Chapter 1

ARMA Models

1.1 AR(p) Processes

We have learned Wold decomposition in the previous lecture. Let e_t be strictly stationary ergodic white noise. The ARMA are the classical approach to model a univariate time series.

- MA(1)

$$y_t = \mu + e_t + \theta e_{t-1}$$

mean: $E[y_t] = \mu$
variance: $var(y_t) = \theta^2 + 1$
autocovariance: $E[e_t e_{t-1}] = \theta$

- MA(∞)

$$y_t = \mu + \sum_{j=1}^{\infty} b_j e_{t-j}$$

where $b_0 = 1$

- AR(1)

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + e_t$$

mean: $E[y_t] = \frac{\alpha_0}{1-\alpha_1}$
variance: $var(y_t) = \frac{\sigma^2}{1-\alpha_1^2}$
MA(∞) regression: $y_t = \mu + \sum_{j=1}^{\infty} \alpha_1^j e_{t-j}$, where $\mu = \frac{\alpha_0}{1-\alpha_1}$.

To facilitate the notation, we introduce the **lag operator** L. Its effect is to push any time series observation one period to the past. That is, $\mathrm{L}x_t = x_{t-1}$. An AR(1) can be written as

$$(1 - \alpha \mathrm{L}) y_t = \alpha_0 + e_t$$
$$y_t = (1 - \alpha \mathrm{L})^{-1} (\alpha_0 + e_t).$$

For stationarity, the AR coefficient $|\alpha| < 1$. If $\alpha = 1$, it becomes a unit root process, which is very different from stationary time series.

- AR(p)

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + ... + +\alpha_p y_{t-p} + e_t$$

lag operator: $(1 - \alpha(\mathrm{L})) y_t = \alpha_0 + e_t$, where $\alpha(z)$ is a polynomial. Stationarity requires that : all roots of $1 - \alpha(z) = 0$ are strictly outside of the unit circle. That is, all the p roots (on the complex plain) must have their modulus strictly greater than 1.

1.2 ARMA and ARIMA Processes

- ARMA: $(1-\alpha(\mathrm{L}))\, y_t = b(\mathrm{L})\, e_t$
- ARIMA(p,d,q): $(1-\alpha(\mathrm{L}))\,(1-\mathrm{L})^d\, y_t = b(\mathrm{L})\, e_t$

1.3 Estimation and Asymptotic Distribution

Estimate AR: take $X_t = (1, y_{t-1}, ..., y_{t-p})$, run OLS:

$$\hat{\alpha} = \left(\frac{X'X}{n}\right)^{-1} \frac{X'y}{n}$$

Theorem 1.1. *If y_t is strictly stationary, ergodic, $E\left[y_t^2\right] < \infty$, then $\hat{\alpha} \xrightarrow{p} \alpha$ and $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$*

Asymptotic normality: If e_t is MDS, with $\mathscr{F}$ including X_t, then

$$E\left[X_t e_t \mid \mathscr{F}_{t-1}\right] = X_t E\left[e_t \mid \mathscr{F}_{t-1}\right] = 0$$

then $\sqrt{n}\,(\hat{\alpha}-\alpha) \xrightarrow{p} N\left(0, Q^{-1} \sum Q^{-1}\right)$,
where $Q = E\left[X_t X_t'\right]$ and $\sum = E\left[X_t X_t' e_t^2\right]$.
Under conditional homoskedasticity $E\left[e_t^2 \mid \mathscr{F}_{t-1}\right] = \sigma^2$, then the variance is simplified to

$$\begin{aligned} \sum &= E\left[X_t X_t' e_t^2\right] = E\left[X_t X_t' E\left[e_t^2 \mid \mathscr{F}_{t-1}\right]\right] \\ &= E\left[X_t X_t'\right] \sigma^2 = Q\sigma^2 \end{aligned}$$

then $\sqrt{n}\,(\hat{\alpha}-\alpha) \xrightarrow{p} N\left(0, Q^{-1}\sigma^2\right)$
Without MDS, $z_t = X_t e_t$ can be serially correlated, we need to estimate the long-run variance $\Omega = \sum_{\ell=-\infty}^{\infty} E\left[X_t X_{t-\ell}' e_t e_{t-\ell}\right]$

1.4 Model Selection

$$\mathrm{AIC} = \log \hat{\sigma}^2 + 2\frac{p}{n}$$

$$\mathrm{BIC} = \log \hat{\sigma}^2 + \frac{p}{n} \log n$$

1.5 Regression with Time Series Data

Observe $(y_t, X_t)_{t=1}^{T}$, want to run regression

$$y_t = X_t' \beta + e_t$$

where X_t can include lagged dependent variables.
AR(p)
By the definition of projection, $E\left[X_t e_t\right] = 0$
The OLS estimator is $\hat{\beta} = (X'X)^{-1} X'y$
The uncorrelation is necessary for asymptotic normality.
If we impose MDS, $E\left[e_t \mid \mathscr{F}_{t-1}\right] = 0$, where $\mathscr{F}_{t-1}$ is adapted to (X_t, e_{t-1})

then we have MDS CLT, because

$$E\left[X_t e_t \mid \mathscr{F}_{t-1}\right] = X_t E\left[e_t \mid \mathscr{F}_{t-1}\right] = X_t \cdot 0 = 0$$

is also MDS.

