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

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

(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 $\mathbf{a}^{\mathbf{T}}\mathbf{y}$ and $E(z) = E(\mathbf{a}^{\mathbf{T}}\mathbf{y}) = \mathbf{a}^{\mathbf{T}}E(\mathbf{y}) = \mathbf{a}^{\mathbf{T}}\mu$ 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 $var(z) = var(\mathbf{a}^{\mathbf{T}}\mathbf{y}) = \mathbf{a}^{\mathbf{T}}\Sigma\mathbf{a}$ 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

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

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

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

(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

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

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

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

Now, by Theorem 4.4D we obtain

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

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

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

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

$$\begin{aligned} \text{cov}(\mathbf{s}|\mathbf{t}) &= \mathbf{\Sigma_s} - \mathbf{\Sigma_{st}\Sigma_t^{-1}\Sigma_{ts}} \\ &= \begin{pmatrix} 4 & -1 \\ -1 & 5 \end{pmatrix} - \begin{pmatrix} 2 & 2 \\ 3 & -4 \end{pmatrix} \begin{pmatrix} 2/10 & 1/10 \\ 1/10 & 3/10 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 2 & -4 \end{pmatrix} \\ &= \begin{pmatrix} 6/5 & 2/5 \\ 2/5 & 4/5 \end{pmatrix} \end{aligned}$$

Therefore,

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

(g) ρ_{13}

Sol.

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

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

Sol. From the result of (f),

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

(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

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

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, $\mathrm{var} y_i = \sigma^2$ for all i, and $\mathrm{cov}(y_i,\ y_j) = \sigma^2\rho$ 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.

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

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

Sol. Since $\mathbf{y} \sim N_n(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$ 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 $\mathbf{A}\boldsymbol{\Sigma}$ is idempotent matrix, since

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

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

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

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

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

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

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

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

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

$$\mathbf{X}_1^T\mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{X}_1^T\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 = \mathbf{X}_1^T\mathbf{y} \tag{1}$$

$$\mathbf{X}_2^T\mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{X}_2^T\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 = \mathbf{X}_2^T\mathbf{y} \tag{2}$$

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.

Sol. Using $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$ and $s^2 = 1/(n-k-1)(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})^T(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})$, The calculation results were follows:

```
> model<-y~x1+x2+x3+x4
> utils::str(m<-model.frame(model,Data))
> X<-model.matrix(model,m)
> size<-dim(X)
> n<-size[1]
> k<-size[2]-1
> beta<-solve(t(X)%*%X)%*%t(X)%*%y
> sigma.hat<-(n-k-1)^(-1)*t(y-X%*%beta)%*%(y-X%*%beta)
```

Therefore,

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} 1.015 \\ -0.029 \\ 0.216 \\ -4.320 \\ 8.975 \end{pmatrix}, \qquad s^2 = 7.453$$

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

Sol. Since $\text{cov}(\hat{\boldsymbol{\beta}}) = s^2(\mathbf{X}^T\mathbf{X})^{-1}$,

```
> XX.inv<-solve(t(X)%*%X)
> cov.beta<-as.numeric(sigma.hat)*XX.inv
```

Therefore,

$$\text{cov}(\hat{\boldsymbol{\beta}}) = \begin{pmatrix} 3.464 & 0.014 & -0.064 & -1.162 & 1.072 \\ 0.014 & 0.008 & -0.002 & -0.163 & 0.078 \\ -0.064 & -0.002 & 0.005 & 0.104 & -0.125 \\ -1.162 & -0.163 & 0.104 & 8.128 & -7.204 \\ 1.072 & 0.078 & -0.125 & -7.204 & 7.687 \end{pmatrix}$$

(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).

Sol. $\hat{\boldsymbol{\beta}}_1 = \mathbf{S}_{xx}^{-1}\mathbf{s}_{yx}$, where $\mathbf{S}_{xx} = \mathbf{X}_c^T\mathbf{X}_c/(n-1)$, $\mathbf{s}_{yx} = \mathbf{X}_c^T\mathbf{y}/(n-1)$, $\mathbf{X}_c = (\mathbf{I} - n^{-1}\mathbf{J})\mathbf{X}_1$ and $\hat{\beta}_0 = \bar{y} - \mathbf{s}_{yx}^T\mathbf{S}_{xx}^{-1}\bar{\mathbf{x}}$. R code is following:

```
> I.<-diag(rep(1,n))
> J<-matrix(1,n,n)
```

```
> X.c<-(I.-n^(-1)*J)%*%X[,2:(k+1)]
> S.xx<-t(X.c)%*%X.c/(n-1)
> s.yx<-t(X.c)%*%y/(n-1)
> barx<-apply(X[,2:(k+1)],2,mean)
> beta.1<-solve(S.xx)%*%s.yx
> beta.0<-mean(y)-t(beta.1)%*%barx
```

Therefore,

$$\hat{\boldsymbol{\beta}}_1 = \begin{pmatrix} -0.029 \\ 0.216 \\ -4.320 \\ 8.975 \end{pmatrix}, \qquad \hat{\beta}_0 = 1.015$$

which is the same result of (a).

