Cálculo Numérico / Métodos Numéricos

Sistemas lineares
Método dos Gradientes

14 Nov 2008 . 16:52

Método dos Gradientes

- Os métodos do tipo gradiente para resolver o sistema $Ax = b$ têm como idéia básica minimizar a função de x seguinte:

$$F(x) = \tfrac{1}{2}\, x^T A x - b^T x$$

$$\min F(x) = \tfrac{1}{2}\, x^T A x - b^T x = \tfrac{1}{2}\,(Ax, x) - (b, x).$$

- A matriz A deve ser simétrica ($A^T = A$) e definida positiva ($x^T A x > 0$, para $x \neq 0$).

Método dos Gradientes – exemplo

- Considere n = 2

$$F(x)=\frac{1}{2}\begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}-\begin{pmatrix}b_1 & b_2\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}$$

$$F(x)=\frac{1}{2}\begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}a_{11}x_1+a_{12}x_2\\ a_{21}x_1+a_{22}x_2\end{pmatrix}-b_1x_1-b_2x_2$$

$$F(x)=\frac{1}{2}\left(a_{11}x_1^2+a_{12}x_2x_1+a_{21}x_1x_2+a_{22}x_2^2\right)-b_1x_1-b_2x_2$$

- Como A é simétrica temos que:

$$F(x)=\frac{1}{2}a_{11}x_1^2+a_{12}x_2x_1+\frac{1}{2}a_{22}x_2^2-b_1x_1-b_2x_2$$

Método dos Gradientes – revisão cálculo

- Do cálculo sabemos que um ponto P = (x_1, x_2) tal que o grad F(P) = 0 é chamado de ***ponto estacionário*** de F(x).

$$grad(x) = \begin{pmatrix} \dfrac{\partial F}{\partial x_1} \\ \dfrac{\partial F}{\partial x_2} \end{pmatrix}$$

- Seja A a matriz dada por:

$$A(x) = \begin{pmatrix} \dfrac{\partial F}{\partial x_1 \partial x_1} & \dfrac{\partial F}{\partial x_1 \partial x_2} \\ \dfrac{\partial F}{\partial x_2 \partial x_1} & \dfrac{\partial F}{\partial x_2 \partial x_2} \end{pmatrix}$$

Assim, se:

- A) A(P) : definida positiva, então P é **ponto de mínimo**;
- B) A(P) : definida negativa, então P é ponto de máximo;
- C) A(P) : indefinida, então P é ponto de cela.

Método dos Gradientes

- Por hipótese, para o método dos Gradientes, A deve ser simétrica e definida positiva, logo, um ponto estacionário P, ou seja, um ponto em que grad(P) = 0, é um **ponto de mínimo**.

- Voltando ao exemplo temos:

$$F(x) = \frac{1}{2}\mathrm{a}_{11}x_1^2 + \mathrm{a}_{12}x_2x_1 + \frac{1}{2}\mathrm{a}_{22}x_2^2 - \mathrm{b}_1x_1 - \mathrm{b}_2x_2$$

Sistema linear.

$$grad\ F(x) = \begin{pmatrix} \mathrm{a}_{11}x_1 + \mathrm{a}_{12}x_2 - \mathrm{b}_1 \\ \mathrm{a}_{12}x_1 + \mathrm{a}_{22}x_2 - \mathrm{b}_2 \end{pmatrix} = 0 \Leftrightarrow \boxed{\begin{array}{l} \mathrm{a}_{11}x_1 + \mathrm{a}_{12}x_2 = \mathrm{b}_1 \\ \mathrm{a}_{12}x_1 + \mathrm{a}_{22}x_2 = \mathrm{b}_2 \end{array}}$$

Método dos Gradientes

- Sendo assim, grad F(x) = 0, significa $Ax = b$, isto é, a solução do sistema de equações lineares minimiza a função quadrática e vice-versa.

