

Programación no lineal

Jesús Getán y Eva Boj

Facultat d'Economia i Empresa
Universitat de Barcelona

Marzo de 2014

Programación no lineal

- Formulación general del problema

- Óptimos locales

- Condición de regularidad

- Condiciones de Kuhn-Tucker

Regla para resolver el problema general

- Algoritmo básico: La Regla

- Una visión geométrica práctica de la solución

- Un ejemplo académico clarificador

- Condición suficiente

- Continuación del ejemplo académico

Óptimos globales

$$\left. \begin{array}{l} \text{Max } f(x_1, \dots, x_n) \\ g_1(x_1, \dots, x_n) \leq 0. \\ g_2(x_1, \dots, x_n) \leq 0. \\ \dots\dots\dots \\ g_m(x_1, \dots, x_n) \leq 0. \end{array} \right\} \quad \text{con} \quad \left\{ \begin{array}{l} f, g_i \in \mathcal{C}^2(D), \\ D \subset \mathbb{R}^n \text{ abierto.} \end{array} \right. \quad (1)$$

El conjunto de soluciones factibles es:

$$K = \{\vec{x} \in \mathbb{R}^n \mid g_1(\vec{x}) \leq 0, \dots, g_m(\vec{x}) \leq 0\}.$$

El conjunto factible consta de interior, vértices y frontera que no es vértice.

Las funciones $f(\vec{x})$ y $g_i(\vec{x})$ con $i = 1, \dots, m$, pueden ser lineales o no lineales.

Las restricciones pueden ser con $\geq$ o $=$, con objeto de

(i) estandarizar el problema y facilitar su estudio

y

(ii) resolución

haremos la siguiente tabla de conversión:

Si son de la forma	$g_i(\vec{x}) \leq b_i,$	se cambia por	$g_i(\vec{x}) - b_i \leq 0.$
Si son de la forma	$g_i(\vec{x}) \geq b_i,$	se cambia por	$b_i - g_i(\vec{x}) \leq 0.$
Si son de la forma	$g_i(\vec{x}) = b_i,$	se cambian por	$g_i(\vec{x}) - b_i \leq 0.$ $b_i - g_i(\vec{x}) \leq 0.$

El estudio que hacemos a continuación es paralelo al que hemos hecho para los programas matemáticos con restricciones de igualdad. Primero, definimos los conceptos de restricciones saturadas.

Definition

Dado el programa (1). El punto $\vec{x}^o \in K$ se dice que satura una restricción $g(\vec{x})$ (o que es una restricción activa) si y sólo si al sustituirlo en la restricción cumple $g(\vec{x}^o) = 0$.

Nótese que, si un punto factible satura una restricción éste la cumple con el signo igual, en caso contrario, si no satura la restricción o no es activa éste será $<$.

Cosideramos un punto $\vec{x}^o \in K$ que es óptimo para el programa (1).

El punto $\vec{x}^o \in K$ saturará algunas restricciones, sea Γ el conjunto de índices de las restricciones saturadas o activas por $\vec{x}^o$.

Definition

Dado el programa (1). Un punto $\vec{x}^o \in K$ es regular si y sólo si los gradientes de las restricciones saturadas o activas en $\vec{x}^o$ forman un conjunto de vectores linealmente independiente.

Dicho de otra manera, si Γ es conjunto de índices de las restricciones saturadas (o activas) y $|\Gamma|$ es el número de restricciones saturadas, tenemos

$$\text{rango } Jg(\vec{x}^o) = \text{rango} \begin{pmatrix} \nabla g_i(\vec{x}^o) \\ \vdots \\ \nabla g_k(\vec{x}^o) \end{pmatrix} = |\Gamma|.$$

$$k \in \Gamma$$

Como punto de partida utilizaremos el problema de programación no lineal en forma canónica (1) y consideraremos el punto $\vec{x}^0 \in K$ como máximo local del problema (1).

Sabemos, por el teorema de Lagrange, que si $\vec{x}^o$ es regular y máximo local de (1) y satura o hace activas $k = |\Gamma|$ restricciones, entonces

existen $\lambda_1^o, \dots, \lambda_k^o \in R$, únicos con $\lambda_i^o > 0$, $\forall i \in \Gamma$ tales que

$$\nabla f(\vec{x}^o) = \lambda_1^o \nabla g_1(\vec{x}^o) + \dots + \lambda_k^o \nabla g_k(\vec{x}^o).$$

Nuestro problema es encontrar el punto donde la función alcanza el valor máximo, por esta razón, no conocemos las restricciones saturadas con antelación.

