ECON 710B - Problem Set 7

Alex von Hafften*

3/23/2021

13.1

Take the model:

$$\begin{aligned} Y &= X'\beta + e \\ E[Xe] &= 0 \\ e^2 &= Z'\gamma + \eta \\ E[Z\eta] &= 0 \end{aligned}$$

Find the method of moments estimators $(\hat{\beta}, \hat{\gamma})$ for (β, γ).

The moment conditions are:

$$\begin{aligned} \begin{pmatrix} E[Xe] \\ E[Z\eta] \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \implies \begin{pmatrix} E[X(Y - X'\beta)] \\ E[Z((Y - X'\beta)^2 - Z'\gamma)] \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \implies \begin{pmatrix} E[g_1(\beta, \gamma)] \\ E[g_2(\beta, \gamma)] \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \text{where } g_1(\beta, \gamma) &= XY - XX'\beta, \\ g_2(\beta, \gamma) &= Z(Y - X'\beta)^2 - ZZ'\gamma \end{aligned}$$

Replacing with the sample moment:

$$\frac{1}{n}\sum_{i=1}^{n}(X_iY_i - X_iX_i'\hat{\beta}) = 0 \implies \hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^{n} X_iX_i\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} X_iY_i\right)$$

$$\frac{1}{n}\sum_{i=1}^{n}(Z_i(Y_i - X_i'\hat{\beta})^2 - Z_iZ_i'\hat{\gamma}) = 0 \implies \hat{\gamma} = \left(\frac{1}{n}\sum_{i=1}^{n} Z_iZ_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i(Y_i - X_i'\hat{\beta})^2\right)$$

*I worked on this problem set with a study group of Michael Nattinger, Andrew Smith, and Ryan Mather. I also discussed problems with Sarah Bass, Emily Case, Danny Edgel, and Katherine Kwok.

13.2

Take the model $Y = X'\beta + e$ with $E[e|Z] = 0$. Let β_{gmm} be the GMM estimator using the weight matrix $W_n = (Z'Z)^{-1}$. Under the assumption $E[e^2|Z] = \sigma^2$ show that

$$\sqrt{n}(\hat{\beta}_{gmm} - \beta) \to_d N(0, \sigma^2(Q'M^{-1}Q)^{-1})$$

where $Q = E[ZX']$ and $M = E[ZZ']$.

We can rewrite $\hat{\beta}_{gmm}$ as:

$$\begin{aligned}
\hat{\beta}_{gmm} &= (X'ZW_nZ'X)^{-1}(X'ZW_nZ'Y) \\
&= (X'Z(nW_n)Z'X)^{-1}(X'Z(nW_n)Z'Y) \\
&= (X'ZV_nZ'X)^{-1}(X'ZV_nZ'Y)
\end{aligned}$$

where $V_n = (n^{-1}Z'Z)^{-1}$. Notice that

$$n^{-1}Z'Z \to_p E[Z'Z]$$

by law of large numbers, so by CMT:

$$V_n = (n^{-1}Z'Z)^{-1} \to_p E[Z'Z]^{-1} \equiv W$$

Notice that $M = W^{-1}$. If $E[e^2|Z] = \sigma^2$, then

$$\Omega = E[ZZ'e^2] = E[ZZ'E[e^2|Z]] = \sigma^2 E[ZZ'] = \sigma^2 M = \sigma^2 W^{-1}$$

By Theorem 13.3, we know that $\sqrt{n}(\hat{\beta}_{gmm} - \beta) \to_d N(0, V_\beta)$ where

$$\begin{aligned}
V_\beta &= (Q'WQ)^{-1}(Q'W\Omega WQ)(Q'WQ)^{-1} \\
&= (Q'WQ)^{-1}(Q'W\sigma^2 W^{-1}WQ)(Q'WQ)^{-1} \\
&= \sigma^2(Q'WQ)^{-1}(Q'WQ)(Q'WQ)^{-1} \\
&= \sigma^2(Q'WQ)^{-1} \\
&= \sigma^2(Q'M^{-1}Q)^{-1}
\end{aligned}$$

13.3

Take the model $Y = X'\beta + e$ with $E[Ze] = 0$. Let $\tilde{e} = Y - X'\hat{\beta}$ where $\hat{\beta}$ is consistent for β (e.g. a GMM estimator with some weight matrix). An estimator of the optimal GMM weight matrix is

$$\hat{W} = \left(\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \tilde{e}_i^2\right)^{-1}$$

Show that $\hat{W} \to_p \Omega^{-1}$ where $\Omega = E[ZZ'e^2]$.

