

397



Universidad Nacional Autónoma
de México

Facultad de Ciencias

OPTIMIZACION.
Métodos y Elementos de Teoría

T E S I S

Que para obtener el Título de

A C T U A R I A

p r e s e n t a

Anne Marie Cajuste Bontemps

TESIS CON
FALLA DE ORIGEN

México, D. F.

1989



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 está 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.

INDICE

INTRODUCCION	1
CAPITULO I. LOCALIZACION DE PUNTOS EXTREMOS	2
A. EXTREMOS SIN RESTRICCIONES. CASO UNIVARIADO	2
1. Teorema del valor medio	4
2. Máximos y mínimos relativos	7
B. EXTREMOS SIN RESTRICCIONES. CASO MULTIVARIADO	12
1. Condiciones necesarias y suficientes para extremos	12
2. Método del análisis canónico	15
C. MAXIMOS Y MINIMOS RESTRICTOS. METODO DE LOS MULTIPLICADORES DE LANGRANGE	19
CAPITULO II. EXTREMOS DE FUNCIONES POR METODOS NUMERICOS	23
A. EXTREMOS DE FUNCIONES EN UNA VARIABLE	23
1. Búsqueda en 3 puntos del intervalo	24
2. Funciones convexas	25
B. EXTREMOS DE FUNCIONES MULTIVARIADAS SIN RESTRICCIONES	28
1. Método del ascenso acelerado	28
2. Método de Newton-Raphson	30
3. Programa SAS para la aplicación del método de Newton-Raphson	38
C. OPTIMIZACION MULTIVARIABLE CON RESTRICCIONES. METODO DE NEWTON-RAPHSON	40
CAPITULO III. AJUSTE DE FUNCIONES DE RESPUESTA	42
A. EL MODELO LINEAL GENERAL	43
1. Modelos lineales especiales	43
2. Modelos linealizables	45
B. MODELOS NO LINEALES	45
1. El modelo semilogarítmico	45
2. El modelo logarítmico	48
3. La logística	49

C. AJUSTE POR MINIMOS CUADRADOS	49
1. Ajuste de modelos de regresión lineal	52
2. Análisis de varianza	54
3. Ajuste de modelos de regresión no lineal	55
D. EJEMPLOS ILUSTRATIVOS	57
CAPITULO IV. ILUSTRACIONES DE LA TECNICA	72
A. OPTIMOS CON POLINOMIOS ENTEROS	72
1. Polinomio de segundo grado	72
2. Polinomio de segundo grado en varias variables	73
3. Ilustración de la técnica del análisis canónico	75
B. OPTIMOS CON POLINOMIOS DE EXPONENTES FRACCIONARIOS	79
1. Polinomios pseucadrático	79
2. Rendimiento e ingreso máximo con un polinomio pseucadrático	81
C. MAXIMOS Y MINIMOS RESTRICITOS	86
1. Máximo restringido con una función potencial	86
2. Costo mínimo restringido a un nivel fijo de producción	88
3. Rendimiento máximo restringido a un crédito fijo	91
CAPITULO V. RESUMEN	95
CAPITULO VI. REFERENCIAS	96

INTRODUCCION

En muchos campos de la investigación, cuando se desea cubrir simultáneamente los aspectos matemáticos, estadísticos y prácticos del problema de la localización de óptimos económicos, el investigador confronta la carencia de una referencia especializada que cubra los aspectos anteriores del propio problema. Las referencias existentes sobre el tema, desarrollan el tema ignorando algunos de los aspectos fundamentales del problema. Así por ejemplo, si el énfasis es en la teoría, los autores se concentran en hacer desarrollos matemáticos, omitiendo la fase estadística de la investigación, e ilustrando su discusión, con ejemplos de pizarrón muy lejanos a lo que pudiera ser una situación de la vida real. Cuando el énfasis se hace sobre los aspectos aplicados del problema, la teoría es prácticamente ignorada por completo, y muchas veces no es posible garantizar que se ha logrado una situación óptima.

Por las razones anteriores, el presente trabajo tiene como objetivo primordial, lograr combinar, más equitativamente, los tres aspectos fundamentales de la localización de condiciones óptimas. Tal propósito pretende lograrse en esta tesis, resumiendo algunos resultados matemáticos y de cálculo numérico, así como el uso de las herramientas estadísticas, y en particular, aquélla de la regresión.

Una característica especial que se manifiesta en esta tesis, es el empleo de paquetes de cómputo electrónico, comprobando la excelencia del paquete estadístico SAS (Statistical Analysis System-Sistema para Analisis Estadístico).

CAPITULO I

LOCALIZACION DE PUNTOS EXTREMOS.

Una de las aplicaciones de las derivadas es la determinación de los valores del máximo y mínimo de una función.

En este capítulo se consideran algunos teoremas que desarrollan la teoría de las condiciones requeridas para la ocurrencia de un máximo o de un mínimo, para funciones de una o más variables con y sin restricciones; esto se hace con el apoyo de algunos ejemplos ilustrativos. En el siguiente capítulo se presentará la ilustración del cálculo de puntos extremos de funciones por métodos numéricos.

A. EXTREMOS SIN RESTRICCIONES. CASO UNIVARIADO.

Se dice que la función f tiene un máximo relativo en un punto c , si existe una vecindad $N(c)$ de c , tal que, para toda x en $N(c)$ y en el dominio de f :

$$f(x) \leq f(c).$$

Análogamente, se dice que la función f tiene un mínimo relativo en un punto c , si existe una vecindad $N(c)$ de c , tal que, para toda x en $N(c)$ y en el dominio de f :

$$f(x) \geq f(c).$$

Los máximos y mínimos relativos de una función, se denominan valores extremos de la función.

Una función continua definida sobre un intervalo cerrado, tiene un máximo y un mínimo sobre el propio intervalo.

Teorema 1.1.

Si:

- i) La función f tiene un valor extremo en c ;
- ii) f está definida en una vecindad de c ; y
- iii) $f'(c)$ existe, entonces $f'(c)=0$.

Demostración:

Supongamos que f tiene un máximo relativo en c . Existe un número $\eta > 0$ tal que f está definida sobre $\langle c-\eta, c+\eta \rangle$ y, para toda $x \in \langle c-\eta, c+\eta \rangle$, $f(x) \leq f(c)$. Luego para toda $h \in \langle 0, \eta \rangle$,

$$\frac{f(c+h) - f(c)}{h} \leq 0,$$

y, por tanto, $f'(c) \leq 0$. Para toda $h \in \langle -\eta, 0 \rangle$, se tiene:

$$\frac{f(c+h) - f(c)}{h} \geq 0,$$

y, por tanto, $f'(c) \geq 0$. Como $f'(c)$ existe por hipótesis, $f'(c) = f'(c)$ y, por tanto, $f'(c) = 0$. ■

Con este teorema se deduce lo siguiente: los valores extremos de una función definida sobre un intervalo pueden ocurrir sólo en los puntos críticos; éstos son los puntos del intervalo donde la derivada es cero o no existe, y también los puntos extremos del intervalo si es que pertenecen al propio intervalo (Hasser et al. 1979).

Se puede concluir que, una condición necesaria para que una función tenga un extremo relativo (máximo o mínimo) en un punto c , es que c sea un punto crítico; la inversa no es necesariamente cierta.

Para la determinación de los valores máximos y mínimos de una función se procede como sigue: si el número de puntos críticos es finito, calculamos el valor de la función en cada uno de los puntos críticos y el mayor de estos valores que se obtenga es el máximo, el menor es el mínimo (Hasser et al. 1979).

Ejemplo 1.1.

Encuéntrense los valores máximos y mínimos de la función:

$$f(x) = x - 3x^{1/3}$$

sobre el intervalo $[-1, 2]$.

Solución:

Para determinar los puntos críticos de f sobre $[-1, 2]$ se

calcula la primera derivada de f :

$$f'(x) = 1 - x^{-2/3}, \quad x \neq 0$$

Se observa que f' no existe en $x = 0$ y que ésta se anula sólo en $x = \pm 1$. Por lo tanto los puntos críticos de f sobre el intervalo $[-1, 2]$ son $-1, 0, 1, 2$. Calculando el valor de f en cada uno de los puntos críticos se tiene:

$$f(-1) = 2,$$

$$f(0) = 0,$$

$$f(1) = -2,$$

$$f(2) = 2 - 3\sqrt[3]{2} > -2.$$

Por consiguiente, $f(-1) = 2$ es el máximo y $f(1) = -2$ es el valor mínimo de la función f sobre $[-1, 2]$.

1. Teorema del valor medio.

Con el teorema del valor medio se podrán desarrollar criterios que nos permitan probar si un punto crítico es un máximo o un mínimo relativo de una función, en el caso en que la función no esté definida sobre un intervalo cerrado.

Un caso particular del teorema del valor medio es el llamado Teorema de Rolle (Hasser et al. 1979).

Teorema 1.2.

(Teorema de Rolle). Supóngase que f es continua sobre $[a, b]$, donde $a < b$, y diferenciable en $\langle a, b \rangle$, además $f(a) = f(b)$. Entonces existe un punto $c \in \langle a, b \rangle$ tal que $f'(c) = 0$.

Demostración:

Si $f(x) > 0$ para alguna $x \in \langle a, b \rangle$, sea c un punto sobre $[a, b]$ donde f alcanza su máximo, entonces como $f(c) > 0$, mientras que $f(a) = f(b)$, tenemos que $c \in \langle a, b \rangle$. Como f está definida en la vecindad $\langle a, b \rangle$ de c y $f'(c)$ existe, por el Teorema 1.1. se tiene

que $f'(c) = 0$.

Si $f(x) < 0$ para alguna $x \in \langle a, b \rangle$, hacemos que c sea un punto de $\langle a, b \rangle$ donde f tome su valor mínimo. entonces razonando como en el caso anterior se tiene que $f'(c) = 0$. ■

La condición de que f' exista en todo punto de $\langle a, b \rangle$ es necesaria para la validez del Teorema de Rolle (Hasser et al. 1979).

Teorema 1.3.

(Teorema del Valor Medio). Si f es continua en $[a, b]$ donde $a < b$ y diferenciable sobre $\langle a, b \rangle$, entonces existe un punto c en $\langle a, b \rangle$ tal que:

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Demostración:

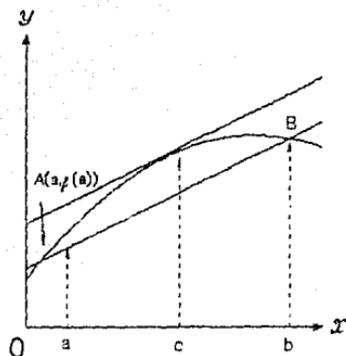


Figura 1.1

Con relación a la Figura 1.1, considérese la ecuación de la recta que pasa por los puntos a y b , a saber:

$$y - f(a) = \frac{f(b) - f(a)}{b - a} (x - a),$$

o bien:

$$y = \frac{f(b) - f(a)}{b - a} (x - a) + f(a).$$

También se define la función $F(x)$, como la distancia vertical entre un punto $(x, f(x))$ en la gráfica de la función f y el punto correspondiente en la recta secante que pasa por a y b , es decir:

$$F(x) = f(x) - \frac{f(b) - f(a)}{b - a} (x - a) - f(a).$$

Se tiene que la función F es continua en el intervalo cerrado $[a, b]$, puesto que es la suma de f y un polinomio lineal, los cuales son continuos allí; también se tiene que f es diferenciable en $\langle a, b \rangle$ y, además, de la ecuación para $F(x)$, se observa que $F(a)=0$ y $F(b)=0$. Dado lo anterior, vemos que se cumplen las condiciones para la validez del Teorema de Rolle; por la conclusión de dicho teorema, se establece que existe una c en el intervalo abierto $\langle a, b \rangle$, tal que $F'(c)=0$; de aquí, se tiene:

$$F'(x) = f'(x) - \frac{f(b) - f(a)}{b - a},$$

$$\Rightarrow F'(c) = f'(c) - \frac{f(b) - f(a)}{b - a} = 0.$$

Por lo tanto, existe un número c en $\langle a, b \rangle$ tal que:

$$f'(c) - \frac{f(b) - f(a)}{b - a} = 0,$$

de donde:

$$f'(c) = \frac{f(b)-f(a)}{b-a} \quad \blacksquare$$

2. Máximos y mínimos relativos.

El siguiente teorema se demuestra en Hassler et al. (1977):

Teorema 1.4.

Si f es continua sobre un intervalo $[a,b]$ y $f'(x) \geq 0$ sobre $\langle a,b \rangle$, entonces f es no decreciente sobre $[a,b]$. Análogamente, si f es continua sobre $[a,b]$ y $f'(x) \leq 0$ en $\langle a,b \rangle$, entonces f es no creciente sobre $\langle a,b \rangle$.

Teorema 1.5.

Si c es un punto crítico de f y si existe un intervalo $[a,b]$ con $c \in \langle a,b \rangle$ tal que f es continua sobre $[a,b]$ y:

- i) $f'(x) \geq 0$ para $x \in \langle a,c \rangle$ y $f'(x) \leq 0$ para $x \in \langle c,b \rangle$, entonces f tiene un máximo relativo en c ;
- ii) $f'(x) \leq 0$ para $x \in \langle a,c \rangle$ y $f'(x) \geq 0$ para $x \in \langle c,b \rangle$, entonces f tiene un mínimo relativo en c ;
- iii) $f'(x) > 0$ para $x \in \langle a,c \rangle$ y $f'(x) > 0$ para $x \in \langle c,b \rangle$, o $f'(x) < 0$ para $x \in \langle a,c \rangle$ y $f'(x) < 0$ para $x \in \langle c,b \rangle$, entonces f no tiene ni un máximo ni un mínimo relativo en c .

Demostración: Sólo el inciso i)

Como $f'(x) \geq 0$ para $x \in \langle a,c \rangle$, f es no decreciente sobre $[a,c]$. De donde:

$$f(x) \leq f(c) \text{ para toda } x \in \langle a,c \rangle.$$

Como $f'(x) \leq 0$ para $x \in \langle c,b \rangle$, f es no creciente sobre $[c,b]$. De donde:

$$f(x) \leq f(c) \text{ para toda } x \in \langle c,b \rangle.$$

Así pues, $f(x) \leq f(c)$ para toda $x \in \langle a,b \rangle$ y f tiene un

máximo relativo en c .

Ejemplo 1.2.

Encuéntrense los máximos y los mínimos relativos de la función f definida por:

$$f(x) = \frac{1}{2} x^2(x-4)^{1/3}, \quad -\infty < x < \infty.$$

Solución:

Como:

$$\begin{aligned} f'(x) &= \frac{1}{6} x^2(x-4)^{-2/3} + x(x-4)^{1/3} \\ &= \frac{1}{6} (x-4)^{-2/3} [x^2 + 6x^2 - 24x] \\ &= \frac{1}{6} x(7x-24)(x-4)^{-2/3}, \quad x \neq 4 \end{aligned}$$

f' no existe en 4 y es cero en 0 y en $\frac{24}{7}$.

Investigando el valor de f' alrededor de estos puntos críticos, se tiene:

$$f'(x) > 0, \quad x \in \langle -\infty, 0 \rangle;$$

$$f'(x) < 0, \quad x \in \left\langle 0, \frac{24}{7} \right\rangle;$$

$$f'(x) > 0, \quad x \in \left\langle \frac{24}{7}, 4 \right\rangle \cup \langle 4, \infty \rangle.$$

Utilizando el Teorema 1.5, f tiene un máximo relativo en 0 y un mínimo relativo en 4. Esta función no tiene ni un máximo ni un mínimo absoluto (Hasser et al. 1979).

Teorema 1.6.

Supóngase que f es diferenciable en una vecindad $N(c)$ de c , $f'(c) = 0$ y que $f''(c)$ existe:

- i) Si $f''(c) < 0$, entonces f tiene un máximo relativo en c .
- ii) Si $f''(c) > 0$, entonces f tiene un mínimo relativo en c .

Demostración:

Si:

$$f''(c) = \lim_{x \rightarrow c} \frac{f'(x) - f'(c)}{x - c} < 0,$$

entonces existe un intervalo $\langle a, b \rangle$, que contiene a c , tal que:

$$\frac{f'(x) - f'(c)}{x - c} = \frac{f'(x)}{x - c},$$

para toda $x \in \langle a, c \rangle \cup \langle c, b \rangle$.

Se puede suponer que $\langle a, b \rangle \subset N(c)$. Entonces f es continua sobre $[a, b]$ y,

$$f'(x) > 0 \text{ para toda } x \in \langle a, c \rangle,$$

$$f'(x) < 0 \text{ para toda } x \in \langle c, b \rangle.$$

Por lo tanto, por el Teorema 1.5 f tiene un máximo relativo en c .

Obsérvese que los Teoremas 1.5 y 1.6 proporcionan sendos métodos para encontrar los valores extremos de una función.

Ejemplo 1.3.

Examinar las siguientes funciones con respecto a sus puntos extremos:

$$i) f(x) = x^3 - 3x, \quad x \in \mathbb{R}.$$

Solución:

Para puntos críticos $f'(x) = 0$; de aquí, se tiene:

$$f'(x) = 3x^2 - 3 \quad \text{y} \quad f''(x) = 6x;$$

de donde resulta $x^2 = 1$ al igualar $f'(x)$ con cero, ecuación que tiene dos soluciones:

$$x = 1 \quad \text{ó} \quad x = -1.$$

Al analizar estos resultados la derivada es cero sólo para $x = 1$ y $x = -1$, que son nuestros puntos extremos. Evaluando f en los puntos anteriores, se tiene:

$$f(1) = 1^3 - 3(1) = -2$$

y:

$$f(-1) = (-1)^3 - 3(-1) = 2.$$

Por consiguiente, f tiene un máximo en $x = -1$ y un mínimo en $x = 1$, puesto que $f''(-1) = -6$ y $f''(1) = 6$.

$$(i) f(x) = 4x - x^2, \quad x \in \mathbb{R}.$$

Solución:

Puesto que $f'(x) = 4 - 2x$, al igualar esta derivada con cero, se tiene en $x = 2$ el valor extremo único; éste es un máximo, puesto que $f''(x) = -2$ para toda $x \in \mathbb{R}$, y por lo tanto en $x = 2$. El máximo de f es $f(2) = 4(2) - 2(2) = 4$.

$$(ii) f(x) = x^4 + x^2, \quad x \in \mathbb{R}.$$

Solución:

$$\begin{aligned} f'(x) &= 4x^3 + 2x \\ &= 2x(2x^2 + 1) = 0. \end{aligned}$$

En este caso tenemos un valor extremo en $x = 0$, el cual es un mínimo puesto que $f'(x) < 0$ para $x \in \langle -\infty, 0 \rangle$ y $f'(x) > 0$ para $x \in \langle 0, \infty \rangle$.

El valor extremo es $f(0) = 0^4 + 0^2 = 0$. ■

Con los resultados anteriores se muestra el uso de los Teoremas 1.4 y 1.5, donde es condición necesaria para que f tenga un punto máximo o mínimo en x_0 , que $f'(x_0) = 0$. La inversa de este teorema es falsa; es decir, es posible que si $f'(x_0) = 0$, la

función puede no tener un extremo en $x = x_0$. El siguiente ejemplo ilustra esta posibilidad:

Ejemplo 1.4.

Analícese la siguiente función y determinense, si existen, sus puntos extremos:

$$f(x) = x^3, \quad x \in \mathbb{R}.$$

Solución:

Puesto que $f'(x) = 3x^2$, $f'(x) = 0$ implica $x = 0$. De aquí, se tiene que el valor de la derivada es positivo para $x \neq 0$, y sólo se hace nulo para $x = 0$. Por el Teorema 1.5 $f(x) = x^3$, $x \in \mathbb{R}$, no tiene un extremo en $x = 0$. Así, el ejemplo muestra que un punto para el cual la derivada se hace cero, puede no conducir a un valor máximo o mínimo de la función.

Teorema 1.7.

Si en un punto estacionario x_0 de $f(x)$ las primeras $(n-1)$ derivadas se anulan y $f^{(n)}(x) \neq 0$, entonces en $x = x_0$ ocurre que $f(x)$ tiene:

- i) un punto de inflexión si n es impar, y
- ii) un punto extremo si n es par. Este punto extremo será un máximo si $f^{(n)}(x_0) < 0$ y un mínimo si $f^{(n)}(x_0) > 0$.

Ejemplo 1.5.

Analícese las siguientes funciones:

i) $f(x) = x^4 \quad x \in \mathbb{R}.$

Solución:

$$f'(x) = 4x^3 = 0,$$

lo cual proporciona el punto estacionario $x_0 = 0$. Ahora bien:

$$f'(0) = f''(0) = f'''(0) = 0,$$

pero:

$$f^{(4)}(0) = 24 > 0,$$

por tanto en $x_0 = 0$ se tiene un punto mínimo.

$$i) f(x) = x^3 \quad x \in \mathbb{R}.$$

Solución:

Esta función proporciona a $x_0 = 0$ como punto estacionario. Ya que $f^{(n)}(0)$ no es cero cuando $n = 3$, se tiene entonces que $x_0 = 0$ es un punto de inflexión.

