la medición de concentración entre 0.4 y 0.5, las mediciones de temperatura dominarán inclusive si las mediciones de temperatura no sean más importantes que las mediciones de concentración para el monitoreo del proceso.

EL auto-escalamiento estandariza las variables del proceso de tal forma que asegura que cada variables le es dada igual peso antes de la aplicación del método de monitoreo del proceso. Esto consiste en dos pasos. El primer paso es sustraer cada variable por su media de muestra; ya que el objetivo es capturar la variación de los dato de la media. El segundo paso es dividir cada variable de los datos de la media centrada por su desviación estándar. Este paso escala cada variable a variación unitaria, garantizando que las variables del proceso con altas varianzas no sean dominantes. Cuando es aplicado el auto-escalamiento a nuevos datos del proceso, la media a ser restada y la desviación estándar a ser dividida son tomadas del conjunto de entrenamiento.

Valores fuera de rango son valores aislados medidos que son erróneos. Dichos valores pueden influenciar significativamente la estimación de los parámetros estadísticos y otros parámetros relacionados a una medida dada. Remover los valores fuera de rango del conjunto de entrenamiento puede mejorar significativamente la estimación de los parámetros y debe ser un paso esencial en el pre-tratamiento de datos. Obviamente los valores fuera de rango pueden ser removidos mediante el trazado e inspeccionando visualmente los datos de valores fuera de rango. Métodos más rigurosos basados en umbrales estadísticos pueden ser empleados para remover valores fuera de rango, y un método para hacerlo usando la estadística T^2 es discutido más adelante. Por simple presentación, es asumido que los datos han sido pre-tratados, a menos que se indique lo contrario.

A.8.1.3 MONITOREO ESTADÍSTICO UNIVARIANTE

La disciplina de la estadística enseña cómo razonar de manera lógica y tomar decisiones informadas en presencia de incertidumbre y variación. Sin estas dos últimas, habría poca necesidad de los métodos estadísticos o de los profesionales de la estadística. Si cada pieza de un tipo específico tuviera exactamente la misma vida, si todos los resistores que produce determinado fabricante tuvieran la misma resistencia, si las determinaciones de ph en muestras de suelo de determinado lugar fueran idénticas etc., entonces se conocería toda la información que se desea en una sola observación.

- ***Poblaciones, muestras y procesos***

Los ingenieros y los científicos tratan de formar constante con conjunto de hechos, o **datos**, tanto en su práctica profesional como en la vida cotidiana. La disciplina de la estadística proporciona métodos para organizar y resumir datos, y para sacar conclusiones con base a la información que contiene los datos.

Una investigación tiene por lo común que ver con una colección muy bien definida de objetos que forman una **población** de interés. En un estudio, la población podría consistir en todas las capsulas de gelatina, de determinado tipo, producidas durante un periodo específico. Cuando se cuenta con la misma información para todos los objetos de la población, se tiene lo que se llama **censo**. Las restricciones de tiempo, dinero y demás recursos insuficientes, hacen impráctico o imposible levantar un censo. En vez de eso, se selecciona de alguna manera prescrita un

subconjunto de la población, una **muestra**. Por consiguiente, se podría obtener como base una muestra de rodamientos de una corrida de producción determinada, con el fin de investigar si éstos se apegan a las especificaciones de manufactura.

Por lo regular, el interés se centra sólo en ciertas características de los objetos de una población: la cantidad de defectos en la superficie de cada tramo de tubo, el espesor de pared de cada cápsula, el género de un graduado en ingeniería, la edad a la que se graduó, entre otras.

Una característica puede ser categórica, como el género o tipo de defecto, o bien puede ser de naturaleza numérica. Una **variable** es cualquier característica cuyo valor cambia entre los objetos de la población. Las variables se representaran como letras minúsculas del abecedario.

Los datos son el resultado de observaciones efectuadas en una sola variable, o bien, de observar de modo simultaneo dos o más variables. Un conjunto de datos **univariados** consiste en observaciones de una sola variable.

Los datos **bivariados** se obtienen cuando se observa dos variables. Se tienen datos **multivariados** cuando se hacen observaciones en más de dos variables.

A.8.1.4 MEDIDAS DE LOCALIZACIÓN

Los resúmenes visuales de datos son herramientas excelentes para obtener impresiones y puntos de vista preliminares. El análisis más formal de los datos requiere calcular e interpretar medidas de resumen numéricas; es decir, a partir de los datos, se intenta extraer varias cantidades de resumen: números que puedan servir para caracterizar el conjunto de datos y mostrar algunas de sus propiedades sobresalientes. El interés fundamental tiene que ver con los datos numéricos.

Supóngase, entonces, que el conjunto de datos es de la forma x_1, x_2, \dots, x_n , donde cada x_i es un número. ¿Qué características de este conjunto son más interesantes y merecen mayor atención? Una característica importante de un conjunto de números es su localización y, en particular, su centro.

- **La Media**

Para un determinado conjunto de números x_1, x_2, \dots, x_n , la media de tendencia central más familiar y útil es la *medida*, o promedio aritmético del conjunto. Debido a que casi siempre se considera a las x_i , como partes constituyentes de una muestra, se hace referencia al promedio aritmético como la *media muestral* y se denota por \bar{x} .

La **media muestral** \bar{x} de un conjunto de observaciones x_1, x_2, \dots, x_n está dada por

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} \quad (8.1)$$

El numerador de \bar{x} se puede escribir de manera más formal como $\sum x_i$ donde la suma es sobre las observaciones de la muestra.

Del mismo modo que \bar{x} representa el valor promedio de las observaciones de una muestra, se puede calcular el promedio de los valores de la población. Este promedio se llama **media poblacional** y se denota mediante la letra griega μ . Cuando hay N valores en la población (una población finita), entonces $\mu =$ (suma de los N valores de la población) $/N$

- **La Mediana**

La palabra mediana es sinónimo de “mitad”, y la mediana muestral es de hecho el valor medio cuando las observaciones se ordenan de menor a mayor. Si las observaciones se denotan por x_1, \dots, x_n , se emplea el símbolo \tilde{x} para representar la mediana muestral.

La **mediana muestral** se obtiene al ordenar primero las n observaciones de menor a mayor (incluso valores repetidos de manera que cada observación muestral aparezca en la lista ordenada). Entonces,

$$\tilde{x} \left\{ \begin{array}{l} \text{Único valor} \\ \text{medio si } n \\ \text{es impar} \\ \text{Promedio de} \\ \text{los dos} \\ \text{valores} \\ \text{medios si } n \\ \text{es par} \end{array} \right. = \begin{array}{l} = \frac{(n+1)^{\text{ésimo}}}{2} \text{ Valor ordenado} \\ \cdot \\ \cdot \\ \cdot \\ = \text{promedio de los valores ordenados} \\ \left(\frac{n}{2}\right)^{\text{ésimo}} \text{ y } \left(\frac{n}{2}+1\right)^{\text{ésimo}} \end{array}$$

Análogo a \tilde{x} como el valor medio de la muestra es un valor medio de la población, la **mediana poblacional**, denotada por $\tilde{\mu}$ al igual que con \bar{x} y μ .

La media μ y la mediana $\tilde{\mu}$ de la población en general no son idénticas. Si la distribución de la población tiene un sesgo positivo o negativo como se observa en la figura:

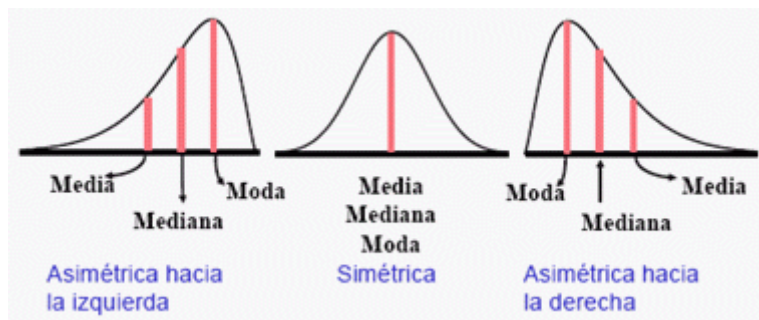


Figura 68. Simetría de los Datos según el sesgo

Fuente: Probabilidad y estadística para ingeniería y ciencias. Jay L. Devore.

A.8.1.5 MEDIDAS DE VARIABILIDAD

Medir una característica de tendencia central sólo proporciona información parcial de un conjunto de datos o distribución. Las distintas muestras o poblaciones pueden tener medidas idénticas de tendencia central, pero diferir entre sí en otros aspectos importantes. En la figura 69. se muestran los diagramas de puntos de tres muestras con la misma media y mediana; sin embargo, el grado de dispersión en torno al centro es distinto para las tres muestras. La primera muestra tiene la mayor cantidad de variabilidad, la tercera tiene cantidad más pequeña y la segunda es intermedia con las otras dos en este proceso.

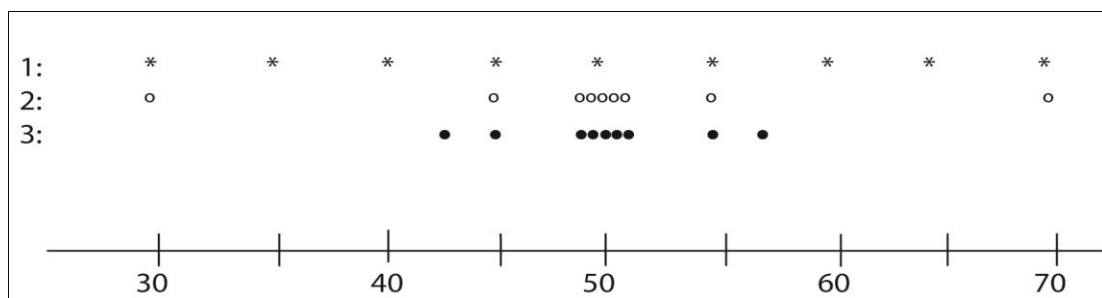


Figura 69. Muestras con medidas de tendencia central idénticas, pero cantidades distintas de variabilidad

Fuente: Probabilidad y estadística para ingeniería y ciencias. Jay L. Devore.

- **Medidas de variabilidad para datos muestrales**

La medida más sencilla de la variabilidad en una muestra es el **rango o recorrido**, que es la distancia entre valores máximos y mínimos de la muestra. Observe que el valor del recorrido para la muestra 1 en la figura 8.4 es mucho

mayor que para la muestra 3, lo que refleja más variabilidad en la primera muestra que en la tercera. Sin embargo, una desventaja del recorrido es que depende sólo de las dos observaciones más extremas, y no toma en cuenta las posiciones de los $n-2$ valores restantes. Las muestras 1 y 2, en la figura 8.4, tienen recorridos idénticos, pero cuando se toman en cuenta las observaciones entre los dos extremos, hay mucho menos variabilidad o dispersión en la segunda muestra que en la primera.

Las principales medidas de variabilidad tienen que ver con las **desviaciones a partir de la media**, $x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x}$. Es decir, las desviaciones respecto a la media se obtienen restando \bar{x} de cada una de las n observaciones de la muestra. Una desviación es positiva si la observación es mayor que la media (está a la derecha de la media en el eje de medición), y negativa si es menor que la media. Si es pequeña la magnitud de las desviaciones, entonces todas las x_i quedan lejos de \bar{x} lo que indica una mayor cantidad de variabilidad. Una forma simple de combinar las desviaciones es una sola cantidad es promediarlas (sumarlas y dividir las entre n). Sin embargo, hay un problema con esta sugerencia:

$$\text{suma de las desviaciones} = \sum_{i=1}^n (x_i - \bar{x}) = 0 \quad (8.2)$$

¿Cómo cambiar las desviaciones a cantidades no negativas para que las desviaciones positivas y negativas no se contrarresten entre sí cuando se combinan? Una posibilidad es trabajar con valores absolutos de las desviaciones y calcular la desviación absoluta promedio, pero debido a que la operación del valor absoluto genera varias dificultades teóricas, considere en su lugar desviaciones elevadas al cuadrado $(x_1 - \bar{x})^2, (x_2 - \bar{x})^2, \dots, (x_n - \bar{x})^2$ y en vez de usar la

desviación promedio elevada al cuadrado $\sum_{i=1}^n (x_i - \bar{x})^2/n$, por varias razones se divide mejor la suma de las desviaciones al cuadrado entre $n - 1$ y no entre n .

La **varianza muestral**, que se denota por s^2 se define por

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{S_{xx}}{n-1} \quad (8.3)$$

La **desviación estándar muestral**, representada por s , es la raíz cuadrada positiva de la varianza:

$$s = \sqrt{s^2} \quad (8.4)$$

Las variables del proceso pueden tener diferentes tipos de distribución de probabilidad. Más sin embargo, si una variable es influenciada por varias entradas que tienen distintas distribuciones de probabilidad, entonces la probabilidad de distribución de las variables del proceso se enfoca a una distribución asintótica Normal (Gaussiana).

Tabla 8.1 Población y muestras estadísticas

Estadística	Población (tamaño m_p)	Muestra (tamaño m_s)
Media	$\mu = \frac{1}{m_p} \sum_{i=1}^{m_p} x_i$	$\bar{x} = \frac{1}{m_x} \sum_{i=1}^{m_s} x_i$
Varianza	$\sigma^2 = \frac{1}{m_p} \sum_{i=1}^{m_p} (x_i - \mu)^2$	$S^2 = \frac{1}{m_s - 1} \sum_{i=1}^{m_s} (x_i - \bar{x})^2$
Rango	$R_i = \max(x_i) - \min(x_i)$	$i = 1, \dots, m_p \text{ O } m_s$

Las características de la población que sigue la distribución normal son resumidas mediante su media y su varianza. La varianza se puede también deducir a partir del rango de variables para pequeños tamaños de muestra. La convención en la suma y representación de la media de valores es

$$\bar{x} = \frac{1}{m} \sum_{j=1}^m x_{ij}, \quad (8.5) \quad \bar{x}_{..} = \frac{1}{mn} \sum_{i=1}^n \sum_{j=1}^m x_{ij}, \quad (8.6)$$

Donde n es el número de muestras (grupos) y m es el número de observaciones en la muestra (tamaño de muestra). Los subíndices “.” Indican el índice que se utiliza en promedio. Cuando no hay ambigüedad, valores promedio son denotados usando $\bar{x}_y \bar{\bar{x}}$. La población y muestras estadísticas para variables que tienen una distribución normal son dadas en la tabla 8.1

En procesos químicos, es común que una simple medición del proceso o un producto variable es hecho en un instante de muestra. La falta de múltiples observaciones limita el uso de los clásicos gráficos charts. La simple observación de cada tiempo de muestra y la existencia de mediciones erráticas aleatorias han hecho que las técnicas SPM basadas en sumas acumuladas, que medias móviles y rangos móviles sean atractivos para la evaluación de desempeño. A menudo, las decisiones tienen que ser hechas sobre poblaciones basadas en la información de la muestra. Una *hipótesis estadística* es una suposición acerca de la población. Es expresado como un estado de los parámetros de la probabilidad de distribución de la población. Procedimientos que permiten toma de decisiones ya sea aceptar o rechazar la hipótesis son llamados *pruebas de hipótesis*. Por ejemplo, si la igualdad de la media de la variable μ a un valor a esta por ser probada, la hipótesis es:

Hipótesis nula: $H_0: \mu = a$ (8.6)

Hipótesis alternativa $H_1: \mu \neq a$ (8.7)

Dos clases de error pueden ser cometidos cuando se prueba una hipótesis: rechazando una hipótesis cuando es verdadera, y aceptando una hipótesis cuando esta es falsa. La primera es llamada Tipo 1 o error α . La segunda es llamada Tipo II o error β . Este es el riesgo del consumidor porque un producto defectuoso no ha sido detectado y es enviado al consumidor. Esto puede ser resumido como,

Tipo I (α) error $P \{rechazada H_0 | H_0\}$

(Riesgo productor): $P \{falla en ser rechazada H_0 | H_0 \text{ es falso}\}$

En el desarrollo de las gráficas SPM¹⁹, primero α es seleccionada para computarizar el límite de confianza para probar la hipótesis. Después, un ensayo de procedimiento es designado para obtener un valor pequeño de β , si es posible. α Es una función del tamaño de la muestra y es reducida a medida que el tamaño de la muestra aumenta. Figura 70. muestra gráficamente los errores α y β para una variable que tiene distribución normal. En el grafico superior, en área bajo la curva a la izquierda de la línea que denota el valor $x = a$ es el error α . En la figura inferior, la media de x se ha desplazado de x_1 a x_2 . El área a la derecha de la línea $x = a$ denota el error β .

¹⁹ Por sus siglas en inglés, Statistical Process Monitoring.

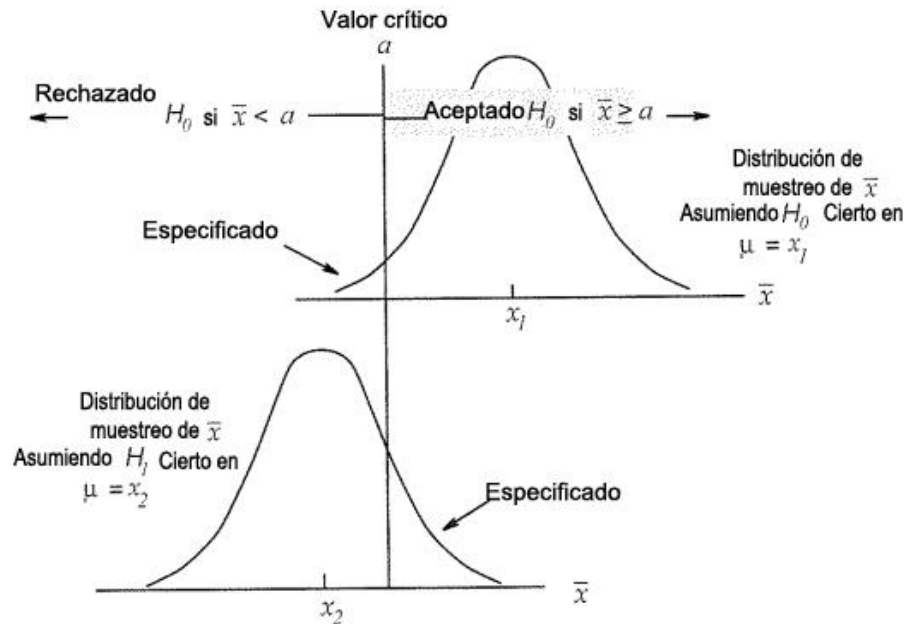


Figura 70. Errores tipo I (α) y tipo II (β)

Fuente: *Chemical Process Performance Evaluation*. Heinz Heinemann.

El valor para el error α puede ser computarizado para gráficas simples SPC²⁰ tal como la gráfica Shewhart usando derivaciones teóricas. Incrementando el umbral (desplazando la línea vertical hacia la derecha de la figura 70) decrece la rata de falsas alarmas pero incrementa la rata de detecciones pérdidas. Los intentos de disminuir las ratas de falsa alarma son usualmente acompañados con incrementos en la rata de detecciones pérdidas, la única manera de conseguir esta compensación es recolectando más datos, o reducir la variabilidad normal del proceso (ej. Mediante la instalación de sensores de alta precisión). El valor del error tipo I, también llamado **nivel de significancia** α , especifica el grado de compensación entre la rata de falsas alarmas y la rata de detecciones pérdidas.

²⁰ De las siglas en inglés, Statistical Process Control

Como un ejemplo específico, asumir para la hipótesis nula que cualquier desviación de las variables del proceso x de un valor deseado μ son debidas a mediciones inherentes y variabilidades del proceso descritas por la distribución normal con desviación estándar σ :

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad (8.7)$$

La hipótesis alternativa es que $x \neq \mu$. Asumiendo que la hipótesis nula es verdadera, la probabilidad que x esté en intervalos certeros es:

$$Pr\{x < (\mu - C_{\alpha/2}\sigma)\} = \alpha/2 \quad (8.8)$$

$$Pr\{x > (\mu + C_{\alpha/2}\sigma)\} = \alpha/2 \quad (8.9)$$

$$Pr\{x(\mu - C_{\alpha/2}\sigma) \leq x \leq (\mu + C_{\alpha/2}\sigma)\} = 1 - \alpha \quad (8.10)$$

Donde $C_{\alpha/2}$ es la desviación normal estándar correspondiente al porcentaje $(1 - \alpha/2)$. La desviación normal estándar es calculada usando la Función de Densidad Acumulativa (CDF o FDA) .

A.8.1.6 TÉCNICAS SPM UNIVARIANTES

Las técnicas SPM usadas para monitorear una sola variable incluye gráficas Shewhart, suma acumulada (CUSUM), promedio móvil (MA), y promedio móvil exponencialmente ponderado (EWMA).

- **Gráficas de control Shewhart y Técnicas SPM univariantes**

Las gráficas Shewhart indican que una especial (asignable) causa de variación es presentada cuando los puntos de la muestra ploteada está fuera de los límites de control, una *prueba gráfica de hipótesis* es realizada mediante el ploteo de media de la muestra, y el rango de desviación estándar; y comparándolos con sus límites de control. Una gráfica Shewhart es diseñada especificando la *línea central* (CL), el límite de control superior (UCL) y el límite de control inferior (LCL)

Dos gráficas shewhart (media de muestra y desviación estándar o el rango) son ploteadas simultáneamente. La media de al muestra es inspeccionada para evaluar variaciones entre las muestras (variabilidad del proceso con el tiempo) ploteando la media de la gráfica de Shewhart (*\bar{x} gráfica*, \bar{x} representa la media de x). Sin embargo, se tiene que asegurar que no hay cambios significativos dentro de las variaciones de las muestras que puedan dar una impresión errónea entre las variaciones de las muestras. Los valores medios en momentos $k-2$ y $k-1$ en la figura 71. lucen similares pero dentro de las variaciones de muestra dentro de momento $k-1$ es significativamente diferente que la dada en la muestra en el momento $k-2$. Por lo tanto, es engañoso afirmar que el estado entre las variaciones de la muestra es despreciable y el nivel del proceso es constante t.

El *gráfico de Rango* (*gráfico R*), o el gráfico de la desviación estándar, (*gráfico S*) es usado para monitorear en *muestras* de variaciones del proceso o propagación (variabilidad del proceso en un tiempo dado). La propagación del proceso debe ser

en-control para una interpretación adecuada de la gráfica \bar{x} . La gráfica \bar{x} debe ser usada junto a la gráfica de propagación.

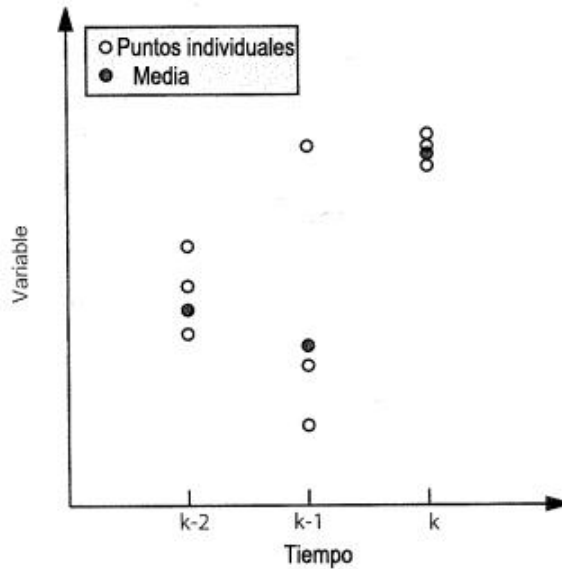


Figura 71. Diagrama punto de observaciones individuales de una variable

Fuente: *Chemical Process Performance Evaluation*. Heinz Heinemann.

Las suposiciones de las gráficas chart son:

- La distribución de los datos es aproximadamente normal.
- El tamaño de los grupos de muestra son iguales.
- Todas las muestras de grupo son ponderadas por igual.
- Las observaciones son independientes.

Un enfoque univariante estadístico para limitar detección-sensores puede ser usado para determinar los umbrales para cada **observación de variables** (una variable de proceso observada mediante la lectura de un sensor), donde dichos umbrales definen el límite para operaciones control-adentro y una violación de

dichos límites con datos on-line indicarán una falla. Este enfoque es típicamente empleado usando **Los Gráficos de control Shewhart** ver figura 72. y ha sido referida ha sido denominado como **límite de detección-sensor** y **Limite valor comprobado**. Los valores de los límites de control superiores e inferiores en los gráficos Shewhart son críticos para minimizar la rata de **Falsas alarmas** y la rata de **detecciones pérdidas**. Una **falsa alarma** es un indicador de falla, cuando en realidad una falla no ha ocurrido; una **detección pérdida** no es una indicación de falla, aunque una falla haya ocurrido. Para la detección de fallas, existe un equilibrio inherente entre minimizar las falsas alarmas y ratas de detecciones pérdidas. Umbrales con límites estrechos para una variable en observación resulta en altas falsas alarmas y pequeñas ratas de detecciones pérdidas, mientras límites que son muy separados resultan en una rata baja de falsas alarmas y detecciones pérdidas altas.

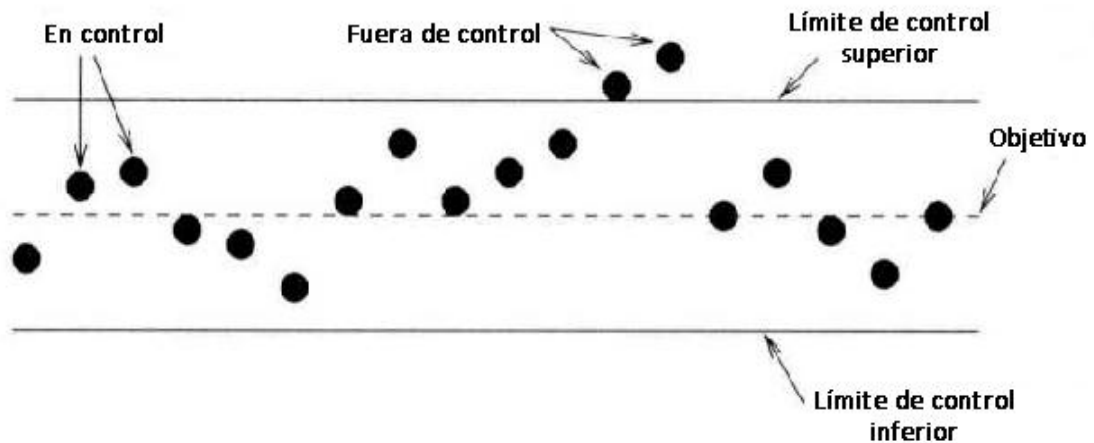


Fig 72. Ilustración de la gráfica de control de Shewhart. Los puntos negros son observaciones.

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

Los esquemas de monitoreo de procesos basados en gráficos de Shewhart pueden no proveer adecuadas tasas de falsas alarmas y detecciones fallidas. Dichas tasas pueden ser mejoradas mediante mediciones que incorporen observaciones de múltiples instancias de tiempo consecutivas, tales como cartas **suma acumulativa** (CUSUM) y **promedio móvil exponencialmente-pesado** (EWMA). Para una tasa de falsa alarma dada. Dichos métodos pueden incrementar la sensibilidad a fallas sobre la mediciones usando los gráficos de Shewhart y, en consecuencia decreciendo la tasa de pérdidas de detección, pero a expensas de detección retrasada, que es el la cantidad de tiempo gastada entre el comienzo de la falla y el tiempo de detección. Esto sugiere que el las gráficas CUSUM y EWMA se adaptan mejor en fallas que producen cambios pequeños persistentes del proceso, y los gráficos de Shewhart son mejores para la detectar fallas que producen grandes cambios súbitos.

Los gráficos estadísticos univariantes (Shewhart, CUSUM, y EWMA) determinan los umbrales para cada variables de observación individualmente sin considerar la información contenida en otras variables.

A.8.2 ESTADÍSTICA MULTIVARIABLE

En los últimos años, el control estadístico de procesos multivariados se ha convertido en una herramienta extremadamente útil para el control simultáneo de varias características de calidad relacionadas.

Estadísticas multivariantes es un término general para una serie de diferentes métodos estadísticos. La característica definitoria de estos métodos es que tienen

el objetivo de entender un conjunto de datos al considerar un grupo de variables juntas en lugar de centrarse en una sola variable a la vez.

El análisis multivariante es la parte de la estadística y del análisis de datos que estudia, analiza, representa e interpreta los datos que resulten de observar un número $p > 1$ de variables estadísticas sobre una muestra de n individuos. Las variables observables son homogéneas y correlacionadas, sin que alguna predomine sobre las demás. La información estadística en análisis multivariante es de carácter multidimensional, por lo tanto la geometría, el cálculo matricial y las distribuciones multivariantes juegan un papel fundamental. La información multivariante es una matriz de datos, pero a menudo, en el análisis multivariante la información de entrada consiste en matrices de distancias o similitudes, que miden el grado de discrepancia entre los individuos. Comenzaremos con las técnicas que se basan en matrices de datos.

A.8.2.1 ESTADÍSTICO DE HOTELLING (T^2)

La siguiente discusión se da una idea de cómo el concepto de distancia estadística, ya que define con el estadístico T^2 , se utiliza en el desarrollo de los procedimientos de control multivariable. Supongamos que denotamos una observación multivariante de las variables p en forma vectorial como $X' = (x_1, x_2, \dots, x_p)$. Nuestra principal preocupación es en el procesamiento de la información disponible sobre estas variables p . Un enfoque es utilizar las técnicas gráficas, que suelen ser excelentes para esta tarea, pero los puntos de trazado en un espacio p -dimensional ($p > 3$) es muy limitada. Esta restricción impide la visualización global de la situación multifactorial. Otro método para examinar la información proporcionada en una observación p -dimensional es reducir el vector

de datos multivariados para una estadística univariante sola. Si la estadística resultante contiene información sobre todas las variables p , puede ser interpretada y utilizada en la toma de decisión es en cuanto a la situación de un proceso.

Supongamos que un proceso genera observaciones bivariantes no correlacionadas, (x_1, x_2) , y se desea que los represente gráficamente. Es común para construir un gráfico de dispersión de dos dimensiones de los puntos. Además, supongamos que hay un interés en la determinación de la distancia un punto en particular es desde el punto medio. La distancia entre dos puntos se mide siempre como un solo número valor. Esto es cierto independientemente de cuántas dimensiones (variables) están involucradas en el problema. La habitual línea recta (euclidiana) de distancia mide la distancia entre dos puntos por el número de unidades que los separan. Cuadrado de la distancia línea recta, llamada D , entre un punto (x_1, x_2) y la media población al punto $(\mu_1; \mu_2)$ se define como:

$$(X_1 - \mu_1)^2 + (X_2 - \mu_2)^2 = D^2 \quad (8.11)$$

Note que hemos tomado la observación bivalente $(x_1; x_2)$, y convirtiéndola a un número sencillo D , la distancia de observación es desde el punto promedio si esa distancia D , es construida, todos los puntos tienen la misma distancia desde el punto promedio, pueden ser representados como un círculo con centro en el punto promedio y radio D . también cualquier punto localizado dentro del círculo tiene una distancia al punto promedio menor que D .

Infortunadamente para determinar esa distancia no se tiene en cuenta la consideración de las diferencias en la variación de las dos variables como los

valores de sus varianzas. Para corregir estas deficiencias se consideran los valores estandarizados.

$$\frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} = SD^2 \quad (8.12)$$

El valor de SD es conocido como distancia estadística. Para una construcción del valor de SD todos los puntos satisfacen que tienen la misma distancia estadística desde el punto promedio. Supongamos que (x_1, x_2) se puede describir en forma conjunta por una distribución normal bivariada. Bajo este supuesto, la distancia estadística entre este punto y el vector de medias (μ_1, μ_2) es el valor de la parte variable del exponente de la función de probabilidad normal bivariada

$$f(x_1, x_2) = \left[\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-p^2}} \right] e^{-\frac{(SD)^2}{2}} \quad (8.13)$$

Donde $-\infty < x_i < \infty$, para $i = 1, 2$, y $\sigma_i > 0$ representa la desviación estándar de x_i . El valor de SD^2 está dada por:

$$\frac{1}{(1-p^2)} \left[\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - 2p \left(\frac{x_1 - \mu_1}{\sigma_1} \right) \left(\frac{x_2 - \mu_2}{\sigma_2} \right) + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right] = SD^2 \quad (8.14)$$

Donde ρ representa la correlación entre las dos variables con $-1 \leq \rho \leq 1$.

Ecuaciones para el contorno de una densidad normal bivariada se obtienen mediante la fijación del valor de SD. Esto se puede ver geoméricamente mediante

el examen de la función de dos variables de probabilidad normal se presenta en la Figura 73.

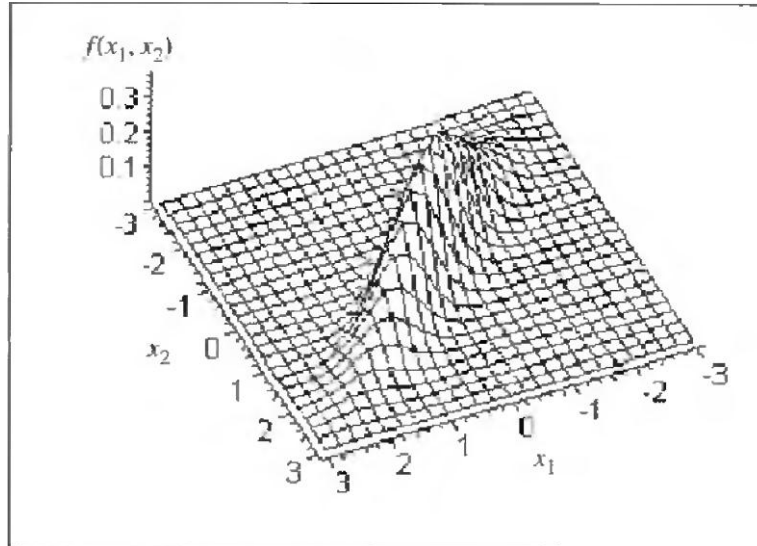


Figura.73. Una función de dos variables de probabilidad normal $\sigma_1=1$, $\sigma_2=1$, $\rho=0,8$.

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

El lugar o la ruta, del punto $X'=(x_1, x_2)$ que viajan alrededor de la función de probabilidad a una altura constante es una elipse. Elipses de densidad constante se conocen como curvas de nivel y se pueden determinar matemáticamente para contener una cantidad fija de probabilidad. Por ejemplo, los contornos de un 75% y 95% para la función normal bivariada ilustra en la Figura 75. Los contornos elípticos representan todos los puntos que tienen la misma distancia estadística o el valor estadístico T^2 .

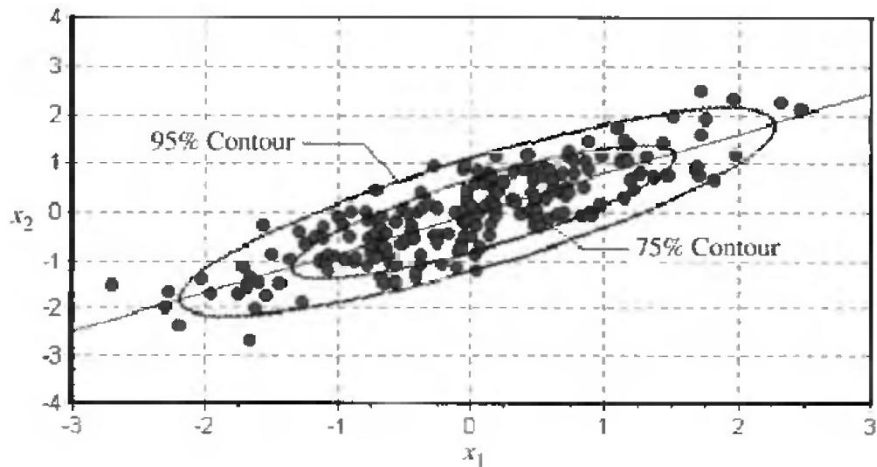


Figura. 74. Un Bivariado contorno normal contiene un 75% y el 95% de la probabilidad

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

A.8.2.2 DESARROLLO DEL MÉTODO ESTADÍSTICO HOTELLING

Observamos que los datos en el conjunto de entrenamiento, que consisten en la observación de las variables m y n observaciones para cada variable, se apilan en una matriz de $X \in \mathbb{R}^{n \times m}$

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \quad (8.15)$$

Entonces la covarianza de matriz de la muestra del conjunto de entrenamiento es igual a:

$$S = \frac{1}{n-1} X^T X. \quad (8.16)$$

Una descomposición de valores propios de la matriz S

$$S = V\Lambda V^T \quad (8.17)$$

Revela la correlación estructural de la matriz de covarianza, donde Λ es diagonal y V es ortogonal ($V^T V = I$, donde I es la matriz identidad) La proyección $\mathbf{y} = V^T \mathbf{x}$ de un vector de observación $\mathbf{x} \in \mathbb{R}^m$ separa el espacio de observación en un conjunto de variables correlacionadas correspondientes a elementos de \mathbf{y} . La varianza del i -ésimo elemento de \mathbf{y} es igual al i -ésimo eigen valor en la matriz Λ . Asumiendo que \mathbf{S} es invertible, y teniendo en cuenta la definición:

$$\mathbf{Z} = \Lambda^{-1/2} V^T \mathbf{X} \quad (8.18)$$

Entonces la estadística de la T^2 de Hotelling viene dada por:

$$T^2 = \mathbf{Z}^T \mathbf{Z} \quad (8.19)$$

La matriz V rota los ejes principales de la matriz de covarianza de \mathbf{x} para que se correspondan directamente con los elementos de \mathbf{y} , y Λ escala de los elementos de \mathbf{y} para producir un conjunto de variables con variación de la unidad correspondiente a los elementos de \mathbf{z} . La conversión de la matriz de covarianza se demuestra gráficamente en la Figura 75. para un espacio de observación de dos dimensiones ($m = 2$)

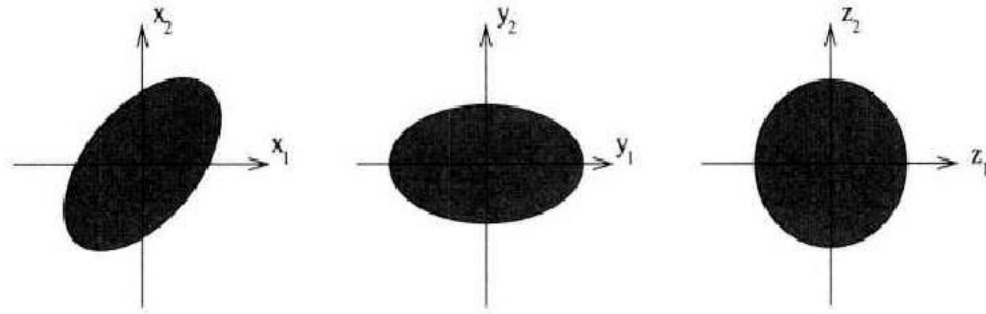


Figura 75. Ilustración grafica de la conversión de la covarianza para la estadística T^2

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

A.8.2.3 LOS UMBRALES PARA LA ESTADÍSTICA T^2

Umbral adecuado para el estadístico T^2 en función del nivel de importancia, α , se puede determinar si se asume que las observaciones son una muestra aleatoria de una distribución normal multivariante. Si se supone además que la media muestral del vector y matriz de covarianza para las operaciones normales son iguales al vector de medias reales y la matriz de covarianza respectivamente, entonces el estadístico T^2 sigue una distribución χ^2 con m grados de libertad

$$T^2_{\alpha} = \chi^2_{\alpha}(m) \quad (8.20)$$

El conjunto de $T^2 < T^2_{\alpha}$, es una región de confianza elíptica en el espacio de observación, como se ilustra en la Figura 76. para dos variables de proceso de $m=2$, la aplicación de la Ecuación anterior; a los datos del proceso produce una región de confianza definida como de control, mientras que un vector de observación proyectado fuera de esta región indica que existe un fallo.

