Financial Time Series

Lecture 2

Linear time series

- A time series $\{r_t\}$ is linear if it can be written as

$$r_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i}$$

Here $\mu = E[r_t]$, $\psi_0 = 1$ and $\{a_t\}$ is white noise, i.e. independent with mean zero and constant variance ${\sigma^2}_a$

- Note that $\{r_t\}$ is (strictly) stationary by definition

Linear time series

- So for the series $\{r_t\}$ we have
- $Var(r_t) = \sigma^2{}_a \sum_{i=0}^{\infty} (\psi_i)^2$
- $\gamma_l = Cov(r_t, r_{t-l}) = \sigma^2{}_a \sum_{i=0}^{\infty} \psi_i \psi_{i+l}$
- $\rho_l = \frac{\gamma_l}{\gamma_0} = \frac{\sum_{i=0}^{\infty} \psi_i \psi_{i+l}}{\sum_{i=0}^{\infty} (\psi_i)^2}$

Autoregressive models

- If we for an observed series find significant one lag autocorrelation it may be useful to model the series as

$$r_t = \varphi_0 + \varphi_1 r_{t-1} + a_t$$

Here $\{a_t\}$ is white noise, i.e. independent with mean zero and constant variance $\sigma^2{}_a$

AR(1)

- So for the AR(1) series $\{r_t\}$ as on the previous slide

- $Var(r_t) = \gamma_0 = \frac{{\sigma^2}_a}{1-(\varphi_1)^2}$

- $\gamma_l = Cov(r_t, r_{t-l}) = \varphi_1 \gamma_{l-1}, l > 0$

- $\rho_l = \varphi_1 \rho_{l-1} = (\varphi_1)^l, l > 0$

AR-processes

- A zero mean AR(p)-process is defined by

$$r_t = \varphi_1 r_{t-1} + \cdots + \varphi_p r_{t-p} + a_t$$

- May be useful in modeling e.g., stock returns

- To be able to do this we need to estimate $\varphi_1, \dots, \varphi_p$…

MA models

- The series defined by

$$r_t = \mu + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q},$$

is called an MA(q) series/process.

- Stationary by definition

MA(q)

- It holds that
- $E[r_t] = \mu$
- $Var(r_t) = \left(1 + \sum_{i=1}^{q} (\theta_i)^2\right) \sigma^2{}_a$
- $\gamma_l = \{\mu = 0\} = \sigma^2{}_a \sum_{k=|l|}^{q} \left(-\theta_{k-|l|}\right) (-\theta_k)$

where $-\theta_0 = 1$ and $\theta_k = 0, k > q$

Reflection

- Note that the autocorrelation cuts off after lag q

- Note also that for an MA(1) we may use $a_t = r_t + \theta_1 a_{t-1}$ and write

$$a_t = r_t + \theta_1 r_{t-1} + (\theta_1)^2 r_{t-2} + \cdots$$

- In order for this to make sense it must hold that $|\theta_1| < 1$ in which case the MA(1) is invertible

Choosing AR or MA or both?

- The choice of model may be made using autocorrelation functions (ACF) and partial autocorrelation functions (PACF)
- Typically, a slowly decaying ACF and an ACF with ”q peaks and then zero” indicate AR and MA(q) models, respectively
- Below we will explain PACF

Estimation of AR-parameters

- Estimation may be done in several different ways, but we need to decide on what number, p, of parameters to use
- This can be done using partial autocorrelations (PACF)

$$\alpha_h = \varphi_{hh}$$

where φ_{hh} is the parameter φ_h in the $AR(h)$ model

Intuition

- For an AR(p) process we may think of the the PACF at lag 5, say, as the correlation between X_t and X_{t+5} not explained by their common correlations with $X_{t+1}, \ldots, X_{t+4}$

- If the PACF at a certain lag is statistically significant the AR model of choice will have order no smaller than that lag

