Ch. 20 Processes with Deterministic Trends

1 Traditional Asymptotic Results of OLS

Suppose a linear regression model with stochastic regressor given by

$$Y_t = \mathbf{x}_t'\boldsymbol{\beta} + \varepsilon_t, \quad t = 1, 2, ..., T, ; \ \boldsymbol{\beta} \in \mathbb{R}^k, \tag{1}$$

or in matrix form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

We are interested in the asymptotic properties such as consistency and limiting distribution [1] of the OLS estimator of $\boldsymbol{\beta}$; $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ as $T \to \infty$, under simple traditional assumptions.

[1] $\mathfrak{Proposition}$:
Given restriction on the dependence, heterogeneity, and moments of a sequence of random variables (you may think this sequence as a sample of size T) $\{Z_t\}$,

$$\bar{Z}_T - \bar{\mu}_T \xrightarrow{a.s.} 0,$$

where

$$\bar{Z}_T \equiv \frac{1}{T}\sum_{t=1}^{T} Z_t \ \ and \ \ \bar{\mu}_T \equiv E(\bar{Z}_T).$$

$\mathfrak{Proposition}$:
Given restriction on the dependence, heterogeneity, and moments of a sequence of random variables (you may think this sequence as a sample of size T) $\{Z_t\}$,

$$\frac{(\bar{Z}_T - \bar{\mu}_T)}{(\bar{\sigma}_T/\sqrt{T})} = \frac{\sqrt{T}(\bar{Z}_T - \bar{\mu}_T)}{\bar{\sigma}_T} \xrightarrow{L} N(0,1),$$

where

$$\bar{Z}_T \equiv \frac{1}{T}\sum_{t=1}^{T} Z_t, \ \ \bar{\mu}_T \equiv E(\bar{Z}_T), \ and \ \ \bar{\sigma}_T^2/T \equiv var(\bar{Z}_T) \ \ (that\ is\ \bar{\sigma}_T^2 = \frac{var(\sum_{t=1}^{T} Z_t)}{T}).$$

To see why this notation, notice that $Var(\bar{Z}_T) = \frac{Var(\sum Z_t)}{T^2} = \frac{Var(\sum Z_t)/T)}{T} = \frac{\bar{\sigma}_T^2}{T}$, that is, we assume $Var(\sum Z_t)$ is $O(T^1)$.

1.1 Independent Identically Distributed Observation

1.1.1 Consistency

To prove consistency of $\hat{\boldsymbol{\beta}}$, we use Kolmogorov's laws of large number of Ch 4. Rewrite

$$\begin{aligned}\hat{\boldsymbol{\beta}}-\boldsymbol{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\\ &= \left(\frac{\mathbf{X}'\mathbf{X}}{T}\right)^{-1}\left(\frac{\mathbf{X}'\boldsymbol{\varepsilon}}{T}\right)\\ &= \left(\frac{\sum_{t=1}^{T}\mathbf{x_t}\mathbf{x_t}'}{T}\right)^{-1}\left(\frac{\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t}{T}\right),\end{aligned}$$

we have the following result.

Theorem:

In addition to (1), suppose that

(a). $\{(\mathbf{x}_t',\varepsilon_t)'\}_{(k+1)\times 1}$ is an $i.i.d.$ sequences;

(b). $(i).E(\mathbf{x}_t\varepsilon_t)=\mathbf{0}$
$(ii).E|X_{ti}\varepsilon_t|<\infty, \quad i=1,2,...,k.$

(c). $(i)E|X_{ti}|^2<\infty, \quad i=1,2,...,k;$
$(ii)\mathbf{M}\equiv E(\mathbf{x_t}\mathbf{x_t}') \quad \textit{is positive definite};$

Then $\hat{\boldsymbol{\beta}}\xrightarrow{a.s}\boldsymbol{\beta}$.

Proof:

It is obvious that from these assumptions we have

$$\left(\frac{\mathbf{X}'\boldsymbol{\varepsilon}}{T}\right)=\left(\frac{\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t}{T}\right)\xrightarrow{a.s}E\left(\frac{\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t}{T}\right)=\mathbf{0}$$

and

$$\left(\frac{\mathbf{X}'\mathbf{X}}{T}\right)=\left(\frac{\sum_{t=1}^{T}\mathbf{x_t}\mathbf{x_t}'}{T}\right)\xrightarrow{a.s}E\left(\frac{\sum_{t=1}^{T}\mathbf{x_t}\mathbf{x_t}'}{T}\right)=\mathbf{M}. \quad (2)$$

Therefore

$$\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\xrightarrow{a.s}\mathbf{M}^{-1}\mathbf{0}=\mathbf{0},$$

or

$$\hat{\boldsymbol{\beta}}\xrightarrow{a.s}\boldsymbol{\beta}.$$

Remark:
(A). Assumption (b i) is talking about of the mean of this *i.i.d.* sequences ($X_{ti}\varepsilon_t, i = 1, 2, ..., k$), see Proposition 3.3 of White, 2001, p.32) and (bii) is about its first moment exist.
(B). Assumption (ci) guarantee its ($X_{ti}X_{tj}$) first moment exist by Cauchy-Schwarz inequality and (cii) is talking about of the mean of this *i.i.d.* ($X_{ti}X_{tj}, i = 1, 2, .., k; j = 1, 2, ..., k$) sequence.

An existence of the first moment is what is need for LLN of *i.i.d.* sequence. See p.15 of Ch.4.

1.1.2 Asymptotic Normality

To prove asymptotic normality of $\hat{\boldsymbol{\beta}}$, we use Kolmogorov's LLN and Lindeberg-Lévy's central limit theorem of Ch 4. Rewrite

$$\begin{aligned}\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) &= \left(\frac{\mathbf{X}'\mathbf{X}}{T}\right)^{-1} \sqrt{T}\left(\frac{\mathbf{X}'\boldsymbol{\varepsilon}}{T}\right) \\ &= \left(\frac{\sum_{t=1}^{T}\mathbf{x_t}\mathbf{x_t'}}{T}\right)^{-1} \sqrt{T}\left(\frac{\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t}{T}\right),\end{aligned}$$

we have the following result.

Theorem:
In addition to (1), suppose
(a). $\{(\mathbf{x}_t', \varepsilon_t)'\}$ is an i.i.d. sequences;

$$\text{(b).}\quad \begin{array}{lll} (i) & E(\mathbf{x}_t\varepsilon_t) = \mathbf{0}, & \\ (ii) & E|X_{ti}\varepsilon_t|^2 < \infty, & i = 1, 2, ..., k \\ (iii) & \mathbf{V}_T \equiv Var(T^{-1/2}\mathbf{X}'\boldsymbol{\varepsilon}) = \mathbf{V} & \textit{is positive definite} \end{array}$$

$$\text{(c).}\quad \begin{array}{ll} (i)\mathbf{M} \equiv E(\mathbf{x_t}\mathbf{x_t'}) & \textit{is positive definite}; \\ (ii)E|X_{ti}|^2 < \infty, & i = 1, 2, ..., k; \end{array}$$

Then $\mathbf{D}^{-1/2}\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{L} N(\mathbf{0}, \mathbf{I})$, where $\mathbf{D} \equiv \mathbf{M}^{-1}\mathbf{V}\mathbf{M}^{-1}$.

