Economics 662D1

October 14, 2021 **R. Davidson**

Midterm Examination

Your completed exam should be uploaded to myCourses by 12.00 on October 15. Please submit two files per student: one, which should be a PDF file, with your written answers, and another, which may or may not be a simple text file, with your computer code. These files must be all your own work. You may make use of whatever non-human resources you wish, but you must not ask for or receive any help from any other person.

All students in this course have the right to submit in English or in French any written work that is to be graded.

Tou(te)s les étudiant(e)s qui suivent ce cours on le droit de soumettre tout travail écrit en français ou en anglais.

Academic Integrity statement [approved by Senate on 29 January 2003]:

McGill University values academic integrity. Therefore all students must understand the meaning and consequences of cheating, plagiarism and other academic offences under the Code of Student Conduct and Disciplinary Procedures.

L'université McGill attache une haute importance à l'honnêteté académique. Il incombe par conséquent à tou(te)s les étudiant(e)s de comprendre ce que l'on entend par tricherie, plagiat et autres infractions académiques, ainsi que les conséquences que peuvent avoir de telles actions, selon le Code de conduite de l'étudiant(e) et des procédures disciplinaires.

Answer all five questions in this exam.

Faites tous les cinq exercices de cet examen.

1. The class of linear unbiased estimators considered by the Gauss-Markov Theorem can be written as $\hat{\boldsymbol{\beta}} = \boldsymbol{A}\boldsymbol{y}$, with $\boldsymbol{A}\boldsymbol{X} = \mathbf{I}$. Show that this class of estimators is in fact identical to the class of estimators of the form

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{W}^\top \boldsymbol{X})^{-1} \boldsymbol{W}^\top \boldsymbol{y}, \tag{1}$$

where $\boldsymbol{W}$ is a matrix of exogenous variables such that $\boldsymbol{W}^\top \boldsymbol{X}$ is nonsingular.

The vector of fitted values given by $\hat{\boldsymbol{\beta}}$ is $\boldsymbol{X}\hat{\boldsymbol{\beta}}$. Express this vector as the product of a matrix and $\boldsymbol{y}$. Show that the matrix is a projection matrix, that is idempotent, but not symmetric, so that it is an oblique projection rather than an orthogonal projection. Characterise the subspace that is the image of the projection, and also the subspace all the vectors in which are annihilated by this oblique projection.

2. In each of the following regressions, y_t is the dependent variable, x_t and z_t are explanatory variables, and α, β, and γ are unknown parameters.

(a) $y_t = \alpha + \beta x_t + \gamma/x_t + u_t$

(b) $y_t = \alpha + \beta x_t + x_t/\gamma + u_t$

(c) $y_t = \alpha + \beta x_t + z_t/\gamma + u_t$

(d) $y_t = \alpha + \beta x_t + z_t/\beta + u_t$

(e) $y_t = \alpha + \beta x_t z_t + u_t$

(f) $y_t = \alpha + \beta\gamma x_t z_t + \gamma z_t + u_t$

(g) $y_t = \alpha + \beta\gamma x_t + \gamma z_t + u_t$

(h) $y_t = \alpha + \beta x_t + \beta x_t^2 + u_t$

(i) $y_t = \alpha + \beta x_t + \gamma x_t^2 + u_t$

(j) $y_t = \alpha + \beta\gamma x_t^3 + u_t$

(k) $y_t = \alpha + \beta x_t + (1 - \beta) z_t + u_t$

(l) $y_t = \alpha + \beta x_t + (\gamma - \beta) z_t + u_t$

For each of these regressions, is it possible to obtain a least-squares estimator of the parameters? In other words, are the parameters of each of these models identified? If not, explain why not. If so, can the estimator be obtained by ordinary (that is, linear) least squares? If it can, write down the regressand and regressors for the linear regression to be used.

3. Perform a set of simulation experiments in which vectors $\boldsymbol{y}$ are generated by the DGP

$$\boldsymbol{y} = \boldsymbol{u}, \quad \boldsymbol{u} \sim \mathrm{N}(0, 1),$$

so that the elements of $\boldsymbol{y}$ are independent standard normal. For each sample size $n = 10, 20, 50, 100, 200, 500, 1000, 2000, 5000$, generate a vector $\boldsymbol{y}$ and construct a graph of its empirical distribution function (EDF). (Recall the definition of the EDF of the elements of a vector: The EDF of the elements of an n-vector is a discrete distribution with n possible points. These points are the n elements of the vector, $y_1, y_2, \ldots, y_n$. Each point is assigned the same probability, which is just $1/n$.)

