Linear Regression with weakly dependent data

Consider the usual regression model

$$y_t = x_t'\beta + u_t$$

Consistency

(*) $\{x_t u_t\}$ and $\{x_t x_t'\}$ are also α-mixing of the same size, by Proposition 1.

Provided

(a) $\{(x_t', u_t)\}$ is α-mixing of any size

(b) $\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} x_t u_t = 0$

(c) $E|x_{tj}|^{2+\eta} < \Delta$ for all t and $j = 1, \dots, k$ and some $\eta > 0$

(d) $E|u_t|^{2+\eta} < \Delta$ for some $\eta > 0$

(e) $M_n = \frac{1}{n}\sum_{t=1}^{n} E x_t x_t'$ is uniformly positive definite over n.

Then $\hat{\beta}_n - \beta = o_p(1)$

proof:

We start from the moment condition

$$0 = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E x_t u_t = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E x_t (y_t - x_t'\beta)$$

$$\Rightarrow \quad \beta = \left(\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E x_t x_t'\right)^{-1} \left(\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E x_t y_t\right)$$

Sample analogue: drop the "E"!

$$\hat{\beta}_n = \left(\frac{1}{n}\sum_{t=1}^{n} x_t x_t'\right)^{-1} \left(\frac{1}{n}\sum_{t=1}^{n} x_t y_t\right)$$

$$= \beta + \left(\frac{1}{n}\sum_{t=1}^{n} x_t x_t'\right)^{-1} \frac{1}{n}\sum_{t=1}^{n} x_t u_t$$

$$= \beta + \left(\frac{1}{n}\sum_{t=1}^{n} x_t x_t' - M_n + M_n\right)^{-1} \frac{1}{n}\sum_{t=1}^{n} x_t u_t$$

$E\|x_t x_t'\|^{1+\eta} \leq (E\|x_t\|^{2+2\eta} E\|x_t\|^{2+2\eta})^{1/2}$ by Cauchy-Schwarz

WLLN $\{x_t x_t'\}$

$$= \beta + \left(o_p(1)\, M_n^{-1} + M_n M_n^{-1}\right)^{-1} M_n^{-1} \frac{1}{n}\sum_{t=1}^{n} x_t u_t$$

$$= \beta + \left(o_p(1)\, O(1) + I_k\right)^{-1} O(1) \frac{1}{n}\sum_{t=1}^{n} x_t u_t$$

$$= \beta + [I_K + o_p(1)]\, O(1)\, \frac{1}{n}\sum_{t=1}^{n} X_t u_t$$

$$= \beta + [I_K + o_p(1)]\, O(1) \left[\frac{1}{n}\sum_{t=1}^{n} X_t u_t - \frac{1}{n}\sum_{t=1}^{n} EX_t u_t + \frac{1}{n}\sum_{t=1}^{n} EX_t u_t \right]$$

$E\|X_t u_t\|^{1+\eta} \le (E\|X_t\|^{2+2\eta} E|u_t|^{2+2\eta})^{1/2}$ by Cauchy-Schwarz

WLLN $\{X_t u_t\}$

$$= \beta + [I_K + O_p(1)]\, O(1) [\, O_p(1) + o(1)\,]$$

$$= \beta + I_K\, O(1)\, o_p(1) + o_p(1)\, O(1)\, o_p(1)$$

$$= \beta + o_p(1).$$

■

<u>Asymptotic Normality</u>

⊛ Reminder: conditions for CLT in Proposition 6

(i) α-mixing size $-p/p-2$

(ii) $E\|X_{nt}\|^p \le \Delta$ for all t, n

(iii) $\Omega_n = Var\left(\frac{1}{\sqrt{n}}\sum X_{nt}\right)$ is uniformly p.d. and $p > 2$.

Provided

(a) $\{(X_t' u_t)\}$ is α-mixing of size $-p/p-2$

(b) $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} EX_t u_t = o(1)$

(c) $E|X_{tj}|^{2p} \le \Delta$ for all t and $j = 1, \dots, K$

(d) $E|u_t|^{2p} \le \Delta$ for all t

(e) $M_n = \frac{1}{n}\sum_{t=1}^{n} EX_t X_t'$ is uniformly positive definite over n

(f) $\Omega_n = Var\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} X_t u_t\right)$ is uniformly positive definite over n

Then

$$V_n^{-1/2}\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{d} N(0, I_K)$$

where $V_n = M_n^{-1}\Omega_n M_n^{-1}$.

