

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE QUIMICA

86

TECNICAS DE ESTIMACION DE PARAMETROS
PARA MODELOS MATEMATICOS DE VELOCIDAD
DE REACCION QUIMICA



GUILLERMO SERGIO DEL CASTILLO VELIZ

JOSE FRANCISCO PICHARDO RANGEL

INGENIERO QUIMICO

1 9 7 7



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Tesis 1977
M-~~87~~ 88
FECHA _____
PROC _____



JURADO ASIGNADO ORIGINALMENTE SEGUN EL TEMA 43

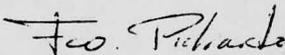
Presidente	Prof. LEOPOLDO RODRIGUEZ SANCHEZ
V o c a l	" JUAN GUERRERO MORALES
Secretario	" CARITINO MORENO PADILLA
1er. Suplente	" GILDARDO ACOSTA RUIZ
2o. Suplente	" CARLOS BAZAN VILLEGAS

Sitio donde se desarrolló el tema: ³⁴ Centro de Servicios de
Cómputo y Biblioteca de la Facultad de Química, U.N.A.M.

Nombre y firma de los sustentantes:



GMO. SERGIO DEL CASTILLO VELIZ



JOSE FCO. PICHARDO RANGEL

Nombre y firma del asesor del tema:



ING. CARITINO MORENO PADILLA

Dedico cariñosamente esta tesis a quienes con su ayuda y -
estímulo, lograron que terminara mis estudios profesionales.

A mi Madre

Margarita Rangel.

A mi Padre

Francisco Pichardo.

A mi Esposa

Beatriz

A mi Hermana

Ana

Francisco .

Este trabajo lo dedico con todo cariño a mi padre

Guillermo del Castillo Victoria.

Guillermo.

Nuestro agradecimiento al
Prof. Caritino Moreno Padilla
por su valiosa ayuda durante
la realización de este trabajo.

"El sentido que se puede vislumbrar
no es el sentido eterno.
El nombre que se puede nombrar
no es el nombre eterno.
Al otro lado de lo que se puede nombrar se encuentra
el principio del mundo.
A este lado de lo que se puede nombrar está
el nacimiento de los seres."

del libro chino Tao-te-king.

Índice.

CAPITULO 1

1 Introducción 1

CAPITULO 2

2 Generalidades sobre técnicas de estimación de parámetros. 4

2.1 Técnica de Mínimos Cuadrados 4

2.1-1 Técnica de Búsqueda Directa 6

2.1-2 Técnica Simplex 8

2.1-3 Técnica de Gauss 11

2.1-4 Técnicas de Gradientes 15

2.1-5 Técnica de Marquardt 19

2.1-6 Técnica de Powell 23

2.1-7 Técnica Spiral 25

2.2 Técnicas Probabilísticas 30

2.2-1 Técnica de Máxima Semejanza 30

a) Algoritmo Maximum Likelihood (Máxima Semejanza) 34

b) Algoritmo Maximum-Simplex 34

c) Algoritmo Maximum-Gradientes Simples 35

d) Algoritmo Maximum-Gradientes Conjugados 35

e) Algoritmo Maximum-Powell 35

f) Algoritmo Maximum-Rosenbrock 38

2.2-2 Técnica Bayesiana 44

CAPITULO 3

3 Modelos de velocidad de reacción 50

3.1 Modelos Homogéneos	52
3.1-1 Reacciones Irreversibles	53
3.1-2 Reacciones Reversibles	56
3.1-3 Sistemas Reaccionantes Complejos	59
a) Reacciones en paralelo	59
b) Reacciones consecutivas	60
3.2 Modelos Heterogéneos	61
CAPITULO 4	
4 Comparación de técnicas de estimación de parámetros	69
CAPITULO 5	
5 Conclusiones	126
APENDICE	
Programas	130
1. Técnica de Gauss	130
2. Técnica de Gradientes Simples	130
3. Técnica de Gradientes Conjugados	130
4. Técnica de Marquardt	130
5. Técnica Spiral	130
6. Técnica Maximum-Simplex	134
7. Técnica Bayes-Simplex	137
8. Técnica Bayes-Rosenbrock	141
BIBLIOGRAFIA	146

Introducción.

En la experimentación común el modelo de un proceso es desconocido, pero las variables principales de éste pueden ser conocidas. Es necesario desarrollar varios modelos empíricos ó teóricos y seleccionar el más adecuado para representar el proceso.

El objeto de un experimento es encontrar alguna relación --- existente entre las variables experimentales; tal relación será el modelo en esta investigación y dependerá de constantes ó parámetros no definidos numéricamente. Una vez seleccionado el modelo es necesario ajustarlo, ó sea, encontrar los valores estimados -- mas próximos para los parámetros involucrados en dicho modelo, pa ra obtener una mejor representación de los datos experimentales.

Las técnicas para estimar los parámetros en base a la forma del modelo pueden dividirse en : técnicas lineales y técnicas no lineales. Para modelos lineales el proceso de cálculo es directo y no es necesario un procedimiento iterativo. Las técnicas no lineales son iterativas, ó sea, los valores estimados de los parámetros deben satisfacer un criterio de convergencia por medio de un algoritmo. Este trabajo tratará sobre estas últimas técnicas.

Los parámetros a estimar serán los coeficientes que se encuentran en un modelo matemático de un proceso. El valor estimado de un parámetro es el mas cercano a su valor real; el estimador para un parámetro desconocido es una función de los elementos de una muestra aleatoria, siendo el mismo también una variable aleatoria. Ejemplificando:

P = parámetro que interesa estimar.

V = valor estimado del parámetro.

S = estimador del parámetro.

Los conceptos anteriores no deben ser confundidos y se desea por supuesto que $|V - P|$ sea muy pequeño. Por lo tanto para considerar qué valores estimados de parámetros son buenos, es necesario que presenten las siguientes características:

a) No sesgado.

El valor estimado V de un parámetro P es no sesgado si su valor esperado $E(V)$ es igual al valor P .

b) Consistencia.

Un estimador es consistente si los valores estimados tienden a aproximarse a los valores P tanto como el tamaño de la muestra sea incrementado; ó sea, $E[(V - P)^2]$ tiende a cero mientras el tamaño de la muestra tiende a infinito.

c) Eficiencia.

La comparación de las variancias de los valores estimados, - indicará cual es el mas eficiente, es decir, el que tenga la menor variancia. Este criterio debe considerarse junto a de no sesgado, ya que un valor estimado ligeramente sesgado con una variancia pequeña puede preferirse en un momento dado a un valor estimado no sesgado con una gran variancia.

d) Suficiencia.

Si el valor estimado V del parámetro P es suficiente, no habrá otro valor estimado de P que suministre información adicional acerca de P ; este concepto es equivalente al requerimiento de mínima variancia, es decir, un valor estimado suficiente es necesariamente mas eficiente y también, en consecuencia, consistente.

Se ha encontrado en la literatura que los métodos más importantes para la estimación de parámetros en modelos no lineales -- son: las Técnicas de Mínimos Cuadrados y las Técnicas Probabilísticas. Este trabajo estará enfocado a la aplicación de ambas técnicas.

CAPITULO 2

Generalidades sobre técnicas de estimación de parámetros.

2.1 Técnicas de Mínimos Cuadrados.

Del análisis de un proceso se determina la forma funcional de un modelo y es posible estimar los parámetros del mismo a partir de la información experimental que se posea. Para la estimación de parámetros, si el modelo es no lineal, es factible aplicar el método de mínimos cuadrados para ajustar el modelo a los datos.

Conociendo la forma funcional de la ecuación para el modelo, considérese lo siguiente:

y = valores de variables dependientes observados aleatorios.

x = valores de variables independientes no aleatorios controlables

P = valor real del parámetro en el modelo.

e = algún tipo de error experimental no observable.

V = valores estimados de los parámetros.

r = valores de variables dependientes predichos a partir del modelo (respuesta), $E(y/x)$.

El subíndice u representará el u -ésimo experimento, siendo $u = 1, 2, \dots, n$.

Tanto " y " como " x " se consideran variables continuas, ó sean números reales en algún rango finito ó infinito.

Sea el modelo (en notación de matrices):

$$\underline{r} = f(\underline{x}; \underline{P}) \quad \dots(2-1)$$

donde \underline{P} representa la matriz $p \times 1$ de valores estimados de los parámetros, siendo p el número de parámetros estimados; es decir,

$$x = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1q} \\ x_{21} & x_{22} & \dots & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nq} \end{bmatrix} \quad P = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_p \end{bmatrix}$$

donde q es el número de variables no aleatorias independientes -- controlables, siendo n mayor que p .

Cada valor observado de y_u para un conjunto dado de x , se relaciona al valor esperado de y_u por:

$$y_u = f(x_u; P) + e_u \quad \dots(2-2)$$

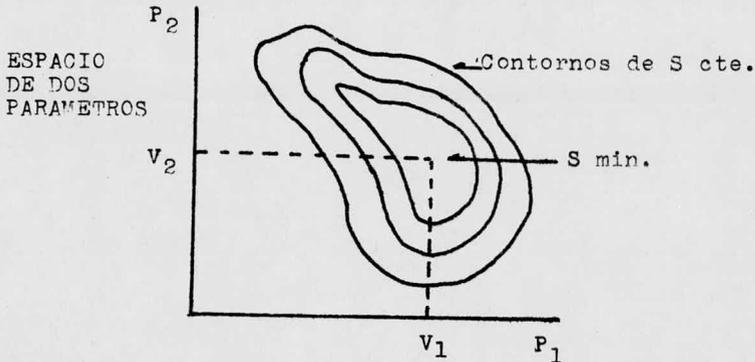
La técnica usual de mínimos cuadrados encuentra valores estimados V para los parámetros P para los cuales la suma de cuadrados

$$S(V) = \sum_{u=1}^n (y_u - r_u)^2 \quad \dots(2-3)$$

sea mínima.

Los errores pueden considerarse como variables independientes distribuidas normalmente teniendo un valor esperado de cero y una variancia constante. Es decir, la función de semejanza (likelihood) $L(V | y, x)$ (ver sección 2.2-1) tiene el mismo contorno que la función que expresa la suma de cuadrados de las desviaciones entre los valores observados "y" y los valores "r". En este caso la técnica se reduce a un problema lineal en los parámetros, y tal contorno es elipsoidal, mientras que si el modelo es no lineal, el contorno es distorsionado de acuerdo a la severidad de la no linealidad. Sin embargo aún con modelos no lineales los contornos son cercanamente elípticos en la vecindad inmediata del mínimo -- $S(V)$.

Enseruida se muestra la interpretación geométrica del espacio de parámetros aplicado a un modelo no lineal.



Típicamente la superficie del contorno es bastante atenuada en algunas direcciones y elongada en otras.

Las técnicas más usuales en mínimos cuadrados para la determinación de parámetros en modelos no lineales son:

- 1.- Técnica de Búsqueda Directa.
- 2.- Técnica Simplex.
- 3.- Técnica de Gauss.
- 4.- Técnica de Gradientes.
- 5.- Técnica de Marquardt.
- 6.- Técnica de Powell.
- 7.- Técnica Spiral.

2.1-1 Técnica de Búsqueda Directa.

La técnica de búsqueda directa, que tiene su fundamento en el algoritmo desarrollado por R. Hooke y T. A. Jeeves que puede conducir a buenos valores de partida a ser usados en otras técnicas más elaboradas, se utilizó específicamente en este trabajo para obtener valores estimados de las funciones objetivas de acuerdo a los modelos estudiados, para la técnica de Gradientes Simples y para Gradientes conjugados, las cuales se describen en la-

sección (2.1-4).

Esta técnica no requiere el cálculo de derivadas, sin embargo es lenta en comparación con métodos derivativos, especialmente cuando el número de parámetros es muy grande. El algoritmo procede de la siguiente forma. Un punto base es escogido, es decir, valores iniciales para los parámetros y un incremento para cada parámetro; en este punto la función objetivo es evaluada. Enseguida se realizan búsquedas locales en cada dirección desplazando un parámetro P_i^0 una distancia s_i a cada lado y evaluando la función objetivo para ver si una función de valor menor es obtenida. Si no hay decremento en la función, en caso de estar buscando el mínimo, el tamaño de paso es reducido y búsquedas nuevas son hechas a partir del punto previo obtenido. Si el valor de la función objetivo por el contrario ha decrecido, $P_i^0 + \Delta P_i^0$ es adoptado como un nuevo valor estimado de P_i , ó sea una nueva búsqueda es llevada a cabo y el valor de la función se verifica nuevamente. Esta expansión continúa mientras el valor de la función objetivo de--- crezca. Si un movimiento exploratorio conduce a una nueva dirección exitosa, entonces los ΔP_i 's son reducidos gradualmente hasta que una nueva dirección pueda ser definida ó cada ΔP_i sea menor que alguna tolerancia seleccionada. Para un valor muy pequeño de ΔP_i , según el criterio personal, habiéndose mejorado la función objetivo se llega a un punto óptimo, especificándose también un número de ciclos determinado, esto es, se satisface un criterio de convergencia.

2.1-2 Técnica Simplex. (ref. 2)

Esta técnica no requiere derivación y utiliza patrones regulares de búsqueda, representados por una figura geométrica. Por ejemplo, si se tienen dos parámetros debe considerarse un triángulo equilátero, ó sea tres puntos, mientras que si son tres parámetros se tiene un tetraedro, es decir, cuatro puntos.

Dicha técnica es útil en la búsqueda de un máximo ó un mínimo de funciones objetivas restringidas ó no restringidas, siendo eficaz especialmente cuando es mayor el número de parámetros en el modelo.

Describiendo el algoritmo para una función que contiene k - parámetros los cuales son considerados como variables independientes para maximizarla, se tiene la tabla siguiente:

VERTICE	COORDENADAS				...
					P_k
1	P_1	P_2	P_3		$\dots P_k$
2	$P_1 + \Delta P_1$	P_2	P_3		$\dots P_k$
3	$P_1 + 1/2 \Delta P_1$	$P_2 + \Delta P_2$	P_3		$\dots P_k$
4	$P_1 + 1/2 \Delta P_1$	$P_2 + 1/2 \Delta P_2$	$P_3 + \Delta P_3$		$\dots P_k$
:					
:					
k+1	$P_1 + 1/2 \Delta P_1$	$P_2 + 1/2 \Delta P_2$	$P_3 + 1/2 \Delta P_3$		$\dots P_k + \Delta P_k$

Los incrementos $\Delta P_1, \Delta P_2, \dots, \Delta P_k$ son seleccionados arbitrariamente. Ejemplificando para una función sencilla que contiene dos parámetros:

$$y = P_1^2 / P_2 + P_2$$

Dando como incrementos

$$\Delta P_1 = 2.0 \qquad \Delta P_2 = 4.0$$

y teniendo como valores iniciales $P_1 = 2.0$ y $P_2 = 2.0$ se forma la siguiente tabla:

	Coordenadas		
Vértice	P_1	P_2	y
1	2.0	2.0	4.0
2	4.0	2.0	10.0
3	3.0	6.0	7.5

Para obtener las coordenadas del nuevo vértice (4), es decir, para la reflexión del simplex, se hace lo siguiente.

Dado que en el vértice (1) es donde se tiene el menor valor de la función:

$$P_1^{(4)} = 2(P_1^{(2)} + P_1^{(3)}/k) - P_1^{(1)}$$

$$P_1^{(4)} = 2(4 + 3/2) - 2 = 5.0$$

Se resta el valor de la coordenada correspondiente al valor mínimo de la función "y". En este caso es $P_1^{(1)}$.

Similarmente,

$$P_2^{(4)} = \frac{2(P_2^{(2)} + P_2^{(3)})}{k} - P_2^{(1)}$$

$$P_2^{(4)} = 2\left(\frac{2 + 6}{2}\right) - 2 = 6$$

Por lo tanto:

$$y^{(4)} = (5)^2/6 + 6 = 10.1$$

Puesto que el valor de $y^{(4)}$ es mayor que los valores $y^{(2)}$, $y^{(3)}$, el incremento fué bien seleccionado. En caso contrario debe reducirse.

Enseguida se sustituyen las coordenadas del vértice (1) por las del nuevo vértice (4):

Vértice	Coordenadas		
	P ₁	P ₂	y
4	5	6	10.1
2	4	2	10.0
3	3	6	7.5

La búsqueda continúa de la misma forma obteniéndose:

Vértice	Coordenadas		
	P ₁	P ₂	y
4	5	6	10.1
2	4	2	10.0
5	6	2	20.0

Vértice	Coordenadas		
	P ₁	P ₂	y
4	5	6	10.1
6	7	6	14.16
5	6	2	20.0

Por lo tanto el máximo obtenido para la función es 20.0 --- puesto que este valor es mayor que 14.16 que fué calculado para el vértice (6).

De trabajos realizados respecto a esta técnica, se ha constatado la poca eficiencia de usarla individualmente. En el presente trabajo la técnica Simplex se utilizó para efectuar la maximización de la función de máxima semejanza (likelihood), (ver 2.2-1) y la maximización de la función de densidad de probabilidad dada por el teorema de Bayes (ver sección 2.2-2).

2.1-3 Técnica de Gauss.

Esta técnica llamada también Gauss-Newton, Gauss Seidel, -- (ref. 2), lineariza el modelo en una serie truncada de Taylor para hacer uso de un análisis lineal, y lograr obtener la suma de mínimos cuadrados de las desviaciones por medio de una secuencia iterativa de cálculos. Se suponen valores iniciales para los parámetros, y ciclicamente son obtenidos nuevos valores estimados por un método que tiene su fundamento en el algoritmo de Newton-Raphson; los cálculos son repetidos hasta que un criterio de convergencia sea satisfecho. El método ha probado ser efectivo si se suponen buenos valores iniciales para los parámetros.

Sea la función de mínimos cuadrados:

$$S(V) = \sum_{u=1}^n (y_u - r_u)^2$$

definida anteriormente. Expandiendo la ecuación del modelo --- $r = f(x;P)$ por series de Taylor y reteniendo solo los términos lineales se tiene:

$$\begin{aligned} r &= r_0 + (\partial r / \partial P_1)_0 (P_1 - V_1) + (\partial r / \partial P_2)_0 (P_2 - V_2) + \dots + \\ &\quad (\partial r / \partial P_p)_0 (P_p - V_p) \\ &= r_0 + \sum_{j=1}^p (\partial r / \partial P_j)_0 \Delta V_j \end{aligned} \quad \dots(2-4)$$

El subscripto "0" designa cantidades evaluadas en los valores iniciales supuestos. El modelo linearizado es sustituido en la función de mínimos cuadrados. Para minimizar $S(V)$ se toman -- las derivadas parciales de tal función con respecto a cada coeficiente (V_1, V_2, \dots, V_p) e igualando a cero:

$$\frac{\partial \sum_{u=1}^n (y_u - (r_u)_0 - \sum_{j=1}^p (\partial r_u / \partial P_j)_0 \Delta V_j)^2}{\partial (\Delta V_j)} = 0 \quad \dots(2-5)$$

lo cual da p ecuaciones normales:

$$2 \sum_{u=1}^n (y_u - (r_u)_0 - (\partial r_u / \partial P_1)_0 \Delta V_1 - (\partial r_u / \partial P_2)_0 \Delta V_2 + \dots) (\partial r_u / \partial P_1)_0 = 0$$

.....

$$2 \sum_{u=1}^n (y_u - (r_u)_0 - (\partial r_u / \partial P_1)_0 \Delta V_1 - (\partial r_u / \partial P_2)_0 \Delta V_2 + \dots) (\partial r_u / \partial P_p)_0 = 0$$

Si $E_u = y_u - (r_u)_0$, estas ecuaciones lineales pueden escribirse de la siguiente manera:

$$\Delta V_1 \sum_{u=1}^n (\partial r_u / \partial P_1)_0 (\partial r_u / \partial P_1)_0 + \Delta V_2 \sum_{u=1}^n (\partial r_u / \partial P_2)_0 (\partial r_u / \partial P_1)_0 + \dots + \Delta V_p \sum_{u=1}^n (\partial r_u / \partial P_p)_0 (\partial r_u / \partial P_1)_0 = \sum_{u=1}^n E_u (\partial r_u / \partial P_1)_0 \quad \dots(2-6)$$

.....

$$\Delta V_1 \sum_{u=1}^n (\partial r_u / \partial P_1)_0 (\partial r_u / \partial P_p)_0 + \Delta V_2 \sum_{u=1}^n (\partial r_u / \partial P_2)_0 (\partial r_u / \partial P_p)_0 + \dots + \Delta V_p \sum_{u=1}^n (\partial r_u / \partial P_p)_0 (\partial r_u / \partial P_p)_0 = \sum_{u=1}^n E_u (\partial r_u / \partial P_p)_0$$

Los valores buscados son ΔV_j ; introduciendo la notación de matrices:

$$(x_{uj}) = \partial r_u(x_u; \underline{V}) / \partial P_j$$

donde $u = 1, 2, \dots, n$; $j = 1, 2, \dots, p$

siendo además

$$\underline{x}_u = (x_{u1}, x_{u2}, \dots, x_{uq})$$

$$\underline{v} = (v_1, v_2, \dots, v_p)$$

$$\underline{x} = \begin{bmatrix} (\partial r_1 / \partial P_1)_o & \dots & (\partial r_1 / \partial P_p)_o \\ \vdots & & \vdots \\ (\partial r_n / \partial P_1)_o & \dots & (\partial r_n / \partial P_p)_o \end{bmatrix}$$

ó sea una matriz n x p.

$$\underline{v} = \begin{bmatrix} \Delta v_1 \\ \Delta v_2 \\ \vdots \\ \Delta v_p \end{bmatrix}$$

$$\text{Si } (\underline{x}^T \underline{x}) = \underline{A} = \begin{bmatrix} \sum_{u=1}^n (\partial r_u / \partial P_1)_o (\partial r_u / \partial P_1)_o \dots \sum_{u=1}^n (\partial r_u / \partial P_p)_o (\partial r_u / \partial P_1)_o \\ \vdots \\ \sum_{u=1}^n (\partial r_u / \partial P_1)_o (\partial r_u / \partial P_p)_o \dots \sum_{u=1}^n (\partial r_u / \partial P_p)_o (\partial r_u / \partial P_p)_o \end{bmatrix}$$

$$\underline{z} = (\underline{x}^T \underline{E}) = \begin{bmatrix} \sum_{u=1}^n E_u (\partial r_u / \partial P_1)_o \\ \vdots \\ \sum_{u=1}^n E_u (\partial r_u / \partial P_p)_o \end{bmatrix}$$

La ecuación (2-6) escrita en notación de matrices es :

$$(\underline{x}^T \underline{x}) \underline{v} = (\underline{x}^T \underline{E}) \dots (2-7)$$

$$\delta \quad \underline{A} \underline{V} = \underline{Z} \quad \dots(2-8)$$

de modo que

$$\underline{V} = \underline{C} \underline{Z} \quad \dots(2-9)$$

donde $\underline{C} = (\underline{A})^{-1}$

Teniendo el cálculo del vector \underline{V} con la ecuación (2-9) un nuevo valor estimado para P_j es obtenido, repitiendo el cálculo con V_j , ó sea el valor estimado de P_j . Las derivadas en la matriz \underline{x} pueden evaluarse analíticamente ó numéricamente. La siguiente ecuación es utilizada para calcular V_j :

$$v_j^{(n+1)} = v_j^{(n)} + T_j^{(n)} \Delta v_j^{(n)} \quad \dots(2-10)$$

$$j = 1, 2, \dots, p$$

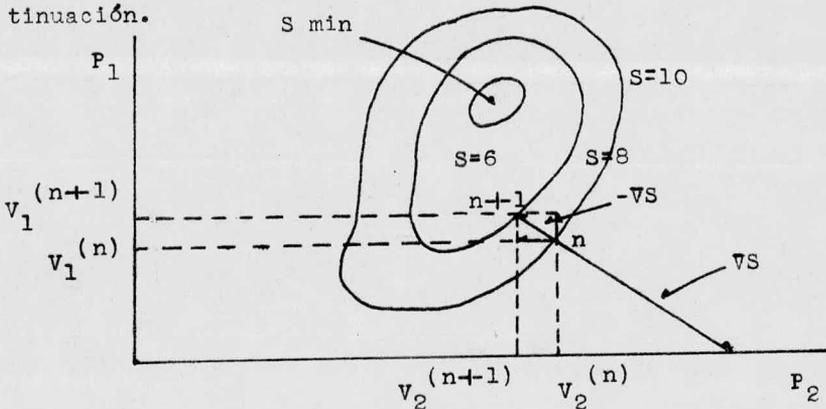
Siendo $T_j = 1$ para el método de Gauss, llamado "factor de -- aceleración", es decir, acelera la búsqueda del valor mínimo de $S(V)$; prácticamente Δv_j determina la dirección de tal búsqueda en el espacio paramétrico, mientras que la longitud de cada paso está dada por el factor T_j de aceleración.

De esta forma, vectores sucesivos \underline{V} son calculados hasta -- que cada Δv_j sea lo suficientemente pequeño y $S(V)$ se aproxime a cero también en cuyo caso la convergencia será alcanzada. Si no se llega a la convergencia, los valores de $\underline{v}^{(0)}$ (valores iniciales ó punto de arranque) son reemplazados por nuevos valores, repitiéndose el proceso.

2.1-4 Técnicas de Gradientes.

Existe otra forma de minimizar la suma de cuadrados de las desviaciones $S(V)$ y es linearizar la misma función objetivo. Esta técnica la lleven a cabo los métodos de Gradientes Simples -- (ref. 3) y Gradientes Conjugados (ref. 4).

El método de Gradientes Simples fué propuesto por R. Fletcher y W. J. D. Powell. El gradiente de $S(V)$, es decir, $\nabla S(V)$, es un vector perpendicular a la superficie $S(V)$ en el espacio paramétrico que se extiende en la dirección del máximo incremento en la función de mínimos cuadrados en un punto dado; el negativo de dicho gradiente se extiende en la dirección descendente (steepest descent). La interpretación geométrica de $S(V)$, $\nabla S(V)$ y $-\nabla S(V)$ en un espacio de dos parámetros P_1 y P_2 se muestra a continuación.



Los contornos de $S(V)$ constante van incrementando su valor a partir del mínimo $S(V)$. Expandiendo $S(V)$ en una serie truncada de Taylor:

$$S(V) \approx S(V)_0 + \sum_{j=1}^p \left(\frac{\partial S(V)}{\partial P_j} \right)_0 (P_j - V_j^{(0)})$$

Las magnitudes de los componentes de $-\nabla S(V)$ son:

$$-\nabla S(V) \Big|_{\underline{y}^{(0)}} = -(\partial S(V)/\partial P_1)_0 \partial P_1 - (\partial S(V)/\partial P_2)_0 \partial P_2 - \dots \\ - (\partial S(V)/\partial P_p)_0 \partial P_p \quad \dots(2-11)$$

que se evalúan en $\underline{y}^{(0)}$ corresponden a los términos de la expansión de primer orden de $S(V)$ en el espacio paramétrico, sirviendo dichos componentes para determinar la dirección de búsqueda en las técnicas de gradientes. Se ha supuesto que $S(V)$ es continua con un solo mínimo. Evaluando los componentes del vector $-\nabla S(V)$ puede desarrollarse una técnica iterativa de cálculos de la siguiente forma:

Se calculan analíticamente ó numéricamente los componentes de $-\nabla S(V)$ evaluándolos en $\underline{y}^{(0)}$. La dirección (steepest descent) es determinada para lo que se calcula el vector unitario en el punto de partida.

O sea

$$\frac{-\nabla S(V)}{\|-\nabla S(V)\|} = \frac{-\partial S(V)/\partial P_1 \partial P_1 - \partial S(V)/\partial P_2 \partial P_2 - \dots}{((-\partial S(V)/\partial P_1)^2 + (-\partial S(V)/\partial P_2)^2 + \dots)^{1/2}} \\ = \frac{-\partial S(V)/\partial P_u}{\left(\sum_{j=1}^n (\partial S(V)/\partial P_j)^2\right)^{1/2}} \quad \dots(2-12) \\ u = 1, 2, \dots, n$$

Donde los componentes del vector unitario evaluados en $\underline{y}^{(0)}$, establecen la dirección de búsqueda para el mínimo de $S(V)$. Cada nuevo ciclo $\underline{y}_j^{(n)}$ es calculado a partir de un ciclo previo de la ecuación (2-10) empezando con $\underline{y}^{(0)}$. La secuencia de cálculos iterativos termina cuando el valor $S(V)$ es menor que algún valor especificado según el criterio que se considere, donde la convergencia habrá sido alcanzada.

