

DIE ANALYSE VON  
ARIMA-PROZESSEN

Wolfgang POLASEK

Forschungsbericht Nr. 140

Juni 1979



# I N H A L T S Ü B E R S I C H T

---

	Seite
1. Die Grundlagen der Box-Jenkins-Methode	1
1.1. Einführung	1
1.2. Operatorenrechnung	7
2. Stochastische Prozesse	8
2.1. Stochastische Prozesse und Autokorrelationsfunktion	8
2.2. Wichtige stochastische Prozesse	11
3. Stationäre Prozesse	16
3.1. Stationäre und invertible ARMA(p,q)-Prozesse und ihre Darstellung	16
3.2. Autoregressive Prozesse	25
3.3. Spezielle autoregressive Prozesse	36
3.4. Moving Average-Prozesse	48
3.5. Autoregressive Moving Average-Prozesse	57
4. Lineare Nichtstationäre Modelle	64
4.1. Allgemeines	64
4.2. Darstellungsformen der ARIMA-Prozesse	68
4.3. Der ARIMA(1,1,1)-Prozeß	81
4.4. Der IMA(1,1)-Prozeß	86
4.5. Der IMA(2,2)-Prozeß	91
4.6. Lineare Differenzgleichung und ARIMA-Prozesse	96
5. Prognose	108
5.1. Theorie der Prognose	108
5.2. Varianz und Korrelation von Prognosen	117
5.3. Darstellungsformen der Prognosefunktion	125
5.4. Die Prognose autoregressiver Prozesse	141
5.5. Prognose von ARIMA(p,d,0)-Prozessen	145
5.6. Die Prognose gemischter und IMA-Prozesse	148
A1 Appendix: Partielle Autokorrelationsfunktion	158
A2 Appendix: Momente von nichtlinearen Funktionen von Zufallsgrößen	171



## V O R W O R T

-----

Diese Arbeit hat ihren Ursprung in einer Vorlesung über die Box-Jenkins-Methode, die ich im Sommersemester 1977 am Institut für Höhere Studien in Wien gehalten habe. Dabei habe ich mich im wesentlichen an das Buch von G.E.P. Box und G.M. Jenkins (1970) gehalten, daneben habe ich die Arbeiten von Granger-Newbold (1970,1977), C.Chatfield (1975) und O.D.Anderson (1976) herangezogen. Im Aufbau folgt diese Arbeit dem Buch von Box-Jenkins, jedoch ohne die Kapitel über Schätzung, Identifikation, saisonale Modelle, Transferfunktionsmodelle, Feedback und Kontrollschemas. In den beiden Appendices werden die partielle Autokorrelationsfunktion aus der allgemeinen Definition der partiellen Autokorrelationskoeffizienten abgeleitet, und das IAZ-System (Interaktives Zeitreihensystem für den Box-Jenkins-Modellbau) vorgestellt. Dieses Programmpaket wurde für den interaktiven Terminalbetrieb am Institut für Höhere Studien adaptiert und weiterentwickelt, und dient u.a. zur vierteljährlichen Prognose der wichtigsten ökonomischen Zeitreihen Österreichs.

Für die Darstellung des Textes habe ich die Definition-Satz-Beispiel-Gliederung gewählt, nicht so sehr wegen der mathematischen Stringenz, sondern zur besseren Übersicht des umfangreichen Stoffes, da für die Identifikation von ARIMA-Prozessen die genaue Kenntnis der wichtigsten Prozesse und deren Eigenschaften von wesentlicher Bedeutung ist.

Im ersten Kapitel erfolgt die Einführung mit der Darstellung der wichtigsten Operatoren und ihren Eigenschaften. Im zweiten Kapitel werden allgemein stochastische Prozesse besprochen und die Klasse der ARIMA-Prozesse definiert. Im dritten Kapitel werden die wichtigsten stationären Prozesse (ARMA-Prozesse) und im vierten Kapitel die wichtigsten nichtstationären (ARIMA)-Prozesse dargestellt. Das fünfte Kapitel gibt eine kurze Einführung in die Prognosetheorie und beschreibt die Eigenschaften der Prognosefunktion der wichtigsten ARIMA-Prozesse.

Für wertvolle Hinweise bin ich den Professoren W. Oberhofer (Regensburg) und M. Deistler (Wien) zu besonderem Dank verpflichtet. Danken möchte ich auch meiner Gattin Edith für das aufopfernde Schreiben des Manuskripts.



## 1. DIE GRUNDLAGEN DER BOX-JENKINS-METHODE

### 1.1. Einführung

#### 1.1.1. Die Bedeutung der Box-Jenkins-Methode

##### Modellbau:

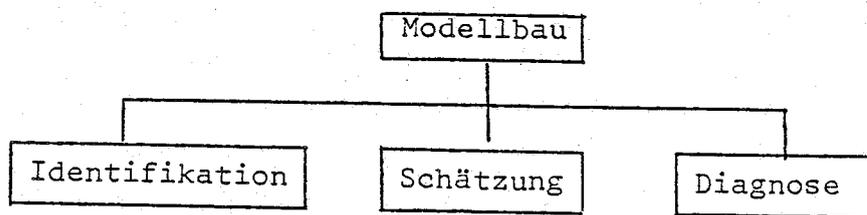
Die Analyse von ARIMA-Prozessen hat sich seit ihrer umfassenden Beschreibung im Buch von Box-Jenkins (1970) als die Methode in der Zeitreihenanalyse entwickelt. Die Gründe für diese Entwicklung sind vielfältig. Obwohl die Darstellung von Zeitreihen als stochastische Differenzgleichungen schon länger bekannt ist, genauso wie die Behandlung von gleitenden Durchschnitten von Zufallsprozessen, gelangte erst in der Darstellung von Box-Jenkins die Methode zu eigenständiger Bedeutung und ist heute als Box-Jenkins-Methode\*) in der Anwendung der Zeitreihenanalyse nicht mehr wegzudenken.

Aufgrund der Prognosequalitäten haben sich ARIMA-Prozesse in der Praxis durchgesetzt, und wegen der dynamischen Eigenschaften hat die Theorie seither in Richtung Transfermodelle und multivariate Modelle einen rapiden Aufschwung erlebt.

Die Gründe hierfür sind folgende:

Zunächst einmal ist zu bemerken, daß die Box-Jenkins-Methode nicht allein die Analyse von einer Klasse stochastischer Prozesse umfaßt (ARIMA-Prozesse), sondern in der Organisation des Verfahrens in Form des iterativen Modellbaus beispielgebend war.

Die drei Säulen des Box-Jenkins-Modellbaus sind:



\*) Genaugenommen müßte man von der Box-Jenkins-Methode zur Schätzung von "Yule-Slutsky-Prozessen" (i.e. AR-MA-Prozesse) sprechen (vgl. Wold (1953)).

Der iterative Modellbau ist komplizierter als bisherige Anwendungen von Methoden in der Zeitreihenanalyse, der Aufwand lohnt sich jedoch. Das ist in Zusammenhang mit der Entwicklung in der Spektralanalyse zu sehen. Die Methoden im Frequenzbereich sind sehr gut geeignet, einen Gesamteindruck über das Verhalten einer Zeitreihe zu geben, sind aber etwas dürftig in der Ausbeute der Detailinformationen. Analysen im Zeitbereich sind umfangreicher und werden leicht unübersichtlich, haben aber den Vorteil, Details genauer spezifizieren zu können. Daraus erklärt sich, warum ein Modellbau im Zeitbereich aufwendiger ist und längeren "try and error"-Prozessen unterworfen ist. Dem steht jedoch der Vorteil gegenüber, im "fine-tuning" des Modellbaus besser abzuschneiden.

#### ARIMA-Prozesse als Vorstufe für Transfermodelle

In der Diskussion über Zeitreihenmodelle tritt oft das Argument auf, Zeitreihenmodelle wären zu "technisch" und wären substanzwissenschaftlichen Interpretationen nicht so zugänglich wie z.B. Regressionsmodelle. Dieser Eindruck mag zwar entstehen, aber nur auf den ersten Blick. Die Box-Jenkins-Methode ist bei genauerer Untersuchung denselben Interpretationen zugänglich wie Regressionsmodelle, nur in allgemeinerer Weise.

Die Box-Jenkins-Methode in Form der Transferfunktionsmodelle kann als Erweiterung der Regressionstheorie für Zeitreihenvariable aufgefaßt werden (dynamische Regression).

Der Zugang zu Transferfunktionsmodellen ist schwieriger, weil Zeitreihen im allgemeinen einer speziellen Struktur folgen, die zum Teil auf endogene Vorgänge der sequentiellen Beobachtung zurückgehen. Wir wollen diesen Teil der Struktur autodynamisch bezeichnen. Der Transferfunktionsmodellbau geht nun folgendermaßen vor sich: Zuerst wird mit Hilfe der univariaten Theorie die autodynamische Struktur der Zeitreihenvariablen spezifiziert und eliminiert; Zwischen den so bereinigten Variablen wird nun die eigentliche Beziehung in der sogenannten Transferfunktion ermittelt, die die Dynamik zwischen den Reihen beschreibt.

Daher ist die Analyse von ARIMA-Prozessen eine notwendige Voraussetzung für die Analyse multivariater Beziehungen von Zeitreihenvariablen.

Die Anwendungsgebiete dieses Systems des Modellbaus lauten wie folgt:

- (1) Prognose von zukünftigen Werten auf Grund von laufenden und vergangenen Werten der Zeitreihe
- (2) Die Bestimmung einer Transfer-Funktion eines linearen Systems. Zwischen der abhängigen Variable  $z_t$  (Outputprozeß) und der unabhängigen Variable  $x_t$  (Inputprozeß) wird eine dynamische Input-Output-Beziehung in Form der sogenannten "Transferfunktion" geschätzt. Diese Methode kann auch auf mehrere Inputzeitreihen erweitert werden. Die Anwendung liegt besonders in der Prognose mit Hilfe von "leading indicators".
- (3) Interventionsanalysen: Mit Hilfe von Dummy-Zeitreihen werden Auswirkungen von Eingriffen in das Zeitreihensystem geschätzt.
- (4) "Feed forward" und "Feed back"-Kontrolle  
Ausgangspunkt dieser Analyse im Rahmen der Kontrolltheorie ist ein Transferfunktionsmodell mit wirksamer Transferfunktion der Kontrollvariable. Weicht die Outputzeitreihe von einem vorgegebenem Ziel zu weit ab, tritt die Kontrollvariable in Kraft. Dazu ist eine optimale Kontrollgleichung zu schätzen. Optimal heißt in diesem Fall: Es ist jene Gleichung zu finden, die den kleinsten mean square error im Outputprozeß liefert.
- (5) Systeme von Transfermodellen  
Einzelne Transferfunktionsbeziehungen werden zu Systemen zusammengefaßt. Die Analyse dieser Vektorprozesse ist der multivariaten Zeitreihenanalyse vorbehalten.

Auf Grund des Herausfilterns von gemeinsamen univariaten Strukturen in Zeitreihenregressionsmodellen kommt man automatisch in den meisten Fällen auf eine geringe Anzahl von Variablen. Es gehen nur jene Zeitreihenvariablen in eine Transferbeziehung ein, die signifikante Beiträge in die gemeinsame Beziehung einbringen. Diese können ebenfalls dynamisch (zeitverzögert) sein, sie sind jedoch nicht autodynamisch, sondern die Dynamik wird in der Transferfunktion spezifiziert.

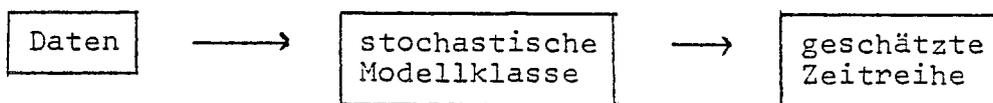
#### Vereinigung von verschiedenen Zeitreihenkonzepten

Weiters tritt die Box-Jenkins-Theorie als "Schmelztiegel" für mehrere statistische Zeitreihenmethoden auf. So können dynamische Ansätze zugleich in nichtstationären und saisonalen Zeitreihen behandelt werden. Bisherige Ansätze sahen in Nichtstationarität

und Saisonalität oft kaum zu vereinbarende Probleme. In der Box-Jenkins-Methode ist der Übergang zu kürzeren Erhebungsperioden in Zeitreihen, wenn sie systematische Muster aufweisen (Saison), ein Informationsgewinn. Man erhält kleinere Varianzen der Residuen und kleinere Prognosefehler; auch Änderungen werden schneller entdeckt. Dies muß z.B. bei Regressionsverfahren nicht immer der Fall sein. Durch Differenzenbildung bestimmter Ordnungen ist es gelungen, das Phänomen der Nichtstationarität konstruktiv in den Griff zu bekommen. Daneben können all diese Probleme sinnvoll mit dynamischen und sonstigen Strukturen (Kontrollvariablen, leading indicators) verknüpft sein.

### 1.1.2. Die Grundidee des BJ-Modellbaus

Ausgangspunkt unserer Betrachtungen sind die Daten und die Klasse von statistischen Modellen, die wir anwenden wollen.



Es gibt eine Vielzahl von Modellen der stochastischen Modellklasse, die auf beobachtete Daten passen. Welche ist im Zweifel zu nehmen? Zur Lösung dieses Problems haben Box-Jenkins das sogenannte "Prinzip der Parsimonität" für den iterativen Modellbau eingeführt.

#### Prinzip der Parsimonität

Man wähle jenes Modell, das die kleinste Anzahl von Parametern für eine adäquate Darstellung benötigt. Man kann das Prinzip der Parsimonität auch als Prinzip der "Sparsamkeit" in bezug auf Parameter nennen.

#### Def. 1.1.1.: Adäquate Darstellung

Eine adäquate Darstellung einer Zeitreihe durch ein Modell  $M$  ist dann erreicht, wenn die folgende Menge gleichwertiger Kriterien möglichst gut erfüllt ist:

- 1) Die geschätzten Residuen folgen einem "white-noise"-Prozeß
  - 1.1. Es gibt keine Autokorrelation  
(das Spektrum der Residuen ist "white noise")
  - 1.2. Die Residuen sind homoskedastisch
- 2) Die geschätzten Parameter sind auf  $\alpha$ -% Niveau signifikant
- 3) Die Quadratsumme der Fehler (geschätzte Residuen) ist minimal
- 4) Der einstufige Prognosefehler für die letzte Periode ist minimal

Beispiel 1.1.:

Es seien  $z_t$  die beobachtete Zeitreihe und  $a_t$  die geschätzten Fehler. Es wurden zwei Modelle geschätzt:

$$(a) \quad z_t = a_t - \theta_1 a_{t-1} \quad (1.1.1)$$

Parameter:  $\theta_1$

$$(b) \quad z_t = a_t - \theta'_1 a_{t-1} - \theta'_2 a_{t-2} - \theta'_3 a_{t-3} \quad (1.1.2)$$

Parameter:  $\theta'_1, \theta'_2, \theta'_3$

Sind beide "adäquat" geschätzt, so nehme man wegen des Prinzips der Parsimonität Prozeß (a), d.h. erfüllen beide Prozesse die Kriterien der adäquaten Darstellung etwa gleich gut.

Beispiel 1.2.:

$$(a) \quad z_t = a_t - \theta'_1 a_{t-1} - \theta'_2 a_{t-2} \quad (1.1.3)$$

Parameter:  $\theta_1, \theta_2$

$$(b) \quad z_t - \phi_1 z_{t-1} = a_t - \theta_1 a_{t-1} \quad (1.1.4)$$

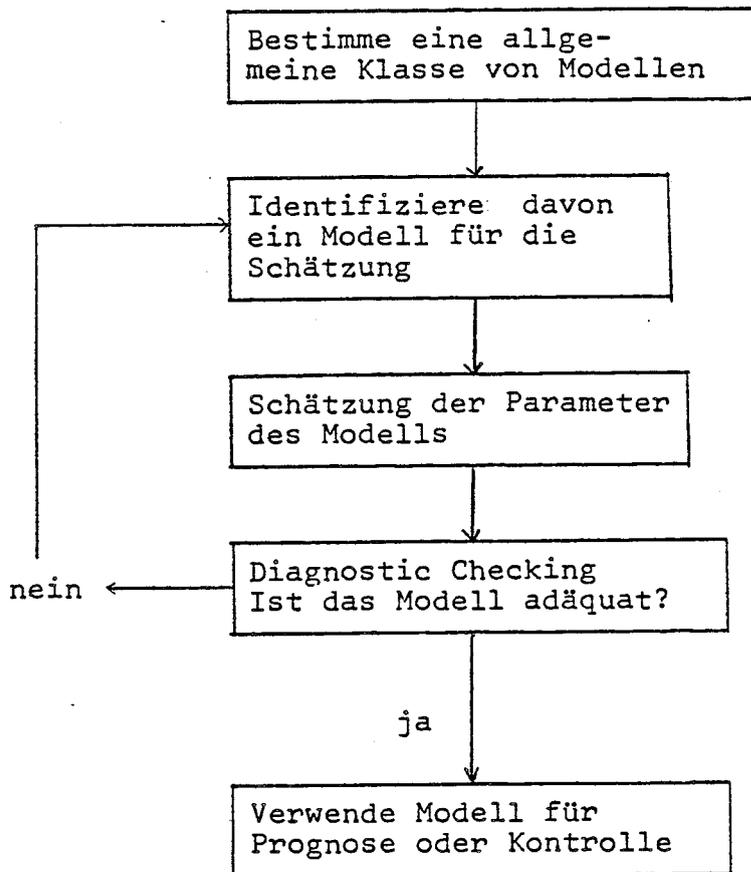
Parameter:  $\phi_1, \theta_1$

Beide Prozesse haben die gleiche Anzahl von Parametern, gehören jedoch getrennten Modellklassen an. Falls es keine Auswahlkriterien mehr gibt, die zwischen diesen beiden Prozessen entscheiden können, sind beide Darstellungen zulässig.

### 1.1.3. Iterativer Modellbau

Die Bedeutung der BJ-Methode liegt nicht nur in der Beschreibung von Zeitreihenmodellen mit Hilfe von ARIMA-Prozessen (bzw. deren Schätzung), sondern auch in der Philosophie des Modellbaus. Durch iterativen Modellbau versucht man, ein möglichst gutes Zeitreihenmodell zu erhalten. Als Voraussetzung dafür muß es eine Modellklasse geben, in der durch Iteration und ein Selektionskriterium ein derartiges Finden von Modellen möglich ist.

Der iterative Modellbau läßt sich durch folgendes Flußdiagramm darstellen:



## 1.2. Operatorenrechnung

Im folgenden seien kurz die Operatoren zusammengestellt, die u.a. zur kompakten Notation von ARIMA-Modellen dienen.

$z_t$  sei eine Zeitreihe mit  $t \in \mathbb{Z}$ .

Der Backwardshiftoperator  $B$  verschiebt eine Zeitreihe um ein Element in die Vergangenheit.

$$Bz_t = z_{t-1} \quad (1.2.1)$$

und

$$B^m z_t = z_{t-m} \quad m \in \mathbb{N}$$

Der Forewardshiftoperator  $F$  verschiebt eine Zeitreihe um ein Element in die Zukunft

$$Fz_t = z_{t+1} \quad (1.2.2)$$

und genauso

$$F^m z_t = z_{t+m}$$

$F$  kann als inverser Operator zu  $B$  angesehen werden:

$$F(Bz_t) = z_t = B(Fz_t)$$

Der Differenzenoperator  $\nabla$  gibt die Differenzen zu vergangenen Elementen an

$$\nabla z_t = z_t - z_{t-1} \quad (1.2.3)$$

und man definiert rekursiv die Differenz  $d$ -ter Ordnung

$$\nabla^d z_t = \nabla \nabla^{d-1} z_t \quad d > 1 \quad (1.2.4)$$

Der Summationsoperator  $S$  gibt die Summe der Elemente bis zum Zeitpunkt  $t$  an, falls die Summe existiert.

$$Sz_t = \sum_{j=0}^{\infty} z_{t-j} = \sum_{j=-\infty}^{\infty} z_t \quad (1.2.5)$$

wobei je nach Bedarf die "Rückwärts-" oder "Vorwärtsschreibweise" verwendet wird. Analog versteht man unter  $S^d z_t = SS^{d-1} z_t$ . Als inverser Operator zu  $S$  kann  $\nabla^{-1}$  ("Antidifferenz") angesehen werden, obwohl  $\nabla^{-1}$  bis auf einen konstanten Faktor nicht eindeutig ist (vgl. Brand (1966), 377).

## 2. STOCHASTISCHE PROZESSE

---

### 2.1. Stochastische Prozesse und Autokorrelationsfunktion

Das wahrscheinlichkeitstheoretische Modell für beobachtete Zeitreihen bilden die stochastischen Prozesse.

Def. 2.1.1.: Stochastische Prozesse

Sei  $\{X_t\}$  eine Familie<sup>\*)</sup> von Zufallsvariablen und  $t \in Z$ , der Zeitindex, stammt aus den ganzen Zahlen. Dann heißt

$$S = \{X(t), t \in T\} := X_t$$

stochastischer Prozeß. Die Zufallsvariablen sind auf dem gemeinsamen Stichprobenraum  $\Omega$  definiert.

Als Realisation eines stochastischen Prozesses versteht man das Eintreten eines bestimmten  $\omega$ , das alle Realisationen der Familie von Zufallsvariablen  $X_n$  umfaßt.

Für ein bestimmtes  $\omega$  wird die Folge  $X_t(\omega)$  festgelegt und für alle  $\omega \in \Omega$  erhalten wir eine Familie solcher Folgen. In der Zeitreihenanalyse beobachtet man meistens nur eine Realisation  $X_t = X_t(\omega)$ .

Für die Analyse von Zeitreihen sind hauptsächlich stationäre und davon die schwach stationären Prozesse von Relevanz.

---

\*) "Familie" läßt im Gegensatz zu "Menge" gleiche Elemente zu.

Def. 2.1.2.: Schwach stationärer (kovarianzstationärer) Prozeß

Gegeben sei ein stochastischer Prozeß  $X_t$ .  $X_t$  heißt kovarianzstationär, falls

a)  $E(X_t) = \mu$

b)  $\text{Cov}(X_t, X_s) < \infty \quad t, s \in T$

c)  $\text{Cov}(X_t, X_s) = f(|t-s|)$

d.h. die Kovarianz für beliebige  $t$  und  $s$  ist unabhängig vom jeweiligen Zeitpunkt und eine Funktion der Länge des Abstandes  $|t-s|$ .

Def. 2.1.3.: Autokovarianzfunktion (ACovF)

Unter der Autokovarianzfunktion verstehen wir die Folge

$$\begin{aligned} \gamma_k &= \text{Cov}(z_t, z_{t+k}) \\ &= E(z_{t-\mu})(z_{t+k-\mu}) \quad k=0,1,2,\dots \end{aligned} \quad (2.1.1)$$

für  $k=0$  gilt

$$\gamma_0 = E(z_{t-\mu})^2 = \sigma_z^2 \quad (2.1.2)$$

$\gamma_k$  ist die Kovarianz der Zufallsgröße  $z_t$  zum Zeitpunkt  $t$  mit der Zufallsgröße  $z_{t+k}$  zum Zeitpunkt  $t+k$ .

Def. 2.1.4.: Autokorrelationsfunktion (ACF)

Aus der ACovF erhält man die Autokorrelationsfunktion durch

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad k=0,1,2,\dots \quad (2.1.3)$$

Es gilt  $\rho_0 = 1$  und  $|\rho_k| < 1$  (2.1.4)

Die ACF ist im Gegensatz zur ACovF ein dimensionsloses Maß für die Abhängigkeit zweier Zeitreihenwerte, d.h. sie hängt nicht vom Maßstab der Zeitreihe ab (Skalierung).

### Die Schätzung der Autokovarianzfunktion

Aus einer Realisation eines stochastischen Prozesses soll die Autokovarianzfunktion geschätzt werden. Man bezeichnet diese auch als empirische Autokovarianzfunktion

$$c_k = \hat{\gamma}_k = \frac{1}{N-k} \sum_{i=1}^{N-k} (z_i - \bar{z})(z_{i+k} - \bar{z}) \quad k=0,1,2,\dots \quad (2.1.5)$$

### Schätzung der Autokorrelationsfunktion

Analog erhält man die empirische ACF durch

$$r_k = \hat{\rho}_k = \frac{c_k}{c_0} \quad (2.1.6)$$

### Autokovarianz- und Autokorrelationsmatrizen

#### Def. 2.1.5.: Autokovarianzmatrix

Eine Autokovarianzmatrix der Ordnung  $n$  setzt sich aus Autokovarianzen in sogenannter Toeplitz-Struktur zusammen.

$$P_n = \begin{pmatrix} \gamma_0 & \gamma_1 & \dots & \dots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \dots & \dots & \gamma_{n-2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \gamma_{n-1} & \dots & \dots & \dots & \gamma_0 \end{pmatrix} \quad (2.1.7)$$

#### Def. 2.1.6.: Autokorrelationsmatrix

$$P_n = \begin{pmatrix} 1 & \rho_1 & \dots & \dots & \rho_{n-1} \\ \rho_1 & 1 & \dots & \dots & \rho_{n-2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \rho_{n-1} & \dots & \dots & \dots & 1 \end{pmatrix} \quad (2.1.8)$$

Es gilt die Beziehung

$$\Gamma_n = \sigma_z^2 P_n$$

Ersetzt man  $\gamma_k$  durch  $c_k$  und  $\rho_k$  durch  $r_k$ , so erhält man die empirischen Autokovarianz- und Autokorrelationsmatrizen  $C_n$  und  $R_n$ .

## 2.2. Wichtige stochastische Prozesse

Def. 2.2.1.: "White noise"-Prozeß oder reiner Zufallsprozeß

Ein diskreter Prozeß  $a_t$  heißt "white noise"-Prozeß ("weißes Rauschen"), wenn die Zufallsgröße  $a_t$  unabhängige identisch verteilte Variablen sind. Mittelwert und Varianz sind per definitionem konstant.

$$E(a_t) = \mu$$

$$\text{Var}(a_t) = \sigma^2 \quad (2.2.1)$$

$$\gamma_k = \text{Cov}(a_t, a_{t+k}) = 0 \quad k \in \mathbb{Z}$$

Mittelwert und ACovF hängen nicht von  $t$  ab, somit ist der Prozeß stationär der 2. Ordnung; darüber hinaus ist er sogar streng stationär.

Die ACF ist gegeben durch

$$\rho_k = \begin{cases} 1 & k = 0 \\ 0 & k \in \mathbb{Z} \end{cases} \quad (2.2.2)$$

Def. 2.2.2.: Random walk-Prozeß

Ein Prozeß  $\{X_t\}$  heißt random walk, falls gilt

$$X_t - X_{t-1} = a_t \quad \forall t \quad (2.2.4)$$

d.h. die ersten Differenzen eines random walk-Prozesses  $a_t$  sind ein white noise-Prozeß (mit Mittelwert  $\mu$  und Varianz  $\sigma_a^2$ ).

Beginnt ein random walk-Prozeß bei  $X_0=0$ , so ist  $X_1=a_1$  und allgemein

$$X_t = \sum_{i=1}^t a_i \quad t=1,2,\dots \quad (2.2.5)$$

In diesem Fall lauten die ersten beiden Momente

$$\begin{aligned} E(X_t) &= t\mu \\ \text{Var}(X_t) &= t\sigma_a^2 \end{aligned} \quad (2.2.6)$$

Da Mittelwert und Varianz Funktionen der Zeit  $t$  sind, ist der Prozeß nicht stationär; die ersten Differenzen sind dagegen stationär.

Def. 2.2.3.: Autoregressiver (AR(p))-Prozeß

Ein stochastischer Prozeß  $\{z_t\}$  heißt autoregressiver Prozeß der Ordnung  $p$  ( $p \in \mathbb{N}$  und  $\phi_p \neq 0$ ), wenn die laufenden Werte durch  $p$  vergangene aufeinanderfolgende Werte erklärt werden:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t \quad (2.2.7)$$

Der Störterm  $a_t$  folgt einem white noise-Prozeß.

Def. 2.2.4.: AR-Operator

Das Operatorpolynom der Ordnung  $p$  heißt AR-Operator:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad \phi_i \in \mathbb{R}, \phi_p \neq 0 \quad (2.2.8)$$

Damit schreiben wir den AR(p)-Prozeß kurz

$$\phi(B)z_t = a_t \quad (2.2.9)$$

Zur Schätzung dieses Modells sind  $p+2$  Parameter notwendig  $(\mu, \phi_1, \dots, \phi_p, \sigma_a^2)$ .

Def. 2.2.5.: Moving Average (MA(q))-Prozeß

Ein stochastischer Prozeß heißt Moving Average Prozeß der Ordnung  $q$  ( $q \in \mathbb{N}$  und  $\theta_q \neq 0$ ), falls der laufende Wert der Zeitreihe  $z_t$  durch eine gewichtete Summe der vergangenen Störungen beeinflusst wird.

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2.2.10)$$

$\theta_i \in \mathbb{R}$  und  $\{a_t\}$  ist der white noise-Prozeß der Störungen.

Def. 2.2.6.: MA-Operator

Das Operatorpolynom vom Grad  $q$  heißt MA-Operator:

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (2.2.11)$$

Somit schreiben wir den MA-Prozeß kurz als

$$z_t = \theta(B)a_t \quad (2.2.12)$$

Die Anzahl der Parameter ist  $q+2$  ( $\mu, \theta_1, \dots, \theta_q, \sigma_a^2$ ).

Mit ARMA(p,q)-Prozessen bezeichnet man gemischte Autoregressive Moving Average-Prozesse.

Def. 2.2.7.: ARMA(p,q)-Prozeß

Der ARMA-Prozeß setzt sich gleichzeitig aus autoregressiven und Moving Average-Gliedern zusammen:

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2.2.13)$$

Mit MA- und AR-Operatorpolynomen lautet der Prozeß kurz

$$\phi(B)z_t = \theta(B)a_t \quad (2.2.14)$$

Anzahl der Parameter:  $p+q+2$  ( $\mu, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma_a^2$ ).

ARIMA(p,d,q)-Prozesse (Autoregressive integrierte Moving-Average-Prozesse) bezeichnen eine Klasse von nichtstationären Prozessen, die aus ARMA-Prozessen erhalten werden kann. Wie noch später gezeigt wird, kann man diese Klasse von nichtstationären Zeitreihen durch Differenzenbildung stationär machen.

Ist  $z_t$  nicht stationär, so existiert ein  $d \in \mathbb{N}$ , sodaß

$$w_t = \nabla^d z_t = (1-B)^d z_t \quad (2.2.15)$$

ein stationärer Prozeß ist.

Operatorschreibweise: Zur Darstellung von nichtstationären Zeitreihen verwenden wir den verallgemeinerten AR-Operator:

Def. 2.2.8.: Unter dem ARI-Operatorpolynom verstehen wir den verallgemeinerten AR-Operator

$$\varphi(B) = \phi(B)(1-B)^d \quad d \in \mathbb{N} \quad (2.2.16)$$

Def. 2.2.9.: ARIMA(p,d,q)-Prozeß

Ein ARIMA(p,d,q)-Prozeß ist ein ARMA(p,q)-Prozeß mit verallgemeinertem AR-Operator:

$$\varphi(B)z_t = \theta(B)a_t \quad (2.2.17)$$

oder in getrennter Schreibweise:

$$\phi(B)w_t = \theta(B)a_t \quad (2.2.18)$$

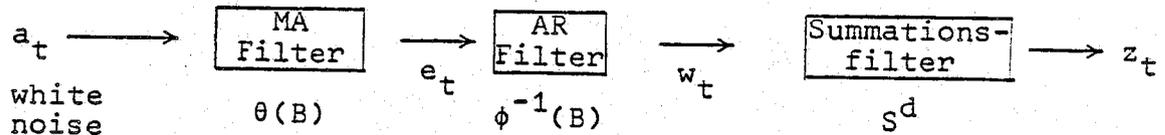
$$w_t = \nabla^d z_t \quad (2.2.19)$$

Anmerkung: Das Wort "integriert" sollte vielmehr "summiert" heißen und leitet sich von der Tatsache her, daß der inverse Operator zum Differenzenoperator  $\nabla$  der Summationsoperator  $S$  ist.

$$\begin{aligned} w_t &= \nabla^d z_t \\ z_t &= S^d w_t \end{aligned} \quad (2.2.20)$$

Black-Box-Schema für den ARIMA-Prozeß

Der allgemeine ARIMA-Prozeß kann als Summe dreier hintereinandergeschalteter Filtermodelle gedacht werden:



mit den jeweiligen Teilprozessen  $e_t = \theta(B)a_t$ ,  $w_t = \phi^{-1}(B)e_t$ ,  $z_t = S^d w_t$ , bzw.

$$z_t = S^d \phi^{-1}(B) \theta(B) a_t \quad (2.2.21)$$

Def. 2.2.10.: Spektrum einer Zeitreihe

Sei  $\gamma_k$  die ACovF einer Zeitreihe  $\{z_t\}$ , dann ist das Spektrum  $SP(f)$  gegeben durch

$$SP(f) = 2\gamma_0 + 4 \sum_{k=1}^{\infty} \gamma_k \cos 2\pi f k \quad 0 \leq f \leq 1/2$$

Das Spektrum ist eine spezielle Form der Fouriertransformation, da  $\gamma_k$  gerade und diskret ist. Dieselbe Information, die im Zeitbereich durch  $\gamma_k$  gegeben ist, wird im Frequenzbereich durch das Spektrum aufbereitet, das als eine Art Varianzanalyse über die Frequenzen  $f=1/T$  (T Periodenlänge) angesehen werden kann.

### 3. STATIONÄRE PROZESSE

-----

#### 3.1. Stationäre und invertible ARMA(p,q)-Prozesse und ihre Darstellung

---

Zunächst unterscheiden wir zwei Darstellungen von ARMA(p,q)-Prozessen und die daran geknüpften Stationaritäts- und Invertibilitätsbedingungen, sowie ihre Beziehungen zueinander.

##### 3.1.1. Die Random-Shock-Form und Stationarität

Def. 3.1.1.: Die RS-Form (Random-Shock-Form)

eines ARMA(p,q)-Prozesses ist die Darstellung als lineares Filtermodell:

$$z_t = \psi(B)a_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} \quad (3.1.1.)$$

In der "Random-Shock"-Form wird der Prozeß  $z_t$  aus einer gewichteten Summe vergangener Störterme (Random Shocks)  $a_t$  erklärt ( $a_t$  ist ein white-noise-Prozeß).

$\psi(B)$  ist die Transferfunktion des linearen Filters und erlaubt zwei Interpretationen:

- a) Die Koeffizienten  $\psi(B)$  geben an, welches Gewicht vergangene Störglieder auf den laufenden Wert der Zeitreihe haben;
- b) Die Transferfunktion kann als erzeugende Funktion\*) der Filterkoeffizienten  $\psi_j$  aufgefaßt werden

$$\psi(s) = \sum_{j=0}^{\infty} \psi_j s^j \quad s \in \mathbb{R} \quad (3.1.2.)$$

Um die Konvergenz dieser Potenzreihe zu gewährleisten, definieren wir wie folgt:

\*) Box-Jenkins (1970) verwenden nicht explizit diese Unterscheidung zwischen Transferfunktion und erzeugender Funktion. Der Backshiftoperator B ist zugleich Dummy-Variable der entsprechenden erzeugenden Funktion.

Def. 3.1.2.: Stationaritätsbedingung

Der Prozeß  $z_t = \psi(B)a_t$  ist stationär, falls die Transferfunktion

$$\psi(s) = \sum_{j=0}^{\infty} \psi_j s^j \quad \text{für } |s| < 1 \quad (3.1.3)$$

konvergiert.

3.1.2. Die invertierte Form und Invertibilitätsbedingung

Def. 3.1.3.: Die  $\pi$ -Form (invertierte Form)

eines ARMA(p,q)-Prozesses ist die Darstellung in der Form

$$\begin{aligned} z_t &= \pi_1 z_{t-1} + \pi_2 z_{t-2} + \dots + a_t = \\ &= \sum_{j=1}^{\infty} \pi_j z_{t-j} + a_t \end{aligned} \quad (3.1.4)$$

bzw.

$$\pi(B)z_t = a_t \quad (3.1.5)$$

mit

$$\begin{aligned} \pi(B) &= 1 - \pi_1 B - \pi_2 B^2 - \dots \\ &= 1 - \sum_{i=1}^{\infty} \pi_i B^i \end{aligned} \quad (3.1.6)$$

Analog zu (3.1.2) ist  $\pi(s)$  die erzeugende Funktion der  $\pi$ -Koeffizienten.

Die "duale" Form zur RS-Form ist ersichtlich. In der  $\pi$ -Darstellung wird der laufende Wert einer Zeitreihe durch eine Gewichtung vergangener Werte plus einem Störterm erklärt. der Name "invertierte Form" kommt daher, daß das  $\pi$ -Polynom als inverses des  $\psi$ -Polynoms dargestellt werden kann (vgl. Satz 3.1.1.).

An die invertierte Form schließt sich die

Def. 3.1.4.: Invertibilitätsbedingung

Ein linearer stochastischer Prozeß heißt invertierbar, falls die erzeugende Funktion der  $\pi_j$ -Koeffizienten

$$\pi(s) = 1 - \theta_1 s - \theta_2 s^2 - \dots \quad (3.1.7)$$

für  $|s| < 1$  (innerhalb des Einheitskreises) konvergiert.

Die Dualität von Stationaritäts- und Invertibilitätsbedingung ist damit offensichtlich.

Interpretation der Invertibilitätsbedingung

Die Invertibilitätsbedingung erlaubt die eindeutige Beschreibung von Moving Average-Prozessen. Zwei Arten von Motivationen sollen diesen Problembereich illustrieren:

a) Interpretation an Hand der ACF eines MA-Prozesses

Die Invertibilitätsbedingung erlaubt uns, zwischen folgenden beiden MA(1)-Prozessen zu unterscheiden<sup>\*</sup>):

$$z_t = a_t - \theta a_{t-1} \quad (3.1.8)$$

$$z'_t = a_t - \frac{1}{\theta} a_{t-1} \quad \theta \neq 0 \quad (3.1.9)$$

Die beiden Prozesse haben verschiedene ACovF, aber gleiche ACF: Berechnen wir die ACovF  $\gamma'_k$  analog zu Beispiel 3.1.1., so erhalten wir

$$\begin{aligned} \gamma'_0 &= \left(1 + \frac{1}{\theta^2}\right) \sigma_a^2 \\ \gamma'_1 &= -\frac{1}{\theta} \sigma_a^2 \\ \gamma'_k &= 0 \quad k > 2 \end{aligned} \quad (3.1.10)$$

---

\* Durch die Forderung  $|\theta| < 1$  gewährleistet man Eindeutigkeit.

Die ACF ist für den Prozeß  $z_t$

$$\rho_1 = \frac{-\theta}{(1+\theta^2)\sigma_a^2} = \frac{-\theta}{1+\theta^2} \quad (3.1.11)$$

und für den Prozeß  $z'_t$

$$\rho'_1 = \frac{-\frac{1}{\theta}\sigma_a^2}{(1+\frac{1}{\theta^2})\sigma_a^2} = \frac{-\theta}{1+\theta^2} \quad (3.1.12)$$

ist jedoch gleich.

b) Interpretation durch die  $\pi$ -Form von MA-Prozessen

Dazu betrachten wir

Beispiel 3.1.1.: MA(1)-Prozeß

Der MA(1)-Prozeß ist gegeben durch

$$z_t = a_t - \theta a_{t-1} = (1-\theta B)a_t \quad (3.1.13)$$

uns kann als einfachster Spezialfall der RS-Form mit Transferfunktion  $\psi_1 = -\theta$  und  $\psi_j = 0, j \geq 2$ , angesehen werden. Gesucht wird die  $\pi$ -Form des MA(1)-Prozesses.

Unter Berücksichtigung der geometrischen Reihe

$$\sum_{j=0}^k q_j = \frac{1-q^{k+1}}{1-q} \quad (3.1.14)$$

berechnen wir die  $\pi$ -Form des Modells

$$\begin{aligned} a_t &= (1-\theta B)^{-1} z_t \\ &= (1+\theta B + \theta^2 B^2 + \dots + \theta^k B^k)(1-\theta^{k+1} B^{k+1})^{-1} z_t \end{aligned} \quad (3.1.15)$$

Daraus erhalten wir

$$\begin{aligned} (1+\theta B + \dots + \theta^k B^k) z_t &= (1-\theta^{k+1} B^{k+1}) a_t \\ z_t &= -\theta z_{t-1} - \theta^2 z_{t-2} - \dots - \theta^k z_{t-k} + a_t - \theta^{k+1} a_{t+k-1} \end{aligned} \quad (3.1.16)$$

Da  $k$  in (3.1.15) beliebig gewählt werden kann, wiederholen wir diesen Vorgang und lassen  $k \rightarrow \infty$  gehen. Doch nur unter der Voraussetzung  $|\theta| < 1$  existiert der Grenzübergang und führt zu

$$z_t = -\theta z_{t-1} - \theta^2 z_{t-2} - \dots + a_t \quad (3.1.17)$$

d.h. die  $\pi$ -Form des MA(1)-Prozesses hat die Koeffizienten

$$\pi_j = -\theta^j \quad (3.1.18)$$

Dies legt  $|\theta| < 1$  als Invertibilitätsbedingung nahe.

Diesen Vorgang beschreiben wir formal dadurch, daß wir (3.1.13) mit dem inversen Operator  $\psi^{-1} = (1 - \theta B)^{-1}$  multiplizieren. Wir haben

$$a_t = (1 - \theta B)^{-1} z_t \quad (3.1.19)$$

Ist  $|\theta| < 1$ , so kann man  $(1 - \theta B)^{-1}$  in einer Reihe entwickeln.

$$\begin{aligned} a_t &= (1 + \theta B + \theta^2 B^2 + \dots) z_t = \\ &= z_t + \theta z_{t-1} + \theta^2 z_{t-2} + \dots \end{aligned} \quad (3.1.20)$$

bzw.

$$\begin{aligned} z_t &= -\theta z_{t-1} - \theta^2 z_{t-2} - \dots + a_t \\ &= -\sum_{j=1}^{\infty} \theta^j z_{t-j} + a_t \end{aligned} \quad (3.1.21)$$

Aus diesem Beispiel sehen wir, daß die Invertibilitätsbedingung  $|\theta| < 1$  die Existenz des  $\pi(B)$ -Polynoms sichert. Dies kann auch inhaltlich interpretiert werden. Würde man einen MA(1)-Prozeß mit  $|\theta| \geq 1$  zulassen, so bedeutet dies für die  $\pi$ -Form des Modelles (3.1.16), daß der laufende Wert einer Zeitreihe durch vergangene Werte bestimmt wird, deren Gewichte umso größer werden, je länger man in die Vergangenheit zurückgeht. Durch die Bedingung  $|\theta| < 1$  nehmen die Gewichte vergangener Perioden geometrisch ab.

### 3.1.3. Die Beziehung zwischen der RS-Form und der $\pi$ -Form

Satz 3.1.1.: Für die Operatorpolynome der RS- und der  $\pi$ -Form gilt folgende Beziehung:

$$\pi(B) = \psi^{-1}(B) = \frac{1}{\psi(B)} \quad (3.1.22)$$

bzw.

$$\psi(B)\pi(B) = 1 \quad (3.1.23)$$

Ist der ARMA(p,q)-Prozeß stationär, dann existiert wegen der Stationaritätsbedingung (3.1.3)  $\psi^{-1}(B)$ .

Beweis: Wir multiplizieren die  $\pi$ -Form  $a_t = \pi(B)z_t$  mit  $\psi(B)$  und erhalten

$$\psi(B)a_t = \psi(B)\pi(B)z_t \quad (3.1.24)$$

$$z_t = \psi(B)\pi(B)z_t \quad (3.1.25)$$

da  $z_t = \psi(B)a_t$  gerade die RS-Form ist.

Für die Operatorpolynome gilt daher die Beziehung (3.1.7).

Die tatsächliche Berechnung der  $\pi$ -Koeffizienten erfolgt durch Koeffizientenvergleich in (3.1.7)

$$(1 - \pi_1 B - \pi_2 B^2 - \dots)(1 + \psi_1 B + \psi_2 B^2 + \dots) = 1 \quad (3.1.26)$$

Daraus können umgekehrt die  $\psi_j$  aus den  $\pi_j$  berechnet werden.