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

Sol. Using (7.56) and (7.59) in the textbook, R code is following:

```
> SSE<-t(y-X%*%beta)%*%(y-X%*%beta)
> SST<-t(y)%*%y-size[1]*mean(y)^2
> SSR<-SST-SSE
> R.sq<-SSR/SST
> R.sq.adj<-((n-1)*R.sq-k)/(n-k-1)
```

Therefore,

$$R^2 = 0.926, \qquad R_a^2 = 0.915$$

In addition, the result of `lm()` for the data is following:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.0150	1.8613	0.55	0.5900
x1	-0.0286	0.0906	-0.32	0.7546
x2	0.2158	0.0677	3.19	0.0036
x3	-4.3201	2.8510	-1.52	0.1413
x4	8.9749	2.7726	3.24	0.0032

Table 1. Output table of regression: Residual standard error: 2.73 on 27 degrees of freedom Multiple R-squared: 0.9261, Adjusted R-squared: 0.9151 F-statistic: 84.54 on 4 and 27 DF, p-value: 7.249e-15

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
x1	1	1857.11	1857.11	249.18	0.0000
x2	1	494.43	494.43	66.34	0.0000
x3	1	90.63	90.63	12.16	0.0017
x4	1	78.09	78.09	10.48	0.0032
Residuals	27	201.23	7.45		

Table 2. ANOVA table of regression

Linear Models in Statistics: HW4

201060072: Boncho Ku

8.12 Find the expected mean square corresponding to the numerator of the F-statistic in (8.20) in Example 8.2(b).

Sol. By Theorem 8.2D,

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

where $\boldsymbol{\beta}_1$ is $(k-h-1)\times 1$ and $\boldsymbol{\beta}_2$ is $h\times 1$ vector, respectively, and $\mathbf{X}_1$ is $n\times(k-h-1)$ and $\mathbf{X}_2$ is $n \times h$ matrix, respectively. By using Theorem 5.2A and the proof of Theorem 7.9C(i), the expected mean square is expressed as

$$E[SS(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)] = \sigma^2 h + \boldsymbol{\beta}_2^T[\mathbf{X}_2^T\mathbf{X}_2 - \mathbf{X}_2^T\mathbf{X}_1(\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{X}_2]\boldsymbol{\beta}_2$$

Here in Example 8.20(b), substituting $h = 1,\ \boldsymbol{\beta}_2 = \beta_k,\ \mathbf{X}_2 = \mathbf{x}_k : (n \times 1)$ to the above equation,

$$\therefore E[SS(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)] = \sigma^2 + \beta_k^2[\mathbf{x}_k^T\mathbf{x}_k - \mathbf{x}_k^T\mathbf{X}_1(\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{x}_k]$$

8.30 (a) If $\bar{y}_0 = \sum_{i=1}^q y_{0i}/q$ is the mean of q future observations at $\mathbf{x}_0$, show that a $100(1-\alpha)\%$ prediction interval for $\bar{y}_0$ is given by

$$\mathbf{x}_0^T\hat{\boldsymbol{\beta}} \pm t_{\alpha/2,n-k-1}s\sqrt{1/q + \mathbf{x}_0^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0}$$

Sol. Let $\hat{y}_0$ be the predicted value for y_0 by $\mathbf{x}_0^T\hat{\boldsymbol{\beta}}$, and y_0 and $\hat{y}_0$ are independent, since $\bar{y}_0$ and $\hat{y}_0$ are also independent. $\bar{y}_0$ can be shown as $\bar{y}_0 = \sum_{i=1}^q y_{0i}/q = \sum_{i=1}^q(\mathbf{x}_0^T\boldsymbol{\beta} + \epsilon_{0i})/q$. To find the expectation and variance of $\bar{y}_0 - \hat{y}_0$,

$$
\begin{aligned}
E(\bar{y}_0 - \hat{y}_0) &= E(\bar{y}_0) - E(\hat{y}_0) \\
&= \frac{1}{q}E\big[\sum_{i=1}^q(\mathbf{x}_0^T\boldsymbol{\beta} + \epsilon_{0i})\big] - E(\mathbf{x}_0^T\hat{\boldsymbol{\beta}}) \\
&= \mathbf{x}_0^T\boldsymbol{\beta} + \frac{1}{q}E\big[\sum_{i=1}^q \epsilon_{0i}\big] - \mathbf{x}_0^T\boldsymbol{\beta} \\
&= \mathbf{x}_0^T\boldsymbol{\beta} + 0 - \mathbf{x}_0^T\boldsymbol{\beta} \\
&= 0 \\
\mathrm{var}(\bar{y}_0 - \hat{y}_0) &= \mathrm{var}(\bar{y}_0) + \mathrm{var}(\hat{y}_0) \\
&= \mathrm{var}\big[\frac{1}{q}\sum_{i=1}^q(\mathbf{x}_0^T\boldsymbol{\beta} + \epsilon_{0i})\big] + \mathrm{var}(\mathbf{x}_0^T\hat{\boldsymbol{\beta}}) \\
&= \frac{1}{q^2}\Big\{\mathrm{var}\big[\sum_{i=1}^q \mathbf{x}_0^T\boldsymbol{\beta}\big] + \mathrm{var}\big[\sum_{i=1}^q \epsilon_{0i}\big]\Big\} + \sigma^2\mathbf{x}_0^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0 \\
&= \frac{1}{q^2}q\sigma^2 + \sigma^2\mathbf{x}_0^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0 \\
&= \sigma^2(1/q + \mathbf{x}_0^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0)
\end{aligned}
$$

