Ch. 2

☆ Matrix Calculus A.9 Hansen

1. Conditioning Theorem

$E(g(x)y|x)=g(x)E(y|x)$ $E(g(x)y)=E(g(x)E(y|x))$ Tips: $E(h(x)e)=0$

Specifically, $E(xe)=0$

2. CEF Function

Def. $m(x)=E(y|x)=\int_R y f_{y|x}(y|x)dy$ CEF Error $e=y-m(x) \Leftrightarrow y=m(x)+e$ $[E(e|x)=E(e)=0]$ By Def.

(1) Def. $\sigma^2=Var(e)=E(e^2)$ σ^2: regression variance $Var(y)\geq Var(y-E(y|x_1))\geq Var(y-E(y|x_1,x_2))$

(2) Best Predictor: $\arg\min E(y-g(x))^2=m(x)$ [Unconditional] P1 信息越多，方差(误差)越小。

(3) Def. Conditional Variance $\sigma^2(x)=E(e^2|x)$

(4) Rescaled error: $\varepsilon=\frac{e}{\sigma(x)}$ $y=m(x)+\sigma(x)\varepsilon$

(5) Linear CEF: Def. $E(y|x)=m(x)=x'\beta$ with $\nabla m(x)=\beta$, $\begin{cases} y=x'\beta+e \\ E(e|x)=0\end{cases}$ Linear CEF Model

3. Linear Projection Tips: 设假设 $m(x)=x'\beta$，则无法有 $E[e|x]=0$.

When CEF is not linear, need approximation.

Def. $P(y|x)=x'\beta$ s.t. $S(\beta)=E(y-x'\beta)^2$ and $\beta=\arg\min S(\beta)$. $\Rightarrow \beta=\underbrace{(E(xx'))^{-1}}_{Q_{xx}^{-1}}\underbrace{E(xy)}_{Q_{xy}}$, P2 $E(xe)=E(e)=0$. $E(x_j e)=0$ ☆

(1) Linear Predictor Error: $e=y-x'\beta$ $\sigma^2=E(e^2)=Q_{yy}-Q_{yx}Q_{xx}^{-1}Q_{xy}=Q_{yy\cdot x}$ P3*

(2) Linear Projection Model in demean variables: $y=x'\beta+\alpha+e$, $\alpha=\mu_y-\mu_x'\beta$, $\beta=Var^{-1}(x)Cov(x,y)$

(3) Decomposition: $\beta_1=\frac{E(u_1 y)}{Eu_1^2}$ P4*: OVB

(4) CEF与linear projection: ① linear CEF是linear projection，反之否 ② LP是CEF的approximation (不一定)

③ $E(xe)=0$ 对LP, CEF都成立，但CEF更强，有$E(e|x)=0$ ④ CEF是linear时，二者等价。

(5) Random Coefficient Model: $y=x'\eta$ η: individual specific random, independent of x. Actual effect on obs.

$\Rightarrow E(y|x)=x'\beta$, $Var(y|x)=x'\Sigma x$ when $\beta=E(\eta)$, $\Sigma=Var(\eta)$.

(6) Causal Effects: $C(x_1,x_2,u)=\nabla_1 h(x_1,x_2,u)$ $ACE(x_1,x_2)=E(C(x_1,x_2,u)|x_1,x_2)=\int_R \nabla_1 h(x_1,x_2,u) f(u|x_1,x_2)d$

$\nabla_1 m(x_1,x_2)=ACE(x_1,x_2)$ when CIA holds. Tips: 注意区别CIA与Mean-Independence

P5*

(7) Reverse Regression

Ch. 3

1. OLS Estimator

$S_n(\beta)=\frac{1}{n}\sum(y_i-x_i'\beta)^2$, $\hat\beta=\arg\min S_n(\beta) \Rightarrow \hat\beta=(\sum x_i x_i')^{-1}(\sum x_i y_i)$ or $\hat\beta=(X'X)^{-1}(X'y)$ $(\sum x_i x_i'=X'X, \sum x_i y_i=X'y)$

(1) Residual: $\hat y_i=x_i'\hat\beta$, $\hat e_i=y_i-\hat y_i=y_i-x_i'\hat\beta$ $\sum x_i\hat e_i=0$, $\sum\hat e_i=0$ [Holds true for all linear regressions]

$X'\hat e=0$ 包括Partition ⑥ $X=[X_1,X_2]$

(2) P的性质: $\underset{n\times n}{P}=X(X'X)^{-1}X'$ ① $Py=X\hat\beta=\hat y$ ② $PZ=Z\,(Z=X\Gamma)$, $PX=X$ ③ Symmetric and Idempotent

