Problem Set 6

Artschil Okropiridse*

March 14, 2024

Lecture 12 Generalized methods of moments

Exercise 1

(Ch. 13, ex. 13.6) Derive the constrained GMM estimator (13.16).

Solution

We begin by defining the moment condition

$$g(\beta) = Z(Y - X'\beta)$$
$$\mathbb{E}[Z(Y - X'\beta)] = 0$$

In the just-identified case we would estimate this by setting the following expression to zero and backing out β.

$$\bar{g}_n(\beta) = \frac{1}{n}(\mathbf{Z}'\mathbf{Y} - \mathbf{Z}'\mathbf{X}\beta)$$

In the more general GMM case, we can't do that. As there is more equations then unknowns. So we define the GMM criterion function

$$J(\beta) = \frac{1}{2}(\mathbf{Z}'\mathbf{Y} - \mathbf{Z}'\mathbf{X}\beta)'\mathbf{W}(\mathbf{Z}'\mathbf{Y} - \mathbf{Z}'\mathbf{X}\beta)$$

Note that pre-multiplying by $\frac{1}{2}$ does not change the min problem. But reduces the notataional clutter by a bit.

The constraint GMM estimator then solves

$$\min J(\beta)$$
$$\text{s.t. } \mathbf{R}'\beta = c.$$

*many thanks to Jakob Beuschlein

We can solve this using the Lagrangian

$$\begin{aligned}\mathcal{L} &= J(\beta) + \lambda'(\mathbf{R}'\beta - c)\\ &= n(\mathbf{Z'Y} - \mathbf{Z'X}\beta)'\mathbf{W}(\mathbf{Z'Y} - \mathbf{Z'X}\beta) + \lambda'(\mathbf{R}'\beta - c)\end{aligned}$$

The derivative wrt λ yields

$$\mathbf{R}'\beta = c$$

and differentiating wrt β gives us

$$\begin{aligned}\frac{\partial}{\partial\beta}\mathcal{L} = (\mathbf{X'ZWZ'}(\mathbf{X}\widehat{\beta}^{\mathbf{cGMM}} - \mathbf{Y})) + \mathbf{R}\lambda = 0\\ (\mathbf{X'ZWZ'X}\widehat{\beta}^{\mathbf{cGMM}}) + \mathbf{R}\lambda = \mathbf{X'ZWZ'Y}\end{aligned}$$

Then multiply both sides by $\mathbf{R}'(\mathbf{X'ZWZ'X})^{-1}$ to get

$$\begin{aligned}\mathbf{R}'(\mathbf{X'ZWZ'X})^{-1}(\mathbf{X'ZWZ'X}\beta) + \mathbf{R}'(\mathbf{X'ZWZ'X})^{-1}\mathbf{R}\lambda =\\ \mathbf{R}'(\mathbf{X'ZWZ'X})^{-1}\mathbf{X'ZWZ'Y}\\ \implies \mathbf{R}'\beta + \mathbf{R}'(\mathbf{X'ZWZ'X})^{-1}\mathbf{R}\lambda = \mathbf{R}'\widehat{\beta}^{GMM}\\ \implies \mathbf{R}'(\mathbf{X'ZWZ'X})^{-1}\mathbf{R}\lambda = \mathbf{R}'\widehat{\beta}^{GMM} - \mathbf{c}\\ \implies \lambda = \left(\mathbf{R}'(\mathbf{X'ZWZ'X})^{-1}\mathbf{R}\right)^{-1}(\mathbf{R}'\widehat{\beta}^{GMM} - c)\end{aligned}$$

