Question #1.

Solution 1. *For part [A]:*

We want to get the singular value decomposition of $\mathbf{A}$*:*

We have

$$\mathbf{AA}' = \begin{bmatrix} 10 & -5 \\ 2 & -11 \\ 6 & -8 \end{bmatrix} \begin{bmatrix} 10 & 2 & 6 \\ -5 & -11 & -8 \end{bmatrix} = \begin{bmatrix} 125 & 75 & 100 \\ 75 & 125 & 100 \\ 100 & 100 & 100 \end{bmatrix}$$

So we have:

$$\begin{aligned}
det(\gamma\mathbf{I}_3 - \mathbf{AA}') &= det\Big(\begin{bmatrix} \gamma-125 & -75 & -100 \\ -75 & \gamma-125 & -100 \\ -100 & -100 & \gamma-100 \end{bmatrix}\Big) \\
&= (\gamma-125)\cdot det\Big(\begin{bmatrix} \gamma-125 & -100 \\ -100 & \gamma-100 \end{bmatrix}\Big) + 75det\Big(\begin{bmatrix} -75 & -100 \\ -100 & \gamma-100 \end{bmatrix}\Big) \\
&\qquad - 100det\Big(\begin{bmatrix} -75 & -100 \\ \gamma-125 & -100 \end{bmatrix}\Big) \\
&= (\gamma-125)\Big[(\gamma-125)(\gamma-100) - 10,000\Big] + 75\Big[-75(\gamma-100) - 10,000\Big] \\
&\qquad - 100\Big(7,500 + 100(\gamma-125)\Big) \\
&= (\gamma-125)^2(\gamma-100) - 10,000(\gamma-125) - 5,625(\gamma-100) - 750,000 \\
&\qquad - 750,000 - 10,000(\gamma-125) \\
&= (\gamma-125)^2(\gamma-100) - 20,000(\gamma-125) - 5,625(\gamma-100) - 1,500,000 \\
&= (\gamma^2 - 250\gamma + 15,625)(\gamma-100) - 20,000(\gamma-125) - 5,625(\gamma-100) \\
&\qquad - 1,500,000 \\
&= \gamma^3 - 100\gamma^2 - 250\gamma^2 + 25,000\gamma + 15,625\gamma - 1,562,500 - 20,000\gamma \\
&\qquad + 2,500,000 - 5,625\gamma + 256,500 - 1,500,000 \\
&= \gamma^3 - 350\gamma^2 + 15,000\gamma \\
&= \gamma(\gamma^2 - 350\gamma + 15,000) \\
&= \gamma(\gamma-300)(\gamma-50)
\end{aligned}$$

So the 3 *eigen values for* $\mathbf{AA}'$ *are* $\gamma_1 = 300, \gamma_2 = 50$ *and* $\gamma_3 = 0$*.*

We intentionally made them in descending order, because later when we verify these results in SAS it is also the algorithm SAS is using. It is easier for us to get consistent results this way.

Now we look for the **unit** *eigenvectors for each eigenvalue. We solve them by doing the Gauss elimination on the rows:*

When $\gamma_1 = 300$*, we have:*

$$\gamma_1\mathbf{I} - \mathbf{AA}' = \begin{bmatrix} 175 & -75 & -100 \\ -75 & 175 & -100 \\ -100 & -100 & 200 \end{bmatrix} \Longrightarrow \begin{bmatrix} -1 & -1 & 2 \\ 3 & -7 & 4 \\ 7 & -3 & -4 \end{bmatrix} \Longrightarrow \begin{bmatrix} -1 & -1 & 2 \\ 0 & -10 & 10 \\ 0 & -10 & 10 \end{bmatrix}$$
$$\Longrightarrow \begin{bmatrix} -1 & -1 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Suppose $\mathbf{x} = (x_1, x_2, x_3)'$ *is one of the eigenvectors, then the above row operation implies that* $x_2 = x_3$ *and* $x_1 = -x_2 + 2x_3 = x_3$*.*

So a candidate of the eigenvector would be $(1, 1, 1)'$*, and hence we got the unit eigenvector:*

$$\mathbf{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

A quick comment here: the choice of eigenvectors are not unique, but we generally make the first element of the eigenvector positive, in order to get consistent results when we verify the correctness of the SVD later.