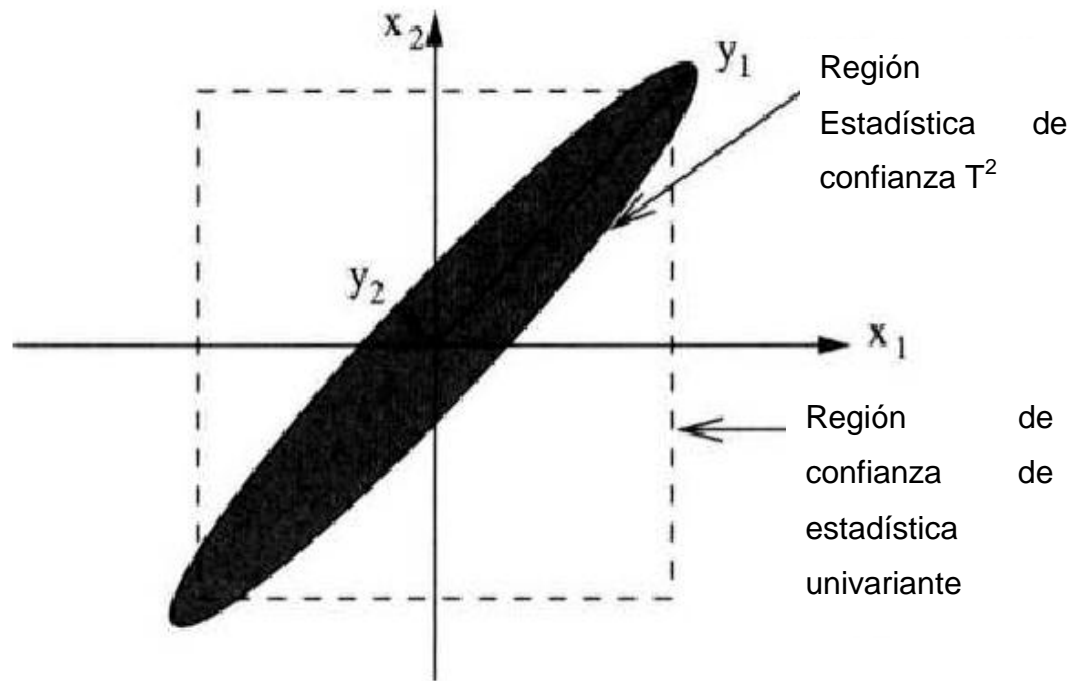


Figura 76. Una comparación de la región en estado de control usando estadística T^2 y la estadística univariante para dos variables de proceso.

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

Dado un nivel de significación, α , la Figura 76. ilustra el conservatismo eliminado mediante el empleo de la estadística de T^2 en comparación con el enfoque estadístico univariante. A medida que el grado de correlación entre las variables de proceso aumenta, la región de confianza se hace más elíptica alargada y por tanto reduce el porcentaje de conservación mediante el uso de las estadísticas T^2 . Cuando la matriz de covarianza real para el estado en el control no se conoce, entonces es estimada a partir de la matriz de covarianza de la muestra \mathbf{S} , los fallos pueden ser detectados por observaciones realizadas fuera del conjunto de entrenamiento utilizando el umbral dado por:

$$T^2_{\alpha} = \frac{m(n-1)(n+1)}{n(n-m)} F_{\alpha}(m, n-m) \quad (8.21)$$

Donde $F_{\alpha}(m, n-m)$ es el superior $100\alpha\%$, punto crítico de la distribución F con m y $n-m$ grados de libertad.

A.8.2.4 REQUERIMIENTOS DE DATOS.

La calidad y cantidad de los datos en el conjunto de entrenamiento tienen una gran influencia en la efectividad del estadístico T^2 , como una herramienta en el monitoreo de procesos. Una cuestión importante concerniente con el campo de entrenamiento es ¿cuántos datos se necesitan para aplicar la estadística poblacional de covarianza para m variables de observación? Esta pregunta es respondida aquí por determinación de la acumulación de los datos necesarios para producir un valor umbral suficientemente cercano al umbral obtenido si hubiésemos obtenido datos infinitos en campo de entrenamiento. Para dar un nivel de significancia α , un umbral basado en infinitas observaciones en el campo de entrenamiento o equivalentemente un conocimiento exacto de la matriz de covarianza puede ser computado usando la ecuación (8.21) y el umbral para n observaciones en el campo de entrenamiento es calculado usando la ecuación (8.20) el error relativo producido por estos dos valores de umbral es:

$$\epsilon = \frac{\frac{m(n-1)(n+1)}{n(n-m)} F_{\alpha}(m, n-m) - X^2_{\alpha}(m)}{X^2_{\alpha}(m)} \quad (8.22)$$

La tabla siguiente muestra los requerimientos de datos para la ecuación 8.22 para varios números de variables observaciones.

Tabla 4. Requerimientos de datos para la ecuación 8.22

Número de variables de observación	Requerimientos de datos n
1	19
2	30
3	41
4	52
5	63
10	118
25	284
50	559
100	1110
200	2210

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

El estadístico de Hotelling (T^2) es muy utilizado en Control Estadístico de Procesos para detectar si el proceso se encuentra fuera de control, pero no ofrece asistencia como herramienta de identificación. Para tal fin se han propuesto diferentes técnicas que calculan la contribución de cada variable al valor del estadístico cuando éste supera el correspondiente valor crítico. Algunas estrategias operan en el espacio definido por las variables originales y otras emplean variables latentes.

La técnica utiliza el estadístico de Hotelling para detectar la presencia de biases (errores sistemáticos de magnitud constante). El método del vecino más próximo en condiciones de control estadístico se emplea para identificar los sensores defectuosos. El procedimiento se basa exclusivamente en el uso de datos

históricos, por lo tanto no se requiere el modelado del proceso, ni tampoco conocer la función de densidad de probabilidad de los errores de las mediciones. En contraposición con los métodos basados en principios fundamentales, es posible identificar errores sistemáticos en todas las mediciones, sin la necesidad de distinguir entre las redundantes y no redundantes.

A.9. MÉTODOS USADOS CON DATA DRIVEN

A.9.1 ANÁLISIS DE COMPONENTES PRINCIPALES

A.9.1.1 INTRODUCCIÓN

Para la proyección de los datos dentro de un espacio de pocas dimensiones que aproxime las características del estado del proceso, técnicas de reducción de dimensiones pueden simplificar enormemente y mejorar los procedimientos de monitoreo de proceso. El Análisis de Componentes Principales [PCA]²¹ es una de las técnicas de reducción de dimensiones. Esta produce una representación de menores dimensiones preservando la estructura de correlación entre el proceso y las variables, y es óptima en términos de captura de variabilidad de los datos.

La aplicación de PCA como una herramienta de reducción dimensional para el proceso de monitoreo industrial ha sido estudiada por varios ingenieros industriales y académicos. Para algunas aplicaciones, la gran variabilidad de los datos puede ser capturada en 3 dimensiones, y la variabilidad del proceso puede ser visualizada con una simple gráfica. Esta gráfica de visualización y la estructura extraída de los datos multidimensionales ayuda a los operadores e ingenieros en la interpretación de tendencias significativas de los datos del proceso.

²¹ Por sus siglas en inglés, Principal Component Analysis.

Para los caso en que la mayoría de las variaciones de los datos no pueden ser capturada en dos o tres dimensiones, se han desarrollado métodos para automatizar el procedimiento de monitoreo de procesos. La aplicación de PCA en estos métodos es motivado por uno o más de tres factores. *Primero*, PCA puede producir representaciones de menores dimensiones de los datos, las cuales son la mejor generalización de los datos independientes del conjunto de entrenamiento que utilizar toda la dimensión del espacio de observación, y por lo tanto, mejoran la habilidad de detectar y diagnosticar fallas. *Segundo*, la estructura abstracta de PCA puede ser útil en la identificación de cualquier variable responsable de falla y/o la mayoría de las variables afectadas. *Tercero*, PCA puede separar la observación espacial dentro de un sub-espacio que capture las tendencias sistemáticas de los procesos y un sub-espacio que contenga esencialmente el ruido aleatorio. Dado que es ampliamente aceptado que ciertas fallas afectan sobre todo a uno de los dos sub-espacios, aplicando una medida desarrollada para uno de los sub-espacios y otra para el otro, puede incrementarse la sensibilidad del proceso de monitoreo para las fallas en general. Los tres atributos mencionados para PCA serán discutidos a profundidad más adelante en este capítulo.

El propósito de este capítulo es el de describir los métodos de PCA para el proceso de monitoreo. En la sección 9.1.2 se definirá PCA y en la sección 9.1.3 se discutirán los diferentes métodos que serán usados para determinar el orden de la representación de PCA. En las secciones 9.1.4, 9.1.5 y 9.1.6 se discutirá sobre el desarrollo de PCA para la detección, identificación, diagnóstico, respectivamente.

A.9.1.2 ANÁLISIS DE COMPONENTES PRINCIPALES

PCA es una técnica lineal de reducción de dimensionalidad, óptima en términos de captura de la variabilidad de los datos. Esta determina un conjunto de vectores ortogonales, llamados vectores de carga, regidos por la cantidad de variables explicadas en las direcciones del vector de carga. Dado un conjunto de entrenamiento de n observaciones y m variables de proceso apilado dentro de una matriz X como en (9.1),

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdot & x_{1m} \\ x_{21} & x_{22} & \cdot & x_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{n1} & x_{n2} & \cdot & x_{nm} \end{bmatrix}; \quad (9.1)$$

los **vectores de carga** son calculados resolviendo los puntos estacionarios del problema de optimización

$$\max_{V \neq 0} \frac{V^T X^T X V}{V^T V} \quad (9.2)$$

Donde $V \in \mathbb{R}^m$. Los puntos estacionarios de (9.2) pueden ser computados por medio de descomposición singular de valores (SVD, por sus siglas en inglés)

$$\frac{1}{\sqrt{n-1}} X = U \Sigma V^T \quad (9.3)$$

donde $U \in \mathbb{R}^{n \times n}$ y $V \in \mathbb{R}^{m \times m}$ y son matrices unitarias, y la matriz $\Sigma \in \mathbb{R}^{n \times m}$ contiene los **valores singulares** reales no negativos de magnitud decreciente a lo largo de

la diagonal ($\sigma_1 \geq \sigma_2 \geq \sigma_3 \dots \dots \geq \sigma_{\min(m,n)} \geq 0$), y cero los elementos fuera de la diagonal. Los vectores de carga son ortonormales a los vectores columna de la matriz V , y la varianza de los datos de entrenamiento proyectados a lo largo de la i^{th} columna de V es igual a la σ_i^2 . Resolver la ecuación (9.3) es equivalente a encontrar la descomposición de eigenvalores de la matriz de covarianza S ,

$$S = \frac{1}{n-1} X^T X = V \Lambda V^T \quad (9.4)$$

donde la diagonal de la matriz $\Lambda = \Sigma^T \Sigma \in \mathbb{R}^{m \times m}$ contiene los **eigenvalores** reales no negativos de magnitud decreciente ($\lambda_1 \geq \lambda_2 \geq \lambda_3 \dots \dots \geq \lambda_m \geq 0$) y el i^{th} eigenvalor es igual al cuadrado del i^{th} valor singular ($\lambda_i = \sigma_i^2$).

Para la óptima captura de la variabilidad de los datos mientras se minimiza el efecto del ruido que corrompe la representación PCA, los vectores de carga correspondientes al valor singular más grande normalmente son retenidos. La motivación de reducir la dimensionalidad de la representación PCA es análoga al argumento dado en la sección 8.1.3 para patrones de clasificación. Seleccionando las columnas $P \in \mathbb{R}^{m \times a}$ de la matriz de carga para el correspondiente vector de carga asociado con el primer valor singular α , las proyecciones de las observaciones en X dentro de un espacio de menores dimensiones en la **matriz score**,

$$T = XP \quad (9.5)$$

Y la proyección de T dentro de un espacio de observación de m dimensiones,

$$\hat{X} = TP^T \quad (9.6)$$

La diferencia entre X y \hat{X} es la matriz residual E :

$$E = X - \hat{X} \quad (9.7)$$

La **matriz residual** captura la variación en el espacio de observación abarcado por vector de carga asociado con los $m - \alpha$ valores singulares más pequeños. Los sub-espacios abarcados por \hat{X} y E son llamados **espacios score** y **espacios residuales**, respectivamente. El sub-espacio contenido en la matriz E tiene una pequeña relación entre la señal y el ruido, y el remover este espacio de X puede producir una mejor aproximación de la representación del proceso.

Definiendo t_i como la i^{th} columna de T en el conjunto de entrenamiento, las siguientes propiedades pueden ser mostradas:

1. $\text{Var}(t_1) \geq \text{Var}(t_2) \geq \dots \geq \text{Var}(t_a)$.
2. $\text{Media}(t_i) = 0; \forall i$.
3. $T_i^T t_k = 0; \forall i \neq k$.
4. No existen ninguna otra expansión ortogonal de los componentes α que capturen mayor variación de los datos.

Un nuevo vector observación (columna) en el conjunto de datos de prueba, $x \in \mathbb{R}^m$, puede ser proyectado dentro de un espacio score de menores dimensiones $t_i = x^T p_i$, donde p_i es el i^{th} vector de carga (ver figura 77). La variable transformada t_i es también llamada **componente principal** de x . Para distinguir entre las variables transformadas y la observación transformada, las variables transformadas se llamaran **componentes principales** y la observaciones transformadas individuales serán llamadas **scores**. Las propiedades estadísticas nombradas arriba permiten que cada uno de los scores sea monitoreado por separado usando el procedimiento de estadística univariable discutido en la sección 8.1.3. Con los vectores proyectados dentro de un espacio de bajas dimensiones usando PCA, solamente son necesarias α variables para monitorear comparado con m variables necesarias si no se usará PCA. Cuando son

suficientes los datos recolectados en el conjunto de prueba se podrán formar los vectores score t_1, t_2, \dots, t_a . Si estos vectores no satisfacen las cuatro propiedades de la lista de arriba, el conjunto de datos fue tomado en su mayoría en condiciones de operación diferentes a las del conjunto de entrenamiento. Esta abstracción de la estructura de los datos multidimensionales es un componente clave de la contribución del método para la identificación de fallas discutida en la sección 9.1.5.

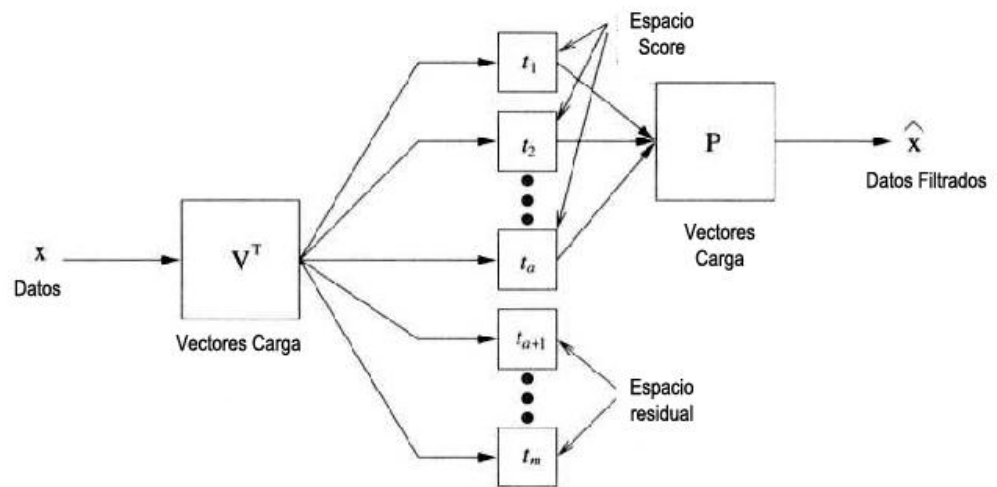


Figura 77. La proyección del vector de observación x dentro del espacio score y residual, y la computación de la observación filtrada \hat{x} .

Fuente: Fuente: Fault Detection and Diagnosis in Industrial Systems. L.H Chiang, E.L Russell and R.D. Braatz.

La aplicación de PCA será ilustrada usando los datos Fisher de clases de lirios.²²El conjunto de datos consiste en tres clases, donde cada clase contiene $m = 4$ medidas y $n = 50$ observaciones (ver Tabla 9.1 y Figura 78).

²²R.A. Fisher. The use of multiple measurements in taxonomic problems. *Ann. Eugenics*, 7:179-188, 1936.

Y.Q. Cheng, Y.M. Zhuang, an J. Y Yang. Optimal Fiseher discriminant analysis using the Rank decomposition. *Pattern Recognition*, 25:101-111, 1992.

Los datos de la clase 3 fueron usados para construir la matriz X como en (9.1). Después de auto-escalar²³ X y resolver (9.4), tenemos:

$$\Lambda = \begin{bmatrix} 1.92 & 0 & 0 & 0 \\ 0 & 0.96 & 0 & 0 \\ 0 & 0 & 0.88 & 0 \\ 0 & 0 & 0 & 0.24 \end{bmatrix} \quad (9.8)$$

y,

$$V = \begin{bmatrix} 0.64 & -0.29 & 0.052 & -0.71 \\ 0.64 & -0.23 & 0.25 & 0.69 \\ 0.34 & 0.33 & -0.88 & 0.11 \\ 0.25 & 0.87 & 0.42 & -0.09 \end{bmatrix} \quad (9.9)$$

La varianza total de X proyectada a lo largo de V es igual a la traza de Λ , la cual es 4. El i^{th} valor en la diagonal de Λ indica la cantidad de varianza capturada por el i^{th} componente principal. Si solo un componente es retenido (por ejemplo, $\alpha = 1$), entonces $(1.92/4) * 100\% = 48\%$ del total de la varianza es capturada. Para $\alpha = 2$, el 72% de la varianza total es capturada. Entonces tomando $\alpha = 2$, la matriz de carga P es igual a las primeras dos columnas de V:

²³ Es común cuando se trabaja con variables que se expresan en unidades distintas que los valores difieran demasiado en magnitud, lo cual afecta significativamente la interpretación de los resultados. Por esta razón es apropiado estandarizar o auto-escalar los datos, lo que significa que a cada valor se le resta la media y se divide en la desviación estándar correspondiente, lo que nos da un nuevo conjunto de datos con media cero y varianza igual a 1.

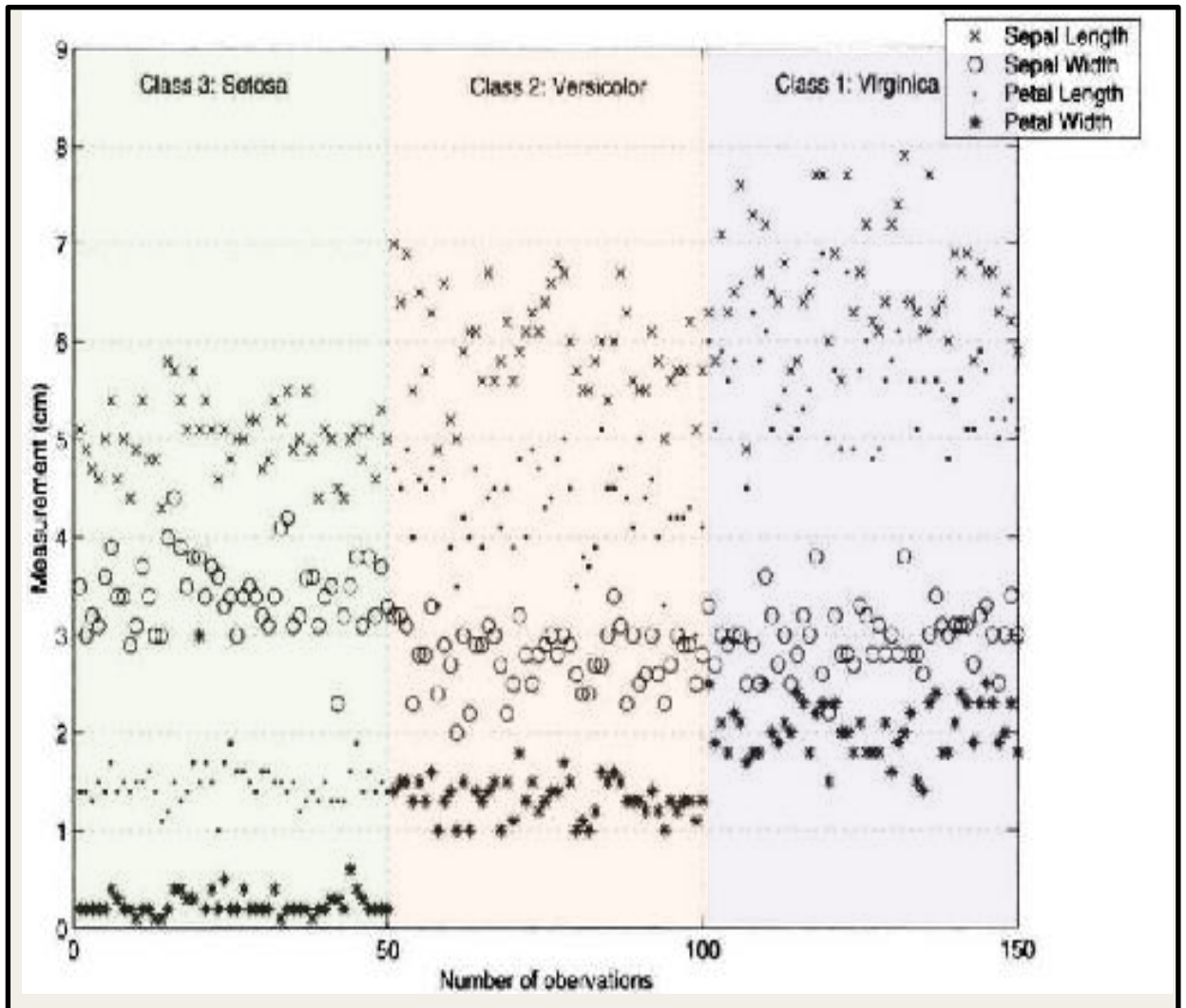


Figura 78 Gráfica de los datos de Fisher

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

Tabla 5 Estadística de los datos Fisher.

Clase 1: Lirio Virginica	Media	Desviación Estándar	Rango
Longitud de Sépalo	6.59	0.64	4.9-7.9
Ancho de Sépalo	2.98	0.32	2.2-3.8
Longitud de Pétalo	5.55	0.55	4.5-6.9
Ancho de Pétalo	2.03	0.27	1.4-2.5
Clase 2: Lirio Versicolor	Media	Desviación Estándar	Rango
Longitud de Sépalo	5.94	0.52	4.9-7.0
Ancho de Sépalo	2.77	0.31	2.0-3.4
Longitud de Pétalo	4.29	0.47	3.0-5.1
Ancho de Pétalo	1.33	0.20	1.0-1.8
Clase 3: Lirio Setosa	Media	Desviación Estándar	Rango
Longitud de Sépalo	5.01	0.35	4.3-5.8
Ancho de Sépalo	3.43	0.38	2.3-4.4
Longitud de Pétalo	1.46	0.17	1.0-1.9
Ancho de Pétalo	0.3	0.40	0.1-3.0

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

$$P = \begin{bmatrix} 0.64 & -0.29 \\ 0.64 & -0.23 \\ 0.34 & 0.33 \\ 0.25 & 0.87 \end{bmatrix} \quad (9.10)$$

La matriz score T es calculada de acuerdo a (9.5). La ventaja de retener solamente dos componentes principales es que la variabilidad del proceso puede ser visualizada por graficando t_2 versus t_1 (ver Figura 9.3).

Es fácil de verificar que $\text{Var}(t_1) \geq \text{Var}(t_2)$ observando que la variación a lo largo del eje horizontal es mucho mayor que la variación en el eje vertical para los datos de la clase 3 en la figura 79. La elipsoide y los datos para la clase 3 están centrados en el origen, lo cual indica que la media $(t_1) = \text{media}(t_2) = 0$. Es sencillo verificar que t_1 y t_2 son ortogonales el uno del otro.

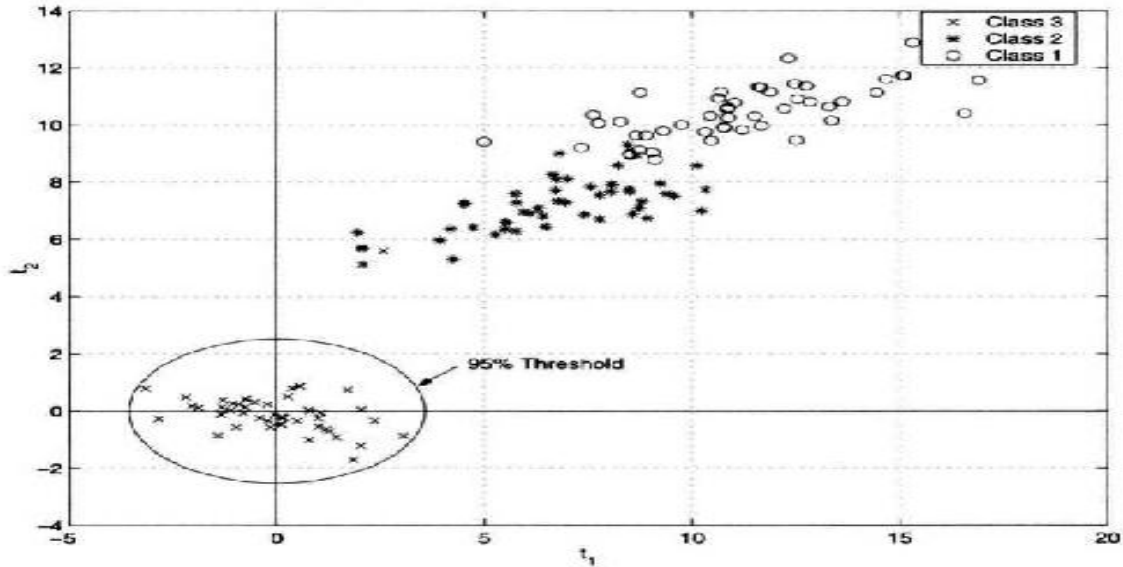


Figura 79. Las proyecciones de los datos Fisher para las tres clases sobre los dos primeros vectores de carga PCA.

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

Un umbral define una región de confianza elíptica para los datos que pertenecen a la clase 3 (el cálculo de este umbral será descrito en la sección 9.1.4). En este ejemplo, la estadística predice que hay un 95% de probabilidad que los datos de la clase 3 estén dentro de la elipse. Se muestra claramente en la figura 9.3 que PCA es capaz de separar la clase 3 de la Clase 1 y 2, excepto por el punto $(t_1, t_2) = (2.5, 5.6)$.

A.9.1.3 REDUCCIÓN DE ORDEN

Es comúnmente aceptado y con varias asunciones teóricamente justificado que la porción del espacio PCA correspondiente a los valores singulares más grandes describen la mayoría de las *variaciones sistemáticas o estables* que ocurren en un proceso, y la porción del espacio PCA correspondiente a los valores singulares más pequeños describe un *ruido aleatorio*. Para determinar apropiadamente el número de vectores de carga, α , que se deben mantener en el modelo PCA, las variaciones sistemáticas pueden ser desacopladas de las variaciones aleatorias, y los dos tipos de variaciones pueden ser monitoreados separadamente, como se discutirá en la sección 9.1.4. Existen varias técnicas para determinar el valor de la reducción de orden α , pero parece que no hay una técnica dominante. Los métodos para determinar α descritos aquí son:

1. La prueba de variación porcentual
2. Prueba de acumulación
3. Análisis paralelo, y
4. El estadístico PRESS.

El método de la **variación porcentual** determina α calculando el número más pequeño de vectores de carga necesarios que correspondan a un porcentaje mínimo específico del total de la varianza. (Recordemos que la varianza asociada con el i^{th} vector de carga es igual a el cuadrado del valor singular, σ_i^2 .) Debido a que este porcentaje mínimo es escogido arbitrariamente, este puede ser tan alto o tan bajo como se requiera en una aplicación particular.

La **prueba de acumulación** asume que la varianza, σ_1^2 , correspondiente a un ruido aleatorio forma un perfil lineal. La dimensión del espacio score α es determinado localizando σ_1^2 donde el perfil no es tan lineal. La identificación de este cambio puede ser ambigua, y así, este método es difícil de automatizar. Especialmente cuando varios cambios de la linealidad ocurren en el perfil.

El **análisis paralelo** determina la dimensionalidad por la comparación entre el perfil de varianza y el obtenido asumiendo variables de observación independientes. La reducción de orden es determinada como el punto en el cual los dos perfiles se cruzan. Este enfoque asegura que las correlaciones significativas son capturadas en el espacio score, y es particularmente atractivo desde el punto de vista intuitivo y es fácil de automatizar.

La dimensión del espacio score puede ser determinada usando un procedimiento de **validación cruzada** con la **Predicción acumulada de los residuos de los cuadrados (PRESS)**,

$$PRESS(i) = \frac{1}{mn} \|X - \hat{X}\|_F^2 \quad (9.11)$$

donde i es el número de vectores de carga retenidos para el cálculo de X y $\|X - \hat{X}\|_F^2$ (la raíz cuadrada de la suma de los cuadrados de todos los elementos). Para la implementación de esta técnica, el conjunto de entrenamiento es dividido en dos grupos. La estadística PRESS para un grupo es computado basado en varias dimensiones del espacio score, i , usando todos los otros grupos. Esto es repetido para cada grupo, y el valor i asociado con el promedio mínimo de la estadística PRESS determina la dimensión del espacio score.

A.9.1.4 DETECCIÓN DE FALLAS

Como se discutió en la sección 8.2, la estadística T^2 conocida puede ser usada, el umbral estadístico T^2 (8.20) es

$$T_a^2 = X_a^2(a) \quad (9.14)$$

Cuando la matriz de covarianza real es estimada de la matriz de covarianza de ejemplo, el umbral T^2 estadístico de (8.21) es

$$T_\alpha^2 = \frac{a(n-1)(n+1)}{n(n-a)} F_\alpha^2(a, n-a) \quad (9.15)$$

Para detectar los puntos fuera del conjunto de entrenamiento, el umbral de (8.22) es

$$T_\alpha^2 = \frac{(n-1)^2 \left(\frac{a}{n-a-1}\right) F_\alpha(a, n-a-1)}{n \left(1 + \left(\frac{a}{n-a-1}\right)\right) F_\alpha(a, n-a-1)} \quad (9.16)$$

Debido a que la estadística T^2 en (9.13) no se afecta por las inexactitudes en los pequeños valores singulares de la matriz de covarianza, es capaz de representar el comportamiento del proceso normal y medida aproximada de la detección de falla cuando se compara la estadística T^2 en (9.12). Usando los argumentos en la sección 9.1.3, la estadística T^2 (9.13) puede ser interpretada como la medida de las variaciones sistemáticas del proceso, y una violación del umbral podría indicar que las variaciones están fuera de control.

Por ejemplo en la última sección, tenemos $n= 50$ y $a= 2$. De acuerdo con la tabla de distribución F^{24} , $F_{0,05}(2,48) = 3.19$. El umbral $T\alpha$ es igual a 6.64 de acuerdo con (9.14). La región de confianza elíptica, como muestra la figura 79, está dada por

$$T^2 = x^T P \Sigma_{\alpha}^{-2} P^T x \leq 6.64, \quad (9.17)$$

Con

$$\Sigma_{\alpha}^2 = \begin{bmatrix} 1.92 & 0 \\ 0 & 0.96 \end{bmatrix} \quad (9.18)$$

La ecuación

$$t = P^T x \quad (9.19)$$

ubica esta región dentro de la elipse en la figura 9.3.

Insertando (9.19) en (9.17) da

$$t^T \Sigma_{\alpha}^{-2} t \leq 6.64 \quad (9.20)$$

o

$$\frac{t_1^2}{1.92} + \frac{t_2^2}{0.96} \leq 6.64 \quad (9.21)$$

donde $t = [t_1 \ t_2]^T$.

Los datos de la clase 1 y 2 son usados para ilustrar que el modelo PCA es capaz de detectar los datos que no pertenecen a la clase 3. El conjunto de datos de la clase 1 y 2 primero se auto-escalán de acuerdo a la media y la desviación

²⁴ W.W. Hines and D. C. Montgomery. Probability and Statistics in Engineering and Management Science. John Wiley & Sons, New York, 3rd Edition, 1990.

estándar calculados para la clase 3. La ecuación (9.5) fue usada para calcular la matriz score para la clase 1 y 2. Como se muestra en la figura 9., la media de cada vector score para las clases 1 y 2 no es cero. De hecho, todos los puntos de estas clases están fuera de la región elíptica de confianza, indicando que los datos de las clases 1 y 2 son de hecho diferentes de los datos de la clase 3.

La estadística T^2 en (9.12) es en general sensible a inexactitudes en el espacio PCA correspondiente a los más pequeños valores singulares debido a que miden directamente las variaciones a lo largo de los vectores de carga. En otras palabras, esto es medir directamente los scores correspondientes a los pequeños valores singulares.

A.9.1.5 IDENTIFICACIÓN DE FALLAS

Una vez la falla ha sido detectada, el siguiente paso es determinar la causa de estado fuera de control. La tarea de diagnóstico de la falla puede ser un reto cuando el número de las variables del proceso es grande, y el proceso está altamente integrado. También, muchas de las variables medibles pueden desviarse de sus puntos de referencia solo por un corto periodo de tiempo cuando la falla ocurre, debido a que los lazos de control las regresan a sus puntos de referencia (aunque la falla sea persistente en el sistema). Este tipo de comportamiento de sistemas puede disfrazar la falla, haciendo difícil automatizar el algoritmo de diagnóstico para un correcto aislamiento de la falla presente en el sistema.

El objetivo de la identificación de falla es determinar cuáles variables de observación son más relevantes para el diagnóstico de falla, enfocando así a los operadores de planta y a los ingenieros en los sub-sistemas que tienen mayor probabilidad de falla. La asistencia prestada por el esquema de identificación de falla para la localización de las anomalías puede incorporar efectivamente a los operadores y los ingenieros en el proceso de monitoreo y reducir significativamente el tiempo de recuperación del control de las operaciones.

Tradicionalmente, las técnicas estadísticas univariantes fueron empleadas para la identificación de fallas. Dado un vector de observación x , los errores normalizados para cada variable x_j fueron calculados como

$$e_j = (x_j - \mu_j)/s_j \quad (9.24)$$

donde μ_j es la media y s_j es la desviación estándar de la j^{th} variable. Estos errores normalizados son ilustrados en una misma gráfica, y con los umbrales basados en el nivel de importancia se detectan las variables fuera de control, como se discutió en la sección 8.1.3. Sin embargo, las técnicas de estadística univariante para la identificación de la falla pueden dejar variables afuera que son responsables de la anomalía debido a que la técnica no toma en cuenta correlaciones entre las variables del proceso, o puede dar lecturas de alarma para muchas variables de las cuales los ingenieros obtienen poca información sobre las principales variables de interés.

Las gráficas de contribución son un enfoque PCA o de identificación de falla que toman en cuenta las correlaciones espaciales, mejorando así las técnicas estadísticas univariantes. El enfoque es basado en la cuantificación de la contribución de cada variable de proceso en el score individual de la representación PCA, y para cada variable de proceso reduce a las contribuciones

solo de aquellos score responsables del estado fuera de control. El procedimiento es aplicado en respuesta a la violación de T^2 , y esto se resume como sigue:

1. Revisar los score normalizados $(t_i/\sigma_i)^2$ para la observación x y determinar los score $r < \alpha$ responsables del estado fuera de control. Por ejemplo, aquellos score con $(t_i / \sigma_i)^2 > (1/\alpha)(T_\alpha^2)$. (Recordando que t_i es el score de la observación proyectada sobre el i^{th} vector de carga, y σ_i es el correspondiente valor singular)
2. Calcular la contribución de cada variable x_j en los scores fuera de control t_i

$$cont_{i,j} = \frac{t_i}{\sigma_i^2} p_{i,j} (x_j - \mu_j) \quad (9.25)$$

Donde $p_{i,j}$ es el $(i, j)^{\text{th}}$ elemento de la matriz de carga P .

