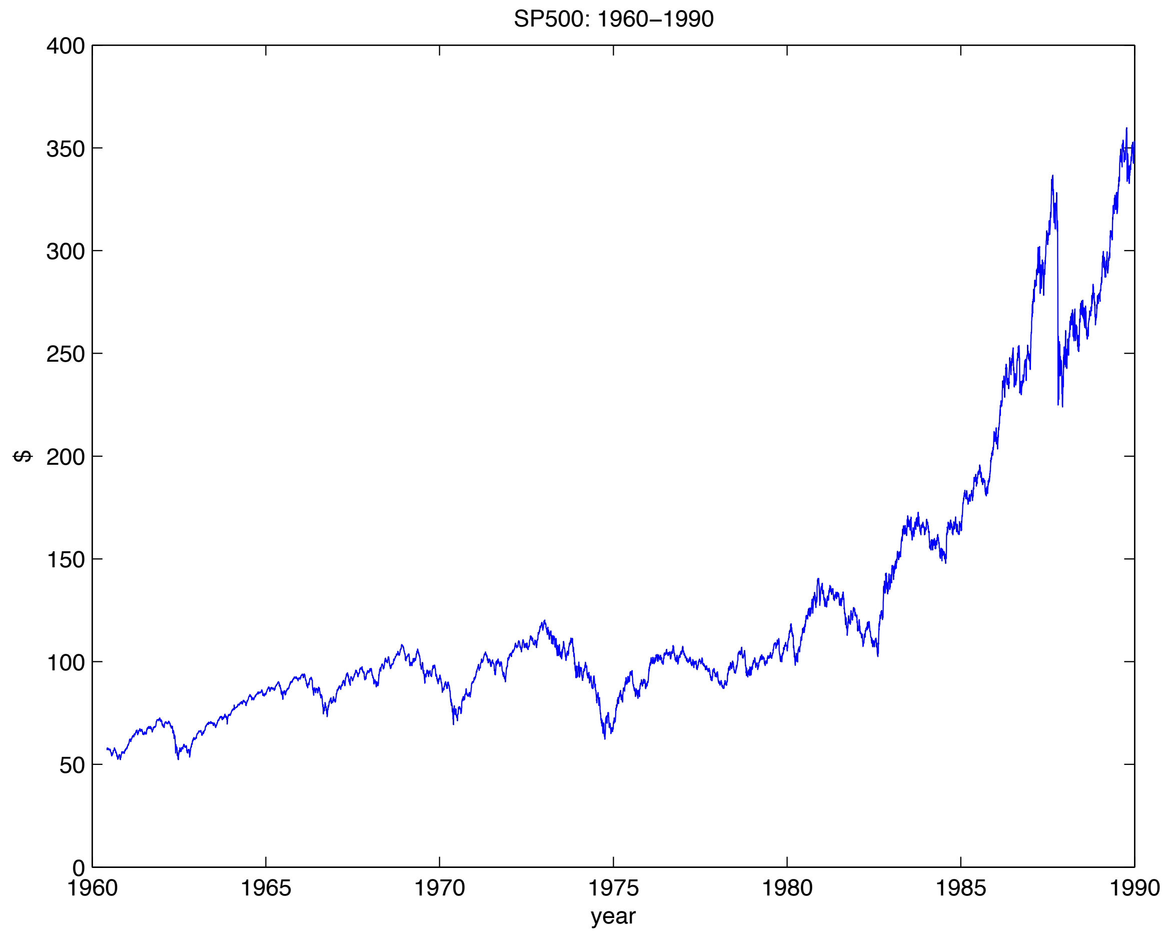
Introduction to Time Series Analysis. Lecture 1.

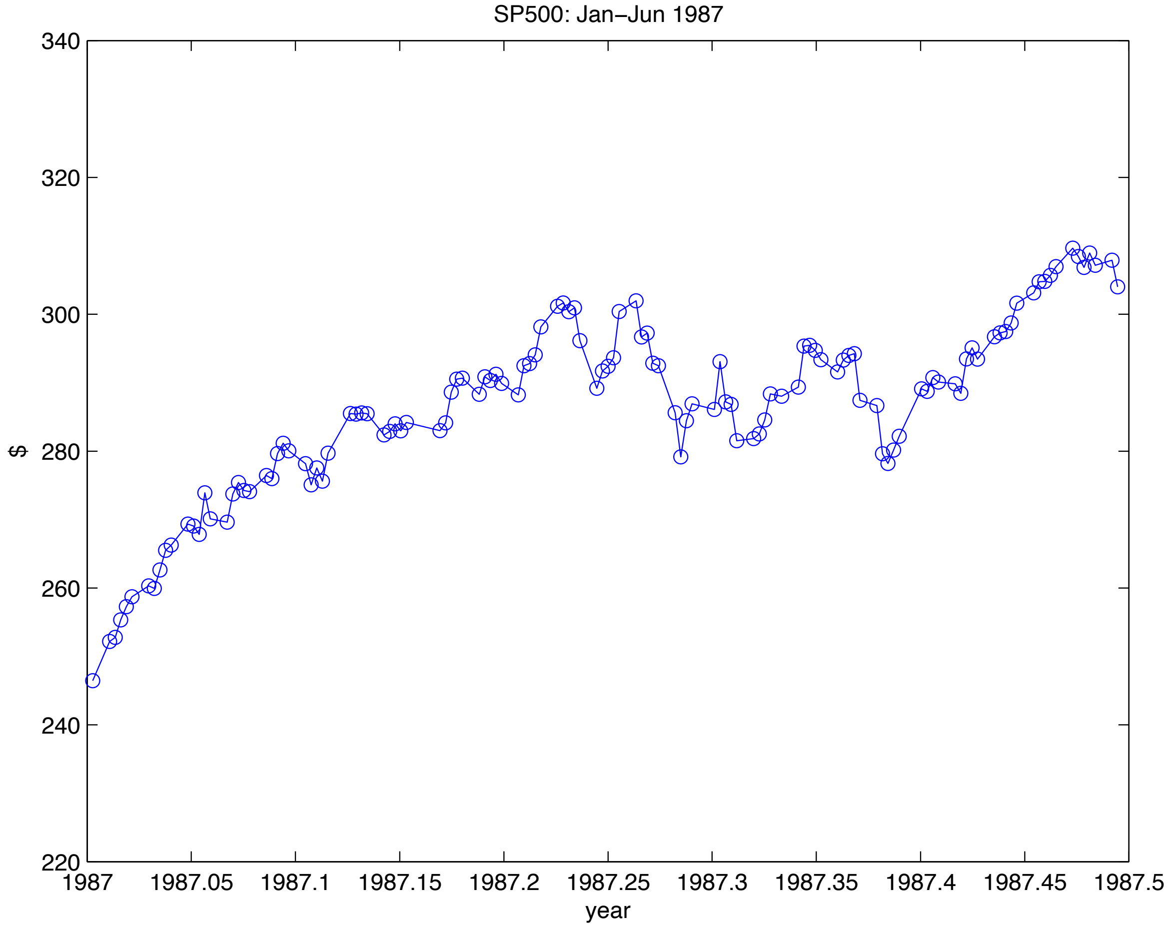
Peter Bartlett

1. Organizational issues.
2. Objectives of time series analysis. Examples.
3. Overview of the course.
4. Time series models.
5. Time series modelling: Chasing stationarity.

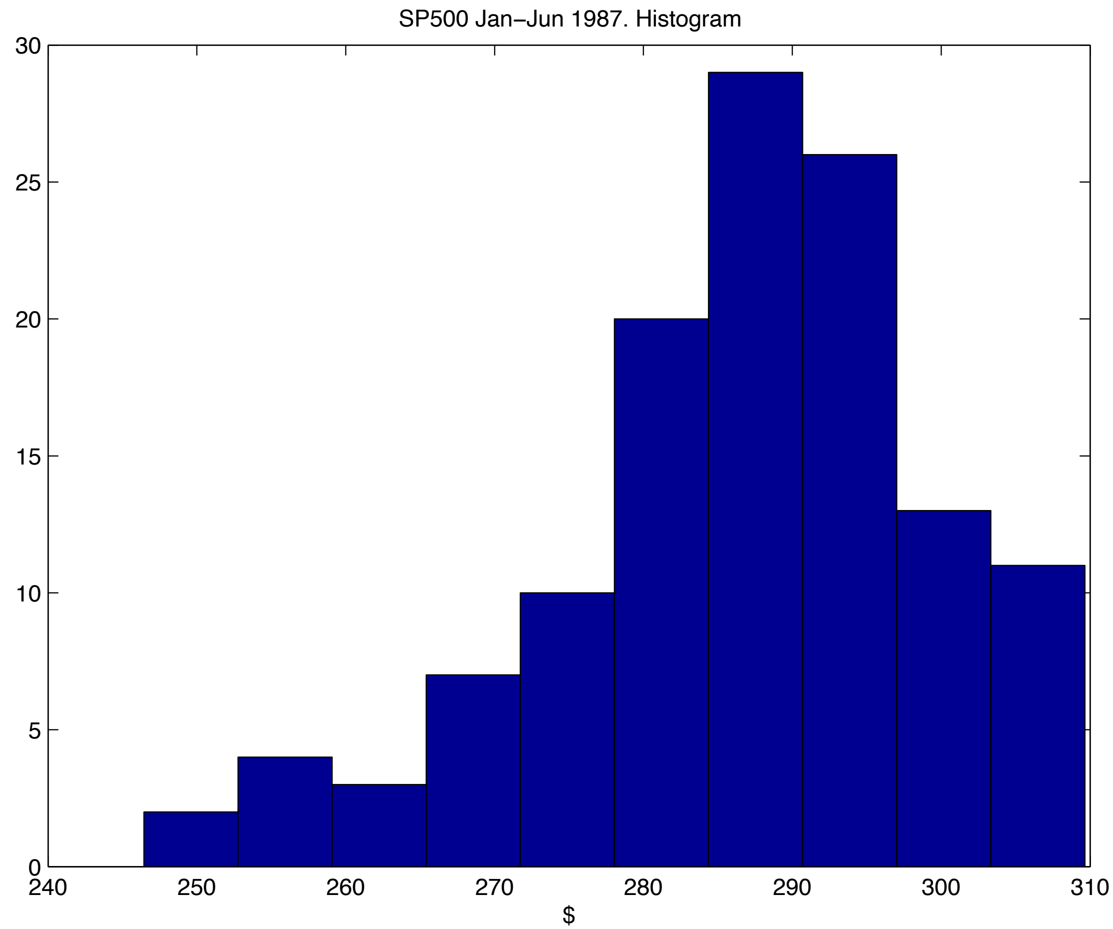
A Time Series



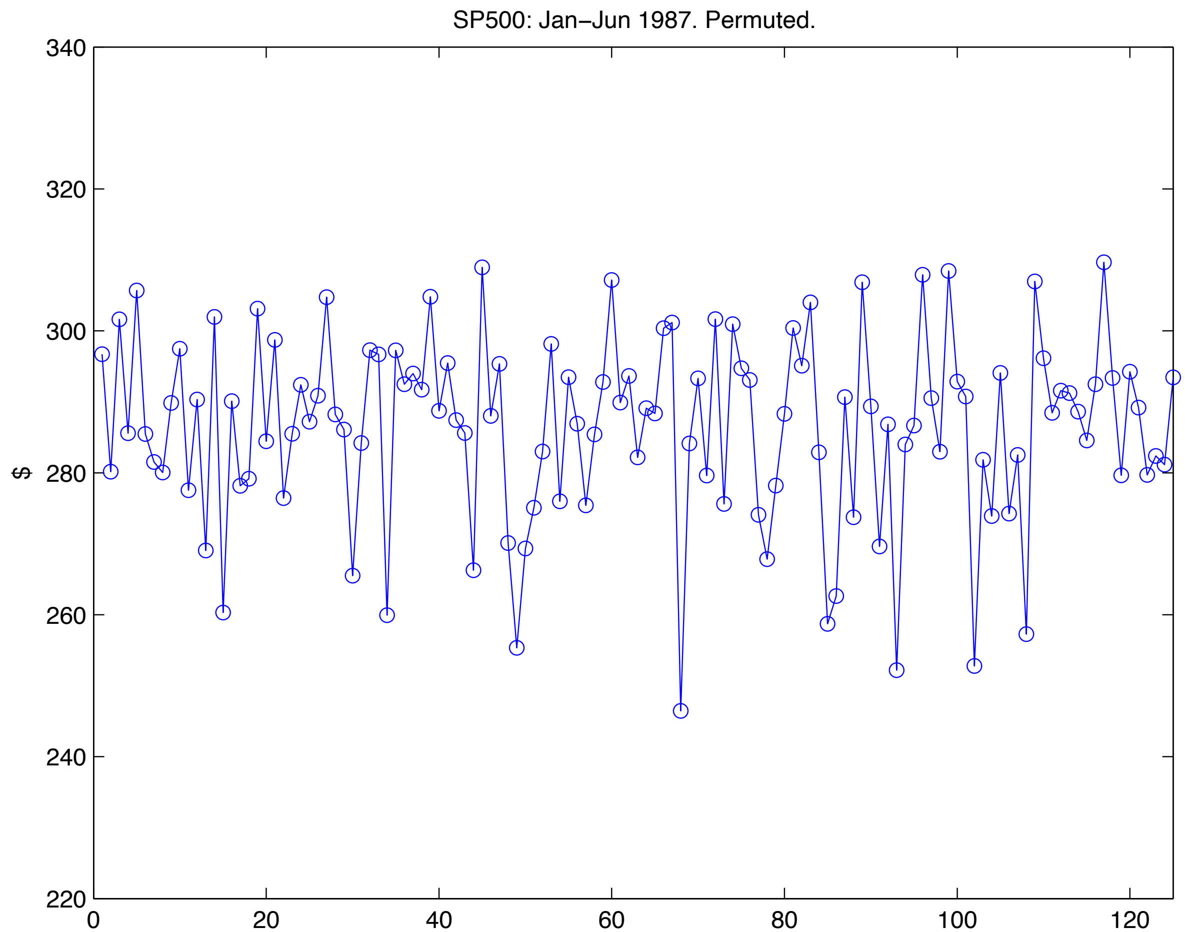
A Time Series



A Time Series



A Time Series



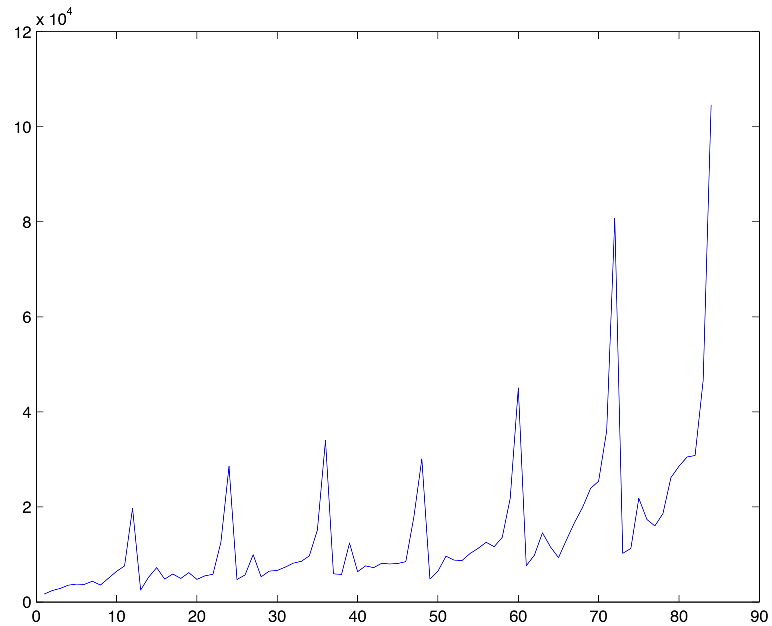
Objectives of Time Series Analysis

1. Compact description of data.
2. Interpretation.
3. Forecasting.
4. Control.
5. Hypothesis testing.
6. Simulation.

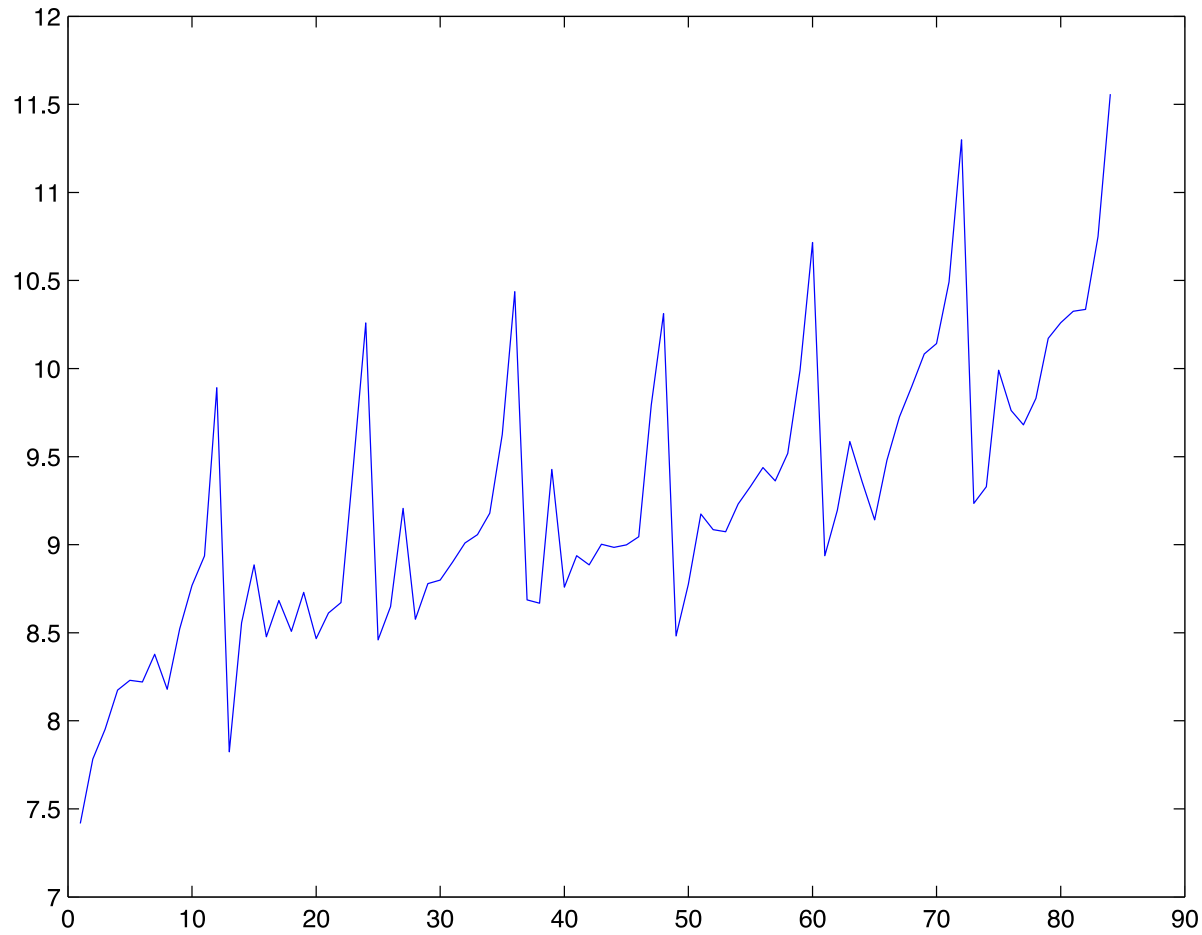
Classical decomposition: An example

Monthly sales for a souvenir shop at a beach resort town in Queensland.

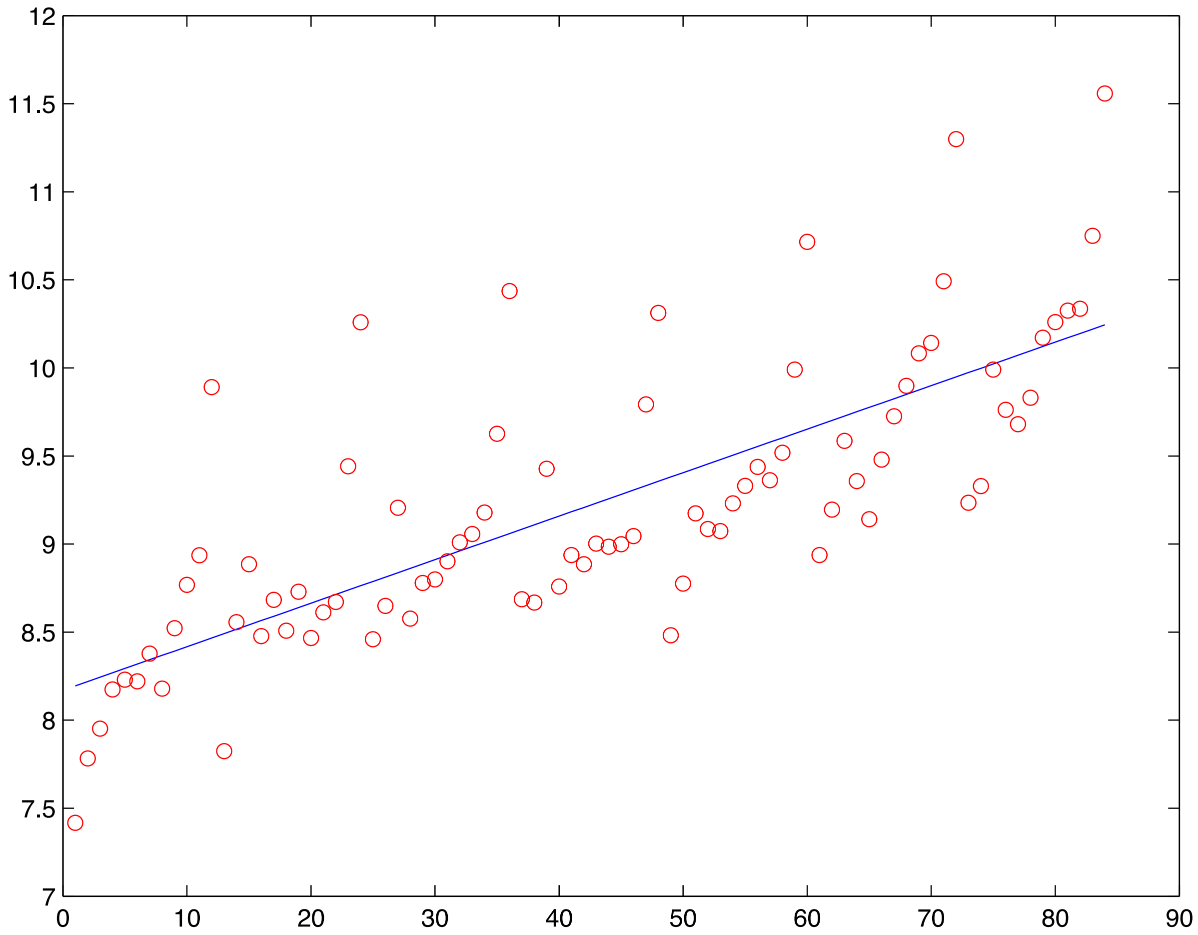
(Makridakis, Wheelwright and Hyndman, 1998)



