

## Tema 4

# Optimización no lineal



### 4.1. Introducción

En este tema vamos a considerar la optimización de problemas que no cumplen las condiciones de linealidad, bien en la función objetivo, bien en las restricciones. En general, se puede expresar un problema de Programación No Lineal (PNL) de la manera siguiente: encontrar los valores de las variables  $(x_1, x_2, \dots, x_n)$  que

$$\begin{array}{ll} \text{Max. (Min.)} & z = f(x_1, x_2, \dots, x_n) \\ \text{s.a.} & g_1(x_1, x_2, \dots, x_n)(\leq, =, \geq)b_1 \\ & g_2(x_1, x_2, \dots, x_n)(\leq, =, \geq)b_2 \\ & \dots \\ & g_m(x_1, x_2, \dots, x_n)(\leq, =, \geq)b_m \\ & x_i \geq 0, \forall i \end{array}$$

El proceso requiere resolver sistemas de ecuaciones e inecuaciones que se complican mucho incluso para problemas muy simples. No obstante, para los problemas de programación clásica la situación es mucho más sencilla.

En general, un problema de programación no lineal tiene infinitas soluciones factibles, de entre las cuales hemos de encontrar la solución óptima. Vamos a ver que, bajo ciertas hipótesis, ésta ha de cumplir necesariamente ciertas condiciones, lo que en muchos casos nos reducirá la búsqueda a analizar un número finito de puntos.

### 4.2. Programación convexa

Las funciones cóncavas y convexas representan un papel fundamental en la Teoría de la Optimización ya que pueden garantizarnos la globalidad de los óptimos locales. Por ello vamos a iniciar este apartado introduciendo el concepto de función cóncava y convexa

para, posteriormente, introducir condiciones que nos permitan reconocer si una función es cóncava o convexa dependiendo de sus propiedades de diferenciabilidad.

Una función  $f(x)$  es **cóncava** en  $x = a$ , si en alguna pequeña región cercana al punto  $[a, f(a)]$  el gráfico de la función se ubica completamente debajo de su línea tangente.

Una función es **convexa** en  $x = a$ , si en un área cercana a  $[a, f(a)]$  el gráfico esta completamente arriba de su línea tangente.

Una segunda derivada positiva en  $x = a$  denota que la función es convexa en  $x = a$ . Análogamente, una segunda derivada negativa en  $x = a$  denota que la función es cóncava en  $x = a$ .

Figura 4.1: Función convexa en  $x = a$ .

Figura 4.2: Función concava  $x = a$ .

Un **extremo relativo** es un punto en el cual una función esta a un máximo o mínimo. Para ello, la función debe estar en un punto en el cual no esta creciendo ni decreciendo, y por lo tanto, su primera derivada debe ser igual a cero o indefinida. Un punto en el dominio de una función donde la derivada iguala a cero o es indefinida es llamado punto o **valor crítico**.

Figura 4.3: Extremos relativos.

Un **punto de inflexion** es un punto en el gráfico donde la función cruza su línea tangente y cambia de concavo a convexo y viceversa. Los puntos de inflexión pueden ocurrir solo donde la segunda derivada iguala a cero o es indefinida. Es decir,  $f''(a) = 0$ .

Sea la función  $f(x)$ , los pasos o condiciones para obtener el (los) máximo(s) o mínimo(s) relativo(s) serán:

1. Identificar los puntos críticos. Calcular la primera derivada e igualarla a 0,  $\frac{f'(x)}{dx} = 0$ .
2. Calcular la segunda derivada, evaluar los puntos críticos, y revisar los signos. Esta condición es llamada “condición suficiente”. Si un punto crítico es  $a$ , entonces:

- $f''(a) < 0$ , la función es cóncava en  $a$ , por lo que  $a$  es un máximo relativo.
- $f''(a) > 0$ , la función es convexa en  $a$ , por lo que  $a$  es un mínimo relativo.
- $f''(a) = 0$ , el test es inconcluso y es necesario realizar el test de las *derivadas sucesivas*:
  - Si el primer valor diferente de cero de una derivada de orden superior, cuando se evalúa un punto crítico es una derivada de grado impar (tercer, quinto, etc.) la función es un punto de inflexión.
  - Si el primer valor diferente de cero de una derivada de orden superior, cuando es evaluado en un punto crítico es una derivada de grado par, entonces la función es un extremo relativo en  $a$ . Si esta derivada tiene valor negativo entonces la función es cóncava en  $a$  (y por lo tanto, es un máximo relativo). Caso contrario, la función es convexa y presenta un mínimo relativo en  $a$ .

