Método de direcciones conjugadas

Estos apuntes están escritos con ayuda de Lourdes Fabiola Uribe Richaud y Juan Esaú Trejo Espino.

Objetivos. Estudiar la idea del método de direcciones conjugadas y demostrar que este método converge en n pasos.

1. De un sistema de ecuaciones lineales a un problema de minimización (repaso). Ya sabemos que si A es una matriz simétrica positiva definida, entonces el problema de resolver el sistema $Ax = b$ es equivalente al problema de minizar la función

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

2. Lema (sobre la minimización sobre una recta, repaso). Sea A, b, f los mismo objetos que en el párrafo anterior. Si $y \in \mathbb{R}^n$ y $p \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$, entonces la función $g \colon \mathbb{R} \to \mathbb{R}$ definida mediante la regla

$$g(\alpha) \coloneqq f(y + \alpha p)$$

alcanza su valor mínimo en el punto

$$\alpha = \frac{p^\top (b - Ay)}{p^\top A p}. \tag{2}$$

3. Lema (sobre la ortogonalidad del residuo nuevo a la dirección anterior, repaso). Sean A, b, y, p, f, g como en el lema anterior, sea α el número definido mediante la fórmula (2), y sea

$$z = y + \alpha p.$$

Entonces

$$p \perp (b - Az).$$

Primera demostración. Cálculo directo:

$$\begin{aligned} p^\top (b - Az) &= p^\top (b - A(y + \alpha p)) = p^\top (b - Ay) - \alpha p^\top A p \\ &= p^\top (b - Ay) - \frac{p^\top (b - Ay)}{p^\top A p} p^\top A p = 0. \end{aligned}$$ □

Segunda demostración. Recordamos que

$$g'(\alpha) = p^\top (\nabla f)(y + \alpha p) = p^\top \big(A(y + \alpha p) - b \big).$$

Si α es el punto mínimo de la función g y $z = y + \alpha p$, entonces $g'(\alpha) = 0$ y

$$p^\top (Az - b) = 0,$$

esto es, $p \perp (b - Az)$. □

4. El producto interno asociado a una matriz. Estamos suponiendo que $A \in \mathcal{M}_n(\mathbb{R})$ es una matriz simétrica positiva definida. Entonces la función de dos argumentos

$$\langle x, y \rangle_A \coloneqq x^\top A y$$

es un producto interno en $\mathbb{R}^n$.

5. Vectores *A*-conjugados. Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz simétrica positiva definida y sean $p^{(0)}, \ldots, p^{(m)} \in \mathbb{R}^n$. Se dice que la lista de vectores $p^{(0)}, \ldots, p^{(m)}$ es *A-conjugada* o *A-ortogonal* si

$$\forall j, k \in \{0, \ldots, m\} \qquad (j \neq k) \quad \Longrightarrow \quad \langle p_j, p_k \rangle_A = 0.$$

6. Independencia lineal de vectores *A*-conjugados. Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz simétrica positiva definida y sea $p^{(0)}, \ldots, p^{(m)} \in \mathbb{R}^n$ una lista A-conjugada de vectores no nulos. Entonces la lista $p^{(0)}, \ldots, p^{(m)}$ es linealmente independiente.

Demostración. Es un hecho general sobre vectores no nulos y ortogonales con respecto a un producto interno. Escribamos la demostración para nuestra situación particular, cuando el producto interno es inducido por la matriz A. Supongamos que $\lambda_0, \ldots, \lambda_m \in \mathbb{R}$ tales que

$$\sum_{j=0}^{m} \lambda_j p^{(j)} = \mathbf{0}_n.$$

Mostremos que para cada $k \in \{0, \ldots, m\}$ el escalar λ_k es cero. Multiplicamos ambos lados de la igualdad por $(p^{(k)})^\top A$, del lado izquierdo, y utilizamos propiedades del producto de matrices:

$$\sum_{0=1}^{m} \lambda_j (p^{(k)})^\top A p^{(j)} = 0.$$

Si $j \neq k$, entonces $(p^{(k)})^\top A p^{(j)} = 0$. Para $j = k$ obtenemos $(p^{(k)})^\top A p^{(k)} > 0$. Por eso la igualdad se convierte en

$$\lambda_k (p^{(k)})^\top A p^{(k)} = 0$$

e implica que $\lambda_k = 0$. $\square$

7. Teorema sobre la convergencia del método de direcciones conjugadas. Sean $A \in \mathcal{M}_n(\mathbb{R})$ una matriz real simétrica positiva definida, $b \in \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$. Además sean $p^{(k)}, \ldots, p^{(n-1)}$ algunos vectores no nulos A-conjugados. Definimos escalares $\alpha_0, \ldots, \alpha_{n-1} \in \mathbb{R}$ y vectores $x^{(1)}, \ldots, x^{(n)} \in \mathbb{R}^n$ mediante las siguientes fórmulas recursivas:

$$\alpha_s \coloneqq \frac{(p^{(s)})^\top (b - Ax^{(s)})}{(p^{(s)})^\top A p^{(s)}}, \qquad x^{(s+1)} \coloneqq x^{(s)} + \alpha_s p^{(s)} \qquad (s = 0, \ldots, n-1).$$

Entonces $Ax^{(n)} = b$.

