Day 10 - Autoregressive models

Introduction

We continue our discussion of basic stochastic models, introducing the generating autoregressive model of order p.

```
# packages
library(tidyverse)
library(lubridate)
```

Autoregressive models

Note

A time series $\{x_t\}$ is an ____ or AR(p) if

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \cdots + \alpha_p x_{t-p} + w_t$$

where $\{w_t\}$ is white noise and the α_i are the model parameters. It can be shown that the AR(p) model can be expressed as a polynomial of order p in the backshift operator:

$$\theta_p(B)x_t = (1 - \alpha_1 B - \alpha_2 B^2 - \cdots - \alpha_p B^p)x_t = w_t$$

A few notes on AR(p) models:

- The ______________________________ is a special case of AR(1) with $\alpha_1 = 1$.

- The _______________________________________ model is a special case of an AR process with $\alpha_i = \alpha(1-\alpha)^i$ as p approaches infinity.

- A prediction at time t is given by

$$\hat{x}_t = \hat{\alpha}_1 x_{t-1} + \hat{\alpha}_2 x_{t-2} + \cdots + \hat{\alpha}_p x_{t-p}$$

- Model parameters are estimated by minimizing the _____________________________.

Is an AR(p) process stationary?

Note

The equation $\theta_p(B) = 0$ is called the __________________________________. The roots of the __________________________________ may be used to determine whether an AR(p) process is stationary.

If all roots of the __________________________________ exceed 1 in magnitude, the model is stationary. You may use the `polyroot` function in R to find the roots of polynomials.

Example: Determine whether the AR(1) model $x_t = \frac{1}{2}x_{t-1} + w_t$ is stationary.

Example: Determine whether the AR(2) model $x_t = x_{t-1} - \frac{1}{4}x_{t-2} + w_t$ is stationary.

Example: Determine whether the AR(2) model $x_t = \frac{1}{2}x_{t-1} - \frac{1}{2}x_{t-2} + w_t$ is stationary.

AR(1) processes

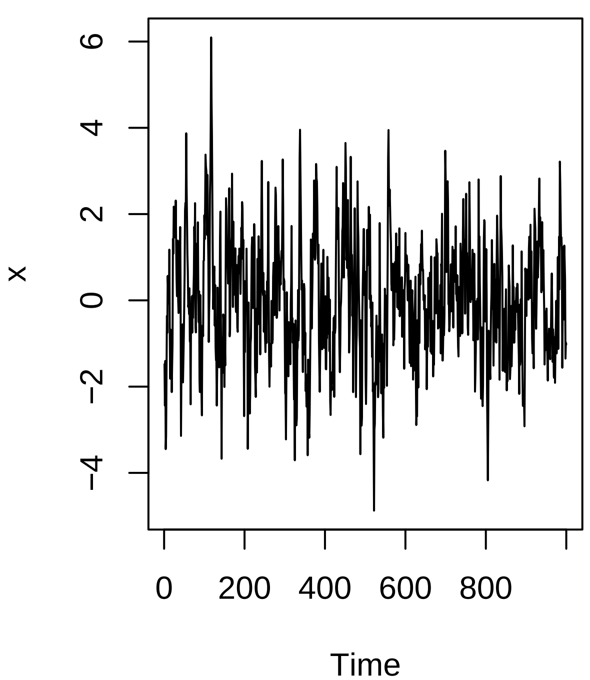
A time series x_t is an AR(1) process if

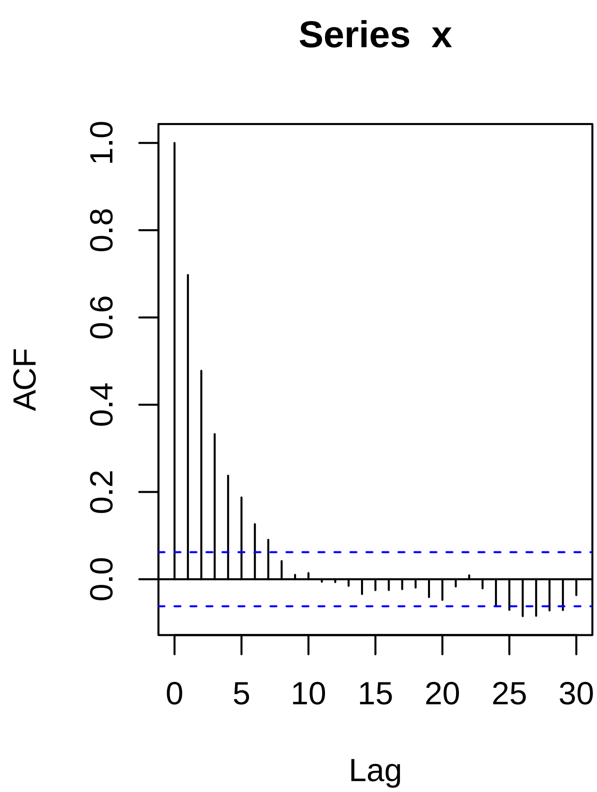
The second-order properties of an AR(1) process are:

$$\mu(t) = 0$$
$$\gamma_k = \frac{\alpha^k \sigma^2}{(1-\alpha^2)}$$
$$\rho_k = \alpha^k$$

