优化方法 HW 2.

1.

$$A^TA=\begin{bmatrix}35 & 44 & 63\\ 44 & 56 & 80\\ 63 & 80 & 125\end{bmatrix} \qquad |\lambda I-A^TA|=0 \Rightarrow \lambda_{max}=205.42$$

$$\lambda_{min}=0.21$$

$$\Rightarrow \quad \mathrm{cond}(A)=\sqrt{\frac{\lambda_{max}}{\lambda_{min}}}=31.55$$

2.

$$A(X)=(X_{11}+X_{12}-X_{31}+2X_{33})\,y$$

$$\langle A(X),y\rangle = \mathrm{tr}(X^TA_0)\cdot y \qquad \cancel{A_0=\begin{bmatrix}1&0&-1\\1&0&0\\0&0&2\end{bmatrix}} \qquad A_0=\begin{bmatrix}1&1&0\\0&0&0\\-1&0&2\end{bmatrix}$$

$$=\mathrm{tr}(X^TA_0y)$$

$$=\langle X, A_0y\rangle \qquad \Rightarrow A^*(y)=A_0y=\begin{bmatrix}y&y&0\\0&0&0\\-y&0&2y\end{bmatrix}$$

3.

$$\langle A(X),Y\rangle = \mathrm{tr}([A(X)]^TY)=\mathrm{tr}\left(\begin{bmatrix}X_{11}&X_{13}&X_{22}\\X_{12}&X_{21}&X_{23}\end{bmatrix}\begin{bmatrix}Y_{11}&Y_{12}\\Y_{21}&Y_{22}\\Y_{31}&Y_{32}\end{bmatrix}\right)$$

$$=X_{11}Y_{11}+X_{13}Y_{21}+X_{22}Y_{31}+X_{12}Y_{12}+X_{21}Y_{22}+X_{23}Y_{32}.$$

$$X^T=\begin{bmatrix}X_{11}&X_{21}\\X_{12}&X_{22}\\X_{13}&X_{23}\end{bmatrix} \qquad X^T\begin{bmatrix}Y_{11}&Y_{12}&Y_{21}\\Y_{22}&Y_{31}&Y_{32}\end{bmatrix} \;//$$

$$\therefore A^*(Y)=\begin{bmatrix}Y_{11}&Y_{12}&Y_{21}\\Y_{22}&Y_{31}&Y_{32}\end{bmatrix}$$

开 闭 有界 紧

4. a. 开. 闭. 有界. 紧 $\quad C_1^\circ=\phi \quad \bar{C_1}=\phi \quad \partial C_1=\phi$

b. 开. 闭. 无界. 不紧 $\quad C_2^\circ=\mathbb{R}^n \quad \bar{C_2}=\mathbb{R}^n \quad \partial C_2=\phi.$

c. 不开. 不闭. 有界. 不紧 $\quad C_3^\circ=(0,1)\cup(2,3)\cup(4,5) \quad \bar{C_3}=[0,1]\cup[2,3]\cup[4,5]$

$$\partial C_3=\{0,1,2,3,4,5\}.$$

d. × × × × $\quad C_4^\circ=\{(x,y)^T \mid x>0, y>0\} \quad \bar{C_4}=\{(x,y)^T \mid x\ge 0, y\ge 0\}$

$$\partial C_4=\{(x,y)^T \mid \cancel{x=0 \ y=0}\} = \{(x,0)^T \mid x\ge 0\} \cup \{(0,y)^T \mid y\ge 0\}$$

e. 开 闭 紧 凸

× ✓ × × $\quad C_5^\circ=\phi \quad \bar{C}_5=C_5 \quad \partial C_5=C_5$

f. × × ✓ × $\quad C_6^\circ=\phi \quad \bar{C}_6=C_6\cup\{0\} \quad \partial C_6=C_6\cup\{0\}$

g × × ✓ × $\quad C_7^\circ=\phi \quad \bar{C}_7=C_7 \quad \partial C_7=C_7\cup\{(0,y)\mid y\in[-1,1]\}$

$\cup\{(0,y)\mid y\in[-1,1]\}$.