Since t statistics is distributed as

$$
t = \frac{\bar{y}_0 - \hat{y}_0}{s\sqrt{1/q + \mathbf{x}_0^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0}} \quad \sim \quad t_{n-k-1}
$$

Therefore, $100(1-\alpha)\%$ prediction interval is

$$
\mathbf{x}_0^T\hat{\boldsymbol{\beta}} \pm t_{\alpha/2,n-k-1}s\sqrt{1/q + \mathbf{x}_0^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0}
$$

(b) Show that for simple linear regression, the prediction interval for $\bar{y}_0$ in part (a) reduces to

$$\hat{\beta}_0 + \hat{\beta}_1 x_0 \pm t_{\alpha/2,n-2}\, s\sqrt{1/q + 1/n + (x_0 - \bar{x})^2 / \sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

Sol. For simple linear regression, $\boldsymbol{\beta}$, $\mathbf{x}_0$, and $\mathbf{X}$ can be expressed as,

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \quad \mathbf{x}_0 = (1\ x_0)^T, \quad \mathbf{X} = (\mathbf{j}\ \mathbf{x}), \quad \mathbf{j} = (1 \cdots 1)^T, \quad \mathbf{x} : n \times 1$$

From the result of Example 7.3.1(b),

$$(\mathbf{X}^T\mathbf{X})^{-1} = \frac{1}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} \begin{pmatrix} \sum_i x_i^2 & -\sum_i x_i \\ -\sum_i x_i & n \end{pmatrix}$$

Then,

$$\begin{aligned}
\mathbf{x}_0^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0 &= \frac{\sum x_i^2 - 1/n\left(\sum x_i\right)^2 + 1/n\left(\sum x_i\right)^2 - 2x_0\sum x_i + nx_0^2}{n\sum x_i^2 - \left(\sum x_i\right)^2} \\
&= \frac{\sum x_i^2 - n\bar{x}^2 + n\bar{x}^2 - 2x_0\sum x_i + nx_0^2}{n\left(\sum x_i^2 - n\bar{x}^2\right)} \\
&= \frac{\sum(x_i - \bar{x})^2 + n(x_0 - \bar{x})^2}{n\sum(x_i - \bar{x})^2} \\
&= \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum(x_i - \bar{x})^2}
\end{aligned}$$

Since $\mathbf{x}_0^T\hat{\boldsymbol{\beta}} = \hat{\beta}_0 + \hat{\beta}_1 x_0$, the result of (a) for simple linear regression ($k = 1$) is expressed as

$$\hat{\beta}_0 + \hat{\beta}_1 x_0 \pm t_{\alpha/2,n-2} s \sqrt{1/q + 1/n + (x_0 - \bar{x})^2 / \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

8.37 Use the gas vapor data in Table 7.3.

(a) Test the overall regression hypothesis $H_0: \boldsymbol{\beta}_1 = \mathbf{0}$ using (8.5) [or (8.22)] and (8.23).

Sol. From the result of 7.53 in HW3, ANOVA table for overall regression is

```
> Fval<-(SSR/k)/(SSE/(n-k-1))
> tmp<-anova(lm1)
> overall<-matrix(0,3,5)
> overall[,1]<-c(k,n-k-1,n-1)
> overall[,2]<-c(sum(tmp[1:4,2]),tmp[5,2],sum(tmp[,2]))
> overall[,3]<-c(sum(tmp[1:4,2])/4,tmp[5,2]/27,NA)
> overall[,4]<-c(Fval,NA,NA)
> overall[,5]<-c(1-pf(Fval,k,n-k-1),NA,NA)
> overall<-as.data.frame(overall)
> row.names(overall)<-c("Beta","Residual","Total")
> names(overall)<-names(tmp)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Beta	4.000	2520.272	630.068	84.540	0.000
Residual	27.000	201.228	7.453		
Total	31.000	2721.500			

Table 3. Result of ANOVA table for overall regression for vapor dataset

(b) Test H_0 : $\beta_1 = \beta_3 = 0$, that is, that x_1 and x_3 do not make a significant contribution above and beyond x_2 and x_4.

Sol. Test on subset of β's using Theorem 8.2C. Let $\boldsymbol{\beta}_2 = (\beta_1\ \beta_3)^T$ and $\boldsymbol{\beta}_1 = (\beta_2\ \beta_4)^T$. Then test H_0 : $\boldsymbol{\beta}_2 = 0$. The result of ANOVA table can be created by using the following scripts in R.

```
> h<-2
> lm2<-lm(y~x2+x4,data=Data)
> SSR.all<-overall[1,2]
> SSR.star<-sum(anova(lm2)[1:2,2])
> Fval.sub<-((SSR.all-SSR.star)/h)/(SSE/(n-k-1))
> p.sub<-1-pf(Fval.sub,h,n-k-1)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Due to $\boldsymbol{\beta}_2$ adjusted for $\boldsymbol{\beta}_1$	2	37.159	18.579	2.493	0.102
Residual	27	201.228	7.453		
Total	31	2721.500			

Table 4. Result of ANOVA table for $\boldsymbol{\beta}_2$ adjusted for $\boldsymbol{\beta}_1$

(c) Test H_0 : $\beta_j = 0$ for $j = 1, 2, 3, 4$ using t_j in (8.40). Use $t_{.05/2}$ for each test and also use a Bonferroni approach based on $t_{.05/8}$ (or compare the p-value to .05/4).

