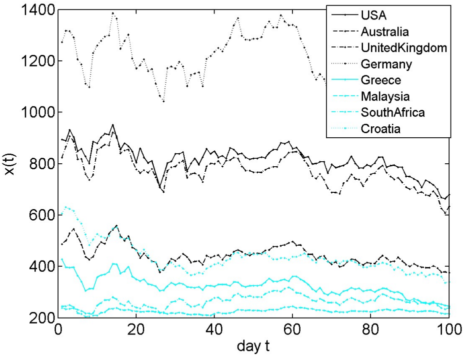
Correlation, complexity, and coupling measures of time series

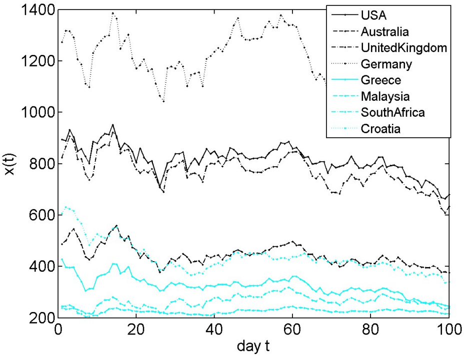
Dimitris Kugiumtzis

November 4, 2020

Time Series, time dependence

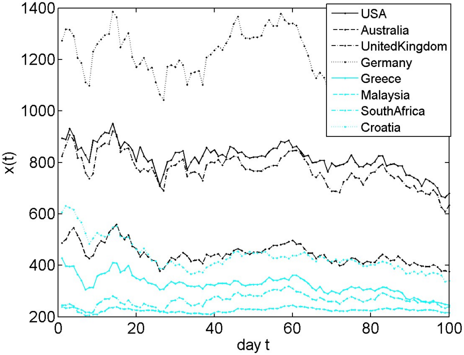


Time Series, time dependence



Time dependence? $\Longrightarrow$ Test for independence (see below)

Time Series, time dependence



Time dependence? $\Longrightarrow$ Test for independence (see below)

Time dependence

- The time series is a function of time
- The time series is a realization of a stochastic process / dynamical system

Stationarity see [1]: Sec 2.1-2.4, [2] Chp 1 and Sec 2.1

Stationary time series: the statistics do not change by the shift of time

Stationarity see [1]: Sec 2.1-2.4, [2] Chp 1 and Sec 2.1

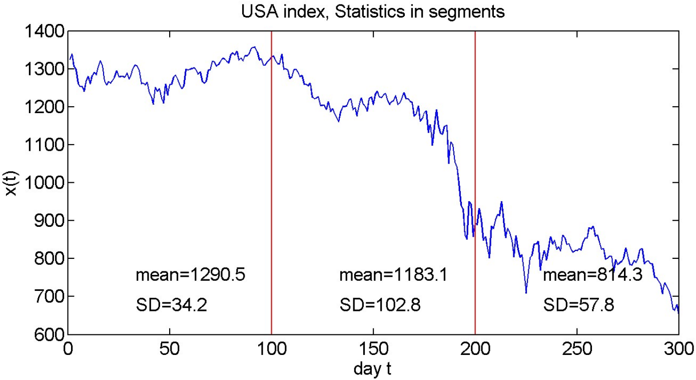
Stationary time series: the statistics do not change by the shift of time

Different definitions: strict stationarity, weak stationarity.

Stationarity see [1]: Sec 2.1-2.4, [2] Chp 1 and Sec 2.1

Stationary time series: the statistics do not change by the shift of time

Different definitions: strict stationarity, weak stationarity.



Time series decomposition see [1]: Sec 2.5-2.6 and [2]: Chp 1

Non-stationarity may be due to deterministic trend or periodicity.

Time series decomposition see [1]: Sec 2.5-2.6 and [2]: Chp 1

Non-stationarity may be due to deterministic trend or periodicity.

Time series decomposition: $y_t = \mu_t + s_t + x_t$,

Time series decomposition see [1]: Sec 2.5-2.6 and [2]: Chp 1

Non-stationarity may be due to deterministic trend or periodicity.

Time series decomposition: $y_t = \mu_t + s_t + x_t$,

μ_t : trend component, slowly varying function of time

Non-stationarity may be due to deterministic trend or periodicity.

Time series decomposition: $y_t = \mu_t + s_t + x_t$,

μ_t : trend component, slowly varying function of time

s_t : periodic / seasonal component, periodic function of time

Time series decomposition see [1]: Sec 2.5-2.6 and [2]: Chp 1

Non-stationarity may be due to deterministic trend or periodicity.

Time series decomposition: $y_t = \mu_t + s_t + x_t$,

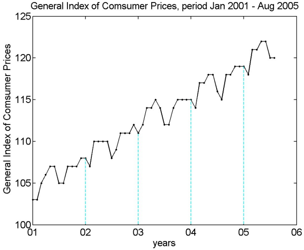
μ_t : trend component, slowly varying function of time

s_t : periodic / seasonal component, periodic function of time

x_t: residual component, stationary time series.

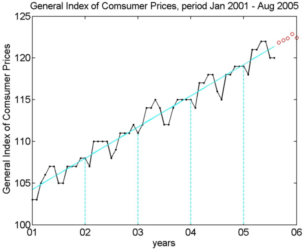
Example: Monthly Greek General Index of Consumer Prices

$\{y_t\}_{t=1}^{56}$



Example: Monthly Greek General Index of Consumer Prices

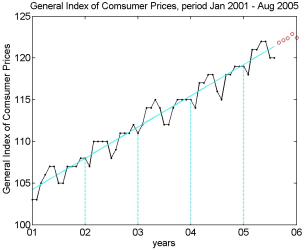
$\{y_t\}_{t=1}^{56} \quad \mu_t = 103.9 + 0.31t$



...

Example: Monthly Greek General Index of Consumer Prices

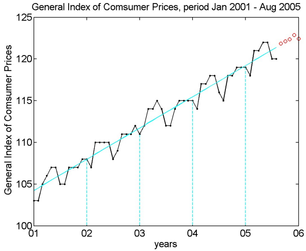
$\{y_t\}_{t=1}^{56}$ $\mu_t = 103.9 + 0.31t$



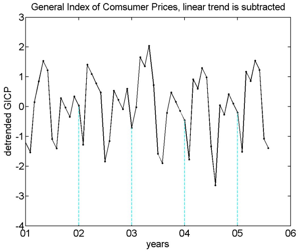
$\ldots \; y'_t = y_t - \mu_t$

Example: Monthly Greek General Index of Consumer Prices

$\{y_t\}_{t=1}^{56} \quad \mu_t = 103.9 + 0.31t$

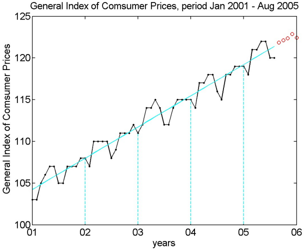


$\ldots \; y'_t = y_t - \mu_t$

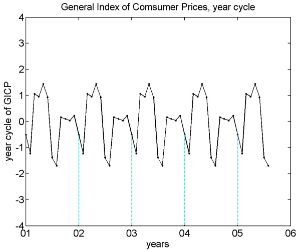


Example: Monthly Greek General Index of Consumer Prices

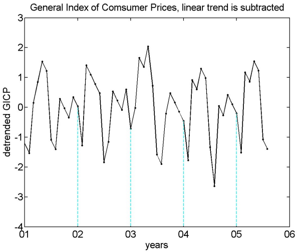
$\{y_t\}_{t=1}^{56}$ $\mu_t = 103.9 + 0.31t$



$\{s_t\}_{t=1}^{12}$

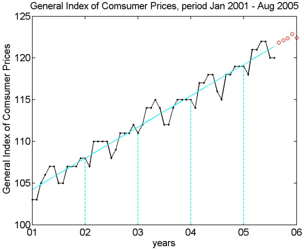


... $y'_t = y_t - \mu_t$

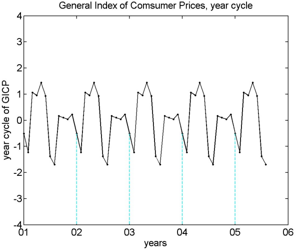


Example: Monthly Greek General Index of Consumer Prices

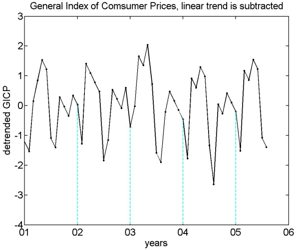
$\{y_t\}_{t=1}^{56} \quad \mu_t = 103.9 + 0.31t$



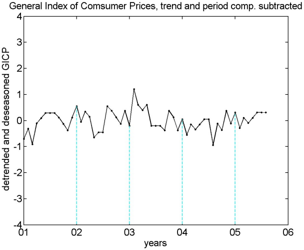
$\{s_t\}_{t=1}^{12}$



$\dots \quad y'_t = y_t - \mu_t$

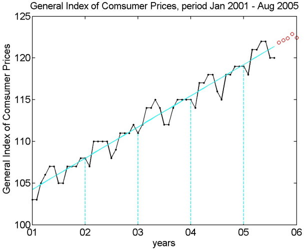


$x_t = y'_t - s_t = y_t - \mu_t - s_t$

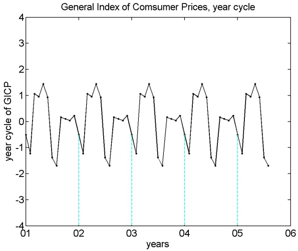


Example: Monthly Greek General Index of Consumer Prices

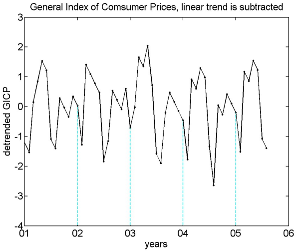
$\{y_t\}_{t=1}^{56} \quad \mu_t = 103.9 + 0.31t$



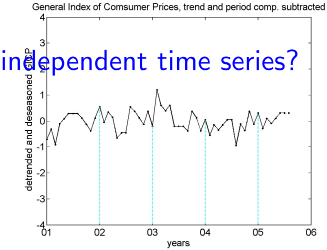
$\{s_t\}_{t=1}^{12}$



$\ldots \quad y'_t = y_t - \mu_t$



$x_t = y'_t - s_t = y_t - \mu_t - s_t$

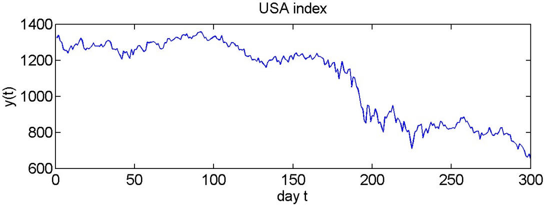


Removing stochastic trends

non-stationary time series with stochastic trends $\{y_t\}$

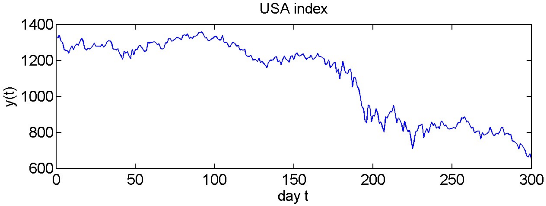
Removing stochastic trends

non-stationary time series with stochastic trends $\{y_t\}$



Removing stochastic trends

non-stationary time series with stochastic trends $\{y_t\}$



First differences:

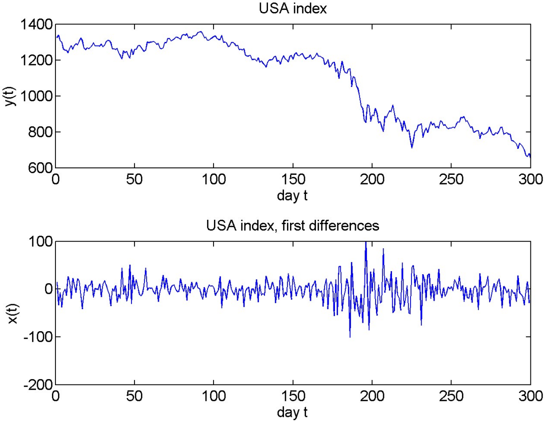
$x_t = y_t - y_{t-1}$

Removing stochastic trends

non-stationary time series with stochastic trends $\{y_t\}$

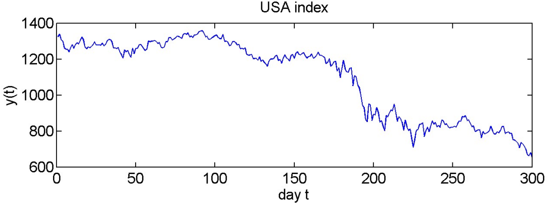
First differences:

