

# **Zeitreihenanalyse**

Fallbeispiele aus der Meteorologie

**Diplomarbeit**  
**in der Studienrichtung Technische Mathematik**  
zur Erlangung des akademischen Grades  
Diplom-Ingenieurin

eingereicht an der  
**Fakultät für Mathematik, Informatik und Physik**  
**der Universität Innsbruck**

von  
**Dorothea Huber**

Betreuer der Diplomarbeit:  
Univ.-Prof. Dr. Norbert Netzer

Innsbruck, August 2009

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>4</b>
1.1	Problemstellung . . . . .	4
1.2	Überblick . . . . .	4
<b>2</b>	<b>Deskriptive Zeitreihenanalyse</b>	<b>6</b>
2.1	Definitionen . . . . .	6
2.2	Das additive Komponentenmodell . . . . .	7
2.2.1	Box-Cox-Transformationen . . . . .	7
2.3	Das standardisierende Komponentenmodell . . . . .	9
2.4	Schätzen von Trend und saisonaler Komponente . . . . .	10
2.4.1	Trendtest nach Mann . . . . .	10
2.4.2	Schätzen der saisonalen Komponente . . . . .	10
2.4.3	Schätzen der saisonalbedingten Standardabweichung . . . . .	11
2.5	Praktischer Teil . . . . .	12
2.5.1	Graphische Darstellung . . . . .	12
2.5.2	Trendtest nach Mann . . . . .	13
2.5.3	Additives Komponentenmodell . . . . .	14
2.5.4	Standardisierendes Komponentenmodell . . . . .	17
2.5.5	Bestimmtheitsmaß . . . . .	18
<b>3</b>	<b>ARMA-Prozesse</b>	<b>19</b>
3.1	Grundlagen-Definitionen . . . . .	20
3.2	Stationäre und invertierbare Arma-Prozesse . . . . .	22
3.3	Schätzer für $\mu$ , $\gamma(h)$ und $\rho(h)$ . . . . .	23
3.4	Prognose mit ARMA-Modellen . . . . .	25
3.5	Parameterschätzung für ARMA-Prozesse . . . . .	30
3.5.1	Berechnung der Startwerte . . . . .	31
3.5.2	Hannan-Rissanen Algorithmus . . . . .	31
3.5.3	Maximum-Likelihood-Schätzer . . . . .	33
3.6	Identifikation des Modells . . . . .	35
3.6.1	Identifikation von p und q anhand der Autokorrelation und der partiellen Autokorrelation . . . . .	35
3.6.2	Informationskriterien . . . . .	40
3.7	Nichtstationäre Zeitreihen . . . . .	41
3.7.1	Der ARIMA-Prozess . . . . .	42
3.7.2	Der SARIMA-Prozess . . . . .	42

3.8	Praktischer Teil . . . . .	44
3.8.1	ARMA-Modelle für das additive Komponentenmodell . . . . .	44
3.8.2	ARMA-Modelle für das standardisierende Komponentenmodell . . . . .	45
3.8.3	Modellüberprüfung . . . . .	46
<b>4</b>	<b>Bivariate Zeitreihenanalyse</b>	<b>48</b>
4.1	Erwartungswert und Korrelation von bivariaten Prozessen . . . . .	48
4.2	Unabhängigkeit von zwei Zeitreihenvariablen . . . . .	49
4.3	Richtung des linearen Zusammenhangs zweier Variablen . . . . .	50
4.4	Bivariate ARMA-Prozesse . . . . .	50
4.4.1	Parameterschätzung bivariater Autoregressiver Prozesse . . . . .	50
4.4.2	Identifikation des autoregressiven Modells . . . . .	52
4.5	Transferfunktionenmodell . . . . .	52
4.6	Verfahren zur Identifikation von Transferfunktionenmodellen . . . . .	54
4.7	Vorhersage aufgrund eines Transferfunktionenmodells . . . . .	59
4.8	Praktischer Teil . . . . .	60
4.8.1	Tageshöchsttemperatur . . . . .	60
4.8.2	Tagestemperaturunterschied . . . . .	65
4.8.3	Nachttemperaturunterschied . . . . .	66
4.8.4	Nacht tiefsttemperatur . . . . .	67
4.8.5	Überprüfung des Modells . . . . .	68
<b>5</b>	<b>Tagesgang der Temperatur</b>	<b>71</b>
5.1	Tagesgang der Temperatur mit dem Transferfunktionenmodell . . . . .	71
5.2	Temperaturkurve . . . . .	75
	<b>Anhang</b>	<b>79</b>
	<b>Literaturverzeichnis</b>	<b>80</b>

# 1 Einleitung

Die Zeitreihenanalyse findet in sehr vielen Wissensgebieten Anwendung. Der zeitliche Ablauf gewisser Ereignisse wird festgehalten, analysiert und es wird versucht daraus eine Prognose der zukünftigen Entwicklung zu erstellen. So werden beispielsweise in der Medizin Fieberkurven angefertigt, daraus werden unter anderem Rückschlüsse auf den Krankheitsverlauf gezogen. Seismologen registrieren ständig die Erschütterungen der Erdkruste, um so Frühwarnungen für Erdbeben geben zu können [Hartung].

## 1.1 Problemstellung

In dieser Diplomarbeit wird versucht anhand von meteorologischen Zeitreihen die Lufttemperatur (in 2 m Höhe) für den nächsten Tag und für die nächsten Stunden vorherzusagen. Dazu wurden Daten der Wetterstationen Nummer 11121 und 11120 in Innsbruck verwendet. Als Kennwerte für die Temperatur wurde die Tageshöchsttemperatur, der Tagestemperaturunterschied (diurnal temperature range, DTR), der Nachttemperaturunterschied (NTR) und die Nachttiefsttemperatur gewählt. Die Vorhersage dieser Kennwerte kann entweder univariat erfolgen oder unter Miteinbeziehung von einer zweiten Zeitreihe. Für die bivariate Zeitreihenanalyse wurden die Bewölkung und die Globalstrahlung gewählt.

## 1.2 Überblick

Es soll hier ein Überblick über die einzelnen Kapitel bzw. Schritte der Problemlösung gegeben werden.

Das **zweite Kapitel** beschäftigt sich mit der deskriptiven Zeitreihenanalyse. Die Zeitreihe wird als sogenanntes Komponentenmodell betrachtet. Hierbei wird davon ausgegangen, dass sich die Zeitreihe aus einer Trendkomponente, einer zyklischen Komponente, einer saisonalen Komponente und einer irregulären Komponente zusammensetzt. Die Komponenten werden im Wesentlichen mit den Methoden der linearen Regression geschätzt.

Die Zeitreihen der maximalen Tagestemperatur, des Tagestemperaturunterschieds, des Nachttemperaturunterschieds, der minimalen Nachttemperatur sowie der mittleren Globalstrahlung pro Tag werden durch zwei verschiedene Komponentenmodelle beschrieben.

Das **dritte Kapitel** beschäftigt sich mit der sogenannten stochastischen Zeitreihenanalyse. Die 1970 vorgestellte Methode von Box und Jenkins betrachtet die irreguläre Komponente vom Zeitreihenmodell als stochastischen Prozess. Die autoregressiven Prozesse (AR) erklären sich aus sich selbst heraus, die moving average Prozesse (MA) beziehen sich auf einem Fehlerprozess. Möglich ist natürlich auch eine Kombination der beiden, ein sogenannter ARMA-Prozess. Soll ein solcher Prozess an eine Zeitreihe angepasst werden, so müssen nicht nur die Prozessparameter geschätzt werden sondern auch das geeignete Modell identifiziert werden.

Für die Modellidentifikation werden zwei Methoden behandelt: die graphische Methode mittels Autokorrelations- und partieller Autokorrelationsfunktion und die Identifikation mit sogenannten Informationskriterien. Die dazugehörigen Prozessparameter werden mit der Maximum-Likelihood-Schätzung ermittelt.

Für die Prognose von Zeitreihen wird das Innovationen-Algorithmus vorgestellt.

Für nichtstationäre Zeitreihen wurde von Box und Jenkins eine Erweiterung des ARMA-Modells zum SARIMA-Modell entwickelt. Dabei wird der polynomiale Trend durch Differenzenbildung eliminiert und die Saisonalität durch ein ARMA-Modell beschrieben.

Im praktischen Teil werden für die irregulären Komponenten der Zeitreihen aus den Komponentenmodellen ARMA-Modelle identifiziert und ihre Prognosefähigkeit durch Vergleich mit den tatsächlichen Zeitreihenwerten getestet.

Das **vierte Kapitel** befasst sich mit der bivariaten Zeitreihenanalyse. Hierbei soll zunächst die statistische Abhängigkeit nachgewiesen und die Richtung des linearen Zusammenhangs ermittelt werden.

Die zwei behandelten bivariaten Zeitreihenmodelle sind der bivariate autoregressive Prozess (falls sich die Zeitreihen gegenseitig beeinflussen) und das Transferfunktionenmodell. Dieses beschreibt den linearen Zusammenhang von zwei Zeitreihen, wobei eine Zeitreihe als Input oder unabhängige Variable und die andere als Output oder abhängige Variable angesehen wird. Auch dieses Verfahren wurde bereits 1970 von Box und Jenkins entwickelt.

Im praktischen Teil wird das Transferfunktionenmodell zur Vorhersage der Tageshöchsttemperatur, des DTR und des NTR angewandt. Als Input werden die mittlere Globalstrahlung, die mittlere Tages- oder Nachtbewölkung herangezogen.

Im **fünften Kapitel** wird versucht die Temperatur der nächsten Stunden mit Hilfe des Transferfunktionenmodells vorherzusagen. Dabei wird zunächst untersucht, ob die drei Zeitreihen der Windgeschwindigkeit, des Bewölkungsgrades und der Globalstrahlung als Input des Transferfunktionenmodells geeignet sind. Ein signifikanter Zusammenhang kann nur zwischen der Temperatur und der Globalstrahlung festgestellt werden.

Als nächstes wird der Temperaturgang von wolkenlosen bis fast wolkenlosen Tagen mit dem Temperaturgang von fast bedeckten bis bedeckten Tagen graphisch verglichen. Hierbei können vor allem Unterschiede bezüglich des Zeitpunkts der Nachttiefsttemperatur beobachtet werden.

## 2 Deskriptive Zeitreihenanalyse

Ziel der deskriptiven Zeitreihenanalyse ist die Darstellung der Zeitreihe als sogenanntes Komponentenmodell. Hierbei wird davon ausgegangen, dass sich die Zeitreihe aus einer Trendkomponente, einer zyklischen Komponente, einer saisonalen Komponente und einer irregulären Komponente zusammensetzt. Je nachdem wie die Komponenten miteinander verknüpft sind, spricht man von einem additiven, multiplikativen oder gemischten Komponentenmodell. Nachdem Trend und saisonale Komponente geschätzt wurden, wird die irreguläre Komponente aus dem Modell extrahiert, diese soll (schwach) stationär sein, damit sie im nächsten Kapitel mit den Methoden von Box und Jenkins weiter untersucht werden kann.

Die meisten meteorologische Zeitreihen, wie Temperatur oder Globalstrahlung, weisen starke Schwankungen im Jahresverlauf auf. Zusätzlich kann aber auch eine Varianzheterogenität beobachtet werden (vgl. dazu grafische Darstellungen im praktischen Teil). Aus diesem Grund muss beim additiven Komponentenmodell eine varianzstabilisierende Transformation der Zeitreihe durchgeführt werden oder man betrachtet ein alternatives Modell. Bei diesem wird die Zeitreihe aufgrund der saisonalen Komponente und der saisonalen Varianz standardisiert [Wilks].

Die Komponenten beider Zeitreihenmodelle werden mit den Methoden der linearen Regression geschätzt. Die saisonale Komponente und die saisonale Standardabweichung der Zeitreihe werden als trigonometrische Polynome geschätzt.

### 2.1 Definitionen

**Definition 1.** Eine **Zeitreihe** ist eine Folge von zeitlich geordneten Beobachtungen. Wenn die Menge der Beobachtungen diskret ist, spricht man von einer **diskreten Zeitreihe**. Werden hingegen die Beobachtungen über einen Zeitintervall kontinuierlich verzeichnet, spricht man von einer **stetigen Zeitreihe**.

Die Beobachtungen einer diskreten Zeitreihe, die zu den Zeitpunkten  $t_1, t_2, \dots, t_n$  gemacht werden, werden mit  $y(t_1), y(t_2), \dots, y(t_n)$  bezeichnet. Bei diskreten Zeitreihen mit äquidistanten Zeitpunkten verwendet man folgende Notation:

$$y = (y_1, y_2, \dots, y_n)$$

wobei  $y_t$  die Beobachtung zum Zeitpunkt  $t$  und  $n$  die Länge der Zeitreihe ist [Box, Jenkins].

**Definition 2.** In der Theorie der deskriptiven Methoden der Zeitreihenanalyse wird davon ausgegangen, dass sich eine Zeitreihe in einem Trend, einer zyklischen Komponente, einer saisona-

len Komponente und in einer irregulären Komponenten darstellen lässt. Dieser Zugang ist nicht immer der beste, aber sehr nützlich, wenn die Variation der Zeitreihe stark von einem Trend oder Saisonalität beeinflusst wird [Chatfield].

**Trend (T):** Ist die langfristige Veränderung der Zeitreihe in der Zeit, wobei "langfristig" von der Art der Zeitreihe abhängt [Chatfield].

**Saisonale Komponente (S):** Der Verlauf der saisonalen Komponente ist wellenförmig. Sie entsteht durch einen periodischen Zeiteinfluss auf die Beobachtungswerte. Ein solcher periodischer Einfluss ist zum Beispiel der Einfluss der Jahreszeiten [Hartung].

**Zyklische Komponente (Z):** Ist ebenfalls eine wellenförmige Komponente, die aber keine von Außen vorgegebene Periode hat. Die beobachtete Periode kann aber geschätzt werden. Ein Beispiel für eine zyklische Komponente ist der 7-jährige Wirtschaftszyklus [Chatfield].

**Irreguläre Komponente (R):** Nachdem Trend, zyklische Komponente und Saisonalität von der Zeitreihe eliminiert wurden, bleiben die Residuen, die zufällig sein können oder als selbsterklärendes Modell dargestellt werden können [Chatfield]. Die Zeitreihe der Residuen soll stationär sein [Brockwell, Davis. 2002].

**Glatte Komponente (G):** Trend und zyklische Komponente können als glatte Komponente zusammengefasst werden [Hartung].

## 2.2 Das additive Komponentenmodell

Falls die Komponenten als additiv miteinander verknüpft angenommen werden, spricht man vom **additiven Komponentenmodell**

$$\begin{aligned} y_t &= T_t + Z_t + S_t + R_t, \\ &= G_t + S_t + R_t \quad \text{für } t = 1, \dots, n \end{aligned} \tag{2.1}$$

In Fällen von Varianzheterogenität oder falls sich die saisonale Komponente und die irreguläre Komponente proportional zum Trend verändern, wird jedoch zunächst eine varianzstabilisierende Transformation der Zeitreihe, eine sogenannte Box-Cox-Transformation, durchgeführt [Brockwell, Davis. 2002].

### 2.2.1 Box-Cox-Transformationen

Zu den varianzstabilisierenden Transformationen gehören die Box-Cox-Transformationen, welche aus [Pengfei Li] entnommen wurden. Die ursprüngliche Form der Box-Cox-Transformation,

die 1964 von Box und Cox veröffentlicht wurde, war

$$y(\lambda) = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \text{wenn } \lambda \neq 0; \\ \log(y), & \text{wenn } \lambda = 0. \end{cases} \quad (2.2)$$

Eine Erweiterung davon erlaubt auch negative Daten zu transformieren.

$$y(\boldsymbol{\lambda}) = \begin{cases} \frac{(y + \lambda_2)^{\lambda_1} - 1}{\lambda_1}, & \text{wenn } \lambda_1 \neq 0; \\ \log(y + \lambda_2), & \text{wenn } \lambda_1 = 0. \end{cases} \quad (2.3)$$

Wobei  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)'$  ist. In der Praxis wird  $\lambda_2$  so gewählt, dass  $y + \lambda_2 > 0$  ist.

Ziel der Box-Cox-Transformation ist die Zeitreihe so umzuformen, dass sie annähernd normalverteilt ist, d.h.  $y(\lambda) \sim N(X\beta, \sigma^2 I_n)$ . Wobei  $X$  die Datenmatrix der unabhängigen Variablen ist, neben der Zeit können auch andere exogene Variablen auftreten.  $\beta$  ist der Parametervektor der linearen Regression und  $\sigma^2$  ist die Varianz.

Die in dieser Diplomarbeit betrachteten Zeitreihen haben keinen signifikanten Trend und die saisonale Komponente wird als trigonometrisches Polynom dargestellt (vgl. Ausführungen im Abschnitt 2.5). Die Datenmatrix  $X$  ist somit

$$\mathbf{X} = \begin{pmatrix} 1 & \cos(2\pi t_1/365) & \sin(2\pi t_1/365) \\ \vdots & \vdots & \vdots \\ 1 & \cos(2\pi t_n/365) & \sin(2\pi t_n/365) \end{pmatrix} \quad (2.4)$$

### Schätzen des Parameters $\lambda$

Für die Schätzung des Parameters  $\lambda$  wird angenommen, dass  $\mathbf{y}(\lambda) \sim N(X\beta, \sigma^2 I_n)$  ist. Die dazugehörige Dichte für  $\mathbf{y}(\lambda)$  ist

$$f(\mathbf{y}(\lambda)) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{y}(\lambda) - \mathbf{X}\beta)'(\mathbf{y}(\lambda) - \mathbf{X}\beta)\right) \quad (2.5)$$

Sei  $J(\lambda, \mathbf{y})$  die Funktionaldeterminante oder Jacobi-Determinante der Transformation von  $\mathbf{y}$  zu  $\mathbf{y}(\lambda)$ , dann ist die Dichtefunktion für  $\mathbf{y}$  und somit die Likelihoodfunktion für das ganze Modell

$$L(\lambda, \beta, \sigma^2 | \mathbf{y}, \mathbf{X}) = f(\mathbf{y}) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{y}(\lambda) - \mathbf{X}\beta)'(\mathbf{y}(\lambda) - \mathbf{X}\beta)\right) J(\lambda, \mathbf{y}) \quad (2.6)$$

Um davon das Maximum-Likelihood zu erhalten können noch ein paar Vereinfachungen durchgeführt werden. Für fixes  $\lambda$  entspricht die obige Likelihoodfunktion der Likelihoodfunktion einer multiplen linearen Regression. Die Schätzer für  $\beta$  und  $\sigma^2$  sind

$$\tilde{\beta}(\lambda) = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}(\lambda) \quad (2.7)$$

$$\widetilde{\sigma^2}(\lambda) = \frac{\mathbf{y}(\lambda)'(I_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{y}(\lambda)}{n} \quad (2.8)$$



Die Jacobi-Determinante  $J(\lambda, \mathbf{y})$  der Box-Cox-Transformation ist

$$J(\lambda, \mathbf{y}) = \begin{vmatrix} \frac{\partial y_1(\lambda)}{\partial y_1} & \dots & \frac{\partial y_1(\lambda)}{\partial y_n} \\ \frac{\partial y_2(\lambda)}{\partial y_1} & \dots & \frac{\partial y_2(\lambda)}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial y_n(\lambda)}{\partial y_1} & \dots & \frac{\partial y_n(\lambda)}{\partial y_n} \end{vmatrix} = \begin{vmatrix} \frac{\partial y_1(\lambda)}{\partial y_1} & 0 & \dots & 0 \\ 0 & \frac{\partial y_2(\lambda)}{\partial y_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \frac{\partial y_n(\lambda)}{\partial y_n} \end{vmatrix} = \prod_{i=1}^n y_i^{\lambda-1} \quad (2.9)$$

Somit erhält man folgende Vereinfachung des Logarithmus der Likelihoodfunktion (2.5):

$$l(\lambda | \mathbf{y}, \mathbf{X}, \tilde{\boldsymbol{\beta}}(\lambda), \tilde{\sigma}^2(\lambda)) = C - \frac{n}{2} \log(\tilde{\sigma}^2(\lambda)) + (\lambda - 1) \sum_{i=1}^n \log(y_i) \quad (2.10)$$

Sei  $g$  der geometrische Mittelwert der ursprünglichen Zeitreihe  $g = (\prod_{i=1}^n y_i)^{1/n}$ . Die Box-Cox-Transformation wird leicht modifiziert, sodass  $\mathbf{y}(\lambda, g) = \frac{y^{\lambda-1}}{g^{\lambda-1}}$ . Dann ist

$$\begin{aligned} l(\lambda | \mathbf{y}, \mathbf{X}) &= C - \frac{n}{2} \log(s_\lambda^2) + \sum_{i=1}^n \log\left(\frac{y_i^{\lambda-1}}{g^{\lambda-1}}\right) \\ &= C - \frac{n}{2} \log(s_\lambda^2) + \sum_{i=1}^n \log(y_i^{\lambda-1}) - n \log\left(\left(\prod_{i=1}^n y_i\right)^{\frac{1}{n}(\lambda-1)}\right) \\ &= C - \frac{n}{2} \log(s_\lambda^2) \end{aligned}$$

Die Likelihoodfunktion wird maximiert indem man das  $\lambda$  findet welches die Quadratsumme der Residuen

$$s_\lambda^2 = \frac{\mathbf{y}(\lambda, g)'(I_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{y}(\lambda, g)}{n}$$

minimiert.

### 2.3 Das standardisierende Komponentenmodell

Bei einer saisonal bedingte Varianzinhomogenität können die Daten der Zeitreihe auch standardisiert werden, sodass man von folgendem Komponentenmodell ausgeht

$$y_t = T_t + S_t + \sigma_t \cdot R_t, \quad \text{für } t = 1, \dots, n \quad (2.11)$$

Wobei  $\sigma_t$  die Standardabweichung zum Zeitpunkt  $t$  ist, sie ist erfahrungsgemäß im Sommer höher als im Winter [Wilks].

## 2.4 Schätzen von Trend und saisonaler Komponente

Es gibt verschiedene Methoden den Trend und die saisonale Komponente zu schätzen. Für kurze Zeitreihen genügt es oft einen linearen Trend anzunehmen. Denkbar sind aber auch polynomiale Trends oder logarithmische Trends. Für längere Zeitreihen im Bereich der Wirtschaft ist es oft sinnvoll von Wachstumsfunktionen mit einer Sättigungsgrenze auszugehen [Hartung]. Die Trends werden dann mit den Methoden der Regression geschätzt. Da bei den in dieser Diplomarbeit betrachteten Zeitreihen kein Trend erkennbar ist (vgl. Ausführung unten), wird lediglich ein Signifikanztest für einen linearen Trend nach Mann [Hartung] angegeben.

### 2.4.1 Trendtest nach Mann

**Satz 1.** Die Prüfgröße für den approximativen **Trendtest nach Mann** lautet:

$$C = \sum_{t=1}^{n-1} \sum_{j=t+1}^n \operatorname{sgn}(y_j - y_t) \quad (2.12)$$

wobei  $\operatorname{sgn}$  die Signum-Funktion bezeichnet (dh.  $\operatorname{sgn}(y_j - y_t)$  ist gleich 1 falls  $y_j - y_t > 0$ , gleich  $-1$ , falls  $y_j - y_t < 0$  und gleich 0 falls  $y_j - y_t = 0$  ist). Die Größe  $C^*$

$$C^* = C / \sqrt{(n(n-1)(2n+5)/18)} \quad (2.13)$$

ist approximativ standardnormalverteilt.

Getestet wird  $H_0$ : kein Trend gegen  $H_1$ : Trend, dh. die Nullhypothese kann mit einer Irrtumswahrscheinlichkeit von 5% verworfen werden, falls

$$|C^*| > 1,96 \quad (2.14)$$

ist [Hartung].

### 2.4.2 Schätzen der saisonalen Komponente

Die trendbereinigten Zeitreihe  $d_t = y_t - T_t$  setzt sich näherungsweise aus saisonaler und irregulärer Komponente zusammen, d.h.

$$d_t \approx S_t + R_t.$$

Um die saisonale Komponente zu schätzen wird laut [Hartung] das arithmetische Mittel gebildet, da wegen der Konstanz der Saisonsfigur die Größen  $d_t, d_{t+p}, d_{t+2p} \dots$  um einen festen Wert schwanken:

$$\bar{d}_t = \frac{1}{n_t} \sum_{j=0}^{n_t-1} d_{t+jp} \quad \text{für } t = 1, \dots, p \quad (2.15)$$

wobei  $p$  die Länge der Periode und  $n_t$  die Anzahl der Perioden ist. Falls die Zeitreihe noch nicht mittelwertbereinigt ist, normiert man die Größen  $\bar{d}_1, \dots, \bar{d}_t$  noch wie folgt:

$$\hat{s}_t = \bar{d}_t - \frac{1}{p} \sum_{j=1}^p \bar{d}_j \quad \text{für } t = 1, \dots, p \quad (2.16)$$

So kann  $\hat{s}_t = \hat{s}_{t+p} = \hat{s}_{t+2p} = \dots$  als Schätzer für die Saisonskomponenten  $S_t = S_{t+p} = S_{t+2p} = \dots$  angewandt werden, da die irreguläre Komponente um Null schwankt. Die Zeitreihe  $d_t - \hat{s}_t$  ist dann trend- und saisonbereinigt.

**Bemerkung 1.** Falls man nicht von einer starren Saisonsfigur ausgehen möchte, ist es sinnvoll die saisonale Komponenten als trigonometrisches Polynom erster Ordnung darzustellen [Wilks]. Bei einer Periode von 365 Tagen erhält man folgende Darstellung

$$\hat{s}(t) = \alpha \cos\left(\frac{2\pi t}{365}\right) + \beta \sin\left(\frac{2\pi t}{365}\right) \quad (2.17)$$

Die Konstanten  $\alpha$  und  $\beta$  findet man durch Minimieren der Summe der Quadratfehler.

$$\begin{aligned} \alpha &= \frac{2}{365} \sum_{t=1}^n \hat{s}(t) \cos\left(\frac{2\pi t}{365}\right) \\ \beta &= \frac{2}{365} \sum_{t=1}^n \hat{s}(t) \sin\left(\frac{2\pi t}{365}\right) \end{aligned} \quad (2.18)$$

### 2.4.3 Schätzen der saisonalbedingten Standardabweichung

Für das standardisierende Komponentenmodell soll auch noch die periodische Standardabweichung  $\sigma_t$  geschätzt werden. Ein Schätzer dafür ist die empirische Standardabweichung

$$\hat{\sigma}_t = \sqrt{\frac{1}{n_t - 1} \sum_{i=0}^{n_t-1} (y(t + ip) - \overline{y(t)})^2} \quad \text{für } t = 1, \dots, p \quad (2.19)$$

wobei

$$\overline{y(t)} = \frac{1}{n_t} \sum_{j=0}^{n_t-1} y_{t+jp} \quad (2.20)$$

ist. Dabei ist  $p$  die Länge der Periode (365 Tage) und  $n_t$  die Anzahl der Perioden. Auch hier kann die Standardabweichung  $\hat{\sigma}_t$  für  $t = 1, \dots, p$  als trigonometrisches Polynom dargestellt werden.

