

Matrix Calculus in Econometrics - Part I: Lecture VI

Charles B. Moss ¹

¹University of Florida

January 20, 2012

- 1 Readings
- 2 Mathematics in Econometrics
- 3 Mathematics in Econometrics - Calculus
 - Optimization problems in Econometrics
 - Multidimensional Calculus
 - Approximation Functions
- 4 Readings

Readings

- * Khuri, A.I. 1993. *Advanced Calculus with Applications in Statistics*. John Wiley & Sons.
- * Moss, C.B. 2010. *Static and Dynamic Optimization for Agricultural and Resource Economics* Draft Textbook. (Chapter 2 pp. 15-40)
- * Popper, K. 2010. *The Logic of Scientific Discovery* Routledge Classics [Chapter 6: Degrees of Testability, 95-120].

Mathematics in Econometrics - Matrix Methods

- Largely our applications will involve samples on linear equations

$$\begin{aligned}y_1 &= \alpha_0 + \alpha_1 x_{11} + \alpha_2 x_{12} + \epsilon_1 \\y_2 &= \alpha_0 + \alpha_1 x_{21} + \alpha_2 x_{22} + \epsilon_2 \\y_3 &= \alpha_0 + \alpha_1 x_{31} + \alpha_2 x_{32} + \epsilon_3 \\y_4 &= \alpha_0 + \alpha_1 x_{41} + \alpha_2 x_{42} + \epsilon_4 \\y_5 &= \alpha_0 + \alpha_1 x_{51} + \alpha_2 x_{52} + \epsilon_5\end{aligned}\tag{1}$$

Matrix Methods - Continued

- In matrix form

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} - \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ 1 & x_{31} & x_{32} \\ 1 & x_{41} & x_{42} \\ 1 & x_{51} & x_{52} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix} \quad (2)$$

- Implying

$$y - X\alpha = \epsilon \Rightarrow \epsilon(\alpha) = y - X\alpha \quad (3)$$

Mathematics in Econometrics - Calculus

- Optimization - Our estimators are largely (the exception being nonparametric kernel estimators proposed by Härdle) are optimization problems.
- Approximations - Related to asymptotic specifications
- Limits - Many of our statistics are limits or asymptotic tests.

Optimization problems in Econometrics

- Direct optimization problem
 - Minimize Least Squares (Ordinary Least Squares and Generalized Least Squares)
 - Maximum Likelihood
- Indirect optimization problems - Generalized Method of Moments

Optimization in Econometrics - Least Squares

- Starting with the simple linear least squares problem

$$\begin{aligned}
 & \min_{\alpha} \sum_{i=1}^5 \epsilon_i^2(\alpha) \\
 & \Rightarrow \min_{\alpha} \epsilon'(\alpha) \epsilon(\alpha) \\
 & \Rightarrow \min_{\alpha} [y' - \alpha' X'] [y - X\alpha] \\
 & \Rightarrow \min_{\alpha} [y'y - y'X\alpha - \alpha'X'y + \alpha'X'X\alpha] \\
 & \Rightarrow \min_{\alpha} [\alpha'X'X\alpha - 2y'X\alpha] \ni: y'X\alpha = \alpha'X'y
 \end{aligned} \tag{4}$$

Optimization in Econometrics - Maximum Likelihood

- Starting with the Logarithmic Normal Likelihood function for the sample

$$\begin{aligned} \max_{\alpha, \sigma^2} -\frac{5}{2} \ln(\sigma^2) - \sum_{i=1}^5 \frac{(y_i - X_i \cdot \alpha)^2}{2\sigma^2} \\ \Rightarrow \hat{\sigma}_{MLE}^2(\alpha) = \sum_{i=1}^5 \frac{(y_i - X_i \cdot \alpha)^2}{5} \end{aligned} \quad (5)$$

$$\Rightarrow \max_{\alpha} -\frac{5}{2} \ln(\hat{\sigma}_{MLE}^2(\alpha)) - 5 \Leftrightarrow \min_{\alpha} \sum_{i=1}^5 \epsilon_i^2(\alpha)$$

