

OPTIMIZACIÓN CON RESTRICCIONES

CESÁREO RAIMÚNDEZ

1. REGIÓN VIABLE

Región viable conexa es el conjunto de puntos que obedecen a todas las restricciones simultáneamente. La región viable en los problemas regulares se describe generalmente en la forma:

$$\Omega(x) = \begin{cases} x \in \mathbb{R}^n \\ c_j(x) \leq 0 & j \in D \\ c_j(x) = 0 & j \in I \end{cases}$$

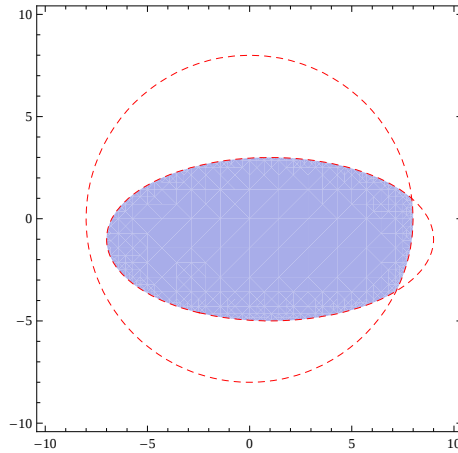


FIGURA 1. Conjunto viable definido por

$$\Omega(x) = \{x_1^2 + x_2^2 - 8^2 \leq 0, (x_1 - 1)^2 + 4(x_2 + 1)^2 - 8^2 \leq 0\}$$

2. OPTIMIZACIÓN CON RESTRICCIONES DE DESIGUALDAD

Los problemas con restricciones de desigualdad pueden transformarse fácilmente en problemas equivalentes con apenas restricciones de igualdad, pudiendo en este caso enfocar el problema de optimización, con el criterio de los multiplicadores de Lagrange. Sea el problema

$$\min_{x \in \Omega} f(x), \quad \Omega = \begin{cases} x \in \mathbb{R}^n \\ g_i(x) \leq 0 \\ i = 1, \dots, m \end{cases}$$

Aquí las restricciones de desigualdad del tipo $g_i(x) \leq 0$ se transformarán en restricciones de igualdad incorporando una variable auxiliar ξ_i de modo que $g_i(x) + \xi_i^2 = 0$. En estas condiciones, formemos la Lagrangiana

$$L(x, \xi, \lambda) = f(x) + \sum_{i=1}^m \lambda_i (g_i(x) + \xi_i^2)$$

cuyas condiciones de punto estacionario son

$$\begin{aligned} \frac{\partial L}{\partial x_j}(x, \xi, \lambda) &= \frac{\partial f}{\partial x_j}(x) + \sum_j \lambda_j \frac{\partial g_j}{\partial x_j}(x) = 0 \\ \frac{\partial L}{\partial \xi_j}(x, \xi, \lambda) &= 2\lambda_j \xi_j = 0 \\ \frac{\partial L}{\partial \lambda_j}(x, \xi, \lambda) &= g_j(x) + \xi_j^2 = 0 \end{aligned}$$

efectuando la expansión en series de Taylor en un entorno del punto (x^k, λ^k, y^k) tendremos:

$$\begin{aligned} L(x, y, \lambda) &= L(x^k, \lambda^k) + \frac{\partial L}{\partial x}(x - x^k) + \frac{\partial L}{\partial \xi}(\xi - \xi^k) + \frac{\partial L}{\partial \lambda}(\lambda - \lambda^k) \\ &+ \frac{1}{2}(x - x^k)^T \frac{\partial^2 L}{\partial x^2}(x - x^k) + \frac{1}{2}(\xi - \xi^k)^T \frac{\partial^2 L}{\partial \xi^2}(\xi - \xi^k) + \frac{1}{2}(\lambda - \lambda^k)^T \frac{\partial^2 L}{\partial \lambda^2}(\lambda - \lambda^k) \\ &+ (x - x^k)^T \frac{\partial^2 L}{\partial x \partial \lambda}(\lambda - \lambda^k) + (\xi - \xi^k)^T \frac{\partial^2 L}{\partial \xi \partial \lambda}(\lambda - \lambda^k) + (\xi - \xi^k)^T \frac{\partial^2 L}{\partial \xi \partial x}(x - x^k) \\ &+ O^3(x - x^k, \xi - \xi^k, \lambda - \lambda^k) \end{aligned}$$