$x_t = y_t - y_{t-1}$

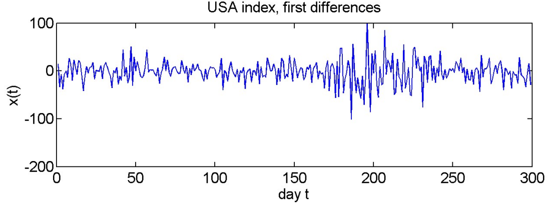


Removing stochastic trends

non-stationary time series with stochastic trends $\{y_t\}$



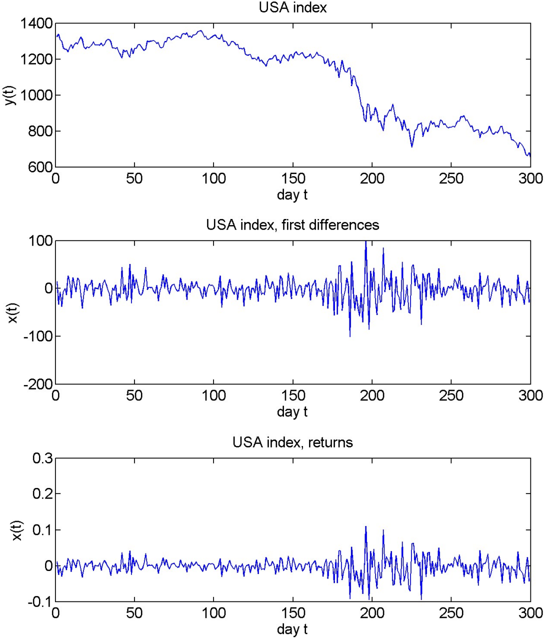
First differences:
$x_t = y_t - y_{t-1}$



Returns:
$x_t = \log(y_t) - \log(y_{t-1})$

Removing stochastic trends

non-stationary time series with stochastic trends $\{y_t\}$



First differences:

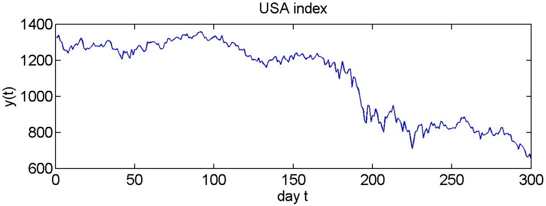
$x_t = y_t - y_{t-1}$

Returns:

$x_t = \log(y_t) - \log(y_{t-1})$

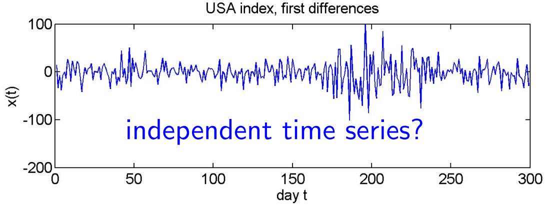
Removing stochastic trends

non-stationary time series with stochastic trends $\{y_t\}$



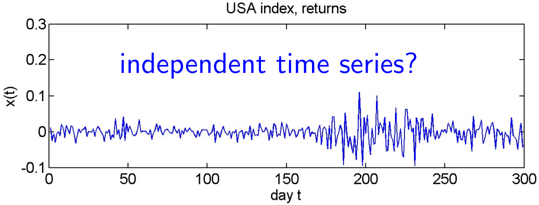
First differences:

$x_t = y_t - y_{t-1}$



Returns:

$x_t = \log(y_t) - \log(y_{t-1})$



Autocovariance and autocorrelation see [1]: Sec 2.7, [2]: Chp 1

Stationary process $\{X_t\}$ and its realization, time series $\{x_t\}_{t=1}^{n}$.

Autocovariance and autocorrelation see [1]: Sec 2.7, [2]: Chp 1

Stationary process $\{X_t\}$ and its realization, time series $\{x_t\}_{t=1}^{n}$.

mean: $\mu_X = \mathrm{E}[X_t]$,

Autocovariance and autocorrelation see [1]: Sec 2.7, [2]: Chp 1

Stationary process $\{X_t\}$ and its realization, time series $\{x_t\}_{t=1}^n$.

mean: $\mu_X = \mathrm{E}[X_t]$, and estimate from $\{x_t\}_{t=1}^n$, $\bar{x} = \frac{1}{n}\sum_{t=1}^n x_t$

Autocovariance and autocorrelation see [1]: Sec 2.7, [2]: Chp 1

Stationary process $\{X_t\}$ and its realization, time series $\{x_t\}_{t=1}^n$.

mean: $\mu_X = \mathrm{E}[X_t]$, and estimate from $\{x_t\}_{t=1}^n$, $\bar{x} = \frac{1}{n}\sum_{t=1}^n x_t$

variance: $\sigma_X^2 = \mathrm{Var}[X_t] = \mathrm{E}[(X_t - \mu_X)^2]$,

Stationary process $\{X_t\}$ and its realization, time series $\{x_t\}_{t=1}^{n}$.

mean: $\mu_X = \mathrm{E}[X_t]$, and estimate from $\{x_t\}_{t=1}^{n}$, $\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$

variance: $\sigma_X^2 = \mathrm{Var}[X_t] = \mathrm{E}[(X_t - \mu_X)^2]$,
and estimate

$$s_X^2 = \frac{1}{n-1}\sum_{t=1}^{n}(x_t - \bar{x})^2$$

Stationary process $\{X_t\}$ and its realization, time series $\{x_t\}_{t=1}^{n}$.

mean: $\mu_X = \mathrm{E}[X_t]$, and estimate from $\{x_t\}_{t=1}^{n}$, $\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$

variance: $\sigma_X^2 = \mathrm{Var}[X_t] = \mathrm{E}[(X_t - \mu_X)^2]$,
and estimate

$$s_X^2 = \frac{1}{n-1}\sum_{t=1}^{n}(x_t - \bar{x})^2$$

autocovariance: $\gamma_X(\tau) = \mathrm{Cov}[X_t, X_{t+\tau}] = \mathrm{E}[(X_t - \mu_X)(X_{t+\tau} - \mu_X)]$,

Autocovariance and autocorrelation see [1]: Sec 2.7, [2]: Chp 1

Stationary process $\{X_t\}$ and its realization, time series $\{x_t\}_{t=1}^{n}$.

mean: $\mu_X = \mathrm{E}[X_t]$, and estimate from $\{x_t\}_{t=1}^{n}$, $\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$

variance: $\sigma_X^2 = \mathrm{Var}[X_t] = \mathrm{E}[(X_t - \mu_X)^2]$,
and estimate

$$s_X^2 = \frac{1}{n-1}\sum_{t=1}^{n}(x_t - \bar{x})^2$$

autocovariance: $\gamma_X(\tau) = \mathrm{Cov}[X_t, X_{t+\tau}] = \mathrm{E}[(X_t - \mu_X)(X_{t+\tau} - \mu_X)]$,
and estimate

$$c_X(\tau) = \hat{\gamma}_X(\tau) = \frac{1}{n-\tau}\sum_{t=1}^{n-\tau}(x_t - \bar{x})(x_{t+\tau} - \bar{x})$$

Stationary process $\{X_t\}$ and its realization, time series $\{x_t\}_{t=1}^n$.

mean: $\mu_X = \mathrm{E}[X_t]$, and estimate from $\{x_t\}_{t=1}^n$, $\bar{x} = \frac{1}{n}\sum_{t=1}^n x_t$

variance: $\sigma_X^2 = \mathrm{Var}[X_t] = \mathrm{E}[(X_t - \mu_X)^2]$,
and estimate

$$s_X^2 = \frac{1}{n-1}\sum_{t=1}^{n}(x_t - \bar{x})^2$$

autocovariance: $\gamma_X(\tau) = \mathrm{Cov}[X_t, X_{t+\tau}] = \mathrm{E}[(X_t - \mu_X)(X_{t+\tau} - \mu_X)]$,
and estimate

$$c_X(\tau) = \hat{\gamma}_X(\tau) = \frac{1}{n-\tau}\sum_{t=1}^{n-\tau}(x_t - \bar{x})(x_{t+\tau} - \bar{x})$$

autocorrelation: $\rho_X(\tau) = \frac{\gamma_X(\tau)}{\gamma_X(0)} = \frac{\gamma_X(\tau)}{\sigma_X^2}$, Stationary process $\{X_t\}$ and its realization, time series $\{x_t\}_{t=1}^n$.

mean: $\mu_X = \mathrm{E}[X_t]$, and estimate from $\{x_t\}_{t=1}^n$, $\bar{x} = \frac{1}{n}\sum_{t=1}^n x_t$

variance: $\sigma_X^2 = \mathrm{Var}[X_t] = \mathrm{E}[(X_t - \mu_X)^2]$,
and estimate

$$s_X^2 = \frac{1}{n-1}\sum_{t=1}^{n}(x_t - \bar{x})^2$$

autocovariance: $\gamma_X(\tau) = \mathrm{Cov}[X_t, X_{t+\tau}] = \mathrm{E}[(X_t - \mu_X)(X_{t+\tau} - \mu_X)]$,
and estimate

$$c_X(\tau) = \hat{\gamma}_X(\tau) = \frac{1}{n-\tau}\sum_{t=1}^{n-\tau}(x_t - \bar{x})(x_{t+\tau} - \bar{x})$$

autocorrelation: $\rho_X(\tau) = \frac{\gamma_X(\tau)}{\gamma_X(0)} = \frac{\gamma_X(\tau)}{\sigma_X^2}$, and estimate

$$r_X(\tau) = \hat{\rho}_X(\tau) = \frac{c_X(\tau)}{c_X(0)} = \frac{c_X(\tau)}{s_X^2}$$

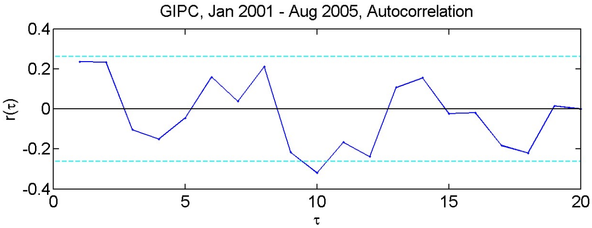
Examples of autocorrelation

Examples of autocorrelation

residuals of GICP

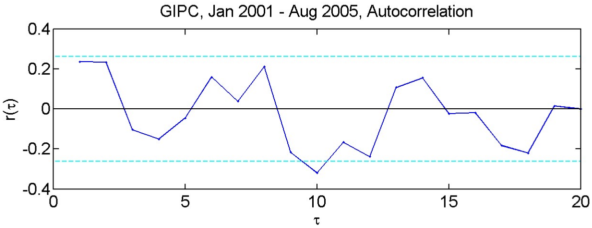
Examples of autocorrelation

residuals of GICP



Examples of autocorrelation

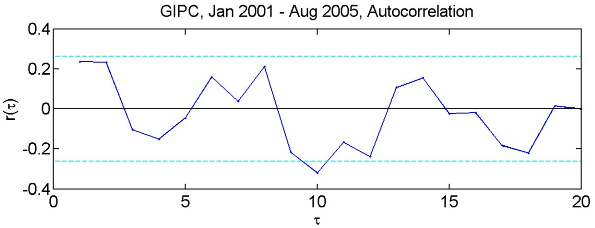
residuals of GICP



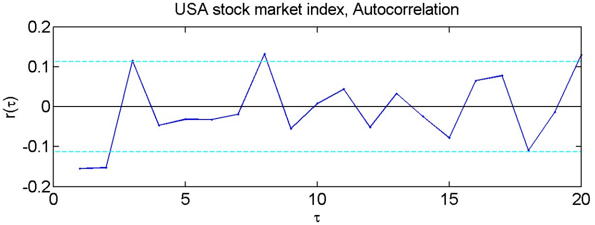
returns of USA stock market index

Examples of autocorrelation

residuals of GICP

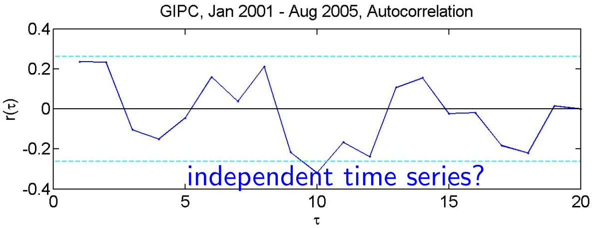


returns of USA stock market index

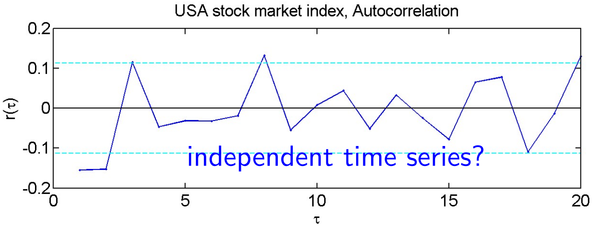


Examples of autocorrelation

residuals of GICP



returns of USA stock market index



Example: Traces of packet arrivals

A trace containing packet arrivals seen on an Ethernet at the Bellcore Morristown Research and Engineering facility, regarding LAN traffic, period 11:25 on August 29, 1989.