Example

- For an AR(1) model, $r_t = \varphi r_{t-1} + a_t$, we know that the ACF is

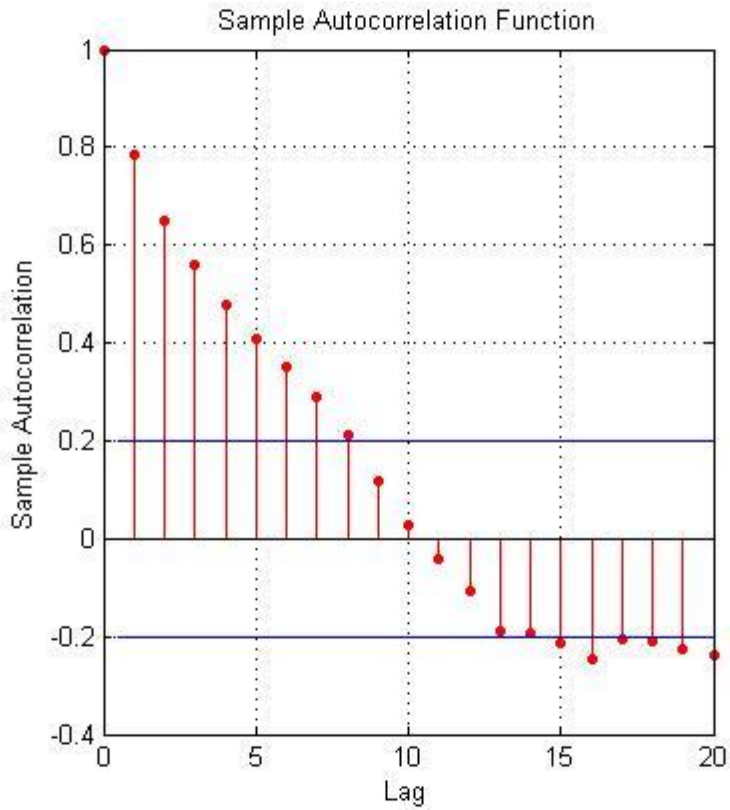
$$\rho_h = \varphi^{|h|}$$

- One can show that the PACF is

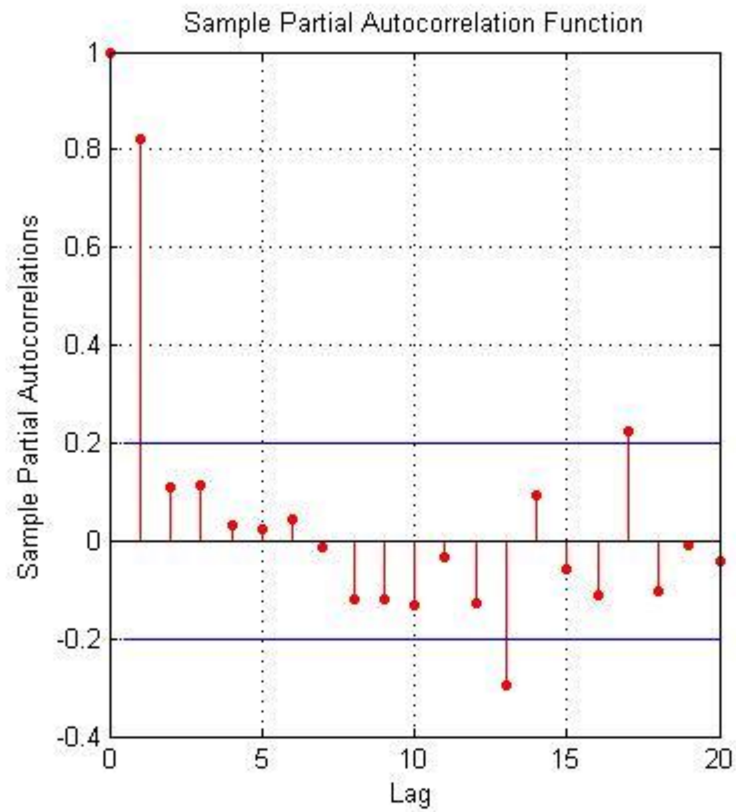
$$\alpha_1 = \varphi \text{ and } \alpha_h = 0 \text{ for } h > 0$$

- So the correlation for larger lags is explained by the lag one correlation

Example AR(1) with $\varphi = 0.8$



ACF



PACF

To find the φ-parameters

- For the $AR(h)$ model we may use the Yule-Walker equations:

$$\gamma = \Gamma\varphi$$

where $\gamma = [\gamma_1 \cdots \gamma_h]^T$ and $\Gamma_{ij} = \gamma_{i-j}$

- Numerically this is done by replacing γ_h with

$$\hat{\gamma}_h = \frac{1}{n}\sum_{j=1}^{n-h}\left(x_{j+h} - \bar{x}\right)\left(x_j - \bar{x}\right)$$

