Linear Filters

Introduction

Linear Filters

Filter

Setting

$$Z_t \longrightarrow \boxed{\text{Filter}} \longrightarrow X_t$$

(Input) (Output)

Examples

Difference: $X_t = Z_t - Z_{t-1}$

5-point moving average smoother: $X_t = \dfrac{Z_{t-2} + Z_{t-1} + Z_t + Z_{t+1} + Z_{t+2}}{5}$

Filter Example:
Difference

Consider the series: $Z_1 = 8, Z_2 = 14, Z_3 = 14, Z_4 = 7$

The difference Filter

$$X_t = Z_t - Z_{t-1}$$

$$X_1 = Z_1 - Z_0 \;\; (Z_0 = ?)$$

$$X_2 = Z_2 - Z_1$$

$$X_3 = Z_3 - Z_2$$

$$\vdots$$

$$X_n = Z_n - Z_{n-1}$$

Differenced data

$$X_1 = Z_1 \quad Z_0 = ??$$

$$X_2 = Z_2 \quad Z_1 = 14 \quad 8 = 6$$

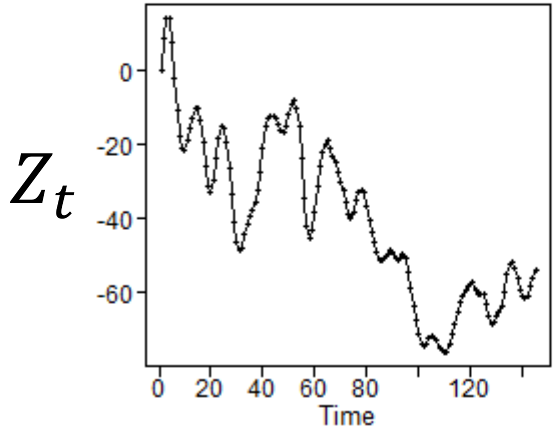
$$X_3 = Z_3 \quad Z_2 = 14 \quad 14 = 0$$

$$X_4 = Z_4 \quad Z_3 = 7 \quad 14 = \quad 7$$

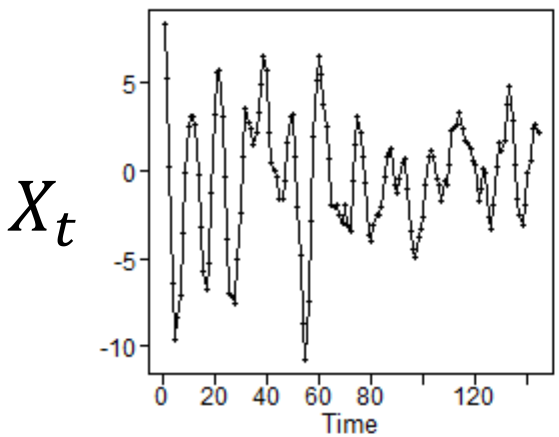
Note:

The differenced data (Z_t) are a realization of length $n-1$.

Original data



Differenced data



Difference removes long-term trending behavior and retains a (monthly) cyclic pattern.

Filter Example: 5-Point Moving Average

Consider the series:

$Z_1 = 8, Z_2 = 14, Z_3 = 14, Z_4 = 7, Z_5 = 1, Z_6 = 10, Z_7 = 9, Z_8 = 2$

$$X_t = \frac{Z_{t-2} + Z_{t-1} + Z_t + Z_{t+1} + Z_{t+2}}{5}$$

The 5-point moving average filter

$X_1 = ?$

$X_2 = ?$

$$X_3 = \frac{8 + 14 + 14 + 7 + 1}{5} = 8.8$$

$$X_4 = \frac{14 + 14 + 7 + 1 + 10}{5} = 9.2$$

$$X_5 = \frac{14 + 7 + 1 + 10 + 9}{5} = 8.2$$

$$X_6 = \frac{7 + 1 + 10 + 9 + 2}{5} = 8.2$$

$X_7 = ?$

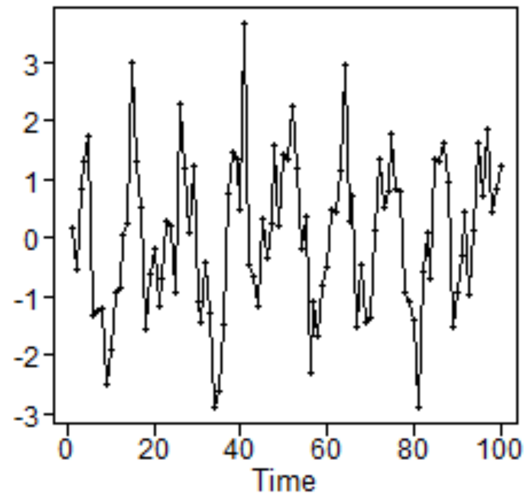
$X_8 = ?$

Note:

The differenced data (Z_t) are a realization of length n–4.

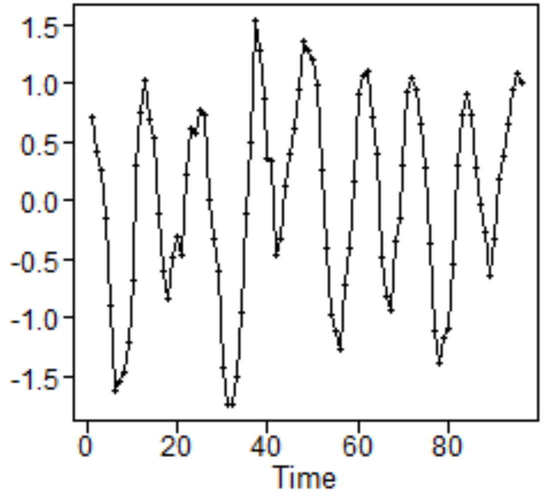
Sine + noise signal

Z_t



Signal above after applying 5-point moving average smoother

X_t



Filters are sometimes used to "filter out" certain frequencies from a set of data.

Types of filters

Low-pass filters:

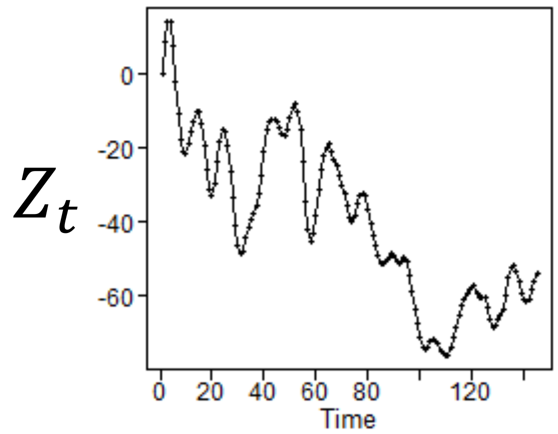
- "Pass" low-frequency behavior and "filter out" higher-frequency behavior
- The 5-point moving average smoother is a low-pass filter

High-pass filters:

- "Pass" high-frequency behavior and "filter out" lower-frequency behavior
- A difference is a high-pass filter

See text for other types of filters.