o en forma más compacta

$$\begin{aligned} L(x, \xi, \lambda) &= L(x^k, \xi^k, \lambda^k) + \begin{pmatrix} \nabla f + \lambda^T \nabla g & 2\lambda^T \xi & \mathbf{g}(x, \xi) \end{pmatrix}_k \begin{pmatrix} x - x^k \\ \xi - \xi^k \\ \lambda - \lambda^k \end{pmatrix} \\ &+ \frac{1}{2} \begin{pmatrix} x - x^k \\ \xi - \xi^k \\ \lambda - \lambda^k \end{pmatrix}^T \begin{pmatrix} \nabla^2 f + \lambda^T \nabla^2 g & 0 & \nabla g \\ 0 & 2\lambda & 2\xi \\ \nabla g^T & 2\xi & 0 \end{pmatrix}_k \begin{pmatrix} x - x^k \\ \xi - \xi^k \\ \lambda - \lambda^k \end{pmatrix} + O^3 \end{aligned}$$

donde $\mathbf{g}(x, \xi) = \{g_i(x) + \xi_i\}$. Calculando los puntos estacionarios de esta aproximación cuadrática de la Lagrangiana, llegamos a:

$$\begin{pmatrix} \nabla f + \lambda^T \nabla g & 2\lambda^T \xi & \mathbf{g}(x, \xi) \end{pmatrix}_k + \begin{pmatrix} \nabla^2 f + \lambda^T \nabla^2 g & 0 & \nabla g \\ 0 & 2\lambda & 2\xi \\ \nabla g^T & 2\xi & 0 \end{pmatrix}_k \begin{pmatrix} x - x^k \\ \xi - \xi^k \\ \lambda - \lambda^k \end{pmatrix} = 0$$

finalmente obteniéndose el proceso iterativo

$$\begin{pmatrix} x \\ \xi \\ \lambda \end{pmatrix}_{k+1} = \begin{pmatrix} x \\ \xi \\ \lambda \end{pmatrix}_k - \begin{pmatrix} \nabla^2 f + \lambda^T \nabla^2 g & 0 & \nabla g \\ 0 & 2\lambda & 2\xi \\ \nabla g^T & 2\xi & 0 \end{pmatrix}_k^{-1} \begin{pmatrix} \nabla f + \lambda^T \nabla g & 2\lambda^T \xi & \mathbf{g}(x, \xi) \end{pmatrix}_k^T$$

2.1. Ejemplo. Sea el problema

$$\min_{x \in \Omega \subset \mathbb{R}} f(x) = x^2 - 4x, \quad \Omega = \{g_1(x) = (x - 4)^2 \leq 1, g_2(x) = x \leq 6\}$$

Inicialmente transformamos la restricción de desigualdad en una restricción de igualdad aumentando la dimensión del espacio de búsqueda. Introduciendo la variable auxiliar y tendremos el problema equivalente:

$$\min_{(x,\xi) \in \Omega \subset \mathbb{R}^2} f(x) = x^2 - 4x, \quad \Omega = \{g_1(x, \xi) = (x - 4)^2 + \xi_1^2 - 1 = 0, \quad g_2(x, \xi) = x + \xi_2^2 - 6 = 0\}$$

La Lagrangiana del problema es:

$$L(x, \xi, \lambda) = x^2 - 4x + \lambda_1((x - 4)^2 + \xi_1^2 - 1) + \lambda_2(x + \xi_2^2 - 6)$$

Calculando los puntos estacionarios de L obtenemos:

$$\begin{aligned} 2x + 2\lambda_1(x - 4) - 4 + \lambda_2 &= 0 \\ 2\lambda_1\xi_1 + 2\lambda_2\xi_2 &= 0 \\ (x - 4)^2 + \xi_1^2 - 1 &= 0 \\ x + \xi_2^2 - 6 &= 0 \end{aligned}$$

de donde sacamos las soluciones

$$\begin{aligned} (x, \xi_1, \xi_2, \lambda_1, \lambda_2)_1 &= (5, 0, 1, -3, 0) \\ (x, \xi_1, \xi_2, \lambda_1, \lambda_2)_2 &= (3, 0, 1.73205, 1, 0) \end{aligned}$$

el primero representando un máximo y el segundo representando un mínimo. Debe observarse que los valores de $\{\lambda_i, \xi_i\}$, $i = \{1, 2\}$ caracterizan cuales las restricciones que son activas: Así $\{\xi_1 = 0, \lambda_1 = 1\}_1$ indican que la restricción g_1 está actuando y $\{\xi_2 = 1, \lambda_2 = 0\}_1$ indican que la restricción g_2 no actúa en la definición del mínimo/máximo. El procedimiento iterativo nos lleva a considerar

$$\begin{aligned} \nabla L &= \begin{pmatrix} 2\lambda_1(x - 4) + 2x - 4 + \lambda_2 & 2\lambda_1\xi_1 & 2\lambda_2\xi_2 & g_1(x, \xi_1) & g_2(x, \xi_2) \end{pmatrix} \\ \nabla^2 L &= \begin{pmatrix} 2(\lambda_1 + 1) & 0 & 0 & 2x - 8 & 1 \\ 0 & 2\lambda_1 & 0 & 2\xi_1 & 0 \\ 0 & 0 & 2\lambda_2 & 0 & 2\xi_2 \\ 2x - 8 & 2\xi_1 & 0 & 0 & 0 \\ 1 & 0 & 2\xi_2 & 0 & 0 \end{pmatrix} \end{aligned}$$

y en consecuencia

$$\begin{pmatrix} x \\ \xi_1 \\ \xi_2 \\ \lambda_1 \\ \lambda_2 \end{pmatrix}_{k+1} = \begin{pmatrix} x \\ \xi_1 \\ \xi_2 \\ \lambda_1 \\ \lambda_2 \end{pmatrix}_k - (\nabla^2 L)_k^{-1} (\nabla L)_k^T$$

