Time-Series Analysis: Stochastic Models

Evgeny Burnaev, Alexey Zaytsev

Skoltech, Moscow, Russia

All models are wrong, but some are useful

- We'll consider some models for time series.
- They have different ideas behind them and can be arbitrary complex.
- But they are able to provide reasonable predictions and are interpretable.

Stations Institute of Science and Technology 2/28

Outline

- Auto Regression (AR)
- 2 Moving average
- Autocovariance function
- Auto Regression (ARMA) process



Skoltech

• Example: AR(1) process (AutoRegression)

$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

 \bullet If $\varphi=1,$ we get nonstationary Random walk process.

Burnaev, ML Such refer effects and the support of t

• Example: AR(1) process (AutoRegression)

$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

• Can we make y_t stationary (i.e. $|\varphi| < 1$)?

$$\mathbb{E}[y_t] = \varphi \mathbb{E}[y_{t-1}] = 0 \text{ (from stationarity)}$$

$$\mathbb{E}[y_t^2] = arphi^2 \mathbb{E}[y_{t-1}^2] + \sigma^2 = rac{\sigma^2}{1-arphi^2}$$
 (from stationarity)

◆ロト ◆個ト ◆量ト ◆量ト ■ からの

Burnaev, ML Skottech

• Example: AR(1) process (AutoRegression)

$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

• Can we make y_t stationary (i.e. $|\varphi| < 1$)?

$$\mathbb{E}[y_t] = \varphi \mathbb{E}[y_{t-1}] = 0 \ \ (\text{from stationarity})$$

$$\mathbb{E}[y_t^2] = arphi^2 \mathbb{E}[y_{t-1}^2] + \sigma^2 = rac{\sigma^2}{1-arphi^2}$$
 (from stationarity)

4□ > <□ > <□ > <□ > <□ > < □ > < □ </p>

• Example: AR(1) process (AutoRegression)

$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

• Can we make y_t stationary (i.e. $|\varphi| < 1$)?

$$\mathbb{E}[y_t] = \varphi \mathbb{E}[y_{t-1}] = 0 \ \ (\text{from stationarity})$$

$$\mathbb{E}[y_t^2] = arphi^2 \mathbb{E}[y_{t-1}^2] + \sigma^2 = rac{\sigma^2}{1-arphi^2}$$
 (from stationarity)

4□ > <</p>
4□ >
4 = >
5
9
0

5/28

• Example: AR(1) process (AutoRegression)

$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

• Can we make y_t stationary (i.e. $|\varphi| < 1$)?

$$\mathbb{E}[y_t] = \varphi \mathbb{E}[y_{t-1}] = 0 \ \ (\text{from stationarity})$$

$$\mathbb{E}[y_t^2] = \varphi^2 \mathbb{E}[y_{t-1}^2] + \sigma^2 = \frac{\sigma^2}{1 - \varphi^2} \quad \text{(from stationarity)}$$

◆ロ > ◆部 > ◆注 > を注 > 注 の < @ >

$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

Assume that y_t is stationary (i.e. $|\varphi| < 1$), then

$$\begin{split} \mathbb{E}[y_t] &= 0, \quad \mathbb{E}[y_t^2] = \frac{\sigma^2}{1 - \varphi^2} \\ \gamma_y(h) &= \operatorname{Cov}\left(\varphi y_{t+h-1} + \varepsilon_{t+h}, y_t\right) = \\ &= \varphi \operatorname{Cov}(y_{t+h-1}, y_t) \\ &= \varphi \gamma_y(h-1) \\ &= \varphi^{|h|} \gamma_y(0) \quad \text{(check for } h > 0 \text{ and } h < 0) \\ &= \frac{\varphi^{|h|} \sigma^2}{1 - \varphi^2} \end{split}$$

