Statistik - 3. Zeitreihenanalyse - Teil 1

Dr. Wolfgang Böhme

Januar 2013

Gliederung

- 1. Einführung
- 2. Deskriptive univariate Verteilungsanalyse
- 3. Deskriptive Korrelations- und Regressionsanalyse
- **4. Deskriptive Zeitreihenanalyse**
- 5. Verhältniszahlen / Indexzahlen
- 6. Einführung in ausgewählte multivariate Analysemethoden

- Statistiksoftware: SPSS

Gliederung

- **4.0 Definitionen**
- 4.1 Gleitende Durchschnitte
- 4.2 Exponetielle Glättung
- 4.3 Trendfunktionen
- 4.4 Filter
- 4.5 Autokorrelation
- 4.6 Ausblick: Multiple Zeitreihenanalyse

Zeitreihen

Eine **Zeitreihe** ist eine zeitlich geordnete Folge von Merkmalsausprägungen eines statistischen Merkmals. In der Regel wird angenommen, dass die Beobachtungswerte in gleichen zeitlichen Abständen vorliegen.
Bezeichnung: $\{y_t,\ t = 1, 2, ..., n\}$

In ökonomischen untersuchungen wird unterschieden:

- Jahresdaten (z.B. Bruttoinlandprodukt)
- Quartalsdaten (z.B. Bruttoinlandprodukt)
- Monatsdaten (z.B. Preisindizes)
- Tagesdaten (z.B. börsennotierte Preise, Zinssätze)

Zeitreihen

Die unterjährige Zeitreihe $\{y_t,\ t = 1, 2, ..., n\}$ ist das Resultat des Zusammenwirkens von folgenden Komponenten:

- glatte Komponente (g_t) (Trendkomponente (T_t), Konjunkturkomponente oder zyklische Komponente (Z_t))
- Saisonkomponente (s_t)
- Irreguläre Komponente (i_t) (nicht auf (g_t) und (s_t) zurückführbar, Residualgröße)

zwei Zeitreihenmodelle (Verknüpfung der Komponenten)

1. additive Verknüpfung:

$$y_t = g_t + s_t + i_t$$

2. multiplikative Verknüpfung:

$$y_t = g_t \cdot s_t \cdot i_t$$

- **4.1 Gleitende Durchschnitte**
- 4.2 Exponetielle Glättung
- 4.3 Trendfunktionen
- 4.4 Filter
- 4.5 Autokorrelation
- 4.6 Ausblick: Multiple Zeitreihenanalyse

4.1 Gleitende Durchschnitte - Gliederung

- **4.1.1 Zentrierte gleitende Durchschnitte**
- 4.1.2 Zurückgreifender gleitender Durchschnitt

Zentrierte gleitende Durchschnitte

Definition: Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, ..., n\}$ und k sei eine ganzzahlige Konstante $(k \geq 0)$, dann heißt die Zeitreihe $\{\overline{y_j},\ j = k+1, k+2, ..., n-k\}$ der n-2k arithmetischen Mittel $\overline{y_j}$:

$$\overline{y_j} = \frac{1}{r} \sum_{i=-k}^{k} y_{i+j}$$

zentrierter gleitender Durchschnitt zum **ungeraden** Stützbereich $r = 2 \cdot k + 1$.

Definition: Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, ..., n\}$ und k sei eine ganzzahlige Konstante $(k \geq 0)$, dann heißt die Zeitreihe $\{\overline{y_j},\ j = k+1, k+2, ..., n-k+1\}$ der n-2k arithmetischen Mittel $\overline{y_j}$:

$$\overline{y_j} = \frac{1}{r} \sum_{i=-k}^{k-1} y_{i+j}$$

zentrierter gleitender Durchschnitt zum **geraden** Stützbereich $r = 2 \cdot k$.