$P'=P$, $P\cdot P=P\Rightarrow$ P.S.D. ④ For $X=1(n\times1)$, $P=\frac{1}{n}11'$, $Py=1\bar y=\begin{bmatrix}\frac{1}{n}\sum y\\ \vdots \\ \frac{1}{n}\sum y\end{bmatrix}$ ⑤ $X=[X_1,X_2]$, $PX_1=X_1$ ⑥ $tr(P)=k=\sum h$

⑦ $h_{ii}=x_i'(X'X)^{-1}x_i$, h_{ii}是P的第i个对角线元素 ⑧ $0\leq h_{ii}\leq 1$

(3) M的性质: ① $M=I_n-P$ ② $MZ=0$ for $Z=X\Gamma$ $(MX=0)$, $X=[X_1,X_2]$, $MX_1=0$ ③ $tr(M)=tr(I_n)-tr(P)=n-k$

④ $My=\hat e$ ⑤ $\underset{n\times n}{M_1}=I_n-P_1=I_n-\frac{1}{n}11'$, $M_1y=y-1\bar y$ (Demeaned value y) ⑥ Symmetric and Idempotent

$M\hat e=\hat e$ ☆

(4) $\hat\sigma^2=\frac{1}{n}\sum\hat e_i^2$, $\hat\sigma^2=\frac{1}{n}\hat e'\hat e=\frac{1}{n}e'Me=\frac{1}{n}y'My$, $\tilde\sigma^2=\frac{1}{n}\sum e_i^2$; $\tilde\sigma^2-\hat\sigma^2\geq0$

(5) For Partition Regression: $X=[X_1,X_2]$, $\hat\beta=\begin{bmatrix}\hat\beta_1\\ \hat\beta_2\end{bmatrix}$, $\hat\beta_1=(X_1'M_2X_1)^{-1}(X_1'M_2y)$, $\hat\beta_2=(X_2'M_1X_2)^{-1}(X_2'M_1y)$

(6) $\hat\beta_2=(\tilde X_2'\tilde X_2)^{-1}(\tilde X_2'\tilde e_1)$, $\tilde X_2=M_1X_2$, $\tilde e_1=M_1y$ (FWL Theorem 证明)

(7) Leave-one-out Error: $\hat\beta_{(-i)}=\hat\beta-(1-h_{ii})^{-1}(X'X)^{-1}x_i\hat e_i$, $\tilde e_i=(1-h_{ii})^{-1}\hat e_i$, $\tilde e=M^*\hat e$, $\tilde\sigma^2=\frac{1}{n}\sum\tilde e_i^2=\frac{1}{n}\sum(1-h_{ii})^{-2}\hat e_i^2$

(8) Normal Regression Model: $e_i|x_i\sim N(0,\sigma^2)$, $y_i|x_i\sim N(x_i'\beta,\sigma^2)$

(9) $R^2=\frac{\sum(\hat y_i-\bar y)^2}{\sum(y_i-\bar y)^2}=1-\frac{\sum\hat e_i^2}{\sum(y_i-\bar y)^2}$

Ch.4

1. Unbiaseness of OLS Estimator

$E(\hat\beta|X)=\beta$, $E(\hat\beta)=\beta$ $\quad \hat\beta=\beta+(X'X)^{-1}X'e$

2. Variance Matrix $\quad * \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix}$

Def. $V_{\hat\beta} \overset{def}{=} Var(\hat\beta|X)$ $\quad$ Def. $Var(e|X)=D=E(ee'|X)=diag(\sigma_1^2,\cdots,\sigma_n^2) \overset{homosk}{=} I_n\sigma^2$

(1) $V_{\hat\beta}=Var(\hat\beta|X)=(X'X)^{-1}X'Var(e|X)X(X'X)^{-1}=(X'X)^{-1}(X'DX)(X'X)^{-1} \overset{homosk}{=} (X'X)^{-1}\sigma^2$ $\quad$ Tips: $X'DX=\sum x_i x_i'\sigma_i^2$

(2) G-M Theorem

(3) Variance of Residuals: $M=I_n-P$ $\quad M^*=diag((1-h_{ii})^{-1})$

① $\hat e=Me$ $\quad \tilde e=M^*\hat e$ $\quad \tilde e=M^*Me$ $\quad \bar e=M^{*\frac12}Me$ $\quad$ ② $\tilde e_i=\frac{\hat e_i}{1-h_{ii}}$ $\quad \bar e_i=\frac{\hat e_i}{\sqrt{1-h_{ii}}}$

