LECTURE 16
TIME SERIES TOPICS

Definitions

Often econometricians have to deal with data sets that come in the form of a *time series* or *stochastic process*, a collection of observations on the same variable (or vector of variables) indexed by the date of measurement of each observation. The data is usually collected at equally spaced dates (daily, weekly, monthly and etc.) and indexed by $t = 1, \ldots, T$:

$$\{Y_t : t = 1, \ldots, T\}.$$

The index t measures passage of time, and Y_t measures evolution of the process in time. It is usually assumed that the observed sample is only a segment of the process that started in the infinite past and will go on indefinitely:

$$\{\ldots, Y_{-1}, Y_0, Y_1, Y_2, \ldots, Y_T, Y_{T+1}, \ldots\}.$$

In the iid or *cross-sectional* cases, we can view the sample as a collection of n independent "copies" of the same random variable/vector. Each Y_i samples from the common probability space, and, therefore, as the sample size increases, sample average $n^{-1}\sum_{i=1}^n Y_i$ reveals the population average EY_i, which is the same for all i's. The time series case differs from the iid framework in the sense that each element in the trajectory $Y_1, Y_2, \ldots, Y_T$ is observed only once. If we are interested in learning about the underlying population model, first, we should be concerned whether different observations share the same population model. It is possible that $EY_t = \mu_t$ is different across t's. The second concern is whether the measurements made at different dates effectively sample from the population model. This might not be the case, if observations are highly correlated. In order to address those issues we introduce following definitions.

Definition 1 *The process Y_t is said to be strictly stationary if $(Y_{t_1}, \ldots, Y_{t_k}) =^d (Y_{t_1+h}, \ldots, Y_{t_k+h})$ for all k, h and $t_1, \ldots t_k$.*

Definition 2 *Suppose that $Var(Y_t) < \infty$ for all t. The process Y_t is said to be (weakly or covariance) stationary if*

$$\begin{aligned} EY_t &= \mu \text{ for all } t, \\ Cov(Y_t, Y_{t-k}) &= Cov(Y_s, Y_{s-k}) \\ &= \gamma(k) \text{ for all } s \text{ and } k. \end{aligned}$$

The function $\gamma(k)$ is called the *autocovariance function.* The stationarity assumption implies that the covariance $Cov(Y_t, Y_s) = \gamma(t-s)$, and therefore, in the scalar case,

$$\begin{aligned} \gamma(k) &= Cov(Y_t, Y_{t-k}) \\ &= Cov(Y_{t-k}, Y_t) \\ &= \gamma(-k). \end{aligned}$$

Similarly, in the vector case, we obtain that

$$\gamma(k) = \gamma(-k)'.$$

Note that for a stationary process, $Var(Y_t) = \gamma(0)$. The *autocorrelation* function is defined as $\rho(k) = \gamma(k)/\gamma(0)$.

Definition 3 *(loose). The process Y_t is said to be ergodic if $\gamma(k) \to 0$ as $k \to \infty$.*

The stationarity assumptions ensure that the observations measured at different time periods share some common underlying model. The ergodicity assumption implies that distant observations are almost uncorrelated.

LLNs for a covariance stationary process

The LLNs for *strictly stationary* and *ergodic* sequence $\{Y_t : t \geq 1\}$ says that if $EY_t = \mu$ (the mean is finite), then, as $T \to \infty$,

$$T^{-1} \sum_{t=1}^{T} Y_t \to_p \mu.$$

We will prove a weaker result. Let $\{Y_t : t \geq 1\}$ be a *covariance stationary* sequence of random variables, and assume that $\sum_{i=1}^{\infty} |\gamma(k)| < \infty$. Then, as $T \to \infty$,

$$T^{-1} \sum_{t=1}^{T} Y_t \to_p EY_t.$$

The condition $\sum_{i=1}^{\infty} |\gamma(k)| < \infty$ above says that the autocovariance function of Y_t has to be absolute summable. For this to hold, it must be true that $\gamma(k) \to 0$ as $k \to \infty$ (Y_t is ergodic). However, contrary to the first result, here we require existence of the second moment.

