Econometrics Assignment-Optional Exercises

- The exercises are optional. The maximum grade that the proposed solutions can attain is 2. The grade is additive to the grade of the final exam
- The instructors retain the right to ask for clarifications on the proposed solutions in due time.
- The proposed solutions can be only delivered electronically and through the open assignment in the eclass of the course. They have to be strictly delivered during the period beginning in 15/01/23, 23:00, and ending in **19**/02/23, 23:00. Any overdue delivery will not be accepted!
- The grade distribution between the QUESTIONS need not be the discrete uniform.

Notation: For a matrix X we denote by $X^\top$ its transpose; all the vectors should understood as column vectors, i.e., a matrix with a single column; for a vector $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$, $\bar{\boldsymbol{x}} = (1/n)\sum_{i=1}^n x_i$.

QUESTION 1

Consider the model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$ and suppose that application of least squares to 20 observations on these variables yields the following results

$$\hat\beta_0 = 0.96587,\ \hat\beta_1 = 0.69914,\ \hat\beta_2 = 1.7769$$

$$\widehat{\text{Cov}}(\hat{\boldsymbol{\beta}}) = \begin{bmatrix} 0.21812 & 0.019195 & -0.050301 \\ 0.019195 & 0.048526 & -0.031223 \\ -0.050301 & -0.031223 & 0.037120 \end{bmatrix} \equiv \hat\sigma^2 (X^\top X)^{-1},$$

where $\hat{\boldsymbol{\beta}} = (\hat\beta_0, \hat\beta_1, \hat\beta_2)^\top$.

(i) Find the total variation, unexplained variation, and explained variation for this model.

(ii) Find 95% interval estimates for β_1 and β_2.

(iii) Use a t-test to test the hypothesis $H_0 : \beta_1 = 1$ against the two-sided alternative.

(iv) Use your answers in (i) to test the joint hypothesis $H_0 : \beta_1 = \beta_2 = 0$.

(v) Test the hypothesis $H_0 : 2\beta_1 = \beta_2$.

QUESTION 2

In a classical linear regression model with 120 observations, the following results are obtained:

$$\hat\beta_0 = 2,\ \hat\beta_1 = 1,\ \hat\beta_2 = 4$$

$$\widehat{\text{Cov}}(\hat{\boldsymbol{\beta}}) = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 4 & -5 \\ 1 & -5 & 9 \end{bmatrix} \equiv \hat\sigma^2 (X^\top X)^{-1},$$

where $\hat{\boldsymbol{\beta}} = (\hat\beta_0, \hat\beta_1, \hat\beta_2)^\top$.

Perform the following tests and be explicit about the critical values, degrees of freedom, etc.

(i) Test whether each true parameter is equal to zero at 5% level against two-sided alternative hypothesis?

(ii) Test that $\beta_1 = \beta_2$ at 5% level.

(iii) Give an expression which can be used to construct a 95% confidence region for β_2 and β_3.

QUESTION 3

(i) Let A an $n \times k$ matrix, $k \leq n$. Show that, if $rank(A) = k$ then $A^\top A$ is a positive definite matrix.

(ii) Let R a $J \times K$ matrix of full row-rank ($rank(R) = J$) and X an $n \times k$ matrix with all the available covariates which is of full column rank. Show that $R(X^\top X)^{-1}R^\top$. *Hint: Consider that the inverse of a positive definite matrix is positive definite.)*

(iii) Show that a variance-covariance matrix is positive semi-definite. Give conditions that ensure that it is positive definite.

(iv) Prove that the following is true when an intercept is included in the regression line:

$$R^2 \equiv \frac{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{[\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})(\hat{y}_i - \bar{\hat{y}})]^2}{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2 \sum_{i=1}^n (y_i - \bar{y})^2} \equiv r^2.$$

Hint: a) you can prove the result by using two different methods, b) for one of them you need to use the fact that $\hat{y}^\top \hat{\epsilon} = \hat{\boldsymbol{\beta}}^\top X^\top \hat{\epsilon} = 0$ *which follows from the Normal equations.*

QUESTION 4

Part A:

Let $\hat{\beta}$ be the OLS estimator of β in the linear model $y = X\beta + \epsilon$, where X is $n \times K$ matrix, and $\tilde{\beta}$ is any $K \times 1$ vector. Prove that

$$(y - X\tilde{\beta})^\top (y - X\tilde{\beta}) \geq (y - X\hat{\beta})^\top (y - X\hat{\beta}).$$

Hint: Show that

$$(y - X\tilde{\beta})^\top (y - X\tilde{\beta}) - (y - X\hat{\beta})^\top (y - X\hat{\beta}) = (\tilde{\beta} - \hat{\beta})^\top X^\top X(\tilde{\beta} - \hat{\beta}),$$

conclude that the right hand side is positive.

Part B:

Let $\mathbf{1}$ the n-dimensional vector of ones and let M_1 be its annihilator matrix, i.e., $M_1 = I_n - \mathbf{1}(\mathbf{1}^\top \mathbf{1})^{-1}\mathbf{1}^\top$. Prove the following

(i) M_1 is symmetric and idempodent.

(ii) $M_1 \mathbf{1} = 0$

(iii) $M_1 \mathbf{1} = \boldsymbol{y} - \bar{y} \cdot \mathbf{1}$, where $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$.

