2. 设 $f(x)=x_1^4+x_1^2+x_2^2$. 给定当前 $x_k=(1,1)^T$, 搜索方向为 $d_k=(-3,-1)^T$ 设 $\rho=0.5$, $\sigma=1$. 分别检验步长因子 α 取值为 0.1, 0.5, 1 时, 是否满足 Goldstein 与 Wolfe 准则.

解: 首先因 $\rho\notin(0,\frac{1}{2})$ 故 Goldstein 准则全不满足.

$$\nabla f=\begin{pmatrix}4x_1^3+2x_1\\2x_2\end{pmatrix}$$

当 $\alpha=0.1$ 时. $f(x_k+\alpha d_k)=1.54\leq 2=f(x_k)+\alpha\rho g_k^Td_k$

$g(x_k+\alpha_k d_k)^Td_k=-10.116\geqslant g_k^Td_k=-20$

故符合 Wolfe 准则.

当 $\alpha=1/2$ 时. $f(x_k+\alpha d_k)=0.5626\geqslant -2. =f(x_k)+\frac{1}{2}\alpha g_k^Td_k.$

故不符合 Wolfe 准则.

当 $\alpha=1$ 时. $f(x_k+\alpha d_k)=20\geqslant -7=f(x_k)+\frac{1}{2}\alpha g_k^Td_k.$

故不符合 Wolfe 准则.

□

3. 证明：若 $\rho<\frac{1}{2}$，则正定二次函数精确线性搜索步长满足 Goldstein 准则。

证：
$$f(x)=\frac{1}{2}x^TGx+b^Tx+c \qquad G\succ 0$$

$$\alpha_k=\min_{\alpha>0} f(x+\alpha_k d_k)$$

$$x_k^TG^Td_k+\alpha d_k^TG^Td_k+b^Td_k=0$$

$$\alpha_k=\frac{-x_k^TG^Td_k-b^Td_k}{d_k^TG^Td_k}$$

$$g_k=Gx_k+b^T$$

$$\alpha_k=\min\{\alpha>0 \mid (G(x_k+\alpha d_k)+b)^Td_k=0\}$$
$$=\frac{-g_k^Td_k}{d_k^TGd_k}$$

$$f(x_k+\alpha_kd_k)=\frac{1}{2}(x_k+\alpha_kd_k)^TG(x_k+\alpha_kd_k)+b^T(x_k+\alpha d_k)+c$$
$$=\frac{1}{2}(x_kGx_k+2\alpha_kd_k^TGx_k+\alpha_k^2d_k^TGd_k)+b^Tx_k+\alpha b^Td_k+c$$
$$=f(x)+\alpha_kd_k^TGx_k+\frac{1}{2}\alpha_k^2d_k^TGd_k+\alpha b^Td_k$$
$$=f(x)+\alpha_kd_k^T(g_k-b^T)+\frac{\alpha_k^2}{2}d_k^TGd_k+\alpha b^Td_k$$
$$=f(x)+\alpha_kd_k^Tg_k+\frac{\alpha_k^2}{2}d_k^TGd_k$$

Goldstein：
$$f(x_k+\alpha_kd_k)\le f(x_k)+\rho\alpha_kg_k^Td_k$$
$$\ge f(x_k)+(1-\rho)\alpha_kg_k^Td_k$$

故 Goldstein 满足 $\Leftrightarrow (1-\rho)\alpha_kg_k^Td_k\le\alpha_kd_k^Tg_k+\frac{\alpha_k^2}{2}d_k^TGd_k\le\rho\alpha_kg_k^Td_k$

左侧 $\Leftrightarrow -\rho g_k^Td_k\le\frac{\alpha_k}{2}d_k^TGd_k \Leftrightarrow -\rho g_k^Td_k\le\frac{1}{2}g_k^Td_k$ 显然成立

右侧 $\Leftrightarrow (\rho-1)g_k^Td_k\ge\frac{1}{2}g_k^Td_k$ 也显然成立

$\square$

第四章

1. $\min f(x) = 10x_1^2 + x_2^2$.

选初始点为 $(0.1,1)^T$. 证明最速下降法只线性收敛的.

首先 $f(x_k) \geq 0 \quad \forall x_k \in \mathbb{R}^2$

而最速下降法与精确线性搜索步长

$d_k = -g_k$ 为下降方向. $\quad \alpha_k = \underset{\alpha>0}{\arg\min} f(x_k + \alpha d_k)$

故 $f_{k+1} < f_k$ 对 $\forall k \in \mathbb{N}^*$, $\forall \mathbb{R}^2/\{0\}$.

则有 f_k 单调下降且有下界, 由单调有界序理知 f_k 收敛.

$$\lim_{k\to\infty} \frac{f_{k+1}}{f_k} = \frac{\lim_{k\to\infty} f_{k+1}}{\lim_{k\to\infty} f_k} = 1. \qquad \square$$

2. $\min f(x) = x_1^2+2x_2^2+4x_1+4x_2$.

设 $x^{(1)} = (0,0)^T$

(a) 证明最速下降法产生的点列会成为

$$x^{(k+1)} = \left(\frac{2}{3^k}-2, \left(-\frac{1}{3}\right)^k-1\right)^T$$

$$G = \begin{pmatrix} 2 & \\ & 4 \end{pmatrix} \quad b = \begin{pmatrix} 4 \\ 4 \end{pmatrix}. \quad f(x) = \frac{1}{2}x^T G x + bx.$$

$g_k = G x_k + b$, 而 $d_k = -g_k$, 故 $\alpha_k = \frac{g_k^T g_k}{g_k^T G g_k} = \frac{(4+2x_1^k)^2+(4+4x_2^k)^2}{2(4+2x_1^k)^2+4(4+4x_2^k)^2}$

$= (2x_1^k+4, 4x_2^k+4)$.