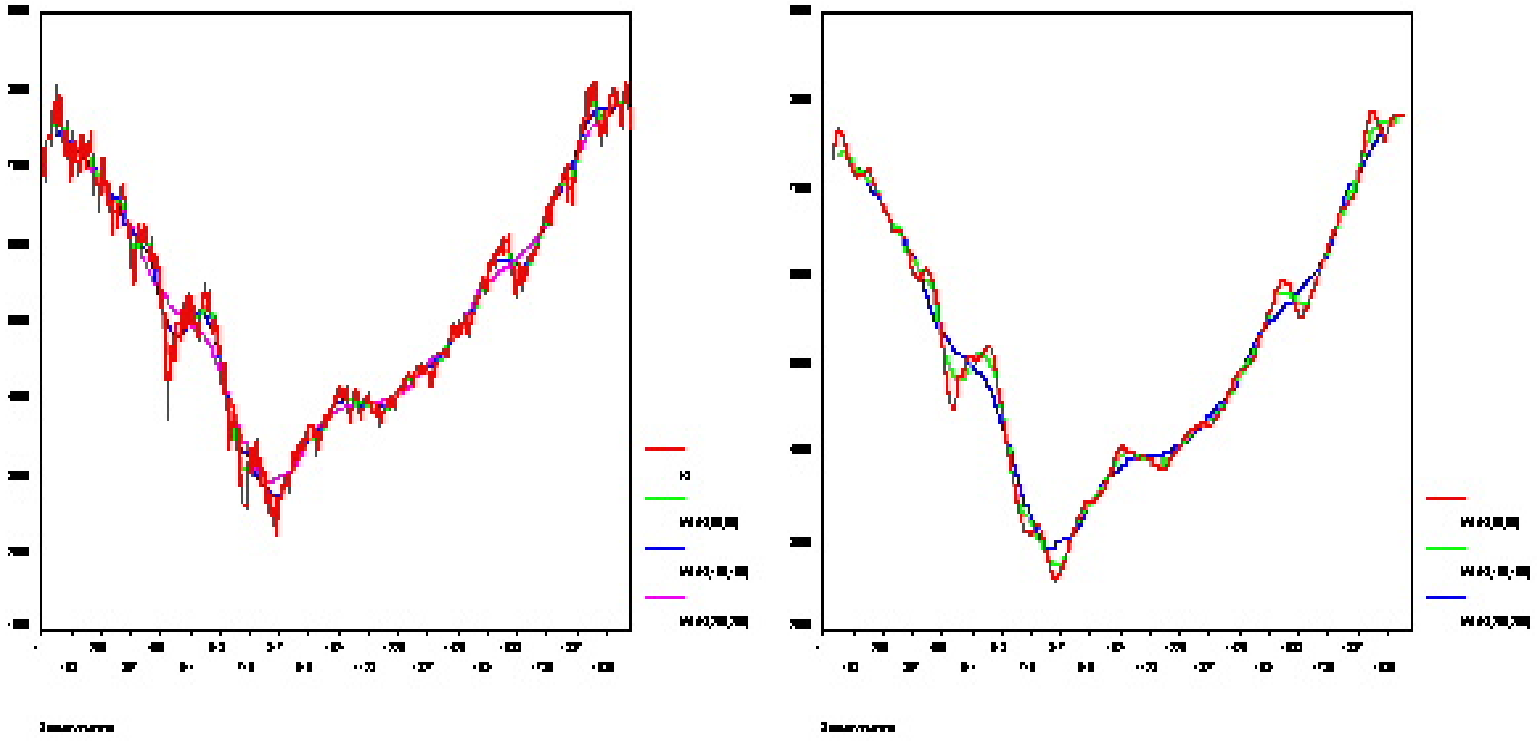
ein einfaches Beispiel

Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, ..., n\} = 1, 2, 4, 7, 9, 10, 11$

1. ungerader Stützbereich: $r = 3$ bzw. $k = 1$
2. gerader Stützbereich:$r = 2$ bzw. $k = 1$
3. Spezialfall: ungerader Stützbereich: $r = 1$ bzw. $k = 0$

ein Beispiel aus der Finanzmathematik

DAX-Entwicklung und zentrierte gleitende Durchschnitte (r = 50,100,200)



4.1 Gleitende Durchschnitte - Gliederung

- 4.1.1 Zentrierte gleitende Durchschnitte
- **4.1.2 Zurückgreifender gleitender Durchschnitt**

Zurückgreifende gleitende Durchschnitte

Definition: Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, ..., n\}$ und k sei eine ganzzahlige Konstante ($k \geq 0$), dann heißt die Zeitreihe $\{\overline{y_j},\ j = 2k + 1, 2k + 2, ..., n\}$ der n-2k arithmetischen Mittel $\overline{y_j}$:

$$\overline{y_j} = \frac{1}{r} \sum_{i=-2k}^{0} y_{i+j}$$

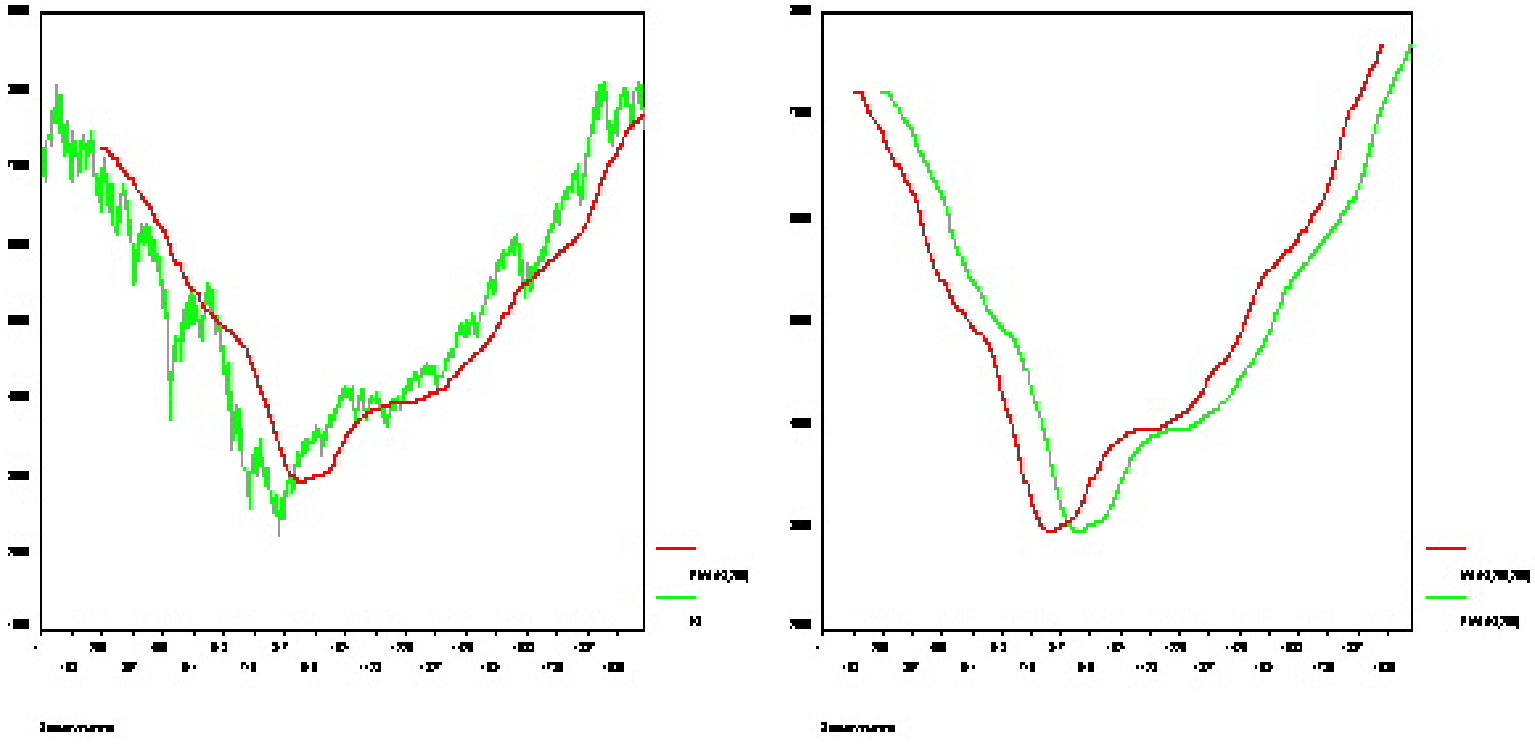
Zurückgreifende gleitender Durchschnitt zum Stützbereich $r = 2 \cdot k + 1$.

Beispiel: Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, ..., n\} = 1, 2, 4, 7, 9, 10, 11$

1. Stützbereich: $r = 3$ bzw. $k = 1$
2. Spezialfall: Stützbereich: $r = 1$ bzw. $k = 0$

ein Beispielaus der Finanzmathematik

DAX-Entwicklung und zurückgesetze gleitende Durchschnitte (r = 50,100,200)



- 4.1 Gleitende Durchschnitte
- **4.2 Exponetielle Glättung**
- 4.3 Trendfunktionen
- 4.4 Filter
- 4.5 Autokorrelation
- 4.6 Ausblick: Multiple Zeitreihenanalyse

4.2 Exponetielle Glättung - Gliederung

- **4.2.1 Grundmodell**
- 4.2.2 Berücksichtigung des Linearen Trends
- 4.2.3 Berücksichtigung der Saisonkomponente

Grundmodell

Intension: einfaches Prognosemodell für Ein-Schrittprognose (für y_{n+1})

Definition: Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, ..., n\}$. Das einfache Prognosemodell

$$y_t^* = \alpha y_t + (1 - \alpha) y_{t-1}^*$$

heißt **Grundmodell der exponentiellen Glättung**.

