

Capítulo 7

Teoría Lagrangiana

7.1. Condiciones para problemas con restricciones de igualdad.

Se desea resolver el problema

$$\begin{aligned} P : \text{mín} \quad & f(x) \\ \text{s.a:} \quad & h_i(x) = 0 \quad i = 1, 2, \dots, m \end{aligned} \quad (7.1)$$

donde $f: \mathbb{R}^n \rightarrow \mathbb{R}$ y $h_i: \mathbb{R}^n \rightarrow \mathbb{R}$ son funciones diferenciables con continuidad. Por simplicidad $h(x) = (h_1(x), \dots, h_m(x))$.

A este problema (P) le asociamos una función, denominada función lagrangiana $L(x, \lambda) = f(x) + \lambda' h(x)$ que tiene la importante propiedad de que si x^* es un mínimo local de (P), entonces existe un único multiplicador λ^* de forma que:

$$\nabla f(x^*) + \nabla h(x^*) \lambda^* = 0.$$

Nótese que si la intersección de las restricciones $h_i(x) = 0$ formaran un recinto convexo entonces podríamos utilizar las técnicas presentadas en el capítulo 3 y 4.

Lo que se va a desarrollar es una herramienta alternativa a los métodos iterados del tema anterior que permitirá identificar los mínimos locales del problema (7.1).

Teorema 7.1.1 *Sea x^* un mínimo local de f con $h(x) = 0$. Supongamos que las restricciones verifican $\nabla h_1(x^*), \dots, \nabla h_m(x^*)$ son linealmente independientes (x^* es regular). Entonces existe un único vector $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)$ tal que:*

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) = 0.$$

Si f y h son dos veces diferenciables con continuidad se tiene

$$y'(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 h_i(x^*))y \geq 0 \quad \forall i = 1, \dots, m \quad \forall y \in V(x^*)$$

$$V(x^*) = \{y : \nabla h_i(x^*)' y = 0, \forall i = 1, \dots, m\}.$$

La interpretación geométrica de la condición del teorema de multiplicadores es que si x^* es mínimo local de (P), el gradiente de la función objetivo debe pertenecer al subespacio generado por los gradientes de las restricciones en x^* . Esto es, $\nabla f(x^*) + \nabla h(x^*)\lambda^{*'} = 0$. Igualmente se puede interpretar como que el gradiente de la función objetivo en el óptimo debe ser ortogonal al espacio de variaciones factibles de primer orden $V(x^*)$.

La condición suficiente nos la da el siguiente teorema

Teorema 7.1.2 Sean f y h dos veces diferenciables con continuidad y sean x^* y λ^* tales que

- i) $\nabla_x L(x^*, \lambda^*) = 0, \quad \nabla_\lambda L(x^*, \lambda^*) = 0$
- ii) $y' \nabla_{xx}^2 L(x^*, \lambda^*) y > 0, \quad \forall y \neq 0, \quad \nabla h(x^*)' y = 0$

entonces x^* es un mínimo local estricto de f sujeto a $h(x) = 0$. De hecho, existen $\gamma > 0, \varepsilon > 0$ tales que:

$$f(x) \geq f(x^*) + \frac{\gamma}{2} \|x - x^*\|^2 \quad \forall x, h(x) = 0$$

y

$$\|x - x^*\| < \varepsilon.$$

Los multiplicadores tienen una interpretación económica muy interesante. Se pueden entender como la tasa de cambio del valor objetivo como función del lado derecho de las restricciones (nivel de las restricciones).

Lo veremos en el caso de un problema con una única restricción lineal.

$$\begin{array}{ll} \text{mín} & f(x) \\ \text{sa} & a'x = b. \end{array} \quad (7.2)$$

Supongamos que x^*, λ^* son la solución óptima y el multiplicador óptimo. Si cambiamos b a $b + \Delta b$ el mínimo cambiará a $x^* + \Delta x$. Δb y Δx están relacionados mediante

$$a'(x^* + \Delta x) = b + \Delta b,$$

entonces como $a'x^* = b$ se tiene

$$a' \Delta x = \Delta b.$$

Usemos la condición de Lagrange $\nabla f(x^*) = -\lambda^* a'$

$$\begin{aligned} \Delta \text{coste} &= f(x^* + \Delta x) - f(x^*) \\ &= \nabla f(x^*)' \Delta x + o(\|\Delta x\|) \\ &= -\lambda^* a' \Delta x + o(\|\Delta x\|) \\ &= -\lambda^* \Delta b + o(\|\Delta x\|) \end{aligned}$$

por tanto

$$\lambda^* = \frac{-\Delta \text{coste}}{\Delta b}$$

hasta el orden 1.