Example: Traces of packet arrivals

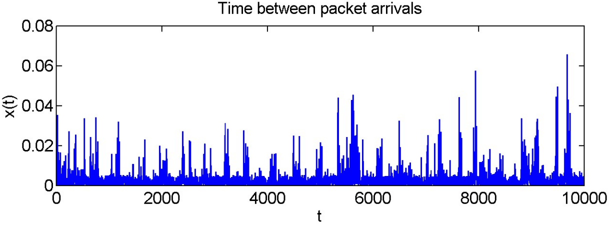
A trace containing packet arrivals seen on an Ethernet at the Bellcore Morristown Research and Engineering facility, regarding LAN traffic, period 11:25 on August 29, 1989.

The time series data are the times between successive arrivals.

Example: Traces of packet arrivals

A trace containing packet arrivals seen on an Ethernet at the Bellcore Morristown Research and Engineering facility, regarding LAN traffic, period 11:25 on August 29, 1989.

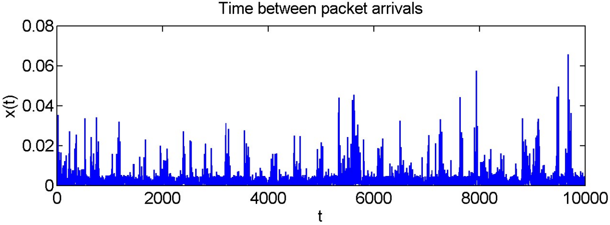
The time series data are the times between successive arrivals.



Example: Traces of packet arrivals

A trace containing packet arrivals seen on an Ethernet at the Bellcore Morristown Research and Engineering facility, regarding LAN traffic, period 11:25 on August 29, 1989.

The time series data are the times between successive arrivals.

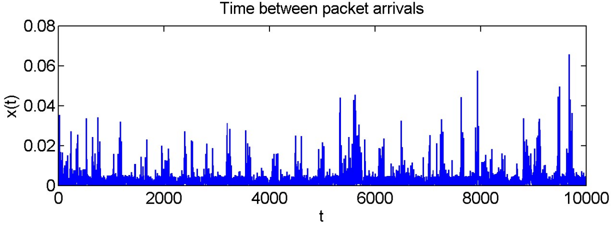


A small data window

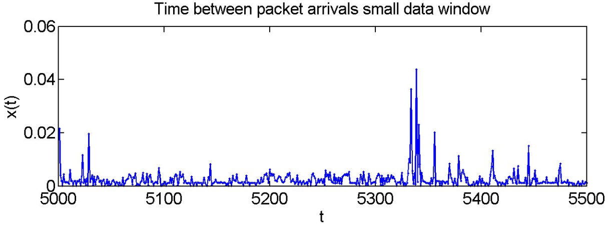
Example: Traces of packet arrivals

A trace containing packet arrivals seen on an Ethernet at the Bellcore Morristown Research and Engineering facility, regarding LAN traffic, period 11:25 on August 29, 1989.

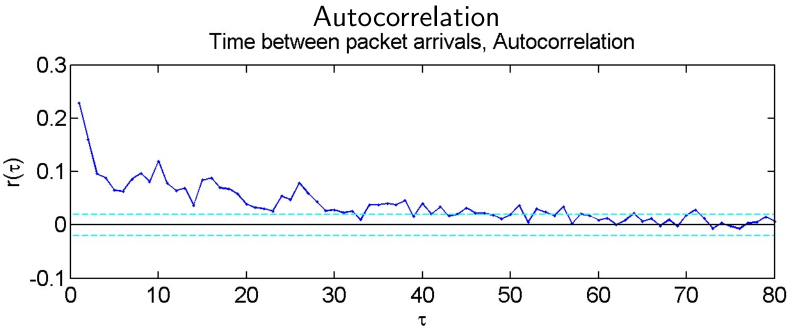
The time series data are the times between successive arrivals.



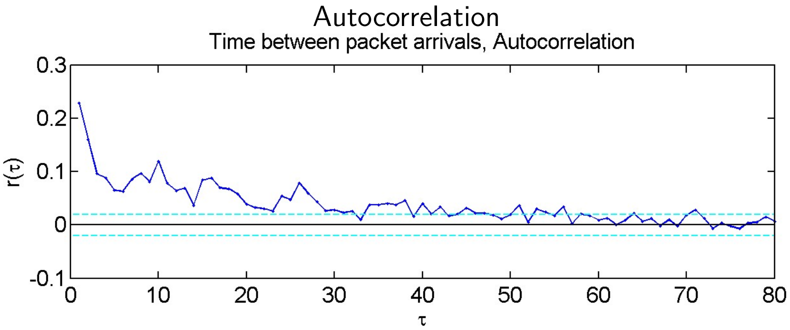
A small data window



Example: Traces of packet arrivals

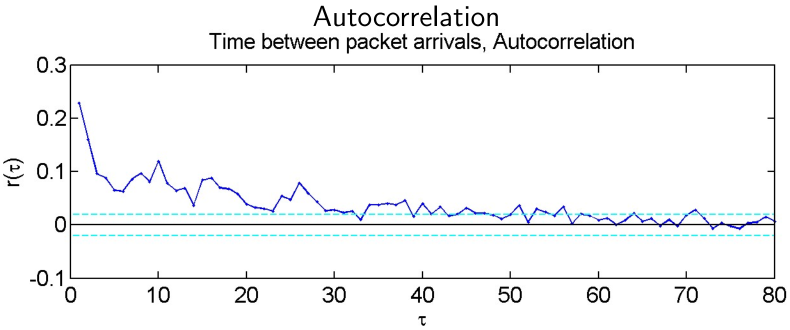


Example: Traces of packet arrivals



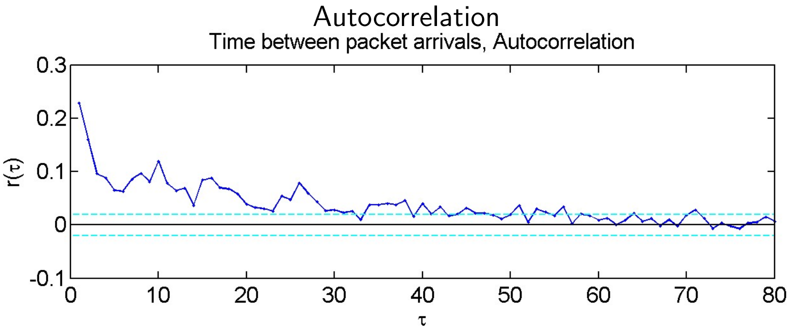
- Are the residuals of GIPC correlated?

Example: Traces of packet arrivals

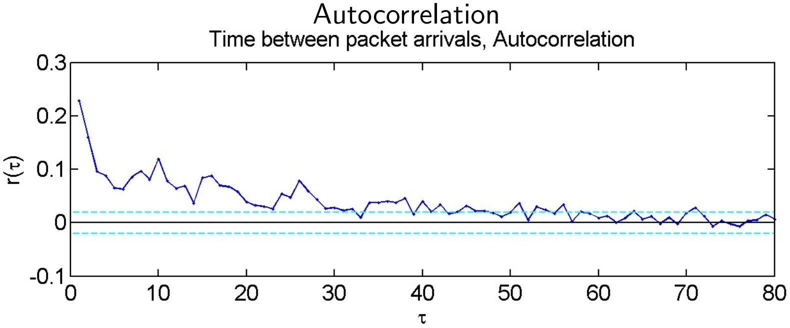


- Are the residuals of GIPC correlated?
- Are the returns of USA stock marker index correlated?

Example: Traces of packet arrivals



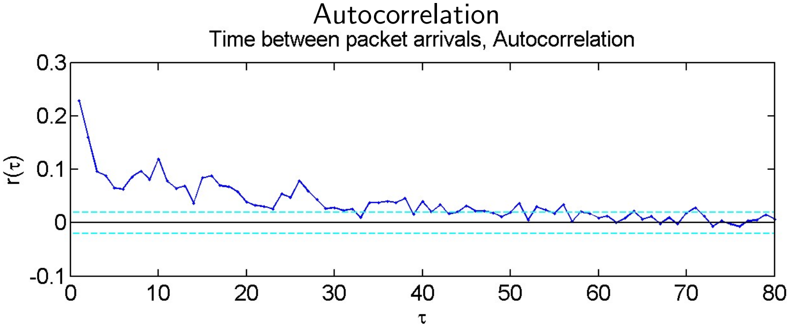
- Are the residuals of GIPC correlated?
- Are the returns of USA stock marker index correlated?
- Are the times between packet arrivals correlated?



- Are the residuals of GIPC correlated?
- Are the returns of USA stock marker index correlated?
- Are the times between packet arrivals correlated?

Are these autocorrelations statistically significant?

Example: Traces of packet arrivals



- Are the residuals of GIPC correlated?
- Are the returns of USA stock marker index correlated?
- Are the times between packet arrivals correlated?

Are these autocorrelations statistically significant?

Are the time series independent?

Statistical test for independence see [1]: Sec 2.7-2.8, [2]: Chp 1

Completely random time series: it consists of a series of independent observations having the same distribution.

Statistical test for independence see [1]: Sec 2.7-2.8, [2]: Chp 1

Completely random time series: it consists of a series of independent observations having the same distribution.

For large n and random time series, $r(\tau) \sim \mathrm{N}(0, 1/n)$.

Completely random time series: it consists of a series of independent observations having the same distribution.

For large n and random time series, $r(\tau) \sim \mathrm{N}(0, 1/n)$.

$r(\tau)$ is significant if it is outside the limits $\pm z_{\alpha/2}\sqrt{1/n}$ (for $\alpha = 0.05$, $\pm 2\sqrt{1/n}$).

Statistical test for independence see [1]: Sec 2.7-2.8, [2]: Chp 1

Completely random time series: it consists of a series of independent observations having the same distribution.

For large n and random time series, $r(\tau) \sim \mathrm{N}(0, 1/n)$.

$r(\tau)$ is significant if it is outside the limits $\pm z_{\alpha/2}\sqrt{1/n}$ (for $\alpha = 0.05$, $\pm 2\sqrt{1/n}$).

... but for random time series and $\tau = 1, \ldots, K$, where say $K = 20$, on average one $r(\tau)$ will be outside the limits

Statistical test for independence see [1]: Sec 2.7-2.8, [2]: Chp 1

Completely random time series: it consists of a series of independent observations having the same distribution.

For large n and random time series, $r(\tau) \sim \mathrm{N}(0, 1/n)$.

$r(\tau)$ is significant if it is outside the limits $\pm z_{\alpha/2}\sqrt{1/n}$
(for $\alpha = 0.05$, $\pm 2\sqrt{1/n}$).

... but for random time series and $\tau = 1, \ldots, K$, where say $K = 20$, on average one $r(\tau)$ will be outside the limits

$\Longrightarrow$ Portmanteau test collecting all $r(\tau)$ (as modified by Ljung and Box):

Statistical test for independence see [1]: Sec 2.7-2.8, [2]: Chp 1

Completely random time series: it consists of a series of independent observations having the same distribution.

For large n and random time series, $r(\tau) \sim \mathrm{N}(0, 1/n)$.

$r(\tau)$ is significant if it is outside the limits $\pm z_{\alpha/2}\sqrt{1/n}$ (for $\alpha = 0.05$, $\pm 2\sqrt{1/n}$).

... but for random time series and $\tau = 1, \ldots, K$, where say $K = 20$, on average one $r(\tau)$ will be outside the limits

$\Longrightarrow$ Portmanteau test collecting all $r(\tau)$ (as modified by Ljung and Box):

1. statistic $Q = n(n+2)\sum_{\tau=1}^{K} \frac{r^2(\tau)}{n-\tau}$.

Statistical test for independence see [1]: Sec 2.7-2.8, [2]: Chp 1

Completely random time series: it consists of a series of independent observations having the same distribution.

For large n and random time series, $r(\tau) \sim \mathrm{N}(0, 1/n)$.

$r(\tau)$ is significant if it is outside the limits $\pm z_{\alpha/2}\sqrt{1/n}$ (for $\alpha = 0.05$, $\pm 2\sqrt{1/n}$).

... but for random time series and $\tau = 1, \ldots, K$, where say $K = 20$, on average one $r(\tau)$ will be outside the limits

$\Longrightarrow$ Portmanteau test collecting all $r(\tau)$ (as modified by Ljung and Box):

1. statistic $Q = n(n+2)\sum_{\tau=1}^{K} \frac{r^2(\tau)}{n-\tau}$.
2. $Q \sim \mathcal{X}_K^2$.

Statistical test for independence see [1]: Sec 2.7-2.8, [2]: Chp 1

Completely random time series: it consists of a series of independent observations having the same distribution.

For large n and random time series, $r(\tau) \sim \mathrm{N}(0, 1/n)$.

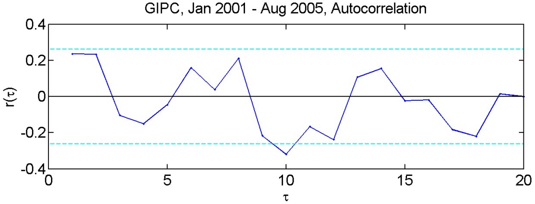
$r(\tau)$ is significant if it is outside the limits $\pm z_{\alpha/2}\sqrt{1/n}$ (for $\alpha = 0.05$, $\pm 2\sqrt{1/n}$).

