Advanced Econometrics

Exam Summer Term 2009, July 28, 2009

Prof. Dr. Nikolaus Hautsch
Prof. Dr. Melanie Schienle
Institute for Statistics and Econometrics
Humboldt-Universität zu Berlin

You have to answer 2 out of 3 problems within 90 minutes (plus 10 minutes "reading time"). If you answer all questions, only the first 2 problems will be taken into account.

You may answer in English or in German. But please stick to one language.

You find necessary underlying results and formulas in the appendix.

Do your best to write legibly. Exams or parts of exams which cannot be read with reasonable effort will not be graded.

Good luck!

Problem 1: GMM Estimation

Let $z_i = (y_i, w_i)$ be a vector of i.i.d. endogenous and predetermined variables for observation $i = 1, \ldots, n$ and let $\theta_0 \in \Theta$ denote the k-dimensional true parameter vector to be estimated. Moreover, let the $(q \times 1)$ vector of functions $\psi(z_i, \theta)$ with $q \geq k$ denote an unconditional moment function with

$$\mathrm{E}[\psi(z_i, \theta_0)] = 0$$

and sample moment function

$$\psi_n(\theta) := \frac{1}{n} \sum_{i=1}^{n} \psi(z_i, \theta).$$

Furthermore, let W_n be a $q \times q$ symmetric positive definite (weighting) matrix with $W_0 := \operatorname*{plim}_{n \to \infty} W_n$.

a) Formulate the objective function of the GMM estimator, $\hat{\theta}_n$, and compute the first order condition of the underlying optimization problem. 3 P

b) Using the mean value theorem show that the first order condition can be rewritten as

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = -(A_n' W_n A_n^*)^{-1} A_n' W_n b_n,$$

where

$$A_n := \frac{1}{n} \sum_{i=1}^{n} \frac{\partial \psi(z_i, \hat{\theta}_n)}{\partial \theta'}, \quad A_n^* := \frac{1}{n} \sum_{i=1}^{n} \frac{\partial \psi(z_i, \theta^*)}{\partial \theta'}, \quad b_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi(z_i, \theta_0),$$

where $\theta^* = \theta_0 + \lambda(\hat{\theta}_n - \theta_0)$ with $0 \leq \lambda \leq 1$.

Hint: Make use of the mean value expansion for a continuously differentiable function $f(x)$ as given by

$$f(x) = f(x_0) + \frac{\partial f(\bar{x})}{\partial x}(x - x_0),$$

where $\bar{x}$ denotes the mean value with $x_0 \leq \bar{x} \leq x$. 5 P

c) Which convergence properties for A_n, A_n^*, b_n, and W_n must hold in order to obtain the asymptotic normality result shown in the Appendix? 2 P

d) Give the optimal form W_0^* of the weighting matrix W_0 (without proof) and interpret it intuitively. Suggest a consistent estimator for W_0^* and explain briefly which underlying assumption justifies its consistency. 2 P

e) Now assume that

$$\text{Cov}\left[\psi(z_i,\theta_0),\psi(z_{i-j},\theta_0)\right]\begin{cases}\neq 0 & \text{if } j=1,\\ =0 & \text{if } j\neq 1.\end{cases}$$

What is the optimal form of W_0^* in this case? Suggest again a consistent estimator. 4 P

f) Assume a linear model

$$y_i = x_i'\beta + \varepsilon_i$$

with potentially endogenous regressors x_i and instruments w_i satisfying $\mathrm{E}[w_i\varepsilon_i] = 0$.

(i) Show that the GMM identification condition corresponds to the rank condition

$\mathrm{E}[w_i x_i']$ *is of full column rank.*

(ii) Show that the rank condition in (i) is equivalent to the condition that the number of predetermined variables excluded from the equation must be greater or equal to the number of (included) endogenous regressors.

4 P

Problem 2: Pseudo Maximum Likelihood

Let $l(y;\theta)$ denote the likelihood function in dependence of data y and parameters θ. Moreover, define the Fisher information $\mathcal{I}(\theta)$ as the covariance matrix of the score vector $\partial \ln l(y;\theta)/\partial\theta$.

a) Given two density functions $f(y)$ and $f^*(y)$, define the Kullback discrepancy between f and f^* as

$$I(f|f^*) = \mathrm{E}^*[\ln f^*(y)/f(y)],$$

where E^* denotes the expectation computed based on f^*.

Show that $I(f|f^*) \geq 0$ and that $I(f|f^*) = 0$ if and only if $f = f^*$. 3 P

b) Show that the expected score is equal to zero,

$$\mathrm{E}\left[\frac{\partial \ln l(y;\theta)}{\partial\theta}\right] = 0.$$

Hint: Integrate both sides of the equation $\int l(y;\theta)dy = 1$ *with respect to* θ *and exploit the fact that for any function* $g(\cdot)$*, we have* $\int g(y;\theta)l(y;\theta)dy = \mathrm{E}[g(y;\theta)]$. 3 P

c) Show that

$$\mathcal{I}(\theta) = \mathrm{E}\left[-\frac{\partial^2 \ln l(y;\theta)}{\partial\theta\partial\theta'}\right].$$