How can we simulate an AR(1) process?

```r
set.seed(10062024)
x <- w <- rnorm(1000)
for(t in 2:1000) x[t] <- 0.7 * x[t-1] + w[t]
par(mfrow = c(1, 2))
plot(x, type = "l", xlab = "Time")
acf(x)
```

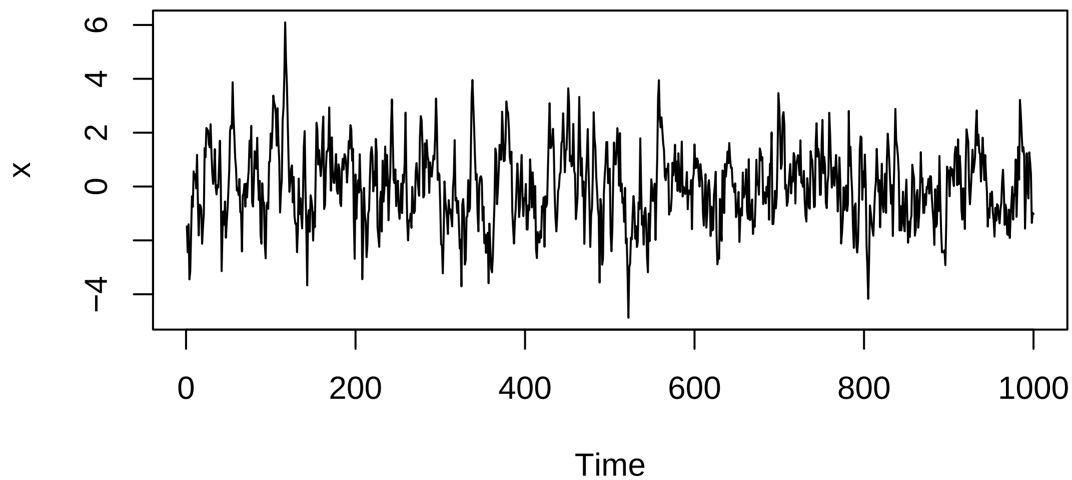


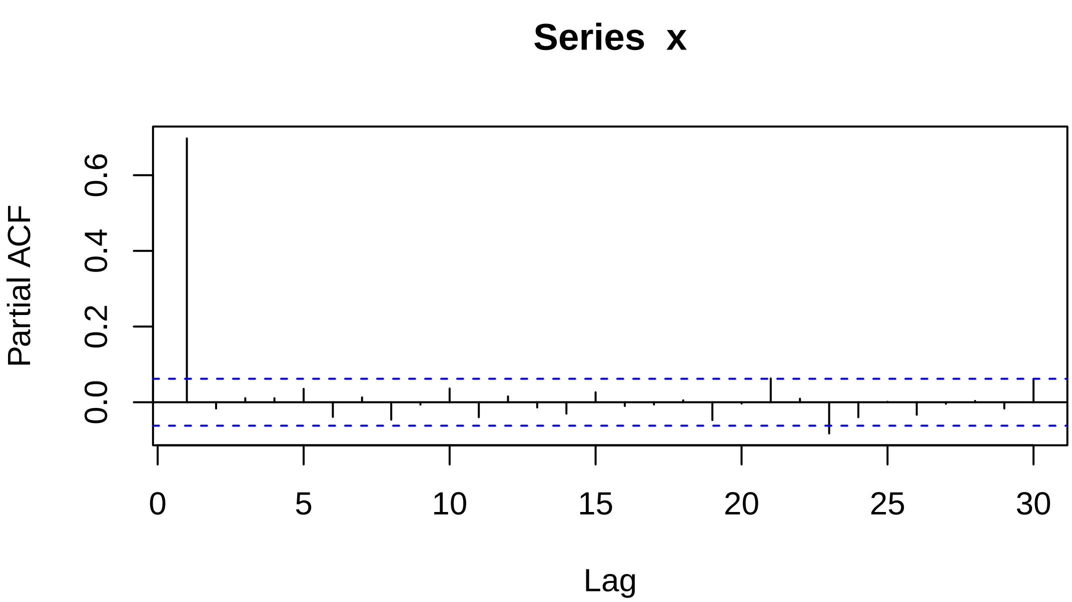


Note

The ____ at lag k is the correlation that results after removing the effect of any correlation due to terms at shorter lags.

```
par(mfrow = c(2, 1))
plot(x, type = "l", xlab = "Time")
pacf(x)
```





Fitting an AR(p) process

To fit an AR(p) process, we use the `ar` function in R. To select the order of the AR process, R minimizes the AIC.

