CAPITULO 4

"Las abejas..., en virtud de una cierta intuición geométrica…, saben que el hexágono es mayor que el cuadrado y que el triángulo, y que podrá contener más miel con el mismo gasto de material"

Pappus de Alejandría.

OPTIMIZACION DE FUNCIONES ESCALARES

4.1 FORMULA DE TAYLOR.

En el capítulo anterior usamos el plano tangente a una superficie para hacer aproximaciones de la misma en la vecindad de un punto. Ahora ampliemos nuestro campo de aplicación y veamos como podemos hacer este tipo de aproximaciones a funciones escalares de orden superior, como por ejemplo funciones cuadráticas.

Para funciones derivables de una variable la fórmula de Taylor nos permite disponer de un polinomio de grado "n" para esta aproximación, para funciones escalares de $R^n \to R$ resulta complejo disponer de un polinomio de grado mayor a 2; revisemos, primero, la fórmula de Taylor para una función derivable de variable real en una vecindad de x_0.

Un polinomio de grado "n" es de la forma:

$$P_n(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots\dots\dots\dots + a_nx^n$$

Como este polinomio lo usaremos para aproximar una función de variable real en la vecindad de un punto, podemos escribir:

$$f(x) \approx a_0 + a_1x + a_2x^2 + a_3x^3 + \dots\dots\dots\dots + a_nx^n \qquad \textbf{4-1}$$

Escribamos 4-1 en el punto $x - x_0$, que esta dentro de una vecindad cualquiera de x:

$$f(x) \approx a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + a_3(x-x_0)^3 + \dots\dots + a_n(x-x_0)^n \qquad \textbf{4-2}$$

Veamos la forma de encontrar los coeficientes del polinomio en la ecuación 4-2:

$$f(x_0) = a_0 + 0 \quad \Rightarrow \quad a_0 = f(x_0)$$

$$f'(x) = a_1 + 2a_2(x-x_0) + 3a_3(x-x_0)^2 + \dots\dots + na_n(x-x_0)^{n-1}$$

$$f''(x) = 2a_2 + (2)(3)a_3(x-x_0) + \dots\dots + (n-1)(n)a_n(x-x_0)^{n-2}$$

$$f'''(x) = (2)(3)a_3 + \dots\dots + (n-2)(n-1)(n)a_n(x-x_0)^{n-3}$$

. . . .

$f^{(n)}(x) = n!a_n$. De aquí podemos obtener los coeficientes:

$$f'(x_0) = a_1 + 0 \quad \Rightarrow \quad a_1 = f'(x_0)$$

$$f''(x_0) = 2a_2 + 0 \quad \Rightarrow \quad a_2 = \frac{f''(x_0)}{2!}$$

$$f'''(x_0) = 6a_3 + 0 \quad \Rightarrow \quad a_3 = \frac{f'''(x_0)}{3!}$$

.
.
.

$$f^{(n)}(x_0) = n!a_n \quad \Rightarrow \quad a_n = \frac{f^{(n)}(x_0)}{n!}.$$

Ahora podemos escribir la ecuación 4-2 de la siguiente forma:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \ldots.. + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

Que para hacer la igualdad le agregamos un error de aproximación:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \ldots.. + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + R_n(x, x_0) \qquad \textbf{4-3}$$

Donde el error $R_n(x, x_0)$ para x cercano a x_0 es:

$R_n(x, x_0) = \int_{x_0}^{x} \frac{(x - t)^n}{n!} f^{(n+1)}(t)dt$, y debe ser un valor pequeño entre más cerca se encuentre x de x_0; se dice "*un infinitésimo de orden superior a n*" que significa:

$$\lim_{x \to x_0} \frac{R_n(x, x_0)}{(x - x_0)^n} = 0.$$

Nos interesa expresar la ecuación 4-3 aplicada a una función escalar $f(X): R^n \to R$, diferenciable en una vecindad de $X_0 \in R^n$. Como el diferencial de segundo orden, para este caso, es una matriz cuadrada, resulta complejo expresar la fórmula 4-3 para un orden mayor al segundo; por cuanto el diferencial de tercer orden es el diferencial de la matriz cuadrada y resulta complejo expresarlo; peor todavía los diferenciales de orden mayor.

Para $f(X): R^n \to R$, el diferencial de primer orden es el vector gradiente, y $(X - X_0)$ se lo puede expresar como el vector:

$(X - X_0) = (x_1 - x_{01}, x_2 - x_{02}, \ldots\ldots, x_n - x_{0n})$, entonces:

Fórmula de Taylor de primer orden:

Sea $f(X): U \subset R^n \to R$; diferenciable en $X_0 \in R^n$, la fórmula de Taylor al primer orden para $f(X)$ en una vecindad de X_0, se puede escribir de la forma:

$$f(X) = f(X_0) + \nabla f(X_0) \bullet (X - X_0) + R_1(X, X_0)$$

Donde $R_1(X, X_0)$ es un infinitésimo de orden superior al primero y tiene la propiedad que:

$$\underset{X \to X_0}{Lim} \frac{R_1(X, X_0)}{\|X - X_0\|} = 0$$

Considerando $X - X_0$ como vector o matriz columna podemos escribir en forma similar a la ecuación 4-3 la fórmula de Taylor de segundo orden:

Fórmula de Taylor de segundo orden:

Sea $f(X): U \subset R^n \to R$; doblemente diferenciable en $X_0 \in R^n$, la fórmula de Taylor al segundo orden para $f(X)$ en una vecindad de X_0, se puede escribir de la forma:

$$f(X) = f(X_0) + \nabla f(X_0) \bullet (X - X_0) + \tfrac{1}{2} H[f(X_0)][X - X_0] \bullet (X - X_0) + R_2(X, X_0)$$

Donde $R_2(X, X_0)$ es un infinitésimo de orden superior al segundo y tiene la propiedad que:

$$\underset{X \to X_0}{Lim} \frac{R_2(X, X_0)}{\|X - X_0\|^2} = 0$$

$\left[X - X_0\right]$ es el vector diferencia expresado como matriz columna de la siguiente forma:

$$\left[X - X_0\right] = \begin{bmatrix} x_1 - x_{01} \\ x_2 - x_{02} \\ . \\ . \\ . \\ \\ x_n - x_{0n} \end{bmatrix}$$, matriz columna de dimensión $n \times 1$

$H\left[f(X)\right]$ es la matriz segunda derivada y se la llama matriz Hessiana, definida de la siguiente forma:

Definición de Matriz Hessiana:

Sea $f(X): U \subset R^n \to R$; de tipo C^2 en una vecindad de X_0 la matriz $n \times n$:

$$H\left[f(x)\right] = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1{}^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2{}^2} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \frac{\partial^2 f}{\partial x_3 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n{}^2} \end{bmatrix}$$

Se llama matriz Hessiana de la función escalar y representa su diferencial de segundo orden.

Entonces, la misma fórmula de Taylor al segundo orden expresada en forma de sumatorias se la escribiría de la siguiente forma:

$$f(x) = f(x_o) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_i - x_{0i}) + \tfrac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_i - x_{0i})(x_j - x_{0j}) + R_2(x, x_0)$$

La matriz Hessiana tiene las siguientes características:

- Es una matriz cuadrada de dimensión *n x n*.
- Tiene las derivadas parciales dobles en la diagonal principal y las mixtas en la triangular superior e inferior.
- Es simétrica, porque las derivadas parciales mixtas son iguales.
- Hace el papel de la segunda derivada en funciones de variable real.

El producto $H[f(X_0)][X - X_0]$, representa un producto de dos matrices, de dimensiones $(n \times n)$ por $(n \times 1)$ y resulta un vector de n componentes que se multiplica escalar mente con el vector diferencia $X - X_0$ y forma el tercer término de la fórmula de Taylor al segundo orden.

$$f(X) = f(X_0) + \nabla f(X_0) \bullet (X - X_0) + \overbrace{\frac{1}{2}\underbrace{H[f(X_0)][X - X_0]}_{\text{producto matricial}} \bullet (X - X_0)}^{\text{producto escalar}} + R_2(X, X_0)$$

Si la función escalar esta en R^2:

$$f : R^2 \to R$$

$$H[f(x,y)] = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \qquad (X - X_o) = (x - x_o ; y - y_o)$$

Ejemplo 4-1 Escribir la fórmula de Taylor al 1er. orden para la función: $f(x,y) = Sen(x + 2y)$ en una vecindad de $(0,0)$

Solución:

$$f(x,y) = Sen(x + 2y)$$

$$\nabla f = (Cos(x + 2y), 2Cos(x + 2y))$$

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

$$\nabla f\big|_{(0,0)} = (1,2)$$

$$f(x,y) = 0 + (1,2) \bullet (x-0, y-0) + R_1$$

$$f(x,y) = x + 2y + R_1$$

Ejemplo 4-2 Escribir la fórmula de Taylor al 2º orden para la función: $f(x,y) = e^{xy} sen(x+y)$ en una vecindad de $P(0, \frac{\pi}{2})$

Solución:

$$\nabla f = \left[ye^{xy} sen(x+y) + e^{xy}\cos(x+y),\quad xe^{xy} sen(x+y) + e^{xy}\cos(x+y)\right]$$

$$H[f] = \begin{bmatrix} \begin{array}{l} y^2e^{xy}sen(x+y) + ye^{xy}\cos(x+y) + \\ ye^{xy}\cos(x+y) - e^{xy}sen(x+y) \end{array} & \begin{array}{l} e^{xy}sen(x+y) + xye^{xy}sen(x+y) + \\ ye^{xy}\cos(x+y) + xe^{xy}\cos(x+y) - \\ e^{xy}sen(x+y) \end{array} \\ \begin{array}{l} e^{xy}sen(x+y) + xye^{xy}sen(x+y) + \\ xe^{xy}\cos(x+y) + ye^{xy}\cos(x+y) - \\ e^{xy}sen(x+y) \end{array} & \begin{array}{l} x^2e^{xy}sen(x+y) + xe^{xy}\cos(x+y) + \\ xe^{xy}\cos(x+y) - e^{xy}sen(x+y) \end{array} \end{bmatrix}$$

$$\nabla f\big|_{(0,\pi/2)} = (\tfrac{\pi}{2}, 0)$$

$$H[f]_{(0,\pi/2)} = \begin{bmatrix} \frac{\pi^2}{4} - 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$X - X_0 = (x, y - \tfrac{\pi}{2})$$

$$[X - X_0] = \begin{bmatrix} x \\ y - \frac{\pi}{2} \end{bmatrix}$$

$$H[f(X_0)][X - X_0] = (\tfrac{\pi^2}{4}x - x, -y + \tfrac{\pi}{2})$$

$$f(0, \tfrac{\pi}{2}) = 1$$

$$\nabla f(X_0) \cdot (X - X_0) = \tfrac{\pi}{2}x$$

$$H[f(X_0)][X - X_0] \bullet (X - X_0) = (\tfrac{\pi^2}{4}x - x)x + (y - \tfrac{\pi}{2})(-y + \tfrac{\pi}{2})$$
$$f(x, y) = 1 + \tfrac{\pi}{2}x + \tfrac{1}{2}(\tfrac{\pi^2}{4} - 1)x^2 + \tfrac{1}{2}(-y^2 + \pi y - \tfrac{\pi^2}{4}) + R_2(x, y)$$

$$f(x, y) = \tfrac{\pi^2 - 4}{8}x^2 - \tfrac{1}{2}y^2 + \tfrac{\pi}{2}x + \tfrac{\pi}{2}y + (1 - \tfrac{\pi^2}{8}) + R_2(x, y)$$