3. CONDICIONES DE KUHN-TUCKER

Las condiciones de Kuhn-Tucker son condiciones necesarias para caracterizar un punto estacionario en un problema con restricciones, como punto de mínimo local. Se enuncian como sigue: Dado un punto estacionario x^* para que sea un mínimo local es necesario que se verifique

$$\begin{aligned} \nabla f(x^*) + \sum_i \lambda_i \nabla g_i(x^*) &= 0 \\ \lambda_i g_i(x^*) &= 0 \\ g_i(x^*) &\leq 0 \\ \lambda_i &\geq 0 \end{aligned}$$

4. CONDICIONES DE KARUSH-KUHN-TUCKER

Las condiciones de Karush-Kuhn-Tucker (KKT) son necesarias para que un punto sea óptimo local. Es una generalización del método de los multiplicadores de Lagrange, incorporando las restricciones de desigualdad. Sea el problema de optimización.

$$\min_{x \in \mathbb{R}^n} f(x) \text{ con condiciones } g_i(x) \leq 0 \wedge h_j(x) = 0, \{i = 1, \dots, m; j = 1, \dots, l\}$$

4.1. Condiciones necesarias. Supongamos que $\{f, g_i, h_j\} : \mathbb{R}^n \rightarrow \mathbb{R}$ tengan derivada continua en el punto $x^* \in S \subset \mathbb{R}^n$. Si x^* es un mínimo local entonces existen constantes $\lambda \geq 0, \mu_i \geq 0$ y ν_j para $i = 1, \dots, m$ y $j = 1, \dots, l$ tales que:

$$\begin{aligned} \lambda + \sum_{i=1}^m \mu_i + \sum_{j=1}^l |\nu_j| &> 0 \\ \lambda \nabla f(x^*) + \sum_{i=1}^m \mu_i \nabla g_i(x^*) + \sum_{j=1}^l \nu_j \nabla h_j(x^*) &= 0 \\ \forall i, \mu_i g_i(x^*) &= 0 \end{aligned}$$

4.2. Condiciones suficientes. En algunos casos estas condiciones son también suficientes. Supongamos que $\{f, g_i\} : \mathbb{R}^n \rightarrow \mathbb{R}$ tengan derivada continua en el punto $x^* \in S \subset \mathbb{R}^n$ y sean convexas. Sean $\{h_j\} : \mathbb{R}^n \rightarrow \mathbb{R}$ funciones afines. Si existen constantes $\mu_i \geq 0$ y ν_j tales que:

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^m \mu_i \nabla g_i(x^*) + \sum_{j=1}^l \nu_j \nabla h_j(x^*) &= 0 \\ \forall i, \mu_i g_i(x^*) &= 0 \end{aligned}$$

entonces x^* es un punto de mínimo global.

4.3. Ejemplo. Sea el problema (no convexo)

$$\begin{aligned} \min f(x) &= 10 - (4x_1^2 + 4x_1x_2 + 2x_2^2 + 3x_1) \\ g_1(x) = x_1^2 + x_2^2 &\leq 8^2 \\ g_2(x) = (x_1 - 1)^2 + 4(x_2 + 1)^2 &\leq 8^2 \end{aligned}$$

por la figura se nota que posibles puntos candidatos son:

$$\left(\begin{array}{ccccc} \mu_1 \rightarrow 2.78139 & \mu_2 \rightarrow 1.89614 & x_1 \rightarrow 7.93836 & x_2 \rightarrow 0.991178 & f \rightarrow -299.324 \\ \mu_1 \rightarrow 5.47395 & \mu_2 \rightarrow -2.61859 & x_1 \rightarrow 7.1716 & x_2 \rightarrow -3.54516 & f \rightarrow -140.681 \\ \mu_1 \rightarrow 0 & \mu_2 \rightarrow 0.121554 & x_1 \rightarrow -1.83158 & x_2 \rightarrow 2.74106 & f \rightarrow 7.1311 \\ \mu_1 \rightarrow 0 & \mu_2 \rightarrow 0.347387 & x_1 \rightarrow 2.20693 & x_2 \rightarrow -4.95422 & f \rightarrow -21.457 \\ \mu_1 \rightarrow 0 & \mu_2 \rightarrow 3.84252 & x_1 \rightarrow -6.66435 & x_2 \rightarrow -2.14649 & f \rightarrow -214.096 \end{array} \right)$$