By the weak law of large numbers and the continuous mapping theorem:

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \tilde{e}_i^2 &= \frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' (Y_i - X_i'\hat{\beta})^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' Y_i^2 - 2\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' Y_i X_i'\hat{\beta} + \frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' X_i'\hat{\beta} X_i'\hat{\beta} \\
&\to_p E[ZZ'Y^2] - 2E[ZZ'YX'\beta] + E[ZZ'X'\beta X'\beta] \\
&= E[ZZ(Y - X'\beta)^2] \\
&= E[ZZe^2]
\end{aligned}$$

Again, by the continuous mapping theorem:

$$\hat{W} = \left(\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \tilde{e}_i^2\right)^{-1} \to_p E[ZZ'e^2]^{-1}$$

13.4

In the linear model estimated by GMM with general weight matrix W the asymptotic variance of $\hat{\beta}_{gmm}$ is

$$V = (Q'WQ)^{-1}Q'W\Omega WQ(Q'WQ)^{-1}$$

(a) Let V_0 be this matrix when $W = \Omega^{-1}$. Show that $V_0 = (Q'\Omega^{-1}Q)^{-1}$.

$$\begin{aligned} V_0 &= (Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}\Omega\Omega^{-1}Q(Q'\Omega^{-1}Q)^{-1} \\ &= (Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}Q(Q'\Omega^{-1}Q)^{-1} \\ &= (Q'\Omega^{-1}Q)^{-1} \end{aligned}$$

(b) We want to show that for any W, $V - V_0$ is positive semi-definite (for then V_0 is the smaller possible covariance matrix and $W = \Omega^{-1}$ is the efficient weight matrix). To do this start by finding matrices A and B such that $V = A'\Omega A$ and $V_0 = B'\Omega B$.

$$\begin{aligned} V &= (Q'WQ)^{-1}Q'W\Omega WQ(Q'WQ)^{-1} \\ &= A'\Omega A \\ A &:= WQ(Q'WQ)^{-1} \\ A' &= (WQ(Q'WQ)^{-1})' \\ &= ((Q'WQ)')^{-1}Q'W' \\ &= (Q'WQ)^{-1}Q'W \end{aligned}$$

Since W is symmetric $\implies$ $Q'WQ$ is symmetric.

$$\begin{aligned} V_0 &= (Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}\Omega\Omega^{-1}Q(Q'\Omega^{-1}Q)^{-1} \\ &= B'\Omega B \\ B &:= \Omega^{-1}Q(Q'\Omega^{-1}Q)^{-1} \\ B' &= (\Omega^{-1}Q(Q'\Omega^{-1}Q)^{-1})' \\ &= (Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1} \end{aligned}$$

(c) Show that $B'\Omega A = B'\Omega B$ and therefore that $B'\Omega(A - B) = 0$.

$$\begin{aligned} B'\Omega A &= [(Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}]\Omega[WQ(Q'WQ)^{-1}] \\ &= (Q'\Omega^{-1}Q)^{-1}Q'WQ(Q'WQ)^{-1} \\ &= (Q'\Omega^{-1}Q)^{-1} \\ &= V_0 \\ &= (Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}\Omega\Omega^{-1}Q(Q'\Omega^{-1}Q)^{-1} \\ &= B'\Omega B \end{aligned}$$

(d) Use the expressions $V = A'\Omega A$, $A = B + (A - B)$, and $B'\Omega(A - B) = 0$ to show that $V \geq V_0$.

$$
\begin{aligned}
V &= A'\Omega A \\
&= (B + (A - B))'\Omega(B + (A - B)) \\
&= B'\Omega B + B'\Omega(A - B) + (A - B)'\Omega B + (A - B)'\Omega(A - B) \\
&= V_0 + (A - B)'\Omega(A - B)
\end{aligned}
$$

$(A - B)'\Omega(A - B)$ is positive semi-definite, so $V \geq V_0$.

13.11

As a continuation of Exercise 12.7 derive the efficient GMM estimator using the instrument $Z = (X X^2)'$. Does this differ from 2SLS and/or OLS?