B. EXTREMOS SIN RESTRICCIONES. CASO MULTIVARIADO.

Un punto $\underline{x}_0 = (x_1, \dots, x_1, \dots, x_n)$ es un punto de máximo si $f(\underline{x}_0 + \underline{h}) \leq f(\underline{x}_0)$ para toda $\underline{h} = (h_1, \dots, h_1, \dots, h_n)$ tal que $|h_j|$ es suficientemente pequeña para toda $j = 1, 2, \dots, n$.

De manera similar, se define a \underline{x}_0 como un punto de mínimo, cuando $f(\underline{x}_0 + \underline{h}) \geq f(\underline{x}_0)$, para toda \underline{h} con las propiedades anteriores.

El gradiente tiene un papel importante en la identificación de los máximos y mínimos. Si un punto de una función con pendiente cero (gradiente) no es un extremo (máximo o mínimo), entonces debe ser un punto de inflexión (Taha 1976).

1. Condiciones necesarias y suficientes para extremos.

Existen teoremas que establecen condiciones necesarias y suficientes para los extremos de una función $f(\underline{x})$ en n variables. En los resultados que siguen se supondrá que existen las primeras y segundas derivadas parciales de $f(\underline{x})$, siendo éstas, a su vez, funciones continuas en \underline{x} .

Teorema 1.8.

Una condición necesaria para que \underline{x}_0 sea un punto extremo de $f(\underline{x})$, es que:

$$\nabla f(\underline{x}_0) = \phi.$$

donde $\nabla f = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right]$, es el gradiente de f y

$\phi = (0, 0, \dots, 0)$, es un vector de ceros.

De aquí, debe entenderse que en cualquier punto extremo, la condición $\nabla f(\underline{x}_0) = \phi$, tiene que satisfacerse, es decir, el valor del gradiente debe ser el vector cero. Tal condición es satisfecha también por los puntos de silla y de inflexión. Consecuentemente, la condición establecida por el Teorema 1.8 es necesaria pero no suficiente, para la existencia de un extremo.

Los puntos obtenidos a partir de la solución de $\nabla f(\underline{x}_0) = \phi$, son llamados por Taha (1976), puntos estacionarios.

Teorema 1.9.

Una condición suficiente para que un punto estacionario sea extremo es que la matriz hessiana (o de Hesse) H evaluada en \underline{x}_0 sea:

- i) positiva definida cuando \underline{x}_0 es un punto de mínimo, y
- ii) negativa definida cuando \underline{x}_0 es un punto de máximo.

Recuérdese que una matriz cuadrada ($n \times n$) es: positiva definida si los valores de los determinantes menores principales de la matriz son todos positivos. Por otra parte, una matriz cuadrada es: negativa definida si el valor del k -ésimo determinante menor principal de la matriz tiene el signo de $(-1)^k$, $k=1, \dots, n$. Recuérdese además que la matriz H de una función, es la matriz de las segundas derivadas.

Ejemplo 1.6.

i) Demuéstrase que la función:

$$f(x_1, x_2, x_3) = x_1 + 2x_2 + x_2x_3 - x_1^2 - x_2^2 - x_3^2, \quad (x_1, x_2, x_3) \in \mathbb{R}^3$$

obtiene un máximo absoluto.

Solución:

La condición necesaria para la ocurrencia de un extremo en el punto $\underline{x}_0 = (x_{10}, x_{20}, x_{30})$, es que $\nabla f(\underline{x}_0) = \phi$. De aquí:

$$\frac{\partial f}{\partial x_1} = 1 - 2x_1 = 0, \quad (1.1)$$

$$\frac{\partial f}{\partial x_2} = x_1 - 2x_2 = 0, \quad (1.2)$$

$$\frac{\partial f}{\partial x_3} = 2 + x_2 - 2x_3 = 0. \quad (1.3)$$

De la ecuación (1.1) se obtiene $x_1 = 1/2$. Por otra parte, resolviendo simultáneamente las ecuaciones (1.2) y (1.3), resulta: $x_2 = 2/3$ y $x_3 = 4/3$. Así, si f tiene un extremo, éste debe ocurrir en el punto $\underline{x}_0 = (1/2, 2/3, 4/3)$. Demostraremos que éste es un máximo absoluto.

Para establecer la condición de suficiencia, evaluaremos H , la matriz Hessiana, en $\underline{x}_0 = (1/2, 2/3, 4/3)$. Obsérvese que:

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix}$$

para todo vector $\underline{x} = (x_1, x_2, x_3)$ y, obviamente, para el vector $\underline{x}_0 = (1/2, 2/3, 4/3)$.

A continuación obtendremos, sucesivamente los determinantes menores principales de $H|_{\underline{x}_0}$:

$$|H_1| = |-2| = -2.$$

$$|H_2| = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4.$$

$$|H_g| = \begin{vmatrix} -2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{vmatrix} = -8 \quad (\text{Aplicando la regla de Sarrus para determinantes de } 3 \times 3).$$

Los valores -2 , 4 y -8 alternan en signos que coinciden con el signo de $(-1)^k$, $k=1,2,3$. Lo cual implica que la matriz $H|_{\underline{x}_0}$ es no negativa definida, demostrándose por el Teorema 1.9 que f tiene un máximo absoluto. El valor máximo de f , es:

$$f\left(\frac{1}{2}, \frac{2}{3}, \frac{4}{3}\right) = \frac{1}{2} + 2\left(\frac{2}{3}\right) + \frac{2}{3}\left(\frac{4}{3}\right) - \left(\frac{1}{2}\right)^2 - \left(\frac{2}{3}\right)^2 - \left(\frac{4}{3}\right)^2 = \frac{1}{4}.$$

(ii) Demuestre que la función:

$$f(x_1, x_2, x_3) = 2x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 + 2x_2x_3, \quad (x_1, x_2, x_3) \in \mathbb{R}^3$$

tiene un mínimo absoluto en \mathbb{R}^3

Solución:

Derivando a f con respecto a x_1, x_2 y x_3 , e igualando las derivadas a cero, se tiene:

$$\frac{\partial f}{\partial x_1} = 4x_1 + 2x_2 = 0,$$

$$\frac{\partial f}{\partial x_2} = 4x_2 + 2x_1 + 2x_3 = 0,$$

$$\frac{\partial f}{\partial x_3} = 6x_3 + 2x_2 = 0;$$

sistema que tiene una solución única: $x_1 = 0$, $x_2 = 0$, $x_3 = 0$, es decir, $\underline{x}_0 = (0, 0, 0)$.

Estableciendo la condición de suficiencia, se tiene:

$$H|_{\underline{x}_0} = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 2 & 6 \end{pmatrix}.$$

cuyos determinantes menores principales, son:

$$|H_1| = |4| = 4.$$

$$|H_2| = \begin{vmatrix} 4 & 2 \\ 2 & 4 \end{vmatrix} = 12.$$

$$|H_3| = \begin{vmatrix} 4 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 2 & 6 \end{vmatrix} = 56. \quad (\text{Después de aplicar la regla de Sarrus}).$$

Los valores 4, 12, y 56, todos positivos, implican que $H|_{x_0}$ es una matriz positiva definida; por lo tanto $\underline{x}_0 = (0,0,0)$, representa un punto mínimo; obsérvese que $f(0,0,0) = 0$.

Si $H|_{\underline{x}_0}$ no es positiva definida o negativa definida, \underline{x}_0 debe ser un punto silla.

Puede darse el caso donde esto no es concluyente, es decir, \underline{x}_0 puede ser o no un punto extremo, y el desarrollo de una condición de suficiencia llega a ser difícil; en algunos casos el método del análisis canónico que se describirá enseguida puede llevar a información más concluyente.

2. Método del análisis canónico.

Considérese el polinomio de segundo grado:

$$Q = \beta + \sum_{i=1}^p \beta_i Z_i + \sum_{i=1}^p \beta_{ii} Z_i^2 + \sum_{i=1}^p \sum_{j=1}^p \beta_{ij} Z_i Z_j. \quad (1.4)$$

Definanse los vectores Z y β , y la matriz B , por las relaciones:

$$Z = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_p \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} \beta_{11} & \beta_{12}/2 & \dots & \beta_{1p}/2 \\ \beta_{21}/2 & \beta_{22} & \dots & \beta_{2p}/2 \\ \vdots & \vdots & \dots & \vdots \\ \beta_{p1}/2 & \beta_{p2}/2 & \dots & \beta_{pp} \end{bmatrix}$$

En términos de los elementos anteriores, el polinomio Q dado por, (1.4), puede escribirse:

$$Q = \beta_0 + Z' \beta + Z' B Z. \quad (1.5)$$

Puede demostrarse, véase Myers (1971), que $\partial Q / \partial Z_1 = 0$, $\partial Q / \partial Z_2 = 0, \dots, \partial Q / \partial Z_p = 0$, cuando Z es el vector:

$$Z_0 = -B^{-1} \beta / 2. \quad (1.6)$$

El vector $Z_0 = (Z_{10}, Z_{20}, \dots, Z_{p0})$, donde las primeras derivadas de Q con respecto a Z_1, Z_2, \dots, Z_p se anulan, recibe el nombre de punto estacionario. Un poco de álgebra puede demostrar que el valor de Q cuando $Z = Z_0$, es:

$$Q_0 = \beta_0 + Z_0' \beta / 2. \quad (1.7)$$

Al introducir la transformación $Z = v + Z_0$ en (1.5), es decir, al mover el origen de coordenadas al punto estacionario, entonces Q toma la forma:

$$Q = Q_0 + v' B v.$$

Por otra parte, al introducir la relación $v = M w$, donde M es tal que: $M' B M = \Lambda$, es una matriz diagonal, con las raíces características de B, $\lambda_1, \lambda_2, \dots, \lambda_p$, en la diagonal principal, entonces Q puede expresarse como sigue:

$$Q = Q_0 + w' \Lambda w = Q_0 + \sum_{i=1}^p \lambda_i w_i^2.$$

Claramente, Q_0 es un máximo cuando todas las λ_i son negativas y un mínimo cuando todas las λ_i son positivas; si hay raíces características de B de ambos signos, entonces Q_0 no es un máximo ni un mínimo. (Myers 1971).

Ejemplo 1.7

Demuestre que el polinomio cuadrático:

$$f(x_1, x_2) = 8x_1x_2 + 3x_2^2, \quad (x_1, x_2) \in \mathbb{R}^2$$

no tiene puntos extremos.

Solución:

Con la notación introducida, claramente:

$$\beta = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & 4 \\ 4 & 3 \end{bmatrix}.$$

De aquí, las coordenadas del punto estacionario, son:

$$Z_0 = -B^{-1}\beta/2 = -\begin{bmatrix} 0 & 4 \\ 4 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \end{bmatrix} / 2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix};$$

y puesto que las raíces características de B, satisfacen la ecuación característica $|B - \lambda I| = 0$, entonces:

$$\left| \begin{bmatrix} 0 & 4 \\ 4 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = \begin{vmatrix} (0-\lambda) & 4 \\ 4 & (3-\lambda) \end{vmatrix} = 0.$$

Así se tiene:

$$\lambda^2 - 3\lambda + 16 = 0,$$

y las raíces características de B, son: $\lambda_1 = 5.772$ y $\lambda_2 = -2.77$. Estas difieren en signo, condición que implica que f no tiene máximo ni mínimo en \mathbb{R}^2 , ocurriendo un punto de silla en:

$$Z_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

C. MAXIMOS Y MINIMOS RESTRICTOS
METODO DE LOS MULTIPLICADORES DE LANGRANGE.

Sea $f(\underline{x}) = f(x_1, x_2, \dots, x_p)$ una función real definida en \mathbb{R}^p y considerese el problema de encontrar sus puntos extremos cuando se satisfacen ciertas restricciones en x_1, x_2, \dots, x_p .

Definase el vector $\underline{g}(\underline{x})$ como sigue:

$$\underline{g}(\underline{x}) = \left[\theta_1(\underline{x}), \theta_2(\underline{x}), \dots, \theta_m(\underline{x}) \right]'$$

donde $\theta_j(\underline{x}) = \theta_j(x_1, x_2, \dots, x_p)$, $j=1, 2, \dots, m$. El problema de encontrar los extremos de f , sujeta a las restricciones $\theta_1(\underline{x}) = 0$, $\theta_2(\underline{x}) = 0, \dots, \theta_m(\underline{x}) = 0$, puede resolverse con cierta facilidad, introduciendo el vector de constantes $\underline{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_p]$, conocido como el vector de "multiplicadores de Lagrange".

La ecuación $\frac{\partial f}{\partial \underline{x}} - \underline{\lambda} \frac{\partial \underline{g}}{\partial \underline{x}} = \underline{\phi}'$, un vector $1 \times p$ de ceros, satisface las condiciones necesarias para puntos estacionarios. Tomando las derivadas parciales con respecto a todas las x_i , en la ecuación anterior, se tiene:

$$\frac{\partial}{\partial x_j} (f - \underline{\lambda} \underline{g}) = 0, \quad j=1, 2, \dots, p,$$

donde las ecuaciones resultantes, junto con las ecuaciones de restricción $\underline{g} = \underline{\phi}$, con $\underline{\phi}$ un vector $p \times 1$ de ceros, proporcionan los valores factibles de \underline{x} y $\underline{\lambda}$ que satisfacen las condiciones necesarias para puntos extremos. Lo anterior, así definido, es el método conocido como método de Lagrange; el método identifica los puntos estacionarios de problemas de optimización con restricciones de igualdad.

Sea:

$$H^0 = \begin{bmatrix} O & P \\ P' & Q \end{bmatrix}_{(m+p) \times (m+p)}$$

donde O , es una matriz $m \times m$ de ceros;

$$P = \begin{bmatrix} \nabla \theta_1(x) \\ \vdots \\ \nabla \theta_m(x) \end{bmatrix}_{m \times n}$$

con $\nabla \theta_i(x)$, siendo el vector hilera de las primeras derivadas de θ_i , con respecto a x_1, x_2, \dots, x_n ; y:

$$Q = \left\| \frac{\partial^2 L(x, \lambda)}{\partial x_i \partial x_j} \right\|_{p \times p}, \text{ para toda } i \text{ y } j$$

siendo $L(x, \lambda) = f(x) - \lambda g(x)$, la función de Lagrange. Q es entonces la matriz de segundas derivadas de L .

La matriz H^B es la llamada matriz hessiana de la frontera.

Dado el punto estacionario (x_0, λ_0) para la función de Lagrange $L(x, \lambda)$ y la matriz hessiana en la frontera H^B evaluada en (x_0, λ_0) , entonces x_0 será un punto de:

(i) Máximo si, comenzando con el determinante menor principal de orden $(2m+1)$, los últimos $(p-m)$ determinantes menores principales de H^B , forman una configuración de signos alternos, comenzando con el de $(-1)^{m+1}$.

(ii) Mínimo si, comenzando con el determinante menor principal de orden $(2m+1)$, los últimos $(p-m)$ determinantes menores principales de H^B tienen el signo de $(-1)^m$.

Con los resultados anteriores, se tiene que: un punto estacionario puede ser un punto extremo sin satisfacer las condiciones anteriores; es decir, estas condiciones son suficientes, pero no necesarias (Taha, 1977).

Existen otras condiciones que son tanto necesarias como suficientes. Defínase la matriz:

$$\Delta = \begin{bmatrix} 0 & P \\ P' & Q - \mu I \end{bmatrix}$$

evaluada en el punto estacionario (x_0, λ_0) . P y Q se definen como

antes; μ es un parámetro desconocido.

Al considerar el determinante $|\Delta|$, cada una de las $(n-m)$ raíces μ_i del polinomio $|\Delta| = 0$, deben ser:

- i) Positivas, si \underline{x}_0 es un punto mínimo.
- ii) Negativas, si \underline{x}_0 es un punto máximo.

Ejemplo 1.8.

Considere el problema de minimizar:

$$f(\underline{x}) = x_1^2 + x_2^2 + x_3^2, \text{ sujeta a las restricciones:}$$

$$g_1(\underline{x}) = x_1 + x_2 + 3x_3 - 2 = 0,$$

y

$$g_2(\underline{x}) = 5x_1 + 2x_2 + x_3 - 5 = 0.$$

Solución:

La función de Lagrange es:

$$L(\underline{x}, \underline{\lambda}) = x_1^2 + x_2^2 + x_3^2 - \lambda_1(x_1 + x_2 + 3x_3 - 2) - \lambda_2(5x_1 + 2x_2 + x_3 - 5).$$

Derivando se obtienen las siguientes condiciones necesarias, para la ocurrencia del punto estacionario:

$$\frac{\partial L}{\partial x_1} = 2x_1 - \lambda_1 - 5\lambda_2 = 0,$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - \lambda_1 - 2\lambda_2 = 0,$$

$$\frac{\partial L}{\partial x_3} = 2x_3 - 3\lambda_1 - \lambda_2 = 0,$$

$$\frac{\partial L}{\partial \lambda_1} = -(x_1 + x_2 + 3x_3 - 2) = 0,$$

$$\frac{\partial L}{\partial \lambda_2} = -(5x_1 + 2x_2 + x_3 - 5) = 0.$$

La solución de este sistema de ecuaciones simultáneas produce:

$$\underline{x}_0 = (x_{10}, x_{20}, x_{30}) = (0.81, 0.35, 0.28),$$

$$\underline{\lambda} = (\lambda_{10}, \lambda_{20}) = (0.0887, 0.3067).$$

H^B es la matriz:

$$H^B = \left(\begin{array}{cc|cc} 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 5 & 2 & 1 \\ \hline 1 & 5 & 2 & 0 & 0 \\ 1 & 2 & 0 & 2 & 0 \\ 3 & 1 & 0 & 0 & 2 \end{array} \right)$$

Puesto que en este problema $m=2$, $z_{m+1} = z_{2+1} = 5$; debe comprobarse que los últimos $p-m = 3-2 = 1$, determinantes menores principales de H^B , tienen el signo de $(-1)^m = (-1)^2 = +1$. Es decir, en este caso es suficiente probar que $|H^B|$, es un número positivo.

De aquí:

$$|H^B| = 460 > 0,$$

lo cual implica que en $\underline{x}_0 = (0.81, 0.35, 0.28)$, ocurre un mínimo.

CAPITULO II

EXTREMOS DE FUNCIONES POR METODOS NUMERICOS.

El cálculo de extremos de una función no lineal sin restricciones, en una o más variables puede ser, en muchos casos, difícil.

Los métodos numéricos que se describirán en este capítulo, juegan un papel importante en la optimización de dichos problemas. Además, para una mejor comprensión de la teoría desarrollada para los métodos, se presentarán varios ejemplos ilustrativos.

Cabe mencionar que, en la última parte de este capítulo, se expone sin mucho detalle, un procedimiento de optimización multivariable con restricciones.

A. EXTREMOS DE FUNCIONES EN UNA VARIABLE.

Un problema no lineal sin restricciones en una variable, tiene la forma:

$$\text{optimicése } z = f(x),$$

donde $f(x)$ es una función (no lineal) de la única variable " x " y la búsqueda del valor óptimo (máximo o mínimo) se realiza sobre todo el conjunto de números reales.

En la práctica, rara vez resulta fructífera la localización de óptimos empleando el cálculo, ya sea porque se desconozca analíticamente la función objetivo, lo cual hace imposible la diferenciación, o bien, que los puntos estacionarios no puedan obtenerse algebraicamente. En tales casos, se emplean métodos numéricos para aproximar la localización de algunos de los óptimos locales dentro de una tolerancia aceptable (Bronson 1984).

Estos métodos que son técnicas de búsqueda secuencial, empiezan con un intervalo finito, sobre el cual, se considera que la función objetivo es unimodal; es decir; se considera que el intervalo incluye uno y sólo un punto en el cual $f(x)$ tiene un máximo o un mínimo local.

Las técnicas reducen entonces, sistemáticamente, el intervalo alrededor del óptimo local, hasta que éste queda confinado dentro de los límites aceptables. Esta reducción se

efectúa evaluando secuencialmente la función objetivo en puntos seleccionados, y empleando después la propiedad unimodal, para eliminar porciones del último intervalo.

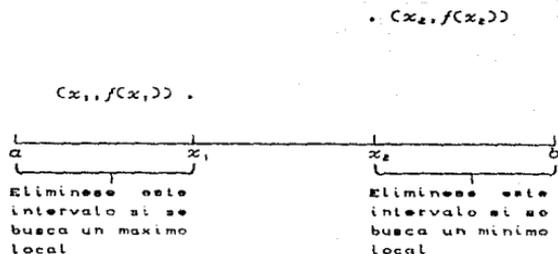


Figura 2.1.

Considérese, por ejemplo, la Figura 2.1 que muestra los valores de la función objetivo en los puntos x_1 y x_2 . Si se sabe que un mínimo local es el único valor extremo en $[a, b]$ entonces este mínimo debe estar a la izquierda de x_2 , ya que $f(x)$ ha empezado a incrementarse a partir de este punto y, por la propiedad unimodal, debe continuar incrementándose a la derecha de él. Por lo tanto, el subintervalo $(x_2, b]$ puede descartarse.