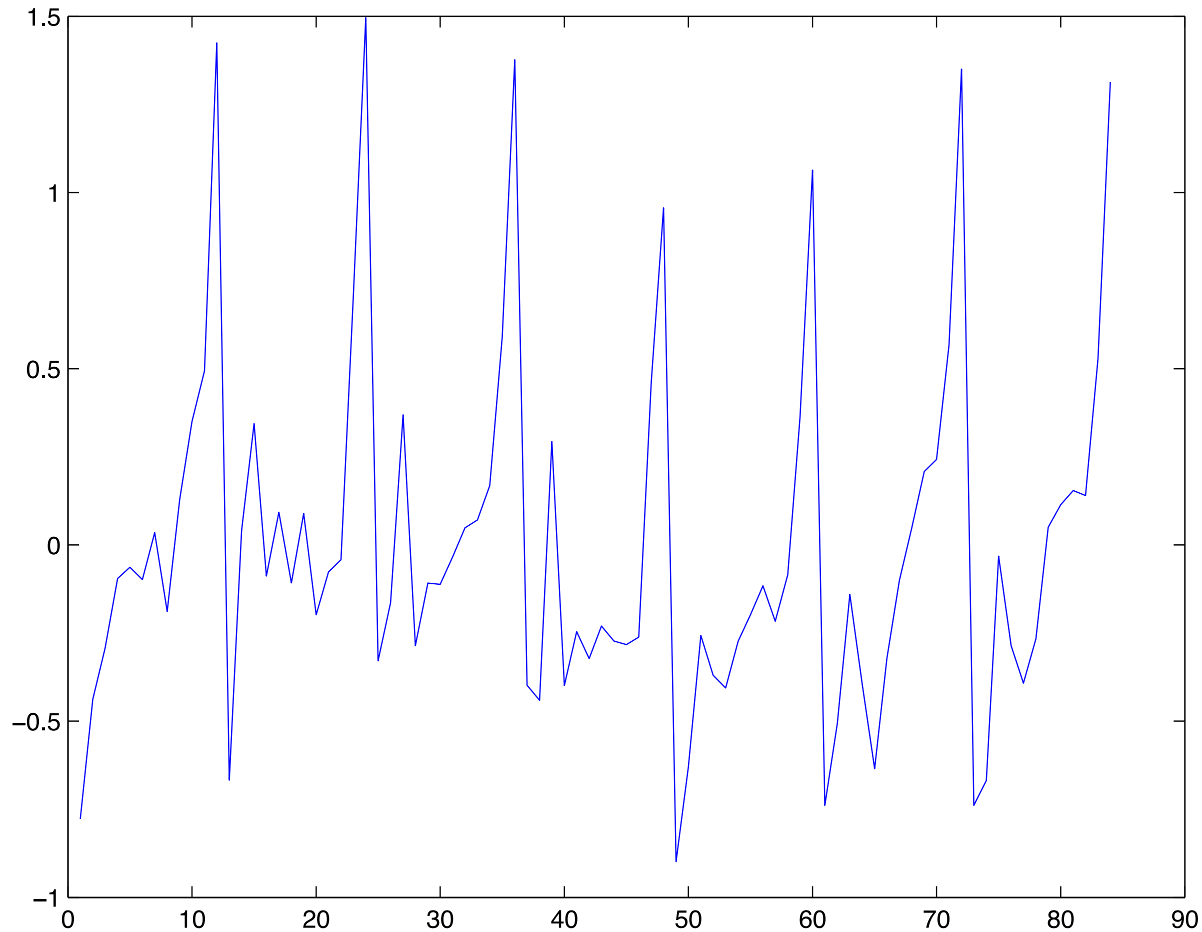
Transformed data



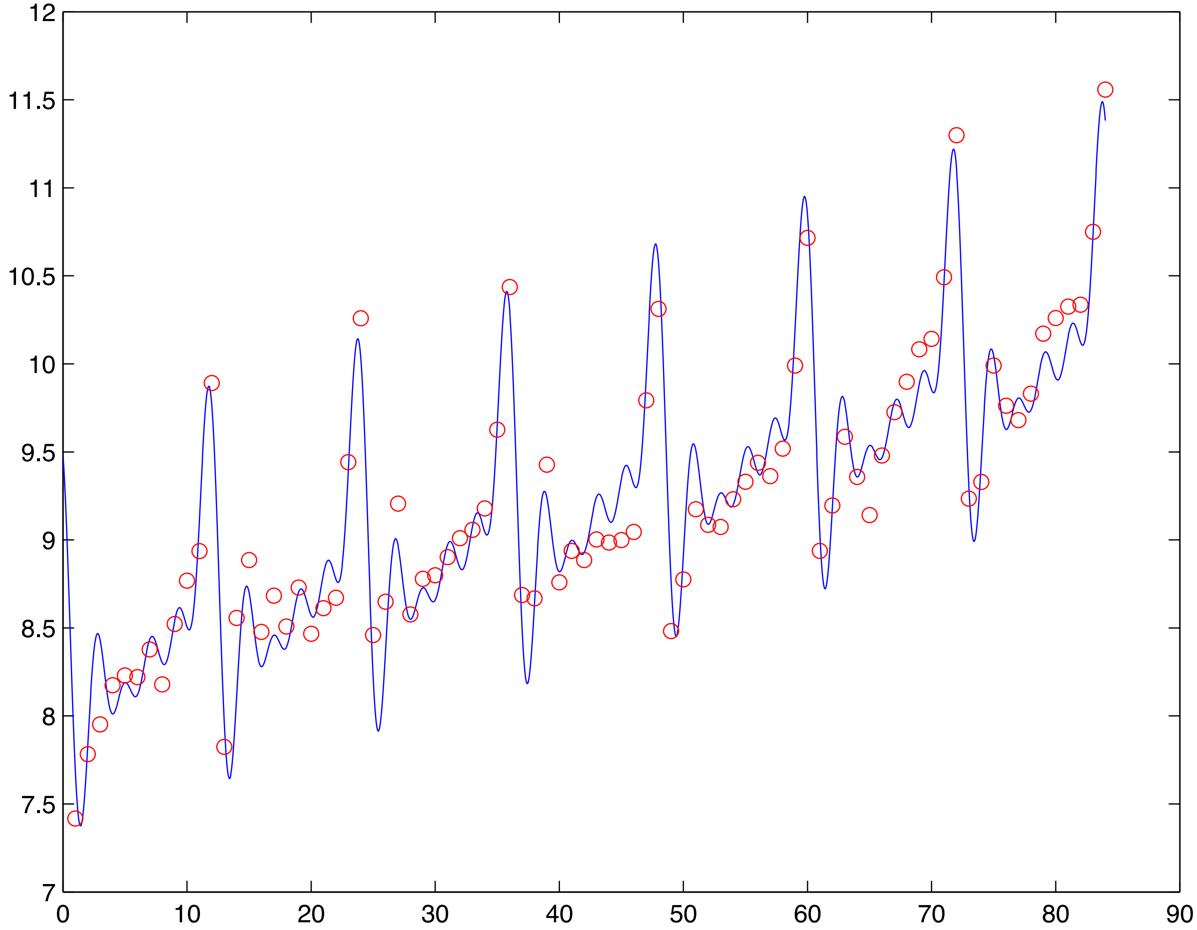
Trend



Residuals



Trend and seasonal variation

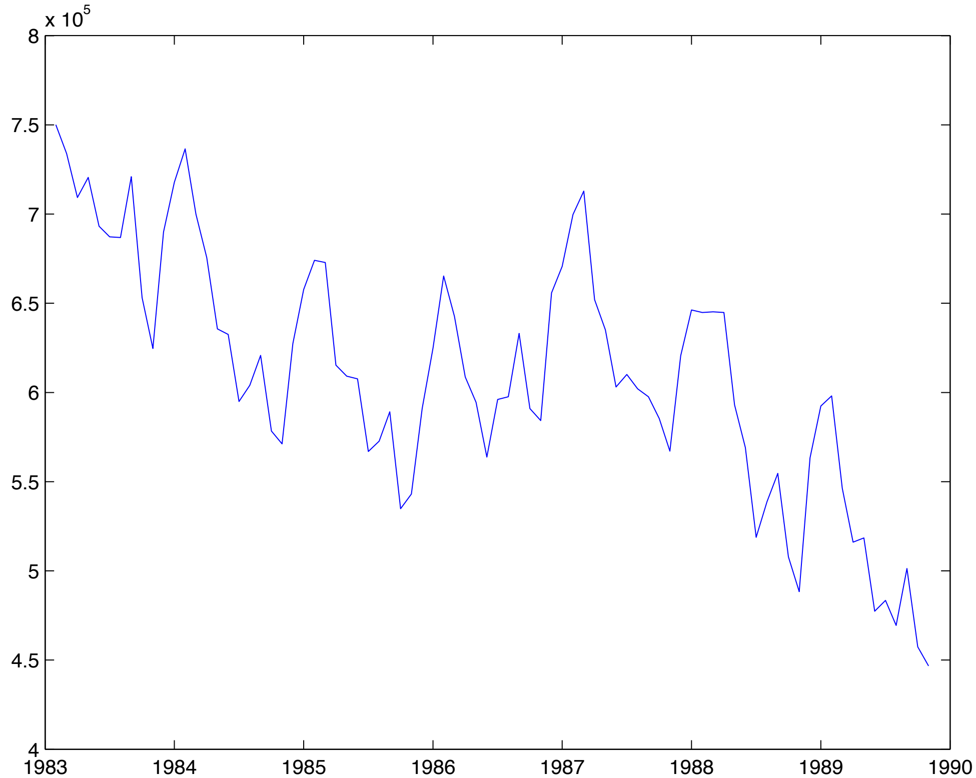


Objectives of Time Series Analysis

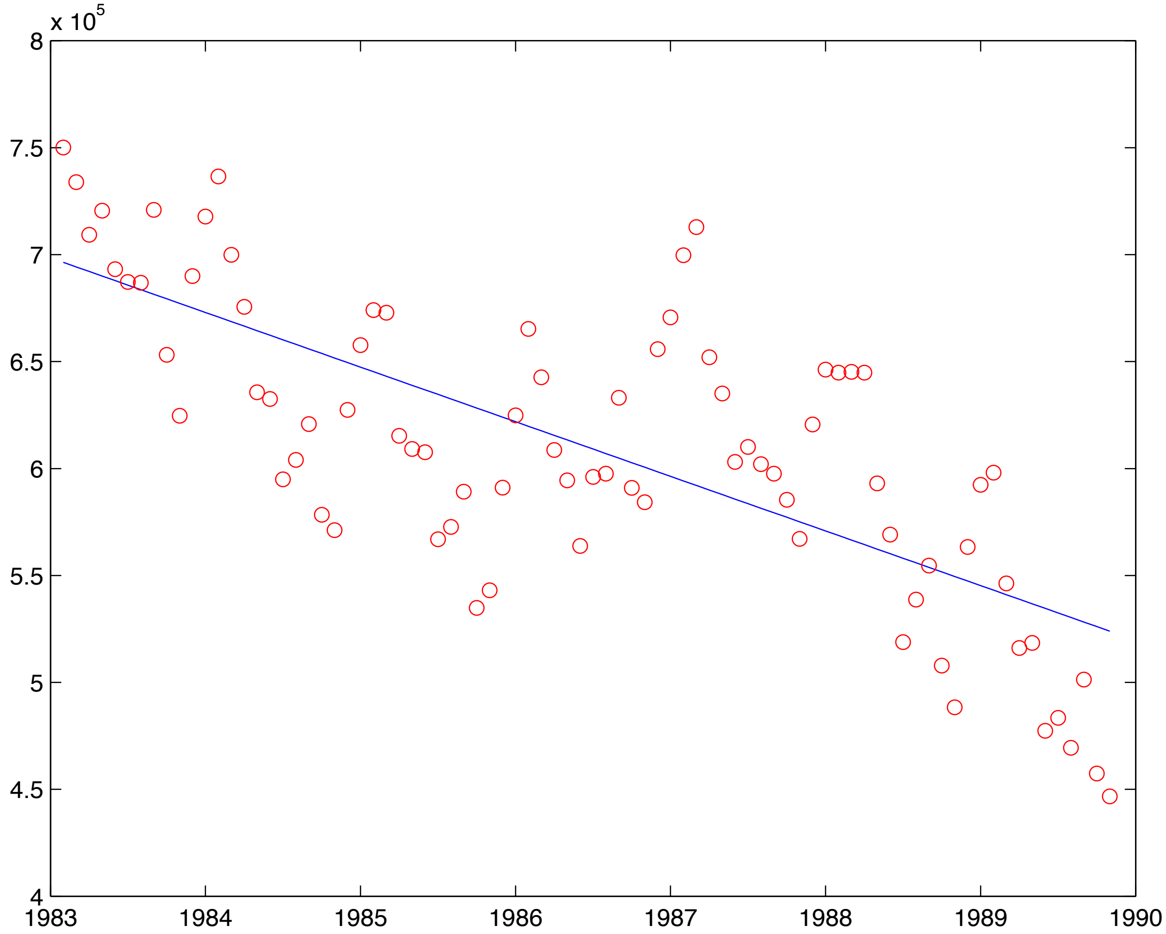
1. Compact description of data.
   Example: Classical decomposition: $X_t = T_t + S_t + Y_t$.
2. Interpretation. Example: Seasonal adjustment.
3. Forecasting. Example: Predict sales.
4. Control.
5. Hypothesis testing.
6. Simulation.

Unemployment data

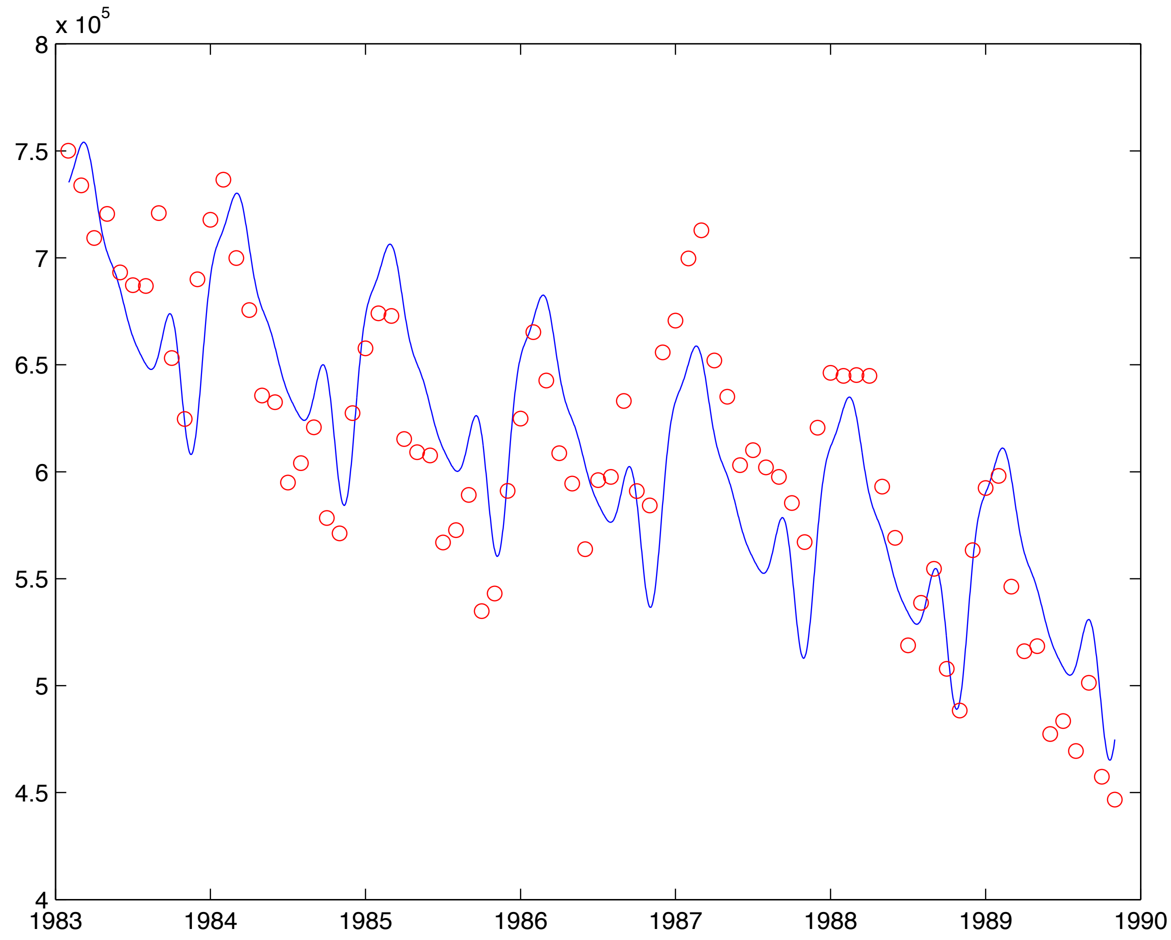
Monthly number of unemployed people in Australia. (Hipel and McLeod, 1994)



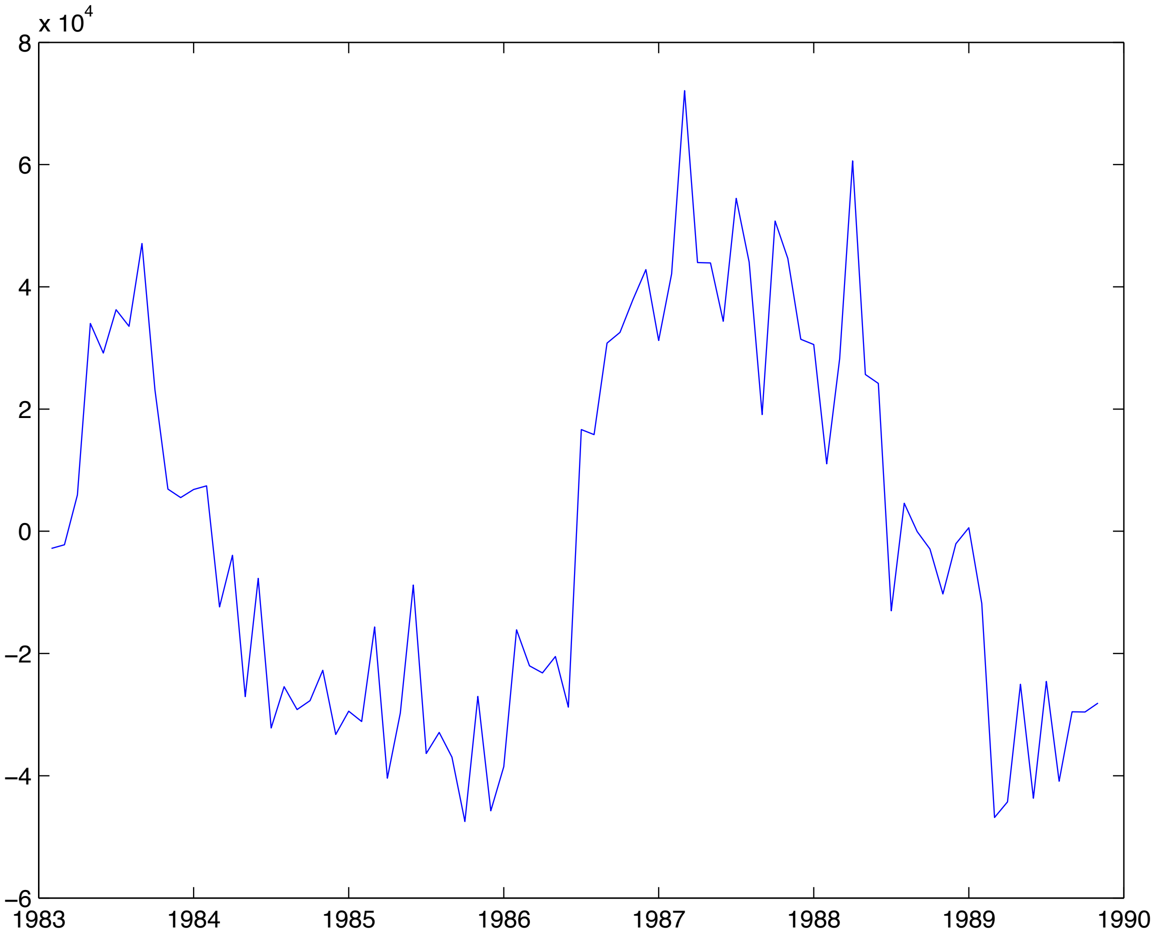
Trend



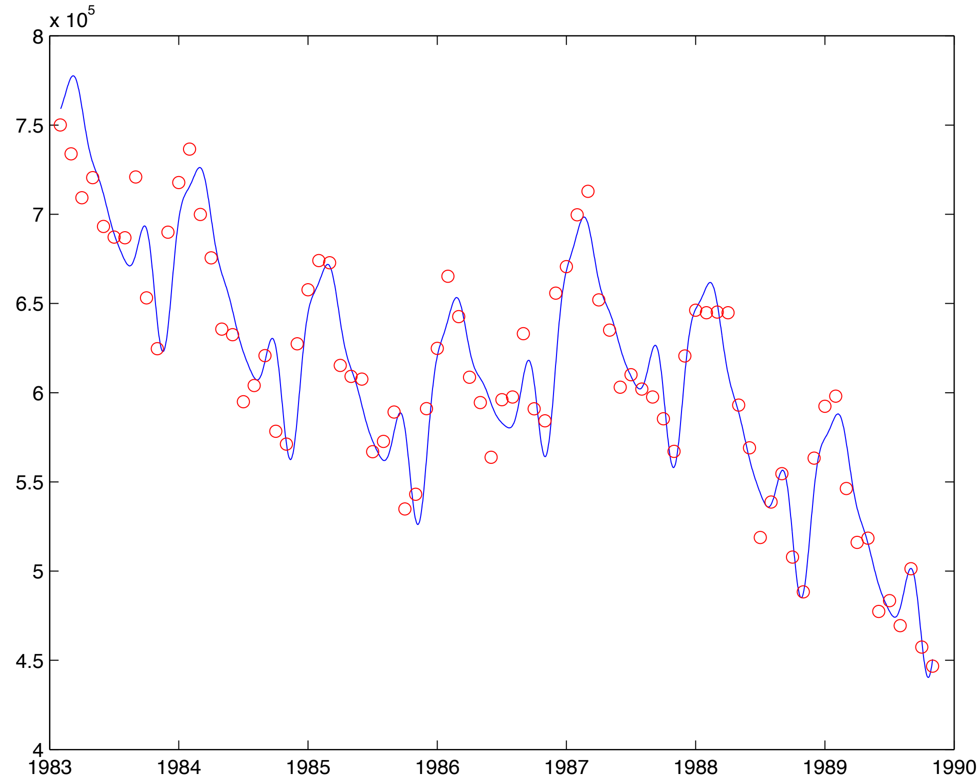
Trend plus seasonal variation



Residuals



Predictions based on a (simulated) variable



Objectives of Time Series Analysis

1. Compact description of data:

$$X_t = T_t + S_t + f(Y_t) + W_t.$$

2. Interpretation. Example: Seasonal adjustment.
3. Forecasting. Example: Predict unemployment.
4. Control. Example: Impact of monetary policy on unemployment.
5. Hypothesis testing. Example: Global warming.
6. Simulation. Example: Estimate probability of catastrophic events.

