Introduction to Econometrics 2: Recitation 7

Seung-hun Lee

Columbia University

March 4th, 2020

Regressing with an overall intercept

- Consider the following model

$$y_{it} = \mu + \rho y_{it-1} + x_{it}'\beta + \alpha_i + u_{it}$$

  where μ term represents an 'overall' constant
- When we estimated ρ and β before, we relied on first-differencing at the expense of losing time-invariant terms
- How do we find μ?
  1. Two-step Approach
  2. Simultaneous Approach

  In both cases, assume $E[\alpha_i] = 0$ (if not, demeaning required)

$$y_{it} = \underbrace{\mu + E(\alpha_i)}_{\mu_0} + \rho y_{it-1} + x_{it}'\beta + \underbrace{\alpha_i - E(\alpha_i)}_{=\alpha_{0i}} + u_{it}$$

Regressing with an overall intercept: Two-step

- We first obtain the estimate $\hat{\rho}$ and $\hat{\beta}$ using an AB estimator.
- Then, we can rearrange the equation into

$$y_{it} - \hat{\rho} y_{it-1} - x_{it}' \hat{\beta} = \mu + \alpha_i + u_{it}$$

- Given that the mean of α_i and u_{it} is zero, we can use the sample mean of the left-hand-side to obtain the estimate for μ
- Therefore,

$$\hat{\mu} = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it} - \hat{\rho} y_{it-1} - x_{it}' \hat{\beta} \right)$$

Regressing with an overall intercept: Simultaneous

- Note that $\alpha_i + u_{it}$ has a zero mean
- Define $\mathbf{y}_i, \mathbf{y}_{i,-1}$ and $\mathbf{x}_i$ as stacked-up vector for i with T observations.
- We obtain

$$\begin{aligned}\mathbf{y}_i =& 1_T\mu + \rho\mathbf{y}_{i,-1} + \mathbf{x}_i\beta + \underbrace{1_T\alpha_i + \mathbf{u}_i}_{=\mathbf{v}_i} \\ =& 1_T\mu + \mathbf{w}_i\delta + \mathbf{v}_i\end{aligned}$$

- We make use of the following moment condition

$$E[\mathbf{v}_i] = E[1_T\alpha_i + \mathbf{u}_i] = E[\mathbf{y}_i - \mathbf{w}_i\delta - 1_T\mu] = 0$$

- Along with the $E[Z_i'\Delta\mathbf{u}_i] = 0$ condition, we can jointly write

$$\begin{bmatrix} E(\mathbf{v}_i) \\ E(Z_i'\Delta\mathbf{u}_i) \end{bmatrix} = E\left[\begin{bmatrix} I_T & 0 \\ 0 & Z_i \end{bmatrix}' \begin{bmatrix} \mathbf{v}_i \\ \Delta\mathbf{u}_i \end{bmatrix}\right]$$

Regressing with an overall intercept: Simultaneous

- Define $\bar{Z}_i = \begin{bmatrix} I_T & 0 \\ 0 & Z_i \end{bmatrix}$, $\bar{W}_i = \begin{bmatrix} I_T & \mathbf{w}_i \\ 0 & \Delta \mathbf{w}_i \end{bmatrix}$, $\bar{Y}_i = \begin{bmatrix} \mathbf{y}_i \\ \Delta \mathbf{y}_i \end{bmatrix}$. Then $\begin{bmatrix} \mathbf{v}_i \\ \Delta \mathbf{u}_i \end{bmatrix} = \bar{Y}_i - \bar{W} \begin{bmatrix} \mu \\ \delta \end{bmatrix}$.
- The moment condition and the resulting estimator for $\begin{bmatrix} \mu \\ \delta \end{bmatrix}$ is ($\bar{V}_n$ is the weighting matrix)

$$E \left[\bar{Z}_i' (\bar{Y}_i - \bar{W}_i) \begin{bmatrix} \mu \\ \delta \end{bmatrix} \right] = 0$$

$$\begin{bmatrix} \hat{\mu} \\ \hat{\delta} \end{bmatrix} = \left[\sum_i (\bar{W}_i' \bar{Z}_i) \bar{V}_n \sum_i \bar{Z}_i' \bar{W}_i \right]^{-1} \sum_i (\bar{W}_i' \bar{Z}_i) \bar{V}_n \sum_i \bar{Z}_i' \bar{Y}_i$$

Subset of Regressors Uncorrelated with FE

- Return to the model where

$$y_{it} = \mu + \rho y_{it-1} + x_{it}'\beta + \alpha_i + u_{it}$$

- Now that a subset of x_{it} is exogenous with respect to α_i.
- Also assume that regressors are strictly exogenous
- Denote such variable as $x_{it}^{(1)}$. Then the following holds

$$E[x_{is}^{(1)}\alpha_i] = 0 \implies E[x_{is}^{(1)}v_{it}] = 0 \implies E[x_{is}^{(1)}(y_{it} - w_{it}'\delta - \mu)] = 0 \, \forall s, t$$

