Linear Models in Statistics: HW1

201060072: Boncho Ku

2.47 Let

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$

(a) Find a symmetric generalized inverse for $\mathbf{A}$.

Sol. Since the first row of $\mathbf{A}$ is expressed as the sum of the last second and third rows of $\mathbf{A}$, and the second row is neither a multiple of the first and third, the rank of $\mathbf{A}$ is 2. Let submatrix $\mathbf{C_1}$ be as follows:

$$\mathbf{C_1} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

Then the inverse matrix of $\mathbf{C_1}$ is calculated as

$$\mathbf{C_1^{-1}} = \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 1.0 \end{pmatrix}$$

By substituting $\mathbf{C_1}$ to $(\mathbf{C_1^{-1}})^{\mathbf{T}}$, therefore, the symmetric generalized inverse matrix $\mathbf{A_1^-}$ can be derived as

$$\mathbf{A_1^-} = \begin{pmatrix} 0.5 & -0.5 & 0.0 \\ -0.5 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 \end{pmatrix}$$

(b) Find a nonsymmetric generalized inverse for $\mathbf{A}$.

Sol. Let the submatrix of $\mathbf{A}$ be $\mathbf{C_2}$ as follows,

$$\mathbf{C_1} = \begin{pmatrix} 2 & 2 \\ 2 & 0 \end{pmatrix}$$

Then the inverse matrix of $\mathbf{C_2}$ is

$$\mathbf{C_2^{-1}} = \begin{pmatrix} 0.0 & 0.5 \\ 0.5 & -0.5 \end{pmatrix}$$

The nonsymmetric generalized inverse $\mathbf{A_2^{-1}}$ is

$$\mathbf{A_2^{-}} = \begin{pmatrix} 0.0 & 0.0 & 0.0 \\ 0.0 & 0.5 & 0.0 \\ 0.5 & -0.5 & 0.0 \end{pmatrix}$$

2.76 For the positive definite matrix

$$\mathbf{A} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

calculate the eigenvalues and eigenvectors and find the square root matrix $\mathbf{A^{1/2}}$ as in (2.108). Check by showing $(\mathbf{A^{1/2}})^{\mathbf{2}} = \mathbf{A}$.

Sol. By solving characteristic equation $\mathbf{Ax} = \lambda\mathbf{x}$

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{pmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{pmatrix} = (\lambda-3)(\lambda-1) = 0$$

$$\therefore \quad \lambda_1 = 3, \lambda_2 = 1$$

Let $\mathbf{D} = \mathbf{diag}(\lambda_1, \lambda_2) = \mathbf{diag}(3, 1)$ and eigenvectors corresponding to each eigenvalue can be calculated as

$$\mathbf{x}_{\lambda=\mathbf{3}} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \quad , \quad \mathbf{x}_{\lambda=\mathbf{1}} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

And let $\mathbf{H}$ be an orthogonal matrix consisting of eigenvectors of $\mathbf{A}$, then the square root matrix $\mathbf{A^{1/2}}$ can be derived by spectral decompositon,

$$\begin{aligned} \mathbf{A^{1/2}} &= \mathbf{HD^{1/2}H^T} \\ &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1+\sqrt{3} & 1-\sqrt{3} \\ 1-\sqrt{3} & 1+\sqrt{3} \end{pmatrix} \end{aligned}$$

For checking $(\mathbf{A^{1/2}})^\mathbf{2} = \mathbf{A}$,

$$\begin{pmatrix} 1+\sqrt{3} & 1-\sqrt{3} \\ 1-\sqrt{3} & 1+\sqrt{3} \end{pmatrix} \begin{pmatrix} 1+\sqrt{3} & 1-\sqrt{3} \\ 1-\sqrt{3} & 1+\sqrt{3} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = \mathbf{A}$$

3.20 Let $\mathbf{y} = (y_1, y_2, y_3)^T$ be a random vector and covariance matrix

$$\mu = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 3 & 10 \end{pmatrix}$$

(a) Let $z = 2y_1 - 3y_2 + y_3$. Find $E(z)$ and $var(z)$.

