ECMT3150: The Econometrics of Financial Markets

1a. Linear Time Series Analysis

Simon Kwok
University of Sydney

Semester 1, 2021

Outline

Autoregressive (AR) Model

$\{y_t\}$ follows an $AR(p)$ model if

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t,$$

where $\{\varepsilon_t\}$ is $wn(0, \sigma_\varepsilon^2)$.
Suppose $\{y_t\}$ is covariance stationary. The *unconditional* (long run) *mean* is:

$$\mu \equiv E(y_t) = \frac{\phi_0}{1 - \sum_{i=1}^{p} \phi_i}. \tag{1}$$

AR Model

Write the $AR(p)$ model in the *demeaned form*:

$$\begin{aligned} y_t &= \mu + u_t, \\ u_t &= \sum_{i=1}^{p} \phi_i u_{t-i} + \varepsilon_t. \end{aligned} \tag{2}$$

Define the p^{th}-order polynomial function
$\phi(x) = \phi_1 x + \phi_2 x^2 + \cdots + \phi_p x^p$.
Let L be the lag operator such that $Lu_t = u_{t-1}$, $L^i u_t = u_{t-i}$, and $Lc = c$ for any constant c.
We can rewrite (2) as

$$[1 - \phi(L)]\, u_t = \varepsilon_t.$$

Autocovariance and Autocorrelation Functions

Suppose $\{y_t\}$ is covariance stationary.

- For a given integer j, the lag-j *autocovariance function* is defined as
$$\gamma_j = Cov(y_t, y_{t-j}).$$
In particular, the variance is $\gamma_0 = Var(y_t)$.
- For a given integer j, the lag-j *autocorrelation function* (ACF) is defined as
$$\rho_j = Corr(y_t, y_{t-j}) = \frac{Cov(y_t, y_{t-j})}{Var(y_t)} = \frac{\gamma_j}{\gamma_0}.$$
By convention, $\rho_0 \equiv 1$.

Both the autocovariance function and ACF are symmetric functions: $\gamma_j = \gamma_{-j}$ and $\rho_j = \rho_{-j}$.

AR(1) Model

$AR(1)$ model:

$$\begin{aligned} y_t &= \mu + u_t, \\ u_t &= \phi_1 u_{t-1} + \varepsilon_t. \end{aligned} \tag{3}$$

Q: What are the autocovariance and autocorrelation functions of y_t?

Multiply both sides of (3) by u_{t-j} for $j \geq 1$, and take expectations:

$$\begin{aligned} E(u_t u_{t-j}) &= \phi_1 E(u_{t-1} u_{t-j}) + E(\varepsilon_t u_{t-j}) \\ \gamma_j &= \phi_1 \gamma_{j-1}. \end{aligned} \tag{4}$$

Note that $E(\varepsilon_t u_{t-j}) = E[E(\varepsilon_t|\mathcal{F}_s)u_{t-j}] = 0$. (4) is the *Yule-Walker equation* of AR(1).

Set $j = 1$: $\gamma_1 = \phi_1 \gamma_0$. As $\gamma_0 = Var(u_t) = \frac{\sigma_\varepsilon^2}{1-\phi_1^2}$, we can solve for γ_1: $\gamma_1 = \frac{\sigma_\varepsilon^2 \phi_1}{1-\phi_1^2}$.

For $j > 1$, $\gamma_j = \phi_1 \gamma_{j-1}$. By recursive substitution, we have $\gamma_j = \frac{\sigma_\varepsilon^2 \phi_1^j}{1-\phi_1^2}$.

By symmetry, the lag-j autocovariance function is $\gamma_j = \frac{\sigma_\varepsilon^2 \phi_1^{|j|}}{1-\phi_1^2}$ for all integers j.

ACF of AR(1)

Recall that the ACF is $\rho_j = \frac{\gamma_j}{\gamma_0}$.
The ACF of $AR(1)$ is

$$\begin{aligned} \rho_j &= \frac{\sigma_\varepsilon^2 \phi_1^{|j|}}{1-\phi_1^2} \Bigg/ \frac{\sigma_\varepsilon^2}{1-\phi_1^2} \\ &= \phi_1^{|j|} \end{aligned}$$

for all integers j.

- When $0 < \phi_1 < 1$, the ACF decays smoothly to zero.
- When $-1 < \phi_1 < 0$, the ACF alternates between positive and negative values, but its magnitude decays smoothly to zero.

