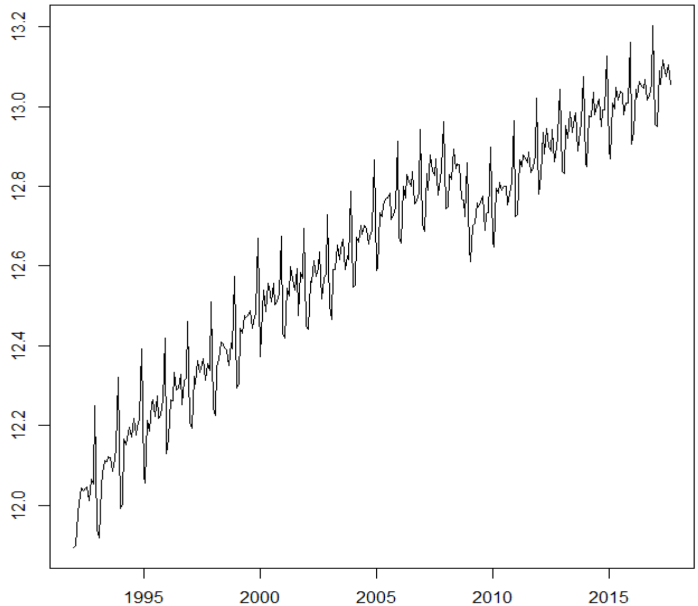
Exercise: Download the monthly not seasonally adjusted **Retail and Food Services Sales** from FRED, import this time series into R, and plot the log series.

- Create a working directory, say **C:\Projects\Retail**.

- Download the data as a csv file (**RSAFSNA.csv**) from FRED into the directory (**C:\Projects\Retail**).

- Launch R and enter the following commands.

```
setwd("C:/Projects/Retail")
H <- read.csv("RSAFSNA.csv",header=TRUE)
y <- ts(data=log(H[,2]),start=1992+0/12,frequency=12)
  # Jan 1992=1992+0/12, Feb 92=1992+1/12, …
  # frequency=number of observations per unit of time
par(mar=c(2,2,1,1)); plot(y)
```



The series exhibits a broken linear trend and a **seasonal pattern** with a maximum in November and a minimum in December or January.

In the simplest case, a trend of a not seasonally adjusted time series may be described as the sum of a linear trend and a periodic function h with period p, i.e.,

$$\mu_t = Ey_t = a + bt + h(t),$$

where

$$h(t) = h(t+p)$$

and $$h(1) + \ldots + h(p) = 0.$$ [1]

Exercise: Show that SP

$$h(1)+\ldots+h(p)=0 \Rightarrow h(t+1)+\ldots+h(t+p)=0 \ \forall t.$$

In the presence of a seasonal trend component $h(t)$ with period p, the symmetric filter

$$z_t = \begin{cases} \frac{1}{2k+1}\left(y_{t-k} + \ldots + y_{t+k}\right), & p = 2k+1 \\ \frac{1}{2k}\left(\frac{1}{2} y_{t-k} + y_{t-k+1} + \ldots + y_{t+k-1} + \frac{1}{2} y_{t+k}\right), & p = 2k \end{cases}$$

is an unbiased estimator for the linear trend component.

[1] Without this restriction the parameter a would be redundant.

(i) $p=2k+1$:

$$Ez_t = \frac{1}{2k+1}\sum_{j=-k}^{k}(a + b(t+j) + h(t+j))$$

$$= a + bt + \frac{1}{2k+1}(\underbrace{h(t-k) + \ldots + h(t+k)}_{=0})$$

$$= a + bt$$

(ii) $p=2k$:

$$Ez_t = \frac{1}{2k}\sum_{j=-k+1}^{k-1}(a + b(t+j) + h(t+j))$$

$$+ \frac{1}{2k}\frac{1}{2}(a + b(t-k) + h(t-k) + a + b(t+k) + h(t+k))$$

$$= \frac{2(k-1)+1}{2k}(a+bt) + \frac{1}{2k}(h(t-k+1) + \ldots + h(t+k-1))$$

$$+ \frac{1}{2k}\frac{2}{2}(a+bt) + \frac{1}{2k}\frac{1}{2}(h(t-k+\underbrace{2k}_{=p}) + h(t+k))$$

$$= a + bt + \frac{1}{2k}(\underbrace{h(t-k+1) + \ldots + h(t+k)}_{=0})$$

$$= a + bt$$ SO

Once the (locally) linear trend is estimated by the symmetric MA filter

$$z_t\ ,\ t=\mathrm{k}+1,,\ldots,n-\mathrm{k},$$

the **seasonal factors** $h(t)$, t=1,..,p can be estimated by

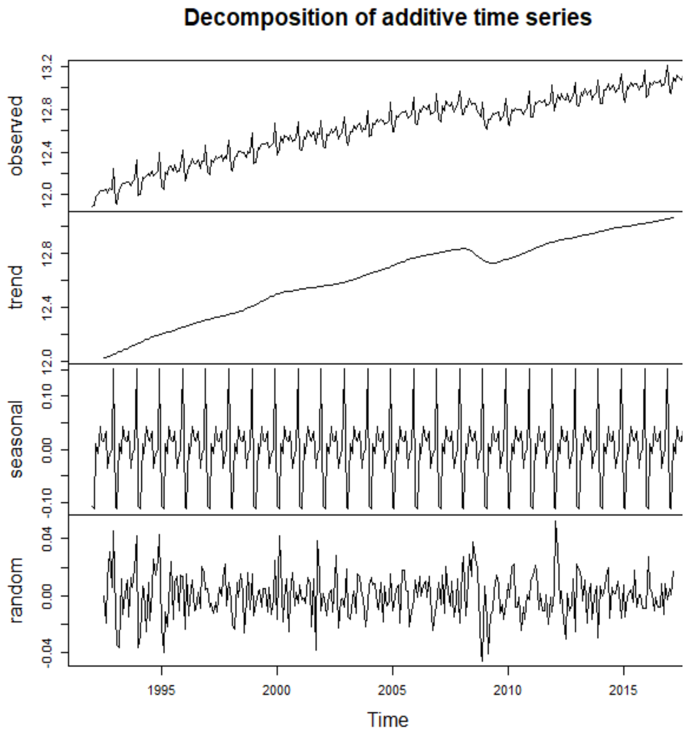
$$\hat{h}_t = h_t^* - mean(h_1^*,\ldots,h_p^*)\,,$$

where

$$h_t^* = mean(y_{t+k} - z_{t+k}, y_{t+k+p} - z_{t+k+p}, y_{t+k+2p} - z_{t+k+2p},\ldots).$$

Exercise: Use the R function **decompose** to decompose the observed monthly time series **y.log** into a trend component, a seasonal component, and a random component.

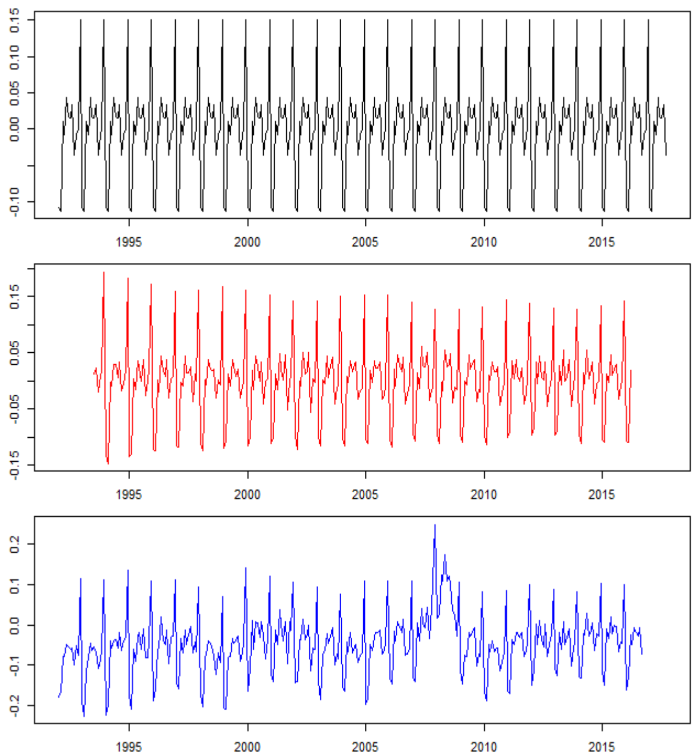
y.dec <- decompose(y); plot(y.dec)