Sol. Let $\mathbf{a}$ be a 3×1 constant vector with $\mathbf{a} = (2, -3, 1)^{\mathbf{T}}$. Then z is equal to $\mathbf{a^T y}$ and $E(z) = E(\mathbf{a^T y}) = \mathbf{a^T} E(\mathbf{y}) = \mathbf{a^T} \mu$ by Theorem 3.6D.

$$\therefore E(z) = \begin{pmatrix} 2 & -3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} = 8$$

The variance of z is $var(z) = var(\mathbf{a^T y}) = \mathbf{a^T} \Sigma \mathbf{a}$ by Theorem 3.6C.

$$\therefore var(z) = \begin{pmatrix} 2 & -3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 3 & 10 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} = 2$$

(b) Let $z_1 = y_1 + y_2 + y_3$ and $z_2 = 3y_1 + y_2 - 2y_3$. Find $E(\mathbf{z})$ and $\mathrm{cov}(\mathbf{z})$, where $\mathbf{z} = (z_1, z_2)^{\mathbf{T}}$

Sol. Since z_1 and z_2 are linear combinations of $\mathbf{y}$ with constant coefficent vectors, $\mathbf{a_1} = (1, 1, 1)^{\mathbf{T}}$ and $\mathbf{a_2} = (3, 1, 2)^{\mathbf{T}}$, respectively. Let $\mathbf{A}$ be a 2×3 matrix consisting of $\mathbf{a_1^T}$ and $\mathbf{a_2^T}$. Then

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 1 & -2 \end{pmatrix}$$

By Theorem 3.6D, $E(\mathbf{z})$ and $\text{cov}(\mathbf{z})$ can be calculated as

$$\begin{aligned} E(\mathbf{z}) &= \mathbf{A}\mu = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 1 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ -4 \end{pmatrix} \\ \text{cov}(\mathbf{z}) &= \mathbf{A}\Sigma\mathbf{A^T} = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 3 & 10 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 1 & 1 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} 21 & -14 \\ -14 & 45 \end{pmatrix} \end{aligned}$$

3.21 Let $\mathbf{y}$ be a random vector and covarance matrix μ and Σ as given in Problem 3.19 and define $\mathbf{w} = (w_1, w_2, w_3)^{\mathbf{T}}$ as follows:

$$\begin{aligned} w_1 &= 2y_1 - y_2 + y_3 \\ w_2 &= y_1 + 2y_2 - 3y_3 \\ w_3 &= y_1 + y_2 + 2y_3 \end{aligned}$$

(a) Find $E(\mathbf{w})$ and $\text{cov}(\mathbf{w})$.

Sol. Define a matrix $\mathbf{B}$ as,

$$\mathbf{B} = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & -3 \\ 1 & 1 & 2 \end{pmatrix}$$

Then $E(\mathbf{w})$ and $\text{cov}(\mathbf{w})$ are calculated as

$$\begin{aligned} E(\mathbf{w}) &= E(\mathbf{By}) = \mathbf{B}E(\mathbf{y}) = \mathbf{B}\boldsymbol{\mu} \\ &= \begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & -3 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ -10 \\ 6 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \text{cov}(\mathbf{w}) &= \text{cov}(\mathbf{By}) = \mathbf{B\Sigma B^T} \\ &= \begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & -3 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 3 & 10 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ -1 & 2 & 1 \\ 1 & -3 & 2 \end{pmatrix} = \begin{pmatrix} 6 & -14 & -18 \\ -14 & 67 & -49 \\ 18 & -49 & 57 \end{pmatrix} \end{aligned}$$

(b) Using $\mathbf{z}$ as defined in Problem 3.19(b), find $\text{cov}(\mathbf{z}, \mathbf{w})$.

Sol. By Theorem 3.6D,

$$\begin{aligned} \text{cov}(\mathbf{z}, \mathbf{w}) &= \text{cov}(\mathbf{Ay}, \mathbf{By}) = \mathbf{A\Sigma B^T} \\ &= \begin{pmatrix} 1 & 1 & 1 \\ 3 & 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 3 & 10 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ -1 & 2 & 1 \\ 1 & -3 & 2 \end{pmatrix} = \begin{pmatrix} 11 & -25 & 34 \\ -8 & 53 & -31 \end{pmatrix} \end{aligned}$$

Linear Models in Statistics: HW2

201060072: Boncho Ku

4.16 Suppose $\mathbf{y}$ is $N_4(\mu, \Sigma)$, where

$$\mu = \begin{pmatrix} 1 \\ 2 \\ 3 \\ -2 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 4 & 2 & -1 & 2 \\ 2 & 6 & 3 & 2 \\ -1 & 3 & 5 & -4 \\ 2 & -2 & -4 & 4 \end{pmatrix}$$

