Optimización con restricciones de igualdad

Jesús Getán y Eva Boj

Facultat d'Economia i Empresa
Universitat de Barcelona

Marzo de 2014

$$\left.\begin{array}{ll} \text{Opt} & f\,(x_1,\ldots,x_n) \\ \text{sujeta a:} & g_1(x_1,\ldots,x_n) = b_1 \\ & g_2(x_1,\ldots,x_n) = b_2 \\ & \ldots\ldots\ldots \\ & g_m(x_1,\ldots,x_n) = b_m \end{array}\right\} \text{ con } \left\{\begin{array}{l} m < n. \\ f\ ,\ g_i \in \mathcal{C}^2(D). \\ D \subset \mathbb{R}^n \ \text{ abierto.} \end{array}\right. \tag{1}$$

El conjunto factible es el definido por:

$$F = \{\vec{x} \in \mathbb{R}^n \mid g_1(\vec{x}) = b_1\,,\ldots\,,\, g_m(\vec{x}) = b_m\}.$$

La condición $m < n$ nos dice que el número de ecuaciones debe ser menor que el número de incognitas, ya que, en caso contrario el sistema puede no tener solución o una única solución y en consecuencia, resolver el programa no tiene sentido.

Método directo de solución

El método directo de solución por eliminación de variables consiste en reducir el problema con n variables y m restricciones a un problema de $n - m$ variables sin restricciones mediante la resolución del sistema de m ecuaciones de las restricciones.

La descripción formal del proceso es la siguiente:

Paso 1: En el sistema de ecuaciones que describe el conjunto factible tenemos m variables no básicas $\vec{x}_{NB}$ y $n-m$ variables básicas $\vec{x}_B$, despejamos las variables no básicas en función de las básicas,

$$\left.\begin{array}{l} g_1(x_1,\ldots,x_n)=b_1 \\ g_2(x_1,\ldots,x_n)=b_2 \\ \ldots\ldots\ldots \\ g_m(x_1,\ldots,x_n)=b_m \end{array}\right\} \Rightarrow \vec{x}_{NB}=\Phi\left(\vec{x}_B\right).$$

Paso 2:

En la función objetivo, sustituimos las variables no básicas por las básicas,

$$f(x_1, \ldots, x_n) = f(\vec{x}_B, \vec{x}_{NB}) = f(\vec{x}_B, \Phi(\vec{x}_B)).$$

Nuestro problema se ha reducido a un problema sin restricciones, lo resolvemos y encontramos la solución $\vec{x}_B^*$, entonces, la solución es

$$(\vec{x}_B, \Phi(\vec{x}_B)) = (\vec{x}_B^*, \vec{x}_{NB}^*).$$

Ejemplo:

En el problema

$$\begin{array}{ll} \min & x^2 + y^2 + z^2 - 4x - 8y - 16z \\ \text{sujeto a:} & x + y + z = 2 \\ & x + 2y = 0 \end{array}$$

despejamos las ($m = 2$ variables no básicas del sistema) variables x e z en función de la variable y ($n - m = 3 - 2 = 1$ variables básicas) y sustituimos en la función objetivo reduciendo el número de variables de la función objetivo a 1. Es decir,

$$\left.\begin{array}{l} x+y+z=2 \\ x+2y=0 \end{array}\right\} \Rightarrow \begin{array}{l} x=-2y \\ z=2+y \end{array} \Rightarrow$$

la función objetivo queda reducida a $f(y)=6y^2-12y-28$,

y resolviendo el problema como un programa de óptimos sin restricciones, obtenemos como resultado que $y=1$ y la solución del problema queda

$$(x^*, y^*, z^*) = (-2y, y, 2+y) = (-2, 1, 3).$$

Ejemplo:

En el problema

$$\begin{array}{ll} \text{Opt} & (x-1)^2 + y^2 \\ \text{sujeto a:} & y^2 = 4x \end{array} \tag{2}$$

Encontrar la solución gráficamente y por sustitución.

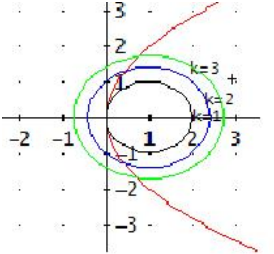

Figure: Dibujo problema

Al resolver el problema anterior ¿sale la solución correcta? En caso negativo, ¿por qué razón?

Dada la formulación general del problema de optimización con restricciones de igualdad, definimos la función de Lagrange correspondiente mediante:

$$L(x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_m) = f(\vec{x}) - \sum_{i=1}^{m} \lambda_i (g_i(\vec{x}) - b_i),$$

donde a los valores λ_i se les llama multiplicadores de Lagrange o también precios sombra.

