System Identification, Lecture 5

Kristiaan Pelckmans (IT/UU, 2338)

Course code: 1RT875, Report code: 61806,
F, FRI Uppsala University, Information Technology

03 Februari 2010

Lecture 5

- Instrumental Variables Methods (IVM) (Ch. 8)

Main Idea- modify the LS method to be consistent also for correlated disturbances

Least Squares Revisited

The LS estimate

$$\hat{\theta} = \left(\frac{1}{N} \sum_{t=1}^{N} \varphi(t)\varphi^T(t) \right)^{-1} \left(\frac{1}{N} \sum_{t=1}^{N} \varphi(t)t(t) \right)$$

has estimation error (when $N \to \infty$)

$$(\hat{\theta} - \theta_0) = \left(E[\varphi(t)\varphi^T(t)] \right)^{-1} E\left[\varphi^T(t)\epsilon(t) \right]$$

Consequently, for $\hat{\theta} - \theta_0 \to 0_n$, one needs

$$E[\phi^T(t)\epsilon(t)] = 0_n,$$

which is satisfied (essentially) only if $\epsilon(t)$ is white noise. Hence the LS estimate is not necessarily consistent for correlated noise sources!

Cure:

- PEM. Model the noise.
  - Applicable for general model structures.
  - In general very good properties of the estimates.
  - Computationally quite demanding.
- Instrumental Variable Method (IVM). Do not model the noise.
  - Maintain the simple OLS structure.
  - Computationally simple and efficient.
  - Consistent for correlated noise.
  - Less robust and statistical efficient than PEM.

The IV Method

Introduce a time series $(z(t))_t \subset \mathbb{R}^n$ with entries uncorrelated to the noise sequence $(\epsilon(t))_t$. Then one has for $N \to \infty$ that (second moments)

$$0_n = \frac{1}{N}\sum_{t=1}^{N} z(t)\epsilon(t) = \frac{1}{N}\sum_{t=1}^{N} z(t)\left(y(t) - \varphi(t)\theta_0\right)(t)$$

which yields (if inverse exists)

$$\hat{\theta}_z = \left(\frac{1}{N}\sum_{t=1}^{N} z(t)\varphi^T(t)\right)^{-1}\left(\frac{1}{N}\sum_{t=1}^{N} z(t)t(t)\right)$$

The time-series $z(t)$ are denoted as **instruments**. Note that the OLS is obtained when $\varphi(t) = z(t)$.

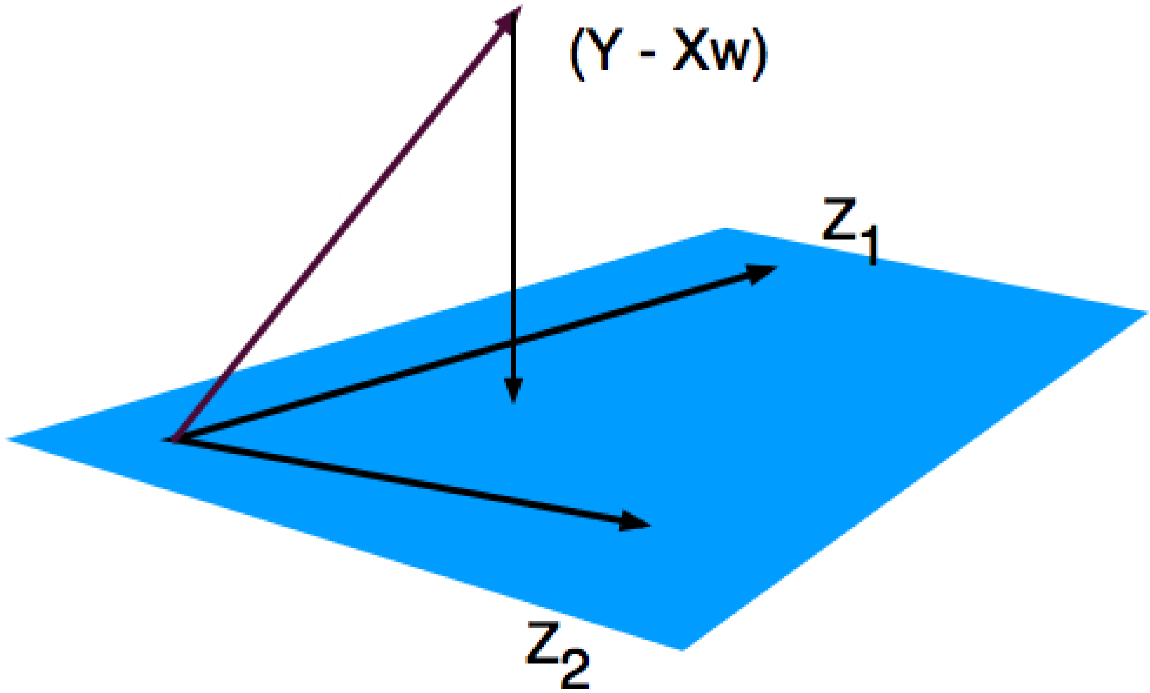

Figure 1: Instrumental Variable as Modified Projection.