## 2.5 Praktischer Teil

Für diese Diplomarbeit wurden Daten der Wetterstationen Nummer 11121 und 11120 in Innsbruck verwendet. Die gemessenen Daten, wie Lufttemperatur, Bodentemperatur, Niederschlagsmenge, Windrichtung und Windgeschwindigkeit, relative Luftfeuchte, Sonnenscheindauer sowie Globalstrahlung, liegen in 10-Minuten-Intervallen vor.

Die beobachteten Daten, wie Stärke und Art der Bewölkung, des Niederschlags und des Nebels, werden hingegen jede Stunde, mit Ausnahme einiger Nachtstunden verzeichnet.

Die Globalstrahlung wird in Millivolt  $mV$  angegeben. Und die betrachtete Lufttemperatur (in 2 m Höhe) ist in Zehntel-Grad Celsius angeführt.

Als relevante Bewölkung wird die Menge der tiefen Wolken angenommen. Die Stärke der Bewölkung wird als Gesamtbedeckung des Himmels in Achteln angegeben.

0/8	wolkenlos
1/8	fast wolkenlos
2/8	leicht bewölkt
3/8	
4/8	wolkig
5/8	
6/8	stark bewölkt
7/8	fast bedeckt
8/8	bedeckt

Für die Unterscheidung zwischen Tag und Nacht wurde der effektive Sonnenaufgang und Sonnenuntergang gewählt. Die Zeitreihen der maximalen Tagestemperatur ( $y_{Tmax}$ ), des Tagestemperaturunterschieds ( $y_{DTR}$ ), des Nachttemperaturunterschieds ( $y_{NTR}$ ), der minimalen Nachttemperatur ( $y_{Tmin}$ ) und der mittleren Globalstrahlung pro Tag ( $y_{Glo}$ ) werden im Zeitraum vom 3.Juni 2003 bis am 31.Mai 2008 betrachtet. Somit sind die Zeitreihen 1825 Tage lang, also 5 Perioden zu 365 Tagen.

Die fünf Zeitreihen sollen zum einen als additives Komponentenmodell und zum anderen als standardisierendes Modell dargestellt werden. Ziel ist es eine stationäre und somit varianzhomogene irreguläre Komponente der Zeitreihen zu extrahieren, damit diese im nächsten Kapitel mit der Methode von Box-Jenkins weiter untersucht werden kann.

Die Berechnungen und die grafischen Darstellungen wurden mit der Software R durchgeführt.

### 2.5.1 Graphische Darstellung

Der erste und wichtigste Schritt in der Zeitreihenanalyse ist die graphische Darstellung der Zeitreihen [Chatfield]. Diese Graphen geben einen ersten Einblick in Bezug auf Trend und Saisonalität. Bei den fünf Zeitreihen im Zeitraum vom 3.Juni 2003 bis am 31.Mai 2008 ist aus der

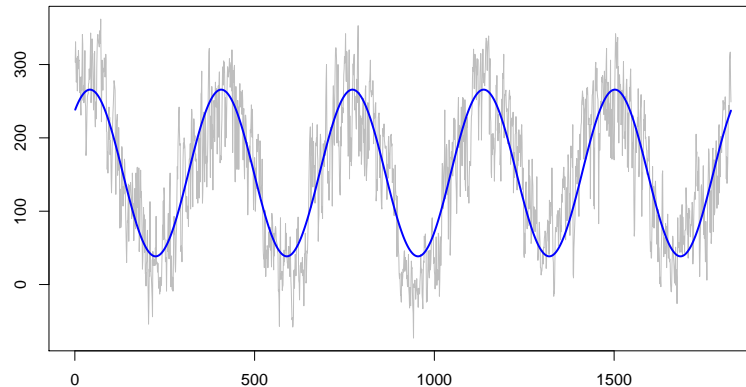


Abbildung 2.1: Tageshöchsttemperatur

Grafik kein Trend erkennbar. Sie sind nämlich mit 5 Jahren für meteorologische Verhältnisse ziemlich kurz. Alle fünf Zeitreihen weisen allerdings saisonale Schwankungen auf. Zudem ist die Varianz nicht homogen. In den Sommermonaten ist die Standardabweichung beispielsweise größere als in den Wintermonaten.

Das Minimum der Tageshöchsttemperatur lag bei  $-7,3$  Grad Celsius, das Maximum bei  $36,2$  und die mittlere Tageshöchsttemperatur bei  $15,2$  Grad Celsius. Der Tagestemperaturunterschied lag zwischen einem Minimum von  $0,6$  und einem Maximum von  $24,2$  Grad Celsius, der Mittelwert war  $10$  Grad Celsius. Der Nachttemperaturunterschied lag hingegen zwischen  $0,3$  und  $19,8$  Grad Celsius mit einem Durchschnitt von  $8,3$  Grad. Die Nachttiefsttemperatur schwankte zwischen  $-17,3$  und  $22,1$  Grad Celsius. Der Durchschnitt war bei  $4$  Grad Celsius. Die mittlere Globalstrahlung pro Tag wurde zwischen  $0,58$  und  $64$   $mV$  gemessen. Der Durchschnitt lag bei  $21$   $mV$ .

### 2.5.2 Trendtest nach Mann

Da für die fünf Zeitreihen kein Trend zu erkennen ist, wird der Trendtest nach Mann durchgeführt. Die Prüfgröße  $C^*$  lautet

$$C = \sum_{t=1}^{n-1} \sum_{j=t+1}^n \operatorname{sgn}(y_j - y_t)$$

$$C^* = C / \sqrt{(n(n-1)(2n+5)/18)}.$$

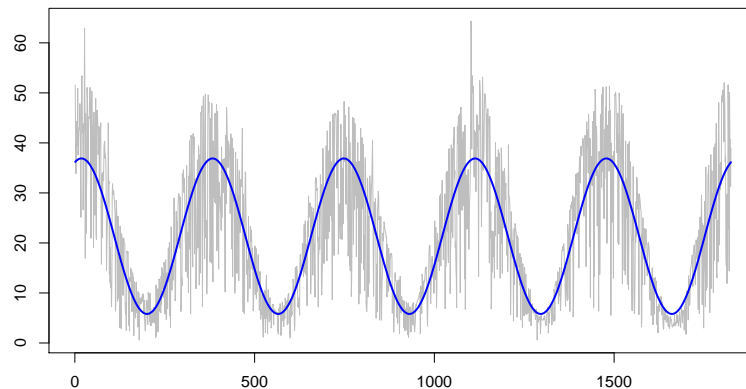


Abbildung 2.2: mittlere Globalstrahlung

$C^*$  ist annähernd standardnormalverteilt.

Die Ergebnisse des Trendtests nach Mann für  $y_{Tmax}$ ,  $y_{DTR}$ ,  $y_{NTR}$ ,  $y_{Tmin}$  und  $y_{Glo}$  sind:

$C^*(y_{Tmax})$	=	0.005
$C^*(y_{DTR})$	=	0.002
$C^*(y_{NTR})$	=	0.002
$C^*(y_{Tmin})$	=	-0.006
$C^*(y_{Glo})$	=	-0.001

Die Nullhypothese,  $H_0$  : kein Trend, kann zu einem Konfidenzintervall von 95% nicht verworfen werden, da die Testgrößen  $|C^*| \leq 1,96$  sind. Bei keiner Zeitreihe wird ein linearer Trend angenommen, dh. die Komponente  $T_t$  wird nur mit dem Mittelwert der Zeitreihe angegeben.

### 2.5.3 Additives Komponentenmodell

Bevor man für die fünf Zeitreihen den Trend und die saisonale Komponente des additiven Zeitreihenmodells

$$y_t = T_t + S_t + R_t, \quad \text{für } t = 1, \dots, n$$

schätzt, müssen varianzstabilisierende Transformationen durchgeführt werden, damit die erhaltene irreguläre Komponente stationär ist.

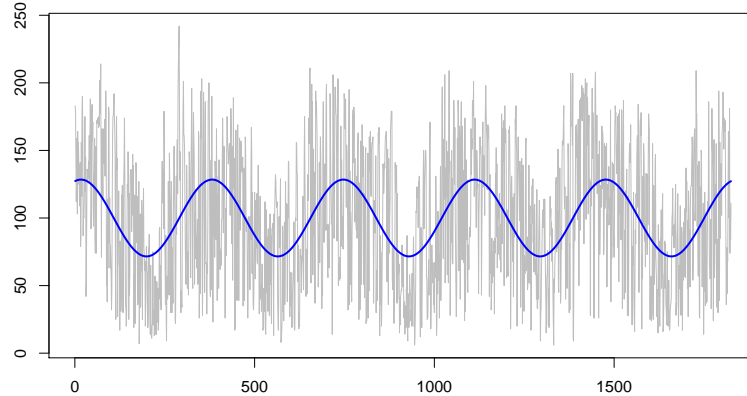


Abbildung 2.3: Tagestemperaturunterschied

### Box-Cox-Transformation

Für die Zeitreihen wird deshalb eine Box-Cox-Transformation durchgeführt.

$$y(\lambda) = \begin{cases} \frac{(y+\lambda_2)^{\lambda_1}-1}{\lambda_1}, & \text{if } \lambda_1 \neq 0; \\ \log(y + \lambda_2), & \text{if } \lambda_1 = 0. \end{cases}$$

Der Parameter  $\lambda_2$  wird so gewählt, dass alle Werte der Zeitreihe größer als Null sind. Für die Maximum-Likelihood-Schätzung des Parameters  $\lambda_1$  muss zunächst festgelegt werden welche Verteilung die transformierte Zeitreihe  $y(\lambda)$  haben soll. Da laut Mann bei keiner Zeitreihe ein signifikanter linearer Trend gefunden wurde, wird die Trendkomponente als Konstanten  $c$  gewählt. Die saisonale Komponente wird hingegen als trigonometrisches Polynom angenommen. Die Box-Cox-transformierte Zeitreihe soll also

$$y(\lambda) \sim N(c + \alpha \cos(2\pi t/365) + \beta \sin(2\pi t/365), \sigma^2 I_n) \quad (2.21)$$

sein. Es sollen nun die  $\lambda$  geschätzt werden, welche die Likelihood-Funktion

$$L(\lambda, \eta, \sigma^2 | \mathbf{y}, \mathbf{t}) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{y}(\lambda) - \mathbf{X}\boldsymbol{\eta})'(\mathbf{y}(\lambda) - \mathbf{X}\boldsymbol{\eta})\right) J(\lambda, \mathbf{y})$$

maximieren. Wobei  $\boldsymbol{\eta} = (c, \alpha, \beta)'$  ist und

$$\mathbf{X} = \begin{pmatrix} 1 & \cos(2\pi t_1/365) & \sin(2\pi t_1/365) \\ \vdots & \vdots & \vdots \\ 1 & \cos(2\pi t_n/365) & \sin(2\pi t_n/365) \end{pmatrix} \quad (2.22)$$

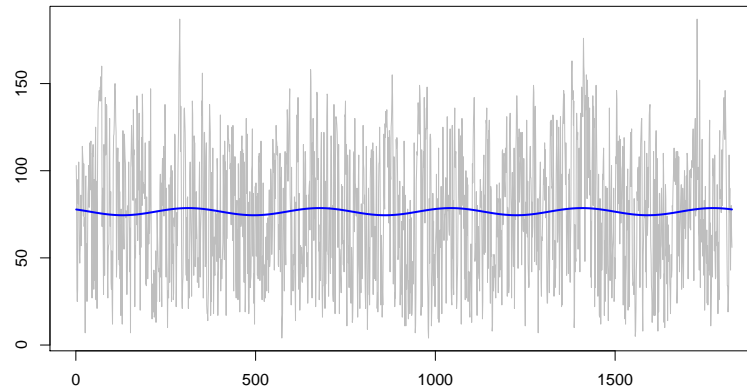


Abbildung 2.4: Nachttemperaturunterschied

Die mit dem Software-Programm R ermittelten  $\lambda$ s und die dazugehörigen Box-Cox-Transformationen sind:

$$y_{DTR}(\lambda) = \frac{y_{DTR}^{0.71} - 1}{0.71}$$

$$y_{NTR}(\lambda) = \frac{y_{NTR}^{0.71} - 1}{0.71}$$

$$y_{Global}(\lambda) = \frac{y_{Global}^{0.42} - 1}{0.42}$$

Für  $y_{Tmax}$  wurde ein  $\lambda$  von 0.95 gefunden, da aber im Konfidenzintervall auch 1 enthalten ist, was keiner Transformation entspricht, wird für  $y_{Tmax}$  keine Transformation durchgeführt.

Bei  $y_{Tmin}$  würde das  $\lambda = 1.43$  sein. Da dieser aber die Varianz der Zeitreihe erhöhen würde, wird auch hier keine Transformation vorgenommen.

### Trend und saisonale Komponente

Als Trend der transformierten Zeitreihen wird deren Mittelwert angenommen. Die saisonale Komponente wird als trigonometrisches Polynom geschätzt. Somit sind die additiven Komponentenmodelle der Box-Cox-transformierten Zeitreihen:

$$y_{Tmax,t} = 152 + 85.08 \cos(2\pi t/365) + 75.37 \sin(2\pi t/365) + R_{tmax1,t}$$

$$y_{DTR,t}(\lambda) = 34 + 7.32 \cos(2\pi t/365) + 2.28 \sin(2\pi t/365) + R_{dtr1,t}$$

$$y_{NTR,t}(\lambda) = 28 + 0.36 \cos(2\pi t/365) - 0.41 \sin(2\pi t/365) + R_{ntr1,t}$$



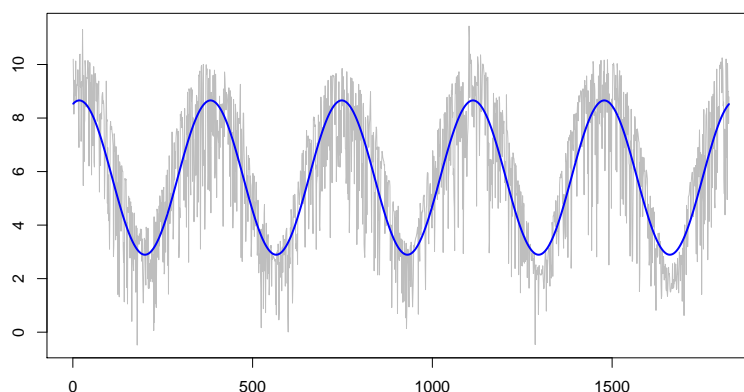


Abbildung 2.5: Box-Cox-transformierte Globalstrahlung

$$y_{Tmin,t} = 45.89 + 65.75 \cos(2\pi t/365) + 70.61 \sin(2\pi t/365) + R_{tmin1,t}$$

$$y_{Glo,t}(\lambda) = 5.78 + 2.74 \cos(2\pi t/365) + 0.88 \sin(2\pi t/365) + R_{glo1,t}$$

## 2.5.4 Standardisierendes Komponentenmodell

Für das standardisierende Modell

$$y_t = T_t + S_t + \sigma_t \cdot R_t, \quad \text{für } t = 1, \dots, n$$

sollen nicht nur Trend und Saisonalität geschätzt werden, sondern auch die Standardabweichung. Auch diese wird als trigonometrisches Polynom dargestellt. Für die fünf betrachteten Zeitreihen wurden folgendes **standardisierendes Modell** gefunden:

$$y_{Tmax,t} = 152.12 + 85.07 \cos(2\pi t/365) + 75.37 \sin(2\pi t/365) + (47.62 + 5.56 \cos(2\pi t/365) - 0.94 \sin(2\pi t/365)) \cdot R_{tmax2,t}$$

$$y_{DTR,t} = 100.01 + 27.21 \cos(2\pi t/365) + 8.14 \sin(2\pi t/365) + (43.86 + 3.46 \cos(2\pi t/365) - 0.87 \sin(2\pi t/365)) \cdot R_{dtr2,t}$$

$$y_{NTR,t} = 76.56 + 1.38 \cos(2\pi t/365) - 1.58 \sin(2\pi t/365) + (35.35 + 0.63 \cos(2\pi t/365) - 1.47 \sin(2\pi t/365)) \cdot R_{ntr2,t}$$

$$y_{Tmin,t} = 45.89 + 65.75 \cos(2\pi t/365) + 70.61 \sin(2\pi t/365) + (33.30 - 7.65 \cos(2\pi t/365) - 6.39 \sin(2\pi t/365)) \cdot R_{tmin2,t}$$

$$y_{Glo,t} = 21.35 + 14.78 \cos(2\pi t/365) + 4.80 \sin(2\pi t/365) + (7.15 + 4.46 \cos(2\pi t/365) + 1.56 \sin(2\pi t/365)) \cdot R_{glo2,t}$$

### 2.5.5 Bestimmtheitsmaß

Als Maß für die Anpassungsgüte des geschätzten Komponentenmodells kann das Bestimmtheitsmaß  $B$  angewandt werden [Hartung], für lineare Regressionsmodelle ist es wie folgt definiert:

$$B = 1 - \frac{\sum_{t=1}^n (y_t - T_t - S_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad \text{mit} \quad \bar{y} = \sum_{t=1}^n y_t \quad (2.23)$$

Es drückt den Anteil der durch das Modell erklärten Varianz aus. Im Falle des additiven Modells wird anstelle von  $y_t$  die Box-Cox-transformierte Zeitreihe  $y(\lambda)$  gesetzt,  $T_t$  und  $S_t$  sind dann Trend und Saisonalität der Box-Cox-transformierten Zeitreihe. Für die betrachteten Zeitreihen ist das Bestimmtheitsmaß

	Bestimmtheitsmaß $B$	
	additives Modell mit Box-Cox-Tr.	standardisierendes Modell
max. Tagestemp.	0.72	0.72
DTR	0.17	0.16
NTR	0.001	0.002
min. Nachttemp.	0.79	0.79
Globalstr.	0.72	0.64

Das Bestimmtheitsmaß ist sehr niedrig, vor allem bei den beiden Temperaturunterschieden.

## 3 ARMA-Prozesse

Möchte man die Temperatur vom nächsten Tag vorhersagen, so kann man in Anlehnung an das vorige Kapitel die aus den vergangenen Jahren ermittelte durchschnittliche Temperatur für den betreffenden Tag vorhersagen. Sinnvoll ist aber auch die Temperatur der letzten Tage in die Prognose miteinzubeziehen. Denn es zeigt sich, dass die Residuen oder Resttemperaturen, die nicht durch Trend und Saisonalität erklärt werden, nicht paarweise unabhängig sind. Sie sind korreliert und die Korrelation hängt nur vom Zeitabstand der Beobachtungen ab. Man spricht von sogenannten autoregressiven Prozessen (AR).

Die moving average Prozesse (MA) beziehen sich hingegen auf einen Fehlerprozess. Man sollte also bei der Prognose zukünftiger Zeireihenwerte die Prognosefehler der Vergangenheit miteinbeziehen. Möglich ist natürlich auch eine Kombination der beiden Prozesse, ein sogenannter ARMA-Prozess.

Aus den Autokorrelationen und partiellen Autokorrelationen zu verschiedenen Zeitabständen (lags) können erste Vermutungen über das Prozessmodell (AR, MA oder ARMA) und dessen Ordnung (wieviele vergangene Beobachtungen oder Fehler miteinbezogen werden sollen) ausgedrückt werden. Außerdem muss auch noch bestimmt werden, wie die Beobachtungen und Fehler gewichtet werden sollen (Parameterschätzung).

Für die Modellidentifikation werden zwei Methoden behandelt: die graphische Methode mittels Autokorrelations- und partieller Autokorrelationsfunktion und die Identifikation mit sogenannten Informationskriterien. Die dazugehörigen Prozessparameter werden mit der Maximum-Likelihood-Schätzung numerisch ermittelt. Für die Berechnung der Startwerte wird die Hannan-Rissanen Methode vorgestellt.

Als Prognosemethode für Zeitreihen wird der Innovationen-Algorithmus beschrieben.

Die ARMA-Modelle werden auf stochastische Prozesse angewandt, deshalb müssen die Zeitreihen zunächst von Trend und Saisonalität bereinigt werden bevor die Methoden von Box und Jenkins verwendet werden können. Dies kann entweder mit dem Komponentenmodell aus dem vorigen Kapitel erfolgen oder man beschreibt die Zeitreihe durch ein SARIMA-Modell.

Falls nicht anders erwähnt wird den Darstellungen [Brockwell, Davis. 1991] und [Brockwell, Davis. 2002] gefolgt.

Im zweiten Teil werden dann die fünf Zeitreihen mit dem Verfahren von Box-Jenkins weiter untersucht.

### 3.1 Grundlagen-Definitionen

Der Grundgedanke von Box und Jenkins war die Zeitreihe als einen stochastischen Prozess zu betrachten.

**Definition 3.** Unter einem **stochastischen Prozess** ist eine Menge von Zufallsvariablen  $\{X_t | t \in T\}$  zu verstehen, die von einer Indexmenge  $T$  abhängt, welche als Parameterraum des Prozesses bezeichnet wird. In der Zeitreihenanalyse bedeutet  $T$  immer eine Menge von Zeitpunkten.

Um die Theorie von Box-Jenkins zu verwenden muss sichergestellt werden, dass die zu untersuchenden Zeitreihen, bzw. die irreguläre Komponente stationär ist.

**Definition 4.** Sei  $\{X_t\}$  eine Zeitreihe mit  $E(X_t^2) < \infty$ . Die **Erwartungswertfunktion** von  $\{X_t\}$  ist

$$\mu_X(t) = E(X_t) \quad (3.1)$$

Die **Kovarianzfunktion** von  $\{X_t\}$  ist

$$\gamma_X(r, s) = Cov(X_r, X_s) = E[(X_r - \mu_X(r))(X_s - \mu_X(s))] \quad (3.2)$$

für alle  $r$  und  $s$ .

**Definition 5.** Die Zeitreihe  $\{X_t\}$  ist (**schwach**) **stationär**, wenn

1.  $\mu_X(t)$  ist unabhängig von der Zeit  $t$  und
2.  $\gamma_X(t+h, t)$  ist unabhängig von  $t$  für alle  $h$

**Bemerkung 2.** Eine Zeitreihe ist streng stationär, falls  $(X_1, \dots, X_n)$  und  $(X_{1+h}, \dots, X_{n+h})$  für alle  $h$  und alle  $n > 0$  dieselbe Verteilung haben.

Im folgenden wird für stationär immer schwach stationär gemeint.

**Definition 6.** Sei  $\{X_t\}$  eine stationäre Zeitreihe. Die **Autokovarianzfunktion** von  $\{X_t\}$  zum lag  $h$  ist

$$\gamma_X(h) = Cov(X_{t+h}, X_t). \quad (3.3)$$

Dabei ist  $\gamma_X(-h) = \gamma_X(h)$ . Die **Autokorrelationsfunktion** von  $\{X_t\}$  zum lag  $h$  ist

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = Corr(X_{t+h}, X_t). \quad (3.4)$$

**Definition 7.** Der stationäre Prozess  $\{Z_t\}$  ist ein **weißes Rauschen**, oder white noise, mit Erwartungswert 0 und Varianz  $\sigma^2$

$$\{Z_t\} \sim WN(0, \sigma^2) \quad (3.5)$$

falls gilt:

$$\gamma_Z(h) = \begin{cases} \sigma^2, & \text{wenn } h = 0 \\ 0, & \text{wenn } h \neq 0 \end{cases} \quad (3.6)$$

Falls die Zufallsvariablen  $Z_t$  paarweise unabhängig und identisch verteilt sind, mit Erwartungswert 0 und Varianz  $\sigma^2$ , dann verwendet man die Notation

$$\{Z_t\} \sim IID(0, \sigma^2) \quad (3.7)$$

**Definition 8.** Der stochastische Prozess  $\{X_t, t = 0, \pm 1, \pm 2, \dots\}$  ist ein **ARMA(p,q)-Prozess**, falls  $\{X_t\}$  stationär ist und für jedes  $t$  gilt

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q} \quad (3.8)$$

wobei  $\{Z_t\} \sim WN(0, \sigma^2)$ .

Ein stochastischer Prozess  $\{X_t\}$  ist ein ARMA(p,q)-Prozess mit Erwartungswert  $\mu$  falls  $\{X_t - \mu\}$  ein ARMA(p,q)-Prozess ist.

Deshalb wird in der weiteren Ausführung o.B.d.A von einem stationären Prozess mit Erwartungswert 0 ausgegangen.

**Bemerkung 3.** Falls  $q = 0$  ist, spricht man von einem reinen autoregressiven Prozess und verwendet die Notation AR(p). Ist hingegen  $p = 0$ , so schreibt man MA(q), dies entspricht einem reinen moving-average Prozess.

**Beispiel 1.** Beim AR(2)-Prozess

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t$$

wird  $X_t$  durch ein gewogenes Mittel seiner zwei Vorgänger  $X_{t-1}$  und  $X_{t-2}$  und einen zufälligen Rest erklärt. Die meisten meteorologischen Zeitreihen sind ein AR(1) oder AR(2)-Modell [Wilks].

**Beispiel 2.** Ein MA(2)-Prozess

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$$

kann als gewogenen Durchschnitt der Ordnung  $q = 2$  eines weißen Rauschens oder Fehlerprozesse interpretiert werden.

**Bemerkung 4.** Die Gleichung (3.8) kann in kompakter Form angegeben werden

$$\phi(B)X_t = \theta(B)Z_t \quad (3.9)$$

wobei  $\phi$  und  $\theta$  Polynome vom Grad  $p$  bzw.  $q$  sind

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \quad (3.10)$$

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q \quad (3.11)$$

und  $B$  ist der Backward Shift Operator ist

$$B^j X_t = X_{t-j}, \quad j = 0, \pm 1, \pm 2, \dots \quad (3.12)$$

Die Polynome  $\phi$  und  $\theta$  werden als **autoregressives** und **moving average Polynom** bezeichnet.

### 3.2 Stationäre und invertierbare Arma-Prozesse

Die in diesem Abschnitt behandelten Sätze und Beweise wurden aus [Box, Jenkins] entnommen.

**Satz 2** (Stationaritätsbedingungen für AR(p)-Prozesse). Ein AR(p)-Prozess

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + Z_t$$

ist **stationär** oder **kausal**, dh. es existieren Konstanten  $\{\psi_j\}$ , sodass  $\sum_{j=0}^{\infty} |\{\psi_j\}| < \infty$  und

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

falls die Nullstellen des autoregressiven Polynoms  $\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$  außerhalb des Einheitskreises liegen.

