Linear Time Series Models : Univariate Time Series

Yanping YI

February 19, 2013

Table of contents

- Financial Time Series: collection of a financial measurement over time
- Example: log return r_t of a stock
- Data: $\{r_1, r_2, \cdots, r_T\}$ (T data points)
- Purpose: What is the information contained in $\{r_t\}_{t=1}^{T}$?

Stationarity

Stationarity: the foundation of time series analysis is stationarity.

- Strict stationarity: distributions are time-invariant. This is a very strong condition that is hard to verify empirically.
- Weak stationarity: first 2 moments are time-invariant. In this course, we are mainly concerned with weakly stationary series.

Weak stationarity

It is also called covariance stationary.

- Mean (or expectation) of returns:

$$E(r_t) = \mu \text{ } for \text{ } all \text{ } t$$

- Variance (variability) of returns:

$$Var(r_t) = E\left[(r_t - \mu)^2\right] = \sigma^2 \text{ } for \text{ } all \text{ } t$$

- Lag-k autocovariance :

$$\gamma_k = cov(r_t, r_{t-k}) = E\left[(r_t - \mu)(r_{t-k} - \mu)\right] \text{ } for \text{ } all \text{ } t \text{ } and \text{ } any \text{ } k$$

a plot of γ_k against k is called the autocovariance function.

Stationary Data

- What does weak stationarity mean in practice?
- Past: time plot of $\{r_t\}_{t=1}^{T}$ fluctuates with constant variation around a fixed level!
- Future: the first 2 moments of future r_t are the same as those of the data so that meaningful inferences (e.g., prediction) can be made.

Why Stationary data?

- To be able to obtain meaningful sample statistics such as means, variances, and correlations with other variables. Such statistics are useful as descriptors of future behavior only if the series is stationary.
- To be able to have a consistent and reliable forecast of economic variables.
- To have a stable and efficient prediction on the movement of economic variables.
- We will study how to deal with non-stationary data in later chapters/lectures.

- In the finance literature, it is common to assume that an asset return series is weakly stationary.
- Serial correlations or autocorrelations play an important role in understanding **linear** time series models.

Autocorrelation Function

Consider a weakly stationary return series r_t .

- Lag-k serial (or auto-) correlation:

$$\rho_k = \frac{cov(r_t, r_{t-k})}{Var(r_t)} = \frac{\gamma_k}{\gamma_0}$$

  a plot of ρ_k against k is called the autocorrelation function (ACF).

- Note: $\rho_0 = 1$ and $\rho_k = \rho_{-k}$ *for* $k \neq 0$. Why?
  Existence of serial correlations implies that the return is predictable, indicating market inefficiency.

Sample Autocorrelation Function

Lag-k sample autocorrelation:

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{T}(r_t - \overline{r})(r_{t-k} - \overline{r})}{\sum_{t=1}^{T}(r_t - \overline{r})^2} \qquad 0 \leq k < T - 1$$

where $\overline{r}$ is the sample mean and T is the sample size.

Sample Autocorrelation Function

- The statistics $\hat{\rho}_1, \hat{\rho}_2, \ldots$ is called the sample autocorrelation function (ACF) of $\{r_t\}$.
- As a matter of fact, a linear time series model can be characterized by its ACF, and linear time series modeling makes use of the sample ACF to capture the **linear** dynamic of the data.

Individual test

- For a given positive integer l, test $H_0 : \rho_l = 0$ against $H_a : \rho_l \neq 0$.
- If $\{r_t\}$ is a stationary Gaussian series satisfying $\rho_j = 0$ for $j \geq l$. $\hat{\rho}_l$ is asymptotically normal with mean zero and variance $\frac{1+2\sum_{i=1}^{l-1}\rho_i^2}{T}$.
- Therefore, the test statistic is

$$t\ ratio = \frac{\hat{\rho}_l}{\sqrt{(1+2\sum_{i=1}^{l-1}\hat{\rho}_i{}^2)/T}}$$

is asymptotically distributed as a standard normal random variable.