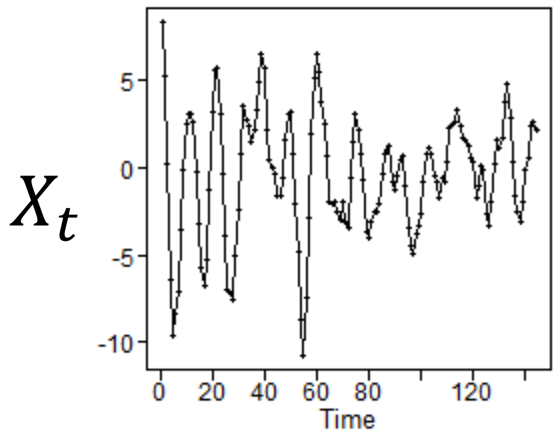
Original data



Low or high pass?

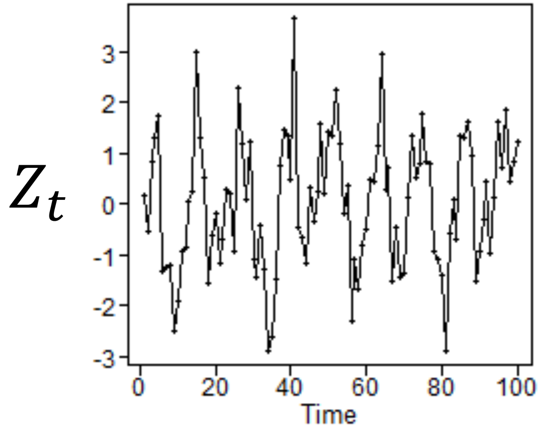
High pass!

Differenced data



Difference removes long-term trending behavior and retains a (monthly) cyclic pattern.

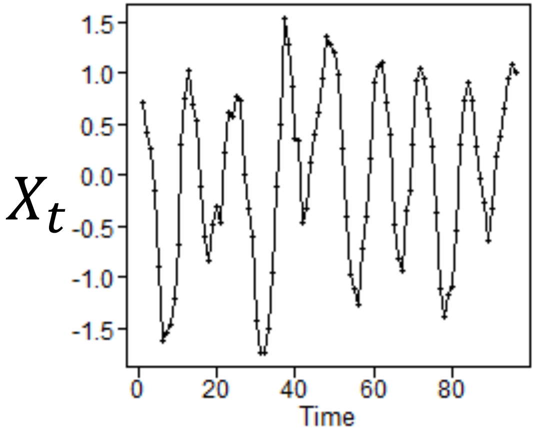
Sine + noise Signal



Low or high pass?

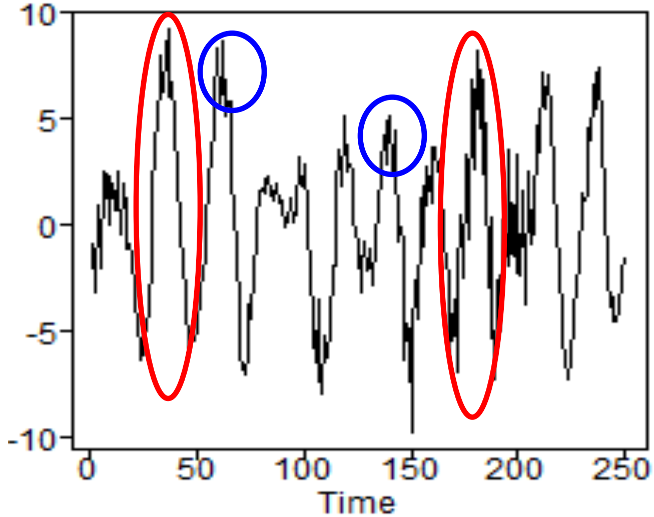
Low pass!

Signal above after applying 5-point moving average smoother

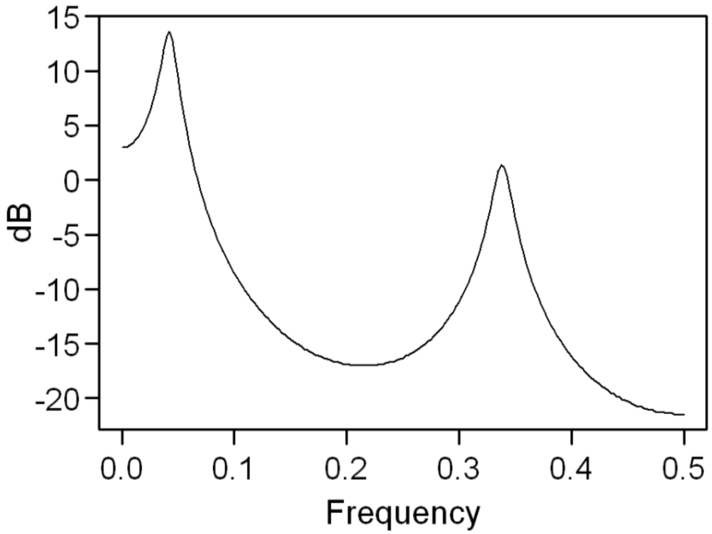


The 5-point moving average filter removes high-frequency behavior and retains the lower frequency signal.

