A Primer on
Financial Time Series Analysis

FNCE 5321
Hang Bai

Overview

- The Linear Model
- Univariate Time Series Models
- Multivariate Time Series Models

The Linear Model

- Linear models of the type below is often used by risk managers,

$$y = a + bx + \varepsilon$$

- Where $E[\varepsilon] = 0$ and x and ε are assumed to be independent.

The Importance of Data Plots

- Linear relationship between two variables can be deceiving
- Consider the four (artificial) data sets in table below
- All four data sets have 11 observations
- Observations in the four data sets are clearly different from each other
- The mean and variance of the x and y variables is exactly the same across the four data sets
- The correlation between x and y are also the same across the four pairs of variables

The Importance of Data Plots

- We also get the same regression parameter estimates in all the four cases
- Figure 3.1 scatter plots *y* against *x* in the four data sets with the regression line included. We see,
- A genuine linear relationship as in the top-left panel
- A genuine nonlinear relationship as in the top-right panel
- A biased estimate of the slope driven by an outlier observation as in the bottom-left panel
- A trivial relationship, which appears as a linear relationship again due to an outlier

Table 3.1

I		II		III		IV	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

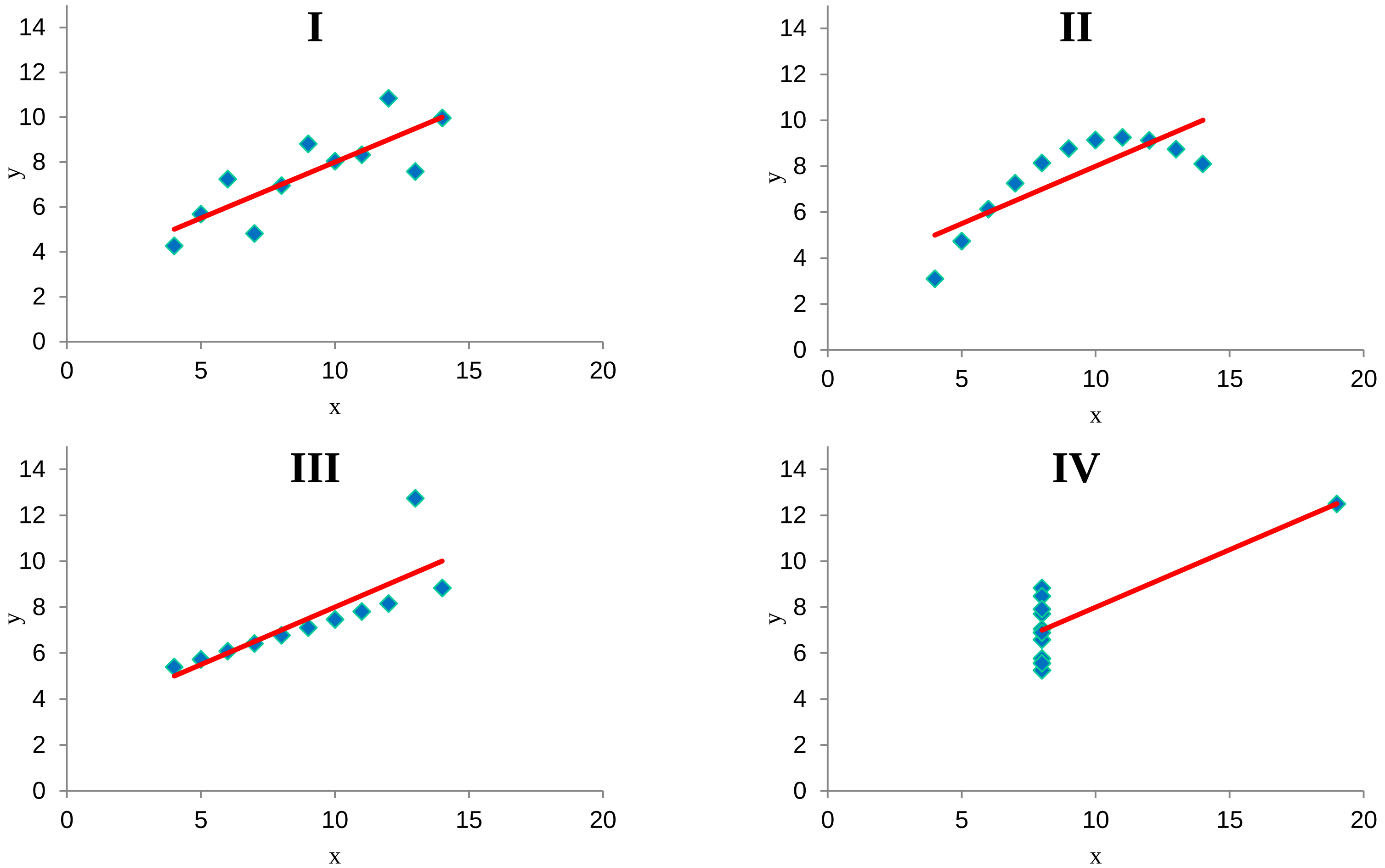
Moments

	I		II		III		IV	
Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	11.0	4.1	11.0	4.1	11.0	4.1	11.0	4.1
Correlation	0.82		0.82		0.82		0.82	

