

18. Saisonbereinigung

18.1 Allgemeines zu periodischen Schwankungen

Eine Zeitreihe $\{S_i\}$ ist **periodisch** mit der Periode τ , falls gilt

$$S_{i+\tau} = S_i$$

Beispiel: Verkehrsaufkommen auf Autobahnen

Es gibt periodische Schwankungen mit mindestens drei Periodendauern:

- Periodendauer 1 Jahr aufgrund Sommer, Winter, Ferienzeit, Feiertagsverkehr,
- Periodendauer 1 Woche aufgrund des Wochenrhythmus: Samstags-, Sonntags- und Werktagsverkehr; letzterer mit Spitzen am Montag und Freitag,
- Periodendauer 1 Tag aufgrund Tag-Nacht-Rhythmus, morgendliche und abendliche Rush-hour etc.

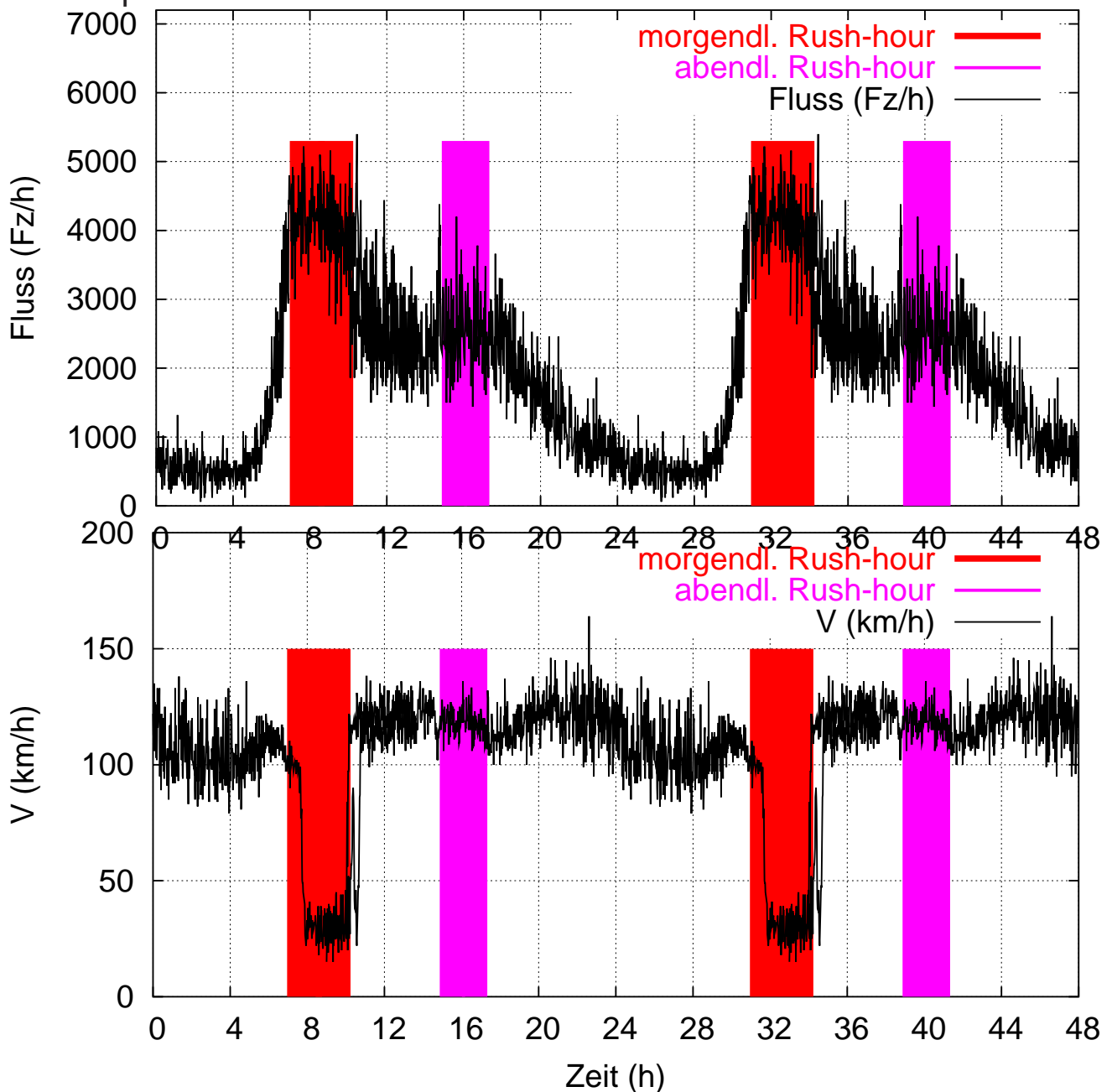
Beispiel 2: Arbeitslosenzahl

Periodendauer 1 Jahr:

- Am meisten Arbeitslose gibt es i.A. im Winter wegen witterungsbedingtem Arbeitsausfall,
- Auch im Sommer gibt es einen kleineren Peak, weil z.B. die neuen Schulabgänger auf den Arbeitsmarkt kommen
- Weitere Saisonalitäten entstehen durch seasonspezifische Arbeitsplätze: Urlaubs-Dienstleistungen, Biergärten, Messezeiten, Erntezeiten etc.

18.1(b) Beispiel 3: Autobahnverkehr

Datenquelle: Induktionsschleifen-Detektor



Gezeigt sind Zeitreihen von Verkehrsfluss (Fahrzeuge/Stunde) und über jeweils eine Minute gemittelte Geschwindigkeit auf der A5 während eines Tages.

Inhaltliche Frage: Die Geschwindigkeit am Morgen bricht im Vergleich zur Rush-hour (Maximum des Verkehrsflusses) verzögert zusammen. Woran könnte dies liegen?

18.1(c) Annahmen

Wir treffen folgende Festlegungen und Annahmen:

1. Im allgemeinen, **nichtstationären** Fall besteht die Zeitreihe, wie üblich, aus Trend-, Konjunktur- und Saisonanteil sowie aus Restschwankungen,

$$y_i = T_i + K_i + S_i + U_i.$$

2. die glatte Komponente (Trend- und Konjunkturanteil) darf nur in Zeiträumen schwanken, die wesentlich größer als die Periode des Saisonanteils sind.
3. Im **stationären Fall** ist der Trendanteil konstant und der Konjunkturanteil verschwindet:

$$T_i = \bar{y}, \quad K_i = 0.$$

4. Der Saisonanteil hat eine feste, *bekannte* Periodendauer τ ; sein arithmetisches Mittel über eine Periode verschwindet.
5. Beliebige Mittelwerte über Restschwankungen verschwinden.

Es ist zweckmäßig, die Periodizität bei der Indizierung der Zeitreihe explizit zu berücksichtigen:

$$y_{ij} := y_{\tau*(i-1)+j}, \quad i = 1, \dots, n, \quad j = 1, \dots, \tau$$

Beispiel jährliche Schwankungen:

i = Jahresindex, j = Monatsindex, Quartalsindex etc, n = Zahl der Jahre, die die Zeitreihe umfasst, τ = 12 Monate, 4 Quartale etc.

