

**MÉTODOS GRÁFICOS DE ANÁLISIS DE DISPERSIÓN DE ÁREAS FACTIBLES
PARA LA OPTIMIZACIÓN DE SISTEMAS DE TORRES DE DESTILACIÓN Y SU
INTERACCIÓN CON REDES DE INTERCAMBIO DE CALOR**

SILVIA JULIANA CARILLO ABRIL

RUTH SMITH FLÓREZ DÍAZ

UNIVERSIDAD INDUSTRIAL DE SANTANDER

FACULTAD DE FISICOQUÍMICA

ESCUELA DE INGENIERIA QUÍMICA

BUCARAMANGA

2011

**MÉTODOS GRÁFICOS DE ANÁLISIS DE DISPERSIÓN DE ÁREAS FACTIBLES
PARA LA OPTIMIZACIÓN DE SISTEMAS DE TORRES DE DESTILACIÓN Y SU
INTERACCIÓN CON REDES DE INTERCAMBIO DE CALOR**

SILVIA JULIANA CARILLO ABRIL

RUTH SMITH FLÓREZ DÍAZ

Tesis de Grado realizada para optar al título de Ingeniero Químico

Director

CRISÓSTOMO BARAJAS FERREIRA

Ingeniero Químico. Msc.

Codirector

CESAR AUGUSTO MAHECHA BOHORQUEZ

Ingeniero Químico. Msc.

UNIVERSIDAD INDUSTRIAL DE SANTANDER

FACULTAD DE FISICOQUÍMICA

ESCUELA DE INGENIERIA QUÍMICA

BUCARAMANGA

2011

AGRADECIMIENTOS

A Dios que en su plan perfecto me dio la oportunidad de venir a esta tierra acompañada de la mejor familia, a mi madre Ana Inés Díaz quien con su esfuerzo y trabajo luchó por que sus hijos fueran profesionales y hoy por fin termina su ardua tarea, a mis hermanos Rahysol, John y Jessica que me ha apoyado en cada paso, a mi cuñado Michael Yeats por su buen corazón, y a mis sobrinitos Natalia, Andrew y Ana por ser fuente de alegría en mi vida.

Este proyecto no existiría sin el apoyo del Instituto Colombiano del Petróleo, gracias por esta oportunidad de desarrollo profesional.

Agradezco al Ingeniero Cesar Augusto Mahecha por la oportunidad que nos dio de realizar esta tesis de investigación, al profesor Crisóstomo Barajas por su respaldo, al Ingeniero Juan Carlos Escobar y al futuro Ingeniero Alexander Díaz Gonzales por su orientación y colaboración, a los Docentes de la escuela de Ingeniería química, que cada día se esfuerzan por compartir su conocimiento y su experiencia, a mi compañera de proyecto Silvia Juliana Carrillo por ser un aliado en esta tarea y por ser una poyo en momentos de desanimo, a mis amigos María Vargas Quintero, Sol Alejandra Calderón, Luzeni Lillo y Oscar Iván Hernández quienes siempre me han apoyado con sus oraciones y a todos los que de corazón compartan esta alegría conmigo y mi familia.

Gracias.

Ruth Smith Flórez Díaz

AGRADECIMIENTOS

A mi Dios y la Virgen Santísima por acompañarme siempre, concediéndome salud y sabiduría en la realización de este proyecto con el que finalizo una etapa muy importante de mi vida.

A San Judas Tadeo que me lleno de paciencia en todas las adversidades que se me presentaron y no me dejó desfallecer.

A mi mamá CLAUDIA ABRIL, a mi papá CESAR CARRILLO y mi hermano DANIEL EDUARDO por iluminar mi vida, ser siempre mi fortaleza y mi apoyo incondicional.

A mi compañera de travesía RUTH SMITH FLÓREZ por sus valiosos aportes su incondicional forma de ser y sobre todo por vivir paso a paso este triunfo Dios la ilumine y proteja siempre.

A los Ingenieros Adriana Joya, Julieth Hernández, Laura Rodríguez, Tania Sandoval, Jesús Pirateque, María Vargas, Kenneth, Diego, Sergio y Fabio por acompañarme y hacer de esta etapa de mi vida inolvidable los quiero y respeto por sus maravillosas formas de ser.

A el Ingeniero Luís Andrés Aguilar por su apoyo incondicional y compañía.

Al Ingeniero Cesar Augusto Mahecha y el profesor Crisóstomo Barajas Ferreira por respaldar esta iniciativa y llevar a feliz término.

Silvia Juliana Carrillo Abril

CONTENIDO

	Pág.
INTRODUCCION.....	19
1. DEFINICIÓN DEL PROBLEMA.....	21
1.1 TORRE DE DESTILACIÓN ATMOSFÉRICA.....	21
1.2 CARGAS DE CRUDO A OPTIMIZAR.....	23
2. OBJETIVOS.....	25
2.1 OBJETIVO GENERAL.....	25
2.2 OBJETIVOS ESPECIFICOS.....	25
3. MARCO REFERENCIAL.....	26
3.1 ASPECTOS GENERALES.....	26
3.2 DISEÑO DE EXPERIMENTOS.....	27
3,2,1, PROPOSITO DE UN DISEÑO EXPERIMENTAL.....	28
3.2.2. PRINCIPIOS BASICOS DEL DISEÑO EXPERIMENTAL.....	28

3.3. DISEÑO SECUENCIAL.....	29
3.3.1. DISEÑO SECUENCIAL DE EXPERIMENTOS.....	29
3.3.2. DISEÑO SECUENCIAL PARA REDUCIR INCERTIDUMBRE ENTRE PARÁMETROS ESTIMADOS	30
3.3.2.1. PROCESOS CON RESPUESTA UNICA.....	31
3.3.2.2. PROCESOS MULTIRESPUESTA.....	38
3.4. PROGRAMACION NO LINEAL.....	42
3.4.1. METODO COMPLEX.....	42
4. RESULTADOS Y ANALISIS.....	47
4.1. TIPO Y FORMA DEL METAMODELO.....	47
4.2. DETERMINACIÓN DEL DISEÑO INICIAL.....	48
4.2.1. VARIABLES DE ENTRADA.....	48
4.2.2. RESTRICCIONES.....	49
4.2.3. VARIABLE DE SALIDA.....	49
4.2.4. RESULTADO DE LA SIMULACION.....	49
4.2.5. VARIABLES INDEPENDIENTES ÓPTIMAS.....	50

5. CONCLUSIONES.....	52
6. RECOMENDACIONES.....	53
BIBLIOGRAFIA.....	54
ANEXO 1.....	57
ANEXO 2.....	61
ANEXO 3.....	63
ANEXO 4.....	64
ANEXO 5.....	65
ANEXO 6.....	66
ANEXO 7.....	80

LISTA DE TABLAS.

	Pág.
Tabla 1. Caracterización de los crudos de la dieta de las UDCs.....	24
Tabla 2. Variables de entrada del metamodelo (curva ASTM D-86 para el diesel de la (TDA1)).....	61
Tabla 3. Variables independientes del metamodelo (curva ASTM D-86 para el diesel de la (TDA1)) reducido a sus variables de mayor incidencia.....	63
Tabla 4. Datos iniciales de las variables independientes (X).....	64
Tabla 5. Datos iniciales de los parámetros del metamodelo (curva ASTM D-86 para el diesel de la (TDA1)) reducido.	65
Tabla 6. Variables independientes óptimas.....	51

LISTA DE FIGURAS.

	Pág.
Figura 1. Diagrama Simplificado del Sistema de UDCs.....	22
Figura 2. Diagrama de la Torre Atmosférica. 1 (TDA 1).....	23
Figura 3. Algoritmo de diseño secuencial.....	41
Figura 4. Procedimiento iterativo método Complex.....	46

LISTA DE ANEXOS.

	Pág.
Anexo 1. DSE (<i>Algoritmo de Diseño Secuencial de Experimentos</i>).....	57
Anexo 2. Tabla Variables de entrada del metamodelo (curva ASTM D-86 para el diesel de la (TDA1)).....	61
Anexo 3. Tabla Variables independientes del metamodelo (curva ASTM D-86 para el diesel de la (TDA1)) reducido a sus variables de mayor incidencia.....	63
Anexo 4. Tabla Datos iniciales de las variables independientes (X).....	64
Anexo 5. Tabla Datos iniciales de los parámetros del metamodelo (curva ASTM D-86 para el diesel de la (TDA1)) reducido.....	65
Anexo 6. Código DSE y subrutinas en MATLAB.....	66
Anexo 7. Ejemplo 8.4-1 Optimal sequential designs for nonlinear model.....	80

RESUMEN

Título: METÓDOS GRÁFICOS DE ANÁLISIS DE DISPERSIÓN DE ÁREAS FACTIBLES PARA LA OPTIMIZACIÓN DE SISTEMAS DE TORRES DE DESTILACIÓN Y SU INTERACCIÓN CON REDES DE INTERCAMBIO DE CALOR*

Autores: Silvia Juliana Carrillo Abril; Ruth Smith Flórez Díaz. **

Palabras Claves: Optimización, Parámetros Diseño secuencial, Metamodelo, Algoritmo, Torre de destilación atmosférica.

Descripción:

Se desarrollo un algoritmo para la optimización de los parámetros de un modelo particular multivariado y no lineal mediante la metodología de diseño secuencial de experimentos, esto con el fin de mejorar los metamodelos creados para la torre de destilación atmosférica1 objeto de estudio y que pertenece a la UDC1 de la refinería de Barrancabermeja de ECOPETROL S.A.

El algoritmo aplicado encuentra los estimados óptimos de los coeficientes, para lo cual requiere valores iniciales de las variables independientes del metamodelo, los estimados iniciales de los coeficientes a optimizar, y un modelo tipo polinomial de segundo grado e interacción binaria. La técnica DSE involucra un proceso de maximización que se realiza mediante una optimización no lineal empleando el método Complex, como resultado obtenemos valores óptimos de las variables independientes que dirigen la búsqueda de los coeficientes y proveen mejores estimaciones.

La optimización de los parámetros de un modelo reviste de cierta complejidad sobre todo si el número de parámetros es elevado, más aun tratándose de la técnica DSE ya que involucra matrices de derivadas parciales, que aportan un alto gasto computacional, es por eso que el desempeño del algoritmo está ligado directamente con el tamaño del modelo. Para el metamodelo de la TDA1 escogido se manejan diez variables de entrada entre flujos, temperaturas y presiones y se cuenta con once parámetros. El tiempo de ejecución usado por el algoritmo DSE para la optimización tanto de las variables independientes como de los coeficientes fue de 4 días, cabe resaltar que esta metodología resulta eficiente en la optimización de metamodelos de menos de 5 variables y disminuye al aumentar la complejidad del modelo siendo insostenible en metamodelos que exceden los 50 parámetros.

*Trabajo de Investigación

**Facultad de Físicoquímicas. Escuela de ingeniería química. Crisóstomo Barajas Ferreira. Cesar Augusto Mahecha Bohórquez.

SUMMARY

Title: GRAPHIC METHODS OF ANALYSIS OF FEASIBLE AREA DISPERSION FOR THE OPTIMIZATION OF TOWER DISTILLATION SYSTEM AND THEIR REACTION WITH HEAT EXCHANGE NETWORKS.*

Authors: Silvia Juliana Carrillo Abril; Ruth Smith Flórez Díaz. **

Key Words: Optimization, Sequential Design, Parameters, Metamodel, Algorithm, Atmospheric Distillation Tower.

Description:

An algorithm was developed for the optimization of parameters of a particular non-linear, multivariate model through the methodology of sequential experimental design, all with the aim of improving the metamodels created by the Atmospheric Distillation Tower 1 object of the study that belongs to the UDC1 of the ECOPETROL S.A. Barrancabermeja refinery.

The applied algorithm finds the optimum estimates of the coefficients, which require initial values for the independent variables of the metamodel, the initial estimates of the coefficients to optimize, and a polynomial type model to the second degree and binary interaction. The DSE technique involves a maximization process that is accomplished through a non linear optimization using the Complex model, resulting in optimum values of the independent variables that direct the search for the coefficients and provide best estimates.

The optimization of a model's parameters is somewhat complex especially if there is a large number of parameters, and even more when the DSE technique is used due to the partial derivative matrices which are involved, contributing to a significant computational investment, and for this reason the performance of the algorithm is directly tied to the size of the model. In order to manage the TDA1 metamodel ten flow, temperature, and pressure injection variables are chosen with eleven parameters. The execution time for the DSE algorithm as well as for the independent variables and coefficients was 4 days, highlighting the fact that this methodology can be efficient for optimizing metamodels with 5 or fewer variables and its efficiency is reduced as the models complexity increases and becomes wholly unsustainable for metamodels with more than 50 parameters.

*Degree Job

**Faculty of Physicochemical Engineering. School of Chemical Engineering. Crisóstomo Barajas Ferreira. Cesar Augusto Mahecha Bohórquez.

INTRODUCCION

Los modelos se utilizan en todos los campos: biología, ingeniería, química, física y economía. Los modelos matemáticos son la forma más conveniente y compacta de reducir la información experimental a una expresión que nos permita representar de manera aproximada la realidad de los procesos y predecir respuestas desconocidas.

Para llegar a tener un modelo óptimo es necesario determinar de manera adecuada y cuidadosa los parámetros que lo integran, por esto existen los métodos de diseño de experimentos que nos permiten establecer el valor de los parámetros del modelo y el intervalo de confianza mediante la búsqueda de estimados que den lugar a las mejores respuestas, a este proceso se le conoce comúnmente como optimización.

