EC475 Problem set 3

LDV models and parametric estimation

Antoine Goujard

19/11/09

Suppose $z \sim \mathrm{N}(\mu_z,\ \sigma_z^2)$.

* Binary case. Let $Q = 1_{z>\lambda}$, then:

- $P(Q=1) = 1 - \Phi(\frac{\lambda-\mu_z}{\sigma_z}) = p = \mathrm{E}(Q) = 1 - P(Q=0)$
- $\mathrm{var}(Q) = p.(1-p)$.

* Truncation case. Let $h(z|Z>c)$ be the pdf of Z at z conditional on $Z>c$ then:

- $h(z|Z>c) = \phi(\frac{z-\mu_z}{\sigma_z}).\frac{1}{1-\Phi(\frac{c-\mu_z}{\sigma_z})}.1_{z>c}$
  Proof: Define the CDF $H(z|Z>c) = P(Z \leq z|Z>c)$ and take derivative wrt z.
- $E(Z|Z>c) = \mu_z + \sigma_z.\frac{\phi(\frac{c-\mu_z}{\sigma_z})}{1-\Phi(\frac{c-\mu_z}{\sigma_z})}$
  Proof: Use the def. of $h(z|Z>c)$ and make the change of variable $u := \frac{z-\mu_z}{\sigma_z}$.

* Selectivity case. Suppose :

$$\boldsymbol{u} := \begin{pmatrix} Z \\ W \end{pmatrix} \sim \mathrm{N}_2 \left(\begin{pmatrix} \mu_Z \\ \mu_W \end{pmatrix}, \begin{pmatrix} \sigma_Z^2 & \sigma_{ZW} \\ \sigma_{ZW} & \sigma_W^2 \end{pmatrix} \right)$$

with $\sigma_{ZW} = \rho.\sigma_Z.\sigma_W$ then:

- $\forall (z,w) \in \mathbb{R}^2$, $h((z,w)|W > c) = \frac{f(z,w)}{P(W>c)}.1_{w>c}$.
  Proof: compute $P(Z \leq z, W \leq w | W > c) = ...$
- If we assume: $E(Z|W) = \mu_Z + \sigma_{ZW}.(\frac{W-\mu_W}{\sigma_W^2})$
- Or, $E(Z|W) = \mu_Z + \rho.\sigma_Z.(\frac{W-\mu_W}{\sigma_W})$
- Then by LIE,
  $E(Z|W > c) = E(E(Z|W)|W > c) = \mu_Z + \rho.\sigma_Z.(\frac{E(W|W>c)-\mu_W}{\sigma_W})$
  From the previous result,
  $E(Z|W > c) = \mu_Z + \rho.\sigma_Z.\frac{\phi(\frac{c-\mu_W}{\sigma_W})}{1-\Phi(\frac{c-\mu_W}{\sigma_W^2})}$

* The previous proof (as in the lecture notes) is based on:

$E(Z|W) = \mu_z + \sigma_{zw}.(\frac{w-\mu_w}{\sigma_w^2})$

Rk: $\frac{\sigma_{zw}}{\sigma_w^2}$ correspond to the plim of the OLS regression of Z on W. It means that for bivariate normal the CEF is the linear projection.
Proof: Let, $U = Z - (\mu_z + \sigma_{zw}.(\frac{w-\mu_w}{\sigma_w^2})) = Z - E^*(Z|W)$

Then:

$$\begin{pmatrix} U \\ W \end{pmatrix} = \begin{pmatrix} -\mu_z + \frac{\sigma_{zw}\mu_w}{\sigma_w^2} \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & -\frac{\sigma_{zw}}{\sigma_w^2} \\ 0 & 1 \end{pmatrix} . \begin{pmatrix} Z \\ W \end{pmatrix}$$

This shows that $(U, W)'$ is a N_2 vector with diag. covar matrix. Hence, U is independent of W. Thus:
$E(Z|W) = E(E^*(Z|W) + U|W) = E^*(Z|W) + E(U|W) =$
$E^*(Z|W) + E(U) = E^*(Z|W)$.

"For estimating LDV models, the GMM approach is preferable to MLE because it only requires assumptions on the moment conditions and not about the full distribution of the data".

eg. The binary choice model:
$y_i^* = \boldsymbol{x}_i'.\beta + \varepsilon_i$ with, $\varepsilon_i \sim (0, 1)$ (latent model)
For some reason, we only observe: $y_i = 1_{y_i^*>0}$

Then, we can assume:

1. $E(y_i|\boldsymbol{x}_i) = P(y_i = 1|\boldsymbol{x}_i) = P(\varepsilon_i > -\boldsymbol{x}_i'.\beta) = \Phi(\boldsymbol{x}_i'.\beta)$.
   which is a GMM assumption.
2. $\varepsilon_i|\boldsymbol{x}_i \sim \mathrm{N}(0, 1)$ which is a MLE assumption.

- Clearly, (1) only requires a correct specification for the CEF, but (1) is already a strong assumption on the shape of the CEF. (rk, as $Y_i \in \{0, 1\}$, (1) gives us <u>all</u> the moments of $Y_i|X_i$.)

- The two types of assumption are equivalent in this case.
- In both cases, this requires a strong prior on F, the Filter function which allows to go from the latent to the observed model.
- So GMM is not really more robust in this case or in general for LDV because it is very difficult to get the correct specification for the CEF without knowing the underlying distribution of the error. So both methods become unreliable when this assumption fails.
- In general, even if weaker identification of β (than MLE or GMM) is feasible, we are ultimately interested in the marginal effects: $m_k = \frac{\partial P(y_i=1|\boldsymbol{x}_i)}{\partial x_k} = \beta_k . \phi(\boldsymbol{x}_i' . \boldsymbol{\beta})$ so identification of β is not enough! We need the correct specification of the filter function.