**Ejemplo 38.** *Obtener el extremo relativo de la siguiente función*

$$f(x) = -7x^2 + 126x - 23$$

Se calcula la primera derivada, se iguala a cero y se despeja el valor de  $x$ :

$$f'(x) = -14x + 126 = 0 \quad x = 9 \text{ (valor crítico)}$$

Se calcula la segunda derivada:

$$f''(x) = -14$$

Se evalúa en el punto crítico obtenido, y se estudia el signo

$$f''(9) = -14 < 0 \Rightarrow \text{máximo relativo}$$

**Ejemplo 39.** *Obtener el extremo relativo de la siguiente función*

$$f(x) = 2x^4 + 16x^3 + 32x^2 + 5$$

Se calcula la primera derivada, se iguala a cero y se despeja el valor de  $x$ :

$$f'(x) = 8x^3 - 48x^2 + 64x = 0 \quad x = 0, x = 2, x = 4 \text{ (puntos críticos)}$$

Se calcula la segunda derivada:

$$f''(x) = 24x^2 - 96x + 64$$

Se evalúa en los puntos críticos obtenidos, y se estudia el signo

$$f''(0) = 64 > 0 \Rightarrow \text{mínimo relativo}$$

$$f''(2) = -32 < 0 \Rightarrow \text{máximo relativo}$$

$$f''(4) = 64 > 0 \Rightarrow \text{mínimo relativo}$$

**Ejemplo 40.** *Obtener el extremo relativo de la siguiente función*

$$f(x) = -(x - 8)^4$$

Se calcula la primera derivada, se iguala a cero y se despeja el valor de  $x$ :

$$f'(x) = -4(x - 8)^3 = 0 \quad x = 8 \text{ (punto crítico)}$$

Se calcula la segunda derivada:

$$f''(x) = -12(x - 8)^2$$

Se evalúa en el punto crítico obtenido, y se estudia el signo

$$f''(8) = 0 \Rightarrow \text{el test requiere derivadas sucesivas}$$

$$f'''(x) = -24(x - 8)$$

$$f'''(8) = 0 \Rightarrow \text{test inconcluso}$$

$$f''''(x) = -24$$

$$f''''(8) = -24 < 0 \Rightarrow \text{máximo relativo}$$

## 4.3. Programación no lineal sin restricciones

Para determinar la concavidad (convexidad) de una función  $f(x_1, \dots, x_n)$  de  $n$  variables es necesario definir algunos conceptos previos.

La **matriz hessiana** asociada a una función  $f(\bar{x}) = (x_1, x_2, \dots, x_n)$  es una matriz cuadrada,  $H_{n \times n}$ , tal que sus elementos son de la forma:

$$h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

Denominamos **hessiano** al determinante asociado a la matriz hessiana.

$$H = \begin{vmatrix} h_{11} & h_{12} & \dots & h_{1n} \\ h_{21} & h_{22} & \dots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \dots & h_{nn} \end{vmatrix}$$

El **menor principal de orden  $i$**  de una matriz  $H_{n \times n}$  es el determinante de cualquier matriz  $i \times i$  que se obtiene al suprimir las  $n - i$  filas y las  $n - i$  columnas correspondientes de la matriz.

- Orden 1:  $|H_1| = h_{11}$
- Orden 2:  $|H_2| = \begin{vmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{vmatrix}$
- Orden 3:  $|H_3| = \begin{vmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{vmatrix}$

Así sucesivamente.

**Teorema:** Sea la función  $f(\bar{x})$  con derivadas parciales de segundo orden continuas para cada punto de  $\bar{x} \in S$  (conjunto convexo de soluciones factibles). Entonces  $f(\bar{x})$  es convexa sobre  $S$  si y sólo si, para cada  $\bar{x} \in S$ , todos los menores principales,  $H_i$ , son no negativos.

**Teorema:** Sea la función  $f(\bar{x})$  con derivadas parciales de segundo orden continuas para cada punto de  $\bar{x} \in S$  (conjunto convexo de soluciones factibles). Entonces  $f(\bar{x})$  es cóncava sobre  $S$  si y sólo si, para cada  $\bar{x} \in S$ , los menores principales,  $H_i$ , no nulos tienen el signo que  $(-1)^i$ .

Estudiaremos cómo obtener una solución óptima (si existe) o un extremo local para el siguiente problema de PNL:

$$\begin{array}{ll} \text{Max. (Min.)} & z = f(x_1, x_2, \dots, x_n) \\ \text{s.a.} & \bar{x} = (x_1, x_2, \dots, x_n), \bar{x} \in \mathbb{R}^n \end{array}$$

Supongamos que existen las primeras y las segundas derivadas parciales de  $f(\bar{x})$  y que son continuas en todos los puntos.

Una condición necesaria para que un punto sea un extremo local para el PNL que Si  $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$  es un extremo local para el PNL sin restricciones, entonces

$$\nabla f(\bar{x}) = \frac{\partial f(\bar{x})}{\partial x_i} = 0, \quad \forall i.$$

Para funciones diferenciables, la condición  $\nabla f(\bar{x}) = 0$  es una condición necesaria para que  $f$  tenga un óptimo local en el punto  $\bar{x}$ , que sin embargo no es una condición suficiente, es decir, pueden existir puntos que anulando el gradiente no sean óptimos locales de  $f$ , lo que motiva la siguiente definición.

Un punto  $\bar{x}$  que satisfaga  $\frac{\partial f(\bar{x})}{\partial x_i} = 0$  es un **punto estacionario** o **crítico** de la función  $f(\bar{x})$ .

- Si  $|H_1(\bar{x})|, |H_2(\bar{x})|, \dots, |H_n(\bar{x})| > 0, (i = 1, 2, \dots, n)$ , entonces un punto estacionario será un **mínimo local** para el PNL sin restricciones.

### 4.3. Programación no lineal sin restricciones

---

- Si  $|H_1(\bar{x})| < 0, |H_2(\bar{x})| > 0, |H_3(\bar{x})| < 0 \dots$ , (en general orden impar negativo y par positivo), entonces un punto estacionario será un **máximo local** para el PNL sin restricciones.
- Si  $H_i(\bar{x}) \neq 0, (i = 1, 2, \dots, n)$ , y no se dan ninguno de los casos anteriores,  $f(\bar{x})$  presenta un punto de inflexión o **punto de silla** en ese punto.