Choice of Instruments

Obviously the choice of instruments is very important: They have to be chosen such that

1. such that $(z(t))_t$ is uncorrelated to $(\epsilon(t))_t$.

2. such that the matrix

$$\mathbf{R}_z = \frac{1}{N}\sum_{t=1}^{N} z(t)\varphi^T(t)$$

   has full rank. In other words, it is crucial that $z(t)$ and $\varphi(t)$ are correlated!

In practice those requirements are satisfied by choosing the instruments as delayed/filtered inputs. A common choice is:

$$z(t) = (-\eta(t-1), \ldots, -\eta(t-n_a), -u(t-1), \ldots, u(t-n_b))^T$$

where

$$C(q^{-1})\eta(t) = D(q^{-1})u(t).$$

In case $C(q^{-1}) = 1$ and $D(q^{-1}) = -q^{-n_b}$ one has

$$z(t) = (u(t-1), \ldots, u(t-n_a-n_b))^T$$

rem. We exploit the assumption that $(u(t))_t$ and $(\epsilon(t))_t$ are uncorrelated.

Extended IV Methods

Recall that the basic IV estimate can be obtained by minimizing

$$\hat{\theta}_{IV} = \operatorname*{argmin}_{\theta} \frac{1}{2} \left\| \sum_{t=1}^{N} z(t)\epsilon_\theta(t) \right\|_2$$

More flexibility is obtained when the instruments $(z(t))_t$ are augmented to dimension n_z (with $n_z > n$). and if we allow for weighting and prefiltering of the residuals by some stable filter $F(q^{-1})$, i.e.

$$\hat{\theta}_{IV} = \operatorname*{argmin}_{\theta} \frac{1}{2} \left\| \sum_{t=1}^{N} z(t)F(q^{-1})\epsilon_\theta(t) \right\|_{\mathbf{Q}}^2$$

where $\mathbf{Q} \in \mathbb{R}^{n_z \times n_z}$ is a positive definite weighting matrix such that $\|x\|_{\mathbf{Q}}^2 = x^T \mathbf{Q} x$.

Working out terms gives the extended IV method:

$$\min_{\theta} \frac{1}{2} \left\| \left(\sum_{t=1}^{N} z(t) F(q^{-1}) \varphi^T(t) \right) \theta - \left(\sum_{t=1}^{N} z(t) F(q^{-1}) y(t) \right) \right\|_{\mathbf{Q}}^2$$

When $F(q^{-1}) = 1$ and $\mathbf{Q} = I_{n_\theta}$, the basic IV method is recovered.

Introduce

$$\begin{cases} \mathbf{R}_N = \frac{1}{N} \sum_{t=1}^{N} z(t) F(q^{-1}) \varphi^T(t) \mathbf{r}_N = \frac{1}{N} \sum_{t=1}^{N} z(t) F(q^{-1}) y(t) \end{cases}$$

Then

$$\begin{aligned} \hat{\theta}_F &= \operatorname*{argmin}_{\theta} \|\mathbf{R}_N \theta - \mathbf{r}_N\|_{\mathbf{Q}}^2 \\ &= \operatorname*{argmin}_{\theta} (\mathbf{R}_N \theta - \mathbf{r}_N)^T \mathbf{Q} (\mathbf{R}_N \theta - \mathbf{r}_N) \\ &= \left(\mathbf{R}_N^T \mathbf{Q} \mathbf{R}_N \right)^{-1} \mathbf{R}_N^T \mathbf{Q} \mathbf{r}_N. \end{aligned}$$

Numerical unstable!

Rem.: $\mathbf{R}_N$ is in general not square.