Sol. In the Table 1. from the result of 7.53 at HW3, t test results for each β_j, $j = 1, 2, 3, 4$ are shown. When we perform tests with α set at .05 and $.05/4 = .0125$, null hypotheses $\beta_2 = 0$ and $\beta_4 = 0$ were rejected.

(d) Using a general linear hypothesis approach, test H_0 : $\beta_1 = \beta_2 = 12\beta_3 = 12\beta_4$, H_{01} : $\beta_1 = \beta_2$, H_{02} : $\beta_2 = 12\beta_3$, H_{03} : $\beta_3 = \beta_4$, and H_{04} : $\beta_1 = \beta_2$ and $\beta 3 = \beta 4$.

Sol. Let $\mathbf{C}$ be a coefficient matrix conrresponding to the hypothesis H_0. Then $\mathbf{C}$ can be expressed as

$$\mathbf{C} = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -12 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

Using the `linearHypothesis()` function from the package `car` in R, general linear hypothesis test for $H_0 :\ \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$ can be performed easily.

```
> library(car)
> C<-matrix(c(0,0,0,1,0,0,-1,1,0,0,-12,1,0,0,-1),3,5)
> linearHypothesis(lm1,C,c(0,0,0))
```

The test result for H_0 is shown in Table 5.

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	30	437.55				
2	27	201.23	3	236.33	10.57	0.0001

Table 5. General linear hypothesis test result for $H_0 :\ \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$

Let $\mathbf{C} = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)^T$. For H_{01} to H_{03}, use each $\mathbf{c}_i$, $i = 1, 2, 3$ and test $H_{0i} :\ \mathbf{c}_i\boldsymbol{\beta} = 0$ separately. To perform tests, R scripts are as follows:

```
> linearHypothesis(lm1,C[1,],c(0))
> linearHypothesis(lm1,C[2,],c(0))
> linearHypothesis(lm1,C[3,],c(0))
```

The test results for $H_{0i} :\ \mathbf{c}_i\boldsymbol{\beta} = 0$ are shown in Table 6. to Table 9.

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	28	227.98				
2	27	201.23	1	26.75	3.59	0.0689

Table 6. General linear hypothesis test result for $H_{01}: \mathbf{c}_1\boldsymbol{\beta} = 0$

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	28	218.52				
2	27	201.23	1	17.29	2.32	0.1393

Table 7. General linear hypothesis test result for $H_{02}: \mathbf{c}_2\boldsymbol{\beta} = 0$

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	28	244.81				
2	27	201.23	1	43.59	5.85	0.0226

Table 8. General linear hypothesis test result for $H_{03}: \mathbf{c}_3\boldsymbol{\beta} = 0$

For H_{04}, use the first and last row of $\mathbf{C}$. Let $\mathbf{A} = (\mathbf{c}_1, \mathbf{c}_3)^T$. Then we test $H_{04}: \mathbf{A}\boldsymbol{\beta} = 0$.

```
> linearHypothesis(lm1,C[c(1,3),],c(0,0))
```

The test result is shown in Table 10.

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	29	407.52				
2	27	201.23	2	206.30	13.84	0.0001

Table 9. General linear hypothesis test result for $H_{04}: \mathbf{A}\boldsymbol{\beta} = 0$

(e) Find confidence intervals for $\beta_1, \beta_2, \beta_3$ and β_4 using both (8.47) and (8.67).

Sol. Using the function `confint()` in R package,

```
> conf.b.tbl<-matrix(0,4,4)

> for (i in 1:4) {
```

```
+         tmp1<-confint(lm1,parm=names(Data)[i+1], level=0.95)
+         tmp2<-confint(lm1,parm=names(Data)[i+1], level=1-0.05/4)
+         conf.b.tbl[i,1:2]<-tmp1
+         conf.b.tbl[i,3:4]<-tmp2
+
+ }
> conf.b.tbl<-as.data.frame(conf.b.tbl)
> row.names(conf.b.tbl)<-c("x1","x2","x3","x4")
> tmp1<-as.data.frame(tmp1)
> tmp2<-as.data.frame(tmp2)
> names(conf.b.tbl)[1:2]<-names(tmp1)
> names(conf.b.tbl)[3:4]<-names(tmp2)
```

$100(1-\alpha)\%$ confidence intervals for β_j are shown as

	2.5 %	97.5 %	0.625 %	99.38 %
β_1	-0.215	0.157	-0.271	0.214
β_2	0.077	0.355	0.035	0.397
β_3	-10.170	1.530	-11.950	3.310
β_4	3.286	14.664	1.555	16.395

Table 10. 95% and Bonferroni $(100-0.05/4)\%$ confidence intervals for each β_j with df=27.

9.10 For the gas vapor data in Table 7.3, compute the diagnostic measures $\hat{y}_i, \hat{e}_i, h_{ii}, r_i, t_i$ and D_i. Display these in a table similar to Table 9.1. Are there outliers or potentially influential observations? Calculate PRESS and compare to SSE.