③ $Var(e|X)=D$ $\quad Var(\hat e|X)=MDM \overset{homosk}{=} M\sigma^2$ $\quad Var(\tilde e|X)=M^*MDMM^* \overset{homosk}{=} M^*MM^*\sigma^2$

$Var(\bar e|X)\overset{homo}{=}M^{*\frac12}MM^{*\frac12}\sigma^2$

When homo: ④ $Var(e_i^2|X)=\sigma^2$ $\quad Var(\hat e_i|X)=(1-h_{ii})\sigma^2$ $\quad Var(\tilde e_i|X)=(1-h_{ii})^{-1}\sigma^2$ $\quad Var(\bar e_i|X)=\sigma^2$

(4) σ^2的估计: ① $\hat\sigma^2=\frac1n\hat e'\hat e$ $\quad E(\hat\sigma^2|X)=\frac{n-k}{n}\sigma^2$ $\quad$ ② $S^2=\frac{1}{n-k}\sum\hat e_i^2$ $\quad E(S^2|X)=\sigma^2$ 无偏 (Under homosk)

③ $\bar\sigma^2=\frac1n\sum\bar e_i^2=\frac1n\sum(1-h_{ii})^{-1}\hat e_i^2$ $\quad E[\bar\sigma^2|X]=\sigma^2$ 无偏 (Under homosk)

(5) Estimation of $\hat\beta$'s variance matrix $\quad$ Tips: $Var(\beta)=E[Var(\beta|X)]+Var[E(\tilde\beta|X)]$

① homosk: $V_{\hat\beta}=(X'X)^{-1}\sigma^2$ $\quad \hat V^0_{\hat\beta}=(X'X)^{-1}S^2$ 无偏

② heterosk: $V_{\hat\beta}=(X'X)^{-1}(X'DX)(X'X)^{-1}$ $\quad D=diag(\sigma_1^2,\cdots,\sigma_n^2)=E(ee'|X)$

$\hat V^{ideal}_{\hat\beta}=(X'X)^{-1}(\sum x_ix_i'e_i^2)(X'X)^{-1}$ $\quad \hat V^W_{\hat\beta}=(X'X)^{-1}(\sum x_ix_i'\hat e_i^2)(X'X)^{-1}$ $\quad \hat V_{\hat\beta}=\frac{n}{n-k}(X'X)^{-1}(\sum x_ix_i'\hat e_i^2)(X'X)^{-1}$

$\tilde V_{\hat\beta}=(X'X)^{-1}(\sum x_ix_i'\tilde e_i^2)(X'X)^{-1}=(X'X)^{-1}(\sum x_ix_i'(1-h_{ii})^{-2}\hat e_i^2)(X'X)^{-1}$ $\quad \bar V_{\hat\beta}=(X'X)^{-1}(\sum x_ix_i'\bar e_i^2)(X'X)^{-1}=(X'X)^{-1}(\sum(1-h_{ii})^{-1}x_ix_i'\hat e_i^2)(X'X)^{-1}$

$\hat V^W_{\hat\beta}<\bar V_{\hat\beta}<\tilde V_{\hat\beta}$

(6) $\bar R^2=1-\frac{S^2}{\hat\sigma_y^2}=1-\frac{(n-1)\sum\hat e_i^2}{(n-k)\sum(y_i-\bar y)^2}$ $\quad \tilde R^2=1-\frac{\sum\tilde e_i^2}{\sum(y_i-\bar y)^2}$

Ch.6

1. Consistency of OLS Estimator

$\hat\beta=\hat Q_{xx}^{-1}\hat Q_{xy}=(\frac1n\sum x_ix_i')^{-1}(\frac1n\sum x_iy_i) \xrightarrow[WLLN,\ CMT]{P} E^{-1}(x_ix_i')E(x_iy_i)=Q_{xx}^{-1}Q_{xy}=\beta$ $\quad \hat Q_{xx}\xrightarrow{P}Q_{xx}$, $\hat Q_{xy}\xrightarrow{P}Q_{xy}$, $\hat Q_{xx}^{-1}\xrightarrow{P}Q_{xx}^{-1}$, $\hat Q_{xe}\xrightarrow{P}0$, $\hat\beta\xrightarrow{P}\beta$ as $n\to\infty$

2. Asymptotic Normality $\quad$ Tips: CLT (Hansen P119) CMT (Hansen P122)

☆ $\sqrt n(\hat\beta-\beta)=(\frac1n\sum x_ix_i')^{-1}(\frac{1}{\sqrt n}\sum x_ie_i)\xrightarrow{d}N(0,V_\beta)$ $\quad V_\beta=Q_{xx}^{-1}\Omega Q_{xx}^{-1}$