Now when $\gamma_2 = 50$*, we have:*

We have:

$$\gamma_2\mathbf{I} - \mathbf{AA}' = \begin{bmatrix} -75 & -75 & -100 \\ -75 & -75 & -100 \\ -100 & -100 & -50 \end{bmatrix} \Longrightarrow \begin{bmatrix} -2 & -2 & -1 \\ 3 & 3 & 4 \\ 3 & 3 & 4 \end{bmatrix} \Longrightarrow \begin{bmatrix} -2 & -2 & -1 \\ 0 & 0 & \frac{5}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

Suppose $\mathbf{x} = (x_1, x_2, x_3)'$ *is one of the eigenvectors, then the above row operation implies that* $x_3 = 0, x_2 = -x_1$ *and hence* $(1, -1, 0)'$ *is one of the eigenvectors. So the unit eigenvector is:*

$$\mathbf{u}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

Finally when $\gamma_3 = 0$*, we have:*

$$\gamma_3\mathbf{I} - \mathbf{AA}' = \begin{bmatrix} -125 & -75 & -100 \\ -75 & -125 & -100 \\ -100 & -100 & -100 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ -3 & -5 & -4 \\ -5 & -3 & -4 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & -1 \\ 0 & 2 & 1 \end{bmatrix}$$
$$\Longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Suppose $\mathbf{x} = (x_1, x_2, x_3)'$ *is one of the eigenvectors, then the above row operation implies that* $x_2 = -\frac{1}{2}x_3$*, and* $x_1 = -x_2 - x_3 = \frac{1}{2}x_3 - x_3 = -\frac{1}{2}x_3$*. So a candidate of the eigenvector would be*

$(1, 1, -2)'$ *and hence the unit eigenvector is:*

$$\mathbf{u}_3 = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{bmatrix}$$

Thus we have found

$$\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{bmatrix}$$

On the other hand, we have:

$$\mathbf{A}'\mathbf{A} = \begin{bmatrix} 10 & 2 & 6 \\ -5 & -11 & -8 \end{bmatrix} \begin{bmatrix} 10 & -5 \\ 2 & -11 \\ 6 & -8 \end{bmatrix} = \begin{bmatrix} 140 & -120 \\ -120 & 210 \end{bmatrix}$$

So

$$\begin{aligned} det(\gamma \mathbf{I}_2 - \mathbf{A}'\mathbf{A}) &= det\Big(\begin{bmatrix} \gamma - 140 & 120 \\ 120 & \gamma - 210 \end{bmatrix} \Big) = (\gamma - 140)(\gamma - 210) - 14,400 \\ &= \gamma^2 - 350\gamma + 29,400 - 14,400 \\ &= \gamma^2 - 350\gamma + 15,000 \\ &= (\gamma - 300)(\gamma - 50) \end{aligned}$$

So we have the two eigenvalues for $\mathbf{A}'\mathbf{A}$ *as* $\gamma_1 = 300, \gamma_2 = 50$.

When $\gamma_1 = 300$*, we have:*

$$\gamma_1 \mathbf{I}_2 - \mathbf{A}'\mathbf{A} = \begin{bmatrix} 160 & 120 \\ 120 & 90 \end{bmatrix} \Longrightarrow \begin{bmatrix} 4 & 3 \\ 4 & 3 \end{bmatrix} \Longrightarrow \begin{bmatrix} 4 & 3 \\ 0 & 0 \end{bmatrix}$$

The row operation above implies that if $\mathbf{x} = (x_1, x_2)'$ *is one of the eigenvectors, then* $4x_1 + 3x_2 = 0$*, and hence one of the candidate would be* $(3, -4)'$*, and hence a unit eigenvector is*

$$\mathbf{v}_1 = \begin{bmatrix} \frac{3}{5} \\ -\frac{4}{5} \end{bmatrix}$$

When $\gamma_2 = 50$*, we have:*

$$\gamma_2 \mathbf{I}_2 - \mathbf{A}'\mathbf{A} = \begin{bmatrix} -90 & 120 \\ 120 & -160 \end{bmatrix} \Longrightarrow \begin{bmatrix} -3 & 4 \\ 3 & -4 \end{bmatrix} \Longrightarrow \begin{bmatrix} -3 & 4 \\ 0 & 0 \end{bmatrix}$$

The above row operation implies that, if $\mathbf{x} = (x_1, x_2)'$ *is an eigenvector, then* $-3x_1 + 4x_2 = 0$*, this implies that a candidate would be* $(4, 3)'$*, and hence the unit eigenvector is:*

$$\mathbf{v}_2 = \begin{bmatrix} \frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$$

Hence we have found:

$$\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2] = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix}$$

We notice that $\mathbf{U}$ *and* $\mathbf{V}$ *have the same two largest eigenvalues* $\gamma_1 = 300$*, and* $\gamma_2 = 50$*, thus we have* $\lambda_1 = \sqrt{\gamma_1} = 10\sqrt{3}$ *and* $\lambda_2 = \sqrt{\gamma_2} = 5\sqrt{2}$*, hence we have also found:*

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 10\sqrt{3} & 0 \\ 0 & 5\sqrt{2} \\ 0 & 0 \end{bmatrix}$$

and thus the SVD(singular value decomplsition) of $\mathbf{A}$ *is:*

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}' = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 10\sqrt{3} & 0 \\ 0 & 5\sqrt{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}$$

To verify with software, we did this with both SAS and R.

