ECON 710A Midterm Cheatsheet

Statistics and linear algebra review

Suppose X and Y are random variables. For any real numbers $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$ and any real number $\varepsilon > 0$.

$$\begin{aligned}
Cov(X,Y) &= E[XY] - E[X]E[Y] \\
E[Y] &= E[E[Y|X]] && \text{(if } E[||Y||] < \infty) \\
Var(Y) &= Var(E[Y|X]) \\
&\quad + E[Var(Y|X)] && \text{(if } E[||Y||^2] < \infty) \\
Pr(|X| \geq \varepsilon) &\leq E[|X|]/\varepsilon && \text{(Markov)} \\
Pr(|X - E[X]| \geq \varepsilon) &\leq Var[X]/\varepsilon^2 && \text{(Chebyshev)} \\
E[|XY|] &\leq E[|X|^p]^{1/p} E[|Y|^q]^{1/q} && \text{(Holder)} \\
E[|XY|]^2 &\leq E[X^2]E[Y^2] && \text{(Cauchy-Schwarz)}
\end{aligned}$$

LLN in $\mathcal{L}^1$ - If $\{X_i\}_{i=1}^\infty$ is a sequence of iid random variables with $E[|X_i|] < \infty$, then $\bar{X}_n \to_p E[X_1]$.

LLN in $\mathcal{L}^2$ - If $\{X_i\}_{i=1}^\infty$ is a sequence of random variables with $E[X_i] = \mu, E[X_i^2] < K$ for some $K \in \mathbb{R}$, and $Cov(X_i, X_j) = 0$ for all $i \neq j$, then $\bar{X}_n \to_p \mu$.

CLT - If $\{X_i\}_{i=1}^\infty$ is a sequence of iid random variables with $E[X_i^2] < \infty$, then $\frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - E[X_1]) \to_d N(0, Var(X_1))$.

Cramer-Wold device - A sequence of random vector $\{W_n\}_{n=1}^\infty$ converge in distribution to the random vector W iff $t'W_n$ converge in distribution to $t'W$ for any nonrandom vector t with $||t|| = 1$.

Block inversion - Consider the matrix $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ where A is square and D is invertible. M is invertible iff $E := A - BD^{-1}C$ is invertible in which case

$$M^{-1} = \begin{bmatrix} E^{-1} & -E^{-1}BD^{-1} \\ -D^{-1}CE^{-1} & D^{-1} + D^{-1}CE^{-1}BD^{-1} \end{bmatrix}$$

Sherman-Morrison - Consider an invertible matrix $A \in \mathbb{R}^{k \times k}$ and vectors $u, v \in \mathbb{R}^k$. $A + uv'$ is invertible iff $1 + v'A^{-1}u \neq 0$ in which case:

$$(A + uv')^{-1} = A^{-1} - \frac{A^{-1}uv'A^{-1}}{1 + v'A^{-1}u}$$

Instrumental Variables

Suppose that $Y = X'\beta_0 + U$ where

1. $E[U|Z] = 0$ (exogeneity)
2. $E[ZX']$ is invertible (relevance)
3. $E[Y^2 + ||X||^2 + ||Z||^2] < \infty$

Indentification:

1. Exogeneity of Z implies that $E[ZU] = E[ZE[U|Z]] = 0$.
2. $E[Z(Y - X'\beta)] = E[ZX'](\beta_0 = \beta)$.
3. Relevance of Z implies that $E[Z(Y - X'\beta)] = 0$ iff $\beta = \beta_0$.

By MOM, $\hat{\beta}^{IV} = (\frac{1}{n} \sum Z_i X_i')^{-1} \frac{1}{n} \sum Z_i Y_i$

IV is biased: $E[\hat{\beta}^{IV} | \mathbb{X}, \mathbb{Z}] = \beta_0 + (\sum Z_i X_i')^{-1} \sum Z_i E[U_i | \mathbb{X}, \mathbb{Z}]$.

