Mathematical Foundations of Time Series Analysis

English

Wolfgang Boehme

wolfgang.boehme@economic-modeling.de

www.economic-modeling.de

25 October 2026

The documentation on this website was developed with the support of the AI Claude (Anthropic).

Table of Contents

1 Introduction **2**
- 1.1 Notation . . . 2

2 Descriptive Statistics **3**
- 2.1 Measures of Location . . . 3
  - 2.1.1 Mean (Arithmetic Mean) . . . 3
  - 2.1.2 Median . . . 3
  - 2.1.3 Mode . . . 3
- 2.2 Measures of Dispersion . . . 3
  - 2.2.1 Variance . . . 3
  - 2.2.2 Standard Deviation . . . 3
- 2.3 Quantiles . . . 3
  - 2.3.1 Interquartile Range (IQR) . . . 3
- 2.4 Shape Measures . . . 4
  - 2.4.1 Skewness . . . 4
  - 2.4.2 Kurtosis (Excess Kurtosis) . . . 4

3 Trend Functions **5**
- 3.1 Linear Regression . . . 5
- 3.2 Polynomial Regression . . . 5
- 3.3 Exponential Regression . . . 5
- 3.4 Logarithmic Regression . . . 6
- 3.5 Goodness-of-Fit Measures . . . 6
  - 3.5.1 Coefficient of Determination (R^2) . . . 6
  - 3.5.2 Adjusted R^2 . . . 6
  - 3.5.3 Root Mean Squared Error (RMSE) . . . 6
  - 3.5.4 Mean Absolute Error (MAE) . . . 6

4 Autocorrelation Analysis **7**
- 4.1 Autocorrelation Function (ACF) . . . 7
- 4.2 Partial Autocorrelation Function (PACF) . . . 7
- 4.3 Ljung-Box Test . . . 7

5 Smoothing Methods **8**
- 5.1 Simple Moving Average (SMA) . . . 8
- 5.2 Exponential Moving Average (EMA) . . . 8
- 5.3 Weighted Moving Average (WMA) . . . 8
- 5.4 Holt-Winters Method . . . 8
- 5.5 LOESS (Locally Estimated Scatterplot Smoothing) . . . 9

6 Time Series Decomposition **10**
- 6.1 Additive Decomposition . . . 10
- 6.2 Multiplicative Decomposition . . . 10
- 6.3 Algorithm . . . 10

7 Filter Methods **11**
7.1 Hodrick-Prescott Filter 11
7.2 Butterworth Filter 11
7.3 Savitzky-Golay Filter 11

8 Seasonal Adjustment **12**
8.1 STL Decomposition 12
8.2 X-13 ARIMA-SEATS (Simplified) 12
8.3 Census I Method 13
8.4 Census II Method 13

9 Complete Example **14**
9.1 Descriptive Statistics 14
9.2 Linear Regression 14
9.3 Seasonal Decomposition (Period = 12) 14
9.4 HP Filter with $\lambda = 14{,}400$ 15

10 Summary **16**
10.1 Method Overview 16
10.2 Recommendations 16
10.3 Further Reading 16

1 Introduction

This document describes the mathematical foundations of all methods implemented in the Time Series Analyser. For each method, the underlying formulae, algorithms and – where appropriate – worked examples are presented.

1.1 Notation

We use the following notation:

- y_t: Observation at time t
- n: Number of observations
- μ: Mean (expected value)
- σ^2: Variance
- σ: Standard deviation

2 Descriptive Statistics

2.1 Measures of Location

2.1.1 Mean (Arithmetic Mean)

$$\mu = \bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t \tag{1}$$

Example: For the time series $y = [100, 102, 101, 105, 103]$:

$$\mu = \frac{100 + 102 + 101 + 105 + 103}{5} = \frac{511}{5} = 102.2$$

2.1.2 Median

The median is the middle value of the sorted time series:

$$\text{Median} = \begin{cases} y_{\frac{n+1}{2}} & \text{if } n \text{ is odd} \\ \dfrac{y_{\frac{n}{2}} + y_{\frac{n}{2}+1}}{2} & \text{if } n \text{ is even} \end{cases} \tag{2}$$

2.1.3 Mode

The mode is the most frequently occurring value in the time series.

