FE570 Financial Markets and Trading
Lecture 5. Linear Time Series Analysis and Its Applications
(Ref. Joel Hasbrouck - *Empirical Market Microstructure*)

Steve Yang

Stevens Institute of Technology

9/25/2012
Outline

1. Moving Average Models
2. ARMA Models
3. Unit-Root Nonstationary Time Series
Moving Average Model

- We start, at least in theory, an AR model with infinite order as:

  \[ r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + ... + a_t. \]  

  \hspace{1cm} (1)

  However, such an AR model is not realistic because it has infinite many parameters.

- One way to make the model practical is to assume that the coefficients \( \phi_i \) satisfy some constraints so that they are determined by a finite number of parameters.

  \[ r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2^2 r_{t-2} + \phi_3^3 r_{t-3} + ... + a_t. \]  

  \hspace{1cm} (2)

  A special case of this idea is: the coefficients depend on a single parameter \( \theta_1 \) via \( \phi_i = -\theta_1^i \) for \( i \geq 1 \).

  To be stationary, \( \theta_1 < 1 \) must hold; otherwise, \( \theta_1^i \) and the series will explode.
Moving Average Model

- The previous representation can be rewritten in a rather compact form:

\[ r_t + \phi_1 r_{t-1} + \phi_2^2 r_{t-2} + \phi_3^3 r_{t-3} + \ldots = \phi_0 + a_t. \]  \hspace{1cm} (3)

- The model for \( r_{t-1} \) is then

\[ r_{t-1} + \phi_1^1 r_{t-2} + \phi_1^2 r_{t-3} + \ldots = \phi_0 + a_{t-1}. \]  \hspace{1cm} (4)

- Multiplying the last equation with \( \theta_1 \) and subtract the result from Eq. (3), we obtain

\[ r_t = \phi_0 (1 - \theta_1) + a_t - \theta_1 a_{t-1}, \]  \hspace{1cm} (5)

which says that, except for the constant term, \( r_t \) is a weighted average of shocks \( a_t \) and \( a_{t-1} \). Therefore, the model is called an MA model of order 1 or MA(1) model for short.
General Moving Average Model

- The general form of an MA(1) model is:

\[ r_t = c_0 + a_t - \theta_1 a_{t-1}, \quad \text{or} \quad r_t = c_0 + (1 - \theta_1 B)a_t, \tag{6} \]

where \( c_0 \) is a constant and \( a_t \) is a white noise series. Similarly, an MA(2) model is in the form

\[ r_t = c_0 + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}, \tag{7} \]

- The general MA(q) model is

\[ r_t = c_0 - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}, \tag{8} \]

or

\[ r_t = c_0 - (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)a_t, \quad \text{where} \quad q > 0. \tag{9} \]

where \( B \) is a back-shift operator.
Properties of MA Models

- MA models are always weakly stationary because they are finite linear combinations of a white noise sequence for which the first two moments are time-invariant.
  - For example, consider MA(1) model, we have
    \[ E(r_t) = c_0 \]
    which is time-invariant.
  - Take the variance of MA(1) model Eq. (6), we have
    \[ \text{Var}(r_t) = \sigma_a^2 + \theta_1^2 \sigma_a^2 = (1 + \theta_1^2)\sigma_a^2. \]
    and we have \( \text{Var}(r_t) \) is time-invariant.
  - For the general MA(q) model, we have
    \[
    E(r_t) = c_0 \\
    \text{Var}(r_t) = (1 + \theta_1^2 + \theta_2^2 + ... + \theta_q^2)\sigma_a^2.
    \]
Autocorrelation Function of MA(q) Model

- Assume for simplicity that $c_0 = 0$ for an $MA(1)$ model, and we have

$$r_t r_{t-\ell} = r_{t-\ell} a_t - \theta_1 r_{t-\ell} a_{t-1}.$$

$$\gamma_1 = -\theta_1 \sigma_a^2,$$ and $$\gamma_\ell = 0,$$ for $\ell > 1$.