Nuestro objeto de estudio es la torre de destilación atmosférica (TDA1) que forma parte de una de las unidades de destilación de crudo (UDC1) presentes en la refinería de Barrancabermeja ECOPETROL S.A., La investigación toma como punto de partida previos estudios^[3] en los que se desarrollaron metamodelos mediante las técnicas de Muestreo por Hipercubo Latino (LHS) y Mínimo Margen de Error del Hipercubo Latino (MBLHD), siendo este segundo nuestra base, en la implementación de la metodología de diseño secuencial de experimentos.

Las torres de destilación atmosférica poseen un número elevado de variables independientes que influyen en su funcionamiento; tales como flujos, temperaturas y presiones. En este caso intervienen 22 variables por lo que es indispensable valerse de una herramienta computacional que agilice el proceso de optimización.

El propósito de los metamodelos es determinar la composición de la mezcla de crudo a cargar y las condiciones de operación de las torres de destilación atmosférica. Su optimización garantiza la obtención de mejores y mayores cantidades de las corrientes valiosas.

Este trabajo tiene como fin el uso de esta nueva técnica para encontrar la región óptima de experimentación que permita dirigir la estimación de los parámetros y hacerla más efectiva.

1. DEFINICIÓN DEL PROBLEMA

El problema analizado en este trabajo, es comúnmente motivo de estudio y consiste en la optimización dirigida al cálculo de variables óptimas de experimentación para funciones limitadas en un dominio específico. [2] En este caso la industria en la que nos movemos es la refinación de petróleo y el proceso en el que trabajamos consiste en las unidades de destilación de crudo, en específico la TDA1.

La razón por la que se usan metamodelos es porque son un punto medio entre modelos rigurosos y lineales que permiten obtener una representación adecuada de la realidad con relativa simplicidad. Los metamodelos diseñados y construidos en el ICP tienen como base los modelos de las plantas de destilación configurados y sintonizados en el simulador comercial PRO/II. Dichos modelos representan de forma rigurosa y adecuada el fenómeno de la destilación de cada una de las unidades por tanto sus predicciones se convirtieron en metamodelos. [3] El interés se centra en encontrar los mejores puntos de operación para el equipo instalado (TDA1) teniendo en cuenta las restricciones de planta.

1.1. Torre de Destilación Atmosférica 1[3]

Forma parte de la Unidad de Destilación de Crudo 1 que consta de 2 torres de destilación atmosférica (TDA1 y TDA2), 2 hornos atmosféricos (HA1 y HA2), 1 torre de vacío (TDV1) y una red de intercambiadores que precalientan crudo (RED1) y es compartida por las dos torres de destilación atmosféricas.

Figura 1. Diagrama Simplificado del Sistema de UDCs.

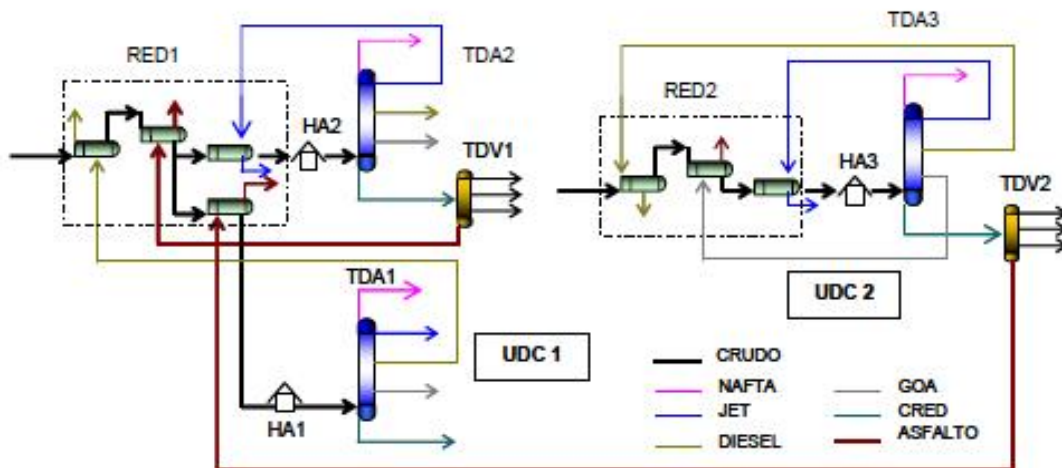


Figura 1. Diagrama Simplificado del Sistema de UDCs

Fuente. (LOPEZ. Diana Carolina, Optimización Simultánea de Composición de Cargas y Variables de Operación de Unidades de Destilación de Crudo, Bucaramanga, Universidad Industrial de Santander. 2008) [3]

La TDA1 está constituida por 30 platos con 3 torres despojadoras de los productos Jet, ACPM y GOA, tiene 3 pumparrounds: cima (UPA), medio (MPA) e inferior (LPA) quienes seden su energía al crudo a través de la integración energética. Su capacidad de procesamiento es de 55000 barriles/día (BPD).

Figura 2. Diagrama de la Torre Atmosférica 1 (TDA 1)

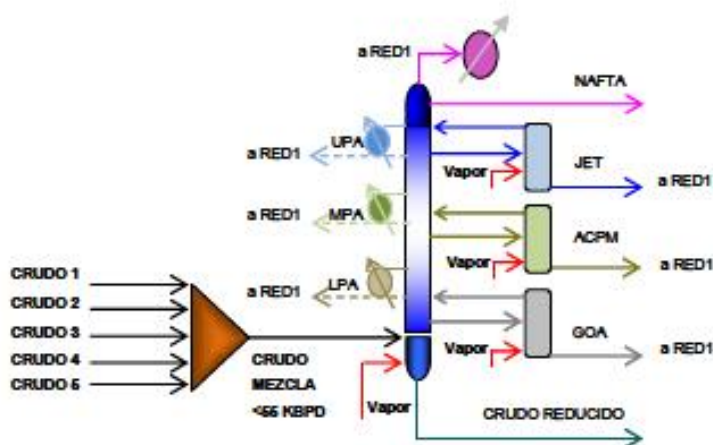


Figura 2. Diagrama de la Torre Atmosférica 1 (TDA1)

Fuente (LOPEZ. Diana Carolina, Optimización Simultánea de Composición de Cargas y Variables de Operación de Unidades de Destilación de Crudo, Bucaramanga, Universidad Industrial de Santander. 2008) [3]

1.2. Cargas de Crudo a Optimizar [3]

Actualmente a la refinería llega un conjunto de crudos que se mezclan para formar la carga de la torre de destilación atmosférica estudiada en este proyecto.

Hoy en día la carga a las 2 UDCs es la mezcla de varios crudos pesados, en este proyecto se usaron los datos proporcionados por el ICP de las cargas de crudo

proyectadas para el año 2015 lo que contempla una alta proporción de crudos extra-pesados.

La composición de la carga a evaluar en el modelo de optimización está conformada por 5 crudos (4 extra-pesados y 1 diluyente-CRUDO1) y sus propiedades se muestran en la tabla 1.

Tabla 1. Caracterización de los crudos de la dieta de las UDCs.

PROPIEDAD	CRUDO 1	CRUDO 2	CRUDO 3	CRUDO 4	CRUDO 5
ASTM-D86 10% (°F)	297,4	611,8	563,6	575,9	476,6
ASTM-D86 30% (°F)	552,6	849,4	736,2	771,3	750,0
ASTM-D86 50% (°F)	760,8	1105,0	880,0	910,0	931,6
ASTM-D86 70% (°F)	955,4	1307,9	1097,8	1176,4	1251,5
ASTM-D86 90% (°F)	1214,7	1483,0	1338,4	1421,1	1463,6
GRAVEDAD API	34,6	8,09	13,16	12,40	13,59
AZUFRE (%Peso)	0,59	3,29	1,27	1,51	2,27
CCR (%Peso)	5,95	18,71	12,73	8,97	16,09
NIQUEL (ppm Peso)	42,40	139,07	26,74	59,95	86,35
VANADIO (ppm Peso)	55,13	533,66	131,55	71,64	352,40
NN (mgKOH/g)	0,901	0,892	0,361	4,669	0,330

Fuente (LOPEZ. Diana Carolina, Optimización Simultánea de Composición de Cargas y Variables de Operación de Unidades de Destilación de Crudo, Bucaramanga, Universidad Industrial de Santander. 2008). [3]

2. OBJETIVOS

2.1. Objetivo General

Crear una herramienta computacional que permita mediante la metodología DSE seleccionar un arreglo óptimo para las variables independientes.

Objetivos Específicos

- Encontrar las condiciones del siguiente experimento, dados unos parámetros y variables iniciales.
- Dado uno de los metamodelos que describen la torre de destilación atmosférica¹ encontrar la región de experimentación óptima que permita hacer más eficiente la búsqueda de los parámetros del mismo.
- Disminuir el proceso de optimización mediante la reducción de la región de experimentación y el direccionamiento de los experimentos.
- Desarrollar una herramienta computacional basada en la metodología DSE que permita encontrar la subsiguiente región de experimentación cada vez que un nuevo vector de parámetros es estimado.

3. MARCO REFERENCIAL.

3.1 ASPECTOS GENERALES [7].

El Diseño de Experimentos tuvo su inicio teórico a partir de 1935 por Sir Ronald A. Fisher, quién sentó la base de la teoría del Diseño Experimental y que a la fecha se encuentra bastante desarrollada y ampliada. Actualmente las aplicaciones son múltiples, especialmente en la investigación de las ciencias naturales, ingeniería, laboratorios y casi todas las ramas de las ciencias sociales.

La experimentación proporciona los datos experimentales, en contraste con los datos de la observación; los datos de la observación se representan como su nombre indica por observaciones de las unidades elementales o de una muestra, y no deben ser cambiados ni modificados por ningún investigador en el curso de la observación. Pasos al planear un experimento:

El método científico sugiere que en el planeamiento de la experimentación se debe tener presente las siguientes etapas:

- a) Definir el problema: En esta etapa se debe determinar los antecedentes, importancia, objetivos, hipótesis a probar y revisión de la bibliografía.
- b) Planeamiento y diseño del experimento: En esta etapa se debe tener en cuenta: Lugar de ejecución del experimento, tamaño de la parcela o unidad

experimental, número de repeticiones por tratamiento, equipos e instrumentos a utilizar y métodos de evaluación de los resultados.

- c) Ejecución del experimento.
- d) Recolección de datos del experimento.
- e) Ordenamiento de la información experimental.
- f) Discusión de los resultados obtenidos.
- g) Análisis económico de los tratamientos que se probaron y utilidad práctica.
- h) Conclusión final y recomendación.

3.2 DISEÑO DE EXPERIMENTOS

Este término se utiliza para planear un experimento de manera que se pueda obtener la información pertinente a un determinado problema que se investiga y así tomar decisiones correctas. El diseño adecuado del experimento es una etapa fundamental de la experimentación, que permite el suministro correcto de datos a posteriori, los que a su vez conducirán a un análisis objetivo y con deducciones válidas del problema.

3.2.1 PROPÓSITO DE UN DISEÑO EXPERIMENTAL

El propósito de un diseño experimental es proporcionar métodos que permitan obtener la mayor cantidad de información válida acerca de una investigación, teniendo en cuenta el factor costo y el uso adecuado del material disponible mediante métodos que permitan disminuir el error experimental.

3.2.2 PRINCIPIOS BASICOS DEL DISEÑO EXPERIMENTAL

Los principios básicos del diseño experimental son: repetición, aleatorización, y control local.

- **Repetición:** Viene a ser la reproducción o réplica del experimento básico (asignación de un tratamiento a una unidad experimental). Las principales razones por las cuales es deseable la repetición son: Primero, por que proporciona una estimación del error experimental, siendo tal estimación confiable a medida que aumenta el número de repeticiones, y segundo permite estimaciones más precisas del tratamiento en estudio.
- **Aleatorización:** Consiste en la asignación al azar de los tratamientos en estudio a las unidades experimentales con el propósito de asegurar que un determinado tratamiento no presente sesgo. Por otro lado la aleatorización hace válidos los procesos de inferencia y las pruebas estadísticas.
- **Control Local (Control del error Experimental):** Consiste en tomar medidas dentro del diseño experimental para hacerlo más eficiente, de tal manera que

pueda permitir la reducción del error experimental y así hacerla más sensible a cualquier prueba de significación.

3.3 DISEÑO SECUENCIAL

En el diseño secuencial, los experimentos se realizan sucesivamente, tomando como base los resultados de la experimentación previa hasta alcanzar unas condiciones óptimas.

3.3.1. DISEÑO SECUENCIAL DE EXPERIMENTOS

Este tipo de diseño de experimentos tiene ciertas ventajas respecto de otros, ya que incluye el modelo matemático del proceso, lo que orienta a la predicción de mejores condiciones de experimentación. Así se tiene la posibilidad de utilizar los recursos con mayor eficacia.

El diseño secuencial es una técnica compleja que requiere el manejo de matrices n -dimensionales de ecuaciones diferenciales con el propósito de seleccionar un arreglo óptimo de las variables para el próximo experimento con la incertidumbre estadística más baja.

El procedimiento para el desarrollo del diseño secuencial está dividido en dos partes:

- a) La discriminación de modelos.

- b) La estimación de parámetros. (Para un modelo $y = a + bx$ este proceso permite estimar los valores de a y b que brindan el mejor ajuste del modelo).

