homework 2.

2021E8014682023 何翔

1. Apply the Newton's Method with step size 1 to following problem:

$$\min_{X=(t_1,t_2)^T\in\mathbb{R}^2} t_1^2+t_2^2+t_1^4$$

with starting point $X_0=(\epsilon,\epsilon)^T$, where $\epsilon>0$ is very small, calculate the next iteration you should find that $\|X_1\|_2=O(\epsilon^3)$

解: $f(X)=t_1^2+t_2^2+t_1^4$

$$\nabla f(X)=\begin{bmatrix}2t_1+4t_1^3\\ 2t_2\end{bmatrix} \qquad \nabla^2 f(X)=\begin{bmatrix}2+12t_1^2 & 0\\ 0 & 2\end{bmatrix}$$

代入初值:

$$\nabla f(X)=\begin{bmatrix}2\epsilon+4\epsilon^3\\ 2\epsilon\end{bmatrix} \qquad \nabla^2 f(X)=\begin{bmatrix}2+12\epsilon^2 & 0\\ 0 & 2\end{bmatrix}$$

牛顿法下一次迭代:

$$X_1=X_0-(\nabla^2 f(X))^{-1}\nabla f(X)=\begin{bmatrix}\epsilon\\ \epsilon\end{bmatrix}-\frac{1}{2(2+12\epsilon^2)}\begin{bmatrix}2 & 0\\ 0 & 2+12\epsilon^2\end{bmatrix}\begin{bmatrix}2\epsilon+4\epsilon^3\\ 2\epsilon\end{bmatrix}$$

$$X_1=\begin{pmatrix}\dfrac{4\epsilon^3}{1+6\epsilon^2}\\ 0\end{pmatrix}$$

显然 $\|X_1\|_2=O(\epsilon^3)$

2. Wirte a program on Quasi-Newton method with exact line search.

$$\min_{x=(t_1,t_2)^T\in\mathbb{R}^2} t_1^2+2t_2^2$$

starting from $X_0=(1,1)^T$, Use BFGS and DFP update formula, respectively. Set the initial matrix $H_0=I$ for both methods.

解：若 $\varepsilon=1e-5$，我们手算给出前两次BFGS结果，与编程验证。

$$P_k=-H_0\nabla f_0=-\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}2\\4\end{bmatrix}=-\begin{bmatrix}2\\4\end{bmatrix}$$

精确线搜索：

$$\nabla f_k=\begin{bmatrix}2t_1\\4t_2\end{bmatrix}\qquad \nabla^2 f_k=\begin{bmatrix}2&0\\0&4\end{bmatrix}$$

$$\alpha_0=-\frac{[2\ 4]\left(-\begin{bmatrix}2\\4\end{bmatrix}\right)}{-[2\ 4]\begin{bmatrix}2&0\\0&4\end{bmatrix}\begin{bmatrix}-2\\-4\end{bmatrix}}=\frac{20}{72}=\frac{5}{18}$$

$$X_1=X_0+\alpha_0P_k=\begin{bmatrix}\frac{4}{9}\\-\frac{1}{9}\end{bmatrix}$$

$$S_1=X_1-X_0=\begin{bmatrix}-\frac{5}{9}\\-\frac{10}{9}\end{bmatrix}\qquad y_1=\nabla f_1-\nabla f_0=\begin{bmatrix}\frac{8}{9}\\-\frac{4}{9}\end{bmatrix}-\begin{bmatrix}2\\4\end{bmatrix}=\begin{bmatrix}-\frac{10}{9}\\-\frac{40}{9}\end{bmatrix}$$