... but for random time series and $\tau = 1, \ldots, K$, where say $K = 20$, on average one $r(\tau)$ will be outside the limits

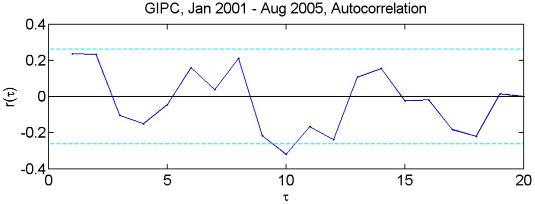
$\Longrightarrow$ Portmanteau test collecting all $r(\tau)$ (as modified by Ljung and Box):

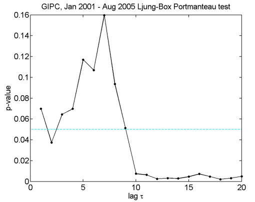
1. statistic $Q = n(n+2)\sum_{\tau=1}^{K} \frac{r^2(\tau)}{n-\tau}$.
2. $Q \sim \mathcal{X}_K^2$.
3. Reject null hypothesis of independence at significance level α if $Q > \mathcal{X}_{K,1-\alpha}^2$ (or compute the corresponding p-value)

Examples of test for independence

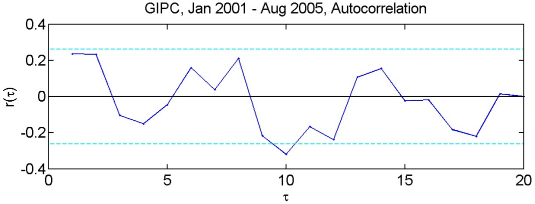


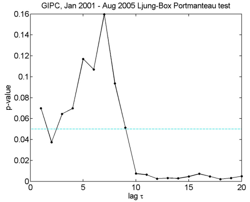
Examples of test for independence

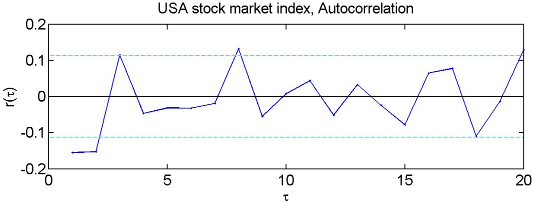




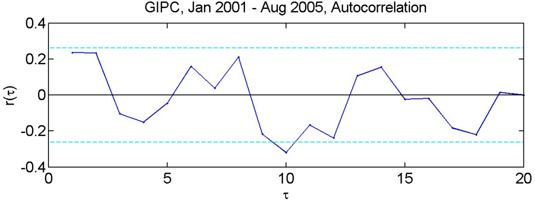
Examples of test for independence

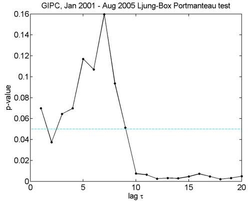


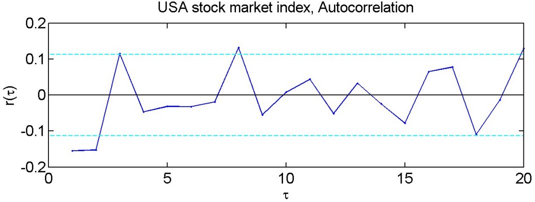


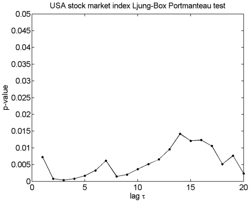


Examples of test for independence

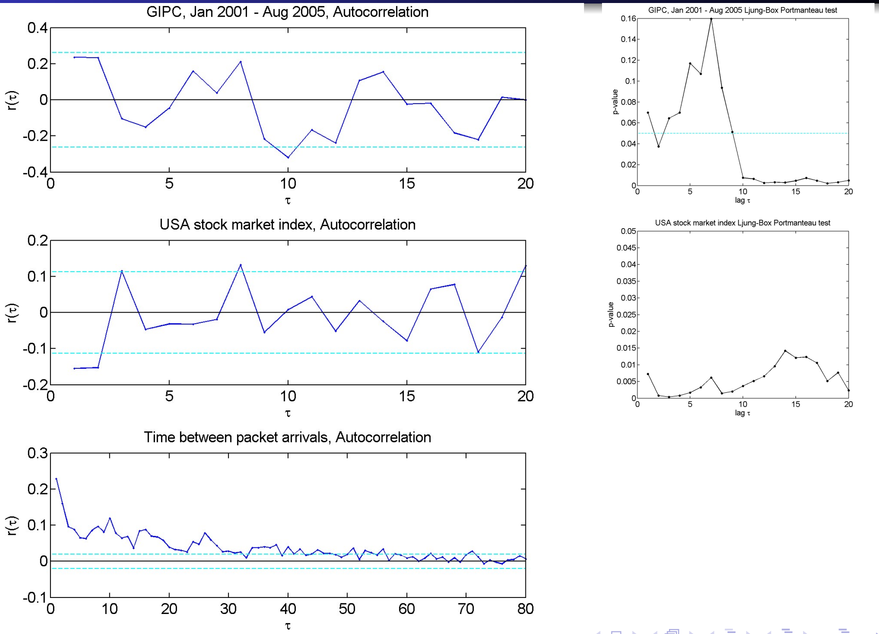




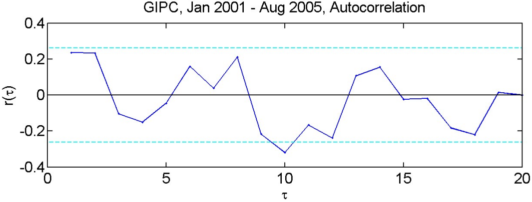


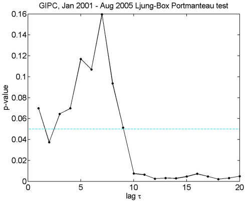


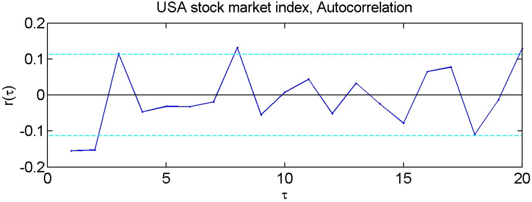
Examples of test for independence

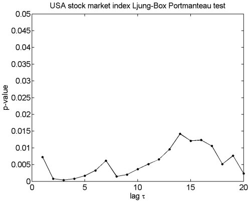


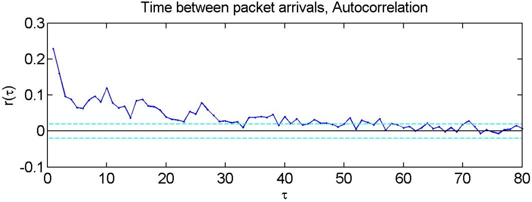
Examples of test for independence

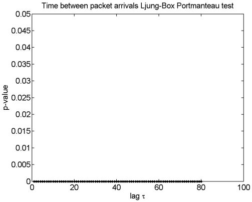












White noise and random walk see [1]: Chp 3, [2]: Chp 1

Two main stochastic processes:

- White noise process: Independent stochastic process, a series of iid variables X_t.

White noise and random walk see [1]: Chp 3, [2]: Chp 1

Two main stochastic processes:

- White noise process: Independent stochastic process, a series of iid variables X_t.

  $X_t \sim \mathrm{WN}(\mu_X, \sigma_X^2)$: white noise with mean μ_X and variance σ_X^2.

White noise and random walk see [1]: Chp 3, [2]: Chp 1

Two main stochastic processes:

- White noise process: Independent stochastic process, a series of iid variables X_t.

  $X_t \sim \text{WN}(\mu_X, \sigma_X^2)$: white noise with mean μ_X and variance σ_X^2.

  Realization of white noise process: a series of iid observations.

Two main stochastic processes:

- White noise process: Independent stochastic process, a series of iid variables X_t.

  $X_t \sim \text{WN}(\mu_X, \sigma_X^2)$: white noise with mean μ_X and variance σ_X^2.

  Realization of white noise process: a series of iid observations.

- Random walk process: at each step a white noise increment is added

$$Y_t = Y_{t-1} + X_t, \qquad X_t \sim \text{WN}(0, \sigma_X^2)$$

  $\mu_Y = \text{E}[Y_t] = 0$ and $\sigma_Y^2 = \text{Var}[Y_t] = t\sigma_X^2$. The variance grows with time.

White noise and random walk see [1]: Chp 3, [2]: Chp 1

Two main stochastic processes:

- White noise process: Independent stochastic process, a series of iid variables X_t.

  $X_t \sim \mathrm{WN}(\mu_X, \sigma_X^2)$: white noise with mean μ_X and variance σ_X^2.

  Realization of white noise process: a series of iid observations.

- Random walk process: at each step a white noise increment is added

$$Y_t = Y_{t-1} + X_t, \qquad X_t \sim \mathrm{WN}(0, \sigma_X^2)$$

  $\mu_Y = \mathrm{E}[Y_t] = 0$ and $\sigma_Y^2 = \mathrm{Var}[Y_t] = t\sigma_X^2$. The variance grows with time.

White noise is a stationary process and random walk a non-stationary process.

Linear autoregressive models see [1]: Chp 3, [2]: Chp 2

A linear stochastic process possessing significant autocorrelation is the linear autoregressive process of order p, AR(p)

$$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \epsilon_t$$

where $\epsilon_t \sim \mathrm{WN}(0, \sigma_\epsilon^2)$ (usually we assume for simplicity $\phi_0 = 0$).

Linear autoregressive models see [1]: Chp 3, [2]: Chp 2

A linear stochastic process possessing significant autocorrelation is the linear autoregressive process of order p, AR(p)

$$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \epsilon_t$$

where $\epsilon_t \sim \text{WN}(0, \sigma_\epsilon^2)$ (usually we assume for simplicity $\phi_0 = 0$).
Random walk is an AR(1) process with $\phi_1 = 1$.

A linear stochastic process possessing significant autocorrelation is the linear autoregressive process of order p, AR(p)

$$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \epsilon_t$$

where $\epsilon_t \sim \text{WN}(0, \sigma_\epsilon^2)$ (usually we assume for simplicity $\phi_0 = 0$).

Random walk is an AR(1) process with $\phi_1 = 1$.

The coefficients are such that the AR(p) process is stationary.

A linear stochastic process possessing significant autocorrelation is the linear autoregressive process of order p, AR(p)

$$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \epsilon_t$$

where $\epsilon_t \sim \text{WN}(0, \sigma_\epsilon^2)$ (usually we assume for simplicity $\phi_0 = 0$).

Random walk is an AR(1) process with $\phi_1 = 1$.

The coefficients are such that the AR(p) process is stationary.

Stationarity condition: the roots of the equation

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p = 0$$

must lie outside the unit circle (roots, which may be complex, are greater than one in modulus).

Linear autoregressive models see [1]: Chp 3, [2]: Chp 2

A linear stochastic process possessing significant autocorrelation is the linear autoregressive process of order p, AR(p)

$$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \epsilon_t$$

where $\epsilon_t \sim \mathrm{WN}(0, \sigma_\epsilon^2)$ (usually we assume for simplicity $\phi_0 = 0$).

Random walk is an AR(1) process with $\phi_1 = 1$.

The coefficients are such that the AR(p) process is stationary.

Stationarity condition: the roots of the equation

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p = 0$$

must lie outside the unit circle (roots, which may be complex, are greater than one in modulus).

Other types of processes: moving average of order q, MA(q), mixed processes ARMA(p, q).

A linear stochastic process possessing significant autocorrelation is the linear autoregressive process of order p, AR(p)

$$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \epsilon_t$$

where $\epsilon_t \sim \text{WN}(0, \sigma_\epsilon^2)$ (usually we assume for simplicity $\phi_0 = 0$).

Random walk is an AR(1) process with $\phi_1 = 1$.

The coefficients are such that the AR(p) process is stationary.

Stationarity condition: the roots of the equation

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p = 0$$

must lie outside the unit circle (roots, which may be complex, are greater than one in modulus).

Other types of processes: moving average of order q, MA(q), mixed processes ARMA(p, q).

To estimate the model AR(p) from a time series, we use least squares to compute the coefficients $\phi_0, \phi_1, \ldots \phi_p$ and σ_ϵ^2.

Exercise 4: Nonparametric test for independence

What if the Q statistic does not follow exactly $\mathcal{X}_K^2$?

Exercise 4: Nonparametric test for independence

What if the Q statistic does not follow exactly $\mathcal{X}_K^2$?
Use resampling (randomization) to form the empirical distribution of Q:
Generate M randomized time series from $\{x_t\}_{t=1}^n$ by random permutation of the samples (destroy the time order but use the same distribution of the original time series). [`matlab: use randperm`]

Exercise 4: Nonparametric test for independence

What if the Q statistic does not follow exactly $\mathcal{X}_K^2$?
Use resampling (randomization) to form the empirical distribution of Q:
Generate M randomized time series from $\{x_t\}_{t=1}^n$ by random permutation of the samples (destroy the time order but use the same distribution of the original time series). [`matlab: use randperm`]

1. Generate a time series $\{x_t\}_{t=1}^n$, $n = 100$.

Exercise 4: Nonparametric test for independence

What if the Q statistic does not follow exactly $\mathcal{X}_K^2$?

