universität
innsbruck

Time Series Analysis

Trends

Trends

Contents

Contents

- Deterministic vs. Stochastic Trends
- Time Series Decomposition
- Estimation of a Constant Mean
- Reliability and Efficiency of Regression Estimates
- Interpreting Regression Output
- Residual Analysis

Trends

Deterministic vs. Stochastic Trends

Deterministic vs. stochastic trends

Trends:

- Mean function over time.
- However, no (widely accepted) precise definition.
- Elusive, may depend on point of view.

Example:

- Random walk, e.g., `rwalk`.
- May be seen as having a trend, however, generated with zero mean.
- Hence, sometimes labeled *stochastic* trend.

Further examples:

- *Deterministic* trend, i.e., pattern continuing forever, e.g., linear or polynomial trend.
- *Seasonal/cyclical* patterns recurring regularly, e.g., every first quarter or every January etc.

Trends

Time Series Decomposition

Time series decomposition

Idea: Time series can be additively decomposed into a (smooth) trend T_t, seasonal/cyclical pattern S_t and error/remainder E_t:

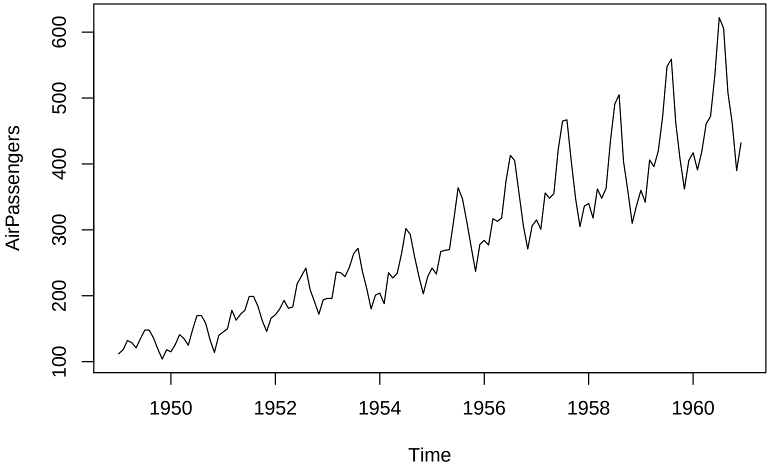
$$Y_t \ = \ T_t \ + \ S_t \ + \ E_t.$$

Alternatively: In case of a multiplicative model $Y_t = T_t \cdot S_t \cdot E_t$, taking logs yields the additive model.

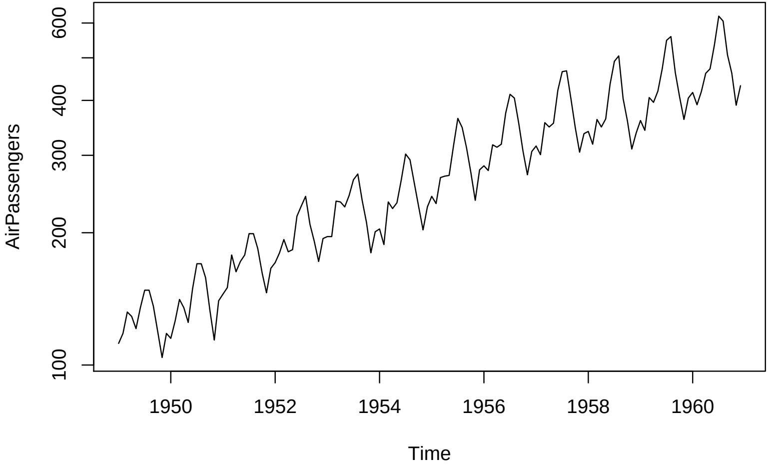
Example:

```
R> data("AirPassengers", package = "datasets")
R> plot(AirPassengers)
R> plot(AirPassengers, log = "y")
R> plot(log(AirPassengers))
```

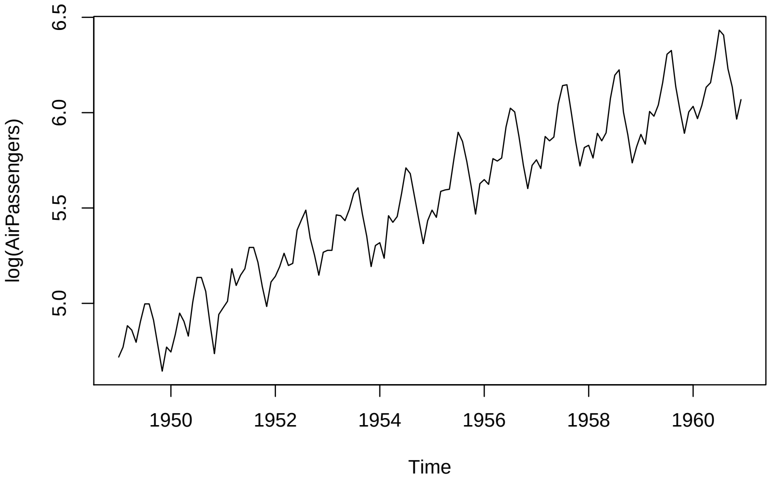
Time series decomposition



Time series decomposition



Time series decomposition



Time series decomposition

Method 1: Estimate T_t and S_t by OLS in a linear regression.