Overview of the Course

1. Time series models
   (a) Stationarity.
   (b) Autocorrelation function.
   (c) Transforming to stationarity.
2. Time domain methods
3. Spectral analysis
4. State space models(?)

Overview of the Course

1. Time series models
2. Time domain methods
   (a) AR/MA/ARMA models.
   (b) ACF and partial autocorrelation function.
   (c) Forecasting
   (d) Parameter estimation
   (e) ARIMA models/seasonal ARIMA models
3. Spectral analysis
4. State space models(?)

Overview of the Course

1. Time series models
2. Time domain methods
3. Spectral analysis
   (a) Spectral density
   (b) Periodogram
   (c) Spectral estimation
4. State space models(?)

Overview of the Course

1. Time series models
2. Time domain methods
3. Spectral analysis
4. State space models(?)
   (a) ARMAX models.
   (b) Forecasting, Kalman filter.
   (c) Parameter estimation.

Time Series Models

A **time series model** specifies the joint distribution of the sequence $\{X_t\}$ of random variables.
For example:

$$P[X_1 \leq x_1, \ldots, X_t \leq x_t] \text{ for all } t \text{ and } x_1, \ldots, x_t.$$

Notation:
$X_1, X_2, \ldots$ is a stochastic process.
$x_1, x_2, \ldots$ is a single realization.

We'll mostly restrict our attention to **second-order properties** only:
$\mathrm{E}X_t, \mathrm{E}(X_{t_1} X_{t_2})$.

Time Series Models

Example: White noise: $X_t \sim WN(0, \sigma^2)$.
i.e., $\{X_t\}$ uncorrelated, $\mathrm{E}X_t = 0$, $\mathrm{Var}X_t = \sigma^2$.

Example: i.i.d. noise: $\{X_t\}$ independent and identically distributed.

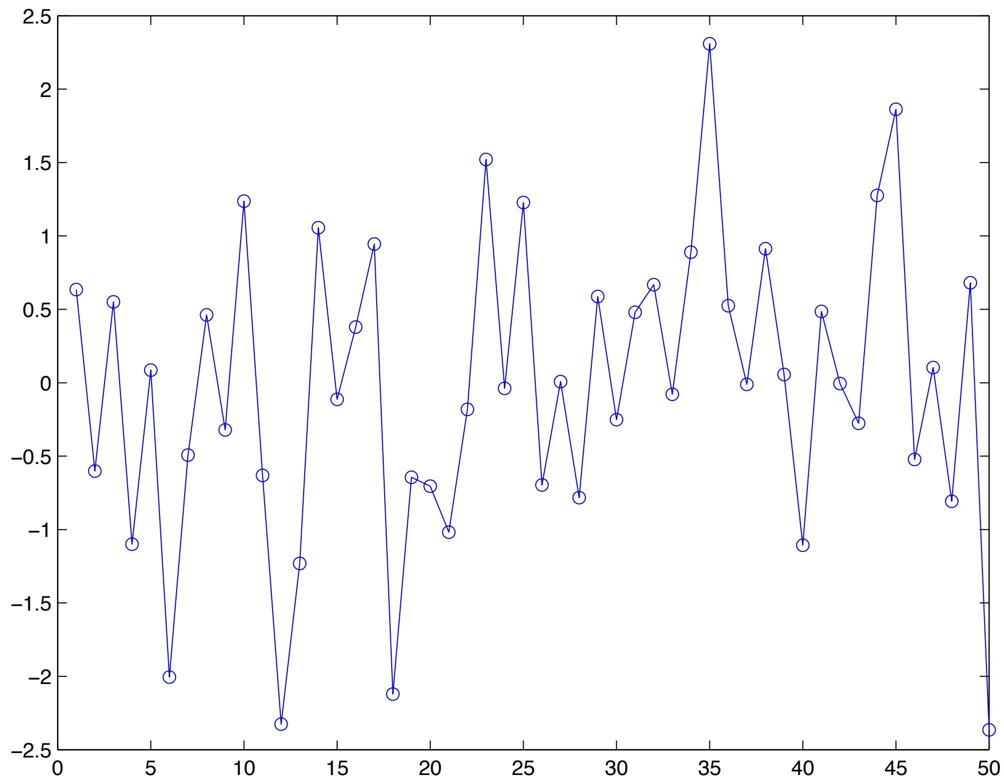
$$P[X_1 \leq x_1, \ldots, X_t \leq x_t] = P[X_1 \leq x_1] \cdots P[X_t \leq x_t].$$

Not interesting for forecasting:

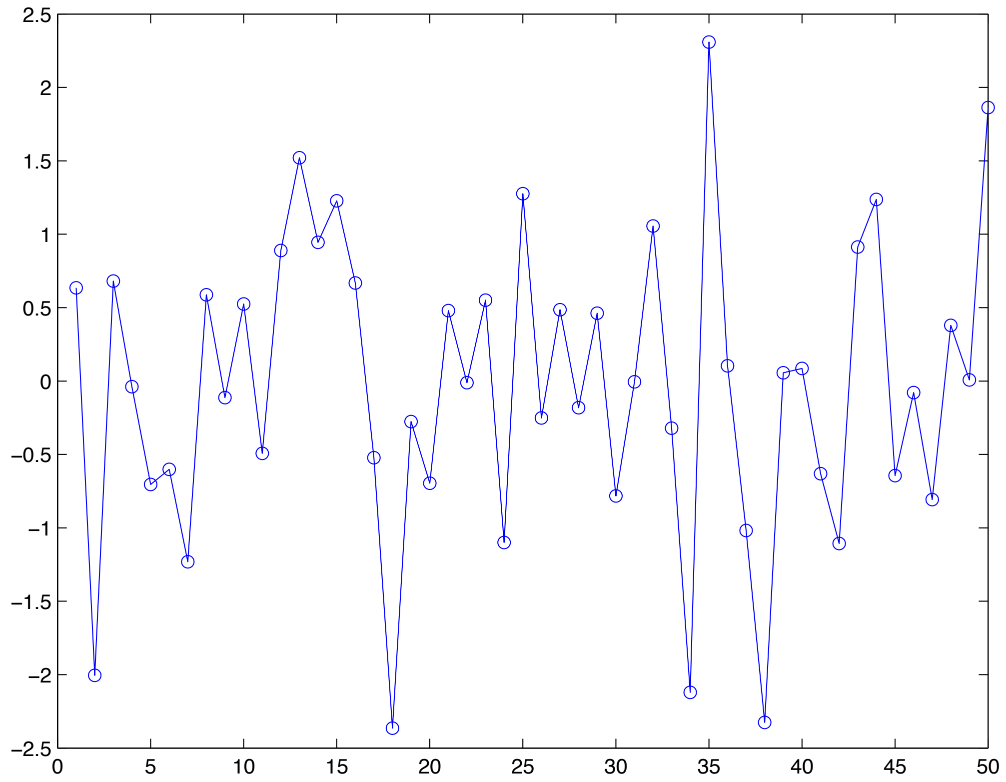
$$P[X_t \leq x_t | X_1, \ldots, X_{t-1}] = P[X_t \leq x_t].$$

Gaussian white noise

$$P[X_t \leq x_t] = \Phi(x_t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_t} e^{-x^2/2} dx.$$

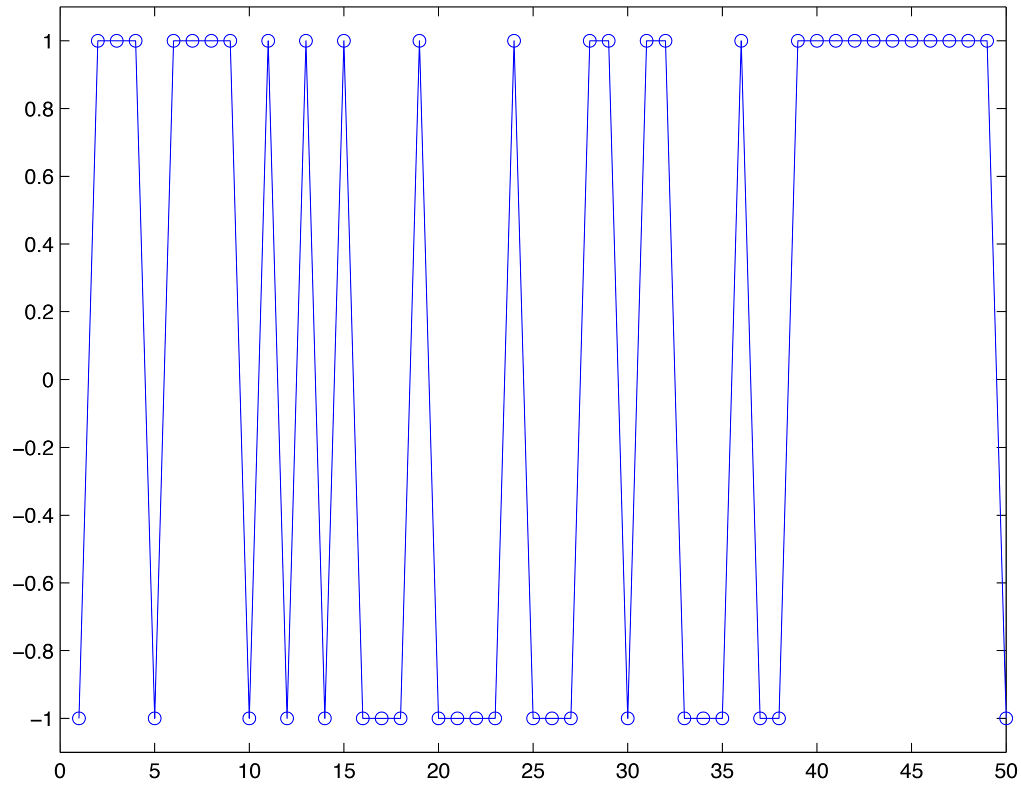


Gaussian white noise



Time Series Models

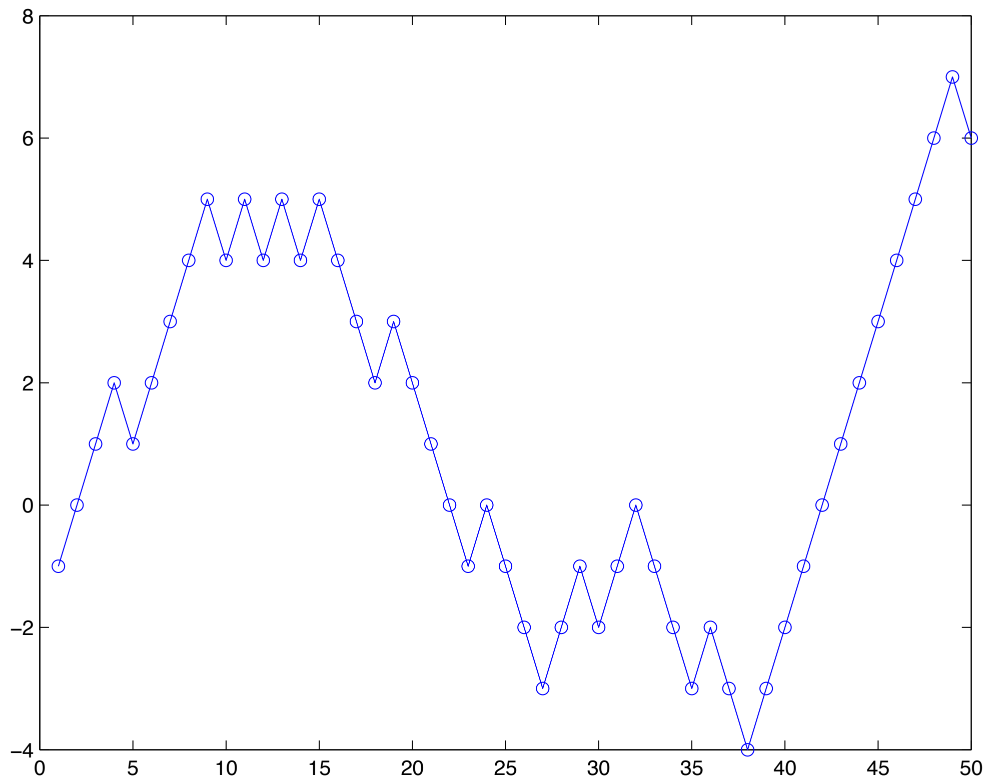
Example: Binary i.i.d. $P[X_t = 1] = P[X_t = -1] = 1/2.$



Random walk

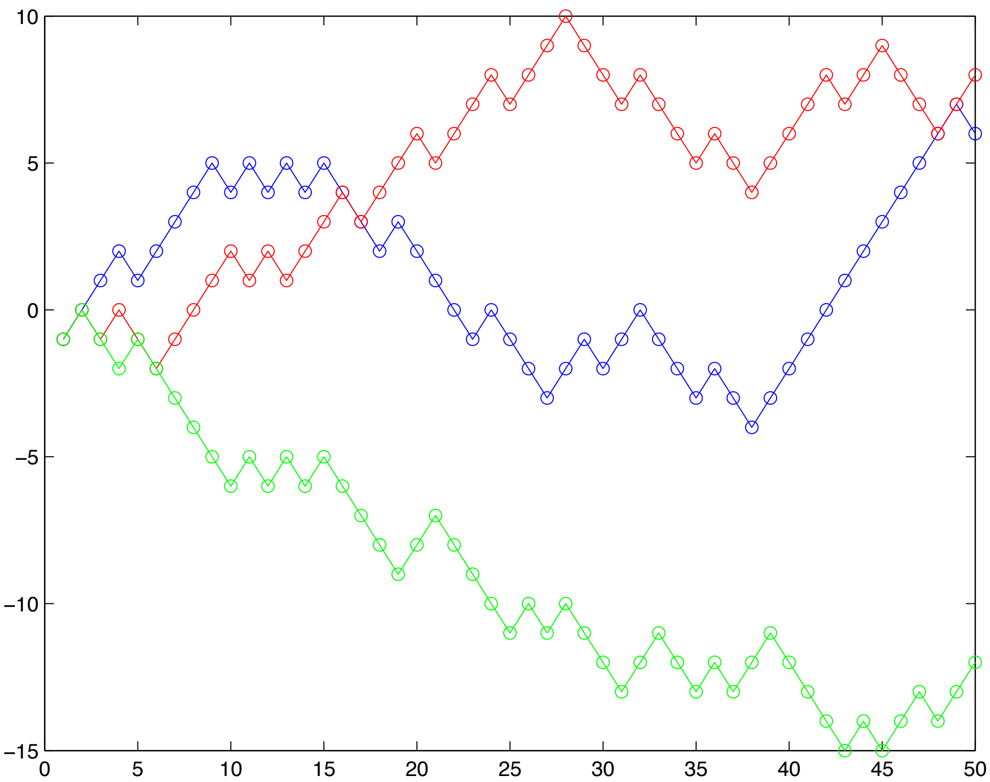
$S_t = \sum_{i=1}^{t} X_i$.

Differences: $\nabla S_t = S_t - S_{t-1} = X_t$.



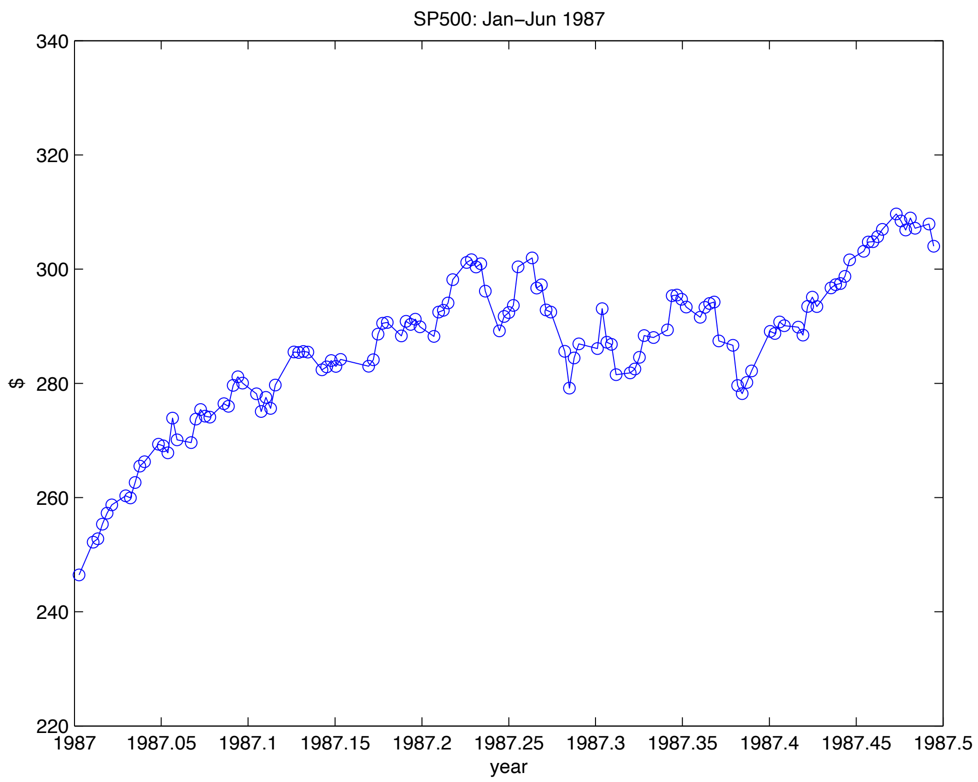
Random walk

$\mathrm{E}S_t$? $\mathrm{Var}S_t$?



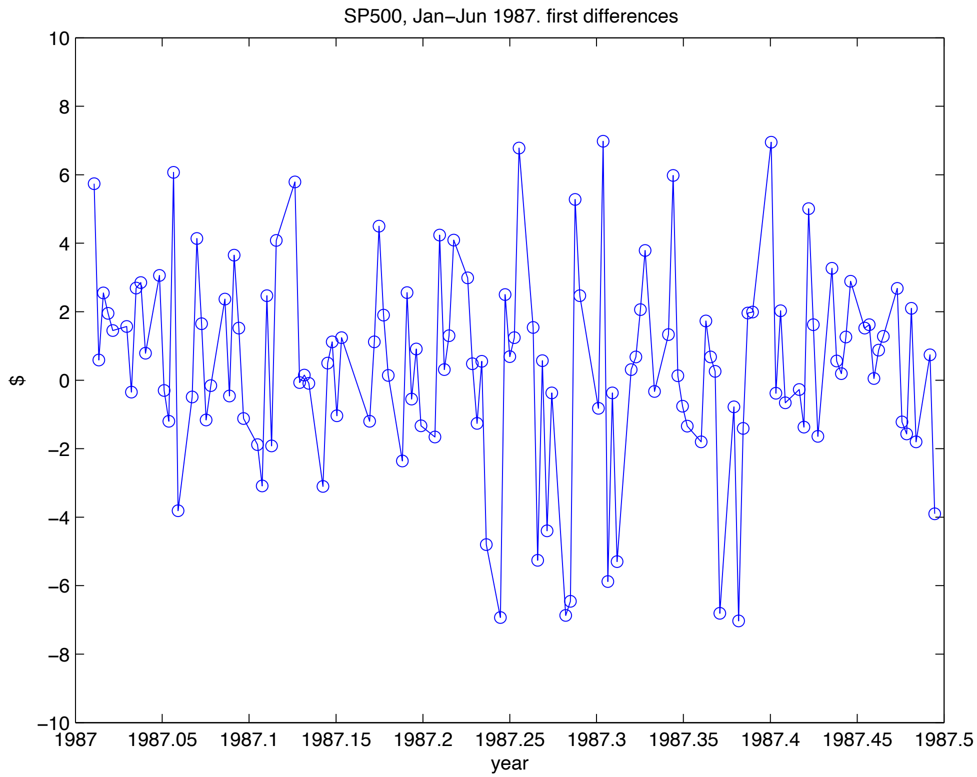
Random Walk

Recall S&P500 data. (Notice that it's smooth)



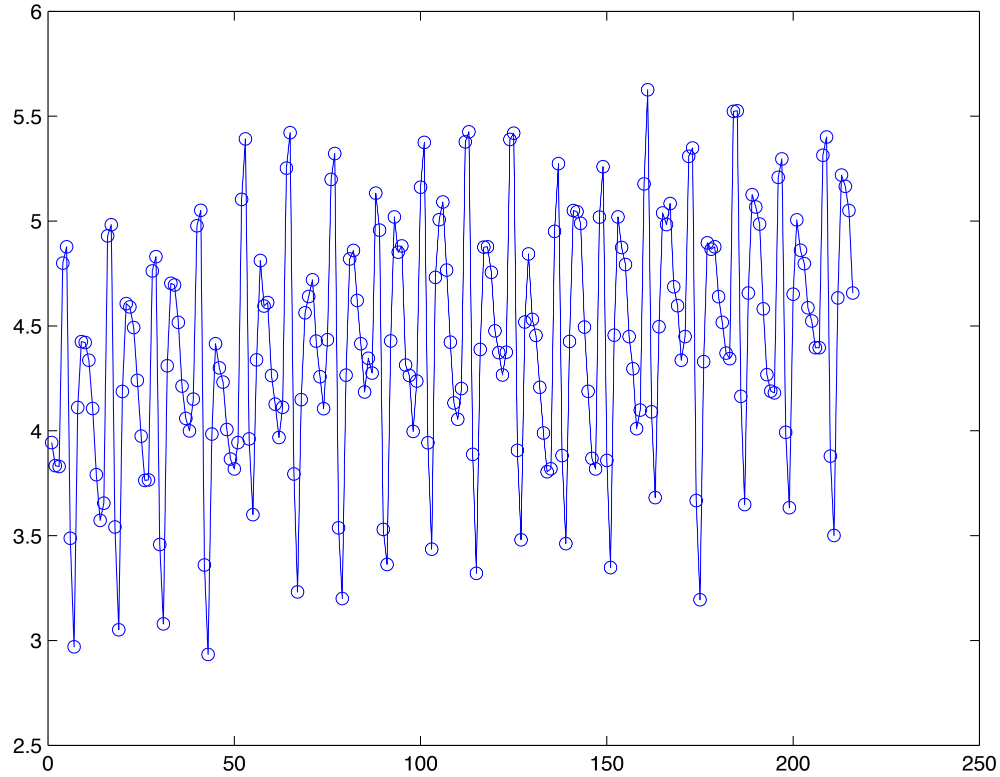
Random Walk

Differences: $$\nabla S_t = S_t - S_{t-1} = X_t.$$



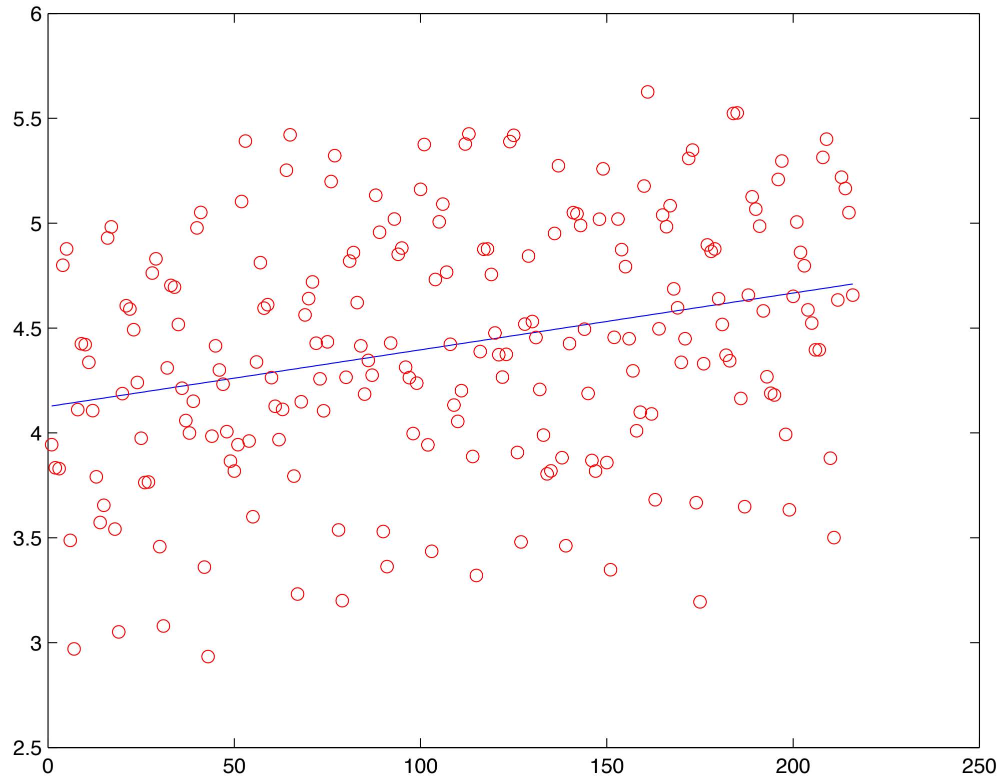
Trend and Seasonal Models

$$X_t = T_t + S_t + E_t = \beta_0 + \beta_1 t + \sum_i \left(\beta_i \cos(\lambda_i t) + \gamma_i \sin(\lambda_i t)\right) + E_t$$



