Programación Matemática

Abel Doñate Muñoz
abel.donate@estudiantat.upc.edu

Contents

1 Programación lineal

1.1 Planteamiento, definiciones y nomenclatura

Un problema de PL (Programación lineal) se puede formular en los siguientes términos:

$$PL \implies \begin{cases} \min_{x\in\mathbb{R}^n} z = c'x & \textbf{función objetivo} \\ Ax \leq b & \textbf{restricciones} \\ l \leq x \leq u & \textbf{cotas} \end{cases}$$

Región factible: P. Puntos que cumplen las restricciones.
Solución factible: $x \in P$.
Conjunto solución: X^* factibles que minimizan la función objetivo.

1.2 Forma estándar

Se trata de replantear las restricciones del problema de PM de manera que:

$$(PL) \begin{cases} \min c'x \\ Ax \leq b \\ l \leq x \leq u \end{cases} \implies (PL)_e \begin{cases} \min c'x \\ A_e x_e = b \\ x_e \geq 0 \end{cases}$$

1.3 Tipos de soluciones

SB (Solución básica)
Cumplen $x = \begin{pmatrix} X_B \\ X_N \end{pmatrix}$, $Ax = b$, pero no tienen porque cumplir $x_i \geq 0$
Se construyen a base de seleccionar $n - m$ variables como no básicas e igualarlas a 0, resolviendo el sistema restante (si la matriz reducida no es singular).

SBF (Solución básica factible)
SB con $x_i \geq 0$

SBD (Solución básica degenerada)
SB con $x_i = 0$ para alguna variable i básica

1.4 Teoremas

Teorema 1 *Existencia de puntos extremos*

$$\text{El poliedro } P \text{ tiene algún punto extremo} \iff P \text{ no contiene ninguna línea}$$

Teorema 2 *Optimalidad de puntos extremos*

Si P contiene algún punto extremo y existe una solución óptima a (PL) $\implies$ existe una solución óptima que es un extremo de P

Teorema 3 *Equivalencia puntos extremos (SBF)*

$$P \text{ en forma estándar} \implies \boxed{x^* \text{ es extremo} \iff x^* \text{ es SBF}}$$

2 Algoritmo del Símplex

El algoritmo consiste en ir saltando de SBF con la condición de que el coste se minimice. Si tenemos la SBF x de base B, para buscar una adyacente y seleccionamos una variable no básica q y hacemos:
En primer lugar miramos a que SBF saltar calculando los costes reducidos:

$$\boxed{r' = c_N' - c_B B^{-1} A_N} \qquad \text{si } r \geq 0 \implies x \text{ es óptimo}$$

Si no es óptimo sea q el primer índice que da negativo:

$$y = x + \theta^* d,\ d = \begin{pmatrix} d_B \\ d_N \end{pmatrix} \implies \boxed{d_B = -B^{-1}A_q},\ \boxed{d_N = \begin{pmatrix} 0 \\ . \\ 1 \\ . \\ 0 \end{pmatrix} \leftarrow q},\ \boxed{\theta^* = \min\left\{ \frac{-x_{B(i)}}{d_{B(i)}} : d_{B(i)} < 0 \right\}}$$

Para hacer los cálculos más eficientes podemos calcular B^{-1} basándonos en la matriz básica anterior:

$$\bar{B}^{-1} = HB^{-1} \implies H = \left[e_1, \cdots, e_{(p-1)}, \eta, \cdots, e_m\right],\ \boxed{\eta_i = \begin{cases} -d_{B(i)}/d_B(p) & \text{si } i \neq p \\ -1/d_B(p) & \text{si } i = p \end{cases}}$$

2.1 Iteración del Simplex

1. Empezamos con base $B = (A_{B(1)}, \ldots, A_{B(m)})$ y la SBF x
2. Calculamos costes reducidos r para todas las coordenadas no básicas.
   - Si $r \geq 0 \implies x$ es Solución. Termina.
   - Si no, elige el primer índice q con $r_q < 0$.
3. Calcula $d = -B^{-1}A_q$
   - Si $d \geq 0 \implies$ Problema ilimitado. Termina.
   - Si no, calcula θ y su índice p
4. Forma una nueva base reemplazando $A_{B(p)}$ por A_q. $y = x + \theta d$ y ve a **2**