Die genaue Beziehung ist in Kapitel 4.2.6. dargestellt, da der Koeffizientenvergleich auch für nichtstationäre Modelle möglich ist.

### 3.1.4. Die Autokovarianzerzeugende Funktion

Im folgenden Abschnitt werden die ACovF und die ACov-erzeugende Funktion für die RS-Form eines ARMA-Prozesses abgeleitet.

Satz 3.1.2.: Gegeben sei die RS-Form (3.1.1), dann lautet die ACovF

$$\gamma_k = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} \quad (3.1.27)$$

wobei  $\sigma_a^2$  die Varianz des white-noise-Prozesses ist.

Beweis:

$$\begin{aligned} \gamma_k &= E(z_t z_{t+k}) \\ &= E \left[ \sum_{j=0}^{\infty} \psi_j a_{t-j} \sum_{i=0}^{\infty} \psi_i a_{t+k-i} \right] \\ &= E \left[ \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \psi_j \psi_i a_{t-j} a_{t+k-i} \right] \\ &= \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} \end{aligned} \quad (3.1.28)$$

Die gemischten Glieder  $a_{t-j} a_{t+k-i}$  liefern nur dann einen Beitrag, wenn  $t-j=t+k-i$  gilt, d.h.  $i=k+j$ .

Als Spezialfall der ACovF erhalten wir für  $k=0$  die Varianz des ARMA(p,q)-Prozesses

$$\text{Var}(z_t) = \gamma_0 = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j^2 \quad (3.1.29)$$

Die Varianz eines linearen stochastischen Prozesses existiert, falls  $\sum_{j=0}^{\infty} \psi_j^2$  konvergiert.

Def. 3.1.3.: Die Autokovarianzerzeugende Funktion (ACovGF)

Gegeben sei ein stochastischer Prozeß  $\{z_t\}$  mit ACovF  $\gamma_k$   $k=1,2,\dots$   
Die ACovGF lautet

$$\gamma(s) = \sum_{k=-\infty}^{\infty} \gamma_k s^k \quad (3.1.30)$$

Satz 3.1.3.: Sei  $\psi(B)$  das  $\psi$ -Polynom eines linearen stationären stochastischen Prozesses. Dann lautet die ACovGF:

$$\begin{aligned}\gamma(s) &= \sigma_a^2 \psi(s)\psi(s^{-1}) \\ &= \sigma_a^2 \psi(s)\psi\left(\frac{1}{s}\right).\end{aligned}\quad (3.1.31)$$

(Box-Jenkins (1970, S.81) schreiben wegen  $s^{-1}=B^{-1}=F$  diese Beziehung als  $\gamma(B)=\sigma_a^2\psi(B)\psi(F)$ ).

Beweis: Setzen wir die ACovF (3.1.14) in Def. 3.1.1 ein, so erhalten wir

$$\begin{aligned}\gamma(s) &= \sum_{k=-\infty}^{\infty} \left( \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} \right) s^k \\ &= \sigma_a^2 \sum_{j=0}^{\infty} \sum_{k=-j}^{\infty} \psi_j \psi_{j+k} s^k\end{aligned}\quad (3.1.32)$$

mit  $\psi_0=0$  für  $j<0$ , da für negative  $j$  die  $\psi_j$  Null sind.

Indextransformation in (3.1.19)  $h=j+k$ , bzw.  $k=h-j$

$$\begin{aligned}\gamma(s) &= \sigma_a^2 \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} \psi_j \psi_h s^{h-j} \\ &= \sigma_a^2 \sum_{h=0}^{\infty} \psi_h s^h \sum_{j=0}^{\infty} \psi_j s^{-j} \\ &= \sigma_a^2 \psi(s)\psi(s^{-1})\end{aligned}\quad (3.1.33)$$

Wie man in (3.1.16) erkennt, erhält man die ACovF, indem man die Koeffizienten der Exponenten von  $s^k$  (bzw.  $B^k$ ) bestimmt.

$$\gamma(B) = \sum_{k=-\infty}^{\infty} \gamma_k B^k = \sigma_a^2 \psi(B)\psi(B^{-1})\quad (3.1.34)$$

Man erhält die Werte der ACovF  $\gamma_k$  als Koeffizienten des Polynoms, das man durch Multiplikation der Transferfunktion  $\psi(B)$  mit  $\psi(B^{-1})$  erhält.

Beispiel 3.1.2.: MA(1)-Prozeß

Der MA(1)-Prozeß hat die Transferfunktion  $\psi(B)=1-\theta B$

$$\begin{aligned}\gamma(B) &= \sigma_a^2 \psi(B)\psi(B^{-1}) \\ &= \sigma_a^2 (1-\theta B)(1-\theta B^{-1}) \\ &= \sigma_a^2 (1-\theta B-\theta B^{-1}+\theta^2 B^0) \\ &= \sigma_a^2 (-\theta B^{-1}+(\theta^2+1)B^0-\theta B) \\ &= \sigma_a^2 (\gamma_{-1} + \gamma_0 + \gamma_1)\end{aligned}\tag{3.1.35}$$

Durch Koeffizientenvergleich der ACovGF erhalten wir die ACovF

$$\begin{aligned}\gamma_0 &= (1+\theta^2)\sigma_a^2 \\ \gamma_1 &= -\theta\sigma_a^2 \\ \gamma_k &= 0 \quad k \geq 2\end{aligned}\tag{3.1.36}$$

### 3.2. Autoregressive Prozesse

#### 3.2.1. Invertibilitätsbedingungen

Ein AR(p)-Prozeß ist in der  $\pi$ -Form eines linearen stochastischen Prozesses definiert

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t \quad (3.2.1)$$

und lautet in der Operatorform

$$\begin{aligned} a_t &= z_t - \phi_1 z_{t-1} - \dots - \phi_p z_{t-p} \\ &= \phi(B) z_t \end{aligned} \quad (3.2.2)$$

wobei  $\phi(B)$  das AR-Polynom ist und

$$\pi(B) = \phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p) \quad (3.2.3)$$

ein endliches Polynom ist. Es sind keine Invertibilitätsbedingungen zu spezifizieren.

Autoregressive Prozesse können als stochastische Differenzgleichungen betrachtet werden. Daher leitet sich folgende Definition ab:

Def. 3.2.1.: Die charakteristische Gleichung

eines autoregressiven Prozesses ist die Gleichung

$$(1 - \phi_1 s - \dots - \phi_p s^p) = \phi(s) = 0 \quad (3.2.4)$$

Die Lösungen  $\alpha_i$ ,  $i=1, \dots, p$ , heißen Nullstellen der charakteristischen Gleichung.

#### 3.2.2. Stationaritätsbedingungen

Die Stationaritätsbedingungen für autoregressive Prozesse können in zwei verschiedenen Arten abgeleitet werden:

- a) für die Koeffizienten (Parameter) der Modelle
- b) für die Wurzeln der charakteristischen Gleichung eines Prozesses

Beide Formen werden aus der RS-Form abgeleitet.  
Zunächst betrachten wir wieder ein Beispiel:

Beispiel 3.2.1.: AR(1)-Prozeß

Um die RS-Form zu erhalten, multiplizieren wir formal mit dem inversen Polynom

$$z_t = (1 - \phi_1 B)^{-1} a_t = \sum_{j=0}^{\infty} \phi_1^j a_{t-j} \quad (3.2.5)$$

oder in Polynomschreibweise

$$\psi(B) = \phi^{-1}(B) = (1 - \phi_1 B)^{-1} = \sum_{j=0}^{\infty} \phi_1^j B^j \quad (3.2.6)$$

a) Stationaritätsbedingungen für die Modellparameter:

Nach Satz 3.1.4. ist der Prozeß dann stationär, falls  $\psi(s)$  für  $|s| \leq 1$  konvergiert.

Die unendliche Reihe (3.2.5) konvergiert nur dann, falls

$$|\phi_1| < 1 \quad (3.2.7)$$

b) Stationaritätsbedingungen für die Wurzeln der charakteristischen Gleichung:

Dazu benötigen wir die Nullstellen der charakteristischen Gleichung des AR(1)-Prozesses

$$(1 - \phi_1 s) = 0 \quad \text{bzw.} \quad s = 1/\phi_1 \quad (3.2.8)$$

Die Wurzel der charakteristischen Gleichung liegt somit außerhalb des Einheitskreises.

Die Stationaritätsbedingungen für den allgemeinen AR(p)-Prozeß erhalten wir ausgehend von der  $\pi$ -Form

$$\phi(B)z_t = a_t$$

Bilden wir formal die RS-Form durch

$$z_t = \phi^{-1}(B)a_t \quad (3.2.9)$$

Wir schreiben dazu das AR-Polynom als Produkt von Linearfaktoren

$$\phi(B) = (1-\alpha_1 B)(1-\alpha_2 B)\dots(1-\alpha_p B) \quad (3.2.10)$$

wobei die  $\alpha_i$  die Nullstellen der zugehörigen charakteristischen Gleichung sind. Damit gilt für  $\phi^{-1}(B)$

$$\begin{aligned} \phi^{-1}(B) &= \prod_{i=1}^p \frac{1}{(1-\alpha_i B)} \\ &= \sum_{i=1}^p \frac{c_i}{(1-\alpha_i B)} \end{aligned} \quad (3.2.11)$$

wobei die  $c_i$  durch die Partialbrüche bestimmt sind. Die Transferfunktion  $\psi(B)$  erhalten wir auch durch

$$\begin{aligned} \phi^{-1}(B) &= \sum_{i=1}^p \frac{1}{(1-\alpha_i B)} \\ &= \sum_{i=1}^p \sum_{k=0}^{\infty} \alpha_i^k B^k \\ &= \sum_{k=0}^{\infty} \psi_k B^k = \psi(B) \end{aligned} \quad (3.2.12)$$

Die Koeffizienten  $\psi_k$  werden durch das Produkt von p Polynomen bestimmt.

- a) Stationaritätsbedingungen für die Modellparameter:  
Soll der Prozeß stationär sein, muß  $\psi(s)=\phi^{-1}(s)$  eine konvergente Reihe für  $|s|<1$  sein.

Aus (3.2.11) und (3.2.12) erkennen wir daß

$$|\alpha_i| < 1 \quad i=1,\dots,p \quad (3.2.13)$$

gilt, d.h. die Koeffizienten der charakteristischen Gleichung liegen innerhalb des Einheitskreises.

Durch Koeffizientenvergleich in (3.2.10) lassen sich die Stationaritätsbedingungen (3.2.13) auch auf die AR-Parameter  $\phi_i$  übertragen. Da diese nichtlineare Funktionen der  $\alpha_i$  sind, ist keine einheitliche Darstellung möglich.

b) Stationaritätsbedingungen für die Wurzeln der charakteristischen Gleichung:

Die Wurzel für jeden Faktor  $(1-\alpha_i B)$  in 3.2.10 ist  $B=s_i=1/\alpha_i$  und wegen  $|\alpha_i|<1$  gilt daher  $|s_i|>1$ . Wir sagen auch: Die Wurzeln der charakteristischen Gleichung (des AR-Polynoms) liegen außerhalb des Einheitskreises.

Zusammenfassung:

Ein AR(p)-Prozeß ist

stationär, falls die Wurzeln von  $\phi(B)$  außerhalb des Einheitskreises liegen, bzw.  $|\phi_i|<1$  gilt.  
invertibel, für alle Parameter ohne Einschränkung

3.2.3. Die ACovF und ACF eines AR(p)-Prozesses

a) Autokovarianzfunktion

Zunächst bestimmen wir die ACovF eines allgemeinen AR(p)-Prozesses

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t \quad (3.2.14)$$

Dazu multiplizieren wir (3.2.14) mit  $z_{t-k}$

$$z_{t-k} z_t = \phi_1 z_{t-k} z_{t-1} + \dots + \phi_p z_{t-k} z_{t-p} + z_{t-k} a_t \quad (3.2.15)$$

und nehmen den Erwartungswert

$$\gamma_{j-k} = \gamma_{k-j} = E(z_{t-k} z_{t-j}) \quad (3.2.16)$$

und erhalten

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \quad (3.2.17)$$

da 
$$E(z_{t-k} a_t) = E\left(\sum_{j=0}^{\infty} a_{t-k-j} a_t\right) = 0 \quad k > 0 \quad (3.2.18)$$

$z_{t-k}$  in der RS-Form dargestellt, enthält nur Störungen  $a_t$  bis zum Zeitpunkt  $t-k$ .

b) Autokorrelationsfunktion

Dazu dividieren wir (3.2.17) durch  $\gamma_0$  und erhalten

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad k > 0 \quad (3.2.19)$$

Definieren wir nun den Backshiftoperator  $B$  auch für Folgen  $\gamma_k$

$$B\gamma_k = \gamma_{k-1} \quad (3.2.20)$$

so erfüllt die ACF dieselbe Differenzgleichungsform wie der zugehörige AR-Prozeß

$$\phi(B)\rho_k = 0 \quad (3.2.21)$$

Das AR-Operatorpolynom ist  $\phi(B)$

$$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p \quad (3.2.22)$$

wirkt in diesem Fall auf die lags  $k$  ((und nicht auf die Zeitpunkte  $t$ )).

3.2.4. Das Verhalten der ACF eines AR(p)-Prozesses

Die in (3.2.19) definierte ACF ist eine Differenzgleichung in  $\rho$ , die man dadurch löst, daß man die Nullstellen der charakteristischen Gleichung bestimmt:

$$\phi(B) = \prod_{i=1}^p (1 - \alpha_i B) = 0 \quad (3.2.23)$$

Wenn keine Mehrfachwurzeln vorkommen, lautet die allgemeine Lösung:

$$\rho_k = A_1 \alpha_1^k + A_2 \alpha_2^k + \dots + A_p \alpha_p^k \quad (3.2.24)$$

wobei die  $A_i$  durch Anfangsbedingungen zu bestimmende Konstante sind. Für die Nullstellen  $\alpha_i$  der charakteristischen Gleichung (3.2.23) gilt wegen der Stationaritätsbedingung (3.2.13) und daher konvergiert  $\rho_k$  für  $k \rightarrow \infty$ . Sind die  $\alpha_i$  alle verschieden, so unterscheiden wir folgende zwei Fälle:

a)  $\alpha_i$  ist reell

Ist eine Wurzel  $\alpha_i$  reell, so liefert in (3.2.24) der Term  $A_i \alpha_i^k$  einen Beitrag zur Autokorrelationsfunktion, der für  $k \rightarrow \infty$  in geometrischer Progression gegen Null geht. Wir bezeichnen dieses Verhalten kurz als "gedämpft exponentiell" (damped exponential).

b) Ein Paar der Wurzeln  $\alpha_i, \alpha_j$  ist komplex

Sind zwei Wurzeln  $\alpha_i$  und  $\alpha_j$  konjugiert komplex, so liefern sie einen Beitrag der Form

$$A_i \alpha_i^k + A_j \alpha_j^k = A d^k \cos(2\pi f k + F) \quad (3.2.25)$$

Wir bezeichnen diesen Beitrag als "gedämpfte Sinusschwingung". Die Konstanten in (3.2.25) bestimmen sich folgendermaßen: Schreiben wir die konjugiert komplexen Wurzeln als

$$\begin{aligned} \alpha_i &= u + iv \\ \alpha_j &= u - iv \end{aligned} \quad (3.2.26)$$

so erhält man daraus die Periode  $T = \frac{1}{f}$  und den Dämpfungsfaktor  $d$  der Sinusschwingung

$$d = \sqrt{u^2 + v^2} \quad (3.2.27)$$

$$\operatorname{tg} \theta = \operatorname{tg}(2\pi f) = \frac{v}{u} \quad (3.2.28)$$

Die aus den Anfangsbedingungen bestimmbaren Konstanten  $A_i$  und  $A_j$  lassen sich umformen zu (vgl. dazu (4.6.61)).

$$A = 2\sqrt{A_i A_j} \quad (3.2.29)$$

$$\operatorname{tg} F = \frac{A_i - A_j}{A_i + A_j} \quad (3.2.30)$$

### Zusammenfassung

Die ACF eines AR(p)-Prozesses besteht allgemein aus einer Kombination von gedämpft exponentiellen Beiträgen und gedämpften Sinusschwingungen.

### 3.2.5. Die Yule-Walker-Gleichungen

Die Yule-Walker-Gleichungen geben die Beziehungen zwischen den Parametern eines autoregressiven Prozesses und dessen ACF an. Die Differenzengleichungsform für die ACF lautet

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad (3.2.31)$$

Substituieren wir der Reihe nach für  $k=1, \dots, p$ , so erhalten wir folgendes lineares Gleichungssystem für die Parameter  $\phi_1 \dots \phi_p$

$$\begin{aligned} \rho_1 &= \phi_1 \rho_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-2} \\ &\vdots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \end{aligned} \quad (3.2.32)$$

Dieses Gleichungssystem bezeichnet man als Yule-Walker-Gleichungen.

In Matrixschreibweise schreiben wir kompakter

$$\underline{\phi}_p^P = \underline{\rho}_p \quad (3.2.33)$$

mit der Autokorrelationsmatrix der Ordnung p

$$\underline{P}_p = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \dots & \dots & 1 \end{pmatrix} \quad (3.2.34)$$

und den Vektoren

$$\underline{\phi}' = (\phi_1 \dots \phi_p) \quad (3.2.35)$$

$$\underline{\rho}'_{(p)} = (\rho_1 \dots \rho_p) \quad (3.2.36)$$

Man beachte, daß  $\underline{P}_p$  nur Autokorrelationen bis zur Ordnung p-1 enthält.

#### Yule-Walker Schätzungen für die Parameter

Die Yule-Walker-Gleichungen (3.2.32) bzw. (3.2.33) kann man zur Schätzung von Parametern des autoregressiven Modells verwenden, indem man die theoretischen Autokorrelationen  $\rho_k$  durch deren Schätzungen  $r_k = \hat{\rho}_k$  verwendet.

Sei

$$\hat{\underline{P}}_p = \begin{pmatrix} 1 & r_1 & \dots & r_{p-1} \\ r_1 & 1 & \dots & r_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & \dots & \dots & 1 \end{pmatrix} \quad (3.2.37)$$

und  $\underline{r}'_{(p)} = (r_1 \dots r_p)$ , so können wir  $\underline{\phi}$  als nichtlineare Funktion der  $r_i$  ausdrücken und erhalten die sogenannten Yule-Walker-Schätzungen

$$\hat{\underline{\phi}} = \hat{\underline{P}}_p^{-1} \underline{r}_{(p)} \quad (3.2.38)$$

Diese Schätzungen haben im allgemeinen schlechte Eigenschaften, sodaß sie nur zur Schätzung von vorläufigen Parameterwerten für eine nichtlineare Schätzung verwendet werden.

### 3.2.6. Die Varianz eines AR(p)-Prozesses

Dazu gehen wir auf die Autokovarianzfunktion (3.2.1) zurück und setzen  $k=0$

$$z_t \cdot z_t = \phi_1 z_{t-1} z_t + \phi_2 z_{t-2} z_t + \dots + \phi_p z_{t-p} z_t + z_t a_t \quad (3.2.39)$$

Nehmen wir die Erwartungswerte auf beiden Seiten und berücksichtigen, daß

$$\begin{aligned} E(z_t a_t) &= E[(\phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t) a_t] = \\ &= \sigma_a^2 \end{aligned} \quad (3.2.40)$$

so erhalten wir

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma_a^2 \quad (3.2.41)$$

Herausheben von

$$\begin{aligned} \gamma_0 = \text{Var}(z_t) &= \sigma_z^2 \quad \text{ergibt} \\ \gamma_0 (1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \dots - \phi_p \rho_p) &= \sigma_a^2 \end{aligned} \quad (3.2.42)$$

oder

$$\begin{aligned} \sigma_z^2 &= \frac{\sigma_a^2}{1 - \rho_1 \phi_1 - \dots - \rho_p \phi_p} \\ &= \sigma_a^2 \left( \sum_{i=0}^p \rho_i \phi_i \right)^{-1} \end{aligned} \quad (3.2.43)$$

### 3.2.7. Die partielle Autokorrelationsfunktion

Die PACF ist ein wichtiges Hilfsmittel für die Entscheidung, welche Ordnung ein autoregressiver Prozeß hat. Das Problem ist teilweise jenem der multiplen Regression ähnlich, wenn es zu entscheiden gilt, wieviele unabhängige Variable in die Regression miteinbezogen werden.

Im folgenden geben wir die gängige Herleitung der PACF wieder, während in Appendix A1 aus der Definition des partiellen Korrelationskoeffizienten (mit Hilfe bedingter Verteilungen) gezeigt wird, daß  $\phi_{kk}$  tatsächlich wieder Korrelationskoeffizienten sind.

Man erhält die PACF als spezielle Koeffizienten des folgenden Systems von Gleichungen

$$\rho_j = \phi_{k1}\rho_{j-1} + \dots + \phi_{kk}\rho_{j-k} \quad j=1,2,\dots,k \quad (3.2.44)$$

für  $k=2, k=3, \dots$

Ist der Prozeß tatsächlich von zweiter Ordnung, dann ist der Koeffizient  $\phi_{22}$  in (3.2.44) ungleich Null, während  $\phi_{33}=0$ , falls  $k=3$  ist.

Um herauszufinden, welche Ordnung ein AR-Prozeß hat, lösen wir für  $k=1,2,\dots$  die Gleichungssysteme (3.2.44).

Für  $k=1$  haben wir eine Gleichung

$$\rho_1 = \phi_{11}\rho_0 \quad (3.2.45)$$

Für  $k=2$  erhalten wir ein Paar von Gleichungen

$$\begin{aligned} \rho_1 &= \phi_{21}\rho_0 + \phi_{22}\rho_1 \\ \rho_2 &= \phi_{21}\rho_0 + \phi_{22}\rho_1 \end{aligned} \quad (3.2.46)$$

Allgemein erhalten wir die  $k$  Yule-Walker-Gleichungen.

$$\begin{pmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & & \rho_{k-2} \\ \vdots & & \cdot & \cdot & \vdots \\ \rho_{k-1} & & & & 1 \end{pmatrix} \cdot \begin{pmatrix} \phi_{k1} \\ \vdots \\ \phi_{kk} \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \vdots \\ \rho_k \end{pmatrix} \quad (3.2.47)$$

oder in Matrixschreibweise

$$P_k \phi_k = \rho(k) \quad (3.2.48)$$

mit

$$\phi_k' = (\phi_{k1}, \dots, \phi_{kk})$$

$$\rho(k) = (\rho_1, \dots, \rho_k)$$

Wir lösen nun die Yule-Walker-Gleichungen der Reihe nach für  $k=1,2,\dots$

$$k=1 \quad \phi_{11} = \rho_1$$

$$k=2 \quad \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_2 & 1 \end{vmatrix}} = \frac{\rho_1^2 - \rho_2}{\rho_1^2 - 1} \quad (3.2.49)$$

$$k=3 \quad \phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}$$

Allgemein bekommt man  $\phi_{kk}$  durch die Cramer'sche Regel (vgl. 1.42). Im Nenner steht die Determinante der Matrix  $P_k$ , im Zähler steht die Determinante der Matrix  $P_k$ , in der die letzte Spalte gegen den  $\rho_k$ -Vektor vertauscht wurde.

Def. 3.2.2.: Die PACF

Die Funktion  $k \rightarrow \phi_{kk}$ ,  $k=0,1,2,\dots$ , heißt partielle Autokorrelationsfunktion (PACF).  $\phi_{kk}$  ist dabei die Folge der Koeffizienten von  $\rho_{j-k}$  in der Yule-Walker-Gleichung  $k$ -ter Ordnung (3.2.44).

Ein AR(p)-Prozeß der Ordnung  $p$  hat demnach eine PACF

$$\phi_{kk} \begin{cases} \neq 0 & k \leq p \\ = 0 & k > p \end{cases} \quad (3.2.50)$$

die einen "cut off" nach dem lag  $p$  hat.

Die Schätzung der PACF

Dazu ersetzen wir die theoretische ACF  $\rho_i$  durch die Schätzung  $r_k$

$$r_j = \hat{\phi}_{k1} r_{j-1} + \hat{\phi}_{k2} r_{j-2} + \dots + \hat{\phi}_{kk} r_{j-k} \quad j=1,2,\dots,k \quad (3.2.51)$$

Zur numerischen Berechnung gibt es eine einfache rekursive Methode von J. Durbin (1960) (vgl. Box-Jenkins, S.82 ff.).

### Standard-Fehler der partiellen Autokorrelationsfunktion

Quenouille (1949) zeigte, daß unter der Annahme, daß der zugrundeliegende Prozeß autoregressiv von der Ordnung  $p$  ist, die geschätzten Koeffizienten der PACF approximativ verteilt sind, mit Varianz

$$\text{Var}(\hat{\phi}_{kk}) \approx \frac{1}{n} \quad k > p+1 \quad (3.2.52)$$

d.h. die Standardabweichung ist  $\text{ST.D}(\hat{\phi}_{kk}) \sim 1/\sqrt{n}$ , wobei  $Z_t$ ,  $t=1, \dots, n$  beobachtet wurde.

### 3.3. Spezielle autoregressive Prozesse

Im folgenden behandeln wir die in der Praxis am häufigsten vorkommenden AR( $p$ )-Prozesse, den AR(1)- und den AR(2)-Prozeß.

#### 3.3.1. Der AR(1)-Prozeß

Der autoregressive Prozeß erster Ordnung kann als einfacher Markovprozeß betrachtet werden: Der derzeitige Wert der Zeitreihe hängt nur vom letzten Wert ab (plus einer zufälligen Störung).

Die  $\pi$ -Form lautet

$$z_t = \phi_1 z_{t-1} + a_t \quad (3.3.1)$$

bzw.

$$\phi(B)z_t = (1 - \phi_1 B)z_t = a_t \quad (3.3.2)$$

wobei  $\phi(B)$  das AR(1)-Polynom ist.

Die RS-Form lautet

$$\begin{aligned} z_t &= a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \dots \\ &= (1 - \phi_1 B)^{-1} a_t = \psi(B) a_t \end{aligned} \quad (3.3.3)$$

Da das charakteristische Polynom nur aus einem Faktor  $(1 - \alpha_1 B) = (1 - \phi_1 B)$  besteht, lautet die Stationaritätsbedingung für den Modellparameter  $|\phi_1| < 1$  ( $-1 < \phi_1 < 1$ ).

Die ACF erfüllt die Differenzgleichung erster Ordnung

$$\rho_k = \phi_1 \rho_{k-1} \quad k > 0 \quad (3.3.4)$$

mit der Anfangsbedingung  $\rho_0 = 1$  wird sie zu

$$\rho_k = \phi_1^k \quad k \geq 0 \quad (3.3.5)$$

Die PACF ist

$$\begin{aligned} \phi_{11} &= \rho_1 \\ \phi_{22} &= \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \frac{\phi_1^2 - \phi_1^2}{1 - \phi_1^2} = 0 \end{aligned} \quad (3.3.6)$$

Die Korrelation von  $z_t$  mit  $z_{t+2}$  geht vollständig auf den gemeinsamen Zusammenhang mit  $z_{t+1}$  und damit auf  $\rho_1$  zurück.

Als besondere Charakteristik ist hervorzuheben, daß

$$\rho_1 = \phi_{11} = \phi_1 \quad (3.3.7)$$

d.h. der Wert der ACF und PACF am lag 1 ist gleich dem AR-Parameter. Die ACF verhält sich gedämpft exponentiell, d.h. die logarithmierte ACF verhält sich wie eine Gerade:

$$\ln |\rho_k| = k \ln |\phi_1| \quad (3.3.8)$$

und die PACF hat einen "cut off" am lag 1.

Die Varianz des AR(1)-Prozesses erhalten wir, indem wir  $p=1$  in Formel (3.2.43) einsetzen

$$\sigma_z^2 = \frac{\sigma_a^2}{1 - \rho_1 \phi_1} = \frac{\sigma_a^2}{1 - \phi_1^2} \quad (3.3.9)$$

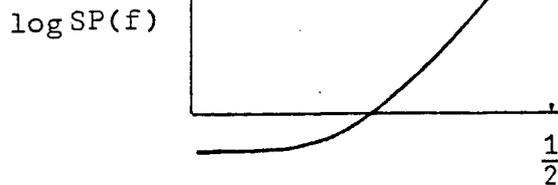
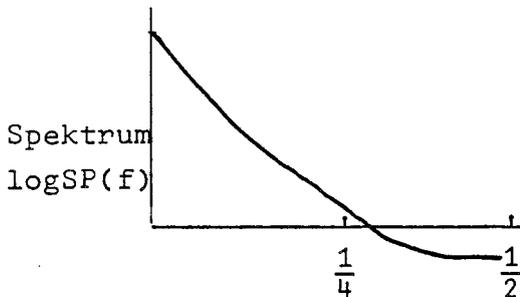
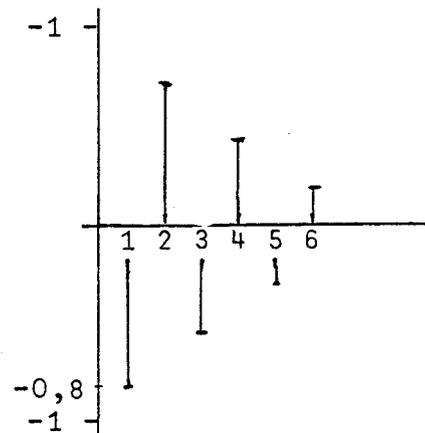
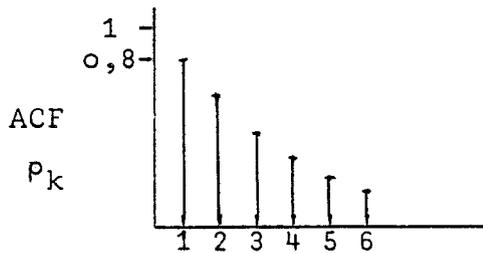
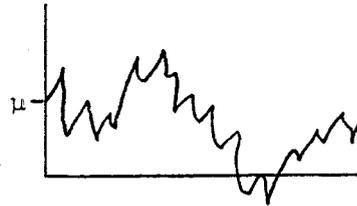
Man beachte: Je größer  $|\phi_1|$ , desto größer wird die Varianz des Prozesses.

Beispiel 3.3.1.: AR(1)-Prozeß

Einen guten Einblick in das Verhalten eines AR(1)-Prozesses erhält man durch einen Vergleich von positiven und negativen Modellparametern

$$z_t = 0,8z_{t-1} + a_t$$

$$z'_t = -0,8z'_{t-1} + a_t$$



Varianz 
$$\sigma_z^2 = \frac{\sigma_a^2}{1-0,8^2} = \sigma_a^2 \frac{1}{0,36} \sim 3\sigma_a^2$$

Der positive Parameter zeigt das Verhalten eines gedämpften exponentiellen Beitrages, der negative Parameter bewirkt die gedämpfte Sinusschwingung zur Periode 2. Der positive Parameter bewirkt ein gleichmäßigeres und geglätteteres Bild der Zeitreihe und der ACF, während der negative Parameter ein zergliederteres und hektischeres Verhalten widerspiegelt. Trotz konträren Verhaltens haben die Prozesse  $z_t$  und  $z'_t$  die gleiche Varianz.

### 3.3.2. Der AR(2)-Prozeß

Der autoregressive Prozeß zweiter Ordnung wird manchmal (oder wie alle AR-Prozesse) Yule-Prozeß genannt, da G.U. Yule (1927) erstmals versuchte, den Zyklus der Sonnenflecken aus der Differenzengleichungsform einer Sinusfunktion zu schätzen.

Die  $\pi$ -Form lautet

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t \quad (3.3.10)$$

bzw.  $\phi(B)z_t = a_t \quad (3.3.11)$

mit dem AR(2)-Polynom

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 \quad (3.3.12)$$

Zur Bestimmung der RS-Form berechnen wir die Nullstellen der charakteristischen Gleichung  $\phi(B)=0$

$$\begin{aligned} 1 - \phi_1 B - \phi_2 B^2 &= (1 - \alpha_1 B)(1 - \alpha_2 B) \\ &= 1 - (\alpha_1 + \alpha_2)B - (-\alpha_1 \alpha_2)B^2 \end{aligned} \quad (3.3.13)$$

Durch Koeffizientenvergleich erhalten wir

$$\phi_1 = \alpha_1 + \alpha_2 \quad (3.3.14)$$

$$\phi_2 = -\alpha_1 \alpha_2 \quad (3.3.15)$$

Nach der Stationaritätsbedingung  $|\alpha_1| < 1$  können wir nun die Stationaritätsbedingung für die Modellparameter ableiten. Zunächst erhalten wir eine Abschätzung für  $\phi_1$  und  $\phi_2$ , indem wir in (3.3.14) und (3.3.15) die Beträge nehmen.

$$|\phi_2| = |-\alpha_1| |\alpha_2| < 1 \quad (3.3.16)$$

und

$$|\phi_1| = |\alpha_1 + \alpha_2| \leq |\alpha_1| + |\alpha_2| < 2 \quad (3.3.17)$$

Für eine gemeinsame Abschätzung beachten wir, daß wegen  $|\alpha_2| < 1$  sowohl  $1 - \alpha_2 > 0$  als auch  $1 + \alpha_2 > 0$  ist. Damit wird jeweils die Ungleichung mit  $|\alpha_1| < 1$  multipliziert.

$$\alpha_1(1 - \alpha_2) < 1 - \alpha_2$$

$$\alpha_1 - \alpha_1\alpha_2 + \alpha_2 < 1 \quad (3.3.18)$$

und

$$\text{damit } \phi_1 + \phi_2 < 1 \quad (3.3.19)$$

und die zweite Ungleichung

$$-(1 + \alpha_2) < \alpha_1(1 + \alpha_2)$$

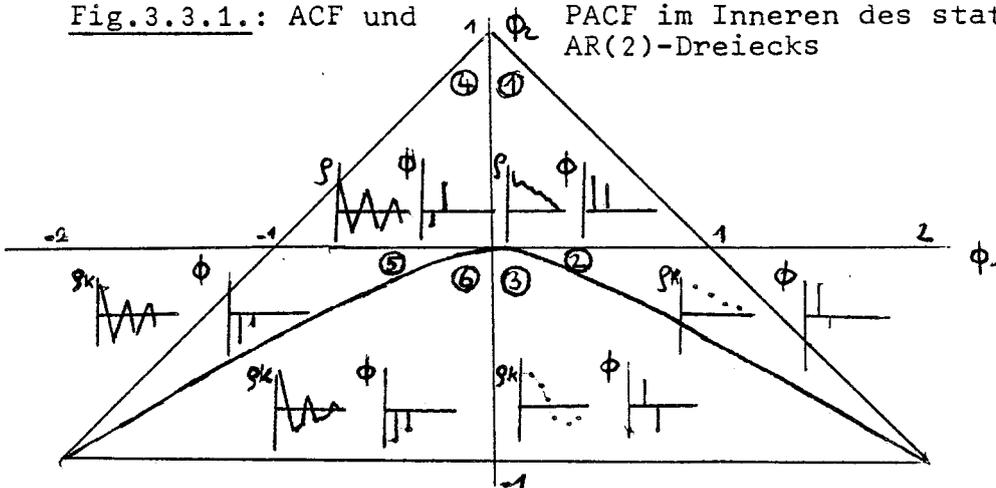
$$-\alpha_2 - \alpha_1 - \alpha_1\alpha_2 < 1 \quad (3.3.20)$$

und

$$\text{damit } \phi_1 - \phi_2 < 1 \quad (3.3.21)$$

Den zulässigen Bereich für die Parameter eines stationären AR(2)-Prozesses kann man graphisch als Dreieck im  $\mathbb{R}^2$  darstellen (vgl. Stralkowski et al. (1970)).

Fig.3.3.1.: ACF und PACF im Inneren des stationären AR(2)-Dreiecks



Den Bereich der komplexen Wurzeln erhält man durch

$$-\phi_2^2 - \phi_1 s + 1 = 0$$

$$s_{1,2} = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}$$

(3.3.22)

Die negative Diskriminante

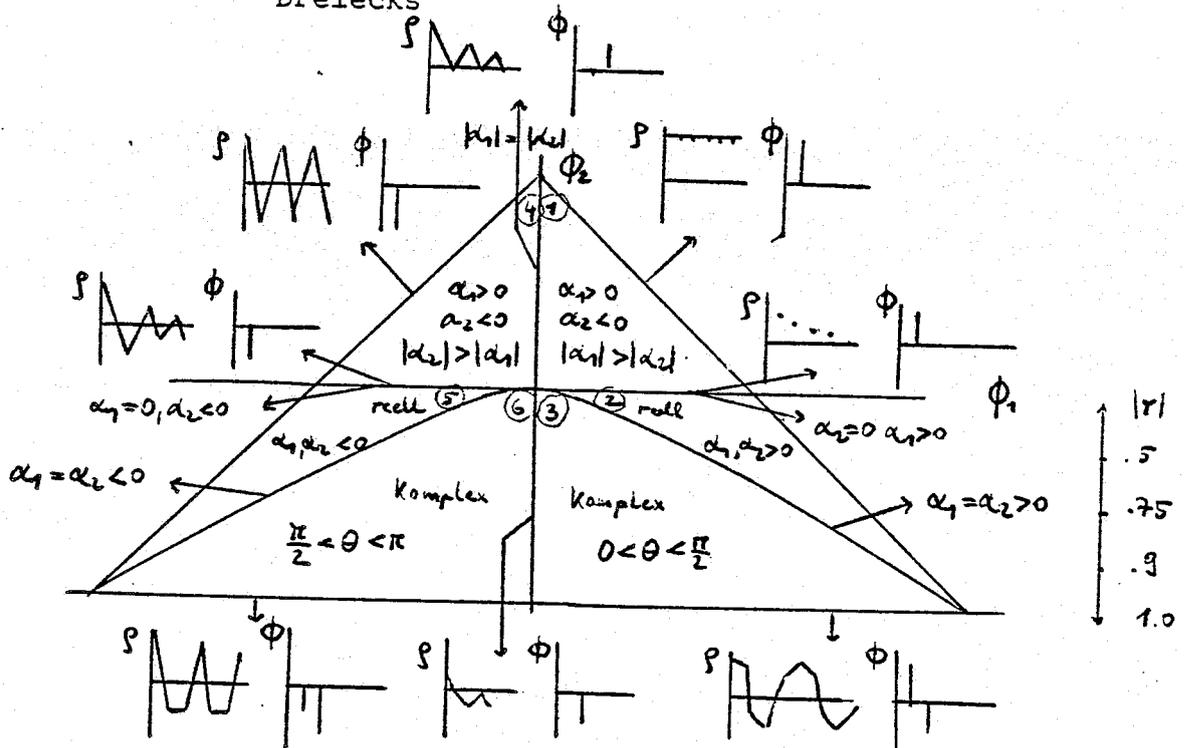
$$\phi_1^2 + 4\phi_2 < 0$$

gibt den Bereich der Parameter an, die für Schwingungsverhalten der Lösungen zuständig sind.

$$\phi_2 < -\frac{\phi_1^2}{4}$$

(3.3.23)

Fig.3.3.2.: ACF und PACF an den Rändern des stationären AR(2)-Dreiecks



### Beschreibung des stationären AR(2)-Dreiecks

Im folgenden sei kurz das Verhalten von ACF und PACF als Funktion der Parameter  $\phi_1$  und  $\phi_2$  beschrieben.

Wegen

$$\begin{aligned}\phi_1 &= \alpha_1 + \alpha_2 \\ \phi_2 &= -\alpha_1\alpha_2\end{aligned}\tag{3.3.24}$$

kommen für  $\phi_1 > 0$  die Bereiche 2 (beide Wurzeln positiv) und der Bereich 1 ( $\alpha_1$  positiv und betragsmäßig größer als  $\alpha_2$ ) in Frage. Entsprechend verkehrt ist das Verhalten für  $\phi_1 < 0$  in den Bereichen 4 und 5. Der Bereich der komplexen Wurzeln (in 3 und 6) ist für positive  $\phi_1$  mit  $0 < \theta < \frac{\pi}{2}$ , für negative  $\phi_1$  mit  $\frac{\pi}{2} < \theta < \pi$  verbunden. Nach der Definition des arccos gilt

$$x \rightarrow \arccos x \quad \text{bzw. } [-1, 1] \rightarrow [0, \pi]\tag{3.3.25}$$

Die Periodenlänge ist gegeben durch

$$T = 2\pi / \arccos(\phi_1 / 2\sqrt{-\phi_2})\tag{3.3.26}$$

$-2 \leq \phi_1 \leq 0$   
 $-1 \leq -\phi_2 \leq 0$

Für den Fall  $\phi_2 = -1$  ist  $r = \sqrt{-\phi_2} = 1$  und die Periodenlänge ist lediglich eine Funktion von  $\phi_1$ .

$$T = 2\pi / \arccos(\phi_1 / 2)\tag{3.3.27}$$

In diesem Fall ist die Sinusschwingung nicht gedämpft.

$$z_t = 2D \cos(\theta t + c) \quad (3.3.28)$$

$\phi_2 = -1$  entspricht der Basislinie des stationären  $(\phi_1, \phi_2)$ -Dreiecks für verschiedene Periodenlängen.

Die Werte der Ordinate sind durch  $\alpha = 0$  und  $\phi_2 = 0$  gekennzeichnet. Wir erhalten hier den Spezialfall des AR(1)-Prozesses.

Die Abzisse stellt die Kombination  $\phi_1 = 0$  ( $\alpha_1 = 0$ ) dar, und wegen  $\cos^{-1}(0) = \frac{\pi}{2}$  ist damit für alle  $\phi_2 < 0$  eine Periodenlänge von 4 verbunden. Weiters gilt  $\alpha_1 = -\alpha_2$  und  $\phi_2 = \alpha_1^2 = \alpha_2^2$ . Für  $\cos^{-1}(1) = 0$  ergeben sich unendliche Periodenlängen. Es muß also gelten

$$\begin{aligned} \phi_1 &= 2\sqrt{-\phi_2} \\ \text{bzw.} \quad \phi_2 &= -\phi_1^2/4 \end{aligned} \quad (3.3.29)$$

Das ist genau die quadratische Begrenzungskurve für den komplexen Bereich (vgl. Fig. 3.3.1.).

Setzen wir (3.3.24) in (3.3.29) ein, so erhalten wir

$$\begin{aligned} \alpha_1 \alpha_2 &= (\alpha_1 + \alpha_2)^2 / 4 \\ (\alpha_1 - \alpha_2)^2 &= 0 \quad \text{bzw.} \quad \alpha_1 = \alpha_2 \end{aligned} \quad (3.3.30)$$

Für  $\phi_1 < 0$  ist  $\cos^{-1}(1) = \pi$ . Wir erhalten die Periodenlänge  $T=2$ .

Der Bereich der komplexen Lösungen ist mit folgenden Periodenlängen verbunden

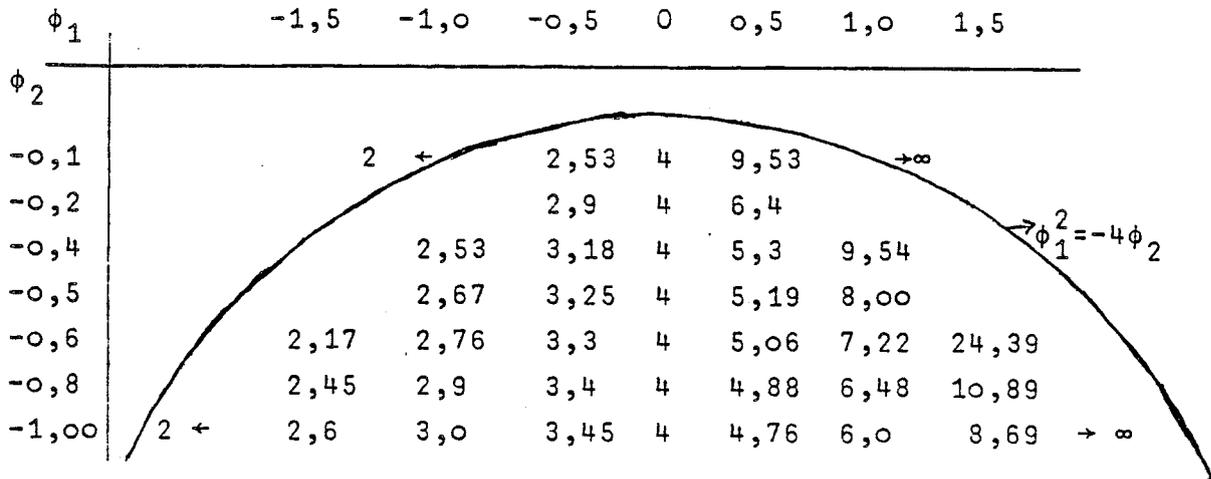


Fig.3.3.3: Periodenlänge des AR(2)-Prozesses für komplexe Wurzeln  
Die Yule-Walker-Gleichungen

Für den Fall  $p=2$  lauten die Yule-Walker-Gleichungen (3.2.29)

$$\rho_1 = \phi_1 + \phi_2 \rho_1 \quad (3.3.31)$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2$$

bzw.  $\underline{\rho}(2) = \underline{P} \underline{\phi} \quad (3.3.32)$

Die Yule-Walker-Schätzungen für  $\phi$  erhalten wir indem wir in (3.3.31)  $\rho$  durch  $r$  ersetzen und nach  $\phi$  auflösen:

$$\hat{\phi}_1 = \frac{r_1(1-r_2)}{1-r_1^2} \quad (3.3.33)$$

$$\hat{\phi}_2 = \frac{r_2 - r_1^2}{1-r_1^2}$$

Mit Hilfe dieser Beziehungen kann man aufgrund geschätzter Autokorrelationen vorläufige Parameter für  $\phi_1$  und  $\phi_2$  angeben. (Box-Jenkins(1976)S.518, geben ein graphisches Diagramm für diese Gleichungen.)