El proceso de discriminación de parámetros se presenta a continuación.

3.3.2. DISEÑO SECUENCIAL PARA REDUCIR INCERTIDUMBRE ENTRE PARAMETROS ESTIMADOS. [5]

Si se parte de un modelo sencillo y adecuado, lo deseable es obtener los mejores estimados para los coeficientes de este. Para lo cual es necesario que los experimentos se planeen apropiadamente, de lo contrario los coeficientes encontrados pueden ser no solo imprecisos sino también altamente correlacionados.

Box .W. G, Hunter, Draper y sus colaboradores participaron en el desarrollo del procedimiento del diseño secuencial para disminuir la incertidumbre en la estimación de parámetros, especialmente en conexión con modelos no lineales.

El procedimiento parte de un modelo que se representa por:

$$Y_i = n_i + \epsilon_i \quad (1)$$

n es el modelo que contiene m parámetros y q variables independientes x , las observaciones se representan por Y_i donde los ϵ_i son errores independientes, normalmente distribuidos, con media cero y una varianza de σ_y^2

Del mismo modo $[X_{ij}] \quad (2)$

Este es el típico elemento de la matriz X que se definirá más adelante.

Los modelos descritos pueden por lo tanto ser de respuesta única o procesos multirespuesta, sus desarrollos son análogos, como se muestra a continuación.

3.3.2.1. PROCESOS CON RESPUESTA UNICA

Se asume que algunas observaciones de las variables independientes han sido tomadas y se desea conocer que valores de estas x deben ser escogidos para la siguiente observación, y que estimaran los parámetros de manera más efectiva en algún sentido.

El procedimiento se desarrolla mediante el uso del teorema de Bayes [5] y asumiendo distribución normal multivariada para β centrado en un vector de estimados b .

El teorema de Bayes escrito en términos de esta notación es:

$$P_{n+1}(\beta|y_{n+1}) = \quad (3)$$

Donde:

= densidad de probabilidad después de n+1 observaciones.

= la anterior densidad de probabilidad de la función β antes de que la última observación o grupo de observaciones ha sido obtenida.

= la función de densidad de probabilidad la cual es la probabilidad de β dado y_{n+1} .

Y donde los subíndices en \mathbf{y} , y \mathbf{p} designan el número de corridas experimentales. La integral en el denominador de la ecuación representa el factor de normalización. La dependencia de la función de densidad posterior tanto en las variables independientes x como en la Y es implícita aunque no específicamente fija. Desde el punto de vista bayesiano la función de probabilidad posterior contiene toda la información disponible acerca de las β 's. Un criterio natural usado en la fijación de los valores de un diseño experimental podría ser seleccionar los valores que producen la distribución posterior más deseable en algún sentido, y en particular los que logran el mayor cambio deseable en la distribución al finalizar el n -ésimo experimento.

Una vez la primera observación n ha sido hecha, se puede escribir una función de probabilidad como sigue:

$$L(\beta|y_n, X, \sigma_Y^2) = \frac{1}{(2\pi)^{n/2} \sigma_Y^n} \quad (4)$$

Para obtener $P_n(\beta)$ y usarlo en el teorema de Bayes es necesario tener la densidad de probabilidad inicial para β antes de la n -ésima observación tomada. Por lo tanto si tenemos $P_0(\beta)$ y $L(\beta|y_n)$ la aplicación de la siguiente ecuación nos puede proporcionar la deseada densidad posterior $P_n(\beta)$, la cual a su vez puede ser usada como la anterior densidad de probabilidad para la subsiguiente observación en el teorema de Bayes.

$$P_n(\beta) = \quad (5)$$

Ahora para calcular $P_0(\beta)$ en lugar de asumir una densidad de probabilidad uniforme, se asumirá que la información anterior está disponible para β de alguna fuente, de modo que la densidad inicial para las β 's puede ser dada por:

$$P_0(\beta) = \frac{1}{(2\pi)^{m/2} [\Omega]^{1/2}} \exp \left[-1/2 (\beta - b^{(0)})^T [\Omega]^{-1} (\beta - b^{(0)}) \right] \quad (6)$$

Donde $b^{(0)}$ es el vector de estimados iniciales para las β 's y Ω es una matriz de covarianza para las β 's en la cual cada elemento está especificado:

(7)

Introduciendo las ecuaciones (6) y (4) dentro del teorema de Bayes se obtiene la deseada función de densidad de probabilidad después de n observaciones.

$$p_n(\beta|y_n) = k_n \exp \left[\frac{-\sum_{i=1}^n Y_i - n_i(\beta, x_i)^2}{2\sigma_Y^2} \right] * \exp \left[-1/2 (\beta - b^{(0)})^T [\Omega]^{-1} (\beta - b^{(0)}) \right] \quad (8)$$

Donde K_n representa el factor de normalización apropiado el cual no necesita ser específicamente determinado aún.

Ahora si deseamos realizar n^* observaciones adicionales de Y, la ecuación (8) se convierte en la distribución anterior para el teorema de Bayes; ahora la nueva función de probabilidad será análoga, excepto porque envuelve n^* productos y la suma va desde $n+1$ hasta $n+n^*$ observaciones, es decir sobre las nuevas observaciones.

$$L(\beta|y_{n+n^*}) = \frac{1}{(2\pi)^{n^*/2} \sigma_Y^{n^*}} \exp \quad (9)$$

La introducción de las dos ecuaciones anteriores dentro de la ecuación (3) da como resultado la deseada densidad de probabilidad posterior para las β 's:

$$p_{n+n^*}(\beta|y_{n+n^*}) = k_{n+n^*} \exp\left[\frac{\sum_{i=1}^{n+n^*} y_i - n_i(\beta, x_i)^2}{2\sigma_Y^2}\right] * \exp\left[-1/2(\beta - b^{(0)})^T [\Omega]^{-1}(\beta - b^{(0)})\right] \quad (10)$$

Donde k_{n+n^*} es el nuevo factor de normalización. Esta ecuación es usada para decidir los valores de la x 's en la matriz de diseño; Draper y Hunter maximizan esta ecuación respecto a las β 's y con respecto a los valores ajustados de las x 's ($x_{n+1}, \dots, x_{n+n^*}$), ya que después de que se conocen los valores de σ_Y^2 y de Ω es posible maximizar mediante métodos iterativos esta ecuación.

Como siempre deben realizarse suposiciones razonables que la computación requiere y que ayudan a la reducción del problema. Si se asume la expansión de la ecuación anterior está se convierte mediante una serie de Taylor truncada a partir de b^* en:

$$p_{n+n^*}(\beta|y_{n+n^*}) = k_{n+n^*} \exp\left[-\frac{[Y-n(b,x)]^T [Y-n(b,x)]}{2\sigma_Y^2}\right] \exp\left[-1/2\sigma_Y^2(\beta - b)^T X^T X(\beta - b) + (\beta - b^{(0)})^T \sigma_Y^2 [\Omega]^{-1}(\beta - b^{(0)})\right] \quad (11)$$

El factor de normalización k_{n+n^*} ahora es de gran interés y esta ecuación es la combinación de dos densidades de probabilidad multinormales.

Ahora podemos suponer que k_{n+n^*} es análogo a:

$$\Delta^{n+n^*} \quad (12)$$

Donde los símbolos $| |$ indican que es un determinante; si la densidad de probabilidad posterior de la ecuación P_{n+n^*} es maximizada respecto de β y el vector de nuevas observaciones $(x_{n+1}, \dots, x_{n+n^*})$, β dependerá de $b^{(0)}$ y de $b^{(n+n^*)}$ y no puede ser determinada antes de adoptar la observación. Pero cualquiera que sea la observación si β resulta estar en el extremo del último exponente del exponencial de la ecuación de P_{n+n^*} entonces se convierte en cero, de esta manera se puede concluir que k_{n+n^*} y a su vez el gradiente dado por la ecuación $\frac{\partial P_{n+n^*}}{\partial \beta}$ determinan el valor de P_{n+n^*} , consecuentemente se debería maximizar el gradiente como criterio para determinar los valores $(x_{n+1}, \dots, x_{n+n^*})$.

Se presenta una dificultad debido a que $(X^T X)$ contienen β , el vector de parámetros desconocidos, por tanto sustuiremos el estimado más reciente de β por $b^{(n)}$ en los elementos de X , en cuyo caso para evitar la confusión la matriz X puede llamarse $X^{(n)}$.

Ahora en la ecuación P_{n+n^*} la matriz X con n observaciones ya hechas y n^* observaciones por hacer será:

$$\begin{bmatrix} X_n \\ X_{n^*} \end{bmatrix} \quad (13)$$

Y los elementos individuales en X están dados por la ecuación (2):

$$[X_{ij}]$$

Ahora si los elementos de la matriz $[X^T X]$ se representan como A_{ij} , la evaluación de un modelo con dos parámetros quedaría así:

$$\Delta = \det \begin{Bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Bmatrix} \quad (14)$$

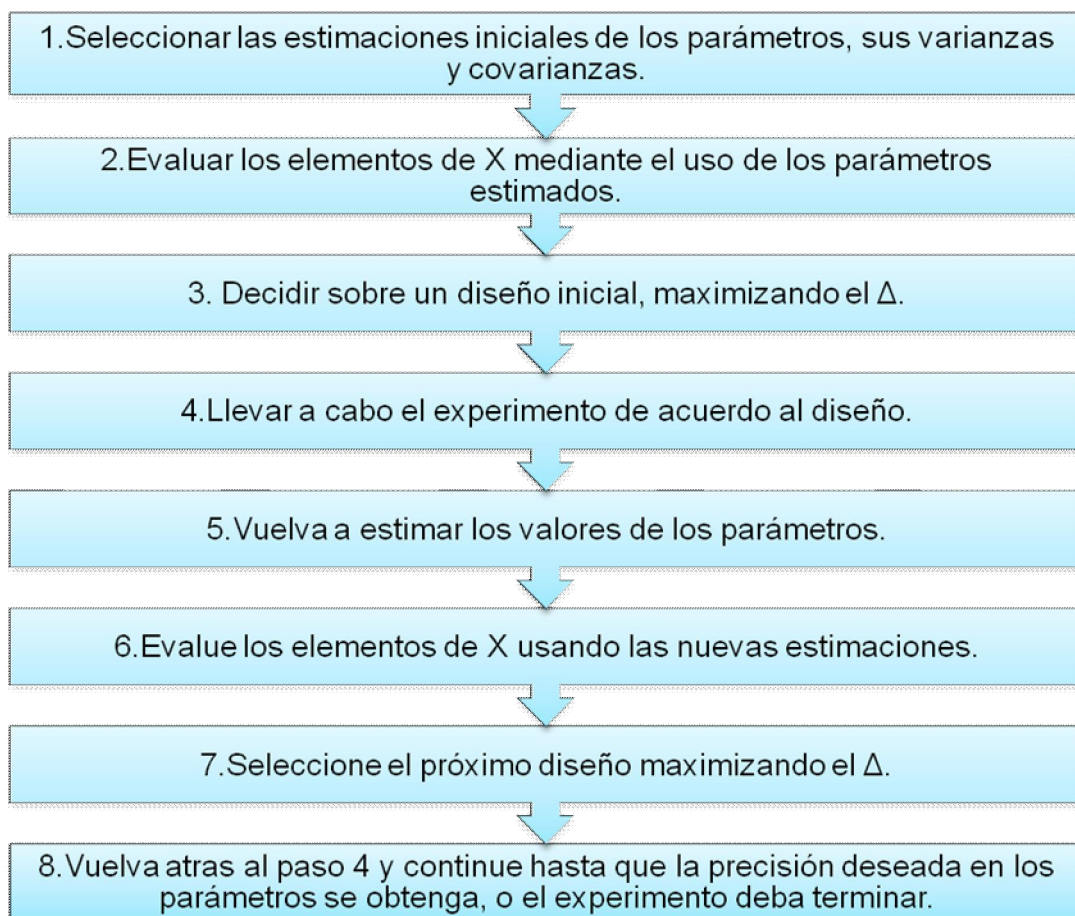
Nosotros llamaremos ρ el coeficiente de correlación entre β_1 y β_2 , y supondremos $\omega_{11} = \omega_1^2$ y $\omega_{22} = \omega_2^2$. Entonces:

$$\Delta = \det \begin{bmatrix} A_{11} + \frac{\sigma_y^2}{\omega_1^2(1-\rho^2)} & A_{12} - \frac{\sigma_y^2 \rho}{\omega_1 \omega_2 (1-\rho^2)} \\ A_{21} - \frac{\sigma_y^2 \rho}{\omega_1 \omega_2 (1-\rho^2)} & A_{22} + \frac{\sigma_y^2}{\omega_2^2(1-\rho^2)} \end{bmatrix} = \frac{(A_{11}A_{22} - A_{12}^2) + (A_{11}\frac{\sigma_y^2}{\omega_2^2} + A_{22}\frac{\sigma_y^2}{\omega_1^2} + 2A_{12}\frac{\sigma_y^2 \rho}{\omega_1 \omega_2})}{(1-\rho^2) + \left(\frac{\sigma_y^2}{\omega_1 \omega_2}\right)^2} \quad (15)$$

Si se supone $\omega_1 = \omega_2 = 0$ esto significaría que la estimación inicial de $b_1^{(0)}$ y $b_2^{(0)}$ tiene varianza cero y por tanto son certezas. Así el gradiente es máximo no importa cuáles sean las A_{ij} y no se necesitan experimentos adicionales. Otro caso extremo sería si $\omega_1 = \omega_2 = \infty$ porque entonces el $\Delta_1 = \det(X^T X) = A_{11}A_{22} - A_{12}^2$ el cual es el criterio desarrollado por Box y Lucas, ellos sugirieron minimizar la región de confianza asociada con el parámetro β para encontrar los valores de las variables independientes; sabiendo que la región de confianza es inversamente proporcional al determinante, podemos decir que mediante la maximización de este la región de confianza puede minimizarse.