<u>proof</u>:

From the previous proposition we get

$$\sqrt{n}(\hat{\beta}_n - \beta) = (I_K + o_p(1))\, M_n^{-1} \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \{X_t u_t - EX_t u_t\} + o_p(1)$$

$$= M_n^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\{X_t u_t - EX_t u_t\} + o_p(1)\, O(1) \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\{X_t u_t - EX_t u_t\} + o_p(1)$$

$$= \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \underbrace{M_n^{-1}\{X_t u_t - EX_t u_t\}}_{\xi_{nt}} + o_p(1)\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\{X_t u_t - EX_t u_t\} + o_p(1)$$

$$= \underbrace{\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \xi_{n,t}}_{A} + \underbrace{o_p(1) \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{x_t u_t - E x_t u_t\}}_{B} + o_p(1)$$

- We will deal with B first. We're interested in the process $\{x_t u_t - E x_t u_t\}$, so we check for the conditions

(i) $\{x_t u_t - E x_t u_t\}$ is a measurable function of $\{x_t', u_t\}$ so by Proposition 1 this is α-mixing of size $-p/p-2$.

(ii) $$E\| x_t u_t - E x_t u_t \|^p \underset{\text{Minkowski's Inequality}}{\leq} \{ (E\| x_t u_t\|^p)^{1/p} + (E\| x_t u_t\|^p)^{1/p} \}^p$$

$$\underset{\text{Cauchy-Schwarz}}{\leq} 2^p \, (E\|x_t\|^{2p})^{1/2} \, (E\|u_t\|^{2p})^{1/2}$$

$$< \infty$$

(iii) $\mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} [x_t u_t - E x_t u_t] \right) = \mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_t u_t \right) = \Omega_n$ which is uniformly p.d. by assumption.

Then, by the CLT $\quad \Omega_n^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{x_t u_t - E x_t u_t\} = O_p(1)$.

We write B as:

$$o_p(1)\, \Omega_n^{-1/2} \Omega_n \, \Omega_n^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{x_t u_t - E x_t u_t\} = o_p(1)\, \Omega_n^{-1/2} \Omega_n \, O_p(1)$$

$$= o_p(1) \, O(1) \, O_p(1)$$

$$= o_p(1).$$

- Now we can deal with A. We're interested in the array $\{ M_n^{-1} (x_t u_t - E x_t u_t) \}$, so we will check for the conditions

(i) $\{ M_n^{-1} (x_t u_t - E x_t u_t) \}$ is a measurable function of $\{x_t', u_t\}$ so by Proposition 1 it's α-mixing of size $-p/p-2$.

(ii) $$E\| M_n^{-1} (x_t u_t - E x_t u_t) \| \leq \|M_n^{-1}\| \; E\| x_t - E x_t u_t\|$$

$$\leq O(1) \quad O(1)$$

$$< \infty$$

(iii) $\mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} M_n^{-1} (x_t u_t - E x_t u_t) \right) = M_n^{-1'} \Omega_n M_n^{-1}$ must be uniformly positive definite.

This requires that for arbitrary x s.t. $\|x\| = 1$

$$x' M_n^{-1'} \Omega_n M_n^{-1} x > 0$$

$$x' M_n^{-1} \Omega_n M_n^{-1} x = x' M_n^{-1} C_n \Delta_n C_n M_n^{-1} x$$

$$= y_n' \Delta_n y_n$$

$$= \sum_{i=1}^{K} e_{ni} y_{ni}^2$$

$$\geq \underline{e_n} \, \| y_n \|^2$$

$$\geq \delta \, x' M_n^{-1} C_n C_n' M_n^{-1} x$$

$$= \delta \, x' M_n^{-2} x$$

$$= \delta \, x' D_n P_n^{-2} D_n x$$

$$= \delta \sum_{i=1}^{K} d_{ni}^{-2} w_{ni}^2$$

$$\geq \frac{\delta}{\bar{d}_n^2} \| x' D_n \|$$

$$= \frac{\delta}{\bar{d}_n^2}$$

> 0 because M_n uniformly p.d. $\sup_n \bar{d}_n < K$ for some $K > 0$.