Proof:
It is obvious that from these assumptions we have

$$\sqrt{T}\left(\frac{\mathbf{X}'\boldsymbol{\varepsilon}}{T}\right) = T^{-1/2}\mathbf{X}'\boldsymbol{\varepsilon} \xrightarrow{L} N(\mathbf{0}, Var(T^{-1/2}\mathbf{X}'\boldsymbol{\varepsilon}) \equiv N(\mathbf{0}, \mathbf{V})$$

and

$$\left(\frac{\mathbf{X}'\mathbf{X}}{T}\right) = \left(\frac{\sum_{t=1}^{T}\mathbf{x_t}\mathbf{x_t}'}{T}\right) \xrightarrow{a.s} E\left(\frac{\sum_{t=1}^{T}\mathbf{x_t}\mathbf{x_t}'}{T}\right) = \mathbf{M}.$$

Therefore

$$\begin{aligned}\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) &\xrightarrow{L} \mathbf{M}^{-1} \cdot N(\mathbf{0}, \mathbf{V}) \\ &\equiv N(\mathbf{0}, \mathbf{M}^{-1}\mathbf{V}\mathbf{M}^{-1}),\end{aligned}$$

or

$$(\mathbf{M}^{-1}\mathbf{V}\mathbf{M}^{-1})^{-1/2}\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{L} N(\mathbf{0}, \mathbf{I}).$$

Remark:
(A). Assumption (bi) is talking about of the mean of this $i.i.d.$ sequences $(X_{ti}\varepsilon_t, i = 1, 2, ..., k)$, (bii.b) is about its second moment exist which is needed for the application of Lindeberg-Lévy's central limit theorem (see p.22 of Ch. 4) and (biii) is to standardize the random vector $T^{-1/2}(\mathbf{X}'\boldsymbol{\varepsilon})$ so that the asymptotic distribution is unit multivariate normal.
(B). Assumption (ci) is talking about of the mean of this $i.i.d.$ $(X_{ti}X_{tj}, i = 1, 2, .., k; j = 1, 2, ..., k)$ sequence and (cii) guarantee its first moment exist by Cauchy-Schwarz inequality. An existence of the first moment is what is need for LLN of $i.i.d.$ sequence. See p.15 of Ch.4.

1.2 Independent Heterogeneously Distributed Observation

1.2.1 Consistency

To prove consistency of $\hat{\boldsymbol{\beta}}$, we use revised Markov laws of large number of Ch 4. Rewrite

$$\begin{aligned}
\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} &= (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\boldsymbol{\varepsilon} \\
&= \left(\frac{\mathbf{X}'\mathbf{X}}{T}\right)^{-1}\left(\frac{\mathbf{X}'\boldsymbol{\varepsilon}}{T}\right) \\
&= \left(\frac{\sum_{t=1}^{T}\mathbf{x_t}\mathbf{x_t'}}{T}\right)^{-1}\left(\frac{\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t}{T}\right),
\end{aligned}$$

we have the following result.

Theorem:

In addition to (1), suppose
(a). $\{(\mathbf{x}_t', \varepsilon_t)'\}$ is an independent sequences;
(b).
(i) $E(\mathbf{x}_t\varepsilon_t) = \mathbf{0}$;
(ii) $E|X_{ti}\varepsilon_t|^{1+\delta} < \Delta < \infty,\ for\ some\ \delta > 0,\ \ i = 1, 2, ..., k$;
(c).
(i) $\mathbf{M_T} \equiv E(\mathbf{X}'\mathbf{X}/T)$ is positive definite;
(ii) $E|X_{ti}^2|^{1+\delta} < \Delta < \infty,\ for\ some\ \delta > 0,\ \ i = 1, 2, ..., k$;

Then $\hat{\boldsymbol{\beta}} \xrightarrow{a.s} \boldsymbol{\beta}$.

Proof:

It is obvious that from these assumptions we have

$$\left(\frac{\mathbf{X}'\boldsymbol{\varepsilon}}{T}\right) = \left(\frac{\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t}{T}\right) \xrightarrow{a.s} E\left(\frac{\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t}{T}\right) = \mathbf{0}$$

and

$$\left(\frac{\mathbf{X}'\mathbf{X}}{T}\right) = \left(\frac{\sum_{t=1}^{T}\mathbf{x_t}\mathbf{x_t'}}{T}\right) \xrightarrow{a.s} E\left(\frac{\sum_{t=1}^{T}\mathbf{x_t}\mathbf{x_t'}}{T}\right) = \mathbf{M_T}.$$

Therefore

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \xrightarrow{a.s} \mathbf{M}_T^{-1}\mathbf{0} = \mathbf{0},$$

or

$$\hat{\boldsymbol{\beta}} \xrightarrow{a.s} \boldsymbol{\beta}.$$

Remark:

1. Assumption (ii.a) is talking about of the mean of this independent sequences $(X_{ti}\varepsilon_t, i = 1, 2, ..., k)$ and (ii.b) is about its $(1+\delta)$ moment exist.
2. Assumption (iii.a) is talking about of the limits of almost sure convergence of $\frac{\mathbf{X}'\mathbf{X}}{T}$ and (iii.b) guarantee its $(1+\delta)$ moment exist of $(X_{ti}X_{tj}, i = 1, 2, .., k; j = 1, 2, ..., k)$ by Cauchy-Schwarz inequality.

An existence of the $(1+\delta)$ moment is what is need for LLN of independent sequence. See p.15 of Ch.4.

1.2.2 Asymptotic Normality

To prove asymptotic normality of $\hat{\boldsymbol{\beta}}$, we use revised Markov's LLN and Liapounov and Lindeberg-Feller's central limit theorem of Ch 4. Rewrite

$$\begin{aligned}
\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) &= \left(\frac{\mathbf{X}'\mathbf{X}}{T}\right)^{-1} \sqrt{T}\left(\frac{\mathbf{X}'\boldsymbol{\varepsilon}}{T}\right) \\
&= \left(\frac{\sum_{t=1}^{T} \mathbf{x_t}\mathbf{x_t'}}{T}\right)^{-1} \sqrt{T}\left(\frac{\sum_{t=1}^{T} \mathbf{x}_t\varepsilon_t}{T}\right),
\end{aligned}$$

we have the following result.

Theorem:

In addition to (1), suppose

(a). $\{(\mathbf{x}_t', \varepsilon_t)'\}$ is an independent sequences;

(b).