The optimal weight matrix is:

$$\Omega = E[ZZ'e^2] = E\left[\begin{pmatrix} X \\ X^2 \end{pmatrix} \begin{pmatrix} X & X^2 \end{pmatrix} e^2\right] = \begin{pmatrix} E[X^2e^2] & E[X^3e^2] \\ E[X^3e^2] & E[X^4e^2] \end{pmatrix}$$

We can estimate the optimal weight matrix as:

$$\hat{\Omega} = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n X_i^2 e_i^2 & \frac{1}{n}\sum_{i=1}^n X_i^3 e_i^2 \\ \frac{1}{n}\sum_{i=1}^n X_i^3 e_i^2 & \frac{1}{n}\sum_{i=1}^n X_i^4 e_i^2 \end{pmatrix}$$

$$\hat{\Omega}^{-1} = \frac{1}{\frac{1}{n}\sum_{i=1}^n X_i^2 e_i^2 \frac{1}{n}\sum_{i=1}^n X_i^4 e_i^2 - (\frac{1}{n}\sum_{i=1}^n X_i^3 e_i^2)^2} \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n X_i^4 e_i^2 & -\frac{1}{n}\sum_{i=1}^n X_i^3 e_i^2 \\ -\frac{1}{n}\sum_{i=1}^n X_i^3 e_i^2 & \frac{1}{n}\sum_{i=1}^n X_i^2 e_i^2 \end{pmatrix}$$

The formula for the efficient GMM is:

$$\hat{\beta}_{gmm} = (X'Z\hat{\Omega}^{-1}Z'X)^{-1}(X'Z\hat{\Omega}^{-1}Z'Y) = ...$$

13.13

Take the linear model $Y = X'\beta + e$ with $E[Ze] = 0$. Consider the GMM estimator $\hat{\beta}$ of β. Let $J = n\bar{g}_n(\hat{\beta})'\hat{\Omega}^{-1}\bar{g}_n(\hat{\beta})$ denote the test of overidentifying restrictions. Show that $J \to_d \chi^2_{\ell-k}$ as $n \to \infty$ by demonstrating each of the following.

(a) Since $\Omega > 0$, we can write $\Omega^{-1} = CC'$ and $\Omega = C'^{-1}C^{-1}$ for some matrix C.

By the spectral decomposition, $\Omega = H\Lambda H'$ where $H'H = I_k$ and Λ is diagonal with strictly positive diagonal elements and thus Λ is positive definite:[1]

$$\Omega = H\Lambda H' = H\Lambda^{1/2}\Lambda^{1/2}H'$$

Notice that $\Omega^{-1} = (H\Lambda H')^{-1} = H\Lambda^{-1}H'$. Define $C := H\Lambda^{-1/2}$. Thus,

$$CC' = H\Lambda^{-1/2}(H\Lambda^{-1/2})' = H\Lambda^{-1/2}\Lambda^{-1/2}H' = H\Lambda^{-1}H' = \Omega^{-1}$$

and $\Omega = C'^{-1}C^{-1}$.

(b) $J = n(C'\bar{g}_n(\hat{\beta}))'(C'\hat{\Omega}C')^{-1}C'\bar{g}_n(\hat{\beta})$.

$$J = n\bar{g}_n(\hat{\beta})'\hat{\Omega}^{-1}\bar{g}_n(\hat{\beta}) = n\bar{g}_n(\hat{\beta})'C'C'^{-1}\hat{\Omega}^{-1}C'^{-1}C'\bar{g}_n(\hat{\beta}) = n(C'\bar{g}_n(\hat{\beta}))'(C'\hat{\Omega}C')^{-1}C'\bar{g}_n(\hat{\beta})$$

(c) $C'\bar{g}_n(\hat{\beta}) = D_nC'\bar{g}_n(\beta)$ where $\bar{g}_n(\beta) = \frac{1}{n}Z'e$ and

$$D_n = I_\ell - C'(\frac{1}{n}Z'X)((\frac{1}{n}X'Z)\hat{\Omega}^{-1}(\frac{1}{n}Z'X))^{-1}(\frac{1}{n}X'Z)\hat{\Omega}^{-1}C'^{-1}$$

$$\begin{aligned}
C'\bar{g}_n(\hat{\beta}) &= C'\frac{1}{n}Z'(Y - X'\hat{\beta}) \\
&= C'\frac{1}{n}Z'(Y - X'(X'Z\hat{\Omega}^{-1}Z'X)^{-1}(X'Z\hat{\Omega}^{-1}Z'Y)) \\
&= C'\frac{1}{n}Z'(I - X'(X'Z\hat{\Omega}^{-1}Z'X)^{-1}(X'Z\hat{\Omega}^{-1}Z'))(X'\beta + e) \\
&= D_nC'\bar{g}_n(\beta)
\end{aligned}$$

[1] By the spectral decomposition, $A = H\Lambda H'$ where $H'H = I_k$ and Λ is diagonal with non-negative diagonal elements. All diagonal elements of Λ are strictly positive iff $A > 0$ (Theorem A.4 (4) in appendix A.10 pg 944 of Hansen, Econometrics). Furthermore,

$$\Lambda^{1/2} = \begin{bmatrix} \lambda_1^{1/2} & 0 & \dots & 0 \\ 0 & \lambda_2^{1/2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_k^{1/2} \end{bmatrix} \implies \Lambda = \Lambda^{1/2}\Lambda^{1/2}$$

(d) $D_n \to_p I_\ell - R(R'R)^{-1}R'$ where $R = C'E[ZX']$.