*Beweis.* Es wird zunächst ein AR(1)-Prozess betrachtet

$$(1 - \phi_1 B)X_t = Z_t$$

Dieser kann geschrieben werden als

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + Z_t = \phi_1(\phi_1 X_{t-2} + Z_t) + Z_t = \dots = \sum_{j=0}^{\infty} \phi_1^j Z_{t-j} = \sum_{j=0}^{\infty} \phi_1^j B^j Z_t \\ X_t &= (1 - \phi_1 B)^{-1} Z_t \end{aligned}$$

Damit  $\psi(B) = \sum_{j=0}^{\infty} \phi_1^j B^j$  für  $|B| \leq 1$  konvergiert, muss  $|\phi_1| < 1$  sein. Äquivalent dazu muss die Nullstellen von  $\phi(B) = (1 - \phi_1 B)$  größer als 1 sein. Für den allgemeinen Fall eines AR(p)-Prozesses werden die Nullstellen des Polynoms berechnet.

$$\phi(B) = (1 - G_1 B)(1 - G_2 B) \dots (1 - G_p B)$$

Mittels Partialbruchzerlegung erhält man

$$X_t = \sum_{i=1}^p \frac{K_i}{(1 - G_i B)} Z_t$$

Der AR(p)-Prozess ist also stationär, wenn die  $|G_i| < 1$  sind. □

**Satz 3** (Invertierbarkeitsbedingungen für MA(q)-Prozesse). Ein MA(q)-Prozesse

$$X_t = (1 + \theta_1 B + \dots + \theta_q B^q) Z_t$$

ist **invertierbar**, dh. es existiert eine Folge  $\{\pi_j\}$  mit  $\sum_{j=0}^{\infty} |\pi_j| < \infty$  und

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \text{ für alle } t$$

falls die Nullstellen des moving average Polynoms  $\phi(z) = 1 + \theta_1 z + \dots + \theta_q z^q$  außerhalb des Einheitskreises liegen.

Für kausale ARMA(p,q)-Prozesse  $\phi(B)X_t = \theta(B)Z_t$  existieren also Konstanten  $\{\psi_j\}$  mit  $\sum_{j=0}^{\infty} |\psi_j| < \infty$ , sodass

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

ist. Die  $\{\psi_j\}$  können dann mittels Koeffizientenvergleich der Polynome

$$(1 - \phi_1 z - \dots - \phi_p z^p)(\psi_0 + \psi_1 z + \dots) = 1 + \theta_1 z + \dots + \theta_q z^q$$

ermittelt werden.

$$\begin{aligned} 1 &= \psi_0 \\ \theta_1 &= \psi_1 - \psi_0 \phi_1 \\ &\vdots \\ \theta_j &= \psi_j - \sum_{k=1}^p \phi_k \psi_{j-k} \end{aligned} \tag{3.13}$$

### 3.3 Schätzer für $\mu$ , $\gamma(h)$ und $\rho(h)$

**Satz 4.** Der **Schätzer für den Erwartungswert**  $\mu$  eines stationären Prozesses ist der Mittelwert:

$$\bar{X}_n = \frac{1}{n} (X_1 + X_2 + \dots + X_n) \tag{3.14}$$

**Satz 5.** Der **Schätzer für die Autokovarianz**  $\gamma(h)$  und **Autokorrelation**  $\rho(h)$  zum lag  $h$  eines stationären Prozesses sind:

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (X_t - \bar{X}_n)(X_{t+|h|} - \bar{X}_n) \tag{3.15}$$

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \tag{3.16}$$

**Bemerkung 5.** Der Schätzer für die Autokovarianz  $\hat{\gamma}(h)$  ist nicht erwartungstreu, dh. er besitzt einen bias. Der alternative Schätzer

$$\tilde{\gamma}(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} (X_t - \bar{X}_n)(X_{t+h} - \bar{X}_n)$$

hat zwar einen geringeren bias, dafür aber in der Regel eine größere Varianz [Stier].

**Bemerkung 6.** Die Stichprobenautokorrelationsfunktion  $\hat{\rho}_k = (\hat{\rho}(1), \dots, \hat{\rho}(k))'$  spielt eine wichtige Rolle bei der Auswahl von geeigneten ARMA-Modellen für stochastische Prozesse. Für große  $n$  ist  $\hat{\rho}_k$  annähernd normalverteilt

$$\hat{\rho}_k \approx N(\rho_k, n^{-1}W),$$

wobei  $\rho_k = (\rho(1), \dots, \rho(k))'$  und  $W$  die Kovarianzmatrix ist. Ihr  $(i, j)$ -ter Eintrag ist durch die Bartlett-Formel gegeben.

$$w_{ij} = \sum_{k=-\infty}^{\infty} \{ \rho(k+i)\rho(k+j) + \rho(k-i)\rho(k+j) + 2\rho(i)\rho(j)\rho^2(k) - \rho(i)\rho(k)\rho(k+j) - 2\rho(j)\rho(k)\rho(k+i) \} \quad (3.17)$$

Mit einfachen Umformungen erhält man

$$w_{ij} = \sum_{k=1}^{\infty} \{ \rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k) \} \cdot \{ \rho(k+j) + \rho(k-j) - 2\rho(j)\rho(k) \} \quad (3.18)$$

**Beispiel 3.** Falls  $\{X_t\} \sim IID(0, \sigma^2)$  ist  $\rho(h) = 0$  für  $|h| > 0$ . Die Einträge der Kovarianzmatrix  $W$  sind dann nach (3.18):

$$\begin{aligned} w_{ij} &= \sum_{k=1}^{\infty} \underbrace{\{ \rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k) \}}_{=0} \cdot \underbrace{\{ \rho(k+j) + \rho(k-j) - 2\rho(j)\rho(k) \}}_{=0} \\ &= \sum_{k=1}^{\infty} \rho(k-i)\rho(k-j) \\ w_{ij} &= \begin{cases} 1, & \text{wenn } i = j \\ 0, & \text{sonst} \end{cases} \end{aligned}$$

Für große Stichproben sind die geschätzten Autokorrelationen von  $\{X_t\} \sim IID(0, \sigma^2)$  annähernd normalverteilt mit einer Varianz von  $n^{-1}$ .



### 3.4 Prognose mit ARMA-Modellen

Der beste lineare Prädiktor (im Sinne von minimalem quadratischen Abstand) von  $X_{n+1}$  aufgrund der Beobachtungen  $\{X_1, \dots, X_n\}$  ist die Projektion von  $X_{n+1}$  auf den abgeschlossenen linearen Untervektorraum  $H_n := \bar{s}p\{X_1, \dots, X_n\}$ .

Der Einschritt-Prädiktor  $\hat{X}_{n+1}$  ist also:

$$\hat{X}_{n+1} = \begin{cases} 0, & \text{wenn } n = 0; \\ P_{\bar{s}p\{X_1, \dots, X_n\}} X_{n+1}, & \text{wenn } n > 1. \end{cases} \quad (3.19)$$

Da  $\hat{X}_{n+1} \in H_n$ , kann der Einschritt-Prädiktor als Linearkombination von  $X_1, \dots, X_n$  geschrieben werden.

Mit dem Gram-Schmidtschen Orthogonalisierungsverfahren kann  $H_n$  auch als

$$H_n := \bar{s}p\{X_1 - \hat{X}_1, X_2 - \hat{X}_2, \dots, X_n - \hat{X}_n\} \quad (3.20)$$

geschrieben werden, mit  $\hat{X}_1 = 0$ , sodass

$$\hat{X}_{n+1} = \sum_{j=1}^n \theta_{nj} (X_{n+1-j} - \hat{X}_{n+1-j}) \quad (3.21)$$

Die Koeffizienten  $\theta_{n1}, \dots, \theta_{nn}$  können mit Hilfe der Autokovarianzfunktion

$$\kappa(i, j) = E(X_i X_j) = \langle X_i, X_j \rangle \quad (3.22)$$

berechnet werden.

**Satz 6** (Der Innovationen Algorithmus). Falls der Erwartungswert von  $\{X_t\}$  null ist und die  $(n \times n)$  Matrix  $[\kappa(i, j)]_{i,j=1}^n$  nicht singular ist für alle  $n = 1, 2, \dots$ , dann ist der Einschritt-Prädiktor  $\hat{X}_{n+1}, n \geq 0$

$$\hat{X}_{n+1} = \begin{cases} 0, & \text{wenn } n = 0, \\ \sum_{j=1}^n \theta_{nj} (X_{n+1-j} - \hat{X}_{n+1-j}), & \text{wenn } n \geq 1, \end{cases} \quad (3.23)$$

Der mittlere quadratische Fehler  $v_n = \|X_{n+1} - \hat{X}_{n+1}\|^2, n \geq 1$  und die Koeffizienten  $\theta_{nj}$  können mit der folgenden Rekursion berechnet werden

$$\begin{aligned} v_0 &= \kappa(1, 1), \\ \theta_{n, n-k} &= v_k^{-1} \left( \kappa(n+1, k+1) - \sum_{j=0}^{k-1} \theta_{k, k-j} \theta_{n, n-j} v_j \right), \quad k = 0, 1, \dots, n-1, \\ v_n &= \kappa(n+1, n+1) - \sum_{j=0}^{n-1} \theta_{n, n-j}^2 v_j. \end{aligned} \quad (3.24)$$

Man löst die Rekursion in der Reihenfolge  $v_0, \theta_{11}, v_1, \theta_{22}, \theta_{21}, v_2, \dots$

*Beweis.*  $\{X_1 - \hat{X}_1, X_2 - \hat{X}_2, \dots, X_n - \hat{X}_n\}$  mit  $\hat{X}_1 := 0$  ist orthogonal. So kann  $\hat{X}_{n+1}$  folgendermaßen berechnet werden

$$\hat{X}_{n+1} = \sum_{j=1}^n \theta_{nj} (X_{n+1-j} - \hat{X}_{n+1-j})$$

Wird der Skalarprodukt von  $\hat{X}_{n+1}$  und  $X_{k+1} - \hat{X}_{k+1}$  gebildet, erhält man für  $0 \leq k < n$

$$\begin{aligned} \langle \hat{X}_{n+1}, X_{k+1} - \hat{X}_{k+1} \rangle &= \left\langle \sum_{j=1}^n \theta_{nj} (X_{n+1-j} - \hat{X}_{n+1-j}), X_{k+1} - \hat{X}_{k+1} \right\rangle \\ &= \langle \theta_{n,n-k} (X_{n+1-n+k} - \hat{X}_{n+1-n+k}), X_{k+1} - \hat{X}_{k+1} \rangle \\ &= \theta_{n,n-k} \|X_{k+1} - \hat{X}_{k+1}\|^2 = \theta_{n,n-k} v_k \end{aligned}$$

Da für  $k = 0, \dots, n-1$   $(X_{n+1} - \hat{X}_{n+1}) \perp (X_{k+1} - \hat{X}_{k+1})$  sind die Koeffizienten  $\theta_{n,n-k}$

$$\begin{aligned} \theta_{n,n-k} &= v_k^{-1} \langle X_{n+1}, X_{k+1} - \hat{X}_{k+1} \rangle \\ &= v_k^{-1} \left( \langle X_{n+1}, X_{k+1} \rangle - \langle X_{n+1}, \hat{X}_{k+1} \rangle \right) \\ &= v_k^{-1} \left( \kappa(n+1, k+1) - \langle X_{n+1}, \sum_{j=1}^k \theta_{kj} (X_{k+1-j} - \hat{X}_{k+1-j}) \rangle \right) \\ &= v_k^{-1} \left( \kappa(n+1, k+1) - \sum_{j=0}^{k-1} \theta_{k,k-j} \underbrace{\langle X_{n+1}, X_{j+1} - \hat{X}_{j+1} \rangle}_{v_j \theta_{n,n-j}} \right) \\ \theta_{n,n-k} &= v_k^{-1} \left( \kappa(n+1, k+1) - \sum_{j=0}^{k-1} \theta_{k,k-j} \theta_{n,n-j} v_j \right) \end{aligned}$$

Laut Projektionssatz ist

$$v_n = \|X_{n+1} - \hat{X}_{n+1}\|^2 = \|X_{n+1}\|^2 - \|\hat{X}_{n+1}\|^2 = \kappa(n+1, n+1) - \sum_{k=0}^{n-1} \theta_{n,n-k}^2 v_k$$

□

Dieser Satz kann auf die Prognose von ARMA-Prozessen angewandt werden. Dabei wird der ARMA-Prozess

$$\phi(B)X_t = \theta(B)Z_t, \quad \{Z_t\} \sim WN(0, \sigma^2), \quad (3.25)$$

leicht umgeformt zu

$$W_t = \begin{cases} \sigma^{-1}X_t, & t = 1, \dots, m, \\ \sigma^{-1}\phi(B)X_t, & t > m, \end{cases} \quad (3.26)$$

wobei  $m = \max(p, q)$  ist. Sei  $\theta_0 := 1$  und  $\theta_j := 0$  für  $j > q$ . Es wird angenommen, dass  $p \geq 1$  und  $q \geq 1$ , falls dies nicht zutrifft werden die entsprechenden Koeffizienten null gesetzt. Die Autokovarianzen  $\kappa(i, j) = E(W_i W_j)$  sind dann

1. für  $1 \leq i, j \leq m$

$$\kappa(i, j) = E(W_i W_j) = E(\sigma^{-1}X_i, \sigma^{-1}X_j) = \sigma^{-2}\gamma_X(i - j)$$

2. für  $i \leq m < j \leq 2m$

$$\begin{aligned} \kappa(i, j) &= E(W_i W_j) = E(\sigma^{-1}X_i, \sigma^{-1}\phi(B)X_j) \\ &= \sigma^{-2}E[X_i, (X_j - \phi_1 X_{j-1} - \dots - \phi_p X_{j-p})] \\ &= \sigma^{-2}E[\gamma(j - i) - \phi_1 \gamma(j - i - 1) - \dots - \phi_p \gamma(j - i - p)] \end{aligned}$$

3. für  $i, j > m$

$$\begin{aligned} \kappa(i, j) &= E(W_i W_j) = E[\sigma^{-1}\phi(B)X_i, \sigma^{-1}\phi(B)X_j] = \sigma^{-2}E[\theta(B)Z_i, \theta(B)Z_j] \\ &= \sigma^{-2}E[\theta_0 Z_i + \theta_1 Z_{i-1} + \dots + \theta_q Z_{i-q}, \theta_0 Z_j + \theta_1 Z_{j-1} + \dots + \theta_q Z_{j-q}] \\ &= \sigma^{-2}\sigma^2 \sum_{r=0}^q \theta_r \theta_{r+|i-j|} \end{aligned}$$

Zusammenfassend

$$\kappa(i, j) = \begin{cases} \sigma^{-2}\gamma_X(i - j), & 1 \leq i, j \leq m, \\ \sigma^{-2}[\gamma_X(i - j) - \sum_{r=1}^p \phi_r \gamma_X(r - |i - j|)], & \min(i, j) \leq m < \max(i, j) \leq 2m \\ \sum_{r=0}^q \theta_r \theta_{r+|i-j|}, & \min(i, j) > m \\ 0, & \text{sonst.} \end{cases} \quad (3.27)$$

Mit den Innovationen-Algorithmus können nun die Einschnitt-Prädiktoren berechnet werden

$$\hat{W}_{n+1} = \begin{cases} \sum_{j=1}^n \theta_{nj}(W_{n+1-j} - \hat{W}_{n+1-j}), & 1 \leq n < m, \\ \sum_{j=1}^q \theta_{nj}(W_{n+1-j} - \hat{W}_{n+1-j}), & n \geq m, \end{cases} \quad (3.28)$$

Die Koeffizienten  $\theta_{nj}$  und der mittlere quadratische Fehler  $r_n = E(W_{n+1} - \hat{W}_{n+1})^2$  können rekursiv aus (3.24) berechnet werden, wobei  $\hat{X}_1 = \hat{W}_1 = 0$  gewählt wird.

Aus der Linearität der Projektion folgt

$$\hat{W}_t = \begin{cases} \sigma^{-1}\hat{X}_t, & t = 1, \dots, m, \\ \sigma^{-1}[\hat{X}_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p}], & t > m \end{cases} \quad (3.29)$$

und

$$X_t - \hat{X}_t = \sigma(W_t - \hat{W}_t) \quad (3.30)$$

Somit sind die Einschnitt-Prädiktoren für ein ARMA-Prozess

$$\hat{X}_{n+1} = \begin{cases} \sum_{j=1}^n \theta_{nj}(X_{n+1-j} - \hat{X}_{n+1-j}), & 1 \leq n < m \\ \phi_1 X_n + \dots + \phi_p X_{n+1-p} + \sum_{j=1}^q \theta_{nj}(X_{n+1-j} - \hat{X}_{n+1-j}), & n \geq m \end{cases} \quad (3.31)$$

und

$$E(X_{n+1} - \hat{X}_{n+1})^2 = \sigma^2 E(W_{n+1} - \hat{W}_{n+1})^2 = \sigma^2 r_n \quad (3.32)$$

**Beispiel 4.** (Vorhersage eines ARMA(1,1)-Prozesses) Der kausale ARMA(1,1)-Prozess

$$X_t - 0.8X_{t-1} = Z_t + 0.5Z_{t-1} \quad Z_t = WN\{0, 1\} \quad (3.33)$$

wird als AR( $\infty$ )-Prozess dargestellt

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j} \quad Z_t = WN\{0, 1\} \quad (3.34)$$

Die  $\psi_j$  können mittels Koeffizientenvergleich der Polynome  $\phi(z) = 1 - \phi_1 z$  und  $\theta(z) = 1 + \theta_1 z$  ermittelt werden.

$$\begin{aligned} (1 + \theta_1 z) &= (1 - \phi_1 z)(\psi_0 + \psi_1 z + \psi_2 z^2 + \dots) \\ \psi_0 &= 1 \\ \psi_1 &= \theta_1 + \psi_0 \phi_1 = \theta_1 + \phi_1 = 0.5 + 0.8 = 1.3 \\ \psi_2 &= \theta_1 \phi_1 + \phi_1^2 = 0.5 \cdot 0.8 + 0.8^2 = 1.04 \\ &\dots \end{aligned}$$

Die Autokovarianzen des ARMA(1,1)-Prozesses zu den verschiedenen lags findet man indem man beide Seiten der Gleichung (3.33) mit  $X_{t-k}$  für  $k = 0, 1, \dots$  multipliziert und anschließend den Erwartungswert bildet.

$$\begin{aligned} X_t - 0.8X_{t-1} &= Z_t + 0.5Z_{t-1} \\ E[X_t X_t] - 0.8E[X_t X_{t-1}] &= E[X_t(Z_t + 0.5EZ_{t-1})] \\ &= E[(\psi_0 Z_t + \psi_1 Z_{t-1} + \dots)(Z_t + 0.5Z_{t-1})] \\ \gamma(0) - 0.8\gamma(1) &= \sigma^2(\psi_0 + 0.5\psi_1) = 1.65 \\ E[X_{t-1} X_t] - 0.8E[X_{t-1} X_{t-1}] &= E[X_{t-1}(Z_t + 0.5EZ_{t-1})] \\ &= E[(\psi_0 Z_{t-1} + \psi_1 Z_{t-2} + \dots)(Z_t + 0.5Z_{t-1})] \\ \gamma(1) - 0.8\gamma(0) &= 0.5\sigma^2\psi_0 = 0.5 \end{aligned}$$

Für den lag  $k \geq 2$  gilt

$$\begin{aligned} E[X_{t-k}X_t] - 0.8E[X_{t-k}X_{t-1}] &= E[X_{t-k}Z_t] + 0.5E[X_{t-k}Z_{t-1}] \\ \gamma(k) - 0.8\gamma(k-1) &= 0 \end{aligned}$$

Die Autokovarianzen  $\gamma(0)$  und  $\gamma(1)$  können nun mit einem Gleichungssystem gelöst werden

$$\begin{pmatrix} \gamma(0) \\ \gamma(1) \end{pmatrix} = \begin{pmatrix} 1 & -0.8 \\ -0.8 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1.65 \\ 0.5 \end{pmatrix} = \begin{pmatrix} 5.694444 \\ 5.055556 \end{pmatrix}$$

Die Einträge der Kovarianzmatrix des ARMA(1,1)-Prozesses mit  $m = \max(p, q) = 1$  sind laut (3.27)

$$\kappa(i, j) = \begin{cases} \sigma^{-2}\gamma(i-j), & i = j = 1, \\ \sigma^{-2}[\gamma(i-j) - \phi_1\gamma(1 - |i-j|)], & \min(i, j) = 1, \max(i, j) = 2 \\ \sum_{r=0}^1 \theta_r\theta_{r+|i-j|}, & i, j > 1 \\ 0, & \text{sonst.} \end{cases} \quad (3.35)$$

Setzt man nun die bekannten Größen ein, sind die Kovarianzeinträge

$$\begin{aligned} \kappa(1, 1) &= \sigma^{-2}\gamma(0) = 5.694444 \\ \kappa(1, 2) &= \kappa(2, 1) = \sigma^{-2}(\gamma(1) - \phi_1\gamma(0)) = 0.50 \\ \kappa(i, j) &= 1 + \theta_1^2 = 1.25 && \text{für } i = j > 1 \\ \kappa(i, j) &= \theta_0\theta_1 = \theta_1 = 0.5 && \text{für } |i - j| = 1, \min(i, j) > 1 \\ \kappa(i, j) &= 0 && \text{sonst} \end{aligned}$$

Aus dem Innovationen-Algorithmus (3.24) folgen dann die  $v_j$  und  $\theta_{nj}$

$$\begin{aligned} v_0 &= \kappa(1, 1) = 5.694444 \\ \theta_{11} &= v_0^{-1}(\kappa(2, 1)) = 0.08780488 \\ v_1 &= \kappa(2, 2) - \theta_{11}^2 v_0 = 1.206098 \\ &\quad \theta_{22} = v_0^{-1}(\kappa(3, 1)) = 0 \\ \theta_{21} &= v_1^{-1}(\kappa(3, 2) - \theta_{11}\theta_{22}v_0) = 0.41456 \\ v_2 &= \kappa(3, 3) - \theta_{22}^2 v_0 - \theta_{21}^2 v_1 = 1.04272 \\ &\quad \theta_{33} = v_0^{-1}\kappa(4, 1) = 0 \\ &\quad \theta_{32} = v_1^{-1}(\kappa(4, 2) - \theta_{11}\theta_{33}v_0) = 0 \\ \theta_{31} &= v_2^{-1}(\kappa(4, 3) - \theta_{22}\theta_{33}v_0 - \theta_{21}\theta_{32}) \\ &= v_2^{-1}\kappa(4, 3) = 0.5v_2^{-1} = 0.4795151 \\ v_3 &= \kappa(4, 4) - \theta_{33}^2 v_0 - \theta_{32}^2 v_1 - \theta_{31}^2 v_2 = 1.010242 \end{aligned}$$

Die Einschnitt-Prädiktoren sind laut (3.31), mit  $\hat{X}_1 = 0$

$$\begin{aligned}\hat{X}_{n+1} &= \phi_1 X_n + \theta_{n1}(X_n - \hat{X}_n) \\ &= 0.8X_n + \theta_{n1}(X_n - \hat{X}_n)\end{aligned}$$

und

$$E(X_{n+1} - \hat{X}_{n+1})^2 = v_n \tag{3.36}$$

**Bemerkung 7.** Anhand des Beispiels sieht man, dass  $\theta_{n1}$  gegen  $\theta_1$  und  $v_n$  gegen  $\sigma^2$  konvergieren. Dies gilt im Allgemeinen, falls  $\{X_t\}$  invertierbar ist, für  $n \rightarrow \infty$ ,

$$\begin{aligned}E(X_n - \hat{X}_n - Z_n)^2 &\rightarrow 0 \\ \theta_{nj} &\rightarrow \theta_j, \quad j = 1, \dots, q \\ r_n &\rightarrow 1\end{aligned}$$

Deshalb wird oft auch folgender Vorhersagealgorithmus (mit  $\hat{X}_1 = 0$ ) verwendet (vgl. [Stier] und [Hartung])

$$\hat{X}_{n+1} = \phi_1 X_n + \dots + \phi_p X_{n+1-p} + \sum_{j=1}^q \theta_j (X_{n+1-j} - \hat{X}_{n+1-j}) \tag{3.37}$$

### h-Schritt-Prognose von ARMA(p,q)-Prozessen

Die Mehrschrittprognose eines ARMA(p,q)-Prozesses ist analog zur Einschnittprognose (3.31). Die von den  $n$  Beobachtungen  $X_1, \dots, X_n$  ausgehende  $h$ -Schrittprognose  $P_n X_{n+h}$  wird rekursiv ( $P_n X_{n+1}, P_n X_{n+2}, P_n X_{n+3}, \dots$ ) berechnet, wobei die fehlenden Werte  $X_{n+1}, \dots, X_{n+h-1}$  durch die Prognosen  $P_n X_{n+1}, P_n X_{n+2}, \dots$  ersetzt werden. Die  $h$ -Schrittprognose  $P_n X_{n+h}$  wird also rekursiv mit

$$P_n X_{n+h} = \begin{cases} \sum_{j=h}^{n+h-1} \theta_{n+h-1,j} (X_{n+h-j} - \hat{X}_{n+h-j}), & 1 \leq h \leq m - n \\ \sum_{i=1}^p \phi_i P_n X_{n+h-i} + \sum_{j=h}^{n+h-1} \theta_{n+h-1,j} (X_{n+h-j} - \hat{X}_{n+h-j}), & h > m - n \end{cases}$$

berechnet.

## 3.5 Parameterschätzung für ARMA-Prozesse

Die Schätzung der Parameter  $\phi_1, \dots, \phi_p$  und  $\theta_1, \dots, \theta_q$  bei bekannter Ordnung  $p$  und  $q$  gliedert sich im Wesentlichen in zwei Schritten. Zunächst werden die Startwerte mittels Minimie-

nung der Summe der Quadratfehler berechnet und im zweiten Schritt wird eine Maximum-Likelihood-Schätzung durchgeführt.

### 3.5.1 Berechnung der Startwerte

Für die Berechnung der Startwerte existieren eine Reihe von Schätzverfahren. Am einfachsten lassen sich die Parameter von reinen AR-Prozessen angeben, hier verwendet man die sogenannten **Yule-Walker-Gleichungen**.

Für reine MA-Prozesse wird beispielsweise der **Innovationen-Algorithmus** verwendet, wobei die Autokorrelationen  $\kappa_i, j$  durch die empirischen Autokorrelationen ersetzt werden.

Bei gemischten ARMA(p,q)-Modellen wird hingegen der **Hannan-Rissanen Algorithmus** angewendet.

### 3.5.2 Hannan-Rissanen Algorithmus

Ein reines Autoregressiv-Modell  $AR(p)$  hat die Form einer linearen Regression mit den Koeffizienten  $\phi_1, \dots, \phi_p$ . Daher ist es naheliegend eine einfache Regressionsanalyse mit der Methode der kleinsten Quadrate anzuwenden.

