1. Let $Y_i = \beta_0 + \beta_1(x_{i1} - \bar{x}_1) + \beta_2(x_{i2} - \bar{x}_2) + \epsilon_i$. Suppose $E[\epsilon] = 0, \mathrm{var}[\epsilon] = \sigma^2 I_n$. Let $\hat{\beta}_1$ be the least squares estimate of β_1. Show that

$$\mathrm{var}[\hat{\beta}_1] = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i1 - \bar{x}_1)^2(1 - r_{12}^2)},$$

where r is the correlation coefficient between x_1 and x_2. State the effects of using highly correlated predictors x_1 and x_2.

Solution. We first observe that the matrix X has the form

$$X = \begin{bmatrix} 1 & x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 \\ 1 & x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 \end{bmatrix}$$

We also know that $\mathrm{var}\,\hat{\beta} = \sigma^2(X'X)^{-1}$. By computation, we find that

$$X'X = \begin{bmatrix} n & \sum(x_{i1} - \bar{x}_1) & \sum(x_{i2} - \bar{x}_2) \\ \sum(x_{i1} - \bar{x}_1) & \sum(x_{i1} - \bar{x}_1)^2 & \sum(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \\ \sum(x_{i2} - \bar{x}_2) & \sum(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) & \sum(x_{i2} - \bar{x}_2)^2 \end{bmatrix}$$

But since we are only showing results that involve x_1 and x_2, then we can only consider a subset of the matrix X that involves both of these variable. Let A be the matrix,

$$A = \begin{bmatrix} \sum(x_{i1} - \bar{x}_1)^2 & \sum(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \\ \sum(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) & \sum(x_{i2} - \bar{x}_2)^2 \end{bmatrix}.$$

Since A is a $2x2$ matrix, we easily find its inverse by

$$A^{-1} = \frac{1}{|A|}\begin{bmatrix} \sum(x_{i2} - \bar{x}_2)^2 & -\sum(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \\ -\sum(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) & \sum(x_{i1} - \bar{x}_1)^2 \end{bmatrix}.$$

Recall, the covariance matrix is given by $\mathrm{cov}[\hat{\beta}] = \sigma^2 A^{-1}$. It follows that,

$$\begin{aligned}
\mathrm{var}[\hat{\beta}_1] &= \frac{\sigma^2 \sum(x_{i2} - \bar{x}_2)^2}{\sum(x_{i1} - \bar{x}_1)^2 \sum(x_{i2} - \bar{x}_2)^2 - [\sum(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2)]^2} \\
&= \frac{\sigma^2}{\sum(x_{i1} - \bar{x}_1)^2 - \frac{[\sum(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2)]^2}{\sum(x_{i2} - \bar{x}_2)^2}} \\
&= \frac{\sigma^2}{\sum(x_{i1} - \bar{x}_1)^2 \left(1 - \frac{[\sum(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2)]^2}{\sum(x_{i1} - \bar{x}_1)^2 \sum(x_{i2} - \bar{x}_2)^2}\right)} \\
&= \frac{\sigma^2}{\sum(x_{i1} - \bar{x}_1)^2(1 - r^2)},
\end{aligned}$$

where r^2 is the correlation coefficient between x_1 and x_2. We see that highly correlated pairs of variables will cause the denominator to get very small and thus, increase the variance.

2. Consider the full rank version of the one way ANOVA model,

$$Y^{2J\times 1} = \begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1J} \\ Y_{21} \\ \vdots \\ Y_{2J} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \epsilon_{11} \\ \vdots \\ \epsilon_{1J} \\ \epsilon_{21} \\ \vdots \\ \epsilon_{1J} \end{bmatrix} = X^{2J\times 2}\beta^{2\times 1} + \epsilon^{2J\times 1}$$

a. Interpret the parameters.
Solution. The Ys are the observed values. The X matrix are the indicator variables. The μs are the grouping variable. The ϵs are the residuals.
To find $\hat{\beta}$, we do

$$\begin{aligned} \hat{\beta} &= (X'X)^{-1}X'Y \\ &= \begin{bmatrix} \frac{1}{J} & 0 \\ 0 & \frac{1}{J} \end{bmatrix} \begin{bmatrix} \sum y_{1i} \\ \sum y_{2i} \end{bmatrix} \\ &= \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \end{bmatrix} \end{aligned}$$

We know $X\hat{\beta}$ is unique since X has full rank.
Through computations, we find

$$P = X(X'X)^{-1}X' = \begin{bmatrix} 1/J & \cdots & 1/J & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 1/J & \cdots & 1/J & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1/J & \cdots & 1/J \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 1/J & \cdots & 1/J \end{bmatrix}$$