3. Cuando $cont_{i,j}$ es negativa, es ajustada a cero.
4. Calcular el *total de la contribución* de la j^{th} variable de proceso, x_j ,

$$CONT_j = \sum_{i=1}^r cont_{i,j} \quad (9.26)$$

5. Graficar $CONT_j$ para todas la m variables del proceso, x_j , en una sólo gráfico.

Las variables responsable de la falla puede ser priorizadas u ordenadas por el valor total de la contribución $CONT_j$, y los operadores de planta y los ingenieros pueden inmediatamente enfocarse en estas variables con altos valores de $CONT_j$ y usar las propiedades conocidas para determinar la causa del estado fuera de control. Mientras en general el enfoque de contribución de variables puede ser aplicado a una porción de espacio de observación correspondiente a los m -

avalores singulares más pequeños, esto no es práctico debido a que los valores de contribución total $CONT_j$, podrían ser sensibles a los pequeños valores singulares.

Wise y otros²⁵ desarrollaron un enfoque PCA para la identificación de falla el cual se basa en la cuantificación de la variación total de cada variable de proceso en el espacio residual. Asumiendo que $m - a$ valores singulares pequeños son todos iguales, la varianza de cada variable x_j dentro del espacio residual puede ser estimada así

$$\hat{s}_j^2 = \sum_{i=a+1}^p p_{i,j} \sigma_1^2 \quad (9.27)$$

Dadas q nuevas observaciones, la varianza de la j^{th} variable fuera del espacio del modelo PCA puede ser probada en

$$s_j^2 / \hat{s}_j^2 > F_\alpha(q - a - 1, n - a - 1) \quad (9.28)$$

Y podría indicar una variable fuera de control, donde s_j^2 y \hat{s}_j^2 son las varianzas estimadas de la j^{th} variable para el nuevo conjunto de entrenamiento, respectivamente, y $F_\alpha(q - a - 1, n - a - 1)$ es el $(1-\alpha)$ límite porcentual usando la distribución F ²⁶. La ecuación (9.28) es la prueba de la hipótesis nula, siendo esta $s_j = \hat{s}_j$ y la hipótesis alternativa unilateral siendo $s_j > \hat{s}_j$. Esta última es aceptada (por ejemplo, la hipótesis nula es rechazada) si (9.28) cumple. La mayoría de las

²⁵B.M. Wise, N. L. Ricker, D. J. Velkmap, and D.F. Veltkamp. Upset and sensor failure detection in multivariate processes. Technical report, Eigenvector Research, Manson, Washington, 1989.

²⁶W.W. Hines and D. C. Montgomery. Probability and Statistics in Engineering and Management Science. John Wiley & Sons, New York, 3rd Edition, 1990.

veces, la variable responsable de la falla posee un valor de varianza más grande que la correspondiente al conjunto de entrenamiento (por ejemplo, $s_j > \hat{s}_j$). Sin embargo, esto no siempre es cierto, por ejemplo un sensor roto puede dar lecturas constantes, indicando que $s_j < \hat{s}_j$. Esto motiva el uso de la prueba de la hipótesis bilateral, con la hipótesis nula siendo $s_j = \hat{s}_j$ y la hipótesis bilateral siendo $s_j \neq \hat{s}_j$. Concluimos que $s_j \neq \hat{s}_j$ si

$$s_j^2 / \hat{s}_j^2 > F_{\alpha/2}(q - a - 1, n - a - 1) \quad (9.29)$$

O

$$\hat{s}_j^2 / s_j^2 > F_{\alpha/2}(q - a - 1, n - a - 1) \quad (9.30)$$

Además, un gran cambio en la media dentro del espacio residual ocurre si

$$\frac{\mu_j - \hat{\mu}_j}{\hat{s}_j \sqrt{\frac{1}{q-a} + \frac{1}{n-a}}} > t_{\frac{\alpha}{2}}(q + n - 2a - 2) \quad (9.31)$$

O

$$\frac{\mu_j - \hat{\mu}_j}{\hat{s}_j \sqrt{\frac{1}{q-a} + \frac{1}{n-a}}} < -t_{\frac{\alpha}{2}}(q + n - 2a - 2) \quad (9.32)$$

Donde μ_j y $\hat{\mu}_j$ son las medias de x_j para el nuevo conjunto de observaciones, respectivamente, y $t_{\frac{\alpha}{2}}(q + n - 2a - 2)$ es la $(1 - \frac{\alpha}{2})$ límite porcentual usando la t distribución. Las ecuaciones 9.31 y 9.32 son probados en la hipótesis nula, siendo

esta $\mu_j = \hat{\mu}_j$ y la hipótesis alternativa que es $\mu_j \neq \hat{\mu}_j$. La hipótesis alternativa es aceptada si (9.31) o (9.32) cumple [120].

Las variables responsables del estado fuera de control, detectado por la estadística Q, pueden ser identificadas usando (9.28), (9.31), y (9.32). Además, las variables pueden ser priorizadas usando los valores de la expresión (9.28), (9.31), y (9.32) donde la variable con el mayor valor es la que da la prioridad. En [346], las fallas en los sensores son detectadas e identificadas usando (9.28), (9.31), y (9.32). Otros métodos basados en PCA se desarrollaron específicamente para detectar fallas en sensores y son discutidas en otro lugar.

El enfoque de identificación de fallas usando (9.28), (9.31) y (9.32) requiere un grupo de $q \gg 1$ observaciones. Como se discutió en la sección 8.1.3, las medidas basadas en varias observaciones consecutivas son capaces de incrementar la robustez y la sensibilidad sobre las medidas basadas solo en una única observación, pero resulta en un tiempo de respuesta lento para procesos con grandes cambios. Una medida de identificación de falla basada en un vector de observación en un único instante de tiempo es el error normalizado

$$RES_j = r_j / \hat{s}_j \quad (9.33)$$

Donde r_j es la j^{th} variable del vector residual. Los valores de (9.33) pueden ser usados para priorizar las variables donde con el error normalizado más alto se da la prioridad. La medida de (9.33), cuando se compara con (9.28), (9.31) y (9.32) es capaz de indicar el estado corriente del proceso inmediatamente después de un gran cambio de proceso de manera más aproximada.

A.9.1.6 DIAGNÓSTICO DE FALLA

En la sección previa se discutió los métodos de identificación de falla, los cuales identifican las variables asociadas con la falla de sub-sistemas. Aunque estos métodos ayudan al diagnóstico de fallas, estos pueden tomar una gran cantidad de tiempo y proceso de especialización por parte de los operadores de planta y los ingenieros antes de que la falla sea apropiadamente diagnosticada. Mucho de este tiempo y especialización puede ser eliminado usando esquema enfocado al diagnóstico. Esto es construir por separado un modelo PCA para cada unidad del proceso. Una falla asociada con una unidad del proceso particular se asume que ocurre si el modelo PCA para esta indica que el proceso está fuera de control. Mientras este enfoque reduce la causa de una operación anormal de operación, no diagnosticará inequívocamente la causa. La diferencia entre la técnica de *aislamiento de falla* (la cual se basa en clasificación no supervisada) de la técnica de diagnóstico (la cual es basada en clasificación supervisada) se describe a continuación.

Varias investigaciones han proporcionado técnicas para usar componentes principales para el diagnóstico de falla. El enfoque simple para construir un sencillo modelo PCA y definir regiones de bajas dimensiones las cuales clasifican si una falla particular ha ocurrido. Este enfoque improbable que sea efectivo cuando un número significativo de fallas puede ocurrir.

Un enfoque PCA el cual pueda manipular un gran número de fallas usa un modelo PCA que desarrolla un modelo basado en los datos recolectados durante cada

falla específica, y luego aplica la estadística Q , T^2 u otras para cada modelo PCA para predecir cual falla o fallas es más probable que ocurra. Este enfoque es una combinación de componentes principales y análisis discriminante, pero que no se va a discutir en este documento.

A.9.2 ANÁLISIS DISCRIMINANTE DE FISHER.

A.9.2.1 INTRODUCCION

En el enfoque de clasificación de patrones para el diagnóstico de fallas descrito anteriormente, se describió como la reducción de dimensionalidad del paso de extracción de características puede ser un factor clave en la disminución de la tasa de error en la clasificación cuando un sistema de clasificación de patrones es aplicado a nuevos datos (datos independientes del conjunto de entrenamiento). La reducción de dimensionalidad es especialmente importante cuando la dimensionalidad del espacio de observación es grande mientras que los números de observaciones en las clases son relativamente pequeños. Un enfoque PCA para la reducción de dimensionalidad se discutió anteriormente. Aunque PCA contiene ciertas propiedades óptimas en términos de la detección de fallas, no es tan buena para el diagnóstico de fallas porque no toma en cuenta la información entre las clases cuando se determina la representación con menor dimensionalidad. El análisis discriminante de Fisher (FDA), es una técnica de reducción de dimensionalidad que ha sido extensivamente estudiada en la literatura de clasificación de patrones, toma en cuenta la información entre las clases y posee ventajas sobre el PCA para el diagnóstico de fallas.

A.9.2.2 ANALISIS DISCRIMINANTE DE FISHER

Para el diagnóstico de fallas, los datos recolectados de la planta durante fallas específicas son categorizados en clases, donde cada clase contiene data representando una falla particular. FDA es una técnica lineal de reducción de dimensionalidad, optima en términos de maximizar la separación entre estas clases. Determina un conjunto de vectores transformación lineal, ordenados en términos de maximizar la dispersión entre las clases mientras minimiza la dispersión dentro de cada clase.

Se define n como el número de observaciones, m como el número de variables medidas, p como el número de clases, y n_j como el número de observaciones en la j -ésima clase. Se representa el vector de variables medidas para la i -ésima observación como x_i

Para entender el análisis discriminante de Fisher, primero necesitamos definir varias matrices que cuantifiquen el total de dispersión, la dispersión dentro de las clases y la dispersión entre clases. La matriz de total de dispersiones:

$$S_t = \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^T \quad (9.34)$$

Donde \bar{x} es el vector total de la media

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (9.35)$$

Con x_j definida como el conjunto de vectores x_i los cuales pertenecen a la clase j , la matriz de dispersión dentro de cada clase para la clase j es

$$S_j = \sum_{x_i \in x_j} (x_i - \bar{x}_j)(x_i - \bar{x}_j)^T \quad (9.36)$$

Donde \bar{x}_j es el vector de la media para la clase j :

$$\bar{x}_j = \frac{1}{n_j} \sum_{x_i \in x_j} x_i \quad (9.37)$$

La matriz de dispersión dentro de las clases es

$$S_w = \sum_{j=1}^p S_j \quad (9.38)$$

Y la matriz de dispersión entre las clases es

$$S_b = \sum_{j=1}^p n_j (x_j - \bar{x})(x_j - \bar{x})^T \quad (9.39)$$

La matriz de dispersión total es igual a la suma de la matriz de dispersión entre clases y la matriz de dispersión dentro de las clases

$$S_t = S_b + S_w \quad (9.40)$$

El objetivo del primer vector FDA es el de maximizar la dispersión entre clases mientras minimiza la dispersión dentro de las clases:

$$\max_{v \neq 0} \frac{v^T S_b v}{v^T S_w v} \quad (9.41)$$

Asumiendo que S_w es invertible donde $V \in R^m$. el segundo vector FDA es computado para maximizar la dispersión entre clases mientras minimiza la dispersión dentro de las clases en medio de todos los ejes perpendiculares al primer vector FDA, y así sucesivamente para los vectores FDA restantes. Los vectores FDA son iguales a los eigenvectores W_k del problema de eigenvalor generalizado

$$S_b W_k = \lambda_k S_w W_k \quad (9.42)$$

Donde los eigenvalores λ_k indican el grado de separabilidad global en medio de las clases mediante la proyección de los datos sobre W_k . Debido a que la dirección y no la magnitud de W_k

Es importante, la norma euclideana (raíz cuadrada de la suma de los cuadrados de cada elemento) de W_k puede escogerse igual a 1 ($\|W_k\| = 1$). Los vectores FDA pueden ser computados de el problema de eigenvalor generalizado mientras que S_w sea invertible. Esto casi siempre será verdad si el numero de observaciones n es significativamente mayor que el número de mediciones m (el caso en la práctica).

El primer vector FDA es el eigenvector asociado con el eigenvalor más grande, el segundo vector FDA es el eigenvector asociado con el segundo eigenvalor más grande, y así sucesivamente. Un valor grande del eigenvalor indica que cuando los datos en las clases son proyectados sobre el eigenvector asociado W_k hay una gran separación global de las medias de las clases relativas a las varianzas de las clases, y consecuentemente, un mayor grado de separación en medio de las clases a lo largo de la dirección W_k . ya que el rango de S_b es menor que p , habrá por lo menos $p-1$ eigenvalores los cuales no son iguales a cero, y el FDA provee un ordenamiento útil de los eigenvectores solo en estas direcciones.

Es más útil escribir la meta del FDA de forma más explícita en términos de una transformación lineal. Se define la matriz $W_p \in R^{m \times (p-1)}$ con los vectores FDA $p-1$ como columnas. Luego la transformación lineal de los datos del espacio m -dimensional a el espacio $(p-1)$ - dimensional se describe por

$$Z_i = W_p^T x_i \quad (9.43)$$

Donde $Z_i \in R^{(p-1)}$. FDA computa la matriz W_p con datos x_1, \dots, x_n para la p clases óptimamente separadas cuando se proyectan sobre el espacio $p-1$ dimensional.

En el caso donde $p = 2$, esto es equivalente a proyectar los datos sobre una línea en la dirección del vector W para el cual los datos proyectados están lo mas separados.

A.9.2.3 REDUCCION DE ORDEN

La reducción de dimensionalidad no sería necesaria si la matriz de covarianza y el vector de media se conocieran exactamente. Errores en la matriz de covarianza de la muestra ocurren en la práctica, sin embargo la reducción de dimensionalidad proporcionada por el FDA permite reducir la tasa de error en la mala clasificación cuando el sistema de patrón de clasificación es aplicado a datos nuevos. Un método popular para seleccionar el orden de reducción es validación cruzada. Este enfoque separa los datos en múltiples conjuntos: el conjunto de entrenamiento, y el conjunto de prueba (o validación). El procedimiento de reducción de dimensionalidad es aplicado a los datos en el conjunto de entrenamiento, y luego su rendimiento es evaluado aplicando el modelo de dimensión-reducida a los datos en el conjunto de prueba para cada reducción de orden. El orden de reducción es seleccionado para optimizar el rendimiento basado en el conjunto de prueba. Por ejemplo, si la meta es el diagnostico de fallas, el orden del modelo reducido debería especificarse minimizando la tasa de error de clasificación del conjunto de prueba.

Validación cruzada no es siempre práctica en las aplicaciones de diagnóstico de fallas debido a que puede no haber suficientes datos para separar en dos conjuntos. En esta situación, es deseable determinar el orden de la reducción de dimensionalidad utilizando todos los datos del conjunto de entrenamiento.

El error de un modelo puede ser minimizado escogiendo el número de parámetros que compensen óptimamente errores de sesgado y las contribuciones de varianza en el error cuadrático medio. En un esfuerzo por minimizar el error cuadrado medio, el criterio:

$$(\text{Término de error de predicción}) + (\text{Término de complejidad del sistema}) \quad (9.44)$$

ha sido minimizado para determinar el modelo de orden apropiado. El criterio de información Akaike (AIC), popularmente aplicado en el sistema de identificación para la selección óptima del modelo de orden, puede ser derivado de la forma (9.44)

En (9.44) el término de predicción de error es una función de un modelo estimado de parámetros y los datos del conjunto de entrenamiento, y el término de complejidad del sistema es una función del número de parámetros independientes y la cantidad de datos en el conjunto de entrenamiento. En la identificación del sistema, el término de predicción de error es generalmente escogido como el promedio cuadrado del error de predicción para el modelo, pero en general, la selección del término complejidad es subjetivo.

Una fortaleza del AIC es que descansa solo en la información en un conjunto de datos (los datos de entrenamiento), a diferencia de la validación cruzada la cual requiere además datos adicionales o una partición del conjunto de datos

originales. Un criterio de la forma (9.44) puede desarrollarse para seleccionar automáticamente el orden de FDA utilizando la información solo en el conjunto de entrenamiento. El orden puede ser determinado mediante el cómputo de la dimensionalidad a que minimiza el criterio de información.

$$f_m(a) + \frac{a}{\tilde{n}} \quad (9.45)$$

Donde $f_m(a)$ es la tasa de error de clasificación (la proporción de errores de clasificación, el cual varía de 0 a 1) para el conjunto de entrenamiento mediante la proyección de los datos sobre el primer vector FDA a , y \tilde{n} es el número promedio de observaciones por clase. La tasa de error de clasificación $f_m(a)$, indica la cantidad de información contenida en el primer vector FDA a beneficiosa para el patrón de clasificación. Mientras que la tasa de error de clasificación del conjunto de entrenamiento típicamente disminuye a medida que a crece, para nuevos datos (datos independientes del conjunto de entrenamiento), la tasa de error de clasificación inicialmente disminuye y luego se incrementa por arriba un cierto orden debido al sobreajuste de los datos. El término de complejidad del modelo $\frac{a}{\tilde{n}}$ es añadido en (9.45) para penalizar el incremento de dimensionalidad. El escalamiento del orden de reducción a por el número promedio de observaciones en cada clase es el mismo, $n_j = \tilde{n}$. Se puede mostrar utilizando álgebra simple, que la inclusión de $\frac{a}{\tilde{n}}$ en (9.45) asegura que el procedimiento de la selección de orden produzca un valor para a menor que o igual a \tilde{n} . En otras palabras, esta restricción previene el modelo de baja dimensionalidad de tener una alta dimensionalidad mayor que la que justifica el número de observaciones por clase.

El término de complejidad del modelo $\frac{a}{\tilde{n}}$ puede también ser interpretado en términos del número total de errores de clasificación por clase. Definiendo $m(a)$ como el número total de errores de clasificación en el conjunto de entrenamiento

para el orden a y asumiendo $n_j = \tilde{n}$, el criterio de información 9.45 puede ser escrito como

$$\frac{m(a)}{p\tilde{n}} + \frac{a}{\tilde{n}} \quad (9.46)$$

Donde $n = p\tilde{n}$ es el número total de observaciones. Consideremos el caso donde se va a determinar si una reducción de orden $a + 1$ debería preferirse sobre un orden de reducción de a . Utilizando el criterio de información 9.46 y recordando que un valor más pequeño para el criterio de información es preferido, una reducción de orden de $a + 1$ se prefiere si

$$\frac{m(a+1)}{p\tilde{n}} + \frac{a+1}{\tilde{n}} < \frac{m(a)}{p\tilde{n}} + \frac{a}{\tilde{n}} \quad (9.47)$$

Esto es equivalente a $\frac{m(a)}{p} - \frac{m(a+1)}{p} > 1$ (9.48)

El término de complejidad no permite que la reducción de orden sea incrementada disminuyendo el número de errores de clasificación, pero solo si el decrecimiento en el número total de errores de clasificación por clase es mayor que 1.

El análisis anterior indica que el escalamiento de a en el término de modelo de complejidad $\frac{a}{\tilde{n}}$ en el criterio de información 9.45 es razonable.

A.9.2.4 DETECCION Y DIAGNOSTICO DE FALLAS

Cuando FDA es aplicado al patrón de clasificación, la técnica de reducción de dimensionalidad es aplicada a los datos en todas las clases simultáneamente. De forma más precisa, se denota $W_a \in R^{m \times a}$ como la matriz que contiene a los eigenvectores W_1, W_2, \dots, W_a computado de 9.42.

La función discriminante es,

$$g_j(x) = -\frac{1}{2}(x - \bar{x}_j)^T W_a \left(\frac{1}{n_j-1} W_a^T S_j W_a \right)^{-1} W_a^T (x - \bar{x}_j) + \ln(p_i) - \frac{1}{2} \ln \left[\det \left(\frac{1}{n_j-1} W_a^T S_j W_a \right) \right] \quad (9.49)$$

Donde S_j , \bar{x}_j , y n_j son definidos en 9.36 y 9.37. En contraste a PCA, FDA utiliza la información de la clase para computar el espacio reducido dimensionalmente, así que la función discriminante aprovecha esa información de la clase a un nivel mayor de lo que se ha hecho con PCA. En contraste a PCA, FDA usa todas las p fallas clases de información cuando evalúa la función discriminante o cada clase. FDA también puede ser utilizado para detectar fallas definiendo una clase adicional de datos, recolectados durante condiciones de operación normal, a las clases de fallas. La capacidad de la detección de fallas mediante 9.49 depende de la similitud entre los datos de las condiciones de operación normal y los datos de las clases de fallas en los conjuntos de entrenamiento. Cuando existe una transformación W tal que los datos de las condiciones normales de operación pueden ser razonablemente separados de las otras clases de fallas, utilizar FDA para la detección de fallas producirá pequeñas tasas de detección perdidas para las clases de fallas conocidas. La ecuación 9.49 no toma en cuenta las fallas asociadas con los datos por fuera del menor espacio dimensional definido por los vectores FDA, de forma que 9.49 puede que no detecte estas clases de fallas. Es mejor utilizar 9.49 con un FDA basado en el residuo estadístico, los cuales juntos

pueden detectar ambas clases asociadas con los datos dentro del espacio definido por los vectores FDA, y fallas asociadas con los datos fuera de este espacio. Esta unión usa de dos FDA estadística es similar a la unión usada del PCA T^2 . La ventaja de utilizar la estadística FDA en lugar de la estadística FDA es que la información para clasificación de las fallas se puede tomar en cuenta para mejorar la habilidad de detectar fallas. La desventaja de la estadística FDA es que requiere que la información para la clasificación de las fallas defina su más bajo espacio dimensional (definido por W).

Solo los primeros eigenvectores $p - 1$ en FDA maximizan la dispersión entre las clases mientras minimiza la dispersión dentro de cada clase. El resto de los eigenvectores $m - p + 1$ correspondientes a los eigenvalores cero no son ordenados por el FDA objetivo. EL RANKING de estos eigenvectores generalizados está determinado por el paquete de software implementando el algoritmo de descomposición de eigenvalores, el cual no ordena los eigenvectores en una manera necesariamente útil para la clasificación. Sin embargo, más que $p - 1$ dimensiones en un espacio de menor dimensión puede ser útil para la clasificación, y un procedimiento para la selección de vectores más allá del primer $p - 1$ FDA vectores pueden ser útiles. Aquí se describen dos métodos los cuales usan PCA para computar los vectores adicionales para la clasificación.

Un método es utilizar FDA para el espacio definido por el primer $p - 1$ eigenvectores, y para usar los vectores PCA1 para el resto de los vectores $m - p + 1$, ordenadas de los vectores PCA asociados con la mayor variabilidad a los vectores asociados con la menor variabilidad.

Si la orden de reducción $a \leq p - 1$, la ecuación 9.49 se usa directamente. Si $a \geq p$, la función discriminante alternativa es

$$g_j(x) = -\frac{1}{2}(x - \bar{x}_j)^T W_{mix,a} \left(\frac{1}{n_j - 1} W_{mix,a}^T S_j W_{mix,a} \right)^{-1} W_{mix,a}^T (x - \bar{x}_j) - \frac{1}{2} \ln \left[\det \left(\frac{1}{n_j - 1} W_{mix,a} S_j W_{mix,a} \right) \right] + \ln(p_i) \quad (9.50)$$

Donde $W_{mix,a} = [W_{p-1} P_{a-p+1}]$ y P_{a-p+1} es el primer $a - p + 1$ columnas del PCA1 matriz de carga P. cuando este método es utilizado para el diagnóstico de fallas, se referenciará como el método FDA/PCA1. Hay que recordar que las varianzas asociadas con los vectores de carga en PCA son clasificados en orden descendente. Dado que los vectores de PCA1 pueden ser útiles en un procedimiento de clasificación, incorporando el primer $a - p + 1$ vector de carga PCA1 en el método FDA/PCA1 puede proveer información adicional para la discriminación entre clases.

Otro método para definir un adicional $m - p + 1$ vectores es aplicar PCA1 al espacio residual de FDA, definido por

$$R = X(I - W_{p-1} W_{p-1}^T) \quad (9.51)$$

Como antes, si la reducción de orden $a < p - 1$ la ecuación 9.49 se usa directamente. Si $a > p$ la función alternativa 9.51 se utiliza con $W_{mix,a} = [W_{p-1} P_{a-p+1}]$, donde P_{a-p+1} es el primer $a - p + 1$ columna de la matriz de carga PCA1 cuando PCA es aplicado a R. este método para el diagnóstico de fallas se referirá como el método FDA/PCA2.

Comparación de PCA y FDA

Las técnicas de reducción de dimensionalidad de PCA y FDA son comparadas via análisis teóricos y gráficos para el caso donde PCA es aplicado a todos los datos en todas las clases juntas. Esto resalta las diferencias geométricas entre los dos procedimientos de reducción de dimensionalidad. La forma en que FDA puede resultar en un diagnóstico de fallas superior al alcanzado por PCA se muestra a continuación.

Los problemas de optimización para PCA y FDA han sido enunciados matemáticamente en 9.1 y 5.8, respectivamente. Se puede mostrar que los vectores de carga de PCA y los vectores FDA pueden ser calculados mediante el cómputo de puntos estacionarios de los problemas de optimización:

$$\max_{v \neq 0} \frac{v^T S_t v}{v^T v} \quad (9.52)$$

Y

$$\max_{v \neq 0} \frac{v^T S_t v}{v^T S_w v} \quad (9.53)$$

Respectivamente. Las ecuaciones anteriores 9.52 y 9.53 indican que los vectores PCA y FDA son idénticos para el caso cuando $S_w = \sigma I$ donde $\sigma > 0$. Un caso en el cual esta situación ocurre si los datos en cada clase pueden ser descritos mediante una bola distribuida uniforme (circulo en espacio 2D y una esfera en espacio 3D), incluso si las bolas son de diferentes tamaños. Diferencias entre las dos técnicas pueden ocurrir solo si la elongación en los datos utilizados para describir cualquiera de las clases. Estas formas elongadas ocurren altamente correlacionadas al conjunto de datos (fig. 80), típico para datos recolectados de procesos industriales. Por ello, cuando se aplica PCA y FDA en la misma forma a los datos del proceso, los vectores de carga PCA y los vectores FDA se esperan que sean significativamente diferentes, y los objetivos difieren, 9.52 y 9.53,

sugieren que FDA será significativamente mejor para la discriminación entre las clases de fallas.

La fig. 80. ilustra una diferencia entre PCA y FDA que puede ocurrir cuando la distribución de los datos en las clases es un tanto elongado. El primer vector FDA y PCA de carga están casi perpendiculares, y la transformación lineal de los datos sobre el primer vector FDA es mucho más capaz de separar los datos en las dos clases que la transformación lineal de los datos sobre el vector PCA de carga.

Las transformaciones lineales de Fisher sobre los datos sobre los primeros dos vectores PCA y FDA de carga se muestran en la fig. 81.

La matriz de dispersión dentro de clases y la matriz de dispersión entre clases son calculadas como

$$S_w = \begin{bmatrix} 56.8 & 37 & 16.4 & 9.17 \\ 37.3 & 88.4 & 10.1 & 17.1 \\ 16.4 & 10.1 & 8.75 & 4.64 \\ 9.17 & 17.1 & 4.64 & 22.8 \end{bmatrix} \quad (9.54)$$

Y

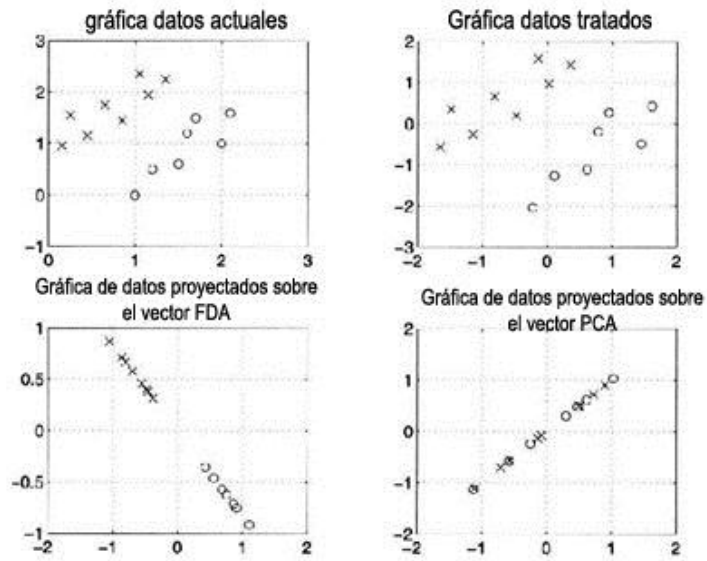


Figura 80. una comparación de la transformación lineal PCA y FDA de los datos en clases “x” y “o” sobre el primer vector de carga FDA y PCA.

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

$$S_b = \begin{bmatrix} 92.2 & -55.7 & 113 & 108 \\ -55.7 & 60.6 & -75.3 & -65.6 \\ 113 & -75.2 & 140 & 133 \\ 108 & -65.6 & 132 & 126 \end{bmatrix} \quad (9.55)$$

Respectivamente. Resolviendo 9.42, tenemos $p - 1 = 2$ eigenvectores asociados con eigenvalores diferentes de cero, los cuales son:

$$W_1 = \begin{bmatrix} 0.15 \\ 0.12 \\ -0.96 \\ -0.18 \end{bmatrix} \quad (9.56)$$

Y

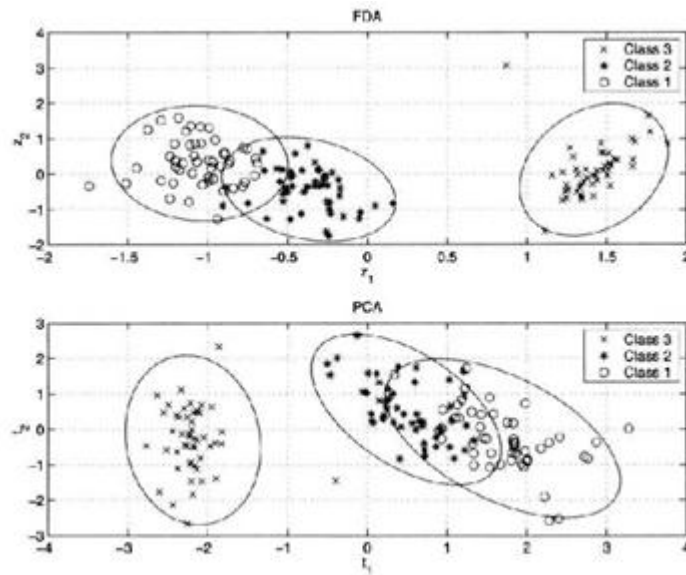


Figura 81. transformación lineal de Fisher de los datos para tres clases sobre los primeros dos vectores de carga FDA y PCA respectivamente.

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

$$W_2 = \begin{bmatrix} -0.13 \\ -0.70 \\ -0.15 \\ 0.68 \end{bmatrix} \quad (9.57)$$

Y los eigenvalores correspondientes son $\lambda_1 = 27$ y $\lambda_2 = 0.24$, respectivamente. Un gran valor de λ_1 indica que hay una gran separación de las medias de cada clase relativa con las clases de las varianzas z_1 . En efecto los valores promedios de z_1 para las tres clases son -1.0, -0.37, y 1.42. El valor pequeño de λ_2 indica que la separación total de las clases medias relativas a las varianzas de las clases es menor en la dirección z_2 . Los valores promedios de z_2 para las tres clases son 0.30, -0.43, y 0.12.

El 95% de la región de confianza elíptica para cada clase puede ser aproximada resolviendo el umbral con T_i^2 sustituyendo en 6.64. Los datos en la intersección de las dos regiones elípticas pueden resultar en un error de clasificación. El grado de traslape entre las regiones de confianza para las clases 1 y 2 para PCA tanto para FDA (49 puntos vs 17 puntos), indicando que las tasas de error de clasificación para PCA serían mayores.

Mientras la región de confianza puede ser utilizada para ilustrar el rendimiento de la clasificación cualitativa para FDA, la función discriminante 9.49 pueden ser utilizadas para determinar las tasas exactas de error en la clasificación para los datos del experimento. Los resultados son ilustrados en la tabla 6. para los diferentes órdenes de reducción FDA. Aunque los datos de la clase 1 y la clase 2 se traslapan cierta región, la función discriminante es capaz de clasificar casi todos los datos correctamente. En realidad, no más de 3 valores de 50 datos son mal clasificados sin importar el orden de la selección.

Tabla 6. Tasa de error de clasificación para los datos de entrenamiento para FDA

Orden (a)	1	2	3	4
Error de clasificación clase 1	0.06	0.02	0.02	0.02
Error de clasificación clase 2	0.06	0.06	0.04	0.06
Error de clasificación clase 3	0	0	0	0
Error de clasificación global	0.04	0.027	0.02	0.027

Fuente: Fault Detection and Diagnosis in Industrial Systems. L.H Chiang, E.L Russell and R.D. Braatz.

Este ejemplo es efectivo para ilustrar la diferencia de los objetivos entre PCA y FDA. Mediante la comparación de los límites de los ejes horizontal y vertical y la inspección visual de los datos, es claro que el lapso de la transformación lineal de PCA es mayor que la transformación lineal FDA. Mientras PCA es mejor para separar los datos como un todo, FDA es mejor para separar los datos entre las clases (*, o). Esto es evidente en el grado de traslape entre las regiones de datos "*" y "o" en las dos gráficas, en las cuales los data points "*" y "o" ligeramente se traslapan para la transformación lineal FDA, mientras hay una clara mezcla de los datos para la transformación lineal PCA.

Todos los datos de Fisher fueron utilizados para el entrenamiento de los modelos FDA y PCA en los ejemplos previos. Una comparación mucho más precisa de PCA y FDA es entrenar las técnicas con un conjunto de datos (el conjunto de entrenamiento de datos), luego aplicarlos a un nuevo conjunto de datos (los datos de prueba). En este ejemplo dos quintos de los datos de Fisher (20 observaciones para cada clase, para un total de 60 observaciones) fueron utilizados para el entrenamiento, mientras el resto de los datos (30 observaciones para cada clase, para un total de 90 observaciones) fueron utilizados para la prueba. Las tasas totales de error en la clasificación de los datos de entrenamiento y los datos testeados utilizando los métodos de diagnóstico de fallas data-driven son mostrados en la tabla 9.3 y 9.4, respectivamente.

La tasa global de error de clasificación para FDA, FDA/PCA1, y FDA/PCA2 fueron los mismos a un dado orden de reducción. Los vectores FDA correspondiente a los dos eigenvalores diferentes de cero son muy efectivos en la discriminación de las tres clases. At $a = 2$, el error global de clasificación para el conjunto de prueba es 0.0667 (84 de 90 datos fueron clasificados correctamente).

Para $a < p$, los métodos FDA tuvieron una menor tasa de error en la clasificación que el método PCA. Esto concuerda con comentarios anteriores que el FDA puede realizar un mejor trabajo en el diagnóstico de fallas que el PCA,

especialmente en ordenes bajos de reducción. A cualquier orden de reducción, PCA1 dio una tasa de error de clasificación menor que PCAm. Esto soporta la discusión que PCA1 generalmente produciría una mejor representación del PCA para el diagnóstico de fallas. Para $a = 4$, todos los métodos, excepto para PCAm, dio la misma tasa de error en la clasificación.

Para ilustrar la dependencia de los números de datos utilizados en el conjunto de entrenamiento en la eficiencia de la clasificación, otro ejemplo fue correr utilizando 120 observaciones en el conjunto de entrenamiento y 30 observaciones en el conjunto de prueba. La tasa de error en la clasificación global para los datos de entrenamiento y los datos de prueba se muestran en la tabla 9 y 10, respectivamente.

Este ejemplo muestra que, con más datos en el conjunto de entrenamiento, la tasa de error en la clasificación global en el conjunto de prueba para todos los métodos es significativamente menor. Este ejemplo muestra las mismas tendencias que todos los métodos FDA superan los métodos PCA, y que PCA1 es mejor que el PCAm.

Note que este conjunto de datos es un ejemplo a relativamente baja escala, en el cual la reducción de dimensionalidad no fue necesaria para proveer bajas tasas de error de clasificación. El beneficio de la reducción de dimensionalidad es más aparente para la clasificación de nuevos datos de sistemas a gran escala, en los cuales el entrenamiento de datos son insuficientes (casopráctico en la industria).

Tabla 7. tasa de error de clasificación para los datos de entrenamiento, utilizando varios métodos para el diagnostico de fallas data-driven (60 observaciones en el conjunto de entrenamiento)

Orden (a)	1	2	3	4
FDA	0	0	0	0
FDA/PCA1	0	0	0	0
FDA/PCA2	0	0	0	0
PCA1	0.083	0.033	0.020	0
PCAm	0.28	0.20	0.15	0.13
MS	-	-	-	0

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

Tabla 8. tasa de error de clasificación global de los datos de prueba, utilizando varios métodos de diagnostico de fallas data-driven (60 observaciones en el conjunto de entrenamiento)

Orden(a)	1	2	3	4
FDA	0.067	0.067	0.078	0.033
FDA/PCA1	0.067	0.067	0.078	0.033
FDA/PCA2	0.067	0.067	0.078	0.033
PCA1	0.10	0.10	0.044	0.033
PCAm	0.17	0.18	0.11	0.11
Ms	-	-	-	0.033

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

A.9.2.5 FDA DINÁMICO

El método de patrón de clasificación para el diagnóstico de fallas puede ser extendido para tomar en cuenta las serial (temporal) correlaciones en los datos, mediante el aumento del vector de observación y apilando la matriz de datos. Este método se referirá como FDA dinámico. Esto permite al sistema de clasificación de patrones utilizar más información en la clasificación de las observaciones. Ya que la información contenida en el vector de observación aumentado es un súper conjunto de la información contenida en un solo vector de observación, es de esperar desde un punto de vista teórico que el enfoque de vector aumentado puede resultar en mejor rendimiento. Sin embargo, la dimensionalidad del problema es incrementada y apilando los datos, donde la magnitud del incremento depende del número de retrasos h . esto implica que más datos pueden ser requeridos para determinar el vector de media y la matriz de covarianza al mismo nivel de precisión para cada clase. En la práctica, aumentar el vector de observación es esperado que tenga mejor rendimiento cuando hay la correlación serial significativa y hay suficientes datos para justificar una mayor dimensionalidad. Ya que la cantidad de datos n es generalmente fijo, realizando la reducción de dimensionalidad utilizando FDA llega a ser incluso más crítica para el sistema de clasificación de patrones cuando el número de retrasos h es grande.