Bemerkungen:

1. y_t^* = Prognosewert für y_{t+1}
2. α - Damping factor (Gewichte) (α nahe 1 - kaum Glättung) $\Rightarrow$ Beispiel
3. Wahl des Parameter α nach dem Kriterium $\sum_{t=2}^{n} (y_t - y_{t-1}^*)^2 = MIN$
4. $y_t^* = F(y_t, y_{t-1}^*)$, d.h. hängt ab von den Beobachtung y_t und dem Schätzwert der Vorperiode y_{t-1}^*
5. $y_t^* = \alpha y_t + \alpha(1 - \alpha) y_{t-1} + (1 - \alpha)^2 y_{t-2} + ... + (1 - \alpha)^{t-1} y_1^*$ $\Rightarrow$ exponentiell abnehmende Gewichte (aktuelle Beobachtungen erhalten ein stärkeres Gewicht)
6. $y_t = F(y_t, y_{t-1}, ..., y_2, y_1^*)$ $\Rightarrow$ Anfangswert muß gesetzt werden

Grundmodell

Wahl des Anfangswertes: wird unterschiedlich in Statistiksoftware gehandhabt:

1. SPSS: Durchschnitt aller Werte:

$$y_1^* = \frac{1}{n} \sum_{t=1}^{n} y_t$$

2. EVIEWS: Durchschnitt der ersten 4 Werte:

$$y_1^* = \frac{1}{n} \sum_{t=1}^{4} y_t$$

ein einfaches Beispiel

Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, ..., n\} = 1, 2, 4, 7, 9, 10, 11$

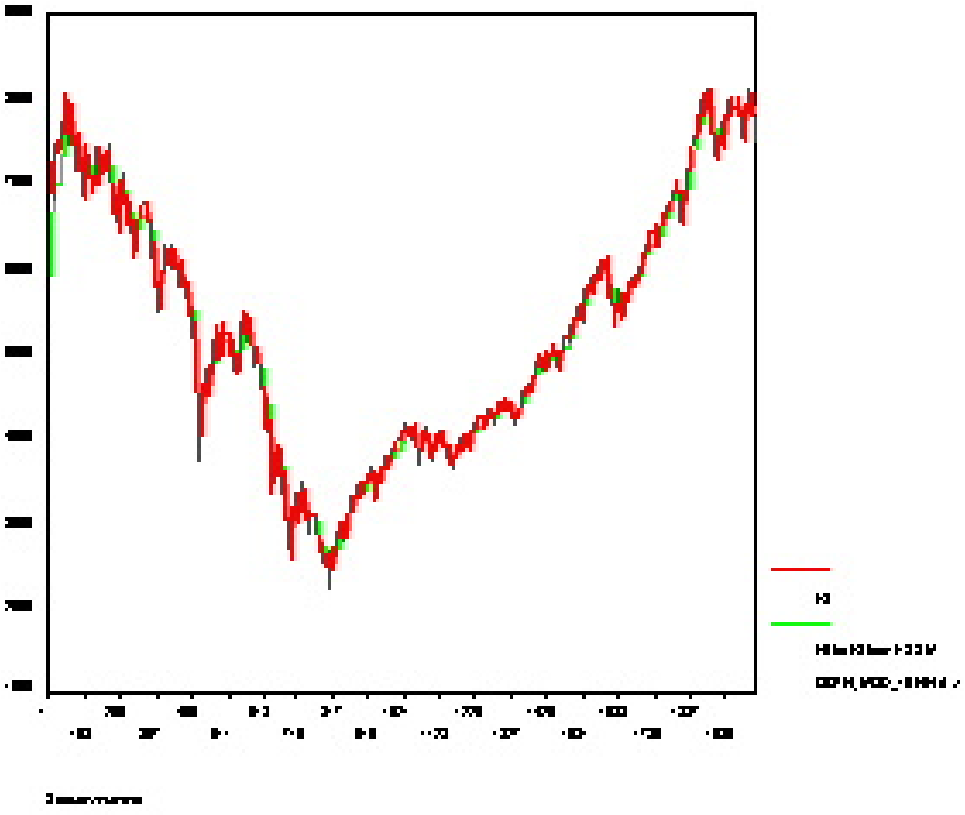
Anfangswert: $y_1^* = \frac{1}{n} \sum_{t=1}^{n} y_t$

