CS 660: Mathematical Foundations of Analytics

Dr. Francis Parisi

Pace University

Spring 2018

Time Series Analysis

- Time Series and Stochastic Processes Basics
- Stationary Processes
- ARMA Models
- Nonstationary and Seasonal Time Series
- Multivariate Time Series
- EWMA, ARCH and GARCH

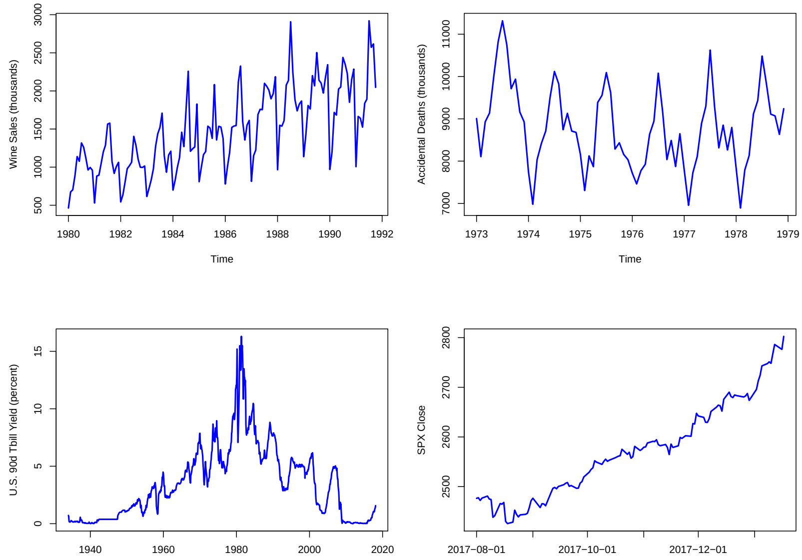
Time Series Analysis

Time Series and Stochastic Processes Basics

1. A **time series** is a sequence of data recorded in time
2. Examples of time series
   - Australian Red Wine Sales
   - Monthly Accidental Deaths
   - U.S. Interest Rates
   - Stock Prices

Time Series Analysis

Examples of time series



Time Series Analysis

Time series may

- Be random noise
- Have a trend
- Have a seasonal component
- Have both a trend and a seasonal component

Time Series Analysis

- Modeling for time series is conceptually similar to other modeling problems
- One major distinction is that usually the next value of the series is highly related to the most recent values
- This dependence decays with time

Time Series Analysis

- When we model a time series we first try simplify it as much as possible
- If the series has properties that change over time (like the mean or the variance), we preprocess the series to make it stationary (statistical properties constant in time)
- First we note that often there are both trends and seasonality in many data sets
- Linear trends are removed by doing a first difference on the time series, where one models the differences between the values at each time step instead of the value at each time step

Time Series Analysis

- Seasonality can be identified and removed as follows
- First one identifies the natural periodicity of the data. This can be done in a variety of ways, such as
  (1) through expert understanding of the dynamics
  (2) through statistical analysis using different window lengths or
  (3) through frequency analysis looking for the fundamental frequency of frequencies
- Once we identify the seasonal component we subtract it from the data

Time Series Analysis

- Most time series modeling handles many of these processes via the Autoregressive Integrated Moving Average (ARIMA) methodology
- The partial autocorrelation function explicitly accounts for the simple statistical linear correlation to the past values
- The ARIMA model training is to learn the not so obvious relationships
- The moving average process helps mitigate noise

Time Series Analysis

- The ARIMA methodology is inherently linear
- Most mainline forecasting limits itself to linear modeling
- Much like the widespread use of linear regressions
- The Generalized Autoregressive Conditional Heteroscedastic (GARCH) methodology is essentially built around autocorrelation and a moving average

Time Series Analysis

- We often deal with multiple time series problems
- One simultaneously predicts the output of multiple time series using all the time series as inputs
- This category of problems is a straightforward extension of the single time series prediction methodology
- Multiple time series are the inputs to each of the single prediction models for each separate time series
- This is known as multivariate time series analysis, also vector autoregressive or VAR

Time Series Analysis

General form of a time series is

$$X_t = m_t + S_t + Y_t \tag{7.1}$$

where m_t is the trend, S_t is the seasonal component, and Y_t is a mean 0 **stationary** series

Our goal: Estimate m_t, S_t and eliminate them to get a mean 0, stationary series

Time Series Analysis

Definition

Let $\{X_t\}$ be a time series with $\mathbb{E}\,X_t^2 < \infty$. The **mean function** of $\{X_t\}$ is

$$\mu_X(t) = \mathbb{E}(X_t)$$

The **covariance function** of $\{X_t\}$ is

$$\gamma_X(r,s) = \mathrm{Cov}(X_r, X_s) = \mathbb{E}[(X_r - \mu_X(r))(X_s - \mu_X(s))]$$

for all $r, s, t \in \mathbb{Z}$.

Time Series Analysis

Definition (2)

$\{X_t\}$ is **(weakly) stationary** if

(i) $\mu_X(t)$ is independent of t

and

(ii) $\gamma_X(t+h,t)$ is independent of t for each h

- **Strict stationarity** of a time series $\{X_t, t=0,\pm 1,\ldots\}$ is defined by the condition that $(X_1,\ldots,X_n)$ and $(X_{1+h},\ldots,X_{n+h})$ have the same joint distributions for all integers h and $n>0$
- It is easy to check that if $\{X_t\}$ is strictly stationary and $\mathbb{E}X_t^2<\infty$ for all t, then $\{X_t\}$ is also weakly stationary

Time Series Analysis

Given condition (ii), whenever we we use covariance function with reference to a stationary series, we'll mean the function γ_X of one variable defined by

$$\gamma_X(h) := \gamma_X(h,0) = \gamma_X(t+h,t)$$

The function $\gamma_X(h)$ is the autocovariance function at lag h

Time Series Analysis

Definition

Let $\{X_t\}$ be a stionary time series. the **autocovariance function** (ACVF) of $\{X_t\}$ is

$$\gamma_X(h) = \mathrm{Cov}(X_{t+h}, X_t)$$

The **autocorrelation function** (ACF) of $\{X_t\}$ is

$$\rho_X(h) \equiv \frac{\gamma_X(h)}{\gamma_X(0)} = \mathrm{Cor}(X_{t+h,X-t})$$

If $\{X_t\}$ is iid noise and $\mathbb{E}(X_t^2) = \sigma^2 < \infty$, then $\mathbb{E}(Xt) = 0$ for all t. By the assumed independence,

$$\gamma_X(t+h,h) = \begin{cases} \sigma^2, & \text{if } h = 0, \\ 0, & \text{if } h \neq 0 \end{cases} \tag{7.2}$$

which does not depend on t

Therefore iid noise with finite second moment is stationary and we denote it as

$$\{X_t\} \sim \mathrm{IID}(0, \sigma^2)$$

to indicate that the random variables X_t are independent and identically distributed random variables, each with mean 0 and variance σ^2.

