Metrics Class notes

Class 2

September 16, 2020

Last class we had

$$y_i = e(x_i, \epsilon_i; \beta) \tag{1}$$

where $\epsilon_i \sim f(\cdot|x_i;\gamma)$ and $\theta \equiv (\beta, \gamma)$. The conditional mean of y_i given x_i is

$$E[y_i|x_i;\theta] = \int e(x_i, \epsilon_i; \beta) f(\epsilon_i|x_i;\gamma) d\epsilon_i \equiv m(x_i;\theta) \tag{2}$$

Provided θ_0 is identified from $E[y_i|x_i;\theta]$ then NLLS

$$E[y - m(x;\theta)]^2 = E\{E[y - m(x;\theta)]^2|x\} = E\{[m(x;\theta_0) - m(x;\theta)]^2\} \geq 0 \tag{3}$$

Equality holds if and only if $\theta = \theta_0$. So,

$$\hat{\theta}_{NLLS} = argmax \frac{1}{N}\sum_{i=1}^{N}[y_i - m(x_i;\theta)]^2 \tag{4}$$

But this is **impractical**. First, $e()$ might be hard to compute itself (i.e. no closed form solution). Second, Integrating (2) can be hard if ϵ is multidimensional. It is **computational impossible** to approximate an integral with more than four dimensions.

Note: If I want to integrate an univariate $\int_K f(x)dx$ over a compact set K, there is:

- Mathematical approximation: Solve it.
- Simmulation methods: $\int_K \frac{f(x)}{g(x)} g(x)dx \equiv E[f(x)/g(x)]$ if $x \sim g(\cdot)$. This can be approximated by $\frac{1}{S}\sum_{s=1}^{S} f(x_s)/g(x_s)$ where $x_s \sim g(\cdot)$ iid. Note that the choice of $g(\cdot)$ is important because it affects the noise of the simulated object.

Let's go back to our problem. We want a practical way to estimate $\epsilon_i \sim f(\cdot|x_i;\gamma)$. Now is when simulation based methods become relevant. There were two cases where NLLS were impractical:

- $e(\cdot)$ tractable (closed form): SNLLS

$$m(x_i;\theta) \approx \frac{1}{S}\sum^{S} e(x_i,\epsilon_{is};\theta)\frac{f(\epsilon_{is}|x_i;\theta)}{g(\epsilon_{is}|x_i)} \equiv m^s(x_i;\theta) \tag{5}$$

and $m_s(x_i;\theta) = e(\cdot,\epsilon_s)\frac{f(\epsilon_{is}|x_i,\theta)}{g(\epsilon_{is}|x_i)}$.

$$m(x_i;\theta) = \int e(x_i,\epsilon_i,\theta)f(\epsilon_i|x_i,\theta)d\epsilon_i = \int e(x_i,\epsilon_i,\theta)\frac{f(\epsilon_i|x_i,\theta)}{g(\epsilon_i|x_i)}g(\epsilon_i|x_i)d\epsilon_i \tag{6}$$

where $\epsilon_i \sim g(\cdot|x_i)$ iid. Our first instinct is to do

$$\min_{\theta}\frac{1}{N}[y_i - m^s(x_i,\theta)]^2 \tag{7}$$

dont do this. We should do,

$$\min_{\theta}\frac{1}{N}\sum_{i=1}^{N}[y_i - m(x_i;\theta)]^2 - \frac{1}{S(S-1)}\sum_{s=1}^{S}[m_s(x_i;\theta) - m^S(x_i,\theta)]^2 \tag{8}$$

- $e(\cdot)$ intractable: We need an **unbiased** simulator of the mean. That is,

$$m(x,\theta) = E[u|x;\theta] = \int uf(u|x;\theta)du \tag{9}$$

where $u \sim f(\cdot|x;\theta)$ known (In the First Price Auction Paper this was the second highest valuation).

$$m(x_i,\theta) \simeq \frac{1}{S}\sum_{s=1}^{S}u_{is}\frac{f(u_{is}|x_i;\theta)}{g(u_{is}|x_{is})} \equiv m^s(x_i;\theta) \tag{10}$$

Notes: draws must be independent of θ.

Givoanni's question: What about MLE? To do MLE we need the density of $e(\cdot)$. How can we compute $f(y_i|x_i;\theta)$? We need a lot, in a simple example

$$\begin{aligned}
y_i &= exp(\epsilon_i).\ \epsilon_i \sim f(\cdot)\\
F_y(y) = Pr(y \leq y) = Pr[exp(\epsilon_i) \leq y] = Pr[\epsilon_i \leq log(y)] &= F_\epsilon(log(y))\\
\Rightarrow f_y(y) &= \frac{1}{y}f_\epsilon(log(y))
\end{aligned}$$

We need a jacobian... not funny.

0.1 Pakes and Pollard

We had our problem $y = e(x, \epsilon, \beta)$ where $\epsilon \sim f(\cdot|x;\gamma)$, $\theta = (\beta, \gamma)$. We look at the first moment

$$E[y|x;\theta] = \int e(x, \epsilon; \beta) f(\epsilon|x;\gamma) d\epsilon \equiv m(x, \theta) \tag{11}$$

So,

$$E\{y - m(x;\theta_0)\}|x\} = 0 \Rightarrow E\{\psi(x)[y - m(x;\theta_0)]\} = 0 \tag{12}$$

which is that the conditional moment equal zero imply unconditional moments are equal to zero too. $\psi(x)$ are called instruments sometimes. We can write the RHS as

$$\int \psi(x)[y - m(x;\theta)]dP(y, x) = 0 \iff \theta = \theta_0 \tag{13}$$

Pakes and Pollard use the notation $G(\theta) = \int h(x;\theta)dP(x)$, with $h(x;\theta) = \psi(x)[y - m(x;\theta)]$. And $G(\theta) = 0$ if and only if $\theta = \theta_0$. But $h(\cdot)$ is intractable. So we do

$$\hat{G}_N(\theta) \equiv \frac{1}{N}\sum_{i=1}^{N} h(x_i, \theta) \tag{14}$$

and do GMM,

$$\hat{\theta}_{GMM} = \arg\min_{\theta} ||\frac{1}{N}\sum_{i=1}^{N} h(x_i;\theta)|| \tag{15}$$

but h is intractable! Lets simmulate stuff. In page 1028, the paper states

$$h(x, \theta) = \int H(x, \xi, \theta)P(d\xi|x) \tag{16}$$