Southons Institute of Science and Technology 6/28

$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

Assume that y_t is stationary (i.e. $|\varphi| < 1$), then

$$\begin{split} \mathbb{E}[y_t] &= 0, \quad \mathbb{E}[y_t^2] = \frac{\sigma^2}{1 - \varphi^2} \\ \gamma_y(h) &= \operatorname{Cov}\left(\varphi y_{t+h-1} + \varepsilon_{t+h}, y_t\right) = \\ &= \varphi \operatorname{Cov}(y_{t+h-1}, y_t) \\ &= \varphi \gamma_y(h-1) \\ &= \varphi^{|h|} \gamma_y(0) \quad \text{(check for } h > 0 \text{ and } h < 0) \\ &= \frac{\varphi^{|h|} \sigma^2}{1 - \varphi^2} \end{split}$$

Skoltech
Salous rozzas et Science and Technology 6/28

$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

Assume that y_t is stationary (i.e. $|\varphi| < 1$), then

$$\begin{split} \mathbb{E}[y_t] &= 0, \ \mathbb{E}[y_t^2] = \frac{\sigma^2}{1 - \varphi^2} \\ \gamma_y(h) &= \operatorname{Cov}\left(\varphi y_{t+h-1} + \varepsilon_{t+h}, y_t\right) = \\ &= \varphi \operatorname{Cov}(y_{t+h-1}, y_t) \\ &= \varphi \gamma_y(h-1) \\ &= \varphi^{|h|} \gamma_y(0) \ \text{(check for } h > 0 \ \text{and } h < 0 \\ &= \frac{\varphi^{|h|} \sigma^2}{1 - \varphi^2} \end{split}$$

Southout Institute of Science and Technology 6/28

$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

Assume that y_t is stationary (i.e. $|\varphi| < 1$), then

$$\begin{split} \mathbb{E}[y_t] &= 0, \ \mathbb{E}[y_t^2] = \frac{\sigma^2}{1 - \varphi^2} \\ \gamma_y(h) &= \operatorname{Cov}\left(\varphi y_{t+h-1} + \varepsilon_{t+h}, y_t\right) = \\ &= \varphi \operatorname{Cov}(y_{t+h-1}, y_t) \\ &= \varphi \gamma_y(h-1) \\ &= \varphi^{|h|} \gamma_y(0) \ \text{(check for } h > 0 \ \text{and } h < 0 \\ &= \frac{\varphi^{|h|} \sigma^2}{1 - \varphi^2} \end{split}$$

Burnaev, ML Salar interest flowers (6/28

$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

Assume that y_t is stationary (i.e. $|\varphi| < 1$), then

$$\begin{split} \mathbb{E}[y_t] &= 0, \ \mathbb{E}[y_t^2] = \frac{\sigma^2}{1 - \varphi^2} \\ \gamma_y(h) &= \operatorname{Cov}\left(\varphi y_{t+h-1} + \varepsilon_{t+h}, y_t\right) = \\ &= \varphi \operatorname{Cov}(y_{t+h-1}, y_t) \\ &= \varphi \gamma_y(h-1) \\ &= \varphi^{|h|} \gamma_y(0) \ \text{(check for } h > 0 \ \text{and } h < 0 \\ &= \frac{\varphi^{|h|} \sigma^2}{1 - \varphi^2} \end{split}$$

SKOITECH alkone brookste of Science and Technology 6/28

$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

Assume that y_t is stationary (i.e. $|\varphi| < 1$), then

$$\begin{split} \mathbb{E}[y_t] &= 0, \ \mathbb{E}[y_t^2] = \frac{\sigma^2}{1 - \varphi^2} \\ \gamma_y(h) &= \operatorname{Cov}\left(\varphi y_{t+h-1} + \varepsilon_{t+h}, y_t\right) = \\ &= \varphi \operatorname{Cov}(y_{t+h-1}, y_t) \\ &= \varphi \gamma_y(h-1) \\ &= \varphi^{|h|} \gamma_y(0) \ \text{(check for } h > 0 \ \text{and } h < 0) \\ &= \frac{\varphi^{|h|} \sigma^2}{1 - \varphi^2} \end{split}$$

