STAT 32950

Robert Winter

HW #3

1a. Data: $(2,12)$, $(8,9)$, $(6,9)$, $(8,10)$. $\mu_0 := \begin{bmatrix} 7 \\ 11 \end{bmatrix}$

1a(i) Calculate $\bar{x}$.

- $$\bar{x} = \begin{bmatrix} (2+8+6+8)/4 \\ (12+9+9+10)/4 \end{bmatrix} = \begin{bmatrix} 24/4 \\ 40/4 \end{bmatrix} = \begin{bmatrix} 6 \\ 10 \end{bmatrix} = \bar{x}$$

1a(ii) Find S.

- $$S = \frac{1}{4-1} \begin{bmatrix} \sum_{i=1}^{4} (x_{i1}-\bar{x}_1)^2 & \sum_{i=1}^{4} (x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2) \\ \sum_{i=1}^{4} (x_{i2}-\bar{x}_2)(x_{i1}-\bar{x}_1) & \sum_{i=1}^{4} (x_{i2}-\bar{x}_2)^2 \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} (2-6)^2+(8-6)^2+(6-6)^2+(8-6)^2 & (2-6)(12-10)+(8-6)(9-10)+(6-6)(9-10)+(8-6)(10- \\ (12-10)(2-6)+(9-10)(8-6)+(9-10)(6-6)+(10-10)(8-6) & (12-10)^2+(9-10)^2+(9-10)^2+(10-10)^2 \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} 16+4+0+4 & (-4)(2)+(2)(-1)+(0)(-1)+(2)(0) \\ (2)(-4)+(-1)(2)+(-1)(0)+(0)(2) & 4+1+1+0 \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} 24 & -8-2 \\ -8-2 & 6 \end{bmatrix}$$

$$\Rightarrow S = \begin{bmatrix} 8 & -10/3 \\ -10/3 & 2 \end{bmatrix}$$

1a(iii) Obtain S^{-1}.

- $$S^{-1} = \frac{1}{\det S} \begin{bmatrix} S_{22} & -S_{12} \\ -S_{21} & S_{11} \end{bmatrix}$$

$$= \frac{1}{(8)(2)-(-\frac{10}{3})(-\frac{10}{3})} \begin{bmatrix} 2 & \frac{10}{3} \\ \frac{10}{3} & 8 \end{bmatrix}$$

$$= \frac{1}{16-\frac{100}{9}} \begin{bmatrix} 2 & \frac{10}{3} \\ \frac{10}{3} & 8 \end{bmatrix}$$

$$= \frac{9}{44} \begin{bmatrix} 2 & \frac{10}{3} \\ \frac{10}{3} & 8 \end{bmatrix}$$

$$\Rightarrow S^{-1} = \begin{bmatrix} 9/22 & 15/22 \\ 15/22 & 18/11 \end{bmatrix}$$

1a(iv) Evaluate Hotelling's T^2.

- $T^2 = (\bar{x} - \mu_0)^T \left(\frac{S}{n}\right)^{-1} (\bar{x} - \mu_0)$

$$= \left(\begin{bmatrix} 6 \\ 10 \end{bmatrix} - \begin{bmatrix} 7 \\ 11 \end{bmatrix}\right)^T \cdot 4S^{-1} \left(\begin{bmatrix} 6 \\ 10 \end{bmatrix} - \begin{bmatrix} 7 \\ 11 \end{bmatrix}\right)$$

$$= \begin{bmatrix} -1 \\ -1 \end{bmatrix}^T \cdot 4 \begin{bmatrix} \frac{9}{22} & \frac{15}{22} \\ \frac{15}{22} & \frac{18}{11} \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

$$= 4 \begin{bmatrix} -1 & -1 \end{bmatrix} \begin{bmatrix} -\frac{9}{22} - \frac{15}{22} \\ -\frac{15}{22} - \frac{18}{11} \end{bmatrix}$$

$$= 4 \begin{bmatrix} -1 & -1 \end{bmatrix} \begin{bmatrix} -\frac{24}{22} \\ -\frac{51}{22} \end{bmatrix}$$

$$= 4\left(\frac{24}{22} + \frac{51}{22}\right)$$

$$= 4\left(\frac{75}{22}\right)$$

$$\Rightarrow T^2 = \frac{150}{11}$$

1a(v) Specify the distribution of T^2 under $H_0: \mu = \mu_0$.