- Define $h_i = vec(x_i^{(1)}) = \begin{bmatrix} x_{i1}^{(1)} \\ ... \\ x_{iT}^{(1)} \end{bmatrix}$. Then the above moment condition is equivalent to

$$E[(I_T \otimes h_i)\mathbf{v}_i] = 0$$

Subset of Regressors Uncorrelated with FE

- We can also make use of the fact that $E[\mathbf{v}_i] = 0$, which is $E[(I_T \otimes 1)\mathbf{v}_i] = 0$.
- This would give us the required moment condition for equations in levels (to back out μ), expressed as

$$E\left[\left(I_T \otimes \begin{bmatrix} 1 \\ h_i \end{bmatrix}\right)\mathbf{v}_i\right] = 0$$

- Another set of required moment condition comes from Z_i matrix defined for strictly exogenous case with $\mathbf{x}'_{iT}$ replaced with $vec(x_i^{(2)})'$.
- This gives us $E[Z_i'\Delta\mathbf{u}_i] = 0$. The GMM estimator has the same expression as in the previous case but with I_T in $\bar{Z}_i$ replaced with $\left(I_T \otimes \begin{bmatrix} 1 \\ h_i \end{bmatrix}\right)$.

Nonlinear Moments

- Consider the model

$$y_{it} = \rho y_{it-1} + \underbrace{\alpha_i + u_{it}}_{=v_{it}}$$

- The assumptions on our variables were
  - $E(\alpha_i) = 0, E(u_{it}) = 0, E(\alpha_i u_{it}) = 0 \ \forall i, t$
  - $E(u_{it} u_{is}) = 0 \ t \neq s$
  - $E(y_{i0} u_{it}) = 0 \ \forall i$ and $t = 1, .., T$
- These three assumptions imply $E[\Delta v_{it-1} v_{it}] = 0 \ t = 3, .., T$.
  - Note that $\Delta v_{it-1} = u_{it-1} - u_{it-2} = \Delta u_{it-1}$, since $v_{it} = \alpha_i + u_{it}$.
  - Since fixed effects is not correlated with the error terms and there is no serial correlation, $E[\Delta v_{it-1} v_{it}] = 0$ holds. In other words,

$$E[(\Delta y_{it-1} - \rho \Delta y_{it-2})(y_{it} - \rho y_{it-1})] = 0$$

Nonlinear Moments

- The sample analogue of this would be

$$\frac{1}{n}\sum_{i=1}^{n}(\Delta y_{it-1}y_{it} - \rho\Delta y_{it-1}y_{it-1} - \rho\Delta y_{it-2}y_{it} + \rho^2\Delta y_{it-2}y_{it-1})$$

- Notice the ρ^2 term here. Because of this, the moment condition becomes nonlinear (quadratic, to be exact).
- If GMM estimation is used, the objective function involves ρ^4 and FOC would involve ρ^3. So it is tricky to work with.

Nonlinear Moments

- In addition, assume that there is a time series homoskedasticity, or $E[u_{it}^2] = \sigma_u^2$.
- Then the following $T-1$ additional moments

$$E[v_{it}^2] - E[v_{it-1}^2] = 0, \ \ t = 2, .., T$$
$$\iff E[(y_{it} - \rho y_{it-1})^2] - E[(y_{it-1} - \rho y_{it-2})^2] = 0$$

- Similar to the previous case, the objective function involves ρ^4 and FOC has ρ^3.
- This is again, tricky to work with.
- There is one way to 'linearize' the moment conditions.

Mean Stationarity

- Mean stationarity refers to the situation where the mean of a variable is time-invariant. One case where this can hold is as follows. Suppose

$$y_{i0} = \frac{\alpha_i}{1-\rho} + e_{i0}$$

- Then $E[y_{i0}|\alpha_i] = \frac{\alpha_i}{1-\rho}$.
- Under the data generating process

$$y_{it} = \rho y_{it-1} + \alpha_i + u_{it}$$

we can show that means stationarity holds

$$\begin{aligned} E[y_{i1}|\alpha_i] &= E[\rho y_{i0} + \alpha_i + u_{i1}|\alpha_i] \\ &= \rho E[y_{i0}|\alpha_i] + \alpha_i + E[u_{i1}|\alpha_i] \\ &= \frac{\rho\alpha_i}{1-\rho} + \alpha_i = \frac{\alpha_i}{1-\rho} \end{aligned}$$

- We can reiterate and get the same result for $E[y_{it}|\alpha_i]$ for $t = 1, ..., T$

Mean Stationarity: Linearizing Moment Conditions

- Start with the first nonlinear moment $E[\Delta v_{it-1} v_{it}] = 0$ or $E[(\Delta y_{it-1} - \rho \Delta y_{it-2})(y_{it} - \rho y_{it-1})] = 0$. By mean stationarity, we get

$$E[y_{it} - y_{it-1} | \alpha_i] = E[\Delta y_{it} | \alpha_i] = 0 \implies E[\alpha_i \Delta y_{it}] = 0 \; \forall t$$