skortecn
store or of the control of

$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

Assume that y_t is stationary (i.e. $|\varphi| < 1$), then

$$\begin{split} \mathbb{E}[y_t] &= 0, \ \mathbb{E}[y_t^2] = \frac{\sigma^2}{1 - \varphi^2} \\ \gamma_y(h) &= \operatorname{Cov}\left(\varphi y_{t+h-1} + \varepsilon_{t+h}, y_t\right) = \\ &= \varphi \operatorname{Cov}(y_{t+h-1}, y_t) \\ &= \varphi \gamma_y(h-1) \\ &= \varphi^{|h|} \gamma_y(0) \ \text{(check for } h > 0 \ \text{and } h < 0) \\ &= \frac{\varphi^{|h|} \sigma^2}{1 - \varphi^2} \end{split}$$

SKOITECH Authono Protitate at Science and Technology 6/28

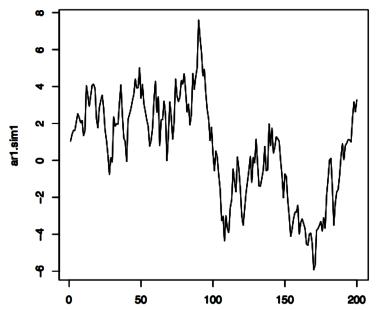
$$y_t = \varphi y_{t-1} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

Assume that y_t is stationary (i.e. $|\varphi| < 1$), then

$$\begin{split} \mathbb{E}[y_t] &= 0, \ \ \mathbb{E}[y_t^2] = \frac{\sigma^2}{1 - \varphi^2} \\ \gamma_y(h) &= \operatorname{Cov}\left(\varphi y_{t+h-1} + \varepsilon_{t+h}, y_t\right) = \\ &= \varphi \operatorname{Cov}(y_{t+h-1}, y_t) \\ &= \varphi \gamma_y(h-1) \\ &= \varphi^{|h|} \gamma_y(0) \ \ \text{(check for } h > 0 \ \text{and } h < 0) \\ &= \frac{\varphi^{|h|} \sigma^2}{1 - \varphi^2} \end{split}$$

Skoltech
Sations trattate of Science and Technology 6/28

AR(1): $\varphi = 0.95$



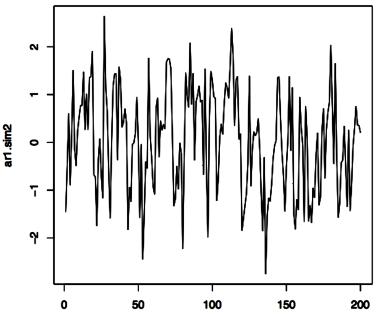
200

7/28

Burnaev, ML

Skoltech

AR(1): $\varphi = 0.5$



Burnaev, ML

Skoltech Skolou reduce of Science and Technology

8/28

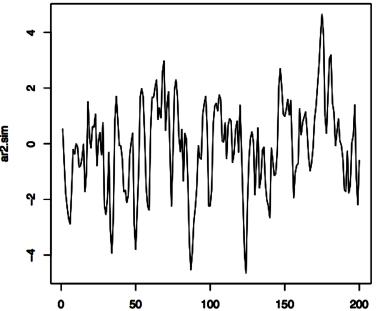
AR(2) process

• AR(2) process (AutoRegression)

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

Burnaev, ML Skoltech

AR(2): $\varphi_1=0.9$, $\varphi_2=0.2$



• AR(p) process (AutoRegression)

$$y_t = \sum_{i=1}^{p} \varphi_i y_{t-i} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

