C A P Í T U L O 6

P R O G R A M A C I Ó N N O L I N E A L

6.1.- Introducción a la Programación No Lineal

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,…, x_n) que

$$
\begin{array}{ll}
\text{Max (min) } z = f(x_1, x_2, \ldots, x_n) & \\
\quad \text{st} & \\
\qquad g_1(x_1, x_2, \ldots, x_n)\ (\le, =, \ge)\ b_1 & \\
\qquad g_2(x_1, x_2, \ldots, x_n)\ (\le, =, \ge)\ b_2 & \\
\qquad \ldots & \qquad\qquad (6.1) \\
\qquad g_m(x_1, x_2, \ldots, x_n)\ (\le, =, \ge)\ b_m & \\
\qquad\qquad x_i \ge 0,\ \forall i &
\end{array}
$$

Antes de entrar de lleno en el tema, veamos una serie de conceptos sobre funciones cóncavas y convexas.

6.2.- Funciones cóncavas y convexas

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 luego más tarde introducir condiciones que nos permitan reconocer si una función es cóncava o convexa dependiendo de sus propiedades de diferenciabilidad.

Definición 6.1. Diremos que una función f(x) es *estrictamente cóncava* en un conjunto S convexo si todo segmento que une dos puntos de la gráfica esta estrictamente por debajo de la gráfica.

Definición 6.2. Diremos que una *función es cóncava* (no estricta) si no todas las cuerdas que unen puntos de la gráfica en dicho intervalo quedan estrictamente por debajo.

Definición 6.3. Sea f(x) una función definida en un intervalo de R, diremos que dicha *función es convexa* en el intervalo si todo segmento que une dos puntos de la gráfica queda por encima de la gráfica. Si siempre queda estrictamente por encima, decimos que la función es estrictamente convexa.

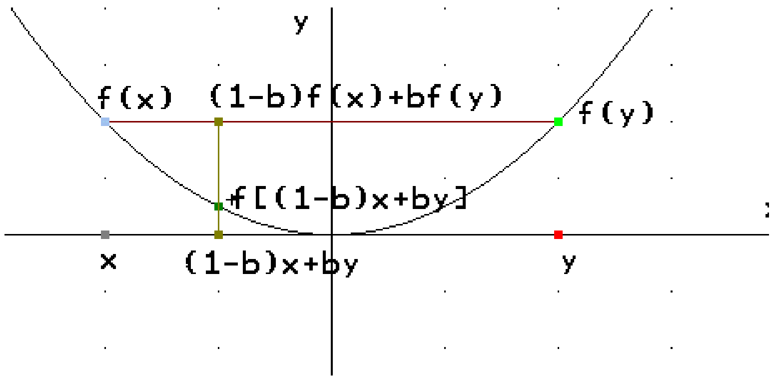
Definición 6.4. Sea S un subconjunto convexo de R^n y sea f: S→ R. Diremos que f(x) es una *función convexa* en S si para cualquier par de puntos x, y∈S y para cualquier b∈[0,1]

$$f[(1-b)x+by] \leq (1-b)f(x)+bf(y)$$

Definición 6.5. f(x) es una *función estrictamente convexa* en S si para cualquier par de puntos x, y∈S y para cualquier b∈[0,1]

$$f[(1-b)x+by] < (1-b)f(x)+bf(y)$$

En la figura siguiente se representa gráficamente una función convexa:



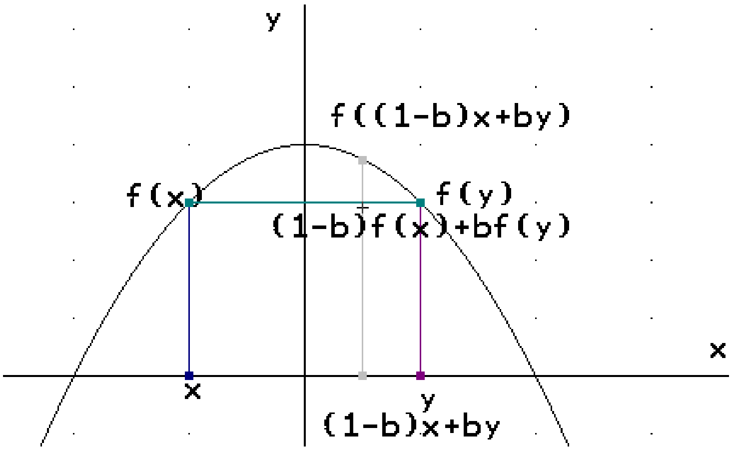
Definición 6.6. f(x) es una función cóncava en S si para cualquier par de puntos x, y∈S y para cualquier b∈[0,1]

$$f[(1-b)x+by] \geq (1-b)f(x)+bf(y)$$

Definición 6.7. f(x) es una *función estrictamente cóncava* en S si para cualquier par de puntos x, y∈S y para cualquier b∈[0,1]

$$f[(1-b)x+by] > (1-b)f(x)+bf(y)$$

En la figura siguiente se representa gráficamente una función cóncava:



Ejercicio 6.1. Demostrar por definición que $f(x)=x^2$ es estrictamente convexa en R.

Para hacer esta demostración, consideremos dos puntos cualesquiera de R, por ejemplo x_1,x_2 . Tenemos que demostrar que

$$f((1-b\ )x_1+ b\ x_2) < (1- b\ )f(x_1)+ b\ f(x_2)$$

$$((1- b\ )x_1+ b\ x_2)^2=(1- b\ )^2x_1^2+ b^2x_2^2+2(1- b)bx_1x_2$$

y comparar con

$$(1- b\ )x_1^2+ b\ x_2^2$$

Obsérvese que no es fácil demostrar la concavidad o convexidad de una función por definición. Por ello es necesario o conveniente disponer de unas condiciones necesarias y suficientes que nos permitan determinar si una función es cóncava o convexa estudiando otros elementos más operativos.

Propiedad 6.1. Si f es cóncava en S, entonces -f es convexa en S y si f es convexa en S, entonces -f es cóncava en S.

Propiedad 6.2. Si f es estrictamente convexa en S, entonces -f es estrictamente cóncava en S y, si f es estrictamente cóncava en S, entonces -f es estrictamente convexa en S .

Propiedad 6.3. Si f_i, i=1,..., n son convexas en S, entonces $\sum_{i=1}^{n} \alpha_i f_i$ con $\alpha_i \geq 0$ es convexa en S. Del mismo modo, si f_i, i=1,...,n son cóncavas en S, entonces $\sum_{i=1}^{n} \alpha_i f_i$ con $\alpha_i \geq 0$ es cóncava en S.

Demostración.

Demostramos sólo la primera: Si f_i es convexa en S para todo i, entonces para cualquier par de puntos x, y∈S y para cualquier b∈[0,1]

$$f_i[(1-b)x+by] \leq (1-b)f_i(x)+bf_i(y)$$

Esto quiere decir que para cualquier $\alpha_i > 0$

$$\alpha_i f_i[(1-b)x+by] \leq \alpha_i((1-b)f_i(x)+bf_i(y))$$

entonces

$$\sum_{i=1}^{n} \alpha_i f_i([(1-b)x+by]) \leq \sum_{i=1}^{n} \alpha_i((1-b)f_i(x)+bf_i(y))$$

y esto último es cierto si y sólo si

$$\sum_{i=1}^{n} \alpha_i f_i([(1-b)x+by]) \leq (1-b)\sum_{i=1}^{n} \alpha_i f_i(x) + b\sum_{i=1}^{n} \alpha_i f_i(y)$$

con lo que queda demostrado que es convexa.

Propiedad 6.4. El producto de funciones cóncavas (convexas) no ha de ser necesariamente una función cóncava (convexa).

Teorema 6.1. Sea S un subconjunto convexo de R^n, sea f:S→ R. Entonces:

a) Si f es convexa en S ⇒ para todo α∈ R, se verifica que el conjunto

   $A_\alpha = \{x \in S \mid f(x) \leq \alpha\}$ es un conjunto *convexo*.

b) Si f es cóncava en S $\Rightarrow$ para todo $\alpha \in R$ se verifica que el conjunto

$B_\alpha=\{x\in S \ / \ f(x) \geq \alpha \}$ es un conjunto *convexo.*

Demostración.

a) Consideremos un número α real cualquiera. Como f es convexa en un conjunto S (convexo de R^n), vamos a demostrar que el conjunto A_α es convexo.

Consideremos dos puntos cualesquiera $x, y\in A_\alpha$ y sea un $b\in[0,1]$

Como dichos puntos son del conjunto A_α, entonces verifican que

$$f(x) \leq \alpha, \quad f(y) \leq \alpha \quad \text{pues } x, y\in A_\alpha$$

Como además f es convexa se verifica que

$$f[(1-b)x+by] \leq (1-b)f(x)+bf(y) \leq (1-b)\alpha+b\alpha=\alpha$$

luego se cumple que $\forall b\in[0, 1]$ y $\forall x, y\in A_\alpha$, $f[(1-b)x+by] \in A_\alpha$, por tanto el conjunto A_α es convexo.

b) Se demuestra de forma análoga.