Ejemplo 4-3 Encontrar la fórmula de Taylor al segundo orden para la función $f(x, y) = e^x Cosy$ en una vecindad de $(0,0)$ y analizar su bondad de aproximación en $(0.1,0.1)$

Solución:

$$\nabla f(x, y) = (e^x Cosy, -e^x Seny)$$
$$\nabla f(0,0) = (1,0)$$
$$f(0,0) = 1$$

$$H[f(x, y)] = \left.\begin{bmatrix} e^x \cos y & -e^x seny \\ -e^x seny & -e^x \cos y \end{bmatrix}\right|_{(0,0)} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$[X - X_o] = \begin{bmatrix} x - 0 \\ y - 0 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$H[f(x_o, y_o)]\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}$$

$$f(x, y) = f(x_o, y_o) + \nabla f(x_o, y_o) \bullet ((x, y) - (x_o, y_o)) +$$
$$\frac{1}{2}[H[f(x_o, y_o)][(x, y) - (x_o, y_o)]] \bullet [(x, y) - (x_o, y_o)] + R$$

$$f(x, y) = 1 + (1,0) \bullet (x, y) + \tfrac{1}{2}(x, -y) \bullet (x, y) + R_2$$
$$= 1 + x + \tfrac{1}{2}(x^2 - y^2) + R_2$$
$$f(x, y) = 1 + x + \tfrac{1}{2}x^2 - \tfrac{1}{2}y^2 + R_2$$

$$f(0.1, 0.1) = 1.099649$$

$$f(x, y)\Big|_{(0.1,0.1)} = 1.1$$

$$R_2 \approx 1.1 - 1.099649 \approx 0.000351$$

Como la fórmula de Taylor sirve para encontrar la aproximación de una función escalar cualquiera a una función lineal o cuadrática en una vecindad de un punto dado, es muy útil para demostrar teoremas o expresar definiciones o aplicaciones especiales que sin esta fórmula darían mucha fatiga realizarlas como es el caso de las aplicaciones de la matriz Hessiana para calificar valores extremos, que lo veremos más adelante. El ejemplo 4-4 demuestra la regla de la cadena enunciada en la sección 3.5, teorema 3-10 y que dejamos pendiente su demostración en el capítulo anterior.

Ejemplo 4-4 Demostrar la regla de la cadena enunciada en el teorema 3-10, sección 3.5.

Solución: Sea: $g(x_0) = y_0$; $g(x_0 + h)$, es el valor de la función en cualquier punto x y a una distancia h de x_0. Entonces usando la aproximación de Taylor al primer orden tenemos:

$$g(x_0 + h) = g(x_0) + Dg(x_0)h + R_g(x_0, h)$$

Haciendo: $k = Dg(x_0)h + R_g(x_0, h)$, tenemos:

$g(x_0 + h) = g(x_0) + k$; por otro lado podemos ver que:

$$\lim_{h \to 0} k = 0 .$$

Si se cumple la propuesta de la regla de la cadena:

$$D(f \circ g)(x_0) = Df(g(x_0))Dg(x_0)$$

Entonces debemos hacer ver que $f \circ g$ es diferenciable en x_0.

$$R_{f \circ g}(x_0, h) = (f \circ g)(x_0 + h) - (f \circ g)(x_0) - Df(g(x_0))Dg(x_0)h$$

O lo que es lo mismo:

$$R_{f \circ g}(x_0, h) = f(g(x_0 + h)) - f(g(x_0)) - Df(g(x_0))Dg(x_0)h$$

$R_{f\circ g}(x_0,h)=f(g(x_0)+k))-f(g(x_0))-Df(g(x_0))Dg(x_0)h$
$R_{f\circ g}(x_0,h)=f(y_0+k))-f(y_0)-Df(y_0)Dg(x_0)h$

Para demostrar que $f\circ g$ es diferenciable en x_0 debemos hacer ver que $\underset{\|h\|\to 0}{Lim}\dfrac{\|R_{f\circ g}(x_0,h)\|}{\|h\|}=0$

Una fórmula de Taylor para $f(y_0+k)$ al primer orden es:

$f(y_0+k)=f(y_0)+Df(y_0)k+R_f(y_0,k)$, entonces:

$R_{f\circ g}(x_0,h)=f(y_0)+Df(y_0)k+R_f(y_0,k)-f(y_0)-Df(y_0)Dg(x_0)h$
$R_{f\circ g}(x_0,h)=Df(y_0)k+R_f(y_0,k)-Df(y_0)Dg(x_0)h$, como:
$k=Dg(x_0)h+R_g(x_0,h)$, entonces:

$R_{f\circ g}(x_0,h)=Df(y_0)\lfloor Dg(x_0)h+R_g(x_0,h)\rfloor+R_f(y_0,k)-Df(y_0)Dg(x_0)h$

$R_{f\circ g}(x_0,h)=Df(y_0)Dg(x_0)h+Df(y_0)R_g(x_0,h)+R_f(y_0,k)-Df(y_0)Dg(x_0)h$
$R_{f\circ g}(x_0,h)=Df(y_0)R_g(x_0,h)+R_f(y_0,k)$, o lo que es lo mismo:

$\|R_{f\circ g}(x_0,h)\|=\|Df(y_0)R_g(x_0,h)+R_f(y_0,k)\|$

Utilizando la desigualdad triangular:

$\|R_{f\circ g}(x_0,h)\|\le\|Df(y_0)R_g(x_0,h)\|+\|R_f(y_0,k)\|$

Utilizando la propiedad matricial: $\|Ax\|\le M\|x\|$, donde A es una matriz y M un número cualquiera, tenemos:

$\|R_{f\circ g}(x_0,h)\|\le M\|R_g(x_0,h)\|+\|R_f(y_0,k)\|$, dividiendo todo para $\|h\|>0$:

$0\le\dfrac{\|R_{f\circ g}(x_0,h)\|}{\|h\|}\le M\dfrac{\|R_g(x_0,h)\|}{\|h\|}+\dfrac{\|R_f(y_0,k)\|}{\|h\|}$; por otro lado:

$$\frac{\left\| R_f(y_0,k) \right\|}{\|h\|} = \frac{\left\| R_f(y_0,k) \right\|}{\|k\|} \frac{\|k\|}{\|h\|}$$

$$\frac{\left\| R_f(y_0,k) \right\|}{\|h\|} = \frac{\left\| R_f(y_0,k) \right\|}{\|k\|} \frac{\left\| Dg(x_0)h + R_g(x_0,h) \right\|}{\|h\|}$$

Aplicando la desigualdad triangular y matricial antes expuesta:

$$\frac{\left\| R_f(y_0,k) \right\|}{\|h\|} \le \frac{\left\| R_f(y_0,k) \right\|}{\|k\|} \left[\frac{\left\| Dg(x_0)h \right\|}{\|h\|} + \frac{\left\| R_g(x_0,h) \right\|}{\|h\|} \right]$$

$$\frac{\left\| R_f(y_0,k) \right\|}{\|h\|} \le \frac{\left\| R_f(y_0,k) \right\|}{\|k\|} \left[\widetilde{M} \frac{\|h\|}{\|h\|} + \frac{\left\| R_g(x_0,h) \right\|}{\|h\|} \right]$$

$$\frac{\left\| R_f(y_0,k) \right\|}{\|h\|} \le \frac{\left\| R_f(y_0,k) \right\|}{\|k\|} \left[\widetilde{M} + \frac{\left\| R_g(x_0,h) \right\|}{\|h\|} \right]$$, remplazando:

$$0 \le \frac{\left\| R_{f \circ g}(x_0,h) \right\|}{\|h\|} \le M \frac{\left\| R_g(x_0,h) \right\|}{\|h\|} + \frac{\left\| R_f(y_0,k) \right\|}{\|k\|} \left[\widetilde{M} + \frac{\left\| R_g(x_0,h \right\|}{\|h\|} \right]$$

Como:

$\lim\limits_{\|h\| \to 0} \frac{\left\| R_g(x_0,h) \right\|}{\|h\|} = 0$, por ser $g(x)$ diferenciable en x_0, y:

$\lim\limits_{\|k\| \to 0} \frac{\left\| R_f(y_0,k) \right\|}{\|k\|} = 0$, por ser $f(x)$ diferenciable en x_0.

Entonces:

$$\lim_{\|h\| \to 0} \frac{\left\| R_{f \circ g}(x_0,h) \right\|}{\|h\|} = 0 \cdot$$

Esto demuestra que $f \circ g$ es diferenciable en x_0, y el teorema de la regla de la cadena queda demostrado. ∀

4.2 EXTREMOS DE FUNCIONES ESCALARES.

Sea $f(x):U \subset R^n \to R$; $x_o \in U$, $\beta_n(x_o,\rho)$ una vecindad de x_o en R^n; $\rho > 0$, si $f(x) \leq f(x_o)$ $\forall x \in \beta_n(x_o,\rho)$ entonces se dice que en x_o hay un ***valor máximo*** local de $f(x)$ que es $f(x_0)$.

Si; por el contrario, $f(x) \geq f(x_o)$ $\forall x \in \beta_n(x_o,\rho)$ entonces se dice que en x_o hay un ***valor mínimo local*** de $f(x)$ que es $f(x_0)$; si $f(x)$ es tal que en ciertas direcciones es un máximo local y en otras direcciones es un mínimo local, entonces se lo llama ***punto de silla***.

A los extremos de una función escalar se los llama también valores óptimos de la función escalar o extremos relativos de la función escalar.

Teorema 4-1 (condición necesaria de óptimo)

> Sea $f(x):U \subset R^n \to R$, $x_0 \in U$. Si en x_o $f(x)$ tiene un extremo local, entonces: $\nabla f(x_0) = 0$

Se llama condición necesaria de óptimo porque:

$$\begin{bmatrix} \text{Si en } x_o \text{ hay un} \\ \text{extremo local} \end{bmatrix} \quad \Rightarrow \quad [\nabla f(x_o) = 0]$$

$$[\nabla f(x_o) = 0] \quad \not\Rightarrow \quad \begin{bmatrix} \text{En } x_0 \text{ exista un} \\ \text{extremo local} \end{bmatrix}$$

⇨ Demostración:

Supongamos que $\nabla f(x_0) \neq 0$.

La derivada direccional de $f(x)$ en x_0, en la dirección del gradiente:

$$f'(xo;\nabla f(xo)) = (\nabla f(x_0) \bullet \nabla f(x_0) = \|\nabla f(xo)\|^2 > 0$$

$$f'(xo;\nabla f(xo)) = \lim_{h\to 0} \frac{f(xo + h\nabla f(xo)) - f(xo)}{h} > 0$$

Por lo tanto:

$\lim\limits_{h\to 0} f(xo + h\nabla f(xo)) - f(xo) > 0$, que implica:

$f(xo + h\nabla f(xo)) > f(xo) \quad \Rightarrow$ en x_0 no hay un máximo.