En resumen el procedimiento general para el diseño secuencial de experimentos consiste en:

Figura 3. Algoritmo de diseño secuencial.



Fuente (HIMMELBLAU, David M., *Process Analysis by Statistical Methods*. John Wiley & Sons Inc (June 1970). Pages 496. ISBN-10: 047139985X). [5]

3.3.2.2 PROCESOS MULTIRESPUESTA.

Continuando con el desarrollo del Diseño Secuencial de Experimentos. La nueva característica es que se observan respuestas múltiples para un modelo que comprende varias ecuaciones de formas conocidas como:

$$\eta_1 = \beta_1 e^{-\beta_2 X} \quad (16)$$

$$\eta_2 = \frac{\beta_1 X_1}{\beta_2 X_2 + \beta_3} \quad (17)$$

Designaremos cada respuesta por:

$$Y_{ri} = \eta_{ri} + \epsilon_{ri} \quad 1 \leq r \leq v \quad (18)$$

Donde el índice r se refiere al número de modelos y el índice i a la cantidad de conjuntos de datos. El modelo η_r será una función de los m parámetros, β como la anterior y las variables independientes $X_{r,ik}$, donde el índice k designa el número de variables. Los errores ϵ_{ri} tienen:

$$\varepsilon\{\epsilon_{ri}\} = 0, \quad \varepsilon\{\epsilon_{ri}\epsilon_{sj}\} = 0 \quad \text{para } i \neq j$$

$$\varepsilon\{\epsilon_{ri}\epsilon_{rj}\} = 0, \quad \varepsilon\{\epsilon_{ri}\epsilon_{ri}\} = \sigma_{rr} \equiv \sigma_r^2$$

$$\varepsilon\{\epsilon_{ri}\epsilon_{si}\} = \sigma_{rs} = \sigma_{sr} \quad \text{para } r \neq s$$

De este modo las observaciones en un experimento para un modelo r y otro s pueden ser correlacionadas, la matriz de covarianza entre modelos cuyos elementos son σ_{rs} será denotada por Γ ; y los elementos de la inversa σ^{rs} serán denotados como Γ^{-1} . Note que para experimentos diferentes los errores son independientes.

Como en el caso anterior seguiremos los argumentos de Drapper y Hunter, y asumiremos que las observaciones de Y son representadas por una distribución normal multivariable después de n experimentos, la función de probabilidad para β será:

$$L(\beta/Y_n X, \sigma^{rs}) = \frac{|\Gamma^{-1}|^{n/2}}{(2\pi)^{nr/2}} \exp \left[-\frac{1}{2} \sum_{r=1}^v \right] \quad (19)$$

Donde:

$$(20)$$

Indica la sumatoria de las sumas de los cuadrados de las desviaciones de las Y_{ri} 's observadas en los n experimentos.

También como en el caso de respuesta única se asumió que la información a priori disponible después de n experimentos adicionales lleva a una densidad exactamente igual que la ecuación (6). Introduciendo la densidad anterior y la función de probabilidad de la ecuación (19) dentro del teorema de Bayes se obtiene la densidad posterior análoga a la ecuación (9). Para seleccionar n^* valores adicionales de la variable independiente, como en el caso anterior de respuesta simple, la densidad posterior puede ser usada como una densidad anterior, ambas con una función de probabilidad que corresponde con la ecuación (19). (Excepto que antes la suma en i es desde $i=n+1$ hasta $i=n+n^*$); entonces el teorema de Bayes para obtener la densidad final después de $n+n^*$ experimentos es:

$$P_{n+n^*}(\beta/Y_{n+n^*}) = \left[\frac{|\Omega|^{-1/2} |\Gamma^{-1}|^{-(n+n^*)/2}}{(2\pi)^{(m+(n+n^*)v)/2}} * \exp\left(-\frac{1}{2} \sum_{r=1}^v \sum_{s=1}^v \sigma^{rs} V_{rs}^{(n+n^*)}\right) * \exp\left(-\frac{1}{2} (\beta - b^{(0)})^T \Omega^{-1} (\beta - b^{(0)})\right) \right] \quad (21)$$

El resto del desarrollo es exactamente igual que en el caso de respuesta única. Para maximizar P_{n+n^*} con respecto a β y el vector de nuevos valores de $X_{ri} \equiv X_{r,i}$ para $i=n+1, \dots, n+n^*$, ahora expandiendo cada modelo obtenemos:

$$\eta_r(\beta, x_{ri}) \cong \eta_r(b^*, x_{ri}) + \sum_{j=1}^m (\beta_j - b_j^*) X_{r,ij} \quad (22)$$

Donde
$$X_{r,ij} = \frac{\partial \eta_r(b^*, x_{ri})}{\partial \beta_j} \quad X_{r,i} = \begin{bmatrix} X_{r,i1} \\ \vdots \\ X_{r,iq} \end{bmatrix} \quad (23)$$

Si reemplazamos b^* con b la máxima probabilidad estimada de β después de $(n+n^*)$ experimentos, el producto en cruz de la doble sumatoria quedaría así:

$$\sum_{r=1}^v \sum_{s=1}^v \sigma^{rs} V_{rs}^{(n+n^*)} = \sum_{r=1}^v \sum_{s=1}^v \sigma^{rs} \sum_{i=1}^{n+n^*} \left[\partial r_i - \sum_{j=1}^m (\beta_j - b_j^*) X_{r,ij} \right] * \left[\partial s_i - \sum_{j=1}^m (\beta_j - b_j^*) X_{s,ij} \right] = \sum_{r=1}^v \sum_{s=1}^v \sigma_i^{rs} \sum_{i01}^{n+n^*} \partial r_i * \partial s_i + \sum_{r=1}^v \sum_{s=1}^v (\beta - b)^T (\sigma^{rs} X_r^T X_s) (\beta - b) \quad (24)$$

Donde:
$$X_r = \begin{bmatrix} X_{r,11} & X_{r,12} & \dots & X_{r,im} \\ X_{r,21} & X_{r,22} & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ X_{r,(n+n^*)1} & X_{r,(n+n^*)2} & \dots & X_{r,(n+n^*)m} \end{bmatrix} \quad (25)$$

$$\partial r_i = Y_{r_i} - \eta_{r_i}(b, X_{r_i}) \quad (26)$$

Para obtener el diseño de los próximos n^* experimentos, debemos maximizar el determinante:

$$\Delta = \left| \sum_{r=1}^v \sum_{s=1}^v \sigma^{rs} X_r^T X_s + \Omega^{-1} \right| \quad (27)$$

Lo cual es proporcional al cuadrado del factor normalización. Todas las derivadas en X_r pueden ser evaluadas en $b^{(n)}$; y los elementos σ_{rs} se presumen conocidos.

Para dar un ejemplo específico del determinante para una situación con múltiple respuesta, tomaremos el caso de 2 modelos $v=2$ con tres coeficientes cada uno ($m=3$), $n=0$ (no hay experimentos completos todavía) todos los w 's = cero y $n^*=4$ (se planean cuatro experimentos) entonces:

$$\Delta = \left| \sigma^{11} (X_{1,n+n^*})^T (X_{1,n+n^*}) + \sigma^{12} (X_{1,n+n^*})^T (X_{2,n+n^*}) + \sigma^{21} (X_{2,n+n^*})^T (X_{1,n+n^*}) + \sigma^{22} (X_{2,n+n^*})^T (X_{2,n+n^*}) \right| \quad (28)$$

Donde:

$$X_{r,n+n^*} = \begin{bmatrix} X_{r,11} & X_{r,12} & X_{r,13} \\ X_{r,21} & X_{r,22} & X_{r,23} \\ X_{r,31} & X_{r,32} & X_{r,33} \\ X_{r,41} & X_{r,42} & X_{r,43} \end{bmatrix} \quad (29)$$

Para el modelo r y $X_{r,ij}$ han sido definidos anticipadamente. Y los estimados σ_{11} , σ_{22} , σ_{12} , pueden ser obtenidos de la réplica de datos.

3.4. PROGRAMACIÓN NO LINEAL. [6][9]

TÉCNICAS DE OPIMIZACIÓN RESTRINGIDA

3.4.1. MÉTODO COMPLEX.

En 1965 Box amplió el método Simplex de minimización no restringida para solucionar problemas de minimización restringida.

$$\text{Minimize } F(x) \quad \text{Sujeto a: } G(x) \leq 0 \quad (30)$$

En general se satisface la restricción $G(x) \leq 0$, pero no se satisface la restricción correspondiente a $G(x) = 0$.

Este método no puede manipular restricciones de igualdades no lineales.

La idea básica del método simplex consiste en la formación de una secuencia de figuras geométricas con $k = n + 1$ vértices de un espacio n -dimensional llamado simplex, es también ésta la idea general del método complex, una secuencia de figuras geométricas teniendo $k \geq n + 1$ vértices que se forman para encontrar un punto mínimo dentro de las restricciones. El método asume que un punto inicial que satisface todas las restricciones está disponible.

PROCEDIMIENTO ITERATIVO.

- a) Encuentre $k \geq n+1$ puntos, cada uno de los cuales satisface todas las m restricciones; en la práctica actual nosotros iniciamos con solo un punto posible x_1 y el resto de $k-1$ puntos son encontrados uno a uno por el uso de números aleatorios generados en el rango de 0 a 1 dentro de la ecuación:

$$x_{i,j} = x_i^l + r_{i,j}(x_i^u - x_i^l) \quad i = 1,2,3, \dots, n \quad j = 1,2,3, \dots, \quad (31)$$

- b) Donde $x_{i,j}$ es el i -ésimo componente del punto x_j y $r_{i,j}$ es un número aleatorio en el intervalo de 0 a 1 lo que hace que los puntos x_2, x_3, \dots, x_k generados de acuerdo con la ecuación (31) satisfagan las restricciones explícitas, pero pueden no satisfacer las restricciones implícitas.
- c) Por lo tanto tan pronto como un nuevo punto x_j es generado se verifica si satisface todas las restricciones, si en algún momento x_j viola alguna de las restricciones implícitas el punto x_j se mueve parcialmente medio camino hacia el centroide del resto de puntos.
- d) El centroide de todos los puntos esta dado por:

$$x_c = \frac{1}{j-1} \sum_{l=1}^{j-1} x_l \quad (32)$$

- e) Si el punto de prueba x_j continua violando las restricciones implícitas, el proceso de mover parcialmente hacia el centroide continua hasta encontrar un punto x_j posible. Finalmente se encontrará un conjunto de puntos posibles x_j

mediante la repetida utilización de este procedimiento. Al final tendremos los puntos $x_1, x_2, x_3, \dots, x_k$.

- f) La función objetivo es evaluada en cada uno de los K puntos (vértices), si el vértice x_h corresponde a el menor valor de la función se usa el proceso de reflexión para encontrar un nuevo punto x_r como sigue:

$$x_r = (1+\alpha)x_c - \alpha x_h \quad (33)$$

- g) Donde $\alpha \geq 1$ y x_c es el centroide de todos los puntos excepto x_h , el centroide se calcula de la siguiente manera:

$$(34)$$

- h) Como el problema es restringido el punto x_r debe ser probado como punto posible, dentro de las restricciones del problema; si el punto x_r es posible y $F(x_r) > F(x_h)$ el punto x_h es reemplazado por x_r y volvemos al paso 2.
- i) Si $F(x_r) < F(x_h)$ se calcula un nuevo punto de ensayo x_r mediante la reducción del valor de α a la mitad en la ecuación [c] y se prueba para saber si satisface la relación $F(x_r) > F(x_h)$, si aún $F(x_r) < F(x_h)$ el procedimiento de encontrar un nuevo valor de x_r con una reducción del valor de α se repite. Este procedimiento se repite si es necesario hasta que α llegue a ser menor que un valor pre-descrito de 10^{-6} , si un nuevo punto x_r con $F(x_r) > F(x_h)$ no puede obtenerse con un pequeño valor de α el punto x_r se descarta y el procedimiento completo de reflexión se reinicia mediante el uso del punto x_p (el cual equivale al segundo mayor valor de la función) en lugar de x_h .
- j) Si en cualquier momento el punto reflejado x_r encontrado en el paso 3 viola alguna de las restricciones implícitas el punto se mueve cierta distancia hacia el centroide mediante la siguiente ecuación:

$$x_{new} = \frac{1}{2} \quad (35)$$

- k) Hasta que este se convierta en un punto posible, esto hará que el método progrese hacia el punto óptimo a lo largo del complex sin que colapse dentro de su centro.
- l) Cada vez que el peor punto x_h del complex actual es reemplazado por un nuevo punto el complex se modifica y por lo tanto se debe chequear la convergencia del proceso.