Proof. First, by the Markov's inequality,

$$P\left(\left|T^{-1} \sum_{t=1}^{T} Y_t - EY_t\right| > \varepsilon\right) \leq \frac{E\left|T^{-1} \sum_{t=1}^{T} Y_t - EY_t\right|^2}{\varepsilon^2}.$$

We need to show that the denominator on the right-hand side of the above expression converges to zero as $T \to \infty$.

$$\begin{aligned}
E\left|T^{-1} \sum_{t=1}^{T} Y_t - EY_t\right|^2 &= E\left|T^{-1} \sum_{t=1}^{T} (Y_t - EY_t)\right|^2 \\
&= T^{-2} \sum_{t=1}^{T} \sum_{s=1}^{T} (Y_t - EY_t)(Y_s - EY_s) \\
&= T^{-2} \sum_{t=1}^{T} \sum_{s=1}^{T} \gamma(t-s) \\
&= T^{-2} \left(T\gamma(0) + 2 \sum_{j=1}^{T-1} (T-j)\gamma(j)\right).
\end{aligned}$$

In order to verify the last equality, note that $\sum_{t=1}^{T} \sum_{s=1}^{T} \gamma(t-s)$ is given by the sum of all elements in the following $T \times T$ matrix.

$$\begin{pmatrix}
\gamma(0) & \gamma(1) & \ldots & \gamma(T-2) & \gamma(T-1) \\
\gamma(1) & \gamma(0) & \ldots & \gamma(T-3) & \gamma(T-2) \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
\gamma(T-2) & \gamma(T-3) & \ldots & \gamma(0) & \gamma(1) \\
\gamma(T-1) & \gamma(T-2) & \ldots & \gamma(1) & \gamma(0)
\end{pmatrix}$$

Thus, since $(T-j)/T < 1$ for $1 \leq j \leq T$, we have that

$$\begin{aligned} E\left|T^{-1}\sum_{t=1}^{T} Y_t - EY_t\right|^2 &\leq T^{-1}\left(\gamma(0) + 2\sum_{j=1}^{T-1}\frac{T-j}{T}\gamma(j)\right) \\ &\leq T^{-1}\left(\gamma(0) + 2\sum_{j=1}^{T-1}\frac{T-j}{T}|\gamma(j)|\right) \\ &\leq T^{-1}\left(\gamma(0) + 2\sum_{j=1}^{T-1}|\gamma(j)|\right) \\ &\leq T^{-1}\left(\gamma(0) + 2\sum_{j=1}^{\infty}|\gamma(j)|\right) \\ &\to 0, \end{aligned}$$

since by the assumption $\sum_{j=1}^{\infty}|\gamma(j)| < \infty$.

The result can be extended to sequences of random p-vectors. In this case, $\gamma(j)$ is a $p \times p$ matrix. Its element (r,s) is given by $Cov(Y_{r,t}, Y_{s,t-j})$. In the vector case, we require that, for all $1 \leq r, s \leq p$, $\sum_{j=0}^{\infty}|Cov(Y_{r,t}, Y_{s,t-j})| < \infty$. In the vector case, the long-run variance covariance matrix is given by

$$\gamma(0) + \sum_{j=1}^{\infty}\gamma(j) + \sum_{j=1}^{\infty}\gamma(j)'.$$

Examples

White noise

Suppose that

$$\begin{aligned} EU_t &= 0 \text{ for all } t, \\ EU_t^2 &= \sigma^2 \text{ for all } t, \\ EU_tU_s &= 0 \text{ for all } s,t \text{ such that } s \neq t. \end{aligned} \tag{1}$$

Such a process is called *white noise.* Sometimes condition (1) is replaced by a stronger one, which says that U_s and U_t are independent for all $s \neq t$. In this case, we say that U_t is *independent white noise.* Naturally, a white noise process is stationary and ergodic.