Individual test

- If for simplicity, we assume that $\{r_t\}$ is an iid sequence, how should the t ratio look like?
- What is the asymptotic variance of $\hat{\rho}_l$ for all $l \neq 0$? The confidence interval of $\hat{\rho}_l$?

Joint test (Ljung-Box statistics/Portmanteau Test):

- $H_0 : \rho_l = \cdots = \rho_m = 0$ against $H_a : \rho_i \neq 0$ for some $i \in \{1, \cdots, m\}$.
- Under the assumption that $\{r_t\}$ is an iid sequence with certain moment conditions

$$Q(m) = T(T+2)\sum_{l=1}^{m} \frac{\hat{\rho}_l^{\,2}}{T-l} \xrightarrow{\mathcal{D}} \chi^2_m$$

  where $\xrightarrow{\mathcal{D}}$ denotes "converge in distribution".
- Decision rule: reject H_0 if $Q(m) > \chi^2_m(\alpha)$, where $\chi^2_m(\alpha)$ denotes the $100(1-\alpha)th$ percentile of a chi-squared distribution with m degrees of freedom.

- To open a workfile, click File/Open → 'Eviews Workfile . . . '
- To open a foreign data file (for example, a txt file or an excel file) as a workfile, click File/Open → 'Foreign Data as Workfile . . . '
- ACF and Portmanteau Test: Quick → Series Statistics → Correlogram

A useful notation in Time Series analysis

- Definition: $Br_t = r_{t-1}$ or $Lr_t = r_{t-1}$
- $B^2 r_t = B(Br_t) = Br_{t-1} = r_{t-2}$.
- B (or L) means time shift! Br_t is the value of the series at time $t-1$

- What are the important statistics in practice?
- **Conditional** quantities, not unconditional

A proper perspective: at a time point t

- Available data (information set): $\{r_1, r_2, \cdots, r_{t-1}\} \equiv F_{t-1}$
- The return is decomposed into two parts as

$$\begin{aligned} r_t &= predictable\ part\ + not\ predictable\ part \\ &= function\ of\ elements\ of\ F_{t-1} + a_t \end{aligned}$$

In other words, given information F_{t-1}

$$\begin{aligned} r_t &= \mu_t + a_t \\ &= E(r_t|F_{t-1}) + \sigma_t \epsilon_t \end{aligned}$$

- μ_t : conditional mean of r_t
- a_t: shock or innovation at time t
- ϵ_t: an iid sequence with mean zero and variance 1
- σ_t : conditional standard deviation (commonly called volatility in finance)

- Traditional Time Series modeling is concerned with μ_t
- Model for μ_t: **mean equation**
- Volatility modeling concerns σ_t
- Model for σ_t^2 : **volatility equation**

Univariate Time Series analysis serves two purposes

- a model for μ_t (Tsay, Chapter 2)
- understanding models for σ_t^2 : properties, forecasting, etc. (Tsay, Chapter 3)

White Noise

- $\{\epsilon_t\}$ is called a white noise if

$$\begin{aligned} E(\epsilon_t) &= 0 \quad \textit{mean zero} \\ E(\epsilon_t^2) &= \sigma^2 \quad \textit{variance } \sigma^2 \\ E(\epsilon_t \epsilon_\tau) &= 0, \;\; \textit{for } t \neq \tau \quad \textit{uncorrelated across time} \end{aligned}$$

- If in addition, $\{\epsilon_t\}$ is independent across time, then it is called independent white noise.
- If furthermore, $\epsilon_t \sim N(0, \sigma^2)$, then we have the Gaussian white noise process.

Linear Time Series

$\{r_t\}$ is linear if

- the predictable part is a linear function of F_{t-1}
- $\{a_t\}$ are independent and have the same distribution. (iid)
- Mathematically, it means r_t can be written as

$$r_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i}$$

where μ is a constant, $\psi_0 = 1$ and $\{a_t\}$ is an iid sequence with mean zero and well-defined distribution. In the economic literature, a_t is the shock (or innovation) at time t and $\{\psi_i\}$ are the impulse responses of r_t.