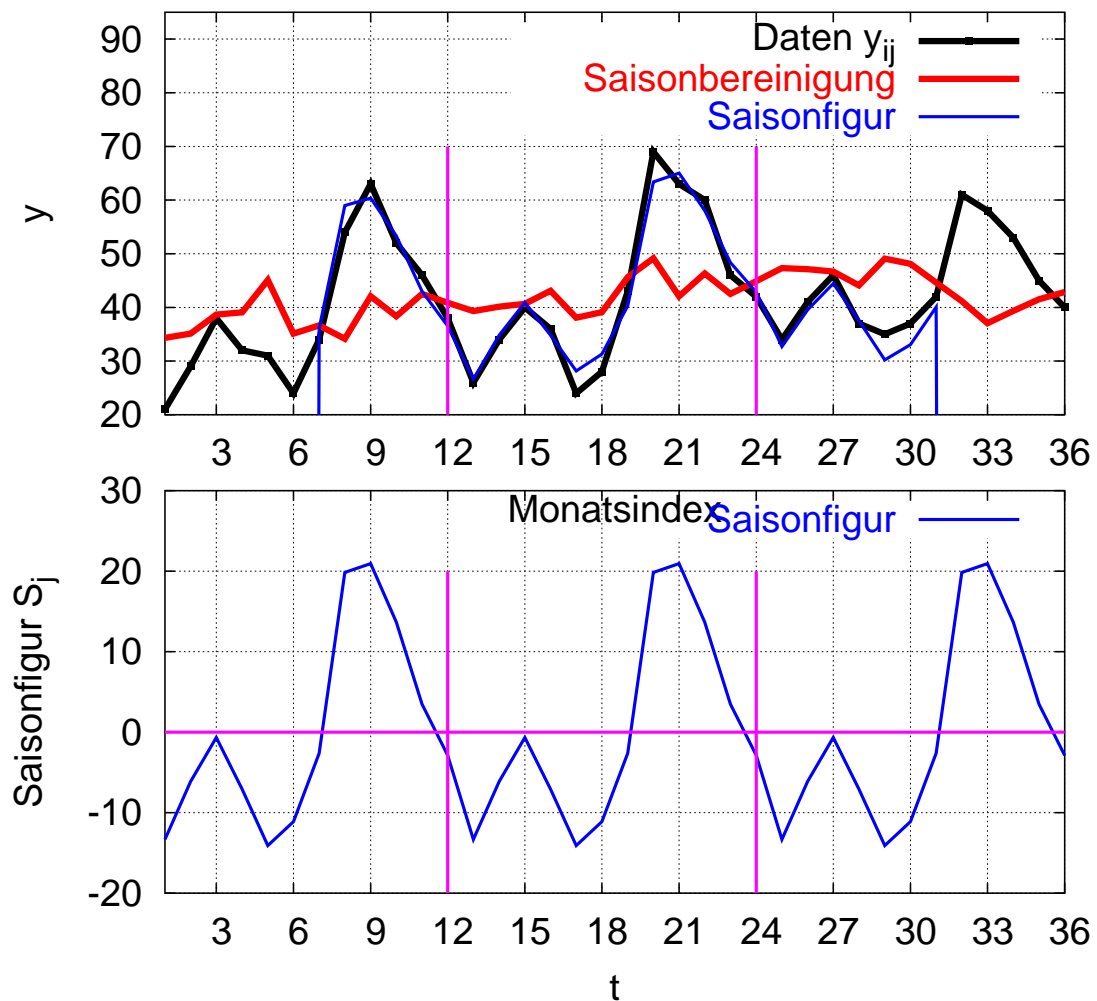
18.2 Algorithmus im stationären Fall

Die **Saisonfigur** S_j (die Saisonkomponente über eine Periode) erhält man dann über das Mittel aus allen Datenpunkten mit festen "Monatsindex" j :

$$S_j = \frac{1}{n} \sum_{i=1}^n (y_{ij} - \bar{y}), \quad \bar{y} = \frac{1}{n\tau} \sum_{i=1}^n \sum_{j=1}^{\tau} y_{ij}.$$

mit $y_{ij} = \bar{y} + S_j + U_{ij}$ (Stationarität!) erhält man damit

- die **saisonbereinigte Zeitreihe** $z_{ij} = y_{ij} - S_j = \bar{y} + U_{ij}$,
- die **irreguläre Komponente** $U_{ij} = y_{ij} - \bar{y} - S_j$,
- sowie den **deterministischen Anteil** $y_{ij} - U_{ij} = \bar{y} + S_j$.



18.2(b) Aufgaben und Beispiele

In den ersten drei Monaten dieses Jahres ist die Geburtenrate um 3.9% im Vergleich zu den ersten drei Monate des Vorjahres gefallen. Es spricht für eine gewisse Ungeduld der Statistiker. Vielleicht sollten sie beim nächsten Mal 9 Monate warten

Kommentar in einem Fernsehmagazin

1. Leiten Sie die Formel für die Saisonfigur her!
2. Verschwindet, wie gefordert, das Mittel der Saisonfigur über eine Periode?
3. Eine Automobilfirma habe in den letzten 3 Jahre folgende Absatzzahlen (in 1000 PKW):

Jahr	1. Quartal	2. Quartal	3. Quartal	4. Quartal
1999	200	260	210	150
2000	220	270	220	140
2001	210	280	200	160

- Berechnen Sie die Saisonfigur S_j , $j = 1, \dots, 4$
- Steigt im 4. Quartal 2001 der saisonbereinigte Absatz (i) gegenüber dem Vorquartal, (ii) gegenüber demselben Quartal des Vorjahres an oder fällt sie?
- Machen Sie eine Prognose für das 2. Quartal 2002!