conforme se puede observar en la tabla, el punto $x^* = (7.93836, 0.991178)$ es el único que obedece las condiciones de KKT ya que para este punto vale $\nabla f(x^*) = -\mu_1 \nabla g_1(x^*) - \mu_2 \nabla g_2(x^*)$ con $\mu_1 = 2.78139 > 0$, $\mu_2 = 1.89614 > 0$ y también que $\mu_1 g_1(x^*) + \mu_2 g_2(x^*) = 0$

5. TÉCNICAS DE TRANSFORMACIÓN

Si las restricciones $g_i(X)$ son funciones explícitas de las variables de decisión y son sencillas puede que sea posible efectuar una transformación en las variables independientes de forma que las restricciones sean automáticamente satisfechas, transformando así un problema con restricciones de desigualdad en uno sin restricciones. Así por ejemplo,

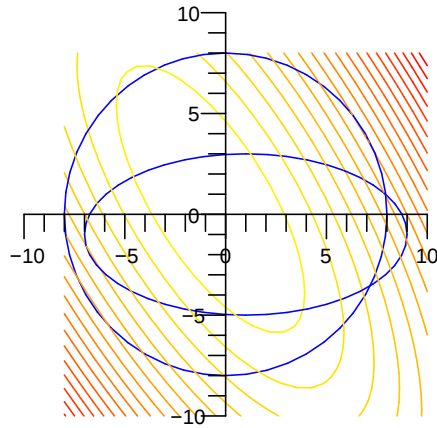


FIGURA 2. ilustración del problema

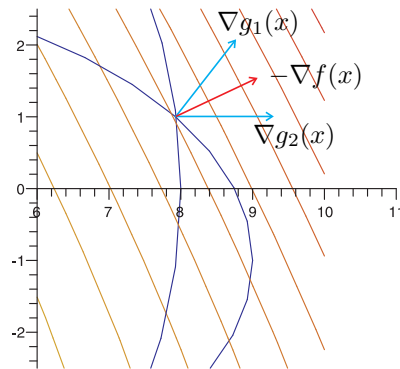


FIGURA 3. detalle del problema

- Las restricciones del tipo

$$l_i \leq x_i \leq u_i$$

pueden transformarse en las relaciones equivalentes

$$x_i = l_i + (u_i - l_i) \sin^2 y_i$$

donde y_i es la nueva variable independiente

- Las restricciones del tipo

$$x_i \geq 0$$

pueden transformarse en

$$x_i = y_i^2, \quad x_i = e^{y_i}, \quad \dots$$

Ejemplo: Determinar las dimensiones de un prisma rectangular con el volumen más grande cuando la suma de largura altura y anchura está restringida a un máximo

de 60 unidades y la largura no puede ultrapasar las 36 unidades. Este problema puede formularse como:

$$\max_{X \in \Omega(X)} x_1 x_2 x_3$$

donde

$$\Omega(X) = (x_1 + x_2 + x_3 \leq 60) \bigcap (0 \leq x_1 \leq 36) \bigcap (0 \leq x_2) \bigcap (0 \leq x_3)$$

introduciendo las nuevas variables $y_1 = x_1$, $y_2 = x_2$, $y_3 = x_1 + x_2 + x_3$ el problema se puede transformarse haciendo

$$\begin{aligned} y_1 &= 36 \sin^2 z_1 \\ y_2 &= 60 \sin^2 z_2 \\ y_3 &= 60 \sin^2 z_3 \end{aligned}$$

y el problema transformado, queda ahora

$$V = x_1 x_2 x_3 = y_1 y_2 (y_3 - y_1 - y_2) = 2160 \sin^2 z_1 \sin^2 z_2 (60 \sin^2 z_3 - 36 \sin^2 z_1 - 60 \sin^2 z_2)$$

y las condiciones necesarias de mínimo son

$$\begin{aligned} \frac{\partial V}{\partial z_1} &= 259200 \sin z_1 \cos z_1 \sin^2 z_2 (\sin^2 z_3 - \frac{6}{5} \sin^2 z_1 - \sin^2 z_2) &= 0 \\ \frac{\partial V}{\partial z_2} &= 518400 \sin^2 z_1 \sin z_2 \cos z_2 (\frac{1}{2} \sin^2 z_3 - \frac{3}{10} \sin^2 z_1 - \sin^2 z_2) &= 0 \\ \frac{\partial V}{\partial z_3} &= 259200 \sin^2 z_1 \sin^2 z_2 \sin z_3 \cos z_3 &= 0 \end{aligned}$$

de donde se saca $\cos z_3 = 0$, $\sin^2 z_1 = \frac{5}{9}$, $\sin^2 z_2 = \frac{1}{3}$ que resultan en los valores $x_1 = x_2 = x_3 = 20$.

