STAT 8260 Homework #1
Solutions

1. $\underset{\sim}{x} = (2,1,1,1,4)^T, \quad \underset{\sim}{y} = (3,6,1,5,7)^T, \quad \underset{\sim}{w} = (-1,2,4,-4,0)^T$

$$\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle = 2(3) + (1)(1) + (1)(6) + 1(5) + 4(7) = 46$$
$$\|\underset{\sim}{x}\|^2 = \langle \underset{\sim}{x}, \underset{\sim}{x} \rangle = 2^2 + 1^2 + 1^2 + 1^2 + 4^2 = 23$$
$$\|\underset{\sim}{y}\|^2 = \langle \underset{\sim}{y}, \underset{\sim}{y} \rangle = 3^2 + 1^2 + 6^2 + 5^2 + 7^2 = 120$$

$$\hat{\underset{\sim}{y}} = p(\underset{\sim}{y} | \underset{\sim}{x}) = \frac{\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle}{\|\underset{\sim}{x}\|^2} \underset{\sim}{x} = \frac{46}{23} \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 2 \\ 2 \\ 8 \end{pmatrix}$$

$$\underset{\sim}{y} - \hat{\underset{\sim}{y}} = (3,6,1,5,7)^T - (4,2,2,2,8)^T = (-1,4,-1,3,-1)^T$$

$$\langle \underset{\sim}{x}, \underset{\sim}{y} - \hat{\underset{\sim}{y}} \rangle = 2(-1) + 1(-1) + 1(4) + 1(3) + 4(-1) = 0 \Rightarrow \underset{\sim}{x} \perp (\underset{\sim}{y} - \hat{\underset{\sim}{y}})$$

$$\|\underset{\sim}{y} - \hat{\underset{\sim}{y}}\|^2 = (-1)^2 + (-1)^2 + 4^2 + 3^2 + (-1)^2 = 28$$
$$\|\hat{\underset{\sim}{y}}\|^2 = 4^2 + 2^2 + 2^2 + 2^2 + 8^2 = 92$$
$$\Rightarrow \|\hat{\underset{\sim}{y}}\|^2 + \|\underset{\sim}{y} - \hat{\underset{\sim}{y}}\|^2 = 92 + 28 = 120 = \|\underset{\sim}{y}\|^2$$

b) $\langle \underset{\sim}{w}, \underset{\sim}{x} \rangle = -1(2) + 4(1) + 2(1) - 4(1) + 0(4) = 0$

$$\underset{\sim}{z} = 3\underset{\sim}{x} + 2\underset{\sim}{w} = 3\begin{pmatrix} 2 & 1 & \\ 1 & 1 & 4 \end{pmatrix} + 2\begin{pmatrix} -1 & 4 & \\ 2 & -4 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 11 & \\ 7 & -5 & 12 \end{pmatrix}$$

$$\|\underset{\sim}{z}\|^2 = 4^2 + 11^2 + 7^2 + (-5)^2 + 12^2 = 355$$
$$\|\underset{\sim}{x}\|^2 = 23, \quad \|\underset{\sim}{w}\|^2 = (-1)^2 + 4^2 + 2^2 + (-4)^2 = 37$$
$$9\|\underset{\sim}{x}\|^2 + 4\|\underset{\sim}{w}\|^2 = 9(23) + 4(37) = 355 = \|\underset{\sim}{z}\|^2$$

Which must be true because of the Pythagorean Theorem (P.T.)

c) $\underline{i}_{A_1} = (1,1,0,0,0)^T, \ \underline{i}_{A_2} = (0,0,1,1,0)^T, \ \underline{i}_{A_3} = (0,0,0,0,1)^T$

$$p(\underline{y} \mid \underline{i}_{A_1}) = \underbrace{(\text{avg of } y_1, y_2)}_{=(3+6)/2 = 4.5} \times \underline{i}_{A_1} = (4.5, 4.5, 0, 0, 0)^T$$

Similarly, $p(\underline{y}, \underline{i}_{A_2}) = (0,0,3,3,0)^T, \ p(\underline{y}, \underline{i}_{A_3}) = (0,0,0,0,7)^T$