• Stationarity condition: all roots of the polynomial $\Phi(z)$ lie inside the unit circle, $|z_i|<1$,

$$\Phi(z) = 1 - \sum_{i=1}^{p} \varphi_i z^{p-i}$$

ロト (個) (注) (注) 注 の(の

tage of Science and Technology 11/28

AR(p) process

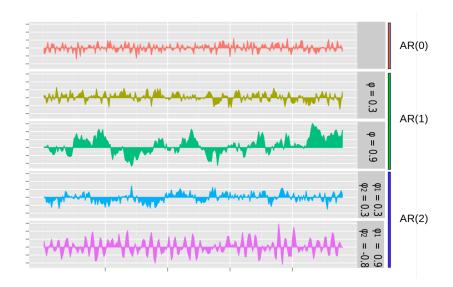
• AR(p) process (AutoRegression)

$$y_t = \sum_{i=1}^p \varphi_i y_{t-i} + \varepsilon_t, \ \{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$$

• Stationarity condition: all roots of the polynomial $\varPhi(z)$ lie inside the unit circle, $|z_i|<1$,

$$\Phi(z) = 1 - \sum_{i=1}^{p} \varphi_i z^{p-i}.$$

Institute of Science and Technology 11/28



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

$$y_t = \varepsilon_t + \theta \varepsilon_{t-1},$$

$$\{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2),$$

 $\{arepsilon_t\}$ is a white noise process. We have $\mathbb{E}[y_t]=0$ and

$$\begin{split} \gamma_y(t+h,t) &= \mathbb{E}(y_{t+h}y_t) \\ &= \mathbb{E}[\left(\varepsilon_{t+h} + \theta\varepsilon_{t+h-1}\right)\left(\varepsilon_t + \theta\varepsilon_{t-1}\right)] \\ &= \begin{cases} \sigma^2(1+\theta^2) & \text{if } h = 0, \\ \sigma^2\theta & \text{if } h = \pm 1, \\ 0 & \text{otherwise} \end{cases} \end{split}$$

• Thus $\{y_t\}_{t\geq 1}$ is stationary

STATE OF THE STATE

Example: MA(1) process (Moving Average)

$$y_t = \varepsilon_t + \theta \varepsilon_{t-1},$$

$$\{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2),$$

 $\{arepsilon_t\}$ is a white noise process. We have $\mathbb{E}[y_t]=0$ and

$$\begin{split} \gamma_y(t+h,t) &= \mathbb{E}(y_{t+h}y_t) \\ &= \mathbb{E}[(\varepsilon_{t+h} + \theta \varepsilon_{t+h-1}) \left(\varepsilon_t + \theta \varepsilon_{t-1}\right)] \\ &= \begin{cases} \sigma^2(1+\theta^2) & \text{if } h = 0, \\ \sigma^2\theta & \text{if } h = \pm 1, \\ 0 & \text{otherwise} \end{cases} \end{split}$$

• Thus $\{y_t\}_{t\geq 1}$ is stationary

Burnaev, ML Subscript of General Process 13/28

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

$$y_t = \varepsilon_t + \theta \varepsilon_{t-1},$$

$$\{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2),$$

 $\{arepsilon_t\}$ is a white noise process. We have $\mathbb{E}[y_t]=0$ and

$$\begin{split} \gamma_y(t+h,t) &= \mathbb{E}(y_{t+h}y_t) \\ &= \mathbb{E}[\left(\varepsilon_{t+h} + \theta\varepsilon_{t+h-1}\right)\left(\varepsilon_t + \theta\varepsilon_{t-1}\right)] \\ &= \begin{cases} \sigma^2(1+\theta^2) & \text{if } h = 0, \\ \sigma^2\theta & \text{if } h = \pm 1, \\ 0 & \text{otherwise} \end{cases} \end{split}$$

