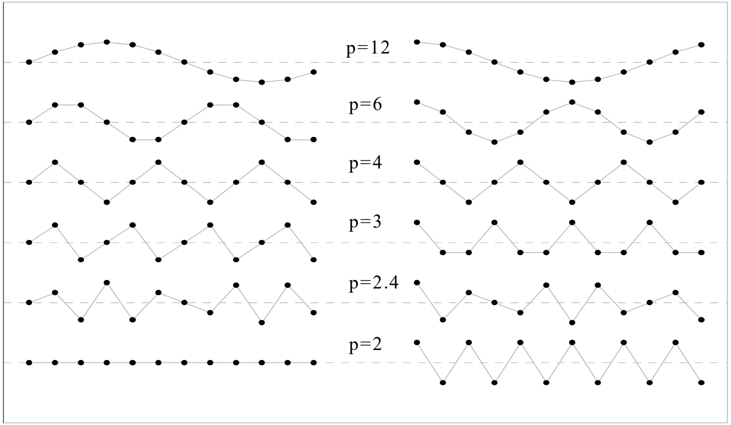
Trigonometric functions in discrete time:

$\sin(\frac{2\pi}{p}t)$ $\cos(\frac{2\pi}{p}t)$



Clearly,

$$g(t)=\cos(\frac{2\pi}{2}t)=\cos(\pi t)=(-1)^t$$

is the most rapid oscillation we can observe.

The frequency π (one half a cycle per sampling interval) is known as **Nyquist frequency**.

If one observation is sampled per unit of time, a periodic function with period n completes one full cycle in the observation period, a periodic function with period $n/2$ completes two full cycles in the observation period, etc. In general, a periodic function with period n/k completes k full cycles in the observation period. The frequency $\omega_k=2\pi k/n$ that is associated with the period n/k is called the k-th **Fourier frequency.**

We consider only those Fourier frequencies ω_k that are less than or equal to the Nyquist frequency π, i.e., $k\leq m=[n/2]$.

If the improper Fourier frequency $2\pi 0/n=0$ is also included, there are always n non-vanishing sines and cosines.

If n is even and $m=n/2$, we have

$$1, \cos(\frac{2\pi\cdot 1}{n}t), \ldots, \cos(\frac{2\pi\cdot m}{n}t), \sin(\frac{2\pi\cdot 1}{n}t), \ldots, \sin(\frac{2\pi\cdot(m-1)}{n}t).$$

If n is odd and $m=(n-1)/2$, we have

$$1, \cos(\frac{2\pi\cdot 1}{n}t), \ldots, \cos(\frac{2\pi\cdot m}{n}t), \sin(\frac{2\pi\cdot 1}{n}t), \ldots, \sin(\frac{2\pi\cdot m}{n}t).$$

Note that $\cos(\frac{2\pi\cdot 0}{n}t)$ equals 1 and $\sin(\frac{2\pi\cdot 0}{n}t)$ vanishes.

If n is even, then $\sin(\frac{2\pi\cdot m}{n}t)=\sin(\pi t)$ vanishes too.

Exericse: Show that $\hat{b} = \sum_{t=1}^{n} x_t y_t \Big/ \sum_{t=1}^{n} x_t^2$ minimizes the sum of squared errors $\sum_{t=1}^{n} (y_t - b x_t)^2$. CB

Using only Fourier frequencies $\omega_k = 2\pi k/n$, $k \in T \subseteq \{0,\dots,m\}$, in the trigonometric regression model

$$y_t = \sum_{k\in T} (A_k \cos(\omega_k t) + B_k \sin(\omega_k t)) + u_t$$

has the advantage that the regressors are orthogonal[1], i.e.,

$$C_k^{\mathrm{T}} S_j = 0, \text{ if } 0 \le j,k \le m,$$

$$S_k^{\mathrm{T}} S_j = C_k^{\mathrm{T}} C_j = 0, \text{ if } 0 \le j \ne k \le m,$$

where

$$C_k = \begin{pmatrix} \cos(\frac{2\pi k}{n} \cdot 1) \\ \vdots \\ \cos(\frac{2\pi k}{n} \cdot n) \end{pmatrix}, \; S_j = \begin{pmatrix} \sin(\frac{2\pi j}{n} \cdot 1) \\ \vdots \\ \sin(\frac{2\pi j}{n} \cdot n) \end{pmatrix}.$$

[1] See Appendix B.

