HW03 Solution.

1) Suppose $P \in \mathbb{R}^{m\times m}$ is a projector.

Show that $\mathrm{Null}(I-P) = \mathrm{range}(P)$.

Sol. 1) $\mathrm{Null}(I-P) \subset \mathrm{range}(P)$.

Let $x \in \mathrm{Null}(I-P)$. Then $(I-P)x = 0$.

$$x - Px = 0$$

$$x = Px.$$

$$\Rightarrow \quad x \in \mathrm{range}(P).$$

2) $\mathrm{range}(P) \subset \mathrm{Null}(I-P)$.

Let $x \in \mathrm{range}(P)$. Then $Px = x$ as P is a projector.

$$\Rightarrow \quad x - Px = 0.$$

$$(I-P)x = 0.$$

$$\Rightarrow \quad x \in \mathrm{Null}(I-P).$$

2) First, find

$$A^TA = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}.$$

$$\Rightarrow \quad A^TA = \frac{1}{10-4}\begin{bmatrix} 5 & -2 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} 5/6 & -2/6 \\ -2/6 & 2/6 \end{bmatrix}.$$

Then,

$$P = A(A^TA)^{-1}A^T$$

$$= \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 5/6 & -2/6 \\ -2/6 & 2/6 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}.$$

$$= \begin{bmatrix} 1/6 & 2/6 \\ -2/6 & 2/6 \\ +5/6 & -2/6 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5/6 & 2/6 & 1/6 \\ 2/6 & 2/6 & -2/6 \\ 1/6 & -2/6 & 5/6 \end{bmatrix}$$

②

$$P\begin{bmatrix}1\\2\\3\end{bmatrix} = \frac{1}{6}\begin{bmatrix}5 & 2 & 1\\2 & 2 & -2\\1 & -2 & 5\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix} = \frac{1}{6}\begin{bmatrix}12\\0\\12\end{bmatrix} = \begin{bmatrix}2\\0\\2\end{bmatrix}.$$

3) $P=\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$.

Observe that the x-axis is spamed by the vector $q=\begin{bmatrix}1\\0\end{bmatrix}$.

$\Rightarrow$ The orthogonal projection ontol the x-axis is:

$$\begin{bmatrix}1\\0\end{bmatrix}\begin{bmatrix}1 & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} = P.$$

The complementary of P is

$$I-P=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} - \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} = \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}.$$

The y-axis is spanned by $\begin{bmatrix}0\\1\end{bmatrix}$, and the orthogonal projection onto y-axis is

$$\begin{bmatrix}0\\1\end{bmatrix}\begin{bmatrix}0 & 1\end{bmatrix} = \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}.$$

4) Suppose A is ~~inter~~ invertible. (A^TA is symmetric).

Take any vector $x \neq 0$. Then

$$x^T(A^TA)x = (Ax)^T(Ax). = \|Ax\|_2^2 > 0$$

since $Ax \neq 0$.

$\therefore$ A^TA is positive definite.

5) Consider any vector $\begin{bmatrix}x\\y\\z\end{bmatrix} \neq \vec{0}$ in $\mathbb{R}^3$. Then

$$\begin{bmatrix}x & y & z\end{bmatrix}\begin{bmatrix}2 & -1 & 0\\-1 & 2 & -1\\0 & -1 & 2\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}2x-y & -x+2y-z & -y+2z\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}$$

$$= 2x^2-xy-xy+2y^2-yz-zy+2z^2$$

$$= 2x^2-2xy+2y^2-2yz+2z^2$$

M is p.d. $\Leftarrow$ $= x^2+(x-y)^2+(y-z)^2+z^2 > 0.$

③.

6) Let $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$. Then $V = \mathrm{Col}(A)$.

a) $A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}$

$\Rightarrow (A^T A)^{-1} = \frac{1}{6} \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix}$.

The orthogonal projector onto V is

$$P = A(A^T A)^{-1} A^T$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \frac{1}{6} \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

$$= \frac{1}{6} \begin{bmatrix} 5 & -3 \\ 2 & 0 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

$$= \frac{1}{6} \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}$$

b) $P \begin{bmatrix} 8 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} 8 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 40 \\ 16 \\ -8 \end{bmatrix}$

c) $\begin{bmatrix} 8 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{6} \begin{bmatrix} 40 \\ 16 \\ -8 \end{bmatrix} = \begin{bmatrix} 8/6 \\ -16/6 \\ 8/6 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 8 \\ -16 \\ 8 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 4 \\ -8 \\ 4 \end{bmatrix}$.

$\Rightarrow \left\| \frac{1}{3} \begin{bmatrix} 4 \\ -8 \\ 4 \end{bmatrix} \right\|_2 = \frac{1}{3} \left\| \begin{bmatrix} 4 \\ -8 \\ 4 \end{bmatrix} \right\|_2 = \frac{1}{3} \sqrt{4^2 + 8^2 + 4^2}$

$= \frac{1}{3} \sqrt{96}$

7) $X = \begin{bmatrix} -2 & 0 & 2 \\ -2 & 0 & 2 \end{bmatrix}$.

a) The first principal component is $v = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

(You don't ~~have to~~ really need to solve any eigenproblem to see it. Why?).

b) The ~~three~~ coordinates of the three points ~~of~~ after projection are:

$$\left\langle \begin{bmatrix} -2 \\ -2 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \right\rangle = -\frac{2}{\sqrt{2}} - \frac{2}{\sqrt{2}} = -\frac{4}{\sqrt{2}} = -2\sqrt{2}$$

$$\left\langle \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \right\rangle = 0.$$

and $\left\langle \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \right\rangle = 2\sqrt{2}$

These coordinates have mean 0.

$\Rightarrow$ their variance $= \frac{1}{3}\left[(-2\sqrt{2})^2 + (0)^2 + (2\sqrt{2})^2\right]$.

$$= \frac{16}{3}.$$

c) The reconstruction errors are 0 since all three points are located on the direction of the first principal component.

8) Let $x_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $x_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, and $x_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$.

$$f(\alpha_1, \alpha_2) = \sum_{i=1}^{3} \| x_i - \alpha_1 b_1 - \alpha_2 b_2 \|_2^2 .$$

Let $y_i = x_i - \alpha_1 b_1 - \alpha_2 b_2$. Then $f = \sum_{i=1}^{3} \|y_i\|_2^2$

$$\frac{\partial f}{\partial \alpha_1} = \frac{\partial f}{\partial y_1} \cdot \frac{\partial y_1}{\partial \alpha_1} + \frac{\partial f}{\partial y_2} \frac{\partial y_2}{\partial \alpha_1} + \frac{\partial f}{\partial y_3} \cdot \frac{\partial y_3}{\partial \alpha_1}$$

chain rule

$$= (2 y_1^T)(-b_1) + (2 y_2^T)(-b_1) + (2 y_3^T)(-b_1)$$

$$= -2 (y_1 + y_2 + y_3)^T b_1 .$$

$$= -2 (x_1 - \alpha_1 b_1 - \alpha_2 b_2 + x_2 - \alpha_1 b_1 - \alpha_2 b_2 + x_3 - \alpha_1 b_1 - \alpha_2 b_2)^T b_1$$

$$= -2 (x_1 + x_2 + x_3 - 3\alpha_1 b_1 - 3\alpha_2 b_2)^T b_1 .$$

Similarly,

$$\frac{\partial f}{\partial \alpha_2} = -2 (x_1 + x_2 + x_3 - 3\alpha_1 b_1 - 3\alpha_2 b_2)^T b_2 .$$