6. PROGRAMACIÓN LINEAL

Hay un capítulo muy importante en la optimización que trata de los problemas lineales en la función objetivo y lineales también en las restricciones. El paradigma típico de programación lineal puede enunciarse como:

$$\begin{aligned} \min_{x \in \Omega} f(x) &= c_1 x_1 + c_2 x_2 + \cdots + c_n x_n = \sum_{i=1}^n c_i x_i \\ \Omega &= \left\{ \begin{array}{lll} g_1(x) &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &\leq b_1 \\ g_2(x) &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &\leq b_2 \\ \vdots &\vdots &\vdots \\ g_m(x) &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &\leq b_m \end{array} \right. \end{aligned}$$

modelos con esta estructura se encuentran en diversas áreas del conocimiento (ingeniería, economía, biología, etc). En este tipo de problemas la solución siempre se encuentra en la frontera ya que el gradiente de la función objetivo es constante.

$$\nabla f(x) = (c_1 \quad c_2 \quad \cdots \quad c_n)$$

Cuando $\Omega \neq \emptyset$ la solución se localiza necesariamente sobre la frontera de Ω . En los casos más sencillos (no degenerados) se sitúa en uno de los vértices del la superficie poliédrica inmersa en $\mathbb{R}^n$. Estos vértices son los puntos viables (están contenidos en) los

$$\binom{m}{n} = \frac{m!}{n!(m-n)!}$$

Ejemplo Un fabricante de disfraces posee estoques de 32m de dril, 22m de seda y 30m de satén y pretende producir dos modelos de disfraz. El primer modelo (M_1) consume 4m de dril, 2m de seda y 2m de satén. El segundo modelo (M_2) consume 2m de dril, 4m de seda y 6m de satén. Si M_1 se vende a 60 euros y M_2 a 100 euros, cuantas piezas de cada tipo deben producirse para obtener la receta máxima ? La solución pasa por la elaboración del modelo que se presenta como:

$$\begin{aligned} \max_x f(x) &= 60x_1 + 100x_2 \\ \text{sujeto a:} \\ 4x_1 + 2x_2 &\leq 32 \\ 2x_1 + 4x_2 &\leq 22 \\ 2x_1 + 6x_2 &\leq 30 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

Este modelo, para ser exacto, debería considerar valores enteros para las variables de decisión, lo que requeriría un procedimiento diferente al que adoptaremos. Relajaremos esta hipótesis admitiendo valores fraccionarios para las variables.

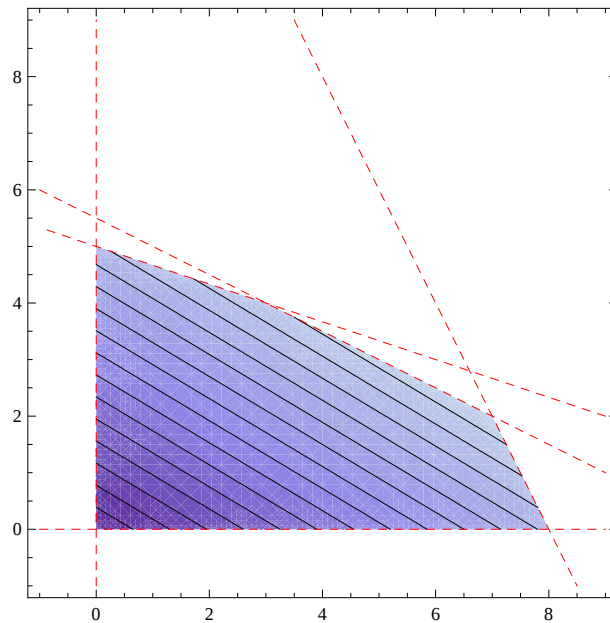


FIGURA 4. ilustración del problema

La solución se observa en el punto $x_1^* = 7$, $x_2^* = 2$. En este punto se verifican las condiciones de optimalidad de KKT, pues

$$\begin{aligned} \nabla f + \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 &= 0 \\ \text{o sea } \begin{pmatrix} -60 & -100 \end{pmatrix} + \lambda_1 \begin{pmatrix} 4 & 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 & 4 \end{pmatrix} &= 0 \\ \lambda_1 g_1(x^*) + \lambda_2 g_2(x^*) &= 0 \end{aligned}$$

de donde se saca $\lambda_1 = 3.333$, $\lambda_2 = 23.333$

Es digno de nota que si

$$f(x) = k(2x_1 + 4x_2) \text{ para } k > 0$$

el máximo de la función se alcanzaría sobre un segmento de recta. Las soluciones de esta índole, se llaman soluciones degeneradas.