Sol. Residuals and influence measures for the gas vapor data

```
> #predicted value
```

```
> y.hat<-predict(lm1,type="response")
> #residuals
> e<-lm1$residuals
> #calculate hat matrix
> H=X%*%solve((t(X)%*%X))%*%t(X)
> h.diag<-diag(H)
> #studentized residual
> std.res<-residuals(lm1)/(summary(lm1)$sigma*sqrt(1-h.diag))
> #studentized deleted residuals
> std.d.res<-studres(lm1)
> #Cook's distance
> cooksd<-cooks.distance(lm1)
> #PRESS
> press<-sum((e/(1-h.diag))^2)
> infl.tbl<-data.frame(y.hat,e,h.diag,std.res,std.d.res,cooksd)
> # Cook's D plot and Influence Plot
> par(mfrow=c(2,1))
> plot(lm1, which=4, cook.levels=cutoff)
> influencePlot(lm1,id.method="identify",ylim=c(-3,3),id.col="red",
+                 main="Influence Plot")
```

Obs.	$\hat{y}_i$	$\hat{e}_i$	h_{ii}	r_i	t_i	D_i
1	27.861	1.139	0.197	0.466	0.459	0.011
2	23.764	0.236	0.219	0.098	0.096	0.001
3	25.880	0.120	0.179	0.049	0.048	0.000
4	23.961	-1.961	0.289	-0.852	-0.847	0.059
5	28.419	-1.419	0.128	-0.557	-0.549	0.009
6	21.671	-0.671	0.121	-0.262	-0.258	0.002
7	31.778	1.222	0.053	0.460	0.453	0.002
8	34.218	-0.218	0.042	-0.082	-0.080	0.000
9	31.983	0.017	0.055	0.006	0.006	0.000
10	33.334	0.666	0.039	0.249	0.244	0.000
11	21.544	-1.544	0.124	-0.604	-0.597	0.010
12	32.154	3.846	0.040	1.438	1.468	0.017
13	33.729	0.271	0.072	0.103	0.101	0.000
14	23.982	-0.982	0.191	-0.400	-0.394	0.008
15	19.713	4.287	0.418	2.058	2.200	0.609
16	32.841	-0.841	0.060	-0.318	-0.312	0.001
17	40.761	-0.761	0.285	-0.330	-0.324	0.009
18	44.386	1.614	0.493	0.831	0.826	0.134
19	52.917	2.083	0.243	0.877	0.873	0.049
20	52.018	-0.018	0.224	-0.007	-0.007	0.000
21	32.377	-3.377	0.177	-1.364	-1.387	0.080
22	23.155	-1.155	0.169	-0.464	-0.457	0.009
23	36.586	-5.586	0.227	-2.328	-2.555	0.319
24	47.909	-2.909	0.185	-1.180	-1.189	0.063
25	32.609	4.391	0.087	1.683	1.746	0.054
26	31.894	5.106	0.109	1.981	2.103	0.096
27	30.225	2.775	0.124	1.086	1.090	0.033
28	31.593	-4.593	0.102	-1.775	-1.854	0.071
29	34.399	-0.399	0.068	-0.151	-0.149	0.000
30	19.324	-0.324	0.091	-0.124	-0.122	0.000
31	19.623	-3.623	0.102	-1.400	-1.427	0.044
32	19.393	2.607	0.086	0.999	0.999	0.019

Table 11. Influence measures for the vapor dataset, $PRESS = 310.443$, $SSE = 201.228$

According to the result of influence measures for vapor dataset shown in Table 11., the observation number 15, 18, and 23 are highly suspected to be outliers of influental observation compared to the other observations due to their abnormal values of $\hat{e}_i$, h_{ii}, r_i,t_i and D_i. I consider that graphical approach may reveil my suspection more clearly by plotting Cook's distance and influence plot, which represents h_{ii}, t_i and D_i simultaneously.

