时间序列分析

Time Series Analysis

Lecture 2

Review: Characteristics of TS

- Stationary
- Non-stationary
- Seasonality
- Level Shift
- Variance Changes
- Intervention
- Bivariate Leading-Lagging Relationship
- ...
- GOAL: model these different patterns for prediction (forecasting)

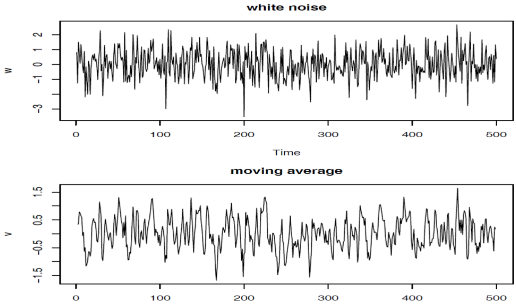
Today's Topics

1. Examples of Time Series Models
2. Measure of Time Series Dependence
3. Stationarity, Ergodicity, LLN and CLT
4. Estimation of Covariances
5. Estimation of Long-run Covariances
6. Nonstationarity
7. Classical Regression with TS Data
8. Time Regressions

1. Time Series Statistical Models

- White Noise
- Moving Average
- Autoregression
- Random Walk
- ARIMA

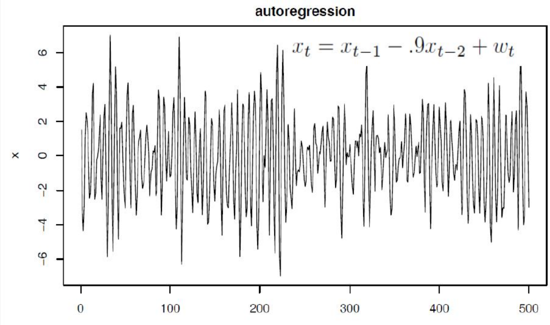
White Noise and Moving Average




```
w = rnorm(500,0,1)                       # 500 N(0,1) variates
v = filter(w, sides=2, rep(1/3,3))      # moving average
par(mfrow=c(2,1))
plot.ts(w, main="white noise")
plot.ts(v, main="moving average")
```

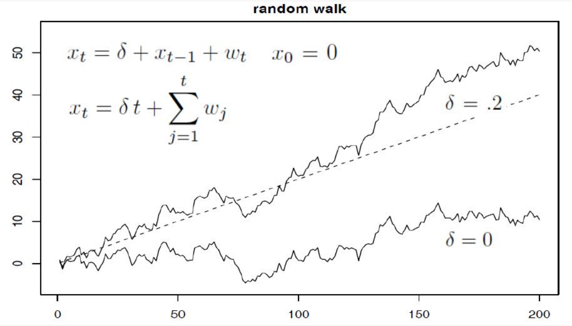
$$v_t = \frac{1}{3}\left(w_{t-1} + w_t + w_{t+1}\right)$$

Autoregression




```
w = rnorm(550,0,1)  # 50 extra to avoid startup problems
x = filter(w, filter=c(1,-.9), method="recursive")[-(1:50)]
plot.ts(x, main="autoregression")
```

Random Walk




```
set.seed(154)                    # so you can reproduce the results
w = rnorm(200,0,1);  x = cumsum(w)    # two commands in one line
wd = w +.2;    xd = cumsum(wd)
plot.ts(xd, ylim=c(-5,55), main="random walk")
lines(x); lines(.2*(1:200), lty="dashed")
```

2. Measures of Time Series Dependence

- Autocovariance function

$$\gamma_x(s,t) = cov(x_s, x_t) = E[(x_s - \mu_s)(x_t - \mu_t)]$$

- Autocorrelation function (ACF)

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s), \gamma(t,t)}}$$

- Cross-covariance function

$$\gamma_{xy}(s,t) = cov(x_s, y_t) = E[(x_s - \mu_{xs})(y_t - \mu_{yt})]$$

- Cross-correlation function (CCF)

$$\rho_{xy}(s,t) = \frac{\gamma_{xy}(s,t)}{\sqrt{\gamma_x(s,s), \gamma_y(t,t)}}$$

3. Stationarity, Ergodicity, LLN and CLT

- Strict Stationarity
- Weak/covariance Stationarity
- Ergodicity
- Examples of Stationary Processes
- Law of Large Numbers and Central Limit Theorems for Ergodic Stationary Processes

Strict Stationarity

A stochastic process $\{\mathbf{z}_i\}$ $(i = 1, 2, 3, \cdots)$ is (strictly) stationary if, for any given finite integer r and for any set of subscripts, $i_1, i_2, \cdots, i_r$, the joint distribution of $(\mathbf{z}_i, \mathbf{z}_{i_1}, \mathbf{z}_{i_2}, \cdots, \mathbf{z}_{i_r})$ depends only on $i_1 - i, i_2 - i, i_3 - i, \cdots, i_r - i$ but not on i.