- Under H_0,

$$T^2 \sim_{H_0} \frac{(n-1)p}{n-p} F_{p,n-p}$$

$$= \frac{(4-1)(2)}{4-2} F_{2,4-2}$$

$$\Rightarrow T^2 \sim_{H_0} 3F_{2,2}$$

- Equivalently, $\frac{1}{3}T^2 \sim_{H_0} F_{2,2}$

1b. Let $C = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. Transform $\{x_j\}$ into $y_j = Cx_j$.

Let $\mu_0^* = C\mu_0 = \begin{bmatrix} 18 \\ 4 \end{bmatrix}$.

- We write

$$X = \begin{bmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{bmatrix}$$

- Then $Y^T = CX^T$

$$= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 8 & 6 & 8 \\ 12 & 9 & 9 & 10 \end{bmatrix}$$

$$= \begin{bmatrix} 2+12 & 8+9 & 6+9 & 8+10 \\ 12-2 & 9-8 & 9-6 & 10-8 \end{bmatrix}$$

$$= \begin{bmatrix} 14 & 17 & 15 & 18 \\ 10 & 1 & 3 & 2 \end{bmatrix}$$

1b(i) Calculate $\bar{y}$.

- $\bar{y} = \begin{bmatrix} (14+17+15+18)/4 \\ (10+1+3+2)/4 \end{bmatrix} = \begin{bmatrix} 64/4 \\ 16/4 \end{bmatrix} = \begin{bmatrix} 16 \\ 4 \end{bmatrix} = \bar{y}$

1b(ii) Derive $S_y = CSC^T$.

- First, see attached sheet for proof that $S_y = CSC^T$.
- $S_y = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 & \frac{-10}{3} \\ \frac{-10}{3} & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

$$= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 - \frac{10}{3} & -8 - \frac{10}{3} \\ 2 - \frac{10}{3} & \frac{10}{3} + 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{14}{3} & \frac{-34}{3} \\ \frac{-4}{3} & \frac{16}{3} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{14}{3} - \frac{4}{3} & \frac{-34}{3} + \frac{16}{3} \\ -\frac{14}{3} - \frac{4}{3} & \frac{34}{3} + \frac{16}{3} \end{bmatrix}$$

→

$$= \begin{bmatrix} 10/3 & -18/3 \\ -18/3 & 50/3 \end{bmatrix}$$

$$\Rightarrow S_y = \begin{bmatrix} 10/3 & -6 \\ -6 & 50/3 \end{bmatrix}$$

1b(iii) Evaluate Hotelling's T^2 for $\{y_j\}$ under H_0: $\mu_y = \mu_0^*$.

- $T^2 = (\bar{y} - \mu_0^*)^T \left(\frac{S_y}{n}\right)^{-1} (\bar{y} - \mu_0^*)$

$$= \left(\begin{bmatrix} 16 \\ 4 \end{bmatrix} - \begin{bmatrix} 18 \\ 4 \end{bmatrix}\right)^T \cdot 4 S_y^{-1} \left(\begin{bmatrix} 16 \\ 4 \end{bmatrix} - \begin{bmatrix} 18 \\ 4 \end{bmatrix}\right)$$

$$= 4 \begin{bmatrix} -2 \\ 0 \end{bmatrix}^T \cdot \frac{1}{\det(S_y)} \begin{bmatrix} \frac{50}{3} & 6 \\ 6 & \frac{10}{3} \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

$$= \frac{4}{(\frac{10}{3})(\frac{50}{3}) - (-6)(-6)} \begin{bmatrix} -2 & 0 \end{bmatrix} \begin{bmatrix} -\frac{100}{3} + 0 \\ -12 + 0 \end{bmatrix}$$

$$= \frac{4}{\frac{500}{9} - 36} \begin{bmatrix} -2 & 0 \end{bmatrix} \begin{bmatrix} -\frac{100}{3} \\ -12 \end{bmatrix}$$

$$= \frac{36}{176} \cdot \frac{200}{3}$$

$$\Rightarrow T^2 = \frac{150}{11}$$ — the same as in Part (a)!

