De la solución de un sistema de ecuaciones lineales a la minimización de una forma cuadrática con términos lineales

Objetivos. Mostrar que el problema $Ax = b$, donde A es una matriz real simétrica estrictamente positiva definida, es equivalente al problema de minimización de una función (de una forma cuadrática con términos lineales).

1. Lema (propiedad simétrica de la forma bilineal asociada a una matriz simétrica). Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz real simétrica, y sean $x, y \in \mathbb{R}^n$. Entonces

$$x^\top Ay = y^\top Ax.$$

Demostración. El producto $x^\top Ay$ es un número, o sea una matriz 1×1, la cual coincide con su transpuesta:

$$x^\top Ay = \left(x^\top Ay\right)^\top = y^\top Ax. \qquad \square$$

2. Teorema (de un sistema de ecuaciones lineales a un problema de optimización). Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz real simétrica, y sea $b \in \mathbb{R}^n$. Definimos $f\colon \mathbb{R}^n \to \mathbb{R}$ mediante la fórmula

$$f(x) = \frac{1}{2}x^\top Ax - x^\top b. \tag{1}$$

Denotemos por u al vector $A^{-1}b$. Entonces para cada $x \in \mathbb{R}^n$

$$f(x) = f(u) + \frac{1}{2}(x-u)^\top A(x-u).$$

En particular, si A es estrictamente positiva definida, entonces $f(x) > f(u)$ para cada $x \in \mathbb{R}^n \setminus \{0\}$.

Demostración.

$$f(x) - f(u) = \frac{1}{2}x^\top Ax - \frac{1}{2}u^\top Au - (x-u)^\top b$$

Usamos la igualdad $b = Au$:

$$= \frac{1}{2}x^\top Ax - \frac{1}{2}u^\top Au - (x^\top - u^\top)Au = \frac{1}{2}x^\top Ax - x^\top Au + \frac{1}{2}u^\top Au$$

Como A es simétrica, $x^\top Au = u^\top Ax$ por el Lema 1

$$= \frac{1}{2}\left(x^\top Ax - x^\top Au - u^\top Ax + u^\top Au\right) = \frac{1}{2}(x-u)^\top A(x-u). \qquad \square$$

3. El gradiente y la matriz hessiana de la función f. Ya sabemos calcular el gradiente y la matriz hessiana de las formas lineales y cuadráticas. Para la función (1) obtenemos

$$(\operatorname{grad} f)(x) = Ax - b, \qquad H_f(x) = A.$$

4. Idea de demostración analítica. Como $H_f(x) = A > 0$ para cada $x \in \mathbb{R}^n$, la función f es estrictamente convexa. Además su gradiente se anula solamente en el punto $u = A^{-1}b$. Por lo tanto el punto u es un punto mínimo global estricto de la función f.

5. Lineas de nivel de la función f, para $n = 2$. Se $n = 2$, entonces la función f se puede escribir como

$$f(x_1, x_2) = \frac{1}{2}A_{1,1}x_1^2 + A_{1,2}x_1x_2 + \frac{1}{2}A_{2,2}x_2^2 - b_1x_1 - b_2x_2.$$

La condición que $A > 0$ es equivalente a las condiciones

$$A_{1,1} > 0, \qquad A_{1,1}A_{2,2} - A_{1,2}^2 > 0.$$

Para cada $c \in \mathbb{R}$ el conjunto

$$\{x \in \mathbb{R}^2 : \ f(x) = c\}$$

es una elipse (o es vacío). El punto $u = A^{-1}b$ es el centro de esta elipse.

6. El caso multidimensional. Para n general, los conjuntos

$$\{x \in \mathbb{R}^2 : \ f(x) = c\}$$

son ciertas superficies en $\mathbb{R}^n$, llamadas *hyperelipses* o *elipsoides n-dimensionales*.