Si un máximo local es el único valor extremo en $[a, b]$, entonces debe estar a la derecha de x_1 , y el subintervalo $(a, x_1]$ puede descartarse. (Bronson 1984).

Hay diversas técnicas de búsqueda secuencial, pero en esta investigación se hará referencia a una, en la siguiente sección. Véase Bronson (1984).

1. Búsqueda en 3 puntos del intervalo.

El intervalo a considerar se divide en 4 partes y se evalúa la función objetivo para los 3 puntos equidistantes interiores. Enseguida se determina el punto interior que produzca el mejor valor de la función objetivo (en caso de empate, selecciónese

arbitrariamente uno de los puntos), y el subintervalo con centro en ese punto, formado por dos cuartas partes del intervalo inicial, reemplaza al propio intervalo inicial.

La búsqueda en 3 puntos del intervalo es el procedimiento más eficiente de búsqueda a espacios iguales, cuando se desea lograr una tolerancia previamente dada, con un número mínimo de evaluaciones funcionales. También es una de las búsquedas secuenciales más fáciles de programar en computadora.

2. Funciones convexas.

Los procedimientos de búsqueda garantizan la aproximación a los óptimos globales en un intervalo de búsqueda, sólo cuando la función objetivo es unimodal allí. Generalmente, en la práctica no se sabe si una función objetivo particular es unimodal sobre un intervalo dado. Por lo tanto, cuando se aplica un procedimiento de búsqueda bajo esta situación, no existe la seguridad de que este procedimiento descubra el óptimo deseado, (Bronson 1984).

Una función $f(x)$ es convexa en el intervalo φ (finito o infinito), si para 2 puntos cualesquiera x_1, x_2 en φ y para toda α en $(0,1)$:

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha)f(x_2) \quad (2.1)$$

Si (2.1) se cumple con la desigualdad inversa, entonces $f(x)$ es cóncava. Así, una función que no es convexa, es cóncava y viceversa. Las funciones cóncavas y convexas son unimodales.

Teorema 2.1.

Si $f(x)$ tiene derivada de segundo orden en φ , entonces $f(x)$ es convexa en φ , sólo si $f''(x) \geq 0$ para toda x en φ . $f(x)$ es cóncava sólo si $f''(x) \leq 0$ para toda x en φ .

Teorema 2.2.

Si $f(x)$ es convexa en φ , entonces cualquier mínimo local en φ es mínimo global en φ . Si $f(x)$ es cóncava en φ , entonces cualquier máximo local en φ es un máximo global en φ .

Si (2.1) se cumple con la desigualdad estricta, excepto en $\alpha=0$ y $\alpha=1$, la función es estrictamente convexa. Tal función tiene una segunda derivada estrictamente positiva y cualquier mínimo local (y por lo tanto global) es único.

Ejemplo 2.1.

Usese la búsqueda en 3 puntos del intervalo para aproximar la localización del máximo de $f(x) = x(5\pi - x)$ en $[0, 20]$ con una aproximación $\epsilon=1$.

Solución:

Ya que $f''(x) = -2 < 0$ en cualquier punto, se tiene del Teorema 2.1 que $f(x)$ es cóncava y, por lo tanto, unimodal, en $[0, 20]$. Se tiene entonces la seguridad de que la búsqueda en 3 puntos del intervalo será convergente hacia el máximo global. Localizaremos el máximo por aproximaciones sucesivas:

Primera iteración:

Dividiendo $[0, 20]$ en 4 partes iguales, se tiene que $x_1 = 5$, $x_2 = 10$ y $x_3 = 15$ son los tres puntos interiores. Entonces:

$$f(x_1) = x_1(5\pi - x_1) = 5(5\pi - 5) = 53.54,$$

$$f(x_2) = x_2(5\pi - x_2) = 10(5\pi - 10) = \underline{57.08},$$

$$f(x_3) = x_3(5\pi - x_3) = 15(5\pi - 15) = 10.62.$$

Ya que x_2 es el punto que genera el mayor valor de la función objetivo, se toma al intervalo $[5, 15]$ con centro en x_2 , como el nuevo intervalo de interés.

Segunda iteración:

Se divide $[5, 15]$ en cuatro partes iguales, con $x_4 = 7.5$, $x_2 = 10$ y $x_3 = 12.5$, como los 3 puntos interiores. Así:

$$f(x_4) = x_4(5\pi - x_4) = (7.5)(5\pi - 7.5) = 61.56,$$

$$f(x_2) = x_2(5\pi - x_2) = 10(5\pi - 10) = 57.08 \text{ (como antes),}$$

$$f(x_3) = x_3(5\pi - x_3) = (12.5)(5\pi - 12.5) = 40.10.$$

Ya que x_4 es el punto interior que da el mayor valor de $f(x)$, se toma el intervalo $[5,10]$ con centro en x_4 , como el nuevo intervalo de interés.

Tercera iteración:

Se divide $[5,10]$ en cuatro partes iguales, con $x_6 = 6.25$, $x_4 = 7.5$ y $x_2 = 8.75$ como los nuevos puntos interiores, entonces:

$$f(x_6) = (6.25)(5\pi - 6.25) = 59.11,$$

$$f(x_4) = 61.56 \text{ como antes,}$$

$$f(x_2) = (8.75)(5\pi - 8.75) = 60.88;$$

como x_4 da el mayor valor de $f(x)$, se toma el intervalo $[6.25, 8.75]$ con centro x_4 , como el nuevo intervalo de interés.

Cuarta iteración:

Dividiendo $[6.25, 8.75]$ en cuatro partes iguales, se generan $x_8 = 6.875$, $x_4 = 7.5$ y $x_0 = 8.125$ como nuevos puntos interiores, entonces:

$$f(x_8) = (6.875)(5\pi - 6.875) = 60.73,$$

$$f(x_4) = 61.56 \text{ (como antes),}$$

$$f(x_0) = (8.125)(5\pi - 8.125) = 61.61.$$

Se tiene ahora que x_0 es el punto interior que da el mayor valor de la función objetivo; se toma por lo tanto el subintervalo $[7.5, 8.75]$ con centro en x_0 , como el nuevo intervalo de interés; pero el punto medio de este intervalo, está dentro de la tolerancia previamente establecida, $\epsilon = 1$, de aquí que se tome como punto óptimo :

$$x^* = x_0 = 8.125 \quad \text{con } z^* = f(x_0) = 61.61.$$

B. EXTREMOS DE FUNCIONES MULTIVARIADAS SIN RESTRICCIONES.

En esta sección se examinará brevemente, el problema de encontrar extremos de funciones multivariadas, sin restricciones.

1. Método del ascenso acelerado.

Se examina la función, dándole valores arbitrarios (con una cierta secuencia), empleando cualquier información previa acerca de la posible ubicación del máximo.

Sea entonces $\underline{x}_0 = (x_{10}, x_{20}, \dots, x_{p0})$ el vector inicial (de acuerdo a los valores arbitrarios dados, el punto que maximice a la función, será nuestro vector inicial). Determinense después los vectores $\underline{x}_1, \underline{x}_2, \underline{x}_3, \dots$ por la relación iterativa:

$$\underline{x}_{k+1} = \underline{x}_k + \lambda_k^* \nabla f \Big|_{\underline{x}_k}$$

Aquí, λ_k^* es un escalar positivo que maximiza $f \left[\underline{x}_k + \lambda \nabla f \Big|_{\underline{x}_k} \right]$.

Es mejor si λ_k^* representa un máximo global; sin embargo, un máximo local sirve. El proceso iterativo termina cuando la diferencia entre los valores de la función objetivo para dos vectores sucesivos, es menor que la tolerancia previamente establecida.

El último vector \underline{x} calculado se vuelve la aproximación final para \underline{x}^* .

Cabe mencionar, en cuanto al valor de la tolerancia, que entre más pequeña sea, se tendrá un valor más aproximado del valor máximo (Bronson 1984).

Ejemplo 2.2.

Usese el método del ascenso acelerado para resolver el siguiente problema:

$$\text{Maximícese } f = -(x_1 - \sqrt{5})^2 - (x_2 - \pi)^2 - 10.$$

Solución:

Al realizar un muestreo sobre varios puntos de la función objetivo sobre la región $-10 \leq x_1, x_2 \leq 10$, se obtiene el valor máximo para J , -18.45 y ocurre en $\underline{x}_0 = [0, 5]^T$; se toma como la aproximación inicial para \underline{x} .

El gradiente de la función objetivo es:

$$\nabla f = \begin{bmatrix} -2cx_1 - \sqrt{5} \\ -2cx_2 - \pi \end{bmatrix}$$

Primera iteración:

$$\underline{x}_0 + \lambda \nabla f \Big|_{\underline{x}_0} = \begin{bmatrix} 0 \\ 5 \end{bmatrix} + \begin{bmatrix} -2(-\sqrt{5}) \\ -2(5 - 3.1416) \end{bmatrix};$$

De aquí:

$$\begin{aligned} f \left(\underline{x}_0 + \lambda \nabla f \Big|_{\underline{x}_0} \right) &= -(4.4721\lambda - 2.2361)^2 - (5 - 3.7168\lambda - 3.1416)^2 - 10 \\ &= -33.74 \lambda^2 + 33.74 \lambda - 18.42 \end{aligned}$$

Se tiene que esta función de λ asume un máximo global en $\lambda_0 = 0.5$, puesto que:

$$\begin{aligned} f(\lambda) &= -33.74 \lambda^2 + 33.74 \lambda - 18.42, \\ f'(\lambda) &= 67.48\lambda + 33.74. \end{aligned}$$

$$\lambda = \frac{33.74}{67.48} = 0.5.$$

Ahora considérese:

$$\underline{x}_1 = \underline{x}_0 + \lambda \nabla f \Big|_{\underline{x}_0} = \begin{bmatrix} 6.597 - 8.722(0.5) \\ 5.69 - 6.499(0.5) \end{bmatrix} = \begin{bmatrix} 2.236 \\ 3.142 \end{bmatrix}$$

con $f(\underline{x}_1) = -10.00$. Ya que la diferencia entre $f(\underline{x}_0) = 18.45$ y $f(\underline{x}_1) = -10.00$ es significativa, se continúa con las iteraciones.

Segunda iteración:

$$\begin{aligned} \underline{x}_1 + \lambda \nabla f \Big|_{\underline{x}_1} &= \begin{bmatrix} 2.236 \\ 3.142 \end{bmatrix} + \lambda \begin{bmatrix} -2(2.236 - \sqrt{5}) \\ -2(3.142 - \pi) \end{bmatrix} \\ &= \begin{bmatrix} 2.236 + .0001\lambda \\ 3.142 - .008\lambda \end{bmatrix}; \end{aligned}$$

De aquí:

$$\begin{aligned} f \left(\underline{x}_1 + \lambda \nabla f \Big|_{\underline{x}_1} \right) &= -(2.236 + .0001\lambda - \sqrt{5})^2 - (3.142 - .0008\lambda - \pi)^2 \\ &= -(6.500\lambda^2 - 8.382\lambda + 10^9) 10^{-7}. \end{aligned}$$

Tenemos $\lambda_1^* = 0.4909$, y de aquí:

$$\underline{x}_2 = \underline{x}_1 + \lambda_1^* \nabla f \Big|_{\underline{x}_1} = \begin{bmatrix} 2.236 + .0001(0.4909) \\ 3.142 - .0008(0.4909) \end{bmatrix} = \begin{bmatrix} 2.236 \\ 3.142 \end{bmatrix}.$$

Como $\underline{x}_1 = \underline{x}_2$ hasta la aproximación empleada, se acepta

$\underline{x}^* = [2.236, 3.142]^T$, punto en el cual $f^* = -10.00$.

2. Método de Newton-Raphson.

Seleccione un vector inicial \underline{x}_0 , como en el método del ascenso acelerado. Los vectores $\underline{x}_1, \underline{x}_2, \underline{x}_3, \dots$, se determinan iterativamente mediante la relación:

$$\underline{x}_{k+1} = \underline{x}_k - \left[H_f \Big|_{\underline{x}_k} \right]^{-1} \nabla f \Big|_{\underline{x}_k}.$$

La regla para detenerse es la misma que la empleada en el método del ascenso acelerado.

El método Newton-Raphson convergerá a un máximo local, si H_f es negativa definida sobre alguna vecindad ϵ alrededor del máximo y si \underline{x}_0 queda dentro de esta vecindad.

Debe mencionarse que si H_f es negativa definida, H_f^{-1} existe y es negativa definida.

Si no se selecciona correctamente el vector inicial \underline{x}_0 , el método puede no converger en absoluto. En ambos casos, el proceso iterativo se da por terminado y se inicia nuevamente con una mejor aproximación inicial.

Ejemplo 2.3.

i) Usando el método de Newton-Raphson maximícese:

$$z = -(x_1 - \sqrt{5})^2 (x_2 - \pi)^2 - 10,$$

con una tolerancia de 0.05.

Solución:

Considérese el muestreo de la función z en los puntos de la Figura 2.2

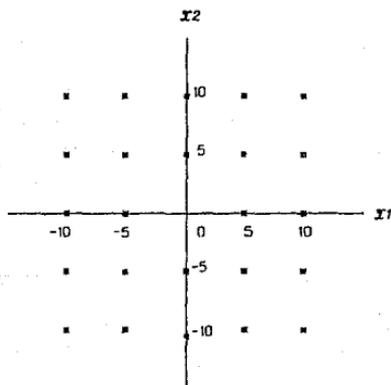


Figura 2.2

En los puntos seleccionados z toma los valores de la Tabla 2.1

x_1	x_2	z
0	0	-24.8000
0	5	-18.4536
0	10	-62.0977
5	0	-27.5089
5	5	-21.0929
5	10	-64.6771
10	0	-90.1482
10	5	-78.7929
10	10	-117.8164
-5	0	-72.2802
-5	5	-65.8148
-5	10	-109.8984
-10	0	-169.5909
-10	5	-163.1750
-10	10	-206.7591
0	-5	-81.2855
0	-10	-187.7014
-5	-5	-128.6462
-5	-10	-235.0621
-10	-5	-226.0068
-10	-10	-332.4228
5	-5	-83.9348
5	10	-190.3407
10	-5	-136.5641
10	-10	-242.9800

Tabla 2.1 Conjunto de valores de z obtenidos al sustituir los puntos seleccionados en la ecuación:

$$z = -(x_1 - \sqrt{5})^2 (x_2 - 11)^2 - 10.$$

Como punto inicial se elige $\underline{x}_0 = [0, 5]$, con $f(\underline{x}_0) = -18.453$.

El vector gradiente, la matriz Hessiana y la inversa de la matriz Hessiana para esta función objetivo, son, respectivamente:

$$\nabla f = \begin{bmatrix} -2(x_1 - \sqrt{5}) \\ -2(x_2 - 11) \end{bmatrix}, \quad H_f = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}, \quad H_f^{-1} = \begin{bmatrix} -0.5 & 0 \\ 0 & -0.5 \end{bmatrix}.$$

para toda x_1 y x_2 .

Primera iteración

Se tiene:

$$\nabla f|_{\underline{x}_0} = \begin{bmatrix} -2(0 - \sqrt{5}) \\ -2(5 - \pi) \end{bmatrix} = \begin{bmatrix} 4.472 \\ -3.716 \end{bmatrix},$$

entonces:

$$\begin{aligned} \underline{x}_1 &= \underline{x}_0 - \left(H_f|_{\underline{x}_0} \right)^{-1} \nabla f|_{\underline{x}_0} \\ &= \begin{bmatrix} 0 \\ 5 \end{bmatrix} - \begin{bmatrix} -0.5 & 0 \\ 0 & -0.5 \end{bmatrix} \begin{bmatrix} 4.472 \\ -3.716988 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix} - \begin{bmatrix} -2.236 \\ 1.858 \end{bmatrix} = \begin{bmatrix} 2.236 \\ 3.142 \end{bmatrix}, \end{aligned}$$

substituyendo \underline{x}_1 en z , resulta:

$$z = f(\underline{x}_1) = -10.00.$$

Como se obtiene:

$$f(\underline{x}_1) - f(\underline{x}_0) = -10.00 - (-18.453) = 8.453 > 0.5,$$

se continúan las iteraciones.

Segunda iteración

Se tiene sucesivamente:

$$\nabla f|_{\underline{x}_1} = \begin{bmatrix} -2(2.236 - \sqrt{5}) \\ -2(3.142 - \pi) \end{bmatrix} = \begin{bmatrix} -0.0001 \\ 0.0008 \end{bmatrix},$$

$$\underline{x}_2 = \underline{x}_1 - \left(H_f|_{\underline{x}_1} \right)^{-1} \nabla f|_{\underline{x}_1},$$

$$= \begin{bmatrix} 2.236 \\ 3.142 \end{bmatrix} - \begin{bmatrix} -0.5 & 0 \\ 0 & -0.5 \end{bmatrix} \begin{bmatrix} -0.0001 \\ 0.0008 \end{bmatrix} = \begin{bmatrix} 2.236 \\ 3.142 \end{bmatrix}.$$

Substituyendo \underline{x}_2 en z , se tiene:

$$z = f(\underline{x}_2) = -10.00.$$

Puesto que $f(\underline{x}_2) - f(\underline{x}_1) = 0 < 0.05$, se toma entonces:

$$\underline{x}^* = \underline{x}_2 = [2.236, 3.142]', \text{ con } z^* = f(\underline{x}_2) = -10.00.$$

(i) Usando el método de Newton-Raphson maximícese
 $z = -\text{sen}x_1x_2 + \text{cos}(x_1 - x_2)$
 con una tolerancia de 0.05.

Solución:

Considérese el muestreo de la función z en los puntos de la
 Figura 2.3.

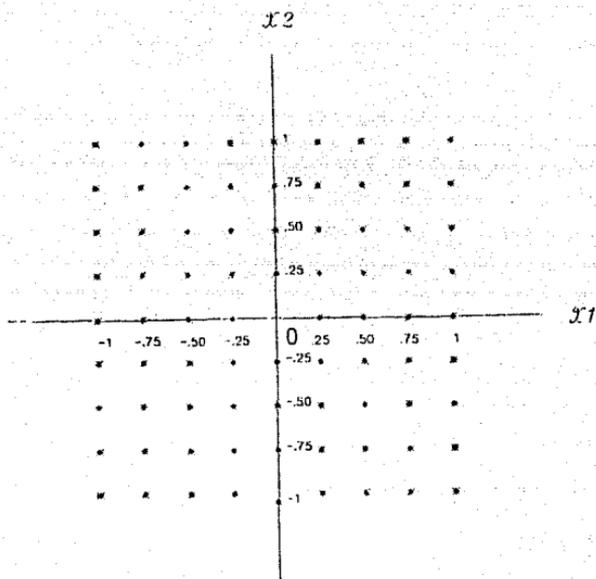


Figura 2.3

En los puntos seleccionados z toma los valores de la Tabla 2.2.

x_1	x_2	z
- 1	1	1.0168
- .75	1	1.0126
- .50	1	1.0083
- .25	1	1.0041
0	1	.9998
.25	1	.9955
.50	1	.9912
.75	1	.9869
1	1	.9825
- 1	.75	1.0126
- .75	.75	1.0094
- .50	.75	1.0063
- .25	.75	1.0031
0	.75	.9999
.25	.75	.9966
.50	.75	.9934
.75	.75	.9901
1	.75	.9864
- 1	.50	1.0083
- .75	.50	1.0063
- .50	.50	1.0042
- .25	.50	1.0020
0	.50	.9999
.25	.50	.9978
.50	.50	.9956
.75	.50	.9934
1	.50	.9912
- 1	.25	1.0041
- .75	.25	1.0031
- .50	.25	1.0020
- .25	.25	1.0010
0	.25	.9999
.25	.25	.9989
.50	.25	.9978
.75	.25	.9966
1	.25	.9955
- 1	0	.9998
- .75	0	.9999
- .50	0	.9999
- .25	0	.9999
0	0	1
.25	0	.9999
.50	0	.9999
.75	0	.9999
1	0	.9998

Tabla 2.2. Conjunto de valores de z obtenidos al sustituir los puntos seleccionados en la ecuación:
 $z = \sin x_1 + \cos(x_1 - x_2)$.

continuación....

- 1	-. 25	. 0055
-. 75	-. 25	. 0066
-. 50	-. 25	. 0078
-. 25	-. 25	. 0089
0	-. 25	. 0099
. 25	-. 25	1. 0010
. 50	-. 25	1. 0020
. 75	-. 25	1. 0031
1	-. 25	1. 0041
- 1	-. 50	. 0012
-. 75	-. 50	. 0034
-. 50	-. 50	. 0056
-. 25	-. 50	. 0078
0	-. 50	. 0099
. 25	-. 50	1. 0020
. 50	-. 50	1. 0042
. 75	-. 50	1. 0063
1	-. 50	1. 0089
- 1	-. 75	. 0069
-. 75	-. 75	. 0061
-. 50	-. 75	. 0034
-. 25	-. 75	. 0066
0	-. 75	. 0099
. 25	-. 75	1. 0031
. 50	-. 75	1. 0063
. 75	-. 75	1. 0094
1	-. 75	1. 0126
- 1	- 1	. 0025
-. 75	- 1	. 0069
-. 50	- 1	. 0012
-. 25	- 1	. 0055
0	- 1	. 0099
. 25	- 1	1. 0041
. 50	- 1	1. 0027
. 75	- 1	1. 0026
1	- 1	1. 0188

Tabla 2.2. Conjunto de valores de z obtenidos al sustituir los puntos seleccionados en la ecuación:
 $z = -\text{sen } x_1 x_2 + \text{cos}(x_1 - x_2)$.

