Homework 6

1. $p(t) = x_0 + x_1 t + x_2 t^2$

t	-2	-1	0	1	2
y	2	-10	0	2	1

let A be a 5×3 matrix and y be 5×1 vector

$$A = \begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \quad y = \begin{pmatrix} 2 \\ -10 \\ 0 \\ 2 \\ 1 \end{pmatrix}$$

$$A^TA = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \end{pmatrix}\begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{pmatrix}$$

$$A^Ty = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \end{pmatrix}\begin{pmatrix} 2 \\ -10 \\ 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -5 \\ 10 \\ 4 \end{pmatrix}$$

$$\det(A^TA) = 5\times10\times34 + 0 + 0 - 10\times10\times10 - 0 - 0 = 700$$

$M_{11} = 10\times34 - 0 = 340$; $M_{12} = -(0-0) = 0$; $M_{13} = 0 - 10\times10 = -100$

$M_{21} = -(0-0) = 0$; $M_{22} = 5\times34 - 10\times10 = 70$; $M_{23} = -(0-0) = 0$

$M_{31} = 0 - 10\times10 = -100$; $M_{32} = -(0-0) = 0$; $M_{33} = 5\times10 - 0 = 50$

$$(A^TA)^{-1} = \frac{1}{\det(A^TA)}\operatorname{adj}(A^TA) = \frac{1}{700}\begin{pmatrix} 340 & 0 & -100 \\ 0 & 70 & 0 \\ -100 & 0 & 50 \end{pmatrix}$$

Thus, $A^TAx = A^Ty$

$$x = (A^TA)^{-1}A^Ty = \frac{1}{700}\begin{pmatrix} 340 & 0 & -100 \\ 0 & 70 & 0 \\ -100 & 0 & 50 \end{pmatrix}\begin{pmatrix} -5 \\ 10 \\ 4 \end{pmatrix}$$

$$= \frac{1}{700}\begin{pmatrix} -2100 \\ 700 \\ 700 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix}$$

Therefore, $p(t) = -3 + t + t^2$

2. $$\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}\begin{pmatrix} r \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{cases} Ir + Ax = b \\ A^Tr + 0\cdot x = 0 \end{cases} \Rightarrow \begin{cases} Ir = b - Ax \\ A^Tr = 0 \end{cases}$$

Since $r = b - Ax$, we have $A^T(b-Ax)=0$

$A^TAx = A^Tb$. This is the solution to the least squares optimization problem of $\min\|Ax-b\|^2$

Therefore, $\min\|Ax-b\|^2$ can be obtained from solving $\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}\begin{pmatrix} r \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$

3. $f(x) = \frac{1}{2}x^TQx - c^Tx$

(a) $\nabla f(x) = \frac{\partial}{\partial x}\left(\frac{1}{2}x^TQx - c^Tx\right) = Qx - c$

$Hf(x) = Q^T = Q$

(b) $f(x) \approx f(a) + \nabla f(a)^T(x-a) + \frac{1}{2}(x-a)^T\nabla^2 f(a)(x-a)$

$\approx \frac{1}{2}a^TQa - c^Ta + (Qa-c)^T(x-a) + \frac{1}{2}(x-a)^TQ(x-a)$

(c) Since r^* is the optimal step size, we have $r^* = \arg\min_r f(x - r(\nabla f)(x))$

$f(x - r\nabla f(x)) = \frac{1}{2}(x - r\nabla f(x))^TQ(x - r\nabla f(x)) - c^T(x - r\nabla f(x))$

$= \frac{1}{2}(x^TQ - r\nabla f(x)^TQ)(x - r\nabla f(x)) - c^Tx + rc^T\nabla f(x)$

$= \frac{1}{2}(x^TQx - rx^TQ\nabla f(x) - r\nabla f(x)^TQx + r^2\nabla f(x)^TQ\nabla f(x)) - c^Tx + rc^T\nabla f(x)$

$= \left(\frac{1}{2}x^TQx - c^Tx\right) + \frac{1}{2}r^2\nabla f(x)^TQ\nabla f(x) - rx^TQ\nabla f(x) + rc^T\nabla f(x)$

$= f(x) + \frac{1}{2}r^2(Qx-c)^TQ(Qx-c) - r(x^TQ(Qx-c) + c^T(Qx-c))$

$\frac{\partial}{\partial r}f(x - r\nabla f(x)) = r(Qx-c)^TQ(Qx-c) - (x^TQQx - x^TQc + c^TQx - c^Tc) = 0$

$\Rightarrow r^* = \dfrac{x^TQQx - c^Tc}{(Qx-c)^TQ(Qx-c)} = \dfrac{(Qx-c)^T(Qx-c)}{(Qx-c)^TQ(Qx-c)}$

4. Suppose nonzero vectors $d_1, \cdots, d_n \in R^n$ are not linearly independent, then

$\exists\, \alpha_1, \cdots, \alpha_n$ not all zero, such that $\alpha_1 d_1 + \alpha_2 d_2 + \cdots + \alpha_n d_n = \sum_{j=1}^{n}\alpha_j d_j = 0$