Maximum Likelihood Continued

- For systems of equations

$$\begin{aligned} & \max_{\alpha} -\frac{T}{2} \ln |\Sigma| - \frac{1}{2} \sum_{i=1}^T (y_i - X_i \alpha)' \Sigma^{-1} (y_i - X_i \alpha) \\ \Rightarrow \max_{\alpha} & -\frac{T}{2} \ln \left| \hat{\Sigma}_{MLE}(\alpha) \right| - Tm \ni: \Sigma_{MLE}(\alpha) = \sum_{i=1}^T \epsilon_i(\alpha) \epsilon_i(\alpha)' \end{aligned} \quad (6)$$

Optimization in Econometrics - Generalized Method of Moments

- The Generalized Method of Moments is developed around a general orthogonality principle

$$E[z_t \epsilon_t(\alpha)] = 0. \quad (7)$$

- Where z_t is a vector of instrumental variables (including the independent variables).
- We solve for the values of α using the minimand

$$Q_T(\alpha) = \left\{ \frac{1}{T} \epsilon(\alpha)' Z \right\}' W_T \left\{ \frac{1}{T} Z' \epsilon(\alpha) \right\} \quad (8)$$

Optimization in Econometrics - Applied Optimization

- My optimization book develops the necessary and sufficient conditions for optimality based on a Taylor series expansion (pp. 18-20)
 - Taking the Second-order Taylor expansion of an unknown function yields

$$f(x) - f(x^*) = f'(x^*)(x - x^*) + \frac{1}{2}f''(x^*)(x - x^*)^2 \quad (9)$$

- In this formulation, a maximum occurs at x^* if $f(x) - f(x^*) \leq 0$ for $x \in N(x^*)$. This requires $f'(x^*) = 0$ and $f''(x^*) \leq 0$ (Necessary).
- A minimum occurs at x^* if $f(x) - f(x^*) \geq 0$ for $x \in N(x^*)$. This requires $f'(x^*) = 0$ and $f''(x^*) \geq 0$ (Necessary).

Optimization in Econometrics - Multivariate Optimality

- Extending the univariate function to multivariate space we have

$$f(x) - f(x^*) = \nabla_x f(x^*) dx + \frac{1}{2} dx' \nabla_{xx}^2 f(x^*) dx \leq 0 \quad (10)$$

- The first-order necessary condition is that the norm of the gradient vector vanishes ($\|\nabla_x f(x^*)\| \rightarrow 0$).
- The second-order necessary condition requires that the Hessian matrix be negative semi-definite ($x' \nabla_{xx}^2 f(x^*) x \leq 0 \forall x$).

Optimization in Econometrics - Maximum Likelihood

- Applying these techniques to maximum likelihood

$$L(\alpha) = -\frac{N}{2} - \frac{1}{2} \sum_{i=1}^N (y_i - X_i \alpha)' \Sigma^{-1} (y_i - X_i \alpha)$$

$$\nabla_{\alpha} L(\alpha) = -\frac{1}{2} \sum_{i=1}^N [-X_i' \Sigma^{-1} (y_i - X_i \alpha) - (y_i - X_i \alpha)' \Sigma^{-1} X_i]$$

$$\nabla_{\alpha} L(\alpha) = \sum_{i=1}^N X_i' \Sigma^{-1} \epsilon_i(\alpha) \quad \ni: \epsilon_i(\alpha) = y_i - X_i \alpha, \Sigma = \sum_{i=1}^N \epsilon_i(\alpha) \epsilon_i(\alpha)' \quad (11)$$

Multidimensional Calculus

- Khuri - Chapter 7
- **Definition 7.3.1** Let $\mathbf{f} : D \rightarrow R^M$, where $D \subset R^n$, and let $a \in D$. Then $\mathbf{f}(x)$ is continuous at a if

$$\lim_{x \rightarrow a} \mathbf{f}(x) = \mathbf{f}(a) \quad (12)$$

where x remains in D as it approaches a .