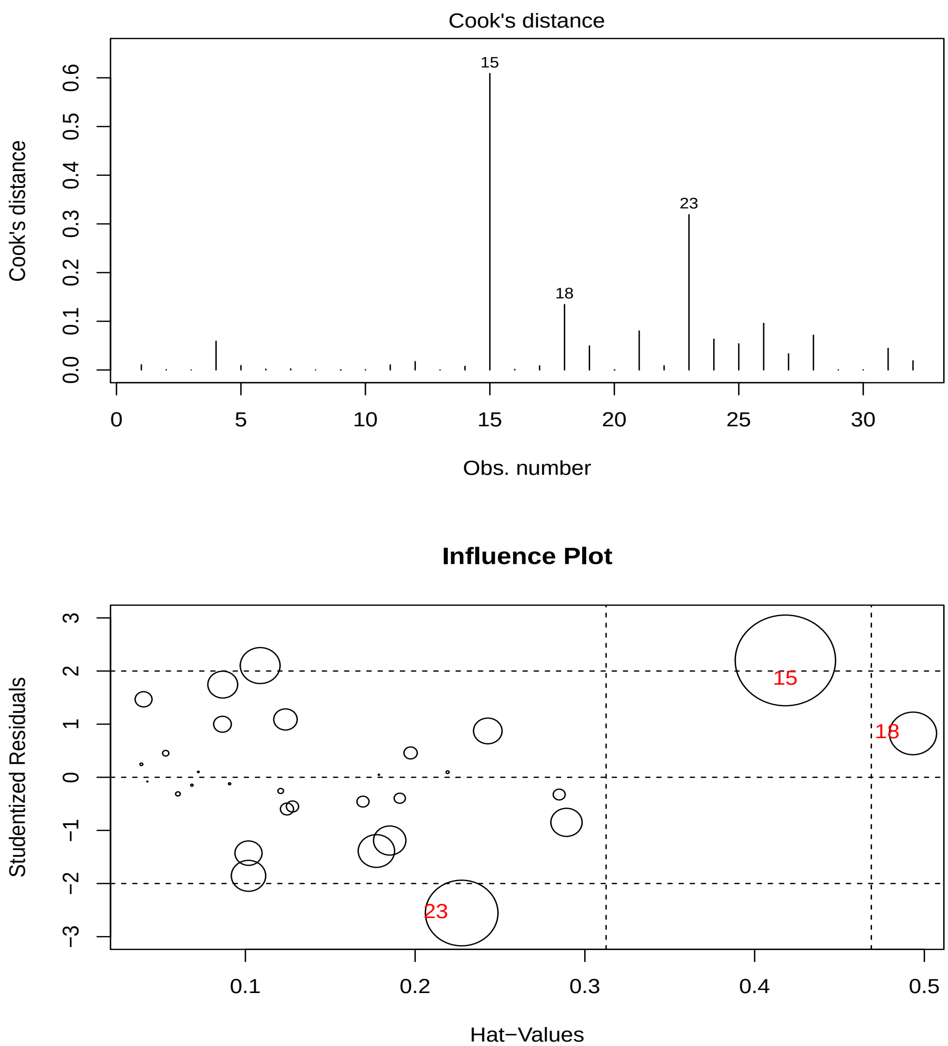

Figure 1. The result of Cook's D plot and Influence Plot: Circles in Influence Plot are proportional to Cook's distance

Linear Models in Statistics: HW5

201060072: Boncho Ku

12.19 Redo Example 11.4 with the reparameterization

$$\boldsymbol{\gamma} = \begin{pmatrix} \mu + \tau_1 \\ \tau_1 - \tau_2 \end{pmatrix}$$

Find $\mathbf{Z}$ and $\mathbf{U}$ by inspection and show that $\mathbf{ZU} = \mathbf{X}$. Then show that $\mathbf{Z}$ can be obtained as $\mathbf{Z} = \mathbf{XU}^T(\mathbf{UU}^T)^{-1}$.

Sol. Let $\mu_i = \mu + \tau_i$. Then $\mu + \tau_1 = \mu_1$ and $\tau_1 - \tau_2 = \mu_1 - \mu_2$. Therefore, design matrix for $\boldsymbol{\gamma}$, $\mathbf{Z}$ can be written as

$$\mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & -1 \\ 1 & -1 \end{pmatrix}$$

Since $\boldsymbol{\gamma} = \mathbf{U}\boldsymbol{\beta}$,

$$\boldsymbol{\gamma} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \end{pmatrix} = \mathbf{U}\boldsymbol{\beta}, \quad \mathbf{U} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

$$\therefore \mathbf{ZU} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \mathbf{X}$$

Consequently, $\mathbf{Z}$ can be obtained from the equation $\mathbf{Z} = \mathbf{X}\mathbf{U}^T(\mathbf{U}\mathbf{U}^T)^{-1}$. Therefore,

$$\mathbf{U}\mathbf{U}^T = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$\mathbf{Z} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 2/3 & -1 \\ -1 & 2/3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & -1 \\ 1 & -1 \end{pmatrix}$$

12.27 Consider the model $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon_{ijk}, \quad i = 1, 2, \quad , j = 1, 2, \quad k = 1, 2.$

(a) Write $\mathbf{X}^T\mathbf{X}$, $\mathbf{X}^T\mathbf{y}$, and the normal equation.

Sol. The given model is for the 2^3 factorial design and the design matrix can be written as

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}, \quad \mathbf{X}^T\mathbf{X} = \begin{pmatrix} 8 & 4 & 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 0 & 2 & 2 & 2 & 2 \\ 4 & 0 & 4 & 2 & 2 & 2 & 2 \\ 4 & 2 & 2 & 4 & 0 & 2 & 2 \\ 4 & 2 & 2 & 0 & 4 & 2 & 2 \\ 4 & 2 & 2 & 2 & 2 & 4 & 0 \\ 4 & 2 & 2 & 2 & 2 & 0 & 4 \end{pmatrix}$$

$$\mathbf{X}^T\mathbf{y} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \end{pmatrix} = \begin{pmatrix} y_{...} \\ y_{1..} \\ y_{2..} \\ y_{.1.} \\ y_{.2.} \\ y_{..1} \\ y_{..2} \end{pmatrix}$$

The normal equation can be expressed as

$$\begin{pmatrix} 8 & 4 & 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 0 & 2 & 2 & 2 & 2 \\ 4 & 0 & 4 & 2 & 2 & 2 & 2 \\ 4 & 2 & 2 & 4 & 0 & 2 & 2 \\ 4 & 2 & 2 & 0 & 4 & 2 & 2 \\ 4 & 2 & 2 & 2 & 2 & 4 & 0 \\ 4 & 2 & 2 & 2 & 2 & 0 & 4 \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix} = \begin{pmatrix} y_{...} \\ y_{1..} \\ y_{2..} \\ y_{.1.} \\ y_{.2.} \\ y_{..1} \\ y_{..2} \end{pmatrix}$$

(b) Find a set of linearly independent estimable functions.