1c. Prove that, in general, if C is a $(p \times p)$ invertible matrix, then the transformed data $y_j = Cx_j$ has the same Hotelling's T^2 statistic (under H_0: $\mu_y = C\mu_0$).

- Let $\bar{x}$ be the sample mean of the x_j's and S_x the sample covariance of the x_j's.
- Note that $\bar{y}$, the sample mean of the y_j's, is
$$\bar{y} = \frac{1}{n} \sum_{j=1}^{n} y_j = \frac{1}{n} \sum_{j=1}^{n} (C x_j) = \frac{1}{n} C \sum_{j=1}^{n} x_j = C\left(\frac{1}{n} \sum_{j=1}^{n} x_j\right) = C\bar{x}.$$
- Let Hotelling's T^2 statistic for the x_j's be
$$T_x^2 = (\bar{x} - \mu_0)^T \left(\frac{S_x}{n}\right)^{-1} (\bar{x} - \mu_0).$$

- Then Hotelling's T^2 statistic for the y_j's is

$$\begin{aligned}
T^2_{\bar{y}} &= (\bar{y}-\mu_y)^T\left(\frac{S_y}{n}\right)^{-1}(\bar{y}-\mu_y) \\
&= (C\bar{x}-C\mu_0)^T \cdot n\, S_y^{-1}\,(C\bar{x}-C\mu_0) \\
&= (C(\bar{x}-\mu_0))^T \cdot n(CS_xC^T)^{-1}(C(\bar{x}-\mu_0)) \\
&= (\bar{x}-\mu_0)^T C^T \cdot n(C^T)^{-1}S_x^{-1}C^{-1}C(\bar{x}-\mu_0) \\
&= n(\bar{x}-\mu_0)^T C^T (C^{-1})^T S_x^{-1} I\,(\bar{x}-\mu_0) \\
&= n(\bar{x}-\mu_0)^T (C^{-1}C)^T S_x^{-1}(\bar{x}-\mu_0) \\
&= n(\bar{x}-\mu_0)^T I^T S_x^{-1}(\bar{x}-\mu_0) \\
&= n(\bar{x}-\mu_0)^T I\, S_x^{-1}(\bar{x}-\mu_0) \\
&= (\bar{x}-\mu_0)^T\left(\frac{S_x}{n}\right)^{-1}(\bar{x}-\mu_0) \\
&= T^2_{\bar{x}},
\end{aligned}$$

as desired. $\square$

1b(ii). contd • First, we show $S_y = CSC^T$.

We write

$$S_y = \frac{1}{4-1}(Y^T - \bar{y}\mathbb{1}^T)(Y^T - \bar{y}\mathbb{1}^T)^T$$
$$= \frac{1}{4-1}(Y^T - \bar{y}\mathbb{1}^T)(Y - \mathbb{1}\bar{y}^T)$$
$$= \frac{1}{4-1}(CX^T - C\bar{x}\mathbb{1}^T)(XC^T - \mathbb{1}(C\bar{x})^T)$$
$$= \frac{1}{4-1}(CX^T - C\bar{x}\mathbb{1}^T)(XC^T - \mathbb{1}\bar{x}^T C^T)$$
$$= \frac{1}{4-1}(CX^T XC^T - CX^T\mathbb{1}\bar{x}^T C^T - C\bar{x}\mathbb{1}^T XC^T + C\bar{x}\mathbb{1}^T\mathbb{1}\bar{x}^T C^T)$$
$$= \frac{1}{4-1}C(X^T XC^T - X^T\mathbb{1}\bar{x}^T C^T - \bar{x}\mathbb{1}^T XC^T + \bar{x}\mathbb{1}^T\mathbb{1}\bar{x}^T C^T)$$
$$= \frac{1}{4-1}C(X^T X - X^T\mathbb{1}\bar{x}^T - \bar{x}\mathbb{1}^T X + \bar{x}\mathbb{1}^T\mathbb{1}\bar{x}^T)C^T$$
$$= \frac{1}{4-1}C(X^T - \bar{x}\mathbb{1}^T)(X - \mathbb{1}\bar{x}^T)C^T$$
$$= \frac{1}{4-1}C(X^T - \bar{x}\mathbb{1}^T)(X^T - \bar{x}\mathbb{1}^T)^T C^T$$
$$= CSC^T.$$

2. X, Y are (2×1) RVs w/ joint covariance matrix

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \Sigma_{11} = \begin{bmatrix} 1 & p \\ p & 1 \end{bmatrix}, \Sigma_{22} = \begin{bmatrix} 1 & q \\ q & 1 \end{bmatrix},$$

$$\Sigma_{12} = \Sigma_{21} = \begin{bmatrix} r & r \\ r & r \end{bmatrix},$$

$p, q, r \in (0, 1)$.