Then, by the CLT $\quad V_n^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} M_n^{-1} (x_t u_t - E x_t u_t) \} \xrightarrow[d]{} N(0,1)$.

Putting it all together yields:

$$V_n^{-1/2} \sqrt{n} (\hat{\beta}_n - \beta) = V_n^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} M_n^{-1} (X_t U_t - E X_t U_t) + V_n^{-1/2} o_p(1)$$

$$V_n^{-1/2} \sqrt{n} (\hat{\beta}_n - \beta) = V_n^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} M_n^{-1} (x_t u_t - E x_t u_t) \} + O(1) \, o_p(1)$$

$$= V_n^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} M_n^{-1} (x_t u_t - E x_t u_t) \} + o_p(1)$$

$$\xrightarrow[d]{} N(0, I_K). \quad \blacksquare$$

Estimation of Asymptotic Variance Matrix

Recall that $V_n = M_n^{-1} \Omega_n M_n^{-1}$ and $M_n = \frac{1}{n}\sum_{t=1}^{n} E X_t X_t'$. Then we can estimate M_n using

$$\hat{M_n} = \frac{1}{n}\hat{\sum} X_t X_t' \quad \text{and hope that} \quad \hat{M_n} - M_n = o_p(1).$$

To estimate Ω_n we need $\Omega_n(h) = \frac{1}{n}\sum_{t=h+1}^{n} E[\underbrace{X_t u_t (X_{t-h} u_{t-h})'}_{\text{because they are centered at 0.}}]$.

Now, our initial estimator could be

$$\tilde{\Omega_n} = \tilde{\Omega_n}(0) + \sum_{h=1}^{n-1} (\tilde{\Omega_n}(h) + \tilde{\Omega_n}(h'))$$

$$\text{where} \quad \tilde{\Omega_n}(h) = \frac{1}{n}\sum_{t=h+1}^{n}[X_t u_t (X_{t-h} u_{t-h})']$$

- Problem: we need to ensure that $(\tilde{\Omega_n}(h) + \tilde{\Omega_n}(h'))$ grow slower than n. A solution would be to allow for autocovariances to grow slower than n.

- New problem: when we truncate we can get non positive definite matrix, so we need to put weights in the sum.

The (infeasible) HAC estimator of variance is

$$\tilde{\Omega_n} = \tilde{\Omega_n}(0) + \sum_{j=1}^{m_n} w(j, m_n)\, (\tilde{\Omega_n}(j) + \tilde{\Omega_n}(j'))$$

Proposition HAC 1.- Suppose that for some $r > 2$ and $\Delta, \delta, C > 0$

(a) $\{(X_t', u_t)\}$ is α-mixing of size $-r/r-2$

(b) $E X_t u_t = 0$ for all t

(c) $E|X_{tj}|^{4r+\delta} \le \Delta$ for all t and all $j = 1, \dots, K$

(d) $E|u_t|^{4r+\delta} \le \Delta$ for all t

(e) $|w(j,m)| \le C$ for all j and m

(g) $\lim_{m\to\infty} w(j,m) = 1$ for all j

(h) $m_n = o(n^{1/4})$.

Then

$$\tilde{\Omega_n} - \Omega_n = o_p(1).$$

Proof: By the Cramer-Wold device it suffices to show that

$$c'(\hat{\Omega}_n - \Omega_n)c = o_p(1) \quad \text{for all } c \in \mathbb{R}^K.$$

Now, define $h_t = c' X_t u_t$ and notice that by Proposition 1 it is α-mixing of size $-p/p-2$. Then we write

$$c'(\hat{\Omega}_n - \Omega_n)c = \underbrace{\frac{1}{n}\sum_{t=1}^{n}(h_t - Eh_t)}_{R_{n,0} = o_p(1) \text{ by LLN}} + \underbrace{2\sum_{j=1}^{m_n} w(j,m_n)\frac{1}{n}\sum(h_t h_{t-j} - Eh_t h_{t-j})}_{R_{n,1} := \text{ regular estimation error of covariances}} + \underbrace{2\sum_{j=1}^{m_n}(w(j,m_n)-1)\frac{1}{n}\sum_{t=h+1}^{n} Eh_t h_{t-j}}_{R_{n,2} := \text{ bias due to using weights}}$$

$$\underbrace{-\,2\sum_{j=m_n+1}^{n-1}\frac{1}{n}\sum_{t=h+1}^{n} Eh_t h_{t-j}}_{R_{n,3} := \text{ bias due to truncation of autocovariances}}$$