Since seasonal patterns often change over time, simple moving average methods with constant seasonal factors are of limited use for the seasonal adjustment of long time series. A simple remedy is to calculate the seasonal factors locally, e.g., for segments of 3 or 5 years.

Exercise: Calculate the seasonal factors for the series **y** locally and compare the results with the estimates obtained by downloading also the seasonally adjusted series from FRED and subtracting the logarithms from **y**.

```
names(y.dec)
"x" "seasonal" "trend" "random" "figure" "type"
h <- y.dec$seasonal; y.res <- y-y.dec$trend; p <- 12
h3 <- (y.res+lag(y.res,k=p)+lag(y.res,k=-p))/3
par(mfrow=c(3,1),mar=c(2,2,1,1))
XL <- range(time(y)) # time range for plots
plot(h,xlim=XL); plot(h3,xlim=XL,col="red")
H.sa <- read.csv("RSAFS.csv",header=TRUE)
y.sa <- ts(data=log(H.sa[,2]),start=1992,frequency=12)
plot(y-y.sa,xlim=XL,col="blue")
```

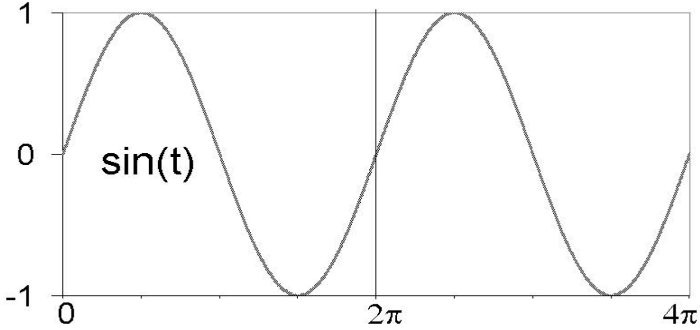


We may also use periodic functions for the description of a seasonal pattern. A flexible periodic function is the **sinusoid**

$$g(t)=R\sin(\omega t+\phi).$$

The **amplitude** R determines the height of the sinusoid, the **frequency** ω determines the number of cycles in an interval of length 2π, and the **phase** ϕ determines the position of the maximum.

For R=1, ω=1, ϕ=0, we obtain the ordinary sine function.



Exercise: Sketch $g(t)=R\sin(\omega t+\phi)$ for

(i) $R=\frac{1}{2}$, ω=2, ϕ=0,

(ii) R=2, $\omega=\frac{1}{2}$, ϕ=0,

(iii) R=1, ω=1, $\phi=\frac{\pi}{2}$. S1

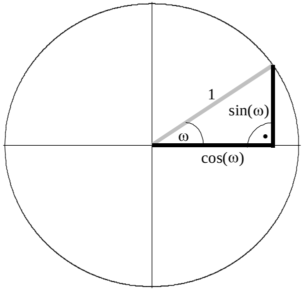
Basic properties of $\sin(\omega)$ and $\cos(\omega)$:

$\sin(\omega+2\pi k)=\sin(\omega)$, $\cos(\omega+2\pi k)=\cos(\omega)$, $k=0,\pm1,\pm2,\pm3,...$

