

Hansmann
Kurzlehrbuch Prognoseverfahren

GABLER Kurzlehrbücher

mit Aufgaben und Lösungen

Herausgeber: Prof. Dr. Dr. h.c. Herbert Jacob, Univ. Hamburg

Prof. Dr. Dietrich Adam, Univ. Münster

Kurzlehrbuch Planung

Prof. Dr. Karl-Werner Hansmann, Hochsch. d. Bundesw. Hamburg

Kurzlehrbuch Prognoseverfahren

Prof. Dr. Wolfgang Hilke, Univ. Freiburg i. Br.

Kurzlehrbuch Bilanzpolitik

Prof. Dr. Dr. h.c. Herbert Jacob, Univ. Hamburg

Kurzlehrbuch Investitionsrechnung

Prof. Dr. Dieter Pressmar, Univ. Hamburg

Kurzlehrbuch Produktionstheorie

9/20.253x7

840202

Prof. Dr. Karl-Werner Hansmann



ZENTRALBIBLIOTHEK
des Fachbereichs Sozial-
wissenschaften der Universität
Hamburg und der Hochschule
für Wirtschaft und Politik

Kurzlehrbuch Prognoseverfahren

Mit Aufgaben und Lösungen

GABLER

CIP-Kurztitelaufnahme der Deutschen Bibliothek

Hansmann, Karl-Werner:
Kurzlehrbuch Prognoseverfahren: mit Aufga-
ben u. Lösungen/Karl-Werner Hansmann. –
Wiesbaden: Gabler, 1983.
(Gablers Kurzlehrbücher)
ISBN 3-409-13444-1

© Betriebswirtschaftlicher Verlag Dr. Th. Gabler GmbH, Wiesbaden 1983
Umschlaggestaltung: Horst Koblitz, Wiesbaden
Satz: Satzstudio RES, R.-E. Schulz, Dreieich
Druck und Buchbinderei: Wilhelm & Adam, Heusenstamm
Alle Rechte vorbehalten. Auch die fotomechanische Vervielfältigung des Werkes
(Fotokopie, Mikroskopie) oder von Teilen daraus bedarf der vorherigen Zustimmung
des Verlages.
Printed in Germany.

ISBN 3 409 13444 1

Vorwort

Jede betriebswirtschaftliche Entscheidung beruht auf Daten, deren zukünftige Entwicklung prognostiziert werden muß. Die Prognose ist daher zentraler Bestandteil der Entscheidungsvorbereitung. In diesem Lehrbuch wird versucht, Studierende und Praktiker mit den Prognoseverfahren vertraut zu machen, die in Betriebswirtschaftslehre und Unternehmenspraxis Verwendung finden und in ihrer Gesamtheit das Gebiet der ökonomischen Prognoserechnung möglichst weitgehend abdecken.

Dabei werden drei Ziele verfolgt:

- Das Buch soll Mittler sein zwischen den Lesern, die sich in die Prognoserechnung einarbeiten wollen und den oft schwer verständlichen Originalarbeiten über einzelne Prognoseverfahren.
- Der Leser soll vertiefte Kenntnisse der Methoden erlangen und sie an Beispielen nachvollziehen können.
- Die Fertigkeit zur konkreten Anwendung von Prognoseverfahren soll durch Aufgaben eingeübt werden.

Dem ersten Ziel dient der weitgehende Verzicht auf mathematische Beweisführungen, dem zweiten eine detaillierte Darstellung insbesondere der anspruchsvollen Prognoseverfahren, kombiniert mit empirischen Anwendungsbeispielen, die alle mit dem am Lehrstuhl des Verfassers entwickelten computergestützten Prognosesystem durchgerechnet wurden.

Ein vertieftes theoretisches Verständnis ist jedoch für den Einsatz von Prognoseverfahren noch nicht hinreichend. Ihre Anwendung muß eingeübt werden. Diesem Ziel dient der umfangreiche Übungsteil mit Aufgaben und Lösungen zu den einzelnen Kapiteln des Lehrbuchs. Die Aufgaben können mit einem Taschenrechner gelöst werden, so daß der Einsatz von Prognose-Software nicht erforderlich ist. Grundkenntnisse in Statistik und (für Kapitel G.) in Matrizendarstellung, wie sie im Grundstudium der Wirtschaftswissenschaften erworben werden, reichen zur Durcharbeitung des Lehrbuchs aus.

Mein herzlicher Dank gilt meinen Mitarbeitern am Institut für Industriebetriebsforschung der Hochschule der Bundeswehr Hamburg, den Herren Dipl.-Kfm. Gunter Paetow, Dr. Ulrich Raubach, Dipl.-Kfm. Andreas Roggon und Dipl.-Kfm. Wolfgang Zetsche für ihre vielfältige Hilfe, insbesondere beim Entwurf der Aufgaben und Lösungen, bei der Erstellung von Prognose-Software, beim Durchrechnen der Anwendungsbeispiele und bei den Zeichnungen. Auch ihre kritischen Anregungen und wertvollen Hinweise habe ich dankbar entgegengenommen.

Herrn Prof. Dr. Dr. h.c. H. Jacob danke ich für die Anregung zu diesem Buch und für die Aufnahme des Werks in die von ihm herausgegebene Schriftenreihe betriebswirtschaftlicher Kurzlehrbücher.

Nicht zuletzt danke ich ganz besonders Frau Elfriede Walsemann für ihren unermüdlichen Einsatz und ihre mühevollen Arbeit beim Fertigstellen des druckfertigen Manuskripts.

KARL-WERNER HANSMANN

Inhaltsverzeichnis

	Seite
Vorwort	5
A. Grundlagen der Prognoserechnung	11
I. Prognosebegriff und Prognoseproblematik	11
II. Aufbau und Arten von Prognosemodellen	12
1. Qualitative und quantitative Prognosemodelle	12
2. Univariate und multivariate Prognosemodelle	12
3. Kurz- und langfristige Prognosemodelle	12
4. Weitere Klassifikationen	13
III. Der zeitliche Ablauf einer Prognose	13
IV. Beurteilungsmaße für Prognosen	14
1. Ex-ante-Beurteilung	14
2. Ex-post-Beurteilung	14
a) Die mittlere absolute Abweichung	15
b) Die mittlere quadratische Abweichung bzw. die Wurzel aus der mittleren quadratischen Abweichung	15
c) Der Ungleichheitskoeffizient von Theil	15
d) Fehlermaße zur qualitativen Beurteilung von Prognosen	16
B. Heuristische Prognoseverfahren	18
I. Überblick	18
II. Die Szenario-Technik	18
1. Kennzeichnung	18
2. Anwendungsbereiche	19
3. Anwendungsbeispiel: Energiebedarf der BRD für die Jahre 2000 und 2030	19
4. Kritik	21
III. Die Delphi-Methode	22
1. Grundprinzipien	22
2. Ablauf der Delphi-Prognose	22
3. Anwendungsbeispiel: Prognose der Veränderung des Aktien- kursindex des Statistischen Bundesamtes	23
4. Kritik des Verfahrens	25
C. Prognosen auf der Grundlage der exponentiellen Glättung	27
I. Gleitende Durchschnitte	27

II.	Das konstante Modell der exponentiellen Glättung	28
1.	Grundgedanken	28
2.	Die Zielfunktion des Verfahrens	29
3.	Die Bedeutung des Glättungsparameters α	31
4.	Anwendung des Modells auf den Verbrauch an Spirituosen in der BRD	31
III.	Das Trendmodell der exponentiellen Glättung	33
1.	Zielfunktion und Prognosegleichung	33
2.	Die Prognose des Verbrauchs von Dieselkraftstoff in der BRD	36
IV.	Die dynamische Anpassung des Glättungsparameters α	39
1.	Die Methoden von Chow und Smith	40
2.	Die Prognose des Dieselkraftstoffverbrauchs mit dynamischer Anpassung von α	41
V.	Erweiterung der exponentiellen Glättung	43
1.	Exponentielle Glättung höherer Ordnung	43
2.	Mehr-Parameter-Modelle	44
VI.	Kritische Zusammenfassung	44
D.	Prognosen bei saisonbehafteten Zeitreihen	46
I.	Überblick	46
II.	Das Saisonverfahren von Winters	47
1.	Die exponentielle Glättung als Basis der Parameterschätzung	47
2.	Die Berechnung des Saisonfaktors	48
3.	Die Prognose der monatlichen Übernachtungen im Reiseverkehr	49
4.	Kritische Zusammenfassung	52
III.	Anwendung der Spektralanalyse	53
1.	Stationäre stochastische Prozesse	53
2.	Ermittlung von Korrelogramm und Spektrum	54
3.	Die Lag-Fenster	58
4.	Spektralanalyse der Auflage einer Publikumszeitschrift	59
5.	Kritische Zusammenfassung	61
E.	Prognosen mit autoregressiven Methoden	63
I.	Grundgedanken der Zeitreihenanalyse	63
II.	Das Box-Jenkins-Verfahren (ARIMA-Modelle)	65
1.	Autoregressive Prozesse (AR-Modelle)	65
2.	Moving-average-Prozesse (MA-Modelle)	69
3.	Nicht-stationäre Prozesse	72
4.	Identifizierung des geeigneten ARIMA-Modells	74
a)	Kennzeichnung von ARIMA-Modellen	74
b)	Identifizierung des passenden Modelltyps	75

5. Schätzung der Parameter	81
a) Anfangsschätzungen	81
b) Das nicht-lineare Hauptschätzverfahren	83
6. Prüfung des Prognosemodells	88
7. Aktienkursprognose mit einem ARIMA-Modell	90
8. Kritik am Box-Jenkins-Verfahren	96
III. Das Verfahren des adaptiven Filterns	98
1. Beziehungen zum Box-Jenkins-Verfahren	98
2. Die dynamischen Anpassungsgleichungen	98
3. Steuerung der Schrittweite durch eine Lernkonstante	100
4. Aktienkursprognose durch adaptives Filtern	102
5. Kritische Zusammenfassung	103
F. Prognosen mit Wachstums- und Sättigungsmodellen	104
I. Langfristige Trendextrapolation	104
II. Die Exponentialfunktion	108
1. Prognosegleichung und Parameterschätzung	108
2. Anwendungsbeispiel: Prognose der Weltbevölkerung	109
III. Die logistische Funktion	111
1. Ableitung der logistischen Anpassungsgleichung	111
2. Schätzung der Parameter	113
3. Anwendungsbeispiel: Prognose der Ausstattung deutscher Privathaushalte mit Farbfernsehern	114
IV. Die Gompertz-Funktion	117
1. Ableitung der Anpassungsgleichung	117
2. Schätzung der Parameter	117
3. Anwendungsbeispiele	
a) Ausstattung privater Haushalte mit Waschmaschinen	118
b) PKW je 1000 Erwachsene Wohnbevölkerung	120
V. Kritische Zusammenfassung	123
G. Prognosen mit der multiplen Regressionsanalyse	125
I. Grundgedanken kausaler Prognosemethoden	125
II. Das multiple Regressionsmodell	126
1. Grundannahmen	126
2. Schätzung der Regressionsparameter	128
3. Prüfkriterien für die Regressionsschätzung	130
a) Das Bestimmtheitsmaß	130
b) Parametertests für den Schätzvektor	132
4. Autokorrelation und Multikollinearität	135
5. Die Prognosegleichungen	137

III. Prognose des Aktienkursindex mit einem multiplen Regressionsmodell	138
IV. Kritische Zusammenfassung	140
H. Kriterien zur Auswahl geeigneter Prognoseverfahren	141
 Aufgaben und Lösungen	
Zu Kapitel A.: Grundlagen der Prognoserechnung – Aufgaben 1–4	145
Zu Kapitel B.: Heuristische Prognoseverfahren – Aufgaben 5–7	146
Zu Kapitel C.: Prognosen auf der Grundlage der exponentiellen Glättung – Aufgaben 8–12	149
Zu Kapitel D.: Prognosen bei saisonbehafteten Zeitreihen – Aufgaben 13–16	157
Zu Kapitel E.: Prognosen mit autoregressiven Methoden – Aufgaben 17–21 .	159
Zu Kapitel E.III.: Prognosen mit dem adaptiven Filtern – Aufgabe 22–23	163
Zu Kapitel F.: Prognosen mit Wachstums- und Sättigungsmodellen – Aufgaben 24–26	166
Zu Kapitel G.: Prognosen mit der multiplen Regressionsanalyse – Aufgaben 27–29	171
<i>Literaturverzeichnis</i>	177
<i>Sachregister</i>	180

A. Grundlagen der Prognoserechnung

I. Prognosebegriff und Prognoseproblematik

Prognosen werden in vielen Wissenschaften erstellt, so daß es nicht erstaunlich ist, in der wissenschaftlichen Literatur eine beträchtliche Reihe verschiedener Prognosedefinitionen zu finden. Gemeinsam ist das Bemühen, aus der Analyse vergangener Ereignisse Gesetzmäßigkeiten zu gewinnen, die einen Schluß auf zukünftige Ereignisse ermöglichen.

Im Anschluß an Theil (1966) soll im folgenden eine Definition zugrunde gelegt werden, die der Zielsetzung dieses Lehrbuchs in besonderem Maße entspricht.

Eine Prognose ist eine Aussage über ein oder mehrere zukünftige Ereignisse, die

1. auf Beobachtungen

2. auf einer Theorie

beruht.

Mit der ersten Bedingung wird sichergestellt, daß jede Prognose auf der Analyse der Vergangenheit basiert, d. h. empirisch fundiert sein muß und kein bloßes „Tippen“ sein darf.

Die zweite Bedingung erfordert eine sachlogische *Begründung* der Prognose und die Angabe von *Prämissen*, unter denen sie abgegeben wird; die zugrunde liegende Theorie muß nicht unbedingt umfangreich oder sehr detailliert ausgearbeitet sein.

Die in diesem Lehrbuch behandelten Prognoseprobleme sind ausschließlich *ökonomischer*, d. h. betriebswirtschaftlicher bzw. gesamtwirtschaftlicher Natur. Die erstellten Prognosen sind daher *Aussagen* über zukünftige ökonomische Ereignisse. Die *Beobachtungen* der Vergangenheit liegen in der Form von *Zeitreihen* vor, und die Theorie bzw. sachlogische Begründung ist in das wirtschaftswissenschaftliche Theoriensystem eingebettet.

Unter einer *Zeitreihe* verstehen wir eine Menge von Beobachtungswerten, die in gleichem zeitlichen Abstand aufeinander folgen. Sie wird symbolisch wie folgt dargestellt:

$$\{x_t\} \quad (t = 1, \dots, T)$$

wobei x_t den zum Zeitpunkt t (oder im Zeitraum t) beobachteten Wert der Zeitreihe angibt.

Die sorgfältige Analyse einer Zeitreihe soll anschließend eine Prognose des künftigen Verhaltens dieser Reihe ermöglichen. Eine solche Prognose ist jedoch nur sinnvoll, wenn unterstellt werden kann, daß die im bisherigen Verlauf der Zeitreihe aufgedeckten Gesetzmäßigkeiten für die Zukunft weiter gültig sind.

Diese Prämisse wird *ZEITSTABILITÄTSHYPOTHESE* genannt. Streng genommen ist sie in der ökonomischen Realität nie vollständig erfüllt, was eigentlich jede Prognose

verbietet. Da aber ökonomische Prognosen als Grundlage für unternehmenspolitische und wirtschaftliche Entscheidungen notwendig sind, hilft nur der pragmatische Standpunkt weiter, Prognosen dann zu erstellen, wenn man Grund zu der Annahme hat, die Zeitstabilitätshypothese sei *im wesentlichen* erfüllt. Man nimmt somit von vornherein Prognosefehler in Kauf, versucht sie jedoch mit Hilfe geeigneter Verfahren zu minimieren.

II. Aufbau und Arten von Prognosemodellen

Ein *Prognosemodell* ist ein System, das die beobachteten Werte der zu prognostizierenden Variablen und evtl. anderer Variablen als Elemente enthält und nach bestimmten Regeln miteinander verknüpft, um als Ergebnis der Verknüpfung Prognosewerte zu erhalten.

Die Vielfalt verschiedener Prognosemodelle läßt sich durch Klassifizierung auf eine überschaubare Zahl von Modelltypen reduzieren, von denen im folgenden die wichtigsten gegenübergestellt werden.

1. Qualitative und quantitative Prognosemodelle

Erfolgt die Verknüpfung mit Hilfe mathematischer Operationen (Addition usw.) in Gestalt eines Gleichungssystems, so liegt ein quantitatives Modell vor. Quantitative Modelle stehen im Zentrum dieses Lehrbuchs (Kapitel C. bis G.).

Qualitative Prognosemodelle verknüpfen die Variablen verbalargumentativ und werden z. B. zur Vorhersage politischer Entwicklungstendenzen verwendet. Sie sind Gegenstand von Kapitel B.

2. Univariate und multivariate Prognosemodelle

Univariate Prognosemodelle beziehen nur die Zeitreihe der zu prognostizierenden Variablen in die Untersuchung ein. Die Zeit ist hier der einzige „erklärende“ Faktor für die Prognose. Die in den Kapiteln C. bis F. dargestellten Verfahren benutzen univariate Modelle, während die Regressionsanalyse (Kapitel G.) als multivariates Verfahren die zu prognostizierende Variable mit anderen Variablen verknüpft, die aufgrund einer theoretischen Begründung als *kausale* (verursachende) Variablen angesehen werden.