(a) The joint marginal distribution of y_1 and y_3

Sol. Let $\mathbf{a_1} = (1\ 0\ 1\ 0)^T$. Then joint marginal distribution of y_1 and y_3 can be expressed as $\mathbf{z_1} = \mathbf{a_1}^T\mathbf{y}$. The mean and variance of $\mathbf{z_1}$ is

$$\begin{aligned}
E(\mathbf{z_1}) &= E(\mathbf{a_1}^T\mathbf{y}) = \mathbf{a_1}^T\mu \\
&= \begin{pmatrix} 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\
\mathrm{cov}(\mathbf{z_1}) &= \mathbf{a_1}^T\Sigma\mathbf{a_1} = \begin{pmatrix} 4 & -1 \\ -1 & 5 \end{pmatrix}
\end{aligned}$$

Therefore, joint marginal distribution of y_1 and y_3 is

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = N_2 \left[\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 & -1 \\ -1 & 5 \end{pmatrix} \right]$$

(b) The marginal distribution of y_2

Sol. Let $\mathbf{a_2} = (0\ 1\ 0\ 0)^T$. Then marginal distribution of y_2 is

$$y_2 \sim N(2, 6)$$

(c) The distribution of $z = y_1 + 2y_2 - y_3 + 3y_4$

Sol. Let $\mathbf{a_3} = (1\ 2\ -1\ 3)^T$. Then $z = \mathbf{a_3}^T \mathbf{y} \sim N(\mathbf{a_3}^T \mu, \mathbf{a_3}^T \Sigma \mathbf{a_3})$. Therefore, the distribution of z is expressed as

$$\mathbf{a_3}^T \mu = \begin{pmatrix} 1 & 2 & -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ -2 \end{pmatrix} = -4$$

$$\mathbf{a_3}^T \Sigma \mathbf{a_3} = \begin{pmatrix} 1 & 2 & -1 & 3 \end{pmatrix} \begin{pmatrix} 4 & 2 & -1 & 2 \\ 2 & 6 & 3 & 2 \\ -1 & 3 & 5 & -4 \\ 2 & -2 & -4 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -1 \\ 3 \end{pmatrix} = 79$$

$$\therefore z \sim N_1(-4,\ 79)$$

(d) The joint distribution of $z_1 = y_1 + y_2 - y_3 - y_4$ and $z_2 = -3y_1 + y_2 + 2y_3 - 2y_4$

Sol. Let $\mathbf{z_2} = (z_1\ z_2)^T$. Then the distribution of $\mathbf{z_2}$ $N(\mathbf{A}\mu, \mathbf{A}\Sigma\mathbf{A}^T)$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 & -1 \\ -3 & 1 & 2 & -2 \end{pmatrix}$$

The mean vector and covariance matrix of $\mathbf{z_2}$ are calculated as

$$\mathbf{A}\mu = \begin{pmatrix} 1 & 1 & -1 & -1 \\ -3 & 1 & 2 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 9 \end{pmatrix}$$

$$\mathbf{A}\Sigma\mathbf{A}^T = \begin{pmatrix} 1 & 1 & -1 & -1 \\ -3 & 1 & 2 & -2 \end{pmatrix} \begin{pmatrix} 4 & 2 & -1 & 2 \\ 2 & 6 & 3 & 2 \\ -1 & 3 & 5 & -4 \\ 2 & -2 & -4 & 4 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ 1 & 1 \\ -1 & 2 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 11 & -6 \\ -6 & 154 \end{pmatrix}$$

Therefore, the joint distribution of z_1 and z_2 is

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = N_2 \left[\begin{pmatrix} 2 \\ 9 \end{pmatrix}, \begin{pmatrix} 11 & -6 \\ -6 & 154 \end{pmatrix} \right]$$

(e) $f(y_1,\ y_2 | y_3,\ y_4)$

Sol. The vector $\mathbf{y}$ can be partitioned with $\mathbf{v}$ and $\mathbf{w}$, where $\mathbf{v} = (y_1\ y_2)^T$ and

$\mathbf{w} = (y_3\ y_4)^T$. Then

$$\boldsymbol{\mu}_{\mathbf{v}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix},\ \boldsymbol{\mu}_{\mathbf{w}} = \begin{pmatrix} 3 \\ -2 \end{pmatrix},\ \Sigma_{\mathbf{v}} = \begin{pmatrix} 4 & 2 \\ 2 & 6 \end{pmatrix},\ \Sigma_{\mathbf{w}} = \begin{pmatrix} 5 & -4 \\ -4 & 4 \end{pmatrix},\ \Sigma_{\mathbf{vw}} = \begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix}$$