- Trend stationary and difference stationary

Weak/Covariance Stationarity

A stochastic process $\{\mathbf{z}_i\}$ is weakly (or covariance) stationary if:

(i) $E(\mathbf{z}_i)$ does not depend on i, and

(ii) $Cov(\mathbf{z}_i, \mathbf{z}_{i-j})$ exists, is finite, and depends only on j but not on i (for example, $Cov(\mathbf{z}_1, \mathbf{z}_5)$ equals $Cov(\mathbf{z}_{12}, \mathbf{z}_{16})$).

• The relative, not absolute, position in the sequence matters for the mean and covariance of a covariance-stationary process. Evidently, if a sequence is (strictly) stationary and if the variance and covariances are finite, then the sequence is weakly stationary (hence the term "strict").

Autocovariances and Autocorrelation

The j-th order autocovariance, denoted $\mathbf{\Gamma}_j$, is defined as

$$\begin{aligned}\mathbf{\Gamma}_j &\equiv Cov(\mathbf{z}_i, \mathbf{z}_{i-j}) \quad (j = 0, 1, 2, \cdots). \\ \mathbf{\Gamma}_j &= \mathbf{\Gamma}'_{-j}.\end{aligned}$$

Autocovariance matrix:

$$Var(z_i, z_{i+1}, \cdots, z_{i+n-1}) = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{n-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \gamma_{n-2} & \cdots & \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_{n-1} & \cdots & \gamma_2 & \gamma_1 & \gamma_0 \end{pmatrix}$$

j-th order autocorrelation coefficient

$$\equiv \rho_j \equiv \frac{\gamma_j}{\gamma_0} = \frac{Cov(z_i, z_{i-j})}{Var(z_i)} \quad (j = 1, 2, \cdots)$$

White Noise

A very important class of weakly stationary processes is a white noise process, a process with zero mean and no serial correlation:

$$\text{stationary process } \{\mathbf{z}_i\} \text{ is white noise if}$$
$$E(\mathbf{z}_i) = 0 \text{ and } Cov(\mathbf{z}_i, \mathbf{z}_{i-j}) = 0 \text{ for } j \neq 0.$$

Example 2.4 (a white noise process that is not strictly stationary[6]): Let w be a random variable uniformly distributed in the interval $(0, 2\pi)$, and define

$$z_i = \cos(iw) \quad (i = 1, 2, \dots).$$

It can be shown that $\mathrm{E}(z_i) = 0$, $\mathrm{Var}(z_i) = 1/2$, and $\mathrm{Cov}(z_i, z_j) = 0$ for $i \neq j$. So $\{z_i\}$ is white noise. However, clearly, it is not an independent white noise process. It is not even strictly stationary.

Ergodicity and LLN

A stationary process $\{z_i\}$ is said to be ergodic if, for any two bounded functions $f : \mathbb{R}^k \to \mathbb{R}$ and $g : \mathbb{R}^l \to \mathbb{R}$,

$$\lim_{n\to\infty} \left| E[f(z_i, \cdots, z_{i+k}) g(z_{i+n}, \cdots, z_{i+n+l})] \right|$$
$$= \lim_{n\to\infty} \left| E[f(z_i, \cdots, z_{i+k})] \right| \left| E[g(z_{i+n}, \cdots, z_{i+n+l})] \right|$$

Ergodic Theorem: (See, e.g., Theorem 9.5.5 of Karlin and Taylor (1975).) Let $\{z_i\}$ be a stationary and ergodic process with $E(z_i) = \mu$. Then,

$$\bar{z}_n \equiv \frac{1}{n} \sum_{i=1}^{n} z_i \overset{a.s.}{\to} \mu.$$

Martingales

Let x_i be an element of $\mathbf{z}_i$. The scalar process $\{x_i\}$ is called a martingale with respect to $\{\mathbf{z}_i\}$ if

$$E(x_i|\mathbf{z}_{i-1}, \mathbf{z}_{i-2}, \cdots, \mathbf{z}_1) = x_{i-1} \textit{ for } i \geq 2.$$