Note the

- We need to invert the ”gamma matrix” to find $\hat{\varphi}$
- Choosing $1/n$ instead of the ”intuitive” $1/(n-h)$ in the sample autocovariance will guarantee invertibility

To determine the order p

- We need to decide up to which order the PACF is statistically significant

- At 95% confidence we want our estimate of φ_{hh} within $\pm 1.96/\sqrt{T}$ where T is the no. of observations

- In matlab you can use the ”parcorr” function

Checking the model

- We want residuals to behave like WN
- So, for our fitted model

$$\hat{r}_t = \hat{\varphi}_1 r_{t-1} + \cdots + \hat{\varphi}_p r_{t-p}$$

we create the residuals

$$e_t = \hat{r}_t - r_t$$

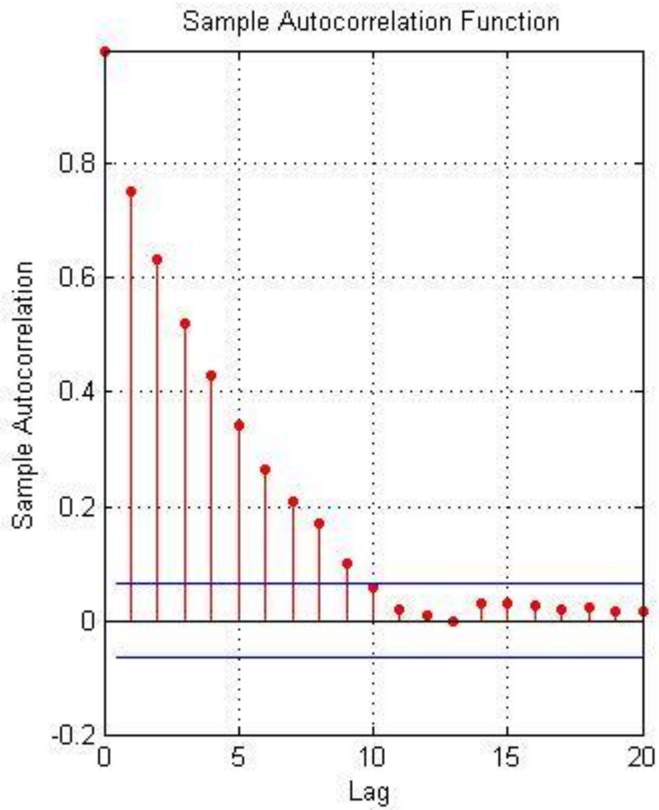
Checking the model

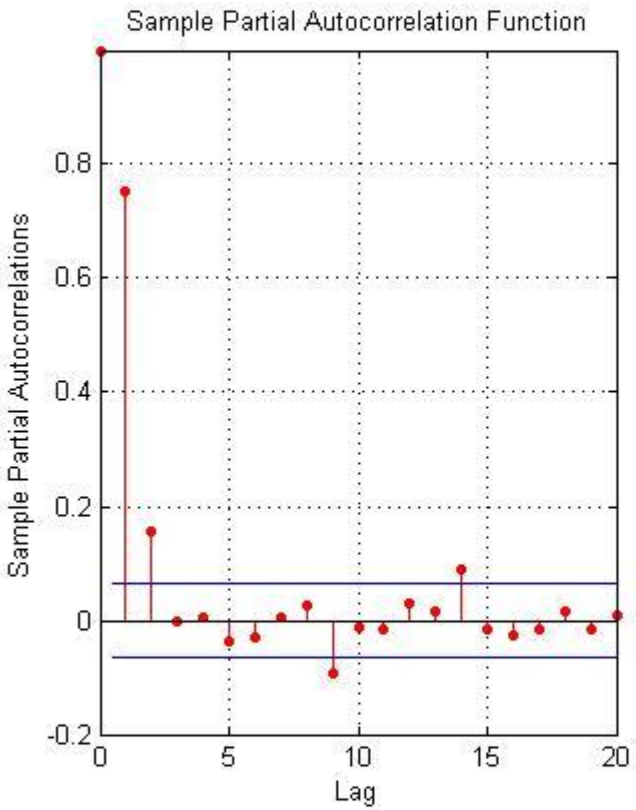
- Apply graphical, Ljung-Box or Portmeanteau procedures

- From the residuals we also get an estimate of the WN variance: $\hat{\sigma}^2{}_a = \frac{1}{T-p}\sum_{i=p+1}^{T} e_i{}^2$