- Exemplo: Seja o sistema linear $Ax = b$, dado por:

$$\begin{array}{lll} 100x_1 & +1x_2 & =1 \\ 1x_1 & +100x_2 & =100 \end{array}$$

- Calcule a função quadrática dada por $F(x) = \frac{1}{2} x^T A x - b^T x$ e mostre que o ponto de mínimo desta função é a solução do sistema dado.

Método dos Gradientes

$$F(x) = \frac{1}{2}\begin{pmatrix} x_1 & x_2 \end{pmatrix}\begin{pmatrix} 100 & 1 \\ 1 & 100 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 1 & 100 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} =$$

$$F(x) = \frac{1}{2}\begin{pmatrix} x_1 & x_2 \end{pmatrix}\begin{pmatrix} 100x_1 + x_2 \\ x_1 + 100x_2 \end{pmatrix} - x_1 - 100x_2 =$$

$$F(x) = \frac{1}{2}(100x_1^2 + x_1x_2 + x_1x_2 + 100x_2^2) - x_1 - 100x_2 =$$

$$F(x) = 50x_1^2 + x_1x_2 + 50x_2^2 - x_1 - 100x_2$$

Método dos Gradientes

$$F(x) = 50x_1^2 + x_1x_2 + 50x_2^2 - x_1 - 100x_2$$

$$grad\ \mathrm{F(x)} = \begin{pmatrix} 100x_1 + x_2 - 1 \\ x_1 + 100x_2 - 100 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \boxed{\begin{array}{ll} 100x_1 + x_2 & = 1 \\ x_1 + 100x_2 & = 100 \end{array}}$$

Sistema original

Solução: $x_1 = 0 \quad e \quad x_2 = 1$

$$F(x) = 0 + 0 + 50 - 0 - 100 = -50$$

Método dos Gradientes

- A matriz Hessiana para o exemplo é dada por

$$\begin{pmatrix} \dfrac{\partial F}{\partial x_1 \partial x_1} & \dfrac{\partial F}{\partial x_1 \partial x_2} \\ \dfrac{\partial F}{\partial x_2 \partial x_1} & \dfrac{\partial F}{\partial x_2 \partial x_2} \end{pmatrix} = \begin{pmatrix} 100 & 1 \\ 1 & 100 \end{pmatrix}$$

- Como essa matriz é simétrica e definida positiva, então o ponto x = (0 1) é um ponto de mínimo da função F(x).

Método dos Gradientes

- Note que:

$$grad\ F(x) = Ax - b$$

- Para o valor que minimiza a função F(x) temos que:

$$grad\ F(x) = Ax - b = 0$$

- No entanto, em outros pontos de F(x):

$$grad\ F(x) = Ax - b \neq 0$$

- Ou ainda, tomando menos o valor do gradiente, temos a diferença entre b e Ax. (Sabemos o quão distante a solução x está de resolver o sistema Ax=b).

- Em outras palavras, o resíduo da solução x é dado por:

$$-grad\ F(x) = b - Ax = r$$

Direção

- O vetor gradiente aponta a direção (sentido) de crescimento máximo de uma função. Portanto, é natural que, nos passos de busca do mínimo de uma função caminhemos na direção contrária ao gradiente, isto é, para um dado ponto $x^{(k)}$ temos:

$$x^{(k+1)} = x^{(k)} - s_k grad\ F(x^{(k)})$$

- Em que a direção é dada por

$$p = -grad\ F(x^{(k)}) = -(Ax^{(k)} - b) = b - Ax^{(k)} = r^{(k)}$$

Passo

- Direção: $p = -grad\ F(x^{(k)}) = -(Ax^{(k)} - b) = b - Ax^{(k)} = r^{(k)}$

$$x^{(k+1)} = x^{(k)} - s_k grad\ F(x^{(k)})$$

passo: o quanto vamos caminhar neste direção.