In SAS, we run the svd function under proc iml as the following:

```
proc iml;
    a = {10 -5,
          2 -11,
          6 -8};
    call svd(u, q, v, a);
    print u, q, v;
quit;
```

The output is:

```
          u

0.5773503 0.7071068
0.5773503 -0.707107
0.5773503         0
          q

   17.320508
   7.0710678
          v

      0.6       0.8
     -0.8       0.6
```

This matches with our manual computation above since $\frac{1}{\sqrt{3}} \simeq 0.5773503$ *and* $\frac{1}{\sqrt{2}} \simeq 0.7071068$*. Also* $10\sqrt{3} \simeq 17.32051$ *and* $5\sqrt{2} \simeq 7.071068$*.*

In R, we simply check that the manually computed $\mathbf{U}, \boldsymbol{\Lambda}$ *and* $\mathbf{V}'$ *above do multiply to get back to* $\mathbf{A}$*:*

```r
#Q1, check the correctness of SVD for A

A= matrix(
  c(10, -5, 2, -11, 6, -8),
  nrow=3,
  ncol=2,
  byrow = TRUE)
U = matrix(
    c(1/sqrt(3), 1/sqrt(2), 1/sqrt(6),
      1/sqrt(3), -1/sqrt(2), 1/sqrt(6),
      1/sqrt(3), 0, -2/sqrt(6)
      ),
    nrow = 3,
    ncol = 3,
    byrow = TRUE
  )
V = matrix(
   c(0.6,0.8,-0.8,0.6),
   nrow=2,
   ncol=2,
   byrow=TRUE
)
Lam = matrix(
  c(10*sqrt(3),0,
    0,5*sqrt(2),
    0,0),
  nrow=3,
  ncol=2,
  byrow=TRUE
)
U%*%Lam%*%t(V)

```

the output is:

```
> U%*%Lam%*%t(V)
     [,1] [,2]
[1,]   10   -5
[2,]    2  -11
[3,]    6   -8
> |
```

Thus our SVD is correct.

For part [B]:

For part (i):

To check that $\mathbf{G}$ *is a generalized inverse of* $\mathbf{A}$*, we need to show that* $\mathbf{AGA} = \mathbf{A}$*.*

We have:

$$
\begin{aligned}
\mathbf{AGA} &= \left(\mathbf{M}\begin{bmatrix} \mathbf{W}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{Q}^T\right)\left(\mathbf{Q}\begin{bmatrix} \mathbf{W}_r^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{M}^T\right)\left(\mathbf{M}\begin{bmatrix} \mathbf{W}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{Q}^T\right) \\
&= \mathbf{M}\begin{bmatrix} \mathbf{W}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\Big(\underbrace{\mathbf{Q}^T\mathbf{Q}}_{=\mathbf{I}_n}\Big)\begin{bmatrix} \mathbf{W}_r^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\Big(\underbrace{\mathbf{M}^T\mathbf{M}}_{=\mathbf{I}_n}\Big)\begin{bmatrix} \mathbf{W}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{Q}^T \\
&= \mathbf{M}\left(\begin{bmatrix} \mathbf{W}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{W}_r^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right)\begin{bmatrix} \mathbf{W}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{Q}^T \\
&= \mathbf{M}\begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{W}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{Q}^T \\
&= \mathbf{M}\begin{bmatrix} \mathbf{W}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{Q}^T = \mathbf{A}
\end{aligned}
$$

Thus $\mathbf{G}$ *is a generalized inverse of* $\mathbf{A}$.

For part (ii):

We need to show $\mathbf{AG} = \left(\mathbf{AG}\right)'$.

We have:

$$\begin{aligned}
\mathbf{AG} &= \mathbf{M}\begin{bmatrix} \mathbf{W}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{Q}^T\mathbf{Q}\begin{bmatrix} \mathbf{W}_r^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{M}^T \\
&= \mathbf{M}\begin{bmatrix} \mathbf{W}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{W}_r^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{M}^T \\
&= \mathbf{M}\begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{M}^T
\end{aligned}$$

So

$$\begin{aligned}
\left(\mathbf{AG}\right)' &= \left(\mathbf{M}^T\right)^T\left(\begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right)^T\mathbf{M}^T \\
&= \mathbf{M}\begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{M}^T \\
&= \mathbf{AG}
\end{aligned}$$

Thus completed the proof.

