MATH402 - Homework #2 Sample Solution Key

Q1 The Sherman-Morrison-Woodbury formula

$$\tilde{A} = A + ab^T, \qquad \tilde{A}^{-1} = A^{-1} - \frac{A^{-1}ab^TA^{-1}}{1+b^TA^{-1}a}$$

$$\tilde{A} = B_{k+1} = \underbrace{B_k}_{A} + \underbrace{\frac{(y_k - B_k s_k)}{(y_k - B_k s_k)^T s_k}}_{a} \underbrace{(y_k - B_k s_k)^T}_{b^T}$$

Applying the Sherman-Morrison-Woodbury formula

$$(B_{k+1})^{-1} = (B_k)^{-1} - \frac{(B_k)^{-1} a b^T (B_k)^{-1}}{1 + b^T (B_k)^{-1} a}$$

$$(B_{k+1})^{-1} = H_{k+1}, \qquad (B_k)^{-1} = H_k$$

$$H_{k+1} = H_k - \frac{H_k \frac{(y_k - B_k s_k)}{(y_k - B_k s_k)^T s_k} (y_k - B_k s_k)^T H_k}{1 + (y_k - B_k s_k)^T (H_k) \frac{(y_k - B_k s_k)}{(y_k - B_k s_k)^T s_k}}$$

Note that $H_k^T = H_k$, $B_k^T H_k = B_k^T H_k^T = I$

$$H_{k+1} = H_k - \frac{\frac{H_k (y_k - B_k s_k)(y_k - B_k s_k)^T H_k}{\cancel{(y_k - B_k s_k)^T s_k}}}{\frac{(y_k - B_k s_k)^T s_k + (y_k - B_k s_k)^T (H_k)(y_k - B_k s_k)}{\cancel{(y_k - B_k s_k)^T s_k}}}$$

Note that $(y_k - B_k s_k)^T s_k$ is scalar!

$$= H_k - \frac{(H_k y_k - \overbrace{H_k B_k}^{I} s_k)(y_k^T H_k - s_k^T \overbrace{B_k^T H_k}^{I})}{(y_k - B_k s_k)^T (s_k + H_k y_k - \underbrace{H_k B_k}_{I} s_k)}$$

$$= H_k - \frac{(H_k y_k - s_k)(y_k^T H_k - s_k^T)}{(y_k - B_k s_k)^T (H_k y_k)}$$

$$= H_k + \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{-y_k^T H_k y_k + s_k^T \underbrace{B_k^T H_k}_{I} y_k} = \boxed{H_k + \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{(s_k - H_k y_k)^T y_k}}$$

Q2| a) $f \in C^1(\mathbb{R}^n)$ has Lipschitz continuous gradient with Lipschitz constant L. Then, we have

$$\|\nabla f(y) - \nabla f(x)\| \leqslant L\|y-x\|, \quad L>0, \ \forall x,y$$

We know from calculus

$$H(b) - H(a) = \int_a^b h(Q)\, dQ. \qquad (1)$$

Let $h(\alpha) = \langle \nabla f(x+\alpha(y-x)), y-x \rangle$ be a function in α and $dQ = d\alpha$. By using $H(b) = f(y)$, $H(a) = f(x)$, Equation (1), and definite integral of $h(\alpha)$ from 0 to 1, we obtain

$$f(y) - f(x) = \int_0^1 \langle \nabla f(x+\alpha(y-x)), y-x \rangle d\alpha,$$

$$= \int_0^1 \langle \nabla f(x+\alpha(y-x)) - \nabla f(x) + \nabla f(x), y-x \rangle d\alpha,$$

α value is independent of $\nabla f(x)$, therefore we can obtain

$$f(y) - f(x) = \langle \nabla f(x), y-x \rangle + \int_0^1 \langle \nabla f(x+\alpha(y-x)) - \nabla f(x), y-x \rangle d\alpha,$$

$$|f(y) - f(x) - \langle \nabla f(x), y-x \rangle| = \left| \int_0^1 \langle \nabla f(x+\alpha(y-x)) - \nabla f(x), y-x \rangle d\alpha \right|,$$

$$\leqslant \int_0^1 |\langle \nabla f(x+\alpha(y-x)) - \nabla f(x), y-x \rangle| d\alpha,$$

By using Cauchy Schwarz

$$\leqslant \int_0^1 \|\nabla f(x+\alpha(y-x)) - \nabla f(x)\| \cdot \|y-x\| d\alpha,$$