Tabla 9. Tasa de error de clasificación para los datos de entrenamiento utilizando varios métodos de diagnóstico data-driven (120 observaciones en el conjunto de entrenamiento)

Orden (a)	1	2	3	4
FDA	0.05	0.033	0.033	0.033
FDA/PCA1	0.05	0.033	0.033	0.033
FDA/PCA2	0.05	0.033	0.033	0.033
PCA1	0.092	0.10	0.050	0.033
PCAm	0.20	0.19	0.067	0.067
Ms	-	-	-	0.033

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

Tabla 10. Tasa de error en la clasificación para los datos de prueba utilizando varios métodos de diagnóstico de fallas data-driven(120 observaciones en el conjunto de entrenamiento)

Orden(a)	1	2	3	4
FDA	0	0	0	0
FDA/PCA1	0	0	0	0
FDA/PCA2	0	0	0	0
PCA1	0.033	0.033	0	0
PCAm	0.067	0.067	0.067	0.067
Ms	-	-	-	0

Fuente: *Fault Detection and Diagnosis in Industrial Systems*. L.H Chiang, E.L Russell and R.D. Braatz.

A.9.3 MÍNIMOS CUADRADOS PARCIALES

A.9.3.1 INTRODUCCIÓN

Mínimos cuadrados parciales (PLS), también conocido como proyección a estructuras latentes, es una técnica de reducción dimensional para maximizar la covarianza entre la matriz (independiente) predictor X y la matriz (dependiente) predicha Y para cada componente del espacio reducido. Una aplicación popular de PLS es seleccionar la matriz Y para contener sólo datos de calidad del producto que pueden incluir inclusive datos de medidas off-line, y la matriz X para contener todas las demás variables del proceso. Tales modelos inferenciales

(también conocidos como sensores suaves) pueden ser usados para la predicción on-line de la calidad de los datos del producto, para la incorporación en algoritmos de control del monitoreo del proceso, como también como para monitoreo de proceso. PLS discriminante selecciona la matriz X para contener todas las variables del proceso y selecciona la matriz Y para enfocar PLS en las tareas de diagnóstico de fallas.

PLS calcula los vectores *carga* y *score* que son correlacionados con el bloque predictor mientras se describe una gran cantidad de la variación en el bloque predictor. Si el bloque predictor tiene una sola variable, el método de reducción de dimensional es conocido como PLS1; si el bloque predictor tiene múltiples variables, el método de reducción dimensional es conocido como PLS2. PLS requiere pasos de calibración y predicción. El algoritmo más popular usado en PLS para computar los parámetros en el paso de calibración es conocido como **mínimos cuadrados parciales no-iterativos** (NIPALS). Otro algoritmo, conocido como SIMPLS, puede también ser usado. Como se mencionó, los bloques predictores usados en PLS discriminante y en otras aplicaciones de PLS son diferentes. En quimiometría y aplicaciones en control de procesos, donde PLS es más comúnmente aplicado, las variables que se predicen son usualmente mediciones de variables de calidad de los productos. En clasificación de patrones, donde PLS discriminante es usado, las variables que se predicen son variables ficticias (1 o 0) donde '1' indica un miembro de en-clase mientras '0' indica un miembro que no es de en-clase. En el paso de predicción del PLS discriminante, análisis discriminante es usado para determinar la clase predicha.

A.9.3.2 ALGORITMOS PLS

PLS requiere de una matriz $X \in R^{n \times m}$ y una matriz $Y \in R^{n \times p}$, donde m es el número de variables predictoras (el número de mediciones en cada observación), n es el número total de observaciones en el conjunto de entrenamiento, y p es el número de variables de observación en Y . Cuando Y es seleccionada para contener únicamente las variables de calidad del producto, entonces p es el número de variables de calidad del producto. Cuando Y es seleccionado como en PLS discriminante, p es el número de clases de falla.

En PLS discriminante, datos diagnosticados son necesarios en la calibración. De ayuda en la descripción de PLS discriminante, los datos en X serán ordenados de una manera particular. Con p clases de falla, suponiendo que son n_1, n_2, \dots, n_p observaciones para cada variable en clases $1, 2, \dots, p$ respectivamente. Recoger el grupo de datos de entrenamiento en la matriz $X \in R^{n \times m}$, como se muestra en ec. (9.58):

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n_1} & x_{n_2} & \cdots & x_{nm} \end{bmatrix} \quad (9.58)$$

Tal que la primera fila n_1 contiene datos de la falla 1, la segunda filas n_2 contiene datos de la segunda falla, y así sucesivamente. En total, hay $n_1 + n_2 + \cdots n_p = n$ filas. Existen dos métodos, conocidos como PLS1 y PLS2, para modelar el bloque que se predice. En PLS1, cada una de las p variables que se predicen y se modelan por separado, resultando en un modelo para cada clase. En PLS2, todas las variables que se predicen son modeladas simultáneamente.

En PLS2, el bloque que se predice $Y \in R^{n \times p}$ contiene p variables de calidad de producto; en PLS2 discriminante, el bloque que se predice $Y \in R^{n \times p}$ es

$$Y = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \dots & \ddots & \dots & \vdots \\ \vdots & \dots & \ddots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad (9.59)$$

Donde cada columna en Y corresponde a una clase. Cada elemento de Y es llenado ya sea con *uno* o *cero*. Los primeros n_1 elementos de la columna 1 son llenados con un '1', el cual indica que las primeras n_1 filas de X son datos de la falla 1. En el PLS1 discriminante, el algoritmo es corrido p veces, cada uno con el mismo X , pero para cada columna individual de Y en (9.59).

Como ya se mencionó el pre-tratamiento de datos es aplicado primero, entonces X y Y son centradas por la media y escaladas. La matriz X es descompuesta en una matriz $scoreT \in R^{n \times m}$:

$$X = TP^T + E. \quad (9.60)$$

La matriz producto TP^T puede ser expresada como la suma del producto de los vectores $scoret_j$ (la j^{th} columna de T) y los vectores $cargap_j$:

$$X = \sum_{j=1}^a \mathbf{t}_j \mathbf{p}_j^T + \mathbf{E}. \quad (9.61)$$

De manera similar, Y es descompuesto en una matriz $scoreU \in R^{n \times a}$, una matriz $carga Q \in R^{p \times a}$ mas una matriz residual $F \in R^{n \times p}$:

$$Y = UQ^T + \bar{F} \quad (9.62)$$

La matriz producto UQ^T puede ser expresada como la suma del producto de los vectores *score* \mathbf{u}_j (la j^{th} columna de U) y los vectores *carga* \mathbf{q}_j (la j^{th} columna de Q):

$$Y = \sum_{j=1}^a \mathbf{u}_j \mathbf{q}_j^T + \bar{F} \quad (9.63)$$

La descomposición en (9.61) y (9.63) tienen la misma forma como la usada en PCA. Las matrices X y Y son representadas como la suma de una serie de matrices de rango uno. Si a se establece igual a $\min(m, n)$, entonces E y F son cero y PLS se reduce a mínimos cuadrados ordinarios. Fijando a menor que $\min(m, n)$ se reduce ruido y colinealidad. La meta de PLS es determinar los vectores de *carga* y *score* que están correlacionados con Y al describir una gran cantidad de la variación en X .

PLS regresa el estimado Y vector *score* $\hat{\mathbf{u}}_j$ al X vector *score* \mathbf{t}_j mediante:

$$\hat{\mathbf{u}}_j = b_j \mathbf{t}_j \quad (9.64)$$

Donde b_j es el coeficiente de regresión. En forma matricial, esta relación puede ser escrita como:

$$\hat{U} = TB \quad (9.65)$$

Donde $B \in R^{a \times a}$ es la matriz de regresión lineal con $B_{jj} = b_j$, y U tiene $\hat{\mathbf{u}}_j$ como sus columnas. Sustituyendo U de (9.65) en (9.62), y teniendo en cuenta que esto modificará la matriz residual, da:

$$Y = TBQ^T + F \quad (9.66)$$

Donde F es la matriz de predicción de error. La matriz B es seleccionada tal que la norma 2 de F (el valor máximo singular de F) $\|F\|_2$, es minimizada. Los vectores *score* \mathbf{t}_j y $\hat{\mathbf{u}}_j$ son calculados para cada factor PLS ($j = 1, 2, \dots, a$) tal que la

covarianza entre X y Y es maximizada en cada factor. En PLS1, pasos similares son hechos, resultando en:

$$\mathbf{y}_i = T_i B_i \mathbf{q}_i^T + \mathbf{f}_i \quad (9.67)$$

Donde $\mathbf{y}_i \in R^n$ es la i^{th} columna de Y , $T_i \in R^{n \times a}$ es la matriz score, $B_i \in R^{a \times a}$ es la matriz de regresión, $\mathbf{q}_i \in R^a$ es el vector carga, y $\mathbf{f}_i \in R^n$ es el vector de predicción de error. Ya que hay p columnas en Y , el rango de i es de 1 a p .

Ahora si las matrices de score y carga para X y Y fueron calculadas separadamente, entonces sus sucesivos vectores score pueden estar débilmente interrelacionados, por lo tanto la regresión (9.65) la cual relaciona X y Y resultará en una pobre relación de reducción dimensional. El algoritmo NIPALS es un enfoque iterativo para computar vectores score modificados tal que resulta en componentes rotados que conllevan a una mejora en la regresión en (9.65). Esto se hace mediante el uso de vectores score de Y en el cálculo de los vectores score para X , y viceversa.

Para el caso de PLS2, el algoritmo NIPALS computa los parámetros usados en (9.68) a (9.78). EL primer paso es la regresión entre X y Y , que son escaladas a fin de tener media cero y varianza unitaria para cada variable. Inicializar el algoritmo NIPALS usando $E_0 = X$ y $F_0 = Y, j = 1$, y \mathbf{u}_j igual a cualquier columna de F_{j-1} . Ecuaciones 9.68- 9.71 son iterativamente calculadas hasta la convergencia, que es determinada comparando \mathbf{t}_j con su valor de una previa iteración (la nomenclatura $\|\cdot\|$ se refiere al vector 2-norm, también conocida como norma euclidiana).

$$\mathbf{w}_j = \frac{E_{j-1}^T \mathbf{u}_j}{\|E_{j-1}^T \mathbf{u}_j\|} \quad (9.68)$$

$$\mathbf{t}_j = E_{j-1} \mathbf{w}_j \quad (9.69)$$

$$\mathbf{q}_j = \frac{F_{j-1}^T \mathbf{t}_j}{\|F_{j-1}^T \mathbf{t}_j\|} \quad (9.70)$$

$$\mathbf{u}_j = F_{j-1} \mathbf{q}_j \quad (9.71)$$

Proceder a (9.71) si converge; regresar a (9.68) si no. Matemáticamente determinar \mathbf{t}_1 , \mathbf{u}_1 , y \mathbf{w}_1 de (9.68) a (9.71) es igual que determinar iterativamente los eigenvectores de $X X^T Y Y^T$, $Y Y^T X X^T$, y $X^T Y Y^T X$ asociado con el eigenvalor más grande, respectivamente.

En el segundo paso, \mathbf{p}_j es calculado como

$$\mathbf{p}_j = \frac{E_{j-1}^T \mathbf{t}_j}{\mathbf{t}_j^T \mathbf{t}_j} \quad (9.72)$$

Los valores finales para \mathbf{p}_j , \mathbf{t}_j , y \mathbf{w}_j son escalados por la norma de $\mathbf{p}_{j,old}$:

$$\mathbf{p}_{j,new} = \frac{\mathbf{p}_{j,old}}{\|\mathbf{p}_{j,old}\|} \quad (9.73)$$

$$\mathbf{t}_{j,new} = \mathbf{t}_{j,old} \|\mathbf{p}_{j,old}\| \quad (9.74)$$

$$\mathbf{w}_{j,new} = \mathbf{w}_{j,old} \|\mathbf{p}_{j,old}\| \quad (9.75)$$

También es común aplicar el escalamiento de (9.73) a (9.75) en el algoritmo, los escalonamiento no son absolutamente necesarios. En particular, los vectores *score* \mathbf{t}_j usados para relacionar X a Y en (7.3.7) son ortogonales en cualquiera de los casos.

Una vez que \mathbf{u}_j y \mathbf{t}_j son calculados usando la expresión de arriba, el coeficiente de regresión b_j que relaciona los dos vectores pueden ser calculados de:

$$b_j = \frac{\mathbf{u}_j^T \mathbf{t}_j}{\mathbf{t}_j^T \mathbf{t}_j} \quad (9.76)$$

Las matrices residuales E_j y F_j necesarias para la siguiente iteración son calculadas de:

$$E_j = E_{j-1} - \mathbf{t}_j \mathbf{p}_j^T \quad (9.77)$$

Y

$$F_j = F_{j-1} - b_j \mathbf{t}_j \mathbf{q}_j^T \quad (9.78)$$

Esto remueve la varianza asociada con los vectores *score* y *carga* ya calculados antes de calcular los vectores *carga* y *score* para la nueva iteración. El procedimiento entero es repetido para el siguiente factor (comúnmente llamado variable latente) $(j+1)$ empezando de (9.68) hasta $j=\min(m, n)$.

Predicciones basadas en el modelo PLS pueden ser calculadas directamente del vector observación y $\mathbf{p}_j, \mathbf{q}_j, \mathbf{w}_j$, y b_j por $j = 1, 2, \dots, \min(m, n)$. También hay un enfoque alternativo donde las predicciones son obtenidas de la matriz de regresión B_{2j}

$$B_{2j} = W_j (P_j^T W_j)^{-1} (T_j^T T_j)^{-1} T_j^T F_0 \quad (9.79)$$

Donde las matrices $P_j \in R^{\min(m,n) \times j}$, $T_j \in R^{n \times j}$, y $W_j \in R^{\min(m,n) \times j}$ son formadas de apilar los vectores $\mathbf{p}_j, \mathbf{t}_j$, y \mathbf{w}_j , respectivamente. Esta matriz es guardada para $j = 1, 2, \dots, \min(m, n)$.

El algoritmo NIPALS para PLS1 es calculado usando (9.80) a (9.85). Se inicia el algoritmo NIPALS usando $E_0 = X, j = 1$, y se fija $i = 1$. La siguiente ecuación es usada:

$$\mathbf{w}_{i,j} = \frac{E_{j-1}^T \mathbf{y}_i}{\|E_{j-1}^T \mathbf{y}_i\|} \quad (9.80)$$

$$\mathbf{t}_{i,j} = E_{j-1} \mathbf{w}_{i,j} \quad (9.81)$$

$$\mathbf{p}_{i,j} = \frac{E_{j-1}^T \mathbf{t}_{i,j}}{\mathbf{t}_{i,j}^T \mathbf{t}_{i,j}} \quad (9.82)$$

Después del re-escalado de $\mathbf{p}_{i,j}$, $\mathbf{t}_{i,j}$, y $\mathbf{w}_{i,j}$ de manera similar como se usó en (7.3.16) a (7.3.18), el coeficiente de regresión $b_{i,j}$ es calculado de

$$b_{i,j} = \frac{\mathbf{y}_i^T \mathbf{t}_{i,j}}{\mathbf{t}_{i,j}^T \mathbf{t}_{i,j}} \quad (9.83)$$

Los residuos de la siguiente iteración son calculados como sigue

$$E_j = E_{j-1} - \mathbf{t}_{i,j} \mathbf{p}_{i,j}^T \quad (9.84)$$

$$\mathbf{f}_{i,j} = \mathbf{f}_{i,j-1} - b_{i,j} \mathbf{t}_{i,j} q_{i,j} \quad (9.85)$$

Donde $\mathbf{f}_{0,i} = \mathbf{y}_i$ y $j = 1$. El procedimiento entero es repetido para la variable latente siguiente ($j + 1$) empezando de (9.80) hasta $j = \min(m, n)$. Después de todos los parámetros para $i = 1$ son calculados, el algoritmo es repetido para $i = 2, 3, \dots, p$.

Predicciones basadas en modelos PLS pueden ser calculadas directamente del vector de observación y el $b_{i,j}$. Alternativamente, las predicciones son obtenidas de la matriz de regresión $B1_j$

$$B_{1j} = [\mathbf{b}_{1,j} \mathbf{b}_{2,j} \dots \mathbf{b}_{p,j}] \quad (9.86)$$

Donde

$$\mathbf{b}_{i,j} = W_{i,j} (P_{i,j}^T W_{i,j})^{-1} (T_{i,j}^T T_{i,j})^{-1} T_{i,j}^T \mathbf{f}_{0,j} \quad (9.87)$$

Las matrices $P_{i,j} \in R^{\min(m,n) \times j}$, $W_{i,j} \in R^{\min(m,n) \times j}$, $T_{i,j} \in R^{n \times j}$ son formadas de apilar los vectores $\mathbf{p}_{i,j}$, $\mathbf{w}_{i,j}$, y $\mathbf{t}_{i,j}$, respectivamente.

A.9.3.3 REDUCCIÓN DE ORDEN

Es importante tener un número apropiados de factores PLS seleccionados en orden de obtener una buena predicción, ya que un número demasiado alto (el valor máximo teórico para a es el rango de X) causará ampliación del ruido y pobre desempeño del monitoreo del proceso. Una manera estándar de determinar la reducción de orden apropiada, se denota como c , es aplicar la validación cruzada usando la **predicción de suma de cuadrados residuales** (PRESS). El orden c es fijado de manera que sea el orden en que PRESS sea mínimo. La debilidad de este enfoque es que requiere que los datos sean divididos en dos partes (el grupo de entrenamiento y de prueba), con los vectores PLS computados basados únicamente en los datos del grupo de entrenamiento.

En el caso de diagnóstico de fallas, un enfoque alternativo es seleccionar el valor c que minimice el criterio de información. Para determinar c , los vectores PLS son construidos usando la totalidad de los datos, después los vectores PLS son aplicados a todos los datos para calcular errores de clasificación para cada opción de reducción de orden, donde la rata de errores de clasificación es definida para que sea la razón del número de clases incorrectamente asignadas al total de

clasificaciones hechas (el número de observaciones en el grupo de entrenamiento).

Para cada factor $j = 1, 2, \dots, \min(m, n)$, el vector score estimado $\hat{\mathbf{t}}_j$ y la matriz residual E_j son

$$\hat{\mathbf{t}}_j = E_{j-1} \mathbf{w}_j \quad (9.88)$$

$$E_j = E_{j-1} - \hat{\mathbf{t}}_j \mathbf{p}_j^T \quad (9.89)$$

Donde $E_0 = X$. Para calcular una predicción del bloque que se predice $Y_{train2,a}$ del set de entrenamiento usando PLS2 con a componentes PLS:

$$Y_{train2,a} = \mathbf{F}_J = \sum_{j=1}^a b_j \hat{\mathbf{t}}_j \mathbf{q}_j^T \quad (9.90)$$

Para PLS1, la predicción del bloque que se predice $Y_{train1,a}$ del grupo de entrenamiento usando PLS1 con a componentes PLS es calculado con

$$Y_{train1,a} = [\mathbf{y}_{train1,a} \mathbf{y}_{train2,a} \dots \mathbf{y}_{trainp,a}] \quad (9.91)$$

Donde

$$\mathbf{y}_{traini,a} = \mathbf{f}_{i,j} = \sum_{j=1}^a b_i \hat{\mathbf{t}}_{i,j} \mathbf{q}_{i,j} \quad (9.92)$$

Alternativamente, una predicción usando PLS2 con a PLS componentes es dada por la ecuación de regresión:

$$Y_{train2,a} = X B_{2a} \quad (9.93)$$

La ecuación anterior es también usada para la predicción alternativa de PLS1 mediante el reemplazo de B_{2a} con B_{1a} .

A.9.3.4 DETECCIÓN DE FALLAS, Y DIAGNÓSTICO

Un modo común de usar PLS es aplicarlo de igual manera que PCA, seleccionando la matriz Y que sea las variables de la calidad del producto. Monitoreando los score PLS de este modo tiene la ventaja sobre los score PCA y es que los score PLS sólo monitorean variaciones en X de las cuales se sabe que están relacionadas con las variables de la calidad del producto. Todas las técnicas de detección, identificación, y diagnosis de fallas para PCA pueden ser aplicadas en exactamente el mismo modo para PLS (ej. Incluidos la Q y T^2 estadística, gráficos de contribución, y análisis discriminante).

En el uso de PLS discriminante, las filas de Y_{train} no tiene la forma $[0,0, \dots, 1 \dots, 0,0]$, que requiere un método para asignar las clases c_k a cada observación k . Un método es asignar c_k para que correspondan al índice de columna cuyo elemento es el más cercano a uno. U segundo método es asignar c_k para que corresponda a la columna cuyo elemento tenga el máximo valor.

El termino **sobrestimación** se refiere al caso donde el elemento de Y_{train} para un miembro en-clase > 1 o el elemento de Y_{train} para un miembro no-clase > 0 . **Subesestimado** es cuando el elemento de Y_{train} para un miembro en-clase < 1 o el elemento de Y_{train} para un miembro no-clase < 0 . Ambos métodos de asignación dan clasificaciones precisas en el caso ideal, que es, cuando ninguno de los elementos de Y_{train} son sobrestimados ni subestimados, y en el caso donde la totalidad de los elementos de Y_{train} subestimados. Si la totalidad de los elementos de Y_{train} son sobrestimados, entonces el primer método de asignación

puede dar altas ratas de clasificaciones erróneas, mientras el segundo método de asignación tenderá a dar buenas clasificaciones. El segundo método de clasificación es preferido por ésta mayor utilidad.

Si alguno de los elementos de Y_{train} son subestimados mientras otros son sobrestimados, cualquiera de los anteriores métodos de asignación se desempeñaran pobremente. Un método para resolver éste problema es llevar cuenta de los subestimados y sobrestimados de Y a un segundo ciclo del algoritmo PLS. El algoritmo NIPALS es corrido por segunda vez para PLS1 y PLS2 reemplazando y_i por y_{train2} , respectivamente. Para distinguir entre el método PLS *normal* y éste método ajustado, PLS1 y PLS2 son denotados como PLS1_{adj} y PLS2_{adj}, respectivamente. El Y que se predice del grupo de entrenamiento usando PLS1_{adj} y PLS2_{adj}, denotado como $Y_{train1,adj}$ y $Y_{train2,adj}$, son obtenidos de modo similar como PLS1 y PLS2, respectivamente.

La efectividad del algoritmo puede ser determinado aplicándolo a un grupo de prueba $X_{test} \in R^{r \times m}$. El bloque que se predice Y_{test1} del grupo de prueba usando PLS1 es calculado usando (9.88) a (9.90) y (9.91) a (9.92) reemplazando X con X_{test} mientras el bloque que se predice Y_{test2} del grupo de prueba usando PLS2 es calculado usando (9.88) a (9.90) reemplazando X con X_{test} . Los bloques que se predicen $Y_{test1,adj}$ y $Y_{test2,adj}$ usando PLS1_{adj} y PLS2_{adj} respectivamente, son obtenidos similarmente.

A.9.3.5 COMPARACIÓN ENTRE PCA Y PLS

Para diagnóstico de fallas, un bloque que se predice Y no es usado en PCA, en cambio una transformación lineal en X es desmpeñada en X tal que los vectores PCA que tienen la mayor clasificación retienen la mayoría de la variación en X . Los score retenidos pueden ser usados con el análisis discriminante para clasificación. La desventaja del enfoque PCA es que los vectores PCA con mayor clasificación pueden no contener el poder discriminatorio necesario para el diagnóstico de fallas.

PCA maximiza la varianza en X mientras PLS maximiza la covarianza entre X y Y . Especificando que en Y incluya la información de fallas como es en PLS discriminante, los vectores PLS son computados de manera tal que entreguen una representación dimensional-baja la cual es correlacionada con diferencias en clases de fallas. Por lo tanto pocos de los vectores PLS discriminantes deberían ser requeridos y bajas ratas en clasificaciones fallidas son obtenidas. Ya que PLS discriminante aprovecha información de las fallas cuando construye estos modelos de baja-dimensión, se espera que el PLS discriminante pueda proveer mejores diagnósticos de fallas que PCA. Sin embargo, esto no es siempre cierto.

A.9.4 ANÁLISIS DE VARIACIÓN CANÓNICA (CVA)

Mediante el análisis de correlación canónica se pueden obtener diferentes grupos de variables como combinación lineal de las variables de input, que maximizan la correlación con otro grupo de variables que representan al conjunto de los outputs. Sobre la base de estas variables, previamente transformadas para facilitar los

análisis de eficiencia, se calcularán los índices de eficiencia relacionados con cada uno de los pares de variables canónicas a través de modelos de tipo estocástico y datos de panel.

El análisis de correlación canónica puede ser un instrumento adecuado para tratarlos supuestos en los que se dispone de información de un conjunto numeroso de variables que pueden agruparse en dos grupos, el primero formado por las variables explicativas y el segundo por las explicadas; y además, dentro de cada grupo se observa un elevado grado de correlación entre las variables.

Formalizando el contenido del anterior párrafo, se estudiaría un fenómeno aleatorio de la realidad, del que puede diferenciarse un primer bloque de aspectos, que son aquellos sobre los que se desea explicar su comportamiento, representado por un conjunto de p variables, denotadas por la matriz Y , de orden $(n \cdot p)$; y un segundo bloque de características, que se consideran explicativas de las variables de la matriz Y , representadas por un conjunto de q variables, denotadas por la matriz X , de orden $(n \cdot q)$.

Si el conjunto de variables de la matriz X es numeroso, y existen correlaciones elevadas entre muchas de ellas, la posibilidad para resolver este inconveniente es eliminar variables que estén altamente correlacionadas, seleccionando las que estén menos correlacionadas entre sí y que además mantengan una influencia sobre las variables endógenas fácilmente interpretable desde un punto de vista económico. Una segunda alternativa es la aplicación de un análisis factorial, recomendable cuando el número de variables con altas correlaciones es elevado y resulta difícil determinar cuáles de ellas constituyen las causas determinantes de la variabilidad de las variables endógenas, cuyo objetivo es reducir la dimensión de la matriz de variables explicativas. El análisis factorial se caracteriza por construir variables que son combinación lineal de las originales, de modo que se maximiza el poder explicativo de las causas comunes que ocasionan la

variabilidad del conjunto de variables explicativas. Además, este análisis garantiza que las variables generadas por este procedimiento no estén correlacionadas entre sí, lo que asegura la inexistencia de multi-colinealidad entre estas *variables ficticias*, pudiendo ser utilizadas como regresores en un modelo de regresión.

El inconveniente de este procedimiento es que no se garantiza que estas *variables ficticias* conserven el poder explicativo sobre las variables que se tratan de explicar. La realización de un análisis factorial sobre cada uno de dos conjuntos de variables (endógenas por un lado y explicativas por otro) garantiza la no correlación entre grupo, pero no una elevada correlación entre grupos de variables explicativas y grupos de explicadas.

El análisis de correlación canónica proporciona variables que garantizan la generación de un conjunto reducido de variables endógenas, y otro de variables explicativas, de manera que ambos grupos estén altamente correlacionados, mientras que las causas comunes entre las variables que forman parte de cada grupo son nulas. Es recomendable su utilización cuando tanto el número de variables endógenas como explicativas es elevado, y existen elevadas correlaciones de las variables de cada grupo entre sí.²⁷

Sean $X = (X_1, \dots, X_p)$, $Y = (Y_1, \dots, Y_q)$ dos vectores aleatorios de dimensiones p y q . Planteemos el problema de encontrar dos variables compuestas

$$U = Xa = a_1X_1 + \dots + a_pX_p, \quad (9.94)$$

$$V = Yb = b_1Y_1 + \dots + b_qY_q, \quad (9.95)$$

Siendo

$$a = (a_1, \dots, a_p)', \quad (9.96)$$

²⁷ EL ANÁLISIS DE CORRELACIÓN CANÓNICA COMO INSTRUMENTO PARA LA EVALUACIÓN DE LA EFICIENCIA.

Moreno Sáez, Alfredo. Trillo del Pozo, David.

$$b = (b_1, \dots, b_p)' \quad (9.97)$$

Tales que la correlación entre ambas $cor(U, V)$, sea máxima.

Mediante este método de reducción de variables para el citado análisis se parte de las matrices X e Y que representan el conjunto de regresores y variables endógenas respectivamente, siendo la matriz: $Z = [X \ Y]$, de dimensión $n \cdot (q + p)$, que tendrá la siguiente matriz de varianzas y covarianzas:

$$S = \begin{Bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{Bmatrix} \quad (9.98)$$

Donde S_{11} es la matriz de varianzas y covarianzas de las q variables de la matriz X , S_{12} es la matriz de covarianzas entre las variables de la matriz X y la matriz Y , S_{22} , es la matriz de varianzas y covarianzas de las p variables de la matriz Y , mientras que S_{21} coincide con la transpuesta de S_{12} .

Podemos suponer

$$var(U) = a' S_{11} a = 1, \quad var(V) = b' S_{22} b = 1.$$

Así el problema se reduce a:

$$\text{Maximizar } a' S_{12} b \text{ restringido a } a' S_{11} a = b' S_{22} b = 1.$$

Los vectores de coeficientes a, b que cumplen esta condición son los primeros vectores canónicos. La máxima correlación entre U, V es la primera correlación canónica r_1 .

Los primeros vectores canónicos satisfacen las ecuaciones

$$S_{12} S_{22}^{-1} S_{21} a = \lambda S_{11} a, \quad (9.99)$$

$$S_{21} S_{11}^{-1} S_{12} b = \lambda S_{22} b, \quad (9.100)$$

Los vectores canónicos normalizados por $a'S_{11}a = b'S_{22}b = 1$, están relacionados por

$$a = \lambda^{-1/2} S_{11}^{-1} S_{12} b, \quad (9.101)$$

$$b = \lambda^{-1/2} S_{22}^{-1} S_{21} a, \quad (9.102)$$

y la primera correlación canónica es $r_1 = \sqrt{\lambda_1}$, donde λ_1 es el primer valor propio de $S_{11}^{-1} S_{12} S_{22}^{-1} S_{21}$.

La correlación es $r_1 = a'S_{12}b$ y como $1 = \lambda^{-1/2} a'S_{12}b$ deducimos que $r_1^2 = \lambda_1$. De hecho, las ecuaciones en valores y vectores propios tienen otras soluciones.

Concretamente hay $m = \min\{p, q\}$ parejas de vectores canónicos $a_1, b_1, \dots, a_m, b_m$, que proporcionan las variables y correlaciones canónicas

$$U_1 = Xa_1, V_1 = Yb_1, r_1 = \text{cor}(U_1, V_1), \quad (9.103)$$

$$U_2 = Xa_2, V_2 = Yb_2, r_2 = \text{cor}(U_2, V_2), \quad (9.104)$$

$$U_m = Xa_m, V_m = Yb_m, r_m = \text{cor}(U_m, V_m), \quad (9.105)$$

Supongamos $r_1 > r_2 > \dots > r_m$. Entonces:

1. Tanto las variables canónicas U_1, \dots, U_m como las variables canónicas V_1, \dots, V_m están incorrelacionadas.
2. La primera correlación canónica $r_1 = \text{cor}(U_1, V_1)$ es la máxima correlación entre una combinación lineal de X y una combinación lineal de Y .
3. La segunda correlación canónica $r_2 = \text{cor}(U_2, V_2)$ es la máxima correlación entre las combinaciones lineales de X incorrelacionadas con U_1 y las combinaciones lineales de Y incorrelacionadas con V_1 .
4. $\text{cor}(U_i, V_j) = 0$ si $i \neq j$.

A.9.4.1 ALGORITMO CVA.

Esta sección describe el modelo de estado-espacio lineal y el algoritmo CVA para identificar modelos de estado espacio proveniente de los datos.

Un aspecto importante del algoritmo CVA es la separación del pasado y el futuro. En el instante particular de tiempo $t \in (1, \dots, n)$, el vector que contiene la información proveniente del pasado es:

$$\mathbf{p}_t = [\mathbf{y}_{t-1}^T, \mathbf{y}_{t-2}^T, \dots, \mathbf{u}_{t-1}^T, \mathbf{u}_{t-2}^T, \dots]^T, \quad (9.106)$$

Y el vector que contiene la información de salida en el presente y futuro es

$$\mathbf{f}_t = [\mathbf{y}_t^T, \mathbf{y}_{t+1}^T, \dots]^T. \quad (9.107)$$

Asumiendo que los datos son generados de un modelo de estado lineal con un número finito de estados k , los elementos del vector de estado \mathbf{x}_t , son iguales a un campo de k combinaciones lineales del pasado.

$$\mathbf{x}_t = J_k \mathbf{p}_t \quad (9.108)$$

Donde

$J_k \in \mathcal{R}^{k \times m_p}$ es una matriz constante con $m_p < \infty$. El vector de estado \mathbf{x}_t tiene como propiedad, que la probabilidad condicional del futuro \mathbf{f}_t condicionado en el pasado \mathbf{p}_t es igual a la probabilidad condicional del futuro \mathbf{f}_t condicionado en el estado \mathbf{x}_t

$$P(\mathbf{f}_t | \mathbf{p}_t) = P(\mathbf{f}_t | \mathbf{x}_t). \quad (9.109)$$

En otras palabras, el estado proporciona mucha información tanto como los datos del pasado como los valores de salida del futuro.

Esto también indica que solo un número finito de combinaciones del pasado afectan las salidas del futuro. Esta propiedad de vector estado puede ser extendida para incluir futuros datos de entrada:

$$P(\mathbf{f}_t | \mathbf{q}_t) | \mathbf{p}_t = P(\mathbf{f}_t | \mathbf{q}_t) | \mathbf{x}_t \quad (9.110)$$

Donde

$$\mathbf{q}_t = [\mathbf{u}_t^T, \mathbf{u}_{t+1}^T, \dots]^T \quad (9.111)$$

Dejando que la memoria de orden k , $\mathbf{m}_t \in \mathcal{R}^k$, este en un campo de k combinaciones lineales del pasado \mathbf{p}_t

$$\mathbf{m}_t = C_k \mathbf{p}_t \quad (9.112)$$

Donde $C_k \in \mathcal{R}^{k \times m}$. El término memoria es usado en lugar de estado porque el vector \mathbf{m}_t puede necesariamente no contener toda la información del pasado (en ocasiones, la dimensionalidad de k puede no ser suficiente para capturar toda la información del pasado). La meta de los procesos de identificación es proporcionar la predicción óptima de los datos de salida del futuro basados en el pasado y el estado normal. Ahora en un proceso real, el estado real de orden k es desconocido, entonces la salida de datos del futuro son predichas basadas en la memoria común:

$$f_t(\mathbf{m}_t) = \Sigma_{fm} \Sigma_{mm}^{-1} \mathbf{m}_t \quad (9.113)$$

Donde f_t es la predicción lineal óptima del futuro f_t basado en la memoria \mathbf{m}_t .

El algoritmo CVA computa la matriz óptima para C_k , esta es la matriz que maximiza la predicción del error promedio.

La estimación óptima para la matriz C_k es igual a J_k , donde J_k es la primera columna k de la matriz J

El óptimo orden k almacenado es $m_t^{opt} = J_k p_t$.

La estructura de la solución indica que el óptimo almacenamiento para el orden k es un sub espacio del almacenamiento óptimo para el orden $k + 1$. El almacenamiento óptimo para un orden k proporcionado, corresponde al primer estado k del sistema, y ese estado es referido como el estado CVA.

A.9.5 APLICACIONES

A.9.5.1 INTRODUCCIÓN

En las primeras secciones de este capítulo varios procesos estadísticos de monitoreo son comparados para aplicaciones en la simulación de una planta industrial. Los métodos idealmente serían ilustrados con los datos recolectados durante fallas específicas conocidas de un proceso industrial real, pero este tipo de datos no está disponible para una planta industrial a gran escala. En cambio, muchos académicos estudian el desarrollo de procesos de monitoreo basados en

los datos recolectados de simulaciones por computador de un proceso industrial. Los métodos de monitoreo de procesos son probados sobre los datos recolectados de la simulación para el Proceso de Tennessee Eastman (TEP²⁸, por sus siglas en inglés). El TEP ha sido ampliamente usado por la comunidad de monitoreo de procesos como una fuente de datos para comparar varios enfoques. El TEP se conforma de cinco unidades principales: un reactor, un condensador, un compresor, un separador vapor/líquido y torre de destilación; y estos contienen ocho componentes: A, B, C, D, E, F, G y H, que se irán describiendo a lo largo de esta sección según sea necesario. (Ver variables en la Tabla 5 y esquema del proceso en la Figura 82).

Los métodos de monitoreo de procesos son probados sobre datos generados por el código de simulación de TEP, operando bajo lazo cerrado con la estructura de control de la planta entera. En la simulación original de TEP se permitía pre-programar 20 fallas que se introducían selectivamente en el proceso. Nosotros agregamos una simulación de falla adicional, teniendo así un total de 21 fallas como se muestra en la tabla 11 Adicional a los aspectos ya mencionados del proceso, el desarrollo del proceso de monitoreo es dependiente de la forma en la que se recolectan los datos, intervalo de muestreo y el tamaño del conjunto de datos.

²⁸El TEP fue creado por la compañía Química Eastman para dar un auténtico proceso industrial que permita evaluar los procesos de control y los métodos de monitoreo.