Univariate Linear Time Series Models

- Autoregressive (AR) models
- Moving-average (MA) models
- Mixed ARMA models

Important Properties of a Model

- Stationarity condition
- Basic properties: mean, variance, serial dependence
- Empirical model building: specification, estimation, & checking
- Forecasting

Motivating Example

Quarterly growth rate of U.S. real gross national product (GNP), seasonally adjusted, from the second quarter of 1947 to the first quarter of 1991. An AR(3) model for the data is

$$r_t = 0.005 + 0.35 r_{t-1} + 0.18 r_{t-2} - 0.14 r_{t-3} + a_t \quad \hat{\sigma}_a = 0.01$$

where a_t denotes a white noise with variance σ_a^2. Given r_n, r_{n-1}, r_{n-2}, we can predict r_{n+1} as

$$\hat{r_{n+1}} = 0.005 + 0.35 r_n + 0.18 r_{n-1} - 0.14 r_{n-2}$$

Motivating Example

How do we specify this model from the data? Is it adequate for the data? What are the implications of the model? These are the questions we shall address in this lecture.

AR(1) Model

- Form: $r_t = \phi_0 + \phi_1 r_{t-1} + a_t$, where ϕ_0 and ϕ_1 are real numbers, which are referred to as "parameters" (to be estimated from the data in an application). For example,

$$r_t = 0.005 + 0.2 r_{t-1} + a_t$$

- Stationarity: necessary and sufficient condition $|\phi_1| < 1$. Why?
- Mean : $E(r_t) = \frac{\phi_0}{1-\phi_1}$

AR(1) Model

- Alternative representation: Let $E(r_t) = \mu$ be the mean of r_t so that $\mu = \frac{\phi_0}{1-\phi_1}$. Plugging in the model, we have

$$(r_t - \mu) = \phi_1(r_{t-1} - \mu) + a_t \tag{1}$$

  This model also has two parameters (μ and ϕ_1) . It explicitly uses the mean of the series.
- Variance : $Var(r_t) = \frac{\sigma_a^2}{1-\phi_1^2}$
- Autocorrelations: $\rho_1 = \phi_1$, $\rho_2 = \phi_1^2$, etc. In general $\rho_k = \phi_1^k$ and ACF ρ_k decays exponentially as k increases

Forecast: minimum squared error

- Suppose we are at the time index n and are interested in forecasting r_{n+l}, where $l \geq 1$. Let $\hat{r}_n(l)$ be the forecast of r_{n+l} using the minimum squared error loss function, i.e. the forecast $\hat{r}_n(l)$ is chosen such that

$$\hat{r}_n(l) = \arg\min_g E\left[(r_{n+l} - g)^2 | \mathcal{F}_n\right]$$

  where g is a function of the information available at time n (inclusive), that is, a function of $\mathcal{F}_n$.
- The answer is $\hat{r}_n(l) = E\left[r_{n+l} | \mathcal{F}_n\right]$!

AR(1) Model

Forecast: Suppose the forecast origin is n. For simplicity, we shall use the model representation in (1) and write $x_t = r_t - \mu$. The model then becomes $x_t = \phi_1 x_{t-1} + a_t$. Note that forecast of r_t is simply the forecast of x_t plus μ.

- 1-step ahead forecast at time n: $\hat{x}_n(1) = \phi_1 x_n$.
- 1-step ahead forecast error: $e_n(1) = x_{n+1} - \hat{x}_n(1) = a_{n+1}$. Thus, a_{n+1} is the un-predictable part of x_{n+1}. It is the shock at time n + 1!
- Variance of 1-step ahead forecast error: $Var(e_n(1)) = Var(a_{n+1}) = \sigma_a^2$.

