1. Consider the GM model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{X}$ is $n \times p$ with rank $r \leq p$, $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $\text{cov}(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}$. Let $\widehat{\boldsymbol{\beta}}$ denote a least squares estimator of $\boldsymbol{\beta}$ and suppose that $\boldsymbol{\lambda}'\boldsymbol{\beta}$ is estimable.
(a) Assume $\mathbf{a}'\mathbf{Y}$ is an unbiased estimator of $\boldsymbol{\lambda}'\boldsymbol{\beta}$. What the Gauss-Markov Theorem say about $\text{var}(\mathbf{a}'\mathbf{Y})$ when compared to $\text{var}(\boldsymbol{\lambda}'\widehat{\boldsymbol{\beta}})$?
(b) Let $\widehat{\boldsymbol{\mu}} = \mathbf{X}\widehat{\boldsymbol{\beta}}$ and $\widetilde{\boldsymbol{\mu}} = \mathbf{A}'\mathbf{Y}$ for a nonrandom $n \times n$ matrix $\mathbf{A}$. Assume that $E(\mathbf{A}'\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$. Show that $\text{cov}(\widetilde{\boldsymbol{\mu}}) - \text{cov}(\widehat{\boldsymbol{\mu}})$ is nnd.
(c) Let $\widetilde{\boldsymbol{\mu}} = \mathbf{A}'\mathbf{Y}$ and assume that $E(\mathbf{A}'\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$. Show that

$$\sum_{i=1}^{n} \text{var}(\widetilde{\mu}_i) \geq r\sigma^2,$$

where $\widetilde{\mu}_i$ is the ith component of $\widetilde{\boldsymbol{\mu}}$.

2. Suppose that $\mathbf{V}$ is pd. Prove that $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} - (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$ is nnd.

3. Consider the linear model defined by

$$\begin{aligned} Y_1 &= 2\theta + \epsilon_1 \\ Y_2 &= \theta + \epsilon_2, \end{aligned}$$

where $\epsilon_1 = 2Z_1 - Z_2$ and $\epsilon_2 = Z_1 + 2Z_2$, and Z_1 and Z_2 are independent random variables with zero mean and constant variance σ^2.
(a) Write this model in $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ form. Find $E(\mathbf{Y})$ and $\text{cov}(\mathbf{Y})$.
(b) Compute the ordinary least squares (OLS) estimator of θ.
(c) Compute the generalized least squares (GLS) estimator of θ.
(d) Show that the OLS and GLS estimators are both unbiased.
(e) Compute the variance of both estimators and compare.

4. Suppose that $Y_1, Y_2, ..., Y_n$ is an iid $\mathcal{U}(0, 2\theta)$ sample, where $\theta > 0$. Define $\epsilon_i = Y_i - \theta$, for $i = 1, 2, ..., n$.
(a) Find the mean and covariance matrix of $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, ..., \epsilon_n)'$.
(b) Show that $\mathbf{Y} = (Y_1, Y_2, ..., Y_n)'$ follows a Gauss-Markov model.
(c) Find the BLUE of θ, say $\widehat{\theta}_{\text{OLS}}$. Give both a matrix expression and an expression in terms of simple summary statistics.
(d) Find c so that $\widehat{\theta} = cY_{(n)}$, where $Y_{(n)} = \max\{Y_1, Y_2, ..., Y_n\}$, is unbiased for θ and compute the variance of $\widehat{\theta}$.
(e) Compare the variances of $\widehat{\theta}_{\text{OLS}}$ and $\widehat{\theta}$. Are you surprised? Explain your findings in light of the Gauss-Markov Theorem.

5. Suppose that $Y_1 \sim \text{beta}(\alpha, \beta)$, where $\alpha = \beta = 2$. Let $Y_2 = 1 - Y_1$, $Y_3 = Y_1Y_2$, and $\mathbf{Y} = (Y_1, Y_2, Y_3)'$. Compute $E(\mathbf{Y}'\mathbf{A}\mathbf{Y})$, where $\mathbf{A}$ is the matrix in HW 2, Problem 10.

6. Consider the linear regression model

$$\begin{aligned}\mathbf{Y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \\ &= \mathbf{1}\beta_0 + \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\epsilon},\end{aligned}$$

where $\mathbf{Y}$ is $n \times 1$, $\mathbf{1}$ is an $n \times 1$ vector of ones, $\mathbf{X}_1$ is an $n \times p$ matrix of (full) rank p, $\mathbf{X} = (\mathbf{1}\ \ \mathbf{X}_1)$ is an $n \times (p+1)$ matrix with $\mathbf{1}'\mathbf{X}_1 = \mathbf{0}$, β_0 is a fixed (scalar) parameter, $\boldsymbol{\beta}_1$ is a $(p \times 1)$ vector of fixed parameters, and $\boldsymbol{\beta} = (\beta_0, \boldsymbol{\beta}_1')'$. Suppose that $\boldsymbol{\epsilon}$ has zero mean and covariance matrix

$$\mathbf{V} = \sigma^2\{(1-\rho)\mathbf{I} + \rho\mathbf{J}\},$$

where $\mathbf{J} = \mathbf{1}\mathbf{1}'$, $\mathbf{I}$ is an $n \times n$ identity matrix, $-1 < \rho < 1$, and $\sigma^2 > 0$.
(a) Show that

$$\mathbf{V}^{-1} = \{\sigma^2(1-\rho)\}^{-1}(\mathbf{I} + b\mathbf{J}),$$

where $b = -\rho/(1-\rho+n\rho)$.
(b) Suppose that ρ is known. Show that the ordinary least squares (OLS) and the generalized least squares (GLS) estimators for $\boldsymbol{\beta} = (\beta_0, \boldsymbol{\beta}_1')'$ are identical.
(c) Find the covariance matrix of the OLS estimator of $\boldsymbol{\beta}_1$.
(d) Let $\widehat{\mathbf{Y}}$ denote the vector of predicted values obtained by fitting the model by OLS, and let $\widehat{\mathbf{e}} = \mathbf{Y} - \widehat{\mathbf{Y}}$. Show that the residual mean square

$$\text{MSE} = \frac{\widehat{\mathbf{e}}'\widehat{\mathbf{e}}}{n-p-1}$$

is an unbiased estimator for $\sigma^2(1-\rho)$.