2.2 Measures of Dispersion

2.2.1 Variance

$$\sigma^2 = \frac{1}{n}\sum_{t=1}^{n}(y_t - \mu)^2 \tag{3}$$

2.2.2 Standard Deviation

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \mu)^2} \tag{4}$$

Example: For $y = [100, 102, 101, 105, 103]$ with $\mu = 102.2$:

$$\begin{aligned} \sigma^2 &= \frac{(100 - 102.2)^2 + (102 - 102.2)^2 + \cdots + (103 - 102.2)^2}{5} \\ &= \frac{4.84 + 0.04 + 1.44 + 7.84 + 0.64}{5} = \frac{14.8}{5} = 2.96 \\ \sigma &= \sqrt{2.96} \approx 1.72 \end{aligned}$$

2.3 Quantiles

2.3.1 Interquartile Range (IQR)

$$Q_1 = 25\text{th percentile} \tag{5}$$

$$Q_3 = 75\text{th percentile} \tag{6}$$

$$\text{IQR} = Q_3 - Q_1 \tag{7}$$

2.4 Shape Measures

2.4.1 Skewness

Measures the asymmetry of the distribution:

$$\text{Skewness} = \frac{\frac{1}{n}\sum_{t=1}^{n}(y_t - \mu)^3}{\sigma^3} \tag{8}$$

Interpretation:

- Skewness $= 0$: Symmetric distribution
- Skewness > 0: Right-skewed (long right tail)
- Skewness < 0: Left-skewed (long left tail)

2.4.2 Kurtosis (Excess Kurtosis)

Measures the "peakedness" of the distribution:

$$\text{Kurtosis} = \frac{\frac{1}{n}\sum_{t=1}^{n}(y_t - \mu)^4}{\sigma^4} - 3 \tag{9}$$

Interpretation (excess kurtosis):

- Kurtosis $= 0$: Normal distribution
- Kurtosis > 0: Leptokurtic (more peaked, heavier tails)
- Kurtosis < 0: Platykurtic (flatter, lighter tails)

3 Trend Functions

3.1 Linear Regression

The linear trend function has the form:

$$\hat{y}_t = a + b \cdot t \tag{10}$$

The parameters are estimated by ordinary least squares:

$$b = \frac{n \sum_{t=1}^{n} t \cdot y_t - \sum_{t=1}^{n} t \sum_{t=1}^{n} y_t}{n \sum_{t=1}^{n} t^2 - \left(\sum_{t=1}^{n} t\right)^2} \tag{11}$$

$$a = \frac{\sum_{t=1}^{n} y_t - b \sum_{t=1}^{n} t}{n} \tag{12}$$

Example: For $y = [100, 102, 105, 108, 110]$ with $t = [0, 1, 2, 3, 4]$:

$$b = \frac{5 \cdot 1130 - 10 \cdot 525}{5 \cdot 30 - 100} = \frac{5650 - 5250}{150 - 100} = \frac{400}{50} = 2.5$$
$$a = \frac{525 - 2.5 \cdot 10}{5} = \frac{500}{5} = 100$$

Trend function: $\hat{y}_t = 100 + 2.5t$

3.2 Polynomial Regression

General form of a polynomial of degree p:

$$\hat{y}_t = a_0 + a_1 t + a_2 t^2 + \cdots + a_p t^p = \sum_{k=0}^{p} a_k t^k \tag{13}$$

The coefficients are determined by solving the normal equations:

$$\mathbf{X}^T \mathbf{X}\, \mathbf{a} = \mathbf{X}^T \mathbf{y} \tag{14}$$

where $\mathbf{X}$ is the Vandermonde matrix:

$$\mathbf{X} = \begin{pmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^p \\ 1 & t_2 & t_2^2 & \cdots & t_2^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & t_n^2 & \cdots & t_n^p \end{pmatrix} \tag{15}$$