2. Yes. Proof: $p(c\,\underline{y} \mid \underline{x}) = \dfrac{\langle c\underline{y}, \underline{x}\rangle}{\|x\|^2}\underline{x} = \dfrac{c\langle \underline{y}, \underline{x}\rangle}{\|x\|^2}\underline{x} = c\,p(\underline{y} \mid \underline{x})$ for all $c \in \mathbb{R}$

$$p(\underline{y} \mid c\underline{x}) = \frac{\langle \underline{y}, c\underline{x}\rangle}{\|c\underline{x}\|^2}(c\underline{x}) = \frac{\cancel{c^2}\langle \underline{y}, \underline{x}\rangle}{\cancel{c^2}\|\underline{x}\|^2}\underline{x} = p(\underline{y} \mid \underline{x})$$

3. Let $Q(b) = \|\underline{y} - b\underline{x}\|^2 = \sum_i (y_i - bx_i)^2$

$$\frac{\partial Q(b)}{\partial b} = -2\sum (y_i - bx_i)x_i = 2b\sum x_i^2 - 2\sum x_i y_i$$

setting $\dfrac{\partial Q(b)}{\partial b} = 0$ and solving for b, we get $b = \dfrac{\sum x_i y_i}{\sum x_i^2} = \dfrac{\langle \underline{y}, \underline{x}\rangle}{\|\underline{x}\|^2}$

This is a minimum because $\dfrac{\partial^2 Q(b)}{\partial b^2} = 2\sum x_i^2 > 0.$

#4 a) $\underline{X}_1 = (1,1,1,0)^T$, $\underline{X}_2 = (1,1,0,0)^T$, $\underline{X}_3 = (1,0,0,1)^T$, $\underline{X}_4 = (10,3,0,7)^T$

$V_2 = \mathcal{L}(\underline{X}_1, \underline{X}_2) = C(\underline{A}_2)$ where $\underline{A}_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}$

$V_3 = \mathcal{L}(\underline{X}_1, \underline{X}_2, \underline{X}_3) = C(\underline{A}_3)$, where $\underline{A}_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$V_4 = \mathcal{L}(\underline{X}_1, \underline{X}_2, \underline{X}_3, \underline{X}_4) = C(\underline{A}_4)$ where $\underline{A}_4 = \begin{pmatrix} 1 & 1 & 1 & 10 \\ 1 & 1 & 0 & 3 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 7 \end{pmatrix}$

b) $V_2 = \mathcal{L}(\underline{X}_1, \underline{X}_2)$ has a spanning set of two vectors $\underline{X}_1, \underline{X}_2$.
$\underline{X}_1$, and $\underline{X}_2$ are linearly independent because

$$b_1 \underline{X}_1 + b_2 \underline{X}_2 = \underline{0} \Rightarrow \begin{pmatrix} b_1 + b_2 \\ b_1 + b_2 \\ b_1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow b_1 = b_2 = 0$$

so $\underline{X}_1, \underline{X}_2$ form a basis for $V_2 \Rightarrow \dim(V_2) = 2$

Similarly, $\underline{X}_1, \underline{X}_2, \underline{X}_3$ are linearly independent because

$$\underline{0} = b_1 \underline{X}_1 + b_2 \underline{X}_2 + b_3 \underline{X}_3 \Rightarrow \begin{pmatrix} b_1 + b_2 + b_3 \\ b_1 + b_2 \\ b_1 \\ b_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow b_1 = b_2 = b_3 = 0$$

$\Rightarrow \dim(V_3) = 3$

c.) $V_2 = \left\{ \begin{pmatrix} b_1 + b_2 \\ b_1 + b_2 \\ b_1 \\ 0 \end{pmatrix} \middle| \, b_1, b_2 \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} a \\ a \\ b \\ 0 \end{pmatrix} \middle| \, a, b \in \mathbb{R} \right\}$

which can be spanned by the basis $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}$

#4 c) (Continued)

~~...~~ Similarly, $V_3 = \left\{ \begin{pmatrix} a+c \\ a \\ b \\ c \end{pmatrix} \middle| a,b,c \in \mathbb{R} \right\}$

$\Rightarrow \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ forms a basis for V_3

d) $\underline{z} = (-1, 1, 0, 1)^T$

e) $\underline{x}_4 \perp \underline{z}$ since $\langle \underline{x}_4, \underline{z} \rangle = -10 + 3 + 0 + 7 = 0$