We can obtain

$$\|\nabla f(x+\alpha(y-x)) - \nabla f(x)\| \leqslant L\|\alpha(y-x)\| \leqslant L|\alpha|\|y-x\| = L\alpha\|y-x\|,$$

by using Lipschitz continuous gradient. The range of integral is from 0 to 1, therefore absolute sign in α can be removed.

$$|f(y) - f(x) - \langle \nabla f(x), y-x \rangle| \leqslant \|y-x\|^2 \int_0^1 L\alpha\, d\alpha = \frac{L}{2}\|y-x\|^2$$

If we use $p = y-x$, then we have

$$|f(x+p)-f(x)-\nabla f^T(x)p| \leq \frac{L}{2}\|p\|^2$$

$$f(x+p) \leq f(x)+\nabla f^T(x)p+\frac{L}{2}\|p\|^2, \quad \forall x,p \in \mathbb{R}^n$$

$\blacksquare$

b) Let $p=-\alpha \nabla f(x)$, where $\alpha>0$. Then we can use the inequality given in part a.

$$f(x-\alpha \nabla f(x)) \leq f(x)-\alpha \nabla f^T(x)\nabla f(x)+\frac{L}{2}\|\alpha \nabla f(x)\|^2$$
$$\leq f(x)-\alpha \nabla f^T(x)\nabla f(x)+\frac{L\alpha^2}{2}\|\nabla f(x)\|^2$$
$$=f(x)-\alpha\langle \nabla f(x), \nabla f(x)\rangle+\frac{L\alpha^2}{2}\|\nabla f(x)\|^2$$
$$=f(x)-\alpha\|\nabla f(x)\|^2+\frac{L\alpha^2}{2}\|\nabla f(x)\|^2$$

$$\Rightarrow f(x-\alpha\nabla f(x))-f(x) \leq \alpha\left(\frac{L}{2}\alpha-1\right)\|\nabla f(x)\|^2$$

$$f(x)-f(x-\alpha\nabla f(x)) \geq \alpha\left(1-\frac{L}{2}\alpha\right)\|\nabla f(x)\|^2$$

$\blacksquare$

c) $p_k=-\nabla f(x_k)$ in the steepest descent method. Then,

$x_{k+1}=x_k-\alpha\nabla f(x_k)$

We know that

$$f(x_k)-f(x_{k+1}) \geq \alpha\left(1-\frac{L}{2}\alpha\right)\|\nabla f(x_k)\|^2 \quad (*)$$

By steepest descent with exact line search method

$x_{k+1}=x_k-\alpha_k\nabla f(x_k)$,

where $\alpha_k \in \underset{\alpha \geq 0}{\operatorname{argmin}} f(x_k-\alpha\nabla f(x_k))$

We know that $\alpha \in \left(0,\frac{2}{L}\right)$ from $(*)$ and $\alpha\left(1-\frac{L}{2}\alpha\right)>0$

$\alpha \in \left(0,\frac{2}{L}\right) \Rightarrow$ guaranteed descent

To obtain maximum value of $\alpha\left(1-\frac{L}{2}\alpha\right)$, we can use

$$\frac{\partial}{\partial\alpha}\left(\alpha-\frac{L}{2}\alpha^2\right)=1-\alpha L=0 \Rightarrow \alpha L=1 \Rightarrow \boxed{\alpha=1/L}$$

By definition of α_k and $(*)$,

$$f\left(x_k-\alpha_k\nabla f(x_k)\right)\leqslant f\left(x_k-\frac{1}{L}\nabla f(x_k)\right)$$

Thus,

$$f(x_k)-f(x_{k+1})\geqslant f(x_k)-f\left(x_k-\frac{1}{L}\nabla f(x_k)\right)$$
$$\geqslant \frac{1}{L}\left(1-\frac{L}{2}\cdot\frac{1}{L}\right)\|\nabla f(x_k)\|^2=\frac{1}{2L}\|\nabla f(x_k)\|^2 \quad \blacksquare$$

d) i) We know that

$$f(x_k)-f(x_{k+1})\geqslant\frac{1}{2L}\|\nabla f(x_k)\|^2,$$

from the part c. Then, we have

$$f(x_k)\geqslant f(x_{k+1})+\frac{1}{2L}\|\nabla f(x_k)\|^2$$

We know that $L>0$. Then, we can obtain

$$f(x_k)\geqslant f(x_{k+1}),$$

which means that $\{f(x_k)\}$ sequence is nonincreasing. The only condition for $f(x_k)=f(x_{k+1})$ is $\nabla f(x_k)=0$. Therefore, we can obtain

$$f(x_{k+1})<f(x_k) \text{ for any } k\geqslant 0 \text{ unless } \nabla f(x_k)=0. \quad \blacksquare$$

ii) We know that

$$f(x_k)-f(x_{k+1})\geqslant\frac{1}{2L}\|\nabla f(x_k)\|^2.$$

$$\lim_{k\to\infty} f(x_k)=f(x^*)$$

If $x_k=x^*$, then $x_{k+1}=x^* \Longrightarrow \lim_{k\to\infty} f(x_k)-f(x_{k+1})=0$