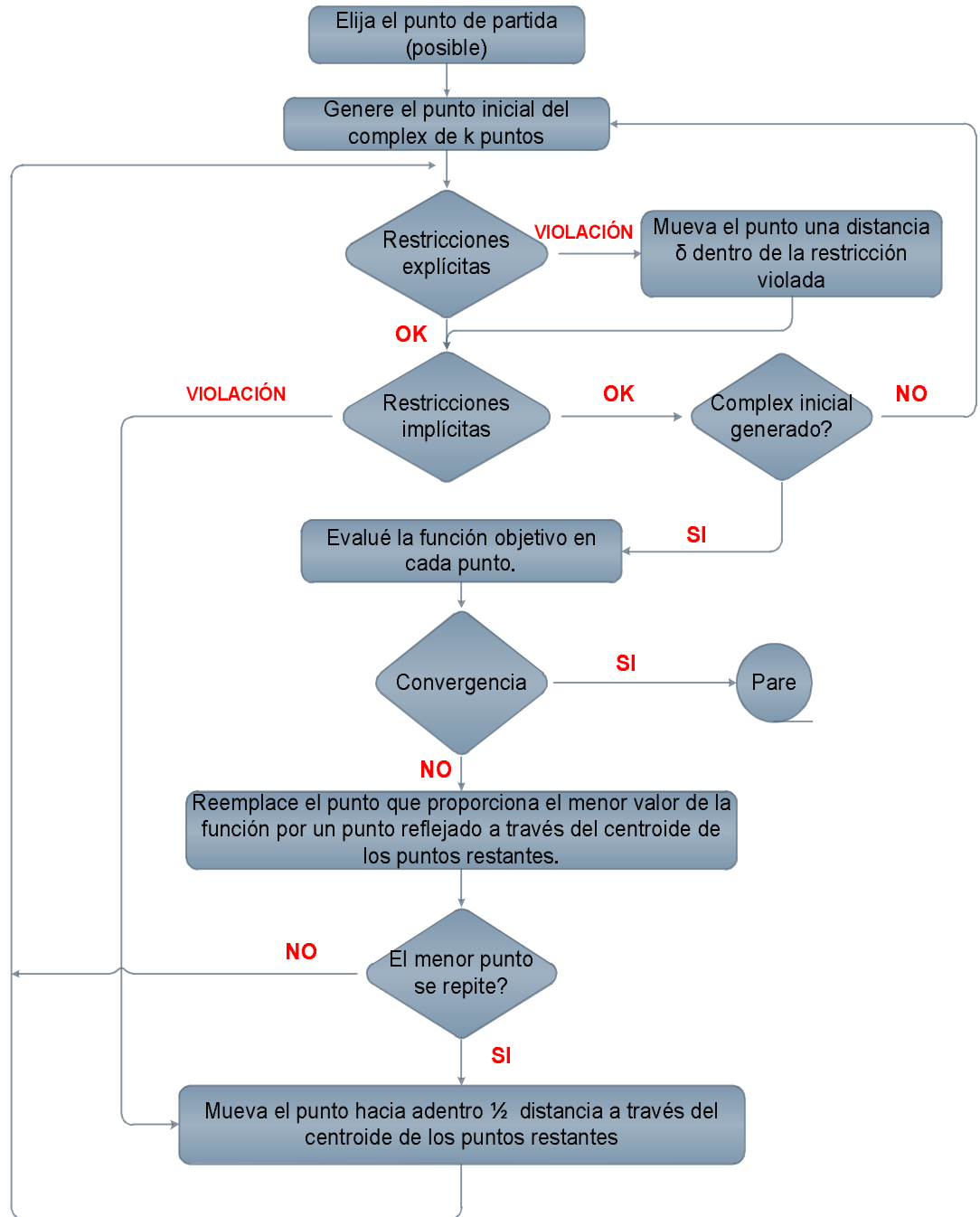
Se asumirá convergencia cuando cualquiera de las dos condiciones se cumple:

- El complex se encoje a un pequeño tamaño específico por ejemplo cuando la distancia entre cualquiera de los dos vértices de x_1, x_2, \dots, x_k llega a ser menor que un valor pre descrito ϵ_1 .
- La desviación estándar del valor de la función llega a ser suficientemente pequeña, por ejemplo cuando:

$$\left\{ \frac{1}{k} \sum_{j=1}^k [F(x) - F(x_j)]^2 \right\}^{1/2} \leq \epsilon \quad (36)$$

Donde x es el centroide de todos los k vértices del complex actual y ϵ es un número pequeño específico

Figura 4. Procedimiento iterativo método Complex



FUENTE: CARRILLO ABRIL Silvia J; FLÓREZ DÍAZ Ruth S. Algoritmo BOX.

4. RESULTADOS Y ANÁLISIS.

El objetivo de este trabajo de investigación es el desarrollo de un código computacional que mediante la técnica de diseño secuencial de experimentos permita encontrar la región de experimentación óptima para el subsiguiente experimento cada vez que un vector de parámetros β es encontrado. La coherencia del código DSE desarrollado fue probada usando el ejemplo reportado en la literatura, (HIMMELBLAU, David M., Process Analysis by Statistical Methods. Ejemplo 8.4-1 Optimal Sequential Designs For a Nonlinear Model) que se reporta en el Anexo 7.

En específico la búsqueda se hizo para los metamodelos creados por el Instituto Colombiano del Petróleo que describen las dos UDC de la refinería de Barrancabermeja.

Por lo tanto los resultados registrados en este capítulo corresponden a la región de experimentación encontrada para uno de los 184 metamodelos de la torre de destilación atmosférica 1 (TDA1).

Particularmente se analizaron los parámetros del metamodelo creado para la variable de salida: curva de destilación ASTM D86 95% para el diesel. [3]

4.1. TIPO Y FORMA DEL METAMODELO.

Los metamodelos desarrollados por el ICP son del tipo polinómico de segundo grado con errores promedio inferiores a 6.6%, 3% y 1,5% en el cálculo de flujos, propiedades de caracterización y temperaturas de corrientes atmosféricas respectivamente. [3], estos en específico son metamodelo con interacción binaria. Como se describe a continuación:

$$\hat{y}(X) = \hat{\beta}_0 + \sum_{i=1}^m \hat{\beta}_i \hat{x}_i + \sum_{i=1}^m \hat{\beta}_{ii} \hat{x}_{ii}^2 + \sum_{i=1}^{m-1} \sum_{j=i+1}^m \hat{\beta}_{ij} \hat{x}_{ij}$$

Donde:

= Es el vector de variables de salida.

= Es el vector de variables independientes normalizado.

4.2. DETERMINACION DEL DISEÑO INICIAL

Para la determinación de la matriz de experimentación óptima se utilizó el software Matlab, en el cual se elaboró un código de diseño secuencial de experimentos el cual parte de: un modelo, una matriz de datos entrada de las variables independientes y suposiciones iniciales para los coeficientes del modelo.

4.2.1. Variables de entrada.

Se sabe que la torre de destilación atmosférica 1 consta de 22 variables de entrada reportadas en el Anexo 2, compuestas por los flujos de los crudos componentes de la carga (F_Crudo1, F_Crudo2, F_Crudo3, F_Crudo4, F_Crudo5) las variables de operación (Temperaturas, Presiones, Flujos de vapor, etc.) y una variable de control de calidad de los productos JET y ACPM (PFE de ASTM D-86).

Se presento una reducción en el número de variables de 22 a 10 variables específicas, que se reportan en el Anexo 3, esto gracias a que en el diseño de experimentos realizado por el ICP los parámetros que acompañaban estas 10 variables fueron diferentes de cero.

4.2.2. Restricciones.

La región de experimentación usada está restringida por el diseño de los equipos en los que inciden las variables de entrada, por lo que las variables están limitadas a una región de optimización específica, reportada en el Anexo 2.

4.2.3. Variable de salida.

La variable de salida en este caso corresponde con la curva ASTM D-86 para el diesel, y los datos iniciales de las variables y los parámetros se tomaron de investigaciones previas [3].

El metamodelo al cual se le determinará los valores de las variables independientes que se usaran como diseño inicial para la búsqueda de los parámetros óptimos es: no lineal, y simplificado a partir de modelos rigurosos en PRO/II y diseño de experimentos (MBLHD).

4.2.4. Resultados de la simulación.

El metamodelo a optimizar es:

$$Y=B1+B2 *xi1+B3 *xi9+B4 *xi10+B5 *xi4^2+B6 *xi9^2+B7 *xi2 *xi3+B8 *xi4 *xi8+B9 *xi5 *xi7+B10 *xi6 *xi9+B11 *xi9 *xi10$$

Para este metamodelo las x_{ij} se reportan en el anexo 3 corresponden a diferentes flujos y propiedades que intervienen en la torre de destilación atmosférica 1 y que se encontraron son representativas la variable dependiente Y corresponde con la curva de destilación ASTM D-86 para el Diesel.

La optimización se da de acuerdo con el procedimiento iterativo descrito en el Anexo1. La matriz de datos iniciales de la variable independiente se encuentra reportada en el Anexo 4. Consta de 11 datos iniciales para cada variable X_{ij} .

El vector de estimados iniciales para los 11 parámetros del metamodelo se tomó de datos suministrados por el ICP y se reporta en el Anexo 5.

Para estos datos iniciales, se simuló el algoritmo DSE en Matlab.

El metamodelo inicial usado corresponde a:

$$Y = 545,5 + 4,8 \cdot x_{i1} + 36,8 \cdot x_{i9} + 3,3 \cdot x_{i10} + 3,3 \cdot x_{i4}^2 + 3,5 \cdot x_{i9}^2 + 3,6 \cdot x_{i2} \cdot x_{i3} + 3,5 \cdot x_{i4} \cdot x_{i8} + 3,4 \cdot x_{i5} \cdot x_{i7} + 2,9 \cdot x_{i6} \cdot x_{i9} + 5,1 \cdot x_{i9} \cdot x_{i10}.$$

Esta ecuación corresponde al metamodelo creado por el ICP mediante técnicas de diseño de experimentos diferentes al DSE (Diseño secuencial de experimentos).

Lo que se pretende es crear para este modelo una región de experimentación óptima que reduzca la incertidumbre cuando se desee optimizar los parámetros del mismo, debido a que el Instituto Colombiano del petróleo cuenta con una gran cantidad de datos iniciales se pretende reducir este grupo de datos mediante la optimización de las variables independientes, este proceso se realiza por medio de la maximización del determinante de la matriz de derivadas parciales respecto a cada parámetro de la ecuación inicial, este procedimiento es el que describe el diseño secuencial de experimentos y toma los coeficientes o β como constantes, mientras se varían los valores de las variables independientes. Este proceso se repite hasta que el valor de la ecuación, que en este caso corresponde con el determinante de la matriz $|X \cdot X^T|$ es máximo.

4.2.5. Variables independientes óptimas.

Tabla 6. Valores de X para el siguiente experimento encontrados con la metodología DSE.

Variable	F_Diluyente	F_San Fernando	F_Castilla	T_MPAAt m1	T_Conden1	F_VaporA tm1	F_Vapor T1B	F_Vapor T1C	T_CorteJet1	T_CorteDiesel1
xmin	3812,47	11,20	94,31	260,07	0,11	72599,88	16857,58	9614,39	500,77	700,06
xmax	47869,06	2229,37	30524,65	329,93	114,66	191400,12	35980,82	23985,61	569,79	755,38
CONDICIONES DE EXPERIMENTACION OPTIMAS	47853,66	1540,66	25302,14	270,64	0,17	191348,48	35973,05	9615,73	553,45	751,06
	47853,90	599,72	26555,78	325,71	28,72	81146,65	23363,82	23322,31	520,43	707,87
	3834,15	1812,55	8827,90	270,41	49,54	160761,73	16861,63	23982,24	530,18	740,93
	36880,27	1870,89	103,76	260,10	91,47	158687,70	35972,86	19446,70	516,78	755,35
	44741,86	1814,22	2008,13	312,89	77,64	168402,08	24519,19	14425,44	569,76	748,76
	3812,70	1609,01	109,47	320,53	0,13	133048,46	23605,35	12181,28	535,92	731,97
	35623,00	2207,47	4046,80	329,91	26,15	191386,63	30443,23	18206,23	510,21	707,50
	9359,27	1731,58	102,09	313,65	87,27	167029,60	16859,17	16971,67	500,79	746,14
	22510,86	574,29	30492,76	277,97	0,12	99821,63	18083,73	23982,46	549,19	715,92
	47856,66	107,63	10298,21	309,12	71,45	72627,82	16862,27	23046,31	510,40	724,93
	14280,55	1425,15	30518,31	329,92	34,71	104280,07	33785,71	13211,40	523,06	700,15

FUENTE: CARRILLO ABRIL Silvia J; FLÓREZ DÍAZ Ruth S. Variables óptima.

5. CONCLUSIONES.

Se obtuvo una matriz de variables independientes óptima dado que se maximizó el determinante de la matriz $X^T X$ y con esto se minimiza la región de confianza del modelo, lo que me garantiza que el punto máximo se encontrara dentro de la matriz reportada.

Se concluyó que aunque el método de Diseño Secuencial de Experimentos es una técnica muy funcional que permite obtener información valiosa a medida que se determinan nuevos puntos experimentales, presenta algunas limitaciones cuando el número de parámetros es muy elevado, debido a que requiere una mayor cantidad de memoria, para el almacenamiento de la información.

La metodología resulta eficiente en la optimización de la región de experimentación de modelos con un número de parámetros inferior a 10 por lo que resultaría conveniente analizar los metamodelos de la TDA1 y reducir el número de variables independientes. De esta manera se reduciría el tiempo de cómputo que es lo que resta atractivo a la metodología, pues se comprobó que el código funciona veloz y correctamente a menor número de parámetros.