2a. Derive ρ_i^*.

- $A = \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$

$$= \begin{bmatrix} 1 & p \\ p & 1 \end{bmatrix}^{-1} \begin{bmatrix} r & r \\ r & r \end{bmatrix} \begin{bmatrix} 1 & q \\ q & 1 \end{bmatrix} \begin{bmatrix} r & r \\ r & r \end{bmatrix}$$

$$= \frac{1}{1-p^2} \begin{bmatrix} 1 & -p \\ -p & 1 \end{bmatrix} r \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \frac{1}{1-q^2} \begin{bmatrix} 1 & -q \\ -q & 1 \end{bmatrix} r \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$= \frac{r^2}{(1-p^2)(1-q^2)} \begin{bmatrix} 1-p & 1-p \\ 1-p & 1-p \end{bmatrix} \begin{bmatrix} 1-q & 1-q \\ 1-q & 1-q \end{bmatrix}$$

$$= \frac{r^2}{(1+p)(1-p)(1+q)(1-q)} \begin{bmatrix} (1-p)(1-q)+(1-p)(1-q) & (1-p)(1-q)+(1-p)(1-q) \\ (1-p)(1-q)+(1-p)(1-q) & (1-p)(1-q)+(1-p)(1-q) \end{bmatrix}$$

$$= \frac{r^2 \cdot 2(1-p)(1-q)}{(1+p)(1-p)(1+q)(1-q)} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$= \frac{2r^2}{(1+p)(1+q)} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

- Let $s = \frac{2r^2}{(1+p)(1+q)}$, so $A = \begin{bmatrix} s & s \\ s & s \end{bmatrix}$.
- Then $\det(A - \rho^2 I) = 0$

$$\Rightarrow \det\left(\begin{bmatrix} s & s \\ s & s \end{bmatrix} - \begin{bmatrix} \rho^2 & 0 \\ 0 & \rho^2 \end{bmatrix}\right) = 0$$

$$\Rightarrow \det\left(\begin{bmatrix} s-\rho^2 & s \\ s & s-\rho^2 \end{bmatrix}\right) = 0$$

$$\Rightarrow (s-\rho^2)^2 - s^2 = 0$$

$$\Rightarrow (s-\rho^2+s)(s-\rho^2-s) = 0$$

$$\Rightarrow (2s-\rho^2)(-\rho^2) = 0$$

$\longrightarrow$

$\Rightarrow \rho^2 = 0$ or $\rho^2 = 2s = \frac{4r^2}{(1+p)(1+q)} > 0$, since $p, q, r \in (0,1)$

- Meanwhile,

$$B = \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$$
$$= \begin{bmatrix} 1 & q \\ q & 1 \end{bmatrix}^{-1} \begin{bmatrix} r & r \\ r & r \end{bmatrix} \begin{bmatrix} 1 & p \\ p & 1 \end{bmatrix}^{-1} \begin{bmatrix} r & r \\ r & r \end{bmatrix}$$
$$= \frac{1}{1-q^2} \begin{bmatrix} 1 & -q \\ -q & 1 \end{bmatrix} r \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \frac{1}{1-p^2} \begin{bmatrix} 1 & -p \\ -p & 1 \end{bmatrix} r \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$
$$= \frac{r^2}{(1-q^2)(1-p^2)} \begin{bmatrix} 1-q & 1-q \\ 1-q & 1-q \end{bmatrix} \begin{bmatrix} 1-p & 1-p \\ 1-p & 1-p \end{bmatrix}$$
$$= \frac{r^2}{(1+q)(1-q)(1+p)(1-p)} \begin{bmatrix} (1-q)(1-p)+(1-q)(1-p) & (1-q)(1-p)+(1-q)(1-p) \\ (1-q)(1-p)+(1-q)(1-p) & (1-q)(1-p)+(1-q)(1-p) \end{bmatrix}$$
$$= \frac{r^2 \cdot 2(1-q)(1-p)}{(1+q)(1-q)(1+p)(1-p)} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$
$$= \frac{2r^2}{(1+q)(1+p)} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$
$$= A,$$

so B shares A's eigenvalues.