3. Kurz- und langfristige Prognosemodelle

Bei dieser Unterscheidung dient die Länge des *Prognosezeitraums* als Kriterium. Leider besteht in der Literatur keine Einigkeit über die Abgrenzung der Fristen. Als Inhalts-

punkt sei hier die Abgrenzung von Makridakis/Wheelwright (1978a, S. 542) wiedergegeben:

<i>Prognosemodell</i>	<i>Prognosezeitraum</i>
kurzfristig	bis 3 Monate
mittelfristig	3 Monate bis 2 Jahre
langfristig	über 2 Jahre

Es sei jedoch darauf hingewiesen, daß jeder Abgrenzung eine gewisse Willkür anhaftet, und darüber hinaus die Länge des Prognosezeitraums auch von dem zeitlichen Abstand der Beobachtungswerte (Wochen, Monate, Quartale, Jahre) abhängen kann.

4. Weitere Klassifikationen

In quantitativen multivariaten Prognosemodellen können eine, aber auch mehrere interdependente Variablen gleichzeitig prognostiziert werden. Da man zur Prognose einer einzigen Variablen eine Prognosegleichung benötigt, heißen diese Modelle Ein-Gleichungsmodelle, die übrigen *Mehr-Gleichungsmodelle*. Beide Modelltypen können *lineare* oder *nicht-lineare* Gleichungen enthalten. Bevorzugtes Anwendungsgebiet von Mehr-Gleichungsmodellen sind gesamtwirtschaftliche Prognosen des Bruttosozialprodukts, des privaten Konsums, der privaten Investitionen usw. Sie stützen sich in der Regel auf Regressionsmodelle aus dem Instrumentarium der Ökonometrie.

III. Der zeitliche Ablauf einer Prognose

Es empfiehlt sich allgemein, den Prognosevorgang in einzelne zeitlich aufeinander folgende Schritte zu zerlegen, wobei der Zerlegungsgrad von der Bedeutung der Prognose für das Unternehmen bestimmt wird. Im folgenden sei ein Ablaufschema von Brockhoff (1977, S. 49) wiedergegeben, das sich durch eine sehr feingliedrige Unterteilung des Prognoseablaufs auszeichnet, die aus Zeit- und Kostengründen auch vergrößert werden kann.

1. Darstellung des Prognosegegenstandes
2. Formulierung eines Erklärungsmodells des Prognosegegenstandes (theoretische Hypothesenbildung)
3. Untersuchung der Zeitstabilitätshypothese (Eignet sich das Erklärungsmodell als Prognosemodell?)
4. Prüfung der Meßbarkeit der Variablen
5. Suche nach meßbaren Daten (Zeitreihen)
6. Test des Erklärungsmodells
7. Anwendung eines Prognoseverfahrens
8. Kritische Beurteilung der Prognoseergebnisse

9. Einsatz der Ergebnisse in der Planung

10. Ex-post Kontrolle der Prognoseergebnisse zur Gewinnung neuer Hypothesen.

Die Punkte 9. und 10. werden häufig nicht mehr zur eigentlichen Prognoseaufgabe gerechnet, sie sollten jedoch auf jeden Fall zur Wahrung des Systemzusammenhangs und zur Rückkoppelung mit der Prognoseaufgabe verbunden werden.

IV. Beurteilungsmaße für Prognosen

Jede Prognose soll in irgendeiner Form die Grundlage für eine *Entscheidung* liefern. Zur Beurteilung der Prognose sind daher folgende Punkte wichtig:

1. Bedeutung der Entscheidung,
2. Prognosekosten,
3. Prognosequalität.

Dabei ist die Bedeutung der Entscheidung Bestimmungsfaktor für die Prognosekosten und -qualität. Zu den Prognosekosten gehören Kosten für die Entwicklung und EDV-Implementierung der Prognoseverfahren, Dokumentation, Rechenzeit, Speicherbedarf sowie Datenbeschaffung und -aufbereitung.

Im folgenden beschränken wir uns auf die Beurteilung der *Prognosequalität*, doch darf man in der Praxis die beiden anderen Punkte nicht vernachlässigen. Die Prognosequalität läßt sich auf zwei Arten beurteilen:

1. Ex-ante-Beurteilung

Hierbei wird die Prognose überprüft, *bevor* beobachtete Werte im Prognosezeitraum vorliegen. Die Beurteilung erstreckt sich z. B. auf die Auswahl der kausalen Variablen, die zugrunde gelegte Funktionsform (Linearität ist in der Realität nicht gewährleistet), die Länge des Beobachtungszeitraumes sowie eventuelle nachhaltige Veränderungen der Zeitreihenstruktur im Beobachtungszeitraum.

Es liegt jedoch auf der Hand, daß eine *abschließende* Beurteilung der Prognosequalität nur durch Vergleich der prognostizierten mit den tatsächlich eintreffenden Werten – also im nachhinein – möglich ist.

2. Ex-post-Beurteilung

Die Prognosequalität wird ex post überwiegend statistisch gemessen, wobei der Prognosefehler im Mittelpunkt steht. Der Prognosefehler ist die Differenz von tatsächlichem Wert x_t und prognostiziertem Wert \hat{x}_t und sei im folgenden mit e_t bezeichnet.

$$\text{Prognosefehler: } e_t = x_t - \hat{x}_t \quad (t = 1, \dots, T) \\ \text{(Prognosezeitraum)}$$

Aus dieser Kennzahl lassen sich weitere Größen ableiten, deren wichtigste die folgenden sind:

– Absoluter Prognosefehler	$ e_t = x_t - \hat{x}_t $
– Quadratischer Prognosefehler	$e_t^2 = (x_t - \hat{x}_t)^2$
– Relativer absoluter Prognosefehler	$\frac{ e_t }{x_t} = \frac{ x_t - \hat{x}_t }{x_t}$

Diese Prognosefehler benutzt man zur Konstruktion von statistischen Fehlermaßen, um damit die Prognosegüte zu beurteilen. Schwarze (1980, S. 317 ff.) hat die gebräuchlichsten Maße zusammengestellt und kommentiert. Von den dort zu findenden Maßen (mehr als 30) werden wir in diesem Buch nur vier quantitative und ein qualitatives Fehlermaß vorstellen und nachstehend näher erläutern.

a) Die mittlere absolute Abweichung (MAA)

$$\text{MAA} = \frac{1}{T} \sum_{t=1}^T |x_t - \hat{x}_t| \quad T = \text{Länge des Prognosezeitraums}$$

Dieses Fehlermaß verhindert das Saldieren positiver und negativer Abweichungen im Prognosezeitraum und gewichtet alle Prognosefehler gleich.

b) Die mittlere quadratische Abweichung (MQA) bzw. die Wurzel aus der mittleren quadratischen Abweichung (WMQA)

$$\text{MQA} = \frac{1}{T} \sum_{t=1}^T (x_t - \hat{x}_t)^2 \quad \text{WMQA} = \sqrt{\frac{1}{T} \sum_{t=1}^T (x_t - \hat{x}_t)^2}$$

Im Unterschied zu MAA erhalten hier große Abweichungen (durch das Quadrieren) ein höheres Gewicht, was häufig erwünscht ist. Beruht die große Abweichung jedoch nur auf einer singulären Störung („Ausreißer“), dann ist die hohe Gewichtung ungerechtfertigt und die Verwendung von MAA angezeigt. Der Vorteil von WMQA gegenüber MQA ist die gleiche Maßeinheit bei Fehlermaß und Zeitreihenwerten.

c) Der Ungleichheitskoeffizient von Theil (U)

Alle bisher besprochenen Fehlermaße sind *nicht normiert*. Bei voller Übereinstimmung von Prognose und Realität nehmen sie zwar alle den Wert Null an und erlauben somit

das Urteil „Ideale Prognose“. Es fehlt jedoch ein kritischer Wert, der zwischen guten und schlechten Prognosen diskriminiert. Diesen Nachteil versucht Theil (1971, S. 28 und 1978, S. 368) mit der Konstruktion des folgenden Maßes zu umgehen

$$U = \sqrt{\frac{\sum_{t=1}^T (x_t - \hat{x}_t)^2}{\sum_{t=1}^T (x_t - x_{t-1})^2}}$$

x_0 ist der letzte beobachtete Wert vor Beginn des Prognosezeitraums

Auch U nimmt den Wert Null an, wenn „ideal“ prognostiziert wird. Zur Abgrenzung der guten von den schlechten Prognosen dient nun der Wert $U = 1$, der sich ergibt, wenn man nach folgender Regel „naiv“ prognostiziert:
 „der letzte Beobachtungswert dient als Prognosewert“

$$\hat{x}_t = x_{t-1} \quad (t = 1, \dots, T)$$

Man rechnet zwar mit keiner Änderung der Zeitreihe, paßt sich jedoch in jedem Zeitpunkt t der letzten Änderung an. Dies ist sicherlich das einfachste Prognoseverfahren und führt zu $U = 1$, wie sich durch Einsetzen in die U-Formel zeigen läßt. Man kann nun festlegen:

Ein Prognoseverfahren ist gut, wenn es einen kleineren Ungleichheitskoeffizienten liefert als die „naive“ Prognose, d. h. $U < 1$.

Dabei sind auch Werte knapp unterhalb von Eins schon als recht gut einzustufen, weil U die Prognosefehler zwischen $U = 0$ und $U = 1$ in der Regel *nicht* linear abbildet. Dies ist zugleich der Hauptnachteil des Ungleichheitskoeffizienten.

Trotzdem erscheint mir das Theil'sche Fehlermaß geeignet zur Ex-post-Beurteilung von Prognosen und den ihnen zugrunde liegenden Prognoseverfahren. Wir werden aber in diesem Buch auch die übrigen Maße verwenden, um sie empirisch vergleichen zu können.

d) Fehlermaße zur qualitativen Beurteilung von Prognosen

Ein sinnvolles Kriterium für die Güte einer Prognose kann u. U. die Häufigkeit sein, mit der die Änderung der Tendenz einer Zeitreihe (Wendepunkt) *richtig* vorhergesagt wird (vgl. Schwarze 1980, S. 319). Da es hierbei nicht auf die *Größe*, sondern nur auf das *Vorzeichen* des Prognosefehlers ankommt, nennt man solche Kriterien qualitative Fehlermaße.

Ein Wendepunkt der Zeitreihe liegt vor, wenn gilt

$$x_t - x_{t-1} > 0 \text{ und } x_{t+1} - x_t < 0$$

oder

$$x_t - x_{t-1} < 0 \text{ und } x_{t+1} - x_t > 0.$$

Man kann nun gemäß Tabelle 1 die Anzahl der *nicht* vorhergesagten Wendepunkte (B) zur Anzahl der vorhergesagten, aber *nicht* eingetretenen Wendepunkte (C) addieren und durch die Summe aus der Anzahl vorhergesagter und nicht vorhergesagter Wendepunkte (A+B+C) dividieren.

Tabelle 1

Anzahl von Wendepunkten	vorhergesagt	nicht vorhergesagt
eingetreten	A	B
nicht eingetreten	C	D

Dieses Wendepunktfehlermaß nimmt bei richtigen Prognosen ($B = 0$ und $C = 0$) den Wert *Null* und für den Fall, daß kein eingetretener Wendepunkt vorhergesagt wurde ($A = 0$), den Wert *Eins* an, ist also ebenfalls normiert.

Damit sei die Erläuterung der Fehlermaße abgeschlossen. Mit der Kenntnis vom Aufbau der Prognosemodelle, zeitlichem Ablauf der Prognose und der Beurteilungsmaße sind die Grundlagen für die folgende Beschreibung einzelner Prognoseverfahren und ihre Anwendung auf empirisches Datenmaterial erarbeitet.

B. Heuristische Prognoseverfahren

I. Überblick

Heuristische Prognoseverfahren sind ein Sammelbegriff für zweckmäßige, methodisch erarbeitete Prognoseregeln (von heuriskein (gr.) = systematisch suchen) ohne schematisiertes Prognosemodell.

Sie zeichnen sich durch folgende Eigenschaften aus (Hansmann 1979, S. 229):

1. Die der Prognose zugrunde liegende Theorie ist nur schwach ausgebildet und/oder enthält viele subjektive, d. h. nicht unmittelbar nachprüfbare Elemente.
2. Die statistisch-mathematischen Instrumente treten in ihrer Bedeutung für die Prognose zurück.
3. Der Einsatz von „Experten“, deren Erfahrungen auf Spezialgebieten für die Prognose nutzbar gemacht werden, ist stark verbreitet.

Wegen 2. werden die heuristischen Prognoseverfahren häufig als qualitative Prognoseverfahren bezeichnet. Dies ist jedoch nicht ganz zutreffend, da in Teilbereichen des Prognoseproblems durchaus quantitative Methoden, z. B. Computer-Simulationen in langfristigen Prognosen des Energiebedarfs, zur Anwendung kommen.

Die Eigenschaften der heuristischen Prognoseverfahren legen es nahe, sie vor allem für langfristige Prognosen (10 bis 50 Jahre) und in der Futurologie (vgl. z. B. Kahn 1976) einzusetzen. Vielfach werden aber auch „Schnellprognosen“ mit diesen Verfahren erstellt, wenn keine Zeit oder Mühe für die Beschaffung und Analyse von Zeitreihen aufgewendet werden soll.

Das Spektrum der heuristischen Prognoseverfahren ist umfangreich. Als bekannteste Verfahren seien erwähnt:

1. Expertenbefragungen
2. Analogieschlüsse (Bruckmann, 1978)
3. die Delphi-Methode
4. die Methode des morphologischen Kastens (Zwicky, 1971)
5. die Szenario-Technik.

Im Rahmen dieses Buches sollen als typische Verfahren die Szenario-Technik und die Delphi-Methode vorgestellt werden.

II. Die Szenario-Technik

1. Kennzeichnung

Das Szenario ist die Beschreibung der zukünftigen Entwicklung des Prognosegegenstandes bei alternativen Rahmenbedingungen.

Aus der gründlichen Analyse der Gegenwart und dem Aufzeigen verschiedener Entwicklungsmöglichkeiten ergibt sich ein Gesamtbild für den Prognosezeitraum, das meist qualitativ – mit einigen quantitativen Daten versehen – beschrieben wird.

2. Anwendungsbereiche

Die Szenario-Technik ist vor allem im Bereich der langfristigen Wirtschafts-, Energie- und Technologieprognose beheimatet, wird aber auch zur Prognose kultureller und gesellschaftlicher Entwicklungsprozesse herangezogen. Anschauungsmaterial bietet das Buch „The next 200 years“ von Hermann Kahn (1976).

3. Anwendungsbeispiel: Energiebedarf der BRD für die Jahre 2000 und 2030

Die vom 8. Deutschen Bundestag eingesetzte Enquete-Kommission hat in ihrem Bericht „Zukünftige Kernenergie-Politik“ (Bundestagsdrucksache 8/4341) ein Szenario vorgestellt, das vier alternative, zukünftige Entwicklungen (Pfade) von Energiebedarf und Energieversorgung enthält.

Als Determinanten der *Energienachfrage* werden

- die Bevölkerungsentwicklung,
- die Strukturentwicklung der Volkswirtschaft,
- das Wachstum des Bruttosozialprodukts und
- Energieeinsparmaßnahmen

angesehen. Die mögliche Entwicklung dieser Größen in den vier Pfaden wird plausibilitätsgemäß gestützt und für die in einem Modell der linearen Optimierung abgebildeten Sachzusammenhänge exogen vorgegeben. Bezüglich der Determinanten und der alternativen Pfade werden folgende Annahmen getroffen:

- In allen vier Pfaden wird von einer einheitlichen Bevölkerungsentwicklung ausgegangen. Es werden die Projektionen des Statistischen Bundesamtes zugrunde gelegt (im Jahre 2000 etwa 57 Millionen, im Jahre 2030 etwa 50 Millionen Personen).
- Für die Strukturentwicklung der Volkswirtschaft wird in den Pfaden 1 und 2 ein „mittlerer Strukturwandel“, in den Pfaden 3 und 4 ein „starker Strukturwandel“ unterstellt. Der Strukturwandel bezieht sich auf die Bereiche
 - Grundstoff- und Produktionsgüterindustrie
 - übrige Industrie und
 - Dienstleistungsbereiche.
- Mittlerer Strukturwandel beinhaltet die Annahmen, daß die Grundstoffproduktion halb so stark wächst wie die übrige Industrie, und der Dienstleistungsbereich mit 10 % gegenüber dem Bruttosozialprodukt überproportional zunimmt.

Starker Strukturwandel ist durch eine gleichbleibende Grundstoffproduktion, gleiches Wachstum der übrigen Industrie und des Bruttosozialproduktes, sowie ein überproportionales Wachstum des Dienstleistungsbereiches von 20 % gekennzeichnet.