- **Theorem 7.3.1** Let $f : D \rightarrow R$ be a real-valued continuous function defined on a closed and bounded set $D \subset R^n$. Then, there exists points p and q in D for which

$$\begin{aligned} f(p) &= \sup_{x \in D} f(x) \\ f(q) &= \inf_{x \in D} f(x) \end{aligned} \quad (13)$$

Multidimensional Calculus - Jacobian

- Consider the function $f : R^2 \rightarrow R^2$, for example the simple linear function

$$\begin{aligned} y_1 &= g_1(x) = a_{11}x_1 + a_{12}x_2 \\ y_2 &= g_2(x) = a_{21}x_1 + a_{22}x_2 \end{aligned} \quad (14)$$

- The Jacobian which encapsulates the mapping from one space to another can be written as

$$\nabla_x f(x) = J_r(x) = \begin{bmatrix} \frac{\partial g_1(x)}{\partial x_1} & \frac{\partial g_1(x)}{\partial x_2} \\ \frac{\partial g_2(x)}{\partial x_1} & \frac{\partial g_2(x)}{\partial x_2} \end{bmatrix} \quad (15)$$

Approximation Functions

- Typically at this point students have been exposed to functional analysis in the guise of expansions several times with some degree of rigor.
 - One typical form used in this lecture is the Taylor series approximation

$$f(x) = f(x^*) + \sum_{i=1}^{\infty} \frac{1}{i!} \left. \frac{\partial^i f(x)}{\partial x^i} \right|_{x \rightarrow x^*} \quad (16)$$

- This expression holds with equality.
- We are usually interested in a second order expansion which results in a slight reformulation of the equality

$$f(x) = f(x^*) + \left. \frac{\partial f(x)}{\partial x} \right|_{x \rightarrow x^*} (x - x^*) + \frac{1}{2} \left. \frac{\partial^2 f(x)}{\partial x^2} \right|_{x \rightarrow x^*} (x - x^*)^2 + R(x, x^*)$$

$$R(x, x^*) = \frac{1}{6} \left. \frac{\partial^3 f(x)}{\partial x^3} \right|_{x \rightarrow x^*} (x - \tilde{x})^3 \text{ for some } \tilde{x} \in [x, x^*] \quad (17)$$

- Again, this equation holds with equality

Approximation Functions - Taylor Series

- This formulation only becomes an approximation in the fact that we don't know what $\tilde{x}$ is in the interval, only that it is bounded (if the derivative is bounded).
- We have already used this approximation in its multivariate form in this lecture

$$f(x) \approx f(x^*) + \nabla_x f(x^*)(x - x^*) + \frac{1}{2}(x - x^*)' \nabla_{xx}^2 f(x^*)(x - x^*) \quad (18)$$

- I really don't know what the third moment of this function would look like.

Fourier Approximations

- The Fourier approximation is somewhat more complex - it states that a function can be approximated by a combination of sines and cosines

$$f(x) \approx \frac{1}{2}\alpha_0 + \sum_{i=1}^{\infty} [\alpha_i \cos(\omega_i x) + \beta_i \sin(\omega_i x)] \quad (19)$$

- This approximation is related to a distribution's characteristic function (see AEB 6571 - Lecture 15 or the my Mathematical Statistics book pp. 83-86).
- The characteristic function for the univariate distribution is

$$f(\lambda) = \exp\left(i\lambda\mu - \frac{\lambda^2\sigma^2}{2}\right) \quad (20)$$

Multivariate Normal Characteristic Function

- Expanding the univariate normal to the multivariate case yields

$$\phi(\lambda) = E[\exp(i\lambda'X)] = \exp\left[i\lambda'\mu - \frac{1}{2}\lambda'\Sigma\lambda\right] \quad (21)$$

Readings - Next Time

- * 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)