Recall: Data with Two Frequencies

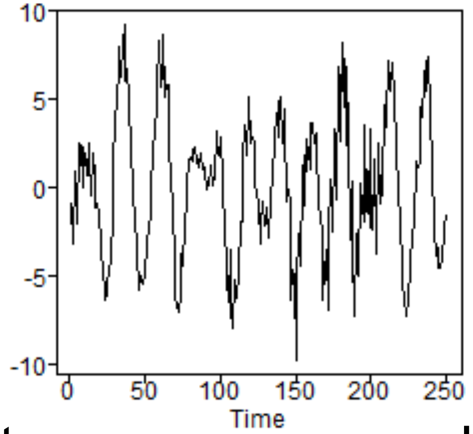


Spectral density showing frequencies at about **.05** and **.35**



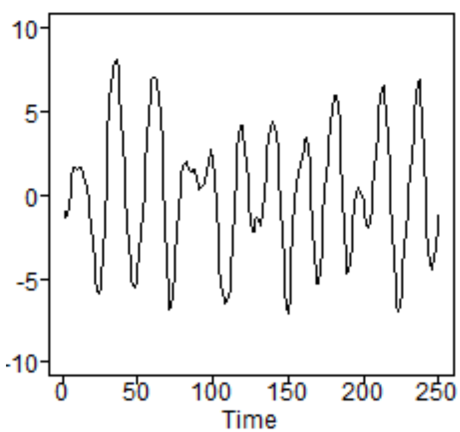
That is, the time series has cycles of length about **20** and **3.**

Original Data Containing Two Frequencies



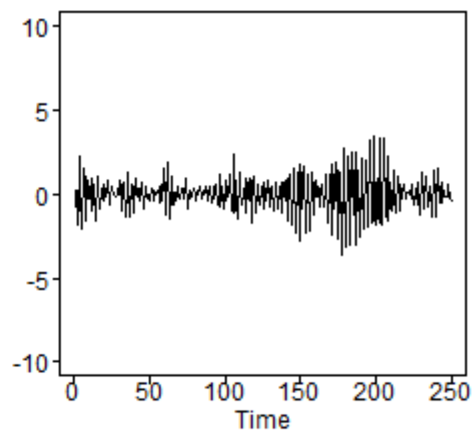
Data (Fig. 1.21a)

Low-pass filtered data



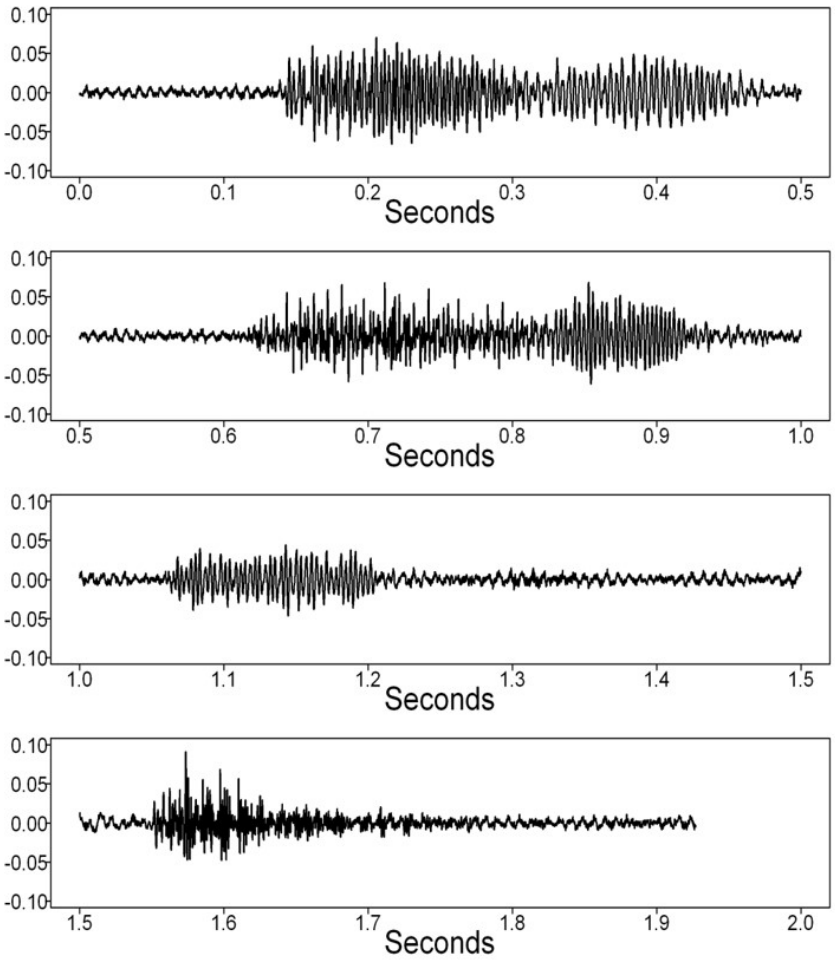
ma = filter(fig1.21a,rep(1,5))/5
plot(ma,type = "l")

High-pass filtered data

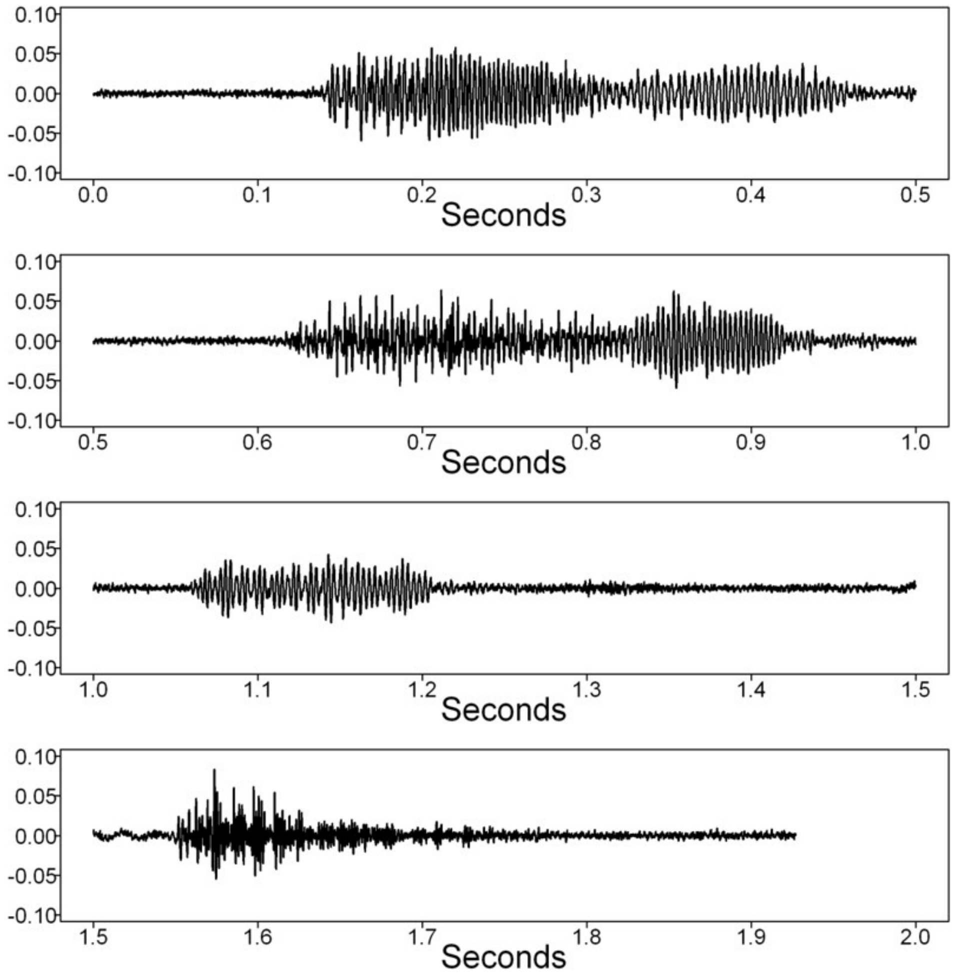


dif = dif(fig1.21a,lag = 1)
plot(dif,type = "l")