Moving average models

Suppose that

$$Y_t = U_t + \theta_1 U_{t-1} + \ldots + \theta_q U_{t-q},$$

where U_t is a white noise. Such a process is called *moving average of order* q and denoted by MA(q). We will show that a moving average process is covariance stationary. First, $EY_t = 0$ for all t (a non-zero mean can be obtained if we include a nonzero constant on the right-hand side of the above equation). Next,

$$
\begin{aligned}
\gamma(0) &= \sigma^2 \left(1 + \theta_1^2 + \ldots \theta_q^2\right). \\
\gamma(1) &= EY_tY_{t-1} \\
&= E\left(U_t + \theta_1 U_{t-1} + \ldots + \theta_q U_{t-q}\right)\left(U_{t-1} + \theta_1 U_{t-2} + \ldots + \theta_q U_{t-1-q}\right) \\
&= \theta_1 EU_{t-1}^2 + \theta_2\theta_1 EU_{t-2}^2 + \ldots + \theta_q\theta_{q-1} EU_{t-q}^2 \\
&= \sigma^2\left(\theta_1 + \theta_2\theta_1 + \ldots + \theta_q\theta_{q-1}\right). \\
&\quad \ldots \\
\gamma(q) &= \sigma^2\theta_q. \\
\gamma(k) &= 0 \text{ for } k > q.
\end{aligned}
$$

Hence, MA(q) is covariance stationary and ergodic.

Autoregressive models

Suppose that

$$Y_t = \beta_1 Y_{t-1} + \ldots \beta_p Y_{t-p} + U_t,$$

where U_t is a white noise process. Such process is called autoregression of order p and denoted by AR(p). Non-zero mean processes can be modelled by including an intercept in the above equation.

Consider the case of AR(1). Write

$$
\begin{aligned}
Y_t &= \beta Y_{t-1} + U_t \\
&= \beta^2 Y_{t-2} + \beta U_{t-1} + U_t \\
&= \beta^t Y_0 + \sum_{j=0}^{t-1} \beta^j U_{t-j}.
\end{aligned}
$$

Assume that $|\beta| < 1$. Then, $\lim_{j\to\infty} \beta^j = 0$, and, therefore,

$$Y_t = \sum_{j=0}^{\infty} \beta^j U_{t-j}.$$

This is called the MA(∞) representation of a AR(1) process. We have,

$$
\begin{aligned}
\gamma(0) &= \sigma^2 \sum_{j=0}^{\infty} \beta^{2j} \\
&= \frac{\sigma^2}{1-\beta^2}, \\
\gamma(1) &= \sigma^2 \sum_{j=0}^{\infty} \beta^j \beta^{j+1} \\
&= \beta\sigma^2 \sum_{j=0}^{\infty} \beta^{2j} \\
&= \frac{\beta\sigma^2}{1-\beta^2}. \\
\gamma(k) &= \sigma^2 \sum_{j=0}^{\infty} \beta^j \beta^{j+k} \\
&= \frac{\beta^k\sigma^2}{1-\beta^2}.
\end{aligned}
$$

We have that the sequence of autocovariances $\gamma(j)$ is independent of t, and, therefore, AR(1) is weakly stationary. Further, it is ergodic since $\lim_{j\to\infty}\gamma(j)=0$. The long-run variance of the AR(1) process is given by

$$\begin{aligned} & \frac{\sigma^2}{1-\beta^2}+2\sum_{j=1}^{\infty}\frac{\beta^j\sigma^2}{1-\beta^2} \\ = & \frac{\sigma^2}{1-\beta^2}\left(1+2\beta\sum_{j=0}^{\infty}\beta^j\right) \\ = & \frac{\sigma^2}{1-\beta^2}\left(1+2\frac{\beta}{1-\beta}\right) \\ = & \frac{\sigma^2}{(1-\beta)^2}. \end{aligned}$$