AR(1) Model

- 2-step ahead forecast: $\hat{x}_n(2) = \phi_1 \hat{x}_n(1) = \phi_1^2 x_n$.
- 2-step ahead forecast error:
  $e_n(2) = x_{n+2} - \hat{x}_n(2) = a_{n+2} + \phi_1 a_{n+1}$.
- Variance of 2-step ahead forecast error:
  $Var(e_n(2)) = (1 + \phi_1^2)\sigma_a^2$, which is greater than or equal to $Var(e_n(1))$, implying that uncertainty in forecasts increases as the number of steps increases.

AR(1) Model

- Behavior of multi-step ahead forecasts. In general, for the l-step ahead forecast at n, we have $\hat{x}_n(l) = \phi_1^l x_n$,
- the forecast error $e_n(l) = a_{n+l} + \phi_1 a_{n+l-1} + \cdots + \phi_1^{l-1} a_{n+1}$,
- and the variance of forecast error $Var(e_n(l)) = (1 + \phi_1^2 + \cdots + \phi_1^{2(l-1)})\sigma_a^2$.

AR(1) Model

- In particular, as $l \to \infty$,

$$\hat{x}_n(l) \to 0, \quad i.e., \quad \hat{r}_n(l) \to \mu.$$

This is called the **mean-reversion** of the AR(1) process. The variance of forecast error approaches

$$Var(e_n(\infty)) = \frac{\sigma_a^2}{1-\phi_1^2} = Var(r_t)$$

In practice, it means that for the long-term forecasts serial dependence is not important. The forecast is just the sample mean and the uncertainty is simply the uncertainty about the series.

AR(1) Model

Half-life : A common way to quantify the speed of mean reversion is the half-life, which is defined as the number of periods needed so that the magnitude of the forecast becomes half of that of the forecast origin. For an AR(1) model, this mean

$$\hat{x}_n(k) = \frac{1}{2}x_n$$

Thus $\phi_1^k x_n = \frac{1}{2}x_n$. Consequently, the half-life of the AR(1) model is $k = \frac{\log(0.5)}{\log(|\phi_1|)}$. For example, if $\phi_1 = 0.5$, then $k = 1$. If $\phi_1 = 0.9$, then $k \approx 6.58$.

AR(2) Model

- Form: $r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + a_t$, or

$$(1 - \phi_1 B - \phi_2 B^2) r_t = \phi_0 + a_t$$

- Mean : $E(r_t) = \frac{\phi_0}{1-\phi_1-\phi_2}$
- ACF : $\rho_0 = 1, \; \rho_1 = \frac{\phi_1}{1-\phi_2}$,

$$\rho_l = \phi_1 \rho_{l-1} + \phi_2 \rho_{l-2}, \quad l \geq 2$$

AR(2) Model

- Stationarity condition: the roots of **characteristic equation** $1 - \phi_1 x - \phi_2 x^2 = 0$ all lie outside the unit circle.

$$x_1, x_2 = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}$$

- Inverses of the two solutions $\omega_i = x_i^{-1}$ are referred to as the **characteristic roots** of the AR(2) model.
- If both ω_i are real valued, then the AR(2) model can be factored as $(1 - \omega_1 B)(1 - \omega_2 B) r_t = \phi_0 + a_t$, i.e. an AR(1) model operates on top of another AR(1) model.

AR(2) Model

- If $\phi_1^2 + 4\phi_2 < 0$, then both ω_i are complex numbers,

$$\omega = \sqrt{-\phi_2}\left(\frac{\phi_1}{2\sqrt{-\phi_2}} \pm \frac{\sqrt{-\phi_1^2 - 4\phi_2}}{2\sqrt{-\phi_2}} i\right)$$

Then r_t shows characteristics of business cycles with average length

$$k = \frac{2\pi}{\cos^{-1}\left[\phi_1/(2\sqrt{-\phi_2})\right]}$$

where the cosine inverse is stated in radian.

AR(2) Model

- Forecasts: Similar to AR(1) models
- Do it after class as an exercise!