QUESTION 5

Part A:

Consider the linear model $y = X\beta + \epsilon$, where y is n-dimensional vector, X is $n \times K$ matrix and $\epsilon \sim N(0, \sigma^2 I_n)$. Let $\hat{\epsilon}$ the n-dimensional vector with residuals find the exact distribution of $\hat{\epsilon}^\top \hat{\epsilon}/\sigma^2$.

Part B:

Let $\tilde{\sigma}^2$ be the maximum likelihood estimator of σ^2 -recall that it does not 'correct' for the degrees of freedom. Show that

$$\tilde{\sigma}^2 \xrightarrow{m.s.} \sigma^2.$$

Is $\tilde{\sigma}^2$ a consistent estimator of σ^2?

QUESTION 6

In many practical fields (e.g., finance) it is common to encounter noise whose distribution has much heavier tails than any Gaussian could give us. One way to model this is with t distributions. Consider, therefore, the model where $y = \beta_0 + \beta_1 X + \epsilon$, where y and X are n-dimensional vectors and $\epsilon/\sigma \sim t_\nu$; ϵ and X are independent random vectors as well as independent across observations. That is, rather than having a Gaussian distribution, the noise follows a t distribution with degrees of freedom (after scaling).

(i) Write down the log-likelihood function. *Hint: If a random variable Z follows the t_ν distribution then its probability density function is given by*

$$f_Z(z) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{z^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

(ii) Find the derivatives of this log-likelihood with respect to the four parameters β_0, β_1, σ^2 and ν. Simplify as much as possible. *Hint: You may have to express your answer in terms of the gamma function and its derivatives.*

(iii) Can you solve for the maximum likelihood estimators of the four parameters? If not, why not? If you can, do they match the least-squares estimators again? If they don't match, how do they differ?

QUESTION 7

Consider the (generalized) linear model written as:

$$y = X\beta + \epsilon, \;\; E(\epsilon|X) = 0, \;\; E(\epsilon\epsilon^\top|X) = \sigma^2\Omega.$$

In the following suppose Ω is known and in all cases X is taken as given.

(i) What is the covariance matrix, $\mathrm{Cov}(\tilde{\beta}, \tilde{\beta} - \hat{\beta})$, where $\tilde{\beta}$ is the GLS estimator and $\hat{\beta}$ is the OLS estimator?

(ii) What is the covariance matrix of the OLS and GLS estimators of β?

(iii) What is the covariance matrix of the OLS residual vector $\hat{\epsilon} = y - X\hat{\beta}$?

(iv) What is the covariance matrix of the HLS residual vector $\tilde{\epsilon} = y - X\tilde{\beta}$?

(v) What is the covariance matrix of the OLS and GLS residual vectors, i.e, $\mathrm{Cov}(\hat{\epsilon}, \tilde{\epsilon}|X)$?

QUESTION 8

For an arbitrary (semi-)parametric statistical model, and a criterion function $c_n(z_n, \theta)$ emerging from it, remember that an OE (optimization based estimator) obeying this structure was defined by

$$\theta_n \in \arg\min_{\theta\in\Theta} c_n(z_n, \theta),$$

for $\Theta \subseteq \mathbb{R}^p$ the underlying parameter space. Consider the following generalization of the definition: for u_n a random variable (that can among others represent some sort of optimization error), and such that $\mathbb{P}(u_n \geq 0) = 1$, the estimator is defined by

$$c_n(z_n, \theta_n) = \inf_{\theta\in\Theta} c_n(z_n, \theta) + u_n.$$

Extend the Weak Consistency Theorem that we have proven in the lectures, so as to hold for this generalized definition (Hint: would it be helpful if you required some sort of asymptotic negligibility for the optimization error?).

QUESTION 9

For the (generalized) linear model of QUESTION 7, assuming now that $\beta \in \Theta \subseteq \mathbb{R}^p$, where the parameter space is not necessarily equal to $\mathbb{R}^p$, we have that the OLSE is defined by $\beta_n \in \beta \in \Theta \frac{1}{n}(y - X\beta)^\top (y - X\beta)$. We have also defined that for V a strictly positive definite $p \times p$ matrix, and in a potentially weaker structure, the IVE that satisfies $\beta_n^\star(V) \in \arg\min_{\beta \in \Theta} (\frac{1}{n} X^\top (y - X\beta))^\top V (\frac{1}{n} X^\top (y - X\beta))$. We have essentially proven that when $\Theta = \mathbb{R}^p$, $\mathbb{P}(\beta_n = \beta_n^\star(V)) = 1$. Find out whether this holds in the more general case where Θ is a non-empty convex subset of $\mathbb{R}^p$. (Hint: would it be helpful if the optimization of a convex function over a convex strict subset of $\mathbb{R}^p$ can be performed via unrestricted optimization of the same function and some appropriate projection?)

QUESTION 10

Suppose that $(X_n)_{n \in \mathbb{N}}$ is a sequence of random variables such that $X_n \xrightarrow{d} X \sim N(0, 1)$. Please find (even only heuristically) the asymptotic distribution of $\theta_n = \arg\min_{\theta \in [0, +\infty)} (X_n - \theta)^2$(Hint: you can among others use the hint of the previous question).