(i) $E(\mathbf{x}_t\varepsilon_t) = \mathbf{0}$;

(ii) $E|X_{ti}\varepsilon_t|^{2+\delta} < \Delta < \infty,\ for\ some\ \delta > 0,\ \ i = 1, 2, ..., k$;

(iii) $\mathbf{V}_T \equiv Var(T^{-1/2}\mathbf{X}'\boldsymbol{\varepsilon})$ is positive definite;

(c).

(i) $\mathbf{M} \equiv E(\mathbf{X}'\mathbf{X}/T)$ is positive definite;

(ii) $E|X_{ti}^2|^{1+\delta} < \Delta < \infty,\ for\ some\ \delta > 0,\ \ i = 1, 2, ..., k$;

Then $\mathbf{D}_T^{-1/2}\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{L} N(\mathbf{0}, \mathbf{I})$, where $\mathbf{D}_T \equiv \mathbf{M}_T^{-1}\mathbf{V}_T\mathbf{M}_T^{-1}$.

Proof:

It is obvious that from these assumptions we have

$$\sqrt{T}\left(\frac{\mathbf{X}'\boldsymbol{\varepsilon}}{T}\right) = T^{-1/2}\mathbf{X}'\boldsymbol{\varepsilon} \xrightarrow{L} N(\mathbf{0}, Var(T^{-1/2}\mathbf{X}'\boldsymbol{\varepsilon}) \equiv N(\mathbf{0}, \mathbf{V}_T)$$

and

$$\left(\frac{\mathbf{X}'\mathbf{X}}{T}\right) = \left(\frac{\sum_{t=1}^T \mathbf{x_t}\mathbf{x_t}'}{T}\right) \xrightarrow{a.s} E\left(\frac{\sum_{t=1}^T \mathbf{x_t}\mathbf{x_t}'}{T}\right) = \mathbf{M}_T.$$

Therefore

$$\begin{aligned}\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) &\xrightarrow{L} \mathbf{M}_T^{-1}N(\mathbf{0}, \mathbf{V}_T)\\ &\equiv N(\mathbf{0}, \mathbf{M}_T^{-1}\mathbf{V}_T\mathbf{M}_T^{-1}),\end{aligned}$$

or

$$(\mathbf{M}_T^{-1}\mathbf{V}_T\mathbf{M}_T^{-1})^{-1/2}\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{L} N(\mathbf{0}, \mathbf{I}).$$

Remark:

(A). Assumption (bi) is talking about of the mean of this independent sequences $(X_{ti}\varepsilon_t, i = 1, 2, ..., k)$, (bii) is about its $(2+\delta)$ moment exist which is needed for the application of Liapounov's central limit theorem (see p.23 of Ch. 4) and (biii) is to standardize the random vector $T^{-1/2}(\mathbf{X}'\boldsymbol{\varepsilon})$ so that the asymptotic distribution is unit multivariate normal.

(B). Assumption (ci) is talking about of the limits of almost sure convergence of $\frac{\mathbf{X}'\mathbf{X}}{T}$ and (iii.b) guarantee its $(1+\delta)$ moment exist of $(X_{ti}X_{tj}, i = 1, 2, .., k; j = 1, 2, ..., k)$ by Cauchy-Schwarz inequality. An existence of the $(1+\delta)$ moment is what is need for LLN of independent sequence. See p.15 of Ch.4.

From results above, the asymptotic normality of OLS estimator depend crucial on the existence of at least second moments of the regressors X_{ti} and from that we have LLN such that $\frac{\mathbf{X}'\mathbf{X}}{T} \xrightarrow{a.s} E\left(\frac{\sum_{t=1}^T \mathbf{x_t}\mathbf{x_t}'}{T}\right) = \mathbf{M}_T = O(1)$. As we have seen from last chapter that a $I(1)$ variables does not have a finite second moments, therefore when the regressor is a unit root process, then traditional asymptotic results for OLS estimator would not apply. However, there is a case that the regressor is not stochastic, but it violate the condition that $\frac{\mathbf{X}'\mathbf{X}}{T} \xrightarrow{a.s} E\left(\frac{\sum_{t=1}^T \mathbf{x_t}\mathbf{x_t}'}{T}\right) = \mathbf{M}_T = O(1)$, as we will see in the following that the asymptotic normality still valid though the rate convergence to the normality changes.

2 Processes with Deterministic Time Trends

The coefficients of regression models involving unit roots or deterministic time trends are typically estimated by OLS. However, the asymptotic distributions of the coefficient estimates cannot be calculated in the same way as are those for regression models involving stationary variables. Among other difficulties, the estimates of different parameters will in general have different asymptotic rate of convergence.

2.1 Asymptotic Distribution of OLS Estimators of the Simple Time Trend Model

Consider the OLS estimation of the parameters of a simple time trend,

$$Y_t = \alpha + \delta t + \varepsilon_t, \tag{3}$$

for ε_t a white noise process. If $\varepsilon_t \sim N(0, \sigma^2)$, the model (3) satisfies the classical assumption and the standard OLS t or F statistics would have exact small-sample t or F distributions. On the other hand, if ε_t is non-Gaussian, then a slightly different technique for finding the asymptotic distribution of the OLS estimates of α and δ would to be used from that employed in last section.

Write (3) in the form of the standard regression model,

$$Y_t = \mathbf{x}_t'\boldsymbol{\beta} + \varepsilon_t,$$

where

$$\begin{aligned} \mathbf{x}_t' &\equiv \begin{bmatrix} 1 & t \end{bmatrix} \\ \boldsymbol{\beta} &\equiv \begin{bmatrix} \alpha \\ \delta \end{bmatrix}. \end{aligned}$$

Let $\hat{\boldsymbol{\beta}}_T$ denote the OLS estimate of $\boldsymbol{\beta}$ based on a sample of size T, the deviation of $\hat{\boldsymbol{\beta}}_T$ from the true value can be expressed as

$$(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}) = \left[\sum_{t=1}^{T} \mathbf{x}_t\mathbf{x}_t'\right]^{-1} \left[\sum_{t=1}^{T} \mathbf{x}_t\varepsilon_t\right]. \tag{4}$$

To find the asymptotic distribution of $(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta})$, the approach in last section was to multiply (4) by $\sqrt{T}$, resulting in

$$\sqrt{T}(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}) = \left[(1/T)\sum_{t=1}^{T} \mathbf{x}_t\mathbf{x}_t'\right]^{-1} \left[(1/\sqrt{T})\sum_{t=1}^{T} \mathbf{x}_t\varepsilon_t\right]. \tag{5}$$

The usual assumption was that $(1/T)\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'$ converge in probability to nonsingular matrix $\mathbf{M}$ while $(1/\sqrt{T})\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t$ converges in distribution to a $N(0,\mathbf{V})$ random variables, implying that $\sqrt{T}(\hat{\boldsymbol{\beta}}_T-\boldsymbol{\beta})\xrightarrow{L}N(\mathbf{0},(\mathbf{M^{-1}VM^{-1}}))$.

