SOLVING THE GMM ESTIMATION PROBLEM

ETHAN LIGON

1. Solving the GMM Estimation Problem

Consider the GMM Criterion associated with a weighting matrix $\boldsymbol{A}$ (so not necessarily efficient!):

$$b_N^A = \arg\min_{b \in B} N \boldsymbol{g}_N(b)^\top \boldsymbol{A} \boldsymbol{g}_N(b).$$

Provided $g_j(b)$ is continuously differentiable the GMM estimator will satisfy the first order conditions:

$$\frac{1}{N} \sum_{j=1}^{N} \frac{\partial g_j(b)}{\partial b^\top} \boldsymbol{A} \boldsymbol{g}_N(\beta) = \mathbf{0}_k;$$

Let $\boldsymbol{Q}_N(b) = \frac{\partial g_j(b)}{\partial b^\top}$, an $\ell \times k$ matrix. Then provided β is in the interior of B we'll have

$$\mathbb{E}\boldsymbol{Q}_N(\beta)^\top \boldsymbol{A} \boldsymbol{g}_N(\beta) = \mathbf{0}_k.$$

Let $\mathbb{E}\boldsymbol{Q}_N(\beta) = \boldsymbol{Q}$. What's the limiting distribution of the GMM estimator?

2. Optimally Weighted GMM

The optimal weights for the moment conditions are given by the inverse of $\Omega = \mathbb{E} g_j(\beta) g_j(\beta)^\top = \boldsymbol{S}^\top \boldsymbol{S}$. So, taking $\boldsymbol{A} = \Omega^{-1}$, we obtain

$$b_N = \arg\min_{b \in B} N \boldsymbol{g}_N(b)^\top \Omega^{-1} \boldsymbol{g}_N(b),$$

3. Feasible Optimally Weighted GMM

Problem: In general the matrix Ω is a function of the unknown parameter vector β. This problem is analogous to the typical problem with GLS, which is that we don't know the covariance matrix of disturbances. This suggests analogous solutions.

(1) Feasible Estimators

Two-step GMM: Analogous to FGLS: Get a "first-step" estimate $\hat{\beta}$, construct $\hat{\Omega}(\hat{\beta})$, then do a second stage.

. *Date*: April 20, 2020.

Iterated GMM: Analogous to iterated FGLS (which converges to Quasi-MLE). Just keep going until things hopefully converge.

Continuously Updated GMM: Analogous to Gaussian MLE. Write weighting matrix as function of b. Has better small-sample properties? But criterion function may be ill-behaved, difficult to minimize.

4. Linear GMM

The linear case is particularly simple to compute, because estimates of the matrix $\boldsymbol{Q}$ don't depend on parameters. We'll focus on the IV problem (but this generalizes straight-forwardly to any linear regression problem):

$$y = X\beta + u \qquad \mathbb{E}(Z^\top u) = 0.$$

Here $g_j(b) = Z_j(y_j - X_j b)$ and $\boldsymbol{g}_N(b) = \boldsymbol{Z}^\top(\boldsymbol{y} - \boldsymbol{X}b)/N$, then sample average of the $g_j(b)$s. And note that $\boldsymbol{Q}_N(b) = \partial\boldsymbol{g}_N(b)/\partial b^\top = \boldsymbol{Z}^\top\boldsymbol{X}$.