De la información obtenida en la corrida del modelo, se puede decir que los datos obtenidos para los nuevos estimados del siguiente experimento son aleatorios y se requiere la justificación de los datos mediante comparaciones con datos reales de planta.

6. RECOMENDACIONES.

Se recomienda probar otra metodología de optimización no lineal, ya que se encontró que el método Complex al utilizar valores iniciales aleatorios presenta mayores posibilidades de no convergencia, asimismo es un método con alto gasto computacional que hace ineficiente el proceso de optimización en metamodelos grandes.

Para establecer si la metodología de diseño secuencial de experimentos implementada en esta tesis es adecuada es necesario que con los resultados de las variables independientes obtenidos se realicen la simulación de planta real en PRO II, que permita continuar el proceso de optimización de las variables.

El usuario encontrara dentro del código implementado en Matlab ayudas que le guiaran dentro del proceso, existen constantes intrínsecas al código que se tomaron de la literatura y que pueden variarse concienzudamente con el fin de personalizar el código a casos particulares, diferentes al estudiado en esta tesis.

Se recomienda usar la metodología DSE solo para metamodelos con un número de parámetros inferiores a 25, y se recomienda el uso de maquinas que tengan más de 4G de RAM, y preferiblemente de 64 bits.

BIBLIOGRAFIA.

- [1] HIMMELBLAU, David y BISCHOFF, Kenneth., Análisis y Simulación de Procesos. Austin, Texas: Editorial Reverte, 1992.
- [2] FLOUDAS, I.G. AKROTIRIANKIS, S. CARATZOULAS, C.A. MEYER, J. Kallrath. [Global optimization in the 21st century: Advances and challenges](#), Computer Aided Chemical Engineering, Volume 18, 2004, Pages 23-51C.A.
- [3] LOPEZ, Diana C., Optimización Simultanea de Composición de Cargas y Variables de Operación en Unidades de Destilación de Crudo. 2008.
- [4] <http://www.ecopetrol.com.co/especiales/RevistaInnova3ed/innovaciones.htm>
- [5] HIMMELBLAU, David M., Process Analysis by Statistical Methods. John Wiley & Sons Inc (June 1970). Pages 496. ISBN-10: 047139985X
- [6] EDGAR, Thomas F. y HIMMELBLAU, David M., Optimization of chemical processes. New York: McGraw-Hill, 1988. 652 p.: il. ISBN 0070189919.
- [7] DE ZAN, Arturo T., Experimental Sequential Designs for Logistic Regression Models: Diseños experimentales secuenciales para modelos logísticos de regresión. Revista Colombiana de Estadística. Bogotá. Volumen 31, No 2 (Diciembre, 2008); pp. 261 a 291. Disponible en: <http://www.emis.de/journals/RCE/V31/v31n2a09DeZan.pdf> .
- [8] AGARWAL, Anil K. y BRISK, Michael L., Sequential Experimental Design for Precise Parameter Estimation: Use of Reparameterization. Ind. Eng. Chem. Process Des. Dev. 1985, 24, Pages 203-207.
- [9] Cochran&Cox (William G. Cochran y Gertrude M Cox), Diseños experimentales, Editorial Trillas, S. A. México, D F edición 1965.
- [10] MARÍN Diazaraque, Juan Miguel. Profesor Titular de la Universidad Carlos III de Madrid. Introducción al Diseño de Experimentos [en línea].
<<http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/Disenno/IntroDE.pdf>>

[11] SINGIRESU S. Rao. Engineering Optimization Theory and Practice. Third Edition. School of Mechanical Engineering Purdue University West Lafayette, Indiana. John Wiley & Sons Inc.

[12] COELLO Coello, Carlos A. Optimización en ingeniería [PDF/Adobe Acrobat en línea] <http://delta.cs.cinvestav.mx/~ccoello/optimizacion/clase4-opt-2009.pdf>

[13] SÁNCHEZ, Omar. Normalización de variables [PDF/Adobe Acrobat en línea] <<http://omarsanchez.net/documents/normalizacion.pdf>>

[14] CAMPILLO Rodríguez, Rafael. INTRODUCCIÓN A MATLAB: Métodos Numéricos en Matlab. [PDF/Adobe Acrobat en línea]. http://www.fimexalapa.org/maestros/rcampillo/archivos/Matlab_Tema5.pdf

[15] ACEVEDO Duarte, Leonardo; LÓPEZ Cárdenas, Diana Carolina y MAHECHA Bohórquez, Cesar Augusto. Análisis jerárquico de variables de operación de una torre de destilación de crudo usando datos de planta, modelos de simulación y diseño de experimentos. Disponible en: REVISTA ION Vol. 21(2008) < <http://revistas.uis.edu.co/index.php/revistaion/article/view/7>>

[16] RUIDIAZ Fiallo, Meira Liliana y SALEDO Ardila, Carlos Armando Diseño secuencial de experimentos aplicado a un reactor de planta piloto tipo riser. Bucaramanga, 2007, 82 h. Tesis de grado (Ingeniero Químico) Universidad Industrial de Santander. Facultad de fisicoquímicas, Escuela de Ingeniería Química. Disponible en: Base de datos segundo piso.

[17] OJEDA Delgado, Karina Angélica y SÁNCHEZ Tuirán, Eduardo Luis Desarrollo de software para diseño secuencial de experimentos aplicados al proceso de hidrosulfurización de fracciones pesadas del petróleo. Bucaramanga 2005, 110 h. Tesis de grado (Ingeniero Químico) Universidad Industrial de Santander. Facultad de fisicoquímicas, Escuela de Ingeniería Química. Disponible en: Base de datos segundo piso.

[18] REIS dos Santos, M. Isabel y REIS dos Santos, Pedro M. Sequential experimental desings for nonlinear regression metamodels in simulation. Disponible en: ELSEVIER [en línea] publicado el 3 de julio de 2008 <<http://web.tagus.ist.utl.pt/~maria.santos/smpt08press.pdf>>

[19] AGARWAL, Anil K y BRISK, Michael L. El diseño secuencial de experimentos para la estimación Precisa de Parámetros. 1 Uso de la Reparametrización Ind. Eng. Chem. Process Des. Dev. 1985, 24, 203- 207. 207 210

[20] AGARWAL, Anil K y BRISK, Michael L. El diseño secuencial de experimentos para la estimación Precisa de Parámetros. 2 Criterio de Diseño. Ind. Eng. Chem. Process Des. Dev. 1985, 24, 207 – 210.

[21] JOHNSON, Eric W. y BROCKMAN, Jay B. Medición y Análisis de Procesos de diseño secuencial. En: ACM Transactions on Design Automation of Electronic Systems (TODAES) Volumen 3 N° 1, (Enero 1998).

[22] KEYS, Anthony C. y REES, Loren Paul. A sequential-design metamodeling strategy for simulation Optimization. En: Computers & Operations Research Volumen 31, N°11, (Septiembre 2004), Paginas 1911-1932.

[23] MARTINEZ CIFUENTES, Víctor. Simulación de procesos en ingeniería química. 1era edición. México D.F.: Editorial Plaza y Valdés, 2000. 187p. ISBN 9688567558, 9789688567555

ANEXO 1.

DSE (*Algoritmo de Diseño Secuencial de Experimentos*)

A. PROPÓSITO.

Este programa encuentra los estimados óptimos para los parámetros de una ecuación no lineal multivariable sujeta a restricciones lineales mediante el uso combinado de la metodología de Diseño secuencial de experimentos, y el método de maximización de Box (Método Complex).

Optimice $G(x)$ con parámetros $\beta_1, \beta_2, \dots, \beta_j$

Sujeto a $G_k \leq x_k \leq$

Las variables implícitas son funciones dependientes de las variables explícitas independientes las restricciones son constantes.

B. MÉTODO.

El procedimiento se basa en el método "DSE" de Box, W.G.Hunter, Draper y otros colaboradores quienes hicieron importantes contribuciones en el desarrollo del procedimiento de diseño secuencial para reducir la incertidumbre en la estimación de parámetros.

Este método es una técnica de búsqueda secuencial la cual ha probado ser efectiva en la solución de problemas con funciones objetivo no lineales. El procedimiento se divide en dos partes, la primera es la búsqueda de la ecuación a optimizar que resulta ser el determinante de la matriz de derivadas parciales respecto a β (parámetros) de la ecuación inicial; la segunda parte es la optimización o maximización del determinante, este procedimiento se realizó mediante el Algoritmo Complex. Este encuentra el máximo global dado al hecho de que el set inicial de puntos es aleatorio dentro de la región posible.

El algoritmo completo procede de la siguiente manera:

Se supone que un sencillo y adecuado modelo ha sido seleccionado y que deseamos encontrar los mejores estimados para sus parámetros para ello debemos:

1. Ingresamos la ecuación con el número de parámetros y variables.

$$n = n(\beta, x) \quad (1)$$

2. Luego derivamos parcialmente la ecuación respecto de cada β y vamos creando la matriz de derivadas parciales.

$$[X_{ij}] = \frac{dn_i(b, x_i)}{d\beta_j} \quad j = 1, 2, 3, \dots \quad (2)$$

$$\begin{bmatrix} X_{1,1} & \dots & X_{1,m} \\ \vdots & \ddots & \vdots \\ X_{n+n^*,1} & \dots & X_{n+n^*,m} \end{bmatrix} \quad (13)$$

3. Posteriormente se halla la transpuesta de la matriz $X = (X^T)$ y multiplicamos estas dos matrices para hallar Δ^{n+n^*} mediante la siguiente ecuación.

Donde:

- σ_Y^2 es la varianza del modelo Y.
- Ω^{-1} Es una matriz de covarianza para los β en la cual cada elemento esta especificado.

4. Este es el último paso en la técnica de diseño secuencial de experimentos y consiste en hallar el determinante de la ecuación anterior.

(12)

5. Dado que hasta ahora no hemos introducido ningún dato numérico en el procedimiento el determinante que obtenemos no consiste en un número sino que será una multiplicación de variables, esta será la ecuación que vamos a maximizar mediante el algoritmo Complex.

6. Encuentre $k \geq n+1$ puntos, cada uno de los cuales satisface todas las m restricciones; en la práctica actual nosotros iniciamos con solo un punto posible x_1 y el resto de $k-1$ puntos son encontrados uno a uno por el uso de números aleatorios generados en el rango de 0 a 1 dentro de la ecuación:

$$x_{i,j} = x_i^l + r_{i,j}(x_i^u - x_i^l) \quad i = 1,2,3, \dots, n \quad j = 1,2,3, \dots, k$$

7. Cuando un nuevo punto x_j es generado se verifica si satisface las restricciones explícitas si no se mueve el punto una distancia δ hacia el centroide.

$$x_c = \frac{1}{j-1} \sum_{l=1}^{j-1} x_l$$

8. Si se viola una restricción implícita se mueve el punto hacia el centroide mediante la ecuación $x_{ij_{new}} = \frac{1}{2}(x_c + x_{ij})$. Al final tendremos los puntos $x_1, x_2, x_3, \dots, x_k$.

9. La función se evalúa en todos los punto x_{ij} , y se le da el nombre de x_h al que corresponde con el menor valor de la función, este punto es el que se refleja mediante la ecuación $x_r = (1+\alpha)x_c - \alpha x_h$ (33)

10. Si el punto es factible se pregunta $F(x_r) > F(x_h)$ si la respuesta es si entonces x_r pasa a ser el nuevo x_h y se regresa al paso 9. Si la respuesta es no entonces se halla un nuevo punto x_r reduciendo α a la mitad y se prueba esta reducción continua hasta encontrar un x_r que satisfaga $F(x_r) > F(x_h)$ y sea remplazado en el paso 9, el proceso termina cuando los valores de la

función se repiten beta unidades en un número lambda de iteraciones consecutivas.

11. Si en algún momento alfa llega a ser menor de 10^{-6} se descartan los puntos randomicos y el proceso se reinicia.

C. DESCRIPCIÓN DEL PROGRAMA

El código consiste de un programa principal (**DSE**) y siete subrutinas (**BOX**, **CENTR**, **CHECK**, **FUNC**, **GenerarFunciones**, **DeterminanteLU**, y el **Encabezado**).

El programa principal DSE se encarga de encontrar el determinante de la matriz $X^T X$ y llamar a las subrutinas. La subrutina DeterminanteLU acompaña al programa principal y construye una matriz LU a partir de la matriz $X^T X$, con el fin de facilitar el cálculo del Determinante que se hace lento cuando las ecuaciones son complejas o se tienen muchos parámetros. El determinante calculado es una ecuación que pasa a la subrutina BOX encargada de maximizar, se asumen los parámetros β constantes y luego se maximiza nuevamente el determinante partiendo de los valores de la variable dependiente calculados, esto es con x's constantes y los parámetros β variables.

Para la maximización, la subrutina BOX utiliza: CENTR que se encarga de hallar el centroide de los puntos x_{ij} , CHEK que se encarga de evaluar las restricciones, FUNC que es la subrutina dirigida a obtener los valores de la función evaluada.

Por último la subrutina Encabezado posee la información de autoría, y nos da la posibilidad de dar indicaciones generales del manejo del código.

ANEXO 2.

Tabla 2. Variables de entrada del metamodelo (curva ASTM D-86 para el reflujo medio de la (TDA1)).