The conditioning set $(\mathbf{z}_{i-1}, \mathbf{z}_{i-2}, \cdots, \mathbf{z}_1)$ is often called the information set at point (date) $i-1$. $\{x_i\}$ is called simply a martingale if the information set is its own past values $(x_{i-1}, \cdots, x_1)$. If $\mathbf{z}_i$ includes x_i, then x_i is a martingale, because

$$\begin{aligned} E(x_i|x_{i-1}, \cdots, x_1) &= E[E(x_i|\mathbf{z}_{i-1}, \cdots, \mathbf{z}_1)|x_{i-1}, \cdots, x_1] \\ &= E(x_{i-1}|x_{i-1}, \cdots, x_1) = x_{i-1}. \end{aligned}$$

Example 2.5 (Hall's Martingale Hypothesis): Let $\mathbf{z}_i$ be a vector containing a set of macroeconomic variables (such as the money supply or GDP) including aggregate consumption c_i for period i. Hall's (1978) martingale hypothesis is that consumption is a martingale with respect to $\{\mathbf{z}_i\}$:

$$\mathrm{E}(c_i \mid \mathbf{z}_{i-1}, \mathbf{z}_{i-2}, \ldots, \mathbf{z}_1) = c_{i-1}.$$

This formalizes the notion in consumption theory called "consumption smoothing": the consumer, wishing to avoid fluctuations in the standard of living, adjusts consumption in date $i-1$ to the level such that no change in subsequent consumption is anticipated.

Random Walks

An important example of martingale is a random walk. Let $\{\mathbf{g}_i\}$ be a vector independent white noise process (so it is i.i.d. with mean 0 and finite variance matrix). A random walk, $\{\mathbf{z}_i\}$, is a sequence of cumulative sums:

$$\mathbf{z}_1 = \mathbf{g}_1, \ \mathbf{z}_2 = \mathbf{g}_1 + \mathbf{g}_2, \cdots, \mathbf{z}_i = \mathbf{g}_1 + \mathbf{g}_2 + \cdots + \mathbf{g}_i, \cdots$$

The first difference of a random walk is independent white noise.

Martingale Difference Sequences

A vector process $\{\mathbf{g}_i\}$ with $E(\mathbf{g}_i) = 0$ is called a martingale difference sequence (m.d.s.) or martingale differences if the expectation conditional on its past values, too, is zero:

$$E(\mathbf{g}_i | \mathbf{g}_{i-1}, \cdots, \mathbf{g}_1) = 0 \quad \textit{for} \quad i \geq 2.$$

A martingale difference sequence has no serial correlation.

PROOF. First note that we can assume, without loss of generality, that $j \geq 1$. Since the mean is zero, it suffices to show that $\mathrm{E}(\mathbf{g}_i \mathbf{g}'_{i-j}) = \mathbf{0}$. So consider rewriting it as follows.

$$\begin{aligned}
&\mathrm{E}(\mathbf{g}_i \mathbf{g}'_{i-j}) \\
&= \mathrm{E}[\mathrm{E}(\mathbf{g}_i \mathbf{g}'_{i-j} \mid \mathbf{g}_{i-j})] \quad \text{(by the Law of Total Expectations)} \\
&= \mathrm{E}[\mathrm{E}(\mathbf{g}_i \mid \mathbf{g}_{i-j}) \mathbf{g}'_{i-j}] \quad \text{(by the linearity of conditional expectations).}
\end{aligned}$$

Now, since $j \geq 1$, $(\mathbf{g}_{i-1}, \ldots, \mathbf{g}_{i-j}, \ldots, \mathbf{g}_1)$ includes $\mathbf{g}_{i-j}$. Therefore,

$$\begin{aligned}
&\mathrm{E}(\mathbf{g}_i \mid \mathbf{g}_{i-j}) \\
&= \mathrm{E}[\mathrm{E}(\mathbf{g}_i \mid \mathbf{g}_{i-1}, \ldots, \mathbf{g}_{i-j}, \ldots, \mathbf{g}_1) \mid \mathbf{g}_{i-j}] \quad \text{(by the Law of Iterated Expectations)} \\
&= \mathbf{0}.
\end{aligned}$$

ARCH Processes

An example of martingale differences, frequently used in analyzing asset returns, is an autoregressive conditional heteroskedastic (ARCH) process introduced by Engle (1982). A process $\{\mathbf{g}_i\}$ is said to be an ARCH process of order 1 (ARCH(1)) if it can be written as

$$\mathbf{g}_i = \sqrt{\zeta + \alpha \mathbf{g}_{i-1}^2} \cdot \varepsilon_i,$$

where $\{\varepsilon_i\}$ is i.i.d. with mean zero and unit variance.