El método de gradientes conjugados selecciona de igual forma un punto de partida. La dirección (steepest descent) es determinada análogamente especificando las componentes del vector dirección (forma normalizada) en el punto de arranque:

$$\frac{-\nabla S(V)}{\|-\nabla S(V)\|} = \frac{-\partial S(V)/\partial P_1}{\left(\sum_{j=1}^n (\partial S(V)/\partial P_j)^2\right)^{1/2}} \quad \dots(2-13)$$

Una búsqueda en una dimensión es llevada a cabo a lo largo de la dirección de (steepest descent) utilizando la ecuación --- (2-10). Cuando un mínimo es obtenido a lo largo de esta dirección una nueva "dirección conjugada" es evaluada en el nuevo punto -- con los siguientes componentes:

$$\frac{-\nabla S(V)}{\|-\nabla S(V)\|} = \frac{-(\partial S(V)/\partial P_u)^{(n)} + B^{(n-1)} (-\nabla S(V)/\|-\nabla S(V)\|)_u^{(n-1)}}{\left(\sum_{j=1}^n \left(-(\partial S(V)/\partial P_j)^{(n)} + B^{(n-1)} (-\nabla S(V)/\|-\nabla S(V)\|)_j^{(n-1)}\right)^2\right)^{1/2}} \quad \dots(2-14)$$

donde

$$B^{(n-1)} = \frac{\sum_{u=1}^n \left((\partial S(V)/\partial P_u)^{(n)}\right)^2}{\sum_{u=1}^n \left((\partial S(V)/\partial P_u)^{(n-1)}\right)^2} \quad \dots(2-15)$$

Una búsqueda unidimensional es desarrollada en esta dirección. Cuando un mínimo es obtenido se efectúa un chequeo de convergencia completa; si la convergencia es alcanzada el proceso termina. En caso de no obtenerse convergencia, nuevas componentes vectoriales de dirección conjugada deben calcularse en el punto mínimo a partir de la búsqueda simple en una dimensión. El proceso continúa hasta obtener convergencia ó hasta que p + 1 direcciones hayan sido exploradas. En caso de que un ciclo de --

$p + 1$ direcciones se haya completado, un nuevo ciclo debe iniciarse y consistirá de una dirección (steepest descent) y " p " -- direcciones conjugadas. Lo escrito anteriormente fué desarrollado para gradientes conjugados por R. Fletcher y C. M. Reeves.

Para el estudio de los modelos cinéticos considerados, se evaluaron sus parámetros minimizando $S(V)$, con los algoritmos de Fletcher y Powell para Gradientes Simples y Fletcher y Reeves para Gradientes Conjugados, que como pudo observarse anteriormente requieren derivadas de la función objetiva de mínimos cuadrados con respecto a cada parámetro. Se emplearon en ambos casos derivadas analíticas.

2.1-5 Técnica de Marquardt.

Donald W. Marquardt ha desarrollado un método (ref. 5) de vecindad máxima, en que se relacionan mediante una interpolación óptima, la técnica de Gauss y la técnica de Gradientes. Esta interpolación está basada en la vecindad máxima en la que la serie truncada de Taylor da una representación adecuada del modelo no lineal, lo cual origina un grupo de ecuaciones algebraicas no lineales que deben resolverse. Considérese el modelo dado por la ecuación (2-1); nuevamente el problema será calcular aquellos valores estimados de los parámetros que minimicen la ecuación (2-3). Para tal propósito Marquardt postuló lo siguiente.

Sea $\lambda \geq 0$ arbitrariamente en la ecuación (2-8) una matriz diagonal se agrega a la matriz \underline{A} y \underline{V} satisface la ecuación siguiente:

$$(\underline{A} + \lambda \underline{I}) \underline{V} = \underline{Z} \quad \dots(2-16)$$

donde \underline{I} es la matriz de identidad. Entonces λ es agregada a cada término de la diagonal principal de la matriz \underline{A} . Puede verse que cuando λ se aproxima ó tiende a infinito, la técnica de Marquardt es idéntica al método de gradientes, es decir $\lambda \underline{I} \gg \underline{A}$ y \underline{V} es calculada como:

$$\underline{V} = (1/\lambda) \underline{Z} \quad (\text{para cualquier iteración}) \quad \dots(2-17)$$

Cuando $\lambda = 0$, la técnica se reduce al método de Gauss en cuyo caso \underline{V} se calculará como en la ecuación (2-9). Los valores intermedios de lambda representan una combinación de ambas técnicas; en general lambda decrece a medida que las iteraciones avanzan. Se trata de encontrar un valor pequeño de lambda, donde las condiciones sean de tal forma que el método no modificado de Gauss converga adecuadamente. Los valores grandes de lambda deben

ser utilizados solamente donde sea necesario satisfacer la condición de que $S(V)$ en el ciclo $R + 1$ deba ser menor que $S(V)$ en el ciclo R ésimo, ó sea

$$S(V)^{(R+1)} < S(V)^{(R)} \quad \dots(2-18)$$

Es esencial seleccionar $\lambda^{(R)}$ tal que satisfaga la relación anterior. Mediante alguna forma de tanteo y error podrá encontrarse un valor $\lambda^{(R)}$ que satisfaciendo dicha relación, - resultará en una rápida convergencia del algoritmo hacia los valores mínimos cuadrados de las desviaciones. Es deseable a cada iteración minimizar $S(V)$ en (aproximadamente) la máxima vecindad en la que la función linearizada, dé una representación adecuada de la función no lineal. Debe por lo tanto buscarse el utilizar un valor pequeño de $\lambda^{(R)}$ de acuerdo a la estrategia para seleccionar $\lambda^{(R)}$ esto es importante principalmente en los últimos pasos del procedimiento de convergencia, cuando los valores a tantear estén en la vecindad inmediata del mínimo, donde - los contornos de $S(V)$ son asintóticamente elípticos y la expansión lineal de la función necesita una buena aproximación en solamente una muy pequeña región. Se define la estrategia para seleccionar λ de la siguiente forma:

Sea $\gamma > 1$

Sea $\lambda^{(R-1)}$ que representa el valor λ de la iteración anterior.

Inicialmente sea $\lambda^{(0)} = 10^{-2}$

Debe calcularse $S(\lambda^{(R-1)})$ y $S(\lambda^{(R-1)}/\gamma)$

a) Si $S(\lambda^{(R-1)}/\gamma) \leq S^{(R)}$, sea $\lambda^{(R)} = \lambda^{(R-1)}/\gamma$

b) Si $S(\lambda^{(R-1)}/\gamma) > S^{(R)}$, y $S(\lambda^{(R-1)}) \leq S^{(R)}$, sea $\lambda^{(R)} = \lambda^{(R-1)}$

c) si $s(\lambda^{(R-1)}/\gamma) > s^{(R)}$, y $s(\lambda^{(R-1)}) > s^{(R)}$ deberá incrementarse lambda por medio de multiplicación sucesiva por γ hasta que para alguna W muy pequeña, $s(\lambda^{(R-1)}\gamma^W) \leq s^{(R)}$. Con esto último se considera $\lambda^{(R)} = \lambda^{(R-1)}\gamma^W$.

Es conveniente hacer notar que los componentes del vector unitario en la dirección óptima, en la técnica de Gradientes, -- pueden multiplicarse por el tamaño de paso $h^{(n)}$ con lo cual se obtiene:

$$\underline{v}^{(n)} = \frac{h^{(n)}(-\nabla S(v^{(n)}))}{\|-\nabla S(v^{(n)})\|}$$

$$= \begin{bmatrix} \frac{-\partial S(v^{(n)})}{\partial P_1} \\ \vdots \\ \frac{-\partial S(v^{(n)})}{\partial P_p} \end{bmatrix} \frac{h^{(n)}}{\|-\nabla S(v^{(n)})\|} \quad \dots(2-19)$$

Debido a que

$$\frac{\partial S(v^{(n)})}{\partial P_j} = -2 \sum_{u=1}^n (y_u - r_u(v^{(n)})) \frac{\partial r_u(v^{(n)})}{\partial P_j}$$

$$= -z_j^{(n)}$$

es el elemento negativo típico en la matriz \underline{z} , se tiene que

$$\underline{z}^{(n)} = \begin{bmatrix} \frac{-\partial S(v^{(n)})}{\partial P_1} \\ \vdots \\ \frac{-\partial S(v^{(n)})}{\partial P_p} \end{bmatrix} \quad \dots(2-20)$$

Y por consiguiente, para en el caso en que lambda tiende a infinito, se aproxima lo siguiente:

$$\lambda \| \nabla S(v^{(n)}) \| / h^{(n)} \quad \dots(2-21)$$

Se observa que (2-16) comprende ambos métodos (Gradientes y Gauss). Por medio de este algoritmo es posible obtener una vecindad factible y casi siempre dentro de un factor determinado por ν , la máxima vecindad en la que las series de Taylor dan una representación adecuada. La selección de ν es arbitraria, $\nu = 10$ se ha encontrado que es un buen valor. En general, la condición (c) mencionada anteriormente rara vez es encontrada, exceptuando el caso en que existen grandes correlaciones entre los valores estimados de los parámetros, lo cual origina valores grandes de λ .

Se utilizaron derivadas numéricas de la función objetiva de mínimos cuadrados con respecto a cada uno de los parámetros para llevar a cabo la minimización de dicha función y calcular los valores de los parámetros en el mínimo.

2.1-6 Técnica de Powell.

M.J.D. Powell (ref. 6) propone la minimización de la función objetivo de mínimos cuadrados sin requerir derivadas analíticas, ó sea, las derivadas se aproximan por diferencias numéricas. La técnica, que es iterativa, trata de modificar el método de Gauss para reducir las dificultades originadas al resolver un conjunto de ecuaciones lineales en cada iteración, por lo que se recurre a una inversión de matriz iterativa para matrices simétricas, la cual cambia solamente un renglón y una columna de la matriz \underline{A} (ver sección 2.1-3) en cada etapa del proceso.

El algoritmo propuesto es el siguiente. Se escoge un punto de partida (valores iniciales de los parámetros a estimar) y un conjunto de componentes vectoriales de dirección $\underline{d}_{u,j}$ paralelamente a los ejes de coordenadas:

$$\underline{d}_1 = (1, 0, 0, \dots, 0)$$

$$\underline{d}_2 = (0, 1, 0, \dots, 0)$$

⋮
⋮
⋮

$$\underline{d}_p = (0, 0, 0, \dots, 1)$$

donde $u = 1, 2, \dots, p$; $j = 1, 2, \dots, p$

Considerando las ecuaciones desarrolladas por la técnica de Gauss, es decir las ecuaciones (2-7):

$$(\underline{x}^T \underline{x}) \underline{y} = (\underline{x}^T \underline{E})$$

en donde las derivadas se evalúan por diferencias numéricas aproximadas. Estas ecuaciones deben resolverse para obtener \underline{y} y tal vector es seguidamente utilizado para calcular un nuevo vector dirección con los componentes siguientes:

$$d_u, \text{ nueva} = \frac{\Delta v_u}{\sqrt{\sum_{j=1}^p \Delta v_j^2}}$$

donde $u = 1, 2, \dots, p$ y Δv_u son los elementos de \underline{v} .

A continuación se realiza una búsqueda en una dimensión en la dirección $\underline{d}_{\text{nueva}}$, por medio de la relación:

$$v_u (\text{nueva}) = v_u (\text{anterior}) - s p_u (\text{nueva})$$

donde $u = 1, 2, \dots, p$; s es la distancia avanzada en la dirección $\underline{d}_{\text{nueva}}$. Una vez que el mínimo en una dimensión se ha obtenido, se lleva a cabo una prueba de convergencia total. En caso de llegar a la convergencia, el proceso es detenido, pero si no se llega a ella, uno de los vectores previos de dirección debe reemplazarse por un nuevo vector de dirección. El vector que se reemplaza es el que posea el índice correspondiente al máximo de productos: $|g_u \Delta v_u|$, donde $u = 1, 2, \dots, p$ y g_u son los elementos de la matriz $(\underline{x}^T \underline{E})$, descrita en la sección (2.1-3).

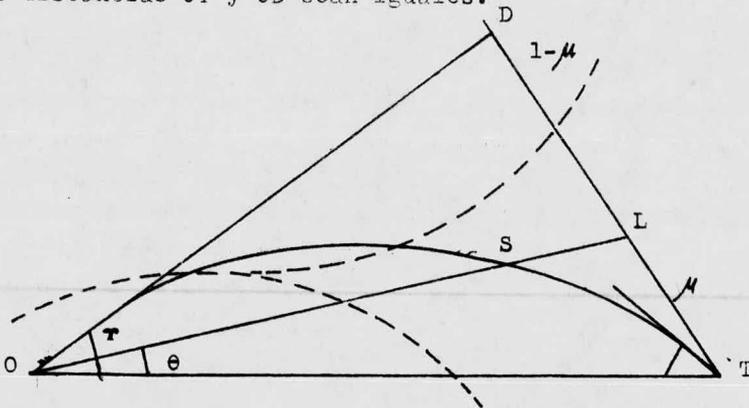
Los valores de las derivadas en la matriz \underline{x} para la nueva dirección son calculados por diferencias numéricas aproximadas, usando valores obtenidos de la búsqueda realizada en una dimensión. Las ecuaciones de Gauss son actualizadas entonces para la nueva dirección, es decir, un renglón y una columna en $(\underline{x}^T \underline{x})$ y un elemento en $(\underline{x}^T \underline{E})$, resolviéndose nuevamente para obtener el vector \underline{v} y repitiéndose todo el procedimiento hasta llegar a la convergencia.

2.1-7 Técnica Spiral.

La técnica Spiral desarrollada por A. Jones (ref. 7) busca aproximadamente la misma área que la técnica de Marquardt, pero los puntos buscados son generados por adición de vectores en lugar de inversión de matrices. Se asume en esta técnica que los pasos usados en el método de Gauss (es decir, un movimiento a un nuevo punto en el espacio paramétrico tan pronto como una suma reducida de cuadrados ha sido encontrada) son preferibles a la minimización lineal para la estimación de parámetros y que el esfuerzo requerido para calcular gradientes (analíticos ó numéricos es digno de atención. Considerese el modelo matemático representado por la ecuación (2-1).

El problema es nuevamente encontrar un valor estimado V de P , para el que la suma de cuadrados representada por la ecuación (2-3) sea un mínimo.

La idea básica del algoritmo Spiral es que una suma mínima de cuadrados puede siempre ser encontrada en el plano definido por el punto de la serie de Taylor y la línea direccional (steep est descent) en el punto base. En la figura siguiente, "O" es el punto base, T es el punto de las series de Taylor y OD es la dirección (steepest descent); el punto D es escogido de modo que las distancias OT y OD sean iguales.



Puesto que la suma de cuadrados debe decrecer inicialmente a lo largo de OD y ya que la aproximación de la serie de Taylor predice una suma reducida de cuadrados en el punto T, entonces es razonable suponer que valores mínimos de la suma de cuadrados podrán ser encontrados en alguna parte en el área OTD. La estrategia total indica que el punto base para la siguiente iteración debe estar tan lejos como sea posible del punto "O", pero que el número de evaluaciones de la superficie de mínimos cuadrados sea llevada a un mínimo.

Teniendo presente lo anterior, el siguiente punto base deberá ser tomado en el primer punto encontrado que muestre una reducción en la suma de cuadrados. Claramente el primer punto a ser investigado debe ser el punto de la serie de Taylor T. Si esto no es logrado entonces la validez de la aproximación lineal al modelo en "O" no se extiende tan lejos como T. Entonces el valle de suma de cuadrados debè curvarse en una de las dos direcciones mostradas como curvas punteadas en la figura. Puesto que la estrategia total tiende a dar puntos base en el trazo externo del valle es razonable asumir que el valle se está alejando de la línea OT. Para tantear e interceptar el valle, el espiral OTS es buscado; ésta curva se aleja de T en un ángulo beta dentro del área OTD y se mueve hacia "O" tangencialmente a OD. La ecuación mas adecuada para esta espiral, expresada en coordenadas polares, teniendo a "O" como origen, es:

$$r = r_0(1 - \theta \cos \beta - (1 - \gamma \cos \beta)(\theta/\gamma)^2) \quad \dots(2-22)$$

donde r es la distancia OS y r_0 es la distancia OT. La secuencia

de puntos S en la espiral que se investiga, es calculada a partir de una secuencia de puntos L, generadas sobre la línea TD -- tal que L divide a TD en la relación μ a $(1 - \mu)$; los valores sucesivos de μ son calculados de la relación auxiliar

$$\mu_{n+1} = 2\mu_n / (1 - \mu_n) \quad \dots(2-23)$$

que ha sido escogida para asegurar que los puntos se acerquen entre sí conforme se aproximen a D. Si las coordenadas de L son -- (E, θ) , entonces están dadas por las relaciones:

$$\tan \theta = \frac{\mu \text{sen } \gamma}{1 - \mu - \mu \text{sen } \gamma} \quad \dots(2-24)$$

$$E = \frac{r_0 \mu \text{sen } \gamma}{\text{sen } \theta} \quad \dots(2-25)$$

Entonces las coordenadas de S en el espacio paramétrico referido a "O" como origen, están dadas en términos de r, las coordenadas de T y d, las coordenadas de D por la relación siguiente

$$s = (r/E)(\mu d + (1 - \mu)t) \quad \dots(2-26)$$

donde d son incrementos en la dirección steepest descent y t son incrementos de la serie de Taylor. Como se observa en la figura, γ es el ángulo entre la serie de Taylor y la dirección steepest descent y s son incrementos de espiral.

Esta última ecuación es la base para el algoritmo Spiral; los puntos de búsqueda sucesivos en el algoritmo de Marquardt -- son generados por una inversión de matriz mientras que en ésta técnica son generados por una suma de dos vectores.

Puesto que la espiral es buscada solamente en puntos discretos es posible saltar de un extremo del valle al otro, sin encon

trar una suma de cuadrados mas pequeña que aquella en el punto--base. Para superar éste obstáculo, una interpolación puede efectuarse donde tres sumas de cuadrados consecutivas a lo largo de la espiral sean convexas hacia abajo, siendo considerada la suma de cuadrados como una función de μ . Puesto que pueden ocurrir mínimos locales aparentes, no hay un punto que refine esta interpolación y así, si la suma de cuadrados en el valor interpolado de μ no es menor que en el punto base, la búsqueda a lo largo de la espiral se continúa hasta que un valor máximo previamente especificado de μ haya sido alcanzado. En esta etapa el punto de serie de Taylor se reemplaza por el punto medio a lo largo de la línea OT y otra espiral es buscada a partir de este punto. Sin embargo si la suma de cuadrados en este punto medio es menor que aquella en la serie de Taylor original, es conveniente interpolar a lo largo de la dirección de serie de Taylor para considerar la posibilidad de que el valle de suma de cuadrados cruce la línea OT. Si el valor de la suma de cuadrados en este punto interpolado es menor que en el punto medio, un nuevo espiral es buscado a partir del punto interpolado.

La técnica permite normalmente, investigar cuatro espirales. Si un valor mínimo de la suma de cuadrados no ha sido alcanzado a lo largo de estos espirales, entonces la dirección de steepest descent debe buscarse. Esto ha ocurrido solamente cuando se presentan problemas para los que existen correlaciones muy altas entre los parámetros. Puede deberse también a la selección de malos valores de arranque, especialmente cuando se emplean derivadas numéricas.

Para la estimación de parámetros en los ejemplos cinéticos-considerados, se emplearon en este trabajo derivadas analíticas-

calculadas de la siguiente fórmula

$$\delta r_u / \delta P_j = \frac{r(P_j + e_j) - r(P_j)}{e_j} \quad \dots (2-26)$$

donde $e_j = \Delta P_j$ si $|\Delta P_j| > B$

y $e_j = \Delta$ si $|\Delta P_j| < B$

donde Δ , y B tienen los valores de 5×10^{-5} y 10^{-10} respectivamente.

El criterio de convergencia para el algoritmo Spiral está basado en el cambio relativo en los parámetros predicho por el punto de serie de Taylor. Específicamente la convergencia se alcanza cuando

$$\text{Max} \left| \frac{t_u}{v_u + \tau} \right| < \epsilon$$

donde r_u son los valores predichos por el modelo para la velocidad de reacción, v_u son los valores estimados de los parámetros y t_u son los correspondientes incrementos de serie de Taylor. -- Los valores de τ y ϵ se tomaron como 0.001 y 0.0001 respectivamente en la estimación de los parámetros.

2.2 Técnicas Probabilísticas.

Las técnicas Probabilísticas involucran funciones de probabilidad para describir el comportamiento de variables aleatorias discretas, funciones de densidad de probabilidad para variables aleatorias continuas, siendo de tales técnicas las mas usuales:

1. Técnica de Máxima Semejanza.

2. Técnica Bayesiana.

2.2-1 Técnica de Máxima Semejanza.

Suponiendo que un valor estimado suficiente existe, la técnica de Máxima Semejanza (Maximum Likelihood), propuesta primeramente por R. A. Fisher en 1912, conducirá a tal valor estimado; en esta técnica se iguala el valor estimado al valor del parámetro desconocido lo cual maximiza la probabilidad de los resultados de muestra observados. O sea da el valor estimado que es el más "semejante" para el parámetro desconocido. Se obtiene de esta forma valores estimados más eficientes que pueden ser sesgados, (ver capítulo 1).

Con objeto de definir la función de Máxima Semejanza (Maximum Likelihood), supóngase que $p(y; P_1, P_2, \dots, P_k)$ es una función de densidad de probabilidad de forma conocida para la variable aleatoria Y , es decir, una probabilidad condicional donde el modelo representa a Y para los valores de los parámetros. Tal función contiene uno ó mas parámetros P_1, P_2, \dots, P_k . Supóngase además que no se conocen cuales son los valores de P_1, P_2, \dots, P_k . La estimación puede hacerse agrupando aleatoriamente valores de las variables aleatorias Y , (y_1, y_2, \dots, y_n) y seleccionando los valores de P_1, P_2, \dots, P_k , ahora considerados como variables ---

aleatorias, que maximizan la función de Máxima Semejanza $L(P_1, P_2, \dots, P_k/y_1, y_2, \dots, y_n)$. Tales valores, V_1, V_2, \dots, V_k , se conocen como estimadores de máxima probabilidad. La técnica selecciona aquellos valores de P_1, P_2, \dots, P_k que son los más probables para dar origen a la muestra de valores observados. Al hacer una estimación de máxima probabilidad, se supone que la forma de la densidad de probabilidad está asociada con los valores buscados, donde todos los valores posibles de P_1 son igualmente probables antes de la experimentación.

La función de Máxima Semejanza para los parámetros, teniendo una observación, es justamente la densidad de probabilidad en la cual la observación se ha considerado como un número fijo y los parámetros como las variables, de modo que

$$L(P_1, P_2, \dots, P_k/y_1) = p(y_1; P_1, P_2, \dots, P_k) \quad \dots(2-28)$$

donde el subíndice "y" designa el valor de la observación respectiva que está involucrada en la función de densidad de probabilidad. La función de Máxima Semejanza para los parámetros, basándose en diversas observaciones, es el producto de las funciones individuales, si las funciones son independientes, ó sea:

$$p(y_1 \cap y_2) = p(y_1) p(y_2), \text{ con lo que}$$

$$L(P_1, P_2, \dots, P_k/y_1, y_2, \dots, y_n) = \prod_{u=1}^n L(P_1, P_2, \dots, P_k/y_u) =$$

$$p(y_1; P_1, P_2, \dots, P_k) p(y_2; P_1, P_2, \dots, P_k), \dots, p(y_n; P_1, P_2, \dots, P_k) \quad \dots(2-29)$$

Se seleccionarán como valores estimados de P los valores --

que maximizan la función de Máxima Semejanza para valores y_1, y_2, \dots, y_n en el tratamiento matemático para calcular los valores de P . Es mas conveniente trabajar tomando logaritmos a la ecuación (2-29):

$$\ln L = \ln p(y_1; P_1, P_2, \dots, P_k) + \ln p(y_2; P_1, P_2, \dots, P_k) + \dots, \ln p(y_n; P_1, P_2, \dots, P_k) = \sum_{u=1}^n \ln p(y_u; P_1, P_2, \dots, P_k) \quad \dots(2-30)$$

El logaritmo de la función de Máxima Semejanza puede maximizarse con respecto al vector \bar{P} , igualando a cero las derivadas parciales, con respecto a cada uno de los parámetros, ó sea:

$$\frac{\partial \ln L}{\partial P_1} = \frac{\partial \sum_{u=1}^n \ln p(y_u; P_1, P_2, \dots, P_k)}{\partial P_1} = 0$$

$$\frac{\partial \ln L}{\partial P_2} = \frac{\partial \sum_{u=1}^n \ln p(y_u; P_1, P_2, \dots, P_k)}{\partial P_2} = 0$$

⋮

$$\frac{\partial \ln L}{\partial P_k} = \frac{\partial \sum_{u=1}^n \ln p(y_u; P_1, P_2, \dots, P_k)}{\partial P_k} = 0 \quad \dots(2-31)$$

Resolviendo las ecuaciones anteriores se obtienen los valores estimados buscados (V_1, V_2, \dots, V_k) y los métodos de resolución son iterativos la mayoría de las veces.

Mientras mayor sea el número de observaciones, ó sea que tienda a infinito, el valor esperado del parámetro, tiende a ser

igual al valor real del mismo, ó sea es un estimador asintótico:

$$\lim_{n \rightarrow \infty} E(V_u) = P_u \quad \dots(2-32)$$

Se ha encontrado (ref. 2) que la función de densidad de probabilidad normal es la mas usada en la técnica de Máxima Semejanza:

$$p(y_u; P_1, P_2, \dots, P_k) = (1/\sigma_y \sqrt{2\pi}) \exp(-(1/2 \sigma_y^2)(y_u - r_u)^2) \quad \dots(2-33)$$

Sustituyendo esta función de densidad en la ecuación (2-29) se forma la función de Máxima Semejanza (Maximum Likelihood):

$$L(\sigma_y^2, P_1, P_2, \dots, P_k/y, x) = L = \prod_{u=1}^n (1/\sigma_y \sqrt{2\pi}) \exp((-1/2 \sigma_y^2)(y_u - r_u)^2) \quad \dots(2-34)$$

tomando logaritmos en la ecuación (2-34) se tiene:

$$\ln L = -n \ln \sqrt{2\pi} - (n/2) \ln \sigma_y^2 - \left(\sum_{u=1}^n (y_u - r_u)^2 / 2\sigma_y^2 \right) \quad \dots(2-35)$$

Esta función se maximiza con respecto al vector de parámetros, como se indicó anteriormente.

En el presente trabajo se ha resuelto el conjunto de ecuaciones, maximizando $\ln L$, por medio de diversos algoritmos, de los cuales se dará una breve descripción. Tales algoritmos son:

- a) Maximum Likelihood (Máxima Semejanza).
- b) Maximum-Simplex.
- c) Maximum-Gradientes Simples.

- d) Maximum-Gradients Conjugados.
- e) Maximum-Powell.
- f) Maximum-Rosenbrock.

El capítulo 4 compara los resultados obtenidos por los diferentes algoritmos mencionados, al evaluar parámetros desconocidos en diversos modelos cinéticos.

La desviación estandar de los valores predichos ó calculados por el modelo, de los valores observados experimentalmente, es considerada para todos los algoritmos como otro parámetro a estimar, para llegar así a un valor de desviación mas apropiado. El valor de partida se calculó con la expresión:

$$\sigma_y = \sqrt{\frac{n}{\sum_{u=1}^n (y_u - \bar{y})^2} (n - 1)} \quad \dots(2-36)$$

- a) Algoritmo **Maximum Likelihood** (Máxima Semejanza).

