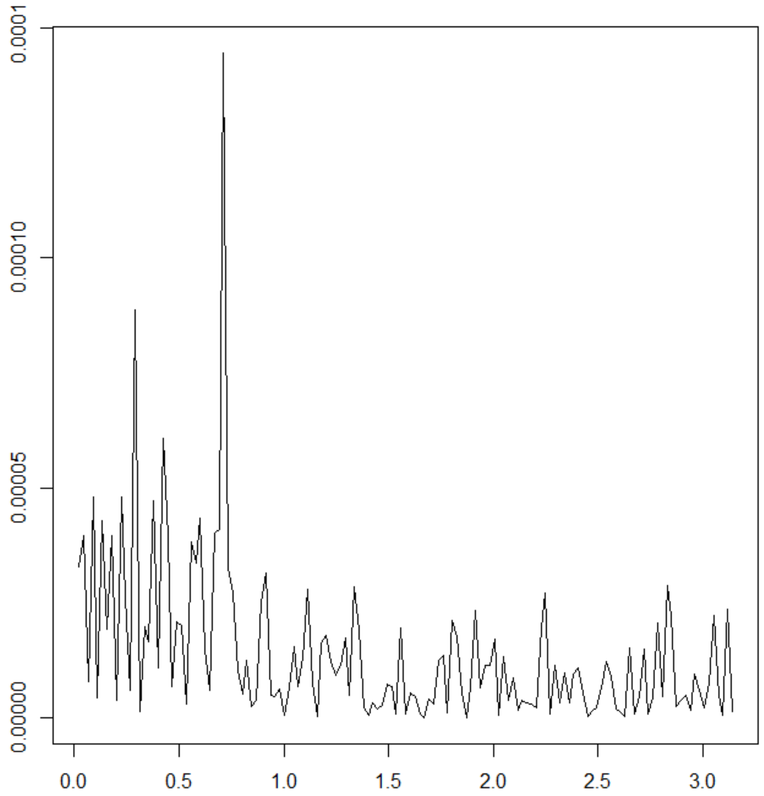
Exercise: Download the quarterly **Real Gross Domestic Product, 3 Decimal** from FRED, import it into R, and plot the periodogram of the differenced log series.

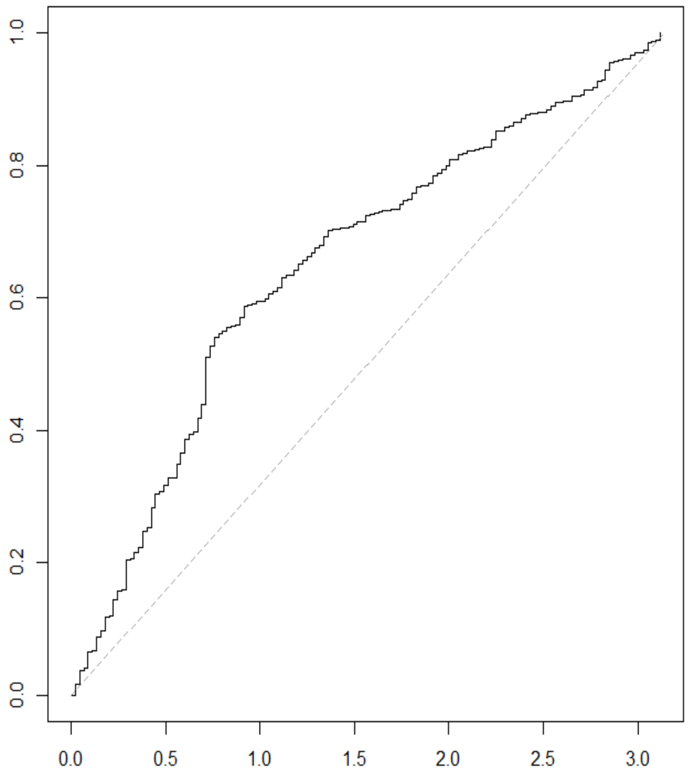
```
D <- read.csv("GDPC1.csv"); N <- nrow(D)
d <- as.Date(D[,1]); y <- log(D[,2])
r <- y[2:N]-y[1:(N-1)]; n <- N-1  # log returns

m <- floor(n/2); f <- (2*pi/n)*(1:m)
ft <- fft(r)[2:(m+1)]  # excl. k=0,m+1,m+2,...,n-1
pg <- (1/(2*pi*n))*(Mod(ft))^2 # Mod = modulus
par(mar=c(2,2,1,1)); plot(f,pg,type="l")
```

Since there are no strict periodicities in the seasonally adjusted GDP, the periodogram does not exhibit any isolated, sharp peaks. However, the cluster of large periodogram ordinates in the low frequency range is an indication of the presence of business cycles stretching over several quarters (or even years).