Como esto no es cierto, y en x_0 si hay un máximo, encontramos un absurdo que demuestra que evidentemente $\nabla f(x_0) = 0$.

Ejemplo 4-5 Encontrar los extremos relativos de la función $f(x,y) = x^2 + y^2$

Solución:

$$\nabla f(x,y) = \begin{bmatrix} 2x & 2y \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

$$\begin{cases} 2x = 0 \\ 2y = 0 \end{cases} \Rightarrow (x,y) = (0,0) \qquad ;$$

Como se ve en la *figura 4-1* el paraboloide tiene un valor mínimo en el punto $(0,0)$ que es $f(0,0) = 0$ y además es un extremo absoluto de esta función. ✓

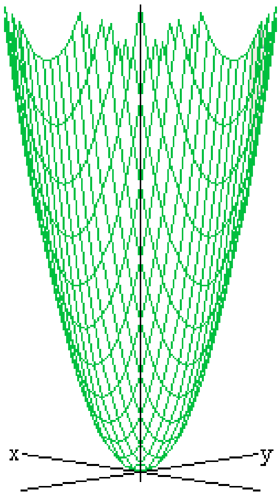

Figura 4-1

4.3 LA MATRIZ HESSIANA COMO CALIFICADORA DE LA NATURALEZA DE EXTREMOS LOCALES.

La matriz Hessiana es una matriz cuadrada y simétrica, como lo vimos en la sección 4-1, y puede estar expresada en forma diagonal o no; cuando está en forma diagonal es porque tiene valores diferentes de cero sólo en la diagonal principal y el resto de valores son cero, cuando esto no se cumple la matriz no es diagonal.

Si la matriz Hessiana es diagonal quiere decir que todas las derivadas parciales mixtas de la función son cero y las dobles no; este razonamiento es muy importante para poder demostrar la forma como sirve la matriz Hessiana para calificar la naturaleza de los valores extremos. Ahora, si la matriz Hessiana es diagonal es importante el signo que tienen los valores que están en la diagonal principal; esto nos lleva hacer la siguiente clasificación de la matriz Hessiana:

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^{\,2}} & 0 & 0 & \cdots & 0 \\ 0 & \frac{\partial^2 f}{\partial x_2^{\,2}} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \frac{\partial^2 f}{\partial x_n^{\,2}} \end{bmatrix} \Rightarrow \boxed{\begin{array}{c}\textbf{Matriz Hessiana} \\ \textbf{diagonal de dimensión} \\ n\times n\end{array}}$$

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^{\,2}} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^{\,2}} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \frac{\partial^2 f}{\partial x_3 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^{\,2}} \end{bmatrix} \Rightarrow \boxed{\begin{array}{c}\textbf{Matriz Hessiana no} \\ \textbf{diagonal de dimensión} \\ n\times n\end{array}}$$

Si la matriz Hessiana es diagonal se la clasifica en las siguientes categorías de acuerdo al signo de los elementos de la diagonal principal:

1. Si $\forall i,\ \frac{\partial^2 f}{\partial x_i^{\,2}} > 0$, (todos los términos de la diagonal principal son positivos) se dice que es ***"DEFINIDA POSITIVA".***
2. Si $\forall i,\ \frac{\partial^2 f}{\partial x_i^{\,2}} < 0$, (todos los términos de la diagonal principal son negativos) se dice que es ***"DEFINIDA NEGATIVA".***
3. Si $\forall i,\ \frac{\partial^2 f}{\partial x_i^{\,2}} \geq 0$, (todos los términos de la diagonal principal son no negativos) se dice que es ***"SEMI DEFINIDA POSITIVA".***
4. Si $\forall i,\ \frac{\partial^2 f}{\partial x_i^{\,2}} \leq 0$, (todos los términos de la diagonal principal son no positivos) se dice que es ***"SEMI DEFINIDA NEGATIVA".***
5. Términos no negativos y no positivos en la diagonal principal se dice que es: ***"NO DEFINIDA NI POSITIVA NI NEGATIVA"***.

Cuando la matriz Hessiana no es diagonal se deben calcular sus autovalores o valores característicos, que por su naturaleza de ser simétrica son reales, y representan los términos de la diagonal principal de la matriz diagonalizada. Entonces, para ubicarla en cualquiera de las categorías anteriores se lo hará en función del signo de los autovalores que serían los términos de la diagonal principal, así:

Sea λ_i el i-ésimo auto valor de la matriz:

1. Si $\forall i,\ \lambda_i > 0$, (todos los autovalores son positivos) se dice que es ***"DEFINIDA POSITIVA".***
2. Si $\forall i,\ \lambda_i < 0$, (todos los autovalores son negativos) se dice que es ***"DEFINIDA NEGATIVA".***
3. Si $\forall i,\ \lambda_i \geq 0$, (todos los autovalores son no negativos) se dice que es ***"SEMI DEFINIDA POSITIVA".***
4. Si $\forall i,\ \lambda_i \leq 0$, (todos los autovalores son no positivos) se dice que es ***"SEMI DEFINIDA NEGATIVA".***
5. Autovalores no negativos y no positivos se dice que es: ***"NO DEFINIDA NI POSITIVA NI NEGATIVA".***

Teorema 4-2 (Criterio para calificar la naturaleza de los valores extremos)

Sea $f(x): U \subset R^n \to R$, de tipo $C^2(U)$, $x_0 \in U$. Si en x_o $f(x)$ tiene un extremo local y $H[f(x_0)]$ es su matriz Hessiana definida en x_o, entonces:

1. Si $H[f(x_0)]$ es definida positiva, entonces en x_0 hay un valor mínimo de la función $f(x_0)$.
2. Si $H[f(x_0)]$ es definida negativa, entonces en x_0 hay un valor máximo de la función $f(x_0)$.
3. Si $H[f(x_0)]$ es semi-definida positiva, entonces en x_0 "puede existir" un valor mínimo de la función $f(x_0)$.
4. Si $H[f(x_0)]$ es semi-definida negativa, entonces en x_0 "puede existir" un valor máximo de la función $f(x_0)$.
5. Si $H[f(x_0)]$ es no definida, entonces en x_0 hay un punto de silla de la función $f(x_0)$.

⇨ Demostración:

Como en x_0 hay un extremo local de la función, del *teorema 4-1* tenemos que $\nabla f(x_0) = 0$.

Una aproximación cuadrática de la función escalar en una vecindad de x_0 está dada por:

$$f(X) = f(X_0) + \nabla f(X_0) \bullet (X - X_0) + \tfrac{1}{2} H[f(X_0)][X - X_0] \bullet (X - X_0) + R_2(X, X_0)$$

Que por haber en x_0 un extremo local la aproximación queda:

$$f(X) = f(X_0) + \tfrac{1}{2} H[f(X_0)][X - X_0] \bullet (X - X_0) + R_2(X, X_0)$$

O lo que es lo mismo:

$$f(X) - f(X_0) = \tfrac{1}{2} H[f(X_0)][X - X_0] \bullet (X - X_0) + R_2(X, X_0)$$

Si $H[f(x_0)]$ es definida positiva; $\tfrac{1}{2} H[f(X_0)][X - X_0] \bullet (X - X_0) > 0$

Y por supuesto $\tfrac{1}{2} H[f(X_0)][X - X_0] \bullet (X - X_0) + R_2(X, X_0) > 0$, por ser $R_2(X, X_0)$ un infinitésimo; entonces:

$f(X) - f(X_0) > 0$ y eso prueba que en $f(x_0)$ hay un mínimo local.

De igual forma:

Si $H[f(x_0)]$ es definida negativa; $\tfrac{1}{2} H[f(X_0)][X - X_0] \bullet (X - X_0) < 0$

Y por supuesto $\tfrac{1}{2} H[f(X_0)][X - X_0] \bullet (X - X_0) + R_2(X, X_0) < 0$, por ser $R_2(X, X_0)$ un infinitésimo; entonces:

$f(X) - f(X_0) < 0$ y eso prueba que en $f(x_0)$ hay un máximo local.

Para el caso de que la matriz Hessiana sea semi-definida positiva o negativa el razonamiento es así:

$$\tfrac{1}{2} H\left[f(X_0)\right]\left[X - X_0\right] \bullet (X - X_0) \geq 0$$

$$\tfrac{1}{2} H\left[f(X_0)\right]\left[X - X_0\right] \bullet (X - X_0) \leq 0$$

Esto no garantiza que:

$$\tfrac{1}{2} H\left[f(X_0)\right]\left[X - X_0\right] \bullet (X - X_0) + R_2(X, X_0) \geq 0$$

$$\tfrac{1}{2} H\left[f(X_0)\right]\left[X - X_0\right] \bullet (X - X_0) + R_2(X, X_0) \leq 0$$

Por cuanto en este caso la presencia del infinitésimo si afecta el signo de este término y esto no permite aseverar que se pueda tratar de un mínimo o máximo, respectivamente, sino solo afirmar que "puede tratarse" de estos extremos.

Con el razonamiento anterior es obvio que si $H\left[f(x_0)\right]$ es no definida entonces:

$f(X) - f(X_0) \leq 0$ o $f(X) - f(X_0) \geq 0$ y por supuesto se trata de un punto de silla.

En el caso del *ejemplo 4-5* la matriz Hessiana es:

$$H\left[f(x, y)\right] = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Definida positiva lo que indica que en el punto $(0,0)$ hay un valor mínimo de la función y es un extremo absoluto por cuanto la matriz Hessiana ni siquiera depende del valor $(0,0)$ para ser definida positiva sino que lo es en todo el dominio de la función.

Con los teoremas 1 y 2 podemos hacer un resumen del procedimiento tradicional para encontrar y calificar los valores extremos de una función escalar, este procedimiento lo presentamos en la *figura 4-2*:

Hay que tomar en cuenta en los ejercicios que si la matriz Hessiana es semi-definida, debemos probar en todas las direcciones posibles antes de concluir que se trata de un valor máximo o mínimo; en el caso de semi-definida es más fácil negar que se trata de un valor extremo que afirmar que es un valor extremo; por cuanto en el primer caso se trata de probar un cuantificador de existencia mientras que en el segundo caso se trata de probar un cuantificador universal y por lo tanto siempre quedará la duda de lo afirmado.