Question #2.

Solution 2. *equation* (7.57) *says:*

$$E[R^2] = \frac{k}{n-1}$$

under the null hypothesis that $\beta_1 = \beta_2 = \cdots = \beta_k = 0$.

We prove it by finding the distribution of R^2.

According to Theorem 8.1*d(page 187) on the textbook by Rencher and Schaalje, under the null hypothesis of* $\beta_1 = \beta_2 = \cdots = \beta_k = 0$, *we have:*

$$F = \frac{SSR/k}{SSE/(n-k-1)} \sim F(k, n-k-1)$$

Since $R^2 = \frac{SSR}{SST}$, we try to write F as a function of R^2 by noticing that:

$$\begin{aligned}
F &= \frac{SSR/k}{SSE/(n-k-1)} \sim F(k, n-k-1) \\
&= \frac{SSR/k}{(SST-SSR)/(n-k-1)} \\
&= \frac{1}{\frac{(SST-SSR)/(n-k-1)}{SSR/k}} \\
&= \frac{1}{\frac{kSST}{(n-k-1)SSR} - \frac{k}{n-k-1}} \\
&= \frac{1}{\left(\frac{k}{n-k-1}\right)\left[\frac{1}{R^2} - 1\right]}
\end{aligned}$$

So we can alternatively write R^2 as a function of F:

$$\begin{aligned}
&\implies \frac{1}{R^2} - 1 = \frac{1}{\frac{k}{n-k-1}F} \\
&\implies \frac{1}{R^2} = \frac{1 + \frac{k}{n-k-1}F}{\frac{k}{n-k-1}F} \\
&\implies R^2 = \frac{\frac{k}{n-k-1}F}{1 + \frac{k}{n-k-1}F} = \frac{kF}{(n-k-1) + kF}
\end{aligned}$$

but this implies that

$$R^2 \sim Beta(\frac{k}{2}, \frac{n-k-1}{2})$$

which is a beta distribution with parameter $\alpha = \frac{k}{2}, \beta = \frac{n-k-1}{2}$. an hence

$$\begin{aligned}
E[R^2] &= \frac{\alpha}{\alpha+\beta} = \frac{\frac{k}{2}}{\frac{k}{2} + \frac{n-k-1}{2}} \\
&= \frac{k}{n-1} \;\; equation(7.57) \\
Var[R^2] &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \\
&= \frac{\frac{k}{2}\frac{n-k-1}{2}}{(\frac{k}{2} + \frac{n-k-1}{2})^2(\frac{k}{2} + \frac{n-k-1}{2} + 1)} \\
&= \frac{\frac{1}{4}k(n-k-1)}{\frac{1}{4}(n-1)^2\frac{1}{2}(n-1+2)} \\
&= \frac{2k(n-k-1)}{(n-1)^2(n+1)}
\end{aligned}$$

for R^2_{adj}, we have the following definition:

$$R^2_{adj} = \frac{(n-1)R^2 - k}{n-k-1}$$

So

$$\begin{aligned}
E[R^2_{adj}] &= \frac{(n-1)E[R^2] - k}{n-k-1} = \frac{(n-1)\cdot\frac{k}{n-1} - k}{n-k-1} = 0 \\
Var[R^2_{adj}] &= \frac{(n-1)^2}{(n-k-1)^2} Var[R^2] = \frac{(n-1)^2}{(n-k-1)^2} \cdot \frac{2k(n-k-1)}{(n-1)^2(n+1)} \\
&= \frac{2k}{(n-k-1)(n+1)}
\end{aligned}$$

Question #3.