For $\mathbf{x}_t$ given in (5), we note that

$$\frac{1}{T^{v+1}}\sum_{t=1}^{T}t^v\rightarrow\frac{1}{v+1}, \tag{6}$$

implying that

$$\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'=\begin{bmatrix}\sum 1 & \sum t\\ \sum t & \sum t^2\end{bmatrix}=\begin{bmatrix}T & T(T+1)/2\\ T(T+1)/2 & T(T+1)(2T+1)/6\end{bmatrix}\equiv\begin{bmatrix}O(T^1) & O(T^2)\\ O(T^2) & O(T^3)\end{bmatrix}. \tag{7}$$

In contrast to the usual results as (2), the matrix $(1/T)\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'$ in (5) diverges. To obtain converge and nondegenerates limiting distribution, we can think of premultiplying and postmultiplying $\left[\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right]$ by the matrix

$$\boldsymbol{\Upsilon}_T^{-1}=\begin{bmatrix}T^{1/2} & 0\\ 0 & T^{3/2}\end{bmatrix}^{-1},$$

and obtains

$$\begin{aligned}\left\{\boldsymbol{\Upsilon}_T^{-1}\left[\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right]\boldsymbol{\Upsilon}_T^{-1}\right\} &= \left\{\begin{bmatrix}[T^{-1/2} & 0\\ 0 & T^{-3/2}\end{bmatrix}\begin{bmatrix}\sum 1 & \sum t\\ \sum t & \sum t^2\end{bmatrix}\begin{bmatrix}[T^{-1/2} & 0\\ 0 & T^{-3/2}\end{bmatrix}\right\}\\ &= \begin{bmatrix}T^{-1}\sum 1 & T^{-2}\sum t\\ T^{-2}\sum t & T^{-3}\sum t^2\end{bmatrix}\rightarrow\mathbf{Q},\end{aligned}$$

where

$$\mathbf{Q}=\begin{bmatrix}1 & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{3}\end{bmatrix} \tag{8}$$

according to (6).

Turning next to the second term in (4) and premultiplying it by $\boldsymbol{\Upsilon}_T^{-1}$,

$$\boldsymbol{\Upsilon}_T^{-1}\left[\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t\right]=\begin{bmatrix}T^{-1/2} & 0\\ 0 & T^{-3/2}\end{bmatrix}\begin{bmatrix}\sum\varepsilon_t\\ \sum t\varepsilon_t\end{bmatrix}=\begin{bmatrix}(1/\sqrt{T})\sum\varepsilon_t\\ (1/\sqrt{T})\sum(t/T)\varepsilon_t\end{bmatrix}. \tag{9}$$

We now prove the asymptotic normality of (9) under standard assumption about ε_t. Suppose that ε_t is *i.i.d.* with mean zero, variance σ^2, and finite fourth moment. Then the first element of the vector in (9) satisfies

$$(1/\sqrt{T})\sum\varepsilon_t\xrightarrow{L}N(0,\sigma^2)$$

by the Linderberg-Lévy CLT.

For the second element of the vector in (9), observe that $\{(t/T)\varepsilon_t\}$ is a martingale difference sequence that satisfies the definition on p.13 of Ch. 4. Specifically, its variance is

$$\sigma_t^2 = E[(t/T)\varepsilon_t]^2 = \sigma^2 \cdot (t^2/T^2),$$

where

$$(1/T)\sum_{t=1}^{T}\sigma_t^2 = \sigma^2(1/T^3)\sum_{t=1}^{T}t^2 \to \sigma^2/3.$$

Furthermore, to apply CLT of a martingale difference sequence, we need to show that $(1/T)\sum_{t=1}^{T}[(t/T)\varepsilon_t]^2 \xrightarrow{p} \sigma^2/3$ as the condition (iii) on page 26 of Ch. 4. To prove this, notices that

$$\begin{aligned}
E\left((1/T)\sum_{t=1}^{T}[(t/T)\varepsilon_t]^2 - (1/T)\sum_{t=1}^{T}\sigma_t^2\right)^2 &= E\left((1/T)\sum_{t=1}^{T}[(t/T)\varepsilon_t]^2 - (1/T)\sum_{t=1}^{T}(t/T)^2\sigma^2\right)^2 \\
&= E\left((1/T)\sum_{t=1}^{T}(t/T)^2(\varepsilon_t^2-\sigma^2)\right)^2 \\
&= (1/T)^2\sum_{t=1}^{T}(t/T)^4 E(\varepsilon_t^2-\sigma^2)^2 \\
&= E(\varepsilon_t^2-\sigma^2)^2\left(1/T^6\sum_{t=1}^{T}t^4\right) \to 0,
\end{aligned}$$

according to (6) and fourth moment of ε_t exist by assumption.

This imply that

$$(1/T)\sum_{t=1}^{T}[(t/T)\varepsilon_t]^2 - (1/T)\sum_{t=1}^{T}\sigma_t^2 \xrightarrow{m.s} 0,$$

which also imply that

$$(1/T)\sum_{t=1}^{T}[(t/T)\varepsilon_t]^2 \xrightarrow{p} \sigma^2/3.$$

Hence, from Theorem ? (p.26 of Ch. 4), $(1/\sqrt{T})\sum_{t=1}^{T}(t/T)\varepsilon_t$ satisfies the CLT:

$$(1/\sqrt{T})\sum_{t=1}^{T}(t/T)\varepsilon_t \xrightarrow{L} N(0,\sigma^2/3).$$

Finally, consider the joint distribution of the two element in the (2×1) vector described by (9). Any linear combination of these elements takes the form

$$(1/\sqrt{T})\sum_{t=1}^{T}[\lambda_1 + \lambda_2(t/T))]\varepsilon_t.$$

Then $[\lambda_1 + \lambda_2(t/T))]\varepsilon_t$ is also a martingale difference sequence with positive variance given by $\sigma^2[\lambda_1^2 + 2\lambda_1\lambda_2(t/T)) + \lambda_2^2(t/T)^2]$ satisfying

$$\begin{aligned}(1/T)\sum_{t=1}^{T}\sigma^2\left[\lambda_1^2 + 2\lambda_1\lambda_2(t/T)) + \lambda_2^2(t/T)^2\right] &\longrightarrow \sigma^2[\lambda_1^2 + 2\lambda_1\lambda_2\left(\frac{1}{2}\right) + \lambda_2^2\left(\frac{1}{3}\right)] \\ &= \sigma^2\left[\begin{array}{cc}\lambda_1 & \lambda_2\end{array}\right]\left[\begin{array}{cc}1 & 1/2 \\ 1/2 & 1/3\end{array}\right]\left[\begin{array}{c}\lambda_1 \\ \lambda_2\end{array}\right] \\ &= \sigma^2\boldsymbol{\lambda}'\mathbf{Q}\boldsymbol{\lambda}\end{aligned}$$

for $\boldsymbol{\lambda} \equiv (\lambda_1, \lambda_2)'$ and $\mathbf{Q}$ the matrix in (8). Furthermore, we can show that

$$(1/T)\sum_{t=1}^{T}[\lambda_1 + \lambda_2(t/T))]^2\varepsilon_t^2 \xrightarrow{p} \sigma^2\boldsymbol{\lambda}'\mathbf{Q}\boldsymbol{\lambda}.$$