Lastly, we put this back in the FOC and solve for $\widehat{\beta}^{cGMM}$:

$$\begin{aligned}(\mathbf{X'ZWZ'X}\widehat{\beta}^{\mathbf{cGMM}}) + \mathbf{R}\left(\mathbf{R}'(\mathbf{X'ZWZ'X})^{-1}\mathbf{R}\right)^{-1}(\mathbf{R}'\widehat{\beta}^{GMM} - c) = \mathbf{X'ZWZ'Y}\\ \widehat{\beta}^{cGMM} = (\mathbf{X'ZWZ'X})^{-1}(\mathbf{X'ZWZ'Y})\dots\\ \widehat{\beta}^{cGMM} = \widehat{\beta}^{GMM} - (\mathbf{X'ZWZ'X})^{-\mathbf{1}}\mathbf{R}(\mathbf{R}'(\mathbf{X'ZWZ'X})^{-\mathbf{1}}\mathbf{R})(\mathbf{R}'\widehat{\beta}^{\mathbf{GMM}} - \mathbf{c})\end{aligned}$$

We are done!

Exercise 2

(Ch. 13, ex. 13.12) In the linear model $Y = X'\beta + e$ with $\mathbb{E}[Xe] = 0$, the GMM criterion function for β is

$$J(\beta) = \frac{1}{n}(\mathbf{Y} - \mathbf{X}\beta)'\mathbf{X}\hat{\mathbf{\Omega}}^{-1}\mathbf{X}'(\mathbf{Y} - \mathbf{X}\beta) \tag{*}$$

where $\hat{\mathbf{\Omega}} = n^{-1}\sum_i X_i X_i'\hat{e}_i^2,\ \hat{e}_i^2 = y_i - X_i'\beta$ are the OLS residuals, and $\hat{\beta} = (\mathbf{X'X})^{-1}\mathbf{X'Y}$ is the least squares estimator. The GMM estimator for β subject to the restriction $r(\beta) = 0$ is

$$\tilde{\beta} = \arg\min_{r(\beta)=0} J_n(\beta).$$

The GMM test statistic (the distance statistic) of the hypothesis $r(\beta) = 0$ is

$$D = J(\tilde{\beta}) = \min_{r(\beta)=0} J(\beta). \tag{+}$$

(a) Show that you can rewrite $J(\beta)$ in (*) as

$$J(\beta) = n(\hat{\beta} - \beta)'\hat{\mathbf{V}}_{\beta}^{-1}(\hat{\beta} - \beta)$$

and thus $\tilde{\beta}$ is the same as the minimum distance estimator.

(b) Show that under linear hypotheses the distance statistic D in (+) equals the Wald statistic.

Solution

(a) First note that $\mathbf{Y} - \mathbf{X}\beta = \mathbf{e}$ (this follows given the model in the question), so we can plug into (*):

$$\begin{aligned} J(\beta) &= \frac{1}{n}(\mathbf{Y} - \mathbf{X}\beta)'\mathbf{X}\hat{\mathbf{\Omega}}^{-1}\mathbf{X}'(\mathbf{Y} - \mathbf{X}\beta) \\ &= \frac{1}{n}\mathbf{e}'\mathbf{X}\hat{\mathbf{\Omega}}^{-1}\mathbf{X}'e \\ &= \frac{1}{n}e'\mathbf{X}(\tfrac{1}{n}\mathbf{X}'\mathbf{X})^{-1}(\tfrac{1}{n}\mathbf{X}'\mathbf{X})\hat{\mathbf{\Omega}}^{-1}(\tfrac{1}{n}\mathbf{X}'\mathbf{X})(\tfrac{1}{n}\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'e. \end{aligned}$$

In the last equality we have multiplied by $\mathbf{X}'\mathbf{X}$ and its inverse, on both sides. Then note that we know from the simple OLS variance:

$$\hat{\mathbf{V}}_{\beta} = (\tfrac{1}{n}\mathbf{X}'\mathbf{X})^{-1}\hat{\mathbf{\Omega}}(\tfrac{1}{n}\mathbf{X}'\mathbf{X})^{-1}$$

and therefore

$$\frac{1}{n}\mathbf{e}'\mathbf{X}(\tfrac{1}{n}\mathbf{X}'\mathbf{X})^{-1}(\tfrac{1}{n}\mathbf{X}'\mathbf{X})\hat{\mathbf{\Omega}}^{-1}(\tfrac{1}{n}\mathbf{X}'\mathbf{X})(\tfrac{1}{n}\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e} = n((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e})'\hat{\mathbf{V}}_{\beta}^{-1}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}).$$