By WLLN,

$$
\begin{aligned}
D_n &= I_\ell - C'(\frac{1}{n}Z'X)((\frac{1}{n}X'Z)\hat{\Omega}^{-1}(\frac{1}{n}Z'X))^{-1}(\frac{1}{n}X'Z)\hat{\Omega}^{-1}C'^{-1} \\
&\to_p I_\ell - C'E[Z'X](E[X'Z]\Omega^{-1}E[Z'X])^{-1}E[X'Z]\Omega^{-1}C'^{-1} \\
&= I_\ell - C'E[Z'X](E[X'Z]CC'E[Z'X])^{-1}E[X'Z]C \\
&= I_\ell - R(R'R)^{-1}R'
\end{aligned}
$$

(e) $n^{1/2}C'\bar{g}_n(\beta) \to_d u \sim N(0, I_\ell)$.

Based on CLT,

$$
\begin{aligned}
n^{1/2}C'\bar{g}_n(\beta) &= n^{1/2}C'\frac{1}{n}Z'e \\
&= C'\frac{1}{\sqrt{n}}Z'e \\
&\to_d C'N(0,\Omega) \\
&= N(0, C'\Omega C) \\
&= N(0, C'C'^{-1}C^{-1}C) \\
&= N(0, I_\ell)
\end{aligned}
$$

(f) $J \to_d u'(I_\ell - R(R'R)^{-1}R')u$.

Notice that $I_\ell - R(R'R)^{-1}R'$ is idempotent:

$$
(I_\ell - R(R'R)^{-1}R')(I_\ell - R(R'R)^{-1}R')' = I_\ell - R(R'R)^{-1}R' - R(R'R)^{-1}R' + R(R'R)^{-1}R'R(R'R)^{-1}R' = I_\ell - R(R'R)^{-1}R'
$$

Thus, by the CMT:

$$
\begin{aligned}
J &= (\sqrt{n}C'\bar{g}_n(\beta))'D_n'(C'\hat{\Omega}C')^{-1}C'D_nC'\sqrt{n}\bar{g}_n(\beta) \\
&\to_d u'(I_\ell - R(R'R)^{-1}R')'(C'\Omega C')^{-1}(I_\ell - R(R'R)^{-1}R')u \\
&= u'(I_\ell - R(R'R)^{-1}R')'(C'C'^{-1}C^{-1}C')^{-1}(I_\ell - R(R'R)^{-1}R')u \\
&= u'(I_\ell - R(R'R)^{-1}R')'(I_\ell - R(R'R)^{-1}R')u \\
&= u'(I_\ell - R(R'R)^{-1}R')u
\end{aligned}
$$

(g) $u'(I_\ell - R(R'R)^{-1}R')u \sim \xi^2_{\ell-k}$. [Hint: $I_\ell - R(R'R)^{-1}R'$ is a projection matrix.]

...

13.18

The observations are i.i.d., $(Y_i, X_i, Q_i : i = 1, ..., n)$, where X is $k \times 1$ and Q is $m \times 1$.The model is $Y = X'\beta + e$ with $E[Xe] = 0$ and $E[Qe] = 0$. Find the efficient GMM estimator for β.

Since $E[Xe] = 0$ and $E[Qe] = 0$, we can use $Z = \begin{pmatrix} X & Q \end{pmatrix}^{-1}$ as a instrument. Thus, the optimal weighting matrix is:

$$\Omega = E\left[\begin{pmatrix} X \\ Q \end{pmatrix} \begin{pmatrix} X' & Q' \end{pmatrix} e \right] = \begin{pmatrix} E[XX'e] & E[XQ'e] \\ E[QX'e] & E[QQ'e] \end{pmatrix}$$

A consistent estimator for Ω is:

$$\hat{\Omega} = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} X_i X_i' e_i & \frac{1}{n}\sum_{i=1}^{n} X_i Q_i' e_i \\ \frac{1}{n}\sum_{i=1}^{n} Q_i X_i' e_i & \frac{1}{n}\sum_{i=1}^{n} Q_i Q_i' e_i \end{pmatrix}$$

The efficient GMM estimator:

$$\hat{\beta} = (X'Z\hat{\Omega}^{-1}Z'X)^{-1}X'Z\hat{\Omega}^{-1}Z'Y$$

13.19

You want to estimate $\mu = E[Y]$ under the assumption that $E[X] = 0$, where Y and X are scalar and observed from a random sample. Find an efficient GMM estimator for μ.

