Einheit 11

Explorative Zeitreihenanalyse

Inhalt

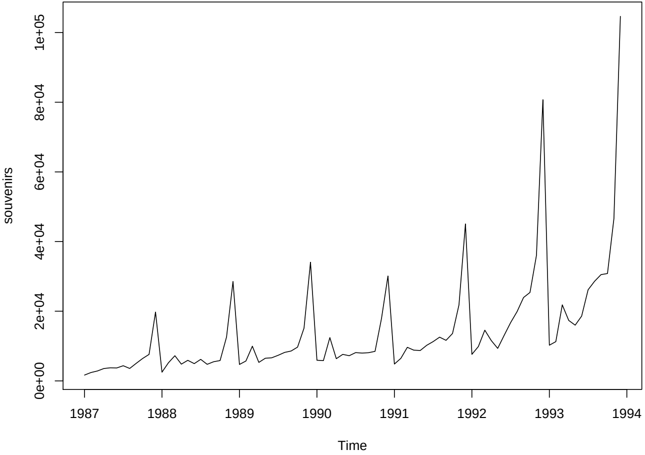
- Beschreiben von Zeitreihen
- Klassische Zeitreihenzerlegung
- Trend, Saison, irreguläre Komponente
- Trendmodelle, Glättung, Trendbereinigung
- Berechnung der Saisonkomponenten, Saisonbereinigung
- Prognose, Maße für die Prognosegüte
- Prognoseverfahren von Holt und von Holt-Winters

Glättung und Prognose von Umsätzen

Im Januar 1987 eröffnet ein Souvenirgeschäft auf dem Kai des Strandbadeorts Maroochydore in Queensland, Australien. Die Verkäufe variieren über die Saison zusammen mit der Anzahl der Touristen. Insbesonders zu Weihnachten und zum örtlichen Surf-Festival, das jeden März seit 1988 veranstaltet wird, kommen die Urlauber an den Strand. Über die Jahre werden Geschäftsräume, Belegschaft und Produktpalette erweitert.
Hier untersuchen wir die monatliche Zeitreihe der Umsätze von Anfang 1988 bis Ende 1993.

Fragen: Wie entwickelt sich der langfristige Trend?
Besitzen die Verkaufszahlen eine regelmäßiges Muster, das sich jedes Jahr wiederholt? Liegt ein saisonales Verhalten vor?
Wie können die systematischen Komponenten zur Prognose der zukünftigen Entwicklung verwendet werden?

Umsätze in Maroochydore, Queensland, AUS



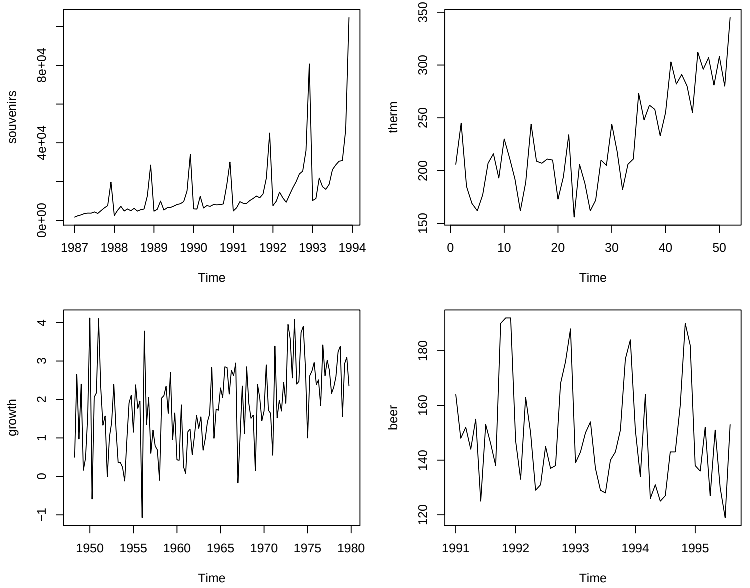
Beispiele für Reihen

Im Gegensatz zu Querschnittsdaten stellen Zeitreihen der Zeit nach geordnete Beobachtungen dar. Viele Daten in der Wirtschaft fallen in dieser Form an.

Die mittel- bis langfristige Entwicklung kann recht unterschiedlich sein:

- Umsatzentwicklung eines Strandshops in Queensland in AUS Dollar, Monatsdaten, Jänner 1988 - Dezember 1993
- Wochenproduktion von Thermostaten in Stück, 52 Wochen
- Einkommenswachstum in Iowa in Prozent, 2. Quartal 1948 – 4. Quartal 1978
- Bierproduktion in Australien in Megaliter, Monatsdaten, Jänner 1991 – August 1995

Zeitreihenplots



Beschreiben von Zeitreihen

- Die Strandumsätze steigen exponentiell mit starken saisonalen Ausschlägen.
- Die Thermostatnachfrage zeigt in der ersten Hälfte der Beobachtungsperiode einen leichten, linearen Rückgang und steigt in der zweiten Hälfte stark (linear) an.
- Die Wachstumsreihe hat ebenfalls keinen stabilen Trend. Die mittelfristige Entwicklung wird von starken unregelmäßigen Schwankungen überlagert.
- Die Bierproduktion verläuft leicht abfallend. Die saisonalen Schwankungen (großer Durst im australischen Sommer, wenig Durst im Winter) sind stark ausgeprägt.

Die regelmäßige Wiederholung eines ähnlichen Verlaufs mit fixer Periode nennen wir **saisonales Muster**. Bei Monatsdaten hat die Periode die Länge $s = 12$ Monate, bei Quartalsdaten $s = 4$ Quartale.