Examples:

- Linear trend: $T_t = \beta_1 + \beta_2 t$. (Corresponds to exponential trend in multiplicative model.)
- Polynomial trend of order p: $T_t = \beta_1 + \beta_2 t + \dots + \beta_{p+1} t^p$.
- Seasonal "dummies": $S_t = \beta_1 I_{\mathrm{Jan}}(t) + \dots + \beta_{12} I_{\mathrm{Dec}}(t)$.
- Harmonic pattern: $S_t = \beta_1 \cos(2\pi t) + \beta_2 \sin(2\pi t)$.

In R: Straightforward estimation via `lm()`.

Furthermore: Convenience functionality for time series regression in `dynlm()` from *dynlm*.

- Time series properties retained, e.g., in residuals.
- Functions in formula like `season()` for season dummies.
- Also: `d()` (for differences) and `L()` (for lags), details later.

Time series decomposition

Illustration: Linear trend plus month dummies.

$$\begin{aligned}\texttt{log(AirPassengers)}_t &= \beta_1 + \beta_2 \cdot \tilde{t} + \beta_3 \cdot I_{\mathsf{Feb}}(t) + \\ &\quad \dots + \beta_{13} \cdot I_{\mathsf{Dec}}(t) + E_t,\end{aligned}$$

where $\tilde{t} = t/12$ is the time in years.

```
R> library("dynlm")
R> ap <- log(AirPassengers)
R> ap_lm <- dynlm(ap ~ trend(ap) + season(ap))
R> summary(ap_lm)
Time series regression with "ts" data:
Start = 1949(1), End = 1960(12)

Call:
dynlm(formula = ap ~ trend(ap) + season(ap))

Residuals:
     Min       1Q   Median       3Q      Max
-0.15637 -0.04102  0.00368  0.04407  0.13232
```

Time series decomposition

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    4.72678    0.01889  250.18  < 2e-16
trend(ap)      0.12083    0.00143   84.40  < 2e-16
season(ap)Feb -0.02205    0.02421   -0.91   0.3640
season(ap)Mar  0.10817    0.02421    4.47  1.7e-05
season(ap)Apr  0.07690    0.02421    3.18   0.0019
season(ap)May  0.07453    0.02422    3.08   0.0025
season(ap)Jun  0.19668    0.02422    8.12  3.0e-13
season(ap)Jul  0.30062    0.02422   12.41  < 2e-16
season(ap)Aug  0.29132    0.02422   12.03  < 2e-16
season(ap)Sep  0.14669    0.02423    6.05  1.4e-08
season(ap)Oct  0.00853    0.02423    0.35   0.7254
season(ap)Nov -0.13519    0.02424   -5.58  1.3e-07
season(ap)Dec -0.02132    0.02425   -0.88   0.3808