King Kong eats grass (with fan noise)

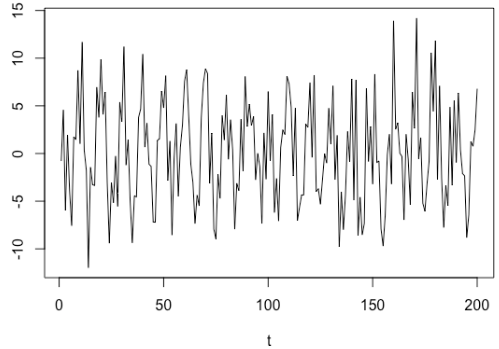


Filtered KK eats grass (fan noise filtered out)



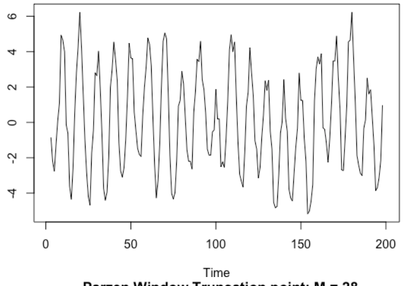
Let's Try One!

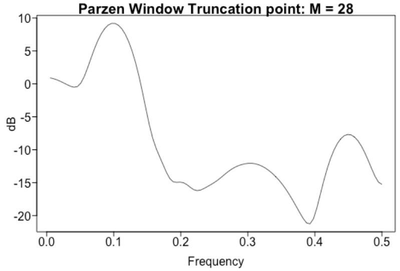
Given the plot of the realization and the plot of the filtered data below, which was the result of the low-pass filter?



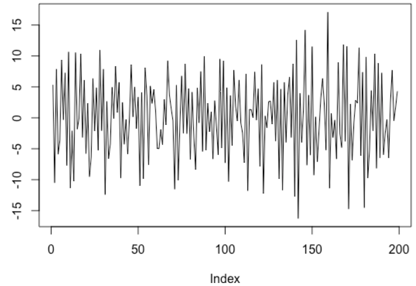
Realization = gen.sigplusnoise.wge(200,coef = c(5,3),freq = c(.1,.45), vara = 10, sn = 1)

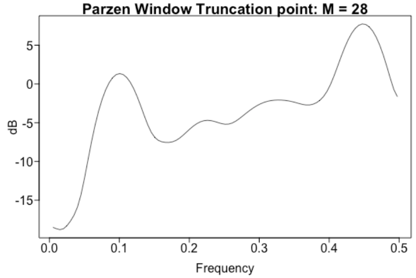
A





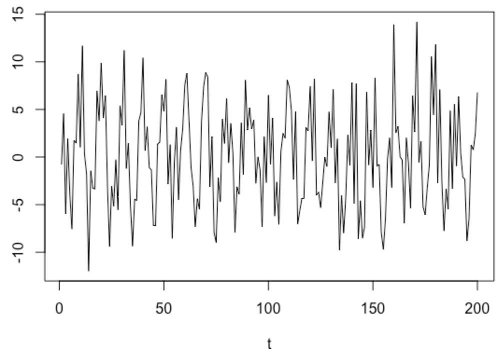
B





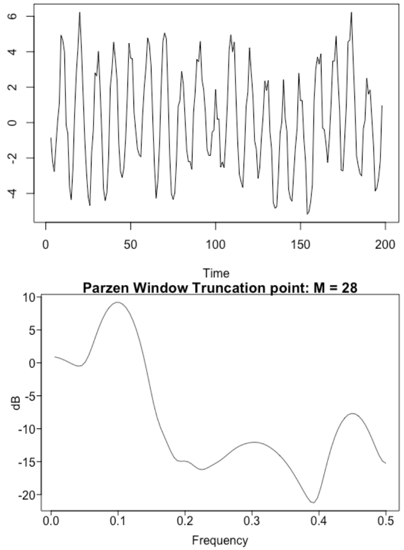
Let's Try One!

Given the plot of the realization and the plot of the filtered data below, which was the result of the low-pass filter?



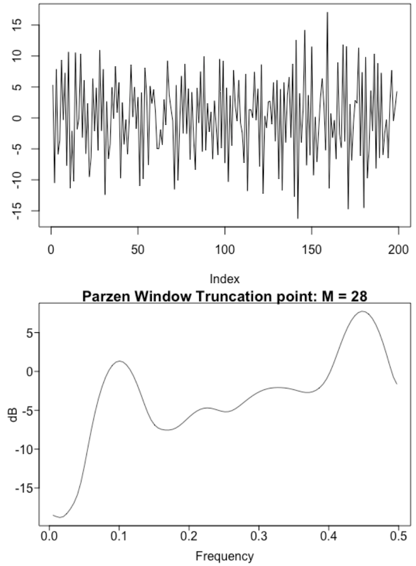
Realization = gen.sigplusnoise.wge(200,coef = c(5,3),freq = c(.1,.45), vara = 10, sn = 1)

A



```
ma = filter(Realization,rep(1,5))/5
plot(ma,type = "l")
parzen.wge(ma[is.na(ma)])
```

B



```
dif = diff(Realization,lag  = 1)
plot(dif,type = "l")
parzen.wge(dif[is.na(dif)])
```

DataScience@SMU

DataScience@SMU

General Linear Process (GLP)

Possibly real-world example

- Mikio slide with “filtering”
- Feature engineering/cleaning
- Response cleaning?
- Wayne had another idea I forgot about. (It may have been the response cleaning.)
- Maybe an interview if we can find one. Greg and Smart cover?

Also… live session… introduce Butterworth Filter exercise.