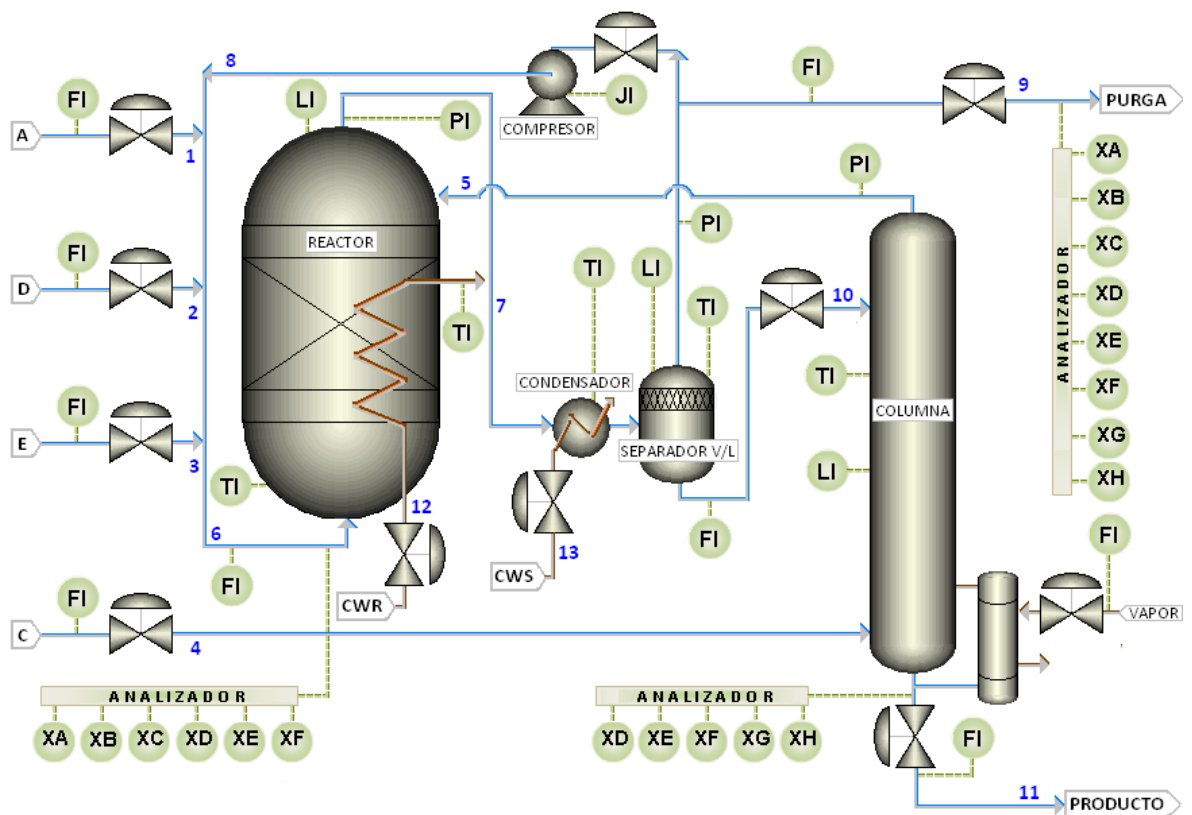


Figura 82 Diagrama de flujo del proceso TEP con estructura de control secundaria.

Fuente: Tesis de maestría en ciencias. Laboratorio Virtual de Procesos. Cinda Luz Sandoval Torres.

El propósito de este capítulo es describir el conjunto de datos y las medidas presentes en el proceso de monitoreo usadas para comparar los métodos presentados en este capítulo. En la sección 9.2 se describe como los datos del entrenamiento y el conjunto de prueba son generados por TEP.

Tabla 11. Fallas del proceso

Variable IDV ²⁹	Descripción	Tipo
(1)	Relación de alimentación A/C, Composición constante B (Corriente 4)	Escalón
(2)	Composición B, Relación constante A/C (Corriente 4)	Escalón
(3)	Temperatura de alimentación D (corriente 2)	Escalón
(4)	Temperatura de entrada del agua fría del reactor	Escalón
(5)	Temperatura de entrada del agua fría del condensador	Escalón
(6)	Pérdida de alimentación A (corriente 1)	Escalón
(7)	Pérdida de cabeza de presión- Capacidad reducida C (corriente 4)	Escalón
(8)	Composición de alimentación A, B, C	Variación aleatoria
(9)	Temperatura de alimentación (corriente 2) D	Variación aleatoria
(10)	Temperatura de alimentación (Corriente 4) C	Variación aleatoria
(11)	Temperatura de entrada del agua fría del reactor	Variación aleatoria
(12)	Temperatura de entrada del agua fría del condensador	Variación aleatoria
(13)	Cinética de reacción	Muy lenta
(14)	Válvula del agua de enfriamiento del reactor	Variación de apertura
(15)	Válvula del agua de enfriamiento del condensador	Variación de apertura
(16)	Desconocida	
(17)	Desconocida	
(18)	Desconocida	

²⁹ Variable perturbada (Disturbance Variables)

(19)	Desconocida	
(20)	Desconocida	
(21)	La válvula de la corriente 4 fue ajustada para permanecer en la posición de estado estable	Posición constante

Fuente: Chiang, Russel and Braatz Fault .Detection and Diagnosis in Industrial Systems

Una discusión sobre como la selección del intervalo de muestras y el tamaño de las mismas afecta el método de monitoreo de procesos se hace en la sección 9.3 y 9.4, respectivamente. En la sección 9.5 se discute la selección del retraso y orden de cada método. En las secciones 9.6, 9.7 y 9.8 se presenta las medidas investigadas para la detección, identificación y diagnóstico de fallas respectivamente. El método de monitoreo de proceso usado para estos propósitos (cubiertos en las primeras secciones de este capítulo) son resumidas en las tablas 15-17 en la cual se muestran como los métodos son relacionados.

A.9.5.2 CONJUNTO DE DATOS

Los datos del conjunto de entrenamiento y de prueba incluidas todas las variables manipuladas y medidas (ver Tablas 12-14), excepto la velocidad de agitación del agitador del reactor para un total de $m= 52$ variables de observación. (La velocidad de agitación no fue incluida puesto que no fue manipulada). Un vector de observación en un instante particular del tiempo está dado por

$$X= [XMEAS(1), \dots , XMEAS (41), XMV(1), \dots , XMV (11)]^T. \quad (9.114)$$

Las observaciones fueron simuladas con una integración escalonada de 1 segundo, y esto no produce ninguna inexactitud numérica. Aunque algunas variables se muestrean continuamente mientras otras tienen algunos retrasos,

esto simplifica la implementación para usar el mismo intervalo de muestreo para cada variable cuando los datos se recogen para el cálculo de múltiples variables medidas del proceso de control. Un intervalo de muestreo de 3 minutos fue usado para recoger los datos simulados para el conjunto de prueba y de entrenamiento.

Tabla 12 Variables Manipuladas

<i>Variable</i>	<i>Descripción</i>
XMV^{30}	
(1)	Flujo de alimentación D (corriente 2)
(2)	Flujo de alimentación E (corriente 3)
(3)	flujo de alimentación A (corriente 1)
(4)	Flujo total de alimentación (corriente 4)
(5)	Válvula de recirculación del compresor
(6)	Válvula de purga (corriente 9)
(7)	Flujo líquido del separador (corriente 10)
(8)	Flujo del producto líquido del extractor (corriente 11)
(9)	Válvula de vapor del extractor
(10)	Flujo de agua fría del reactor
(11)	Flujo de agua fría del condensador
(12)	Velocidad de Agitación

Fuente: Chiang, Russel and Braatz. Fault Detection and Diagnosis in Industrial Systems

³⁰Variables Manipuladas (Manipulated Variables)

Tabla 13. Mediciones del proceso (Intervalos de muestra de 3 minutos)

Variable XMEAS ³¹	Descripción	Unidades
1	Alimentación A (corriente 1)	kscmh
2	Alimentación D (corriente 2)	Kg/hr
3	Alimentación E (corriente 3)	Kg/hr
4	Alimentación total (corriente 4)	kscmh
5	Flujo de recirculación (corriente 8)	kscmh
6	Velocidad de alimentación al reactor (corriente 6)	kscmh
7	Presión del reactor	kPa
8	Nivel del reactor	%
9	Temperatura del reactor	°C
10	Velocidad de corriente de la purga (corriente 9)	kscmh
11	Temperatura del separador	°C
12	Nivel del separador	%
13	Presión del separador	Kpa
14	Flujo de inferior del separador (corriente 10)	m ³ /hr
15	Nivel de la columna	%
16	Presión de la columna	kPa
17	Flujo inferior de la columna (corriente 11)	m ³ /hr
18	Temperatura de la columna	°C

³¹ Variables medidas (Measured Variables)

19	Flujo de vapor a la columna	Kg/hr
20	Trabajo del compresor	KW
21	Temperatura del agua de enfriamiento a la entrada del reactor	°C
22	Temperatura del agua de enfriamiento a la entrada del condensador	°C

Fuente: Chiang, Russel and Braatz *Fault Detection and Diagnosis in Industrial Systems*

Tabla 14 Mediciones de composiciones (Unidades: mol %)

XMEAS	Medida del proceso	Corriente	Intervalos de muestreo
23	Componente A	6	6
24	Componente B	6	6
25	Componente C	6	6
26	Componente D	6	6
27	Componente E	6	6
29	Componente A	9	6
30	Componente B	9	6
31	Componente C	9	6
32	Componente D	9	6
33	Componente E	9	6
34	Componente F	9	6
35	Componente G	9	6

36	Componente H	9	6
37	Componente D	11	15
38	Componente E	11	15
39	Componente F	11	15
40	Componente G	11	15
41	Componente H	11	15

Fuente: Chiang, Russel and Braatz . Fault Detection and Diagnosis in Industrial Systems

Los datos en el conjunto de entrenamiento se constituyen de 22 simulaciones diferentes, donde las cantidades se cambian aleatoriamente entre cada corrida. Una simulación (Falla 0) fue generada sin falla; otra simulación (Falla 21) fue generada ajustando la posición de la válvula de la corriente 4 en la posición de equilibrio; y, cada una de las otras 20 simulaciones (Falla 1-20) fue generada bajo una falla diferente, y fueron listadas en la tabla 5. El tiempo de simulación para cada corrida fue de 25 horas. Las simulaciones empezaron sin falla, y la falla se introdujo a 1 hora de simulación dentro de la corrida. El número total de observaciones que se generó para cada corrida fue $n= 500$, pero solo 480 de ellas fueron recogidas después de la introducción de la falla. Es decir, solamente 480 de las observaciones se usaron para construir las mediciones del proceso de monitoreo.

Los datos en el conjunto de prueba también se constituyen de 22 simulaciones diferentes, donde las cantidades son cambiadas aleatoriamente en cada corrida. Estas simulaciones corresponden directamente a las corridas del conjunto de entrenamiento (Falla 0-21). La simulación empieza sin falla, y la falla se introduce

cuando la simulación lleva 8 horas. El número total de observaciones generadas para cada corrida fue de $n= 960$.

A.9.5.3 INTERVALO DE MUESTREO

La cantidad de tiempo en el que se recogen datos sobre la calidad de los procesos industriales ya sea en operaciones con o sin control es usualmente definido en la práctica. Normalmente, solo existe una pequeña porción del tiempo de operación donde este puede ser determinado con confianza sin que el dato sea dañado y no haya fallas en el proceso. También, los supervisores del proceso generalmente no permiten fallas para permanecer en el proceso por largos periodos de tiempo con el propósito de producir datos útiles en algoritmos de diagnóstico de falla.

Típicamente los datos recogidos durante operaciones de falla son almacenados en bases de datos históricas en las cuales los ingenieros o los operarios diagnostican las fallas, algunas veces después de que estas ocurran, y luego ingresan la información dentro de la base de datos histórica. La cantidad de esos datos disponibles es normalmente fija y el intervalo de muestreo para el método de monitoreo de proceso necesita ser determinada.

Esto es conveniente para detectar, identificar y diagnosticar fallas tan pronto como sea posible. Esto sugiere una alta tasa de muestreo. También, dado un tiempo fijo $T = n\Delta t$, es beneficioso desde el punto de vista de información tener una muestra tan rápido como sea posible ($\Delta t \rightarrow 0, n \rightarrow \infty$). En términos del procesos de monitoreo, sin embargo, hay tres posibles problemas con muestrear tan rápido como se pueda. Por la cantidad de datos producidos, los requerimientos

computacionales pueden exceder la capacidad disponible. Además, el ajuste del modelo se puede concentrar a las frecuencias más altas, donde el ruido de medición es predominante. Al identificar un modelo ARX³² por medio del enfoque de mínimos cuadrados, Ljung muestra como sesgo (bias)³³ se desplaza cuando el muestreo es de alta frecuencia. Este cambio del sesgo para rápidas tasas de muestreo puede ser indeseable, especialmente si las fallas principales afectan baja frecuencia dinámica del proceso. Finalmente, la estadística que ignora la correlación serial generalmente se desarrollará pobremente para cortos tiempos de muestreo.

Escoger el intervalo de muestras para el proceso de monitoreo es usualmente seleccionado basado en el juicio del ingeniero. Para sistemas de identificación, una regla de oro es establecer el intervalo de muestre para una décima de la constante de tiempo del proceso. Considerando que muchas de las constantes de tiempo del problema de Tennessee Eastman son bajo lazo cerrado aparecen cerca de las 2 horas (Ver figura 83), es aconsejable desde el punto de vista de un sistema de identificación para las muestras un intervalo de 12 minutos. Esto, sin embargo, no aprovecha la instrumentación del proceso, el cual es capaz de tomar muestras mucho más rápido. Un intervalo de muestreo de 3 minutos fue seleccionado aquí para permitir una rápida detección, identificación, y diagnóstico de falla, y para permitir una buena comparación entre las técnicas que pueden tomar en cuenta o ignorar las correlaciones seriales. Además, el mismo intervalo de muestreo ha sido usado en otras aplicaciones del proceso de monitoreo para el TEP.

³² AutoRegressive model with eXternal input

³³Error sistemático.

Un enfoque alternativo podría ser promediar cada medición en un periodo de tiempo antes de usar los datos en un algoritmo de control de procesos. Esto y técnicas similares a “moving window” generalmente reducen la variabilidad del

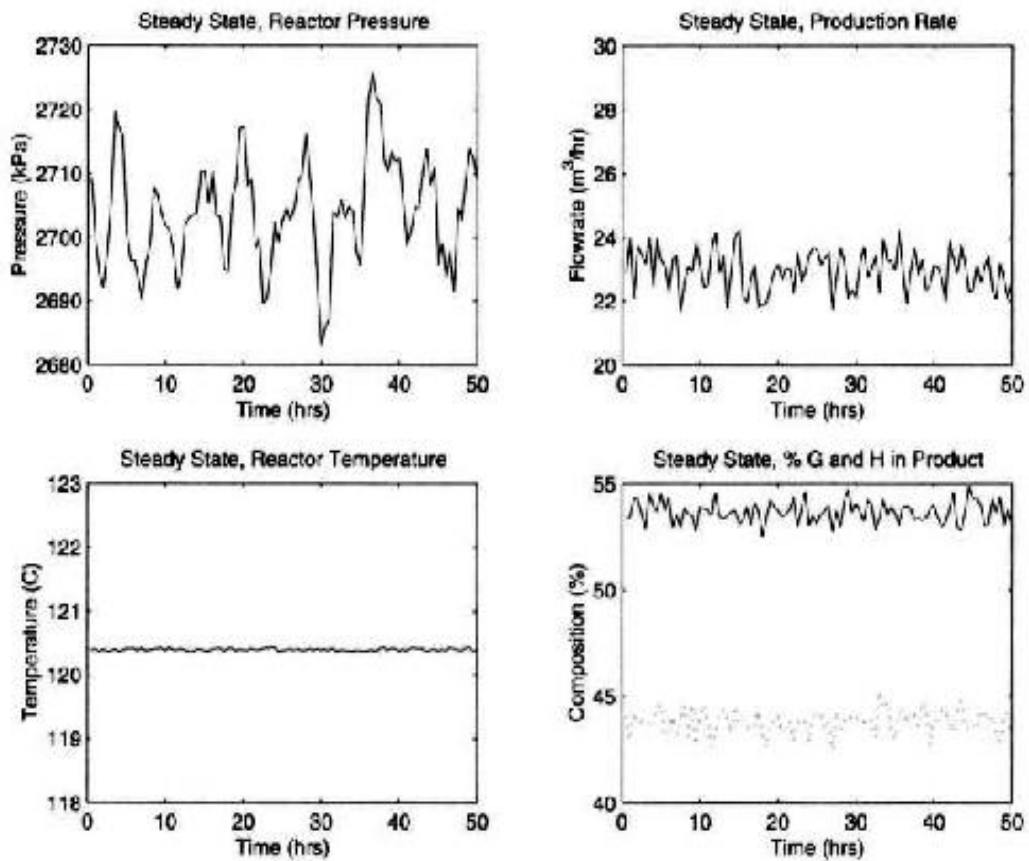


Figura 83. Simulación en lazo cerrado para estado estable sin fallas. Las líneas sólidas y punteadas en la gráfica inferior derecha representan las composiciones de G y H respectivamente.

Fuente: Chiang, Russel and Braatz. *Fault Detection and Diagnosis in Industrial Systems*

proceso y así mismo producen más sensibilidad en el método de monitoreo. Sin embargo, esto puede convertirse en un retraso en la detección de falla. Wise¹² y colaboradores señalan que el ancho de la ventana (es decir, el número de datos usados para computar el promedio) tiene un efecto importante en el rendimiento. En general, una ventana “ancha” permite la detección de cambios muy pequeños, pero no responde tan rápidamente como las ventanas “estrechas”.

A.9.5.4 TAMAÑO DE LA MUESTRA

Como se mencionó en la sección previa, el tiempo total abarcado por el conjunto de entrenamiento es generalmente limitado. En los casos donde el tiempo total $T = n\Delta t$ es fijado, la selección del intervalo de muestreo Δt y el tamaño de la muestra n no pueden ser desacoplados. Así, el efecto del intervalo de muestreo sobre el tamaño de la muestra debería considerarse cuando seleccionamos el intervalo, y viceversa. Una consideración importante para el tamaño de la muestra es el número total de parámetros independientes que contiene el modelo que se están identificando. Esto es deseable para tener un número de parámetros de modelo que sea mucho más pequeño que el número total de variables del proceso m multiplicado por el total de número de observaciones n .

Debido a que los datos de este libro son simulados por el TEP, el tamaño de la muestra no está limitado por T y puede ser considerado separado del intervalo de muestreo. Normalmente se recomienda un tiempo de simulación entre 24 y 48 horas para alcanzar el efecto total de las fallas. Con un intervalo igual a 3 minutos, 24 a 48 horas de simulación contienen $n = 480$ a 960 observaciones. Las simulaciones (ver Figura 84.) sugiere que una corrida que contiene 24 horas de

simulación es suficiente para capturar los desplazamientos significativos en los datos producidos por la falla.

El suficiente tamaño de la muestra de entrenamiento $n=480$ puede ser determinado examinando el número total de parámetros independientes asociados con el orden de varios procesos de monitoreo (ver tabla 15). El número total de estados en el proceso con lazo cerrado es $K=61$; 50 estados de lazo abierto más 11 estados de controladores PI. Para el modelo de espacio estable de orden $K=61$ con 11 entradas y 41 salidas, el número de parámetros independientes M_k es igual a 6985 de acuerdo con (9.115)

$$M_k = k(2m_y + m_u) + m_u m_y + \frac{m_y(m_y + 1)}{2} \quad (9.115)$$

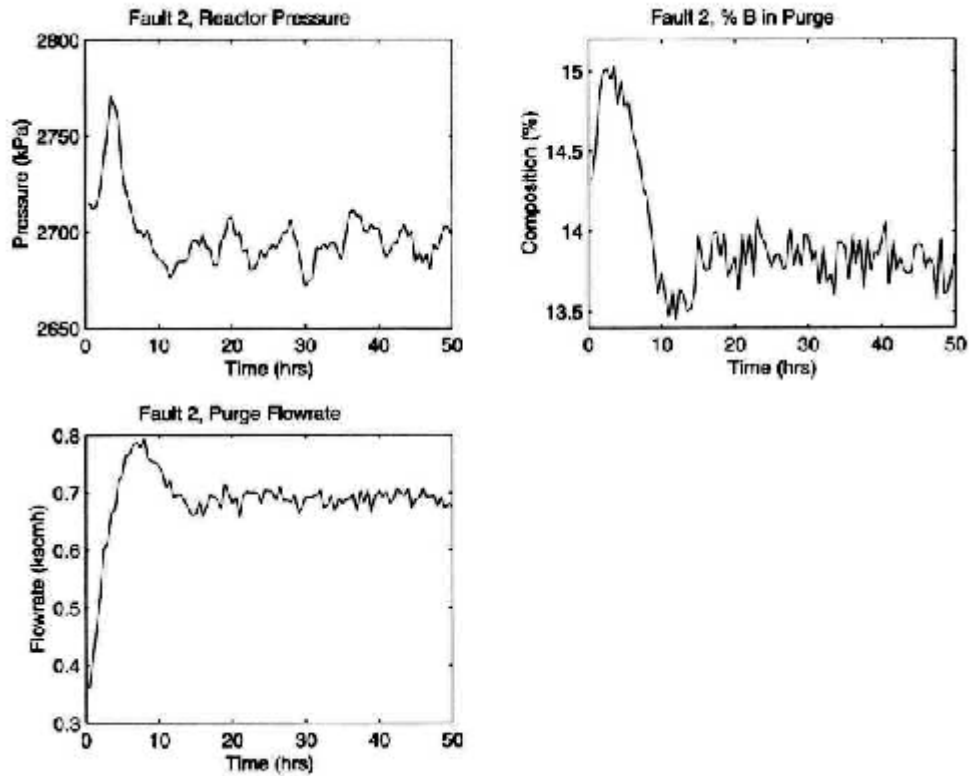


Figura 84.. Simulación en lazo cerrado para cambio escalonado de la composición B (IDV(2) en la tabla 9.7)

Fuente: Chiang, Russel and Braatz. *Fault Detection and Diagnosis in Industrial Systems*

Para la detección de fallas usando PCA basado en estadística T^2 , el número de parámetros estimados M_a es igual al número de grados independientes de libertad de la matriz producto en (9.116), lo cual se calculo así:

$$M_a = \frac{a+2am-a^2}{2} \quad (9.116)$$

Para $a = 51$, el número de parámetros independientes es 1377. Para detección de fallas usando CVA basado en estadística T_s^2 , el número de parámetros estimados

M_k es igual al número de grados independientes de libertad de $J_k^T J_k$, lo cual es calculado así:

$$M_k = \frac{k+2kh-k^2}{2} \quad (9.117)$$

Para $h=2$ y $k=61$, el número de parámetros independientes es 4029. El número total de datos en el conjunto de entrenamiento es igual a $nm = (480)(52) = 24.960$. El requerimiento mínimo absoluto para aplicar PCA, CVA o modelo de espacio estable en un orden dado es que el número de datos sea más grande que el número de parámetros independientes en el modelo. La relación de los números de datos con el número de parámetros independientes es $nm/M_k = (480)(52) / 6985 = 3.57$ para modelos de espacio estable, $nm/M_k = 18.1$ para modelo basado en PCA, $nm/M_k = 5.53$ para modelo basado en CVA. Con todas las variables en igualdad de condiciones (por ejemplo, nivel de ruido), cuanto mayor que uno sea la relación, mejor será la aproximación del modelo. Para este conjunto de datos, todas las relaciones son mayores que uno, indicando que el tamaño de los datos de entrenamiento ($n = 480$) es suficiente para aplicar PCA, CVA, y modelos de espacio estable. La reducción de orden puede todavía dar un modelo de más alta calidad, dependiendo del nivel de ruido. El modelo de espacios estables requiere el mayor número de parámetros independientes, seguido por CVA, y PCA. Un modelo PCA de un orden dado ha reducido significativamente los parámetros independientes, pero no toma en cuenta las correlaciones seriales.

Tabla 15 El número de parámetros independientes estimado para varios modelos y ordenes.

<i>Orden</i> [*]	<i>Entradas</i>	<i>Salidas</i>	<i>Estado Estable</i> ^{**}	<i>PCA</i> ^{***}	<i>CVA</i> ^{****}
1	11	41	1405	52	104
11	11	41	2335	517	1089
21	11	41	3265	882	1974
31	11	41	4195	1147	2759
41	11	41	5125	1312	3444
51	11	41	6055	1377	4029
61	11	41	6985	-	4514

Fuente: Chiang, Russel and Braatz. Fault Detection and Diagnosis in Industrial Systems

* El orden es igual a a y el orden k para modelo de espacio estable y CVA

** El número de parámetros está basado en (9.115)

*** El número de parámetros está basado en (9.116)

**** El número de parámetros está basado en (9.117), usando $h=2$ retrasos

A.9.5 SELECCIÓN DE RETRASO Y ORDEN

El número de retrasos incluido en los métodos de control DPCA, DFDA, y CVA, pueden afectar significativamente el rendimiento. Lo mejor es escoger para el número de retrasos el mínimo necesario para captar la dinámica exacta del proceso. Escogiendo el número de retrasos más grandes que lo necesario puede decrecer substancialmente la robustez de las medidas del proceso de control, la dimensionalidad extra captura ruido adicional, lo cual puede dificultar la

caracterización con datos limitados. El procedimiento usado por este libro sigue la sugerencia de Larimore de seleccionar retrasos h tales que minimicen el criterio de pequeña muestra usando un modelo ARX. Esto asegura que el número de retrasos es lo suficientemente grande para captar estadísticamente toda la información significativa en los datos.

Como se describió en la primera parte de este capítulo, la selección de la reducción de orden es crítica para desarrollar medidas eficientes para el proceso de control.

Aunque es popularmente referido en la literatura, el método de validación cruzada no es usado aquí por ninguno de los métodos de monitoreo de procesos. La validación cruzada es computacionalmente costosa cuando se trata con varios conjuntos de datos grandes. Y más importante, puede haber problemas con esta técnica cuando existen correlaciones seriales en los datos.

A.10. METODOS BASADOS EN CONOCIMIENTO

A.10.1 INTRODUCCIÓN

Surge como método alternativo al enfoque analítico o basado en modelos cualitativo, donde este enfoque para sistemas a gran escala se convierte en demasiado costoso y requiere bastante tiempo de desarrollo.

Un método alternativo para el monitoreo del proceso es utilizar los métodos basados en conocimiento tales como análisis de tendencias cualitativo (QTA), sistemas expertos, y reconocimiento de patrones. Estas técnicas se basan en modelos cualitativos, los cuales se pueden obtener a través de un modelamiento causal del sistema, conocimiento experto, de una detallada descripción del sistema o ejemplos fallas-síntomas. Las técnicas de análisis causal se basan en el modelamiento causal de las relaciones falla-síntoma. Las relaciones cualitativas y semicuantitativas en estos modelos causales se pueden obtener sin utilizar conocimiento extraído de los principios físicos del proceso. Las técnicas de análisis causal como SDG y el árbol de síntomas son principalmente utilizados para el diagnóstico de fallas. Los sistemas expertos se utilizan para imitar el razonamiento humano de expertos cuando diagnostican fallas. La experiencia de un dominio experto puede ser formulada en términos de reglas, las cuales se pueden combinar con el conocimiento de los conocimientos derivados de los principios físicos del proceso o una descripción estructural del sistema para el diagnóstico de las fallas. Los sistemas expertos son capaces de capturar asociaciones humanas para el diagnóstico que nos son fácilmente transcritas en modelos matemáticos o causales. Las técnicas de reconocimiento de patrones utilizan asociaciones entre patrones de datos clases de fallas sin un modelo explícito de los estados internos del proceso o estructura. Ejemplos incluyen redes neuronales artificiales y mapas auto-organizados. Estas técnicas se relacionan a los métodos utilizados en data-driven en términos de modelar las relaciones entre

patrones de datos y clases de fallas. Las técnicas data-driven son técnicas de reducción de dimensionalidad basadas en una rigurosa estadística multivariable, mientras que las redes neuronales actúan como cajas negras que aprenden los patrones basados completamente en las sesiones de entrenamiento. Cada uno de los enfoques data-driven, analítico y basado en conocimiento tienen fortalezas y limitaciones. La incorporación de varias técnicas para el monitoreo del proceso puede ser beneficioso en muchas aplicaciones.

A.10.2 ANALISIS CUALITATIVO DE TENDENCIAS (QTA)

Las plantas modernas son ricas en datos y pobres en información. Una vasta cantidad de datos del proceso disponible se puede utilizar para evaluar el estado del proceso utilizando características relevantes presentes en los datos medidos.

Las tendencias cualitativas (ej. Incremento, constante, etc.) son la representación más natural de características utilizadas ampliamente en el diagnóstico de fallas.

Cada sistema de diagnóstico que utilice las tendencias del proceso para conseguir la clasificación de las fallas tiene tres componentes:

- i) Un lenguaje para la representación de las tendencias tales como triangular episodes (Cheung y Stephanopoulos, 1990), lenguaje basado en primitivas (Janusz y Venkatasubramanian, 1991) y elementos lineales por tramos(MAh et al., 1992)
- ii) Una metodología para extraer las tendencias tal como método basado en wavelet (Bakshi y Stephanopoulos, 1994^a), uso del wavelet, redes neuronales y el método basado en B-splines
- iii) Una metodología de clasificación y una base de conocimiento para mapear las tendencias-sensor en fallas tales como árbol de decisión, árbol ponderado de síntomas, y árbol de fallas.

El lenguaje basado en primitivas es utilizado para la representación cualitativa de tendencias, usualmente se usan siete primitivas : A(0,0), B(+,+),C(+,0),D(+,-),E(-,+),F(-,0),G(-,-) donde los signos son la primera y segunda derivada respectivamente, Funciona representando las señales medidas como una secuencia básica de formas (tales como aumento, constante, disminución ...) llamadas primitivas. Esta secuencia de primitivas constituyen una tendencia; como se muestra en la figura 85, las primitivas B,D,E, y G son primitivas no lineales. El conjunto de extracción de tendencias más paridad de tendencias es denominado análisis cualitativo de tendencias (QTA).

En otras palabras análisis de tendencias cualitativo QTA (qualitative trend analysis) es una técnica semi-cuantitativa de data-driven que ha sido aplicada para el monitoreo de procesos y FDD.

QTA es una técnica atractiva debido a su origen intuitivo, simplicidad, y habilidad para tratar con sistemas a gran escala.

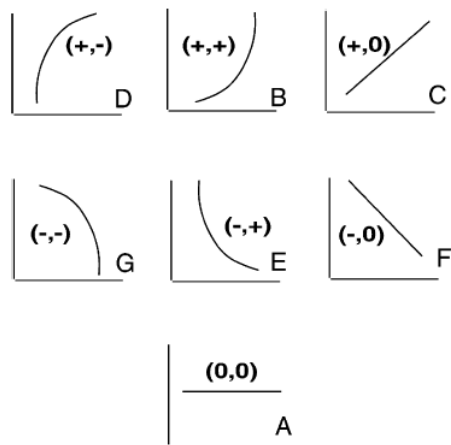


Figura 85 lenguaje fundamental: primitivas

Fuente: Autores

Cualquier señal regular puede ser representada por una secuencia de las siete primitivas.

Limitación.

Aunque QTA provee un diagnóstico rápido y preciso, el incremento de la complejidad computacional de QTA con el incremento del número de sensores utilizados para el diagnóstico.

QTA permite extraer tendencias y calcular similitudes. La idea básica del QTA es representar la señal medida como una tendencia.

El problema con la implementación on-line para un sistema a gran escala es que una gran cantidad de datos a ser procesados, necesitan distinguir entre eventos operacionales y eventos anormales, definición vaga de la operación normal y anotaciones incompletas, entre otras.

Existen dos sub-tareas en QTA

- i) Extracción de tendencias
- ii) Similitud entre tendencias (evaluación de la similitud entre dos primitivas, entre dos tendencias para el mismo sensor y evaluación de la similitud total o índice de confianza).

Para la extracción de tendencias se utilizan algoritmos.

QTA transforma las variables del proceso en descripciones de sus tendencias en una forma explícita y significativa, en tiempo real. Las primitivas son utilizadas en una base de conocimiento para realizar el diagnóstico de fallas.

- i) Utilizando el algoritmo reducción del intervalo a la mitad (interval halving)

La extracción de tendencias mediante el algoritmo reducción del intervalo a la mitad funciona ajustando un polinomio ya sea constante, de primer orden o de segundo orden (en esa secuencia) a los datos y la reducción a la mitad del intervalo es el error ajustado es significativo comparado con el ruido presente en la señal.

Una vez el polinomio es ajustado sobre determinado intervalo, una primitiva es asignada basándose en el signo de la primera y segunda derivada.

Luego el procedimiento de reducción del intervalo a la mitad es aplicado a los datos remanentes hasta que la señal completa es transformada en una secuencia de primitivas.

El algoritmo funciona ajustando ya sea un polinomio de orden cero (cte.) de primer orden o de segundo orden (en esa respectiva secuencia) a la data y reducción a la mitad del intervalo del intervalo si el error ajustado es significativo comparado al ruido presente en la señal.

Una vez el polinomio es ajustado sobre cierto intervalo, se asigna una primitiva según el signo de la primera y segunda derivada.

Luego el procedimiento de la reducción del intervalo a la mitad es aplicado a los datos restantes, hasta transformar la señal completa en una secuencia de primitivas.

Similitud fuzzy entre tendencias.

Requiere el cálculo de

- i) Similitud fuzzy entre dos primitivas
- ii) medida de la similitud entre dos tendencias para el mismo sensor
- iii) Similitud multivariable total o índice de confianza $C.I = \min(S_1, S_2, \dots, S_n)$
n= numero sensores

Extracción de tendencias on –line

El algoritmo básico de reducción del intervalo a la mitad es capaz de extraer primitivas no-lineales y no necesita concatenación.

Las primitivas deberían evolucionar a medida que más datos entran.

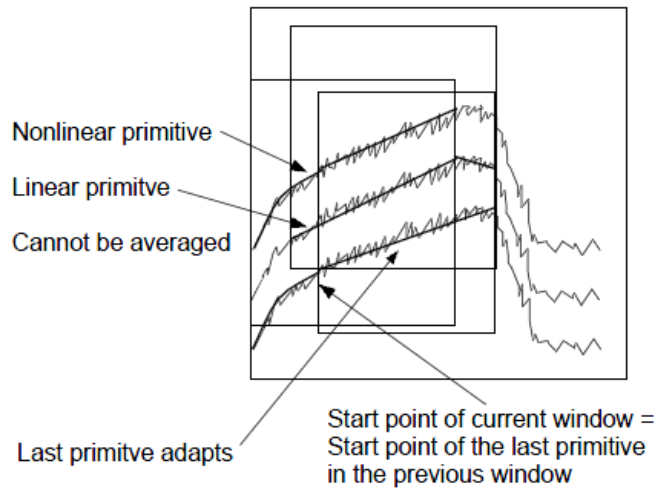


Figura. 86. selección adaptativa de la ventana

Fuente: Autores

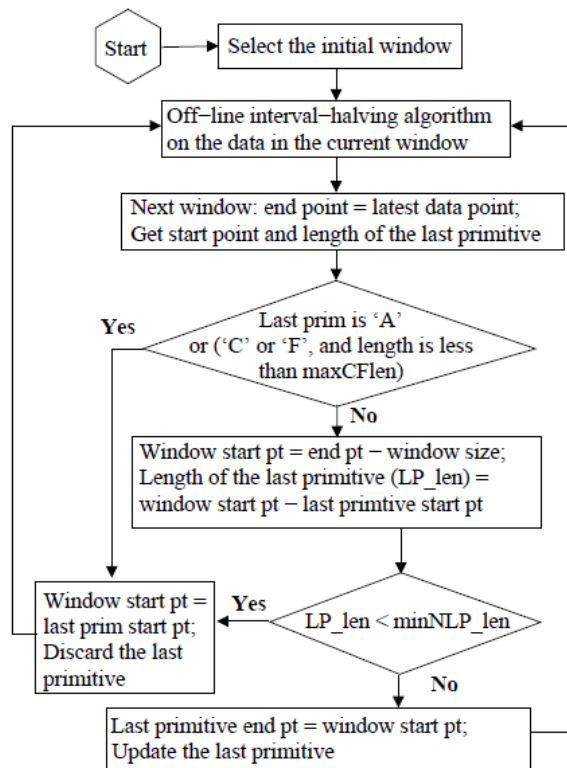


Figura. 87 diagrama de flujo para la extracción de tendencias on-line

Fuente: Autores

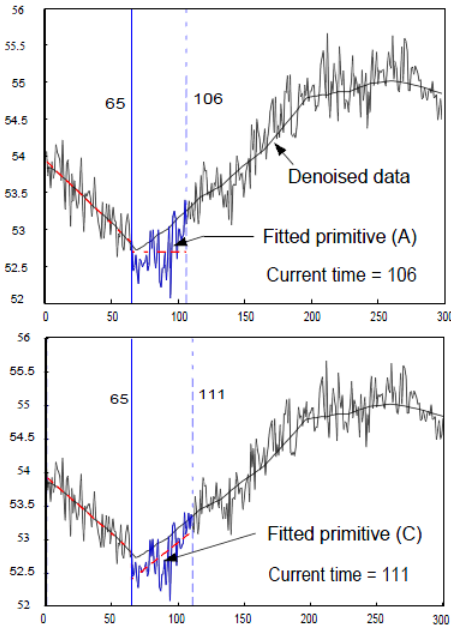


Figura. 88. Instantánea durante dos ventanas consecutivas

Fuente: Autores

A.10.2.1 BASE DE CONOCIMIENTO PARA QTA.