VARIABLES DE ENTRADA	DESCRIPCION	UNIDAD	MAX	MIN
F_Diluyente	Flujo volumétrico de diluyente en carga	BPD	3812	47869
F_San Fernando	Flujo volumétrico de crudo 2 en carga	BPD	11	2229
F_Rubiales	Flujo volumétrico de crudo 3 en carga	BPD	124	10206
F_Nare-Teca	Flujo volumétrico de crudo 4 en carga	BPD	49	9389
F_Castilla	Flujo volumétrico de crudo 5 en carga	BPD	94	30525
T_PrecalT1	Temperatura del crudo antes del horno	°F	430	500
T_HornoAtm1	Temperatura del crudo a la salida del horno	°F	660	710
T_CimaAtm1	Temperatura de la corriente de cima de la torre	°F	270	320
P_CimaAtm1	Presión de cima de la torre	Psig	13	20
F_MPAAtm1	Flujo volumétrico de reflujo medio	BPD	10591	17544
F_LPAAtm1	Flujo volumétrico de reflujo inferior	BPD	10239	17992
T_UPAAtm1	Temperatura del retorno del reflujo cima	°F	75	135
T_MPAAtm1	Temperatura del retorno del reflujo medio	°F	260	330
T_LPAAtm1	Temperatura del retorno del reflujo inferior	°F	460	500

VARIABLES DE ENTRADA	DESCRIPCION	UNIDAD	MAX	MIN
T_Conden1	Temperatura del condensador	°F	0	115
P_Conden1	Presión del condensador	Psig	9	18
F_VaporAtm1	Flujo de vapor de despojo a torre	Lb/día	72600	191400
F_VaporT1A	Flujo de vapor a despojadora de JET (T-A)	BPD	16858	35904
F_VaporT1B	Flujo de vapor a despojadora de ACPM (T-B)	BPD	16858	35981
F_VaporT1C	Flujo de vapor a despojadora de GOA (T-C)	BPD	9614	23986
T_CorteJet1	PFE ASTM D-86 del JET	°F	501	570
T_CorteDiesel1	PFE ASTM D-86 del ACPM	°F	700	755

Fuente: Instituto Colombiano del Petróleo. (ICP) (LOPEZ. Diana Carolina, Optimización Simultánea de Composición de Cargas y Variables de Operación de Unidades de Destilación de Crudo, Bucaramanga, Universidad Industrial de Santander. 2008). [3].

ANEXO 3.

Tabla 3. Variables independientes del metamodelo (curva ASTM D-86 para el reflujo medio de la (TDA1)) reducido a sus variables de mayor incidencia.

VARIABLES DE ENTRADA	DESCRIPCION	VARIABLE	UNIDAD	MAX	MIN
F_Diluyente	Flujo volumétrico de diluyente en carga	Xi1	BPD	3812	47869
F_San Fernando	Flujo volumétrico de crudo 2 en carga	Xi2	BPD	11	2229
F_Castilla	Flujo volumétrico de crudo 5 en carga	Xi3	BPD	94	30525
T_MPAAtm1	Temperatura del retorno del reflujo medio	Xi4	°F	260	330
T_Conden1	Temperatura del condensador	Xi5	°F	0	115
F_VaporAtm1	Flujo de vapor de despojo a torre	Xi6	Lb/día	72600	191400
F_VaporT1B	Flujo de vapor a despojadora de ACPM (T-B)	Xi7	BPD	16858	35981
F_VaporT1C	Flujo de vapor a despojadora de GOA (T-C)	Xi8	BPD	9614	23986
T_CorteJet1	PFE ASTM D-86 del JET	Xi9	°F	501	570
T_CorteDiesel1	PFE ASTM D-86 del ACPM	Xi10	°F	700	755

Fuente: Instituto Colombiano del Petróleo. (ICP) (LOPEZ. Diana Carolina, Optimización Simultánea de Composición de Cargas y Variables de Operación de Unidades de Destilación de Crudo, Bucaramanga, Universidad Industrial de Santander. 2008). [3]

ANEXO 4.

Tabla 4. Datos iniciales de las variables independientes (X).

Variable	F_Diluyente	F_SanFer	F_Castilla	T_MPAAtm1	T_Conden1	F_VaporA tm1	F_Vapor T1B	F_Vapor T1C	T_Corte Jet1	T_CorteDie sel1
xmin	3812	11	94	260	0	72600	16858	9614	501	700
xmax	47869	2229	30525	330	115	191400	35981	23986	570	755
DATOS INICIALES	28719,0	11,4	13915,6	329,9	68,2	143160,0	31756,8	9902,4	567,1	704,9
	18995,9	11,2	12207,4	290,4	82,9	164760,1	19699,2	9758,4	554,0	743,6
	32035,8	32,9	9533,4	275,5	1,0	112680,0	31065,6	19348,8	540,4	717,2
	36408,9	38,2	8591,4	271,4	4,7	90119,9	30528,0	21393,6	551,2	715,3
	30712,3	42,1	9330,3	303,3	101,5	73319,9	27340,8	13531,2	560,4	707,6
	24024,0	52,6	19748,1	292,3	97,4	185640,1	27494,4	21220,8	548,5	712,3
	25870,5	43,8	5813,4	269,7	95,6	160680,1	23232,0	13214,4	517,8	710,1
	25487,9	63,0	12758,5	281,6	48,0	158040,1	35212,8	13761,6	532,8	741,5
	25719,0	68,2	15613,9	273,1	73,3	173160,1	29875,2	14971,2	565,6	713,5
	36591,8	84,8	5737,5	291,3	70,5	164040,1	18432,0	9844,8	524,6	702,6
24140,8	63,4	5218,3	260,6	105,5	74519,9	28377,6	21595,2	536,2	713,6	

ANEXO 5.

Tabla 5. Datos iniciales de los parámetros del metamodelo (curva ASTM D-86 para el reflujo medio de la (TDA1)) reducido.

PARAMETRO	ESTIMADO INICIAL
B1	545,5123075
B2	4,8335495
B3	36,82146084
B4	3,258680844
B5	3,322550425
B6	3,504103757
B7	3,611794043
B8	3,516164856
B9	3,410513334
B10	2,940873749
B11	5,101587745

Fuente: Instituto Colombiano del Petróleo. (ICP) (LOPEZ. Diana Carolina, Optimización Simultánea de Composición de Cargas y Variables de Operación de Unidades de Destilación de Crudo, Bucaramanga, Universidad Industrial de Santander. 2008). [3]

ANEXO 6.

- Código DSE

```
clear all; dcs1 = 1;
while dcs1 == 1
    Encabezado();
    % ***** Ingreso de los valores iniciales *****
    r = input('Número de ecuaciones ');           % Número de ecuaciones
    nPar = input('Numero de Parametros ');        % Número de Observaciones
    nObs = nPar;                                  % Número de Parámetros
    nVar = input('Numero de Variables ');         % Número de Variables
    disp(' ');
    Mats_Func = sym(zeros(nPar,nPar,r));
    for k = 1 : r
        % ***** Ingreso de la ecuación *****
        Y = fileread('D:\Users\Power\Documents\MATLAB\Ruth\Datos\Ecuacion.txt'); % Ecuación
        simbólica cargada de un .txt
        disp(['Ecuación... ', Y]);
        disp(' ');
        % ***** Generación de las ecuaciones simbólicas *****
        Vcts_Y = sym(zeros(1,nPar));
        for i = 1 : nPar
            Vcts_Y(i) = GenerarFunciones(Y,i);
        end
        % ***** Generación de los valores simbólicos *****
        Mats_Xij = sym(zeros(nPar,nVar));          % Matriz de X simbólicas
        for i = 1 : nPar
            for j = 1 : nVar
                Mats_Xij(i,j) = sym(sprintf('x%d%d',i,j));
            end
        end
        Vcts_Bj = sym(zeros(nPar,1));              % Vector de B simbólicos
        for i = 1 : nPar
            Vcts_Bj(i) = sym(sprintf('B%d',i));
```

```

end
% ***** Derivación de las funciones con respecto a cada B *****
Mats_dB = sym(zeros(nPar)); % Matriz de derivadas simbólicas
for i = 1 : nPar
    for j = 1 : nPar
        Mats_dB(i,j) = diff(Vcts_Y(i),Vcts_Bj(j));
    end
end
disp(Mats_Xij);
disp(Vcts_Bj);
disp(Mats_dB);
Mats_Func(:,k) = Mats_dB;
end
% ***** Ingreso de los valores iniciales de Sigam y Omega *****
Mats_Sigma = sym(zeros(r)); % Matriz de Sigma
for i = 1 : r
    for j = 1 : r
        Mats_Sigma(i,j) = '1';
    end
end
Mats_Omega = sym(zeros(nPar)); % Matriz de Omega
for i = 1 : nPar
    for j = 1 : nPar
        Mats_Omega(i,j) = '0';
    end
end
% ***** Se construye la matriz para calcular el determinante *****
Mats_X = sym(zeros(nPar)); % Matriz para hallar el determinante
for i = 1 : r
    for j = 1 : r
        Mats_X = Mats_X + ((Mats_Sigma(i,j)*(Mats_Func(:,i)).*Mats_Func(:,j))+Mats_Omega);
    end
end
disp('Matriz para hallar el determinante');
disp(' ');

```

```

disp(Mats_X);
% ***** Calculo del determinante *****
Dets_X = DeterminanteLU(Mats_X); % Determinante Simbolica
disp(' ');
disp('Determinante ok');
disp(' ');
% Obtención de los datos Iniciales
Vct_Bi = csvread('D:\Users\Power\Documents\MATLAB\Ruth\Datos\Vct_Bi.dat'); % Vector B
inicial
Mat_Xi = csvread('D:\Users\Power\Documents\MATLAB\Ruth\Datos\Mat_Xi.dat'); % Matriz X
inicial
disp(Mat_Xi);
disp(Vct_Bi);
% ***** Ingreso de los limites 1 *****
Vct_LimG1 = csvread('D:\Users\Power\Documents\MATLAB\Ruth\Datos\Vct_G1.dat'); %
Vectores de limites inferiores
Vct_LimH1 = csvread('D:\Users\Power\Documents\MATLAB\Ruth\Datos\Vct_H1.dat'); %
Vectores de limites superiores
% ***** Maximización de la función *****
BOX(Dets_X,Mats_Xij,Vcts_Bj,Mat_Xi,Vct_Bi,Vct_LimG1,Vct_LimH1);
% ***** Ingreso de los limites 2 *****
Vct_LimG2 = csvread('D:\Users\Power\Documents\MATLAB\Ruth\Datos\Vct_G2.dat'); %
Vectores de limites inferiores
Vct_LimH2 = csvread('D:\Users\Power\Documents\MATLAB\Ruth\Datos\Vct_H2.dat'); %
Vectores de limites superiores
% ***** Maximización de la función *****
BOX(Dets_X,Vcts_Bj,Mats_Xij,Vct_Bi,Mat_Xi,Vct_LimG2,Vct_LimH2);
dcs1 = 0;
while (dcs1 ~= 1) && (dcs1 ~= 2)
    dcs1 = input('Continuar (1) ; Salir (2) ');
end
end
end

```

- Código Box

```

function BOX(F,Mat_Xsym,Vct_Bsym,Mat_Xini,Vct_Bini,Vct_G,Vct_H)
% **** Función encargada de maximizar el determinante ****
format long;
% **** Ingreso de los valores iniciales ****
% alpha = input('factor de reflexion (1.3) ');
% beta = input('Parametro beta de convergencia ');
% gama = input('Parametro gama de convergencia ');
% delta = input('Factor de reflexion ');
% itmax = input('maximo de Iteraciones ');
% symG1_imp = input('Función implícita... ','s');
% symG2_imp = input('Función implícita... ','s');
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% *****Valores Recomendados*****
alpha = 1.3;
beta = 0.1;
gama = 5;
delta = 0.001;
itmax = 1000;
symG1_imp = '1';
symG2_imp = '0';
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% **** Creación de la matriz ****
[nf,nc] = size(Mat_Xini);
N = nf * nc; % Columnas de la matriz del complex inicial
K = 2 * N; % Filas de la matriz del complex inicial
Mat_Xij = zeros(K,N);
% Matriz del complex inicial

for i = 1 : nf
    for j = 1 : nc
        Mat_Xij(1,j+(nc*(i-1))) = Mat_Xini(i,j);
    end
end

```

```

end
% % ***** Ingreso de los limites *****
% Vct_G = csvread('D:\lex\Programación\Matlab\Ruth\Datos\Vct_G.dat');      % Vectores de
limites inferiores
% Vct_H = csvread('D:\lex\Programación\Matlab\Ruth\Datos\Vct_H.dat');      % Vectores de
limites superiores
Mat_R = rand(K-1,N);                % Matriz de números randómicos
% ***** Genera la matriz del complex inicial *****
for i = 1 : K-1
    for j = 1 : N
        Mat_Xij(i+1,j) = Vct_G(j) + Mat_R(i,j)*(Vct_H(j)-Vct_G(j));
    end
end
% ***** Evaluacion de la convergencia *****
cont1 = 0;
cont2 = 0;
conver = 0;
while conver ~= 1 && cont1 <= itmax
    alpha1 = alpha;
    % Factor de reflexión
    cont1 = cont1 + 1;
    % Contador de iteraciones
    conver = 0;
    % Convergencia
    % ***** Comprobación de las restricciones *****
    for i = 1 : K
        Vct_Xc = CENTR(N,K,Mat_Xij,i);
    % Vector centroide de X
        Mat_Xij(i,:) =
CHECK(N,Vct_G,Vct_H,Mat_Xij(i,:),delta,Vct_Xc,symG1_imp,symG2_imp,Mat_Xsym,Vct_Bsym,Vc
t_Bini,nf,nc);
    end
    % ***** Evalua la función *****
    Vct_F = zeros(K,1);
% Vector de las funciones evaluadas