ACF of AR(1)

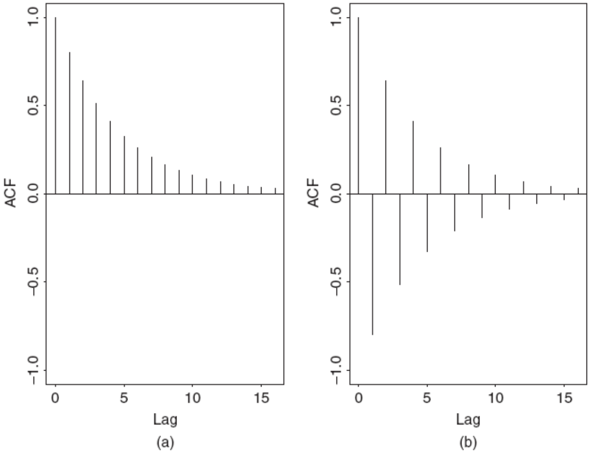

Figure 2.3 Autocorrelation function of an AR(1) model: (a) for $\phi_1 = 0.8$ and (b) for $\phi_1 = -0.8$.

ACF of AR(2)

Let $y_t \sim AR(2)$: $y_t = \mu + u_t$ and $u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \varepsilon_t$.
Ex: Show that the autocovariance function is

$$\begin{aligned}
\gamma_0 &= \frac{\sigma_\varepsilon^2}{D}(1-\phi_2), \\
\gamma_1 &= \frac{\sigma_\varepsilon^2}{D}\phi_1, \\
\gamma_j &= \phi_1\gamma_{j-1} + \phi_2\gamma_{j-2} \text{ for } j \geq 2, \\
\gamma_{-j} &= \gamma_j,
\end{aligned}$$

where $D \equiv (1+\phi_2)(1+\phi_1-\phi_2)(1-\phi_1-\phi_2)$, and that the ACF is

$$\begin{aligned}
\rho_1 &= \frac{\gamma_1}{\gamma_0} = \frac{\phi_1}{1-\phi_2}, \\
\rho_j &= \phi_1\rho_{j-1} + \phi_2\rho_{j-2} \text{ for } j \geq 2, \\
\rho_{-j} &= \rho_j.
\end{aligned}$$

ACF of AR(2)

The polynomial function of $AR(2)$ is $\phi(x) = \phi_1 x + \phi_2 x^2$.
Consider the polynomial equation

$$\begin{aligned} 1 - \phi(x) &= 0 \\ 1 - \phi_1 x - \phi_2 x^2 &= 0. \end{aligned}$$

The equation has two roots $x = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}$.

- When the roots are real ($\phi_1^2 + 4\phi_2 \geq 0$), the ACF decays smoothly.
- When the roots are complex ($\phi_1^2 + 4\phi_2 < 0$), the ACF is oscillating with an average period $k = \frac{2\pi}{\cos^{-1}(\phi_1/2\sqrt{-\phi_2})}$.
  Let the complex roots be $a \pm ib$. The period k can be solved from $\cos(\frac{2\pi}{k}) = \frac{a}{\sqrt{a^2+b^2}}$, where $a = \frac{\phi_1}{-2\phi_2}$ and $b = \frac{\sqrt{-\phi_1^2 - 4\phi_2}}{-2\phi_2}$.

ACF of AR(2)

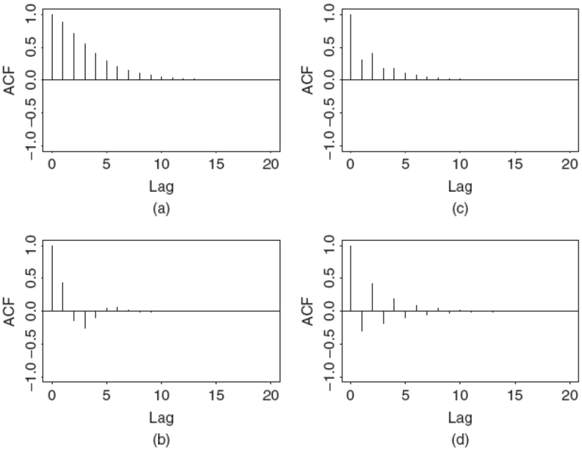

Figure 2.4 Autocorrelation function of an AR(2) model: (a) $\phi_1 = 1.2$ and $\phi_2 = -0.35$, (b) $\phi_1 = 0.6$ and $\phi_2 = -0.4$, (c) $\phi_1 = 0.2$ and $\phi_2 = 0.35$, and (d) $\phi_1 = -0.2$ and $\phi_2 = 0.35$.