Primero requiere la extracción de tendencias de todas las fallas conocidas

Descripción de las fallas almacenadas, descripción de los sensores utilizados para la extracción de tendencias

Rata de muestreo, valor normal, ruido etc. El punto de inicio y final de los datos utilizados para la extracción de tendencias para cada sensor en cada escenario

de falla, los parámetros globales relacionados con la extracción de tendencias online como la longitud default de la ventana.

- *Similitud entre tendencias FUZZY*

Paridad de tendencias requiere el cálculo de lo siguiente:

- a) Similitud fuzzy match entre dos primitivas
- b) Medida de la similitud entre dos tendencias (el tiempo promedio ponderado de la similitud entre las primitivas involucradas en diferentes intervalos de tiempo) para el mismo sensor.
- c) Similitud multivariable (global) o índice de confianza (C.I) entre dos escenarios (dado por $C.I = \min (S_1, S_2, \dots ,S_n)$ donde n es el número de sensores y S_k es la medida de similitud entre las tendencias del k-esimo sensor en los dos escenarios)

Para realizar el diagnostico de fallas mediante QTA, las fallas almacenadas en la base de conocimiento son ordenados en orden decreciente de su C.I. la falla con el mayor C.I es la falla más, probable en ocurrir. Un bajo valor de C.I (por debajo de 0.5) para todas las fallas, indica la ocurrencia de una nueva falla.

Para grandes plantas se debe considerar la complejidad computacional.

Además se espera que los sensores que están físicamente localizados cerca de la falla origen mostrarían una desviación de su región de operación normal antes que los sensores más lejanos.

Tales sensores son útiles para la detección de la falla y deberían ser escogidos para estimar la medida de similitud.

La selección óptima de los sensores correspondientes a cada falla es importante.

- i) Consistencia entre fallas similares : los sensores deberían mostrar tendencias similares para múltiples manifestaciones de la misma falla
- ii) Discriminación de otras fallas: los sensores deberían ser escogidos de forma que provean la máxima discriminación de otra falla diferente.

Después de la detección es muy importante estimar el tiempo en el cual la falla ocurrió para que una porción apropiada de la infinita secuencia de primitivas pueda ser extraída. Si el tiempo de ocurrencia de la falla no es estimado adecuadamente entonces se obtendrá una medida de similitud pobre incluso para tendencias muy similares.

La metodología para la estimación del tiempo de ocurrencia de la falla es llamada backtracking.

Actividades principales en el FD online

- Detección de la falla
- Estimación del tiempo de ocurrencia de la falla
- similitud entre tendencias computacionalmente eficiente
- Aprendizaje y actualización del QTA-KB

A.10.2.2 DIAGNOSTICO DE FALLAS ON-LINE CON QTA

En la extracción de tendencias off-line, el algoritmo de reducción del intervalo a la mitad es aplicado en la totalidad de los datos. Durante la implementación on-line, más y más datos se encuentran disponibles de las mediciones de los sensores. Las primitivas obtenidas en el tiempo actual pueden no representar correctamente la tendencia en un tiempo futuro. Así un elemento clave de un algoritmo para la extracción de tendencias on-line es que las primitivas deberían actualizarse a medida que más datos se encuentren disponibles. Por un lado, la extracción de tendencias no puede ser efectuada sobre todos los datos disponibles en cada tiempo de muestreo. Por otro lado, si solo se encontraran tendencias correspondientes a los nuevos datos, se terminaría asignando solo primitivas "A" ya que ninguna tendencia útil está contenida en pocas muestras. Así el data set

para la extracción de tendencias debería abarcar de los datos pasados y los nuevos datos disponibles.

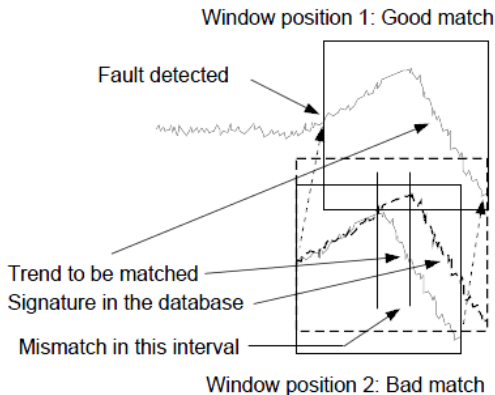


Figura 89. importancia del tiempo de ocurrencia de la falla

Fuente: Autores

Esto crea la pregunta de ¿cómo se debería escoger el segmento de datos para la extracción de tendencias? Y ¿deberían los segmentos de datos escogidos para dos tendencias consecutivas traslaparse? ¿Se puede calcular una primitiva promedio en la región de traslape? Etc.

Para un diagnóstico de fallas en tiempo real se necesita considerar lo siguiente:

- i) Construir una base de conocimiento de fallas -síntoma
- ii) Detectar la ocurrencia de un evento anormal (detección de fallas)
- iii) Extraer la porción relevante de la tendencia de una larga secuencia arbitraria de primitivas.(el tiempo de ocurrencia de la falla juega un rol importante)
- iv) Computo de la medida de la similitud eficiente
- v) Capacidad de aprendizaje/actualización de la base de conocimiento si una nueva falla es manualmente diagnosticada por el operador.

Extracción de tendencias on-line.

Dada los datos del sensor, el algoritmo básico de reducción del intervalo a la mitad puede ser aplicado en la totalidad de los datos en forma off-line para extraer tendencias. La extracción de tendencias debe realizarse sobre un marco (window) de datos, que la ventana debería moverse a medida que más datos se encuentran disponibles y las tendencias extraídas en la ventana actual deben ser combinadas de forma inteligente con las tendencias que ya han sido extraídas.

Antes de tratar generar un algoritmo para llevar a cabo el pre-procesamiento adicional (y posible post-procesamiento) para la extracción de tendencias on-line, analizaremos brevemente otra metodología off-line y su implementación on-line para la extracción de tendencias basada en B-spline.

Para implementar el algoritmo basado en B-Spline (el cual extrae primitivas lineales) para la extracción de tendencias on-line utilizando el enfoque de ventana-deslizante, el algoritmo off-line es aplicada a una ventana que contiene 2^k+1 muestras (k es un entero positivo) después de medir p mas muestras, la ventana se desliza por p datos (el tamaño de la ventana no cambia). Es similar al enfoque de la ventana deslizante para on-line eliminación de ruido utilizando el análisis wavelet.

Para obtener una lista consolidada de primitivas lineales hasta el tiempo actual, las primitivas en la ventana actual se combinan con la antigua lista de primitivas lineales consolidadas. Ya que las primitivas que se combinan son lineales, un valor promedio se puede utilizar como punto de inicio de la ventana actual. Después de este instante de tiempo, las antiguas primitivas son reemplazadas por

las primitivas en la ventana actual. Cada tendencia de tiempo es extraída, después de actualizar la lista de primitivas lineales, son concatenados para obtener primitivas de mayor orden. Las primitivas de alto orden no se actualizan directamente.

Las características principales para la extracción de tendencias del algoritmo off-line reducción del intervalo a la mitad se describirán a continuación. El algoritmo reducción del intervalo a la mitad es capaz de extraer primitivas no lineales directamente y no se necesita la concatenación. Esto quiere decir que si el algoritmo reducción del intervalo a la mitad identifica una primitiva no lineal luego no hay necesidad de permitir a esta primitiva no lineal de evolucionar más excepto cuando la longitud de la primitiva es muy pequeña.

Esto es muy bueno ya que la concatenación requiere un parámetro por cada sensor (magnitud del umbral) y no es transparente.

Dado que uno debería preservar la habilidad de extraer primitivas no –lineales y evitar la concatenación, a diferencia de la implementación on-line de la metodología basada en B-Spline, promediando en el punto inicial de la ventana actual no es una opción. De hecho, promediar convertiría los coeficientes del polinomio en inútiles(los cuales pueden utilizarse para la compresión de datos).

De forma similar, concatenación no es permitida. Aun, las primitivas deberían evolucionar a medida que entran más datos. Dos ideas claves para alcanzar este efecto se explican a continuación.

Evolución sin restricción de la última primitiva de la ventana actual.

Por la naturaleza del algoritmo reducción del intervalo a la mitad, a medida que más muestras se encuentran disponibles, solo la última primitiva evoluciona. Esta asunción puede ser violada para intervalos muy pequeños de tiempo (donde una de las primitivas antes de la última primitiva podría ser una primitiva 'A') debido a la naturaleza del algoritmo de menor número de parámetros y encajar la primitiva más simple. Tales violaciones locales no tienen mucho efecto en la medida de la similitud.

Si la última primitiva es una primitiva lineal ('A', 'C' o 'F') entonces se le permite evolucionar hasta que se convierte en una primitiva no lineal o se vuelve muy larga (de modo que una evolución mayor requeriría demasiados cálculos). Si la última primitiva es no lineal entonces no se le permitiría evolucionar más, excepto cuando su longitud es muy pequeña.

Parámetros y selección de la ventana.

Tres parámetros importantes son la longitud de la ventana por defecto, la longitud de la primitiva más pequeña (minNLP_len) y la longitud de la primitiva lineal más larga (maxCFlen). Para escoger la ventana, el punto final de la ventana actual debería coincidir con la última muestra disponible. El punto inicial de la ventana se puede escoger para que coincida con el punto inicial de la última primitiva. Con esta regla en efecto, el tamaño de la ventana cambia de forma adaptativa.

Una excepción a esta regla es que si la última primitiva es no lineal y es muy larga entonces se puede escoger el punto inicial de la ventana de forma que mantenga el tamaño default de la ventana. Dependiendo sobre cómo se selecciona la ventana, se debería guardar una grabación de las primitivas que no evolucionarían

más y así la lista más actualizada de primitivas puede obtenerse simplemente añadiendo las primitivas en la actual (o última) ventana a la lista existente de primitivas que no evolucionen. El procedimiento esquemático de la selección de la ventana se muestra en la fig.87.. un diagrama de flujo para todo el algoritmo se muestra en la fig 88.

La fig. 89. muestra una instantánea de la extracción de tendencias en dos ventanas consecutivas para una simple señal. Si uno fuera utilizar el algoritmo reducción del intervalo a la mitad en una ventana deslizante de longitud fija, la última primitiva (en el tiempo actual) todavía será 'A'. Además, cálculos innecesarios serían realizados sobre cierta porción de la primera primitiva.

A.10.2.3 CONSTRUYENDO LA BASE DE CONOCIMIENTO QTA

El desarrollo del QTA-KB primeramente requiere la extracción de tendencias para todas las fallas conocidas. Algunos de los parámetros (los cuales son universales para una planta dada) relacionada con QTA-KB son- descripción de las fallas almacenadas (incluyendo si la falla realmente corresponde a un escenario anormal), descripción de los sensores los cuales son usados para la extracción de tendencias, tasa de muestreo, valor normal, ruido... los puntos inicial y final de los datos utilizados para la extracción de tendencias para cada sensor en cada escenario de falla, los parámetros globales relacionados a la extracción de tendencias on-line tales como longitud de la ventana por defecto etc. Para plantas a pequeña escala, donde el número de sensores es pequeño (menos de 100), es posible computar la medida de similitud para todos los sensores con todos los escenarios de fallas en tiempo –real pero para grandes plantas, puede ser inviable. Para grandes plantas, se debe considerar el tema de la complejidad

computacional. Aparte de la inviabilidad computacional, se espera que los sensores que se encuentran localizados físicamente más cerca a la falla originaria deberían mostrar una desviación de la región de operación normal (NOR) antes que los sensores que están localizados más lejos. Así, tales sensores son útiles en la detección de la ocurrencia de la falla y deberían ser escogidos para la estimación de la medida de similitud. Así, la selección óptima de sensores (que deberían ser usados para el cálculo de la medida de la similitud) correspondiente a cada falla es muy importante. Otros elementos en la selección óptima de los sensores son: i) consistencia entre fallas similares: los sensores deberían mostrar tendencias similares para múltiples manifestaciones de la misma falla y ii) discriminación de otras fallas: los sensores deben ser escogidos de forma que provean la máxima discriminación de otras fallas diferentes. Además, cuando se escoge un número menor de sensores, de forma más probable el sensor seleccionado mostraría una evolución rápida. Esto resulta en robustez en el cálculo de similitud, particularmente durante la etapa incipiente de la evolución de la falla, la duración en la cual se está interesado en el diagnóstico de la falla. Ya que el índice de confianza asignado a la falla es el mínimo de la medida de similitud para los sensores dedicados a diagnosticar la falla, si todos los sensores son escogidos arbitrariamente para la evaluación cada falla entonces el correcto diagnóstico será retrasado. Para resumir estas ideas, los sensores dedicados para diagnosticar varias fallas deberían escogerse con respecto a tres criterios:

- i) Consistencia entre fallas similares
- ii) Dinámica rápida para la falla
- iii) Discriminación de otras fallas.

El procedimiento para la detección óptima de sensores se discute a continuación:

- 1) extraer para todos los sensores para todos los escenarios de falla.
- 2) Computar la matriz de similitud global que contiene la medida de similitud de cada sensor para todos los pairs de fallas. Se realizan los pasos 3 a 7 para cada falla.

- 3) Identificar el conjunto de fallas que son similares a esta falla, y el conjunto de fallas que son diferentes.
- 4) En cuanto al conjunto de fallas similares, para cada sensor, extraer la medida de similitud de la matriz de similitud y tomar el mínimo. Así el sensor que muestre la máxima similitud estará en el inicio de la lista (lista 1).
- 5) Con respecto a los pares de esta falla con las diferentes fallas, extraer la medida de similitud para todos los sensores, tomar el máximo para cada sensor, y organizar los sensores en orden creciente del máximo. Así los sensores que muestran menor medida de similitud (máxima discriminación) estarían al inicio de la lista. (lista 2)
- 6) Rank los sensores en orden decreciente de la velocidad de evolución (el sensor que evoluciona más rápido debe estar al inicio) (lista 3).
- 7) Prepara una nueva lista para seleccionar los sensores de la lista 1 y la lista 3 (organizarlos de acuerdo a el criterio ponderado). Esta nueva lista es una lista organizada de sensores en orden decreciente de la consistencia y la velocidad de evolución. Si el valor del criterio ponderado es igual para dos o más sensores en esta lista luego ordenar los de acuerdo a la lista 2 (discriminación decreciente). Ahora selecciona sensores de la nueva lista una tras otra hasta que la falla puede ser resuelta de todas las otras fallas diferentes.

Por supuesto, en el procedimiento de arriba, un enfoque a prueba de fallos debería ser adoptado de forma tal que un número suficiente de sensores son escogidos para asegurar la discriminación de las otras fallas. En algunos casos raros, donde los datos utilizados para el desarrollo de la base de conocimiento no es recolectada apropiadamente, consistencia con fallas similares y discriminación de diferentes fallas pueden estar en conflicto.

En tales casos, una base de conocimiento robusta y confiable no se puede desarrollar.

Diagnóstico de fallas on-line

Durante la implementación on-line, las primitivas son continuamente extraídas. Cuando una falla ocurre los sensores comienzan a desviarse de su valor normal. Así, el primer paso es la detección de fallas. En la detección de fallas basadas en QTA, la presencia de una primitiva diferente de A y la desviación del indicador NOR indica la presencia de una falla. Una metodología multivariante adecuada, también puede ser utilizada para la detección de fallas. Después de la detección, es muy importante estimar el tiempo en el cual ocurrió la falla de forma que una porción apropiada de la secuencia infinita de primitivas puede ser extraída. Si el tiempo de ocurrencia de la falla no se estima adecuadamente, una medida de la similitud pobre puede ser obtenida (figura 5) incluso para tendencias muy similares. Debido a excesiva complejidad computacional, la medición de la similitud con cambios en las tendencias no pueden ser evaluadas en tiempo real y por lo tanto, una buena estimación del tiempo de ocurrencia de fallas es requerido. La metodología utilizada para la estimación del tiempo de ocurrencia de la falla es llamada backtracking. Como se muestra en la figura 88, una vez una falla es detectada, intentamos ajustar una primitiva "A" en el último intervalo. Si una "A" puede ser ajustado luego hay una pequeña variación la cual significa que la falla ocurrida en algún tiempo anterior a este intervalo. So la ventana de estimación es estirada hacia atrás y este procedimiento es repetido sobre la ventana estirada hasta que una primitiva "A" no puede ser ajustada, hay suficiente variación en los datos en la ventana de estimación. Este procedimiento siempre termina cuando la última primitiva no sea una primitiva "A". Las tendencias son recalculadas para los datos después del tiempo de ocurrencia de la falla. Se puede observar que este método no toma en cuenta el tiempo de retraso pero no es un problema porque, para asegurar la estimación robusta de la medida de la similitud, la misma metodología puede utilizarse durante el desarrollo del QTA-KB. A medida que la falla evoluciona, más y más sensores se desvían de su NOR. Estos sensores son utilizados para la estimación de la medición de la similitud y C.I para varias fallas. Esto

asegura que la medida de similitud para un sensor de evolución lenta o un sensor que no muestra suficiente desviación no resultaría en un incorrecto C.I para la falla actual. En resumen, las principales actividades envueltas en el diagnóstico de fallas online son:

- Detección de fallas
- Estimación del tiempo de ocurrencia de la falla
- similitud entre tendencias eficientemente computacional
- Aprendizaje y actualización del QTA-KB

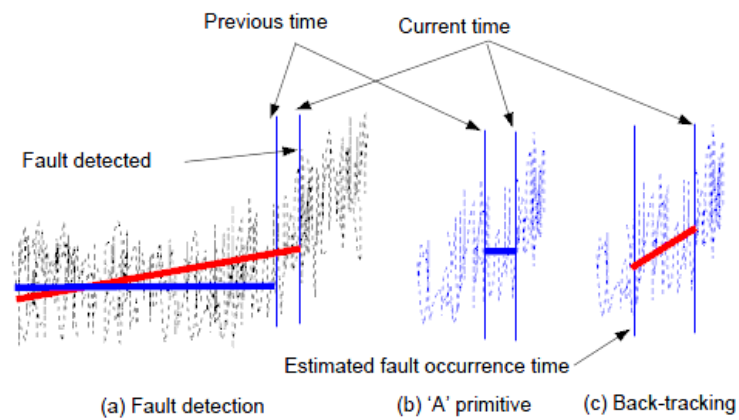


Figura 90. estimación del tiempo de ocurrencia de la falla

Fuente: Autores

A.10.2 SISTEMAS EXPERTOS

Varias aplicaciones de diagnósticos de fallas en el área de la ingeniería han hecho uso de los sistemas expertos. **Los sistemas expertos** son técnicas basadas-conocimiento que son cercanas al estilo en que los humanos resolvemos los problemas. Un sistema experto bien desarrollado es capaz de representar conocimiento experto existente, acomodarse a bases de datos existentes, acumular nuevo conocimiento, hacer inferencias lógicas, hacer recomendaciones, y tomar decisiones con razonamiento. La principal ventaja de usar sistemas expertos es que expertos no necesitan estar presentes para consultas.

Los componentes básicos de un sistema experto incluyen una base de conocimiento, y un motor de inferencia, y un sistema de interface humano/experto. La base de conocimiento puede ser obtenida vía **conocimiento superficial** (basado en testimonio heurístico y experto) y/o **conocimiento profundo** (basado en modelos estructurales, matemáticos, o de comportamiento). Varios tipos esquemas de representación de conocimiento pueden ser usados incluidas las reglas de producción, marcos, y redes semánticas. La veracidad y exhaustividad de la información almacenada en las bases de conocimiento específica el desempeño factible por el sistema experto, para beneficiarse de la nueva base de experiencia y conocimiento, la base de conocimiento necesita también ser actualizada periódicamente. El motor de inferencia dirige el uso de la base de conocimiento. Los mecanismos de inferencia incluyen encadenamiento adelantado-atrasado, métodos de pruebas de hipótesis, búsqueda heurística, reglas-meta, y redes neuronales artificiales. La interface del sistema humano experto debe traducir la entrada de los usuarios a lenguaje computacional y

presentar conclusiones y explicaciones al usuario en un modo de fácil entendimiento.

A.10.2.1 SISTEMA EXPERTO DE CONOCIMIENTO SUPERFICIAL

Un ingeniero con experiencia o del dominio de experto es capaz de diagnosticar fallas en un tiempo mucho más corto que un operador inexperto porque el personal con experiencia ha acumulado conocimiento experimental. Para asistir al personal para diagnosticar fallas, la experiencia de expertos puede ser formulada como un conjunto de reglas IF-THEN, las cuales pueden ser usadas para construir un sistema experto. Esto es referido como **Sistema experto de conocimiento superficial** (también conocido como sistemas de razonamiento empírico). Los métodos no dependen de un entendimiento funcional de la mecánica o física del sistema.

Las ventajas del sistema experto de conocimiento superficial son que su flexibilidad y conclusiones pueden ser fácilmente verificadas y explicadas. Unen observaciones a conclusiones directamente; es por ello, que el conocimiento superficial puede ser aplicado en áreas donde los principios fundamentales o descripciones completas del sistema son escasas, pero soluciones heurísticas son aceptadas.

Los resultados de un sistema experto de conocimiento superficial dependen fuertemente de la adecuación del conocimiento incorporado dentro del sistema experto. Sin embargo, heurístico no garantiza cualquier solución del problema de diagnóstico de fallas, especialmente para situaciones que no han sido del dominio de expertos (ej. Conocimiento fuera del dominio de la experiencia). Como mínimo,

un sistema experto de conocimiento superficial bien desarrollado debería ser capaz de ofrecer soluciones que sean lo suficientemente buenas la mayoría del tiempo.

La mayor dificultad de aplicar sistemas expertos de conocimiento superficial está en el paso de **adquisición de conocimiento**, el cual es el paso de adquirir conocimiento adecuado del dominio de expertos y traducirlo a programas de computadoras. Primero, expertos en los sectores pueden no estar disponibles para únicos escenarios de operación y para plantas nuevas o adaptadas.

Segundo, cuando expertos en el sector están disponibles, puede que ellos no entiendan o sean capaces de explicar claramente como ellos resuelven el problema. Cada sistema experto es una aplicación específica. Desarrollar un sistema experto efectivo desde cero puede consumir tiempo y costo para un sistema a gran escala.

A.10.2.2 SISTEMAS EXPERTOS DE CONOCIMIENTO PROFUNDO

En contraste con los sistemas expertos de conocimiento superficial, los sistemas expertos de conocimiento profundo están basados en un modelo tal como se haría con fundamentos ingenieriles, una descripción estructural del sistema, o una completa descripción del comportamiento de estos componentes en casos de normalidad y de falla. Sistemas expertos de conocimiento profundo son también conocidos como basados en modelos, razonamiento funcional, o sistemas expertos de diagnóstico de principios básicos. Para situaciones nuevas o únicas,

sistemas expertos de conocimiento profundo comúnmente ofrecen información de gran utilidad para diagnóstico de fallas. Conocimiento profundo es comúnmente necesario cuando un problema particularmente difícil es confrontado o una explicación para el proceso de diagnóstico es requerida.

Conocimiento profundo envuelve el uso de razonamiento en información casual y funcional. Conocimiento de los principios que gobiernan el proceso puede ser usado en sistemas expertos de conocimiento profundo. Ecuaciones que gobiernan basadas en leyes físicas proveen un conjunto de restricciones en los valores de las variables del proceso. Violaciones significativas de estas restricciones son indicador de fallas en el proceso. Cada restricción es asociada con el conjunto de fallas que causan violaciones de las restricciones.

Otro método para desarrollar un sistema experto de conocimiento profundo es usar razonamiento causal en el SDG; combinado dichas reglas produce todas las candidatas de fallas posibles. Similar a las técnicas analíticas que dependen fuertemente en principios básicos, un sistema experto de conocimiento profundo es también difícil de desarrollar para sistema complejo a gran escala cuyo modelo matemático no pueda estar disponible.

A.10.2.3 COMBINACIÓN DE SISTEMAS EXPERTOS DE CONOCIMIENTO-SUPERFICIAL Y DE CONOCIMIENTO-PROFUNDO

Un ingeniero experimentado usa una combinación de técnicas para diagnosis, incluyendo una familiaridad con la documentación del sistema, un entendimiento

funcional de los componentes del sistema, un entendimiento de las interrelaciones del sistema, conocimiento del historial de fallas de los componentes, junto a heurísticas numerosas. Esto sugiere que el conocimiento superficial y el conocimiento profundo deberán ser combinados en un sistema experto. El razonamiento de conocimiento profundo es comúnmente necesario para suplementar el conocimiento superficial.

Aunque es costoso obtener un modelo de principios básicos para sistemas a gran escala, modelos de componentes individuales son usualmente disponibles. Dicha información puede ser combinada con conocimiento superficial en orden de diagnosticar fallas efectivamente. Un método para combinar conocimiento superficial y profundo es convertir el conocimiento profundo en normas de producción.

A.10.2.4 TÉCNICAS DE APRENDIZAJE AUTOMÁTICO

La mayor dificultad de utilizar conocimiento superficial es la adquisición de conocimiento. Expertos son usualmente mejores colectando y archivando casos que en expresar la experiencia y casos explícitamente en reglas de producción.

Una manera de resolver éste problema es usar técnicas de aprendizaje automático, en el que conocimiento es automáticamente extraído de los datos.

Información simbólica puede ser integrada en un algoritmo de aprendizaje de red neuronal artificial. Ya que un sistema de aprendizaje permite para extracción de conocimiento y conocimiento previo ser codificado en forma de reglas. Reglas fuzzy pueden ser también usadas para extraer conocimiento de los datos.

A.10.2.5 REPRESENTACIÓN DEL CONOCIMIENTO

La forma simple de representar conocimiento en un sistema experto es usando una serie de reglas IF-THEN para representar el conocimiento experto del sistema. La mayoría de los sistemas expertos industriales usan sistemas basados en reglas, los cuales son compuestos de reglas base, memoria de trabajo, y un intérprete de reglas. La **regla base** es comúnmente particionado en grupos de reglas, llamados apiladores **clusters de reglas**. Cada apilador de reglas codifica el conocimiento requerido para desempeñar cierta tarea. La **memoria de trabajo** es una base de datos que posee datos de entrada, hipótesis inferidas, e información interna sobre el programa. Un **intérprete de reglas** es el mecanismo para seleccionar reglas y evaluar reglas. Las ventajas de los sistemas basados en reglas son que ellos aplican una representación homogénea del conocimiento, permiten que se incremente el conocimiento mediante la adición de reglas, y permiten interacciones no planeadas pero útiles.

En adicción a los sistemas basados en reglas que representen el conocimiento experto, una red semántica puede ser también usada. Una red semántica es un método de representación de conocimiento cuyos conceptos son representados como nodos en la red y las relaciones son representadas como arcos dirigidos (ver figura 91). Debe tener una manera de asociar los significados con la red. Una

manera de hacerlo es asociar un conjunto de programas que operen con descripción en la representación.

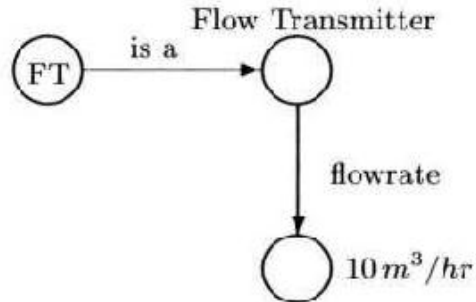


Figura 91. Red semántica representado el conocimiento "FT es un transmisor de flujo con una rata de $10m^3/hr$ "

Fuente: Fault detection and diagnosis in industrial systems.

Alternativamente, marcos pueden ser usados para la representación de conocimiento. Un marco es una colección de redes de nodos semánticas que juntas proveen una representación estructural de un objeto, acto, o evento. Los marcos pueden ser enlazados en jerarquías para mostrar la relación entre objetos de dominio, mientras que las reglas pueden sólo describir indirectamente los objetos que comprenden el dominio.

A.10.2.6 MOTOR DE INFERENCIA

Un motor de inferencia usa un mecanismo de inferencia para reunir la información necesaria (del conocimiento base u del usuario) para dibujar inferencia o conclusiones para el proceso involucrado, y presentar dichas inferencias o conclusiones con explicaciones o bases. El enfoque más comúnmente usado en

los mecanismos de inferencia son de encadenamiento backward/forward (hacia adelante/hacia atrás). En encadenamiento backward, el sistema trabaja hacia adelante para tentativas conclusiones u objetivos para hallar evidencia de apoyo. Encadenamiento backward comienza seleccionando una hipótesis particular, las reglas son examinadas para ver si la hipótesis es consecuente. Si es así, la premisa (también llamada condición, patrón. O antecedente) forma el siguiente conjunto de hipótesis. El procedimiento es continuar hasta que alguna hipótesis sea falsa o todas las hipótesis sean verdaderas en los datos.

En encadenamiento hacia adelante el sistema, las razones del sistema se adelantan de un set de hechos conocidos para inferir conclusiones. El diseño del sistema es una aplicación de encadenamiento-hacia-adelante donde el sistema experto comienza con los requerimientos conocidos, investiga las posibles disposiciones, y hace recomendaciones. La combinación de encadenamiento hacia-adelante Y hacia-atrás son comunes en varias aplicaciones.

Agentes.

Un agente denota por lo general un *sistema* que procesa información y produce una salida de una entrada. Estos agentes pueden estar clasificados de muchas maneras distintas.

En ciencia computacional clásica, los agentes del software son primeramente empleados (Fig. 92.) en este caso consisten de un programa que calcula un resultado de la entrada de un usuario.

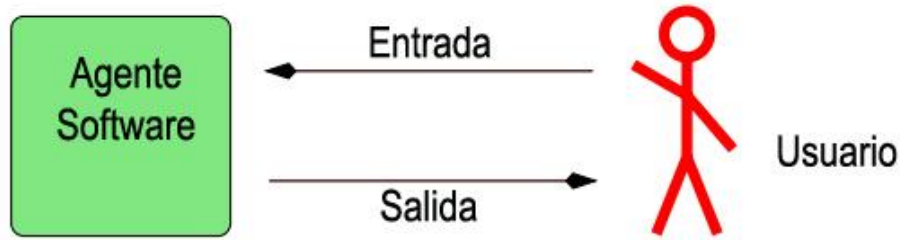


Figura 92. Un agente software con interacción de usuario

Fuente: Introduction to artificial intelligence. Wolfgang Ertel

En robótica, por otro lado, agentes hardware (también llamados robots) son empleados por lo que adicionalmente tiene sensores a su disposición ver (Fig. 93). El agente puede percibir su ambiente con los sensores. Con los actuadores realiza acciones y cambia su ambiente.

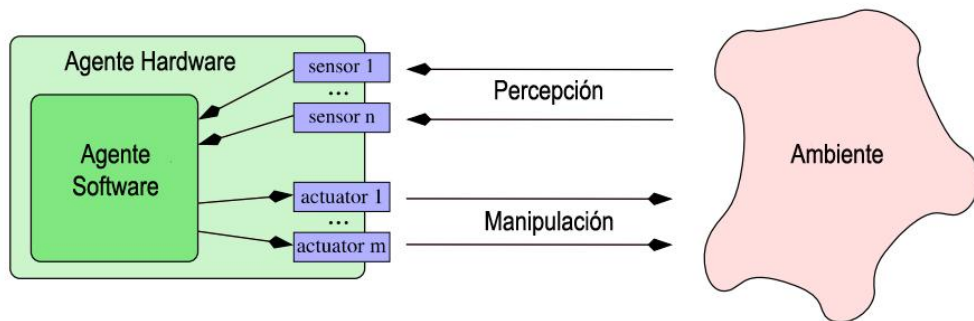


Figura 93. Un agente hardware

Fuente: Introduction to artificial intelligence. Wolfgang Ertel

Sistemas Basados en Conocimiento.

Un agente es un programa que implementa un mapeo de percepciones a acciones. Para aplicaciones complejas en donde el agente debe ser capaz de depender en un gran número de información amontonada lo que significa una **tarea difícil, programar el agente puede costar bastante y no ser claro como proceder.**

Primero se separa *conocimiento* del sistema o programa, que usa el conocimiento, por ejemplo, la búsqueda de conclusiones, consulta de respuesta, o salir con un plan. Este mecanismo es llamado *mecanismo de inferencia*. El conocimiento es almacenado en una *base de conocimiento* (KB). La adquisición de conocimiento en la base de conocimiento es denotado como *conocimiento de ingeniería* y está basado en varias fuentes del saber tales como expertos humanos, el conocimiento ingenieril, y bases de datos. En la Fig. 94. se visualiza la arquitectura general de un sistema basado en conocimiento.

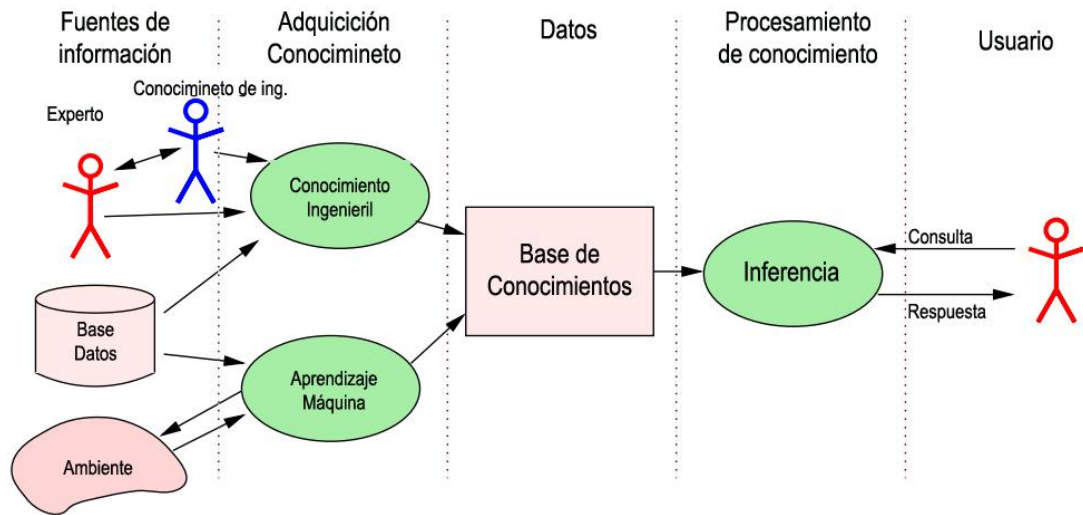


Figura 94. Arquitectura general de un sistema basado en conocimiento.

Fuente: Introduction to artificial intelligence. Wolfgang Ertel

Irse por la separación entre conocimiento e inferencia tiene varias ventajas cruciales. La separación de conocimiento e inferencia le da al segundo la posibilidad de ser implementado en amplio camino de aplicaciones independientes del sistema de inferencia en uso. Sin esta clara separación, conocimiento y procesamiento de pasos de inferencia estarán entrelazados y cualquier chance de conocimiento será costoso. Existen varios tipos de lenguaje, *cálculos proporcionales* y *lógica predicada de primer-orden* (PLS1), *lógica probabilística*, *lógica fuzzy* y *árboles de decisiones*.

Lógica Proposicional.

En lógica proposicional, como el nombre sugiere, las proposiciones son conectadas mediante operadores lógicos. La declaración “*la calle está mojada*” es una proposición, como es “*está lloviendo*” esas dos proposiciones pueden ser conectadas para formar una nueva proposición

Si está lloviendo la calle esta mojada.

Escrito más formalmente

Está lloviendo \Rightarrow La calle esta mojada

Esta notación tiene la ventaja que las preposiciones elementales aparecen de nuevo en forma sin alterar.

Sintaxis.

Definición 10.1 Sean $Op = \{ \neg, \wedge, \vee, \Rightarrow, \Leftrightarrow, (,) \}$ conjunto de operadores y Σ un conjunto de símbolos. Los conjuntos $Op\Sigma$ y $\{t, f\}$ disjuntos por pares. Σ es llamado la *firma* y sus elementos son las *variables de proposición*. El conjunto de fórmulas de lógica proposicional es ahora recursivamente definido:

- t y f son formulas (atómicas).
- Todas las variables de proposición, que son todos los elementos de Σ , son fórmulas (atómicas).
- Si A y B son formulas, entonces $\neg A, (A), A \wedge B, A \vee B, A \Rightarrow B, A \Leftrightarrow B$ sin también formulas.

Definición 10.2 Los símbolos y operadores se leen de la siguiente manera:

t : “verdadero”

f : “falso”

$\neg A$: “no A” (negación)

$A \wedge B$: “A y B” (conjunción)

$A \vee B$: “A o B” (Disyunción)

$A \Rightarrow B$: “si A entonces B” (Implicación)

$A \Leftrightarrow B$: “A si y solo si B” (equivalencia)

Semántica.

En lógica proposicional hay dos valores ciertos: t para “verdad” y f para “falso”. Empezamos con un ejemplo y nos respondemos si la formula $A \wedge B$ son verdaderos. La respuesta es: depende en si las variables A y B son verdaderas. Por ejemplo, si A representa “está lloviendo hoy” y B representa “está haciendo frío hoy” y ambas son verdaderas, entonces $A \wedge B$ y ambas son ciertas, entonces $A \wedge B$ es verdadero. Si, sin embargo, B representa “está haciendo calor hoy” (y esto es falso), entonces $A \wedge B$ es falso.

Diseño de un sistema experto basado en reglas.

La figura 95. muestra los módulos más importantes que conforman un sistema experto basado en reglas. El usuario interactúa con el sistema experto mediante una *interface de usuario* que simplifica y esconde mucha de la complejidad del sistema (ej. La estructura interna de la base de reglas). Los sistema expertos emplean una variedad de estilos de interfaz, incluyendo preguntas y respuestas,

menú-driven, lenguaje natural, o interfaces gráficas, donde la decisión final en el tipo de interface es comprometida entre las necesidades de los usuarios y los requerimientos de la base de conocimiento y el sistema de inferencia.