- Das Wachstum des Bruttosozialprodukts hängt wesentlich vom eintretenden Strukturwandel ab. 47 % des Energiebedarfs sind mit der Entwicklung des Bruttosozialprodukts gekoppelt. 53 % des Energiebedarfs werden als unabhängig vom Bruttosozialprodukt angenommen und einheitlich für alle Pfade exogen vorgegeben.
- Das Wirtschaftswachstum wird in durchschnittlichen jährlichen Steigerungsraten des Bruttosozialprodukts angegeben, wobei Zeiträume vor und nach dem Jahr 2000 unterschieden werden.

	vor 2000	nach 2000
Pfad 1	+ 3,3 %	+ 1,4 %
Pfad 2	+ 2,0 %	+ 1,1 %
Pfad 3	+ 2,0 %	+ 1,1 %
Pfad 4	+ 2,0 %	+ 1,1 %

- Das Maß der Energieeinsparung wird für die einzelnen Pfade normativ gesetzt, indem jeweils ein technisch realisierbarer *und* gewünschter Standard der Energieeinsparung unterstellt wird. Die Einsparraten (ER) sind auf den technischen Standard von 1980 bezogen. In Abhängigkeit von den wirtschaftlichen Sektoren werden für die einzelnen Pfade folgende Einsparraten im Zeitraum 1980 bis 2030 angenommen.

Pfad 1	ER bis zu 25 % := Trendsparen
Pfad 2	ER bis zu 52 % := starkes Sparen
Pfad 3	ER bis zu 70 % := sehr starkes Sparen
Pfad 4	ER bis zu 80 % := extrem starkes Sparen

In Tabelle 2 sind die Ergebnisse des Szenarios auszugsweise wiedergegeben. Der für die Jahre 2000 und 2030 und pro Pfad errechnete Energiebedarf (letzte Zeile) verteilt sich auf die vier Energieträger Öl/Gas, Kohle, Uran und regenerative Energiequellen, wie z. B. Wind, Biogas, Treibstoffe aus Biomasse. Die Tabellenwerte sind einheitlich in 10^6 t SKE (Mega-Tonnen Steinkohleeinheiten; 1 kg SKE = 8,141 kWh) wiedergegeben.

Deutlich erkennbar weist Tabelle 2 die Wirkungen gesellschaftlicher Haltungen gegenüber der Energieproblematik aus. Während in den Pfaden 1 und 2 die Kernenergie toleriert wird, erscheint sie in den Pfaden 3 und 4 als gesellschaftlich nicht durchsetzbar. Die Pfade geben insgesamt ein breites Spektrum von Entwicklungen des Primärenergiebedarfs wieder. Im Vergleich zu den realisierten Werten des Jahres 1978 werden bis zum Zeithorizont im Jahre 2030 sowohl Steigerungen als auch Abnahmen des Energiebedarfs für möglich gehalten.

Das Szenario der Enquete-Kommission zeichnet sich durch seine Transparenz und Nachvollziehbarkeit aus. Darüber hinaus wird es durch die Darstellung alternativer Entwicklungen der Zielsetzung eines Szenarios gerecht.

Tabelle 2

Szenario „Primärenergiebedarf und Aufteilung auf alternative Energieträger“
[in 10⁶ t SKE]

Energieträger	1978	2000				2030			
		Pfad 1	Pfad 2	Pfad 3	Pfad 4	Pfad 1	Pfad 2	Pfad 3	Pfad 4
Stein- und Braunkohle	105	175	145	145	130	210	160	160	145
Erdöl und Erdgas	265	250	190	190	165	250	130	130	65
Kernenergie	12	135	70	–	–	290	210	–	–
regenerative Energiequellen	8	40	40	40	50	50	50	70	100
Primärenergiebedarf	390	600	445	375	345	800	550	360	310

4. Kritik

Vorzug der Szenario-Technik ist die Beachtung der allgemeinen Interdependenz der Einflußfaktoren. Es werden alternative Entwicklungstendenzen nach vielen Richtungen ausgelotet. Das Prognoseumfeld und die Randbedingungen sind Bestandteil des Szenarios.

Als *Nachteil* muß die Fülle von subjektiven Einschätzungen angesehen werden. Im Anwendungsbeispiel sind die Annahmen über die Einsparraten für Energie ähnlich heikel wie die Vermutungen über die Verfügbarkeit regenerativer Energiequellen.

Trotzdem hat das Szenario bei komplexen Prognoseproblemen, die keine einfache Trendextrapolation gestatten, durchaus seinen Wert, insbesondere dann, wenn es, wie im Anwendungsbeispiel, als ein fundiertes Ausleuchten von Zukunftsperspektiven betrachtet wird.

III. Die Delphi-Methode

1. Grundprinzipien

Die Delphi-Methode ist eine spezielle Form der *Gruppenprognose*, die von Wissenschaftlern der amerikanischen RAND Corporation Anfang der 60er Jahre entwickelt wurde (Helmer/Gordon 1967, Geschka in Bruckmann 1978, S. 27 ff.).

Das Prognoseverfahren hat folgende charakteristische Eigenschaften (Hansmann 1979, S. 232):

1. Die Prognose wird von *Experten* abgegeben, die sich mit unterschiedlichen Aspekten des Prognoseproblems beschäftigt haben.
2. Die Experten lernen sich *nicht* kennen (*Anonymität*).
3. Die Prognose vollzieht sich in mehreren Runden (Iterationen), wobei von einer Runde zur nächsten eine kontrollierte *Informationsrückkoppelung* stattfindet.
4. Die einzelnen Prognosen der Expertengruppe werden *statistisch ausgewertet*.

Durch die anonyme Befragung soll der aus Gruppendiskussionen bekannte Effekt verhindert werden, daß dominierende Gruppenmitglieder die Prognosen der anderen beeinflussen. Auf der anderen Seite wird die Gefahr von isolierten Einzelprognosen, die infolge der Anonymität naheliegt, durch die Informationsrückkoppelung zwischen den Befragungsrunden vermieden, denn das Ergebnis einer Prognoserunde wird den Teilnehmern in statistischer Form mitgeteilt, so daß sie über zusätzliche Informationen verfügen.

2. Ablauf der Delphi-Prognose

Nachdem der Expertengruppe der Prognosegegenstand mitgeteilt und die Delphi-Technik erläutert worden ist, beginnt die eigentliche Prognoseprozedur.

Runde 1

Die Experten geben schriftlich und anonym ihre erste Prognose über den Prognosegegenstand ab. Der Projektleiter ermittelt aus den abgegebenen Prognosen den Median, das untere und obere Quartil und die Quartilspanne.

Runde 2

Jeder Teilnehmer erhält die statistischen Kennzahlen der ersten Runde. Er gibt nun unter Berücksichtigung dieser Information eine zweite Prognose ab. Liegt sie außerhalb der Quartilspanne, ist sie kurz zu begründen. Der Projektleiter ermittelt aus den zweiten Prognosen wiederum Median, Quartile und Quartilspanne.

Runde 3

Die neuen statistischen Kennzahlen *und* evtl. Begründungen für „extreme“ Prognosen werden den Teilnehmern zugestellt. Anschließend prognostizieren die Experten zum dritten Mal, und der Projektleiter wertet die Angaben statistisch aus. Der *Median* ist der endgültige *Prognosewert*, die Quartilspanne wird als Streuung interpretiert.

Dieser Ablauf stellt das Grundschema dar, das sich selbstverständlich auf weitere Runden (hoher Zeitaufwand!) ausdehnen läßt. Im Normalfall genügen jedoch drei Runden, um die gewünschte Konvergenz der Meinungen zu erreichen.

3. Anwendungsbeispiel: Prognose der Veränderung des Aktienkursindex des Statistischen Bundesamtes

Für eine empirische Anwendung der Delphi-Methode konnte der Verfasser Wertpapierexperten von acht Hamburger Kreditinstituten gewinnen. Prognosegegenstand war die prozentuale Veränderung des Aktienkursindex des Statistischen Bundesamtes zwischen Dezember 1981 und Dezember 1982 (Prognosezeitraum = ein Jahr; Aktienkursindex 31.12.81 = 99,8, 31.12.82 = 107,1, Veränderung = 7,3).

Alle Regeln der Delphi-Methode, insbesondere die Anonymität, wurden eingehalten. Drei Prognoserunden fanden im Zeitraum Januar/April 1982 statt. Die Expertenschätzungen sowie die Ergebnisse der statistischen Auswertung jeder Runde werden im folgenden chronologisch wiedergegeben:

Als statistische Kennzahlen wurden errechnet:

Z := Median

Q_u := unteres Quartil

Q_o := oberes Quartil

Q_m := Quartilspanne = $Q_o - Q_u$

\bar{x} := arithmetisches Mittel

s := Standardabweichung

R := Spannweite

In der 1. Runde schätzten die Experten folgende Werte x_i ($i=1, \dots, 8$):

Prozentuale Änderung des Aktienkursindex (1. Runde)

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
+ 5%	+ 6,2%	+ 8%	+ 10%	+ 11%	+ 11,1%	+ 14%	+ 15,2%

Daraus errechneten sich die

statistischen Kennzahlen für die Schätzwerte der 1. Runde

Z	Q _u	Q _o	Q _m	\bar{x}	s	R
10,5%	6,2%	14%	7,8%	10,06%	3,56%	10,2%

Als Gruppenurteil (Median Z) ergab sich ein prozentualer Anstieg des Aktienkursindex von 10,5 %, was einem erwarteten Indexwert von 110,3 zum 31.12.82 entspricht.

In der 2. Runde waren Schätzwerte außerhalb der Quartilspanne verbal zu begründen. Folgende Schätzwerte x_i wurden abgegeben:

Prozentuale Änderung des Aktienkursindex (2. Runde)

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
+ 6,2%	+ 8%	+ 10%	+ 10%	+ 11%	+ 11,1%	+ 15%	+ 15,2%

Die über $Q_o = 14\%$ liegenden Schätzwerte wurden mit einem erwarteten sinkenden Zinsniveau in den USA begründet. Es wurde besonders hervorgehoben, daß die Deutsche Börse extrem empfindlich auf Zinsänderungen in den USA reagiere.

Aus den Angaben errechneten sich die folgenden

statistischen Kennzahlen für die Schätzwerte der 2. Runde

Z	Q _u	Q _o	Q _m	\bar{x}	s	R
10,5%	8%	15%	7,0%	10,81%	3,10%	9%

Während der Median Z im Vergleich zur 1. Runde mit 10,5 unverändert bleibt, signalisiert das arithmetische Mittel \bar{x} der 2. Runde mit 10,81 eine insgesamt optimistischere Einschätzung der prozentualen Änderung des Aktienkursindex. Darüber hinaus zeichnet sich in den Werten der Quartilspanne Q_m , der Standardabweichung s sowie der Spannweite R eine schwache Konvergenz der Meinungen ab.

In der abschließenden 3. Runde wurden folgende Werte x_i von den Experten geschätzt:

Prozentuale Änderung des Aktienkursindex (3. Runde)

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
+ 7,7%	+ 8%	+ 10%	+ 10%	+ 11%	+ 14%	+ 15,2%	+ 16,2%

Daraus errechneten sich als

statistische Kennzahlen für die Schätzwerte der 3. Runde

Z	Q _u	Q _o	Q _m	\bar{x}	s	R
10,5%	8%	15,2%	7,2%	11,51%	3,24%	8,2%

Das Ergebnis der Delphi-Prognose zeigt, daß der Median während der drei Befragungsrunden konstant bleibt, das arithmetische Mittel jedoch stetig ansteigt. Auch die beiden Quartile nehmen im Verlauf der Befragung höhere Werte an. Die Experten schätzten also die Aktienkursentwicklung Anfang April etwas optimistischer ein als Ende Januar. Der Median ist allerdings nicht sensibel genug, um diese leichte Veränderung bereits anzuzeigen.

Eine Konvergenz der Expertenschätzungen – gemessen an der Quartilspanne bzw. der Standardabweichung – ist nur von der ersten zur zweiten Runde festzustellen. In der dritten Runde divergieren die Meinungen wieder etwas. Dies dürfte in erster Linie auf Änderungen der gesamtwirtschaftlichen Daten (Unternehmenserträge, Zins- und Dollarkursentwicklung) zurückzuführen sein, die die Expertenmeinungen unterschiedlich beeinflußt haben.

Bemerkenswert ist die Tatsache, daß alle Experten einen Anstieg des Aktienkursniveaus bis zum Jahresende 1982 erwarten, über das Ausmaß jedoch recht unterschiedlicher Ansicht sind, was durch die Spannweite von immerhin 8,2 Prozentpunkten nach der dritten Runde zum Ausdruck kommt.

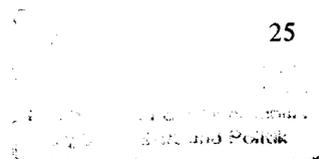
4. Kritik des Verfahrens

Die wichtigsten *Vorzüge* des Verfahrens seien hier noch einmal zusammengefaßt (vgl. Hansmann 1979, S. 232/233):

1. Durch die anonyme Befragung unterbleibt ein Gruppendruck.
2. Die abgegebenen Prognosen werden über mehrere Runden hinweg durch die Gruppe kontrolliert (durch statistische Kenngrößen).
3. Die schriftlichen Begründungen extremer Schätzungen liefern zusätzliche Informationen.
4. Die Informationsrückkoppelung begünstigt die Konvergenz auseinanderstrebender Expertenmeinungen.

Neben diesen unmittelbar einzusehenden Vorzügen hat die Delphi-Methode jedoch bisweilen nicht wünschenswerte Effekte, die gerade durch diese Form der Befragung begünstigt werden.

1. Durch die anonyme Befragung kann unter den Teilnehmern kein direkter Lernprozeß stattfinden.
2. Das starre Befragungsschema läßt für neue Entwicklungen während der Befragung nicht genügend Raum.
3. Die Informationsrückkoppelung kann vorsichtige Schätzungen verstärken; die Teilnehmer weichen dem Begründungszwang aus.
4. Die Delphi-Methode erfordert in der Regel einen hohen Zeitbedarf, der allerdings von der Gruppengröße abhängt. Die Delphi-Gruppe sollte aus statistischen Gründen



mindestens acht Teilnehmer aufweisen. Schon bei dieser Gruppengröße erfordert die Prognose mehrere Monate. Für 20 Teilnehmer veranschlagt man ein halbes Jahr, was für manche Prognosegegenstände schon unvertretbar sein kann.

Die bisher vorliegenden empirischen Erfahrungen mit der Delphi-Methode beweisen nicht generell ihre Überlegenheit gegenüber den normalen Gruppendiskussionen (Brockhoff 1977, S. 82 f.), doch wird die Leistungsfähigkeit des Verfahrens für mittel- bis langfristige Prognosen grundsätzlich anerkannt.

C. Prognosen auf der Grundlage der exponentiellen Glättung

Nach den heuristischen Prognoseverfahren wenden wir uns nunmehr den quantitativen, auf der Analyse von Zeitreihen aufbauenden Prognoseverfahren zu und beginnen mit dem einfachsten und in der Praxis weit verbreiteten Verfahren: der exponentiellen Glättung (engl. exponential smoothing).

Die exponentielle Glättung wird überwiegend auf quantitative univariate kurzfristige Prognosemodelle im Sinne unserer Klassifikation von Kapitel A.II. (S. 12 ff.) angewendet. Die zukünftigen Werte einer Zeitreihe $\{x_t\}$ werden also aus der Zeitreihe heraus ohne Beachtung anderer (kausaler) Einflußfaktoren prognostiziert. Dazu verwendet man *eine einzige*, in den Zeitreihenwerten *lineare* Prognosegleichung.

Die Ableitung und inhaltliche Interpretation der Prognosegleichung geschieht am einfachsten über das Verfahren der „Gleitenden Durchschnitte“, dessen wichtigste Eigenschaften kurz erörtert werden sollen (vgl. hierzu Hansmann 1980 a, S. 17 ff.).

I. Gleitende Durchschnitte

Ziel dieses Verfahrens ist die Ausschaltung zufallsbedingter Unregelmäßigkeiten im Verlauf einer Zeitreihe. Um die systematischen Komponenten der Zeitreihe klarer hervortreten zu lassen, wird das arithmetische Mittel der n letzten Zeitreihenwerte gebildet und dadurch ein „Filtereffekt“ hervorgerufen.

Bezeichnet man den *Gegenwartszeitpunkt* mit t , das arithmetische Mittel der Zeitreihe zum Zeitpunkt t mit \bar{x}_t und die Periodenlänge des gleitenden Durchschnitts mit n , so ergibt sich folgende Definitionsgleichung:

$$(1) \quad \bar{x}_t = \frac{x_t + x_{t-1} + x_{t-2} + \dots + x_{t-(n-1)}}{n}$$

Diese Gleichung läßt sich kürzer schreiben, wenn man beachtet, daß

$$(2) \quad \bar{x}_{t-1} = \frac{x_{t-1} + x_{t-2} + \dots + x_{t-n}}{n}$$

gilt. (2) in (1) eingesetzt ergibt

$$(3) \quad \bar{x}_t = \bar{x}_{t-1} + \frac{x_t - x_{t-n}}{n}$$

Der jeweils neueste gleitende Durchschnitt ergibt sich also aus dem vorhergehenden, indem man den aktuellen Zeitreihenwert x_t , gewichtet mit $1/n$, hinzufügt und den ältesten Wert x_{t-n} , mit demselben Gewicht versehen, abzieht.

Unter der Voraussetzung, daß die Zeitreihe *keinen Trend* aufweist und die *Zeitstabilitätshypothese* (S. 11) gilt, kann der gleitende Durchschnitt \bar{x}_t als Prognosewert \hat{x}_{t+1} der nächsten Periode verwendet werden.