[Explorative Zeitreihenanalyse]

Die klassische Zeitreihenzerlegung

Zerlegung einer Reihe in 3 Komponenten

Wir setzen eine Reihe aus 3 Komponenten zusammen:

- $T_t \ldots$ Trend und/oder Zyklus
- $S_t \ldots$ Saison
- $E_t \ldots$ irreguläre Komponente, Fehler

Das Modell kann **additiv** oder **multiplikativ** sein:

$$Y_t = T_t + S_t + E_t \qquad \text{oder} \qquad Y_t = T_t \cdot S_t \cdot E_t$$

- **Additiv**: Der saisonale Effekt ist keine Funktion des Niveaus der Reihe. Vgl. `beer` Reihe.
- **Multiplikativ**: Der saisonale Effekt wird durch T_t verstärkt. Vgl. `souvenirs` Reihe.

Durch Logarithmierung können wir das multiplikative Modell in das additive überführen: $\log(Y_t) = \log(T_t) + \log(S_t) + \log(E_t)$.

Berechnung der Komponenten (additives Modell)

Wir ermitteln die systematischen Komponenten T_t und S_t in 3 Schritten:

1) Berechnung der Trendkomponente, T_t:
Z.B. linearer, exponentieller Trend, mittelfristiger Trend durch Glättung

2) Berechnung der trendbereinigten (detrended) Reihe $TB_t = (Y_t - T_t)$.

3) Ermittlung der Saisonkomponente, S_t, aus $(Y_t - T_t)$.

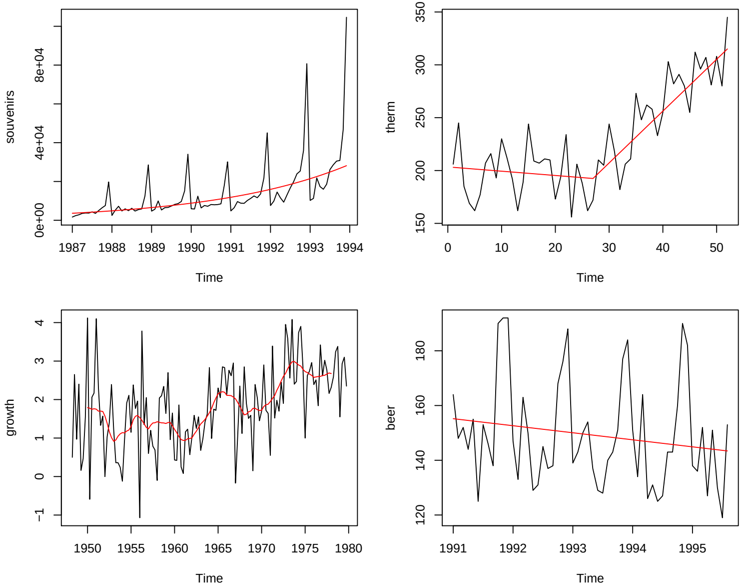
Als Rest ergibt sich die irreguläre Komponente, E_t.

$$E_t = Y_t - (T_t + S_t)$$

[Explorative Zeitreihenanalyse]

Trend

Der Trend der Reihen



Linearer und exponentieller Trend

- Ein **linearer Ternd** mit dem Interzept a, der Steigung b hat die Form

$$Y_t = a + b \cdot t + u_t, \qquad t = 1, 2, \ldots, n$$

  t ist die Trendvariable, $t = 1, 2, 3, \ldots$, und u_t die irreguläre Komponente. Das Modell kann mit der Methode der kleinsten Quadrate, KQ, geschätzt werden.

- **Exponentielles Wachstum** mit der Wachstumsrate β

$$Y_t = \alpha \cdot e^{\beta \cdot t} \cdot v_t$$

  kann durch Logarithmieren in eine linearen Funktion übergeführt werden und mit KQ geschätzt werden. $a = \log(\alpha)$.

$$\log(Y_t) = a + \beta \cdot t + u_t$$

Der Trend in der Reihe `souvenirs`

Wir logarithmieren `souvenirs` und schätzen einen linearen Trend.

Abhängige Variable: log(souvenirs) n=84

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	8.1702	0.11322	72.2	5.69e-76
t	0.0247	0.00231	10.7	3.52e-17

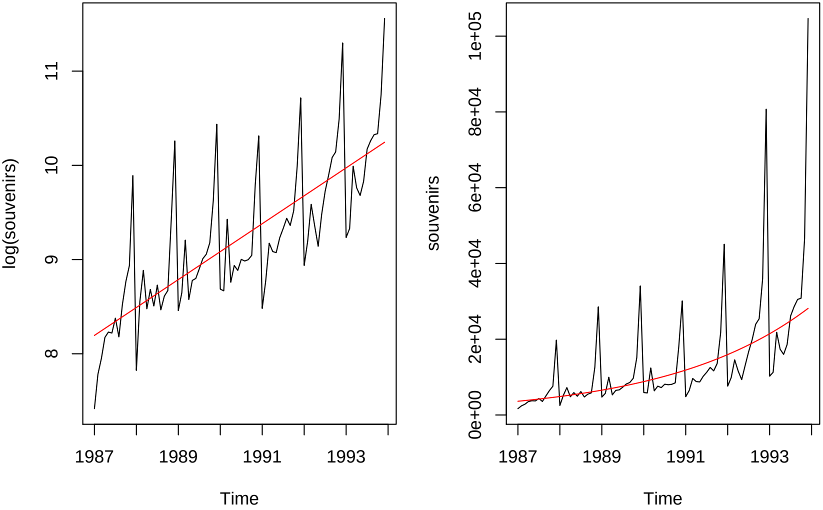
R^2=0.582

Das Modell für `souvenirs` lautet

$$\texttt{souvenirs}_t = 3534.1 \cdot e^{0.025 \cdot \texttt{t}} \cdot v_t$$

Der Faktor α ergibt aus $\exp(8.1702) = 3534.1$.