Time Series Analysis

White Noise

If $\{Xt\}$ is a sequence of uncorrelated random variables, each with zero mean and variance σ^2, then $\{X_t\}$ is stationary with the same covariance function as the *iid* noise

Such a sequence is referred to as **white noise** with mean 0 and variance σ^2
This is indicated by the notation

$$\{X_t\} \sim \mathrm{WN}(0, \sigma^2)$$

Every $\mathrm{IID}(0, \sigma^2)$ sequence is $\mathrm{WN}(0, \sigma^2)$ but not the converse

Time Series Analysis

First-order Moving Average – $MA(1)$ process

Consider the series defined by the equation

$$X_t = Z_t + \theta Z_{t1}, t = 0, \pm 1, \dots$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$ and θ is a real-valued constant

We see that $\mathbb{E}\, X_t = 0, \mathbb{E}\, X_t^2 = \sigma^2(1+\theta^2) < \infty$, and

$$\gamma_X(h+t, t) = \begin{cases} \sigma^2(1+\theta^2), & \text{if } h = 0, \\ \sigma^2\theta, & \text{if } h = \pm 1, \\ 0, & \text{if } |h| > 1. \end{cases}$$

Time Series Analysis

Definition 2 is satisfied, and $\{Xt\}$ is stationary

The autocorrelation function of $\{Xt\}$ is

$$\rho_X(h) = \begin{cases} 1, & \text{if } h = 0, \\ \theta/(1+\theta^2), & \text{if } h = \pm 1, \\ 0, & \text{if } |h| > 1. \end{cases}$$

Let's assume now that $\{Xt\}$ is a stationary series satisfying the equation

$$X_t = \phi X_{t-1} + Z_t, t = 0, \pm 1, \dots$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2), |\phi| < 1$, and Z_t is uncorrelated with X_s for each $s < t$.

By taking expectations on each side and using the fact that $\mathbb{E}\, Z_t = 0$, we see that $\mathbb{E}\, X_t = 0$

Time Series Analysis

To find the autocorrelation function of $\{X_t\}$ we multiply each side of the equation by X_{t-h}, $(h > 0)$ and then take expectations, lastly observing that $\gamma(h) = \gamma(-h)$ we get

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \phi^{|h|},\ h = 0, \pm 1, \ldots$$

It follows from the linearity of the covariance function in each of its arguments and the fact that Z_t is uncorrelated with $X_{t-1,}$ that

$$\gamma_X(0) = \mathrm{Cov}(X_t, X_t) = \mathrm{Cov}(\phi X_{t-1} + Z_t, \phi X_{t-1} + Z_t) = \phi^2 \gamma_X(0) + \sigma^2$$

and so

$$\gamma_X(0) = \frac{\sigma^2}{1 - \phi^2}$$

- Normally we don't start with a model but with data
- To determine the degree of dependence among the observed data we use the sample ACF
- The sample ACF provides an estimate of the ACF of $\{X_t\}$

Time Series Analysis

Definition

Let $x_1, \ldots, x_n$ be observations of a time series. The **sample mean** of $x_1, \ldots, x_n$ is

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

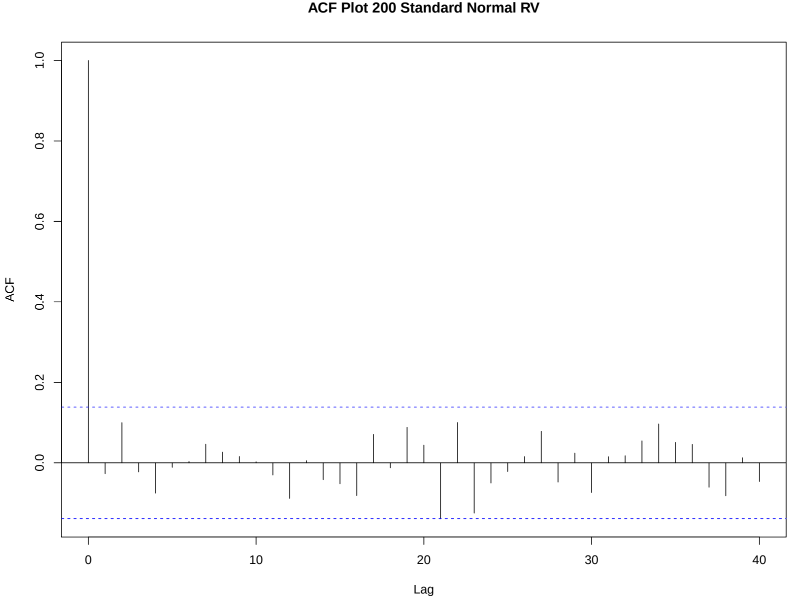
The **sample autocovariance function** is

$$\hat{\gamma}(h) := n^{-1} \sum_{i=1}^{n-|h|} (x_{t+|h|-\bar{x}})(x_t - \bar{x}), \qquad -n < h < n.$$

The **sample autocorrelation function** is

$$\hat{\rho} = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}, \qquad -n < h < n.$$