Die Yule-Walker-Gleichungen (3.3.31) nach  $\rho_1, \rho_2$  aufgelöst ergeben nichtlineare Funktionen der Parameter  $\phi$

$$\begin{aligned}\rho_1 &= \frac{\phi_1}{1-\phi_2} \\ \rho_2 &= \phi_2 + \frac{\phi_1^2}{1-\phi_2}\end{aligned}\tag{3.3.34}$$

Da die Autokorrelationsmatrix  $\underline{P}_3$  positiv definit sein muß, tritt nur ein bestimmter Bereich von  $(-1,1) \times (-1,1)$  als zulässiger Bereich der Autokorrelationen  $(\rho_1, \rho_2)$  auf. Es gelten die Beziehungen

$$\begin{aligned}|\rho_1| < 1 \quad \text{und} \quad |\rho_2| < 1 \\ \rho_1^2 > \frac{1}{2}(\rho_2 + 1)\end{aligned}\tag{3.3.35}$$

Die letzte Ungleichung erhält man, da die Determinante von  $\underline{P}_3$  positiv sein muß

$$|\underline{P}_3| = \begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix} = 1 + 2\rho_1^2\rho_2 - \rho_2^2 - 2\rho_1^2 > 0\tag{3.3.36}$$

Da  $(1-\rho_2) > 0$ , können wir kürzen

$$\begin{aligned}1 - \rho_2^2 &> 2\rho_1^2(1 - \rho_2) \\ 1 + \rho_2 &> 2\rho_1^2 \\ \rho_1^2 &< \frac{1}{2}(1 + \rho_2^2)\end{aligned}\tag{3.3.37}$$

Die Autokorrelationsfunktion

Aus der allgemeinen ACF (3.3.19) erhalten wir für  $p=2$

$$\rho_k = \phi_1 \rho_k + \phi_2 \rho_{k-2} \tag{3.3.38}$$

Für  $k=1,2$  erhalten wir einerseits die Yule-Walker-Gleichungen (3.3.31), andererseits ist die ACF die Lösung einer Differenzgleichung 2. Ordnung

$$\rho_k = C_1 \alpha_1^k + C_2 \alpha_2^k \quad \alpha_1 \neq \alpha_2 \tag{3.3.39}$$

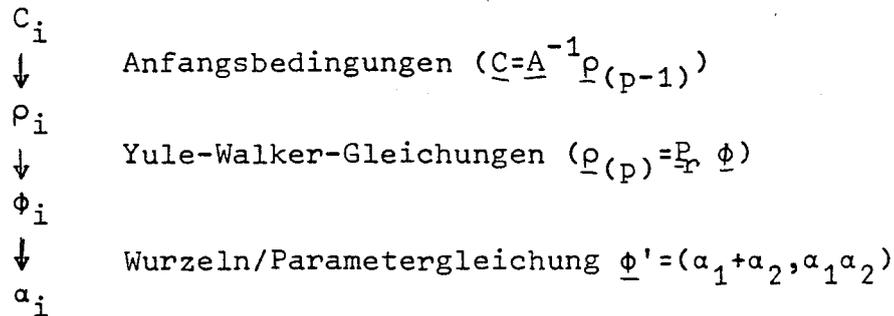
wobei die  $\alpha_i$  die Wurzeln der charakteristischen Gleichung sind. Die  $C_i$  sind Konstante, die sich aus den Anfangsbedingungen ermitteln lassen, und ebenfalls als Funktionen von  $\alpha_i$  darstellbar sind. Ziel ist es, die ACF lediglich als Funktion der  $\alpha_i$  auszudrücken.

Satz 3.3.1.:

Die ACF eines AR(2)-Prozesses als Funktion der Wurzeln  $\alpha_i$  ist gegeben durch

$$\rho_k = \frac{\alpha_1(\alpha_1^2 - 1)\alpha_1^k + \alpha_2(1 - \alpha_1^2)\alpha_2^k}{(1 + \alpha_1\alpha_2)(\alpha_2 - \alpha_1)} \quad \begin{matrix} \alpha_1 \neq \alpha_2 \\ k > 0 \end{matrix} \tag{3.3.40}$$

Der Beweis fordert einige algebraische Umformungen und folgt folgendem funktionellem Zusammenhang zwischen den Parametern eines AR-Prozesses (wir verwenden dazu Vektornotation):



Setzen wir in (3.3.39)  $k=0,1$ , so erhalten wir die Anfangsbedingungen

$$\rho_0 = C_1 + C_2 \quad (3.3.41)$$

$$\rho_1 = C_1 \alpha_1 + C_2 \alpha_2$$

bzw. mit  $\underline{\rho}'_{(1)} = (\rho_0, \rho_1)$  und  $\underline{c}' = (C_1, C_2)$  und der Matrix

$$\underline{A} = \begin{pmatrix} 1 & 1 \\ \alpha_1 & \alpha_2 \end{pmatrix} \quad \alpha_1 \neq \alpha_2 \quad (3.3.42)$$

$$\underline{\rho}'_{(1)} = \underline{A} \cdot \underline{c} \quad (3.3.43)$$

Als Umkehrfunktion erhalten wir (wegen  $\alpha_1 \neq \alpha_2$ )

$$\underline{c} = \underline{A}^{-1} \underline{\rho}'_{(1)} = \frac{1}{\alpha_2 - \alpha_1} \begin{pmatrix} \alpha_2 & -1 \\ -\alpha_1 & 1 \end{pmatrix} \underline{\rho}'_{(1)} \quad (3.3.44)$$

Als nächstes bestimmen wir aus den Yule-Walker-Gleichungen (3.3.32)  $\underline{\phi}' = (\phi_1, \phi_2)$  als nichtlineare Funktionen der  $\rho_i$ .

$$\underline{\phi} = \underline{P}^{-1} \underline{\rho}'_{(2)} \quad (3.3.45)$$

Von diesen nichtlinearen Funktionen bestimmen wir die Umkehrfunktion, wobei wir nur die erste Komponente brauchen

$$\underline{\rho}'_{(2)} = g(\underline{\phi}) \quad (3.3.46)$$

Speziell ist

$$\rho_1 = \frac{\phi_1}{1 - \phi_2} = \frac{\alpha_1 + \alpha_2}{1 + \alpha_1 \alpha_2} \quad (3.3.47)$$

wenn man die nichtlinearen Wurzelparametergleichungen

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_2 \\ -\alpha_1 \alpha_2 \end{pmatrix} \quad (3.3.48)$$

berücksichtigt. Durch Einsetzen des Ergebnisses (3.3.43) in (3.3.40) erhalten wir schließlich die Konstanten  $(C_1, C_2)$  als Funktion der Wurzeln  $\alpha_i$  alleine und damit (3.3.34).

Anmerkung: Für komplexere Wurzeln lautet die ACF

$$\rho_k = \frac{(\text{sign} \phi_1)^k d^k \sin(2\pi f_0 k + F)}{\sin F} \quad (3.3.49)$$

mit den transformierten Konstanten (3.2.26) bis (3.2.30).

### 3.4. Moving Average-Prozesse

#### 3.4.1. Der allgemeine MA(q)-Prozeß

Der MA(q)-Prozeß ist als endlicher "Durchschnitt" von zufälligen Störungen definiert. Slutsky (1927) zeigte, daß damit ebenfalls zyklische Phänomene beschrieben werden können (daher der Name "Slutsky-Prozesse").

$$z_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (3.4.1)$$

mit dem MA-Operatorpolynom  $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ .

#### a) Invertibilitätsbedingungen

Um den MA(q)-Prozeß in der  $\pi$ -Form darzustellen, benötigen wir die Invertibilitätsbedingung nach Def. 3.1.4, d.h. die Wurzeln von  $\pi(B)$  liegen innerhalb des Einheitskreises  $|B| < 1$ .

Um die invertierte Form

$$a_t = \theta^{-1}(B)z_t \quad (3.4.2)$$

herzuleiten, bestimmen wir die Wurzeln des Moving-Average-Polynoms

$$\theta(B) = \prod_{j=1}^q (1 - \beta_j B) \quad (3.4.3)$$

Für  $\pi(B)$  erhalten wir durch Partialbruchzerlegung (mit den Konstanten  $C_j$ )

$$\pi(B) = \theta^{-1}(B) = \sum_{j=1}^q \frac{C_j}{(1 - \beta_j B)} \quad (3.4.4)$$

Die Invertibilitätsbedingung lautet somit

$$|\beta_j| < 1 \quad j=1, \dots, q$$

b) Stationaritätsbedingungen

Da mit  $\Psi(B) = \theta(B)$  in (3.3.2) ein endliches Polynom vorliegt, ist jeder MA(q)-Prozeß stationär, d.h.

$$\Psi(s) = \sum_{j=0}^{\infty} \psi_j s^j$$

konvergiert für alle  $|s| < 1$ .

c) Varianz

$$\begin{aligned} \sigma_z^2 &= \gamma_0 = E(a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q})^2 = \\ &= (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_a^2 \end{aligned} \quad (3.4.5)$$

d) Autokovarianzfunktion ACovF

Wir setzen in Def. 2.5.1 ein und erhalten

$$\begin{aligned}
 \gamma_k &= E(z_t, z_{t+k}) = \\
 &= E(a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}) (a_{t-k} - \theta_1 a_{t-k-1} - \dots - \theta_q a_{t-k-q}) = \\
 &= E a_{t-k} (a_t - \dots - \theta_q a_{t-q}) - \\
 &\quad - \theta_1 a_{t-k-1} (a_t - \dots - \theta_q a_{t-q}) - \\
 &\quad - \theta_q a_{t-k-q} (a_t - \dots - \theta_q a_{t-q}) \qquad (3.4.6)
 \end{aligned}$$

Man sieht a) für  $k > q$  ist  $\gamma_k = 0$   
 b) für  $k < q$  ist  $\gamma_k \neq 0$

In folgender Tabelle erkennt man das Bildungsgesetz

	$a_{t-k}$	$a_{t-k-1}$	$a_{t-k-2}$	$a_{t-k-3}$	$\dots$	$a_{t-k-q}$
$a_t$	1	$-\theta_1$	$-\theta_2$	$-\theta_3$	$\dots$	$-\theta_q$
$a_{t-1}$	$-\theta_1$	$-\theta_1^2$	$\theta_1 \theta_2$	$\theta_1 \theta_3$	$\dots$	$\theta_1 \theta_q$
$a_{t-2}$	$-\theta_2$	$\theta_2 \theta_1$	$\theta_2^2$	$\theta_2 \theta_3$	$\dots$	$\theta_2 \theta_q$
$a_{t-3}$	$-\theta_3$	$\theta_3 \theta_1$	$\theta_3 \theta_2$	$\theta_3^2$	$\dots$	$\theta_3 \theta_q$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$a_{t-q}$	$-\theta_q$	$\theta_q \theta_1$	$\theta_q \theta_2$	$\theta_q \theta_3$	$\dots$	$\theta_q^2$

Kürzer schreibt man die ACF als

$$\gamma_k = \begin{cases} -\theta_k + \theta_1 \theta_{k+1} + \theta_2 \theta_{k+2} + \dots + \theta_{q-k} \theta_q) \sigma_a^2 & k=1 \dots q \\ 0 & k > q \end{cases} \quad (3.4.7)$$

Anmerkung: Die Yule-Walker-Gleichungen sind für den MA(q)-Prozeß nicht linear und außer dem Fall  $k=1$  nicht einfach lösbar.

### 3.4.2. Der MA(1)-Prozeß

Der MA(1)-Prozeß wurde in einer Reihe von Beispielen zur Illustration komplexerer Probleme verwendet (vgl. Beispiele 3.1.1., 3.1.2), sodaß wir im folgenden Abschnitt die Ergebnisse zusammenstellen.

Nach Def. 2.4.6. mit  $p=1$  lautet der MA(1)-Prozeß

$$\begin{aligned} z_t &= a_t - \theta_1 a_{t-1} \\ &= (1 - \theta_1 B) a_{t-1} \end{aligned} \quad (3.4.8)$$

Die Invertibilitätsbedingung lautet  $|\theta_1| < 1$  (vgl. Kapitel 3.1.5) und die Autokovarianzfunktion (vgl. Beispiel 3.1.2)

$$\begin{aligned} \gamma_0 &= (1 + \theta_1^2) \sigma_a^2 \\ \gamma_1 &= -\theta_1 \\ \gamma_k &= 0 \quad k \geq 2 \end{aligned} \quad (3.4.9)$$

Die ACF ist damit

$$\rho_k = \begin{cases} -\theta_1 / (1 + \theta_1^2) & k=1 \\ 0 & \text{sonst} \end{cases} \quad (3.4.10)$$

Für  $\rho_1$  gilt wegen  $|\theta_1| < 1$  folgende Ungleichung

$$|\rho_1| \leq \frac{|\theta_1|}{1 + \theta_1^2} \leq \frac{1}{2} \quad (3.4.11)$$

Für  $k=1$  in (3.4.10) erhalten wir eine quadratische Gleichung in  $\rho_1$  ( $\rho_1 \neq 0$ ) und dem Parameter des MA(1)-Prozesses

$$\theta_1^2 + \frac{\theta_1}{\rho_1} + 1 = 0 \quad (3.4.12)$$

Nach dem Vieta'schen Wurzelsatz gilt für die Wurzeln

$$\theta_1' \theta_1'' = 1 \quad (3.4.13)$$

Ist  $|\theta_1'| < 1 \Rightarrow |\theta_1''| > 1$ , d.h. nur eine der Wurzeln erfüllt die Invertibilitätsbedingung.

Beispiel 3.4.1.:

Die quadratische Gleichung (3.4.12) aufgelöst ergibt

$$\theta_1', '' = \frac{-1 \pm \sqrt{1 - 4\rho_1^2}}{2\rho_1} \quad (3.4.14)$$

Ist  $\rho_1 = -0,4$ , so ist  $\theta_1' = 2$  und  $\theta_1'' = 0,5$ . Nur  $\theta_1''$  entspricht der Invertibilitätsbedingung.

Die PACF

Setzen wir (3.3.10) in die Def. 3.2.2. ein, so erhalten wir

$$\begin{aligned} \phi_{11} &= \rho_1 \\ \phi_{22} &= \frac{\rho_1^2 - \rho_2}{\rho_1^2 - 1} = \frac{-\theta_1^2}{(1 - \theta_1^2)} (1 - \theta_1^2) \end{aligned} \quad (3.4.15)$$

Allgemein kann man zeigen, daß

$$\phi_{kk} = \frac{-\theta_1^k}{(1 - \theta_1^{2(k+1)})} (1 - \theta_1^2) \quad (3.4.16)$$

Man erkennt, daß  $|\phi_{kk}| < |\theta_1|^k$ , d.h. das Verhalten der PACF wird durch ein gedämpftes exponentielles Verhalten dominiert. Wir unterscheiden 2 Fälle

- (1)  $\varphi_1 > 0 \rightarrow \theta_1 < 0$  die PACF wechselt Vorzeichen
- (2)  $\varphi_2 < 0 \rightarrow \theta_1 > 0$  die PACF ist negativ

Man beachte die Dualität des AR(1)-Prozesses und des MA(1)-Prozesses bezüglich ihres Verhaltens der ACF und der PACF

	ACF $\varrho_k$	PACF $\phi_{kk}$
MA(1)	cut off	gedämpft exp.
AR(1)	gedämpft exp.	cut off

Wobei man unter "cut off" versteht, daß nur an der Stelle  $k=1$  ein Beitrag  $\neq 0$  vorliegt, für  $k \geq 0$  sind die Werte Null.

### 3.4.3. Der MA(2)-Prozeß

Der MA(2)-Prozeß ist definiert durch

$$\begin{aligned}
 z_t &= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} = \\
 &= (1 - \theta_1 B - \theta_2 B^2) a_t
 \end{aligned}
 \tag{3.4.17}$$

Er ist stationär für alle  $\theta_1, \theta_2 \in \mathbb{R}$

Die Invertibilitätsbedingungen für die Parameter erhalten wir durch Wurzeln von

$$(1 - \theta_1 B - \theta_2 B^2) = 0
 \tag{3.4.18}$$

Durch dieselben Umformungen wie für den AR(2)-Prozeß (vgl. (3.3.16) bis (3.3.21)) erhalten wir

$$\begin{aligned} |\theta_2| < 1, \quad |\theta_1| < 2 \\ \theta_2 + \theta_1 < 1 \\ \theta_2 - \theta_1 < 1 \end{aligned} \tag{3.4.19}$$

Varianz:  $\gamma_0 = \sigma_a^2 (1 + \theta_1^2 + \theta_2^2)$  (3.4.20)

ACF:

$$\begin{aligned} \rho_1 &= \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} = -\theta_1 + \theta_1\theta_2 \\ \rho_2 &= \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} \\ \rho_k &= 0 \quad k \geq 3 \end{aligned} \tag{3.4.21}$$

Unter Verwendung der Invertibilitätsbedingungen (3.4.19) und der ACF (3.4.21) liegen die Autokorrelationen eines MA(2)-Prozesses im folgenden Bereich

$$\begin{aligned} \rho_2 + \rho_1 &\geq -0,5 \\ \rho_2 - \rho_1 &\geq -0,5 \\ \rho_1^2 &\leq 4\rho_2(1-2\rho_2) \end{aligned} \tag{3.4.22}$$

Die PACF für einen MA(2)-Prozeß ist kompliziert abzuleiten. Sie wird dominiert von der Summe zweier gedämpfter Exponentialanteile, falls die Wurzeln von (3.4.18) reell sind, und von einer gedämpften Sinusschwingung, falls die Wurzeln komplex sind. Sie verhält sich so wie die ACF eines AR(2)-Prozesses.

An folgenden Abbildungen ist die Dualität des MA(2)- und AR(2)-Prozesses zu erkennen (vgl. Stralkowski et al. (1970)).

Fig.3.4.1.: Invertibles

MA(2)-Dreieck

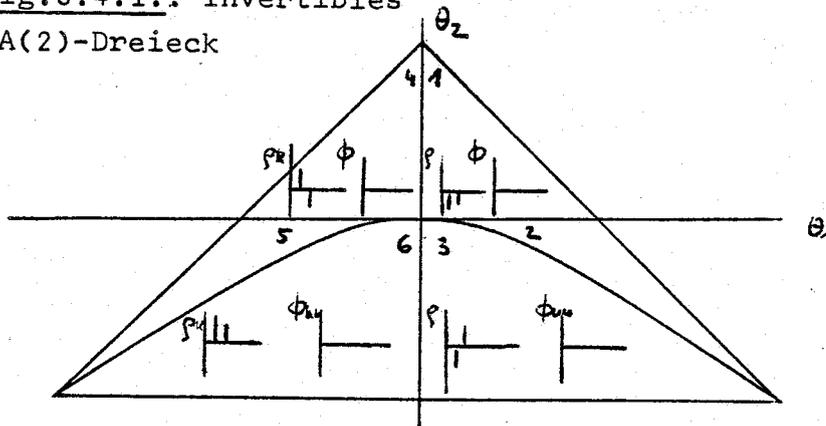
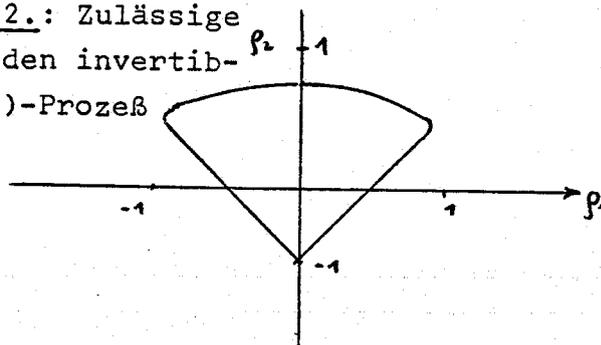


Fig.3.4.2.: Zulässige

ACF für den invertib-

len MA(2)-Prozeß



### 3.4.4. Dualität zwischen MA- und AR-Prozessen

Fassen wir die Ergebnisse der letzten Kapitel zusammen, so konzentrieren sich die Dualitätseigenschaften auf vier Punkte:

- (1) Ein stationärer AR-Prozeß kann als endliche gewichtete Summe vergangener  $z_t$  oder als unendliche gewichtete Summe vergangener  $a_t$  dargestellt werden.

$$\phi(B)z_t = a_t$$

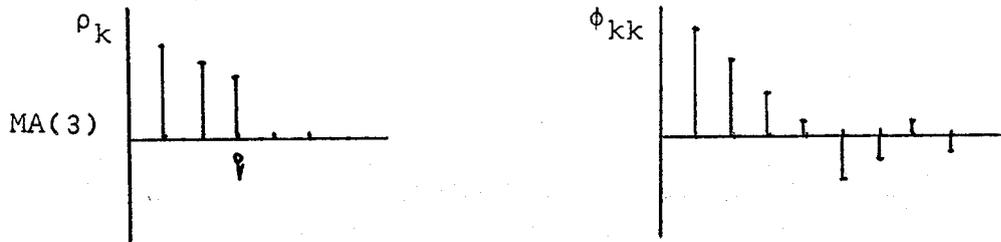
$$z_t = \phi^{-1}(B)a_t$$

Ein invertibler MA-Prozeß kann als endliche gewichtete Summe vergangener  $a_t$  oder als unendliche gewichtete Summe vergangener  $z_t$  dargestellt werden.

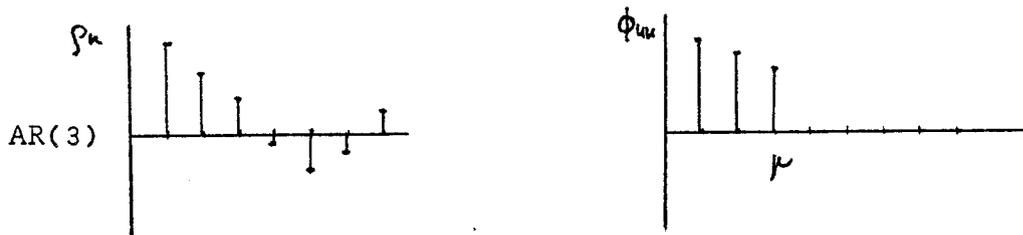
$$\theta^{-1}(B)z_t = a_t$$

$$z_t = \theta(B)a_t$$

- (2) Die ACF eines MA(q)-Prozesses ist Null für lags  $k > q$  ("cut off"), die PACF erstreckt sich nach  $+\infty$  und wird durch gedämpfte Exponential- oder Sinusschwingungen dominiert.



Die ACF eines AR(p)-Prozesses erstreckt sich nach  $+\infty$  und wird durch gedämpfte Exponential- und Sinusschwingungen dominiert, während die PACF Null für lags  $k > p$  ist.



- (3) Ein AR(p)-Prozeß ist immer invertierbar; um stationär zu sein, müssen die Wurzeln von  $\phi(B)=0$  außerhalb des Einheitskreises liegen (bzw.  $|\alpha_i| < 1$ ).

Ein MA(q)-Prozeß ist immer stationär; um invertierbar zu sein, müssen die Wurzeln von  $\theta(B)=0$  außerhalb des Einheitskreises liegen (bzw.  $|\beta_i| < 1$ ).

- (4) Das Spektrum eines MA-Prozesses besitzt eine inverse Beziehung zum Spektrum des korrespondierenden AR-Prozesses.

### 3.5. Autoregressive Moving Average Prozesse

#### 3.5.1. Der ARMA(p,q)-Prozeß

Nach Def. 2.7. lautet der ARMA(p,q)-Prozeß

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$
$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (3.5.1)$$

$$\phi(B) z_t = \theta(B) a_t$$

wobei  $\phi(B)$  das AR-Polynom vom Grad p und  
 $\theta(B)$  das MA-Polynom vom Grad q ist.

Ein gemischter Prozeß kann auf zwei Arten dargestellt werden:

(1) Als ein AR(p)-Prozeß in  $z_t$

$$\phi(B) z_t = e_t \quad (3.5.2)$$

mit einem MA(q)-Prozeß in  $e_t$

$$e_t = \theta(B) a_t \quad (3.5.3)$$

(2) Als ein MA(q)-Prozeß in  $b_t$

$$z_t = \theta(B) b_t \quad (3.5.4)$$

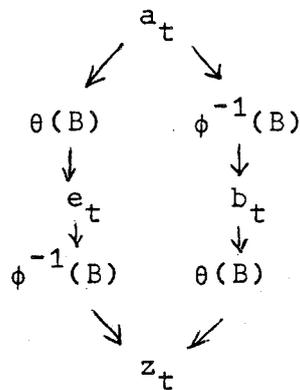
mit einem AR(p)-Prozeß in  $b_t$

$$\phi(B) b_t = a_t \quad (3.5.5)$$

sodaß

$$\phi(B) z_t = \theta(B) \phi(B) b_t = \theta(B) a_t \quad (3.5.6)$$

Die zweifache Darstellung sei in folgender Abbildung verdeutlicht



Die Stationaritätsbedingungen beziehen sich nur auf den AR(p)-Anteil, und für die Wurzeln der charakteristischen Gleichung

$$\phi(B) = \prod_{i=1}^p (1 - \alpha_i B) = 0 \quad (3.5.7)$$

gilt  $|\alpha_i| < 1$ ,  $i=1, \dots, p$

Die Invertibilitätsbedingungen beziehen sich nur auf den MA(q)-Anteil, und für die Wurzeln der Gleichung

$$\theta(B) = \prod_{i=1}^q (1 - \beta_i B) = 0 \quad (3.5.8)$$

gilt  $|\beta_j| < 1$ ,  $j=1, \dots, q$

Autokovarianzfunktion

Multiplizieren wir (3.4.1) mit  $z_{t-k}$  und nehmen Erwartungswerte so erhalten wir

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p} + \gamma_{za(k)} - \theta_1 \gamma_{za(k-1)} - \dots - \theta_q \gamma_{za(k-q)} \quad (3.5.9)$$

wobei  $\gamma_{za(k)} = E z_{t-k} a_t$  (3.5.10)

die Kreuzkovarianzfunktion der Zeitreihe  $z_t$  mit dem Störterm  $a_t$  ist.

Es gilt

$$\begin{aligned} \gamma_{za}(0) &= E z_t a_t \\ &= E a_t (\phi_1 z_{t-1} + \dots + a_t^{-\theta} \phi_1 a_{t-1}) \\ &= \sigma_a^2 \end{aligned} \quad (3.5.11)$$

Da die  $z_{t-k}$  nur  $a_t$ -Glieder bis zum Zeitpunkt  $t-k$  enthalten, gilt für  $\gamma_{za}(k)$

$$\gamma_{za}(k) = \begin{cases} \neq 0 & k < 0 \\ \sigma_a^2 & k = 0 \\ 0 & k > 0 \end{cases} \quad (3.4.12)$$

Wir erhalten damit für die ACovF

$$\gamma_k = \phi_1 \gamma_k + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \quad k \geq q+1 \quad (3.5.13)$$

Die ACF lautet

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad k \geq q+1 \quad (3.5.14)$$

bzw.  $\phi(B) \rho_k = 0 \quad k > q+1$

wobei B auf k wirkt.

Die ersten q Autokorrelationen  $\rho_1, \rho_2, \dots, \rho_q$  sind Funktionen sowohl von  $\theta$  als auch von  $\phi$ .

$$\rho_k = f(\theta, \phi) \quad 1 \leq k \leq q$$

Für das Verhalten der ACF sind zwei Fälle zu unterscheiden

- (1)  $p > q$  Die ACF besteht aus einer Mischung von gedämpften Exponential- und gedämpften Sinusschwingungen.

(2)  $p \neq q$  In diesem Falle gibt es  $q-p-1$  Anfangswerte  $\rho_0, \rho_1, \dots, \rho_{q-p}$  die nicht einem allgemeinen Schema folgen. Für  $j > q-p$  verhält sich die ACF wie im Fall (1).

### Varianz

Für  $k=0$  erhalten wir in (3.4.13)

$$\gamma_0 = \phi_1 \gamma_1 + \dots + \phi_p \gamma_p + \sigma^2 a^{-\theta_1} \gamma_{za(-1)} \dots a^{-\theta_q} \gamma_{za(-q)} \quad (3.5.15)$$

Da die Varianz von  $\gamma_1, \dots, \gamma_p$  abhängt, muß sie für jeden Prozeß zusammen mit  $p$  Gleichungen (3.5.9) gelöst werden.

### PACF:

Schreiben wir (3.5.1) in der Form

$$a_t = \theta^{-1}(B) \phi(B) z_t \quad (3.5.16)$$

so wird das Polynom

$$\pi(B) = \theta^{-1}(B) \phi(B) \quad (3.5.17)$$

durch  $p$  Werte  $\phi_p$  bestimmt, jedoch durch  $\infty$  viele Werte aufgrund reziproker  $\theta_j$ . Daher erstreckt sich die PACF ebenfalls nach  $+\infty$ . Sie ist der PACF eines MA( $q$ )-Prozesses ähnlich, die zusätzlich wegen  $\phi(B)$  durch gedämpfte Exponential- und Sinusschwingungen überlagert wird. Die Stärke hängt von der Anzahl und der Größe der Parameter  $\phi_p$  ab.

### 3.5.2. Der ARMA(1,1)-Prozeß

Der wichtigste ARMA( $p, q$ )-Prozeß für die Praxis (vgl. Prinzip der Parsimonität) ist der ARMA(1,1)-Prozeß. Nach (3.5.1) lautet dieser

$$\begin{aligned} z_t - \phi_1 z_{t-1} &= a_t - \theta_1 a_{t-1} \\ (1 - \phi_1 B) z_t &= (1 - \theta_1 B) a_t \end{aligned} \quad (3.5.18)$$

Aus der Stationaritätsbedingung folgt

$$|\phi_1| < 1 \quad (3.5.19)$$

Aus der Invertibilitätsbedingung folgt

$$|\theta_1| < 1 \quad (3.5.20)$$

Die ACovF

$$\begin{aligned} \gamma_0 &= \phi_1 \gamma_1 + \sigma_a^2 - \theta_1 \gamma_{za}(-1) \\ \gamma_1 &= \phi_1 \gamma_0 - \theta_1 \sigma_a^2 \\ \gamma_k &= \phi_1^k \gamma_{k-1} \quad k \geq 2 \end{aligned} \quad (3.5.21)$$

Zunächst berechnen wir die Kreuzkovarianzfunktion  $\gamma_{za}(-1)$   
Dazu multiplizieren wir (3.5.19) mit  $a_{t-1}$  und nehmen Erwartungswerte

$$\begin{aligned} E(z_t a_{t-1} - \phi_1 z_{t-1} a_{t-1}) &= E(e_t a_{t-1} - \theta_1 a_{t-1} a_{t-1}) \\ \gamma_{za}(-1) - \phi_1 \gamma_{za}(0) &= 0 - \theta_1 \sigma_a^2 \\ \gamma_{za}(-1) &= \phi_1 \sigma_a^2 - \theta_1 \sigma_a^2 \\ &= \sigma_a^2 (\phi_1 - \theta_1) \end{aligned} \quad (3.5.22)$$

Setzen wir (3.5.22) in die ACovF (3.5.21) ein, so erhalten wir für die Varianz

$$\begin{aligned} \gamma_0 &= \phi_1 (\phi_1 \gamma_0 - \theta_1 \sigma_a^2) + \sigma_a^2 - \theta_1 (\phi_1 - \theta_1) \sigma_a^2 \\ \gamma_0 - \phi_1^2 \gamma_0 &= \sigma_a^2 - \phi_1 \theta_1 \sigma_a^2 - \theta_1 \phi_1 \sigma_a^2 + \theta_1^2 \sigma_a^2 \\ \gamma_0 &= \frac{1 - 2\phi_1 \theta_1 + \theta_1^2}{1 - \phi_1^2} \sigma_a^2 \end{aligned} \quad (3.5.23)$$

und für die erste Autokovarianz

$$\begin{aligned}
 \gamma_1 &= \phi_1 \sigma_a^2 \frac{1 - 2\phi_1\theta_1 + \theta_1^2}{1 - \phi_1^2} - \theta_1 \sigma_a^2 \\
 &= \frac{\phi_1 - 2\phi_1^2\theta_1 + \theta_1^2\phi_1 - \theta_1 + \theta_1\phi_1^2}{1 - \phi_1^2} \sigma_a^2 \\
 &= \frac{\phi_1 - \theta_1 - \phi_1^2\theta_1 + \theta_1^2\phi_1}{1 - \phi_1^2} \sigma_a^2 \\
 &= \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 - \phi_1^2} \sigma_a^2 \tag{3.5.26}
 \end{aligned}$$

Die ACF lautet somit

$$\rho_1 = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 - 2\phi_1\theta_1} \tag{3.5.27}$$

$$\rho_2 = \phi_1 \rho_1$$

$$\rho_k = \phi_1 \rho_{k-1} = \phi_1^{k-1} \rho_{k-1}$$

Der zulässige ACF-Bereich ist

Wegen  $\rho_2 = \phi_1 \rho_1$  und  $|\phi_1| < 1$

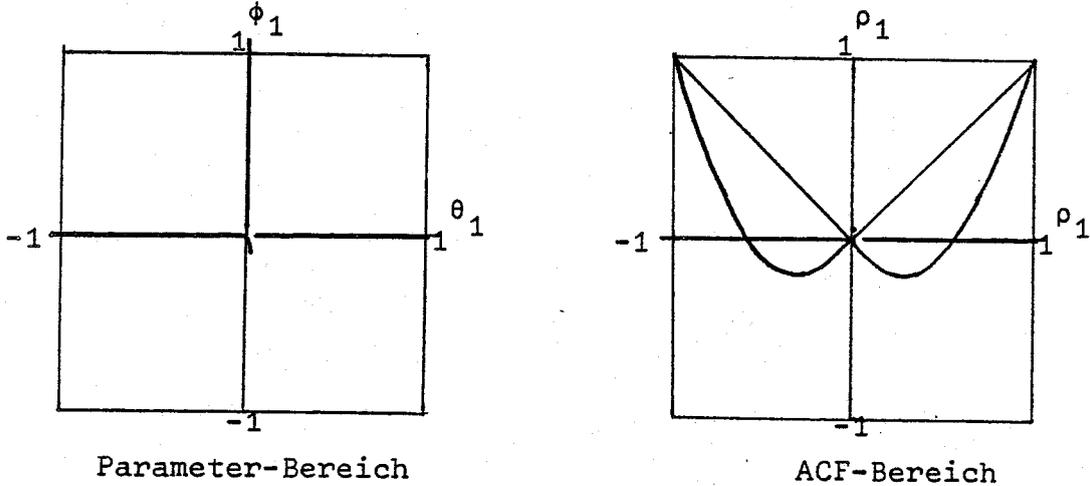
$$|\rho_2| = |\phi_1| |\rho_1|$$

$$|\rho_2| < |\rho_1|$$

$$\rho_2 > \rho_1(2\rho_1 + 1) \quad \rho_1 < 0 \tag{3.5.28}$$

$$\rho_2 > \rho_1(2\rho_1 - 1) \quad \rho_1 > 0$$

Fig.3.5.1. Stationärer und invertibler ARMA(1,1)-Prozeß



Die Ungleichungen (3.5.28) geben an, welche Kombination von  $\rho_1$  und  $\rho_2$  für einen zulässigen ARMA(1,1)-Prozeß möglich sind.

Die PACF eines ARMA(1,1)-Prozesses besteht aus einem einzigen Anfangswert  $\phi_{11} = \rho_1$ . Danach verhält sie sich wie die PACF eines MA(1)-Prozesses, d.h. wird durch einen gedämpften Exponentialanteil dominiert; wir unterscheiden zwei Fälle für  $\theta_1$  (obere und untere Hälfte) und zwei Fälle für  $\theta_1 - \phi_1$  (oberes oder unteres Dreieck):

(1)

$\theta_1 > 0$	$\theta_1 - \phi_1 < 0$	Ia	$\phi_{11} > 0$	$\phi_{kk}$ ged.exp.	$k > 1$
	$\theta_1 - \phi_1 > 0$	Ib, II	$\phi_{11} < 0$	$\phi_{kk}$ ged.exp.	$k > 1$

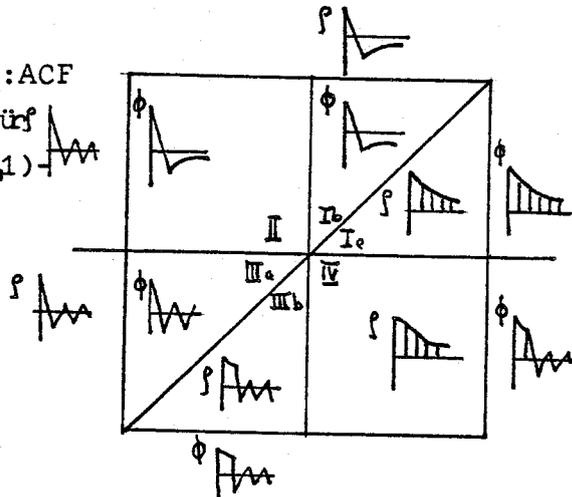
(2)

$\theta_1 < 0$	$\theta_1 - \phi_1 > 0$	IIIa	$\phi_{11} < 0$	$\phi_{kk}$ ged.oszill.	$k > 1$
	$\theta_1 - \phi_1 < 0$	IIIb, IV	$\phi_{11} > 0$	$\phi_{kk}$ ged.oszill.	$k > 1$

Die römischen Ziffern bezeichnen die Quadranten in folgender Abbildung

Fig.3.5.2.: ACF

und PACF für den ARMA(1,1) Prozeß



## 4. LINEARE NICHTSTATIONÄRE MODELLE

=====

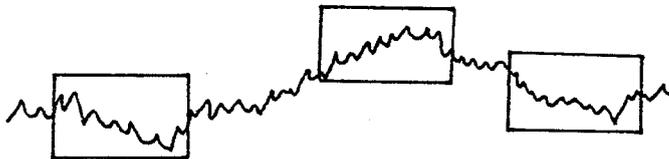
### 4.1. Allgemeines

#### 4.1.1. Einführende Beispiele

Eine grundlegende Annahme für stationäre Prozesse ist die Zeitunabhängigkeit des Mittelwertes.

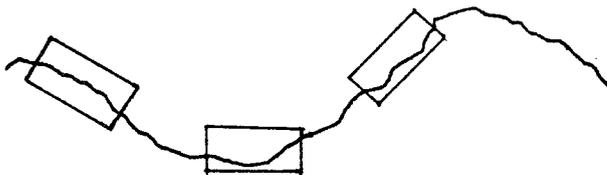
Viele empirische Zeitreihen verhalten sich so, als hätten sie keinen konstanten Mittelwert. Trotzdem zeigen sie eine Form der Homogenität. Dazu betrachten wir zwei Beispiele:

##### Beispiel 4.1.1.:



Diese Zeitreihe zeigt eine Nichtstationarität im Mittelwert an. Im jeweiligen Kästchen, nur durch das Niveau unterschieden, herrschen scheinbar dieselben Gesetze, die das Bild der Zeitreihe ausmachen.

##### Beispiel 4.1.2.:



Diese Zeitreihe zeigt ebenfalls ein Bild der Homogenität, jedoch sind hier nicht nur die Niveaus, sondern auch die Anstiege verschieden. Das wird durch die schiefe Lage der Kästchen angezeigt.

Der Lösungsansatz für die Box-Jenkins-Methode lautet:

Nichtstationarität im Mittelwert wird durch einmaliges Differenzenbilden eliminiert.

Nichtstationarität in Mittelwert und Anstieg wird durch zweimaliges Differenzenbilden eliminiert.

#### 4.1.2. Nichtstationarität als Grenzfall der Stationarität

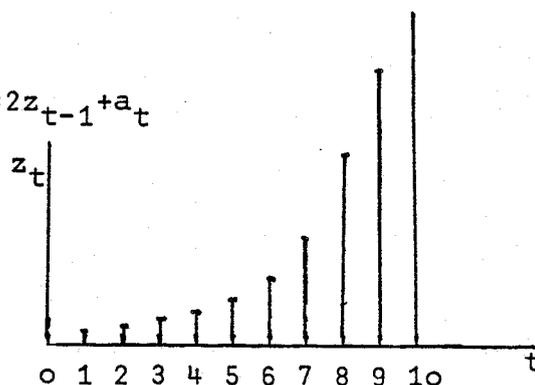
Es gibt eine Vielzahl von verschiedenen Arten der Nichtstationarität. Nichtstationarität in Mittelwert und Anstieg sind gerade diejenigen, die sich als Spezialfall von ARMA-Prozessen ergeben:

Beispiel 4.1.3.:  $z_t = 2z_{t-1} + a_t$

Diesen Prozeß können wir als AR(1)-Prozeß auffassen, der nicht der Stationaritätsbedingung  $|\phi| < 1$  genügt. Wir studieren das Verhalten für  $\phi=2$ , indem wir eine Stichprobe vom Umfang  $n=10$  aus  $a_t \sim N(0,1)$  wählen und den Startwert  $z_0=0,7$  vorgeben.

t	0	1	2	3	4	5	6	7	8	9	10
$a_t$		0.1	-1.1	0.2	-2.0	-0.2	-0.8	0.8	0.1	0.1	-0.9
$z_t$	0.7	1.5	1.9	4.0	6.0	11.8	22.8	46.4	92.5	185.9	370.9

Graph von  $z_t = 2z_{t-1} + a_t$



Der Prozeß explodiert für  $|\phi| > 1$ .

Betrachten wir nun den Fall, daß die Wurzeln von  $\phi(B)=0$  auf dem Einheitskreis liegen.

Der AR(1)-Prozeß mit  $\phi=1$  wird zum Differenzenoperator

$$(1-\phi B) = 1-B = \nabla \quad (4.1.1)$$

Der dazugehörige Prozeß ist der Random-walk-Prozeß (vgl. Kapitel 2.3.2 ). Der Random-walk-Prozeß ist somit als ein einfaches Modell von Nichtstationarität im Mittelwert auszudrücken.

Allgemein kann man nichtstationäre ARMA-Prozesse durch den verallgemeinerten AR-Operator (ARI-Operator) nach Def. 2.4.9. beschreiben

$$\begin{aligned} \gamma(B) &= \phi(B)\nabla^d = \\ &= \phi(B)(1-B)^d \end{aligned} \quad (4.1.2)$$

wobei  $\phi(B)$  ein stationärer AR-Operator der Ordnung  $p$  ist und  $\nabla^d$  den nichtstationären Anteil in Form der  $d$ -ten Differenz  $\nabla^d$  ausdrückt.

Nach Def. 2.2.9. schreiben wir ARIMA( $p,d,q$ )-Prozesse kurz als

$$\gamma(B)z_t = \theta(B)a_t \quad (4.1.3)$$

wobei  $\gamma(B)$  ein nichtstationärer Operator nach (4.1.2) und  $\theta(B)$  ein MA-Operator der Ordnung  $q$  ist. ARIMA( $p,d,q$ ) ist demnach eine Abkürzung für Autoregressive integrierte Moving-Average-Prozesse mit  $p$  autoregressiven Gliedern,  $d$  Differenzen und  $q$  Moving Average Gliedern.

