Method of moments

$Y_i = g(X_i;\theta)$ is not always the case

Suppose $\max_{Y_1,Y_2} \{u(Y_1,Y_2,X;\theta_0) : p_1Y_1 + p_2Y_2 = I\}$

$$\frac{\frac{\partial u}{\partial Y_1}}{\frac{\partial u}{\partial Y_2}} = \frac{p_1}{p_2}$$ optimality condition. Implicit function of Y_1, Y_2, X, and θ.

$$E_0\left[\frac{\partial u}{\partial Y_1}\Big/\frac{\partial u}{\partial Y_2} - \frac{p_1}{p_2}\right] = 0$$

Sample analog: $\hat{\theta}_n$ solves

$$\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\frac{\partial u(Y_{1i},Y_{2i},X_i;\hat{\theta}_n)}{\partial Y_{1i}}}{\frac{\partial u(Y_{1i},Y_{2i},X_i;\hat{\theta}_n)}{\partial Y_{2i}}} - \frac{p_1}{p_2}\right] = 0$$

Suppose $\varphi(Y,X;\theta)$ is s.t.

$E_0[\varphi(Y,X;\theta)] = 0$ iff $\theta = \theta_0$

Where $\left.\begin{array}{l}\varphi \text{ is } p\times 1\\ \theta_0 \text{ is } k\times 1\end{array}\right\}$ and $p > k$

We will not necessarily have a solution to

$\frac{1}{n}\sum_{i=1}^{n}\varphi(Y_i,X_i;\hat{\theta}_n) = 0$ (too many equations in too few unknowns.)

How can we use all the available information?

We will collapse the problem into something manageable

Define $\underset{p\times 1}{m_n(\theta)} = \frac{1}{n}\sum_{i=1}^{n} \varphi(Y_i, X_i; \theta)$

Collapsing this into a scalar problem:

$\underset{1\times p}{m_n(\theta)'} \underset{p\times p}{I_p} \underset{p\times 1}{m_n(\theta)}$ — this works, but we can do better

$Q_n(\theta) = \underset{1\times p}{m_n(\theta)'} \underset{p\times p}{V_n^{-1}} \underset{p\times 1}{m_n(\theta)}$ For some $\underline{V_n} \xrightarrow{p} V$ stochastic
nonstochastic.
V_n, V nonsingular $\forall n$.

We want to $\min Q_n(\theta)$ (ie we want $m_n(\hat\theta_n)$ close to 0).

$\hat\theta_n = \underset{\theta\in\Theta}{\operatorname{argmin}}\ Q_n(\theta)$

$\hat\theta_n$ will satisfy $\frac{\partial Q_n(\hat\theta_n)}{\partial\theta} = 0$ (k equations and k unknowns)

__assumption__ $E_0\left[\sup_{\theta\in\Theta} |\varphi(Y_i, X_i; \theta)|\right] < +\infty$

Then $m_n(\theta) \xrightarrow{p} E_0[\varphi(Y_i, X_i; \theta)]$ uniformly over Θ.

Thus $Q_n(\theta) \equiv m_n(\theta)' V_n^{-1} m_n(\theta) \xrightarrow{p} E_0[\varphi(Y_i, X_i; \theta)]' V^{-1} E_0[\varphi(Y_i, X_i; \theta)]$
$= Q_0(\theta)$

uniformly over Θ.

and therefore $\hat\theta_n = \underset{\theta\in\Theta}{\operatorname{argmin}}\ Q_n(\theta)$ and $\theta_0 = \underset{\theta\in\Theta}{\operatorname{argmin}}\ Q_0(\theta)$

gives us that $\hat\theta_n \xrightarrow{p} \theta_0$

(*) If $f_n \to f$, $g_n \to g$ uniformly, does $f_n g_n \to fg$ uniformly?

Quick review of vector calculus.

Let $h(m) = m' V_n^{-1} m$

$$\Rightarrow \frac{\partial h(m)}{\partial m'} = \underset{1\times p}{2m'} \underset{p\times p}{V_n^{-1}}$$

$$\Rightarrow \frac{\partial h(m(\theta))}{\partial \theta'} = \frac{\partial h(m)}{\partial m'} \frac{\partial m'}{\partial \theta'} = \underset{1\times p}{2m'} \underset{p\times p}{V_n^{-1}} \underset{p\times k}{\frac{\partial m_n(\theta)}{\partial \theta'}} = 0$$