6.1. Solución con la OPTIMIZATION toolbox de MATLAB. La función de MATLAB que resuelve el problema de programación lineal es

`x = linprog(c,A,b)`

que corresponden al modelo

$$\min_x cx \text{ con las restricciones } Ax \leq b$$

las matrices son:

`c = -[60 100];`

`A = [4 2; 2 4; 2 6; -1 0; 0 -1];`

`b = [32 22 30 0 0]';`

y la solución

`[x,valor,flag] = linprog(c,A,b)`

`x = [7.000 2.000]'`, `valor = -620.0000`, `flag = 1`

6.2. Método de Karmarkar. En 1984 Karmarkar propuso un nuevo método de resolución de problemas de programación lineal de grandes dimensiones que se mostró muy eficaz. El método es de punto interior ya que selecciona nuevas direcciones de búsqueda estrictamente en puntos interiores a la región viable. Este hecho contrasta con el método Simplex tradicional que evoluciona a lo largo de las aristas entre vértices de la frontera de la región viable. Para problemas de grande dimensión (150000 variables y 12000 restricciones) el método de Karmarkar es asta 50 veces más rápido que el método simplex en obtener la solución. [1]

7. TÉCNICAS DE BARRERA

Las técnicas de barrera o de penalidad transforman el problema de optimización original en una formulación alternativa de forma que la solución al problema con restricciones se obtiene a través de la solución de una secuencia convergente de problemas de mínimo sin restricciones. Así consideremos el problema básico como

$$\min F(X) \text{ con } X \in \Omega(X) = \bigcap_{i=1}^m g_i(X) \leq 0$$

Este problema se transforma en un problema equivalente sin restricciones construyendo una función del tipo

$$\phi_k(X, r_k) = f(X) + r_k \sum_{j=1}^n G_j(g_j(X))$$

donde $G_j(\cdot)$ es una función conocida como *función de penalidad* para la restricción $g_j(\cdot)$ y $r_k > 0$ es conocido como *parámetro de penalidad*.

Los métodos de barrera para tratar de restricciones de desigualdad se dividen en dos categorías:

- Métodos interiores

$$G_j(g_j) = -\frac{1}{g_j}$$

- Métodos exteriores

$$G_j(g_j) = \log(-g_j)$$

o también

$$G_j(g_j) = \max(0, g_j), \quad G_j(g_j) = (\max(0, g_j))^2$$

A guisa de ejemplo, mostramos la realización con penalidades del problema

$$\min x^2 \text{ tal que } (x \geq 1) \cap (x \leq 4)$$

La barrera interna se realiza con

$$\phi_k(x, r_k) = x^2 - \frac{1}{r_k} \left(\frac{1}{1-x} + \frac{1}{x-4} \right)$$

resultando en las gráficas que se pueden observar en la figura 5

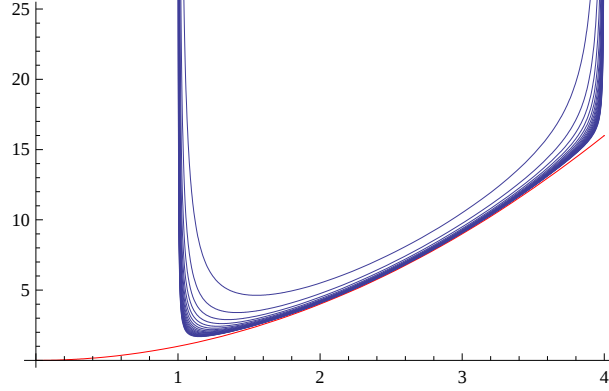


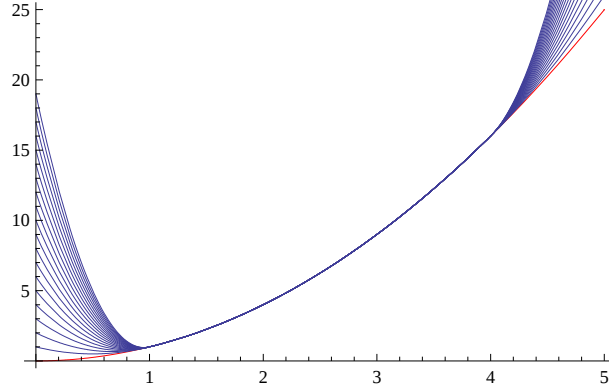
FIGURA 5. Evolución de ϕ_k interior

Análogamente la barrera externa se realiza con

$$\phi_k(x, r_k) = x^2 + r_k ((\max(1-x))^2 + (\max(x-4))^2)$$

resultando en las gráficas que se pueden observar en la figura 6

El método de la barrera externa tiene la ventaja de poder empezar el proceso de búsqueda desde un punto no perteneciente a la región viable.