With Phillips (1983) setup, $E[\hat{\beta}^{IV} | \mathbb{X}, \mathbb{Z}] = \beta_1 + \frac{\sigma_{UV}}{\sigma_V^2} \frac{\nu}{\frac{\pi_1}{sd(\hat{\pi}_1)} + \nu}$ where $\nu \sim N(0,1)$. [$\frac{\pi_1}{sd(\hat{\pi}_1)}$ is signal-to-noise ratio]

IV is consistent. For asymptotic normality, we need finite fourth moments. We can use Cramer-Wold device to establish asymptotic normality of the numerator.

$\Omega^{IV} = E[ZX']^{-1} E[ZZ'U^2] E[ZX']^{-1}$

$\hat{\Omega}^{IV} = (\frac{1}{n} \sum Z_i X_i')^{-1} \frac{1}{n} \sum Z_i Z_i' \hat{U}_i^2 (\frac{1}{n} \sum Z_i X_i')^{-1}$

F-statistic threshold for IV is 10.

Under homoskedasticity, the optimal instrument is $h^*(Z) = E[X|Z]$.

Two-square least squares:

- The relevance assumption becomes $E[ZX']$ has rank equal to the dimension of X and $E[ZZ']$ is invertible.
- Estimate π in $X_1 = Z'\pi + V$ by OLS.
- Define $h(Z, \hat{\pi}) = (Z'\hat{\pi}, X_2')'$.
- $\hat{\beta}^{2SLS} = (\frac{1}{n} \sum h(Z_i, \hat{\pi}) X_i')^{-1} \frac{1}{n} \sum h(Z_i, \hat{\pi}) Y_i$
- $\hat{\beta}^{2SLS} = (\frac{1}{n} \sum h(Z_i, \hat{\pi}) h(Z_i, \hat{\pi})')^{-1} \frac{1}{n} \sum h(Z_i, \hat{\pi}) Y_i$
- Small F-statistics can lead to substantial bias in 2SLS
- F-statistics threshold of 18 for 6 instruments.
- With multiple weak instruments, limited maximum likelihood estimator (LIML) is prefered to 2SLS.

Local average treatment effects can be used for heterogeneous coefficient models ($Y = X'\beta_0 \cdot U$ where $E[U|Z] = 1$ and $E[ZX']$ is invertible).

Time series

$$
\begin{aligned}
Y_t &= \alpha_0 + X_t'\delta_0 + U_t && \text{static}\\
Y_t &= \alpha_0 + X_t'\delta_0 + ... + X_{t-s}'\delta_s + U_t && \text{distributed lag}\\
Y_t &= \alpha_0 + Y_{t-1}\rho_1 + ... + Y_{t-p}\rho_p + U_t && \text{AR(p)}\\
Y_t &= \alpha_0 + \alpha_1 t + U_t && \text{linear trend}\\
log(Y_t) &= \beta_0 + \beta_1 t + U_t && \text{exponential trend}\\
Y_t &= \alpha_0 + \alpha_1 1\{t/12 \text{ is an integer }\} + U_t && \text{seasonality}\\
Y_t &= \varepsilon_t + \theta_1\varepsilon_{t-1} + ... + \theta_q\varepsilon_{t-q} && \text{MA(q)}
\end{aligned}
$$

A sequence of stochastic vectors $\{Z_t\}_{t=1}^{Z}$ is strictly stationary if $(Z_t, ..., Z_{t+k}) \sim (Z_1, ..., Z_{1+k})$ for all t and k.

If $\{Z_t\}_{t=1}^{Z}$ is strictly stationary and $\tilde{Z}_t = \rho(Z_t, Z_{t-1}, ...)$, then $\{\tilde{Z}_t\}$ is strictly stationary.

The time series $\{Y_t\}_{t=1}^{T}$ with $E[Y_t^2] < \infty$ for all t is covariance stationary if $E[Y_t] = \mu$ for all t and $Cov(Y_t, Y_{t+k}) = \gamma(k)$ for all t and some function γ.

γ is the autocovariance function and it is symmetric $\gamma(k) = \gamma(-k)$.

Suppose $Y_t = W_t'\beta + U_t$ where $\{(Y_t, W_t')\}_{t=1}^{T}$ is strictly stationary (W_t can include lags of X_t and Y_t),

1. $E[U_t|W_t] = 0$ (contemporaneous exogeneity)
2. $E[W_tW_t']$ and $\frac{1}{T}\sum_{t=1}^{T} W_tW_t'$ are invertible (no multicollinearity).