Der Trend in der Reihe `souvenirs`



Gleitender Durchschnitt, k MA

Ist der Trend nicht stabil, können **gleitende Durchschnitte** (moving averages, MA) verwendet werden, um die mittelfristige Entwicklung sichtbar zu machen.
Wir wählen $k = 2 \cdot m + 1$ (ungerade) benachbarte Beobachtungen, berechnen daraus das arithmetische Mittel und ordnen diesen Wert dem mittleren Beobachtunszeitpunkt zu. Dies wiederholen wir für alle Zeitpunkte, für die das möglich ist.

$$T_t = \frac{1}{2m+1}(Y_{t-m} + \ldots + Y_t + \ldots + Y_{t+m})$$

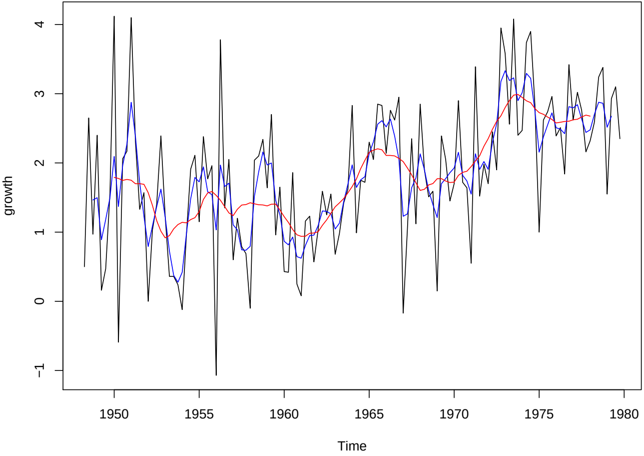
$t = m+1, m+2, \ldots, n-m$.
Wir nennen T_t den $(2m+1)$-gliedrigen Durchschnitt.

Die Stärke der Glättung hängt von der Wahl von k ab:

- Für große k (bzw. m) wird Y_t stark,
- für kleine k (bzw. m) wird wenig geglättet.

Moving average Glättung von `growth`

$\texttt{BLAU}$: $k = 5 = 2 \cdot 2 + 1 \quad (m = 2)$
$\texttt{ROT}$: $k = 15 = 2 \cdot 7 + 1 \quad (m = 7)$



Gewichteter moving average

Saisonale Daten hingegen weisen oft eine gerade Periodenlänge auf: 4 für Quartalsdaten, 12 für Monatsdaten. Wir verwenden hier statt eines k MA einen geeigneten **gewichteten moving average**.

- Für Quartalsdaten, $s = 4$:

$$T_t = (0.5 \cdot Y_{t-2} + Y_{t-1} + Y_t + Y_{t+1} + 0.5 \cdot Y_{t+2})/4$$

  Die 5 Gewichte $\{1/8, 1/4, 1/4, 1/4, 1/8\}$ summieren sich zu 1.

- Für Monatsdaten, $s = 12$:

$$T_t = (0.5 \cdot Y_{t-6} + Y_{t-5} + \ldots + Y_t + \ldots + Y_{t+5} + 0.5 \cdot Y_{t+6})/12$$

  Die Summe der 13 Gewichte $\{1/24, 1/12, \ldots, 1/12, 1/24\}$ ist 1.
  Z.B. für den Juli-Wert gehen die Jännermonate zweier aufeinderfolgender Jahre mit je 50% in die Berechnung ein.

Rechenbeispiele

Gegeben ist die Reihe Y_t der Länge $n = 8$.

t	1	2	3	4	5	6	7	8
Y_t	1	2	1	3	4	0	1	2

- Der 3-gliedrige gleitende Durchschnitt für $t = 2$ ist

$$(1 + 2 + 1)/3 = 4/3$$

- Der 5-gliedrige gleitende Durchschnitt für $t = 5$ ist

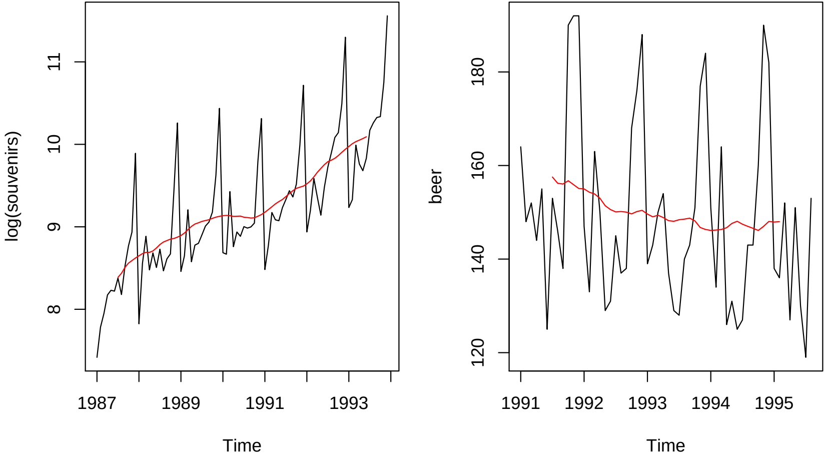
$$(1 + 3 + 4 + 0 + 1)/5 = 9/5$$

- Nehmen wir an, es liegen Quartalsdaten vor, $s = 4$. Der gewichtete moving average für $t = 5$ ist

$$(0.5 \cdot 1 + 3 + 4 + 0 + 0.5 \cdot 1)/4 = 8/4$$

Glättung saisonaler Reihen

Wir glätten $\log(\texttt{souvenirs})$ und `beer` mittels des saisonalen gewichteten Durchschnitts, $s = 12$. (Vgl. Folie 15)