Under MDS

$$\sqrt{n}\left(\hat{\beta}-\beta\right) \xrightarrow{p} N\left(0, Q_X^{-1} \sum Q_X^{-1}\right)$$

where $\Omega = E\left[X_t X_t' e_t^2\right]$

Under $E\left[X_t e_t\right]=0$, we need conditions about the α-mixing coefficient, then

$$\sqrt{n}\left(\hat{\beta}-\beta\right) \xrightarrow{p} N\left(0, Q_X^{-1} \sum Q_X^{-1}\right)$$

where Ω is the long-run variance of $\{X_t e_t\}$.

1.6 Regression with Deterministic Trend

$y_t = T_t + u_t$, where T_t is a deterministic trend and u_t is a random error term.

Example 1.1. $T_t = \beta_0 + \beta_1 t$ (linear trend) or $T_t = \beta_0 + \beta_1 t + +\beta_2 t^2$ (quadratic trend)

Fact:

$$\frac{1}{n^{1+r}} \sum_{t=1}^{n} t^r = \frac{1}{n} \sum_{t=1}^{n} \left(\frac{t}{n}\right)^r \to \int_0^1 x^r dx = \frac{1}{1+r} x^{r+1} \mid_0^1 = \frac{1}{1+r}$$

Thus, $\frac{1}{n^2}\sum_{t=1}^n t = \frac{1}{2}$, $\frac{1}{n^3}\sum_{t=1}^n t^2 = \frac{1}{3}$

OLS estimator

$$\hat{\beta}-\beta = \left(X'X\right)^{-1} X'u = \begin{pmatrix} n & \sum_{t=1}^n t \\ \sum_{t=1}^n t & \sum_{t=1}^n t^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum u_t \\ \sum t u_t \end{pmatrix}$$

Let $D_n = \begin{pmatrix} n^{\frac{1}{2}} & 0 \\ 0 & n^{\frac{3}{2}} \end{pmatrix}$

$$\begin{aligned}
D_n\left(\hat{\beta}-\beta\right) &= D_n \begin{pmatrix} n & \sum_{t=1}^n t \\ \sum_{t=1}^n t & \sum_{t=1}^n t^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum u_t \\ \sum t u_t \end{pmatrix} \\
&= D_n \begin{pmatrix} n & \sum_{t=1}^n t \\ \sum_{t=1}^n t & \sum_{t=1}^n t^2 \end{pmatrix}^{-1} D_n D_n^{-1} \begin{pmatrix} \sum u_t \\ \sum t u_t \end{pmatrix} \\
&= \left(D_n^{-1} \begin{pmatrix} n & \sum_{t=1}^n t \\ \sum_{t=1}^n t & \sum_{t=1}^n t^2 \end{pmatrix} D_n^{-1}\right)^{-1} \begin{pmatrix} \frac{1}{\sqrt{n}} \sum u_t \\ \frac{1}{n^{3/2}} \sum t u_t \end{pmatrix} \\
&= \begin{pmatrix} 1 & \frac{1}{n^2}\sum_{t=1}^n t \\ \frac{1}{n^2}\sum_{t=1}^n t & \frac{1}{n^3}\sum_{t=1}^n t^2 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{\sqrt{n}} \sum u_t \\ \frac{1}{n^{3/2}} \sum t u_t \end{pmatrix}
\end{aligned}$$

The denominator

$$\begin{pmatrix} 1 & \frac{1}{n^2}\sum_{t=1}^n t \\ \frac{1}{n^2}\sum_{t=1}^n t & \frac{1}{n^3}\sum_{t=1}^n t^2 \end{pmatrix} \to \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}$$

The numerator is

$$\begin{pmatrix} \frac{1}{\sqrt{n}} \sum u_t \\ \frac{1}{n^{3/2}} \sum t u_t \end{pmatrix} = \frac{1}{\sqrt{n}} \sum \begin{pmatrix} 1 \\ \frac{t}{n} \end{pmatrix} u_t = \frac{1}{\sqrt{n}} \sum X_t u_t$$

where $X_t = \begin{pmatrix} 1 \\ \frac{t}{n} \end{pmatrix}$

$$var\left(\frac{1}{\sqrt{n}}\sum X_t u_t\right) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j' E\left[u_i u_j\right]$$

In the special case when u_i is a white noise,

$$\begin{aligned} var\left(\frac{1}{\sqrt{n}}\sum X_t u_t\right) &= \left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j'\right)\sigma^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}\begin{pmatrix} 1 & \frac{t}{n} \\ \frac{t}{n} & \frac{t^2}{n^2} \end{pmatrix}\sigma^2 \xrightarrow{d} \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}\sigma^2 \end{aligned}$$

`Zhentao Shi. Mar 21, 2023. Transcribed by Shu Shen.`