Regression

	I	II	III	IV
a	3.00	3.00	3.00	3.00
b	0.50	0.50	0.50	0.50

Figure 3.1
Scatter Plot of Four Data Sets with Regression Lines



Univariate Time Series Models

- It studies the behavior of a single random variable observed over time

- These models forecast the future values of a variable using past and current observations on the same variable

Autocorrelation

- Autocorrelation measures the dependence between the current value of a time series variable and the past value of the same variable.
- The autocorrelation for lag τ is defined as

$$\rho_\tau \equiv Corr\left[R_t, R_{t-\tau}\right] = \frac{Cov\left[R_t, R_{t-\tau}\right]}{\sqrt{Var\left[R_t\right] Var\left[R_{t-\tau}\right]}} = \frac{Cov\left[R_t, R_{t-\tau}\right]}{Var\left[R_t\right]}$$

- It captures the linear relationship between today's value and the value τ days ago

Autocorrelation

- Assuming $\{R_1, R_2, ..., R_T\}$ represents the series of returns, the sample autocorrelation measures the linear dependence between today's return, R_t, and the return τ days ago, $R_{t-\tau}$

$$\widehat{\rho}_\tau = \frac{\frac{1}{T-\tau}\sum_{t=\tau+1}^{T}\left(R_t - \overline{R}\right)\left(R_{t-\tau} - \overline{R}\right)}{\frac{1}{T}\sum_{t=1}^{T}\left(R_t - \overline{R}\right)^2}, \qquad \tau = 1, 2, ..., m < T$$

- To see the dynamics of a time series it is very useful to plot the autocorrelation function which plot $\widehat{\rho}_\tau$ on the vertical axis against τ on the horizontal axis.

Autocorrelation

- The statistical significance of a set of autocorrelations can be formally tested using the Ljung-Box statistic.
- It tests the null hypothesis that the autocorrelation for lags 1 through m are all jointly zero via

$$LB(m) = T(T+2)\sum_{\tau=1}^{m}\frac{\widehat{\rho}_{\tau}^{2}}{T-\tau} \sim \chi_m^2$$

- Where χ_m^2 denotes the chi-squared distribution. CHIINV(.,.) can be used in Excel to find the critical values.

Autoregressive (AR) Models

- If a pattern is found in the autocorrelations then we want to match that pattern in our forecasting model.
- The simplest model for this purpose is the autoregressive model of order 1, which is defined as

$$R_t = \phi_0 + \phi_1 R_{t-1} + \varepsilon_t$$

- Where, $E[\varepsilon_t] = 0$, $Var[\varepsilon_t] = \sigma_\varepsilon^2$, and $R_{t-\tau}$ and ε_t are assumed to be independent for all $\tau > 0$

Autoregressive (AR) Models

- The condition mean forecast for one period ahead under this models is,

$$E\left(R_{t+1}|R_t\right) = E\left(\phi_0 + \phi_1 R_t + \varepsilon_{t+1}|R_t\right) = \phi_0 + \phi_1 R_t$$

- By using the AR formula repeatedly we can write,

$$\begin{aligned} R_{t+\tau} &= \phi_0 + \phi_1 R_{t+\tau-1} + \varepsilon_{t+\tau} \\ &= \phi_0 + \phi_1^2 R_{t+\tau-2} + \phi_1 \varepsilon_{t+\tau-1} + \varepsilon_{t+\tau} \\ &\ldots \\ &= \phi_0 + \phi_1^\tau R_t + \phi_1^{\tau-1} \varepsilon_{t+1} + \ldots + \phi_1 \varepsilon_{t+\tau-1} + \varepsilon_{t+\tau} \end{aligned}$$

Autoregressive (AR) Models

- The multistep forecast in the AR(1) model is therefore given by

$$E\left(R_{t+\tau} | R_t\right) = \phi_0 + \phi_1^{\tau} R_t$$

- If $|\phi_1| < 1$ then the (unconditional) mean is given by, $E\left(R_t\right) = E\left(R_{t-1}\right) = \mu$, which in the AR(1) model implies

$$\begin{aligned}
E\left(R_t\right) &= \phi_0 + \phi_1 E\left(R_{t-1}\right) + E\left(\varepsilon_t\right) \\
\mu &= \phi_0 + \phi_1 \mu, \text{ and so} \\
E\left(R_t\right) &= \mu = \frac{\phi_0}{1 - \phi_1}
\end{aligned}$$

Autoregressive (AR) Models

- When $|\phi_1| < 1$, $Var(R_t) = Var\left(R_{t-1}\right)$.
- The (unconditional) variance is similarly,