• Thus $\{y_t\}_{t\geq 1}$ is stationary

Burnaev, ML

Gushan wozzar et Convoice and To-Orology 13/28

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

$$y_t = \varepsilon_t + \theta \varepsilon_{t-1},$$

$$\{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2),$$

 $\{arepsilon_t\}$ is a white noise process. We have $\mathbb{E}[y_t]=0$ and

$$\begin{split} \gamma_y(t+h,t) &= \mathbb{E}(y_{t+h}y_t) \\ &= \mathbb{E}[(\varepsilon_{t+h} + \theta \varepsilon_{t+h-1}) \left(\varepsilon_t + \theta \varepsilon_{t-1}\right)] \\ &= \begin{cases} \sigma^2(1+\theta^2) & \text{if } h = 0, \\ \sigma^2\theta & \text{if } h = \pm 1, \\ 0 & \text{otherwise} \end{cases} \end{split}$$

• Thus $\{y_t\}_{t\geq 1}$ is stationary

Example: MA(1) process (Moving Average)

$$y_t = \varepsilon_t + \theta \varepsilon_{t-1},$$

$$\{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2),$$

 $\{arepsilon_t\}$ is a white noise process. We have $\mathbb{E}[y_t]=0$ and

$$\begin{split} \gamma_y(t+h,t) &= \mathbb{E}(y_{t+h}y_t) \\ &= \mathbb{E}[\left(\varepsilon_{t+h} + \theta\varepsilon_{t+h-1}\right)\left(\varepsilon_t + \theta\varepsilon_{t-1}\right)] \\ &= \begin{cases} \sigma^2(1+\theta^2) & \text{if } h = 0, \\ \sigma^2\theta & \text{if } h = \pm 1, \\ 0 & \text{otherwise} \end{cases} \end{split}$$

• Thus $\{y_t\}_{t\geq 1}$ is stationary

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

$$y_t = \varepsilon_t + \theta \varepsilon_{t-1},$$

$$\{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2),$$

 $\{arepsilon_t\}$ is a white noise process. We have $\mathbb{E}[y_t]=0$ and

$$\begin{split} \gamma_y(t+h,t) &= \mathbb{E}(y_{t+h}y_t) \\ &= \mathbb{E}[\left(\varepsilon_{t+h} + \theta\varepsilon_{t+h-1}\right)\left(\varepsilon_t + \theta\varepsilon_{t-1}\right)] \\ &= \begin{cases} \sigma^2(1+\theta^2) & \text{if } h = 0, \\ \sigma^2\theta & \text{if } h = \pm 1, \\ 0 & \text{otherwise} \end{cases} \end{split}$$

• Thus $\{y_t\}_{t\geq 1}$ is stationary

Skoltech

MA(q) process (Moving Average)

$$y_t = \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i},$$

$$\{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2),$$

- $\{\varepsilon_t\}$ is a white noise process.
- ullet For MA we have limit range of covariances between $y_t.$

Skoltech
Stations Insitate of Science and Technology 14/28

Sample autocovariance function

$$\widehat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (y_{t+|h|} - \overline{y})(y_t - \overline{y})$$

- ullet pprox sample covariance of $(y_1,y_{h+1}),\ldots,\,(y_{n-h},y_n)$, except that
 - we normalize by n instead of n-h, and
 - we subtract the full sample mear
- We estimate sample variance and obtain sample autocorrelation function

Skoltech

15/28

Sample autocovariance function

$$\widehat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (y_{t+|h|} - \overline{y})(y_t - \overline{y})$$

- $\bullet \approx$ sample covariance of $(y_1, y_{h+1}), \ldots, (y_{n-h}, y_n)$, except that
 - we normalize by n instead of n-h, and
 - we subtract the full sample mean
- We estimate sample variance and obtain sample autocorrelation function

Skoltech

15/28

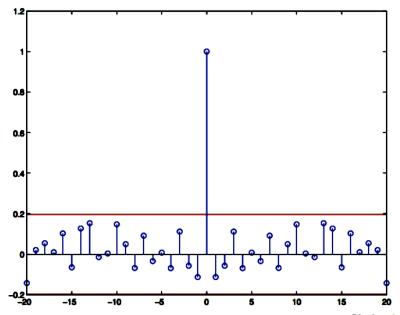
Sample autocovariance function

$$\widehat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (y_{t+|h|} - \overline{y})(y_t - \overline{y})$$