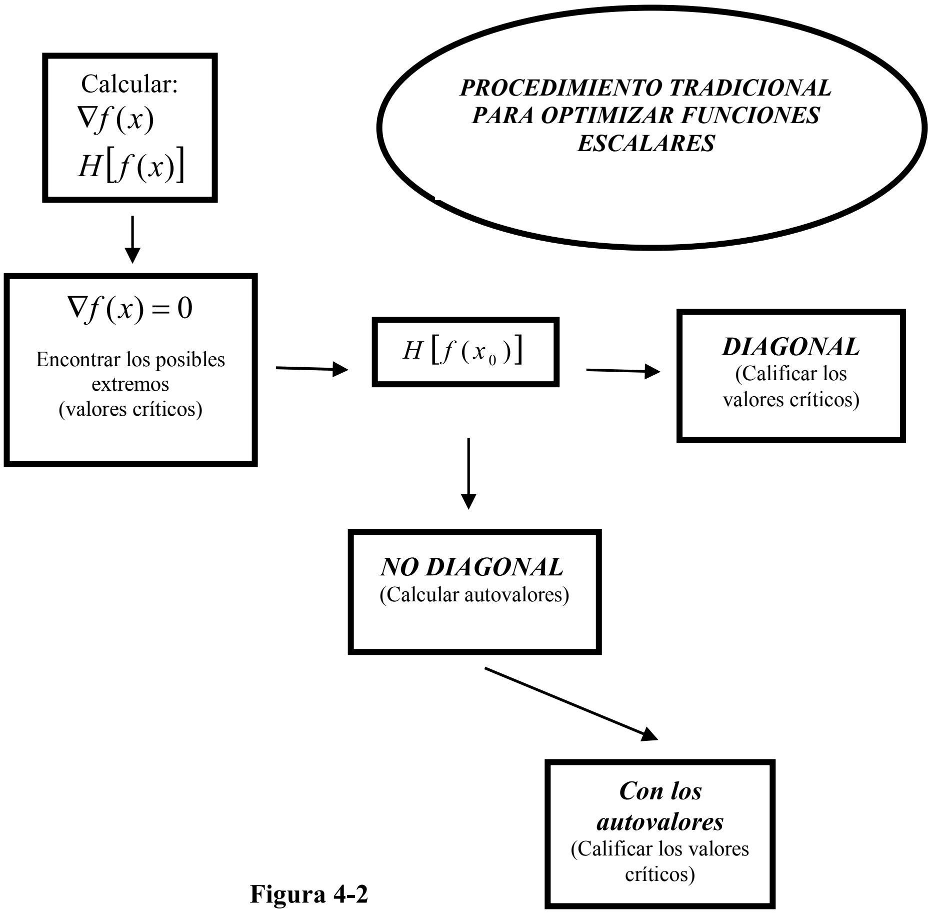

Figura 4-2

Ejemplo 4-6 Encontrar los extremos relativos de la función $f(x, y) = x^2 - y^2$

Solución: $\nabla f(x, y) = (2x, -2y)$

$\nabla f(x, y) = (0,0) \Rightarrow (x, y) = (0,0)$

$H[f(x, y)] = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$, no definida, por lo tanto en $(0,0)$ hay un punto de silla

Autovalores de una matriz

Dada la matriz cuadrada A; se desea resolver las ecuaciones $A\vec{v} = \lambda\vec{v}$ donde λ son los auto valores o valores propios de la matriz A y $\vec{v}$ son los auto vectores o vectores propios de la matriz A correspondientes a λ.

La solución se da si existe al menos un numero real λ y un vector $\vec{v} \neq 0$ que resuelva la ecuación. Esto es:

$$[A - \lambda I]\vec{v} = 0 \text{ ; (vector cero)} \qquad \textbf{4-4}$$

La ecuación 4-4 representa un sistema homogéneo de n ecuaciones con n incógnitas que se resuelve de acuerdo a la regla de Cramer.

Regla de CRAMER

Si en un sistema homogéneo el determinante del sistema es diferente de 0 $\Rightarrow$ que el sistema tiene solución, única y es la solución cero. Si en un sistema homogéneo el determinante del sistema es igual a 0 $\Rightarrow$ que el sistema tiene infinitas soluciones diferentes de la solución nula.

De la regla de CRAMER para que este sistema homogéneo de la ecuación 4-4 tenga "n" soluciones diferentes de cero, tiene que cumplirse que:

$$\det[A - \lambda I] = 0 \qquad \textbf{4-5}$$

La condición 4-5 da una ecuación en λ de grado "n" de cuya solución se obtendrá "n" raíces entre reales e imaginarias. Estas raíces son los autovalores de A.

Ejemplo 4-7 Encontrar los autovalores de la matriz: $A = \begin{bmatrix} 7 & 3 & 0 \\ 3 & 7 & 4 \\ 0 & 4 & 7 \end{bmatrix}$

Solución: $$A - \lambda I = \begin{bmatrix} 7 & 3 & 0 \\ 3 & 7 & 4 \\ 0 & 4 & 7 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\det\begin{bmatrix} 7-\lambda & 3 & 0 \\ 3 & 7-\lambda & 4 \\ 0 & 4 & 7-\lambda \end{bmatrix} = 0$$

$$(7-\lambda)\left((7-\lambda)^2 - 16\right) - 3\left(3(7-\lambda)\right) = 0$$
$$(7-\lambda)\left(49 - 14\lambda + \lambda^2 - 16\right) - 9(7-\lambda) = 0$$
$$(7-\lambda)\left(33 - 14\lambda + \lambda^2\right) - 9(7-\lambda) = 0$$
$$(7-\lambda)\left(24 - 14\lambda + \lambda^2\right) = 0$$
$$(7-\lambda)(\lambda-12)(\lambda-2) = 0$$

$$\begin{cases} \lambda_1 = 7 \\ \lambda_2 = 12 \\ \lambda_3 = 2 \end{cases}$$

Regresando a lo nuestro, esta matriz es definida positiva ✔

Ejemplo 4-8 Encontrar los valores extremos del campo escalar

$$f(x,y) = x^2 + y^2 - 2x + 4y + 5$$

Solución: $\nabla f = \begin{bmatrix} 2x-2 & 2y+4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}$

$$2x - 2 = 0 \ \Rightarrow\ x = 1$$
$$2y + 4 = 0 \ \Rightarrow\ y = -2$$

Valor crítico: (1,-2)

$H\left[f(x,y)\right] = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$; definida positiva

$\therefore$ en (1,-2) hay un mínimo de la función que es: $f(1,-2) = 0$ ✔

Ejemplo 4-9 Encontrar los valores extremos del campo escalar

$$f(x, y) = x^2 - 2xy + y^2 + x^3 + y^3$$

Solución:

$$\nabla f(x, y) = \begin{bmatrix} 2x - 2y + 3x^2 & -2x + 2y + 3y^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{cases} 2x - 2y + 3x^2 = 0 \\ -2x + 2y + 3y^2 = 0 \end{cases}$$

$$3(x^2 + y^2) = 0$$

$$\Rightarrow (0,0) \; Valor \quad crítico$$

$$H[f(x, y)] = \begin{bmatrix} 2 + 6x & -2 \\ -2 & 2 + 6y \end{bmatrix}$$

$$H[f(0,0)] = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$; como no es diagonal, calculamos los autovalores:

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

$$(2 - \lambda)^2 - 4 = 0$$

$$4 - 4\lambda + \lambda^2 - 4 = 0$$

$$\lambda(\lambda - 4) = 0$$

$$\lambda_1 = 0$$

$$\lambda_2 = 4 \; ;$$

Semi definida positiva

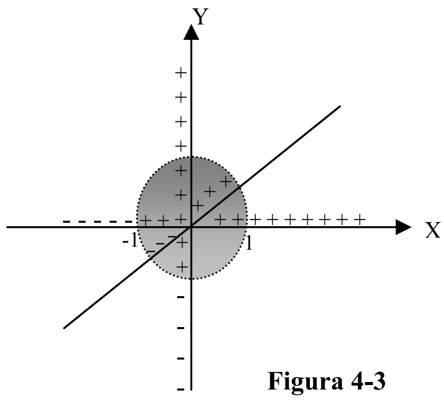

Figura 4-3

Como es semi definida no podemos aseverar que se trate de un mínimo; procedemos a comprobar esto en distintas direcciones.

En la dirección: $y = 0 \quad eje \;\; "X"$; $f(x,0) = x^2 + x^3$

Como se ve en la figura 4-3, en una pequeña vecindad de (0,0) si se trata de un mínimo.

Como se aprecia en la misma figura lo mismo pasa para la dirección del eje "Y" $x = 0$; pero no para el caso de la dirección

$y = x$, donde se aprecia que no cumple, esto nos hace concluir que la función no tiene extremos.

$$x = 0 \quad eje \ \text{"}Y\text{"}; \ f(0, y) = y^2 + y^3$$

$$y = x \qquad ; \ f(x, x) = 2x^3 \qquad \checkmark$$

Ejemplo 4-10 Encontrar los valores extremos del campo escalar:

$$f(x, y) = x^3 - x^2 y + y^2 - x^2$$

Solución:

$$\nabla f = \begin{bmatrix} 3x^2 - 2xy - 2x & -x^2 + 2y \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{cases} 3x^2 - 2xy - 2x = 0 \\ -x^2 + 2y = 0 \end{cases}$$

$$y = \frac{x^2}{2} \Rightarrow 3x^2 - 2x\left(\frac{x^2}{2}\right) - 2x = 0$$

$$3x^2 - x^3 - 2x = 0$$

$$x(3x - x^2 - 2) = 0$$

$$x = 0 \qquad 3x - x^2 - 2 = 0$$

$$x = 2; \quad x = 1$$

Existen tres puntos críticos: $(0,0); \quad (1, \frac{1}{2}); \quad (2,2)$

$$Hf = \begin{bmatrix} 6x - 2y - 2 & -2x \\ -2x & 2 \end{bmatrix} \qquad Hf\Big|_{(0,0)} = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}$$

No definida, $\Rightarrow$ (0,0) es un punto de silla

$$Hf\Big|_{\left(1,-\frac{1}{2}\right)} = \begin{bmatrix} 3 & -2 \\ -2 & 2 \end{bmatrix}; \qquad \det\begin{bmatrix} 3-\lambda & -2 \\ -2 & 2-\lambda \end{bmatrix} = 0$$

$$(3-\lambda)(2-\lambda) - 4 = 0$$

$$6-5\lambda+\lambda^2-4=0$$
$$\lambda^2-5\lambda+2=0$$

$$\lambda=\frac{5\pm\sqrt{25-8}}{2}; \quad \lambda_1=4.56; \quad \lambda_2=0.43$$; definida (+)

En $(1,\frac{1}{2})$ hay un mínimo.

$$Hf\big|_{(2,2)}=\begin{bmatrix}6 & -4\\ -4 & 2\end{bmatrix}; \qquad \det\begin{bmatrix}6-\lambda & -4\\ -4 & 2-\lambda\end{bmatrix}=0$$

$$(6-\lambda)(2-\lambda)-16=0$$
$$12-8\lambda+\lambda^2-16=0$$
$$\lambda^2-8\lambda-4=0$$

$$\lambda=\frac{8\pm\sqrt{64+16}}{2}; \qquad \lambda_1=8.47; \quad \lambda_2=-0.47$$; no definida

En $(2,2)$ hay un punto de silla.