Para salvar esta situación, añadiremos el resto de las restricciones (las no saturadas o no activas) y diremos que $\lambda_i^o = 0$ para todas ellas y con el fin de garantizar la elección de los signos de λ introducimos la siguiente ecuación

$$\lambda_i^o g_i(\vec{x}^o) = 0 \text{ para todo } i = 1, 2, \dots, m.$$

Esta relación nos dice que

- (i) Si el punto $\vec{x}^o$ satura la restricción i ($g_i(\vec{x}^o) = 0$) $\Rightarrow \lambda_i^o > 0$.
- (ii) Si el punto $\vec{x}^o$ no satura la restricción i ($g_i(\vec{x}^o) < 0$) $\Rightarrow \lambda_i^o = 0$.

Condición necesaria de máximo:

Theorem

Dado el programa (1) con $\vec{x}^o$ regular.

Si $\vec{x}^o$ es máximo local $\Rightarrow$ debe haber multiplicadores

$\lambda_1^o, \dots, \lambda_m^o \in R$, tales que satisfagan

- a) $\nabla f(\vec{x}^o) - \lambda_1^o \nabla g_1(\vec{x}^o) - \dots - \lambda_m^o \nabla g_m(\vec{x}^o) = 0$,
- b) $\lambda_i^o g_i(\vec{x}^o) = 0$, $\forall i \in \{1, \dots, m\}$,
- c) $\lambda_i^o \geq 0$, $\forall i \in \{1, \dots, m\}$,
- d) $g_i(\vec{x}^o) \leq 0$, $\forall i \in \{1, \dots, m\}$.

Al igual que en los problemas con restricciones de igualdad, el multiplicador λ_i^o se podría considerar como el precio sombra para la i -ésima restricción de (1). Observamos que también **debemos** incluir en las restricciones las de no negatividad (i.e. $x_i \geq 0$).

La regla

El enfoque básico consiste en convertir el problema (1) en uno sin restricciones mediante la función de Lagrange.

Lo haremos por pasos.

Paso 1. Escribir la función lagrangiana

$$L(\vec{x}; \vec{\lambda}) = f(\vec{x}) - \sum_{i=1}^m \lambda_i g_i(\vec{x}).$$

Nótese que están todas las restricciones. Los λ_i son los multiplicadores de Lagrange asociados a las restricciones.

Paso 2. Calculamos el Jacobiano de las restricciones

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

Paso 3. Calcular el gradiente de la función lagrangiana.

$\nabla_x L(\vec{x}; \vec{\lambda})$ e igualar a cero.

$$\frac{\partial L(\vec{x}; \vec{\lambda})}{\partial x_i} = \frac{\partial f(\vec{x})}{\partial x_i} - \sum_{i=1}^m \lambda_i \frac{g_i(\vec{x})}{\partial x_i} = 0 \text{ para todo } i \in \{1, \dots, n\}.$$

Calcularemos las derivadas parciales con respecto a cada una de las variables de decisión y las igualaremos a cero.

Paso 4. Imponemos las condiciones de holgura complementaria

$$\lambda_i \geq 0 \text{ para todo } i \in \{1, \dots, m\}.$$

Paso 5. Decisiones sobre el signo de λ_i

$$\lambda_i g_i(\vec{x}) = 0 \text{ para todo } i = 1, 2, \dots, m.$$

Paso 6. Escribimos las condiciones de factibilidad de la solución

$$g_i(\vec{x}) \leq 0 \text{ para todo } i = 1, 2, \dots, m.$$

Los posibles resultados del problemas serán la o las soluciones del siguiente sistema:

$$\left. \begin{aligned} \frac{\partial f(\vec{x})}{\partial x_i} - \sum_{i=1}^m \lambda_i \frac{g_i(\vec{x})}{\partial x_i} &= 0 \text{ para todo } i \in \{1, \dots, n\}. \\ \lambda_i &\geq 0 \text{ para todo } i \in \{1, \dots, m\}. \\ \lambda_i g_i(\vec{x}) &= 0 \text{ para todo } i = 1, 2, \dots, m. \\ g_i(\vec{x}) &\leq 0 \text{ para todo } i = 1, 2, \dots, m. \end{aligned} \right\} \rightarrow$$

→ El $\vec{x}^o$ que es candidato a óptimo.