In addition, since $\underline{x}_4$ can be written $\sum_{i=1}^{3} b_i \underline{x}_i + c\underline{z}$,

$\langle \underline{x}_4, \underline{z} \rangle = 0 \Rightarrow \langle \sum_{i=1}^{3} b_i \underline{x}_i + c\underline{z}, \underline{z} \rangle = 0$

$\Rightarrow \sum_{i=1}^{3} b_i \underbrace{\langle \underline{x}_i, \underline{z} \rangle}_{=0} + c \langle \underline{z}, \underline{z} \rangle = 0$

$\Rightarrow c\|\underline{z}\|^2 = 0 \Rightarrow c = 0$ because $\underline{z} \neq \underline{0}$

$\Rightarrow \underline{x}_4 \in V_3$ and $\dim(V_4) = 3$.

f) V_3 is the set of all vectors orthogonal to $\underline{z}$. That is, the set of all vectors where the first component is equal to the sum of the 2^{nd} and 4^{th} components.

#5 a) $\underline{y} = (2,1,0,7,9,3)^T$ $\underline{a}_1 = (1,1,1,0,0,0)^T$, $\underline{a}_2 = (0,0,0,1,1,0)^T$

$\underline{a}_3 = (0,0,0,0,0,1)^T$

$V = \mathcal{L}(\underline{a}_1, \underline{a}_2, \underline{a}_3)$ $\underline{a}_1, \underline{a}_2, \underline{a}_3$ form an orthogonal basis for V

$\Rightarrow \hat{\underline{y}} = p(\underline{y}|V) = p(\underline{y}, \underline{a}_1) + p(\underline{y}, \underline{a}_2) + p(\underline{y}|\underline{a}_3)$

$= \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 8 \\ 8 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 8 \\ 8 \\ 3 \end{pmatrix}$

#5. a) (continued) $\underline{y}-\hat{\underline{y}} = (1,0,-1,-1,1,0)^T$

$$\|\underline{y}\|^2 = 2^2+7^2+3^2+1^2+9^2+0^2 = 144$$

$$\|\underline{y}-\hat{\underline{y}}\|^2 = (1)^2+(-1)^2+0^2+0^2+1^2+(-1)^2 = 4$$

$$\|\hat{\underline{y}}\|^2 = 1^2+8^2+3^2+1^2+8^2+1^2 = 140$$

b.) $$\hat{\underline{y}} = \begin{pmatrix} (y_{11}+y_{12}+y_{13})/3 \\ (y_{11}+y_{12}+y_{13})/3 \\ (y_{11}+y_{12}+y_{13})/3 \\ (y_{21}+y_{22})/2 \\ (y_{21}+y_{22})/2 \\ y_{31} \end{pmatrix}$$

#6. $\underline{x}_1 = (1,1,1,1)^T$, $\underline{x}_2 = (4,1,3,4)^T$, $\underline{y} = (1,9,5,5)^T$, $V = \mathcal{L}(\underline{x}_1, \underline{x}_2)$

a) $\underline{x}_1, \underline{x}_2$ are LIN, so they form a basis for V, but not an orthonormal basis. Orthonormalize with Gram-Schmidt:

Let $\underline{v}_1 = \underline{x}_1$, $\underline{v}_2 = \underline{x}_2 - p(\underline{x}_2|\underline{x}_1) = \underline{x}_2 - \dfrac{\langle \underline{x}_2, \underline{x}_1\rangle}{\|\underline{x}_1\|^2}\underline{x}_1$

$$= \begin{pmatrix}4\\1\\3\\4\end{pmatrix} - \frac{12}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} = \begin{pmatrix}1\\-2\\0\\1\end{pmatrix}$$

$$\Rightarrow p(\underline{y}|V) = p(\underline{y}|\underline{v}_1) + p(\underline{y}|\underline{v}_2)$$

$$= \begin{pmatrix}5\\5\\5\\5\end{pmatrix} + \begin{pmatrix}-2\\4\\0\\-2\end{pmatrix} = \begin{pmatrix}3\\9\\5\\3\end{pmatrix}, \quad \underline{e} = \underline{y}-\hat{\underline{y}} = \begin{pmatrix}-2\\0\\0\\2\end{pmatrix}$$

b.) $\hat{y}_1 = p(y|\underline{x}_1) = \begin{pmatrix}5\\5\\5\\5\end{pmatrix}$ from part 1(a)