Cuando algún  $|H_i(\bar{x})| = 0$  no se puede concluir nada, y habría que utilizar otras técnicas para determinarlo. Para  $n = 2$  el criterio se mejora en el sentido de que si  $|H_1(\bar{x})| = 0$  y  $|H_2(\bar{x})| < 0$ , entonces  $(\bar{x})$  es un punto de silla.

**Ejemplo 41.** Obtener el extremo de  $z = x_1 + 2x_3 + x_2x_3 - x_1^2 - x_2^2 - x_3^2$ .

La condición necesaria para que exista extremo es:  $\frac{\partial z}{\partial x_i} = 0$ , para  $i = 1, 2, 3$ .

$$\frac{\partial z}{\partial x_1} = 1 - 2x_1 = 0, \quad \frac{\partial z}{\partial x_2} = x_3 - 2x_2 = 0, \quad \frac{\partial z}{\partial x_3} = 2 + x_2 - 2x_3 = 0$$

Resolviendo el sistema anterior, obtenemos, para el extremo, el punto objeto de estudio

$$\bar{x} = \left( \frac{1}{2}, \frac{2}{3}, \frac{4}{3} \right).$$

Condición suficiente:

$$H = \begin{pmatrix} \frac{\partial^2 z}{\partial x_1^2} & \frac{\partial^2 z}{\partial x_1 \partial x_2} & \frac{\partial^2 z}{\partial x_1 \partial x_3} \\ \frac{\partial^2 z}{\partial x_2 \partial x_1} & \frac{\partial^2 z}{\partial x_2^2} & \frac{\partial^2 z}{\partial x_2 \partial x_3} \\ \frac{\partial^2 z}{\partial x_3 \partial x_1} & \frac{\partial^2 z}{\partial x_3 \partial x_2} & \frac{\partial^2 z}{\partial x_3^2} \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix}$$

Estudio de los menores principales:

Orden 1:

$$|H_1| = -2 < 0 \Rightarrow \text{orden impar negativo}$$

Orden 2:

$$|H_2| = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4 > 0 \Rightarrow \text{orden par positivo}$$

Orden 3:

$$|H_3| = \begin{vmatrix} -2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{vmatrix} = -8 + 2 = -6 < 0 \Rightarrow \text{orden impar negativo}$$

Por tanto, en el punto  $\bar{x} = \left( \frac{1}{2}, \frac{2}{3}, \frac{4}{3} \right)$  tenemos un máximo. Como el valor de los menores principales no depende del punto, en ese punto tenemos un máximo global.

**Ejemplo 42.** Estudiar los extremos relativos de  $f(x_1, x_2) = x_1^2 + x_1x_2 + x_2^3 - 3x_1 - 2x_2 + 1$ .

Por la condición necesaria de extremo relativo, se tiene que:

$$\left. \begin{aligned} \frac{\partial f}{\partial x_1} &= 2x_1 + x_2 - 3 = 0 \\ \frac{\partial f}{\partial x_2} x_1 + 3x_2^2 - 2 &= 0 \end{aligned} \right\} \quad a = \left(\frac{5}{4}, \frac{1}{2}\right); \quad b = \left(\frac{5}{3}, -\frac{1}{3}\right)$$

Los puntos  $a$  y  $b$  verifican la condición necesaria de extremo.

Veamos la condición suficiente, para lo que se obtienen, a continuación, las derivadas parciales segundas de  $f$  en cada uno de los puntos que cumplen la condición necesaria de extremo,

$$H = \begin{pmatrix} \frac{\partial^2 z}{\partial x_1^2} & \frac{\partial^2 z}{\partial x_1 \partial x_2} \\ \frac{\partial^2 z}{\partial x_2 \partial x_1} & \frac{\partial^2 z}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 6x_2 \end{pmatrix}$$

Estudio de los menores principales:

Orden 1:

$$|H_1| = 2$$

Orden 2:

$$|H_2| = \begin{vmatrix} 2 & 1 \\ 1 & 6x_2 \end{vmatrix}$$

Comprobamos los signos de los menores principales para cada punto:

■ Punto  $a$

- $|H_1(a)| = 2 > 0$
- $|H_2(a)| = \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = 5 > 0$

Por lo que  $a$  es un mínimo local.

■ Punto  $b$

- $|H_1(b)| = 2 > 0$
- $|H_2(b)| = \begin{vmatrix} 2 & 1 \\ 1 & -2 \end{vmatrix} = -5 < 0$

Por lo que  $b$  es un punto de silla.