We have two moment conditions:

$$\begin{pmatrix} E[Y-\mu] \\ E[X] \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} E[g_1(\mu)] \\ E[g_2(\mu)] \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

where $g_1(\mu) = Y - \mu$ and $g_2(\mu) = X$. Therefore,

$$g_i(\mu) = \begin{pmatrix} Y_i - \mu \\ X_i \end{pmatrix}$$

$$\bar{g}_n(\mu) = \begin{pmatrix} \bar{Y} - \mu \\ \bar{X} \end{pmatrix}$$

The optimal weighting matrix is $W = \Omega^{-1}$ where:

$$\Omega = E\left[\begin{pmatrix} Y-\mu \\ X \end{pmatrix} \begin{pmatrix} Y-\mu & X \end{pmatrix} \right] = \begin{pmatrix} Var(Y) & Cov(Y,X) \\ Cov(Y,X) & Var(X) \end{pmatrix}$$

$$\Omega^{-1} = \frac{1}{Var(Y)Var(X) - Cov(Y,X)^2} \begin{pmatrix} Var(X) & -Cov(Y,X) \\ -Cov(Y,X) & Var(Y) \end{pmatrix}$$

The efficient GMM estimator minimizes the following:

$$\begin{aligned}
J(\mu) &= \bar{g}_n(\mu)' \Omega^{-1} \bar{g}_n(\mu) \\
&= \begin{pmatrix} \bar{Y} - \mu & \bar{X} \end{pmatrix} \frac{1}{Var(Y)Var(X) - Cov(Y,X)^2} \begin{pmatrix} Var(X) & -Cov(Y,X) \\ -Cov(Y,X) & Var(Y) \end{pmatrix} \begin{pmatrix} \bar{Y} - \mu \\ \bar{X} \end{pmatrix} \\
&= \frac{1}{Var(Y)Var(X) - Cov(Y,X)^2} \begin{pmatrix} (\bar{Y}-\mu)Var(X) - \bar{X}Cov(Y,X) & -(\bar{Y}-\mu)Cov(Y,X) + \bar{X}Var(Y) \end{pmatrix} \begin{pmatrix} \bar{Y} - \mu \\ \bar{X} \end{pmatrix} \\
&= \frac{Var(X)(\bar{Y}-\mu)^2 - 2Cov(X,Y)\bar{X}(\bar{Y}-\mu) + Var(Y)\bar{X}^2}{Var(Y)Var(X) - Cov(Y,X)^2}
\end{aligned}$$

FOC of $J(\hat{\mu})$:

$$\begin{aligned}
\frac{-2Var(X)(\bar{Y}-\hat{\mu}) + 2Cov(X,Y)\bar{X}}{Var(Y)Var(X) - Cov(Y,X)^2} &= 0 \\
\implies Var(X)(\bar{Y}-\hat{\mu}) &= Cov(X,Y)\bar{X} \\
\implies \hat{\mu} &= \bar{Y} - \frac{Cov(X,Y)}{Var(X)}\bar{X}
\end{aligned}$$

Replace $Cov(X,Y)$ and $Var(X)$ with estimators:

$$\hat{\mu} = \bar{Y} - \frac{\hat{Cov}(X,Y)}{\hat{Var}(X)}\bar{X}$$

13.28

Continuation of Exercise 12.25, which involved estimation of a wage equation by 2SLS.

(a) Re-estimate the model in part (a) by efficient GMM. Do the results change meaningfully?

. . .

(b) Re-estimate the model in part (d) by efficient GMM. Do the results change meaningfully?

. . .

(c) Report the J statistic for overidentification.

. . .

17.15

In this exercise you will replicate and extend the empirical work reported in Arellano and Bond (1991) and Blundell and Bond (1998). Arellano-Bond gathered a dataset of 1031 observations from an unbalanced panel of 140 U.K. companies for 1976-1984 and is in the datafile `AB1991` on the textbook webpage. The variables we will be using are log employment (N), log real wages (W), and log capital (K). See the description file for definitions.

(a) Estimate the panel AR(1) $K_{it} = \alpha K_{it-1} + u_i + v_t + \varepsilon_{it}$ using Arellano-Bondone-step GMM with clustered standard errors. Note that the model includes year fixed effects.

...

(b) Re-estimate using Blundell-Bondone-step GMM with clustered standard errors.

...

(c) Explain the difference in the estimates. ...