Métodos Numéricos de Optimización con restricciones.

Bustos Jordi
Práctica I

September 9, 2025

Theorem 1. Dada la función

$$f(x,y) = (x - y^2)\left(x - \tfrac{1}{2}y^2\right).$$

Mostrar que $(0,0)$ es el único punto estacionario que NO es minimizador, sin embargo

$$\forall d \in \mathbb{R}^n : \quad f(x^*) \leq f(x^* + td) \quad \text{para todo } t \text{ pequeño.}$$

Proof. Es claro que f es una función polinómica y por lo tanto es infinitamente diferenciable en todo $\mathbb{R}^2$ con:

$$\begin{aligned}\nabla f(x,y) &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \\ &= \left(2x - \tfrac{3}{2}y^2, -3xy + 2y^3\right).\end{aligned}$$

Luego,

$$\nabla f(x,y) = 0 \Leftrightarrow \begin{cases} 2x - \frac{3}{2}y^2 = 0 \\ -3xy + 2y^3 = 0 \end{cases}$$

Resolviendo el sistema se obtiene que el único punto estacionario es $(0,0)$. Además,

$$\begin{aligned}H &= \frac{\partial^2 f(0,0)}{\partial x^2} \cdot \frac{\partial^2 f(0,0)}{\partial y^2} - \left(\frac{\partial^2 f(0,0)}{\partial (x,y)}\right)^2 \\ & 2 \cdot 0 - 0 = 0\end{aligned}$$

Por lo tanto, el criterio de la matriz hessiana no nos permite concluir si $(0,0)$ es un mínimo local, máximo local o punto silla. Si analizamos la función por el camino $x = y^2$ se obtiene que $f(y^2, y) = \frac{1}{2} \cdot (y^4 - 3y^3)$ que tiene un máximo local en $y = 0$. Por otro lado, el camino $y = 0$ da $f(x,0) = x^2$ que tiene un mínimo local en $x = 0$. Por lo tanto, $(0,0)$ no puede ser minimizador local.
Sea $d = (d_1, d_2) \in \mathbb{R}^2$ y $t > 0$ suficientemente chico. Entonces

$$\begin{aligned}f(0,0) &= 0 \\ f(t \cdot d_1, t \cdot d_2) &= t^2(d_1^2 - \frac{3}{2}d_1 d_2^2 t + \frac{1}{2}t^2 d_2^4)\end{aligned}$$

Como los términos entre paréntesis son positivos para t suficientemente chico, se concluye que $f(0,0) \leq f(t \cdot d_1, t \cdot d_2)$. □

Theorem 2. Sea

$$f(x,y) = \left(y - x^2\right)^2 + x^5.$$

Mostrar que f tiene un único punto estacionario que no es maximizador ni minimizador de f.

Proof. Análogamente al punto anterior calculemos el gradiente de f:

$$\begin{aligned}\nabla f(x,y) &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \\ &= \left(-4x(y-x^2)+5x^4, 2(y-x^2)\right).\end{aligned}$$

Luego,

$$\nabla f(x,y) = 0 \Leftrightarrow \begin{cases} -4x(y-x^2)+5x^4 = 0 \\ 2(y-x^2) = 0 \end{cases}$$

Resolviendo el sistema se obtiene que el único punto estacionario es $(0,0)$. Además, $H = 0$ y por lo tanto el criterio de la matriz hessiana no nos permite concluir si $(0,0)$ es un mínimo local, máximo local o punto silla. Si analizamos la función por el camino $y = x^2$ se obtiene que $f(x,x^2) = x^5$ que toma valores tanto positivos como negativos en cualquier vecindad de $x = 0$. Por lo tanto $(0,0)$ no puede ser ni minimizador ni maximizador local. □

Theorem 3. Si x^* es un minimizador local de (PC) entonces x^* es un minimizador global de f.

Proof. Supongamos que x^* es minimizador local, pero no global. Entonces podemos encontrar un punto $z \in \mathbb{R}^n$ con $f(z) < f(x^*)$. Consideremos el segmento de línea que une x^* a z, es decir,

$$x = \lambda z + (1-\lambda)x^*, \quad \text{para algún } \lambda \in (0,1].$$

Por la propiedad de convexidad de f, tenemos

$$f(x) \leq \lambda f(z) + (1-\lambda)f(x^*) < f(x^*).$$

Cualquier vecindad $\mathcal{N}$ de x^* contiene un trozo del segmento de línea, por lo que siempre habrá puntos $x \in \mathcal{N}$ en los que se satisfaga la segunda ecuación. Por lo tanto, x^* no es un minimizador local. □

Theorem 4. Sea $f : \mathbb{R}^n \to \mathbb{R}$ diferenciable en Ω convexo. Probar que f es convexa si y sólo si

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

(b) $\nabla f(x)^T (y-x) \leq \nabla f(y)^T (y-x)$ $\quad\forall x, y \in \Omega.$

Proof. Veamos (a) primero. Sea f convexa, para $x, y \in \Omega$ y $t \in (0,1)$ fijo, definido $d = y - x$ tenemos $x + td \in \Omega$ y por la convexidad de f

$$f(x+td) = f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$$

Por lo tanto,

$$\begin{aligned}f(y) - f(x) &\geq \lim_{t \to 0^+} \frac{f(x+td) - f(x)}{t} \\ &= \nabla f(x)^T d \\ &= \nabla f(x)^T (y-x)\end{aligned}$$

Para la vuelta, consideremos $z = (1-t)x + ty$ y notemos que

$$f(x) \geq f(z) + \nabla f(z)^T (x-z) \text{ y } f(y) \geq f(z) + \nabla f(z)^T (y-z)$$