$$AIC = 2 \cdot (-\text{log-likelihood} + \text{number of parameters})$$

```r
par(mfrow = c(1,1))
ar_fit <- ar(x)
str(ar_fit)
```

```
List of 15
 $ order       : int 1
 $ ar          : num 0.697
 $ var.pred    : num 1.03
 $ x.mean      : num 0.00372
 $ aic         : Named num [1:31] 664.22 0 1.72 3.6 5.49 ...
  ..- attr(*, "names")= chr [1:31] "0" "1" "2" "3" ...
 $ n.used      : int 1000
 $ n.obs       : int 1000
 $ order.max   : num 30
 $ partialacf  : num [1:30, 1, 1] 0.6974 -0.0167 0.0108 0.0107 0.0355 ...
 $ resid       : num [1:1000] NA -1.411 0.298 -2.471 -0.748 ...
 $ method      : chr "Yule-Walker"
 $ series      : chr "x"
 $ frequency   : num 1
 $ call        : language ar(x = x)
 $ asy.var.coef: num [1, 1] 0.000515
 - attr(*, "class")= chr "ar"
```

```r
ar_fit$aic
```

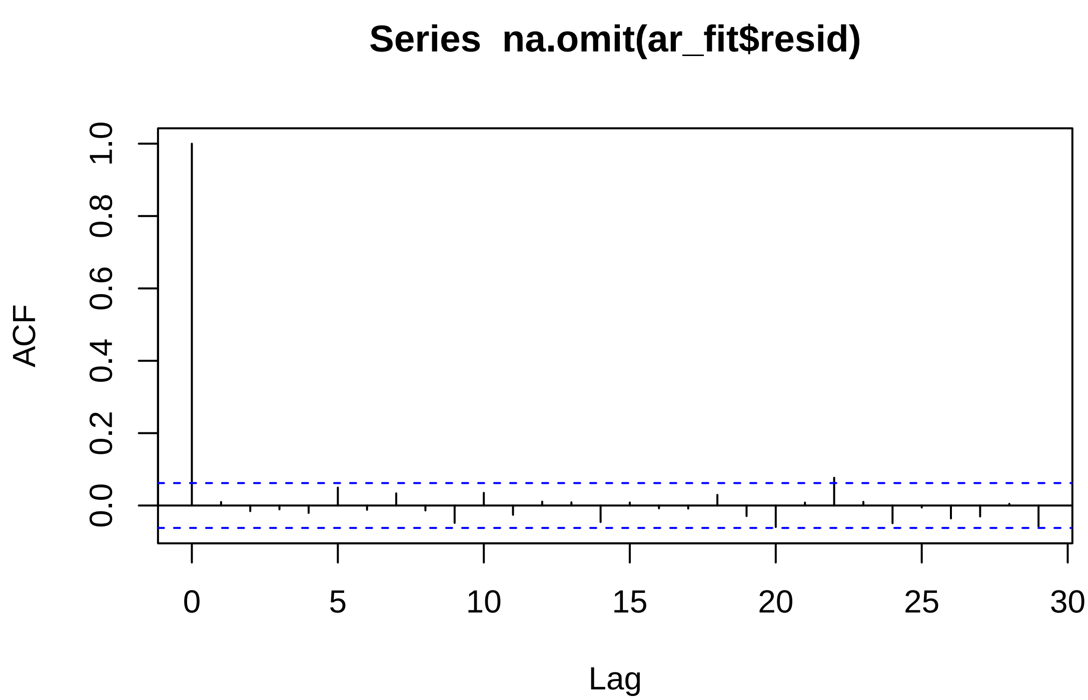
```
         0          1          2          3          4          5          6 
664.217434   0.000000   1.720656   3.603433   5.489964   6.232324   6.709250 
         7          8          9         10         11         12         13 
  8.541464   8.401815  10.359473  11.043122  11.467784  13.227779  15.034625 
        14         15         16         17         18         19         20 
 16.127797  17.433967  19.334440  21.296706  23.269581  23.045314  25.032714 
        21         22         23         24         25         26         27 
 23.063698  24.975371  20.131901  20.554231  22.552658  23.453832  25.436742 
        28         29         30 
 27.424293  29.143135  27.388364 
```

```r
predict(ar_fit, n.ahead = 10)
```

```
$pred
Time Series:
Start = 1001
End = 1010
Frequency = 1
 [1] -0.70807211 -0.49267756 -0.34246365 -0.23770604 -0.16464917 -0.11370008
 [7] -0.07816872 -0.05338951 -0.03610876 -0.02405734

$se
Time Series:
Start = 1001
End = 1010
Frequency = 1
 [1] 1.015897 1.238542 1.333456 1.377255 1.398061 1.408069 1.412911 1.415260
 [9] 1.416401 1.416955
```

```r
acf(na.omit(ar_fit$resid))
```



Closing remarks

- The stochastic models discussed this week (white noise, random walks, and AR(p) processes) are not very useful for forecasting on their own, and are unlikely to compete with procedures like Holt-Winters.
- In practice, these stochastic models are combined with other techniques (regression, moving average, and integrated moving averages) to construct powerful forcasting techniques.
- The AR(p) process is one component of the ARIMA model, which we will discuss towards the end of the semester.