OLS needs strict exogeneity ($E[U_t|\mathbb{W}] = 0$) to be unbiased.

Contemporaneous exogeneity can fail if W_t includes lags of Y_t and U_t is serially correlated (e.g., ARMA(1, 1)), thus we have settings where OLS is asymptotically normal:

1. Models that include dynamic structure in outcome variables with iid errors. (White robust SE)
2. Models that exclude dynamic structure in outcome variables with serial dependent errors. (HAC or Newey-West SE)

To establish asymptotic normality for OLS of AR models, standard LLN and CLT do not suffice.

Martingale CLT: If $\{Z_t\}_{t=1}^{T}$ is strictly stationary with $E[Z_1^2] < \infty$, $E[Z_t|Z_{t-1}, ..., Z_1] = 0$ and $\frac{1}{T}\sum Z_t^2 \to_p E[Z_1^2]$, then $\frac{1}{\sqrt{T}}\sum Z_t \to_d N(0, E[Z_1^2]$ as $T \to \infty$.

Panel

Static error components model for $\{(Y_t, X_t')\}_{t=1}^{T}$: $Y_t = X_t'\beta_0 + U_t$ where $U_t = \alpha + \varepsilon_t$.

1. $E[\varepsilon_t|X_1, ..., X_T] = 0$
2. $E[X_tX_t']$ is invertible.
3. $E[\alpha^4 + ||X_t||^4 + \varepsilon_t^4] < \infty$ for all t.
4. $\{\{(Y_{it}, X_{it})\}_{t=1}^{T}\}_{i=1}^{n}$ is a random sample from the distribution of $\{(Y_t, X_t')\}_{t=1}^{T}$. (Note that α may depend on i).

Random effects assumptions:

1. $E[\alpha] = 0$, $\sigma_\alpha^2 = E[\alpha^2]$, and α is independent of $\{(X_t', \varepsilon_t)\}_{t=1}^{T}$.
2. $\{\varepsilon_t\}_{t=1}^{T}$ is white noise ($Cov(\varepsilon_t, \varepsilon_s|X_1, ..., X_t) = \sigma^2 1\{s = t\}$).

Strict exogeneity and independence between X_t and α results in an abundance of moment conditions to consider.

Pooled OLS focuses on contemporaneous moment conditions:

$\hat{\beta}^{OLS} = (\frac{1}{n}\sum_i\sum_t X_{it}X_{it}')^{-1}\frac{1}{n}\sum_i\sum_t X_{it}Y_{it}$.

FGLS uses all moment conditions:

$\hat{\beta}^{GLS} = (\frac{1}{n}\sum_i\sum_t \tilde{X}_{it}X_{it}')^{-1}\frac{1}{n}\sum_i\sum_t \tilde{X}_{it}Y_{it}$

where $\tilde{X} = X_{it} - \frac{T\hat{\sigma}_\alpha^2}{\hat{\sigma}^2 + T\hat{\sigma}_\alpha^2}\bar{X}_i$ and $\bar{X}_i = \frac{1}{T}\sum_t X_{it}$.

Fixed effects impose no further assumptions on α, thus allow for dependence between $(X_1, ..., X_t)$ and α.

$\hat{\beta}^{FE} = (\frac{1}{n}\sum_i\sum_t(X_{it} - \bar{X}_i)X_{it}')^{-1}\frac{1}{n}\sum_i\sum_t(X_{it} - \bar{X}_i)Y_{it}$

$\sqrt{n}(\hat{\beta}^{FE} - \beta_0) \to_d N(0, \bar{H}^{-1}\Omega\bar{H}^{-1})$ $\bar{H} = E[\sum_t(X_t - \bar{X})(X_t - \bar{X})']$ $\Omega = E[(\sum_t(X_t - \bar{X})\varepsilon_t)(\sum_s(X_s - \bar{X})\varepsilon_s)']$

Estimating ε_t is difficult because we don't have an estimator for α, but we know that

$\Omega = E[(\sum_t(X_t - \bar{X})U_t)(\sum_s(X_s - \bar{X})U_s)']$ where $U_t = \alpha + \varepsilon_t$, so $\hat{U}_t = Y_t - X_t'\hat{\beta}^{FE}$.