Este algoritmo consiste en la maximización de $\ln L$, simplemente derivándola con respecto a cada uno de los parámetros desconocidos en el modelo considerado, igualando a cero las derivadas y resolviendo el sistema de ecuaciones originado por el método de Newton Rapson.

- b) Algoritmo Maximum-Simplex.

La ecuación (2-35) en este caso es maximizada mediante el método Simplex, el cual no requiere derivación y utiliza patrones regulares de búsqueda, representados por una figura geométrica. Por ejemplo, si se tienen dos parámetros debe considerarse un triángulo equilátero, ó sea, tres puntos. Mientras que si son tres parámetros se tiene un tetraedro, es decir, cuatro puntos.

La técnica Simplex ha sido explicada en la sección 2.1-2.

c) Algoritmo Maximum-Gradientes Simples.

El logaritmo natural de la función de Máxima Semejanza se maximiza utilizando el algoritmo de Fletcher y Powell (ref.3) -- que involucra gradientes simples, descrito en la sección 2.1-4, y donde son evaluados los parámetros de los modelos cinéticos para un máximo de la función objetiva. En este caso se emplearon derivadas numéricas respecto a cada uno de los parámetros.

d) Algoritmo Maximum-Gradientes Conjugados.

La función de Máxima Semejanza en su forma logarítmica es maximizada mediante el algoritmo de Fletcher y Reeves (ref. 4) - que usa gradientes conjugados y que también se ha descrito en la sección 2.1-4.

Análogamente al algoritmo anterior se emplearon derivadas numéricas de la función objetiva.

e) Algoritmo Maximum-Powell.

Dicho algoritmo desarrollado por M. J. D. Powell (ref. 8) - fué empleado para la maximización de la función objetiva $\ln L$.

Este algoritmo de Powell encuentra el mínimo ó máximo de -- una función F no-lineal conteniendo P_1, P_2, \dots, P_k parámetros desconocidos sin utilizar derivadas. Básicamente consta de lo siguiente:

a) Se selecciona un punto de partida $\underline{P}_0^{(0)}$. Las direcciones iniciales de búsqueda $\underline{M}_i^{(0)}$ (donde $i = 1, 2, \dots, k$) son paralelas a los ejes coordenados.

b) Se realiza una serie de búsquedas para un solo parámetro en las k direcciones iniciales, por medio de aproximación cuadrática. Esto es:

- i) Se evalúa la función objetiva en el punto de partida.
- ii) Se incrementa el parámetro (que se está considerando como variable independiente) una distancia ΔP y se evalúa nuevamente la función objetiva. Si se obtiene un avance en la búsqueda del máximo, el tamaño de paso se duplica para la siguiente evaluación de la función, pero si no es así, debe regresarse en la dirección y localizarse el siguiente punto a una distancia $-\Delta P$ del punto de partida.
- iii) Después de lo anterior, el tamaño de paso se duplica cuando se avanza en la búsqueda del máximo y se reduce a la mitad si la evaluación de la función no mejora.
- iv) Una vez que se encuentra un máximo local, se obtienen los puntos: P_q, P_{q-1}, P_{q-2} (donde "q" indica la etapa correspondiente al máximo local) y se localiza un punto adicional:

$$P_{q+1} = P_{q-1} + \Delta P/2$$

reteniéndose los tres mejores puntos que pueden llamarse P_1, P_2 y P_3 .

- v) Una ecuación cuadrática f es ajustada a los tres puntos escogidos. Un nuevo óptimo local al igualar $\frac{\partial f}{\partial P} = 0$ es el siguiente:

$$P' = \frac{1}{2} \left[\frac{(P_2^2 - P_3^2)F(P_1) + (P_3^2 - P_1^2)F(P_2) + (P_1^2 - P_2^2)F(P_3)}{(P_2 - P_3)F(P_1) + (P_3 - P_1)F(P_2) + (P_1 - P_2)F(P_3)} \right]$$

c) Se obtienen entonces los puntos siguientes:

$\underline{P}_k^{(q)}$ = último punto obtenido de la serie de búsquedas para un solo parámetro.

$\underline{P}_t^{(q)}$ = punto expandido = $2\underline{P}_k^{(q)} - \underline{P}_o^{(q)}$; donde $\underline{P}_o^{(q)}$ es el punto de partida para la iteración.

aquí "q" indica la etapa (iteración) y aumenta para cada conjunto nuevo de dirección de búsqueda.

d) Se evalúa la función objetivo para ver si en el punto expandido $\underline{P}_t^{(q)}$, se obtiene un valor mejor que el del punto de partida - $\underline{P}_o^{(q)}$. Si no se mejora, el último punto $\underline{P}_k^{(q)}$ será el nuevo punto de partida y una nueva serie de búsquedas de un solo parámetro - debe hacerse en las mismas direcciones de antes, por lo tanto:

$$\underline{P}_o^{(q+1)} = \underline{P}_k^{(q)}$$

$$\underline{M}_i^{(q+1)} = \underline{M}_i^{(q)}$$

donde $i = 1, 2, \dots, k$.

Si la función objetivo $F_t^{(q)}$ en el punto expandido, es un avance en la búsqueda del máximo sobre la función en el punto de partida $F_o^{(q)}$, se realiza la siguiente prueba:

$$(F_o^{(q)} - 2F_k^{(q)} + F_t^{(q)})(F_o^{(q)} - F_k^{(q)} - \Delta)^2 \geq \frac{\Delta(F_o^{(q)} - F_t^{(q)})^2}{2}$$

siendo $\Delta = \left| F_M^{(q)} - F_{M-1}^{(q)} \right|$

Si se satisface, las direcciones antiguas de búsqueda son conservadas y una nueva serie de búsquedas de un solo parámetro-

principia como anteriormente se indicó. En caso de que la prueba no se satisfaga, una búsqueda de un solo parámetro se realiza en la dirección $\underline{N}^{(q)}$:

$$\underline{N}^{(q)} = \underline{P}_k^{(q)} - \underline{P}_0^{(q)}$$

que termina cuando el mejor valor $\underline{P}_0^{(q+1)}$ se ha hallado. Enseguida se escogen nuevas direcciones de búsqueda de la forma siguiente:

$$\underline{M}_i^{(q+1)} = \underline{M}_i^{(q)} \quad \text{para } i = 1, 2, \dots, M-1$$

$$\underline{M}_i^{(q+1)} = \underline{M}_{i-1}^{(q)} \quad \text{para } i = M, \dots, k-1$$

$$\underline{M}_k^{(q+1)} = \underline{N}^{(q)}$$

y principia una nueva serie de búsquedas de un solo parámetro.

e) Se llegará a la convergencia una vez que los valores para los parámetros entre dos iteraciones sucesivas sean menores que ciertos límites predispuestos, ó sea:

$$\left| P_i^{(q)} - P_i^{(q-1)} \right| < \epsilon$$

donde $i = 1, 2, \dots, k$.

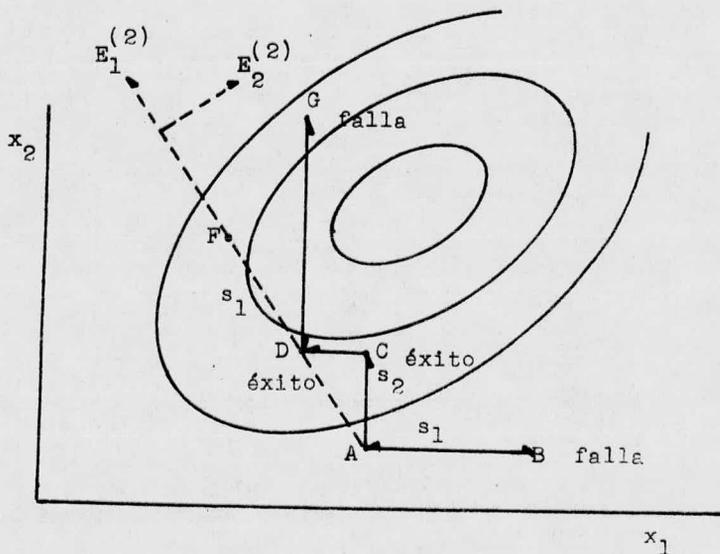
f) Algoritmo Maximum-Rosenbrock.

El algoritmo de H. H. Rosenbrock es utilizado para encontrar el máximo de la función $\ln L$, (ref. 9).

La técnica de búsqueda de Rosenbrock asume una función uni-

modal, ó sea, varios conjuntos de valores de partida para los pa-
rámetros a estimar, considerados como variables independientes, -
deberían ser usados, si es conocido que más de un mínimo existe -
ó si la forma de la superficie no es conocida.

Se procede de la siguiente forma. Un punto de partida debe-
escogerse así como un tamaño de paso para cada dirección de bús-
queda: se tienen entonces "p" tamaños de paso, s_1, s_2, \dots, s_p , -
que son conocidos. Inicialmente se escogen estas direcciones de
búsqueda que corresponden a los ejes coordenados. Una etapa for-
mada por una serie de ciclos puede entonces realizarse; princi-
pia en el punto de partida, por ejemplo, el punto A de la si-
guiente figura, al buscar un máximo se explora primero la direc-
ción $E_1^{(1)} = (1,0)$ en el punto B, a una distancia s_1 del punto de
partida A.



Búsqueda de Rosenbrock en dos dimensiones

Para esta figura $y(B) < y(A)$, de modo que el paso es fallido. En tal caso de falla, la distancia de movimiento s_1 es reemplazada para el ciclo siguiente por $-\beta s_1$, siendo $0 < \beta < 1$ sugiriendo Rosenbrock un valor de $\beta = 1/2$. La primera etapa continúa buscando en la dirección $E_2^{(1)}$ en el punto C, a una distancia s_2 de A. Si hay éxito se reemplaza s_2 por αs_2 en el siguiente ciclo de búsqueda, siendo $\alpha > 1$; Rosenbrock sugiere $\alpha = 3$. Se ensaya otra vez en la dirección $E_1^{(1)}$ obteniéndose el punto D, que está a una distancia s_1 del último punto exitoso (punto C).

Evaluando la función objetivo, se encuentra que $y(D) > y(C)$. Moviéndose nuevamente en la dirección $E_2^{(1)}$ y probando el punto G a una distancia αs_2 de D, $y(G) < y(D)$, ó sea hay falla en la -- prueba y D es retenida como el mejor punto en dicha etapa. Se obtuvo al menos un éxito y una falla en cada una de las direcciones $E_1^{(1)}$ y $E_2^{(1)}$. Enseguida Rosenbrock recomienda empezar una segunda etapa, cambiando la dirección de búsqueda a $E_1^{(2)}$ y $E_2^{(2)}$ de forma que después de varios cambios de dirección, la nueva E_1 -- tienda a apoyarse a lo largo de la dirección del avance mas rápido siendo E_2 perpendicular a tal dirección.

Un grupo de cálculos para un conjunto determinado de direcciones es una etapa y cuando ha habido al menos un éxito y una falla en cada dirección, termina la etapa, cambiándose enseguida las direcciones. Se definen las siguientes funciones para el grupo de direcciones de la segunda etapa:

$$A_1^{(1)} = d_1 E_1^{(1)} + d_2 E_2^{(1)} \quad \dots(2-37)$$

ó sea,

$$A_{i,1}^{(1)} = d_{1,i} m_{i,1}^{(1)} + d_{2,i} m_{i,2}^{(1)} \quad \dots(2-38)$$

para $i = 1, 2$

Donde d_1 y d_2 son las sumas algebraicas de los movimientos exitosos en las direcciones $E_1^{(1)}$ y $E_2^{(1)}$ respectivamente. Las $A_{i,j}^{(1)}$ están definidas en términos de las componentes direccionales y de las magnitudes de los movimientos exitosos de la primera etapa.

Entonces para el caso de la figura anterior:

$$d_1 = -\beta s_1 \quad (\text{movimiento con éxito de C a D en la dirección } E_1^{(1)})$$

$$d_2 = s_2 \quad (\text{movimiento con éxito de A a C en la dirección } E_2^{(1)})$$

$$\text{Si } E_1^{(1)} = (1,0) \text{ y } E_2^{(1)} = (0,1)$$

$$\text{se tiene } A_{1,1}^{(1)} = -\beta s_1 \text{ y } A_{2,1}^{(1)} = s_2$$

$$\text{es decir, } A_1^{(1)} = (-\beta s_1, s_2)$$

Las componentes direccionales necesarias para la segunda etapa de la búsqueda son:

$$M_{i,1}^{(2)} = \frac{A_{i,1}^{(1)}}{\sqrt{(A_{1,1}^{(1)})^2 + (A_{2,1}^{(1)})^2}} \quad \dots (2-39)$$

para $i = 1, 2$

las cuales definen la nueva dirección

$$E_1^{(2)} = (m_{1,1}^{(2)}, m_{2,1}^{(2)})$$

Para una segunda dirección de búsqueda, que es perpendicular a $E_1^{(2)}$, se define

$$A_{2,2}^{(1)} = d_2 E_2^{(1)} \quad \dots(2-40)$$

Para el caso de la figura,

$$A_{1,2}^{(1)} = 0 \quad \text{y} \quad A_{2,2}^{(1)} = s_2 \quad \text{o} \quad A_2^{(1)} = (0, s_2)$$

Se define la nueva dirección $E_2^{(2)}$ con la condición de que:

$$E_1^{(2)} \cdot E_2^{(2)} = 0$$

que es necesario para que las dos direcciones sean ortogonales.-

$E_1^{(2)}$ se ha normalizado de modo que:

$$E_1^{(2)} \cdot E_1^{(2)} = \sum_{i=1}^n m_{i,1}^{(2)} m_{i,1}^{(2)} = 1$$

Debe notarse que el vector $E_1^{(2)}$, que es una nueva dirección de búsqueda en la segunda etapa, corresponde a una línea que conecta el punto base en el principio de la primera etapa con el punto terminal en esa etapa, que es la suma vectorial de los movimientos exitosos.

Si $s_1 = s_2$, la ecuación (2-39) determinará la dirección $E_1^{(2)}$

$$m_{1,1}^{(2)} = \frac{-\beta s_1}{\sqrt{(\beta s_1)^2 + (s_2)^2}} = \frac{-\beta}{\sqrt{\beta^2 + 1}}$$

$$y \quad m_{2,1}^{(2)} = \frac{s_2}{\sqrt{(\beta s_1)^2 + (s_2)^2}} = \frac{1}{\sqrt{\beta^2 + 1}}$$

La dirección $E_2^{(2)}$ está dada a su vez por:

$$m_{1,2}^{(2)} = \frac{1}{\sqrt{\beta^2 + 1}}$$

$$y \quad m_{2,2}^{(2)} = \frac{\beta}{\sqrt{\beta^2 + 1}}$$

$E_2^{(2)}$ es perpendicular a la dirección $E_1^{(2)}$ ya que:

$$m_{1,1}^{(2)} m_{1,2}^{(2)} + m_{2,1}^{(2)} m_{2,2}^{(2)} = 0$$

ó vectorialmente:

$$E_1^{(2)} \cdot E_2^{(2)} = 0$$

Como se ilustra en la figura, un nuevo conjunto de direcciones de búsqueda en la segunda etapa se obtiene. Moviéndose -- del punto D en dirección $E_1^{(2)}$, una distancia s_1 al punto F, se -- obtiene una falla, puesto que la función objetiva disminuye. Nuevamente en D, se prueba en la dirección $E_2^{(2)}$, una distancia s_2 -- (no mostrada en la figura) y también falla. Estando aún en el -- punto D, se efectúa un movimiento a partir de él, una distancia -- βs_1 en la dirección $E_1^{(2)}$. Otra vez falla. Ahora, moviéndose -- una distancia $\beta^2 s_1$ en dirección $E_1^{(2)}$ se tiene que la prueba puede tener éxito, dependiendo del valor de β , ya que el movimiento es a lo largo de la línea que une D con F, una distancia frac

cional (β^2). Esto puede originar un nuevo punto base y un tanteo exitoso en la dirección $E_1^{(2)}$. El proceso de cálculo continuaría hasta que se obtuviese un tanteo exitoso en la dirección $E_2^{(2)}$. En tal caso, nuevas direcciones de búsqueda se escogerían nuevamente y principiaría la tercer etapa.

El final del cálculo lo dan un determinado número de movimientos ó la obtención de algún pequeño límite $d_1^2 + d_2^2$ previamente fijo; esto da el criterio de convergencia.

2.2-2 Técnica Bayesiana.

La técnica Bayesiana para la estimación de parámetros utilizando información previa. Este conocimiento previo puede provenir de consideraciones teóricas, de resultados de experimentos anteriores ó de suposiciones del experimentador. Normalmente la técnica Bayesiana supone previamente una distribución de probabilidad para un parámetro desconocido P en algún espacio paramétrico F ; esta distribución es actualizada por medio del teorema de Bayes para obtener la distribución de probabilidad posterior.

Para entender el teorema de Bayes, considérese un conjunto de eventos ó sucesos E_1, E_2, \dots, E_n , que son mutuamente exclusivos: ($E_i \cap E_j = \emptyset$; $i \neq j$) y $E_1 \cup E_2 \cup \dots \cup E_n = S$ entonces los E_1, E_2, \dots, E_n forman una partición de S (espacio muestra). Con lo anterior, para algún otro evento $A \subset S$, se tiene que la probabilidad de que el evento E_1 ocurra, dado que el evento A ya ha ocurrido, expresada por $p(E_1/A)$, es igual al producto de la probabilidad de que E_1 ocurra independientemente de que A tenga lugar, por la probabilidad de que A ocurra, dado que E_1 ha tenido lugar, dividido por la probabilidad de ocurrencia de A . O sea:

$$p(E_1/A) = \frac{p(A/E_1) p(E_1)}{\sum_{i=1}^n p(A/E_i) p(E_i)} \quad \dots(E-41)$$

Puesto que $p(E_1)$ es el grado de certidumbre de que el suceso E_1 ocurra, es llamado probabilidad previa. $p(E_1/A)$ es el grado de certidumbre de que el evento E_1 ocurra, dando como evidencia adicional el evento A , y se llama probabilidad posterior. ----- $p(A/E_1)$ indica la probabilidad de Máxima Semejanza (Maximum Likelihood) de que el evento A ocurra dado que el evento E_1 es verdadero. Esta es una probabilidad condicional, interpretada en la técnica Bayesiana como una probabilidad de Máxima Semejanza (Maximum Likelihood) ó sea, $L(E_1/A)$.

Una vez explicado el teorema de Bayes, se aplicará a la selección de modelos y estimación de parámetros involucrados, calculando las probabilidades requeridas en la técnica Bayesiana.

Sea H_u la hipótesis de que algún modelo j y algún conjunto de valores de parámetros V sean correctos. Usando el teorema de Bayes.

$$p(H_u/y) \propto p(H_u) p(y/H_u) \quad \dots(2-42)$$

siendo "y" el vector de observaciones $n \times 1$, y "n" el número de ensayos experimentales u observaciones numéricas disponibles.

En esta expresión, la cantidad de la izquierda es llamada la probabilidad posterior de H_u y es su probabilidad según los datos observados. El primer factor de la derecha es la probabilidad previa y se puede escribir como $p(M_j) p(V_j/M_j)$. Para establecer la distribución de probabilidad previa para los modelos, se considera la probabilidad del modelo j como $p(M_j)$, donde $j = 1,$

2, ..., n, siendo la $\sum_{j=1}^m p(M_j) = 1$.

El segundo factor de la derecha de la ecuación (2-42) es llamado factor de Máxima Semejanza (Maximum Likelihood), como ya se ha visto, quedando especificado por los datos y la hipótesis-Hu y tiene el valor numérico dado por $g_j(y/V_j)$, la función de densidad de probabilidad para "y" asumiendo el modelo "j". Para todas las posibles Hu se tiene lo siguiente:

$$p(Hu/y) = \frac{p(M_j) p(V_j/M_j) g_j(y/V_j)}{\sum_i \sum_{V_i} p(M_i) p(V_i/M_i) g_i(y/V_i)} \quad \dots(2-43)$$

donde $i = 1, 2, \dots, n$

La probabilidad posterior de un modelo es calculada actualizando la última ecuación para todas las posibilidades de parámetros para ese modelo.

$$p(M_j/y) = \frac{p(M_j) \sum_{V_j} p(V_j/M_j) g_j(y/V_j)}{\sum_i (p(M_i) \sum_{V_i} p(V_i/M_i) g_i(y/V_i))} \quad \dots(2-44)$$

La probabilidad posterior de un conjunto de valores de parámetros es hallada dividiendo el segundo término de la ecuación (2-43) entre el segundo término de la ecuación (2-44):

$$p(P/M, y) = \frac{p(V_j/M_j) g_j(y/V_j)}{\sum_{V_j} p(V_j/M_j) g_j(y/V_j)} \quad \dots(2-45)$$

En ambas ecuaciones el primer factor en el numerador, del lado derecho es una probabilidad previa y el segundo factor es de Máxima Semejanza (Maximum Likelihood).

Por consiguiente los valores numéricos de los parámetros -- con probabilidad más alta, serán los valores estimados más cercanos a los valores **reales** de los parámetros desconocidos.

Para variables continuas, el teorema de Bayes se expresa -- en términos de funciones de densidad de probabilidad. Así la --- ecuación (2-41) expresada en términos de un conjunto de valores-observados de la variable aleatoria "y" y de los parámetros desconocidos queda:

$$p_n(P/y_n) = \frac{L(P/y_n) p_0(P)}{\int_{-\infty}^{\infty} L(P/y_n) p_0(P) dP} \quad \dots(2-46)$$

Se ~~asume~~ generalmente una distribución normal para el vector P centrado en un vector de valores estimados V. Donde:

$p_n(P/y_n)$ = Densidad de probabilidad posterior para P después de obtener n observaciones.

$p_0(P)$ = Función de densidad de probabilidad previa para el -- vector P.

$L(P/y_n)$ = Función de densidad de probabilidad, que es la fun--- ción de Máxima Semejanza (Maximum Likelihood) del vec tor P dado los valores observados de "y".

El denominador de la ecuación (2-46) es un factor normali-- zante, es decir la integración de todos los términos posibles y- semejantes al numerador.

Una vez que se han realizado n observaciones, una función de Máxima Semejanza (Likelihood) puede escribirse:

$$L(P/y_n) = (1/2\pi)^{n/2} \sigma_y^{-n} \exp \left[- \frac{\sum_{u=1}^n (y_u - r_u(x, P))^2}{2 \sigma_y^2} \right] \dots(2-47)$$

Como ya se indicó en la sección 2.1-1, conforme a los trabajos de Draper y Hunter, se asume que información previa acerca del vector P está disponible, y la densidad inicial para tal vector está dada por:

$$P_0(P) = \frac{1}{(2\pi)^{n/2} |\Omega|^{1/2}} \exp \left[-1/2(\underline{P}-\underline{P}_0)^T \Omega^{-1} (\underline{P}-\underline{P}_0) \right] \dots(2-48)$$

Siendo P_0 el vector de valores estimados iniciales y Ω la matriz covariancia de dimensiones $k \times k$, para el vector P .

Sustituyendo las ecuaciones (2-47) y (2-48) en (2-46) se obtiene la función de densidad de probabilidad dada por el teorema de Bayes una vez que se han efectuado las determinaciones experimentales.

$$p_n(P/y_n) = F \exp \left[- \frac{\sum_{u=1}^n (y_u - r_u(P, x))^2}{2 \sigma_y^2} \right] \times \exp \left[-1/2(\underline{P}-\underline{P}_0)^T \Omega^{-1} (\underline{P}-\underline{P}_0) \right] \dots(2-49)$$

donde F representa el factor normalizante.

Sacando logaritmos en ambos lados de la ecuación:

$$\ln p_n(P/y_n) = \ln F - \left[\frac{\sum_{u=1}^n (y_u - r_u(P, x))^2}{2 \sigma_y^2} \right] - \left[1/2(\underline{P}-\underline{P}_0)^T \Omega^{-1} (\underline{P}-\underline{P}_0) \right] \dots(2-50)$$

$r_n(P, x)$ representa los diversos modelos cinéticos estudiados y los parámetros involucrados en los mismos, se estimaron maximizando la ecuación (2-50) con respecto a P para obtener así los valores estimados "y" con mayor probabilidad dadas las "n" observaciones experimentales. Fué necesario seleccionar, además de los valores iniciales de los parámetros, una matriz covariancia- Ω , dándole valores arbitrarios a los elementos de acuerdo a un arreglo numérico de la misma y se especificó también la desviación estandar σ_y como:

$$\sigma_y = \sqrt{\frac{\sum_{u=1}^n (y_u - \bar{y})^2}{n-1}} \quad \dots(2-51)$$

El factor normalizante F fué considerado como otro parámetro de valor real desconocido, puesto que se trata de una constante en la ecuación (2-50), por tanto se tomó un valor arbitrario como valor inicial y la densidad de probabilidad posterior $p_n(P/y_n)$ se optimizó también para dicho factor llegando a obtenerse, dependiendo del modelo considerado, factores del orden de 800000 a 1000000.

La función dada por la ecuación (2-50), se maximizó empleando los siguientes algoritmos:

- a) Bayes-Simplex.
- b) Bayes-Gradientes Simples.
- c) Bayes-Gradientes Conjugados.
- d) Bayes-Powell.
- e) Bayes-Rosenbrock.

Los cuales ya fueron descritos esencialmente en la sección 2.1-1, siendo desde luego el $\ln p_n(P, y_n)$ la función objetiva y procediendo análogamente.

CAPITULO 3

Modelos de velocidad de reacción.

Existen un mecanismo y una velocidad de reacción que sugieren información de como una entidad química se transforma en otra distinta.

Los diferentes procesos químicos individuales que ocurrirán al mismo tiempo o en forma consecutiva constituirán el mecanismo de reacción y finalmente darán como resultado una determinada reacción química. Es por lo tanto importante poseer una expresión para la velocidad de reacción lo más congruente posible con el sistema reaccionante que se tenga, lo cual implica haber seleccionado el modelo de velocidad de reacción adecuado.

Los valores de la velocidad de reacción pueden variar desde cero hasta cercanos al infinito, pero generalmente se presentan entre estos extremos. Los parámetros de operación tales como presión, temperatura, concentración, influirán de alguna forma en la velocidad de reacción, lo cual deberá conocerse, para efectuar predicciones correctas. Como base a los modelos para los sistemas reaccionantes, se considera como definición de la velocidad de reacción, la variación en moles de uno de los componentes con respecto al tiempo, para un diferencial del volumen del sistema reaccionante, cuando es homogéneo; para sistemas reaccionantes heterogéneos se utilizan por conveniencia unidades de moles/tiempo x masa ó área cuando hay una fase sólida presente. Tal variación será negativa si es concerniente a un reactivo y positiva si lo es a un producto. Se tiene entonces:

$$N = CV$$

...(3-1)

donde N es el número de moles de un componente de un sistema reaccionante, C es la concentración del mismo y V es el volumen.

La velocidad de reacción es:

$$r = \frac{1}{V} \frac{d(CV)}{dt} \quad \dots(3-2)$$

Si las condiciones de densidad en el sistema reaccionante -- permanecen constantes se tendrá para la reacción



la expresión de velocidad

$$r_A = -dC_A/dt$$

$$r_B = -dC_B/dt$$

$$r_C = dC_C/dt$$

Relacionándose tales variaciones de la forma siguiente:

$$-\frac{1}{a} \frac{dC_A}{dt} = -\frac{1}{b} \frac{dC_B}{dt} = \frac{1}{c} \frac{dC_C}{dt} \quad \dots(3-3)$$

y tomando en cuenta que $r \propto C_A^\alpha C_B^\beta$ se tiene que introduciendo la -- constante de proporcionalidad

$$r = -\frac{dC_A}{dt} = k C_A^\alpha C_B^\beta \quad (3-4)$$

donde

$$\gamma = \alpha + \beta$$

k = constante de velocidad de reacción

α = orden de la reacción con respecto a A

β = orden de la reacción con respecto a B

γ = orden global ó total de la reacción

Cuando es posible conocer el mecanismo de una reacción química se puede establecer una ecuación de velocidad adecuada, esto es, un modelo adecuado. Con este modelo es necesario estimar el valor de los parámetros que estén involucrados en él, para lo cual se compararán las diferentes técnicas de estimación mencionadas en el capítulo 2.