Neben der Form (4.1.3) schreiben wir den ARIMA(p,d,q)-Prozeß auch

$$\phi(B)w_t = \theta(B)a_t \quad (4.1.4)$$

mit

$$w_t = \nabla^d z_t \quad (4.1.5)$$

Weiters gilt die Beziehung

$$\nabla^d z_t = \nabla^d(z_{t+\mu}) \quad (4.1.6)$$

weshalb für  $d > 0$  keine Prozesse als Abweichungen vom Mittelwert verwendet werden müssen.

Für spezielle ARIMA-Prozesse definieren wir folgende "parsimonische" Abkürzungsnotation:

Def. 4.1.1.: ARI- und IMA-Prozesse

- a) ARIMA(0,d,q)-Prozesse bezeichnen wir kurz als IMA(p,q)-Prozesse (integrierte Moving Average-Prozesse).
- b) ARIMA(p,d,0)-Prozesse bezeichnen wir kürzer als ARI(p,d)-Prozesse (autoregressive Integrierte Prozesse

Betrachten wir die Beziehung (4.1.5), so sehen wir, daß die Analyse von ARIMA-Prozessen die Annahme unterstellt, daß die d-te Differenz eines nichtstationären Prozesses als stationärer, invertierbarer ARIMA-Prozeß dargestellt werden kann.

Um einen differenzierten Prozeß in der ursprünglichen Form wieder darstellen zu können, wenden wir den entsprechenden Summationsoperator an.

$$z_t = S^d w_t \quad (4.1.7)$$

In dieser Form erkennt man, daß man nichtstationäre ARIMA-Prozesse dadurch erhält, daß man stationäre ARMA-Prozesse d-mal summiert. Die Autoregressiven "integrierten" Moving-Average-Prozesse sollten daher eher den Ausdruck "summiert" in ihrem Titel tragen.

Anmerkung:

Box-Jenkins (1970, S.92 ) weisen darauf hin, daß man eine Erweiterung der ARIMA-Prozesse dadurch erhält, daß man in ihrer Summationsdarstellung (4.1.7) eine deterministische Komponente in Form der sogenannten "Trendkonstanten" einfügt. Man kann dadurch ein Polynom d-ten Grades in die Analyse von ARIMA-Prozessen einfügen.

$$z_t = S^d(w_t + \theta_0) \quad (4.1.8)$$

bzw.

$$\psi(B)z_t = \theta_0 + \theta(B)a_t \quad (4.1.9)$$

Ist  $\theta_0 \neq 0$ , so erhält man

$$E(w_t) = E(V^d z_t) = \frac{\theta_0}{1 - \phi_1 - \dots - \phi_p} \quad (4.1.10)$$

Ist z.B.  $d=1$ , so erhält man durch  $\theta_0 \neq 0$  neben dem stochastischen Trend einen deterministischen in Form eines linearen Trends.

4.2. Darstellungsformen der ARIMA-Prozesse

Neben den beiden von den ARMA-Prozessen bekannten Darstellungsformen, der RS-Form und der  $\pi$ -Form (invertierten Form) unterscheiden wir bei den ARIMA-Prozessen noch zwei zusätzliche Darstellungsformen: Die Differenzengleichungsform (D-Form) und die Truncated Random Stock Form (TRS-Form). Für IMA(p,q)-Prozesse gibt es noch zusätzlich die Darstellung in der Differenzenoperatorform (V-Form). Zusätzlich kann man bei ARIMA(p,d,q)-Prozesse mit  $d>0$  die  $\pi$ -Form als "normierte"  $\pi$ -Form beschreiben.

#### 4.2.1. Die Differenzgleichungsform

Def. 4.2.1.: D-Form

Unter der D-Form des ARIMA-Prozesses

$$\phi(B)\nabla^d z_t = \theta(B)a_t \quad (4.2.1)$$

verstehen wir die Schreibweise

$$\psi(B)z_t = \theta(B)a_t \quad (4.2.2)$$

d.h.

$$z_t = \psi_1 z_{t-1} + \dots + \psi_{p+d} z_{t-p-d} - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} + a_t \quad (4.2.3)$$

Die D-Form ist die "Standardform" für ARIMA-Prozesse.

Anmerkung: Ist  $\psi^*(B)$  die zum ARIMA-Prozeß gehörige Transferfunktion des ARMA(p,q)-Prozesses, so besteht zwischen der Transferfunktion die Beziehung

$$\psi(B) = S^d \psi^*(B)$$

bzw. für den ARIMA-Prozeß

$$z_t = S^d \psi^*(B) a_t$$

#### 4.2.2. Die Random Shock-Form

Analog zu den ARMA-Prozessen können wir  $z_t$  als Funktion der vergangenen Störungen ansetzen.

Def. 4.2.2.: Die Random-Shock-Form (RS-Form) lautet

$$\begin{aligned} z_t &= a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \\ &= \sum_{j=0}^{\infty} \psi_j a_{t-j} \\ &= \psi(B) a_t \end{aligned} \quad (4.2.4)$$

Satz 4.2.1.: Man beachte, daß für nichtstationäre Modelle die Summe (4.2.4) nicht konvergiert.

Die  $\psi_j$ -Koeffizienten erhält man durch Koeffizientenvergleich der Polynome

$$\varphi(B)\Psi(B) = \theta(B) \quad (4.2.5)$$

Beweis:

Wir multiplizieren (4.2.4) mit  $\psi(B)$

$$\varphi(B)z_t = \varphi(B)\Psi(B)a_t \quad (4.2.6)$$

Setzen wir (4.2.2) ein, erhalten wir

$$\theta(B)a_t = \varphi(B)\Psi(B)a_t \quad (4.2.7)$$

(4.2.5) in ausgeschriebener Form lautet

$$(1 - \varphi_1 B - \dots - \varphi_{p+d} B^{p+d})(1 + \psi_1 B + \psi_2 B^2 + \dots) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \quad (4.2.8)$$

Für  $j > \max(p+d-1, q)$  erfüllen die  $\psi$ -Koeffizienten die Differenzgleichung, die durch den verallgemeinerten Autoregressiven Operator  $\psi(B)$  definiert ist

$$\varphi(B)\psi_j = \phi(B)(1-B)^d \psi_j = 0 \quad (4.2.9)$$

wobei  $B$  auf den Index  $j$  wirkt. Für genügend große  $j$  bestehen die  $\psi_j$ -Koeffizienten aus einer Mischung von Polynomen, gedämpften exponentiellen und gedämpften Sinusschwingungen.

#### 4.2.3. Die invertierte Form des Modells

Def. 4.2.3.: Die invertierte Form ( $\pi$ -Form) eines ARIMA-Prozesses ist gegeben durch

$$a_t = \pi(B)z_t \quad (4.2.10)$$

mit

$$\pi(B) = (1 - \pi_1 B - \pi_2 B^2 - \dots) \quad (4.2.11)$$

$$\text{bzw. } z_t = \pi_1 z_{t-1} + \pi_2 z_{t-2} + \dots + a_t \quad (4.2.12)$$

$z_t$  wird dadurch als unendliche Summe vergangener  $z_t$  plus einer zufälligen Störung  $a_t$  dargestellt.

Satz 4.2.1.:

$\pi(B)$  konvergiert innerhalb des Einheitskreises und man erhält die  $\pi_j$  durch Koeffizientenvergleich von

$$\varphi(B) = \theta(B)\pi(B) \quad (4.2.13)$$

Beweis:

Multiplizieren wir (4.2.10) mit  $\theta(B)$ , so erhalten wir

$$\theta(B)a_t = \theta(B)\pi(B)z_t \quad (4.2.14)$$

Darin Formel (4.2.2) eingesetzt, ergibt

$$\varphi(B)z_t = \theta(B)\pi(B)z_t \quad (4.2.15)$$

und daraus die Behauptung (4.2.13).

Weiters existiert  $\theta^{-1}(B)$ , da der zugehörige ARMA-Prozeß invertierbar ist, und es gilt

$$\pi(B) = \theta^{-1}(B)\varphi(B) \quad (4.2.16)$$

d.h. die erzeugende Funktion der  $\pi(s)$  konvergiert innerhalb  $|s| \leq 1$

Verhalten der  $\pi$ -Koeffizienten

(4.2.13) ausgeschrieben lautet

$$(1 - \varphi_1 B - \dots - \varphi_{p+d} B^{p+d}) = (1 - \theta_1 B - \dots - \theta_q B^q)(1 - \pi_1 B - \pi_2 B^2 - \dots) \quad (4.2.17)$$

Für  $j > \max(p+d, q)$  (4.2.18)

erfüllen die  $\pi$ -Gewichte die Differenzengleichung des MA-Operators

$$\theta(B)\pi_j = 0 \quad (4.2.19)$$

wobei B auf j wirkt.

Für genügend große j verhalten sich die  $\pi$ -Koeffizienten wie die ACF eines AR-Prozesses, d.h. sie werden durch gedämpfte Exponentialanteile und/oder gedämpfte Sinusschwingungen dominiert.

Satz 4.2.2.:

Für  $d \geq 1$  ist die Summe der  $\pi$ -Koeffizienten eins.

Beweis:

Dazu schreiben wir (4.2.13) in der Form

$$\phi(B)(1-B)^d = \theta(B)\pi(B) \quad (4.2.20)$$

und setzen  $B=1$ , so erhalten wir

$$0 = \theta(1)\pi(1) \quad (4.2.21)$$

Nun ist

$$\begin{aligned} \theta(1) &= (1-\theta_1-\theta_2-\dots-\theta_q) = \\ &= \prod_{j=1}^q (1-\beta_j) \neq 0 \end{aligned} \quad (4.2.22)$$

da die Wurzeln von  $\theta(B)$  außerhalb des Einheitskreises liegen.

Daher gilt

$$\pi(1) = (1-\pi_1-\pi_2-\dots) = 0 \quad (4.2.23)$$

bzw.

$$\sum_{j=0}^{\infty} \pi_j = 1 \quad (4.2.24)$$

und damit folgt die Behauptung.

Def. 4.2.4.: Normierte D-Form

Wegen Satz 4.2.2. schreiben wir die D-Form für  $d \geq 1$  auch als

$$\begin{aligned} z_t &= \pi_1 z_{t-1} + \pi_2 z_{t-2} + \dots + a_t \\ &= \bar{z}_{t-1}(\pi) + a_t \end{aligned} \quad (4.2.25)$$

wobei

$$\bar{z}_{t-1}(\pi) = \sum_{j=1}^{\infty} \pi_j z_{t-j} \quad (4.2.26)$$

ein gewogener Durchschnitt der vergangenen  $z_t$  bis zum Zeitpunkt  $t-1$  ist.

4.2.4. Die Truncated Random Shock Form (TRS-Form)

Die TRS-Form ist mit der RS-Form eines ARIMA-Modelles stark verbunden und wird in spezieller Form für die Darstellung der Prognosefunktion verwendet. Die TRS-Form ist eine endliche Form der RS-Form für den "truncation point"  $k$  <sup>+)</sup> .

Def. 4.2.5.: TRS-Form

Unter der TRS-Form eines ARIMA-Prozesses verstehen wir die Schreibweise

$$z_t = C_k(t-k) + \sum_{j=k+1}^t \psi_{t-j} a_j \quad (4.2.27)$$

wobei  $C_k(t-k)$  die abgeschnittene (truncated) unendliche Summe der RS-Form ist.

$$C_k(t-k) = \sum_{j=-\infty}^k \psi_{t-j} a_j \quad (4.2.28)$$

Wird die RS-Form eines ARIMA-Prozesses als unendliche gewichtete Summe der vergangenen Störungen dargestellt, so wird in der TRS-Form  $z_t$  als gewogene endliche Summe der letzten  $t-k$  Störungen dargestellt, wobei die Störungen vor dem variablen Startzeitpunkt  $k$  in der Funktion  $C_k(t-k)$  zusammengefaßt werden.

<sup>+)</sup>  "Truncation point" könnte man am besten noch als "Stützpunkt" (von "abstutzen") bezeichnen.

Satz 4.2.3.: TRS-Form als Lösung der ARIMA Differenzengleichung

Die TRS-Form des ARIMA-Prozesses

$$\varphi(B)z_t = \theta(B)a_t \quad (4.2.29)$$

betrachtet die Zeitreihe  $z_t$  als Lösung der stochastischen Differenzengleichung mit

$$z_t = C_k(t-k) + I_k(t-k) \quad (4.2.30)$$

wobei  $C_k(t-k)$  die allgemeine Lösung der homogenen Differenzengleichung

$$\varphi(B)C_k(t-k) = 0 \quad (4.2.31)$$

und  $I_k(t-k)$  eine partikuläre Lösung der inhomogenen Differenzengleichung ist

$$\varphi(B)I_k(t-k) = \theta(B) \quad (4.2.32)$$

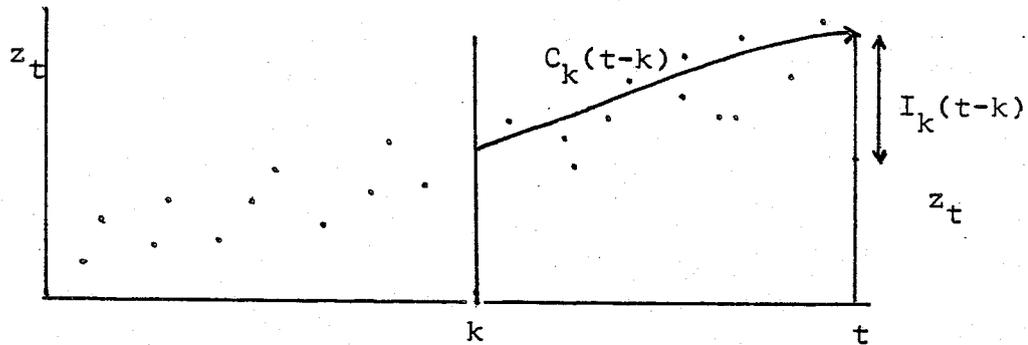
(wobei B auf t wirkt und nicht auf k)...

Die partikuläre Lösung lautet

$$I_k(s-k) = \begin{cases} 0 & s \leq k \\ \sum_{j=k+1}^s \psi_{s-j} a_j = \sum_{l=0}^{s-k-1} \psi_l a_{t-l} & s > k \end{cases} \quad (4.2.33)$$

Der Beweis findet sich in Kapitel 4.6.

Zur Illustration betrachten wir folgende Abbildung



Die TRS-Form besagt, daß der Wert einer Zeitreihe in zwei Anteile aufgespalten werden kann.

- a) Die homogene Lösung  $C_k(t-k)$  ist jener Anteil, der zum Zeitpunkt  $k$  bereits fest vorgegeben ist. Sie enthält alle Information der Beobachtungen der Zeitreihe vor dem Zeitpunkt  $k_0=k$ , der für die Erklärung des Wertes  $z_t$  relevant ist. Sie gibt an, in welche Richtung sich der Prozeß entwickelt, falls der Einfluß der Störungen  $a_t$  zum Zeitpunkt  $k$  "abgedreht" würde.
- b) Die partikuläre Lösung  $I_k(t-k)$  gibt jenen Anteil an der Beobachtung  $z_t$  an, der zum Zeitpunkt  $k$  nicht vorhersehbar ist. Sie faßt den Einfluß der Störungen ab dem Zeitpunkt  $k$  zusammen.

Die homogene Lösung  $C_k(t-k)$  kann auch in der Form geschrieben werden

$$\begin{aligned}
 C_k(t-k) &= b_0^{(k)} + b_1^{(k)}(t-k) + b_{d-1}^{(k)}(t-k)^{d-1} + b_d^{(k)} G_{d-1}^{t-k} + b_{p+d-1}^{(k)} G_p^{t-k} \\
 &= G_0^{(t-k)} \sum_{j=0}^{d-1} A_j (t-k)^j + \sum_{i=1}^p D_i c_i^{t-k}
 \end{aligned}
 \tag{4.2.34}$$

wobei

$$b_i^{(k)} = \begin{cases} G_0 A_i & i \leq d-1 \\ G_{d-i+1} D_{d-i+1} & i \geq d \end{cases}
 \tag{4.2.35}$$

#### 4.2.5. Die Differenzenoperatorform

Ziel der  $\nabla$ -Form ist es, das MA-Polynom eines integrierten Moving Average-Prozesses so umzuformen, daß durch Anwendung des Summationsoperators die RS-Form (und damit auch die TRS-Form) leicht erhältlich wird.

Def. 4.2.6.: Die  $\nabla$ -Form

Die  $\nabla$ -Form eines IMA(p,q)-Prozesses lautet

$$\begin{aligned}\nabla^d z_t &= \nabla^d a_t + \lambda(\nabla) a_{t-1} \\ &= (\nabla + \lambda(\nabla)B) a_t\end{aligned}\tag{4.2.36}$$

wobei

$$\lambda(\nabla) = \lambda_0 + \lambda_1 \nabla + \dots + \lambda_{q-1} \nabla^{q-1}\tag{4.2.37}$$

ein Operatorpolynom in  $\nabla$  der Ordnung  $q-1$  ist.

Die RS-Form erhält man durch Anwendung des Summationsoperators  $S^d$

$$z_t = a_t + (\lambda_0 S^d + \dots + \lambda_{q-1}) a_{t-1}\tag{4.2.38}$$

Die Koeffizienten  $\lambda$  erhalten wir durch Koeffizientenvergleich von

$$\nabla^d + \lambda(\nabla)B = \theta(B)\tag{4.2.39}$$

4.2.6. Die Bestimmung der  $\psi_j$ - und  $\pi_j$ -Koeffizienten

In Kapitel 3.1.3. haben wir die Beziehung zwischen der RS- und der  $\pi$ -Form eines allgemeinen ARMA(p,q)-Prozesses beschrieben. Im folgenden erweitern wir diese Beziehung auf ARIMA(p,d,q)-Prozesse und bestimmen die Formeln zur Berechnung der  $\psi_j$ - und der  $\pi_j$ -Koeffizienten. Diese Koeffizienten sind für die Darstellung der Prognosefunktion von großer Bedeutung (vgl. Kapitel 5.1.), die  $\psi_j$ -Koeffizienten werden weiters noch für die Berechnung der Varianzfunktion und für das Updaten von Prognosen verwendet.

a) Die Berechnung der  $\psi_j$ -Koeffizienten

Nach Satz 4.2.1. erhalten wir die  $\psi$ -Koeffizienten durch Koeffizientenvergleich von

$$(1 - \gamma_1 B - \dots - \gamma_{p'} B^{p'}) (1 + \psi_1 B + \psi_2 B^2 + \dots) = (1 - \theta_1 B - \dots - \theta_q B^q) \quad (4.2.40)$$

wobei  $p' = p + d$  ist. Ausführlich angeschrieben lautet diese Gleichung

$$\begin{aligned} 1 - \gamma_1 B - \gamma_2 B^2 - \dots - \gamma_{p+d} B^{p+d} + \\ \psi_1 B - \gamma_1 \psi_1 B^2 - \dots - \psi_1 \gamma_{p+d} B^{p+d+1} \\ \psi_2 B^2 - \gamma_1 \psi_2 B^2 - \dots \\ \dots = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \end{aligned} \quad (4.2.41)$$

Koeffizientenvergleich der  $B^j$  ergibt

$$\begin{aligned} B^0: \quad \psi_0 &= 1 = \theta_0 = 1 \\ B^1: \quad \psi_1 - \gamma_1 &= -\theta_1 \\ B^2: \quad \psi_2 - \gamma_2 - \gamma_1 \psi_1 &= -\theta_2 \end{aligned} \quad (4.2.42)$$

nach aufsteigenden  $\psi_j$  rekursiv ausgedrückt

$$\begin{aligned} \psi_1 &= \gamma_1 - \theta_1 \\ \psi_2 &= \gamma_1 \psi_1 + \gamma_2 - \theta_2 \\ &\vdots \\ \psi_j &= \gamma_1 \psi_{j-1} + \dots + \gamma_{p+d} \psi_{j-p-d} - \theta_j \quad j \leq q \end{aligned} \quad (4.2.43)$$

Für  $k = \text{Max}(q, p+d-1)$

erfüllen die  $\psi_j$  die Differenzengleichung

$$\psi_j = \gamma_1 \psi_{j-1} + \gamma_2 \psi_{j-2} + \dots + \gamma_{p+d} \psi_{j-p-d} \quad j \leq k \quad (4.2.44)$$

bzw.

$$\gamma(B) \psi_j = 0 \quad j > k \quad (4.2.45)$$

Beispiel 4.2.1.: ARIMA(1,1,0)-Prozeß

Man bestimme die Koeffizienten für  $\phi_1 = 0,8$

$$(1 - 0,8B)(1 - B)z_t = a_t$$

$$(1 - 1,8B + 0,8B^2)z_t = a_t$$

Koeffizientenvergleich für die  $\psi_j$   $\gamma_1 = 1,8, \gamma_2 = -0,8$

$$(1 - 1,8B + 0,8B^2)(1 + \psi_1 B + \psi_2 B^2 + \dots) = 1$$

entweder direkter Koeffizientenvergleich oder unter Anwendung von Formel (4,2.44)

a) direkter Koeffizientenvergleich:

$$\begin{aligned} & 1 - 1,8B + 0,8B^2 + \\ & + \psi_1 B - 1,8\psi_1 B^2 + 0,8\psi_1 B^3 + \\ & \quad + \psi_2 B^2 - 1,8\psi_2 B^3 + 0,8\psi_2 B^4 + \\ & \quad \quad + \psi_3 B^3 - 1,8\psi_3 B^4 + 0,8\psi_3 B^5 + \\ & \quad \quad \quad + \psi_4 B^4 + \dots = 1 \end{aligned}$$

$$\begin{aligned} B^0 & \quad \psi_0 = 1 \\ B^1 & \quad -1,8 + \psi_1 = 0 \\ B^2 & \quad 0,8 - 1,8\psi_1 + \psi_2 = 0 \\ B^3 & \quad 0,8\psi_1 - 1,8\psi_2 + \psi_3 = 0 \\ B^4 & \quad 0,8\psi_2 - 1,8\psi_3 + \psi_4 = 0 \\ B^j & \quad 0,8\psi_{j-2} - 1,8\psi_{j-1} + \psi_j = 0 \end{aligned}$$

oder

$$\begin{aligned}\psi_0 &= 1 \\ \psi_1 &= 1,8 & j=1 \\ \psi_j &= 1,8B\psi_{j-1} - 0,8\psi_{j-2} & j=2,3\end{aligned}$$

b) nach Formel (4.2.44) gilt

$$k = \text{Max}(q, p+d-1) = \text{Max}(0, 1) = 1$$

Dann ist

$$\begin{aligned}\psi_0 &= 1 \\ \text{und} \quad \psi_1 &= \gamma_1 - \theta_1 = 1,8\end{aligned}$$

$$\text{für } j > k=1 \quad p+d=2$$

$$\psi_j = \gamma_1 \psi_{j-1} + \gamma_2 \psi_{j-2}$$

oder die  $\gamma_i$  eingesetzt ergibt die Differenzengleichung in  
verschiedener Form

$$\psi_j = 1,8\psi_{j-1} - 0,8\psi_{j-2} \quad j > k=1$$

$$(1 - 1,8B + 0,8B^2)\psi_j = 0$$

$$\gamma(B)\psi_j = 0 \quad j > k=1$$

wobei B auf j wirkt.

c) speziell:

$$\psi_0 = 1 \quad \psi_1 = 1,8 \quad \gamma_1 = 1,8 \quad \gamma_2 = -0,8$$

$$\psi_2 = 1,8 \cdot 1,8 + (-0,8) \cdot 1 = 3,24 - 0,8 = 2,44$$

$$\psi_3 = 1,8 \cdot 2,44 + (-0,8) \cdot 1,8 = 2,95$$

b) Die Berechnung der  $\pi_j$ -Koeffizienten

Analog dem Satz 3.1.1. gilt für die ARIMA-Prozesse die Beziehung

$$\psi(B)\pi(B) = 1$$

jedoch existiert im Unterschied zu (3.1.9) das inverse Polynom  $\psi^{-1}(B)$  nicht. Die Berechnung der  $\pi_j$ -Koeffizienten erfolgt durch Koeffizientenvergleich

$$(1 + \psi_1 B + \psi_2 B^2 + \dots)(1 - \pi_1 B - \pi_2 B^2 - \dots) = 1$$

Die Koeffizienten lauten

$$\begin{aligned} \psi_1 &= \pi_1 \\ \psi_2 &= \psi_1 \pi_1 + \pi_2 \\ \psi_3 &= \psi_2 \pi_1 + \psi_1 \pi_2 + \pi_3 \\ &\vdots \\ \psi_j &= \sum_{i=1}^j \psi_{j-i} \pi_i \quad j \geq 1 \end{aligned} \quad (4.2.46)$$

Umgekehrt erhält man daraus für die  $\pi_j$ -Koeffizienten

$$\begin{aligned} \pi_1 &= \psi_1 \\ \pi_2 &= \psi_2 - \psi_1 \pi_1 \\ \pi_3 &= \psi_3 - \psi_2 \pi_1 - \psi_1 \pi_2 \\ &\vdots \\ \pi_j &= \psi_j - \sum_{i=1}^{j-1} \psi_{j-i} \pi_i \quad j \geq 1 \end{aligned} \quad (4.2.47)$$

### 4.3. Der ARIMA(1,1,1)-Prozeß

#### 4.3.1. D-Form

Der ARIMA(1,1,1)-Prozeß ist nach Def. 4.1.2. gegeben durch

$$(1-\phi B)(1-B)z_t = (1-\theta_1 B)a_t \quad (4.3.1)$$

mit verallgemeinertem AR-Operator (vgl. Def. 4.1.1)

$$\begin{aligned} \varphi(B) &= (1-\phi B)(1-B) \\ &= (1+(1-\phi)B+\phi B^2) \end{aligned} \quad (4.3.2)$$

mit

$$\begin{aligned} \varphi_0 &= 1 \\ \varphi_1 &= 1-\phi \\ \varphi_2 &= \phi \\ \varphi_k &= 0 \quad k \geq 3 \end{aligned} \quad (4.3.3)$$

Die D-Form lautet damit

$$z_t = (1+\phi)z_{t-1} - \phi z_{t-2} + a_t - \theta a_{t-1} \quad (4.3.4)$$

#### 4.3.2. RS-Form

Die RS-Form erhält man durch Koeffizientenvergleich

$$(1-(1+\phi)B+\phi B^2)(1+\psi_1 B+\psi_2 B^2+\dots) = (1-\theta B) \quad (4.3.5)$$

Es ist

$$\begin{aligned} B & \quad -(1+\phi)+\psi_1 = -\theta \\ B^2 & \quad \phi-\psi_1(1+\phi)+\psi_2 = 0 \\ & \quad \vdots \\ B^k & \quad \psi_{k-2}\phi+\psi_{k-1}(1+\phi)+\psi_k = 0 \end{aligned} \quad (4.3.6)$$

bzw.

$$\varphi(B)\psi_k = 0 \quad (4.3.7)$$

Das entspricht Formel (4.2.9) mit  $j > \max(1+1-1, 1) = 1$ .

Wir lösen nun die Differenzgleichung (4.3.7) mit Hilfe der charakteristischen Gleichung

$$(1-\phi B)(1-B) = (1-G_1 B)(1-G_2 B) = 0 \quad (4.3.8)$$

mit  $G_1 = 1/\phi$  und  $G_2 = 1$  (4.3.9)

Damit lautet die allgemeine Lösung

$$\begin{aligned} \psi_j &= A_0 G_1^j + A_1 G_2^j \\ &= A_0 + A_1 \phi^j \end{aligned} \quad (4.3.10)$$

Die Konstanten werden durch die Anfangswerte bestimmt:

$$\psi_0 = A_0 + A_1 = 1 \quad (4.3.11)$$

$$\begin{aligned} \psi_1 &= A_0 + A_1 \phi = \\ &= 1 + \phi - \theta \end{aligned} \quad (4.3.12)$$

mit  $\psi_1$  aus (4.3.6). Wir erhalten nun

$$A_1 = \frac{\theta - \phi}{1 - \phi} \quad (4.3.13)$$

$$A_0 = \frac{1 - \theta}{1 - \phi} \quad (4.3.14)$$

Die RS-Form des ARIMA(1,1,1)-Prozesses lautet nun

$$\begin{aligned} z_t &= \sum_{j=0}^{\infty} \psi_j a_{t-j} \\ &= \sum_{j=0}^{\infty} (A_0 + A_1 \phi^j) a_{t-j} \end{aligned} \quad (4.3.15)$$

In dieser Form erkennt man, daß wegen  $|\phi| < 1$  für  $j \rightarrow \infty$   $\phi_j \rightarrow 0$  geht. Das bedeutet, daß die Störungen eines (1,1,1)-Prozesses ein annähernd konstantes Gewicht  $A_0$  bekommen, je größer  $j$  ist, d.h. je länger man in die Vergangenheit zurückschreitet.

### 4.3.3. Die $\pi$ -Form

Nach Satz 4.2.1. erhalten wir die  $\pi$ -Koeffizienten durch Koeffizientenvergleich von

$$\varphi(B) = \theta(B)\pi(B) \quad (4.3.16)$$

bzw. durch

$$\pi(B) = \varphi(B)\theta^{-1}(B)$$

$$(1-\pi_1 B - \pi_2 B^2 - \dots) = (1-(1+\phi)B + \phi^2 B^2)(1+\theta B + \theta^2 B^2 + \dots) \quad (4.3.17)$$

durch Koeffizientenvergleich erhalten wir

$$\begin{aligned} B \quad \pi_1 &= (1-\theta) + \phi \\ B^2 \quad \pi_2 &= (1-\theta)(\theta-\phi) \\ B^j \quad \pi_j &= \theta^{j-2}(1-\theta)(\theta-\phi) \quad j \geq 2 \end{aligned} \quad (4.3.18)$$

Dazu kommen wir auch wegen (4.2.18)  $j > \max(2, 1) = 2$  und (4.2.19)

$$\begin{aligned} (1-\theta B)\pi_j &= 0 \\ \pi_j &= \theta^{j-2}\pi_2 \end{aligned} \quad (4.3.19)$$

Die  $\pi$ -Form lautet somit

$$z_t = (1-\theta+\phi)z_{t-1} + (1-\theta)(\theta-\phi) \sum_{j=2}^{\infty} \theta^{j-2} z_{t-j} + a_t \quad (4.3.20)$$

Nach Satz 4.2.2. ist die Summe der  $\pi$ -Koeffizienten eins:

$$\begin{aligned} \sum_{j=1}^{\infty} \pi_j &= (1-\theta) + \phi + (1-\theta)(\theta-\phi) \sum_{j=2}^{\infty} \theta^{j-2} \\ &= (1-\theta) + \phi + (\theta-\phi) \\ &= 1 \end{aligned} \quad (4.3.21)$$

Beispiel 4.3.1.:

Zur Illustration betrachten wir einen ARIMA(1,1,1)-Prozeß mit den Koeffizienten  $\phi = -0,3$  und  $\theta = 0,5$

Die D-Form lautet

$$(1+0,3B)(1-B)z_t = (1-0,5B)a_t$$

$$(1-0,7B-0,3B^2)z_t = (1,05B)a_t$$

bzw.

$$z_t = 0,7z_{t-1} + 0,3z_{t-2} + a_t - 0,5a_{t-1}$$

Die  $\pi$ -Gewichte und  $\psi$ -Gewichte lauten

j	0	1	2	3	4	5	6	7
$\pi_j$	1	0,2	0,4	0,2	0,1	0,05	0,025	
$\psi_j$	1	0,58	0,454	0,41	0,41	0,4	0,4	

$$\pi_1 = (1-\theta) + \phi = 0,2$$

$$\pi_2 = (1-\theta)(\theta - \phi) = 0,5 \cdot 0,8 = 0,4$$

$$\pi_j = 0,5\pi_{j-1} \quad j > 2$$

$$\psi_j = A_0 + A_1 \phi^j$$

$$A_0 = \frac{1-\theta}{1-\phi} = \frac{0,5}{1,3} \quad 0,4$$

$$A_1 = \frac{\theta-\phi}{1-\phi} = \frac{0,8}{1,3} \quad 0,6$$

$$\psi_j = 0,4 + 0,6(-0,3)^j$$

$$\psi_0 = 1$$

Aus der Größenordnung der  $\pi$ -Koeffizienten erkennt man, daß der Großteil (in diesem Beispiel 95%) der Information über vergangene  $a_t$  in den letzten Werten (in diesem Fall 5) der  $z_t$  zusammengefaßt werden kann.

#### 4.3.4. Die TRS-Form

Nach Def. 4.2.5 erhalten wir die TRS-Form durch

$$z_t = C_k(t-k) + \sum_{j=k+1}^t \psi t^{-j} a_j \quad (4.3.22)$$

Wir schreiben die RS-Form (4.3.15) des ARIMA (1,1,1)-Prozesses in Vorwärtssummation

$$z_t = \sum_{j=-\infty}^t (A_0 + A_1 \phi^{t-j}) a_j \quad (4.3.23)$$

Die TRS-Form lautet somit

$$z_t = C_k(t-k) + \sum_{j=k+1}^t (A_0 + A_1 \phi^{t-j}) a_j \quad (4.3.24)$$

mit

$$C_k(t-k) = \sum_{j=-\infty}^k (A_0 + A_1 \phi^{t-j}) a_j \quad (4.3.25)$$

wobei  $A_0$  und  $A_1$  die nach (4.3.13) bzw. (4.3.14) bestimmten Konstanten sind.

$C_k(t-k)$  ist die Lösung der homogenen Differenzgleichung

$$(1-\phi B)(1-B) = 0 \quad (4.3.26)$$

und kann daher in der Form (vgl. 4.2.34) geschrieben werden:

$$C_k(t-k) = b_0^{(k)} + b_1^{(k)} \phi^{t-k} \quad (4.3.27)$$

wobei die  $b_i^{(k)}$  Konstante sind, die vom jeweiligen Zeitpunkt  $k$  abhängen

$$b_0^{(k)} = A_0 \sum_{j=-\infty}^k a_j \quad (4.3.28)$$

$$b_1^{(k)} = A_1 \sum_{j=-\infty}^k \phi^{k-j} a_j \quad (4.3.29)$$

#### 4.4. Der IMA(1,1)-Prozeß

In den folgenden beiden Kapiteln behandeln wir die wichtigsten integrierten Moving Average-Prozesse: Den IMA(1,1) und den IMA(2,2)-Prozeß.

##### 4.4.1. D-Form

Nach Def. 4.1.2. lautet der IMA(1,1)-Prozeß

$$(1-B)z_t = (1-\theta B)a_t \quad (4.4.1)$$

Der verallgemeinerte AR-Operator besteht nur aus dem Differenzenoperator

$$\psi(B) = 1-B = \nabla \quad (4.4.2)$$

Die D-Form des Prozesses lautet

$$z_t = z_{t-1} + a_t - \theta a_{t-1} \quad (4.4.3)$$

Aufgrund des zugehörigen invertierbaren MA(1)-Prozesses gilt die Invertibilitätsbedingung

$$|\theta| < 1 \quad (4.4.4)$$

##### 4.4.2. V-Form

Nach Def. 4.2.6. setzen wir die Form

$$\begin{aligned} (1-\theta B) &= 1-B+\lambda_0 B = \\ &= 1+(\lambda_0-1)B \end{aligned} \quad (4.4.5)$$

Wir erhalten somit

$$\lambda = \lambda_0 = 1-\theta \quad (4.4.6)$$

Die V-Form lautet

$$\nabla z_t = \nabla a_t + (1-\theta)a_{t-1} = \nabla a_t + \lambda a_{t-1} \quad (4.4.7)$$

#### 4.4.3. Die RS-Form

Durch Anwendung von S auf beiden Seiten von (4.4.7) erhalten wir die RS-Form

$$\begin{aligned}
 z_t &= \lambda S a_{t-1} + a_t \\
 &= \lambda \sum_{j=-\infty}^{t-1} a_j + a_t \\
 &= \sum_{j=0}^{\infty} \psi_j a_{t-j}
 \end{aligned} \tag{4.4.8}$$

Die  $\psi$ -Koeffizienten lauten

$$\begin{aligned}
 \psi_0 &= 1 \\
 \psi_j &= \lambda = (1-\theta) \quad j \geq 1
 \end{aligned} \tag{4.4.9}$$

#### 4.4.4. Die TRS-Form

Spalten wir die Form (4.2.8) gemäß Def. 4.2.5 auf, erhalten wir als Truncated Random Shock-Form

$$z_t = C_k(t-k) + \lambda \sum_{j=k+1}^{t-1} a_j + a_t \tag{4.4.10}$$

In diesem Fall gilt

$$b_0^{(k)} = C_k(t-k) \tag{4.4.11}$$

und ist die allgemeine Lösung der homogenen Differenzgleichung

$$(1-B)C_k(t-k) = 0 \tag{4.4.12}$$

und (4.2.11) ist

$$b_0^{(k)} = \lambda S a_k = \lambda \sum_{j=-\infty}^k a_j \tag{4.4.13}$$

Für dieses Modell ist die allgemeine Lösung einfach eine Konstante

$$C_k(t-k) = b_o^{(k)} t^0 = b_o^{(k)} \quad (4.4.14)$$

oder anders ausgedrückt: Ein Polynom mit dem Grad 0.  $b_o^{(k)}$  stellt das laufende "Niveau" des Prozesses dar und hängt von der speziellen Wahl des Ursprungs (Ausgangszeitpunktes)  $k_o = k$  ab. Wird der Ursprung von  $k_o = k$  zu  $k_o = k-1$ , so ändert sich die allgemeine Lösung  $b_o$  gemäß

$$b_o^{(k)} = b_o^{(k-1)} + \lambda a_k \quad (4.4.15)$$

Das neue Niveau setzt sich aus der alten plus einer Gewichtung des letzten Fehlers zusammen.

#### 4.4.5. Die invertierte Form

Nun drücken wir das Modell als Funktion der vergangenen  $z_t$  aus:

$$a_t = \pi(B)z_t \quad (4.4.16)$$

und formen  $\pi(B)$  um

$$\begin{aligned} \pi(B) &= \frac{1-B}{1-\theta B} \\ &= \frac{(1-\theta B) - (1-\theta)B}{1-\theta B} \\ &= 1 - (1-\theta)B \left( \sum_{j=0}^{\infty} (\theta B)^j \right) \\ &= 1 - (1-\theta)(B + \theta B^2 + \theta^2 B^3 + \dots) \end{aligned} \quad (4.4.17)$$

Damit sind die  $\pi_j$  bestimmt

$$\begin{aligned} \pi_j &= (1-\theta)\theta^{j-1} \\ &= \lambda(1-\lambda)^{j-1} \quad j=1,2,3,\dots \end{aligned} \quad (4.4.18)$$

Die normierte  $\pi$ -Form lautet

$$z_t = \bar{z}_{t-1}(\pi) + a_t \quad (4.4.19)$$

mit

$$\bar{z}_{t-1}(\pi) = \lambda \sum_{j=1}^{\infty} (1-\lambda)^{j-1} z_{t-j} \quad (4.4.20)$$

wobei sich die  $\pi_j$  wieder auf 1 summieren.

Andererseits gilt aus dem Vergleich von (4.4.19) mit der RS-Form (4.4.8) die Beziehung

$$\bar{z}_{t-1}(\pi) = \lambda S a_{t-1} \quad (4.4.21)$$

und damit folgt für die allgemeine Lösung der TRS-Form

$$b_o^{(k)} = \bar{z}_k(\pi) \quad (4.4.22)$$

#### 4.4.6. Exponentielles Glätten als IMA(1,1)-Prozeß

Der ARIMA(0,1,1)-Prozeß ist äquivalent mit exponentiellem Glätten, wobei die Gewichte der vergangenen Beobachtungen lauten

$$\lambda \quad \lambda(1-\lambda) \quad \lambda(1-\lambda)^2 \quad \lambda(1-\lambda)^3 \quad \dots$$

Diese fallen exponentiell, d.h. mit einer geometrischen Progression ab. ( $\lambda = 1-\theta$ )

Durch Umformung von (4.4.20) erhalten wir die bekannte Rekursionsformel für das exponentielle Glätten

$$\bar{z}_t(\pi) = \lambda z_t + (1-\lambda) \bar{z}_{t-1}(\pi) \quad (4.4.23)$$

Dieser Ausdruck zeigt, daß man das neue Niveau des Prozesses (zum Zeitpunkt t) durch einen gewogenen Durchschnitt des letzten Niveaus und der neuen Beobachtung erhält.

Da dieser Prozeß nicht stationär ist, besitzt er keinen Mittelwert (befindet sich nicht im "Gleichgewicht"). Als Ersatz für den Mittelwert ist das jeweilige "Niveau" des Prozesses anzusehen. Dieses Niveau ist Startpunkt eines "random-walk"-Prozesses, dessen Zuwächse  $a_t$  zwar im Mittel null sind, jedoch sonst im Verhalten nicht vorhersehbar ist.

Das "Updaten" des Niveaus kann auch durch folgende Beziehung gesehen werden

$$\bar{z}_t(\pi) = \bar{z}_{t-1}(\pi) + \lambda a_t$$

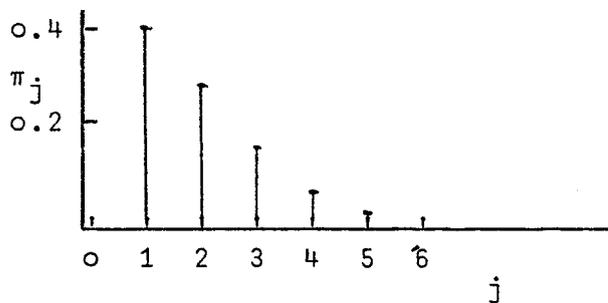
Das neue Niveau des Prozesses setzt sich aus dem alten plus einem gewissen Anteil der neuen Störung  $a_t$  zusammen.

Beispiel 4.4.1.: IMA(1,1)-Prozeß

Berechnung der  $\pi$ -Gewichte für

$$\lambda = 0.4 \quad (\theta = 0.6)$$

j	0	1	2	3	4
$\pi_j$	1	0,4	0,24	0,144	0,0864



4.4.7. Spezialfälle des IMA(1,1)-Prozesses

Der Invertibilitätsbereich für  $\lambda$  lautet

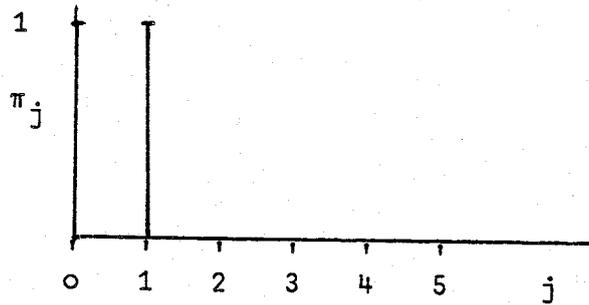
$$0 < \lambda < 2 \tag{4.4.24}$$

Dazu betrachten wir folgende drei Spezialfälle

a) Fall  $\lambda=1$  ( $\theta=0$ )

Die  $\pi$ -Koeffizienten sind

$$\begin{aligned} \pi_1 &= 1 \\ \pi_j &= 0 \quad j > 1 \end{aligned} \tag{4.4.25}$$



b) Fall  $\lambda \rightarrow 0$  ( $\theta \rightarrow 1$ )

Die  $\pi_j$ -Gewichte fallen immer langsamer ab und der Prozeß erstreckt sich immer weiter in die Vergangenheit.

Beispiel 4.4.2.: IMA(1,1)-Prozeß mit  $\lambda=0,1$  ( $\theta=0,9$ )

j	1	2	3	4	5
$\pi_j$	0.1	0.09	0.081	0.0729	0.06561

c) Fall  $\lambda=0$  ( $\theta=1$ )

In diesem Fall wird der Prozeß zu

$$(1-B)z_t = (1-B)a_t \quad (4.4.26)$$

bzw.

$$z_t = a_t + \mu \quad (4.4.27)$$

wobei  $\mu$  der Mittelwert aller vergangenen Werte ist.

