TA session2# 9

Jun Sakamoto

January 10,2019

Contents

1 GMM linear model

Consider the following model.

$$Z'y = Z'X\beta + Z'u \tag{1}$$

where Z, y, X, β and u are $n \times r, n \times 1, n \times k, k \times 1$ and $n \times 1$ matrices and vectors, and $r \geq k$. Rewrite the above regression as follows

$$y^* = X^*\beta + u^* \tag{2}$$

where $y^* = Z'y, X^* = Z'X$ and $u^* = Z'u$. Then,

$$E(u^*) = 0, V(u^*) = \sigma^2 Z'Z = \sigma^2\Omega \tag{3}$$

GMM estimator is obtained by using GLS, that is,

$$\begin{aligned}\beta_{GMM} =&(X^{*'}\Omega^{-1}X^*)^{-1}X^{*'}\Omega^{-1}y^* \\ =&(X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y\end{aligned} \tag{4}$$

Let us assume the following condition,

$$\begin{aligned}&\frac{1}{n}Z'Z \to_p M_{zz} < \infty \\ &\frac{1}{n}Z'X \to_p M_{zx} < \infty \\ &\frac{1}{n}Z'u \to_p 0\end{aligned} \tag{5}$$

Under those assumptions, GMM estimator is consistent. Thus,

$$
\begin{aligned}
\beta_{GMM} =&(X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y \\
=&(X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'(X\beta+u) \\
=&\beta+(X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'u \\
=&\beta+(\frac{1}{n}X'Z(\frac{1}{n}Z'Z)^{-1}\frac{1}{n}Z'X)^{-1}\frac{1}{n}X'Z(\frac{1}{n}Z'Z)^{-1}\frac{1}{n}Z'u \\
\to_p&\beta+(M_{zx}M_{zz}^{-1}M_{zx})^{-1}M_{zx}M_{zz}^{-1}\times 0=\beta
\end{aligned} \tag{6}
$$

By the same logic of IV method,

$$
\begin{aligned}
\sqrt{n}(\beta_{GMM}-\beta) =&(\frac{1}{n}X'Z(\frac{1}{n}Z'Z)^{-1}\frac{1}{n}Z'X)^{-1}\frac{1}{n}X'Z(\frac{1}{n}Z'Z)^{-1}\frac{1}{\sqrt{n}}Z'u \\
\to&N(0,\sigma^2(M_{zx}M_{zz}^{-1}M_{zx})^{-1})
\end{aligned} \tag{7}
$$

In practice, we use the following distribution.

$$
\beta_{GMM}\sim N(\beta,s^2(X'Z(Z'Z)^{-1}Z'X)^{-1}) \tag{8}
$$

where $s^2=\frac{1}{n-k}(y-X\beta_{GMM})'(y-X\beta_{GMM})$

Empirical Example1 Hayashi's textbook pp.250.
Consider the wage equation

$$
LW=\beta S+\gamma IQ+\delta h+e \tag{9}
$$

where LW is log wages, S is schooling, and h is control variables. Estimate the wage equation by GMM. 3results are characterized by below.

1. Schooling is exogenous variable and weighting matrix is $(Z'Z)^{-1}$
2. Schooling is exogenous variable and weighting matrix which considered heteroscedasticity
3. Schooling is endogenous variable and weighting matrix which considered heteroscedasticity

