Economics 468

October 19, 2021 **R. Davidson**

Midterm Examination

Your completed exam should be uploaded to myCourses by 15.00 on October 20. Please submit two files per student: one, which should be a PDF file, with your written answers, and another, which can 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. You will find the data to be used for this question at the URL:
https://russell-davidson.arts.mcgill.ca/data/e468.midterm.20.dat

There are 100 observations, representing 25 years of quarterly data. The first column contains the observations for the variable Y, for disposable income. The second column is for C, consumption expenditure.

Regress C on a constant and Y, and give the interpretation of the estimated coefficients under the assumption that the model is correctly specified.

Generate four seasonal indicator variables $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, $\boldsymbol{s}_3$, and $\boldsymbol{s}_4$. Now regress C on these four seasonal variables and Y. Give the interpretation of the results of this new regression. Explain why there is no constant in the set of regressors.

Next, regress C on the seasonal variables only, and save the residuals. Do the same for Y. Regress the residuals from C on the residuals from Y, with no other regressor at all. Explain what all this means, and show that the parameter estimate from this last regression is equal to one of the estimates from an earlier regression.

Now generate three seasonal dummies, as follows:

$$\boldsymbol{s}'_1 = \boldsymbol{s}_1 - \boldsymbol{s}_4, \quad \boldsymbol{s}'_2 = \boldsymbol{s}_2 - \boldsymbol{s}_4, \quad \boldsymbol{s}'_3 = \boldsymbol{s}_3 - \boldsymbol{s}_4.$$

Regress C on a constant, these three new variables, and Y. Interpret the estimated coefficient of the constant, and show that its numerical value is compatible with your earlier results.

3. Consider the autoregressive model

$$y_t = \beta_1 + \beta_2 y_{t-1} + u_t, \qquad t = 2, \ldots, n.$$

Write down the OLS estimating equations for the parameters β_1 and β_2, and show that they are unbiased under the assumption that the disturbances u_t, $t = 1, \ldots, n$ are innovations.

Consider a DGP for which $\beta_1 = 2$, $\beta_2 = 0.5$, and $u_t \sim \text{NID}(0, 1)$, and $y_0 = 1.5$.. Carry out a simulation experiment in order to estimate the bias of the OLS estimator of β_2, and that of the OLS estimator of β_1 for this DGP, with sample sizes $n = 20, 50, 100, 200, 500$. Do your simulation results suggest that either of the OLS estimators is (i) unbiased, or (ii) consistent? If the estimators appear biased, use your simulation results to estimate the bias. How do these estimates vary with the sample size?

If you find that the estimators are biased, how can you reconcile this with the fact, shown in the first part of this question, that the estimating equations are unbiased?

4. In section 4.9 of the textbook, three versions of the goodness-of-fit measure R^2 are defined: uncentred, centred, and adjusted. It is pointed out that the last of these may in some circumstances be negative.

Conduct an experiment in which you investigate the way in which these three measures change when more and more redundant regressors are used in OLS regression.

Specifically: generate a regressand $\boldsymbol{y}$ consisting of 20 IID realisations from the standard normal N(0,1) distribution. Then, for $k = 1, \ldots, 10$, regress $\boldsymbol{y}$ on a constant $\boldsymbol{\iota}$, and a set of regressors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$, where these are all generated, exactly as $\boldsymbol{y}$ is generated, as NID(0,1).

Report the values of the three measures for each value of $k = 1, \ldots, 10$. Explain what is going on.

5. Partition the $n \times k$ regressor matrix $\boldsymbol{X}$ as $\boldsymbol{X} = [\boldsymbol{X}_1 \quad \boldsymbol{X}_2]$, where $\boldsymbol{X}_1$ has k_1 columns, $\boldsymbol{X}_2$ has k_2 columns, with $k_1 + k_2 = k$. Let $\boldsymbol{P_X}$ be the $n \times n$ orthogonal projection matrix on to the span $\mathcal{S}(\boldsymbol{X})$ of the k columns of $\boldsymbol{X}$, and let $\boldsymbol{P}_1$ the orthogonal projection on to $\mathcal{S}(\boldsymbol{X}_1)$, the span of the k_1 columns of $\boldsymbol{X}_1$. Show that $\boldsymbol{P_X}\boldsymbol{P}_1 = \boldsymbol{P}_1\boldsymbol{P_X} = \boldsymbol{P}_1$.

Use that result to show that $\boldsymbol{P_X} - \boldsymbol{P}_1$ is an orthogonal projection matrix.

Show further that $\boldsymbol{P_X} - \boldsymbol{P}_1 = \boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}$, the orthogonal projection on to the span $\mathcal{S}(\boldsymbol{M}_1\boldsymbol{X}_2)$ of the columns of $\boldsymbol{M}_1\boldsymbol{X}_2$. How many of these columns are there?