## 4.4. Programación no lineal con restricciones

### 4.4.1. Restricciones de igualdad: Multiplicadores de Lagrange

Se plantea un nuevo problema, el de optimizar una función sujeta a una restricción de igualdad:

$$\begin{array}{ll}\text{Max. (Min.)} & z = f(x_1, x_2, \dots, x_n) \\ \text{s.a.} & g_1(x_1, x_2, \dots, x_n) = b_1 \\ & g_2(x_1, x_2, \dots, x_n) = b_2 \\ & \dots \\ & g_m(x_1, x_2, \dots, x_n) = b_m\end{array}$$

Para resolver el problema anterior, en primer lugar igualamos a cero cada restricción:

$$\begin{array}{l}b_1 - g_1(x_1, x_2, \dots, x_n) = 0 \\ b_2 - g_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ b_m - g_m(x_1, x_2, \dots, x_n) = 0\end{array}$$

luego le asociamos a cada restricción un multiplicador  $\lambda_i$  ( $i = 1, 2, \dots, m$ ) y formamos la función lagrangiana

$$L(\bar{x}, \bar{\lambda}) = f(x_1, x_2, \dots, x_n) + \sum_{i=1}^m \lambda_i [b_i - g_i(x_1, x_2, \dots, x_n)]$$

donde  $\lambda_i$  ( $i = 1, 2, \dots, m$ ) son constantes (desconocidas) denominadas **multiplicadores de Lagrange**. La función lagrangiana se trata de una función  $L$  cuyas variables son las variables principales del problema más una nueva variable por cada restricción. Puesto que la restricción es siempre igual a cero, el producto  $\lambda_i [b_i - g_i(x_1, x_2, \dots, x_n)]$  también es igual a cero y la suma de tal término no cambia el valor de la función objetivo.

**Condición necesaria de extremo:** Los valores críticos (para los cuales la función es optimizada) son obtenidos tomando las derivadas parciales de  $L$  con respecto a cada una de las tres variables independientes e regalándolas a cero. Es decir, simultáneamente:

$$\begin{array}{l}\frac{\partial L}{\partial x_i} = 0, (i = 1, 2, \dots, n) \\ \frac{\partial L}{\partial \lambda_i} = 0, (i = 1, 2, \dots, m)\end{array}$$

Obteniendo el valor crítico  $x^* = (x_1^*, x_2^*, \dots, x_n^*, \lambda_1^*, \dots, \lambda_m^*)$ .

**Condición suficiente de extremo:**



### 1. Caso de dos variables

- a) **Método de la diferencial segunda:** El problema de la existencia y el carácter del extremo condicional se resuelve averiguando el signo de la segunda diferencial de la función de Lagrange (particularizada para el valor de  $\lambda$  obtenido al resolver el sistema de la condición necesaria).
- b) **Método del Hessiano:** Se halla el hessiano de la función de Lagrange y se evalúa en el punto crítico correspondiente, y solo se puede concluir en el caso de que sea positivo.

$$H = \begin{pmatrix} \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} \\ \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} \end{pmatrix} \Rightarrow \text{si } |H(x_1^*, x_2^*)| > 0 \rightarrow \begin{cases} \frac{\partial^2 L}{\partial x_1^2} > 0 \text{ mínimo condicional} \\ \frac{\partial^2 L}{\partial x_1^2} < 0 \text{ máximo condicional} \end{cases}$$

Es decir, si el hessiano es positivo hay extremo (mínimo o máximo dependiendo del signo de  $\frac{\partial^2 L}{\partial x^2}$ ). En los demás casos hay duda y habría que recurrir a otros métodos.

### 2. Caso de tres o más variables

Se calculan los siguientes determinantes evaluados en los puntos críticos:

$$\Delta_3 = \begin{vmatrix} 0 & \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} \\ \frac{\partial g}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} \\ \frac{\partial g}{\partial x_2} & \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} \end{vmatrix}; \quad \Delta_4 = \begin{vmatrix} 0 & \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} & \frac{\partial g}{\partial x_3} \\ \frac{\partial g}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \frac{\partial^2 L}{\partial x_1 \partial x_3} \\ \frac{\partial g}{\partial x_2} & \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} & \frac{\partial^2 L}{\partial x_2 \partial x_3} \\ \frac{\partial g}{\partial x_3} & \frac{\partial^2 L}{\partial x_3 \partial x_1} & \frac{\partial^2 L}{\partial x_3 \partial x_2} & \frac{\partial^2 L}{\partial x_3^2} \end{vmatrix}; \quad \dots; \Delta_{n+1}$$

- a) Si todos los determinantes tienen **signo negativo**, entonces la función tiene un **mínimo** condicionado en  $x^*$ .
- b) Si los determinantes **tienen signo alterno** (comenzando con un valor positivo), entonces la función tiene un **máximo** condicionado en  $x^*$ .
- c) Si todos los determinantes son **distintos de cero**, pero no cumplen ninguna de las condiciones anteriores, entonces la función **no posee extremo** condicionado en  $x^*$ .
- d) Si algún determinante es **nulo**, no se puede concluir nada.

**Ejemplo 43.** Consideramos el siguiente problema:

$$\begin{aligned} \text{Max.} \quad & z = 500x + 1000y \\ \text{s.a.} \quad & x^2 + \frac{y^2}{4} = 1700 \end{aligned}$$

#### 4.4. Programación no lineal con restricciones

---

La función lagrangiana es  $L(x, y; \lambda) = 500x + 1000y + \lambda(1700 - x^2 - \frac{y^2}{4})$ .

Buscamos los valores críticos derivando con respecto a cada variable e igualamos a cero.

$$\begin{aligned}\frac{\partial L}{\partial x} = 0 &\Rightarrow 500 - 2x\lambda = 0 \\ \frac{\partial L}{\partial y} = 0 &\Rightarrow 1000 - \frac{1}{2}y\lambda = 0 \\ \frac{\partial L}{\partial \lambda} = 0 &\Rightarrow 1700 - x^2 - \frac{y^2}{4} = 0\end{aligned}$$

A continuación se resuelve el sistema de ecuaciones obteniendo los puntos críticos  $x_1^* = (10, 80)$  con  $\lambda_1 = 25$  y  $x_2^* = (-10, -80)$  con  $\lambda_2 = -25$ .