Estimating by RATS.
Result1

```
Linear Regression - Estimation by GMM
Dependent Variable LW
Usable Observations                    758
Degrees of Freedom                     745
Mean of Dependent Variable    5.6867387863
Std Error of Dependent Variable 0.4289493543
Standard Error of Estimate    0.3302676854
Sum of Squared Residuals      81.262174293
J-Specification(3)                 74.1649
Significance Level of J          0.0000000
Durbin-Watson Statistic             1.7208

    Variable                        Coeff      Std Error      T-Stat      Signif
************************************************************************************
1.  S                          0.076835442  0.013185921      5.82708  0.00000001
2.  IQ                        -0.001401432  0.004113143     -0.34072  0.73331404
3.  EXPR                       0.031233938  0.006693110      4.66658  0.00000306
4.  TENURE                     0.048999777  0.007343684      6.67237  0.00000000
5.  RNS                       -0.100681117  0.029588671     -3.40269  0.00066726
6.  SMSA                       0.133597277  0.026324545      5.07501  0.00000039
7.  Y66                        4.436784464  0.289950362     15.30188  0.00000000
8.  Y67                        4.415770982  0.293999764     15.01964  0.00000000
9.  Y68                        4.525883796  0.286858823     15.77739  0.00000000
10. Y69                        4.644032862  0.296708747     15.65182  0.00000000
11. Y70                        4.670615269  0.309123900     15.10920  0.00000000
12. Y71                        4.671336935  0.302109595     15.46239  0.00000000
13. Y73                        4.772811156  0.302499921     15.77789  0.00000000
```

Result2

```
Linear Regression - Estimation by Instrumental Variables
With Heteroscedasticity/Misspecification Adjusted Standard Errors
Dependent Variable LW
Usable Observations                    758
Degrees of Freedom                     745
Mean of Dependent Variable    5.6867387863
Std Error of Dependent Variable 0.4289493543
Standard Error of Estimate    0.3302676854
Sum of Squared Residuals      81.262174293
J-Specification(3)                 71.5752
Significance Level of J          0.0000000
Durbin-Watson Statistic             1.7208

    Variable                        Coeff      Std Error      T-Stat      Signif
************************************************************************************
1.  S                          0.076835442  0.013296885      5.77845  0.00000001
2.  IQ                        -0.001401432  0.004155593     -0.33724  0.73593599
3.  EXPR                       0.031233938  0.006728753      4.64186  0.00000345
4.  TENURE                     0.048999777  0.007419060      6.60458  0.00000000
5.  RNS                       -0.100681117  0.029911276     -3.36599  0.00076269
6.  SMSA                       0.133597277  0.026589325      5.02447  0.00000050
7.  Y66                        4.436784464  0.293344054     15.12485  0.00000000
8.  Y67                        4.415770982  0.297636143     14.83614  0.00000000
9.  Y68                        4.525883796  0.290049068     15.60386  0.00000000
10. Y69                        4.644032862  0.300356739     15.46172  0.00000000
11. Y70                        4.670615269  0.312069317     14.96660  0.00000000
12. Y71                        4.671336935  0.305381496     15.29673  0.00000000
13. Y73                        4.772811156  0.305920948     15.60145  0.00000000
```

Result3

```
Linear Regression - Estimation by Instrumental Variables
With Heteroscedasticity/Misspecification Adjusted Standard Errors
Dependent Variable LW
Usable Observations                  758
Degrees of Freedom                   745
Mean of Dependent Variable      5.6867387863
Std Error of Dependent Variable 0.4289493543
Standard Error of Estimate      0.3853599722
Sum of Squared Residuals        110.63421957
J-Specification(2)                   11.2947
Significance Level of J            0.0035269
Durbin-Watson Statistic               1.8032

    Variable                       Coeff      Std Error      T-Stat      Signif
*******************************************************************************
1.  S                           0.176980773  0.020966861      8.44098  0.00000000
2.  IQ                         -0.010049394  0.004953785     -2.02863  0.04249603
3.  EXPR                        0.048729196  0.008180713      5.95659  0.00000000
4.  TENURE                      0.042330673  0.009630671      4.39540  0.00001106
5.  RNS                        -0.105322483  0.033960937     -3.10128  0.00192684
6.  SMSA                        0.124568446  0.031222519      3.98970  0.00006616
7.  Y66                         4.069138570  0.339509669     11.98534  0.00000000
8.  Y67                         4.019250991  0.344587276     11.66396  0.00000000
9.  Y68                         4.113533133  0.337028355     12.20530  0.00000000
10. Y69                         4.214657968  0.350230931     12.03394  0.00000000
11. Y70                         4.232791698  0.362089659     11.68990  0.00000000
12. Y71                         4.169772647  0.356916670     11.68276  0.00000000
13. Y73                         4.175477510  0.360696265     11.57616  0.00000000
```

Empirical Example2 “How does the European Central Bank react to oil prices?” Guillaume and Julien Licheron Economics Letters 116,pp445-447

We investigate the potential transmission effect of monetary policy in the Economic and Monetary Union (EMU) from an empirical point of view. An extended Taylor rule to evaluate the sensitivity of the European Central Bank (ECB) to oil price fluctuations is estimated with GMM. We construct several indicators of oil prices to assess whether the effect of oil prices in the ECB interest-rate setting process is asymmetric and/or nonlinear.