Bei dieser – einfachen – Prognose erhält jeder Zeitreihenwert – unabhängig von seiner Aktualität – dasselbe Gewicht $1/n$. Wir hatten in Kapitel A. jedoch darauf hingewiesen, daß die Zeitstabilitätshypothese in der Realität nur näherungsweise gilt. Die Struktur von Zeitreihen ändert sich allmählich im Zeitablauf. Will man dieser Veränderung Rechnung tragen, so sollten die jüngeren Zeitreihenwerte ein stärkeres Gewicht erhalten als die älteren, weil sie die gegenwärtige Zeitreihenstruktur besser widerspiegeln.

Die exponentielle Glättung ist ein Verfahren, das diese Überlegungen in bestimmter Weise operationalisiert.

II. Das konstante Modell der exponentiellen Glättung

1. Grundgedanken

Das Verfahren wurde 1959 von Brown entwickelt, indem er zwei Überlegungen kombinierte (Brown 1963):

- Das Gewicht der Zeitreihenwerte für die Prognose soll mit zunehmendem „Alter“ der Werte abnehmen.
- Der Prognosefehler der Gegenwart (e_t) wird bei den folgenden Prognosen berücksichtigt.

Aus diesen beiden Überlegungen läßt sich das konstante Modell der exponentiellen Glättung ableiten (vgl. Hansmann 1980 a, S. 18 ff.).

Nach der zweiten Überlegung lautet der Prognosewert für die nächste Periode ($t + 1$)

$$(4) \quad \hat{x}_{t+1} = \hat{x}_t + \alpha e_t = \hat{x}_t + \alpha(x_t - \hat{x}_t) \quad (0 < \alpha \leq 1)$$

wobei $e_t = x_t - \hat{x}_t$ der Prognosefehler zum Gegenwartszeitpunkt t ist. Durch einfache Umformulierung von (4) gewinnt man die Grundformel der exponentiellen Glättung

$$(5) \quad \hat{x}_{t+1} = \alpha x_t + (1 - \alpha)\hat{x}_t \quad (0 < \alpha \leq 1)$$

Man benötigt zur Ermittlung des Prognosewertes nur den vorhergehenden Prognosewert, den letzten Zeitreihenwert und den sog. *Glättungsparameter* α . Die Wahl von α ist nach Gleichung (4) der Bruchteil des Prognosefehlers, der in die neue Prognose eingehen soll. Welcher Bruchteil ist nun geeignet?

Diese Frage ist nicht leicht zu beantworten. Hüttner berichtet über gute Erfahrungen mit α -Werten zwischen 0,1 und 0,3 (Hüttner 1982, S. 99). Weitere Möglichkeiten zur Wahl von α werden wir in Abschnitt IV., S. 39 ff., kennenlernen.

Es bleibt noch zu zeigen, daß Grundformel (5) auch die erste Überlegung von Brown umfaßt, nämlich die Zeitreihenwerte mit abnehmenden Gewichten zu versehen.

Wir schreiben (5) für die Zeitpunkte (bzw. Perioden) $t - 1$ und $t - 2$

$$(5a) \quad \hat{x}_t = \alpha x_{t-1} + (1 - \alpha) \hat{x}_{t-1}$$

$$(5b) \quad \hat{x}_{t-1} = \alpha x_{t-2} + (1 - \alpha) \hat{x}_{t-2}$$

und setzen nun (5a) in (5) ein

$$(6) \quad \hat{x}_{t+1} = \alpha x_t + (1 - \alpha) (\alpha x_{t-1} + (1 - \alpha) \hat{x}_{t-1})$$

Einsetzen von (5b) in (6) und Zusammenfassen ergibt

$$(6a) \quad \hat{x}_{t+1} = \alpha x_t + \alpha(1 - \alpha) x_{t-1} + \alpha(1 - \alpha)^2 x_{t-2} + (1 - \alpha)^3 \hat{x}_{t-2}$$

Hier zeigt sich bereits die Struktur der Gewichtung der Zeitreihenwerte: *Exponentielle Abnahme* mit zunehmendem Alter. Durch rekursives Einsetzen der Grundgleichung (5) für \hat{x}_{t-i} ($i = 0, 1, \dots, t-1$) bis zum Beginn der Zeitreihe ergibt sich die allgemeine Form

$$(7) \quad \hat{x}_{t+1} = \alpha \sum_{i=0}^{t-1} (1 - \alpha)^i x_{t-i} + (1 - \alpha)^t \hat{x}_1$$

in der $\alpha(1 - \alpha)^i$ das Gewicht des Reihenwertes x_{t-i} darstellt. Von dieser exponentiellen Abnahme der Gewichte hat das Verfahren seinen Namen erhalten.

Bei großem Zeitreihenumfang kann das Restglied $(1 - \alpha)^t \hat{x}_1$ vernachlässigt werden, da es gegen Null konvergiert. Schon 10 Reihenwerte ergeben mit $\alpha = 0,5$ ein Restglied von $0,5^{10} \hat{x}_1$. Der Einfluß von \hat{x}_1 auf den Prognosewert ist demnach kleiner als 0,1% der ersten Schätzung.

Wächst die Anzahl der Zeitreihenwerte über alle Grenzen, so strebt die Summe aller Gewichte für $0 < \alpha \leq 1$ gegen Eins, eine „angenehme“ Eigenschaft, die auch das Verfahren der Gleitenden Durchschnitte aufweist:

$$(8) \quad \lim_{t \rightarrow \infty} \alpha \sum_{i=0}^{t-1} (1 - \alpha)^i = \lim [\alpha + \alpha(1 - \alpha) + \alpha(1 - \alpha)^2 + \dots] = \frac{\alpha}{1 - (1 - \alpha)} = 1$$

2. Die Zielfunktion des Verfahrens

Das oben abgeleitete konstante Modell der exponentiellen Glättung ist auf Zeitreihen anwendbar, die keinem Trend und keinen Saisonschwankungen folgen. Um seine Verwandtschaft mit statistischen Zeitreihenverfahren aufzuzeigen, wollen wir unter diesen beiden Voraussetzungen die Zielfunktion des Verfahrens betrachten und beweisen, daß der nach (5) ermittelte Prognosewert \hat{x}_{t+1} die Summe der exponentiell gewogenen quadratischen Abweichungen der Zeitreihenwerte vom Prognosewert minimiert.

Eine Zeitreihe ohne Trend und Saison kann als Konstante S aufgefaßt werden, die von einer zufälligen Störvariablen u additiv überlagert wird. Die Störvariable enthält eine

Vielzahl von unabhängig wirkenden schwachen Einflußfaktoren, die einzeln nicht erfaßbar sind und sich im Mittel gegenseitig neutralisieren. Es gilt daher die Gleichung

$$(9) \quad x_{t-i} = S + u_{t-i} \quad (i = 0, 1, \dots, t-1)$$

und für u_{t-i} :

$$E[u_{t-i}] = 0 \quad \text{und} \quad E[u_{t-i} \cdot u_{t-j}] = \begin{cases} \sigma_u^2 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$$

$E[u_{t-i}]$ ist der Erwartungswert und σ_u^2 die Varianz der Störvariablen u .

Bei richtiger Schätzung von S für den Vergangenheitszeitraum kann der Schätzwert \hat{S} als Prognosewert \hat{x}_{t+1} verwendet werden. Man sucht also das Minimum folgender Zielfunktion:

$$(10) \quad Z = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i \cdot u_{t-i}^2 = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i (x_{t-i} - S)^2 \rightarrow \min!$$

Nullsetzen der ersten Ableitung nach S ergibt

$$(11) \quad \frac{dZ}{dS} = -2\alpha \sum_{i=0}^{t-1} (1-\alpha)^i (x_{t-i} - S) = 0$$

Hieraus läßt sich, da die zweite Ableitung nach S positiv ausfällt, der Wert für S bestimmen, der die Zielfunktion minimiert

$$(12) \quad \hat{S} = \frac{\alpha \sum_{i=0}^{t-1} (1-\alpha)^i x_{t-i}}{\alpha \sum_{i=0}^{t-1} (1-\alpha)^i}$$

Da der Nenner von (12) bei großem Zeitreihenumfang gegen Eins strebt und in diesem Fall auch das Restglied in Gleichung (7) vernachlässigt werden darf, stimmen die Gleichungen (7) und (12) für genügend lange Zeitreihen überein

$$(12a) \quad \hat{x}_{t+1} = \hat{S} = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i x_{t-i} = \alpha x_t + (1-\alpha) \hat{x}_t$$

Somit minimiert \hat{x}_{t+1} die Summe der exponentiell gewogenen Abweichungsquadrate der Zeitreihenwerte vom Prognosewert.

3. Die Bedeutung des Glättungsparameters α

α legt, wie wir gesehen haben, die Gewichte der einzelnen Zeitreihenwerte fest. Ein großer Wert für α gewichtet den letzten Zeitreihenwert x_t gegenüber den früheren Werten sehr stark. Bei $\alpha = 1$ ist gemäß Prognosegleichung (12a) der Prognosewert \hat{x}_{t+1} gleich x_t und die übrige Vergangenheit wird „vergessen“. Ein hohes α ist also nur bei einer erwarteten grundlegenden Änderung der Zeitreihe (sog. Strukturbruch) sinnvoll.

Ein geringer Wert für α gewichtet die Vergangenheit höher als den letzten Zeitreihenwert, bei $\alpha \rightarrow 0$ würde er überhaupt nicht berücksichtigt. Es empfiehlt sich daher ein kleines α , wenn der letzte Zeitreihenwert als singulärer „Ausrutscher“ angesehen werden kann.

Diese Überlegungen zeigen, daß es nicht möglich ist, ein optimales α ex ante festzulegen. Man erhält bessere Prognosen, wenn α dynamisch verändert und an die Entwicklung der Zeitreihe angepaßt wird (vgl. Abschnitt IV., S. 39 ff.).

4. Anwendung des Modells auf den Verbrauch an Spirituosen in der BRD

Als Anwendungsbeispiel für das konstante Modell wird der Pro-Kopf-Verbrauch von Spirituosen (gemessen in Litern) in der Bundesrepublik Deutschland für den Zeitraum 1970 bis 1981 zugrunde gelegt und eine Prognose für das Jahr 1982 errechnet. Die Zeitreihe läßt weder einen klaren Trend noch einen Saisonzyklus erkennen, so daß das konstante Modell angebracht ist. Die Werte sind dem ifo-schnelldienst (Heft 7/8, 35. Jg., 1982, S. 6) entnommen.

Tabelle 3a

Periode	t	x_t (l)	Periode	t	x_t (l)
1970	1	6,80	1976	7	8,40
1971	2	7,90	1977	8	6,80
1972	3	6,60	1978	9	7,90
1973	4	7,20	1979	10	7,90
1974	5	7,20	1980	11	8,00
1975	6	7,60	1981	12	7,70

Gesucht ist der Prognosewert \hat{x}_{1982} , als Gegenwartszeitraum t gilt das Jahr 1981. Für α wählen wir den Wert 0,3. Damit das Verfahren initialisiert werden kann, benötigen wir nach Gleichung (5) bzw. (7) einen Schätzwert für das Jahr 1970 (\hat{x}_1), den wir (willkürlich) gleich dem Beobachtungswert $x_1 = 6,8$ setzen.

Mit den Prognosegleichungen (5) oder (7) kann nunmehr die Prognose erstellt werden. (5) hat den Vorteil der Fortschreibung aller Schätzwerte und wird in EDV-Programmen benutzt, (7) ist numerisch aufwendiger.

In der folgenden Tabelle sind die tatsächlichen Zeitreihenwerte, die fortgeschriebenen Schätzwerte und die Schätzfehler wiedergegeben.

Tabelle 3 b

t-i	x_{t-i}	\hat{x}_{t-i}	e_{t-i}
1	6,80	6,80	0,00
2	7,90	6,80	1,10
3	6,60	7,13	-0,53
4	7,20	6,97	0,23
5	7,20	7,04	0,16
6	7,60	7,09	0,51
7	8,40	7,24	1,16
8	6,80	7,59	-0,79
9	7,90	7,35	0,55
10	7,90	7,52	0,38
11	8,00	7,63	0,37
12	7,70	7,74	-0,04
13	—	7,73	—

MAA	(mittlere absolute Abweichung)	= 0,53
WMQA	(Wurzel aus mittlerer quadrierter Abweichung)	= 0,63
U	(Theil'scher Ungleichheitskoeffizient)	= 0,81

Der Prognosewert für das Jahr 1982 beträgt 7,73 l. Die willkürliche Festsetzung von $\hat{x}_1 = 6,8$ wirkt sich mit einem Gewicht von $0,7^{12} \approx 1,4\% = 0,11$ auf die Prognose aus, d. h. die Forderung „genügend großer Zeitreihenumfang“ ist praktisch erfüllt.

Bei der Errechnung der Fehlermaße wurden die Werte x_1 , \hat{x}_1 und e_1 nicht berücksichtigt, da sie der Initialisierung dienen. Der Beobachtungszeitraum weist damit $T = 11$ Werte auf. Die errechneten Größen sind auf zwei Stellen nach dem Komma gerundet. Der Ungleichheitskoeffizient von Theil ($U = 0,81$) läßt erkennen, daß das Verfahren dem Beispiel angemessen ist und ein besseres Ergebnis als die „naive“ Prognose erzielt.

III. Das Trendmodell der exponentiellen Glättung

1. Zielfunktion und Prognosegleichung

Das konstante Modell der exponentiellen Glättung versagt, wenn die Zeitreihe einen Trend oder einen Saisonzyklus aufweist. Der Prognosewert hinkt infolge seiner Konstruktion als gewogener Durchschnitt aller vergangenen Zeitreihenwerte dem aktuellen Trend hinterher, eine Eigenschaft, die zu schlechten Prognosen führt. Um das Prinzip der exponentiellen Glättung auch für Zeitreihen mit *Trend* nutzbar zu machen (Saisonzyklen werden im Kapitel D. behandelt), gehen wir davon aus, daß die Zeitreihe in erster Näherung durch einen linearen Prozeß dargestellt werden kann, der – wie im konstanten Modell – von einer Störvariablen u additiv überlagert ist. Der Prozeß wird durch die beiden Parameter a (Grundwert) und b (Trendwert, d. h. Änderung des Grundwertes pro Periode) charakterisiert und durch folgende Gleichung beschrieben:

$$(13) \quad x_{t-i} = a - b \cdot i + u_{t-i} \quad (i = 0, 1, \dots, t-1)$$

Das Minuszeichen des Trendwertes wurde hier aus rein didaktischen Gründen gewählt, um das retrograde Vorgehen von der Gegenwart in die Vergangenheit anzudeuten.

Wie im konstanten Modell minimieren wir auch jetzt die Summe der exponentiell gewogenen Abweichungsquadrate, um die Parameter a und b zu schätzen.

$$(14) \quad Z = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i u_{t-i}^2 = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i (x_{t-i} - a + bi)^2 \rightarrow \min!$$

Zur einfacheren Schreibweise erhalten die Gewichte das Symbol g_{t-i}

$$(15) \quad g_{t-i} = \alpha(1-\alpha)^i \quad (i = 0, 1, \dots, t-1)$$

Die beiden partiellen Ableitungen von Z nach a und b werden gleich Null gesetzt.

$$(16a) \quad \frac{\partial Z}{\partial a} = -2 \sum_{i=0}^{t-1} g_{t-i} (x_{t-i} - a + bi) = 0$$

$$(16b) \quad \frac{\partial Z}{\partial b} = 2 \sum_{i=0}^{t-1} g_{t-i} (x_{t-i} - a + bi) i = 0$$

Daraus ergeben sich folgende Bestimmungsgleichungen für a und b ¹

¹ In Übereinstimmung mit der ökonomischen Literatur werden in diesem Buch nicht nur die Prognosewerte sondern auch die mit dem Schätzverfahren geschätzten Parameterwerte (hier: \hat{a} und \hat{b} mit einem „^“ versehen).

$$(17a) \quad \hat{a} \cdot \sum_{i=0}^{t-1} g_{t-i} - \hat{b} \cdot \sum_{i=0}^{t-1} g_{t-i} \cdot i = \sum_{i=0}^{t-1} g_{t-i} \cdot x_{t-i}$$

$$(17b) \quad \hat{a} \cdot \sum_{i=0}^{t-1} g_{t-i} \cdot i - \hat{b} \cdot \sum_{i=0}^{t-1} g_{t-i} \cdot i^2 = \sum_{i=0}^{t-1} g_{t-i} \cdot x_{t-i} \cdot i$$

Die Auswertung der in (17a) und (17b) vorkommenden Summen ist für längere Zeitreihen außerordentlich mühsam. Der folgende Ansatz vereinfacht die Bestimmungsgleichungen für a und b , gilt jedoch nur für *lange* Zeitreihen, bei denen man unbedenklich annehmen kann, daß der Zeitindex i von 0 bis ∞ läuft. Der Ansatz ist ausführlich in (Hansmann 1980 b, S. 323–325) dargestellt.

Zunächst wissen wir aus Gleichung (8), daß der Faktor von \hat{a} in (17a) gleich Eins gesetzt werden kann, weil die unendliche geometrische Reihe mit $0 < \alpha \leq 1$ absolut konvergiert.