Aber auch im Fall eines  $ARMA(p, q)$ -Prozesses, wo  $X_t$  nicht nur von  $X_{t-1}, \dots, X_{t-p}$  sondern auch von  $Z_{t-1}, \dots, Z_{t-q}$  abhängt, ist es möglich die Methode der kleinsten Quadrate anzuwenden. Hier werden die unbeobachteten  $Z_{t-1}, \dots, Z_{t-q}$  durch die geschätzten  $\hat{Z}_{t-1}, \dots, \hat{Z}_{t-q}$  ersetzt. Die Parameter  $\phi$  und  $\theta$  werden dann durch die einfache Regressionsanalyse mit der abhängigen Variablen  $X_t$  und den unabhängigen Variablen  $X_{t-1}, \dots, X_{t-p}, \hat{Z}_{t-1}, \dots, \hat{Z}_{t-q}$  bestimmt.

**1.Schritt.** Für die Zeitreihe wird ein  $AR(m)$ -Modell mit  $m > \max(p, q)$  gewählt. Zunächst werden die Parameter  $(\hat{\phi}_{m1}, \dots, \hat{\phi}_{mm})$  mittels Yule-Walker-Gleichungen geschätzt.

Für ein  $AR(m)$ -Modell  $X_t = \phi_{m1}X_{t-1} + \dots + \phi_{mm}X_{t-m} + Z_t$  ergeben sich folgende Kovarianzen zum lag 0.

$$\begin{aligned} X_t &= \phi_{m1}X_{t-1} + \dots + \phi_{mm}X_{t-m} + Z_t \\ X_t X_t &= \phi_{m1}X_{t-1}X_t + \dots + \phi_{mm}X_{t-m}X_t + X_t Z_t \\ E(X_t X_t) &= \phi_{m1}E(X_{t-1}X_t) + \dots + \phi_{mm}E(X_{t-m}X_t) + E(Z_t X_t) \\ \gamma(0) &= \sum_{j=1}^m \phi_{mj} \gamma(j) + \sigma^2 \end{aligned}$$

Denn  $E(Z_t X_t) = E[Z_t \sum_{j=0}^{\infty} \psi_j Z_{t-j}] = \psi_0 E[Z_t Z_t] = \sigma^2$ , da  $\{Z_t\}$  weißes Rauschen und  $\psi_0 = 1$  (vgl.(3.13)) ist .

Die Kovarianzen zum lag  $k > 0$  sind

$$\begin{aligned}
 X_t &= \phi_{m1}X_{t-1} + \dots + \phi_{mm}X_{t-m} + Z_t \\
 X_t X_{t-k} &= \phi_{m1}X_{t-1}X_{t-k} + \dots + \phi_{mm}X_{t-m}X_{t-k} + Z_t X_{t-k} \\
 E(X_t X_{t-k}) &= \phi_{m1}E(X_{t-1}X_{t-k}) + \dots + \phi_{mm}E(X_{t-m}X_{t-k}) + E\left(Z_t \sum_{j=0}^{\infty} \psi_j Z_{t-k-j}\right) \\
 \gamma(k) &= \sum_{j=1}^m \phi_{mj} \gamma(k-j)
 \end{aligned}$$

Somit sind die Autokorrelationen zum lag  $k$

$$\begin{aligned}
 \rho(0) &= 1 \\
 \rho(k) &= \sum_{j=1}^m \phi_{mj} \rho(k-j) \quad \text{für } k > 0
 \end{aligned}$$

Und die empirischen Yule-Walker-Gleichungen

$$\begin{aligned}
 \hat{\rho}(0) &= 1 \\
 \hat{\rho}(k) &= \sum_{j=1}^m \phi_{mj} \hat{\rho}(k-j)
 \end{aligned}$$

Um die Parameter  $\phi_{m1}, \dots, \phi_{mm}$  zu schätzen wird das Gleichungssystem mit den empirisch ermittelten Autokorrelationen

$$\begin{pmatrix} \hat{\rho}(1) \\ \hat{\rho}(2) \\ \vdots \\ \hat{\rho}(m) \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & \hat{\rho}(1) & \hat{\rho}(2) & \dots & \hat{\rho}(m-1) \\ \hat{\rho}(1) & 1 & \hat{\rho}(1) & \dots & \hat{\rho}(m-2) \\ \vdots & \vdots & \vdots & & \vdots \\ \hat{\rho}(m-1) & \hat{\rho}(m-2) & \hat{\rho}(m-3) & \dots & 1 \end{pmatrix}}_{R(m)} \begin{pmatrix} \phi_{m1} \\ \phi_{m2} \\ \vdots \\ \phi_{mm} \end{pmatrix}$$

gelöst. Die Matrix  $R(m)$  invertierbar, da für stationäre Prozesse gilt  $\rho(h) \rightarrow 0$  für  $h \rightarrow \infty$ . Somit hat die Matrix vollen Rang.

$$\begin{pmatrix} \phi_{m1} \\ \phi_{m2} \\ \vdots \\ \phi_{mm} \end{pmatrix} = R^{-1}(m) \begin{pmatrix} \hat{\rho}(1) \\ \hat{\rho}(2) \\ \vdots \\ \hat{\rho}(m) \end{pmatrix}$$



Nun können die Residuen

$$\hat{Z}_t = X_t - \hat{\phi}_{m1}X_{t-1} - \dots - \hat{\phi}_{mm}X_{t-m}, \quad t = m + 1, \dots, n \quad (3.38)$$

gebildet werden.

**2.Schritt.** Nachdem die Residuen  $\hat{Z}_t$  für  $t = m + 1, \dots, n$  berechnet wurden, wird der Parametervektor  $\beta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)'$  mittels Regressionsanalyse mit der abhängigen Variablen  $X_t$  und den unabhängigen Variablen  $(X_{t-1}, \dots, X_{t-p}, \hat{Z}_{t-1}, \dots, \hat{Z}_{t-q})$  für  $t = m + 1, \dots, n$  geschätzt. Dabei wird die Summe der Quadratfehler minimiert:

$$S(\beta) = \sum_{t=m+1+q}^n (X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} - \theta_1 \hat{Z}_{t-1} - \dots - \theta_q \hat{Z}_{t-q})^2 \quad (3.39)$$

Dies ergibt den Hannan-Rissanen-Schätzer

$$\hat{\beta} = (Z'Z)^{-1}Z'X_n \quad (3.40)$$

mit  $X_n = (X_{m+1+q}, \dots, X_n)'$  und  $Z$  ist die  $(n - m - q) \times (p + q)$  Matrix

$$Z = \begin{bmatrix} X_{m+q} & X_{m+q-1} & \cdots & X_{m+q+1-p} & \hat{Z}_{m+q} & \hat{Z}_{m+q-1} & \cdots & \hat{Z}_{m+1} \\ X_{m+q+1} & X_{m+q} & \cdots & X_{m+q+2-p} & \hat{Z}_{m+q+1} & \hat{Z}_{m+q} & \cdots & \hat{Z}_{m+2} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ X_{n-1} & X_{n-2} & \cdots & X_{n-p} & \hat{Z}_{n-1} & \hat{Z}_{n-2} & \cdots & \hat{Z}_{n-q} \end{bmatrix}$$

(Falls  $p = 0$ , enthält  $Z$  nur die letzten  $q$  Spalten.)

Der Hannan-Rissanen-Schätzer für die Varianz des weißen Rauschens ist

$$\hat{\sigma}_{HR}^2 = \frac{S(\hat{\beta})}{n - m - q}$$

**Bemerkung 8.** Hannan und Rissanen haben einen dritten Schritt zur Verbesserung des Schätzers vorgesehen. In den meisten Software-Programmen wird aber Maximum-Likelihood für den entgeltigen Schätzer verwendet.

### 3.5.3 Maximum-Likelihood-Schätzer

Es wird angenommen, dass die Zeitreihe  $\{X_t\}$  normalverteilt mit Erwartungswert 0 und Kovarianzmatrix  $\Gamma_n = E(\mathbf{X}_n \mathbf{X}_n')$  ist, wobei  $\mathbf{X}_n = (X_1, \dots, X_n)'$  ist und  $\Gamma_n$  als nicht singulär angenommen wird.

Die Likelihoodfunktion von  $\mathbf{X}_n$  ist

$$L(\Gamma_n) = (2\pi)^{-n/2} (\det \Gamma_n)^{-1/2} \exp\left(-\frac{1}{2} \mathbf{X}_n' \Gamma_n^{-1} \mathbf{X}_n\right) \quad (3.41)$$

Aus den Innovationen-Algorithmus wissen wir, dass

$$\hat{\mathbf{X}}_n = B(\mathbf{X}_n - \hat{\mathbf{X}}_n) \quad (3.42)$$

wobei  $B$  die  $n \times n$ -Matrix

$$B = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ \theta_{1,1} & 0 & 0 & \dots & 0 \\ \theta_{2,2} & \theta_{2,1} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \theta_{n-1,n-1} & \theta_{n-1,n-2} & \theta_{n-1,n-3} & \dots & 0 \end{pmatrix} \quad (3.43)$$

ist. Es gilt

$$\begin{aligned} \mathbf{X}_n &= \mathbf{X}_n - \hat{\mathbf{X}}_n + \hat{\mathbf{X}}_n = \mathbf{X}_n - \hat{\mathbf{X}}_n + B(\mathbf{X}_n - \hat{\mathbf{X}}_n) \\ &= (B + I)(\mathbf{X}_n - \hat{\mathbf{X}}_n) = C(\mathbf{X}_n - \hat{\mathbf{X}}_n) \end{aligned} \quad (3.44)$$

Die  $n \times n$ - Diagonalmatrix

$$D = \text{diag}(v_0, v_1, \dots, v_{n-1}) \quad (3.45)$$

ist die Kovarianzmatrix von  $(\mathbf{X}_n - \hat{\mathbf{X}}_n)$ , daher gilt

$$\begin{aligned} \Gamma_n &= E[\mathbf{X}_n \mathbf{X}_n'] = E[C(\mathbf{X}_n - \hat{\mathbf{X}}_n)(\mathbf{X}_n - \hat{\mathbf{X}}_n)'C'] \\ &= CE[(\mathbf{X}_n - \hat{\mathbf{X}}_n)(\mathbf{X}_n - \hat{\mathbf{X}}_n)']C' = CDC' \end{aligned} \quad (3.46)$$

Aus (3.44) und (3.46) folgt

$$\begin{aligned} \mathbf{X}_n' \Gamma_n^{-1} \mathbf{X}_n &= (\mathbf{X}_n - \hat{\mathbf{X}}_n)' C' (CDC')^{-1} C (\mathbf{X}_n - \hat{\mathbf{X}}_n) \\ &= (\mathbf{X}_n - \hat{\mathbf{X}}_n)' D^{-1} (\mathbf{X}_n - \hat{\mathbf{X}}_n) = \sum_{j=1}^n (X_j - \hat{X}_j)^2 / v_{j-1} \end{aligned} \quad (3.47)$$

$C$  ist eine untere Dreiecksmatrix mit Diagonaleinträgen 1 und  $D$  eine Diagonalmatrix. Die Determinante ist deshalb

$$\det \Gamma_n = (\det C)^2 (\det D) = v_0 v_1 \dots v_{n-1} \quad (3.48)$$

Aus diesen Gründen kann die Likelihoodfunktion (3.41) als

$$L(\Gamma_n) = (2\pi)^{-n/2} (v_0 v_1 \dots v_{n-1})^{-1/2} \exp \left( -\frac{1}{2} \sum_{j=1}^n (X_j - \hat{X}_j)^2 / v_{j-1} \right) \quad (3.49)$$

geschrieben werden.

Die Einschnitt-Prädiktoren  $\hat{X}_j$  und die dazugehörigen mittleren Quadratfehler  $v_{j-1}$  können mit Hilfe des Innovationen-Algorithmus bestimmt werden. Da  $v_{j-1} = \sigma^2 r_{j-1}$  gilt

$$L(\phi, \theta, \sigma) = (2\pi\sigma^2)^{-n/2} (r_0 r_1 \cdots r_{n-1})^{-1/2} \exp \left( -\frac{1}{2\sigma^2} \sum_{j=1}^n (X_j - \hat{X}_j)^2 / r_{j-1} \right) \quad (3.50)$$

### 3.6 Identifikation des Modells

Die Problemstellung der Modellidentifikation ist die Bestimmung der Ordnungen  $p$  und  $q$  des ARMA-Prozesses. Laut [Box, Jenkins] können die Ordnungen anhand der Autokorrelation- und der partiellen Autokorrelationsfunktion zum lag  $h$  geschätzt werden. Dabei werden beispielsweise jene Autokorrelationen als signifikant zur 5%-Irrtumswahrscheinlichkeit angesehen, welche außerhalb des Intervalls  $[-1.96/n^{-1/2}, 1.96/n^{-1/2}]$  liegen. Dies ist nämlich der Konfidenzintervall für  $\{X_t\} \sim IID(0, \sigma^2)$ . Es ist nicht ratsam große  $p$  und  $q$  zu wählen, da mit ansteigender Anzahl von Parametern die Vorhersagen zunehmend ungenauer werden.

#### 3.6.1 Identifikation von $p$ und $q$ anhand der Autokorrelation und der partiellen Autokorrelation

Die Identifikation der Ordnungen mit Hilfe der Autokorrelationsfunktion und der partiellen Autokorrelationsfunktion wird in zwei Beispielen veranschaulicht.

**Beispiel 5.** Es sei bekannt, dass die Zeitreihe  $X_t$  ein MA(2)-Prozess

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$$

ist. Es wird nun untersucht welche die Autokovarianzen, bzw. Autokorrelationen zu den lags 0 bis 3 sind.

- $\gamma(0)$  und  $\rho(0)$

$$\begin{aligned} E[X_t X_t] &= E[(Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2})(Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2})] \\ &= E[Z_t Z_t] + E[\theta_1 Z_{t-1} \theta_1 Z_{t-1}] + E[\theta_2 Z_{t-2} \theta_2 Z_{t-2}] \\ \gamma(0) &= \sigma^2 (1 + \theta_1^2 + \theta_2^2) \\ \rho(0) &= 1 \end{aligned}$$

- $\gamma(1)$  und  $\rho(1)$

$$\begin{aligned} E[X_{t-1}X_t] &= E[(Z_{t-1} + \theta_1 Z_{t-2} + \theta_2 Z_{t-3})(Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2})] \\ &= E[Z_{t-1}\theta_1 Z_{t-1}] + E[\theta_1 Z_{t-2}\theta_2 Z_{t-2}] \\ \gamma(1) &= \sigma^2(\theta_1 + \theta_1\theta_2) \\ \rho(1) &= \frac{\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2} \end{aligned}$$

- $\gamma(2)$  und  $\rho(2)$

$$\begin{aligned} E[X_{t-2}X_t] &= E[(Z_{t-2} + \theta_1 Z_{t-3} + \theta_2 Z_{t-4})(Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2})] \\ &= E[Z_{t-2}\theta_2 Z_{t-2}] \\ \gamma(2) &= \sigma^2\theta_2 \\ \rho(2) &= \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2} \end{aligned}$$

- Autokovarianz zum lag 3  $\gamma(3)$

$$\begin{aligned} E[X_{t-3}X_t] &= E[(Z_{t-3} + \theta_1 Z_{t-4} + \theta_2 Z_{t-5})(Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2})] \\ \gamma(3) &= \rho(3) = 0 \end{aligned}$$

Schätzt man nun die Autokorrelationen der Zeitreihe  $X_t$

$$\begin{aligned} \hat{\gamma}(h) &= \frac{1}{n} \sum_{t=1}^{n-|h|} (X_t - \bar{X}_n)(X_{t+|h|} - \bar{X}_n) \\ \hat{\rho}(h) &= \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \end{aligned}$$

und stellt fest, dass die Autokorrelationen nur bis zum lag 2 signifikant von Null verschieden sind, so identifiziert man die Zeitreihe als MA(2)-Prozess. Signifikant sind jene Autokorrelationen, die außerhalb des Intervalls  $[-1.96/n^{-1/2}, 1.96/n^{-1/2}]$  liegen. Die geschätzte Autokorrelationsfunktion von  $X_t = Z_t - 0.4Z_{t-1} + 0.7Z_{t-2}$  ist in Abbildung 3.1 dargestellt.

**Beispiel 6.** Für den AR(1)-Prozess

$$X_t = \phi_1 X_{t-1} + Z_t$$

sind die Autokovarianzen bzw. Autokorrelationen zu den verschiedenen lags

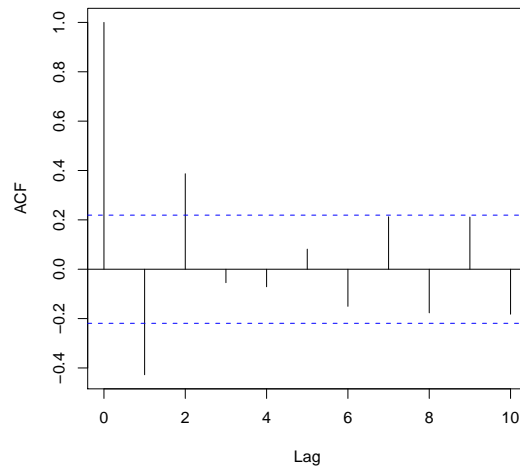


Abbildung 3.1: Autokorrelationsfunktion von  $X_t = Z_t - 0.4Z_{t-1} + 0.7Z_{t-2}$

- $\gamma(0)$  und  $\rho(0)$

$$\begin{aligned}
 E(X_t X_t) &= E[(\phi_1 X_{t-1} + Z_t)(\phi_1 X_{t-1} + Z_t)] \\
 &= E[(\phi_1 X_{t-1})(\phi_1 X_{t-1})] + E[Z_t Z_t] \\
 \gamma(0) &= \phi_1^2 \gamma(0) + \sigma^2 \\
 &= \frac{\sigma^2}{1 - \phi_1^2} \\
 \rho(0) &= 1
 \end{aligned}$$

- $\gamma(1)$  und  $\rho(1)$

$$\begin{aligned}
 E(X_{t-1} X_t) &= E[X_{t-1}(\phi_1 X_{t-1} + Z_t)] \\
 &= E[(\phi_1 X_{t-1} X_{t-1})] \\
 \gamma(1) &= \phi_1 \gamma(0) \\
 \rho(1) &= \phi_1
 \end{aligned}$$

- $\gamma(k)$  und  $\rho(k)$

$$\begin{aligned}
 E(X_{t-k} X_t) &= E[X_{t-k}(\phi_1 X_{t-1} + Z_t)] \\
 &= E[(\phi_1 X_{t-k} X_{t-1})] \\
 \gamma(k) &= \phi_1 \gamma(k-1) = \dots = \phi_1^k \gamma(0) \\
 \rho(k) &= \phi_1^k
 \end{aligned}$$

Die geschätzte Autokorrelationsfunktion von  $X_t = -0.8X_{t-1} + Z_t$  ist in Abbildung 3.2 dargestellt.

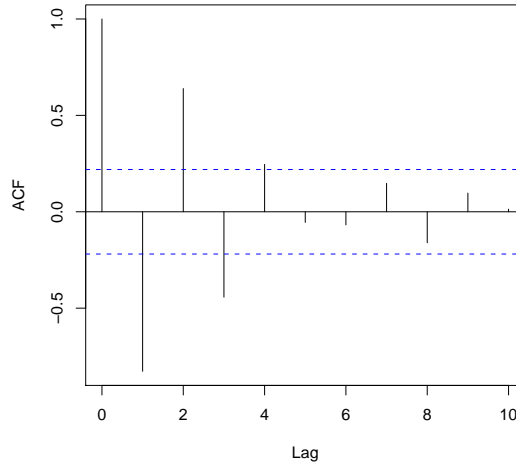


Abbildung 3.2: Autokorrelationsfunktion von  $X_t = -0.8X_{t-1} + Z_t$

Aussagekräftiger für einen AR(p)-Prozess ist allerdings die partielle Autokorrelation  $\alpha(k)$  zum lag  $k$ , diese beschreibt die Korrelation zwischen  $X_1$  und  $X_{k+1}$ , unter Ausschluss des Einflusses von  $X_2, \dots, X_k$ .

**Definition 9.** Die **partielle Autokorrelationsfunktion**  $\alpha(\cdot)$  einer stationären Zeitreihe mit Erwartungswert 0 ist definiert als

$$\alpha(1) = \text{Corr}(X_2, X_1) \tag{3.51}$$

$$\alpha(k) = \text{Corr}(X_{k+1} - P_{\{X_2, \dots, X_k\}}X_{k+1}, X_1 - P_{\{X_2, \dots, X_k\}}X_1) \quad k \geq 2 \tag{3.52}$$

wobei  $P_{\{X_2, \dots, X_k\}}$  die Projektion ist.

Die partielle Korrelation ist also die Korrelation der Residuen zweier linearer Regressionen mit den unabhängigen Variablen  $X_2, \dots, X_k$  und der abhängigen Variable  $X_{k+1}$  bzw.  $X_1$ .

Mit Durbin-Levinson kann gezeigt werden, dass die partielle Autokorrelation leicht berechnet werden kann. Der beste lineare Prädiktor für  $X_{n+1}$  ist

$$\hat{X}_{n+1} = P_{\{X_2, \dots, X_n\}}X_{n+1} + a(X_1 - P_{\{X_2, \dots, X_n\}}X_1) \tag{3.53}$$

wobei  $a$  die Projektion von  $X_{n+1}$  auf  $X_1 - P_{\{X_2, \dots, X_n\}}X_1$  ist

$$a = \frac{\langle X_{n+1}, X_1 - P_{\{X_2, \dots, X_n\}}X_1 \rangle}{\|X_1 - P_{\{X_2, \dots, X_n\}}X_1\|^2} \quad (3.54)$$

Da für die Autokovarianz  $\gamma(h) = \gamma(-h)$  gilt, ist

$$\begin{aligned} P_{\{X_2, \dots, X_n\}}X_1 &= \sum_{j=1}^{n-1} \phi_{n-1,j} X_{j+1} \\ P_{\{X_2, \dots, X_n\}}X_{n+1} &= \sum_{j=1}^{n-1} \phi_{n-1,j} X_{n+1-j} \end{aligned} \quad (3.55)$$

Setzt man diese zwei Gleichungen in (3.53) ein, erhält man

$$\hat{X}_{n+1} = aX_1 + \sum_{j=1}^{n-1} (\phi_{n-1,j} - a\phi_{n-1,n-j})X_{n+1-j} \quad (3.56)$$

Vergleicht man diese Koeffizienten mit der Darstellung

$$\hat{X}_{n+1} = P_{\{X_1, X_2, \dots, X_n\}}X_{n+1} = \sum_{j=1}^n \phi_{n,j} X_{n+1-j} \quad (3.57)$$

so ist

$$\phi_{nn} = a \quad (3.58)$$

Nun kann gezeigt werden, dass

$$\phi_{nn} = \text{Corr}(X_{n+1} - P_{\{X_2, \dots, X_n\}}X_{n+1}, X_1 - P_{\{X_2, \dots, X_n\}}X_1) \quad (3.59)$$

Da  $\phi_{nn} = a$  und  $P_{\{X_2, \dots, X_n\}}X_{n+1} \perp (X_1 - P_{\{X_2, \dots, X_n\}}X_1)$  ist, folgt

$$\begin{aligned} \phi_{nn} &= \frac{\langle X_{n+1}, X_1 - P_{\{X_2, \dots, X_n\}}X_1 \rangle}{\|X_1 - P_{\{X_2, \dots, X_n\}}X_1\|^2} \\ &= \frac{\langle X_{n+1} - P_{\{X_2, \dots, X_n\}}X_{n+1}, X_1 - P_{\{X_2, \dots, X_n\}}X_1 \rangle}{\|X_1 - P_{\{X_2, \dots, X_n\}}X_1\|^2} \end{aligned}$$

Aus (3.55) und (3.55) folgt  $\|X_1 - P_{\{X_2, \dots, X_n\}}X_1\|^2 = \|X_{n+1} - P_{\{X_2, \dots, X_n\}}X_{n+1}\|^2$  und somit

$$\begin{aligned} \phi_{nn} &= \frac{\langle X_{n+1} - P_{\{X_2, \dots, X_n\}}X_{n+1}, X_1 - P_{\{X_2, \dots, X_n\}}X_1 \rangle}{\|X_{n+1} - P_{\{X_2, \dots, X_n\}}X_{n+1}\|^2} \\ &= \text{Corr}(X_{n+1} - P_{\{X_2, \dots, X_n\}}X_{n+1}, X_1 - P_{\{X_2, \dots, X_n\}}X_1) \end{aligned}$$

Den partiellen Autokorrelationskoeffizienten zum lag  $k$  findet man also, indem man die Parameter der lineare Regression  $\hat{X}_{k+1} = \phi_{k1}X_k + \dots + \phi_{kk}X_1$  berechnet. Der letzte Koeffizient  $\phi_{kk}$  ist dann die partielle Autokorrelation zum lag  $k$ .

Aus dem Innovationen-Algorithmus weiß man, dass der beste lineare Prädiktor für ein AR(p)-Modell

$$\hat{X}_{t+1} = \phi_1 X_t + \dots + \phi_p X_{t+1-p} \tag{3.60}$$

ist. Daher sieht man, dass für AR(p) die partielle Autokorrelation zum lag  $p$  gleich  $\phi_p$  ist und für  $k > p$  ist die partielle Autokorrelation  $\alpha(k) = \phi_{kk} = 0$ .

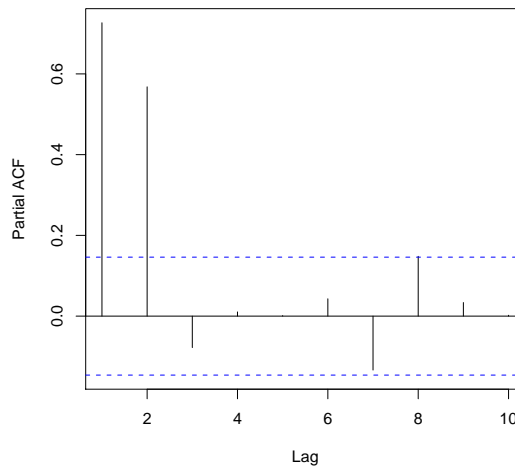


Abbildung 3.3: Partielle Autokorrelationsfunktion von  $X_t = 0.3X_{t-1} + 0.6X_{t-2} + Z_t$

Um reine MA-Prozesse zu identifizieren schaut man sich also die geschätzte Autokorrelationsfunktion an. Ist diese für alle lags  $k > q$  nicht signifikant (im Intervall  $[-1.96n^{-1/2}, 1.96n^{-1/2}]$ ), so handelt es sich um einen MA(q)-Prozess.

