

Matrix Calculus in Econometrics - Part II: Lecture VII

Charles B. Moss ¹

¹University of Florida

January 25, 2012

- 1 Readings
- 2 Simple Regression
 - Matrix Calculus
 - Matrix Notation
- 3 Matrix Derivatives
- 4 Readings - Next Time

Readings

- * Dhrymes, P.J. 2000. *Mathematics for Econometrics* Third Edition. Springer-Verlag. (Chapter 5).
- * Greene, W.H. 2012. *Econometric Analysis* Seventh Edition. Prentice Hall. (Appendix A: Matrix Algebra pp. 973-1014).
- Magnus, J.R. and H. Neudecker. 1988. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons. (Chapters 8 and 9).
- * Popper, K. 2010. *The Logic of Scientific Discovery* Routledge Classics. (Chapter 7: Simplicity, 121-132)
- * Schott, J.R. 1997. *Matrix Analysis for Statistics*. John Wiley & Sons. (Chapter 8)

Simple Regression

- Starting with the simple regression

$$y = \begin{bmatrix} 0.00674 \\ 0.01134 \\ 0.01989 \\ 0.01834 \end{bmatrix} = \begin{bmatrix} 1.0 & 0.08377 \\ 1.0 & -0.02591 \\ 1.0 & 0.01241 \\ 1.0 & -0.016457 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = X\beta \quad (1)$$

$$\begin{bmatrix} 0.00674 \\ 0.01134 \\ 0.01989 \\ 0.01834 \end{bmatrix} = \begin{bmatrix} \beta_0 + 0.08377\beta_1 \\ \beta_0 - 0.02591\beta_1 \\ \beta_0 + 0.01241\beta_1 \\ \beta_0 - 0.16457\beta_1 \end{bmatrix}$$

Simple Regression - Continued

- Defining the residual

$$\begin{aligned} \epsilon &= \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{bmatrix} = \begin{bmatrix} 0.00674 \\ 0.01134 \\ 0.01989 \\ 0.01834 \end{bmatrix} - \begin{bmatrix} \beta_0 + 0.08377\beta_1 \\ \beta_0 - 0.02591\beta_1 \\ \beta_0 + 0.01241\beta_1 \\ \beta_0 - 0.16457\beta_1 \end{bmatrix} \\ &= \begin{bmatrix} 0.00674 - \beta_0 - 0.08377\beta_1 \\ 0.01134 - \beta_0 + 0.02591\beta_1 \\ 0.01989 - \beta_0 - 0.01241\beta_1 \\ 0.01834 - \beta_0 + 0.16457\beta_1 \end{bmatrix} \end{aligned} \quad (2)$$

Simple Regression - Continued

- Defining the variance (objective function of ordinary least squares)

$$\begin{aligned} \epsilon' \epsilon &= \begin{bmatrix} 0.00674 - \beta_0 - 0.08377\beta_1 \\ 0.01134 - \beta_0 + 0.02591\beta_1 \\ 0.01989 - \beta_0 - 0.01241\beta_1 \\ 0.01834 - \beta_0 - 0.16457\beta_1 \end{bmatrix}' \begin{bmatrix} 0.00674 - \beta_0 - 0.08377\beta_1 \\ 0.01134 - \beta_0 + 0.02591\beta_1 \\ 0.01989 - \beta_0 - 0.01241\beta_1 \\ 0.01834 - \beta_0 - 0.16457\beta_1 \end{bmatrix} \\ &= \sum_{i=1}^4 (y_i - \beta_0 - \beta_1 X_{i2}) (y_i - \beta_0 - \beta_1 X_{i2}) = \\ &\quad \sum_{i=1}^4 (y_i - X_{i.}\beta) (y_i - X_{i.}\beta) \end{aligned} \tag{3}$$

Simple Regression - Matrix Calculus

- Following the derivation of the first-order conditions, consider the derivative of the least squares regression function in vector space

$$\nabla_{\beta} (\epsilon' \epsilon) = \begin{bmatrix} \frac{\partial \epsilon' \epsilon}{\partial \beta_0} & \frac{\partial \epsilon' \epsilon}{\partial \beta_1} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial \sum_{i=1}^4 (y_i - X_{i.}\beta) (y_i - X_{i.}\beta)}{\partial \beta_0} & \frac{\partial \sum_{i=1}^4 (y_i - X_{i.}\beta) (y_i - X_{i.}\beta)}{\partial \beta_1} \end{bmatrix} \quad (4)$$

Simple Regression - Matrix Calculus - Continued

$$\begin{aligned}\nabla_{\beta} (\epsilon' \epsilon) &= \begin{bmatrix} 2 \sum_{i=1}^4 (y_i - X_i \cdot \beta) (-1) & 2 \sum_{i=1}^4 (y_i - X_i \cdot \beta) (-X_{i2}) \end{bmatrix} \\ &= \begin{bmatrix} -\sum_{i=1}^4 (y_i - X_i \cdot \beta) & -\sum_{i=1}^4 (y_i - X_i \cdot \beta) X_{i2} \end{bmatrix} + \\ &\quad \begin{bmatrix} -\sum_{i=1}^4 (y_i - X_i \cdot \beta) & -\sum_{i=1}^4 (y_i - X_i \cdot \beta) X_{i2} \end{bmatrix} \end{aligned} \tag{5}$$