- This result, along with $E[y_{i1} u_{is}] = 0$ for $s \geq 2$ implies that $E[\Delta y_{i1} u_{is}] = 0$ for $s \geq 2$.
- Thus, $E[\Delta y_{i1} v_{is}] = 0$ $(s \geq 2)$. So we have $E[\Delta y_{i1} v_{i2}] = 0$.
- For $s \geq 3$, we can use the nonlinear moment condition $E[\Delta v_{is-1} v_{is}] = 0$ to back out

$$\begin{aligned} E[\Delta v_{is-1} v_{is}] = 0 &\implies E[(\Delta y_{is-1} - \rho \Delta y_{is-2}) v_{is}] = 0 \\ (s = 3) &\implies E[(\Delta y_{i2} - \rho \Delta y_{i1}) v_{i3}] = 0 \\ &= E[\Delta y_{i2} v_{i3}] = 0 \; (\because E[\Delta y_{i1} v_{is}] = 0 \; (s \geq 2)]) \end{aligned}$$

Repeat the similar process to ultimately get $E[\Delta y_{it-1} v_{it}] = 0$

Mean Stationarity: Linearizing Moment Conditions

- So the lagged differences of the instruments qualify as instruments for equation in levels.
- This is a Blundell-Bond estimator, or what is known as a system GMM estimation.
- Combine the above moment condition with the usual $E[Z_i' \Delta \mathbf{u}_i] = 0$ to get a joint moment condition

$$E\left[Z_i^{+'} \begin{pmatrix} \Delta \mathbf{u}_i \\ \mathbf{v}_i \end{pmatrix} \right] = 0$$

  where $Z_i^{+}, \mathbf{v}_i$ are as defined in the recitation note

Mean Stationarity: Linearizing Moment Conditions

- As for the homoskedastic case, we can use the mean stationarity condition to obtain a linear moment condition

$$
\begin{aligned}
E[v_{it}^2] - E[v_{it-1}^2] &= E[(y_{it} - \rho y_{it-1})v_{it} - (y_{it-1} - \rho y_{it-2})v_{it-1}] \\
&= E[y_{it}v_{it} - y_{it-1}v_{it-1}] = 0 \ (t = 2, ..., T)
\end{aligned}
$$

- The remaining terms can be written as

$$
\begin{aligned}
E[-\rho y_{it-1}(\alpha_i + u_{it}) + \rho y_{it-2}(\alpha_i + u_{it-1})] &= E[-\rho \Delta y_{it-1}\alpha_i + \rho y_{it-2}u_{it-1} - \rho y_{it-1}u_{it}] \\
&= -0 + 0 - 0 = 0
\end{aligned}
$$

  The mean stationarity justifies the first zero

- Therefore, using the combined moment condition

$$
E\left[Z_{iH}^{+'} \begin{pmatrix} \Delta \mathbf{u}_i \\ \mathbf{v}_i \end{pmatrix} \right] = 0
$$

  we can obtain a GMM estimator of ρ ($Z_{iH}^{+}, \mathbf{v}_i$ are defined in reci notes)

Factor Analysis

Framework

- Consider the following model

$$y_{it} = \mu_i + x_{it}'\beta + \lambda_i' f_t + u_{it}$$

  where λ_i is a vector of factor loadings and f_t is a vector of factors.
- Each can be written

$$\lambda_i = \begin{bmatrix} \lambda_{i1} \\ \dots \\ \lambda_{ir} \end{bmatrix}, f_t = \begin{bmatrix} f_{1t} \\ \dots \\ f_{rt} \end{bmatrix}$$

  where r is usually a small number.
- We will assume that only y_{it} and x_{it} is observable.

Factor Analysis

Framework

- In fact, we can see that this format is a generalized version of the additive fixed effect we have seen so far
  - We can back out the two-way fixed effects model
  - We can capture unobserved individual traits that can vary with time.
  - We can also capture entity-level responses to a common shock at certain time.
- If it is the case that we know β, then we can write

$$y_{it} - x_{it}'\beta = \mu_i + \lambda_i' f_t + u_{it}$$

  and estimate pure factor models.
- If we know what $\mu_i + \lambda_i' f_t$ is, we write

$$y_{it} - \mu_i - \lambda_i' f_t = x_{it}'\beta + u_{it}$$

  which becomes a standard model.

Factor Analysis

Framework

- If we need to determine the parameters of interest all at once, we can do a LASSO-type regularized regression in the following sense.

$$\min \sum_i \sum_t (y_{it} - x_{it}'\beta - l_{it})^2 + \tau \|L\|_*$$

  where l_{it} is the $\lambda_i' f_t$ and L is the matrix of l_{it}'s.
- Note that we are applying a nuclear norm here.
- In this context, what we are doing is to minimize the sum squared residuals but with the penalty that applies when the rank of L is large.
- Basically, we are treating β, λ_i, f_t as a parameter to be estimated, whereas for the static dynamic model, our interest was on β