Solution 3. *Translate the question equivalently, we are trying to prove that, for a matrix denoted by $(\mathbf{X'X})^{(-)}$,*

$$\begin{aligned}
&(\mathbf{X'X})(\mathbf{X'X})^{(-)}(\mathbf{X'X}) = \mathbf{X'X} \\
\iff\ &(\mathbf{X'X})(\mathbf{X'X})^{(-)}\mathbf{X'y} = \mathbf{X'}y
\end{aligned}$$

First we assume that $(\mathbf{X'X})^{(-)}$ is the generalized inverse of $(\mathbf{XX'})$, which is to say:

$$(\mathbf{X'X})(\mathbf{X'X})^{(-)}(\mathbf{X'X}) = \mathbf{X'X}$$

Then we have:

$$(\mathbf{X'X})(\mathbf{X'X})^{(-)}(\mathbf{X'X})\hat{\beta} = (\mathbf{X'X})\hat{\beta} = \mathbf{X'y}$$

but for the left hand side above, we can also replace $(\mathbf{X'X})\hat{\beta}$ by $\mathbf{X'y}$, thus we have:

$$(\mathbf{X'X})(\mathbf{X'X})^{(-)}\mathbf{X'y} = \mathbf{X'y}$$

which means $(\mathbf{X'X})^{(-)}\mathbf{X'y}$ is a solution for $(\mathbf{X'X})\hat{\beta} = \mathbf{X'y}$.

On the other direction, we assume that

$$(\mathbf{X'X})(\mathbf{X'X})^{(-)}\mathbf{X'y} = \mathbf{X'y}$$

We can substitute $\mathbf{X'y}$ in terms of the left hand side of the equation into the left hand side itself, thus we have:

$$(\mathbf{X'X})(\mathbf{X'X})^{(-)}(\mathbf{X'X})(\mathbf{X'X})^{(-)}\mathbf{X'y} = \mathbf{X'y}$$

Compare the above two equations we have:

$$\Big[(\mathbf{X'X})(\mathbf{X'X})^{(-)}(\mathbf{X'X}) - (\mathbf{X'X})\Big](\mathbf{X'X})^{(-)}\mathbf{X'y} = 0$$

Since the equation holds true for any possible value of $\mathbf{y}$, hence we must have:

$$(\mathbf{X'X})(\mathbf{X'X})^{(-)}(\mathbf{X'X}) = \mathbf{X'X}$$

which is to say $(\mathbf{X'X})^{(-)}$ is a generalized inverse of $\mathbf{X'X}$.

Thus completed the proof.

Question #4.

Solution 4. *Before we head into any sub question here, we need to be aware that $a\mathbf{I}_n + b\mathbf{J}_n$ as a variance matrix of a normal distribution, it should be positive definite.*

It is well known from linear algebra results that $\mathbf{J}_n$ is a semi-definite matrix with two eigen values, one is $\lambda_1 = n$ with multiplicity 1 and the other is $\lambda_2 = 0$ with multiplicity 0, so there exists orthogonal matrix $\mathbf{D}$ such that

$$\mathbf{J}_n = \mathbf{D}\begin{bmatrix} n & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}\mathbf{D}^T$$

Thus we have:

$$\begin{aligned} a\mathbf{I}_n + b\mathbf{J}_n &= a\mathbf{D}\mathbf{D}^T + b\mathbf{D}\begin{bmatrix} n & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}\mathbf{D}^T \\ &= \mathbf{D}\begin{bmatrix} a+bn & & & \\ & a & & \\ & & \ddots & \\ & & & a \end{bmatrix}\mathbf{D}^T \end{aligned}$$

So a sufficient and necessary condition for $a\mathbf{I}_n + b\mathbf{J}_n$ to be positive definite is that:

$$a > 0 \text{ and } a + bn > 0$$

when we are doing the following problems, our range of a and b should comply with this condition.

For part (i):

we have:

$$\bar{\mathbf{y}}^2 = \left(\mathbf{y}'\frac{\mathbf{j}}{n}\right)\cdot\left(\frac{\mathbf{j}'}{n}\mathbf{y}\right) = \mathbf{y}'\frac{\mathbf{j}\mathbf{j}'}{n^2}\mathbf{y} = \mathbf{y}'\frac{\mathbf{J}_n}{n^2}\mathbf{y}$$

which is a quadratic form with $\mathbf{A} = \frac{\mathbf{J}_n}{n^2}$. We also know $\mathbf{y}$ is multivariate normal with

$$\Sigma = a\mathbf{I}_n + b\mathbf{J}_n$$

Thus we have:

$$\begin{aligned} \mathbf{A\Sigma} &= \frac{\mathbf{J}_n}{n^2}(a\mathbf{I}_n + b\mathbf{J}_n) \\ &= \frac{a}{n^2}\mathbf{J}_n + \frac{b}{n^2}\mathbf{J}_n^2 \\ &= \frac{a}{n^2}\mathbf{J}_n + \frac{b}{n}\mathbf{J}_n \ \textit{(since } \frac{\mathbf{J}_n}{n} \textit{ is idempotent)} \\ &= \frac{\mathbf{J}_n}{n}(\frac{a}{n} + b) \end{aligned}$$

In order for $\bar{\mathbf{y}}^2 = \mathbf{y}'\frac{\mathbf{J}_n}{n^2}\mathbf{y} = \mathbf{y}'\mathbf{A}\mathbf{y}$ to have χ^2 distribution, we need $\mathbf{A\Sigma}$ to be idempotent.

