Matemática Numérica II

Trabalho de Março

Catarina Silva

Pergunta 1

Seja $\boldsymbol{R}: \boldsymbol{D} \subset \mathbb{R}^{\boldsymbol{n}} \to \mathbb{R}^{\boldsymbol{m}}$ uma função continuamente diferenciável numa vizinhança de $\boldsymbol{x_k}$. Denotemos por $\boldsymbol{J}(\boldsymbol{x}) \in \mathbb{R}^{\boldsymbol{m}\times\boldsymbol{n}}$ a matriz Jacobiana, em $\boldsymbol{x}$, da função vectorial $\boldsymbol{R}$. Prove que se a caracteristicca de $\boldsymbol{J}(\boldsymbol{x_k})$ for igual a n então o passo Gauss-Newton

$$\boldsymbol{p_k} = -(\boldsymbol{J}(\boldsymbol{x_k})^{\boldsymbol{T}}\boldsymbol{J}(\boldsymbol{x_k}))^{-\boldsymbol{1}}\boldsymbol{J}(\boldsymbol{x_k})^{\boldsymbol{T}}\boldsymbol{R}(\boldsymbol{x_k})$$

Em direcção de descida para a função $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{R}(\boldsymbol{x})^{\boldsymbol{T}}\boldsymbol{R}(\boldsymbol{x})/\boldsymbol{2}$.

- -

O problema anterior é traduzido pela seguinte implicação

$$Car(n) = n \Rightarrow \quad o\ passo\ Gauss - Newton \quad p_k = -(J(x_k)^T J(x_k))^{-1} J(x_k)^T R(x_k)$$
$$\text{é}\ uma\ direcção\ de\ descida\ para\ a\ função\ f(x) = R(x)^T R(x)/2.$$

- Uma das condições para que o método de Gauss-Newton esteja bem definido, este tem de gerar iterações de modo que $J(x_k)^T J(x_k)$ seja não singular.

- **Proposição:** Seja $f: D \subset \mathbb{R}^n \to \mathbb{R}$ uma função direcção de continuamente diferenciável em D aberto e $x_* \in D$. A direcção de descida em x_* se $-\nabla f(x_*)^T p > 0$.

Provemos primeiro que está bem definida.

- Consideremos então λ valor próprio de $A^T A$, então $Av = \lambda v$, para v vector próprio associado λ.

Deste modo $\quad v^T A^T A v = v^T \lambda v \Leftrightarrow (Av)^T Av = \lambda v^T v, \qquad com\ v^T v = \|v\|^2$

$$\Leftrightarrow \lambda = \frac{\|Av\|^2}{\|v\|^2},\ onde\ \frac{\|Av\|^2}{\|v\|^2} > 0\ (1)$$

- $((J(x_k)^T J(x_k))_{\mathrm{ij}} = \left[\frac{d\mathrm{r_i}}{d\mathrm{x_i}}, \dots, \frac{d\mathrm{r_i}}{d\mathrm{x_n}}\right] \begin{bmatrix} \frac{d\mathrm{r_j}}{d\mathrm{x_i}} \\ \cdot \\ \cdot \\ \cdot \\ \frac{d\mathrm{r_j}}{d\mathrm{x_n}} \end{bmatrix} = ((J(x_k)^T J(x_k))_{\mathrm{ji}}$ ou seja, é Simétrica (2)

Logo todos os valores próprios de $J(x_k)^T J(x_k)$ são positivos por (1). Como é simétrica por (2) concluímos que $J(x_k)^T J(x_k)^{-1}$ é definida positiva (3).

Provemos agora a direcção de descida, ou seja que $-\nabla f(x_*)^T p > 0$.

$$-\nabla f(x)^T p_k = (J(x)^T R(x)^T (J(x)^T J(x))^{-1} J(x)^T R(x) \quad > 0 \quad \text{por (3)}$$

1. **Considere a função de Rosenbrock**

$$\boldsymbol{f(x_1, x_2) = 100\left(x_2 - {x_1}^2\right)^2 + (1 - x_1)^2} \ .$$

a) **Mostre que $\boldsymbol{x_* = [1\ 1]^T}$ é o único minimizante local de $\boldsymbol{f}$ e que $\boldsymbol{\nabla^2 f(x_*)}$ é definida positiva.**

Provemos então o pedido, para tal é necessário ter em conta a seguinte definição

Se $f : D \subset \mathbb{R}^n \to \mathbb{R}$, D aberto, for duas vezes continuamente diferenciável em D, se $x^* \in D$ tal que $\nabla f(x^*) = 0$ e $\nabla^2 f(x^*)$ for definida positiva, então x^* é um minimizante local de f.

