The Key for Homework 1

2022 年 10 月 14 日

1. 证明. 必要性：
   $\forall(x_1,y_1),(x_2,y_2)\in epi(f)$，以及$\alpha\in[0,1]$，有$f(x_1)\le y_1$,$f(x_2)\le y_2$。由凸函数的性质：

$$\begin{aligned} f(\alpha x_1+(1-\alpha)x_2) &\le \alpha f(x_1)+(1-\alpha)f(x_2) \\ &\le \alpha y_1+(1-\alpha)y_2, \end{aligned}$$

   故 $(\alpha x_1+(1-\alpha)x_2,\alpha y_1+(1-\alpha)y_2)\in epi(f)$，$epi(f)$为凸集。

   充分性：
   假设$epi(f)$为凸集。$\forall x_1,x_2\in dom(f)$,有$(x_1,f(x_1)),(x_2,f(x_2))\in epi(f)$, $\forall\alpha\in[0,1]$,由凸集的性质得$(\alpha x_1+(1-\alpha)x_2,\alpha f(x_1)+(1-\alpha)f(x_2))\in epi(f)$即：

$$f(\alpha x_1+(1-\alpha)x_2)\le \alpha f(x_1)+(1-\alpha)f(x_2),$$

   $f(x)$为凸函数得证。 □

2. 证明. 2.1(局部全局定理): 设x^*为$f(x)$的一个局部极小值点，由$f(x)$的可微性，x^*一定为稳定点。有

$$f'(x^*)=0,$$

   由凸函数的性质，我们得到：

$$f(y)\ge f(x^*)+f'(x^*)^\top(y-x^x)=f(x^*),$$

   其中$y\in dom(f)$。从而 x^*为全局极小值点。

   2.2(反证): 假设$y^*\ne x^*$为另外的极小值点，由2.1的结论可知，y^*也是全局极小值点，故有$f(x^*)=f(y^*)$,$\forall\alpha\in(0,1)$，由$f(x^*)$严格凸:

$$f(\alpha x^*+(1-\alpha)y^*)<\alpha f(x^*)+(1-\alpha)f(y^*)=f(x^*).$$

   这与x^*为全局极小值点矛盾，原问题得证。

   □

3. 解：易得三个子问题中的迭代序列均收敛到0。

3.1

$$\lim_{k\to\infty}\frac{2^{-(k+1)}}{2^{-k}}=\frac{1}{2}$$

由$0<\frac{1}{2}<1$,序列线性收敛。

3.2

$$\begin{aligned}\lim_{k\to\infty}\frac{(k+1)^{-(k+1)}}{k^{-k}}&=\lim_{k\to\infty}\frac{(k)^k}{(k+1)^{(k+1)}}\\&=\lim_{k\to\infty}\frac{1}{k+1}\times(\frac{k}{k+1})^k\\&=0\times\frac{1}{e}=0\end{aligned}$$

序列超线性收敛。

3.3

$$\lim_{k\to\infty}\frac{a^{2^{k+1}}}{(a^{2^k})^2}=\lim_{k\to\infty}a^{2^{k+1}-2^{k+1}}=1$$

序列为二阶收敛。

4. 证明. 由假设可知

$$f(x_k+\alpha d_k)\leq f(x_k)+\alpha d_k^T\nabla f(x_k)+\frac{\alpha^2}{2}M\|d_k\|^2$$

对一切$\alpha>0$都成立.令

$$\bar{\alpha}=-d_k^T\nabla f(x_k)/(M\|d_k\|^2),$$

则有

$$\begin{aligned}f(x_k)-f(x_k+\alpha_k d_k)&\geq f(x_k)-f(x_k+\bar{\alpha}d_k)\\&\geq-\bar{\alpha}d_k^T\nabla f(x_k)-\frac{\bar{\alpha}^2}{2}M\|d_k\|^2\\&=\frac{1}{2}\frac{(d_k^T\nabla f(x_k))^2}{M\|d_k\|^2}\\&=\frac{1}{2M}\|\nabla f(x_k)\|^2\frac{(d_k^T\nabla f(x_k))^2}{\|d_k\|^2\|\nabla f(x_k)\|^2}\\&=\frac{1}{2M}\|\nabla f(x_k)\|^2cos^2\langle d_k,-\nabla f(x_k)\rangle.\end{aligned}$$