Se calcula el hessiano de la función de Lagrange

$$H = \begin{pmatrix} -2\lambda & 0 \\ 0 & -1/2\lambda \end{pmatrix}$$

Se calcula el determinante del hessiano evaluándolo en cada punto crítico:

$$\lambda = 25 \quad |H(10, 80)| = \begin{vmatrix} -50 & 0 \\ 0 & -25 \end{vmatrix} = 1250 > 0 \Rightarrow \frac{\partial^2 L}{\partial x^2} = -50 < 0 \rightarrow \text{máximo condicional}$$

$$\lambda = -25 \quad |H(-10, -80)| = \begin{vmatrix} 50 & 0 \\ 0 & 25 \end{vmatrix} = 1250 > 0 \Rightarrow \frac{\partial^2 L}{\partial x^2} = 50 > 0 \rightarrow \text{mínimo condicional}$$

**Ejemplo 44.** Consideramos el siguiente problema:

$$\begin{aligned}\text{Max.} \quad & z = 2x - 3y + z \\ \text{s.a.} \quad & x^2 + y^2 + z^2 = 9\end{aligned}$$

La función lagrangiana es:

$$L(x, y, z; \lambda) = 2x - 3y + z + \lambda(9 - x^2 - y^2 - z^2)$$

Buscamos los valores críticos derivando con respecto a cada variable e igualamos a cero.

$$\frac{\partial L}{\partial x} = 0 \Rightarrow 2 - 2\lambda x = 0 \tag{4.1}$$

$$\frac{\partial L}{\partial y} = 0 \Rightarrow -3 - 2\lambda y = 0 \tag{4.2}$$

$$\frac{\partial L}{\partial z} = 0 \Rightarrow 1 - 2\lambda z = 0 \tag{4.3}$$

$$\frac{\partial L}{\partial \lambda} = 0 \Rightarrow -x^2 - y^2 - z^2 + 9 = 0 \tag{4.4}$$

A continuación se resuelve el sistema de ecuaciones:

- Despejamos  $\lambda$  de las ecuaciones (4.5) y (4.6):

$$2 - 2\lambda x = 0 \rightarrow \lambda = \frac{1}{x} \quad -3 - 2\lambda y = 0 \rightarrow \lambda = -\frac{3}{2y}$$

Igualando  $\lambda$  se obtiene:

$$\frac{1}{x} = -\frac{3}{2y} \rightarrow y = -\frac{3}{2x} \quad (a)$$

- Despejamos  $\lambda$  de las ecuaciones (4.5) y (4.7):

$$2 - 2\lambda x = 0 \rightarrow \lambda = \frac{1}{x} \quad 1 - 2\lambda z = 0 \rightarrow \lambda = \frac{1}{2z}$$

Igualando  $\lambda$  se obtiene:

$$\frac{1}{x} = \frac{1}{2z} \rightarrow z = \frac{x}{2} \quad (b)$$

- Ahora reemplazamos (a) y (b) en (4.8) y resolvemos:

$$\begin{aligned} -x^2 - y^2 - z^2 + 9 &= 0 \rightarrow x^2 + \left(-\frac{3}{2x}\right)^2 + \left(\frac{x}{2}\right)^2 - 9 = 0 \rightarrow x^2 + y^2 + z^2 - 9 = 0 \\ &\rightarrow x = \pm\sqrt{18/7} = \pm 1,6 \end{aligned}$$

Por lo tanto las soluciones son:

$$x_1^* = (1,6, -2,41, 0,8) \quad x_2^* = (-1,6, 2,41, -0,8)$$

Se calculan los siguientes determinantes:

$$\Delta_3 = \begin{vmatrix} 0 & 2x & 2y \\ 2x & -2\lambda & 0 \\ 2y & 0 & -2\lambda \end{vmatrix}; \quad \Delta_4 = \begin{vmatrix} 0 & 2x & 2y & 2z \\ 2x & -2\lambda & 0 & 0 \\ 2y & 0 & -2\lambda & 0 \\ 2z & 0 & 0 & -2\lambda \end{vmatrix}$$

Para  $x_1^*$  se tiene que  $\Delta_3 = 41,84 > 0$  y  $\Delta_4 = -46,94 < 0$ , por lo que es un máximo.

Para  $x_2^*$  se tiene que  $\Delta_3 = -41,84 < 0$  y  $\Delta_4 = -65,94 < 0$ , por lo que es un mínimo.

**Ejemplo 45.** Consideramos el siguiente problema:

$$\begin{aligned} \text{Max.} \quad & z = x^2 + y^2 + z^2 \\ \text{s.a.} \quad & x + y + z = 1 \end{aligned}$$

#### 4.4. Programación no lineal con restricciones

---

La función lagrangiana es  $L(x, y, z; \lambda) = x^2 + y^2 + z^2 + \lambda(1 - x + y + z)$ .