Bei AR-Prozessen betrachtet man hingegen die empirische partielle Autokorrelationsfunktion. Sind die partiellen Autokorrelationen für den lag  $k > p$  nicht signifikant, so ist die vorliegende Zeitreihe ein AR(p)-Prozess.

### 3.6.2 Informationskriterien

Bei gemischten ARMA(p,q)-Modellen wird es allerdings schwierig nur mit Hilfe der Autokorrelations- und der partiellen Autokorrelationsfunktion eine Aussage über die Ordnung zu machen. Effektiver sind deshalb sogenannte Informationskriterien, wie AIC (Akaike's Information Criterion), AICC (bias-corrected version of AIC) oder BIC (Bayesian Information Criterion). Es werden jene  $p$  und  $q$  gewählt für welche die Informationskriterien minimal sind.



### AIC (Akaike's Information Criterion)

Der AIC basiert auf den Kulback-Leibler-Index und wurde 1973 von Akaike wie folgt definiert:

$$AIC(\beta) := -2 \ln L_X(\beta, S_X(\beta)/n) + 2(p + q + 1) \quad (3.61)$$

wobei  $L_X$  die Likelihoodfunktion,  $\beta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$  und  $S_X$  die Quadratsumme der Residuen ist.  $S_X(\beta)/n$  ist ein Schätzer für die Varianz vom weißen Rauschen.

### AICC (bias-corrected version of AIC)

Die modifizierte Version vom AIC lautet:

$$AICC(\beta) := -2 \ln L_X(\beta, S_X(\beta)/n) + 2(p + q + 1)n/(n - p - q - 2) \quad (3.62)$$

Der AIC und der AICC neigen dazu die Ordnung  $p$  zu überschätzen. Eine Alternative zu diesen beiden Kriterien ist der

### BIC (Bayesian Information Criterion)

$$BIC := (n - p - q) \ln[n\hat{\sigma}^2/(n - p - q)] + n(1 + \ln \sqrt{2\pi}) + (p + q) \ln \left[ \left( \sum_{t=1}^n X_t^2 - n\hat{\sigma}^2 \right) / (p + q) \right] \quad (3.63)$$

wobei  $\hat{\sigma}^2$  der Maximum-Likelihood-Schätzer der Varianz vom weißen Rauschen ist.

## 3.7 Nichtstationäre Zeitreihen

Bei nichtstationären Zeitreihen kann man Trend und Saisonalität entweder durch die Methoden der deskriptiven Zeitreihenanalyse herausfiltern oder man wendet Erweiterungen des ARMA-Modells an.

Bei nichtstationären Zeitreihen mit einem Trend setzt man ARIMA-Modelle ein.

### 3.7.1 Der ARIMA-Prozess

Falls die Zeitreihe einen polynomialen Trend aufweist kann dieser durch Differenzenbildung eliminiert werden.

Hat eine Zeitreihe  $Y_1, \dots, Y_n$  beispielsweise einen linearen Trend  $Y_t = a + bt + X_t$ , wobei  $X_t$  ein stationärer Prozess mit Erwartungswert 0 ist, so kann der Trend durch den Übergang zu den ersten Differenzen eliminiert werden.

$$(1 - B)Y_t = Y_t - Y_{t-1} = a + bt + X_t - a - b(t-1) - X_{t-1} = b + (X_t - X_{t-1})$$

Die Zeitreihe kann nun als stationär angesehen werden. Auf die mittelwertbereinigte Zeitreihe  $(1 - B)Y_t$  kann nun ein ARMA(p,q)-Modell angewendet werden [Hamilton]. Die Zeitreihe  $Y_t$  wird als ARIMA(p,d,q)-Modell mit  $d=1$  bezeichnet.

Hat eine Zeitreihe hingegen einen polynomialen Trend vom Grad  $d$ , so kann sie durch bilden der  $d$ -ten Differenzen

$$(1 - B)^d Y_t = (1 - B)^{d-1} (1 - B) Y_t$$

stationär gemacht werden. Anschließend kann wieder ein ARMA-Prozess auf die mittelwertbereinigte differenzierte Zeitreihe angepasst werden [Hartung].

**Beispiel 7.** Die Zeitreihe  $Y_t$  habe einen quadratischen Trend, d.h.  $Y_t = a + bt + ct^2 + X_t$ , so kann dieser durch Bilden der zweiten Differenzen eliminiert werden

$$\begin{aligned} (1 - B)Y_t &= Y_t - Y_{t-1} = a + bt + ct^2 + X_t - a - b(t-1) - c(t-1)^2 - X_{t-1} \\ &= b + 2ct - c + (X_t - X_{t-1}) \\ (1 - B)(1 - B)Y_t &= b + 2ct - c(X_t - X_{t-1}) - (b + 2c(t-1) - c) - (X_{t-1} - X_{t-2}) \\ &= 2c + (X_t - 2X_{t-1} + X_{t-2}) \end{aligned}$$

**Definition 10.**  $\{Y_t\}$  ist ein **ARIMA(p,d,q)-Prozess**, falls der Trend der Zeitreihe ein Polynom vom Grad  $d$  ist und somit  $X_t := (1 - B)^d Y_t$  ein kausaler ARMA(p,q)-Prozess ist.

$$\phi(B)(1 - B)^d Y_t = \theta(B)Z_t \quad \{Z_t\} \sim WN(0, \sigma^2) \quad (3.64)$$

### 3.7.2 Der SARIMA-Prozess

Bei einer Zeitreihe mit konstanter Saisonsfigur der Periodenlänge  $s$ , kann ebenfalls die Differenzenmethode zur Elimination der Saisonskomponente angewandt werden. Dabei geht man

von der Zeitreihe  $Y_1, \dots, Y_n$  zu den saisonalen Differenzen

$$(1 - B^s)Y_t = Y_t - Y_{t-s}$$

über [Hartung].

Falls die Saisonsfigur als nicht konstant angesehen werden muss, so kann man ein ARMA(P,Q)-Modell

$$X_t = \Phi_1 X_{t-s} + \dots + \Phi_P X_{t-sP} + U_t + \Theta_1 + \dots + \Theta_Q U_{t-sQ} \quad (3.65)$$

einsetzen. Geht man von einer Zeitreihe mit Monatsdaten aus, so bezieht sich der autoregressive Anteil für den Monat März nicht auf die  $P$  vorangegangenen Monaten sondern auf die Märzmonate der  $P$  vorangegangenen Jahren. In der kompakten Form ist die obere Gleichung

$$\Phi(B^s)X_t = \Theta(B^s)U_t \quad (3.66)$$

Auf die Folge  $U_t$  kann man schließlich noch ein ARMA(p,q)-Modell

$$\phi(B)U_t = \Theta(B)Z_t \quad \{Z_t\} \sim WN(0, \sigma^2) \quad (3.67)$$

anwenden, sodass sich unter Einbezug der Differenzen  $(1 - B)^d$  und saisonalen Differenzen  $(1 - B^s)^D$  der saisonale ARIMA-Prozess ergibt.

**Definition 11. (Der SARIMA( $p, d, q$ )  $\times$  ( $P, D, Q$ ) $_s$  Prozess)**  $\{Y_t\}$  ist ein saisonaler ARIMA( $p, d, q$ )  $\times$  ( $P, D, Q$ ) $_s$  Prozess mit Periodenlänge  $s$ , falls der differenzierte  $X_t := (1 - B)^d(1 - B^s)^D Y_t$  ein stationärer Prozess

$$\phi(B)\Phi(B^s)X_t = \theta(B)\Theta(B^s)Z_t \quad \{Z_t\} \sim WN(0, \sigma^2), \quad (3.68)$$

ist [Brockwell, Davis. 1991].

Falls die Saisonsfigur als konstant identifiziert wurde sind  $P = 0$ ,  $Q = 0$  und  $D = 1$ .

### 3.8 Praktischer Teil

Für die maximale Tagestemperatur, den Tagestemperaturunterschied, die minimale Nachttemperatur und den Nachttemperaturunterschied wird die irreguläre Komponente der beiden Zeitreihenmodelle mit den Methoden von Box und Jenkins weiter untersucht. Die Voraussetzung für Box-Jenkins sind Stationarität der Zeitreihe. Zu diesem Zwecke wurden bereits im ersten Kapitel entsprechende Komponentenmodelle mit eventuellen varianzstabilisierenden Umformungen durchgeführt.

Für den stochastischen Prozess wird nun jenes ARMA-Modell gesucht, welches das AIC minimiert. Die Modellidentifikation und Parameterschätzung wurden mit dem Software-Programm R durchgeführt. R lässt außerdem nur stationäre und invertierbare Modelle zu, dh. die Wurzeln des autoregressiven und moving-average Polynoms müssen außerhalb des Einheitskreises liegen.

#### 3.8.1 ARMA-Modelle für das additive Komponentenmodell

Beim additiven Komponentenmodell mit Box-Cox-Transformation wird für die irreguläre Komponente zunächst das laut AIC beste ARMA-Modell ermittelt. Obwohl das Bestimmtheitsmaß nicht das Kriterium für die Auswahl des Modells ist, wird es als Gütemaß angewandt. Außerdem wird überprüft, ob die Residuen weißes Rauschen sind.

1. Für die Modellidentifikation der irreguläre Komponente der maximalen Tagestemperatur  $R_{\max 1}$  wird die empirische Autokorrelationsfunktion und die partielle Autokorrelationsfunktion analysiert (vgl. Abbildung 3.4). Man sieht, dass die partiellen Autokorrelation

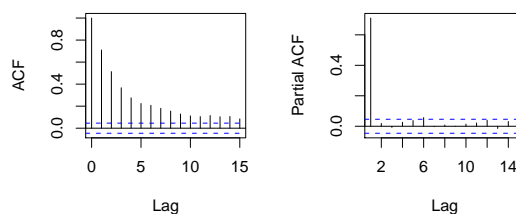


Abbildung 3.4: ACF und PACF von  $R_{\max 1}$

für die lags  $> 1$  nicht signifikant sind, somit geht man von einem AR(1)-Modell aus. Die partielle Autokorrelation zum lag 1 ist 0.71, somit ist  $\phi_1 = 0.71$ . Die Varianz des weißen

Rauschens ist laut Beispiel 6

$$\begin{aligned}\hat{\sigma}^2 &= \hat{\gamma}(0)(1 - \hat{\phi}_1^2) \\ &= 2452(1 - 0.71^2) = 1216\end{aligned}$$

Dieses Modell wird auch laut AIC als bestes Modell identifiziert.  
Das Bestimmtheitsmaß  $B$  dieses Modells ist

$$B = 1 - \frac{\hat{\sigma}_{WN}^2}{\hat{\sigma}_{Rtmax1}^2} = 1 - \frac{1216}{2453} = 0.50$$

2. Auch die Restkomponente des Tagestemperaturunterschiedes  $Rdtr1$  ist ein AR(1)-Modell mit  $\phi_1 = 0.4594$  und  $\sigma^2 = 118.7$

$$\begin{aligned}(1 - \phi_1 B)Rdtr1_t &= Z_t & Z_t &\sim WN(0, \sigma^2) \\ (1 - 0.46B)Rdtr1_t &= Z_t & Z_t &\sim WN(0, 118.7)\end{aligned}$$

Das Bestimmtheitsmaß  $B$  liegt nur bei 0.21.

3. Für die irreguläre Komponente des Nachttemperaturunterschieds  $Rntr1$  wird ein AR(1)-Modell identifiziert. Die Parameter sind  $\phi_1 = 0.5031$  und  $\sigma^2 = 84.82$ .

$$\begin{aligned}(1 - \phi_1 B)Rntr1_t &= Z_t & Z_t &\sim WN(0, \sigma^2) \\ (1 - 0.50B)Rntr1_t &= Z_t & Z_t &\sim WN(0, 84.82)\end{aligned}$$

Das Bestimmtheitsmaß  $B$  liegt hier auch nur bei 0.25.

4. Für die Restkomponente der Nachttiefsttemperatur  $Rtmin1$  wird auch ein AR(1)-Modell angenommen. Der Parameter  $\phi_1$  ist 0.6481. Die Varianz des weißen Rauschen ist  $\sigma^2 = 723.4$ .

$$\begin{aligned}(1 - \phi_1 B)Rtmin1_t &= Z_t & Z_t &\sim WN(0, \sigma^2) \\ (1 - 0.65B)Rtmin1_t &= Z_t & Z_t &\sim WN(0, 723)\end{aligned}$$

Das Bestimmtheitsmaß  $B$  liegt hier bei 0.42.

### 3.8.2 ARMA-Modelle für das standardisierende Komponentenmodell

1. Die Restkomponente des standardisierenden Komponentenmodells für die Tageshöchsttemperatur  $Rtmax2$  ist ein AR(1)-Modell mit  $\phi_1 = 0.7132$  und  $\sigma^2 = 0.5273$

$$\begin{aligned}(1 - \phi_1 B)Rtmax2_t &= Z_t & Z_t &\sim WN(0, \sigma^2) \\ (1 - 0.71B)Rtmax2_t &= Z_t & Z_t &\sim WN(0, 0.5273)\end{aligned}$$

Das Bestimmtheitsmaß des AR(1)-Modells ist 0.51.

2. Für die irreguläre Komponente des Tagestemperaturunterschieds  $Rdtr2$  wird als AR(1)-Modell mit  $\phi_1 = 0.4786$  und  $\sigma^2 = 0.8373$  identifiziert.

$$\begin{aligned}(1 - \phi_1 B)Rdtr2_t &= Z_t & Z_t &\sim WN(0, \sigma^2) \\ (1 - 0.48B)Rdtr2_t &= Z_t & Z_t &\sim WN(0, 0.84)\end{aligned}$$

Das Bestimmtheitsmaß  $B$  ist 0,23.

3. Der Restkomponente des Nachttemperaturunterschieds wird wiederum ein AR(1)-Modell angepasst. Der Parameter  $\phi_1$  ist 0.5245. Die Varianz des weißen Rauschens ist  $\sigma^2 = 0.78$

$$\begin{aligned}(1 - \phi_1 B)Rntr2_t &= Z_t & Z_t &\sim WN(0, \sigma^2) \\ (1 - 0.52B)Rntr2_t &= Z_t & Z_t &\sim WN(0, 0.78)\end{aligned}$$

Das Bestimmtheitsmaß  $B$  liegt bei 0.28.

4. Für die irreguläre Komponente der Nachttiefsttemperatur  $Rtmin2$  wird ein ARMA(2,1)-Modell identifiziert. Die Parameter sind  $\phi_1 = 1.4006$ ,  $\phi_2 = -0.4624$ , und  $\theta_1 = -0.8029$ . Die Varianz des weißen Rauschens ist  $\sigma^2 = 0.6725$

$$\begin{aligned}(1 - \phi_1 B - \phi_2 B^2)Rtmin2_t &= (1 + \theta_1 B)Z_t & Z_t &\sim WN(0, \sigma^2) \\ (1 - 1.41B + 0.47B^2)Rtmin2_t &= (1 - 0.81B)Z_t & Z_t &\sim WN(0, 0.67)\end{aligned}$$

Das Bestimmtheitsmaß  $B$  liegt bei 0.39.

### 3.8.3 Modellüberprüfung

Für alle Zeitreihen werden nun die durch den ARMA-Prozess, den Trend und die Saisonalität vorhergesagten Werte mit den tatsächlichen Zeitreihenwerten verglichen.

#### Modellüberprüfung des additiven Komponentenmodells

Die Einschnitt-Prädiktoren der ARMA-Prozesse werden mit  $\hat{R}_t$  bezeichnet. Somit ist das Vorhersagemodell für das additive Komponentenmodell

$$\begin{aligned}\hat{y}(\lambda) &= c + \alpha \cos(2\pi t/365) + \beta \sin(2\pi t/365) + \hat{R}_t \\ \hat{y}_t &= (\lambda_1 \hat{y}(\lambda) + 1)^{1/\lambda_1} - \lambda_2\end{aligned}\tag{3.69}$$

Als Merkmale für die Prognosequalität der ARMA-Modelle werden der maximale Fehler und die Wahrscheinlichkeit, dass der Prognosefehler  $< 4^\circ C$  ist, gewählt.

Als Vergleich für die Prognosequalität des ARMA-Modells wird die Prognosequalität der deskriptiven Zeitreihenanalyse herangezogen. Für die vier Zeitreihen sind die Ergebnisse in der folgenden Tabelle aufgelistet.

	ARMA-Modell		deskriptive Z.	
	max. Fehler	$\pm 4^\circ C$	max. Fehler	$\pm 4^\circ C$
Tmax	17° C	78%	15° C	54%
DTR	13° C	66%	15° C	54%
NTR	9.6° C	79%	11° C	67%
TMIN	11° C	86%	14° C	75%

### Modellüberprüfung des standardisierenden Komponentenmodells

Das Vorhersagemodell für das standardisierende Modell ist

$$\hat{y}_t = \hat{T}_t + \hat{S}_t + \hat{\sigma}_t \hat{R}_t \tag{3.70}$$

wobei  $\hat{R}_t$  die Einschnitt-Prädiktoren des ARMA-Modells sind. Die Ergebnisse sind

	ARMA-Modell		deskriptive Z.	
	max. Fehler	$\pm 4^\circ C$	maximaler Fehler	$\pm 4^\circ C$
Tmax	17° C	78%	15° C	54%
DTR	14° C	67%	14° C	54%
NTR	10° C	78%	11° C	67%
TMIN	11° C	87%	14° C	75%

### Modellkontrollmethode der Meteorologie

In der Meteorologie ist es üblich eine Vorhersagemethode mit der Methode "morgen wird es wie heute" oder Persistenz-Vorhersage zu vergleichen. Für diese Methode erhält man folgende Werte

	max. Fehler	$\pm 4^\circ C$
Tmax	20° C	76%
DTR	20° C	63%
NTR	13° C	74%
TMIN	12° C	81%

Mit der Methode von Box und Jenkins konnten also keine relevante Verbesserungen gegenüber der Persistenzmethode erreicht werden, zumal diese Methode Ähnlichkeit mit einem AR(1)-Modell hat.

## 4 Bivariate Zeitreihenanalyse

Bisher wurden nur univariate stochastische Prozesse betrachtet. Oft ist es aber sinnvoll Zeireihen für mehrere Variablen, multivariate Zeitreihen, zu betrachten.

Dabei soll die Abhängigkeit zwischen den Variablen untersucht werden und gegebenenfalls die Art der Beziehung zwischen den Variablen ermittelt werden. Die Variablen können sich einseitig oder gegenseitig beeinflussen.

Multivariate ARMA-Modelle ergeben sich durch Verallgemeinerung der univariaten Modelle. An Stelle von einzelnen stochastischen Prozessen treten vektorielle Prozesse, an Stelle von (skalaren) Parametern treten Parametermatrizen.

Ich werde mich auf bivariate Prozesse beschränken, wobei ich auf den autoregressiven Prozess AR(p) näher eingehen werde, da die meisten univariaten meteorologischen Zeitreihen eine Realisation autoregressive Prozessen darstellen.

Falls sich die Variablen der Zeitreihen einseitig beeinflussen, wird nicht ein AR(p)-Modell sondern ein sogenanntes Transferfunktionenmodell verwendet. Hierbei wird eine Zeitreihenvariable als Input (oder unabhängig) und die andere als Output (oder abhängig) betrachtet.

### 4.1 Erwartungswert und Korrelation von bivariaten Prozessen

**Definition 12.** Bei bivariaten stochastischen Prozessen  $\mathbf{X}_t = (X_{t1}, X_{t2})'$  ist der **Vektor des Erwartungswerts** zum Zeitpunkt  $t$

$$\boldsymbol{\mu}_t := E\mathbf{X}_t = \begin{pmatrix} EX_{t1} \\ EX_{t2} \end{pmatrix} \quad (4.1)$$

und die **Kovarianzmatrix**

$$\Gamma(t+h, t) := Cov(\mathbf{X}_{t+h}, \mathbf{X}_t) = \begin{pmatrix} cov(X_{t+h,1}, X_{t1}) & cov(X_{t+h,1}, X_{t2}) \\ cov(X_{t+h,2}, X_{t1}) & cov(X_{t+h,2}, X_{t2}) \end{pmatrix} \quad (4.2)$$

Analog zur univariaten Zeitreihe vom vorigen Kapitel, ist die bivariate Zeitreihe  $\{\mathbf{X}_t\}$  (schwach) stationär, falls der Erwartungswertvektor und die Kovarianzmatrix unabhängig vom Zeitpunkt  $t$  sind. In diesem Fall verwendet man folgende Notation

$$\boldsymbol{\mu} = E\mathbf{X}_t = \begin{pmatrix} EX_{t1} \\ EX_{t2} \end{pmatrix} \quad (4.3)$$



und

$$\Gamma(h) = Cov(\mathbf{X}_{t+h}, \mathbf{X}_t) = \begin{pmatrix} \gamma_{11}(h) & \gamma_{12}(h) \\ \gamma_{21}(h) & \gamma_{22}(h) \end{pmatrix} \quad (4.4)$$

Die beiden Kovarianzen  $\gamma_{12}(h)$  und  $\gamma_{21}(h)$  sind dabei nicht gleich, sondern  $\gamma_{12}(h) = \gamma_{21}(-h)$ .

Als **Schätzer** für die zwei Komponenten des Erwartungswertvektors wird wiederum das Stichprobenmittel genommen:

$$\bar{X}_1 = \frac{1}{n} \sum_{t=1}^n X_{t1}, \quad \bar{X}_2 = \frac{1}{n} \sum_{t=1}^n X_{t2} \quad (4.5)$$

Der Schätzer für Einträge der Kovarianzmatrix ist

$$\hat{\gamma}_{ij}(h) = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-h} (X_{t+h,i} - \bar{X}_i)(X_{t,j} - \bar{X}_j), & \text{für } 0 \leq h \leq n-1; \\ \hat{\gamma}_{ij}(-h), & \text{für } -n+1 \leq h < 0. \end{cases} \quad (4.6)$$

Die Korrelation  $\rho_{ij}(h)$  zum lag  $h$  zwischen  $X_{t+h,i}$  und  $X_{t,j}$  wird dann als

$$\hat{\rho}_{ij}(h) = \frac{\hat{\gamma}_{ij}(h)}{\sqrt{\hat{\gamma}_{ii}(0)\hat{\gamma}_{jj}(0)}} \quad (4.7)$$

geschätzt [Brockwell, Davis. 2002].

## 4.2 Unabhängigkeit von zwei Zeitreihenvariablen

Die Unabhängigkeit von zwei Zeitreihenvariablen testet man laut [Brockwell, Davis. 2002], indem man zunächst auf beide Variablen getrennt ein ARMA-Modell anpasst. Anschließend werden die Residuen oder white noise der beiden Prozesse verglichen. Dazu wird die Kreuzkorrelationsfunktion geschätzt. Ist die geschätzte Kreuzkorrelation für die zwei white noise  $Z_1$  und  $Z_2$

$$\hat{\rho}_{12}(h) = \frac{\hat{\gamma}_{12}(h)}{\sqrt{\hat{\gamma}_{11}(0)\hat{\gamma}_{22}(0)}} = \frac{cov(Z_{t+h,1}, Z_{t2})}{\sigma_{Z1}\sigma_{Z2}}$$

für kein  $h$  signifikant, d.h. innerhalb des Intervalls  $[-1.96n^{-1/2}, 1.96n^{-1/2}]$ , dann sind die Zeitreihenvariablen als unabhängig zu betrachten.

Bei Abhängigkeit ( $\exists h \hat{\rho}_{12}(h) \neq 0$ ) kann die Richtung des linearen Zusammenhangs ermittelt werden.

### 4.3 Richtung des linearen Zusammenhangs zweier Variablen

Die Kreuzkorrelationsfunktion  $\hat{\rho}_{12}(h)$  zwischen den Variablen  $X_{t+h,1}$  und  $X_{t2}$  misst hingegen die Stärke und die Richtung des linearen Zusammenhangs

- Ist  $\hat{\rho}_{12}(h)$  für kein  $h < 0$  signifikant, dann ist  $X_{t2}$  kausal für  $X_{t1}$ , bzw.  $X_{t1}$  ist von  $X_{t2}$  abhängig.
- Ist umgekehrt  $\hat{\rho}_{12}(h)$  für kein  $h > 0$  signifikant, dann wird  $X_{t1}$  als kausal für  $X_{t2}$  bezeichnet.
- $X_{t1}$  und  $X_{t2}$  beeinflussen sich gegenseitig, falls  $\hat{\rho}_{12}(h)$  für gewisse  $h > 0$  und  $h < 0$  ist.

Wenn sich die white-noise-Zeitreihen gegenseitig beeinflussen ist es ratsam von einem bivariaten ARMA(p,q)-Prozess auszugehen. Bei einseitiger Beeinflussung wird hingegen ein Transferfunktionenmodell angewendet.

### 4.4 Bivariate ARMA-Prozesse

**Definition 13.** Eine multivariate Zeitreihe  $\{X_t\}$  ist ein **ARMA(p,q)-Prozess**, falls  $\{X_t\}$  stationär ist und zu jedem Zeitpunkt  $t$  gilt

$$X_t - \Phi_1 X_{t-1} - \dots - \Phi_p X_{t-p} = Z_t + \Theta_1 Z_{t-1} + \dots + \Theta_q Z_{t-q} \quad \{Z_t\} \sim WN(\mathbf{0}, \Sigma) \quad (4.8)$$

$\Phi_1, \dots, \Phi_p$  und  $\Theta_1, \dots, \Theta_q$  sind im bivariaten Fall  $2 \times 2$ - Matrizen und  $\Sigma$  ist die Kovarianzmatrix von  $\{X_{t1}\}$  und  $\{X_{t2}\}$  zum lag 0 [Brockwell, Davis. 1991].

Im Wesentlichen sind die Verfahren zur Parameterschätzung für univariate Zeitreihen auf multivariate Zeitreihen übertragbar.

Da man die meisten univariaten meteorologischen Zeitreihen als autoregressiven Prozess darstellen kann (siehe Ausführung Kapitel 3), werde ich mich im Folgenden auf bivariate AR-Prozesse beschränken.

#### 4.4.1 Parameterschätzung bivariater Autoregressiver Prozesse

Für den bivariaten kausalen AR(p)-Prozess

$$X_t = \Phi_1 X_{t-1} + \dots + \Phi_p X_{t-p} + Z_t, \quad Z_t \sim WN(\mathbf{0}, \Sigma) \quad (4.9)$$

können die Matrizen  $\Phi_1, \dots, \Phi_p$  und  $\Sigma$  mittels Yule-Walker-Gleichungen geschätzt werden. Im multivariaten Fall spricht man vom Whittle's Algorithmus [Brockwell, Davis. 1991].

Eine weitere Möglichkeit für die Parameterschätzung eines autoregressiven Modells ist die lineare Regression [Hamilton]. Diese wird mit einem Beispiel veranschaulicht.