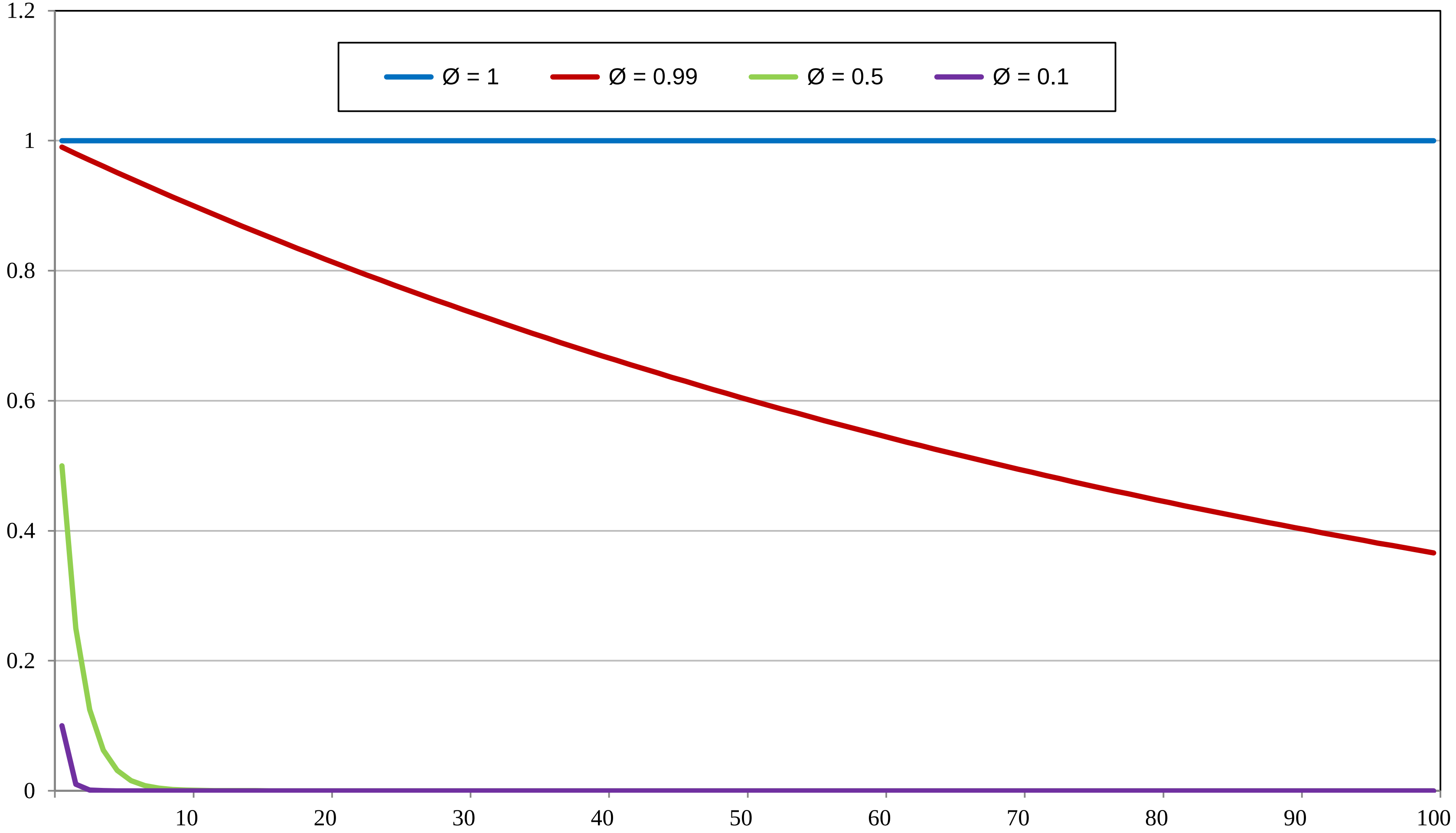
$$\begin{aligned} Var(R_t) &= \phi_1^2 Var\left(R_{t-1}\right) + Var\left(\varepsilon_t\right), \text{ so that} \\ Var(R_t) &= \frac{\sigma_\varepsilon^2}{1-\phi_1^2} \end{aligned}$$

Autoregressive (AR) Models

- To derive the ACF for AR(1) model without loss of generality we can assume that $\mu = 0$
- Then we would get,

$$
\begin{aligned}
R_t &= \phi_1 R_{t-1} + \varepsilon_t, \text{ and} \\
R_t R_{t-\tau} &= \phi_1 R_{t-1} R_{t-\tau} + \varepsilon_t R_{t-\tau}, \text{ and so} \\
E\left(R_t R_{t-\tau}\right) &= \phi_1 E\left(R_{t-1} R_{t-\tau}\right), \text{ which implies} \\
\rho_\tau &= \phi_1 \rho_{\tau-1}, \text{ so that} \\
\rho_\tau &= \phi_1^\tau \rho_0 = \phi_1^\tau
\end{aligned}
$$