In the case of orthogonal regressors, the LS estimates

$$\left(\hat{A}_0, \hat{A}_1, \hat{B}_1, \dots\right)^{\mathrm{T}}$$

$$= \left((C_0, C_1, S_1, \dots)^{\mathrm{T}} (C_0, C_1, S_1, \dots)\right)^{-1} (C_0, C_1, S_1, \dots)^{\mathrm{T}} y$$

$$= \underbrace{\begin{pmatrix} C_0^{\mathrm{T}} C_0 & C_0^{\mathrm{T}} C_1 & C_0^{\mathrm{T}} S_1 & \cdots \\ C_1^{\mathrm{T}} C_0 & C_1^{\mathrm{T}} C_1 & C_1^{\mathrm{T}} S_0 & \cdots \\ S_1^{\mathrm{T}} C_0 & S_1^{\mathrm{T}} C_1 & S_1^{\mathrm{T}} S_1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}^{-1}}_{=diag(n,n/2,n/2,\dots)} \begin{pmatrix} C_0^{\mathrm{T}} y \\ C_1^{\mathrm{T}} y \\ S_1^{\mathrm{T}} y \\ \vdots \end{pmatrix}$$

obtained from the full model are identical to the estimates

$$\hat{A}_0 = \sum_{t=1}^{n} \cos(\omega_0 t) y_t \Big/ \sum_{t=1}^{n} \cos^2(\omega_0 t) = \tfrac{1}{n} \sum_{t=1}^{n} \cos(\omega_0 t) y_t = \bar{y},$$

$$\hat{A}_1 = \sum_{t=1}^{n} \cos(\omega_1 t) y_t \Big/ \sum_{t=1}^{n} \cos^2(\omega_1 t) = \tfrac{2}{n} \sum_{t=1}^{n} \cos(\omega_1 t) y_t,$$

$$\hat{B}_1 = \sum_{t=1}^{n} \sin(\omega_1 t) y_t \Big/ \sum_{t=1}^{n} \sin^2(\omega_1 t) = \tfrac{2}{n} \sum_{t=1}^{n} \sin(\omega_1 t) y_t, \dots$$

obtained from simple models with only one regressor. CR

Exercise: Show that $\sum_{t=1}^{n}\cos^2(\frac{2\pi k}{n}t)=n$, if $k=\frac{n}{2}$. CN

Exercise: If n is even, then $m=n/2$ and $\omega_m=\pi$. But since $\sin(\pi t)=0$, only $\cos(\pi t)$ can be used as a regressor. Show that the LS estimate of the parameter A_m is given by

$$\hat{A}_m=\frac{1}{n}\sum_{t=1}^{n}y_t(-1)^t.$$

CM

The **periodogram** $I(\omega)$ of a time series $y_1,\ldots,y_n$ is for a Fourier frequency ω_k with $k\in\{1,2,\ldots,[\frac{n-1}{2}]\}$ defined by

$$I(\omega_k)=\frac{n}{8\pi}(\underbrace{\hat{A}_k^2+\hat{B}_k^2}_{\hat{R}_k^2})$$

$$=\frac{n}{8\pi}((\frac{2}{n}\sum_{t=1}^{n}y_t\cos(\omega_k t))^2+(\frac{2}{n}\sum_{t=1}^{n}y_t\sin(\omega_k t))^2)$$

$$=\frac{1}{2\pi n}((\sum_{t=1}^{n}y_t\cos(\omega_k t))^2+(\sum_{t=1}^{n}y_t\sin(\omega_k t))^2).$$

Apart from an unimportant scaling factor, $I(\omega_k)$ is just the squared estimate of the amplitude of a sinusoid with frequency ω_k, hence its size indicates how important that particular frequency is.

Defining the periodogram for an arbitrary frequency ω by

$$I(\omega)=\frac{1}{2\pi n}((\sum_{t=1}^{n}y_t\cos(\omega t))^2+(\sum_{t=1}^{n}y_t\sin(\omega t))^2),$$

it can be written in complex form as

$$I(\omega)=\frac{1}{2\pi n}\left|\sum_{t=1}^{n}y_t\cos(\omega t)+i\sum_{t=1}^{n}y_t\sin(\omega t)\right|^2$$

$$=\frac{1}{2\pi n}\left|\sum_{t=1}^{n}y_t(\cos(\omega t)+i\sin(\omega t))\right|^2$$

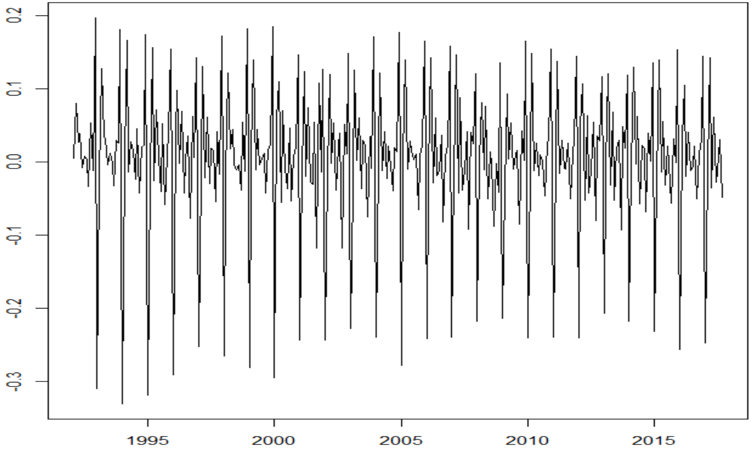
$$=\frac{1}{2\pi n}\left|\sum_{t=1}^{n}y_te^{i\omega t}\right|^2.$$