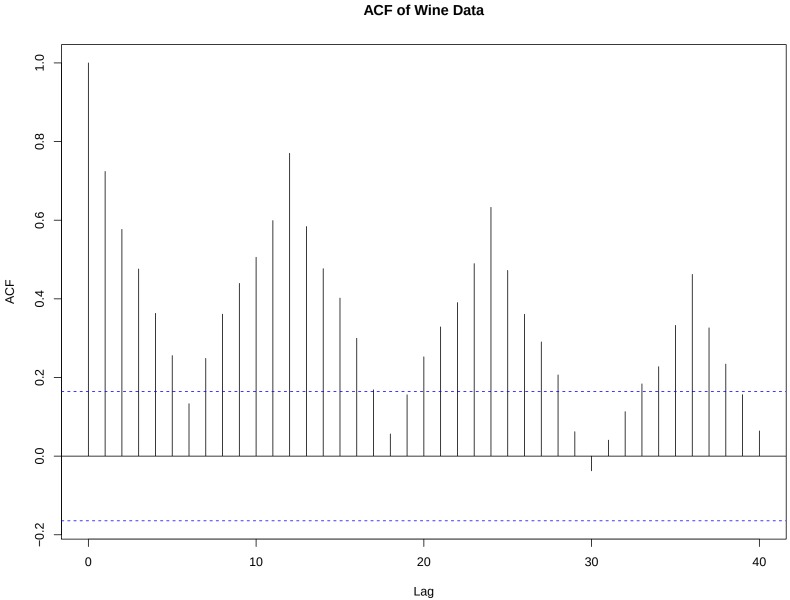
Time Series Analysis



Time Series Analysis

- For data containing a trend, $|\hat{\rho}(h)|$ will exhibit slow decay as h increases
- For data with a substantial deterministic periodic component, $|\hat{\rho}(h)|$ will exhibit similar behavior with the same periodicity

Time Series Analysis



Time Series Analysis

- Recall the general form of our time series (7.1) with a trend, seasonal component and noise
- To analyze the time series we first need to plot the data and inspect the series
- If the series displays a trend and or seasonal component, we teansform the data before further modeling

Time Series Analysis

A common way to de-trend a series without seasonality is by differencing

We define the lag-1 difference operator ∇ by

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t$$

where B is the backard shift operator

$$BX_t = X_{t-1}$$

Time Series Analysis

Powers of the operators B and ∇ are defined in the obvious way, i.e., $B^j(X_t) = X_{t-j}$ and $\nabla^j(X_t) = \nabla(\nabla^{j-1}(X_t)), j \geq 1$, with $\nabla^0(X_t) = X_t$

Polynomials in B and ∇ are manipulated in precisely the same way as polynomial functions of real variables

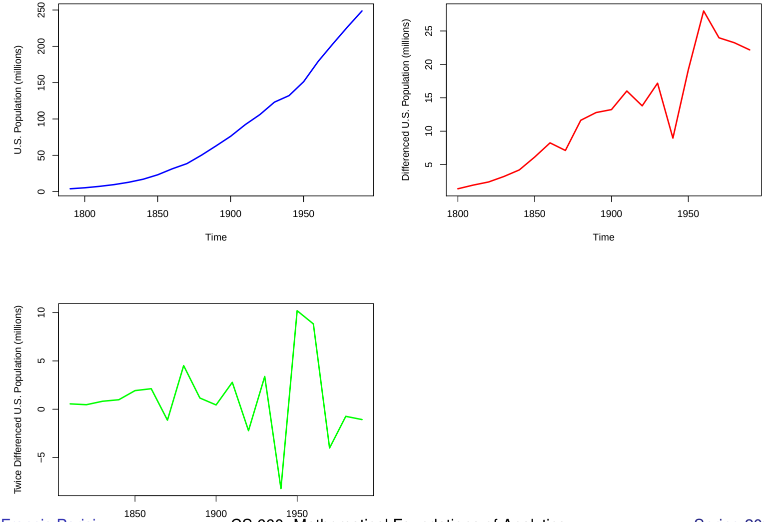
For example,

$$\begin{aligned}\nabla^2 X_t &= \nabla(\nabla(X_t)) = (1-B)(1-B)X_t = (1-2B+B^2)X_t \\ &= X_t - 2X_{t-1} + X_{t-2}\end{aligned}$$

is how we would create a twice-differenced series

Time Series Analysis

Figure 7.1: U.S. Population data twice-differenced to remove trend



Earlier we introduced and defined the **autocovariance function** $\gamma(h)$, and the **autocorrelation function** $\rho(h)$

Basic Properties of $\gamma(\cdot)$:

$$\gamma(0) \geq 0,$$
$$|\gamma(h)| \leq \gamma(0) \text{ for all } h,$$

and $\gamma(\cdot)$ is even, i.e.,

$$\gamma(h) = \gamma(-h) \text{ for all } h$$

Time Series Analysis

We also talked about (weakly) stationary time series and now look at **strictly** stationary time series

$\{X_t\}$ is a **strictly stationary time series** if

$$(X_1, \ldots, X_n)' \stackrel{d}{=} (X_{1+h}, \ldots, X_{n+h})'$$

for all integers h and $n \geq 1$. (The symbol $\stackrel{d}{=}$ means the two random vectors are equal in distribution, or have the same joint distribution function)

Time Series Analysis

Properties of s Strictly Stationary Time Series

(a) The random variables X_t are identically distributed

(b) $(X_t, \dots, X_{t+h})' \stackrel{d}{=} (X_1, \dots, X_{1+h})'$ for all integers t and h

(c) $\{X_t\}$ is weakly stationary if $\mathbb{E}\, X_t^2 < \infty$ for all t

(d) Weak stationary does *not* imply strict stationary (but the other way around)

(e) An *iid* sequence is strictly stationary

We've looked at $\mathrm{AR}(p)$ and $\mathrm{MA}(q)$ processes and now consider the $\mathrm{ARMA}(p,q)$ process

Definition

The time series $\{X_t\}$ is an $\mathbf{ARMA(1,1)}$ **process** if it is stationary and satisfies the following for all t

$$X_t - \phi X_{t-1} = Z_t + \theta Z_{t-1}, \tag{7.3}$$

where $\{Z_t\} \sim \mathrm{WN}(0,\sigma^2)$ and $\phi + \theta \neq \theta$.

Time Series Analysis

Using the backward shift operator B, we can rewrite (7.3) as

$$\phi(B)X_t = \theta(B)Z_t$$

where $\phi(B)$ and $\theta(B)$ are the linear filters

$$\phi(B) = 1 - \phi B \text{ and } \theta(B) = 1 + \theta B,$$

respectively.