Si $\vec{x}^o$ es un punto factible del problema entonces,

$$f(\vec{x}^o) = L(\vec{x}^o; \vec{\lambda}^o).$$

En lo que sigue, escribiremos para denotar el vector gradiente de una función $g_1(\vec{x})$:

$$\nabla g_1(\vec{x}) = \left(\frac{\partial g_1}{\partial x_1}, \frac{\partial g_1}{\partial x_2}, \cdots, \frac{\partial g_1}{\partial x_n}\right),$$

y para denotar al Jacobiano de la función $\vec{g}(\vec{x})$:

$$J\vec{g}(\vec{x}) = \begin{pmatrix} \nabla g_1(\vec{x}) \\ \vdots \\ \nabla g_m(\vec{x}) \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1(\vec{x})}{\partial x_1} & \cdots & \frac{\partial g_1(\vec{x})}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_m(\vec{x})}{\partial x_1} & \cdots & \frac{\partial g_m(\vec{x})}{\partial x_n} \end{pmatrix}.$$

Una definición importante es la siguiente:

Definition

Al punto $\vec{x}^o$ se le llama regular si y sólo si los vectores

$$\nabla g_1(\vec{x}^o)\,,\ \ \nabla g_2(\vec{x}^o)\,,\ldots,\ \nabla g_m(\vec{x}^o)$$

son linealmente independientes.

También se puede decir que el rango de la matriz Jacobiana

$$J\vec{g}(\vec{x}^o) = \begin{pmatrix} \nabla g_1(\vec{x}^o) \\ \vdots \\ \nabla g_m(\vec{x}^o) \end{pmatrix}$$

sea igual al número de restricciones m .

Sea la función lagrangiana

$$L(x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_m) = f(\vec{x}) - \lambda_1 (g_1(\vec{x}) - b_1) - \cdots - \lambda_m (g_m(\vec{x}) - b_m) \, .$$

$$L(\vec{x}; \vec{\lambda}) = f(\vec{x}) - \sum_{j=1}^{m} \lambda_j \left(g_j(\vec{x}) - b_j \right) .$$

Denotemos como $\nabla L(\vec{x}; \vec{\lambda})$ el gradiente de la función de Lagrange.

Theorem

Dado un punto $\vec{x}^o \in D$ *que es regular.*

En $\vec{x}^o$ *la función* $f(\vec{x})$ *alcanza un óptimo local, entonces* $\nabla L(\vec{x}^o; \vec{\lambda}^o) = \vec{0}$ *para cierto* $\vec{\lambda}^o \in \mathbb{R}^m$.

Denotemos como $HL(\vec{x};\vec{\lambda})$ la matriz hessiana de la función de Lagrange. Desarrollada,

$$HL(\vec{x};\vec{\lambda}) = \begin{pmatrix} \frac{\partial^2 L(\vec{x};\vec{\lambda})}{\partial^2 x_1} & \cdots & \frac{\partial^2 L(\vec{x};\vec{\lambda})}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 L(\vec{x};\vec{\lambda})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 L(\vec{x};\vec{\lambda})}{\partial^2 x_n} \end{pmatrix}.$$

Theorem

Sea dado un punto $\vec{x}^o$ que es regular y que satisface la condición necesaria ($\nabla L(\vec{x}^o; \vec{\lambda}^o) = \vec{0}$).

Si para todo $\vec{v} \in R^n \backslash \{0\}$ tal que

$$\left.\begin{array}{c} \nabla g_1(\vec{x}^o)\vec{v} = 0 \\ \cdots\cdots \\ \nabla g_m(\vec{x}^o)\vec{v} = 0 \end{array}\right\}$$

se cumple que

$$\vec{v}^T HL(\vec{x}^o; \vec{\lambda}^o)\vec{v} < 0$$

entonces $\vec{x}^o$ es máximo local del programa.

Al vector $\vec{v}$ se le conoce como vector de dirección factible.

Theorem

Sea dado un punto $\vec{x}^o$ que es regular y satisface la condición necesaria $\left(\nabla L(\vec{x}^o;\vec{\lambda}^o)=\vec{0}\right)$.
Si para todo $\vec{v}\in R^n\backslash\{0\}$ tal que

$$\left.\begin{array}{c}\nabla g_1(\vec{x}^o)\vec{v}=0\\ \cdots\cdots\\ \nabla g_m(\vec{x}^o)\vec{v}=0\end{array}\right\}$$

se cumple que

$$\vec{v}^T HL(\vec{x}^o;\vec{\lambda}^o)\vec{v}>0$$

entonces, $\vec{x}^o$ es mínimo local del programa.