AR(p) Model

- Form: $r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \cdots + \phi_p r_{t-p} + a_t$, or

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) r_t = \phi_0 + a_t$$

- Mean : $E(r_t) = \frac{\phi_0}{1-\phi_1-\phi_2-\cdots-\phi_p}$
- Stationarity condition: the roots of **characteristic equation** $1 - \phi_1 x - \phi_2 x^2 - \cdots - \phi_p x^p = 0$ all lie outside the unit circle. Equivalently, stationarity requires that all characteristic roots of the model (inverses of the solutions) are less than 1 in modulus.
- For a stationary AR(p) series, the ACF satisfies

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) \rho_l = 0, \qquad \textit{for } l > 0$$

Building an AR Model : Order Specification

- Partial ACF: (naive, but effective) Use consecutive fittings
  See Text (p. 46-47) for details
  Key feature: PACF cuts off at lag p for an AR(p) model.
- Identification (model selection) would typically not be done using acf's or pacf's.
- We want to form a parsimonious model, because
  1. variance of estimators is inversely proportional to the number of degrees of freedom.
  2. models which are profligate might be inclined to fit to data specific features

Building an AR Model : Order Specification

- This gives motivation for using information criteria, which embody 2 factors
  1. a term which is a function of the RSS
  2. some penalty for adding extra parameters

  The object is to choose the number of parameters which minimizes the information criterion.

Building an AR Model : Order Specification

- Akaike information criterion:

$$AIC(l) = \underbrace{\log(\tilde{\sigma}_l^2)}_{goodness\ \ of\ \ fit} + \underbrace{\frac{2l}{T}}_{penalty\ \ function}$$

  for an AR(l) model, where $\tilde{\sigma}_l^2$ is the MLE of residual variance. Find the AR order with **minimum** AIC for $l \in \{0, \cdots, P\}$.

- BIC criterion:

$$BIC(l) = \underbrace{\log(\tilde{\sigma}_l^2)}_{goodness\ \ of\ \ fit} + \underbrace{\frac{l\log(T)}{T}}_{penalty\ \ function}$$

Building an AR Model : Order Specification

- Hannan-Quinn criterion (HQIC):

$$HQIC(l) = \underbrace{\log(\tilde{\sigma}_l^2)}_{goodness\ \ of\ \ fit} + \underbrace{\frac{2l\log(\log(T))}{T}}_{penalty\ \ function}$$

- BIC embodies a much stiffer penalty term than AIC, while HQIC is somewhere in between.

Building an AR Model : Order Specification

Which IC should be preferred if they suggest different model orders?

- BIC is strongly consistent but (inefficient). In other words, BIC will asymptotically deliver the correct model order.
- AIC is not consistent, and will typically pick bigger models. But it is generally more efficient. The average variation in selected model orders from different samples within a given population will be greater in the context of BIC than AIC.

Building an AR Model : Estimation

Least squares method or maximum likelihood method

Building an AR Model : Model checking

- Fitted model: $\hat{r}_t = \hat{\phi}_0 + \hat{\phi}_1 r_{t-1} + \cdots + \hat{\phi}_p r_{t-p}$.
- Residual : $\hat{a}_t = r_t - \hat{r}_t$.
- Residual should be close to white noise if the model is adequate. Use Ljung-Box statistics of residuals, but degrees of freedom is $m - p$, where p is the number of AR coefficients used in the model.

Building an AR Model : Eviews Commands

- Example 2.1
- Estimation : Quick → Estimate Equation → type "gnp c ar(1) ar(2) ar(3)" → LS -Least Squares (NLS and ARMA)
- Model checking: View → Residual Tests
- Forecast : Table 2.2/ Figure 2.7
  Forecast icon → estimation sample (1926m01-2007m12) → forecast sample (2008m01-2008m12)→ Dynamic
- Resize the sample : workfile → Proc → Structure/Resize

Moving-average (MA) model

- Model with finite memory!
- Some daily stock returns have minor serial correlations and can be modeled as MA or AR models.