Passo

- Dado um $x^{(k)}$, queremos o melhor $x^{(k+1)}$ possível. Ele será obtido andando-se a quantia certa na direção oposta ao gradiente:

$$\min F(x^{(k+1)}) = \min F(x^{(k)} + s_k r^{(k)})$$

- Quantia certa:
  s_k que minimiza : $F(x^{(k)} + s_k r^{(k)})$

Método dos Gradientes

Queremos minimizar:

$$\min F(x^{(k+1)}) = \min F(x^{(k)} + s_k r^{(k)})$$

Simplificando a notação, temos que:

$$F(x+sr) = \frac{1}{2}(x+sr)^T A\ (x+sr) - b^T (x+sr)$$

Derivando em relação a **s** temos:

$$F'(x+sr) = \frac{1}{2} r^T A(x+sr) + \frac{1}{2}(x+sr)^T Ar - b^T r$$

Método dos Gradientes

$$F'(x+sr) = \frac{1}{2} r^T A(x+sr) + \frac{1}{2}(x+sr)^T Ar - b^T r$$

$$F'(x+sr) = sr^T Ar + r^T (Ax - b) = sr^T Ar - r^T r$$

Para encontrar s tomamos a derivada igual a zero:

$$F'(x+sr) = sr^T Ar - r^T r = 0$$

$$s = \frac{r^T r}{r^T Ar}$$

Método dos Gradientes

Tomando a segunda derivada de F(x+sr):

$$F''(x+sr) = r^T Ar$$

Uma vez que a matriz A é definida positiva sabemos que:

$$F''(x+sr) = r^T Ar > 0$$

Logo, o ponto obtido anteriormente é um ponto de mínimo.

Método dos Gradientes - Algoritmo

- Dados A, b, max e Erro
- 1) $x^{(0)}=0$
- 2) $k = 0$
- 3) $r = b - Ax^{(k)}$
- 4) $s = r^T r / r^T A r$
- 5) $x^{(k+1)} = x^{(k)} + s\, r$
- 6) Se $||x_{k+1}-x_k||_\infty / ||x_{k+1}||_\infty <$ Erro então faça solução $= x^{(k+1)}$ e PARE
- 7) $k = k+1$
- 8) Se ($k<max$) então volte ao Passo 3.
- 9) Senão escreva que a solução $x^{(k+1)}$ e o erro. PARE.

Método dos Gradientes - Exemplo

- Usando o método dos gradientes resolva o sistema dado por:

$$\begin{pmatrix} 10 & 1 & 0 \\ 1 & 10 & 1 \\ 0 & 1 & 10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 11 \\ 11 \\ 1 \end{pmatrix}$$

- Com precisão 10^{-1} e max = 3 iterações.

Método dos Gradientes - Exemplo

- Solução

$$x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad k = 0 \qquad r^{(0)} = b - Ax^{(0)} = \begin{pmatrix} 11 \\ 11 \\ 1 \end{pmatrix}$$ (direção)

$$s = \frac{r^{(0)T} r^{(0)}}{r^{(0)T} A r^{(0)}} = \frac{243}{2694} = 0{,}0902$$ (passo)

$$x^{(1)} = x^{(0)} + s r^{(0)} = \begin{pmatrix} 0{,}9922 \\ 0{,}9922 \\ 0{,}0902 \end{pmatrix}$$

Método dos Gradientes - Exemplo

$$x^{(1)} = \begin{pmatrix} 0{,}9922 \\ 0{,}9922 \\ 0{,}0902 \end{pmatrix} \qquad k = 1 \qquad r^{(1)} = b - Ax^{(1)} = \begin{pmatrix} 0{,}0858 \\ -0{,}0044 \\ -0{,}8542 \end{pmatrix}$$

$$s = \frac{r^{(1)T} r^{(1)}}{r^{(1)T} A r^{(1)}} = \frac{0{,}8070}{8{,}0769} = 0{,}0999$$

$$x^{(2)} = x^{(1)} + s r^{(1)} = \begin{pmatrix} 1{,}0007 \\ 0{,}9917 \\ 0{,}0009 \end{pmatrix}$$

$$\frac{\|x^2 - x^1\|_\infty}{\|x^2\|_\infty} = 0.09$$