Example

- Given some time series, we find





Example

- Based on the plots it seems reasonable to fit an AR(2)-process

- To do this, using the Yule-Walker equations, we need the sample covariances at lags 0,1,2

- Using the cov command in matlab we find the elements of the estimated covariance matrix, $\hat{\Gamma}$ and the estimated covariance vector $\hat{\gamma}$

Example

- We have $\hat{\Gamma}_{11} = \hat{\Gamma}_{22} = \hat{\gamma}_0 = 2.3456$, $\hat{\Gamma}_{12} = \hat{\Gamma}_{21} = \hat{\gamma}_1 = 1.7667$ and $\hat{\gamma}_2 = 1.4882$
- We may invert $\hat{\Gamma}$ using the inv command
- Finally, we find
$$\hat{\varphi} = \hat{\Gamma}^{-1}\hat{\gamma} = [0.6362\ 0.1553]^T$$

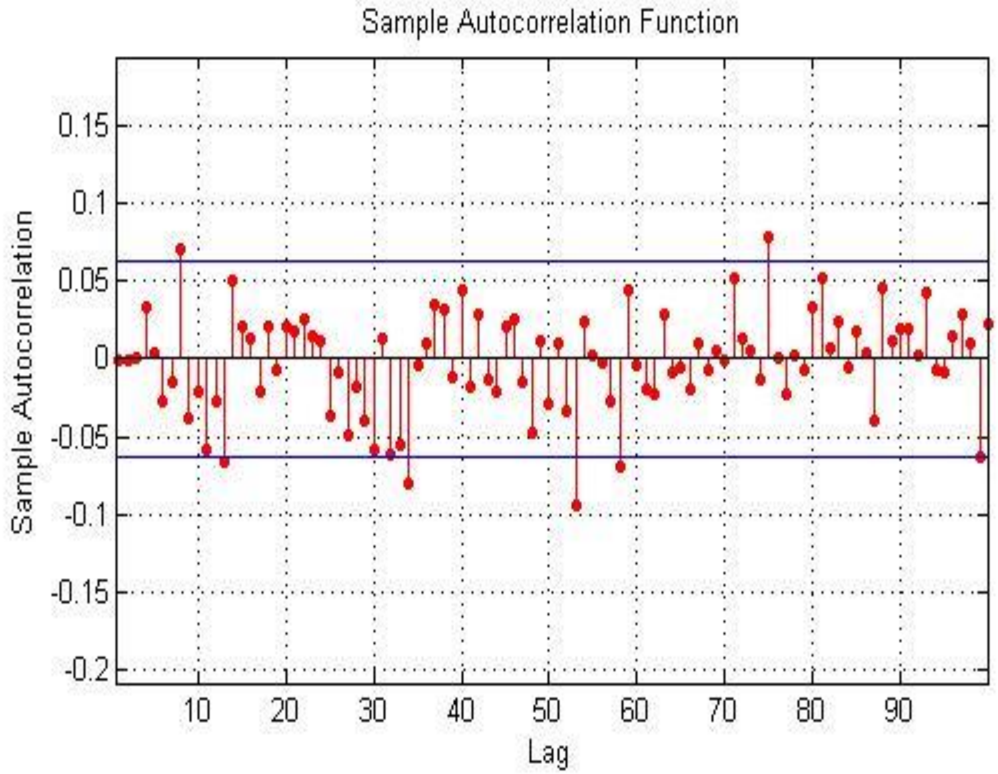
Example

- So our estimated series, i.e., predictions are

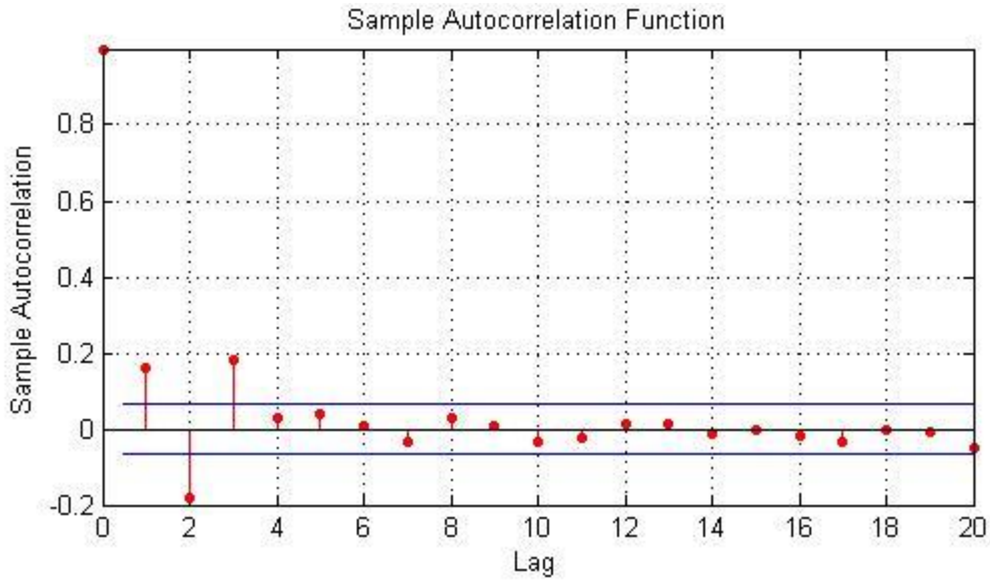
$$\hat{r}_t = 0.6362 r_{t-1} + 0.1553 r_{t-2}$$

- If the fit is good, we will see uncorrelated residuals $e_t = r_t - \hat{r}_t$

Example


Sample Autocorrelation Function
Sample Autocorrelation
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
10
20
30
40
50
60
70
80
90
Lag

Identifying MA



Identifying MA

- Above, we see the ACF for an MA(3)

- Generally, we may use the ACF and decide on using MA if the ACF is significant up to a certain lag

- The last lag at which the ACF is significant is q

Reflection

- To identify the number of parameters to use in AR estimation we use the PACF

- To identify the number of parameters to use in MA estimation we use the ACF

Estimating MA parameters

- In the example above, we use the ACF to decide that an MA(3) seems reasonable

- To use this for predictions, we need to estimate θ_1, θ_2 and θ_3

- This can be done using the auto covariance function, ML or ”The innovations algorithm” or…

Estimation using the autocovariance