Now, by Theorem 4.4D we obtain

$$\begin{aligned}
E(\mathbf{v}|\mathbf{w}) &= \boldsymbol{\mu}_{\mathbf{v}} + \boldsymbol{\Sigma}_{\mathbf{vw}}\boldsymbol{\Sigma}_{\mathbf{w}}^{-1}(\mathbf{w} - \boldsymbol{\mu}_{\mathbf{w}}) \\
&= \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 5/4 \end{pmatrix}\left(\begin{pmatrix} y_3 \\ y_4 \end{pmatrix} - \begin{pmatrix} 3 \\ -2 \end{pmatrix}\right) \\
&= \begin{pmatrix} y_3 + 3/2y_4 + 1 \\ y_3 + 1/2y_4 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
\operatorname{cov}(\mathbf{v}|\mathbf{w}) &= \boldsymbol{\Sigma}_{\mathbf{v}} - \boldsymbol{\Sigma}_{\mathbf{vw}}\boldsymbol{\Sigma}_{\mathbf{w}}^{-1}\boldsymbol{\Sigma}_{\mathbf{wv}} \\
&= \begin{pmatrix} 4 & 2 \\ 2 & 6 \end{pmatrix} - \begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 5/4 \end{pmatrix}\begin{pmatrix} -1 & 3 \\ 2 & -2 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}
\end{aligned}$$

Therefore,

$$f(y_1,\ y_2|y_3,\ y_4) \sim N_2\left[\begin{pmatrix} y_3 + 3/2y_4 + 1 \\ y_3 + 1/2y_4 \end{pmatrix},\ \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}\right]$$

(f) $f(y_1,\ y_3|y_2,\ y_4)$

Sol. The vector $\mathbf{y}$ can be partitioned with $\mathbf{s}$ and $\mathbf{t}$, where $\mathbf{s} = (y_1\ y_3)^T$ and $\mathbf{t} = (y_2\ y_4)^T$. Then

$$\mu_{\mathbf{s}} = \begin{pmatrix} 1 \\ 3 \end{pmatrix},\ \mu_{\mathbf{t}} = \begin{pmatrix} 2 \\ -2 \end{pmatrix},\ \Sigma_{\mathbf{s}} = \begin{pmatrix} 4 & -1 \\ -1 & 5 \end{pmatrix},\ \Sigma_{\mathbf{t}} = \begin{pmatrix} 6 & -2 \\ -2 & 4 \end{pmatrix},\ \Sigma_{\mathbf{st}} = \begin{pmatrix} 2 & 2 \\ 3 & -4 \end{pmatrix}$$

Now, by Theorem 4.4D we obtain

$$\begin{aligned} E(\mathbf{s}|\mathbf{t}) &= \mu_{\mathbf{s}} + \boldsymbol{\Sigma}_{\mathbf{st}}\boldsymbol{\Sigma}_{\mathbf{t}}^{-1}(\mathbf{t} - \mu_{\mathbf{t}}) \\ &= \begin{pmatrix} 1 \\ 3 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ 3 & -4 \end{pmatrix} \begin{pmatrix} 2/10 & 1/10 \\ 1/10 & 3/10 \end{pmatrix} \begin{pmatrix} y_2 - 2 \\ y_4 + 2 \end{pmatrix} \\ &= \begin{pmatrix} 3/5y_2 + 4/5y_4 + 7/5 \\ 1/5y_2 - 9/10y_4 + 4/5 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \mathrm{cov}(\mathbf{s}|\mathbf{t}) &= \boldsymbol{\Sigma}_{\mathbf{s}} - \boldsymbol{\Sigma}_{\mathbf{st}}\boldsymbol{\Sigma}_{\mathbf{t}}^{-1}\boldsymbol{\Sigma}_{\mathbf{ts}} \\ &= \begin{pmatrix} 4 & -1 \\ -1 & 5 \end{pmatrix} - \begin{pmatrix} 2 & 2 \\ 3 & -4 \end{pmatrix} \begin{pmatrix} 2/10 & 1/10 \\ 1/10 & 3/10 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 2 & -4 \end{pmatrix} \\ &= \begin{pmatrix} 6/5 & 2/5 \\ 2/5 & 4/5 \end{pmatrix} \end{aligned}$$

Therefore,

$$f(y_1,\ y_3|y_2,\ y_4)\ \sim\ N_2\left[\begin{pmatrix} 3/5y_2+4/5y_4+7/5 \\ 1/5y_2-9/10y_4+4/5 \end{pmatrix},\ \begin{pmatrix} 6/5 & 2/5 \\ 2/5 & 4/5 \end{pmatrix}\right]$$

(g) ρ_{13}

Sol.