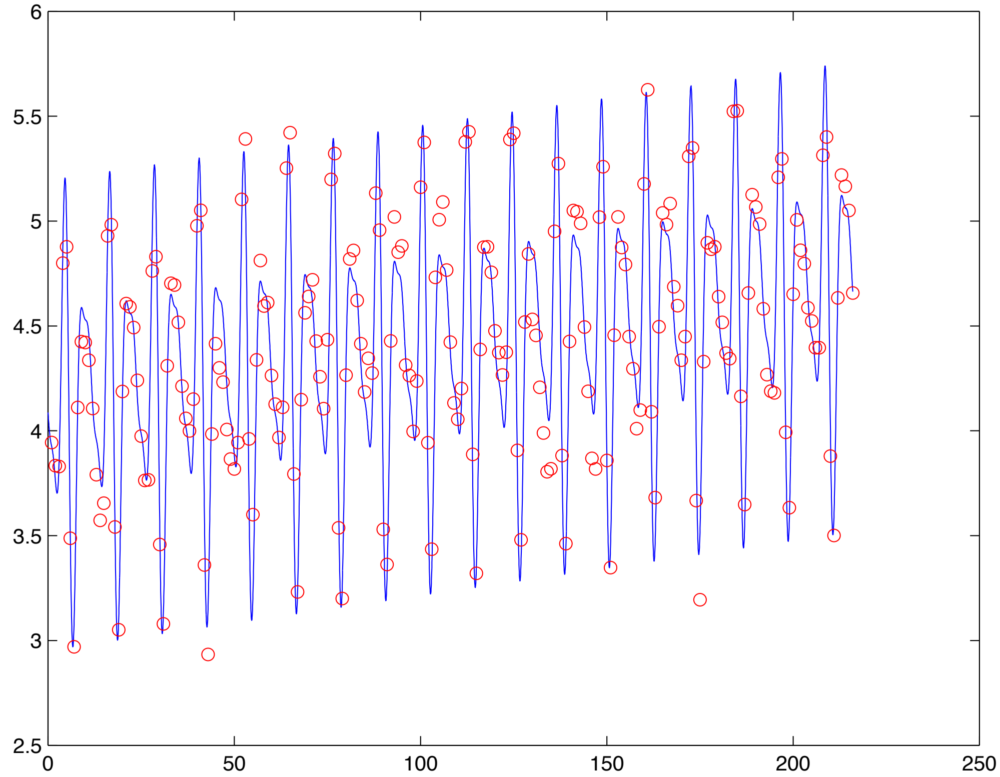
Trend and Seasonal Models

$$X_t = T_t + E_t = \beta_0 + \beta_1 t + E_t$$

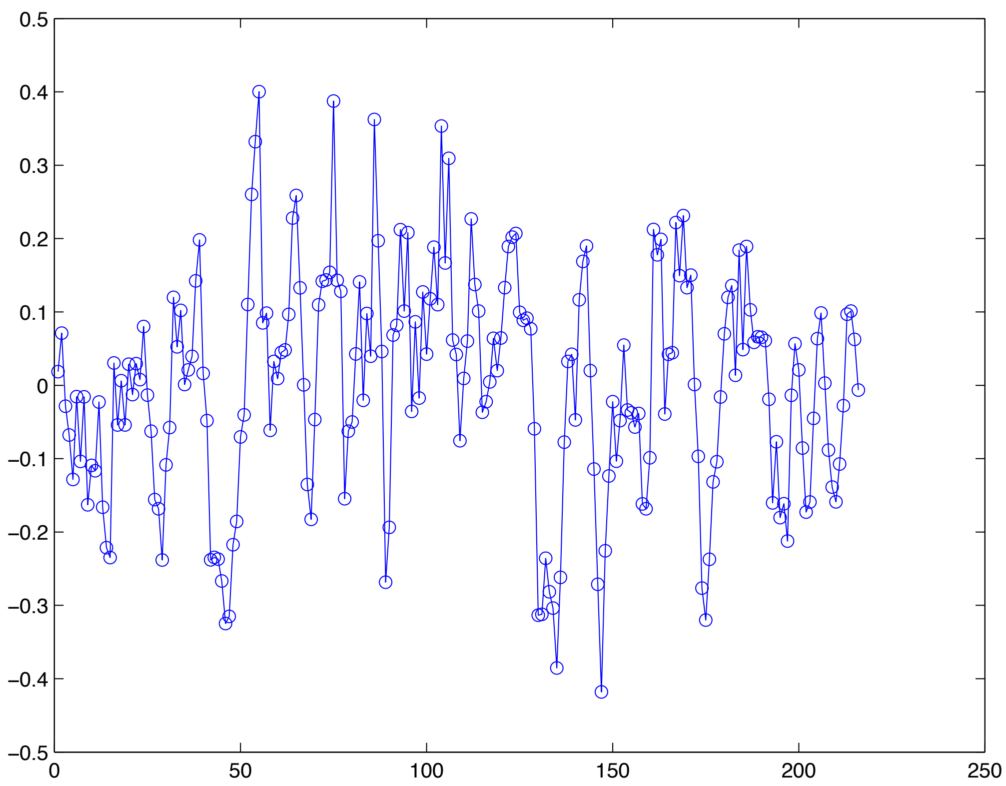


Trend and Seasonal Models

$$X_t = T_t + S_t + E_t = \beta_0 + \beta_1 t + \sum_i \left(\beta_i \cos(\lambda_i t) + \gamma_i \sin(\lambda_i t)\right) + E_t$$



Trend and Seasonal Models: Residuals



Time Series Modelling

1. Plot the time series.

   Look for trends, seasonal components, step changes, outliers.

2. Transform data so that residuals are **stationary**.

   (a) Estimate and subtract T_t, S_t.

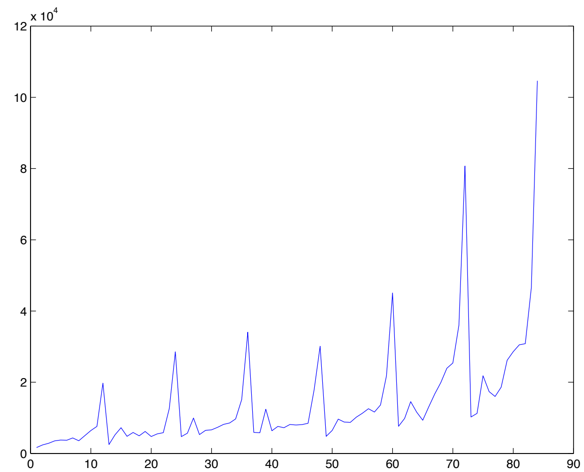
   (b) Differencing.

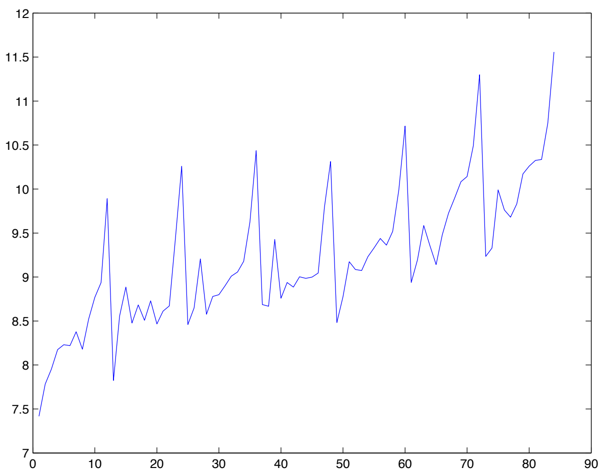
   (c) Nonlinear transformations (log, $\sqrt{\cdot}$).

3. Fit model to residuals.

Nonlinear transformations

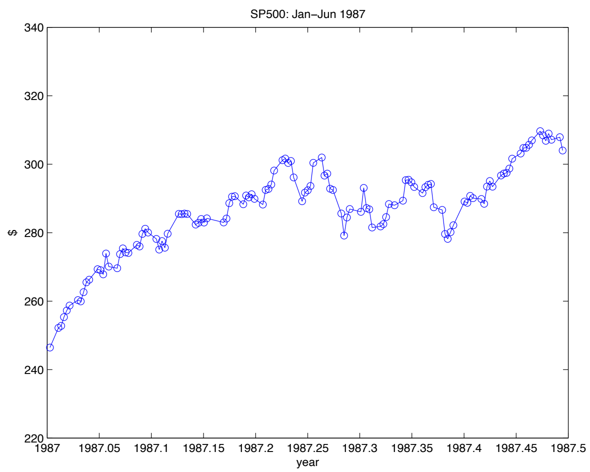
Recall: Monthly sales. (Makridakis, Wheelwright and Hyndman, 1998)

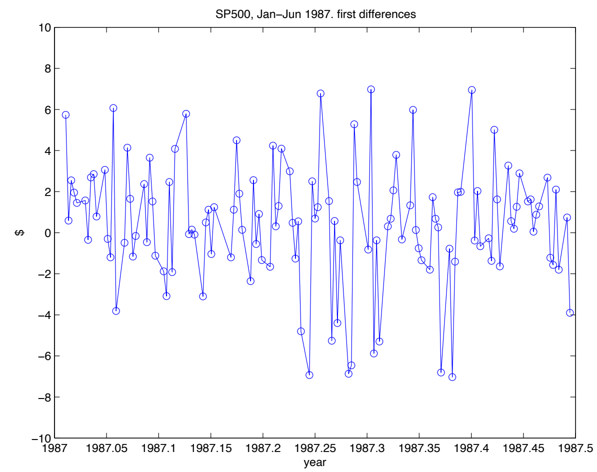




Differencing

Recall: S&P 500 data.





Differencing and Trend

Define the lag-1 **difference operator**, (think 'fi rst derivative')

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

where B is the **backshift** operator, $BX_t = X_{t-1}$.

- If $X_t = \beta_0 + \beta_1 t + Y_t$, then

$$\nabla X_t = \beta_1 + \nabla Y_t.$$

- If $X_t = \sum_{i=0}^{k} \beta_i t^i + Y_t$, then

$$\nabla^k X_t = k!\beta_k + \nabla^k Y_t,$$

  where $\nabla^k X_t = \nabla(\nabla^{k-1} X_t)$ and $\nabla^1 X_t = \nabla X_t$.

Differencing and Seasonal Variation

Define the lag-s **difference operator**,

$$\nabla_s X_t = X_t - X_{t-s} = (1 - B^s) X_t,$$

where B^s is the backshift operator applied s times, $B^s X_t = B(B^{s-1} X_t)$ and $B^1 X_t = B X_t$.

If $X_t = T_t + S_t + Y_t$, and S_t has period s (that is, $S_t = S_{t-s}$ for all t), then

$$\nabla_s X_t = T_t - T_{t-s} + \nabla_s Y_t.$$

Least Squares Regression

Model:

$$
\begin{aligned}
X_t &= \beta_0 + \beta_1 t + W_t \\
&= \begin{pmatrix} 1 & t \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + W_t,
\end{aligned}
$$

$$
\underbrace{\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_T \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & T \end{pmatrix}}_{Z} \underbrace{\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}}_{\beta} + \underbrace{\begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_T \end{pmatrix}}_{w}
$$

Least Squares Regression

$$x = Z\beta + w.$$

Least squares: choose β to minimize $\|w\|^2 = \|x - Z\beta\|^2$.

Solution $\hat{\beta}$ satisfies the *normal equations*:

$$\nabla_\beta \|w\|^2 = 2Z'(x - Z\hat{\beta}) = 0.$$

If $Z'Z$ is nonsingular, the solution is unique:

$$\hat{\beta} = (Z'Z)^{-1}Z'x.$$

Least Squares Regression

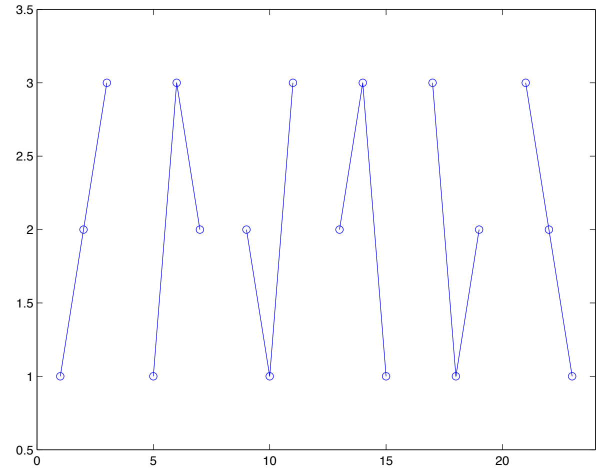
Properties of the least squares solution ($\hat{\beta} = (Z'Z)^{-1}Z'x$):

- Linear.
- Unbiased.
- For $\{W_t\}$ i.i.d., it is the linear unbiased estimator with smallest variance.

Other regressors Z: polynomial, trigonometric functions, piecewise polynomial (splines), etc.

Testing i.i.d.: Turning point test

$\{X_t\}$ i.i.d. implies that X_t, X_{t+1} and X_{t+2} are equally likely to occur in any of six possible orders:



(provided X_t, X_{t+1}, X_{t+2} are distinct).

Four of the six are **turning points**.

Testing i.i.d.: Turning point test

Define $T = |\{t : X_t, X_{t+1}, X_{t+2}$ is a turning point$\}|$.

$\mathrm{E}T = (n-2)2/3$.

Can show $T \sim AN(2n/3, 8n/45)$.

Notation: $X \sim AN(\mu, \sigma^2) \Leftrightarrow \frac{X-\mu}{\sigma} \stackrel{d}{\to} N(0,1)$.

Reject (at 5% level) the hypothesis that the series is i.i.d. if

$$\left| T - \frac{2n}{3} \right| > 1.96 \sqrt{\frac{8n}{45}}.$$

Tests for positive/negative correlations at lag 1.

Testing i.i.d.: Difference-sign test

$$S = |\{i : X_i > X_{i-1}\}| = |\{i : (\nabla X)_i > 0\}|.$$

$$\mathrm{E}S = \frac{n-1}{2}.$$

Can show $S \sim AN(n/2, n/12)$.

Reject (at 5% level) the hypothesis that the series is i.i.d. if

$$\left| S - \frac{n}{2} \right| > 1.96\sqrt{\frac{n}{12}}.$$

Tests for trend.

(But a periodic sequence can pass this test...)

Testing i.i.d.: Rank test

$$N = |\{(i,j) : X_i > X_j \text{ and } i > j\}|.$$

$$\mathrm{E}N = \frac{n(n-1)}{4}.$$

Can show $N \sim AN(n^2/4, n^3/36)$.

Reject (at 5% level) the hypothesis that the series is i.i.d. if

$$\left| N - \frac{n^2}{4} \right| > 1.96\sqrt{\frac{n^3}{36}}.$$

Tests for linear trend.

Testing if an i.i.d. sequence is Gaussian: qq plot

Plot the pairs $(m_1, X_{(1)}), \ldots, (m_n, X_{(n)})$,
where $m_j = \mathrm{E}X_{(j)}$,
$X_{(1)} < \cdots < X_{(n)}$ are order statistics from $N(0,1)$ sample of size n, and
$X_{(1)} < \cdots < X_{(n)}$ are order statistics of the series $X_1, \ldots, X_n$.

Idea: If $X_i \sim N(\mu, \sigma^2)$, then

$$\mathrm{E}X_{(j)} = \mu + \sigma m_j,$$

so $(m_j, X_{(j)})$ should be *linear*.

There are tests based on how far correlation of $(m_j, X_{(j)})$ is from 1.

Stationarity

$\{X_t\}$ is **strictly stationary** if
for all $k, t_1, \ldots, t_k, x_1, \ldots, x_k$, and h,

$$P(X_{t_1} \leq x_1, \ldots, X_{t_k} \leq x_k) = P(x_{t_1+h} \leq x_1, \ldots, X_{t_k+h} \leq x_k).$$

i.e., shifting the time axis does not affect the distribution.

We shall consider **second-order properties** only.

Mean and Autocovariance

Suppose that $\{X_t\}$ is a time series with $\mathrm{E}[X_t^2] < \infty$.

Its **mean function** is

$$\mu_t = \mathrm{E}[X_t].$$

Its **autocovariance function** is

$$\begin{aligned}\gamma_X(s,t) &= \mathrm{Cov}(X_s, X_t) \\ &= \mathrm{E}[(X_s - \mu_s)(X_t - \mu_t)].\end{aligned}$$

Weak Stationarity

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

1. μ_t is independent of t, and
2. For each h, $\gamma_X(t+h, t)$ is independent of t.

In that case, we write

$$\gamma_X(h) = \gamma_X(h, 0).$$

Stationarity

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

$$\begin{aligned}\rho_X(h) &= \frac{\gamma_X(h)}{\gamma_X(0)} \\ &= \frac{\text{Cov}(X_{t+h}, X_t)}{\text{Cov}(X_t, X_t)} \\ &= \text{Cor}(X_{t+h}, X_t).\end{aligned}$$

Stationarity

Example: i.i.d. noise, $\mathrm{E}[X_t] = 0, \mathrm{E}[X_t^2] = \sigma^2$. We have

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

Thus,

1. $\mu_t = 0$ is independent of t.
2. $\gamma_X(t+h, t) = \gamma_X(h, 0)$ for all t.

So $\{X_t\}$ is stationary.

Similarly for any white noise (uncorrelated, zero mean), $X_t \sim WN(0, \sigma^2)$.

Stationarity

Example: Random walk, $S_t = \sum_{i=1}^{t} X_i$ for i.i.d., mean zero $\{X_t\}$.
We have $\mathrm{E}[S_t] = 0, \mathrm{E}[S_t^2] = t\sigma^2$, and

$$\begin{aligned}\gamma_S(t+h,t) &= \mathrm{Cov}(S_{t+h}, S_t)\\ &= \mathrm{Cov}\left(S_t + \sum_{s=1}^{h} X_{t+s}, S_t\right)\\ &= \mathrm{Cov}(S_t, S_t) = t\sigma^2.\end{aligned}$$

1. $\mu_t = 0$ is independent of t, but
2. $\gamma_S(t+h,t)$ is not.

So $\{S_t\}$ is not stationary.

An aside: covariances

$$\text{Cov}(X+Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z),$$
$$\text{Cov}(aX, Y) = a\,\text{Cov}(X, Y),$$

Also if X and Y are independent (e.g., $X = c$), then

$$\text{Cov}(X, Y) = 0.$$

Stationarity

Example: MA(1) process (**Moving Average**):

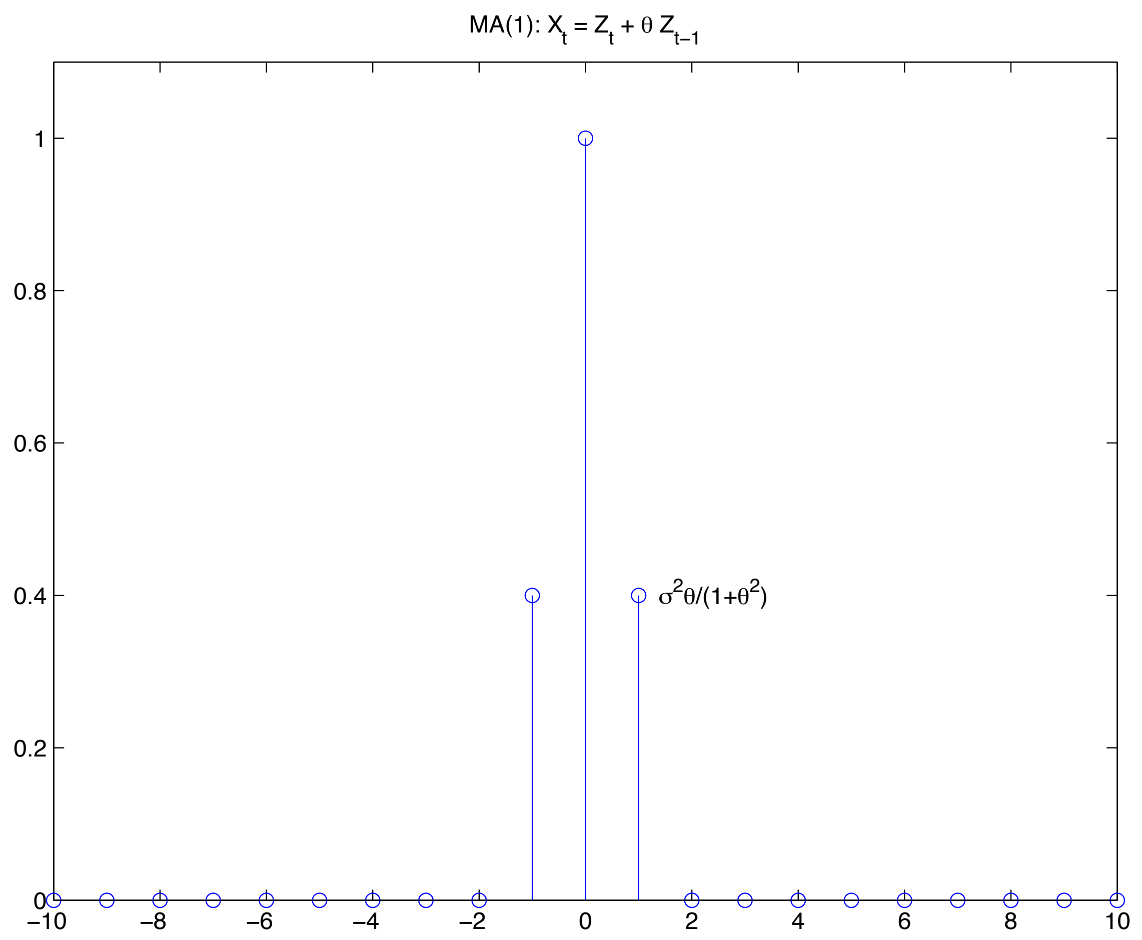
$$X_t = W_t + \theta W_{t-1}, \qquad \{W_t\} \sim WN(0, \sigma^2).$$

We have $\mathrm{E}[X_t] = 0$, and

$$\begin{aligned}\gamma_X(t+h,t) &= \mathrm{E}(X_{t+h}X_t)\\ &= \mathrm{E}[(W_{t+h} + \theta W_{t+h-1})(W_t + \theta W_{t-1})]\\ &= \begin{cases}\sigma^2(1+\theta^2) & \text{if } h = 0,\\ \sigma^2\theta & \text{if } h = \pm 1,\\ 0 & \text{otherwise.}\end{cases}\end{aligned}$$

Thus, $\{X_t\}$ is stationary.