- We know that for an MA(q) process the autocovarince function is

$$\gamma_h = {\sigma^2}_a \sum_{k=|h|}^{q} \left(-\theta_{k-|h|}\right)(-\theta_k)$$

Estimation using the autocovariance

- So for the MA(3), we have

$$\gamma_0 = \sigma^2{}_a\left(1 + \theta_1{}^2 + \theta_2{}^2 + \theta_3{}^2\right)$$

$$\gamma_1 = \sigma^2{}_a(-\theta_1 + \theta_1\theta_2 + \theta_2\theta_3)$$

$$\gamma_2 = \sigma^2{}_a(-\theta_2 + \theta_1\theta_3)$$

$$\gamma_3 = -\sigma^2{}_a\theta_3$$

To solve these equations

- We may use ”fsolve” in matlab replacing γ by $\hat{\gamma}$
- Using 1000 observations of the MA(3) process above (which was in fact $r_t = a_t - (-0.6)a_{t-1} - 0.4a_{t-2} - (-0.2)a_{t-3}$ and $a_t \sim N(0,1)$) we find

$$\hat{\theta}_1 = -1.0603, \hat{\theta}_2 = 0.3520, \hat{\theta}_3 = -0.0982,$$

$$\hat{\sigma}^2 = 0.6620$$

But, using 10000 or 100000 obs…

- For 10000 obs, we get

$$\hat{\theta}_1 = -0.6287, \hat{\theta}_2 = 0.4133, \hat{\theta}_3 = -0.2174,$$

$$\hat{\sigma}^2 = 0.9655$$

- And for 100000

$$\hat{\theta}_1 = -0.6048, \hat{\theta}_2 = 0.4058, \hat{\theta}_3 = -0.1951,$$

$$\hat{\sigma}^2 = 0.9971$$

Innovations algorithm

- Recursive procedure which also gives us an estimate of the WN variance ${\sigma^2}_a$
- Uses the sample covariances at lags up to the same order as the process we want to fit
- Gives ”rough estimates” that may be used as starting values for refined methods, e.g., ML or least squares

Innovations algoritm (in general)

- Start with $\hat{v}_0 = \hat{\gamma}(0)$
- For $m = 1, \dots, n-1$ and $k = 0, \dots, m-1$,proceed using

$$\hat{\theta}_{m,m-k} = \frac{1}{\hat{v}_k}\left[\hat{\gamma}(m-k) - \sum_{j=0}^{k-1} \hat{\theta}_{m,m-j}\hat{\theta}_{k,k-j}\hat{v}_j\right]$$

$$\hat{v}_m = \hat{\gamma}(0) - \sum_{j=0}^{m-1} \hat{\theta}^2{}_{m,m-j}\,\hat{v}_j$$