El vector gradiente y la matriz hessiana para esta función objetivo son, respectivamente:

$$\nabla f = \begin{bmatrix} -x_2 \cos x_1 x_2 - \text{sen}(x_1 - x_2) \\ -x_1 \cos x_1 x_2 - \text{sen}(x_1 - x_2) \end{bmatrix},$$

$$H_f =$$

$$\begin{bmatrix} x_2^2 \operatorname{sen} x_1 x_2 - \cos(x_1 - x_2) & -\cos x_1 x_2 + x_1 x_2 \operatorname{sen} x_1 x_2 + \cos(x_1 - x_2) \\ -\cos x_1 x_2 - x_1 x_2 \operatorname{sen} x_1 x_2 + \cos(x_1 - x_2) & x_1^2 \operatorname{sen} x_1 x_2 - \cos(x_1 - x_2) \end{bmatrix}$$

Se elige $x_0 = (0,0)$ como nuestro valor inicial.

Primera iteración:

Se tiene:

$$\nabla f|_{x_0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad H_f|_{x_0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad (H_f|_{x_0})^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix};$$

entonces:

$$x_1 = x_0 - (H_f|_{x_0})^{-1} \nabla f|_{x_0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

con $f(x_1) = 1$, y como $f(x_1) - f(x_0) = 1.000 - 1 = 0 < 0.5$, setoma:
 $x^* = x_1 = [0,0]^T$ y $z^* = f(x_1) = 1$.

Se elige ahora a $x_0 = (-1,1)$ como nuestro valor inicial.

Segunda iteración:

$$\nabla f|_{x_0} = \begin{bmatrix} 1.0346 \\ 0.985 \end{bmatrix}, \quad H_f|_{x_0} = \begin{bmatrix} -1.0167 & 0.0169 \\ 0.0169 & -1.0167 \end{bmatrix},$$

$$(H_f|_{x_0})^{-1} = \begin{bmatrix} -.9837 & -.0163 \\ -.0163 & -.9837 \end{bmatrix};$$

entonces:

$$\begin{aligned} x_1 &= x_0 - (H_f|_{x_0})^{-1} \nabla f|_{x_0} \\ &= \begin{bmatrix} -1 \\ 1 \end{bmatrix} - \begin{bmatrix} -.9837 & -.0163 \\ -.0163 & -.9837 \end{bmatrix} \begin{bmatrix} 1.0346 \\ 0.985 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ -1 \end{bmatrix} - \begin{bmatrix} -1.033 \\ -0.986 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned}$$

con $f(x_1) = 1$. Como: $f(x_1) - f(x_0) = 1.000 - 0.9977$
 $= 0.0023 < 0.05$, se toma:

$$x^* = x_1 = (0,0)' \quad \text{y} \quad z^* = f(x_1) = 1.$$

3. Programas SAS para la aplicación del método de Newton-Raphson

El procedimiento IML de SAS para microcomputadoras, ofrece alternativas de programación de rutinas iterativas de cómputo. Como un ejemplo, considérese el programa de la Tabla 2.3, el cual aplica el método de Newton-Raphson para encontrar el máximo de la función:

$$z = -\text{sen}x_1 x_2 + \text{cos}(x_1 - x_2).$$

Esta función fue evaluada manualmente, como se presenta en el ejemplo 2.3. El programa requiere la alimentación de expresiones SAS, muy parecidas a las instrucciones de los lenguajes BASIC, para la función objetivo, la matriz Hessiana y el gradiente.

Para el uso del procedimiento IML, el lector puede consultar los manuales SAS, del SAS Institute (1985).

En el programa de la Tabla 2.3, $F(x)$ es el gradiente de la función objetivo y $J(x)$ es la matriz Hessiana.

Abajo se presenta el resultado que produce la computadora:

EL PUNTO MAXIMO ES:

X
2.236068
3.141516

Se observa que este resultado concuerda con el punto óptimo obtenido en el ejemplo 2.3.

Tabla 2.3. Programa SAS para la aplicación del Método de Newton-Raphson

```

START NEWTON;
RUN FUN;
DO ITER=1 TO MAXITER
  WHILE (MAX(ABS(F)) > CONVERGE);
  RUN DERIV;
  DELTA=-SOLVE(J,F)
  X=X+DELTA
  RUN FUN;
  END;
FINISH;
MAXITER=2;
CONVERGE=0.05;
START FUN;
X1=X|1; X2=X|2;
F=(-2*(X1-(SQRT 5))//
  (-2*(X2-1));
FINISH;
START DERIV;
J=(-2 || 0)//
  (0 || -2);
FINISH
DO;
PRINT "EL PUNTO MAXIMO ES:";
X=(0,5);
RUN NEWTON;
PRINT X;
END;

```

**C. OPTIMIZACION MULTIVARIABLE CON RESTRICCIONES
METODO DE NEWTON-RAPHSON**

Sea \underline{x} el vector: $\underline{x} = [x_1, x_2, \dots, x_n]'$. La forma estándar de un problema para el cálculo de extremos de una función no lineal con restricciones de igualdad es:

$$\text{Maximizar : } z = f(\underline{x}),$$

sujeta a las restricciones:

$$\begin{aligned} g_1(\underline{x}) &= 0, \\ g_2(\underline{x}) &= 0, \\ &\vdots \\ g_m(\underline{x}) &= 0, \end{aligned} \quad (2.2)$$

con $m < n$ (menos restricciones que variables).

Para resolver el problema, se forma la función de Lagrange:

$$L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m) = f(\underline{x}) - \sum_{i=1}^m \lambda_i g_i(\underline{x}). \quad (2.3)$$

Al tomar derivadas de L , se obtiene el siguiente sistema de ecuaciones:

$$\begin{aligned} \frac{\partial L}{\partial x_j} &= 0, & j &= 1, 2, \dots, n, \\ \frac{\partial L}{\partial \lambda_i} &= 0, & i &= 1, 2, \dots, m; \end{aligned} \quad (2.4)$$

el cual debe resolverse simultáneamente.

Describiremos enseguida el método de Newton-Raphson para la solución del sistema (2.4).

Por hipótesis:

$$L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m) \equiv L(\underline{z}),$$

es un sistema no lineal de ecuaciones, el cual resulta generalmente imposible de resolver analíticamente. Sin embargo, como las soluciones de (2.4), son los puntos estacionarios de $L(\underline{z})$, y dado que los extremos de $L(\underline{z})$ son algunos de los puntos

estacionarios, deberá ser posible lograr aproximarse al extremo derecho de $L(z)$, que corresponde a la solución óptima del problema original. La forma iterativa de lograrlo se realiza a través de la relación:

$$z_{k+1} = z_k - \left[H_L | z_k \right]^{-1} \nabla | z_k .$$

Este enfoque tiene valor limitado porque es muy difícil determinar un z_0 adecuado.

Para un z_0 incorrecto, el método de Newton-Raphson puede diverger o converger en el extremo "erróneo" de $L(z)$. También es posible que el método converga cuando no existe la solución óptima (Bronson 1984).

CAPITULO III

AJUSTE DE FUNCIONES DE RESPUESTA.

En la determinación de óptimos, uno de los aspectos fundamentales es lograr el ajuste de la función de respuesta propuesta. Esencialmente, se considera el valor de la respuesta explicado como una función continua de los niveles de los factores; así: si η es el nivel de la respuesta, entonces:

$$\eta = f(X_1, X_2, \dots, X_p) \quad (3.1)$$

f puede ser de estructura matemática muy variada y propuesta por el investigador, en tanto que, el fenómeno que estudia a las variables $\eta, X_1, X_2, \dots, X_p$, debe ser, necesariamente, un fenómeno observable.

Ahora, en general, el valor observado $Y_i, i=1, 2, \dots, n$ y lo que pronostica f para los valores $X_{11}, X_{12}, \dots, X_{1p}$, correspondientes a la observación, no coinciden con exactitud en la mayoría de los fenómenos objeto de estudio, de modo que podemos escribir:

$$Y_i - f(X_{11}, X_{12}, \dots, X_{1p}) = e_i, \quad i=1, 2, \dots, n, \quad (3.2)$$

donde e_i , el error de observación, puede considerarse como una variable aleatoria. Usualmente se supone $E(e_i) = 0$, $E(e_i^2) = \sigma^2$, y $E(e_i, e_j) = 0$ si $i \neq j$, siendo σ^2 una constante desconocida, y por el Teorema del Límite Central, los errores e_i , se consideran distribuidos normalmente.

Para aclararlo un poco, del fenómeno en estudio se hacen n observaciones, teniéndose, como resultado, n vectores de la forma:

$$(Y_i, X_{11}, X_{12}, \dots, X_{1p}), \quad i=1, 2, \dots, n; \quad (3.3)$$

con estos valores observados, el siguiente paso es lograr el ajuste de los datos a la función f .

En este capítulo, expondremos brevemente la estructura matemática más común de las funciones de respuesta propuestas por

los investigadores, refiriéndose particularmente, al modelo lineal general y a algunos modelos no lineales de mayor empleo en las aplicaciones, así como también, a los métodos estadísticos apropiados para el ajuste de los modelos propuestos.

A. EL MODELO LINEAL GENERAL

En muchas situaciones f es una función lineal de los parámetros, es decir, f es de la forma:

$$f(X_{11}, X_{12}, \dots, X_{1p}) = \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \dots + \beta_p X_{1p}, \quad (3.4)$$

$i = 1, 2, \dots, n, \quad (n \geq p+1)$

donde $\beta_0, \beta_1, \dots, \beta_p$ son parámetros desconocidos que pueden estimarse a partir de las observaciones. De aquí, si (3.4) se substituye en (3.2), se obtiene:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + e_i, \quad (3.5)$$

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

$$E(e_i) = 0, \quad E(e_i^2) = \sigma^2, \quad E(e_i e_j) = 0 \quad i \neq j.$$

(3.5) es el llamado *modelo general de las hipótesis lineales*, o bien, *modelo lineal general*. De acuerdo con la estructura de este modelo, los términos Y_i son variables aleatorias, pero las cantidades $X_{i1}, X_{i2}, \dots, X_{ip}$ deben ser valores numéricos conocidos; es decir, las X_{ij} , $i = 1, 2, \dots, n$, $j = 1, 2, \dots, p$ son constantes numéricas y no variables aleatorias.

1. Modelos lineales especiales.

El caso más simple ocurre cuando Y_i se aproxima por una función lineal de los niveles de los factores, como ocurre en la expresión (3.5). En otras ocasiones, f es un polinomio de grado entero en uno o más factores; por ejemplo, en el caso de un solo factor, puede ocurrir que f sea un polinomio de grado m en X_{11} , como sigue:

$$f(X_{11}, X_{12}, \dots, X_{1p}) = \beta_0 + \beta_1 X_{11} + \beta_2 X_{11}^2 + \dots + \beta_m X_{11}^m, \quad (3.6)$$

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

Puesto que al ser dado el nivel X_{1i} del factor en estudio, automáticamente es posible calcular los valores $X_{11}^2, \dots, X_{11}^p$, puede verse que haciendo:

$$X_1 = X_{11}, \quad X_1^2 = X_{12}, \dots, \quad X_1^p = X_{1p},$$

al substituir f por el lado derecho de (3.6), en términos de la equivalencia anterior, se obtiene la expresión (3.5) del modelo lineal general. En el caso en que f es un polinomio de segundo grado en los niveles de los factores, se tiene la expresión general:

$$f(X_{11}, X_{12}, \dots, X_{1p}) = \beta_0 + \sum_{j=1}^p \beta_j X_{1j} + \sum_{j=1}^p \beta_j^2 X_{1j}^2 + \dots + \sum_{j=1}^p \sum_{k=1}^p \beta_{jk} X_{1j} X_{1k}, \quad (3.7)$$

nuevamente, puede verse que conocidos los valores X_{1j} , pueden calcularse de inmediato las cantidades X_{1j}^2 y $X_{1j} X_{1k}$, obteniéndose un modelo lineal en los parámetros, que sería un caso particular de (3.5).

Los polinomios con exponentes fraccionarios conocidos sobre los niveles de los factores, también ocurren como un caso particular del modelo lineal general. Considérese, por ejemplo, el modelo:

$$Y_i = \beta_0 + \beta_1 X_{11}^{0.6} + \beta_2 X_{12}^{0.7} + \beta_3 X_{11}^{1.2} + \beta_{22} X_{12}^{1.4} + \beta_{12} X_{11}^{0.6} X_{12}^{0.7} + e_i,$$

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

en este caso, si se dan los niveles X_{11} y X_{12} , automáticamente es posible calcular las cantidades $X_{11}^{0.6}$, $X_{12}^{0.7}$, $X_{11}^{1.2}$, $X_{12}^{1.4}$, $X_{11}^{0.6} X_{12}^{0.7}$, para todas las observaciones, teniéndose de nuevo a f como una función lineal de los parámetros.

2. Modelos Linealizables.

Los modelos linealizables son aquellos que, mediante una transformación, dan lugar a un modelo lineal; entre los más conocidos están los modelos tipo Cobb-Douglas, cuya forma general es la siguiente:

$$Y_i = \alpha Q_{i,1}^{\beta_1} Q_{i,2}^{\beta_2} \dots Q_{i,p}^{\beta_p} \eta_i, \quad i=1,2,\dots,n, \quad (3.8)$$

donde Y_i es la respuesta, $Q_{i,1}, Q_{i,2}, \dots, Q_{i,p}$ son los niveles ensayados u observados de las variables explicativas, $\alpha, \beta_1, \beta_2, \dots, \beta_p$ son los parámetros del modelo, y las η_i son los errores (multiplicativos) del propio modelo. Tomando logaritmos en (3.8), se obtiene:

$$\log(Y_i) = \log(\alpha) + \beta_1 \log(Q_{i,1}) + \beta_2 \log(Q_{i,2}) + \dots + \beta_p \log(Q_{i,p}) + \log(\eta_i), \quad (3.9)$$

que es un modelo lineal en $\beta_0, \beta_1, \beta_2, \dots, \beta_p$, después de hacer $\log(\alpha) = \beta_0$; las $\log(\eta_i)$ serían los errores del modelo. (3.8) es entonces un modelo linealizabile, puesto que, una transformación logarítmica, produce un modelo lineal, aquél dado por (3.9).

B. MODELOS NO LINEALES

Cuando f no es una función lineal de los parámetros, entonces se obtiene un modelo de regresión no lineal, siempre que no exista una transformación que linearice el modelo. Consideraremos en esta sección, algunos modelos de regresión no lineal.

1. El modelo semilogarítmico.

En los fenómenos relacionados con el crecimiento de seres vivos o poblaciones, o bien, en el contexto económico en los fenómenos relacionados con el crecimiento de una empresa o de un capital, si el nivel de la variable dependiente $Y = f(X)$, donde X es la variable explicativa, tiene una velocidad de crecimiento que

es proporcional al valor actual de Y, puede escribirse:

$$\frac{dY}{dX} = \beta Y. \quad (3.10)$$

Separando variables e integrando, se obtiene:

$$\int \frac{dY}{Y} = \int \beta dX,$$

de donde resulta:

$$\ln(Y) = \beta X + C, \quad (3.11)$$

donde Ln denota al logaritmo natural y C es una constante de integración. Tomando exponenciales en ambos miembros de (3.11), se tiene:

$$Y = e^{\beta X + C} = \alpha e^{\beta X}, \quad (3.12)$$

donde se ha escrito $e^C = \alpha$. Al hacer observaciones del fenómeno, Y no resulta exactamente igual al valor observado, ocurriendo un error de observación ϵ (usamos este símbolo para el error, para no confundirlo con e, la base de los logaritmos naturales), pudiendo escribirse:

$$Y = \alpha e^{\beta X} + \epsilon. \quad (3.13)$$

Véase la Figura 3.1.

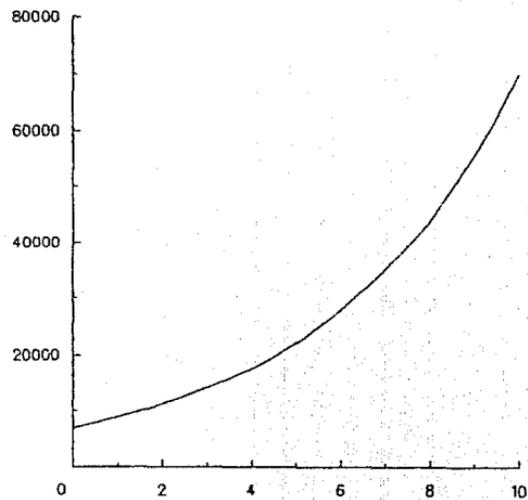
(3.13) es un modelo no lineal en los parámetros α y β , que es imposible de ser linearizado a través de una transformación.

Obsérvese que si en (3.13), el término de error es multiplicativo, se obtendría el modelo:

$$Y = \alpha e^{\beta X} \epsilon, \quad (3.14)$$

el cual podría transformarse a un modelo lineal en los parámetros, mediante una transformación logarítmica. Así, si en (3.14) se toman logaritmos naturales en ambos miembros, se obtiene:

Modelo Semilogaritmico



— $7025 \cdot e^{(0.23 \cdot x)}$

Figura 3.1

$$\ln(Y) = \ln(\alpha) + \beta X + \ln(e). \quad (3.15)$$

(3.15) es un modelo lineal en los parámetros $\ln(\alpha)$ y β . (3.13) y (3.14) no son, entonces, modelos equivalentes. Las diferencias entre estos modelos han sido establecidas en el trabajo de Castillo M. (1978).

2. El modelo logarítmico.

Supóngase que en una función de producción la rapidez de variación de la cantidad de producto Y , con respecto a la cantidad de insumo X , es directamente proporcional a la razón entre producto e insumo. Es decir:

$$\frac{dY}{dX} = \beta \frac{Y}{X}. \quad (3.16)$$

En este modelo X y Y sólo pueden tomar valores positivos, y: (i) si β es positivo, el modelo puede representar el crecimiento acelerado de una empresa, proporcionalmente a la razón Y/X , lo cual incrementa las diferencias con otras empresas; (ii) si β es negativo, el fenómeno se invierte, y la rapidez de variación de la cantidad de producto con respecto al insumo, decrece proporcionalmente a la razón Y/X .

Separando variables e integrando en (3.16), se obtiene:

$$\int \frac{dY}{Y} = \beta \int \frac{dX}{X}; \quad (3.17)$$

de donde resulta:

$$\ln(Y) = \beta \ln(X) + C,$$

donde C es una constante de integración. Tomando exponentiales en (3.17), se tiene:

$$Y = e^{\beta \ln(X) + C} = e^C \left[e^{\ln(X)} \right]^\beta = \alpha X^\beta, \quad (3.18)$$

donde se ha escrito $e^C = a$.

En términos de las observaciones, al introducir el elemento aleatorio de error, (3.18) puede escribirse:

$$Y = \alpha X^\beta + \epsilon. \quad (3.19)$$

Véase la Figura 3.2.

3. La logística.

Cuando la velocidad de crecimiento de Y con respecto a X , es de la forma:

$$\frac{dY}{dX} = kY(A-Y); \quad (3.20)$$

al integrar la relación anterior, se obtiene una expresión con estructura:

$$Y = \frac{A}{1 + ae^{-A\beta X}}, \quad (3.21)$$

donde A es el valor máximo que puede tomar Y . Al referir las observaciones con relación al valor teórico anterior, se obtiene el modelo:

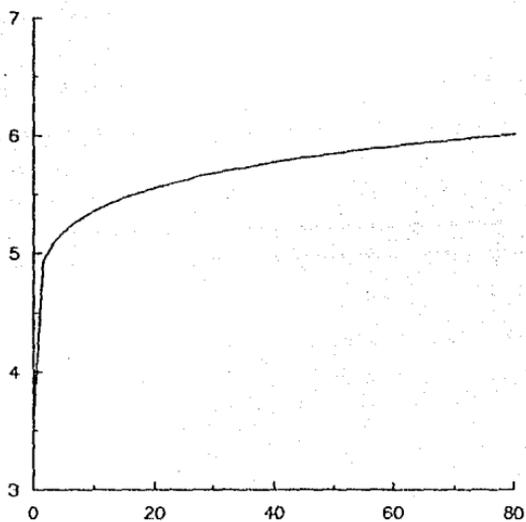
$$Y = \frac{A}{1 + ae^{-A\beta X}} + \epsilon. \quad (3.22)$$

La función (3.21) se conoce como la *logística*. Véase Figura la 3.3.