Parameter α:

1. $\alpha = 0,01$
2. $\alpha = 0,3$
3. $\alpha = 0,5$
4. $\alpha = 0,9$

ein Beispiel aus der Finanzmathematik

DAX-Entwicklung und Grundmodell der exp. Glättung



Exponetielle Glättung - Gliederung

- 4.2.1 Grundmodell
- **4.2.2 Berücksichtigung des Linearen Trends**
- 4.2.3 Berücksichtigung der Saisonkomponente

Double smoothing - Glättung nach HOLT

Intension: Berücksichtigung des linearen Trends

Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, ..., n\}$. Grundmodell:

$$y_t^* = \alpha y_t + (1 - \alpha) y_{t-1}^*$$

Einführung einer Trendkomponente $T_t \implies$ modifizierter Ansatz:

$$\begin{aligned} y_t^* &= \alpha y_t + (1 - \alpha)(y_{t-1}^* + T_{t-1}) \\ T_t &= \beta(y_t^* - y_{t-1}^*) + (1 - \beta) T_{t-1} \end{aligned}$$

Bemerkungen:

1. Anfangswerte y_1^* und T_1 müssen gegeben sein
2. $y_t^* =$ Prognosewert für y_{t+1}

ein einfaches Beispiel

Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, ..., n\} = 1, 2, 4, 7, 9, 10, 11$

Anfangswert: y_1^*, T_1 aus SPPS

Parameter α:

1. $\alpha = 0,01$
2. $\alpha = 0,3$
3. $\alpha = 0,5$
4. $\alpha = 0,9$

ein Beispiel aus der Finanzmathematik

DAX-Entwicklung und exp. Glättung nacch HOLT

Results of EXSMOOTH procedure for Variable DAX
MODEL= HOLT (Linear trend, no seasonality)

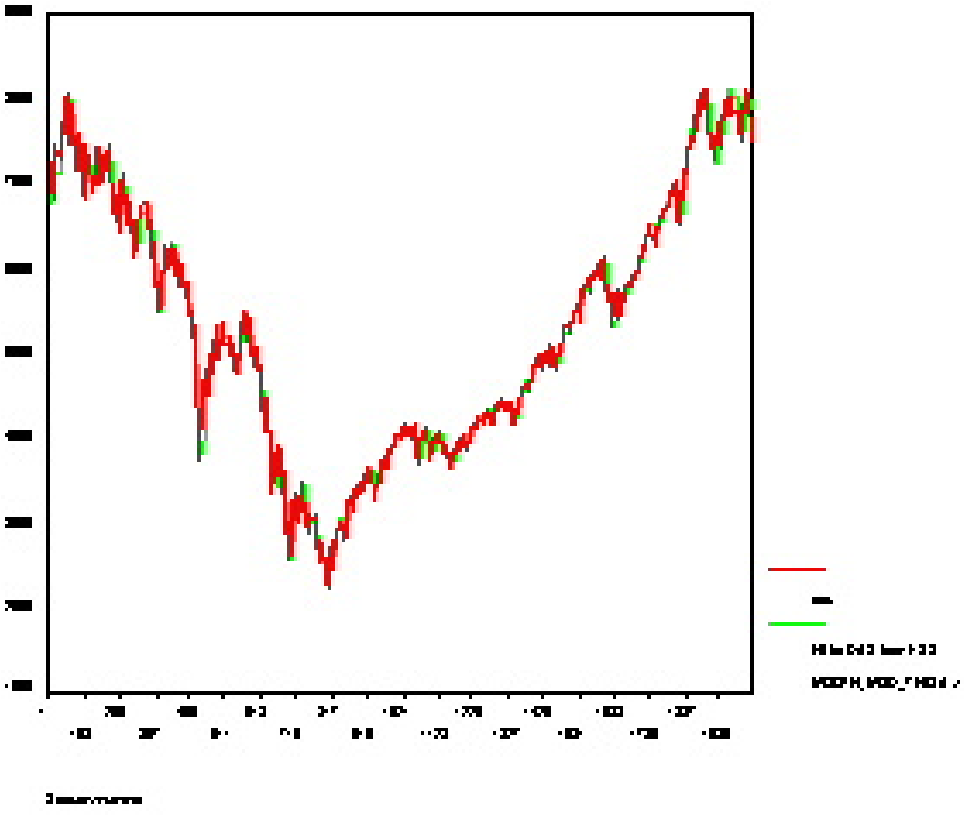
Initial values:	Series	Trend
	6958.01928	.24144

DFE = 2025.