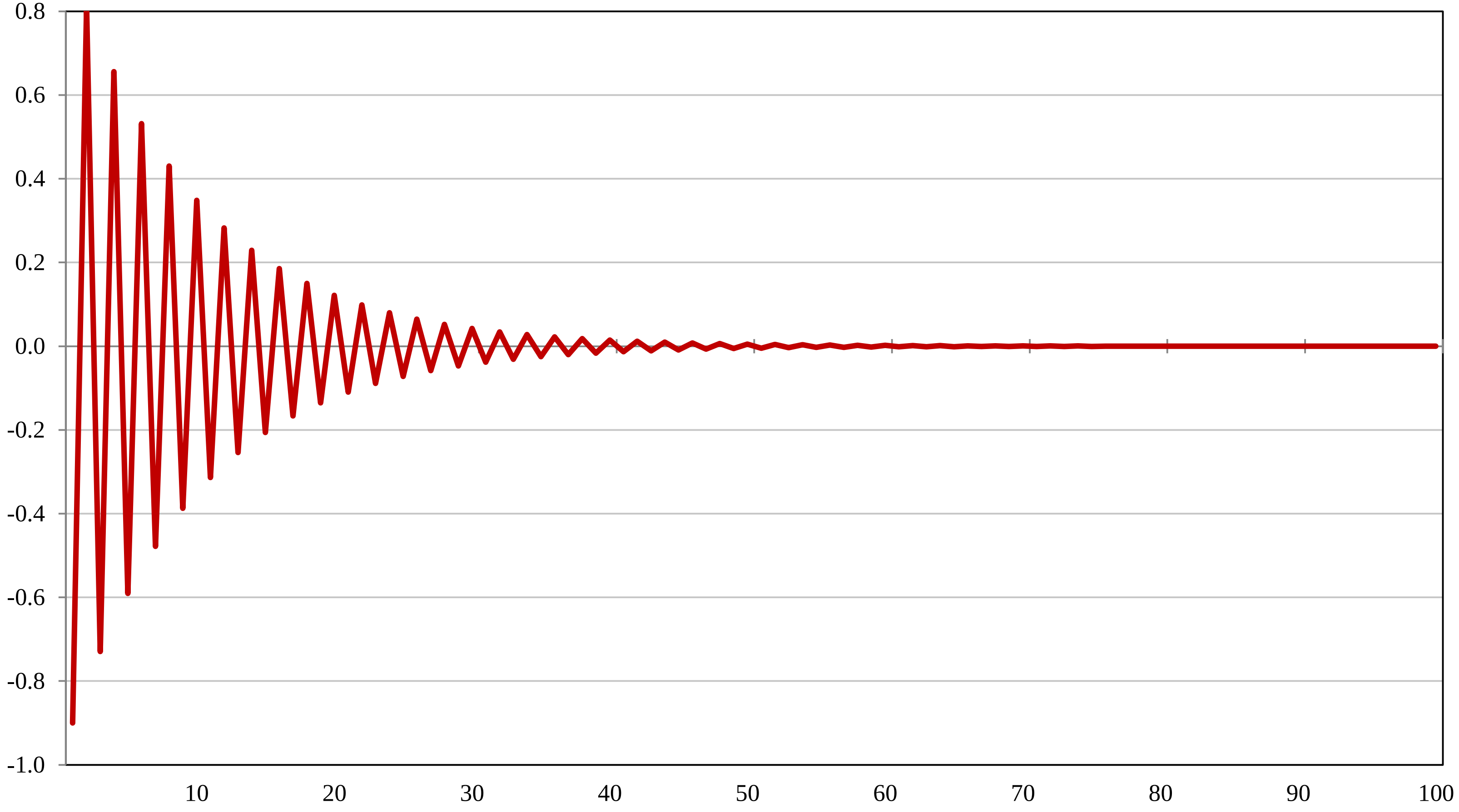
Figure 3.2
Autocorrelation Functions for AR(1) Models with Positive ϕ_1



Autoregressive (AR) Models

- Figure 3.2 shows examples of the ACF in AR(1) models
- When ϕ_1 <1 then the ACF decays to zero exponentially
- The decay is much slower when ϕ_1 = 0.99 than when it is 0.5 or 0.1
- When ϕ_1 =1 then the ACF is flat at 1. This is the case of a random walk

Figure 3.3
Autocorrelation Functions for AR(1) Models with Positive ϕ_1 =-0.9



Autoregressive (AR) Models

- Figure 3.3 shows the ACF of an AR(1) when ϕ_1=-0.9
- When ϕ_1 <0 then the ACF oscillates around zero but it still decays to zero as the lag order increases
- The ACFs in Figure 3.2 are much more common in financial risk management than are the ACFs in Figure 3.3

Autoregressive (AR) Models

- The simplest extension to the AR(1) model is the AR(2) model defined as,

$$R_t = \phi_0 + \phi_1 R_{t-1} + \phi_2 R_{t-2} + \varepsilon_t$$

- The ACF of the AR(2) is

$$\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2}, \text{ for } \tau > 1$$

- Because for example,

$$E\left(R_t R_{t-3}\right) = \phi_1 E\left(R_{t-1} R_{t-3}\right) + \phi_2 E\left(R_{t-2} R_{t-3}\right)$$

- So that,

$$\rho_3 = \phi_1 \rho_2 + \phi_2 \rho_1$$

Autoregressive (AR) Models

- In order to derive the first lag autocorrelation note that the ACF is symmetric around $\tau = 0$ meaning that,

$$Corr\left(R_t, R_{t-\tau}\right) = Corr\left(R_t, R_{t+\tau}\right) \text{ for all } \tau$$

- We therefore get that

$$E\left(R_t R_{t-1}\right) = \phi_1 E\left(R_{t-1} R_{t-1}\right) + \phi_2 E\left(R_{t-2} R_{t-1}\right)$$