2.2 Teoremas

Supondremos siempre P_e no vacío y de rango completo, x será SBF y $y = x + \theta^* d$

Teorema 4 *y SBF*

- Si $d_B \ngeq 0 \implies y$ es SBF
- Si $d_B \geq 0 \implies$ no existe θ tal que y sea SBF

Teorema 5 *Condiciones de optimalidad SBF*

- Si $r \geq [0] \implies x$ es óptima
- Si x es SBF óptima y no degenerada $\implies r \geq 0$

2.3 Fase I

Para encontrar la primera SBF hacemos símplex del siguiente problema de PL:

$$x = \begin{pmatrix} x_{[n]} \\ x_{[m]} \end{pmatrix}, \quad c_I' = (0, \ldots, 0|1, \ldots, 1) \implies (PL) \begin{cases} \min c_I' x \\ \begin{pmatrix} A|I \end{pmatrix} x = b \\ x \geq 0 \end{cases}$$

Este problema tiene una SBF trivial en $x[n] = 0, x_{[m]} = b$, por lo que podemos partir de ahí para realizar el símplex. Al ser factilbe tiene una solución óptima.

- Si $z^* > 0 \implies (PL)_e$ es infactible
- Si $z^* = 0 \implies (PL)_e$ es factible
  - Si $\mathcal{B}_I^* \subset \{1, \ldots, n\} \implies \mathcal{B}_I^*$ es SBF de $(PL)_e$
  - Si $\mathcal{B}_I^* \not\subset \{1, \ldots, n\} \implies \mathcal{B}_I^*$ es SBF degenerada de $(PL)_I$ a partir de la cual se puede obtener una SBF de $(PL)_e$

3 Simplex Dual

Podemos asociar a cada problema de símplex primal su problema dual:

$$(P)\begin{cases}\min c'x\\ Ax \ ? \ b\\ x \ ! \ 0\end{cases} \implies (D)\begin{cases}\max b'\lambda\\ A'\lambda \ !^T \ c\\ \lambda \ ? \ 0\end{cases}$$

3.1 Teoremas del Símplex Dual

Teorema 8 *Teorema débil de Dualidad*
Sea x SF del primal y λ SF del Dual

$$b'\lambda \leq c'x$$

Teorema 10 *Teorema fuerte de Dualidad*
Sean x^*, λ^* los óptimos del primal y dual, en caso de que existan:

$$b'\lambda^* = c'x^*$$

Teorema 11 *Teorema de Folga complementaria*
Sea x SF del primal y λ SF del Dual, son óptimas si y solo si

$$\begin{cases}\lambda_i(a_i'x - b_i) & = 0, \qquad \forall i\\ (c_j - \lambda' A_j)x_j & = 0 \qquad \forall j\end{cases}$$

3.2 Implicaciones del Dual

- (P) óptimo $\implies$ (D) óptimo
- (P) ilimitado $\implies$ (D) infactible
- (P) infactible $\implies$ (D) $\begin{cases}\text{infactible}\\ \text{ilimitado}\end{cases}$

3.3 Factibilidad Primal y Dual

Factibilidad Primal $\implies x_B = B^{-1}b \geq 0$
Factibilidad Dual $\implies r_N' = c_N' - c_B B^{-1} A_N = c_N' - \lambda' A_N \geq 0$

Si una solución es factible primal y dual, entonces es óptima.

4 Programación Lineal Entera (PLE)

Un problema de Programación Lineal Entera (PLE) es un problema de Programación Lineal donde algunas variables solo pueden tomar valores enteros.

Definimos la Relajación Lineal (RL) como el problema de PL asociado al PLE donde hacemos las variables enteras reales.