Their model relies on a Taylor rule to describe the behaviour of Central Banks.

$$i_t^* = \bar{i}_t + \beta(\pi_t - \pi^*) + \gamma(y_t - y^*) \tag{10}$$

where i_t is the equilibrium nominal interest rate, π_t is the inflation rate, y_t is the output growth rate. π^* and y^* are target rate.
They added smoothing parameter ρ and change of oil price. Then, their estimate equation is

$$i_t = \alpha_1 + \alpha_2 i_{t-1} + \alpha_3(\pi_t - \pi^*) + \alpha_4(y_t - y^*) + \alpha_5 \Delta o_t + e_t \tag{11}$$

Table 1
Estimation results.

	[1]	[2]	[3]	[4]	[5]
Constant	**3.357*****	**0.144*****	0.081	0.082	0.067
	(0.089)	(0.051)	(0.049)	(0.050)	(0.067)
i_{t-1}		**0.960*****	**0.975*****	**0.964*****	**0.976*****
		(0.016)	(0.015)	(0.018)	(0.020)
$(\pi_t - \pi^*)$	**1.956*****	0.107	**0.121***	**0.176****	**0.142***
	(0.291)	(0.074)	(0.064)	(0.081)	(0.080)
$(y_t - y_t^*)$	**0.383***	**0.235*****	**0.195*****	**0.211*****	**0.203*****
	(0.220)	(0.042)	(0.037)	(0.040)	(0.055)
Δo_t			**0.0012*****		
			(0.0003)		
Δo_t^+				**0.0022****	
				(0.001)	
Δo_t^-				−0.0025	
				(0.003)	
NOPI					**0.0211****
					(0.009)
NOPD					0.0067
					(0.009)
Implied coefficients					
ρ	–	**0.960**	**0.975**	**0.964**	**0.976**
β	**1.956**	2.675	**4.840**	**4.889**	**5.917**
γ	**0.383**	**5.875**	**7.800**	**5.861**	**8.458**
λ	–	–	**0.048**	–	–
λ^+	–	–	–	**0.061**	**0.879**
λ^-	–	–	–	−0.069	0.279
Observations	117	118	118	118	118
Adjusted R2	0.117	0.980	0.982	0.979	0.976
Hansen *J*-test	20.325	3.093	7.082	5.658	5.282
P-value	[0.000]	[0.378]	[0.132]	[0.130]	[0.152]

2 GMM non linear model

Now, let us consider the general case of orthogonality condition such that

$$E[h(\theta; w)] = 0 \tag{12}$$

where θ is a $k \times 1$ vector of parameter. $h(\theta; w)$ is a $r \times 1$ vector for $r \geq k$. Let $w_i = (y_i, x_i)$ be the ith observed data. Define $g(\theta; w)$ as

$$g(\theta; w) = \frac{1}{n}\Sigma_{i=1}^{n} h(\theta; w) \tag{13}$$

where $w = w_1, w_2, ..., x_n$. In the same way as the GMM estimator in linear case, we define the GMM estimator $\hat{\theta}$, which minimizes

$$g(\theta; w)' W_n g(\theta; w) \tag{14}$$

with respect to θ. Let us consider the asymptotic distribution of GMM estimator in general case under two assumption holds, that is

$$\text{Assumption1:} \hat{\theta} \rightarrow \theta$$
$$\text{Assumption2:} \sqrt{n} g(\theta; w) \rightarrow N(0, S)$$

where

$$S = lim_{n\rightarrow\infty} V(\sqrt{n} g(\theta; w)) \tag{15}$$

The first order condition of GMM is

$$\frac{\partial g(\theta; w)'}{\partial \theta} W_n g(\theta; w) = 0 \tag{16}$$

The GMM estimator, denoted by $\hat{\theta}$, satisfies the above equation. Therefore, we have the following

$$\frac{\partial g(\hat{\theta}; w)'}{\partial \theta} W_n g\hat{\theta}; w) = 0 \tag{17}$$