- Which in turn implies that,

$$\rho_1 = \phi_1 + \phi_2 \rho_1$$

$$\rho_1 = \frac{\phi_1}{1 - \phi_2}$$

Autoregressive (AR) Models

- The general AR(p) model is simply defined as

$$E_t\left(R_{t+1}\right) \equiv E\left(R_{t+1} | R_{t,} R_{t-1,} \ldots\right) = \phi_0 + \phi_1 R_t + \ldots + \phi_p R_{t+1-p}$$

- The τ day ahead forecast can be built using

$$E_t\left(R_{t+\tau}\right) = \phi_0 + \sum_{i=1}^{p} \phi_i E_t(R_{t+\tau-i})$$

- Which is called the chain rule of forecasting.
- Note that when $\tau < i$ then,

$$E_t(R_{t+\tau-i}) = R_{t+\tau-i}$$

Autoregressive (AR) Models

- The partial autocorrelation function (PACF) gives the marginal contribution of an additional lagged term in AR models of increasing order.
- First estimate a series of AR models of increasing order:

$$
\begin{aligned}
R_t &= \phi_{0,1} + \phi_{1,1} R_{t-1} + \varepsilon_{1t} \\
R_t &= \phi_{0,2} + \phi_{1,2} R_{t-1} + \phi_{2,2} R_{t-2} + \varepsilon_{2t} \\
R_t &= \phi_{0,3} + \phi_{1,3} R_{t-1} + \phi_{2,3} R_{t-2} + \phi_{3,3} R_{t-3} + \varepsilon_{3t} \\
&\vdots \quad \vdots
\end{aligned}
$$

Autoregressive (AR) Models

- The PACF is now defined as the collection of the largest order coefficients,

$$\{\phi_{1,1}, \phi_{2,2}, \phi_{3,3}, \ldots\}$$

- Which can be plotted against the lag order just as we did for the ACF.
- The optimal lag order p in the AR(p) can be chosen as the largest p such that $\phi_{p,p}$ is significant in the PACF.
- Note that in the AR models the ACF decays exponentially whereas the PACF drops abruptly.

Moving Average Models

- In AR models the ACF dies off exponentially but in finance there are cases such as bid-ask spreads where the ACFs die off abruptly.
- These require a different type of model.
- We can consider MA(1) model defined as

$$R_t = \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

  - Where ε_t and ε_{t-1} are independent of each other and $E\left[\varepsilon_t\right] = 0$.
  - Note that

$$E\left[R_t\right] = \theta_0 \quad \text{and} \quad Var(R_t) = (1+\theta_1^2)\sigma_\varepsilon^2$$

Moving Average Models

- To derive the ACF of the MA(1) assume without loss of generality that $\theta_0 = 0$ we then have,

$$\begin{aligned}
R_t &= \varepsilon_t + \theta_1 \varepsilon_{t-1} \text{ which implies} \\
R_{t-\tau} R_t &= R_{t-\tau}\varepsilon_t + \theta_1 R_{t-\tau}\varepsilon_{t-1}, \text{ so that} \\
E\left(R_{t-1} R_t\right) &= \theta_1 \sigma_\varepsilon^2, \text{ and} \\
E\left(R_{t-\tau} R_t\right) &= 0, \text{ for } \tau > 1
\end{aligned}$$

- Using the variance expression from before, we get

$$\rho_1 = \frac{\theta_1}{1+\theta_1^2}, \text{ and } \quad \rho_\tau = 0, \text{ for } \tau > 1$$

Moving Average Models

- The MA(1) model must be estimated by numerical optimization of the likelihood function.
  - First set the unobserved $\varepsilon_0 = 0$
  - Second, set parameter starting values for θ_0,θ_1, and σ_ε^2.
  - We can use the average of R_t for θ_0, use 0 for θ_1 and use the sample variance of R_t for σ_ε^2
- Now compute time series of residuals via

$$\varepsilon_t = R_t - \theta_0 - \theta_1 \varepsilon_{t-1}, \text{ with } \varepsilon_0 = 0.$$

Moving Average Models

- If we assume that ε_t is normally distributed then

$$f\left(\varepsilon_t\right)=\frac{1}{\left(2\pi\sigma_\varepsilon^2\right)^{1/2}}\exp\left(-\frac{\varepsilon_t^2}{2\sigma_\varepsilon^2}\right)$$

- Since ε_t are assumed to be independent over time we have,

$$\begin{aligned} f\left(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_T\right) &= f\left(\varepsilon_1\right)f\left(\varepsilon_2\right)\ldots f\left(\varepsilon_T\right) \\ &= \prod_{t=1}^{T}\frac{1}{\left(2\pi\sigma_\varepsilon^2\right)^{1/2}}\exp\left(-\frac{\varepsilon_t^2}{2\sigma_\varepsilon^2}\right) \end{aligned}$$

Moving Average Models

- We can use an iterative search (using for example Solver of Excel) to find the parameters' ($\theta_0, \theta_1, \sigma_\varepsilon^2$) estimates for the MA(1)

$$
\begin{aligned}
L\left(R_1, \ldots, R_T | \theta_0, \theta_1, \sigma_\varepsilon^2\right) &= \prod_{t=1}^{T} \frac{1}{\left(2\pi\sigma_\varepsilon^2\right)^{1/2}} \exp\left(-\frac{\varepsilon_t^2}{2\sigma_\varepsilon^2}\right) \\
\text{where } \varepsilon_t &= R_t - \theta_0 - \theta_1 \varepsilon_{t-1}, \text{ with } \varepsilon_0 = 0
\end{aligned}
$$

- Once the parameters are estimated we can use the model for forecasting. The conditional mean forecast is,

$$
\begin{aligned}
E\left(R_{t+1} | R_t, R_{t-1}, \ldots\right) &= \theta_0 + \theta_1 \varepsilon_t \\
E\left(R_{t+\tau} | R_t, R_{t-1}, \ldots\right) &= \theta_0, \text{for } \tau > 1
\end{aligned}
$$

Moving Average Models

- The general MA(q) model is defined,

$$R_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

- The ACF for MA(q) is non-zero for the first q lags and then drops abruptly to zero.