18.3 Saisonbereinigung nichtstationärer Zeitreihen

Ansätze für die Saisonkomponente

1. **Additives Modell:** Ebenso wie im stationären Fall kann die Zeitreihe als *Summe* der vier Komponenten aufgefasst werden:

$$y_i = T_i + K_i + S_i + U_i$$

- Der Trendanteil ist häufig linear steigend oder fallend
- Der Konjunkturanteil schwankt unregelmäßig mit Periodendauern sehr viel größer als die Saisonkomponente

2. **Multiplikatives Modell:** Die Zeitreihe kann als *Produkt* der vier Komponenten aufgefasst werden:

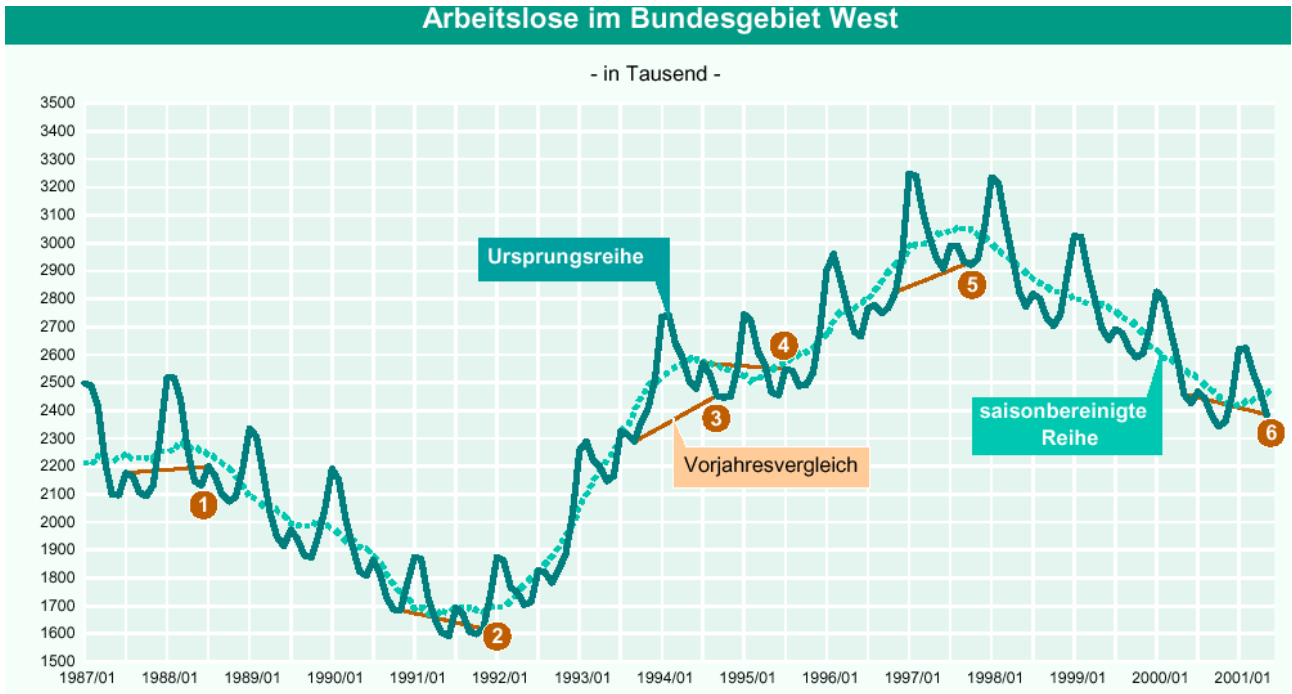
$$y_i = T_i K_i S_i U_i$$

- Die Komponenten K_i , S_i und U_i geben keine absoluten Anteile sondern *relative* Schwankungen an. Die S_i heißen in diesem Zusammenhang auch **Saisonindexziffern**.
- Beispiele: offene Stellen, Aktienkurse, Devisenkurse, Autoumsätze etc. *Die meisten Wirtschaftsdaten gehorchen diesem Modell!*

3. **Gemischt additiv-multiplikatives Modell:** Die Schwankungen steigen mit y_i , aber im Gegensatz zum multiplikativen Modell weniger als proportional.