Tenemos las siguientes implicaciones:

- RL óptimo $\implies$ PLE $\begin{cases}\text{óptimo}\\ \text{infactible}\end{cases}$
- RL infactible $\implies$ PLE infactible
- RL ilimitado $\implies$ PLE $\begin{cases}\text{óptimo (no si } A \text{ tiene coeficientes racionales)}\\ \text{infactible}\\ \text{ilimitado}\end{cases}$

4.1 Branch and Bound

Definimos K_{PEi} como una partición disjunta de PE menores tal que la suma es PE. Trivialmente se cumple:

- $x^*_{PE} = x^*_{PE} \iff i = argminc'x^*_{PEi}$
- $K_{PE} = \emptyset \implies K_{PEi} = \emptyset$
- $z^*_{PE} \to -\infty \iff \exists i:\ z^*_{PEi} \to \infty$
- Podemos anidar las separaciones creando problemas PE más pequeños.

4.1.1 Algoritmo del Branch and bound

Sea un PE con $L =$ PE, $k = 1, z^* = \infty, z^*_{PE1} = -\infty$
Mientras que $L \neq \emptyset$:

- Seleccionamos un problema $PEj \in L$
- Resolvemos $RLj \implies x^*_{RLj}, z^*_{PEj} \leftarrow z^*_{RLj}$
- Si (PLj es ilimitado) ó ($K_{RLj} = \emptyset$) ó ($z^*_{PEj} \geq z^*$) ó ($x^*_{RLj} \in K_{PEj}$) :
  - $L \leftarrow L - \{PEj\}$
  - Si $x^*_{RLj} \in K_{PEj}$:
    * Si $z^*_{RLj} \leq z^* \implies x^* \leftarrow x^*_{RLj}, z^* \leftarrow z^*_{RLj}$
    * Si $z^* \leq z^*_{PEi}$ para algún PEi **cosas a completar**
- Si no:
  - Separamos $PEj = PE(k+1) \cup PE(k+2)$
  -

5 Optimización no lineal

5.1 Formulación general

$$(PO)\begin{cases}\min f(x) \\ \text{s.a. } h(x) = 0 \\ \qquad g(x) \leq 0\end{cases} \quad , \qquad h(x) = \begin{pmatrix} h_1(x) \\ \vdots \\ h_m(x)\end{pmatrix}, \quad g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_p(x)\end{pmatrix}$$

5.2 Teoremas e implicaciones

Teorema de Taylor

$$f(x + \alpha d) = f(x) + \alpha \nabla f(x)^T d + \frac{1}{2}\alpha^2 d^T \nabla^2 f(x) d + O(\alpha^2)$$

Teorema de Weierstrass
Si $f : \Omega \to \mathbb{R}$ continua y Ω compacto $\implies$ tiene un mínimo

Conjuntos y funciones convexas
Ω es convexo $\iff \alpha x + (1-\alpha)y \in \Omega, \quad \alpha \in [0,1] \quad \forall x, y \in \Omega$
f es convexa en $\Omega \iff f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y), \quad \alpha \in [0,1] \quad \forall x, y \in \Omega$

Condición necesaria y suficiente de convexidad con ∇f
f es convexa en Ω convexo si solo si

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \quad \forall x, y \in \Omega$$

Condición necesaria y suficiente de convexidad con $\nabla^2 f$
f es convexa en Ω convexo si solo si $\nabla^2 f$ semidefinida positiva en Ω ($\nabla^2 f \succ 0$)

Condición necesaria conjuntos convexos
Si D convexo, $g_i : D \to \mathbb{R}$ convexa $\implies$ $S_c = \{x : x \in D, g_i(x) \leq c_i\}$ es convexo.

Problemas de optimización convexos
Si Ω convexo y f convexa $\implies$ el problema de optimización es convexo
En un problema convexo un mínimo local es también global

Condición suficiente para ser mínimo
Si :

- $\nabla f(x^*) = 0$
- $\nabla^2 f(x^*) \succ 0$

entonces x^* es un mínimo local de f en un entorno abierto de x^*

5.3 Convergencia de el algoritmo

Para buscar la α se puede hacer una optimización en una variable restringiendo en la dirección de descenso. Si buscamos el global se dice que es exacta, si encontramos un local, inexacta.