Continued

As $k \to \infty$, we have

$$0 \geq \frac{1}{2L} \|\nabla f(x_k)\|^2 \quad \text{and} \quad L > 0,$$

$$\Rightarrow 0 \geq \|\nabla f(x_k)\|^2$$

$$\Rightarrow \nabla f(x_k) \to 0 \quad \text{as} \quad k \to \infty \quad \blacksquare$$

Q3] a) For $i=j$, we have

$p_i^T A p_j = p_i^T A p_i > 0$ (A is symmetric positive definite)

Now, we should check $i \neq j$ case. We know that

$$p_j = \alpha_0 p_0 + \alpha_1 p_1 + \dots + \alpha_{j-1} p_{j-1}$$

$$p_i^T A p_j = p_i^T A(\alpha_0 p_0 + \alpha_1 p_1 + \dots + \alpha_{j-1} p_{j-1})$$

$$= \alpha_0 \underbrace{p_i^T A p_0}_{=0} + \alpha_1 \underbrace{p_i^T A p_1}_{=0} + \dots + \alpha_{j-1} \underbrace{p_i^T A p_{j-1}}_{=0} = 0 \quad (\text{for } i > j)$$

$\to$ since p_i's are A conjugate.

$$p_i^T A p_j = p_i^T A(\alpha_0 p_0 + \alpha_1 p_1 + \dots + \alpha_i p_i + \dots + \alpha_{j-1} p_{j-1})$$

$$= \alpha_0 \underbrace{p_i^T A p_0}_{=0} + \alpha_1 \underbrace{p_i^T A p_1}_{=0} + \dots + \alpha_i \underbrace{p_i^T A p_i}_{>0} + \dots + \alpha_{j-1} \underbrace{p_i^T A p_{j-1}}_{=0} > 0 \quad (\text{for } i < j)$$

As a result, we can obtain $p_i^T A p_j \geq 0 \ \forall i,j$ $\blacksquare$

b) $b = Ax_i - r_i \qquad \alpha_i = \dfrac{b^T p_i}{p_i^T A p_i}$

$$\alpha_i = \frac{(Ax_i - r_i)^T p_i}{p_i^T A p_i} = \frac{x_i^T A p_i - r_i^T p_i}{p_i^T A p_i}$$

($A^T = A$ since A is symmetric)

$$x_i = x_0 + \alpha_0 p_0 + \alpha_1 p_1 + \dots + \alpha_{i-1} p_{i-1}$$

$$x_i^T A p_i = (x_0 + \alpha_0 p_0 + \alpha_1 p_1 + \dots + \alpha_{i-1} p_{i-1})^T A p_i$$

$$= x_0^T A p_i + \alpha_0 \underbrace{p_0^T A p_i}_{=0} + \dots + \alpha_{i-1} \underbrace{p_{i-1}^T A p_i}_{=0}$$

$= x_0^T A p_i = 0$ (since $x_0 \in \mathbb{R}^n$ is any arbitrary starting point we can choose $x_0 = [0, 0, \dots, 0]^T$)

Continued $\to$

Then, we have

$$\alpha_i = -\frac{r_i^T p_i}{p_i^T A p_i} = -\frac{r_i^T \overbrace{(-r_i+\beta_{i-1} p_{i-1})}^{= p_i \text{ (algorithm from lecture notes)}}}{p_i^T A p_i} = \frac{r_i^T r_i - r_i^T \beta_{i-1} p_{i-1}}{p_i^T A p_i}$$

We know that the residuals are orthogonal to directions:

$r_i^T p_j = 0 \Rightarrow$ Then, $\beta_{i-1} r_i^T p_{i-1} = 0$

As a result, we obtain

$$\alpha_i = \frac{r_i^T r_i}{p_i^T A p_i}$$

■

Q4) min $x_1 x_2$ s.t. $4x_1^2 + x_2^2 - 4 = 0$

a) $L(x,\lambda) = f(x) - \sum_{j=1}^{m} \lambda_j c_j(x)$

$$L(x,\lambda) = x_1 x_2 - \lambda_1 (4x_1^2 + x_2^2 - 4)$$

$$\nabla L(x,\lambda) = \begin{bmatrix} x_2 - 8\lambda_1 x_1 \\ x_1 - 2\lambda_1 x_2 \\ 4 - 4x_1^2 - x_2^2 \end{bmatrix}$$

b) $\nabla L(x,\lambda) = 0$ for stationary points

$$\left.\begin{aligned} x_2 - 8\lambda_1 x_1 &= 0 \\ x_1 - 2\lambda_1 x_2 &= 0 \\ 4 - 4x_1^2 - x_2^2 &= 0 \end{aligned}\right\} \quad \left.\begin{aligned} x_2 &= 8\lambda_1 x_1 \\ x_1 &= 2\lambda_1 x_2 \end{aligned}\right\} \quad \begin{aligned} &x_1 = 0 \ x_2 = 0 \quad 4 - 4x_1^2 - x_2^2 \neq 0 \text{ (not possible)} \\ &x_1 = 16\lambda_1^2 x_1 \Rightarrow \boxed{\lambda_1 = \frac{1}{4}} \text{ or } \boxed{\lambda_1 = -\frac{1}{4}} \end{aligned}$$

for $\lambda_1 = \frac{1}{4} \Rightarrow x_2 = 2x_1 \Rightarrow 4 - 4x_1^2 - 4x_1^2 = 0 \Rightarrow x_1^2 = \frac{1}{2} \Rightarrow \boxed{\begin{aligned} x_1 &= -\tfrac{1}{\sqrt{2}} \\ x_2 &= -\sqrt{2} \end{aligned}}$ or $\boxed{\begin{aligned} x_1 &= \tfrac{1}{\sqrt{2}} \\ x_2 &= \sqrt{2} \end{aligned}}$

for $\lambda_1 = -\frac{1}{4} \Rightarrow x_2 = -2x_1 \Rightarrow 4 - 4x_1^2 - 4x_1^2 = 0 \Rightarrow x_1^2 = \frac{1}{2} \Rightarrow \boxed{\begin{aligned} x_1 &= -\tfrac{1}{\sqrt{2}} \\ x_2 &= \sqrt{2} \end{aligned}}$ or $\boxed{\begin{aligned} x_1 &= \tfrac{1}{\sqrt{2}} \\ x_2 &= -\sqrt{2} \end{aligned}}$

Then, we have 4 stationary points such that

$$(\bar{x_1}, \bar{x_2}, \bar{\lambda_1})_1 = \left(-\frac{1}{\sqrt{2}}, -\frac{2}{\sqrt{2}}, \frac{1}{4}\right)$$

$$(\bar{x_1}, \bar{x_2}, \bar{\lambda_1})_2 = \left(\frac{1}{\sqrt{2}}, \frac{2}{\sqrt{2}}, \frac{1}{4}\right)$$

$$(\bar{x_1}, \bar{x_2}, \bar{\lambda_1})_3 = \left(-\frac{1}{\sqrt{2}}, \frac{2}{\sqrt{2}}, -\frac{1}{4}\right)$$

$$(\bar{x_1}, \bar{x_2}, \bar{\lambda_1})_4 = \left(\frac{1}{\sqrt{2}}, -\frac{2}{\sqrt{2}}, -\frac{1}{4}\right)$$

c) $$\nabla^2 L(x,\lambda) = \begin{bmatrix} -8\lambda_1 & 1 & -8x_1 \\ 1 & -2\lambda_1 & -2x_2 \\ -8x_1 & -2x_2 & 0 \end{bmatrix}$$

$$\nabla^2 L(x_0,\lambda) = \begin{bmatrix} -8 & 1 & 0 \\ 1 & -2 & 2 \\ 0 & 2 & 0 \end{bmatrix} \qquad \nabla L(x_0,\lambda) = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}$$

$$\nabla^2 L(x_0,\lambda) \begin{bmatrix} \Delta x_k \\ \Delta \lambda_k \end{bmatrix} = -\nabla L(x_0,\lambda)$$

$$\begin{bmatrix} -8 & 1 & 0 \\ 1 & -2 & 2 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta \lambda_1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -3 \end{bmatrix} \Rightarrow$$