$$\rho_{13} = -\frac{1}{\sqrt{4}\sqrt{5}} = -\frac{\sqrt{5}}{10}$$

(h) $\rho_{13\cdot 24}$

Sol. From the result of (f),

$$\rho_{13\cdot 24} = \frac{2/5}{\sqrt{6/5}\sqrt{4/5}} = \frac{1}{\sqrt{6}}$$

(i) $f(y_1|y_2,\ y_3,\ y_4)$

Sol. Let $\mathbf{x} = (y_2\ y_3\ y_4)^T$. Then

$$\mu_{y_1} = 1,\ \boldsymbol{\mu}_{\mathbf{x}} = \begin{pmatrix} 2 \\ 3 \\ -2 \end{pmatrix},\ \sigma^2_{y_1} = 4,\ \boldsymbol{\Sigma}_{\mathbf{x}} = \begin{pmatrix} 6 & 3 & -2 \\ 3 & 5 & -4 \\ -2 & -4 & 4 \end{pmatrix},\ \boldsymbol{\Sigma}_{y_1,\mathbf{x}} = \begin{pmatrix} 2 & -1 & 2 \end{pmatrix}$$

Now, conditional mean and variance of y_1 given $\mathbf{x}$ is

$$\begin{aligned}
E(y_1|\mathbf{x}) &= \mu_{y_1} + \boldsymbol{\Sigma}_{y_1,\mathbf{x}}\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x}}) \\
&= 1 + \begin{pmatrix} 2 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1/4 & -1/4 & -1/8 \\ -1/4 & 5/4 & 9/8 \\ -1/8 & 9/8 & 21/16 \end{pmatrix} \begin{pmatrix} y_2 - 2 \\ y_3 - 3 \\ y_4 + 2 \end{pmatrix} \\
&= \frac{2y_2 + 2y_3 + 5y_4}{4} + 1 \\
\operatorname{cov}(y_1|\mathbf{x}) &= \sigma_{y_1}^2 - \boldsymbol{\Sigma}_{y_1,\mathbf{x}}\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\boldsymbol{\Sigma}_{\mathbf{x},y_1} \\
&= 4 - \begin{pmatrix} 2 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1/4 & -1/4 & -1/8 \\ -1/4 & 5/4 & 9/8 \\ -1/8 & 9/8 & 21/16 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} \\
&= 1 \\
\therefore\ & f(y_1|y_2,\ y_3,\ y_4) \sim N(\frac{2y_2 + 2y_3 + 5y_4}{4} + 1, 1)
\end{aligned}$$

5.26 Suppose $\mathbf{y}$ is $N_n(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = \mu\mathbf{j}$ and

$$\boldsymbol{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}$$

Thus $E(y_i) = \mu$ for all i, $\text{var} y_i = \sigma^2$ for all i, and $\operatorname{cov}(y_i,\ y_j) = \sigma^2\rho$ for all $i \neq j$; that is, the y's are equicorrelated.

(a) Show that $\boldsymbol{\Sigma}$ can be written in the form of $\boldsymbol{\Sigma} = \sigma^2[(1-\rho)\mathbf{I} + \rho\mathbf{J}]$.

Sol.

$$\begin{aligned}\mathbf{\Sigma} &= \sigma^2[\mathbf{I} + \rho\mathbf{J} - \rho\mathbf{I}] \\ &= \sigma^2[(1-\rho)\mathbf{I} + \rho\mathbf{J}]\end{aligned}$$

(b) Show that $\sum_{i=1}^{n}(y_i - \bar{y})^2/[\sigma^2(1-\rho)]$ is $\chi^2(n-1)$.

Sol. Since $\mathbf{y} \sim N_n(\boldsymbol{\mu},\ \mathbf{\Sigma})$ and

$$\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{\sigma^2(1-\rho)} = \frac{\mathbf{y}^T(\mathbf{I_n} - \mathbf{P})\mathbf{y}}{\sigma^2(1-\rho)}$$

where $\mathbf{I}_n$ is $n \times n$ identity matrix and $\mathbf{P} = 1/n\mathbf{J}_n$. Let $\mathbf{A} = \mathbf{I_n} - \mathbf{P}$, then $\mathbf{A\Sigma}$ is idempotent matrix, since

$$\begin{aligned}\mathbf{A\Sigma} &= \frac{\sigma^2}{\sigma^2(1-\rho)}(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n)[(1-\rho)\mathbf{I}_n + \rho\mathbf{I}_n] \\ &= \frac{1}{1-\rho}[(1-\rho)\mathbf{I}_n + (\rho - \frac{1}{n} + \frac{\rho}{n} - \frac{n\rho}{n})\mathbf{J}_n] \\ &= \frac{1}{1-\rho}[(1-\rho)\mathbf{I}_n - \frac{(1-\rho)}{n}\mathbf{J}_n] \\ &= \mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\end{aligned}$$