Residual standard error: 0.0593 on 131 degrees of freedom
Multiple R-squared:  0.983,	Adjusted R-squared:  0.982
F-statistic:  649 on 12 and 131 DF,  p-value: <2e-16
```

Time series decomposition

Method 2: Linear filtering.

$$\{\tilde{Y}_t\} = \left\{ \sum_{j=-k}^{\ell} a_j Y_{t-j} \right\},$$

where $a_j, j = -k, \ldots, \ell$ are the filter weights. Appropriate choice of a_j can eliminate certain components, e.g., trend or season.

Example: Moving average of order q

$$\{\tilde{Y}_t\} = \left\{ \frac{1}{2q+1} \sum_{j=-q}^{q} Y_{t-j} \right\}.$$

Refinement: Weighted moving average for removing seasonality, e.g., $a = (1/8, 1/4, 1/4, 1/4, 1/8)$ for quarterly series.

Time series decomposition

More generally: Time series decomposition of seasonal series with frequency s via linear filters.

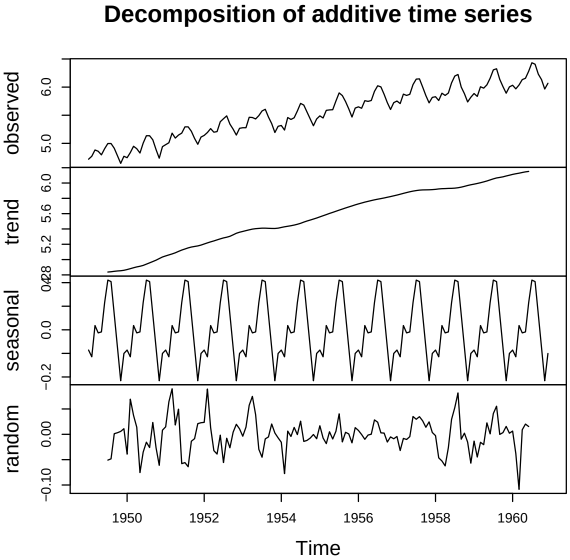
1. Estimate trend component T_t by weighted moving average with $s + 1$ weights $a = (0.5, 1, \dots, 1, 0.5)/s$.
2. Compute detrended series $Y_t - T_t$ and estimate seasonality figure by taking averages over common periods, e.g., all detrended January observations etc.

In R: `decompose()` for regular “`ts`” series.

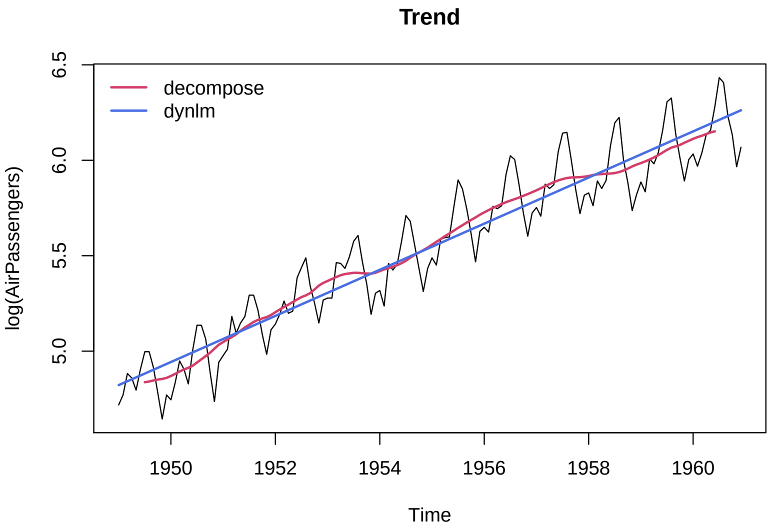
Illustration:

```
R> ap_dec <- decompose(ap)
R> plot(ap_dec)
```

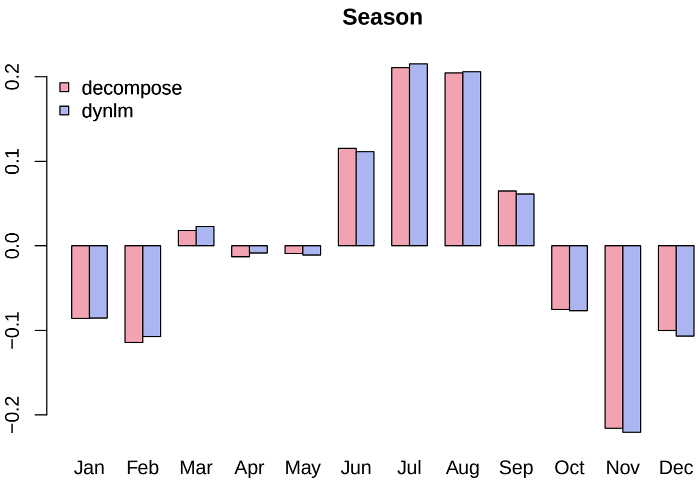
Time series decomposition



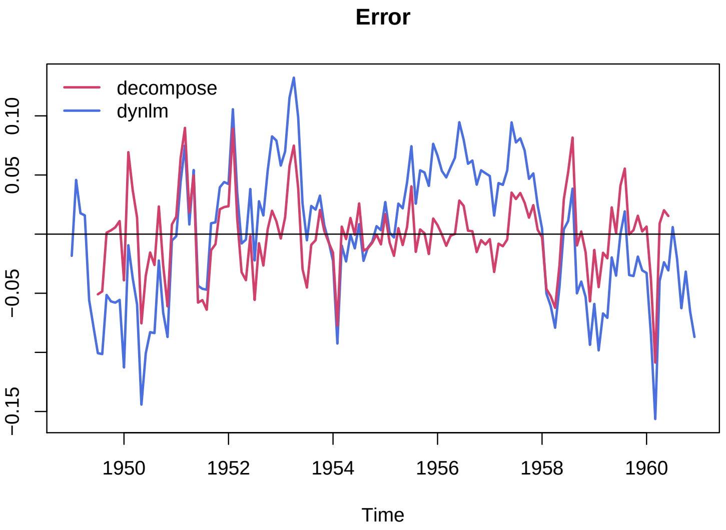
Time series decomposition



Time series decomposition



Time series decomposition



Time series decomposition

Comparison:

- In case of stable linear trend and stable additive seasonality, both methods lead to similar results. Then, linear regression is more efficient.
- Decomposition with linear filters can also deal with nonlinear trends. Seasonality has to be stable, though.
- Improved version of decomposition uses LOESS smoothing (instead of ordinary moving averages) both for trend and season component and can thus also deal with variations in the seasonality figure.

Further filters: In economics many filters are popular, e.g., X11, X12, X13, Hodrick-Prescott, Baxter-King, ... Choosing appropriate filters is an important topic in the time series literature in engineering.

Time series decomposition

In R:

- `stl()` for STL (season-trend decomposition via LOESS). For regular "`ts`" series.
- `filter()` for moving and recursive filters with arbitrary weights. For regular "`ts`" series or plain vectors.
- `rollapply()` in package *zoo* for applying arbitrary rolling functions, e.g., `mean()`, `median()`, `sd()`, etc. Generic function with methods for "`ts`", "`zoo`", and default (essentially numeric vectors).

Question: What are the theoretical properties of these estimation/decomposition/smoothing methods?

Answer: More difficult than in cross-section data with uncorrelated observations. Hence, consider simple cases first.

Trends

Estimation of a Constant Mean

Estimation of a constant mean

Consider: Assume

$$Y_t = \mu + X_t$$

where $\mathrm{E}(X_t) = 0$ for all t.

Clear: Unbiased estimate for μ is empirical mean

$$\bar{Y} = \frac{1}{n} \sum_{t=1}^{n} Y_t.$$

Question: What is the precision of $\bar{Y}$?

Answer: Further assumptions necessary. Assume that $\{Y_t\}$ (or equivalently $\{X_t\}$) is stationary with autocorrelation function ϱ_k.

Estimation of a constant mean

Then:

$$
\begin{aligned}
\mathrm{Var}(\bar{Y}) &= \frac{\gamma_0}{n}\left[\sum_{k=-n+1}^{n-1}\left(1-\frac{|k|}{n}\right)\varrho_k\right] \\
&= \frac{\gamma_0}{n}\left[1+2\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)\varrho_k\right].
\end{aligned}
$$

Example: $\{X_t\}$ white noise. Then all $\varrho_k = 0$ for $k > 0$ and $\mathrm{Var}(\bar{Y}) = \gamma_0/n$.

Example: $\{X_t\}$ stationary moving average $X_t = e_t - 0.5e_{t-1}$. Then $\varrho_1 = -0.4$ and $\varrho_k = 0$ for all $k > 1$.

Estimation of a constant mean

This yields

$$\begin{aligned}\mathrm{Var}(\bar{Y}) &= \frac{\gamma_0}{n}\left[1+2\left(1-\frac{1}{n}\right)(-0.4)\right] \\ &= \frac{\gamma_0}{n}\left[1-0.8\left(1-\frac{1}{n}\right)\right] \\ &\approx 0.2\,\frac{\gamma_0}{n}\end{aligned}$$

for large n.

Remarks:

- Negative correlation leads to higher precision.
- Conversely, positive correlation makes estimation more difficult.
- Mixed positive and negative correlations can have different net effects.

Estimation of a constant mean

Typically: Autocorrelation function is assumed to decay quickly enough

$$\sum_{k=0}^{\infty} |\varrho_k| < \infty.$$

Then, for large n

$$\mathrm{Var}(\bar{Y}) \approx \frac{\gamma_0}{n} \left[\sum_{k=-\infty}^{\infty} \varrho_k \right].$$

Example: Assume $\varrho_k = \phi^{|k|}$ for all k with $-1 < \phi < 1$. Then

$$\mathrm{Var}(\bar{Y}) \approx \frac{1+\phi}{1-\phi} \frac{\gamma_0}{n}.$$

Estimation of a constant mean

Example: $\{X_t\}$ random walk $X_t = \sum_{j=1}^{t} e_j$. Then

$$
\begin{aligned}
\operatorname{Var}(\bar{Y}) &= \frac{1}{n^2} \operatorname{Var}\left(\sum_{i=1}^{n} Y_i\right) \\
&= \frac{1}{n^2} \operatorname{Var}\left(\sum_{i=1}^{n} \sum_{j=1}^{i} e_j\right) \\
&= \frac{1}{n^2} \operatorname{Var}(e_n + 2e_{n-1} + \cdots + ne_1) \\
&= \frac{\sigma_e^2}{n^2} \sum_{k=1}^{n} k^2 \;=\; \sigma_e^2\,(2n+1)\,\frac{n+1}{6n},
\end{aligned}
$$

which increases with n.

Hence: Different estimation technique required for nonstationary series.

Trends

Reliability and Efficiency of Regression Estimates

Reliability and efficiency

Now: Assume that $\{Y_t\}$ has a deterministic trend μ_t with

$$Y_t \;=\; \mu_t \;+\; X_t,$$

where $\{X_t\}$ is a zero-mean stationary process with autocovariance γ_k and autocorrelation ϱ_k.