Propiedad 6.5. *Región factible* para un problema de PNL es el conjunto de valores $x=(x_1, x_2,\ldots, x_n)$ que satisfacen las restricciones de (6.1).

Propiedad 6.6. *Máximo local*: sea S un conjunto convexo tal que $x_1, x_2\in S$, $f(x_1) \geq f(x_2)\ \forall x_2 \Rightarrow x_1$ sería un máximo local.

Propiedad 6.7. *Mínimo local*: sea S un conjunto convexo tal que $x_1, x_2\in S$, tal que $f(x_1) \leq f(x_2)\ \forall x_2 \Rightarrow x_1$ sería un mínimo local.

Propiedad 6.8. Sea el problema Maximización de z en un problema de programación no lineal (PNL). Si para todo x existe x' tal que $f(x') \geq f(x)$, x' sería la solución óptima (cumpliendo la *Propiedad 6.6*).

Propiedad 6.9. Sea el problema Minimización de z en un problema de programación no lineal (PNL). Si para todo x existe x' tal que $f(x')\leq f(x)$, sería la solución óptima (cumpliendo la *Propiedad 6.7*).

Teorema 6.2. Consideremos un PNL de maximización. Supóngase que la región factible, S, para el PNL es un conjunto convexo. Si f(x) es cóncava sobre S, entonces cualquier máximo local de PNL es una solución óptima para el problema de PNL.

Demostración

Supongamos que existe $\bar{x}$ que es máximo local y que no es solución óptima, es decir, $\exists$ x / f(x) > f ($\bar{x}$).

Podemos formar combinación lineal: $f(\alpha\bar{x} + (1-\alpha)\, x) \geq \alpha\, f(\bar{x}) + (1-\alpha)\, f(x)$

Como f(x) > f ($\bar{x}$), tenemos

$$f(\alpha\bar{x} + (1-\alpha)\, x) \geq \alpha\, f(x) + (1-\alpha)\, f(x) > \alpha\, f(\bar{x}) + (1-\alpha)\, f(\bar{x}) = f(\bar{x})$$

Podemos tomar $\alpha \approx 1 \Rightarrow \alpha\, \bar{x} + (1-\alpha)x \approx \bar{x}$ no puede ser un máximo local, en contra de la hipótesis inicial. Luego, llegamos a una contradicción, por lo que el teorema es cierto. Es decir, si es máximo local, entonces es una solución óptima para el PNL.

Definición 6.8. La matriz Hessiana asociada a una función f(x)= (x_1, x_2,…, x_n) es una matriz cuadrada, $H_{n \times n}$, tal que sus elementos h_{ij} son de la forma:

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

Definición 6.9. Denominamos Hessiano al determinante asociado a la matriz Hessiana. En $\Re^2$, el hessiano es H: $\Re^2 \rightarrow \Re$

$$H = \begin{vmatrix} \dfrac{\partial^2 z}{\partial x_1^2} & \dfrac{\partial^2 z}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 z}{\partial x_1 \partial x_2} & \dfrac{\partial^2 z}{\partial x_2^2} \end{vmatrix}$$

Para un problema de Maximización con dos variables, las condiciones que se han de verificar son:

$$\frac{\partial z}{\partial x_1} = 0, \quad \frac{\partial^2 z}{\partial x_1^2} < 0$$

$$\frac{\partial z}{\partial x_2} = 0, \quad H(\bar{x}) > 0$$

donde $\bar{x}$ es el punto de estudio correspondiente al máximo.

Para un problema de Minimización con dos variables, las condiciones que se han de verificar son:

$$\frac{\partial z}{\partial x_1} = 0, \ \frac{\partial^2 z}{\partial x_1^2} > 0$$

$$\frac{\partial z}{\partial x_2} z = 0, \ H(\bar{x}) > 0$$

donde $\bar{x}$ es el punto de estudio correspondiente al mínimo.

Definición 6.10. El *menor principal de orden i* de una matriz H_{nxn} es el determinante de cualquier matriz ixi que se obtiene al suprimir las n-i filas y las n-i columnas correspondientes de la matriz.

Definición 6.11. El *menor principal dominante de orden i* de una matriz H_{nxn} es el determinante de cualquier matriz ixi que se obtiene al suprimir las n-i últimas filas y columnas de la matriz.

Teorema 6.3. Sea la función f(x) con derivadas parciales de segundo orden continuas para cada punto de $x \in S$ (conjunto convexo de soluciones factibles). Entonces f(x) es convexa sobre S si y sólo si, para cada $x \in S$, todos los menores principales, H_i, son no negativos.

Teorema 6.4. Sea la función f(x) con derivadas parciales de segundo orden continuas para cada punto $x \in S$ (conjunto convexo de soluciones factibles). Entonces f(x) es cóncava sobre S si y sólo si, para cada $x \in S$, los menores principales, H_i, no nulos tienen el signo que $(-1)^i$.

6.3. Optimización sin restricciones en R^n

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

$$\begin{aligned} &\text{Max (min) } z = f(x_1, x_2, \ldots, x_n) \\ &\text{st} \\ &\quad x = (x_1, x_2, \ldots, x_n), \ x \in R^n \end{aligned} \qquad (6.2)$$

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

Una condición necesaria para que un punto $\bar{x}$ sea un extremo local para el PNL (6.2) nos la proporciona el teorema siguiente:

Teorema 6.5. Si $\bar{x} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ es un extremo local para (6.2), entonces $\frac{\partial f(\bar{x})}{\partial x_i} = 0, \forall i$.

Definición 6.12. Un punto $\bar{x}$ que satisfaga $\frac{\partial f(\bar{x})}{\partial x_i} = 0$ es punto estacionario de la función f(x).

Teorema 6.6. Si $H_i(\bar{x}) > 0$, (i=1, 2, ..., n), entonces un punto estacionario $\bar{x}$ será un mínimo local para (6.2).

Teorema 6.7. Si $H_i(\bar{x}) \neq 0$ (i=1, 2, ..., n) y tiene el signo que $(-1)^k$, entonces un punto estacionario $\bar{x}$ será un maximo local para (6.2).

Teorema 6.8. Si $H_i(\bar{x}) \neq 0$ (i=1, 2, ..., n) y no se dan ninguno de los casos anteriores (*teoremas 6.6 y 6.7*), f(x) presenta un punto de inflexión en ese punto $\bar{x}$.

Ejercicio 6.2. Obtener el extremo de $z = x_1 + 2x_3 + x_2 x_3 - x_1^2 - x_2^2 - x_3^2$

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

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

Resolviendo el sistema anterior, obtenemos, para el extremo, el punto objeto de estudio $(1/2, 2/3, 4/3) = x_0$

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:

$$| H_{1x1} | = -2 < 0 \qquad \Rightarrow \text{signo } (-1)^1$$

Orden 2:

$$| H_{2x2} | = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4 > 0 \qquad \Rightarrow \text{signo } (-1)^2$$

Orden 3:

$$| H_{3x3} | = \begin{vmatrix} -2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{vmatrix} = -8+2 = -6 < 0 \qquad \Rightarrow \text{signo } (-1)^3$$

Por tanto, tenemos el signo de $(-1)^i$. Así pues, en el punto x_0, tenemos un posible máximo. Como el valor de los menores principales no depende del punto, en ese punto tenemos un máximo.

6.4.- Búsqueda de la sección Áurea

Consideremos una función f(x) para la cual la derivada de f(x), f'(x), es difícil de obtener o no existe. El problema de PNL a resolver es

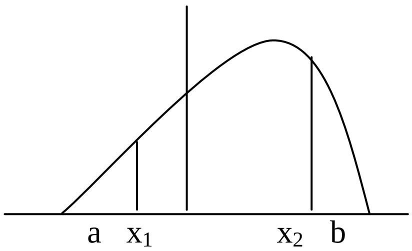
$$\text{Max } z = f(x)$$

$$\text{st } \ a \leq x \leq b \tag{6.3}$$

Definición 6.13. Una función f(x) es unimodal en el intervalo [a,b] si se verifica que existe un punto $\bar{x}$, tal que f(x) es creciente en [a, $\bar{x}$] y decreciente en [$\bar{x}$, b].