推导 (1) $\frac{1}{\sqrt n}\sum x_ie_i\xrightarrow{d}N(0,\Omega)$, $\Omega=E(x_ix_i'e_i^2)$ $\quad \therefore \sqrt n(\hat\beta-\beta)\xrightarrow{d}Q_{xx}^{-1}N(0,\Omega)=N(0,Q_{xx}^{-1}\Omega Q_{xx}^{-1})$

V_β: asymptotic covariance matrix of $\hat\beta$ $\quad$ ☆ $V_\beta\approx nV_{\hat\beta}$ $\quad \hat V_\beta=n\hat V_{\hat\beta}$

(2) Under homosk: $cov(x_ix_i',e_i^2)=0\Rightarrow\Omega=E(x_ix_i')E(e_i^2)=Q_{xx}\sigma^2$, $V_\beta=Q_{xx}^{-1}\Omega Q_{xx}^{-1}=Q_{xx}^{-1}\sigma^2\equiv V_\beta^0$

(3) Consistency of Error Variance Estimator: $\hat\sigma^2\xrightarrow{P}\sigma^2$, $S^2\xrightarrow{P}\sigma^2$

(4) Consistent Estimation of asymptotic covariance matrix:

① Under homosk: $\hat V^0_\beta=\hat Q_{xx}^{-1}S^2\xrightarrow[CMT]{P}V^0_\beta$ $\quad$ Proof☆

② Under heterosk: $\hat V^W_\beta=\hat Q_{xx}^{-1}\hat\Omega\hat Q_{xx}^{-1}$ $(\hat\Omega=\frac1n\sum x_ix_i'\hat e_i^2)$ $[\hat V^W_\beta=n\hat V^W_{\hat\beta}]$

Tips: Selector Matrix: $r(\beta)=R'\beta$, $R=\begin{pmatrix}I\\0\end{pmatrix}$, $V_\theta=V_{ii}$

3. Function of Parameters

(1) $\hat\theta=r(\hat\beta)\xrightarrow{CMT}\hat\theta\xrightarrow{P}\theta$ $\quad \sqrt n(\hat\theta-\theta)\xrightarrow{d}N(0,V_\theta)$ $\quad V_\theta=R'V_\beta R$, $R=\frac{\partial}{\partial\beta}r(\beta)'$; $\hat V_\theta=\hat R'\hat V_\beta\hat R=n\hat V_{\hat\theta}=n\hat R'\hat V_{\hat\beta}\hat R$.

$\hat\beta\xrightarrow{P}\beta$

(2) $t_n(\theta)=\frac{\hat\theta-\theta}{S(\hat\theta)}\xrightarrow{d}N(0,1)$ $\quad C_n=[\hat\theta-c\cdot S(\hat\theta),\ \hat\theta+c\cdot S(\hat\theta)]$ $\quad$ Reg Interval: $[x'\hat\beta\pm1.96\sqrt{x'\hat V_{\hat\beta}x}]$ $\quad m(x)=x'\beta$

(3) Wald Statistic: $\theta=r(\beta)$: $R^k\to R^q$ $\quad W_n(\theta)=(\hat\theta-\theta)'\hat V_{\hat\theta}^{-1}(\hat\theta-\theta)=n(\hat\theta-\theta)'\hat V_\theta^{-1}(\hat\theta-\theta)$

Ch.7

(1) Constrained Least Squares

$\underset{q\times k}{R'}\underset{k\times1}{\beta}=\underset{q\times1}{c}$ k个回归子, q个约束条件, 线性

$\tilde{\beta}_{cls}=\underset{R'\beta=c}{\arg\min} SSE_n(\beta)$ $SSE_n(\beta)=\sum(y_i-x_i'\beta)^2$ $\Rightarrow \tilde{\beta}_{cls}=\hat{\beta}-\underbrace{(X'X)^{-1}}_{n\hat{Q}_{xx}^{-1}}R[R'\underbrace{(X'X)^{-1}}_{n\hat{Q}_{xx}^{-1}}R]^{-1}(R'\hat{\beta}-c)$ * 证明

$\tilde{\sigma}^2_{cls}=\frac{1}{n}\sum\tilde{e}_i^2$ $S^2_{cls}=\frac{1}{n-k+q}\sum\tilde{e}_i^2$ $E(S^2_{cls}|X)=\sigma^2$