Einfache exponentielle Glättung für Reihen ohne Trend und ohne Saison

Die geglättete Reihe heißt nun L_t, wie level (Niveau). Sie wird rekursiv als gewichtetes Mittel aus der aktuellen Beobachtung und der Glättung der Vorperiode berechnet.

$$L_t = \alpha \cdot Y_t + (1 - \alpha) \cdot L_{t-1}, \quad 0 < \alpha \leq 1$$

- Ist $\alpha = 1.00$, ist die geglättete Reihe gleich der ursprünglichen. Es wird nicht geglättet.
- Ist α groß, z.B. 0.75, so ist der aktuelle Wert sehr wichtig, der Einfluss des geglätteten Wertes der Vorperiode (in dem auch die gesamte Vergangenheit der Reihe eingeht) gering. Es wird wenig geglättet.
- Ist α klein, z.B. 0.10, so kann der aktuelle Y-Wert die Glättung nur wenig beeinflussen. Es wird stark geglättet.

Rechenbeispiel zur einfachen exponentiellen Glättung

Gegeben ist die Reihe Y_t der Länge $n = 8$ und $\alpha = 0.6$.

$$L_t = 0.6 \cdot Y_t + 0.4 \cdot L_{t-1}$$

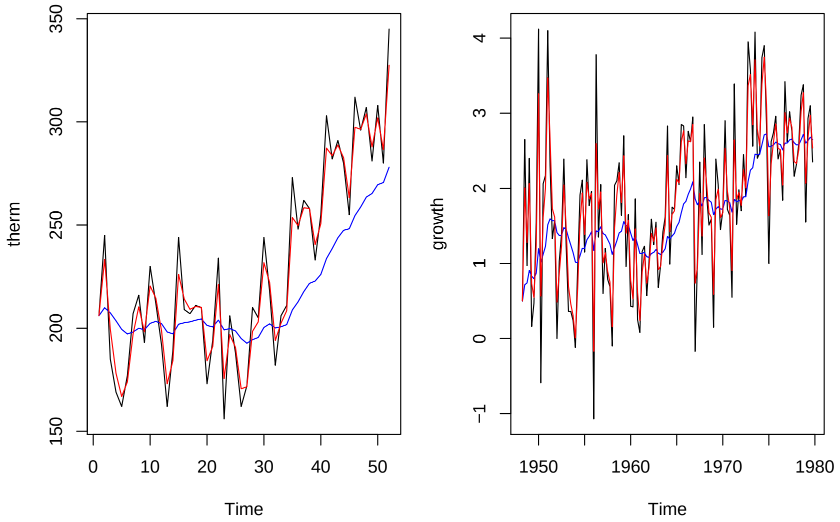
Der Anfangswert für L_1 sei $L_1 = Y_1 = 1$.

t	1	2	3	4	5	6	7	8
Y_t	1	2	1	3	4	0	1	2
L_t	*1*	1.6	1.24	2.296	3.318	1.327	1.131	1.652

$$L_2 = 0.6 \cdot Y_2 + 0.4 \cdot L_1 = 0.6 \cdot 2 + 0.4 \cdot 1 = 1.6$$
$$L_3 = 0.6 \cdot Y_3 + 0.4 \cdot L_2 = 0.6 \cdot 1 + 0.4 \cdot 1.6 = 1.24$$
$$\ldots$$
$$L_8 = 0.6 \cdot Y_8 + 0.4 \cdot L_7 = 0.6 \cdot 2 + 0.4 \cdot 1.131 = 1.652$$

Exponentielle Glättung von `therm` und `growth`

`BLAU:` $\alpha = 0.1$ `ROT:` $\alpha = 0.7$



Exponentielle Glättung von `therm` und `growth`

Zu `therm`:
Die Glättung mit $\alpha = 0.1$ liegt im ansteigenden Bereich systematisch unterhalb der Beobachtungen. Daher ist sie nicht geeignet.
Die andere Variante glättet kaum, verfolgt aber gut den Pfad.

Zu `growth`:
Die Glättung mit $\alpha = 0.1$ erfasst weitgehend die wesentlichen Bewegungen. Die Variante $\alpha = 0.7$ glättet kaum, da sie jeden Ausschlag fast zu Gänze nachbildet.

Trendbereinigung

Die trendbereinigte Reihe, TB_t, ergibt sich aus Y_t nach Abzug des Trends.

$$TB_t = Y_t - T_t$$

Rechenbeispiel:
Gegeben ist Y_t und die bereits berechnete Trendkomponente T_t.

t	1	2	3	4	5	6	7	8
Y_t	2	3	-1	3	6	5	4	9
T_t	1	2	3	4	5	6	7	8
TB_t	1	1	-4	-1	1	-1	-3	1

[Explorative Zeitreihenanalyse]

Saison

Saisonkomponenten

Wir berechnen die **Saisonkomponente** sehr einfach. Die Zeit t wird für Monatsdaten als $t = j.m$ notiert. j bezeichnet das Jahr, m den betreffenden Monat. Der Effekt des Monats m (fix) ergibt sich als der durchschnittlicher Effekt des Monats m in der trendbereinigten Reihe, TB_t, über alle Jahre.

$$S_m = S_{j.m} = (TB_{1.m} + TB_{2.m} + \ldots + TB_{J.m})/J \qquad j = 1, \ldots, J$$

J bezeichnet die Anzahl der Jahre für die eine Beobachtung des Monats m zur Verfügung steht.