5. a. 极限是 $x^*=1$. $\quad e_k=5\times10^{-2k}$ $\quad r=1 \quad C=\frac{1}{100}$

b. $x^*=0 \quad e_k=x_k=2^{-2^k} \quad \frac{e_{k+1}}{e_k^2}=\frac{2^{-2^{k+1}}}{2^{-2^{k+1}}}=1$

$\therefore r=2. \quad C=1.$

c. $x^*=0 \quad e_k=x_k=3^{-k^2} \quad \frac{e_{k+1}}{e_k^r}=\frac{3^{-(k+1)^2}}{3^{-k^2\cdot r}}=3^{k^2r-(k+1)^2}$

$\Rightarrow r=1 \quad C=0$

d. $x^*=1$ $\quad$ k倍收敛更慢. $\quad \|x_k-x^*\|\le 2^{-k}=e_k$.

$$\frac{e_{k+1}}{e_k^r}=\frac{2^{-(k+1)}}{2^{-kr}}=2^{kr-(k+1)} \Rightarrow r=1 \quad C=\frac{1}{2}$$

6. a. $\frac{\partial f(\vec{x})}{\partial x_i}= \cancel{p\,\|\vec{x}\|_p}\cdot\frac{1}{p}\left(\sum|x_i|^p\right)^{\frac{1}{p}-1}\cdot p|x_i|^{p-1}\operatorname{sign}(x_i)$

$$\nabla f(x)=\left(\sum|x_i|^p\right)^{\frac{1}{p}-1}\cdot\left(|x_i|^{p-1}\operatorname{sign}(x_i)\right)_{n\times1}.$$

$$=\left(\left(\frac{|x_j|}{\|\vec{x}\|_p}\right)^{p-1}\cdot\operatorname{sgn}(x_j)\right)_{n\times1}$$

b.

b. $\frac{df(x)}{dx} = \cancel{ab^Tx + (a^Tx)\cdot b} \; a^T\cdot(b^Tx) + (a^Tx)\cdot b^T.$

$\nabla f(x) = (b^Tx)a + (a^Tx)b$

$\nabla^2 f(x) = ab^T + ba^T$

c. $\nabla f(x) = A^T(Ax-b)$

$\nabla^2 f(x) = A^TA$

d. $\nabla f(x) = R^T \nabla(\overset{h(Rx)}{\cancel{u^Tg(Rx)}})$

$h(Rx) = f(x) = u^Tg(Rx) \qquad h(x) = u^Tg(x).$

$\nabla h(x) = u\cdot \nabla g(x) = u \odot g'(x) \qquad g'(x) = (g_0'(x_1), \cdots g_0'(x_n))^T.$

$\Rightarrow \nabla f(x) = R^Tu \odot g'(Rx)$ 　　⊙ 逐元素乘.

$\nabla^2 f(x) = R^Tu \, g''(Rx) R.$ 　　$g''(x) = (g_0''(x_i))^T$

e. $f(x) = \frac{1}{2} tr((xx^T-I)^T(xx^T-I))$

$df(x) = \frac{1}{2} d\,tr(\quad) = \frac{1}{2} tr(d((xx^T-I)^T(xx^T-I)))$

$= \frac{1}{2} tr[(d(xx^T-I)^T)\cdot(xx^T-I) + (xx^T-I)^T d(xx^T-I)] = tr(d(xx^T-I)^T d(xx^T-I)))$

$tr(d(xx^T-I)) = tr(dx\cdot x^T + x\,dx^T) = tr(2x\,dx^T) = 2tr(x^T dx)$

$\Rightarrow \nabla f(x) = \cancel{2x(xx^T-I)^T} \; 2((xx^T-I)^Tx^T)^T = 2x(x^Tx-I).$

$\nabla^2 f(x) = 2(x^Tx-I) + 4xx^T$

f. $f(X) = tr((X^TAX)^T (X^TAX))$

$\cancel{df(X) = d\,tr(X^TAX) = tr(dX^TAX)}$

$\cancel{= tr((dX^T)AX + X^TAdX)}$

$= \cancel{tr(dX^TA^TdX + X^TAdX)}$

$= \cancel{2\,tr(X^TAdX)}.$

$\cancel{\Rightarrow \nabla f(X) = 2X^TA}$ $\quad df(X) = tr(d(X^TAX)^T(X^TAX) + (X^TAX)^T d(X^TAX))$

$\cancel{\nabla^2 f(X) =}$ $\quad d(X^TAX) = dX^T \cdot AX + X^TA\,dX.$

代入有 $df(X) = tr(2(X^TAX)X^TA\,dX).$

$\Rightarrow \nabla f(X) = (2X^TAXX^TA)^T = 2AXX^TAX$

$\cancel{\nabla^2 f(X) = d\nabla f(X)_{(k,l)} = tr(d(AXX^TAX)_{(k,l)}) =}$

$\cancel{\nabla^2 f(X) -}$ $\quad d\nabla f(X) = 2A\,dX(X^TAX) + 2AX[(dX)^TAX + X^TA\,dX].$

$tr(XY) = tr(YX).$

g. $df(X) = d\,tr(AXB) = tr(dAXB) = tr(A\,dX\,B) = tr(BA\,dx)$

$\Rightarrow \nabla f(X) = BA$

$\nabla^2 f(X) = 0$ (矩阵)