Finally,

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{Y} - \mathbf{X}'\beta) = \hat{\beta} - \beta.$$

So we get that

$$\frac{1}{n}(\mathbf{Y} - \mathbf{X}'\beta)'\mathbf{X}\hat{\mathbf{\Omega}}^{-1}\mathbf{X}'(\mathbf{Y} - \mathbf{X}'\beta) = n(\hat{\beta} - \beta)'\hat{\mathbf{V}}_{\beta}^{-1}(\hat{\beta} - \beta).$$

(b) Under the linear restriction $\mathbf{R}'\beta = \mathbf{0}$ the restricted feasible efficient GMM estimator is given by

$$\tilde{\beta} = \hat{\beta} - \hat{\mathbf{V}}_{\beta}\mathbf{R}(\mathbf{R}'\hat{\mathbf{V}}_{\beta}\mathbf{R})^{-1}(\mathbf{R}'\beta - \mathbf{0}).$$

which you can show by solving $\tilde{\beta} = \arg\min_{\mathbf{R}'\beta=0} J(\tilde{\beta})$. Plugging this into (+) gives us

$$\begin{aligned}
D &= n(\hat{\mathbf{V}}_\beta \mathbf{R}(\mathbf{R}'\hat{\mathbf{V}}_\beta \mathbf{R})^{-1}\mathbf{R}'\beta)'\hat{\mathbf{V}}_\beta^{-1}(\hat{\mathbf{V}}_\beta \mathbf{R}(\mathbf{R}'\hat{\mathbf{V}}_\beta \mathbf{R})^{-1}\mathbf{R}'\beta) \\
&= n(\hat{\mathbf{V}}_\beta \mathbf{R}(\mathbf{R}'\hat{\mathbf{V}}_\beta \mathbf{R})^{-1}\mathbf{R}'\beta)'\mathbf{R}(\mathbf{R}'\hat{\mathbf{V}}_\beta \mathbf{R})^{-1}\mathbf{R}'\beta \\
&= n\beta'\mathbf{R}(\mathbf{R}'\hat{\mathbf{V}}_\beta \mathbf{R})^{-1}\mathbf{R}'\hat{\mathbf{V}}_\beta \mathbf{R}(\mathbf{R}'\hat{\mathbf{V}}_\beta \mathbf{R})^{-1}\mathbf{R}'\beta \\
&= n(\mathbf{R}'\beta)'(\mathbf{R}'\hat{\mathbf{V}}_\beta \mathbf{R})^{-1}(\mathbf{R}'\beta) = W_n
\end{aligned}$$

Exercise 3

(Ch. 13, ex. 13.27) Continuation of Exercise 12.22, based on the empirical work reported in Acemoglu, Johnson, and Robinson (2001).

(a) Re-estimate the model in part (j) by efficient GMM. se the 2SLS estimates as the first step for the weight matrix and then calculate the GMM estimator using the weight matrix without further iteration. Report the estimates.

(b) Calculate and report the J statistic for overidentification.

(c) Compare the GMM and 2SLS estimates. Discuss your findings.

Solution

R solution at the end of the document.

Lecture 13 Time series

Exercise 4

(Ch. 14, ex. 14.8) Suppose $y_t = y_{t-1} + e_t$ with e_t iid $N(0, 1)$ and $y_0 = 0$. Find $\text{Var}[y_t]$. Is y_t stationary?

