Complements on classical linear model *

Jean-Marie Dufour †
McGill University

First version: April 1995
Revised: February 2002, July 2011
This version: July 2011
Compiled: November 22, 2011, 19:08

*This work was supported by the William Dow Chair in Political Economy (McGill University), the Bank of Canada (Research Fellowship), a Guggenheim Fellowship, a Konrad-Adenauer Fellowship (Alexander-von-Humboldt Foundation, Germany), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds de recherche sur la société et la culture (Québec).

† William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 519, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 8879; FAX: (1) 514 398 4938; e-mail: jean-marie.dufour@mcgill.ca . Web page: http://www.jeanmariedufour.com

Contents

1. Formulas for partitioned regression

$$
\begin{aligned}
y &= X\beta + \varepsilon \\
&= (X_1,\, X_2)\begin{pmatrix}\beta_1 \\ \beta_2\end{pmatrix} + \varepsilon \\
&= X_1\beta_1 + X_2\beta_2 + \varepsilon
\end{aligned}
\tag{1.1}
$$

$$
\hat{\beta} = (X'X)^{-1}X'y = \begin{pmatrix}\hat{\beta}_1 \\ \hat{\beta}_2\end{pmatrix}
\tag{1.2}
$$

where

$$
\begin{aligned}
X &: T\times k\ ,\ X_1 : T\times k_1\ ,\ X_2 : T\times k_2\ , \\
\beta_1 &: k_1\times 1\ ,\ \beta_2 : k_2\times 1\ ,\ k = k_1 + k_2\,.
\end{aligned}
$$

$$
\begin{aligned}
\hat{\beta}_1 &= (X_1'X_1)^{-1}X_1'y - (X_1'X_1)^{-1}X_1'X_2D^{-1}X_2'M_1y \\
&= b_1 - (X_1'X_1)^{-1}X_1'X_2D^{-1}X_2'M_1y
\end{aligned}
\tag{1.3}
$$

where

$$
b_1 = (X_1'X_1)^{-1}X_1'y\,, \tag{1.4}
$$

$$
M_1 = I_T - X_1(X_1'X_1)^{-1}X_1'\,, \tag{1.5}
$$

$$
D = X_2'M_1X_2\,; \tag{1.6}
$$

$$
\hat{\beta}_2 = D^{-1}X_2'M_1y = (X_2'M_1X_2)^{-1}X_2'M_1y\,; \tag{1.7}
$$

$$
\hat{\beta}_1 = (X_1'M_2X_1)^{-1}X_1'M_2y \tag{1.8}
$$

where

$$
M_2 = I_T - X_2(X_2'X_2)^{-1}X_2'\,. \tag{1.9}
$$

For further discussion, the reader may con consult Schmidt (1976) and Seber

(1977).

2. Updating formulas for linear regressions

$$y_t = x_t'\beta + \varepsilon_t \quad , \quad t = 1, ..., T \tag{2.1}$$

where

$$x_t : k \times 1\,, \tag{2.2}$$

$$Y_r = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{pmatrix} \quad , \quad X_r = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_r' \end{bmatrix} \quad , \quad r = k, k+1, ..., T\,. \tag{2.3}$$

$$b_r = (X_r'X_r)^{-1}X_r'Y_r \tag{2.4}$$

is the estimator of β based on the first r observations. Then the following updating formulas hold [see Brown, Durbin and Evans (1975)] :

$$b_r = b_{r-1} + (X_r'X_r)^{-1} x_r \left(y_r - x_r'b_{r-1}\right)\,,\; k+1 \le r \le T\,, \tag{2.5}$$

$$(X_r'X_r)^{-1} = (X_{r-1}'X_{r-1})^{-1} - \frac{\left(X_{r-1}'X_{r-1}\right)^{-1} x_r x_r' \left(X_{r-1}'X_{r-1}\right)^{-1}}{1 + x_r'\left(X_{r-1}'X_{r-1}\right)^{-1} x_r}\,. \tag{2.6}$$

Further,

$$\begin{aligned} V(b_r) - V(b_{r-1}) &= \sigma^2(X_r'X_r)^{-1} - \sigma^2(X_{r-1}'X_{r-1})^{-1} \\ &= -\sigma^2 \frac{\left(X_{r-1}'X_{r-1}\right)^{-1} x_r x_r' \left(X_{r-1}'X_{r-1}\right)^{-1}}{1 + x_r'\left(X_{r-1}'X_{r-1}\right)^{-1} x_r} \end{aligned} \tag{2.7}$$

is a negative semidefinite matrix.

3. Orthogonal decompositions of least squares estimators

Consider $\hat{\beta}$ and $\hat{\beta}_0$, respectively the unrestricted estimator of β and the restricted estimator of β under the constraint $R\beta = r$:

$$\hat{\beta} = (X'X)^{-1} X'y\,, \tag{3.1}$$

$$\hat{\beta}_0 = \hat{\beta} + Q_R\big[r - R\hat{\beta}\big] \tag{3.2}$$

where

$$Q_R = (X'X)^{-1} R'\big[R\,(X'X)^{-1} R'\big]^{-1}. \tag{3.3}$$

Then, we see easily that

$$\begin{aligned} R\hat{\beta} - r &= R\big[\beta + (X'X)^{-1} X'\varepsilon\big] - r && (3.4)\\ &= (R\beta - r) + R_X\varepsilon && (3.5) \end{aligned}$$

where

$$R_X = R\,(X'X)^{-1} X'\,, \tag{3.6}$$

$$\begin{aligned} \hat{\beta} - \hat{\beta}_0 &= Q_R\big[R\hat{\beta} - r\big] \\ &= Q_R\big[\,(R\beta - r) + R_X\varepsilon\big] \\ &= Q_R\,(R\beta - r) + Q_R R_X\varepsilon \\ &= Q_R\,(R\beta - r) + Q\varepsilon && (3.7) \end{aligned}$$