(2) Minimum Distance

① $J_n(\beta)=n(\hat{\beta}-\beta)'\underset{k\times k}{W_n}(\hat{\beta}-\beta)$ $\tilde{\beta}_{md}=\underset{R'\beta=c}{\arg\min} J_n(\beta)$ 本质: Weighted Euclidean distance between β and $\hat{\beta}$

$\tilde{\beta}_{md}=\hat{\beta}-W_n^{-1}R(R'W_n^{-1}R)^{-1}(R'\hat{\beta}-c)$ *证明 ☆当 $W_n=\hat{Q}_{xx}$, $\tilde{\beta}_{md}=\tilde{\beta}_{cls}$. ☆*(证明)

② Asymptotic Distribution: $\tilde{\beta}_{md}=\hat{\beta}-W_n^{-1}R(R'W_n^{-1}R)^{-1}(R'\hat{\beta}-c)\Rightarrow\sqrt{n}(\tilde{\beta}_{md}-\beta)=[I_k-\underbrace{W_n^{-1}}_{\xrightarrow{p}W^{-1}}R(R'W_nR)^{-1}R']\underbrace{\sqrt{n}(\hat{\beta}-\beta)}_{\xrightarrow{d}N(0,V_\beta)}$

$\xrightarrow{d}N(0,\overset{\Sigma(W)}{\cancel{V_\beta(W)}})$ $\Sigma(W)=[\cdot\cdot]V_\beta[\cdot\cdot]$

③ Efficient Minimum Distance Estimator

For $W_n=\hat{V}_\beta^{-1}$ $(W=V_\beta^{-1})$, $\tilde{\beta}_{emd}=\hat{\beta}-\hat{V}_\beta R(R'\hat{V}_\beta R)^{-1}(R'\hat{\beta}-c)$ $\sqrt{n}(\tilde{\beta}_{emd}-\beta)\xrightarrow{d}N(0,V_\beta^*)$ $V_\beta^*=V_\beta-V_\beta R(R'V_\beta R)^{-1}R'V_\beta$

$\therefore V_\beta^*\le V_\beta$, the EMD estimator has smaller variance than unrestricted one. And $V_\beta^*\le V_\beta(W)$, $\forall W$.

(3) Misspecification

$R'\beta=c^*$ ← real value $\tilde{\beta}_{md}\xrightarrow{p}\beta^*_{md}=\beta-\underbrace{W^{-1}R(R'W^{-1}R)^{-1}(c^*-c)}_{\text{asymptotic bias}}\neq\beta$

Let $\beta_n^*=\beta-W_n^{-1}R(R'W_n^{-1}R)^{-1}(c^*-c)$, $\sqrt{n}(\tilde{\beta}_{md}-\beta_n^*)=(I-W_n^{-1}R(R'W_n^{-1}R)^{-1}R')\sqrt{n}(\hat{\beta}-\beta)\xrightarrow{d}N(0,\Sigma(W))$

Asymptotic distribution centered at the pseudo-true values

Ch.8

(1) Wald Test

$\theta\in R^q$ $\sqrt{n}(\hat{\theta}-\theta)\xrightarrow{d}N(0,\underset{q\times q}{V_\theta})$ 一般式: $\theta=r(\beta)$, H_0: $\theta=\theta_0$ $\hat{\theta}=r(\hat{\beta})$

H_0: $\theta=0$ $W_n=n\hat{\theta}'\hat{V}_\theta^{-1}\hat{\theta}\xrightarrow{d}\chi^2_q$ $W_n=(\hat{\theta}-\theta_0)'\hat{V}_\theta^{-1}(\hat{\theta}-\theta_0)\cdot n=(\hat{\theta}-\theta_0)'\hat{V}_{\hat\theta}^{-1}(\hat{\theta}-\theta_0)$

☆☆当 $W_n=\hat{V}_\beta^{-1}$, $R'\beta=\theta_0$ 时(即EMD, 线性约束). $\hat{V}_\theta=\hat{R}'\hat{V}_\beta\hat{R}$, $\hat{R}=\frac{\partial}{\partial\beta}r(\hat{\beta})'$.

* $J_n^*=W_n$, J_n为minimum distance criterion. 证明见 Hansen P195

Under homoskedasticity: $F=\frac{W_n}{q}$, $qF\xrightarrow{d}\chi^2_q$.

Ch.9

(1) NLS

$m(x,\theta)=E(y_i|x_i=x)$, m is non-linear function of θ.