In case of white noise, all periodogram ordinates should be approximately of the same size and the standardized cumulative periodogram should increase linearly from 0 to 1. Bartlett's test for white noise (which is included in the R package **hwwntest**) uses the maximum deviation of the standardized cumulative periodogram and this straight line as test statistic.

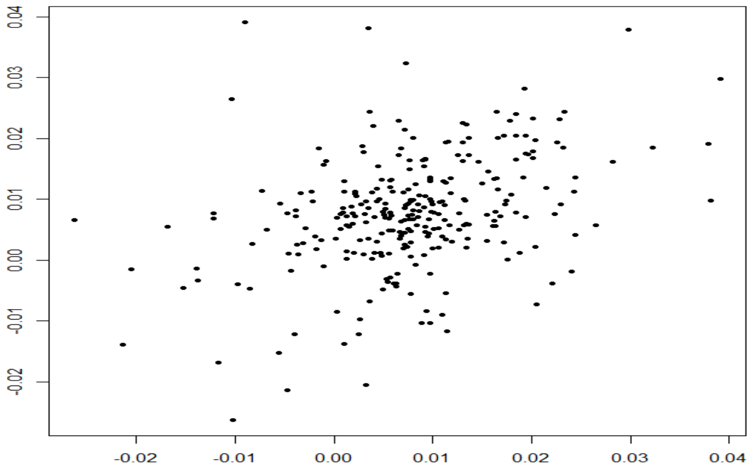
```
par(mfrow=c(1,1),mar=c(2,2,1,1))
M <- 1:floor((n-1)/2)                    # pi is excluded
h <- cumsum(pg[M])/sum(pg[M])
      # standardized cumulative periodogram
plot(c(0,f[M]),c(0,h),type="s")       # 0 is included
                                         # "s": stair steps
lines(c(0,pi),c(0,1),lty=2,col="darkgray")
```



The presence of business cycles of several quarters duration implies a tendency for good quarters to be followed by good quarters and bad quarters to be followed by bad quarters.

Exercise: Compare the growth rates of successive quarters.

```
plot(r[1:(n-1)],r[2:n],pch=20)
```



In this **scatter diagram**, the GDP growth rate in quarter t is plotted against the GDP growth rate in quarter t-1.

The graph suggests that successive growth rates are positively correlated.

Given a time series $x_1,\ldots,x_n$, we could try to measure the strength of the relationship between successive observations by calculating the sample covariance from the pairs $(x_1,x_2),\ldots,(x_{n-1},x_n)$.

The **sample autocovariance** (at lag 1)

$$\hat{\gamma}(1) = \frac{1}{n}\sum_{t=1}^{n-1}(x_t - \bar{x})(x_{t+1} - \bar{x}),$$

differs from a conventional sample covariance in two ways.

Firstly, the sum is divided by n although it contains less than n terms.

Secondly, the two sample means occurring in each term are not calculated from the samples $x_1,\ldots,x_{n-1}$ and $x_2,\ldots,x_n$, respectively, but from the whole sample $x_1,\ldots,x_n$.

Sample moments calculated from a given time series $x_1,\ldots,x_n$ are only meaningful, if the data generating process satisfies certain conditions. For example, the sample mean $\bar{x}$ does not have a meaningful interpretation, unless all x_t have the same mean.

A stochastic process x is said to be (**weakly**) **stationary** if all x_t have the same mean and the same variance, and

$$\mathrm{Cov}(x_s,x_{s-k})=\mathrm{Cov}(x_t,x_{t-k}) \;\forall s,t,k\in\mathbf{Z}.$$

If x is white noise, it clearly satisfies the first two conditions of stationarity. In addition, it follows from

$$\mathrm{Cov}(x_s,x_t)=0 \;\forall s\neq t$$

that

$$\mathrm{Cov}(x_s,x_{s-k})=\begin{Bmatrix} Var(x_t) & if\ k=0 \\ 0 & if\ k\neq 0\end{Bmatrix}=\mathrm{Cov}(x_t,x_{t-k}) \;\forall s,t,k\in\mathbf{Z},$$

hence white noise is stationary.

The **autocovariance function** and the a**utocorrelation function** of a weakly stationary process x are defined by

$$\gamma(k)=\mathrm{Cov}(x_t,x_{t-k}),\ k=0,\pm1,\pm2,\ldots,$$ [1]

and $$\rho(k)=\gamma(k)/\gamma(0),\ k=0,\pm1,\pm2,\ldots,$$

respectively.