Solution

Let's take expectations:

$$\begin{aligned}
\mathbb{E}[y_t] &= \mathbb{E}[y_{t-1}] + 0 \\
&= \mathbb{E}[y_{t-2}] \\
&= \dots \\
&= \mathbb{E}[y_0] = 0
\end{aligned}$$

Now for the the variance we get:

$$\begin{aligned}
\text{Var}[y_t] &= \mathbb{E}\left[(y_t - \mathbb{E}[y_t])^2\right] \\
&= \mathbb{E}\left[y_t^2\right] \\
&= \mathbb{E}\left[(y_{t-1} + e_t)^2\right] \\
&= \mathbb{E}\left[(y_{t-2} + e_{t-1} + e_t)^2\right] \\
&= \dots \\
&= \mathbb{E}\left[(y_0 + e_1 + ... + e_t)^2\right] \\
&= \mathbb{E}\left[(\sum^t e_i)^2\right] \\
&= t
\end{aligned}$$

Note that we have no cross correlations given iid errors.

This implies that the distribution of y_t increases as t increases. So the series is NOT stationary.

Exercise 5

(Ch. 14, ex. 14.9) Take the AR(1) model with no intercept $y_t = \alpha_1 y_{t-1} + e_t$.

(a) Find the impulse response function $b_j = \frac{\partial}{\partial e_t} y_{t+j}$.

(b) Let $\hat{\alpha}_1$ be the least squares estimator of α_1. Find an estimator for b_j.

(c) Let $s(\hat{\alpha}_1)$ be s standard error for $\hat{\alpha}_1$. Use the delta method to find a 95% asymptotic confidence interval for b_j.

Solution

(a) The IRF is the change in y_{t+j} due to a shock at time t - so $\frac{\partial y_{t+j}}{\partial e_t} = b_j$. We can approach this problem iteratively. For $j = 0$ we get

$$b_0 = \frac{\partial}{\partial e_t} y_t = \frac{\partial}{\partial e_t} (\alpha_1 y_{t-1} + e_t) = 1.$$

Then, since

$$y_{t+1} = \alpha_1 y_t + e_{t+1} = \alpha_1^2 y_{t-1} + \alpha_1 e_t + e_{t+1}$$

$$b_1 = \frac{\partial}{\partial e_t} y_{t+1} = \alpha_1.$$

Generally,

$$y_{t+j} = \alpha_1^{j+1} y_t + \alpha_1^j e_t + \alpha_1^{j-1} e_{t+1} + ... + e_{t+j}$$

and hence

$$b_j = \alpha_1^j.$$

(b) We can simply use the plug-in estimator $\hat{b}_j = (\hat{a}_1)^j$.

(c) The derivative of $g_j(x) = x^j$ is give by $g_j'(x) = jx^{j-1} = \mathbf{G}$. Therefore, by the delta method

$$\begin{aligned} \sqrt{n}(\hat{b}_j - b_j) &\to^d \mathbf{G} * N(0, s^2) \\ &\to^d \mathbf{N}(0, G^2 s^2) \\ &\to^d N\left(0, (jx^{j-1})^2 s^2(\hat{a}_1)\right)\right) \end{aligned}$$

as $n \to \infty$. The 95 % CI is then given by

$$\text{CI} = \left[\hat{b}_j - 1.96 \times j|\hat{a}_1^{j-1}|s(\hat{a}_1), \hat{b}_j + 1.96 \times j|\hat{a}_1^{j-1}|s(\hat{a}_1)\right]$$

Exercise 6

(Ch. 14, ex. 14.18) Take the quarterly series *pnfix* (nonresidential real private fixed investment) from FRED-Q.

(a) Transform the seres into quarterly growth rates.

(b) Estimate an AR(4) model. Report using heteroskedastic-consistent standard errors.

(c) Repeat using Newey-West standard errors, using $M = 5$.

(d) Comment on the magnitude and interpretation of the coefficients.

(e) Calculate (numerically) the impulse response function for $j = 1, ..., 10$.