Los sistemas reaccionantes de acuerdo a las fases presentes se clasifican en sistemas homogéneos y sistemas heterogéneos. La constante de velocidad de reacción dependerá de variables tales como la temperatura, la concentración de los componentes y características del catalizador presente aún cuando el sistema se encuentre en fase homogénea. Se considera la variación de la constante de velocidad de reacción k con la temperatura según la ecuación de Arrhenius:

$$k = A e^{-E_a/RT} \quad \dots(3-5)$$

Siendo A el factor de frecuencia, T la temperatura absoluta, E_a la energía de activación, que puede ser el resultado de las energías de activación de los procesos elementales en una reacción química; R es la constante general del estado gaseoso.

Expresando la ecuación (3-5) en forma logarítmica:

$$\ln k = \ln A - E_a/RT \quad \dots(3-6)$$

que conduce a una relación lineal entre $1/T$ y k , pudiéndose ajustar fácilmente un conjunto de datos experimentales por medio de mínimos cuadrados, para obtener los valores de la energía de activación y el factor de frecuencia. La teoría de colisiones y la del estado de transición han aportado también expresiones para la evaluación de la constante de la velocidad de reacción.

Se describen a continuación los modelos más comunes en sistemas reaccionantes homogéneos y heterogéneos.

3.1 Modelos Homogéneos.

Para escoger un modelo adecuado a un sistema reaccionante homogéneo puede efectuarse generalmente una comparación entre datos obtenidos experimentalmente como son las diferentes concentraciones de los componentes de una mezcla reaccionante determinadas

al cabo de distintos intervalos de tiempo y las diferentes ecuaciones para los modelos de velocidad de reacción con que se cuenta, de tal forma que pueda determinarse que modelo es el más adecuado para la reacción en estudio.

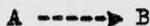
El establecer un modelo para la velocidad de reacción puede dificultarse aún cuando se tengan condiciones de temperatura constante y de composición uniforme, ó sea densidad constante. Dos métodos de comparación son: el Método Diferencial y el Método Integral. El primero requiere la diferenciación de los datos de concentración, obtenidos experimentalmente con respecto al tiempo, a partir de lo cual se obtendrán valores de velocidad de reacción que deben confrontarse con los valores de la misma predichos por el modelo.

En el segundo método se debe integrar la ecuación propuesta para la velocidad de reacción, de forma que se obtengan en función del tiempo, valores de concentración evaluados a partir del modelo.

Los modelos para sistemas homogéneos se describen brevemente a continuación.

3.1-1 Reacciones Irreversibles.

Para reacciones de orden cero:

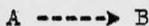


donde $-dC_A/dt = k_0$; integrando esta expresión se obtiene

$$C_A = C_{A0} - k_0 t \quad \dots(3-7)$$

con $C_A = C_{A0}$ y $t = 0$ inicialmente.

Para reacciones de primer orden:



en la cual $-dC_A/dt = kC_A$, se obtiene por integración:

$$\ln C_A/C_{A0} = -kt \quad \dots(3-8)$$

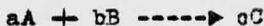
Para reacciones de segundo orden:



donde $-dC_A/dt = kC_A^2$; integrando se tiene:

$$1/C_A - 1/C_{A_0} = kt \quad \dots(3-9)$$

También de segundo orden se considera:



$$-dC_A/dt = kC_A C_B$$

Inicialmente:

$$C_A = C_{A_0}$$

$$C_B = C_{B_0}$$

$$t = 0$$

El balance de materia en cualquier tiempo es:

$$b(C_{A_0} - C_A) = a(C_{B_0} - C_B) \quad \dots(3-10)$$

por lo tanto

$$C_B = C_{B_0} - \frac{b}{a} (C_{A_0} - C_A)$$

definiendo la conversión en base al reactivo limitante:

$$x = \frac{C_{A_0} - C_A}{C_{A_0}} \quad \text{ó} \quad C_A = C_{A_0} (1-x) \quad \dots(3-11)$$

y sustituyendo C_A en la ecuación anterior

$$C_B = C_{B_0} - \frac{b}{a} C_{A_0} x$$

la velocidad de reacción es por lo tanto:

$$-dC_A/dt = C_{A_0} \frac{dx}{dt} = kC_A C_B$$

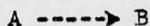
con las expresiones de C_A y C_B se tiene

$$-dC_A/dt = C_{A_0} \frac{dx}{dt} = k(C_{A_0}(1-x))(C_{B_0} - \frac{b}{a} C_{A_0} x)$$

la forma integrada será:

$$\ln \frac{C_{A_0}(1-x)}{C_{B_0} - \frac{b}{a} C_{A_0} x} = \ln C_A/C_B = \frac{(bC_{A_0} - aC_{B_0})}{a} kt + \ln C_{A_0}/C_{B_0} \quad \dots(3-12)$$

Para reacciones de tercer orden:



$$-dC_A/dt = kC_A^3$$

Con forma integrada:

$$1/2 \left(\frac{1}{C_A^2} - \frac{1}{C_{A0}^2} \right) = kt \quad \dots(3-13)$$

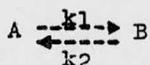
Es factible encontrar también reacciones de orden fraccionario. Asimismo, debe tenerse presente que el orden de una reacción es un número empírico obtenido mediante la experimentación, mientras que la molecularidad es un concepto que indica teóricamente, las especies que intervienen en tal sistema reaccionante, y ambos conceptos no deben forzosamente estar relacionados.

Los parámetros importantes en los modelos mencionados, son los órdenes de reacción con respecto a un componente ó totales, especialmente en la selección de modelos y la constante de velocidad de reacción, la cual puede evaluarse también como ya se indicó. La temperatura y la presión son variables controlables en el sistema, no así la concentración de los componentes (excepto cuando se propicia el exceso de alguno de los reactivos). El tiempo es una variable independiente.

La constante de velocidad y los órdenes de reacción respecto a cada uno de los reactivos son estimados en este trabajo utilizando las diversas técnicas de estimación ya descritas en el capítulo 2 (ver capítulo 4).

3.1-2 Reacciones Reversibles.

Considerando la reacción reversible de primer orden:



el modelo puede ser representado por

$$r = -dC_A/dt = k_1 C_A - k_2 C_B \quad \dots(3-14)$$

y para fines de integración deberá expresarse una de las variables en función de la otra, para lo cual se realiza un balance de materia:

$(C_{A_0} - C_A) = (C_B - C_{B_0})$; donde $t=0$, $C_A=C_{A_0}$ y $C_B=C_{B_0}$ en las condiciones iniciales. Por lo tanto:

$$C_B = C_{A_0} + C_{B_0} - C_A$$

sustituyendo en el modelo de velocidad

$$-dC_A/dt = (k_1 + k_2)C_A - k_2 C_{A_0} - C_{B_0} \quad \dots(3-15)$$

introduciendo la constante de equilibrio $K = k_1/k_2$, se tiene

$$\begin{aligned} -dC_A/dt &= (k_1 + \frac{k_1}{K})C_A - \frac{k_1}{K}(C_{A_0} + C_{B_0}) \\ &= k_1 \left[\frac{K+1}{K} C_A - \frac{1}{K}(C_{A_0} + C_{B_0}) \right] \quad \dots(3-16) \end{aligned}$$

además

$$K = C_{B_e}/C_{A_e} = \frac{C_{A_0} + C_{B_0} - C_{A_e}}{C_{A_e}}$$

$$C_{A_e}(K + 1) = C_{A_0} + C_{B_0}$$

sustituyendo en ecuación (3-16):

$$\begin{aligned} -dC_A/dt &= k_1 \left[\frac{K+1}{K} C_A - \frac{K+1}{K} C_{A_e} \right] \\ &= k_1 \frac{K+1}{K} (C_A - C_{A_e}) \end{aligned}$$

Llamando

$$C_A' = C_A - C_{Ae}$$

puede expresarse:

$$-dC_A/dt = k_R C_A' \quad \text{que corresponde a una forma de ecuación reversible de primer orden.}$$

...(3-17)

En donde $k_R = k_1 \left(\frac{K+1}{K} \right)$

Integrando (3-17):

$$\ln C_{A0}'/C_A = k_R t$$

$$\ln \frac{C_{A0} - C_{Ae}}{C_A - C_{Ae}} = k_R t; \quad \text{para reacciones de primer orden.}$$

...(3-18)

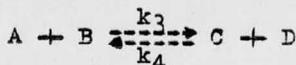
Otra forma integrada del modelo de velocidad de reacción puede obtenerse integrando

$$r = -dC_A/dt = (k_1+k_2)C_A - k_2(C_{A0}+C_{B0}) \quad \dots(3-19)$$

por lo tanto

$$\frac{1}{k_1+k_2} \ln \frac{(k_1+k_2)C_{A0} - k_2(C_{A0}+C_{B0})}{(k_1+k_2)C_A - k_2(C_{A0}+C_{B0})} = t \quad \dots(3-20)$$

Para reacciones reversibles de segundo orden:



el modelo puede ser representado por

$$r = -dC_A/dt = k_3 C_A C_B - k_4 C_C C_D \quad \dots(3-21)$$

Aplicando el concepto de conversión ya indicado:

$$r = -dC_A/dt = C_{A0} \frac{dx}{dt} = k_3 (C_{A0} - C_{A0}x)(C_{B0} - C_{A0}x) - k_4 (C_{C0} + C_{A0}x)(C_{D0} + C_{A0}x)$$

...(3-22)

Siendo

$K = k_3/k_4$; $k_4 = k_3/K$ y sustituyendo en la ecuación anterior, se tiene

$$C_{A_0} dx/dt = k_3((C_{A_0} - C_{A_0}x)(C_{B_0} - C_{A_0}x) - \frac{1}{K}(C_{C_0} + C_{A_0}x)(C_{D_0} + C_{A_0}x)) \quad \dots(3-23)$$

simplificando:

$$dx/dt = h + ix + jx^2$$

donde

$$h = k_3(C_{B_0} - \frac{1}{K} \frac{C_{C_0} C_{D_0}}{C_{A_0}})$$

$$i = -k_3((C_{A_0} + C_{B_0}) + \frac{1}{K}(C_{C_0} + C_{D_0}))$$

$$j = (k_3 - \frac{1}{K}) C_{A_0}$$

Tomando como condiciones iniciales

$$C_A = C_{A_0}$$

$$C_B = C_{B_0}$$

en $t = 0$

$$C_C = C_{C_0}$$

$$C_D = C_{D_0}$$

se obtiene la forma integrada del modelo de velocidad de reacción

$$\ln \frac{(2jx/i - (i^2 - 4hj)^{1/2}) + 1}{(2jx/i + (i^2 - 4hj)^{1/2}) + 1} = (i^2 - 4hj)^{1/2} t \quad \dots(3-24)$$

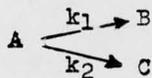
Los parámetros involucrados se han agrupado para la simplificación del modelo, siendo por lo tanto h, i, j, los parámetros -- importantes.

3.1-3 Sistemas Reaccionantes Complejos.

Los modelos para sistemas reaccionantes complejos pueden involucrar un alto grado de dificultad, ya que más de una sola reacción puede llevarse a cabo simultáneamente ó en varias etapas. Se enumerarán los casos mas importantes.

3.1-3a Reacciones en Paralelo.

Considérese el sistema siguiente que se realiza en paralelo ó simultáneamente.



Si las reacciones son de primer orden e irreversibles, el modelo se representa

$$r = -dC_A/dt = k_1 C_A + k_2 C_A = (k_1 + k_2) C_A \quad \dots(3-25)$$

Si $k_1 + k_2 = k_3$, entonces

$$-dC_A/dt = k_3 C_A \quad \dots(3-26)$$

La forma integrada es:

$$\ln C_A/C_{A_0} = -k_3 t$$

rearrreglando

$$\ln C_A - \ln C_{A_0} = -k_3 t ; \ln C_A = \ln C_{A_0} - k_3 t$$

$$C_A = C_{A_0} \exp(-k_3 t) \quad \dots(3-27)$$

además $dC_B/dt = k_1 C_A$

ó sea $dC_B/dt = k_1 C_{A_0} \exp(-k_3 t)$

integrando:

$$\int_{C_{B_0}}^{C_B} dC_B = \int_0^t k_1 C_{A_0} \exp(-k_3 t) dt$$

$$C_B - C_{B_0} = -(k_1/k_3) C_{A_0} \exp(-k_3 t)$$

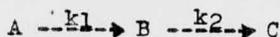
$$C_B = C_{B0} + (k_1/k_3)C_{A0} (1 - \exp(-k_3t)) \quad \dots(3-28)$$

Análogamente:

$$C_C = C_{C0} + (k_2/k_3)C_{A0} (1 - \exp(-k_3t)) \quad \dots(3-29)$$

3.1-3b Reacciones Consecutivas.

Los modelos para reacciones consecutivas ó en serie se establecen según la reacción



Si las concentraciones iniciales de los componentes A, B y C son igual a cero en el tiempo cero, los modelos de velocidad de reacción están representados por:

$$r = -dC_A/dt = k_1C_A \quad \dots(3-30)$$

$$dC_B/dt = k_1C_A - k_2C_B \quad \dots(3-31)$$

$$dC_C/dt = k_2C_B \quad \dots(3-32)$$

La expresión para C_A es

$$C_A = C_{A0} \exp(-k_1t) \quad \dots(3-33)$$

sustituyendo en la ecuación (3-31):

$$dC_B/dt = k_1C_{A0}(\exp(-k_1t)) - k_2C_B$$

$$dC_B/dt + k_2C_B = k_1C_{A0} \exp(-k_1t)$$

La ecuación integrada del modelo es:

$$C_B = \frac{k_1C_{A0}}{k_2-k_1} (\exp(-k_1t) - \exp(-k_2t)) \quad \dots(3-34)$$

sustituyendo en la ecuación (3-32) e integrando:

$$\int_{C_{C0}}^{C_C} dC_C = \int_0^t \left(k_2 \frac{k_1C_{A0}}{k_2-k_1} \right) (\exp(-k_1t) - \exp(-k_2t)) dt$$

$$C_C = C_{A0} \left(1 + \frac{1}{k_1-k_2} (k_2 \exp(-k_1t) - k_1 \exp(-k_2t)) \right) \quad \dots(3-35)$$

Habiendo mencionado los casos mas importantes en sistemas homogéneos, se describirán modelos para sistemas heterogéneos.

3.2 Modelos Heterogéneos.

Los sistemas reaccionantes heterogéneos involucran fronteras al estar presentes diferentes fases. Un catalizador podrá estar presente en una fase igual ó distinta a la de los reactivos, sin embargo en el caso mas común de un sistema heterogéneo se presenta un fluido (reactivos) y un catalizador en fase sólida, llevándose a cabo tres etapas principalmente: adsorción química de cuando menos uno de los reactivos del sistema reaccionante en el sólido, la reacción de dicho reactivo en la superficie y la desorción del producto de la superficie del catalizador sólido.

Una de las clases de modelos utilizados para el ajuste de datos cinéticos en sistemas heterogéneos es el modelo de potencias análogo al ya mencionado en sistemas homogéneos:

$$r = k_1 C_A^\alpha C_B^\beta \quad \dots(3-36)$$

Y si los reactivos están en fase gaseosa puede expresarse en función de las presiones parciales de los componentes:

$$r = k_1 P_A^\alpha P_B^\beta \quad \dots(3-37)$$

Los métodos que se usan generalmente para analizar una serie de datos con este modelo, se desarrollan primero para un solo componente reactivo, para el que la ecuación anterior puede escribirse:

$$r = k_1 P_A^\alpha \quad \dots(3-38)$$

La cual si se combina con la ecuación de continuidad para el reactor y se integra, puede aplicarse entonces el método integral ya descrito para comprobar que los órdenes de la reacción son los adecuados.

De otra manera, velocidades de reacción pueden ser obtenidas

experimentalmente, en cuyo caso podrá utilizarse el método diferencial para probar que tan adecuados son los órdenes de reacción asumidos.

En cualquier caso se trata generalmente de reorganizar el modelo en una forma lineal y de graficar los datos para observar -- que tanto difieren de una gráfica lineal. Algunos métodos han empleado tiempos de vida fraccionales, curvas dimensionales y transformaciones estadísticas. Estas y otras técnicas numéricas tales como mínimos cuadrados, se asemejan en que realizan una transformación algebraica de variables de modo que la forma de una superficie de respuesta puede interpretarse aproximadamente. No obstante, es preferible de ser posible, tratar al modelo en su forma original para que no sea tan grande el error en la estimación de los parámetros.

En el caso en que varios reactivos forman parte del sistema-reaccionante, ó sea aplicando como modelo la ecuación (3-37), un método de análisis sugiere la medición de datos de velocidades de reacción, teniendo presentes altas concentraciones de todos los reactivos, exceptuando a uno de ellos, de modo que dichas concentraciones permanezcan constantes prácticamente, mientras se desarrolla la reacción. De esta forma el orden del reactivo limitante puede ser determinado por uno de los métodos usuales para un solo componente. Tal método, conocido como de aislamiento, determina ciertos órdenes respecto a los reactivos, pero esta determinación sólo comprende una región limitada del espacio experimental y el modelo obtenido en estas condiciones deberá usarse cautelosamente, puesto que en las condiciones experimentales, las concentraciones correspondientes a los reactivos no variaron todas simultáneamente. Los parámetros que puedan ser estimados en el modelo, podrán quizá ser utilizados como valores iniciales aceptables, cuando se

use alguna de las técnicas de iteraciones no lineal, la cual pueda operar adecuadamente los datos que se obtengan, no obstante -- que estén variando las concentraciones de los componentes reactivos.

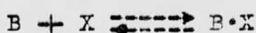
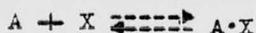
Asimismo, la ecuación (3-37) puede ser aplicada para el ajuste de datos experimentales junto con las velocidades iniciales de reacción, esto es, se varían las concentraciones iniciales de cada uno de los reactivos, mientras las restantes son mantenidas -- constantes; de tal forma los datos experimentales pueden analizarse para un solo componente reactivo y en comparación con el método anterior, las concentraciones pueden ser cercanamente semejantes sin algunas de ellas se encuentren en exceso. Se obtienen así velocidades de reacción en un rango entero de composiciones con -- varios componentes reactivos, pero podrán diferir de velocidades -- obtenidas experimentalmente, cuando se manifiesten conversiones -- grandes motivadas por la presencia de subproductos durante la --- reacción, que afectarán desde luego la velocidad de reacción.

Otro método más generalizado para tales estudios, se basa en los modelos de Hougen y Watson (Langmuir-Hinshelwood), en el cual deben analizarse los numerosos modelos posibles con objeto de poder descartar aquellos que sean los mas improbables; considérese -- al respecto lo siguiente.

Para una reacción cualquiera fluido-sólido:



puede asumirse un mecanismo:



(adsorción de dos componentes reactivos)



(reacción en la superficie)



(desorción del producto)

donde \bar{X} simboliza un centro activo del catalizador. Así, para el caso de que el paso controlante sea la reacción en la superficie, la adsorción y desorción se considera que tienen lugar casi en el equilibrio. La velocidad en la superficie es:

$$r_s = \frac{k_s}{C_x} (\bar{C}_A \bar{C}_B - \frac{1}{K_s} \bar{C}_C C_V) \quad \dots(3-39)$$

siendo \bar{C}_A , \bar{C}_B y \bar{C}_C las concentraciones de los componentes A, B y C respectivamente, adsorbidos; C_V es la concentración de centros activos que no tienen reactivo ó producto adsorbidos, K_s es la constante de equilibrio para la reacción en la superficie y C_x es la concentración total de centros activos:

$$C_x = C_V + \bar{C}_A + \bar{C}_B + \bar{C}_C \quad \dots(3-40)$$

Considerando el equilibrio para la adsorción y desorción, las concentraciones de equilibrio serán

$$\bar{C}_{Aeq} = K_A C_A C_V \quad \dots(3-41)$$

$$\bar{C}_{Beq} = K_B C_B C_V \quad \dots(3-42)$$

$$\bar{C}_{Ceq} = K_C C_C C_V \quad \dots(3-43)$$

donde C_A , C_B y C_C son las concentraciones de los componentes (en el fluido) A, B y C respectivamente en la superficie del catalizador, K_A y K_B son las constantes de equilibrio de adsorción de los reactivos y K_C es la constante de equilibrio de la desorción del producto C. Sustituyendo en ecuación (3-39) se obtiene la ecuación de velocidad en la superficie:

$$r_s = \frac{k_s}{C_x} (K_A K_B C_A C_B C_V^2 - \frac{K_C}{K_s} C_C C_V^2) \quad \dots(3-44)$$

Relacionando las ecuaciones (3-41), (3-42) y (3-43) con la ecuación (3-40) y despejando la concentración de centros activos sin componente adsorbido:

$$C_V = C_x / (1 + K_A C_A + K_B C_B + K_C C_C) \quad \dots(3-45)$$

de ecuaciones (3-44) y (3-45):

$$r = \frac{k_s C_x (K_A K_B C_A C_B - (K_C / K_s) C_C)}{(1 + K_A C_A + K_B C_B + K_C C_C)^2} \quad \dots(3-46)$$

Para la reacción total:

$$\text{constante de equilibrio} = K = C_C / (C_A C_B) \quad \dots(3-47)$$

sin embargo el término siguiente es el resultado de la combinación con las ecuaciones (3-41), (3-42) y (3-43):

$$K = \frac{\bar{C}_C / (K_C C_V)}{(\bar{C}_A / K_A C_V) (\bar{C}_B / K_B C_V)} = (K_A K_B / K_C) (C_V \bar{C}_{Ceq} \sqrt{\bar{C}_{Aeq} \bar{C}_{Beq}}) \quad \dots(3-48)$$

en donde el término $C_V \bar{C}_{Ceq} \sqrt{\bar{C}_{Aeq} \bar{C}_{Beq}}$ corresponde a la constante de equilibrio K_s , para la adsorción de dos reactivos. Por lo tanto

$$K = \frac{K_A K_B K_s}{K_C} \quad \dots(3-49)$$

sustituyendo en la ecuación del modelo (3-46):

$$r = k \frac{(C_A C_B - (C_C / K))}{(1 + K_A C_A + K_B C_B + K_C C_C)^2} \quad \dots(3-50)$$

donde k representa el producto $(k_s C_x K_A K_B)$.

La ecuación anterior representa el modelo de velocidad de reacción en función de las concentraciones en el fluido, en el caso en que el paso controlante es la reacción química en la superficie, con los dos reactivos A y B adsorbidos. Así los parámetros importantes que deben estimarse son k , K (que puede ser estimada termodinámicamente), las constantes de adsorción K_A , K_B y K_C de desorción.

Cuando se tiene un solo componente reactivo que se adsorbe, es decir, se tiene $A \rightleftharpoons X$, el denominador de la ecuación del modelo, queda elevado a la primera potencia y se reduce a la simplificación:

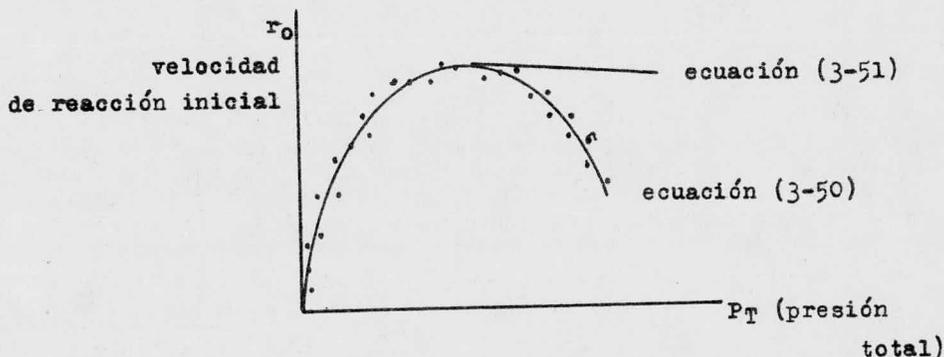
$$r = k' \frac{C_A}{K_C C_C}$$

donde $k' = k_s C_x K_A$ y la reacción es únicamente: $A \rightleftharpoons C$

Otro caso que puede presentarse es una adsorción muy débil - de todos los componentes reaccionantes, con lo que el término --- $(1 + K_A C_A + K_B C_B + K_C C_C)^2$ equivaldría a la unidad y la ecua--- ción del modelo de velocidad en tal situación se expresaría como--- para una reacción en fase homogénea

$$r = k(C_A C_B - \frac{1}{K} C_C)$$

Desde luego la presión parcial de cada uno de los componen--- tes puede sustituirse en vez de sus concentraciones respectivas,--- en fase gaseosa, como usualmente se registran los datos cinéticos a partir de los cuales la velocidad inicial de reacción se grafica como función de la presión total, como por ejemplo en la si--- guiente figura, donde las curvas de velocidad inicial sugieren --- que un modelo con dos componentes reactivos adsorbidos es el más--- probable para un conjunto de datos experimentales como el que se muestra:



donde la ecuación (3-51) sería el modelo para un sólo componente--- adsorbido:

$$r = \frac{k (C_A C_B - (1/K) C_C)}{(1 + K_A C_A + K_C C_C)} \dots (3-51)$$

y $k = k_s C_X K_A$

Cuando la etapa controlante es la adsorción de uno de los -- reactivos, la adsorción de los reactivos restantes, la reacción -- en la superficie del catalizador y la desorción de los productos-- se llevarán a cabo en equilibrio, en cuyo caso la expresión para el modelo de velocidad puede constituirse considerando la ecuación de adsorción del reactivo A, por ejemplo,

$$r_{ads} = k_{ads} (C_A C_V - (1/K_A) \bar{C}_A) \quad \dots(3-52)$$

donde la concentración adsorbida en el catalizador del reactivo A, se evalúa en base a la expresión

$$K_B = C_V \bar{C}_C / \bar{C}_A \bar{C}_B$$

puesto que se tienen las condiciones de equilibrio mencionadas. -- Asimismo estando en equilibrio la adsorción de B,

$$(\bar{C}_B)_{eq} = K_B C_B C_V \quad (\text{ecuación 3-42})$$

y la desorción del producto C, se tiene,

$$(\bar{C}_C)_{eq} = K_C C_C C_V \quad (\text{ecuación 3-43})$$

con las cuales se puede establecer la concentración adsorbida de A como:

$$\bar{C}_A = C_V \bar{C}_C / K_B \bar{C}_B = C_V K_C C_C C_V / K_B K_B C_B C_V$$

ó

$$\bar{C}_A = C_V K_C C_C / K_B K_B C_B$$

Sustituyendo el valor de K_B de la ecuación (3-49) en la expresión anterior:

$$\bar{C}_A = C_V K_C C_C K_A K_B / K_C K_B C_B$$

ó

$$\bar{C}_A = C_V C_C K_A / K_C B \quad \dots(3-53)$$

la cual puede introducirse en la ecuación de adsorción para el -- reactivo A, obteniéndose la ecuación de velocidad:

$$r = k_{ads} C_V (C_A - C_C/KC_B) \quad \dots(3-54)$$

que es el modelo cuando el paso controlante es la adsorción de A. A partir de los valores de equilibrio para \bar{C}_B y \bar{C}_C , dados por las ecuaciones (3-42) y (3-43) y \bar{C}_A por la ecuación (3-53) junto con la ecuación (3-40), se despeja de esta última el valor de C_V y se sustituye en la ecuación (3-54), dando como resultado la ecuación final del modelo para este caso:

$$r = k'' \frac{(C_A - (1/K)(C_C/C_B))}{1 + K_B C_B + (K_A/K)(C_C/C_B) + K_C C_C} \quad \dots(3-55)$$

donde $k'' = k_{ads} C_X$

De manera análoga, pueden presentarse infinidad de combinaciones entre mecanismos de reacción y pasos controlantes de los mismos, que originarán otros tantos modelos de ecuaciones de velocidad de reacción, cuyos parámetros es necesario calcular por algunas de las técnicas de estimación. En el capítulo 4 de este trabajo se estudian algunos modelos cinéticos, estimándose sus parámetros por las técnicas descritas en el capítulo 2.