- $R_{n,2}$: We want to use the covariance inequalities, so we need to show that $\sup_t E|h_t|^p < \infty$. To see this

$$E|h_t|^p \le \|c\| \; E\|X_t u_t\|^p$$

$$\le \|c\| \; (E\|X_t\|^{2p})^{1/2} \; (E\|u_t\|^{2p})^{1/2}$$

$$< \infty$$

Then

$$|R_{n,2}| \le \sum_{j=1}^{m_n} |w(j,m_n) - 1| \; \frac{1}{n}\sum_{t=h+1}^{n} Eh_t h_{t-j}$$

$$\underset{\text{Proposition 4}}{\le} \sum_{j=1}^{m_n} |w(j,m_n) - 1| \; \frac{1}{n}\sum_{t=h+1}^{n} K \; \alpha(j)^{1-2/p}$$

$$= \sum_{j=1}^{m_n} |w(j,m_n) - 1| \; \frac{1}{n}\sum_{t=h+1}^{n} K \; j^{(-\frac{p}{p-2} - \varepsilon)(\frac{p-2}{p})}$$

$$= \sum_{j=1}^{m_n} |w(j,m_n) - 1| \; \frac{1}{n}\sum_{t=h+1}^{n} K \; j^{-1-\eta} \qquad \text{for some } \eta > 0$$

$$\le \sum_{j=1}^{m_n} |w(j,m_n) - 1| \; \frac{1}{n} K j^{-1-\eta} \; (n - (h+1) + 1)$$

$$\le K \sum_{j=1}^{\infty} |w(j,m_n) - 1| \; j^{-1-\eta}$$

$$\lim_{n\to\infty} |R_{n,2}| \le \lim_{n\to\infty} K \sum_{j=1}^{\infty} |w(j,m_n) - 1| \; j^{-1-\eta}$$

$$= K \sum_{j=1}^{\infty} \left| \lim_{n\to\infty} w(j,m_n) - 1 \right| j^{-1-\eta}$$

Dominated Convergence Theorem

$K(C+1)\, j^{-1-\eta}$ can be the dominating function and it's integrable/summable.

$$= 0 .$$

- $R_{n,3}$: We will use the same idea as in the previous part.

$$|R_{n,3}| \leq \sum_{j=m_n+1}^{n-1} \frac{1}{n} \sum_{t=1}^{n} |E h_t h_{t-j}|$$

$$\underset{\text{Using bounds computed in } R_{n,2}}{\leq} K \sum_{j=m_n+1}^{n-1} j^{-1-\eta}$$

$$\leq -\frac{1}{\eta} n^{-\eta} K + \frac{K}{\eta} m_n^{-\eta}$$

$$\lim_{n\to\infty} |R_{n,3}| \leq \lim_{n\to\infty} -\frac{K}{\eta} n^{-\eta} + \lim_{n\to\infty} \frac{K}{\eta} m_n^{-\eta}$$

$$= 0 .$$

- $R_{n,1}$: before we deal with this I will write this term again to see why this can be difficult to check.

$$R_{1,n} := \sum_{j=1}^{m_n} w(j,m_n) \frac{1}{n} \sum_{t=j+1}^{n} \underbrace{(h_t h_{t-j} - E h_t h_{t-j})}$$

Call this process Z_{jt}. Moreover, notice that $Z_{jt} = g(h_t, h_{t-j})$ so by Proposition 1 it is α-mixing and $\alpha_{Z_j}(\ell) \leq \alpha_h(\ell - j)$ for all $\ell = j+1, j+2, \dots$

You will see why this is important later. Mark this as ⊛.

This number must be positive!