Estimation: Given a parametrization of $\mu_t = \mu(t, \beta)$ (e.g., linear, polynomial, seasonal, . . .), the parameter β are estimated by OLS.

Question: What are the properties of $\hat{\beta}$?

Reliability and efficiency

Example: Seasonal means $\mu(t, \beta) = \sum_{j=1}^{s} \beta_j \cdot I_j(t)$.
OLS estimates are simply seasonal averages

$$\hat{\beta}_j \;=\; \frac{1}{N} \sum_{i=0}^{N-1} Y_{j+s \cdot i},$$

where $j = 1, \ldots, s$ is the number of seasonal means (e.g., months) and N is the number of periods (e.g., years). Hence, assuming we only observe full periods, $n = N \cdot s$ is the number of observations.

Then: Analogous to the constant mean case

$$\mathrm{Var}(\hat{\beta}_j) \;=\; \frac{\gamma_0}{N} \left[1 + 2 \sum_{k=1}^{N-1} \left(1 - \frac{k}{N} \right) \varrho_{s \cdot k} \right].$$

Note that only the seasonal autocorrelations $\varrho_{s \cdot k}$ enter.

Reliability and efficiency

Example: Linear time trend $\mu(t, \beta) = \beta_1 + \beta_2 \cdot t$.
OLS estimate of slope is

$$\hat{\beta}_2 \;=\; \frac{\sum_{t=1}^{n}(t-\bar{t})Y_t}{\sum_{t=1}^{n}(t-\bar{t})^2}.$$

Then: As $\hat{\beta}_2$ is a linear combination of the Y_t

$$\mathrm{Var}(\hat{\beta}_2) \;=\; \frac{12\,\gamma_0}{n\,(n^2-1)}\,\left[1+\frac{24}{n\,(n^2-1)}\sum_{s=2}^{n}\sum_{t=1}^{s-1}(t-\bar{t})(s-\bar{t})\varrho_{s-t}\right],$$

using $\sum_{t=1}^{n}(t-\bar{t})^2 = n\,(n^2-1)/12$.

Reliability and efficiency

Special case: Assume $\varrho_k = 0$ for $k > 1$, then

$$\begin{aligned} \mathrm{Var}(\hat{\beta}_2) &= \frac{12\,\gamma_0}{n\,(n^2-1)} \left[1 + 2\varrho_1 \left(1 - \frac{3}{n}\right)\right] \\ &\approx \frac{12\,\gamma_0\,(1+2\varrho_1)}{n\,(n^2-1)}, \end{aligned}$$

if n is “large”.

Again: Variance is increased/reduced by positive/negative correlation in the disturbances (compared to the uncorrelated case).

Reliability and efficiency

Comparison: Ordinary least squares (OLS) vs. generalized least squares (GLS).

- GLS estimates are BLUE (best linear unbiased estimates).
- GLS requires full knowledge of covariance for $\{X_t\}$.
- BLUE can be approximated by iteratively estimating μ_t and covariance matrix of $\{X_t\}$.
- For deterministic trends (e.g., combinations of polynomial, seasonal, harmonic) and wide class of stochastic processes $\{X_t\}$, OLS is asymptotically as efficient as BLUE.
- However, OLS estimation of $Y_t = \beta Z_t + X_t$ may be inefficient or biased, even if n is large, if Z_t is also stochastic.

Interpreting regression output

Clear: If μ_t is correctly specified and $\{X_t\}$ is (Gaussian) white noise, then regression output can be interpreted "as usual".

Remarks: For more general $\{X_t\}$ caution may be necessary.

- If $\{X_t\}$ has constant variance it can be estimated via

$$\hat{\gamma}_0 = \frac{1}{n-p} \sum_{t=1}^{n} (Y_t - \hat{\mu}_t)^2,$$

  where $\hat{\mu}_t = \mu(t, \hat{\beta})$ and β is p-dimensional.
- Standard errors of $\hat{\beta}$ are only consistent if $\{X_t\}$ is white noise. Heteroskedasticity and autocorrelation consistent (HAC) standard errors may sometimes be used instead.
- Associated t statistics only have t_{n-p} distribution if $\{X_t\}$ is Gaussian white noise. Otherwise, normality may hold asymptotically (given consistent scaling).

Trends

Residual Analysis

Residual analysis

Idea: If trend μ_t is correctly specified (and consistently estimated), then the residuals

$$\hat{X}_t = Y_t - \hat{\mu}_t$$

should behave approximately like the stochastic component $\{X_t\}$.

Hence: Check assumptions about $\{X_t\}$ by assessing $\{\hat{X}_t\}$.

Furthermore: Consider scaled versions, e.g., standardized residuals $\hat{X}_t / \sqrt{\hat{\gamma}_0}$. R provides `rstudent()` for computing studentized residuals, i.e., scaled by leave-one-out standard deviations and hat values.

Residual analysis

Question: Is the trend μ_t correctly specified? Or, equivalently, does $\{X_t\}$ have a zero mean?

Clear: $\sum_{t=1}^{n} \hat{X}_t = 0$.

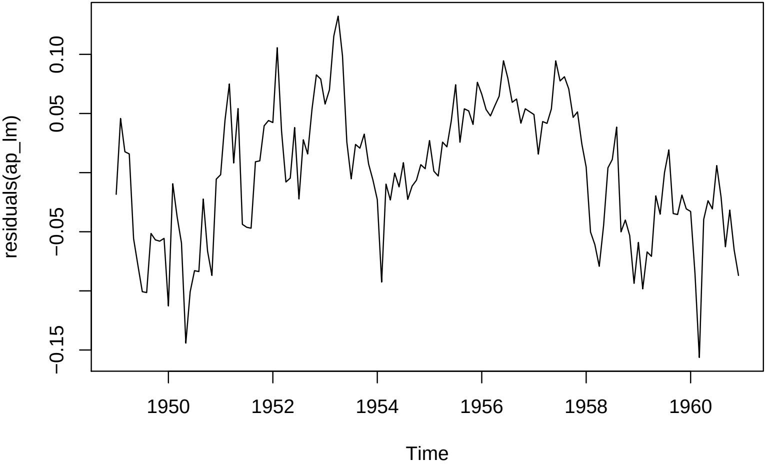
But: Are are there local deviations, i.e., remaining time trends?

Solution: Time series of (studentized) residuals with smoothed trend. Or scatter plot of residuals $\hat{X}_t$ against fitted values $\hat{\mu}_t$.

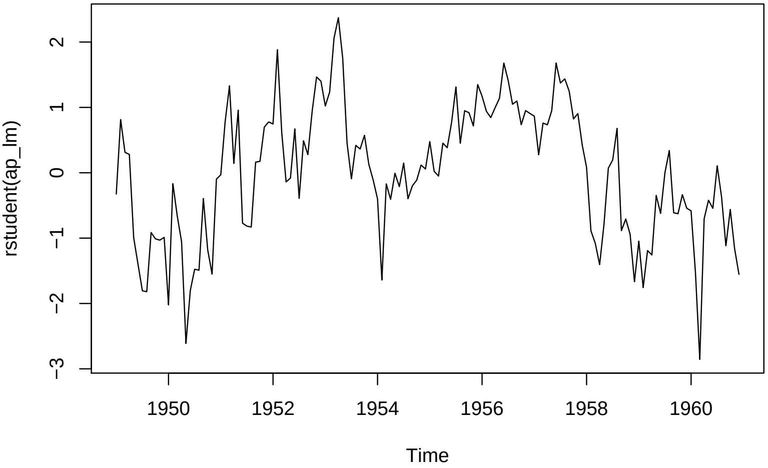
Illustration: Linear regression with season and trend for airline data.