It is easy to show that $\mathbf{g}_i$ is an m.d.s.

$$\begin{aligned} E(\mathbf{g}_i | \mathbf{g}_{i-1}, \mathbf{g}_{i-2}, \cdots, \mathbf{g}_1) &= 0 \\ E(\mathbf{g}_i^2 | \mathbf{g}_{i-1}, \mathbf{g}_{i-2}, \cdots, \mathbf{g}_1) &= \zeta + \alpha \mathbf{g}_{i-1}^2. \end{aligned}$$

Different Formulation of Lack of Serial Dependence

(1) $\{\mathbf{g}_i\}$ is independent white noise.

$\Rightarrow$ (2) $\{\mathbf{g}_i\}$ is stationary m.d.s. with finite variance.

$\Rightarrow$ (3) $\{\mathbf{g}_i\}$ is white noise.

The CLT for Ergodic Stationary M.D.S.

Ergodic Stationary Martingale Differences CLT (Billingsley, 1961): Let $\{\mathbf{g}_i\}$ be a vector martingale difference sequence that is stationary and ergodic with $E(\mathbf{g}_i\mathbf{g}_i') = \Sigma$, and let $\bar{\mathbf{g}} \equiv \frac{1}{n}\sum_{i=1}^{n}\mathbf{g}_i$. Then,

$$\sqrt{n}\bar{\mathbf{g}} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{g}_i \xrightarrow{d} N(0, \Sigma).$$

LLN for Covariance-Stationary Processes (Anderson, 1971)

(a) $\bar{\mathbf{y}} \equiv \frac{1}{n}\sum_{t=1}^{n} \mathbf{y}_t \overset{m.s.}{\rightarrow} \mu$ if each diagonal element of $\mathbf{\Gamma}_j$ goes to zero as $j \rightarrow \infty$, and

(b) $\lim_{n\rightarrow\infty} Var(\sqrt{n}\bar{\mathbf{y}})$ (which is the long-run covariance matrix of $\{\mathbf{y}_t\}$) equals $\sum_{j=-\infty}^{\infty} \mathbf{\Gamma}_j$ if $\mathbf{\Gamma}_j$ is summable (i.e., if each component of $\mathbf{\Gamma}_j$ is summable).

Gordin's (1969) CLT for Zero-mean Ergodic Stationary Processes

Gordin's condition on ergodic stationary processes

(a) $E(\mathbf{y}_t \mathbf{y}_t')$ exists and is finite.

(b) $E(\mathbf{y}_t | \mathbf{y}_{t-j}, \mathbf{y}_{t-j-1}, \cdots) \stackrel{m.s.}{\rightarrow} 0$ as $j \rightarrow \infty$.

(c) $\sum_{j=0}^{\infty} [E(\mathbf{r}_{tj}' \mathbf{r}_{tj})]^{1/2}$ is finite, where
$\mathbf{r}_{tj} \equiv E(\mathbf{y}_t | \mathbf{y}_{t-j}, \mathbf{y}_{t-j-1}, \cdots) - E(\mathbf{y}_t | \mathbf{y}_{t-j-1}, \mathbf{y}_{t-j-2}, \cdots)$.

Then,

$$\sqrt{n}\bar{\mathbf{y}} \stackrel{d}{\rightarrow} N(\mathbf{0}, \sum_{j=-\infty}^{\infty} \boldsymbol{\Gamma}_j).$$

4. Estimation of Autocovariances

- Sample autocovariance function

$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-h} (x_{t+h} - \bar{x})(x_t - \bar{x}),$$

  with $\hat{\gamma}(-h) = \hat{\gamma}(h)$ for $h = 0, 1, \cdots, n-1$.

