Métodos Numéricos II

Segundo de Grado en Matemáticas, Curso 2018/19

Facultad de Ciencias de la Universidad de Málaga
Dpto. de Análisis Matemático, Estadística e Investigación Operativa, y Matemática Aplicada
Profs. Francisco José Palma Molina (Área de Conocimiento de Matemática Aplicada)

Métodos de tipo descenso.

Se considera el sistema lineal $A\,X = B$, donde los datos del problema son la matriz $A \in \mathcal{M}_n(\mathbb{R})$ que suponemos (simétrica y) definida positiva y $B \in \mathbb{R}^n$. Evidentemente al ser A inversible este sistema tiene solución única, que denotamos por $\bar{X} \in \mathbb{R}^n$.

Ejercicio. Definida la función

$$\begin{array}{rcl} J\colon \mathbb{R}^n & \to & \mathbb{R} \\ X & \mapsto & J(X) = X^t\,A\,X - 2\,X^t\,B\,, \end{array}$$

demostrar que posee un único mínimo global, que es $\bar{X}$.

Resolución. Observamos que

$$\begin{array}{rcl} J(X) & = & \langle A\,X, X\rangle - 2\,\langle B, X\rangle \\ & = & \displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}\,x_i\,x_j - 2\,\sum_{i=i}^{n} x_i\,b_i\,, \end{array}$$

con lo que

$$\frac{\partial\, J(X)}{\partial\, x_k} = 2\,\left(\sum_{j=1}^{n} a_{k,j}\,x_j - b_k\right)\,,$$

y por tanto

$$\nabla\, J(X) = 2\,(A\,X - B)\,.$$

Por otro lado

$$\frac{\partial^2\, J(X)}{\partial\, x_l\,\partial\, x_k} = 2\,a_{k,l}\,,$$

de donde se obtiene que

$$H\,J(X) = 2\,A\,.$$

Dado que la matriz A es definida positiva, se deduce que la función J es estrictamente convexa en todo $\mathbb{R}^n$, por lo que su extremo, de existir, es mínimo global y estricto y queda caracterizado por $\nabla\, J(X) = 0$. Habida cuenta la caracterización de $\nabla\, J(X)$, se llega a que existe un único mínimo global que es justamente $\bar{X}$, es decir, única solución del sistema lineal propuesto. ■

Para cada $X \in \mathbb{R}^n$, consideramos ahora el vector (función) error absoluto

$$E(X) = E_X = X - \bar{X}\,,$$

y el vector (función) residuo

$$R(X) = R_X = B - A\,X\,.$$

Ejercicio. Demostrar que $\nabla\, J(X) = -2\,A\,E_X$. ■

Ejercicio. Definida la función

$$\begin{array}{rcl} F\colon \mathbb{R}^n & \to & \mathbb{R} \\ X & \mapsto & F(X) = {E_X}^t\,A\,E_X\,, \end{array}$$

demostrar que posee un único mínimo global, que es $\bar{X}$.

Resolución. Basta observar que

$$F(X) = J(X) + \bar{X}^t\,A\,\bar{X}\,,$$

es decir, las funciones J y F se diferencian en una constante, por lo que poseen los mísmos extremos (aunque evidentemente estos extremos alcancen valores diferentes). Un cálculo inmediato permite comprobar que $F(\bar{X}) = 0$. ■

Los métodos iterativos de tipo descenso para resolver el sistema lineal $A\,X = B$, tratan de construir una sucesión $\{X_k\}_{k\in\mathbb{N}}$ convergente hacia $\bar{X}$, bajo el criterio de que en cada iteración $F(X_{k+1}) < F(X_k)$, y siendo

$$X_{k+1} = X_k + \alpha_k\,P_k\,,$$

donde el vector P_k recibe el nombre de dirección de descenso de la iteración, y α_k paso en la dirección de descenso.

Ejercicio. Demostrar que en un método iterativo de tipo descenso, fijada la dirección de descenso P_k, el paso óptimo α_k, es decir, aquel tal que

$$F(X_k + \alpha_k\,P_k) = \min_{\alpha\in\mathbb{R}}\, F(X_k + \alpha\,P_k)$$

se obtiene cuando

$$\alpha_k = \frac{{P_k}^t\,R_k}{{P_k}^t\,A\,P_k} = \frac{\langle R_k, P_k\rangle}{\langle A\,P_k, P_k\rangle}\,.$$

(Mediante $\langle\cdot,\cdot\rangle$ notamos el producto escalar ordinario de $\mathbb{R}^n$.)