Def. $\hat{\theta}=\arg\min\frac{1}{n}\sum(y_i-m(x_i,\theta))^2$ FOC: $\sum m_\theta(x_i,\hat{\theta})\hat{e}_i=0$, $\frac{\partial}{\partial\theta}m(x,\theta)=m_\theta(x,\theta)$.

Asymptotic Distribution: $\sqrt{n}(\hat{\theta}-\theta)\xrightarrow{d}N(0,V_\theta)$

(2) GLS

Def. $\tilde{\beta}=(X'D^{-1}X)^{-1}X'D^{-1}y$ But D cannot be observed. We need FGLS.

$E[e_i^2|x_i]=z_i'\alpha$ z_i: submatrix of x_i. $\therefore e_i^2=z_i'\alpha+\xi_i$. $\hat{\alpha}=(\frac{1}{n}\sum z_iz_i')^{-1}(\frac{1}{n}\sum z_ie_i^2)$ $\tilde{\sigma}_i^2=z_i'\hat{\alpha}$.

But e_i cannot be observed. $\therefore\tilde{\alpha}=(\frac{1}{n}\sum z_iz_i')^{-1}(\frac{1}{n}\sum z_i\hat{e}_i^2)$, $\tilde{\sigma}_i^2=z_i'\tilde{\alpha}$. $\tilde{\alpha}$ and $\hat{\alpha}$ 渐近等价。

$\therefore$ FGLS: $\frac{y_i}{\tilde{\sigma}_i}=(\frac{x_i}{\tilde{\sigma}_i})'\beta+\frac{e_i}{\tilde{\sigma}_i}$. $\tilde{\beta}=(X'\tilde{D}^{-1}X)^{-1}X'\tilde{D}^{-1}y$, $\tilde{D}=\mathrm{diag}\{\tilde{\sigma}_1^2,\ldots,\tilde{\sigma}_n^2\}$.

(3) LAV

Def. $\theta = \arg\min_\theta E|y-\theta|$, $\theta = med(y)$ medium.

Least Absolute Deviation: $\min_{g(\cdot)} E[|y_i - g(x_i)|] \Rightarrow$ Solution $med(y_i|x_i)$. when $g(x_i)=x_i'\beta$, $\tilde{\beta}_{LAD}$, $\sqrt{n}(\hat{\beta}_{LAD}-\beta) \xrightarrow{d} N(0,V)$ ↱ Asymptotic Var

(4) Quantile Regression

① Quantile: Def1. $Q_\tau = \inf\{u: F_u(u) \geq \tau\}$ τ^{th} Quantile. Def2. $Q_\tau = \arg\min_\theta E[\rho_\tau(U-\theta)]$, $\rho_\tau(q) = \begin{cases} -q(1-\tau) & q<0 \\ q\tau & q\geq 0 \end{cases} = q(\tau - 1(q<0))$.

② Quantile Reg Model: $y_i = x_i'\beta_\tau + e_i$, $Q_\tau(x) = \beta_\tau' x$ $\sqrt{n}(\hat{\beta}_\tau - \beta) \xrightarrow{d} N(0, V_\tau)$.

Tips: Leibniz Rule

Ch. 13

(1) Over-identified Linear Model

$y_i = x_i'\beta + e_i$ $x_i: k\times 1$ $g_i(\beta) = z_i(y_i - x_i'\beta)$ $J_n(\beta) = n\bar{g}_n(\beta)'W_n\bar{g}_n(\beta)$

$E(z_ie_i)=0$ $z_i: l\times 1$, $l\geq k$ $\bar{g}_n(\beta) = \frac{1}{n}\sum z_i(y_i - x_i'\beta) = \frac{1}{n}(Z'y - Z'X\beta)$

Def. $\hat{\beta}_{GMM} = \arg\min_\beta J_n(\beta) = \arg\min_\beta n\bar{g}_n(\beta)'W_n\bar{g}_n(\beta) \Rightarrow$ solution: $\hat{\beta}_{GMM} = [(X'Z)W_n(Z'X)]^{-1}(X'Z)W_n(Z'y)$.