For each graph, in addition to the graph of the EDF, plot the cumulative distribution function (CDF) Φ of the standard normal distribution. If you think it suitable, you can group the graphs of more than one sample size in one plot, along with the graph of Φ. The abscissa for the graphs should run from -3 to +3, and the graphs themselves constructed by evaluating the EDFs and the standard normal CDF at around 100 intermediate points.

4. The file located at the URL

https://russell-davidson.arts.mcgill.ca/e662/e662.midterm.21.dat

contains 20 observations on four variables. Construct a 20×5 matrix $\boldsymbol{X}$ with a vector of ones, $\boldsymbol{\iota}$, and these four variables as columns. Then evaluate the matrix $\boldsymbol{X}^\top\boldsymbol{X}$, which should be symmetric and positive definite.

The main purpose of this exercise is to compute a matrix $\boldsymbol{A}$ such that $\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{X}^\top\boldsymbol{X}$. The easiest way to do so is to use Crout's algorithm to compute an upper- or lower-triangular matrix $\boldsymbol{A}$. Wikipedia and/or Youtube will give you the information you need to implement the algorithm. Once you have computed $\boldsymbol{A}$, you should then check that the requirement that $\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{X}^\top\boldsymbol{X}$ is satisfied.

You may be brave enough to do the necessary calculations by hand, or with the help of a calculator. In that case, please give enough details of your calculations for us to know just what you did. If you use a computer, then include the code in the code file that you will upload.

5. The file found at

https://russell-davidson.arts.mcgill.ca/data/consumption.data

contains data in 5 columns. What is in each column is explained at the end of the file. Let $\boldsymbol{C}$ denote the variable called Personal Consumption Expenditure, and $\boldsymbol{Y}$ denote Personal Disposable Income. Let $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, $\boldsymbol{s}_3$, and $\boldsymbol{s}_4$ be the usual four seasonal indicator variables, and, as in the textbook, let $\boldsymbol{s}_1' = \boldsymbol{s}_1 - \boldsymbol{s}_4$ $\boldsymbol{s}_2' = \boldsymbol{s}_2 - \boldsymbol{s}_4$, and $\boldsymbol{s}_3' = \boldsymbol{s}_3 - \boldsymbol{s}_4$. As usual, $\boldsymbol{\iota}$ is a vector each element of which is equal to 1, and, in addition, $\boldsymbol{T}$ is a time trend.

You are asked to run (by OLS) and report the results of six regressions, as follows:

$$\begin{aligned}
&\boldsymbol{C} = \beta_1\boldsymbol{\iota} + \beta_2\boldsymbol{Y} + \boldsymbol{u},\\
&\boldsymbol{C} = \gamma_1\boldsymbol{s}_1 + \gamma_2\boldsymbol{s}_2 + \gamma_3\boldsymbol{s}_4 + \beta_2\boldsymbol{Y} + \boldsymbol{u},\\
&\boldsymbol{C} = \beta_1\boldsymbol{\iota} + \gamma_1\boldsymbol{s}_1 + \gamma_2\boldsymbol{s}_2 + \gamma_3\boldsymbol{s}_3 + \beta_2\boldsymbol{Y} + \boldsymbol{u},\\
&\boldsymbol{C} = \beta_1\boldsymbol{\iota} + \alpha_1\boldsymbol{s}_1' + \alpha_2\boldsymbol{s}_2' + \alpha_3\boldsymbol{s}_3' + \beta_2\boldsymbol{Y} + \boldsymbol{u},\\
&\boldsymbol{C} = \beta_1\boldsymbol{\iota} + \theta\boldsymbol{T} + \beta_2\boldsymbol{Y} + \boldsymbol{u},\\
&\boldsymbol{C} = \beta_1\boldsymbol{\iota} + \theta\boldsymbol{T} + \gamma_1\boldsymbol{s}_1 + \gamma_2\boldsymbol{s}_2 + \gamma_2\boldsymbol{s}_3 + \beta_2\boldsymbol{Y} + \boldsymbol{u}
\end{aligned}$$

A coefficient denoted with a symbol may well have different estimates in different regressions. Note also, that the notation $+\boldsymbol{u}$ is purely conventional: it simply means that, when the regressions are run, there will be nonzero residuals.

Determine which of these regressions yield the same fitted values and residuals, and explain why. Also explain why some regressions yield different fitted values and residuals. In addition, make any comments that occur to you about the results of the regressions.