Time Series Analysis

Extending to $p, q > 1$

Definition

$\{X_t\}$ is an $\mathrm{ARMA}(p, q)$ process if $\{X_t\}$ is stationary and if for every t,

$$X_t - \phi_1 X_{t1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t1} + \cdots + \theta_q Z_{t-q}, \qquad (7.4)$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$ and the polynomials $(1 - \phi_1 z - \cdots - \phi_p z^p)$ and $(1 + \theta_1 z + \cdots + \theta_q z^q)$ have no common factors

Time Series Analysis

The process $\{Xt\}$ is said to be an $\mathrm{ARMA}(p, q)$ process with mean μ if $\{X_t - \mu\}$ is an $\mathrm{ARMA}(p, q)$ process
It is convenient to use the more concise form

$$\phi(B)X_t = \theta(B)Z_t$$

where the $\phi(\cdot)$ and $\theta(\cdot)$ are the p^{th} and q^{th} degree polynomials

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$$

and

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$$

Time Series Analysis

Existence and Uniqueness

A stationary solution $\{X_t\}$ of (7.4) exists (and is also the unique stationary solution) if and only if

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0, \text{ for all } |z| = 1$$

An $\mathrm{ARMA}(p,q)$ process $\{X_t\}$ is **causal**, or a **causal function** of $\{Z_t\}$, if there exist constants $\{\psi_j\}$ such that $\sum_{j=0}^{\infty} |\psi_j| < \infty$ and

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad \text{for all } t.$$

Causality is equivalent to the condition

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0, \text{ for all } |z| \leq 1$$

We can write a causal model as an infinite-order moving average model $\mathrm{MA}(\infty)$

An $\mathrm{ARMA}(p,q)$ process $\{X_t\}$ is **invertible** if there exist constants $\{\pi_j\}$ such that $\sum_{j=0}^{\infty} |\pi_j| < \infty$ and

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \text{ for all } t.$$

Invertibility is equivalent to the condition

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q \neq 0 \text{ for all } |z| \leq 1$$

We can write a invertible model as an infinite-order autoregressive model $\mathrm{AR}(\infty)$

Time Series Analysis

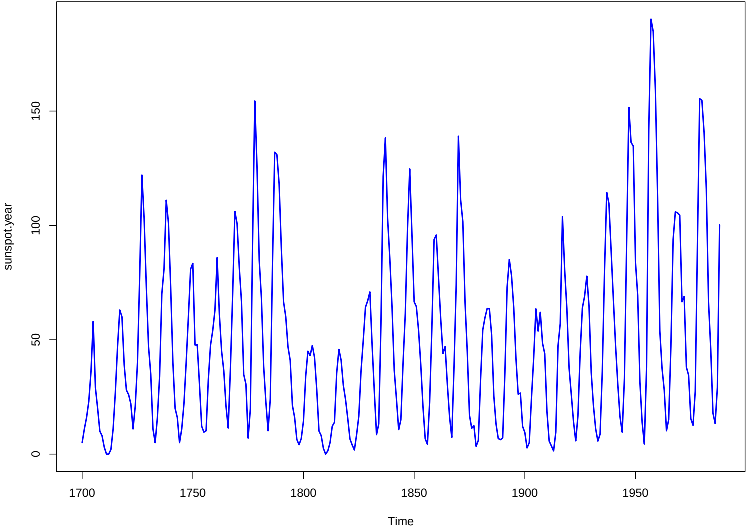
Earlier we introduced the autocorrelation function

The **partial** ACF measures the correlation between the residuals of X_t and X_{t+k} after the regressions on $X_{t+1}, \ldots, X_{t+k-1}$

The PACF of a causal $\mathrm{AR}(p)$ process is zero for lags greater than p so the PACF is helpful in identifying the order of the process

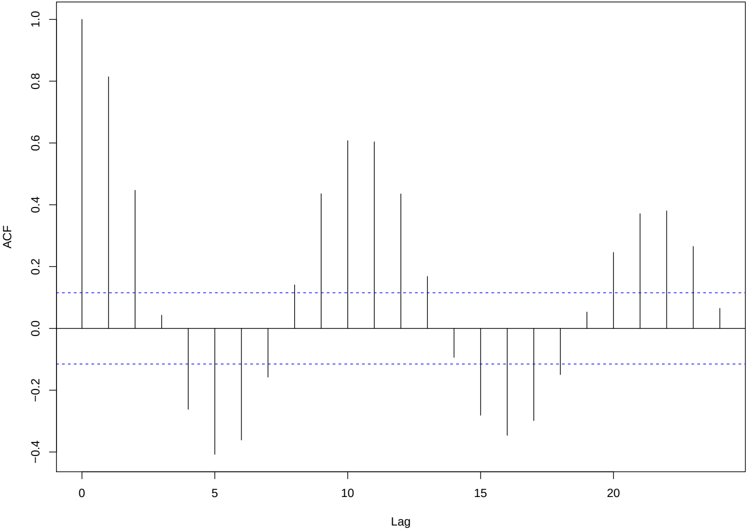
Time Series Analysis

```
ts.plot(sunspot.year, col="blue", lwd=2)
```



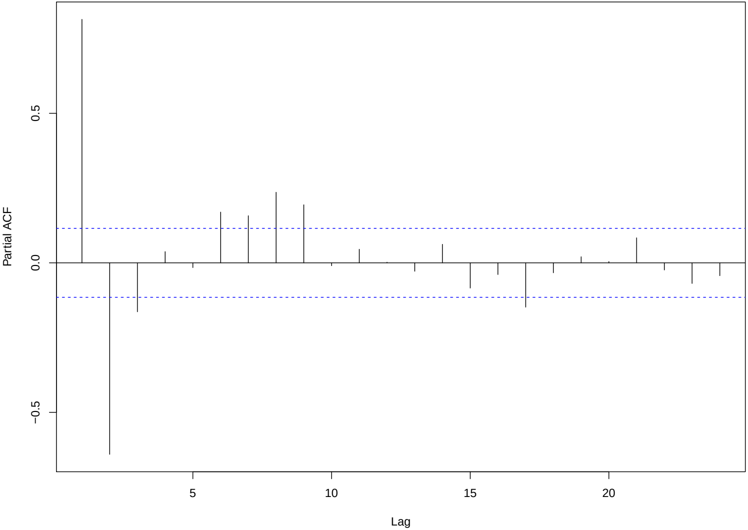
Time Series Analysis

```
acf(sunspot.year, main="")
```