- $\bullet \approx$ sample covariance of $(y_1, y_{h+1}), \ldots, (y_{n-h}, y_n)$, except that
 - we normalize by n instead of n-h, and
 - we subtract the full sample mean
- We estimate sample variance and obtain sample autocorrelation function

Skoltech

Sample ACF for Gaussian noise



Summary for sample ACF

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

0	Lime	series:

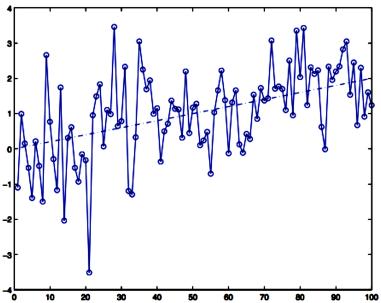
- White
- Trend
- Periodic
- MA(q)
- AR(p)

Sample ACF:

- ightarrow zero
- \rightarrow Slow decay
- \rightarrow Periodic
- \rightarrow Zero for |h| > q
- $\rightarrow\,$ Decays to zero exponentially

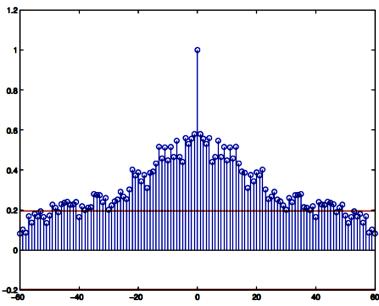
Skoltech

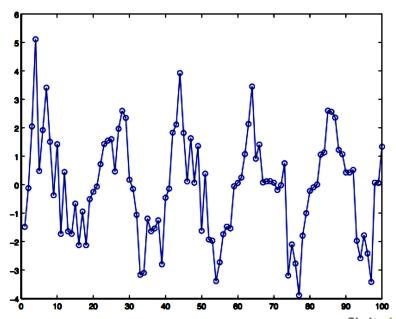
17/28



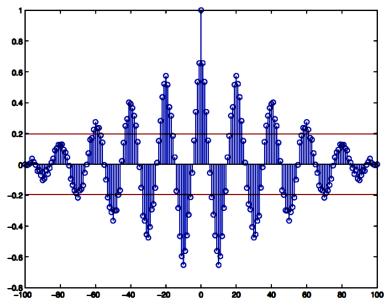
Skoltech Sustana trattare et Science and Technology

Sample ACF: Trend

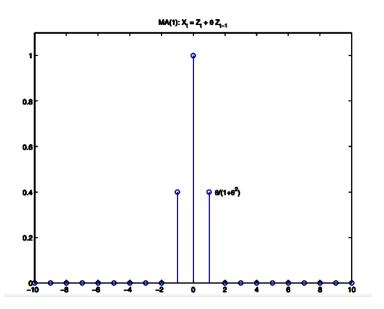


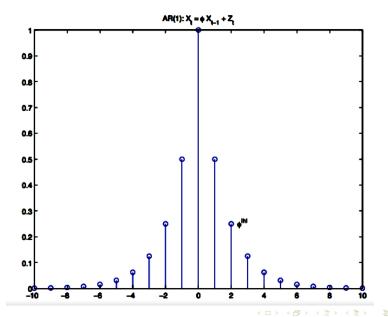


Sample ACF: Trend



ACF: MA(1)





Outline

- Auto Regression (AR)
- 2 Moving average
- 3 Autocovariance function
- 4 Auto Regression (ARMA) process

• An **ARMA(p,q) process** $\{y_t\}_{t\geq 1}$ is a stationary process that satisfies

$$y_t-\varphi_1y_{t-1}-\ldots-\varphi_py_{t-p}=\varepsilon_t+\theta_1\varepsilon_{t-1}+\ldots\theta_q\varepsilon_{t-q},$$
 where $\{\varepsilon_t\}_{t\geq 1}\sim WN(0,\sigma^2)$