**Beispiel 8** (Down Jones und All Ordinaries Indices). Die zwei Zeitreihen des Dow Jones Index in New York  $\{X_{t1}\}$  und des australischen All Ordinaries Indices  $\{X_{t2}\}$  wurden aus [Brockwell, Davis. 2002] entnommen. In den Zeitreihen ist die prozentuelle Änderung des Indekurses zum Vortag angegeben. Wegen der Zeitverschiebung zwischen Sydney und New York schließen die Märkte nicht gleichzeitig, in Sydney ist aber der Schlusskurs des Dow Jones bei Börsenöffnung bekannt. Die Daten umfassen jeweils 250 Einträge.

Zunächst werden die Autokorrelationen und die Kreuzkorrelationen der beiden Zeitreihen mit R dargestellt (vgl. Abbildung 4.1). Man sieht, dass nur die Kreuzkorrelation zwischen All Or-

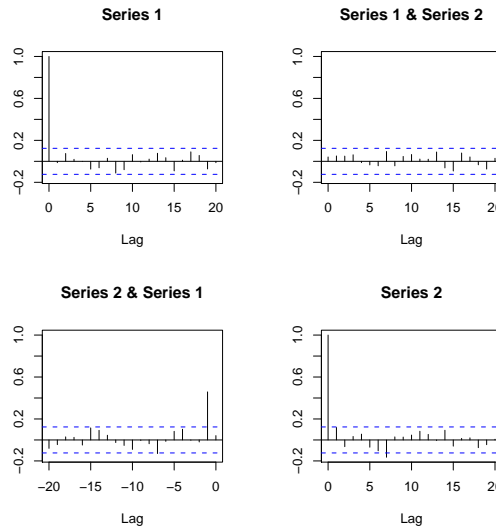


Abbildung 4.1: Autokorrelationen und Kreuzkorrelationen von Dow Jones (Series 1) und All Ordinaries (Series 2)

dinaries und Dow Jones  $\hat{\rho}_{21}(h) = \hat{\rho}_{12}(-h)$  zum lag 1 signifikant ist. Deshalb wird ein bivarites AR(1)-Modell  $\mathbf{X}_t = \Phi_1 \mathbf{X}_{t-1} + \mathbf{Z}_t$  angewendet. Für die Schätzung der Parameter  $\Phi_1$  und  $\Sigma$  wird die multiple lineare Regression verwendet. Die Datenmatrix  $X$  und die abhängige Variable  $Y$  sind dann

$$X = \begin{pmatrix} 1 & X_{11} & X_{12} \\ \vdots & \vdots & \vdots \\ 1 & X_{n-1,1} & X_{n-1,2} \end{pmatrix} \quad Y = \begin{pmatrix} X_{21} & X_{22} \\ \vdots & \vdots \\ X_{n,1} & X_{n,2} \end{pmatrix}$$

Der Parametervektor wird mit  $b$  bezeichnet.

Für die multiple lineare Regression  $Y = Xb$  soll nun  $b$  geschätzt werden. Durch Ableiten und Nullsetzen der Summe der Quadratfehler

$$S = (Y - Xb)'(Y - Xb) \quad (4.10)$$

erhält man die sogenannte Normalengleichung und den Schätzer für den Parametervektor  $b$

$$\hat{b} = (X'X)^{-1}X'Y \quad (4.11)$$

Für den Dow Jones und den All Ordinaries Index sind dann die Parameter:

$$\hat{b} = \begin{pmatrix} 0.027 & 0.012 \\ -0.015 & 0.671 \\ 0.036 & 0.099 \end{pmatrix} \quad (4.12)$$

Und

$$\mathbf{X}_t = \begin{pmatrix} 0.027 \\ 0.012 \end{pmatrix} + \begin{pmatrix} -0.015 & 0.036 \\ 0.671 & 0.099 \end{pmatrix} \mathbf{X}_{t-1} + \mathbf{Z}_t \quad \mathbf{Z}_t \sim WN(\mathbf{0}, \Sigma) \quad (4.13)$$

Nun muss nur mehr die Kovarianzmatrix  $\Sigma$  geschätzt werden

$$\hat{\Sigma} = \frac{1}{n-2}(Y - X\hat{b})'(Y - X\hat{b}) = \begin{pmatrix} 0.368 & 0.023 \\ 0.023 & 0.603 \end{pmatrix} \quad (4.14)$$

#### 4.4.2 Identifikation des autoregressiven Modells

Auch für den multivariaten Fall kann zur Identifikation des autoregressiven Modells **Akaike's Informationskriterium** angewendet werden. Für eine m-variate Zeitreihe ist dieser

$$AICC = -2 \ln L(\Phi - 1, \dots, \Phi_p, \Sigma) + \frac{2(pm^2 + 1)nm}{nm - pm^2 - 2} \quad (4.15)$$

Es wird das AR(p)-Modell ausgewählt, bei dem das AICC minimal ist [Brockwell, Davis. 1991].

### 4.5 Transferfunktionenmodell

Das Transferfunktionenmodell oder transfer noise model beschreibt den Zusammenhang zwischen zwei Zeitreihen  $\{X_{t,1}\}$  und  $\{X_{t,2}\}$ , wobei  $\{X_{t,2}\}$  von  $\{X_{t,1}\}$  abhängig ist, aber das Umgekehrte gilt nicht.  $\{X_{t,1}\}$  kann als unabhängige Variable oder Input betrachtet werden und  $\{X_{t,2}\}$  als Output oder abhängige Variable.

Mit diesem Modell wird versucht bessere Ergebnisse für die Vorhersage der Temperatur zu erzielen. Es wird entweder die Globalstrahlung oder die Bewölkung als Input verwendet.

Es wurde den Ausführungen von [Brockwell, Davis. 2002] gefolgt.

Man betrachtet folgendes Transferfunktionenmodell zwischen dem Input  $\{X_{t,1}\}$  und dem Output  $\{X_{t,2}\}$ :

$$X_{t,2} = \sum_{j=0}^{\infty} \tau_j X_{t-j,1} + N_t \quad (4.16)$$

wobei  $T = \{\tau_j, j = 0, 1, \dots\}$  ein zeitinvarianter linearer Filter ist.  $\{N_t\}$  ist ein stationärer Prozess mit Erwartungswert 0, der mit dem Inputprozess  $\{X_{t,1}\}$  unkorreliert ist. Weiters wird angenommen, dass  $\{X_{t,1}\}$  eine stationäre Zeitreihe mit Erwartungswert 0 ist. Dann ist der bivariate Prozess  $\{X_{t1}, X_{t2}\}$  auch stationär.

Multipliziert man beide Seiten von (4.16) mit  $X_{t-k,1}$  und nimmt dann den Erwartungswert davon, erhält man die Gleichung:

$$\begin{aligned} X_{t,2}X_{t-k,1} &= \sum_{j=0}^{\infty} \tau_j X_{t-j,1}X_{t-k,1} + X_{t-k,1}N_t \\ E(X_{t,2}X_{t-k,1}) &= \sum_{j=0}^{\infty} \tau_j E(X_{t-j,1}X_{t-k,1}) \\ \gamma_{21}(k) &= \sum_{j=0}^{\infty} \tau_j \gamma_{11}(k-j) \end{aligned}$$

Wenn der Input ein weißes Rauschen wäre, z.B.  $\{X_{t1}\} \sim WN(0, \sigma_1^2)$  mit  $\gamma_{11}(k-j) = 0$  für  $k \neq j$  und  $\gamma_{11}(0) = \sigma_1^2$ , dann wäre

$$\tau_k = \gamma_{21}(k) / \sigma_1^2 \quad (4.17)$$

und man könnte  $\{\tau_j, j = 0, 1, \dots\}$  leicht aus den Kreuzkorrelationen schätzen. Aus diesem Grund ist es sinnvoll den Inputprozess vorzuweißen ("prewhitening"). Dies vereinfacht die Identifikation eines geeigneten Transferfunktionenmodells und liefert erste Schätzer  $\hat{\tau}_k$  für die Koeffizienten  $\tau_k$ .

Falls der Input  $\{X_{t1}\}$  als ARMA(p,q) Prozess dargestellt werden kann,

$$\phi(B)X_{t1} = \theta(B)Z_t, \quad \{Z_t\} \sim WN(0, \sigma_Z^2) \quad (4.18)$$

dann erhält man durch Anwendung des Filters  $\pi(B) = \phi(B)\theta^{-1}(B)$  auf  $\{X_{t1}\}$  die geweißte Reihe  $\{Z_t\}$ .

Wendet man nun den Operator  $\pi(B)$  auf beide Seiten von (4.16) an, erhält man

$$\pi(B)X_{t,2} = \sum_{j=0}^{\infty} \tau_j \pi(B)X_{t-j,1} + \pi(B)N_t \quad (4.19)$$

Sei  $Y_t = \pi(B)X_{t,2}$  und  $N'_t = \pi(B)N_t$ , dann ist das Transferfunktionenmodell

$$Y_t = \sum_{j=0}^{\infty} \tau_j Z_{t-j} + N'_t. \quad (4.20)$$

$\{N'_t\}$  ist wiederum ein stationärer Prozess mit Erwartungswert 0, der mit  $\{Z_t\}$  unkorreliert ist. Multipliziert man nun beide Seiten mit  $Z_{t-i}$  und bildet wiederum den Erwartungswert erhält man

$$\begin{aligned} Y_t Z_{t-i} &= \sum_{j=0}^{\infty} \tau_j Z_{t-j} Z_{t-i} + N'_t Z_{t-i} \\ E(Y_t Z_{t-i}) &= \sum_{j=0}^{\infty} \tau_j E(Z_{t-j} Z_{t-i}) + E(N'_t Z_{t-i}) \\ \gamma_{YZ}(i) &= \sum_{j=0}^{\infty} \tau_j \gamma_{ZZ}(i-j) \end{aligned}$$

$\{Z_t\}$  ist weißes Rauschen, deshalb ist  $\gamma_{ZZ}(i-j) = 0$  für  $i \neq j$ , daraus folgt:

$$\begin{aligned} \gamma_{YZ}(i) &= \tau_i \gamma_{ZZ}(0) = \tau_i \sigma_Z^2 \\ \tau_i &= \frac{\gamma_{YZ}(i)}{\sigma_Z^2} = \frac{\rho_{YZ}(i) \sigma_Y \sigma_Z}{\sigma_Z^2} = \frac{\rho_{YZ}(i) \sigma_Y}{\sigma_Z} \end{aligned} \quad (4.21)$$

Und die  $\tau_i$  können aus der Kreuzkorrelationsfunktion und den Varianzen von den vorgeweißten In- und Outputs geschätzt werden.

Dies führt zu folgendem Verfahren zur Identifikation von Transferfunktionenmodellen

## 4.6 Verfahren zur Identifikation von Transferfunktionenmodellen

1. Für den Input-Prozess  $\{X_{t1}\}$  wird ein ARMA-Modell identifiziert

$$\hat{\phi}(B)X_{t1} = \hat{\theta}(B)\hat{Z}_t \quad (4.22)$$

$\hat{\phi}$  und  $\hat{\theta}$  sind die Maximum-Likelihood-Schätzer der Autoregressiv- und moving-average Parameter.  $\hat{\sigma}_Z^2$  ist der Schätzer der Varianz von  $\{Z_t\}$ . Wendet man den Operator  $\hat{\pi}(B) = \hat{\phi}(B)\hat{\theta}^{-1}(B)$  auf  $\{X_{t1}\}$  an erhält man die Residuen  $\{\hat{Z}_t\}$ .

2. Der Operator  $\hat{\pi}(B)$  wird nun auf die Output-Reihe  $\{X_{t2}\}$  angewandt, die erhaltenen Residuen werden als  $\{\hat{Y}_t\}$  bezeichnet, ihre geschätzte Varianz als  $\hat{\sigma}_Y^2$ .
3. Nun wird die Kreuzkorrelationsfunktion  $\rho_{YZ}(h)$  zwischen  $\{\hat{Y}_t\}$  und  $\{\hat{Z}_t\}$  geschätzt. Um ermitteln zu können welche Kreuzkorrelationen signifikant sind, wird die Bartlett-Formel verwendet.

Unter der Annahme, dass  $\{\hat{Z}_t\} \sim WN(0, \hat{\sigma}_Z^2)$  und  $\{(\hat{Y}_t, \hat{Z}_t)'\}$  stationär multivariat normalverteilt ist gilt:

$$nVar(\hat{\rho}_{YZ}(h)) \sim 1 - \rho_{YZ}^2(h) \left[ 1.5 - \sum_{k=-\infty}^{\infty} (\rho_{YZ}^2(k) + \rho_{YY}^2(k)/2) \right] + \sum_{k=-\infty}^{\infty} [\rho_{YZ}(h+k)\rho_{YZ}(h-k) - 2\rho_{YZ}(h)\rho_{YZ}(h+k)\rho_{YY}^2(k)]$$

Unter der Nullhypothese  $H_0$ , dass  $\rho_{YZ}(h) = 0$ ,  $h \notin [a, b]$  wobei  $a$  und  $b$  ganze Zahlen sind, ist

$$Var(\hat{\rho}_{YZ}(h)) \sim \frac{1}{n} \quad \text{für } h \notin [a, b] \quad (4.23)$$

Ein Vergleich der geschätzten Kreuzkorrelationen zum lag  $h$   $\hat{\rho}_{YZ}(h)$  mit den Grenzen  $\pm 1.96n^{-1/2}$  gibt Aufschluss darüber, für welche lags  $h$  die Kreuzkorrelation signifikant unterschiedlich von 0 ist.

4. Die Schätzer für  $\tau_h$  zum lag  $h$  bei welchem  $\hat{\rho}_{YZ}(h)$  signifikant ist, sind

$$\hat{\tau}_h = \hat{\rho}_{YZ}(h)\hat{\sigma}_Y/\hat{\sigma}_Z \quad (4.24)$$

Für die übrigen  $h$  sind die Schätzer  $\hat{\tau}_h = 0$ . Sei  $m > 0$  der größte Wert von  $h$ , für welcher  $\hat{\tau}_h$  ungleich Null ist und  $b \geq 0$  der kleinste Wert. Der Parameter  $b$  des Filters  $\{\hat{\tau}_j\}$  bezeichnet dabei die reine Verzögerung [Stier].

Falls  $m$  sehr groß ist und die Koeffizienten  $\{\hat{\tau}_j\}$  annähernd einer Differenzgleichung der Form

$$\hat{\tau}_j - v_1\hat{\tau}_{j-1} - \dots - v_p\hat{\tau}_{j-p} = 0, \quad j \geq b + p \quad (4.25)$$

entsprechen, kann  $\hat{T}(B) = \sum_{j=b}^m \hat{\tau}_j B^j$  mit weniger Parametern dargestellt werden

$$\hat{T}(B) = w_0(1 - v_1B - \dots - v_pB^p)^{-1}B^b \quad (4.26)$$

Box und Jenkins empfehlen  $\hat{T}(B)$  als rationale Funktion zu wählen.

$$\hat{T}(B) = B^b w(B)v^{-1}(B) = B^b(w_0 + w_1B + \dots + w_qB^q)(1 - v_1B - \dots - v_pB^p)^{-1}$$

Die Grade der Polynome sind meist schwer zu bestimmen. Aber auch hier gilt, dass man mit möglichst wenigen Parametern versuchen soll  $\hat{T}(B)$  zu approximieren. Die Parameter werden mittels Koeffizientenvergleich der Polynome bestimmt. In diesem Fall ist  $m = \max(q + b, p)$ .

5. Die Noise-Reihe  $N_t, t = m + 1, \dots, n$  wird geschätzt als

$$\hat{N}_t = X_{t2} - \hat{T}(B)X_{t1} \quad (4.27)$$

Für  $t \leq m$  wird  $N_t = 0$  gesetzt.

6. Falls die Noise-Reihe nicht weißes Rauschen ist, soll ein ARMA( $p_2, q_2$ )-Modell für  $N_t$  gefunden werden

$$\phi^{(N)}(B)N_t = \theta^{(N)}(B)W_t \quad \{W_t\} \sim WN(0, \sigma_W^2) \quad (4.28)$$

Dabei wird die Reihe  $\hat{N}_{m+1}, \dots, \hat{N}_n$  betrachtet, da die ersten  $m$  Werte Null sind.

7. Das identifizierte Transferfunktionenmodell ist nun

$$\begin{aligned} X_{t2} &= \hat{T}(B)X_{t1} + \hat{N}_t \\ &= \sum_{j=b}^m \hat{\tau}_j B^j X_{t1} + \theta^{(N)}(B)\phi^{(N)-1}(B)W_t \\ \phi^{(N)}(B)X_{t2} &= \sum_{j=b}^m \hat{\tau}_j B^j \phi^{(N)}(B)X_{t1} + \theta^{(N)}(B)W_t \end{aligned} \quad (4.29)$$

Falls  $\hat{T}(B)$  als rationale Funktion dargestellt wurde, ist das Transferfunktionenmodell

$$\begin{aligned} X_{t2} &= \hat{T}(B)X_{t1} + \hat{N}_t \\ &= B^b w(B)v^{-1}(B)X_{t1} + \theta^{(N)}(B)\phi^{(N)-1}(B)W_t \\ \phi^{(N)}(B)v(B)X_{t2} &= B^b w(B)\phi^{(N)}(B)X_{t1} + v(B)\theta^{(N)}(B)W_t \end{aligned} \quad (4.30)$$

Die Residuen des Transferfunktionenmodells  $\hat{W}_t(w, v, \phi^{(N)}, \theta^{(N)})$  können für  $t > m^* = \max(p_2 + p, b + p_2 + q)$  berechnet werden, für  $t \leq m^*$  ist  $\hat{W}_t = 0$ .

Die Parameter  $w, v, \phi^{(N)}$  und  $\theta^{(N)}$  können mittels der Methode der kleinsten Quadrate nochmals geschätzt werden. Die Summe der Quadratfehler

$$\sum_{m^*+1}^n \hat{W}_t^2(w, v, \phi^{(N)}, \theta^{(N)}) \quad (4.31)$$

soll dabei minimiert werden. Außerdem soll sichergestellt werden, dass die Wurzeln der Polynome  $\phi^{(N)}(B)$ ,  $\theta^{(N)}(B)$  und  $v(B)$  außerhalb des Einheitskreises liegen.

8. Um die Güte des Transferfunktionenmodells zu überprüfen werden die Residuen  $\{\hat{W}_t, t > m^*\}$  und  $\{\hat{Z}_t, t > m^*\}$  betrachtet. Die Reihen sollen unkorreliert und weißes Rauschen sein.

Es sind nämlich verschiedene Fehlspezifikationen des Modells möglich: das noise-Modell, die Transferfunktion oder beides sind fehlspezifiziert. Im einzelnen sind folgende Fälle zu unterscheiden:

- a) Die Transferfunktion  $\hat{T}(B)$  ist fehlspezifiziert. Sei

$$X_{t2} = \hat{T}(B)X_{t1} + \theta^{(N)}/\phi^{(N)}W_t$$

das korrekt spezifiziert Modell und bezeichne  $\hat{T}_0(B)$  die fehlspezifiziert Transferfunktion. Als Konsequenz dieser Fehlspezifikation ergibt sich ein fehlspezifiziertes noise-Modell  $\theta_0^{(N)}/\phi_0^{(N)}V_t$ . Somit ist

$$X_{t2} = \hat{T}_0(B)X_{t1} + \theta_0^{(N)}/\phi_0^{(N)}V_t$$



Daraus folgt

$$\hat{T}(B)X_{t1} + \theta^{(N)}(B)/\phi^{(N)}(B)W_t = \hat{T}_0(B)X_{t1} + \theta_0^{(N)}(B)/\phi_0^{(N)}(B)V_t$$

und:

$$\begin{aligned} V_t &= \frac{\phi_0^{(N)}(B)}{\theta_0^{(N)}(B)} [\hat{T}(B) - \hat{T}_0(B)]X_{t1} + \frac{\phi_0^{(N)}(B)}{\theta_0^{(N)}(B)} \frac{\theta^{(N)}(B)}{\phi^{(N)}(B)} W_t \\ &= \frac{\phi_0^{(N)}(B)}{\theta_0^{(N)}(B)} [\hat{T}(B) - \hat{T}_0(B)] \frac{\theta(B)}{\phi(B)} Z_t + \frac{\phi_0^{(N)}(B)}{\theta_0^{(N)}(B)} \frac{\theta^{(N)}(B)}{\phi^{(N)}(B)} W_t \end{aligned}$$

Somit ist  $V_t$  mit den Residuen des Input  $Z_t$  kreuzkorreliert, dies gilt sogar wenn das noise-Modell korrekt spezifiziert ist.

- b) Das noise-Modell ist fehl-, die Transferfunktion aber korrekt spezifiziert. In diesem Fall ist die Kreuzkorrelation gleich Null für alle lags.

Deshalb ist es sinnvoller zuerst die Kreuzkorrelation zu überprüfen.

**Beispiel 9.** Es werden wieder die Zeitreihen des Dow Jones und des All Ordinaries betrachtet. Auf die zwei Zeitreihen wird zunächst jeweils ein ARMA-Modell angewandt, dabei ist laut AIC die Zeitreihe des Dow Jones weißes Rauschen.

Für den All Ordinaries Index wird hingegen laut AIC ein MA(1)-Modell identifiziert.

$$X_{t,AO} = 0.14Z_{t-1} + Z_t \quad \{Z_t\} \sim WN(0, 0.76) \quad (4.32)$$

Dies kann man auch anhand der Autokorrelation und partiellen Autokorrelation der Zeitreihe erkennen (vgl. Abbildung 4.2).

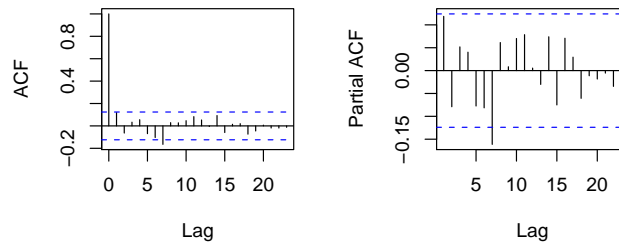


Abbildung 4.2: ACF und PACF des All Ordinaries

Für die Stärke und Richtung der beiden Indices wird

1. die Kreuzkorrelation der Residuen des All Ordinaries und des Dow Jones (ist bereits weißes Rauschen) und

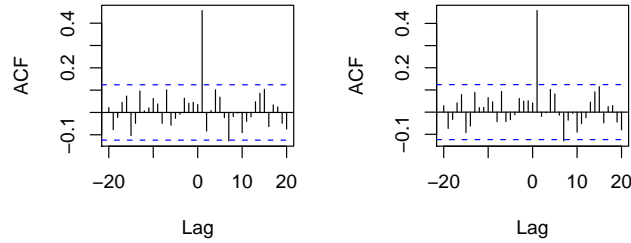


Abbildung 4.3: CCF der Residuen und CCF der Zeitreihen

2. die Kreuzkorrelation der beiden Zeitreihen gebildet (vgl. Abbildung 4.3).

Man sieht, dass die Kreuzkorrelationen der Residuen und der eigentlichen Zeitreihen nahezu identisch sind. Außerdem kann beobachtet werden, dass sich eine zeitliche Verzögerung von einem lag, bzw. Tag ergibt.

Deshalb wird ein Transferfunktionenmodell mit Input Dow Jones, Output All Ordinaries und Zeitverzögerung  $b = 1$  angewandt. Da der Input bereits weißes Rauschen darstellt, müssen die Zeitreihen nicht vorgeweißt werden und die Transferfunktion kann anhand der Kreuzkorrelation der beiden Zeitreihen geschätzt werden (vgl. zweite Abbildung in 4.3). Die Kreuzkorrelation zum lag 1 ist 0.458. Somit ist  $\hat{\tau}_0$

$$w_0 = \frac{\hat{\rho}_{AO,DJ} \hat{\sigma}_{AO}}{\hat{\sigma}_{DJ}} = \frac{0,458 \cdot 0,88}{0,60} = 0,672 \quad (4.33)$$

und die vorläufig identifizierte Transferfunktion

$$\hat{T}(B) = w_0 B^b = 0,672 B \quad (4.34)$$

Für die Noise-Folge

$$\begin{aligned} N_t &= X_{t,AO} - \hat{T}(B)X_{t,DJ} \\ &= X_{t,AO} - w_0 X_{t-1,DJ} \quad \text{für } t \geq 2 \end{aligned}$$

und  $N_t = 0$  für  $t = 1$  wird aufgrund des AIC das ARMA(1,1)-Modell

$$(1 + 0,24B)N_t = 0,40W_{t-1} + W_t \quad (4.35)$$

identifiziert. Schließlich könnten die Koeffizienten nochmals geschätzt werden, indem man die Summe der Fehlerquadrate minimiert, sie ändern sich aber nicht wesentlich. Somit lautet das Transferfunktionenmodell

$$(1 + 0,24B)X_{t,AO} = 0,67(1 + 0,24B)X_{t-1,DJ} + (1 + 0,40B)W_t \quad (4.36)$$

## 4.7 Vorhersage aufgrund eines Transferfunktionenmodells

Um die Vorhersage  $P_n X_{n+h,2}$  des Outputs zu berechnen wird das Transferfunktionenmodell (4.30) umgeschrieben.

$$\begin{aligned}\phi^{(N)}(B)v(B)X_{t2} &= B^b w(B)\phi^{(N)}(B)X_{t1} + v(B)\theta^{(N)}(B)W_t \\ A(B)X_{t2} &= B^b U(B)X_{t1} + V(B)W_t\end{aligned}$$

wobei  $A, U$  und  $V$  Polynome sind:

$$\begin{aligned}A(B) &= 1 + A_1 B + \dots + A_a B^a & a &= p_2 + p \\ U(B) &= U_0 + U_1 B + \dots + U_u B^u & u &= p_2 + q \\ V(B) &= 1 + V_1 B + \dots + V_v B^v & v &= q_2 + p\end{aligned}$$

Wird der Operator  $\tilde{P}_n$  auf diese Gleichung angewandt mit  $t = n + h$  erhält man

$$\tilde{P}_n X_{n+h,2} = - \sum_{j=1}^a A_j \tilde{P}_n X_{n+h-j,2} + \sum_{j=0}^u U_j \tilde{P}_n X_{n+h-b-j,1} + \sum_{j=h}^v V_j W_{n+h-j} \quad (4.37)$$

Die letzte Summe ist Null falls  $h > v$  ist. Die Einschnitt-Prädiktoren  $\tilde{P}_n X_{n+h-b-j,1}$  erhält man indem man die univariaten Zeitreihe  $X_{t,1}$  aufgrund ihres ARMA(p,q)-Modells vorhersagt. Die Werte  $W_j, j \leq n$  werden mit den geschätzten Werten  $\hat{W}_j$  vom Schritt 7 der Identifikation des Transferfunktionsmodells ersetzt.