■

3. Now consider the non full rank version of the one way ANOVA model,

$$Y^{2J\times 1} = \begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1J} \\ Y_{21} \\ \vdots \\ Y_{2J} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \epsilon_{11} \\ \vdots \\ \epsilon_{1J} \\ \epsilon_{21} \\ \vdots \\ \epsilon_{1J} \end{bmatrix} = X^{2J\times 3}\beta^{3\times 1} + \epsilon^{2J\times 1}$$

a. Interpret the parameters. Answer the rest of the questions for question 3 from the sheet.
Solution. The Ys are the observed values. The X matrix are the indicator variables. The μ is the common component of the groups, and the αs are the grouping variables. The ϵs are the residuals. We know that $\hat{\beta}$ will not be unique since X is not of full rank.
To compute P, we work it out on paper to get:

$$\begin{aligned} P &= X(X'X)^{-}X' \\ &= \begin{bmatrix} 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \end{bmatrix} \begin{pmatrix} \frac{2}{9j} & \frac{1}{9j} & \frac{1}{9j} \\ \frac{1}{9j} & \frac{5}{9j} & -\frac{4}{9j} \\ \frac{1}{9j} & -\frac{4}{9j} & \frac{5}{9j} \end{pmatrix} \begin{bmatrix} 1 & \cdots & 1\cdots & 1 \\ 1 & \cdots & 0\cdots & 0 \\ 0 & \cdots & 1\cdots & 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{3J} & \frac{2}{3J} & 0 \\ \vdots & \vdots & \vdots \\ \frac{1}{3J} & -\frac{1}{3J} & \frac{2}{3J} \\ \vdots & \vdots & \vdots \\ \frac{1}{3J} & -\frac{1}{3J} & \frac{2}{3J} \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1\cdots & 1 \\ 1 & \cdots & 0\cdots & 0 \\ 0 & \cdots & 1\cdots & 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{J} & \cdots & \frac{1}{3J} & \cdots & \frac{1}{3J} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & \frac{1}{3J} & \cdots & \frac{1}{3J} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & \frac{1}{3J} & \cdots & \frac{1}{3J} \end{bmatrix} \end{aligned}$$

4. Let $Y \sim N(X\beta, \sigma^2 I)$. Suppose X has full rank.

a. Find $Var[S^2]$. To find this, we use the fact that $\frac{(n-p)S^2}{\sigma^2} \sim \chi^2_{n-p}$ and $\text{var}[\chi^2_{n-p}] = 2(n-p)$. Thus,

$$\begin{aligned}
\text{var}[S^2] &= \text{var}[\frac{\sigma^2}{(n-p)} \cdot \frac{(n-p)}{\sigma^2} S^2] \\
&= \frac{\sigma^4}{(n-p)^2} \text{var}[\frac{(n-p)S^2}{\sigma^2}] \\
&= \frac{\sigma^4}{(n-p)^2} \cdot 2(n-p) \\
&= \frac{2\sigma^4}{(n-p)}
\end{aligned}$$

■

b. Evaluate $E[(Y'AY - \sigma^2)^2]$ for $A = \frac{1}{n-p+2}(I - X(X'X)^{-1}X')$

Solution. Let $A = \frac{1}{n-p+2}(I - X(X'X)^{-1}X')$. Let $R = (I - X(X'X)^{-1}X')$. Then $Y'AY - \sigma^2$ becomes $\frac{1}{n-p+2} Y'RY - \sigma^2$. Thus,

$$\begin{aligned}
E[(Y'AY)^2] &= \text{Var}[(Y'AY)^2] + E[(Y'AY)^2]^2 \\
&= \text{Var}[\left(\frac{Y'RY}{n-p+2}\right)] + E[\left(\frac{Y'RY}{n-p+2}\right)]^2 \\
&= \frac{1}{(n-p+2)^2} \text{Var}[Y'RY] + \left[\frac{1}{(n-p+2)} E[Y'RY] - \sigma^2\right]^2 \\
&= \frac{1}{(n-p+2)^2} \left(\frac{\sigma^4(\mu_4 - 3\sigma^4)}{\sigma^4} a'a + 2\sigma^4 tr(R^2)\right) + \left[\frac{1}{(n-p+2)}(n-p)\sigma^2 - \sigma^2 \frac{(n-p+2)}{(n-p+2)}\right]^2 && \text{by thms 1.5, 1.6} \\
&= \frac{1}{(n-p+2)^2} \left(2\sigma^4(n-p)\right) + \left[\frac{\sigma^2(n-p-n+p-2)}{n-p+2}\right]^2 && \text{since } \mu_4 = 3\sigma^4 \\
&= \frac{1}{(n-p+2)^2} \left(2\sigma^4(n-p)\right) + \left[\frac{\sigma^2(-2)}{n-p+2}\right]^2 \\
&= \frac{2\sigma^4(n-p) + 4\sigma^4}{(n-p+2)^2} \\
&= \frac{2\sigma^4(n-p+2)}{(n-p+2)^2} \\
&= \frac{2\sigma^4}{(n-p+2)} && \text{which is the mean squared error.}
\end{aligned}$$

□

4c. Let $\hat{\sigma}^2 = Y'AY$. Show that $\text{Var}[\hat{\sigma}^2] \leq \text{Var}[S^2]$,and thus S^2 does not have minimum mean squared error among estimates of σ^2.

Solution. We will use computations similar to the ones on 4b to show the inequality. Consider,

$$\begin{aligned}
\text{Var}[\hat{\sigma}^2] &= \text{Var}[Y'AY] \\
&= \frac{1}{(n-p+2)^2}\,\text{Var}[Y'RY] \\
&= \frac{1}{(n-p+2)^2} 2\sigma^4(n-p) \qquad \text{shown in 4b} \\
&\leq \frac{2\sigma^4}{(n-p+2)^2} \\
&\leq \frac{2\sigma^4}{(n-p)} \\
&= \text{Var}[S^2] \qquad \text{shown in 4a.}
\end{aligned}$$

□