ARMA Models

- We can combine AR and MA models.
- ARMA models often enables us to forecast in a parsimonious manner.
- ARMA(1,1) is defined as

$$R_t = \phi_0 + \phi_1 R_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

- The mean of the ARMA(1,1) times series is

$$E\left[R_t\right] = \phi_0 + \phi_1 E\left[R_{t-1}\right] = \phi_0 + \phi_1 E\left[R_t\right]$$

- When $|\phi_1| < 1$, $E(R_t) = \dfrac{\phi_0}{1 - \phi_1}$
- R_t will tend to fluctuate around the mean
- R_t is mean-reverting in this case

ARMA Models

- Using the fact that $E\left[R_t \varepsilon_t\right] = \sigma_\varepsilon^2$, variance is

$$Var\left[R_t\right] = \phi_1^2 Var\left[R_t\right] + \theta_1^2 \sigma_\varepsilon^2 + \sigma_\varepsilon^2 + 2\phi_1 \theta_1 \sigma_\varepsilon^2$$

- Which implies that,

$$Var\left(R_t\right) = \frac{\left(1 + 2\phi_1\theta_1 + \theta_1^2\right)\sigma_\varepsilon^2}{1 - \phi_1^2}$$

- The first order autocorrelation is given from

$$E\left[R_t R_{t-1}\right] = \phi_1 E\left[R_{t-1} R_{t-1}\right] + \theta_1 E\left[\varepsilon_{t-1} R_{t-1}\right] + E\left[\varepsilon_t R_{t-1}\right]$$

- In which we assume again that $\phi_0 = 0$.

ARMA Models

- We can write,

$$\rho_1 Var(R_t) = \phi_1 Var(R_t) + \theta_1 \sigma_\varepsilon^2$$

- So that,

$$\rho_1 = \phi_1 + \frac{\theta_1 \sigma_\varepsilon^2}{Var(R_t)}$$

- For higher order autocorrelation the MA term has no effect and we get the same structure as in the AR(1),

$$\rho_\tau = \phi_1 \rho_{\tau-1}, \text{ for } \tau > 1$$

- The general ARMA(p,q) model is,

$$R_t = \phi_0 + \sum_{i=1}^{p} \phi_i R_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Random Walks, Unit Roots, and ARIMA

- Let S_t be the closing price of an asset and let $s_t = \ln(S_t)$ so that the log returns are defined by $R_t \equiv \ln(S_t) - \ln(S_{t-1}) = s_t - s_{t-1}$
- The random walk (or martingale) model is now defined as

$$s_t = s_{t-1} + \varepsilon_t$$

- By iteratively substituting in lagged log prices we can write,

$$\begin{aligned} s_t &= s_{t-2} + \varepsilon_{t-1} + \varepsilon_t \\ s_t &= s_{t-\tau} + \varepsilon_{t-\tau+1} + \varepsilon_{t-\tau+2} + \ldots + \varepsilon_t \end{aligned}$$

Random Walks, Unit Roots, and ARIMA

- In Random Walk model the conditional mean and variance are given by,

$$E_t(s_{t+\tau}) = s_t$$

$$Var_t(s_{t+\tau}) = \tau\sigma_\varepsilon^2$$

- Equity returns typically have a small positive mean corresponding to a small positive drift in the log price. This motivates RW with drift:

$$s_t = \mu + s_{t-1} + \varepsilon_t$$

- Substituting in lagged prices back to time 0,

$$s_t = t\mu + s_0 + \varepsilon_t + \varepsilon_{t-1} + \ldots + \varepsilon_1$$

Random Walks, Unit Roots, and ARIMA

- s_t follows an ARIMA(p,1,q) model if the first difference, $s_t - s_{t-1}$, follows a mean reverting ARMA(p,q) model.
- In this case we say that s_t has a unit root.
- The random walk model has a unit root as well because $s_t - s_{t-1} = \varepsilon_t$ which is a ARMA(0,0) model

Pitfall #1: Spurious Mean-Reversion

- Consider the AR(1) model

$$\begin{aligned} s_t &= \phi_1 s_{t-1} + \varepsilon_t \Leftrightarrow \\ s_t - s_{t-1} &= (\phi_1 - 1)s_{t-1} + \varepsilon_t \end{aligned}$$

- Note, when $\phi_1 = 1$ AR(1) model has a unit root and becomes the random walk model
- The OLS estimator contains a small sample bias in dynamic models
- In an AR(1) model when the true ϕ_1 coefficient is close or equal to 1, the finite sample OLS estimate will be biased downward.
- This is known as the Hurwitz bias or the Dickey-Fuller bias