Rechenbeispiel

Gegeben ist die trendbereinigte Quartalsreihe TB_t der Länge $n = 12$.
$s = 4$, $J = 3$, $q = 1, 2, 3, 4$.

t	1	2	3	4	:	5	6	7	8	:	9	10	11	12
j	1	1	1	1	:	2	2	2	2	:	3	3	3	3
q	1	2	3	4	:	1	2	3	4	:	1	2	3	4
TB_t	5	1	-3	1	:	4	0	-2	2	:	3	-1	-1	3
S_t	4	0	-2	2	:	4	0	-2	2	:	4	0	-2	2

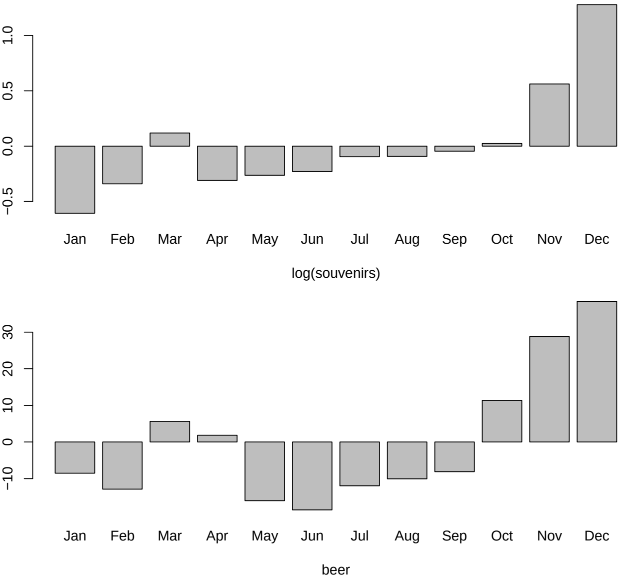
Die Quartalseffekte sind

$$S_1 = (5 + 4 + 3)/3 = 4 \qquad S_2 = (1 + 0 - 1)/3 = 0$$
$$S_3 = (-3 - 2 - 1)/3 = -2 \qquad S_4 = (1 + 2 + 3)/3 = 2$$

Die Saisonkomponente ist

$$S_t = S_{j.q} = S_q$$

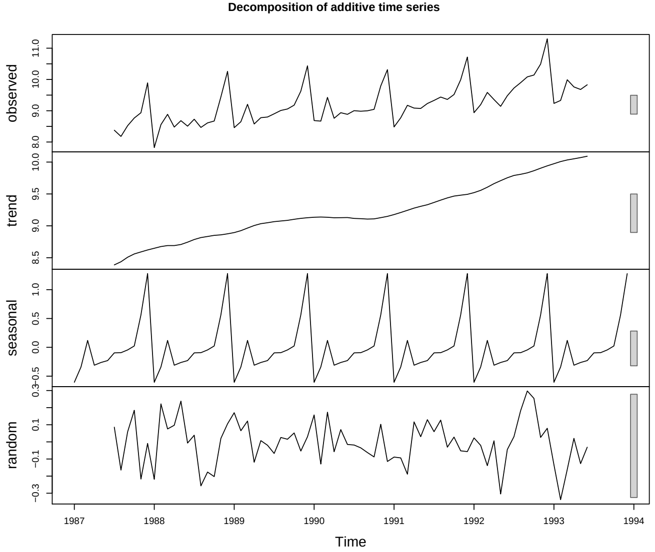
Saisoneffekte von `log(souvenirs)` und `beer`



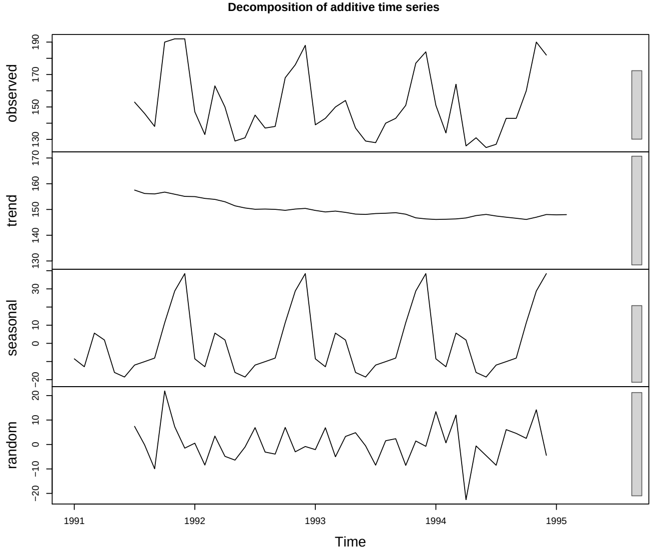
[Explorative Zeitreihenanalyse]

Zusammenfassung der Komponentenzerlegung

Komponentenzerlegung von `log(souvenirs)`



Komponentenzerlegung von `beer`



[Explorative Zeitreihenanalyse]

Saisonbereinigung

Saisonbereinigung

Ausgeprägte saisonale Schwankungen erschweren die mittel- und längerfristige Entwicklung zu beurteilen. Daher werden oft nur saisonbereinigte Reihen publiziert. Vgl. BIP.

Die **saisonbereinigte Reihe** (seasonally adjusted, sa) zu Y erhalten wir, indem wir die Saisonkomponente von der beobachteten Reihe abziehen.

$$(Y_t - S_t) = T_t + E_t$$

Wir unternehmen dazu folgende Schritte:

1) Berechnung der Trendkomponente, T_t:
Z.B. linearer, exponentieller Trend, kein Trend oder Glättung

2) Berechnung der trendbereinigten (detrended) Reihe $TB_t = (Y_t - T_t)$.

3) Ermittlung der Saisonkomponente, S_t, aus $(Y_t - T_t)$.

4) Bereinigen der Reihe Y_t um die Saison: $Y_t^{sa} = (Y_t - S_t)$.