(k equations in k unknowns) and

$$\frac{\partial m_n(\theta)}{\partial \theta'} = \begin{bmatrix} \frac{\partial}{\partial \theta_1} m_{n,1}(\theta), & \dots, & \frac{\partial}{\partial \theta_k} m_{n,1}(\theta) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial \theta_1} m_{n,p}(\theta), & \dots, & \frac{\partial}{\partial \theta_k} m_{n,p}(\theta) \end{bmatrix}$$

$$\Rightarrow \frac{\partial h(m(\hat{\theta}_n))}{\partial \theta} = \left[\frac{\partial h(m(\hat{\theta}_n))}{\partial \theta'}\right]' = 2\frac{\partial m_n(\hat{\theta}_n)}{\partial \theta}(V_n^{-1})' m_n(\hat{\theta}_n) = 0$$

What is the asymptotic distribution?

Looking first at $m_n(\hat{\theta}_n)$, we use the mean value theorem:

$$m_n(\hat{\theta}_n) = m_n(\theta_0) + \frac{\partial m_n(\tilde{\theta})}{\partial \theta'}(\hat{\theta}_n - \theta_0) \text{ for some } \tilde{\theta} \in |\hat{\theta}_n, \theta_0|$$

Recall that $m_n(\hat{\theta}_n) = o_p(1)$

$$\Rightarrow \sqrt{n}\, m_n(\hat{\theta}_n) = \sqrt{n}\, m_n(\theta_0) + \frac{\partial m_n(\tilde{\theta})}{\partial \theta'}\sqrt{n}(\hat{\theta}_n - \theta_0)$$

Recall $2\frac{\partial m_n(\hat{\theta}_n)}{\partial \theta}(V_n^{-1})' m_n(\hat{\theta}_n) = 0$

$$\Rightarrow 2\frac{\partial m_n(\hat{\theta}_n)}{\partial \theta}(V_n^{-1})' \sqrt{n}\, m_n(\hat{\theta}_n) = 0$$

$$0 = 2\,\frac{\partial m_n(\hat\theta_n)}{\partial \theta}(V_n^{-1})'\left[\sqrt{n}\,m_n(\theta_0) + \frac{\partial m_n(\hat\theta)}{\partial \theta'}\sqrt{n}\,(\hat\theta_n - \theta_0)\right]$$

$$= 2\,\frac{\partial m_n(\hat\theta_n)}{\partial \theta}(V_n^{-1})'\sqrt{n}\,m_n(\theta_0) + 2\,\frac{\partial m_n(\hat\theta_n)}{\partial \theta}\frac{\partial m_n(\hat\theta)}{\partial \theta'}\sqrt{n}\,(\hat\theta_n - \theta_0)$$

$$\Rightarrow \sqrt{n}\,(\hat\theta_n - \theta_0) = \left[\frac{\partial m_n(\hat\theta_n)}{\partial \theta}\frac{\partial m_n(\hat\theta)}{\partial \theta'}\right]^{-1}\frac{\partial m_n(\hat\theta_n)}{\partial \theta}(V_n^{-1})'\sqrt{n}\,m_n(\theta_0)$$

assuming $\frac{\partial m_n(\hat\theta_n)}{\partial \theta}\frac{\partial m_n(\hat\theta)}{\partial \theta'}$ is invertible.

assume $E_0\left[\sup_{\theta \in \Theta}\left|\frac{\partial}{\partial \theta_l}\varphi_j(Y_i, X_i; \theta)\right|\right] < +\infty \quad \forall j = 1, \dots, p,\ l = 1, \dots, k$

i.e. $\left\| E_0\left[\sup_{\theta \in \Theta}\left|\frac{\partial}{\partial \theta} m_n(\theta)\right|\right]\right\| < +\infty$

$$\Rightarrow \frac{\partial m_{nj}}{\partial \theta_l}(\theta) = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial \theta_l}\varphi_j(Y_i, X_i; \theta)$$

$$\xrightarrow{p} E_0\left[\frac{\partial}{\partial \theta_l}\varphi_j(Y_i, X_i; \theta)\right] \quad \text{uniformly in } \theta.\ \forall j = 1, \dots, p\ \ \forall l = 1, \dots, k$$

$$\Rightarrow \frac{\partial m_n(\theta)}{\partial \theta'} \xrightarrow{p} E_0\left[\frac{\partial}{\partial \theta'}\varphi(Y_i, X_i; \theta)\right] \quad \text{uniformly in } \theta$$

and thus $\frac{\partial m_n(\hat\theta_n)}{\partial \theta'} \xrightarrow{p} E_0\left[\frac{\partial}{\partial \theta'}\varphi(Y_i, X_i; \theta_0)\right] \equiv A(\theta_0)$