Resolução de sistemas de equações lineares: Método do Gradiente

Marina Andretta

ICMC-USP

24 de março de 2015

Método do Gradiente

Estamos interessados em resolver o sistema linear

$$Ax = b,$$

com $A \in \boldsymbol{R}^{n \times n}$ e $b \in \boldsymbol{R}^n$ dados.

Quando a matriz A é esparsa, podemos não tê-la representada, mas dispor apenas de rotinas que calculem o produto de A por um dado vetor v. Neste caso, não é possível usar os métodos diretos ou os Métodos Jacobi-Richardson ou Gauss-Seidel.

Método do Gradiente

Para isso, usaremos o Método do Gradiente. A ideia básica deste método é trocar o problema de encontrar a solução do sistema $Ax = b$ pelo problema de encontrar o minimizador da função quadrática

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

Funções quadráticas

Para ver que estes problemas podem ser equivalentes, vejamos quais as condições para que um ponto x seja minimizador de f.

Funções quadráticas

Do cálculo temos que se um ponto x é minimizador de uma função g,

- gradiente de $g(x) = \nabla g(x) = 0$ e
- Hessiana de $g(x) = \nabla^2 g(x)$ semidefinida positiva.

Mais ainda, se $\nabla^2 g(x)$ for definida positiva, x é um minimizador local de g.

Funções quadráticas

Dada a função $f(x) = \frac{1}{2}x^T Ax - b^T x$, temos que

$$\nabla f(x) = Ax - b,$$

$$\nabla^2 f(x) = A.$$

Note que o minimizador x^* de f deve satisfazer

$$\nabla f(x^*) = Ax^* - b = 0 \Rightarrow Ax^* = b.$$

Mas só temos garantia de que o ponto x^* é minimizador de f se A for simétrica e definida positiva.

Assim, se A for simétrica e definida positiva, podemos aplicar o Método do Gradiente e encontrar o minimizador da função f, que será solução do sistema $Ax = b$.

Método do Gradiente

Para minimizar a função quadrática f, partimos de um ponto inicial x_0 e, a cada iteração k, escolhemos uma direção d_k que faz com que o valor da função diminua de valor a partir de x_k.

Escolhida a direção d_k, escolhemos o tamanho do passo α_k que daremos na direção d_k, a partir de x_k.

Assim, o ponto x_{k+1} é dado por

$$x_{k+1} = x_k + \alpha_k d_k.$$

Método do Gradiente

Como, a partir de um ponto x_k, a função f aumenta na direção de $\nabla f(x_k)$, definiremos a direção d_k como

$$d_k = -\nabla f(x_k).$$

Definiremos o resíduo do sistema linear no ponto x_k como

$$r_k = Ax_k - b = \nabla f(x_k) \Rightarrow d_k = -r_k.$$

Método do Gradiente

Para escolher o tamanho de passo α_k, vamos calcular o valor que minimiza a função f na direção $d_k = -r_k$. Ou seja, vamos calcular α_k tal que

$$\phi(\alpha) = f(x_k + \alpha d_k)$$

seja mínima.

Método do Gradiente

Note que

$$\phi(\alpha) = f(x_k + \alpha d_k) = \frac{1}{2}(x_k + \alpha d_k)^T A(x_k + \alpha d_k) - b^T(x_k + \alpha d_k),$$

com x_k e d_k conhecidos.

Como $\phi(\alpha)$ é uma parábola convexa, para encontrar o valor de α que minimiza ϕ basta calcular a derivada de ϕ em relação a α e igualá-la a 0.

Método do Gradiente

Com isso, temos que

$$\alpha_k = \frac{d_k^T d_k}{d_k^T A d_k} = \frac{r_k^T r_k}{r_k^T A r_k}.$$

Algoritmo

Método do Gradiente: dadas a dimensão n, uma matriz $A \in \mathbb{R}^{n \times n}$ simétrica definida positiva, um vetor $b \in \mathbb{R}^n$, uma aproximação inicial x_0, uma tolerância $TOL > 0$ e o número máximo de iterações $MAXIT$, devolve x_k uma aproximação da solução de $Ax = b$ ou emite uma mensagem de erro.

Passo 1: Faça $k \leftarrow 0$.

Passo 2: Faça $r_k \leftarrow Ax_k - b$.

Passo 3: Se $\|r_k\| < TOL$, então

devolva x_k como solução e pare.

Passo 4: Faça $\alpha_k \leftarrow \frac{r_k^T r_k}{r_k^T A r_k}$.

Passo 5: Faça $x_{k+1} \leftarrow x_k - \alpha_k r_k$.

Algoritmo - continuação

Passo 6: Se $\|x_{k+1} - x_k\| < TOL$ ou $\frac{\|x_{k+1} - x_k\|}{\|x_{k+1}\|} < TOL$, então

devolva x_{k+1} como solução e pare.

Passo 7: Se $k > MAXIT$ então

Escreva "o método falhou após $MAXIT$ iterações" e pare.

Passo 8: Faça $k \leftarrow k + 1$ e volte para o Passo 2.

Exemplo

Vamos usar o Método dos Gradientes Conjugados para encontrar a solução do sistema linear

$$\begin{pmatrix} 4 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$$

com precisão 10^{-1}, usando o ponto inicial $x_0 = (0,0)^T$.

Exemplo

Na primeira iteração ($k = 0$), temos que

$$r_0 = Ax_0 - b = -b = \begin{pmatrix} -5 \\ -4 \end{pmatrix}.$$

Como $\|r_0\|_\infty = 5 > 10^{-1}$, calculamos

$$\alpha_0 = \frac{r_0^T r_0}{r_0^T A r_0} = \frac{41}{188} \approx 0.2181,$$

$$x_1 = x_0 - \alpha_0 r_0 = \begin{pmatrix} 1.0905 \\ 0.8724 \end{pmatrix}.$$

Exemplo

Calculando os erros absoluto e relativo, temos que

$$\|x_1 - x_0\|_\infty = 1.0905 > 10^{-1},$$

$$\frac{\|x_1 - x_0\|_\infty}{\|x_1\|_\infty} = 1 > 10^{-1}.$$

Exemplo

Na segunda iteração ($k = 1$), temos que

$$r_1 = Ax_1 - b = \begin{pmatrix} 5.2344 \\ 3.7077 \end{pmatrix} - \begin{pmatrix} 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 0.2344 \\ -0.2923 \end{pmatrix}.$$

Como $\|r_1\|_\infty = 0.2923 > 10^{-1}$, calculamos

$$\alpha_1 = \frac{r_1^T r_1}{r_1^T A r_1} \approx \frac{0.1404}{0.3390} \approx 0.4142,$$

$$x_2 = x_1 - \alpha_1 r_1 \approx \begin{pmatrix} 0.9934 \\ 0.9935 \end{pmatrix}.$$

Exemplo

Calculando os erros absoluto e relativo, temos

$$\|x_2 - x_1\|_\infty \approx 0.1211 > 10^{-1},$$

$$\frac{\|x_2 - x_1\|_\infty}{\|x_2\|_\infty} \approx 0.1219 > 10^{-1}.$$

Exemplo

Na terceira iteração ($k = 2$), temos que

$$r_2 = Ax_2 - b = \begin{pmatrix} 4.9671 \\ 3.9739 \end{pmatrix} - \begin{pmatrix} 5 \\ 4 \end{pmatrix} = \begin{pmatrix} -0.0329 \\ -0.0261 \end{pmatrix}.$$

Como $\|r_2\|_\infty = 0.0329 < 10^{-1}$, paramos com $x_2 = (0.9934, 0.9935)^T$ como solução aproximada do sistema.