Pitfall #1: Spurious Mean-Reversion

- Econometricians are skeptical about technical trading analysis as it attempts to find dynamic patterns in prices and not returns
- Asset prices are likely to have a ϕ_1 very close to 1
- But it is likely to be estimated to be lower than 1, which in turn suggests predictability
- Asset returns have a ϕ_1 close to zero and its estimate does not suffer from bias
- Dynamic patterns in asset returns is much less likely to produce false evidence of predictability than is dynamic patterns in asset prices

Testing for Unit Roots

- Asset prices often have a ϕ_1 very close to 1
- We need to determine whether $\phi_1 = 0.99$ or 1 because the two values have very different implications for long term forecasting
- $\phi_1 = 0.99$ implies that the asset price is predictable whereas $\phi_1 = 1$ implies it is not
- Consider the AR(1) model with and without a constant term

$$
\begin{aligned}
s_t &= \phi_0 + \phi_1 s_{t-1} + \varepsilon_t \\
s_t &= \phi_1 s_{t-1} + \varepsilon_t
\end{aligned}
$$

Testing for Unit Roots

- Unit root tests have been developed to assess the null hypothesis

$$H_0 : \phi_1 = 1$$

- against the alternative hypothesis that

$$H_A : \phi_1 < 1$$

- When the null hypothesis H_0 is true, so that $\phi_1 = 1$, the unit root test does not have the usual normal distribution even when T is large
- OLS estimation of ϕ_1 to test $\phi_1 = 1$ using the usual t-test, likely leads to rejection of the null hypothesis much more often than it should

Multivariate Time Series Models

- Multivariate time series analysis consider risk models with multiple related risk factors or models with many assets
- This section will introduce the following topics:

➢ Time series regressions

➢ Spurious relationships

➢ Cointegration

➢ Cross correlations

➢ Vector autoregressions

➢ Spurious causality

Time Series Regression

- The relationship between two time series can be assessed using the regression analysis
- But the regression errors must be scrutinized carefully
- Consider a simple bivariate regression of two highly persistent series
- Example: the spot and futures price of an asset

$$s_{1t} = a + bs_{2t} + e_t$$

- To diagnose a time series regression model, we need to plot the ACF of the regression errors, e_t.

Time Series Regression

- If ACF dies off only very slowly, then we need to first-difference each series and run the regression

$$(s_{1t} - s_{1t-1}) = a + b\,(s_{2t} - s_{2t-1}) + e_t$$

- Now use the ACF on the residuals of the new regression and check for ACF dynamics
- The AR, MA, or ARMA models can be used to model any dynamics in e_t.
- After modeling and estimating the parameters in the residual time series, e_t, the entire regression model including a and b can be reestimated using MLE.

Pitfall #2: Spurious Regression

- Consider two completely unrelated times series—each with a unit root
- They are likely to appear related in a regression that has a significant *b* coefficient
- Let s_{1t} and s_{2t} be two independent random walks

$$\begin{aligned} s_{1t} &= s_{1t-1} + \varepsilon_{1t} \\ s_{2t} &= s_{2t-1} + \varepsilon_{2t} \end{aligned}$$

- where ε_{1t} and ε_{2t} are independent of each other and independent over time.
- True value of *b* is zero in the time series regression

$$s_{1t} = a + bs_{2t} + e_t$$

Pitfall #2: Spurious Regression

- However standard t-tests will tend to conclude that *b* is nonzero when in truth it is zero.
- This problem is known as spurious regression
- So, use ACF to detect spurious regression
- If the relationship between s_{1t} and s_{2t} is spurious then the error term, e_t; will have a highly persistent ACF and the regression in first differences will not show a significant estimate of *b*

$$(s_{1t} - s_{1t-1}) = a + b\,(s_{2t} - s_{2t-1}) + e_t$$

Cointegration

- Relationships between variables with unit roots are not always spurious.
- A variable with a unit root is also called integrated
- If two variables that are both integrated have a linear combination with no unit root then we say they are cointegrated.
- Examples: long-run consumption and production in an economy
- The spot and the futures price of an asset that are related via a no-arbitrage condition.

Cointegration

- The pairs trading strategy consists of two stocks whose prices tend to move together.
- If prices diverge then we buy the temporarily cheap stock and short sell the temporarily expensive stock and wait for the typical relationship between the prices to return
- Such a strategy hinges on the stock prices being cointegrated
- Consider a simple bivariate model where

$$
\begin{aligned}
s_{1t} &= \phi_0 + s_{1,t-1} + \varepsilon_{1t} \\
s_{2t} &= b s_{1t} + \varepsilon_{2t}
\end{aligned}
$$

Cointegration

- Note that s_{1t} has a unit root and that the level of s_{1t} and s_{2t} are related via b.
- Assume that ε_{1t} and ε_{2t} are independent of each other and independent over time.
- The cointegration model can be used to preserve the relationship between the variables in the long-term forecasts

$$
\begin{aligned}
E\left(s_{1,t+\tau} | s_{1t}, s_{2t}\right) &= \phi_0 \tau + s_{1t} \\
E\left(s_{2,t+\tau} | s_{1t}, s_{2t}\right) &= b\phi_0 \tau + b s_{1t}
\end{aligned}
$$