The SSE is:	Alpha	Gamma	SSE
	.1000000	.1000000	51472963.140

ein Beispielaus der Finanzmathematik

DAX-Entwicklung und exp. Glättung nacch HOLT



Exponetielle Glättung - Gliederung

- 4.2.1 Grundmodell
- 4.2.2 Berücksichtigung des Linearen Trends
- **4.2.3 Berücksichtigung der Saisonkomponente**

Double smoothing - Glättung nach HOLT

Intension: Berücksichtigung des liearen Trends und einer Saisonkomponente

Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, \ldots, n\}$ und die Trendkomponente T_t. Glättung nach HOLT:

$$\begin{aligned} y_t^* &= \alpha y_t + (1-\alpha)(y_{t-1}^* + T_{t-1}) \\ T_t &= \beta(y_t^* - y_{t-1}^*) + (1-\beta)T_{t-1} \end{aligned}$$

Einführung einer Trendkomponente $c_t \implies$ modifizierter Ansatz:
Additive Saisonkomponente:

$$\begin{aligned} y_t^* &= \alpha(y_t - c_{t-s}) + (1-\alpha)(y_{t-1}^* + T_{t-1}) \\ T_t &= \beta(y_t^* - y_{t-1}^*) + (1-\beta)T_{t-1} \\ c_t &= \gamma(y_t - y_t^*) + (1-\gamma)c_{t-s} \end{aligned}$$

Bemerkungen:

1. s = 12, wenn Monatsdaten
2. s = 4, wenn Quartalsdaten ...
3. auch multiplikative Verknüpfung möglich $\implies$ modifizierter Ansatz des Gleichungssystems

Double smoothing - Glättung nach HOLT

Multiplikative Saisonkomponente:

$$
\begin{aligned}
y_t^* &= \alpha \frac{y_t}{c_{t-s}} + (1-\alpha)(y_{t-1}^* + T_{t-1}) \\
T_t &= \beta(y_t^* - y_{t-1}^*) + (1-\beta)\, T_{t-1} \\
c_t &= \gamma \frac{y_t}{y_t^*} + (1-\gamma) c_{t-s}
\end{aligned}
$$

Bemerkungen:

1. s = 12, wenn Monatsdaten
2. s = 4, wenn Quartalsdaten ...
3. multiplikative Verknüpfung = Standardeinstellung bei SPSS

- 4.1 Gleitende Durchschnitte
- 4.2 Exponetielle Glättung
- **4.3 Trendfunktionen**
- 4.4 Filter
- 4.5 Autokorrelation
- 4.6 Ausblick: Multiple Zeitreihenanalyse

4.4 Trendfunktionen - Gliederung

- **4.3.1 Definition**
- 4.3.2 Ein Beispiel

Definition

Definition: Es seien X und Y zwei metrisch skalierte Merkmale. Die Ausprägungen dieser zwei Merkmale wurden an n statistischen Einheiten erhoben:

$$(x_1, y_1),\ (x_2, y_2),\ ...,(x_n, y_n)$$

Eine Funktion $y^* = f(x)$ zur Beschreibung der statistischen Abhängigkeit des Merkmals Y vom Merkmal X heißt **einfache Regression** von Y auf X.

Spezialfall:

Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, ..., n\}$. Eine Funktion $y^* = f(t)$ zur Beschreibung der statistischen Abhängigkeit des Merkmals Y von der Zeit t heißt **Trendfunktion**

$$\begin{aligned} y &= f(x) + u = y^* + u \\ y &= f(t) + u = y^* + u \end{aligned}$$

4.4 Trendfunktionen - Gliederung

- 4.3.1 Definition
- **4.3.2 Ein Beispiel**

Beispiel Wechselkurs USD EUR

Excel : Trendfunktion Wechselkurs USD EUR

2 Zeiträume:

1. 15.8 bis 30.11 linearer Trend
2. 2.1 bis 30.4 Polynom 2. Grades

Modellvergleich Beispiel Wechselkurs USD EUR

folgende Modelltypen:

1. linear
2. quadratisch

Auswahlkriterien:

1. Bestimtheitsmaß
2. korrigiertes Bestimmtheitsmaß

$\Rightarrow$ SPSS

- 4.1 Gleitende Durchschnitte
- 4.2 Exponetielle Glättung
- 4.3 Trendfunktionen
- 4.4 **Filter**
- 4.5 Autokorrelation
- 4.6 Ausblick: Multiple Zeitreihenanalyse

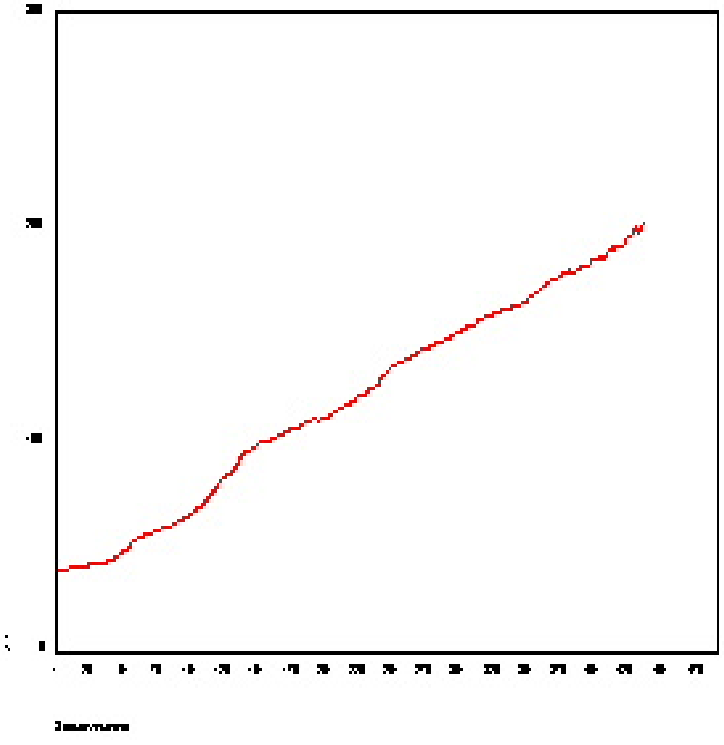
4.4 Filter - Gliederung

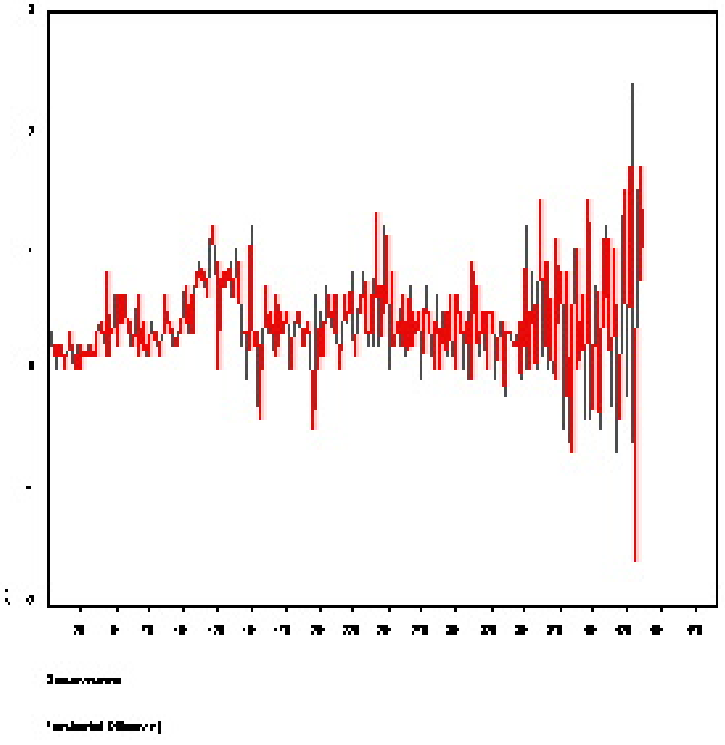
- **4.4.1 Differenzenfilter**
- 4.4.2 Saisonale Differenzenfilter

Differenzenfilter

Intensions: Trendbereinigung

Einführendes Beispiel: CPI USA





Differenzenfilter

Definition: Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, ..., n\}$.