Queremos então encontrar os extremos, e se estes existiram serão os nossos pontos críticos. Calculemos então os pontos do domínio onde o vector gradiente se anula.

$$\nabla f(x_1, x_2) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 400x_1^3 - 400x_1x_2 + 2x_1 - 2 \\ 200(x_2 - x_1^2) \end{bmatrix}$$

$$\nabla f(x_1, x_2) = 0 \iff \begin{bmatrix} 400x_1^3 - 400x_1x_2 + 2x_1 - 2 \\ 200(x_2 - x_1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff \begin{cases} x_1 = 1 \\ x_2 = 1 \end{cases}$$

Determine o minimizante de f usando o método de Newton e método de BFGS. Considere $x_0 = [-1.2\ 1]^T$ e $B_0 = (\nabla^2 f(x_0))^{-1}$. Faça uma tabela de valores $\|\ x_k - x_*\ \|$ até 40 iterações para os dois métodos.

b) Se usarmos o método de descida mais rápida para resolver o problema da alínea anterior observa-se que são necessárias 5264 iterações para reduzir a norma do gradiente para 10^{-5}. Quantas iterações dos métodos de Newton e BFGS são necessárias para fazer o mesmo?

Pergunta 3

Prove que nas condições do Teorema da convergência local do método de Newton e do segundo Teorema da Aula 5, temos que (a) e (b) são equivalentes

a) $\lim_{k\to\infty} \frac{\|J(x_k)(p_k - s_k)\|}{\|s_k\|} = 0$ b) $\lim_{k\to\infty} \frac{\|Jp_k - s_k\|}{\|s_k\|} = 0$

Onde $s_k = -A_k^{-1}F(x_k)$, $p_k = -J(x_k)^{-1}F(x_k)$. As matrizes A_k são as do método de Broyden.

- -

Como queremos provar a equivalência, ou seja,

$$\lim_{k\to\infty} \frac{\|J(x_k)(p_k - s_k)\|}{\|s_k\|} = 0 \Leftrightarrow \lim_{k\to\infty} \frac{\|p_k - s_k\|}{\|s_k\|} = 0$$

teremos de provar nos dois sentidos.

Pelo teorema temos da aula 5:

$$\big(A_k - J(x_k)\big)s_k = A_k s_k - J(x_k)s_k = -F(x_k) - J(x_k)s_k = J(x_k)(p_k - s_k)$$

Provemos então no sentido $\Leftarrow$

$$\lim_{k\to\infty} \frac{\|J(x_k)(p_k - s_k)\|}{\|s_k\|} \le \lim_{k\to\infty} \|J(x_k)\| \frac{\|(p_k - s_k)\|}{\|s_k\|}$$

E por hipótese temos

$$\lim_{k\to\infty} \frac{\|p_k - s_k\|}{\|s_k\|} = 0$$

Basta provar que $J(x_k)$ é limitada, para que $\|J(x_k)\|$ seja limitada.
Então pelo teorema da aula 5, J é continua à Lipschitz, isto é,

$$\|J(x_k) - J(x_*)\| \le \gamma\|x_k - x_*\| \,.$$

Como $|\|J(x_k)\| - \|J(x_*)\|| \le \|J(x_k) - J(x_*)\|$ e

$$|\|J(x_k)\| - \|J(x_*)\|| \le \gamma\|x_k - x_*\| \,.$$

Temos então que $\|J(x_k)\| - \|J(x_*)\| \le \gamma\|x_k - x_*\|$

Logo $|\|J(x_k)\|| \le |\|J(x_*)\|| + \gamma\|x_k - x_*\|$, ou seja limitada como queríamos.

Provemos então no sentido ⇒

$$\left(A_k - J(x_k)\right)s_k = J(x_k)(p_k - s_k)$$

$$p_k - s_k = J(x_k)^{-1}(A_k - J(x_k))s_k$$

$$\lim_{k\to\infty}\frac{\|(p_k - s_k)\|}{\|s_k\|} = \lim_{k\to\infty}\frac{\left\|J(x_k)^{-1}(A_k - J(x_k))s_k\right\|}{\|s_k\|} \leq \lim_{k\to\infty}\|J(x_k)^{-1}\|\frac{\|(A_k - J(x_k))s_k\|}{\|s_k\|}$$

$$\lim_{k\to\infty}\frac{\|(A_k - J(x_k))s_k\|}{\|s_k\|} = 0$$

Logo é igual a 0.