We have:

$$
\begin{aligned}
(\mathbf{A\Sigma})^2 &= \Big(\frac{\mathbf{J}_n}{n}(\frac{a}{n}+b)\Big)^2 = \Big(\frac{\mathbf{J}_n}{n}\Big)^2\Big(\frac{a}{n}+b\Big)^2 \\
&= \frac{\mathbf{J}_n}{n}\Big(\frac{a}{n}+b\Big)^2
\end{aligned}
$$

So in order to have

$$\mathbf{A\Sigma} = (\mathbf{A\Sigma})^2$$

we need

$$\frac{a}{n}+b = \Big(\frac{a}{n}+b\Big)^2$$

If $\frac{a}{n}+b = 0$, then $a+bn = 0$, which violate the necessary condition we discussed above that $a+bn > 0$ in order for $a\mathbf{I}_n + b\mathbf{J}_n$ to be positive definite.

So the only choice we have is

$$\frac{a}{n}+b = 1 \text{ and } a > 0(\text{ discussed above for why})$$

also, when $\frac{a}{n}+b = 1$, we have $a+bn = n > 0$, which complies with the condition we derived above.

So the final answer for this question is, we need:

$$\frac{a}{n}+b = 1 \text{ and } a > 0$$

in order for $\bar{\mathbf{y}}^2$ to be chi square distributed. This also implies that $b < 1$ otherwise we could not guarantee $a > 0$, but otherwise, the choice of b is arbitrary.

For part [ii]:

We have that:

$$
\begin{aligned}
\hat{\sigma}^2 &= \frac{\sum y_i^2 - n\bar{\mathbf{y}}^2}{n} = \frac{1}{n}\mathbf{y}'\mathbf{y} - \bar{\mathbf{y}}^2 \\
&= \mathbf{y}'\frac{\mathbf{I}_n}{n}\mathbf{y} - \mathbf{y}'\frac{\mathbf{J}_n}{n^2}\mathbf{y} \\
&= \mathbf{y}'\Big(\frac{\mathbf{I}_n}{n} - \frac{\mathbf{J}_n}{n^2}\Big)\mathbf{y}
\end{aligned}
$$

So $\hat{\sigma}^2$ is a quadratic form $\mathbf{y}'\mathbf{A}\mathbf{y}$ with $\mathbf{A} = \frac{\mathbf{I}_n}{n} - \frac{\mathbf{J}_n}{n^2}$. Meanwhile the variance matrix for $\mathbf{y}$ is $\mathbf{\Sigma} = a\mathbf{I}_n + b\mathbf{J}_n$, so we have:

$$
\begin{aligned}
\mathbf{A\Sigma} &= \Big(\frac{\mathbf{I}_n}{n} - \frac{\mathbf{J}_n}{n^2}\Big)\Big(a\mathbf{I}_n + b\mathbf{J}_n\Big) \\
&= \frac{a}{n}\mathbf{I}_n + \frac{b}{n}\mathbf{J}_n - \frac{a}{n^2}\mathbf{J}_n - \frac{b}{n}\mathbf{J}_n \\
&= \frac{a}{n}\mathbf{I}_n - \frac{a}{n^2}\mathbf{J}_n \\
&= \frac{a}{n}\Big(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\Big)
\end{aligned}
$$

Thus

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

So in order to have $\mathbf{A\Sigma} = (\mathbf{A\Sigma})^2$*, we need*

$$\frac{a}{n}(1 - \frac{a}{n}) = 0$$

So we need

$$
\begin{aligned}
&\frac{a}{n} = 0 \Longrightarrow a = 0(\textit{ which we can't do since } a > 0) \\
&\textit{or } 1 - \frac{a}{n} = 0 \Longrightarrow a = n
\end{aligned}
$$

Since we also need $a + bn > 0$*, while here we need* $a = n$*, this implies that*

$$n + bn > 0 \Longrightarrow n(b+1) > 0 \Longrightarrow b > -1$$

So the final answer for this one is that we need $a = n$ *and* $b > -1$ *for* $\hat{\sigma}^2$ *to be chi square distributed.*

For part (iii):

We have:

$$cn\bar{\mathbf{y}}^2 = cny'\frac{\mathbf{jj}'}{n^2}\mathbf{y} = \mathbf{y}' \cdot \frac{c\mathbf{J}_n}{n} \cdot \mathbf{y}$$

So this is a quadrati form with $\mathbf{A} = \frac{c\mathbf{J}_n}{n}$.