Innovations algoritm

- The estimate of the WN variance is $\hat{v}_m$

Innovations algorithm

- So, to fit an AR(3), we need $\hat{\gamma}(0), \hat{\gamma}(1)$, …

- Using the cov command in matlab we use the covariance matrix with covariances for 99 lags

Innovations algorithm

- So $\hat{v}_0 = \hat{\gamma}(0) = 1.4337$
- $\hat{\theta}_{11} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = 0.1588$
- $\hat{v}_1 = \hat{\gamma}(0) - \hat{\theta}^2{}_{11}\hat{v}_0 = 1.3975$
- $\hat{\theta}_{22} = \frac{\hat{\gamma}(2)}{\hat{\gamma}(0)} = -0.1811$
- $\hat{\theta}_{21} = \frac{1}{\hat{v}_1}\left[\hat{\gamma}(1) - \hat{\theta}_{22}\hat{\theta}_{11}\hat{v}_0\right] = 0.1924$
- $\hat{v}_2 = \hat{\gamma}(0) - \hat{\theta}^2{}_{22}\hat{v}_0 - \hat{\theta}^2{}_{21}\hat{v}_1 = 1.3349$

Innovations algorithm

- $\hat{\theta}_{33} = \frac{\hat{\gamma}(3)}{\hat{\gamma}(0)} = 0.1802$
- $\hat{\theta}_{32} = \frac{1}{\hat{v}_1}\left[\hat{\gamma}(2) - \hat{\theta}_{33}\hat{\theta}_{11}\hat{v}_0\right] = -0.2151$
- $\hat{\theta}_{31} = \frac{1}{\hat{v}_2}\left[\hat{\gamma}(1) - \hat{\theta}_{33}\hat{\theta}_{22}\hat{v}_0 - \hat{\theta}_{32}\hat{\theta}_{21}\hat{v}_1\right] = 0.2490$
- $\hat{v}_3 = \hat{\gamma}(0) - \hat{\theta}^2{}_{33}\hat{v}_0 - \hat{\theta}^2{}_{32}\hat{v}_1 - \hat{\theta}^2{}_{31}\hat{v}_2 = 1.2397$

Innovations algorithm

- Proceeding in 99 steps (until estimates ”do not fluctuate too much”) we find

$\hat{\theta}_{99,1} = 0.4216$, $\hat{\theta}_{99,2} = -0.2902$, $\hat{\theta}_{99,3} = 0.1885$ and

$$\hat{v}_{99} = 1.2290$$

Predictions

- For the AR(3) example, with the estimates from the autocovariance procedure based on 10000 obs

$$\hat{r}_t = 0.6287a_{t-1} - 0.4133a_{t-2} + 0.2174a_{t-3}$$

- How do we find (estimate) the WN components?