- Using the prior result and the fact that $\text{Var}(r_t) = (1 + \theta_1^2) \sigma_a^2$, we have

$$\rho_0 = 1, \ \rho_1 = \frac{-\theta_1}{1 + \theta_1^2}, \ \rho_\ell = 0,$$ for $\ell > 1$

- For MA(2) model, the autocorrelation coefficients are:

$$\rho_1 = \frac{-\theta_1 + \theta_1 \theta_2}{1 + \theta_1^2 + \theta_2^2}, \ \rho_2 = \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}, \ \rho_\ell = 0,$$ for $\ell > 2$
**Autocorrelation Function of MA(q) Model**

- We can conclude that for MA(1) model, the lag-1 ACF is not zero, but all higher order ACFs are zeros.
- For the MA(2) model, the ACF cuts off at lag-2.
- For an MA(q) model, the lag-$\ell$ ACF is not zero, but $\rho_\ell = 0$ for $\ell > q$. Consequently, an MA(q) series is only linearly related to its first q lagged values - ”finite-memory” model.

**Invertibility**

- Rewriting a zero-mean MA(1) model as $a_t = r_t + \theta_1 a_{t-1}$, and

  $$a_t = r_t + \theta_1 r_{t-1} + \theta_1^2 r_{t-2} + \theta_1^3 r_{t-3} + \ldots$$

- Intuitively, $\theta_1^j$ should go to zero as $j$ increases because the remote return $r_{t-j}$ should have very little impact on the current shock, if any. We require $|\theta_1| < 1$, so that MA(1) is invertable.
Identifying MA(q) Order

- For a time series \( r_t \) with ACF \( \rho_\ell \), if \( \rho_q \neq 0 \), but \( \rho_\ell = 0 \) for \( \ell > q \), then \( r_t \) follows an MA(q) model.

Estimation

- Maximum likelihood estimation is commonly used to estimate MA models. There are two approaches for evaluating the likelihood function of an MA model: *conditional likelihood method* vs. *exact likelihood method*. (We are not going to get into the details, but just note that the latter is preferred over the former).

Forecasting Using MA Models

- Assume that the forecast origin is \( h \) and let \( F_h \) denote the information available at time \( h \). For the \( l \)-step ahead forecast of an MA(1) process, the model says

\[
r_{h+1} = c_0 + a_{h+1} - \theta_1 a_h.
\]
Forecasting Using MA Models

- Taking the conditional expectation, we have

\[ \hat{r}(1) = E(r_{h+1}|F_h) = c_0 - \theta_1 a_h, \]
\[ e_h(1) = r_{h+1} - \hat{r}(1) = a_{h+1} \]

- The variance of the 1-step ahead forecast error is
\[ \text{Var}[e_h(1)] = \sigma_a^2. \]

- In practice, the quantity \( a_h \) can be obtained in several ways. For instance, assume that \( a_0 = 0 \), then \( a_1 = r_1 - c_0 \), and we can compute \( a_t \) for \( a \leq t \leq h \) recursively by using
\[ a_t = r_t - c_0 + \theta_1 a_{t-1}. \]

- For the 2-step ahead forecast, from the equation
\[ r_{h+2} = c_0 + a_{h+2} - \theta_1 a_{h+1}. \]
Forecasting Using MA Models

- Taking the conditional expectation, we have

\[ \hat{r}(2) = E(r_{h+2}|F_h) = c_0, \]
\[ e_h(2) = r_{h+2} - \hat{r}(2) = a_{h+2} - \theta_1 a_{h+1}. \]

- The variance of the 2-step ahead forecast error is

\[ \text{Var}[e_h(2)] = (1 + \theta_1^2)\sigma^2, \]
which is the variance of the model and is greater than or equal to that of the 1-step ahead forecast error. More generally, \( \hat{r}_h(\ell) = c_0 \) for \( \ell \geq 2 \).