Therefore, the rank of $\mathbf{A}$, $r = n - 1$ by Theorem 2.13D. To find λ, which is given by

$$\lambda = \frac{\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}}{2\sigma^2(1-\rho)} = 0$$

By Theorem 5.5A,

$$\frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{[\sigma^2(1-\rho)]} \quad \sim \quad \chi^2(n-1)$$

5.29 if $\mathbf{y}$ is $N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = \begin{pmatrix} 3 \\ -2 \\ 1 \\ 4 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & -4 \\ 0 & 0 & -4 & 6 \end{pmatrix}$$

Find a matrix $\mathbf{A}$ such that $\mathbf{y}^T\mathbf{A}\mathbf{y}$ is $\chi^2(4,\ \frac{1}{2}\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu})$. What is $\lambda = \frac{1}{2}\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}$?

Sol. To suffice $\mathbf{y}^T\mathbf{A}\mathbf{y} \sim \chi^2(4,\ \frac{1}{2}\boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu})$, $\mathbf{A}\boldsymbol{\Sigma}$ has to be an idempotent matrix such that $(\mathbf{A}\boldsymbol{\Sigma})(\mathbf{A}\boldsymbol{\Sigma}) = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma} = \mathbf{A}\boldsymbol{\Sigma}$ with 4×4 full rank symmetric matrix $\mathbf{A}$. To find $\mathbf{A}$,

$$\mathbf{A} = \mathbf{A}(\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma})(\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma})^{-1} = \mathbf{A}\boldsymbol{\Sigma}(\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma})^{-1} = \boldsymbol{\Sigma}^{-1}$$

$$\therefore \lambda \;=\; \frac{1}{2}\begin{pmatrix} 3 & -2 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 2 & \frac{3}{2} \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 1 \\ 4 \end{pmatrix} = 27$$

5.30 Suppose $\mathbf{y}$ is $N_3(\mu, \sigma^2\mathbf{I})$ and let

$$\mu = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}, \quad \mathbf{A} = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$$

(a) What is the distribution of $\mathbf{y}^T\mathbf{A}\mathbf{y}/\sigma^2$?

Sol. To verify $\mathbf{A}$ is idempotent,

$$\mathbf{A}^2 = \mathbf{A}\mathbf{A} = \frac{1}{9}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

$\text{rank}(\mathbf{A}) = \text{tr}(\mathbf{A}) = \frac{2+2+2}{3} = 2$. Therefore, $\mathbf{y}^T\mathbf{A}\mathbf{y}/\sigma^2 \sim \chi^2(2,\ \mu^T\mathbf{A}\mu/2\sigma^2)$, where

$$\frac{1}{2}\mu^T\mathbf{A}\mu = \frac{1}{6}\begin{pmatrix} 3 & -2 & 1 \end{pmatrix}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} = \frac{38}{6}$$

(b) Are $\mathbf{y}^T\mathbf{A}\mathbf{y}$ and $\mathbf{B}\mathbf{y}$ independent?

Sol.

$$\begin{aligned} \mathbf{B}\mathbf{A} &= \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix} \\ &\neq \mathbf{O} \end{aligned}$$

By Theorem 5.6A, therefore, $\mathbf{y}^T\mathbf{A}\mathbf{y}$ and $\mathbf{B}\mathbf{y}$ are **not independent**.

(c) Are $\mathbf{y}^T\mathbf{A}\mathbf{y}$ and $y_1 + y_2 + y_3$ independent?

Sol. Let $\mathbf{a} = (1\ 1\ 1)^T$, then $y_1+y_2+y_3 = \mathbf{a}^T\mathbf{y}$. To verify the independence between $\mathbf{y}^T\mathbf{A}\mathbf{y}$ and $\mathbf{a}^T\mathbf{y}$,

$$\begin{aligned} \mathbf{a}^T\mathbf{A} &= \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} \\ &= \mathbf{O} \end{aligned}$$

Therefore, $\mathbf{y}^T\mathbf{A}\mathbf{y}$ and $\mathbf{a}^T\mathbf{y}$ are independent.

Linear Models in Statistics: HW3

201060072: Boncho Ku

7.23 Show that $(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) = (\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})^T(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}) + (\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})^T\mathbf{X}^T\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ as in (7.53) in the proof of Theorem7.6C.