- Thus, the largest shared eigenvalue of A & B is

$$\rho_1^{*2} = \frac{4r^2}{(1+p)(1+q)}$$
$$\Rightarrow \rho_1^* = \frac{2r}{\sqrt{(1+p)(1+q)}}$$

is the largest canonical correlation b/w X & Y.

2b. Derive the canonical variate pairs $(U_1, V_1) = (a_1^T X, b_1^T Y)$ corresponding to ρ_1^*, w/ normalization $a_1^T \Sigma_{11} a_1 = 1$, $b_1^T \Sigma_{22} b_1 = 1$.

- Since ρ_1^{*2} is an eigenvalue of A, we have
$$A a_1 = \rho_1^{*2} a_1.$$
- The corresponding eigenspace is
$$E = \mathrm{null}(A - \rho_1^{*2} I)$$
$$= \mathrm{null}\left(\begin{bmatrix} \frac{2r^2}{(1+p)(1+q)} & \frac{2r^2}{(1+p)(1+q)} \\ \frac{2r^2}{(1+p)(1+q)} & \frac{2r^2}{(1+p)(1+q)} \end{bmatrix} - \begin{bmatrix} \rho_1^{*2} & 0 \\ 0 & \rho_1^{*2} \end{bmatrix} \right)$$
$$= \mathrm{null}\left(\begin{bmatrix} \frac{1}{2}\rho_1^{*2} & \frac{1}{2}\rho_1^{*2} \\ \frac{1}{2}\rho_1^{*2} & \frac{1}{2}\rho_1^{*2} \end{bmatrix} - \begin{bmatrix} \rho_1^{*2} & 0 \\ 0 & \rho_1^{*2} \end{bmatrix} \right)$$
$$= \mathrm{null}\left(\begin{bmatrix} -\frac{1}{2}\rho_1^{*2} & \frac{1}{2}\rho_1^{*2} \\ \frac{1}{2}\rho_1^{*2} & -\frac{1}{2}\rho_1^{*2} \end{bmatrix} \right)$$
$$= \mathrm{null}\left(\frac{\rho_1^{*2}}{2} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \right)$$
- So, $a_1 = (a_{11}, a_{12})$ satisfies
$$\frac{\rho_1^{*2}}{2} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} -a_{11} + a_{12} \\ a_{11} - a_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} -a_{11} + a_{12} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow a_{11} = a_{12},$$
so $a_1 = (a_{11}, a_{11})$ for some $a_{11} \in \mathbb{R}$.
- In particular,
$$a_1^T \Sigma_{11} a_1 = 1 \Rightarrow \begin{bmatrix} a_{11} & a_{11} \end{bmatrix} \begin{bmatrix} 1 & p \\ p & 1 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{11} \end{bmatrix} = 1$$
$$\Rightarrow \begin{bmatrix} a_{11} & a_{11} \end{bmatrix} \begin{bmatrix} a_{11}(1+p) \\ a_{11}(1+p) \end{bmatrix} = 1$$
$$\Rightarrow a_{11}^2(1+p) + a_{11}^2(1+p) = 1$$
$$\Rightarrow 2a_{11}^2(1+p) = 1$$

$$\Rightarrow a_{11}^2 = \frac{1}{2(1+p)}$$

$$\Rightarrow a_{11} = \frac{1}{\sqrt{2(1+p)}}$$

- Thus, $a_1 = \begin{bmatrix} 1/\sqrt{2(1+p)} \\ 1/\sqrt{2(1+p)} \end{bmatrix}$

- Symmetrically,

$$Bb_1 = \rho_1^{*2} b_1,$$

with corresponding eigenspace

$$E = \text{null}(B - \rho_1^{*2} I)$$
$$= \text{null}\left(\frac{\rho_1^{*2}}{2}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\right) \quad \text{since } B = A$$

- So, $b_1 = (b_{11}, b_{12})$ satisfies

$$\frac{\rho_1^{*2}}{2}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} b_{11} \\ b_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow b_{11} = b_{12} \quad \text{by symmetry with before.}$$