- Similarly, for an MA(2) model, we have

\[ r_{h+\ell} = c_0 + a_{h+\ell} - \theta_1 a_{h+\ell-1} - \theta_2 a_{h+\ell-2}. \]
\[ \hat{r}_h(1) = c_0 - \theta_1 a_h - \theta_2 a_{h-1}, \]
\[ \hat{r}_h(2) = c_0 - \theta_2 a_h, \]
\[ \hat{r}_h(\ell) = c_0, \text{ for } \ell > 2 \]
Forecasting Using MA Models

- In general, for an MA(q) model, multi-step ahead forecasts go to the mean after first q steps.

Summary for AR and MA Models

- For MA models, the ACF is useful in specifying the order because the ACF cuts off at lag q for an MA(q) series.
- For AR models, the PACF is useful in order determination because the PACF cuts off at lag p for an AR(p) process.
- An MA series is always stationary, but for an AR series to be stationary, all of its characteristic roots must be less than 1 in modulus.
- For stationary series, the multi-step ahead forecasts converge to the mean of the series and the variances of forecast errors converge to the variances of the series.
Building an MA Model

- Specification: Use sample ACF. Sample ACFs are all small after lag q for an MA(q) series.
- Constant term? Check the sample mean.
- Estimation: use maximum likelihood method (Exact method is preferred, but it is more computing intensive.)
- Model checking: examine residuals (to be white noise)
- Forecast: use the residuals as \( \{a_t\} \), which can be obtained from the data and fitted parameters, to perform forecasts.

- R examples to demonstrate ACF and PACF, and build MA models.
  1. Simulated AR1, AR2, MA1, and MA2
  2. IBM stock return, and E-Mini S&P Simulated time series analysis.
ARMA Models

- In some applications, the AR and MA models discussed in the previous sections become cumbersome because one may need a high-order model with many parameters to adequately describe the dynamic structure of the data.

- To overcome this difficulty, the autoregressive moving-average (ARMA) models are introduced. An ARMA model combines the idea of AR and MA models into a compact form so that the number of parameters used is kept small.

ARMA(1,1) Model

- A time series follows an ARMA(1,1) model if it satisfies:

\[ r_t - \phi_1 r_{t-1} = \phi_0 + a_t - \theta_1 a_{t-1} \]  \hspace{1cm} (10)

where \( \{a_t\} \) is a white noise series.
Properties of ARMA(1,1) Models

- Taking expectation of Eq. 10, we have

\[ E(r_t) - \phi_1 E(r_{t-1}) = \phi_0 + E(a_t) - \theta_1 E(a_{t-1}). \]  

(11)

Because \( E(a_i) = 0 \) for all \( i \), the mean of \( r_t \) is

\[ E(r_t) = \mu = \frac{\phi_0}{1 - \phi_1}. \]  

(12)

provided that the series is weakly stationary.

- Next, we consider the autocovariance function of \( r_t \). We multiply Eq. 10 by \( a_t \) and take expectation (simplify \( \phi_0 = 0 \))

\[ E(r_t a_t) = E(a_t^2) - \theta_1 E(a_t a_{t-1}) = E(a_t^2) = \sigma_a^2. \]

\[ r_t = \phi_1 r_{t-1} + a_t - \theta_1 a_{t-1}. \]
Properties of ARMA(1,1) Models

- Taking variance of prior equation:

\[ Var(r_t) = \phi_1^2 Var(r_{t-1}) + \sigma_a^2 + \theta_1^2 \sigma_a^2 - 2\phi_1 \theta_1 E(r_{t-1}a_{t-1}). \]

- If the series \( r_t \) is weakly stationary, then \( Var(r_t) = Var(r_t - 1) \) and we have

\[ Var(r_t) = \frac{(1 - 2\phi_1 \theta_1 + \theta_1^2)\sigma_a^2}{1 - \phi_1^2}. \]

we need \( \phi_1^2 < 1 \)