Buscamos los valores críticos derivando con respecto a cada variable e igualamos a cero.

$$\frac{\partial L}{\partial x} = 0 \Rightarrow 2x - \lambda = 0 \quad (4.5)$$

$$\frac{\partial L}{\partial y} = 0 \Rightarrow 2y - \lambda = 0 \quad (4.6)$$

$$\frac{\partial L}{\partial z} = 0 \Rightarrow 2z - \lambda = 0 \quad (4.7)$$

$$\frac{\partial L}{\partial \lambda} = 0 \Rightarrow 1 - x - y - z = 0 \quad (4.8)$$

A continuación se resuelve el sistema de ecuaciones obteniendo el punto crítico

$$x^* = \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{2}{3} \right)$$

Se calculan los siguientes determinantes evaluados en el punto crítico:

$$\Delta_3 = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 0 & 2 \end{vmatrix} = -4; \quad \Delta_4 = \begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{vmatrix} = -12$$

Como todos los determinantes tienen signo negativo, entonces la función tiene un mínimo condicionado en  $x^* = \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{2}{3} \right)$ .

#### Caso particular: Función objetivo convexa o cóncava y restricciones lineales:

- Supongamos un problema de **maximización** con restricciones de igualdad. Si  $f(\bar{x})$  es una **función cóncava** y si cada  $g_i(\bar{x})$  es una función lineal, entonces cualquier punto  $(\bar{x}, \bar{\lambda})$  que satisfagan las restricciones proporcionará una solución óptima  $\bar{x} = (x_1, x_2, \dots, x_n)$  para el problema considerado.
- Supongamos un problema de **minimización** con restricciones de igualdad. Si  $f(\bar{x})$  es una **función convexa** y si cada  $g_i(\bar{x})$  es una función lineal, entonces cualquier punto  $(\bar{x}, \bar{\lambda})$  que satisfaga las restricciones proporcionará una solución óptima  $\bar{x} = (x_1, x_2, \dots, x_n)$  para el problema considerado.

**Ejemplo 46.** Una compañía planifica gastar 10000 euros en publicidad. Cuesta 3000 euros un minuto de publicidad en la televisión y 1000 euros un minuto de publicidad en la radio. Si la empresa compra  $x$  minutos de publicidad en la televisión e  $y$  minutos de publicidad en la radio, su ingreso, en miles de euros, está dado por  $f(x, y) = -2x^2 - y^2 + xy + 8x + 3y$ . Plantear y resolver el problema de manera que la empresa maximice sus ingresos.

La formulación del problema no lineal queda de la forma:

$$\begin{array}{ll} \text{Max.} & -2x^2 - y^2 + xy + 8x + 3y \\ \text{s.a.} & 3x + y = 10 \end{array}$$

Formamos la función lagrangiana

$$L(x, y, \lambda) = -2x^2 - y^2 + xy + 8x + 3y + \lambda(10 - 3x - y)$$

Operando tenemos

$$\frac{\partial L}{\partial x} = -4x + y + 8 - 3\lambda = 0 \Rightarrow y = 4x - 8 + 3\lambda$$

$$\frac{\partial L}{\partial y} = -2y + x + 3 - \lambda = 0 \Rightarrow x = 2y - 3 + \lambda$$

$$\frac{\partial L}{\partial \lambda} = (10 - 3x - y) = 0 \Rightarrow 3x + y = 10$$

Vamos a poner las expresiones de  $x$  e  $y$  en función de  $\lambda$ . Sustituyendo  $x$  en  $y = 4x - 8 + 3\lambda$ , tenemos:

$$y = 4(2y - 3 + \lambda) - 8 + 3\lambda \Rightarrow 7y = 20 - 7\lambda \Rightarrow y = 20/7 - \lambda$$

Llevando este valor de  $y$  a  $x = 2y - 3 + \lambda$ , tenemos:

$$x = 2y - 3 + \lambda = 2(20/7 - \lambda) - 3 + \lambda = 19/7 - \lambda$$

Sustituyendo los valores de  $x$  e  $y$  en  $3x + y = 10$ , tenemos:

$$3x + y = 10 = 3(19/7 - \lambda) + 20/7 - \lambda = 11 - 4\lambda \Rightarrow \lambda = 1/4$$

Sustituyendo este valor de  $\lambda = 1/4$  en las expresiones de  $x$  e  $y$ , tenemos:

$$\begin{array}{l} y = 20/7 - \lambda = 20/7 - 1/4 = 73/28 \\ x = 19/7 - \lambda = 19/7 - 1/4 = 69/28 \end{array}$$

La matriz hessiana asociada con la función  $f(x, y) = -2x^2 - y^2 + xy + 8x + 3y$  es

$$H(x, y) = \begin{pmatrix} -4 & 1 \\ 1 & -2 \end{pmatrix}$$

El menor principal de primer orden es negativo y  $|H_{2 \times 2}| = 7 > 0$ , luego  $f(x, y)$  es una función cóncava. La restricción es lineal. Entonces, concluimos que la solución es óptima (máximo) para el problema no lineal.

Así, la empresa tendría que comprar 69/28 minutos (2 min. y 27 seg.) de tiempo de televisión y 73/28 minutos (2 min. y 36 seg.) de tiempo de radio; que le generará un ingreso de 15017.85 euros.

### Interpretación económica de los multiplicadores de Lagrange

Los multiplicadores representan el cambio en la función objetivo para un cambio unitario en la cota de cada restricción. En programación lineal éstos recibían el nombre de variables duales o precios sombra.

