Generalised Method of Moments

(Wooldridge, chapter 8)

1 Introduction

Again we consider the model,

$$y = \mathbf{x}\beta + e \tag{1}$$

where $\mathbf{x}$ is a $k*1$ vector of explanatory variables and some are endogenous so that the OLS1 assumption, $E(\mathbf{x}'e) = 0$, does not hold.

And again, we assume there exists a set of variables $\mathbf{z}$ of size $l \geqslant k$ that satisfy the 2SLS assumptions, which we now call the GMM assumptions,

GMM1: $E(\mathbf{z}'e) = 0$

GMM2: $\text{rank}\left(E(\mathbf{z}'\mathbf{x})\right) = k$

When $l = k$ the IV estimator is the solution of the sample counterpart of the moment

$$E\left(\mathbf{z}'e\left(b^{IV}\right)\right) = 0 \tag{2}$$

However, if $l > k$ this defines a set of l equations to determine k parameters. Thus, the system has no solution (over-identification).

2 The GMM solution

The GMM procedure aims at solving the moment conditions (2) as closely as possible. Closeness is measured in terms of a weighted squared error:

$$\begin{aligned} \min_{b^{GMM}} \quad & \left(\sum_{i=1}^{N} \mathbf{z}_i'\left(y_i - \mathbf{x}_i b^{GMM}\right)\right)' W \left(\sum_{i=1}^{N} \mathbf{z}_i'\left(y_i - \mathbf{x}_i b^{GMM}\right)\right) \\ & = \left(Z'\left(\mathbf{y} - Xb^{GMM}\right)\right)' W \left(Z'\left(\mathbf{y} - Xb^{GMM}\right)\right) \end{aligned}$$

where W is an $l*l$ matrix of weights. The additional assumption of GMM is the following:

GMM3: W is a non-random, symmetric and positive definite matrix.

The purpose of introducing W is to produce an estimator as precise as possible, as we will see below. The foc for the above minimisation problem are

$$X'ZW(Z'\mathbf{y} - Z'Xb^{GMM}) = 0$$

which under (GMM2) and (GMM3) can be explicitly solved for b^{GMM}:

$$b^{GMM} = (X'ZWZ'X)^{-1}X'ZWZ'\mathbf{y}$$

3 Asymptotic properties of the GMM estimator

Typically W is not known and instead we need to estimate it. In such case, we replace (GMM3) by the alternative assumption,

GMM3': A consistent estimator $\widehat{W}$ of W exists, where W is a non-random, symmetric and positive definite matrix.

Consistency of GMM Under (GMM1), (GMM2) and (GMM3'), the GMM estimator is consistent. To see why, write

$$\begin{aligned} b^{GMM} &= (X'Z\widehat{W}Z'X)^{-1}X'Z\widehat{W}Z'\mathbf{y} \\ &= \beta + \left[\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i'\mathbf{z}_i\right)\widehat{W}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{z}_i'\mathbf{x}_i\right)\right]^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i'\mathbf{z}_i\right)\widehat{W}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{z}_i'e_i\right) \end{aligned}$$

The LLN ensures that,

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i'\mathbf{z}_i \xrightarrow{p} E(\mathbf{x}_i'\mathbf{z}_i)$$

which has rank k under (GMM2). (GMM3') ensures that $\widehat{W} \xrightarrow{p} W$ where W is pd. Thus,

$$\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i'\mathbf{z}_i\right)\widehat{W}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{z}_i'\mathbf{x}_i\right) \xrightarrow{p} E(\mathbf{x}_i'\mathbf{z}_i)WE(\mathbf{z}_i'\mathbf{x}_i)$$

which has rank k and is, therefore, invertible.

Using the same reasoning,

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i'\mathbf{z}_i\widehat{W} \xrightarrow{p} E(\mathbf{x}_i'\mathbf{z}_i)W$$

The LLN ensures that,

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{z}_i'e_i \quad \overset{p}{\to} \quad E(\mathbf{z}_i'e_i)$$

which under (GMM1) equals zero.

Thus $b^{GMM} \overset{p}{\to} \beta$.