C. AJUSTE POR MÍNIMOS CUADRADOS

La función de respuesta propuesta, se ajusta por el método de mínimos cuadrados. Cuando la función de respuesta es lineal, de modo que puede establecerse como un caso particular del modelo general lineal, el ajuste por mínimos cuadrados conduce a un sistema de ecuaciones lineales en los estimadores, que se conoce como el "sistema de ecuaciones normales"; estas se resuelven con cierta facilidad y dan lugar a estimadores con propiedades óptimas. En el caso de funciones de respuesta no lineales, el método de mínimos cuadrados se aplica por aproximaciones

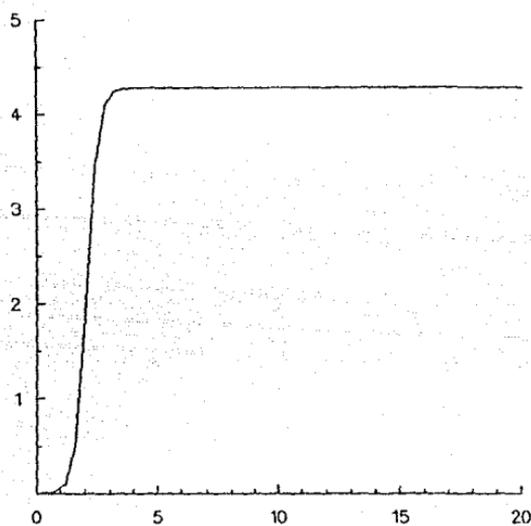
Modelo Logaritmico



$$3.575 + 1.268 \cdot x^{0.14928}$$

Figura 3.2

Modelo Logístico



$$— 4.28/(1+1.7*e^{(-4.28*X)})$$

Figura 3.3

sucesivas. Ampliaremos estas ideas, a lo largo de la presente sección.

1. Ajuste de modelos de regresión lineal.

Considérense las n observaciones que se ajustan al modelo lineal general, escritas en notación matricial:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1p} \\ 1 & X_{21} & X_{22} & \dots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \quad (3.23)$$

donde $E(e_j) = 0$, $E(e_j^2) = \sigma^2$, $E(e_i e_j) = 0$ si $i \neq j$, con $n \geq p+1$. Supondremos, adicionalmente, que los errores e_i se distribuyen normalmente.

Si se escribe:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = Y, \quad \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1p} \\ 1 & X_{21} & X_{22} & \dots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix} = X, \quad (3.24)$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} = \beta, \quad \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = e,$$

entonces (3.23) se reduce a la forma compacta:

$$Y = X\beta + e. \quad (3.25)$$

Supondremos que todos los vectores columna de la matriz X , son linealmente independientes, produciéndose "el modelo general de regresión múltiple". Entonces, la aplicación del método de mínimos cuadrados, requiere encontrar un vector $\hat{\beta}$ que haga mínima, la suma de cuadrados de los errores, $\sum_{i=1}^n e_i^2 = e'e$. Es decir, se trata de encontrar el vector $\beta = \hat{\beta}$, tal que, la expresión:

$$e'e = (Y - X\beta)'(Y - \beta), \quad (3.26)$$

sea un mínimo absoluto. Encontrar tal vector β es un problema matemático, cuya solución conduce al sistema de ecuaciones normales:

$$X'X \hat{\beta} = X'Y, \quad (3.27)$$

sistema que produce la solución $\hat{\beta} = (X'X)^{-1}X'Y$.

El conjunto de elementos en $\hat{\beta}$, son, por el Teorema de Gauss-Markoff, los mejores estimadores lineales insesgados. Son lineales, porque se escriben como una función lineal de las Y_i, e_i ; son insesgados, porque $E(\hat{\beta}_i) = \beta_i, i=0, 1, \dots, p$; y son los mejores estimadores, en el sentido de mínima varianza. Para la demostración del Teorema de Gauss-Markoff, véase Martínez Garza y Castillo M. (1987).

Insertando el vector $\hat{\beta} = (X'X)^{-1}X'Y$, en lugar de β en la expresión (3.26), el mínimo de la suma de cuadrados de los errores, es:

$$\begin{aligned} (Y - X\hat{\beta})'(Y - X\hat{\beta}) &= Y'Y - Y'X\hat{\beta} - \hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta} \\ &= Y'Y - \hat{\beta}'X'Y + (\hat{\beta}'X'X - Y'X)\hat{\beta} \\ &= Y'Y - \hat{\beta}'X'Y, \end{aligned}$$

puesto que $(\hat{\beta}'X'X - Y'X)' = \phi_n$, un vector $n \times 1$ de ceros.

Por las propiedades de los errores $Y'Y - \hat{\beta}'X'Y = SCE$,

dividida entre σ^2 , se distribuye como una variable ji-cuadrada con $n-1-p$ grados de libertad, independientemente de $\hat{\beta}$. Además, si $H_0: \beta_1 = \beta_2 = \dots = \beta_p$, es cierta, entonces $(\hat{\beta}'X'Y - n\bar{Y}^2) / \sigma^2$ se distribuye como una variable ji-cuadrada con p grados de libertad.

2. Análisis de varianza.

Usualmente, en problemas aplicados el investigador se interesa en probar la hipótesis: $H_0 = \beta_1 = \beta_2 = \dots = \beta_p = 0$ vs. H_a : al menos una β_i es distinta de cero. Sin entrar en los detalles del problema, la prueba se realiza construyendo el análisis de varianza de la Tabla 3.1, en la cual los cuadrados medios se calculan dividiendo las sumas de cuadrados entre sus

Tabla 3.1 Análisis de Varianza de la Regresión

Fuentes de Variación	Grados de Libertad	Sumas de Cuadrados	Cuadrados medios	F Calculada
Regresión	p	$\hat{\beta}'X'Y - n\bar{Y}^2$	CMRegr	$\frac{\text{CMRegr}}{S^2}$
Error	$n-1-p$	$Y'Y - \hat{\beta}'X'Y = \text{SCE}$	$\text{CME} = S^2$	
Total	$n-1$	$Y'Y - n\bar{Y}^2$		

correspondientes grados de libertad. La F calculada (F_{calc}) se compara con un valor tabulado, al nivel de significancia elegido, y si ésta es mayor que el valor tabulado, entonces se rechaza H_0 , en favor de H_a .

Los métodos de computación electrónica facilitan enormemente el ajuste de modelos de regresión múltiple. Paquetes estadísticos como el SAS (Statistical Analysis System - Sistema para Análisis Estadístico), o el SPSS (Statistical Package for Social Science - Paquete Estadístico para Ciencias Sociales), tienen programadas las rutinas de cálculo de la técnica de regresión. Ilustraremos el uso del paquete SAS en la siguiente sección, con algunos ejemplos.

3. Ajuste de modelos de regresión no lineal.

En el caso general un modelo no lineal puede representarse simbólicamente, en la forma:

$$Y_i = h(\beta_1, \beta_2, \dots, \beta_p; X_{i1}, X_{i2}, \dots, X_{iq}) + e_i, \quad (3.28)$$

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

donde $\beta_1, \beta_2, \dots, \beta_p$, son los parámetros del modelo, y $X_{i1}, X_{i2}, \dots, X_{iq}$, son números reales conocidos.

El método de Gauss-Newton para el ajuste de un modelo no lineal, por mínimos cuadrados, utiliza el Teorema de Taylor; véase Widdor (1961). Emplea un conjunto de valores arbitrarios iniciales de los parámetros, tomando de la expansión de la función h en serie de Taylor alrededor de los valores anteriores, los primeros términos. Se obtiene así, un modelo lineal que permite derivar una primera estimación de los parámetros. Escríbase:

$$\hat{\beta}_1 = \beta_{10} + \Delta\beta_1, \quad \hat{\beta}_2 = \beta_{20} + \Delta\beta_2, \dots, \hat{\beta}_p = \beta_{p0} + \Delta\beta_p; \quad (3.29)$$

donde $\beta_{10}, \beta_{20}, \dots, \beta_{p0}$ es el conjunto de valores iniciales propuestos para los parámetros; de aquí, se tiene, con la notación acostumbrada:

$$h(\hat{\beta}_1, \dots, \hat{\beta}_p; X_{i1}, \dots, X_{iq}) = h(\beta_{10}, \dots, \beta_{p0}; X_{i1}, \dots, X_{iq}) + \sum_{j=1}^p h'_j(\beta_{10}, \dots, \beta_{p0}; X_{i1}, \dots, X_{iq}) \Delta\beta_j + R_i, \quad (3.30)$$

donde $h'_j = \partial h / \partial \beta_j$, y R_i es el residuo de la i ésima observación, obtenido al aproximar la función h por un polinomio de primer grado en las $\Delta\beta_j$.

El método de mínimos cuadrados se aplica para estimar las $\Delta\beta_j$, aproximando la función h , por el lado derecho de (3.30). Es decir, se tendría la suma de cuadrados de los errores:

$$Q_{(1)} = \sum_{i=1}^n [Y_i - h(\beta_{10}, \dots, \beta_{p0}; X_{i1}, \dots, X_{iq}) - \sum_{j=1}^p h'_j(\beta_{10}, \dots, \beta_{p0}; X_{i1}, \dots, X_{iq}) \Delta\beta_j]^2 \quad (3.31)$$

(3.31) es un modelo lineal en las $\Delta\beta_j$, el cual puede ajustarse por los métodos de regresión múltiple. Aquí termina la primera aproximación del procedimiento.

El procedimiento anterior se repite, tomando como valores iniciales en lugar de las β_{j0} , los valores $\beta_{j0} + \hat{\Delta}\beta_j$, siendo los $\hat{\Delta}\beta_j$, los valores que minimizan el valor de $Q_{(1)}$, de la relación (3.31). Sea:

$$\beta_{j0} + \hat{\Delta}\beta_j = \beta_{j1}, \quad j=1, 2, \dots, p, \quad (3.32)$$

y escribáse de nuevo:

$$\hat{\beta}_j = \beta_{j1} + \Delta_j\beta_j, \quad j=1, 2, \dots, p, \quad (3.33)$$

Se tiene así, expandiendo el valor de h en serie Taylor, alrededor de $\beta_{11}, \beta_{12}, \dots, \beta_{1p}$, la expresión para la suma de cuadrados:

$$Q_{(2)} = \sum_{i=1}^n [Y_i - h(\beta_{11}, \dots, \beta_{p1}; X_{i1}, \dots, X_{iq}) - \sum_{j=1}^p h'_j(\beta_{11}, \dots, \beta_{p1}; X_{i1}, \dots, X_{iq}) \Delta\beta_j]^2,$$

de donde se derivan nuevas estimaciones para $\Delta\beta_j$. Aquí termina la segunda aproximación del procedimiento.

El proceso se continúa, hasta que las $\hat{\Delta}_j\beta_j$ sean prácticamente iguales a cero. Si esto ocurre después de m etapas aproximativas, se toma $\hat{\beta}_j = \beta_{jm}$, $j=1, 2, \dots, p$.

Uso del PROC NLIN de SAS

Algunos modelos de regresión no lineal, se ajustan fácilmente por métodos de computación electrónica. Por ejemplo el procedimiento NLIN de SAS, logra el ajuste de modelos no lineales con distintos algoritmos de cálculo. Varios de ellos requieren la alimentación de las derivadas parciales del modelo con respecto a los parámetros, por parte del usuario del sistema computacional; sin embargo, se tiene un método que no requiere de esta información para lograr el ajuste, pero toma más tiempo de máquina que los anteriores para hacerlo. Ilustraremos el empleo de estas técnicas, en la siguiente sección.

D. EJEMPLOS ILUSTRATIVOS

El ajuste a modelos de regresión lineal y no lineal, a través de paquetes de computación electrónica, se ilustrará con algunos ejemplos en la presente sección.

Ejemplo 3.1.

Considérese el conjunto de datos de la Tabla 3.2, que presenta los rendimientos de caña a partir de 12 fórmulas de fertilización, en promedio de 5 experimentos realizados en la zona de abastecimiento del Ingenio Xicoténcatl, en el Estado de Tamaulipas.

El programa SAS de la Tabla 3.3, realiza el ajuste de los datos al modelo:

$$Y_i = \beta_0 + \beta_1 N_i + \beta_2 N_i^2 + \beta_3 P_i + e_i, \quad i=1,2,\dots,n. \quad (3.31)$$

donde Y_i denota al rendimiento de caña en toneladas por hectárea. Se usan en el programa, los siguientes nombres SAS: R para el rendimiento de caña en $\text{Kg}/60\text{m}^2$, N, P y K para las dosis de nitrógeno, fósforo y potasio en $\text{Kg}/\text{Ha.}$, y Y para el rendimiento de caña en toneladas por hectárea, el cual no es otra cosa, sino una transformación lineal de $Y=R/6$. Obviamente, el ajuste se logra a través del procedimiento GLM de SAS, puesto que se trata de un

Tabla 3.2. Fertilizantes en Xicoténcatl.

Fórmula	N (Kkg. del nutriente/Ha.)	P	K	Rendimiento de caña ₂ (kg. por parcela de 60m ²)
1	0	0	0	434.90
2	200	0	0	499.15
3	0	200	0	482.05
4	200	200	0	559.70
5	0	0	200	434.55
6	200	0	200	508.25
7	0	200	200	446.75
8	200	200	200	537.90
9	100	100	100	502.80
10	300	100	100	558.40
11	100	300	100	581.15
12	100	100	300	533.65

Tabla 3.3. Programa SAS. Ajuste de un Modelo de Regresión

```

DATA UNO; INPUT N P K R;
Y=R/L; CARDS;
0 0 0 434.90
200 0 0 499.15
0 200 0 482.05
200 200 0 559.70
0 0 200 434.55
200 0 200 508.25
0 200 200 446.75
200 200 200 537.90
100 100 100 502.80
300 100 100 558.40
100 300 100 581.15
100 100 300 533.65
PROC GLM; MODEL Y= N NMN P;
PROC PRINT; RUN;

```

modelo lineal (con los parámetros). La Tabla 3.4, presenta los resultados producidos por la computadora. Para la descripción y el uso del SAS, el lector puede consultar Martínez Garza (1980), o cualquiera de los manuales SAS del SAS Institute, por ejemplo SAS Institute (1982).

Para probar $H_0 = \beta_1, \beta_2, \dots, \beta_p$ en oposición a $H_a =$ al menos una β_j es distinta de cero, con $\alpha = 0.01$, obtenemos $F_{0.01, p, n-p} = 7.591$, valor tabulado de F y, puesto que $F_{c, 1} > 7.591$, se rechaza H_0 con $\alpha = 0.01$ y se concluye que al menos una de las variables del modelo produce algún efecto significativo sobre el rendimiento.

La ecuación de regresión estimada es:

$$\hat{Y} = 74.7255 + 0.10243N - 0.0001924N^2 + 0.02840P. \quad (3.35)$$

Ejemplo 3.2.

Considérese el conjunto de datos de la Tabla 3.5, que es la misma que la Tabla 3.2 presentada en el ejemplo 1, con la diferencia de que se reportan desglosados para las 5 localidades.

El programa SAS de la Tabla 3.6 realiza el ajuste de los datos al modelo:

$$Y_i = \beta_0 + \beta_1 N_i + \beta_2 N_i^{0.8} + \beta_3 P_i^{1.6} + \beta_4 P_i + e_i, \quad (3.36)$$

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

Realizando la siguiente transformación:

$$U_i = N_i^{0.8} \quad \text{y} \quad V_i = P_i^{1.6}$$

se obtiene a Y como una función lineal de los parámetros; es decir:

$$Y_i = \beta_0 + \beta_1 N_i + \beta_2 U_i + \beta_3 V_i + \beta_4 P_i + e_i, \quad (3.37)$$

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

El ajuste al modelo (3.37) también se logra con el procedimiento GLM de SAS, puesto que se trata de un modelo lineal.

De los resultados producidos por la computadora reportados en la Tabla 3.7, se obtuvo un valor de $F_{c, 1} = 14.50$. De acuerdo con

Tabla 3.4. Ajuste del Modelo $Y_i = \beta_0 + \beta_1 N_i + \beta_2 N_i^2 + \beta_3 P_i + \epsilon_i$.

SAS
GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE: Y

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE
MODEL	3	575.13629414	192.04543105	8.35
ERROR	11	253.70344315	22.9730367	PR > F
CORRECTED TOTAL	11	750.03964629		0.0076

R-SQUARE	C.V.	ROOT MSE	Y-MEAN
0.758034	5.4765	4.79457302	84.43402773

SOURCE	DF	TYPE I SS	F VALUE	PR > F
N	1	413.60546875	18.21	0.0027
N*N	1	74.17515625	3.27	0.1132
P	1	83.35557814	3.67	0.0934

SOURCE	DF	TYPE III SS	F VALUE	PR > F
N	1	119.27231295	5.27	0.0517
N*N	1	21.37062056	1.30	0.3835
P	1	83.35557814	3.67	0.0934

PARAMETER	ESTIMATE	T FOR H0: PARAMETER=0	PR > T	STD ERROR OF ESTIMATE
INTERCEPT	74.72553763	33.47	0.0001	2.45250141
N	0.10243145	2.28	0.0517	0.04682764
N*N	-0.00019231	-1.15	0.3835	0.00016737
P	0.02840188	1.90	0.0934	0.01491533

Tabla 3.5. Rendimientos de Caña de Ton/Ha en 5 Localidades de la Zona Cañera de Xicoténcatl, Tamps.

Tratamiento	N	P	K					
				1	2	3	4	5
1	0	0	0	82.656	73.021	73.392	59.892	73.455
2	100	0	0	89.031	89.646	82.142	78.851	76.288
3	0	200	0	75.031	71.937	77.309	85.267	92.163
4	200	200	0	89.490	97.646	98.517	98.934	81.830
5	0	0	200	77.531	73.604	59.309	65.309	86.372
6	200	0	200	75.990	89.729	77.226	88.809	91.780
7	0	200	200	78.906	66.687	78.392	71.392	76.913
8	200	200	200	81.573	98.062	99.726	86.976	81.913
9	100	100	100	86.948	87.104	80.892	88.101	75.955
10	300	100	100	98.198	91.104	92.976	106.517	76.538
11	100	300	100	96.323	89.646	96.059	94.226	108.038
12	100	100	300	81.531	85.021	97.267	88.934	91.955

Tabla 3.6. Programa SAS. Ajuste de un Modelo de Regresión

```

DATA UNO; INPUT TRAT LOC N P K R;
U= N**0.8;
V= P**1.6;
CARDS;
1 1 0 0 0 82.656
1 2 0 0 0 73.021
1 3 0 0 0 73.392
1 4 0 0 0 59.892
1 5 0 0 0 73.455
:::
:::
11 4 100 300 100 94.226
11 5 100 300 100 108.038
12 1 100 100 300 81.531
12 2 100 100 300 85.021
12 3 100 100 300 97.267
12 4 100 100 300 88.934
12 5 100 100 300 91.955
PROC GLM; MODEL Y= N U V P; RUN;

```

Nota: Las instrucciones U= N**0.8 y V= P**1.6, generan las variables SAS U y V, que se emplean como regresoras en la instrucción MODEL.

Tabla 3.5. Rendimientos de Caña de Ton/Ha en 5 Localidades de la Zona Cañera de Xicoténcatl, Tamps.

Tratamiento	N	P	K					
				1	2	3	4	5
1	0	0	0	82.656	73.021	73.392	59.892	73.455
2	100	0	0	89.031	89.646	82.142	78.851	76.288
3	0	200	0	75.031	71.937	77.309	85.267	92.163
4	200	200	0	89.490	97.646	98.517	98.934	81.830
5	0	0	200	77.531	73.604	59.309	65.309	86.372
6	200	0	200	75.990	89.729	77.226	88.809	91.788
7	0	200	200	78.906	66.687	78.392	71.392	76.913
8	200	200	200	81.573	98.062	99.726	86.976	81.913
9	100	100	100	86.948	87.104	80.892	88.101	75.955
10	300	100	100	98.198	91.104	92.976	106.517	76.538
11	100	300	100	96.323	89.646	96.059	94.226	108.038
12	100	100	300	81.531	85.021	97.267	88.934	91.955

Tabla 3.6. Programa SAS. Ajuste de un Modelo de Regresión

```

DATA UNO; INPUT TRAT LOC N P K R;
U= N**0.8;
V= P**1.6;
CARDS;
1 1 0 0 0 82.656
1 2 0 0 0 73.021
1 3 0 0 0 73.392
1 4 0 0 0 59.892
1 5 0 0 0 73.455
...
11 4 100 300 100 94.226
11 5 100 300 100 108.038
12 1 100 100 300 81.531
12 2 100 100 300 85.021
12 3 100 100 300 97.267
12 4 100 100 300 88.934
12 5 100 100 300 91.955
PROC GLM; MODEL Y= N U V P; RUN;

```

Nota: Las instrucciones U= N**0.8 y V= P**1.6, generan las variables SAS U y V, que se emplean como regresoras en la instrucción MODEL.