General Linear Process (GLP)

Before discussing autoregressive (AR) processes, moving average (MA), and autoregressive-moving average (ARMA) processes, we briefly introduce the ***general linear process.***

A GLP is a linear filter with ***white noise input.***

$$a_t \longrightarrow \boxed{\text{Filter}} \longrightarrow X_t$$

(Input) (Output)

$$\sum_{j=0}^{\infty} \psi_j a_{t-j} = X_t - \mu$$

General Linear Process (GLP)

A GLP is an infinite linear combination of white-noise terms.

- Expressed mathematically as:

$$X_t - \mu = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

$$X_t \longleftarrow \boxed{\text{Filter}} \longleftarrow a_t$$

- where the ψ_j's are called the ψ-weights

- AR, MA, and ARMA processes are all special cases of GLPs.
- This will be very useful later when we study confidence intervals for forecasts.
- See Woodward et al. (2017) for details.

DataScience@SMU

AR(1) | Introduction and Key Result

Autoregressive (AR) Models

In this unit, we will introduce and study the well-known autoregressive (AR) models.

Possibly Interview Here **(or somewhere soon)**? Describing the use in industry of Autoregressive models.

Also…Shall we mention that the AR is a Linear Filter? add a graphic?

$$Z_t \longrightarrow \boxed{\text{AR Model}} \longrightarrow X_t$$

(Input) (Output)

Autoregressive Models

AR(p) p = 1

- In this unit, we will introduce and study the well-known autoregressive (AR) models.
- Autoregressive models tend to be very useful for describing stationary data that move forward in time.
- We start with the simplest AR model, the AR(1).

Autoregressive Model of Order 1: AR(1)

$$\boxed{X_t = \beta + \varphi_1 X_{t-1} + a_t} \qquad \text{where } \beta = (1-\varphi_1)\mu$$

Notes:

- β is called the *moving average constant.*
- The AR(1) model says that the value of the process at time t (i.e., X_t) depends on the value of the process at time $t-1$ plus a random noise component (and a constant).
- This is a sensible way to describe the way a time series might progress in time.
- This is similar to the simple linear regression model, but in this case, the “independent variable” is a value of the dependent variable at a prior time period.

AR(1) model: $X_t = (1-\varphi_1)\mu + \varphi_1 X_{t-1} + a_t$

Key result

An AR(1) process is stationary if and only if $|\varphi_1| < 1$.

Examples

Stationary

$$X_t = 10 + .8X_{t-1} + a_t$$

$$X_t = 19 + .4X_{t-1} + a_t$$

***Not* stationary**

$$X_t = 10 + 1.4X_{t-1} + a_t$$

$$X_t = 12 + 1X_{t-1} + a_t$$

DataScience@SMU

AR(1) | Conditions 2 and 3

Variance and Autocorrelations

Stationary AR(1): Conditions 1, 2, and 3

Condition 1: Expected value

$$E[X_t] = \mu$$

Mean does not depend on *t*.

Condition 2: Variance

$$\sigma_X^2 = \frac{\sigma_a^2}{1-\varphi_1^2}$$

Note that the variance is finite as long as $|\phi_1| < 1$, and also note that it does not depend on *t*!

Condition 3: Autocorrelations

$$\rho_k = \varphi_1^k, k \geq 0$$

Note that the ρ_k decreases exponentially with *k* and only depends on *k* and not *t*.

Bonus: Spectral density

$$S_X(f) = \frac{\sigma_a^2}{\sigma_X^2}\left(\frac{1}{\left|1-\varphi_1 e^{-2\pi i f}\right|^2}\right)$$

Monotonically increasing or decreasing in *f* depending on φ_1

DataScience@SMU

AR(1) | Zero Mean Form

Stationary AR(1) Facts

2. Note that if $\quad = 0,$ the AR(1) model takes the form

$$X_t = \ {}_1X_{t \ \ 1} + a_t$$

- this is called the "zero mean" form of the model, and we sometimes use it for convenience

- the "zero-mean" form of the model is sometimes written in the alternative form

$$X_t - \varphi_1 X_{t-1} = a_t$$

DataScience@SMU

AR(1) | Backshift Operator

Stationary AR(1) Facts

3. The AR(1) is sometimes expressed using the ***backshift operator*** defined by:

$$BX_t = X_{t-1}$$

Using the backshift operator, the model

$$X_t - \varphi_1 X_{t-1} = a_t$$

can be written as

$$X_t - \varphi_1 B X_t = a_t$$

or $$(1 - \varphi_1 B) X_t = a_t$$

or $$\varphi(B) X_t = a_t$$

where $\varphi(B)$ is the operator

$$\varphi(B) = 1 - \varphi_1 B$$

DataScience@SMU

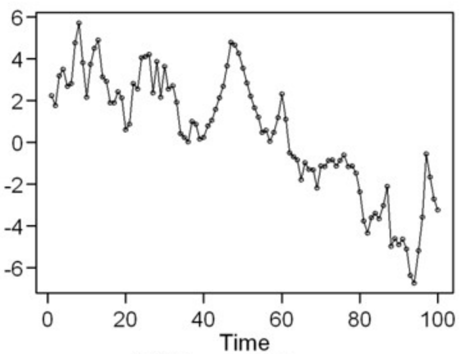
AR(1)