That is this martingale difference sequence $[\lambda_1 + \lambda_2(t/T))]\varepsilon_t$ could apply CLT. Thus, any linear combination of the two elements in the vector in (9) is asymptotically Gaussian, implying that bivariate Gaussian distribution:

$$\left[\begin{array}{c}(1/\sqrt{T})\sum\varepsilon_t \\ (1/\sqrt{T})\sum(t/T)\varepsilon_t\end{array}\right] \xrightarrow{L} N(\mathbf{0}, \sigma^2\mathbf{Q})$$

form Cramer-Wold device and the fact that $E\{[(1/\sqrt{T})\sum\varepsilon_t][(1/\sqrt{T})\sum(t/T)\varepsilon_t]\} = \frac{1}{T^2}\sigma^2\sum t \to \frac{1}{2}\sigma^2$.

Collecting results we have

$$\begin{aligned}\left[\begin{array}{c}T^{1/2}(\hat{\alpha}_T - \alpha) \\ T^{3/2}(\hat{\delta}_T - \delta)\end{array}\right] &= \boldsymbol{\Upsilon}_T\left[\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right]^{-1}\left[\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t\right] \\ &= \boldsymbol{\Upsilon}_T\left[\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right]^{-1}\boldsymbol{\Upsilon}_T\boldsymbol{\Upsilon}_T^{-1}\left[\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t\right] \\ &= \left\{\boldsymbol{\Upsilon}_T^{-1}\left[\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right]\boldsymbol{\Upsilon}_T^{-1}\right\}^{-1}\left\{\boldsymbol{\Upsilon}_T^{-1}\left[\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t\right]\right\} \\ &\xrightarrow{L} \mathbf{Q}^{-1}\cdot N(\mathbf{0}, \sigma^2\mathbf{Q}) \\ &\equiv N(\mathbf{0}, \sigma^2\mathbf{Q}^{-1}).\end{aligned}$$

It turns out that the *OLS* estimators $\hat{\alpha}_T$ and $\hat{\delta}_T$ have different asymptotic rates of convergence. Note that the time trend estimator $\hat{\delta}_T$ is *superconsistent*–not only $\hat{\delta}_T \xrightarrow{p} \delta$, but even when multiplied by T, we still have

$$T(\hat{\delta}_T - \delta) \xrightarrow{p} 0.$$

2.2 Asymptotic Distribution of *OLS* Estimators for an Autoregressive Process Around a Deterministic Time Trend

The same principles can be used to study a general autoregressive process around a deterministic trend:

$$Y_t = \alpha + \delta t + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \varepsilon_t. \tag{10}$$

It is assumed that ε_t is *i.i.d.* with mean zero, variance σ^2, and finite fourth moment, and that roots of

$$1 - \phi_1 z - \phi_2 z^2 - ... - \phi_p z^p = 0$$

lie outside the unit circle. Consider a sample of size $T + p$ observations on Y, and let $\hat{\alpha}, \hat{\delta}, \hat{\phi}_{1,T}, ..., \hat{\phi}_{p,T}$ denote coefficients based on OLS estimation of (10) for $t = 1, 2, ..., T$.

Remark: The regressor $Y_{t-i}, i = 1, ..., p$ in (10) is a trend-stationary process (it is nonstationary itself !). To remove the nonstationarity of the regressors for the validity of LLN ($\mathbf{X}'\mathbf{X}/T \xrightarrow{p} E(\mathbf{X}'\mathbf{X}/T)$), we may transform the regressor in terms of zero-mean covariance-stationary process by subtracting time trend from each regressor.

2.2.1 A useful Transformation of the Regressors

By adding and subtracting $\phi_j[\alpha + \delta(t - j)]$ for $j = 1, 2, ..., p$ on the right side, the regression model (10) can be equivalent be written as (for each regressor Y_{t-j}, it has a constant α and trend $\delta(t - j)$)

$$\begin{aligned} Y_t &= \alpha(1 + \phi_1 + \phi_2 + ... + \phi_p) - \delta(\phi_1 + 2\phi_2 + ... + p\phi_p) \\ &\quad +\delta(1 + \phi_1 + \phi_2 + ... + \phi_p)t \\ &\quad +\phi_1[Y_{t-1} - \alpha - \delta(t - 1)] + \phi_2[Y_{t-2} - \alpha - \delta(t - 2)] + ... \\ &\quad +\phi_p[Y_{t-p} - \alpha - \delta(t - p)] + \varepsilon_t \end{aligned}$$

or

$$Y_t = \alpha^* + \delta^* t + \phi_1^* Y_{t-1}^* + \phi_2^* Y_{t-2}^* + ... + \phi_p^* Y_{t-p}^* + \varepsilon_t. \tag{11}$$

where

$$\begin{aligned}
\alpha^* &= \alpha(1+\phi_1+\phi_2+...+\phi_p)-\delta(\phi_1+2\phi_2+...+p\phi_p)\\
\delta^* &= \delta(1+\phi_1+\phi_2+...+\phi_p)\\
\phi_j^* &= \phi_j \quad for\ j=1,2,...,p
\end{aligned}$$

and

$$Y_{t-j}^* = Y_{t-j}-\alpha-\delta(t-j) \qquad for\ j=1,2,...,p.$$

The original regression model (10) can be written

$$Y_t = \mathbf{x}_t'\boldsymbol{\beta}+\varepsilon_t, \tag{12}$$

where

$$\mathbf{x}_t = \begin{bmatrix} Y_{t-1}\\ Y_{t-2}\\ .\\ .\\ .\\ Y_{t-p}\\ 1\\ t \end{bmatrix} \qquad \boldsymbol{\beta} = \begin{bmatrix} \phi_1\\ \phi_2\\ .\\ .\\ .\\ \phi_p\\ \alpha\\ \delta \end{bmatrix}.$$

The algebraic transformation in arriving at (11) could then be described as rewriting (12) in the form

$$Y_t = \mathbf{x}_t'\mathbf{G}'[\mathbf{G}']^{-1}\boldsymbol{\beta}+\varepsilon_t = [\mathbf{x}_t^*]'\boldsymbol{\beta}^*+\varepsilon_t, \tag{13}$$

where

$$\mathbf{G}' \equiv \begin{bmatrix}
1 & 0 & . & . & . & 0 & 0 & 0\\
0 & 1 & . & . & . & 0 & 0 & 0\\
. & . & . & . & . & . & . & .\\
. & . & . & . & . & . & . & .\\
. & . & . & . & . & . & . & .\\
0 & 0 & . & . & . & 1 & 0 & 0\\
-\alpha+\delta & -\alpha+2\delta & . & . & . & -\alpha+p\delta & 1 & 0\\
-\delta & -\delta & . & . & . & -\delta & 0 & 1
\end{bmatrix},$$

$$[\mathbf{G}']^{-1} \equiv \begin{bmatrix} 1 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & 1 & . & . & . & 0 & 0 & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & 1 & 0 & 0 \\ \alpha-\delta & \alpha-2\delta & . & . & . & \alpha-p\delta & 1 & 0 \\ \delta & \delta & . & . & . & \delta & 0 & 1 \end{bmatrix},$$