Differenzen - Filter: $\Delta y_t = y_t - y_{t-1}$

Andere gebräuchliche Differenzenfilter:

Differenzen der Logarithmen (LogReturns):

$\Delta \widehat{y_t} = ln(y_t) - ln(y_{t-1}) = ln(\frac{y_t}{y_{t-1}})$

Differenzen 2. Ordnung:

$\Delta \widehat{y_t} = \Delta y_t - \Delta y_{t-1} = y_t - y_{t-1} - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-1})$

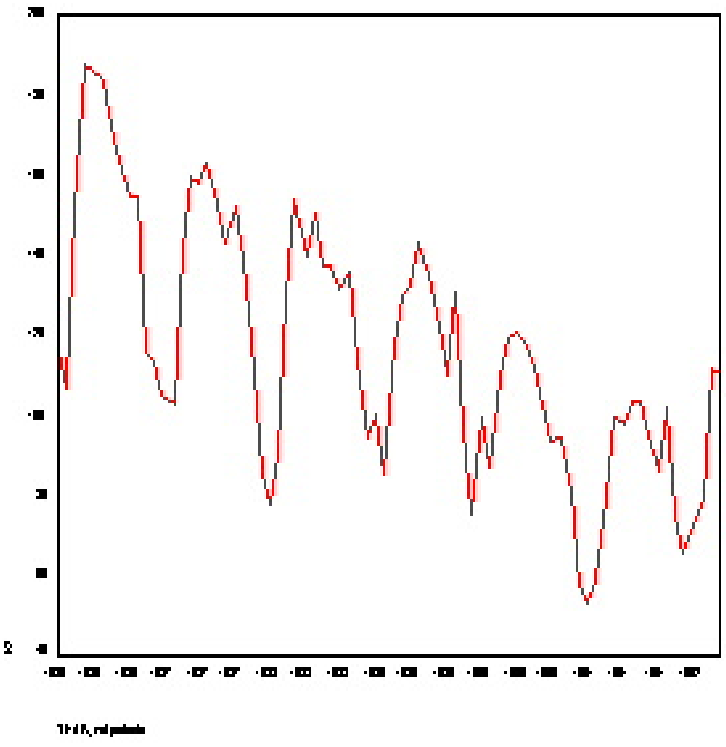
4.4 Filter - Gliederung

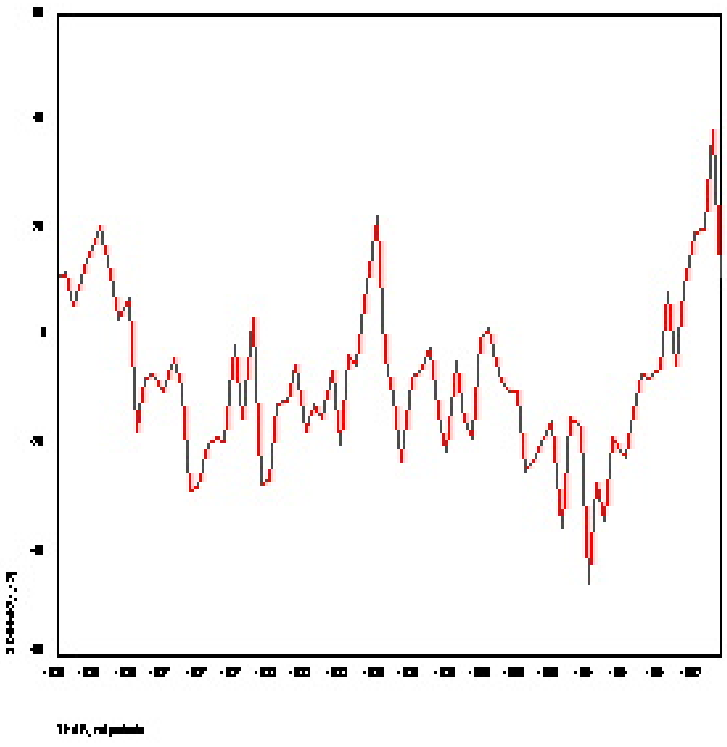
- 4.4.1 Differenzenfilter
- **4.4.2 Saisonale Differenzenfilter**

Saisonale Differenz

Intensions: Saisonalen Effekt herausfiltern

Einführendes Beispiel: Housing Starts in USA 1959-1990





Saisonale Differenz

Definition: Gegeben sei die Zeitreihe $\{y_t,\ t = 1, 2, ..., n\}$.

Identifizierte Periodizität s

Differenzen - Filter: $\Delta y_t = y_t - y_{t-s}$

1. s = 12, wenn Monatsdaten
2. s = 4, wenn Quartalsdaten ...