ACF of the MA(1) process



Stationarity

Example: AR(1) process (**AutoRegressive**):

$$X_t = \phi X_{t-1} + W_t, \qquad \{W_t\} \sim WN(0, \sigma^2).$$

Assume that X_t is stationary and $|\phi| < 1$. Then we have

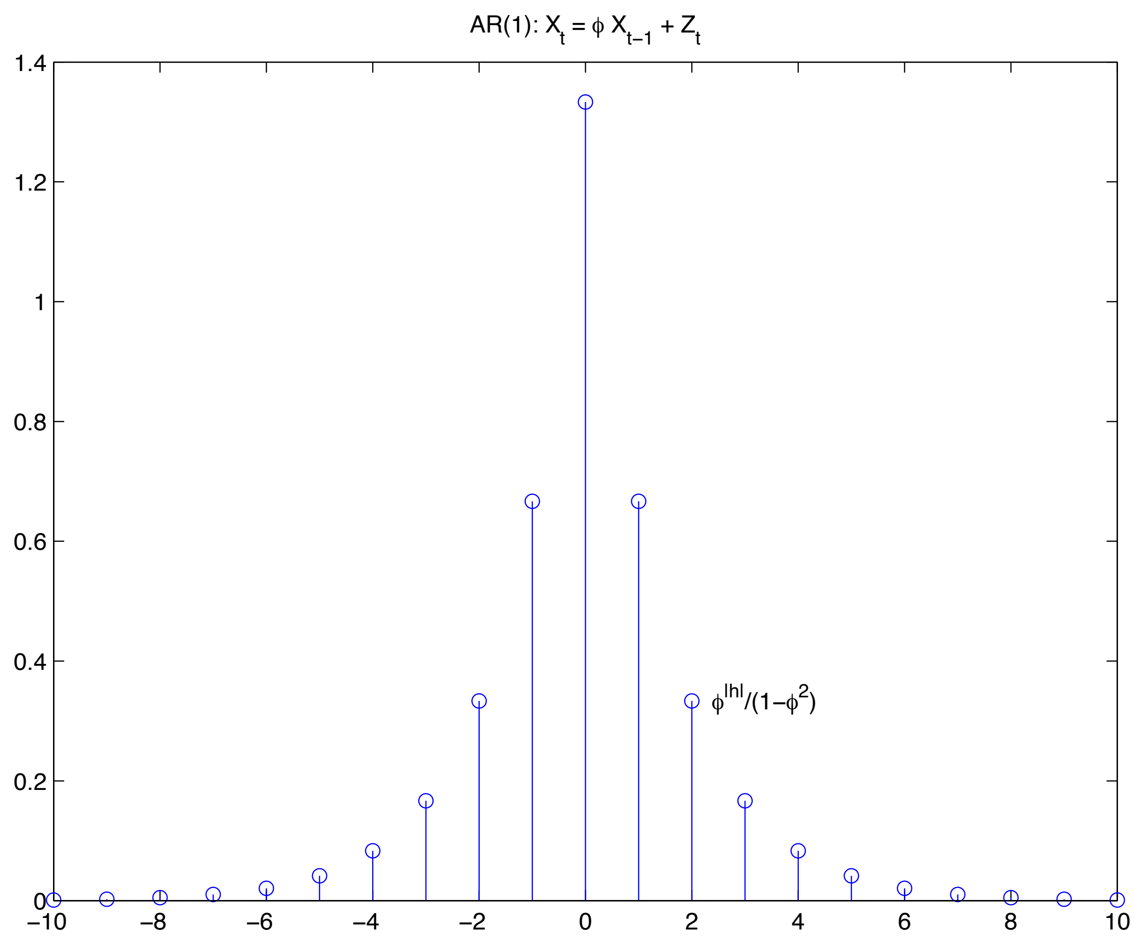
$$\begin{aligned} \mathrm{E}[X_t] &= \phi \mathrm{E} X_{t-1} \\ &= 0 \qquad \text{(from stationarity)} \\ \mathrm{E}[X_t^2] &= \phi^2 \mathrm{E}[X_{t-1}^2] + \sigma^2 \\ &= \frac{\sigma^2}{1-\phi^2} \qquad \text{(from stationarity)}, \end{aligned}$$

Stationarity

Example: AR(1) process, $X_t = \phi X_{t-1} + W_t, \qquad \{W_t\} \sim WN(0, \sigma^2)$.
Assume that X_t is stationary and $|\phi| < 1$. Then we have

$$
\begin{aligned}
\mathrm{E}[X_t] &= 0, \qquad \mathrm{E}[X_t^2] = \frac{\sigma^2}{1-\phi^2} \\
\gamma_X(h) &= \mathrm{Cov}(\phi X_{t+h-1} + Z_{t+h}, X_t) \\
&= \phi \mathrm{Cov}(X_{t+h-1}, X_t) \\
&= \phi \gamma_X(h-1) \\
&= \phi^{|h|} \gamma_X(0) \qquad \text{(check for } h > 0 \text{ and } h < 0\text{)} \\
&= \frac{\phi^{|h|}\sigma^2}{1-\phi^2}.
\end{aligned}
$$

ACF of the AR(1) process



Linear Processes

An important class of stationary time series:

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

where $\{W_t\} \sim WN(0, \sigma_w^2)$

and μ, ψ_j are parameters satisfying

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty.$$

Linear Processes

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

We have

$$\mu_X = \mu$$

$$\gamma_X(h) = \sigma_w^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{h+j}. \qquad \text{(why?)}$$

Examples of Linear Processes: White noise

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

Choose μ,

$$\psi_j = \begin{cases} 1 & \text{if } j = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\{X_t\} \sim WN(\mu, \sigma_W^2)$. (why?)

Examples of Linear Processes: MA(1)

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

Choose $\mu = 0$

$$\psi_j = \begin{cases} 1 & \text{if } j = 0, \\ \theta & \text{if } j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then $X_t = W_t + \theta W_{t-1}$. (why?)

Examples of Linear Processes: AR(1)

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

Choose $\mu = 0$

$$\psi_j = \begin{cases} \phi^j & \text{if } j \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then for $|\phi| < 1$, we have $X_t = \phi X_{t-1} + W_t$. (why?)

Estimating the ACF: Sample ACF

For observations $x_1, \ldots, x_n$ of a time series,

the **sample mean** is $$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t.$$

The **sample autocovariance function** is

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

The **sample autocorrelation function** is

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}.$$

Estimating the ACF: Sample ACF

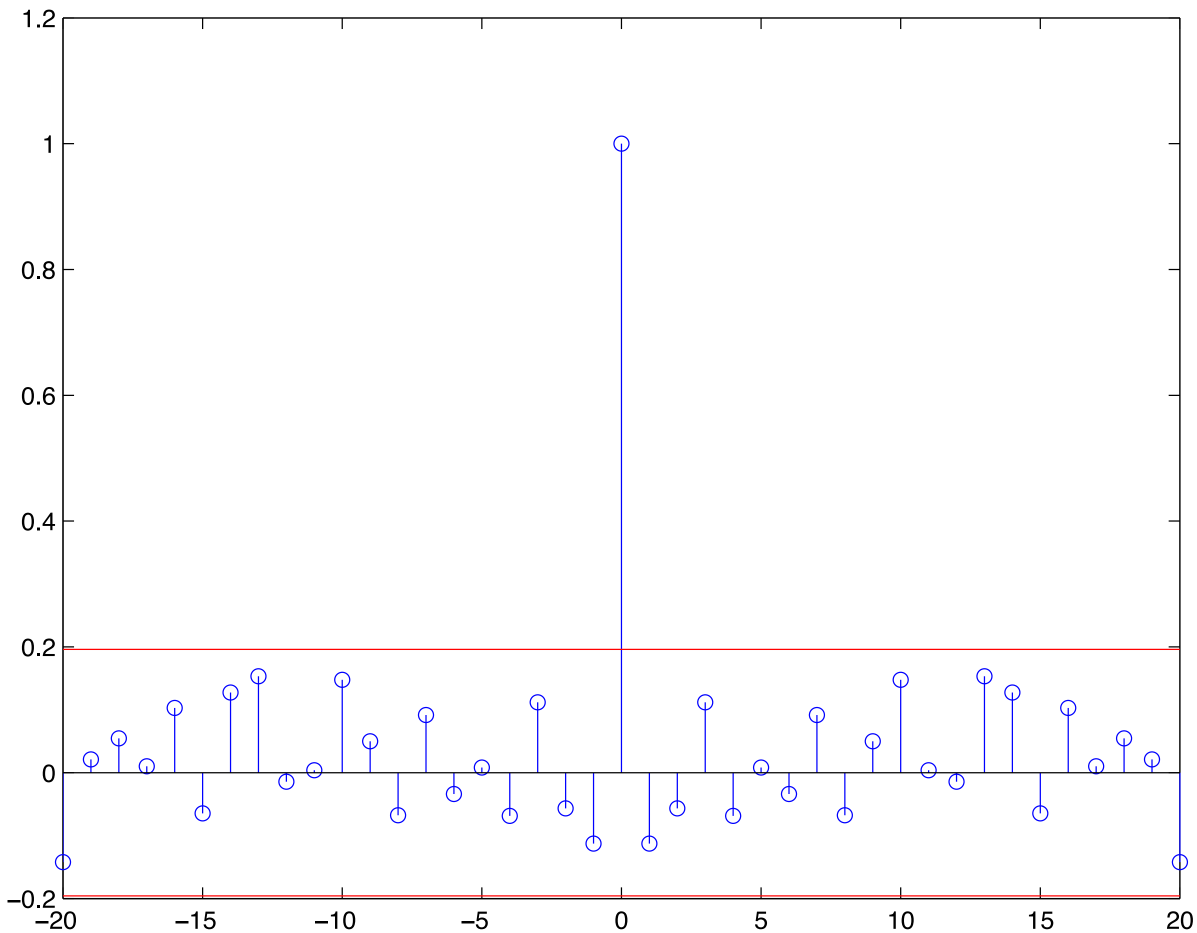
Sample autocovariance function:

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}).$$

$\approx$ the sample covariance of $(x_1, x_{h+1}), \ldots, (x_{n-h}, x_n)$, except that

- we normalize by n instead of $n - h$, and
- we subtract the full sample mean.

Sample ACF for white Gaussian (hence i.i.d.) noise

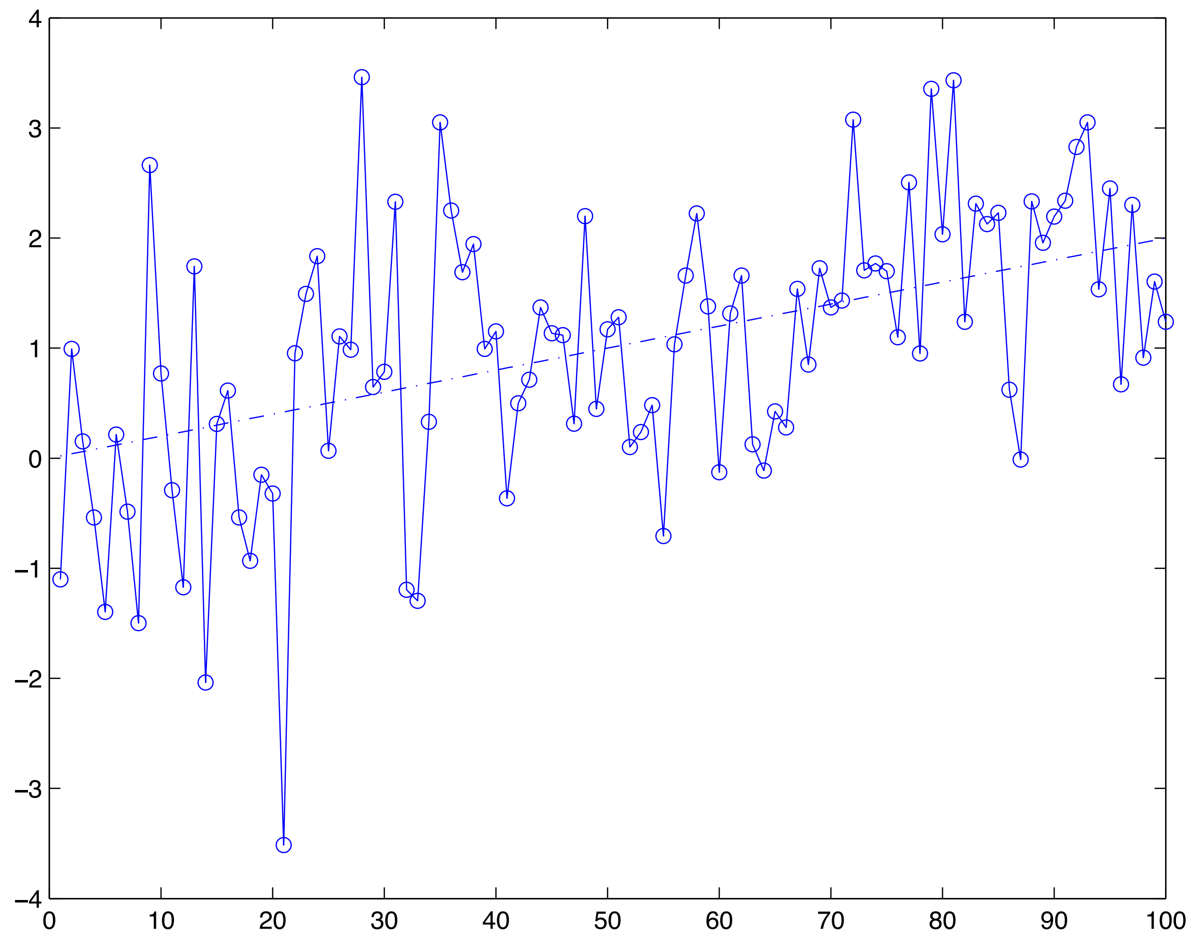


Sample ACF

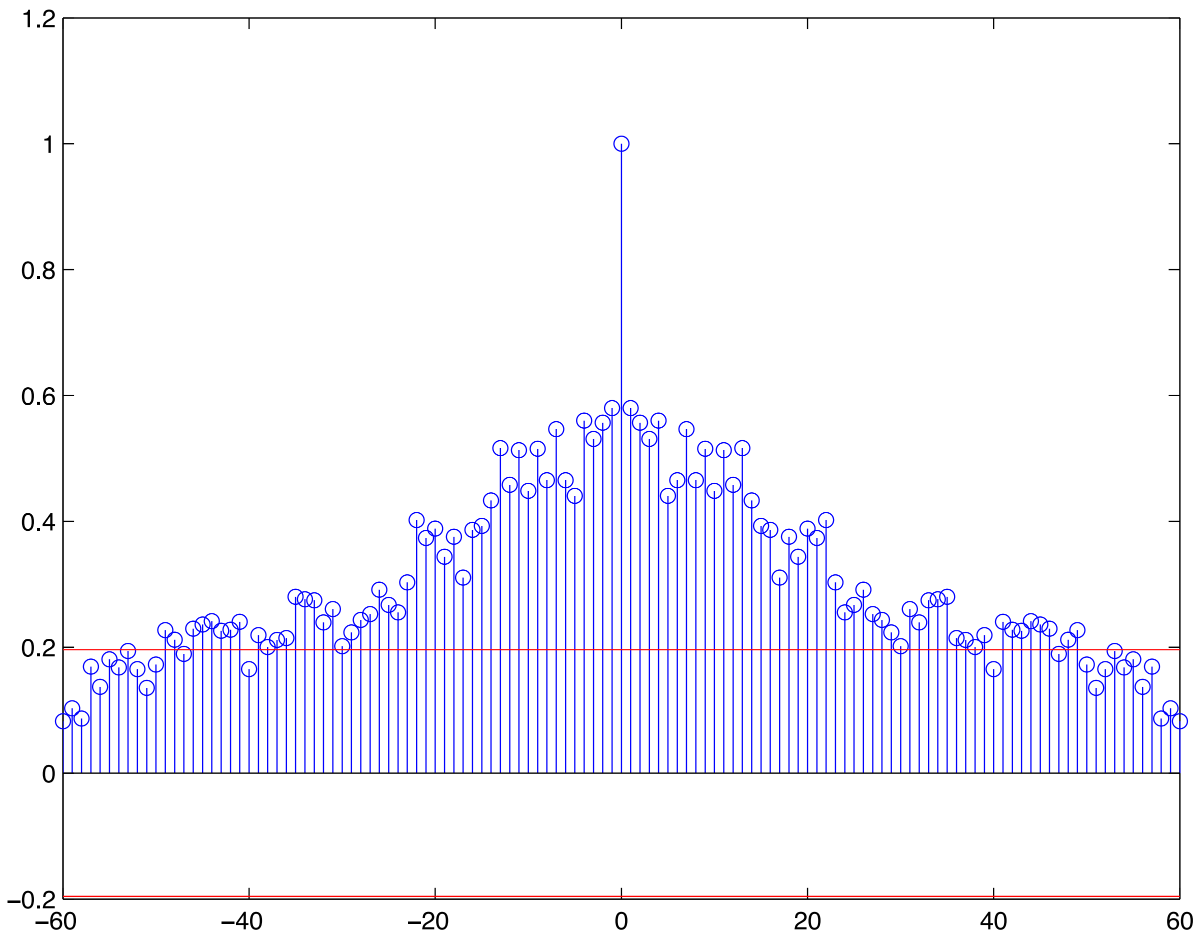
We can recognize the sample autocorrelation functions of many non-white (even non-stationary) time series.

Time series:	**Sample ACF:**
White	zero
Trend	Slow decay
Periodic	Periodic
MA(q)	Zero for $\|h\| > q$
AR(p)	Decays to zero exponentially