φ_1 Positive

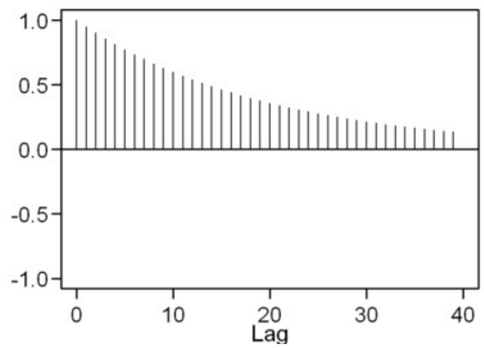
AR(1) Models: φ_1 Positive

$$\boxed{\varphi_1 = .95, \text{ i.e. } X_t - .95X_{t-1} = a_t}$$

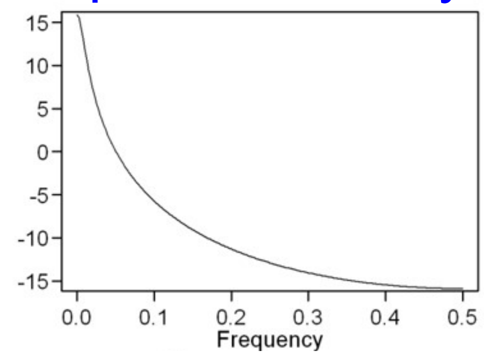
Data



Autocorrelations

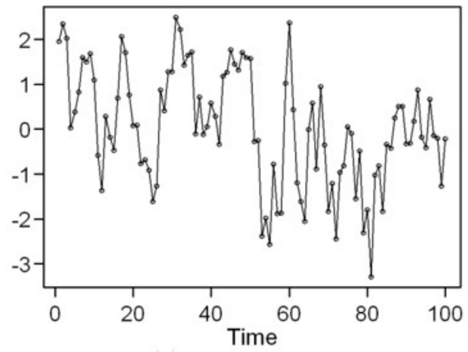


Spectral density

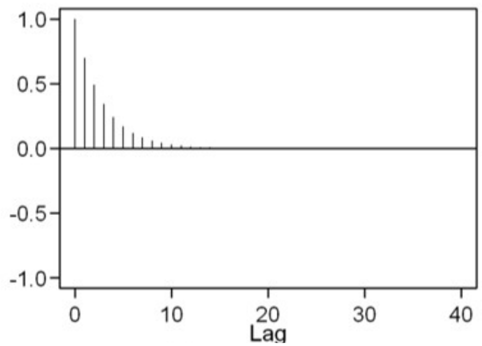


$$\boxed{\varphi_1 = .7, \text{ i.e. } X_t - .7X_{t-1} = a_t}$$

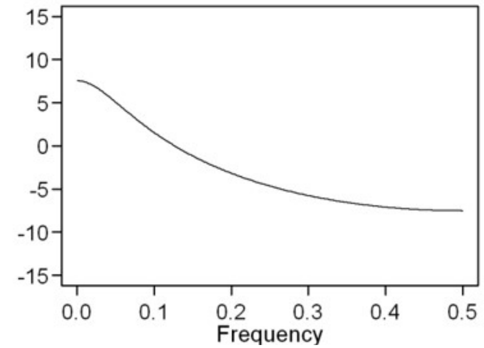
Data



Autocorrelations



Spectral density



AR(1) Models: φ_1 Positive

Summary

Realizations seem to be “wandering,” aperiodic in nature.

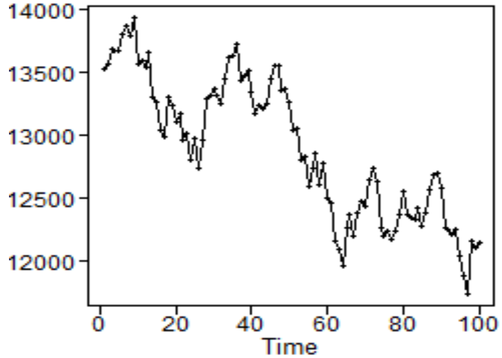
Autocorrelations are damped exponentials.

Spectral densities have peaks at $f = 0$, which is consistent with the behavior of the realizations.

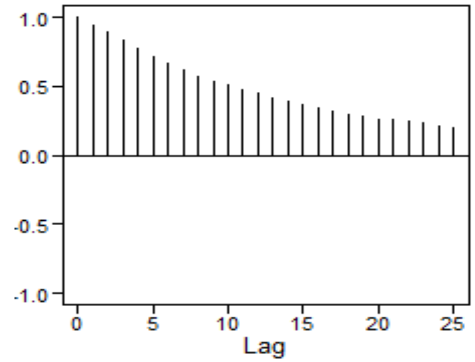
Example: DOW Jones Industrial Average

DOW Jones Average over Two Consecutive 100-Day Periods

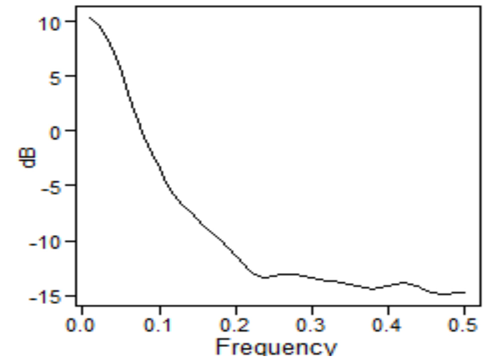
Dow Jones average



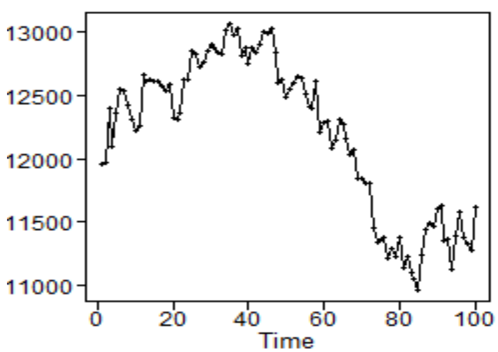
Autocorrelations



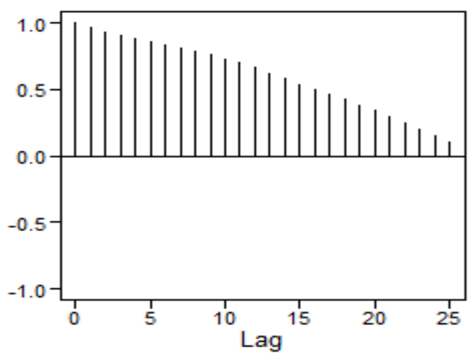
Spectral density



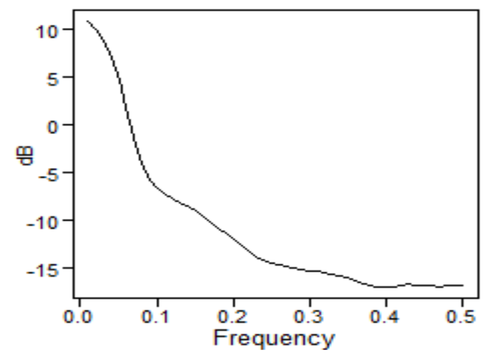
Dow Jones average



Autocorrelations



Spectral density



Note: Both data sets have characteristics similar to an AR(1) with $\varphi_1 \approx .95$.

DataScience@SMU

AR(1)

φ_1 Negative

AR(1) Models: φ_1 Negative

Summary

Realizations seem to be "oscillating," that is, if X_t is above the mean, then the strong tendency is for X_{t+1} to be below the mean and so on.

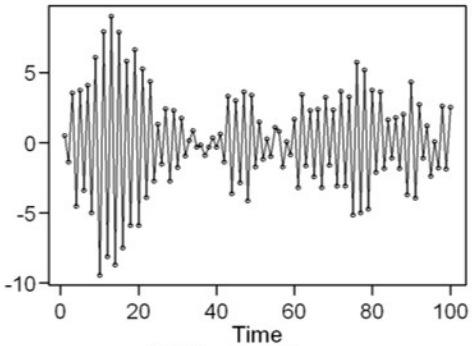
Autocorrelations are damped, oscillating exponentials. For example, for $\varphi_1 = -.95$, then $\rho_1 = -.95$, which is consistent with the behavior described above for realizations.