Auch der Faktor von \hat{b} in (17a) konvergiert absolut

$$(18) \quad \sum_{i=0}^{t-1} g_{t-i} \cdot i = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i \cdot i = \frac{1-\alpha}{\alpha}$$

weil sich die Summe $S_u = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i \cdot i$ folgendermaßen umformen läßt:

$$\begin{aligned} S_u &= \alpha(1-\alpha) + 2\alpha(1-\alpha)^2 + 3\alpha(1-\alpha)^3 + \dots \\ (1-\alpha) S_u &= \alpha(1-\alpha)^2 + 2\alpha(1-\alpha)^3 + 3\alpha(1-\alpha)^4 + \dots \\ S_u - (1-\alpha) S_u &= \alpha(1-\alpha) + \alpha(1-\alpha)^2 + \alpha(1-\alpha)^3 + \dots = \frac{\alpha(1-\alpha)}{\alpha} \\ S_u &= \frac{(1-\alpha)}{\alpha} \end{aligned}$$

Es bleibt noch die rechte Seite von (17a) zu betrachten.

Den Ausdruck

$$\sum_{i=0}^{t-1} g_{t-i} \cdot x_{t-i}$$

haben wir im konstanten Modell mit \hat{S} bezeichnet. Im Unterschied zu einem noch zu entwickelnden Glättungsoperator fügen wir den Index 1 an und zusätzlich den Zeitpunkt t , an dem er berechnet wird.

$$(19) \quad \hat{S}_t^1 = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i x_{t-i} = \alpha x_t + (1-\alpha) \hat{S}_{t-1}^1$$

\hat{S}_t^1 wird Glättungsoperator erster Ordnung genannt. Damit haben wir die Bestimmungsgleichung (17a) vereinfacht zu

$$(20) \quad \hat{a} - \frac{(1-\alpha)}{\alpha} \hat{b} = \hat{S}_t^1$$

Die vereinfachte Schreibweise der Bestimmungsgleichung (17b) ist etwas aufwendig herzuleiten. Es sollen deshalb hier nur die Ergebnisse wiedergegeben werden. Die vollständige Herleitung findet man in (Hansmann 1980 b, S. 324–325).

Der Faktor von \hat{b} in (17b) vereinfacht sich zu

$$(21) \quad \sum_{i=0}^{t-1} g_{t-i} \cdot i^2 = \frac{(1-\alpha)(2-\alpha)}{\alpha^2}$$

Die rechte Seite von (17b) lässt sich schreiben

$$(22) \quad \sum_{i=0}^{t-1} g_{t-i} \cdot x_{t-i} \cdot i = \frac{\hat{S}_t^2}{\alpha} - \hat{S}_t^1$$

wobei \hat{S}_t^2 ein Glättungsoperator zweiter Ordnung mit folgender Definition ist:

$$(23) \quad \hat{S}_t^2 = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i \hat{S}_{t-i}^1 = \alpha \hat{S}_t^1 + (1-\alpha) \hat{S}_{t-1}^2$$

Damit kann man (17b) wie folgt schreiben

$$(24) \quad \frac{1-\alpha}{\alpha} \hat{a} - \frac{(1-\alpha)(2-\alpha)}{\alpha^2} \hat{b} = \frac{\hat{S}_t^2}{\alpha} - \hat{S}_t^1$$

Aus (20) und (24) ergeben sich nun die Schätzwerte für die Parameter a und b

$$(25a) \quad \hat{a} = 2 \hat{S}_t^1 - \hat{S}_t^2$$

$$(25b) \quad \hat{b} = \frac{\alpha}{1-\alpha} (\hat{S}_t^1 - \hat{S}_t^2)$$

Man benötigt also zur numerischen Auswertung nur den Glättungsparameter α und die beiden Glättungsoperatoren \hat{S}_t^1 und \hat{S}_t^2 , die sich aus den Gleichungen (19) und (23) rekursiv errechnen lassen, wenn die Anfangswerte \hat{S}_0^1 und \hat{S}_0^2 bekannt sind.

Das Trendmodell ist daher in der *Anwendung* trotz der komplizierten Herleitung von (25a) und (25b) nicht viel aufwendiger als das konstante Modell.

Die Prognosegleichung des Trendmodells für m Perioden im voraus lautet also

$$(26) \quad \hat{x}_{t+m} = \hat{a}_t + \hat{b}_t \cdot m \quad (m = 1, 2, \dots)$$

Sie wird im folgenden Anwendungsbeispiel benutzt.

2. Die Prognose des Verbrauchs von Dieselkraftstoff in der BRD

Als Anwendungsbeispiel betrachten wir die Zeitreihe des jährlichen Dieselkraftstoffverbrauchs in der Bundesrepublik Deutschland während der Jahre 1960 bis 1980 (entnommen aus: Verkehr in Zahlen 1981, S. 230–231). Die Daten sind in Tabelle 4 wiedergegeben und in Abbildung 1 zusammen mit den Prognosewerten graphisch dargestellt.

Tabelle 4

Verbrauch von Dieselkraftstoff in der BRD p.a. (in Mio-Tonnen)

Jahr	t	x_t
1960	1	3,160
1961	2	3,715
1962	3	3,933
1963	4	4,116
1964	5	4,568
1965	6	4,800
1966	7	5,174
1967	8	5,050
1968	9	5,558
1969	10	5,750
1970	11	6,320
1971	12	6,532
1972	13	6,900
1973	14	7,330
1974	15	6,750
1975	16	7,060
1976	17	7,650
1977	18	8,350
1978	19	8,900
1979	20	9,650
1980	21	9,880

Aus Abbildung 1 ist zu ersehen, daß der Verbrauch einen langfristig steigenden, annähernd linearen Trend aufweist, so daß wir das Trendmodell der exponentiellen Glättung anwenden können.

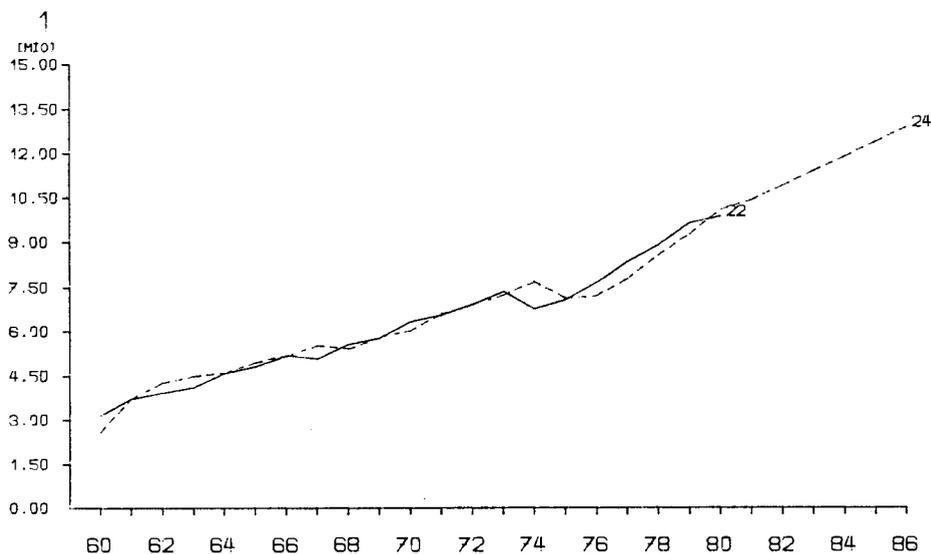


Abbildung 1
Verbrauch von Dieseldieselkraftstoff in der BRD p.a. (Zeitreihenwerte und Prognosewerte)

Die durchgezogene Linie entspricht den wahren Zeitreihenwerten, die gestrichelte Linie den Prognosewerten. Die Zeitreihe enthält 21 Werte, die in Mio-Tonnen angegeben sind. Als Glättungsparameter wird $\alpha = 0,5$ gewählt. Die Prognose soll für ein Jahr im voraus erstellt werden.

Zunächst sind die Anfangswerte \hat{S}_0^1 und \hat{S}_0^2 der Glättungsoperatoren zu bestimmen. Dazu setzt man $x_2 - x_1 = \hat{b}_0$ und $x_1 - \hat{b}_0 = \hat{a}_0$, weil man über keine älteren Daten verfügt. \hat{a}_0 und \hat{b}_0 werden in (25a) und (25b) eingesetzt und daraus \hat{S}_0^1 und \hat{S}_0^2 bestimmt. Dabei ergeben sich folgende Werte:

$$\hat{a}_0 = 2,605 \quad \hat{b}_0 = 0,555 \quad \hat{S}_0^1 = 2,050 \quad \hat{S}_0^2 = 1,495$$

Da man die Zeitreihenwerte von 1960 (x_1) und 1961 (x_2) für die Initialisierung des Verfahrens verwendet hat, erfolgt die erste richtige Prognose für 1962 (Periode $t = 3$). Die Glättungsoperatoren werden gemäß (19) und (23) fortgeschrieben.

$$\begin{aligned}\hat{S}_1^1 &= 0,5 \cdot 3,160 + 0,5 \cdot 2,050 = 2,605 & \hat{a}_1 &= 3,160 \\ \hat{S}_1^2 &= 0,5 \cdot 2,605 + 0,5 \cdot 1,495 = 2,050 & \hat{b}_1 &= 0,555 \\ \hat{S}_2^1 &= 0,5 \cdot 3,715 + 0,5 \cdot 2,605 = 3,160 & \hat{a}_2 &= 3,715 \\ \hat{S}_2^2 &= 0,5 \cdot 3,160 + 0,5 \cdot 2,050 = 2,605 & \hat{b}_2 &= 0,555\end{aligned}$$

Mit dem Grundwert \hat{a}_2 und dem Trendwert \hat{b}_2 – berechnet auf der Basis des Jahres 1961 – prognostiziert man mit (26) den Dieselkraftstoffverbrauch für 1962 ($m = 1$)

$$\hat{x}_{t+m} = \hat{x}_3 = 3,715 + 0,555 = 4,27$$

d. h. 4,27 Mio-Tonnen. Der wahre Wert liegt hingegen bei 3,933 Mio. Tonnen, so daß ein Prognosefehler von 337.000 Tonnen zu registrieren ist.

Für die Prognose des Jahres 1963 ($t = 4$) wird der neuste Zeitreihenwert $x_3 = 3,933$ verwendet. Die Glättungsoperatoren und Parameter erhalten folgende Werte:

$$\begin{aligned}\hat{S}_3^1 &= 0,5 \cdot 3,9330 + 0,5 \cdot 3,160 = 3,5465 & \hat{a}_3 &= 4,0172 \\ \hat{S}_3^2 &= 0,5 \cdot 3,5465 + 0,5 \cdot 2,605 = 3,0758 & \hat{b}_3 &= 0,4707\end{aligned}$$

Die Prognosegleichung (26) liefert für 1963 den Prognosewert

$$\hat{x}_4 = 4,0172 + 0,4707 = 4,4879$$

während der tatsächliche Zeitreihenwert $x_4 = 4,116$ beträgt.

In gleicher Weise lassen sich – jährlich fortschreitend – die Prognosewerte für die Jahre 1964–1981 ermitteln. Sie sind in der folgenden Tabelle 5 zusammengestellt.

Ohne Berücksichtigung des Anfangsjahres 1960 erhält man folgende Gütemaße der Prognose:

MAA	(mittlere absolute Abweichung)	= 0,25
WMQA	(Wurzel aus mittlerer quadrierter Abweichung)	= 0,34
U	(Theil'scher Ungleichheitskoeffizient)	= 0,783

Soll eine Prognose für mehrere Jahre im voraus erstellt werden, so wird das letzte Jahr der Zeitreihe (1980) als Basis benutzt ($t = 21$) und für die Prognosewerte Gleichung (26) mit $m > 1$ zugrunde gelegt. Man erhält damit eine lineare Fortschreibung der Prognose (vgl. Abbildung 1).

Als Prognosewert für 1984 ergibt sich zum Beispiel

$$\hat{x}_{21+4} = \hat{x}_{25} = 9,9369 + 4 \cdot 0,4937 = 11,91 \text{ Mio. Tonnen.}$$

Da die Prognose ab 1980 nicht mehr an jährliche Änderungen der Zeitreihe angepaßt

Tabelle 5

Jahr	t	x_t	\hat{x}_t	e_t	\hat{a}_t	\hat{b}_t
1960	1	3,1600	—	—	3,1600	0,5550
1961	2	3,7150	3,7150	0,0000	3,7150	0,5550
1962	3	3,9330	4,270	-0,3370	4,0173	0,4708
1963	4	4,1160	4,4880	-0,3720	4,2090	0,3778
1964	5	4,5680	4,5867	-0,0187	4,5727	0,3731
1965	6	4,8000	4,9457	-0,1457	4,8364	0,3366
1966	7	5,1740	5,1731	0,0009	5,1738	0,3369
1967	8	5,0500	5,5106	-0,4606	5,1652	0,2217
1968	9	5,5580	5,3869	0,1711	5,5152	0,2645
1969	10	5,7500	5,7797	-0,0297	5,7574	0,2571
1970	11	6,3200	6,0145	0,3055	6,2436	0,3334
1971	12	6,5320	6,5771	-0,0451	6,5433	0,3222
1972	13	6,9000	6,8654	0,0346	6,8914	0,3308
1973	14	7,3300	7,2222	0,1078	7,3030	0,3578
1974	15	6,7500	7,6608	-0,9108	6,9777	0,1301
1975	16	7,0600	7,1078	-0,0478	7,0719	0,1181
1976	17	7,6500	7,1901	0,4599	7,5350	0,2331
1977	18	8,3500	7,7681	0,5819	8,2045	0,3786
1978	19	8,9000	8,5831	0,3169	8,8208	0,4578
1979	20	9,6500	9,2786	0,3714	9,5571	0,5507
1980	21	9,8800	10,1078	-0,2278	9,9369	0,4937
1981	22	—	10,4310	—	—	—

werden kann, nimmt die Genauigkeit mit zunehmendem Prognosehorizont ab. Für langfristige Voraussagen sind daher andere Verfahren, die in Kapitel F., S. 104 ff., beschrieben werden, anzuwenden.

IV. Die dynamische Anpassung des Glättungsparameters α

Wir hatten auf S. 31 erkannt, daß ein konstanter Wert für α den Anforderungen einer sich verändernden Zeitreihenstruktur nicht gerecht wird. Selbst wenn wir α so bestimmen, daß die Summe der quadrierten Ex-post-Prognosefehler während des Beobachtungszeitraums minimiert wird²

$$(27) \quad Z = \sum_{i=0}^{t-1} e_{t-i}^2 = \sum_{i=0}^{t-1} (x_{t-i} - \alpha x_{t-i-1} - (1-\alpha) \hat{x}_{t-i-1})^2 \rightarrow \min!$$

2 (27) stellt den Ansatz für das konstante Modell dar.

indem wir Z nach α ableiten und das Minimum iterativ errechnen, muß dieser Wert für die Zukunft nicht der beste sein. Es empfiehlt sich daher, α Jahr für Jahr neu an die Entwicklung anzupassen. Zwei dynamische Ansätze sollen im folgenden betrachtet werden.

1. Die Methoden von Chow und Smith

Das einfachste Verfahren zur Anpassung von α wurde von Chow (1965, S. 314 ff.) vorgeschlagen. Danach wird jede Prognose mit drei verschiedenen Werten von α durchgeführt, z. B. $\alpha_1 = 0,1$ $\alpha_2 = 0,15$ $\alpha_3 = 0,2$. Stellt sich am Ende des Prognosezeitraums heraus, daß mit $\alpha_1 = 0,1$ das genaueste Prognoseergebnis erzielt wurde, so wird der α -Katalog nach unten verschoben und die nächste Prognose mit $\alpha_1 = 0,05$ $\alpha_2 = 0,1$ $\alpha_3 = 0,15$ erstellt. Wäre hingegen $\alpha_3 = 0,2$ der beste Wert gewesen, so würde mit $\alpha_1 = 0,15$ $\alpha_2 = 0,2$ und $\alpha_3 = 0,25$ weiter prognostiziert. α wird also immer in der Richtung angepaßt, die die beste Prognose geliefert hat.

In praktischen Untersuchungen hat sich nun gezeigt, daß α sehr *oszillieren* kann und die Ausschläge der Prognosewerte nicht gut gedämpft werden. Das Verfahren kann daher nur als ein erster Ansatz zur dynamischen Anpassung von α angesehen werden.

Im folgenden wird das Verfahren von Smith (1974, S. 421 ff.) näher erläutert, mit dem der Verfasser – in einer eigenen Programmversion – bei verschiedenartigen Anwendungsbeispielen die besten Erfahrungen gemacht hat. Das Verfahren baut auf einer Arbeit von Trigg und Leach (1967, S. 53 ff.) auf.

Smith benutzt zur Anpassung von α zwei Größen

– den mittleren Prognosefehler
$$MA_t = \frac{1}{t} \sum_{i=0}^{t-1} e_{t-i}$$

– den mittleren absoluten Prognosefehler
$$MAA_t = \frac{1}{t} \sum_{i=0}^{t-1} |e_{t-i}|$$

allerdings exponentiell gewichtet, da beide Prognosefehler mit exponentieller Glättung fortgeschrieben werden:

$$(27) \quad MA_t = \beta e_t + (1-\beta)MA_{t-1} \quad (0 < \beta < 1)$$

$$(28) \quad MAA_t = \beta |e_t| + (1-\beta)MAA_{t-1} \quad (0 < \beta < 1)$$

Der Glättungsparameter α_t zum Zeitpunkt t ergibt sich als Quotient dieser Prognosefehler

$$(29) \quad \alpha_t = \frac{|MA_t|}{MAA_t}$$

Diese Definition von α_t beruht auf folgender Überlegung:

Falls die Prognosefehler nur auf Zufallseinflüssen beruhen, gilt $E[e_{t-i}] = 0$ und MA_t wird nahe bei Null liegen, weil sich die Zufallseinflüsse kompensieren. Das Prognosemodell ist also richtig und braucht nicht durch den jüngsten Zeitreihenwert korrigiert zu werden. Ein kleiner Wert für α_t ist angemessen und wird durch (29) auch gewährleistet.