Use resampling (randomization) to form the empirical distribution of Q:

Generate M randomized time series from $\{x_t\}_{t=1}^n$ by random permutation of the samples (destroy the time order but use the same distribution of the original time series). [`matlab: use randperm`]

1. Generate a time series $\{x_t\}_{t=1}^n$, $n = 100$.
2. Use parametric and nonparametric Portmanteau test (e.g. for $K = 5$, $M = 1000$) [`matlab: use portmanteauLB.m` from the course files].

Exercise 4: Nonparametric test for independence

What if the Q statistic does not follow exactly $\mathcal{X}_K^2$?

Use resampling (randomization) to form the empirical distribution of Q:

Generate M randomized time series from $\{x_t\}_{t=1}^n$ by random permutation of the samples (destroy the time order but use the same distribution of the original time series). [`matlab: use randperm`]

1. Generate a time series $\{x_t\}_{t=1}^n$, $n = 100$.
2. Use parametric and nonparametric Portmanteau test (e.g. for $K = 5$, $M = 1000$) [`matlab: use portmanteauLB.m` from the course files].
3. Repeat the tests 100 times. Are the proportions of rejection the same for the two test types?

Exercise 4: Nonparametric test for independence

What if the Q statistic does not follow exactly $\mathcal{X}_K^2$?
Use resampling (randomization) to form the empirical distribution of Q:
Generate M randomized time series from $\{x_t\}_{t=1}^n$ by random permutation of the samples (destroy the time order but use the same distribution of the original time series). [`matlab: use randperm`]

1. Generate a time series $\{x_t\}_{t=1}^n$, $n = 100$.
2. Use parametric and nonparametric Portmanteau test (e.g. for $K = 5$, $M = 1000$) [`matlab: use portmanteauLB.m` from the course files].
3. Repeat the tests 100 times. Are the proportions of rejection the same for the two test types?

The following types of time series will be generated in (1):

1. White noise with normal distribution [`matlab: use randn`]

Exercise 4: Nonparametric test for independence

What if the Q statistic does not follow exactly $\mathcal{X}_K^2$?
Use resampling (randomization) to form the empirical distribution of Q:
Generate M randomized time series from $\{x_t\}_{t=1}^n$ by random permutation of the samples (destroy the time order but use the same distribution of the original time series). [`matlab: use randperm`]

1. Generate a time series $\{x_t\}_{t=1}^n$, $n = 100$.
2. Use parametric and nonparametric Portmanteau test (e.g. for $K = 5$, $M = 1000$) [`matlab: use portmanteauLB.m` from the course files].
3. Repeat the tests 100 times. Are the proportions of rejection the same for the two test types?

The following types of time series will be generated in (1):

1. White noise with normal distribution [`matlab: use randn`]
2. White noise with log-normal distribution [`matlab: use lognrnd`].

Exercise 4: Nonparametric test for independence

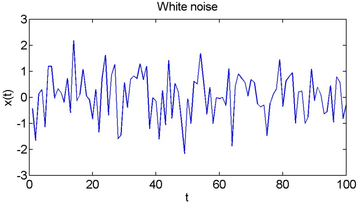
What if the Q statistic does not follow exactly $\mathcal{X}_K^2$?
Use resampling (randomization) to form the empirical distribution of Q:
Generate M randomized time series from $\{x_t\}_{t=1}^n$ by random permutation of the samples (destroy the time order but use the same distribution of the original time series). [`matlab: use randperm`]

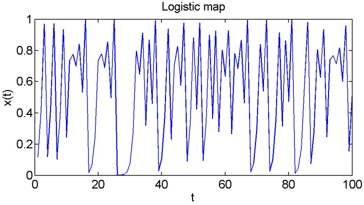
1. Generate a time series $\{x_t\}_{t=1}^n$, $n = 100$.
2. Use parametric and nonparametric Portmanteau test (e.g. for $K = 5$, $M = 1000$) [`matlab: use portmanteauLB.m` from the course files].
3. Repeat the tests 100 times. Are the proportions of rejection the same for the two test types?

The following types of time series will be generated in (1):

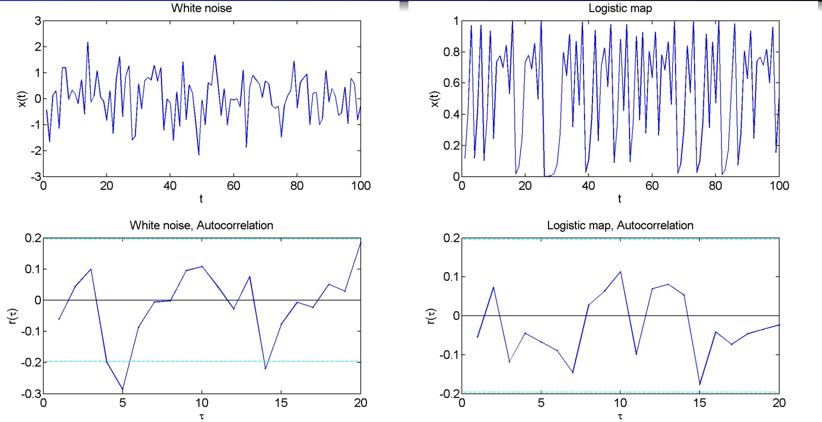
1. White noise with normal distribution [`matlab: use randn`]
2. White noise with log-normal distribution [`matlab: use lognrnd`].
3. AR(1) with $\phi_0 = 0$ and $\phi_1 = 0.2, 0.4$ and 0.6 [`matlab: use ARm.m` from the course files]

Nonlinear correlations see [3], Chp 1,2,3

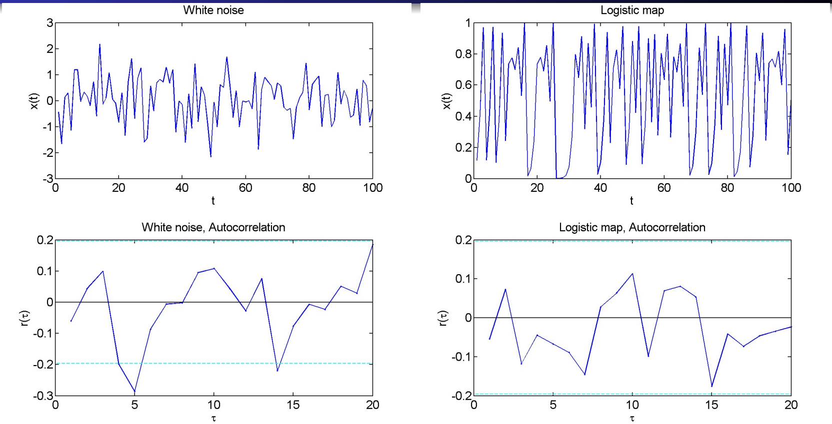




Nonlinear correlations see [3], Chp 1,2,3

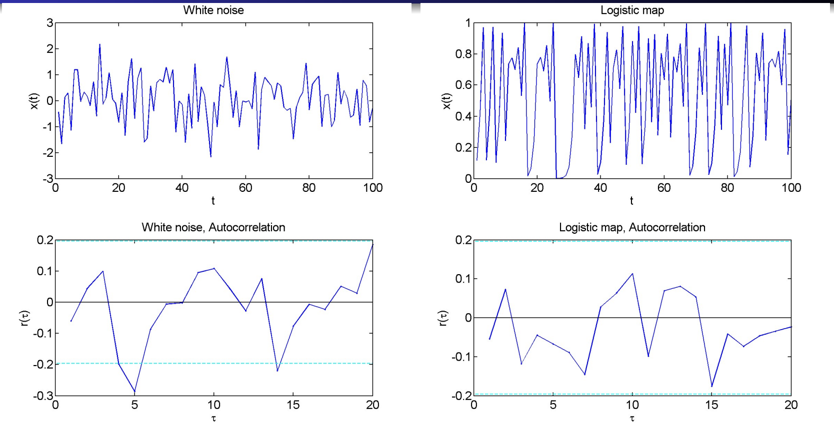


Nonlinear correlations see [3], Chp 1,2,3



Logistic map:
$X_t = 4X_{t-1}(1 - X_{t-1})$,
nonlinear dynamical system

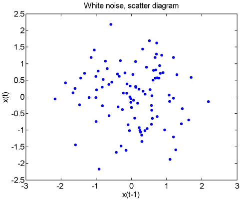
Nonlinear correlations see [3], Chp 1,2,3

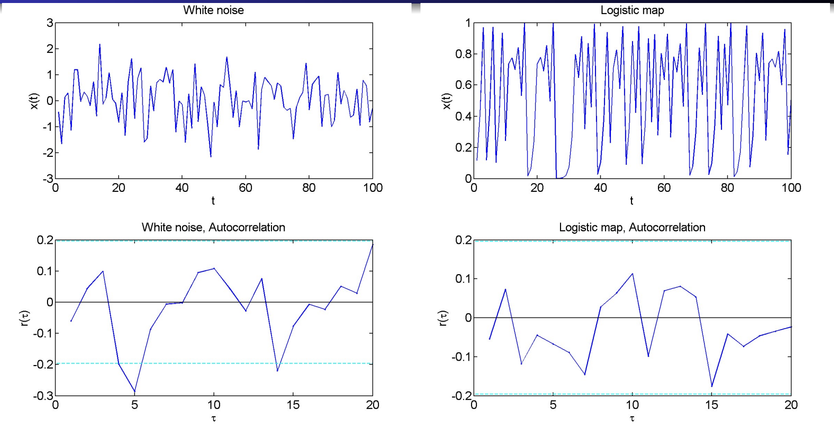


Logistic map:

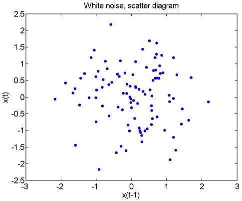
$X_t = 4X_{t-1}(1 - X_{t-1})$,

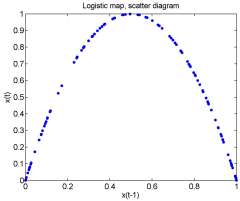
nonlinear dynamical system





Logistic map:
$X_t = 4X_{t-1}(1 - X_{t-1})$,
nonlinear dynamical system





Measure of nonlinear autocorrelation see [3]: Sec 9.2

Autocovariance: $\gamma_X(\tau) = \text{Cov}[X_t, X_{t+\tau}] = \text{E}[(X_t - \mu_X)(X_{t+\tau} - \mu_X)]$,

Measure of nonlinear autocorrelation see [3]: Sec 9.2

Autocovariance: $\gamma_X(\tau) = \text{Cov}[X_t, X_{t+\tau}] = \text{E}[(X_t - \mu_X)(X_{t+\tau} - \mu_X)]$,

Extend joint moment of order one $\text{E}[X_t X_{t+\tau}]$ to higher order joint moments

Measure of nonlinear autocorrelation see [3]: Sec 9.2

Autocovariance: $\gamma_X(\tau) = \text{Cov}[X_t, X_{t+\tau}] = \text{E}[(X_t - \mu_X)(X_{t+\tau} - \mu_X)]$,
Extend joint moment of order one $\text{E}[X_t X_{t+\tau}]$ to higher order joint moments
$\Longrightarrow$ nonlinear measures.

Measure of nonlinear autocorrelation see [3]: Sec 9.2

Autocovariance: $\gamma_X(\tau) = \text{Cov}[X_t, X_{t+\tau}] = \text{E}[(X_t - \mu_X)(X_{t+\tau} - \mu_X)]$,
Extend joint moment of order one $\text{E}[X_t X_{t+\tau}]$ to higher order joint moments $\Longrightarrow$ nonlinear measures.

Entropy: information from each sample of X (assuming discrete X)

$$\text{H}(X) = \text{E}[\log p_X(x)] = \sum_x p_X(x) \log p_X(x).$$

Measure of nonlinear autocorrelation see [3]: Sec 9.2

Autocovariance: $\gamma_X(\tau) = \mathrm{Cov}[X_t, X_{t+\tau}] = \mathrm{E}[(X_t - \mu_X)(X_{t+\tau} - \mu_X)]$,
Extend joint moment of order one $\mathrm{E}[X_t X_{t+\tau}]$ to higher order joint moments
$\Longrightarrow$ nonlinear measures.

Entropy: information from each sample of X (assuming discrete X)

$$\mathrm{H}(X) = \mathrm{E}[\log p_X(x)] = \sum_x p_X(x) \log p_X(x).$$

Mutual information: information for Y knowing X and vice versa

$$\mathrm{I}(X,Y) = \mathrm{H}(X) + \mathrm{H}(Y) - \mathrm{H}(X,Y) = \sum_{x,y} p_{XY}(x,y) \log \frac{p_{XY}(x,y)}{p_X(x) p_Y(y)}$$

Measure of nonlinear autocorrelation see [3]: Sec 9.2

Autocovariance: $\gamma_X(\tau) = \mathrm{Cov}[X_t, X_{t+\tau}] = \mathrm{E}[(X_t - \mu_X)(X_{t+\tau} - \mu_X)]$,
Extend joint moment of order one $\mathrm{E}[X_t X_{t+\tau}]$ to higher order joint moments $\Longrightarrow$ nonlinear measures.