El corazón del sistema experto es la *base de conocimiento general*, que contiene el conocimiento de resolución de problemas de la aplicación en particular. En un sistema experto basado en reglas este conocimiento es representado en la forma de reglas *if ... then ...*,

El *motor de inferencia* aplica el conocimiento a la solución del problema actual. Es esencialmente un intérprete de la base de conocimiento

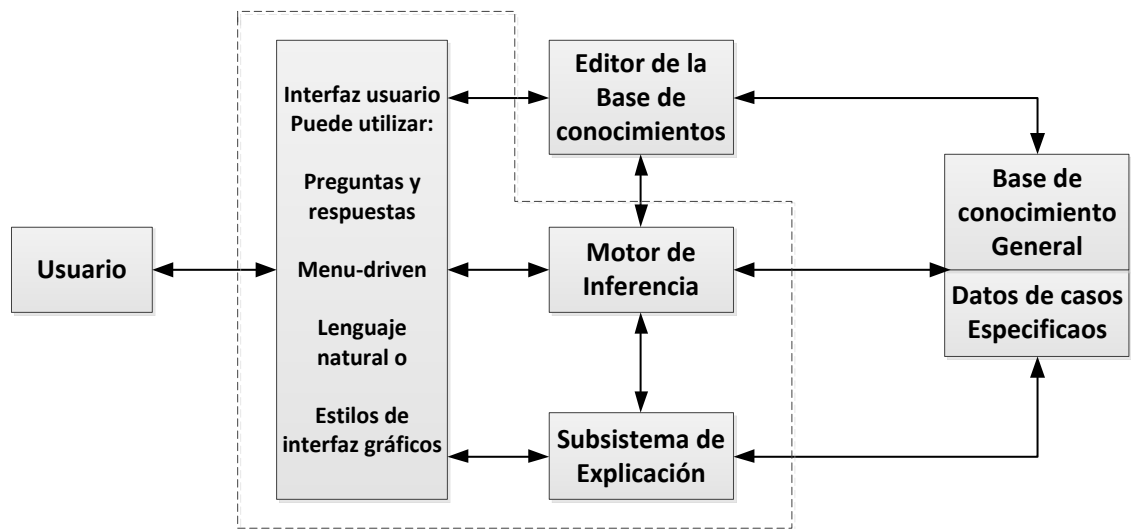


Figura 95. Arquitectura de un sistema experto típico

Fuente: Artificial Intelligence structures and Strategies for complex problems solving. George F. Luger

En el sistema de producción, el motor de inferencia desempeña el acto de reconocimiento del ciclo en control. El procedimiento que implementa el ciclo de control está separado de las reglas de producción por su cuenta. Es importante mantener la separación entre la base de conocimiento y el motor de inferencia por varias razones

- la separación entre el conocimiento para la resolución de problemas y el motor de inferencia hace posibles representar el conocimiento de manera más natural. Las reglas if-then, por ejemplo están más cercanas a la manera en que los humanos describen su resolución de problemas que incorporar este conocimiento en un código de bajo nivel computacional.
- La separación de conocimiento y control, junto con la modalidad provista por las reglas y otras representaciones estructurales usadas en la construcción de bases de conocimiento.. permite hacer cambios en una parte de la base de conocimiento sin crear efectos secundarios en otras partes del programa.
- La separación del conocimiento y los elementos de control permite el mismo control y software de interface sea usado en una variedad de sistemas.

El programa debe hacer seguimiento de los de datos *del caso-específico*. Los hechos, las conclusiones, y otras informaciones relevantes para el caso en consideración.

El subsistema de explicación permite al programa explicar razonablemente al usuario. Dichas explicaciones incluyen justificaciones para las conclusiones del sistema.

Varios programas incluyen el *editor de la base de conocimiento* que ayuda al programador a ubicar y corregir los errores en el desempeño del programa, ellos también ayuda en la adición de nuevo conocimiento, y a mantener las correctas reglas de sintaxis, y realizan consistentemente revisiones a las actualizaciones de la base de conocimiento.

El uso de un armazón de sistemas expertos puede reducir el tiempo en diseño e implementación de un programa considerablemente.

A.10.2.7 SELECCIONADO PROBLEMAS PARA DESARROLLAR SISTEMAS EXPERTOS

Los sistemas expertos por lo general envuelven una considerable inversión de dinero y esfuerzo humano. Los intentos de resolver un problema que sea muy complejo, muy poco entendido, o de otra manera inadecuado para la tecnología disponible.

1. La necesidad para la solución justifica el costo y esfuerzo en construir un sistema experto
2. La experiencia humana no está disponible en todas la situaciones donde es necesaria.
3. El problema puede ser resuelto usando técnicas de razonamiento simbólico.
4. El dominio del problema está bien estructurado y no requiere razonamiento de sentido común
5. El problema no puede ser resuelto usando métodos tradicionales de computación

6. Existan expertos cooperativos y estén articulados.
7. El problema está apropiadamente dimensionado y el alcance es adecuado. Es importante que el problema no exceda las capacidades de la tecnología que se dispone

A.10.2.8 EL PROCESO DE CONOCIMIENTO INGENIERIL

Las primeras persona envuelta en la construcción de un sistema experto son los ingenieros con conocimiento técnico, los expertos en el tema, y usuarios finales

Los ingenieros de conocimiento lenguaje AI y la representación experta, su tarea principal es seleccionar las herramientas de software y hardware para el proyecto, ayudan a los expertos en el tema a articular el conocimiento necesario, e implementan dicho conocimiento correcta y eficientemente en la base de conocimiento.

Los sistemas expertos son construidos mediante aproximaciones progresivas, con los errores del programa conluciendo a correcciones o adiciones a la base de conocimiento. La Figura 96. representa un diagrama de flujo describiendo el ciclo de desarrollo del programa de exploración

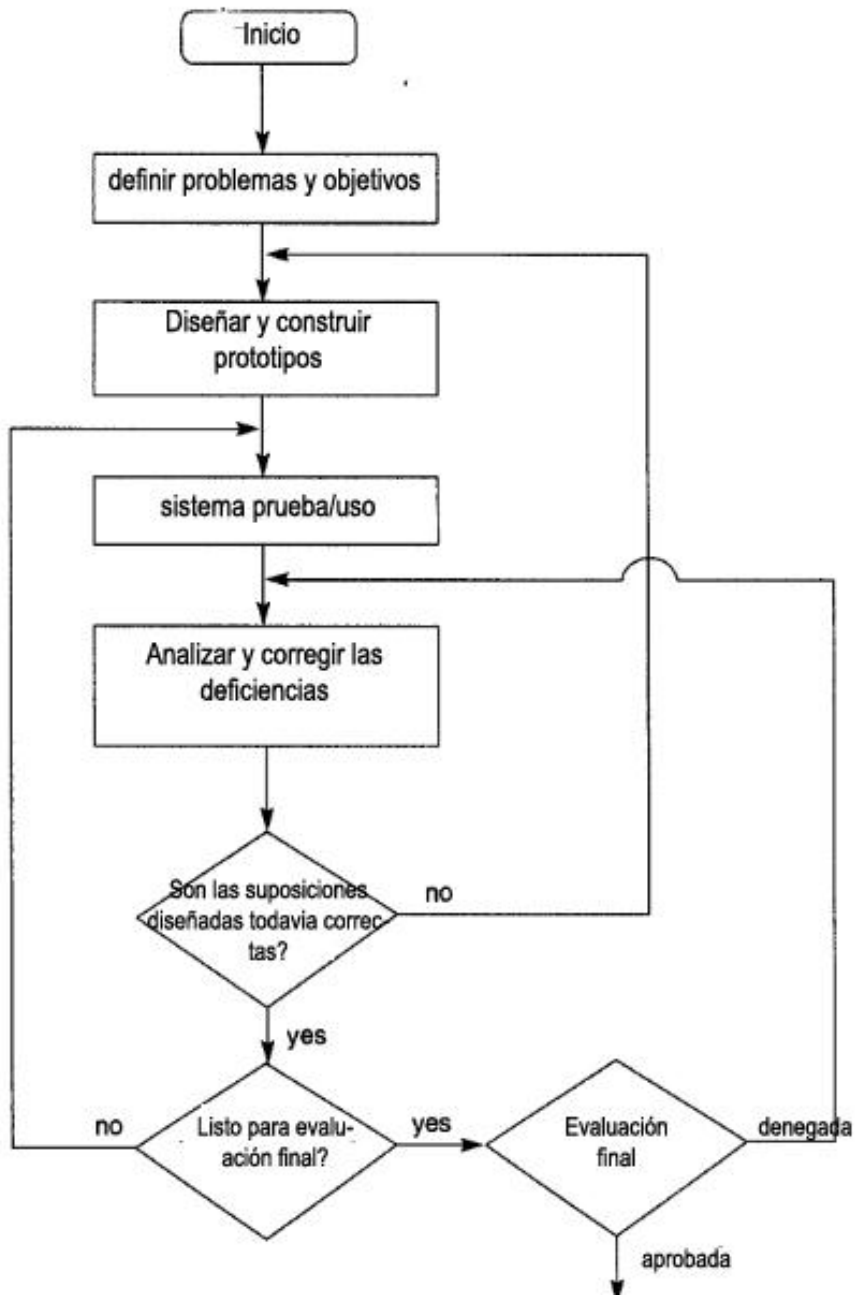


Figura 96. Diagrama de flujo del ciclo de desarrollo del programa de exploración.

Fuente: Artificial Intelligence structures and Strategies for complex problems solving. George F. Luger

A.10.2.9 MODELOS CONCEPTUALES Y SU ROL EN LA ADQUISICIÓN DE CONOCIMIENTO

La figura 97. representa un modelo simplificado del proceso de adquisición de conocimiento que servirá como una primera aproximación útil para entender los problemas envueltos en adquirir y formalizar desempeño de humanos expertos. Los expertos humanos, trabajando en un área de aplicación, operan con dominio del conocimiento, habilidad, y practicidad. Este conocimiento es por lo general impreciso, y sólo parcialmente verbalizados. El ingeniero de conocimiento debe traducir esta experiencia informal en un lenguaje ajustable a un sistema computacional.

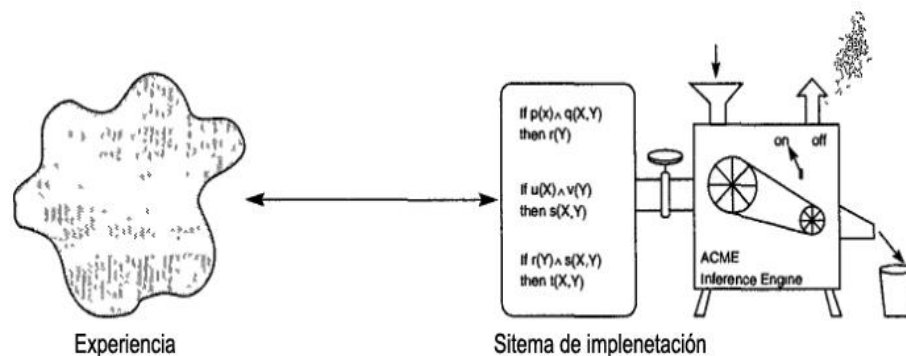


Figura 97. representación de un modelo simplificado del proceso de adquisición de conocimiento

Fuente: Artificial Intelligence structures and Strategies for complex problems solving. George F. Luger

El modelo simplificado de la figura 97. ayuda a articular varios problemas en adquisición de conocimiento:

1. Las habilidades humanas son comúnmente inaccesibles a la conciencia humana.

2. La habilidad humana comúnmente toma la forma de *cómo hacer frente* en una situación en vez de conocer *qué* caracterización racional de la situación puede haber, del desarrollo de mecanismos calificados de rendimiento en lugar de un entendimiento fundamental de que son esos mecanismos.
3. Comúnmente se piensa de la adquisición de conocimiento como obtener conocimiento de los hechos de una realidad objetiva, el llamado “mundo-real”. Tanto en la práctica como la teoría han demostrado, que la habilidad humana representa modelos individuales o comunitarios del mundo.
4. Cambios en la experiencia, No sólo hace ganar a expertos humanos conocimientos nuevos, pero también el conocimiento existente puede estar sujeto reformulaciones radicales, como lo demuestra controversias en curso tanto en campos científicos.

Consecuentemente, el conocimiento ingenieril es difícil y debe ver que abarca el ciclo de vida del sistema experto. Para simplificar esta tarea, es útil tener, como se muestra en la figura 98, *un modelo conceptual* que se encuentre entre la experiencia humana y el programa implementado. Por modelo conceptual, se refiere a la evolución de las ideas del ingeniero de conocimientos en el dominio del conocimiento. Aunque esto es completamente diferente al del dominio de los expertos, es este modelo el que en verdad determina la construcción de la base de conocimiento formal.

A causa de la complejidad y múltiples fuentes de ambigüedad en el problema, no se debe tomar este estado intermediario por garantizado. El ingeniero de conocimiento debe documentar y hacer públicas sus asunciones del dominio mediante metodologías ingenieriles de software comunes.

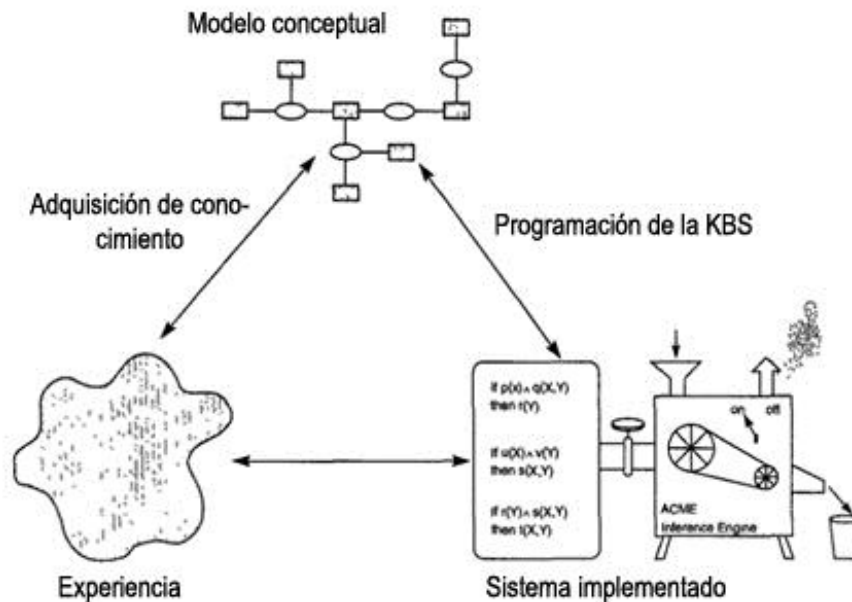


Figura 98. El rol de modelos mentales o conceptuales en solucionadores de problemas

Fuente: Artificial Intelligence structures and Strategies for complex problems solving. George F. Luger

A.10.2.10 EL SISTEMA DE PRODUCCIÓN Y SOLUCIÓN DE PROBLEMAS MEDIANTE GOAL-DRIVEN

La arquitectura de un sistema experto basado en reglas debe ser entendida en términos del sistema de producción para solución de problemas. Los pares de acciones-condiciones son representados como reglas *if ...then ...*, con las premisas de la reglas, la parte *if* corresponde a la condición, y la conclusión *then*. El motor de inferencia implementa el ciclo de acto de reconocimiento del sistema de producción; este sistema puede ser tanto data-driven o goal-driven.

En un sistema experto goal-driven, la expresión objetiva es inicialmente ubicada en la memoria de trabajo. En el sistema coinciden reglas de conclusiones con las de objetivo, seleccionado una regla y ubicando sus premisas en la memoria de trabajo. Esto corresponde a una descomposición de los objetivos del problema en otras más simples. El proceso continua en la siguiente iteración del sistema de producción, con dichas premisas convirtiendo los nuevos objetivos para que coincidan con la reglas de conclusiones. Así, el sistema funciona desde el objetivo original hasta que todos los sub-objetivos en la memoria de trabajo sean verdaderos, indicando que la hipótesis ha sido verificada.

En un sistema experto, los sub-objetivos son comúnmente solucionados haciendo preguntas al usuario por información. La mayoría de estructuras de sistemas expertos permiten al diseñador del sistema especificar que sub-objetivos pueden ser solucionados por preguntas al usuario.

Como un ejemplo más detallado de solución de problemas por goal-driven, considere el siguiente pequeño sistema experto para el diagnóstico de problema en un automóvil:

Regla 1: if

Al motor le llega combustible, y

El motor no se enciende

Then

El problema es de la bujía

Regla 3: if

el motor no enciende, y

Regla 2: if

el motor no se enciende, y

las luces no se encienden

el problema es batería o cables

Regla 4: if

hay gasolina en el tanque, y

las luces se encienden

hay gasolina en el carburador

Then then

El problema es el motor de encendido

al motor le está llegando gasolina.

Para correr esta base de conocimiento bajo un régimen de control de objetivo-dirigido, reemplazar el objetivo de alto nivel, **el problema es X**, en la memoria de trabajo como lo muestra la figura 99. es una variable que puede coincidir con cualquier frase, por ejemplo **el problema es batería o cables**; quedará entrelazado a la solución cuando el problema es resuelto.

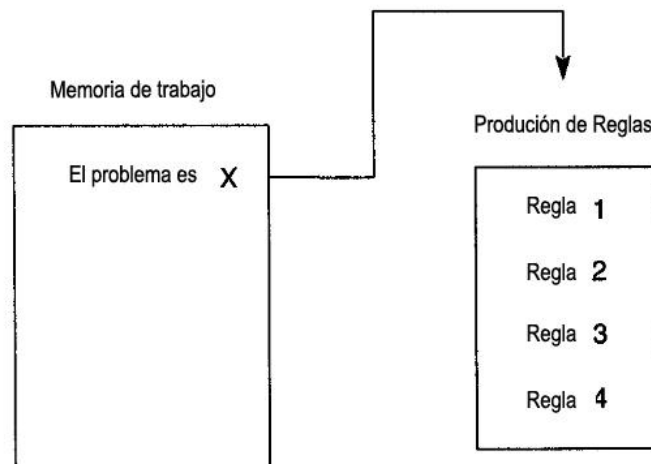


Figura 99. El sistema de producción al comienzo de la consulta del ejemplo de diagnóstico de un automóvil

Fuente: Artificial Intelligence structures and Strategies for complex problems solving. George F. Luger

Tres reglas se acomodan a la expresión en la memoria de trabajo: regla 1, regla 2, y regla 3. Si se resuelve el conflicto a favor de la regla con número más bajo, entonces la regla 1 se activará. Esto causa que **X** este sujeto al valor **bujías** y la premisa de la regla 1 sea reemplazado en la memoria de trabajo como en la

Figura 100. El sistema ha optado por explorar la posible hipótesis que las bujías estén mal.

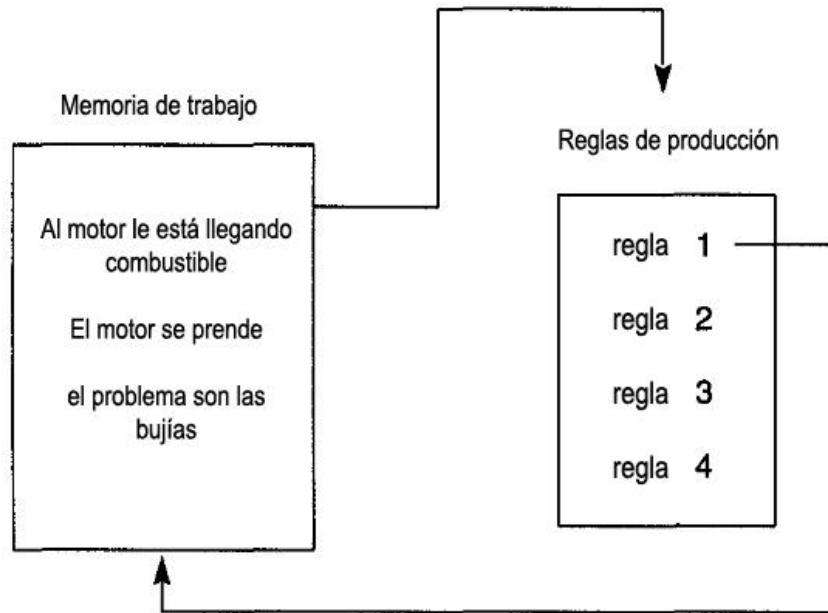


Figura 100. El sistema de producción después que la regla 1 ha sido lanzada

Fuente: Artificial Intelligence structures and Strategies for complex problems solving. George F. Luger

Note que hay dos premisas para la regla 1, ambas que deben ser satisfechas para probar la conclusión verdadera. Hay entonces **y** ramas de la búsqueda gráfica representando la descomposición del problema (encontrando si el problema es de bujías) en dos sub-problemas (encontrando si al motor le está llegando gasolina y si el motor enciende), haciendo que la premisa se reemplace en la memoria de trabajo como lo muestra la figura 101.

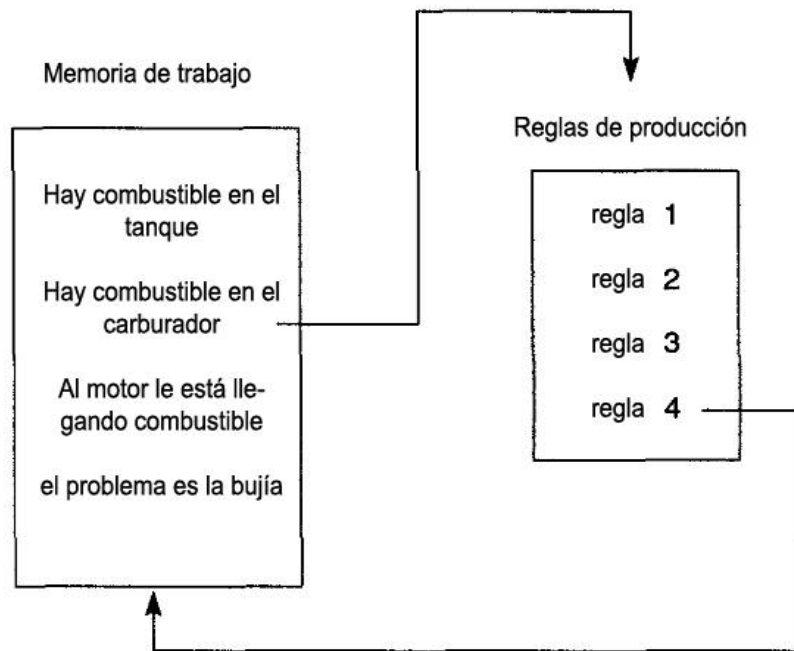


Figura 101. Sistema de producción después que la regla 4 ha sido lanzada

Fuente: Artificial Intelligence structures and Strategies for complex problems solving. George F. Luger

En este punto, hay tres entradas en la memoria de trabajo que no coinciden con regla de conclusión. Nuestro sistema experto entonces, en esta situación, indagará al usuario directamente acerca de sub-objetivos. Si el usuario confirma los tres como ciertos, el sistema experto determinará acertadamente que el carro no arrancará porque las bujías están mal.

A.10.3 RECONOCIMIENTO DE PATRONES.

Algunos métodos basados en datos, analíticos y basados en conocimiento, incorporan técnicas de reconocimiento de patrones en algunas aplicaciones. Algunos métodos de reconocimiento de patrones, para monitoreo de procesos usa la relación entre los patrones de los datos y las clases de fallas sin modelar el estado del proceso interno o su estructura explícita.

Estas aplicaciones incluyen las redes neuronales artificiales (ANN), y los mapas auto-organizativos. Las aplicaciones para el reconocimiento de patrones, están basadas en razonamiento inductivo a través de la generalización de una base de datos recopilada. Estas técnicas son usadas cuando los datos son abundantes, pero cuando el conocimiento experto es escaso. La meta en este trabajo es describir las redes neuronales artificiales y los mapas auto-organizativos, ya que estos dos son las más populares técnicas de reconocimiento de patrones.

Elementos del reconocimiento de patrones:

- Patrones: tras los procesos de segmentación, extracción de características y descripción, cada objeto queda representado por una colección (posiblemente ordenada y estructurada) de descriptores, denominada *patrón*. En los problemas de reconocimiento, cada patrón se supone perteneciente a una categoría o clase, *ci*. El sistema de reconocimiento debe asignar cada objeto (de interés) a su categoría.
- Reconocimiento o clasificación: proceso por el que se asigna una “etiqueta”, que representa una clase, a un patrón concreto.
- Clase: conjunto de entidades que comparten alguna característica que las diferencia de otras.

- Clase de rechazo: conjunto de entidades que no se pueden etiquetar como ninguna de las clases del problema
- Extractor de características: subsistema que extrae información relevante para la clasificación a partir de las entidades cuantificables.
- Clasificador: subsistema que utiliza un vector de características de la entidad cuantificable y lo asigna a una de m clases
- Evaluación del error de clasificación: “error de clasificación”, “tasa de error empírica”, “tasa de rechazo empírica”, “conjunto de datos independientes”.
- Falso rechazo (falso negativo) y falsa aceptación (falso positivo): para problemas de 2 clases, estas definiciones reflejan la importancia de una decisión contra la opuesta. El sistema de clasificación se puede “sintonizar” para que trabaje ponderando un tipo de error sobre el otro.

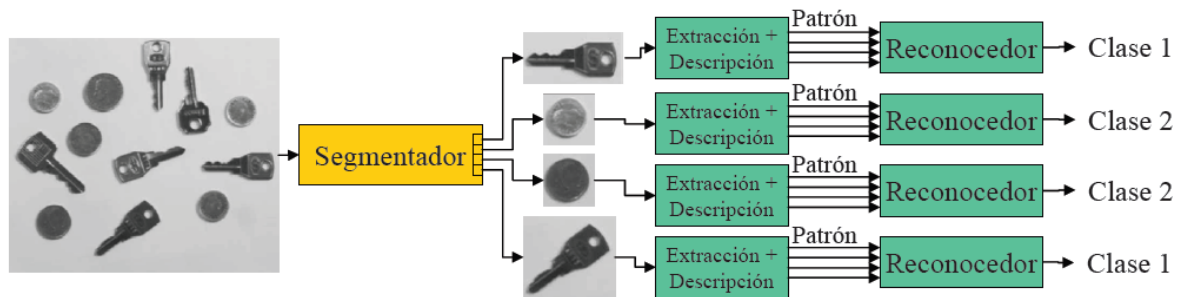


Figura 102. Ejemplo de esquema general para el reconocimiento de patrones.

Fuente: Reconocimiento de Patrones. José Luis Alba -Universidad de Vigo. Jesús Cid -Universidad Carlos III de Madrid. Última revisión: mayo de 2006

Patrones.

Para el reconocimiento automático, es importante hallar patrones que describan objetos de una misma clase, presenten características similares. O también patrones que describan objetos de diferentes clases que presenten características diferenciadas. Entre los tipos de patrones podemos encontrar:

- Vectores: $x = (x_1, x_2, \dots, x_n)^T$
- Cadenas
- Árboles

El mecanismo de generación de patrones se puede representar de forma probabilística:

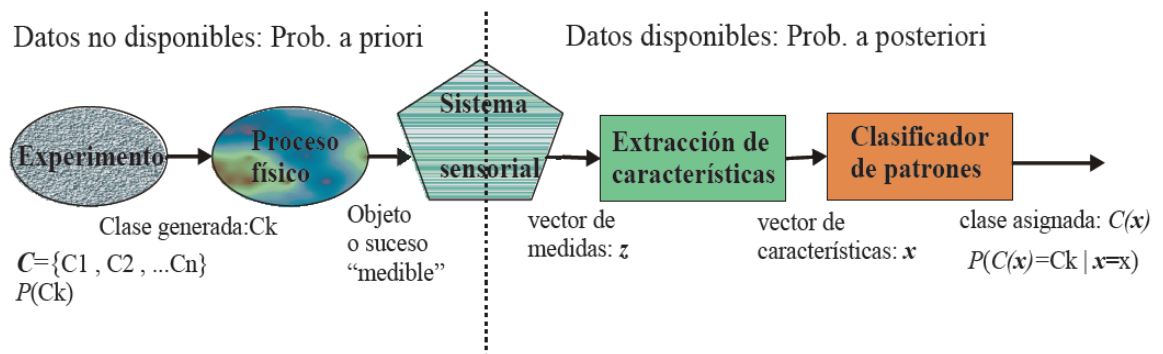


Figura 103. Esquema para el mecanismo de generación de patrones.

Fuente: Reconocimiento de Patrones. José Luis Alba -Universidad de Vigo. Jesús Cid -Universidad Carlos III de Madrid. Última revisión: mayo de 2006

El análisis previo pone de manifiesto que, para diseñar un clasificador estadísticamente óptimo a partir de un conjunto de datos de entrenamiento, pueden adoptarse tres estrategias:

- Estimar $P(c_k|x)$

- Estimar $P(x|c_k)$ y $P(c_k)$
- Determinar una función discriminante que proporcione la frontera de decisión óptima.³⁴

A.10.3.1 REDES NEURONALES.

A.10.3.1.1 Definiciones de una red neuronal.

Existen numerosas formas de definir a las redes neuronales; desde las definiciones cortas y genéricas hasta las que intentan explicar más detalladamente qué son las redes neuronales. Por ejemplo:

- 1) Una nueva forma de computación, inspirada en modelos biológicos.
- 2) Un modelo matemático compuesto por un gran número de elementos procesales organizados en niveles.
- 3) Es un sistema de computación compuesto por un gran número de elementos simples, elementos de procesos muy interconectados, los cuales procesan información por medio de su estado dinámico como respuesta a entradas externas.
- 4) Redes neuronales artificiales son redes interconectadas masivamente en paralelo de elementos simples (usualmente adaptativos) y con organización jerárquica, las cuales intentan interactuar con los objetos del mundo real del mismo modo que lo hace el sistema nervioso biológico.

³⁴ Reconocimiento de Patrones. José Luis Alba -Universidad de Vigo. Jesús Cid -Universidad Carlos III de Madrid. Última revisión: mayo de 2006

Ventajas que ofrecen las redes neuronales.

Debido a su constitución y a sus fundamentos, las redes neuronales artificiales presentan un gran número de características semejantes a las del cerebro. Esto hace que ofrezcan numerosas ventajas y que este tipo de tecnología se esté aplicando en múltiples áreas. Entre las ventajas se incluyen:

- _ Aprendizaje Adaptativo. Capacidad de aprender a realizar tareas basadas en un entrenamiento o en una experiencia inicial.
- _ Auto-organización. Una red neuronal puede crear su propia organización o representación de la información que recibe mediante una etapa de aprendizaje.
- _ Tolerancia a fallos. La destrucción parcial de una red conduce a una degradación de su estructura; sin embargo, algunas capacidades de la red se pueden retener, incluso sufriendo un gran daño.
- _ Operación en tiempo real. Los cómputos neuronales pueden ser realizados en paralelo; para esto se diseñan y fabrican máquinas con hardware especial para obtener esta capacidad.
- _ Fácil inserción dentro de la tecnología existente. Se pueden obtener chips especializados para redes neuronales que mejoran su capacidad en ciertas tareas. Ello facilitará la integración modular en los sistemas existentes.

Aprendizaje adaptativo.

La capacidad de aprendizaje adaptativo es una de las características más atractivas de redes neuronales. Esto es, aprenden a llevar a cabo ciertas tareas mediante un entrenamiento con ejemplos ilustrativos. Como las redes neuronales pueden aprender a diferenciar patrones mediante ejemplos y entrenamientos, no es necesario elaborar modelos a priori ni necesidad de especificar funciones de distribución de probabilidad. Las redes neuronales son sistemas dinámicos auto-

adaptativos. Son adaptables debido a la capacidad de autoajuste de los elementos procesales (neuronas) que componen el sistema. Son dinámicos, pues son capaces de estar constantemente cambiando para adaptarse a las nuevas condiciones.

En el proceso de aprendizaje, los enlaces ponderados de las neuronas se ajustan de manera que se obtengan ciertos resultados específicos. Una red neuronal no necesita un algoritmo para resolver un problema, ya que ella puede generar su propia distribución de pesos en los enlaces mediante el aprendizaje. También existen redes que continúan aprendiendo a lo largo de su vida, después de completado su período de entrenamiento. La función del diseñador es únicamente la obtención de la arquitectura apropiada. No es problema del diseñador el cómo la red aprenderá a discriminar. Sin embargo, sí es necesario que desarrolle un buen algoritmo de aprendizaje que le proporcione a la red la capacidad de discriminar, mediante un entrenamiento con patrones.

Auto-organización.

Las redes neuronales emplean su capacidad de aprendizaje adaptativo para auto-organizar la información que reciben durante el aprendizaje y/o la operación.

Mientras que el aprendizaje es la modificación de cada elemento procesal, la auto-organización consiste en la modificación de la red neuronal completa para llevar a cabo un objetivo específico. Cuando las redes neuronales se usan para reconocer ciertas clases de patrones, ellas auto-organizan la información usada. Por ejemplo, la red llamada backpropagation, creará su propia representación característica, mediante la cual puede reconocer ciertos patrones.

Esta auto-organización provoca la generalización: facultad de las redes neuronales de responder apropiadamente cuando se les presentan datos o situaciones a las que no había sido expuesta anteriormente. El sistema puede generalizar la entrada para obtener una respuesta. Esta característica es muy importante cuando se tiene que resolver problemas en los cuales la información de entrada no es muy clara; además permite que el sistema dé una solución, incluso cuando la información de entrada está especificada de forma incompleta.

Tolerancia a fallos.

Las redes neuronales fueron los primeros métodos computacionales con la capacidad inherente de tolerancia a fallos. Comparados con los sistemas computacionales tradicionales, los cuales pierden su funcionalidad cuando sufren un pequeño error de memoria, en las redes neuronales, si se produce un fallo en un número no muy grande de neuronas y aunque el comportamiento del sistema se ve influenciado, no sufre una caída repentina. Hay dos aspectos distintos respecto a la tolerancia a fallos:

- a) Las redes pueden aprender a reconocer patrones con ruido, distorsionados o incompletos. Esta es una tolerancia a fallos respecto a los datos.
- b) Las redes pueden seguir realizando su función (con cierta degradación) aunque se destruya parte de la red.

La razón por la que las redes neuronales son tolerantes a los fallos es que tienen su información distribuida en las conexiones entre neuronas, existiendo cierto grado de redundancia en este tipo de almacenamiento. Por lo tanto, la mayoría de las interconexiones entre los nodos de la red tendrán sus valores en

función de los estímulos recibidos, y se generará un patrón de salida que represente la información almacenada.

Operación en tiempo real.

Una de las mayores prioridades, casi en la totalidad de las áreas de aplicación, es la necesidad de realizar procesos con datos de forma muy rápida. Las redes neuronales se adaptan bien a esto debido a su implementación paralela. Para que la mayoría de las redes puedan operar en un entorno de tiempo real, la necesidad de cambio en los pesos de las conexiones o entrenamiento es mínima.

A.10.3.1.2 Elementos Básicos.

El elemento básico de un sistema neuronal biológico es la neurona. Un sistema neuronal biológico está compuesto por millones de neuronas organizadas en capas. En la emulación de dicho sistema neuronal biológico, por medio de un sistema neuronal artificial, se puede establecer una estructura jerárquica similar a la existente en el cerebro. El elemento esencial será la neurona artificial, la cual se organizará en capas.

Varias capas constituirán una red neuronal. Finalmente una red neuronal junto con los interfaces de entrada y salida constituirá el sistema global de proceso

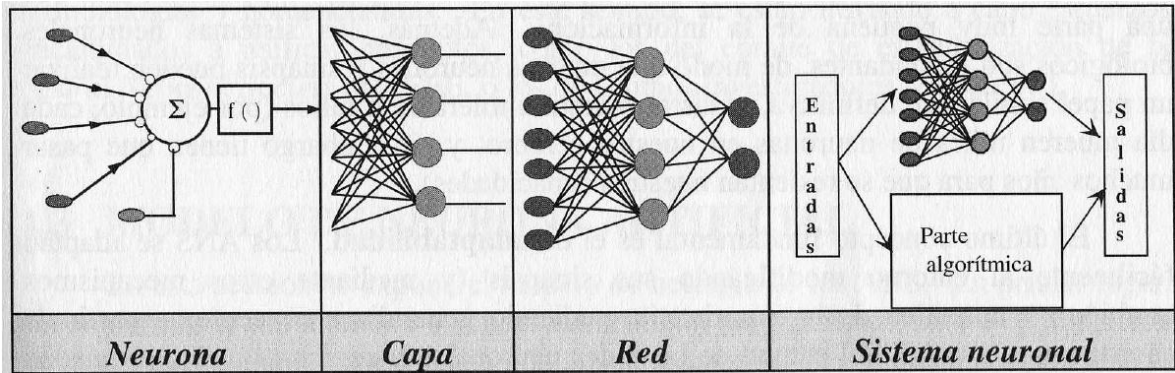


Figura 104. Sistema global de proceso de una red neuronal.

Fuente: Redes Neuronales. Pedro Larrañaga, Iñaki Inza, AbdelmalikMoujahid. Departamento de Ciencias de la Computación e Inteligencia Artificial. Universidad del País Vasco-EuskalHerrikoUnibertsitatea.

A continuación se puede ver, en la Figura 105, un esquema de una red neuronal:

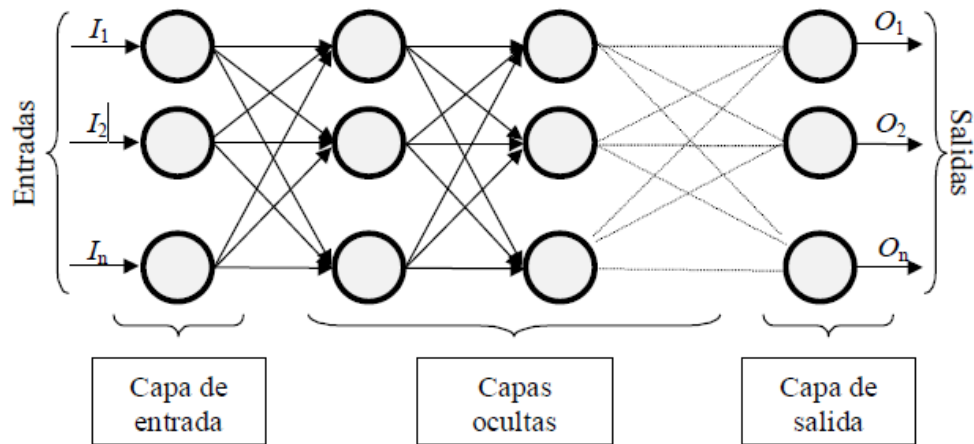


Figura 105. Ejemplo de una red neuronal totalmente conectada.