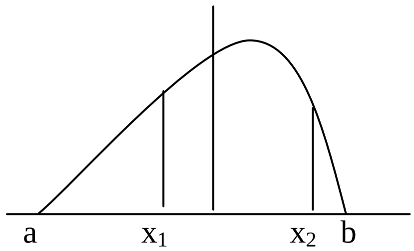
En la búsqueda de la *Sección Áurea*, elegimos dos puntos, x_1, x_2, de manera razonada, de forma que podamos estudiar los siguientes casos:

CASO 1: $f(x_1) < f(x_2)$



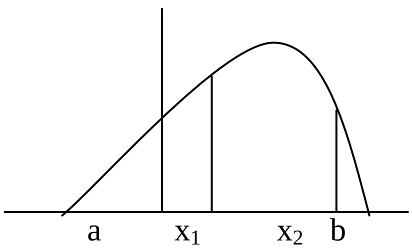
$I = (x_1, b]$

CASO 2: $f(x_1) = f(x_2)$



$I = [a, x_2)$ ó $I = (x_1, b]$

CASO 3: $f(x_1) > f(x_2)$



$I = [a, x_2)$

Sección Áurea se denomina al coeficiente r, el cual se determina de la siguiente forma:

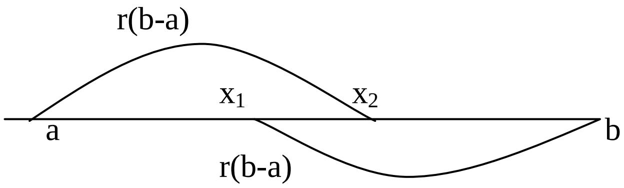
$$\frac{1\ (longitud\ de\ todo\ el\ segmento)}{r\ (longitud\ de\ la\ parte\ más\ larga\ del\ segmento)} =$$

$$= \frac{r\ (longitud\ de\ la\ parte\ más\ larga\ del\ segmento)}{1-r\ (longitud\ de\ la\ parte\ más\ corta\ del\ segmento)}$$

$$r^2 = 1-r \Rightarrow r^2+r-1=0, \quad r = 0{,}618$$

En cada caso, podemos demostrar que la solución óptima para (6.1) estará en un subconjunto [a, b].

La búsqueda de la *Sección Áurea* empieza con la evaluación de f(x) en los puntos x_1 y x_2, donde x_1=b-r(b-a) y x_2=a+r(b-a).



Definiciones 6.14.

L_k : longitud del intervalo de incertidumbre después de realizar k iteraciones.

I_k : intervalo de incertidumbre después de realizar k iteraciones.

$b - x_1 = r(b-a) \Rightarrow x_1 = b - r(b-a)$

$x_2 - a = r(b-a) \Rightarrow x_2 = a + r(b-a)$

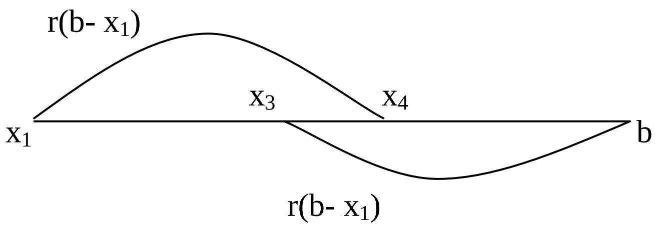
x_1 y x_2 quedan perfectamente definidos.

$L_1 = r(b-a)$

CASO 1: $f(x_1) < f(x_2)$

$L_1 = r(b-a)$

$I_1 = (x_1, b]$, $b - x_1= r(b-a)$, $r^2 = 1-r$



$x_3 = b - r(b- x_1) = b – r\, r(b-a) = b - r^2 (b-a)$

$x_3 = b - (1-r) (b-a)= x_2$

$x_4 = x_1 + r(b- x_1) = x_1 + r\, r(b-a) = x_1 + r^2 (b-a)$

$L_2 = b - x_3 = r(b – x_1) = r^2 (b-a)$, $I_2 = (x_3, b]$

Siguiendo el mismo proceso, llegamos a una expresión general para la longitud del intervalo de incertidumbre después de realizadas *k* iteraciones del algoritmo:

$L_k = r^k\,(b\text{-}a)$, $\qquad r^k = \dfrac{L_k}{b-a}$

Tomando logaritmos

$$k \log r = \log L_k - \log (b\text{-}a)$$

$$k = \frac{\log L_k - \log(b-a)}{\log r}$$

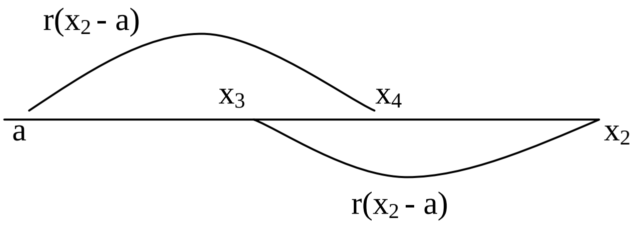
Poniendo la condición de que la longitud de incertidumbre sea menor o igual que un valor prefijado, ε, tenemos que el número de iteraciones, *k*, viene dado por

$$L_k \le \varepsilon \Rightarrow k = \frac{\log \varepsilon - \log(b-a)}{\log r}$$

El valor que tomamos para k es el entero más próximo por exceso, [k]+1.

CASO 2 : $f(x_1) \ge f(x_2)$

$L_1 = r(b\text{-}a)$
$I_1 = [a, x_2)$, $\quad x_2\text{-}a = r(b\text{-}a)$, $\ r^2 = 1\text{-}r$



$x_3 = x_2 - r(x_2 - a) = x_2 - r^2\,(b\text{-}a)$
$x_4 = a + r\,(x_2\text{-}a) = a + r^2\,(b\text{-}a)$
$x_4 = a + (1\text{-}r)(b\text{-}a) = a\text{+}b\text{-}a\text{-}r(b\text{-}a) = b\text{-}r(b\text{-}a) = x_1$
$x_1 = b - r\,(b\text{-}a)$
$L_2 = x_4 - a = r(x_3 - a) = r^2\,(b\text{-}a)$

Siguiendo el mismo proceso, llegamos a una expresión general para la longitud del intervalo de incertidumbre después de realizadas *k* iteraciones del algoritmo:

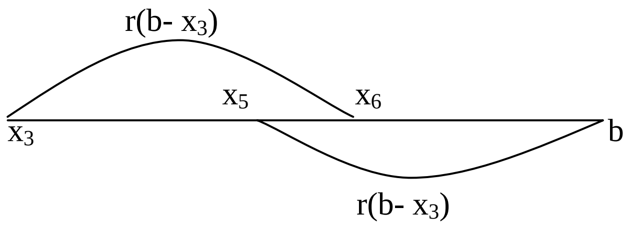
$L_k = r^k\,(b-a)$, $\quad r^k = \dfrac{L_k}{b-a}$

$L_k \le \varepsilon \Rightarrow k = \dfrac{\log\varepsilon - \log(b-a)}{\log r}$;

El valor que tomamos para k es el entero más próximo por exceso, [k].

CASO 3: $f(x_1) < f(x_2)$
Caso 3.1: $f(x_3) < f(x_4)$
Caso 3.2: $f(x_3) \ge f(x_4)$

Caso 3.1:
$f(x_1) < f(x_2)$ $\quad \Rightarrow L_1 = r(b-a);\ I_1 = (x_1, b]$
$f(x_3) < f(x_4)$



$L_2 = r\,(b - x_1)$, $\quad L_2 < L_1$
$I_2 = (x_3, b]$

$x_5 = b - r\,(b- x_3)$
$x_6 = x_3 + r\,(b- x_3)$

y así sucesivamente.

Ejemplo: Max z =f(x)
st
$-1 \le x \le 1$

¿Cuánto tiene que valer k para que el error sea $\varepsilon \le 10^{-3}$?

$k = \dfrac{\log 10^{-3} - \log 2}{\log 0'618} = 15'79 \Rightarrow k=16$

6.5.- Método del gradiente

El método del gradiente se aplicará para aproximar un punto estacionario de una función cuando sea difícil encontrar ese punto estacionario.

En este apartado, se estudiará cómo obtener una solución óptima (si existe) para un problema de PNL sin restricciones:

$$\begin{aligned} &\text{Max (min) } z = f(x) \\ &\quad \text{st } \ x \in \Re^n \end{aligned}$$

Definición 6.15. El gradiente es un vector cuyas componentes son derivadas parciales de la función f(x) que queremos estudiar.