• Given n observations, in case of AR(p) process the parameters can be estimated by least-squares

$$\widehat{\varphi} = \arg\min_{\varphi} \sum_{t=p+1}^{n} [y_t - \varphi_1 y_{t-1} - \dots - \varphi_p y_{t-p}]^2$$

In matrix form for

$$\mathbf{X} = \begin{bmatrix} y_{n-1} & y_{n-2} & \dots & y_{n-p-1} \\ y_{n-2} & y_{n-3} & \dots & y_{n-p-2} \\ \vdots & \vdots & \vdots & \vdots \\ y_p & y_{p-1} & \dots & y_1 \end{bmatrix}, \ \mathbf{y} = \begin{bmatrix} y_n \\ y_{n-1} \\ \vdots \\ y_{p+1} \end{bmatrix}$$

then

$$\widehat{\varphi} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}$$



• An **ARMA(p,q) process** $\{y_t\}_{t\geq 1}$ is a stationary process that satisfies

$$y_t-\varphi_1y_{t-1}-\ldots-\varphi_py_{t-p}=\varepsilon_t+\theta_1\varepsilon_{t-1}+\ldots\theta_q\varepsilon_{t-q},$$
 where $\{\varepsilon_t\}_{t\geq 1}\sim WN(0,\sigma^2)$

 Given n observations, in case of AR(p) process the parameters can be estimated by least-squares

$$\widehat{\varphi} = \arg\min_{\varphi} \sum_{t=p+1}^{n} [y_t - \varphi_1 y_{t-1} - \dots - \varphi_p y_{t-p}]^2$$

In matrix form for

$$\mathbf{X} = \begin{bmatrix} y_{n-1} & y_{n-2} & \cdots & y_{n-p-1} \\ y_{n-2} & y_{n-3} & \cdots & y_{n-p-2} \\ \vdots & \vdots & \vdots & \vdots \\ y_p & y_{p-1} & \cdots & y_1 \end{bmatrix}, \ \mathbf{y} = \begin{bmatrix} y_n \\ y_{n-1} \\ \vdots \\ y_{p+1} \end{bmatrix}$$

ther

$$\widehat{\varphi} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}$$



• An **ARMA(p,q) process** $\{y_t\}_{t\geq 1}$ is a stationary process that satisfies

$$y_t - \varphi_1 y_{t-1} - \ldots - \varphi_p y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots \theta_q \varepsilon_{t-q},$$
 where $\{\varepsilon_t\}_{t \ge 1} \sim WN(0, \sigma^2)$

 Given n observations, in case of AR(p) process the parameters can be estimated by least-squares

$$\widehat{\varphi} = \arg\min_{\varphi} \sum_{t=p+1}^{n} [y_t - \varphi_1 y_{t-1} - \dots - \varphi_p y_{t-p}]^2$$

In matrix form for

$$\mathbf{X} = \begin{bmatrix} y_{n-1} & y_{n-2} & \dots & y_{n-p-1} \\ y_{n-2} & y_{n-3} & \dots & y_{n-p-2} \\ \vdots & \vdots & \vdots & \vdots \\ y_p & y_{p-1} & \dots & y_1 \end{bmatrix}, \ \mathbf{y} = \begin{bmatrix} y_n \\ y_{n-1} \\ \vdots \\ y_{p+1} \end{bmatrix}$$

then

$$\widehat{\varphi} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}$$



SKOITECH Authoric Fredrickste et Science and Technology 25/28

- AR models assume that the relation between past and future is linear
- Nonlinear Auto Regressive (NAR) formulation