Time Series Analysis

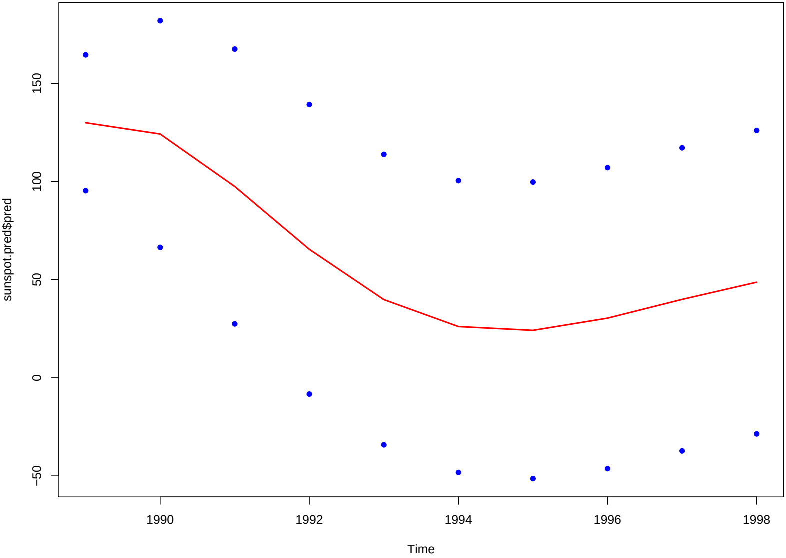
```
acf(sunspot.year, type = "partial", main="")
```



Time Series Analysis

```
sunspot.ar <- ar(sunspot.year, order.max = 2)
sunspot.pred <- predict(sunspot.ar, n.ahead = 10)

ylow = min(sunspot.pred$pred-1.96*sunspot.pred$se)
yhigh = max(sunspot.pred$pred+1.96*sunspot.pred$se)

plot(sunspot.pred$pred, type="l", col="red", lwd=2,
    ylim = c(ylow, yhigh))

points(sunspot.pred$pred+1.96*sunspot.pred$se,
     col="blue", pch=16)

points(sunspot.pred$pred-1.96*sunspot.pred$se,
     col="blue", pch=16)
```

Time Series Analysis



Recall the $\mathrm{AR}(1)$ model has the form

$$X_t = \phi X_{t-1} + Z_t$$

where $\{Z_t\} \sim WN(0, \sigma^2)$

We can write this as

$$X_t = \phi(\phi X_{t-2} + Z_{t-1}) + Z_t = \phi^2 X_{t-2} + \phi Z_{t-1} + Z_t$$

Recursive substitution leads to

$$X_t = Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} + \phi^3 Z_{t-3} + \cdots$$

Time Series Analysis

Taking expectations on both sides we get $\mathbb{E}\,X_t = 0$

$$
\begin{aligned}
\gamma(h) &= \mathrm{Cov}(X_t, X_{t-h}) = \mathbb{E}[X_t X_{t-h}] - \mathbb{E}[X_t]\,\mathbb{E}[X_{t-1}] \\
&= \mathbb{E}[X_t X_{t-h}] \text{ since } \mathbb{E}\,X_t = 0, \forall t \\
&= \mathbb{E}[(\phi X_{t-1} + Z_t) X_{t-h}] = \phi\,\mathbb{E}[X_{t-1} X_{t-h}] \\
&= \phi\,\mathbb{E}[(\phi X_{t-2} + Z_{t-1}) X_{t-h}] = \phi^2\,\mathbb{E}[X_{t-2} X_{t-h}] \\
&\vdots \\
&= \phi^h\,\mathbb{E}[X_{t-h} X_{t-h}] = \phi^h \mathrm{Var}\left(X_{t-h}\right) = \phi^h \frac{\sigma^2}{1-\phi^2}
\end{aligned}
$$

- Most real world time series are not stationary Oftentimes, differencing the data can help to achieve stationarity This leads to the class of autoregressive integrated moving average models, or ARIMA

Time Series Analysis

Definition

If d is a nonnegative integer, then $\{X_t\}$ is an $\mathrm{ARIMA}(p, d, q)$ process if $Y_t := (1 - B)^d X_t$ is a causal $\mathrm{ARMA}(p, q)$ process.

General form of the difference equation is

$$\phi * (B) X_y \equiv \phi(B)(1 - B)^d X_t = \theta(B) Z_t,$$

where $Z_t \sim WN(0, \sigma^2)$

Time Series Analysis

Example

Suppose $\{X_t\}$ is $\text{ARIMA}(1, 1, 0)$

$$(1 - \phi B)(1 - B)X_t = Z_t$$

which gives us $(1 - \phi B - B + \phi B^2)X_t = Z_t$ or

$$X_t - \phi X_{t-1} - X_{t-1} + \phi X_{t-2} = Z_t \qquad (7.5)$$

If we let $Y_t = X_t - X_{t-1}$ then (7.5) becomes

$$Y_t = \phi Y_{t-1} + Z_t$$

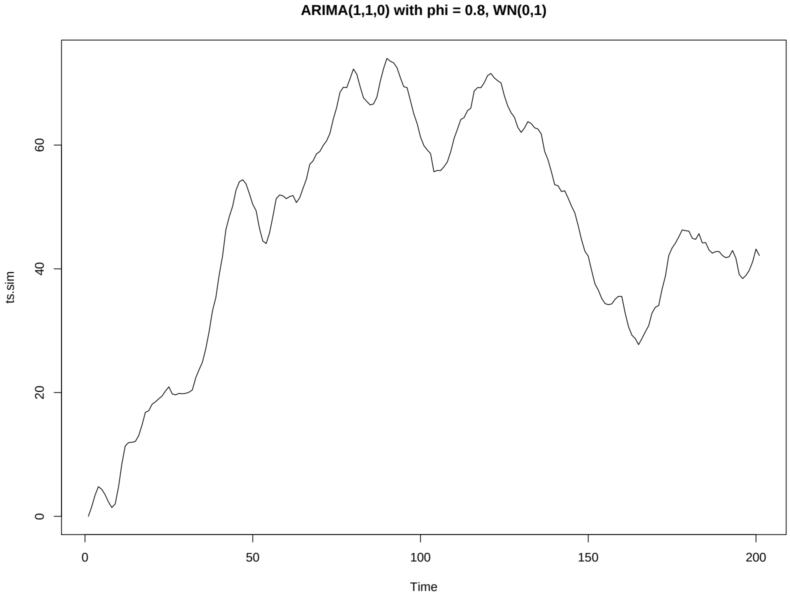
which we recognize as an $\text{AR}(1)$ model