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n}\right)$$

El vector normalizado del gradiente es: $u = \dfrac{\nabla f(x)}{\|\nabla f(x)\|}$

El gradiente nos proporciona la dirección de máximo cambio.

$$x_0 \xrightarrow{\nabla f(x)} x_1 \ ; \quad \|\nabla f(x_1)\| < \varepsilon$$

Los pasos a seguir son:

i) Fijamos un punto inicial x_0 y en la dirección del gradiente, nos desplazamos hasta otro punto x_1.
ii) Fijamos el grado de aproximación a la solución a través de ε ($\|\nabla f(x_1)\| < \varepsilon$).
iii) Si se cumple la condición, x_1 se encontrará cercano a un punto estacionario ($\dfrac{\partial f(x)}{\partial x} = 0$).
iv) Si no se cumple, buscaremos de forma iterativa un punto x_j en el que se verifique $\|\nabla f(x_j)\| < \varepsilon$.

$$\begin{aligned} &\text{Max } z = f(x_0 + t_0 \nabla f(x_0)) \Rightarrow \ x_0 \xrightarrow{\nabla f(x)} x_1: \ x_1 = x_0 + t_0 \nabla f(x_0) \\ &\quad \text{st } \ t_0 \geq 0 \end{aligned}$$

Nos preguntamos ¿ es $\|\nabla f(x_1)\| < \varepsilon$?
Si lo es, hemos terminado.

En caso contrario

$$\text{Max } z=f(x_1 + t_1\nabla f(x_1)) \Rightarrow \quad x_1 \xrightarrow{\nabla f(x)} x_2: \; x_2= x_1 + t_1\nabla f(x_1)$$
$$\text{st} \quad t_1 \geq 0$$
$$¿\text{ Es } \|\nabla f(x_2)\| < \varepsilon \text{ ?}$$

Si lo es, hemos terminado.

En caso contrario, seguimos con el proceso hasta encontrar un punto que lo verifique.

El proceso anterior se repite hasta cumplir la condición de parada,($\|\nabla f(x_i)\| < \varepsilon$).

Ejercicio 6.3. Utilizando el método del gradiente aproxime la solución de

$$\text{Max } z = f(x_1, x_2) = -(x_1-1)^2 - (x_2-2)^2$$
$$\text{st} \; (x_1, x_2) \in R^2$$

Para resolverlo elegimos de forma arbitraria el punto $x_0 = (1, 1)$ y calculamos el gradiente en ese punto

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}\right) = (-2(x_1-1)-2(x_2-2))$$

$$x_1 = x_0 + t_0 \nabla f(x_0) = (1, 1)+ t_0 (0, 2)= (1, 1+2 t_0)$$

$$\text{Max } z = f(x_0 + t_0 \nabla f(x_0))= \text{Max } f(1, 1+2 t_0)$$

$$f(1, 1+2 t_0)= (2 t_0 -1)^2 = f(t_0)$$

Derivando e igualando a cero la derivada, $f'(t_0)= 0$, tenemos

$$f'= 4(2 t_0 -1)= 8 t_0 - 4 = 0, \; t_0 = 1/2$$

Desplazándonos en la dirección del gradiente hacia el nuevo punto, x_1, tenemos

$$x_1 = x_0 + t_0 \nabla f(x_0)= (1, 1)+ ½ (0, 2)= (1, 2)$$

Calculando $\nabla f(1, 2)= (0, 0)$, nos indica que hemos finalizado. Al ser $f(x_1, x_2)$ una función cóncava, hemos encontrado la solución óptima del problema.

6.6.- Multiplicadores de Lagrange

Se pueden utilizar los multiplicadores de Lagrange para resolver problemas de PNL en los cuales las restricciones son de igualdad.

$$\begin{array}{lll}
\text{Max (min) } z = f(x) = f(x_1, x_2,\dots, x_n) & & \\
\text{s.t.} & & \\
g_1(x_1, x_2,\dots, x_n) = b_1 & \Rightarrow & g_1^* = b_1 - g_1(x) = 0 \\
g_2(x_1, x_2,\dots, x_n) = b_2 & \Rightarrow & g_2^* = b_2 - g_2(x) = 0 \\
\dots\dots & & \dots\dots \\
g_m(x_1, x_2,\dots, x_n) = b_m & \Rightarrow & g_m^* = b_m - g_m(x) = 0
\end{array} \qquad (6.3)$$

Para resolver el problema anterior, asociamos a cada restricción un multiplicador λ_i (i=1, 2, ..., m) y formamos la función lagrangiana

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

donde λ_i (i=1, 2, ..., m) son constantes (desconocidas) denominadas *multiplicadores de Lagrange*.

Buscaremos una solución de $L(x, \lambda)$. Para ello, según se ha visto en el apartado 6.2, la condición necesaria para que $L(x, \lambda)$ tenga un máximo o un mínimo en el punto $(\bar{x}, \bar{\lambda}) = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n, \bar{\lambda}_1, \bar{\lambda}_2, ..., \bar{\lambda}_m)$ es

$$\frac{\partial L}{\partial x_i} = 0, \ (i=1, 2, \dots, n)$$
$$\frac{\partial L}{\partial \lambda_i} = 0, \ (i=1, 2, \dots, m) \qquad (6.4)$$

Teorema 6.9. Supongamos que (6.3) es un problema de maximización. Si f(x) es una función cóncava y si cada $g_i(x)$ es una función lineal, entonces cualquier punto $(\bar{x}, \bar{\lambda})$ que satisfaga (6.4) proporcionará una solución óptima $x=(x_1, x_2,\dots, x_n)$ para el problema (6.3).

Teorema 6.10. Supongamos que (6.3) es un problema de minimización. Si f(x) es una función convexa y si cada $g_i(x)$ es una función lineal, entonces cualquier punto $(\bar{x}, \bar{\lambda})$ que satisfaga (6.4) proporcionará una solución óptima $x=(x_1, x_2,\dots, x_n)$ para el problema (6.3).

Ejercicio 6.4. Comprobar que el punto (4,-2;4,8) es estacionario para la función lagrangiana asociada al problema:

$$\begin{aligned}
&\text{Max } z = (x-2)^2+(y-2)^2\\
&\text{st}\\
&\quad -x+y^2 \leq 0\\
&\quad x+y \leq 2
\end{aligned}$$

Ejercicio 6.5. Una compañía planifica gastar 10.000 euros en publicidad. Cuesta 3.000 euros un minuto de publicidad en la televisión y 1.000 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.

Solución:

La formulación del problema no lineal queda de la forma:

$$\begin{aligned}
&\text{Max } z =-2x^2- y^2+xy+8x+3y\\
&\text{st}\\
&\quad 3x+y=10
\end{aligned}$$

Formamos la 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 λ .
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

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

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

$$\bar{y} = 20/7-\lambda = 20/7-1/4 = 73/28$$

$$\bar{x} = 19/7-1/4 = 69/28$$

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

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

Los menores principales de primer orden son negativos y $H_2=7 > 0$, luego f(x, y) es una función cóncava. La restricción es lineal, por el *Teorema 6.9,* concluimos que la solución es óptima para el problema no lineal.

Así, la empresa tendría que comprar 69/28 minutos de tiempo de televisor y 73/28 minutos de tiempo de radio.

6.6.1. Interpretación de los multiplicadores de Lagrange

Estamos interesados en estudiar en qué medida varía el valor de la función objetivo en $\bar{x}$ cuando variamos el recurso b.

Tenemos que en el punto $(\bar{x}, \bar{\lambda}) = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n, \bar{\lambda}_1, \bar{\lambda}_2, ..., \bar{\lambda}_m)$ el valor de la función objetivo, z =f(x), es óptimo. Si tomamos derivadas parciales en ese punto respecto a los recursos b_i, tenemos:

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i\, [b_i - g_i(x)], \qquad \frac{\partial f(\bar{x})}{\partial b_i} = \bar{\lambda}_i \quad (i=1, 2, \ldots, m)$$

lo que nos indica que los *multiplicadores de Lagrange* vienen a representar los precios marginales (precios sombra) ya mencionados al estudiar la interpretación económica del DUAL.

En el ejercicio anterior $\bar{\lambda}$ =1/4, el gasto de un cantidad extra pequeña, Δb (en miles de euros) aumentaría los ingresos de la empresa en aproximadamente 0.25Δb.

6.7.- Condiciones de Kuhn -Tucker

En esta sección se estudian las condiciones necesarias y suficientes para que el punto $\bar{x} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ sea una solución óptima para el PNL, con restricciones de desigualdad.

Las condiciones de *Kuhn-Tucker* sólo son aplicables si las $g_i(x)$ (i=1, 2, ..., m) satisfacen ciertas *condiciones de regularidad*:

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

Para aplicar los resultados de este apartado, todas las restricciones del PNL tienen que ser del tipo $\leq$.