PACF

Consider the following regressions:

$$
\begin{aligned}
y_t &= \phi_{0,1} + \boxed{\phi_{1,1}} y_{t-1} + e_{1t} \\
y_t &= \phi_{0,2} + \phi_{1,2} y_{t-1} + \boxed{\phi_{2,2}} y_{t-2} + e_{2t} \\
y_t &= \phi_{0,3} + \phi_{1,3} y_{t-1} + \phi_{2,3} y_{t-2} + \boxed{\phi_{3,3}} y_{t-3} + e_{3t} \\
&\vdots
\end{aligned}
$$

The lag-j population *partial autocorrelation function* (PACF) is defined as $\phi_{j,j}$ for each $j = 1, 2, \ldots$.
The lag-j sample PACF is the ordinary least squares (OLS) estimate $\hat{\phi}_{j,j}$.
If $\{y_t\}$ follows an AR(p) process, then $\hat{\phi}_{p,p} \to \phi_{p,p}$ as $T \to \infty$. In particular, $\hat{\phi}_{\ell,\ell} \to 0$ as $T \to \infty$ for $\ell > p$. The asymptotic variance of $\hat{\phi}_{\ell,\ell}$ is $\frac{1}{T}$ for $\ell > p$.

AR Model Estimation

Given the time series data $\{y_t\}_{t=1}^{T}$, we can estimate an AR(p) model by *conditional least squares* method. Conditional on the first p values: $y_1, \ldots, y_p$, we run the regression

$$y_t = \phi_0 + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t$$

for $t = p+1, p+2, \ldots, T$.
Number of observations $=$ $T - p$.
Number of parameters $= p + 1$.
Degrees of freedom $= (T - p) - (p + 1) = T - 2p - 1$.
Let $\hat{\phi}_0, \ldots, \hat{\phi}_p$ be the OLS coefficient estimates. The residual series is $\{\hat{\varepsilon}_t\}_{t=p+1}^{T}$, where

$$\hat{\varepsilon}_t = y_t - \hat{\phi}_0 - \hat{\phi}_1 y_{t-1} - \ldots - \hat{\phi}_p y_{t-p}.$$

If $\{\varepsilon_t\}$ is homoskedastic and serially uncorrelated, the variance of ε_t is consistently estimated by the sample variance of the residuals

$$\hat{\sigma}_{\varepsilon}^2 = \frac{1}{T - 2p - 1} \sum_{t=p+1}^{T} \hat{\varepsilon}_t^2.$$

AR Model Selection

- Akaike information criterion:
  $AIC = \frac{-2}{T}\ln(\text{likelihood}) + \frac{2}{T}(\text{no. of parameters})$.
- The Schwarz–Bayesian information criterion:
  $BIC = \frac{-2}{T}\ln(\text{likelihood}) + \frac{\ln(T)}{T}(\text{no. of parameters})$.

For an $AR(p)$ model with Gaussian errors, AIC and BIC become

$$\begin{aligned} AIC &= \ln(\tilde{\sigma}_p^2) + \frac{2p}{T} + \text{constant}, \\ BIC &= \ln(\tilde{\sigma}_p^2) + \frac{p\ln(T)}{T} + \text{constant}, \end{aligned}$$

where $\tilde{\sigma}_p^2$ is the maximum likelihood estimate of the error variance: $\tilde{\sigma}_p^2 = \frac{1}{T}\sum_{t=p+1}^{T}\hat{\varepsilon}_t^2$.

Goodness-of-fit

- $R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{\sum_{t=p+1}^{T} \hat{\varepsilon}_t^2}{\sum_{t=p+1}^{T} (y_t - \bar{y})^2}$.
  By SST = SSR + SSE, we have $0 \leq R^2 \leq 1$.
- Adjusted $R^2 = 1 - \frac{\frac{1}{T-2p-1} \sum_{t=p+1}^{T} \hat{\varepsilon}_t^2}{\frac{1}{T-p-1} \sum_{t=p+1}^{T} (y_t - \bar{y})^2} = 1 - \frac{\hat{\sigma}_\varepsilon^2}{\hat{\sigma}_y^2}$, where $\hat{\sigma}_y^2$ is the sample variance of y_t.
  Adjusted R^2 may fall outside the [0,1] interval.