Las soluciones de (1) tienen que estar en el conjunto factible K , más concretamente en uno de los conjuntos siguientes:

$$\left\{ \begin{array}{l} O_1 \text{ conjunto de puntos del interior de } K. \\ O_2 \text{ conjunto de puntos de la frontera que no sean vértices de } K. \\ O_3 \text{ conjunto de puntos que son vértices de } K. \end{array} \right.$$

Esta idea, nos permitirá localizar las soluciones e interpretar su posición geométrica en el problema.

Encontrar la solución del problema

$$\begin{array}{ll}\max & 3x_1 + x_2, \\ \text{s.a:} & x_1^2 + x_2^2 - 5 \leq 0, \\ & x_1 - x_2 - 1 \leq 0.\end{array}$$

Paso 1.

$$L(x_1, x_2 : \lambda_1, \lambda_2) = 3x_1 + x_2 - \lambda_1 (x_1^2 + x_2^2 - 5) - \lambda_2 (x_1 - x_2 - 1).$$

Paso 2.

$$Jg(\vec{x}^o) = \begin{pmatrix} 2x_1 & 2x_2 \\ 1 & -1 \end{pmatrix}$$

Paso 3.

$$\frac{\partial L(\vec{x}; \vec{\lambda})}{\partial x_1} = 3 - 2\lambda_1 x_1 - \lambda_2 = 0,$$

$$\frac{\partial L(\vec{x}; \vec{\lambda})}{\partial x_2} = 1 - 2\lambda_1 x_2 + \lambda_2 = 0.$$

Paso 4.

$$\lambda_1 \geq 0, \lambda_2 \geq 0.$$

Paso 5.

$$\begin{aligned}\lambda_1 (x_1^2 + x_2^2 - 5) &= 0, \\ \lambda_2 (x_1 - x_2 - 1) &= 0.\end{aligned}$$

Paso 6.

$$\begin{aligned}x_1^2 + x_2^2 - 5 &\leq 0, \\ x_1 - x_2 - 1 &\leq 0.\end{aligned}$$

Escribimos el sistema de inecuaciones que tenemos que resolver en una matriz

$$\begin{array}{lll}
 3 - 2\lambda_1 x_1 - \lambda_2 = 0 & \lambda_2 \geq 0 & x_1^2 + x_2^2 - 5 \leq 0 \\
 1 - 2\lambda_1 x_2 + \lambda_2 = 0 & \lambda_1 (x_1^2 + x_2^2 - 5) = 0 & x_1 - x_2 - 1 \leq 0 \\
 \lambda_1 \geq 0 & \lambda_2 (x_1 - x_2 - 1) = 0 &
 \end{array}$$

Nos ayudamos a determinar el caso que tenemos que resolver mediante un diagrama basado en que la solución debe saturar ($\lambda \neq 0$) o no saturar ($\lambda = 0$) las restricciones.

λ_1		λ_2	Significado	La solución está
$\neq 0$		$\neq 0$	satura las dos restricciones	en un vértice
		$= 0$	satura la 1 y no satura la 2	en frontera no vértice
$= 0$		$\neq 0$	no satura la 1 y satura la 2	en frontera no vértice
		$= 0$	no satura la 1 y no satura la 2	en el interior

Mediante un dibujo del conjunto factible y las curvas de nivel de la función objetivo sabemos que la solución está en un vértice, en consecuencia, el caso a estudiar es el $\lambda_1 \neq 0$ y $\lambda_2 \neq 0$.

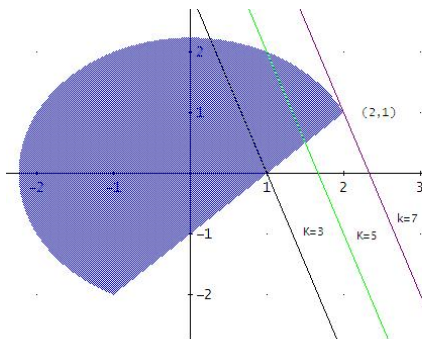


Figure: gráfica

Sustituimos la información en las ecuaciones, en nuestro caso, que las restricciones se saturan

$$\begin{array}{lll} 3 - 2\lambda_1 x_1 - \lambda_2 = 0 & \lambda_2 > 0 & x_1^2 + x_2^2 - 5 = 0 \\ 1 - 2\lambda_1 x_2 + \lambda_2 = 0 & \lambda_1 (x_1^2 + x_2^2 - 5) = 0 & x_1 - x_2 - 1 = 0 \\ \lambda_1 > 0 & \lambda_2 (x_1 - x_2 - 1) = 0 & \end{array}$$