Este resultado es cierto en general para restricciones no necesariamente lineales, como puede verse en la proposición ??.

Resolver el problema siguiente haciendo uso de la teoría lagrangiana:

$$\min \sum_{i=1}^n x_i$$

s.a.

$$x_1 x_2 \dots x_n = 1, x_i > 0, i = 1, \dots, n$$

Probar haciendo uso del resultado anterior la desigualdad:

$$(x_1 x_2 \dots x_n)^{\frac{1}{n}} \leq \frac{\sum_{i=1}^n x_i}{n}$$

$$L(x_1, \lambda) = \sum_{i=1}^n x_i + \lambda (\prod_{i=1}^n x_i - 1)$$

$$\nabla L(x_i, \lambda) = \begin{pmatrix} 1 + \lambda \prod_{i \neq 1}^n x_i \\ 1 + \lambda \prod_{i \neq 2}^n x_i \\ \dots \\ 1 + \lambda \prod_{i \neq n}^n x_i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix} \Rightarrow \lambda = \frac{-1}{\prod_{i \neq 1} x_i} = \frac{-1}{\prod_{i \neq 2} x_i} = \dots = \frac{-1}{\prod_{i \neq n} x_i} \Rightarrow x_i = x_j, \forall i, j$$

Como $x_1 x_2 \dots x_n = 1$ entonces $x_i^n = 1, \forall i$ y $x_i = \sqrt[n]{1}, \forall i$

Como $\lambda = \frac{-1}{\prod_{i \neq 1} x_i} = \frac{-1}{x_i^{n-1}}$ entonces $\lambda = -\sqrt[n]{1}$

El valor mínimo es $\sum_{i=1}^n x_i = n \sqrt[n]{1} = n$ ya que $x_i > 0, \forall i$ y entonces $x_i = 1, \forall i$

Para probar la desigualdad $(x_1 x_2 \dots x_n)^{\frac{1}{n}} \leq \frac{\sum_{i=1}^n x_i}{n}$

Resolvemos el problema:

$$\min \sum_{i=1}^n x_i$$

s.a.

$$x_1 x_2 \dots x_n = u, x_i > 0, i = 1, \dots, n$$

$$L(x_1, \lambda) = \sum_{i=1}^n x_i + \lambda (\prod_{i=1}^n x_i - u)$$

$$\nabla h(x) = \left[\prod_{j \neq 1} x_j, \dots, \prod_{j \neq n} x_j \right]$$

$$y' \nabla h(x^*) = [y_1 \dots y_n] \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \sum_{i=1}^n y_i = 0.$$

$$\nabla^2 L(x^*, \lambda^*) = \begin{bmatrix} 0 & -1 & \dots & -1 \\ -1 & 0 & & \\ \vdots & & \ddots & \\ -1 & & & 0 \end{bmatrix}$$

$$\frac{\partial^2 L}{\partial x_i \partial x_j} = \lambda \prod_{k \neq i, j} x_k$$

$$y \in V(x^*) \quad y' \nabla^2 h(x^*, \lambda^*) y = \left[-\sum_{k \neq 1} y_k, \dots, \sum_{k \neq n} y_k \right]$$

$$= -\sum_{k=1}^n y_i \sum_{k \neq i} y_k = -\sum_{i=1}^n y_i (-y_i) \neq 0.$$

7.2. Condiciones para problemas con desigualdades

$$\min \quad f(x) \tag{7.3}$$

$$sa: \quad h_1(x) = 0, \dots, h_m(x) = 0$$

$$g_1(x) \leq 0, \dots, g_r(x) \leq 0$$

que es equivalente al problema

$$\begin{aligned} \text{mín} \quad & f(x) \\ \text{sa :} \quad & h(x) = 0 \\ & g(x) \leq 0 \end{aligned} \tag{7.4}$$

En un punto x , $A(x) = \{j : g_j(x) = 0\}$ es el conjunto de restricciones activas.

