Método del gradiente para resolver sistemas de ecuaciones lineales

Objetivos. Deducir las fórmulas del “método de gradiente”. Más precisamente, es el método del descenso en el sentido del antigradiente.

1. 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

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

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.

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

Usamos la igualdad $b = Ax$:

$$\begin{aligned}
&= \frac{1}{2} x^\top A x - \frac{1}{2} u^\top A u - (x-u)^\top A u \\
&= \frac{1}{2}(x-u)^\top A x + \frac{1}{2} u^\top A x - \frac{1}{2} u^\top A u - (x-u)^\top A u \\
&= \frac{1}{2}(x-u)^\top A x + \frac{1}{2} u^\top A(x-u) - (x-u)^\top A u
\end{aligned}$$

Aplicamos la simetría de A:

$$\begin{aligned}
&= \frac{1}{2}(x-u)^\top A x - \frac{1}{2}(x-u)^\top A u \\
&= \frac{1}{2}(x-u)^\top A(x-u).
\end{aligned}$$

$\square$

2. Gradiente de la función f. Ya sabemos calcular el gradiente de las formas lineales y cuadráticas. En nuestro caso,

$$(\operatorname{grad} f)(x) = Ax - b.$$

3. Derivada direccional de la función $\boldsymbol{f}$. Sean x y p dos vectores fijos. Definimos $g\colon \mathbb{R} \to \mathbb{R}$ mediante la regla

$$g(\alpha) = f(x + \alpha p).$$

Entonces

$$g'(\alpha) = \alpha p^\top A p - p^\top (b - Ax). \tag{1}$$

Demostración. Método I. Usar la fórmula para la derivada direccional:

$$g'(\alpha) = p^\top \, (\operatorname{grad} f)(x + \alpha p).$$

Método II. Primero calcular la función g:

$$\begin{aligned} g(\alpha) &= \frac{1}{2}(x + \alpha p)^\top A (x + \alpha p) - (x + \alpha p)^\top b \\ &= \frac{1}{2}\alpha^2 p^\top A p + \alpha \, \big(pAx^\top - p^\top b\big) + \frac{1}{2} x^\top A x - x^\top b. \end{aligned}$$

Luego sacar la derivada de g:

$$g'(\alpha) = \alpha p^\top A p - p^\top (b - Ax).$$ □

4. Ideas del método del gradiente.

- Construir una sucesión de puntos $x^{(0)}, x^{(1)}, x^{(2)}, \ldots$, en cada paso intentando disminuir el valor de $f(x)$.

  $$x^{(s+1)} \coloneqq x^{(s)} + \alpha_s p^{(s)},$$

  donde el vector $p^{(s)}$ y el número $\alpha_s > 0$ se eligen de cierta manera.
- Poner $p^{(s)}$ igual al antigradiente de la función f en el punto $x^{(s)}$.
- Eligir α_s minimizando la función f sobre la recta $x^{(s)} + \alpha p^{(s)}$.

5. Elegir la dirección.

$$p^{(s)} \coloneqq -(\operatorname{grad} f)(x^{(s)}) = b - Ax^{(s)}.$$

La expresión $b - Ax^{(s)}$ es el *residuo* $r^{(s)}$ del problema $Ax = b$ en el punto $x^{(s)}$.

6. Elegir el paso. Definimos la función $g\colon \mathbb{R} \to \mathbb{R}$,

$$g(\alpha) \coloneqq f(x^{(s)} + \alpha r^{(s)}).$$

Ya hemos deducido una fórmula (1) para la derivada de g. La función g alcanza su mínimo en el punto

$$\alpha_s = \frac{(r^{(s)})^\top r^{(s)}}{(r^{(s)})^\top A r^{(s)}}.$$

7. El cambio del residuo en un paso.

$$r^{(s+1)} - r^{(s)} = Ax^{(s)} - Ax^{(s+1)} = -\alpha_s A r^{(s)}.$$

8. Algoritmo del antigradiente.

Entrada: A, b, ε, $smax$.

Salida: x, s.

$s \leftarrow 0;$

$x \leftarrow 0;$

$r \leftarrow b;$

Mientras $\quad (\|r\| \geq \varepsilon) \quad \wedge \quad (s < smax)$:

$\quad p \leftarrow Ar;$

$\quad \alpha \leftarrow \dfrac{r^\top r}{r^\top A r};$

$\quad x \leftarrow x + \alpha r;$

$\quad r \leftarrow r - \alpha A r;$

$\quad s \leftarrow s + 1;$

Regresar x y s.