```
R> plot(residuals(ap_lm))
R> plot(rstudent(ap_lm))
R> lines(rollapply(rstudent(ap_lm), 36, mean), col = 2)
R> plot(ap_lm, which = 1)
```

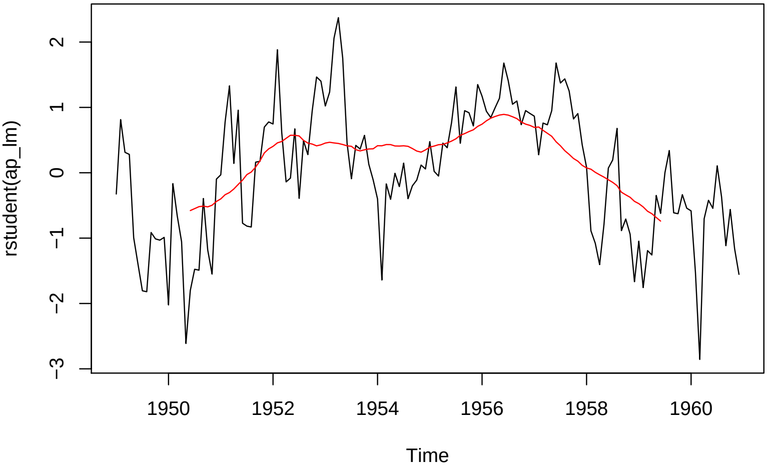
Residual analysis



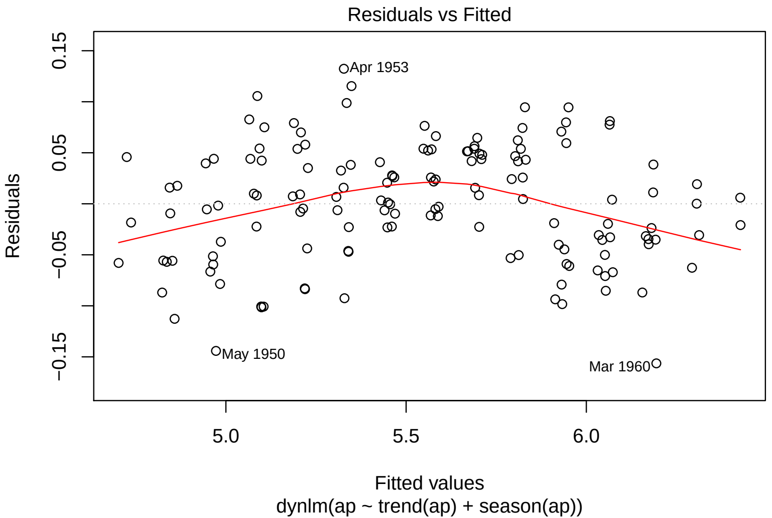
Residual analysis



Residual analysis



Residual analysis



Residual analysis

Question: Is $\{X_t\}$ (approximately) normal?

Answer:

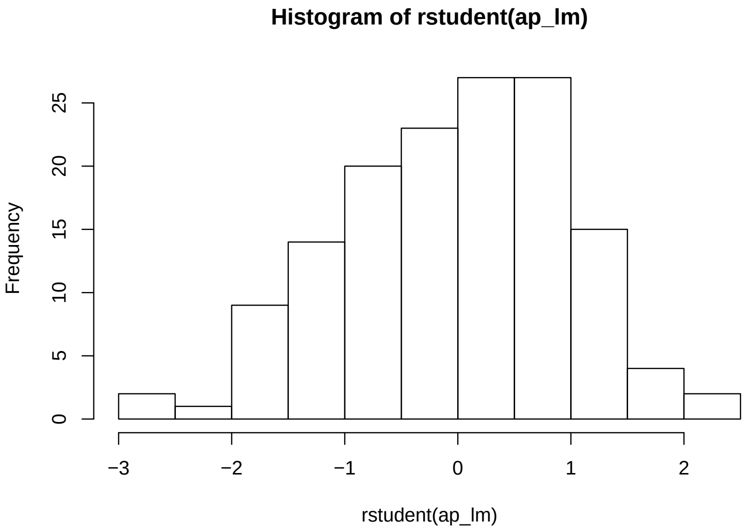
- Histogram of (standardized) residuals.
- QQ plot of empirical quantiles from residuals agings theoretical quantiles from standard normal distribution.
- Significance tests for normality, e.g., Shapiro-Wilk or Anderson-Darling test etc. *Caveat:* Normality is “only” null hypothesis.

Illustration:

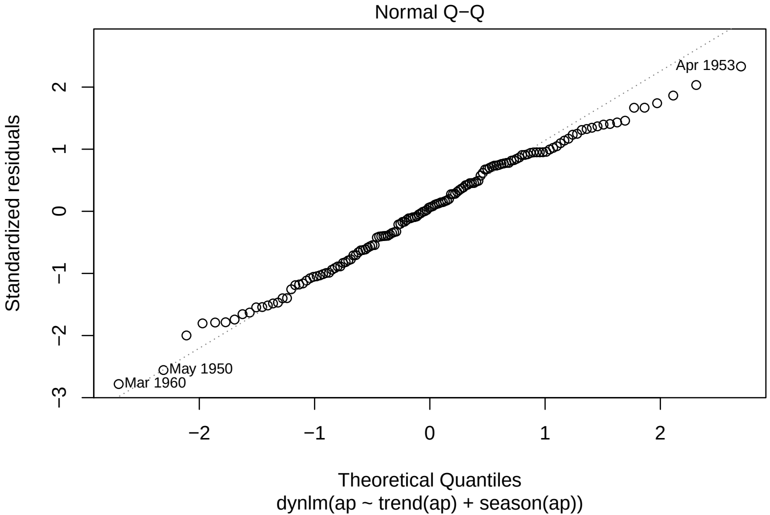
```
R> hist(rstudent(ap_lm))
R> plot(ap_lm, which = 2)
R> shapiro.test(residuals(ap_lm))
        Shapiro-Wilk normality test