$\hat{y}_2 = p(y|\underline{x}_2) = \frac{\langle y, x_2 \rangle}{\|\underline{x}_2\|^2} \underline{x}_2 = \frac{4+9+\cancel{8}15+20}{42}\begin{pmatrix}4\\1\\3\\4\end{pmatrix} = \frac{8}{7}\begin{pmatrix}4\\1\\3\\4\end{pmatrix}$

$\hat{y}_1 + \hat{y}_2 = \begin{pmatrix}5\\5\\5\\5\end{pmatrix} + \frac{8}{7}\begin{pmatrix}4\\1\\3\\4\end{pmatrix} \neq \begin{pmatrix}1\\9\\5\\5\end{pmatrix} = \underline{y}$

c) Every vector in V can be written as a linear combination of $\underline{x}_1, \underline{x}_2$ so it is sufficient to verify that $\underline{e} \perp \underline{x}_1$ and $\underline{e} \perp \underline{x}_2$

$\langle \underline{e}, \underline{x}_1 \rangle = -2 + 2 = 0, \quad \langle e, \underline{x}_2 \rangle = -8 + 8 = 0$

d) $\|\underline{y}\|^2 = 132, \|\hat{\underline{y}}\|^2 = 124, \|\underline{y} - \hat{\underline{y}}\|^2 = 8$ so P.T. holds $(132 = 124 + 8)$

$\|\hat{\underline{y}}\|^2 = \underline{y}^T P \underline{y} = \underline{y}^T \underline{X}(\underline{X}^T\underline{X})^{-1}\underline{X}^T\underline{y}$ where $\underline{X} = \begin{pmatrix}1 & 4\\1 & 1\\1 & 3\\1 & 4\end{pmatrix}$

$\underline{X}^T\underline{y} = \begin{pmatrix}1 & 1 & 1 & 1\\4 & 1 & 3 & 4\end{pmatrix}\begin{pmatrix}1\\9\\5\\5\end{pmatrix} = \begin{pmatrix}20\\48\end{pmatrix}, \quad \underline{X}^T\underline{X} = \begin{pmatrix}1 & 1 & 1 & 1\\4 & 1 & 3 & 4\end{pmatrix}\begin{pmatrix}1 & 4\\1 & 1\\1 & 3\\1 & 4\end{pmatrix} = \begin{pmatrix}4 & 12\\12 & 42\end{pmatrix}$

$\Rightarrow \|\hat{\underline{y}}\|^2 = \begin{pmatrix}20\\48\end{pmatrix}^T \underbrace{\begin{pmatrix}4 & 12\\12 & 42\end{pmatrix}^{-1}}_{= \frac{1}{4(42) - 12^2}\begin{pmatrix}42 & -12\\-12 & 4\end{pmatrix}} \begin{pmatrix}20\\48\end{pmatrix} = 124$

#6 (e) let $\underline{X}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\underline{X}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$. Then $\underline{X}_1, \underline{X}_2, \underline{X}_3, \underline{X}_4$ are linearly independent.

To orthogonalize $\underline{X}_1, \underline{X}_2, \underline{X}_3, \underline{X}_4$, let $\underline{V}_1 = \underline{X}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

$$V_2 = \underline{X}_2 - p(\underline{X}_2 | \underline{X}_1) = \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix}$$

$$\underline{V}_3 = \underline{X}_3 - p(\underline{X}_3 | \underline{V}_2) - p(\underline{X}_3 | \underline{V}_1)$$

$$= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} - 0 \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/4 \\ -1/4 \\ 3/4 \\ -1/4 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} -1 \\ -1 \\ 3 \\ -1 \end{pmatrix}$$

$$\underline{V}_4 = \underline{X}_4 - p(\underline{X}_4 | \underline{V}_3) - p(\underline{X}_4 | \underline{V}_2) - p(\underline{X}_4 | \underline{V}_1)$$

$$= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \frac{(-1/4)}{12/16} \begin{pmatrix} -1/4 \\ -1/4 \\ 3/4 \\ -1/4 \end{pmatrix} - \frac{1}{6} \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \frac{1}{12} \begin{pmatrix} -1 \\ -1 \\ 3 \\ -1 \end{pmatrix} - \frac{1}{6} \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 0 \\ 0 \\ +1/2 \end{pmatrix}$$