Resolvemos el sistema por sustitución

$$\left. \begin{array}{l} x_1^2 + x_2^2 - 5 = 0 \\ x_1 - x_2 - 1 = 0 \end{array} \right\} \rightarrow \text{nos da dos soluciones}$$

si $x_2 = -2$ entonces $x_1 = -1$.

si $x_2 = 1$ entonces $x_1 = 2$.

Resolvemos el sistema siguiente para cada caso

$$\left. \begin{aligned} 3 - 2\lambda_1 x_1 - \lambda_2 &= 0 \\ 1 - 2\lambda_1 x_2 + \lambda_2 &= 0 \end{aligned} \right\}$$

Caso $(x_1, x_2) = (-1, -2)$

$$\rightarrow \left. \begin{array}{l} 3 + 2\lambda_1 - \lambda_2 = 0 \\ 1 + 4\lambda_1 + \lambda_2 = 0 \end{array} \right\} \rightarrow \text{la solución es } \lambda_1 = -\frac{2}{3}, \lambda_2 = \frac{5}{3}$$

contradiendo que estos valores tienen que ser positivos, por tanto, abandonamos este caso.

Caso $(x_1, x_2) = (2, 1)$

$\rightarrow \begin{cases} 3 - 4\lambda_1 - \lambda_2 = 0 \\ 1 - 2\lambda_1 + \lambda_2 = 0 \end{cases} \rightarrow$ la solución es $\lambda_1 = \frac{2}{3}, \lambda_2 = \frac{1}{3}$ siendo los dos positivos.

Como la matriz Jacobiana era funcional, no hemos podido saber su rango, sustituyendo la solución encontrada en ella y calculando su rango

rango de $\begin{pmatrix} 4 & 2 \\ 1 & -1 \end{pmatrix}$ es 2, que coincide con el número de restricciones saturadas, luego el punto es regular y como no hay ninguna contradicción con las inecuaciones, consideraremos que el candidato a máximo es

$$(x_1^o, x_2^o; \lambda_1^o, \lambda_2^o) = \left(2, 1; \frac{2}{3}, \frac{1}{3}\right).$$

Faltaría comprobar que efectivamente es el máximo.

Theorem

Dado el programa (1) y un punto $\vec{x}^o$ que es regular y que cumple la condición necesaria de optimalidad.

Si para todo vector $\vec{v} \in \mathbb{R}^N$ distinto del vector nulo tal que

$$\vec{v}^T \cdot \nabla g_j(\vec{x}^o) = 0 \text{ para todo } j \in \Gamma$$

y $\vec{v}^T H_{\vec{x}} L(\vec{x}^o; \vec{\lambda}^o) \vec{v}$ es definida negativa, entonces $\vec{x}^o$ es máximo local.

El conjunto Γ es de índices de las restricciones saturadas. Nótese que el vector v es ortogonal a todos los gradientes de las restricciones que son saturadas por $\vec{x}$ y la matriz hessiana está calculada únicamente para las variables $\vec{x}$.

En definitiva nos dice que la forma cuadrática (el hessiano) restringida a las direcciones factibles $\vec{v}$ tiene que ser definida negativa.

En el caso de que la forma cuadrática $H_{\vec{x}}L(\vec{x}^o; \vec{\lambda}^o)$ sea definida negativa, la forma cuadrática restringida a $\vec{v}$ también lo será, en consecuencia podemos decir directamente que el punto $\vec{x}^o$ es máximo local.