4.5. Der IMA(2,2)-Prozeß

4.5.1. Die Differenzgleichungsform

Nach Def. 4.1.2 lautet der IMA(2,2)-Prozeß

$$\nabla^2 z_t = (1-\theta_1 B - \theta_2 B^2) a_t \quad (4.5.1)$$

bzw. ausgeschrieben

$$z_t = 2z_{t-1} - z_{t-2} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \quad (4.5.2)$$

Dieser Prozeß ist für die Darstellung von stochastischen Trends sehr gut geeignet.

Der zugehörige MA(2)-Prozeß erfüllt die Invertibilitätsbedingungen

$$-1 < \theta_2 < 1$$

$$\theta_2 + \theta_1 < 1 \quad (4.5.3)$$

$$\theta_2 - \theta_1 < 1$$

#### 4.5.2. Die V-Form

Nach Def. 4.2.6 setzen wir das MA(2)-Polynom an als

$$(1 - \theta_1 B - \theta_2 B^2) = (\lambda_0 + \lambda_1 V) B + V^2 \quad (4.5.4)$$

und erhalten durch Koeffizientenvergleich

$$\lambda_0 = 1 - \theta_1 - \theta_2 \quad (4.5.5)$$

$$\lambda_1 = 1 + \theta_2$$

die umgekehrte Beziehung lautet

$$\theta_1 = 2 - \lambda_0 - \lambda_1 \quad (4.5.6)$$

$$\theta_2 = \lambda_1 - 1$$

Damit lautet die V-Form

$$V^2 z_t = (\lambda_0 + \lambda_1 V) a_{t-1} + V^2 a_t \quad (4.5.7)$$

Die Invertibilitätsbedingungen für die  $\lambda_i$  lauten

$$\lambda_0 > 0$$

$$0 < \lambda_1 < 2 \tag{4.5.8}$$

$$2\lambda_1 + \lambda_0 < 4$$

da

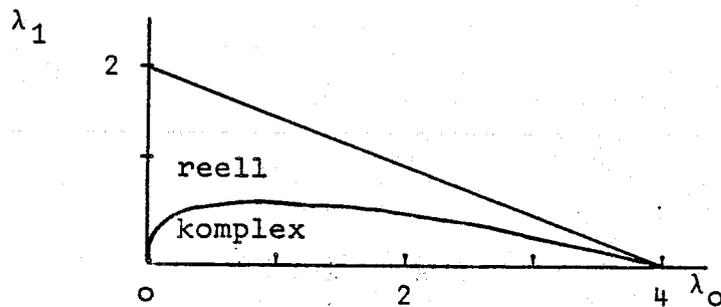
$$\theta_2 + \theta_1 = 1 - \lambda_0 < 1$$

$$-1 < \lambda_1 - 1 < 1 \tag{4.5.9}$$

$$\theta_2 - \theta_1 = 2\lambda_1 - 3 + \lambda_0 < 1$$

gilt.

Im  $R^2$  erhalten wir für  $\lambda_0$  und  $\lambda_1$  folgenden zulässigen Bereich



### Spezialfälle

Setzen wir in (4.5.7)  $\lambda_0 = 0$  und kürzen durch  $\nabla$ , so erhalten wir

$$\nabla z_t = \lambda_1 a_{t-1} + \nabla a_t \tag{4.5.10}$$

Dies ist jedoch genau die  $\nabla$ -Form (4.4.7) des IMA(1,1)-Prozesses mit  $\lambda = \lambda_1$ . Für die D-Form gibt es eine derartige einfache Beziehung zwischen diesen beiden Prozessen nicht.

#### 4.5.3. Die RS-Form

Durch zweimaliges Summieren in der  $\nabla$ -Form (4.5.7) erhalten wir

$$z_t = \lambda_1 S a_{t-1} + \lambda_0 S^2 a_{t-1} + a_t \quad (4.5.11)$$

Wegen (1.2.30) können wir dies auch schreiben

$$\begin{aligned} z_t &= \lambda_1 \sum_{j=0}^{\infty} a_{t-1-j} + \lambda_0 \sum_{j=1}^{\infty} j a_{t-j} + a_t \\ &= a_t + \sum_{j=1}^{\infty} (\lambda_1 + \lambda_0 j) a_{t-j} \end{aligned} \quad (4.5.12)$$

Die  $\psi$ -Koeffizienten lauten somit

$$\begin{aligned} \psi_0 &= 1 \\ \psi_j &= \lambda_0 + j \lambda_1 \quad j \geq 1 \end{aligned} \quad (4.5.13)$$

#### 4.5.4. Truncated Random Shock-Form

Nach Def. 4.2.5 lautet

$$z_t = C_k(t-k) + \lambda_0 \sum_{j=k+1}^{t-1} a_j + \lambda_1 \sum_{i=k+1}^{t-1} \sum_{j=k+1}^i a_j + a_t \quad (4.5.14)$$

$C_k(t-k)$  ist die Lösung der homogenen Gleichung

$$(1-B)^2 C_k(t-k) = 0 \quad (4.5.15)$$

mit

$$\begin{aligned} C_k(t-k) &= b_0^{(k)} + b_1^{(k)}(t-k) \\ &= ((\lambda_0 - \lambda_1) S a_k + \lambda_1 S^2 a_k) + (\lambda_1 S a_k)(t-k) \end{aligned} \quad (4.5.16)$$

$C_k(t-k)$  ist ein Polynom vom Grad 1 in  $(t-k)$ , d.h. der Prozeß verhält sich etwa wie eine lineare Funktion, wobei die Koeffizienten  $b_i^{(k)}$  vom jeweiligen truncation point  $k$  abhängen.

Veränderung des Truncation Punktes

Betrachtet man die Differenzen  $b_i^{(k)} - b_i^{(k-1)}$ , so erhält man

$$b_0^{(k)} = b_0^{(k-1)} + b_1^{(k-1)} + \lambda_0 a_k \quad (4.5.17)$$

$$b_1^{(k)} = b_1^{(k-1)} + \lambda_1 a_k \quad (4.5.18)$$

Der "Anstieg" des Prozesses  $b_1^{(k)}$  setzt sich aus dem alten Anstieg plus einer  $\lambda_1$ -Gewichtung der neuen Störung zusammen.  $b_0^{(k)}$  entspricht dem absoluten Glied und die neue Störung wird mit  $\lambda_0$  gewichtet. Durch die Akkumulierung von Störgliedern seit dem Zeitpunkt  $k$  ersieht man in (4.5.14), daß dadurch größere Abweichungen von der Gerade (4.5.16) vorkommen.

4.5.5. Die invertierte Form

Die normierte  $\pi$ -Form

$$\begin{aligned} z_t &= \sum_{j=1}^{\infty} \pi_j z_{t-j} + a_t \\ &= \bar{z}_{t-1}(\pi) + a_t \end{aligned} \quad (4.5.19)$$

erhalten wir durch Koeffizientenvergleich der Polynome

$$(1 - 2B + B^2) = (1 - \theta_1 B - \theta_2 B^2)(1 - \pi_1 B - \pi_2 B^2 - \dots) \quad (4.5.20)$$

$$\pi_1 = 2 - \theta_1$$

$$\pi_2 = \theta_1(2 - \theta_1) - (1 + \theta_2) \quad (4.5.21)$$

$$(1 - \theta_1 B - \theta_2 B^2)\pi_j = 0 \quad j \geq 3$$

In  $\lambda$  ausgedrückt, sind die Anfangswerte

$$\begin{aligned} \pi_1 &= \lambda_0 + \lambda_1 \\ \pi_2 &= \lambda_0 + 2\lambda_1 - (\lambda_0 + \lambda_1)^2 \end{aligned} \quad (4.5.22)$$

### Verhalten der $\pi$ -Koeffizienten

Sind die Wurzeln der charakteristischen Gleichung

$$1 - \theta_1 B - \theta_2 B^2 = 0$$

- a) reell, so sind die  $\pi$ -Gewichte eine Kombination von zwei Exponentialfolgen
- b) komplex, so verhalten sich die  $\pi$ -Gewichte wie eine gedämpfte Sinusschwingung

## 4.6. Lineare Differenzgleichung und ARIMA-Prozesse

### 4.6.1. Die allgemeine Lösung

Gegeben sei die lineare Differenzgleichung (mit konstanten Koeffizienten)

$$z_t - \varphi_1 z_{t-1} - \varphi_2 z_{t-2} - \dots - \varphi_p z_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (4.6.1)$$

bzw. in Operatorform

$$\varphi(B)z_t = \theta(B)a_t \quad (4.6.2)$$

wobei das Operatorpolynom  $\varphi(B)$  die Ordnung  $p' = p + d$  hat.  
Für die Lösungsfolgen gilt folgender

#### Satz 4.6.1.

Die allgemeine Lösung der Differenzgleichung (4.6.1) mit Anfangswerten zum Zeitpunkt  $k > t$  lautet:

$$z_t = C_k(t-k) + I_k(t-k) \quad (4.6.3)$$

$C_k(t-k)$  ist die Lösung der homogenen Gleichung  
(complementary function)

und  $I_k(t-k)$  ist eine spezielle (partikuläre) Lösung der inhomogenen Differenzgleichung

$$C_k(t-k) = C_0^{t-k} \sum_j A_j (t-k)^j + \sum_i^P D_i C_i^{t-k} \quad (4.6.4)$$

$$I_k(t-k) = \sum_{j=k+1}^t \psi_{s-j} a_j \quad (4.6.5)$$

Beweis der allgemeinen Lösung (4.6.3)

Die partikuläre Lösung  $z_t^I$  der Differenzgleichung (4.6.2) erfüllt

$$\varphi(B)z_t^I = \theta(B)a_t \quad (4.6.6)$$

Bilden wir die Differenzen von (4.6.2) und (4.6.6)

$$\varphi(B)(z_t - z_t^I) = 0 \quad (4.6.7)$$

und setzen

$$z_t^C = z_t - z_t^I$$

so erhalten wir

$$\varphi(B)z_t^C = 0 \quad (4.6.8)$$

Die gesuchte Lösungsfolge setzt sich zusammen aus

$$z_t = z_t^C + z_t^I \quad (4.6.9)$$

Denn setzen wir (4.6.9) in (4.6.2) ein, so erhalten wir

$$\varphi(B)(z_t^C + z_t^I) = \theta(B) \quad (4.6.10)$$

$$\varphi(B)z_t^C + \varphi(B)z_t^I = \theta(B) \quad (4.6.11)$$

und da (4.6.7) und (4.6.8) vorausgesetzt waren, ist (4.6.9) die allgemeine Lösung, die sich aus der allgemeinen Lösung der homogenen Gleichung zusammensetzt

$$z_t^C = C_k(t-k) \quad (4.6.12)$$

und einer speziellen Lösung der inhomogenen Gleichung

$$z_t^I = I_k(t-k) \quad (4.6.13)$$

Die Bezeichnungen  $C_k(t-k)$  und  $I_k(t-k)$  sollen darauf hindeuten, daß die Lösungen vom jeweils gewählten Zeitpunkt  $k$  (truncation point) und den damit verbundenen verschiedenen Anfangsbedingungen abhängen.

#### 4.6.2. Die Lösung der homogenen Gleichung

Für die Lösung der homogenen Gleichung  $C_k(t-k)$

$$\varphi(B)z_t^C = 0 \quad (4.6.14)$$

spielt der verallgemeinerte AR-Operator eine große Rolle

$$\varphi(B) = \phi(B)(1-B)^d \quad (4.6.15)$$

Zur Lösung der Differenzgleichung benötigen wir die Nullstellen der charakteristischen Gleichung

$$\varphi(B) = (1-C_1B) \dots (1-C_pB) = 0 \quad (4.6.16)$$

Je nach Art der Wurzeln unterscheiden wir zwei Fälle der Lösungen:

Fall 1: Die Wurzeln  $C_1 \dots C_p$ , sind alle verschieden, d.h.  $\varphi(B)$  kann höchstens die Form

$$\varphi(B) = \phi(B)(1-B) \quad (4.6.17)$$

annehmen, und alle Wurzeln von  $\phi(B)$  sind verschieden.

Satz 4.6.2.

Die allgemeine homogene Lösung zu (4.6.14) zum Startwert  $t_0=k$  und den daraus bestimmbar Konstanten  $A_i$  lautet

$$z_t^C = A_1 C_1^{t-k} + A_2 C_2^{t-k} + \dots + A_{p'} C_{p'}^{t-k} \quad (4.6.18)$$

wobei die  $C_i^{-1}$  die Wurzeln der charakteristischen Gleichung  $\varphi(B)=0$  sind.

Beweis:

Wir setzen (4.6.18) in (4.6.14) ein

$$\varphi(B) \left( \sum_{i=1}^{p'} A_i C_i^{t-k} \right) = \sum_{i=1}^{p'} \varphi(B) A_i C_i^{t-k} = 0 \quad (4.6.19)$$

und betrachten ein einzelnes Glied von (4.6.19), und da B auf t wirkt

$$B^{p'} C^t = B^{t-p} \quad (4.6.20)$$

gilt

$$\begin{aligned} \varphi(B) A_i C_i^{t-k} &= (1 - \varphi_1 B - \dots - \varphi_{p'} B^{p'}) A_i C_i^{t-k} \\ &= A_i C_i^{t-k-p'} (C_i^{p'} - \varphi_1 C_i^{p'-1} - \dots - \varphi_{p'}) \end{aligned} \quad (4.6.21)$$

(4.6.19) ist dann null, wenn jeder einzelne Summand (4.6.21) null ist. Das ist aber der Fall, wenn

$$C_i^{p'} - \varphi_1 C_i^{p'-1} - \dots - \varphi_{p'} = 0 \quad (4.6.22)$$

Das ist aber gerade die charakteristische Gleichung in  $C_i$ , d.h. wenn  $C_i$  eine Wurzel ist, dann ist (4.6.19) erfüllt.

Beispiel 4.6.1.

Wir zeigen am Beispiel des AR(2)-Prozesses einen konstruktiven Weg, den Lösungsansatz (4.6.18) zu erhalten. Die homogene Lösung des AR(2)-Prozesses lautet

$$(1 - \phi_1 B - \phi_2 B^2) z_t = 0 \quad (4.6.23)$$

In die Wurzeln der charakteristischen Gleichung faktorisiert

$$(1 - C_1 B)(1 - C_2 B) z_t = 0 \quad (4.6.24)$$

Wir lösen die Differenzgleichung 2. Ordnung durch zweimalige Reduktion auf Differenzgleichung 1. Ordnung:

a) reduzierte Lösung 1. Ordnung

Setzen wir vorübergehend

$$y_t = (1 - C_2 B) z_t \quad (4.6.25)$$

und setzen in (4.6.24) ein, so erhalten wir

$$(1 - C_1 B) y_t = 0 \quad (4.6.26)$$

bzw.

$$y_t = C_1 y_{t-1} \quad (4.6.27)$$

Wir haben die Differenzgleichung 2. Ordnung auf eine der 1. Ordnung zurückgeführt und lösen diese

$$\begin{aligned} y_t &= C_1 y_{t-1} = C_1^2 y_{t-2} = \dots = C_1^k y_{t-k} = \\ &= C_1^{t-k} y_k \end{aligned} \quad (4.6.28)$$

wobei  $y_k$  der Startwert der Lösungsfolge ist, d.h. einer Konstanten nach Formel (4.6.18) entspricht.

$$D_1 = y_k \quad (4.6.29)$$



Betrachten wir wieder den Fall eines AR(2)-Prozesses mit zwei gleichen Wurzeln, dann können wir  $C_0$  in (4.6.31) einsetzen und erhalten

$$z_t = C_0^{t-k} z_k + D_1 C_0^{t-k} (t-k) \quad (4.6.35)$$

bzw.

$$z_t = (A_0 + A_1(t-k)) C_0^{t-k} \quad (4.6.36)$$

Allgemein kann man nun zeigen, daß wenn  $d$  gleiche Wurzeln  $C_0$  auftreten, die Lösung lautet:

$$z_t = A_0 + A_1(t-k) + \dots + A_{d-1}(t-k)^{d-1} C_0^{t-k} \quad (4.6.37)$$

Dies ist ein Polynom vom Grad  $d-1$  in  $(t-k)$ .

Werden die gleichen Wurzeln nur durch den Differenzenoperator bestimmt, so ist  $C_0=1$ , und die Lösung lautet

$$z_t = A_0 + A_1(t-k) + \dots + A_{d-1}(t-k)^{d-1} \quad (4.6.38)$$

### Fall 3: Komplexe Wurzeln

Für die Lösungsdarstellung bei komplexen Wurzeln diskutieren wir o.B.d.A. den AR(2)-Prozeß.

Komplexe Wurzeln eines AR(2)-Prozesses treten dann auf, wenn die Wurzeln der charakteristischen Gleichung<sup>+) )</sup>

$$(1 - \phi_1 B - \phi_2 B^2) = (1 - C_1 B)(1 - C_2 B) \quad (4.6.39)$$

eine negative Diskriminante

$$\frac{\phi_1^2}{4} + \phi_2 < 0 \quad -1 \leq \phi_2 \leq 0 \quad (4.6.40)$$

aufweisen. In diesem Fall sind die Wurzeln konjugiert komplex.

---

+) Die in der Theorie der Differenzgleichung übliche Form ist der "positive" Ansatz

$$y_t + a y_{t-1} + b y_{t-2} = 0$$

Daher unterscheiden sich die Formeln für die Lösungen in den Vorzeichen.

$$\begin{aligned} C_1 &= u + iv \\ C_2 &= u - iv \end{aligned} \quad (4.6.41)$$

wobei  $i = \sqrt{-1}$  die imaginäre Einheit ist und

$$u = \frac{\phi_1}{2} \quad (4.6.42)$$

$$v = \frac{1}{2} \sqrt{-\phi_2^2 - 4\phi_1} \quad (4.6.43)$$

Die konjugiert komplexen Wurzeln werden einfacher in Polarkoordinaten geschrieben

$${}_1C_2 = r e^{\pm i\omega} = r(\cos\omega \pm i\sin\omega) \quad (4.6.44)$$

mit

$$r = \sqrt{u^2 + v^2} \quad (4.6.45)$$

$$\operatorname{tg}\omega = \frac{u}{v} \quad (4.6.46)$$

Das Verhältnis von Realteil und Imaginärteil der konjugiert komplexen Wurzeln bestimmt den Winkel  $\omega$  und damit die Periodenlänge der Lösung.

Wegen (4.6.45) kann man  $r$  als Funktion der Parameter ausdrücken

$$r = \sqrt{\frac{\phi_1^2}{4} - \frac{\phi_2^2}{4} - \phi_2} = \sqrt{-\phi_2} \quad (4.6.47)$$

Zur Darstellung des Winkels als Funktion der Parameter ist die Formel über  $\cos\omega$  die einfachste

$$\cos\omega = \frac{u}{r} = \frac{\phi_1}{2\sqrt{-\phi_2}} \quad 0 \leq \omega \leq \pi \quad \phi_2 < 0 \quad (4.6.48)$$

Die Wurzeln  $C_1$  und  $C_2$  sind als konjugiert komplexe Wurzeln verschieden und bilden daher zwei linear unabhängige Fundamentallösungen. Die homogene Lösung lautet

$$Y_t^c = A_1 r^t (\cos\omega + i\sin\omega)^t + A_2 r^t (\cos\omega - i\sin\omega)^t \quad (4.6.49)$$

Nach dem Satz von De Moivre ist

$$(\cos\omega \pm i\sin\omega)^t = \cos t\omega \pm i\sin t\omega \quad (4.6.50)$$

Die Formel (4.6.49) kann man durch folgenden Satz für konjugiert komplexe Variable auf reelle Form vereinfachen:

Satz 4.6.3.: Produkte konjugierter komplexer Zahlen

Gegeben seien die konjugiert komplexen Variablen

$$\begin{aligned} z &= a + ib & \bar{z} &= a - ib \\ y &= c + id & \bar{y} &= c - id \end{aligned} \quad (4.6.51)$$

dann gilt

$$\begin{aligned} \text{a)} \quad zy + \bar{z}\bar{y} &= 2(ac - bd) = 2r_1 r_2 \cos(\alpha - \beta) \\ \text{b)} \quad zy - \bar{z}\bar{y} &= 2i(ad + bc) = 2ir_1 r_2 \sin(\alpha + \beta) \\ \text{c)} \quad z\bar{y} + y\bar{z} &= 2(ac + bd) = 2r_1 r_2 \cos(\alpha + \beta) \\ \text{d)} \quad z\bar{y} - y\bar{z} &= 2i(bc - ad) = 2ir_1 r_2 \sin(\alpha - \beta) \end{aligned} \quad (4.6.52)$$

Beweis:

$$\begin{aligned} \text{a)} \quad zy + \bar{z}\bar{y} &= (a+ib)(c+id) + (a-ib)(c-id) \\ &= ac - bd + i(bc+ad) + ac - bd - i(ad+bc) \\ &= 2(ac - bd) \end{aligned} \quad (4.6.53)$$

$$\begin{aligned} \text{c)} \quad z\bar{y} + y\bar{z} &= ac + bd + i(bc - ad) + ac + bd - i(bc - ad) \\ &= 2(ac + bd) \end{aligned} \quad (4.6.54)$$

$$\text{Ist } z = r_1 e^{-i\omega_1} \quad \text{und} \quad y = r_2 e^{-i\omega_2}$$

so ist

$$\begin{aligned} zy + \bar{z}\bar{y} &= 2r_1 r_2 (\cos\omega_1 \cos\omega_2 - \sin\omega_1 \sin\omega_2) \\ &= 2r_1 r_2 \cos(\alpha - \beta) \end{aligned} \quad (4.6.55)$$

$$zy - \bar{z}\bar{y} = 2r_1 r_2 i \sin(\alpha + \beta) \quad (4.6.56)$$

Setzen wir die konstanten  $A_i$  in (4.6.49) als konjugiert komplex an

$$A_1 = G(\cos F + i \sin F) = y \quad (4.6.57)$$

$$A_2 = G(\cos F - i \sin F) = \bar{y}$$

und kombinieren diese nach (4.6.51)

$$z = r^t (\cos \omega t + i \sin \omega t) \quad (4.6.58)$$

so ist

$$zy + \bar{z}\bar{y} = 2Gr^t \cos(\omega t - F) \quad (4.6.59)$$

Den Winkel  $\omega$  können wir in der Form

$$\omega = 2\pi f = \frac{2\pi}{T} \quad (4.6.60)$$

schreiben und erhalten so die Periodenlänge  $T$  der gedämpften Sinusschwingung mit dem Dämpfungsfaktor  $r^t$ .

Umgekehrt erhalten wir die Konstanten  $2G$  und  $F$  aus (4.6.57) analog (4.6.46) durch

$$\operatorname{tg} F = \frac{A_1 - A_2}{A_1 + A_2} \quad (4.6.61)$$

da

$$A_1 + A_2 = 2G \cos F \quad (4.6.62)$$

$$A_1 - A_2 = 2G i \sin F$$

gilt und wegen  $|z| = z\bar{z}$  erhält man analog zu (4.6.45)

$$G = \sqrt{A_1 A_2} \quad (4.6.63)$$

### Die allgemeine homogene Lösung

Fassen wir die vorhergehenden Fälle zusammen, so erhalten wir für die allgemeine homogene Lösung

$$C_k(t-k) = C_0^{t-k} \sum_{j=0}^{d-1} A_j (t-k)^j + \sum_{i=1}^p D_i C_i^{t-k} \quad (4.6.64)$$

Wie man aus (4.6.64) erkennt, setzt sich die allgemeine homogene Lösung aus einer Mischung von gedämpften exponentiellen Gliedern  $C_i^{t-k}$  (für alle  $C_i$ ) und gedämpften Sinusschwingungen der Form  $d^{t-k} \sin(2\pi fkt + F)$  (für jedes Paar konjugiert komplexer Lösungen) und Polynomanteilen der Form  $(t-k)^j$  der Ordnung  $d-1$  (für  $d$  gleiche Wurzeln) zusammen.



da die RS-Form  $z_t = \psi(B)a_t$  gilt. Wir schreiben nun das Produkt der beiden Polynome der Identität (4.6.70) (linke Seite) ausführlich an

$$(1 - \psi_1 B - \psi_2 B^2 - \dots - \psi_{p'} B^{p'}) \psi(B) a_t =$$

$$= \psi(B) a_t - \psi_1 \psi(B) a_{t-1} - \dots - \psi_{p'} \psi(B) a_{t-p'} \quad (4.6.71)$$

oder

$$\begin{pmatrix} a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \\ -\psi_1 (a_{t-1} + \psi_1 a_{t-2} + \dots) \\ \vdots \\ -\psi_{p'} (\dots a_{t-p'}) \end{pmatrix} \begin{pmatrix} + \psi_{t-k-1} a_{k+1} + \sum_{j=k}^{\infty} \psi_{t-j} a_j \\ + \psi_{t-k-2} a_{k+1} + \sum_{j=k}^{\infty} \psi_{t-i-j} a_j \\ \dots \\ + \psi_{t-k-p'-1} a_{k+1} + \sum_{j=k}^{\infty} \psi_{t-p'-j} a_j \end{pmatrix} \quad (4.6.72)$$

Daran können wir erkennen, daß gerade der endliche Teil des obigen Summentableaus mit Formel (4.6.69) identisch ist. Die ersten  $q+1$  Spalten des Summentableaus summieren sich auf

$$a_t, -\theta_1 a_{t-1}, \dots, -\theta_q a_{t-q} \quad (4.6.73)$$

da die  $\psi_j$ -Koeffizienten gerade so definiert waren (vgl. (4.2.44)). Damit ist gezeigt, daß (4.6.67) eine partikuläre Lösung der Differenzgleichung ist.

## 5. P R O G N O S E

### 5.1. Theorie der Prognose

Zur Spezifikation des Prognoseproblems für ARIMA-Prozesse benötigen wir einige Resultate aus der Prognosetheorie. Statistische Modelle von Zeitreihen dienen einerseits der Analyse und Beschreibung, andererseits dienen sie als Prognosemodell. Zur Ermittlung der Prognosefunktion haben wir folgende Punkte zu berücksichtigen:

- 1) Zeitraum und Zeitpunkt der Prognose
- 2) Wahl eines Prognosekriteriums
- 3) Wünschenswerte Eigenschaften der Prognosefunktion.

#### 5.1.1. Zeitraum und Zeitpunkt der Prognose

Beobachtet werde eine Zeitreihe  $x_t$ ,  $t=1, \dots, n$  als Realisation eines stochastischen Prozesses, und gesucht sind Prognosen zum Zeitpunkt  $t_0=t$ . Zumeist ist  $t=n$ , jedoch verwendet man Prognosezeitpunkte  $t < n$ , um das Prognoseverhalten mit aktuellen Werten zu vergleichen (Prognosechecking). Der Zeitraum  $L$  einer Prognose gibt an, wieviele Werte der Zeitreihe vom Prognosezeitpunkt  $t$  vorhergesagt werden sollen. Nach der Länge des Zeitraumes  $L$  unterscheiden wir einstufige und mehrstufige Prognosen.

#### Def. 5.1.1.: Prognosefunktion

Unter der Prognosefunktion verstehen wir die Prognosen  $\hat{z}_{t+l}$  eines zukünftigen Wertes der Zeitreihe  $z_{t+l}$  als Funktion von  $l=1, 2, \dots, L$  für einen bestimmten Zeitraum  $L$ . Wir nennen  $l$  die "leadzeit" und  $\hat{z}_{t+l}$  die Prognose zur leadzeit  $l$ .

Für  $l=1$  erhalten wir die einstufigen Prognosen.

### 5.1.2. Prognosekriterium

Prognosekriterien verwendet man zur Bestimmung der Güte und Optimalität von Prognosefunktionen. Das am häufigsten verwendete Kriterium ist das des erwarteten mittleren quadratischen Fehlers bzw. mean square errors (MSE-Kriterium). Granger (1969) und Granger/Newbold (1977) diskutierten einige alternative Kriterien. Alle Prognosekriterien schließen an eine Bewertung des Prognosefehlers an.

Def. 5.1.2.: Prognosefehler

Unter dem Prognosefehler  $e_{t+1}$  einer Prognosefunktion verstehen wir

$$e_{t+1} = \hat{z}_{t+1} - z_{t+1} \quad (5.1.1)$$

Ist  $z_t$  ein stochastischer Prozeß, so ist der Prognosefehler  $e_{t+1}$  ebenfalls eine Zufallsvariable.

Das MSE-Kriterium unterstellt eine quadratische Verlustfunktion

$$E e_{t+1}^2 = \text{Min!} \quad (5.1.2)$$

### 5.1.3. Eigenschaften der Prognosefunktion

In der Theorie der Prognosen werden einige Eigenschaften für Prognosefunktionen gefordert (vgl. Granger/Newbold 1977, S.111ff.). Für die Prognose des ARIMA-Prozesses verwenden wir folgende drei Charakteristika:

#### a) Informationsmenge und Autogenität

Jede Prognose basiert auf einer bestimmten Anzahl von Voraussetzungen, Theorien, Daten und sonstigem Wissen, die wir kurz als Informationsmenge bezeichnen wollen.

Def. 5.1.3.: Informationsmenge  $I_t$

Unter der Informationsmenge  $I_t$  verstehen wir die Informationsmenge zum Zeitpunkt  $t$ , das sind alle Eigenschaften, Daten, usw. die bis zum Zeitpunkt  $t$  erhältlich sind. Unter der autogenen Informationsmenge  $I_t$  verstehen wir die gesamte Zeitreihe  $z_t$ ,  $t=1, \dots, n$ , als Stichprobe eines stochastischen Prozesses, mit bekannten Eigenschaften.

Eine andere Informationsmenge könnte z.B. aus mehreren Zeitreihen bestehen, oder nur die letzten Werte einer Zeitreihe verwenden. Für Vergleiche verschiedener Prognoseprozeduren ist es deshalb wichtig, die verwendeten Informationsmengen genau zu spezifizieren.

Für die Prognose von ARIMA(p,d,q)-Prozessen verwenden wir eine autogene Informationsmenge  $I_t$  mit den geschätzten Parametern des Modells bis zum Zeitpunkt  $t$ .

b) Zeitinvarianz

Eine Prognosefunktion heißt zeitinvariant, falls sie nur eine Funktion der leadzeit  $l$  ist, und nicht von den Prognosezeitpunkten abhängt.

Ist  $I_t$  die verwendete Informationsmenge, so ist

$$\hat{z}_{t+1} = f(l; I_t) \quad (5.1.3)$$

und nicht

$$\hat{z}_{t+1} = f_t(l; I_t)$$

Eine einstufige Prognose zum Zeitpunkt  $t$  hat dieselbe (funktionale) Form wie die einstufige Prognose  $k$  Zeitpunkte früher oder später.

c) Linearität

Das Problem der Wahl der besten Prognosefunktion (aufgrund eines Prognosekriteriums) kann man durch Annahme einer linearen Prognosefunktion ziemlich erleichtern. Das bedeutet, daß die Prognosen nur eine lineare Funktion der Daten in der Informationsmenge  $I_t$  sind. Unter dieser Einschätzung betrachten wir nur lineare Prognosen. Die Vorteile von linearen Prognosefunktionen sind so groß, daß nur in Ausnahmefällen nichtlineare Prognosen verwendet werden. So werden durch ökonometrische Modelle oft bestimmte nichtlineare funktionale Beziehungen vorgegeben (vgl. Granger/Newbold (1977)).

Eine autogene, zeitinvariante, lineare Prognosefunktion hat die Form

$$\hat{z}_{t+1} = \sum_{j=1}^n \alpha_j(1) z_{t-j} \quad (5.1.4)$$

Die theoretische Annahme des Modells sind in den Koeffizienten  $\alpha_j$  enthalten.

Für die Prognose von ARIMA-Prozessen bezeichnen wir die Prognosefunktion mit

$$\hat{z}_t(1) = \hat{z}_{t+1} \quad (5.1.5)$$

Da die  $z_t$  eines ARIMA-Prozesses durch die Transferfunktion  $\psi(B)$  eindeutig als Funktion der  $a_t$  dargestellt werden kann, können wir die Prognosefunktion  $\hat{z}_t(1)$  als lineare Funktion der vergangenen Störungen ansetzen

$$\hat{z}_t(1) = \sum_{j=1}^n \beta_j(1) a_{t-j} \quad (5.1.6)$$

Im nächsten Kapitel bestimmen wir die  $\beta_j(1)$  nach dem MSE-Kriterium für die ARIMA-Prozesse und in Kapitel 5.3.3. die  $\alpha_j(1)$  nach Formel (5.1.4), die mit der  $\pi$ -Form von ARIMA Prozessen berechnet werden kann.

#### 5.1.4. MSE-Prognosen

Im folgenden Abschnitt wollen wir eine lineare, zeitinvariante Prognosefunktion für die ARIMA-Prozesse ableiten, wobei wir das MSE-Kriterium zugrundelegen.

Gegeben sei der allgemeine ARIMA(p,d,q)-Prozeß

$$\mathcal{Y}(B)z_t = \theta(B)a_t \quad (5.1.7)$$

Für die Prognosefunktion als Funktion der  $a_t$  benötigen wir das Modell in der RS-Form (vgl. 3.1.1.)

$$\begin{aligned} z_t &= \psi(B)a_t \\ &= \sum_{j=0}^{\infty} \psi_j a_{t-j} \\ &= \sum_{j=-\infty}^t \psi_{t-j} a_j \end{aligned} \quad (5.1.8)$$

#### Satz 5.1.1.: MSE-Prognosekoeffizienten

Die Koeffizienten der linearen zeitinvarianten Prognosefunktion

$$\hat{z}_t(l) = \sum_{j=1}^n \beta_j(l) a_{t-j} \quad (5.1.9)$$

für den ARIMA(p,d,q)-Prozeß lauten

$$\beta_j(l) = \psi_{l+j} \quad l > 0, j \geq 0 \quad (5.1.10)$$

#### Beweis:

Vom Prognosezeitpunkt  $t$  aus gesehen, stellen wir den künftigen Wert  $z_{t+1}$  in der RS-Form dar

$$\begin{aligned} z_{t+1} &= \sum_{j=0}^{\infty} \psi_j a_{t+1-j} = \\ &= \sum_{j=-\infty}^{t+1} \psi_{t+1-j} a_j \end{aligned} \quad (5.1.11)$$

Das MSE-Kriterium lautet

$$E(e_t(l)^2) = \text{Min!}$$

$$\begin{aligned} E e_t(l)^2 &= E(z_{t+1} - z_t(l))^2 \\ &= E\left(\sum_{j=-\infty}^{t+1} \psi_{t+1-j} a_t - \sum_{j=-\infty}^t \beta_{t-j} a_t\right)^2 \\ &= E\left(\sum_{j=t+1}^{t+1} \psi_{t+1-j} a_j - \sum_{j=-\infty}^t (\psi_{t+1-j} - \beta_{t-j}) a_t\right)^2 = \\ &= \sum_{j=t+1}^{t+1} \psi_{t+1-j}^2 \sigma_a^2 - \sum_{j=-\infty}^t (\psi_{t+1-j} - \beta_{t-j})^2 \sigma_a^2 \end{aligned} \quad (5.1.12)$$

da  $E(a_i a_j) = 0$  für  $i \neq j$  ist.

Schreiben wir (5.1.12) in Rückwärtssummation, so erhalten wir

$$E e_t(l)^2 = (1 + \psi_1^2 + \dots + \psi_{l-1}^2) \sigma_a^2 + \sigma_a^2 \sum_{j=0}^{\infty} (\psi_{1+j} - \beta_j)^2 \quad (5.1.13)$$

Diese Summe erzielt ihr Minimum, falls der zweite Faktor Null wird, d.h.

$$\beta_j = \psi_{1+j} \quad j=0, 1, 2, \dots \quad (5.1.14)$$

Bezeichnen wir mit

$$\psi_1(B) = \psi_1 + \psi_{1+1} B + \psi_{1+2} B^2 + \dots \quad l=1, 2, \dots \quad (5.1.15)$$

so lautet die MSE-Prognose in RS-Form einfach

$$z_t(l) = \psi_1(B) a_t \quad l=1, 2, \dots \quad (5.1.16)$$

Ausgeschrieben lautet die Prognosefunktion

$$z_t(1) = \sum_{j=0}^{\infty} \psi_{j+1} a_{t-j} = \psi_1 a_t + \psi_2 a_{t-1} + \psi_3 a_{t-2} + \dots$$

$$z_t(2) = \sum_{j=0}^{\infty} \psi_{j+2} a_{t-j} = \psi_2 a_t + \psi_3 a_{t-1} + \psi_4 a_{t-2} + \dots \quad (5.1.17)$$

$$\vdots$$

$$z_t(l) = \sum_{j=0}^{\infty} \psi_{j+l} a_{t-j} = \psi_l a_t + \psi_{l+1} a_{t-1} + \dots$$

Man beachte, daß sich mit größer werdenden leadzeiten die Indices der  $\psi_j$ -Koeffizienten immer weiter nach hinten verschieben.

Man erhält die Prognosen für die leadtime  $l$  dadurch, daß man die vergangenen Abweichungen mit Gewichten beginnend mit  $\psi_1$  bewertet und aufsummiert. Als Orientierung für die verschiedenen Indices in den Vorwärts- und Rückwärtssummationen diene folgende Tabelle

l	0	1	. . .	l-1	l	l+1
t	t+1	t+1-1		t+1	t	t-1

Der Prognosefehler im Falle der MSE-Prognose lautet

$$\begin{aligned}
 e_t(1) &= \sum_{j=t+1}^{t+1} \psi_{t+1-j} a_j \\
 &= \sum_{j=0}^{1-1} \psi_j a_{t+1-j}
 \end{aligned}
 \tag{5.1.18}$$

Damit kann ein zukünftiger Wert  $z_{t+1}$  folgendermaßen einfach zerlegt werden

$$\begin{aligned}
 z_{t+1} &= \sum_{j=0}^{1-1} \psi_j a_{t+1-j} + \sum_{j=0}^{\infty} \psi_{j+1} a_{t-j} \\
 &= e_t(1) + \hat{z}_t(1)
 \end{aligned}
 \tag{5.1.19}$$

Die RS-Form des zukünftigen Wertes  $z_{t+1}$  setzt sich aus der zum Zeitpunkt  $t$  "bekannten" Summe  $\hat{z}_t(1)$  und der "unbekannten" Summe  $e_t(1)$  zusammen.

Weiters gilt für MSE-Prognosen, daß jede Linearkombination von MSE-Prognosen wieder eine MSE-Prognose ist.

Ist  $\hat{z}_t(l)$  eine MSE-Prognose von  $z_{t+1}$ ,  $l=1, \dots, L$ , so ist

$$\sum_{l=1}^L \hat{z}_t(l) \text{ eine MSE-Prognose von } \sum_{l=1}^L z_{t+1}$$

Dies ist für die Prognose von saisonalen Zeitreihen von Bedeutung, wenn neben den saisonalen Prognosen auch aggregierte (z.B. jährliche) Prognosen notwendig sind.

#### 5.1.5. Prognosen als bedingte Erwartungswerte

Um zukünftige (noch unbekannte) Störungen und bereits bekannte Störungen (realisierte Zufallsvariable) in einer Funktion zu kombinieren, bedient man sich des Konzeptes des bedingten Erwartungswertes.

Def. 5.1.4.: Bedingter Erwartungswert der Störungen

Sei  $\{a_t\}$  ein "white-noise"-Prozeß, so verstehen wir unter dem bedingten Erwartungswert zum Zeitpunkt  $t$

$$E_t[a_j] = \begin{cases} a_j & j \leq t \\ 0 & j > t \end{cases} \quad (5.1.20)$$

Vom Zeitpunkt  $t$  aus betrachtet, sind die erwarteten Werte von  $a_t$ , die noch nicht realisiert sind, Null, und die bedingte Erwartung der bereits vergangenen Zufallsgrößen sind gleich ihrer realisierten Werte.

Satz 5.1.2.

a) Der bedingte Erwartungswert des Prognosefehlers ist Null

$$E_t[e_t(1)] = 0 \quad (5.1.21)$$

b) Der bedingte Erwartungswert einer zukünftigen Beobachtung  $z_{t+1}$  ist gleich der Prognosefunktion zum Zeitpunkt  $t$

$$E_t[z_{t+1}] = \hat{z}_t(1) \quad (5.1.22)$$

Beweis:

a) Wir berechnen den bedingten Erwartungswert des Prognosefehlers (5.1.18)

$$\begin{aligned} E_t[e_t(1)] &= E_t \sum_{j=0}^{l-1} \psi_j a_{t+1-j} = \\ &= E_t[a_{t+1}] + \psi_1 E_t[a_{t+1-1}] + \dots + \psi_{l-1} E_t[a_{t+1}] = \\ &= 0 \end{aligned}$$

$$\begin{aligned} \text{b)} \quad E_t[z_{t+1}] &= E_t[e_t(1)] + E_t[\hat{z}_t(1)] = \\ &= \sum_{j=0}^{\infty} \psi_{j+1} E_t[a_{t-j}] \\ &= \hat{z}_t(1) \end{aligned}$$

Zur Vereinfachung der Schreibweise für bedingte Erwartungswerte verwenden Box-Jenkins (1976) folgende

Def. 5.1.5.: Abkürzungen

Für die bedingten Erwartungswerte der  $a_t$  und  $z_t$  verwenden wir folgende Abkürzungen

$$[a_{t+1}] = E_t[a_{t+1}] \quad (5.1.23)$$

$$[z_{t+1}] = E_t[z_{t+1}] = E[z_{t+1} | z_t, z_{t-1}, \dots] \quad (5.1.24)$$

Satz 5.1.3.: Bedingte D-Form

Die bedingten Erwartungswerte genügen der D-Form des ARIMA-Prozesses

$$\psi(B)[z_{t+1}] = \theta(B)[a_{t+1}] \quad (5.1.25)$$

Beweis:

$$\begin{aligned} E_t[\psi(B)z_{t+1}] &= E_t\left[\sum_{j=0}^{p+d} \psi_j z_{t+1-j}\right] \\ &= \sum_{j=0}^{p+d} \psi_j E_t[z_{t+1-j}] \\ &= \psi(B)[z_{t+1-j}] \end{aligned} \quad (5.1.26)$$

Ebenso ist

$$\begin{aligned} E_t[\theta(B)a_{t+1}] &= E_t\left[\sum_{j=0}^q \theta_j a_{t+1-j}\right] \\ &= \sum_{j=0}^q \theta_j E_t[a_{t+1-j}] \\ &= \theta(B)[a_{t+1-j}] \end{aligned} \quad (5.1.27)$$

Damit gilt für die Prognosefunktion

$$\psi(B)\hat{z}_t(1) = \theta(B)[a_{t+1}] \quad (5.1.28)$$

wobei B auf 1 wirkt und für  $l > q$  wird (5.1.28) zu

$$\psi(B)\hat{z}_t(1) = 0$$

Damit verbinden wir eine erweiterte Definition von  $\hat{z}_t(1)$ .  
Wegen  $\hat{z}_t(1) = E_t[z_{t+1}]$  gilt für negative 1

$$\hat{z}_t(-1) = z_{t-1} \quad (5.1.29)$$

## 5.2. Varianz und Korrelation von Prognosen

Im folgenden Abschnitt soll der Prognosefehler  $e_t(1)$  der MSE-Prognose auf Unabhängigkeit, Varianz und Korrelation überprüft werden.

### 5.2.1. Die Varianz der Prognosen

Nach Satz 5.1.2 ist die Prognose unverzerrt, d.h. der bedingte Erwartungswert des Prognosefehlers ist Null

$$E_t e_t(1) = E_t(z_{t+1} - \hat{z}_t(1)) = 0 \quad (5.2.1)$$

#### Satz 5.2.1.: Varianz des Prognosefehlers

Die Varianz des Prognosefehlers lautet

$$\text{Var}(e_t(1)) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 \quad (5.2.2)$$

Beweis:

$$\begin{aligned} \text{Var}(e_t(1)) &= \text{Var}\left(\sum_{j=0}^{l-1} \psi_j a_{t+1-j}\right) = \\ &= \sum_{j=0}^{l-1} \psi_j^2 \text{Var}(a_{t+1-j}) = \\ &= \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 \end{aligned} \quad (5.2.3)$$

Wegen (5.1.17) ist die Varianz des Prognosefehlers gleich der Varianz der Prognose. Das führt zu folgender

Def. 5.2.1.: Varianzfunktion

Unter der Varianzfunktion  $V(l)$  verstehen wir die Varianzen der Prognosefunktion als Funktion von  $l$

$$V(l) = \text{Var}(e_t(l)^2) \quad (5.2.4)$$

Die Varianzfunktion ist eine monotone Funktion von  $l$ .  
Es gilt

$$\text{Var } e_t(l+1) > \text{Var } e_t(l) \quad (5.2.5)$$

da die Varianz mit dem Quadrat der  $\psi_j$ -Koeffizienten wächst. Weiters ist im einzelnen die Frage zu überprüfen, ob die Varianzfunktion gegen einen endlichen Wert konvergiert für  $l \rightarrow \infty$ , oder ob sie "explodiert".