For $|k|<n$ the autocovariance $\gamma(k)$ is estimated by

$$\hat{\gamma}(k)=\tfrac{1}{n}\sum_{t=1}^{n-|k|}(x_t-\bar{x})(x_{t+|k|}-\bar{x})$$

(**sample autocovariance at lag *k***) and the autocorrelation $\rho(k)$ by

$$\hat{\rho}(k)=\hat{\gamma}(k)/\hat{\gamma}(0)$$

(**sample autocorrelation at lag *k***).

[1] Because of the stationarity of x, this definition does not depend on the choice of t.

It is shown in Appendix B that the periodogram of a sample $x_1,\ldots,x_n$ can for a Fourier frequency

$$0<\omega_k=\frac{2\pi k}{n}<\pi$$

be written in terms of its sample autocovariances:

$$I(\omega_k)=\frac{1}{2\pi}\left(\hat{\gamma}(0)+2\sum_{j=1}^{n-1}\hat{\gamma}(j)\cos(\omega_k j)\right)$$

The periodogram may therefore be regarded as a sample analogue of the function

$$f(\omega)=\frac{1}{2\pi}\left(\gamma(0)+2\sum_{j=1}^{\infty}\gamma(j)\cos(\omega j)\right)$$

$$=\frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma(j)\cos(\omega j)+i\frac{1}{2\pi}\underbrace{\sum_{j=-\infty}^{\infty}\gamma(j)\sin(\omega j)}_{=0}$$

$$=\frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma(j)e^{-i\omega j}$$

which is called the **spectral density** of the process x.

It follows from $\cos(-\omega)=\cos(\omega)$ that f is an even function, i.e., $f(-\omega)=f(\omega)$. Moreover f is periodic with period 2π. Hence, we need to consider $f(\omega)$ only on the interval $[0,\pi]$.

The spectral density of a zero-mean white noise is a constant function, i.e.,

$$f(\omega)=\frac{1}{2\pi}\gamma(0).$$

To test the null hypothesis that a given time series is a realization of white noise we may either check whether its periodogram is roughly constant (**frequency domain tests**) or, equivalently, whether its sample autocovariances are close to zero (**time domain tests**).

Suppose that $z_1,\ldots,z_n$ are independent random variables with mean μ=0 and variance σ^2>0.

Then the mean and the variance of the statistic

$$\widehat{\gamma}(1)=\frac{1}{n}\sum_{t=1}^{n-1}(z_t-\underbrace{\mu}_{=0})(z_{t+1}-\underbrace{\mu}_{=0})\text{ [2]}$$

$$=\frac{1}{n}\sum_{t=1}^{n-1}z_t z_{t+1}$$

are given by

$$E\widehat{\gamma}(1)=\frac{1}{n}\sum_{t=1}^{n-1}E(z_t z_{t+1})$$

$$=\frac{1}{n}\sum_{t=1}^{n-1}\underbrace{E(z_t)}_{=0}\underbrace{E(z_{t+1})}_{=0}$$

$$=0,$$

[2] In contrast to the sample autocovariance, which uses the sample mean, this statistic uses μ.

$$\mathrm{Var}(\widehat{\gamma}(1))$$

$$=E(\widehat{\gamma}(1))^2-(\underbrace{E\widehat{\gamma}(1)}_{=0})^2$$

$$=E\left(\frac{1}{n}\sum_{t=1}^{n-1}z_t z_{t+1}\right)^2$$

$$=\frac{1}{n^2}E\left(\sum_{t=1}^{n-1}z_t z_{t+1}\sum_{t=1}^{n-1}z_t z_{t+1}\right)$$

$$=\frac{1}{n^2}E\,(z_1z_2\,(z_1z_2+z_2z_3+z_3z_4+z_4z_5+\ldots)+$$
$$z_2z_3\,(z_1z_2+z_2z_3+z_3z_4+z_4z_5+\ldots)+\ldots\,\}$$

$$=\frac{1}{n^2}E\,\{(z_1^2z_2^2+z_1z_2^2z_3+z_1z_2z_3z_4+z_1z_2z_4z_5+\ldots)+$$
$$(z_1z_2^2z_3+z_2^2z_3^2+z_2z_3^2z_4+z_2z_3z_4z_5+\ldots)+\ldots\,\}$$