Si x^* es un mínimo local de (7.4) es también un mínimo local del problema en el que las restricciones inactivas en x^* se han eliminado. Por otra parte en el entorno de x^* las restricciones activas se pueden considerar como igualdades de manera que x^* será un mínimo local de:

$$\begin{aligned} \text{mín} \quad & f(x) \\ \text{sa :} \quad & h(x) = 0 \\ & g_j(x) = 0, \quad \forall j \in A(x^*) \end{aligned} \tag{7.5}$$

Por tanto si x^* es regular (los gradientes son l.i.) existirán multiplicadores $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)$ y $\mu_j^* \quad \forall j \in A(x^*)$ tales que:

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) + \sum_{j \in A(x^*)} \mu_j^* \nabla g_j(x^*) = 0$$

y si asociamos $\mu_j^* = 0 \quad \forall j \notin A(x^*)$ tendríamos una definición similar a la del caso con igualdades.

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^r \mu_j^* \nabla g_j(x^*) &= 0 \\ \mu_j^* &= 0 \quad \forall j \notin A(x^*) \end{aligned}$$

Además se puede argumentar intuitivamente que los μ_j^* son no negativos. En efecto, si consideramos las restricciones $g_j(x) \leq b_j$ con $b_j > 0$ el valor óptimo del problema será menor pues la región factible ha crecido. Por tanto como sabemos que $\mu_j^* = \frac{-\Delta \text{ coste debido a } b_j}{b_j} > 0$ (pues $(\Delta \text{ coste debido a } b_j) \leq 0$ y los $b_j > 0$)

A continuación daremos la demostración rigurosa

Teorema 7.2.1 Sea x^* un mínimo local de (7.4) y x^* es regular. Entonces existen multiplicadores únicos $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)$, $\mu^* = (\mu_1^*, \dots, \mu_r^*)$ tales que

$$\begin{aligned} \nabla_x L(x^*, \lambda^*, \mu^*) &= 0, \\ \mu_j^* &\geq 0 \quad \forall j \\ \mu_j^* &= 0, \quad \forall j \notin A(x^*) \end{aligned}$$

donde $A(x^*) = \{j : g_j(x^*) = 0\}$. Si además las funciones son dos veces diferenciables con continuidad se verifica que

$$y' \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) y \geq 0, \quad \forall y \in V(x^*)$$

donde $V(x^*) = \{y | \nabla h_i(x^*)' y = 0 \quad \forall i, \quad \nabla g_j(x^*)' y = 0 \quad \forall j \in A(x^*)\}$

Proposición 7.2.2 (Condición suficiente de segundo orden) *Supongamos que f , h , y g son dos veces diferenciables con continuidad, y sean $x^* \in \mathbb{R}^n$, $\lambda^* \in \mathbb{R}^m$ y $\mu^* \in \mathbb{R}^r$ satisfaciendo*

$$\nabla_x L(x^*, \lambda^*, \mu^*) = 0, \quad h(x^*) = 0, \quad g(x^*) \leq 0, \quad (7.6)$$

$$\mu_j^* \geq 0, \quad j = 1, \dots, r \quad (7.7)$$

$$\mu_j^* = 0, \quad j \notin A(x^*), \quad (7.8)$$

$$y' \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) y > 0, \quad \text{para todo } y \neq 0 \quad \text{con } y \in V(x^*) \quad (7.9)$$

donde

$$V(x^*) = \{y \mid \nabla h_i(x^*) y' = 0, i = 1, \dots, m, \nabla g_j(x^*) y' = 0, j \in A(x^*)\}. \quad (7.10)$$

Asumimos también que

$$\mu_j^* > 0, \quad \forall j \in A(x^*). \quad (7.11)$$

Entonces x^* es un mínimo local estricto de f sujeto a $h(x) = 0, g(x) \leq 0$.

La condición (7.11) es conocida como la *condición de complementariedad estricta*.