Midterm Exam (3pm-5pm on April, 13, 2022)

Student ID: _____ **Name:** _____

Problems:

1. Let U and W be subspaces in $\mathbb{R}^d$ satisfying $u^T w = 0$ for any $u \in U$ and $w \in W$. Let $U \oplus W = \{\boldsymbol{u} + \boldsymbol{w} : \boldsymbol{u} \in V, \ \ \boldsymbol{w} \in W\}$. Prove or disprove that

$$\Pi(\ \cdot\ |U \oplus W) = \Pi(\ \cdot\ |U) + \Pi(\ \cdot\ |W),$$

where for any subspace S in $\mathbb{R}^d$, $\Pi(\cdot\,|S)$ is the projection operator on S, that is, for any vector $\boldsymbol{v} \in \mathbb{R}^d$, $\Pi(\boldsymbol{v}\,|S)$ is the projection vector of $\boldsymbol{v}$.

2. Consider the following multiple linear regression: $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{Y} = (Y_1, \ldots, Y_n)^\top$, $\mathbf{X} = (\mathbf{1}, \mathbf{x}_1, \ldots, \mathbf{x}_p)^\top$ with $\mathbf{1} = (1, \ldots, 1)^\top$ and $\mathbf{x}_j = (x_{1j}, \ldots, x_{nj})^\top$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)^\top$ and $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)^\top \sim N_n(\mathbf{0}, \sigma^2 I_n)$. Assume that $\mathbf{X}^\top\mathbf{X}$ is invertible.

   (a) Prove that the LSE $\hat{\boldsymbol{\beta}}$ is BLUE (Best Linear Unbiased Estimator).

   (b) Prove that the above (a) implies $\mathrm{var}(\boldsymbol{\lambda}^\top\tilde{\boldsymbol{\beta}}) - \mathrm{var}(\boldsymbol{\lambda}^\top\hat{\boldsymbol{\beta}}) \geq 0$ for any $\boldsymbol{\lambda} \in \mathbb{R}^{p+1}$.

3. This problem consists of three parts.

   (a) Consider the multiple linear regression model

$$Y = X_1\beta_1 + X_2\beta_2 + \epsilon,$$

   where Y is $n \times 1$, X_1 is $n \times p_1$, β_1 is $p_1 \times 1$, X_2 is $n \times p_2$, β_2 is $p_2 \times 1$, $\epsilon \sim N_n(0_n, \sigma^2 I_n)$ is $n \times 1$, and $X_1^T X_1$ and $X_2^T X_2$ are invertible. Suppose that in fact $\beta_2 = 0$, in other words, the model used by the experimenter is an overfitted model and the true model is

$$Y = X_1\beta_1 + \epsilon.$$

   Let $\hat{\sigma}_o^2$ denote the usual estimate of variance based on the overfitted model, i.e., $\hat{\sigma}_o^2 = Y^T(I - P)Y/(n - p_1 - p_2)$, where P is the projection onto the space spanned by the columns of X_1 and the columns of X_2. Prove that $\hat{\sigma}_o^2$ is an unbiased estimate of σ^2 even if the smaller model is true.

   (b) Find 95% confidence interval for σ^2 based on the overfitted model.

   (c) Let $\hat{\sigma}_r^2$ be the estimate of σ^2 based on the reduced (and correct) model

$$Y = X_1\beta_1 + \epsilon.$$

   Show that the expected length of the confidence interval for σ^2 based on the reduced model is smaller than under the overfitted model.

4. Consider the following regression model

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \qquad i = 1, 2, \ldots, n.$$

The predicted responses $\hat{\mathbf{Y}} = (\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_n)^\top$ with $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$. Let $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)^\top$, $\mathbf{1} = (1, 1, \ldots, 1)^\top$, and $\mathbf{x} = (x_1, x_2, \ldots, x_n)^\top$.

(a) Show that $\mathbf{Y} - \hat{\beta}_0\mathbf{1} - \hat{\beta}_1\mathbf{x}$ is orthogonal to the linear subspace $\mathcal{C}_{\mathbf{1},\mathbf{x}}$ (in $\mathbb{R}^n$) spanned by the vectors $\mathbf{1}$ and $\mathbf{x}$ in a sense that $\mathbf{Y} - \hat{\beta}_0\mathbf{1} - \hat{\beta}_1\mathbf{x}$ is orthogonal to any vector $\mathbf{v} \in \mathcal{C}_{\mathbf{1},\mathbf{x}}$ (Do not use properties of projection when showing it).
(Hint: two vectors $v = (v_1, \ldots, v_n)^\top, w = (w_1, \ldots, w_n)^\top$ in $\mathbb{R}^n$ are orthogonal if $v^\top w = 0$ with $v^\top w = \sum_{i=1}^n v_i w_i$.)

(b) Show that $R^2 = (\cos\theta)^2$, where θ is the angle between $Y - \bar{Y} \cdot \mathbf{1}$ and $\hat{Y} - \bar{Y} \cdot \mathbf{1}$.

5. Suppose that the variables are all standardized and orthogonal in the sense that

$$\sum_{i=1}^n x_{ij} = 0, \quad n^{-1}\sum_{i=1}^n x_{ij}^2 = 1, \quad \sum_{i=1}^n x_{ij}x_{ik} = 0$$

for $1 \le j \ne k \le p$.

(a) Find the least square estimate $\hat{\beta}_j$ and its variance $\mathrm{var}(\hat{\beta}_j)$.

(b) Based on general linear hypothesis, perform the following hypothesis testing in details:

$$H_0 : A\beta = 0 \quad \text{vs} \quad H_1 : \text{not } H_0,$$

where $\beta = (\beta_1, \cdots, \beta_p)^T$ and A is a following $3 \times p$ matrix:

$$\begin{pmatrix} 1 & 0 & 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & 1 & 0 & \cdots & 0 \\ 2 & 1 & 1 & -1 & 0 & \cdots & 0 \end{pmatrix}.$$