Für  $h = 1$  erhält man

$$\tilde{P}_n X_{n+1,2} = - \sum_{j=1}^a A_j X_{n+1-j,2} + \sum_{j=0}^u U_j \tilde{P}_n X_{n+1-b-j,1} + \sum_{j=1}^v V_j W_{n+1-j}$$

Die Gleichung (4.37) wird rekursiv gelöst für die Prädiktoren  $\tilde{P}_n X_{n+1,2}, \tilde{P}_n X_{n+2,2}, \tilde{P}_n X_{n+3,2} \dots$

**Bemerkung 9.** Hier wurde eine sogenannte Vorhersage mit unendlich vielen vergangenen Werten gemacht. Im Gegensatz zur Vorhersage mit dem Innovationenalgorithmus wird hier die Konvergenz von  $\theta_{nj} \rightarrow \theta_j$  für  $n \rightarrow \infty$  ( $j = 1, \dots, q$ ) ausgenutzt und  $\theta_{nj}$  mit  $\theta_j$  ersetzt.

## 4.8 Praktischer Teil

In diesem Abschnitt wird versucht eine Vorhersage der maximalen Temperatur und des Tages-temperaturunterschieds anhand der mittleren Globalstrahlung und der mittleren Bewölkung zu machen. D.h. es wird ein bivariater stochastischer Prozess betrachtet.

Für die minimale Nachttemperatur und den Nachttemperaturunterschied wird nur die mittlere Bewölkung während der Nacht herangezogen, da die Globalstrahlung während der Nacht Null ist.

Da für beide Komponentenmodelle ähnliche Ergebnisse gefunden wurden, wird hier nur die Ausführung des standardisierenden Modells gezeigt.

Für die mittlere Tages- und Nachtbewölkung konnten weder Trend noch Saisonalität festgestellt werden, sodass man von einem stationären Prozess ausgehen kann.

### 4.8.1 Tageshöchsttemperatur

#### Transferfunktionenmodell zwischen Tageshöchsttemperatur und Globalstrahlung

Es werden die stationären Restkomponenten der Temperatur und der Globalstrahlung des standardisierenden Modells herangezogen. Zunächst soll die Abhängigkeit der beiden Zeitreihen gezeigt werden. Zu diesem Zweck wird auf beide Zeitreihen ein ARMA-Prozess angewendet

$$\begin{aligned}(1 - 0.71B)Rtmax2_t &= Z_{t1} & Z_{t1} &\sim WN(0, 0.5273) \\ (1 - 0.51B)Rglo2_t &= (1 - 0.23B)Z_{t2} & Z_{t2} &\sim WN(0, 0.9615)\end{aligned}$$

und dann die Kreuzkorrelationsfunktion  $\hat{\rho}_{Z_1, Z_2}(k)$  der beide white noise gebildet. Dabei ist  $\hat{\rho}_{Z_1, Z_2}(k)$  die geschätzte Korrelation zwischen  $Z_{t+k,1}$  und  $Z_{t,2}$ . Man sieht (vgl. Abbildung 4.4), dass die Kreuzkorrelation zum lag 0 signifikant ist. Als zweiter Schritt soll die Richtung des linearen Zusammenhangs überprüft werden. Zu diesem Zweck wird die Kreuzkorrelation zwischen  $Rtmax2$  und  $Rglo2$  betrachtet (vgl. Abbildung 4.5). Da die Kreuzkorrelationen für lag  $k \geq 0$  größer sind als jene für  $h < 0$ , und nur  $\hat{\rho}_{X_1, X_2}(-1)$  signifikant ist, wird  $Rtmax2$  als abhängig von  $Rglo2$  angenommen.

Es wird ein Transferfunktionen-Modell mit Input  $Rglo2$  und Output  $Rtmax2$  angewendet. Die einzelnen Schritte werden im Folgenden kurz dargestellt.

1. Auf den Input-Prozess  $Rglo2$  wird laut AIC ein ARMA(1,1)-Prozess angewendet

$$(1 - 0.51B)Rglo2_t = (1 - 0.23B)Z_{t2}$$

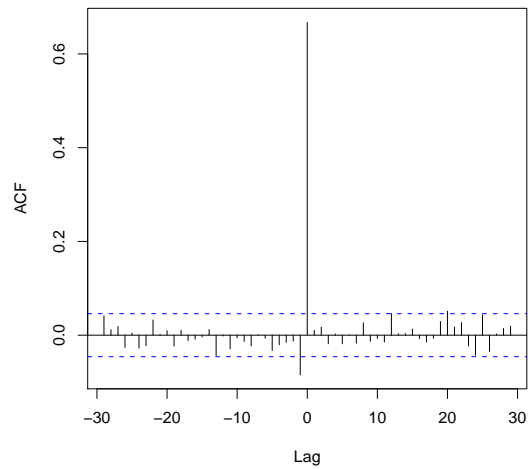


Abbildung 4.4: Kreuzkorrelationsfunktion zwischen den white noise von  $Rtmox2_t$  und  $Rglo2_t$

2. Auf den Output-Prozess wird nun auch ein ARMA(1,1)-Prozess mit  $\phi_1 = 0.51$  und  $\theta_1 = -0.23$  angewendet.

$$Y_t = \frac{(1 - 0.51B)}{(1 - 0.23B)} Rtmox2_t$$

Man erhält die Residuen  $Y_t$ .

3. Nun wird die (bereinigte) Kreuzkorrelation zwischen den Residuen des Inputs  $Z_t$  und den vorgeweißten Residuen des Outputs  $Y_t$  geschätzt (vgl. erste Abbildung in 4.6). Vernachlässigt man die signifikante Kreuzkorrelation zum lag -1, so sind mindestens 6 Kreuzkorrelationen signifikant.
4. Die geschätzte Transferfunktion  $\hat{T}(B) = \sum_{j=0}^3 \hat{\tau}_j B^j$  wird daher als rationale Funktion

$$\hat{T}(B) = \frac{w_0}{1 - v_1 B}$$

dargestellt. Mittels Koeffizientenvergleich erhält man  $w_0$  und  $v_1$

$$\begin{aligned} w_0 &= (1 - v_1 z)(\hat{\tau}_0 + \hat{\tau}_1 z + \hat{\tau}_2 z^2 + \dots) \\ w_0 &= \hat{\tau}_0 \\ 0 &= \hat{\tau}_1 - \hat{\tau}_0 v_1 \Rightarrow v_1 = \hat{\tau}_1 / \hat{\tau}_0 \end{aligned}$$

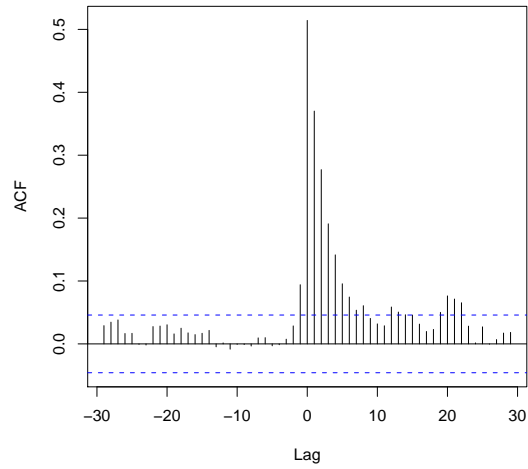


Abbildung 4.5: Kreuzkorrelationsfunktion zwischen  $Rtmax2_t$  und  $Rglo2_t$

Die Koeffizienten  $\tau_0$  und  $\tau_1$  sind

$$\tau_0 = \frac{\hat{\rho}_{YZ}(0)\hat{\sigma}_Y}{\hat{\sigma}_Z} = 0.46$$

$$\tau_1 = \frac{\hat{\rho}_{YZ}(1)\hat{\sigma}_Y}{\hat{\sigma}_Z} = 0.19$$

und die Koeffizienten der rationalen Funktion  $w_0$  und  $v_1$  sind

$$w_0 = \hat{\tau}_0 = 0.46$$

$$v_1 = \hat{\tau}_1/\hat{\tau}_0 = 7.94/17.35 = 0.41$$

5. Nun kann das weiße Rauschen des Transferfunktionenmodells für  $t > 1$  geschätzt werden

$$\hat{N}_t = Rtmax2 - \frac{w_0}{1 - v_1 B} Rglo2$$

$$(1 - v_1 B)\hat{N}_t = (1 - v_1 B)Rtmax2 - w_0 Rglo2$$

$$\hat{N}_t = v_1 \hat{N}_{t-1} + Rtmax2_t - v_1 Rtmax2_{t-1} - w_0 Rglo2_t$$

$$= 0.41 \hat{N}_{t-1} + Rtmax2_t - 0.41 Rtmax2_{t-1} - 0.46 Rglo2_t$$

wobei  $\hat{N}_1 = 0$  gesetzt wird.

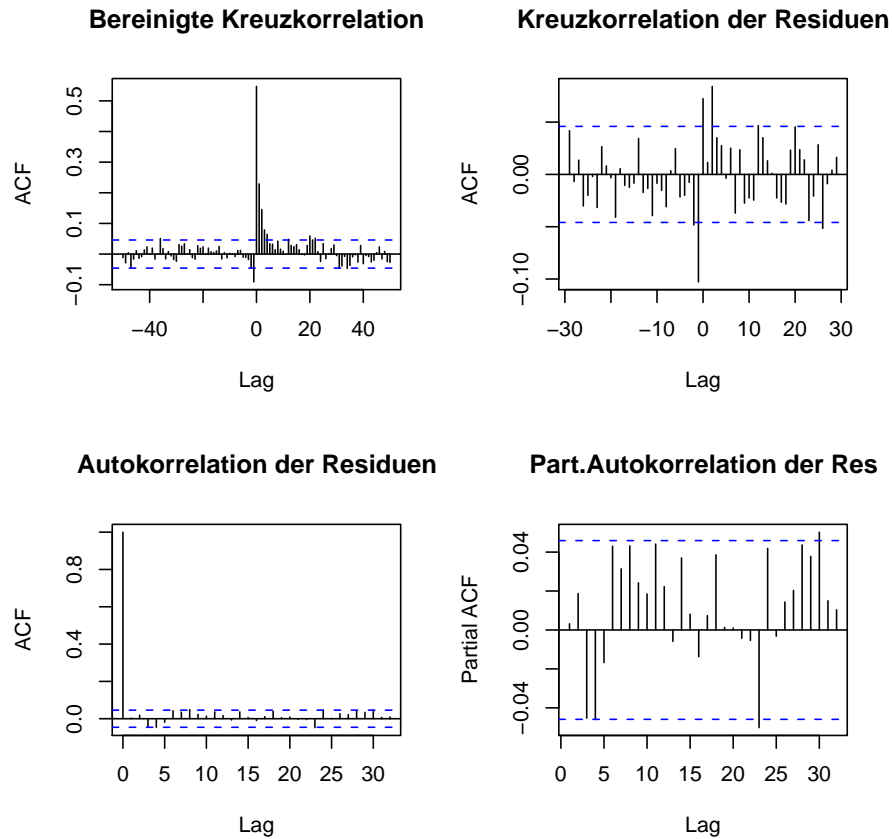


Abbildung 4.6: Transferfunktionenmodell mit Input  $Rglo2$  und Output  $Rtmax2$

6. Schließlich wird auf  $\{N_t\}$  ein ARMA-Prozess angewendet. Es zeigt sich, dass ein AR(1)-Modell bereits gute Ergebnisse für das Transferfunktionenmodell liefert.

$$\begin{aligned} (1 - \phi_1 B)\hat{N}_t &= W_t \\ (1 - 0.77B)\hat{N}_t &= W_t \end{aligned}$$

7. Das vorläufige Modell ist nun

$$\begin{aligned} Rtmax2_t &= \hat{T}(B)Rglo2_t + \hat{N}_t \\ Rtmax2_t &= \frac{w_0}{1 - v_1 B}Rglo2_t + \frac{1}{1 - \phi_1 B}W_t \\ Rtmax2_t &= \frac{0.46}{1 - 0.41B}Rglo2_t + \frac{1}{1 - 0.77B}W_t \\ (1 - 0.77B)(1 - 0.41B)Rtmax2_t &= 0.46(1 - 0.77B)Rglo2_t + (1 - 0.41B)W_t \end{aligned}$$

Für dieses Modell sind die  $\hat{W}_t$  für  $t > 2$

$$\begin{aligned}\hat{W}_t = & 0.41\hat{W}_{t-1} + Rtmax2_t - 1.18Rtmax2_{t-1} + 0.32Rtmax2_{t-2} \\ & - 0.46Rglo2_t + 0.35Rglo2_{t-1}\end{aligned}$$

wobei  $\hat{W}_1 = \hat{W}_2 = 0$  gesetzt werden. Die Parameter  $\phi_1$ ,  $w_0$  und  $v_1$  können nochmals geschätzt werden, indem man die Summe der Quadratfehler  $\sum_{t=3}^n \hat{W}_t^2$  minimiert. Für dieses Modell ändern sich die Koeffizienten jedoch nicht.

8. Anschließend werden noch als Kontrolle die Autokorrelation und partielle Autokorrelation der Residuen  $\{\hat{W}_t, t > 2\}$ , sowie die Kreuzkorrelation zwischen  $\{\hat{W}_t, t > 2\}$  und den Residuen des Input  $\{\hat{W}_t, t > 2\}$  gebildet (siehe Abbildung 4.6).

Das Transferfunktionenmodell für die Restkomponenten der maximalen Temperatur und der Globalstrahlung ist also:

$$(1 - 0.77B)Rtmax2_t = 0.46 \frac{1 - 0.77B}{1 - 0.42B} Rglo2_t + W_t$$

### Transferfunktionenmodell zwischen Tageshöchsttemperatur und Bewölkung

Für die mittlere Tagesbewölkung  $Bew_{Tag}$  konnten weder Trend noch Saisonalität festgestellt werden. Der Mittelwert lag bei 3.27, also etwa bei eine Bewölkungsanteil von 3/8. Für die mittelwertbereinigte Zeitreihe wurde laut AIC ein AR(1)-Modell

$$(1 - 0.40B)(Bew_{Tag,t} - 3.27) = Z_t \quad Z_t \sim WN(0, 4.646) \quad (4.38)$$

ermittelt. Das Bestimmtheitsmaß dieses Modells

$$B = 1 - \frac{\sigma_Z^2}{\sigma_{Bew_{Tag}}^2} = 0.16 \quad (4.39)$$

liegt aber nur bei 16% , sodass man eventuell auch von einem weißen Rauschen ausgehen könnte.

Auch für die maximale Temperatur in Abhängigkeit von der mittleren Tagesbewölkung, wird zunächst die Kreuzkorrelation der Residuen der beiden auf die zwei Zeitreihen abgewendeten ARMA-Modellen gebildet um die Abhängigkeit festzustellen (vgl. erste Abbildung in 4.7). Man sieht, dass die Bewölkung und die Temperatur negativ korreliert sind. Anschließend wird die Kreuzkorrelation der beiden Zeitreihen angeschaut, um die Richtung der Abhängigkeit zu ermitteln (zweite Abbildung in 4.7). Auch hier wird davon ausgegangen, dass die Bewölkung



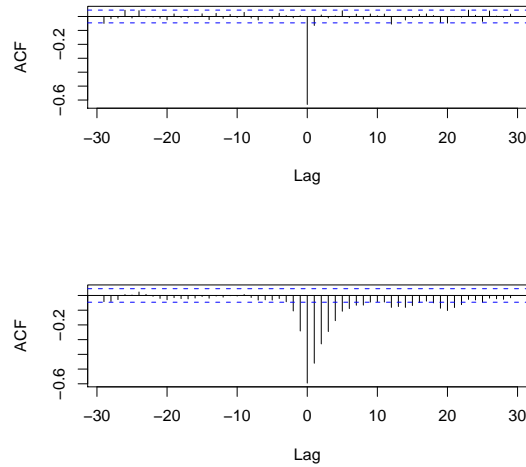


Abbildung 4.7: CCF der Residuen und der Zeitreihen ( $Rtmax2$  und  $Bew_{Tag}$ )

die Temperatur beeinflusst und somit werden die zwei signifikanten Korrelationen zu negativen lags vernachlässigt. Die Transferfunktion wird wieder als rationale Funktion

$$T(B) = \frac{w_0}{1 - v_1 B}$$

gewählt. Das Transferfunktionenmodell für  $Rtmax2$  und  $Bew_{Tag}$  ist

$$(1 - 0.75B)Rtmax2_t = -0.21 \frac{1 - 0.75B}{1 - 0.40B} (Bew_{Tag,t} - 3.27) + (1 - 0.11B)W_t \quad (4.40)$$

## 4.8.2 Tagestemperaturunterschied

### Transferfunktionenmodell zwischen Tagestemperaturunterschied und Globalstrahlung

Auch für den Tagestemperaturunterschied wurde zunächst die Abhängigkeit von der Globalstrahlung anhand der Kreuzkorrelationsfunktionen der Residuen der beiden Restkomponenten  $Rdtr2$  und  $Rglo2$  überprüft (vgl. Abb. 4.8). Die Richtung der Abhängigkeit wurde in der zweiten Abbildung von (4.8) ermittelt. Anschließend wurde das Transferfunktionenmodell für  $Rdtr2$  und  $Rglo2$  ermittelt

$$(1 - 0.81B)Rdtr2_t = 0.70 \frac{1 - 0.81B}{1 - 0.36B} Rglo2_t + (1 - 0.54B)W_t. \quad (4.41)$$

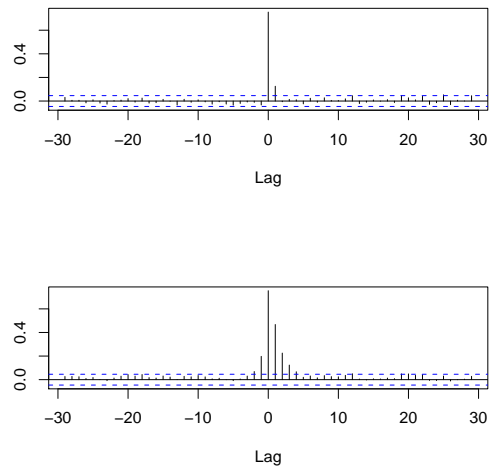


Abbildung 4.8: CCF der Residuen und der Zeitreihen ( $Rdtr2$  und  $Rglo2$ )

### Transferfunktionenmodell zwischen Tagestemperaturunterschied und mittlerer Bewölkung

Bei den Kreuzkorrelationen zwischen Tagestemperaturunterschied und mittlerer Bewölkung sieht man zwar die eindeutig negative Korrelation zwischen den zwei Residuen, die Richtung der Abhängigkeit ist allerdings nicht mehr eindeutig (vgl. dazu Abbildung 4.9).

Trotzdem wird auch hier ein Transferfunktionenmodell angewandt, da die Kreuzkorrelationen zu den positiven lags größer sind als die zu den negativen.

Das Transferfunktionenmodell für  $Rdtr2$  und  $Bew_{Tag}$  lautet somit

$$(1 - 0.86B)Rdtr2_t = -0.33 \frac{1 - 0.86B}{1 - 0.19B} (Bew_{Tag,t} - 3.27) + (1 - 0.67B)W_t \quad (4.42)$$

### 4.8.3 Nachttemperaturunterschied

#### Transferfunktionenmodell zwischen Nachttemperaturunterschied und mittlerer Bewölkung

Wolken verhindern in der Nacht Entweichen von Wärme, deshalb ist anzunehmen, dass die nächtliche Abkühlung zur mittleren Bewölkung während der Nacht negativ korreliert ist [Dai, Trenberth, Karl]. Auch bei diesen Kreuzkorrelationen ist allerdings die Richtung nicht klar ersichtlich (vgl. Abbildung 4.10). Das Transferfunktionenmodell für  $Rntr2$  und  $Bew_{Nacht}$

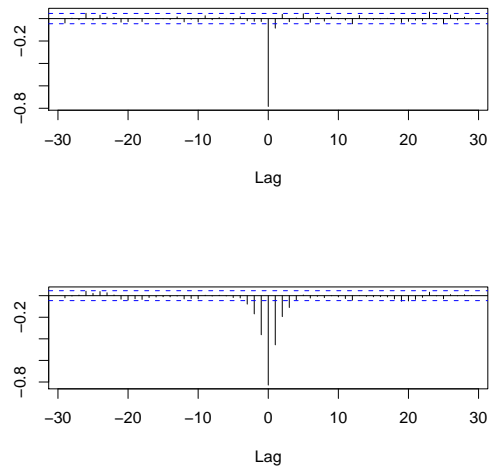


Abbildung 4.9: CCF der Residuen und der Zeitreihen ( $Rdtr2$  und  $Bew_{Tag}$ )

lautet

$$(1 - 0.48B)Rntr2_t = -0.27 \frac{1 - 0.48B}{1 - 0.25B} (Bew_{Nacht,t} - 3.62) + (1 - 0.25)W_t \quad (4.43)$$

#### 4.8.4 Nachttiefsttemperatur

Für die minimale Nachttemperatur werden die Kreuzkorrelationen von den Residuen und die Korrelationen der Zeitreihen zunächst für die 4 Jahreszeiten getrennt betrachtet, da die Kreuzkorrelationen für die gesamten 5 Jahre nicht eindeutig, bzw widersprüchlich waren. Die Kreuzkorrelationen der Residuen für das Jahr 2005 sind in der Abb. 4.11 dargestellt. Es konnten nur für den Winter brauchbare Kreuzkorrelationen gefunden werden, sodass nur für den Winter 04/05 ein Transferfunktionenmodell angewendet wird. Dieses lautet

$$(1 - 0.74B)(Rntr2_t + 0.49) = 0.20(1 - 0.74B)(Bew_{Nacht,t} - 3.89) + W_t \quad (4.44)$$

Die Transferfunktionenmodelle der übrigen 4 Winter sind ähnlich.

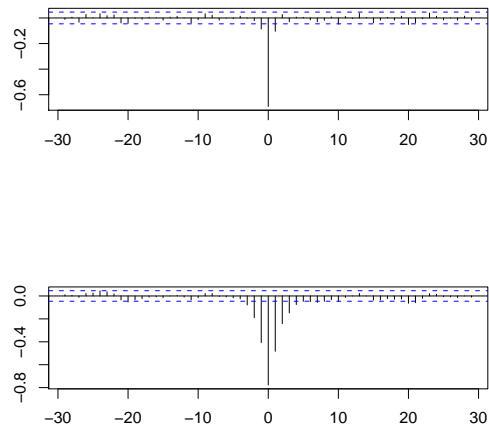


Abbildung 4.10: CCF der Residuen und der Zeitreihen ( $Rntr2$  und  $Bew_{Nacht}$ )

### 4.8.5 Überprüfung des Modells

#### Saisonale Unterschiede

Die Transferfunktionen

$$T(B) = \frac{w_0}{1 - v_1 B}$$

werden nun auf saisonale Unterschiede überprüft. Dabei geht man von meteorologischen Jahreszeiten aus. Der Winter umfasst die Monate Dezember bis Februar, der Frühling die Monate März bis Mai, der Sommer die Monate Juni bis August und der Herbst September bis einschließlich November. Vor allem bei  $w_0$  wurden große Unterschiede zwischen Sommer und Winter beobachtet.

Output und Input	$w_0$ im Sommer			$w_0$ im Winter		
	min	max	Mittel	min	max	Mittel
Tmax-Glo	0.51	0.71	0.58	0.10	0.32	0.16
Tmax-Bew	-0.42	-0.30	-0.34	-0.10	-0.01	-0.07
DTR-Glo	0.62	0.88	0.73	0.48	0.67	0.58
DTR-Bew	-0.50	-0.35	-0.41	-0.29	-0.19	-0.25
NTR-Bew	-0.38	-0.27	-0.32	-0.27	-0.18	-0.22
TMIN-Bew	0.02	0.06	0.04	0.09	0.24	0.17

Die beiden Temperaturunterschiede sind in allen Jahreszeiten von der Globalstrahlung bzw. Bewölkung abhängig. Bei der Tageshöchsttemperatur wurde im Winter kein signifikanter Zu-

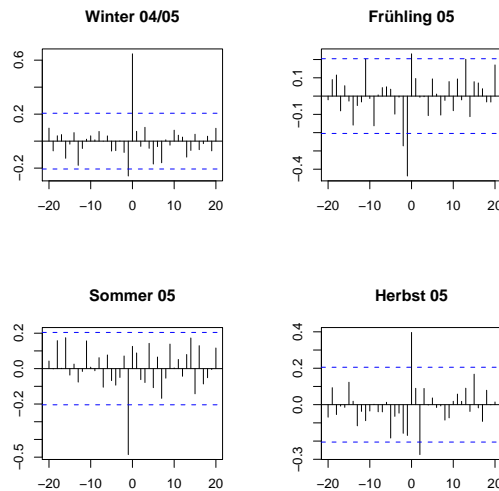


Abbildung 4.11: Kreuzkorrelation der Residuen ( $Rntr2$  und  $Bew_{Nacht}$ )

sammenhang zur Globalstrahlung oder Bewölkung gefunden. Die Nachtstempertemperatur ist hingegen nur im Winter signifikant von der mittleren Nachtbewölkung abhängig.

Die Abhängigkeit von Zeitreihen drückt nur die Abhängigkeit der beiden white noise aus. Zwei Zeitreihen können die gleiche Saisonalität haben, mit beispielsweise höheren Sommer- als Winterwerten, und trotzdem als unabhängig gelten.

### Prognosequalität

Nun werden die gefundenen Transferfunktionenmodelle auf ihre Prognosequalität überprüft. Dabei wird die Einschnittprognose, also die Prognose für den nächsten Tag, mit dem tatsächlich eintretenden Wert verglichen. Bei keinem Transferfunktionenmodell wurde eine zeitliche Verzögerung festgestellt. Um eine Einschnittprognose für den Output zu machen, müsste man auch eine Einschnittprognose für den Input machen. Dies kann anhand des ARMA-Modell des Inputs geschehen.