- To obtain the autocovariance function of \( r_t \), we assume \( \phi_0 = 0 \) and multiply the model in 10 by \( r_{t-\ell} \) to obtain:

\[ r_t r_{t-\ell} - \phi_1 r_{t-1} r_{t-\ell} = a_t r_{t-\ell} - \theta_1 a_{t-1} r_{t-\ell}. \]
Properties of ARMA(1,1) Models

- For $\ell = 1$, taking expectation

$$\gamma_1 - \phi_1 \gamma_0 = -\theta_1 \sigma_a^2,$$

$$\gamma_\ell - \phi_1 \gamma_{\ell-1} = 0, \text{ for } \ell > 1.$$

- In terms of ACF, we have

$$\rho_1 = \phi_1 - \frac{\theta_1 \sigma_a^2}{\gamma_0}, \rho_\ell = \phi_1 \rho_{\ell-1}, \text{ for } \ell > 1.$$

- The ACF of an ARMA(1,1) model does not cut off at any finite lag.

- The stationary condition of an ARMA(1,1) model is the same as that of an AR(1) model, and the ACF of an ARMA(1,1) exhibits a pattern similar to that of an AR(1) model except that the pattern starts at lag 2.
General ARMA Models:

- A general ARMA(p,q) model is in the form

\[ r_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} - \sum_{i=1}^{q} \theta_i a_{t-i}, \]  

(13)

where \( \{a_t\} \) is a white noise series and \( p \) and \( q \) are no-negative integers. The AR and MA models are special cases of ARMA(p,q) model. Using the back-shift operator, the model can be written as

\[ (1 - \phi_1 B - \ldots - \phi_p B^p) r_t = \phi_0 + (1 - \theta_1 B - \ldots - \theta_q B^q) a_t. \]

The polynomial \( 1 - \phi_1 B - \ldots - \phi_p B^p \) is the AR polynomial of the model. Similarly, \( 1 - \theta_1 B - \ldots - \theta_q B^q \) is the MA polynomial. We require that there are no common factors between the AR and MA polynomials.
Identifying ARMA Models:

- The ACF and PACF are not informative in determining the order of an ARMA model. There are two ways to determine the order of ARMA models:
  - *The Extended Autocorrelation Function (EACF)* - If we can obtain a consistent estimate of the AR component of an ARMA model, then we can derive the MA component. From the derived MA series, we can use the ACF to identify the order of the MA component.
  - *The Information Criteria (AIC or BIC)* - Typically, for some prespecified positive integers $P$ and $Q$, one computes AIC (or BIC) for ARMA$(p,q)$ models, where $0 \leq p \leq P$ and $0 \leq q \leq Q$, and selects the model that gives the minimum AIC (or BIC).

- Once an ARMA$(p,q)$ model is specified, its parameters can be estimated by either the conditional or exact likelihood method.
- The Ljung-Box statistics of the residuals can be used to check the adequacy of a fitted model.
Forecasting Using ARMA Models:

- Denote the forecast origin by \( h \) and the available information by \( F_h \). The 1-step ahead forecast of \( r_{h+1} \) becomes:

\[
\hat{r}_h(1) = E(r_{h+1}|F_h) = \phi_0 + \sum_{i=1}^{p} \phi_i r_{h+1-i} - \sum_{i=1}^{q} \theta_i a_{h+1-i}.
\]

and the forecast error is \( e_h(1) = r_{h+1} - \hat{r}_h(1) = a_{h+1} \). The variance of 1-step ahead forecast error is \( \text{Var}[e_h(1)] = \sigma_a^2 \).

- For the \( \ell \)-step ahead forecast, we have:

\[
\hat{r}_h(\ell) = E(r_{h+\ell}|F_h) = \phi_0 + \sum_{i=1}^{p} \phi_i \hat{r}_h(\ell - i) - \sum_{i=1}^{q} \theta_i a_{h}(\ell - i).
\]

where it is understood that \( \hat{r}_h(\ell - i) = r_{h+\ell-i} \) if \( \ell - i \leq 0 \) and \( a_h(\ell - i) = 0 \) if \( \ell - i > 0 \) and \( a_h(\ell - i) = a_{h+\ell-i} \) if \( \ell - i \leq 0 \).
Forecasting Using ARMA Models:

- Thus, the multi-step ahead forecasts of an ARMA model can be computed recursively. The associated forecast error is

\[ \epsilon_h(\ell) = r_{h+\ell} - \hat{r}_h(\ell - i). \]

Three Model Representations for an ARMA Model

- The three representations of a stationary ARMA(p,q) models serve different purposes. Knowing these representations can lead to a better understanding of the model.