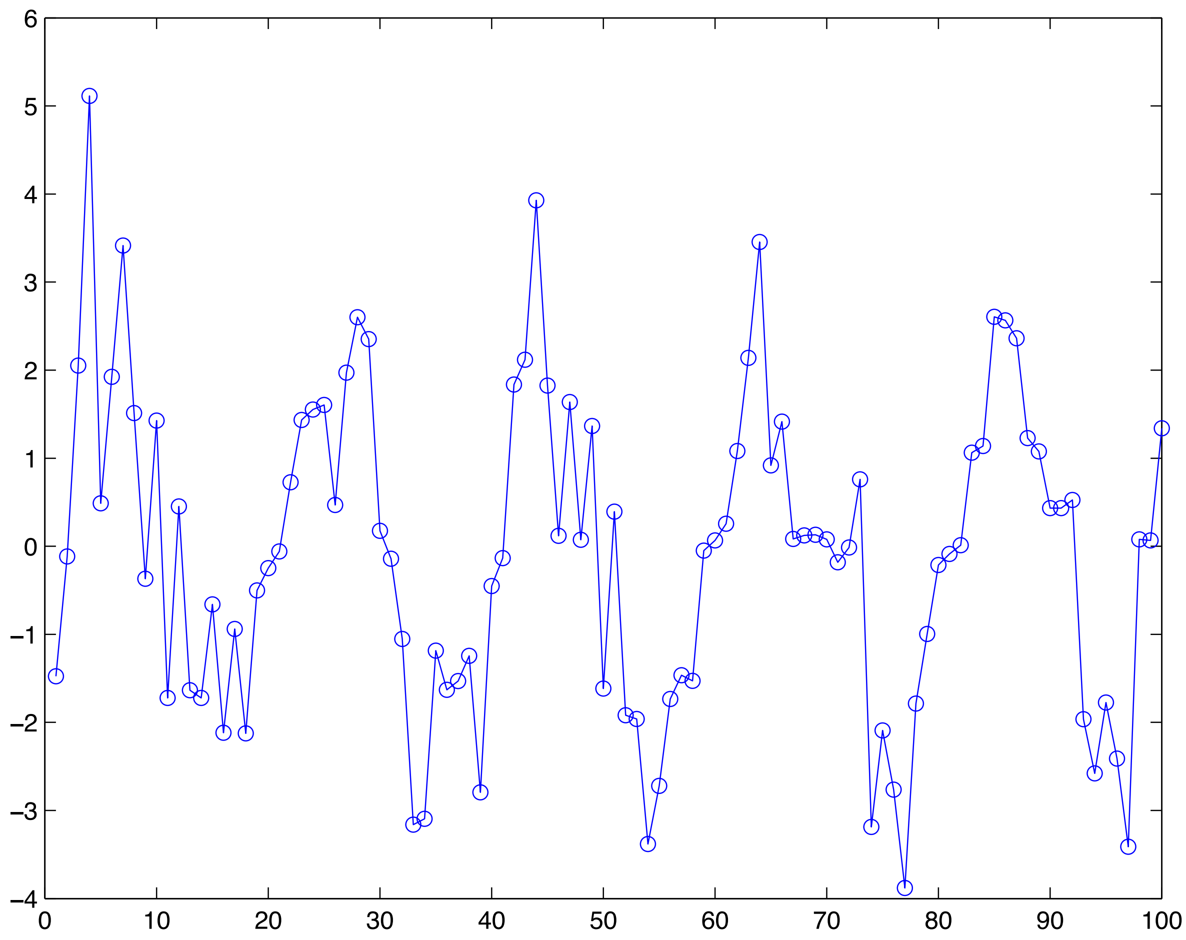
Sample ACF: Trend



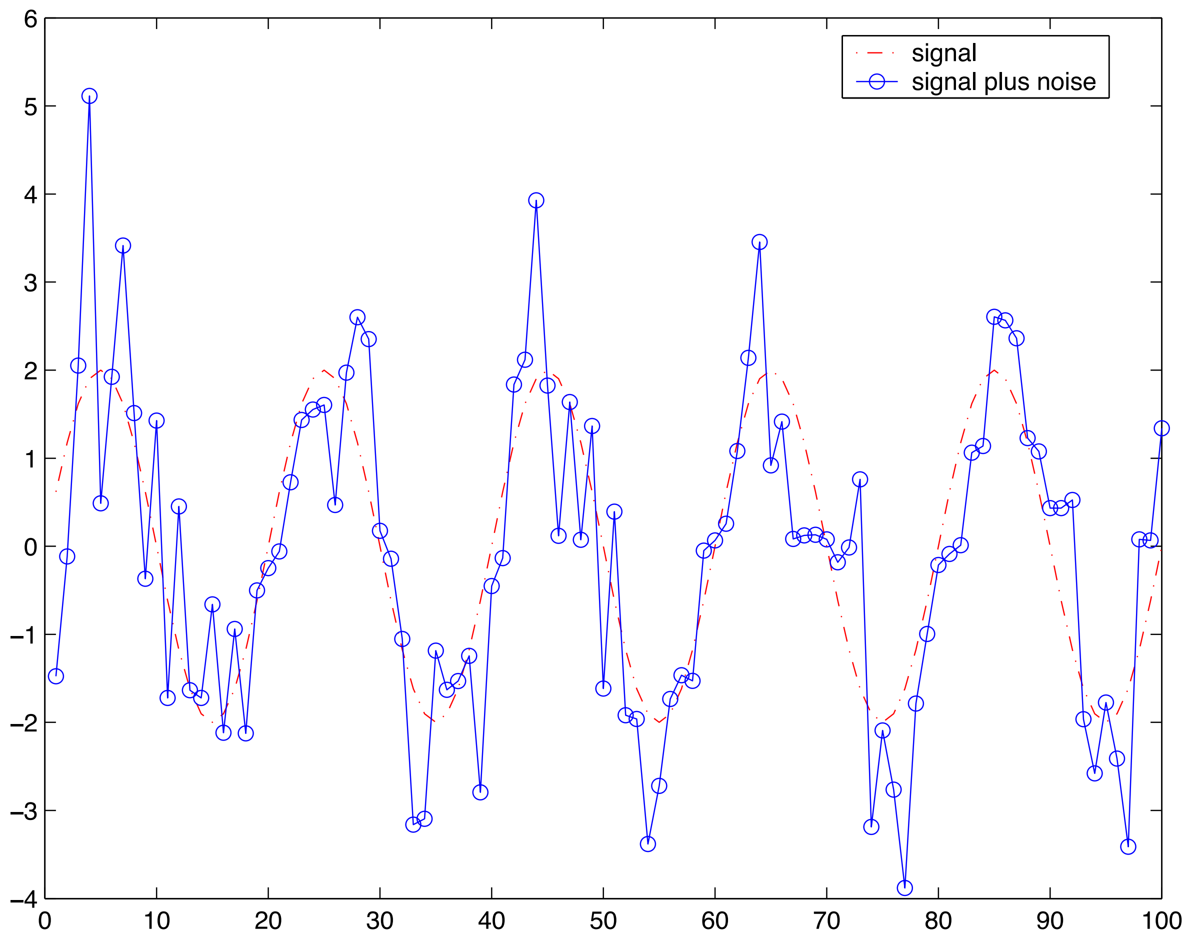
Sample ACF: Trend



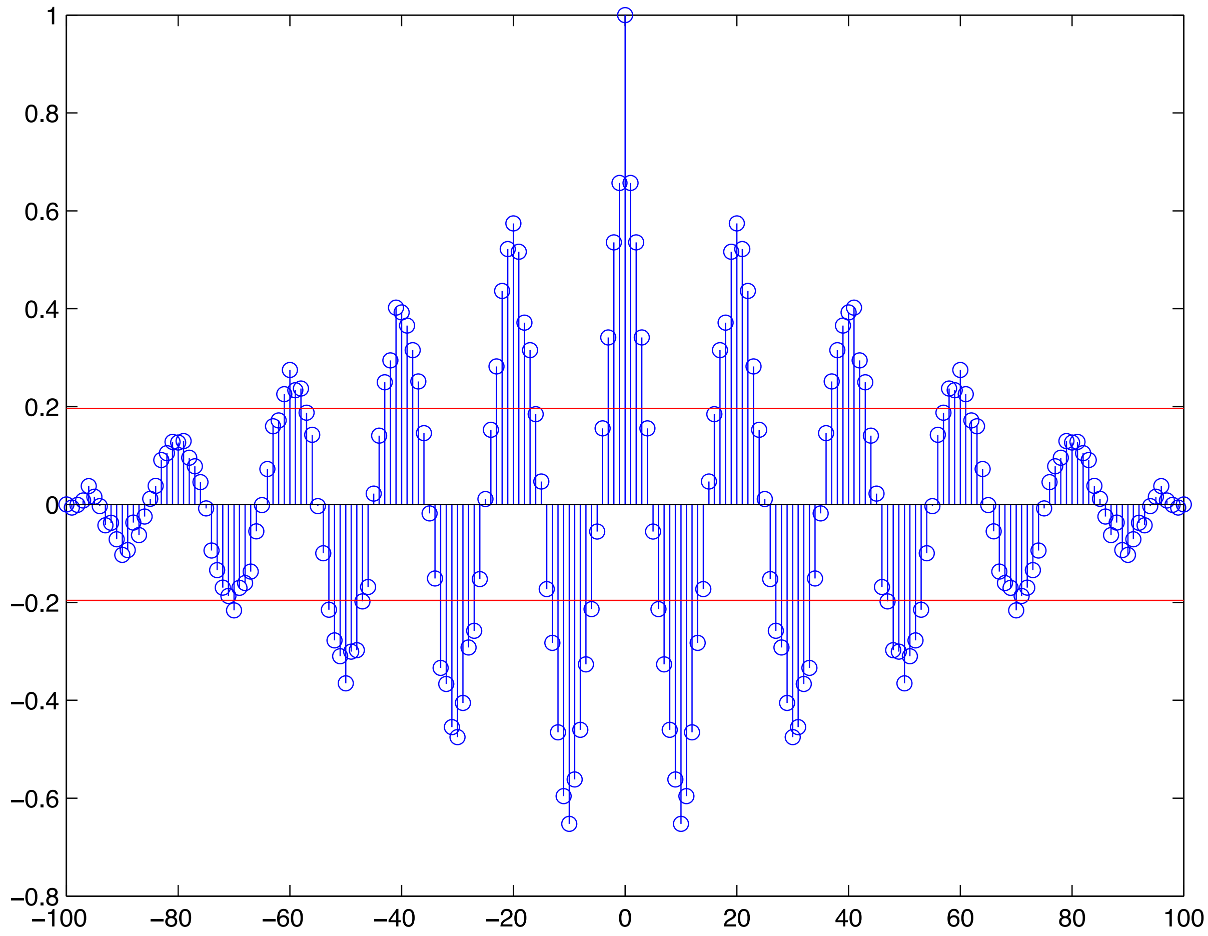
Sample ACF: Periodic



Sample ACF: Periodic

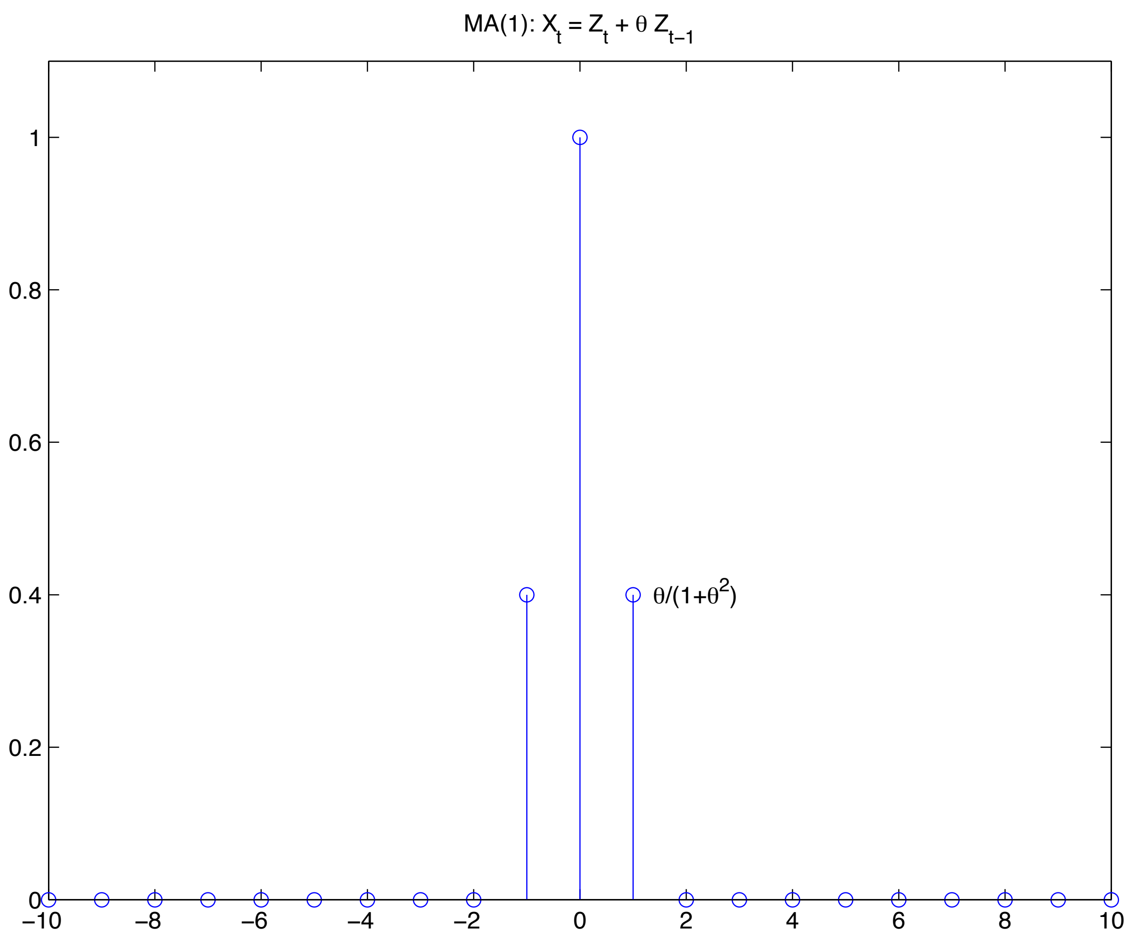


Sample ACF: Periodic

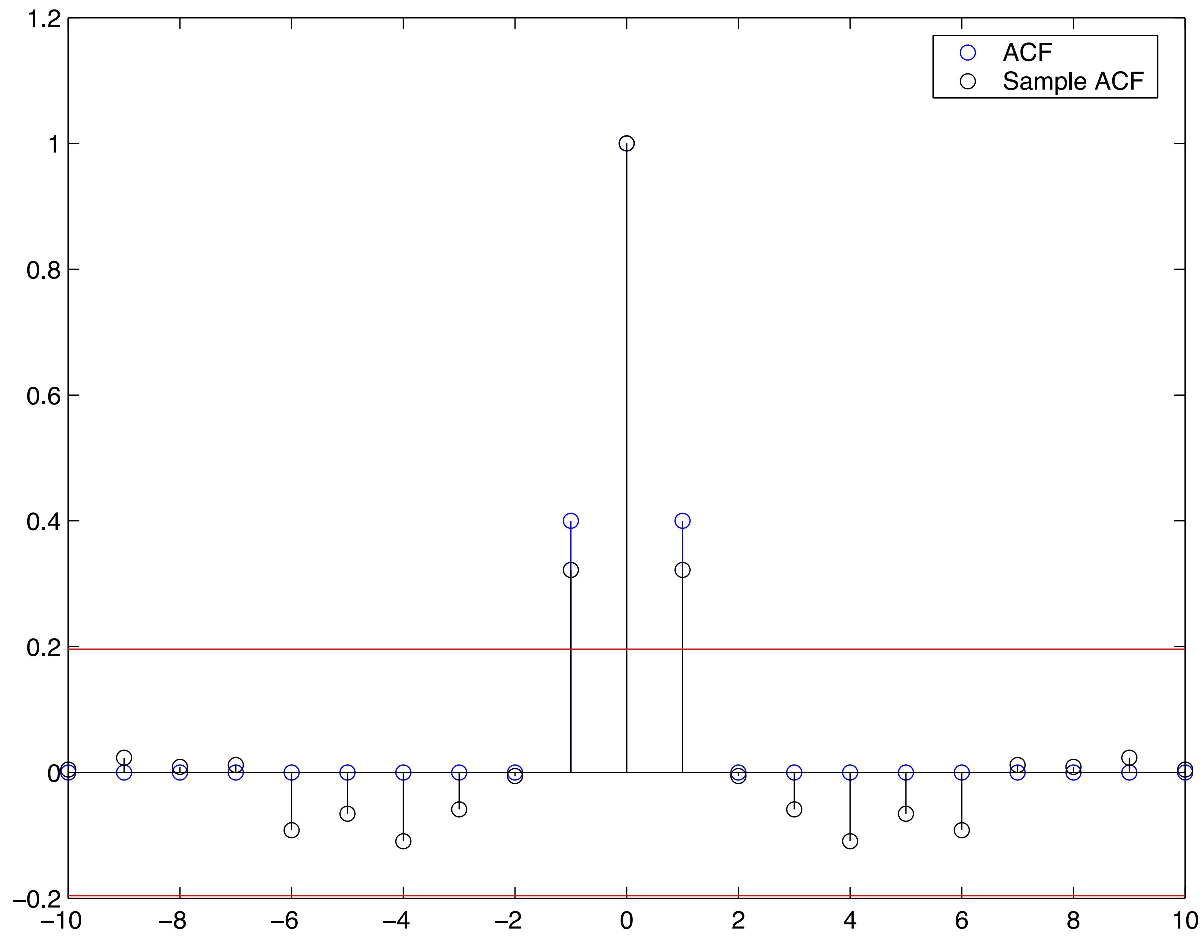


(why?)

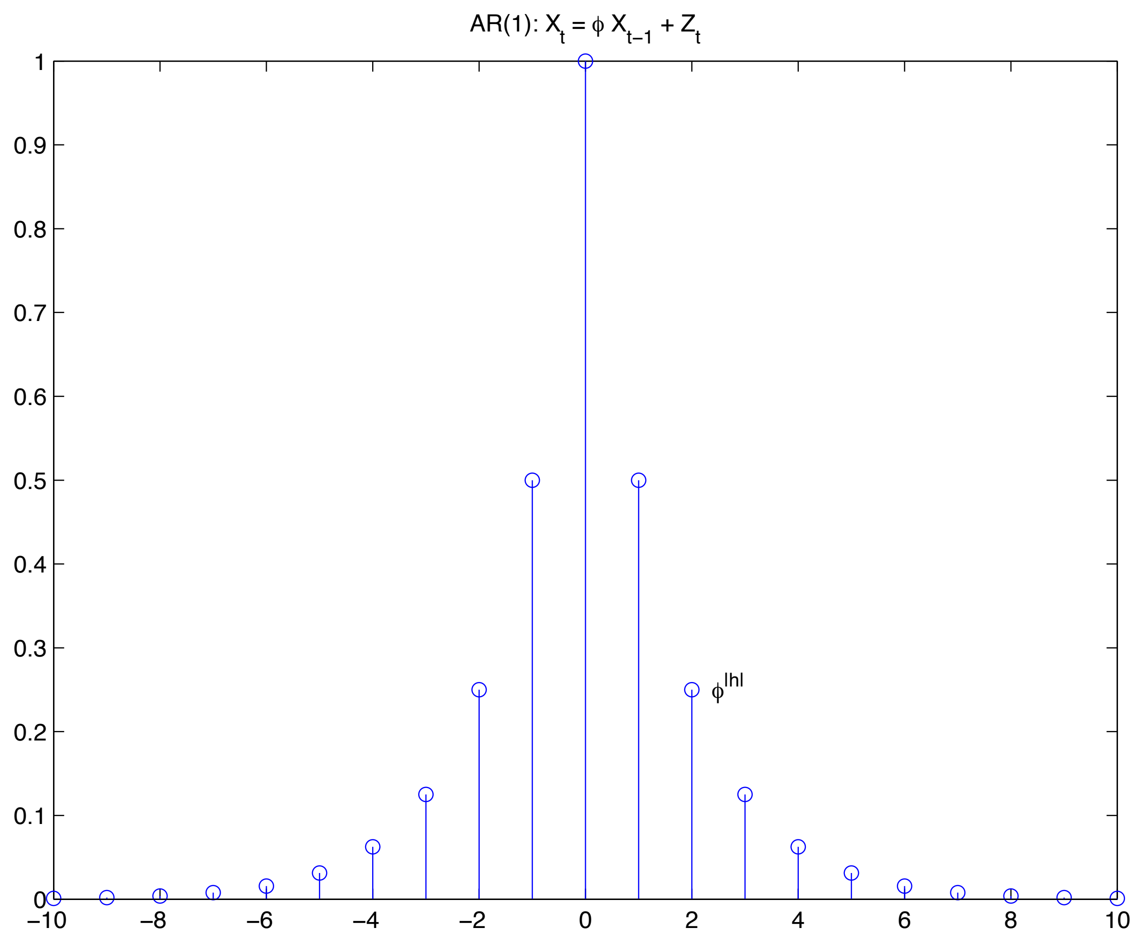
ACF: MA(1)



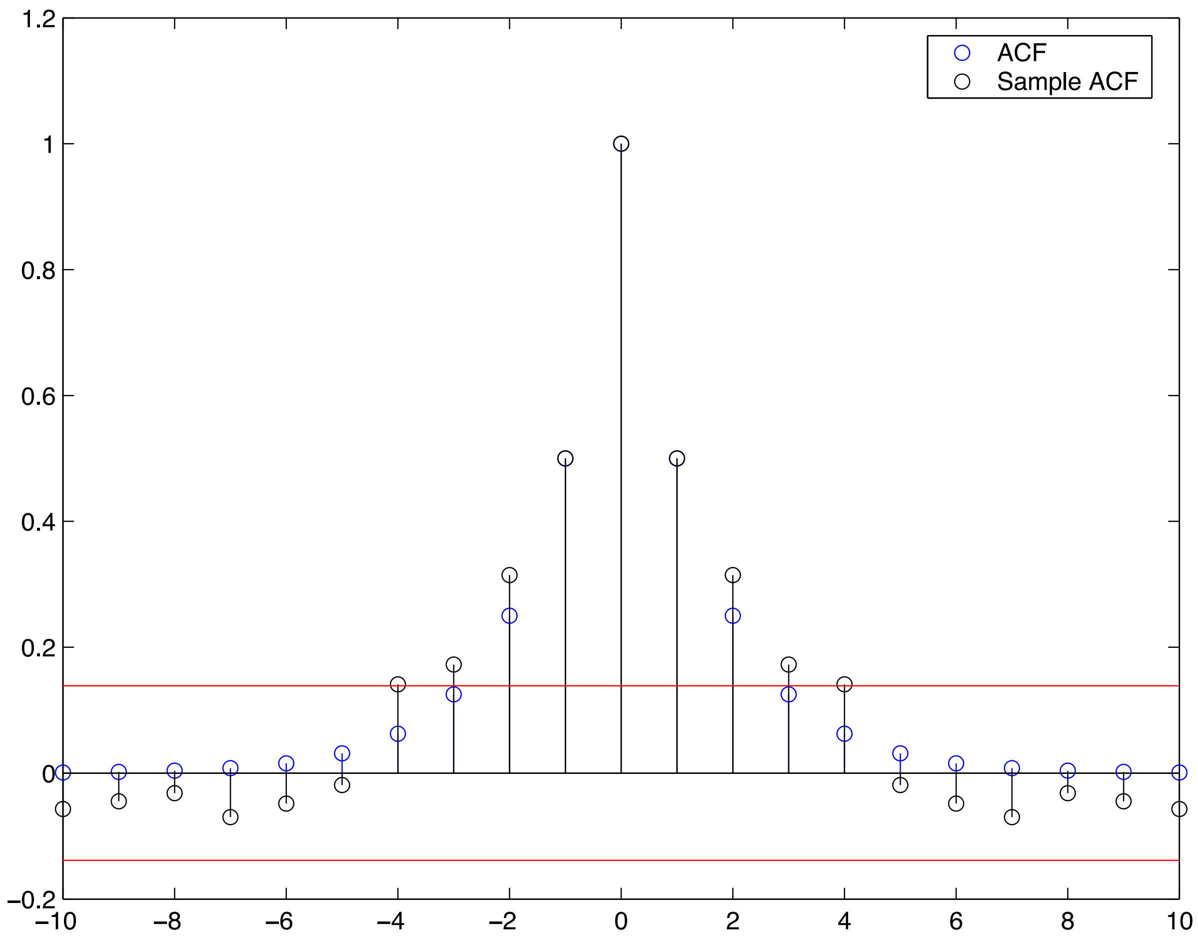
Sample ACF: MA(1)



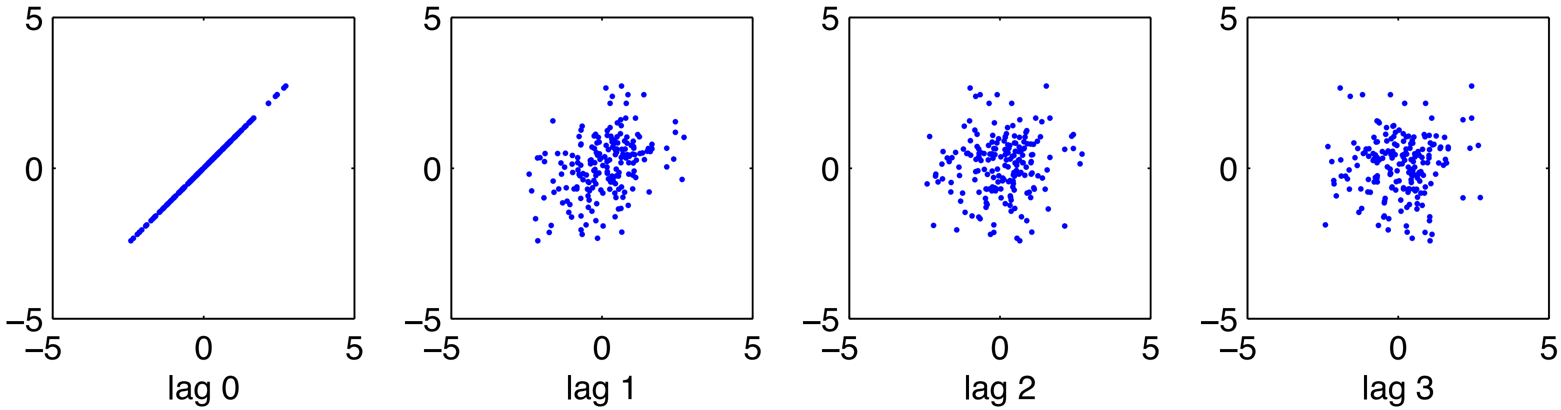
ACF: AR(1)


AR(1): $X_t = \phi X_{t-1} + Z_t$
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−10
−8
−6
−4
−2
0
2
4
6
8
10
$\phi^{|h|}$

Sample ACF: AR(1)



ACF of a MA(1) process



Properties of the autocovariance function

For the autocovariance function γ of a stationary time series $\{X_t\}$,

1. $\gamma(0) \geq 0$, (variance is non-negative)
2. $|\gamma(h)| \leq \gamma(0)$, (from Cauchy-Schwarz)
3. $\gamma(h) = \gamma(-h)$, (from stationarity)
4. γ is positive semidefinite.