~~f) Projecting $\underline{y}$ onto $\mathcal{L}(V_1, V_2, V_3, V_4) = \mathbb{R}^4$ we get $\underline{y} = 5\underline{V}_1 - 2\underline{V}_2 + 0\underline{V}_3 + 4\underline{V}_4$~~

~~Projecting $\hat{\underline{y}}$ onto " we get $\hat{\underline{y}} = 5\underline{V}_1 - 2\underline{V}_2 + 0\underline{V}_3 + 0\underline{V}_4$.~~

#7. a) $\underline{v}_1 = \underline{j}_n$, $\underline{v}_2 = \underline{x} - p(\underline{x} \mid \underline{j}_n) = \underline{x} - \bar{x}\,\underline{j}_n$

b.) $\hat{y} = p(y \mid V) = a_0 \underline{j}_n + a_1 \underline{x}^*$

$$\hat{y} = p(y \mid j_n) + p(y \mid \underline{x}^*) = \bar{y}\,\underline{j}_n + \frac{\sum_i (x_i - \bar{x}\underline{j}_n) y_i}{\sum_i (x_i - \bar{x} j_n)^2} \underline{x}^*$$

$$= \underbrace{\bar{y}}_{a_0}\,\underline{j}_n + \underbrace{\frac{S_{xy}}{S_{xx}}}_{a_1} \underline{x}^*$$

c) $\Rightarrow \hat{y} = \bar{y}\,\underline{j}_n + \frac{S_{xy}}{S_{xx}}(\underline{x} - \bar{x}\,\underline{j}_n)$

$$= \underbrace{\left(\bar{y} - \frac{S_{xy}}{S_{xx}}\bar{x}\right)}_{b_0} \underline{j}_n + \underbrace{\frac{S_{xy}}{S_{xx}}}_{b_1} \underline{x}$$

d) $\|\hat{y}\|^2 = \|a_0 j_n + a_1 \underline{x}^*\|^2 = \sum_i (a_0 1 + a_1 x_i^*)^2$

$$= \sum (a_0^2 + a_0 a_1 x_i^* + a_1^2 (x_i^*)^2) = \underbrace{\sum a_0^2}_{= n\bar{y}^2} + a_0 a_1 \underbrace{\sum x_i^*}_{=0} + a_1^2 \underbrace{\sum (x_i^*)^2}_{= S_{xx}}$$

$$= n\bar{y}^2 + \frac{S_{xy}^2}{S_{xx}^2} S_{xx} = n\bar{y}^2 + \frac{S_{xy}^2}{S_{xx}}$$

or, by a formula on p. ~~[illegible]~~ 28 of our notes ~~[illegible]~~ ~~the square doesn't belong in the denominator [illegible] Ch. 8)~~:

$$\|p(y \mid V)\|^2 = \frac{\langle y, v_1 \rangle^2}{\|v_1\|^2} + \frac{\langle y, v_2 \rangle^2}{\|v_2\|^2} = \frac{(n\bar{y})^2}{n} + \frac{S_{xy}^2}{S_{xx}}$$

$$= n\bar{y}^2 + \frac{S_{xy}^2}{S_{xx}}$$

#7. d.) (continued) By the P.T., $\|\underline{y}-\hat{\underline{y}}\|^2 = \|\underline{y}\|^2 - \|\hat{\underline{y}}\|^2$

$$= \sum_i y_i^2 - \left(n\bar{y}^2 + \frac{S_{xy}^2}{S_{xx}}\right)$$

$$= \underbrace{\left(\sum_i y_i^2 - n\bar{y}^2\right)}_{= S_{yy}} - \frac{S_{xy}^2}{S_{xx}}$$

$$= S_{yy} - \frac{S_{xy}^2}{S_{xx}}$$

e) $$(\underline{X}^T\underline{X}) = \begin{pmatrix} \underline{j}_n^T \\ \underline{X}^T \end{pmatrix} (\underline{j}_n, \underline{X}) = \begin{pmatrix} \underline{j}_n^T \underline{j}_n & \underline{j}_n^T \underline{X} \\ \underline{X}^T \underline{j}_n & \underline{X}^T\underline{X} \end{pmatrix} = \begin{pmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{pmatrix}$$