Sol. Since rank($\mathbf{X}$) $=$ 4, there exist 4 linearly independent estimable functions found by reduced row echelon form of $\mathbf{X}$,

$$\mathbf{X} \sim \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Therefore, esitmable functions are given by

$$\mu + \alpha_1 + \beta_1 + \gamma_1,\ \alpha_1 - \alpha_2,\ \beta_1 - \beta_2,\ \gamma_1 - \gamma_2$$

(c) Define appropriate side conditions and find the resulting solution to the normal equation.

Sol. Side conditions for the given model are

$$\sum_{i=1}^{2} \alpha_i = 0, \ \sum_{j=1}^{2} \beta_j = 0, \ \sum_{k=1}^{2} \gamma_k = 0$$

Then a set of non-estimable function is defined by

$$\mathbf{T}\boldsymbol{\beta} \quad = \quad \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \boldsymbol{\beta} = 0$$

$$\mathbf{T}^T\mathbf{T} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \quad (\mathbf{X}^T\mathbf{X} + \mathbf{T}^T\mathbf{T}) = \begin{pmatrix} 8 & 4 & 4 & 4 & 4 & 4 & 4 \\ 4 & 5 & 1 & 2 & 2 & 2 & 2 \\ 4 & 1 & 5 & 2 & 2 & 2 & 2 \\ 4 & 2 & 2 & 5 & 1 & 2 & 2 \\ 4 & 2 & 2 & 1 & 5 & 2 & 2 \\ 4 & 2 & 2 & 2 & 2 & 5 & 1 \\ 4 & 2 & 2 & 2 & 2 & 1 & 5 \end{pmatrix}$$

$$(\mathbf{X}^T\mathbf{X}+\mathbf{T}^T\mathbf{T})^{-1} = \begin{pmatrix} 28/32 & -1/4 & -1/4 & -1/4 & -1/4 & -1/4 & -1/4 \\ -1/4 & 3/8 & 1/8 & 0 & 0 & 0 & 0 \\ -1/4 & 1/8 & 3/8 & 0 & 0 & 0 & 0 \\ -1/4 & 0 & 0 & 3/8 & 1/8 & 0 & 0 \\ -1/4 & 0 & 0 & 1/8 & 3/8 & 0 & 0 \\ -1/4 & 0 & 0 & 0 & 0 & 3/8 & 1/8 \\ -1/4 & 0 & 0 & 0 & 0 & 1/8 & 3/8 \end{pmatrix}$$

$$\begin{aligned}
\hat{\mu} &= \frac{28}{32}y_{...} - \frac{24}{32}y_{...} = \frac{1}{8}y_{...} = \bar{y}_{...}, \\
\hat{\alpha}_1 &= -\frac{1}{4}y_{...} + \frac{3}{8}y_{1..} + \frac{1}{8}y_{2..} = \frac{1}{4}y_{1..} - \frac{1}{8}y_{...} = \bar{y}_{1..} - \bar{y}_{...}, \\
\hat{\alpha}_2 &= -\frac{1}{4}y_{...} + \frac{1}{8}y_{1..} + \frac{3}{8}y_{2..} = \frac{1}{4}y_{2..} - \frac{1}{8}y_{...} = \bar{y}_{2..} - \bar{y}_{...}, \\
\hat{\beta}_1 &= \bar{y}_{.1.} - \bar{y}_{...}, \quad \hat{\beta}_2 = \bar{y}_{.2.} - \bar{y}_{...}, \\
\hat{\gamma}_1 &= \bar{y}_{..1} - \bar{y}_{...}, \quad \hat{\gamma}_2 = \bar{y}_{..2} - \bar{y}_{...}
\end{aligned}$$

(d) Show that $H_0 : \alpha_1 = \alpha_2$ is testable. Find $\hat{\boldsymbol{\beta}}\mathbf{X}^T\mathbf{y} = \mathrm{SS}(\mu, \alpha, \beta, \gamma)$ and $\hat{\boldsymbol{\beta}}_2\mathbf{X}_2^T\mathbf{y} = \mathrm{SS}(\mu, \beta, \gamma)$.

Sol. Since $\alpha_1 - \alpha_2$ is a linearly independent estimable function $\boldsymbol{\lambda}^T\boldsymbol{\beta}$, $H_0 : \alpha_1 - \alpha_2 =$

