Método do Gradiente para resolução de sistemas lineares

Cálculo Numérico

Instituto de Ciências Matemáticas e de Computação - ICMC - USP

Jeinny Peralta
Luiz Henrique Cherri

14/09/2012

Introdução

- Considere a matriz A de dimensão $n \times n$ e os vetores b e x de dimensão n;
- Dado o sistema linear $Ax = b$, desejamos obter a solução x deste sistema;
- O Método do Gradiente é utilizado para a resolução deste sistema;
- A ideia básica do método é encontrar o mínimo da função $f(x) = \frac{1}{2}x^T Ax - b^T x$.

Questão...

Quais condições são necessárias sobre uma função f para que um ponto x seja mínimo local desta função?

Resposta

Do cálculo temos que um ponto mínimo de uma função é aquele que satisfaz:

1. Gradiente$(f(x)) = \nabla f(x) = 0$;
2. Hessiana$(f(x)) = \nabla^2 f(x)$ definida positiva.

A primeira condição garante que o ponto seja estacionário, e a segunda garante que este seja um mínimo local.

Questão...

Então, quais condições são necessárias para que o ponto x^* seja o minimizador da função $f(x) = \frac{1}{2}x^T Ax - b^T x$?

Resposta

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

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

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

- Assim, para podermos aplicar o Método do Gradiente e encontrar o mínimo da função a matriz A deve ser simétrica e definida positiva;
- Sendo A simétrica e definida positiva, encontrar a solução do sistema linear $Ax = b$ é equivalente a encontrar o ponto x que satisfaz $\nabla f(x) = Ax - b = 0$, ou seja, o mínimo da função f.

Ideia do método

- O Método do Gradiente é um método iterativo. Portanto, é necessário partirmos de um ponto inicial $x^{(0)}$;
- Assim, dado o sistema $Ax = b$ e um ponto inicial $x^{(0)}$ nosso objetivo é construir uma sequência $\{x^{(k)}\}$ que convirja para a solução do sistema.

Note que, $\{x^{(k)}\}$ convergir para a solução do sistema é equivalente a $\{||r^{(k)}||\}$ convergir para zero, sendo,

$$r^{(k)} = b - Ax^{(k)} = -\nabla f(x^{(k)}),$$

o resíduo no ponto $x^{(k)}$.

Ideia do método - Direção

Para que a sequência $\{||r^{(k)}||\}$ tenda a zero, tomamos uma direção $d^{(k)}$ e atualizamos $x^{(k)}$ nesta direção, assim temos:

$$x^{(k+1)} = x^{(k)} + \alpha \overrightarrow{d^{(k)}},$$

de forma que $||r^{(k+1)}|| < ||r^{(k)}||$.
É natural tomarmos a direção $\overrightarrow{d^{(k)}}$ como sendo a direção oposta a $\nabla f(x^{(k)})$ devido ao fato que a função f cresce na direção do gradiente, ou seja, $\overrightarrow{d^{(k)}} = \overrightarrow{r^{(k)}}$.

Ideia do método - Tamanho do passo

Como $x^{(k+1)} = x^{(k)} + \alpha \overrightarrow{r^{(k)}}$ só nos resta definir α de forma que minimize $f(x^{(k+1)})$ na direção $\overrightarrow{r^{(k)}}$. Isso é equivalente a encontrar α tal que,

$$\frac{\partial f(x^{(k+1)})}{\partial \alpha} = 0.$$

Note que,

$$f(x^{(k+1)}) = \frac{1}{2}(x^{(k)} + \alpha r^{(k)})^T A(x^{(k)} + \alpha r^{(k)}) - b^T(x^{(k)} + \alpha r^{(k)}).$$

Ideia do método - Tamanho do passo

Com algumas manipulações algébricas temos,

$$\begin{aligned} f(x^{(k+1)}) &= \frac{1}{2}(x^{(k)})^T A x^{(k)} - b^T x^{(k)} + \alpha (r^{(k)})^T A x^{(k)} \\ &\quad + \frac{\alpha^2}{2}(r^{(k)})^T A r^{(k)} - \alpha b^T r^{(k)}. \end{aligned}$$

Assim,

$$\begin{aligned} \frac{\partial f(x^{(k+1)})}{\partial \alpha} &= (r^{(k)})^T A x^{(k)} + \alpha (r^{(k)})^T A r^{(k)} - b^T r^{(k)} \\ &= 0. \end{aligned}$$

Ideia do método - Tamanho do passo

Isolando α obtemos,

$$\alpha = \frac{(r^{(k)})^T r^{(k)}}{(r^{(k)})^T A r^{(k)}},$$

que é o melhor passo que pode ser dado nesta direção.