```

```

for i = 1 : K
    Vct_F(i) = FUNC(F,Mat_Xsym,Vct_Bsym,Mat_Xij(i,:),Vct_Bini,nf,nc);
end
% ***** Valor y posición de la menor función evaluada *****
[fmin,p_fmin] = min(Vct_F);
% ***** *****
Vct_Xc = CENTR(N,K,Mat_Xij,p_fmin);
% ***** Refleja los puntos de Xh *****
Vct_Xr = zeros(1,N); %
for i = 1 : N
    Vct_Xr(i) = alpha*(Vct_Xc(i)-Mat_Xij(p_fmin,i))+Vct_Xc(i);
end
% ***** Comprobación de las restricciones de los X reflejados*****
Vct_Xr =
CHECK(N,Vct_G,Vct_H,Vct_Xr,delta,Vct_Xc,symG1_imp,symG2_imp,Mat_Xsym,Vct_Bsym,Vct_Bi
ni,nf,nc);
% ***** Comparacion de las funciones de Xh y Xr *****
Fh = FUNC(F,Mat_Xsym,Vct_Bsym,Mat_Xij(p_fmin,:),Vct_Bini,nf,nc); %
Fr = FUNC(F,Mat_Xsym,Vct_Bsym,Vct_Xr,Vct_Bini,nf,nc); %
if (Fr > Fh)
    Mat_Xij(p_fmin,:) = Vct_Xr;
    disp('Fr > Fh');
else
    disp('Fr <= Fh');
    p = p_fmin;
    while Fr <= Fh && alpha1 >= 0.000001
        alpha1 = alpha1/2;
        for i = 1 : N
            Vct_Xr(i) = alpha1*(Vct_Xc(i)-Mat_Xij(p,i))+Vct_Xc(i);
        end
        Vct_Xr =
CHECK(N,Vct_G,Vct_H,Vct_Xr,delta,Vct_Xc,symG1_imp,symG2_imp,Mat_Xsym,Vct_Bsym,Vct_Bi
ni,nf,nc);
        Fh = FUNC(F,Mat_Xsym,Vct_Bsym,Mat_Xij(p_fmin,:),Vct_Bini,nf,nc);
        Fr = FUNC(F,Mat_Xsym,Vct_Bsym,Vct_Xr,Vct_Bini,nf,nc);

```

```

end
if alpha1 < 0.000001
    disp('Alpha es menor a 0.000001');
    % ***** BUscA el segundo menor *****
    Vct_Faux = sort(Vct_F);
    for i = 1 : K
        if Vct_F(i) == Vct_Faux(2)
            p_2fmin = i;
            Vct_Xc = CENTR(N,K,Mat_Xij,p_2fmin);
            for j = 1 : N
                Vct_Xr(j) = alpha*(Vct_Xc(j)-Mat_Xij(p_2fmin,j))+Vct_Xc(j);
            end
            Vct_Xr =
CHECK(N,Vct_G,Vct_H,Vct_Xr,delta,Vct_Xc,symG1_imp,symG2_imp,Mat_Xsym,Vct_Bsym,Vct_Bi
ni,nf,nc);
            Frp = FUNC(F,Mat_Xsym,Vct_Bsym,Vct_Xr,Vct_Bini,nf,nc);
        end
    end
    if (Frp > Fh)
        Mat_Xij(p_fmin,:) = Vct_Xr;
        Fr = Frp;
    else
        alpha1 = alpha;
        p = p_2fmin;
        disp(' Frp < Fh ');
        while Frp <= Fh && alpha1 >= 0.000001
            alpha1 = alpha1/2;
            for i = 1 : N
                Vct_Xr(i) = alpha1*(Vct_Xc(i)-Mat_Xij(p_2fmin,i))+Vct_Xc(i);
            end
            Vct_Xr =
CHECK(N,Vct_G,Vct_H,Vct_Xr,delta,Vct_Xc,symG1_imp,symG2_imp,Mat_Xsym,Vct_Bsym,Vct_Bi
ni,nf,nc);
            Frp = FUNC(F,Mat_Xsym,Vct_Bsym,Vct_Xr,Vct_Bini,nf,nc);
        end
    end
end

```


- Subrutina CENTR

```
function Vct_Xc = CENTR(N,K,Mat_X,m)
% ***** Función encargada de calcular el centroide *****
Vct_Xc = zeros(1,N);           % Vector centroide
for i = 1 : N
    suma = 0;
    for j = 1 : K
        if j ~= m
            suma = suma + Mat_X(j,i);
        end
    end
    Vct_Xc(i) = suma/(K-1);
end
```

- Subrutina CHECK

```
function Vct_Xnew =
CHECK(N,Vct_Linf,Vct_Lsup,Vct_X,delta,Vct_Xc,symG1,symG2,Mat_Xsym,Vct_Bsym,Vct_Bini,nf,
nc)
% ***** Función encargada de comprobar el cumplimiento de las restricciones *****
% ***** Comprueba el cumplimiento de las restricciones explícitas *****
for i = 1 : N
    while Vct_X(i) < Vct_Linf(i) || Vct_X(i) > Vct_Lsup(i)
        if Vct_X(i) < Vct_Linf(i)
            Vct_X(i) = Vct_X(i) + delta;
        end
        if Vct_X(i) > Vct_Lsup(i)
            Vct_X(i) = Vct_X(i) - delta;
        end
    end
end
```

```

        end
    end
end
% ***** Comprueba el cumplimiento de las restricciones implícitas *****
G1 = FUNC(symG1,Mat_Xsym,Vct_Bsym,Vct_X,Vct_Bini,nf,nc);
G2 = FUNC(symG2,Mat_Xsym,Vct_Bsym,Vct_X,Vct_Bini,nf,nc);
while G1 <= G2
    disp('Implicita');
    Vct_X = (Vct_X + Vct_Xc)/2;
    G1 = FUNC(symG1,Mat_Xsym,Vct_Bsym,Vct_X,Vct_Bini,nf,nc);
    G2 = FUNC(symG2,Mat_Xsym,Vct_Bsym,Vct_X,Vct_Bini,nf,nc);
    disp(Vct_X);
    disp(G1);
    disp(G2);
    pause;
end
Vct_Xnew = Vct_X;

```

- Subrutina DeterminanteLU

```
function DetsX = DeterminanteLU(MatsX)
```

```
% ***** Función encargada de hallar el determinante por medio del metodo LU *****
```

```
m = size(MatsX);
```

```
MatL = sym(zeros(m));
```

```
MatU = sym(zeros(m));
```

```
for i = 1 : m
```

```
    for j = 1 : m
```

```

if j == i

    MatL(i,j) = 1;

    sum = 0;

    for k = 1 : i - 1

        sum = sum + MatL(i,k) * MatU(k,j);

    end

    MatU(i,j) = MatsX(i,j) - sum;

end

if j > i

    sum = 0;

    for k = 1 : i - 1

        sum = sum + MatL(i,k) * MatU(k,j);

    end

    MatU(i,j) = MatsX(i,j) - sum;

end

if j < i

    sum = 0;

    for k = 1 : j - 1

        sum = sum + MatL(i,k) * MatU(k,j);

    end

    MatL(i,j) = (1 / MatU(j,j)) * (MatsX(i,j) - sum);

end

```

```

    end

end

disp(MatU(i,j));

disp(MatL(i,j));

DetsX = 1;

for i = 1 : m

    DetsX = DetsX * MatU(i,i);

end

disp('Determinante LU');

disp(DetsX);

```

- Subrutina GenerarFunciones

```

function VctFunc = GenerarFunciones(Func,n)
% ***** Función encargada de generar las funciones simbólicas *****
new = int2str(n);
nc = length(Func);
VctFunc = Func; % Vector de funciones simbólicas
strFunc = "";
for i = 1 : nc
    if VctFunc(i) == 'i'
        strFunc = strcat(strFunc,new);
    else

```

```

        strFunc = strcat(strFunc,VctFunc(i));
    end
end
VctFunc = strFunc;

```

- Subrutina FUNC

```

function valF = FUNC(F,Mat_sym1,Mat_sym2,Vct_1,Mat_2,m,n)

% ***** Función encargada de evaluar la función dada por el determinante

Mat_1 = zeros(m,n);

cnt = 0;

for i = 1 : m

    for j = 1 : n

        cnt = cnt + 1;

        Mat_1(i,j) = Vct_1(1,cnt);

    end

end

Mat_sym = [Mat_sym1,Mat_sym2];

% Matriz con los datos simbólicos

Mat = [Mat_1,Mat_2];

% Matriz con los datos numéricos

valF = double(subs(F,{Mat_sym},{Mat}));

```

% Valor de la función evaluada

- Subrutina Encabezado

```
function Encabezado()
% ***** Función encargada de generar el encabezado de las presentaciones *****
clc;
disp('***** ICP-UNIVERSIDAD INDUSTRIAL DE SANTANDER *****');
disp(' ');
disp('***** Autores: RUTH SMITH FLOREZ DIAZ *****');
disp('***** SILVIA CARRILLO ABRIL *****');
disp(' ');
disp('***** ESTIMACION DE VARIABLES MEDIANTE LA METODOLOGIA DSE *****');
disp(' ');
```

Anexo 7.

Ejemplo 8.4-1 Optimal Sequential Designs for Nonlinear Model.

***** ICP-UNIVERSIDAD INDUSTRIAL DE SANTANDER *****

***** Autores: RUTH SMITH FLOREZ DIAZ *****

***** SILVIA CARRILLO ABRIL *****

*ESTIMACION DE PARAMETROS MEDIANTE LA METODOLOGIA DSE *

Número de ecuaciones 1

Numero de Parámetros 2

Numero de Variables 2

Ecuación

$$(1-\exp(B1*xi1))*(1-B2*xi2)$$

Matriz inicial

[x11, x12]

[x21, x22]

Vector de Parámetros inicial

B1

B2

Matriz de derivadas parciales

$$[x11*\exp(B1*x11)*(B2*x12 - 1), x12*(\exp(B1*x11) - 1)]$$

$$[x21*\exp(B1*x21)*(B2*x22 - 1), x22*(\exp(B1*x21) - 1)]$$

Matriz para hallar el determinante.

$$[\begin{matrix} & & & x11^2 \cdot \exp(2 \cdot B1 \cdot x11) \cdot (B2 \cdot x12 - 1)^2 + \\ x21^2 \cdot \exp(2 \cdot B1 \cdot x21) \cdot (B2 \cdot x22 - 1)^2, & x11 \cdot x12 \cdot \exp(B1 \cdot x11) \cdot (B2 \cdot x12 - 1) \cdot (\exp(B1 \cdot x11) - 1) + \\ & x21 \cdot x22 \cdot \exp(B1 \cdot x21) \cdot (B2 \cdot x22 - 1) \cdot (\exp(B1 \cdot x21) - 1) \end{matrix}]$$

$$[\begin{matrix} x11 \cdot x12 \cdot \exp(B1 \cdot x11) \cdot (B2 \cdot x12 - 1) \cdot (\exp(B1 \cdot x11) - 1) + \\ x21 \cdot x22 \cdot \exp(B1 \cdot x21) \cdot (B2 \cdot x22 - 1) \cdot (\exp(B1 \cdot x21) - 1), \\ x12^2 \cdot (\exp(B1 \cdot x11) - 1)^2 + x22^2 \cdot (\exp(B1 \cdot x21) - 1)^2 \end{matrix}]$$

Determinante LU

$$\begin{aligned} & (x11^2 \cdot \exp(2 \cdot B1 \cdot x11) \cdot (B2 \cdot x12 - 1)^2 + x21^2 \cdot \exp(2 \cdot B1 \cdot x21) \cdot (B2 \cdot x22 - 1)^2) \cdot (x12^2 \cdot (\exp(B1 \cdot x11) - 1)^2 - (x11 \cdot x12 \cdot \exp(B1 \cdot x11) \cdot (B2 \cdot x12 - 1) \cdot (\exp(B1 \cdot x11) - 1) + x21 \cdot x22 \cdot \exp(B1 \cdot x21) \cdot (B2 \cdot x22 - 1) \cdot (\exp(B1 \cdot x21) - 1))^2) / \\ & (x11^2 \cdot \exp(2 \cdot B1 \cdot x11) \cdot (B2 \cdot x12 - 1)^2 + x21^2 \cdot \exp(2 \cdot B1 \cdot x21) \cdot (B2 \cdot x22 - 1)^2) + x22^2 \cdot (\exp(B1 \cdot x21) - 1)^2 \end{aligned}$$

El método tiene convergencia

1.0e+005 *

0.000010999976311	0.000000000011845	0.000010999998501
0.001439999194719	9.095382953050368	

1.0e+005 *

0.000010999976398	0.000000000007315	0.000010999998399
0.001439999114047	9.095382013779936	

1.0e+005 *

0.000010999976375 0.000000000004285 0.000010999998413
0.001439999128465 9.095382193475922

1.0e+005 *

0.000010999976376 0.000000000002842 0.000010999998425
0.001439999151179 9.095382534638695

1.0e+005 *

0.000010999976347 0.000000000000806 0.000010999998435
0.001439999186343 9.095382927759289

1.0e+005 *

0.000010999976317 0.000000000002342 0.000010999998520
0.001439999193801 9.095383132286242