FIGURA 6. Evolución de ϕ_k exterior

8. PROGRAMACIÓN LINEAL SECUENCIAL

Dado el problema general

$$\min_{x \in \Omega \subset \mathbb{R}^n} f(x), \quad \Omega = \left\{ \begin{array}{l} g_i(x) \leq 0 \\ h_j(x) = 0 \end{array} \right\}$$

en que las funciones f, g_i, h_j son suficientemente regulares, se puede plantear el problema lineal aproximado en cada punto, y progresar a través de estas aproximaciones, al punto óptimo. Para esto, en el punto actual x_k se linealiza el modelo como sigue:

$$\min_{x \in \Omega \subset \mathbb{R}^n} f(x_k) + \nabla f(x_k)(x - x_k), \quad \Omega = \left\{ \begin{array}{l} g_i(x_k) + \nabla g_i(x_k)(x - x_k) \leq 0 \\ h_j(x_k) + \nabla h_j(x_k)(x - x_k) = 0 \end{array} \right\}$$

8.1. Ejercicio. Resolver el modelo

$$\min_{x \in \Omega \subset \mathbb{R}^2} 24.6615x_1x_2, \quad \Omega = \left\{ \begin{array}{l} \frac{0.31831}{x_1x_2} - 1 \leq 0 \\ 1 - 0.0130226x_1^3x_2 \geq 0 \\ 1 - 0.2x_1 \leq 0 \\ 0.01x_1 - 1 \leq 0 \\ 1 - 2x_2 \leq 0 \\ 0.2x_2 - 1 \leq 0 \end{array} \right\}$$

La etapa de linealización en el punto x_k queda:

$$\begin{aligned} \min_{x \in \Omega \subset \mathbb{R}^2} f_k(x) &= 24.6615(x_1^k x_2 + x_1 x_2^k - x_1^k x_2^k) \\ \Omega &= \left\{ \begin{array}{l} -0.31831(x_1^k x_2 + x_1 x_2^k) + 0.95493x_1 x_2^k - (x_1^k)^2 (x_2^k)^2 \leq 0 \\ 1 + (x_1^k)^3 (0.0390678x_2^k - 0.0130226x_2) - 0.0390678x_1 (x_1^k)^2 x_2^k \geq 0 \\ 1 - 0.2x_1 \leq 0 \\ 0.01x_1 - 1 \leq 0 \\ 1 - 2x_2 \leq 0 \\ 0.2x_2 - 1 \leq 0 \end{array} \right\} \end{aligned}$$

que aún puede arreglarse en una forma más conveniente

$$\min_x c_k x \quad \text{con las restricciones} \quad A_k x + b_k \leq 0$$

donde

$$\begin{aligned}
 c_k &= 24.6615(x_2^k, x_1^k)^T \\
 A_k &= \begin{pmatrix} -\frac{0.31631}{(x_1^k)^2 x_2^k} & -\frac{0.31631}{x_1^k (x_2^k)^2} \\ -0.0390678(x_1^k)^2 x_2^k & -0.0130226(x_1^k)^3 \\ -0.2 & 0 \\ 0.01 & 0 \\ 0 & -2 \\ 0 & 0.2 \end{pmatrix} \\
 b_k &= \begin{pmatrix} \frac{0.95493}{x_1^k x_2^k} - 1 \\ -1 - 0.390678(x_1^k)^3 x_2^k \\ 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}
 \end{aligned}$$

9. INTRODUCCIÓN A LA PROGRAMACIÓN CUADRÁTICA

El problema típico de programación cuadrática puede presentarse como:

$$\min_{x \in \Omega} f(x) = c^T x + \frac{1}{2} x^T Q x, \quad \Omega = \begin{cases} Ax \leq b \\ x \geq 0 \end{cases}$$

Formando la Lagrangiana para el problema tendremos

$$L(x, \lambda) = c^T x + \frac{1}{2} x^T Q x + \lambda^T (Ax - b)$$

Las condiciones de KKT para el punto de mínimo son:

$$\begin{aligned}
 c^T + x^T Q + \lambda^T A &\geq 0 \\
 Ax - b &\leq 0 \\
 \lambda^T (Ax - b) &= 0 \\
 x &\geq 0 \\
 \lambda &\geq 0
 \end{aligned}$$

Este problema puede ser propuesto como un modelo de programación lineal, incorporando las variables artificiales y, v como sigue:

$$\min_{x \in \Omega} f(x) = y + v, \quad \Omega = \begin{cases} Qx + A^T \lambda - y + c^T = 0 \\ Ax + v = b \\ x \geq 0 \\ \lambda \geq 0 \\ y \geq 0 \\ v \geq 0 \\ \lambda^T x = 0 \\ \lambda v = 0 \end{cases}$$