Rechenbeispiel

Im obigen Rechenbeispiel zu den Saisonkomponenten auf Folie 27 ist nun auch Y_t gegeben. (Hier: $Y_t = TB_t + t$.) Die saisonalen Effekte haben wir bereits berechnet.

t	1	2	3	4	:	5	6	7	8	:	9	10	11	12
Y_t	6	3	0	5	:	9	6	5	10	:	12	9	10	15
S_t	4	0	-2	2	:	4	0	-2	2	:	4	0	-2	2
Y_t^{sa}	2	3	2	3	:	5	6	7	8	:	8	9	12	13

Saisonbereinigungsverfahren, die in statistischen Zentralämter Verwendung finden, sind u.A.:

- X-12-ARIMA Verfahren des U. S. Census Bureau, oder
- TRAMO SEATS bei Eurostat.

[Explorative Zeitreihenanalyse]

Prognose

Beobachtungen und Vorhersagen

Beobachtungsperiode						:	Prognoseperiode		
Y_1	Y_2	Y_3	$\dots$	Y_{n-1}	Y_n	:	Y_{n+1}	$\dots\dots$	Y_{n+r}
historische Daten						:	zukünftige Beobachtungen		

Wir kennen die Daten für die Periode $t = 1, \dots, n$ und möchten Vorhersagen für die zukünftigen Zeitpunkte $t = n + 1, \dots, n + r$ berechnen.

Bezeichnungen:

$Y_n(r)$	bezeichnet die Vorhersage für die zukünftige Beobachtung Y_{n+r}, r-Schritt Prognose
n	gibt den Vorhersagezeitpunkt an,
r	den Vorhersagehorizont.

Beobachtungen und Vorhersagen

$Y_n(r)$: Ist r klein sprechen wir von einer kurzfristigen, sonst von einer mittel- oder langfristigen Prognose.

$Y_{2008/9}(3)$ Vorhersage für 2008/12 (Dez 2008) auf Basis der Beobachtungen bis einschließlich 2008/9, eine 3-Schritt Prognose.

Vorhersagen für Okt 2008:

$Y_{2008/6}(4)$ Vorhersage für 2008/10 (Okt 2008) auf Basis der Beobachtungen bis einschließlich 2008/6, eine 4-Schritt Prognose.

$Y_{2008/9}(1)$ Vorhersage für 2008/10 (Okt 2008) auf Basis der Beobachtungen bis einschließlich 2008/9, eine 1-Schritt Prognose.

Prognosefehler

Die Differenz zwischen tatsächlicher, zukünftiger Beobachtung und der Vorhersage bezeichnen wir als Prognosefehler.

$$r\text{-Schritt Prognosefehler} = Y_{n+r} - Y_n(r)$$

Zum Bewerten der Prognosen $Y_n(j), j = 1, \ldots, r$ einer Reihe eignet sich der **mean square error**, MSE,

$$MSE = \frac{1}{r} \sum_{j=1}^{r} [Y_{n+j} - Y_n(j)]^2$$

oder der **root mean square error**, RMSE,

$$RMSE = \sqrt{MSE}$$

Rechenbeispiel zum MSE und RMSE

Angenommen wir verwenden 2 verschiedene Verfahren, A und B, um eine Reihe Y_t, $t = 1, \ldots, n$, vorherzusagen und notieren uns die Vorhersagewerte. Nach 3 Perioden (und 3 weiteren Beobachtungen) wollen wir feststellen, welche Methode die besseren Vorhersagen geliefert hat. Wir verwenden dazu den RMSE. $r = 3$.

	t	1	$\ldots$	$n-1$	n	:	$n+1$	$n+2$	$n+3$
	Y_t	5	$\ldots$	-1	4	:	*3*	*2*	*3*
	j					:	1	2	3
Verf. A	$Y_n(j)$					:	3	4	2
Verf. B	$Y_n(j)$					:	2	3	3

$$RMSE_A = \sqrt{(3-3)^2 + (2-4)^2 + (3-2)^2} = 2.236$$

$$RMSE_B = \sqrt{(3-2)^2 + (2-3)^2 + (3-3)^2} = 1.414$$

Verfahren B hat den kleineren RMSE und ist daher vorzuziehen.

Prognose bei stabilem Trend ohne Saison

Liegt eine Reihe mit einem stabilen Trend vor, können wir die Trend-Regression für mittel- bis langfristige Prognosen verwenden. Für $\texttt{log(souv}_t\texttt{)}$ haben wir die Gleichung (ohne saisonale Effekte zu berücksichtigen) bereits geschätzt.

$$\texttt{log(souv}_t\texttt{)} = 8.1702 + 0.0247 \cdot \texttt{t} + \hat{u}_t, \quad t = 1, \ldots, 84$$

Für die Prognose ersetzen wir u_t durch $E(u_t) = 0$. Damit erhalten wir die Prognosefunktion

$$\widehat{\texttt{log(souv}_t\texttt{)}} = 8.1702 + 0.0247 \cdot \texttt{t}$$

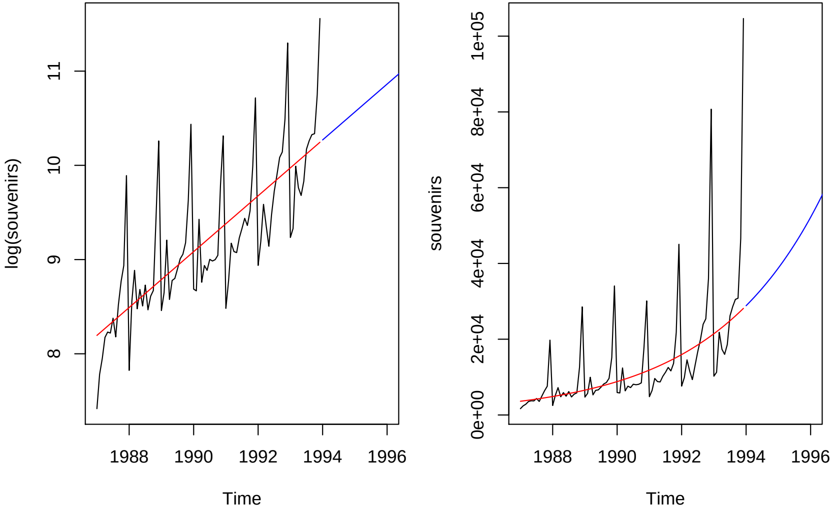
in die wir für $t = 85, 86, \ldots, 120$ einsetzen.