#### 4.4.2. Restricciones de desigualdad: Condiciones de Kuhn-Tucker (CKT)

En esta sección se estudian las condiciones necesarias y suficientes para que el punto  $\bar{x} = (x_1, x_2, \dots, x_n)$  sea una solución óptima para el PNL, con restricciones de desigualdad. Son una generalización del método de los multiplicadores de Lagrange para restricciones de desigualdad.

Considérese el siguiente problema:

$$\begin{aligned} \text{Min.} \quad & z = f(\bar{x}) \\ \text{s.a.} \quad & g_1(\bar{x}) \leq b_1 \\ & g_2(\bar{x}) \leq b_2 \\ & \dots \\ & g_m(\bar{x}) \leq b_m \end{aligned} \tag{4.9}$$

En primer lugar se transforman las restricciones de desigualdad a restricciones de igualdad:

$$g_j(\bar{x}) \leq b_j \Rightarrow g_j(\bar{x}) + s_j = b_j$$

De acuerdo a la técnica de los multiplicadores de Lagrange se construye la función lagrangiana:

$$L(\bar{x}, \bar{\lambda}) = f(\bar{x}) + \sum_{j=1}^m \lambda_j [b_j - g_j(x_j) - s_j]$$

Los puntos que minimizan  $f$  sujeta a las restricciones  $g_j \leq b_j$  están dentro de los puntos críticos de  $L$ :

- Que hacen cero las parciales con respecto a las variables  $x_i$  ( $i = 1, \dots, n$ ):

$$\frac{\partial L}{\partial x_i} = 0 \text{ para cada } i = 1, \dots, n.$$

- Que hacen cero las parciales con respecto a las variables  $\lambda_j$  ( $j = 1, \dots, m$ ):

$$\frac{\partial L}{\partial \lambda_j} = g_j(\bar{x}) + s_j \Leftrightarrow g_j(\bar{x}) \leq b_j$$

- Que hacen cero las parciales con respecto a las variables  $s_j$  ( $j = 1, \dots, m$ ):

$$\frac{\partial L}{\partial s_j} = \lambda_j s_j \Leftrightarrow \lambda_j (b_j - g_j(x_j)) = 0$$

Lo anterior se resume en el siguiente teorema que indica las condiciones que deben satisfacer los puntos que minimizan la función sujeta a las restricciones y que se conoce **condiciones de Kuhn-Tucker (CKK)**:

**Teorema 7.** *Supongamos un problema de minimización. Si  $x^*$  es un óptimo, entonces deben existir números reales llamados multiplicadores  $\lambda_j$  **no negativos** tales que  $(x^*, \lambda_1, \dots, \lambda_m)$  es un punto crítico para  $L$ . Es decir, que se cumple:*

1.  $\frac{\partial L}{\partial x_i} = 0$  para cada  $i = 1, \dots, n$ .
2.  $g_j(\bar{x}) \leq b_j$  para cada  $j = 1, \dots, m$
3.  $\lambda_j \geq 0$  para cada  $j = 1, \dots, m$
4.  $\lambda_j (b_j - g_j(\bar{x})) = 0$ .

Cuando el problema es de **maximización**, se opera de la misma forma, pero utilizando  $-f(x)$ . De esta forma el langragiano será:

$$L(\bar{x}, \bar{\lambda}) = -f(\bar{x}) + \sum_{j=1}^m \lambda_j [b_j - g_j(x_j) - s_j]$$

Cuando se tenga condiciones de no negatividad, éstas se tratarán como una restricción más, transformándolas para ello en una restricción  $\leq$ :

$$x_i \geq 0 \Rightarrow -x_i \leq 0$$

La forma de operar las condiciones de KKT ha de ser la siguiente: buscamos es el punto  $x^*$  y para ello se resuelve el sistema de ecuaciones en las variables  $x_i$  y  $\lambda_j$ , obtenido de aplicar las CKT. Las soluciones se revisan una a una para ver cual de ella cumple que los  $\lambda_j$  son no negativos y que también se cumplen las restricciones  $g_j(\bar{x}) < b_j$  en los puntos encontrados. Normalmente se realiza una tabla donde se hace la verificación.

Obsérvese también es posible trabajar el problema de maximización resolviendo el problema de minimización pero conservando aquellos puntos que tengan los valores de los multiplicadores no positivos.

Se puede comprobar que las tablas para minimización y para maximización son idénticas salvo que los valores de los multiplicadores están cambiados de signo. Por tanto, la estrategia conveniente para optimizar una función sujeta a restricciones de desigualdad por el método de las condiciones de KKT será:

#### 4.4. Programación no lineal con restricciones

---

1. Plantear el problema como si se tratara sólo de minimización y resolver el sistema de ecuaciones correspondientes.
2. Eliminar aquellos puntos encontrados que no satisfacen las restricciones  $g_j \leq b_j$ .
3. Eliminar aquellos puntos que tienen a la vez multiplicadores positivos y negativos.
4. Para minimización, escoger dentro de aquellos puntos que tienen multiplicadores no negativos aquél que tienen la menor evaluación de la función objetivo.
5. Para maximización, escoger dentro de aquellos puntos que tienen multiplicadores no positivos aquél que tienen la mayor evaluación de la función objetivo.

