Metrics Class notes

Class 3

September 18, 2020

The model

$$y_i = e(z_i, \epsilon_i, \theta)$$

where $\epsilon \sim f(\cdot|z, \theta)$. This condition can arise from FOCs, for example. We care about the first conditional moment (also about others, but first lets try with this one):

$$E[y|z, \theta] = \int e(z_i, \epsilon, \theta) f(\epsilon|z, \theta) d\epsilon \equiv m(z, \theta)$$

The NLLS estimator is

$$\hat{\theta}_{NLLS} = argmin_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^{n} [y_i - m(z, \theta)]^2$$

which is valid iff there exist a unique θ_0 is the unique minimizator

$$\theta_0 = argmin E[y - m(z, \theta)]^2 \tag{1}$$

so θ_0 solves

$$\mathbb{E}\left[\frac{\partial m(z, \theta)}{\partial \theta}[y - m(z, \theta)]\right] = 0$$

and

$$\frac{1}{n} \sum^{N} \frac{\partial m(z_i, \theta)}{\partial \theta}[y_i - m(z_i, \theta)] = 0$$

both are of dimension $k \times 1$. We **have** to use the k moments for NLLS. Actually you can prove that if we care only about the first conditional moment then this k moments are the most efficient (actually $dm/d\theta$ are the optimal instruments, thats the reason for efficiency).

As we discussed last classes, $m(z, \theta)$ is not always computable. In that case we go with simulation

$$m(z,\theta) = \int e(z_i,\epsilon,\theta)f(\epsilon|z,\theta)d\epsilon = \int [e(z_i,\epsilon,\theta)f(\epsilon|z,\theta)/g(\epsilon|z)]g(\epsilon|z)d\epsilon$$

which we approximate with

$$m^s(z,\theta) = \frac{1}{S}\sum e(z_i,\epsilon,\theta)\frac{f(\epsilon|z,\theta)}{g(\epsilon|z)}$$

that was last classes.

GMM

Suppose we care only about the first moment $E[y|z,\theta]$. We have

$$\mathbb{E}[y - m(z,\theta_0)|z] = 0$$

This conditional moment imply plenty of unconditional moments, for example

$$\mathbb{E}[\psi(z)[y - m(z;\theta_0)]] = 0$$

and we pick $\psi(z)$. In NLLS it was equal to the optimal instrument (score), but in GMM we can pick another which will be less efficient. So the GMM estimator θ_{GMM} solve

$$||\frac{1}{n}\sum_{i=1}^{n}\psi(z_i)[y_i - m(z_i,\theta)]||$$

Now imagine that we care about the second moment

$$\mathbb{E}[y^2|z,\theta] = \int e^2(z,\epsilon,\theta)f(\epsilon|z,\theta)d\epsilon \equiv m_2(z,\theta)$$

Now we have

$$\mathbb{E}[y - m(z;\theta_0)|z] = 0 \Rightarrow \mathbb{E}[\psi_1(z)[y - m(z;\theta_0)]] = 0$$
$$\mathbb{E}[y^2 - m^2(z;\theta_0)|z] = 0 \Rightarrow \mathbb{E}[\psi_2(z)[y^2 - m^2(z;\theta_0)]] = 0$$

and the GMM estimator θ_{GMM} solves

$$\begin{cases} \frac{1}{n}\sum^n \psi_1(z_i)[y_i - m(z_i,\theta)] \\ \frac{1}{n}\sum^n \psi_2(z_i)[y_i^2 - m_2(z_i,\theta)] \end{cases} = 0$$

Now the question is how do we choose ψ to get the most efficient estimation.

Pakes and Pollard SMM (1989)

They consider

$$\underbrace{G(\theta)}_{\text{k unconditional moments}} = \int h(x,\theta)dP(x) \equiv \mathbb{E}[h(x,\theta)]$$

with $G(\theta_0) = 0$. In particular

$$h(x,\theta) = \begin{bmatrix} \psi_1[y - m(z,\theta)] \\ \psi_2[y^2 - m_2(z,\theta)] \end{bmatrix}$$

[I GOT LOST HERE...]

The problem that we have is not the integral $\int dP$ because it can be always approximated by the empirical mean. Actually, θ_{GMM} solves $||n^{-1}\sum h(x,\theta)||$. The problem is to compute $h(\cdot)$ so we have to simulate.

Main assumptions:

- **Unbiased simulator:** $h(x,\theta) = \int H(x,\zeta,\theta)dP(\zeta|x)$ which is a demanding assumption since it is related with m. Lets look at the first moment

$$\begin{aligned} h_1(x,\theta) &= \psi_1(z)[y - m(z,\theta)] \\ &= \psi_1(z)y - \psi_1(z)m(z,\theta) \\ &= \psi_1(z)y - \psi_1(z)\underbrace{\int M(x,\zeta,\theta)dP(\zeta|z)}_{=\int e(z,\epsilon,\theta)f(\epsilon|z,\theta)d\epsilon} \end{aligned}$$

Example: this assumtion implies that

$$m(z_i,\theta) \simeq \frac{1}{S}\sum^S M(z_i,\zeta_{is},\theta)$$

where $z_{is} \sim P(\cdot|z)$ iid.

Coming back to the general case,

$$\hat{h}^S(x_i, \theta) = \frac{1}{S} \sum^{S} H(x_i, \zeta_{is}, \theta)$$

where $\zeta_{is} \sim P(\cdot|x_i)$. And SMM estimator $\hat{\theta}^S_{GMM}$ solves

$$||\frac{1}{n} \sum^{N} \hat{h}^S(x_i, \theta)|| = 0$$

In the example:

$$\begin{aligned} h(x, \theta) &= \psi(z)[y - m(z; \theta] \\ \hat{h}^S(x, \theta) &= \psi(z_i)[y_i - \frac{1}{S} \sum^{S} [y - H(z_i, \zeta_{is}, \theta)] \end{aligned}$$

Forget about the whole section 2, do not read. Now we jump to section 3.

Theorem 3.1 Consistency. They show that $\theta^S_{GMM} \xrightarrow{p} \theta_0$ under some assumptions.

Corollary 3.2 Consistency. They show that $\theta^S_{GMM} \xrightarrow{p} \theta_0$ under stronger assumptions.

Theorem 3.2 Asymptotic normality (to do inference). The asy variance of the estimator is

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \to \mathcal{N}(0, (\Gamma'\Gamma)^{-1}\Gamma' V \Gamma (\Gamma'\Gamma)^{-1}$$

[He talked more about the paper now...]