Economics 240A, Section 3: Short and Long Regression (Ch. 17) and the Multivariate Normal Distribution (Ch. 18)

Michael R. Roberts
Department of Economics and
Department of Statistics
University of California at Berkeley

March 26, 2000

1 Introduction

This handout reviews some of the key points regarding regression algebra and the multivariate normal distribution. It follows closely Goldberger Ch.'s 17 and Ch. 18.

2 Short and Long Regressions

The basic set-up is

$$y = X\beta + \varepsilon = X_1\beta_1 + X_2\beta_2 + \varepsilon \tag{1}$$

where we have partitioned the $n \times k$ matrix X into two submatrices $X_1 \in R^{n\times k_1}$ and $X_2 \in R^{n\times k_2}$.

We can think of two regressions:

1. a short one

$$y = X_1\beta_1 + \varepsilon \tag{2}$$

   and,

2. a long one

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon \tag{3}$$

I'll use the same notation as Goldberger so let b_i be a vector of OLS parameter estimates for the subvector β_i in a long regression and b_i^* be the OLS estimates of β_i in a short regression. And, let e be the residuals from the long regression and e^* be the residuals from the short regression.

Exercise 1:

Let b_1^* be the OLS estimates of β_1 from regression 2 and b_1 be the OLS estimates of β_1 from regression 3. Show

$$b_1^* = b_1 + \left(X_1'X_1\right)^{-1} X_1'X_2b_2 \tag{4}$$

Exercise 2:

Letting e^* be the residuals from 2 show that

$$e^* = M_1X_2b_2 + e$$

In words, what is M_1X_2?

Exercise 3:

Show that

$$e^{*\prime}e^* = b_2'X_2'M_1X_2b_2 + e'e$$

and interpret this result. What implication does this have for the fit of the long regression relative to the short regression?

Result 1 *Some exceptions*

1. *If* $b_2 = 0$ *then* $b_1^* = b_1$ *and* $e^* = e$.
2. *If* $X_1'X_2 = 0$ *then* $b_1^* = b_1$ *but* $e^* \neq e$.

3 Frisch-Waugh-Lovell

Problem 2 proves the Frisch-Waugh-Lovell theorem which can be thought of as an alternative way of getting at the OLS estimator of β_2.

1. Regress each column of X_2 on X_1 and save the corresponding set of residuals in a matrix, X_2^*.
2. Regress y on X_1 and save its residual as y^*.(In fact, this step is unnecessary and Goldberger refers to this as a double residual regression. *Exercise: 4 Prove that this step is in fact unnecessary*)
3. Regress y^* on X_2^* and the resulting coefficient vector is the same as the OLS coefficients from the original regression in 1.

To see this consider regressing y on $M_1 X_2$ $(= X_2^*$ in Goldberger$)$. The coefficient vector is

$$\begin{aligned} c_2^* &= \left(X_2' M_1 X_2\right)^{-1} X_2' M_1 y \\ &= \left(X_2' M_1 X_2\right)^{-1} X_2' M_1 \left(X_1 b_1 + X_2 b_2 + e\right) \\ &= \left(X_2' M_1 X_2\right)^{-1} X_2' M_1 \left(X_2 b_2 + e\right) \qquad (M_1 X_1 = 0) \\ &= b_2 \qquad \text{(cancelling and noting } M_1 e = e \text{ and } X_2' e = 0) \end{aligned}$$

For some applications see section 17.4.

4 The CR Model

Recall the set-up

$$\begin{aligned} E\left(y\right) &= X\beta = X_1 \beta_1 + X_2 \beta_2 \\ V\left(y\right) &= \sigma^2 I \\ X &: \text{ full rank and nonstochastic} \end{aligned}$$

4.1 The Parameters

Exercise 5: (Omitted Variables Bias):

Show that the estimated coefficients from the short regression (2) b_1^* are biased.

Exercise 6:

What is the variance of the short regression coefficents b_1^* and what is its relation relative to the variance of the long regression coefficients b_1?

4.2 The Residuals

Exercise 7:

Find the expectation and variance of the short regression residual vector e^*.

Exercise 8:

Find the expectation of the sum of squared residuals, $e^{*\prime} e^*$.

5 The Normal Distribution

You better become REAL familiar with this. There are just a zillion different properties that the normal (univariate and multivariate) distribution has. Here's a short list of some things that might be worth knowing:

5.1 Univariate Normal Distribution

1. $X\tilde{}N\left(\mu,\sigma^2\right)$ means X has a univariate normal distribution with mean parameter μ and σ^2. The density is of course
$$f_X\left(x\right)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2}\frac{\left(x-\mu\right)^2}{\sigma^2}\right\}$$
which is often denoted by $\phi\left(x\right)$ and there is no closed form for the corresponding distribution, $\Phi\left(x\right)$

2. The distribution is symmetric implying
$$\Phi\left(-x\right)=1-\Phi\left(x\right)$$
This is easily seen by thinking of the area under the normal density.