data:  residuals(ap_lm)
W = 0.99, p-value = 0.3
```

Residual analysis



Residual analysis



Residual analysis

Question: Is $\{X_t\}$ white noise? Are the X_t uncorrelated?

Answer: Explore empirical autocorrelations in residuals.

- Visualize empirical autocorrelation function.
- Test hypothesis that selected $\varrho_k = 0$ (typically for $k = 1$), e.g., Durbin-Watson test or Breusch-Godfrey test, among many others.

If answer is no:

- Leave correlation unspecified but correct inference, e.g., using HAC standard errors (Newey-West or Andrews' kernel HAC standard errors).
- Employ different model, e.g., GLS instead of OLS.

Residual analysis

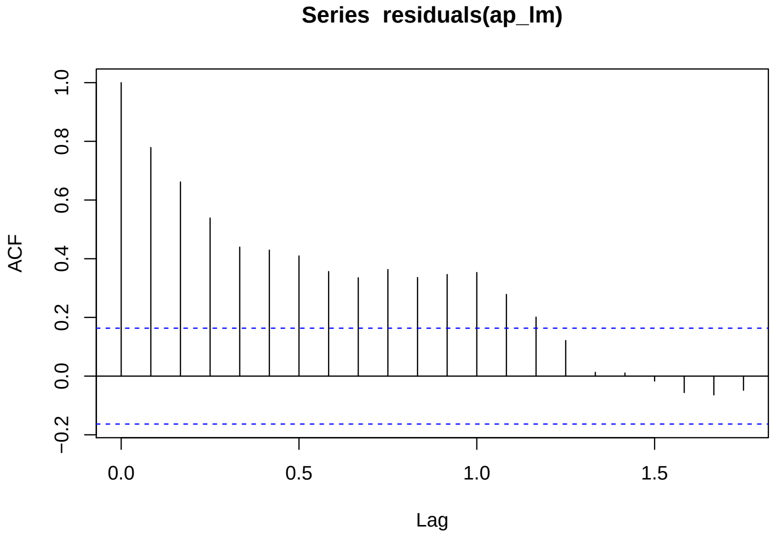
Definition: The *sample autocorrelation function* of a series $\{Y_t\}$ is

$$r_k = \frac{\sum_{t=k+1}^{n}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}.$$

Remarks:

- Assumes stationarity.
- Hence, "grand mean" $\bar{Y}$ and "grand sum of squares" are employed both for Y_t and lagged Y_{t-k}.
- Denominator contains n terms while numerator only has $n - k$.
- Approximate standard error for r_k from white noise series is $1/\sqrt{n}$.
- Plot of r_k vs. k is called *correlogram*. Often complemented with pointwise confidence intervals.

Residual analysis



Residual analysis