CAPITULO 4

Comperación de técnicas de estimación de parámetros.

En este capítulo se exponen los resultados obtenidos mediante computación electrónica, al estimar los parámetros de algunos modelos cinéticos, aplicando las técnicas de Búsqueda Directa, Mínimos Cuadrados y Probabilísticas descritas en el capítulo 2.

Para tal objetivo se describen los ejemplos estudiados y se tabulan los resultados obtenidos por las diferentes técnicas, las cuales se enuncian enseguida.

Búsqueda Directa.

Técnica 1.- Algoritmo de Hooke-Jeeves, (H-J).

Mínimos Cuadrados.

Técnica 2.- Algoritmo de Gauss, (GAUSS).

Técnica 3.- Algoritmo de Gradientes Simples, (GS).

Técnica 4.- Algoritmo de Gradientes Conjugados, (GC).

Técnica 5.- Algoritmo de Marquardt, (MARQ).

Técnica 6.- Algoritmo de Powell, (POW).

Técnica 7.- Algoritmo Spiral, (SPI).

Probabilísticas.

Técnica 8.- Algoritmo de Maximum Likelihood, (MAXL).

Técnica 9.- Algoritmo Maximum-Simplex, (MAXS).

Técnica 10.- Algoritmo Maximum-Gradientes Simples, (MAXGS).

Técnica 11.- Algoritmo Maximum-Gradientes Conjugados, (MAXGC).

Técnica 12.- Algoritmo Maximum-Powell, (MAXPOW).

Técnica 13.- Algoritmo Maximum-Rosenbrock, (MAXROS).

Técnica 14.- Algoritmo Bayes-Simplex, (BAYS).

Técnica 15.- Algoritmo Bayes-Gradientes Simples, (BAYGS).

Técnica 16.- Algoritmo Bayes-Gradientes Conjugados, (BAYGC).

Técnica 17.- Algoritmo Bayes-Powell, (BAYPOW).

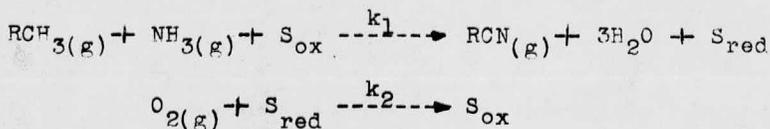
Técnica 18.- Algoritmo Bayes-Rosenbrock, (BAYROS).

EJEMPLO NUMERO 1.

Amonoxidación en fase vapor de 3-Picolina.

R. Prasad y A. K. Kar (ref. 10) estudiaron la cinética de oxidación y amonólisis de 3-Picolina en un reactor de flujo diferencial, utilizando un catalizador de pentóxido de vanadio-óxido de molibdeno-óxido de aluminio en el rango de 240-336°C.

En primer lugar se consideró la conversión de 3-Picolina a Nicotinonitrilo y el mecanismo asumido fué:



donde R es

Y siendo S_{ox} un centro activo de oxígeno adsorbido y S_{red} es el centro reducido de oxígeno ó el centro vacío. La ecuación de velocidad que representó los datos mas adecuadamente fué:

$$r_N = \frac{k_1 P_P P_A}{1 + \frac{3k_1 P_P P_A}{2k_2 P_O^{1/2}}}$$

Los datos experimentales utilizados en la estimación de los-
parámetros k_1 y k_2 son:

Presión parcial, atm			mol/hr g	
p_P	p_A	p_O	r_N	
0.004616	0.4072	0.2307	0.001201	
0.002056	0.4081	0.2313	0.000389	
0.006782	0.3002	0.2825	0.001291	T = 276°C
0.007339	0.3537	0.2431	0.001559	
0.008747	0.3002	0.2825	0.001562	

VALORES INICIALES DE LOS PARAMETROS
mol/hr g atm

SUMA RESIDUAL INICIAL DE CUADRADOS

$$k_1 = 0.676$$

$$k_2 = 0.091$$

$$5.3346817 \times 10^{-8}$$

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS		SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		k_1	k_2		
BUSQUEDA DIRECTA	H-J	0.676	0.091	5.3346817×10^{-8}	3.150
MINIMOS CUADRADOS	GAUSS	0.606487	846.075	2.242253×10^{-8}	8.403
	GS	0.640851503	0.07840861	$2.41800517 \times 10^{-8}$	4.372
	GC	0.60693865	41.514608	2.2420815×10^{-8}	3.692
	MARQ	0.58440525	100.00000	3.3561249×10^{-8}	4.548
	POW	0.61840039	0.1820000	$5.33468168 \times 10^{-8}$	7.752
	SPI	0.666793334	0.09099657	$4.14240000 \times 10^{-8}$	4.216
PROBABI- LISTICAS	MAXL	0.057838368	-0.00016755	$9.17326826 \times 10^{-6}$	3.095
	MAXS	0.9730	0.870172	3.28981×10^{-8}	3.524
	MAXGS	-----	-----	-----	---
	MAXGC	0.67599953	0.09099476	5.33434×10^{-8}	4.337
	MAXPOW	-----	-----	-----	---
	MAXROS	0.61065453	0.70558761	2.25345×10^{-8}	4.130
	BAYS	-----	-----	-----	---
	BAYGS	-----	-----	-----	---
	BAYGC	-----	-----	-----	---
	BAYPOW	-----	-----	-----	---
	BAYROS	0.64942682	0.04607812	3.07481×10^{-8}	4.845

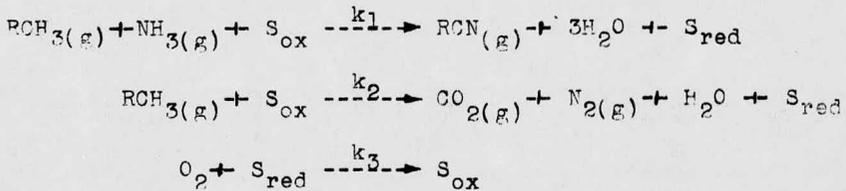
En este modelo, la técnica de Búsqueda Directa no produjo ningún avance en la minimización de la suma residual de cuadrados, mientras que de las técnicas de Mínimos Cuadrados, el método de Gradientes Conjugados llegó a un menor valor en 3.692 seg, el cual representa también el menor tiempo comparado con las otras técnicas de Mínimos Cuadrados. En las técnicas Probabilísticas, la técnica Maximum-Rosenbrock obtuvo la menor suma residual en 4.130 seg; se observan asimismo diversas técnicas que no llegan a la convergencia por exceso de tiempo de proceso ó porque la desviación estandar obtenida era negativa (recuérdese que es un parámetro a estimar en el modelo de las técnicas probabilísticas).

Comparando las dos mejores técnicas para este ejemplo, se considera que el método de Gradientes Conjugados es el más adecuado debido a que presenta la mínima suma residual de cuadrados y el menor tiempo de proceso.

EJEMPLO NUMERO 2.

Amonoxidación en fase vapor de 3-Picolina.

La segunda fase del experimento de R. Prasad y A. K. Kar -- (ref. 10) comprendió la conversión total de 3-Picolina.



donde R es

La ecuación de velocidad que representó satisfactoriamente - los datos fué:

$$r_T = \frac{k_1' p_P}{1 + \frac{k_1' p_P}{k_2' p_{\text{O}}^{1/2}}}$$

Los datos experimentales usados para la estimación de k_1' y k_2' son:

Presión parcial, atm		mol/hr g	
p_P	p_{O}	r_T	
0.003488	0.1459	0.002354	
0.005021	0.1457	0.003268	
0.006119	0.1456	0.003876	T = 336°C
0.008088	0.1453	0.004641	
0.006791	0.2825	0.004476	

VALORES INICIALES DE LOS PARAMETROS
mol/hr g atm

SUMA RESIDUAL INICIAL DE CUADRADOS

$$K_1 = 0.840$$

$$K_2 = 0.039$$

$$1.2826031 \times 10^{-8}$$

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS		SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		K_1	K_2		
BUSQUEDA DIRECTA	H-J	0.840	0.039	1.2826031×10^{-8}	3.294
MINIMOS CUADRADOS	GAUSS	0.819640	0.0413919	1.091343×10^{-8}	3.176
	GS	0.81963963	0.041892011	1.091341×10^{-8}	3.880
	GC	0.81901027	0.041834469	1.0913975×10^{-8}	3.945
	MARQ	0.81864092	0.041891780	1.091341×10^{-8}	4.605
	POW	0.81774536	0.041743523	$1.28260311 \times 10^{-8}$	8.081
	SPI	0.81862903	0.041893598	1.0913×10^{-8}	4.381
PROBABI- LISTICAS	MAXL	3.57600686	4.663637×10^{-9}	$7.28179631 \times 10^{-5}$	3.074
	MAXS	0.631865	0.83126	3.35114×10^{-8}	3.593
	MAXGS	-----	-----	-----	---
	MAXGC	0.83999715	0.038905734	1.27249×10^{-8}	4.013
	MAXPOW	0.84000000	0.039000000	1.28260×10^{-8}	5.205
	MAXROS	0.83606351	0.039215393	1.21247×10^{-8}	3.675
	BAYS	0.817389	0.0421094	5.69471×10^{-8}	3.397
	BAYGS	-----	-----	-----	---
	BAYGC	0.84045798	0.038255923	1.34539×10^{-8}	4.506
	BAYPOW	-----	-----	-----	---
	BAYROS	0.83583244	0.039136137	1.21942×10^{-8}	3.772

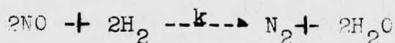
Para el ejemplo número 2, la técnica de Búsqueda Directa no mejoró la suma residual inicial de cuadrados. De las técnicas de Mínimos Cuadrados, las técnicas de Gradientes tanto Simples como Conjugados obtienen resultados aceptables para la suma residual, en el menor tiempo. En las técnicas Probabilísticas destacan los algoritmos de Maximum-Rosenbrock y Bayes-Rosenbrock que llegan a los mejores valores para la suma residual de cuadrados en 3.675 seg y 3.772 seg respectivamente. Las técnicas de Maximum-Gradientes Simples, Bayes-Gradientes Simples, y Bayes-Powell no alcanzan la convergencia debido a que en un punto dado en el cálculo, el logaritmo de la desviación estandar es - invalido.

Al comparar el método de Gradientes Simples con el de Maximum-Rosenbrock, se aprecia que el primero obtiene el mejor valor para la suma residual pero tarda 3.880 seg, mientras que el segundo aún cuando obtiene un valor mayor en la suma residual, - tarda 3.675 seg.

EJEMPLO NUMERO 3.

Reducción no catalítica del óxido nítrico.

Esta reacción (ref. 11) ha sido estudiada en la región de 1100°K pudiendo expresarse la reacción como:



El modelo de velocidad es:

$$r = k p_{\text{NO}}^{\alpha} p_{\text{H}_2}^{\beta}$$

Los datos experimentales son:

Presión parcial, mm Hg		mm Hg/seg	
p_{H_2}	p_{NO}	r_o	
400	359	1.5	
400	300	1.03	
400	152	0.25	$T = 826^\circ\text{C}$
289	400	1.6	
205	400	1.1	
147	400	0.79	

VALORES INICIALES DE LOS PARAMETROS

$k = 1.5 \times 10^{-5} = 1.30 = 1.20$

SUMA RESIDUAL INICIAL DE CUADRADOS

1.6639254×10^6

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS			SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		k				
BUSQUEDA DIRECTA	H-J	1.5×10^{-5}	1.30	0.64375	0.133041970	3.039
MINIMOS CUADRADOS	GAUSS	6.97933×10^{-9}	2.36246	0.889326	1.554647×10^{-2}	8.753
	GS	1.04607	-9.20631×10^{-5}	-8.87024×10^{-5}	1.21549644	4.447
	GC	3.19397×10^{-8}	1.80	1.20	1.9135381	3.802
	MARQ	6.97288×10^{-9}	2.362617	0.88934137	1.5546229×10^{-2}	4.375
	POW	3.46944×10^{-17}	1.80079015	1.19915279	1.663856×10^6	8.451
	SPI	-5.89378×10^{-6}	2.15704963	0.84953178	3.1272×10^5	4.028
PROBABILISTICAS	MAXL	-----	-----	-----	-----	---
	MAXS	-----	-----	-----	-----	---
	MAXGS	-----	-----	-----	-----	---
	MAXGC	-----	-----	-----	-----	---
	MAXPOW	-2.25	-0.450015	-1.050015	2.37180	4.826
	MAXROS	-----	-----	-----	-----	---
	BAYS	-----	-----	-----	-----	---
	BAYGS	-----	-----	-----	-----	---
	BAYGC	-----	-----	-----	-----	---
	BAYPOW	-----	-----	-----	-----	---
	BAYROS	-----	-----	-----	-----	---

En este caso, la técnica de Búsqueda Directa si redujo la suma residual inicial. Los algoritmos de Gauss y Marquardt, de Mínimos Cuadrados, reducen mas aún la suma residual inicial de cuadrados, aunque requieren mayor tiempo de proceso. De las técnicas Probabilísticas, únicamente la técnica de Maximum-Powell obtiene resultados, pero el valor obtenido de la constante cinética de velocidad es negativa; los algoritmos probabilísticos restantes presentan el problema ya mencionado de invalidar los logaritmos de los parámetros. Tampoco se alcanza la convergencia en su forma exponencial.

La técnica de Marquardt se ha seleccionado como la mejor para este modelo por presentar la mínima suma residual y el menor tiempo de proceso.

EJEMPLO NUMERO 4.

Reducción catalítica del óxido nítrico.

La reducción catalítica del óxido nítrico fué estudiada por R. J. Ayen y M. S. Peters (ref. 12), por medio de un reactor de flujo que se operó diferencialmente a presión atmosférica. Las presiones parciales del óxido nítrico y del hidrógeno se variaron de 0.005 a 0.05 atm y los datos se tomaron a 375°C, 400°C y 425°C, usando $ZnCr_2O_4-CuO-ZnO-Cr_2O_3$ como catalizador. La reacción es:



teniendo como modelo de velocidad de reacción:

$$r = \frac{k_1 p_{H_2} p_{NO}}{(1 + k_2 p_{NO} + k_3 p_{H_2})^2}$$

Se utilizaron los siguientes datos experimentales:

Presión parcial, atm		gmol/min g cat.
p_{H_2}	p_{NO}	r
0.00922	0.0500	0.0000160
0.0136	0.0500	0.0000256
0.0197	0.0500	0.0000327
0.0280	0.0500	0.0000364
0.0291	0.0500	0.0000348
0.0389	0.0500	0.0000446
0.0485	0.0500	0.0000475
0.0500	0.00918	0.0000147
0.0500	0.0184	0.0000248
0.0500	0.0298	0.0000345
0.0500	0.0378	0.0000406
0.0500	0.0491	0.0000475

VALORES INICIALES DE LOS PARAMETROS atm^{-1}

SUMA RESIDUAL INICIAL DE CUADRADOS

$K_1 = 0.13741$ $K_2 = 14.64$ $K_3 = 19.0$

2.97123×10^{-11}

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS			SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg.)
		K_1	K_2	K_3		
BUSQUEDA DIRECTA	H-J	0.13741	14.74	19.10	2.7507×10^{-11}	3.298
MINIMOS CUADRADOS	GAUSS	0.126620	13.1865	18.4871	2.615228×10^{-11}	7.697
	GS	0.137410805	14.640	19.0000	1.468805×10^{-8}	4.064
	GC	0.12820238	13.423642	18.582891	2.618227×10^{-11}	3.964
	MARQ	0.13741	14.64	19.0	$2.8712339 \times 10^{-11}$	4.663
	POW	0.126546979	13.2014106	18.5096823	$2.8712339 \times 10^{-11}$	8.268
	SPI	0.126633956	13.137506	18.4892045	2.6152×10^{-11}	3.832
PROBABILISTICAS	MAXL	0.003785495	-20.3713377	0.552221	1.9342061×10^0	2.979
	MAXS	0.139164	15.0957	19.1665	4.33334×10^{-8}	3.619
	MAXGS	4.0106×10^{-10}	1.164×10^{-10}	0.0	1.46445×10^{-8}	4.968
	MAXGC	0.13473057	14.640098	19.000035	2.88931×10^{-11}	4.636
	MAXPOW	2.3692796	17.530676	21.310659	$4.2362199 \times 10^{-06}$	5.079
	MAXROS	0.13734896	14.413433	19.535156	2.6504×10^{-11}	3.540
	BAYS	0.132573	13.7687	19.1948	3.23630×10^{-08}	3.555
	BAYGS	-----	-----	-----	-----	---
	BAYGC	0.13510771	14.64	18.939896	2.30964×10^{-11}	4.933
	BAYPOW	-----	-----	-----	-----	---
	BAYROS	0.13685818	14.275071	19.58033	2.64396×10^{-11}	4.309

La técnica de Búsqueda Directa, mejora la suma residual -- inicial de cuadrados en 3.998 seg. Para Mínimos Cuadrados se en cuenta que las técnicas de Gauss y Spiral alcanzan el menor va lor de suma residual, sin embargo, la técnica Spiral requiere - solamente 3.832 seg, mientras que la técnica de Gauss tarda --- 7.697 seg. Dentro de las técnicas Probabilísticas destacan los- algoritmos de Maximum-Rosenbrock y Bayes-Rosenbrock, obteniendo el primero la mayor suma residual de ambos pero en menos tiempo, mientras que el segundo obtiene la menor suma residual en un -- tiempo mayor.

Comparando los tres tipos de técnicas, puede observarse -- que el tiempo menor de proceso corresponde a la técnica de Bús- queda Directa, pero la Técnica Spiral converge a un valor menor de la suma residual en 3.832 seg.

EJEMPLO NUMERO 5.

Síntesis del cloruro de etilo.

Esta reacción fué analizada por George Thodos y L. F. Stutzman (ref. 13), con el objeto de establecer el mecanismo de reacción. Se utilizó metano como diluyente inerte; el catalizador fué $(ZrOCl_2 \cdot 8H_2O)$ en una solución diluida de HCl. Se utilizó el método diferencial, y se experimentó en la región de 350°F. - La reacción es:



El modelo de velocidad considerado es:

$$r_A = \frac{c(p_A p_B - p_R / K_P)}{(1 + K_A p_A + K_B p_B + K_R p_R + K_I p_I)^2}$$

donde K_P se tomó como 35.5 en base a los trabajos desarrollados por Rudkovskii, Trifel y Frost.

Los datos experimentales son:

r_A	Presión parcial, atm			
	p_I	p_A	p_B	p_R
0.000262	7.005	0.300	0.370	0.149
0.000260	7.090	0.416	0.215	0.102
0.000252	7.001	0.343	0.289	0.181
0.000216	9.889	0.511	0.489	0.334
0.000263	10.169	0.420	0.460	0.175
	CH ₄	C ₂ H ₄	HCl	C ₂ H ₅ Cl

VALORES INICIALES DE LOS PARAMETROS

g mol/min g-eqH⁺

$k_2 = 0.115$

$k_D = 0.743$

$K_M = 1.28$

$K_W = 13.25$

g mol/min g-eqH⁺

$k_3 = 1.0 \times 10^{-9}$

SUMA RESIDUAL INICIAL DE CUADRADOS

$5.55332084 \times 10^{-2}$

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS					SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		k_2	K_D	K_M	K_W	k_3		
BUSQUEDA DIRECTA	H-J	0.059	0.68699	1.336	13.25	1.0×10^{-9}	1.1181162×10^{-3}	3.171
MINIMOS CUADRADOS	GAUSS	-----	-----	-----	-----	-----	-----	---
	GS	-0.014905	-0.18508	-0.000038	-0.00002	0.151120	2.7483568×10^{-3}	4.628
	GC	0.682282	0.04444	1.247392	13.12125	0.036482	7.6170224×10^{-4}	4.195
	MARQ	0.442718	1.35652	8.019641	10.05877	0.155555	$3.8635889 \times 10^{-16}$	4.907
	POW	0.115	0.743	1.28	13.25	1.0×10^{-9}	5.5533208×10^{-2}	7.697
	SPI	0.114994	0.74299	1.28	13.25	-5.2×10^{-7}	5.6709×10^{-2}	4.809
PROBABILISTICAS	MAXL	-----	-----	-----	-----	-----	-----	---
	MAXS	0.297516	0.11180	0.979497	15.3965	0.030558	9.40995×10^{-8}	3.652
	MAXGS	-----	-----	-----	-----	-----	-----	---
	MAXGC	0.113695	0.74269	1.279765	13.24997	0.014380	1.46061×10^{-3}	4.115
	MAXPOW	0.115	0.743	1.28	13.25	1.0×10^{-9}	1.47328×10^{-3}	5.662
	MAXROS	0.115	0.743	1.28	13.25	1.0×10^{-9}	1.47328×10^{-3}	3.883
	BAYS	0.72858	0.04006	-0.200779	11.0196	-0.041021	5.76020×10^{-8}	3.332
	BAYGS	-----	-----	-----	-----	-----	-----	---
	BAYGC	0.11499	0.743	1.280021	13.25	-0.000257	1.47312×10^{-3}	4.335
	BAYPOW	-----	-----	-----	-----	-----	-----	---
	BAYROS	0.13031	0.24242	1.092011	4.48922	0.030118	7.41908×10^{-4}	4.433

En el presente ejemplo se encuentra que la técnica de Búsqueda Directa mejora la suma residual inicial en solamente 2.70%. El método de Gradientes Conjugados alcanza la menor suma residual para Mínimos Cuadrados en 4.467 seg. En las técnicas Probabilísticas, el algoritmo Bayes-Rosenbrock logra la menor suma residual en 4.733 seg.

Las técnicas que no convergen, se exceden en el tiempo de proceso, al calcular el valor de la variable dependiente por medio del modelo.

De las técnicas mencionadas, el algoritmo de Gradientes -- Conjugados alcanza el mejor valor para la suma residual de cuadrados, pero tarda 4.467 seg, es decir, mas tiempo de proceso -- que la técnica de Búsqueda Directa.

EJEMPLO NUMERO 6.

Formación de fosgeno.

La cinética de la formación catalítica de fosgeno (ref. 14) fué estudiada diferencialmente usando un catalizador de carbón - activado. Se hicieron mediciones de la velocidad de reacción en el rango de temperaturas de 31 a 99°C. El modelo de velocidad -- considerando la reacción en la superficie como paso controlante -- se expresa:

$$r = \frac{K_1 p_{CO} p_{Cl_2}}{(1 + K_2 p_{Cl_2} + K_3 p_{COCl_2})^2}$$

para la reacción:



Los datos experimentales reportados son:

r gr mol/hr gr cat.	Presión parcial, atm			T = 30.6°C
	p _{CO}	p _{Cl₂}	p _{COCl₂}	
0.00414	0.406	0.352	0.226	
0.00440	0.396	0.363	0.231	
0.00241	0.310	0.320	0.356	
0.00245	0.287	0.333	0.376	
0.00157	0.253	0.218	0.522	
0.00390	0.610	0.113	0.231	
0.00200	0.179	0.608	0.206	

VALORES INICIALES DE LOS PARAMETROS

SUMA RESIDUAL INICIAL DE CUADRADOS

$K_1 = 0.58 \quad K_2 = 2.70 \quad K_3 = 1.58$

4.93388×10^{-4}

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS			SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		K_1	K_2	K_3		
BUSQUEDA DIRECTA	H-J	0.18	2.95	1.98	$1.59293419 \times 10^{-7}$	3.570
MINIMOS CUADRADOS	GAUSS	0.16919	2.73129	1.82337	1.433633×10^{-7}	9.167
	GS	0.1634	2.76353215	1.63742646	$1.54030774 \times 10^{-7}$	5.996
	GC	0.1683173	2.7273783	1.8181057	1.4336954×10^{-7}	3.732
	MARQ	0.16919891	2.731372	1.8235067	1.4336629×10^{-7}	4.453
	POW	0.174	2.70674007	1.63938974	$4.93387538 \times 10^{-4}$	7.584
	SPI	0.169191346	2.73130474	1.823388617	1.4337×10^{-7}	4.695
PROBABILISTICAS	MAXL	-----	-----	-----	-----	---
	MAXS	-----	-----	-----	-----	---
	MAXGS	-----	-----	-----	-----	---
	MAXGC	0.58	2.70	1.58	4.93388×10^{-4}	4.460
	MAXPOW	0.83383042	2.9538304	1.8338304	9.19586×10^{-4}	5.436
	MAXROS	-----	-----	-----	-----	---
	BAYS	-----	-----	-----	-----	---
	BAYGS	-----	-----	-----	-----	---
	BAYGC	0.58	2.70	1.579999999	4.93388×10^{-4}	4.588
	BAYPOW	-----	-----	-----	-----	---
	BAYROS	-----	-----	-----	-----	---

Se llegó a un menor valor de la suma residual inicial de cuadrados por el método de Búsqueda Directa. En Mínimos Cuadrados sobresale la técnica de Marquardt que mejora el valor obtenido de la suma residual por el algoritmo de búsqueda Directa. Las técnicas Probabilísticas no fueron muy eficaces en este caso, puesto que únicamente tres de ellas alcanzan la convergencia, siendo la mejor la técnica Maximum-Gradientes Conjugados.

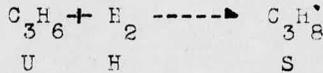
El resto de las técnicas Probabilísticas no convergen por manejar magnitudes muy grandes para la computadora.

Se ha seleccionado el método de Marquardt como el mejor para el presente modelo, que minimiza la suma residual al menor valor en 4.453 seg.

EJEMPLO NUMERO 7.

Hidrogenación de propileno.

M. V. Sussman y Charles Potter (ref. 15) investigaron la velocidad de hidrogenación del propileno con un catalizador Cu-MgO mediante un reactor de flujo diferencial. El rango de temperaturas fué de 48-76°C a una atmosfera de presión. La reacción que describe el proceso es:



Uno de los dos mecanismos posibles es el siguiente, asumiendo que el hidrógeno es adsorbido atómicamente, estando adsorbido también el propileno para la reacción en la superficie como paso controlante.

$$r = \frac{k_H K_U p_H p_U}{(1 + K_U p_U + K_S p_S)^3}$$

Datos experimentales:

Presión parcial, atm			gr mol/hr gr cat.	
p_H	p_U	p_S	r	
0.5442	0.3153	0.1405	0.03474	
0.4995	0.2960	0.2045	0.02800	
0.4741	0.2772	0.2487	0.02347	T = 76°C
0.6216	0.3607	0.0177	0.04520	
0.4423	0.2956	0.2621	0.02301	
0.3582	0.3427	0.2991	0.01631	

VALORES INICIALES DE LOS PARAMETROS

SUMA RESIDUAL INICIAL DE CUADRADOS

$k_H = 0.508$ $K_U = 3.03$ $K_S = 1.51$

2.83688×10^{-4}

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS			SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		k_H	K_U	K_S		
BUSQUEDA DIRECTA	H-J	0.608	2.83	1.31	$3.678777773 \times 10^{-6}$	3.045
MINIMOS CUADRADOS	GAUSS	0.610148	2.80926	1.24184	3.516023×10^{-6}	8.678
	GS	0.610149861	2.80927833	1.24183198	$3.51602491 \times 10^{-6}$	5.010
	GC	0.60625549	2.7742701	1.2290408	3.5163134×10^{-6}	4.294
	MARQ	0.61001848	2.8081112	1.2413993	3.5160258×10^{-6}	5.122
	POW	0.586393459	2.77552581	1.208	$2.83687665 \times 10^{-4}$	7.611
	SPI	0.610148704	2.80927207	1.24184496	3.5160×10^{-6}	4.228
PROBABILISTICAS	MAXL	0.017480130	2662.56577	0.0000098108	5.380220×10^{-3}	2.677
	MAXS	-----	-----	-----	-----	---
	MAXGS	-----	-----	-----	-----	---
	MAXGC	0.50782407	3.0299557	1.5098816	2.84289×10^{-4}	4.680
	MAXPOW	0.508	3.03	1.51	2.83688×10^{-4}	4.963
	MAXROS	-----	-----	-----	-----	---
	BAYS	0.628408	2.97814	1.29129	6.84121×10^{-8}	3.405
	BAYGS	-----	-----	-----	-----	---
	BAYGC	0.50830091	3.0299985	0.5103024	8.72586×10^{-5}	4.925
	BAYPOW	-----	-----	-----	-----	---
	BAYROS	0.59478505	2.6799786	1.1825368	3.54552×10^{-6}	4.974

Para este modelo se encontró que la técnica de Búsqueda -- Directa redujo la suma residual inicial de cuadrados. La técnica Spiral obtuvo la menor suma residual en el menor tiempo dentro de las técnicas de Mínimos Cuadrados. En las técnicas Probabilísticas destaca el algoritmo Bayes-Simplex que llega al menor valor de la suma residual en 3.405 seg. Las técnicas que no alcanzaron la convergencia se excedieron en el tiempo de proceso.