Multiplicando la primera por $(1-t)$ y la segunda por t obtenemos

$$(1-t)f(x) + tf(y) \geq f\left((1-t)x + ty\right)$$

Completando la demostración de (a). Para (b), consideremos primero que f es convexa y utilizando (a) obtenemos

$$\begin{aligned} f(y) &\geq f(x) + (y-x)^T \nabla f(x) \\ f(x) &\geq f(y) + (x-y)^T \nabla f(y) \end{aligned}$$

Por lo tanto,

$$\begin{aligned} f(y) + f(x) &\geq f(x) + f(y) + (y-x)^T \nabla f(x) + (x-y)^T \nabla f(y) \\ 0 &\geq (y-x)^T \nabla f(x) + (x-y)^T \nabla f(y) \\ (y-x)^T \nabla f(y) &\geq (y-x)^T \nabla f(x) \end{aligned}$$

Recíprocamente, si f cumple (b), dados $x, y \in \Omega$ consideremos $\xi = (1-t)x + ty$ con $t \in (0,1)$. Entonces,

$$f(y) - f(x) = \nabla f(\xi)^T (y-x) \quad \text{por el teorema del valor medio.}$$

Si aplicamos la hipótesis a x, ξ se tiene que:

$$\begin{aligned} \nabla f(x)^T (\xi - x) &\leq \nabla f(\xi)^T (\xi - x) \\ \nabla f(x)^T (y - x) &\leq \nabla f(\xi)^T (y - x) \quad \text{ya que } \xi - x = t(y-x) \end{aligned}$$

Combinando con la igualdad del teorema del valor medio obtenemos

$$\begin{aligned} f(y) - f(x) = \nabla f(\xi)^T (y-x) &\geq \nabla f(x)^T (y-x) \\ f(y) &\geq f(x) + \nabla f(x)^T (y-x) \end{aligned}$$

Luego, vale (a) y por lo tanto f es convexa. □

Theorem 5. Sea $D \neq \emptyset$ un conjunto cerrado y convexo. Sea $x \in \mathbb{R}^n$. Entonces:

1. Para todo $y \in D$ se tiene que
$$(x - P_D(x))^T (y - P_D(x)) \leq 0.$$

2. Si $\bar{z} \in D$ cumple que
$$(x - \bar{z})^T (y - \bar{z}) \leq 0, \quad \forall y \in D$$
entonces $\bar{z} = P_D(x)$.

3. La función $P : \mathbb{R}^n \to D$ definida como $P(x) = P_D(x)$ cumple la siguiente propiedad:
$$\|P(x) - P(y)\| \leq \|x - y\|, \quad \forall x, y \in \mathbb{R}^n$$
lo que implica que la proyección P es una función continua.

Proof. Consideremos un punto arbitrario $y \in D$. Dado $t \in (0,1)$, por la convexidad de D, tenemos que $(1-t)P_D(x) + ty \in D$. Entonces,

$$\|x - P_D(x)\| \leq \|x - (1-t)P_D(x) - ty\| = \|x - P_D(x) + t(P_D(x) - y)\|$$

Luego,

$$\begin{aligned}\|x - P_D(x)\|^2 &\leq \|x - P_D(x) + t(P_D(x) - y)\|^2 \\ &= \|x - P_D(x)\|^2 + 2t(x - P_D(x))^T(P_D(x) - y) + t^2\|P_D(x) - y\|^2\end{aligned}$$

Como $t > 0$, tenemos que $2(x - P_D(x))^T(y - P_D(x)) \leq t\|P_D(x) - y\|^2$. Haciendo $t \to 0$ obtenemos

$$(x - P_D(x))^T(y - P_D(x)) \leq 0$$

que es lo que queríamos demostrar.
Supongamos que existe $\bar{z} \in D$ que cumple la desigualdad anterior. Entonces, dado $y \in D$

$$\begin{aligned}\|x - \bar{z}\|^2 - \|x - y\|^2 &= -2x^T\bar{z} + \bar{z}^T\bar{z} + 2x^Ty - y^Ty \\ &= (y - \bar{z})^T(2x - y - \bar{z}) \\ &= (y - \bar{z})^T(2(x - \bar{z}) - (y - \bar{z})) \leq 0\end{aligned}$$

usando la igualdad algebraica $\|a\|^2 - \|b\|^2 = (a - b)^T(a + b)$. Esto prueba que $\bar{z} = P_D(x)$.
Finalmente, para probar la tercera parte, consideremos $p := P(x)$ y $q := P(y)$. Aplicando la parte (1) convenientemente obtenemos

$$\begin{aligned}&(x - p)^T(q - p) \leq 0 \\ &(y - q)^T(p - q) \leq 0 \\ &\implies (x - p)^T(q - p) + (y - q)^T(p - q) \leq 0\end{aligned}$$

Además $(y - q)^T(p - q) = -(y - q)^T(q - p)$. Reordenando los términos obtenemos

$$\begin{aligned}(x - p)^T(q - p) - (y - q)^T(q - p) &\leq 0 \\ (x - y - (p - q))^T(q - p) &\leq 0 \\ (x - y - (p - q))^T(p - q) &\geq 0 \\ (x - y)^T(p - q) &\geq \|p - q\|^2\end{aligned}$$

Recordemos que $(p - q)^T(p-q) = \|p-q\|^2$. Además, por la desigualdad de Cauchy-Schwarz, $(x - y)^T(p-q) \leq \|x - y\| \cdot \|p - q\|$. Por lo tanto,

$$\|p - q\|^2 \leq \|x - y\| \cdot \|p - q\| \implies \|p - q\| \leq \|x - y\|$$

que es lo que queriamos probar. □