Demostración. Denotemos por u a la solución exacta del sistema: $u \coloneqq A^{-1}b$. Nuestro objetivo es demostrar que $x^{(n)} = u$. Los vectores $p^{(0)}, \ldots, p^{(n-1)}$ forman una base de $\mathbb{R}^n$, por eso el vector $u - x^{(0)}$ se puede expandir en una combinación lineal de la forma

$$u - x^{(0)} = \sum_{j=0}^{n-1} \lambda_j p^{(j)}, \tag{3}$$

donde $\lambda_0, \ldots, \lambda_{n-1}$ son algunos números reales. Por otro lado, la definición de los vectores $x^{(1)}, \ldots, x^{(n)}$ implica que

$$x^{(s)} - x^{(0)} = \sum_{j=0}^{s-1} \alpha_j p^{(j)}. \tag{4}$$

Vamos a demostrar que $\lambda_j = \alpha_j$ para cada j; esto implicará que $x^{(n)} - x^{(0)} = u - x^{(0)}$ y $x^{(n)} = u$.

Sea $s \in \{0, \ldots, n-1\}$. Multipliquemos la igualdad (3) por $(p^{(s)})^\top A$ del lado izquierdo y apliquemos la hipótesis que $p^{(0)}, \ldots, p^{(n-1)}$ son A-conjugados:

$$(p^{(s)})^\top A(u - x^{(0)}) = \lambda_s (p^{(s)})^\top A p^{(s)}.$$

Luego notamos que $A(u - x^{(0)}) = Au - Ax^{(0)} = b - Ax^{(0)}$ y obtenemos

$$\lambda_s = \frac{(p^{(s)})^\top (b - Ax^{(0)})}{(p^{(s)})^\top A p^{(s)}}.$$

Por otro lado, simplifiquemos le numerador del cociente que aparece en la definición del escalar α_s:

$$(p^{(s)})^\top A x^{(s)} = (p^{(s)})^\top A x^{(0)} + \sum_{j=1}^{s-1} \alpha_j (p^{(s)})^\top A p^{(j)} = (p^{(s)})^\top A x^{(0)}.$$

Hemos demostrado que $\lambda_s = \alpha_s$. $\square$

8. Teorema sobre la ortogonalidad de los residuos y direcciones en el método de direcciones conjugadas. Sean $A \in \mathcal{M}_n(\mathbb{R})$ una matriz real simétrica positiva definida, $b \in \mathbb{R}^n$, $x^{(0)} \in \mathbb{R}^n$. Sean $p^{(0)}, \ldots, p^{(n-1)}$ algunos vectores no nulos A-ortogonales. Definimos escalares $\alpha_0, \ldots, \alpha_{n-1} \in \mathbb{R}$ y vectores $x^{(1)}, \ldots, x^{(n)} \in \mathbb{R}^n$ mediante las siguientes fórmulas recursivas:

$$\alpha_s := \frac{(p^{(s)})^\top (b - Ax^{(s)})}{(p^{(s)})^\top A p^{(s)}}, \qquad x^{(s+1)} := x^{(s)} + \alpha_s p^{(s)} \qquad (s = 0, \ldots, n-1).$$

Entonces $(p^{(j)})^\top (b - Ax^{(k)}) = 0$ para cualesquiera $j, k \in \{0, \ldots, n\}$ con $j < k$.

Demostración. Denotemos $b - Ax^{(k)}$ por $r^{(k)}$. Entonces

$$r^{(s)} = r^{(s-1)} - \alpha_{s-1} A p^{(s-1)}.$$

De la definición de los escalares α_s se obtiene inmediatamente que $p^{(s-1)} \perp r^{(s)}$ (véase el Lema 3). Demostremos por inducción sobre s que para cada $s \in \{1, \ldots, n\}$ y cada $j \in \{0, \ldots, s-1\}$ se tiene $(p^{(j)})^\top r^{(s)} = 0$.

Base de inducción: $s = 1$. En este caso el único valor de j es 0, así que $j = s - 1$, y la afirmación se cumple por el Lema 3.

Supongamos que $j \leq s - 2$. Entonces

$$(p^{(j)})^\top r^{(s)} = (p^{(j)})^\top (r^{(s-1)} - \alpha_{s-1} A p^{(s-1)}) = (p^{(j)})^\top r^{(s-1)} - \alpha_{s-1} (p^{(j)})^\top A p^{(s-1)}.$$

El primer sumando es 0 por la hipótesis de inducción, y el segundo sumando es 0 por la hipótesis que las direcciones son A-conjugadas. Concluimos que $p^{(j)} \perp r^{(s)}$. Hemos demostrado el paso de inducción. $\square$