Tabla 3.7. Ajuste del Modelo $Y_i = \beta_0 + \beta_1 N_i + \beta_2 N_i^{0.8} + \beta_3 P_i^{1.6} + \beta_4 P_i + e_i$

SAS
GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE: Y				
SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE
MODEL	4	3446.91446936	861.72861734	14.57
ERROR	55	3268.52057057	59.42401027	PR > F
CORRECTED TOTAL	59	6715.23503993		0.0001
R-SQUARE	C.V.	ROOT MSE	Y MEAN	
0.513298	0.1299	7.70859706	84.53376667	
SOURCE	DF	TYPE I SS	F VALUE	PR > F
N	1	2004.57342849	33.72	0.0001
U	1	533.70150409	9.91	0.0027
V	1	245.13029624	4.77	0.0304
P	1	85.50724054	0.14	0.7066
SOURCE	DF	TYPE III SS	F VALUE	PR > F
N	1	136.82213673	2.30	0.1343
U	1	158.91034464	4.31	0.0427
V	1	13.04253944	0.22	0.6313
P	1	8.50724054	0.14	0.7066
PARAMETER	ESTIMATE	T FOR HO: PARAMETER=0	PR > T	STD ERROR OF ESTIMATE
INTERCEPT	71.47707627	34.05	0.0001	2.10027402
N	-0.17037260	-1.52	0.1343	0.11232604
U	0.62591756	2.08	0.0427	0.23524727
V	0.03072488	0.47	0.6417	0.06475334
P	0.11770151	0.38	0.7066	0.06677583

los grados de libertad del modelo y del error $F_{1, 3, 0.05}^* = 2.051$, se tiene $F_{(1, 3)} > 2.051$, por lo que se rechaza la hipótesis nula y se concluye que al menos uno de los parámetros β es distinto de 0. La ecuación de regresión estimada es:

$$\hat{Y} = 71.47707 - 0.17007N + 0.69598U + 0.0007848V + 0.0177005P \quad (3.38)$$

Ejemplo 3.3.

Se desea ajustar el siguiente modelo:

$$Y_i = \beta_0 + \beta_1 N_i + \beta_2 P_i + \beta_3 N_i^2 + \beta_4 P_i^2 + \beta_5 N_i P_i + e_i, \quad (3.39)$$

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

haciendo uso de los datos de la Tabla 3.5. Tal ajuste se logra con el programa SAS de la Tabla 3.8.

Tabla 3.8. Programa SAS. Ajuste de un Modelo de Regresión

```
DATA UNO; INPUT TRAT LOC N P K R;
CARDS;
1 1 0 0 0 82.656
1 2 0 0 0 73.021
1 3 0 0 0 73.342
1 4 0 0 0 59.892
1 5 0 0 0 73.455
:::
11 4 100 300 100 44.286
11 5 100 300 100 108.038
12 1 100 100 300 81.531
12 2 100 100 300 85.021
12 3 100 100 300 97.267
12 4 100 100 300 88.934
12 5 100 100 300 91.955
PROC RSREG; MODEL Y=N P; RUN;
```

Nota: RSREG calcula automáticamente las variables para los efectos cuadráticos puros y mixtos, de allí que la instrucción MODEL sea simplemente MODEL Y= N P;

El procedimiento utilizado en este caso fue el procedimiento RSREG de SAS que ajusta polinomios de segundo grado, debido a que se quiere estimar el efecto de términos cuadráticos y de interacción sobre el rendimiento. Se puede omitir la declaración de las variables N2, P2, y NP, puesto que el procedimiento RSREG las calcula automáticamente.

Las hipótesis nula y alterna siguen siendo las mismas que en los ejemplos anteriores. Los valores estimados para los β_i son dados por la Tabla 3.9; como $F_{\alpha, 1, 10} = 10.28 > F_{1, 10, 0.05}^{\alpha} = 1.981$ se rechaza H_0 , concluyendo que existe significancia estadística del modelo. En la parte inferior de la Tabla 3.9, el procedimiento RSREG proporciona las raíces y vectores característicos de la matriz B (la matriz B se define en la Sección 2 del Capítulo I), y el punto crítico, indicando que dicho punto es un punto de silla.

La ecuación de regresión estimada es:

$$\hat{Y} = 72.4541 + 0.10103882N + 0.1857882P - 0.000184999N^2 + 0.000073584P^2 + 0.000064277NP. \quad (3.40)$$

Ejemplo 3.4.

Considérese el conjunto de datos de la Tabla 3.10, que presenta los rendimientos de un experimento llevado a cabo por el Dr. W. L. Nelson, del departamento de Agronomía de la Universidad Estatal de Carolina del Norte, sobre papa irlandesa con fertilizante fosfatado a los niveles de 0, 40, 80, 120 y 160 libras por acre, de superfosfato triple.

El rendimiento medio obtenido se expresa en libras por parcela de $1/65$ de acre.

El programa SAS de la Tabla 3.11, realiza el ajuste de los datos al modelo:

$$Y = A + BR^X + e. \quad (3.41)$$

por varios procedimientos, a saber: GAUSS, MARQUARDT, GRADIENT y DUD.

Tabla 3.9. Ajuste del Modelo $Y_i = \beta_0 + \beta_1 N_i + \beta_2 P_i + \beta_{11} N_i^2 + \beta_{22} P_i^2 + \beta_{12} N_i P_i$
a Través de RSREG de SAS.

SAS

RESPONSE SURFACE FOR VARIABLE Y

RESPONSE MEAN	94.43357
ROOT MSE	7.37572
R-SQUARE	0.5014686
COEF OF VARIATION	0.093253

REGRESSION	DF	TYPE III SS	R-SQUARE	F-RATIO	PROB
LINEAR	2	3152.25344	0.4694	25.42	0.0001
QUADRATIC	2	122.64762	0.0294	1.61	0.2139
CROSSPRODUCT	1	16.52638802	0.0025	0.22	0.6377
TOTAL REGRESS	5	3367.47944	0.5015	10.87	0.0001
RESIDUAL	DF	SS	MEAN SQUARE		
TOTAL ERROR	64	3347.75560	61.6954466		

PARAMETER	DF	ESTIMATE	STD DEV	T-RATIO	PROB
INTERCEPT	1	72.45610000	2.48598904	29.10	0.0001
N	1	0.12133832	3.35477184	3.62	0.0001
P	1	0.31527682	0.33457784	0.94	0.6377
N*N	1	-0.000129959	0.00123378	-1.06	0.1326
P*P	1	0.002344277	0.00124434	1.88	0.0777
N*P	1	0.000073584	0.00113317	0.60	0.5534

FACTOR	DF	SS	MEAN SQUARE	F-RATIO	PROB
N	2	2305.33	768.4501	12.40	0.0001
P	2	996.0985	332.3328	5.36	0.0028

SOLUTION FOR OPTIMUM RESPONSE

FACTOR	CRITICAL VALUE
N	235.64735
P	-215.57026

Nota: Obsérvese que RSREG de SAS aplica la técnica del análisis canónico y diagnóstica la naturaleza del punto crítico de $Y = \hat{\beta}_0 + \hat{\beta}_1 N + \hat{\beta}_2 P + \hat{\beta}_{11} N^2 + \hat{\beta}_{22} P^2 + \hat{\beta}_{12} NP$.

PREDICTED VALUE AT OPTIMUM 92.57316

EIGENVALUES	EIGENVECTORS	N	P
0.00007751861	0.12133832	0.9925876	
-0.000149834	0.000073584	-0.121332	

SOLUTION HAS A SADDLE POINT

Este último procedimiento no requiere la alimentación de las derivadas de Y con respecto a A y B, los parámetros del modelo.

El procedimiento NLIN requiere el empleo de varias instrucciones, a saber: PARAMETERS, DER. y MODEL.

Tabla 3.10. Fertilizante de fosfato en un cultivo de papa irlandesa.

Superfosfato Triple	0	40	80	120	160
Nivel (x)	0	1	2	3	4
Rendimiento (R)	229.1	231.8	254.	250.1	249.6

La instrucción PARAMETERS introduce valores iniciales para los parámetros, siendo en este caso $A = 251.5$, $B = -22.4$ y $R = 0.455$. Estos valores iniciales deben ser sugeridos por el investigador por algún procedimiento aproximado, véase Draper y Smith (1967).

Las instrucciones DER. especifican las derivadas del modelo y la instrucción MODEL indica el modelo a utilizar, en términos de una expresión SAS.

Los resultados del SAS se reportan en la Tabla 3.12. La mínima suma de cuadrados de los errores es de 131.786, la cual ocurre cuando $A = 255.53$, $B = -28.308$ y $R = 0.5744$.

El método GRADIENT no produjo convergencia.

Tabla 3.11. Programa SAS. Ajuste de un Modelo no lineal

```

DATA UNO; INPUT X Y;
CARDS;
0 229.1
1 231.8
2 254.2
3 250.6
4 249.6
PROC NLIN METHOD=GAUSS;
PARAMETERS A=251.5 B=-22.4 R=0.455;
MODEL Y=A+B*R**X;
DER,A=1;
DER,B=R**X;
DER,R=B**X*R**X*(X-1);
PROC NLIN METHOD=MARQUARDT;
PARAMETERS A=251.5 B=-22.4 R=0.455;
MODEL Y=A+B*R**X;
DER,A=1;
DER,B=R**X;
DER,R=B**X*R**X*(X-1);
PROC NLIN METHOD=GRADIENT;
PARAMETERS A=251.5 B=-22.4 R=0.455;
MODEL Y=A+B*R**X;
DER,A=1;
DER,B=R**X;
DER,R=B**X*R**X*(X-1);
PROC NLIN METHOD=DUD;
PARAMETERS A=251.5 B=-22.4 R=0.455;
MODEL Y=A+B*R**X;
PROC PRINT;
RUN;

```

Tabla 3.12. Ajuste de un Modelo de Regresión no Lineal a Tráves del PROC NLIN de SAS.

MÉTODO DE MARQUARDT				
NON-LINEAR LEAST SQUARES ITERATIVE PHASE				
DEPENDENT VARIABLE: Y				
METHOD: MARQUARDT				
ITERATION	A	B	C	RESIDUAL SS
0	251.500000	-22.50000000	0.45500000	142.55651329
1	255.695023	-29.334872115	0.572312215	120.260488074
2	255.568675	-27.261021201	0.540241222	120.231283935
3	253.316115	-25.451788511	0.547629092	120.231030030
4	253.429554	-25.117101127	0.577162272	121.257167822
5	255.471769	-28.253068261	0.570506827	121.257205027
6	255.564667	-28.224160201	0.575426909	121.266128721
7	255.509418	-28.200707444	0.572784937	121.266128721
8	255.562002	-27.217059195	0.574721221	121.265822121
9	255.522025	-28.201026349	0.574122226	121.265822121
10	255.522013	-28.210261817	0.574548272	121.265822121
11	255.522167	-28.204667751	0.574321265	121.265822121
12	255.522517	-28.204268911	0.574422522	121.265822121

NOTE: CONVERGENCE CRITERION MET.

MÉTODO DE MARQUARDT				
NON-LINEAR LEAST SQUARES SUMMARY STATISTICS				
DEPENDENT VARIABLE: Y				
SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	
REGRESSION	3	225001.42400	75000.47469	
RESIDUAL	2	121.26582	60.63291	
UNCORRECTED TOTAL	5	225003.68982		
(CORRECTED TOTAL)	4	565.23200		
PARAMETER	ESTIMATE	ASYMPTOTIC STD. ERROR	ASYMPTOTIC 95% CONFIDENCE INTERVAL	ASYMPTOTIC 95% CONFIDENCE INTERVAL
A	255.5225169	17.224792109	191.2222500	320.6666262
B	-28.2042422	17.021279655	-121.2948122	45.17011809
C	0.5743226	0.572246123	-1.4544651	0.26119226

ASYMPTOTIC CORRELATION MATRIX OF THE PARAMETERS

CORR	A	B	C
A	1.0000		
B	-0.8016	1.0000	
C	0.0536	-0.2522	1.0000

Tabla 3.12. Continuación.

METHOD OF GAUSS				
NON-LINEAR LEAST SQUARES ITERATIVE PHASE				
DEPENDENT VARIABLE Y		METHOD OF GAUSS-NEWTON		
ITERATION	A	B	C	RESIDUAL SS
0	251.500000	-27.5000000	0.454000000	147.68681200
1	256.172413	-28.20702875	0.476648810	130.91067165
2	253.821213	-25.32953132	0.565717205	135.02722666
3	255.693796	-28.24722646	0.451384725	131.91714651
4	255.380083	-27.12924638	0.539357147	131.79710753
5	255.610014	-27.36672770	0.571346974	131.79572608
6	255.470033	-27.26572165	0.672023623	131.79603048
7	255.560004	-28.32052652	0.575222174	131.79622784
8	255.512224	-28.32222222	0.672023644	131.79622784
9	255.561221	-27.31640322	0.574720117	131.79622784
10	255.523220	-28.28162677	0.574212132	131.79622784
11	255.523143	-28.22822222	0.574322211	131.79622784
12	255.522101	-28.22222222	0.574422211	131.79622784
13	255.522652	-28.22422222	0.574227617	131.79622784

NOTE: CONVERGENCE CRITERION MET.

METHOD OF GAUSS			
NON-LINEAR LEAST SQUARES SUMMARY STATISTICS		DEPENDENT VARIABLE Y	
SOURCE	DEGREE OF FREEDOM	MEAN SQUARE	
REGRESSION	3	29526.47468	9842.15823
RESIDUAL	11	11.79532	1.07206
UNCORRECTED TOTAL	4	29642.26999	
(CORRECTED TOTAL)	4	545.26000	
PARAMETER	ESTIMATE	ASYMPTOTIC STD. ERROR	ASYMPTOTIC 95 % CONFIDENCE INTERVAL LOWER UPPER
A	255.5226522	17.328167211	101.37102653 209.67422217
B	-28.2226520	17.322145541	-111.20514963 54.76035505
C	0.5742276	0.477147222	-1.46466693 2.61279012

ASYMPTOTIC COVARIANCE MATRIX OF THE PARAMETERS

CORR	A	B	C
A	1.0000	-0.0015	-1.0000
B	-0.0015	1.0000	-0.7500
C	-1.0000	-0.7500	1.0000

Tabla 3.12. Continuación.

MÉTODO DE DUD				
NON-LINEAR LEAST SQUARES ITERATIVE PHASE				
DEPENDENT VARIABLE: Y				
METHOD: DUD				
ITERATION	A	B	R	RESIDUAL SS
-4	251.500000	-22.400000000	0.455000000	146.584516705
-3	275.650000	-22.400000000	0.455000000	3404.702754712
-2	251.500000	-24.640000000	0.455000000	147.665205677
-1	251.500000	-24.400000000	0.503500000	145.312302644
0	251.500000	-24.400000000	0.455000000	147.665005667
1	225.884740	-24.625002202	0.560706125	124.756408480
2	225.467843	-24.300000086	0.615641020	123.133027604
3	220.458917	-20.372076775	0.611549364	122.618251053
4	226.535876	-20.102084827	0.611897581	122.093271071
5	225.274744	-20.126447702	0.567800270	121.860787150
6	225.344072	-20.240527052	0.557905440	121.815174464
7	225.377664	-20.261655050	0.560716100	121.809070720
8	225.465176	-20.352027540	0.567012134	121.800007116
9	225.657300	-20.255000130	0.573753433	121.805402000
10	225.617865	-20.410620043	0.574414293	121.780002041
11	225.611565	-20.400712420	0.574219645	121.780002041
12	225.615075	-20.395272144	0.574232044	121.780002041
13	225.614766	-20.396677427	0.574232044	121.780002041
14	225.597855	-20.4271220145	0.574231000	121.780002041
15	225.527672	-20.300703120	0.574234841	121.780002041
16	225.524381	-20.303047171	0.574244734	121.780002041
17	225.520234	-20.307550551	0.574230136	121.780002041
18	225.520066	-20.306362224	0.574235050	121.780002041
19	225.520237	-20.306045478	0.574236022	121.780002041

NOTE: CONVERGENCE CRITERION MET.

Tabla 3.12. Continuación.

MÉTODO DE MÍN			
NON-LINEAR LEAST SQUARES SUMMARY STATISTICS			DEPENDENT VARIABLE
SOURCE	DF	SUM OF SQUARES	MEAN SQUARE
REGRESSION	3	295934.42409	98644.77469
RESIDUAL	2	121.70592	60.85296
UNCORRECTED TOTAL	5	295936.21000	
(CORRECTED TOTAL)	4	545.30200	

PARAMETER	ESTIMATE	ASYMPTOTIC STD. ERROR	ASYMPTOTIC 95% CONFIDENCE INTERVAL	
			LOWER	UPPER
A	255.5222369	17.249030507	181.1118326	320.74664020
B	-28.7060657	17.110537147	-101.02744476	45.31521364
B	0.5743682	0.473964522	-1.45495927	2.61369522

ASYMPTOTIC CORRELATION MATRIX OF THE PARAMETERS

PARAMETER	A	B	B
A	1.0000	-0.8017	0.0320
B	-0.8017	1.0000	-0.7400
B	0.0320	-0.7400	1.0000

NOTE: ALL ASYMPTOTIC STATISTICS ARE APPROXIMATE. REFERENCE: BALSTON AND JENNIFER, TECHNOMETRICS, FEBRUARY 1971

CAPITULO IV

ILUSTRACIONES DE LA TECNICA.

Una vez que se han desarrollado los temas anteriores, a saber: 1. Localización de puntos extremos, 2. Extremos de funciones por métodos numéricos y 3. Ajuste de funciones de respuesta, toca ahora ilustrar las técnicas descritas, aplicadas a la solución de algunos problemas típicos. La discusión comenzará, considerando el caso de la localización de condiciones óptimas, cuando la función de respuesta es un polinomio entero; enseguida se procederá considerando funciones de respuesta cada vez más complicadas, cubriendo los casos de polinomios con exponentes fraccionarios, hasta funciones de estructura más complicada, cuyo ajuste requiere el uso de los métodos de la regresión no lineal. En cada caso, si se requiere, se describirán los detalles de las técnicas bajo consideración. También se ilustra la localización de puntos óptimos con restricciones.

A. OPTIMOS CON POLINOMIOS ENTEROS.

Esencialmente, se considerarán ejemplos que empleen funciones de respuesta de segundo grado, puesto que éstas ocurren con mucha frecuencia en estudios económicos.

1. Polinomio de segundo grado.

En el Ejemplo 1 del Capítulo III, la ejecución del programa SAS que manda el ajuste de los rendimientos de caña al polinomio:

$$Y_1 = \beta_0 + \beta_1 N_1 + \beta_{11} N_1^2 + \beta_2 P_1 + e_{11}, \quad (4.1)$$

donde Y_1 es rendimiento de caña en toneladas por hectárea y N_1 y P_1 son las dosis de nitrógeno y de fósforo aplicadas al suelo por hectárea, la computadora produce los resultados de la Tabla 3.1. De esta tabla se obtiene la ecuación de regresión estimada:

$$\hat{Y} = 74.7255 + 0.10243N - 0.0001924N^2 + 0.02840P. \quad (4.2)$$

Suponiendo que la tonelada de caña se liquida al agricultor, a razón de \$31,000 y que los precios del kilogramo de nitrógeno y de fósforo aplicados al suelo, son, respectivamente de \$530 y \$635, entonces el ingreso neto, I, que obtiene el agricultor, es dado por:

$$I = 31000(74.7255 + 0.10243N - 0.0001924N^2 + 0.02840P) - 530N - 635P - C_0, \quad (4.3)$$

siendo C_0 , el costo fijo de producción, cuyo monto es innecesario calcular, puesto que es una constante en este problema. Claramente I es un polinomio de segundo grado en N, y de primer grado en P. Obtendremos la dosis de nitrógeno que maximiza el ingreso neto del agricultor.

Por los resultados del Capítulo I, I es un máximo, cuando:

$$\frac{dI}{dN} = 0 \quad \text{y} \quad \frac{d^2I}{dN^2} < 0.$$

Operando sobre (4.1), se obtiene:

$$\frac{dI}{dN} = 31000(0.10243 - 2 \times 0.0001924N) - 530 = 0,$$

$$\therefore N_{\text{optimo}} = \frac{0.10243 - 530/31000}{2 \times 0.0001924} = 221.75 \text{ kg.}$$

Es fácil ver que $d^2I/dN^2 = 31000 \times (-2 \times 0.0001924) = -11.92 < 0$, lo que implica que $N_{\text{optimo}} = 221.75 \text{ kg/Ha}$, es el punto donde I es máximo.

2. Polinomio de segundo grado en varias variables.

En un experimento realizado en el noroeste de México, reportado por Marinato y Palacios (1979), se obtuvo una función que relaciona el rendimiento del cultivo de trigo, con los niveles de humedad aprovechable consumida del suelo, en tres etapas fenológicas diferentes; del mismo experimento se obtuvo también, la relación entre la cantidad de agua consumida y los mismos niveles de humedad aprovechable. Las funciones estimadas son:

$$Y = 2.32 + 7.95X_1 + 5.18X_2 + 4.86X_3 - 4.78X_1^2 - 4.59X_2^2 - 5.81X_3^2 - 3.64X_1X_2, \quad (4.4)$$

y

$$A = 129.12 - 21.78X_1 - 53.7X_2 - 22.46X_3, \quad (4.5)$$

donde Y es el rendimiento de trigo en toneladas por hectárea, X_i es la proporción de la humedad aprovechable consumida antes del siguiente riego en la etapa fenológica i , $i=1,2,3$, y A es la lámina de agua usada por el cultivo, expresado en centímetros.