- So, $b_1 = (b_{11}, b_{11})$ for some $b_{11} \in \mathbb{R}$.
- In particular,

$$b_1^T \Sigma_{22} b_1 = 1 \Rightarrow [b_{11} \;\; b_{11}]\begin{bmatrix} 1 & q \\ q & 1 \end{bmatrix}\begin{bmatrix} b_{11} \\ b_{11} \end{bmatrix} = 1$$

$$\Rightarrow [b_{11} \;\; b_{11}]\begin{bmatrix} b_{11}(1+q) \\ b_{11}(1+q) \end{bmatrix} = 1$$

$$\Rightarrow b_{11}^2(1+q) + b_{11}^2(1+q) = 1$$

$$\Rightarrow 2b_{11}^2(1+q) = 1$$

$$\Rightarrow b_{11}^2 = \frac{1}{2(1+q)}$$

$$\Rightarrow b_{11} = \frac{1}{\sqrt{2(1+q)}}$$

- Thus, $b_1 = \begin{bmatrix} 1/\sqrt{2(1+q)} \\ 1/\sqrt{2(1+q)} \end{bmatrix}$

- So, if $X=(X_1, X_2)$ & $Y=(Y_1, Y_2)$,

$$U_1 = a_1^T X$$
$$= \begin{bmatrix} \frac{1}{\sqrt{2(1+p)}} & \frac{1}{\sqrt{2(1+p)}} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$
$$= \frac{1}{\sqrt{2(1+p)}} X_1 + \frac{1}{\sqrt{2(1+p)}} X_2,$$

and

$$V_1 = b_1^T Y$$
$$= \begin{bmatrix} \frac{1}{\sqrt{2(1+q)}} & \frac{1}{\sqrt{2(1+q)}} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$
$$= \frac{1}{\sqrt{2(1+q)}} Y_1 + \frac{1}{\sqrt{2(1+q)}} Y_2$$

- That is,

$$(U_1, V_1) = \left(\frac{1}{\sqrt{2(1+p)}} X_1 + \frac{1}{\sqrt{2(1+p)}} X_2, \frac{1}{\sqrt{2(1+q)}} Y_1 + \frac{1}{\sqrt{2(1+q)}} Y_2 \right)$$

5a. Let $X = \begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix}$ be an $(n \times p)$ data matrix of n observations w/ j^{th} obs. $x_j \in \mathbb{R}^p$. Sample covariance matrix can be expressed as $S = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})(x_j - \bar{x})^T$. Show

$$S = \frac{1}{n-1} X^T H X, \quad \text{w/} \quad H = I_n - \tfrac{1}{n} \mathbb{1}_n \mathbb{1}_n^T.$$

- $$\begin{aligned} \frac{1}{n-1} X^T H X &= \frac{1}{n-1} X^T (I_n - \tfrac{1}{n} \mathbb{1}_n \mathbb{1}_n^T) X \\ &= \frac{1}{n-1} X^T (I_n X - \tfrac{1}{n} \mathbb{1}_n \mathbb{1}_n^T X) \\ &= \frac{1}{n-1} (X^T X - \tfrac{1}{n} X^T \mathbb{1}_n \mathbb{1}_n^T X) \\ &= \frac{1}{n-1} (X^T X - \tfrac{1}{n} (\mathbb{1}_n^T X)^T \mathbb{1}_n^T X) \end{aligned}$$
- $$\begin{aligned} X^T X &= [x_1 \cdots x_n] \begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix} \\ &= x_1 x_1^T + \cdots + x_n x_n^T \\ &= \textstyle\sum_{j=1}^n x_j x_j^T \end{aligned}$$
- $$\begin{aligned} \mathbb{1}_n^T X &= [1 \cdots 1] \begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix} \\ &= x_1^T + \cdots + x_n^T \\ &= \textstyle\sum_{j=1}^n x_j^T \end{aligned}$$
- $$\begin{aligned} (\mathbb{1}_n^T X)^T &= (\textstyle\sum_{j=1}^n x_j^T)^T \\ &= \textstyle\sum_{j=1}^n (x_j^T)^T \\ &= \textstyle\sum_{j=1}^n x_j \end{aligned}$$
- So, continuing,