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}) + \varepsilon_t$$

where the missing information is lumped into a noise term $arepsilon_t$

 We will consider this relationship as a particular instance of a dependence

$$y_t = f(\mathbf{x}_t) + \varepsilon, \ y_t \in \mathbb{R}^1, \ \mathbf{x}_t \in \mathbb{R}^p,$$

where

$$\mathbf{x}_t = [y_{t-1}, y_{t-2}, \dots, y_{t-p}]$$

• We train $f(\cdot)$ using ML regression algorithm and a sample (\mathbf{X}, \mathbf{y}) , where

$$\mathbf{X} = \{\mathbf{x}_{p+1}, \mathbf{x}_{p+2}, \dots, \mathbf{x}_T\}$$

and

$$\mathbf{y} = \{y_{n+1}, y_{n+2}, \dots, y_T\}$$

Skoltech

Burnaev, ML

- AR models assume that the relation between past and future is linear
- Nonlinear Auto Regressive (NAR) formulation

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}) + \varepsilon_t,$$

where the missing information is lumped into a noise term ε_t

 We will consider this relationship as a particular instance of a dependence

$$y_t = f(\mathbf{x}_t) + \varepsilon, \ y_t \in \mathbb{R}^1, \ \mathbf{x}_t \in \mathbb{R}^p,$$

where

$$\mathbf{x}_t = [y_{t-1}, y_{t-2}, \dots, y_{t-p}]$$

• We train $f(\cdot)$ using ML regression algorithm and a sample (\mathbf{X}, \mathbf{y}) , where

$$\mathbf{X} = \{\mathbf{x}_{p+1}, \mathbf{x}_{p+2}, \dots, \mathbf{x}_T\}$$

and

$$\mathbf{y} = \{y_{n+1}, y_{n+2}, \dots, y_T\}$$

- AR models assume that the relation between past and future is linear
- Nonlinear Auto Regressive (NAR) formulation

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}) + \varepsilon_t,$$

where the missing information is lumped into a noise term ε_t

 We will consider this relationship as a particular instance of a dependence

$$y_t = f(\mathbf{x}_t) + \varepsilon, \ y_t \in \mathbb{R}^1, \ \mathbf{x}_t \in \mathbb{R}^p,$$

where

$$\mathbf{x}_t = [y_{t-1}, y_{t-2}, \dots, y_{t-p}]$$

• We train $f(\cdot)$ using ML regression algorithm and a sample (\mathbf{X}, \mathbf{y}) , where

$$\mathbf{X} = \{\mathbf{x}_{p+1}, \mathbf{x}_{p+2}, \dots, \mathbf{x}_T\}$$

and

$$\mathbf{y} = \{y_{n+1}, y_{n+2}, \dots, y_T\}$$

- AR models assume that the relation between past and future is linear
- Nonlinear Auto Regressive (NAR) formulation

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}) + \varepsilon_t,$$

where the missing information is lumped into a noise term $arepsilon_t$

 We will consider this relationship as a particular instance of a dependence

$$y_t = f(\mathbf{x}_t) + \varepsilon, \ y_t \in \mathbb{R}^1, \ \mathbf{x}_t \in \mathbb{R}^p,$$

where

$$\mathbf{x}_t = [y_{t-1}, y_{t-2}, \dots, y_{t-p}]$$

• We train $f(\cdot)$ using ML regression algorithm and a sample (\mathbf{X}, \mathbf{y}) , where

$$\mathbf{X} = \{\mathbf{x}_{p+1}, \mathbf{x}_{p+2}, \dots, \mathbf{x}_T\}$$

and

$$\mathbf{y} = \{y_{n+1}, y_{n+2}, \dots, y_T\}$$

26/28

There is a zoo of related models

- ARIMA model differences $y_t y_{t-1}$ instead of y_t
- SARIMA takes into account seasonality
- ARCH autoregression with conditional heterscedasticity
- GARCH generalized ARCH

Skoltech

Conclusions

- Autoregression (AR) and Moving average (MA) models describe different types of relations between neighbour observations of time series
- For autoregression autocorrelations are non-zero for all differences
- For Moving average autocorrelations are zero (and we specify the order to define the range of dependence)
- ARIMA model unites AR and MA models
- Using Autocorrelation function we can identify the right model ARMA(p, q)

Skoltech