(hints: From partitioned inverse rule, $\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{H} & \mathbf{I} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{H} & \mathbf{I} \end{bmatrix}$.)

$$\mathbf{x}_t^* \equiv \mathbf{G}\mathbf{x}_t = \begin{bmatrix} Y_{t-1}^* \\ Y_{t-2}^* \\ . \\ . \\ . \\ Y_{t-p}^* \\ 1 \\ t \end{bmatrix} \qquad \boldsymbol{\beta}^* \equiv [\mathbf{G}']^{-1}\boldsymbol{\beta} = \begin{bmatrix} \phi_1^* \\ \phi_2^* \\ . \\ . \\ . \\ \phi_p^* \\ \alpha^* \\ \delta^* \end{bmatrix}.$$

The OLS estimate of $\boldsymbol{\beta}^*$ based on regression of Y_t on $\mathbf{x}_t^*$ is given by

$$\begin{aligned}
\hat{\boldsymbol{\beta}}^* &= \left[\sum_{t=1}^T \mathbf{x}_t^*[\mathbf{x}_t^*]'\right]^{-1}\left[\sum_{t=1}^T \mathbf{x}_t Y_t\right] \\
&= \left[\mathbf{G}\left(\sum_{t=1}^T \mathbf{x}_t\mathbf{x}_t'\right)\mathbf{G}'\right]^{-1}\mathbf{G}\left(\sum_{t=1}^T \mathbf{x}_t Y_t\right) \\
&= [\mathbf{G}']^{-1}\left(\sum_{t=1}^T \mathbf{x}_t\mathbf{x}_t'\right)^{-1}\mathbf{G}^{-1}\mathbf{G}\left(\sum_{t=1}^T \mathbf{x}_t Y_t\right) \\
&= [\mathbf{G}']^{-1}\left(\sum_{t=1}^T \mathbf{x}_t\mathbf{x}_t'\right)^{-1}\left(\sum_{t=1}^T \mathbf{x}_t Y_t\right) \\
&= [\mathbf{G}']^{-1}\hat{\boldsymbol{\beta}},
\end{aligned}$$

where $\hat{\boldsymbol{\beta}}$ is the coefficients OLS estimation from original data of Y_t on $\mathbf{x}_t$.

The asymptotic distribution of $\hat{\boldsymbol{\beta}}$ can therefore be inferred from

$$\hat{\boldsymbol{\beta}} = \mathbf{G}'\hat{\boldsymbol{\beta}}^*. \tag{14}$$

2.2.2 The Asymptotic Distribution of *OLS* Estimates for the Transformed Regression

To derive the asymptotic distribution $\hat{\boldsymbol{\beta}}^*_T$, we note that

$$\begin{aligned}
\boldsymbol{\Upsilon}_T(\hat{\boldsymbol{\beta}}^*_T-\boldsymbol{\beta}^*) &= \boldsymbol{\Upsilon}_T\left[\sum_{t=1}^{T}\mathbf{x}^*_t[\mathbf{x}^*]'_t\right]^{-1}\left[\sum_{t=1}^{T}\mathbf{x}^*_t\varepsilon_t\right]\\
&= \boldsymbol{\Upsilon}_T\left[\sum_{t=1}^{T}\mathbf{x}^*_t[\mathbf{x}^*]'_t\right]^{-1}\boldsymbol{\Upsilon}_T\boldsymbol{\Upsilon}_T^{-1}\left[\sum_{t=1}^{T}\mathbf{x}^*_t\varepsilon_t\right]\\
&= \left\{\boldsymbol{\Upsilon}_T^{-1}\left[\sum_{t=1}^{T}\mathbf{x}^*_t[\mathbf{x}^*]'_t\right]\boldsymbol{\Upsilon}_T^{-1}\right\}^{-1}\left\{\boldsymbol{\Upsilon}_T^{-1}\left[\sum_{t=1}^{T}\mathbf{x}^*_t\varepsilon_t\right]\right\},
\end{aligned}$$

where

$$\boldsymbol{\Upsilon}_T=\begin{bmatrix}
T^{1/2} & 0 & 0 & . & . & . & 0 & 0 & 0\\
0 & T^{1/2} & 0 & . & . & . & 0 & 0 & 0\\
. & . & . & . & . & . & . & . & .\\
. & . & . & . & . & . & . & . & .\\
. & . & . & . & . & . & . & . & .\\
0 & 0 & 0 & . & . & . & T^{1/2} & 0 & 0\\
0 & 0 & 0 & . & . & . & 0 & T^{1/2} & 0\\
0 & 0 & 0 & . & . & . & 0 & 0 & T^{3/2}
\end{bmatrix}.$$

From (13),

$$\sum_{t=1}^{T}\mathbf{x}^*_t[\mathbf{x}^*]'_t=\begin{bmatrix}
\sum(Y^*_{t-1})^2 & \sum Y^*_{t-1}Y^*_{t-2} & . & . & . & \sum Y^*_{t-1}Y^*_{t-p} & \sum Y^*_{t-1} & \sum tY^*_{t-1}\\
\sum Y^*_{t-2}Y^*_{t-1} & \sum(Y^*_{t-2})^2 & . & . & . & \sum Y^*_{t-2}Y^*_{t-p} & \sum Y^*_{t-2} & \sum tY^*_{t-2}\\
. & . & . & . & . & . & . & .\\
. & . & . & . & . & . & . & .\\
. & . & . & . & . & . & . & .\\
\sum Y^*_{t-p}Y^*_{t-1} & \sum Y^*_{t-p}Y^*_{t-2} & . & . & . & \sum(Y^*_{t-p})^2 & \sum Y^*_{t-p} & \sum tY^*_{t-p}\\
\sum Y^*_{t-1} & \sum Y^*_{t-2} & . & . & . & \sum Y^*_{t-p} & \sum 1 & \sum t\\
\sum tY^*_{t-1} & \sum tY^*_{t-2} & . & . & . & \sum tY^*_{t-p} & \sum t & \sum t^2
\end{bmatrix},$$

and therefore,

$$\boldsymbol{\Upsilon}_T^{-1}\sum_{t=1}^{T}\mathbf{x}^*_t[\mathbf{x}^*]'_t\boldsymbol{\Upsilon}_T^{-1}$$