- *Beispiel:* Arbeitslosenzahl

18.3 (b) Komplexeres Beispiel



Aufgabe:

Extrahieren Sie qualitativ-grafisch aus der dargestellten Arbeitslosen-Zeitreihe die Saisonfigur S und die glatte Komponente $T + K$!

18.4 Additive Saisonkomponente

Das Vorgehen ist prinzipiell wie im stationären Fall, nur dass man die glatte Komponente nicht durch den globalen Mittelwert, sondern durch ein gleitendes Mittel mit Ordnung gleich Periode (*warum?*) ersetzt:

$$\tilde{S}_j = \frac{1}{n} \sum_{i=1}^n (y_{ij} - \bar{y}_{ij}^{(\tau)}),$$

wobei das gleitende Mittel $\bar{y}_{ij} = \bar{y}_{i*\tau+j}$ wie in Kap. 18.1. berechnet wird.

Wegen der i.A. nichtlinearen glatten Komponente $T + K$ ist die Bedingung des verschwindenden Mittelwertes für \tilde{S}_j nicht genau erfüllt. (*warum ist dies bei linearer glatter Komponente der Fall?*) Die **Saisonfigur** erhält man daher erst nach Abzug des Mittelwertes von \tilde{S}_j :

$$S_j = \tilde{S}_j - \frac{1}{\tau} \sum_{j'=1}^{\tau} \tilde{S}_{j'}.$$