(2) Asymptotic Distribution of $\hat{\beta}_{GMM}$

$\hat{\beta}_{GMM} = [(X'Z)W_n(Z'X)]^{-1}(X'Z)W_n(Z'(X\beta+e)) = \beta + [(X'Z)W_n(Z'X)]^{-1}(X'Z)W_nZ'e$ $\Omega = E[z_iz_i'e_i^2] = E(g_ig_i') = E[Z'ee'Z]$.

$$\sqrt{n}(\hat{\beta}_{GMM}-\beta) = [\underbrace{\tfrac{1}{n}(X'Z)W_n(\tfrac{1}{n}Z'X)}_{\xrightarrow{p} Q'W_nQ}]^{-1}\underbrace{(\tfrac{1}{n}X'Z)}_{Q'}\underbrace{W_n}_{W}\underbrace{(\tfrac{1}{\sqrt{n}}Z'e)}_{N(0,\Omega)} \xrightarrow{d} N(0, V_\beta) \quad V_\beta = (Q'WQ)^{-1}(Q'W\Omega WQ)(Q'WQ)^{-1}$$

Optimal: $W_0 = \Omega^{-1} \Rightarrow$ efficient GMM: $\hat{\beta} = (X'Z\Omega^{-1}Z'X)^{-1}X'Z\Omega^{-1}Z'y$, $\sqrt{n}(\hat{\beta}-\beta) \xrightarrow{d} N(0, (Q'\Omega^{-1}Q)^{-1})$.

(3) Estimation of Ω^{-1}.

① Uncentered estimator: $W_n = (\frac{1}{n}\sum \hat{g}_i\hat{g}_i')^{-1}$ $\hat{g}_i = z_i(y_i - x_i'\hat{\beta})$

② Centered estimator: $W_n = (\frac{1}{n}\sum \hat{g}_i^*\hat{g}_i^{*\prime})^{-1} = (\frac{1}{n}\sum \hat{g}_i\hat{g}_i' - \bar{g}_n\bar{g}_n')$ $\hat{g}_i^* = \hat{g}_i - \bar{g}_n$

First Step: Estimate $\hat{\beta}$, $\hat{e}_t = y_i - x_i'\hat{\beta}$ with GMM using $W_n = (Z'Z)^{-1}$;

Second Step: Using $\hat{\beta}$ from first step to estimate $\hat{\Omega}^{-1} = W_n$ (① or ②);

$$\hat{\beta} = (X'Z(\underbrace{\hat{g}'\hat{g} - n\bar{g}_n\bar{g}_n'}_{nW_n})^{-1}Z'X)^{-1}X'Z(\underbrace{\hat{g}'\hat{g} - n\bar{g}_n\bar{g}_n'}_{nW_n})^{-1}Z'y \qquad \underset{l\times n\; n\times l}{\hat{g}'\hat{g}} = \sum \underset{l\times 1\; 1\times l}{\hat{g}_i\hat{g}_i'}$$

(with $\hat{e}_i$ written above the first $\hat{g}$)

(4) Tests

① Over-identification Test

$J_n = n\bar{g}_n'W_n\bar{g}_n = n^2\bar{g}_n'(\hat{g}'\hat{g} - n\bar{g}_n\bar{g}_n')^{-1}\bar{g}_n \xrightarrow{d} \chi^2_{l-k}$ * If $J \geq$ critical value, reject the over-identified model.

② Distance Statistic (Better for non-linear case)

$J_n = n\cdot\bar{g}_n(\beta)'W_n\bar{g}_n(\beta)$ $H_0: h(\beta)=0$; $\hat{\beta} = \arg\min J_n(\beta)$, $\tilde{\beta} = \arg\min_{h(\beta)=0} J(\beta)$

$D_n = J_n(\tilde{\beta}) - J_n(\hat{\beta}) \xrightarrow{d} \chi^2_r$ If h is linear, D = Wald statistic.

(5) Conditional Moment Restrictions

$E(e_i(\beta)|z_i)=0$

Optimal instrument: *

$R_i = E(\frac{\partial}{\partial\beta}e_i(\beta)|z_i)$, $\sigma_i^2 = E(e_i(\beta)^2|z_i)$ $A_i = -\sigma_i^{-2}R_i$ (optimal instrument) $g_i(\beta) = A_ie_i(\beta)$

In linear: $R_i = -E(x_i|z_i)$ $\sigma_i^2 = E(e_i^2|z_i)$ $A_i = \sigma_i^{-2}E(x_i|z_i)$ $g_i(\beta) = \sigma_i^{-2}E(x_i|z_i)\cdot e_i$

If $x_i = z_i$ (linear regression): $g_i = \sigma_i^{-2}\cdot x_i\cdot e_i$ $Eg_i = 0 \Rightarrow E[\frac{x_i}{\sigma_i}\cdot\frac{(y_i - x_i'\beta)}{\sigma_i}] = 0$ GLS.