3. Closed under affine transformations. If $x\tilde{}N\left(\mu,\sigma^2\right)$ then $y=\alpha+\beta x$ is distributed $N\left(\alpha+\beta\mu,\beta^2\sigma^2\right)$.

4. Is uniquely determined by it's first two moments.

5. $$\phi'\left(x\right)=x\phi\left(x\right)$$

6. If Z is standard normal than all odd moments are equal to 0 and
$$E\left(Z^{2k}\right)=\frac{(2k)!}{2^k\cdot k!},\qquad k=1,2,3,...$$
(This can be shown using integration by parts and induction)

5.2 Multivariate Normal Distribution

1. The vector $\mathbf{x}\in\mathbf{R}^n$ is distributed multivariate normal with mean vector $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$ and has the corresponding density
$$f_{\mathbf{X}}\left(\mathbf{x}\right)=\left(2\pi\right)^{-n/2}\times\left|\boldsymbol{\Sigma}\right|^{-1/2}\times\exp\left\{-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}\right)'\boldsymbol{\Sigma}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}\right)\right\}$$
where $|\cdot|$ means determinant.

2. Two normal random variables are independent if and only if they are uncorrelated.

3. Affine transformations of a vector of normal random variables are again normal. So, if $\mathbf{x}\tilde{\ }MVN\left(\boldsymbol{\mu},\boldsymbol{\Sigma}\right)$ then $\mathbf{y}=\mathbf{Hx}+\mathbf{b}$ is distributed multivariate normal with mean $\mathbf{H}\boldsymbol{\mu}+\mathbf{b}$ and variance $\mathbf{H}\boldsymbol{\Sigma}\mathbf{H}'$

4. *Important!!!* Consider a pair of random vectors $\mathbf{x}$ and $\mathbf{y}$ each multivariate normal such that $(\mathbf{x}',\mathbf{y}')$ has mean and covariance matrix given by

$$\boldsymbol{\mu}=\begin{pmatrix}\boldsymbol{\mu}_x\\ \boldsymbol{\mu}_y\end{pmatrix}\text{ and }\boldsymbol{\Sigma}=\begin{pmatrix}\boldsymbol{\Sigma}_{xx} & \boldsymbol{\Sigma}_{xy}\\ \boldsymbol{\Sigma}_{yx} & \boldsymbol{\Sigma}_{yy}\end{pmatrix}$$

respectively. Then the distribution of $\mathbf{x}$ conditional on $\mathbf{y}$ is also multivariate normal with mean

$$\boldsymbol{\mu}_{x|y}=\boldsymbol{\mu}_x+\boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\left(\mathbf{y}-\boldsymbol{\mu}_y\right)$$

and covariance matrix

$$\boldsymbol{\Sigma}_{x|y}=\boldsymbol{\Sigma}_{xx}-\boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx}$$

Note that the conditional covariance matrix does not depend on $\mathbf{y}$ and that while $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_{yy}$ are assumed to be nonsingular, $\boldsymbol{\Sigma}_{yy}^{-1}$ can be replaced by a pseudo inverse.

5.3 Functions of Normal Random Variables

1. Let $\mathbf{x}$ be a $k-$dimensional vector of standard normal random variables. Then $\mathbf{x}'\mathbf{x}$ is distributed χ^2 with k degrees of freedom.

2. Extending the above result, if $\mathbf{x}\in\mathbf{R}^n$ is distributed $MVN\left(\boldsymbol{\mu},\boldsymbol{\Sigma}\right)$ then

$$\left(\mathbf{x}-\boldsymbol{\mu}\right)'\boldsymbol{\Sigma}\left(\mathbf{x}-\boldsymbol{\mu}\right)$$

is distributed χ_n^2

3. If $\mathbf{x}\in\mathbf{R}^n$ is distributed $MVN\left(0,\mathbf{I}\right)$ and M is any nonrandom idempotent matrix with rank $r\leq n$ then $u'Mu$ is distributed χ_r^2.

4. Let $\mathbf{x}\in\mathbf{R}^n$ be distributed $MVN\left(0,\mathbf{I}\right)$. Let M be any nonrandom idempotent matrix with rank $r\leq n$ and let L be a nonrandom matrix such that $LM=0$. Then $a=Mu$ and $b=Lu$ are independent random vectors.

5. Let $v\tilde{\ }\chi_n^2$ and $w\tilde{\ }\chi_d^2$ be two independent chi-square random variables. Then

$$z=\frac{v/n}{w/d}$$

is distributed Snedecor-$F:F\left(n,m\right)$

6. Let $z\tilde{}N(0,1)$ and $w\tilde{}\chi^2_n$ independent of z. Then

$$t = \frac{z}{w/n}$$

has a Student's t-distribution with n degrees of freedom (t_n)

7. If $u\tilde{}t_n$ then $u^2\tilde{}F(1,n)$.