3.3 Exponential Regression

Form:

$$\hat{y}_t = a \cdot e^{bt} \tag{16}$$

Taking logarithms reduces this to a linear problem:

$$\ln(\hat{y}_t) = \ln(a) + bt \tag{17}$$

3.4 Logarithmic Regression

Form:

$$\hat{y}_t = a + b \cdot \ln(t+1) \tag{18}$$

3.5 Goodness-of-Fit Measures

3.5.1 Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2} = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \tag{19}$$

$R^2 \in [0, 1]$; higher values indicate a better fit.

3.5.2 Adjusted R^2

Accounts for the number of parameters:

$$R^2_{\text{adj}} = 1 - \frac{(1 - R^2)(n-1)}{n-k-1} \tag{20}$$

where k is the number of independent variables.

3.5.3 Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \tag{21}$$

3.5.4 Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| \tag{22}$$

4 Autocorrelation Analysis

4.1 Autocorrelation Function (ACF)

The ACF measures the linear dependence between observations separated by lag k:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\frac{1}{n}\sum_{t=1}^{n-k}(y_t - \mu)(y_{t+k} - \mu)}{\frac{1}{n}\sum_{t=1}^{n}(y_t - \mu)^2} \tag{23}$$

where γ_k is the autocovariance at lag k.
Properties:

- $\rho_0 = 1$ (perfect correlation with itself)
- $-1 \leq \rho_k \leq 1$
- $\rho_k = \rho_{-k}$ (symmetry)

4.2 Partial Autocorrelation Function (PACF)

The PACF measures the correlation between y_t and y_{t+k} after removing the influence of all intermediate lags:

$$\phi_{kk} = \text{PACF}(k) \tag{24}$$

Computed via Yule-Walker equations or the Durbin-Levinson algorithm.

4.3 Ljung-Box Test

Tests for significant autocorrelation:

$$Q = n(n+2)\sum_{k=1}^{h}\frac{\rho_k^2}{n-k} \tag{25}$$

Q follows asymptotically a χ^2 distribution with h degrees of freedom.
Hypotheses:

- H_0: No autocorrelation up to lag h
- H_1: Significant autocorrelation present

H_0 is rejected when $Q > \chi^2_{0.05,\,h}$ (5% significance level).

5 Smoothing Methods

5.1 Simple Moving Average (SMA)

Simple moving average with window size m:

$$\hat{y}_t = \frac{1}{m} \sum_{i=0}^{m-1} y_{t-i} \tag{26}$$

Example: For $y = [100, 102, 105, 108, 110]$ with $m = 3$:

$$\hat{y}_3 = \frac{100 + 102 + 105}{3} = 102.33$$
$$\hat{y}_4 = \frac{102 + 105 + 108}{3} = 105.00$$
$$\hat{y}_5 = \frac{105 + 108 + 110}{3} = 107.67$$

5.2 Exponential Moving Average (EMA)

Exponential smoothing with parameter α:

$$\hat{y}_t = \alpha y_t + (1 - \alpha)\hat{y}_{t-1} \tag{27}$$

where $\alpha = \frac{2}{m+1}$ for window size m.
Properties:

- More recent observations receive larger weights
- $\alpha \in (0, 1)$: smaller α = stronger smoothing

5.3 Weighted Moving Average (WMA)

Linearly weighted moving average:

$$\hat{y}_t = \frac{\sum_{i=0}^{m-1}(m - i) \cdot y_{t-i}}{\sum_{i=0}^{m-1}(m - i)} = \frac{\sum_{i=0}^{m-1}(m - i) \cdot y_{t-i}}{\frac{m(m+1)}{2}} \tag{28}$$

5.4 Holt-Winters Method

Double exponential smoothing for series with a trend:

$$\text{Level:} \quad L_t = \alpha y_t + (1 - \alpha)(L_{t-1} + T_{t-1}) \tag{29}$$
$$\text{Trend:} \quad T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \tag{30}$$
$$\text{Forecast:} \quad \hat{y}_{t+h} = L_t + h \cdot T_t \tag{31}$$

where α is the smoothing parameter for the level and β is the smoothing parameter for the trend.