□

5. 证明. $f(x_k+\alpha d_k)$在$\alpha=0$处的斜率为$g_k^Td_k$.

$$\forall\epsilon>0,\exists|\alpha_1-0|<\delta$$

$$s.t|\frac{f(x_k+\alpha d_k)-f(x_k)}{\alpha_1}-g_k^Td_k|<\epsilon$$

则有$f(x_k+\alpha_1 d_k)<f(x_k)+\alpha_1 g_k^T d_k+\epsilon<f(x_k)+\rho\alpha_1 g_k^T d_k$.

假设$f(x_k+\alpha d_k)<f(x_k)+\rho g_k^T d_k\alpha$对$\alpha>0$恒成立.

令$\alpha\to+\infty$,有$f(x_k+\alpha d_k)\to-\infty$.

这与$f(x_k+\alpha d_k)$在$\alpha>0$时有下界相矛盾.

故必存在一个α^*使得$f(x_k+\alpha^* d_k)=f(x_k)+\rho g_k^T d_k\alpha^*$.(不妨设$\alpha^*$为$f(x_k+\alpha d_k)$与$f(x_k)+\rho\alpha_1 g_k^T d_k$在$\alpha>0$范围内的首个交点)

则α^*满足Goldstein准则.

因此，在$[0,\alpha^*]$内都有$f(x_k+\alpha d_k)\geq f(x_k)+\rho g_k^T d_k\alpha$成立.

记$h(x)=f(x_k+\alpha d_k)-f(x_k)-\rho g_k^T d_k$.

很容易知道$h(x)$满足Rolle定理的条件.

故存在$\alpha_0\in[0,\alpha^*]$使得$g(x_k+\alpha d_k)^T d_k=\rho g_k^T d_k>\sigma g_k^T d_k$成立.

则α_0满足Wolfe准则. □

6. 解：算法如下：

第一步：给定$x_0,\epsilon>0,k=0$；

第二步：若终止条件满足，则迭代终止；

第三步：$d_k=-g_k$；

第四步：精确线搜索求α_k；

第五步：$x_{k+1}=x_k+\alpha_k d_k,k=k+1$，转第二步。

以下是对无约束问题的解答：

$\phi(\alpha)=f(x+\alpha d)=(x_1+\alpha d_1)^2+2(x_2+\alpha d_2)^2$

令$\phi^{'}(\alpha)=0$，得$\alpha=-\frac{d_1x_1+2d_2x_2}{d_1^2+2d_2^2};\nabla f(x)=(2x_1,4x_2)^T$

第1次迭代：$\nabla f(x^{(0)})=(2,4)^T,d^{(0)}=-\nabla f(x^{(0)})=(-2,-4)^T,\phi(\alpha)=f(x^{(0)}+\alpha d^{(0)})$,

令$\phi^{'}(\alpha)=0$，得$\alpha_0=\frac{5}{18},x^{(1)}=x^{(0)}+\alpha_0 d^{(0)}=(\frac{4}{9},-\frac{1}{9})^T$

第2次迭代：$\nabla f(x^{(1)})=(\frac{8}{9},-\frac{4}{9})^T,d^{(1)}=-\nabla f(x^{(1)})=(-\frac{8}{9},\frac{4}{9})^T$

$\phi(\alpha)=f(x^{(1)}+\alpha d^{(1)})$，令$\phi^{'}(\alpha)=0$，得$\alpha_1=\frac{5}{12}$，故$x^{(2)}=x^{(1)}+\alpha_1 d^{(1)}=(\frac{2}{27},\frac{2}{27})^T$

如此迭代，达到自己选择的精度ϵ即可。