- The first representation is the ARMA(p,q) model in Eq. 14. This representation is compact and useful in parameter estimation. It is also useful in computing recursively multi-step ahead forecasts of \( r_t \).

- For the other two representations, we use long division of two polynomials.
Three Model Representations for an ARMA Model

- Given the following two polynomials:

\[ \phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i \]
\[ \theta(B) = 1 - \sum_{i=1}^{q} \theta_i B^i \]

we can obtain, by long divisions, that

\[ \frac{\theta(B)}{\phi(B)} = 1 + \psi_1 B + \psi_2 B^2 + \ldots \equiv \psi(B) \quad (14) \]
\[ \frac{\phi(B)}{\theta(B)} = 1 + \pi_1 B + \pi_2 B^2 + \ldots \equiv \pi(B) \quad (15) \]

from the definition \( \psi(B) \pi(B) = 1 \).
Three Model Representations for an ARMA Model

- For instance, if $\theta(B) = 1 - \theta_1 B$ and $\phi(B) = 1 - \phi_1 B$, then

\[
\frac{\phi_0}{\theta_1} = \frac{\phi_0}{1 - \theta_1 - \ldots - \theta_q} \quad \text{and} \quad \frac{\phi_0}{\phi_1} = \frac{\phi_0}{1 - \phi_1 - \ldots - \phi_p}
\]

- **AR Representation** (shows the dependence of the current return $r_t$ on the past return $r_{t-i}$ where $i > 0$)

\[
r_t = \frac{\phi_0}{1 - \theta_1 - \ldots - \theta_q} + \pi_1 r_{t-1} + \pi_2 r_{t-2} + \pi_3 r_{t-3} + \ldots \quad (16)
\]

- **MA Representation** (shows explicitly the impact of the past shock $a_{t-i}(i > 0)$ on the current return $r_t$.)

\[
r_t = \frac{\phi_0}{1 - \phi_1 - \ldots - \phi_p} + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \ldots \quad (17)
\]
Unit-Root Nonstationary Time Series

Consider an ARMA model, if one extends the model by allowing the AR polynomial to have 1 as a characteristic root, then the model becomes the well-known autoregressive integrated moving-average (ARIMA) model. A conventional approach for handling unit-root nonstationarity is to use differencing.

A time series is said to be an ARMA(p, 1, q) process if the change series

\[ c_t = y_t - y_{t-1} = (1 - B)y_t, \]

follows a stationary and invertible ARMA(p, q) model. In finance, price series are commonly believed to be nonstationary, but the log return series is stationary.

\[ r_t = \ln(p_t) - \ln(p_{t-1}) \]
Unit-Root Nonstationary Time Series

- In some scientific field, a time series $y_t$ may contain multiple unit roots and needs to be differenced multiple times to become stationary. More complex model ARIMA($p, d, q$) can be applied.

- **Unit-Root Test**
  
  whether the log price $p_t$ of an asset follows a random walk or a random walk with drift.

  $$p_t = \theta_1 p_t + e_t,$$

  $$p_t = \theta_0 + \theta_1 p_{t-1} + e_t,$$

  the null hypotheses $H_0 : \phi_1 = 1$ versus the alternative hypothesis $H_a : \phi_1 < 1$. This is well-known *Dickey-Fuller* test.

- ARIMA also removes the trend and provides a trend-stationary time series.
Example

Use R ARIMA module to fit an ARMA model

- Identify Model
- Estimate Parameters
- Model Checking
- Forecast Using Fitted Model