Using the Theorem of Mean Value, linearized $g(\theta; w)$ around $\hat{\theta} = \theta$ can be written as follows

$$\begin{aligned} g(\hat{\theta}; w) =& g(\theta; w) + \frac{\partial g(\bar{\theta}; w)'}{\partial \theta}(\hat{\theta} - \theta) \\ =& g(\theta; w) + \bar{D}(\hat{\theta} - \theta) \end{aligned} \tag{18}$$

where $\bar{D} = \frac{\partial g(\bar{\theta}; w)'}{\partial \theta}$, and $\bar{\theta}$ is between $\hat{\theta}$ and θ. Substituting the linear approximation at $\hat{\theta} = \theta$, we obtain

$$
\begin{aligned}
0 &= \hat{D}'W_n g(\hat{\theta}; w) \\
&= \hat{D}'W_n(g(\theta; w) + \bar{D}(\hat{\theta} - \theta)) \\
&= \hat{D}'W_n g(\theta; w) + \hat{D}'W_n\bar{D}(\hat{\theta} - \theta)
\end{aligned} \tag{19}
$$

which can be written as

$$
\hat{\theta} - \theta = -(\hat{D}'W_n\bar{D})^{-1}\hat{D}'W_n g(\theta; w) \tag{20}
$$

From Assumption 1, $\hat{\theta} \to \theta$ implies $\bar{\theta} \to \theta$. Therefore,

$$
\sqrt{n}(\hat{\theta} - \theta) = -(\hat{D}'W_n\bar{D})^{-1}\hat{D}'W_n \times \sqrt{n}g(\theta; w) \tag{21}
$$

Assuming $W_n \to W$, the GMM estimator $\hat{\theta}$ has the following asymptotic distribution

$$
\sqrt{n}(\hat{\theta} - \theta) \to N(0, (D'WD)^{-1}D'WSWD(D'WD)^{-1}) \tag{22}
$$

Note that $\hat{D} \to D, \bar{D} \to D$, and assumption 2 are utilized. If we set the weighting matrix $W_n \to W = S^{-1}$, then this expression can be simplified as

$$
\sqrt{n}(\hat{\theta} - \theta) \to N(0, (D'S^{-1}D)^{-1}) \tag{23}
$$

Let us discuss how to obtain consistent estimator of S. If $h(\theta; w_i), i = 1, .., n$, are mutually independent, S is

$$
\begin{aligned}
S =& V(\sqrt{n}g(\theta; w)) = nE(g(\theta; w)g(\theta; w)') \\
=& nE((\frac{1}{n}\Sigma_{i=1}^n h(\theta; w_i))(\frac{1}{n}\Sigma_{i=1}^n h(\theta; w_i))') \\
=& \frac{1}{n}\Sigma\Sigma E(h(\theta; w_i)h(\theta; w_j)') \\
=& \frac{1}{n}\Sigma E(h(\theta; w_i)h(\theta; w_i))
\end{aligned} \tag{24}
$$

Note that

1. $E(h(\theta; w_i)) = 0$ for all i amd accordingly $E(g(\theta; w)) = 0$
2. $g(\theta; w) = \frac{1}{n}\Sigma_{i=1}^n h(\theta; w_i) = \frac{1}{n}\Sigma_{j=1}^n h(\theta; w_j)$
3. $E(h(\theta; w_i)h(\theta; w_j)') = 0$ for $i \neq j$

The estimator of S, denote by $\hat{S}$ is given by

$$
\hat{S} = \frac{1}{n}\Sigma_{i=1}^n h(\hat{\theta}; w_i)h(\hat{\theta}; w_i)' \to S \tag{25}
$$