5.2.2. Konfidenzintervalle für Prognosen

Zur Berechnung von Konfidenzintervallen für die Prognose benötigen wir eine Verteilungsannahme. Unter der Annahme eines normalverteilten Störprozesses  $\{a_t\}$  sind die bedingten Verteilungen für  $z_t$  ebenfalls normalverteilt

$$p(z_{t+1} | z_t, z_{t-1} \dots) \sim N(\hat{z}_t(l), V(l)) \quad (5.2.6)$$

Da die Varianz der Störungen  $\sigma_a^2$  nicht bekannt ist, gelangt man zu einer Schätzung der Varianzfunktion (5.2.4) durch eine Schätzung von  $\sigma_a^2$  (die geschätzten Residuen sind aus der Schätzphase bekannt).

$$\hat{V}(l) = \hat{\sigma}_a^2 \sum_{j=0}^{l-1} \psi_j^2 \quad (5.2.7)$$

Bezeichnen wir mit  $z_{t+1}(+)$  die obere und mit  $z_{t+1}(-)$  die untere Grenze des Konfidenzintervalles für  $\hat{z}_t(l)$  zur Irrtumswahrscheinlichkeit  $\alpha$ , so erhalten wir

$$\begin{aligned} z_{t+1}(\pm) &= \hat{z}_t(l) \pm t \left( \frac{1+\alpha}{2} \right) \sqrt{\hat{V}(l)} \\ &= \hat{z}_t(l) \pm t \left( \frac{1+\alpha}{2} \right) \hat{\sigma}_a \sqrt{\sum_{j=0}^{l-1} \psi_j^2} \end{aligned} \quad (5.2.8)$$

wobei  $\hat{\sigma}_a = \sqrt{\hat{\sigma}_a^2}$  die geschätzte Standardabweichung und  $t(\alpha)$  die  $\alpha\%$ -Grenzen der Normalverteilung  $P(X < t(\alpha)) = \alpha$  mit  $X \sim N(0,1)$  bedeuten. Das Konfidenzintervall lautet

$$P(z_{t+1}(-) < \hat{z}_t(1) < z_{t+1}(+)) = 1-\alpha \quad (5.2.9)$$

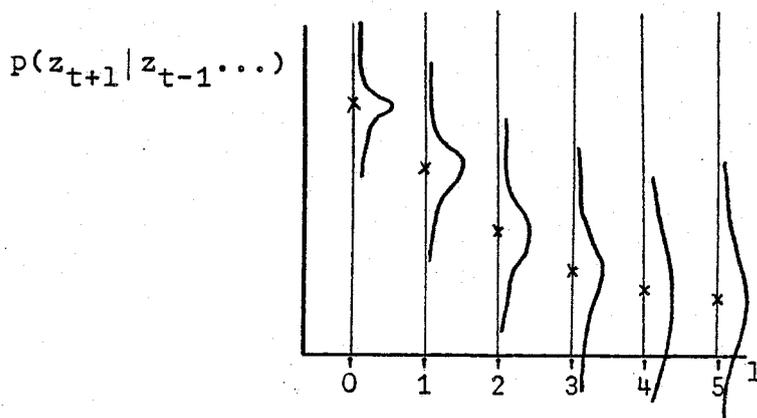
Man beachte, daß die berechneten Konfidenzintervalle nur für die einzelnen Prognosewerte  $\hat{z}_t(1)$  gelten und nicht einen gemeinsamen "Konfidenzbereich" über verschiedene leadzeiten bilden.

Beispiel 5.2.1.: ARIMA(1,1,0)-Prozeß

Für den Zeitraum  $L=7$  berechnen wir aus den  $\psi_1$  die Varianzfunktion  $V(1)$  ( $\hat{\phi}_1=0,8$ )

1	1	2	3	4	5	6	7
$\psi_1$	1.8	2.44	2.95	3.36	3.69	3.95	4.16
$V(1)/\sigma_a^2$	1.00	4.24	10.19	18.96	30.24	43.82	59.48

Man sieht, wie rasch die Varianzfunktion ansteigt. In der folgenden Abbildung sind die zugehörigen bedingten (normalverteilten) Dichten angedeutet



Für  $l=2$  und  $\alpha=50\%$  und  $\alpha=5\%$  seien die Konfidenzintervalle berechnet.

Die geschätzte Standardabweichung beträgt  $\hat{\sigma}_a = 0,134$

a)  $\alpha=50\%$ ,  $t\left(\frac{1+\alpha}{2}\right) = 0,674$

$$z_{t+2}(\pm) = \hat{z}_t(2) \pm 0,674 \cdot 0,134 \sqrt{1+1,8^2} = \hat{z}_t(2) \pm 0,19$$

b)  $\alpha=5\%$ ,  $t\left(\frac{1+\alpha}{2}\right) = 1,96$

$$z_{t+2}(\pm) = \hat{z}_t(2) \pm 1,96 \cdot 0,134 \sqrt{4,24} = \hat{z}_t(2) \pm 0,55$$

### 5.2.3. Unabhängigkeit des Prognosefehlers

Wir betrachten nun einstufige Prognosen und den damit verbundenen einstufigen Prognosefehler  $e_t(1)$  ("one step ahead forecast error").

#### Satz 5.2.2.: Einstufige Unabhängigkeit

Der einstufige Prognosefehler ist unabhängig von den vergangenen Störungen

$$\text{Cov}(e_t(1), a_t) = 0 \quad (5.2.10)$$

#### Beweis:

Der einstufige Prognosefehler ist

$$\begin{aligned} e_t(1) &= z_{t+1} - \hat{z}_t(1) = \\ &= \sum_{j=-\infty}^{t+1} \psi_{t+1-j} a_j - \sum_{j=-\infty}^t \psi_{t+1-j} a_j \\ &= a_{t+1} \end{aligned} \quad (5.2.11)$$

Damit folgt aber

$$\text{Cov}(e_t(1), a_t) = \text{Cov}(a_{t+1}, a_t) = 0 \quad (5.2.12)$$

Für MSE-Prognosen gilt, daß der einstufigen Prognosefehler unkorreliert mit den vergangenen Störungen ist. Denn wäre der einstufige Prognosefehler nicht unkorreliert mit den vergangenen Störungen, so würde das bedeuten, daß der Prognose-

fehler  $e_t(1)$  teilweise aus den vergangenen Störungen  $a_t, a_{t-1}, \dots$  vorhergesagt werden könnte. Der Prognosefehler wäre in diesem Fall  $\hat{a}_{t+1}$  und  $\hat{z}_t(1) + \hat{a}_{t+1}$  wäre eine bessere Prognose für  $z_{t+1}$  als  $\hat{z}_t(1)$ .

Dieser Sachverhalt sei in folgender Tabelle demonstriert:

Prognose für	unkorreliert	korreliert
$a_{t+1}$	0	$\hat{a}_{t+1}$
$z_{t+1}$	$\hat{z}_t(1)$	$\hat{z}_t(1) + \hat{a}_{t+1}$

#### 5.2.4. Die Korrelation von Prognosefehlern

In Kapitel 5.2.3. haben wir gezeigt, daß der einstufige Prognosefehler  $e_t(1)$  unabhängig von den vergangenen Störungen  $a_t$  ist. Die Prognosefehler höherer Stufen sind jedoch nicht unabhängig voneinander. Da Prognosen auf bestimmten Informationsmengen  $I_t$  beruhen, die sich bei nahe beieinanderliegenden Prognosezeitpunkten nicht viel ändern, kann man erwarten, daß die Prognosen im allgemeinen korreliert sind. Wir unterscheiden dabei zwei Arten von Korrelationen.

##### Def. 5.2.2.: Vertikalkorrelation

Unter der Vertikalkorrelation von Prognosen verstehen wir die Korrelation zweier Prognosefunktionen mit gleicher leadzeit zu verschiedenen Prognosezeitpunkten

$$\text{Corr}(e_t(1), e_{t-j}(1)) \quad (5.2.13)$$

##### Satz 5.2.3.

Die Vertikalkorrelation der Prognose für ARIMA-Prozesse lautet

$$\text{Corr}(e_t(1), e_{t-j}(1)) = \frac{\sigma_a^2 \sum_{k=j}^{l-1} \psi_j \psi_{k-j}}{V(1)} \quad \begin{matrix} 0 \leq j < l \\ j > l \end{matrix} \quad (5.2.14)$$

$$= 0$$

wobei  $V(1)$  die Prognosevarianzfunktion nach (5.2.3) ist.

Beweis:

$$\begin{aligned} \text{Es ist } e_t(1) &= z_{t+1} - \hat{z}_t(1) = \\ &= \sum_{i=0}^{l-1} \psi_i a_{t+1-i} \end{aligned} \quad (5.2.15)$$

$$\begin{aligned} e_{t-j}(1) &= z_{t+1-j} - \hat{z}_{t-j}(1) = \\ &= \sum_{i=0}^{l-1} \psi_i a_{t+1-j-i} \end{aligned} \quad (5.2.16)$$

Die Kovarianz lautet somit

$$\begin{aligned} E[e_t(1), e_{t-j}(1)] &= \sum_{i=0}^{l-1} \psi_i a_{t+1-i} \sum_{k=j}^{j+l-1} \psi_{k-j} a_{t+1-k} = \\ &= \sigma_a^2 \sum_{k=j}^{l-1} \psi_k \psi_{k-j} \end{aligned} \quad (5.2.17)$$

Speziell für  $l=1$  lautet die Kovarianz

$$E(e_t(1), e_{t-j}(1)) = E a_t a_{t-j} = 0 \quad (5.2.18)$$

Beispiel 5.2.2.: ARIMA(1,1,0)-Prozeß

Speziell:  $(1-0,8B)\nabla z_t = a_t$

Wir berechnen  $E(e_t(6)e_{t-j}(6)) = \sum_{i=j}^5 \psi_i \psi_{i-j}$

$j=0:$   $E(e_t(6)^2) = V(6) = \sigma_a^2 \sum_{i=0}^5 \psi_i^2 = 43,86\sigma_a^2$

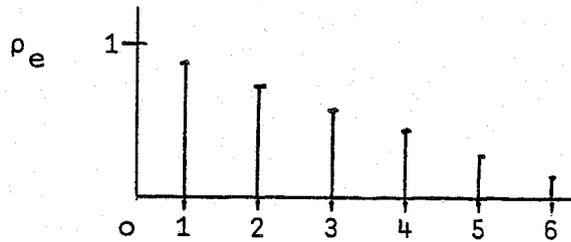
$j=1:$   $E(e_t(6)e_{t-1}(6)) = \sigma_a^2 (1,8+2,44 \cdot 1,8+\dots+3,69 \cdot 3,36) = 35,7\sigma_a^2$

$$\text{Corr}(e_t(6)e_{t-1}(6)) = 43,86/35,7 = 0,81$$

Die vollständige Korrelationsfunktion lautet (durch analoge Berechnung für  $j=0,1,2,\dots$ )

j	0	1	2	3	4	5
$\rho$	1,00	0,81	0,61	0,41	0,23	0,08

Je weiter die Prognosezeitpunkte auseinander liegen, desto kleiner ist die Korrelation der Prognosen.



Def. 5.2.3.: Horizontalkorrelation

Unter der Horizontalkorrelation von Prognosefehlern verstehen wir die Korrelation von verschiedenen Werten einer Prognosefunktion.

$$\text{Corr}(e_t(1), e_t(1+j)) \tag{5.2.19}$$

Satz 5.2.4.

Die Horizontalkorrelation der Prognosefunktion eines ARIMA-Prozesses lautet

$$\text{Corr}(e_t(1), e_t(1+j)) = \frac{\sum_{i=0}^{l-1} \psi_i \psi_{j+i}}{\sqrt{V(1)V(1+j)}} \tag{5.2.20}$$

Beweis:

Es ist 
$$e_t(1) = \sum_{i=0}^{l-1} \psi_i a_{t+1-i} \tag{5.2.21}$$

und

$$\begin{aligned} e_t(1+j) &= z_{t+1+j} - \hat{z}_t(1+j) \tag{5.2.22} \\ &= \sum_{i=0}^{l+j-1} \psi_i a_{t+1+j-i} \end{aligned}$$

Die Kovarianz ist somit

$$\text{Cov}(e_t(1), e_t(1+j)) = \sigma^2 \sum_{i=0}^{l-1} \psi_i \psi_{j+i} \tag{5.2.23}$$

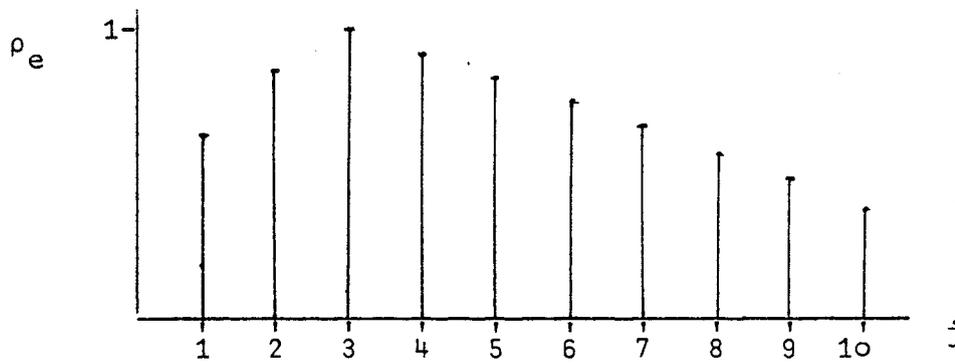
Wie man im folgenden Beispiel erkennen kann, so sind die Prognosen vom selben Prognosezeitpunkt hoch korreliert.

Beispiel 5.2.3.: ARIMA(1,1,0)-Prozeß ( $\phi=0,8$ )

$$E(e_t(3), e_t(5)) = \sigma_a^2 (2,44 + 1,8 \cdot 2,95 + 2,44 \cdot 3,36) = 15,94 \sigma_a^2$$

Analog berechnet man die weiteren Werte und erhält

j	1	2	3	4	5	6	7	8	9	10
Corr[ $e_t(3), e_t(j)$ ]	0,76	0,94	1,00	0,96	0,91	0,85	0,8	0,75	0,71	0,67



Diese beiden Prognosekorrelationen seien in folgender Tabelle der Prognosefehler zusammengefaßt

Prognosezeitpunkte	leadzeiten					
	1	1	2	3	...	1
t	$e_t(1)$	$e_t(2)$	...			$e_t(1)$
t-1	$e_{t-1}(1)$	$e_{t-1}(2)$	.	.	.	$\vdots$
t-j	$e_{t-j}(1)$	...	.	.	.	$e_{t-j-1}(1)$

Prognosefehler von

$z_{t+2}$

$z_{t+1}$

$z_t$

### 5.3. Darstellungsformen der Prognosefunktion

#### 5.3.1. Prognoseformen

Analog zu den Darstellungsformen des ARIMA-Prozesses (vgl. Kapitel 4.2.) unterscheiden wir vier Prognoseformen: Die D-Form, die RS-Form, die TRS-Form und die  $\pi$ -Form. Eine weitere Charakterisierung der Prognosefunktion erfolgt mittels der eventuellen Prognosefunktion, die im Kapitel 5.3.2. gesondert behandelt werden wird.

##### 1) Die Differenzengleichungsform

Die D-Form für den zukünftigen Wert  $z_{t+1}$  erhält man durch

$$\gamma(B)z_{t+1} = \theta(B)a_{t+1} \quad (5.3.1)$$

oder

$$z_{t+1} = \gamma_1 z_{t+1-1} + \dots + \gamma_{p+d} z_{t+1-p-d} - \theta_1 a_{t+1-1} \dots - \theta_q a_{t+1-q} + a_{t+1} \quad (5.3.2)$$

Die Prognosefunktion  $\hat{z}_t(1) = [z_{t+1}]$  erhalten wir durch den bedingten Erwartungswert

$$\gamma(B)[z_{t+1}] = \theta(B)[a_{t+1}] \quad (5.3.3)$$

oder ausgeschrieben

$$\hat{z}_t(1) = [z_{t+1}] = \gamma_1 [z_{t+1-1}] + \dots + \gamma_{p+d} [z_{t+1-p-d}] - \theta_1 [a_{t+1-1}] - \dots - \theta_q [a_{t+1-q}] + [a_{t+1}] \quad (5.3.4)$$

Für  $q > 1$  erfüllen sie die Differenzengleichung

$$\gamma(B)[z_{t+1}] = 0 \quad (5.3.5)$$

2) Die RS-Form

Die RS-Form für den Wert  $z_{t+1}$  lautet

$$z_{t+1} = \sum_{j=-\infty}^{t+1} \psi_{t+1-j} a_j \quad (5.3.6)$$

Entsprechend lautet die Prognosefunktion

$$\hat{z}_t(1) = \sum_{j=-\infty}^{t+1} \psi_{t+1-j} [a_j] \quad (5.3.7)$$

3) Die TRS-Form

Als truncation point für die TRS-Form empfiehlt es sich, den Prognosezeitpunkt  $k=t$  zu nehmen. Wir erhalten somit für den Wert  $z_{t+1}$

$$z_{t+1} = C_t(1) + a_{t+1} + \psi_1 a_{t+1-1} + \dots + \psi_{1-1} a_{t+1} \quad (5.3.8)$$

wobei die allgemeine Lösung der homogenen Differenzgleichung  $C_t(1)$  durch

$$C_t(1) = \sum_{j=-\infty}^t \psi_{t+1-j} a_j = \sum_{j=0}^{\infty} \psi_{1+j} a_{t-j} \quad 1 > q > 0 \quad (5.3.9)$$

gegeben ist. Entsprechend lautet die Prognosefunktion

$$\hat{z}_t(1) = C_t(1) + \sum_{j=t+1}^{t+1} \psi_{t+1-j} [a_j] \quad (5.3.10)$$

4) Die  $\pi$ -Form

Die invertierte Form des ARIMA-Prozesses für den Wert  $z_{t+1}$  lautet

$$z_{t+1} = \sum_{j=1}^{\infty} \pi_j z_{t+1-j} + a_{t+1} \quad (5.3.11)$$

Falls  $d \geq 1$  ist, schreiben wir

$$\bar{z}_{t+1-1}(\pi) = \sum_{j=1}^{\infty} \pi_j z_{t+1-j} \quad (5.3.12)$$

da die Summe der  $\pi_j$  eins ergibt und  $z_{t+1}$  wird in der normierten  $\pi$ -Form dargestellt

$$z_{t+1} = \bar{z}_{t+1-1}(\pi) + a_{t+1}$$

Nehmen wir bedingte Erwartungswerte in (5.3.10), so erhalten wir die Prognosefunktion in  $\pi$ -Form

$$\begin{aligned} \hat{z}_t(1) &= \sum_{j=1}^{\infty} \pi_j [z_{t+1-j}] + [a_{t+1}] = \\ &= \sum_{j=1}^{\infty} \pi_j \hat{z}_t(1-j) \end{aligned} \quad (5.3.13)$$

$$\text{wobei } z_t(-1) = z_{t-1} \text{ gilt.} \quad (5.3.14)$$

Beispiel 5.3.1.: ARIMA(1,1,0)-Prozeß

Prognosen in Differenzgleichungsform.

Gegeben sei der Prozeß mit  $\phi = 0,8$

$$(1-0,8B)(1-B)z_{t+1} = a_{t+1}$$

$$(1-0,8B+0,8B^2)z_{t+1} = a_{t+1}$$

$$z_{t+1} = 1,8z_{t+1-1} - 0,8z_{t+1-2} + a_{t+1}$$

Die Prognosefunktion ist

$$l=1 \quad \hat{z}_t(1) = 1,8z_t - 0,8z_{t-1}$$

$$l=2 \quad \hat{z}_t(2) = 1,8\hat{z}_t(1) - 0,8z_t$$

$\vdots$

$$l=3,4,5.. \quad \hat{z}_t(l) = 1,8\hat{z}_t(l-1) - 0,8\hat{z}_t(l-2)$$

bzw. erfüllt die Differenzgleichung

$$(1-0,8B)(1-B)\hat{z}_t(l) = 0 \quad \text{für } l=3,4,5..$$

### 5.3.2. Die eventuelle Prognosefunktion

Um das allgemeine Verhalten der Prognose verschiedener ARIMA(p,d,q)-Prozesse zu analysieren, führen wir das Konzept der eventuellen Prognosefunktion EFF (eventual forecast function) ein. Sie stützt sich im wesentlichen auf die Differenzgleichungsform von Prognosen und deren Darstellung als homogene Lösung der Differenzgleichung.

Nach Formel (5.3.3) lautet die D-Form für Prognosen

$$\psi(B)\hat{z}_t(l) = \theta(B)[a_{t+1}] \quad (5.3.15)$$

Diese erfüllt für  $l > q$  die Differenzgleichung

$$\psi(B)\hat{z}_t(l) = 0 \quad l > q \quad (5.3.16)$$

Dies führt zu der folgenden Definition

Def. 5.3.1.: Eventuelle Prognosefunktion (EFF)

Sei  $\hat{z}_t(l)$  die Prognosefunktion eines ARIMA-Prozesses, so nennen wir die Lösung der homogenen Gleichung der Prognosefunktion (5.3.16) für  $l > q - p - d$  eventuelle Prognosefunktion (EFF)

$$\begin{aligned} z_t(l) &= \sum_{i=0}^{p+d-1} b_i^{(t)} f_i(l) = && l > q - p - d \\ &= b_0^{(t)} f_0(l) + b_1^{(t)} f_1(l) + \dots + b_{p+d-1}^{(t)} f_{p+d-1}(l) \end{aligned} \quad (5.3.17)$$

Die  $b_i^{(t)}$  sind Konstante, die für jeden Prognosezeitpunkt bestimmt werden. Die  $f_i(l)$  werden aufgrund der Wurzeln der charakteristischen Gleichung bestimmt.

Man beachte, daß die Koeffizienten  $b_i^{(t)}$  der eventuellen Prognosefunktion konstant für einen bestimmten Zeitpunkt  $t$  sind. Wechselt man von einem Prognosezeitpunkt  $t$  zu einem anderen  $t+1$ , so verändern sich die  $b_i^{(t)}$ . Sie sind also adaptiv, und können sich deshalb an denjenigen Teil der Zeitreihe anpassen, den man gerade prognostizieren will.

Verhalten der EFF: Da die EFF Lösung der homogenen Differenzengleichung ist, wird ihr Verhalten durch die Wurzeln der charakteristischen Gleichung (ARI-Operator) bestimmt. Sie setzt sich aus gedämpften Exponential- und Sinusschwingungen, aus Polynomen oder einer Kombination dieser zusammen.

Pivotprognosen: Da die EFF eine homogene Differenzengleichung von der Ordnung  $p'=p+d$  ist, benötigen wir  $p'$  Anfangsbedingungen zur Fixierung der EFF. Wir kommen derart zur

Def. 5.3.2.: Pivotprognose

Sei  $\overset{v}{z}_t(1)$   $1 > q-p-d$  eine EFF. Dann heißen die  $p'=p+d$  ersten Werte der EFF Pivotprognosen

$$\overset{v}{z}_t(j) \quad j=1, \dots, p'=p+d$$

Wir unterscheiden 3 Fälle der Pivotprognosen

Fall 1: Ist  $q > p+d$ , so lauten die Pivotprognosen

$$\overset{v}{z}_t(j) = \hat{z}_t(q-p-d+j) \quad j=1, \dots, p'=p+d \quad (5.3.18)$$

Die ersten  $q$  Werte der Prognosefunktion  $\hat{z}_t(1)$  nennen wir "Komplementärprognose" der EFF  $\overset{v}{z}_t(1)$ , da sie nicht durch die EFF erfaßt werden.

Beispiele: MA( $q$ )-Prozesse, IMA(1,2)-, IMA(2,3)-Prozesse, allgemein IMA( $d,q$ ), falls  $q > d$ .

Fall 2: Ist  $q=p+d$ , so gilt

$$\overset{v}{z}_t(1) = \hat{z}_t(1) \quad 1=1,2,\dots \quad (5.3.19)$$

und die Pivotprognosen sind die ersten  $p'=q$  Werte der Prognosefunktion, d.h. die EFF und die Prognosefunktion fallen zusammen.

Beispiele: IMA(1,1)-, IMA(2,2)-Prozeß, ARMA(1,1)-Prozeß.

Fall 3: Ist  $q < p+d$ , so gibt es  $p'$  Pivotprognosen, von denen  $q' = p' - q$  vor dem Prognosezeitpunkt  $t$  liegen, die restlichen danach

$$\hat{z}_t(j) = \hat{z}_t(-j)$$

und

$$\hat{z}_t(j) \begin{cases} = z_{t-j} & j=1, \dots, q' \\ = z_t(j) & j=q'+1 \dots p' \end{cases} \quad (5.3.20)$$

$$\hat{z}_t(j) \begin{cases} = z_{t-j} & j=1, \dots, q' \\ = z_t(j) & j=q'+1 \dots p' \end{cases} \quad (5.3.21)$$

Beispiel: ARIMA(1,1,1)-Prozeß:

Ist  $0 = q < p+d$ , so liegen alle  $p'$  Pivotprognosen unter  $t$ , d.h. es gilt

$$\hat{z}_t(j) = z_{t-j} \quad j=1, \dots, p' \quad (5.3.22)$$

Beispiele: AR(p)-, ARIMA(p,d,o)-Prozesse.

### Die Rolle des Moving Average Operators

Bestimmt in einem ARIMA(p,d,q)-Prozeß der ARI-Teil im wesentlichen das allgemeine Verhalten der Prognosefunktion (d.i. die EFF), so hat der MA-Teil des Prozesses wesentlichen Einfluß auf die Startwerte der EFF, den Pivotprognosen. In Fall 2 und Fall 3 sind es jeweils die  $q$  ersten Werte der Prognosefunktion, die als Pivotprognose noch den MA-Teil beinhalten. Im Fall 3 kommen dazu noch  $q'$  vergangene Werte der Zeitreihe.

Im Fall 1 ist der MA-Teil so dominierend, daß die ersten  $q$  Werte nicht einer allgemeinen Beschreibung zugänglich sind. Erst die nächsten  $p'$  Werte kommen als Pivotprognosen zum Zug.

### 5.3.3. Die MSE-Prognose in $\pi$ -Form

Das MSE-Kriterium fordert die Minimierung des Quadrates des Prognosefehlers (vgl. (5.1.2))

$$E(e_t(1)^2) = \text{Min!} \quad (5.3.23)$$

Setzen wir dafür die Prognosefunktion in Form (5.1.4) und die  $\pi$ -Form des Modells ein, so erhalten wir

$$\begin{aligned} E(e_t(1)^2) &= E(z_{t+1} - \hat{z}_t(1))^2 = \\ &= E\left(\sum_{j=1}^{\infty} \pi_j z_{t+1-j} + a_{t+1} - \sum_{j=1}^{\infty} \alpha_j(1) z_{t-j}\right)^2 \end{aligned} \quad (5.3.24)$$

Diese Summe ist dann ein Minimum, wenn die beiden Summen gleich sind

$$\sum_{j=1}^{\infty} \pi_j z_{t+1-j} = \sum_{j=1}^{\infty} \alpha_j(1) z_{t-j} \quad (5.3.25)$$

Dazu werden alle  $z_{t+1-j}$ , die über dem Prognosezeitpunkt  $t$  liegen, durch die  $\pi$ -Form ersetzt.

Nach Formel (5.3.13) lautet die Prognose in  $\pi$ -Form

$$\begin{aligned} \hat{z}_t(1) &= \sum_{j=1}^{\infty} \pi_j z_t(1-j) = \\ &= \pi_1 \hat{z}_t(1-1) + \dots + \pi_{1-1} \hat{z}_t(1) + \\ &+ \pi_1 z_t + \pi_1 z_{t-1} + \dots \end{aligned} \quad (5.3.26)$$

Speziell gilt für die einstufige Prognose

$$\hat{z}_t(1) = \pi_1 z_t + \pi_2 z_{t-1} + \pi_3 z_{t-2} + \dots \quad (5.3.27)$$

daß sie eine Funktion der bisherigen Beobachtungen allein ist, wobei die gewichtete Summe die bekannten  $\pi$ -Gewichte benützt.

Entsprechend kann man die Prognose zweiter Stufen lediglich als gewichtete Summe der bisher beobachteten Zeitreihe darstellen:

$$\begin{aligned}
 \hat{z}_t(2) &= \pi_1 \hat{z}_t(1) + \pi_2 z_t + \pi_3 z_{t-1} + \dots \\
 &= \pi_1 \sum_{j=1}^{\infty} \pi_j z_{t-j+1} + \sum_{j=1}^{\infty} \pi_{j+1} z_{t-j+1} = \\
 &= \sum_{j=1}^{\infty} (\pi_1 \pi_j + \pi_{j+1}) z_{t-j+1} = \\
 &= \sum_{j=1}^{\infty} \pi_j^{(2)} z_{t-j+1} \quad (5.3.28)
 \end{aligned}$$

Die  $\pi_j^{(2)}$  sind die Prognosekoeffizienten zweiter Ordnung der  $\pi$ -Form. Setzt man diese Prozedur fort, so kommt man zu

Satz 5.3.1.: Prognosekoeffizienten  $\pi_j^{(1)}$

Die Prognosekoeffizienten  $l$ -ter Ordnung der  $\pi$ -Prognoseform sind gegeben durch

$$\pi_j^{(1)} = \sum_{i=1}^l \psi_{l-i} \pi_{i+j-1} \quad (5.3.29)$$

Ausführlich lautet dieses Gleichungssystem

$$\begin{aligned}
 \pi_1^{(1)} &= \psi_{l-1} \pi_1 + \psi_{l-2} \pi_2 + \dots + \pi_l \\
 \pi_2^{(1)} &= \psi_{l-1} \pi_2 + \psi_{l-2} \pi_3 + \dots + \pi_{l+1} \\
 &\vdots \\
 \pi_j^{(1)} &= \psi_{l-1} \pi_j + \psi_{l-2} \pi_{j+1} + \dots + \pi_{l+j-1}
 \end{aligned} \quad (5.3.30)$$

Beweis: Dazu drücken wir  $\hat{z}_t(1)$  durch die vorhergehende einstufige Prognose  $\hat{z}_{t+1-1}(1)$  aus:

Es ist

$$\hat{z}_t(1) = \sum_{j=0}^{\infty} \psi_{1+j} a_{t-j} \quad (5.3.31)$$

$$\hat{z}_{t+1-1}(1) = \psi_1 a_{t+1-a} + \dots + \psi_{l-1} a_{t+1} + \sum_{j=0}^{\infty} \psi_{1+j} a_{t-j} \quad (5.3.32)$$

Subtrahieren wir beide Ausdrücke, so erhalten wir

$$\hat{z}_t^{(1)} = \hat{z}_{t+1-1}^{(1)} - \psi_1 a_{t+1-1} - \psi_2 a_{t+1-2} - \dots - \psi_{l-1} a_{t+1} \quad (5.3.33)$$

Diese Form drücken wir der Reihe nach in der  $\pi$ -Form aus

$$\begin{aligned} \hat{z}_t^{(1)} = & \pi_1 z_{t+1-1} + \pi_2 z_{t+1-2} + \pi_3 z_{t+1-3} + \dots + \pi_{l-1} z_{t+1} + \pi_l z_t + \sum_{j=1}^{\infty} \pi_{l+j} z_{t-j} \\ & + \psi_1 (-z_{t+1-1} + \pi_1 z_{t+1-2} + \dots + \pi_{l-2} z_{t+1} + \pi_{l-1} z_t + \sum_{j=1}^{\infty} \pi_{l+j-1} z_{t-j}) \\ & + \psi_2 (-z_{t+1-2} + \dots + \pi_{l-3} z_{t+1} + \pi_{l-2} z_t + \sum_{j=1}^{\infty} \pi_{l+j-2} z_{t-j}) \\ & \dots \dots \dots \\ & + \psi_{l-1} (-z_{t+1} + \pi_1 z_t + \sum_{j=1}^{\infty} \pi_{j+1} z_{t-j}) \end{aligned} \quad (5.3.34)$$

Wegen 
$$\psi_j = \sum_{i=1}^j \psi_{j-i} \pi_i \quad (5.3.35)$$

verschwinden die Koeffizienten von  $z_{t+1-1}, \dots, z_{t+1}$ .

Ordnen wir die weiteren Koeffizienten  $z_t, z_{t-1}, \dots$ , so erhalten wir gerade Formel (5.3.29).

Die Prognosekoeffizienten  $\pi_j^{(1)}$  kann man in zwei Rekursionsformeln darstellen

$$\begin{aligned} \pi_j^{(1)} &= \pi_{j+1-1} + \sum_{h=1}^{l-1} \pi_h \pi_j^{(1-h)} \\ &= \pi_{j+1}^{(1-1)} + \psi_{l-1} \pi_j \quad l > 0, j=1, 2, \dots \end{aligned} \quad (5.3.36)$$

wobei man speziell definiert

$$\pi_j^{(1)} = \pi_j$$

Die Prognosefunktion in  $\pi$ -Form lautet schließlich

$$\hat{z}_t(1) = \sum_{j=1}^{\infty} \pi_j^{(1)} z_{t-j+1} \quad (5.3.37)$$

Damit sind die  $\alpha_j(1)$ -Koeffizienten nach (5.1.4) durch

$$\alpha_j(1) = \pi_j^{(1)} \quad (5.3.38)$$

bestimmt.

Beispiel 5.3.3.: IMA(2,2)-Prozeß

Die D-Form ist

$$\nabla^2 z_t = (1 - 0,9B + 0,5B^2) a_t$$

Durch Koeffizientenvergleich

$$(1 - 2B + B^2) = (1 - 0,9B + 0,5B^2)(1 - \pi_1 B - \pi_2 B^2 - \dots)$$

erhalten wir die  $\pi_j$ -Koeffizienten. Daraus berechnen wir nach Formel (5.3.36) die  $\pi^{(2)}$ -Koeffizienten.

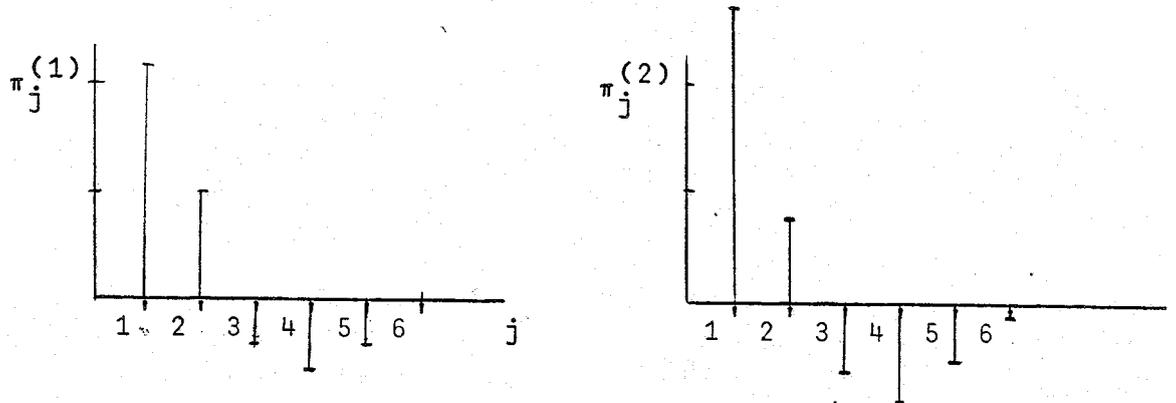
j	1	2	3	4	5	6
$\pi_j = \pi_j^{(1)}$	1,1	0,49	-0,11	-0,34	-0,25	-0,06
$\pi_j^{(2)}$	1,7	0,43	-0,46	-0,63	-0,33	0,01

Die Prognosen als Funktionen der Beobachtungen alleine lauten

$$\hat{z}_t(1) = 1,1z_t + 0,49z_{t-1} - 0,11z_{t-2} - \dots$$

$$\hat{z}_t(2) = 1,7z_t + 0,43z_{t-1} - 0,46z_{t-2} - \dots$$

Graphisch haben die  $\pi_j$ -Koeffizienten folgenden Verlauf



Sind die  $\pi_j > 1$ , so kommt ihnen in der nächsten Stufe ein überproportional größeres Gewicht zu.

#### 5.3.4. Updaten von Prognosen

Das Updaten von Prognosen ist ein vereinfachtes Verfahren zur Revision von Prognosen, wenn eine neue Beobachtung der Zeitreihe bekannt wird. Durch Bewertung des einstufigen Prognosefehlers werden die letzten prognostizierten Werte korrigiert.

Zum Prognose-updating benötigen wir die Prognose in der RS-Form

$$\hat{z}_t(1) = \sum_{j=0}^{\infty} \psi_{1+j} a_{t-j} \quad (5.3.39)$$

Ein zukünftiger Wert  $z_{t+1}$ , der zum Zeitpunkt  $t$  prognostiziert wurde, kann nach Bekanntwerden von  $z_{t+1}$  zum Zeitpunkt  $t+1$  neu prognostiziert werden. Die entsprechenden Prognosen in der RS-Form lauten:

$$\hat{z}_{t+1}(1) = \psi_1 a_{t+1} + \psi_{1+1} a_t + \psi_{1+2} a_{t-1} + \dots \quad (5.3.40)$$

$$\hat{z}_t(1+1) = \psi_{1+1} a_t + \psi_{1+2} a_{t+1} + \dots \quad (5.3.41)$$

Bilden wir die Differenz der beiden Prognosen, so erhalten wir

$$\hat{z}_{t+1}(l) - \hat{z}_t(l+1) = \psi_1 a_{t+1}$$

Damit erhalten wir die einfache Prognoseupdate-Formel

$$\hat{z}_{t+1}(l) = \hat{z}_t(l+1) + \psi_1 a_{t+1} \quad l=1, \dots, L-1 \quad (5.3.42)$$

Die Prognose von  $z_{t+1+1}$  zum Zeitpunkt  $t+1$  wird durch die Prognose zum Zeitpunkt  $t$  dadurch "upgedated", indem man zur alten Prognose den Korrekturfaktor  $\psi_1 a_{t+1}$  dazuaddiert.  $a_{t+1}$  ist dabei der einstufige Prognosefehler.

$$a_{t+1} = z_{t+1} - \hat{z}_t(1) \quad (5.3.43)$$

Die implizite Annahme bei diesem Vorgehen ist die Konstanz der Modellparameter, d.h. die  $\psi$ -Koeffizienten würden sich nur unwesentlich bei einer Neuschätzung ändern. Je länger der letzte Prognosezeitpunkt  $t$  zurückliegt, desto weniger Werte können upgedated werden und desto eher wird man eine Änderung in den Parametern vermuten.

Bei zu großem Abstand der beiden Prognosezeitpunkte leidet auch die Adaptationsfähigkeit der Prognose. Die kurzfristige Anpassung wird gerade durch die Moving-Average-Teile des Modells bestimmt, der in den  $\psi_1$ -Koeffizienten für große  $l$  kaum mehr enthalten ist.

#### Praktisches Vorgehen:

Angenommen, wir stehen am Zeitpunkt  $t$  und haben Prognosen für die Leadtime  $L$ ,  $l=1, \dots, L$

$$\hat{z}_t(1), \dots, \hat{z}_t(L)$$

berechnet. Wir erhalten zum Zeitpunkt  $t+1$  den neuen Wert der Zeitreihe  $z_{t+1}$ .

Wie ändert sich die Prognose durch diesen neuen Wert?

Es gibt zwei Möglichkeiten zur Revision der Prognose:

- 1) Neuschätzen des Modells mit dem zusätzlichen Wert und neuerliche Prognose;
- 2) Updaten der alten Prognose, bzw. Revision der Prognose für die Leadtime L. Dazu berechnen wir nach Formel (5.3.37) den Prognosefehler, und können die neuen Prognosen updaten

$$\hat{z}_{t+1}(l) = \hat{z}_t(l+1) + \psi_1 a_{t+1} \quad \text{für } l=1,2,\dots,L-1$$

Die Prognose für die "neue" Leadzeit L (bzw. "alte" Leadzeit L+1) kann mit Hilfe der Update-Formel nicht berechnet werden.

$\hat{z}_{t+1}(L)$  kann aber durch die D-Form der Prognose ermittelt werden.

Beispiel 5.3.4.: ARIMA(1,1,1)-Prozeß

	$z_{21}$	$z_{22}$	$z_{23}$	$z_{24}$
1	1	2	3	4
$\hat{z}_{20}(1)$	23,16	22,97	22,81	22,69
$\hat{z}_{21}(1)$	[23,1]	22,86	22,67	22,51
$\psi_1$	1,8	2,44	2,95	3,36

Die geschätzten Prognosen  $\hat{z}_{20}(1)$  werden aufgrund des neuen Wertes  $\hat{z}_{21}=23,1$  (in der Tabelle in eckiger Klammer angegeben) upgedated. Der einstufige Prognosefehler lautet

$$\begin{aligned} e_{20}(1) &= a_{21} = z_{21} - \hat{z}_{20}(1) = \\ &= 23,1 - 23,16 = 0,06 \end{aligned}$$

Die einzelnen Rechenschritte lauten

$$\hat{z}_{21}(1) = 22,97 + 1,8(-0,06) = 22,86$$

$$\hat{z}_{21}(2) = 22,81 + 2,44(-0,06) = 22,67$$

$$\hat{z}_{21}(3) = 22,69 + 2,95(-0,06) = 22,51$$

Da der eingetroffene Wert leicht unter der Prognose  $\hat{z}_{20}(1)$  liegt, werden alle Prognosen leicht nach unten korrigiert:

$\hat{z}_{21}(4) = \hat{z}_{25}$  ist nach der Update-Methode nicht zu erhalten. Wir müssen nun auf die DE übergehen.

$$(B)z_{t+1} = \theta(B)a_{t+1}$$

$$(1 - 1,8B + 0,8B^2)z_{t+1} = a_{t+1}$$

$$(1 - 1,8B + 0,8B^2)\hat{z}_t(1) = 0 \quad l > 2, \quad l = 3, 4, 5, \dots$$

d.i.  $\hat{z}_t(1) = 1,8\hat{z}_t(1-1) - 0,8\hat{z}_t(1-2)$

$$\hat{z}_{21}(4) = 1,8\hat{z}_{21}(3) - 0,8\hat{z}_{21}(2)$$

$$= 1,8 \cdot 22,51 - 0,8 \cdot 22,67$$

$$= 22,39$$

### 5.3.5. Updaten der eventuellen Prognosefunktion

Wie wir in Kapitel 5.3.2. gesehen haben, lassen sich die Prognosen durch die EFF ausdrücken

$$\hat{z}_t^v(1) = \sum_{i=0}^{p+d-1} b_i^{(t)} f_i(1) \quad l > q - p - d \quad (5.3.44)$$

Die Koeffizienten  $b_i^{(t)}$  werden aus den Pivotprognosen  $\hat{z}_t^v(j)$ ,  $j=1, \dots, p+d$ , bestimmt. Diese sind vom jeweiligen Prognosezeitpunkt  $t$  abhängig. Wird eine neue Prognose zum Zeitpunkt  $t+1$  durchgeführt, so verschiebt sich die Prognosefunktion; die neuen Pivotprognosen  $\hat{z}_{t+1}^v(j)$  bestimmen neue Koeffizienten  $b_i^{(t+1)}$ . Aus diesem Grunde nennen wir die Prognosefunktion eines ARIMA( $p, d, q$ )-Prozesses adaptiv.

Die adaptiven Koeffizienten  $b_i^{(t)}$  müssen bei aufeinanderfolgenden Prognosezeitpunkten nicht neu aus den Pivotprognosen berechnet werden, sondern können aus den letzten Koeffizienten "upgedated" werden.