Time Series Analysis

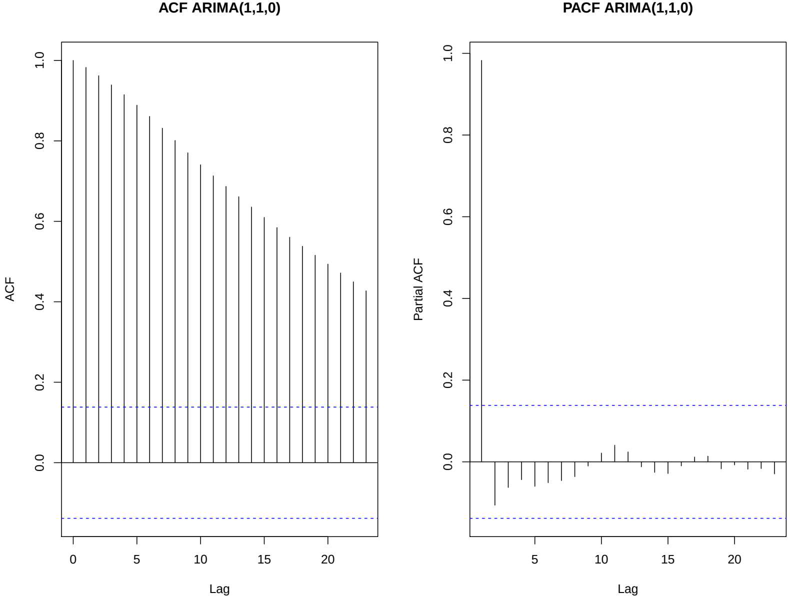
Simulate an $\mathrm{ARIMA}(1, 1, 0)$ process, and plot the data, ACF and PACF

```
ts.sim <- arima.sim(list(order = c(1,1,0),
                    ar = 0.8), n = 200, sd=1)
ts.plot(ts.sim,
        main="ARIMA(1,1,0) with phi = 0.8,
        WN(0,1)")

par(mfrow=c(1,2))
tmp <- acf(ts.sim, main="ACF ARIMA(1,1,0)")
tmp <- acf(ts.sim, type="partial",
           main="PACF ARIMA(1,1,0)")
par(mfrow=c(1,1))
```

Time Series Analysis



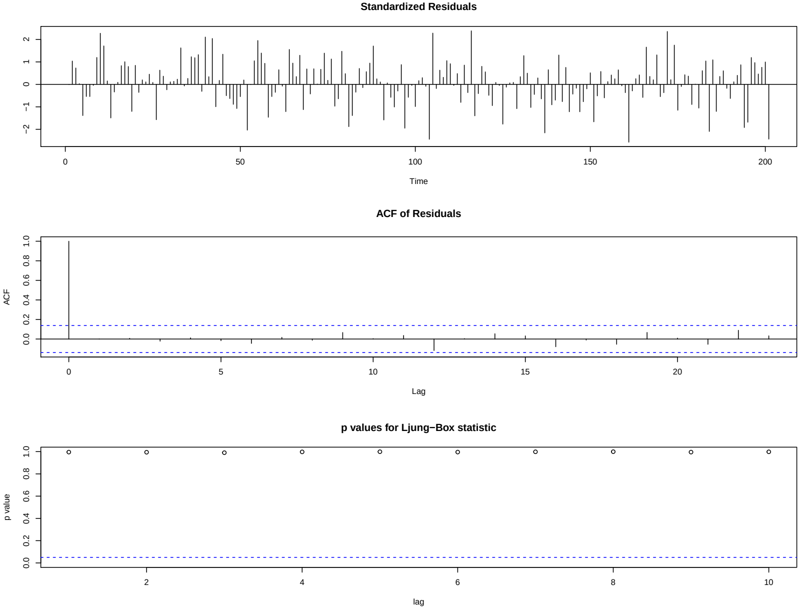
Time Series Analysis



Fitting an ARIMA model to the data

```
#fitting an ARIMA model to data
arima.out <- arima(ts.sim, order=c(1,1,0))
tsdiag(arima.out)
```

Time Series Analysis



Time Series Analysis

Dealing with Non-stationarity

- Visualization – plot the data
- Classical decomposition breaks out trend, seasonal, and random components
- Differencing the data to remove trend and seasonality

Time Series Analysis

- Inspect the ACF and PACF for estimating the order(s) of the model
- Try several models and select the one with the lowest Akaike Information Criterion Corrected, or AICC

$$AICC = -2\log\mathcal{L} + \frac{2(p+q+1)n}{n-p-q-2}$$

- Once we fit the model we need to check the assumptions such as white noise residuals
- Try to keep $p+q$ reasonable; avoid large models

Time Series Analysis

Multivariate Time Series

- Some time series are best treated as components of a vector valued (or multivariate) time series, $\{\mathbf{X}_t\}$
- In such cases, we have both serial dependence and interdependence between series
- The mean vector is given by

$$\boldsymbol{\mu} = \mathbb{E}\,\mathbf{X}$$

and the covariance matrix by

$$\Gamma(h) = \mathbb{E}[\mathbf{X}_{t+h}\mathbf{X}_t^T] - \boldsymbol{\mu}\boldsymbol{\mu}^T$$

Time Series Analysis

- The problem of identifying the MV ARMA model can be avoided is we consider the MV AR model, or the Vector AR (VAR) model
- For a bivariate time series we have

$$\boldsymbol{\mu} = \mathbb{E}\,\mathbf{X} = \begin{bmatrix} \mathbb{E}\,X_{t1} \\ \mathbb{E}\,X_{t2} \end{bmatrix}$$