Furthermore, any function $\gamma : \mathbb{Z} \to \mathbb{R}$ that satisfies (3) and (4) is the autocovariance of some stationary (Gaussian) time series.

Properties of the autocovariance function

A function $f : \mathbb{Z} \to \mathbb{R}$ is *positive semidefinite* if for all n, the matrix F_n, with entries $(F_n)_{i,j} = f(i-j)$, is positive semidefinite.

A matrix $F_n \in \mathbb{R}^{n \times n}$ is positive semidefinite if, for all vectors $a \in \mathbb{R}^n$,

$$a' F a \geq 0.$$

To see that γ is psd, consider the variance of $(X_1, \ldots, X_n)a$.

Properties of the sample autocovariance function

The **sample autocovariance function**:

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

For any sequence $x_1, \ldots, x_n$, the sample autocovariance function $\hat{\gamma}$ satisfies

1. $\hat{\gamma}(h) = \hat{\gamma}(-h)$,
2. $\hat{\gamma}$ is positive semidefinite, and hence
3. $\hat{\gamma}(0) \geq 0$ and $|\hat{\gamma}(h)| \leq \hat{\gamma}(0)$.

Properties of the sample autocovariance function: psd

$$\hat{\Gamma}_n = \begin{pmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \cdots & \hat{\gamma}(n-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & \cdots & \hat{\gamma}(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\gamma}(n-1) & \hat{\gamma}(n-2) & \cdots & \hat{\gamma}(0) \end{pmatrix}$$

$$= \frac{1}{n} M M',$$

$$\text{so } a'\hat{\Gamma}_n a = \frac{1}{n}(a'M)(M'a)$$

$$= \frac{1}{n}\|M'a\|^2$$

$$\geq 0.$$

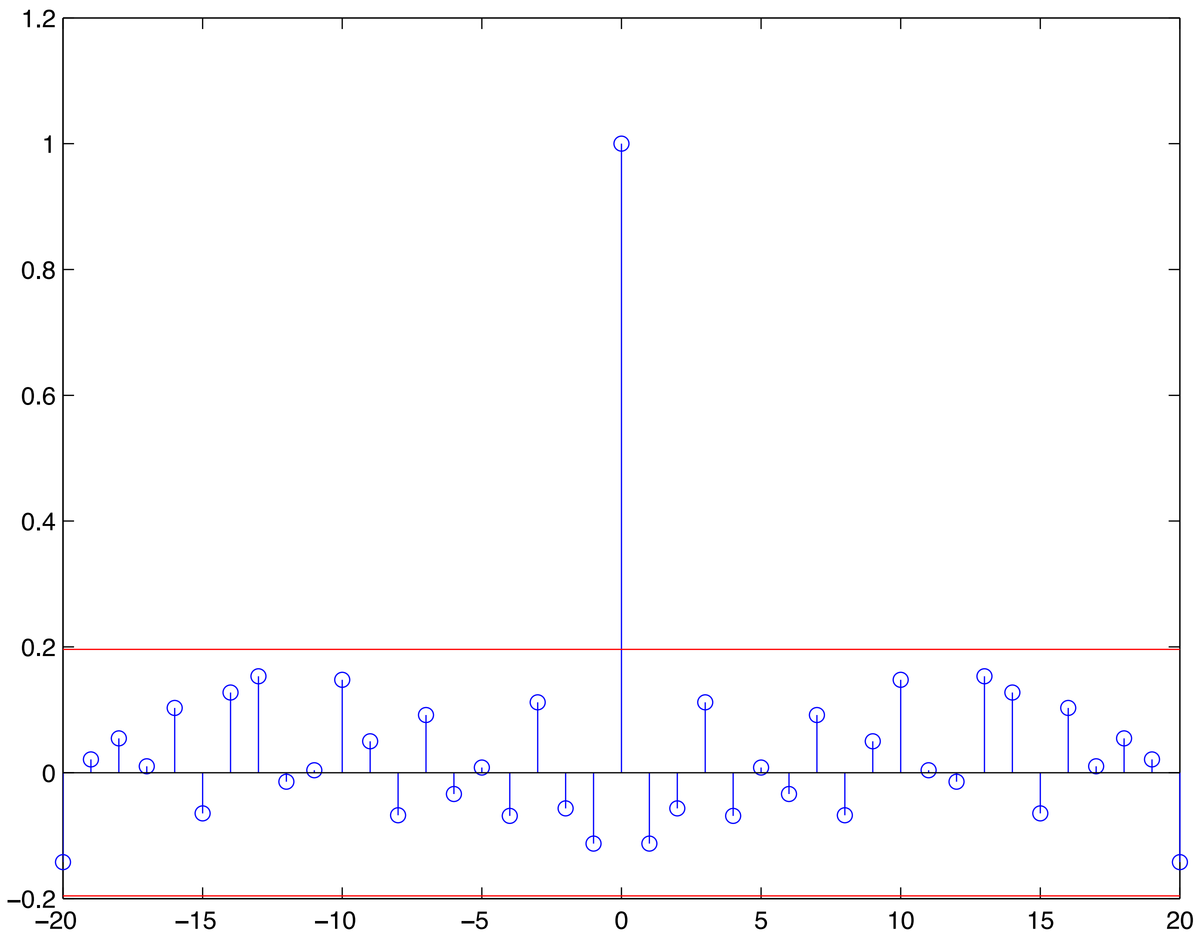
Sample ACF and testing for white noise

If $\{X_t\}$ is white noise, we expect no more than $\approx 5\%$ of the peaks of the sample ACF to satisfy

$$|\hat{\rho}(h)| > \frac{1.96}{\sqrt{n}}.$$

This is useful because we often want to introduce transformations that reduce a time series to white noise.

Sample ACF for white Gaussian (hence i.i.d.) noise



Sample ACF for MA(1)

Recall: $\rho(0) = 1$, $\rho(\pm 1) = \frac{\theta}{1+\theta^2}$, and $\rho(h) = 0$ for $|h| > 1$. Thus,

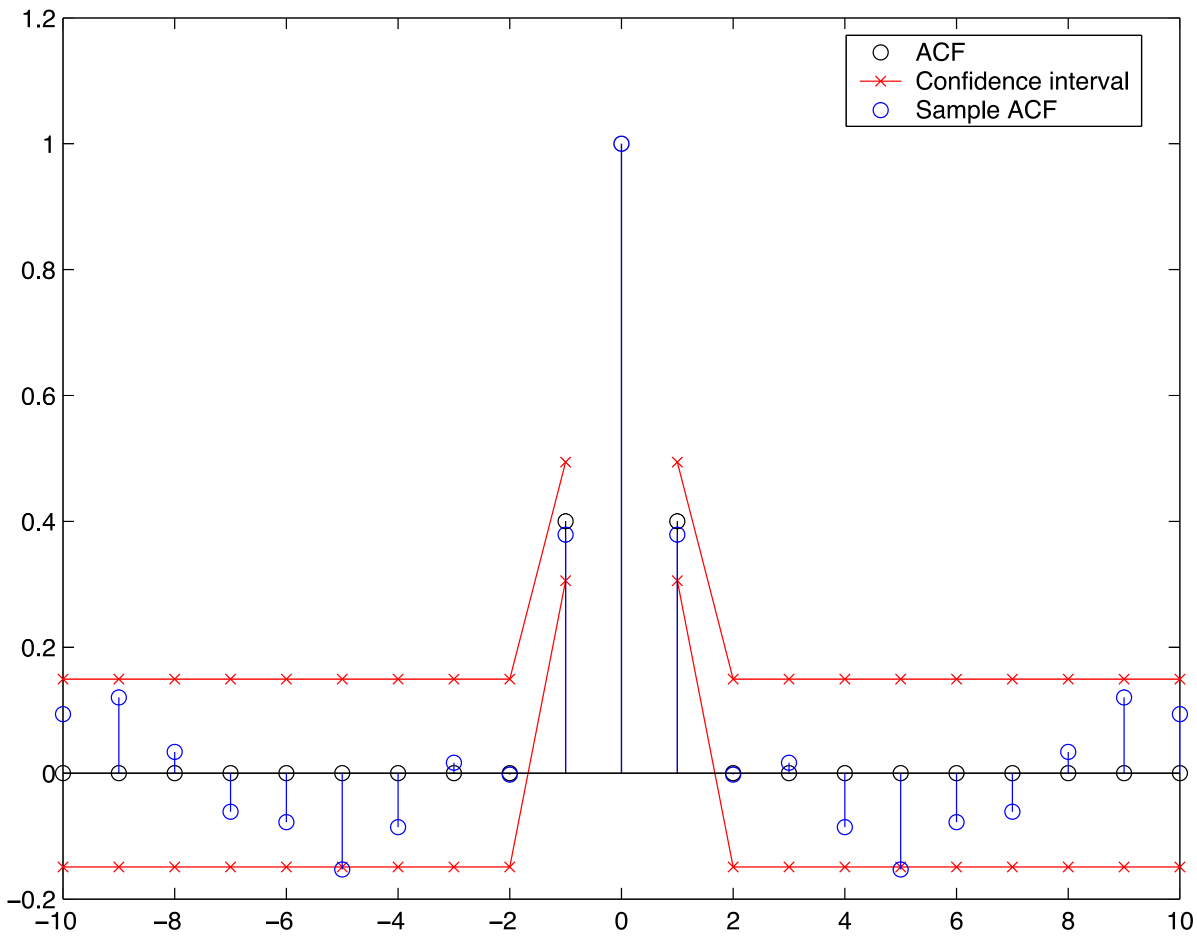
$$V_{1,1} = \sum_{h=1}^{\infty} \left(\rho(h+1) + \rho(h-1) - 2\rho(1)\rho(h)\right)^2 = (\rho(0) - 2\rho(1)^2)^2 + \rho(1)^2$$

$$V_{2,2} = \sum_{h=1}^{\infty} \left(\rho(h+2) + \rho(h-2) - 2\rho(2)\rho(h)\right)^2 = \sum_{h=-1}^{1} \rho(h)^2.$$

And if $\{X_t\}$ is a realization of this MA(1) process, with probability 0.95,

$$|\hat{\rho}(h) - \rho(h)| \leq 1.96\sqrt{\frac{V_{hh}}{n}}.$$

Sample ACF for MA(1)



AR(1) and Causality

Let X_t be the stationary solution to

$$X_t - \phi X_{t-1} = W_t,$$

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

If $|\phi| < 1$,

$$X_t = \sum_{j=0}^{\infty} \phi^j W_{t-j}.$$

$\phi = 1$?

$\phi = -1$?

$|\phi| > 1$?

AR(1) and Causality

If $|\phi| > 1$, $\pi(B)W_t$ does not converge.

But we can rearrange

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

$$\text{as} \qquad X_{t-1} = \frac{1}{\phi}X_t - \frac{1}{\phi}W_t,$$

and we can check that the unique stationary solution is

$$X_t = -\sum_{j=1}^{\infty} \phi^{-j} W_{t+j}.$$

But... X_t depends on **future** values of W_t.

Causality

A linear process $\{X_t\}$ is **causal** (strictly, a **causal function of** $\{W_t\}$) if there is a

$$\psi(B) = \psi_0 + \psi_1 B + \psi_2 B^2 + \cdots$$

with $$\sum_{j=0}^{\infty} |\psi_j| < \infty$$

and $$X_t = \psi(B) W_t.$$

AR(1) and Causality

• Causality is a property of $\{X_t\}$ **and** $\{W_t\}$.

• The AR(1) process defined by $\phi(B)X_t = W_t$ (with $\phi(B) = 1 - \phi B$) is causal iff $|\phi| < 1$, iff the root z_1 of the polynomial $\phi(z) = 1 - \phi z$ satisfies $|z_1| > 1$.

• If $|\phi| > 1$, we can define an equivalent causal model, $X_t - \phi^{-1} X_{t-1} = \tilde{W}_t$, where $\tilde{W}_t$ is a new white noise sequence.

• Is an MA(1) process causal?

MA(1) and Invertibility

Define

$$\begin{aligned} X_t &= W_t + \theta W_{t-1} \\ &= (1 + \theta B) W_t. \end{aligned}$$

If $|\theta| < 1$, we can write

$$\begin{aligned} (1+\theta B)^{-1} X_t &= W_t \\ \Leftrightarrow \qquad (1 - \theta B + \theta^2 B^2 - \theta^3 B^3 + \cdots) X_t &= W_t \\ \Leftrightarrow \qquad \sum_{j=0}^{\infty} (-\theta)^j X_{t-j} &= W_t. \end{aligned}$$

That is, we can write W_t as a *causal* function of X_t.
We say that this MA(1) is **invertible**.

MA(1) and Invertibility

$$X_t = W_t + \theta W_{t-1}$$

If $|\theta| > 1$, the sum $\sum_{j=0}^{\infty}(-\theta)^j X_{t-j}$ diverges, but we can write

$$W_{t-1} = -\theta^{-1}W_t + \theta^{-1}X_t.$$

Just like the noncausal AR(1), we can show that

$$W_t = -\sum_{j=1}^{\infty}(-\theta)^{-j}X_{t+j}.$$

That is, we can write W_t as a linear function of X_t, but it is not causal. We say that this MA(1) is not **invertible**.

Invertibility

A linear process $\{X_t\}$ is **invertible** (strictly, an **invertible function of** $\{W_t\}$) if there is a

$$\pi(B) = \pi_0 + \pi_1 B + \pi_2 B^2 + \cdots$$

$$\text{with} \qquad \sum_{j=0}^{\infty} |\pi_j| < \infty$$

$$\text{and} \qquad W_t = \pi(B) X_t.$$

MA(1) and Invertibility

- Invertibility is a property of $\{X_t\}$ and $\{W_t\}$.
- The MA(1) process defined by $X_t = \theta(B)W_t$ (with $\theta(B) = 1 + \theta B$) is invertible iff $|\theta| < 1$ iff the root z_1 of the polynomial $\theta(z) = 1 + \theta z$ satisfies $|z_1| > 1$.
- If $|\theta| > 1$, we can define an equivalent invertible model in terms of a new white noise sequence.
- Is an AR(1) process invertible?

AR(p): Autoregressive models of order p

An **AR(p) process** $\{X_t\}$ is a stationary process that satisfies

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = W_t,$$

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

Equivalently, $\phi(B)X_t = W_t$,

where $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$.

AR(p): Constraints on ϕ

Recall: For $p = 1$ (AR(1)), $\phi(B) = 1 - \phi_1 B$.

This is an AR(1) model only if there is a *stationary* solution to $\phi(B)X_t = W_t$, which is equivalent to $|\phi_1| \neq 1$.

This is equivalent to the following condition on $\phi(z) = 1 - \phi_1 z$:

$$\forall z \in \mathbb{R},\ \phi(z) = 0 \ \Rightarrow\ z \neq \pm 1$$

$$\text{equivalently, } \forall z \in \mathbb{C},\ \phi(z) = 0 \ \Rightarrow\ |z| \neq 1,$$

where $\mathbb{C}$ is the set of complex numbers.

AR(p): Constraints on ϕ

$$\text{Stationarity:} \qquad \forall z \in \mathbb{C},\ \phi(z) = 0 \Rightarrow |z| \neq 1,$$

where $\mathbb{C}$ is the set of complex numbers.

$\phi(z) = 1 - \phi_1 z$ has one root at $z_1 = 1/\phi_1 \in \mathbb{R}$.
But the roots of a degree $p > 1$ polynomial might be complex.
For stationarity, we want the roots of $\phi(z)$ to avoid the **unit circle**,
$\{z \in \mathbb{C} : |z| = 1\}$.

AR(p): Stationarity and causality

Theorem: A (unique) *stationary* solution to $\phi(B)X_t = W_t$ exists iff

$$|z| = 1 \Rightarrow \phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0.$$

This AR(p) process is *causal* iff

$$|z| \leq 1 \Rightarrow \phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0.$$

ARMA(p,q): Autoregressive moving average models

An **ARMA(p,q) process** $\{X_t\}$ is a stationary process that satisfies

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q},$$

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

- AR(p) = ARMA(p,0): $\theta(B) = 1$.
- MA(q) = ARMA(0,q): $\phi(B) = 1$.

ARMA processes

Can accurately approximate many stationary processes:

For any stationary process with autocovariance γ, and any $k > 0$, there is an ARMA process $\{X_t\}$ for which

$$\gamma_X(h) = \gamma(h), \qquad h = 0, 1, \ldots, k.$$

ARMA(p,q): Autoregressive moving average models

An **ARMA(p,q) process** $\{X_t\}$ is a stationary process that satisfies

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q},$$

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

Usually, we insist that $\phi_p, \theta_q \neq 0$ and that the polynomials

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p, \qquad \theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$$

have no common factors. This implies it is not a lower order ARMA model.

ARMA(p,q): An example of parameter redundancy

Consider a white noise process W_t. We can write

$$\begin{aligned} X_t &= W_t \\ \Rightarrow \qquad X_t - X_{t-1} + 0.25X_{t-2} &= W_t - W_{t-1} + 0.25W_{t-2} \\ (1 - B + 0.25B^2)X_t &= (1 - B + 0.25B^2)W_t \end{aligned}$$

This is in the form of an ARMA(2,2) process, with

$$\phi(B) = 1 - B + 0.25B^2, \qquad \theta(B) = 1 - B + 0.25B^2.$$

But it is white noise.

ARMA(p,q): An example of parameter redundancy

$$\text{ARMA model:} \qquad \phi(B)X_t = \theta(B)W_t,$$

$$\text{with} \qquad \phi(B) = 1 - B + 0.25B^2,$$

$$\theta(B) = 1 - B + 0.25B^2$$

$$X_t = \psi(B)W_t$$

$$\Leftrightarrow \qquad \psi(B) = \frac{\theta(B)}{\phi(B)} = 1.$$

i.e., $X_t = W_t$.

ARMA(p,q): Stationarity, causality, and invertibility

Theorem: If ϕ and θ have no common factors, a (unique) *stationary* solution to $\phi(B)X_t = \theta(B)W_t$
exists iff

$$|z| = 1 \;\Rightarrow\; \phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0.$$

This ARMA(p,q) process is *causal* iff

$$|z| \leq 1 \;\Rightarrow\; \phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0.$$

It is *invertible* iff

$$|z| \leq 1 \;\Rightarrow\; \theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q \neq 0.$$

ARMA(p,q): Stationarity, causality, and invertibility

Example: $(1 - 1.5B)X_t = (1 + 0.2B)W_t.$

$$\phi(z) = 1 - 1.5z = -\frac{3}{2}\left(z - \frac{2}{3}\right),$$
$$\theta(z) = 1 + 0.2z = \frac{1}{5}\left(z + 5\right).$$

1. ϕ and θ have no common factors, and ϕ's root is at $2/3$, which is not on the unit circle, so $\{X_t\}$ is an ARMA(1,1) process.
2. ϕ's root (at $2/3$) is inside the unit circle, so $\{X_t\}$ is *not causal*.
3. θ's root is at -5, which is outside the unit circle, so $\{X_t\}$ is *invertible*.

ARMA(p,q): Stationarity, causality, and invertibility

Example: $(1 + 0.25B^2)X_t = (1 + 2B)W_t.$

$$\phi(z) = 1 + 0.25z^2 = \frac{1}{4}\left(z^2 + 4\right) = \frac{1}{4}(z + 2i)(z - 2i),$$

$$\theta(z) = 1 + 2z = 2\left(z + \frac{1}{2}\right).$$

1. ϕ and θ have no common factors, and ϕ's roots are at $\pm 2i$, which is not on the unit circle, so $\{X_t\}$ is an ARMA(2,1) process.
2. ϕ's roots (at $\pm 2i$) are outside the unit circle, so $\{X_t\}$ is *causal*.
3. θ's root (at $-1/2$) is inside the unit circle, so $\{X_t\}$ is *not invertible*.

Causality and Invertibility

Theorem: Let $\{X_t\}$ be an ARMA process defined by $\phi(B)X_t = \theta(B)W_t$. If $\theta(z) \neq 0$ for all $|z| = 1$, then there are polynomials $\tilde{\phi}$ and $\tilde{\theta}$ and a white noise sequence $\tilde{W}_t$ such that $\{X_t\}$ satisfies $\tilde{\phi}(B)X_t = \tilde{\theta}(B)\tilde{W}_t$, and this is a causal, invertible ARMA process.

So we'll stick to causal, invertible ARMA processes.

Calculating ψ for an ARMA(p,q): matching coefficients

Example: $X_t = \psi(B)W_t \quad \Leftrightarrow \quad (1 + 0.25B^2)X_t = (1 + 0.2B)W_t$

so $\quad 1 + 0.2B = (1 + 0.25B^2)\psi(B)$

$$\Leftrightarrow \quad 1 + 0.2B = (1 + 0.25B^2)(\psi_0 + \psi_1 B + \psi_2 B^2 + \cdots)$$

$$\begin{aligned} \Leftrightarrow \quad 1 &= \psi_0, \\ 0.2 &= \psi_1, \\ 0 &= \psi_2 + 0.25\psi_0, \\ 0 &= \psi_3 + 0.25\psi_1, \\ &\vdots \end{aligned}$$

Calculating ψ for an ARMA(p,q): example

$$\Leftrightarrow \qquad 1 = \psi_0, \qquad 0.2 = \psi_1,$$
$$0 = \psi_j + 0.25\psi_{j-2}.$$

We can think of this as $\theta_j = \phi(B)\psi_j$, with $\theta_0 = 1, \theta_j = 0$ for $j < 0, j > q$.

This is a *first order difference equation* in the ψ_js.
We can use the θ_js to give the initial conditions and solve it using the theory of homogeneous difference equations.

$$\psi_j = \left(1, \tfrac{1}{5}, -\tfrac{1}{4}, -\tfrac{1}{20}, \tfrac{1}{16}, \tfrac{1}{80}, -\tfrac{1}{64}, -\tfrac{1}{320}, \ldots\right).$$

Calculating ψ for an ARMA(p,q): general case

$$\phi(B)X_t = \theta(B)W_t, \qquad \Leftrightarrow \qquad X_t = \psi(B)W_t$$

$$\text{so} \qquad \theta(B) = \psi(B)\phi(B)$$

$$\Leftrightarrow \; 1 + \theta_1 B + \cdots + \theta_q B^q = (\psi_0 + \psi_1 B + \cdots)(1 - \phi_1 B - \cdots - \phi_p B^p)$$

$$\Leftrightarrow \qquad 1 = \psi_0,$$

$$\theta_1 = \psi_1 - \phi_1 \psi_0,$$

$$\theta_2 = \psi_2 - \phi_1 \psi_1 - \cdots - \phi_2 \psi_0,$$

$$\vdots$$

This is equivalent to $\theta_j = \phi(B)\psi_j$, with $\theta_0 = 1, \theta_j = 0$ for $j < 0, j > q$.

Review: Autoregressive moving average models

An **ARMA(p,q) process** $\{X_t\}$ is a stationary process that satisfies

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q},$$

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

Usually, we insist that $\phi_p, \theta_q \neq 0$ and that the polynomials

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p, \qquad \theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$$

have no common factors. This implies it is not a lower order ARMA model.