Las dos últimas relaciones se utilizan en el algoritmo Simplex de la programación lineal para escoger las bases de inicio de búsqueda en el procedimiento. Lo que se pretende aquí es mostrar que el problema de programación cuadrática puede transformarse y resolverse a través de un algoritmo de programación lineal utilizado

en la realización de las condiciones de optimalidad de KKT. El modelo típico de Programación Cuadrática se resuelve utilizando la función

quadprog

dentro del **Optimization Toolbox** de Matlab

10. PROGRAMACIÓN CUADRÁTICA SECUENCIAL

Cualquier problema del tipo

$$\min_{x \in \Omega \subset \mathbb{R}^n} f(x), \quad \Omega = \left\{ \begin{array}{lcl} g_i(x) & \leq & 0 \\ h_j(x) & = & 0 \end{array} \right\}$$

lo suficientemente regular, puede aproximarse por una expansión de series de Taylor en una vecindad del punto actual x_k , y un problema de programación cuadrática equivalente

$$\begin{aligned} \min_{x \in \Omega \subset \mathbb{R}^n} f_{k+1}(x) &= f(x_k) + \nabla f_k(x - x_k) + \frac{1}{2}(x - x_k)^T H_k(x - x_k) \\ \Omega &= \left\{ \begin{array}{lcl} g_i(x_k) + \nabla g_i^k(x - x_k) & \leq & 0 \\ h_j(x_k) + \nabla h_j^k(x - x_k) & = & 0 \end{array} \right\} \end{aligned}$$

11. TOOLBOX DE OPTIMIZACIÓN DE MATLAB

Matlab nos ofrece un toolbox con métodos para determinación de mínimos de funciones con o sin restricciones. Estos métodos pueden ser apoyados en mayor o menor grado por la información sobre la regularidad de las funciones empleadas para describir el modelo. Así, se puede aumentar la eficacia de un método, incorporando las informaciones de las derivadas (gradiente y hessiana). El ejemplo 4.3 puede resolverse utilizando la función **fmincon**. Esta función puede utilizar mas o menos información en la obtención del resultado final. Las restricciones pueden darse agrupadas en las clases:

- Restricciones lineales (afines) de desigualdad: dadas en la forma $A_1 x \leq b_1$.
- Restricciones lineales (afines) de igualdad: dadas en la forma $A_2 x = b_2$.
- Restricciones no lineales de desigualdad: dadas en la forma $\{c_i(x) \leq 0\}$.
- Restricciones no lineales de igualdad: dadas en la forma $\{e_j(x) = 0\}$.
- Límites inferiores y superiores para las variables de decisión: dadas en la forma $l_b \leq x \leq u_b$.

La calidad del procedimiento puede aumentarse dándose la información de los gradientes de la función objetivo y de las restricciones no lineales. En el caso del ejemplo que sigue, tenemos que: $A_1 = \emptyset$, $b_1 = \emptyset$, $A_2 = \emptyset$, $b_2 = \emptyset$, $l_b = \emptyset$, $u_b = \emptyset$. La intención de utilizar la información de los gradientes, debe informarse en la declaración de **options**.

```
%-----
% Punto de partida
%-----
X0 = 10*(rand(1,2)-0.5*ones(1,2));
options = optimset('LargeScale','off');
options = optimset(options,'GradObj','on','GradConstr','on');
%-----
% No hay restricciones del tipo A x <= b
% No hay restricciones del tipo Ae x = be
% No hay restricciones del tipo límite superior x <= ub
% No hay restricciones del tipo límite inferior lb <= x
%-----
lb = [];
ub = [];
A = [];
```

```

b = [];
Ae = [];
be = [];
[X,fval] = fmincon(@objfungrad,X0,A,b,Ae,be,lb,ub,@confungrad,options)
%-----
% Verifica el valor de las restricciones en el punto solución
%-----
[c,ceq] = confungrad(X)

function [f, Df] = objfungrad(X)
x1 = X(1); x2 = X(2);
f = 10-(4*x1^2+4*x1*x2+2*x2^2+3*x1);
Df = [-8*x1 - 4*x2 - 3, -4*x1 - 4*x2]';

function [c,ceq,Dc,Dceq] = confungrad(X)
x1 = X(1); x2 = X(2);
g1 = x1^2+x2^2-64;
g2 = (x1-1)^2+4*(x2+1)^2-64;
Dg1 = [2*x1, 2*x2]';
Dg2 = [2*x1 - 2, 8*x2 + 8]';

c = [g1, g2];
Dc = [Dg1, Dg2];
ceq = [];
Dceq = [];

```

REFERENCIAS

1. *Singiresu S. Rao*
- **Engineering Optimization**
- John Wiley & Sons, Inc ISBN 0-471-55034-5.

ETSII - VIGO
E-mail address: cesareo@uvigo.es