Resolución. En primer lugar observamos que para aligerar la escritura hemos notado $R_k = R_{X_k} = B - A\,X_k$; análogamente pondremos $E_k = E_{X_k} = X_k - \bar{X}$. Para obtener el resultado buscado, basta observar que

$$F(X_k + \alpha\,P_k) = \alpha^2\,{P_k}^t\,A\,P_k - 2\,\alpha\,{P_k}^t\,R_k + {E_k}^t\,A\,E_k\,.$$

La conclusión es inmediata dado el carácter cuadrático de esta función. ■

Ejercicio. Demostrar que en el caso de elección de paso óptimo, se verifica que

$$F(X_{k+1}) = F(X_k)\,(1 - \gamma_k)\,,$$

con $0 \le \gamma_k \le 1$.

Resolución. Un cálculo inmediato permite obtener que

$$\gamma_k = \frac{\langle R_k, P_k\rangle^2}{({R_k}^t\,A^{-1}\,R_k)\,({P_k}^t\,A\,P_k)}\,,$$

y dado el carácter definido positivo de A (y de A^{-1}), se llega a la conclusión buscada. ■

Ejercicio. Demostrar que si

$$0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$$

son los autovalores de la matriz A, entonces para todo $X \in \mathbb{R}^n$ se verifica que

$$\lambda_1\,X^t\,X \le X^t\,A\,X \le \lambda_n\,X^t\,X\,.$$

■

Ejercicio. Demostrar que en el caso de elección de paso óptimo, se verifica que

$$\gamma_k \geq \frac{1}{\mathrm{cond}(A)} \left\langle \frac{1}{\|R_k\|} R_k, \frac{1}{\|P_k\|} P_k \right\rangle^2 .$$

(Todas las normas, y por tanto el condicionamiento, que aparecen hacen referencia a la norma vectorial euclídea o norma-2 y a la norma matricial subordinada correspondiente, norma espectral o norma-2.)

Resolución. De acuerdo con el resultado anterior, se tiene que

$$P_k{}^t A P_k \leq \lambda_n \|P_k\|^2 , \quad \text{y} \quad R_k{}^t A^{-1} R_k \leq \lambda_1^{-1} \|R_k\|^2 ,$$

de donde el resultado pedido. ■

Ejercicio. Demostrar que en los métodos de tipo descenso, toda elección de la dirección de descenso P_k que verifique que

$$\left\langle \frac{1}{\|R_k\|} R_k, \frac{1}{\|P_k\|} P_k \right\rangle^2 \geq \mu > 0$$

con μ independiente de k, implica la convergencia del método iterativo.

Resolución. Se prueba que en este caso

$$F(X_k) \leq \left(1 - \frac{\mu}{\mathrm{cond}(A)}\right)^k F(X_0) .$$

Como $0 < \mu \leq 1$, de donde

$$0 \leq 1 - \frac{\mu}{\mathrm{cond}(A)} < 1 ,$$

se llega a que $\{F(X_k)\}_{k\in\mathbb{R}}$ converge hacia 0. ■

Se deduce entonces que las direcciones de descenso P_k no deben ser ortogonales a R_k y que una buena elección es tomar

$$P_k = R_k = -\frac{1}{2} \nabla J(X_k)$$

pues en ese caso $\mu = 1$. Esto es conocido como método del gradiente.

Finalmente vamos a considerar un método de descenso de tipo gradiente, es decir, en el que $P_k = R_k$, pero con paso fijo, es decir, $\alpha_k = \alpha$ fijo.

Ejercicio. Demostrar que en los método de descenso de tipo gradiente con paso fijo α, la convergencia está asegurada siempre que $\alpha \in (0, \frac{2}{\varrho(A)})$.

Resolución. Basta comprobar que

$$E_k = (I - \alpha A)^k E_0 ;$$

si imponemos ahora que el radio espectral de la matriz $I - \alpha A$ sea menor que 1 se obtiene lo pedido. ■