$$=\frac{1}{n^2}\{(\underbrace{Ez_1^2}_{=\sigma^2}\underbrace{Ez_2^2}_{=\sigma^2}+\underbrace{Ez_1}_{=0}Ez_2^2z_3+\underbrace{Ez_1}_{=0}Ez_2z_3z_4+\ldots)+\ldots\,\}$$

$$=\frac{n\text{-}1}{n^2}\sigma^4$$

$$\sim\frac{\sigma^4}{\mathrm{n}}.$$

RM

Exercise: Suppose that $x_1,\dots,x_n$ are independent random variables with mean μ and variance $\sigma^2>0$. Show that

$$\hat{\rho}(1) = \frac{1}{n\sigma^2}\sum_{t=1}^{n-1}(x_t-\mu)(x_{t+1}-\mu)$$

has mean zero and variance $\frac{n-1}{n^2} \sim \frac{1}{n}$. RS

Regarding the sample autocorrelation

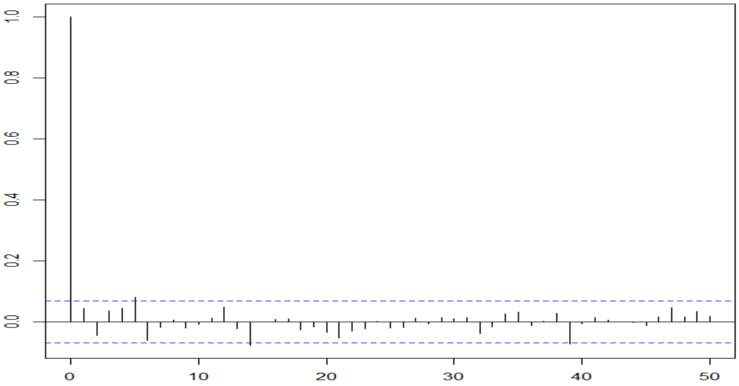
$$\hat{\rho}(1) = \hat{\gamma}(1)/\hat{\gamma}(0) = \frac{1}{n\hat{\gamma}(0)}\sum_{t=1}^{n-1}(x_t-\bar{x})(x_{t+1}-\bar{x})$$

as an approximation of $\hat{\rho}(1)$, we may expect that under the null hypothesis of independence, the mean and the variance of $\sqrt{n}\hat{\rho}(1)$ approximately are given by 0 and 1, respectively.

For large values of n we may also expect that the distribution of $\sqrt{n}\hat{\rho}(1)$ is approximately normal, hence $\hat{\rho}(1)$ should therefore fall between the bounds $\pm 1.96/\sqrt{n}$ approximately with probability 95%.

Exercise: Plot the sample autocorrelation function of the returns on the monthly S&P 500 Index.

```
par(mfrow=c(1,1),mar=c(2,2,1,1)); acf(r,lag.max=50)
```

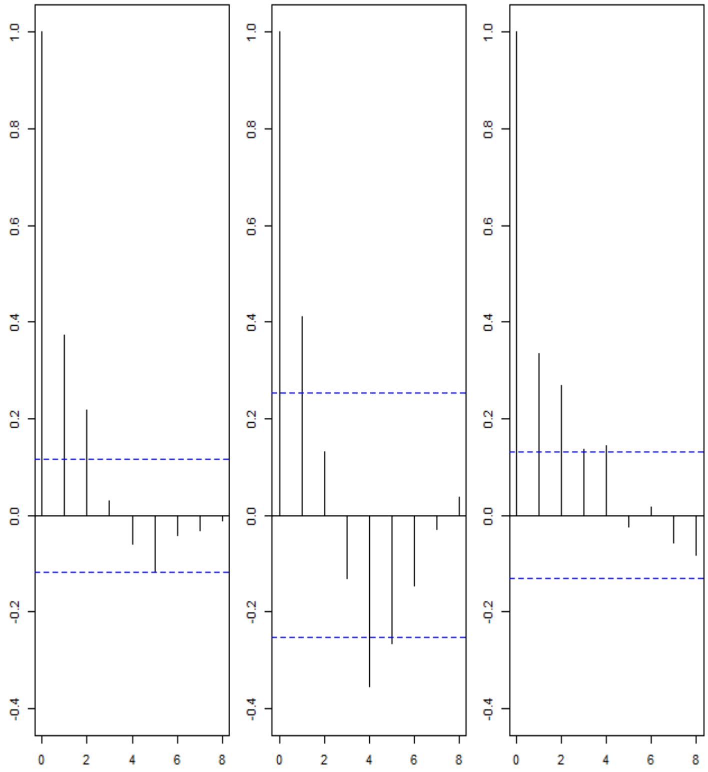


The fact that individual sample autocorrelations fall outside the approximate 95% bounds must be interpreted with caution. The bounds can only be used to assess the significance of one specified sample autocorrelation. Performing significance tests simultaneously for several sample autocorrelations inevitably increases the risk of rejecting a true null hypothesis.