Supóngase que el precio del trigo es $P_Y = \$2900$ por tonelada y que el precio del agua es $P_A = \$40$ por centímetro. De aquí, I, el ingreso neto, es dado por:

$$I = P_Y Y - C_0 - P_A A = 2900Y - C_0 - 40A, \quad (4.6)$$

donde C_0 denota el costo fijo de producción. Encontraremos el punto donde I, dado por (4.6), es un máximo.

Por los resultados del Capítulo I, si la función I, dada por (4.6), tiene un extremo, éste debe ocurrir en el punto donde:

$$\frac{\partial I}{\partial X_1} = 0, \quad \frac{\partial I}{\partial X_2} = 0, \quad \frac{\partial I}{\partial X_3} = 0.$$

De aquí, tomando derivadas parciales en la expresión para I y usando las relaciones (4.4) y (4.5), se tiene, sucesivamente:

$$\begin{aligned} \frac{\partial I}{\partial X_1} &= 2900 \frac{\partial Y}{\partial X_1} - 40 \frac{\partial A}{\partial X_1} \\ &= 2900(7.95 - 9.56X_1 - 3.64X_2) - 40(-21.78) = 0, \end{aligned}$$

$$\begin{aligned} \frac{\partial I}{\partial X_2} &= 2900 \frac{\partial Y}{\partial X_2} - 40 \frac{\partial A}{\partial X_2} \\ &= 2900(5.18 - 9.18X_2 - 3.64X_1) - 40(-53.7) = 0, \end{aligned}$$

$$\begin{aligned} \frac{\partial I}{\partial X_3} &= 2900 \frac{\partial Y}{\partial X_3} - 40 \frac{\partial A}{\partial X_3} \\ &= 2900(4.86 - 11.62X_3) - 40(-22.46) = 0. \end{aligned}$$

Después de algunas simplificaciones, las ecuaciones anteriores se reducen a:

$$9.56X_1 + 3.64X_2 = 8.2504,$$

$$3.64X_1 + 9.18X_2 = 5.9207,$$

$$11.62X_3 = 5.1698,$$

cuya solución numérica es $X_{1,0} = 0.727$, $X_{2,0} = 0.357$ y $X_{3,0} = 0.445$. Por otra parte, operando sobre las primeras derivadas de la función I, se obtiene:

$$\frac{\partial^2 I}{\partial X_1^2} = 2900(-9.56), \quad \frac{\partial^2 I}{\partial X_1 \partial X_2} = 2900(-3.64), \quad \frac{\partial^2 I}{\partial X_1 \partial X_3} = 0,$$

$$\frac{\partial^2 I}{\partial X_2^2} = 2900(-9.18), \quad \frac{\partial^2 I}{\partial X_2 \partial X_3} = 0, \quad \frac{\partial^2 I}{\partial X_3^2} = 2900(-11.62),$$

de modo que la matriz Hessiana, ignorando el factor constante 2900, es:

$$H = \begin{bmatrix} -9.56 & -3.64 & 0 \\ -3.64 & -9.18 & 0 \\ 0 & 0 & -11.62 \end{bmatrix};$$

y sus menores principales son $\Delta_1 = -9.56$, $\Delta_2 = (-9.56)(-9.18) - (-3.64)(-3.64) = 70.84$, $\Delta_3 = -11.62 [(-9.56)(-9.18) - (-3.64)(-3.64)] = -868.82$. Estos alternan en signo como lo requiere el Teorema 1.9 del Capítulo I, de modo que la función I, dada por (4.6), tiene un máximo en el punto (0.727, 0.357, 0.445).

3. Ilustración de la técnica del análisis canónico.

Como vimos en el Capítulo I, la técnica del análisis canónico para determinar los extremos de un polinomio de segundo grado, representa un método muy apropiado en estos casos. Ilustraremos el empleo de la técnica, con los resultados completos

de la investigación sobre fertilizantes, que en forma condensada se usaron en el Ejemplo 3.1. Así, la Tabla 3.5 presenta los resultados medios de 5 experimentos, donde se investiga el efecto del nitrógeno, fósforo y potasio, sobre el rendimiento de caña de azúcar. Por otra parte, el programa de la Tabla 3.8 conduce al modelo de regresión ajustado que sigue:

$$\hat{Y} = 72.4541 + 0.10103882N + 0.01857682P - 0.00018497N^2 \quad (4.7) \\ + 0.00007358P^2 + 0.00006428NP$$

con los precios considerados en la subsección 1, I, el ingreso neto para el agricultor, es dado por:

$$I = 31000(72.4541 + 0.10103882N + 0.01857682P - 0.00018497N^2 \\ + 0.00007358P^2 + 0.00006428NP) - 530N - 635P - C_0.$$

Encontraremos, si existe, el punto máximo de I.

Obsérvese que maximizar I, es equivalente a maximizar $I/31000$, donde:

$$\frac{I}{31000} = 72.4541 + 0.10103882N + 0.01857682P \\ - 0.00018497N^2 + 0.00007358P^2 + 0.00006428NP \\ - 0.17098774N - 0.2048387P - C_0/31000 \\ = 72.4541 + 0.083942046N - 0.00390705P \\ - 0.00018497N^2 + 0.00007358P^2 + 0.00006428NP \\ - C_0/31000.$$

Ahora, de acuerdo con la notación introducida en el Capítulo I, para el análisis canónico, se tiene: $\beta_0 = 72.4541 - C_0/31000$, $\beta_1 = 0.083942046$, $\beta_2 = -0.00390705$, $\beta_{11} = -0.00018497N^2$, $\beta_{22} = 0.00007358$ y $\beta_{12} = 0.00006428 = \beta_{21}$. De aquí, el vector β y la matriz B del propio Capítulo I, son:

$$\beta = \begin{bmatrix} 0.08394204 \\ -0.00390705 \end{bmatrix}, \quad B = \begin{bmatrix} -0.00018497 & 0.00003214 \\ 0.00003214 & 0.00007358 \end{bmatrix};$$

consecuentemente, las coordenadas del punto crítico, son dados por:

$$Z_0 = -B^{-1} \beta / 2 = - \begin{bmatrix} -0.00018497 & 0.00003214 \\ 0.00003214 & 0.00007358 \end{bmatrix}^{-1} \begin{bmatrix} 0.08394204 \\ -0.00390705 \end{bmatrix} \times 0.5$$

$$= \begin{bmatrix} 215.1880 \\ -67.4452 \end{bmatrix}$$

Por otra parte, las raíces características de la matriz B, se obtienen como solución de la ecuación:

$$(-0.00018497 - \lambda)(0.00007358 - \lambda) - (0.00003214)^2 = 0,$$

de la cual se obtiene $\lambda_1 = 0.0000775153$ y $\lambda_2 = -0.00018890$, valores que difieren en signo, indicando que ocurre un punto de silla y no hay máximo ni mínimo de I.

Derivación de una recomendación. Puesto que el ingreso neto I, no se maximiza dentro (ni fuera) de la región de exploración, es necesario producir una recomendación para los agricultores. Una manera de hacerlo, consiste en explorar la frontera de la región experimental, calculando el ingreso neto en varios puntos. El problema es adecuado para su solución por métodos computacionales. Así por ejemplo, el programa SAS de la Tabla 4.1, calcula el valor de I para 31² puntos experimentales, puesto que tanto los valores de N, como los de P, se incrementan a intervalos de 10 Kilogramos. El uso del procedimiento SORT, manda el ordenamiento de los valores de I, de mayor a menor. El primer punto que se obtenga en la impresión de resultados, deberá ser nuestra recomendación. Como resultado del programa anterior, la máquina imprime varias hojas, de las cuales la primera se reproduce en la Tabla 4.2. La recomendación que se puede dar a los agricultores, es la siguiente: N= 280 y P= 300, que produce un ingreso neto de: \$4,045,040.76 por hectárea (valor que incluye el costo fijo de producción).

Tabla 4.1. Programa SAS. Cálculo del Ingreso Neto

```

DATA UNO;
DO N=0 TO 300 BY 10;
DO P=0 TO 300 BY 10;
I= 31000*(72.4541+0.10103682*N+0.01657682*N*P
    -0.0001677*N*N+0.00007358*N*P
    +0.00006428*N*P)-530*N-635*P;
OUTPUT; END; END;
PROC SORT; BY DESCENDING I;
DATA DOS; SET UNO; IF _N_ LE 100;
PROC PRINT DATA=DOS;
PROC PRINT; RUN;
    
```

Tabla 4.2. Resultado de la Tabla 4.1.

		SAS		
OBS	N	P	I	
1	280	300	4045040.76	
2	270	300	4044578.08	
3	290	300	4044356.63	
4	260	300	4042968.58	
5	300	300	4042525.69	
6	250	300	4040212.27	
7	240	300	4036309.15	
8	230	300	4031259.21	
9	220	300	4025062.46	
10	210	300	4017718.89	
11	200	300	4009228.51	
12	190	300	3999591.31	
13	180	290	3994377.50	
14	270	290	3994114.08	
15	290	290	3993494.10	
16	260	290	3992703.85	
17	300	290	3991463.89	

B. OPTIMOS CON POLINOMIOS DE EXPONENTES FRACCIONARIOS

Consideraremos en esta sección, un par de ejemplos donde se localizan puntos óptimos con polinomios de variables con exponentes fraccionarios. Dadas las dificultades para resolver las ecuaciones que establecen las condiciones de primer orden, en el segundo ejemplo se propone un método computacional para definir el punto óptimo.

1. Polinomio pseudocuadrático.

Martínez Garza y Palacios V. (1981) consideran el siguiente problema: en un experimento realizado en el Estado de Sonora, se encontró una función que relaciona el rendimiento de trigo con cantidades de agua y nitrógeno utilizados. La función obtenida es:

$$Y = -40.13 - 8.72A - 0.085N + 30.26A^{0.5} + 1.97N^{0.5} + 0.03A^{0.5}N^{0.5} \quad (4.8)$$

donde Y es el rendimiento de trigo en toneladas por hectárea, A la cantidad de agua usada en millares de metros cúbicos por hectárea y N la dosis de nitrógeno en kilogramos por hectárea; si se supone que el precio del trigo es de \$2900 la tonelada, el precio del agua \$400 el millar de metros cúbicos y el precio del nitrógeno \$12 por kilogramo, ¿cuáles son los niveles óptimos de agua y nitrógeno que permiten maximizar el ingreso?

Solución. Sea I el ingreso en pesos por hectárea. Es fácil ver que una expresión para I, es la siguiente:

$$I = 2900(-40.13 - 8.72A - 0.085N + 30.26A^{0.5} + 1.97N^{0.5} + 0.03A^{0.5}N^{0.5}) - C_0 - 400A - 12N, \quad (4.9)$$

donde C_0 es el costo fijo de producción, el cual es una constante en este problema. (4.9) puede escribirse:

$$\frac{I}{2900} = -40.13 - 8.72A - 0.085N + 30.26A^{0.5} + 1.97N^{0.5} \\ + 0.03A^{0.5}N^{0.5} - \frac{C_0}{2900} - \frac{400}{2900}A - \frac{12}{2900}N,$$

expresión que al substituir A por U^2 y N por V^2 , toma la forma:

$$\frac{I}{2900} = \left[-40.13 - \frac{C_0}{2900} \right] + 30.26U + 1.97V \\ + \left[-8.72 - \frac{400}{2900} \right]U^2 + \left[-0.085 - \frac{12}{2900} \right]V^2 + 0.03UV \\ = \left[-40.13 - \frac{C_0}{2900} \right] + 30.26U + 1.97V \\ -6.85793U^2 - 0.0891379V^2 + 0.03UV. \quad (4.10)$$

Claramente (4.10) es un polinomio cuadrático, cuyo único extremo puede determinarse por la técnica del análisis canónico, tomando:

$$\beta = \begin{bmatrix} 30.26 \\ 1.97 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} -6.85793 & 0.015 \\ 0.015 & -0.0891379 \end{bmatrix}.$$

Consecuentemente, con la notación del Capítulo I, las coordenadas de Z_0 , el punto estacionario, son:

$$Z_0 = \begin{bmatrix} U_0 \\ V_0 \end{bmatrix} = -B^{-1}\beta/z = \begin{bmatrix} 2.2312 \\ 11.4257 \end{bmatrix};$$

además las raíces características de la matriz B son $\lambda_1 = -0.0891$ y $\lambda_2 = -6.658$, ambas negativas, lo cual implica que Z_0 es un máximo, y puesto que $A = U^2$ y $N = V^2$, entonces:

$$A_0 = (2.2312)^2 = 4.978 \text{ millares de metros cúbicos de agua,}$$

y:

$$N_0 = (11.4257)^2 = 130.5 \text{ kilogramos de nitrógeno por hectárea.}$$

Cuando $A = 4.978$ y $N = 130.5$, $I + C_0 = \$14,158.83$, que es el ingreso neto máximo.

2. Rendimiento e ingreso máximo con un polinomio pseudocuadrático.

Marín García (1989) en su examen preliminar para obtener el grado de maestro en ciencias, en el Centro de Hidrociencias del Colegio de Postgraduados, presenta los resultados de un experimento de maíz, en el cual se ensayan dosis de nitrógeno en kilogramos del nutriente por hectárea, N , y densidad de semilla en kilogramos por hectárea, D , sobre el rendimiento del cultivo, Y , expresado en kilogramos de grano por hectárea, los cuales se reportan en la Tabla 4.3. Por métodos de regresión múltiple se obtuvo la relación:

$$Y = 1353.785048 + 26.857304N^{0.6} + 425.725796D^{0.7} - 0.466002N^{1.6} - 15.379641D^{1.4} + 1.499999N^{0.6}D^{0.7} \quad (4.11)$$

Procediendo como en el ejemplo anterior, la transformación $N^{0.6} = U$ y $D^{0.7} = V$, produce el polinomio de segundo grado que sigue:

$$Y = 1353.785048 + 26.857304U + 425.725796V - 0.466002U^2 - 15.379641V^2 + 1.499999UV. \quad (4.12)$$

Usando la técnica del análisis canónico, con:

$$\beta = \begin{bmatrix} 26.857304 \\ 425.725796 \end{bmatrix}, \quad B = \begin{bmatrix} -0.466002 & 0.749995 \\ 0.749995 & -15.379641 \end{bmatrix}.$$

el punto estacionario es dado por:

$$\begin{bmatrix} U_0 \\ V_0 \end{bmatrix} = -B^{-1}\beta/2 = \begin{bmatrix} 55.4437138 \\ 15.5442169 \end{bmatrix}.$$

Tal punto es un máximo, puesto que las raíces características de B , son $\lambda_1 = -15.4172$ y $\lambda_2 = -0.42837$, ambas negativas. En términos de N y de D y por la transformación definida, el rendimiento máximo de Y , ocurre cuando $N = 151.29$ y $D = 55.07$.

Tabla 4.3. Rendimiento de grano (Kg/Ha) obtenido para cada uno de los tratamientos de nitrógeno y presión poblacional bajo estudio

TRAT.	NITROGENO N (Kg/Ha)	PRESION DE POBLACION D (P1/Ha)	R E P E T I C I O N E S			
			I	II	III	IV
1	0	20	2446.52	4139.23	4030.90	4739.58
10	0	60	2830.20	3507.31	4974.31	4103.13
2	0	100	2162.15	1469.58	2703.02	3259.02
5	60	40	4996.88	5258.67	5904.16	6648.95
6	60	80	3931.59	5330.90	5899.65	4816.32
12	120	20	3832.28	3615.63	4455.20	5525.03
9	120	60	3751.03	4048.95	4834.38	5592.70
13	120	100	2785.07	4107.64	4572.57	3999.31
7	180	40	3633.68	4667.35	7140.97	6630.91
8	180	80	5915.46	5335.41	6495.49	8377.77
3	240	20	2807.64	2699.31	5136.79	5714.58
11	240	60	3913.53	4721.52	4784.72	6332.99
4	240	100	3398.95	3322.22	5046.52	5732.64

Cálculo del máximo por métodos numéricos. El máximo de Y , dentro de la región de exploración, a saber: $\{(N, D) | 0 \leq N \leq 240, 20 \leq D \leq 100\}$, se encontrará por métodos computacionales. Esto se hace a través del programa SAS de la Tabla 4.4. Los resultados dados por la computadora, después de examinar más de 10.000 puntos dentro de la región de exploración, se presentan en la Tabla 4.5, la cual sólo presenta la impresión de la primera hoja de los propios resultados. Los valores de N y de D que producen el máximo rendimiento, son: $N = 151$ y $D = 55$, los cuales coinciden con los producidos por el análisis canónico, lo que prueba la bondad del método de cálculo numérico. Se trató de resolver el mismo problema por el método de Newton, sin que hubiera convergencia a una solución.

Una vez que se ha mostrado la bondad del método del cálculo numérico introducido en esta tesis, encontraremos los niveles de N y D , que maximizan el ingreso neto del agricultor, digamos I , cuando el kilogramo de nitrógeno cuesta \$517, el kilogramo de semilla de maíz \$384 y el precio de venta del kilogramo de maíz es de \$ 325, (precios dados por Marín García (1989)).

Con los precios anteriores, es fácil ver que I , el ingreso neto del agricultor, es dado por:

$$I = 325(1353.765048 + 26.857304N^{0.8} + 425.725798D^{0.7} - 0.466002N^{1.0} - 15.379641D^{1.4} + 1.49999N^{0.8}D^{0.7}) - 517N - 384D - 286000,$$

donde la constante 286,000, es el costo fijo de producción, en pesos.

El programa S/S de la Tabla 4.6, que explora el mismo número de puntos que aquél de la Tabla 4.4, produce los resultados de la Tabla 4.7. De acuerdo con esta última tabla, $N = 130$ y $D = 53$, producen un ingreso máximo, dentro de la región de exploración, de \$ 1,447,243.48.

Tabla 4.4. Programa SAS. Cálculo del Rendimiento Máximo

```

DATA UNO;
DO N=0 TO 240 BY 1;
DO D=20 TO 100 BY 1;
Y= 1353.769048+26.857304*(N**0.8)+425.725796*(D**0.7)
    -0.466002*(N**1.6)-15.379641*(D**1.4)
    +1.4999998*(N**0.8)*(D**0.7);
OUTPUT; END; END;
PROC SORT; BY DESCENDING Y;
DATA DOS; SET UNO; IF _N_ LE 200;
PROC PRINT DATA=DOS;
RUN;

```

Tabla 4.5. Resultado de la Tabla 4.4.

SAS				
OBS	N	D	Y	
1	151	55	5619.97	
2	152	55	5619.94	
3	150	55	5619.91	
4	153	55	5619.84	
5	149	55	5619.77	
6	154	55	5619.66	
7	148	55	5619.55	
8	152	54	5619.43	
9	153	56	5619.42	
10	155	55	5619.40	
11	151	56	5619.36	
12	154	56	5619.33	
13	147	55	5619.25	
14	150	54	5619.23	
15	151	54	5619.21	
16	150	56	5619.21	
17	149	54	5619.20	

Tabla 4.6. Programa SAS. Cálculo del Ingreso Neto

```

DATA UNO;
DO N=0 TO 240 BY 1;
DO D=20 TO 100 BY 1;
I=325*(1353.765048+26.857304*(N**0.8)+425.725796*(D**0.7)
  -0.466002*(N**1.6)-15.379641*(D**1.4)
  +1.4999998*(N**0.6)*(D**0.7))-51.7*N-384*D-286000;
OUTPUT; END; END;
PROC SORT; BY DESCENDING I;
DATA DOS; SET UNO; IF _N_ LE 200;
PROC PRINT DATA=DOS;
RUN;

```

Tabla 4.7. Resultado de la Tabla 4.6.

SAS			
OBS	N	D	I
1	130	53	1447243.48
2	129	53	1447230.84
3	131	53	1447227.67
4	128	53	1447189.61
5	132	53	1447183.53
6	129	52	1447143.36
7	128	52	1447133.63
8	130	52	1447124.55
9	127	53	1447119.65
10	133	53	1447111.20
11	127	52	1447095.22
12	131	52	1447077.33
13	126	52	1447027.98
14	126	53	1447020.81
15	134	53	1447010.82
16	132	52	1447001.84
17	125	52	1446931.79

C. MAXIMOS Y MINIMOS RESTRICTOS

En esta sección se presentarán algunos ejemplos de problemas de máximos y mínimos, cuando no hay restricciones entre las variables.