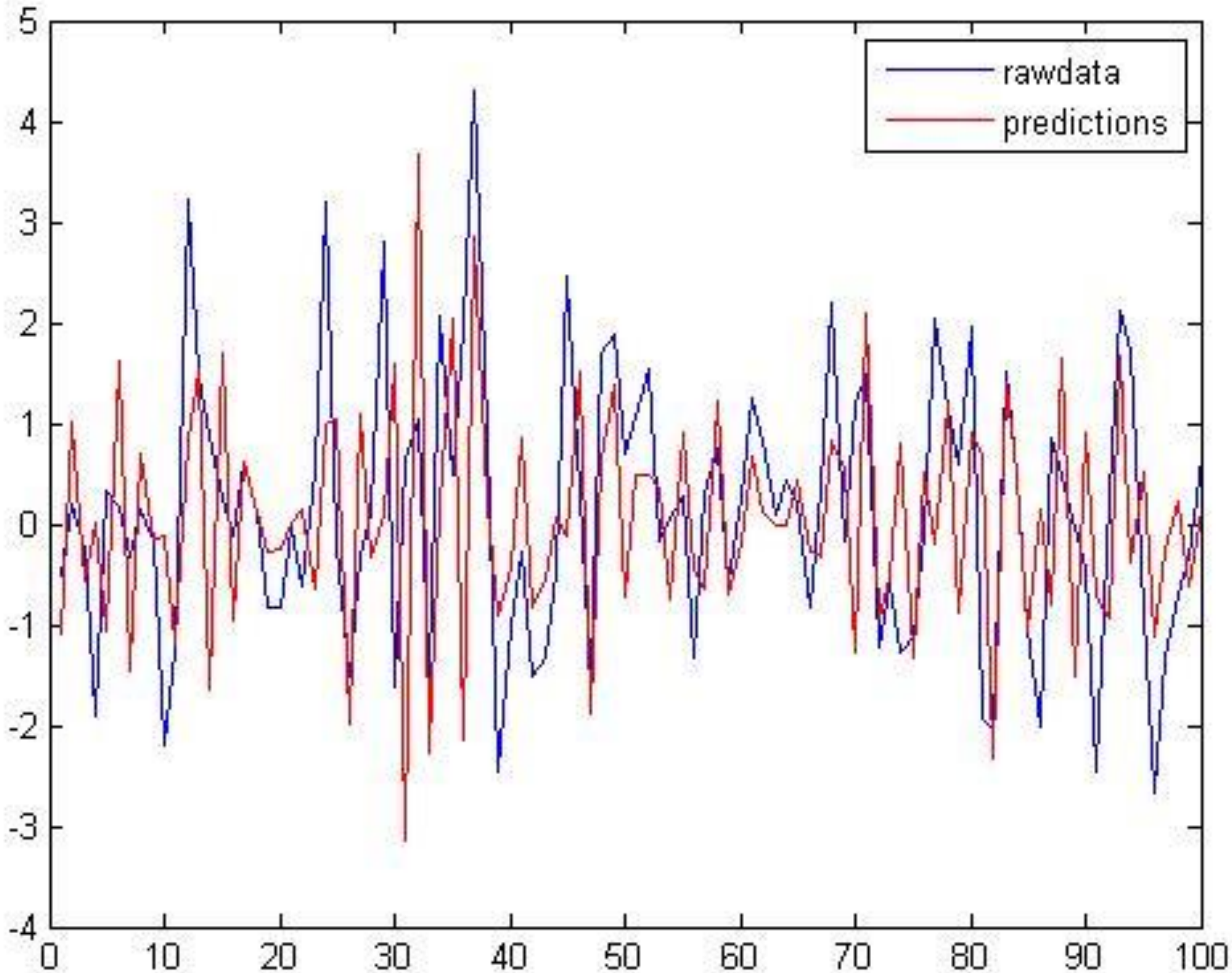
Predictions

- We may use

$$a_1 = r_1$$
$$a_t = r_t - 0.6287 a_{t-1}$$

- Below we find a plot of predictions and observed values

Predictions



Alternative prediction method

- Generally, we are looking for linear predictors

$$\hat{r}_t = \alpha_0 + \sum_{i=1}^{n} \alpha_i r_{t-i}$$

that minimize

$$E[(\hat{r}_t - r_t)^2]$$

Alternative prediction method

- The optimal ”alphas” may be found by numerically minimizing (fminsearch in matlab)

$$\sum_{k=0}^{q+n}\left[\sum_{l=\max\{0,k-q\}}^{\min\{k,n\}} \alpha_l \hat{\theta}_{k-l}\right]^2$$

Characteristic roots and business cycles

- The roots of the characteristic polynomial of AR(p) model:

$$\varphi(z) = 1 - \varphi_1 z - \cdots - \varphi_p z^p$$

can be used to describe long-term behvaiour of the series

Example

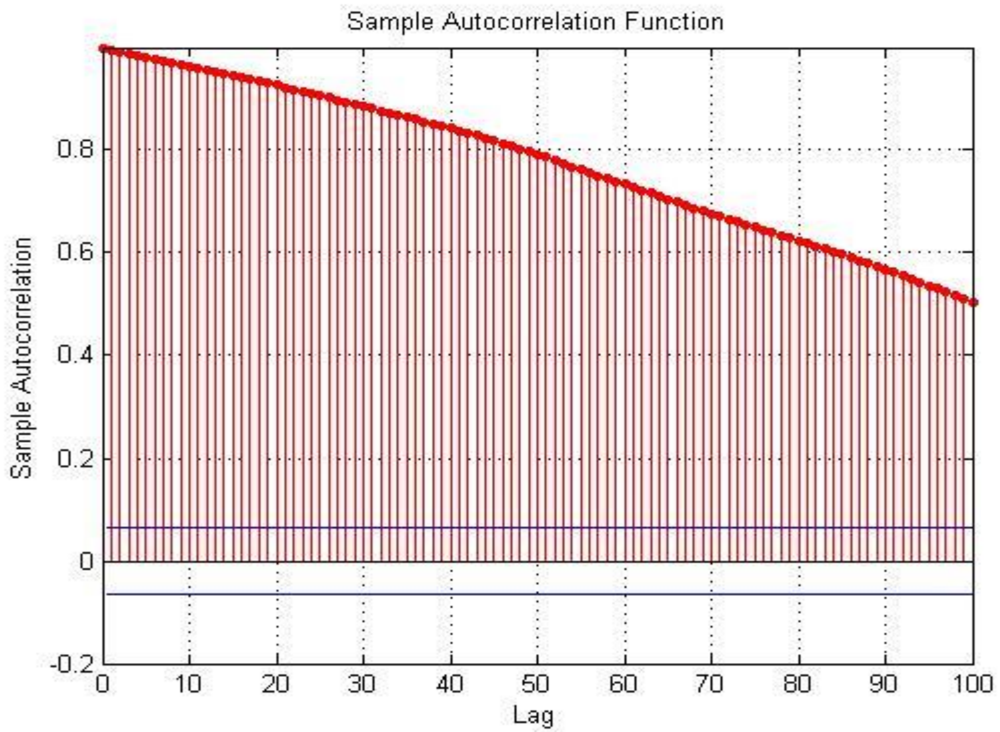
- For an AR(2), we may solve the equation

$$1 - \varphi_1 x - \varphi_2 x^2 = 0$$

- If the roots are real the ACF will behave like a mixture of two exponential decays
- If the roots are complex numbers the ACF will behave like damping sine and cosine waves

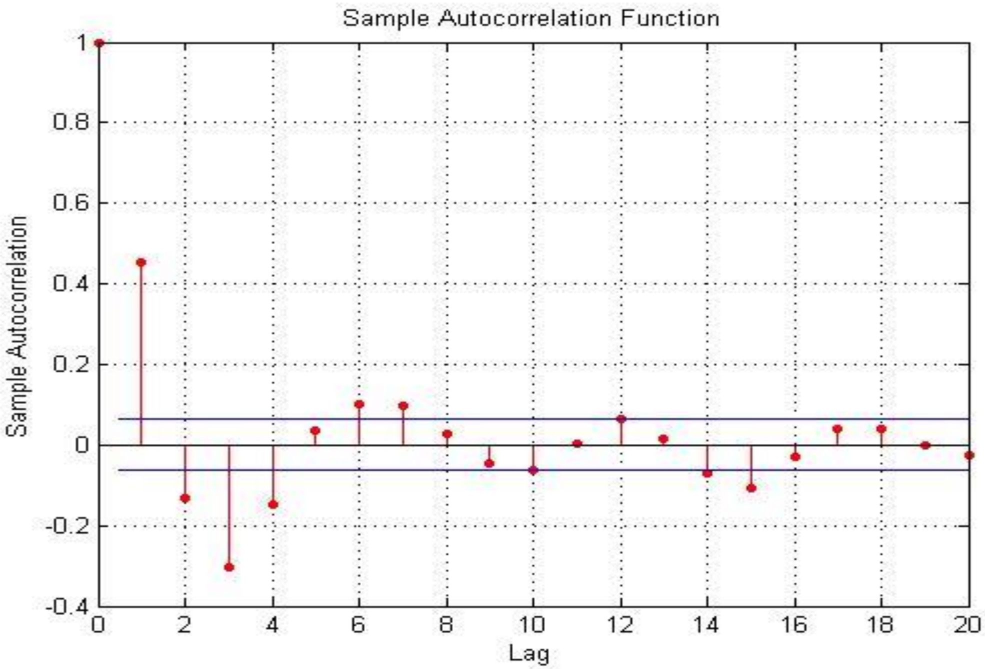
Example

- For $r_t = 0.6r_{t-1} + 0.4r_{t-2} + a_t$, "the roots" are 1 and 2.5



Example

- For $r_t = 0.6r_{t-1} - 0.4r_{t-2} + a_t$ "the roots" are $\frac{3}{4} \pm \frac{\sqrt{31}}{4} i$



Business cycles

- The complex roots indicate business cycles for which the average lengths, k, are given by cycles

$$k = \frac{2\pi}{arccos\left(\varphi_1/\left(2\sqrt{-\varphi_2}\right)\right)}$$