La técnica Bayes-Simplex es considerada como la mas efectiva en dicho modelo.

EJEMPLO NUMERO 8.

Hidrogenación de propileno.

En este ejemplo se evalúan los parámetros de otro posible mecanismo de reacción para la hidrogenación (ref. 18).

Se **asume** en este caso hidrógeno molecularmente adsorbido y propileno adsorbido, teniendo a la reacción en la superficie como paso controlante:

$$r = \frac{k_H K_U p_H p_U}{(1 + K_U p_U + K_S p_S)^2}$$

Los datos experimentales usados para la estimación de los parámetros son los mismos que para el ejemplo número 7.

VALORES INICIALES DE LOS PARAMETROS

SUMA RESIDUAL INICIAL DE CUADRADOS

$k_H = 0.543$ $k_U = 14.10$ $k_S = 7.20$

2.35954×10^{-5}

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS			SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg.)
		k_H	k_U	k_S		
BUSQUEDA DIRECTA	H-J	0.543	13.80	6.90	$1.66995187 \times 10^{-5}$	3.462
MINIMOS CUADRADOS	GAUSS	0.458328	10.475	4.59646	3.572498×10^{-6}	9.755
	GS	0.56441876	14.5424245	6.25559243	$3.64756212 \times 10^{-6}$	5.317
	GC	0.56454329	14.543644	6.2601825	3.6476905×10^{-6}	4.155
	MARQ	0.45739765	10.438696	4.5815636	3.5725127×10^{-6}	4.838
	POW	0.443942708	10.017371	4.3200000	$2.85954135 \times 10^{-5}$	7.344
	SPI	0.458443732	10.4797802	4.59836760	3.5725×10^{-6}	4.363
PROBABILISTICAS	MAXL	0.751563022	21.5141641	9.11790991	$3.85012157 \times 10^{-6}$	3.017
	MAXS	0.556355	14.2375	6.13103	6.51747×10^{-8}	3.980
	MAXGS	-----	-----	-----	-----	---
	MAXGC	0.54299857	14.100000	7.1999998	2.8597×10^{-5}	4.272
	MAXPOW	0.54300000	14.100000	7.2000000	2.85954×10^{-5}	5.276
	MAXROS	0.54300000	13.728699	5.9254086	3.62553×10^{-6}	3.754
	BAYS	0.62024	17.9779	6.02682	3.8284×10^{-8}	3.836
	BAYGS	-----	-----	-----	-----	---
	BAYGC	0.5330004	14.099795	5.5077566	6.07866×10^{-6}	4.329
	BAYPOW	-----	-----	-----	-----	---
	BAYROS	0.57103612	15.351099	5.8040039	6.49086×10^{-6}	4.103

Para el ejemplo número 8 se observa una minimización de --
la suma residual inicial de cuadrados mediante la técnica de --
Búsqueda Directa. Las técnicas de Gauss y Spiral de Mínimos Cua-
drados, alcanzan las menores sumas residuales, correspondiendo--
el menor tiempo de proceso a la técnica Spiral. Nuevamente so--
bresale en las técnicas Probabilísticas el algoritmo Bayes-Sim-
plex por su mínima suma residual en 3.836 seg de proceso.

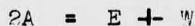
Únicamente se encontraron tres técnicas probabilísticas --
que no llegaron a la convergencia por excederse en el tiempo de
proceso.

Comparando las técnicas mencionadas se ha seleccionado a --
la técnica Bayes-Simplex como la mejor tanto en su suma resi---
dual como en su tiempo de proceso.

EJEMPLO NUMERO 9.

Deshidratación de iso-propanol.

La deshidratación del 2-propanol (ref. 16) fué estudiada -- usando resina de intercambio iónico como catalizador, en fase líquida y en un reactor de flujo, a las temperaturas de 90°, 100° y 110°C, a presión atmosférica. Considere la reacción:



donde A es el alcohol, E es el eter y W agua. Para adsorción de iso-propanol en dos centros activos, controlando la reacción en la superficie se estableció el modelo de velocidad de reacción siguiente:

$$r_1 = \frac{kK_A^2 p_A^2}{(1 + K_A p_A + K_W p_W + K_E p_E)^2}$$

Los datos experimentales reportados son:

Presión parcial, atm			mol/hr gr cat.	
p_A	p_W	p_E	$r \times 10^4$	
0.928	0.0125	0.0025	71.72	
0.601	0.0073	0.0014	71.01	
0.649	0.156	0.00065	34.66	T = 110°C
0.485	0.0062	0.389	77.20	
0.409	0.0025	0.0011	83.20	

VALORES INICIALES DE LOS PARAMETROS
 $k = 0.0093$ $K_A = 17.25$ $K_W = 34.50$ $K_R = 0.0001$
 mol/hr cat atm^{-1}

SUMA RESIDUAL INICIAL DE CUADRADOS

2.80585×10^{-6}

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS				SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		k	K_A	K_W	K_R		
BUSQUEDA DIRECTA	H-J	0.0094	17.2501	34.5001	0.0001	2.720960×10^{-6}	2.984
MINIMOS CUADRADOS	GAUSS	-----	-----	-----	-----	-----	---
	GS	0.00933574	17.2786873	34.6114444	-1.005899	2.4026413×10^{-6}	4.867
	GC	0.00933271	17.278061	34.612532	-0.984577	2.4022085×10^{-6}	3.714
	PARQ	0.00844014	46.774739	100.00	-1.5137232	1.1439532×10^{-6}	4.776
	POW	0.00378201	22.9713476	45.6677483	0.00054	2.8058463×10^{-6}	7.634
	SPI	0.00759846	113.178633	221.596693	-0.0531323	2.0066×10^{-6}	4.277
PROBABILISTICAS	MAXL	-----	-----	-----	-----	-----	---
	MAXS	0.00950096	17.2504	34.5022	0.000170	7.68601×10^{-8}	3.538
	MAXGS	-----	-----	-----	-----	-----	---
	MAXGC	0.00932922	17.25	34.50	-0.109507	3.30224×10^{-6}	4.132
	MAXPOW	0.0093	17.25	34.50	0.0001	2.80585×10^{-6}	5.307
	MAXROS	0.00891013	32.769571	85.752190	-0.167785	1.49439×10^{-6}	3.788
	BAYS	-----	-----	-----	-----	-----	---
	BAYGS	-----	-----	-----	-----	-----	---
	BAYGC	0.0093	17.25	34.50	0.0001	2.80585×10^{-6}	4.278
	BAYPOW	-----	-----	-----	-----	-----	---
	BAYROS	0.00666943	317.25	334.50	-74.7499	8.62051×10^{-6}	3.830

La técnica de Búsqueda Directa redujo la suma residual inicial de cuadrados en 2.984 seg de proceso, sin obtener el mejor valor de la misma. El algoritmo desarrollado por Marquardt obtuvo el menor valor para la suma residual en 4.776 seg. dentro de las técnicas de Mínimos Cuadrados. La técnica Maximum-Simplex - encontró la mínima suma residual en 3.588 seg, entre las técnicas Probabilísticas.

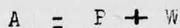
La falta de convergencia en varias técnicas no se analizó porque consumía mucho tiempo.

Como resultado de la comparación de las diferentes técnicas resultó seleccionada la técnica de Maximum-Simplex por haber obtenido la mínima suma residual de cuadrados, en la estimación de los parámetros del ejemplo número 9.

EjemPlo número 10.

Deshidratación del iso-propanol.

Esta reacción se estudió en las mismas condiciones que en el ejemplo número 9 (ref. 16), representándose por:



donde P es propileno.

El mecanismo en este caso considera a la reacción en la superficie como paso controlante, con el alcohol adsorbido en dos centros activos.

$$r_2 = \frac{kK_A p_A}{(1 + K_A p_A + K_W p_W + K_E p_E + K_P p_P)^2}$$

Los datos experimentales son los usados en el ejemplo número 9, pero tomando en cuenta también la presión parcial del propileno.

Presión parcial, atm				mol/hr gr cat.	
p_A	p_W	p_E	p_P	$r_2 \times 10^4$	
0.928	0.0125	0.0025	0.0016	23.49	
0.601	0.0073	0.0014	0.0014	26.70	
0.649	0.156	0.00065	0.00041	10.60	T = 110°C
0.485	0.0062	0.389	0.0016	23.58	
0.409	0.0025	0.0011	0.328	19.90	

VALORES INICIALES DE LOS PARAMETROS
 mol/hr g cat atm^{-1}

$k = 0.0145$ $K_A = 3.44$ $K_W = 20.95$ $K_D = 1.10$ $K_P = 1.50$

SUMA RESIDUAL INICIAL DE CUADRADOS

2.56923×10^{-7}

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS					SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		k	K_A	K_W	K_D	K_P		
BUSQUEDA DIRECTA	H-J	0.0145	3.54	21.05	1.10	1.60	7.1592659×10^{-7}	3.004
MINIMOS CUADRADOS	GAUSS	0.011627	2.25241	9.8839	0.42537	1.13934	1.815724×10^{-11}	3.348
	GS	0.013970	3.43123	20.9023	0.837746	1.97265	1.2884511×10^{-7}	5.151
	GC	0.014541	3.67671	20.7452	1.149582	2.22077	1.1224719×10^{-7}	4.287
	MARQ	0.011623	2.25377	9.8196	0.425558	1.13914	$2.6317307 \times 10^{-11}$	4.942
	POW	0.012006	2.49540	10.5552	0.550	1.17141	2.5692312×10^{-7}	7.451
	SPI	0.011627	2.25242	9.8840	0.425373	1.13934	3.1409×10^{-20}	4.478
PROBABILISTICAS	MAXL	-----	-----	-----	-----	-----	-----	---
	MAXS	0.013747	3.15314	20.5254	0.866774	1.9156	1.61598×10^{-8}	3.312
	MAXGS	-----	-----	-----	-----	-----	-----	---
	MAXGC	0.0145	3.44	20.95	1.10	1.50	2.56923×10^{-7}	4.815
	MAXPOW	0.0145	3.44	20.95	1.10	1.50	2.56923×10^{-7}	5.136
	MAXROS	0.014439	3.98607	15.4655	1.061269	2.34679	3.3318×10^{-8}	4.292
	BAYS	0.014161	3.86117	15.9453	0.853035	2.34092	7.3567×10^{-8}	6.035
	BAYGS	-----	-----	-----	-----	-----	-----	---
	BAYGC	0.014603	3.44	20.9497	1.098670	1.49599	2.75008×10^{-7}	4.571
	BAYPOW	-----	-----	-----	-----	-----	-----	---
	BAYROS	0.013965	3.64988	13.8088	1.019494	2.36038	3.47865×10^{-8}	3.730

Para este caso la técnica de Búsqueda Directa no mejoró -- la suma residual inicial de cuadrados. En las técnicas de Mínimos Cuadrados destacan dos valores para la suma residual, obtenidos por los algoritmos de Gauss y Spiral correspondiendo el mejor valor a esta última técnica. De las técnicas Probabilísticas el mejor valor para la suma residual fué obtenido por el algoritmo de Maximum-Simplex.

Las técnicas que no lograron la convergencia fué debido al tiempo de proceso excedido al estar iterando.

Comparando las dos mejores técnicas para este ejemplo, se considera que el método Spiral es el mas adecuado debido a que presenta la mínima suma residual de cuadrados en un tiempo de proceso de 4.478 seg.

EJEMPLO NUMERO 11.

Deshidratación del iso-propanol.

En este caso se considera nuevamente la reacción:



ya descrita en el ejemplo número 9 (ref. 16), siendo el paso con
trolante la desorción del éter isopropílico. El modelo de veloci
dad de reacción en este caso es:

$$r_1 = \frac{kK(p_A^2/p_W)}{1 + K_A p_A + kK_E(p_A^2/p_W)}$$

Los datos experimentales se encuentran ya tabulados en el -
ejemplo número 9.

VALORES INICIALES DE LOS PARAMETROS

SUMA RESIDUAL INICIAL DE CUADRADOS

mol/hr g oest.
k = 0.0093

K = 92.55

atm⁻¹
K_A = 17.25

K_E = 0.0001

7.94098 x 10¹

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS				SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		k	K	K _A	K _E		
BUSQUEDA DIRECTA	H-J	0.00009999	92.5406	17.2594	0.0095	6.835586 x 10 ⁻⁶	3.887
MINIMOS CUADRADOS	GAUSS	-----	-----	-----	-----	-----	---
	GS	0.12313755	92.5499888	17.2499999	5.26230654	7.03783718 x 10 ⁻⁶	4.628
	GC	0.03545596	94.460958	4.3767267	3.6780568	6.2655114 x 10 ⁻⁷	4.127
	MARQ	0.00390605	36.65949	100.0	0.48361958	6.228947 x 10 ⁻⁷	4.516
	POW	0.01803686	0.3759613	17.2548731	0.0002	7.94097823 x 10 ¹	6.502
	SPI	-0.1504859	-4511.054	-714.5236	-0.144831	2.1783 x 10 ⁻⁴	4.804
PROBABILISTICAS	MAXL	-----	-----	-----	-----	-----	---
	MAXS	-----	-----	-----	-----	-----	---
	MAXGS	-----	-----	-----	-----	-----	---
	MAXGC	0.00929999	92.55	17.25	0.00099987	7.94109 x 10 ¹	4.153
	MAXPOW	-0.020824508	89.475865	17.105543	2.8115325	1.01385 x 10 ⁻³	5.100
	MAXROS	0.00942204	36.053306	240.69802	1.166583	6.20112 x 10 ⁻⁷	4.020
	BAYS	0.00030674	89.7105	18.0297	0.0379692	5.18014 x 10 ⁻⁸	3.751
	BAYGS	-----	-----	-----	-----	-----	---
	BAYGC	0.0093	92.55	17.25	0.0001	7.94098 x 10 ¹	4.985
	BAYPOW	-----	-----	-----	-----	-----	---
	BAYROS	0.00935134	17.8	317.25	1.235336	7.9603 x 10 ⁻⁶	3.868

Se reduce la suma residual inicial de cuadrados mediante la técnica de Búsqueda Directa en 3.887 seg. De las técnicas de Mínimos Cuadrados destaca el algoritmo de Marquardt que obtiene un menor valor de la suma residual en 4.516 seg. El algoritmo Bayes-Simplex se seleccionó como el mejor de los métodos probabilísticos, llevándose un tiempo de proceso de 3.751 seg.

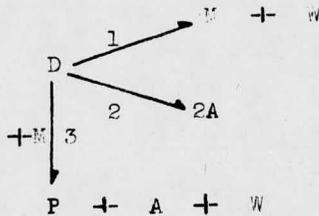
La técnica de Gauss no converge debido a que se excede en el tiempo de proceso, análogamente sucede con algunas técnicas Probabilísticas.

Al comparar las técnicas mas eficaces, para este ejemplo se selecciona a la técnica Bayes-Simplex, que obtuvo en el menor tiempo de proceso, la mínima suma residual.

EJEMPLO NUMERO 12.

Cinética de la conversión del alcohol diacetona a Óxido de mesitilo.

Este estudio fué realizado por R. E. Cunningham y N. O. -- Lercoff (ref. 17) utilizando resina de intercambio iónico como catalizador, estando el sistema en fase líquida y en la región de temperaturas de 30-40°C, en un reactor de flujo y a presión atmosférica. La reacción considerada es:



en la cual D representa al alcohol diacetona (4-Hidroxi-4-metil-2-pentanona), M al Óxido de mesitilo, W al agua, A al acetona, P la forona. Para la primera reacción se estudia el modelo de velocidad de formación del agua que corresponde a reacción en la superficie controlando en un doble centro activo:

$$r_W = \frac{k_1 K_D (y_D - y_M y_W / 0.5304)}{(1 + K_D y_D + K_M y_M + K_W y_W)^2} + k_3 y_D y_M$$

Los datos experimentales son:

fracción mol			gr mol/hr gr-eq	
y_D	y_M	y_W	r_W	
0.980	0.0103	0.0077	0.190	
0.506	0.0095	0.0087	0.177	
0.905	0.0096	0.083	0.119	$T = 40^\circ\text{C}$
0.891	0.066	0.0339	0.116	
0.303	0.688	0.0066	0.0376	

VALORES INICIALES DE LOS PARAMETROS

$g \text{ mol/min } g\text{-eq } H^+$

$k_1 = 0.895$

$K_D = 0.953$

$K_W = 7.86$

$K_M = 7.25$

$g \text{ mol/min } g\text{-eq } H^+$

$k_3 = 0.175$

SUMA RESIDUAL INICIAL DE CUADRADOS

4.14×10^{-3}

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS					SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		k_1	K_D	K_W	K_M	k_3		
BUSQUEDA DIRECTA	H-J	0.645	0.303	7.135	7.25	0.175	6.291523×10^{-3}	2.964
MINIMOS CUADRADOS	GAUSS	0.56407	1.95021	3.26674	51.79	1.08826	9.171065×10^{-14}	9.217
	GS	0.85126	1.25991	7.96406	7.325315	0.13196	7.260660×10^{-5}	4.803
	GC	0.86332	1.33267	8.580780	8.400814	0.14547	1.758759×10^{-5}	4.780
	MAXARQ	0.46934	1.35611	8.593931	9.994893	0.15569	2.382257×10^{-11}	4.636
	POW	0.4475	1.14097	8.223579	8.505093	0.15620	7.184821×10^{-2}	7.371
	SPI	0.38442	1.35567	9.937809	8.580515	0.15581	3.116×10^{-16}	4.413
PROBABILISTICAS	MAXL	-----	-----	-----	-----	-----	-----	---
	MAXS	-----	-----	-----	-----	-----	-----	---
	MAXGS	-----	-----	-----	-----	-----	-----	---
	MAXGC	0.89499	0.95999	7.86	7.25	0.17499	2.3762	4.438
	MAXPOW	-1.34487	-0.11507	5.694760	5.035293	-1.16011	1.11796	5.417
	MAXROS	0.90550	1.26475	9.945542	8.696354	0.15577	2.06579×10^{-5}	4.469
	BAYS	-----	-----	-----	-----	-----	-----	---
	BAYGS	-----	-----	-----	-----	-----	-----	---
	BAYGC	0.89499	0.9530	7.860135	7.249967	0.17337	2.37618	4.230
	BAYPOW	-----	-----	-----	-----	-----	-----	---
BAYROS	0.44330	1.25317	8.900483	8.350669	0.14805	1.52213×10^{-5}	4.455	

La suma residual inicial de cuadrados, en el ejemplo numero 12 no pudo mejorarse al utilizar la técnica de Búsqueda Directa. En contraste con lo anterior, el algoritmo Spiral obtiene una muy buena suma residual en un tiempo de 4.413 seg. Las técnicas Probabilísticas no mostraron gran avance en esta estimación de parámetros, no convergiendo varias de ellas por requerir demasiado tiempo en la evaluación de los parámetros, no obstante la técnica Bayes-Rosenbrock obtuvo una minimización de la suma residual inicial de cuadrados en 4.455 seg.

Como se observa la técnica Spiral llegó al mejor resultado en el menor tiempo de proceso.

EJEMPLO NUMERO 13.

Cinética de la conversión del alcohol diacetona a óxido de mesitilo.

También para la primera reacción del sistema reaccionante-descripto en el ejemplo número 12 (ref. 17), esto es:

$$D = M + W$$

se evalúan los parámetros del modelo correspondiente al caso en que el paso controlante es la reacción en la superficie, estando el reactivo adsorbido en un solo centro activo. Este modelo es:

$$r_W = \frac{k_1 K_D (y_D - y_M y_W / 0.5304)}{(1 + (K_D y_D)^{0.5} + K_M y_M + K_W y_W)^2} + k_3 y_D y_M$$

Los datos experimentales utilizados son los mismos que los del ejemplo número 12.

No se obtuvo un valor adecuado de la suma residual inicial al aplicarse la técnica de Búsqueda Directa en este modelo. En Mínimos Cuadrados destaca el algoritmo Spiral que redujo efectivamente la suma residual inicial en un tiempo de proceso de 4.705 seg. En las técnicas Probabilísticas el mejor valor para la suma residual fué alcanzado por la técnica Bayes-Simplex en 3.523 seg. Se presenta falta de convergencia por la técnica de Gradientes Simples al excederse en el tiempo de proceso empleado al estar evaluando los parámetros del modelo. Algunas técnicas Probabilísticas tampoco convergen debido a que se invalida el término potencial que se encuentra dentro del denominador del modelo.

La comparación de las mejores técnicas en este ejemplo da como resultado la selección de la técnica Spiral.

EJEMPLO NUMERO 14.

Cinética de la conversión del alcohol diacetona a óxido de mesitilo.

La segunda reacción del sistema descrito anteriormente (ref. 17) se representa por el siguiente modelo de velocidad de formación de la acetona en el que el paso controlante es la reacción en la superficie en un doble centro activo.

$$r_A = \frac{2k_2 K_D y_D}{(1 + K_D y_D + K_M y_M + K_W y_W)^2} + k_3 y_D y_M$$

Los datos experimentales son:

fracción mol			gr mol/hr gr-eq	
y_D	y_M	y_W	r_A	
0.980	0.0103	0.0077	0.0506	
0.506	0.0095	0.0087	----	
0.905	0.0096	0.083	0.0210	T = 40°C
0.891	0.066	0.0339	0.0338	
0.303	0.688	0.0066	0.0094	

VALORES INICIALES DE LOS PARAMETROS
atm⁻¹

SUMA RESIDUAL INICIAL DE CUADRADOS

C = 4.67 K_A = 27.42 K_B = 34.54 K_R = 36.76 K_T = 2.20

1.91599 x 10⁻⁹

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS					SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		C	K _A	K _B	K _R	K _T		
BUSQUEDA DIRECTA	H-J	4.57	27.52	34.64	36.86	2.20	1.387378 x 10 ⁻⁹	2.702
MINIMOS CUADRADOS	GAUSS	---	---	---	---	---	-----	---
	GS	4.61874	27.42375	34.5436	36.76217	2.27431	1.3940836 x 10 ⁻⁹	4.711
	GC	20.08447	118.7968	126.9840	81.14913	-0.58417	5.2537362 x 10 ⁻¹⁰	4.467
	MARQ	12.35824	93.40623	100.0000	63.75721	-0.51763	5.4550503 x 10 ⁻¹⁰	4.660
	POW	7.005	35.99548	44.5188	45.66198	2.64967	1.8133464 x 10 ⁻⁹	8.472
	SPI	4.67000	27.41999	34.5399	33.75999	2.19999	1.816 x 10 ⁻⁹	4.475
PROBABILISTICAS	MAXL	---	---	---	---	---	-----	---
	MAXS	4.67	27.42	34.54	36.76	2.20	1.81599 x 10 ⁻⁹	3.658
	MAXGS	---	---	---	---	---	-----	---
	MAXGC	4.57	27.42	34.54	36.76	2.20	1.81599 x 10 ⁻⁹	4.352
	MAXPOW	4.67	27.42	34.54	36.76	2.20	1.81599 x 10 ⁻⁹	5.501
	MAXROS	---	---	---	---	---	-----	---
	BAYS	27304.5	1245490.	357202.	521344.0	-40571.9	332502 x 10 ³	3.925
	BAYGS	---	---	---	---	---	-----	---
	BAYGC	4.66999	27.42	34.5404	34.75974	2.19243	2.00383 x 10 ⁻⁷	4.638
	BAYPOW	---	---	---	---	---	-----	---
	BAYROS	4.57733	27.24107	35.0803	47.25023	2.02232	9.91685 x 10 ⁻¹⁰	4.733

En el presente caso la técnica de Búsqueda Directa si redujo la suma residual inicial; sin embargo el algoritmo de Marquardt obtiene un mejor valor en 4.907 seg, dentro de los métodos de Mínimos Cuadrados, aunque la técnica de Gauss no llega a la convergencia por exceso de tiempo de proceso. El mejor resultado para la suma residual en los métodos Probabilísticos, se obtuvo al aplicar la técnica Bayes-Simplex para la estimación de los parámetros, en 3.332 seg.

La técnica elegida en este modelo es la de Marquardt por obtener el mejor valor para la suma residual.

EJEMPLO NUMERO 15.

Cinética de la conversión del alcohol diacetona a óxido de mesitilo.

Otro modelo cinético para la segunda reacción del sistema en estudio descrito en el ejemplo número 12 (ref. 17), supone que la reacción en la superficie controla y que el reactivo se ha adsorbido en un centro activo. Dicho modelo es el siguiente:

$$r_A = \frac{2k_2 K_D y_D}{(1 + (K_D y_D)^{0.5} + K_M y_M + K_W y_W)^2} + k_3 y_D y_M$$

Para la estimación de los parámetros se usaron los datos tabulados en el ejemplo número 14.

VALORES INICIALES DE LOS PARAMETROS

SUMA RESIDUAL INICIAL DE CUADRADOS

$k_2 = 0.167$ $K_D = 0.275$ $K_M = 0.554$ $K_W = 9.26$ $k_3 = 1.0 \times 10^{-9}$

$8.29578666 \times 10^{-4}$

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS					SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		k_2	K_D	K_M	K_W	k_3		
BUSQUEDA DIRECTA	H-J	0.1673	0.2753	0.554	9.2597	1.0×10^{-9}	8.431475×10^{-4}	3.863
MINIMOS CUADRADOS	GAUSS	-----	-----	-----	-----	-----	-----	---
	GS	-----	-----	-----	-----	-----	-----	---
	GC	0.173871	0.295595	0.55279	9.259926	0.003622	7.802099×10^{-4}	4.565
	MARQ	-----	-----	-----	-----	-----	-----	---
	POW	0.167	0.275	0.554	9.26	1.0×10^{-9}	8.295786×10^{-4}	8.274
	SPI	-----	-----	-----	-----	-----	-----	---
PROBABILISTICAS	MAXL	-----	-----	-----	-----	-----	-----	---
	MAXS	-----	-----	-----	-----	-----	-----	---
	MAXGS	-----	-----	-----	-----	-----	-----	---
	MAXGC	0.167	0.275	0.554	9.26	0.8×10^{-9}	8.29579×10^{-4}	4.205
	MAXPOW	-----	-----	-----	-----	-----	-----	---
	MAXROS	0.167	0.30625	0.3665	5.51	-0.04687	8.2958×10^{-4}	3.886
	BAYS	-----	-----	-----	-----	-----	-----	---
	BAYGS	-----	-----	-----	-----	-----	-----	---
	BAYGC	0.203667	0.29860	0.55361	9.259882	-0.11484	1.34827×10^{-3}	6.058
	BAYPOW	-----	-----	-----	-----	-----	-----	---
	BAYROS	0.190396	0.23469	0.57469	5.231679	0.006628	7.35781×10^{-4}	4.610

Gran cantidad de técnicas fallan en este modelo al no converger debido a la invalidez de la elevación de uno de los parámetros a la potencia de 0.05; esto sucede tanto en las técnicas de Mínimos Cuadrados como en las Probabilísticas. La técnica de Búsqueda Directa no converge adecuadamente, puesto que maximiza la suma residual inicial. En Mínimos Cuadrados, el algoritmo de Gradientes Conjugados llega a un valor menor que el inicial para la suma residual en 4.565 seg. La técnica Bayes-Rosenbrock de las Probabilísticas, llega al mejor valor de la suma residual en el presente modelo, requiriendo 4.610 seg.

EJEMPLO NUMERO 16.

Esterificación isopropanol-ácido acético.