Hint: Compute $\partial^2 \ln l(y;\theta)/(\partial\theta\partial\theta')$ *and re-arrange both sides of the equation in order to exploit again* $\int g(y;\theta)l(y;\theta)dy = \mathrm{E}[g(y;\theta)]$. 4 P

d) Give the definition of the pseudo maximum likelihood (PML) estimator. 2 P

e) Using your insights from (b) and (c) explain why in general for the PML

$$\text{(i)} \qquad 0 \neq \mathrm{E}\left[\frac{\partial \ln l(y;\theta)}{\partial\theta}\right]$$

$$\text{(ii)} \qquad \mathcal{I}(\theta) \neq \mathrm{E}\left[-\frac{\partial^2 \ln l(y;\theta)}{\partial\theta\partial\theta'}\right].$$

2 P

f) Explain briefly the major principle of *consistent* PML estimation for parameters of a conditional mean specification. (No formal proof, but state the underlying necessary relationships.) 2 P

g) Assume a linear regression model with conditionally heteroscedastic errors.

- (i) Formulate the OLS estimator as a PML estimator.
- (ii) Give the explicit form of the inequality considered in (e),(ii).

4 P

Problem 3: Nonparametric Estimation

In the following assume that X and Y are univariate continuously distributed and observations $(X_i, Y_i)_{i=1}^n$ are iid.

For constructing a nonparametric estimator $\hat{F}_{Y|X}$ of a conditional distribution function $F_{Y|X}$ use the fact that

$$F_{Y|X}(y|x) = \mathbb{P}(Y \leq y|X = x) = \mathbb{E}\left(\mathbf{1}_{\{Y \leq y\}}|X = x\right) \tag{1}$$

and then use a conditional mean regression estimator to estimate $F_{Y|X}(y|x)$.

a) State the explicit form of a local constant kernel type estimator $\hat{F}_{Y|X}(y|x)$ according to (1). 1 P

b) Show that the estimator in a) yields indeed a distribution function if the kernel function is a symmetric probability density function. 2 P

c) Under the assumptions above and both $F_{Y|X}$ and the density of the regressor f_X are positive and have second order continuous derivatives with respect to x, give a formal proof that asymptotically for $n \to \infty$ and $nh \to 0$ it is:

$$\mathbb{E}\left(\hat{F}_{Y|X}(y|x)\hat{f}_X(x)\right) = F_{Y|X}(y|x)f_X(x) + h^2 B(y,x) + o(h^2) \tag{2}$$

with $B(y,x) = \frac{1}{2}\kappa_2\left(\frac{\partial^2}{\partial x^2}F_{Y|X}(y|x)f_X(x) + 2\frac{\partial}{\partial x}F_{Y|X}(y|x)\frac{\partial}{\partial x}f_X(x) + F_{Y|X}(y|x)\frac{\partial^2}{\partial x^2}f_X(x)\right)$ and appropriate kernel constant κ_2.

Hint: Use law of iterated expectations before substituting 4 P

d) Use the result in b) to show that

$$\mathbb{E}\left((\hat{F}_{Y|X}(y|x) - F_{Y|X}(y|x))\hat{f}_X(x)\right) =$$
$$= h^2\frac{1}{2}\kappa_2\left(\frac{\partial^2}{\partial x^2}F_{Y|X}(y|x) + 2\frac{\frac{\partial}{\partial x}F_{Y|X}(y|x)\frac{\partial}{\partial x}f_X(x)}{f_X(x)}\right) + o(h^2)\ .$$

2 P

e) Use the above result to determine the asymptotic bias of $\hat{F}_{Y|X}(y|x)$. If you assume that $F_{Y|X}(y|x)$ is very smooth in x. How could you accelerate the speed of convergence of the estimator? State formal conditions and results. Check if the resulting estimator still yields a distribution function as in b) 5 P

f) What are the advantages and disadvantages of employing a local linear instead of a local constant type estimator in this setting? Include results from the lecture in your reasoning without proof. 2 P

g) Now assume the dimension d of regressors is no longer one, but large with $d \gg 1$.

How do you have to modify your estimator? How is the performance of your estimator effected in terms of asymptotic rates of bias and variance? How do you call this phenomenon? Briefly describe a way how to countervail this problem with keeping maximal model flexibility and sketch a way how to estimate in such a scenario. 4 P

Appendix

The asymptotic distribution of the GMM estimator $\hat{\theta}_n$ of parameter θ_0 with unconditional moment function $\psi(z_i, \theta)$ and weighting matrix W_n is given by

$$\sqrt{n}(\hat{\theta}_n(W_n) - \theta_0) \xrightarrow{d} N(0, \Delta_0 V_0 \Delta_0'),$$

where

$$\begin{aligned}
\Delta_0 &\equiv (A_0' W_0 A_0)^{-1} A_0' W_0, \\
A_0 &\equiv \mathrm{E}\left[\frac{\partial \psi(z_i, \theta_0)}{\partial \theta'}\right], \\
V_0 &\equiv \operatorname*{plim}_{n\to\infty} \mathrm{V}\left[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi(z_i, \theta_0)\right] = \operatorname*{plim}_{n\to\infty} \mathrm{V}[\sqrt{n}\psi_n(\theta)], \\
\theta_0 &\equiv \operatorname*{plim}_{n\to\infty}\ \hat{\theta}_n, \\
W_0 &\equiv \operatorname*{plim}_{n\to\infty}\ W_n.
\end{aligned}$$