5.5 LOESS (Locally Estimated Scatterplot Smoothing)

Local polynomial regression with a weighting function:

$$\hat{y}_t = \sum_{i \in N_t} w_i(t)\, y_i \tag{32}$$

Weights are based on the tri-cube function:

$$w_i(t) = \left(1 - \left|\frac{t-i}{d_{\max}}\right|^3\right)^3 \tag{33}$$

where N_t denotes the h nearest neighbours of t and $d_{\max}$ is the maximum distance.

6 Time Series Decomposition

A time series can be decomposed into three components:

- Trend (T_t): Long-term development
- Seasonality (S_t): Periodic fluctuations
- Remainder (R_t): Random fluctuations

6.1 Additive Decomposition

$$y_t = T_t + S_t + R_t \tag{34}$$

Appropriate when the amplitude of seasonal fluctuations is constant.

6.2 Multiplicative Decomposition

$$y_t = T_t \cdot S_t \cdot R_t \tag{35}$$

Appropriate when the amplitude is proportional to the level.

6.3 Algorithm

Step 1: Trend extraction with a centred moving average of period p:

$$T_t = \frac{1}{p} \sum_{i=-\lfloor p/2 \rfloor}^{\lfloor p/2 \rfloor} y_{t+i} \tag{36}$$

Step 2: De-trending:

$$D_t = \begin{cases} y_t - T_t & \text{(additive)} \\ y_t / T_t & \text{(multiplicative)} \end{cases} \tag{37}$$

Step 3: Seasonal component by averaging over periods:

$$S_j = \frac{1}{m} \sum_{i=1}^{m} D_{j+ip} \quad \text{for } j = 1, \ldots, p \tag{38}$$

where m is the number of complete periods.

Step 4: Normalisation:

$$S_j^* = \begin{cases} S_j - \dfrac{1}{p} \sum_{j=1}^{p} S_j & \text{(additive)} \\ S_j \,/\, \sqrt[p]{\prod_{j=1}^{p} S_j} & \text{(multiplicative)} \end{cases} \tag{39}$$

Step 5: Remainder component:

$$R_t = \begin{cases} y_t - T_t - S_t & \text{(additive)} \\ y_t / (T_t \cdot S_t) & \text{(multiplicative)} \end{cases} \tag{40}$$

Example (additive): For monthly data with $p = 12$ and annual data:

$$y_{13} = 105, \quad T_{13} = 100, \quad S_1 = 8$$
$$R_{13} = 105 - 100 - 8 = -3$$

7 Filter Methods

7.1 Hodrick-Prescott Filter

The HP filter decomposes a time series into trend (τ_t) and cycle (c_t):

$$y_t = \tau_t + c_t \tag{41}$$

The trend is determined by minimising the following objective function:

$$\min_{\{\tau_t\}} \left\{ \sum_{t=1}^{n} (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{n-1} \left[(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}) \right]^2 \right\} \tag{42}$$

The first term measures goodness of fit; the second term penalises curvature in the trend.

Parameter λ:

- $\lambda = 100$: Annual data
- $\lambda = 1{,}600$: Quarterly data
- $\lambda = 14{,}400$: Monthly data
- $\lambda \to 0$: Trend follows the data exactly
- $\lambda \to \infty$: Trend becomes linear

Solution by Gaussian elimination of the system:

$$(\mathbf{I} + \lambda \mathbf{K}^T \mathbf{K})\, \boldsymbol{\tau} = \mathbf{y} \tag{43}$$

where $\mathbf{K}$ is the second-difference matrix.