$\sin^2(\omega)+\cos^2(\omega)=1$, $\sin(\frac{\pi}{4})=\cos(\frac{\pi}{4})=\frac{\sqrt{2}}{2}$

$\sin(\alpha\pm\beta)=\sin(\alpha)\cos(\beta)\pm\cos(\alpha)\sin(\beta)$

$\cos(\alpha\pm\beta)=\cos(\alpha)\cos(\beta)\mp\sin(\alpha)\sin(\beta)$

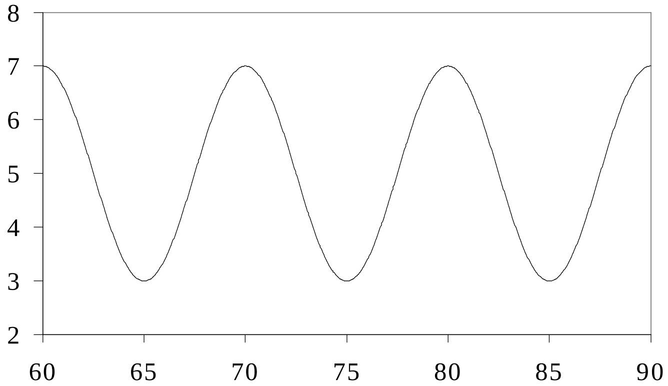


Exercise: $g(t)=R\sin(\omega t+\phi)$ S2

(i) Show that $g(t)$ is periodic with period $p=\frac{2\pi}{\omega}$.

(ii) Determine the value of ω such that $g(t)$ is periodic with period $p=12$.

Exercise: Find the parameters μ, R, and ϕ of the function $g(t)=\mu+R\sin(\omega t+\phi)$ displayed below. S3



Hidden periodicities model: SH

$$y_t=a+bt+R\sin(\omega t+\phi)+u_t,\quad Eu_t=0$$

We often have prior information about ω. For example, in the case of a not seasonally adjusted monthly series, we may expect that the period is 12 and the frequency is $\frac{2\pi}{12}$.

Assuming that ω is known and using

$$\sin(\omega t+\phi)=\sin(\omega t)\cos(\phi)+\cos(\omega t)\sin(\phi)$$

we obtain the linear model

$$y_t=a+bt+\underbrace{R\sin(\phi)}_{A}\cos(\omega t)+\underbrace{R\cos(\phi)}_{B}\sin(\omega t)+u_t.$$

The LS estimates of a, b, A, B are given by

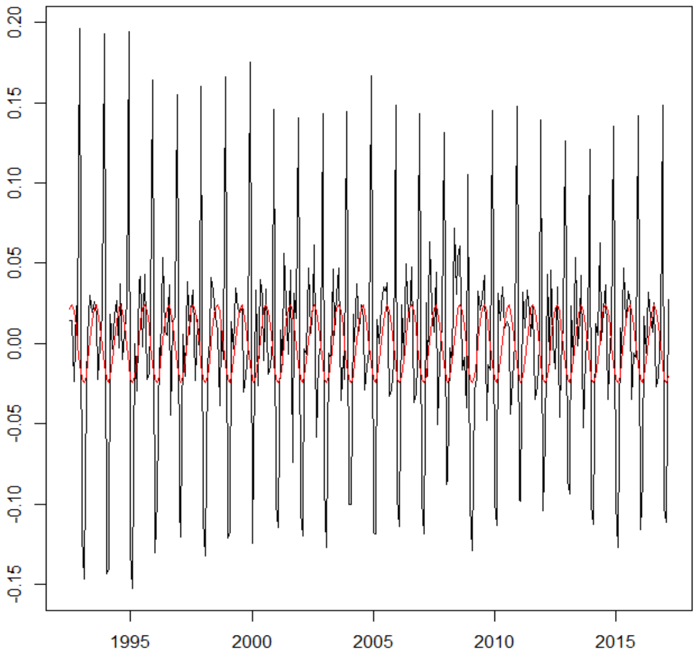
$$(\hat{a},\hat{b},\hat{A},\hat{B})^T=(X^TX)^{-1}X^Ty,$$

where $X=\begin{pmatrix}1 & 1 & \cos(\omega\cdot 1) & \sin(\omega\cdot 1)\\ \vdots & \vdots & \vdots & \vdots\\ 1 & n & \cos(\omega\cdot n) & \sin(\omega\cdot n)\end{pmatrix}$ and $y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$.