$$2\Delta x_2 = -3 \quad \boxed{\Delta x_2 = -\frac{3}{2}}$$

$$-8\Delta x_1 + \Delta x_2 = 1$$

$$-8\Delta x_1 = \frac{5}{2} \Rightarrow \boxed{\Delta x_1 = -\frac{5}{16}}$$

$$\Delta x_1 - 2\Delta x_2 + 2\Delta\lambda_1 = -2$$

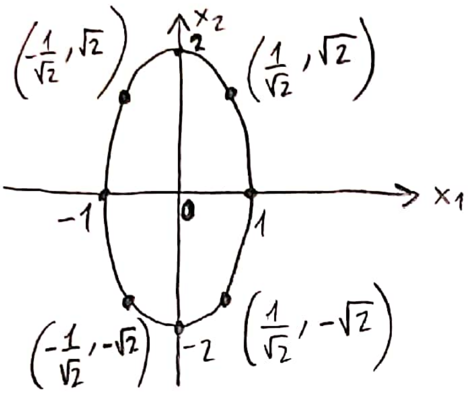
$$2\Delta\lambda_1 = -5 + \frac{5}{16}$$

$$\boxed{\Delta\lambda_1 = -\frac{75}{32}}$$

Continued →

$$\begin{bmatrix} x_1^1 \\ x_1^2 \\ \lambda_1^1 \end{bmatrix} = \begin{bmatrix} x_0^1 \\ x_0^2 \\ \lambda_0^1 \end{bmatrix} + 1. \begin{bmatrix} -5/16 \\ -3/2 \\ -75/32 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} - \begin{bmatrix} 5/16 \\ 3/2 \\ 75/32 \end{bmatrix} = \begin{bmatrix} -5/16 \\ -5/2 \\ -43/32 \end{bmatrix} = \begin{bmatrix} -0.3125 \\ -2.5 \\ -1.34375 \end{bmatrix}$$

d) $c(x) = 4 - 4x_1^2 - x_2^2$



$4 - 4x_1^2 - x_2^2 = 0$

$c(x(\alpha)) = 0$, take $x_1 = \mp\alpha$, then

$4 - 4\alpha^2 - x_2^2 = 0 \Rightarrow x_2 = \mp\sqrt{4-4\alpha^2} = \mp 2\sqrt{(1-\alpha^2)}$

We have 4 feasible paths:

$x_1(\alpha) = (\alpha, 2\sqrt{(1-\alpha^2)})$

$x_2(\alpha) = (\alpha, -2\sqrt{(1-\alpha^2)})$

$x_3(\alpha) = (-\alpha, 2\sqrt{(1-\alpha^2)})$

$x_4(\alpha) = (-\alpha, -2\sqrt{(1-\alpha^2)})$

We can obtain the derivatives wrt α:

$x_1'(\alpha) = \left(1, -\frac{2\alpha}{\sqrt{1-\alpha^2}}\right) \neq (0,0)$

$x_2'(\alpha) = \left(1, \frac{2\alpha}{\sqrt{1-\alpha^2}}\right) \neq (0,0)$

$x_3'(\alpha) = \left(-1, -\frac{2\alpha}{\sqrt{1-\alpha^2}}\right) \neq (0,0)$

$x_4'(\alpha) = \left(-1, \frac{2\alpha}{\sqrt{1-\alpha^2}}\right) \neq (0,0)$

For the 1^{st} stationary point $\left(-\frac{1}{\sqrt{2}}, -\frac{2}{\sqrt{2}}\right)$:

$$\left.\begin{aligned} x_1'\left(-\frac{1}{\sqrt{2}}\right) &= (1, 2) \\ x_2'\left(-\frac{1}{\sqrt{2}}\right) &= (1,-2) \\ x_3'\left(-\frac{1}{\sqrt{2}}\right) &= (-1, 2) \\ x_4'\left(-\frac{1}{\sqrt{2}}\right) &= (-1,-2) \end{aligned}\right\}$$

We can obtain the tangent cone as:

$$T^{o}(x_1^*) = T^{o}\left(-\frac{1}{\sqrt{2}}, -\frac{2}{\sqrt{2}}\right) = \left\{\alpha \begin{bmatrix} 1 \\ -2 \end{bmatrix} : \alpha \in \mathbb{R}\right\}.$$

It can be observed from the graph shown in the previous page.