Solution

R solution at the end of the document

Metrics1_PS7

Artschil Okropiridse

Exercise 3

```
# Prepare data
data <- read.delim("Data/AJR2001.txt")

data$logm2 <- data$logmort0^2
x <- as.matrix(cbind(1,data$risk))
z <- as.matrix(cbind(1,data$logmort0,data$logm2))
y <- as.matrix(data$loggdp)
```

a)

```
# Use 2SLS to get fitted values
iv.a <- ivreg(loggdp ~ risk| logmort0 + logm2, data=data)

# Initial matrices
Omega.hat <- matrix(0, ncol = 3, nrow = 3)
Q.hat <- matrix(0, ncol = 2, nrow = 3)

n <- nrow(y)
for (i in 1:n) {
  e.tilde <- iv.a$residuals[i]
  g.tilde <- as.matrix(z[i,] * e.tilde)
  Omega.hat <- Omega.hat + 1/n * g.tilde %*% t(g.tilde)
  Q.hat <- Q.hat + 1/n * z[i,] %*% t(x[i,])
}

w.hat <- solve(Omega.hat)
beta_gmm <- solve(t(x)%*%z%*%w.hat%*%t(z)%*%x) %*% t(x)%*%z%*%w.hat%*%t(z)%*%y
```

b)

```
# Calculate J statistic
g.bar <- 1/n*(t(z)%*%y - t(z)%*%x%*%beta_gmm)
J <- n*t(g.bar)%*%w.hat%*%g.bar
J
```

```
         [,1]
[1,] 3.747879
```

c)

```
# 2SLS
iv.a$coefficients
```

```
(Intercept)        risk
  3.0188488   0.7722554
```

```
beta_gmm
```

```
          [,1]
[1,] 3.3362346
[2,] 0.7278459
```

The coefficients are very similar. So there is really not all to much to say.

Exercise 6

a) Transform the seres into quarterly growth rates.

```
# Prepare data
data <- read.table("Data/FRED-QD.txt", header=TRUE)['pnfix']

data <- data |>
  mutate(y = pnfix/lag(pnfix, n = 1L) - 1) |>
  mutate(L1.y = lag(y, n = 1L)) |>
```

```r
  mutate(L2.y = lag(L1.y, n = 1L)) |>
  mutate(L3.y = lag(L2.y, n = 1L)) |>
  mutate(L4.y = lag(L3.y, n = 1L)) |>
  na.omit(data) # Drop missing values (e.g. L1 is missing in t = 1)
```

b) Estimate an AR(4) model. Report using heteroskedastic-consistent standard errors.

Under the assumption of no correlation between error terms of different time periods, we can estimate the AR(4) model using the toolbox we already know from cross-sectional data.

```r
# Estimate model via OLS
ols <- lm(y ~ L1.y + L2.y + L3.y + L4.y, data)

# Adj. heteroskedasticity robust variance estimator
v.hc1 <- vcovHC(ols, type="HC1")
se.hc1 <- sqrt(diag(v.hc1))

coeftest(ols, vcov = v.hc1)
```

```
t test of coefficients:

              Estimate Std. Error t value  Pr(>|t|)
(Intercept)  0.0049078  0.0014744  3.3287  0.001019 **
L1.y         0.5005001  0.0765264  6.5402 4.049e-10 ***
L2.y         0.1696361  0.0712133  2.3821  0.018044 *
L3.y        -0.0206539  0.0634170 -0.3257  0.744965
L4.y        -0.0734205  0.0523918 -1.4014  0.162474
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

c)

Since in many settings it is unlikely that that there is actually no autocorrelation in the error terms, our estimator for the asymptotic covariance matrix $\Omega = E[X_t X_t' e_t^2]$ may not be consistent. Instead under general dependence the correct asymptotic covariance matrix is given by

$$\Omega = \sum_{l=-\infty}^{\infty} E[X_{t-l} X_t' e_t e_{t-l}]$$

Note that under stationarity this converges to a finite object.

Let $\Gamma(l) = E[X_{t-l}X_t' e_t e_{t-l}]$ and $\hat{\Gamma}(l)$ an estimator for $\Gamma(l)$. Then the Newey-West estimator for Ω is given by

$$\hat{\Omega}_M = \sum_{l=-M}^{M} \left(1 - \frac{|l|}{M+1}\right) \hat{\Gamma}(l)$$

where we use a finite number of lags to approximate the full dependence structure. Note that if $M = 0$ we get the standard heteroskedasticity-robust estimator.