$$=\begin{bmatrix}
T^{-1}\sum(Y^*_{t-1})^2 & T^{-1}\sum Y^*_{t-1}Y^*_{t-2} & . & . & . & T^{-1}\sum Y^*_{t-1}Y^*_{t-p} & T^{-1}\sum Y^*_{t-1} & T^{-2}\sum tY^*_{t-1}\\
T^{-1}\sum Y^*_{t-2}Y^*_{t-1} & T^{-1}\sum(Y^*_{t-2})^2 & . & . & . & T^{-1}\sum Y^*_{t-2}Y^*_{t-p} & T^{-1}\sum Y^*_{t-2} & T^{-2}\sum tY^*_{t-2}\\
. & . & . & . & . & . & . & .\\
. & . & . & . & . & . & . & .\\
. & . & . & . & . & . & . & .\\
T^{-1}\sum Y^*_{t-p}Y^*_{t-1} & T^{-1}\sum Y^*_{t-p}Y^*_{t-2} & . & . & . & T^{-1}\sum(Y^*_{t-p})^2 & T^{-1}\sum Y^*_{t-p} & T^{-2}\sum tY^*_{t-p}\\
T^{-1}\sum Y^*_{t-1} & T^{-1}\sum Y^*_{t-2} & . & . & . & T^{-1}\sum Y^*_{t-p} & T^{-1}T & T^{-2}\sum t\\
T^{-2}\sum tY^*_{t-1} & T^{-2}\sum tY^*_{t-2} & . & . & . & T^{-2}\sum tY^*_{t-p} & T^{-2}\sum t & T^{-3}\sum t^2
\end{bmatrix}.(15)$$

For the first p rows and columns, the row i, column j elements of the matrix (15)

$$T^{-1}\sum_{t=1}^{T} Y^*_{t-i}Y^*_{t-j} \xrightarrow{p} \gamma^*_{|i-j|}$$

by LLN of covariance stationary process.

The first p element of row $p+1$ (and the first p element of column $p+1$)

$$T^{-1}\sum_{t=1}^{T} Y^*_{t-i} \xrightarrow{p} 0$$

also by LLN of zero-mean covariance stationary process.

The first p element of row $p+2$ (and the first p element of column $p+2$)

$$T^{-1}\sum_{t=1}^{T} (t/T)Y^*_{t-i} \xrightarrow{p} 0$$

from the theorem below. Thus,

$$\mathbf{\Upsilon}_T^{-1}\sum_{t=1}^{T} \mathbf{x}^*_t[\mathbf{x}^*]'_t\mathbf{\Upsilon}_T^{-1} \xrightarrow{p} \mathbf{Q}^* \tag{16}$$

where

$$\mathbf{Q}^* = \begin{bmatrix} \gamma^*_0 & \gamma^*_1 & \gamma^*_2 & . & . & . & \gamma^*_{p-1} & 0 & 0 \\ \gamma^*_1 & \gamma^*_0 & \gamma^*_1 & . & . & . & \gamma^*_{p-2} & 0 & 0 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ \gamma^*_{p-1} & \gamma^*_{p-2} & \gamma^*_{p-3} & . & . & . & \gamma^*_0 & 0 & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & . & . & . & 0 & \frac{1}{2} & \frac{1}{3} \end{bmatrix}.$$

Theorem:
Let Y^*_{t-i} be covariance-stationary with mean zero and absolutely summable autocovariance, then $T^{-2}\sum ty^*_{t-i} \xrightarrow{p} 0,\ i=1,2,...,p.$

Proof:
We show that $E(T^{-2}\sum tY^*_{t-i}-0)^2 \rightarrow 0$ which would imply that $T^{-2}\sum tY^*_{t-i} \xrightarrow{m.s} 0$ and also that $T^{-2}\sum tY^*_{t-i} \xrightarrow{p} 0.$

To see this, since

$$
\begin{aligned}
E(T^{-2}\sum tY^*_{t-i}-0)^2 &= (1/T^4)E[(Y_{1-i}+2Y_{2-i}+...+TY_{T-i})(Y_{1-i}+2Y_{2-i}+...+TY_{T-i})] \\
&= (1/T^4)E\{(Y_{1-i})[(Y_{1-i}+2Y_{2-i}+...+TY_{T-i})] \\
&\quad +(2Y_{2-i})[(Y_{1-i}+2Y_{2-i}+...+TY_{T-i})] \\
&\quad +(3Y_{3-i})[(Y_{1-i}+2Y_{2-i}+...+TY_{T-i})] \\
&\quad +...+(TY_{T-i})[(Y_{1-i}+2Y_{2-i}+...+TY_{T-i})]\} \\
&= (1/T^4)\{[1\cdot 1\gamma_0+1\cdot 2\gamma_1+1\cdot 3\gamma_2+1\cdot 4\gamma_3+...+1\cdot T\gamma_{T-1}] \\
&\quad +[2\cdot 1\gamma_1+2\cdot 2\gamma_0+2\cdot 3\gamma_1+2\cdot 4\gamma_2+...+2\cdot T\gamma_{T-2}] \\
&\quad +[3\cdot 1\gamma_2+3\cdot 2\gamma_1+3\cdot 3\gamma_0+3\cdot 4\gamma_1+3\cdot 5\gamma_2+...+3\cdot T\gamma_{T-3}] \\
&\quad +...+[T\cdot 1\gamma_{T-1}+T\cdot 2\gamma_{T-2}+T\cdot 3\gamma_{T-3}+...+T\cdot T\gamma_0]\} \\
&= (1/T^4)\left\{\sum_{t=1}^{T} t^2\gamma_0+2\sum_{t=1}^{T-1} t(t+1)\gamma_1+2\sum_{t=1}^{T-2} t(t+2)\gamma_2+...+2T\gamma_{T-1}\right\} \\
&= (1/T)\left\{\left[\sum_{t=1}^{T} t^2/T^3\right]\gamma_0+\left[\sum_{t=1}^{T-1} t(t+1)/T^3\right]2\gamma_1\right. \\
&\quad \left.\left[\sum_{t=1}^{T-2} t(t+2)/T^3\right]2\gamma_2+....+\left[T/T^3\right]2\gamma_{T-1}\right\} \\
&= (1/T)\left|\left[\sum_{t=1}^{T} t^2/T^3\right]\gamma_0+\left[\sum_{t=1}^{T-1} t(t+1)/T^3\right]2\gamma_1\right. \\
&\quad \left.\left[\sum_{t=1}^{T-2} t(t+2)/T^3\right]2\gamma_2+....+\left[T/T^3\right]2\gamma_{T-1}\right|,
\end{aligned}
$$

then

$$\begin{aligned}
T\cdot E(T^{-2}\sum tY^*_{t-i}-0)^2 &= \left|\left[\sum_{t=1}^{T} t^2/T^3\right]\gamma_0+\left[\sum_{t=1}^{T-1} t(t+1)/T^3\right]2\gamma_1\right. \\
&\quad \left.\left[\sum_{t=1}^{T-2} t(t+2)/T^3\right]2\gamma_2+....+\left[T/T^3\right]2\gamma_{T-1}\right|, \\
&\leq \left\{\left|\sum_{t=1}^{T} t^2/T^3\right||\gamma_0|+\left|\sum_{t=1}^{T-1} t(t+1)/T^3\right|2|\gamma_1|\right. \\
&\quad \left.\left|\sum_{t=1}^{T-2} t(t+2)/T^3\right|2|\gamma_2|+....+\left|T/T^3\right|2|\gamma_{T-1}|\right\} \\
&\leq \{|\gamma_0|+2|\gamma_1|+2|\gamma_2|+...\} \quad \left(since\ (1/T^{v+1}\sum_{t=1}^{T} t^v \rightarrow 1/(v+1)<1\right) \\
&< \infty.
\end{aligned}$$

So, $E(T^{-2}\sum tY^*_{t-i}-0)^2 \longrightarrow 0$ and therefore $T^{-2}\sum tY^*_{t-i} \xrightarrow{p} 0$ as claimed.