Wold decomposition

The MA(∞) representation plays an important role in time-series analysis due to the result called the Wold decomposition. According to it, if $\{Y_t\}$ is covariance stationary and ergodic process, then there exists the white noise sequence $\{U_t\}$ and the sequence of constants $\{\theta_j\}$ such that

$$\begin{aligned} Y_t &= \mu+\sum_{j=0}^{\infty}\theta_j U_{t-j}, \text{ where} \\ \sum_{j=0}^{\infty}\theta_j^2 &< \infty. \end{aligned}$$

Such a process is called *linear.*

CLT for linear processes

Let Y_t be a linear process:

$$Y_t=\sum_{j=0}^{\infty}\theta_j U_{t-j},$$

where $\{U_t\}$ are iid with $EU_t=0$ and $EU_t^2=\sigma^2<\infty$, the sequence $\{\theta_j\}$ is absolutely summable, and $\sum_{j=0}^{\infty}\theta_j\neq 0$. Then,

$$T^{-1/2}\sum_{t=1}^{T}Y_t\to_d N\left(0,\sum_{j=-\infty}^{\infty}\gamma(j)\right),$$

where $\gamma(j)$ is the j-th autocovariance of Y_t, and, therefore, $\sum_{j=-\infty}^{\infty}\gamma(j)$ is the long-run variance of Y_t. For example, in the scalar case, the long-run variance is given by

$$\sigma^2\left(\sum_{j=0}^{\infty}\theta_j\right)^2.$$

Time-series regression

Suppose that the econometrician observes $\{(Y_t,X_t):t=1,\ldots,T\}$, where

- $Y_t = X_t'\beta + U_t$.
- $\beta \in R^k$ is a vector of unknown regression coefficients.
- $\{X_t\}$ is strictly stationary and ergodic with finite second moments, and EX_tX_t' is positive definite.
- $\{U_tX_t\}$ is a vector process and satisfies the conditions of the CLT for linear processes.

Consider the OLS estimator of β:

$$T^{1/2}\left(\widehat{\beta}_T - \beta\right) = \left(T^{-1}\sum_{t=1}^{T} X_tX_t'\right)^{-1} T^{-1/2}\sum_{t=1}^{T} U_tX_t.$$

By the LLN we have that

$$\left(T^{-1}\sum_{t=1}^{T} X_tX_t'\right)^{-1} \to_p \left(EX_tX_t'\right)^{-1}.$$

The CLT implies that

$$T^{-1/2}\sum_{t=1}^{T} U_tX_t \to_d N\left(0, \Omega\right),$$

where Ω is the long-run variance-covariance matrix of U_tX_t.

$$\Omega = EU_t^2X_tX_t' + \sum_{j=1}^{\infty} EU_tU_{t-j}\left(X_tX_{t-j}' + X_{t-j}X_t'\right).$$

Therefore,

$$T^{1/2}\left(\widehat{\beta}_T - \beta\right) \to_d N(0, V),$$

where

$$V = \left(EX_tX_t'\right)^{-1}\Omega\left(EX_tX_t'\right)^{-1}.$$

In this case, we allow U_t to be heteroskedastic and correlated across t (serially correlated). The result is similar to the iid case, except for the expression for Ω. The asymptotic variance matrix V can be consistently estimated by

$$\widehat{V}_T = \left(T^{-1}\sum_{t=1}^{T} X_tX_t'\right)^{-1}\widehat{\Omega}_T\left(T^{-1}\sum_{t=1}^{T} X_tX_t'\right)^{-1},$$

where $\widehat{\Omega}_T$ is heteroskedasticity and autocorrelation consistent (HAC) estimator of Ω Newey-West (1987, Econometrica):

$$\widehat{\Omega}_T = T^{-1}\sum_{t=1}^{T}\widehat{U}_t^2X_tX_t' + T^{-1}\sum_{l=1}^{L}\sum_{t=l+1}^{T}\frac{l}{L+1}\widehat{U}_t\widehat{U}_{t-l}\left(X_tX_{t-l}' + X_{t-l}X_t'\right),$$

where L is called the truncation parameter. The HAC estimator is consistent if $L \to \infty$ at a suitable rate (slower than T).