Econometrics Notesheet

Michael B. Nattinger

December 11, 2020

1 Stats and Expectations

- LIE: $E[E[Y|X]] = E[Y], E[E[Y|X,Z]|X] = E[Y|X]$
- Conditioning thm: $E[g(X)Y|X] = g(X)E[Y|X], E[g(X)Y] = E[g(X)E[Y|X]]$
- CEF ($m(x) = E[Y|X = x]$) error $e = Y - m(X), E[e|X] = 0 = E[e], E[h(X)e] = 0$.
- conditional mean best predictor: $E[(Y - g(X))^2] \geq E[(Y - m(X))^2]$
- best linear predictor of Y given X $X'\beta$ minimizes $E[(Y - X'\beta)^2]$
- $\beta = (E[XX'])^{-1}E[XY]$
- For $Y = X'\beta + \alpha + e, \alpha = \mu_Y - \mu_X'\beta, \beta = var(X)^{-1}cov(X,Y)$

2 OLS

- $\hat{\beta}_{ols} = (X'X)^{-1}X'Y$ minimizes $\frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i'\beta)^2$
- $X'\hat{e} = 0, P = X(X'X)^{-1}X', M = I_n - X(X'X)^{-1}X', \hat{e} = MY = Me$
- The OLS estimator for slope coefficients is a regression with demeaned data.
- $Y'Y = \hat{Y}'\hat{Y} + \hat{e}'\hat{e}$ so (w/ constant) $(Y - \bar{Y})'(Y - \bar{Y}) = (\hat{Y} - \bar{Y})'(\hat{Y} - \bar{Y}) + \hat{e}'\hat{e}$.
- $\bar{R}^2 = 1 - \frac{(n-1)^{-1}\hat{e}'\hat{e}}{(Y-\bar{Y})'(Y-\bar{Y})}$
- $Y = \beta_1 X_1 + \beta_2 X_2 + e \Rightarrow \hat{\beta}_1 = (X_1'M_2X_1)^{-1}(X_1'M_2Y), \hat{\beta}_2 = (X_2'M_1X_2)^{-1}(X_2'M_1Y)$
- $\beta_1 = (X_1'X_1)^{-1}X_1'(Y - X_2\beta_2)$

3 OLS Features

- (D var-cov mat of e_i's) $V_{\hat{\beta}} = (X'X)^{-1}X'DX(X'X)^{-1}$
- with homoskedastic SE: $V_{\hat{\beta}} = \sigma^2(X'X)^{-1}$
- $var[\hat{\beta}] = E[(X'X)^{-1}X'DX(X'X)^{-1}]$
- Gauss-Markov $var[\tilde{\beta}|X] \geq \sigma^2(X'X)^{-1}$
- GLS: $Y = X\beta + e, E[e|X] = 0, var[e|X] = \Sigma.\tilde{\beta}_{gls} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$
- Generalized G-M: $var[\tilde{\beta}|X] \geq (X'\sigma^{-1}X)^{-1}$
- $\text{MSFE}_n = E[\tilde{e}^2_{n+1}] = \sigma^2 + E[X'_{n+1}V_{\hat{\beta}}X_{n+1}]$
- $\hat{V}_{\hat{\beta}} = (X'X)^{-1}s^2$
- Robust V: $\hat{V}_{\hat{\beta}HC0} = (X'X)^{-1}\left(\sum_{i=1}^{n} X_iX_i'\hat{e}_i^2\right)(X'X)^{-1}, \hat{V}_{\hat{\beta}HC1} = \frac{n}{n-k}\hat{V}_{\hat{\beta}HC0}$

4 Normal regression

In the normal regression model,

- $\hat{\beta}|X \sim N(\beta, \sigma^2(X'X)^{-1})$
- $\hat{e}|X \sim N(0, \sigma^2 M)$
- $\frac{(n-k)s^2}{\sigma^2} \sim \chi^2_{n-k}$ and is independent of $\hat{\beta}$
- T Stat: $T = \frac{\hat{\beta}_j - \beta_j}{s(\hat{\beta}_j)} = \frac{\hat{\beta}_j - \beta_j}{\sqrt{s^2[(X'X)^{-1}]_{j,j}}}$
- $T \sim t_{n-k}$ (Homoskedastic SE only)

5 Asymptotics

- CLT(m.v.) $\sqrt{n}(\bar{Y} - \mu) \to_d N(0, V), V = E[(Y-\mu)'(Y-\mu)]$
- D.M. $\sqrt{n}(g(\hat{\mu}) - g(\mu)) \to_d N(0, G'VG), G = G(\mu), G(u) = \frac{\partial}{\partial u}g(u)'$.