$$g(\alpha) = f(x^k + \alpha d), \qquad \alpha = \arg\min g(\alpha)$$

5.3.1 Tipos de convergencia

- Convergencia lineal (Método del Gradiente en funciones cuadráticas):

$$\exists r \in (0,1) : \frac{||x^{k+1} - x^*||}{||x^k - x^*||} \leq r \quad \forall k > N$$

- Convergencia supralineal ():

$$\lim_{k \to \infty} \frac{||x^{k+1} - x^*||}{||x^k - x^*||} = 0$$

- Convergencia cuadrática ():

$$\exists M : \frac{||x^{k+1} - x^*||}{||x^k - x^*||^2} \leq M \quad \forall k > N$$

5.3.2 Condiciones de Armijo-Wolfe

Condición 1. Decrecimiento suficiente.

$$g(\alpha) \leq f(0) + \alpha c_1 g'(0) \iff f(x^k + \alpha d) \leq f(x^k) + \alpha c_1 \nabla f(x^k) d$$

El valor de $c_1 \in (0,1)$ suele ser 10^{-4}

Condición 2. Condición de curvatura.

$$g'(\alpha) \geq c_2 g'(0) \iff \nabla f(x^k + \alpha d)^T d \geq c_2 \nabla f(x^k)^T d$$

El valor de $c_2 \in (c_1, 1)$ suele ser 0.9

5.3.3 Teorema de Zoutendijk

podemos calcular el ángulo θ_k de la dirección de descenso con respecto a la dirección de máximo descenso como:

$$\cos(\theta_k) = \frac{-\nabla f(x^k)^T d^k}{||\nabla f(x^k)||||d^k||}$$

El teorema asegura que si en el proceso iterativo $x^{k+1} = x^k + \alpha^k d^k$:

- d^k es de descenso
- Se cumple Armijo-Wolfe
- f fitada inferiormente
- $f \in \mathcal{C}^1$ en un abierto N que contiene el conjunto de nivel $L = \{x : f(x) < f(x^0)\}$
- ∇f Lipschitz continuo en $N \iff ||\nabla f(x) - \nabla f(y)|| \leq L||x - y||$

Entonces podemos asegurar $\sum_0^\infty \cos^2 \theta_k ||\nabla f(x^k)||^2 < \infty$

5.4 Tipos de métodos

Existen principalmente dos tipos de métodos de optimización:

- *Métodos de exploración lineal.* Primero calculamos d y luego α obteniendo $x^{k+1} = x^k + \alpha d$
- *Métodos de región de garantía.* Primero fijamos Δ (radio de entorno) y luego calculamos d en el entorno obteniendo $x^{k+1} = x^k + d$.

Utilizaremos el primero en esta asignatura

5.4.1 Método del gradiente

La dirección de descenso es la de máximo descenso $d = -\nabla f(x^k)$.

Este algoritmo asegura la convergencia por la propiedad:

$$\frac{f(x^{k+1}) - f(x^*)}{f(x^k) - f(x^*)} = \frac{||x^{k+1} - x^*||_Q^2}{||x^k - x^*||_Q^2} \leq r^2, \qquad r \in \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}, 1\right), \quad \lambda_i \text{ VAPs de } \nabla^2 f(x^*)$$

Si la función es cuadrática $f(x) = \frac{1}{2}$:

- $\nabla f(x) = Qx - b$
- $\alpha = \dfrac{\nabla f(x)^T \nabla f(x)}{\nabla f(x)^T Q \nabla f(x)}$
- $r = \dfrac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}$

5.4.2 Método de Newton

Se obtiene a partir de la aproximación $f(x^k + d) = f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T \nabla^2 f(x^k) d + R_2(||d||)$

$$m_k(x) = f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T \nabla^2 f(x^k) d$$

Si calculamos el mínimo de esta función tenemos

$$d = -(\nabla^2 f(x^k))^{-1} \nabla f(x^k)$$

Se demuestra que a medida que nos acercamos a x^* se verifica $\alpha^k = 1$ verifica Armijo-Wolfe. Ademas si consideramos:

- $f \in \mathcal{C}^2$ y $\nabla^2 f$ Lipschitz continua en N entorn de x^*
- Se cumplen las conduciones suficientes de optimalidad $\nabla f(x^*) = 0, \nabla^2 f(x^*) \succ 0$
- x^0 suficientemente cercano a x^*