7. Suppose that $\mathbf{X}$ is $n \times p$ with rank r.
(a) For $\mathbf{a}_{n\times 1}$, show that $\mathbf{P}_\mathbf{X}\mathbf{a} = \mathbf{0}$ if and only if $\mathbf{X}'\mathbf{a} = \mathbf{0}$, where $\mathbf{P}_\mathbf{X}$ is the perpendicular projection matrix onto $\mathcal{C}(\mathbf{X})$.
(b) Assume the GM model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ holds; i.e., $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $\text{cov}(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}$. Suppose that $\boldsymbol{\lambda}'\boldsymbol{\beta}$ is estimable. Prove: If $\widehat{\boldsymbol{\beta}}$ solves the normal equations, then $0 < \text{var}(\boldsymbol{\lambda}'\widehat{\boldsymbol{\beta}}) < \infty$.

8. Let $Y_1, Y_2, ..., Y_n$ be uncorrelated observations with $E(Y_i) = \mu$ and $\text{var}(Y_i) = \sigma^2/w_i$, $i = 1, 2, ..., n$, where $w_1, w_2, ..., w_n$ are fixed constants.
(a) Find the BLUE of μ and find the variance of this estimator.
(b) Find the OLS estimator of μ and its variance.

9. Consider the Aitken model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{X}$ is $n \times p$ with rank $r \leq p$, $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $\text{cov}(\boldsymbol{\epsilon}) = \sigma^2\mathbf{V}$, $\mathbf{V}$ known.
(a) Derive the mean and covariance matrix of the ordinary least squares (OLS) estimator.
(b) Derive the mean and covariance matrix of the generalized least squares (GLS) estimator.
(c) If $r = p$, show that $\text{cov}(\widehat{\boldsymbol{\beta}}_{\text{OLS}} - \widehat{\boldsymbol{\beta}}_{\text{GLS}}) = \text{cov}(\widehat{\boldsymbol{\beta}}_{\text{OLS}}) - \text{cov}(\widehat{\boldsymbol{\beta}}_{\text{GLS}})$.

10. Consider the linear regression model $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where $\epsilon_1, \epsilon_2, ..., \epsilon_n$ are uncorrelated zero mean random variables with common variance σ^2.
(a) Show that the least squares estimator of β_1 under the model when $\beta_0 = 0$ is given by

$$\widehat{\beta}_1^* = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}.$$

(b) Find the bias of $\widehat{\beta}_1^*$ when, in fact, $\beta_0 \neq 0$.
(c) Derive the variance of $\widehat{\beta}_1^*$ and the variance of the least squares estimator $\widehat{\beta}_1$ when β_0 is not known. Comment on which estimator has the smallest variance and whether or not this contradicts the Gauss-Markov Theorem.

11. Consider the Aitken model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $\text{cov}(\boldsymbol{\epsilon}) = \sigma^2\mathbf{V}$, $\mathbf{V}$ known and pd. Recall that the GLS estimator $\widehat{\boldsymbol{\beta}}_{\text{GLS}}$ minimizes

$$Q^*(\boldsymbol{\beta}) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}).$$

(a) Show that $Q^*(\widehat{\boldsymbol{\beta}}) = \mathbf{Y}'(\mathbf{I} - \mathbf{A})'\mathbf{V}^{-1}(\mathbf{I} - \mathbf{A})\mathbf{Y}$, where $\mathbf{A} = \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}$. Note that $Q^*(\widehat{\boldsymbol{\beta}})$ is the residual sum of squares in the Aitken model fit.
(b) Suppose that $r(\mathbf{X}) = r$. Under the Aitken model, show that $Q^*(\widehat{\boldsymbol{\beta}})/(n - r)$ is an unbiased estimator of σ^2.

12. Suppose that $\mathbf{Y} = (Y_1, Y_2, Y_3)'$ has a trivariate normal distribution; specifically,

$$\mathbf{Y} \sim \mathcal{N}_3\left\{\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & -1 \\ 0 & -1 & 4 \end{pmatrix}\right\}.$$

(a) Define

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 0 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Compute $E(\mathbf{Y}'\mathbf{A}\mathbf{Y})$.
(b) The variance of a quadratic form $\mathbf{Y}'\mathbf{A}\mathbf{Y}$, $\mathbf{A}$ symmetric, when $\mathbf{Y}$ is multivariate normal with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{V}$, (which it is here), is

$$V(\mathbf{Y}'\mathbf{A}\mathbf{Y}) = 2[\text{tr}(\mathbf{A}\mathbf{V})]^2 + 4\boldsymbol{\mu}'\mathbf{A}\mathbf{V}\mathbf{A}\boldsymbol{\mu}.$$

Find $V(\mathbf{Y}'\mathbf{A}\mathbf{Y})$. Note that this formula is not correct when $\mathbf{Y}$ is not multivariate normal.