For the 2^{nd} stationary point $\left(\frac{1}{\sqrt{2}}, \frac{2}{\sqrt{2}}\right)$:

$$\left.\begin{aligned} x_1'\left(\frac{1}{\sqrt{2}}\right) &= (1,-2) \\ x_2'\left(\frac{1}{\sqrt{2}}\right) &= (1, 2) \\ x_3'\left(\frac{1}{\sqrt{2}}\right) &= (-1,-2) \\ x_4'\left(\frac{1}{\sqrt{2}}\right) &= (-1, 2) \end{aligned}\right\}$$

We can obtain the tangent cone as:

$$T^{o}(x_2^*) = T^{o}\left(\frac{1}{\sqrt{2}}, \frac{2}{\sqrt{2}}\right) = \left\{\alpha \begin{bmatrix} 1 \\ -2 \end{bmatrix} : \alpha \in \mathbb{R}\right\}$$

For the 3^{rd} stationary point $\left(-\frac{1}{\sqrt{2}}, \frac{2}{\sqrt{2}}\right)$:

$$\left.\begin{aligned} x_1'\left(-\frac{1}{\sqrt{2}}\right) &= (1, 2) \\ x_2'\left(-\frac{1}{\sqrt{2}}\right) &= (1,-2) \\ x_3'\left(-\frac{1}{\sqrt{2}}\right) &= (-1, 2) \\ x_4'\left(-\frac{1}{\sqrt{2}}\right) &= (-1,-2) \end{aligned}\right\}$$

We can obtain the tangent cone as:

$$T^{o}(x_3^*) = T^{o}\left(-\frac{1}{\sqrt{2}}, \frac{2}{\sqrt{2}}\right) = \left\{\alpha \begin{bmatrix} 1 \\ 2 \end{bmatrix} : \alpha \in \mathbb{R}\right\}$$

For the 4^{th} stationary point $\left(\frac{1}{\sqrt{2}}, -\frac{2}{\sqrt{2}}\right)$:

$$\left.\begin{aligned} x_1'\left(\frac{1}{\sqrt{2}}\right) &= (1,-2) \\ x_2'\left(\frac{1}{\sqrt{2}}\right) &= (1,2) \\ x_3'\left(\frac{1}{\sqrt{2}}\right) &= (-1,-2) \\ x_4'\left(\frac{1}{\sqrt{2}}\right) &= (-1,2) \end{aligned}\right\}$$

We can obtain the tangent cone as:

$$T^0(x_4^*) = T^0\left(\frac{1}{\sqrt{2}}, -\frac{2}{\sqrt{2}}\right) = \left\{ \alpha \begin{bmatrix} 1 \\ 2 \end{bmatrix} : \alpha \in \mathbb{R} \right\}$$

Now, we can check the null spaces.

$$J(x^*) = \begin{bmatrix} 8x_1 & 2x_2 \end{bmatrix}$$

$$\begin{aligned} Null(J(x_1^*)) &= \{ p \in \mathbb{R}^2 : J(x_1^*)p = 0 \} \\ &= \left\{ p \in \mathbb{R}^2 : \begin{bmatrix} -8/\sqrt{2} & -4/\sqrt{2} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = 0 \right\} \\ &= \left\{ c \begin{bmatrix} 1 \\ -2 \end{bmatrix} : c \in \mathbb{R} \right\} \end{aligned}$$

$$\begin{aligned} Null(J(x_2^*)) &= \{ p \in \mathbb{R}^2 : J(x_2^*)p = 0 \} \\ &= \left\{ p \in \mathbb{R}^2 : \begin{bmatrix} 8/\sqrt{2} & 4/\sqrt{2} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = 0 \right\} \\ &= \left\{ c \begin{bmatrix} 1 \\ -2 \end{bmatrix} : c \in \mathbb{R} \right\} \end{aligned}$$

$$\begin{aligned} Null(J(x_3^*)) &= \{ p \in \mathbb{R}^2 : J(x_3^*)p = 0 \} \\ &= \left\{ p \in \mathbb{R}^2 : \begin{bmatrix} -8/\sqrt{2} & 4/\sqrt{2} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = 0 \right\} \\ &= \left\{ c \begin{bmatrix} 1 \\ 2 \end{bmatrix} : c \in \mathbb{R} \right\} \end{aligned}$$

$$\mathrm{Null}(J(x_4^*)) = \{ p \in \mathbb{R}^2 : J(x_4^*)p = 0 \}$$

$$= \left\{ p \in \mathbb{R}^2 : \begin{bmatrix} 8/\sqrt{2} & -4/\sqrt{2} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = 0 \right\}$$

$$= \left\{ c \begin{bmatrix} 1 \\ 2 \end{bmatrix} : c \in \mathbb{R} \right\}$$