Falls $|MA_t| = MAA_t$ gilt, so bedeutet dies, daß alle Prognosen über den tatsächlichen Zeitreihenwerten oder alle darunter liegen. Das bisherige Prognosemodell ist also falsch und der jüngste Zeitreihenwert die beste Prognose für die Zukunft. α_t sollte in diesem Extremfall gleich Eins sein, was durch Gleichung (29) auch erreicht wird.

Die Realität wird zwischen diesen beiden Extremfällen liegen, so daß α_t zwischen Null und Eins variiert. Dadurch vereinigt die Definition (29) zwei Vorzüge:

- a) α_t wird in jeder Periode neu bestimmt als Funktion der geglätteten Prognosefehler
- b) Die Variation von α_t vollzieht sich in den oben beschriebenen Grenzen ($0 < \alpha_t \leq 1$).

Bei größeren zufallsbedingten Schwankungen der Zeitreihe schwankt α_t ebenfalls stark, was zu einem instabilen Prognosesystem führen kann. Um dies zu verhindern, wird das mit (29) errechnete α_t von Smith exponentiell geglättet.

$$(30) \quad \bar{\alpha}_t = \gamma \alpha_t + (1-\gamma) \bar{\alpha}_{t-1} \quad (0 < \gamma < 1)$$

$\bar{\alpha}_t$ geht in die Prognosegleichung ein. Der Rechenaufwand wird durch die dynamische Anpassung von α nur unwesentlich erhöht, lediglich vier Werte (MA_t , MAA_t , β und γ) sind zusätzlich zu speichern. Geeignete Werte für β und γ sind problemabhängig und müssen aufgrund von numerischen Erfahrungen gewonnen werden. Darauf ist beim Anwendungsbeispiel noch näher einzugehen.

In der Regel zeichnet sich das Verfahren von Smith durch recht gute Dämpfungseigenschaften aus und ist dem einfachen Verfahren der exponentiellen Glättung überlegen. Um einen Vergleich zu ermöglichen, wollen wir es nun an unserem vorigen Anwendungsbeispiel testen.

2. Die Prognose des Dieselkraftstoffverbrauchs mit dynamischer Anpassung von α

Die Daten der Zeitreihe sind Tabelle 4 zu entnehmen. Für $\bar{\alpha}_1$ wählen wir den Wert 0,59 und die anderen Startwerte werden wie folgt besetzt:

$$MA_1 = 0 \quad MAA_1 = 0,2 \quad \beta = 0,5 \quad \gamma = 0,1$$

Das Ergebnis der Prognose mit α -Anpassung ist in Tabelle 6 zusammengestellt.

Tabelle 6

Jahr	t	x_t	\hat{x}_t	e_t	\hat{a}_t	\hat{b}_t	$\bar{\alpha}_t$
1960	1	3,1600	—	—	3,1600	0,5550	0,59
1961	2	3,7150	3,7150	0,0000	3,6920	0,4670	0,53
1962	3	3,9330	4,1590	-0,2260	3,9616	0,3331	0,54
1963	4	4,1160	4,2946	-0,1786	4,1581	0,2967	0,57
1964	5	4,5680	4,4548	0,1132	4,5547	0,3607	0,53
1965	6	4,8000	4,9154	0,1154	4,8121	0,2844	0,53
1966	7	5,1740	5,0964	0,0776	5,1562	0,3045	0,48
1967	8	5,0500	5,4607	-0,4107	5,1471	0,1698	0,51
1968	9	5,5580	5,3169	0,2411	5,5056	0,2491	0,47
1969	10	5,7500	5,7547	-0,0047	5,7401	0,2171	0,43
1970	11	6,3200	5,9572	0,3628	6,1896	0,2559	0,46
1971	12	6,5320	6,4455	0,0865	6,5174	0,3030	0,50
1972	13	6,9000	6,8204	0,0796	6,8917	0,3581	0,53
1973	14	7,3300	7,2498	0,0802	7,3255	0,4234	0,57
1974	15	6,7500	7,7488	-0,9988	6,9456	0,1447	0,60
1975	16	7,0600	7,0903	-0,0303	7,0674	0,1439	0,62
1976	17	7,6500	7,2114	0,4386	7,5892	0,3223	0,59
1977	18	8,3500	7,9115	0,4385	8,2673	0,4441	0,59
1978	19	8,9000	8,7114	0,1886	8,8715	0,5207	0,61
1979	20	9,6500	9,3922	0,2578	9,6174	0,6429	0,64
1980	21	9,8800	10,2603	-0,3803	9,9396	0,5335	0,60
1981	22	—	10,4731	—	—	—	—

MAA	(mittlere absolute Abweichung)	= 0,235
WMQA	(Wurzel aus mittlerer quadrierter Abweichung)	= 0,326
U	(Theil'scher Ungleichheitskoeffizient)	= 0,736

Man erkennt, daß alle drei Gütemaße gegenüber der Prognose mit konstantem α leicht verbessert sind. Eine weitere Steigerung der Genauigkeit durch die dynamische Anpassung von α war nicht möglich. Das mag daran liegen, daß schon die einfache exponentielle Glättung bei dieser Zeitreihe sehr gute Prognoseergebnisse lieferte, die nur schwer zu übertreffen sind (vgl. hierzu die Abbildungen 1 und 2). Bei anderen Zeitreihen ist die Verbesserung deutlicher ausgefallen (vgl. das Beispiel in Hansmann 1980 b, S. 380).

Es soll an dieser Stelle nochmals betont werden, daß die Prognoseergebnisse natürlich sehr stark von der „geschickten“ Wahl der Parameter β und γ abhängen, was beträchtliche numerische Erfahrungen voraussetzt.

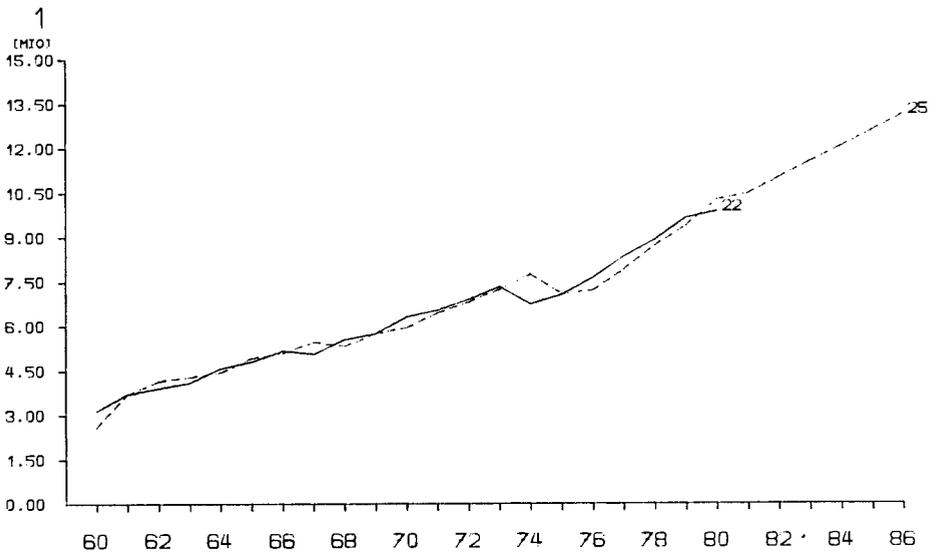


Abbildung 2
Prognose des Dieselkraftstoffverbrauchs mit α -Anpassung

V. Erweiterung der exponentiellen Glättung

1. Exponentielle Glättung höherer Ordnung

Neben dem *konstanten* Modell und dem *linearen Trendmodell* sind weitere Modelle mit Polynomen höherer Ordnung (z. B. ein quadratischer Trend) denkbar.

Brown und Meyer (1961, S. 673 ff.) haben in ihrem Fundamentaltheorem gezeigt, daß die $n + 1$ Koeffizienten eines Polynoms n -ter Ordnung

$$(31) \quad \hat{x}_{t+m} = \hat{a}_0 + \hat{a}_1 m + \hat{a}_2 m^2 + \dots + \hat{a}_n m^n$$

durch eine Linearkombination von $n + 1$ Glättungsoperatoren \hat{S}_t^j ($j = 1, 2, \dots, n + 1$) erster bis $(n + 1)$ -ter Ordnung bestimmt werden können, wobei gilt

$$(32) \quad \hat{S}_t^j = \alpha \hat{S}_t^{j-1} + (1-\alpha) \hat{S}_{t-1}^j \quad (j = 2, 3, \dots, n + 1)$$

In der Praxis werden Prognosen mit der exponentiellen Glättung zweiter oder höherer Ordnung jedoch kaum angewandt, weil

- der *Rechenaufwand* mit höherer Ordnung exponentiell zunimmt,

- die *ökonomische* Interpretation der Koeffizienten nicht mehr möglich bzw. unplausibel ist,
- die *Prognoseergebnisse* nicht oder nur unwesentlich verbessert werden.

Aus diesen Gründen und wegen des mathematischen Aufwands sei auf eine eingehende Darstellung verzichtet (vgl. dazu Hüttner 1982, S. 107).

2. Mehr-Parameter-Modelle

Das Prognosesystem der exponentiellen Glättung wird in seiner Reagibilität nur von α gesteuert. Mehrere Autoren halten dies für unzureichend und versprechen sich eine größere Flexibilität durch die Einführung weiterer Parameter. Holt (1960) hat ein Zwei-Parameter-Modell entwickelt, das sich leicht auf drei Parameter erweitern läßt. Die Prognosegleichung lautet

$$(33) \quad \hat{x}_{t+1} = \hat{x}_t + \alpha e_t + \beta(e_t - e_{t-1}) + \gamma \sum_{i=0}^{t-1} e_{t-i}$$

mit $0 < \alpha \leq 1$ und $\beta > 0, \gamma \geq 0$

In diesem Modell wird neben dem Prognosefehler e_t auch seine zeitliche Änderung und der kumulierte Prognosefehler berücksichtigt. Das konstante Modell von Brown ist als Spezialfall in (33) enthalten, wenn man $\beta = \gamma = 0$ setzt.

Über die Wahl der drei Glättungsparameter liegen bisher wenige Ergebnisse vor. Numerische Erfahrungen des Verfassers zeigen keine entscheidende Verbesserung gegenüber der einfachen exponentiellen Glättung mit α -Anpassung.

Wegen weiterer Informationen und Anwendungsbeispiele sei auf die Arbeit von Scheer (Scheer 1980, insbesondere S. 109 ff.) verwiesen.

VI. Kritische Zusammenfassung

Thesenartig sollen hier noch einmal die herausgearbeiteten Vorzüge und Unzulänglichkeiten der exponentiellen Glättung zusammengefaßt werden.

Folgende Vorteile sind festzuhalten:

- Das Verfahren ist leicht *programmierbar*, da nur elementare Rechenoperationen zur Bestimmung des Prognosewertes ausgeführt werden müssen.
- Der *Speicherbedarf* ist niedrig; man braucht nicht alle vergangenen Zeitreihenwerte, sondern nur den aktuellen Zeitreihenwert, die aktuellen Glättungsoperatoren \hat{S}_t^1 und \hat{S}_t^2 sowie α zu speichern.

- Das Prognosesystem wird durch einen *einzigsten* Parameter (α), der sich *selbst* einer veränderten Zeitreihenstruktur anpassen kann, gesteuert und ist damit willkürlichen Eingriffen weitgehend entzogen.
- Die Verfahrensschritte sind leicht durchschaubar und plausibel, so daß die Prognoseergebnisse und ihre Begründung auch Nicht-Experten verständlich gemacht werden können.

Diese Vorzüge haben zu der außergewöhnlich weiten Verbreitung der exponentiellen Glättung in der Unternehmenspraxis geführt, insbesondere in der Absatz- und Materialbedarfsprognose der Industrieunternehmen.

Demgegenüber dürfen jedoch bestimmte *Unzulänglichkeiten* des Verfahrens nicht unerwähnt bleiben:

- Das Verfahren berücksichtigt keine kausalen Einflußfaktoren auf die Zeitreihe, sondern nur die „Zeit“. Eine Zeitreihe nur aus sich selbst heraus zu prognostizieren, ist aber vom theoretischen Standpunkt aus etwas problematisch, so daß das Verfahren nur bei mangelnder Kenntnis von quantifizierbaren Einflußfaktoren gerechtfertigt erscheint.
- Es kann als unbefriedigend empfunden werden, daß der Glättungsparameter α nicht objektiv bestimmt werden kann, sondern mehr oder weniger dem „Fingerspitzengefühl“ anheimgestellt bleibt. An dieser Problematik kann auch die dynamische Anpassung von α nur graduell etwas ändern, da das Ergebnis vom gewählten Anfangswert α_0 und den Parametern β und γ abhängt, die ihrerseits bestimmt werden müssen. In Kapitel E. werden wir einen anderen Ansatz für dieses Problem kennenlernen.
- Der Haupteinwand gegen die exponentielle Glättung bezieht sich jedoch auf die *Zielfunktion*, die vorschreibt, die Summe der exponentiell gewogenen quadratischen Schätzfehler zu minimieren.

Sind diese Schätzfehler reine Zufallsschwankungen, so erübrigt sich eine Gewichtung, da die Zufallseinwirkungen in jeder Periode die gleiche Verteilung aufweisen. Werden sie jedoch auch von systematischen Faktoren beeinflusst, dann ist das Modell falsch spezifiziert und man müßte zur Regressionsanalyse (Kapitel G.) übergehen.

Dieser Einwand ist jedoch nur bei *strenger* Geltung der Zeitstabilitätshypothese berechtigt. In der Praxis muß man aber immer mit Strukturveränderungen der Zeitreihe rechnen, und dann kann es durchaus sinnvoll sein, die aktuelle Struktur stärker zu gewichten als die frühere Entwicklung.

Die exponentielle Gewichtung der Schätzfehler ist in dieser Hinsicht ein – relativ grober – Mechanismus, der dies bewerkstelligt und das Prognosesystem ständig aktualisiert.

D. Prognosen bei saisonbehafteten Zeitreihen

I. Überblick

Mit den bisher dargestellten Methoden können wir stationäre und trendbehaftete Zeitreihen analysieren. Besonders bei industriellen Absatzreihen treten jedoch häufig zyklische Schwankungen auf, deren Länge sich bestimmten Kalenderzeiträumen zuordnen läßt (Kalenderjahr, Quartal, Monat, Woche).

Schwankungen mit der Länge eines Jahres treten z. B. bei Bier, nicht alkoholischen Erfrischungsgetränken, Speiseeis, Heizöl, Benzin, Publikumszeitschriften, Sonnenschutzmittel, Spielwaren und allen Produkten auf, die ein besonderes Weihnachtsgeschäft haben.

Beispiele für Monatszyklen sind Banküberweisungen oder der Bargeldumlauf in der Wirtschaft.

Zeitreihen mit solchen zyklischen Schwankungen heißen *saisonbehaftete Zeitreihen*. Ihre Analyse und Prognose erfordert besondere Prognoseverfahren, die sog. Saisonverfahren.

Man kann die in der Literatur beschriebenen Verfahren grob in zwei Klassen einteilen:

1. Saisonbereinigung auf der Grundlage gleitender Durchschnitte
2. Abbildung der Saisonkomponente durch Sinus- bzw. Kosinus-Funktionen.

Zu der ersten Gruppe gehören vor allem das ursprüngliche Bundesbankverfahren (Deutsche Bundesbank 1961; Brockhoff 1977, S. 99), die Census-Methode II (Shiskin/Eisenpress 1957), sowie das Verfahren von Winters (1960).

Ausgangspunkt ist die Bildung eines gleitenden Durchschnitts mit der Länge eines vollständigen Saisonzyklus; bei monatlich erhobenen Daten und einer jährlichen Saison werden also 12 Werte im gleitenden Durchschnitt erfaßt. Anschließend wird ein monatlicher Saisonindex als Quotient aus dem Zeitreihenwert dieses Monats und dem gleitenden Durchschnitt ermittelt.

Die Verfahren unterscheiden sich dann nur noch durch die unterschiedliche iterative Verbesserung der Saisonindizes. Stellvertretend für diese Gruppe wird im nächsten Abschnitt das Verfahren von Winters näher erläutert.