Dazu verwenden wir die Update-Formel (5.3.42)

$$\hat{z}_{t+1}(1) = \hat{z}_t(1) + \psi_1 a_{t+1} \quad (5.3.45)$$

Setzen wir in diese die EFF (5.3.17) ein, so erhalten wir

$$\sum_{i=0}^{p+d-1} b_i^{(t+1)} f_i(1) = \sum_{i=0}^{p+d-1} b_i^{(t)} f_i(1) - \psi_1 a_{t+1} \quad l > q-p+d \quad (5.3.46)$$

Lösen wir  $p+d$  dieser Gleichungen für  $l=1, \dots, p+d$ , so erhalten wir die Update-Formel für den  $b_i^{(t)}$ -Koeffizienten

$$b_i^{(t+1)} = \sum_{j=0}^{p+d-1} g_{ij} b_j^{(t)} + h_i a_{t+1} \quad i=0, \dots, p+d-1 \quad (5.3.47)$$

Man beachte, daß die Anpassung der EFF-Koeffizienten nur vom einstufigen Prognosefehler

$$a_{t+1} = z_{t+1} - \hat{z}_t(1) \quad (5.3.48)$$

abhängt. Die Konstanten  $g_{ij}$  sind lediglich Funktionen der  $f_i(1)$ , die jeweils aus den Wurzeln des ARI-Operatorpolynoms bestimmt werden.

### Darstellung in Matrizen

Die Updateformel läßt sich übersichtlicher durch die Matrixschreibweise der EFF darstellen.

$$\begin{aligned} \text{Seien } \underline{b}^{(t)} &= (b_0^{(t)}, \dots, b_{p+d-1}^{(t)})' \\ \underline{\psi}_1 &= (\psi_1, \dots, \psi_{1+p+d})' \end{aligned} \quad (5.3.49)$$

vom Typ  $\underline{b}^{(t)} = (p \times 1)$  und  $\underline{\psi}_1 = (p \times 1)$  mit  $p' = p+d$  die Vektoren der EFF- und  $\psi_1$ -Koeffizienten. Weiters sei  $F_e = (p' \times p')$  die Matrix der Elemente der EFF

$$F_e = \begin{pmatrix} f_0(1) & f_1(1) & \dots & f_{p'-1}(1) \\ f_0(1+1) & f_1(1+1) & & \\ \vdots & & \ddots & \\ f_0(1+p'-1) & f_1(1+p'-1) & & f_{p'-1}(1+p'-1) \end{pmatrix} \quad (5.3.50)$$

so lautet die EFF in Matrixform

$$\underline{z}_t(1) = \underline{F}_e \underline{b}^{(t)} \quad (5.3.51)$$

wobei

$$\underline{z}_t(1) = (z_t(1), \dots, z_t(1+p'-1)) \quad (5.3.52)$$

der Vektor von  $p'$  Werten der EFF für  $1, 1+1, \dots, 1+p'$  ist.

Setzen wir diese in die Update-Formel (5.3.45) ein, so erhalten wir

$$\underline{F}_e \underline{b}^{(t+1)} = \underline{F}_{e+1} \underline{b}^{(t)} + \psi_1 a_t \quad (5.3.53)$$

Für  $|\underline{F}_e| \neq 0$  ist

$$\begin{aligned} \underline{b}^{(t+1)} &= (\underline{F}_e^{-1} \underline{F}_{e+1}) \underline{b}^{(t)} + (\underline{F}_e^{-1} \psi_1) a_t \\ &= \underline{G}' \underline{b}^{(t)} + \underline{h} a_t \end{aligned} \quad (5.3.54)$$

mit

$$\begin{aligned} \underline{G}' &= (\underline{F}_e^{-1} \underline{F}_{e+1}) \\ \underline{h} &= (\underline{F}_e^{-1} \psi_1) \end{aligned} \quad (5.3.55)$$

wobei die  $G = (g_{ij}) \quad i, j=1, \dots, p+d$  und

$$h = (h_i) \quad i=1, \dots, p+d$$

der Form (5.3.47) entsprechen.

In dieser Form lassen sich die Brown'sche Methode der "diskontierten Kleinstquadrate" (discounted least squares)-Prognosen mit den MSE-Prognosen der ARIMA-Prozesse vergleichen (vgl. Brown (1963), Brown & Meyer (1961), Ledolter (1976)).

5.4. Die Prognose autoregressiver Prozesse

Zunächst betrachten wir als Spezialfall autoregressiver Prozesse die ARIMA(0,d,0)-Prozesse für  $d=1$  und  $d=2$ . Anschließend werden die autoregressiven Prozesse erster und zweiter Ordnung besprochen.

5.4.1. ARIMA(0,1,0)-Prozeß

Dieser Prozeß ist der "random-walk"-Prozeß

$$(1-B)z_t = a_t \quad (5.4.1)$$

bzw.

$$z_t = z_{t-1} + a_t$$

Die EFF ist die Lösung von

$$(1-B) \sum_t(1) = 0 \quad l > -d=1 \quad (5.4.2)$$

bzw.

$$\sum_t(1) = b_0^{(t)} = z_t \quad (5.4.3)$$

und überdeckt die Prognosefunktion bis auf die Pivotprognose

$$\sum_t(0) = z_t \quad (5.4.4)$$

Die MSE-Prognose für den random-walk-Prozeß ist der letzte Wert der Zeitreihe (für alle leadzeiten)

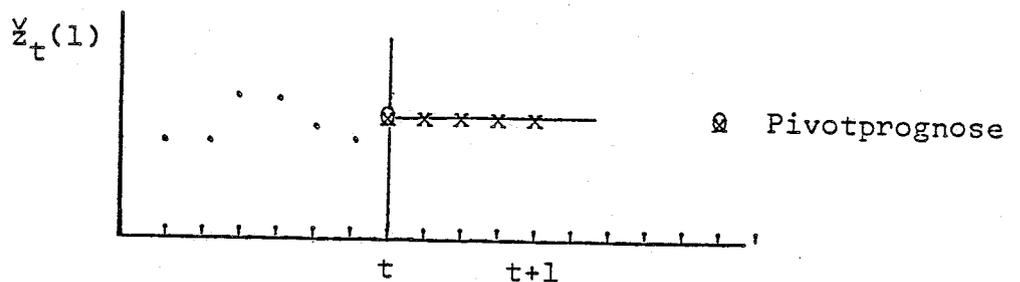


Abb.: Prognose eines "random-walk"-Prozesses

Die Varianzfunktion erhält man für  $\phi \rightarrow 1$  aus der Varianzfunktion eines AR(1)-Prozesses. Man erkennt in (5.4.21), daß die Varianzfunktion für  $\phi \rightarrow 1$  explodiert.

5.4.2. ARIMA(0,2,0)-Prozeß

Die D-Form des Prozesses lautet

$$(1-B)^2 z_t = a_t \quad (5.4.5)$$

Die EFF ist Lösung von

$$(1-B)^2 \check{z}_t(1) = 0 \quad |>-d=-2 \quad (5.4.6)$$

$$\check{z}_t(1) = b_0^{(t)} + b_a^{(t)} 1 \quad (5.4.7)$$

Die Pivotprognosen sind

$$\check{z}_t(0) = z_t \quad (5.4.8)$$

$$\check{z}_t(-1) = z_{t-1} \quad (5.4.9)$$

Daraus lassen sich die  $b_i$ -Koeffizienten bestimmen

$$b_0 - b_1 = z_t \quad b_0 = z_t + b_1 \quad (5.4.10)$$

$$b_0 - 2b_1 = z_{t-1}$$

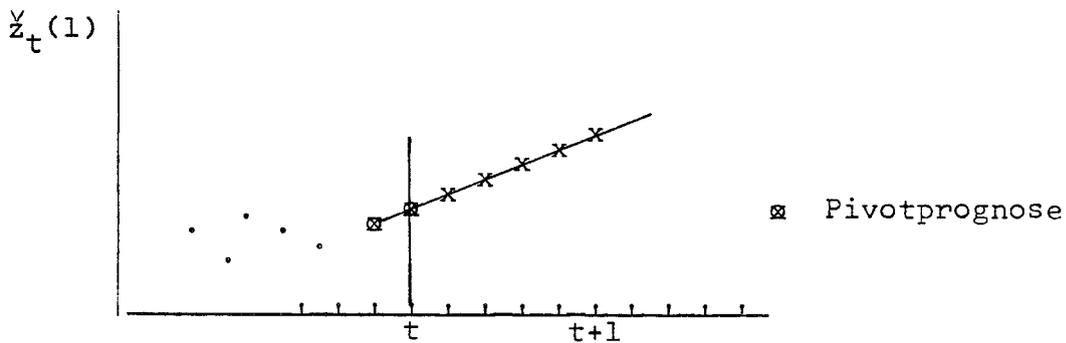
$$b_1 = z_t - z_{t-1} \quad (5.4.11)$$

$$b_0 = 2z_t - z_{t-1}$$

Die EFF lautet somit

$$\check{z}_t(1) = 2z_t - z_{t-1} + (z_t - z_{t-1})l \quad |>-1 \quad (5.4.12)$$

und überdeckt die Pivotfunktion



Die Varianzfunktion explodiert ebenfalls.

### 5.4.3. AR(1)-Prozeß

Die D-Form ist gegeben als

$$(1-\phi B)z_t = a_t \quad (5.4.13)$$

Die EFF ist Lösung von

$$(1-\phi B)\check{z}_t(1) = 0 \quad l > 1-p-d=-1 \quad (5.4.14)$$

bzw.

$$\begin{aligned} \check{z}_t(1) &= \phi \check{z}_t(1-1) \\ &= b_0^{(t)} \phi^l \quad l \geq 0 \end{aligned} \quad (5.4.15)$$

Die Pivotprognose ist

$$\hat{z}_t(0) = b_0^{(t)} = z_t \quad (5.4.16)$$

Daher lautet die EFF

$$\check{z}_t(1) = z_t \phi^l \quad l \geq 0 \quad (5.4.17)$$

Die EFF überdeckt die Prognosefunktion und den letzten Zeitreihenwert. Wegen der Stationaritätsbedingung  $|\phi| < 1$  geht die Prognosefunktion exponentiell gegen Null, bzw. gegen den Mittelwert (es war  $\mu=0$  vorausgesetzt). Startpunkt (Pivot) ist der letzte Wert der Zeitreihe

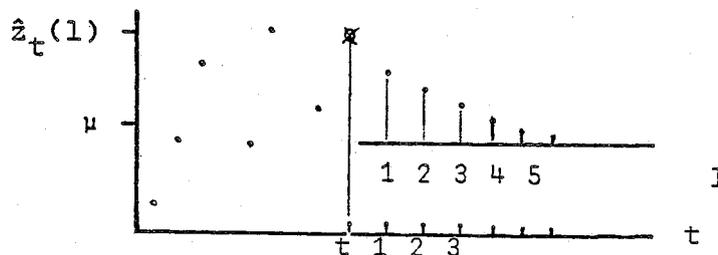


Abb.: Prognosefunktion für den AR(1)-Prozeß ( $\phi > 0$ ).

Für  $\phi < 0$  oszillieren die Prognosen für kurze Zeit (für  $\phi$  nahe -1 etwas länger).

#### Die Varianzfunktion

Für die Berechnung des Prognosefehlers benötigen wir  $z_{t+1}$  in der RS-Form

$$z_{t+1} = a_{t+1} + \phi a_{t+1-1} + \dots + \phi^{l-1} a_{t+1} + \phi^l z_t \quad (5.4.18)$$

Davon (5.4.17) subtrahiert, ergibt

$$\begin{aligned} e_t(1) &= z_t - \hat{z}_t(1) \\ &= a_t + \phi a_{t-1} + \dots + \phi^{l-1} a_{t+1} \end{aligned} \quad (5.4.19)$$

Die Varianz des Prognosefehlers ist

$$\begin{aligned} V(1) &= \text{Var}[e_t(1)] = \sigma_a^2 (1 + \phi^2 + \dots + \phi^{2l-2}) = \\ &= \sigma_a^2 \frac{1 - \phi^{2l}}{1 - \phi^2} \end{aligned} \quad (5.4.20)$$

Bilden wir den Grenzübergang für  $l \rightarrow \infty$

$$\lim_{l \rightarrow \infty} V(1) = \sigma_a^2 \frac{1}{1 - \phi^2} \quad (5.4.21)$$

so erkennt man, daß die Varianz der Prognose eines AR(1)-Prozesses dem konstanten Wert  $\sigma_a^2 (1 - \phi^2)$  zustrebt. Im Unterschied dazu divergiert für  $l \rightarrow \infty$  die Varianz der EFF eines ARIMA(0,1,0)-Prozesses.

#### 5.4.4. AR(2)-Prozeß

Die D-Form für  $z_{t+1}$  lautet

$$(1 - \phi_1 B - \phi_2 B^2) \hat{z}_{t+1} = a_{t+1} \quad (5.4.22)$$

Die Prognosefunktion lautet

$$\begin{aligned} \hat{z}_t(1) &= \phi_1 z_t + \phi_2 z_{t-1} \\ \hat{z}_t(2) &= \phi_1 \hat{z}_t(1) + \phi_2 z_t \\ \hat{z}_t(l) &= \phi_1 \hat{z}_t(l-1) + \phi_2 \hat{z}_t(l-2) \quad l \geq 3 \end{aligned} \quad (5.4.23)$$

Die EFF ist Lösung von

$$(1 - \phi_1 B - \phi_2 B^2) \check{z}_t(1) = 0 \quad l > q - p - d = -2 \quad (5.4.24)$$

bzw.

$$\check{z}_t(1) = b_0^{(t)} \phi_1 + b_1^{(t)} \phi_2 \quad l \geq -1 \quad (5.4.25)$$

Die EFF überdeckt die Prognosefunktion und enthält die beiden letzten Zeitreihenwerte.

Die Pivotprognosen sind

$$\begin{aligned} \hat{z}_t(-1) &= b_0^{(t)} \phi_1^{-1} + b_1^{(t)} \phi_2^{-1} = z_{t-1} \\ \hat{z}_t(0) &= b_0^{(t)} + b_1^{(t)} = z_t \end{aligned} \quad (5.4.26)$$

Die EFF besitzt ein dreifaches Lösungsverhalten

$$\begin{aligned} \hat{z}_t(1) &= b_0^{(t)} \alpha_1^1 + b_1^{(t)} \alpha_2^1 && \alpha_1, \alpha_2 \text{ reell und verschieden} \\ &= (b_0^{(t)} + b_1^{(t)}) \alpha^1 && \alpha_1, \alpha_2 \text{ reell und gleich} \\ &= b_0^{(t)} r^1 \cos(\omega t + b_1^{(t)}) && \alpha_1, \alpha_2 \text{ komplex } (re^{\pm i\omega}) \end{aligned} \quad (5.4.27)$$

Da die EFF wie die ACF derselben Differenzgleichung genügen, so geben die im zulässigen Dreieck (vgl. (3.3.2)) angegebenen verschiedenen ACF auch die verschiedenen Verhalten der Prognosefunktion wieder.

## 5.5. Prognose von ARIMA(p,d,0)-Prozessen

### 5.5.1. Allgemein

Diese Prozesse sind von der Form

$$\phi(B) \nabla^d z_t = a_t \quad (5.5.1.)$$

wobei für den Prozeß

$$w_t = \nabla^d z_t \quad (5.5.2.)$$

die Annahme gilt

$$E w_t = \mu = 0 \quad (5.5.3.)$$

Ist aus der Schätzung ersichtlich, daß diese Annahme nicht zutrifft, gehen wir auf das allgemeinere Modell

$$\phi(B) (\nabla^d z_t - \mu) = a_t \quad (5.5.4.)$$

Über. Das bedeutet, daß neben einem stochastischen Trend (bzw. Polynom) aufgrund des Differenzenoperators  $\nabla^d$  ein zusätzlicher deterministischer Trend (bzw. Polynom) auftritt.

5.5.2. Der ARIMA(1,1,0)-Prozeß

Die D-Form des Prozesses lautet

$$\begin{aligned} (1-\phi B)(\nabla z_t - \mu) &= a_t \\ (1-\phi B)(z_t - z_{t-1} - \mu) &= a_t \\ z_t - z_{t-1} - \mu - \phi z_{t-1} + \phi z_{t-2} + \phi \mu &= a_t \\ z_t &= (1+\phi)z_{t-1} - \phi z_{t-2} + (1-\phi)\mu + a_t \end{aligned} \quad (5.5.5)$$

Zur Berechnung der EFF zerlegen wir das Modell in zwei Teile

$$\begin{aligned} (1-\phi)w_t &= a_t \\ w_t &= \nabla z_t - \mu \end{aligned} \quad (5.5.6)$$

und betrachten zunächst die EFF in den  $w_t$ . Dies ist die EFF eines AR(1)-Prozesses (vgl. (5.4.3)), und sie fällt mit der Prognosefunktion zusammen

$$(1-\phi B)\hat{w}_t(1) = 0 \quad (5.5.7)$$

oder

$$\hat{w}_t(1) = w_t \phi^l \quad l \geq 0 \quad (5.5.8)$$

wobei  $\hat{w}_t(0) = w_t$  die Pivotprognose für den Prozeß  $w_t$  ist. Die EFF für den zweiten Teil des Prozesses lautet

$$\hat{w}_t(1) = \nabla \hat{z}_t(1) - \mu \quad (5.5.9)$$

Setzen wir nun (5.5.8) ein, so erhalten wir

$$\nabla \hat{z}_t(1) - \mu = w_t \phi^l \quad l \geq 0$$

$$\hat{z}_t(1) - \hat{z}_t(1-1) - \mu = \phi^l (z_t - z_{t-1} - \mu) \quad (5.5.10)$$

Daraus erkennt man, daß die prognostizierten Differenzen von ihrem Anfangswert  $z_t - z_{t-1}$  gegen ihren Mittelwert  $\mu$  hin abfallen.

Nun summieren wir die Form (5.4.37) von  $j=1, \dots, l$

$$\sum_{j=1}^l (\nabla \hat{z}_t(1) - \mu) = \sum_{j=1}^l \phi^j (\nabla z_t - \mu) \quad (5.5.11)$$

dann reduziert sich die linke Seite teleskopartig.

$$\hat{z}_t(l) - \hat{z}_t(0) - \mu l = (z_t - z_{t-1} - \mu) \frac{(1 - \phi^l)}{1 - \phi} \quad l \geq 1 \quad (5.5.12)$$

Dabei sind  $\hat{z}_t(0) = z_t$  und  $\hat{z}_t(1)$  die Pivotprognosen.

Für  $l \rightarrow \infty$  nähert sich die EFF asymptotisch der Geraden

$$\hat{z}_t(l) = z_t + \mu l + (z_t - z_{t-1} - \mu) \frac{\phi}{1 - \phi} \quad (5.5.13)$$

wobei  $\mu$  der Anstieg der Geraden ist.

Beispiel 5.4.1.: ARIMA(1,1,0)-Prozeß

Zur Illustration betrachten wir den Prozeß  $(1 - 0,8B)Vz_t = a_t$  mit zwei verschiedenen Mittelwerten

a)  $\mu = 0$  und b)  $\mu = 0,2$

a)  $\mu = 0$

$$\begin{aligned} \hat{z}_t(l) &= z_t + Vz_t \frac{\phi}{1 - \phi} (1 - \phi^l) \\ &= z_t + Vz_t 4 (1 - 0,8^l) \end{aligned}$$

für  $l \rightarrow \infty$  gilt

$$\begin{aligned} \hat{z}_t(l) &= z_t + 4(z_t - z_{t-1}) = \\ &= 5z_t - 4z_{t-1} \end{aligned}$$

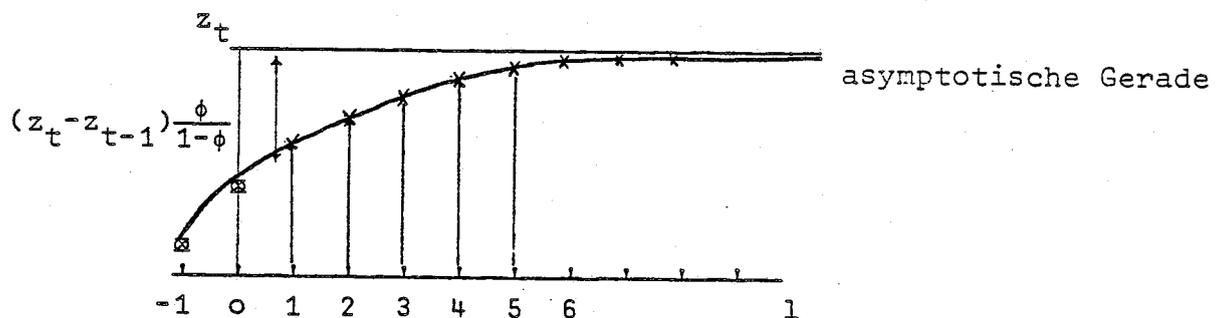


Abb.: Prognosefunktion eines ARI(1,1)-Prozesses mit  $\mu = 0$

b)  $\mu=0,2$

Die Prognosefunktion lautet

$$\hat{z}_t(l) = z_t + 0,2l + (z_t - z_{t-1} - 0,2)4(1-0,8^l) \quad l \geq 1$$

für  $l \rightarrow \infty$  erhalten wir

$$\hat{z}_t(l) = 5z_t - 4z_{t-1} - 0,8 + 0,2l$$

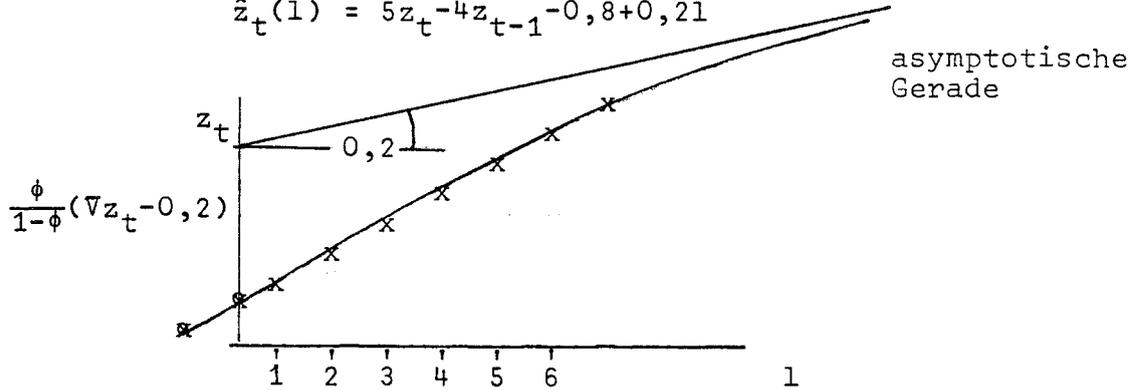


Abb.: Prognosefunktion eines ARI(1,1)-Prozesses mit  $\mu=0,2$

## 5.6. Die Prognose gemischter und IMA-Prozesse

### 5.6.1. Prognose eines ARMA(1,1)-Prozesses

Die D-Form des Prozesses ist (vgl. (3.5.1))

$$(1-\phi B)z_t = (1-\theta B)a_t \quad (5.6.1)$$

und damit lautet die Prognose in D-Form

$$\begin{aligned} \hat{z}_t(l) &= \phi z_t - \theta a_t & l=1 \\ & \phi \hat{z}_t(l-1) & l>1 \end{aligned} \quad (5.6.2)$$

Die EFF ist Lösung von

$$(1-\phi B)\hat{z}_t(l) = 0 \quad (5.6.3)$$

$$\hat{z}_t(l) = b_0^{(t)} \phi^l \quad l > q-p-d=0 \quad (5.6.4)$$

und identisch mit der Prognosefunktion.

Die Pivotprognose ist

$$\hat{z}_t(1) = b_0 \phi = \phi z_t - \theta a_t \quad (5.6.5)$$

Durch den einstufigen Prognosefehler

$$a_t = z_t - \hat{z}_{t-1}(1) = e_t(1) \quad (5.6.6)$$

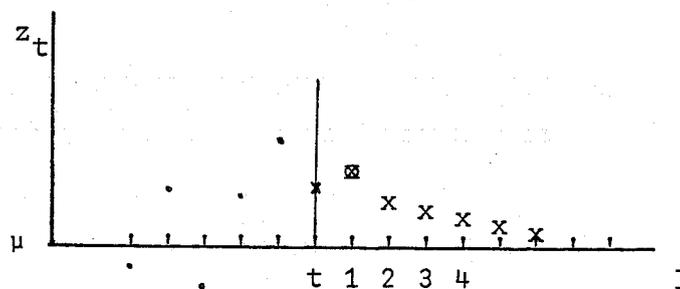
läßt sich die EFF umformen zu

$$\hat{z}_t(1) = \phi^1 [(1-\gamma)z_t + \gamma\hat{z}_{t-1}(1)] \quad (5.6.7)$$

wobei  $\gamma = \theta/\phi$  (5.6.8)

ist. Aus dieser Formersieht man, daß sich die Pivotprognose  $\hat{z}_t(1)$  aus einem gewogenen Durchschnitt der einstufigen Prognose  $\hat{z}_{t-1}(1)$  und dem realisierten Wert  $z_t$  zusammensetzt. Die weiteren Prognosen fallen von diesem Wert exponentiell ab. (Für  $\phi=1$  geht die Prognose in den IMA(1,1)-Prozeß über.)

Das Verhalten der EFF ist dem eines AR(1)-Prozesses ähnlich, nur daß die Prognosewerte erst ab  $\hat{z}_t(1)$  exponentiell abfallen.



### Die Varianzfunktion

Setzen wir die  $\psi_j$ -Koeffizienten (vgl. (4.3.10)) in die Varianzformel (3.5.23) ein, so erhalten wir

$$\begin{aligned} V(1) &= \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 = \\ &= \sigma_a^2 (1 - (\phi - \theta))^2 \left( \frac{1 - \phi^{2l-2}}{1 - \phi^2} \right) \end{aligned} \quad (5.6.9)$$

für  $l \rightarrow \infty$  strebt die Varianzfunktion gegen den Grenzwert

$$V(1) = \sigma_a^2 \frac{1 - 2\phi\theta + \theta^2}{1 - \phi^2} \quad (5.6.10)$$

Die Update-Formel für den ARMA(1,1)-Prozeß lautet

$$\hat{z}_{t+1}(1) = \hat{z}_t(1) + (\phi - \theta)\phi^{l-1} a_{t+1} \quad (5.6.11)$$

5.6.2. Die Prognose des ARIMA(1,1,1)-Prozesses

Die D-Form des Prozesses lautet (vgl. (4.3.1))

$$(1-\phi B)\nabla z_t = (1-\theta B)a_t \quad (5.6.12)$$

und die Prognosefunktion lautet

$$\hat{z}_t(l) = (1+\phi)\hat{z}_t(l-1) - \phi\hat{z}_t(l-2) \quad (5.6.13)$$

Die Pivotprognosen sind

$$\hat{z}_t(0) = z_t \quad (5.6.14)$$

$$\hat{z}_t(1) = (1+\phi)z_t - \phi z_{t-1} - \theta a_t \quad (5.6.15)$$

Die EFF erhalten wir durch Lösung von

$$(1-\phi B)(1-B)\check{z}_t(l) = 0 \quad l > q-p-d = -1 \quad (5.6.16)$$

$$\check{z}_t(l) = b_0^{(t)} + b_1^{(t)}\phi^l \quad l \geq 0 \quad (5.6.17)$$

Die EFF überdeckt die Prognosefunktion  $\hat{z}_t(l)$  und enthält den Wert  $l=0$ .

Bestimmen wir die Konstanten  $b_i^{(t)}$  durch die Pivotprognose, so erhalten wir

$$b_0^{(t)} = z_t + \frac{\phi}{1-\phi}(z_t - z_{t-1}) - \frac{\phi}{1-\phi} a_t \quad (5.6.18)$$

$$b_1^{(t)} = \frac{1}{1-\phi} [\theta a_t - \phi(z_t - z_{t-1})] \quad (5.6.19)$$

Damit lautet die EFF

$$\hat{z}_t(l) = z_t + \phi \frac{(1-\phi^l)}{1-\phi} (z_t - z_{t-1}) - \theta \frac{(1-\phi^l)}{1-\phi} a_t \quad (5.6.20)$$

Für  $l \rightarrow \infty$  strebt die EFF gegen den Grenzwert  $b_0^{(t)}$ .

### 5.6.3. Der IMA(1,1)-Prozeß

Der IMA(1,1)-Prozeß nimmt eine besondere Stellung innerhalb der Prognose der ARIMA-Prozesse ein, da er unabhängig von der Darstellung innerhalb der Klasse der ARIMA-Prozesse als exponentielles Glätten ("exponential smoothing") große praktische Bedeutung erlangt hat.

Die D-Form des Prozesses lautet

$$\nabla z_t = (1-\theta B)a_t \quad (5.6.21)$$

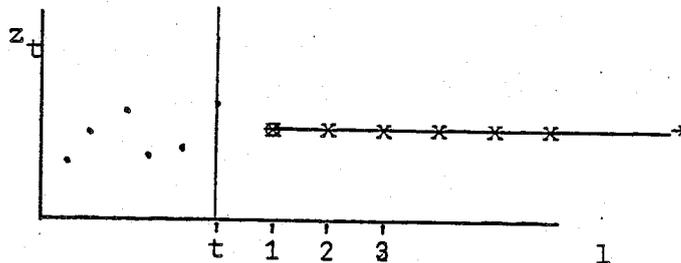
Nehmen wir bedingte Erwartungswerte zum Zeitpunkt  $t+1$ , so erhalten wir die Prognosefunktion

$$\hat{z}_t(l) = \hat{z}_t(l-1) \quad l \geq 1 \quad (5.6.22)$$

mit der Pivotprognose

$$\hat{z}_t(1) = z_t - \theta a_t \quad (5.6.23)$$

Die EFF ist mit der Prognosefunktion identisch und verläuft ausgehend von der Pivotprognose parallel zur Zeitachse.



Die Prognosefunktion für leadzeiten  $l \geq 1$  ist eine sogenannte "no change"-Prognose. Die Bedeutung der Prognose des IMA(1,1)-Prozesses liegt im kurzfristigen Bereich, speziell der einstufigen Prognose. Dazu kann man zwei Prognoseformen ableiten, die sich nur in der Interpretation des Modellparameters  $\lambda = (1-\theta)$  in Bezug auf die Bewertung von Prognosefehler, letzter Prognose und realisiertem Wert unterscheiden.

Ausgangspunkt beider Formen ist der einstufige Prognosefehler

$$\begin{aligned} e_t(1) &= a_{t+1} = \\ &= z_{t+1} - \hat{z}_t(1) \end{aligned} \quad (5.6.24)$$

und die Pivotprognose

$$\hat{z}_t(1) = z_t - \theta a_t \quad (5.6.25)$$

### Prognoseform 1

erhalten wir, indem wir  $a_t$  als den einstufigen Prognosefehler des Prognosezeitpunktes  $t-1$  auffassen und nach  $z_t$  auflösen

$$z_t = \hat{z}_{t-1}(1) + a_t \quad (5.6.26)$$

Setzen wir diesen in die Pivotprognose (5.5.24) ein, so erhalten wir

$$\begin{aligned} \hat{z}_t(1) &= \hat{z}_{t-1}(1) + (1-\theta)a_t \\ &= \hat{z}_{t-1}(1) + \lambda a_t \end{aligned} \quad (5.6.27)$$

wobei

$$\lambda = 1-\theta \quad (5.6.28)$$

der einzige  $\lambda$ -Parameter der  $\nabla$ -Form des IMA(1,1)-Prozesses ist (vgl. (4.4.5)). Wir können daher den Index weglassen. Mit Hilfe der Formel (5.6.27) lassen sich die einstufigen Prognosen wie folgt "fortschreiben": Die neue Prognose  $\hat{z}_t(1)$  erhält man durch Addition der letzten Prognose plus einem mit  $\lambda$  bewerteten Prognosefehler  $e_{t-1}(1) = a_t$ . Der Parameter  $\lambda$  kann somit als Bewertungsfaktor des Prognosefehlers für einstufige Prognosen interpretiert werden. Für einstufige Prognosen ist die Prognoseform 1 äquivalent der Update-Formel für Prognosen (vgl. (5.3.42)).

Prognoseform 2

erhalten wir, indem wir den einstufigen Prognosefehler zum Prognosezeitpunkt  $t-1$

$$a_t = z_t - \hat{z}_{t-1}(1) \tag{5.6.29}$$

direkt in die Pivotprognose einsetzen:

$$\begin{aligned} \hat{z}_t(1) &= (1-\theta)z_t + \theta\hat{z}_{t-1}(1) \\ &= \lambda z_t + (1-\lambda)\hat{z}_{t-1}(1) \end{aligned} \tag{5.6.30}$$

In dieser Form ist die einstufige Prognose ein gewogenes Mittel aus der Prognose zum Zeitpunkt  $t-1$  und der neuen realisierten Beobachtung  $z_t$ . Für die Prognose als Funktion des Parameters  $\lambda$ , bzw.  $\theta$  ergeben sich zwei Extremfälle:

$\lambda \rightarrow 0$  ( $\theta \rightarrow 1$ ) bedeutet, daß die neue Beobachtung nur mit wenig Gewicht in die Prognose eingeht. Man vertraut eher der alten Prognose.

$\lambda \rightarrow 1$  ( $\theta \rightarrow 0$ ) bedeutet, daß die neue Beobachtung mit  $\lambda$ -Prozent in die Prognose eingeht. Der alten Prognose wird zugunsten des neuen Wertes kaum Gewicht beigemessen.

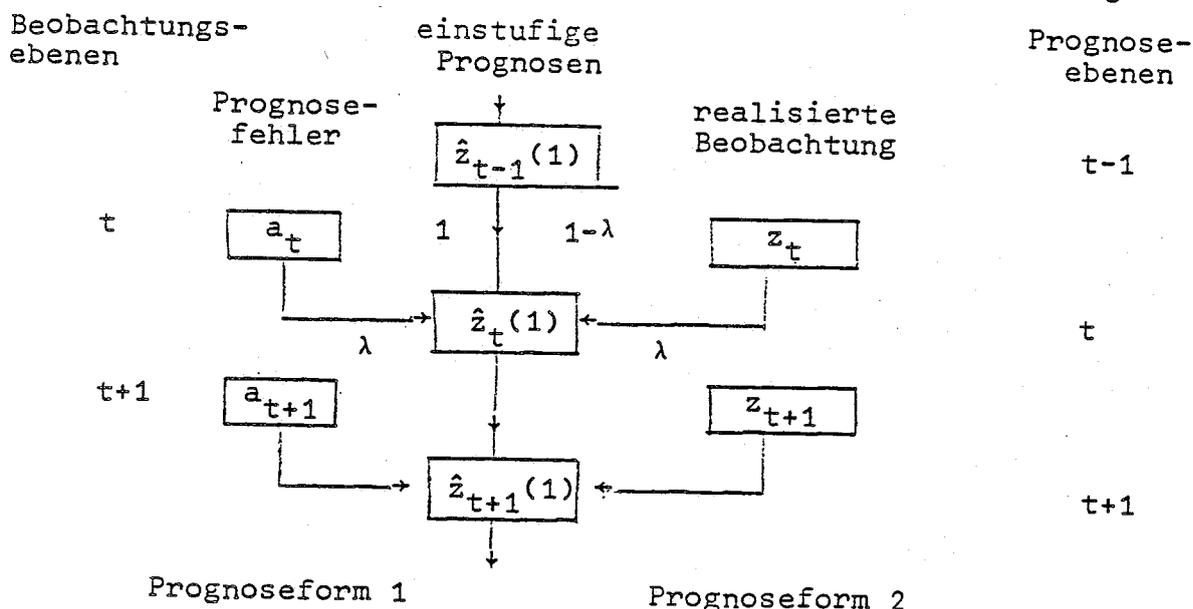


Tabelle: Einstufige Prognosekette für beide Prognoseformen

In obiger Abbildung wird gezeigt, wie man die beiden Prognoseformen miteinander vergleichen kann. Wir unterscheiden in der Kette der einstufigen Prognosen als Horizontalschnitte einerseits Beobachtungsebenen (links) und andererseits Prognoseebenen (rechts). Als Vertikalschnitte der Prognosekette erhalten wir Prognosefehler, Prognosefunktion und Beobachtungen. Prognoseform 1 und Prognoseform 2 verwenden jeweils die linke oder die rechte Vertikale in Kombination mit der Prognosefunktion. An den Kanten sind jeweils die Gewichtungsfaktoren angegeben.

### Prognose in invertierter Form

Wegen der  $\pi$ -Form des Prozesses

$$z_t = \lambda \sum_{j=1}^{\infty} (1-\lambda)^{j-1} z_{t-j} \quad (5.6.31)$$

ist die Prognose in  $\pi$ -Form

$$\hat{z}_t(1) = \lambda \sum_{j=1}^{\infty} (1-\lambda)^{j-1} [z_{t+1-j}] \quad (5.6.32)$$

Für die einstufige Prognose lautet sie

$$\hat{z}_t(1) = \lambda z_t + \lambda(1-\lambda)z_{t-1} + \lambda(1-\lambda)^2 z_{t-2} + \dots \quad (5.6.33)$$

In dieser Form erkennt man an Hand der geometrischen Gewichtungsfaktoren eine Darstellungsform des exponentiellen Glättens.

### Die Varianzfunktion

Da die  $\psi_j = \lambda$ ,  $j=1,2,\dots$  konstant sind, hat die Varianzfunktion die einfache Gestalt

$$\begin{aligned} V(1) &= \sigma_a^2 \sum_{j=0}^{1-1} \psi_j^2 \\ &= \sigma_a^2 (1 + (1-1)\lambda^2) \end{aligned} \quad (5.6.34)$$

Allerdings weist sie für  $1 \rightarrow \infty$  ein explodierendes Verhalten auf.

#### 5.6.4. Der IMA(2,2)-Prozeß

Die D-Form des Prozesses lautet (vgl. (4.5.1))

$$\nabla^2 z_t = (1 - \theta_1 B - \theta_2 B^2) a_t \quad (5.6.35)$$

Nehmen wir bedingte Erwartungswerte zum Zeitpunkt  $z_{t+1}$ , erhalten wir die Prognosefunktion

$$\begin{aligned} \hat{z}_t(1) &= 2z_t - z_{t-1} - \theta_1 a_t - \theta_2 a_{t-2} \\ \hat{z}_t(2) &= 2\hat{z}_t(1) - z_t - \theta_2 a_t \\ \hat{z}_t(1) &= 2\hat{z}_t(1-1) - \hat{z}_t(1-2) \quad l \geq 3 \end{aligned} \quad (5.6.36)$$

Die EFF ist Lösung von

$$(1-B)^2 \check{z}_t(1) = 0 \quad l > p+d-q=0 \quad (5.6.37)$$

$$\check{z}_t(1) = b_0^{(t)} + b_1^{(t)} 1 \quad l > 0 \quad (5.6.38)$$

mit den Pivotprognosen  $\hat{z}_t(1)$  und  $\hat{z}_t(2)$ . Prognosefunktion und EFF fallen bei diesem Modell zusammen.

#### Updaten der Prognose

Die  $\psi_j$ -Koeffizienten in  $\nabla$ -Form lauten (vgl. (4.5.13))

$$\psi_j = \lambda_0 + j\lambda_1 \quad j=1,2,\dots \quad (5.6.39)$$

mit

$$\lambda_0 = 1 + \theta_2 \quad (5.6.40)$$

und

$$\lambda_1 = 1 - \theta_1 - \theta_2 \quad (5.6.41)$$

Die Update-Formel lautet daher

$$\begin{aligned} \hat{z}_{t+1}(1) &= \hat{z}_t(1+1) + \psi_1 a_{t+1} \\ &= \hat{z}_t(1+1) + (\lambda_0 + 1\lambda_1) a_{t+1} \end{aligned} \quad (5.6.42)$$

#### Die Varianzfunktion

Setzen wir die  $\psi_j$  aus (5.6.39) in die Formel für die Varianzfunktion ein, so erhalten wir

$$\begin{aligned} V(1) &= \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 = \\ &= \sigma_a^2 \left( 1 + (l-1)\lambda_0^2 + \frac{1}{6}l(l-1)(2l-1)\lambda_1^2 + \lambda_0\lambda_1 l(l-1) \right) \end{aligned} \quad (5.6.43)$$

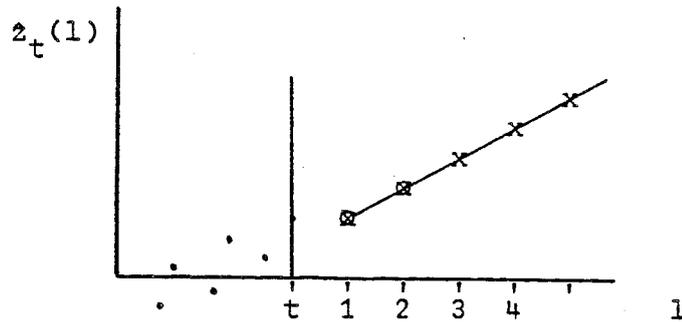


Abb.: Prognosefunktion für den IMA(2,2)-Prozeß

Wie wir in Abschnitt 4.5 gesehen haben, ist der IMA (2,2)-Prozeß zur Darstellung eines stochastischen Trends geeignet. In der Prognosefunktion wird der stochastische Trend zu einer Prognosegeraden, da der stochastische Trend durch Kumulation von zukünftigen Störungen entsteht, die zum Prognosezeitpunkt  $t$  noch nicht bekannt sind. Die Lage der Prognosegeraden wird durch die (adaptiven) Pivotprognosen  $\hat{z}_t(1)$  und  $\hat{z}_t(2)$  bestimmt.

#### 5.6.5. Der IMA(2,3)-Prozeß

Die D-Form lautet

$$V^2 z_t = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3) a_t \quad (5.6.44)$$

für den Zeitpunkt  $t+1$ :

$$z_{t+1} - 2z_{t+1-1} + z_{t+1-2} = a_{t+1} - \theta_1 a_{t+1-1} - \theta_2 a_{t+1-2} - \theta_3 a_{t+1-3} \quad (5.6.45)$$

Durch bedingte Erwartungswerte erhalten wir die Prognosefunktion

$$\begin{aligned}
 \hat{z}_t(1) &= 2z_t - z_{t-1} - \theta_1 a_t - \theta_2 a_{t-1} - \theta_3 a_{t-2} \\
 \hat{z}_t(2) &= 2\hat{z}_t(1) - z_t - \theta_2 a_t - \theta_3 a_{t-1} \\
 \hat{z}_t(3) &= 2\hat{z}_t(2) - \hat{z}_t(1) - \theta_3 a_t \\
 \hat{z}_t(l) &= 2\hat{z}_t(l-1) - \hat{z}_t(l-2) \quad l > 3
 \end{aligned}
 \tag{5.6.46}$$

Die EFF ist für  $l > q - p - d = 3 - 2 = 1$  definiert als

$$\hat{z}(l) = b_0^{(t)} + b_1^{(t)} l \quad l > 1 \tag{5.6.47}$$

Die  $p'=2$  Pivotprognosen sind  $\hat{z}_t(2)$  und  $\hat{z}_t(3)$

Die Gestalt der EFF ist dieselbe wie für den IMA(2,2)-Prozeß, die Prognosegerade wird durch die Werte  $\hat{z}_t(2)$  und  $\hat{z}_t(3)$  festgelegt. Die EFF ist erst für  $l=2$  mit der Prognosefunktion identisch.