MA(1) Model

- Form: $r_t = \mu + a_t - \theta a_{t-1}$.
- Stationarity: always stationary.
- Mean (or expectation): $E(r_t) = \mu$.
- Variance: $Var(r_t) = (1 + \theta^2)\sigma_a^2$.
- Autocovariance:
  Lag 1: $Cov(r_t, r_{t-1}) = -\theta\sigma_a^2$;
  Lag l : $Cov(r_t, r_{t-l}) = 0$ *for* $l > 1$.
  Thus, r_t is not related to $r_{t-2}, r_{t-3} \cdots$.
- ACF: $\rho_1 = \frac{-\theta}{1+\theta^2}$, $\rho_l = 0$ *for* $l > 1$. Finite memory! MA(1) models do not remember what happen two time periods ago.

MA(1) Model : Forecast

Forecast (at origin t = n) :

- 1-step ahead: $\hat{r}_n(1) = \mu - \theta a_n$. Why? Because at time n, a_n is known, but a_{n+1} is not.
- 1-step ahead forecast error: $e_n(1) = a_{n+1}$ with variance σ_a^2.
- Multi-step ahead: $\hat{r}_n(l) = \mu$ *for* $l \geq 2$ Thus, for an MA(1) model, the multi-step ahead forecasts are just the mean of the series. Why? Because the model has memory of 1 time period.
- Multi-step ahead forecast error: $e_n(l) = a_{n+l} - \theta a_{n+l-1}$
- Variance of multi-step ahead forecast error: $(1 + \theta^2)\sigma_a^2 =$ variance of r_t.

MA(1) Model : Invertibility

- Concept: r_t is a proper linear combination of a_t and the past observations $\{r_{t-1}, r_{t-2}, \cdots\}$.
- Why is it important? It provides a simple way to obtain the shock a_t. For an invertible model, the dependence of r_t on r_{t-l} converges to zero as l increases.
- Condition : $|\theta| < 1$.
- Invertibility of MA models is the dual property of stationarity for AR models.

MA(2) Model

- Form: $r_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$ or

$$r_t = \mu + (1 - \theta_1 B - \theta_2 B^2) a_t$$

- Stationary with $E(r_t) = \mu$.
- Variance: $Var(r_t) = (1 + \theta_1^2 + \theta_2^2)\sigma_a^2$.
- ACF: $\rho_1 = \frac{-\theta_1 + \theta_1\theta_2}{1+\theta_1^2+\theta_2^2}$, $\rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}$, $\rho_l = 0$ *for* $l > 2$. ACF cuts off at lag 2.
- Forecasts go the the mean after 2 periods.

MA(q) Model

- Form: $r_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$ or

$$r_t = \mu + (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) a_t, \ where\ q > 0$$

- Stationary with $E(r_t) = \mu$.
- Variance: $Var(r_t) = (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma_a^2$.
- Autocovariance:

$$\gamma_s = \begin{cases} (-\theta_s + \theta_{s+1}\theta_1 + \theta_{s+2}\theta_2 + \cdots + \theta_q \theta_{q-s})\sigma_a^2 & \text{for} \quad s \leq q \\ 0 & \text{for} \quad s > q \end{cases}$$

An MA(q) series is only linearly related to its first q-lagged values and hence is a "finite-memory" model.

Building an MA Model : Specification

- Use sample ACF
- Sample ACFs are all small after lag q for an MA(q) series. (See test of ACF.)

Building an MA Model : Estimation

Estimation: use maximum likelihood method

- Conditional: Assume $a_t = 0$ for $t \leq 0$.
- Exact: Treat a_t for $t \leq 0$ as parameters, estimate them to obtain the likelihood function.
- Exact method is preferred, but it is more computing intensive. For large sample, the improvement is only marginal.

Building an MA Model

- Model checking: examine residual series (to be white noise)
- Forecast: use the residuals $\{\hat{a}_t\}$ as $\{a_t\}$ (which can be obtained from the data and fitted parameters) to perform forecasts.

ARMA(1,1) model

Focus on the ARMA(1,1) model for

- simplicity
- useful for understanding GARCH models in Chapter 3 for volatility modeling.