and

$$\begin{aligned} \hat{\beta}_0 &= \hat{\beta} + (\hat{\beta}_0 - \hat{\beta}) \\ &= \beta + (X'X)^{-1} X'\varepsilon - Q_R\,(R\beta - r) - Q\varepsilon \\ &= \beta + Q_R\,(r - R\beta) + [(X'X)^{-1} X' - Q]\varepsilon && (3.8) \end{aligned}$$

where

$$Q = Q_R R_X = Q_R R\,(X'X)^{-1} X'\,. \tag{3.9}$$

Since

$$R_X X\left(X'X\right)^{-1} = R\left(X'X\right)^{-1} X'X\left(X'X\right)^{-1} = R\left(X'X\right)^{-1}, \tag{3.10}$$

$$R_X R_X' = R\left(X'X\right)^{-1} X'X\left(X'X\right)^{-1} R' = R\left(X'X\right)^{-1} R' \tag{3.11}$$

and

$$\begin{aligned} R_X Q' &= R_X R_X' Q_R' \\ &= R\left(X'X\right)^{-1} R'\left[R\left(X'X\right)^{-1} R'\right]^{-1} R\left(X'X\right)^{-1} \qquad (3.12) \\ &= R\left(X'X\right)^{-1}, \qquad (3.13) \end{aligned}$$

it follows that

$$\begin{aligned} \mathsf{C}(R\hat{\beta} - r,\, \hat{\beta}_0) &= \mathsf{C}\left(R_X\varepsilon,\, [(X'X)^{-1}X' - Q]\varepsilon\right) \\ &= \mathsf{E}\left[R_X\varepsilon\varepsilon'[(X'X)^{-1}X' - Q]'\right] \\ &= \sigma^2 R_X[(X'X)^{-1}X' - Q]' \\ &= \sigma^2 R_X[X\left(X'X\right)^{-1} - Q'] \\ &= \sigma^2[R\left(X'X\right)^{-1} - R\left(X'X\right)^{-1}] = 0\,. \qquad (3.14) \end{aligned}$$

and

$$\begin{aligned} \mathsf{C}\left(\hat{\beta} - \hat{\beta}_0,\, \hat{\beta}_0\right) &= \mathsf{C}\left(Q_R\left[R\hat{\beta} - r\right], \hat{\beta}_0\right) \\ &= Q_R\mathsf{C}\left(R\hat{\beta} - r, \hat{\beta}_0\right) = 0\,. \end{aligned}$$

Thus $\hat{\beta}_0$ and $R\hat{\beta} - r$ are uncorrelated under the assumptions of the classical linear model, and similarly for $\hat{\beta}_0$ and $\hat{\beta} - \hat{\beta}_0$. This holds even if the normality assumption or the restriction $R\beta = r$ do not hold. Consequently, the identity

$$\hat{\beta} = \hat{\beta}_0 + \left(\hat{\beta} - \hat{\beta}_0\right) \tag{3.15}$$

provides a decomposition of $\hat{\beta}$ as the sum of two uncorrelated random vectors, so that

$$\mathsf{V}\left(\hat{\beta}\right) = \mathsf{V}\left(\hat{\beta}_0\right) + \mathsf{V}\left(\hat{\beta} - \hat{\beta}_0\right). \tag{3.16}$$

More explicitly, we have

$$\hat{\beta} = \hat{\beta}_0 + Q_R(r - R\beta) - Q\varepsilon \tag{3.17}$$

where

$$\mathsf{C}\left[\hat{\beta}_0, Qy\right] = \mathsf{C}\left[\hat{\beta}_0, Q\varepsilon\right] = 0\,. \tag{3.18}$$

An interesting special case of the latter results is the one where

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon \tag{3.19}$$

and the restrictions take the form

$$\beta_2 = 0\,, \tag{3.20}$$

with

$$R = [0, I_{k_2}]\,,\ r = 0\,. \tag{3.21}$$

Then

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}, \quad \hat{\beta}_0 = \begin{pmatrix} \hat{\beta}_{10} \\ \hat{\beta}_{20} \end{pmatrix} = \begin{pmatrix} (X_1'X_1)^{-1}X_1'y \\ 0 \end{pmatrix} \tag{3.22}$$

and

$$\hat{\beta}_1 = \hat{\beta}_{10} - Q_{20}R\hat{\beta} = \hat{\beta}_{10} - Q_{20}\hat{\beta}_2 \tag{3.23}$$

where

$$\hat{\beta}_2 = (X_2'M_1X_2)^{-1}X_2'M_1y \tag{3.24}$$

and $\hat{\beta}_2$ is independent of $\hat{\beta}_{10}$.[1]

[1] See Magnus and Durbin (1999) and Danilov and Magnus (2001) for further discussion.

References

Brown, R. L., Durbin, J. and Evans, J. M. (1975), 'Techniques for testing the constancy of regression relationships over time', *Journal of the Royal Statistical Society, Series B* **37**, 149–192.

Danilov, D. and Magnus, J. R. (2001), On the harm that pretesting does, Technical report, CentER, Tilburg University.

Maddala, G. S. (1977), *Econometrics*, McGraw-Hill, New York.

Magnus, J. R. and Durbin, J. (1999), 'Estimation of regression coefficients of interest when other regression coefficients are of no interest', *Econometrica* **67**, 639–643.

Schmidt, P. (1976), *Econometrics*, Marcel Dekker, New York.

Seber, G. A. F. (1977), *Linear Regression Analysis*, John Wiley & Sons, New York.

Theil, H. (1971), *Principles of Econometrics*, John Wiley & Sons, New York.