CP

Exercise: Show that it is sufficient to consider $I(\omega)$ on the interval $0 \le \omega \le \pi$ by proving that $I(\omega)$ is periodic with period 2π and symmetric about the y-axis. CS

Exercise: Revisit the not seasonally adjusted **Retail and Food Services Sales**. Plot the differenced log series.

```
d <- y-lag(y,k=-1); d <- na.omit(d,method="r")
par(mar=c(2,2,1,1)); plot(d,type="l")
```



The sequence

$$\sum_{t=1}^{n} y_t \ e^{-i\omega_k t}, \ k=0,\ldots,n-1,$$

is called the **discrete Fourier transform (DFT)** of the sequence $y_1,\ldots,y_n$.

The R function fft uses a fast algorithm (the **fast Fourier transform**) to calculate the slightly different version

$$\sum_{t=1}^{n} y_t \ e^{-i\omega_k (t-1)}, \ k=0,\ldots,n-1.$$

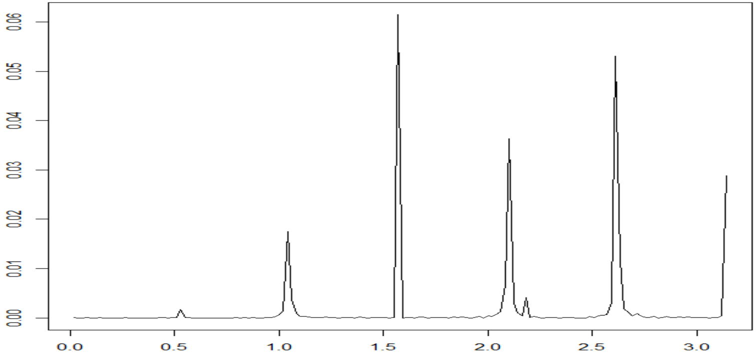
However, we have

$$\left| \sum_{t=1}^{n} y_t e^{-i\omega_k (t-1)} \right|^2 = \left| e^{i\omega_k} \sum_{t=1}^{n} y_t e^{-i\omega_k t} \right|^2$$

$$= \underbrace{\left| e^{i\omega_k} \right|^2}_{=1} \left| \sum_{t=1}^{n} y_t e^{-i\omega_k t} \right|^2 .$$

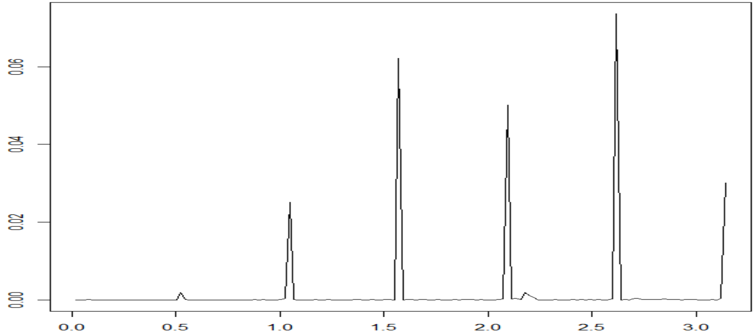
Exercise: Calculate the periodogram of d at the Fourier frequencies ω_k, k=1,…,[n/2] and plot it.

```
n <- length(d); m <- floor(n/2)  # number of frequencies
f <- (2*pi/n)*(1:m)  # vector of frequencies
ft <- fft(d)[2:(m+1)]  # excl. k=0,m+1,m+2,...,n-1
pg <- (1/(2*pi*n))*(Mod(ft))^2 # Mod = modulus
plot(f,pg,type="l")
```



Exercise: Redo the last exercise, but this time omit some observations to guarantee that the seasonal frequencies $2\pi j/12$, j=1,…,6, are Fourier frequencies.

```
r12 <- n%%12  # n modulo 12
          #  the remainder we get when we divide n by 12
n12 <- n - r12
d12 <- d[(r12+1):n] # omit the first r12 observations
          #  length of new series d12 is divisible by 12
m12 <- floor(n12/2); f12 <- (2*pi/n12)*(1:m12)
ft12 <- fft(d12); ft12 <- ft12[2:(m12+1)]
pg12 <- (1/(2*pi*n12))*(Mod(ft12))^2
plot(f12,pg12,type="l")
```



The peaks at the seasonal frequencies are more pronounced than the slightly displaced peaks in the last exercise.