- Sample autocorrelation function

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

  which is approximately normal with mean zero and standard deviation $\sigma_{\hat{\rho}_x(h)} = \frac{1}{\sqrt{n}}$, if it is white noise.

Estimation of Cross-correlation

- Sample cross-covariance function

$$\hat{\gamma}_{xy}(h) = n^{-1} \sum_{t=1}^{n-h} (x_{t+h} - \bar{x})(y_t - \bar{y}),$$

- Sample cross-correlation function

$$\hat{\rho}_{xy}(h) = \frac{\hat{\gamma}_{xy}(h)}{\sqrt{\hat{\gamma}_x(0)\hat{\gamma}_y(0)}}$$

which is approximately normal with mean zero and standard deviation $\sigma_{\hat{\rho}_{xy}(h)} = \frac{1}{\sqrt{n}}$, if at least one of the processes is independent white noise.

5. Consistent Estimation of Long-run Covariances

Assumption (Gordin's condition restricting the degree of serial correlation): $\{\mathbf{g}_i\}$ satisfies Gordin's condition. Its long-run covariance matrix is nonsingular.

Its long-run covariance matrix is

$$\mathbf{S} = \sum_{j=\infty}^{\infty} \mathbf{\Gamma}_j = \mathbf{\Gamma}_0 + \sum_{j=1}^{\infty} (\mathbf{\Gamma}_j + \mathbf{\Gamma}_j'),$$

where $\mathbf{\Gamma}_j$ is the j-th order autocovariance matrix

$\mathbf{\Gamma}_j = E(\mathbf{g}_t \mathbf{g}_{t-j}') \quad (j = 0, \pm 1, \cdots).$

The natural estimator of individual autocovariance is

$$\hat{\mathbf{\Gamma}}_{\mathbf{j}} = \frac{1}{n} \sum_{t=j+1}^{n} \hat{\mathbf{g}}_t \hat{\mathbf{g}}_{t-j}' \quad (j = 0, 1, \cdots, n-1).$$

- If we know a priori that $\mathbf{\Gamma}_j = 0$ for $j > q$ where q is known and finite, then clearly $\mathbf{S}$ can be consistently estimated by

$$\hat{\mathbf{S}} = \hat{\mathbf{\Gamma}}_0 + \sum_{j=1}^{q}(\hat{\mathbf{\Gamma}}_j + \hat{\mathbf{\Gamma}}_j') = \sum_{j=-q}^{q} \hat{\mathbf{\Gamma}}_j \qquad (\textit{recall}: \ \hat{\mathbf{\Gamma}}_{-j} = \hat{\mathbf{\Gamma}}_j').$$

- Otherwise, the kernel-based estimator should be used

$$\hat{\mathbf{S}} = \sum_{j=-n+1}^{n-1} k\left(\frac{j}{q(n)}\right) \cdot \hat{\mathbf{\Gamma}}_j,$$

where k is a kernel function and $q(n)$ is a function of n.

- Newey and West (1987, ECA) suggest use the Bartlett kernel

$$k(x) = \begin{cases} 1 - |x|, & \textit{for } |x| \le 1, \\ 0, & \textit{for } |x| > 1. \end{cases}$$

- For example, for $q(n) = 3$, the kernel-based estimator includes autocovariances up to two (not three) lags:

$$\hat{\mathbf{S}} = \hat{\mathbf{\Gamma}}_0 + \frac{2}{3}(\hat{\mathbf{\Gamma}}_1 + \hat{\mathbf{\Gamma}}_1') + \frac{1}{3}(\hat{\mathbf{\Gamma}}_2 + \hat{\mathbf{\Gamma}}_2')$$

- Andrews (1991, ECA) suggests use the quadratic spectral (QS) kernel

$$k(x) = \frac{25}{12\pi^2 x^2}\Big(\frac{\sin(6\pi x/5)}{6\pi x/5} - \cos 6\pi x/5)\Big)$$

6. Nonstationarity

- Seasonality
- Level shift
- Intervention
- Variance change
- Trend stationary/Difference Stationary
- Unit Root

7. Classical Regression with Time Series Data

- Regression, estimation and testing

$$
\begin{aligned}
y_i &= \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i \qquad (i = 1, 2, \cdots, n) \\
\mathbf{b} - \boldsymbol{\beta} &= \Big(\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i'\Big)^{-1}\Big(\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i\varepsilon_i'\Big)
\end{aligned}
$$

- Information Criteria and Model selection

$$
\log(SSR_j/n) + (j+1)C(n)/n,
$$

For the AIC, $C(n) = 2$, BIC, $C(n) = \log(n)$.