```
# Newey-West variance estimator with M = 0
v.nw.0 <- NeweyWest(ols, lag = 0, prewhite = F, adjust = T)
se.hac.0 <- sqrt(diag(v.nw.0))

# This should equal to HC1 Se's
se.hc1
```

```
(Intercept)        L1.y        L2.y        L3.y        L4.y
 0.00147439  0.07652636  0.07121333  0.06341703  0.05239180
```

```
se.hac.0
```

```
(Intercept)        L1.y        L2.y        L3.y        L4.y
 0.00147439  0.07652636  0.07121333  0.06341703  0.05239180
```

```
# Newey-West variance estimator with M = 5
v.nw.5 <- NeweyWest(ols, lag = 5, prewhite = F, adjust = T)
se.hac.5 <- sqrt(diag(v.nw.5))

# Compare results
se.hc1
```

```
(Intercept)        L1.y        L2.y        L3.y        L4.y
 0.00147439  0.07652636  0.07121333  0.06341703  0.05239180
```

```
se.hac.5
```

```
(Intercept)        L1.y        L2.y        L3.y        L4.y
0.001424423 0.085158514 0.071793930 0.066592649 0.051419961
```

d)

The standard errors seem to be fairly similar indicating that autocorrelation of the error terms does not play an important role here.

e) The impulse response function is the derivative of the outcome at point $t + j$ with respect to shocks ("innovations") to the error term in period t

$$\text{IRF}_j = \frac{\partial y_{t+j}}{\partial e_t}.$$

In the model we consider, naturally $\text{IRF}_0 = 1$. If we write down our model for y_{t+1}

$$y_{t+1} = \alpha_0 + \alpha_1 y_t + \alpha_2 y_{t-1} + \alpha_3 y_{t-2} + \alpha_4 y_{t-3} + e_{t+1}$$

and note that y_{t-k} is independent of e_t for all strictly positive k we can see that

$$\text{IRF}_1 = \alpha_1 \frac{\partial y_t}{\partial e_t} = \alpha_1 \text{IRF}_0 = \alpha_1.$$

For $j = 2$ we get

$$\text{IRF}_2 = \alpha_1 \frac{\partial y_{t+1}}{\partial e_t} + \alpha_2 \frac{\partial y_t}{\partial e_t} = \alpha_1^2 + \alpha_2.$$

Following this logic we get that

$$\text{IRF}_j = \alpha_1 \text{IRF}_{j-1} + \alpha_2 \text{IRF}_{j-2} + \alpha_3 \text{IRF}_{j-3} + \alpha_4 \text{IRF}_{j-4}$$

where $\text{IRF}_{j-k} = 0$ if $k > j$.

```
# Get coefficients from OLS
alpha <- coef(ols)

# Manually set IRF_k = 0 for k < 1 and IRF_1 = 1
IRF <- c(0,0,0,1,rep(NA,10))
for (j in 5:14) {
  IRF[j] <- alpha[2]*IRF[j-1] + alpha[3]*IRF[j-2] + alpha[4]*IRF[j-3] + alpha[5]*IRF[j-4]
}
IRF <- IRF[5:14]
```

```r
# Plot results
j <- seq(1:10)
plot(j,IRF,type='b')
```