ARMA(1,1) model

- Form : $r_t = \phi_0 + \phi_1 r_{t-1} + a_t - \theta_1 a_{t-1}$, where $\{a_t\}$ is a white noise series. Or

$$(1 - \phi_1 B) r_t = \phi_0 + (1 - \theta_1 B) a_t$$

  A combination of an AR(1) on the LHS and an MA(1) on the RHS.
- Stationarity: same as AR(1)
- Invertibility: same as MA(1)
- Mean: as AR(1), i.e. $E(r_t) = \frac{\phi_0}{1-\phi_1}$.

ARMA(1,1) Model

- Variance: $Var(r_t) = \frac{(1-2\phi_1\theta_1+\theta_1^2)\sigma_a^2}{1-\phi_1^2}$
- ACF: Satisfies $\rho_l = \phi_1\rho_{l-1}$, for $l > 1$, but

$$\rho_1 = \phi_1 - \frac{\theta_1\sigma_a^2}{Var(r_t)} \neq \phi_1$$

  This is the difference between AR(1) and ARMA(1,1) models.
- PACF: does not cut off at finite lags.

ARMA(p,q) Models

- Form : $r_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + a_t - \sum_{i=1}^{q} \theta_i a_{t-i}$, where $\{a_t\}$ is a white noise series and p and q are nonnegative integers. Or $(1 - \phi_1 B - \cdots - \phi_p B^p) r_t = \phi_0 + (1 - \theta_1 B - \cdots - \theta_q B^q) a_t$
- Stationarity: same as AR(p)
- Invertibility: same as MA(q)
- Mean: as AR(p), i.e. $E(r_t) = \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p}$.
- The ACF for an ARMA process will display combinations of behavior derived from the AR and MA parts, but for lags beyond q, the ACF will simply be identical to the individual AR(p) model.

Building an ARMA Model

- Specification: Use AIC, BIC, HQIC
- Estimation: conditional or exact likelihood method
- Model checking: examine residual series (to be white noise). Ljung-Box statistics of the residuals Q(m) follows asymptotically a chi-squared distribution with $m - g$ degrees of freedom (if the model is correctly specified), where g denotes the number of AR or MA coefficients fitted in the model.

Building an ARMA Model : Forecast

- MA(1) affects the 1-step ahead forecast. Others are similar to those of AR(1) models.
- Forecasts of an ARMA(p, q) model have similar characteristics as those of an AR(p) model after adjusting for the impacts of the MA component on the lower horizon forecasts.

Building an ARMA Model : Eviews Commands

- Estimation : Quick → Estimate Equation → type "gnp c ar(1) ar(2) ma(1) ma(2)" → LS -Least Squares (NLS and ARMA)
- Model checking: View → Residual Tests
- Forecast : the same as AR model
- Structure of the ARMA portion of the estimated equation: View → ARMA Structure.
  There are three views available: roots, correlogram, and impulse response.

Three model representations:

- ARMA form:
  $(1 - \phi_1 B - \cdots - \phi_p B^p) r_t = \phi_0 + (1 - \theta_1 B - \cdots - \theta_q B^q) a_t$
  compact, useful in estimation and forecasting
- AR representation: (by long division)

$$r_t = \frac{\phi_0}{1 - \theta_1 - \cdots - \theta_q} + \pi_1 r_{t-1} + \pi_2 r_{t-2} + \pi_3 r_{t-3} + \cdots + a_t$$

It tells how r_t depends on its past values.

Three model representations:

- MA representation: (by long division)

$$r_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots$$

  where $\mu = E(r_t) = \frac{\phi_0}{1-\phi_1-\cdots-\phi_p}$. It tells how r_t depends on the past shocks.
- For a stationary series, ψ_i converges to zero as $i \to \infty$. Thus, the effect of any shock is transitory. The MA representation is particularly useful in computing variances of forecast errors.

Assignment 1

- Tsay page 104 : 2.1, 2.2, 2.3, 2.4, 2.13
- Due next week in class