Entropy: information from each sample of X (assuming discrete X)

$$\mathrm{H}(X) = \mathrm{E}[\log p_X(x)] = \sum_x p_X(x) \log p_X(x).$$

Mutual information: information for Y knowing X and vice versa

$$\mathrm{I}(X,Y) = \mathrm{H}(X) + \mathrm{H}(Y) - \mathrm{H}(X,Y) = \sum_{x,y} p_{XY}(x,y) \log \frac{p_{XY}(x,y)}{p_X(x) p_Y(y)}$$

For $X \to X_t$ and $Y \to X_{t+\tau}$, the delayed mutual information:

$$\mathrm{I}(\tau) = \mathrm{I}(X_t, X_{t+\tau}) = \sum_{x_t, x_{t+\tau}} p_{X_t X_{t+\tau}}(x_t, x_{t+\tau}) \log \frac{p_{X_t X_{t+\tau}}(x_t, x_{t+\tau})}{p_{X_t}(x_t) p_{X_{t+\tau}}(x_{t+\tau})}$$

Measure of nonlinear autocorrelation see [3]: Sec 9.2

Autocovariance: $\gamma_X(\tau) = \text{Cov}[X_t, X_{t+\tau}] = \text{E}[(X_t - \mu_X)(X_{t+\tau} - \mu_X)]$,
Extend joint moment of order one $\text{E}[X_t X_{t+\tau}]$ to higher order joint moments $\Longrightarrow$ nonlinear measures.

Entropy: information from each sample of X (assuming discrete X)

$$\text{H}(X) = \text{E}[\log p_X(x)] = \sum_x p_X(x) \log p_X(x).$$

Mutual information: information for Y knowing X and vice versa

$$\text{I}(X,Y) = \text{H}(X) + \text{H}(Y) - \text{H}(X,Y) = \sum_{x,y} p_{XY}(x,y) \log \frac{p_{XY}(x,y)}{p_X(x) p_Y(y)}$$

For $X \to X_t$ and $Y \to X_{t+\tau}$, the delayed mutual information:

$$\text{I}(\tau) = \text{I}(X_t, X_{t+\tau}) = \sum_{x_t, x_{t+\tau}} p_{X_t X_{t+\tau}}(x_t, x_{t+\tau}) \log \frac{p_{X_t X_{t+\tau}}(x_t, x_{t+\tau})}{p_{X_t}(x_t) p_{X_{t+\tau}}(x_{t+\tau})}$$

To estimate $\text{I}(\tau)$ make a partition of $\{x_t\}_{t=1}^n$ and compute probabilities for each cell from the relative frequency.

Mutual information: white noise and logistic map

Computation of $I(\tau)$:

Mutual information: white noise and logistic map

Computation of $I(\tau)$:

1. Equidistant partition (histogram): split $\{x_t\}_{t=1}^{n}$ to b equidistant intervals.

Mutual information: white noise and logistic map

Computation of $I(\tau)$:

1. Equidistant partition (histogram): split $\{x_t\}_{t=1}^{n}$ to b equidistant intervals.
2. Count x_t, $t = \tau + 1, \ldots, n$ in each interval. The same for $x_{t+\tau}$.

Mutual information: white noise and logistic map

Computation of $I(\tau)$:

1. Equidistant partition (histogram): split $\{x_t\}_{t=1}^{n}$ to b equidistant intervals.
2. Count x_t, $t = \tau + 1, \ldots, n$ in each interval. The same for $x_{t+\tau}$.
3. Count pairs $(x_t, x_{t+\tau})$, $t = 1, \ldots, n - \tau$ in each of the b^2 cells.

Mutual information: white noise and logistic map

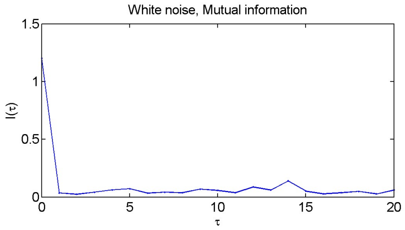
Computation of $I(\tau)$:

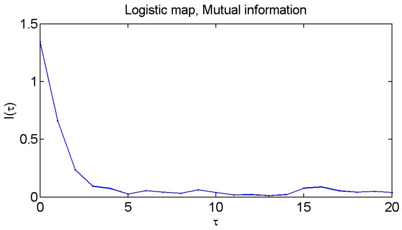
1. Equidistant partition (histogram): split $\{x_t\}_{t=1}^{n}$ to b equidistant intervals.
2. Count x_t, $t = \tau + 1, \ldots, n$ in each interval. The same for $x_{t+\tau}$.
3. Count pairs $(x_t, x_{t+\tau})$, $t = 1, \ldots, n - \tau$ in each of the b^2 cells.
4. The relative frequencies in 2 and 3 are the estimates for $p_{X_t}(x_t)$, $p_{X_{t+\tau}}(x_{t+\tau})$ and $p_{X_t X_{t+\tau}}(x_t, x_{t+\tau})$.

Mutual information: white noise and logistic map

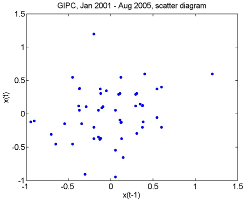
Computation of $I(\tau)$:

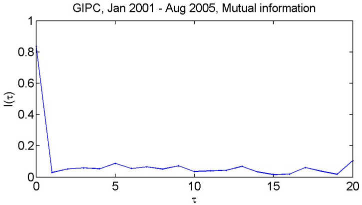
1. Equidistant partition (histogram): split $\{x_t\}_{t=1}^{n}$ to b equidistant intervals.
2. Count x_t, $t = \tau + 1, \ldots, n$ in each interval. The same for $x_{t+\tau}$.
3. Count pairs $(x_t, x_{t+\tau})$, $t = 1, \ldots, n - \tau$ in each of the b^2 cells.
4. The relative frequencies in 2 and 3 are the estimates for $p_{X_t}(x_t)$, $p_{X_{t+\tau}}(x_{t+\tau})$ and $p_{X_t X_{t+\tau}}(x_t, x_{t+\tau})$.



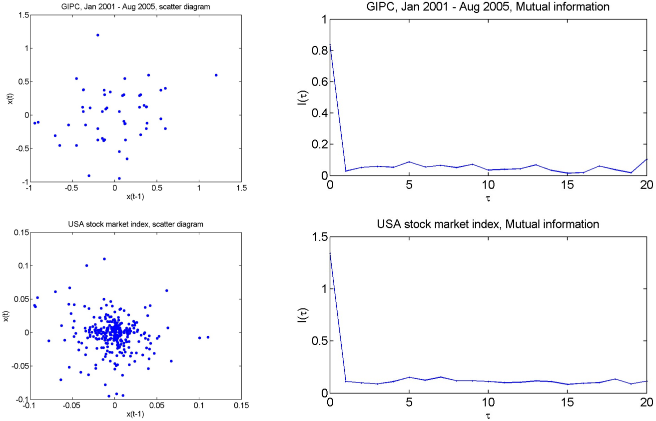


Mutual information: real examples

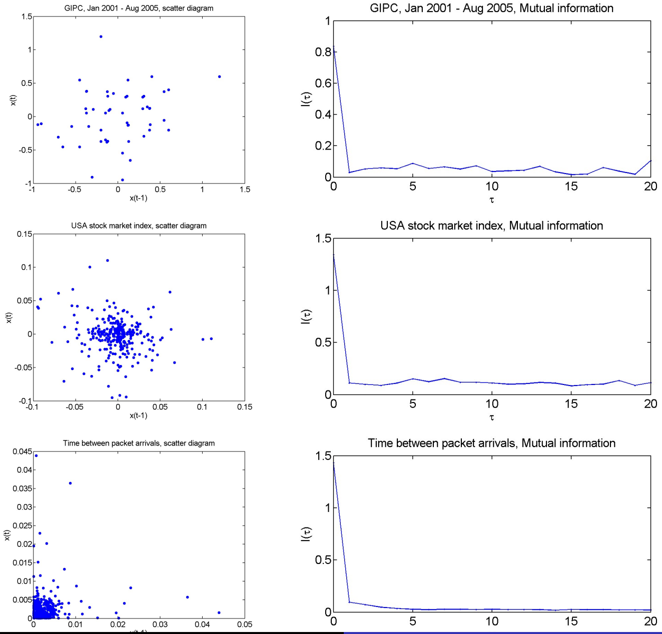




Mutual information: real examples



Mutual information: real examples



Exercise 5: Nonparametric test for zero mutual information

How can we test for zero delayed mutual information ? [to compute $I(\tau)$ in `matlab` use `mutual.m` from the course files]

Exercise 5: Nonparametric test for zero mutual information

How can we test for zero delayed mutual information ? [to compute $I(\tau)$ in `matlab` use `mutual.m` from the course files]

Use resampling (randomization) to form the empirical distribution of $I(\tau)$, as for Exercise 4.

Exercise 5: Nonparametric test for zero mutual information

How can we test for zero delayed mutual information ? [to compute $I(\tau)$ in `matlab` use `mutual.m` from the course files]

Use resampling (randomization) to form the empirical distribution of $I(\tau)$, as for Exercise 4.

Perform the randomization test for the three real time series:

1. Residuals of GICP [course data file `GPIC2001_2005residuals.dat`]

Exercise 5: Nonparametric test for zero mutual information

How can we test for zero delayed mutual information ? [to compute $I(\tau)$ in `matlab` use `mutual.m` from the course files]

Use resampling (randomization) to form the empirical distribution of $I(\tau)$, as for Exercise 4.

Perform the randomization test for the three real time series:

1. Residuals of GICP [course data file `GPIC2001_2005residuals.dat`]
2. Returns of USA stock marker index [course data file `USAreturns.dat`].

Exercise 5: Nonparametric test for zero mutual information

How can we test for zero delayed mutual information ? [to compute $I(\tau)$ in `matlab` use `mutual.m` from the course files]

Use resampling (randomization) to form the empirical distribution of $I(\tau)$, as for Exercise 4.

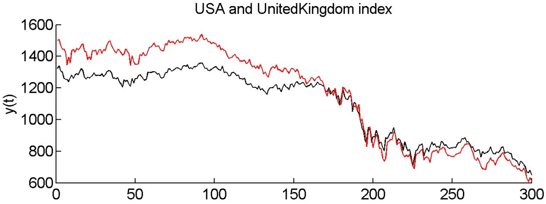
Perform the randomization test for the three real time series:

1. Residuals of GICP [course data file `GPIC2001_2005residuals.dat`]
2. Returns of USA stock marker index [course data file `USAreturns.dat`].
3. Times between packet arrivals [course data file `PacketArrival.dat`]

Time evolution of two stock market indices

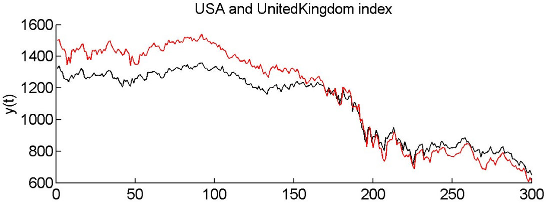
Bivariate time series

Time evolution of two stock market indices



Bivariate time series

Time evolution of two stock market indices

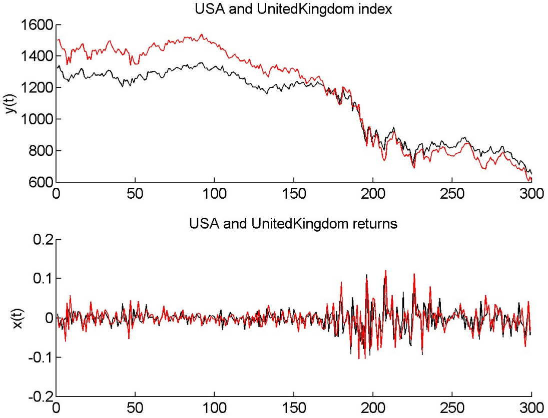


Time evolution of two stock market returns

Bivariate time series

Time evolution of two stock market indices

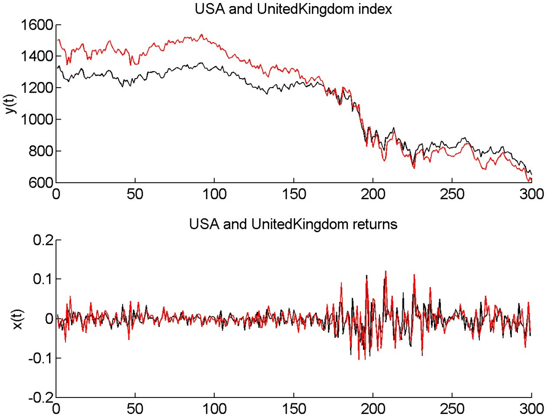
Time evolution of two stock market returns



Bivariate time series

Time evolution of two stock market indices

Time evolution of two stock market returns

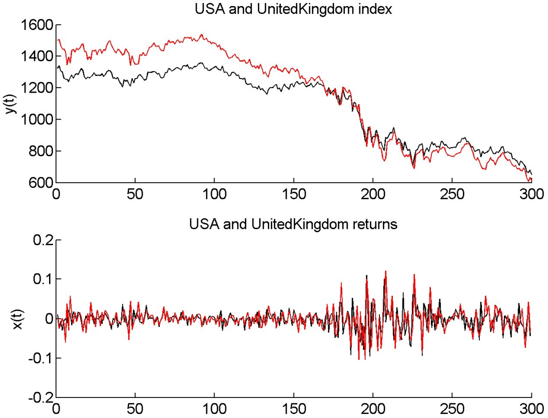


Are there autocorrelations in the two indices?