In this formula, "$j + 1$" enters as the number of parameters.

8. Time Regressions

We consider the regression with time as a regressor,

$$
\begin{aligned}
y_t &= \alpha + \delta \cdot t + \varepsilon_t, \\
y_t &= \mathbf{x}_t' \boldsymbol{\beta} + \varepsilon_t, \\
\mathbf{x}_t &= (1, t)', \quad \boldsymbol{\beta} = (\alpha, \delta)'. \\
\mathbf{b} &\equiv (\hat{\alpha}, \hat{\delta})' = \left(\frac{1}{n} \sum_{t=1}^{n} \mathbf{x}_t \mathbf{x}_t' \right)^{-1} \left(\frac{1}{n} \sum_{t=1}^{n} \mathbf{x}_t y_t' \right). \\
\mathbf{b} - \boldsymbol{\beta} &\equiv (\hat{\alpha} - \alpha, \hat{\delta} - \delta)' = \left(\frac{1}{n} \sum_{t=1}^{n} \mathbf{x}_t \mathbf{x}_t' \right)^{-1} \left(\frac{1}{n} \sum_{t=1}^{n} \mathbf{x}_t \varepsilon_t' \right).
\end{aligned}
$$

$$
\begin{aligned}
\sum_{t=1}^{n} \mathbf{x}_t \mathbf{x}_t' &= \begin{pmatrix} n & n(n+1)/2 \\ n(n+1)/2 & n(n+1)(2n+1)/6 \end{pmatrix} \\
\Upsilon_n &= \begin{pmatrix} \sqrt{n} & 0 \\ 0 & n^{3/2} \end{pmatrix}. \\
\Upsilon_n(\mathbf{b} - \boldsymbol{\beta}) &= \begin{pmatrix} \sqrt{n}(\hat{\alpha} - \alpha) \\ n^{3/2}(\hat{\delta} - \delta) \end{pmatrix} = \Upsilon_n \Big(\sum_{t=1}^{n} \mathbf{x}_t \mathbf{x}_t' \Big)^{-1} \Big(\frac{1}{n} \sum_{t=1}^{n} \mathbf{x}_t \varepsilon_t \Big) \\
&= \Upsilon_n \Big(\sum_{t=1}^{n} \mathbf{x}_t \mathbf{x}_t' \Big)^{-1} \Upsilon_n \Upsilon_n^{-1} \Big(\sum_{t=1}^{n} \mathbf{x}_t \varepsilon_t \Big) \\
&= \Big[\Upsilon_n^{-1} \Big(\sum_{t=1}^{n} \mathbf{x}_t \mathbf{x}_t' \Big) \Upsilon_n^{-1} \Big]^{-1} \Big(\Upsilon_n^{-1} \sum_{t=1}^{n} \mathbf{x}_t \varepsilon_t \Big) \\
&\equiv \mathbf{Q}_n^{-1} \mathbf{v}_n, \\
\mathbf{Q}_n &\equiv \Upsilon_n^{-1} \Big(\sum_{t=1}^{n} \mathbf{x}_t \mathbf{x}_t' \Big) \Upsilon_n^{-1} \qquad \mathbf{v}_n \equiv \Upsilon_n^{-1} \sum_{t=1}^{n} \mathbf{x}_t \varepsilon_t .
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{Q}_n &= \begin{pmatrix} 1 & (n+1)/(2n) \\ (n+1)/(2n) & (n+1)(2n+1)/(6n^2) \end{pmatrix}, \\
\mathbf{v}_n &= \begin{pmatrix} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \varepsilon_t \\ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{t}{n} \varepsilon_t \end{pmatrix}. \\
\mathbf{Q}_n &\xrightarrow{p} \mathbf{Q} \equiv \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix}. \\
\mathbf{v}_n &\xrightarrow{d} N(0, \sigma^2 \mathbf{Q}). \\
\begin{pmatrix} \sqrt{n}(\hat{\alpha} - \alpha) \\ n^{3/2}(\hat{\delta} - \delta) \end{pmatrix} &\xrightarrow{d} N(0, \sigma^2 \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix}^{-1}).
\end{aligned}
$$