```
R> dwtest(ap_lm)

        Durbin-Watson test

data:  ap_lm
DW = 0.43, p-value <2e-16
alternative hypothesis: true autocorrelation is greater than 0
R> bgtest(ap_lm)

        Breusch-Godfrey test for serial correlation of order up
        to 1

data:  ap_lm
LM test = 89, df = 1, p-value <2e-16
```

Residual analysis

```
R> coeftest(ap_lm)
t test of coefficients:

               Estimate Std. Error t value Pr(>|t|)
(Intercept)     4.72678    0.01889  250.18  < 2e-16
trend(ap)       0.12083    0.00143   84.40  < 2e-16
season(ap)Feb -0.02205    0.02421   -0.91   0.3640
season(ap)Mar  0.10817    0.02421    4.47  1.7e-05
season(ap)Apr  0.07690    0.02421    3.18   0.0019
season(ap)May  0.07453    0.02422    3.08   0.0025
season(ap)Jun  0.19668    0.02422    8.12  3.0e-13
season(ap)Jul  0.30062    0.02422   12.41  < 2e-16
season(ap)Aug  0.29132    0.02422   12.03  < 2e-16
season(ap)Sep  0.14669    0.02423    6.05  1.4e-08
season(ap)Oct  0.00853    0.02423    0.35   0.7254
season(ap)Nov -0.13519    0.02424   -5.58  1.3e-07
season(ap)Dec -0.02132    0.02425   -0.88   0.3808
```

Residual analysis

```
R> coeftest(ap_lm, vcov = kernHAC)
t test of coefficients:

                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       4.72678    0.02847  166.04  < 2e-16
trend(ap)         0.12083    0.00474   25.48  < 2e-16
season(ap)Feb    -0.02205    0.01678   -1.31  0.19104
season(ap)Mar     0.10817    0.01967    5.50  1.9e-07
season(ap)Apr     0.07690    0.01721    4.47  1.7e-05
season(ap)May     0.07453    0.01991    3.74  0.00027
season(ap)Jun     0.19668    0.01865   10.55  < 2e-16
season(ap)Jul     0.30062    0.01915   15.70  < 2e-16
season(ap)Aug     0.29132    0.01866   15.61  < 2e-16
season(ap)Sep     0.14669    0.01344   10.92  < 2e-16
season(ap)Oct     0.00853    0.01316    0.65  0.51805
season(ap)Nov    -0.13519    0.01262  -10.71  < 2e-16
season(ap)Dec    -0.02132    0.00684   -3.12  0.00226
```

Residual analysis

```
R> coeftest(ap_lm, vcov = NeweyWest)
t test of coefficients:

              Estimate Std. Error t value Pr(>|t|)
(Intercept)    4.72678    0.02877  164.31  < 2e-16
trend(ap)      0.12083    0.00473   25.54  < 2e-16
season(ap)Feb -0.02205    0.01615   -1.37   0.1745
season(ap)Mar  0.10817    0.01860    5.82  4.4e-08
season(ap)Apr  0.07690    0.01497    5.14  9.9e-07
season(ap)May  0.07453    0.01746    4.27  3.8e-05
season(ap)Jun  0.19668    0.01660   11.85  < 2e-16
season(ap)Jul  0.30062    0.01840   16.34  < 2e-16
season(ap)Aug  0.29132    0.01784   16.33  < 2e-16
season(ap)Sep  0.14669    0.01148   12.78  < 2e-16
season(ap)Oct  0.00853    0.01245    0.69   0.4944
season(ap)Nov -0.13519    0.01190  -11.36  < 2e-16
season(ap)Dec -0.02132    0.00643   -3.32   0.0012
```