Für die erste Diagnose wurden aber die tatsächlichen eintretenden Größen des Inputs verwendet. Die auftretenden Fehler werden mit den Fehlern des univariaten Modells verglichen. Es wird wiederum kontrolliert, wieviele Fehler innerhalb des Intervalls  $(-4, 4)$  Grad liegen

	univariat	bivariat
Rtmax2	78%	90% (mit Rglo2) 88% (mit Bew )
Rdtr2	67%	90% (mit Rglo2) 91% (mit Bew )
Rntr2	78%	94% (mit Bew )

Geht man hingegen von einem Fehlerintervall  $(-2, 2)$  Grad aus so liegen nur mehr ca. 60% der Fehler in diesem Intervall. Das Bestimmtheitsmaß des jeweiligen Transferfunktionenmodells und der Anteil der Fehler sind:

	univariat	bivariat	Bestimmtheitsmaß
Rtmax2	46%	57% (mit Rglo2) 59% (mit Bew )	0.73 0.71
Rdtr2	33%	58% (mit Rglo2) 62% (mit Bew )	0.69 0.73
Rntr2	37%	67% (mit Bew )	0.66

Schließlich werden die Ergebnisse untersucht, die mit einem "echten" Transferfunktionenmodell erzielt wurden. Hierbei wurden für die Einschrittprognose nicht die tatsächlichen Inputwerte für den folgenden Tag hergenommen, sondern mit den Methoden der Zeitreihenanalyse geschätzt. Die Ergebnisse unterscheiden sich jedoch nicht wesentlich von der univariaten Zeitreihenanalyse. Es wird wiederum der Anteil der Fehler im Intervall  $(-2, 2)$  Grad Celsius betrachtet.

	univariat	bivariat
Rtmax2	46%	45% (mit Rglo2) 46% (mit Bew )
Rdtr2	33%	35% (mit Rglo2) 32% (mit Bew )
Rntr2	37%	43% (mit Bew )

**Bemerkung 10.** Bei den betrachteten Transferfunktionenmodellen tritt keine zeitliche Verzögerung auf, sodass für eine Prognose des Outputs auch der Input vorhergesagt werden muss. Deshalb sind die Prognosen sehr schlecht.

Das von [Box, Jenkins] vorgestellte Beispiel zum Transferfunktionenmodell hat aber eine Zeitverzögerung von 3 lags, eine Vorhersage des Inputs ist nicht erforderlich. Ein weiterer Kritikpunkt von Young (1974) und Chatfield (1977) war die ausserordentlich hohe Korrelation zwischen Input und Output [Chatfield].

# 5 Tagesgang der Temperatur

## 5.1 Tagesgang der Temperatur mit dem Transferfunktionenmodell

In diesem Kapitel wird untersucht, ob es möglich ist die Temperatur der nächsten Stunden mit dem Transferfunktionenmodell vorherzusagen. Zu diesem Zweck werden nicht die Zeitreihen in 10-Minuten-Intervallen hergenommen, sondern nur die Beobachtungen zur vollen Stunde.

### Sommerwoche

Da bereits bei den letzten Ausführungen in den Sommermonaten bessere Ergebnisse erzielt wurden als in den anderen Jahreszeiten, wird zunächst die Wochen vom 1. bis zum 7. Juli 2007 betrachtet. Die untersuchten Zeitreihen sind die der Temperatur, der Globalstrahlung, der Bewölkung und der Windgeschwindigkeit. Die fehlenden Werte bei der Zeitreihe der Bewölkung wurden linear interpoliert.

### Saisonale Differenzen

Bei der Temperatur, der Globalstrahlung und der Windgeschwindigkeit ist eine Saisonalität zu sehen (vgl. Abb. 5.1). Bei den Zeitreihen im Stundenintervall wird von einer konstanten Saisonsfigur ausgegangen. D.h. die Zeitreihe wird mit saisonalen Differenzen stationär gemacht (vgl. Kapitel 3). Auf die Zeitreihe  $Y_t$  wird also die Transformation

$$X_t = (1 - B^{24})Y_t = Y_t - Y_{t-24} \quad (5.1)$$

angewandt. Die betrachteten Zeitreihen verkürzen sich dadurch um einen Tag. Bei der Bewölkung wird keine Transformation durchgeführt, da sie bereits als stationär angesehen werden kann.

## 5 Tagesgang der Temperatur

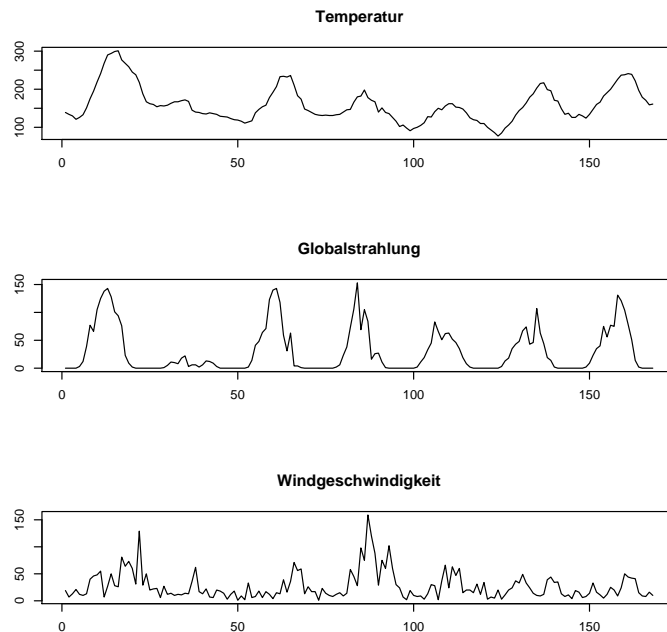


Abbildung 5.1: Temperatur, Globalstrahlung und Windgeschwindigkeit der 1. Juliwoche

### Abhängigkeit der Zeitreihen

Als zweites wird die Abhängigkeit der einzelnen Zeitreihen getestet, dazu werden die Kreuzkorrelationen der Residuen betrachtet (Abb. 5.2). Zwischen Temperatur und Wind wurde keine Korrelation festgestellt. Um eine bessere Aussage über den eventuellen Zusammenhang zu erhalten, wird der Wind grob in Nord-, Ost-, Süd- und Westwind unterteilt. Dabei ist die Windrichtung des Nordwinds zwischen  $315^\circ$  und  $45^\circ$ , da Norden  $0^\circ$  ist. Der Ostwind weht aus  $45^\circ$  bis  $135^\circ$ , der Südwind aus  $135^\circ$  bis  $225^\circ$  und der Westwind aus  $225^\circ$  bis  $315^\circ$ . Aber auch hier kann keine Abhängigkeit festgestellt, zumal man bei einer zweiten Sommerwoche (1. bis 7. August 2007) unterschiedliche signifikante Kreuzkorrelationen gefunden hat.

Bei den Zeitreihen im Jahresverlauf konnte ein Zusammenhang zwischen der Tageshöchsttemperatur und der mittleren Bewölkung festgestellt werden, deshalb werden die sechs Tage (ohne die dazugehörige Nacht) getrennt betrachtet. Der Sonnenaufgang ist für diesen Zeitraum um ca. 5 Uhr und der Sonnenuntergang um ca. 19 Uhr MEZ (Mittleuropäischer Zeit), dies entspricht 4 Uhr und 18 Uhr UTC. Hier konnte ebenfalls kein Zusammenhang festgestellt werden.

Deshalb wird nur ein Transferfunktionenmodell zwischen Temperatur  $T_{W1}$  und Globalstrahlung  $G_{W1}$  betrachtet (vgl. Abbildung 5.3).



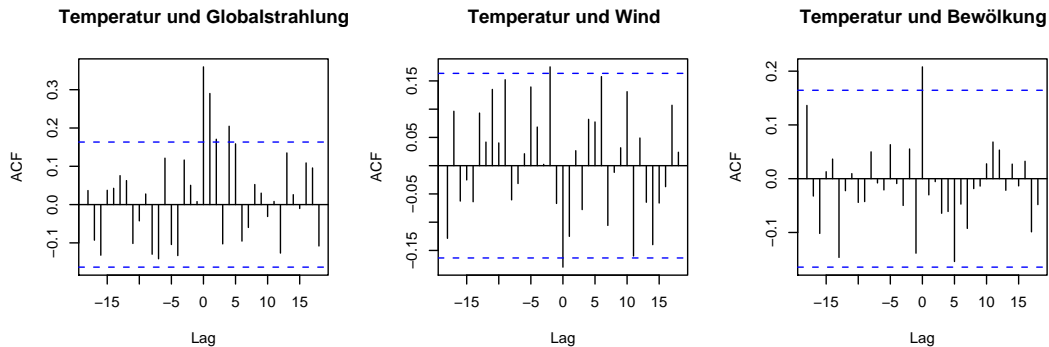


Abbildung 5.2: Kreuzkorrelationen der Residuen

Als Transferfunktionenmodell wurde identifiziert:

$$(1 - B^{24})T_{W1} = (0.18 + 0.23B + 0.24B^2 + 0.15B^3 + 0.23B^4 + 0.22B^5)(1 - B^{24})G_{W1} + \frac{1}{(1 - 0.94B)}W_t$$

Man sieht, dass die Wurzel des autoregressiven Polynoms  $\phi(B) = 1 - 0.94B$  des Transferfunktionenmodells sehr nahe am Einheitskreis ist, sodass sich die Vorhersage sehr stark an die letzte Temperatursausprägung orientieren wird.

Das Transferfunktionenmodell ändert sich auch nicht wenn man nur die Tagstunden heranzieht. Für die Modelldiagnose wird wiederum die Prognosefähigkeit des Modells getestet, die h-Schrittprognosen werden mit den tatsächlich eintretenden Werten verglichen. In diesem Fall wird von einer "echten" Prognose ausgegangen, d.h. die zukünftigen Werte des Inputs werden mit Hilfe des zugehörigen ARMA-Modells prognostiziert.

Bei der Einschrittprognose wurden in 60% der Prognosen ein Fehler  $< 1$  Grad Celsius und 92% kleiner als 2 Grad Celsius gemacht. Bei der 2-Stundenprognose liegen etwa 77% der Prognosen im Intervall  $\pm 2^\circ$ . Im Intervall  $\pm 1^\circ$  liegen aber nur mehr 50%.

Zur Überprüfung des gefundenen Transferfunktionenmodells für eine Sommerwoche wird eine weitere Woche untersucht und zwar vom 1. bis 7. August 2007. Als Transferfunktionenmodell wurde hier identifiziert:

$$(1 - B^{24})T_{W2} = (0.12 + 0.35B + 0.19B^2 + 0.17B^3 + 0.14B^4 + 0.20B^5)(1 - B^{24})G_{W2} + \frac{1}{(1 - 1.22B + 0.30B^2)}W_t$$

Die Koeffizienten sind ähnlich wie bei der Juliwoche und man greift wiederum 5 Stunden zurück. Doch auch hier ist eine Wurzel des autoregressiven Polynoms nahe am Einheitskreis.

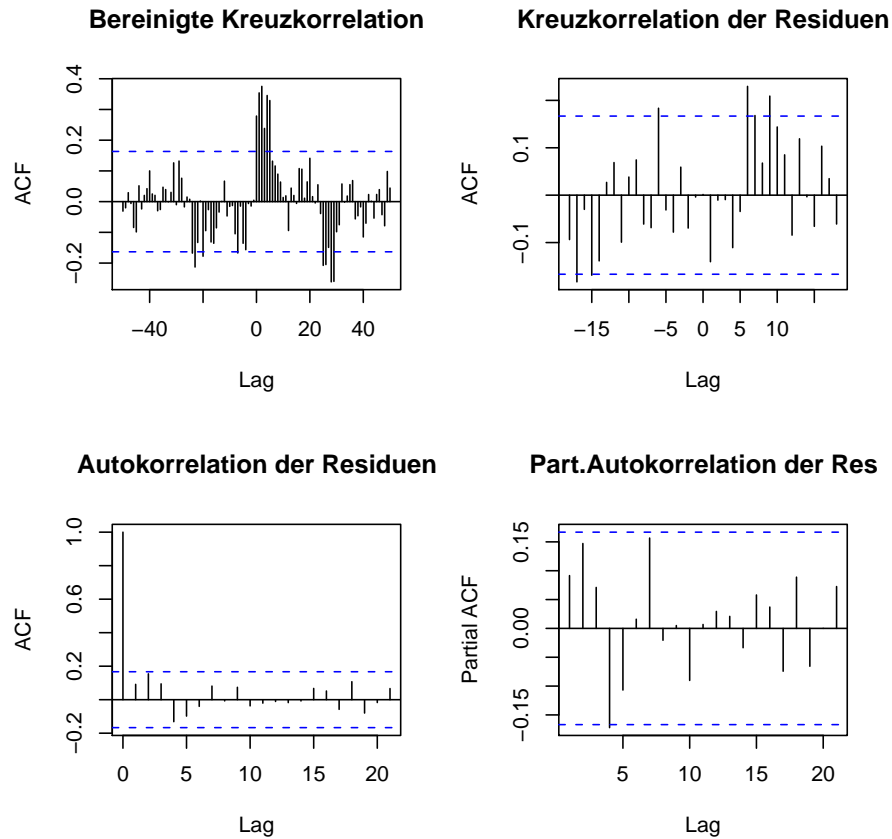


Abbildung 5.3: Auswertung der Transferfunktion für die erste Juliwoche

### Herbstwoche

Da bereits bei den Transferfunktionen für den Jahresverlauf saisonale Unterschiede bemerkt wurden, wird nun als Vergleich für die Sommerwoche die Herbstwoche vom 1. bis zum 7. Oktober betrachtet. Es konnte wieder nur ein Zusammenhang zwischen Temperatur und Globalstrahlung festgestellt gefunden werden. Das zugehörige Transferfunktionenmodell ist

$$(1 - B^{24})T_{W3} = (0.26 + 0.36B)(1 - B^{24})G_{W3} + \frac{1 + 0.31B}{(1 - 0.92B)}W_t$$

Hier greift die Zeitreihe nur mehr eine Stunde zurück, aber auch hier ist die Wurzel des autoregressiven Polynoms sehr nahe am Einheitskreis. Für eine Einstundenprognose liegen 81% der Fehler im Intervall (-1,+1) Grad Celsius. Wird hingegen die Temperatur mit einem Prognosehorizont von 2 Stunden vorhergesagt, so liegen 66% im Intervall  $\pm 1$  Grad Celsius.

### Winterwoche

Nun wird eine Winterwoche betrachtet. Bei den jährlichen Zeitreihen konnte nur ein Zusammenhang zwischen Tagestemperaturunterschied und Globalstrahlung aber nicht zwischen Tageshöchsttemperatur und Globalstrahlung festgestellt werden.

Zwischen der stündlichen Temperatur und der Globalstrahlung konnte kein signifikanter Zusammenhang festgestellt werden.

### Frühlingswoche

Als letzte Woche wird noch eine Woche im Frühling und zwar vom 3. bis 9. April betrachtet. Auch hier trägt die Transferfunktion  $T(B)$  wenig zum Modell bei, denn eine Wurzel des autoregressiven Polynoms des Transferfunktionenmodells ist wiederum sehr nahe am Einheitskreis.

$$(1 - B^{24})T_{W5} = 0.17(1 - B^{24})G_{W5} + \frac{1}{(1 - 1.22B + 0.26B^2)}W_t$$

Dieses Modell bezieht sich nur auf die kontemporäre Globalstrahlung, deshalb sind die Prognosen im Vergleich zu den anderen Wochen schlecht. Bei der Einstundenprognose lagen nur 66% der Fehler im Bereich  $\pm 1$  Grad Celsius und bei der 2-Stundenprognose waren 50% im Bereich  $\pm 1$ .

## 5.2 Temperaturkurve

In Anlehnung an die Artikel von [Betts] und [Dai, Trenberth, Karl] wird der Temperaturgang bei klarem bis leicht bewölktem Himmel mit jenem bei starker Bewölkung verglichen. Dazu werden aus den Daten jene Tage ausgewählt bei denen die mittelhohen Wolken von der Bewölkungsklasse 0 und 1 waren. Vom 1. Juli 2002 bis 31. Mai 2008 verteilen sich die Tage mit klarem Himmel wie folgt:

	Jän.	Febr.	März	April	Mai	Juni	Juli	Aug.	Sept.	Okt.	Nov.	Dez.
Anzahl	24	37	26	19	10	7	12	8	29	31	15	34

Anschließend wurde für jeden Monat der Durchschnitt gebildet.

Die Temperaturgänge für die einzelnen Monate sind in Abb. 5.4 dargestellt. Die Indexskala ist in 10-Minuten-Intervallen angegeben, so hat ein Tag 144 Einträge.

Man sieht, dass die Tageshöchsttemperatur immer in etwa zur gleichen Zeit, nämlich zwischen 13.20 Uhr und 14.30 Uhr erreicht wird. Bei der minimalen Tagestemperatur gibt es hingegen

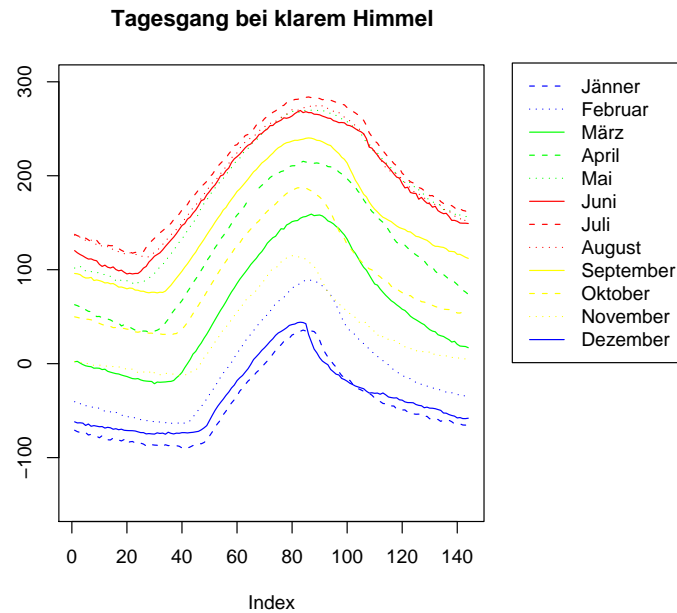


Abbildung 5.4: Tagesgang der Temperatur bei 0 bis 1/8 Gesamtbedeckung

größere Schwankungen, so wird sie im Juni um 3.20 und im Jänner um 7.00 Uhr verzeichnet. Vergleicht man dies mit dem effektiven Sonnenaufgang so wird die Tagestiefsttemperatur im Jänner deutlich vor Sonnenaufgang (zwischen 7.50 und 8.28 Uhr) und im Juni etwa bei Sonnenaufgang (3.50 Uhr) erreicht. Die Zeitangaben sind wiederum in UTC.

Nun wird der Temperaturgang bei fast bedecktem bis bedecktem Himmel angeschaut (Abb. 5.5). Die Verteilung der Tage auf die verschiedenen Monate schaut folgendermaßen aus:

	Jän.	Febr.	März	April	Mai	Juni	Juli	Aug.	Sept.	Okt.	Nov.	Dez.
Anzahl	6	10	10	10	5	0	1	5	9	3	7	16

Die Tageshöchsttemperatur wird schon ab 12.10 bis 13.50 Uhr erreicht, mit Ausnahme vom August, wo die Tageshöchsttemperatur erst um 15.10 Uhr verzeichnet wird. Die Tageshöchsttemperatur wird also zumindest bei diesem Datensatz etwas früher erreicht als bei klarem Himmel. Die Tagestiefsttemperatur wird hingegen meist um Mitternacht verzeichnet, deshalb wird angenommen, dass die Tiefsttemperatur erst in der kommenden Nacht erreicht wird.

Als nächstes wird deshalb der Nachtgang der Temperatur beobachtet, zu diesem Zweck werden Temperaturreihen von 9.00 Uhr bis 9.00 Uhr des darauffolgenden Tages beobachtet. Wiederum wurde zwischen klarem und bedecktem Himmel unterschieden. Die klaren Nächte sind wie folgt auf die Monate verteilt:

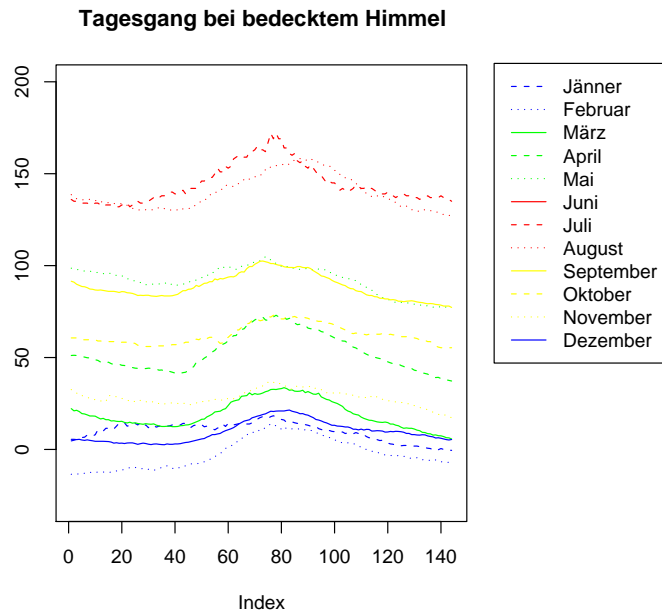


Abbildung 5.5: Tagesgang der Temperatur bei 7/8 bis 8/8 Gesamtbedeckung

	Jän.	Febr.	März	April	Mai	Juni	Juli	Aug.	Sept.	Okt.	Nov.	Dez.
Anzahl	24	32	27	22	8	6	11	6	37	33	17	35

Die Verteilung der Nächte mit bedecktem Himmel ist hingegen

	Jän.	Febr.	März	April	Mai	Juni	Juli	Aug.	Sept.	Okt.	Nov.	Dez.
Anzahl	6	5	8	8	0	0	1	4	6	2	6	16

Im Vergleich zu den Nächten mit klarem Himmel gibt es nur wenige Nächte mit durchgehend bedecktem Himmel. Für diesen Datensatz kann man aber beobachten, dass die Tiefsttemperatur in den Wintermonaten Jänner und Dezember in etwa eine Stunde früher erreicht wird als bei klarem Himmel. In den Sommermonaten konnte hingegen kein Unterschied festgestellt werden.

## 5 Tagesgang der Temperatur

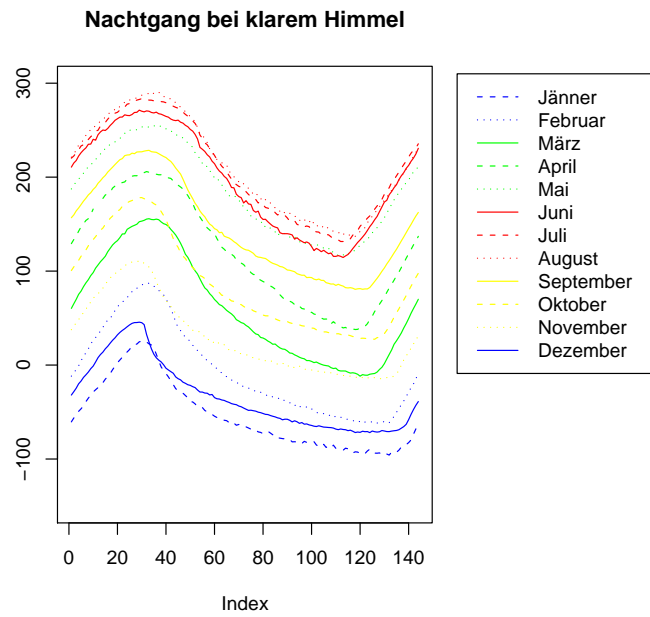


Abbildung 5.6: Nachtgang der Temperatur bei 0 bis 1/8 Gesamtbedeckung

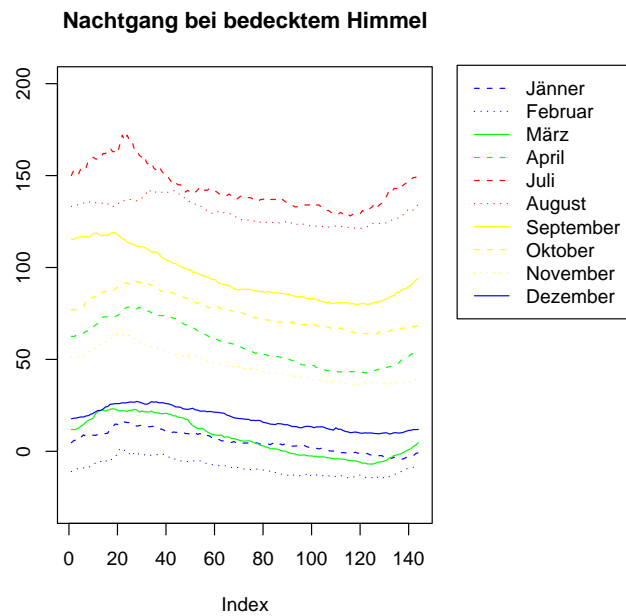


Abbildung 5.7: Nachtgang der Temperatur bei 7/8 bis 8/8 Gesamtbedeckung

# Anhang

Die in dieser Diplomarbeit verwendeten Daten sind auf der beigelegten CD.

Die gemessenen Daten der Lufttemperatur, der Globalstrahlung, der Windgeschwindigkeit und der Windrichtung sind auf der Datei: `tawes11121.dat`

Die beobachteten Daten der Bewölkung sind auf der Datei: `gm_synop_11120.dat`

Die Uhrzeit des Sonnenaufgangs und -untergangs ist auf: `soaufunt.dat`

Die Indices des Dow Jones und All Ordineries sind in der Datei: `djaopc2.dat`

Die Ausarbeitung der Zeitreihen im Jahresverlauf und das Programm für das Transferfunktionsmodell mit der Software R sind auf `jahr.r`

Die Ausarbeitung des 5. Kapitels ist auf den Dateien `woche1.r` bis `woche5.r` sowie auf `stichtage.r` und `stichnächte.r`

# Literaturverzeichnis

- [Betts] *Betts A.*(2006): Radiative scaling of the nocturnal boundary layer and the diurnal temperature range. Erschienen im Journal of geophysical research, Vol. 111.
- [Box, Jenkins] *Box, G.E.P. und Jenkins, G.M.*(1976): Time series analysis, forecasting and control. Holden-Day, San Francisco.
- [Brockwell, Davis. 1991] *Brockwell, P.J. und Davis, R.A.*(1991): Time Series: Theory and Methods, Second Edition. Springer Verlag, New York
- [Brockwell, Davis. 2002] *Brockwell, P.J. und Davis, R.A.*(2002): Introduction to Time Series and Forecasting. Springer Verlag, New York
- [Chatfield] *Chatfield, C.*(1996): The analysis of time series, an introduction. Chapman & Hall, London.
- [Dai, Trenberth, Karl] *Dai A., Trenberth K.E., Karl T.*(August 1999): Effects of Clouds, Soil Moisture, Precipitation, and Water Vapor on Diurnal Temperature Range. Erschienen im Journal of Climate, Volume 12.
- [Hamilton] *Hamilton, J.D.*(1994): Time Series Analysis. Princeton University Press, Princeton.
- [Hartung] *Hartung, J.*(2002): Statistik, 13. Auflage. Oldenbourg, München.
- [Pengfei Li] *Pengfei Li*(2005): Box-Cox Transformations: An Overview. [http://www.stat.uconn.edu/~studentjournal/index\\_files/pengfi\\_s05.pdf](http://www.stat.uconn.edu/~studentjournal/index_files/pengfi_s05.pdf)
- [Stier] *Stier, W.*(2001): Methoden der Zeitreihenanalyse. Springer, Berlin.
- [Wilks] *Wilks, D.S.*(2006): Statistical Methods in the Atmospheric Sciences: An Introduction. Elsevier Inc.