$$\Rightarrow (\underline{X}^T\underline{X})^{-1} = \frac{1}{n\Sigma x_i^2 - (\Sigma x_i)^2} \begin{pmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & n \end{pmatrix} = \frac{1}{nS_{xx}} \begin{pmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & n \end{pmatrix}$$

and $$\underline{X}^T \underline{y} = \begin{pmatrix} \underline{j}_n^T \\ \underline{X}^T \end{pmatrix} \underline{y} = \begin{pmatrix} \underline{j}_n^T \underline{y} \\ \underline{X}^T \underline{y} \end{pmatrix} = \begin{pmatrix} \sum_i y_i \\ \sum_i x_i y_i \end{pmatrix}$$

$$\Rightarrow (\underline{X}^T\underline{X})^{-1}\underline{X}^T\underline{y} = \frac{1}{nS_{xx}} \begin{pmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & n \end{pmatrix} \begin{pmatrix} \Sigma y_i \\ \Sigma x_i y_i \end{pmatrix} = \frac{1}{nS_{xx}} \begin{pmatrix} (\Sigma x_i^2)(\Sigma y_i) - (\Sigma x_i)\Sigma x_i y_i \\ -(\Sigma x_i)(\Sigma y_i) + n\,\Sigma x_i y_i \end{pmatrix}$$

$$= \frac{1}{\cancel{n}S_{xx}} \begin{pmatrix} \cancel{n}\left[\bar{y}\Sigma x_i^2 - \bar{y}\bar{x}^2 n + \bar{y}\bar{x}^2 n - \bar{x}\Sigma x_i y_i\right] \\ \cancel{n}(\Sigma x_i y_i - n\bar{x}\bar{y}) \end{pmatrix}$$

$$= \frac{1}{S_{xx}} \begin{pmatrix} \bar{y} S_{xx} - \bar{x} S_{xy} \\ S_{xy} \end{pmatrix} = \begin{pmatrix} \bar{y} - \bar{x}\dfrac{S_{xy}}{S_{xx}} \\ \dfrac{S_{xy}}{S_{xx}} \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} \quad ✓$$

#7.

f) $a_0 = \bar{y} = 6$, $a_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum x_i y_i - \bar{x}\bar{y}n}{\sum x_i^2 - \bar{x}^2 n} = \frac{46 - 6\left(\frac{6}{4}\right)4}{14 - \left(\frac{6}{4}\right)^2 4} = 2$

$$\hat{\underline{y}} = \begin{pmatrix} 6 \\ 6 \\ 6 \\ 6 \end{pmatrix} + 2\begin{pmatrix} -1.5 \\ -.5 \\ .5 \\ 1.5 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 7 \\ 9 \end{pmatrix}$$

$b_0 = \bar{y} - \frac{S_{xy}}{S_{xx}}\bar{x} = 6 - 2\left(6/4\right) = 3$, $b_1 = a_1 = 2$

$\|y\|^2 = 168$, $\|\hat{y}\|^2 = n\bar{y}^2 + \frac{S_{xy}^2}{S_{xx}} = 4(6)^2 + \frac{(10)^2}{5} = 144 + 20 = 164$

$\|y - \hat{y}\|^2 = 168 - 164 = 4.$

$b_0 \langle y, j_n \rangle + b_1 \langle y, x \rangle = 3(24) + 2(46) = 72 + 92 = 164 = \|\hat{y}\|^2$

$\underline{y} - \hat{\underline{y}} = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$ $\quad \langle y - \hat{y}, j_n \rangle = 0$ and $\langle y - \hat{y}, x \rangle = 0$

$\Rightarrow y - \hat{y} \perp \mathcal{L}(j_n, \underline{x}) = V$

#8 Suppose $\underline{x}_1, \ldots, \underline{x}_k$ is a basis for V and $p(y|V) = \sum_{j=1}^{k} p(y|\underline{x}_j)$ $\forall y \in \Omega$

Prove $\underline{x}_1, \ldots, \underline{x}_k$ are mutually orthogonal. Must show $\langle x_i, x_j \rangle = 0$ $\forall i \neq j$