0 is testable, where $\boldsymbol{\lambda}^T = (0, 1, -1, 0, 0, 0, 0)$. Then,

$$
\begin{aligned}
\mathrm{SS}(\mu,\alpha,\beta,\gamma) &= \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} = \begin{pmatrix} \hat{\mu} & \hat{\alpha}_1 & \ldots & \hat{\gamma}_2 \end{pmatrix} \begin{pmatrix} y_{...} \\ y_{1..} \\ y_{2..} \\ y_{.1.} \\ y_{.2.} \\ y_{..1} \\ y_{..2} \end{pmatrix} \\
&= \hat{\mu} y_{...} + \hat{\alpha}_1 y_{1..} + \hat{\alpha}_2 y_{2..} + \hat{\beta}_1 y_{.1.} + \hat{\beta}_2 y_{.2.} + \hat{\gamma}_1 y_{..1} + \hat{\gamma}_2 y_{..2} \\
&= \frac{1}{8} y_{...}^2 + [\sum_{i=1}^{2} \frac{1}{4} y_{i..}^2 - \frac{1}{8} y_{...}^2] + [\sum_{j=1}^{2} \frac{1}{4} y_{.j.}^2 - \frac{1}{8} y_{...}^2] + [\sum_{k=1}^{2} \frac{1}{4} y_{..k}^2 - \frac{1}{8} y_{...}^2] \\
&= \mathrm{SS}(\mu) + \mathrm{SS}(\alpha) + \mathrm{SS}(\beta) + \mathrm{SS}(\gamma) \\
\mathrm{SS}(\mu,\beta,\gamma) &= \mathrm{SS}(\mu) + \mathrm{SS}(\beta) + \mathrm{SS}(\gamma)
\end{aligned}
$$

(e) Construct an analysis of variance table for the test of $H_0 : \alpha_1 = \alpha_2$.

Sol. ANOVA table for testing $H_0 : \alpha_1 = \alpha_2$

Source	Df	Sum Sq	Mean Sq	F value
$\mathrm{SS}(\alpha \mid \mu,\beta,\gamma)$	1	$\mathrm{SS}(\mu,\alpha,\beta,\gamma) - \mathrm{SS}(\mu,\beta,\gamma) = \mathrm{SS}(\alpha)$	$\mathrm{SS}(\alpha)$	$\mathrm{SS}(\alpha)/\mathrm{MSE}$
Error	8-4=4	$\mathrm{SSE} = \sum_{ijk} y_{ijk}^2 - \mathrm{SS}(\mu,\alpha,\beta,\gamma)$	SSE/4	

13.28 Blood sugar levels were measured on 10 animals from each of five breeds (Daniel 1974, p. 197). The results are in Table 14.

(a) Test the hypothesis of equality of means for the five bleeds.

Sol. Performing one-way ANOVA analysis

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Breed	4	3213.48	803.37	6.59	0.0003
Residuals	45	5485.40	121.90		

Table 12. ANOVA table for testing equal means for the five breeds

(b) Make the followng four comparisons by means of orthogonal contrasts:

$$\text{A, B, C vs. D, E;}\quad \text{A, B vs. C;}\quad \text{A, vs. B;}\quad \text{D, vs. E;}$$

Sol. The hypothesis for testing the given comparisons can be expressed as H_0 : $\mathbf{C}\boldsymbol{\beta} = 0$. The orthogonal contrasts can be expressed as

$$\frac{\mu_A + \mu_B + \mu_C}{3} = \frac{\mu_D + \mu_E}{2}; \quad \frac{\mu_A + \mu_B}{2} = \mu_C; \quad \mu_A = \mu_B; \quad \mu_D = \mu_E$$

$$\mathbf{C} = \begin{pmatrix} 2 & 2 & 2 & -3 & -3 \\ 1 & 1 & -2 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Breed	4	3213.48	803.37	6.59	0.0003
Breed: A,B,C vs. D,E	1	82.16	82.16	0.67	0.4160
Breed: A,B vs. C	1	1025.07	1025.07	8.41	0.0058
Breed: A vs. B	1	5.00	5.00	0.04	0.8404
Breed: D vs. E	1	2101.25	2101.25	17.24	0.0001
Residuals	45	5485.40	121.90		

Table 13. ANOVA table for the orthogonal contrasts

Breed				
A	B	C	D	E
124	111	117	104	142
116	101	142	128	139
101	130	121	130	133
118	108	123	103	120
118	127	121	121	127
120	129	148	119	149
110	122	141	106	150
127	103	122	107	149
106	122	139	107	120
130	127	125	115	116

Table 14. Blood Sugar Levels (mg/100g) for Ten Animals from Each of Five Breeds

13.33 Weight gains in pigs subjected to four different treatments are given in Table 17 (Crapmton and Hopkins 1934).

(a) Test the hypothesis of equal mean treatment effects.

Sol. One-way ANOVA table without contrasts

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Trt	3	3394.47	1131.49	6.57	0.0012
Residuals	36	6197.90	172.16		

Table 15. ANOVA table for testing equal means for the four treatments

(b) Using contrasts, compare treatments 1, 2, 3 vs. 4; 1, 2, vs. 3; and 1 vs. 2.

Sol. Contrast matrix for the given comparison

$$\mathbf{C} = \begin{pmatrix} 1 & 1 & 1 & -3 \\ 1 & 1 & -2 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}$$

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Trt	3	3394.47	1131.49	6.57	0.0012
Trt: 1,2,3 vs. 4	1	1944.07	1944.07	11.29	0.0019
Trt: 1,2 vs. 3	1	72.60	72.60	0.42	0.5202
Trt: 1 vs. 2	1	1377.80	1377.80	8.00	0.0076
Residuals	36	6197.90	172.16		

Table 16. ANOVA table for the orthogonal contrasts

Treatment			
1	2	3	4
165	168	164	185
159	180	156	195
159	180	189	186
167	166	138	201
170	170	153	165
146	161	190	175
130	171	160	187
151	169	172	177
164	179	142	166
158	191	155	165

Table 17. Weight Gain of Pigs Subjected to Four Treatments