Which Moments To Match?

Gallant and Tauchen

Econometric Theory, 1996

Sargent Reading Group Presentation

Greg Kaplan

10/2/2007

Introduction

Q: How to choose moment conditions when using GMM to estimate parameters of a structural model?

A: Use expectation, under structural model, of the score from an auxiliary model, as the vector of moment conditions.

Auxiliary Model:

1. fits the data well
2. likelihood can be computed

Notation

- Data: y, observed sample: $\{\tilde{y}_i\}_{i=1}^n$
- Model: $p\left(y|\rho\right)$, data generated by $p\left(y|\rho^0\right)$
- Auxiliary Model: $f\left(y|\theta\right)$

$$\sqrt{n}X_n \approx N\left(0, V_v\right)$$
$$\Leftrightarrow$$
$$\sqrt{n}X_n V_v^{-\frac{1}{2}} \xrightarrow{d} N\left(0, I\right)$$

Moment Conditions

- Auxiliary model parameter estimates:

$$\tilde{\theta}_n = \arg\max_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^{n} \log f\left(\tilde{y}_i | \theta\right)$$

- Exact:

$$m_n\left(\rho, \tilde{\theta}_n\right) = E\left[\frac{\partial}{\partial \theta} \log f\left(y | \tilde{\theta}_n\right) | \rho\right]$$

- By Simulation:

$$m_n\left(\rho, \tilde{\theta}_n\right) = \frac{1}{S} \sum_{s=1}^{S} \frac{\partial}{\partial \theta} \log f\left(y_s | \tilde{\theta}_n\right)$$

with y_s drawn from $p\left(y|\rho\right)$

- Estimator:

$$\hat{\rho}_n = \arg\min_{\rho \in R} m_n\left(\rho, \tilde{\theta}_n\right) W m_n\left(\rho, \tilde{\theta}_n\right)'$$

Efficient Weighting Matrix

- From definition of $\tilde{\theta}_n$:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f\left(\tilde{y}_i|\tilde{\theta}_n\right)=0$$

- Define the "psuedo-true value", θ_n^0 as

$$m_n\left(\rho^0,\theta_n^0\right)\equiv E\left[\frac{\partial}{\partial\theta}\log f\left(y|\theta_n^0\right)|\rho^0\right]=0$$

$$\Longrightarrow\lim_{n\to\infty}\left(\tilde{\theta}_n-\theta_n^0\right)=0$$

- Expand $\frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f\left(y_i|\tilde{\theta}_n\right)$ around θ_n^0:

$$\begin{aligned}\sqrt{n}\left(\tilde{\theta}_n-\theta_n^0\right) &\approx N\left(0,\left(Q_n^0\right)^{-1}\Omega_n^0\left(Q_n^0\right)^{-1}\right)\\ Q_n^0 &= \frac{\partial}{\partial\theta'}m_n\left(\rho^0,\theta_n^0\right)\\ \Omega_n^0 &= Var\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f\left(\tilde{y}_i|\theta_n^0\right)\right]\end{aligned}$$

- Expand $m_n\left(\rho^0,\tilde{\theta}_n\right)$ around $m_n\left(\rho^0,\theta_n^0\right)=0$:

$$\sqrt{n}m_n\left(\rho^0,\tilde{\theta}_n\right)=\sqrt{n}Q_n^0\left(\tilde{\theta}_n-\theta_n^0\right)+o_p\left(1\right)$$

$$\Longrightarrow\sqrt{n}m_n\left(\rho^0,\tilde{\theta}_n\right)\approx N\left(0,\Omega_n^0\right)$$

- Efficient weighting matrix is $W=\left(\Omega_n^0\right)^{-1}$:

$$\hat{\rho}_n=\arg\min_{\rho\in R}m_n\left(\rho,\tilde{\theta}_n\right)\left(\tilde{\Omega}_n\right)^{-1}m_n\left(\rho,\tilde{\theta}_n\right)'$$

Consistency and Asymptotic Normality

Identification condition: $\exists\ n_0$ such that $\forall\ n \geq n_0$

$$m_n\left(\rho, \theta_n^0\right) = 0$$
$$\Longrightarrow$$
$$\rho = \rho^0$$

Then:

1. $\hat{\rho}_n \xrightarrow{a.s} \rho^0$
2. $\sqrt{n}\left(\hat{\rho}_n - \rho^0\right) \approx N\left(0, \left(M_n^0\right)^{-1} \Omega_n^0 \left(M_n^0\right)^{-1}\right)$
3. $\hat{M}_n \xrightarrow{a.s} M_n^0$

where

$$\begin{aligned} M_n^0 &= \frac{\partial}{\partial \rho'} m_n\left(\rho^0, \theta_n^0\right) \\ \hat{M}_n &= \frac{\partial}{\partial \rho'} m_n\left(\hat{\rho}, \tilde{\theta}_n\right) \end{aligned}$$

Smooth-Embededness

$\exists$ an open neighborhood, R^0, of ρ^0 and twice continuosuly differentiable mapping, $g : R^0 \rightarrow \Theta$ s.t:

$$p\left(y|\rho\right) = f\left(y|g\left(\rho\right)\right) \;\; \forall \; \rho \in R^0$$

Define:

1. $\hat{\rho}_{mcs} = \arg\min_{\rho \in R} \left[\tilde{\theta} - g\left(\rho\right)\right]' \hat{\Omega}_n \left[\tilde{\theta} - g\left(\rho\right)\right]$

2. $\hat{\rho}_{mle} = \arg\max_{\rho \in R} \frac{1}{n} \sum_{i=1}^{n} \log f\left(\tilde{y}_i | g\left(\rho\right)\right)$

$\Longrightarrow \hat{\rho}_n, \hat{\rho}_{mcs}, \hat{\rho}_{mle}$ all have same asymptotic distribution

Final Thoughts

- Choosing an auxiliary model? Semi-non parametric model with linear basis functions
- All results extend to non-stationary time series models of the form

$$\{p_1(x_1|\rho)\}, \{p_t(y_t|x_{t,}\rho)\}_{t=1}^{\infty}$$