Time Series Forecasting

The ℓ-step ahead forecast $\hat{y}_t(\ell)$ is obtained by minimizing the conditional mean squared error:

$$\hat{y}_t(\ell) = \arg\min_g E[(y_{t+\ell} - g)^2 | \mathcal{F}_t].$$

The solution is $\hat{y}_t(\ell) = E[y_{t+\ell} | \mathcal{F}_t]$.
Interpretation: the projection of $y_{t+\ell}$ on the information set at time t.
The forecast error is

$$e_t(\ell) = y_{t+\ell} - \hat{y}_t(\ell).$$

Two types of forecasts in practice:

- Conditional forecast: compute $\hat{y}_t(\ell)$ using estimated parameters, without accounting for parameter uncertainty.
- Unconditional forecast: explicitly account for parameter uncertainty. Wider confidence interval around $\hat{y}_t(\ell)$.

Forecasting with AR model

Suppose $\{y_t\}$ follows a stationary $AR(p)$ with errors $\{\varepsilon_t\}$ $\sim wn(0, \sigma_\varepsilon^2)$.
For linear models, we consider linear projection. Here, $\mathcal{F}_t$ represents the history of $\{y_t\}$ (or equivalently the history of $\{\varepsilon_t\}$) up to time t, i.e., $\mathcal{F}_t = \{y_t, y_{t-1}, \ldots\} = \{\varepsilon_t, \varepsilon_{t-1}, \ldots\}$.
For $\ell = 1$,

$$
\begin{aligned}
\hat{y}_t(1) &= E[y_{t+1}|\mathcal{F}_t] \\
&= E[\phi_0 + \phi_1 y_t + \cdots + \phi_p y_{t-p+1} + \varepsilon_{t+1}|\mathcal{F}_t] \\
&= \phi_0 + \phi_1 y_t + \cdots + \phi_p y_{t-p+1},
\end{aligned}
$$

as $y_t, \ldots, y_{t-p+1}$ are known given $\mathcal{F}_t$, and $E[\varepsilon_{t+1}|\mathcal{F}_t] = E[\varepsilon_{t+1}|\varepsilon_t, \varepsilon_{t-1}, \ldots] = 0.$[1]
The forecast error is $e_t(1) = y_{t+1} - \hat{y}_t(1) = \varepsilon_{t+1}$, with variance $Var[e_t(1)] = \sigma_\varepsilon^2$.

[1] This would be invalid if the condition expectation is a nonlinear projection, which is typically the case for a nonlinear model. A stronger *mds* assumption is required to ensure $E(\varepsilon_{t+1}|\mathcal{F}_t) = 0$.

Forecasting with AR model

For $\ell = 2$,

$$
\begin{aligned}
\hat{y}_t(2) &= E[y_{t+2}|\mathcal{F}_t] \\
&= E[\phi_0 + \phi_1 y_{t+1} + \phi_2 y_t + \cdots + \phi_p y_{t-p+2} + \varepsilon_{t+2}|\mathcal{F}_t] \\
&= \phi_0 + \phi_1 E[y_{t+1}|\mathcal{F}_t] + \phi_2 y_t + \cdots + \phi_p y_{t-p+2} \\
&= \phi_0 + \phi_1 \hat{y}_t(1) + \phi_2 y_t + \cdots + \phi_p y_{t-p+2}.
\end{aligned}
$$

The forecast error is
$e_t(2) = y_{t+2} - \hat{y}_t(2) = \phi_1 e_t(1) + \varepsilon_{t+2} = \phi_1 \varepsilon_{t+1} + \varepsilon_{t+2}$, with variance

$$
\begin{aligned}
Var[e_t(2)] &= \phi_1^2 Var(\varepsilon_{t+1}) + Var(\varepsilon_{t+2}) + 2\phi_1 Cov(\varepsilon_{t+1}, \varepsilon_{t+2}) \\
&= \phi_1^2 \sigma_\varepsilon^2 + \sigma_\varepsilon^2 + 0 \\
&= (1 + \phi_1^2)\sigma_\varepsilon^2.
\end{aligned}
$$

Forecasting with AR model

Q: Let $\ell > p$. What is the ℓ-step ahead forecast of a stationary $AR(p)$ model? What happens to the forecast when $\ell \to \infty$?

A: With $\ell > p$, the ℓ-step ahead forecast is

$$\hat{y}_t(\ell) = \phi_0 + \phi_1 \hat{y}_t(\ell - 1) + \cdots + \phi_p \hat{y}_t(\ell - p). \tag{5}$$

By stationarity, $\hat{y}_t(\ell)$ converges to a limit b as $\ell \to \infty$. The limit must satisfy (5), yielding the solution

$$b = \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p},$$

which is $\mu = E(y_t)$. This is the *mean-reverting property*.