and the covariance matrix by

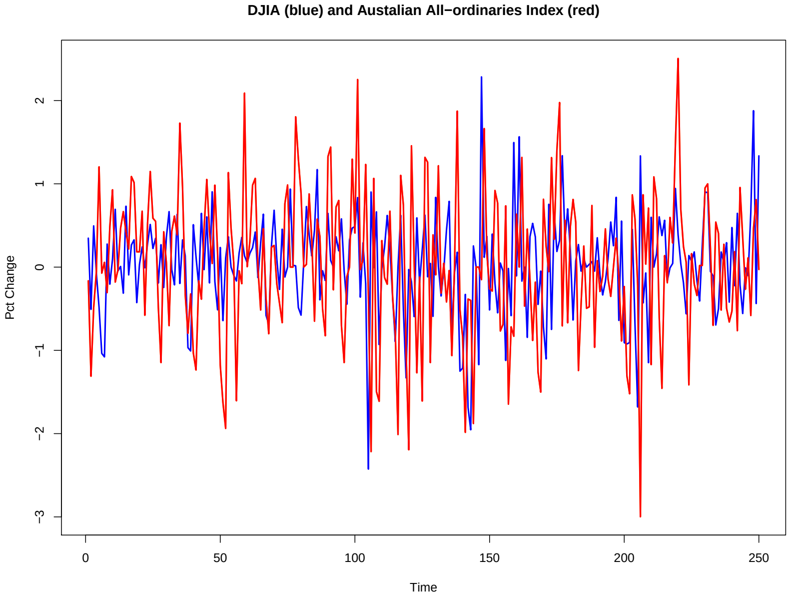
$$\Gamma(t+h,t) = \operatorname{Cov}(\mathbf{X}_{t+h},\mathbf{X}_t) \begin{bmatrix} \operatorname{Cov}(X_{t+h,1}X_{t1}) & \operatorname{Cov}(X_{t+h,1}X_{t2}) \\ \operatorname{Cov}(X_{t+h,2}X_{t1}) & \operatorname{Cov}(X_{t+h,2}X_{t2}) \end{bmatrix}$$

Time Series Analysis

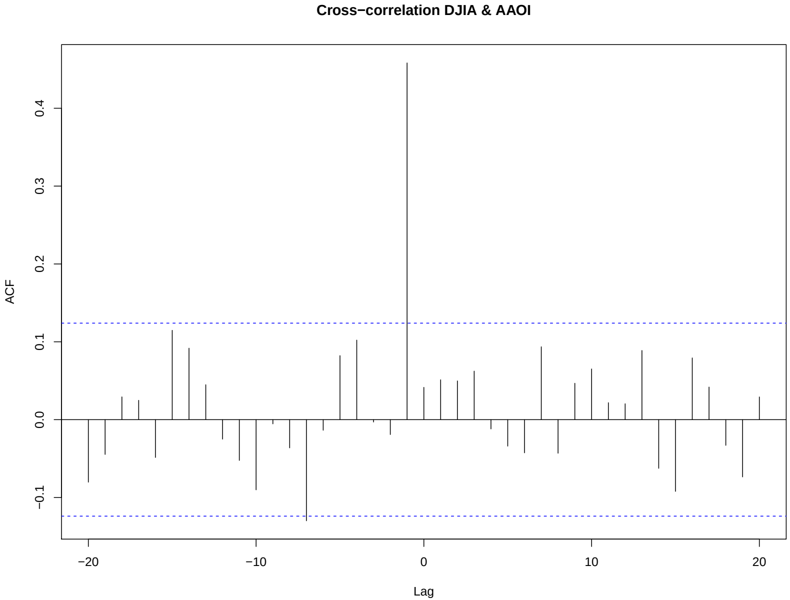
Fitting a bivariate model DJIA and AAOI

```
ts.plot(djaopc2$DJIA, type="l",
        main = "DJIA (blue) and Austalian
        All-ordinaries Index (red)",
        col="blue", lwd=2,
        ylab="Pct Change",
        ylim = c(range(djaopc2$AAOI)))
lines(djaopc2$AAOI, col="red", lwd=2)
tmp <- ccf(djaopc2$DJIA, djaopc2$AAOI,
       main="Cross-correlation DJIA & AAOI")
```

Time Series Analysis



Time Series Analysis



Time Series Analysis

```
dj.ar <- ar(djaopc2, order.max = 10)
dj.ar

Call:
ar(x = djaopc2, order.max = 10)

$ar
, , 1

        DJIA   AAOI
DJIA -0.0148 0.0357
AAOI  0.6589 0.0998

dj.ar$aic
       0        1        2        3        4
54.94030  0.00000  1.68532  8.45426 13.10385
```

Time Series Analysis

In the R output the matrix labeled **`$ar`** gives us the AR coefficients (ϕ_{ij}) for the bivariate model

$$\begin{bmatrix} Y_{t1} \\ Y_{t2} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} Y_{t-1,1} \\ Y_{t-1,2} \end{bmatrix} + \begin{bmatrix} Z_{t1} \\ Z_{t2} \end{bmatrix}$$

Time Series Analysis

Exponentially Weighted Moving Average

- Moving averages are often used to smooth volatility in data display an overall trend
- For example we may see a 60-day moving average overlaid on a chart of daily closing stock prices
- These are typically equally weighted so that the value for each day contributes equally
- Equally weihted moving averages are sensitive to extreme values

Time Series Analysis

- To mitigate the effects of extremes we can weight the observations differently often putting more weight on the most recent observations
- The **exponentially weighted moving average**, EWMA, does this by using weights that decay exponentially
- An n-period EWMA of a series $\{X_t\}$ is

$$\tilde{\mu}_t(n) = \sum_{i=0}^{n-1} w_i y_{i-1}, \qquad w_i = \frac{\lambda^{i-1}}{\sum_{i=0}^{n-1} \lambda^{i-1}}$$

  where $0 < \lambda < 1$ is the decay parameter and y_t is the observation at time t
- The closer λ is to 1, the more weight on the more recent observations

Time Series Analysis

- As $n \to \infty$, $\lambda^n \to 0,\ w_n \to 0$ and $\tilde{\mu}_t(n)$ approaches

$$\tilde{\mu}_t(\lambda) = (1-\lambda)\sum_{i=0}^{\infty} \lambda^i y_{t-i}$$

- EWMA can be computed recursively

$$\tilde{\mu}_t(\lambda) = (1-\lambda)y_t + \lambda\tilde{\mu}_{t-1}(\lambda)$$

- One often uses the first value of the series or an average over some small window as a starting value for $\tilde{\mu}_0(\lambda)$

Time Series Analysis
GARCH

- ARMA and ARIMA models focus on modeling and predicting the conditional mean of a time series
- In some areas, especially financial markets, it is important to model the volatility of the series
- Volatility is key to risk management, portfolio selection, and pricing derivatives, for example
- The class of models for time varying volatility is the generalized autoregressive heteroscedasticity (GARCH) models (also spelled *heteroskedasticity*)