Exercise: Examine the sample autocorrelations of the quarterly GDP.

Nonsignificant sample autocorrelations may be due to nonstationarity. For example, the sample autocorrelation at lag 4 of the quarterly GDP growth rates x_t is slightly negative over the whole postwar period. But a subperiod analysis shows that there is a clear negative relationship between x_t and x_{t-4} in the first 15 years, whereas in the remaining years there is even a positive relationship between x_t and x_{t-4}.

```
par(mfrow=c(1,3),mar=c(2,2,1,1))
acf(r[1:n],lag.max=8,ylim=c(-0.4,1))
acf(r[1:60],lag.max=8,ylim=c(-0.4,1))
acf(r[61:n],lag.max=8,ylim=c(-0.4,1))
```



A promising way to reveal any nonstationarities is to produce a cumulative plot.

For an examination of the relationship between x_t and $x_{t\text{-}1}$, we plot

$$x_2\, x_{2\text{-}1},\ x_2\, x_{2\text{-}1}+x_3\, x_{3\text{-}1},\ x_2\, x_{2\text{-}1}+x_3\, x_{3\text{-}1}+x_4\, x_{4\text{-}1},\ \ldots$$

against time.

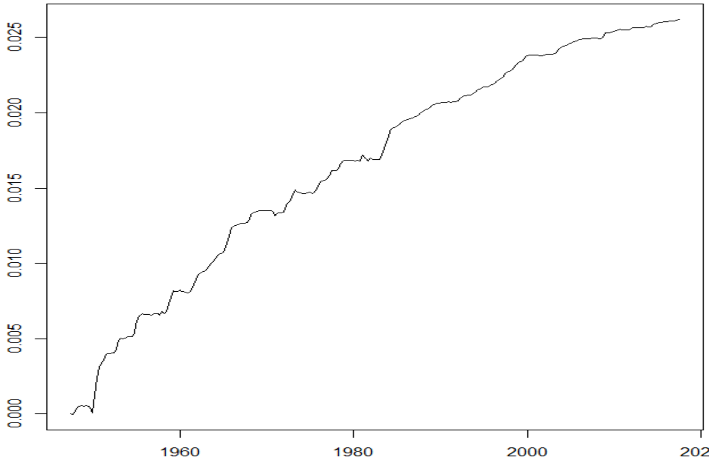
The last value

$$x_2\, x_{2\text{-}1}+x_3\, x_{3\text{-}1}+\ldots+x_n\, x_{n\text{-}1}$$

divided by n is an estimate of the non-central second moment $E\, x_t\, x_{t\text{-}1}$.

In case of stationarity, there should be a roughly linear increase (or decrease) from the first value to the last.

```
par(mfrow=c(1,1),mar=c(2,2,1,1))
plot(d[(1+2):N],cumsum(r[2:n]*r[1:(n-1)]),type="l")
```



There is a break in the noncentral second moment which may be due to a break in the mean or a break in the autocovariance.

A break in the autocovariance may be due to a break in the autocorrelation or a break in the variance[3].

[3] There was a reduction in macroeconomic volatility starting in the 1980s (Great Moderation).

Since periodogram ordinates are nonnegative, it makes sense to use sums of periodogram ordinates as test statistics. In contrast, sample autocorrelations need not have the same sign. We should therefore rather use sums of squared sample autocorrelations.

In case of independence, the sample autocorrelations are approximately normal with mean zero and variance $1/n$.

Hence, the statistic

$$Q = n \sum_{k=1}^{h} \hat{\rho}_k^2$$

will be approximately chi-squared distributed with h degrees of freedom, if the sample autocorrelations are approximately independent.

For large values of n, this statistic is almost identical to the widely used **Ljung-Box statistic**

$$Q_{LB} = n(n+2) \sum_{k=1}^{h} \frac{\hat{\rho}_k^2}{n-k}.$$

<u>Exercise:</u> Suppose that $z_1,\ldots,z_n$ are independent random variables with mean zero and variance σ^2. Show that

$$\mathrm{Cov}\left(\frac{1}{n}\sum_{t=1}^{n-1} z_t z_{t+1}, \frac{1}{n}\sum_{t=1}^{n-2} z_t z_{t+2}\right) = 0$$ RI