So we have:

$$
\begin{aligned}
\mathbf{A\Sigma} &= \frac{c\mathbf{J}_n}{n}(a\mathbf{I}_n + b\mathbf{J}_n) = ac\frac{\mathbf{J}_n}{n} + bc\frac{\mathbf{J}_n^2}{n} \\
&= ac\frac{\mathbf{J}_n}{n} + nbc \cdot \frac{\mathbf{J}_n^2}{n^2} \\
&= ac \cdot \frac{\mathbf{J}_n}{n} + nbc \cdot \frac{\mathbf{J}_n}{n} \\
&= c(a + nb)\frac{\mathbf{J}_n}{n}
\end{aligned}
$$

Hence

$$(\mathbf{A\Sigma})^2 = c^2(a+nb)^2\frac{\mathbf{J}_n^2}{n^2} = c^2(a+nb)^2\frac{\mathbf{J}_n}{n}$$

In order for $(\mathbf{A\Sigma})^2 = \mathbf{A\Sigma}$*, we then need*

$$c(a + nb) = c^2(a+nb)^2$$

and hence

$$c(a+nb)=0 \ \textit{or}\ c(a+nb)=1$$

we can not take the first case since we already discussed $a+nb>0$ *in order for* $a\mathbf{I}_n+b\mathbf{J}_n$ *to be positive definite, also* $c\neq 0$ *otherwise* $cn\bar{\mathbf{y}}^2$ *is a constant equal to* 0 *and can not be chi square distributed.*

So we are left with choice

$$c(a+nb)=1 \Longrightarrow c=\frac{1}{a+nb}$$

Now for $\mathbf{y}^T(\mathbf{I}_n-d\mathbf{J}_n)\mathbf{y}$*, it is a quadraic form with*

$$\mathbf{A}=\mathbf{I}_n-d\mathbf{J}_n$$

So we have:

$$\begin{aligned}\mathbf{A\Sigma}&=(\mathbf{I}_n-d\mathbf{J}_n)(a\mathbf{I}_n+b\mathbf{J}_n)\\&=a\mathbf{I}_n+(b-ad)\mathbf{J}_n-bd\mathbf{J}_n^2\\&=a\mathbf{I}_n+n(b-ad)\frac{\mathbf{J}_n}{n}-n^2bd\frac{\mathbf{J}_n}{n}\\&=a\mathbf{I}_n+\Big[nb-nd(a+nb)\Big]\frac{\mathbf{J}_n}{n}\end{aligned}$$

So we have:

$$\begin{aligned}(\mathbf{A\Sigma})^2&=\Big(a\mathbf{I}_n+\Big[nb-nd(a+nb)\Big]\frac{\mathbf{J}_n}{n}\Big)^2\\&=a^2\mathbf{I}_n+2a\Big[nb-nd(a+nb)\Big]\frac{\mathbf{J}_n}{n}+\Big[nb-nd(a+nb)\Big]^2\frac{\mathbf{J}_n}{n}\end{aligned}$$

In order for $\mathbf{A\Sigma}=(\mathbf{A\Sigma})^2$*, we need*

$$\begin{cases}a\mathbf{I}_n=a^2\mathbf{I}_n\\ nb-nd(a+nb)=2a\Big[nb-nd(a+nb)\Big]+\Big[nb-nd(a+nb)\Big]^2\end{cases}$$

From the first equation, we need $a=a^2$*. Since we need* $a>0$ *to guarantee the positive definiteness of* $a\mathbf{I}_n+b\mathbf{J}_n$*, so we must have* $a=1$*. Plug this result into the second equation, we need*

$$\begin{aligned}&nb-nd(a+nb)=2\Big[nb-nd(a+nb)\Big]+\Big[nb-nd(a+nb)\Big]^2\\&\Longrightarrow\Big[nb-nd(a+nb)\Big]^2+\Big[nb-nd(a+nb)\Big]=0\\&\Longrightarrow\Big(nb-nd(a+nb)\Big)\Big[\Big(nb-nd(a+nb)\Big)+1\Big]=0\end{aligned}$$

so we need either:

$$
\begin{aligned}
& nb - nd(a + nb) = 0 \\
\Longrightarrow\ & d = \frac{b}{a + nb}
\end{aligned}
$$

or we need:

$$
\begin{aligned}
& nb - nd(a + nb) = -1 \\
\Longrightarrow\ & nd(a + nb) = nb + 1 \\
\Longrightarrow\ & d = \frac{nb + 1}{n(a + nb)}
\end{aligned}
$$

So to conclude, in order for $\mathbf{y}^T(\mathbf{I} - d\mathbf{J})\mathbf{y}$ *to be a chi square random variable, we need*

$$d = \frac{b}{a + nb}$$

or

$$d = \frac{nb + 1}{n(a + nb)} = \frac{1}{n}(\ since\ a = 1)$$

and also $a = 1$.