Simple Regression - Matrix Calculus - Continued

- Looking at each half

$$\begin{bmatrix} -\sum_{i=1}^4 (y_i - X_{i.}\beta) & -\sum_{i=1}^4 (y_i - X_{i.}\beta) X_{i2} \end{bmatrix} = \begin{bmatrix} 0.00674 - \beta_0 - 0.08377\beta_1 \\ 0.01134 - \beta_0 + 0.02591\beta_1 \\ 0.01989 - \beta_0 - 0.01241\beta_1 \\ 0.01834 - \beta_0 + 0.16457\beta_1 \end{bmatrix}' \begin{bmatrix} -1.0 & -0.08377 \\ -1.0 & 0.02591 \\ -1.0 & -0.01241 \\ -1.0 & 0.16457 \end{bmatrix} \quad (6)$$

Simple Regression - Matrix Calculus - Continued

$$\begin{bmatrix} 0.00674 - \beta_0 - 0.08377\beta_1 \\ 0.01134 - \beta_0 + 0.02591\beta_1 \\ 0.01989 - \beta_0 - 0.01241\beta_1 \\ 0.01834 - \beta_0 + 0.16457\beta_1 \end{bmatrix}' \begin{bmatrix} -1.0 & -0.08377 \\ -1.0 & 0.02591 \\ -1.0 & -0.01241 \\ -1.0 & 0.16457 \end{bmatrix} \quad (7)$$

$$= [y - X\beta]'(-X)$$

- Therefore the gradient becomes

$$\begin{aligned} \nabla_{\beta} (\epsilon'\epsilon) &= -[y - X\beta]'X - [y - X\beta]'X \\ \Rightarrow -(y'X + \beta'X'X) - (y'X - \beta'X'X) &= 0 \\ \Rightarrow X'y - X'X\beta = 0 \Rightarrow \beta &= (X'X)^{-1} X'y \end{aligned} \quad (8)$$

Matrix Notation

- Returning to the vector formulation $\nabla_{\beta} (\epsilon' \epsilon)$, we use the following results:
 - $\nabla_x (UV) = (\nabla_x U) V + U (\nabla_x V)$
 - $\nabla_x U' = (\nabla_x U)'$
- Substituting for ϵ

$$\nabla_{\beta} [(y - X\beta)' (y - X\beta)] = \nabla_{\beta} [y'y - y'X\beta - \beta'X'y + \beta'X'X\beta]$$

$$\ni: \begin{cases} \nabla_{\beta} (y'X\beta) = y'X \\ \nabla_{\beta} (\beta'X'y) = y'X \\ \nabla_{\beta} (\beta'X'X\beta) = X'X\beta + X'X\beta \end{cases}$$

$$\Rightarrow \nabla_{\beta} [(y - X\beta)' (y - X\beta)] = -2y'X + 2X'X\beta$$

(9)

Matrix Notation - Continued

- A little more detail from Schott p. 330.

$$\begin{aligned}
 g &= x'Ax \Rightarrow dg = d(x'Ax) \\
 dg &= d(x')Ax + x'd(Ax) \\
 dg &= d(x')Ax + x'dx \\
 d(x')Ax &\Rightarrow \{ (dx)'Ax \}' \\
 &\Leftarrow dU' = (dU)' \\
 dg &= x'Adx + x'Adx = 2x'Adx
 \end{aligned} \tag{10}$$

Matrix Derivatives

- $\nabla_X |A| = A^\#$ where $A^\#$ is the adjoint matrix.

$$\begin{aligned}
 |A| &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} (-1)^{1+1} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \\
 &+ a_{21} (-1)^{2+1} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} (-1)^{3+1} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \quad (11) \\
 \frac{\partial |A|}{\partial a_{11}} &= (-1)^{1+1} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}
 \end{aligned}$$

Matrix Derivatives

- $\nabla_X A^{-1} = -A^{-1} (dA) A^{-1}$

$$AA^{-1} = I$$

$$\Rightarrow (dA) A^{-1} + A (dA^{-1}) = 0 \quad (12)$$

$$\Rightarrow A (dA^{-1}) = - (dA) A^{-1}$$

$$dA^{-1} = -A^{-1} (dA) A^{-1}$$

Readings - Next Time

- * Greene, W.H. 2012. *Econometric Analysis* Seventh Edition. (Chapter 3)
- * Hansen, B.E. 2011. *Econometrics* Draft Textbook Website: <http://www.ssc.wisc.edu/bhansen>. (Chapter 4).
- Hayashi, F. 2000. *Econometrics* Oxford University Press (Section 1.2 1.4 pp. 15-46)
- Theil, H. 1971. *Principles of Econometrics* John Wiley & Sons. (Chapter 3)
- * Popper, K. 2010. *The Logic of Scientific Discovery* Routledge Classics [Chapter 8: Probability, 133-208].