Estudiaremos los dos casos siguientes:

Caso 1: x_i sin restricción de signo

$$\begin{aligned} &\text{Max (min) } z = f(x) \\ &\text{s.t.} \\ &\quad g_1(x) \leq b_1 \\ &\quad g_2(x) \leq b_2 \\ &\quad \ldots. \\ &\quad g_m(x) \leq b_m \end{aligned} \tag{6.5}$$

El problema anterior lo transformamos en un problema de *multiplicadores de Lagrange* introduciendo variables de holgura h_i^2 :

$$g_1(x) + h_1^2 = b_1 \quad \Rightarrow g_1^* = (b_1 - (g_1(x) + h_1^2)) = 0$$
$$g_2(x) + h_2^2 = b_2 \quad \Rightarrow g_2^* = (b_2 - (g_2(x) + h_2^2)) = 0$$
$$\ldots$$

$$g_m(x) + h_m^2 = b_m \quad \Rightarrow g_m^* = (b_m - (g_m(x) + h_m^2)) = 0$$

Formamos la función lagrangiana

$$L(x, h_i, \lambda_i) = f(x) + \sum_{i=1}^{m} \lambda_i [b_i - (g_i(x) + h_i^2)]$$

Y, operando, tenemos

$$\frac{\partial L}{\partial x_j} = 0 = \frac{\partial}{\partial x_j} f(x) - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(x)}{\partial x_j}$$

$$\frac{\partial L}{\partial \lambda_i} = 0 = [b_i - (g_i(x) + h_i^2)] \Rightarrow h_i^2 = b_i - g_i(x)$$

$$\frac{\partial L}{\partial h_i} = 0 = -2 h_i\lambda_i \approx h_i^2 \lambda_i = \lambda_i [b_i - g_i(x)]$$

Las condiciones de *Kuhn-Tucker* se establecen como

$$\boxed{\begin{array}{ll} \dfrac{\partial f(\bar{x})}{\partial x_j} \pm \displaystyle\sum_{i=1}^{m} \bar{\lambda}_i \frac{\partial g_i(\bar{x})}{\partial x_j} = 0 & (j=1, 2, ..., n) \\ \bar{\lambda}_i [b_i - g_i(\bar{x})] = 0 & (i=1, 2, ..., m) \\ \bar{\lambda}_i \geq 0 \quad (i=1, 2, ..., m) & (6.6) \end{array}}$$

Tomaremos en la expresión anterior el signo:

- para un problema de maximización.

+ para un problema de minimización.

Tenemos que tener presente que puede haber óptimos locales y globales que no verifiquen las condiciones *Kuhn-Tucker*.

Si un punto $\bar{x}$ satisface las condiciones de K-T este punto es un óptimo local.

Si un punto $\bar{x}$ es un óptimo local no tiene por qué satisfacer las condiciones de K-T.

Los teoremas siguientes nos dan las condiciones necesarias y suficientes (si satisfacen las *condiciones de regularidad)*, para que un punto $\bar{x} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ sea solución de (6.5).

Teorema 6.11. Sea el problema (6.5) de maximizar f(x). Si el punto $\bar{x} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ es una solución óptima de (6.5), entonces $\bar{x}$ tendrá que satisfacer las *m* restricciones del problema de PNL (6.5) y deberán existir multiplicadores $\bar{\lambda}_1$, $\bar{\lambda}_2$, …, $\bar{\lambda}_m$, que verifiquen las condiciones de *Kuhn-Tucker* (6.6).

Teorema 6.12. Sea el problema (6.5) de minimizar f(x). Si el punto $\bar{x} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ es una solución óptima de (6.5), entonces $\bar{x}$ tendrá que satisfacer las *m* restricciones del problema de PNL (6.5) y deberán existir multiplicadores $\bar{\lambda}_1$, $\bar{\lambda}_2$, …, $\bar{\lambda}_m$, que verifiquen las condiciones de *Kuhn-Tucker* (6.6).

Caso 2: x_i con restricción de signo ($x_i \geq 0$)

Cuando se ponen restricciones de no negatividad para las variables, las condiciones de *Kuhn-Tucker* para el PNL (6.5) se establecen

$$
\begin{array}{ll}
\text{Max (min) } z = f(x) & \\
\quad \text{st.} & \\
\quad\quad g_1(x) \leq b_1 & \\
\quad\quad g_2(x) \leq b_2 & \\
\quad\quad \ldots. & \qquad\qquad (6.7) \\
\quad\quad g_m(x) \leq b_m & \\
\quad\quad -x_i \leq 0 \quad \text{para } (i=1, 2, ..., n) &
\end{array}
$$

Operando de forma análoga a como se hizo para el *Caso 1*, teniendo en cuenta que aquí tenemos que introducir dos variables de holgura h_i^2, (i=1, 2, ..., m) y t_j^2 (j=1, 2, ..., n) y dos multiplicadores de Lagrange λ_i y μ_j, escribiremos la función lagrangiana

$$L(x, h, t, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i [b_i - (g_i(x) + h_i^2)] + \sum_{j=1}^{n} \mu_j [x_j - t_j^2)]$$

y obtenemos las condiciones de *Kuhn-Tucker:*

$$
\boxed{
\begin{array}{l}
\dfrac{\partial f(\bar{x})}{\partial x_j} \pm \displaystyle\sum_{i=1}^{m} \bar{\lambda}_i \dfrac{\partial g_i(\bar{x})}{\partial x_j} \pm \bar{\mu}_j = 0 \quad (j=1, 2, ..., n) \\[2ex]
\bar{\lambda}_i [b_i - g_i(\bar{x})] = 0 \qquad (i=1, 2, ..., m) \\[2ex]
\left[\dfrac{\partial f(\bar{x})}{\partial x_j} \pm \displaystyle\sum_{i=1}^{m} \bar{\lambda}_i \dfrac{\partial g_i(\bar{x})}{\partial x_j}\right] \bar{x}_j = 0 \quad (j=1, 2, ..., n)
\end{array}
}
$$

$\bar{\lambda}_i \geq 0$ (i=1, 2, ..., m)		
$\bar{\mu}_j \geq 0$ (j=1, 2, ..., n)		(6.8)

Tomaremos en la expresión anterior el signo:

(-, +) para un problema de *maximización*.
(+, -) para un problema de *minimización*.

Como $\bar{\mu}_j \geq 0$, la primera restricción de (6.8) puede escribirse como

$$\frac{\partial f(\bar{x})}{\partial x_j} - \sum_{i=1}^{m} \bar{\lambda}_i \frac{\partial g_i(\bar{x})}{\partial x_j} \leq 0 \quad \text{(j=1, 2, ..., n) (problema de maximización)}$$

$$\frac{\partial f(\bar{x})}{\partial x_j} + \sum_{i=1}^{m} \bar{\lambda}_i \frac{\partial g_i(\bar{x})}{\partial x_j} \geq 0 \quad \text{(j=1, 2, ..., n) (problema de minimización)}$$

Generalizando los *Teoremas 6.11* y *6.12*, tenemos:

Teorema 6.13. Sea el problema (6.7) de maximizar f(x). Si el punto $\bar{x} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ es una solución óptima de (6.7), entonces $\bar{x}$ tendrá que satisfacer las *m* restricciones del problema de PNL (6.7) y deberán existir multiplicadores $\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_m$; $\bar{\mu}_1, \bar{\mu}_2, \ldots, \bar{\mu}_n$, que verifiquen las condiciones de *Kuhn-Tucker* (6.8).

Teorema 6.14. Sea el problema (6.7) de minimizar f(x). Si el punto $\bar{x} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ es una solución óptima de (6.7), entonces $\bar{x}$ tendrá que satisfacer las *m* restricciones del problema de PNL (6.7) y deberán existir multiplicadores $\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_m$; $\bar{\mu}_1, \bar{\mu}_2, \ldots, \bar{\mu}_n$ que verifiquen las condiciones de *Kuhn-Tucker* (6.8).

Podemos establecer dos teoremas, correspondientes a los *Teorema 6.13* y *6.14*, que recojan las condiciones *necesarias* y *suficientes* para que $\bar{x} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ sea una solución óptima de (6.5) o (6.7).

Teorema 6.15. Sea el problema (6.7) de maximizar f(x). Si f(x) es una función cóncava y las $g_i(x)$ (i=1, 2, ..., m) son funciones convexas, entonces cualquier punto $\bar{x} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ que satisfaga las condiciones de *Kuhn-Tucker* (6.8) y las *m* restricciones del problema de PNL es una solución óptima de (6.7).

Teorema 6.16. Sea el problema (6.7) de minimizar f(x). Si f(x) es una función convexa y las $g_i(x)$ (i=1, 2, ..., m) son funciones convexas, entonces cualquier punto $\bar{x} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ que satisfaga las condiciones de *Kuhn-Tucker* (6.8) y las *m* restricciones del problema de PNL es una solución óptima de (6.7).