Aplicamos al problema (2) el método de Lagrange:

$$L(x, y; \lambda) = (x-1)^2 + y^2 - \lambda(y^2 - 4x).$$

Condición necesaria:

$$\left.\begin{array}{rcl} \frac{\partial L(x,y;\lambda)}{\partial x} & = & 2(x-1) + 4\lambda = 0 \\ \frac{\partial L(x,y;\lambda)}{\partial y} & = & 2y - 2\lambda y = 0 \\ \frac{\partial L(x,y;\lambda)}{\partial \lambda} & = & -y^2 + 4x = 0 \end{array}\right\} 2y(1-\lambda) = 0 \Rightarrow y = 0 \text{ ó } \lambda = 1.$$

Si $y = 0 \Rightarrow x = 0 \Rightarrow -2 + 4\lambda = 0 \Rightarrow \lambda = \frac{1}{2}$. Candidato $(0, 0; \frac{1}{2})$.

Si $\lambda = 1 \Rightarrow 2(x-1) + 4 = 0 \Rightarrow x - 1 = -2 \Rightarrow x = -1 \Rightarrow y^2 = -4$ que no tiene resultado en el campo de los números reales.

Condición suficiente:

$$HL(x,y;\lambda) = \begin{pmatrix} \frac{\partial^2 L(x,y;\lambda)}{\partial^2 x} & \frac{\partial^2 L(x,y;\lambda)}{\partial x \partial y} \\ \frac{\partial^2 L(x,y;\lambda)}{\partial y \partial x} & \frac{\partial^2 L(x,y;\lambda)}{\partial^2 y} \end{pmatrix}.$$

En nuestro caso:

$$HL(x,y;\lambda) = \begin{pmatrix} 2 & 0 \\ 0 & 2-2\lambda \end{pmatrix} \rightsquigarrow HL(0,0;\tfrac{1}{2}) = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix},$$

que es definida positiva, y por tanto el candidato es un mínimo local o condicionado del problema.

Theorem

Sea el problema general (1) que para $\vec{b} = (b_1, \cdots, b_m)$ *la función objetivo posee un óptimo local sobre el conjunto factible en el punto* $(\vec{x}^*; \vec{\lambda}^*)$ *que es regular. Entonces,*

$$\frac{\partial f(\vec{b})}{\partial b_i} = \lambda_i^* \quad \text{para } i = 1, \cdots, m.$$

Las restricciones de igualdad obligan al uso de toda capacidad de los inputs (uso total de los recursos), es decir, no podemos disponer de más unidades de recursos o dejar de disponer algunas unidades.

La pregunta es ¿cuánto estaríamos dispuestos a pagar para que se nos liberase del cumplimiento de la restricción? ó ¿cuál es el precio sombra de la restricción?

Véase un ejemplo.

Un empresario necesita K unidades de capital y L horas de trabajo para producir Q unidades de un producto, según la función de producción

$$Q(K, L) = 60K^{\frac{1}{2}}L^{\frac{1}{3}}$$

El precio del capital es de 20$ por unidad y el coste de una hora de trabajo es de 15$.

(1) ¿Qué cantidad de capital y trabajo producen 4200 unidades de producto con un coste mínimo?

2) ¿Cómo varía aproximadamente el coste mínimo si se han de producir 4250 unidades?

La función de Lagrange: $(\mathcal{L}(K, L; \lambda) = L(K, L; \lambda))$

$$\mathcal{L}(K, L; \lambda) = 20K + 15L - \lambda(60K^{\frac{1}{2}}L^{\frac{1}{3}} - 4200).$$

El punto crítico es $(171.62, 152.55)$ con multiplicador $\lambda = 1.63$.
La hessiana es

$$H\mathcal{L}(K, L, \lambda) = \begin{pmatrix} 15\lambda K^{-\frac{3}{2}}L^{\frac{1}{3}} & -10\lambda K^{-\frac{1}{2}}L^{-\frac{2}{3}} \\ -10\lambda K^{-\frac{1}{2}}L^{-\frac{2}{3}} & \frac{40}{3}\lambda K^{\frac{1}{2}}L^{-\frac{5}{3}} \end{pmatrix},$$

que evaluada en el punto crítico, $H\mathcal{L}(171.62, 152.55)$, es definida positiva y por lo tanto el punto crítico es un mínimo local o condicionado del problema.

El coste mínimo es

$$C(171.62, 152.55) = 20 \times 171.62 + 15 \times 152.55 = 5720.65\$.$$

Si la producción aumenta en 50 unidades, el coste mínimo varia aproximadamente en $1.63 \times 50 = 81.5$ \$.