Zu der zweiten Gruppe gehören das Berliner Verfahren (Nullau et al. 1969), das Verfahren von Harrison (1965) und – mit etwas anderer Zielrichtung – die Spektralanalyse. Die Verfahren dieser Gruppe zeichnen sich durch einen hohen mathematischen Schwierigkeitsgrad (Fourier-Analyse) und großen numerischen Aufwand aus, ohne in ihren Ergebnissen signifikant besser zu sein als die einfacheren Verfahren des gleitenden Durchschnitts. Lewandowski (1974, S. 111 ff.) und Schläger (in: Mertens 1981, S. 75 ff.) berichten z. B. von einem wesentlich höheren Rechenaufwand des Verfahrens von Harrison gegenüber dem von Winters mit nur leicht verbesserten Prognoseergebnissen. Diese Erfahrungen sind vom Verfasser, der für beide Verfahren entsprechende

Software entwickelt und an einer Univac 1100-Anlage implementiert hat, voll bestätigt worden. In einem konkreten Anwendungsfall (Absatzprognose von Publikumszeitschriften) war Harrison Winters sogar unterlegen (Hansmann 1980 c, S. 235).

Leistungsfähiger erscheint aus dieser Gruppe die Spektralanalyse, die im Abschnitt III. dieses Kapitels besprochen wird (vgl. S. 53 ff.).

II. Das Saisonverfahren von Winters

Das Verfahren von Winters zeichnet sich neben seiner mathematischen Einfachheit durch enge Verwandtschaft mit der exponentiellen Glättung aus. Es kann als eine Weiterentwicklung der exponentiellen Glättung für Saisonreihen angesehen werden und baut direkt auf den im vorigen Kapitel gewonnenen Erkenntnissen auf.

1. Die exponentielle Glättung als Basis der Parameterschätzung

Ausgangspunkt des Verfahrens von Winters ist das lineare Trendmodell der exponentiellen Glättung (vgl. S. 33)

$$(26) \quad \hat{x}_{t+m} = \hat{a}_t + \hat{b}_t \cdot m \quad (m = 1, 2, \dots)$$

Das Trendmodell wird nun mit einer Saisonkomponente *multiplikativ* verknüpft, die die Überlagerung des Trends durch die Saison beschreibt:

$$(34) \quad \hat{x}_{t+m} = (\hat{a}_t + \hat{b}_t \cdot m) \cdot \hat{s}_{lj} \quad (m = 1, 2, \dots)$$

\hat{s}_{lj} ist der Saisonfaktor zum Zeitpunkt $l = t - L + m \bmod L$ mit der Nummer $j = (t + m) \bmod L$

Hierbei ist t der Prognosezeitpunkt (Gegenwart), L die Länge des Saisonzyklus und $(m \bmod L)$ der Rest, der nach Division von m durch L bleibt.

Da die Symbolik etwas verwirrend aussieht, sei sie noch durch ein Beispiel erläutert:

Gegenwart $\hat{=}$ Januar 1982 $m \hat{=}$ 3 Monate $L \hat{=}$ 12 Monate

Die Zeitreihe beginnt im Januar 1980 mit $t = 1$, so daß der Januar 1982 $t = 25$ ist. Gesucht ist der Saisonfaktor für April 1982.

$$\begin{aligned} l &= 25 - 12 + 3 \bmod 12 & &= 16 \\ j &= (25 + 3) \bmod 12 & &= 4 \end{aligned}$$

Der Index l zählt also *linear vom Beginn der Zeitreihe*, während j zyklisch (1 bis 12) numeriert. Da der April der 4. Monat eines Jahres ist und der April 1981 der 16. Mo-

nat seit Beginn der Zeitreihe, verwendet man als Saisonfaktor für April 1982:

$$\hat{s}_{16,4} = \hat{s}_{\text{April 1981}}$$

Die Saisonfaktoren hinken also immer einen vollen Saisonzyklus hinter dem Prognosezeitraum her. Warum dies so sein muß, wird nach Ableitung der Parametergleichungen deutlich werden.

Der Grundwert \hat{a}_t wird von Winters durch exponentielle Glättung nach folgender Formel fortgeschrieben: ($0 < \alpha < 1$)

$$(35) \quad \hat{a}_t = (\hat{a}_{t-1} + \hat{b}_{t-1}) + \alpha \left[\frac{x_t}{\hat{s}_{t-L, t \bmod L}} - (\hat{a}_{t-1} + \hat{b}_{t-1}) \right]$$

Auch hier wird wieder die alte Schätzung ($\hat{a}_{t-1} + \hat{b}_{t-1}$) um den Bruchteil α des Prognosefehlers korrigiert, wobei der aktuelle Zeitreihenwert x_t *saisonbereinigt* werden muß, um ihn mit Grund- und Trendwert vergleichen zu können. Da der „neue“ Saisonfaktor noch nicht berechnet worden ist, wird der um L Perioden (einen Zyklus) zurückliegende Saisonfaktor $\hat{s}_{t-L, t \bmod L}$ genommen.

Der Trendwert \hat{b}_t wird wie folgt fortgeschrieben: ($0 < \beta < 1$)

$$(36) \quad \hat{b}_t = \hat{b}_{t-1} + \beta(\hat{a}_t - \hat{a}_{t-1} - \hat{b}_{t-1})$$

Der Prognosefehler müßte hier durch die Differenz von „wahrem“ Trend und geschätztem Trend ausgedrückt werden. Da es aber für den wahren Trend keinen Beobachtungswert gibt, nimmt Winters den neusten, durch die Grundwerte \hat{a}_t und \hat{a}_{t-1} geschätzten Trendwert und glättet ihn exponentiell. Gleichung (35) wird nun zur Bestimmung des aktuellen Saisonfaktors herangezogen.

2. Die Berechnung des Saisonfaktors

Winters aktualisiert den Saisonfaktor durch exponentielle Glättung gemäß der Formel

$$(37) \quad \hat{s}_{tj} = \hat{s}_{t-L,j} + \gamma \left(\frac{x_t}{\hat{a}_t} - \hat{s}_{t-L,j} \right) \quad (0 < \gamma < 1)$$

Die Anpassung erfolgt dabei über den Quotienten x_t/\hat{a}_t , wobei man beachten sollte, daß \hat{a}_t gemäß Gleichung (35) mit einem um L Perioden zurückliegenden Saisonfaktor aktualisiert wurde. Hierin liegt das Dilemma des Winters-Verfahrens: Man müßte eigentlich \hat{a}_t mit dem Quotienten x_t/\hat{s}_{tj} berechnen, hat aber \hat{s}_{tj} nach Gleichung (37) erst zur Verfügung, wenn \hat{a}_t berechnet worden ist.

Als Ausweg nimmt Winters in Gleichung (35)¹ statt \hat{s}_{tj} den letzten verfügbaren Saisonfaktor $\hat{s}_{t-L,j}$, was einen ungünstigen Einfluß auf die Prognosegüte des Verfahrens ha-

¹ Für $m = 0$ gilt $j = t \bmod L$ (vgl. S. 47).

ben kann (vgl. Harrison 1965, S. 102 ff.), da sich die Zeitreihenstruktur während dieser L Perioden verändert haben könnte.

Bevor wir diese Kritik vertiefen, soll jedoch zunächst ein Anwendungsbeispiel mit dem Verfahren von Winters durchgerechnet werden.

3. Die Prognose der monatlichen Übernachtungen im Reiseverkehr

Zur Erläuterung des Verfahrens von Winters soll eine Zeitreihe mit ausgeprägter Saisonstruktur gewählt werden. Es handelt sich um die Anzahl der monatlichen Übernachtungen im Reiseverkehr in der BRD für den Zeitraum Januar 1977 bis Dezember 1980² (n = 48 Zeitreihenwerte). Tabelle 7 enthält die Daten, deren ausgeprägte Saisonstruktur aus Abbildung 3 deutlich hervorgeht.

Der Saisonzyklus beträgt L = 12 Monate. Zur Initialisierung des Verfahrens benötigen wir Startwerte für die Parameter a und b und für 12 Saisonfaktoren. Man erhält die 12 Start-Saisonfaktoren am einfachsten durch Division der monatlichen Zeitreihenwerte durch den Jahresdurchschnitt des ersten Jahres (1977), kann aber auch andere Werte zugrunde legen.

Numerische Erfahrungen am Lehrstuhl des Verfassers haben ergeben, daß folgende Werte für die Glättungsparameter bei dieser Zeitreihe am besten geeignet sind:

$$\alpha = 0,01 \quad \beta = 0,1 \quad \gamma = 0,8$$

Wir zeigen nun beispielhaft die Berechnung der Parameterwerte für den letzten Monat der Zeitreihe (Dezember 1980; t = 48). Mit Hilfe des Computerprogramms haben wir zur Verfügung:

$$\hat{a}_{47} = 9,9942 \quad \hat{b}_{47} = 0,0054 \quad \text{und} \quad \hat{s}_{36,0} = 1,0848 .$$

Nach Gleichung (35) ergibt sich

$$\begin{aligned} \hat{a}_{48} &= (9,9942 + 0,0054) + 0,01 \left[\frac{11,41}{1,0848} - (9,9942 + 0,0054) \right] \\ &= 10,0048 \end{aligned}$$

Der Trendwert folgt aus Gleichung (36) zu

$$\hat{b}_{48} = 0,0054 + 0,1 (10,0048 - 9,9942 - 0,0054) = 0,0059$$

² Quelle: Wirtschaft und Statistik, Hefte 12/1977, 12/1978, 12/1979, 12/1980.

Schließlich liefert Gleichung (37) den neuen Saisonfaktor

$$\hat{s}_{48,0} = 1,0848 + 0,8 \left(\frac{11,41}{10,0048} - 1,0848 \right) = 1,1293$$

In gleicher Weise lassen sich die entsprechenden Werte für die anderen Monate bestimmen.

Die Prognose für den Januar 1981 läßt sich mit Gleichung (34) durchführen, wobei folgendes gilt:

$$m = 1 \quad l = 48 - 12 + 1 \bmod 12 = 37 \quad j = (48 + 1) \bmod 12 = 1$$

$$\hat{s}_{lj} = \hat{s}_{37,1} = 1,1265$$

Tabelle 7

Monatliche Übernachtungen im Reiseverkehr (in Millionen)

Monat	t	x_t	Monat	t	x_t
Jan. 1977	1	9,61	Jan. 1979	25	10,85
Feb. 1977	2	10,00	Feb. 1979	26	10,56
März 1977	3	12,56	März 1979	27	12,97
Apr. 1977	4	15,11	Apr. 1979	28	16,87
Mai 1977	5	19,86	Mai 1979	29	20,90
Juni 1977	6	27,38	Juni 1979	30	27,73
Juli 1977	7	38,66	Juli 1979	31	38,41
Aug. 1977	8	36,80	Aug. 1979	32	37,35
Sep. 1977	9	25,75	Sep. 1979	33	27,00
Okt. 1977	10	17,26	Okt. 1979	34	19,60
Nov. 1977	11	9,23	Nov. 1979	35	10,64
Dez. 1977	12	9,50	Dez. 1979	36	10,87
Jan. 1978	13	10,58	Jan. 1980	37	11,27
Feb. 1978	14	10,79	Feb. 1980	38	11,82
März 1978	15	15,49	März 1980	39	14,35
Apr. 1978	16	13,61	Apr. 1980	40	16,69
Mai 1978	17	21,25	Mai 1980	41	22,64
Juni 1978	18	25,77	Juni 1980	42	28,27
Juli 1978	19	38,04	Juli 1980	43	37,55
Aug. 1978	20	38,48	Aug. 1980	44	36,70
Sep. 1978	21	26,45	Sep. 1980	45	28,43
Okt. 1978	22	17,84	Okt. 1980	46	20,06
Nov. 1978	23	9,92	Nov. 1980	47	10,75
Dez. 1978	24	10,45	Dez. 1980	48	11,41

Als Prognosewert ergibt sich

$$\hat{x}_{\text{Januar 1981}} = \hat{x}_{49} = (10,0048 + 0,0059 \cdot 1) \cdot 1,1265 = 11,2771$$

In Abbildung 3 sind die Ex-ante-Prognosewerte bis einschließlich Juni 1981 ($m = 6$) eingezeichnet. Tabelle 8 faßt die Ex-post-Prognoseergebnisse für das Jahr 1980 (als Ausschnitt) zusammen.

Bezogen auf den Gesamtzeitraum von 1977 bis 1980 weisen die Gütemaße der Ex-post-Prognose folgende Werte auf:

$$\text{MAA} = 0,8 \text{ [Mio.]} \quad \text{WMQA} = 1,08 \text{ [Mio.]} \quad U = 0,17$$

Insbesondere der Theil'sche Ungleichheitskoeffizient U zeigt, daß die Prognose dieser Zeitreihe mit dem Verfahren von Winters sehr gut ist. Ein ähnlicher Eindruck in bezug auf die Übereinstimmung von effektiven und prognostizierten Werten läßt sich auch aus Abbildung 3 gewinnen.

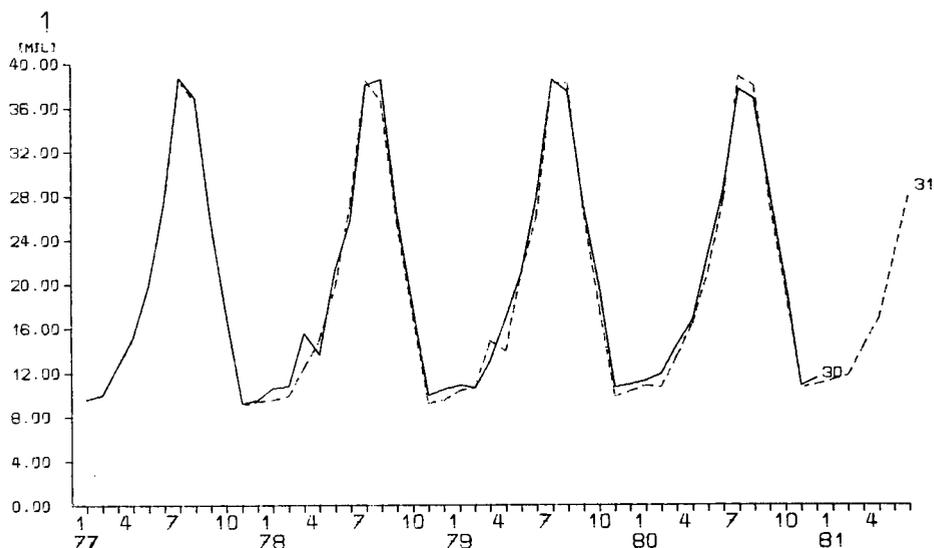


Abbildung 3: Monatliche Übernachtungen im Reiseverkehr [Mio]
– Zeitreihenwerte und Prognosewerte (gestrichelt) –

Tabelle 8

Ex-post-Prognose der Anzahl Übernachtungen im Reiseverkehr für das Jahr 1980
(in Mio)

Monat	t	x_t	\hat{x}_t	e_t	\hat{s}_{tj}
Jan. 1980	37	11,27	10,79	0,48	1,1265
Feb. 1980	38	11,82	10,62	1,20	1,1660
März 1980	39	14,35	13,44	0,91	1,4251
Apr. 1980	40	16,69	16,38	0,31	1,6714
Mai 1980	41	22,64	21,07	1,57	2,2417
Juni 1980	42	28,27	27,62	0,65	2,8234
Juli 1980	43	37,55	38,68	-1,13	3,7894
Aug. 1980	44	36,70	37,84	-1,14	3,7037
Sep. 1980	45	28,43	27,12	1,31	2,8226
Okt. 1980	46	20,06	19,41	0,65	1,9954
Nov. 1980	47	10,75	10,57	0,18	1,0719
Dez. 1980	48	11,41	10,86	0,55	1,1293

4. Kritische Zusammenfassung

Das Verfahren von Winters zeichnet sich durch folgende *Vorzüge* aus:

- Durch die enge Verwandtschaft mit der exponentiellen Glättung sind nur *elementare Rechenoperationen* auszuführen, die leicht programmierbar sind.
- Das Verfahren benötigt wenig Rechenzeit und hat nur einen geringen Speicherbedarf, da nur der letzte Beobachtungswert x_t , die L Saisonfaktoren, Trend- und Grundwert sowie die Glättungsparameter α, β, γ zu speichern sind.
- Drei verschiedene Glättungsparameter verleihen dem Verfahren eine erhebliche Flexibilität und die Möglichkeit, sich an Änderungen von Trend und Saison differenziert anzupassen.

Folgende Gesichtspunkte können sich demgegenüber *nachteilig* bemerkbar machen:

- Bei der Bestimmung der Glättungsparameter ist man häufig auf „trial und error“ angewiesen. Der Einfluß dieser Werte (in ihrer Gesamtheit) auf den Prognosefehler kann kaum simultan abgeschätzt werden. Erfahrungen über dynamische Anpassungen der Parameter sind dem Verfasser nicht bekannt geworden.
- Die Prognose wird mit einem Saisonfaktor erstellt, der L Perioden, d. h. einen ganzen Zyklus, vorher errechnet worden ist. Dieser Mangel an Aktualität kann aufgrund des Prognosesystems nicht behoben werden. Er beeinflußt sicherlich die Prognosegüte, daß er aber „zum Zusammenbruch des ganzen Systems“ (Schläger in Mertens 1981, S. 62) führen kann, konnte der Verfasser bei seinen Anwendungsbeispielen nicht beobachten.

- Die multiplikative Verknüpfung des Saisonfaktors ist für ökonomische Zeitreihen, insbesondere mit Trend, sicherlich am günstigsten. Es mag aber Zeitreihen geben, die mit additiver Verknüpfung besser zu behandeln sind. Hierzu müßte das Winters-Modell umgearbeitet werden.