Para poder aplicar los resultados anteriores es importante que el problema se plantee en la forma que aparece descrita; de no ser así, los signos de los multiplicadores podrían ser diferentes.

i) Debe también verificarse la *hipótesis de regularidad* que obliga a que los vectores gradientes de las restricciones saturadas sean linealmente independientes en el óptimo. De no ser así, el punto podría no verificar el teorema correspondiente, como ocurría en el caso de restricciones de igualdad.
ii) Los multiplicadores de *Kuhn-Tucker* (λ, μ) también se conocen como variables duales del problema.
iii) Para problemas de maximización (minimización) se tendría un resultado similar, con la única diferencia de que los multiplicadores deben ser negativos (cambio de signo).
iv) Estos teoremas sirven para determinar los posibles óptimos de un problema; para ello, habría que resolver un sistema de *n+m* ecuaciones con *n+m* incógnitas. Una vez resuelto este sistema habría que seleccionar aquellas soluciones que verifican las desigualdades. Por analogía con los problemas con restricciones de igualdad, a estos puntos se les llama puntos estacionarios.

A continuación, veremos un caso práctico de localización de posibles óptimos mediante las *condiciones de Kuhn-Tucker*.

Ejemplo 6.3. Encontrar las posibles soluciones del problema

$$\begin{aligned} &\text{Min } z = (x-7)^2+(y-10)^2 \\ &\quad \text{st} \\ &\quad y-8 \leq 0 \\ &\quad (x-10)^2+(y-10)^2-36 \leq 0 \end{aligned}$$

Para resolver el problema, comenzamos por observar que las restricciones están dadas en la forma estándar. Según las condiciones de Kuhn-Tucker, todo mínimo del problema debería verificar el sistema:

$$2(x-7)+ \lambda_2*2(x-10)=0$$

$$2(y-10)+ \lambda_1+\lambda_2*2(y-10)=0$$

$\lambda_1(y-8)=0$

$\lambda_2\,[(x-10)^2+(y-10)^2-36]=0$

Las soluciones de este sistema son:

x	y	λ_1	λ_2
16	10	0	-1.5
4	10	0	-0.5
7	10	0	0
7	8	4	0
15,6569	8	-2,12132	-1,53033
4,34315	8	2,12132	-0,46967

Si de esas soluciones seleccionamos las que corresponden a puntos factibles y con multiplicadores positivos (por tratarse de un problema de minimización), se obtiene un único candidato a mínimo del problema: $x=7$, $y=8$, $\lambda_1=4$, $\lambda_2=0$.

Con las condiciones estudiadas hasta ahora, no se está en disposición de asegurar que dicho punto corresponda realmente a un mínimo local del problema. Se hace necesario, por tanto, el estudio de condiciones suficientes de optimalidad. La siguiente sección se dedica precisamente a eso.

Condiciones suficientes de optimalidad.

Una vez seleccionadas las posibles soluciones de un problema de optimización con restricciones de desigualdad, deben estudiarse condiciones suficientes que permitan decidir si realmente los puntos localizados corresponden a verdaderas soluciones.

Teorema 6.17: Si el problema es convexo (restricciones convexas), entonces las condiciones *necesarias* de Kuhn-Tucker son además *suficientes*.

El ejemplo anterior es un caso de problema convexo, por lo tanto se puede concluir que el punto estacionario encontrado es mínimo y además global.

6.8.- Programación cuadrática

La programación cuadrática es un caso particular de la programación no lineal en el que la función objetivo es cuadrática y las restricciones son lineales.

$$\text{Max (min) } z = C^T X + X^T D\, X$$
$$\text{st}$$
$$AX \leq b$$
$$X \geq 0$$

La función $X^T D\, X$ define una forma cuadrática donde D es simétrica. La matriz D, se define:

>0 definida positiva (mínimo)
≥ 0 semidefinida positiva (cuasimínimo)
<0 definida negativa (máximo)
≤ 0 semidefinida negativa (cuasimáximo)
$=0$ indefinida (puntos de inflexión)

Tendremos un mínimo si los menores principales asociados con D son todos positivos y un máximo si los menores principales tienen el signo $(-1)^i$.

Ejercicio 6.6. Una empresa puede invertir 1000 millones de euros en tres títulos de bolsa. Sea X_i la variable aleatoria que corresponde al interés anual por cada millón de pesetas invertido en el título i-ésimo. Se conocen los valores siguientes: $E(X_1)=0.14$, $E(X_2)=0.11$, $E(X_3)=0.10$; $Var(X_1)=0.2$, $Var(X_2)=0.08$, $Var(X_3)=0.18$; $Cov(X_1,X_2)=0.05$, $Cov(X_1,X_3)=0.02$, $Cov(X_2,X_3)=0.03$.

a) Formule un problema de programación cuadrática que se pueda utilizar para encontrar la cartera de varianza mínima que alcance un interés anual esperado de por lo menos un 10%.
b) Resuélvalo mediante LINDO.

Solución:

a) Definamos: x_j=número de millones de euros invertidos en el título j-ésimo.

$$E\left[\frac{x_1 X_1 + x_2 X_2 + x_2 X_2}{1000}\right] \geq 0.10$$

$$\text{Min } \operatorname{Var}\left[x_1 X_1 + x_2 X_2 + x_2 X_2\right] = \text{Min} z$$

$$\text{Min } z = 0.20\, x_1^2 + 0.08\, x_2^2 + 0.18\, x_3^2 + 0.10 x_1 x_2 + 0.04 x_1 x_3 + 0.06 x_2 x_3$$

$$\text{st} \quad 0.14x_1 + 0.11x_2 + 0.10\, x_3 \geq 100$$
$$x_1 + x_2 + x_3 = 1000$$
$$x_i \geq 0,\ i = 1, 2, 3.$$

En el programa LINDO se introduce de la forma siguiente:

```
Min X1+X2+X3+L1+L2+L3
 st
   0.40X1+0.10X2+0.04X3-0.14L1+L2-L3=>0
```

```
    0.10X1+0.16X2+0.06X3-0.11L1+L2-L3=>0
    0.04X1+0.06X2+0.36X3-0.10L1+L2-L3=>0
           -0.14X1-0.11X2-0.10X3<=-100
                        X1+X2+X3<=1000
                      -X1-X2-X3<=-1000
END
QCP 5
```

b) Salida:

```
      OBJECTIVE FUNCTION VALUE

      1)      63988.77

VARIABLE        VALUE          REDUCED COST
      X1       140.449432          0.000000
      X2       623.595520          0.000000
      X3       235.955063          0.000000
      L1         0.000000         11.853932
      L2         0.000000          0.000000
      L3       127.977531          0.000000

     ROW   SLACK OR SURPLUS     DUAL PRICES
      2)         0.000000       -140.449432
      3)         0.000000       -623.595520
      4)         0.000000       -235.955063
      5)        11.853932          0.000000
      6)         0.000000          0.000000
      7)         0.000000        127.977531
```

En la tabla siguiente se recogen las inversiones:

Título	1	2	3
Euros invertidos	140.4494	623.5955	235.9550

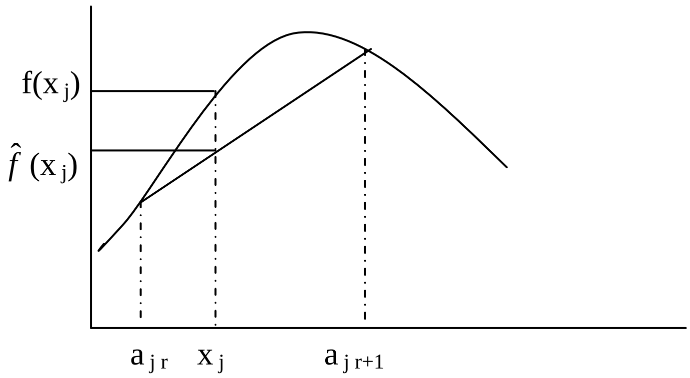
6.9.- Programación separable

A los problemas de programación no lineal se les denomina de *programación separable* cuando las variables de decisión aparecen en términos separados tanto en la función objetivo como en las restricciones.

$$\text{Max (min)}\ \ z = \sum_{j=1}^{n} f_j(x_j) = f_1(x_1) + f_2(x_2) + \ldots + f_n(x_n)$$

$$\text{st}\quad \sum_{j=1}^{n} g_{ij}(x_j) \le b_i \quad (i=1, 2, ..., m)$$

En la figura siguiente representamos una función y su aproximación en un intervalo de un punto x_j, $a_{jr} \leq x_j \leq a_{j\,r+1}$



donde $\hat{f}(x_j)$ es la función lineal que emplearemos para estimar $f(x_j)$.