用数学归纳法. $x^{(2)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{4^2+4^2}{2\cdot 4^2+4\cdot 4^2} \cdot \begin{pmatrix} 4 \\ 4 \end{pmatrix} = \left(-\frac{4}{3}, -\frac{4}{3}\right)^T$ 成立.

当 x^k 成立时. 考虑 x^{k+1}

$$x^{(k+1)} = x^{(k)} + \alpha_k d_k = x^{(k)} - \frac{(4+2x_1^{(k)})^2+(4+4x_2^{(k)})^2}{2(4+2x_1^{(k)})^2+4(4+4x_2^{(k)})^2}\begin{pmatrix} 2x_1^{(k)}+4 \\ 4x_2^{(k)}+4 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{2}{3^k}-2 \\ \left(-\frac{1}{3}\right)^k-1 \end{pmatrix} - \frac{\left(\frac{4}{3^k}\right)^2+\left(4\left(-\frac{1}{3}\right)^k\right)^2}{2\left(\frac{4}{3^k}\right)^2+4\left(4\left(-\frac{1}{3}\right)^k\right)^2}\begin{pmatrix} \frac{4}{3^k} \\ 4\left(-\frac{1}{3}\right)^k \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{3^k}-2 \\ \left(-\frac{1}{3}\right)^k-1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} \frac{4}{3^k} \\ 4\left(-\frac{1}{3}\right)^k \end{pmatrix} = \begin{pmatrix} \frac{2}{3^{k+1}}-2 \\ \left(-\frac{1}{3}\right)^{k+1}-1 \end{pmatrix}. \quad \square$$

(b) 试用牛顿法解以上问题.

$$g_k = \nabla f(x_k) = (2x_1^{(k)}+4, 4x_2^{(k)}+4)^T.$$

$$G_k = \nabla^2 f(x_k) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}. \quad G_k^{-1} = \begin{pmatrix} 1/2 & \\ & 1/4 \end{pmatrix}.$$

$$x^{(2)} = x^{(1)} - \begin{pmatrix} 1/2 & \\ & 1/4 \end{pmatrix}\begin{pmatrix} 4 \\ 4 \end{pmatrix} = -\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

因二次函数正定, 一步收敛. $\begin{pmatrix} -2 \\ -1 \end{pmatrix}$ 为最优解.

$\square$

3. 试用共轭梯度法，求下面二次函数的极小点：

$$f(x)=\frac{3}{2}x_1^2+\frac{1}{2}x_2^2-x_1x_2-2x_1$$

初始点：$x^{(0)}=(-2,4)^T$

解：$f(x)=\frac{1}{2}xGx-bx \qquad G=\begin{pmatrix}3&-1\\-1&1\end{pmatrix} \qquad b=\begin{pmatrix}2\\0\end{pmatrix}$

$$r_0=Gx_0-b=\begin{pmatrix}12\\-6\end{pmatrix}$$

$$d_0=-r_0$$

$$\alpha_0=\frac{r_0^Tr_0}{r_0^TGr_0}=\frac{12^2+6^2}{(12,-6)\begin{pmatrix}3&-1\\-1&1\end{pmatrix}\begin{pmatrix}12\\-6\end{pmatrix}}=\frac{5}{17}$$

$$x_1=\begin{pmatrix}-2\\4\end{pmatrix}+\frac{5}{17}\cdot\begin{pmatrix}-12\\6\end{pmatrix}=\frac{2}{17}\begin{pmatrix}13\\19\end{pmatrix}$$

$$r_1=Gx_1-b=\frac{2}{17}\begin{pmatrix}3&-1\\-1&1\end{pmatrix}\begin{pmatrix}13\\19\end{pmatrix}-\begin{pmatrix}2\\0\end{pmatrix}=\frac{6}{17}\begin{pmatrix}1\\2\end{pmatrix}$$

$$\beta_0=\frac{r_1^Tr_1}{r_0^Tr_0}=\frac{(\frac{6}{17})^2\cdot 5}{12^2+6^2}=\frac{1}{289}$$

$$d_1=-r_1+\beta_0d_0=-\frac{6}{17}\begin{pmatrix}1\\2\end{pmatrix}+\frac{1}{289}\begin{pmatrix}-12\\6\end{pmatrix}=-\frac{30}{289}\begin{pmatrix}3\\7\end{pmatrix}$$

$$\alpha_1=\frac{r_1^Tr_1}{d_1^TGd_1}=\frac{(\frac{6}{17})^2\cdot 5}{(30/289)^2\,(3,7)\begin{pmatrix}3&-1\\-1&1\end{pmatrix}\begin{pmatrix}3\\7\end{pmatrix}}=\frac{17}{10}$$

$$x_2=x_1+\alpha_1d_1=\frac{2}{17}\begin{pmatrix}13\\19\end{pmatrix}+\frac{17}{10}\cdot\frac{30}{289}\begin{pmatrix}3\\7\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$$

$r_2=\begin{pmatrix}0\\0\end{pmatrix}$ 　　故，$x_2=(1,1)^T$ 为所求极小点.

$\square$