$$\rho_1=\frac{1}{\frac{50}{81}+\frac{400}{81}}=\frac{9}{50}$$

$$H_{k+1}=(I-\rho_kS_ky_k^T)H_k(I-\rho_ky_kS_k^T)+\rho_kS_kS_k^T$$

代入可得 $H_1=\begin{bmatrix}-1.04&-0.13\\-0.13&0.28\end{bmatrix}$

同理，$H_2=\begin{bmatrix}\frac{1}{2}&0\\0&\frac{1}{4}\end{bmatrix}$

此时函数最小值为0。

(可看出BFGS收敛快，超线性)，

DFP仅仅是H更新公式不同。

3、Prove that if $J_k^T r_k \neq 0$, the Cauchy point of the trust region subproblem for nonlinear least-squares problem

$$\min q_k(d) = \frac{1}{2}\|J_k d + r_k\|_2^2 \quad s.t.\ \|d\|_2 \leq \Delta_k$$

satisfies:

$$q_k(0) - q_k(s_k^c) \geq \frac{1}{2}\|J_k^T r_k\| \min\left\{\frac{\|J_k^T r_k\|}{\|J_k^T J_k\|}, \Delta_k\right\}$$

证：$s_k^c = \arg\min q_k(d)$.

Lemma 4.3:

The Cauchy point s_k^c satisfies:

$$q_k(0) - q_k(s_k^c) \geq \frac{1}{2}\|g_k\| \min\left(\frac{\|g_k\|}{\|B_k\|}, \Delta_k\right).$$

使用$J_k^T J_k$代替精确Hessian

对信赖域：

$$m_k(p) = f_k + g_k^T p + \frac{1}{2}p^T B_k p \quad \|p\| \leq \Delta_k.$$

$$q_k(d) = \frac{1}{2}\|J_k d + r_k\|_2^2.$$

$$= \frac{1}{2}r_k^T r_k + r_k^T J_k d + \frac{1}{2}d^T J_k^T J_k d.$$

$$\|g_k\| = \|J_k^T r_k\| = \|\nabla f_k\| \qquad \|B_k\| = J_k^T J_k.$$

即 $$q_k(0) - q_k(s_k^c) \geq \frac{1}{2}\|\nabla f_k\| \min\left(\frac{\|J_k^T r_k\|}{\|J_k^T J_k\|}, \Delta_k\right).$$

4. Find the KKT points of the following problem.

$$\min_{X=(t_1,t_2)^T\in\mathbb{R}^2} \quad t_1+t_2$$

$$\text{s.t.} \quad 2-2t_1^2-t_2^2\geq 0, \quad t_2\geq 0.$$

解: $L(X,\lambda)=t_1+t_2-\lambda_1(2-2t_1^2-t_2^2)-\lambda_2 t_2$

由于KKT条件. $\nabla_X L(X,\lambda)=0$. 对于 $\lambda=(\lambda_1,\lambda_2)^T\geq 0$.

$$\begin{cases} 1+4\lambda_1 t_1=0 \\ 1+2\lambda_1 t_2-\lambda_2=0 \end{cases}$$

由于互补性条件.

$$\begin{cases} 2-2t_1^2-t_2^2=0 \\ t_2=0 \end{cases}$$

解得: Case I:

$$X=\begin{pmatrix}1\\0\end{pmatrix} \quad \lambda=\begin{pmatrix}-\frac{1}{4}\\1\end{pmatrix}$$

Case II:

$$X=\begin{pmatrix}-1\\0\end{pmatrix} \quad \lambda=\begin{pmatrix}\frac{1}{4}\\1\end{pmatrix}$$

要求λ所有分量均大于零,所以

$$X=\begin{pmatrix}1\\0\end{pmatrix}, \quad \lambda=\begin{pmatrix}\frac{1}{4}\\1\end{pmatrix}$$

是上述问题的KKT点,

Scanned with CamScanner

5. Write a program to apply a penalty function method to solve.

$$\min_{x=(t_1,t_2)^T\in\mathbb{R}^2} \quad t_1+t_2.$$

$$\text{s.t.} \quad t_1^2+t_2^2-2=0$$

Calculate the iterates generated in first two iterations, namely x_1, x_2.

解：罚函数方法还没有讲，作业先没有做，讲了之后会补上。