2. Reduction to stationarity

Often, time series data are not realizations of a stationary process. A first goal is then the reduction to a stationary process.

Consider the model

$$X_t = \underbrace{T_t}_{\text{trend}} + \underbrace{S_t}_{\text{seasonal}} + \underbrace{I_t}_{\text{"irregular", stationary}}, \quad t \in \mathbb{Z}.$$

For the seasonal effect, we assume a period s:

$$S_t \approx S_{t+s}.$$

2.1. Methods based on differencing

Temporal differencing.
Assume that

$$T_{t-1} \approx T_t, \quad \text{i.e. a slowly varying trend.}$$

If there are no seasonal effects, the first temporal difference

$$\Delta X_t = X_t - X_{t-1} \approx I_t - I_{t-1}$$

is approximately stationary, since it has approximately no trend anymore

Seasonal differencing.
Assume that

$$S_{t-s} \approx S_t \quad \text{i.e. approximate seasonal effect.}$$

If there is no trend, the first seasonal difference

$$\Delta_s X_t = X_t - X_{t-s} \approx I_t - I_{t-s}$$

is approximately stationary, since it has approximately no seasonal effect anymore.

Temporal and seasonal differencing.
Assume that

$$\begin{aligned} T_{t-1} &\approx T_t \quad \text{i.e. a slowly varying trend,} \\ S_{t-s} &\approx S_t \quad \text{i.e. approximate seasonal effect.} \end{aligned}$$

Consider

$$\Delta_s \Delta X_t = (X_t - X_{t-1}) - (X_{t-s} - X_{t-s-1}) \approx \Delta_s \Delta I_t$$

has approximately no trend and no seasonal effect anymore.

Main disadvantage: $\Delta_s \Delta X_t$ has no natural interpretation.

2.2. Methods based on smoothing

We parameterize the seasonal effects as

$$S_t = \alpha_j \text{ for } t = j \text{ modulo } s, \quad \sum_{j=1}^{s} \alpha_j = 0.$$

Conceptually, we can now estimate the trend and seasonal effects as follows.

1. In absence of seasonal effects, i.e. $\alpha_j \equiv 0$:

   $$\hat{T}_t = (2m+1)^{-1}(X_{t-m} + X_{t-m+1} + \ldots + X_t + X_{t+1} + \ldots + X_{t+m}),$$

   where m is a reasonable window width.
   In case of seasonal effects (and assuming $s = 2m$ for simplicity):

   $$\hat{T}_t = s^{-1}(0.5 \cdot X_{t-m} + X_{t-m+1} + \ldots + X_{t+m-1} + 0.5 \cdot X_{t+m}).$$

   Thereby note that $s^{-1}(0.5 \cdot S_{t-m} + S_{t-m+1} + \ldots + S_{t+m-1} + 0.5 \cdot S_{t+m}) = 0$ since $\sum_{j=1}^{s} \alpha_j = 0$. And hence, $\hat{T}_t$ does not contain seasonal effects anymore.

2. Consider

   $$Y_t = X_t - \hat{T}_t \approx S_t + I_t.$$

   Estimate

   $$\hat{\alpha}_j = (k+1)^{-1}(Y_j + Y_{j+s} + \ldots Y_{j+ks}) \text{ where } k \text{ is maximal, i.e. } k \approx [n/s].$$

   Use

   $$\hat{S}_t = \hat{\alpha}_j \text{ for } t = j \text{ modulo } s.$$

3. Estimate trend T_t anew as in step 1, but based on $X_t - \hat{S}_t$ instead; and then, estimate seasonal effect anew as in step 2. Iterate the steps 1. and 2.

A refined version of such an iterative method is implemented in the `R`-package `stl`.

