1. [10] I have two almost identical regressors, $b = 100a$, one is measured in dollars and the other one — in cents. Consider the ridge regression loss function,

$$\text{loss}(\hat\beta_a, \hat\beta_b) = \sum_i (y_i - \hat y_i)^2 + \lambda \cdot (\hat\beta_a^2 + \hat\beta_b^2), \quad \hat y_i = \hat\beta_a a_i + \hat\beta_b b_i.$$

   (a) [7] Find the penalized estimates $\hat\beta_a$ and $\hat\beta_b$ for arbitrary $\lambda > 0$.

   In the regression $\hat y_i = \hat\gamma_1 a_i$ the OLS estimator $\hat\gamma_1$ is equal to 2.

   (b) [3] What are the approximate values of $\hat\beta_a$ and $\hat\beta_b$ in the penalized model for small $\lambda \approx 0$?

2. [10] Elon Musk studies the model $y_i = \beta_0 + \beta_1 a_i + \beta_2 b_i + u_i$. He has 500 observations and he knows the heteroskedasticity form up to an unknown constant. To obtain the most efficient unbiased estimators $\hat\beta$ he has divided each observation by $a_i b_i$.

   (a) [5] What is the variance $\mathbb{V}\text{ar}(u_i \mid X)$ supposed by Elon Musk?

   (b) [5] To check whether the heteroskedasticity is really present Elon Musk estimated the regression

$$\hat v_i = \hat\gamma_0 + \hat\gamma_1 a_i + \hat\gamma_2 b_i + \hat\gamma_3 a_i b_i, \quad R^2 = 0.05,$$

   where v_i are the squared residuals $\hat u_i^2$ from original model.
   What is the statistical conclusion about the heteroskedasticity presence?

   Critical values for 5% significance level: $\chi^2_1 = 3.84$, $\chi^2_2 = 5.99$, $\chi^2_3 = 7.81$, $\chi^2_{498} = 549$, $\chi^2_{499} = 551$, $\chi^2_{500} = 553$.

3. [10] Donald Trump has 403 observations. He estimated the first simple regression:

$$\hat x_i = 2 - 2w_i, \quad R^2 = 0.81, \quad \text{SST} = 100 \quad \text{(Regression A)}.$$

   Than he estimated the second regression,

$$\hat w_i = \hat\gamma_0 + \hat\gamma_1 x_i, \quad \text{SST} = 200 \quad \text{(Regression B)}.$$

   (a) [2] Find R^2 in the regression B.

   Than he estimated the third regression,

$$\hat y_i = 2 + 3x_i + 5w_i, \quad R^2 = 0.6, \quad \text{SST} = 500 \quad \text{(Regression C)}.$$

   (b) [3] Find the variance inflation factors of x_i and w_i in the regression C.

   (c) [5] Find the 95% confidence interval for $\partial y / \partial w$.

4. [10] All regressors in X matrix are standardized. There are 3 regressors and 10000 observations. You partially know the singular value decomposition of X:

$$X = UDV^T, \text{diag}(D) = (5 \cdot 10^5, 10^5, 10^4), V = \begin{pmatrix} 0.000156 & 0.005953 & -0.999982 \\ 0.007737 & 0.999952 & 0.005954 \\ 0.999970 & -0.007737 & 0.000109 \end{pmatrix}.$$

   (a) [3] Find the sample variance of each principal component.

   (b) [3] How much variance in % does the first principal component explain?

   (c) [4] Express the first original regressor as the function of the three principal components.

5. [10] Consider the model $y_i = \beta_0 + \beta_1 x_i + u_i$ with independent and identically distributed observations. The regressor x_i is correlated with u_i and with variable z_i. The variable z_i is uncorrelated with u_i.

   The variable z_i is binary and takes only values 0 or 1.

   The regressor x_i is continuous and all observations may be divided into two groups:

Group	$\bar{y}_z$	$\bar{x}_z$	n_z
$z = 0$	5	4	500
$z = 1$	7	2	800

   Here $\bar{y}_z$, $\bar{x}_z$ and n_z are average values of y_i, x_i and the number of observations for each group.

   (a) [8] Find the instrumental variable estimate $\hat{\beta}_1$.

   (b) [2] Will the estimate change if the labels $z = 0$ and $z = 1$ will be switched?

6. (based on past LSE exams) Consider the simple linear regression model:

$$y_i = \beta_0 + \beta_1 x_i + u_i, \quad i = 1, \ldots, n.$$

   We assume that the errors $\{u_i\}_{i=1}^n$ are independent random variables with zero mean.

   The regressor $\{x_i\}_{i=1}^n$ is stochastic and Gauss-Markov conditions are satisfied. Under these conditions, the OLS estimator for β_1, $\hat{\beta}_1$, is conditionally unbiased. You are not asked to derive $\hat{\beta}_1$.

   (a) [2] Explain the concept of unbiasedness of an estimator.

   (b) [5] Let us consider two other estimators for the slope β_1:

$$\hat{\beta}_1^\circ = \frac{\sum_{i=1}^n (z_i - \bar{z}) y_i}{\sum_{i=1}^n (z_i - \bar{z}) x_i} \quad \text{and} \quad \hat{\beta}_1^* = \frac{\sum_{i=1}^n (z_i - \bar{z}) y_i}{\sum_{i=1}^n (z_i - \bar{z}) z_i}$$

   where $z_i = \sqrt{x_i}$ for all i and $\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i$. Please indicate whether $\hat{\beta}_1^\circ$ and $\hat{\beta}_1^*$ are conditionally unbiased estimators for β_1. Clearly show your derivations.

   (c) [3] Briefly indicate how you would choose between the three estimators, $\hat{\beta}_1$, $\hat{\beta}_1^\circ$, and $\hat{\beta}_1^*$.