Time Series Analysis

- We can model the serial correlation in squared returns using a simple $\mathrm{AR}(p)$ model
- Assuming the returns are stationary and have iid errors with mean zero
- And the assumed variance is time dependent

$$\sigma_t^2 = a_0 + a_1\epsilon_{t-1}^2 + \cdots + a_p\epsilon_{t-p}^2$$

- We have an $\mathrm{ARCH}(p)$ model and we can test for ARCH effects

Time Series Analysis

- If the ARCH test is significant we can model the volatility but this is impractical
- The $\mathrm{GARCH}(p, q)$ model extends $\mathrm{ARCH}(p)$, and when $q = 0$ we have the latter
- Under $\mathrm{GARCH}(p, q)$ the conditional variance σ_t^2 depends on the squared residuals in the p previous periods, and the conditional variance of the q previous periods

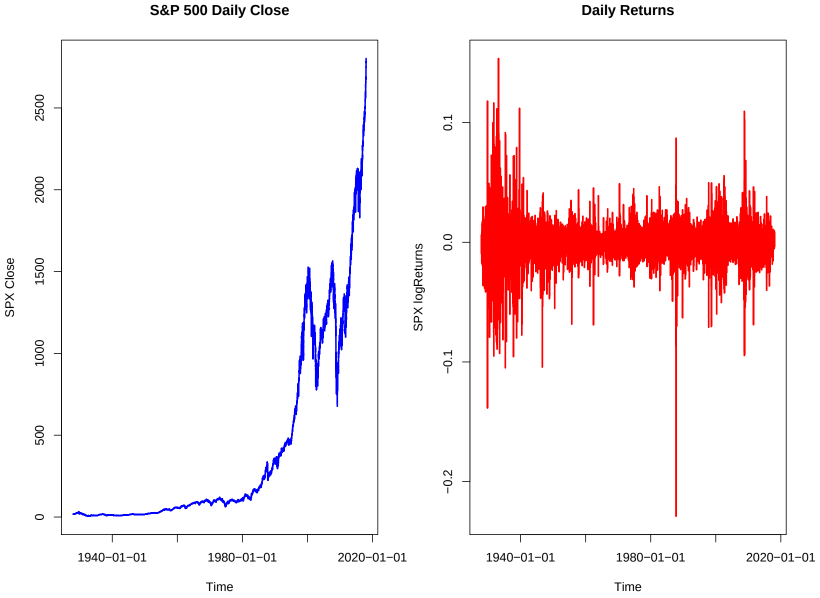
$$y_t = c + \epsilon_t, \qquad \sigma_t^2 = a_0 + \sum_{i=1}^{p} a_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} b_j \sigma_{t-j}^2$$

- For financial time series data, a $\mathrm{GARCH}(1, 1)$ model is usually adequate

Time Series Analysis

```
par(mfrow=c(1,2))
plot(spx.ts, type="l", col = 'blue', lwd =2,
      ylab = "SPX Close",
      main="S&P 500 Daily Close")

spx.ret <- getReturns(spx.ts)
plot(spx.ret, type="l", col = 'red', lwd =2,
      ylab = "SPX logReturns",
      main="Daily Returns")
par(mfrow=c(1,1))
```

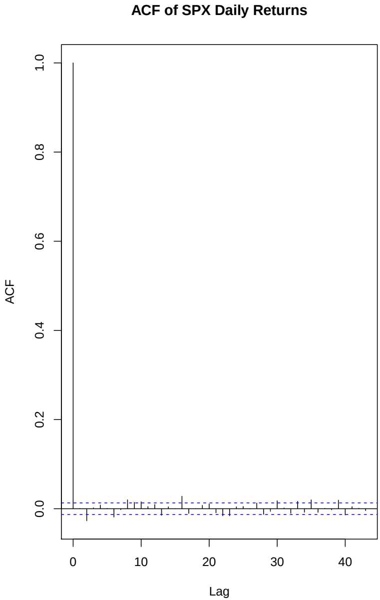
Time Series Analysis

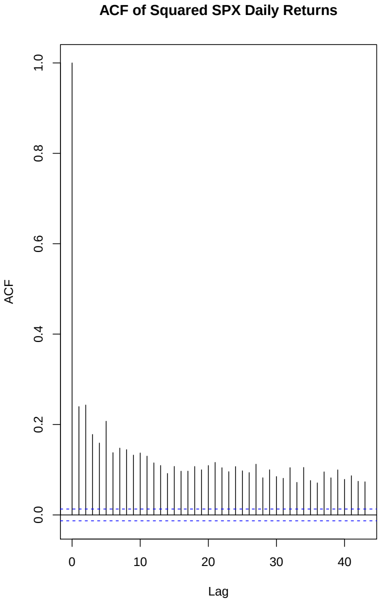


Time Series Analysis

```
par(mfrow=c(1,2))
tmp <- acf(spx.ret,
       main="ACF of SPX Daily Returns")
tmp <- acf(spx.ret^2,
       main="ACF of Squared SPX Daily Returns")
par(mfrow=c(1,1))
```

Time Series Analysis





Time Series Analysis

```
library(MTS)
archTest(spx.ret)

Q(m) of squared series(LM test):
Test statistic:  7117  p-value:  0
Rank-based Test:
Test statistic:  7254.98  p-value:  0
```

`archTest` is significant so we move onto GARCH

Time Series Analysis

```
library(tseries)
spx.gar <- garch(spx.ret)
summary(spx.gar)
```

Time Series Analysis

```
Model:
GARCH(1,1)

Residuals:
Min          1Q    Median        3Q       Max
-10.9355  -0.5108   0.0565   0.6222   6.5912

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)
a0  8.50e-07   5.37e-08     15.8   <2e-16 ***
a1  8.80e-02   1.53e-03     57.4   <2e-16 ***
b1  9.09e-01   1.79e-03    508.4   <2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1
```

Time Series Analysis

Based on the above results our model for the conditional variance is

$$\sigma_t^2 = 0.00000085 + 0.088\epsilon_{t-1}^2 + 0.909\sigma_{t-1}^2$$

The variance at time t relies on the squred errors from the prior period plus the variance from the prior period

References

Brockwell, P. J. and Davis, R. A. (1991).
Time Series: Theory and Methods.
Springer-Verlag, New York, second edition.

Brockwell, P. J. and Davis, R. A. (1996).
Introduction to Time Series and Forecasting.
Springer-Verlag, New York.

Carmona, R. A. (2004).
Statistical Analysis of Financial Data in S-PLUS.
Springer-Verlag.

Wu, J. and Coggeshall, S. (2012).
Foundations of Predictive Analytics.
CRC Press, Boca Raton, FL.

Zivot, E. and Wang, J. (2006).
Modeling Financial Time Series with S-PLUS.
Springer-Verlag, New York, second edition.