Prognose bei stabilem Trend ohne Saison

Die Vorhersagen für die Perioden $85, \ldots, 120$ sind:

t	j	Periode		$\texttt{log(souv}_{84}\texttt{)}(j)$	$\texttt{souv}_{84}(j)$
85	1	Jan 94	$8.1702 + 0.0247 \cdot (85)$	10.2697	28845.2
86	2	Feb 94	$8.1702 + 0.0247 \cdot (86)$	10.2944	29566.6
		$\ldots$	$\ldots$	$\ldots$	$\ldots$
97	13	Jan 95	$8.1702 + 0.0247 \cdot (97)$	10.6551	38797.9
		$\ldots$	$\ldots$	$\ldots$	$\ldots$
120	38	Dez 96	$8.1702 + 0.0247 \cdot (120)$	11.1342	68473.4

Prognose mit stabilem Trend: `souvenirs`



Prognose für Trend ohne Saison: Das Modell von Holt

Das Verfahren von Holt verwendet exponentielle Glättungen für das Niveau der Reihe L_t und für die 1-Perioden Veränderung im Niveau, b_t. Dazu gibt es 2 Glättungsparameter: $0 < \alpha, \beta \leq 1$.

Das Modell lautet:

$$\begin{array}{lrcl} \text{Niveau:} & L_t & = & \alpha \cdot Y_t + (1-\alpha) \cdot (L_{t-1} + b_{t-1}) \\ \text{Trend:} & b_t & = & \beta \cdot (L_t - L_{t-1}) + (1-\beta) \cdot b_{t-1} \\ \text{Prognose:} & Y_t(r) & = & L_t + b_t \cdot r \end{array}$$

mit den Anfangswerten: $L_2 = Y_2$ und $b_2 = L_2 - L_1 = Y_2 - Y_1$.

Im Gegensatz zur einfachen Glättung verwendet man hier die geglätteten Werte, L_t und b_t, in der Prognose für die folgende Periode.

Dieses Verfahren modelliert additives Wachstum. Liegt ein exponentielles Wachstum vor, ersetzt man Y_t durch die logarithmierte Reihe.

Holt: Die Gleichungen

- Die Prognosegleichung für $r = 1$:

  $$Y_t(1) = L_t + b_t \cdot 1$$

  Für die 1-Schritt Prognose wird zum prognostizierten Niveau die prognostizierte Veränderung addiert.

- Die Trendgleichung stellt eine exponentielle Glättung der Veränderung der Niveaus $(L_t - L_{t-1})$ dar.

  $$b_t = \beta \cdot (L_t - L_{t-1}) + (1 - \beta) \cdot b_{t-1}$$

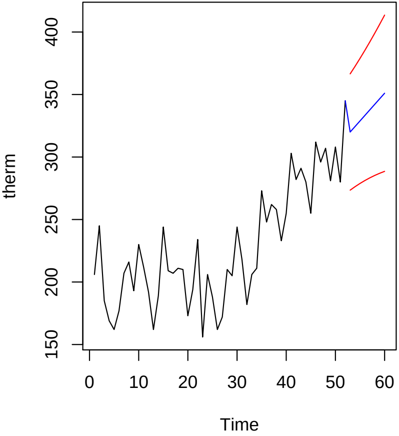
- Die Niveaugleichung glättet Y_t, wobei $(L_{t-1} + b_{t-1})$ als Glättung für das Niveau in $(t - 1)$ verwendet wird.

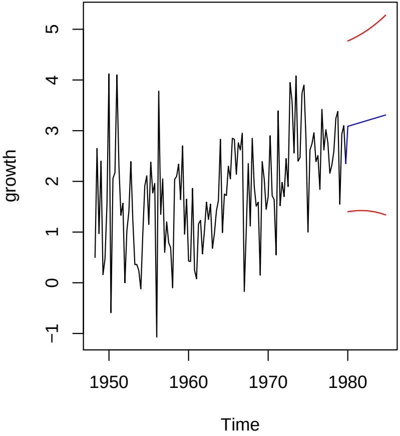
  $$L_t = \alpha \cdot Y_t + (1 - \alpha) \cdot (L_{t-1} + b_{t-1})$$

Die optimalen Werte für α und β ermittelt man durch Minimierung der Summe der quadrierten 1-Schritt Prognosefehler, $\sum_t [Y_t - Y_{t-1}(1)]^2$.

Prognose mit Holt: `therm` und `growth`

α=0.245, β=0.093 α=0.078, β=0.074





BLAU . . . Punktprognosen ROT . . . 95% Prognoseintervalle

Prognose mit Holt: Kommentar

Der Graph der Reihen wird mit den Punktprognosen $Y_n(1), Y_n(2), \ldots,$ $Y_n(8)$ bzw. $Y_n(20)$ fortgesetzt.
Die roten Linien geben die 95% Prognoseintervall-Grenzen an. Sie zeigen, dass bei beiden Reihen eine beträchtliche Unsicherheit über den zukünftigen Verlauf besteht.