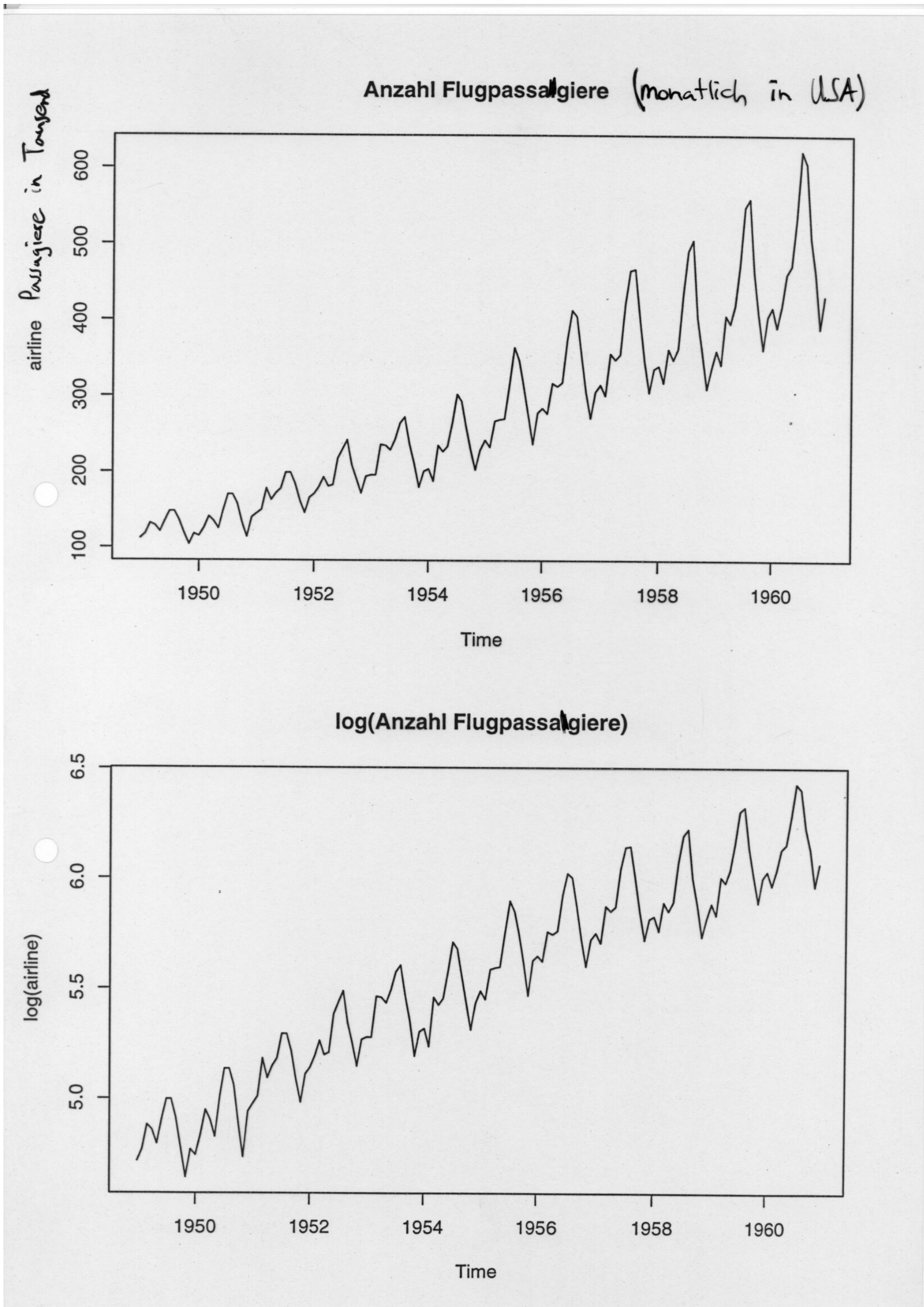

Figure 1: Top: Number of airline passengers per month in the USA (in thousands). Bottom: log-transformed values.