Tenemos que el Lagrangiano es

$$L(x_1, x_2; \lambda_1, \lambda_2) = 3x_1 + x_2 - \lambda_1 (x_1^2 + x_2^2 - 5) - \lambda_2 (x_1 - x_2 - 1).$$

El hessiano en el candidato $(x_1^o, x_2^o; \lambda_1^o, \lambda_2^o) = \left(2, 1; \frac{2}{3}, \frac{1}{3}\right)$ será:

$$H_{\vec{x}}L(\vec{x}; \vec{\lambda}) = \begin{pmatrix} -2\lambda_1 & 0 \\ 0 & -2\lambda_1 \end{pmatrix};$$

$$H_{\vec{x}}L(\vec{x}^o; \vec{\lambda}^o) = \begin{pmatrix} -\frac{4}{3} & 0 \\ 0 & -\frac{4}{3} \end{pmatrix}$$

que es definida negativa, en consecuencia el punto $\vec{x}^o$ es máximo local.

$$\left. \begin{array}{l} \text{Min } f(x_1, \dots, x_n) \\ g_1(x_1, \dots, x_n) \geq 0. \\ g_2(x_1, \dots, x_n) \geq 0. \\ \dots\dots\dots \\ g_m(x_1, \dots, x_n) \geq 0. \end{array} \right\} \quad \text{con} \quad \left\{ \begin{array}{l} f, g_i \in \mathcal{C}^2(D), \\ D \subset \mathbb{R}^n \text{ abierto.} \end{array} \right.$$

(2)

Condición necesaria de mínimo:

Theorem

Dado el programa (2) con $\vec{x}^o$ regular.

Si $\vec{x}^o$ es mínimo local $\Rightarrow$ debe haber multiplicadores

$\lambda_1^o, \dots, \lambda_m^o \in R$, tales que satisfagan

- a) $\nabla f(\vec{x}^o) - \lambda_1^o \nabla g_1(\vec{x}^o) - \dots - \lambda_m^o \nabla g_m(\vec{x}^o) = 0$,
- b) $\lambda_i^o g_i(\vec{x}^o) = 0$, $\forall i \in \{1, \dots, m\}$,
- c) $\lambda_i^o \geq 0$, $\forall i \in \{1, \dots, m\}$,
- d) $g_i(\vec{x}^o) \geq 0$, $\forall i \in \{1, \dots, m\}$.

Observamos que también **debemos** incluir en las restricciones las de no negatividad (i.e. $x_i \geq 0$).

Caso programa cóncavo

Theorem

Dado el programa (1). Si la función f es cóncava, K es un conjunto convexo y $\vec{x}^o \in K$ es regular. Son equivalentes:

1. $\vec{x}^o \in K$ es máximo local.
2. $\vec{x}^o \in K$ es máximo global.
3. $\exists \lambda_1^o, \dots, \lambda_m^o \in \mathbb{R}$, únicos y tales que

$$\left\{ \begin{array}{l} a) \quad \nabla f(\vec{x}^o) - \lambda_1^o \nabla g_1(\vec{x}^o) - \dots - \lambda_m^o \nabla g_m(\vec{x}^o) = 0. \\ b) \quad \lambda_i^o g_i(\vec{x}^o) = 0, \quad \forall i \in \{1, \dots, m\}. \\ c) \quad \lambda_i^o \geq 0, \quad \forall i \in \{1, \dots, m\}. \\ d) \quad g_i(\vec{x}^o) \leq 0, \quad \forall i \in \{1, \dots, m\}. \end{array} \right.$$

El teorema anterior nos asegura que para un programa cóncavo las condiciones necesarias de Kuhn–Tucker son suficientes para máximo y en un programa convexo, las condiciones necesarias de Kuhn–Tucker son suficientes para mínimo.

Caso programa convexo

Theorem

Dado el programa (2). Si la función f es convexa, K es un conjunto convexo y $\vec{x}^o \in K$ es regular. Son equivalentes:

- 1. $\vec{x}^o \in K$ es mínimo local.*
- 2. $\vec{x}^o \in K$ es mínimo global.*
- 3. $\exists \lambda_1^o, \dots, \lambda_m^o \in \mathbb{R}$, únicos y tales que*

$$\left\{ \begin{array}{l} a) \quad \nabla f(\vec{x}^o) - \lambda_1^o \nabla g_1(\vec{x}^o) - \dots - \lambda_m^o \nabla g_m(\vec{x}^o) = 0. \\ b) \quad \lambda_i^o g_i(\vec{x}^o) = 0, \forall i \in \{1, \dots, m\}. \\ c) \quad \lambda_i^o \geq 0, \forall i \in \{1, \dots, m\}. \\ d) \quad g_i(\vec{x}^o) \geq 0, \forall i \in \{1, \dots, m\}. \end{array} \right.$$