答案

2022 年 10 月 23 日

问题1

证明. 记$x_0 = (0.1, 1)^T$,$f'(x) = 2(10x_1, x_2)^T$ 根据最速下降法的迭代公式$x_k = x_{k-1} - \alpha f^{'}(x_{k-1})$，得:

$$\|x_k\|^2 = (1 - \frac{\alpha}{10})^2 (x_{k-1}^1)^2 + (1 - \alpha)^2 (x_{k-1}^2)^2,$$

选取$0 < \alpha < 1$有$\|x_k - 0\| \le (1 - \frac{\alpha}{10})\|x_{k-1} - 0\|$,此时最速下降法线性收敛。

□

问题2

证明. 由利普希兹连续的性质，$\forall x$,

$$f(x - \alpha \nabla f(x)) \le f(x) - \alpha(1 - \frac{L\alpha}{2})\|\nabla f(x)\|^2, \tag{1}$$

记$\tilde{x} = x - \alpha f(x)$并限制$0 < \alpha < \frac{1}{L}$，有:

$$\begin{aligned}
f(\tilde{x}) &\le f(x) - \frac{\alpha}{2}\|\nabla f(x)\|^2 \\
&\le f^* + \nabla f(x)^T (x - x^*) - \frac{\alpha}{2}\|\nabla f(x)\|^2 \\
&= f^* + \frac{1}{2\alpha}(\|x - x^*\|^2 - \|x - x^* - \alpha \nabla f(x)\|^2) \\
&= f^* + \frac{1}{2\alpha}(\|x - x^*\| - \|\tilde{x} - x^*\|^2),
\end{aligned}$$

取$x=x^{i-1}$,则$\tilde{x}=x^i$。并将$i=1,2,\cdots,k$求和得到

$$\begin{aligned}\sum_{i=1}^{k}(f(x^i)-f^*)) &\leq \frac{1}{2\alpha}\sum_{i=1}^{k}(\|x^{i-1}-x^*\|^2-\|x^i-x^*\|^2)\\ &=\frac{1}{2\alpha}(\|x^{)}-x^*\|^2-\|x^k-x^*\|^2)\\ &\leq \frac{1}{2\alpha}\|x^0-x^*\|^2.\end{aligned}$$

由(1)知$f(x^i)$是非增的，所以

$$f(x^k)-f^* \leq \frac{1}{k}\sum_{i=1}^{k}(f(x^i)-f^*) \leq \frac{1}{2k\alpha}\|x^0-x^*\|^2.$$

□

问题3

1. 给点初始点$x_0,\epsilon>0,k=1$。
2. 若终止准则满足，则输出有关信息，停止迭代。
3. $x_{k+1}=x_k-G_K^{-1}g_kk=k+1$,转2

$f^{'}(x)=4x^3-12x^2-12x-16, f^{''}(x)=12(x^2-2x-1)$ 计算过程如下：

1. $x_1=x_0-\frac{f^{'}(x_0)}{f^{''}(x_0)}=5.167;$
2. $x_2=x_1-\frac{f^{'}(x_1)}{f^{''}(x_1)}=4.335;$
3. $x_3=x_2-\frac{f^{'}(x_2)}{f^{''}(x_2)}=4.040;$
4. $x_4=x_3-\frac{f^{'}(x_3)}{f^{''}(x_3)}=4.001$。

问题4 $G_0=\begin{pmatrix}0&1\\1&2\end{pmatrix}$, $G_0+\overline{v}I=\begin{pmatrix}\overline{v}&1\\1&\overline{v}+2\end{pmatrix}$要使$G_0+\overline{v}I$正定只需$\overline{v}>0,|G_0+\overline{v}I|>0$,解得:$\overline{v}>\sqrt{2}-1$令$v_0=1$得

$$d_0=-(G_0+I)^{-1}g_0=-\begin{pmatrix}1&1\\1&3\end{pmatrix}^{-1}\begin{pmatrix}0\\2\end{pmatrix}=\begin{pmatrix}1\\-1\end{pmatrix},$$