For each i, $i = 1, \ldots, k$, $\underline{x}_i = p(\underline{x}_i | V) = \sum_j p(\underline{x}_i | \underline{x}_j) = \sum_j \frac{\langle x_i, x_j \rangle}{\|x_j\|^2} \underline{x}_j$

$= \frac{\langle x_i, x_i \rangle}{\|x_i\|^2} \underline{x}_i + \sum_{i \neq j} \frac{\langle x_i, x_j \rangle}{\|x_j\|^2} \underline{x}_j = \underline{x}_i + \sum_{i \neq j} \frac{\langle x_i, x_j \rangle}{\|x_j\|^2} \underline{x}_j$

$\Rightarrow \underline{x}_i = \underline{x}_i + \sum_{\substack{j=1 \\ j \neq i}}^{k} \frac{\langle \underline{x}_i, \underline{x}_j \rangle}{\|x_j\|^2} \underline{x}_j \Rightarrow \sum_{j \neq i}^{k} \frac{\langle x_i, x_j \rangle}{\|x_j\|^2} \underline{x}_j = 0$

But since $\underline{x}_j$'s are linearly independent $\Rightarrow \left\{ \frac{\langle \underline{x}_i, \underline{x}_j \rangle}{\|\underline{x}_j\|^2} = 0, i \neq j \right\}$

#8 (Cont'd) $\Rightarrow \{ \langle x_i, x_j \rangle = 0,\ i \neq j \}$. This can be repeated for each i.
QED.

#9 Show that for $\underline{W}_{n\times k} = \underline{X}_{n\times k}\, \underline{B}_{k\times k}$ w/ $\underline{B}$ nonsingular,

$$\underline{X}(\underline{X}^T\underline{X})^{-1}\underline{X}^T = \underline{W}(\underline{W}^T\underline{W})^{-1}\underline{W}^T$$

$$\underline{X}(\underline{X}^T\underline{X})^{-1}\underline{X}^T = \cancel{\underline{B}^{-1}\underline{W}((\underline{B}^{-1}\underline{W})^T\underline{B}^{-1}\underline{W})^{-1}(\underline{B}^{-1}\underline{W})^T}$$

$$= \underline{W}\underline{B}^{-1}\left[(\underline{W}\underline{B}^{-1})^T\underline{W}\underline{B}^{-1}\right]^{-1}(\underline{W}\underline{B}^{-1})^T$$

$$= \underline{W}\underline{B}^{-1}\left[(\underline{B}^{-1})^T\underline{W}^T\underline{W}\underline{B}^{-1}\right]^{-1}(\underline{B}^{-1})^T\underline{W}^T$$

$$= \underline{W}\underline{B}^{-1}\underline{B}(\underline{W}^T\underline{W})^{-1}\underline{B}^T(\underline{B}^T)^{-1}\underline{W}^T = \underline{W}(\underline{W}^T\underline{W})^{-1}\underline{W}^T$$

#10. $\underline{x} = (2,-1,-1)^T$, $\underline{x}_1 = (1,1,1)^T$, $\underline{x}_2 = (1,-1,0)^T$

a) $$\underline{P}_{\mathcal{L}(\underline{x})} = \frac{1}{6}\begin{pmatrix}2\\-1\\-1\end{pmatrix}(2\ -1\ -1) = \frac{1}{6}\begin{pmatrix}4 & -2 & -2\\-2 & 1 & 1\\-2 & 1 & 1\end{pmatrix}$$

$$\underline{P}\underline{P} = \frac{1}{36}\begin{pmatrix}4 & -2 & -2\\-2 & 1 & 1\\-2 & 1 & 1\end{pmatrix}\begin{pmatrix}4 & -2 & -2\\-2 & 1 & 1\\-2 & 1 & 1\end{pmatrix} = \frac{1}{36}\begin{pmatrix}24 & -12 & -12\\-12 & 6 & 6\\-12 & 6 & 6\end{pmatrix} = \underline{P}$$

$\underline{P}$ is obviously symmetric

b) $V = \mathcal{L}(\underline{x}_1, \underline{x}_2)$ has orthonormal basis $\left\{\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}\right\}$