$\Rightarrow$ La función tiene un mínimo en $(1,\frac{1}{2})$ y dos puntos de silla en los puntos $(0,0)$; $(2,2)$. ∀

Ejemplo 4-11 Encontrar los extremos de: $f(x,y)=x^4+y^4-2x^2+4xy-2y^2$

Solución:

$$\nabla f=\begin{bmatrix}4x^3-4x+4y\\ 4y^3+4x-4y\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

$$x^3-x+y=0; \quad \Rightarrow \qquad y=x-x^3$$
$$y^3+x-y=0$$
$$(x-x^3)^3+x-x+x^3=0$$
$$x^3(1-x^2)^3+x-x+x^3=0$$

$$x^3(1-3x^2+3x^4-x^6)+x^3=0$$
$$x^3-3x^5+3x^7-x^9+x^3=0$$
$$x^9-3x^7+3x^5-2x^3=0$$
$$x^3(x^6-3x^4+3x^2-2)=0\,; \quad \Rightarrow \qquad x=0; \quad y=0$$
$$(x^6-3x^4+3x^2-1)-1=0$$

$(x^2-1)^3-1=0$; diferencia de cubos

$$(x^2-1-1)(x^4-2x^2+1+x^2-1+1)=0$$
$$(x^2-2)(x^4-x^2+1)=0$$

$(x^4-x^2+1)=0$; no sirve porque son valores imaginarios

$$\Rightarrow \quad x^2=2 \qquad\qquad x=\pm\sqrt{2}$$

Puntos Críticos (0 , 0) ; ($\sqrt{2}$,- $\sqrt{2}$) ; (- $\sqrt{2}$, $\sqrt{2}$)

$$H(x)=\begin{bmatrix} 12x^2-4 & 4 \\ 4 & 12y^2-4 \end{bmatrix}$$

$$H(0,0)=\begin{bmatrix} -4 & 4 \\ 4 & -4 \end{bmatrix}$$

hay que calcular valores característicos

$$\begin{vmatrix} -4-\lambda & 4 \\ 4 & -4-\lambda \end{vmatrix}=0 \qquad\qquad 16+8\lambda+\lambda^2-16=0$$

$$\lambda_1=0 \qquad\qquad \lambda_2=-8$$

La matriz Hessiana es semidefinida negativa; lo que no garantiza que haya un máximo en este punto.

Analicemos en un entorno de (0 ,0); la *figura 4-4* ilustra gráficamente este análisis sobre las rectas $y=x$; y sobre la recta $y=-x$:

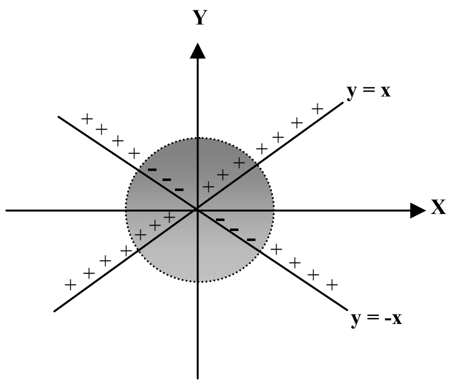

Figura 4-4

$y = x \quad \Rightarrow \quad f(x,x) = 2x^4$; siempre positivos

$y = -x \quad \Rightarrow \quad f(x,-x) = 2x^4 - 8x^2$; negativo dentro de una pequeña vecindad de (0,0).

Esto indica que en (0,0) no existe un máximo.

$$H(\sqrt{2},-\sqrt{2}) = \begin{bmatrix} 20 & 4 \\ 4 & 20 \end{bmatrix}$$

$$\begin{vmatrix} 20-\lambda & 4 \\ 4 & 20-\lambda \end{vmatrix} = 0 \quad \Rightarrow \quad (20-\lambda)^2 - 16 = 0$$

$$\lambda_1 = 24 \qquad \lambda_2 = 16$$

Definido positivo, hay un mínimo

$H(-\sqrt{2},\sqrt{2}) = \begin{bmatrix} 20 & 4 \\ 4 & 20 \end{bmatrix}$; lo mismo, por lo tanto:

Hay mínimos en $(\sqrt{2},-\sqrt{2});(-\sqrt{2},\sqrt{2})$

$f(\sqrt{2},-\sqrt{2}) = (-\sqrt{2},\sqrt{2}) = -8$ ✔

Ejemplo 4-12 Encontrar los extremos de la función:
$f(x, y) = (x - y^2)(2x - y^2) = 2x^2 + y^4 - 3xy^2$

Solución:

$$\nabla f = \begin{bmatrix} 4x - 3y^2 \\ 4y^3 - 6xy \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$H = \begin{bmatrix} 4 & -6y \\ -6y & 12y^2 - 6x \end{bmatrix}$$

$$\nabla f = 0$$
$$4x - 3y^2 = 0 \quad \Rightarrow \quad x = \tfrac{3}{4} y^2$$
$$4y^3 - 6xy = 0\,; \quad 4y^3 - 6(\tfrac{3}{4} y^2) y = 0$$
$$4y^3 - \tfrac{9}{2} y^3 = 0 \quad \Rightarrow \quad x = 0\,; \quad y = 0$$

$$H(0,0) = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{Semidefinida Positiva}$$

Quiere decir que posiblemente se trata de un mínimo.

Veamos que pasa si analizamos la función sobre la recta $y = ax$, la *figura 4-5* representa gráficamente este perfil para un cierto valor de a:

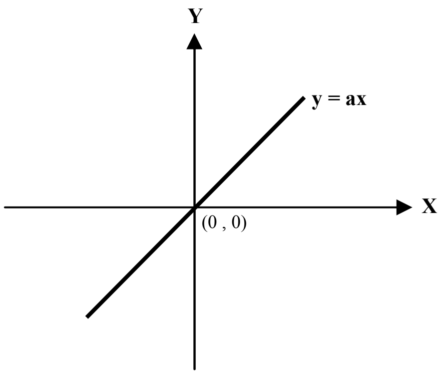

Figura 4-5

$$f(x, ax) = 2x^2 + a^4x^4 - 3a^2x^3$$

$$\hat{f}' = 4x + 4a^4x^3 - 9a^2x^2$$

$$\hat{f}'' = 4 + 12a^4x^2 - 18a^2x$$

$\hat{f}'(0) = 0$; $\hat{f}''(0) = 4 > 0 \Rightarrow$ Indica que la función tiene un mínimo sobre cualquier perfil de la forma $y = ax$.

Sin embargo de esto no es suficiente para asegurar que se trata de un mínimo. La *figura 4-6* representa un análisis gráfico en una pequeña vecindad de (0,0) entre las regiones planas limitadas por las parábolas $x = y^2$ y $x = \frac{1}{2}y^2$:

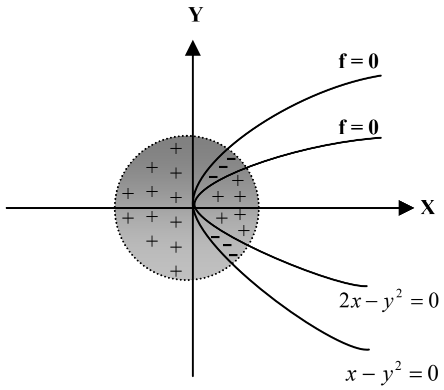

Figura 4-6

Vemos que escribiendo la función en la forma original.
$f(x, y) = (x - y^2)(2x - y^2)$, nos permite análizar los signos de la función en cada una de las subregiones como lo indica la figura.
En esta discución se ve que $(0,0)$ no es un mínimo de la función a pesar de que en los perfiles rectos de la forma $y = ax$ si lo era.

$\therefore$ Esta función no tiene extremos.

Es muy importante notar, como lo dijimos anteriormente y como lo intentan explicar los ejemplos anteriores, que cuando la matriz Hessiana es semidefinida positiva o negativa es bastante difícil demostrar que se trate de mínimos o máximos respectivamente; es más fácil negar esta afirmación si el ejercicio lo permite.

4.4 EXTREMOS CONDICIONADOS.

Es importante que comprendamos lo que se conoce como un problema de extremos condicionados; que lo identificamos como un problema de calcular los valores máximos o mínimos sujetos a un conjunto de restricciones.

Es muy común hablar de problemas condicionados; por ejemplo, en el campo de los negocios, una empresa que produce algunos artículos necesita saber cuales deben ser los niveles de producción de cada uno de los artículos que produce, para maximizar sus utilidades, sujeto a restricciones de capital disponible, materia prima por disponibilidad de proveedores, capacidad de producción instalada, características de mercado y muchas otras restricciones más que son muy comunes en el mundo de los negocios, este ejemplo y muchos más que son frecuentes en este campo son problemas de extremos condicionados.

Un problema condicionado lo podemos formular así:

Optimizar: $z = f(X)$ ***Función objetivo***

Sujeto a: $\begin{cases} g_1(X) = b_1 \\ g_2(X) = b_2 \\ \quad . \\ g_k(X) = b_k \end{cases}$ ***Conjunto de restricciones***

Donde $z = f(X): R^n \to R$, y $g_i(X) = b_i$ para $i = 1,2,......,k$ son superficies de nivel en R^n.

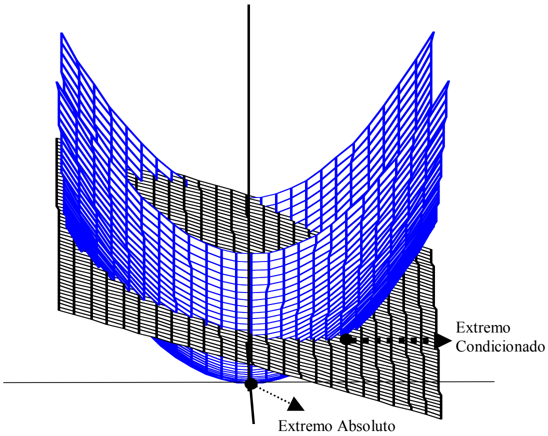

Figura 4-7

La figura 4-7 representa un problema condicionado:

El mínimo de $z = x^2 + y^2$, sujeto a la restricción:

$ax + by + cz + d = 0$ (plano π)

El punto P_1 es el mínimo absoluto de la función y el punto P_2 representa el mínimo condicionado al plano π, como puede apreciar en este gráfico resolver este problema condicionado, geométricamente, es encontrar el mínimo de la función sobre el corte o intersección de la superficie con el plano

El conjunto de restricciones no necesariamente son igualdades, pueden también ser desigualdades de la forma:

$$\begin{cases} g_1(X) \le b_1 \\ g_2(X) \le b_2 \\ \quad . \\ g_k(X) \le b_k \end{cases} \text{ o de la forma } \begin{cases} g_1(X) \ge b_1 \\ g_2(X) \ge b_2 \\ \quad . \\ g_k(X) \ge b_k \end{cases} \text{ o combinando las desigualdades}$$

con las igualdades.

Existen algunos métodos para resolver este tipo de problemas condicionados, de los cuales los más comunes y que estudiaremos en este libro son el de Los **Multiplicadores de Lagrange**, para el caso de que sólo existan restricciones de igualdad y el de las **condiciones de Kuhn Tucker** para el caso de que existan desigualdades en el conjunto de restricciones, analizaremos cada uno de estos métodos por separado.