从而$f(x_0+d_0)=0<f(0,0)=1$,最后由LM算法可知$x_0+d_0=\begin{pmatrix}\frac{2}{\overline{v}^2+2\overline{v}-1}\\\frac{-2\overline{v}}{\overline{v}^2+2\overline{v}-1}\end{pmatrix}$编程计算可得$\overline{v}<0.9004$。

问题 5

证明. 先证$\forall k$，有$H^k y^{k-1} = s^{k-1}$.

$$\begin{aligned}
H^k y^{k-1} &= H^{K-1} y^{k-1} + \frac{(s^{k-1} - H^{k-1} - y^{k-1})(s^{k-1} - H^{K-1} y^{k-1})^T}{(s^{k-1} - H^{K-1} y^{k-1})^T y^{K-1}} y^{k-1} \\
&= H^{k-1} y^{k-1} + s^{k-1} - H^{k-1} y^{k-1} \\
&= s^{k-1},
\end{aligned}$$

考虑$j = k-2$的情况：

$$\begin{aligned}
H^k y^{k-2} &= H^{K-1} y^{k-2} + \frac{(s^{k-1} - H^{k-1} - y^{k-1})(s^{k-1} - H^{K-1} y^{k-1})^T}{(s^{k-1} - H^{K-1} y^{k-1})^T y^{K-1}} y^{k-2} \\
&= s^{k-2} + \frac{(s^{k-1} - H^{k-1} - y^{k-1})(s^{k-1} - H^{K-1} y^{k-1})^T}{(s^{k-1} - H^{K-1} y^{k-1})^T y^{K-1}} y^{k-2},
\end{aligned}$$

由于

$$\begin{aligned}
(s^{k-1} - H^{K-1} y^{k-1})^T y^{k-2} &= (s^{k-1})^T y^{k-2} - (y^{k-1})^T (H^{k-1})^T y^{k-2} \\
&= (s^{k-1})^T y^{k-2} - (y^{k-1})^T H^{k-1} y^{k-2} \\
&= (s^{k-1})^T y^{k-2} - (y^{k-1})^T s^{k-1} = 0
\end{aligned}$$

从而$j = k-2$时成立, 类似可证$\forall j = 1, 2, \cdots, k-1$.成立。 □

问题 6 解：$\nabla f(x) = \begin{pmatrix} 2x_1 \\ 8x_2 \end{pmatrix}, \nabla^2 f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$

(1)$\nabla f(x^0) = \begin{pmatrix} 2 \\ 8 \end{pmatrix}, d^0 = -\nabla f(x^0) = \begin{pmatrix} -2 \\ -8 \end{pmatrix}$

$\phi(\alpha) = f(x^0 + \alpha d^0) = (1-2\alpha)^2 + 4(1-8\alpha)^2$

令$\phi^{'}(\alpha) = 0$，得$\alpha = \frac{68}{520} = 0.13077$

进而有$x^1 = x^0 + \alpha_0 d^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - 0.13077 \begin{pmatrix} 2 \\ 8 \end{pmatrix} = \begin{pmatrix} 0.73846 \\ -0.04616 \end{pmatrix}$

$\nabla f(x^1) = (1.47692, -0.36923)^T, ||\nabla f(x^1)|| = 1.52237$

由于$||\nabla f(x^1)||$较大，因此还需要迭代下去。(2)先确定搜索方向，为此需要按公式计算

$$H_1 = H_0 + \frac{\delta_1 \delta_1^T}{\delta_1^T \gamma_1} - \frac{H_0 \gamma_1 \gamma_1^T H_0}{\gamma_1^T H_0 \gamma_1}$$