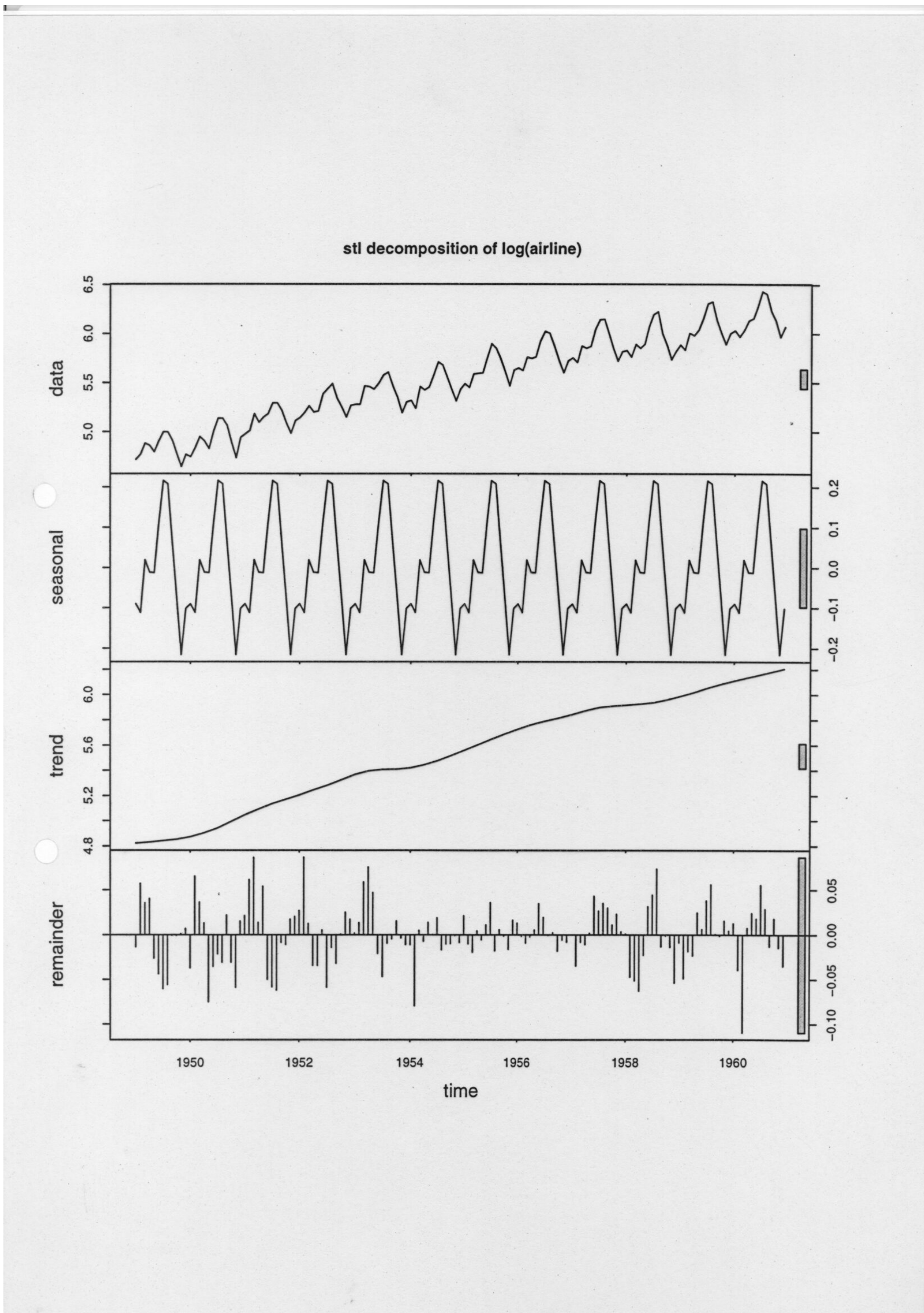

Figure 2: Seasonal-trend decomposition of log-transformed airline passenger data, using R-package stl.