Sol.

$$\begin{aligned}
(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) &= (\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}+\mathbf{X}\hat{\boldsymbol{\beta}}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}+\mathbf{X}\hat{\boldsymbol{\beta}}-\mathbf{X}\boldsymbol{\beta}) \\
&= [(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})^T-(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^T\mathbf{X}^T][(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})-\mathbf{X}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})] \\
&= (\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})^T(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}) + (\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^T\mathbf{X}^T\mathbf{X}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}) - \\
&\quad (\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^T\mathbf{X}^T(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}) - (\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})^T\mathbf{X}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})
\end{aligned}$$

The last two terms of the above equation can be written as

$$\begin{aligned}
-(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^T(\mathbf{X}^T\mathbf{y}-\mathbf{X}^T\mathbf{X}\hat{\boldsymbol{\beta}}) &- [\mathbf{X}^T(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})]^T(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}) = \mathbf{0} \\
(\because \mathbf{X}^T\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^T\mathbf{y}) &
\end{aligned}$$

$$\therefore (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) = (\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})^T(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}) + (\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})^T\mathbf{X}^T\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$$

7.51 Show that $\mathbf{X}_{2\cdot 1} = \mathbf{X}_2 - \hat{\mathbf{X}}_2(\mathbf{X}_1)$ is orthogonal to $\mathbf{X}_1$, that is, $\mathbf{X}_1^T\mathbf{X}_{2\cdot 1} = \mathbf{O}$, as in (7.98).

Sol. Using $\mathbf{X}_{2\cdot1} = \mathbf{X}_2 - \mathbf{X}_1\mathbf{A}$ where $\mathbf{A} = (\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{X}_2$,

$$\begin{aligned}
\mathbf{X}_1^T\mathbf{X}_{2\cdot1} &= \mathbf{X}_1^T[\mathbf{X}_2 - \mathbf{X}_1(\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{X}_2] \\
&= \mathbf{X}_1^T\mathbf{X}_2 - (\mathbf{X}_1^T\mathbf{X}_1)(\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{X}_2 \\
&= \mathbf{X}_1^T\mathbf{X}_2 - \mathbf{X}_1^T\mathbf{X}_2 \\
&= \mathbf{O} \\
&\therefore\ \mathbf{X}_{2\cdot1} \perp \mathbf{X}_1
\end{aligned}$$

7.52 Show that $\hat{\boldsymbol{\beta}}_2$ in (7.101) is the same as in the full fitted model $\hat{\mathbf{y}} = \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{X}_2\hat{\boldsymbol{\beta}}_2$.

Sol. Let $\mathbf{X} = [\mathbf{X}_1\ \ \mathbf{X}_2]$ and $\hat{\boldsymbol{\beta}} = [\hat{\boldsymbol{\beta}}_1\ \ \hat{\boldsymbol{\beta}}_2]$, then normal equation $\mathbf{X}^T\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^T\mathbf{y}$ is expressed as

$$\begin{aligned}
\begin{pmatrix} \mathbf{X}_1^T \\ \mathbf{X}_2^T \end{pmatrix} \begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\beta}}_1 \\ \hat{\boldsymbol{\beta}}_2 \end{pmatrix} &= \begin{pmatrix} \mathbf{X}_1^T \\ \mathbf{X}_2^T \end{pmatrix} \mathbf{y} \\
\begin{pmatrix} \mathbf{X}_1^T\mathbf{X}_1 & \mathbf{X}_1^T\mathbf{X}_2 \\ \mathbf{X}_2^T\mathbf{X}_1 & \mathbf{X}_2^T\mathbf{X}_2 \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\beta}}_1 \\ \hat{\boldsymbol{\beta}}_2 \end{pmatrix} &= \begin{pmatrix} \mathbf{X}_1^T\mathbf{y} \\ \mathbf{X}_2^T\mathbf{y} \end{pmatrix}
\end{aligned}$$

The above matrix becomes

$$\begin{aligned}
\mathbf{X}_1^T\mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{X}_1^T\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 &= \mathbf{X}_1^T\mathbf{y} && (1) \\
\mathbf{X}_2^T\mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{X}_2^T\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 &= \mathbf{X}_2^T\mathbf{y} && (2)
\end{aligned}$$