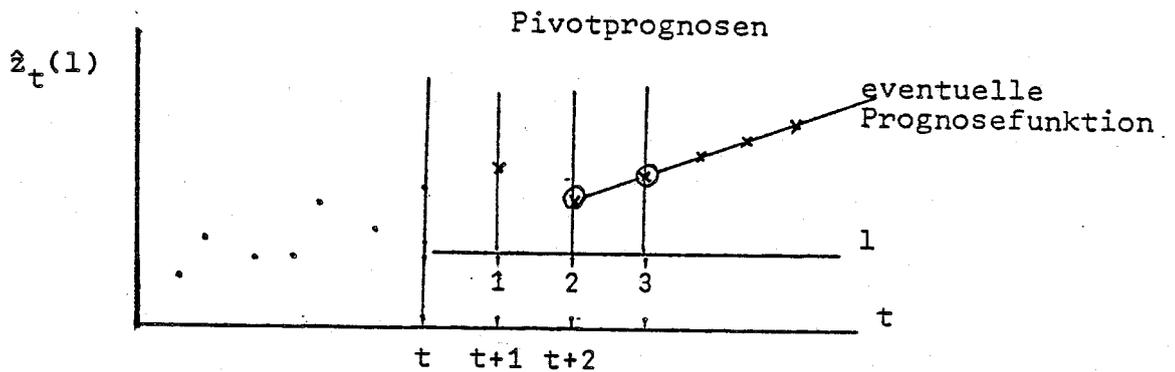


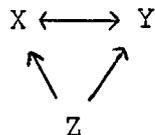
Abb.: Die EFF eines IMA(2,3)-Prozesses

A1 APPENDIX: Partielle Autokorrelationsfunktion

Im folgenden Abschnitt soll eine kurze Einführung in die Idee der partiellen Korrelation und ihre Anwendung als partielle Autokorrelationsfunktion in der Zeitreihenanalyse gegeben werden (vgl. Kendall & Stuart, Vol.II, S.317 ff.). Historisch gesehen wurde das Konzept der partiellen Korrelation für die Normalverteilung entwickelt und wurde dann auf andere Verteilungen übertragen.

A 1.1. Idee der partiellen Korrelation

Seien  $X$  und  $Y$  zwei Zufallsvariable mit gemeinsamer Verteilung  $p(X,Y)$ . Diese sei normalverteilt (bzw. fast normalverteilt) und damit kann man den Korrelationsparameter als Maß für den Zusammenhang zweier Zufallsvariablen definieren. In der Praxis treten oft Schwierigkeiten bei der Interpretation berechneter Korrelationskoeffizienten zweier Zufallsvariablen auf. Wenn zwei Variable miteinander korreliert sind, so kann das deswegen der Fall sein, weil beide Variable jeweils mit einer dritten Variablen  $Z$  zusammenhängen.



Dies kann daran liegen, daß man das Problem des Zusammenhangs zu isoliert gesehen hat und die dritte Variable nicht weiß, oder vergessen hat, oder daß die gefundene Korrelation aus einer multivariaten Untersuchung stammt, wo man bereits Vermutungen hat, daß bestimmte Variable einen bivariaten Zusammenhang mit beeinflussen.

Diese Überlegungen führen zur Frage, ob es nicht möglich ist, ein Korrelationsmaß zu definieren, aus dem der Einfluß einer (oder mehrerer) Variabler ausgeschaltet werden kann.

Für  $n=3$  ist der partielle Korrelationskoeffizient ein Korrelationskoeffizient zwischen  $X$  und  $Y$  bei gegebener Variabler  $Z$  und wird mit  $\rho_{xy.z}$  bezeichnet.

$$\rho_{xy.z} = \frac{\rho_{xz} - \rho_{xy}\rho_{yz}}{(1 - \rho_{xy}^2)(1 - \rho_{yz}^2)} \quad (\text{A 1.1})$$

Der partielle Korrelationskoeffizient  $\rho_{xy.z}$  ist der Korrelationsparameter der bedingten Normalverteilung  $p(X,Y|Z)$  von 3 normalverteilten Zufallsgrößen. Im Gegensatz zu Randverteilungen, wo über den Einfluß der Variablen  $Z$  gemittelt wird, ist der Korrelationskoeffizient der bedingten Verteilung eine Funktion der bedingten Variablen  $Z$ , wobei der Einfluß der Variablen  $Z$  konstant gehalten wird.

#### A 1.2. Interpretation des partiellen Korrelationskoeffizienten

In der Interpretation des partiellen Korrelationskoeffizienten unterscheiden wir drei wichtige Fälle:

(1)  $\rho_{xy} > \rho_{xy.z}$

Ist die normale Korrelation  $\rho_{xy}$  größer als die partielle Korrelation  $\rho_{xy.z}$  (unter Ausschaltung der Variablen  $Z$ ), so kann man daraus schließen, daß der Zusammenhang zwischen  $X$  und  $Y$  zum Teil auf den Einfluß der Variablen  $Z$  auf beide Variablen zurückgeht.

(2)  $\rho_{xy} \gg \rho_{xy.z} \sim 0$

Ist die Korrelation  $\rho_{xy}$  bedeutend größer als die partielle Korrelation  $\rho_{xy.z}$ , die fast Null oder Null ist, so ist der starke Zusammenhang von  $X$  und  $Y$  eine Scheinkorrelation, da der gesamte Zusammenhang von  $X$  und  $Y$  auf die starke Korrelation der dritten Variablen  $Z$  auf jede einzelne Variable  $X$  und  $Y$  zurückzuführen ist.

(3)  $\rho_{xy} < \rho_{xy.z}$

Ist die Korrelation  $\rho_{xy}$  kleiner als die partielle Korrelation  $\rho_{xy.z}$ , so kann man daraus erkennen, daß die dritte Variable Z den Zusammenhang zwischen X und Y gestört hat, bzw. teilweise verdeckt (Z is "masking the correlation"). Man sagt auch, Z "maskiert" die Korrelation von X und Y.

Beispiel A 1.1.

Wir betrachten einige einfache Fälle der paarweisen Korrelationen (CC) dreier Zufallsvariablen und berechnen jeweils die partiellen Korrelationen (PCC)

	CC	Dritte Variable		PCC
	$\rho_{xy}$	$\rho_{xz}$	$\rho_{yz}$	$\rho_{xy.z}$
a)	$\frac{1}{2}$	$\frac{+1}{2}$	$\frac{+1}{2}$	$\frac{1}{3}$
b)	$\frac{1}{2}$	$\frac{+1}{2}$	$\frac{-1}{2}$	1
c)	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	-1
d)	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{3}$
e)	$\frac{1}{2}$	.7	.7	0

Im Fall a) und d) besagt die partielle Korrelation von  $\frac{1}{3}$ , daß jede der drei Variablen gleichen Anteil an der gemeinsamen Korrelation hat.

Im Fall b) und c) ist die dritte Variable jeweils positiv oder negativ mit X und Y korreliert, sodaß der echte Zusammenhang der Variablen X und Y gestört wird.

Im Fall e) ist die Korrelation von 0.5 zwischen X und Y eine Scheinkorrelation, da die dritte Variable (mit nur 0.7) mit jeder Variablen korreliert ist.

Trotzdem muß man auch nach Verwendung von PCC vorsichtig in der Interpretation sein. Auch für den Fall 3 haben wir (statisch gesehen) keine Garantie, daß nun tatsächlich ein starker Zusammenhang von X und Y vorliegt. Dies kann z.B. durch den Einfluß einer weiteren Variable erzeugt sein, die in der Untersuchung übersehen worden war.

Generell kann man sagen: Für die Interpretation der partiellen Korrelation gilt dasselbe wie für den gewöhnlichen Korrelationskoeffizienten: Die Annahme der Kausalität in einer Beziehung zweier Variabler ist immer außer-statistischen Überlegungen überlassen.

### A 1.3. Partielle Autokorrelation

Gegeben sei die Zeitreihe

$$\dots z_t, z_{t+1}, z_{t+2}, \dots \quad t \in Z \quad (\text{A 1.2})$$

Wir berechnen die Autokorrelationen

$$\rho_k = \text{Corr}(z_t, z_{t+k}) = \frac{\text{Cov}(z_t, z_{t+k})}{\text{Var}(z_t)} \quad (\text{A 1.3})$$

Nun bezeichnen wir mit

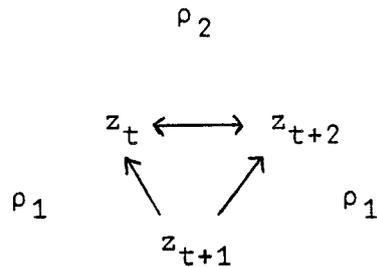
$$\rho_1 = \text{Corr}(z_t, z_{t+1}) = \rho_{xy} \quad (\text{A 1.4})$$

$$\rho_2 = \text{Corr}(z_t, z_{t+2}) = \rho_{xz} \quad (\text{A 1.5})$$

$$\rho_1 = \text{Corr}(z_{t+1}, z_{t+2}) = \rho_{yz} \quad (\text{A 1.6})$$

und stellen uns folgendes Problem:

Zeitreihen sind dadurch gekennzeichnet, daß aufeinanderfolgende Beobachtungen stark miteinander korrelieren. Ist aber  $z_t$  mit  $z_{t+2}$  nur deswegen stark korreliert, weil  $z_{t+2}$  mit  $z_{t+1}$  korreliert ist und andererseits  $z_{t+1}$  mit  $z_t$  korreliert ist?



Mit Hilfe des partiellen Korrelationskoeffizienten mißt man die Korrelation von  $z_t$  mit  $z_{t+2}$ , indem man den Einfluß der Variablen  $z_{t+1}$  konstant hält:

$$\begin{aligned} \rho_{xy.z} &= \frac{\rho_{xz} - \rho_{xy}\rho_{yz}}{\sqrt{(1-\rho_{xy}^2)(1-\rho_{yz}^2)}} \\ &= \frac{\rho_2 - \rho_1^2}{1-\rho_1^2} = \rho_{t,t+2.t+1} \end{aligned} \quad (\text{A 1.7})$$

Der partielle Korrelationskoeffizient beantwortet also die Frage, ob die Autokorrelation zum lag 2 durch die Autokorrelation zum lag 1 erklärt werden kann.

#### A 1.4. Multivariate Normalverteilung

Bevor wir den PCC ableiten, benötigen wir einige Resultate der multivariaten Normalverteilung.

##### Def. A 1.1.: Reguläre Normalverteilung

Sei  $X = (X_1, \dots, X_p)$  ein ( $p$ -dimensionaler) normalverteilter Zufallsvektor. Die gemeinsame Dichte ist

$$p(x) = \frac{1}{(2\pi)^{p/2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)' \Sigma^{-1}(x-\mu)\right) \quad (\text{A 1.8})$$

wobei

$$\mu = E(X) \quad (\text{A 1.9})$$

$$\Sigma = \text{Cov}(X) \quad (\text{A 1.10})$$

der Erwartungswert  $\mu$  und die Kovarianzmatrix  $\Sigma$  ist. Die Verteilung sei regulär, d.h.  $|\Sigma| \neq 0$ . Wir schreiben auch kurz

$$X \sim N(\mu, \Sigma) \quad (\text{A 1.11})$$

Def. A 1.2.: Kofaktor

Die Kofaktoren einer quadratischen Matrix  $\Sigma=(p \times p)$  lauten

$$\Sigma_{ij} = (-1)^{i+j} |\Sigma_{ij}| \quad (\text{A 1.12})$$

wobei  $|\Sigma_{ij}|$  die Determinante der adjungierten Matrix ist. Die adjungierte Matrix  $\Sigma_{ij}$  erhält man durch Streichen der  $i$ -ten Zeile und  $j$ -ten Spalte (ist also vom Typ  $\Sigma_{ij}=(p-1 \times p-1)$ ).

Anmerkung: Die Elemente der inversen Matrix  $\Sigma^{-1} = (\sigma^{ij})$  erhält man durch

$$\sigma^{ij} = \frac{\Sigma_{ji}}{|\Sigma|} \quad (\text{A 1.13})$$

bzw

$$\Sigma^{-1} = \frac{1}{|\Sigma|} \Sigma_{\text{adj}} \quad (\text{A 1.14})$$

wobei

$$\Sigma_{\text{adj}} = \begin{pmatrix} \Sigma_{11} & \dots & \Sigma_{n1} \\ \Sigma_{1n} & \dots & \Sigma_{nn} \end{pmatrix} \quad (\text{A 1.15})$$

Satz A 1.1.: Bedingte Normalverteilung

Sei  $X = N(\mu, \Sigma)$  regulär normalverteilt. Man partitioniere  $X$  und  $\Sigma$  in

$$X = (X_1, X_2) \quad (\text{A 1.16})$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{22} & \Sigma_{21} \end{pmatrix} \quad (\text{A 1.17})$$

Das ist die bedingte Verteilung von  $X_1/X_2$  wieder normalverteilt mit der Dichte

$$p(x_1 | X_2=x_2) \sim N(\mu_{1.2}, \Sigma_{11.2}) \quad (\text{A 1.18})$$

wobei

$$\mu_{1.2} = \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (x_2 - \mu_2) \quad (\text{A 1.19})$$

$$\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}' \quad (\text{A 1.20})$$

Beweis: Siehe Anderson (1958)

Für die Ableitung des PCC verwenden wir die standardisierte Normalverteilung.

Satz A 1.2.: Standardisierte Normalverteilung

Standardisieren wir in einer regulären Normalverteilung jede Komponente analog

$$Z_i = \frac{X_i - \mu_i}{\sigma_i} \quad i=1, \dots, p \quad (\text{A 1.21})$$

so ist  $Z=(Z_1, \dots, Z_p)$  standardisiert normalverteilt mit

$$Z \sim N(\theta, C) \quad (\text{A 1.22})$$

wobei  $\theta=(0, \dots, 0)$  der Nullvektor und C die Korrelationsmatrix der Zufallsvariable ist

$$C = \begin{pmatrix} 1 & \rho_{12} \dots & \rho_{1p} \\ & 1 & \vdots \\ & & \ddots \\ ./\cdot & & & 1 \end{pmatrix} \quad (\text{A 1.23})$$

A 1.5. Die Ableitung des PCC

Im Unterschied zum multiplen Korrelationskoeffizienten ist der PCC wie der gewöhnliche CC ein Maß für den Zusammenhang zweier Zufallsvariablen. Jedoch bezieht der PCC andere Zufallsvariable in Form von bedingten Zufallsvariablen mit in die Analyse ein.

Für den PCC wird die Information von p-dimensionalen Verteilungen immer auf zwei Variablen komprimiert und davon der gewöhnliche CC berechnet. Dazu zeigen wir, daß der gewöhnliche CC als Funktion der Kofaktoren einer zweidimensionalen Kovarianzmatrix erklärt werden kann.

Sei  $\underline{X}=(X_1, X_2)$  ein zweidimensionaler Zufallsvektor. Die Kovarianzmatrix  $\Sigma$  lautet

$$\underline{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \quad (\text{A 1.24})$$

Der CC ist definiert als

$$\begin{aligned} \rho &= \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} \\ &= \frac{\sigma_{21}}{\sqrt{\sigma_{11}\sigma_{22}}} \\ &= \frac{\underline{\Sigma}_{12}}{\sqrt{\underline{\Sigma}_{22}\underline{\Sigma}_{11}}} \end{aligned} \quad (\text{A 1.25})$$

wobei die  $\underline{\Sigma}_{11} \dots \underline{\Sigma}_{22}$  die Kofaktoren der Kovarianzmatrix  $\Sigma$  sind.

### Beispiel A 1.2.

Berechnung von  $\rho_{12,3}$  für drei Zufallsvariable

Sei  $X=(x_1, x_2, x_3)$  ein dreidimensionaler normalverteilter Zufallsvektor mit Korrelationsmatrix

$$\underline{C} = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ & 1 & \rho_{23} \\ ./\cdot & & 1 \end{pmatrix} \quad (\text{A 1.26})$$

Wir partitionieren in  $X=(X_{(1)} X_{(2)})$  mit

$$\begin{aligned} \underline{X}_{(1)} &= (X_1, X_2) \\ \underline{X}_{(2)} &= (X_3) \end{aligned} \quad (\text{A 1.27})$$

und

$$\underline{\Sigma}_{11} = \begin{pmatrix} 1 & \rho_{12} \\ ./\cdot & 1 \end{pmatrix}$$

$$\underline{\Sigma}_{12} = \begin{pmatrix} \rho_{12} \\ \rho_{23} \end{pmatrix} = \underline{\Sigma}'_{21} \quad (\text{A 1.28})$$

$$\underline{\Sigma}_{22} = (1)$$

Wir berechnen die bedingte Kovarianzmatrix

$$\begin{aligned} \underline{\Sigma}_{11.2} &= \underline{\Sigma}_{11} - \underline{\Sigma}_{12} \cdot \underline{\Sigma}_{22}^{-1} \underline{\Sigma}_{21} = \\ &= \begin{pmatrix} 1 & \rho_{12} \\ ./\cdot & 1 \end{pmatrix} - \begin{pmatrix} \rho_{12} \\ \rho_{23} \end{pmatrix} (1) (\rho_{12} \rho_{23}) = \\ &= \begin{pmatrix} 1 - \rho_{13}^2 & \rho_{12} - \rho_{13} \rho_{23} \\ ./\cdot & 1 - \rho_{23}^2 \end{pmatrix} \\ &= \begin{pmatrix} \sigma_{11.3} & \sigma_{12.3} \\ ./\cdot & \sigma_{22.3} \end{pmatrix} \quad (\text{A 1.29}) \end{aligned}$$

Die bedingte Kovarianzmatrix  $\underline{\Sigma}_{11.2}$  ist die Kovarianzmatrix einer zweidimensionalen Verteilung mit Korrelationsparameter  $\rho_{12.3}$ , der nach Formel (A 1.25) aus  $\underline{\Sigma}_{11.2}$  berechnet werden kann

$$\begin{aligned} \rho_{12.3} &= \frac{\sigma_{12.3}}{\sqrt{\sigma_{11.3} \sigma_{22.3}}} \\ &= \frac{\rho_{12} - \rho_{13} \rho_{23}}{\sqrt{(1 - \rho_{13}^2)(1 - \rho_{23}^2)}} \quad (\text{A 1.30}) \end{aligned}$$

Auf dasselbe Resultat kommt man, wenn man aus der Matrix C die Kofaktoren bildet.  $\underline{\Sigma}_{11.2}$  ist dann gerade die Matrix der Kofaktoren

$$\begin{aligned} \underline{C}_{11} &= (-1)^2 \begin{vmatrix} 1 & \rho_{23} \\ \rho_{23} & 1 \end{vmatrix} = 1 - \rho_{23}^2 \\ \underline{C}_{22} &= (-1)^4 \begin{vmatrix} 1 & \rho_{13} \\ \rho_{13} & 1 \end{vmatrix} = 1 - \rho_{13}^2 \quad (\text{A 1.31}) \end{aligned}$$

$$c_{12} = (1-\rho_{13}^2) \begin{vmatrix} \rho_{12} & \rho_{23} \\ \rho_{13} & 1 \end{vmatrix} = -(\rho_{12} - \rho_{13}\rho_{23})$$

Diese Kofaktoren werden nun als Varianzen der bedingten Variablen interpretiert und wir definieren als Korrelationskoeffizienten der bedingten Variablen

$$\begin{aligned} \rho_{12.3} &= - \frac{c_{12}}{\sqrt{c_{11}c_{22}}} \\ &= \frac{\rho_{12} - \rho_{13}\rho_{23}}{\sqrt{(1-\rho_{13}^2)(1-\rho_{23}^2)}} \end{aligned} \quad (\text{A 1.32})$$

Dies legt nun nahe, folgende allgemeine Definition für den PCC einzuführen:

Def. A 1.3.: Partieller Korrelationskoeffizient

Sei  $\underline{X} = (X_1, \dots, X_p)$  eine p-dimensionale reguläre Normalverteilung. Fixieren wir (p-s) Variablen, d.h. partitionieren wir  $\underline{X}_{(1)} = (X_1, X_2)$  und  $\underline{X}_{(2)} = (X_3, \dots, X_p)$ , so lautet der zugehörige partielle Korrelationskoeffizient

$$\rho_{12.34\dots p} = \frac{-c_{12}}{\sqrt{c_{11}c_{22}}} \quad (\text{A 1.33})$$

wobei die  $c_{ij}$  die Kofaktoren von  $\rho_{ij}$  der Korrelationsmatrix C (vgl. (A 1.23)) der standardisierten Zufallsvariablen sind.

Schließlich ist es naheliegend, die Formel (A 1.33) als PCC für beliebig p-dimensional verteilte Zufallsvariable zu verwenden.

A 1.6. Die PACF

Die in der Zeitreihenanalyse wichtige PACF  $\phi_{kk}$  ist eine Folge von speziellen partiellen Korrelationskoeffizienten. Die Zufallsvariablen sind in diesem Fall die um  $k=0,1,2,\dots$  verschobene Zeitreihe

$$x_k = \{z_{t+k}\} \quad t=1,\dots,n-k \quad (\text{A 1.34})$$

Damit erhalten wir folgende Zuordnung

$$\begin{aligned} \phi_{11} &= \rho_{12} \\ \phi_{22} &= \rho_{13.2} \\ \phi_{33} &= \rho_{14.23} \\ &\vdots \\ \phi_{kk} &= \rho_{1k+1.23\dots k} \end{aligned} \quad (\text{A 1.35})$$

Für die Berechnung der PACF ist eine Folge von Autokorrelationsmatrizen  $P_k$  zu bilden

$$P_{k+1} = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{k-1} & \rho_k \\ & 1 & & & \vdots \\ & & \cdot & & \rho_1 \\ & & & \cdot & \\ & \cdot & & & 1 \end{pmatrix} \quad (\text{A 1.36})$$

und die zugehörige PACF lautet

$$\begin{aligned} \phi_{kk} &= \rho_{1k+1.23\dots k} = \\ &= \frac{-C_{1k+1}}{\sqrt{C_{11}C_{k+1,k+1}}} \end{aligned} \quad (\text{A 1.37})$$

Man beachte, daß die Kofaktoren im Falle der Autokorrelationsmatrizen eine spezielle Struktur aufweisen

$$C_{11} = |P_k| \quad (\text{A 1.38})$$

$$C_{k+1,k+1} = |P_k| \quad (\text{A 1.39})$$

Damit lautet

$$\phi_{kk} = \frac{-C_{1,k+1}}{|P_k|} \quad (\text{A 1.40})$$

$C_{1,k+1}$  läßt sich als partitionierte Matrix von  $P_k$  und  $\rho_k$  umfassen

$$C_{1,k+1} = \begin{pmatrix} 1 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_{k-2} & \rho_2 \\ \vdots & & \cdot & \vdots \\ \rho_{k-1} & \dots & \rho_1 & \rho_k \end{pmatrix}$$

$$= -[P_k \ \rho] \quad (\text{A 1.41})$$

wobei  $[P_k \ \rho]$  bedeuten soll, daß die letzten Spalten der Matrix  $P_k$  durch den Vektor  $\rho_k$  ersetzt wird. Damit lautet die PACF

$$\phi_{kk} = \frac{\det[P_{ki} \ \rho_k]}{\det(P_k)} \quad (\text{A 1.42})$$

Damit kann man  $\phi_{kk}$  als Lösung (mittels der Cramer'schen Regel) des Gleichungssystems

$$P_k \phi_k = \rho_k \quad (\text{A 1.43})$$

auffassen, mit

$$\begin{aligned} \phi_k' &= (\phi_{11}, \dots, \phi_{kk}) \\ \rho_k' &= (\rho_1, \dots, \rho_k) \end{aligned} \quad (\text{A 1.44})$$

und der Autokorrelationsmatrix  $P_k$

$$P_k = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-2} \\ \vdots & & \ddots & \vdots \\ \rho_{k-1} & \dots & \rho_1 & 1 \end{pmatrix} \quad (\text{A 1.45})$$

Die wichtigste Anwendung der PACF liegt im folgenden

Satz A 1.3.

Als PACF eines AR(p)-Prozesses erhält man die p autoregressiven Parameter  $\phi_1, \dots, \phi_p$ , d.h.

$$\phi_{kk} = \phi_k \quad k=1, \dots, p \quad (\text{A 1.46})$$

Beweis:

Für  $k=p$  ist das Gleichungssystem (A 1.43) äquivalent mit den Yule-Walker-Gleichungen (vgl. (3.2.33))

$$\underline{P}_p \underline{\phi} = \underline{\rho}_p \quad (\text{A 1.47})$$

mit

$$\underline{\phi} = (\phi_1, \dots, \phi_p)$$

Damit gilt

$$\underline{\phi}_k = \underline{\phi} \quad (\text{A 1.48})$$

APPENDIX A 2:

Momente von nichtlinearen Funktionen von Zufallsgrößen

In vielen praktischen statistischen Problemen ist es notwendig, nichtlineare Funktionen von Zufallsvariablen (und -vektoren) zu betrachten. Außer in Spezialfällen ist es allgemein nicht möglich, die exakte Verteilung dieser nichtlinearen Funktionen abzuleiten, und daher möchte man die Verteilung durch ihre Momente beschreiben. Im folgenden zeigen wir, wie man approximativ den Mittelwert und die Varianz einer nichtlinearen Funktion einer Zufallsgröße bestimmen kann.

A 2.1. Mittelwert einer nichtlinearen Funktion

Gegeben sei ein n-dimensionaler Zufallsvektor  $\underline{X}=(X_1, \dots, X_n)$  mit Mittelwertsvektor  $\underline{\mu}=(\mu_1, \dots, \mu_n)$  und Kovarianzmatrix  $\underline{\Sigma}=(\sigma_{ij})$   $i, j=1, \dots, n$ . Es sei  $g(X_1, \dots, X_n)$  eine nichtlineare Funktion des Zufallsvektors. Diese linearisieren wir durch eine Taylor-Reihe um den Punkt  $a=(a_1, \dots, a_n)$ . Die Approximation erster Ordnung ist

$$g(X_1, X_2, \dots, X_n) \sim g(a_1, \dots, a_n) + \sum_{i=1}^n \left( \frac{\partial g}{\partial X_i} \right)_a (x_i - a_i) \quad (\text{A 2.1})$$

wobei die  $\left( \frac{\partial g}{\partial X_i} \right)_a$  die partiellen Ableitungen von  $g$  nach  $x_i$  an der Stelle  $(X_1, \dots, X_n)=(a_1, \dots, a_n)$  sind. Nehmen wir Erwartungswerte in (A 2.1)

$$E(g(X_1, \dots, X_n)) \sim g(a_1, \dots, a_n) + \sum_{i=1}^n \left( \frac{\partial g}{\partial X_i} \right)_a (E(X_i) - a_i) \quad (\text{A 2.2})$$

und setzen wir  $a_i = \mu_i$ , so verschwindet der zweite Ausdruck und wir erhalten als Approximation erster Ordnung für den Mittelwert einer nichtlinearen Funktion

$$E(g(X_1, \dots, X_n)) \sim g(\mu_1, \dots, \mu_n) \quad (\text{A 2.3})$$

A 2.2. Varianz einer nichtlinearen Funktion

Wegen (A.2.1) ist

$$g(X_1, \dots, X_n) - g(\mu_1, \dots, \mu_n) \sim \sum_{i=1}^n \left( \frac{\partial g}{\partial X_i} \right)_\mu (X_i - \mu_i) \quad (\text{A 2.4})$$

d.h. die approximativen Abweichungen sind approximativ linear in den  $X_i$ .

Die Varianz von Linearformen ist das Quadrat von (A 2.4)

$$\text{Var}[g(X_1, \dots, X_n)] \approx \sum_{i=1}^n \sum_{j=1}^n \left( \frac{\partial g}{\partial X_i} \right)_\mu \left( \frac{\partial g}{\partial X_j} \right)_\mu \text{Cov}(X_i, X_j) \quad (\text{A 2.5})$$

Beispiel A 2.1.:  $g(X) = X^2$

Für  $n=1$  erhalten wir als approximative Varianz

$$\text{Var}(g(X)) \sim \left( \frac{\partial g}{\partial X} \right)_\mu^2 \text{Var}X$$

und wegen

$$\left( \frac{dX^2}{dX} \right)_\mu = 2\mu$$

ist

$$\text{Var}X^2 \sim 4\mu^2 \text{Var}X$$

Beispiel A 2.1.:  $g(X_1, X_2) = \sqrt{X_1^2 + X_2^2}$

Für  $n=2$  erhalten wir

$$\text{Var} g(X_1, X_2) \sim \left( \frac{\partial g}{\partial X_1} \right)_\mu^2 \text{Var}X_1 + \left( \frac{\partial g}{\partial X_2} \right)_\mu^2 \text{Var}X_2 + 2 \left( \frac{\partial g}{\partial X_1} \right)_\mu \left( \frac{\partial g}{\partial X_2} \right)_\mu \text{Cov}(X_1, X_2)$$

Zur weiteren Vereinfachung nehmen wir an, daß  $X_1$  und  $X_2$  unkorreliert sind ( $\text{Cov}(X_1, X_2) = 0$ ), dann ist

$$\left( \frac{\partial g}{\partial X_i} \right)_\mu = (X_1^2 + X_2^2)^{-\frac{1}{2}} X_i$$

$$\text{Var} \sqrt{X_1^2 + X_2^2} \sim \frac{1}{\mu_1^2 + \mu_2^2} (\mu_1^2 \text{Var}X_1 + \mu_2^2 \text{Var}X_2)$$

LITERATURVERZEICHNIS  
-----

- ANDERSON, O.D. (1976) Time Series Analysis and Forecasting.  
The Box-Jenkins-Approach. Butterworth, London
- ANDERSON, R.L. (1942) Distribution of the Serial Correlation  
Coefficient, Ann.math.Statist., 13, 1-13
- ANDERSON, T.W. (1958) An Introduction to Multivariate Analysis.  
Wiley, New York
- ANDERSON, T.W. (1971) The Statistical Analysis of Time Series.  
Wiley, New York
- BARTLETT, M.S. (1946) On the Theoretical Specification of  
Sampling Properties of Autocorrelated Time Series.  
J.R.Stat.Soc. B 8, 27-41
- BARTLETT, M.S. (1966) Stochastic Processes, 2nd edn.,  
Cambridge Univ. Press
- BARTLETT, M.S. (1967) Some Remarks on the Analysis of Time Series,  
Biometrika 54, 25-38
- BATES, J.M. and GRANGER, C.W.J. (1969) The Combination of  
Forecasts, Op.Res.Quart., 20, 451-468
- BERGSTROM, A.R. (1976) Statistical Inference in Continuous  
Time Economic Models. North Holland Publ.Co.,  
Amsterdam
- BOX, G.E.P. and COX, D.R. (1964) An Analysis of Transformations,  
J.R.Statist.Soc., B 26, 211-252
- BOX, G.E.P. and JENKINS, G.M. (1967) Models for Forecasting  
Seasonal and Nonseasonal Time Series, in  
HANNAN (1967), 271-311
- BOX, G.E.P. and JENKINS, G.M. (1970) Time-Series Analysis,  
Forecasting and Control. Holden-Day, San Francisco
- BRAND, L. (1966) Differential and Difference Equations.  
John Wiley, New York

- BROWN, R.G. (1963) Smoothing, Forecasting and Prediction.  
Prentice-Hall, Englewood Cliffs, N.J.
- BROWN, R.G. and MEYER, R.F. (1961) The Fundamental Theorem  
of Exponential Smoothing, Operations Res.9,673 ff.
- CHATFIELD, C. (1965) The Analysis of Time Series. Theory and  
Practice. Chapman & Hall, London.
- CHATFIELD, C. and NEWBOLD, P. (1974) Adaptive Filtering,  
Op.Res.Quart. 25, 494-495.
- CHATFIELD, C. and PROTHERO, D.L. (1973a) Box Jenkins Seasonal  
Forecasting: Problems in a Case Study, J.R.Stat.  
Soc.A, 136, 295-336.
- CHATFIELD, C. and PROTHERO, D.L. (1973b) A Reply to some  
Comments by Box and Jenkins, J.R.Stat.Soc.A,  
136, 345-352.
- CLEVELAND, W.S. (1972) The Inverse Autocorrelations of a Time  
Series and Their Applications, Technometrics, Vol.14  
No.2, 277-298.
- COX, D.R. and MILLER, H.D. (1968) The Theory of Stochastic  
Processes. Wiley, New York.
- CRAMER, H. and LEADBETTER, M.R. (1967) Stationary and Related  
Stochastic Processes. Wiley, New York.
- DAVIS, H.T. (1941) The Analysis of Economic Time Series.  
The Principia Press, Bloomington.
- DHRYMES, P.J. (1971) Distributed Lags. Holden-Day, San Francisco.
- DHRYMES, P.J. (1974) Econometrics. Springer, New York.
- DRAPER, N.R. and SMITH, H. (1966) Applied Regression Analysis.  
Wiley, New York.
- DURBIN, J. (1960) The Fitting of Time Series Models, Rev.Int.  
Inst.Stat. 28, 233.
- DURBIN, J. (1969) Tests for Serial Correlation in Regression  
Analysis Based on the Periodogram of Least Squares  
Residuals, Biometrika 56, 1-16.

- DURBIN, J. (1970) Testing For Serial Correlation in Least Squares Regression When Some of the Regressors Are Lagged Dependent Variables, *Econometrica* 38, 410-421
- DURBIN, J. and WATSON, G.S. (1950;1951;1971) Testing for Serial Correlation in Least Squares Regression, I, II, III, *Biometrika* 37, 409-428; 38, 159-178; 58, 1-19
- FISHMAN, G.S. (1969) *Special Methods in Econometrics*. Harvard Univ.Press, Cambridge, Mass.
- GOLDBERG, S. (1968) *Differenzengleichung und ihre Anwendung in Wirtschaftswissenschaft, Psychologie und Soziologie*. Oldenburg, München
- GRANGER, C.W.J. (1962) The Effect of Varying Month-Lengths on the Analysis of Economic Time Series. Research Memorandum 41, Princeton University
- GRANGER, C.W.J. (1966) The Typical Shape of an Econometric Variable, *Econometrica* 34, 150-161
- GRANGER, C.W.J. and HATANAKA, M. (1964) *Special Analysis of Economic Time Series*. Princeton Univ.Press, Princeton
- GRANGER, C.W.J. and MORGENSTERN, O. (1963) Spectral Analysis of New York Stock Market Prices, *Kyklos* Vol. 16, 1-27
- GRANGER, C.W.J. and NEWBOLD, P. (1974) Spurious Regressions in Econometrics, *J. of Econometrics* 2, 111-120
- GRANGER, C.W.J. and NEWBOLD, P. (1977) *Forecasting Economic Time Series*. Academic Press, New York
- GRENANDER, U. and ROSENBLATT, M. (1957) *Statistical Analysis of Stationary Time Series*. Wiley, New York
- HANNAN, E.J. (1960) *Time Series Analysis*. Methuen, London
- HANNAN, E.J. (1970) *Multiple Time Series*. Wiley, New York
- HARRIS, B. (1967) *Spectral Analysis of Time Series*. Wiley, New York
- HARRISON, P.J. (1965) Short Term Sales Forecasting, *J.R.Stat.Soc.C* 14, 102-139

- HARRISON, P.J. and STEVENS, C.F. (1971) A Bayesian Approach to Short-Term Forecasting. *Op.Res.Quart.* 22, 341-362
- HAUSE, J.C. (1971) Spectral Analysis and the Detection of Lead-Lag Relations, *Amer.Econ.Rev.* 61, 213-217
- JENKINS, G.M. (1954) Tests of Hypotheses in the Linear Autoregressive Model, *Biometrika* 41, 405 ff.
- JENKINS, G.M. (1956) Tests of Hypotheses in the Linear Autoregressive Model, *Biometrika* 43, 186 ff.
- JENKINS, G.M. and WATTS, D.G. (1968) *Spectral Analysis and Its Applications*. Holden-Day, San Francisco
- KENDALL, M.G. (1945) On the Analysis of Oscillatory Time Series, *J.R.Statist.Soc.* 108, 93 ff.
- KENDALL, M.G. (1973) *Time-Series*. Griffin, London
- KENDALL, M.G. and STUART, A. (1966) *The Advanced Theory of Statistics*, Vol. 3, Griffin, London
- KENKEL, J.L. (1974) *Dynamic Linear Economic Models*. Gordon and Breach, New York
- LEDOLTER, J. (1977) Adaptivity and Stability of Time Series Models, *Empirica* 2, 179-195
- LEDOLTER, J. (1976) Condition for the Optimality of Exponential Smoothing Forecast Procedure. IIASA, Laxenburg/Austria
- LEDOLTER, J., SCHEBECK, F. and THURY, G. (1977) Box-Jenkins-Methode. Alternative Verfahren zur Prognose ökonomischer Zeitreihen. *Empirica* 1
- PALM, F. (1977) On Efficient Estimation of the Final Equation Form of a Linear Multiple Time Series Process. Paper, presented on ESEM 1977, Vienna
- PARZEN, E. (1962) *Stochastic Processes*. Holden-Day, San Francisco
- POLASEK, W. (1975) *Statistische Modelle der Zyklenmessung*. Dissertation, Universität Wien

- POLASEK, W. (1977) Das IAZ-System (Interaktives Zeitreihensystem).  
Institutsarbeit Nr. 87, Institut für Höhere Studien,  
Wien
- QUENOUILLE, M.H. (1949) Approximate Tests of Correlation in  
Time Series, J.R.Statist.Soc. B 11, 68 ff.
- QUENOUILLE, M.H. (1957) The Analysis of Multiple Time Series.  
Griffin, London
- RAO, C.R. (1965) Linear Statistical Inference and its Application.  
Wiley, New York
- REID, D.J. (1972) A Comparison of Forecasting Techniques on  
Economic Time-Series, Proc. of the 1971 conference  
on Forecasting in Action, organised by the O.R.Soc.  
and the Soc. for Long Range Planning.
- REINHARDT, F. und SOEDER, H. (1974) dtv-Atlas Mathematik I, II.  
Deutscher Taschenbuchverlag, München
- ROSANOW, J.A. (1975) Stochastische Prozesse. Akademie-Verlag, Berlin
- SCHÄFFER, K.-A. (1976) Beiträge zur Zeitreihenanalyse.  
Vandenhoeck & Ruprecht, Göttingen
- SCHMETTERER, L. (1966) Einführung in die Mathematische Statistik.  
2. Aufl., Springer, Wien
- SCHÖNFELD, P. (1969) Methode der Ökonometrie I, II. Franz Vahlen,  
Berlin
- STOPP, F. (1976) Operatorenrechnung. Teubner, Leipzig
- TINTNER, G. (1940) The Variate Difference Method.  
Bloomington Press, Indiana
- WETZEL, W. (1970) Neuere Entwicklungen auf dem Gebiet der Zeit-  
reihenanalyse. Vandenhoeck & Ruprecht, Göttingen
- WIENER, N. (1949) Extrapolation, Interpolation and Smoothing of  
Stationary Time-Series. M.I.T.Press, Cambridge, Mass.
- WITTING, H. (1966) Mathematische Statistik. Teubner, Stuttgart
- WOLD, H.O. (1965) Bibliography on Time-Series and Stochastic  
Processes. Oliver & Boyd, Edinburgh
- YAGLOM, A.M. (1962) An Introduction to the Theory of Stationary  
Random Functions. Prentice-Hall, Englewood Cliffs, N.J.

VERZEICHNIS DER VERWENDETEN ABKÜRZUNGEN UND SYMBOLE

---

ACF	Autokorrelationsfunktion
ACovF	Autokovarianzfunktion
AR(p)	Autoregressiver Prozeß der Ordnung p
ARI(p,d)	Autoregressiver Integrierter Prozeß der Ordnung p und d
ARIMA(p,d,q)	Autoregressiver integrierter Moving Average Prozeß der Ordnung p,d,q
ARMA(p,q)	Autoregressive Moving Average Prozesse der Ordnung p und q
$a_t$	Zufällige Störung (Random Shock)
$[a_t]$	Bedingter Erwartungswert der $a_t$
$\alpha_i$	Wurzel des ARI-Operators
B	Backwardshiftoperator
BJ	Box-Jenkins
$b_i^{(t)}$	Adaptive Koeffizienten der EFF
$\beta_i$	Wurzel des MA-Operators
C	Komplexe Zahlen
CC	Korrelationskoeffizient
Corr(X,Y)	Korrelation der Zufallsvariablen X und Y
$C(t-k)$	Homogene Lösung
$c_k$	Geschätzte Autokovarianzfunktion
D-Form	Differenzgleichungsform
d	Ordnung des Differenzenoperators (Anzahl der Differenzen)
det	Determinante
$\nabla$	Differenzenoperator
$\nabla$ -Form	Differenzenoperatorform
E	Erwartungswert
EFF	Eventuelle Prognosefunktion
$E_t$	Bedingter Erwartungswert
$e_t^{(l)}$	Prognosefehler der leadzeit l
F	Forwardshiftoperator
F(X)	Verteilungsfunktion der Zufallsvariable X
F(X Y)	Bedingte Verteilung von X, gegeben Y

$\phi(B)$	AR-Operatorpolynom
$\phi_i$	Autoregressive Koeffizienten
$\underline{\phi}_k$	Vektor der PACF
$\phi_{kk}$	Partieller Autokorrelationskoeffizient der Ordnung k
$\psi(B)$	Verallgemeinertes AR-Operatorpolynom
$\psi_i$	Verallgemeinerte autoregressive Koeffizienten
$\Gamma_n$	Autokovarianzmatrix der Ordnung n
$\gamma_{za}$	Kovarianzfunktion von $z_t$ und $a_t$
$\gamma_k$	Autokovarianzfunktion
$I(t-k)$	Partikuläre Lösung
$IMA(d,q)$	Integrierter Moving Average Prozeß der Ordnung d und q
L	Länge des Prognosezeitraumes
l	leadzeit
$\lambda$	Koeffizienten der V-Form
$MA(q)$	Moving Average Prozeß der Ordnung q
$\mu$	Mittelwert der Zeitreihe
$\underline{\mu}_{1.2}$	Mittelwertsvektor der bedingten Verteilung
N	Natürliche Zahlen
$N(\mu, \Sigma)$	Multivariate Normalverteilung mit Mittelwertsvektor $\mu$ und Kovarianzmatrix $\Sigma$
$N(0,1)$	Normierte Normalverteilung
PACF	Partielle Autokorrelationsfunktion
PCC	Partieller Korrelationskoeffizient
$P(A)$	Wahrscheinlichkeit des Ereignisses A
$P_k$	Autokorrelationsmatrix der Ordnung k
$ \underline{P}:\underline{p} $	Partitionierte Matrix (Matrix P, Vektor p)
P	Ordnung des AR-Operators (Anzahl der $\phi_i$ )
$p(X)$	Wahrscheinlichkeitsdichte von X
$p(X Y)$	Bedingte Dichte von X, gegeben Y
$P(x_1, \dots, x_n)$	Wahrscheinlichkeitsdichte des Zufallsvektors $(x_1, \dots, x_n)$
$p'=p+d$	Ordnung des ARI-Operators
$\psi(B)$	Transferfunktion
$\psi_i$	Lineare Filterkoeffizienten
$\pi$ -Form	Invertierte Form
$\pi(B)$	Inverses Polynom, Operatorpolynom der $\pi$ -Form
$\pi_i$	Koeffizienten der $\pi$ -Form

$q$	Ordnung des MA-Operators (Anzahl der $\theta_i$ )
$R$	Reelle Zahlen
$RS$	Random Shock
$r_k$	Geschätzte Autokorrelation der Ordnung $k$
$\rho_k$	Autokorrelation der Ordnung $k$
$\underline{\rho}(k)$	Vektor der ACF
$\rho_{xy}$	Korrelationskoeffizient der Zufallsvariablen $X$ und $Y$
$\rho_{12.3}$	Partieller Korrelationskoeffizient
$S$	Summationsoperator
$SP(f)$	Spektrum zur Frequenz $f = \frac{1}{T}$
$\underline{\Sigma}_{adj}$	Adjungierte Matrix von $\underline{\Sigma}$
$\underline{\Sigma}_{11.2}$	Kovarianzmatrix der bedingten Verteilung
$\sigma^{ij}$	Elemente der inversen Matrix von $\Sigma$
$\sigma_a^2$	Varianz der $a_t$
$\sigma_z^2$	Varianz der $z_t$
$\sigma_{12}$	Kovarianz der Zufallsvariablen $X_1$ und $X_2$
$T$	Indexmenge
$T$	Periodenlänge
$TRS$	Truncated Random Shock
$t(\alpha)$	$\alpha\%$ Grenze der Normalverteilung
$\theta(B)$	MA-Operatorpolynom
$\theta_i$	Moving Average Koeffizienten
$V(l)$	Varianzfunktion der leadzeit $l$
$Var(X)$	Varianz der Zufallsvariable $X$
$\{x_t\}$	Stochastischer Prozeß
$z_t$	Zeitreihe
$\{z_t\}$	Bedingter Erwartungswert der $z_t$
$z_t^C$	Homogene Lösung
$z_t^I$	Partikuläre Lösung
$\bar{z}_t(\pi)$	Normierte $\pi$ -Form
$\hat{z}_t(l)$	Prognosefunktion
$\check{z}_t(l)$	Eventuelle Prognosefunktion
$\mathcal{O}$	Nullvektor