En ese intervalo x_j, se puede expresar como una combinación lineal de los extremos, es decir,

$$x_j = \alpha\, a_{jr} + (1-\alpha)\, a_{j\,r+1}, \qquad f(x_j) = \alpha\, f(a_{jr}) + (1-\alpha)\, f(a_{j\,r+1})$$

Vamos a considerar que todas las variables tienen el mismo recorrido dividido en *k* celdillas.

$$a \leq x_j \leq b; \quad a = a_{j1} \leq a_{j2} \leq \ldots \leq a_{j\,k-1} \leq a_{jk} = b$$

$$x_j = \sum_{r=1}^{k} \alpha_{jr}\, a_{jr}\,; \quad \sum_{r=1}^{k} \alpha_{jr} = 1 \quad (j=1, 2, ..., n)$$

Aplicando la función de estimación a la función objetivo y a las restricciones, tenemos

$$\hat{f}_j(x_j) = \sum_{r=1}^{k} \alpha_{jr}\, f_j(a_{jr})$$

$$\hat{g}_{ij}(x_j) = \sum_{r=1}^{k} \alpha_{jr}\, g_{ij}(a_{jr})$$

Para que la aproximación sea buena, debemos tener en cuenta la *suposición de adyacencia*, que dice que, a lo sumo, dos α_{jr} pueden ser positivos. Si para un j dado (j=1, 2, ..., n), dos α_{jr} son positivos, entonces tienen que ser adyacentes. Es decir, deberán ser positivos α_{jr-1} y α_{jr+1}.

Ahora ya podemos formular el problema:

$$\text{Max (min)}z = \sum_{j=1}^{n} \hat{f}_j(x_j) = \sum_{j=1}^{n}\sum_{r=1}^{k} \alpha_{jr} f_j(a_{jr})$$

$$\text{st} \sum_{j=1}^{n} \hat{g}_{ij}(x_j) = \sum_{j=1}^{n}\sum_{r=1}^{k} \alpha_{jr} g_{ij}(a_{jr}) \le b_i$$

$$\sum_{r=1}^{k} \alpha_{jr} = 1 \quad (j=1, 2, ..., n),$$

$$\alpha_{jr} \ge 0 \; (j=1, 2, ..., n; \; r=1, 2, ..., k)$$

Ejercicio 6.7. Resuelva el problema siguiente:

$$\text{Max } z=30x_1 + 35x_2 - 2x_1^2 - 3x_2^2$$
$$\text{s.t.}$$
$$x_1^2 + 2x_2^2 \le 250$$
$$x_1 + x_2 \le 20$$
$$x_i \ge 0 \; (i=1, 2)$$

Como una de las restricciones es no lineal, sino cuadrática, no podemos aplicar Kuhn-Tucker, por lo que recurriremos a la programación separable.

$$z = \sum_{j=1}^{n} f_j(x_j) = f_1(x_1) + f_2(x_2) + \ldots + f_n(x_n)$$

$$\sum_{j=1}^{n} g_{ij}(x_j) \le b_i \Rightarrow$$

$$g_{11}(x_1) + g_{12}(x_2) + \ldots + g_{1n}(x_n) \le b_1$$
$$\ldots. \qquad \ldots. \qquad \ldots$$
$$g_{m1}(x_1) + g_{m2}(x_2) + \ldots + g_{mn}(x_n) \le b_m$$

$$f_1(x_1) = 30x_1 - 2x_1^2$$
$$f_2(x_2) = 35x_2 - 3x_2^2$$
$$g_{11}(x_1) = x_1^2$$
$$g_{12}(x_2) = 2x_2^2$$
$$g_{21}(x_1) = x_1$$

$g_{22}(x_2) = x_2$

$a_1 \le x_1 \le b_1 \Rightarrow 0 \le x_1 \le 20$
$a_2 \le x_2 \le b_2 \Rightarrow 0 \le x_2 \le 20$

$a_1 = a_{11} = 0;\ a_{12} = 5\ ;\ a_{13} = 10\ ;\ a_{14} = 15\ ;\ a_{15} = 20 = b_1;$
$a_2 = a_{21} = 0;\ a_{22} = 5\ ;\ a_{23} = 10\ ;\ a_{24} = 15\ ;\ a_{25} = 20 = b_2;$

Cambiamos las funciones f_i y g_{ij} por sus estimaciones:

$\hat{f}_1(x_1) = \alpha_{11} f_1(a_{11}) + \alpha_{12} f_1(a_{12}) + \ldots + \alpha_{15} f_1(a_{15})$
$\hat{f}_2(x_2) = \alpha_{21} f_2(a_{21}) + \alpha_{22} f_2(a_{22}) + \ldots + \alpha_{25} f_5(a_{25})$
$x_1 = \alpha_{11} a_{11} + \alpha_{12} a_{12} + \ldots + \alpha_{15} a_{15}$

	a_{21}, a_{11}	a_{22}, a_{12}	a_{23}, a_{13}	a_{24}, a_{14}	a_{25}, a_{15}
	0	5	10	15	20
f_1	0	100	100	0	-200
f_2	0	100	50	-150	-500
g_{11}	0	25	100	225	400
g_{12}	0	50	200	450	800
g_{21}	0	5	10	15	20
g_{22}	0	5	10	15	20

Max $\hat{z} = \hat{f}_1 + \hat{f}_2 = 100\,\alpha_{12} + 100\,\alpha_{13} - 200\,\alpha_{15} + 100\,\alpha_{22} + 50\,\alpha_{23} - 150\,\alpha_{24} - 500\,\alpha_{25}$
s.t.
$25\,\alpha_{12} + 100\,\alpha_{13} + 225\,\alpha_{14} + 400\,\alpha_{15} + 50\,\alpha_{22} + 200\,\alpha_{23} + 450\,\alpha_{24} + 800\,\alpha_{25} \le 250$
$5\,\alpha_{12} + 10\,\alpha_{13} + 15\,\alpha_{14} + 20\,\alpha_{15} + 5\,\alpha_{22} + 10\,\alpha_{23} + 15\,\alpha_{24} + 20\,\alpha_{25} \le 20$
$\alpha_{11} + \alpha_{12} + \alpha_{13} + \alpha_{14} + \alpha_{15} = 1$
$\alpha_{21} + \alpha_{22} + \alpha_{23} + \alpha_{24} + \alpha_{25} = 1$
$\alpha_{jr} \ge 0$ (j=1, 2; r =1, 2, 3, 4, 5)

Para obtener la solución final tendremos en cuenta la condición de adyacencia. Dicha solución es:

$\alpha_{21} = \alpha_{22} = 1;\ x_1 = 5,\ x_2 = 5,\ z = 200$

6.10.- Programación Estocástica

La programación aleatoria o estocástica se presenta cuando alguno o todos los parámetros del problema se pueden describir mediante variables aleatorias. Nuestro objetivo es expresar la naturaleza probabilística del problema en términos determinísticos.

Dado el problema de programación no lineal en forma canónica

$$\text{Max } z = \sum_{i=1}^{n} c_i\, x_i$$

$$\text{st} \sum_{j=1}^{n} a_{i\,j}\, x_j \le b_i$$

$$x_j \ge 0 \quad \forall j = 1, 2, \ldots, n$$

supondremos que los coeficientes del problema son variables aleatorias que siguen distribuciones normales. Los casos que se nos pueden presentar son:

Caso 1: El coeficiente de la función objetivo es aleatorio.

Si el coeficiente c_i es una variable aleatoria con valor esperado $E[c_i]$ y varianza Var (c_i), entonces el problema, en forma determinística, se establece:

$$\text{Max } z = \sum_{i=1}^{n} E[c_i]\, x_i$$

$$\text{st} \sum_{j=1}^{n} a_{i\,j}\, x_j \le b_i$$

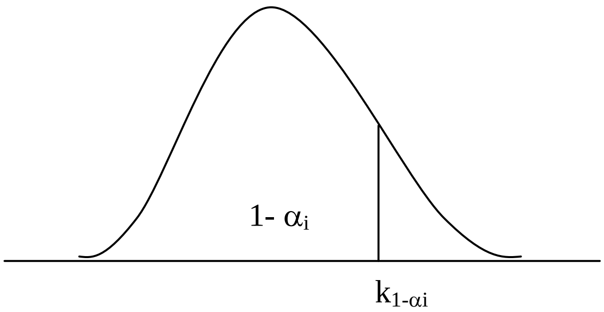
$$x_j \ge 0 \quad \forall j = 1, 2, \ldots, n$$

Caso 2: Los a_{ij} siguen una distribución normal, $a_{ij} \approx N(E[a_{i\,j}], \text{Var}(a_{i\,j}))$. La covarianza de a_{ij} y $a_{i'j'}$ viene dada por Cov $(a_{ij}, a_{i'j'})$.