6 Asymptotics for Least Squares

- $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i e_i \rightarrow_d N(0, \Sigma), \Sigma = E[XX'e^2]$
- $\sqrt{n}(\hat{\beta} - \beta) = (\sum_{i=1}^{n} X_i X_i')' \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i e_i \right) \rightarrow_d N(0, V_\beta)$
- where $V_\beta = Q_{XX}^{-1} \Sigma Q_{XX}^{-1}$
- $\sqrt{n}(\hat{\theta} - \theta) \rightarrow_d N(0, V_\theta), V_\theta = R' V_\beta R, R = \frac{\partial}{\partial \beta} r(\beta)'$
- $s(\hat{\theta}) = \sqrt{\hat{R}' \hat{V}_\beta \hat{R}}$
- T stat (asymptotic) $T(\theta) = \frac{\hat{\theta} - \theta}{s(\theta)} \sim N(0, 1)$
- Wald $W(\theta) = (\hat{\theta} - \theta)' \hat{V}_{\hat{\theta}}^{-1} (\hat{\theta} - \theta) = n(\hat{\theta} - \theta)' \hat{V}_\theta^{-1} (\hat{\theta} - \theta)$

7 Restricted Estimation

7.1 CLS

- CLS: $\tilde{\beta}_{cls} = \arg\min_{R'\beta = c} SSE(\beta)$
- $\frac{\partial}{\partial \beta} \mathcal{L} = -X'Y + X'X \tilde{\beta}_{cls} + R \tilde{\lambda}_{cls} = 0$
- $\frac{\partial}{\partial \lambda} \mathcal{L} = R' \tilde{\beta} - c = 0$
- premultiply first foc: $-R' \hat{\beta} + R' \tilde{\beta}_{cls} + R'(X'X)^{-1} R \tilde{\lambda}_{cls} = 0$
- impose $R' \tilde{\beta}_{cls} = c : \tilde{\lambda}_{cls} = [R'(X'X)^{-1} R]^{-1} (R' \hat{\beta} - c)$
- $\tilde{\beta}_{cls} = \hat{\beta}_{ols} - (X'X)^{-1} R [R'(X'X)^{-1} R]^{-1} (R' \hat{\beta}_{ols} - c)$
- If finite sample properties of CLS come up, see section 8.4 (p.200 - p. 203).

7.2 Minimum Distance

- $J(\beta) = n(\hat{\beta} - \beta)' \hat{W} (\hat{\beta} - \beta)$
- $\tilde{\beta}_{md} = \arg\min_{R'\beta = c} J(\beta)$
- CLS is special case of this when $\hat{W} = \hat{Q}_{XX}$
- $\tilde{\lambda}_{md} = n(R' \hat{W} R)^{-1} (R' \hat{\beta} - c)$
- $\tilde{\beta}_{md} = \hat{\beta} - \hat{W}^{-1} R (R' \hat{W}^{-1} R)^{-1} (R' \hat{\beta} - c)$
- If asymptotics come up see section 8.6 (p. 204 - 205), for variance estimation of cls see 8.7 (206), for variance estimatin of emd see 8.10 (p. 209)

- The asymptotically optimal weight matric is the one which minimizes the asymptotic variance. This is $W = V_\beta^{-1}$.
- $\tilde{\beta}_{emd} = \hat{\beta} - \hat{V}_\beta R(R'\hat{V}_\beta R)^{-1}(R'\hat{\beta} - c)$
- $\sqrt{n}(\tilde{\beta}_{emd} - \beta) \to_d N(0, V_{\beta,emd})$
- $V_{\beta,emd} = V_\beta - V_\beta R(R'V_\beta R)^{-1}R'V_\beta$
- Hausman equality: $avar[\hat{\beta}_{ols} - \tilde{\beta}_{emd}] = avar[\hat{\beta}_{ols}] - avar[\tilde{\beta}_{emd}]$
- nonlinear constraints: no general closed form soln so may have to set up and solve lagrangian with objective functions for CLS (argmin SSE) or MD (argmin J)
- Asymptotics in section 8.14 (p. 217)

8 Hypothesis Testing

- $T = |T(\theta_0)|, T(\theta) = \frac{\hat{\theta}-\theta}{s(\hat{\theta})}$
- $T(\theta_0) \to_d N(0,1)$ under the null
- Multidimensional test statistic is wald $W = W(\theta_0) = (\hat{\theta} - \theta_0)'\hat{V}_{\hat{\theta}}^{-1}(\hat{\theta} - \theta_0)$
- $\hat{V}_{\hat{\theta}} = \hat{R}'\hat{V}_{\hat{\beta}}\hat{R}, R = \frac{\partial}{\partial\beta}r(\hat{\beta})'$
- Linear: $W = (R'\hat{\beta} - \theta_0)'(R'\hat{V}_{\hat{\beta}}R)^{-1}(R'\hat{\beta} - \theta_0)$
- $W \to_d \chi^2_q$ where q is the number of restrictions tested.
- Criterion-based tests (discrepancy between the criterion function minimized with and without the restriction).
- $J = \min_{\beta \in B_0} J(\beta) - \min_{\beta \in B} J(\beta)$ distance statistic.
- for md: test stat is $J = J(\tilde{\beta}_{md})$ (with appropriate $\hat{W}$)
- For emd, $J = W \sim \chi^2_q$ (wald statistic)