We now turn to second element of OLS estimator,

$$\left\{\boldsymbol{\Upsilon}_T^{-1}\left[\sum_{t=1}^{T}\mathbf{x}_t^*\varepsilon_t\right]\right\}=\begin{bmatrix} T^{-1/2}\sum Y^*_{t-1}\varepsilon_t \\ T^{-1/2}\sum Y^*_{t-2}\varepsilon_t \\ . \\ . \\ . \\ T^{-1/2}\sum Y^*_{t-p}\varepsilon_t \\ T^{-1/2}\sum \varepsilon_t \\ T^{-1/2}\sum (t/T)\varepsilon_t \end{bmatrix}=T^{-1/2}\sum_{t=1}^{T}\boldsymbol{\xi}_t,$$

where

$$\boldsymbol{\xi}_t=\begin{bmatrix} \sum Y^*_{t-1}\varepsilon_t \\ \sum Y^*_{t-2}\varepsilon_t \\ . \\ . \\ . \\ \sum Y^*_{t-p}\varepsilon_t \\ \sum \varepsilon_t \\ \sum (t/T)\varepsilon_t \end{bmatrix}.$$

But $\boldsymbol{\xi}_t$ is a martingale difference with variance

$$E(\boldsymbol{\xi}_t\boldsymbol{\xi}_t')=\sigma^2\boldsymbol{Q}_t^*,$$

where

$$\mathbf{Q}_t^* = \begin{bmatrix} \gamma_0^* & \gamma_1^* & \gamma_2^* & . & . & . & \gamma_{p-1}^* & 0 & 0 \\ \gamma_1^* & \gamma_0^* & \gamma_1^* & . & . & . & \gamma_{p-2}^* & 0 & 0 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ \gamma_{p-1}^* & \gamma_{p-2}^* & \gamma_{p-3}^* & . & . & . & \gamma_0^* & 0 & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 1 & t/T \\ 0 & 0 & 0 & . & . & . & 0 & t/T & t^2/T^2 \end{bmatrix}$$

and

$$(1/T)\sum_{t=1}^{T} \mathbf{Q}_t^* \rightarrow \mathbf{Q}^*.$$

Applying the CLT for martingale difference, it can be shown that

$$\boldsymbol{\Upsilon}_T^{-1}\sum_{t=1}^{T} \mathbf{x}_t^* \varepsilon_t \xrightarrow{L} N(\mathbf{0}, \sigma^2 \mathbf{Q}^*). \tag{17}$$

It follow from (16) and (17) that

$$\boldsymbol{\Upsilon}_T(\hat{\boldsymbol{\beta}}^*{}_T - \boldsymbol{\beta}) \xrightarrow{L} N(\mathbf{0}, [\mathbf{Q}^*]^{-1}\sigma^2\mathbf{Q}^*[\mathbf{Q}^*]^{-1}) = N(\mathbf{0}, \sigma^2[\mathbf{Q}^*]^{-1}). \tag{18}$$

2.2.3 The Asymptotic Distribution of *OLS* Estimators for the Original Regression

From (14) we have

$$\begin{bmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ . \\ . \\ . \\ \hat{\phi}_p \\ \hat{\alpha} \\ \hat{\delta} \end{bmatrix} = \begin{bmatrix} 1 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & 1 & . & . & . & 0 & 0 & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & 1 & 0 & 0 \\ -\alpha+\delta & -\alpha+2\delta & . & . & . & -\alpha+p\delta & 1 & 0 \\ -\delta & -\delta & . & . & . & -\delta & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\phi}_1^* \\ \hat{\phi}_2^* \\ . \\ . \\ . \\ \hat{\phi}_p^* \\ \hat{\alpha}^* \\ \hat{\delta}^* \end{bmatrix}.$$

The original *OLS* estimators $\hat{\phi}_i$ are identical to the estimators from the transformed regression $\hat{\phi}_i^*$, the asymptotic distribution of $\hat{\phi}_i$ is given immediately from

(18). The estimator $\hat{\alpha}$ is a linear combination of variables that converges to a Gaussian distribution at rate $\sqrt{T}$, so $\hat{\alpha}$ behave the same way. Specifically, $\hat{\alpha} = \mathbf{g}'_\alpha \hat{\boldsymbol{\beta}}^*_T$, where

$$\mathbf{g}'_\alpha \equiv \begin{bmatrix} -\alpha + \delta & -\alpha + 2\delta & . & . & . & -\alpha + p\delta & 1 & 0 \end{bmatrix},$$

so from (18),

$$\sqrt{T}(\hat{\alpha} - \alpha) \xrightarrow{L} N(\mathbf{0}, \mathbf{g}'_\alpha \sigma^2 [\mathbf{Q}^*]^{-1} \mathbf{g}_\alpha).$$

Finally, the estimator $\hat{\delta}$ is a linear combination of variables converging at different rate:

$$\hat{\delta} = \mathbf{g}'_\delta \hat{\boldsymbol{\beta}}^*_T + \hat{\delta}^*_T,$$

where

$$\mathbf{g}'_\delta = \begin{bmatrix} -\delta & -\delta & . & . & . & -\delta & 0 & 0 \end{bmatrix}.$$

Since $\mathbf{g}'_\delta \hat{\boldsymbol{\beta}}^*_T$ is $O(T^{-1/2})$ and $\hat{\delta}^*_T$ is $O(T^{-3/2})$, $\hat{\delta}$ is $O(T^{-1/2})$ (see p. 3 of Ch 4), therefore,

$$\sqrt{T}(\hat{\delta} - \delta) \xrightarrow{L} N(\mathbf{0}, \mathbf{g}'_\delta \sigma^2 [\mathbf{Q}^*]^{-1} \mathbf{g}_\delta).$$

Thus, each of the elements of $\hat{\boldsymbol{\beta}}_T$ individually is asymptotically Gaussian and is $O(T^{-1/2})$. The asymptotic distribution of the full vector $\sqrt{T}(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta})$ is multivariate Gaussian.