Método de descenso más rápido y del gradiente conjugado para la resolución de sistemas lineales

Elementos de Cálculo Numérico

Segundo Cuatrimestre de 2007

Dado el sistema $Ax = b$ con A matriz simétrica y definida positiva, sabemos que (ver Kincaid, *Análisis Numérico*, pag. 210 en adelante) resolver este sistema es equivalente a minimizar la forma cuadrática

$$q(x) = \langle x, Ax \rangle - 2\langle x, b \rangle .$$

El método del descenso más rápido consiste en dar una sucesió $(x_k)_k$ tal que por todo k, $q(x_{k+1}) \leq q(x_k)$ con igualdad unicamente en algún caso especial. La idea genérica es obtener x_{k+1} a partir de x_k y de otro vector v_k que puede ya estar dado o que se lo vaya construyendo durante la evolución del método. La idea para construir x_{k+1} es la siguiente:
Siendo en x_k y moviendonos en la dirección v_k, minimizamos q restringida a la recta $t \to x_k + tv_k$, es decir minimizamos la función $t \to q(x_k + tv_k)$. Esta función alcanza su mínimo en

$$t_k = \frac{\langle v_k, b - Ax_k \rangle}{\langle v_k, Av_k \rangle}$$

(verifiquelo ! Porque podemos dividir por $\langle v_k, Av_k \rangle$?), y tenemos

$$q(x_k + t_k v_k) = q(x_k) - \frac{\langle v_k, b - Ax_k \rangle^2}{\langle v_k, Av_k \rangle} \leq q(x_k).$$

Definimos ahora x_{k+1} por

$$x_{k+1} = x_k + t_k v_k.$$

La elección de la sucesión (v_k) es la que da origen a los distintos métodos.

Descenso más rápido El método del descenso más rápido se define tomando $v_k = -\nabla q(x_k)$ (o un múltiplo positivo de él). La practicidad del método reside en que el residuo $r_k = b - Ax_k$ apunta en la dirección de $-\nabla q(x_k)$ (comprobarlo). Por lo tanto, podemos tomar $v_k = \frac{r_k}{\|r_k\|}$.

Dirección conjugada La idea es considerar una base $\{u_1, \ldots, u_n\}$ ortogonal en el sentido que $\langle u_i, u_j \rangle_A := \langle u_i, Au_j \rangle = 0$ si $i \neq j$, y luego definir la sucesión $(x_k)_{0 \leq k \leq n}$ por

$$x_k = x_{k-1} + \frac{\langle b - Ax_{k-1}, u_k \rangle}{\langle u_k, Au_k \rangle} u_k \qquad 1 \leq k \leq n$$

arrancando desde un punto x_0 cualquiera. Una gran ventaja de este método es que, en teoría, en n pasos alcanza la solución del sistema, pero en la práctica es muy sensible a los errores.
Un método para encontrar los u_i es aplicar el algoritmo de Gram-Schmidt a la base standard $\{e_1, \ldots, e_n\}$ de $\mathbb{R}^n$ por el producto escalar $\langle ., . \rangle_A$: definimos $u_1 = e_1$ y luego por recurencia

$$u_j = e_j - \sum_{i<j} \langle e_j, u_i \rangle_A u_i$$

por $j = 2 \ldots n$.

Gradiente conjugado Se basa en la idea anterior pero no da a priori una base $\{u_1, \ldots, u_n\}$ ortogonal según el sentido anterior, sino que va construyéndola con la propiedad adicional que la sucesión de residuos $r_k = b - Ax_k$ forman un sistema ortogonal en el sentido usual: $\langle r_i, r_j \rangle = 0$ si $i \neq j$.
El algoritmo es el siguiente: arrancando de un x_0 cualquiera, se define $r_0 = b - Ax_0$ y $v_0 = r_0$. En el paso k-ésimo, ya tenemos definido x_k, r_k y v_k. Para pasar al $k+1$-ésimo se define, siguiendo el orden,

$$\begin{aligned}
t_k &= \frac{\langle r_k, r_k \rangle}{\langle r_k, Av_k \rangle} \\
x_{k+1} &= x_k + t_k v_k \\
r_{k+1} &= r_k - t_k A v_k \\
s_k &= \frac{\langle r_{k+1}, r_{k+1} \rangle}{\langle r_k, r_k \rangle} \\
v_{k+1} &= r_{k+1} + s_k v_k.
\end{aligned}$$

Paramos la iteración cuando r_k es sufiicientemente pequeño en la norma usual.

Ejemplos Resuelva cada uno de los siguientes sistemas por los distintos métodos que se mencionaron, y también por los métodos de Jacobi y Gauss-Seidel. Compara la velocidad de convergencia de todos ellos.

1. $A = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 4 \\ 4 \\ 2 \end{pmatrix}$

2. $Ax = b$ donde A es la matriz de Hilbert de $n \times n$ dada por $a_{ij} = (i + j + 1)^{-1}$ y $b_i = \frac{1}{3} \sum_{j=1}^{n} a_{ij}$ para $1 \leq i \leq n$