Ch.15

(1) Sources of Endogeneity

① Measurement error

Assume (y_i, x_i^*), x_i^* real value, $E(y_i|x_i^*)=x_i^{*\prime}\beta$. $x_i = x_i^* + u_i$ observed value

$y_i = x_i^{*\prime}\beta + e_i \Rightarrow y_i=(x_i-u_i)'\beta+e_i \Rightarrow y_i = x_i'\beta + \underbrace{(e_i-u_i'\beta)}_{v_i}$ $\quad E(x_iv_i)\neq 0 \quad \therefore \hat\beta \xrightarrow{P} \beta^* \neq \beta$.

② Simultaneous equation bias

(2) IV and identification

Def. $\underset{l\times 1}{z_i}$ is IV if $E(z_ie_i)=0$. Conditions: ① $E[z_ie_i]=0$ ② $rank(E[z_ix_i'])=k$.

How to identify β using IV?

First Stage: $\underset{n\times k}{X} = \underset{n\times l}{Z}\,\underset{l\times k}{\Gamma} + \underset{n\times k}{U}$ where $E(z_iu_i')=0$; Reduced Form: $\underset{n\times 1}{y}=(Z\Gamma+U)\underset{k\times 1}{\beta}+e = Z\underbrace{\Gamma\beta}_{\lambda}+(\underbrace{U\beta+e}_{v})$

If $l=k$, $\beta=\Gamma^{-1}\lambda$; If $l>k$, $\forall W>0$, $\beta=(\underset{k\times l}{\Gamma'}\,\underset{l\times l}{W}\,\underset{l\times k}{\Gamma})^{-1}\Gamma'W\lambda$. 前提：$rank(\Gamma)=k$

(3) Estimation

Tips: GMM是IV估计的普遍解，ILS与2SLS均是其特解。

① GMM

$y_i=x_i'\beta+e_i$, $E(z_ie_i)=0$ $\Rightarrow$ $Eg_i(\beta)=0$, $g_i(\beta)=z_i(y_i-x_i'\beta)$.

Using GMM we have: $\hat\beta=(X'ZW_nZ'X)^{-1}X'ZW_nZ'y$

② ILS ~~(前提: k=l, just identified)~~

$\hat\beta_{ILS}=\hat\Gamma^{-1}\hat\lambda = \underbrace{((Z'Z)^{-1}(Z'X))^{-1}}_{\text{X对Z回归, First Stage}}\underbrace{((Z'Z)^{-1}(Z'y))}_{\text{y对Z回归, Reduced Form}} = (Z'X)^{-1}(Z'Z)(Z'Z)^{-1}(Z'y)=(Z'X)^{-1}(Z'y)$

If we use GMM estimation:

$\hat\beta_{GMM}=(X'ZW_nZ'X)^{-1}(X'ZW_nZ'y)=(Z'X)^{-1}W_n^{-1}(X'Z)^{-1}X'ZW_nZ'y=(Z'X)^{-1}Z'y=\hat\beta_{ILS}$

∴当$l=k$，恰好识别时，$\hat\beta_{ILS}=\hat\beta_{GMM}$.

③ 2SLS

$\hat\beta_{2SLS}=(X'\underbrace{Z(Z'Z)^{-1}Z'}_{P}X)^{-1}X'Z(Z'Z)^{-1}Z'y$ Special case of GMM when $W_n=(Z'Z)^{-1}$.

Consider $Z(Z'Z)^{-1}Z'$ to be "Z"'s Projection Matrix. $P\cdot P=P$

$\hat\beta_{2SLS}=(X'PX)^{-1}(X'Py)=(X'P'PX)^{-1}(X'P'y)=(\hat X'\hat X)^{-1}(\hat X'y)$

First Stage: $X = Z\Gamma+U$ Second Stage: $y=\hat X\beta+e$ where $\hat X = Z\hat\Gamma = PX$.

When homoskedasticity, $\Omega=E(z_iz_i'e_i^2)=E[E(z_iz_i'e_i^2|z)]=E[z_iz_i']\cdot\sigma^2$, efficient GMM has $W_n=(Z'Z)^{-1}$, ∴ $\hat\beta_{2SLS}$ is efficient.

①当$W_n=\hat Q_{xx}$, $\tilde\beta_{md}=\tilde\beta_{cls}$
②当$W_n=\hat V_\beta^{-1}$, $\tilde\beta_{md}=\tilde\beta_{emd}$
③当$W_n=\Omega^{-1}$, $\hat\beta_{GMM}$ is efficient
④当$l=k$, $\hat\beta_{GMM}=\hat\beta_{ILS}$
⑤当homo, $\hat\beta_{GMM}=\hat\beta_{2SLS}$ $(W_n=(Z'Z)^{-1})$
⑥约束线性且EMD时，$J_n^*=W_n$.