Las CKK sólo son aplicables si las  $g_i(\bar{x})$ ,  $(i = 1, 2, \dots, m)$  satisfacen ciertas **condiciones de regularidad**:

- i) Ser lineales e independientes.
- ii) Restricciones linealmente independientes: continuas y los gradientes en la solución óptima forman un sistema de vectores linealmente independientes.

**Ejemplo 47.**

$$\begin{array}{ll} \text{Min.} & z = x^2 + y^2 + y - 1 \\ \text{s.a.} & x^2 + y^2 \leq 1 \end{array}$$

Formamos el langragiano:  $L = x^2 + y^2 + y - 1 + \lambda(1 - x^2 - y^2)$ .

Las CKT son:

1.  $\frac{\partial L}{\partial x} = 2x - 2\lambda x = 0$      $\frac{\partial L}{\partial y} = 2y + 1 - 2\lambda y = 0$
2.  $x^2 + y^2 \leq 1$
3.  $\lambda \geq 0$
4.  $\lambda(1 - x^2 - y^2) = 0$

La siguiente tabla muestra las soluciones para  $x$ ,  $y$  y  $\lambda$ :

$x$	$y$	$\lambda$	$g(x, y) = x^2 + y^2$	$z$
0	-1/2	0	1/4	-5/4
0	-1	-1/2	1	-1
0	1	-3/2	1	1

Obsérvese que:



- Los tres puntos cumplen la restricción  $g(x, y) \leq 1$
- Como estamos minimizando, el único punto que cumple la condición  $\lambda \geq 0$  es el primero. Por tanto,  $(0, -1/2)$  es un mínimo de la función.
- Se pueden utilizar los valores con  $\lambda < 0$  para determinar el posible máximo. En este caso los puntos candidatos son  $(0, -1)$  y  $(0, 1)$ . Se debe escoger como máximo aquel que tenga mayor valor de la función objetivo. En este caso,  $(0, 1)$  es el máximo.

## 4.5. Ejercicios

1. Estudie la concavidad o convexidad de las siguientes funciones:

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

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

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

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

2. Encuentra los extremos de los siguientes problemas:

a) Min.  $x^2 + xy + y^2 - 6x + 2$

b) Max.  $x^2 + xy + y^2 + x + 5y$

c) Max.  $3x^2 + 5y^2 + 5z^2 + 2yz + 6zx - 2xy$

d) Max.  $4x - 6y - x^2 - 2y^2$

3. Sea  $(0, 0, 0)$  un punto crítico de la función  $f$  y su hessiano

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

¿El punto  $(0, 0, 0)$  corresponde a un mínimo local, un máximo local o un punto de silla?

4. En el punto  $(1, 0)$  la función  $f(x, y) = x^3 - 2xy^2 - 3x$ , ¿presenta un máximo local, un máximo global o un punto de silla?

5. Dado el PNL:

$$\begin{array}{ll} \text{Max.} & f(\bar{x}) = x_1^2 + 20x_2 \\ \text{s.a.} & x_1 + x_2 = 100 \end{array}$$

Resuélvalo usando el método de Lagrange.

6. Encuentra los extremos globales de los problemas siguientes:

$$\begin{array}{lll} \text{a) Opt.} & 3x^2 + xy + 4y^2 & \text{b) Opt.} & x^2 + y^2 + z^2 + t^2 & \text{c) Opt.} & x^2 + y^2 + z^2 - 2x - 2y \\ \text{s.a.} & 3x + y = 72 & \text{s.a.} & x + y + z + t = 1 & \text{s.a.} & x + y - 2z = 0 \\ & & & & & y - x = 0 \end{array}$$

$$\begin{array}{ll} \text{d) Opt.} & -x^2 + xy + yz - y^2 - z^2 & \text{e) Opt.} & 4x - y \\ \text{s.a.} & x + z = 2 & \text{s.a.} & y = x^2 \\ & x + y = 4 & & \end{array}$$

7. Considera el problema

$$\begin{array}{ll}\text{Max.} & y \\ \text{s.a.} & x^2 + y^2 \leq 6 \\ & x^2 - y \geq 0 \\ & x, y \geq 0\end{array}$$

- a) Escribe las condiciones de Kuhn-Tucker.
- b) Comprueba que el punto óptimo  $(\sqrt{2}, 2)$  cumple las condiciones de Kuhn-Tucker.

8. Considera el problema

$$\begin{array}{ll}\text{Max.} & (x - 2)^2 + (y - 2)^2 \\ \text{s.a.} & x - y \leq 8 \\ & x + y \geq -4 \\ & x, y \geq 0\end{array}$$

- a) Escribe las condiciones de Kuhn-Tucker.
- b) Comprueba si el punto  $(2, -6)$  verifica dichas condiciones.

9. Resolver el siguiente problema:

$$\begin{array}{ll}\text{Max.} & (x - 7)^2 + (y - 10)^2 \\ \text{s.a.} & y \leq 8 \\ & (x - 10)^2 + (y - 10)^2 \leq 36\end{array}$$

10. Resolver el siguiente problema:

$$\begin{array}{ll}\text{Max.} & x^2 + y^2 + y - 1 \\ \text{s.a.} & x^2 + y^2 \leq 1\end{array}$$

11. Resolver el siguiente problema:

$$\begin{array}{ll}\text{Max.} & -x^2 - y^2 + 17x + 35y \\ \text{s.a.} & x + y \leq 17,25 \\ & x, y \geq 0\end{array}$$