Bivariate time series

Time evolution of two stock market indices

Time evolution of two stock market returns



Are there autocorrelations in the two indices?

Are there cross-correlations in the two indices?

Cross-covariance and cross-correlation see [1]: Chp 8, [2]: Chp 7

Stationary processes $\{X_t\}$ and $\{Y_t\}$ and their realizations, time series $\{x_t, y_t\}_{t=1}^{n}$.

Stationary processes $\{X_t\}$ and $\{Y_t\}$ and their realizations, time series $\{x_t, y_t\}_{t=1}^{n}$.

cross-covariance: $\gamma_{XY}(\tau) = \text{Cov}[X_t, Y_{t+\tau}] = \text{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$,

Stationary processes $\{X_t\}$ and $\{Y_t\}$ and their realizations, time series $\{x_t, y_t\}_{t=1}^{n}$.

cross-covariance: $\gamma_{XY}(\tau) = \text{Cov}[X_t, Y_{t+\tau}] = \text{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$, and estimate

$$c_{XY}(\tau) = \hat{\gamma}_{XY}(\tau) = \frac{1}{n-\tau} \sum_{t=1}^{n-\tau} (x_t - \bar{x})(y_{t+\tau} - \bar{y})$$

Stationary processes $\{X_t\}$ and $\{Y_t\}$ and their realizations, time series $\{x_t, y_t\}_{t=1}^{n}$.

cross-covariance: $\gamma_{XY}(\tau) = \text{Cov}[X_t, Y_{t+\tau}] = \text{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$, and estimate

$$c_{XY}(\tau) = \hat{\gamma}_{XY}(\tau) = \frac{1}{n-\tau}\sum_{t=1}^{n-\tau}(x_t - \bar{x})(y_{t+\tau} - \bar{y})$$

cross-correlation: $\rho_{XY}(\tau) = \frac{\gamma_{XY}(\tau)}{\gamma_{XY}(0)} = \frac{\gamma_{XY}(\tau)}{\sigma_X \sigma_Y}$,

Cross-covariance and cross-correlation see [1]: Chp 8, [2]: Chp 7

Stationary processes $\{X_t\}$ and $\{Y_t\}$ and their realizations, time series $\{x_t, y_t\}_{t=1}^{n}$.

cross-covariance: $\gamma_{XY}(\tau) = \text{Cov}[X_t, Y_{t+\tau}] = \text{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$, and estimate

$$c_{XY}(\tau) = \hat{\gamma}_{XY}(\tau) = \frac{1}{n-\tau} \sum_{t=1}^{n-\tau} (x_t - \bar{x})(y_{t+\tau} - \bar{y})$$

cross-correlation: $\rho_{XY}(\tau) = \frac{\gamma_{XY}(\tau)}{\gamma_{XY}(0)} = \frac{\gamma_{XY}(\tau)}{\sigma_X \sigma_Y}$, and estimate

$$r_{XY}(\tau) = \hat{\rho}_{XY}(\tau) = \frac{c_{XY}(\tau)}{c_{XY}(0)} = \frac{c_{XY}(\tau)}{s_X s_Y}$$

Stationary processes $\{X_t\}$ and $\{Y_t\}$ and their realizations, time series $\{x_t, y_t\}_{t=1}^{n}$.

cross-covariance: $\gamma_{XY}(\tau) = \text{Cov}[X_t, Y_{t+\tau}] = \text{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$, and estimate

$$c_{XY}(\tau) = \hat{\gamma}_{XY}(\tau) = \frac{1}{n-\tau}\sum_{t=1}^{n-\tau}(x_t - \bar{x})(y_{t+\tau} - \bar{y})$$

cross-correlation: $\rho_{XY}(\tau) = \frac{\gamma_{XY}(\tau)}{\gamma_{XY}(0)} = \frac{\gamma_{XY}(\tau)}{\sigma_X \sigma_Y}$, and estimate

$$r_{XY}(\tau) = \hat{\rho}_{XY}(\tau) = \frac{c_{XY}(\tau)}{c_{XY}(0)} = \frac{c_{XY}(\tau)}{s_X s_Y}$$

Cross-covariance is not even function: $\gamma_{XY}(\tau) \neq \gamma_{XY}(-\tau)$,

Cross-covariance and cross-correlation see [1]: Chp 8, [2]: Chp 7

Stationary processes $\{X_t\}$ and $\{Y_t\}$ and their realizations, time series $\{x_t, y_t\}_{t=1}^n$.

cross-covariance: $\gamma_{XY}(\tau) = \text{Cov}[X_t, Y_{t+\tau}] = \text{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$, and estimate

$$c_{XY}(\tau) = \hat{\gamma}_{XY}(\tau) = \frac{1}{n-\tau}\sum_{t=1}^{n-\tau}(x_t - \bar{x})(y_{t+\tau} - \bar{y})$$

cross-correlation: $\rho_{XY}(\tau) = \frac{\gamma_{XY}(\tau)}{\gamma_{XY}(0)} = \frac{\gamma_{XY}(\tau)}{\sigma_X \sigma_Y}$, and estimate

$$r_{XY}(\tau) = \hat{\rho}_{XY}(\tau) = \frac{c_{XY}(\tau)}{c_{XY}(0)} = \frac{c_{XY}(\tau)}{s_X s_Y}$$

Cross-covariance is not even function: $\gamma_{XY}(\tau) \neq \gamma_{XY}(-\tau)$, but it holds that $\gamma_{XY}(\tau) = \gamma_{YX}(-\tau)$.

Cross-covariance and cross-correlation see [1]: Chp 8, [2]: Chp 7

Stationary processes $\{X_t\}$ and $\{Y_t\}$ and their realizations, time series $\{x_t, y_t\}_{t=1}^{n}$.

cross-covariance: $\gamma_{XY}(\tau) = \mathrm{Cov}[X_t, Y_{t+\tau}] = \mathrm{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$, and estimate

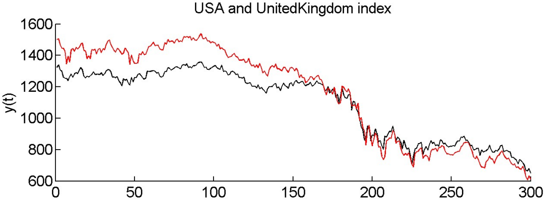
$$c_{XY}(\tau) = \hat{\gamma}_{XY}(\tau) = \frac{1}{n-\tau}\sum_{t=1}^{n-\tau}(x_t - \bar{x})(y_{t+\tau} - \bar{y})$$

cross-correlation: $\rho_{XY}(\tau) = \frac{\gamma_{XY}(\tau)}{\gamma_{XY}(0)} = \frac{\gamma_{XY}(\tau)}{\sigma_X\sigma_Y}$, and estimate

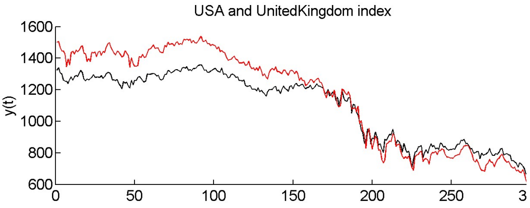
$$r_{XY}(\tau) = \hat{\rho}_{XY}(\tau) = \frac{c_{XY}(\tau)}{c_{XY}(0)} = \frac{c_{XY}(\tau)}{s_X s_Y}$$

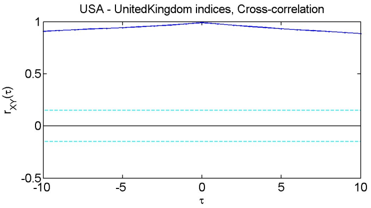
Cross-covariance is not even function: $\gamma_{XY}(\tau) \neq \gamma_{XY}(-\tau)$,
but it holds that $\gamma_{XY}(\tau) = \gamma_{YX}(-\tau)$. Also it holds $|\rho_{XY}(\tau)| \leq 1$.

Example: two world stock indices

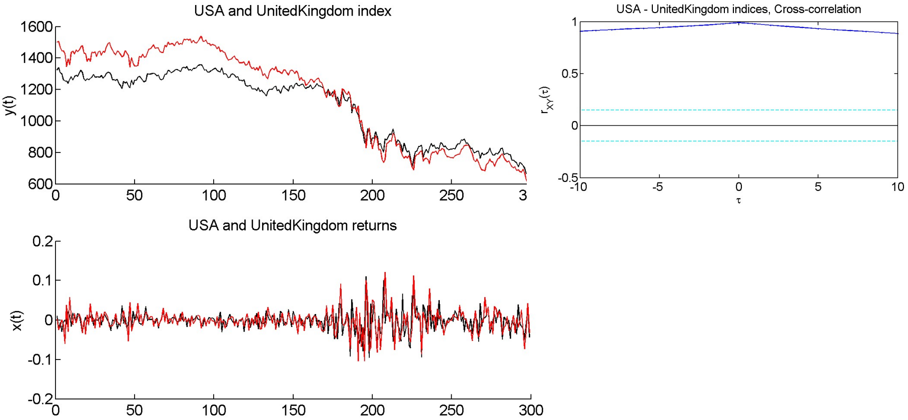


Example: two world stock indices

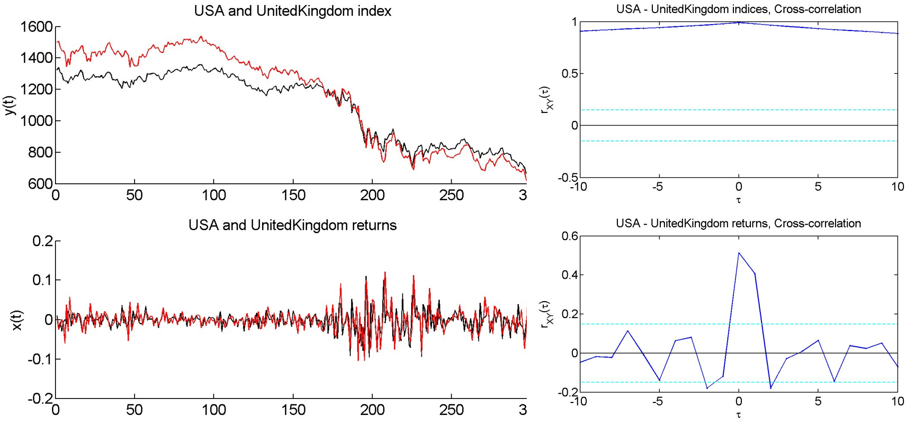




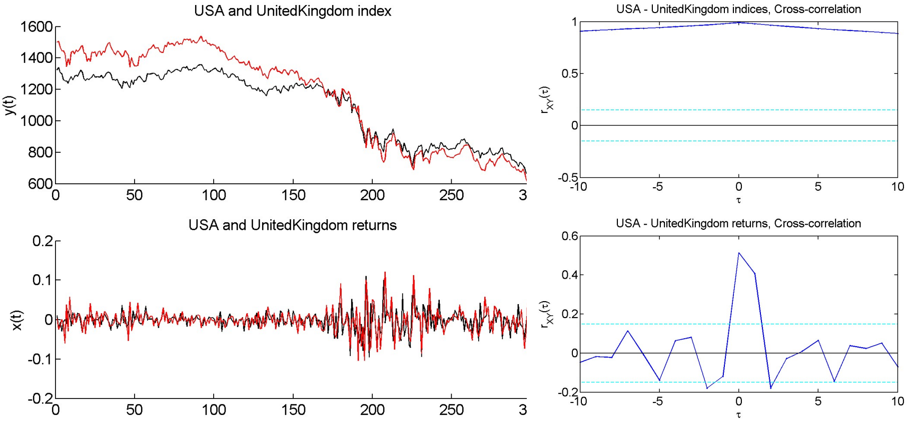
Example: two world stock indices



Example: two world stock indices

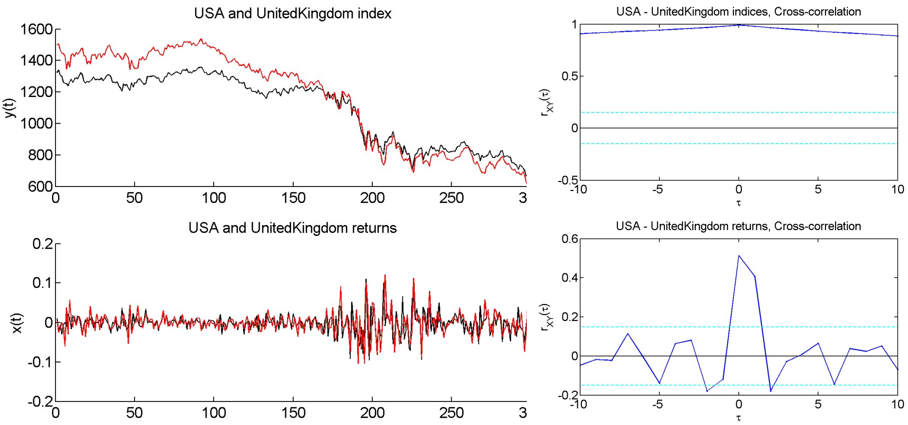


Example: two world stock indices



Limits of significance: $\pm z_{\alpha/2}/\sqrt{n}$.