Spectral densities have peaks at $f = .5$ (i.e. a cycle length of 2). This is consistent with the up-and-down behavior in the realizations.

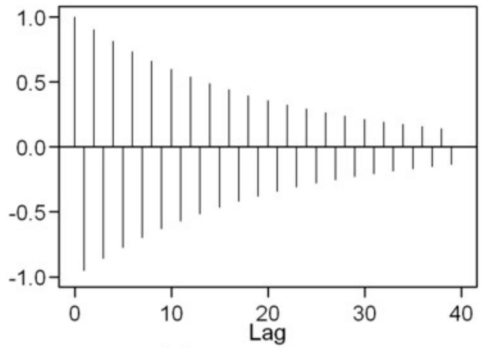
AR(1) Models: φ_1 Negative

$$\boxed{\varphi_1 = -.95, \text{ i.e. } X_t + .95X_{t-1} = a_t}$$

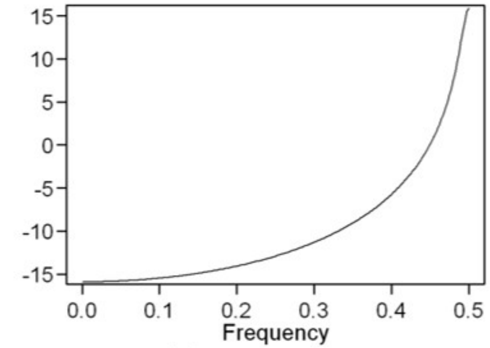
Data



Autocorrelations

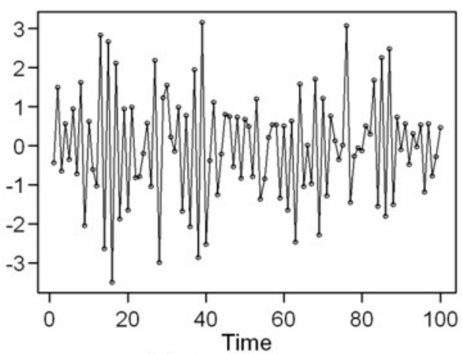


Spectral density

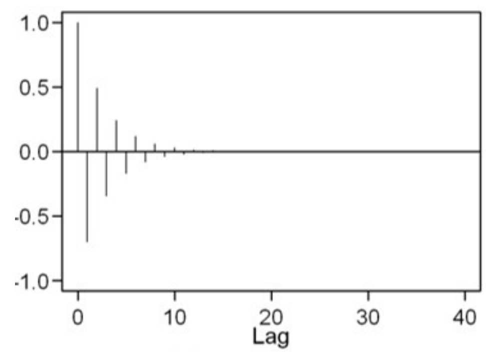


$$\boxed{\varphi_1 = -.7, \text{ i.e. } X_t + .7X_{t-1} = a_t}$$

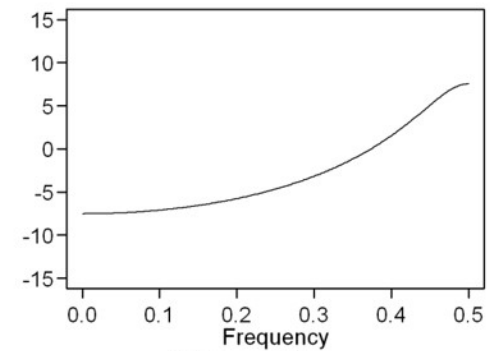
Data



Autocorrelations



Spectral density



DataScience@SMU

TSWGE Demo

`tswge` demo

The following command generates and plots a realization from an AR, MA, or ARMA model.

```
gen.arma.wge(n,phi,theta,vara,sn)
```

Notes:

- For AR(1) models, phi is a constant and theta retains its default value of 0.
- sn=0 (default) generates a new (randomly obtained) realization each time. Setting sn>0 allows you to generate the same realization each time you apply the command with the same sn.

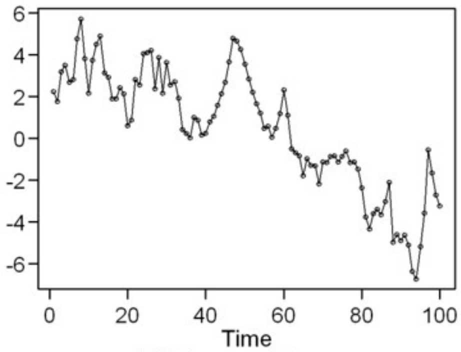
```
gen.arma.wge(n=100,phi=.95)
gen.arma.wge(n=100,phi=-.7)
gen.arma.wge(n=100,theta=.95,sn=5)
gen.arma.wge(n=100,theta=.95,sn=5)
```

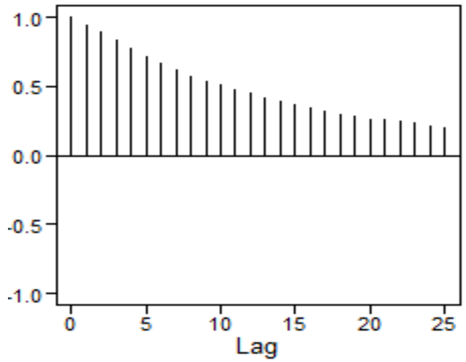
tswge demo

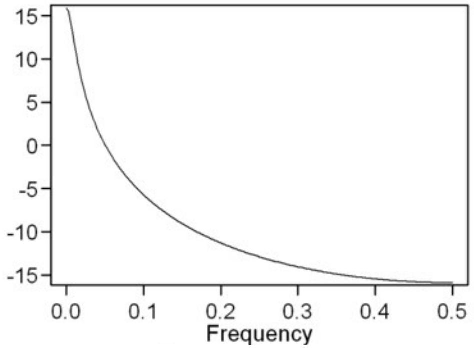
The following command plots a realization of length n (default $= 100$), the true autocorrelations, and the spectral density for an MA, AR, or ARMA model.

```
plotts.true.wge(n,phi,theta,lag.max)
```

Example: `plotts.true.wge(phi=.95)`







tswge Demo

```
# plot realization of length 100, true
# autocorrelations and spectral density
# for an AR(1)
plotts.true.wge(phi=.95)
# generate (and plot) AR(1) realization and
# place it in vector x
x=gen.arma.wge(n=200,phi=.95)
# plot realization in x along with sample
# autocorrelations, periodogram, and Parzen
# window-based spectral estimator
plotts.sample.wge(x)
```