$$\begin{aligned} \frac{1}{n-1} X^T H X &= \frac{1}{n-1} \left(\textstyle\sum_{j=1}^n x_j x_j^T - \tfrac{1}{n} (\textstyle\sum_{j=1}^n x_j)(\textstyle\sum_{j=1}^n x_j^T) \right) \\ &= \frac{1}{n-1} \left(\textstyle\sum_{j=1}^n x_j x_j^T - \tfrac{1}{n} (n\bar{x})(n\bar{x}^T) \right) \\ &= \frac{1}{n-1} \left(\textstyle\sum_{j=1}^n x_j x_j^T - n\bar{x}\bar{x}^T \right) \\ &= \frac{1}{n-1} \left(\textstyle\sum_{j=1}^n x_j x_j^T - n\bar{x}\bar{x}^T - n\bar{x}\bar{x}^T + n\bar{x}\bar{x}^T \right) \\ &= \frac{1}{n-1} \left[\textstyle\sum_{j=1}^n x_j x_j^T - (\textstyle\sum_{j=1}^n x_j)\bar{x}^T - \bar{x}(\textstyle\sum_{j=1}^n x_j^T) + \textstyle\sum_{j=1}^n \bar{x}\bar{x}^T \right] \\ &= \frac{1}{n-1} \left[\textstyle\sum_{j=1}^n x_j x_j^T - \textstyle\sum_{j=1}^n x_j \bar{x}^T - \textstyle\sum_{j=1}^n \bar{x} x_j^T + \textstyle\sum_{j=1}^n \bar{x}\bar{x}^T \right] \\ &= \frac{1}{n-1} \left[\textstyle\sum_{j=1}^n (x_j x_j^T - x_j \bar{x}^T - \bar{x} x_j^T + \bar{x}\bar{x}^T) \right] \\ &= \frac{1}{n-1} \textstyle\sum_{j=1}^n (x_j - \bar{x})(x_j^T - \bar{x}^T) \\ &= \frac{1}{n-1} \textstyle\sum_{j=1}^n (x_j - \bar{x})(x_j - \bar{x})^T \end{aligned}$$

$\rightarrow$

$$\Rightarrow \hat{\Sigma}_{j=1} \frac{1}{n-1} X^T H X = S,$$

as desired. $\square$

5b. Let $W = AY + c$; Y is p-dim RV, A is $(k \times p)$ scalar matrix, $c \in \mathbb{R}^k$ fixed. Show $Cov(W) = A\,Cov(Y)A^T$.

$$
\begin{aligned}
\cdot\ Cov(W) &= \mathbb{E}[(W-\mathbb{E}[W])(W-\mathbb{E}[W])^T] \\
&= \mathbb{E}[(W-\mathbb{E}[W])(W^T-\mathbb{E}[W]^T)] \\
&= \mathbb{E}[WW^T - W\mathbb{E}[W]^T - \mathbb{E}[W]W^T + \mathbb{E}[W]\mathbb{E}[W]^T] \\
&= \mathbb{E}[WW^T] - \mathbb{E}[W\mathbb{E}[W]^T] - \mathbb{E}[\mathbb{E}[W]W^T] + \mathbb{E}[\mathbb{E}[W]\mathbb{E}[W]^T] \\
&= \mathbb{E}[WW^T] - \mathbb{E}[W]\mathbb{E}[W]^T - \mathbb{E}[W]\mathbb{E}[W^T] + \mathbb{E}[W]\mathbb{E}[W]^T \\
&= \mathbb{E}[WW^T] - \mathbb{E}[W]\mathbb{E}[W^T] \\
&= \mathbb{E}[(AY+c)(AY+c)^T] - \mathbb{E}[AY+c]\mathbb{E}[(AY+c)^T] \\
&= \mathbb{E}[(AY+c)(Y^TA^T+c^T)] - (A\mathbb{E}[Y]+c)\mathbb{E}[Y^TA^T+c^T] \\
&= \mathbb{E}[AYY^TA^T + AYc^T + cY^TA^T + cc^T] \\
&\quad - (A\mathbb{E}[Y]+c)(\mathbb{E}[Y^T]A^T + c^T) \\
&= \mathbb{E}[AYY^TA^T] + A\mathbb{E}[Y]c^T + c\mathbb{E}[Y^T]A^T + cc^T \\
&\quad - (A\mathbb{E}[Y]\mathbb{E}[Y^T]A^T + A\mathbb{E}[Y]c^T + c\mathbb{E}[Y^T]A^T + cc^T) \\
&= \mathbb{E}[AYY^TA^T] - A\mathbb{E}[Y]\mathbb{E}[Y^T]A^T \\
&= A\mathbb{E}[YY^T]A^T - A\mathbb{E}[Y]\mathbb{E}[Y^T]A^T \\
&= A(\mathbb{E}[YY^T]A^T - \mathbb{E}[Y]\mathbb{E}[Y^T]A^T) \\
&= A(\mathbb{E}[YY^T] - \mathbb{E}[Y]\mathbb{E}[Y^T])A^T \\
&= A\,Cov(Y)A^T,
\end{aligned}
$$