Se ha estudiado la esterificación del alcohol isopropílico con ácido acético (ref. 18), efectuándose mediciones a tres temperaturas distintas (110°, 116° y 120°C), usando como catalizador resinas de intercambio iónico, en un reactor de flujo en fase gaseosa, a presión atmosférica. Representétese la reacción por



en donde A es el alcohol, B es el ácido, C es el ester y W es - el agua.

El mecanismo considera como paso controlante a la reacción en la superficie, estando los reactivos adsorbidos en un centro activo cada uno. El modelo de velocidad de reacción es:

$$r = \frac{k K_A K_B p_A p_B}{(1 + K_A p_A + K_B p_B + K_C p_C + K_W p_W)^2}$$

Los datos experimentales reportados son:

Presión parcial, atm.				mol/seg g-cat	
p_A	p_B	p_C	p_W	$rx10^5$	
0.589	0.289	0.006	0.017	3.22	
0.391	0.291	0.005	0.015	2.53	$T = 120^\circ C$
0.293	0.490	0.004	0.015	3.46	
0.191	0.191	0.400	0.020	1.23	

mol/seg g cat.

VALORES INICIALES DE LOS PARAMETROS
atm⁻¹

SUMA RESIDUAL INICIAL DE CUADRADOS

k = 0.00154

K_A = 1.0

K_B = 0.33

K_N = 1.12

K_G = 0.01

3.92211 x 10⁻¹¹

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS					SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		k	K _A	K _B	K _N	K _G		
BUSQUEDA DIRECTA	H-J	0.00164	1.000100	0.3301	1.1199	0.01	2.16777 x 10 ⁻¹⁰	3.259
MINIMOS CUADRADOS	GAUSS	-----	-----	-----	-----	-----	-----	---
	GS	0.001590	1.00000	0.3300002	1.119999	0.00999999	3.61915712 x 10 ⁻¹¹	7.399
	GC	0.001007	1.041478	0.5706952	0.848152	0.0082042	2.266196 x 10 ⁻¹¹	4.935
	MARQ	-----	-----	-----	-----	-----	-----	---
	POW	0.00154	1.00000	0.33	1.12	0.01	3.92211 x 10 ⁻¹¹	8.058
	SPI	0.001481	0.67116	0.1630540	0.361201	30.517965	6.9325 x 10 ⁻¹⁸	4.637
PROBABILISTICAS	MAXL	-----	-----	-----	-----	-----	-----	---
	MAXS	0.001590	1.00000	0.329934	1.1201	0.0101097	3.50791 x 10 ⁻⁸	3.750
	MAXGS	-----	-----	-----	-----	-----	-----	---
	MAXGC	0.001539	1.0	0.33	1.12	0.0099998	3.92211 x 10 ⁻¹¹	4.193
	MAXPOW	0.00154	1.0	0.33	1.12	0.01	3.91952 x 10 ⁻¹¹	5.576
	MAXPOS	0.001539	0.905673	0.3378213	-0.0414502	-1.927646	2.01666 x 10 ⁻¹²	4.096
	BAYS	0.001290	0.891155	0.456786	1.02605	-0.0933107	8.97403 x 10 ⁻⁸	3.617
	BAYGS	-----	-----	-----	-----	-----	-----	---
	BAYGC	0.0019052	0.999782	0.2635472	1.1094589	-0.246116	3.82158 x 10 ⁻¹¹	4.826
	BAYPOW	-----	-----	-----	-----	-----	-----	---
	BAYROS	0.0015374	0.993932	0.3310395	-0.0349015	-0.6134073	3.13982 x 10 ⁻¹²	4.858

Nuovamente para este modelo se maximiza la suma residual - inicial de cuadrados por el método de Búsqueda Directa. La técnica Spiral destaca en Mínimos Cuadrados, por su mejor valor de - suma residual, el cual se obtuvo en 4.637 seg de proceso. La -- técnica de Marquardt, también de Mínimos Cuadrados, no funciona puesto que el número de incognitas es mayor que el número de da tos experimentales en este modelo. La técnica de Gauss no alcan za la convergencia por excederse en el tiempo de proceso al es- tar calculando la variable dependiente. Tambien se observa fal- ta de convergencia por exceso de tiempo de proceso en varias té nicas Probabilísticas al estar evaluando la suma residual; sin- embargo la técnica Maximum-Rosenbrock alcanza un buen valor pa- ra la suma residual en 4.096 seg.

La comparación de los diversos métodos de estimación origi na la selección del algoritmo Spiral que alcanza el mejor valor para la suma residual.

EJEMPLO NUMERO 17.

Velocidades iniciales de dismutación del propileno.

M. J. Lewis y G. B. Willis han trabajado en la determinación de la velocidad inicial de dismutación del propileno (ref. 19) empleando como catalizador $\text{CoO-MoO}_3\text{-Al}_2\text{O}_3$ en un reactor de flujo a una temperatura de 436°K .

El modelo para el cálculo de la velocidad inicial es:

$$r = \frac{K_p p^2}{(1 + K_p p)^2}$$

Para la estimación de los parámetros de este modelo se emplearon los datos experimentales siguientes:

Presión total, atm	r, g mol/hr g-cat	
0.939	0.13	
0.939	0.155	
0.939	0.124	
2.939	0.277	T = 436°K
2.939	0.299	
2.939	0.263	

VALORES INICIALES DE LOS PARAMETROS

SUMA RESIDUAL INICIAL DE CUADRADOS

 $k = 0.59$ $K_p = 1.50$ 4.10864×10^{-2}

TECNICAS		VALORES ESTIMADOS DE LOS PARAMETROS		SUMA RESIDUAL DE CUADRADOS	TIEMPO DE PROCESO (seg)
		K	K_p		
BUSQUEDA DIRECTA	H-J	1.41	2.5	2.0299247 $\times 10^{-2}$	2.931
MINIMOS CUADRADOS	GAUSS	0.270896	0.923508	1.199334 $\times 10^{-3}$	9.266
	GS	-0.0143981	0.3104423	1.2887741	4.612
	GC	0.7738283	1.3215854	1.200622 $\times 10^{-3}$	4.530
	MARQ	0.27089625	0.92350879	1.1993333 $\times 10^{-3}$	5.968
	POW	0.354	1.14647238	8.28340277 $\times 10^{-3}$	6.741
	SPI	0.786134819	1.33634319	1.199 $\times 10^{-3}$	4.240
PROBABI- LISTICAS	MAXL	-----	-----	-----	---
	MAXS	0.786048	1.33626	7.95623 $\times 10^{-8}$	3.118
	MAXGS	-----	-----	-----	---
	MAXGC	0.58999797	1.4999988	4.10867 $\times 10^{-2}$	3.811
	MAXPOW	1.5213927	2.272835	1.08977 $\times 10^{-2}$	5.136
	MAXROS	0.59	1.50	4.10864 $\times 10^{-2}$	3.829
	BAYS	-----	-----	-----	---
	BAYGS	-----	-----	-----	---
	BAYGC	0.78385626	1.0532206	5.17565 $\times 10^{-2}$	4.661
	BAYPOW	-----	-----	-----	---
BAYROS	0.80841431	1.339111	1.38305 $\times 10^{-3}$	4.474	

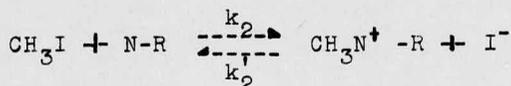
En el ejemplo número 17, la técnica de Búsqueda Directa -- reduce la suma residual inicial en solamente 2.931 seg de proceso. En las técnicas de Mínimos Cuadrados destacan las técnicas de Gauss, Marquardt y Spiral, pero por el menor tiempo de proceso (4.240 seg) puede considerarse al algoritmo Spiral como el mejor. El algoritmo Maximum-Simplex obtiene un mejor valor para la suma residual en un menor tiempo de proceso, (3.118 seg). -- Algunas técnicas probabilísticas no alcanzan la convergencia, - como es el caso del algoritmo Bayes-Simplex que se excede en el tiempo de proceso al maximizar la ecuación dada por el teorema de Bayes.

Como resultado de la comparación de los algoritmos mencionados, se considera que la técnica Maximum-Simplex converge mas aceptablemente.

EJEMPLO NUMERO 18.

Cinética de la reacción entre el yoduro de metilo y la dimetil-p-toluidina.

La reacción que se lleva a cabo en solución de nitro-benceno entre el yoduro de metilo y la dimetil-p-toluidina, en fase líquida, fué estudiada (ref. 20) sumergiendo recipientes sellados conteniendo a los reactivos, en un baño de temperatura constante, formándose una sal cuaternaria de amonio ionizado. La reacción es :



El modelo de velocidad de reacción se expresa como:

$$\log r = \log k_2 + \log ((0.05 - C_I)^{n_1+n_2} - 0.70 (C_I)^{n_3+n_4})$$

donde C_I indica la concentración del ión yoduro.

Los datos empleados en la estimación de los parámetros, obtenidos a partir de una concentración inicial de 0.05 g mol/l de yoduro de metilo y dimetil-p-toluidina son:

C_I g mol/l	r	log r
0.0	1.93×10^{-5}	-4.71
0.00875	1.12×10^{-5}	-4.95
0.0171	0.62×10^{-5}	-5.20
0.0201	0.42×10^{-5}	-5.37
0.0261	0.13×10^{-5}	-5.89

El modelo del ejemplo número 18 presenta extraordinaria dificultad en la evaluación de sus parámetros, tanto en su forma potencial como logarítmica para los tres tipos de técnicas. Dentro de mínimos Cuadrados las técnicas de Gradientes Simples, -- Gradientes Conjugados y Spiral presentan invalidez en el cálculo de las derivadas. Asimismo, la mayor parte de las técnicas - Probabilísticas no llegan a la convergencia cuando se presenta invalidez logarítmica al estar evaluando la variable dependiente para el cálculo de la suma residual.

La técnica de Búsqueda Directa maximizó la suma residual inicial de cuadrados; la técnica de Gauss converge minimizando la suma residual y obtiene valores aceptables para la suma algebraica de los parámetros ($n_1 + n_2$) y ($n_3 + n_4$), requiriendo 8.640 seg de proceso; el algoritmo Maximum-Gradientes Conjugados no consigue ningún avance en la minimización en la suma residual inicial.

Las técnicas mencionadas son las únicas que convergen en este modelo, siendo seleccionada obviamente la técnica de Gauss como la mejor.

NOTA: Los programas correspondientes a las técnicas empleadas en este trabajo para la estimación de parámetros de los diversos modelos cinéticos, fueron codificados en FORTRAN y ejecutados en la computadora Burroughs B-6700 del C.S.C., UNAM.

TABLA DE PUNTUACION DE LAS TECNICAS DE ESTIMACION
DE PARAMETROS EMPLEADAS EN LOS EJEMPLOS CINETICOS

EJEMPLOS

TECNICAS		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	PUNTUACION
BUSQUEDA DIRECTA	H-J																			0
MINIMOS CUADRADOS	GAUSS																		1	1
	GS		1																	1
	GC	1				1														2
	MARQ			1			1								1					3
	POW																			0
	SPI				1						1		1	1				1		5
PROBABILISTICAS	MAXL																			0
	MAXS									1								1		2
	MAXGS																			0
	MAXGC																			0
	MAXPOW																			0
	MAXROS																			0
	BAYS							1	1			1								3
	BAYGS																			0
	BAYGC																			0
	BAYPOW																			0
	BAYROS																	1		1

TOTAL 18

1 = Técnica que funcionó mejor en cada ejemplo.

Conclusiones.

De la comparación efectuada entre las diferentes técnicas - de estimación de parámetros en el capítulo 4, pueden concluirse varios aspectos importantes.

En primer lugar, según el criterio y necesidades del usuario, podrán utilizarse las técnicas que encuentren un valor aceptable de la suma residual de cuadrados en poco tiempo de proceso (hasta 5 seg) ó aquellas técnicas que convergan con mayor exactitud pero que requieran un mayor tiempo de proceso (mas de 5 seg).

Tales situaciones pueden presentarse cuando en un proceso industrial específico, se necesite rapidez en el cálculo y conocimiento aproximado de los parámetros que intervengan en dicho proceso, ó cuando lo que interese en una investigación específica, sea la exactitud en la evaluación de los parámetros del modelo en estudio (análisis de mecanismo de reacción), sin importar el tiempo de proceso.

5.1 Dificultades presentadas por las diferentes técnicas al ser aplicadas a la estimación de parámetros.

Se encontraron dificultades en varias de las técnicas utilizadas en este trabajo para el cálculo de los parámetros, las --- cuales se mencionan a continuación.

1.- En la técnica de Búsqueda Directa se presentó el problema de que se dirigía a un punto en el cual la suma residual de cuadrados era mayor que la inicial, ó en ocasiones no mejoraba la suma residual inicial, al utilizar dicha técnica en ciertos tipos de-

modelos.

2.- Dentro de Mínimos Cuadrados, la técnica de Gauss minimiza la suma residual inicial aceptablemente, como se observa en las tablas de resultados del capítulo 4, aunque requiere mayor tiempo de proceso que cualquier otro tipo de técnica; es por lo tanto recomendable usarla en todos los tipos de modelos.

3.- Respecto al algoritmo de Marquardt es conveniente recordar que no funciona cuando el número de parámetros a estimar es mayor que el número de datos experimentales, es decir, que el número de ecuaciones (ver ejemplo número 16 del capítulo 4).

4.- En general, las técnicas Probabilísticas de Máxima Semejanza presentaron el problema de obtener valores negativos para la desviación estandar, al estar iterando, puesto que la desviación -- fué considerada como un parámetro más a estimar.

5.- Especialmente se registraron problemas cuando se trataba de estimar los parámetros de modelos cinéticos potenciales ó logarítmicos, por originarse exponentes demasiado grandes ó logaritmos de magnitudes negativas (ver ejemplos números 3 y 18 del capítulo 4)

6.- Asimismo se presentó exceso en el tiempo de proceso para diferentes técnicas tanto de Mínimos Cuadrados como Probabilísticas, debido a que la secuencia de iteraciones fué demasiado larga, para el tiempo de proceso límite, que en común se aplicó a todas las técnicas.

5.2 Analisis de la puntuación obtenida para las diferentes técnicas empleadas en la estimación de parámetros.

De la tabla de puntuación del capítulo 4, se observa que la técnica Spiral obtuvo la mayor puntuación, en cuanto a obtener el menor valor para la suma residual de cuadrados en tiempos cor

tos de proceso.

Siguiendo un orden de importancia, se encuentran las técnicas de Marquardt y Bayes-Simplex, que mostraron igual efectividad en la estimación de parámetros.

A continuación se tienen las técnicas de Gradientes Conjugados y Maximum-Simplex que obtuvieron la misma puntuación al ser aplicados.

Finalmente se presentan las técnicas de Gauss, Gradientes - Simples y Bayes-Rosenbrock con la misma frecuencia de efectividad.

5.3 Conclusiones y recomendaciones para el uso de las diversas técnicas seleccionadas, según sea el modelo del que se interesen calcular sus parámetros.

1.- Para modelos homogéneos potenciales irreversibles de la forma:

$$r = k C_A^{\alpha} C_B^{\beta}$$

es recomendable utilizar las técnicas de Marquardt y Gauss de Mínimos Cuadrados (ejemplo número 3 del capítulo 4).

2.- Para modelos homogéneos potenciales reversibles de la forma:

$$r = k C_A^{\alpha} C_B^{\beta} - k' C_C^{\delta}$$

se recomienda utilizar la técnica de Gauss (ejemplo número 18 -- del capítulo 4).

3.- Para modelos heterogéneos irreversibles de la forma:

$$r = \frac{kC_A C_B}{(1 + K_1 C_A + K_2 C_B)^n}$$

se recomienda usar las técnicas: Spiral, Marquardt, Maximum-Simplex, Bayes-Simplex (ejemplos números 1, 2, 4, 6, 7, 8, 9, 10, 14, 15, 16, y 17 del capítulo 4).

4.- Para modelos heterogéneos reversibles de la forma:

$$r = \frac{k(C_A C_B - C_C / K_{eq})}{(1 + K_1 C_A + K_2 C_B + K_3 C_C)^n}$$

es recomendable utilizar las técnicas: Spiral, Gradientes Conjugados y Bayes-Simplex (ejemplos números 5, 11, 12 y 13 del capítulo 4).

Desde luego pueden presentarse variaciones en las formas -- funcionales de los modelos que deseen estudiarse y combinaciones entre dichas formas funcionales en sistemas complejos y en tales casos deberá efectuarse la selección más conveniente de las técnicas estudiadas en este trabajo, según el criterio del usuario.

En el apéndice se enlistan los programas seleccionados en el presente capítulo.

APENDICE.

PROGRAMAS

- 1.- Técnica de Gauss. (Ref. 21)
- 2.- Técnica de Gradientes Simples. (Ref. 21)
- 3.- Técnica de Gradientes Conjugados. (Ref. 21)
- 4.- Técnica de Marquardt. (Ref. 21)
- 5.- Técnica Spiral.

El ejemplo No. 4 se utiliza para ejemplificar la técnica --
Spiral.

Requerimientos del usuario:

- a) Determinar los valores de: N , KP , TAU , EXS , $(BO(J), J=1, KP)$.
- b) Ajustar la proposición DIMENSION de acuerdo al problema en --
particular.
- c) Especificar modelo, derivadas y función objetiva de mínimos --
cuadrados en la subrutina DERIV.
- d) Ajustar formatos de salida según se requiera por el problema--
en particular.
- e) Este programa puede operar para las variables independientes--
 X , W , Z , U , V . Ajustar al número de variables independientes--
según el problema en particular.

```

C   TÉCNICA SPIRAL DE MINIMOS CUADRADOS
C   PROGRAMA PRINCIPAL
C   N#   NUMERO DE DATOS O NUMERO DE ECUACIONES
C   NP   NUMERO DE INCOGNITAS
C   K#   VECTOR DE INCOGNITAS
C   X    VECTOR DE DATOS DE VARIABLE INDEPENDIENTE
C   Y    VECTOR DE DATOS DE VARIABLE INDEPENDIENTE
C   YAU  VECTOR DE DATOS DE VARIABLE DEPENDIENTE
C   TAU  PARAMETRO DE CONVERGENCIA
C   EXS  PARAMETRO DE CONVERGENCIA
C   DF   VECTOR DE DERIVADAS

```

```

DOUBLE PRECISION A*P
REAL MU
DIMENSION A(30,30),T(30),B(30,30),D(30,30),PT(30,30),F(30),REST
12(30),E(30),G(30),H(30),I(30),J(30),K(30),L(30),M(30),N(30)
COMMON X(30),Y(30),YAU(30),Z(30),U(30),V(30)
C   LECTURA DE NUMERO DE DATOS Y DE PARAMETROS
READ (5,*) N#K#

```

```

READ (5,*) (B(K,J), J=1,KP)
READ (5,*) (H(I), I=1,N)
READ (5,*) (T(I), I=1,N)
READ (5,*) TAU,EXS
WRITE (6,003)

```

```

003 FORMAT (1H1,2X,30HALGORITMO DE REGRESION SPIRAL)
WRITE (6,006)
006 FORMAT (12X,35HVALORES I NICIALES DE LOS PARAMETROS)
40  FORMAT (12X,35HBO(,12,4H) = ,1PE16.8)
WRITE (6,007) TAU,EXS
007 FORMAT (12X,6HTAU = ,F6. ,4X,6HEXS = ,F6.4)
WRITE (6,005)

```

```

C   005 FORMAT (12X,35HVALORES DE LOS COEFICIENTES FINALES)
23  CALCULO DE LA MATRIZ A(I,J)
CALL DERIV (DF,N,KP,F0,B0,F1,0,1)

```

```

DO 13 J=1,KP
DO 13 I=1,N
A(I,J)=DF(I,J)
13  F(I,J)=T(KP,J)
DO 10 I=1,KP
DO 10 J=1,N
A(I,J)=0
DO 11 K=1,N
11  A(I,J)=A(I,J)+PT(I,K)*P(K,J)
10  CONTINUE
DO 14 I=1,N
14  REST(I)=Y(I)-F(0)
DO 15 J=1,KP
G(J)=0
DO 12 I=1,N
115 G(J)=PT(J,I)*REST(I)+G(J)
12  CONTINUE
15  CONTINUE
DO 41 J=1,KP
41  A(J,KP+1)=G(J)
DETER=SIMUL (KP,A,T,30)

```

```

IF (DETER=EQ=0) CALL MARC (A,AMBDA,T,KP,30)
ORG=0.
ORT=0.
DO 16 J=1,KP
ORG=G(J)**2*ORG
16  ORT=T(J)**2*ORT
ORG=SQR(ORG)
ORT=SQR(ORT)
DO 17 J=1,KP
D(J)=G(J)*ORT/ORG
17  COSA=D(J)*T(J)/(ORT**2)
30  CUEN=1.
VMAX=0.
DO 18 J=1,KP
H(J)=T(J)/(B(J)+TAU)
18  IF (ABS(H(J))>GT:VMAX) VMAX=ABS(H(J))
IF (VMAX*LT:EXS) GO TO 34
DO 20 J=1,KP
20  B(J)=B0(J)+T(J)

```

```

CALL DERIV (D,F,N,KP,F0,B,FI,2)
IF (E1.F10) GO TO 21
25 DO 22 J=1,KP
22 B(J)=B(J)
F10=F1
GO TO 23
C
21 POSSIBLE INTERPOLACION DE SERIE DE TAYLOR
MU=0.1
MU=UR1
SENGA=SQRT(1.-COSGA**2)
TANGA=SENGA/COSGA
GAMMA=ATAN(TANGA)
27 TANTE=MU*SENGA/(1.-MU*MU*SENGA)
TETA=ATAN(TANTE)
PSI=R0*MU*SENGA/(SIN(TETA))
R=PSI*(1.-TETA+COS(GAMMA/2.))*(TETA/GAMMA)
1**2)
DO 24 J=1,KP
24 S(J)=B(J)*((MU*B(J))+(1.-MU)*T(J))
CALL DERIV (D,F,N,KP,F0,B,FI,2)
IF (E1.F10) GO TO 25
C
25 POSIBLE INTERPOLACION SPIRAL
MU=2.*MU/(1.-MU)
GO TO 27
26 CUENTA=CUENTA+1
IF (CUENTA.GT.4.) GO TO 28
DO 29 J=1,KP
29 T(J)=T(J)/2.
ORT=ORT/2.
GO TO 30
28 ORS=0.
DO 31 J=1,KP
31 ORS=6*(J**2)*GRS
ORS=SQRT(ORS)
DO 32 J=1,KP
32 D(J)=-G(J)*ORS/GRS
DO 3 J=1,KP
37 B(J)=B(J)+D(J)
CALL DERIV (D,F,N,KP,F0,B,FI,2)
IF (E1.L1.F10) GO TO 25

```

```

VMD=0.
DO 33 J=1,KP
33 IF (ABS(D(J)).GT.VMD) VMD=ABS(D(J))
IF (VMD.L1.0.00001) GO TO 34
DO 35 J=1,KP
35 D(J)=D(J)/2.
GO TO 36
34 DO 134 J=1,KP
134 WRITE (6,50) J, B(J)
50 FORMAT (1,2A,2L8(,I2,4H) = ,1PE16.8)
WRITE (6,51) PI
51 FORMAT (1X,22HLA MINIMA RE_PUESTA ES, E15.5/)
CALL EXIT
END

```

```

SUBROUTINE MARQ (A,AMBDA,T,KP,NRC)
DOUBLE PRECISION A,T
DIMENSION A(NRC,NRC),T(NRC)
AMBDA=0.01
DO I=1,KP
1 A(I,I)=A(I,I)+AMBDA
DETER=SMUL (KP,A,T,NRC)
RETURN
END

```

```

SUBROUTINE DERIV(DF,N,KP,F0,B0,F10,LI)
DIMENSION F0(30),B0(30),DF(30,30)
COMMON X(30),W(30),Y(30),Z(30),U(30),V(30)
F10=8.0
DO TO (41)*2)LI
DO 39 I=1,N
F0(I)=(B0(1)*X(I)*W(I))/(1.0+(B0(2)*X(I))+(B0(3)*W(I)))**2
39 CONTINUE
DO 40 Y=1,N
DF(I,1)=X(I)*W(I)/(1.0+(B0(2)*X(I))+(B0(3)*W(I)))**2
DF(I,2)=2.0*((B0(1)*X(I)*W(I)+X(I))/(1.0+(B0(2)*X(I))+(B0(3)*W(I)
1))**3)
DF(I,3)=2.0*((B0(1)+X(I)*W(I))**2)/(1.0+(B0(2)*X(I))+(B0(3)*W(I)
1))**3)
40 CONTINUE
42 DO 142 I=1,N
142 F10=(F0(I)*X(I)*W(I))/(1.0+(B0(2)*X(I))+(B0(3)*W(I)))**2)
142*2)F10
RETURN
END

```

```

FUNCTION SIMUL(N,A,X,NRC)
DOUBLE PRECISION X,EPS
DOUBLE PRECISION PIVOT,DETER,A
DIMENSION IRW(50),JCOL(50),JORD(50),Y(50),A(NRC,NRC),X(NRC)
DETER=1.0
MAX=N+1
DETER=1.0
DO 10 K=1,N
C K=K-1
ELEMENTO PIVOTE
PIVOT=0
DO 11 I=1,N
DO 11 J=1,N
IF (X.EQ.0) GO TO 9
DO 8 ISCAN=1,KMI
DO 8 JSCAN=1,KMI
IF (X.EQ.0) IRW(ISCAN) GO TO 11
IF (J.EQ.0) JCOL(JSCAN) GO TO 11
8 CONTINUE
9 IF (DABS(A(I,J)).LE.DABS(PIVOT)) GO TO 11
PIVOT=A(I,J)
IRW(K)=I
JCOL(K)=J
11 CONTINUE
IF (DABS(PIVOT).GT.EPS) GO TO 13
SIMUL=0.
PRINT 100,PIVOT,EPS
FORMAT (2X,"SALI CON SIMUL=0 ",2X,E25.5,2X,F25.5)
13 IRWK=IRW(K)
JCOLK=JCOL(K)
DETER=DETER*PIVOT
NORMALIZACION
DO 14 J=1,MAX
C A(IRWK,J)=A(IRWK,J)/PIVOT
C ELIMINACION
Y(I)=A(I,JCOLK)=1./PIVOT
DO 18 I=1,N
AIJCK=A(I,JCOLK)
IF (I.EQ.0) IRWK=I
A(I,JCOLK)=AIJCK/PIVOT
DO 17 J=1,MAX
17 IF (J.NE.0) A(I,J)=A(I,J)-AIJCK*A(IRWK,J)
18 CONTINUE
DO 23 I=1,N
IRWI=IRW(I)
JCOLI=JCOL(I)
JORD(IRWI)=JCOLI
23 X(JCOLI)=A(IRWI,MAX)
SIMUL=200.
RETURN
END

```

ALGORITMO DE REGRESION SPIRAL

VALORES INICIALES DE LOS PARAMETROS

BO(1) = 1.37410000E+01

BO(2) = 1.46400000E+01

BO(3) = 1.79000000E+01

TAM = 0.0010 EXS = 0.0001

VALORES DE LOS COEFICIENTES FINALES

B(1) = 1.26613956E+01

B(2) = 1.11875060E+01

B(3) = 1.84882045E+01

LA MINIMA RESCUERDA ES 4245 E-10

6.- Técnica Maximum-Simplex

El ejemplo No. 4 se utiliza para ejemplificar la técnica --
Maximum-Simplex.

Descripción de parámetros:

- NX Número de incógnitas, (incluyendo la desviación estandar).
- STEP Parámetro limitante del tamaño de paso (valor sugerido = :
0.50).
- N Número de datos ó número de ecuaciones.
- X Vector de incógnitas.
- PL Vector de datos de variable independiente.
- ZL Vector de datos de variable independiente.
- YL Vector de datos de variable dependiente.

Requerimientos del usuario.

- a) Determinar los valores de: NX, STEP, N, (X(I), I = 1, NX).
- b) Ajustar la proposición DIMENSION de acuerdo al problema en --
particular.
- c) Especificar modelo y función objetivo de mínimos cuadrados en
la subrutina SPUR.
- d) Este programa puede operar para las variables independientes--
PL, ZL, TL, W1.