METODO DE LOS MULTIPLICADORES DE LAGRANGE

Consideremos el problema:

$$\text{Optimizar } z = f(X)$$

$$\text{Sujeto a: } \begin{cases} g_1(X) = b_1 \\ g_2(X) = b_2 \\ \quad . \\ g_k(X) = b_k \end{cases}$$

Podemos construir una función que asocie la función objetivo del problema original y el conjunto de restricciones de la siguiente manera:

$$L(x_1, x_2, \ldots\ldots, x_n, \lambda_1, \lambda_2, \ldots\ldots\ldots, \lambda_k) = f(X) - \lambda_1(g_1(X) - b_1) - \lambda_2(g_2(X) - b_2) - \ldots\ldots - \lambda_k(g_k(X) - b_k)$$

Esta nueva función no altera el valor de la función objetivo original porque se esta restando el valor: $\sum_{i=1}^{k} \lambda_i (g_i(X) - b_i)$, que es cero por contener todas las restricciones igualadas a cero.

A la función $L(x_1, x_2, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_k) : R^{n+k} \to R$, se la conoce como función de Lagrange y los coeficientes λ_i se los conoce como Multiplicadores de Lagrange, el óptimo de $L(x_1, x_2, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_k)$ es el óptimo del problema condicionado.

Aplicando la condición necesaria de óptimo a la función de Lagrange obtenemos:

$\nabla L = 0$, esto implica:

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} - \lambda_1 \frac{\partial g_1}{\partial x_i} - \lambda_2 \frac{\partial g_2}{\partial x_i} - .... - \lambda_k \frac{\partial g_k}{\partial x_i} = 0; \quad i = 1,2,.....,n$$

$$\frac{\partial L}{\partial \lambda_j} = g_j(X) - b_j = 0; \quad j = 1,2,....,k$$

El método consiste en encontrar los valores de x y de λ que satisfagan las condiciones:

$$\frac{\partial L}{\partial x_i} = 0; \qquad i = 1,2,....,n$$

$$\frac{\partial L}{\partial \lambda_j} = 0; \qquad j = 1,2,....,k$$

Esto lleva a resolver un sistema de $n + k$ ecuaciones con $n + k$ incógnitas, donde por supuesto daremos prioridad al cálculo de las x_i que es lo que generalmente interesa en el problema original.

Ejemplo 4-13 Encontrar los extremos de la función:

$f(x, y, z) = 3x^2 + 4y^2 + z^2$,

sujeto a: $x^2 + y^2 + z^2 - 1 = 0$

Solución:

Es un problema de extremos condicionados con una sola restricción que también se lo puede resolver haciendo una substitución directa de la restricción y trabajando sobre el corte de la siguiente forma:

$\hat{f}(x,y) = 3x^2 + 4y^2 + 1 - x^2 - y^2$ o
$\hat{f}(x,y) = 2x^2 + 3y^2 + 1$,

$\nabla\hat{f}(x,y) = (4x, 6y) = (0,0) \Rightarrow \begin{cases} 4x = 0 \\ 6y = 0 \end{cases}$, la solución es:

$(0,0,1)$, calculemos la matriz Hessiana para probar si es máximo o mínimo:

$H[\hat{f}(x,y)] = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$; definida positiva, lo que indica que el punto antes mencionado es un mínimo.

Ahora procedamos por el método de Los Multiplicadores de Lagrange, la función de Lagrange es de la forma:

$$L(x,y,z,\lambda) = 3x^2 + 4y^2 + z^2 - \lambda(x^2 + y^2 + z^2 - 1)$$
$$\nabla L = 0$$

$$\begin{cases} 6x - 2x\lambda = 0 \\ 8y - 2y\lambda = 0 \\ 2z - 2z\lambda = 0 \\ x^2 + y^2 + z^2 = 1 \end{cases}$$

Existen 6 puntos diferentes de la forma (x,y,z,λ) que satisfacen este sistema:
$(\pm1,0,0,\mp3); (0,\pm1,0,\mp4); (0,0,\pm1,\mp1)$,
probemos estos valores:

$$f(\pm1,0,0) = 3$$
$$f(0,\pm1,0) = 4$$
$$f(0,0,\mp1) = 1$$

Aquí se ve que $(0,0,\pm 1)$ son mínimos del problema condicionado y $(0,\pm 1,0)$ son valores máximos del problema condicionado. ∀

Como podemos observar en el ejemplo anterior 4-13, por substitución directa perdimos los extremos $(0,0,-1)$ que también es un mínimo y los extremos $(0,\pm 1,0)$ que son máximos locales; lo cual indica el cuidado que se debe tener cundo se aplica substitución directa que es aplicable siempre que se tenga una sola restricción de igualdad; pues, si se tiene mas de una restricciones de igualdad este método ya se dificulta por cuento se tendría que trabajar sobre el conjunto solución de todas las restricciones.

Ejemplo 4-14 Encontrar los extremos de la función: $f(x,y) = xy$

sujeto a: $x^2 + y^2 = 1 \Rightarrow y = \sqrt{1-x^2}$

Solución: Por sustitución directa:

$$f(x,y) = x\sqrt{1-x^2} = (x^2 - x^4)^{\frac{1}{2}} = f(x)$$

$$f'(x) = \tfrac{1}{2}\left(x^2 - x^4\right)^{-\frac{1}{2}}\left(2x - 4x^3\right) = 0$$

$$= \frac{2x-4x^3}{2\sqrt{x^2-x^4}} = \frac{1-2x^2}{\sqrt{1-x^2}} = 0 \Rightarrow x = \pm\frac{\sqrt{2}}{2} \Rightarrow y = \pm\frac{\sqrt{2}}{2}$$

Los extremos son:

$$\left(\tfrac{\sqrt{2}}{2},\tfrac{\sqrt{2}}{2}\right),\ \left(-\tfrac{\sqrt{2}}{2},\tfrac{\sqrt{2}}{2}\right),\ \left(\tfrac{\sqrt{2}}{2},-\tfrac{\sqrt{2}}{2}\right),\ \left(-\tfrac{\sqrt{2}}{2},-\tfrac{\sqrt{2}}{2}\right)$$

$f\left(\tfrac{\sqrt{2}}{2},\tfrac{\sqrt{2}}{2}\right) = f\left(-\tfrac{\sqrt{2}}{2},-\tfrac{\sqrt{2}}{2}\right) = \tfrac{1}{2}$, máximos condicionados

$f\left(\tfrac{\sqrt{2}}{2},-\tfrac{\sqrt{2}}{2}\right) = f\left(-\tfrac{\sqrt{2}}{2},\tfrac{\sqrt{2}}{2}\right) = -\tfrac{1}{2}$, mínimos condicionados

Por Lagrange:

$$L(x,y,\lambda) = xy - \lambda\left(x^2 + y^2 - 1\right)$$

$$\nabla L = \begin{bmatrix} y - 2x\lambda & x - 2\lambda y & x^2 + y^2 - 1 \end{bmatrix}$$

$$\begin{cases} y - 2x\lambda = 0 \\ x - 2y\lambda) = 0 \\ x^2 + y^2 = 1 \end{cases}$$; da como solución:

$(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$, $(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$, $(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$, $(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$

Que da la misma solución obtenida por substitución directa. ✓

Ejemplo 4-15 Encontrar la distancia más corta entre la elipse $2x^2 + 3y^2 = 12$ y la recta $x + y = 6$.

Solución:

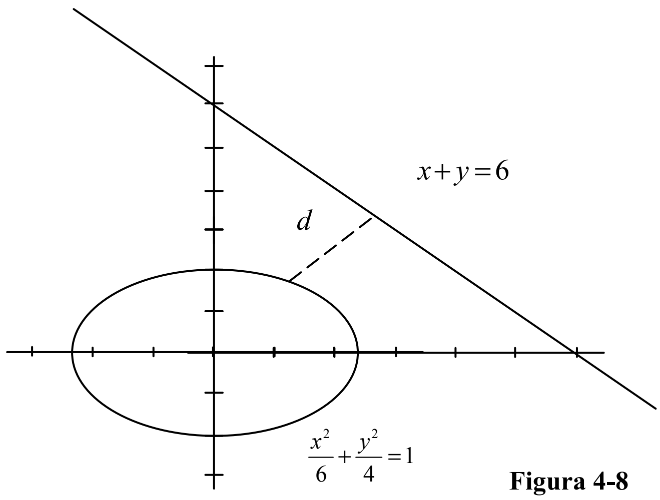

Figura 4-8

Como podemos apreciar en la figura 4-8, considerando que la distancia de un punto a una recta en el plano es: $d = \frac{|ax + by + c|}{\sqrt{a^2 + b^2}}$, se lo puede plantear como un problema condicionado de la forma:

Minimizar : $f(x, y) = \frac{|x + y - 6|}{\sqrt{2}}$

Sujeto a: $2x^2 + 3y^2 = 12$

$$L(x, y, \lambda) = \tfrac{1}{\sqrt{2}}|x + y - 6| - \lambda(2x^2 + 3y^2 - 12)$$

$$\nabla L = \left[\tfrac{\sqrt{2}}{2} - 2x\lambda \quad \tfrac{\sqrt{2}}{2} - 6y\lambda \quad 2x^2 + 3y^2 - 12\right]$$

$$\begin{cases} \frac{\sqrt{2}}{2} - 2x\lambda = 0 \\ \frac{\sqrt{2}}{2} - 6y\lambda = 0 \\ 2x^2 + 3y^2 = 12 \end{cases} \Rightarrow \begin{matrix} x = -1.96 & y = -1.3 & \lambda = +0.09 \\ x = +1.96 & y = +1.3 & \lambda = -0.09 \end{matrix}$$

$f(-1.96,-1.3) = 6.54$

$f(1.96.1.3) = 1.93$

$\Rightarrow$ la mínima distancia entre la elipse y la recta es 1.93.

Ejemplo 4-16 Cuál es el volumen del más grande paralelepípedo que puede ser inscrito en el elipsoide $\frac{x^2}{9} + \frac{y^2}{16} + \frac{z^2}{36} = 1$

Solución:

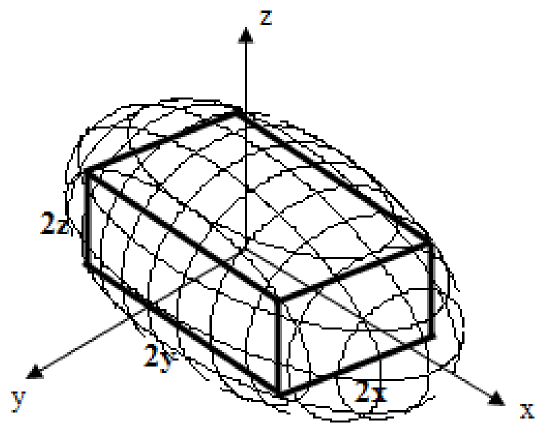

Figura 4-9

Como se puede ver en la figura 4-9, en el sistema cartesiano las dimensiones del paralelepípedo son: $2x \times 2y \times 2z$, el problema condicionado es de la forma:

Maximizar: $8xyz$

Sujeto a: $\frac{x^2}{9} + \frac{y^2}{16} + \frac{z^2}{36} = 1$

$$L(x,y,z,\lambda) = 8xyz - \lambda\left(\frac{x^2}{9} + \frac{y^2}{16} + \frac{z^2}{36} - 1\right)$$

$$\nabla L = \begin{bmatrix} 8yz - \frac{2}{9}x\lambda \\ 8xz - \frac{1}{8}y\lambda \\ 8xy - \frac{1}{18}z\lambda \\ \frac{x^2}{9} + \frac{y^2}{16} + \frac{z^2}{36} - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Esto lleva a resolver el sistema:

$$\begin{cases} 8yz - \frac{2}{9}x\lambda = 0 \\ 8xz - \frac{1}{8}y\lambda = 0 \\ 8xy - \frac{1}{18}z\lambda = 0 \\ \frac{x^2}{9} + \frac{y^2}{16} + \frac{z^2}{36} = 1 \end{cases}$$, hay dos conjuntos de soluciones:

Si $\lambda = 0$; $\begin{matrix} 8yz = 0 \\ 8xz = 0 \\ 8xy = 0 \end{matrix} \Rightarrow \left\{ \begin{matrix} x = 0 & y = 0 & z = 6 \\ x = 0 & y = 4 & z = 0 \\ x = 3 & y = 0 & z = 0 \end{matrix} \right\}$

Todos estos valores dan volumen cero; valores mínimos condicionados; para este caso no nos interesa.