Resumo

Dado o sistema linear $Ax = b$ e um ponto inicial $x^{(0)}$, o Método do Gradiente para a resolução do sistema é definido com:

$$
\begin{aligned}
r^{(k)} &= b - Ax^{(k)}; \\
\alpha &= \frac{(r^{(k)})^T r^{(k)}}{(r^{(k)})^T A r^{(k)}}; \\
x^{(k+1)} &= x^{(k)} + \alpha r^{(k)}.
\end{aligned}
$$

Algoritmo

Dados: A, b, $x^{(0)}$ e a precisão ϵ desejada. O algoritmo do Método do Gradiente é definido abaixo:

1 Faça $k \leftarrow 0$;

2 Faça $r^{(k)} = b - Ax^{(k)}$ e $\alpha = \frac{(r^{(k)})^T r^{(k)}}{(r^{(k)})^T A r^{(k)}}$;

3 Atualize $x^{(k+1)} = x^{(k)} + \alpha r^{(k)}$;

4 Critério de parada:

- Se $\frac{||x^{(k+1)} - x^{(k)}||}{||x^{(k+1)}||} < \epsilon$ pare com a solução $x^{(k+1)}$;
- Se $k > MAX_{it}$ pare. Número máximo de iterações já realizado;

5 Faça $k \leftarrow k + 1$ e volte ao Passo 2.

Exemplo

Usando o Método dos Gradientes, obtenha uma solução aproximada para o sistema linear $Ax = b$ sendo,

$$A = \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix}$$

e

$$b = \begin{bmatrix} 5 \\ 4 \end{bmatrix},$$

a partir do ponto $x^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ e uma precisão $\epsilon = 10^{-1}$.

Resolução

$k = 0.$
Primeira iteração

$$r^{(0)} = b - Ax^{(0)} = b = \begin{bmatrix} 5 \\ 4 \end{bmatrix},$$

$$\alpha = \frac{r^{(0)})^T r^{(0)}}{r^{(0)})^T A r^{(0)}} = 0.22,$$

$$x^{(1)} = x^{(0)} + \alpha r^{(0)} = \begin{bmatrix} 1.09 \\ 0.87 \end{bmatrix}.$$

Verificando o critério de parada:

$$\frac{||x^{(1)} - x^{(0)}||}{||x^{(1)}||} = 1 > 10^{-1}.$$

Resolução

$k = 1$.
Segunda iteração

$$r^{(1)} = b - Ax^{(1)} = \begin{bmatrix} -0.23 \\ 0.30 \end{bmatrix},$$

$$\alpha = \frac{r^{(1)})^T r^{(1)}}{r^{(1)})^T A r^{(1)}} = 0.41,$$

$$x^{(2)} = x^{(1)} + \alpha r^{(1)} = \begin{bmatrix} 0.99 \\ 0.99 \end{bmatrix}.$$

Verificando o critério de parada:

$$\frac{||x^{(2)} - x^{(1)}||}{||x^{(2)}||} = 0.12 > 10^{-1}.$$

Resolução

$k = 2$.

Terceira iteração

$$r^{(2)} = b - Ax^{(2)} = \begin{bmatrix} 0.05 \\ 0.04 \end{bmatrix},$$

$$\alpha = \frac{r^{(1)})^T r^{(1)}}{r^{(1)})^T A r^{(1)}} = 0.22,$$

$$x^{(3)} = x^{(2)} + \alpha r^{(2)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Verificando o critério de parada:

$$\frac{||x^{(2)} - x^{(1)}||}{||x^{(2)}||} = 0.01 < 10^{-1}.$$

Referências

- Slides de aula:
  - Professora Silvia Maria Pereira Grandi dos Santos - USP - São Carlos;
  - Professora Marina Andretta - USP - São Carlos;
- Ana Friedlander - Elementos de Programação não Linear.