7.2 Butterworth Filter

Bidirectional exponential smoothing:

$$\text{Forward:} \quad f_t = \alpha y_t + (1 - \alpha) f_{t-1} \tag{44}$$

$$\text{Backward:} \quad b_t = \alpha y_t + (1 - \alpha) b_{t+1} \tag{45}$$

$$\text{Trend:} \quad \tau_t = \frac{f_t + b_t}{2} \tag{46}$$

7.3 Savitzky-Golay Filter

Local polynomial fitting with window size m and polynomial degree p:

$$\hat{y}_t = \sum_{j=-\lfloor m/2 \rfloor}^{\lfloor m/2 \rfloor} c_j \, y_{t+j} \tag{47}$$

The coefficients c_j are determined by least-squares fitting of a polynomial within the local window.

8 Seasonal Adjustment

8.1 STL Decomposition

STL (Seasonal and Trend decomposition using Loess) is a robust method for decomposing time series.

Algorithm:

Inner loop:

1. **De-trending:** Remove trend with LOESS: $D_t = y_t - T_t$
2. **Cycle-subseries smoothing:** Smooth each subseries of the period
3. **Low-pass filter:** Smooth the seasonal component
4. **Deseasonalising:** $V_t = y_t - S_t$
5. **Trend smoothing:** Smooth V_t with LOESS

Outer loop: Robustification against outliers

Result:

$$y_t = T_t + S_t + R_t \tag{48}$$

8.2 X-13 ARIMA-SEATS (Simplified)

X-13 is the standard method for official statistics.

Simplified algorithm:

Step 1: Trend with moving average of period p:

$$T_t = \frac{1}{p} \sum_{i=-\lfloor p/2 \rfloor}^{\lfloor p/2 \rfloor} y_{t+i} \tag{49}$$

Step 2: Seasonal factors:

$$SF_j = \frac{1}{m} \sum_{i=1}^{m} (y_{j+ip} - T_{j+ip}) \quad \text{for } j = 1, \ldots, p \tag{50}$$

Step 3: Normalisation:

$$SF_j^* = SF_j - \frac{1}{p} \sum_{j=1}^{p} SF_j \tag{51}$$

Step 4: Seasonally adjusted series:

$$SA_t = y_t - SF_{(t \bmod p)} \tag{52}$$

8.3 Census I Method

Ratio-to-moving-average method:

Step 1: Centred moving average:

$$MA_t = \frac{1}{p} \sum_{i=-\lfloor p/2 \rfloor}^{\lfloor p/2 \rfloor} y_{t+i} \tag{53}$$

Step 2: Ratios to the moving average:

$$r_t = \frac{y_t}{MA_t} \tag{54}$$

Step 3: Seasonal indices by averaging:

$$SI_j = \frac{1}{m} \sum_{i=1}^{m} r_{j+ip} \quad \text{for } j = 1, \ldots, p \tag{55}$$

Step 4: Normalisation to period p:

$$SI_j^* = \frac{SI_j}{\frac{1}{p} \sum_{j=1}^{p} SI_j} \cdot p \tag{56}$$

Step 5: Seasonally adjusted series:

$$SA_t = \frac{y_t}{SI^*_{(t \bmod p)}} \tag{57}$$

8.4 Census II Method

Extended version of Census I with multiple passes:

Pass 1: Initial trend estimate with moving average

Pass 2: Seasonal-irregular component:

$$SI_t = y_t - T_t \tag{58}$$

Pass 3: Seasonal component by averaging

Pass 4: Smoothing of the seasonal component with a 3-term moving average:

$$S_j^* = \frac{S_{j-1} + S_j + S_{j+1}}{3} \tag{59}$$

Pass 5: Centring:

$$S_j^{**} = S_j^* - \frac{1}{p} \sum_{j=1}^{p} S_j^* \tag{60}$$

Pass 6: Seasonally adjusted series:

$$SA_t = y_t - S^{**}_{(t \bmod p)} \tag{61}$$

Pass 7: Re-estimation of the trend from SA_t

9 Complete Example

Consider a monthly time series over 2 years (24 observations):