又$\delta_1 = x^1 - x^0 = \begin{pmatrix} -0.26154 \\ -1.04616 \end{pmatrix}, \gamma_1 = \nabla f(x^1) - \nabla f(x^0) = \begin{pmatrix} -0.52308 \\ -8.36923 \end{pmatrix}$
故$H_1 = \begin{pmatrix} 1.00380 & -0.03149 \\ -0.03149 & 0.12697 \end{pmatrix}$
下面再确定步长：
利用$\frac{df(x^1+\alpha d^1)}{d\alpha} = 0$，得到$\alpha_1 = 0.49423$
所以$x^2 = x^1 + \alpha_1 d^1 = \begin{pmatrix} 0.00000 \\ 0.00000 \end{pmatrix}$
由于$||\nabla f(x^2)|| < \epsilon$，因此停止迭代，最优解为$x^* = x^2$.

问题 7

证明. 定理的证明归结于用归纳法证明如下不等式：

$$-\sum_{j=0}^{k} \sigma^j \leq \frac{g_k^T d_k}{||g_k||^2} \leq -2 + \sum_{j=0}^{k} \sigma^j \tag{2}$$

对于所有k成立，其中$\sigma \in (0, \frac{1}{2})$.事实上，若(2)成立，由于

$$\sum_{j=0}^{k} \sigma^j < \sum_{j=0}^{\infty} = \frac{1}{1-\sigma} \tag{3}$$

由此可知(2)式右侧为负，从而

$$g_k^T d_k < 0 \tag{4}$$

下降性质成立. 显然，当$k = 0$时，(2)式成立.今设对任何$k \geq 0$,(2)成立.
由FR算法的迭代过程，有

$$\frac{g_{k+1}^T d_{k+1}}{||g_{k+1}||^2} = -1 + \frac{g_{k+1}^T d_k}{||g_k||^2} \tag{5}$$

则有

$$-1 + \sigma \frac{g_k^T d_k}{||g_k||^2} \leq \frac{g_{k+1}^T d_{k+1}}{||g_{k+1}||^2} \leq -1 - \sigma \frac{g_k^T d_k}{||g_k||^2} \tag{6}$$

再由归纳法假设(2)，有

$$-\sum_{j=0}^{k+1} \sigma^j = -1 - \sigma \sum_{j=0}^{k} \sigma^j \leq \frac{g_{k+1}^T d_{k+1}}{||g_{k+1}||^2} \leq -1 + \sigma \sum_{j=0}^{k} \sigma^j = -2 + \sum_{j=0}^{k+1} \sigma^j. \tag{7}$$

于是，对于$k+1$，不等式(2)成立.

□

问题 8 解：$g(x) = (3x_1 - x_2 - 2, x_2 - x_1)^T, G(x) = \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}$

因$g_0 = (-2, 0)^T \neq 0$，故取$p_0 = (2, 0)^T$，从x_0出发，沿p_0做一维搜索，也就是求

$$min f(x_0 + \alpha p_0) = 6\alpha^2 - 4\alpha$$

的极小点，得步长$\alpha_0 = \frac{1}{3}$.于是得到$x_1 = x_0 + \alpha_0 p_0 = (\frac{2}{3}, 0), g_1 = (0, -\frac{2}{3})^T$.由FR公式得

$$\beta_0 = \frac{g_1^T g_1}{g_0^T g_0} = \frac{1}{9}$$

故$p_1 = -g_1 + \beta_0 p_0 = (\frac{2}{9}, \frac{2}{3})^T$. 从$x_1$出发，沿$p_1$作一维搜索，求

$$min f(x_1 + \alpha p_1) = \frac{4}{27}\alpha^2 - \frac{4}{9}\alpha + \frac{2}{3}$$

的极小点，解得$\alpha_1 = \frac{3}{2}$，于是$x_2 = x_1 + \alpha_1 p_1 = (1, 1)^T$ 此时$g_2 = (0, 0)^T$，故$x^* = x_2 = (1, 1)^T, f^* = -1$