Review: Properties of ARMA(p,q) models

Theorem: If ϕ and θ have no common factors, a (unique) *stationary* solution to $\phi(B)X_t = \theta(B)W_t$
exists iff

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p = 0 \Rightarrow |z| \neq 1.$$

This ARMA(p,q) process is *causal* iff

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p = 0 \Rightarrow |z| > 1.$$

It is *invertible* iff

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q = 0. \Rightarrow |z| > 1.$$

Review: Properties of ARMA(p,q) models

$$\phi(B)X_t = \theta(B)W_t, \qquad \Leftrightarrow \qquad X_t = \psi(B)W_t$$

$$\text{so} \qquad \theta(B) = \psi(B)\phi(B)$$

$$\Leftrightarrow \; 1 + \theta_1 B + \cdots + \theta_q B^q = (\psi_0 + \psi_1 B + \cdots)(1 - \phi_1 B - \cdots - \phi_p B^p)$$

$$\Leftrightarrow \qquad 1 = \psi_0,$$

$$\theta_1 = \psi_1 - \phi_1 \psi_0,$$

$$\theta_2 = \psi_2 - \phi_1 \psi_1 - \cdots - \phi_2 \psi_0,$$

$$\vdots$$

This is equivalent to $\theta_j = \phi(B)\psi_j$, with $\theta_0 = 1, \theta_j = 0$ for $j < 0, j > q$.

Autocovariance functions of linear processes

Consider a linear process $\{X_t\}$ defined by $X_t = \psi(B)W_t$.

$$\begin{aligned}
\gamma(h) &= \mathrm{E}\left(X_t X_{t+h}\right) \\
&= \mathrm{E}\left(\psi_0 W_t + \psi_1 W_{t-1} + \psi_2 W_{t-2} + \cdots\right) \\
&\qquad \times \left(\psi_0 W_{t+h} + \psi_1 W_{t+h-1} + \psi_2 W_{t+h-2} + \cdots\right) \\
&= \sigma_w^2 \left(\psi_0 \psi_h + \psi_1 \psi_{h+1} + \psi_2 \psi_{h+2} + \cdots\right).
\end{aligned}$$

Autocovariance functions of MA processes

Consider an MA(q) process $\{X_t\}$ defined by $X_t = \theta(B)W_t$.

$$\gamma(h) = \begin{cases} \sigma_w^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h} & \text{if } h \leq q, \\ 0 & \text{if } h > q. \end{cases}$$

Autocovariance functions of ARMA processes

ARMA process: $\phi(B)X_t = \theta(B)W_t$.

To compute γ, we can compute ψ, and then use

$$\gamma(h) = \sigma_w^2 \left(\psi_0\psi_h + \psi_1\psi_{h+1} + \psi_2\psi_{h+2} + \cdots \right).$$

Autocovariance functions of ARMA processes

An alternative approach:

$$\begin{aligned}
X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} \\
&= W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q}, \\
\text{so } \mathrm{E}\left((X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p})\, X_{t-h}\right) \\
&= \mathrm{E}\left((W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q})\, X_{t-h}\right), \\
\text{that is, } \gamma(h) - \phi_1 \gamma(h-1) - \cdots - \phi_p \gamma(h-p) \\
&= \mathrm{E}\left(\theta_h W_{t-h} X_{t-h} + \cdots + \theta_q W_{t-q} X_{t-h}\right) \\
&= \sigma_w^2 \sum_{j=0}^{q-h} \theta_{h+j} \psi_j. \qquad \text{(Write } \theta_0 = 1\text{).}
\end{aligned}$$

This is a linear difference equation.

Autocovariance functions of ARMA processes: Example

$$(1+0.25B^2)X_t = (1+0.2B)W_t, \qquad \Leftrightarrow \qquad X_t = \psi(B)W_t,$$

$$\psi_j = \left(1, \frac{1}{5}, -\frac{1}{4}, -\frac{1}{20}, \frac{1}{16}, \frac{1}{80}, -\frac{1}{64}, -\frac{1}{320}, \ldots\right).$$

$$\gamma(h) - \phi_1\gamma(h-1) - \phi_2\gamma(h-2) = \sigma_w^2 \sum_{j=0}^{q-h} \theta_{h+j}\psi_j$$

$$\Leftrightarrow \gamma(h) + 0.25\gamma(h-2) = \begin{cases} \sigma_w^2\left(\psi_0 + 0.2\psi_1\right) & \text{if } h=0, \\ 0.2\sigma_w^2\psi_0 & \text{if } h=1, \\ 0 & \text{otherwise.} \end{cases}$$

Autocovariance functions of ARMA processes: Example

We have the homogeneous linear difference equation

$$\gamma(h) + 0.25\gamma(h-2) = 0$$

for $h \geq 2$, with initial conditions

$$\gamma(0) + 0.25\gamma(-2) = \sigma_w^2\,(1 + 1/25)$$
$$\gamma(1) + 0.25\gamma(-1) = \sigma_w^2/5.$$

We can solve these linear equations to determine γ.

Or we can use the theory of linear difference equations...

Difference equations

Examples:

$$x_t - 3x_{t-1} = 0 \qquad \text{(first order, linear)}$$

$$x_t - x_{t-1}x_{t-2} = 0 \qquad \text{(3rd order, nonlinear)}$$

$$x_t + 2x_{t-1} - x_{t-3}^2 = 0 \qquad \text{(3rd order, nonlinear)}$$

Homogeneous linear diff eqns with constant coefficients

$$
\begin{aligned}
a_0 x_t + a_1 x_{t-1} + \cdots + a_k x_{t-k} &= 0 \\
\Leftrightarrow \qquad \left(a_0 + a_1 B + \cdots + a_k B^k\right) x_t &= 0 \\
\Leftrightarrow \qquad a(B) x_t &= 0 \\
\text{auxiliary equation:} \qquad a_0 + a_1 z + \cdots + a_k z^k &= 0 \\
\Leftrightarrow \qquad (z - z_1)(z - z_2) \cdots (z - z_k) &= 0
\end{aligned}
$$

where $z_1, z_2, \ldots, z_k \in \mathbb{C}$ are the roots of this *characteristic polynomial*.

Thus,

$$
a(B) x_t = 0 \qquad \Leftrightarrow \qquad (B - z_1)(B - z_2) \cdots (B - z_k) x_t = 0.
$$

Homogeneous linear diff eqns with constant coefficients

$$a(B)x_t = 0 \qquad \Leftrightarrow \qquad (B - z_1)(B - z_2) \cdots (B - z_k)x_t = 0.$$

So any $\{x_t\}$ satisfying $(B - z_i)x_t = 0$ for some i also satisfies $a(B)x_t = 0$.

Three cases:

1. The z_i are real and distinct.
2. The z_i are complex and distinct.
3. Some z_i are repeated.

Homogeneous linear diff eqns with constant coefficients

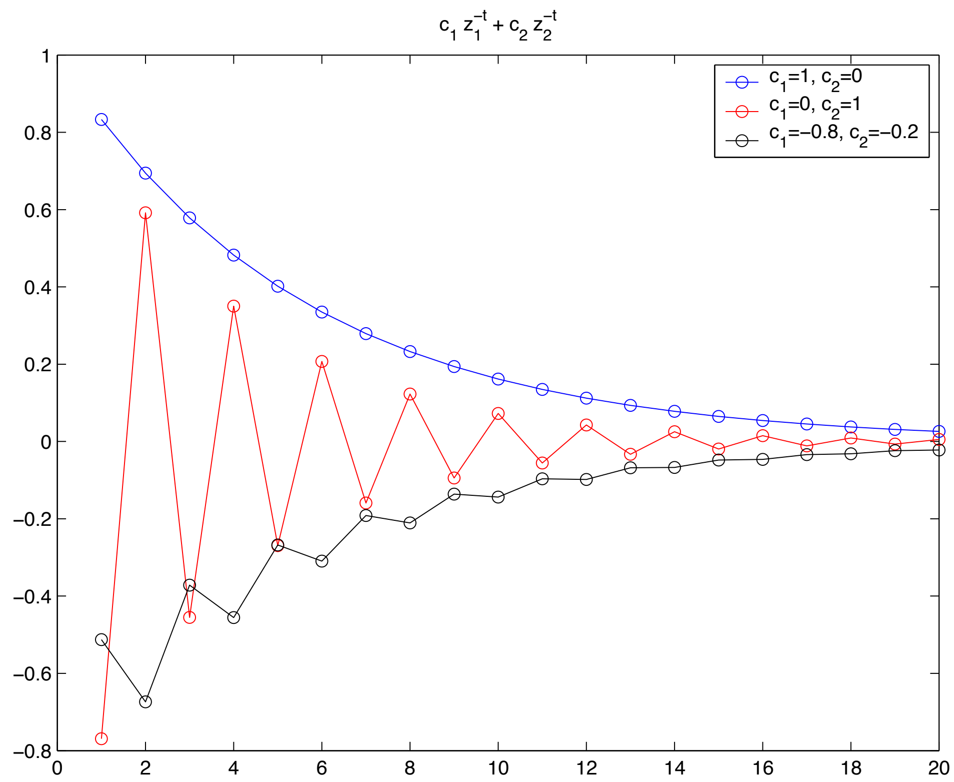
1. The z_i are real and distinct.

$$
\begin{aligned}
& & a(B)x_t = 0 \\
\Leftrightarrow & & (B - z_1)(B - z_2)\cdots(B - z_k)x_t = 0 \\
\Leftrightarrow & & (B - z_1)x_t = 0 \text{ or } (B - z_2)x_t = 0 \text{ or } \cdots \text{ or } (B - z_k)x_t = 0 \\
\Leftrightarrow & & x_t = z_1^{-1}x_{t-1} \text{ or } x_t = z_2^{-1}x_{t-1} \text{ or } \cdots \text{ or } x_t = z_k^{-1}x_{t-1} \\
\Leftrightarrow & & x_t = c_1 z_1^{-t} + c_2 z_2^{-t} + \cdots + c_k z_k^{-t},
\end{aligned}
$$

for some constants $c_1, \ldots, c_k$.

Homogeneous linear diff eqns with constant coefficients

1. The z_i are real and distinct. $z_1 = 1.2, z_2 = -1.3$



Homogeneous linear diff eqns with constant coefficients

2. The z_i are complex and distinct.

As before, $$a(B)x_t = 0$$

$$\Leftrightarrow \qquad x_t = c_1 z_1^{-t} + c_2 z_2^{-t} + \cdots + c_k z_k^{-t}.$$

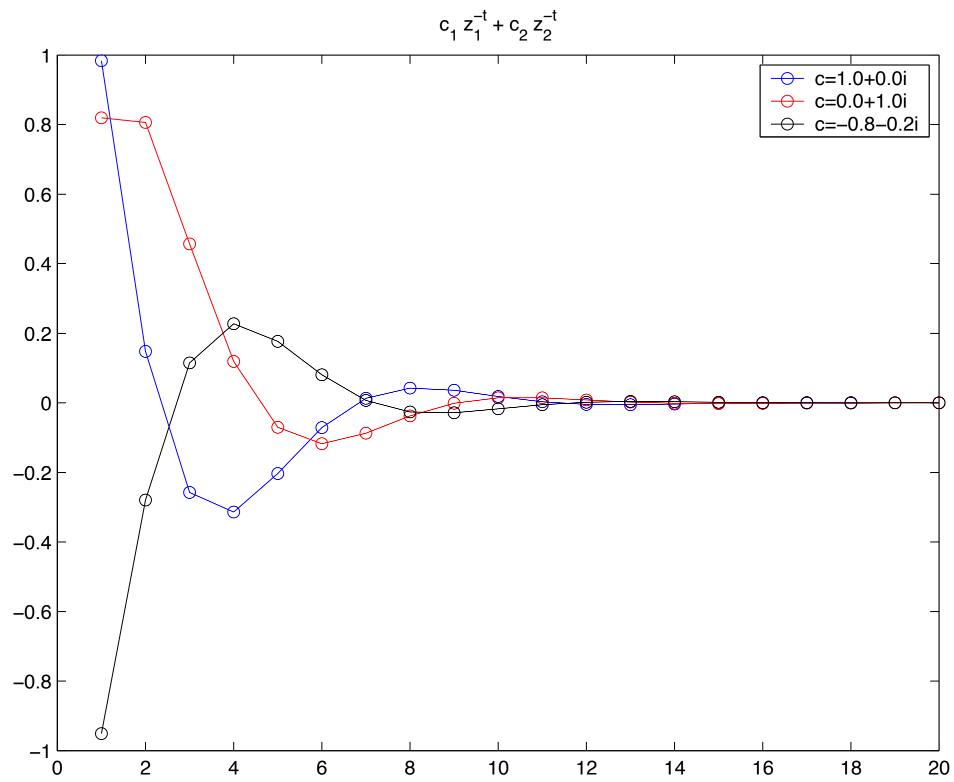
If $z_1 \notin \mathbb{R}$, since $a_1, \ldots, a_k$ are real, we must have the complex conjugate root, $z_j = \bar{z_1}$. And for x_t to be real, we must have $c_j = \bar{c_1}$. For example:

$$\begin{aligned}
x_t &= c\, z_1^{-t} + \bar{c}\, \bar{z_1}^{-t} \\
&= r\, e^{i\theta} |z_1|^{-t} e^{-i\omega t} + r\, e^{-i\theta} |z_1|^{-t} e^{i\omega t} \\
&= r|z_1|^{-t} \left(e^{i(\theta - \omega t)} + e^{-i(\theta - \omega t)} \right) \\
&= 2r|z_1|^{-t} \cos(\omega t - \theta)
\end{aligned}$$

where $z_1 = |z_1| e^{i\omega}$ and $c = re^{i\theta}$.

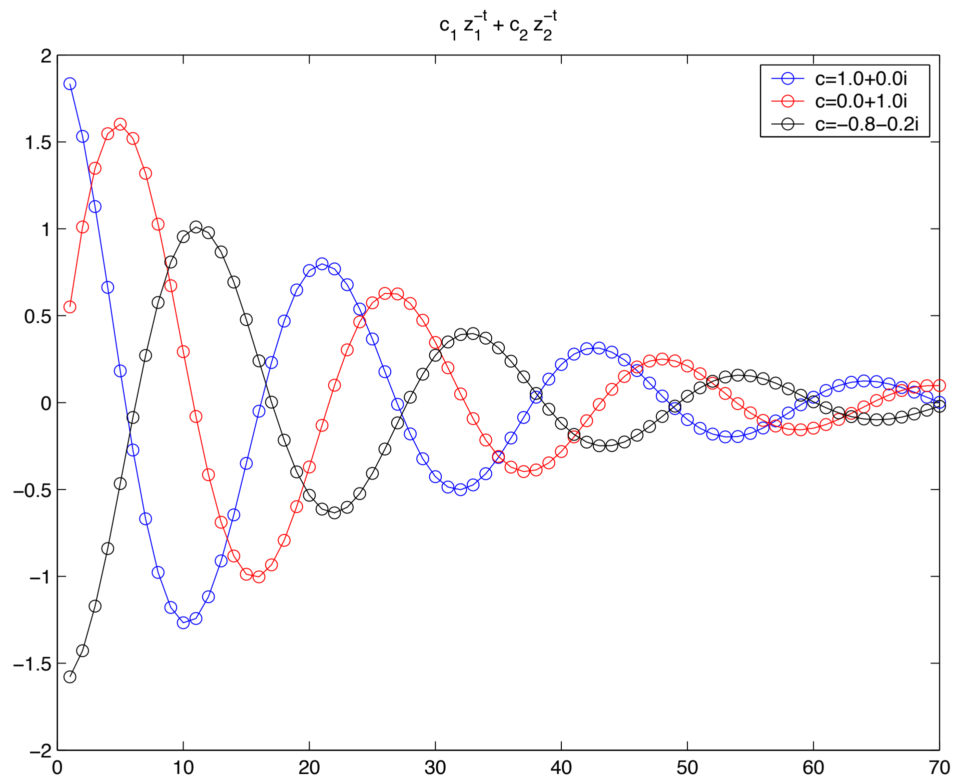
Homogeneous linear diff eqns with constant coefficients

2. The z_i are complex and distinct. $z_1 = 1.2 + i$, $z_2 = 1.2 - i$



Homogeneous linear diff eqns with constant coefficients

2. The z_i are complex and distinct. $z_1 = 1 + 0.1i, z_2 = 1 - 0.1i$



Homogeneous linear diff eqns with constant coefficients

3. Some z_i are repeated.

$$
\begin{aligned}
& & a(B)x_t = 0 \\
& \Leftrightarrow & (B - z_1)(B - z_2)\cdots(B - z_k)x_t = 0.
\end{aligned}
$$

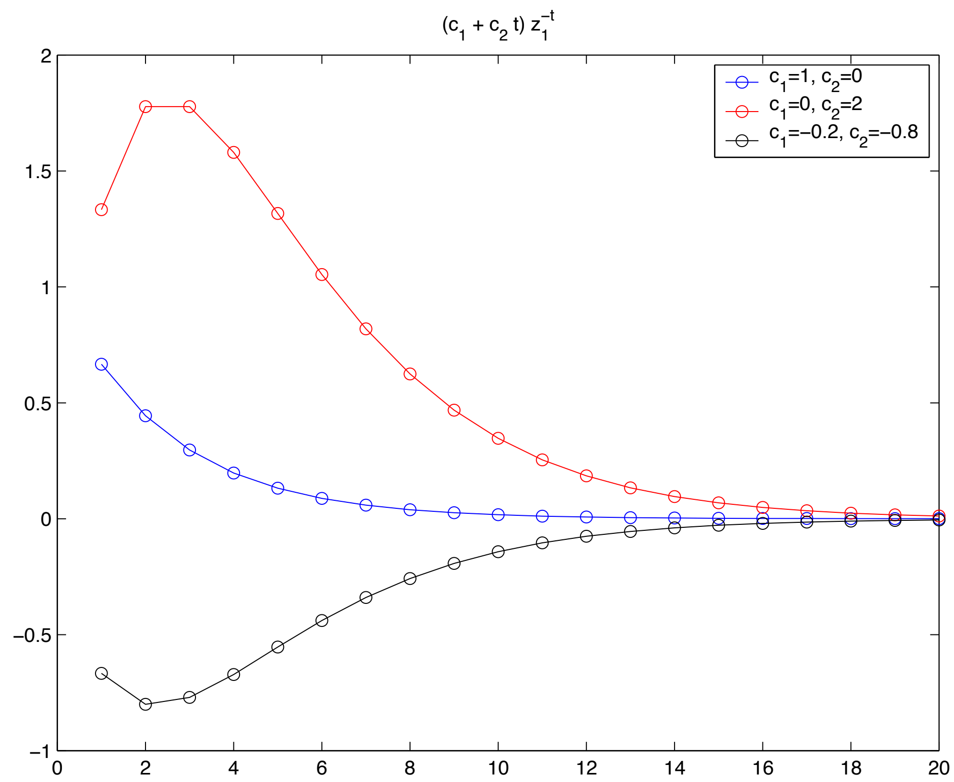
$$
\begin{aligned}
\text{If } z_1 = z_2, & & (B - z_1)(B - z_2)x_t = 0 \\
\Leftrightarrow & & (B - z_1)^2 x_t = 0.
\end{aligned}
$$

We can check that $(c_1 + c_2 t)z_1^{-t}$ is a solution...

More generally, $(B - z_1)^m x_t = 0$ has the solution $\left(c_1 + c_2 t + \cdots + c_{m-1} t^{m-1}\right) z_1^{-t}$.

Homogeneous linear diff eqns with constant coefficients

3. Some z_i are repeated. $z_1 = z_2 = 1.5$.



Solving linear diff eqns with constant coefficients

$$a_0 x_t + a_1 x_{t-1} + \cdots + a_k x_{t-k} = 0,$$
$$\text{with initial conditions} \quad x_1, \ldots, x_k.$$

Auxiliary equation in $z \in \mathbb{C}$: $$a_0 + a_1 z + \cdots + a_k z^k = 0$$
$$\Leftrightarrow \qquad (z - z_1)^{m_1}(z - z_2)^{m_2} \cdots (z - z_l)^{m_l} = 0,$$

where $z_1, z_2, \ldots, z_l \in \mathbb{C}$ are the roots of the characteristic polynomial, and z_i occurs with multiplicity m_i.
Solutions: $c_1(t) z_1^{-t} + c_2(t) z_2^{-t} + \cdots + c_l(t) z_l^{-t}$,
where $c_i(t)$ is a polynomial in t of degree $m_i - 1$.
We determine the coefficients of the $c_i(t)$ using the initial conditions (which might be linear constraints on the initial values $x_1, \ldots, x_k$).

Autocovariance functions of ARMA processes: Example

$$(1 + 0.25B^2)X_t = (1 + 0.2B)W_t, \qquad \Leftrightarrow \qquad X_t = \psi(B)W_t,$$

$$\psi_j = \left(1, \frac{1}{5}, -\frac{1}{4}, -\frac{1}{20}, \frac{1}{16}, \frac{1}{80}, -\frac{1}{64}, -\frac{1}{320}, \ldots\right).$$

$$\gamma(h) - \phi_1\gamma(h-1) - \phi_2\gamma(h-2) = \sigma_w^2 \sum_{j=0}^{q-h} \theta_{h+j}\psi_j$$

$$\Leftrightarrow \gamma(h) + 0.25\gamma(h-2) = \begin{cases} \sigma_w^2\left(\psi_0 + 0.2\psi_1\right) & \text{if } h = 0, \\ 0.2\sigma_w^2\psi_0 & \text{if } h = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Autocovariance functions of ARMA processes: Example

We have the homogeneous linear difference equation

$$\gamma(h) + 0.25\gamma(h-2) = 0$$

for $h \geq 2$, with initial conditions

$$\gamma(0) + 0.25\gamma(-2) = \sigma_w^2\,(1 + 1/25)$$
$$\gamma(1) + 0.25\gamma(-1) = \sigma_w^2/5.$$

Autocovariance functions of ARMA processes: Example

Homogeneous lin. diff. eqn:

$$\gamma(h) + 0.25\gamma(h-2) = 0.$$

The characteristic polynomial is

$$1 + 0.25z^2 = \frac{1}{4}\left(4 + z^2\right) = \frac{1}{4}(z - 2i)(z + 2i),$$

which has roots at $z_1 = 2e^{i\pi/2}$, $\bar{z_1} = 2e^{-i\pi/2}$.
The solution is of the form

$$\gamma(h) = cz_1^{-h} + \bar{c}\bar{z_1}^{-h}.$$

Autocovariance functions of ARMA processes: Example

$z_1 = 2e^{i\pi/2}, \bar{z_1} = 2e^{-i\pi/2}, c = |c|e^{i\theta}.$

We have
$$\begin{aligned}\gamma(h) &= cz_1^{-h} + \bar{c}\bar{z_1}^{-h} \\ &= 2^{-h}\left(|c|e^{i(\theta - h\pi/2)} + |c|e^{i(-\theta + h\pi/2)}\right) \\ &= c_1 2^{-h}\cos\left(\frac{h\pi}{2} - \theta\right).\end{aligned}$$

And we determine c_1, θ from the initial conditions

$$\gamma(0) + 0.25\gamma(-2) = \sigma_w^2\left(1 + 1/25\right)$$
$$\gamma(1) + 0.25\gamma(-1) = \sigma_w^2/5.$$

Autocovariance functions of ARMA processes: Example