$y = [100, 95, 98, 102, 110, 120, 125, 118, 112, 105, 98, 95, 105, 98, 102, 108, 118, 130, 138, 128, 120, 110, 102, 9$

9.1 Descriptive Statistics

$$
\begin{aligned}
n &= 24 \\
\mu &= \frac{1}{24}\sum_{t=1}^{24} y_t = \frac{2653}{24} = 110.54 \\
\sigma^2 &= \frac{1}{24}\sum_{t=1}^{24}(y_t - 110.54)^2 = 156.41 \\
\sigma &= \sqrt{156.41} = 12.51
\end{aligned}
$$

9.2 Linear Regression

With $t = 0, 1, 2, \ldots, 23$:

$$
\begin{aligned}
b &= \frac{24 \cdot 28626 - 276 \cdot 2653}{24 \cdot 4324 - 76176} = \frac{687024 - 732228}{103776 - 76176} = \frac{-45204}{27600} = -1.64 \\
a &= \frac{2653 - (-1.64) \cdot 276}{24} = \frac{3105.64}{24} = 129.4
\end{aligned}
$$

Trend function: $\hat{y}_t = 129.4 - 1.64t$

The negative trend indicates a slight decline over time (which does not match the visual impression – illustrating the limitation of a purely linear fit on seasonal data).

9.3 Seasonal Decomposition (Period = 12)

Step 1: Trend with 12-month moving average (for month 7):

$$T_7 = \frac{1}{12}\sum_{i=1}^{12} y_{i+1} = \frac{1}{12}(95 + 98 + \cdots + 105) = 108.5$$

Step 2: De-trending (additive):

$$D_7 = y_7 - T_7 = 125 - 108.5 = 16.5$$

Step 3: Seasonal components for each month:

$$S_{\text{Jan}} = \frac{(y_1 - T_1) + (y_{13} - T_{13})}{2} \approx -10.5$$

Step 4: Seasonally adjusted series:

$$SA_7 = y_7 - S_7 = 125 - 16.5 = 108.5$$

9.4 HP Filter with $\lambda = 14{,}400$

The system $(\mathbf{I} + \lambda \mathbf{K}^T \mathbf{K})\, \boldsymbol{\tau} = \mathbf{y}$ is solved.

For $t = 7$, for example:

$$\tau_7 \approx 109.2$$

The cycle is then:

$$c_7 = y_7 - \tau_7 = 125 - 109.2 = 15.8$$

10 Summary

10.1 Method Overview

Category	Methods
Statistics	Mean, Median, Mode, Variance, Skewness, Kurtosis
Trend	Linear, Polynomial (degrees 2–6), Exponential, Logarithmic
Autocorrelation	ACF, PACF, Ljung-Box test
Smoothing	SMA, EMA, WMA, Holt-Winters, LOESS
Decomposition	Additive/Multiplicative (trend, seasonal, remainder)
Filters	Hodrick-Prescott, Butterworth, Savitzky-Golay
Seasonal adjustment	STL, X-13, Census I, Census II

10.2 Recommendations

- **Trend identification:** Start with linear regression; increase the polynomial degree only when necessary.
- **Smoothing:** LOESS and Holt-Winters for complex patterns; SMA/EMA for simple cases.
- **Seasonal adjustment:** STL is robust and flexible; X-13 for official statistics.
- **Trend-cycle separation:** HP filter with an appropriate λ.

10.3 Further Reading

- Brockwell, P. J., & Davis, R. A. (2016). *Introduction to Time Series and Forecasting.* Springer.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice.* OTexts.
- Cleveland, R. B., et al. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics*, 6(1), 3–73.
- Hodrick, R. J., & Prescott, E. C. (1997). Postwar US Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking*, 29(1), 1–16.