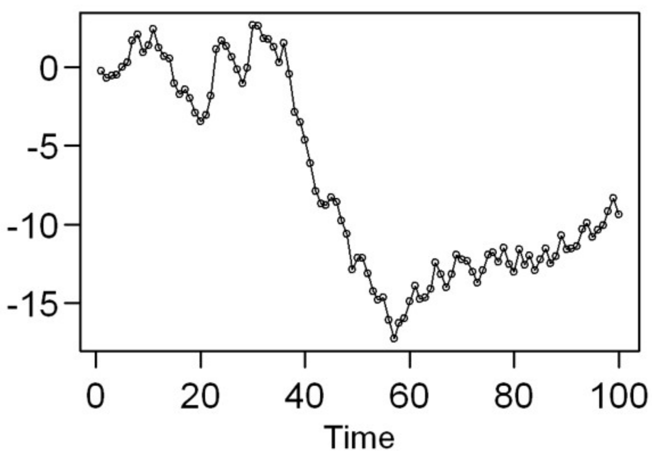
DataScience@SMU

Nonstationary “AR(1)-Type” Models

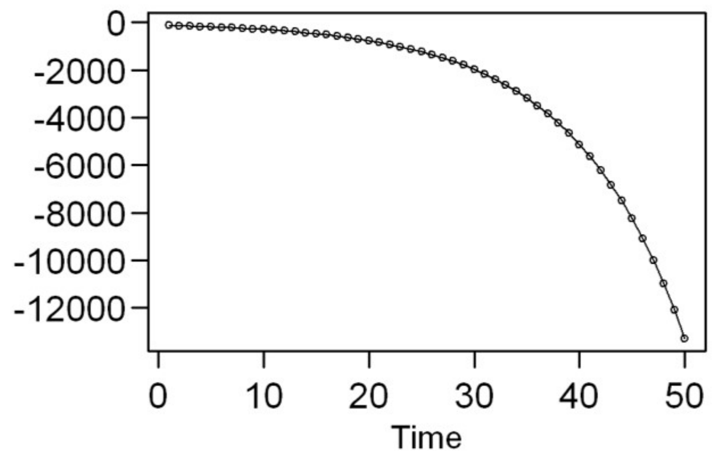
Nonstationary “AR(1)-Type” Models

AR(1) model: $X_t = (1 \quad {}_1) \quad + \quad {}_1 X_{t \quad 1} + a_t$

${}_1 = 1$



${}_1 = 1.1$



Nonstationary "AR(1)-Type" Models

- In both cases, it is not true that $|\varphi_1| < 1$, so both are realizations from *nonstationary* processes.
- When $|\varphi_1| = 1$, the realization looks similar to realizations that are encountered in practice.
  - In fact, somewhat similar to Dow Jones data
  - This is a special case of an *ARIMA* model to be covered later
- When $|\varphi_1| > 1$, the realization is explosive and is not typical of realizations seen in practice.

tswge

gen.arma.wge – is only for stationary models

gen.arima.wge – for ARIMA models like $\varphi_1 = 1$ case

tswge demo

```
gen.arma.wge(n=100,phi=c(.9999))#nearly nonstationary
gen.arma.wge(n=100,phi=c(1)) #error message
gen.arima.wge(n=100,d=1) #ARIMA case (similar to .9999)
gen.arma.wge(n=50,phi=c(1.1))  #error message
```

Functions in **tswge will not** generate realizations from the explosively non-stationary process $X_t = 1.1X_{t-1} + a_t$.

Using R code

```
n=50
x=rep(0,50)
a=rnorm(n)
x[1:50]=0
for(k in 2:n) {
    x[k]=1.1*x[k-1]+a[k]

}
plotts.wge(x)
```

DataScience@SMU

Summary of AR(1)

AR(1) Models: Summary

Realizations are wandering or oscillating depending on whether the root of the characteristic equation is positive or negative, respectively.

Autocorrelations are damped exponentials or damped oscillating exponentials depending on whether the root of the characteristic equation is positive or negative, respectively.

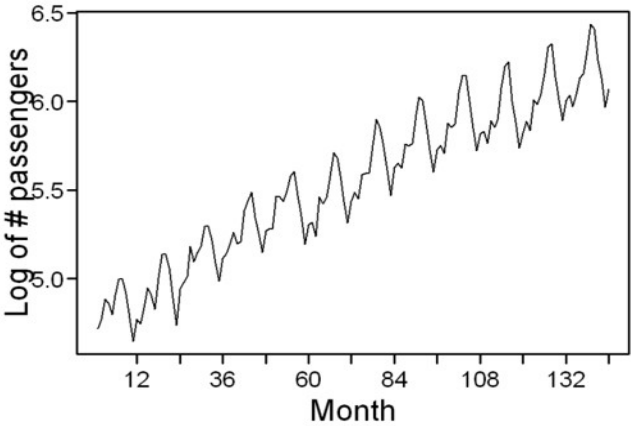
Spectral densities have a peak at $f = 0$ or $f = .5$ depending on whether the root of the characteristic equation is positive or negative, respectively.

DataScience@SMU

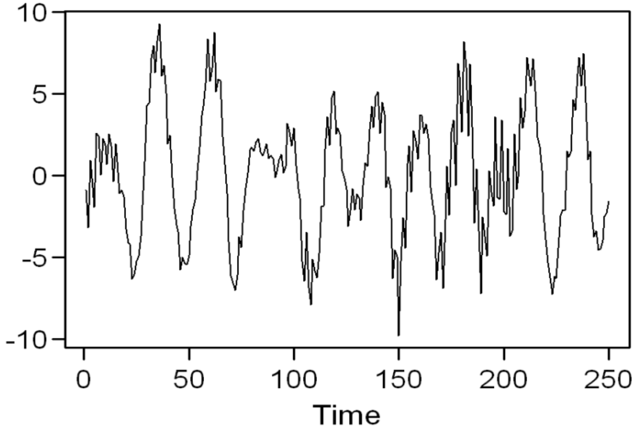
Beyond AR(1) and AR(2)

Some More Complicated Behaviors Cannot Be Explained by an AR(1) or AR(2) Model

Log airline data



Data with two frequencies



There are many behaviors that cannot be explained/modeled with AR(1) or AR(2) models.

Clearly, we need “higher-order” models to explain some behaviors.

- And in some cases, we need non-stationary-type models.
- In response, we will define the general AR(p) model.

DataScience@SMU