#10

b). (cont'd) $P_V = \frac{1}{6}\begin{pmatrix}\sqrt{2} & \sqrt{3}\\ \sqrt{2} & -\sqrt{3}\\ \sqrt{2} & 0\end{pmatrix}\begin{pmatrix}\sqrt{2} & \sqrt{2} & \sqrt{2}\\ \sqrt{3} & -\sqrt{3} & 0\end{pmatrix} = \frac{1}{6}\begin{pmatrix}5 & -1 & 2\\ -1 & 5 & 2\\ 2 & 2 & 2\end{pmatrix}$ symmetric

$$P_V P_V = \frac{1}{36}\begin{pmatrix}5 & -1 & 2\\ -1 & 5 & 2\\ 2 & 2 & 2\end{pmatrix}\begin{pmatrix}5 & -1 & 2\\ -1 & 5 & 2\\ 2 & 2 & 2\end{pmatrix} = \frac{1}{36}\begin{pmatrix}30 & -6 & 12\\ -6 & 30 & 12\\ 12 & 12 & 12\end{pmatrix} = P_V$$

#11.

$$V = \mathcal{L}(\underline{1}_n, \underline{x}^*) = \mathcal{L}\left(\frac{1}{\sqrt{n}}\underline{1}_n, \frac{1}{\sqrt{S_{xx}}}\underline{x}^*\right)$$

orthonormal

$$\Rightarrow P_V = \left(\frac{1}{\sqrt{n}}\underline{1}_n, \frac{1}{\sqrt{S_{xx}}}\underline{x}^*\right)\begin{pmatrix}\frac{1}{\sqrt{n}}\underline{1}_n^T\\ \frac{1}{\sqrt{S_{xx}}}\underline{x}^{*T}\end{pmatrix} = \frac{1}{n}\underline{1}_n\underline{1}_n^T + \frac{1}{S_{xx}}\underline{x}^*\underline{x}^{*T}$$

$$P_{V^\perp} = I - P_V = I - \frac{1}{n}\underline{1}_n\underline{1}_n^T - \frac{1}{S_{xx}}\underline{x}^*\underline{x}^{*T}$$

$$P_{V_1} = P_V - P_{V_0} = \frac{1}{n}\underline{1}_n\underline{1}_n^T + \frac{1}{S_{xx}}\underline{x}^*\underline{x}^{*T} - \frac{1}{n}\underline{1}_n\underline{1}_n^T = \frac{1}{S_{xx}}\underline{x}^*\underline{x}^{*T}$$

12. (2.23 in text) Suppose $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$ are vectors in $\mathbb{R}^n$, where without loss of generality $\underline{a}_1 = \underline{0}$. Then $\underline{a}_1, \ldots, \underline{a}_k$ are linearly dependent because (*) $c_1\underline{a}_1 + c_2\underline{a}_2 + \cdots + c_k\underline{a}_k = \underline{0}$ does not imply $c_1 = c_2 = \cdots = c_k = 0$. For example, we could have $c_1 = 1, c_2 = c_3 = \cdots = c_k = 0$ and still satisfy (*).

13. (2.24 in text) $\underline{A}, \underline{B}$ are both $n \times n$ w/ $\underline{A}\underline{B} = \underline{0}$

show this implies $\underline{A}, \underline{B}$ singular.

or $\underline{A} = \underline{0}$ or $\underline{B} = \underline{0}$

Suppose 1.) $\underline{A} \neq \underline{0}$ and 2.) $\underline{B} \neq \underline{0}$ and 3.) At least one of the matrices $\underline{A}, \underline{B}$ is nonsingular.

Then, by property 5 on rank (p. 15 of the class notes), either $\text{rank}(\underline{B}) = \text{rank}(\underline{A}\underline{B}) = \text{rank}(\underline{0}) = 0$

~~or rank(A) = rank(AB~~

or $\text{rank}(\underline{A}) = \text{rank}(\underline{A}\underline{B}) = \text{rank}(\underline{0}) = 0$

I.e., then either $\underline{B}$ or $\underline{A}$ is singular.

This is a contradiction, implying our assumption was wrong. ~~There fo~~ Therefore,

1.) $\underline{A} = \underline{0}$ or 2.) $\underline{B} = \underline{0}$ or 3.) none of the two matrices $\underline{A}, \underline{B}$, is nonsingular (both singular).