Asymptotic normality of GMM From the above expression for b^{GMM} we can write

$$\sqrt{N}\left(b^{GMM}-\beta\right) \quad = \quad \left[\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i'\mathbf{z}_i\right)\widehat{W}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{z}_i'\mathbf{x}_i\right)\right]^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i'\mathbf{z}_i\right)\widehat{W}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{z}_i'e_i\right) \quad (3)$$

We have seen that

$$\left[\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i'\mathbf{z}_i\right)\widehat{W}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{z}_i'\mathbf{x}_i\right)\right]^{-1} \quad \overset{p}{\to} \quad \left[E(\mathbf{x}_i'\mathbf{z}_i)WE(\mathbf{z}_i'\mathbf{x}_i)\right]^{-1}$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i'\mathbf{z}_i\widehat{W} \quad \overset{p}{\to} \quad E(\mathbf{x}_i'\mathbf{z}_i)W$$

Applying the CLT to the last term in (3) yields

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{z}_i'e_i \quad \overset{a}{\sim} \quad \mathcal{N}(0,\Sigma)$$

where $\Sigma = E(e_i^2\mathbf{z}_i'\mathbf{z}_i)$. But then

$$\sqrt{N}\left(b^{GMM}-\beta\right) \quad \overset{a}{\sim} \quad \mathcal{N}(0,\Omega)$$

where the variance covariance matrix Ω is,

$$\Omega \quad = \quad \left[E(\mathbf{x}'\mathbf{z})WE(\mathbf{z}'\mathbf{x})\right]^{-1}E(\mathbf{x}'\mathbf{z})W\Sigma WE(\mathbf{z}'\mathbf{x})\left[E(\mathbf{x}'\mathbf{z})WE(\mathbf{z}'\mathbf{x})\right]^{-1}$$

4 What is the optimal choice of W?

The optimal choice of W is the one that minimises the variance of the GMM estimator. It can be proved that the optimal choice is (exercise)

$$\begin{aligned} W &= \Sigma^{-1} \\ &= E\left[e^2\mathbf{z}'\mathbf{z}\right]^{-1} \\ &= \text{var}(\mathbf{z}'e)^{-1} \end{aligned}$$

In this case, the asymptotic covariance of $\sqrt{N}\left(b^{GMM}-\beta\right)$ simplifies to,

$$\text{avar}\left(\sqrt{N}\left(b^{GMM}-\beta\right)\right) = \left[E(\mathbf{x}'\mathbf{z})\Sigma^{-1}E(\mathbf{z}'\mathbf{x})\right]^{-1}$$

5 Homoscedastic case

Suppose $E\left[e^2\mathbf{z}'\mathbf{z}\right]=\sigma_e^2 E\left[\mathbf{z}'\mathbf{z}\right]$. This is assumption (2SLS4). In this case

$$\widehat{W} = \left(\frac{\sigma_e^2}{N}Z'Z\right)^{-1}$$

and the GMM estimator is

$$\begin{aligned}
b^{GMM} &= \left(X'Z\left(\frac{\sigma_e^2}{N}Z'Z\right)^{-1}Z'X\right)^{-1}X'Z\left(\frac{\sigma_e^2}{N}Z'Z\right)^{-1}Z'\mathbf{y} \\
&= \left(X'Z\left(Z'Z\right)^{-1}Z'X\right)^{-1}X'Z\left(Z'Z\right)^{-1}Z'\mathbf{y} \\
&= \left(\widehat{X}'X\right)^{-1}\widehat{X}'\mathbf{y} \\
&= \left(\widehat{X}'\widehat{X}\right)^{-1}\widehat{X}'\mathbf{y} \\
&= \beta^{2SLS}
\end{aligned}$$

That is, the 2SLS is efficient under homoscedastic residuals.

6 Implementing GMM

Since generally we do not know W beforehand, we need to follow some steps to produce the GMM estimator. The procedure proposed by Hansen is,

step1: estimate the model using 2SLS (consistent but not efficient under the GMM assumptions);

step2: obtain $\widehat{W}=\left(\frac{1}{N}\sum_{i=1}^{N}\widehat{e}_i^2\mathbf{z}_i'\mathbf{z}_i\right)^{-1}$;

step1: estimate β by GMM using $\widehat{W}$.