Cross-Correlations

- Consider two financial time series, $R_{1,t}$ and $R_{2,t}$
- They can be dependent in three possible ways:
- $R_{1,t}$ can lead $R_{2,t}$ (e.g., $Corr\left(R_{1,t}, R_{2,t+1}\right) \neq 0$)
- $R_{1,t}$ can lag $R_{2,t}$ (e.g., $Corr\left(R_{1,t+1}, R_{2,t}\right) \neq 0$),
- They can be contemporaneously related

  (e.g., $Corr\left(R_{1,t}, R_{2,t}\right)$)
- We use cross-correlation matrices to detect all these possible dynamic relationships
- The sample cross-correlation matrices are the multivariate analogues of the ACF function

Cross-Correlations

- For a bivariate time series, the cross-covariance matrix for lag τ is

$$\Gamma_\tau = \begin{bmatrix} Cov\left(R_{1,t}, R_{1,t-\tau}\right) & Cov\left(R_{1,t}, R_{2,t-\tau}\right) \\ Cov\left(R_{2,t}, R_{1,t-\tau}\right) & Cov\left(R_{2,t}, R_{2,t-\tau}\right) \end{bmatrix}, \tau \geq 0$$

- The two diagonal terms are the autocovariance function of $R_{1,t}$, and $R_{2,t}$, respectively
- In the general case of a *k*-dimensional time series, we have

$$\Gamma_\tau = E\left\{(R_t - E\left[R_t\right])(R_{t-\tau} - E\left[R_t\right])'\right\}, \ \tau \geq 0$$

- where R_t is now a *k* by 1 vector of variables

Cross-Correlations

- Detecting lead and lag effects is important
- For example, when relating an illiquid stock to a liquid market factor.
- The illiquidity of the stock implies price observations that are often stale, which in turn will have a spuriously low correlation with the liquid market factor.
- The stale equity price will be correlated with the lagged market factor and this lagged relationship is used to compute a liquidity-corrected measure of the dependence between the stock and the market

Vector Autoregressions (VAR)

- Consider a first-order Vector Autoregression, call it VAR(1)

$$R_t = \phi_0 + \Phi R_{t-1} + \varepsilon_t, \quad Var(\varepsilon_t) = \Sigma$$

- where R_t is again a k by 1 vector of variables
- The bivariate case is simply

$$\begin{aligned} R_{1,t} &= \phi_{0,1} + \Phi_{11} R_{1,t-1} + \Phi_{12} R_{2,t-1} + \varepsilon_{1,t} \\ R_{2,t} &= \phi_{0,1} + \Phi_{21} R_{1,t-1} + \Phi_{22} R_{2,t-1} + \varepsilon_{2,t} \\ \Sigma &= \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \end{aligned}$$

Vector Autoregressions (VAR)

- In the VAR, $R_{1,t}$ and $R_{2,t}$ are related via their covariance $\sigma_{12} = \sigma_{21}$
- The VAR only depends on lagged variables so, it is immediately useful in forecasting.
- If the variables included on the right-hand-side of each equation in the VAR are the same then the VAR is called unrestricted
- If so, OLS can be used equation-by-equation to estimate the parameters.

Pitfall #3: Spurious Causality

- We may want to see if the lagged value of $R_{2,t}$, namely $R_{2,t-1}$, is causal for the current value of $R_{1,t}$
- If so, $R_{2,t-1}$ can be used in forecasting
- Consider a simple regression of the form

$$R_{1,t} = a + bR_{2,t-1} + e_t$$

- This regression may easily lead to false conclusions if $R_{1,t}$ is persistent and so depends on its own past value

Pitfall #3: Spurious Causality

- To truly assess if $R_{2,t-1}$ causes $R_{1,t}$ we need to check if past $R_{2,t}$ was useful for forecasting current $R_{1,t}$ once the past $R_{1,t}$ has been accounted for
- This question can be answered by running a VAR model:

$$
\begin{aligned}
R_{1,t} &= \phi_{0,1} + \Phi_{11} R_{1,t-1} + \Phi_{12} R_{2,t-1} + \varepsilon_{1,t} \\
R_{2,t} &= \phi_{0,2} + \Phi_{21} R_{1,t-1} + \Phi_{22} R_{2,t-1} + \varepsilon_{2,t}
\end{aligned}
$$

Pitfall #3: Spurious Causality

- Now we can define Granger causality as follows:

$R_{2,t}$ is said to Granger cause $R_{1,t}$ if $\Phi_{12} \neq 0$

$R_{1,t}$ is said to Granger cause $R_{2,t}$ if $\Phi_{21} \neq 0$

- In some cases several lags of $R_{1,t}$ may be needed on the right-hand side of the equation for $R_{1,t}$
- We may need more lags of $R_{2,t}$ in the equation for $R_{2,t}$

Summary

- Financial asset prices and portfolio values can be viewed as examples of very persistent time series
- The three most important issues are
- Spurious detection of mean reversion-erroneously finding that a variable is mean-reverting when it is truly a random walk
- Spurious regression-erroneously finding that a variable x is significant when regressing y on x
- Spurious detection of causality-erroneously finding that the current value of x causes future values of y when in reality it cannot