Fuente: Redes Neuronales: Conceptos Básicos y Aplicaciones. Informática Aplicada a la Ingeniería de Procesos – Orientación I. Carlos Alberto Ruiz.Marta Susana Basualdo. Universidad Tecnológica Nacional – Facultad Regional Rosario,Rosario – Marzo del 2001

La misma está constituida por neuronas interconectadas y arregladas en trescapas (esto último puede variar). Los datos ingresan por medio de la “capa de

entrada”, pasan a través de la “*capa oculta*” y salen por la “*capa de salida*”. Cabe mencionar que la capa oculta puede estar constituida por varias capas.

La distribución de neuronas dentro de la red se realiza formando niveles o capas, con un número determinado de dichas neuronas en cada una de ellas. A partir de su situación dentro de la red, se pueden distinguir tres tipos de capas:

- *De entrada*: es la capa que recibe directamente la información proveniente de las fuentes externas de la red.

- *Ocultas*: son internas a la red y no tienen contacto directo con el entorno exterior. El número de niveles ocultos puede estar entre cero y un número elevado. Las neuronas de las capas ocultas pueden estar interconectadas de distintas maneras, lo que determina, junto con su número, las distintas topologías de redes neuronales.

- *De salidas*: transfieren información de la red hacia el exterior.

En la *Figura 105* se puede ver el ejemplo de la estructura de una posible red multicapa, en la que cada nodo o neurona únicamente está conectada con neuronas de un nivel superior. Notar que hay más conexiones que neuronas en sí; en este sentido, se dice que una red es totalmente conectada si todas las salidas desde un nivel llegan a todos y cada uno de los nodos del nivel siguiente.

A.10.3.1.3 ARQUITECTURAS DE REDES NEURONALES

Se denomina arquitectura a la topología, estructura o patrón de conexionado de una red neuronal. En una red neuronal artificial los nodos se conectan por medio de sinapsis, estando el comportamiento de la red determinado por la estructura de

conexiones sinápticas. Estas conexiones sinápticas son direccionales, es decir, la información solamente puede propagarse en un único sentido (desde la neurona pre sináptica a la pos-sináptica). En general las neuronas se suelen agrupar en unidades estructurales que denominaremos capas. El conjunto de una o más capas constituye la red neuronal.

Se distinguen tres tipos de capas: de entrada, de salida y ocultas. Una capa de entrada, también denominada sensorial, está compuesta por neuronas que reciben datos o señales procedentes del entorno. Una capa de salida se compone de neuronas que proporcionan la respuesta de la red neuronal. Una capa oculta no tiene una conexión directa con el entorno, es decir, no se conecta directamente ni a órganos sensores ni a efectores. Este tipo de capa oculta proporciona grados de libertad a la red neuronal gracias a los cuales es capaz de representar más fehacientemente determinadas características del entorno que trata de modelar. Teniendo en cuenta diversos conceptos se pueden establecer diferentes tipos de arquitecturas neuronales.

A.10.3.1.3.1 Organización de las redes

Así considerando su estructura podemos hablar de redes mono capa, compuestas por una única capa de neuronas, o redes multicapa (Las neuronas se organizan en varias capas). Teniendo en cuenta el flujo de datos, podemos distinguir entre redes unidireccionales (feedforward) y redes recurrentes o realimentadas (feedback). Mientras que en las redes unidireccionales la información circula en un único sentido, en las redes recurrentes o realimentadas la información puede

circular entre las distintas capas de neuronas en cualquier sentido, incluso en el de salida-entrada.³⁵

Redes monocapa.

En las redes monocapa, se establecen conexiones entre las neuronas que pertenecen a la única capa que constituye la red. Las redes monocapas se utilizan generalmente en tareas relacionadas con lo que se conoce como auto asociación (regenerar información de entrada que se presenta a la red de forma incompleta o distorsionada).

Redes multicapa.

Las redes multicapas son aquellas que disponen de un conjunto de neuronas agrupadas en varios (2, 3, etc.) niveles o capas. En estos casos, una forma para distinguir la capa a la que pertenece una neurona, consistiría en fijarse en el origen de las señales que recibe a la entrada y el destino de la señal de salida. Normalmente, todas las neuronas de una capa reciben señales de entrada desde otra capa anterior (la cual está más cerca a la entrada de la red), y envían señales de salida a una capa posterior (que está más cerca a la salida de la red). A estas conexiones se las denomina *conexiones hacia adelante o feedforward*.

Sin embargo, en un gran número de estas redes también existe la posibilidad de conectar la salida de las neuronas de capas posteriores a la entrada de capas

³⁵ Redes Neuronales. Pedro Larrañaga, Iñaki Inza, AbdelmalikMoujahid. Departamento de Ciencias de la Computación e Inteligencia Artificial. Universidad del País Vasco-EuskalHerrikoUnibertsitatea.

anteriores; a estas conexiones se las denomina *conexiones hacia atrás* o *feedback*. Estas dos posibilidades permiten distinguir entre dos tipos de redes con múltiples capas: las redes con conexiones hacia adelante o *redes feedforward*, y las redes que disponen de conexiones tanto hacia adelante como hacia atrás o *redes feedforward/feedback*.³⁶

Conexión entre neuronas.

La conectividad entre los nodos de una red neuronal está relacionada con la forma en que las salidas de las neuronas están canalizadas para convertirse en entradas de otras neuronas. La señal de salida de un nodo puede ser una entrada de otro elemento de proceso, o incluso ser una entrada de sí mismo (*conexión auto-recurrente*). Cuando ninguna salida de las neuronas es entrada de neuronas del mismo nivel o de niveles precedentes, la red se describe como de *conexión hacia delante* (ver *Figura 105*). Cuando las salidas pueden ser conectadas como entradas de neuronas de niveles previos o del mismo nivel, incluyéndose ellas mismas, la red es de *conexión hacia atrás*. Las redes de propagación hacia atrás que tienen lazos cerrados son llamadas *sistemas recurrentes*.

Redes de propagación hacia atrás.

El nombre de "*back propagation*" resulta de la forma en que el error es propagado hacia atrás a través de la red neuronal, en otras palabras el error se propaga hacia

³⁶ *Redes Neuronales: Conceptos Básicos y Aplicaciones. Informática Aplicada a la Ingeniería de Procesos – Orientación I. Carlos Alberto Ruiz. Marta Susana Basualdo. Universidad Tecnológica Nacional – Facultad Regional Rosario, Rosario – Marzo del 2001*

atrás desde la capa de salida. Esto permite que los pesos sobre las conexiones de las neuronas ubicadas en las capas ocultas cambien durante el entrenamiento.

El cambio de los pesos en las conexiones de las neuronas además de influir sobre la entrada global, influye en la activación y por consiguiente en la salida de una neurona. Por lo tanto, es de gran utilidad considerar las variaciones de la función activación al modificarse el valor de los pesos. Esto se llama *sensibilidad* de la función activación, de acuerdo al cambio en los pesos.

A.10.3.1.3.2 Funciones neuronales

Función de entrada.

La neurona trata a muchos valores de entrada como si fueran uno solo; esto recibe el nombre de *entrada global*. Por lo tanto, ahora nos enfrentamos al problema de cómo se pueden combinar estas simples entradas (in_1, in_2, \dots) dentro de la entrada global, gin_i . Esto se logra a través de la *función de entrada*, la cual se calcula a partir del *vector entrada*. La función de entrada puede describirse como sigue:

$$input_i = (in_1 \bullet w_{i1}) * (in_2 \bullet w_{i2}) * \dots (in_n \bullet w_{in})$$

Donde: * representa al operador apropiado (por ejemplo: máximo, sumatoria, productoria, etc.), n al número de entradas a la neurona N_j y w_{ij} al peso.

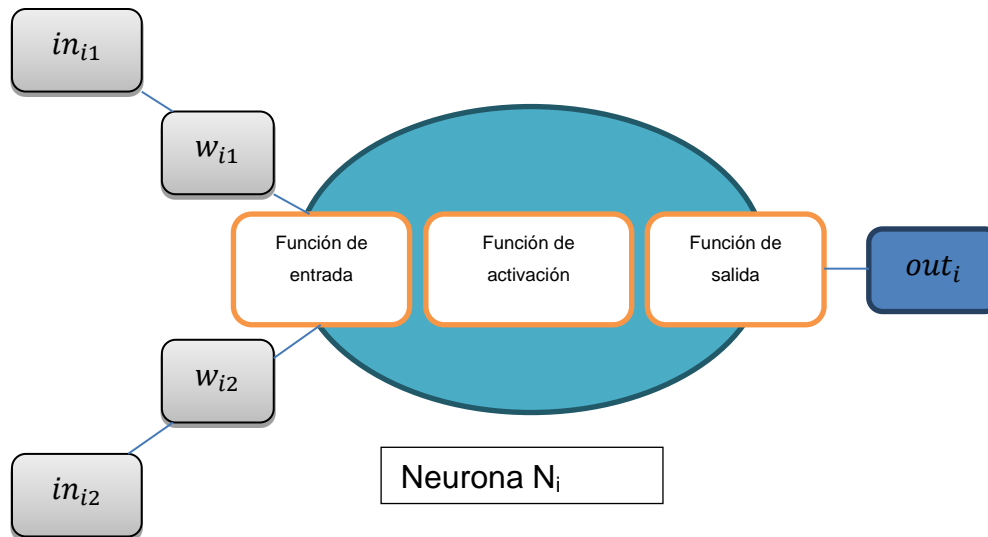


Figura 106. Ejemplo de una neurona con 2 entradas y una salida.

Fuente: *Redes Neuronales: Conceptos Básicos y Aplicaciones. Informática Aplicada a la Ingeniería de Procesos – Orientación I.* Carlos Alberto Ruiz. Marta Susana Basualdo. Universidad Tecnológica Nacional – Facultad Regional Rosario, Rosario – Marzo del 2001

La nomenclatura utilizada en la Figura 106 es la siguiente: in_{i1} = entrada número 1 a la neurona N_i ; w_{i1} = peso correspondiente a in_{i1} ; in_{i2} = entrada número 2 a la neurona N_i ; w_{i2} = peso correspondiente a in_{i2} ; y out_i = salida de la neurona N_i . El conjunto de todas las n entradas $in_i = (in_{i1}, in_{i2}, \dots, in_{in})$ es comúnmente llamado “vector entrada”.

Algunas de las funciones de entrada más comúnmente utilizadas y conocidas son:

1) *Sumatoria de las entradas pesadas*: es la suma de todos los valores de entrada a la neurona, multiplicados por sus correspondientes pesos.

$$\sum_j (n_{ij} w_{ij}), \quad \text{con } j = 1, 2, \dots, n$$

2) *Productoria de las entradas pesadas*: es el producto de todos los valores de entrada a la neurona, multiplicados por sus correspondientes pesos.

$$\prod_j (n_{ij} w_{ij}), \quad \text{con } j = 1, 2, \dots, n$$

3) *Máximo de las entradas pesadas*: solamente toma en consideración el valor de entrada más fuerte, previamente multiplicado por su peso correspondiente.

$$\text{Max}_j (n_{ij} w_{ij}), \quad \text{con } j = 1, 2, \dots, n$$

Función de activación.

Una neurona biológica puede estar activa (excitada) o inactiva (no excitada); es decir, que tiene un “estado de activación”. Las neuronas artificiales también tienen diferentes estados de activación; algunas de ellas solamente dos, al igual que las biológicas, pero otras pueden tomar cualquier valor dentro de un conjunto determinado.

La *función activación* calcula el estado de actividad de una neurona; transformando la entrada global (menos el umbral, θ_i) en un valor (estado) de activación, cuyo rango normalmente va de (0 a 1) o de (-1 a 1). Esto es así, porque una neurona puede estar totalmente inactiva (0 o -1) o activa (1).

La función activación, es una función de la entrada global (gin_i) menos el umbral (θ_i). Las funciones de activación más comúnmente utilizadas se detallan a continuación:

1) *Función lineal*:

$$f(x) = \begin{cases} -1 & x \leq -1/a \\ a * x & -1/a < x < 1/a \\ 1 & x \geq 1/a \end{cases}$$

Con $x = gin_i - \theta_i$, y $a > 0$.

Los valores de salida obtenidos promedio de esta función de activación serán: $a \cdot (gin_i - \theta_i)$, cuando el argumento de $(gin_i - \theta_i)$ esté comprendido dentro del

rango $(-1/a, 1/a)$. Por encima o por debajo de esta zona se fija la salida en 1 ó -1, respectivamente. Cuando $a = 1$ (siendo que la misma afecta la pendiente de la gráfica), la salida es igual a la entrada.

2) *Función sigmoidea:*

$$f(x) = \frac{1}{1 + e^{-gx}}, \quad x = gin_i - \theta_i$$

Los valores de salida que proporciona esta función están comprendidos dentro de un rango que va de 0 a 1. Al modificar el valor de g se ve afectada la pendiente de la función de activación.

3) *Función tangente hiperbólica:*

$$f(x) = \frac{e^{gx} - e^{-gx}}{e^{gx} + e^{-gx}}, \quad x = gin_i - \theta_i$$

Los valores de salida de la función tangente hiperbólica están comprendidos dentro de un rango que va de -1 a 1. Al modificar el valor de g se ve afectada la pendiente de la función de activación.

A.10.3.1.4 ALGORITMOS DE APRENDIZAJE

El proceso de aprendizaje es el proceso por el cual la red neuronal se adapta al estímulo modificando sus pesos y eventualmente produce un resultado esperado. Hay dos tipos de aprendizaje: supervisado y sin supervisión.

A.10.3.1.4.1 Aprendizaje supervisado.

El aprendizaje supervisado incluye un supervisor (o teacher) que le indica a la red cuan cerca esta de aprender y va modificando los pesos neuronales.

El modelo de aprendizaje con profesor (teacher) significa que durante el aprendizaje hay circuitos (o programas) externos a la neurona que comparan el output deseado (T_i) con el actual (O_i) y producen el error entre ellas ($E_i = T_i - O_i$). Usando el error (E_i) un algoritmo de aprendizaje modifica los valores de los pesos uno por uno tratando de minimizar el error, este proceso requiere de muchas iteraciones

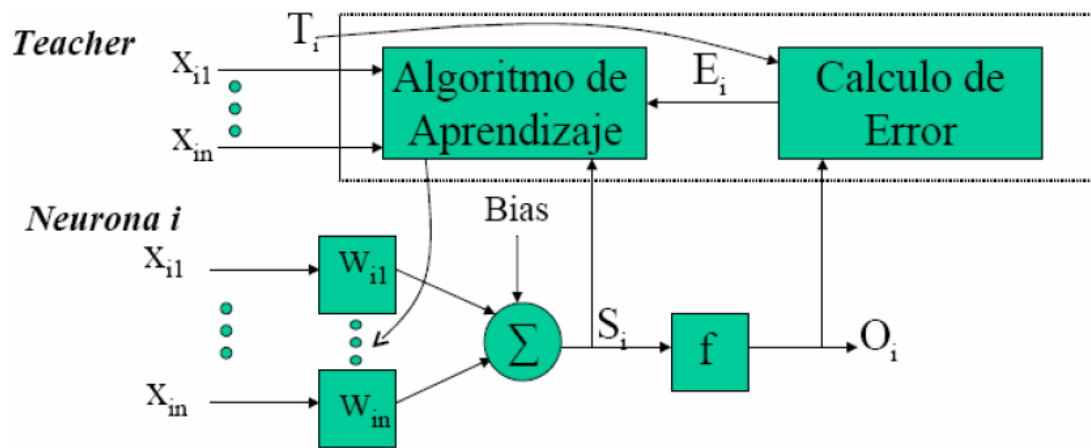


Figura 107. Esquema de aprendizaje supervisado.

Fuente: Introducción a las Redes Neuronales. Tomás Arredondo Vidal. Depto. Electronica UTFSM. 4 mayo 2011

El aprendizaje sin supervisión solamente se forman grupos de clasificación de acuerdo a indicaciones o propiedades, no se calcula un error (E). Se puede usar el error o el gradiente de error para que la red aprenda. La dirección del gradiente es

hacia incrementar el Error por ende se cambian los pesos hacia la dirección inversa.

En este tipo de aprendizaje se suelen considerar, a su vez, tres formas de llevarlo a cabo, que dan lugar a los siguientes aprendizajes supervisados:

- 1) Aprendizaje por corrección de error.
- 2) Aprendizaje por refuerzo.
- 3) Aprendizaje estocástico.

Aprendizaje por corrección de error.

Consiste en ajustar los pesos de las conexiones de la red en función de la diferencia entre los valores deseados y los obtenidos a la salida de la red, es decir, en función del error cometido en la salida.

Aprendizaje por refuerzo.

Se trata de un aprendizaje supervisado, más lento que el anterior, que se basa en la idea de no disponer de un ejemplo completo del comportamiento deseado, es decir, de no indicar durante el entrenamiento exactamente la salida que se desea que proporcione la red ante una determinada entrada.

En el aprendizaje por refuerzo la función del supervisor se reduce a indicar mediante una señal de refuerzo si la salida obtenida en la red se ajusta a la deseada (éxito = +1 o fracaso = -1), y en función de ello se ajustan los pesos basándose en un mecanismo de probabilidades. Se podría decir que en este tipo de aprendizaje la función del supervisor se asemeja más a la de un crítico (que opina sobre la respuesta de la red) que a la de un maestro (que indica a la red la respuesta concreta que debe generar), como ocurría en el caso de supervisión por corrección del error.

Aprendizaje estocástico.

Consiste básicamente en realizar cambios aleatorios en los valores de los pesos de las conexiones de la red y evaluar su efecto a partir del objetivo deseado y de distribuciones de probabilidad. En el aprendizaje estocástico se suele hacer una analogía en términos termodinámicos, asociando a la red neuronal con un sólido físico que tiene cierto estado energético. En el caso de la red, la energía de la misma representaría el grado de estabilidad de la red, de tal forma que el estado de mínima energía correspondería a una situación en la que los pesos de las conexiones consiguen que su funcionamiento sea el que más se ajusta al objetivo deseado. Según lo anterior, el aprendizaje consistiría en realizar un cambio aleatorio de los valores de los pesos y determinar la energía de la red (habitualmente la función energía es una función de Liapunov). Si la energía es menor después del cambio, es decir, si el comportamiento de la red se acerca al deseado, se acepta el cambio; si, por el contrario, la energía no es menor, se aceptaría el cambio en función de una determinada y preestablecida distribución de probabilidades.

A.10.3.1.4.2 Aprendizaje no supervisado.

Las redes con aprendizaje no supervisado (también conocido como auto supervisado) no requieren influencia externa para ajustar los pesos de las conexiones entre sus neuronas. La red no recibe ninguna información por parte del entorno que le indique si la salida generada en respuesta a una determinada entrada es o no correcta.

Estas redes deben encontrar las características, regularidades, correlaciones o categorías que se puedan establecer entre los datos que se presenten en su entrada.

Existen varias posibilidades en cuanto a la interpretación de la salida de estas redes, que dependen de su estructura y del algoritmo de aprendizaje empleado.

En algunos casos, la salida representa el grado de familiaridad o similitud entre la información que se le está presentando en la entrada y las informaciones que se le han mostrado hasta entonces (en el pasado). En otro caso, podría realizar una clusterización (clustering) o establecimiento de categorías, indicando la red a la salida a qué categoría pertenece la información presentada a la entrada, siendo la propia red quien debe encontrar las categorías apropiadas a partir de las correlaciones entre las informaciones presentadas.³⁷

A.10.3.2 MAPAS AUTO-ORGANIZATIVOS (SOM)

Introducidos por TeuvoKohonen como KohonenMaps, Un mapa SOM es un tipos de red neuronal que se entrena sin supervisión para producir una salida discretizada de baja dimensión de las entradas (ej. un mapa). Usan una función de vecindad para preservar propiedades del espacio de entrada. Las redes SOM operan en dos modos, entrenamiento y mapeo, en entrenamiento se elije la

³⁷ *Redes Neuronales: Conceptos Básicos y Aplicaciones. Informática Aplicada a la Ingeniería de Procesos – Orientación I. Carlos Alberto Ruiz.Marta Susana Basualdo. Universidad Tecnológica Nacional – Facultad Regional Rosario,Rosario – Marzo del 2001*

neurona que mas aproxima al vector de entrada como la solución para esa entrada y se repite iterativamente hasta tener neuronas asociadas a cada entrada. En la etapa de mapeo se entrega una entrada y la red la clasifica. El procedimiento para poner un vector desde el espacio de datosal mapa es encontrar el nodo con el vector de pesos más similaral vector de entrada y asignarle los valores del vector de entrada.

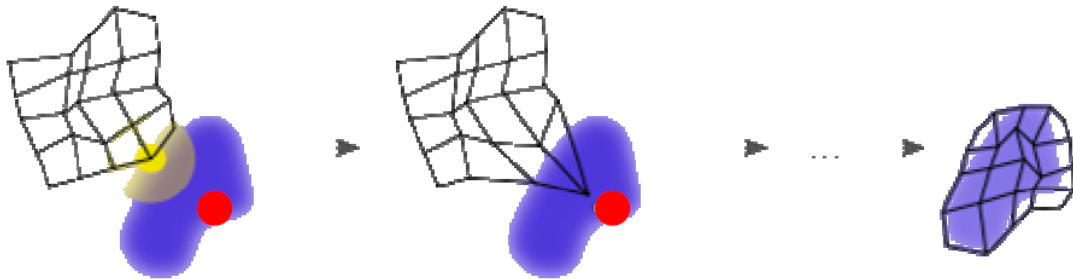


Figura 108. Ilustracion de modo de operar del mapa auto-organizativo

Fuente: Introducción a las RedesNeuronales. Tomás Arredondo Vidal.Depto.Electronica UTFSM.4 mayo 2011

En el ejemplo, se ve como el nodo más cercano (en amarillo) al vector objetivo (en rojo) es asignado ese valor, se puede ponderar el grado de acercamiento de los nodos dada su distancia (ej. Euclideana) al vector (ej. punto) objetivo. Eventualmente la red SOM tiende a tomar los valores de los datos de entrada. Usando aprendizaje competitivo, cuando un ejemplo se muestra a la red, su distancia Euclidiana a todos los vectores de pesos de los nodos se calcula. El nodo (ej. neurona) mas similar se denomina la Best Matching Unit (BMU), los pesos de la BMU y las neuronas cercanas se ajustan en dirección del vector de entrada del ejemplo.

La magnitud del cambio es decreciente con tiempo y distancia ala BMU. La formula de actualización para una neurona con vector de pesos $Wv(t)$ es:

$$Wv(t + 1) = Wv(t) + \theta(v, t) \alpha(t)(D(t) - Wv(t)), \quad (10.1)$$

Donde $\alpha(t)$ es un coeficiente de aprendizaje decrecientemonotonicamente y $D(t)$ es el vector de entrada. La función de cercanía $\theta(v, t)$ depende de la distancia entre elBMU y la neurona v . En los casos más simples es 1 para todas las neuronas cerca de la BMU y 0 para las otras pero una funciónGaussiana es otra posible elección³⁸.

A.10.4 COMBINACIÓN DE VARIAS TÉCNICAS

Cada técnica de monitoreo de procesos tiene sus fortalezas y limitaciones. Se ha hecho esfuerzos por desarrollar esquemas basados en combinación de técnicas basadas en modelos, analíticas y enfoques de data driven. Los resultados muestran que la combinación de múltiples técnicas puede resultar en mejores procesos de control para muchas aplicaciones.

A.10.4.1 REDES NEURONALES Y SISTEMAS EXPERTOS

La mayoría de los métodos basados en conocimientos pueden ser usados en conjunto con otro. Por ejemplo, redes neuronales y sistemas expertos han sido combinados y usados en aplicaciones industriales. Como se muestra en la sección 10.2, la fortaleza de sistemas expertos es su habilidad de imitar el razonamiento humano en resolver problemas de diagnóstico de fallas y su debilidad es el cuello

³⁸Introducción a las RedesNeuronales. Tomás Arredondo Vidal.Depto.Electronica UTFSM.4 mayo 2011

de botella en la adquisición de conocimiento. Como se muestra en la sección 10.3, la fortaleza de las redes neuronales es su habilidad para reconocer patrones basados en los ejemplos de entrenamiento y su debilidad es pobre habilidad para explicar los resultados.

La aplicación más directa a la utilización de redes neuronales para la mejora de los sistemas expertos es tener una red neuronal que sirva como una base de conocimiento para sistemas expertos. Esto permite que el sistema experto adquiera conocimiento de los datos. El entrenamiento puede ser en línea o desarrollado durante un periodo de iniciación. La base de conocimiento puede contener también modelos de sistemas los cuales producen en tiempo real resultados o aprendizaje seguro por medio de redes neuronales y así obtener nuevo conocimiento.

Los sistemas expertos pueden ser usados para mejorar las redes neuronales también. Una aplicación es usar los sistemas expertos como intérprete de redes para ejecutar diagnóstico de falla o evaluar resultados. Un sistema experto puede también ser usado para re-entrenar la red y adaptarla a situaciones desafiantes. Una herramienta combinada de redes neuronales y sistemas expertos fue desarrollada para transformar el diagnóstico de falla. Los resultados fueron que una herramienta que combine una red neuronal artificial y un sistema experto da un mejor rendimiento que usar cada técnica por separado.

A.10.4.2 LÓGICA FUZZY³⁹

³⁹ Se usará el término fuzzy en vez de su traducción (difusa o borrosa) en la mayoría del texto para poder relacionarlo fácilmente con terminología propia de la técnica.

La lógica fuzzy fue desarrollada por primera vez a mediados de 1960 para la representación de conocimiento incierto e impreciso. La lógica difusa es un medio aproximado pero eficaz de describir los complejos sistemas mal definidos mediante el uso de declaraciones clasificadas en lugar de los que son estrictamente verdaderas o falsas. La lógica difusa ha sido ampliamente aplicada en muchas áreas de la ingeniería en los últimos años. Existen muchos libros de lógica difusa⁴⁰.

Descripciones comúnmente usadas en los sistemas de ingeniería como “grande o pequeño” o “alto o bajo” son características de fuzzy. La descripción fuzzy es una conceptualización de valores numéricos que pueden ser cualitativos o significados de operadores. Una variable de proceso puede ser trasladada a un concepto fuzzy mediante una **función de pertenencia** $\mu_A(x)$, la cual mapea cada elemento x del conjunto X del intervalo $[0,1]$. Matemáticamente, esto puede ser definido como:

$$\mu_A(x): X \rightarrow [0,1] \quad (10.2)$$

donde A es un subconjunto fuzzy de X . Cada valor de esta función es llamada un **grado de pertenencia**. Un grado de pertenencia 0 indica no pertenencia, mientras un grado 1 indica pertenencia total en el conjunto A . Un conjunto definido en lógica

⁴⁰ N.K. Kasabov. Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering. MIT Press, Cambridge, Massachusetts, 1996.

L.H. Tsoukalas and R. E. Uhrig, Fuzzy and Neural Approaches in Engineering, John Wiley & Sons, New York, 1997.

Y. Zhang and A. Kandel. Compensatory Genetic Fuzzy Neural Networks and Their Applications. World Scientific, Singapore, 1997.

clásica (comúnmente referida como un **conjunto clásico**) es un caso especial de un conjunto fuzzy, en el cual solamente se permiten dos grados de pertenencia, 0 y 1. Un conjunto fuzzy A definido en X puede escribirse como una colección de parejas ordenadas

$$A = \bigcup_{x \in X} (x, \mu(x)) \tag{10.3}$$

donde cada pareja $(x, \mu(x))$ es llamado **instancia único (singleton)**. Si el conjunto X es discreto, una función de pertenencia puede ser definida por un conjunto finito:

$$A = \bigcup_k (x_k, \mu(x_k)) \tag{10.4}$$

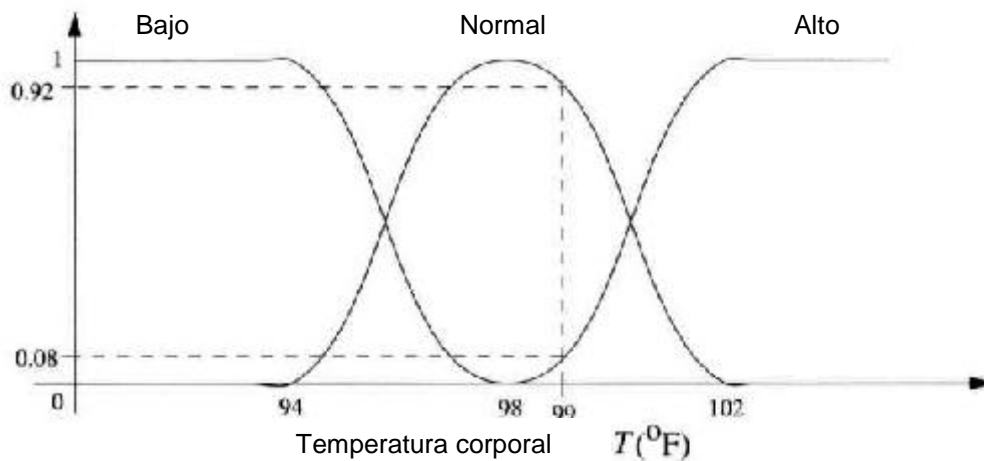


Figura 109. Función de pertenencia representando 3 conjuntos fuzzy para la variable lingüística Temperatura Corporal.

Fuente: Fault Detection and Diagnosis in Industrial Systems. L.H Chiang, E.L Russell and R.D. Braatz.

La lógica fuzzy permite una representación de variables y relaciones en términos lingüísticos. Una **Variable lingüística** es una variable la cual toma valores fuzzy y ha sido explicada lingüísticamente. Estas pueden ser basadas en variables cuantitativas del proceso, por ejemplo, la variable lingüística *temperatura corporal*, la cual puede tomar valores de “baja”, “normal” y “alta”. Cada valor fuzzy puede ser modelado como se muestra en la Figura 109. Por ejemplo, una temperatura corporal de 99°F toma un valor fuzzy de “normal” y un grado de pertenencia de 0.92 por medio de $\mu_{\text{Normal}}(T)$. Esta también puede tomar un valor fuzzy de “alto” y un grado de pertenencia de 0.08 por medio de $\mu_{\text{Alta}}(T)$. Las variables lingüísticas pueden también ser cualitativas, por ejemplo, la variable *certeza* la cual puede tener valores fuzzy como “alta certeza” o “no muy cierto”. El proceso de representación de las variables lingüísticas dentro de un conjunto de valores fuzzy es llamado **cuantificación fuzzy**.

Las funciones de pertenencia mostradas en la figura 109. son definidas basadas en datos estadísticos. Las funciones de pertenencia para “bajo”, “normal” y “alto” son representadas por un función Z (la cual es una función sigmoideal invertida), una función en forma de campana , y una función sigmoideal, respectivamente. Otros tipos de funciones de pertenencia incluyendo trapezoidal, triangular, y funciones de valores singulares pueden ser usadas.

Los sistemas de lógica fuzzy enfrentan la imprecisión de las variables de entrada y salida directamente definiéndolas con números fuzzy y conjuntos fuzzy que pueden ser expresados en términos lingüísticos.

El complejo comportamiento del proceso puede ser descrito en términos generales sin definir exactamente el complejo fenómeno involucrado. Sin embargo, esto es difícil y consume tiempo determinar el correcto conjunto de reglas y funciones de pertenencia para un sistema considerablemente complejo. La puesta a punto de una solución difusa requiere una gran cantidad de tiempo. Para resolver algunos de los hechos, las redes neuronales pueden ser usadas para aprender la mejor función de pertenencia a través del entrenamiento.

A.10.4.3 SISTEMAS EXPERTOS FUZZY

Ha sido observado que el número de reglas SI-ENTONCES requeridas para definir un sistema experto tiende a crecer exponencialmente cuando la complejidad del sistema se incrementa. Cuando el número de reglas SI-ENTONCES sea más grande de 200, es virtualmente imposible de escribir una regla significado que no entre en conflicto con las reglas existentes. Esto ha motivado recientes investigaciones en incorporar lógica fuzzy dentro de sistemas expertos en un intento por reducir el número de reglas requeridas.

Un **sistema experto fuzzy** (también conocido como un **sistema fuzzy**) es definido de la misma forma que un sistema experto ordinario como se describe en la sección 10.2, excepto que usa la lógica fuzzy. Los sistemas expertos fuzzy usan datos fuzzy, reglas fuzzy, y un mecanismo de inferencia fuzzy la cual incluye fuzzificación y defuzzificación.

Los datos de entrada y salida pueden ser fuzzy (como se describió en la sección 10.4.2) o exactos (crisp o clásicos).

Cuando los datos de entrada y los valores de salida son exactos, entonces la fuzzificación, regla fuzzy y el método de inferencia defuzzificación es aplicado. **Fuzzificación** es el proceso para encontrar la función de pertenencia $\mu_A(x)$ y producir la función de pertenencia de salida.

La **defuzzificación** es el proceso de calcular valores numéricos singulares de salida para una variable fuzzy de salida sobre la base de la función de pertenencia inferida para esta variable.

Las reglas fuzzy y las funciones de pertenencia forman la base del sistema de conocimiento. Las **reglas fuzzy** se enfrentan con los valores fuzzy. La regla más popular es la SI-ENTONCES. Estos son argumentos condicionales que describen la dependencia de uno o más variables lingüísticas con otras.

Una debilidad del enfoque fuzzy es la cantidad relativamente grande de reglas fuzzy que se generan. Para reducir este número se requiere describir un sistema complejo, un algoritmo genético de optimización puede ser usado. Alternativamente, un proceso basado en estadística puede analizar la situación y dar la contribución de cada regla a la solución.

La inferencia fuzzy toma las entradas, aplica reglas fuzzy, y produce salidas. La inferencia fuzzy es un método de inferencia que usa relaciones de implicación

fuzzy (por ejemplo, la regla SI-ENTONCES), operadores de composición fuzzy (por ejemplo, MIN, MAX), y un operador (por ejemplo, AND, OR) para enlazar las reglas fuzzy. El proceso de inferencia resulta en inferir nuevos hechos basados en reglas fuzzy y en la información de entrada suministrada.

En general, el gran número de reglas fuzzy, la alta probabilidad de generar conflicto de reglas (es decir, reglas que tienen la misma parte SI pero diferente parte ENTONCES). Para resolver este problema, la regla con el más alto grado de confianza es retenida y la regla con el más bajo grado es descartada. El máximo número de reglas fuzzy generadas en el conjunto de entrenamiento es igual al número de observaciones en este conjunto.

Las tasas de clasificación errónea general para los datos de Fisher son más altas que con los métodos Data-Driven (PCA, PLS y FDA). La eficiencia de las reglas fuzzy depende de la selección de las funciones de pertenencia y del número de valores fuzzy. La puesta a punto de los parámetros podría resultar en mejores resultados de clasificación.

A.10.4.4 REDES NEURONALES FUZZY

La lógica fuzzy puede ser usada con redes neuronales. Una **neurona fuzzy** tiene la misma estructura básica que una neurona artificial, excepto que algunas o todos los componentes y parámetros pueden ser descritos a través de lógica fuzzy. Una **red neuronal fuzzy** es construída con neuronas fuzzy o neuronas estándar pero enfrentada con datos fuzzy. Una red neuronal fuzzy es un modelo conexionista para la implementación y la inferencia de reglas fuzzy. Hay muchas maneras

distintas para fuzzificar una red neuronal, cuyos resultados son una variedad de neuronas fuzzy y redes fuzzy.

Los procesos de entrenamiento en una red neuronal difusa consisten en modificar los parámetros o pesos de la red para obtener un valor óptimo. Esto requiere una colección de pares de entrada- salida y de un índice que exprese el mapa de las entradas y los objetivos correspondientes.

Representación de datos en Sistemas Híbridos (Neuro-fuzzy)⁴¹

Durante los últimos años se ha tenido un incremento importante en investigaciones para sintetizar la lógica Fuzzy con las redes neuronales artificiales (RNA's).

Las RNA's ofrecen ventajas en áreas de aprendizaje, clasificación y optimización, mientras que la lógica Fuzzy en áreas de razonamiento de alto nivel.

Representación difusa de variables que son entradas y salidas de RNA's.

⁴¹ Los términos neuro-fuzzy, neuro-difuso o difuso neuronal se refieren en la literatura a combinaciones híbridas de herramientas lógico difusas y neuronales.

Algunas veces la distribución de las posibles salidas de una RNA requiere un gran número de neuronas, lo cual implica un aumento en la complejidad y el entrenamiento.

Si se considera una representación difusa o lingüística de las variables, se pueden reducir de “n” salidas a solo algunas representadas por conjuntos que agrupan posibles valores.

Algunas aplicaciones de Ingeniería.

- Las redes neuronales fuzzy ambicionan integrar el aprendizaje neuronal con la capacidad de representación de conocimiento de los sistemas difusos.
- Las RNA's son típicamente heterogéneas para un mejor reflejo lógico de un problema dado.
- Las capas y nodos de una red neuro-fuzzy pueden ser interpretadas como una realización de reglas fuzzy SI/ENTONCES.
- Pedrycz y Rocha (1993) modelos neurodifusos usando operadores lógicos (AND, OR, NOT), así como normas T y S.
- Zhang y Chen (1994) utilizan un híbrido que consistía de una red de parámetros estimables, una capa convertidora, y una basada en programación hacia atrás para la clasificación de problemas con características complejas. Esto ha sido aplicado en la clasificación de imágenes de satélites y determinación litográfica.