Si $\lambda \neq 0$; $\lambda(\frac{2}{9}x^2 - \frac{1}{8}y^2) = \lambda(\frac{1}{8}y^2 - \frac{1}{18}z^2) = 0$

$\Rightarrow$ $x = \sqrt{3} \quad y = \frac{4}{3}\sqrt{3} \quad z = 2\sqrt{3}$

El volumen máximo es: $64\sqrt{3}$ unid^3

Ejemplo 4-17 Una caja rectangular abierta en la parte superior tiene un volumen de 32 cm^3, cada lado debe tener las dimensiones tales que la superficie total sea máxima. Encontrar las dimensiones.

Solución: Sean x, y, z las dimensiones de la caja:

Minimizar:

$$f(x, y, z) = 2xz + 2yz + xy$$

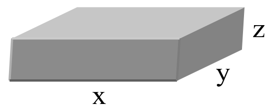

Figura 4-10

Sujeto a: $xyz = 32$

$$L(x, y, z, \lambda) = 2xz + 2yz + xy - \lambda(xyz - 32)$$

$$\nabla L = \begin{bmatrix} 2z + y - yz\lambda \\ 2z + x - xz\lambda \\ 2x + 2y - xy\lambda \\ xyz - 32 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Tenemos que resolver el sistema:

$$\begin{cases} 2z + y - yz\lambda = 0 \\ 2z + x - xz\lambda = 0 \\ 2x + 2y - xy\lambda = 0 \\ xyz = 32 \end{cases} ; \quad \begin{cases} 2xz + xy - xyz\lambda = 0 \\ 2yz + xy - xyz\lambda = 0 \\ 2xz + 2yz - xyz\lambda = 0 \\ xyz = 32 \end{cases}$$

$$\begin{cases} 2xz - 2yz = 0 \\ xy - 2xz = 0 \\ xyz = 32 \end{cases} ; \quad \begin{cases} z(x - y) = 0 \\ x(y - 2z) = 0 \\ xyz = 32 \end{cases}$$

$$x \neq 0; \quad y \neq 0; \quad z \neq 0 \;\Rightarrow\; x = y; \quad y = 2z.$$

$$(y)(y)(\tfrac{y}{2}) = 32 \;\Rightarrow\; y^3 = 64; \quad y = 4$$

$x = 4; \quad y = 4; \quad z = 2$; son las dimensiones

$$Area_{\max} = 48 \text{ cm}^2$$

INTERPRETACIÓN DEL MULTIPLICADOR DE LAGRANGE λ

En la mayoría de las aplicaciones del método de los multiplicadores de Lagrange no es necesario calcular el valor numérico del multiplicador λ; sin embargo ahora analizaremos la importancia de la interpretación del multiplicador λ.

Sea M el valor óptimo de $f(X)$, sujeta a la restricción $g(X) = k$.

Entonces $M = f(X)$ para alguna terna $(x_1, x_2, ..., x_n) \in R^n$ que satisfaga las $n+1$ ecuaciones que resultan de aplicar la condición necesaria de óptimo a la función de Lagrange:

$$\begin{cases} f'_{x_1} = \lambda g'_{x_1} \\ f'_{x_2} = \lambda g'_{x_2} \\ \quad \vdots \\ f'_{x_n} = \lambda g'_{x_n} \\ g(X) = k \end{cases}$$, además las coordenadas de la terna $(x_1, x_2, ..., x_n) \in R^n$,

dependen de k ya que los diferentes niveles de la restricción llevarán por lo general diferentes combinaciones óptimas de x_i; por lo tanto:

$M = f(X)$; donde x_i dependen de k $\Rightarrow$ aplicando la regla de la cadena:

$$\frac{dM}{dk} = \frac{\partial M}{\partial x_1}\frac{dx_1}{dk} + \frac{\partial M}{\partial x_2}\frac{dx_2}{dk} + ... + \frac{\partial M}{\partial x_n}\frac{dx_n}{dk}$$ o lo que es lo mismo:

$$\frac{dM}{dk} = f'_{x_1}\frac{dx_1}{dk} + f'_{x_2}\frac{dx_2}{dk} + ... + f'_{x_n}\frac{dx_n}{dk}$$ por que $M = f(X)$, o:

$$\frac{dM}{dk} = \lambda g'_{x_1}\frac{dx_1}{dk} + \lambda g'_{x_2}\frac{dx_2}{dk} + ... + \lambda g'_{x_n}\frac{dx_n}{dk}$$

$$\frac{dM}{dk} = \lambda(g'_{x_1}\frac{dx_1}{dk} + g'_{x_2}\frac{dx_2}{dk} + ... + g'_{x_n}\frac{dx_n}{dk})$$; o:

$\frac{dM}{dk} = \lambda \frac{dg}{dk}$ aplicando la regla de la cadena,

Como $g(X) = k \Rightarrow \frac{dg}{dk} = 1$ y por supuesto:

$$\boxed{\frac{dM}{dk} = \lambda}$$

Esto quiere decir que λ representa el cambio del valor óptimo de $f(X)$ debido a un incremento unitario de k, que es el margen de la restricción. Visto de otra forma; la variación del valor óptimo de la función con respecto al valor marginal de la restricción.

Ejemplo 4-18 Un fabricante tiene asignado \$60.000,00 para invertir en el desarrollo y la promoción de un nuevo producto. Se ha calculado que si gasta x miles de dólares en desarrollo y y miles de dólares en promoción, se venderán aproximadamente $f(x, y) = 20x^{\frac{3}{2}}y$ unidades del nuevo producto.

a.- ¿Cuánto dinero debe gastar el fabricante en desarrollo y cuánto en promoción para maximizar las ventas?

b.- Supóngase que le aumentan la asignación para invertir en desarrollo y promoción a \$60.200,00. Calcular de qué manera afectará al nivel máximo de ventas los \$200 adicionales.

Solución: **a.**- El problema condicionado será:

Maximizar $f(x, y) = 20x^{\frac{3}{2}}y$

Sujeto a: $x + y = 60$

$$L(x, y, \lambda) = 20x^{\frac{3}{2}}y - \lambda(x + y - 60)$$

$$\nabla L = \begin{bmatrix} 30x^{\frac{1}{2}}y - \lambda \\ 20x^{\frac{3}{2}} - \lambda \\ x + y - 60 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$, esto lleva a resolver el sistema:

$$\begin{cases} 30x^{\frac{1}{2}}y - \lambda = 0 \\ 20x^{\frac{3}{2}} - \lambda = 0 \\ x + y = 60 \end{cases}$$, la solución es: $x = 36 \quad y = 24$

Esto es; para maximizar las ventas, el fabricante debe invertir \$36.000,00 en desarrollo y \$24.000,00 en promoción y venderá aproximadamente 103.680 unidades del nuevo producto.

b.- Como: $\dfrac{dM}{dk} = \lambda$, aplicando diferenciales tenemos:

$\Delta M \approx \dfrac{dM}{dk}\Delta k = \lambda \Delta k$, calculemos λ

$\lambda = 20x^{\frac{3}{2}} = 20(36)^{\frac{3}{2}} = 4.320$, $\Delta k = 0.2$ (miles de dólares)

$\Delta M \approx (4.320)(0.2) = 864$

Lo que quiere decir que las ventas máximas del nuevo producto se incrementarán aproximadamente en 864 unidades, si el presupuesto se aumenta de \$60.000,00 a \$60.200,00 ✓

Hablando de maximización de utilidad sujeto a una restricción presupuestaria, el multiplicador de Lagrange es el cambio aproximado en la utilidad máxima, resultante de un incremento unitario en el presupuesto y los entendidos en esta materia lo conocen como **utilidad marginal del dinero**.

CONDICIONES DE KUHN – TUCKER

Este procedimiento es utilizado cuando el problema condicionado tiene restricciones de desigualdad y esta basado también en el método de Lagrange.

Consideremos el problema:

Maximizar: $z = f(X)$

Sujeto a: $g_i(X) \leq 0$; $i = 1,2,........,m$

Las restricciones de desigualdad pueden transformarse en igualdades aumentándoles una variable no negativa que se la llama variable de holgura, para asegurarnos la no negatividad tomemos esta variable como S_i^2, $i = 1,2,......,m$, entonces el problema queda de la forma:

Maximizar: $z = f(X)$

Sujeto a: $g_i(X) + S_i^2 = 0$; $i = 1,2,........,m$

Este es un problema al que le aplicamos el método de multiplicadores de Lagrange y la función de Lagrange será de la forma:

$$L(x_1, x_2,.., x_n, s_1, s_2,.., s_m, \lambda_1, \lambda_2,.., \lambda_m) = f(X) - \lambda(g(X) + S^2) \quad \textbf{4-6}$$

Donde: $X = (x_1, x_2,..., x_n) \in R^n$

$S = (s_1, s_2,...., s_m) \in R^m$

$\lambda = (\lambda_1, \lambda_2,...., \lambda_m) \in R^m$

$g(X) = (g_1(X), g_2(X),...., g_m(X)) \in R^m$

Dado que: $g_i(X) \leq 0$, una condición necesaria para la optimización es que λ sea no negativa para casos de maximización y que sea no positiva para casos de minimización; esto se justifica de la siguiente manera:

Como vimos anteriormente λ representa la tasa de variación de $f(X)$ con respecto a $g(X)$,

$$\lambda = \frac{\partial f}{\partial g}$$

conforme el lado derecho de la restricción $g(X) \leq 0$ aumenta sobre cero, el <espacio solución se hace menos restringido y, por lo tanto, $f(X)$ no puede disminuir, esto significa que $\lambda \geq 0$. De forma similar para la minimización, conforme aumenta un

recurso, $f(X)$ no puede aumentar, lo que implica que $\lambda \leq 0$. Si las restricciones son igualdades, es decir, $g(X) = 0$ entonces λ es no restringida en signo.