Consideremos que la i-ésima restricción se produce con una probabilidad mínima de $1 - \alpha_i$.

$$P\left(\sum_{j=1}^{n} a_{ij}\, x_j \le b_i\right) \ge 1 - \alpha_i$$

Tomemos $h_i = \sum_{j=1}^{n} a_{ij}\, x_j$, entonces la probabilidad asociada a la i-ésima restricción la escribimos como $P(h_i \le b_i) \ge 1 - \alpha_i$



Tipificando $P(h_i \le b_i)$, tenemos:

$$P\left[\frac{h_i - E[h_i]}{\sqrt{Var(h_i)}} \le \frac{b_i - E[h_i]}{\sqrt{Var(h_i)}}\right] \ge 1 - \alpha_i$$

La varianza asociada a h_i es $Var(h_i) = X^T D_i X$, y su desviación típica

$$\sqrt{Var(h_i)} = \sqrt{X^T D_i X}$$

siendo D_i la matriz de la i-ésima covarianza.

$$D_i = \begin{pmatrix} Var(a_{i1}) & Cov(a_{i1}, a_{i2}) & .. & Cov(a_{i1}, a_{in}) \\ Cov(a_{i2}, a_{i1}) & Var(a_{i2}) & .. & Cov(a_{i2}, a_{in}) \\ .. & ... & .. & .. \\ Cov(a_{in}, a_{i1}) & Cov(a_{in}, a_{i2}) & .. & Var(a_{in}) \end{pmatrix}$$

donde $\left[\frac{h_i - E[h_i]}{\sqrt{Var(h_i)}}\right] \approx N(0, 1)$

por tanto, $\phi\left[\frac{h_i - E[h_i]}{\sqrt{Var(h_i)}}\right]$ = 1 - α_i

El cuantil correspondiente verifica $\frac{b_i - E[h_i]}{\sqrt{Var(h_i)}} \geq k_{1-\alpha i}$

Operando llegamos al resultado

$$E[h_i] + k_{1-\alpha i}\sqrt{Var(h_i)} \leq b_i$$

Sustituyendo en la expresión anterior $h_i = \sum_{j=1}^{n} a_{ij}\, x_j$, tenemos para la i-ésima restricción

$$\sum_{j=1}^{n} E[a_{ij}]\, x_j + k_{1-\alpha i}\sqrt{X^T D_i X} \leq b_i \quad \forall\, i = 1, 2, \ldots, m$$

Si las v.a. a_{ij} son v. a. i. (caso particular del anterior), tenemos

$$D_i = \begin{pmatrix} Var(a_{i1}) & & & 0 \\ & Var(a_{i2}) & .. & \\ & \ldots & .. & \\ 0 & & & Var(a_{in}) \end{pmatrix}$$

Así, podemos escribir:

$$\sum_{j=1}^{n} E[a_{ij}]\, x_j + k_{1-\alpha i}\sqrt{\sum_{j=1}^{n} Var(a_{ij}) x_j^2} \leq b_i$$

Caso 3: Los coeficientes de las restricciones, b_i, son v.a., es decir,

$b_i \approx N(E[b_i], Var(b_i))$

Operando de forma similar al *Caso 2*, fijamos la probabilidad de que ocurra la i-ésima restricción

$$P(b_i \geq \sum_{j=1}^{n} a_{ij} x_j) \geq \alpha_i$$

$$P(b_i \geq \sum_{j=1}^{n} a_{ij} x_j) = 1 - P(b_i \leq \sum_{j=1}^{n} a_{ij} x_j)$$

Tipificando la última expresión y operando

$$1-P\left[\frac{b_i - E[b_i]}{\sqrt{Var(b_i)}} \leq \frac{\sum_{j=1}^{n} a_{ij} x_j - E[b_i]}{\sqrt{Var(b_i)}}\right] \geq \alpha_i$$

$$P\left[\frac{b_i - E[b_i]}{\sqrt{Var(b_i)}} \leq \frac{\sum_{j=1}^{n} a_{ij} x_j - E[b_i]}{\sqrt{Var(b_i)}}\right] \leq 1 - \alpha_i$$

donde $\left[\frac{b_i - E[b_i]}{\sqrt{Var(b_i)}}\right] \approx N(0, 1)$

por tanto, $\phi\left[\frac{\sum_{j=1}^{n} a_{ij} x_j - E[b_i]}{\sqrt{Var(b_i)}}\right] = 1 - \alpha_i$

El cuantil correspondiente verifica $\frac{\sum_{j=1}^{n} a_{ij} x_j - E[b_i]}{\sqrt{Var(b_i)}} \leq k_{1-\alpha i}$

por tanto

$$k_{1-\alpha i}\ \sqrt{Var(b_i)} \geq \sum_{j=1}^{n} a_{ij} x_j - E[b_i]$$

Operando llegamos al resultado para la i-ésima restricción

$$\sum_{j=1}^{n} a_{ij}\, x_j \leq E[b_i] + k_{1-\alpha i}\ \sqrt{Var(b_i)}$$

Caso 4: Los coeficientes a_{ij} y b_i son v. a.

$$\sum_{j=1}^{n} a_{ij}\, x_j \leq b_i, \quad \sum_{i=1}^{n} a_{ij}\, x_j - b_i \leq 0$$

Como todas las a_{ij} y b_i siguen distribuciones normales, es decir,

$$a_{ij} \approx N(E[a_{i\,j}], Var(a_{i\,j})),\ b_i \approx N(E[b_i], Var(b_i))$$

$\sum_{j=1}^{n} a_{ij}\, x_j - b_i$ también sigue una distribución normal. Por tanto, podemos hacer un estudio análogo al *Caso 2*.

Ejercicio 6.7. Una empresa minera dispone de tres canteras, X_1, X_2 y X_3, las cantidades a explotar siguen distribuciones normales: N(150m^3, σ^2=625m^6), N(200m^3, σ^2=2500m^6) y N(100m^3, σ^2=625m^6), respectivamente. Por otra parte, el material debe transportarse a los lugares X_4, X_5, X_6, X_7 y X_8, en donde se requieren respectivamente 100 m^3, 100 m^3, 150 m^3, 250 m^3 y 50 m^3. Debido a las restricciones en la flotilla de camiones, en el equipo de explotación de las canteras y en los caminos de acceso, las capacidades de transporte de cada una de las canteras a los lugares de descarga resultan como se muestra en la siguiente tabla:

	Destino				
Origen	X_4	X_5	X_6	X_7	X_8
X_1	60	40	80	80	0
X_2	60	60	60	80	60
X_3	0	40	30	70	30

Tomar como dato, para las canteras, el cuantil correspondiente a $F_i(2)=0.977250$ (i=1, 2, 3). ¿Cómo se pueden satisfacer las demandas al máximo?¿en qué destinos no se satisface la demanda al ritmo deseado y en qué cantidades?

Resolviendo el problema mediante LINDO, tenemos

i) Planteamiento del problema:

Sea x_j (1=1, 2, ..., 15) la cantidad de producto que enviaremos del origen al destino.

Este ejercicio corresponde al *Caso 3* de programación estocástica

$$\sum_{i=1}^{n} a_{ij}\, x_j \leq E[b_i] + k_{1-\alpha i} \sqrt{Var(b_i)}$$

Las cantidades a explotar son:

$$b_1 = 150+2\sqrt{625} = 200,\ b_2 = 200+2\sqrt{2500} = 300,\ b_3 = 100+2\sqrt{625} = 150$$

```
Max x1+x2+x3+x4+x5+x6+x7+x8+x9+x10+x11+x12+x13+x14+x15
 st
   x1<=60
   x2<=40
   x3<=80
   x4<=80
   x5=0
   x6<=60
   x7<=60
   x8<=60
   x9<=80
   x10<=60
   x11=0
   x12<=40
   x13<=30
   x14<=70
   x15<=30
   x1+x2+x3+x4+x5<=200
   x6+x7+x8+x9+x10<=300
   x11+x12+x13+x14+x15<=150
   x1+x6+x11>=100
   x2+x7+x12>=100
   x3+x8+x13>=150
   x4+x9+x14>=250
   x5+x10+x15>=50
```

ii) Resolución mediante LINDO

```
      OBJECTIVE FUNCTION VALUE

      1)      650.0000

VARIABLE        VALUE          REDUCED COST
      X1        40.000000          0.000000
      X2        20.000000          0.000000
      X3        60.000000          0.000000
      X4        80.000000          0.000000
      X5         0.000000          0.000000
      X6        60.000000          0.000000
      X7        60.000000          0.000000
      X8        60.000000          0.000000
      X9        80.000000          0.000000
     X10        40.000000          0.000000
     X11         0.000000          0.000000
     X12        20.000000          0.000000
     X13        30.000000          0.000000
     X14        70.000000          0.000000
     X15        30.000000          0.000000
```

Resumiendo los resultados en una tabla, tenemos

z=650 m^3	Destino				
Origen	X_4	X_5	X_6	X_7	X_8
X_1	40	20	60	80	0
X_2	60	60	60	80	40
X_3	0	20	30	70	30
Demandado	100	100	150	250	50
Dado	100	100	150	230	70
Demanda no satisfecha	0	0	0	-20	20

El lugar X_7 no satisface sus necesidades, necesita 20 m^3 más y el X_8 tiene un sobrante de 20 m^3.