as desired. $\square$

5c. Let Y be a p-dim RV, W a q-dim RV, $a \in \mathbb{R}^p$, $b \in \mathbb{R}^q$ fixed vectors. Show

$$\mathrm{Cov}(a^T Y, b^T W) = a^T \mathrm{Cov}(Y, W) b = b^T \mathrm{Cov}(W, Y) a.$$

- $$\begin{aligned}
\mathrm{Cov}(a^T Y, b^T W) &= \mathbb{E}[(a^T Y - \mathbb{E}[a^T Y])(b^T W - \mathbb{E}[b^T W])^T] \\
&= \mathbb{E}[(a^T Y - a^T \mathbb{E}[Y])(W^T b - \mathbb{E}[W^T b])] \\
&= \mathbb{E}[a^T Y W^T b - a^T Y \mathbb{E}[W^T b] - a^T \mathbb{E}[Y] W^T b \\
&\quad + a^T \mathbb{E}[Y] \mathbb{E}[W^T b]] \\
&= \mathbb{E}[a^T Y W^T b] - \mathbb{E}[a^T Y \mathbb{E}[W^T] b] \\
&\quad - \mathbb{E}[a^T \mathbb{E}[Y] W^T b] + \mathbb{E}[a^T \mathbb{E}[Y] \mathbb{E}[W^T] b] \\
&= a^T \mathbb{E}[Y W^T] b - a^T \mathbb{E}[Y] \mathbb{E}[W^T] b \\
&\quad - a^T \mathbb{E}[Y] \mathbb{E}[W^T] b + a^T \mathbb{E}[Y] \mathbb{E}[W^T] b \\
&= a^T \mathbb{E}[Y W^T] b - a^T \mathbb{E}[Y] \mathbb{E}[W^T] b \\
&= a^T (\mathbb{E}[Y W^T] b - \mathbb{E}[Y] \mathbb{E}[W^T] b) \\
&= a^T (\mathbb{E}[Y W^T] - \mathbb{E}[Y] \mathbb{E}[W^T]) b \\
&= a^T \mathrm{Cov}(Y, W) b.
\end{aligned}$$

- Since a is $(1 \times p)$, $\mathrm{Cov}(Y, W)$ is $(p \times q)$, and b is $(q \times 1)$, $a^T \mathrm{Cov}(Y, W) b$ is (1×1), hence is its own transpose:

$$\begin{aligned}
a^T \mathrm{Cov}(Y, W) b &= (a^T \mathrm{Cov}(Y, W) b)^T \\
&= b^T \mathrm{Cov}(Y, W)^T (a^T)^T \\
&= b^T (\mathbb{E}[Y W^T] - \mathbb{E}[Y] \mathbb{E}[W^T])^T a \\
&= b^T (\mathbb{E}[Y W^T]^T - \mathbb{E}[W^T]^T \mathbb{E}[Y]^T) a \\
&= b^T (\mathbb{E}[(Y W^T)^T] - \mathbb{E}[(W^T)^T] \mathbb{E}[Y^T]) a \\
&= b^T (\mathbb{E}[W Y^T] - \mathbb{E}[W] \mathbb{E}[Y^T]) a \\
&= b^T \mathrm{Cov}(W, Y) a.
\end{aligned}$$

- Thus,

$$\mathrm{Cov}(a^T Y, b^T W) = a^T \mathrm{Cov}(Y, W) b = b^T \mathrm{Cov}(W, Y) a,$$

as desired. $\square$