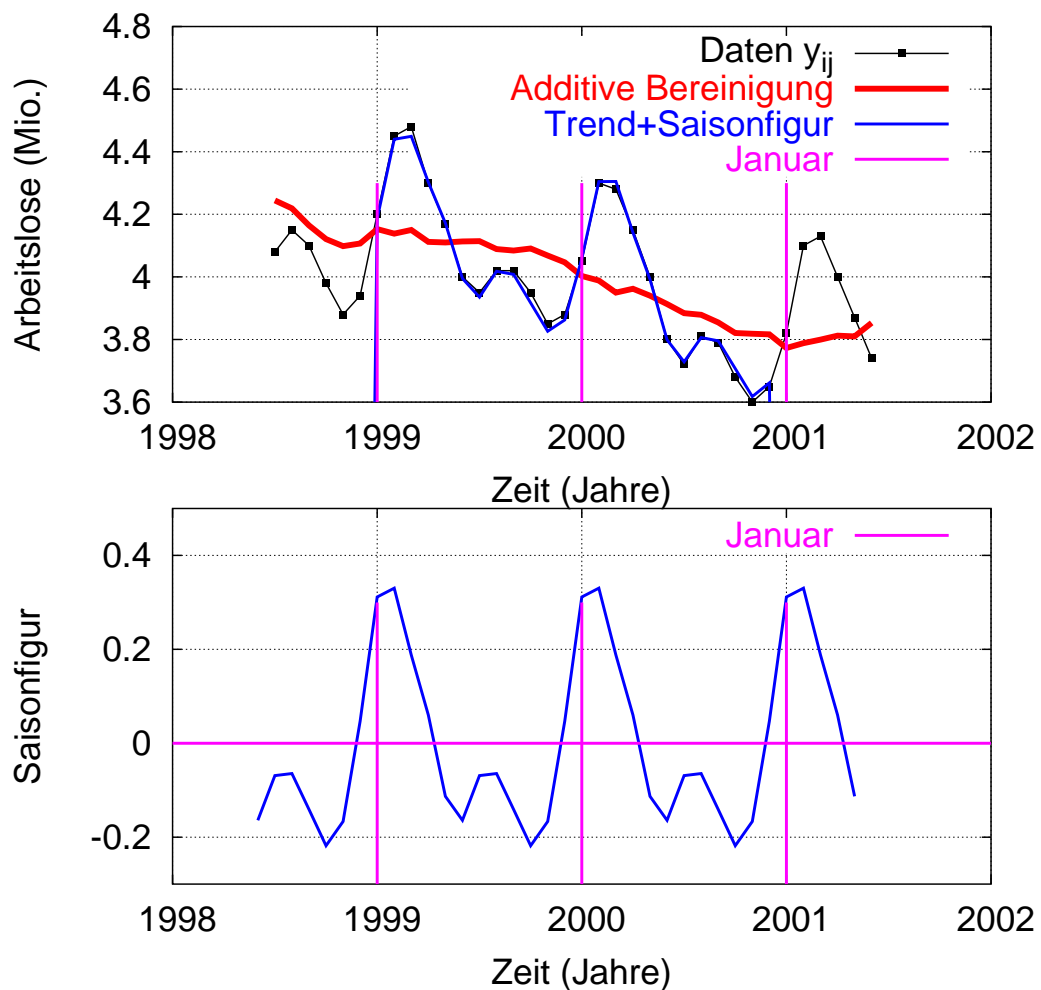
die **saisonbereinigte Zeitreihe**

$$z_{ij} = y_{ij} - S_j$$

ist dann wie im stationären Fall gegeben.

18.4(b) Beispiel: Arbeitslosenzahlen

Datenquelle: www.iab.de/ftp/root/kb1201.pdf



Aufgabe: In den Monaten Januar und Juli 1989 bis 1993 gab es in Westdeutschland folgende Arbeitslosenzahlen in Mio. (die Beschränkung auf Januar und Juli dient der Vereinfachung):

Jahr	1989	1990	1991	1992	1993
Januar	2.4	2.2	1.9	1.9	2.3
Juli	2.0	1.85	1.7	1.85	2.2

Führen Sie die additive Saisonbereinigung durch!

18.5 Multiplikative Saisonkomponente

Analog wie bei der nichtlinearen Regression der Exponentialfunktion wird dies auf den additiven Fall zurückgeführt:

1. Logarithmiere die Zeitreihe:

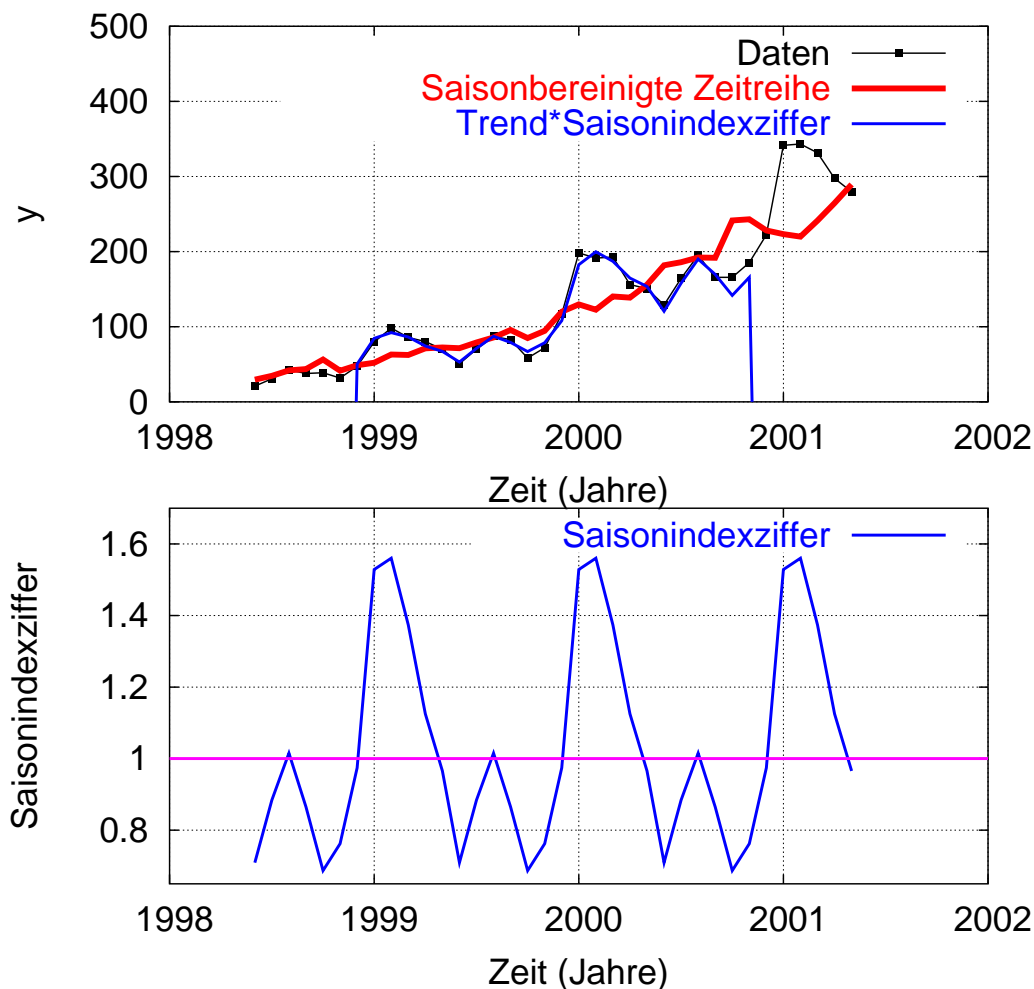
$$\begin{aligned}\hat{y}_i &:= \ln y_i &= \ln(T_i * K_i * S_i * U_i) \\ & &= \ln T_i + \ln K_i + \ln S_i + \ln U_i \\ & &:= \hat{T}_i + \hat{K}_i + \hat{S}_i + \hat{U}_i\end{aligned}$$