1. Máximo restringido con una función potencial.

Martínez G. y Palacios V. (1981) presentan los resultados de un estudio realizado en Brazil en 1971; se encontró que el ingreso bruto esperado por campesino, Y , era una función de la superficie cultivada, S , del número de jornales empleados, T , y otros costos variables, C . A partir de una encuesta sobre predios, se estimó la función:

$$Y = 80.4S^{0.569} T^{0.086} C^{0.9}, \quad (4.13)$$

donde Y se da en cruzeiros, S en hectáreas, T en jornales y C también en cruzeiros. Se sabe que la renta de la tierra es de 100 cruzeiros por hectárea, el jornal de 5 cruzeiros y la tasa de interés al capital de 10%; además, el gasto medio que los campesinos pueden realizar, es del orden de 250 cruzeiros, es decir:

$$100S + 5T + 1.1C = 250. \quad (4.14)$$

Considérese el problema de maximizar (4.13), sujeto a la condición (4.14).

Usando el método de multiplicadores de los Langrange, se establece la función:

$$L = 80.4S^{0.569} T^{0.086} C^{0.9} + \lambda(250 - 100S - 5T - 1.1C).$$

Así, la primera condición, necesaria para el máximo de L , es:

$$\frac{\partial L}{\partial S} = L_1 = 80.4 \times 0.569S^{-0.431} T^{0.086} C^{0.9} - 100\lambda = 0,$$

$$\frac{\partial L}{\partial T} = L_2 = 80.4 \times 0.086S^{0.569} T^{-0.914} C^{0.3} - 5\lambda = 0,$$

$$\frac{\partial L}{\partial C} = L_3 = 80.4 \times 0.3S^{0.569} T^{0.086} C^{-0.3} - 1.1\lambda = 0,$$

$$\frac{\partial L}{\partial \lambda} = 250 - 100S - 5T - 1.1C = 0.$$

De las ecuaciones anteriores se obtiene el sistema:

$$45.7476S^{-0.431} T^{0.086} C^{0.3} = 100\lambda, \quad (4.15a)$$

$$6.0144S^{0.569} T^{-0.914} C^{0.3} = 5\lambda, \quad (4.15b)$$

$$24.12S^{0.569} T^{0.086} C^{-0.7} = 1.1\lambda, \quad (4.15c)$$

$$100S + 5T + 1.1C = 250. \quad (4.15d)$$

Dividiendo miembro a miembro (4.15a) entre (4.15b), resulta:

$$6.6162791S^{-1}T = 20, \quad \therefore \frac{T}{S} = 3.0228471. \quad (4.16)$$

Dividiendo miembro a miembro (4.15a) entre (4.15c), resulta:

$$1.8968667S^{-1}C = 90.909091, \quad \therefore \frac{C}{S} = 47.930979. \quad (4.17)$$

De los resultados (4.16) y (4.17) se obtiene, respectivamente:

$T = 3.0228471S$ y $C = 47.930979S$. Estos valores se insertan en la condición (4.15d), de donde se obtiene: $S_0 = 1.48953$; de aquí, las relaciones anteriores para T y C producen: $T_0 = 4.50262$ y $C_0 = 71.3946$ y el valor de λ es 1.5779 .

Para demostrar que L tiene un máximo en $S_0 = 1.48953$, $T_0 = 4.50262$ y $C_0 = 71.3946$ debe obtenerse la matriz Hessiana de la frontera, que se discute en la Sección C del Capítulo I. Se observa que:

$$L_{11} = 45.7476 \times (-0.431)S^{-1.431} T^{0.086} C^{0.3},$$

$$L_{12} = 45.7476 \times (0.086)S^{-0.431} T^{-0.914} C^{0.3} = L_{21},$$

$$L_{33} = 45.7476 \times (0.3)S^{-0.431} T^{0.086} C^{-0.7} = L_{31},$$

$$L_{2,2} = 8.9144 \times (-0.914) S^{0.369} T^{-1.014} C^{0.9},$$

$$L_{2,3} = 8.9144 \times (0.3) S^{0.369} T^{-0.914} C^{-0.7} = L_{2,1},$$

$$L_{3,3} = 24.12 \times (-0.7) S^{0.369} T^{0.089} C^{-1.7};$$

además, como se tiene la restricción $g(S,T,C) = 250 - 100S - 5T - 1.1C$, entonces:

$$g_1(S,T,C) = -100, \quad g_2(S,T,C) = -5 \quad \text{y} \quad g_3(S,T,C) = -1.1.$$

Evaluando las L_{ij} en el punto $S_0 = 1.48953$, $T_0 = 4.50262$, $C_0 = 71.3945$ (es innecesario hacerlo con las g_{ij} , puesto que son constantes para todo valor de S , T y C); se obtiene entonces:

$$H = \begin{bmatrix} 0 & g_1 & g_2 & g_3 \\ g_1 & L_{11} & L_{12} & L_{13} \\ g_2 & L_{21} & L_{22} & L_{23} \\ g_3 & L_{31} & L_{32} & L_{33} \end{bmatrix} = \begin{bmatrix} 0 & -100 & -5 & -1.1 \\ -100 & -45.6573 & 3.013905 & 0.663037 \\ -5 & 3.013905 & -1.60152 & 0.331519 \\ -1.1 & 0.663037 & 0.0331519 & -0.017018 \end{bmatrix}$$

Es fácil ver que el tercero y el cuarto determinantes menores de H , digamos Δ_3 y Δ_4 , son: $\Delta_3 = +20170.4$ y $\Delta_4 = -172.4$, los cuales alternan en signo, comenzando con el de $(-1)^{1+1} = +1$. Por consiguiente, la función (4.13) tiene un máximo restringido por la condición (4.14), en el punto: $S_0 = 1.49$ hectáreas, $T_0 = 4.5$ jornales, $C_0 = 71.4$ cruzeiros, siendo el valor máximo de Y , $Y_0 = 366.4$ cruzeiros.

2. Costo mínimo restringido a un nivel fijo de producción.

Considérese la ecuación de regresión estimada:

$$Y = 2757.5117 + 12.260788N - 0.057692N^2 + 65.666621D - 0.696958D^2 + 0.105441ND, \quad (4.18)$$

que da el rendimiento de maíz Y , en kilogramos por hectárea, en términos de N y D , la dosis de nitrógeno expresada en kilogramos por hectárea y la cantidad de semilla aplicada por hectárea, expresada en kilogramos, respectivamente, cuando los datos de la

Tabla 4.3, se ajustan a un polinomio cuadrático en N y D. Tal ajuste como hemos visto, se logra con facilidad a través de un programa SAS. Encontraremos los valores de N y de D, que minimizan el costo de producción en pesos por hectárea C, dado por:

$$C = 286000 + 517N + 384D, \quad (4.19)$$

sujeto a la restricción de que Y = 5000 kilogramos de maíz por hectárea.

Usando el método de los multiplicadores de Lagrange, descrito en el Capítulo I, se forma la función:

$$L = 286,000 + 517N + 384D - \lambda [5000 - (2757.5117 + 12.260768N - 0.057692N^2 + 65.666821D - 0.696956D^2 + 0.105441ND)]. \quad (4.20)$$

El mínimo de L (si es que ocurre), se tiene cuando $\partial L / \partial N = 0$, $\partial L / \partial D = 0$, $\partial L / \partial \lambda = 0$, es decir, cuando:

$$\frac{\partial L}{\partial N} = 517 + \lambda(12.260768 - 0.115364N + 0.105441D) = 0,$$

$$\frac{\partial L}{\partial D} = 384 + \lambda(65.666821 - 1.393912D + 0.105441N) = 0, \quad (4.21)$$

$$\frac{\partial L}{\partial \lambda} = (2757.5117 + 12.260768N - 0.057692N^2 + 65.666821D - 0.696956D^2 + 0.105441ND) - 5000 = 0,$$

Sistema que produce la solución aceptable (hay otra inaceptable por que ocurre fuera de la región de exploración):

$$N = 49.474302, \quad D = 44.841415 \quad \text{y} \quad \lambda = -45.831843,$$

después de un álgebra muy laboriosa. La matriz Q de la Sección C del Capítulo I, tiene los siguientes elementos:

$$\frac{\partial^2 L}{\partial N^2} = -0.115364\lambda, \quad \frac{\partial^2 L}{\partial N \partial D} = +0.105441\lambda,$$

$$\frac{\partial^2 L}{\partial D \partial N} = +0.105441\lambda, \quad \frac{\partial^2 L}{\partial D^2} = -1.393912\lambda.$$

Ahora, como se tiene una sola restricción, definida por la relación:

$$g(N,D) = (2757.5117 + 12.260786N - 0.057692N^2 + 65.666821D - 0.696958D^2 + 0.105441ND) - 5000 = 0,$$

entonces:

$$g_1(N,D) = \frac{\partial g}{\partial N} = 12.260786 - 0.115384N + 0.105441D,$$

$$g_2(N,D) = \frac{\partial g}{\partial D} = 65.666821 - 1.393912D + 0.105441N.$$

De aquí, la matriz Hessiana de la frontera (la matriz H^B de la Sección C del Capítulo I), evaluada en el punto estacionario: $N = 49.474302$, $D = 44.841415$, $\lambda = -45.831843$, es dada por:

$$H = \begin{bmatrix} 0 & 11.280367 & 8.3784544 \\ 11.280367 & 5.2862614 & -4.8325554 \\ 8.3784544 & -4.8325554 & 63.88556 \end{bmatrix}.$$

Es fácil ver que $|H^B| = -9413.92 < 0$, que tiene el signo de $(-1)^m = (-1)^1 = -1$, (donde $m=1$, puesto que se tiene una sola restricción) lo que implica que C tiene un mínimo relativo en el punto $(N = 49.47, D = 44.84)$, por el inciso (c) de la Sección C del Capítulo I.

Puede demostrarse que en el punto $(N = 271.32, D = 73.64)$, que produce la solución inaceptable referida en líneas anteriores, ocurre un máximo relativo.

Solución alternativa. Una técnica más práctica para localizar el punto que conduce al costo mínimo de producción, para el problema que se ha venido discutiendo, explota los medios de computación electrónica. Así, cuando el rendimiento de maíz se restringe a un nivel Y_0 de producción en la expresión (4.18), se obtiene la ecuación cuadrática en N que sigue:

$$(-0.057692)N^2 + (12.260786 + 0.105441D)N + (2757.5117 + 65.666821D - 0.696958D^2 - 5000) = 0,$$

ecuación que puede resolverse fácilmente para N, produciendo la solución:

$$N = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \quad (4.21)$$

donde $a = -0.057892$, $b = 12.260786 + 0.105441D$, $c = 2757.5117 + 85.666821D - 0.696956D^2 - 5000$. Este valor de N dado por (4.21) se inserta en (4.19), y la expresión resultante se evalúa para algunos valores de D, dentro del rango experimental. El programa SAS de la Tabla 4.8, hace el trabajo anterior para valores de D entre 20 y 100, a intervalos unitarios, y para valores de Y_0 que van de 3000 a 5000, a intervalos de 100 unidades. La tabla 4.9 da los puntos donde C se maximiza o se minimiza, para los valores propuestos de D.

3. Rendimiento máximo restringido a un crédito fijo.

Si ahora el problema es maximizar Y dado por (4.18), sujeta a la restricción de que el costo de producción C dado por (4.19), se fija a un nivel C_0 (lo cual puede ocurrir cuando el crédito a los agricultores, se restringe a un cierto monto, digamos C_0), entonces se construye la función:

$$L = 2757.5117 + 12.260786N - 0.057892N^2 + 85.666821D - 0.696956D^2 + 0.105441ND - \lambda [C_0 - (286000 + 517N + 384D)]. \quad (4.22)$$

El máximo de L (si es que ocurre), se tiene cuando $\partial L / \partial N = 0$, $\partial L / \partial D = 0$, $\partial L / \partial \lambda = 0$, es decir, cuando:

$$\frac{\partial L}{\partial N} = 12.260786 - 0.115384N + 0.105441D + 517\lambda = 0,$$

$$\frac{\partial L}{\partial D} = 85.666821 - 1.393912D + 0.105441N + 384\lambda = 0,$$

$$\frac{\partial L}{\partial \lambda} = 286000 + 517N + 384D - C_0 = 0.$$

El programa SAS de la Tabla 4.10 resuelve el sistema

Tabla 4.8. Programa SAS. Cálculo del costo mínimo y máximo de producción

```

DATA UNO;
DO D=20 TO 100 BY 1; DO CO=3000 TO 5000 BY 100;
A=-0.057692; B=0.105441*D+12.260786;
C=2757.5117-CO+65.666621*D-0.696956*D*D;
Q=B*B-4*A*C;
IF Q GE 0 THEN N1=(-B+SQRT(Q))/(2*A);
IF Q LT 0 THEN N1=.;
IF Q GE 0 THEN N2=(-B-SQRT(Q))/(2*A);
IF Q LT 0 THEN N2=.;
CS1=286000+517*N1+384*D;
CS2=286000+517*N2+384*D;
OUTPUT; END; END;
PROC SORT; BY CO CS1; PROC PRINT; BY CO;
PROC SORT; BY CO CS2; PROC PRINT; BY CO;
RUN;
    
```

Tabla 4.9. Resultado de la Tabla 4.8.

CO	Mínimo Costo			Máximo Costo		
	D	N1	CS1	D	N2	CS2
3000	91	1.76	314808.28	80	383.61	515041.90
3100	89	0.86	320624.40	80	379.31	512824.73
3200	87	0.22	319522.27	80	374.93	510558.95
3300	86	2.35	320239.55	80	370.44	508241.21
3400	84	2.10	319344.27	80	365.85	505867.78
3500	82	2.12	318585.98	80	361.15	503434.42
3600	79	0.21	316149.36	80	356.31	500936.36
3700	77	0.90	316035.85	80	351.35	498368.15
3800	20	0.55	292298.70	80	346.23	495723.52
3900	23	0.05	294861.62	80	340.95	492995.21
4000	26	0.42	296201.22	80	335.50	490174.74
4100	29	1.59	297960.76	80	329.84	487252.09
4200	34	0.98	299563.45	80	323.97	484215.21
4300	39	2.56	302000.02	80	317.85	481049.75
4400	40	8.18	305589.45	79	312.24	477767.15
4500	40	14.77	309001.14	78	306.38	474351.18
4600	41	20.92	312561.31	77	300.22	470785.84
4700	20	27.39	316291.43	77	293.00	467051.66
4800	43	34.24	320216.95	76	286.16	463129.13
4900	44	41.53	324371.02	75	278.87	458976.53
5000	45	49.35	328798.24	74	271.049	454548.56

anterior de ecuaciones, para N , D y λ . Puede verse que en todos los puntos que satisfacen las condiciones de primer orden, la matriz Hessiana de la frontera, es:

$$H^B = \begin{bmatrix} 0 & -517 & -384 \\ -517 & -0.115384 & 0.105441 \\ -384 & 0.105441 & -1.393912 \end{bmatrix}$$

$|H^B| = 431457.4 > 0$, que es el signo de $(-1)^{m+1} = (-1)^2 = 1$, lo que implica que en todos los puntos que satisfacen las condiciones de primer orden, ocurre un máximo restringido, por el inciso d) de la Sección C del Capítulo I.

Tabla 4.10. Programa SAS. Cálculo del Rendimiento Máximo

```

DATA UNO;
DO CO=30000 TO 50000 BY 10000;
A=0,105441+1,039912*517/384;
B=65,666821*517/384-12,260786;
C=0,105441*517/384+0,115384;
E=(CO-286000)/517; F=-384/517;
D=(B+CE)/(A-C*F); N=E+F*D;
LAMBDA=(12,260786-0,115384*(CO-286000-384*D))/517+(0,105441*D)/(-517);
Y=2757,5117+12,260786*N-0,057692*N*N+65,666821*D-0,676956*D*D
+0,105441*N*D;
OUTPUT; END;
PROC SORT; BY Y;
PROC PRINT DATA=UNO; VAR N D LAMBDA CO Y;
RUN;

```

Tabla 4.11. Resultado de la Tabla 4.10.

SAS

OBS	N	D	LAMBDA	CO	Y
1	351.496	84.0535	0.0375071	50000	3650.05
2	333.855	81.7631	0.0341190	49000	4008.59
3	316.214	79.4727	0.0306490	48000	4332.42
4	16.314	40.5361	-0.0283416	31000	4528.56
5	298.573	77.1823	0.0271790	47000	4621.57
6	33.955	42.8265	-0.0248716	32000	4794.62
7	260.931	74.8919	0.0237089	46000	4876.01
8	51.596	45.1169	-0.0214016	33000	5025.99
9	263.290	72.6915	0.0202389	45000	5095.79
10	69.237	47.4073	-0.0179315	34000	5227.66
11	245.649	70.3111	0.0167689	44000	5280.79
12	86.878	49.6977	-0.0144615	35000	5364.62
13	228.008	68.0207	0.0132988	43000	5431.17
14	104.570	51.9880	-0.0109914	36000	5511.07
15	210.367	65.7304	0.0098288	42000	5546.77
16	122.161	54.2784	-0.0075214	37000	5604.45
17	192.725	63.4400	0.0063587	41000	5627.70
18	179.902	56.5688	-0.0040514	38000	5662.71
19	175.084	61.1496	0.0028867	40000	5673.94
20	157.443	58.8592	-0.0005813	39000	5685.48

CAPITULO V

RESUMEN

En esta tesis se discuten, en forma condensada, los elementos matemáticos, estadísticos y prácticos del problema de la localización de condiciones óptimas.

En el Capítulo I se consideran algunos teoremas que establecen las condiciones de ocurrencia de puntos extremos (máximos o mínimos), en el caso de funciones de una o más variables. Se discuten tanto las situaciones de condiciones extremas sin restricciones, como aquellos en que se imponen restricciones sobre las variables.

En el Capítulo II presenta algunos de los métodos numéricos que desempeñan un papel importante, en relación con el problema de localizar condiciones óptimas. El énfasis aquí, es tratar de resolver el problema cuando es prácticamente imposible encontrar una solución abierta, como ocurre cuando el problema involucra funciones no lineales. Se exponen, entre otros, los métodos de Newton-Raphson y el del ascenso acelerado, que son técnicas de búsqueda secuencial.

Por otra parte, como el ajuste de la función de respuesta, es uno de los aspectos relevantes en la determinación de condiciones óptimas, el Capítulo III cubre muy brevemente, los métodos estadísticos de ajuste para modelos lineales y no lineales. Para tal propósito, se ilustra el uso de los procedimientos GLM, RSREG y NLIN de SAS, en varios ejemplos de naturaleza aplicada, cuando el ajuste de las funciones de respuesta, se realiza por métodos de computación electrónica.

Finalmente, el Capítulo IV presenta una serie de ejemplos ilustrativos, con los cuales se redondea el tema de esta tesis.

En algunos casos, se sugieren algunos artificios para la localización de condiciones óptimas, que constituyen una aportación práctica de este trabajo.

CAPITULO IV

REFERENCIAS

- Bronson, R. 1984. Teoría y Problemas de Investigación de Operaciones. Mc Graw-Hill Book Co., U.S.A. Traducción al español por María de Lourdes Fournier García, Universidad Nacional Autónoma de México.
- Castillo, M. A. 1978. Predicción óptima en el modelo $Y = \alpha x^n e$. Agrociencia No. 33: 139-143.
- Draper, N. R. y Smith, H. 1987. Applied regression analysis. John Wiley, Nueva York.
- Haaser, N. B., Lasalle, J. P. y Sullivan, J. A. 1979. Análisis Matemático. Massachusetts, E.U.A. Traducción al español por Federico Velasco Coba, Facultad de Ciencias, Universidad Veracruzana.
- Marín, G. M. 1989. Crecimiento y rendimiento del maíz CP-581 en relación a la presión de población, aprovechamiento del nitrógeno y balance hídrico bajo condiciones de temporal en la región central costera de Veracruz. Tesis de maestría. Colegio de Postgraduados. Centro de Hidrociencias, Montecillos, México.
- Martínez, G. A. 1974. Métodos Econométricos. Colegio de Postgraduados, Chapingo, México.
- Martínez, G. A., 1980. Introducción al SAS. CEC, Colegio de Postgraduados, Chapingo México.
- Myers, R. H., Response surface methodology, Allyn and Bacon, Boston, 1971.
- SAS Institute. 1982. SAS User's guide Raleigh, N.C.

SAS Institute. 1985. User guide. Raleigh, N.C.

Taha, H. A. 1976. Investigación de operaciones. Traducción al español por José de Jesús Acosta Flores, Universidad Nacional Autónoma de México.

Widder, D. 1981. Advanced Calculus. Prentice Hall, Inc; Englewood Cliffs, N.J.

Bibliografía complementaria.

Martínez, G. y Castillo. 1987. Teoría de la Regresión con aplicaciones agronómicas. Colegio de Postgraduados, Chapingo, México.

Martínez, G. y Palacios, E. 1981. Determinación de óptimos económicos en investigaciones agropecuarias. ALCA (Asociación Latinoamericana de Ciencias Agrícolas). Vol. 16 No. 1: 123-141.

Martínez, G. A. Econometría. Notas de clase. Centro de Estadística y Cálculo; Universidad Autónoma de Chapingo.

Pimentel, G. F. 1981. The Use of Mitscherlich's Regression Law in the analysis of experiments with fertilizers. Biometrics, 9: 498-516.