From (1),

$$\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}_1^T\mathbf{X}_1)^{-1}[\mathbf{X}_1^T\mathbf{y} - \mathbf{X}_1^T\mathbf{X}_2\hat{\boldsymbol{\beta}}_2] \tag{3}$$

and subsititute (3) into (2),

$$\mathbf{X}_2^T\mathbf{X}_{2\cdot 1}\hat{\boldsymbol{\beta}}_2 = \mathbf{X}_2^T\mathbf{y} - \mathbf{X}_2^T\mathbf{X}_1(\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{y} \tag{4}$$

where $\mathbf{X}_{2\cdot 1} = \mathbf{X}_2 - \mathbf{X}_1\mathbf{A},\ \mathbf{A} = (\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{X}_2$. Multiplying $\mathbf{X}_2^T$ to (7.101),

$$\mathbf{X}_2^T\mathbf{X}_{2\cdot 1}\hat{\boldsymbol{\beta}}_2 = \mathbf{X}_2^T[\hat{\mathbf{y}}(\mathbf{X}_1,\ \mathbf{X}_2) - \hat{\mathbf{y}}(\mathbf{X}_1)] \tag{5}$$

where $\hat{\mathbf{y}}(\mathbf{X}_1,\ \mathbf{X}_2) = \hat{\mathbf{y}}$ and $\hat{\mathbf{y}}(\mathbf{X}_1) = \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{X}_1\mathbf{A}\hat{\boldsymbol{\beta}}_2$. To verify that (4) is equal to (5),

$$\begin{aligned}
\mathbf{X}_2^T\mathbf{X}_{2\cdot 1}\hat{\boldsymbol{\beta}}_2 &= \mathbf{X}_2^T[\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 - \mathbf{X}_1\mathbf{A}\hat{\boldsymbol{\beta}}_2] \\
&= \mathbf{X}_2^T[\mathbf{X}_2(\mathbf{X}_2^T\mathbf{X}_{2\cdot 1})^{-1}\mathbf{X}_2^T\mathbf{y} - \mathbf{X}_2(\mathbf{X}_2^T\mathbf{X}_{2\cdot 1})^{-1}\mathbf{X}_1(\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{y} - \mathbf{X}_1\mathbf{A}\hat{\boldsymbol{\beta}}_2]
\end{aligned}$$

Since $(\mathbf{X}_2^T\mathbf{X}_{2\cdot 1} + \mathbf{X}_2^T\mathbf{X}_1\mathbf{A})\hat{\boldsymbol{\beta}}_2 = \mathbf{X}_2^T\mathbf{X}_2\hat{\boldsymbol{\beta}}_2$,

$$\begin{aligned}
\mathbf{X}_2^T\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 &= \mathbf{X}_2^T\mathbf{X}_2(\mathbf{X}_2^T\mathbf{X}_{2\cdot 1})^{-1}[\mathbf{X}_2^T\mathbf{y} - \mathbf{X}_2^T\mathbf{X}_1(\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{y}] \\
\hat{\boldsymbol{\beta}}_2 &= (\mathbf{X}_2^T\mathbf{X}_{2\cdot 1})^{-1}[\mathbf{X}_2^T\mathbf{y} - \mathbf{X}_2^T\mathbf{X}_1(\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{y}] \\
\mathbf{X}_2^T\mathbf{X}_{2\cdot 1}\hat{\boldsymbol{\beta}}_2 &= \mathbf{X}_2^T\mathbf{y} - \mathbf{X}_2^T\mathbf{X}_1(\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{y}
\end{aligned}$$

Which is the same as (4). Therefore, $\hat{\boldsymbol{\beta}}_2$ is the same as $\hat{\boldsymbol{\beta}}_2$ in the full fitted model.

7.53 When gasoline is pumped into the tank of a car, vapors are vented into the atmosphere. An experiment was conducted to determine whether y, the amount of vapor, can be predicted using the following four variables based on initial conditions of the tank and the dispensed gasoline:

$$\begin{aligned}
x_1 &= \text{tank temperature} \quad (^\circ\text{F}), \\
x_2 &= \text{gasoline temperature} \quad (^\circ\text{F}), \\
x_3 &= \text{vapor pressure in tank} \quad (\text{psi}), \\
x_4 &= \text{vapor pressure of gasoline} \quad (\text{psi}).
\end{aligned}$$

(a) Find $\hat{\boldsymbol{\beta}}$ and s^2.

(b) Find an estimate of $\text{cov}(\hat{\boldsymbol{\beta}})$

(c) Find $\hat{\boldsymbol{\beta}}_1$ and $\hat{\beta}_0$ using $\mathbf{S}_{xx}$ and $\mathbf{S}_{yx}$ as in (7.47) and (7.48).

(d) Find R^2 and R_a^2.