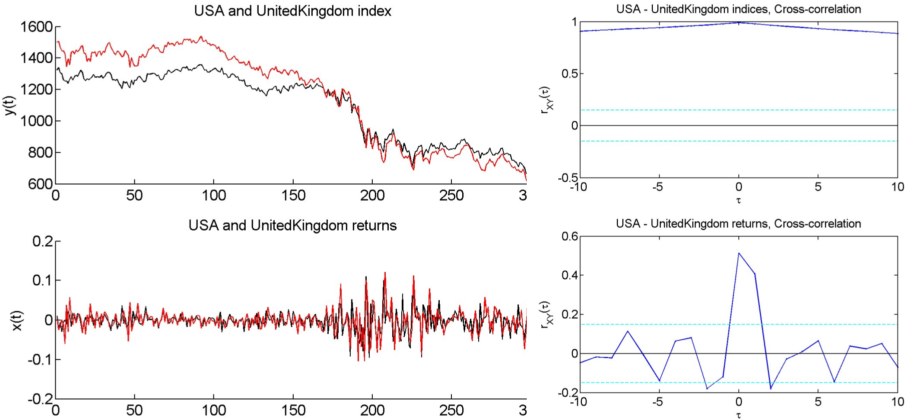
Example: two world stock indices



Limits of significance: $\pm z_{\alpha/2}/\sqrt{n}$.

Large cross-correlations (X: USA, Y: UK):

Example: two world stock indices

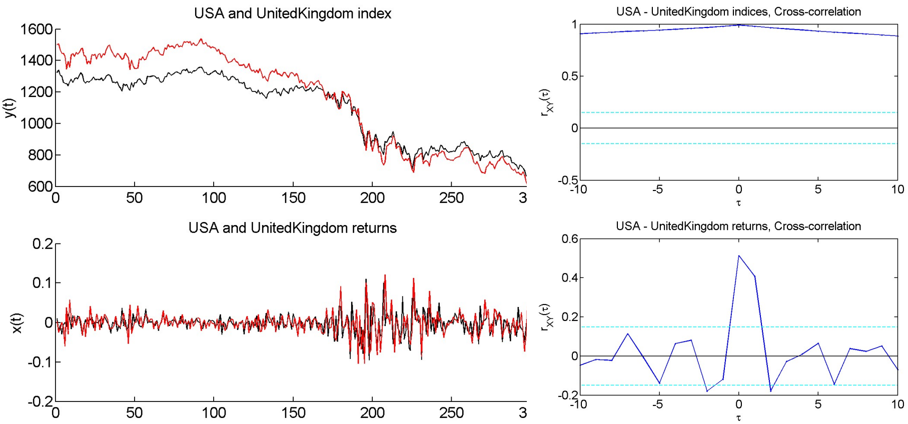


Limits of significance: $\pm z_{\alpha/2}/\sqrt{n}$.

Large cross-correlations (X: USA, Y: UK):

$r_{X,Y}(0) = \mathrm{Corr}(X_t, Y_t)$: USA and UK returns are instantly correlated.

Example: two world stock indices



Limits of significance: $\pm z_{\alpha/2}/\sqrt{n}$.

Large cross-correlations (X: USA, Y: UK):

$r_{X,Y}(0) = \text{Corr}(X_t, Y_t)$: USA and UK returns are instantly correlated.

$r_{X,Y}(1) = \text{Corr}(X_t, Y_{t+1})$: USA return is correlated to UK return a day ahead $\Longrightarrow$ USA returns influence UK returns.

Measure of nonlinear cross-correlation

Cross-covariance: $\gamma_{XY}(\tau) = \text{Cov}[X_t, Y_{t+\tau}] = \text{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$,

Measure of nonlinear cross-correlation

Cross-covariance: $\gamma_{XY}(\tau) = \text{Cov}[X_t, Y_{t+\tau}] = \text{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$,
Extend joint moment of order one $\text{E}[X_t Y_{t+\tau}]$ to higher order joint moments

Measure of nonlinear cross-correlation

Cross-covariance: $\gamma_{XY}(\tau) = \text{Cov}[X_t, Y_{t+\tau}] = \text{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$,
Extend joint moment of order one $\text{E}[X_t Y_{t+\tau}]$ to higher order joint moments $\Longrightarrow$ nonlinear measures.

Measure of nonlinear cross-correlation

Cross-covariance: $\gamma_{XY}(\tau) = \text{Cov}[X_t, Y_{t+\tau}] = \text{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$,
Extend joint moment of order one $\text{E}[X_t Y_{t+\tau}]$ to higher order joint moments $\Longrightarrow$ nonlinear measures.

For $X \to X_t$ and $Y \to Y_{t+\tau}$, the cross-delayed mutual information:

$$\text{I}_{XY}(\tau) = \text{I}(X_t, Y_{t+\tau}) = \sum_{x_t, y_{t+\tau}} p_{X_t Y_{t+\tau}}(x_t, y_{t+\tau}) \log \frac{p_{X_t Y_{t+\tau}}(x_t, y_{t+\tau})}{p_{X_t}(x_t) p_{Y_{t+\tau}}(y_{t+\tau})}$$

Measure of nonlinear cross-correlation

Cross-covariance: $\gamma_{XY}(\tau) = \text{Cov}[X_t, Y_{t+\tau}] = \text{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$,
Extend joint moment of order one $\text{E}[X_t Y_{t+\tau}]$ to higher order joint moments $\Longrightarrow$ nonlinear measures.

For $X \to X_t$ and $Y \to Y_{t+\tau}$, the cross-delayed mutual information:

$$\text{I}_{XY}(\tau) = \text{I}(X_t, Y_{t+\tau}) = \sum_{x_t, y_{t+\tau}} p_{X_t Y_{t+\tau}}(x_t, y_{t+\tau}) \log \frac{p_{X_t Y_{t+\tau}}(x_t, y_{t+\tau})}{p_{X_t}(x_t) p_{Y_{t+\tau}}(y_{t+\tau})}$$

To estimate $\text{I}_{XY}(\tau)$ make a partition of $\{x_t\}_{t=1}^n$, a partition of $\{y_t\}_{t=1}^n$ and compute probabilities for each cell from the relative frequency,

Measure of nonlinear cross-correlation

Cross-covariance: $\gamma_{XY}(\tau) = \text{Cov}[X_t, Y_{t+\tau}] = \text{E}[(X_t - \mu_X)(Y_{t+\tau} - \mu_Y)]$,
Extend joint moment of order one $\text{E}[X_t Y_{t+\tau}]$ to higher order joint moments $\Longrightarrow$ nonlinear measures.

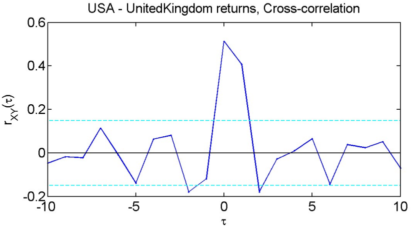
For $X \to X_t$ and $Y \to Y_{t+\tau}$, the cross-delayed mutual information:

$$\text{I}_{XY}(\tau) = \text{I}(X_t, Y_{t+\tau}) = \sum_{x_t, y_{t+\tau}} p_{X_t Y_{t+\tau}}(x_t, y_{t+\tau}) \log \frac{p_{X_t Y_{t+\tau}}(x_t, y_{t+\tau})}{p_{X_t}(x_t) p_{Y_{t+\tau}}(y_{t+\tau})}$$

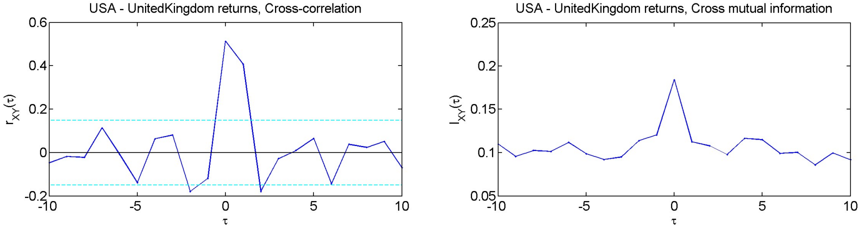
To estimate $\text{I}_{XY}(\tau)$ make a partition of $\{x_t\}_{t=1}^n$, a partition of $\{y_t\}_{t=1}^n$ and compute probabilities for each cell from the relative frequency,

... or better, standardize both time series and use the same partition for each.

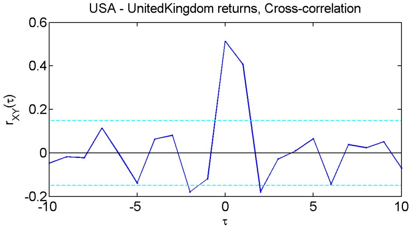
Example: two world stock indices

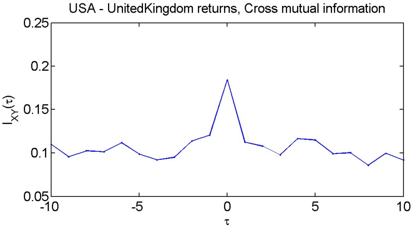


Example: two world stock indices



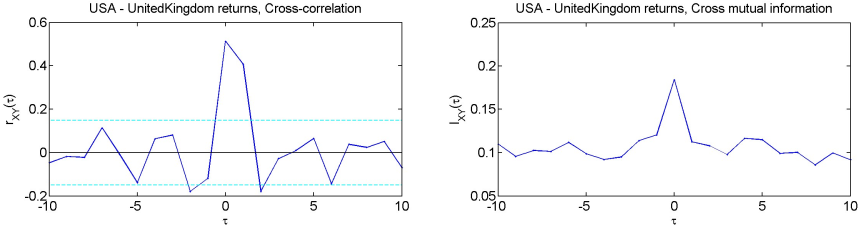
Example: two world stock indices





Limits of significance for $I_{XY}(\tau)$?

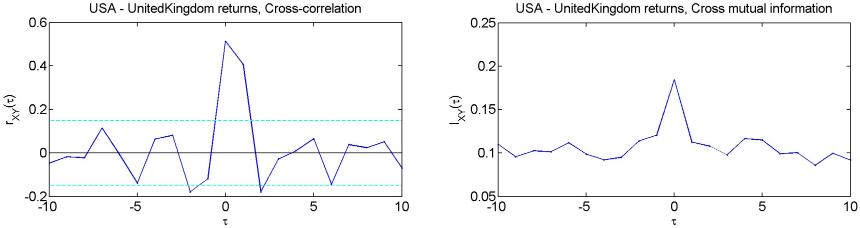
Example: two world stock indices



Limits of significance for $\mathrm{I}_{XY}(\tau)$?

$r_{XY}(0)$ and $\mathrm{I}_{XY}(0)$: USA and UK returns are instantly correlated (linearly and nonlinearly).

Example: two world stock indices



Limits of significance for $I_{XY}(\tau)$?

$r_{XY}(0)$ and $I_{XY}(0)$: USA and UK returns are instantly correlated (linearly and nonlinearly).

$r_{XY}(1)$ large but $I_{XY}(1)$ not large: do USA returns influence UK returns?

Exercise 6: Correlation of two financial indices

Find the correlation between two financial indices

Exercise 6: Correlation of two financial indices

Find the correlation between two financial indices

1. Choose two of the eights markets in file `WorldMarkets.dat` (1. USA, 2. Australia, 3. UnitedKingdom, 4. Germany, 5. Greece, 6. Malaysia, 7. SouthAfrica, 8. Croatia)
2. Compute the cross-correlation between the two indices and between their returns.
3. Decide for the statistical significant cross-correlation between the two markets (use a parametric significance test).

Literature

[1] Chatfield C (2004) *The Analysis of Time Series, An Introduction*, Sixth Edition, Chapman & Hall.

[2] Brockwell PJ and Davis RA (2002) *Introduction to Time Series and Forecasting*, Second Edition, Springer.

[3] Kantz H and Schreiber T (2003) *Nonlinear Time Series Analysis*, Second Edition, Cambridge.