Entonces $\{x^k\}$ converge cuadráticamente a x^* y $\{||\nabla f(x^k)||\}$ converge cuadráticamente a 0

5.5 Modificaciones del método de Newton

- **Descomposición espectral**
  Si $\nabla^2 f(x^k) = VDV^T$ (diagonaliza por el teorema espectral), hacemos:
  $$B_k = V(D+\Delta)V^T = \nabla^2 f(x^k) + E_k$$
- **Adición de una matriz diagonal**
  Buscamos un τ tal que:
  $$B_k = \nabla^2 f(x^k) + \tau I \succ 0$$
- **Factorización de Cholesky**

6 Optimización no lineal con restricciones

6.1 Definiciones

En un punto x la restricción $g_i(x) \leq 0$ puede ser:

- **Activa** si $g_i(x) = 0$
- **Inactiva** si $g_i(x) < 0$

Las restricciones $h_i(x), i = 1, \ldots, m$ son siempre activas.
Denotamos $\mathcal{A}(x) = \{j \in \{1, \ldots, m\} : g_j(x) = 0\}\}$ al conjunto de índices cuya restricción $g_i(x)$ es activa en x.

Denotamos $c_i(x) = 0$ el conjunto de las restricciones activas en un punto dado. Esto es, el conjunto de las h_i y el de las g_i activas.

Decimos que un punto x^* es regular si los gradientes ∇c_i con linealmente independientes.

Definimos el plano tangente $T = \{$derivadas de la curva diferencial que pasa por $x^*\}$

Definimos el subespacio de direcciones $M = \{d : \nabla c(x^*)^T d = 0\}$

Definimos el subespacio de direcciones $M' = \{d : \nabla h(x^*)^T d = 0, \quad \nabla g_j(x^*)^T d = 0 \quad \forall f \in \mathcal{A}(x^*) \cap \{j : \mu_j^* > 0\}\}$

6.2 Multiplicadores de Lagrange

Definimos la función Lagraniana:

$$L(x, \lambda) = f(x) + \lambda^T + h(x)\mu^T g(x)$$

Condiciones de optimalidad:

- Necesaria de 1º orden: $\nabla_x L(x^*, \lambda^*) = 0$
- Suficiente de 2º orden: $d^T \nabla^2_{xx} L(x^*, \lambda^*) d > 0$ para todo $d : \nabla h(x^*)^T d = 0$

6.3 KKT

Teorema previo. x^* es regular en la superficie de S definida por $c(x) = 0 \implies T = M$

Condiciones KKT:

- **Necesarias:**
  - Primer orden: $\begin{cases} h(x^*) = 0, g(x^*) \leq 0 \\ \nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \nabla h(x^*)\lambda^* + \nabla g(x^*)\mu^* = 0 \\ \mu^* \geq 0, \mu^{*T} g(x^*) = 0, \quad (g_i \text{ inactiva } \implies \mu_i^* = 0) \end{cases}$

- Segundo orden: $d^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) d \geq 0 \quad \forall d \in M$

- **Suficientes:**
  - Segundo orden: $d^T \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) d > 0 \quad \forall d \in M'$

6.4 Análisis de sensibilidad

Planteamos el problema de programación no lineal:

$$\begin{aligned} &\min f(x) \\ &\text{s.a. } h(x) = \varepsilon \quad (h_i(x) = \varepsilon_i) \\ &\qquad g(x) \leq \delta \quad (g_j(x) \leq \delta_j) \end{aligned}$$

Si la solución para $\varepsilon = \delta = 0$ es x^* regular y λ^*, μ^* satisfacen KKT, entonces $\forall (\varepsilon, \delta) \in \mathcal{E}_{(0,0)}$ hay una solución $x^* \varepsilon, \delta)$ tal que:

$$\begin{aligned} x^*(0,0) &= x^* \\ \nabla_\epsilon f(x(\epsilon, \delta))|_{(0,0)} &= -\lambda^* \\ \nabla_\delta f(x(\epsilon, \delta))|_{(0,0)} &= -\mu^* \end{aligned}$$