Thus, we have $d_i^TQ\left(\sum_{j=1}^{n}\alpha_j d_j\right) = 0$

Since $d_i^TQd_j = 0$ for all i, j with $i \neq j$, we can get

$$\sum_{j=1}^{n} \alpha_j d_i^T Q d_j = \alpha_i d_i^T Q d_i = 0$$

Because Q is a $n \times n$ symmetric and positive definite matrix

$x^T Q x > 0$, $\forall x \in R^n$, $\quad d_i^T Q d_i > 0$

This means $\alpha_i = 0$ for each i

This result contradicts the assumption that $\alpha_1, \cdots, \alpha_n$ are not all zero

Therefore, $d_1, \cdots, d_n$ are linearly independent in the standard sense

5. (a) $x^* - x^0 = \alpha_0 d_0 + \cdots + \alpha_{n-1} d_{n-1} = \sum_{i=0}^{n-1} \alpha_i d_i$

$d_k^T Q (x^* - x^0) = \alpha_0 d_k^T Q d_0 + \cdots + \alpha_{n-1} d_k^T Q d_{n-1}$

Since $d_i^T Q d_j = 0$ for $i \neq j$, then we have

$$d_k^T Q (x^* - x^0) = \alpha_k d_k^T Q d_k$$

Thus, $\alpha_k = \dfrac{d_k^T Q (x^* - x^0)}{d_k^T Q d_k}$

(b) $x^{k+1} = x^k + \alpha_k d_k \Longrightarrow x^k = x^0 + \alpha_0 d_0 + \cdots + \alpha_{k-1} d_{k-1}$

Substituting $k = n$, we have $x^n = x^0 + \alpha_0 d_0 + \cdots + \alpha_{n-1} d_{n-1}$

$= x^0 + \sum_{i=0}^{n-1} \alpha_i d_i$

Since $x^* - x_0 = \sum_{i=0}^{n-1} \alpha_i d_j$

it follows that $x^* = x^n$

For $x^k = x^0 + \sum_{i=0}^{k-1} \alpha_i d_i$, apply $d_k^T Q$

$$d_k^T Q x^k = d_k^T Q x^0 + d_k^T Q \sum_{i=0}^{k-1} \alpha_i d_i$$

$$= d_k^T Q x^0 + \alpha_0 d_k^T Q d_0 + \cdots + \alpha_{k-1} d_k^T Q d_{k-1}$$

$$= d_k^T Q x^0$$

Thus, $d_k^T Q x^k = d_k^T Q x^0$

(c) $d_k^T Q x^0 = d_k^T Q x^k$

$d_k^T Q (x^* - x^0) = d_k^T Q (x^* - x^k)$

$$= d_k^T(Qx^* - Qx^k)$$

$$= d_k^T(b - Qx^k) \qquad \text{since } g_k = Qx^k - b$$

$$= d_k^T(-g_k) = -d_k^T g_k = -g_k^T d_k$$

Thus, $d_k^T Q(x^* - x^0) = -g_k^T d_k$

since $\alpha_k = \dfrac{d_k^T Q(x^* - x^0)}{d_k^T Q d_k}$ obtained in part (a)

we can get $\alpha_k = \dfrac{-g_k^T d_k}{d_k^T Q d_k}$

6. (a) $Q = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, $d_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $d_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $d_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$

$$d_0^T Q d_1 = (1\ 0\ 0)\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = (1\ 1\ 1)\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 0$$

$$d_0^T Q d_2 = (1\ 0\ 0)\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = (1\ 1\ 1)\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = 0$$

$$d_1^T Q d_2 = (1\ 0\ -1)\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = (1\ 1\ 0)\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = 0$$

Thus, d_0, d_1, d_2 are Q-conjugate

(b) $f(x) = \frac{1}{2}x^T Q x - b^T x$, $Q = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $x^0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$x^{k+1} = x^k + \alpha_k d_k$ with $\alpha_k = -\dfrac{g_k^T d_k}{d_k^T Q d_k}$ where $g_k = Qx^k - b$

$$g_0 = Qx^0 - b = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$$

$$\alpha_0 = -\frac{g_0^T d_0}{d_0^T Q d_0} = -\frac{(-1\ -1\ 0)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}{(1\ 0\ 0)\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}} = -\frac{-1}{1} = 1$$

$$x^1 = x^0 + \alpha_0 d_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$g_1 = Qx^1 - b = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

$$\alpha_1 = -\frac{g_1^T d_1}{d_1^T Q d_1} = -\frac{(0\ -1\ 0)\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}}{(1\ 0\ -1)\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}} = 0$$

$$x^2 = x^1 + \alpha_1 d_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 0\cdot\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$g_2 = Qx^2 - b = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

$$\alpha_2 = -\frac{g_2^T d_2}{d_2^T Q d_2} = -\frac{(0\ -1\ 0)\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}}{(1\ -1\ 0)\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}} = \frac{-1}{(1\ 0\ 0)\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}} = \frac{-1}{1} = -1$$

$$x^3 = x^2 + \alpha_2 d_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-1)\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$