Answer Sheet 2

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

$$\min_{x=(t_1,t_2)^T\in\mathbb{R}^2} \quad t_1^2+t_2^2+t_1^4 \tag{1}$$

with starting point $x_0=(\epsilon,\epsilon)^T$ where $\epsilon>0$ is very small, calculate the next iterate and you should find that $\|x_1\|_2=O(\epsilon^3)$.

Sol. Due to the fact that the gradient is $\nabla f(x)=(2t_1+4t_1^3,2t_2)^T$, and the Hessian is $\nabla^2 f(x)=[2+12t_1^2,0;0,2]$, at the starting point $x_0=(\epsilon,\epsilon)^T$, we have

$$\nabla f_0=(2\epsilon+4\epsilon^3,2\epsilon)^T,\quad \nabla^2 f_0=[2+12\epsilon^2,0;0,2].$$

Then applying Newton's method with step size 1 yields

$$x_1=x_0-(\nabla^2 f_0)^{-1}g_0=(\frac{4\epsilon^3}{1+6\epsilon^2},0)^T.$$

Therefore, $\|x_1\|_2=O(\epsilon^3)$.

2. Write a program on quasi-Newton method with exact line search to solve the problem:

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

starting from $x_0=(1,1)^T$. Use BFGS and DFP update formula, respectively. Set the initial matrix $H_0=I$ for both methods.

3. Prove that if $J_k^T r_k\neq 0$, the Cauchy point of the trust region subproblem

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

satisfies

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

Pf. According to the definition of the Cauchy point s_k^c:

$$s_k^c=\arg\min\ q_k(d)\quad s.t.\ d=-\tau J_k^T r_k,\|d\|\leq\Delta_k,\tau\geq 0,$$

the parameter τ should satisfy

$$\tau_k=\arg\min\phi(\tau)=-\tau\|J_k^T r_k\|^2+\frac{\tau^2}{2}(J_k^T r_k)^T J_k^T J_k(J_k^T r_k),\quad s.t.\ 0\leq\tau\leq\frac{\Delta_k}{\|J_k^T r_k\|}.$$

If $J_k J_k^T r_k=0$,

$$\tau_k=\Delta_k/\|J_k^T r_k\|,\ \text{then}\ q_k(0)-q_k(s_k^c)=\Delta_k\|J_k^T r_k\|;$$

otherwise,

$$\tau_k=\min\left\{\frac{\|J_k^T r_k\|^2}{(J_k^T r_k)^T J_k^T J_k(J_k^T r_k)},\frac{\Delta_k}{\|J_k^T r_k\|}\right\}.$$

Case (i): $\tau_k = \frac{\|J_k^T r_k\|^2}{(J_k^T r_k)^T J_k^T J_k (J_k^T r_k)}$. Then,

$$q_k(0) - q_k(s_k^c) = \frac{\|J_k^T r_k\|^4}{2(J_k^T r_k)^T J_k^T J_k (J_k^T r_k)} \geq \frac{\|J_k^T r_k\|^2}{2\|J_k^T J_k\|}.$$

Case (ii): $\tau_k = \frac{\Delta_k}{\|J_k^T r_k\|}$, which means $\frac{\|J_k^T r_k\|^2}{(J_k^T r_k)^T J_k^T J_k (J_k^T r_k)} \geq \frac{\Delta_k}{\|J_k^T r_k\|}$. Then

$$q_k(0) - q_k(s_k^c) = \Delta_k \|J_k^T r_k\| - \frac{\Delta_k^2 (J_k^T r_k)^T J_k^T J_k (J_k^T r_k)}{2\|J_k^T r_k\|^2} \geq \frac{\Delta_k \|J_k^T r_k\|}{2}.$$

The proof is complete.

4. Find the KKT point(s) of the following problem:

$$\begin{aligned} \min_{x=(t_1,t_2)^T \in \mathbb{R}^2} \quad & t_1 + t_2 \\ \text{s. t.} \quad & 2 - 2t_1^2 - t_2^2 \geq 0, \quad t_2 \geq 0. \end{aligned}$$

Sol. The KKT conditions are

$$\begin{aligned} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = & \quad \lambda_1 \begin{pmatrix} -4t_1 \\ -2t_2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \lambda_1(2 - 2t_1^2 - t_2^2) = 0, & \qquad \lambda_2 t_2 = 0, \\ 2 - 2t_1^2 - t_2^2 \geq 0, & \qquad t_2 \geq 0, \\ \lambda_1 \geq 0, & \qquad \lambda_2 \geq 0. \end{aligned}$$

According to $\lambda_2 t_2 = 0$, we have the following two cases.

(i) $t_2 = 0$. Since $1 = -2\lambda_1 t_2 + \lambda_2$, $\lambda_2 = 1$. Then

$$\lambda_1(2 - 2t_1^2 - t_2^2) = 0, \quad 1 = -4\lambda_1 t_1, \quad \lambda_1 \geq 0$$

yields $t_1 = -1, \lambda_1 = \frac{1}{4}$. Therefore, $(-1, 0)^T$ is a KKT point.

(ii) $\lambda_2 = 0$. Then $1 = -2\lambda_1 t_2$ which contradicts with $t_2 \geq 0, \lambda_1 \geq 0$.

Hence, there is only one KKT point $(-1, 0)^T$.

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

$$\begin{aligned} \min_{x=(t_1,t_2)^T \in \mathbb{R}^2} \quad & t_1 + t_2 \\ \text{s. t.} \quad & t_1^2 + t_2^2 - 2 = 0. \end{aligned}$$

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