Exercise: Show that $R=\sqrt{A^2+B^2}$. SI

Exercise: Remove missing values from the trend residuals y.res and fit a sinusoid with period 12.

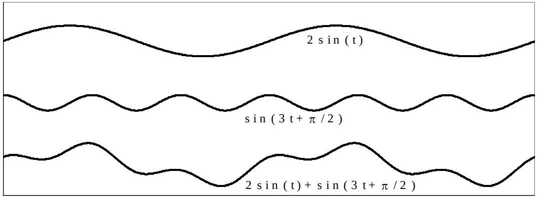
```
y.r <- na.omit(y.res,method="r") # remove NAs
N <- length(y.r); D <- 1:N; omega <- 2*pi/12
lm.sin <- lm(y.r~cos(omega*D)+sin(omega*D))
y.sin <- 0*y.r+lm.sin$fitted.values # result is time series
par(mfrow=c(1,1),mar=c(2,2,1,1))
plot(y.r,type="l"); lines(y.sin,col="red")
```



The fit is very poor. The hidden periodicities model is much too simplistic.

The seasonal trend components of real time series are much more complicated than simple sinusoids, hence we need periodic functions that are more flexible than sinusoids, e.g., sums of sinusoids.

The combination of different sinusoids gives non-sinusoidal oscillations.



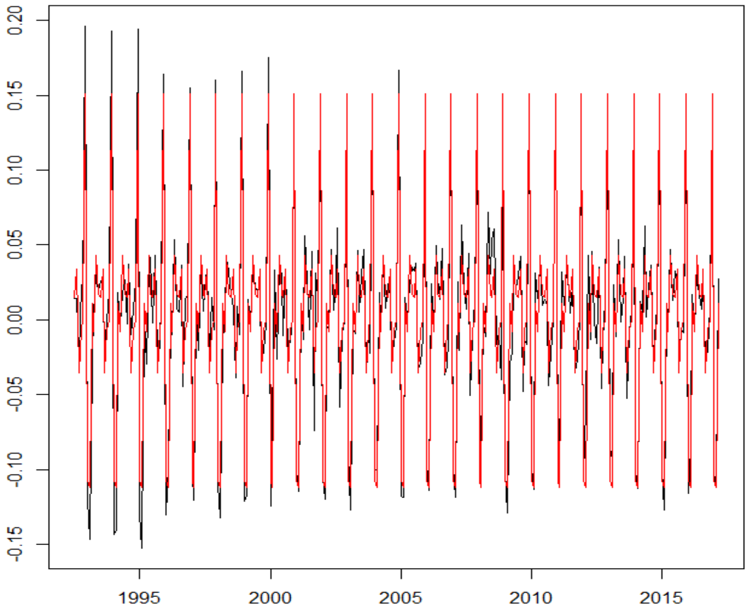
/

A sum of sinusoids with periods p_k=12/k, k=1,…,6, is periodic with period 12, hence it may be used to describe a nonsinusoidal seasonal pattern in a monthly time series:

$$y_t = a + bt + \sum_{k=1}^{6} \left(A_k \cos\left(\frac{2\pi}{p_k} t\right) + B_k \sin\left(\frac{2\pi}{p_k} t\right) \right) + u_t$$

Note: $\sin(\frac{2\pi}{p_6} t) = \sin(\pi t) = 0$ for all $t \in \mathbf{Z}$.

Exercise: Fit a combination of six sinusoids to the trend residuals **y.r**.



The fit is now much better.