2. Bestimme \hat{S}_i mit der additiven Saisonzerlegung
3. Bestimme die **Saisonindexziffern** S_i und die **saisonbereinigte Zeitreihe** z_i durch die Rücktransformation (Exponentenbildung):

$$S_i = e^{\hat{S}_i}, \quad z_i = y_i / S_i.$$

Verständnisfrage: Warum ist bei Wechselkurs-Schwankungen *immer* eine multiplikative Verknüpfung der einzelnen Komponenten sinnvoll?

18.5(b) Beispiel: stark ansteigende Zeitreihe



Aufgaben:

1. Führen Sie für das Beispiel von Kap. 18.2 (Arbeitslose von 1989-1993) eine multiplikative Saisonbereinigung durch! Warum sieht man im Ergebnis kaum einen Unterschied? Wann würden Sie größere Unterschiede erwarten?
2. Berechnen Sie für folgende stark ansteigende Zeitreihe (z.B. Umsatz bei Digitalkameras) die zeitbereinigte Zeitreihe und die Saisonindexziffern. Was sagen die Saisonindexziffern aus?

Jahr	2000	2001	2002	2003
1. Halbjahr	4	9	20	30
2. Halbjahr	3	6	15	18

18.6 Harmonische Analyse

Eine bemerkenswerte mathematische Tatsache ist, dass man *jeden beliebigen* endlichen Datensatz vom Umfang n *exakt* durch eine Summe aus n Sinus- und Kosinus-Schwingungen darstellen kann.

Die Vorfaktoren dieser Schwingungen werden dabei durch die sogenannte **Fouriertransformation** gewonnen. Für die Saisonbereinigung bedeutet dies

- Eine Saisonkomponente mit Periode τ kann als Überlagerung aller Sinus- und Kosinus-Schwingungen angesehen werden, deren Perioden ein *ganzer Bruchteil* von τ (sowie ≥ 2) sind. Für $\tau = 12$ sind dies die Perioden $12, 12/2, 12/3, \dots, 12/6$.
- Als Ergebnis dieser Überlagerung ("**Fourieranalyse**") erhält man folgendes einfache Vorgehen, bei dem sich *alle Details der Fouriertransformation* "herauskürzen":

1. Spalte durch lineare Regression den Trend $(a + bx)$ ab.
2. Berechne den Saisonanteil:

$$S_m = \frac{1}{n} \sum_{i=1}^n (y_{ij} - a - bx_{ij})$$

(Indexkonvention $y_{ij} = y_{\tau*(i-1)+j}$ wie bisher)

3. Die **saisonbereinigte Zeitreihe** ist wie bei der additiven Zerlegung gegeben:

$$z_{ij} = y_{ij} - S_j.$$

(Herleitung für mathematisch Interessierte: Siehe Exkurs zu diesem Abschnitt)

18.6(b) Ergebnis der harmonischen Analyse