Question #5. (Exercise 6.14 from textbook)

Solution 5. *All solutions for this question is computed under proc iml of SAS:*

We manually created .csv file for the data, and read it into SAS:

```
dm'output;clear;log;clear;';

title "eruption data analysis";

ods html close;
ods listing;

libname data  "C:\akira\data";

/*Question 5*/
/*import the data*/
proc import datafile = "C:\akira\data\eruption.csv"
    dbms = dlm
/*  out = data.eruption*/
    out = work.eruption
    replace;
    delimiter = ',';
    getnames = yes;
run;

proc print data=eruption;
run;
```

We then start proc iml and import the data from the dataset we created.

```
proc iml;
use work.eruption;
read all var {y} into Y;
read all var{x} into X;

/*find the number of observations*/
d1 = nrow(X[, 1]);
print d1;

/*create column of 1s and insert into X*/
intercept = j(d1,1,1);
print intercept;
X_design = intercept||X;
print X_design;

/*check the dimension of design matrix*/
d = dimension(X_design);
print d;

/*create labels for matrices*/
cY = {"Y"};
cX = {"Intercept" "X"};
mattrib Y colname=cY X_design colname=cX;

/*check if the label works, this is optional*/
print Y;
print X_design;
```

For part (a):

We compute the estimate of $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)'$ with the following code:

```
/*part (a) OLS for beta*/
hat_beta = inv(t(X_design)*X_design)*t(X_design)*Y;
print hat_beta;
beta_0 = hat_beta[1, 1];
beta_1 = hat_beta[2, 1];
print beta_0;
print beta_1;
```

The output is:

```
beta_0

31.752285

beta_1

11.367804
```

which is our $\hat{\beta}_0$ ad $\hat{\beta}_1$.

. *For part (b):*

We test on hypothesis $H_0 : \beta_1 = 0$ versus $H_1 : \beta_1 \neq 0$ by using the statistic:

$$t = \frac{\hat{\beta}_1}{\sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n-2}} \cdot / \sqrt{\sum(x_i - \bar{x})^2}} = \frac{\hat{\beta}_1}{\sqrt{MSE}/\sqrt{\sum(x_i - \bar{x})^2}}$$

Under the null we have $t \sim t(n-2)$ while here $n = 53$. We have the following code to compute the t value and corresponding two-sided p-value:

```
/*part (b) inference on beta_1*/

/*SSE, MSE, SSX and t*/
SSE = sum((y - X_design*hat_beta)##2);
print SSE;
MSE = SSE/(d1-2);
print MSE;
SSX = sum((x - mean(x))##2);
print SSX;

/*t statistic*/
t = beta_1/sqrt(MSE/SSX);
/*p value for alpha = 0.05*/
p = 2*(1 - probt(abs(t), d1-2));
print "t = ",,,t;
print "two-sided p value = ",,,p;
```

The output is:

```
t

11.108592


p

3.109E-15
```

Since p value is highly significant here, we reject the null and conclude that there is a linear association between the interval to the next eruption and the duration of an eruption.

For part (c):

We use the formula:

$$\hat{\beta}_1 \pm t_{\alpha/2, n-2} \sqrt{MSE/\sum(x_i - \bar{x})^2}$$

with the following SAS code:

```
/*confidence interval for beta_1*/
CI_lower = beta_1 - tinv(0.975, d1-2)*sqrt(MSE/SSX);
CI_upper = beta_1 + tinv(0.975, d1-2)*sqrt(MSE/SSX);
print CI_lower;
print CI_upper;
```

The output is:

```
CI_lower

9.3133749

CI_upper

13.422234
```

Finally to compute r^2, we use the formula:

$$r^2 = \frac{SSR}{SST} = \frac{\sum(\hat{y}_i - \bar{\mathbf{y}})^2}{\sum(y_i - \bar{\mathbf{y}})^2}$$

with the SAS code:

```
/*R-square: coefficient of determination*/
SSR = sum((X_design*hat_beta - mean(y))##2);
SST = sum((y - mean(y))##2);
R_square = SSR/SST;
print R_square;

quit;
```

The output is:

```
R_square

0.7075702
```