We want to show that the object is $o_P(1)$, so we write

$$P\left(\left| \sum_{j=1}^{m_n} \underbrace{w(j,m_n)}_{\leq C} \frac{1}{n} \sum_{t=j+1}^{n} Z_{jt} \right| > \varepsilon \right) \leq P\left(\sum_{j=1}^{m_n} \left| \sum_{t=j+1}^{n} Z_{jt} \right| > \frac{\varepsilon \cdot n}{C} \right)$$

$$\leq P\left(\left| \sum_{t=1+1}^{n} Z_{1t} \right| > \frac{\varepsilon \cdot n}{C \cdot m_n} \right) + \dots + P\left(\left| \sum_{t=m_n+1}^{n} Z_{m_n t} \right| > \frac{\varepsilon \cdot n}{C \cdot m_n} \right)$$

$$= \sum_{j=1}^{m_n} P\left(\left| \sum_{t=j+1}^{n} Z_{jt} \right| > \frac{\varepsilon \cdot n}{C \cdot m_n} \right)$$

$$\underset{\text{Markov's Inequality}}{\leq} \sum_{j=1}^{m_n} \frac{C^2 m_n^2}{\varepsilon^2 n^2} E \left| \sum_{t=j+1}^{n} Z_{jt} \right|^2$$

Claim 1.- If $E|\hat{\sum} Z_{jt}|^2 \le K\cdot n\cdot(j+2)$ then $R_{1,n} = o_p(1)$.

Using this claim we get

$$\le \sum_{j=1}^{m_n} \frac{c^2 m_n^2}{\varepsilon^2 n^2} K\cdot n\cdot (j+2)$$

$$= K \frac{c^2 m_n^2}{n\,\varepsilon^2}\sum_{j=1}^{m_n}(j+2)$$

$$\frac{1}{m_n}\sum_{j=1}^{m_n}(j+2) = \frac{\text{first} + \text{last}}{2}$$

$$= K \frac{c^2 m_n^2}{n\,\varepsilon^2}\left[\frac{(m_n+2)+3}{2}\right] m_n$$

$$\le K \frac{m_n^4}{n}$$

$$= K\frac{1}{n}\, o(n)$$

$$= o(1).$$

To finish the proof we only need to show that the assumption for Claim 1 is true. write

$$E\Big|\sum^{n} Z_{jt}\Big|^2 \underset{\text{Variance formula}}{=} \sum_{t=j+1}^{n} E Z_{jt}^2 + 2\sum_{\ell=1}^{n-j-1}\sum_{t=\ell+j+1}^{n} E(Z_{jt} Z_{jt-\ell})$$

$$\le \sum_{t=j+1}^{n} \sup_t E|h_t|^4 + 2\sum_{\ell=1}^{n-j-1}\sum_{t=\ell+j+1}^{n} E(Z_{jt} Z_{jt-\ell})$$

$$\underset{\text{Cauchy Schwarz}}{\le} \sum_{t=j+1}^{n} \left(\sup_t E|\varepsilon_t|^8 \cdot \sup_t E|c'x_t|^8\right)^{1/2} + 2\sum_{\ell=1}^{n-j-1}\sum_{t=\ell+j+1}^{n} E(Z_{jt} Z_{jt-\ell})$$

$$\le K\cdot n + 2\sum_{\ell=1}^{n-j-1}\sum_{t=\ell+j+1}^{n} E(Z_{jt} Z_{jt-\ell})$$

split the sum

$$= K\cdot n + 2\underbrace{\sum_{\ell=1}^{j}\sum_{t=\ell+j+1}^{n} |E Z_{jt} Z_{jt-\ell}|}_{\text{we cannot use mixing coefficient properties here but we can use Cauchy Schwarz, recall (*)}} + 2\underbrace{\sum_{\ell=j+1}^{n-j-1}\sum_{t=j+1+\ell}^{n} |E Z_{jt} Z_{jt-\ell}|}_{\text{we can use mixing coefficients properties here because } \ell \ge j+1\text{, recall (*)}}$$

Cauchy Schwarz + Proposition 4

$$\le K\cdot n + 2\sum_{\ell=1}^{j}\sum_{t=\ell+j+1}^{n} \big(\underbrace{E|Z_{jt}|^2\, E|Z_{jt-\ell}|^2}_{\text{we just showed these are bounded}}\big)^{1/2} + 2\sum_{\ell=1}^{n-j-1}(n-\ell-j)\,\ell^{-1-\eta}$$

$$\le Kn + K\cdot n\cdot j + 2n\sum_{\ell=1}^{n-j-1}\ell^{-1-\eta}$$

By summability

$$\le K\cdot n + K\cdot n\cdot j + K\cdot n = K\cdot n(j+2). \quad \blacksquare$$