- `therm`: Das Niveau wird zu 24.5% (α) durch den aktuellen Wert bestimmt, der Trend nur zu 9.3% (β). Die Graphik auf Folie 45 zeigt, dass die Prognose im Wesentlichen den Trend der zweiten Hälfte der Beobachtungsperiode berücksichtigt.
- `growth`: Hier spielt das Niveau vergangener Beobachtungen eine große Rolle ($(1-\alpha) = 0.922$). Die Gewichtung der aktuellen lokalen Veränderungen (Trend) ist ähnlich gering wie bei der Reihe `therm` ($\beta = 0.074$).

Prognose für Trend und Saison: Das Modell von Holt-Winters

Das Modell von Holt-Winters hat 3 Glättungsparameter: α, β und γ.
γ ist der Glättungsparameter für die Saisonkomponente.
s gibt die Periodenlänge der Saison an.
Z.B. $s = 4$ bei Quartalsdaten, oder $s = 12$ bei Monatsdaten.

$$
\begin{array}{lrcl}
\text{Niveau:} & L_t & = & \alpha \cdot (Y_t - S_{t-s}) + (1-\alpha) \cdot (L_{t-1} + b_{t-1}) \\
\text{Trend:} & b_t & = & \beta \cdot (L_t - L_{t-1}) + (1-\beta) \cdot b_{t-1} \\
\text{Saison:} & S_t & = & \gamma \cdot (Y_t - L_t) + (1-\gamma) \cdot S_{t-s} \\
\text{Prognose:} & Y_t(r) & = & L_t + b_t \cdot r + S_{t+r-s}, \quad r = 1, \ldots, s \\
 & Y_t(r) & = & L_t + b_t \cdot r + S_{t+r-2s}, \quad r = s+1, \ldots, 2s \\
 & \text{usw.} & &
\end{array}
$$

Holt-Winters: Die Gleichungen

- Die Prognosegleichung für $r = 1$:

$$Y_t(1) = L_t + b_t \cdot (1) + S_{t-s+1}$$

  Die 1-Schritt Prognose setzt sich aus dem geglätteten Niveau, L_t, der geglätteten 1-Perioden Veränderung, b_t, und der Saisonkomponente aus $[(t+1) - s]$ zusammen.

- Die Saisongleichung glättet die saisonalen Effekte in der niveaubereinigten Reihe $(Y_t - L_t)$ und verwendet dabei vergleichbare saisonale Effekte aufeinanderfolgender Jahre.

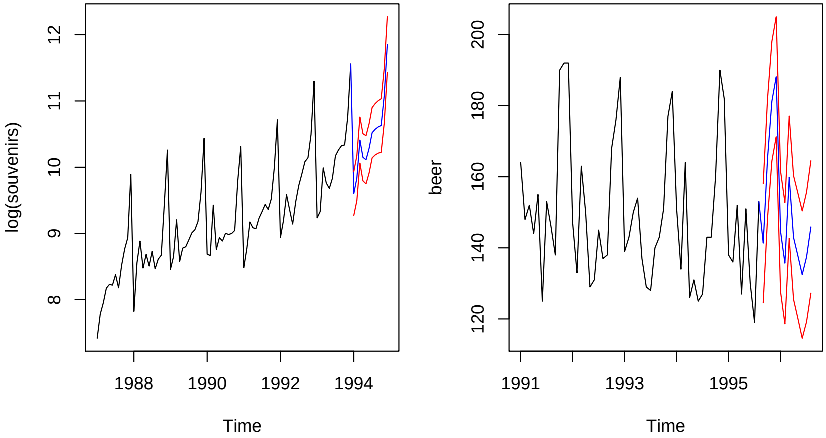
$$S_t = \gamma \cdot (Y_t - L_t) + (1 - \gamma) \cdot S_{t-s}$$

- Die Trendgleichung ist identisch mit der im Holt-Verfahren.
- Die Niveaugleichung glättet die saisonbereinigte Reihe $(Y_t - S_{t-s})$.

$$L_t = \alpha \cdot (Y_t - S_{t-s}) + (1 - \alpha) \cdot (L_{t-1} + b_{t-1})$$

Prognose mit Holt-Winters: `log(souvenirs)` und `beer`

α=`0.218`, β=`0.013`, γ=`0.484` α=`0.025`, β=`0.721`, γ=`0.223`



`BLAU` ... Punktprognosen `ROT` ... 95%Prognoseintervalle

Prognose mit Holt-Winters: Kommentar

- `log(souvenirs)`: Ist der Trend (Steigung) einmal gewählt, wird er kaum mehr angepasst. α von 0.218 zeigt an, dass das aktuelle Niveau nur mit 21.8 Prozent in die Prognose eingeht. Der aktuelle saisonale Ausschlag geht mit einem relativ großen Gewicht von 48.4 Prozent ein.
- `beer`: Hier ist für das Niveau kaum eine Anpassung notwendig. Der aktuelle Trend geht mit 72.1 Prozent (β) in die Prognose ein. Der Beitrag des aktuellen Saisoneffekts ist 22.3 Prozent ($\gamma = 0.223$).

Software:
Es gibt eine Reihe von Programmen, die diese und andere Prognoseverfahren z.T. automatisiert implementiert haben: AutoBox, EViews, ForecastPro, Minitab, R, S, SAS, STATA, u.a.

Zusammenfassung

- Beschreiben von Zeitreihen
- Klassische Zeitreihenzerlegung
- Trend, Saison, irreguläre Komponente
- Trendmodelle, Glättung, Trendbereinigung
- Berechnung der Saisonkomponenten, Saisonbereinigung
- Prognose, Maße für die Prognosegüte
- Prognoseverfahren von Holt und von Holt-Winters

Weiterführende Literatur

Makridakis, S., Wheelwright, S.C. and Hyndman, R.J.: *Forecasting, Methods and Applications*, John Wiley.