Las restricciones de λ son parte de las condiciones de Kuhn-Tucker, las condiciones restantes las definiremos de la función de Lagrange, sacando las derivadas parciales en la ecuación **4-6** con respecto a X, S y λ tenemos:

$$\frac{\partial L}{\partial X} = \nabla f(X) - \lambda \nabla g(X) = 0$$

$$\frac{\partial L}{\partial S_i} = -2\lambda_i S_i = 0; \qquad i = 1,2,3,......,m$$

$$\frac{\partial L}{\partial \lambda} = -(g(X) + S^2) = 0$$

Del segundo conjunto de estas ecuaciones podemos obtener los siguientes resultados:

1. Si λ_i no es cero, entonces $S_i^2 = 0$, lo que significa que el recurso correspondiente es escaso; por lo tanto, se consume por completo (restricción de igualdad)
2. Si $S_i^2 = 0$, entonces $\lambda_i = 0$, lol que significa que el iésimo recurso no es escaso y, en consecuencia, no afecta el valor de $f(X)$

Del segundo y tercer conjunto de estas ecuaciones se infiere que:

$$\lambda_i g_i(X) = 0; \qquad i = 1,2,3,....,m$$

Esta nueva condición repite el argumento anterior , por cuanto $\lambda_i > 0$, implica $g_i(X) = 0$ o $S_i^2 = 0$. De igual manera, si $g_i(X) < 0$, $S_i^2 > 0$ y $\lambda_i = 0$.

Las condiciones necesarias de Kuhn-Tucker para que X y λ sean un punto crítico del problema de maximización se resumen de la siguiente forma:

$$
\boxed{
\begin{array}{ll}
\lambda = 0 & \\
\nabla f(X) - \lambda \nabla g(X) = 0 & \\
\lambda_i g_i(X) = 0, & i = 1,2,3,\ldots,m \\
g(X) < 0 &
\end{array}
}
$$

Se puede demostrar como ejercicio que estas condiciones también se cumplen para el caso de minimización con la excepción de que λ debe ser no positiva. Tanto en la maximización como en la minimización, los multiplicadores de Lagrange que corresponden a las restricciones de igualdad no deben estar restringidos en signo.

Las condiciones de Kuhn-Tucker son necesarias y suficientes si la función objetivo y el espacio solución satisfacen ciertas condiciones con respecto a la convexidad y a la concavidad, que son las siguientes:

Maximización: función obj. $\Rightarrow$ Cóncava; espacio solución $\Rightarrow$ Conjunto convexo

Minimización: función obj. $\Rightarrow$ Convexa; espacio solución $\Rightarrow$ Conjunto convexo

Tomar en cuenta que en la práctica es más fácil demostrar que una función es convexa o cóncava que demostrar que un conjunto es convexo.

Entonces un problema condicionado general, queda definido de la forma:

Maximizar o minimizar $z = f(X)$

$$
\text{Sujeta a:}\quad
\begin{array}{ll}
g_i(X) \leq 0, & i = 1,2,3,\ldots,r \\
g_i(X) \geq 0, & i = r+1,\ldots,p \\
g_i(X) = 0, & i = p+1,\ldots.,m
\end{array}
$$

$$
L(X,S,\lambda) = f(X) - \sum_{i=1}^{r} \lambda_i \left[g_i(X) + S_i^2 \right] - \sum_{i=r+1}^{p} \lambda_i \left[g_i(X) - S_i^2 \right] - \sum_{i=p+1}^{m} \lambda_i g_i(X)
$$

Donde λ_i es el multiplicador asociado con la restricción i.

Ejemplo 4-19 Considerar el siguiente problema condicionado de minimización:

Minimizar: $f(x, y, z) = x^2 + y^2 + z^2$

Sujeta a:
$$\begin{aligned} g_1(x, y, z) &= 2x + y - 5 \le 0 \\ g_2(x, y, z) &= x + z - 2 \le 0 \\ g_3(x, y, z) &= 1 - x \le 0 \\ g_4(x, y, z) &= 2 - y \le 0 \\ g_5(x, y, z) &= -z \le 0 \end{aligned}$$

Solución: Como es un caso de minimización, $\lambda \le 0$ y las condiciones de Kuhn-Tucker se resumen de la siguiente forma:

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) \le 0$$

$$(2x, 2y, 2z) - (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)\begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = 0$$

$$\lambda_1 g_1 = \lambda_2 g_2 = ... = \lambda_5 g_5 = 0$$

$$g(X) \le 0$$

Estas condiciones generan las siguientes ecuaciones:

$$\begin{aligned} \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5 &\le 0 \\ 2x - 2\lambda_1 - \lambda_2 + \lambda_3 &= 0 \\ 2y - \lambda_1 + \lambda_4 &= 0 \\ 2z - \lambda_2 + \lambda_5 &= 0 \\ \lambda_1(2x + y - 5) &= 0 \\ \lambda_2(x + z - 2) &= 0 \\ \lambda_3(1 - x) &= 0 \end{aligned}$$

$$\lambda_4(2-y)=0$$
$$\lambda_5 z=0$$
$$2x+y\le 5$$
$$x+z\le 2$$
$$x\ge 1,\ y\ge 2,\ z\ge 0$$

La solución de este conjunto de ecuaciones es:

$$x=1,\ y=2,\ z=0,\ \lambda_1=\lambda_2=\lambda_5=0,\ \lambda_3=-2,\ \lambda_4=-4$$

Ya que tanto $f(x,y,z)$ como el conjunto $g(x,y,z)\le 0$ son convexos, $L(x,y,z,\lambda_i,S_i)$ debe ser convexa y el pinto crítico encontrado es un mínimo restringido global. ✓

EJERCICIOS

1.- Dada la función $f(x,y)=e^{2x+3y}$

a) Encontrar una fórmula de Taylor de segundo orden para aproximar esta función en una vecindad del punto *(0,0)*.

b) Estime el error de aproximación en *(0.01, -0.03)*.

2.- Dada la función $f(x,y)=e^{(x-1)^2}\cos y$

a) Encontrar una fórmula de Taylor de segundo orden para aproximar esta función en una vecindad del punto *(1,0)*.

b) Estime el error de aproximación en *(1.2, 0.2)*.

3.- Calcular aproximadamente el valor de:

$$\varphi=\frac{0.98}{\sqrt{15.03+\sqrt[3]{0.97}}}$$

4.- Utilice una aproximación de Taylor para estimar el valor de:

$$\frac{e^{0.03}}{\sqrt{2(0.98)^3+2.02}}$$

Estimar el error de aproximación con tres cifras significativas.

5.- Calcular aproximadamente el valor de:

$$\psi = \frac{2.03^{0.98}}{\sqrt[3]{24.97}}$$

6.- Si q es la capacitancia total de tres capacitores conectados en serie, tal que:

$$\frac{1}{q} = \sum_{n=1}^{n=3} \frac{1}{q_n}$$

Si las medidas de los capacitores son $q_1=25\mu F$; $q_2=40\mu F$; $q_3=50\mu F$; con errores del 0.5% en cada caso, estime el error máximo en el valor de q.

7.- Si el radio de un cilindro aumenta en un 1% y la altura en un 2%, determine el porcentaje en el cual cambia el volumen y el área total de la superficie externa.

8.- Determinar y clasificar los puntos críticos de las siguientes funciones:

a) $f(x,y) = x^2y - x - xy^2 + y$

b) $z = (0.5 - x^2 + y^2)e^{1-x^2-y^2}$

c) $f(x,y) = x^3 - x^2y + y^2 - x^2$

d) $f(x,y) = (x - y^2)(2x - y^2)$

e) $f(x,y,z) = \frac{1}{3}x^3 - x + 2 - y^2 + 2y - z^2 + 2z$

f) $f(x,y,z) = e^{-x^2-y^2-z^2+2y+xz}$

g) $z = senx + seny + sen(x+y)$; en la región $0 \le x \le \frac{\pi}{2}, 0 \le y \le \frac{\pi}{2}$

h) $f(w,x,y,z) = w + \frac{x}{w} + \frac{y}{x} + \frac{z}{y} + + \frac{1}{z}$

9.- La suma de tres números es 50. Determinar el valor de cada uno de ellos para que el producto sea máximo.

10. Sean tres números positivos x, y, z determine el máximo producto de estos tres números, si se sabe que su suma es constante.

11. Utilice este resultado para determinar si es verdadera la siguiente proposición:

$$\sqrt[3]{xyz} \geq \frac{x+y+z}{3}$$

12. Hallar el volumen máximo de un sólido rectangular que tiene la propiedad de que la suma de las área de las seis caras es $6a^2$

13. Un paquete en forma rectangular se puede enviar por correo, si la suma de su longitud y el perímetro de una sección transversal perpendicular a la longitud es igual a 34cm. Encuentre las dimensiones del paquete de máximo volumen que puede ser enviado por correo.

14. Demostrar que un triangulo es equilátero si el producto de los senos de sus ángulos es máximo.

15. Determinar el volumen del paralelepípedo rectangular más grande que puede inscribirse en el elipsoide $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$

16. Encuentre los puntos más cercanos al origen de la superficie $xy^3z^2=16$.

17. ¿Cuál es la distancia mínima entre $C:\begin{cases} x^2+y^2-xy-z^2+1=0 \\ x^2+y^2=1 \end{cases}$ y el origen?

18. Determine el área del paralelogramo de máxima área que se puede inscribir en una elipse de ejes 2 y 3.

19. Hallar la distancia más cercana al origen y la curva $C:\begin{cases} 2z=16-x^2-y^2 \\ x+y=4 \end{cases}$

20. Hallar la distancia mínima entre $9x^2+16y^2=144$ y $5x+8y=40$

21. Sea $T(x,y,z)=100+x^2+y^2$ la temperatura en cada punto de la esfera $x^2+y^2+z^2=50$. Hállese la temperatura máxima en la curva de intersección de la esfera y el plano $x-z=0$

22. Cual es la máxima área que debe tener un rectángulo si la longitud de su diagonal debe ser 2.

23. Obtenga los puntos sobre la curva de intersección del elipsoide $x^2+4y^2+9z^2=36$ y el plano $x-4y-z=0$ que están más cerca del origen y calcular la distancia mínima.

24. Encontrar las dimensiones del paralelepípedo de volumen máximo que puede ser inscrito en el sólido limitado por el paraboloide $\frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{z}{c}$ y el plano $z=c$.
25. Hallar los puntos de la superficie $z^2-xy=1$ más próximos al origen.
26. Halle que dimensiones debe tener una caja rectangular de máximo volumen tal que la suma de su largo, ancho y altura debe ser c.
27. La suma de tres números x, y, z es 100, hállelos de tal modo que el producto $x^a y^b z^c$; donde a, b y c son constantes, sea máximo.
28. Hallar el mayor volumen que puede tener una caja rectangular donde el área total de su superficie debe ser igual a “A”.
29. Encontrar las dimensiones de la caja de máximo volumen que se puede construir al recortar cuatro cuadrados en las esquinas de una plancha cuya área es igual a “A”.