TECNICA MAXIMUM-SIMPLEX. PROGRAMA PRINCIPAL

101 DIMENSION (C(5),S(5),Y(5),Z(5),SUM(5),FL(5),ZL(5),TL(5),HL(5)),

102 C(1),Y(1),Z(1),SUM(1),STEL(1),S(1),H(1)

100 REAL(5),ZL(5),TL(5),HL(5),STEL(1)

101 FL(5)=1.0
102 S(5)=1.0
103 Z(5)=1.0
104 Y(5)=1.0
105 TL(5)=1.0
106 HL(5)=1.0
107 STEL(1)=1.0

ALFA=1.0
L1=0.0
CALFA=2.0
L1L1=C.
Y1=Y
Z1=Z
CALL SUBR (Y1,FL,Z1,TL,HL,N)
IF (L1 .GT. S(1)) GO TO 11
IF (L1 .GT. STEL) GO TO 11
IF (L1 .GT. 1.0) GO TO 11

103 FL=AT(CX,1) FUNCTION VALUE, 15X, 3HX1=, 20X, 3HX2=, 20X, 3HX3=, 20X,
104 3HX4=, 10X, 12HEM C. CHANGE)

102 FL=AT(C1,1,1)CX, 31 FUNCTION STARTING VALUE, 120.5, /, "THE X ARRAY IS"

101 /, 5X, 1 (C11,4,2X)

1002 FL=AT(12X,"STEL"=116.0)

H1=1+X
H2=1+X+Z
H3=1+Y+Z
H4=1+Y+Z
CALL ST/IT (Y1,FL,Z1,TL,HL,N)

20 DO 3 I=1,N
DO 4 J=1,NX
X(J)=Y1(I,J)

3 CALL SUBR (Y1,FL,Z1,TL,HL,N)
CONTINUE

21 SUMH=SUM(1)
I=1
DO 7 I=1,N
IF (SUM(I).IF.SUM) GO TO 7
SUMH=SUM(I)

7 CONTINUE
SUMH=SUM(I)
K=1
DO 8 I=1,N
IF (SUMH.IF.SUM(I)) GO TO 8
SUMH=SUM(I)

6 CONTINUE
DO 9 J=1,NX
SUM2=0.

10 DO 10 I=1,N
SUM2=SUM2+Y1(I,J)
X1(I,2,J)=1.0+Y1(I,J)*Z(SUM2-Y1(I,HLX,J))
X1(I,3,J)=(1.0+ALFA)*X1(I,2,J)-ALFA*X1(INDEX,J)
9 X(J)=Y1(I,3,J)

11 I=1
CALL SUBR (Y1,FL,Z1,TL,HL,N)
IF (SUM(I,3).GT.SUM) GO TO 11
IF (SUM(I,3).GT.1.0) GO TO 11
SUMS=SUM(I)

37 DO 11 I=1,N
SUMS=SUM(I)
38 DO 12 I=1,N
IF (SUM(I,3).GT.L1) GO TO 12
IF (SUM(I,3).GT.SUM) GO TO 12
SUMS=SUM(I)

12 CONTINUE
IF (SUM(I,3).GT.SUMS) GO TO 13
GO TO 14
11 DO 14 I=1,N
X1(I,3,J)=(1.0+ALFA)*X1(I,2,J)+ALFA*X1(I,3,J)

```

15 X(J)=Y*(I,J)
   II=I
   CALL SUBT (YL,PL,ZL,TL,WL,N)
   IF (SUM(CI)*IT.SUM(I)) GO TO 16
   GO TO 17
16 X(I,II)=Y*(I,II)
   II=II+1
17 IF (II.EQ.N) GO TO 18
   X(I,II)=PL*(I,II)+X(I,II)+PL*(I,II)*X(I,II)
18 X(J)=Y*(I,J)
   CALL SUBT (YL,PL,ZL,TL,WL,N)
   II=II
   IF (SUM(CI)*SUM(I)) GO TO 16
   GO TO 17
19 X(I,J)=C.*(X(I,J)+X(I,II)*Y)
   GO TO 17
20 X(J)=Y*(I,J)
   II=I
   CALL SUBT (YL,PL,ZL,TL,WL,N)
   GO TO 21
21 X(I,II)=Y*(I,II)
   II=II+1
   CALL SUBT (YL,PL,ZL,TL,WL,N)
   GO TO 22
22 X(J)=X(I,II)*Y
   II=II+1
   CALL SUBT (YL,PL,ZL,TL,WL,N)
   GO TO 23
23 X(J)=X(I,J)
   II=II
   CALL SUBT (YL,PL,ZL,TL,WL,N)
   DIFF=C.
   GO TO 24
24 DIFF=DIFF+(SUM(CI)-SUM(I))*Y
   DIFF=1./X*SUM(I)*DIFF
101 PRINT 1(1,SUM(L*(X1(ROUNT,J), J=1,N)),DIFF)
   FORMAT(2(2Y,F10.6),3(7X,F10.6),12Y,F10.6)

```

IF (DIFF.CF.C.C6(C61)) GO TO 28

GO TO 100

999 CONTINUE

CALL EXIT

END

----- TO THE FOLLOWING FOR EXPERIMENTAL ACTION ON THE I/O OF

SUBROUTINE START (YL,PL,ZL,TL,WL,N)

DIFF=1./X*SUM(CI)*X1(50,50)*X1(50,50)*SUM(50)*PL(50)*ZL(50)*TL(50)

1 X(50),YL(50)

COMMON/DEF Y,X1,X2,STEP,XI,SUM,IN

VF=1

STEP1=STEP/(VN*SQRT(C2))*(SQRT(VN+1.))+VN-1.)

STEP2=STEP/(VN*SQRT(C2))*(SQRT(VN+1.))-1.)

DO 1 J=1,N

1 A(1,J)=C.

DO 2 I=1,N

DO 3 J=1,N

A(I,J)=STEP2

L=1

A(L,I)=STEP1

CONTINUE

DO 3 J=1,N

DO 3 J=1,N

3 X(I,J)=X(J)+A(L,J)

RETURN

END

```

SUBROUTINE SOL (YL, PL, ZL, TL, N, PD)
DIMENSION YL(50), PL(50), ZL(50), TL(50), N(50), PD(50)
NYL(50)
COMPILE / FORT / X(1), STEL, N(1), SIM, TN
FL = 1.0
LI = 30
FL = ((YI(J) - (YCL) + FL(J) * ZI(J)) / (1.0 + Y(3) * FL(J) + X(4) * ZL(J)) ** 2))
1 ** 2) + 1
30 CONTINUE
GUL(1) = 1.0 * AL(1) + 1.0 * (Y(1) ** 2) + FL / (2.0 * X(1) ** 2)
RETURN
END

```

VALOR DE LA FUNCION

X1

X2

-.143634E+03

-.153282E-05

.139164E+00

X3

X4

.150957E+02

.191665E+02

DIFFERENCIA DE CUADRADOS

.433304E-07

7.- Técnica Bayes-Simplex.

El ejemplo No. 4 se utiliza para ejemplificar la técnica -- Bayes-Simplex.

Descripción de parámetros:

- NX Número de incógnitas, (incluyendo el factor normalizante).
- STEP Parámetro limitante del tamaño de paso (valor sugerido = 1.0).
- N Número de datos ó número de ecuaciones.
- XO Vector de incógnitas.
- X Vector de incógnitas.
- W Matriz covariancia inversa para los parámetros a estimar con dimensiones NX x NX.
- PL Vector de datos de variable independiente.
- ZL Vector de datos de variable independiente.
- YL Vector de datos de variable dependiente.
- SIGMA Desviación estandar.

Requerimientos del usuario.

- a) Determinar los valores de: NX, STEP, N, SIGMA, (W(K,J), --
K = 1, NX, J = 1, NX), (XO(I), I = 1, NX), (X(I), I = 1, NX).
- b) Ajustar la proposición DIMENSION de acuerdo al problema en --
particular.
- c) Especificar el modelo en la función objetivo de mínimos cua--
drados en la subrutina SUMR.
- d) Ajustar formatos de salida según se requiera por el problema--
en particular.
- e) Este programa puede operar para las variables independientes--
PL, ZL, TL, WL, VL. Ajustar al número de variables indepen---
dientes según el problema en particular.

B6700/B7700 FORTRAN COMPILATION MARK

```
0 TECNICA BAYES-SIMPLEX. PROGRAMA PRINCIPAL
DIMENSION X1(50,50),X(50),SUM(50),PL(50),ZL(50),TL(50),WL(50),
100 COHMG/ONE7 X,X1,NX,STEP,K1,SUM2,N,XO,H,SIGMA
100 READ (5,/)END=999) NX,STEP,N,SIGMA

DO 200 K=1,NX
200 READ(5,/) (X(K,J),J=1,NX)
READ(5,/) (XO(I),I=1,NX)
READ(5,/) (X(I),I=1,NX)
READ(5,/) (PL(I),I=1,N)
READ(5,/) (ZL(J),J=1,N)
READ(5,/) (YL(J),J=1,N)
WRITE(6,/)NX,STEP,(X(I),I=1,NX)

ALFA=1.0
BETA=0.5
GAMA=2.0
DIFER=0.
XNX=NX
IN=1
CALL SUMR (YL,PL,ZL,TL,WL,VL,N)
PRINT 1002,STEP
1002 FORMAT(2X,STEP=N,F6.2)
IN=103102
102 FORMAT (4X,7HSUM(1)=,E11.4)
WRITE(6,103) (X(I),I=1,NX)
103 FORMAT (/,3(2X,2HX(,12,4H) = ,E11.4))
K1=NX+1
K2=NX+2
K3=NX+3
K4=NX+4
25 CALL START (YL,PL,ZL,TL,WL,VL,N)
DO 3 I=1,K1
DO 4 J=1,NX
4 X(J)=X1(I,J)
IN=1
CALL SUMR (YL,PL,ZL,TL,WL,VL,N)
3 CONTINUE
28 SUMH=SUM(1)
INDEX=1
DO 7 I=2,K1
IF (SUM(I).LE.SUMH) GO TO 7
SUMH=SUM(I)
INDEX=I
7 CONTINUE
SUML=SUM(1)
COUNT=1
DO 8 I=2,K1
IF (SUML.GE.SUM(I)) GO TO 8
SUML=SUM(I)
COUNT=I
8 CONTINUE
DO 9 J=1,NX
SUM2=0
DO 10 I=1,K1
10 SUM2=SUM2+X1(I,J)
X1(K2,J)=1./XNX*(SUM2=X1(INDEX,J))
X1(K3,J)=(1.+ALFA)*X1(K2,J)-ALFA*X1(INDEX,J)

9 X(J)=X1(K3,J)
IN=K3
CALL SUMR (YL,PL,ZL,TL,WL,VL,N)
IF (SUM(K3).LT.SUM1) GO TO 11
IF (INDEX.EQ.1) GO TO 38
SUMS=SUM(1)
GO TO 39
38 SUMS=SUM(2)
DO 12 I=1,K1
IF ((INDEX-I).EQ.0) GO TO 12
```

```

IF (SUM(I).LE(SUMS) GO TO 12
SUMS=SUM(I)
12 CONTINUE
IF (SUM(K3).GT.SUMS) GO TO 13
GO TO 12
11 DO 15 J=1,NX
X1(K4,J)=(I-GAMA)*X1(K2,J)+GAMA*X1(K3,J)
15 X(J)=X1(K4,J)
IN=K4
CALL SUMR (YL,PL,ZL,TL,WL,VL,N)
IF (SUM(K4).LT.SUM1) GO TO 16
GO TO 13
13 IF (SUM(K3).GT.SUMH) GO TO 17
DO 18 J=1,NX
X1(INDEX,J)=X1(K3,J)
17 DO 19 J=1,NX
X1(K4,J)=BETA*X1(INDEX,J)+(1-BETA)*X1(K2,J)
19 X(J)=X1(K4,J)
CALL SUMR (YL,PL,ZL,TL,WL,VL,N)
IN=K4
IF (SUMH.GT.SUM(K4)) GO TO 16
DO 20 I=1,K1
DO 20 J=1,NX
20 X(I,J)=0.5*(X1(I,J)+X1(KOUNT,J))
DO 29 I=1,K1
DO 30 J=1,NX
30 X(J)=X1(I,J)
IN=1
CALL SUMR (YL,PL,ZL,TL,WL,VL,N)
29 CONTINUE
GO TO 26
DO 21 J=1,NX
X1(INDEX,J)=X1(K4,J)
21 X(J)=X1(INDEX,J)
IN=INDEX
CALL SUMR (YL,PL,ZL,TL,WL,VL,N)
GO TO 26
DO 22 J=1,NX
X1(INDEX,J)=X1(K3,J)
22 X(J)=X1(INDEX,J)
IN=INDEX
CALL SUMR (YL,PL,ZL,TL,WL,VL,N)
DO 26 J=1,NX
26 DO 23 J=1,NX
23 X(J)=X1(K2,J)
IN=K2
CALL SUMR (YL,PL,ZL,TL,WL,VL,N)
DIFER=0
DO 24 I=1,K1
24 DIFER=DIFER+(SUM(I)-SUM(K2))**2
DIFER=1./NXX*SQRT(DIFER)
PRINT 101,SUM1,X1(KOUNT,J), J=1,NX,DIFER
101 FORMAT(2(2X,E16.6),3(7X,E16.6),12X,E16.6)
IF (DIFER.GE.0.0000001) GO TO 28
GO TO 100
999 CONTINUE
CALL EXIT
END

```

```

SUBROUTINE START (YL,PL,ZL,TL,WL,VL,N)
DIMENSION A(50,50),X(50,50),SUM(50),PL(50),ZL(50),TL(50),
1 WL(50),VL(50),VL(50)
COMMON /ONE7 X,X1,NX,STEP,K1,SUM,IN,X0,h,SIGMA
VN=NX
STEP1=STEP/(VN*SQRT(2.))*(SQRT(VN+1.))+VN-1.
STEP2=STEP/(VN*SQRT(2.))*(SQRT(VN+1.))-1.
DO 1 J=1,NX
1 A(1,J)=0.
DO 2 I=2,K1
DO 2 J=1,NX
A(I,J)=STEP2
L(I,J)=STEP1
2 CONTINUE
DO 3 I=1,K1
DO 3 J=1,NX
3 X(I,J)=X(J)+A(I,J)
RETURN
END

```

```

SUBROUTINE SUMR (YL,PL,ZL,TL,HL,VL,N)
DIMENSION X1(50,50),X(50),SUM(50),PL(50),ZL(50),TL(50),HL(50),
1 VL(50),YL(50),H(50,50),A(50,50),X0(50)
COMMON/ONE/ X,X1,NX,STEP,K1,SUM,IN,X0,H,SIGMA
DO 10 J=1,NX
DO 10 K=1,NX
10 A(1,J)=X(K)-X0(K)*H(K,J)
RESUL=0.
DO 11 J=1,NX
11 RESUL=A(1,J)*X(J)-X0(J)+RESUL
FL=0.0
DO 12 J=1,N
12 FL=(YL(J)-((X(2)*PL(J)+ZL(J))/(1.0+X(3)*PL(J)+X(4)*ZL(J)**2))**2
1)+FL
30 CONTINUE
SUM(IN)=ALOG(X(1))-(FL/(2.*(SIGMA)**2))-(RESUL/2.)
SUM(IN)=-SUM(IN)
RETURN
END

```

VALOR DE LA FUNCION

X1

X2

-.807276E+01

.999999E+06

.132573E+00

X3

X4

.137687E+02

.191948E+02

DIFERENCIA DE CUADRADOS

.323680E-07

8.- Técnica Bayes-Rosenbrock.

El ejemplo No. 4 se utiliza para ejemplificar la técnica --
Bayes-Rosenbrock.

Descripción de parámetros:

KM Número de incógnitas, (incluyendo el factor normalizante).

MAXK Máximo número de veces que será evaluada la función del teorema de Bayes.

EKAT Máximo número de veces que serán rotados los ejes.

MCYC Número de fallas sucesivas encontradas en todas las direcciones antes de finalizar el programa.

NSTEP Variable de control que determina el tamaño de paso a -- ser utilizado después de cada rotación de los ejes. --
NSTEP = 1 si se utiliza tamaño de paso inicial. NSTEP = 2 si se utiliza tamaño de paso final en la etapa previa.

NN Número de datos ó número de ecuaciones.
EPSY Error que se alcanza antes de finalizar el programa, en
 la función del teorema de Bayes.
ALPHA Factor escalante para el incremento del tamaño de paso.
BETA Factor escalante para la reducción del tamaño de paso.
AKE Vector de valores iniciales para las incógnitas.
XO Vector de valores iniciales para las incógnitas.
EPS Vector de tamaños de paso iniciales para las incógnitas.
SIGMA Desviación estandar.
W Matriz covariancia inversa para los parámetros a estimar
 con dimensiones KM x KM.
PL Vector de datos de variable independiente.
ZL Vector de datos de variable independiente.
YL Vector de datos de variable dependiente.

Requerimientos del usuario.

- a) Determinar los valores de : KM, MAXK, MKAT, MCOYC, NSTEP, NN,
EPSY, ALPHA, BETA, (AKE(I), I = 1, KM), (XO (I), I = 1, KM), --
(EPS(J), J = 1, KM), SIGMA, (W(K, J), K = 1, KM, J = 1, KM).
- b) Ajustar la proposición DIMENSION de acuerdo al problema en --
particular.
- c) Especificar el modelo en la función objetiva de mínimos cua--
drados en la subrutina OBJECT.
- d) Ajustar formatos de salida según se requiera por el problema--
en particular.
- e) Este programa puede operar para las variables indeoendientes--
PL, ZL, TL, WL, VL. Ajustar al número de variables indepen --
dientes según el problema en particular.

B6700/B7700 FORTRAN COMPILATION MARK 2.

```

C TECNICA BAYES-ROSENBROCK. PROGRAMA PRINCIPAL
DIMENSION N(30,30),X0(30)
DIMENSION EPS(20),AKE(20),V(20,20),D(20),BL(20,20),BLEN(20),AJ(20)
1 A(20),AL(20,20),AFK(20),YL(30),PL(30),TL(30),HL(30),VL(30),ZL(30)
COMMON SIGMA,H,R0
READ(5,/) KM,MAXK,MKAT,MCYC,NSTEP,NN

READ(5,/)(AKE(I),I=1,KM)
READ(5,/)(X0(I),I=1,KM)
READ(5,/)(EPS(J),J=1,KM)
READ(5,/) EPSY,ALPHA,BETA
READ(5,/) SIGMA
DO MCC K=1,KM
200 READ(5,/)(CHK(J),J=1,KM)
READ(5,/)(PL(J),J=1,NN)
READ(5,/)(ZL(J),J=1,NN)
READ(5,/) (YL(J),J=1,NN)
CALL ROSENB(AKE,EPS,KM,MAXK,MKAT,MCYC,ALPHA,BETA,V,NSTEP,EPSY,D,
1 BL,BLEN,AJ,E,AL,AFK,YL,PL,ZL,TL,HL,VL,NN)
999 STOP
CALL EXIT
END

```

```

SUBROUTINE ROSENB(AKE,EPS,KM,MAXK,MKAT,MCYC,ALPHA,BETA,V,NSTEP,
1 EPSY,D,BL,BLEN,AJ,E,AL,AFK,YL,PL,ZL,TL,HL,VL,NN)
DIMENSION AKE(KM),D(KM),V(KM,KM),BL(KM,KM),BLEN(KM),EPS(KM),
1 J(KM),E(KM),AL(KM,KM),AFK(KM),TL(KM),PL(KM),ZL(KM),TL(KM),HL(KM),
2 VL(KM))
COMMON /B137/ B125
WRITE(6,999)

```

```

99 FORMAT(1H1,10X,34HROSENBROCK MAXIMIZATION PROCEDURE )
1004 WRITE(6,1004) MAXK,MKAT,MCYC,NSTEP,ALPHA,BETA,EPSY
1004 FORMAT(//,10X,10HPARAMETERS //,2X,7HNAXK = ,1X,4X,7HMKAT = ,1X,4X,
1 7HMCYC = ,1X,4X,7HNSSTEP = ,1X,4X,7HSHALPHA = ,1X,4X,7HBETA = ,
2 5X,2,4X,7HEPSY = ,1X,4X,7H)
DO 98 II=1,KM
DO 98 JJ=1,KM
V(II,JJ) = 0.0
IF(II=JJ) V(II,JJ) = 1.0
98 CONTINUE
CALL OBJECT(AKE,SUMN,KM,YL,PL,ZL,TL,HL,VL,NN)
SUMD = SUMN
DO 97 K=1,KM
AFK(K) = AKE(K)
812 CONTINUE
RK1 = 1
IF (NSTEP=1) 701,700,701
700 GO TO 1000
701 CONTINUE
DO 350 I=1,KM
E(I) = EPS(I)
350 CONTINUE
1000 DO 702 J=1,KM
FBEST = SUMN
AJ(J) = 0.0
IF (A5) 702,703,702
702 GO TO 250
703 CONTINUE
E(I) = EPS(I)
250 D(I) = 0.0
III = 0
397 III = III + I
258 I = 1
259 DO 251 J=1,KM
251 AKE(J) = AKE(J) + E(I) * V(I,J)
CALL OBJECT(AKE,SUMN,KM,YL,PL,ZL,TL,HL,VL,NN)
KAT = KAT + 1
SUMDIF = FBEST - SUMN
IF (ABS(SUMDIF) = EPSY) 704,704,705
704 GO TO 1001
705 CONTINUE
IF (KAT = MAXK) 706,707,707
707 GO TO 1001
706 CONTINUE
IF (SUMN = SUMD) 708,708,709
708 GO TO 253
709 CONTINUE
DO 254 J=1,KM
254 AKE(J) = AKE(J) - E(I) * V(I,J)
E(I) = BETA * E(I)

```

```

1001 WRITE (6,1002) KK1
1002 FORMAT (//,2X,25HTOTAL NUMBER OF STAGES = ,I2)
      WRITE (6,1003) RAT
1003 FORMAT (//,2X,38HTOTAL NUMBER OF FUNCTION EVALUATIONS ,I5)
      WRITE (6,1005) SUMO
1005 FORMAT (//,2X,33HFINAL VALUE OF ABJECT FUNCTION = ,E16,0)
      DO 1007 IX=1,KM
      WRITE (6,1006) IX,AKE(IX)
1006 FORMAT (//,2X,ZHC,12,4H) = ,E16,8)
1007 CONTINUE
300 FORMAT (//,2X,26HDIFERENCIA DE CUADRADOS = ,E15,6)
      WRITE (6,300) B125
      RETURN
      END

```

```

SUBROUTINE OBJECT (AKE,SUMN,KM,YI,PL,ZL,XI,WI,VI,NN)
DIMENSION AKE(KM),VL(CNN),PL(CNN),ZL(CNN),X(30),VL(CNN),WL(CNN),VL(CNN)
COMMON /G4A,N=20
COMMON /B137/ FL
DO 2 I=1,KM
2 K=I+AKE(I)
FL=0.0
DO 30 J=1,NN
30 FL=((YI(J)-((X(2)*PL(J)+ZL(J))/(1.0+X(3)+PL(J)+X(4)+ZL(J)**2))
1+*2)*PL
DO 10 J=1,KM
DO 10 K=1,KM
10 A(I,J)=((X(K)-X0(K))*H(K,J)
RESUL=0.0
DO 11 J=1,KM
11 RESUL=A(I,J)+X(J)-X0(J)+RESUL
SUMN=ALOG(X(1))-((FL/(2.0*(SIGMA)**2)))-(RESUL/2.0)
SUMN=SUMN
RETURN
END

```

ROSENBROCK MAXINIZATION PROCEDURE

```

PARAMETERS
MAXK = 3000      MKAT = 30      MCYC = 50      NSTEP = 2
ALPHA = 1.00     BETA = 0.25    EPSY = 1.0000E-05

```

```

TOTAL NUMBER OF STAGES = 1
TOTAL NUMBER OF FUNCTION EVALUATIONS 3000
FINAL VALUE OF ABJECT FUNCTION = -1.66275918E-01
( , 1 ) = .47500000E+03
( , 2 ) = .13685818E+00
( , 3 ) = .14275071E+02
( , 4 ) = .19580330E+02
DIFERENCIA DE CUADRADOS = .264386E-10

```

BIBLIOGRAFIA

- (1) Hooke R., and T. A. Jeeves "Direct Search Solution of Numerical and Statistical Problems", J. Assoc. Comp. March., 8, 212-229, 1961.
- (2) Himmelblau David M. Process Analysis by Statistical Methods John Wiley and Sons, Inc., New York, 1968.
- (3) Fletcher R. and M. J. D. Powell "A Rapidly Convergent Descent Method for Minimization", Computer J., 6, 163-168, -- 1963.
- (4) Fletcher R. and C. M. Reeves "Function Minimization by Conjugate Gradients", Computer J. 7, 149-154, 1964.
- (5) Marquardt, D. M. "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," J. Soc. Indust. Appl. Math. 11, 431-441, 1963.
- (6) Powell, M. J. D. "A Method for Minimizing a Sum of Squares of Non-Linear Functions Without Calculating Derivatives," Computer J., 7, 303-307, 1965.
- (7) Jones A. "Spiral - A New Algorithm for Non-Linear Parameter Estimation Using Least Squares," Computer J., 13, 301-308, 1970.
- (8) Powell, M. J. D. "An Efficient Method for Finding the Minimum of a Function of Several Variables Without Calculating Derivatives," Computer J., 7, 155-162, 1964.

- (9) Rosenbrock, H. H. "An Automatic Method for Finding the --
Greatest or Least Value of a Function," Computer J., 3, --
175-184, 1960.
- (10) R. Prasad and A. K. Kar "Kinetics and Mechanism of the Va-
por-Phase Ammoxidation," Ind. Eng. Chem., Process Des. Dev.
Vol. 15, No. 1, 170-175, 1976.
- (11) Wolfenden "Numerical Problems in Physical Chemistry," Ox-
ford University Press, New York, 1938.
- (12) R. J. Ayen and M. S. Peters "Catalytic Reduction of Nitric
Oxide," Ind. Eng. Chem., Process Des. Develop., Vol. 1 --
No. 3, 204-207, 1962.
- (13) George Thodos and L. F. Stutzman "Synthesis of Ethyl Chlo-
ride," Industrial and Engineering Chemistry, Vol. 50, --
No. 3, 413-416, 1958.
- (14) Charles Potter and Seymour Baron "Kinetics of the Cataly--
tic Formation of Phosgene", Chemical Engineering Progress, -
Sep., Vol. 47, No. 9, 473-480, 1951.
- (15) M. V. Sussman and Charles Potter "Kinetics of Propylene --
and Ethylene Hydrogenation," Industrial and Engineering --
Chemistry, March, Vol. 46, No. 3, 457-465, 1954.
- (16) J. C. Gottifredi, A. A. Yeramian and R. E. Cunningham "Va-
por-Phase Reactions Catalyzed by Ion Exchange Resins," --
Journal of Catalysis, 12, 245-256, 1968.

- (17) N. O. Lemcoff and R. E. Cunningham "Kinetics of Diacetone-Alcohol Conversion to Mesityl Oxide Catalyzed by Ion Exchange Resin," Journal of Catalysis, 23, 81-92, 1971.
- (18) A. A. Yeramian, J. C. Gottifredi and R. E. Cunningham "Vapor-Phase Reactions Catalyzed by Ion Exchange Resins," -- Journal of Catalysis, 12, 257-262, 1968.
- (19) M. J. Lewis and G. B. Wills "Initial Velocities of Disproportionation of Propylene," Journal of Catalysis, Vol. 15, 140-143, 1969.
- (20) J. M. Smith "Chemical Engineering Kinetics", Mc Graw-Hill-Book Company, New York, 1970.
- (21) James L. Kuester and Joe H. Mize "Optimization Techniques-with Fortran," McGraw-Hill Book Company, New York, 1973.
- (22) Stanley M. Walas "Cinética de Reacciones Químicas," McGraw Hill Book Company, New York, 1965.
- (23) Daniel D. McCracken "A guide to FORTRAN IV programming", - John Wiley and Sons, Inc., New York, 1972.