Theoretical Analysis

Assumptions

1. The system is strictly causal and asymptotically stable.

2. The input is PE of a sufficiently high order.

3. The disturbance is a stationary stochastic process with rational spectral density

$$\epsilon(t) = H(q^{-1})e(t), \qquad E[e^2(t)] = \lambda^2$$

4. The inputs and disturbances are not correlated (open loop).

5. The model θ and the 'true' system θ_0 have the same transfer function if and only if $\theta = \theta_0$ (PI)

6. The instruments and disturbances are uncorrelated.

Given the system

$$y(t) = \varphi(t)\theta_0 + \epsilon(t)$$

Then

$$\begin{aligned}
\mathbf{r}_N &= \frac{1}{N}\sum_{t=1}^{N} z(t)F(q^{-1})y(t) \\
&= \frac{1}{N}\sum_{t=1}^{N} z(t)F(q^{-1})\varphi(t)\theta_0 + \frac{1}{N}\sum_{t=1}^{N} z(t)F(q^{-1})\epsilon(t) \\
&= \mathbf{R}_N\theta_0 + \mathbf{q}_N
\end{aligned}$$

Thus

$$\hat{\theta}_{\mathbf{Q}} - \theta_0 = (\mathbf{R}_N^T \mathbf{Q} \mathbf{R}_N)^{-1} \mathbf{R}_N^T \mathbf{Q} \mathbf{r}_N \to (\mathbf{R}^T \mathbf{Q} \mathbf{R})^{-1} \mathbf{R}^T \mathbf{Q} \mathbf{r}$$

with

$$\left\{ \mathbf{R}_N = E\left[z(t)F(q^{-1})\varphi^T(t)\right] \mathbf{r}_N = E\left[z(t)F(q^{-1})y(t)\right] \right.$$

Therefore the IV estimate will be consistent if

1. $\mathbf{R}$ has full rank. (Inaccurate if $\mathbf{R}$ nearly rank deficient)

2. $E[z(t)F(q^{-1})\epsilon(t)] = 0_n$

Furthermore, the parameter estimation errors are asymptotically gaussian distributed with zero mean and covariance $\mathbf{P}_{IV} \in \mathbb{R}^{n \times n}$, or

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \sim \mathcal{N}(0, \mathbf{P}_{IV})$$

where

$$\mathbf{P}_{IV} = \lambda^2 (\mathbf{R}^T \mathbf{Q} \mathbf{R})^{-1} (\mathbf{R}^T \mathbf{Q} \mathbf{S} \mathbf{Q} \mathbf{R}) (\mathbf{R}^T \mathbf{Q} \mathbf{R})^{-1}$$

and

$$\mathbf{S} = E\left[F(q^{-1})H(q^{-1})z(t)\right] E\left[F(q^{-1})H(q^{-1})z(t)\right]^T$$

For MIMO systems, $\mathbf{S}$ must be modified.

Optimal IVM

The main use for expressing $\mathbf{P}_{IV}$ is for comparison with $\mathbf{P}$. (recall that PEM is efficient for Gaussian disturbances). A good choice of instruments leads to 'optimal' IVM. For example

$$\begin{cases} z(t) = H^{-1}(q^{-1})\tilde{\varphi}(t) \\ F(q^{-1}) = H^{-1}(q^{-1}) \\ \mathbf{Q} = I_n \end{cases}$$

where $\tilde{\varphi}(t)$ is the noise-free part of $\varphi(t)$. Then

$$\mathbf{P}_{IV}^{opt} = \lambda^2 \left(E[(H(q^{-1})\tilde{\varphi}(t))(H(q^{-1})\tilde{\varphi}(t))^T] \right)^{-1}$$

and $\mathbf{P}_{PEM} \leq \mathbf{P}_{IV}^{opt} \leq \mathbf{P}_{IV}$.

Approximative implementation of the optimal IVM

Note that the optimal instruments require knowledge of the 'true' undisturbed outputs, the noise variance and the shaping filter $H(q^{-1})$, hence

1. Use OLS to obtain $\hat{\theta}_N^{(1)}$
2. Estimate $\tilde{\varphi}(t)$ as

$$\tilde{\varphi}^{(1)}(t) = \frac{B(q^{-1}, \hat{\theta}_N^{(1)})}{A(q^{-1}, \hat{\theta}_N^{(1)})} u(t)$$

3. Use the IV with instruments

$$z^{(1)}(t) = \Big(-\tilde{\varphi}^{(1)}(t-1), \ldots, \tilde{\varphi}^{(1)}(t-n_a), u(t-1), \ldots, u(t-n_b$$

4. Estimate $H(q^{-1})$ based on the residuals. Postulate an AR model and use OLS

5. Use the IVM with $F(q^{-1})$

Summary IVM

- The implementation of PEM is computationally too demanding in many cases.
- The comp. convenient OLS is normally bias for such model structures (correlated noise)
- The IV method uses instruments that are uncorrelated with the disturbances to make a OLS-alike formulation.
- The parameters obtained by the IVM are consistent (when choosing the instruments with care). but it has a (slightly) larger variance than PEM estimates.
- Approximately optimal IV methods can be implemented in an iterative way to achieve lowest possible variance of the IV estimates.