Trotz dieser Kritikpunkte erfreut sich das Winters-Verfahren in der Praxis und bei Anbietern von EDV-Programmen großer Beliebtheit.

Wir wollen nun einen anderen Lösungsansatz des Saisonproblems kennenlernen, der eine etwas veränderte Blickrichtung in bezug auf Zeitreihenzyklen mit sich bringen wird, nämlich die Spektralanalyse.

III. Anwendung der Spektralanalyse

Es kann nicht Aufgabe dieses Lehrbuchs sein, eine detaillierte Einführung in die statistischen Grundlagen der Spektralanalyse zu geben. Andererseits ist diese Theorie ein leistungsfähiges Instrument zur Analyse von Zeitreihen, insbesondere zur Aufdeckung mehrerer sich gegenseitig überlagernder Saisonzyklen (Hansmann 1980 c, S. 236 ff.).

Mit folgendem Kompromiß soll dieses Problem angegangen werden: Nur die für das Verständnis unabdingbar notwendigen Begriffe werden – möglichst elementar – eingeführt und der Schwerpunkt der Erörterungen auf die Anwendungen gelegt. Vertiefte Kenntnisse lassen sich durch die Lektüre der Monographien (Leiner 1978, König/Wolters 1972) gewinnen.

1. Stationäre stochastische Prozesse

Ein stochastischer Prozeß kann grob als ein statistischer Vorgang angesehen werden, der sich nach den Gesetzen der Wahrscheinlichkeit zeitlich vollzieht (Box/Jenkins 1976, S. 24).

Nach einer etwas formaleren – jedoch gehaltvolleren – Definition ist ein stochastischer Prozeß eine *Familie von Zufallsvariablen*

$$\{X(t, \omega); t \in T, \omega \in \Omega\}$$

die auf der Menge Ω der Elementarereignisse und auf der Indexmenge T , die z. B. als Menge von Zeitpunkten interpretiert werden kann, definiert ist.

Für einen *festen* Zeitpunkt t ergibt sich die einfache Zufallsvariable $X(\omega | t)$, für *festes* ω bekommen wir eine Funktion der Zeit $X(t | \omega)$, vereinfacht geschrieben $X(t)$, die man als Realisation $\{x_t\}$ des stochastischen Prozesses auffassen kann.

Wir können nunmehr unseren schon bekannten Begriff der Zeitreihe mit diesen neuen Begriffen in Verbindung setzen:

Eine *Zeitreihe* ist eine *Realisation* eines *stochastischen Prozesses*.

Die Anwendung der Spektralanalyse ist nur bei solchen stochastischen Prozessen zu empfehlen, deren gemeinsame *Verteilungsfunktion* im Zeitablauf *unverändert* bleibt. Man nennt solche Prozesse *streng stationär*.

Verteilungsfunktionen sind jedoch oft schwierig zu bestimmen, so daß zur Charakterisierung der Stationarität gewöhnlich die beiden ersten Momente der Verteilungsfunktion herangezogen werden.

Wir bezeichnen einen stochastischen Prozeß als *schwach stationär*, wenn folgende Bedingungen erfüllt sind:

a) Der Erwartungswert des Prozesses ändert sich nicht im Zeitablauf und ist endlich

$$E[X(t)] = \mu < \infty$$

b) Die Varianz des Prozesses ist zeitunabhängig und endlich

$$V[X(t)] = R(0) < \infty$$

c) Die Autokovarianzfunktion des Prozesses ist endlich und nur von der zeitlichen Differenz τ abhängig

$$\text{Cov}[X(t), X(t + \tau)] = R(\tau) < \infty$$

Im folgenden werden wir uns bei der Anwendung der Spektralanalyse mit der schwachen Stationarität des zugrunde liegenden stochastischen Prozesses begnügen.

2. Ermittlung von Korrelogramm und Spektrum

Einen stationären stochastischen Prozeß kann man mit Hilfe seiner *Autokovarianzfunktion*

$$E[(X(t) - \mu)(X(t + \tau) - \mu)] = \text{Cov}[X(t), X(t + \tau)] = R(\tau)$$

oder nach Berechnung der Autokorrelationskoeffizienten durch sein *Korrelogramm*

$$\frac{\text{Cov}[X(t), X(t + \tau)]}{V[X(t)]} = \frac{R(\tau)}{R(0)} = r(\tau)$$

beschreiben. Abbildung 4 zeigt die Autokovarianzfunktion der monatlichen Veränderungen des Aktienkursindex vom Stat. Bundesamt (Reihe 2, Nr. 17) von 1974 bis 1981, die näherungsweise während des betrachteten Zeitraums als Realisation eines stationären stochastischen Prozesses aufgefaßt werden können.

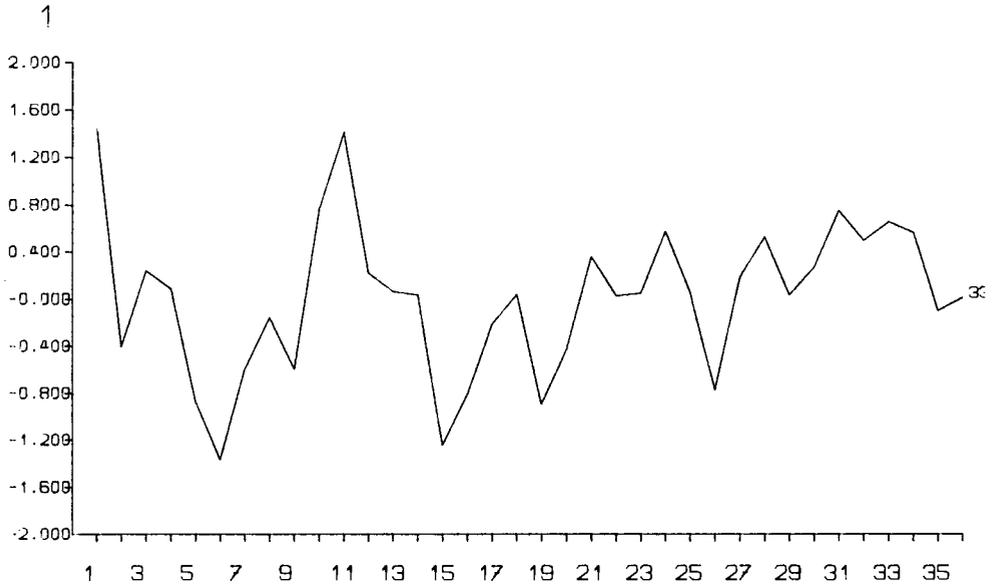


Abbildung 4: Autokovarianzfunktion der monatlichen Veränderungen des Aktienkursindex

Hohe (absolute) Autokorrelationskoeffizienten bei einem bestimmten Wert von τ zeigen an, daß der Prozeß an Zeitpunkten, die τ Zeiteinheiten voneinander entfernt liegen, ähnliches Verhalten aufweist. τ kann daher u. U. die Länge eines Saisonzyklus sein. In Abbildung 4 identifizieren wir z. B. einen 11-Monats-Zyklus, falls die Schätzung R_{11} signifikant ist. Dieser Zyklus wird uns im nächsten Kapitel noch beschäftigen.

Im allgemeinen läßt sich *ein* ausgeprägter Saisonzyklus mit Hilfe der Autokovarianzfunktion identifizieren. Treten jedoch bei einem Prozeß mehrere sich gegenseitig überlappende Saisonzyklen auf, so ist eine Identifizierung der einzelnen Zyklen meistens nicht mehr möglich, da die Autokovarianzen Beiträge von allen Zyklen enthalten.

Hier setzt nun die Spektralanalyse ein. Durch Transformation des Korrelogramms vom *Zeitbereich* in den *Frequenzbereich* ist es möglich, mehrere Zyklen zu isolieren und ihre Periodenlängen zu bestimmen. Diese sog. *Fourier-Transformation* geht folgendermaßen vonstatten (vgl. König/Wolters 1972, S. 38 ff.):

Die Autokovarianzfunktion $R(\tau)$ kann als Stieltjes-Integral geschrieben werden (Wold 1954)

$$(38) \quad R(\tau) = \int_{-\pi}^{\pi} e^{i\lambda\tau} \cdot f(\lambda) d\lambda$$

Hier bedeuten λ die *Kreisfrequenz*, d. h. die auf dem Einheitskreis während einer Periode zurückgelegte Strecke, $f(\lambda)$ das *Spektrum* und i die imaginäre Einheit.

Gleichung (38) läßt sich mit Hilfe der Eulerschen Gleichung umformen in

$$(39) \quad R(\tau) = \int_{-\pi}^{\pi} \cos\lambda\tau \cdot f(\lambda) d\lambda + i \int_{-\pi}^{\pi} \sin\lambda\tau \cdot f(\lambda) d\lambda$$

und weiter, da R eine reellwertige und gerade Funktion ist, in

$$(40) \quad R(\tau) = 2 \int_0^{\pi} \cos\lambda\tau \cdot f(\lambda) d\lambda$$

Dies ist die *erste Grundgleichung* der *Spektralanalyse* für die Berechnung der Autokovarianzen aus dem gegebenen Spektrum. Umgekehrt ermittelt man aus $R(\tau)$ das Spektrum, indem man Gleichung (38) über eine Fourierreihe in

$$(41) \quad f(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} R(\tau) e^{-i\lambda\tau} \quad (-\pi \leq \lambda \leq \pi)$$

umformt. Dafür kann geschrieben werden

$$(42) \quad f(\lambda) = \frac{1}{2\pi} [R(0) + \sum_{\tau=1}^{\infty} R(\tau) (e^{i\lambda\tau} + e^{-i\lambda\tau})]$$

Hieraus ergibt sich nach Anwendung der Eulerschen Gleichung die *zweite Grundgleichung* der *Spektralanalyse*

$$(43) \quad f(\lambda) = \frac{1}{2\pi} [R(0) + 2 \sum_{\tau=1}^{\infty} R(\tau) \cdot \cos\lambda\tau] \quad (0 \leq \lambda \leq \pi)$$

Damit ist der Übergang vom Zeitbereich in den Frequenzbereich vollzogen. Obwohl hierdurch keine zusätzlichen Informationen gewonnen werden, macht Gleichung (40) den Vorteil der Spektraldarstellung deutlich. Setzt man dort $\tau = 0$, so ergibt sich die Varianz des stochastischen Prozesses $V[X(t)] = R(0)$ zu

$$(44) \quad R(0) = 2 \int_0^{\pi} f(\lambda) d\lambda$$

$2f(\lambda)d\lambda$ gibt also den Anteil der Varianz wieder, der im Frequenzband zwischen λ und $\lambda + d\lambda$ liegt. Die Teilvarianzen addieren sich nach (44) zur Gesamtvarianz, sind also unabhängig voneinander darstellbar. Dadurch lassen sich Frequenzen mit hohem Varianz-

anteil identifizieren, die Saisonzyklen zugeordnet werden können. Die Länge eines Saisonzyklaus L ergibt sich aus der Beziehung

$$(45) \quad L = \frac{2\pi}{\lambda}$$

Auch mehrere sich überlappende Saisonzyklen können so durch die Unabhängigkeit ihrer Varianzanteile im Frequenzbereich isoliert werden.

Die Abbildung 5 zeigt das geschätzte Spektrum der monatlichen Veränderungen des Aktienkursindex und ist somit das Pendant zur Abbildung 4.

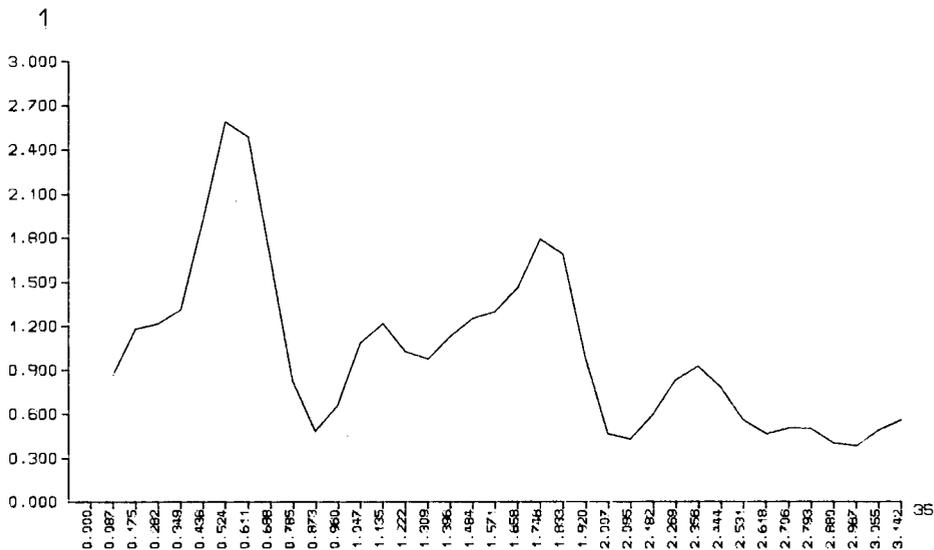


Abbildung 5: Spektrum der monatlichen Veränderungen des Aktienkursindex

Die Fläche unter der Kurve des Spektrums stellt nach Gleichung (44) die Gesamtvarianz dar. Ein großer Teil dieser Varianz befindet sich im Frequenzband $0,4 < \lambda < 0,7$. Setzt man als Mittelwert $\lambda = 0,55$ in Gleichung (45) ein, so ergibt sich eine Zykluslänge L von 11,4 Monaten. Dieser Zyklus wurde auch im Korrelogramm gefunden.

Ein zweiter Zyklus kann bei $\lambda = 1,75$ vermutet werden. Seine Länge beträgt 3,6 Monate. In der Autokovarianzfunktion hebt sich dieser Zyklus jedoch nicht aus dem Gesamtbild heraus.

3. Die Lag-Fenster

Bisher wurde nicht der Frage nachgegangen, wie das Spektrum einer konkret vorliegenden Zeitreihe eigentlich geschätzt wird. Wir kennen ja in der Wirklichkeit *nicht* den stochastischen Prozeß, sondern haben nur eine Realisation – in Form der Zeitreihe $\{x_t\}$ – vorliegen.

Zunächst schätzt man die Autovarianzen des stochastischen Prozesses aus der Zeitreihe mit Hilfe folgender konsistenter Schätzfunktion, wobei \bar{x} das arithmetische Mittel der Zeitreihe ist

$$(46) \quad \hat{R}(\tau) = \frac{1}{T} \sum_{t=1}^{T-\tau} (x_t - \bar{x})(x_{t+\tau} - \bar{x}) \quad \tau = 0, 1, \dots, T-1$$

Werden die $\hat{R}(\tau)$ in Gleichung (43) eingesetzt, wobei man beachten muß, daß die Summation infolge der endlichen Zeitreihe nur bis $T-1$ durchgeführt werden kann, so läßt sich zeigen, daß die Schätzfunktion $\hat{f}(\lambda)$ nicht konsistent ist (König/Wolters 1972, S. 59/60).

Wir erhalten eine konsistente Schätzfunktion, indem wir nur $m < T-1$ Autokovarianzen $\hat{R}(\tau)$ schätzen und sie anschließend mit bestimmten Gewichten $w_m(\tau)$ versehen:

$$(47) \quad \hat{f}(\lambda) = \frac{1}{2\pi} [\hat{R}(0) + 2 \sum_{\tau=1}^m w_m(\tau) \cdot \hat{R}(\tau) \cdot \cos\lambda\tau] \quad (0 \leq \lambda \leq \pi)$$

Es gibt in der Literatur verschiedene Ansätze zur Bestimmung der Gewichte $w_m(\tau)$ – auch als „Lag-Fenster“ bezeichnet –, die hier nicht wiedergegeben werden können (vgl. Leiner 1978, S. 38 ff.; König/Wolters 1972, S. 60 ff.).

Für alle in diesem Buch geschätzten Spektren wird die folgende Gewichtungsfunktion von (Parzen 1961, S. 167 ff.) verwendet, mit der der Verfasser die besten numerischen Erfahrungen gemacht hat:

$$(48) \quad w_m(\tau) = \begin{cases} 1 - 6 \frac{\tau^2}{m^2} \left(1 - \frac{\tau}{m}\right) & \text{für } \tau < \frac{m}{2} \\ 2 \left(1 - \frac{\tau}{m}\right)^3 & \text{für } \frac{m}{2} \leq \tau \leq m \end{cases}$$

Es bleibt noch die Festlegung eines geeigneten Wertes für m . Eine Faustformel besagt, daß m nicht größer als $T/3$ sein sollte. Da dies nur ein sehr grober Anhaltspunkt ist, sollte man mit verschiedenen Werten von m das Spektrum schätzen und die Unterschiede im Verlauf der Spektren untersuchen. Auf diese Weise kann man den Wert von m finden, der die wesentlichen Einzelheiten des Spektrum klar hervortreten läßt.

4. Spektralanalyse der Auflage einer Publikumszeitschrift

Als Anwendungsbeispiel der Spektralanalyse soll die Entwicklung der verkauften Auflage einer deutschen Publikumszeitschrift im Zeitraum von 1978 bis 1981 betrachtet werden. Um eine stationäre Zeitreihe zu erhalten, eliminieren wir den Trend der Zeitreihe durch Bildung der ersten Differenzen ($x_t - x_{t-1}$) für $t = 2 \dots T$ und erhalten folgendes Bild:

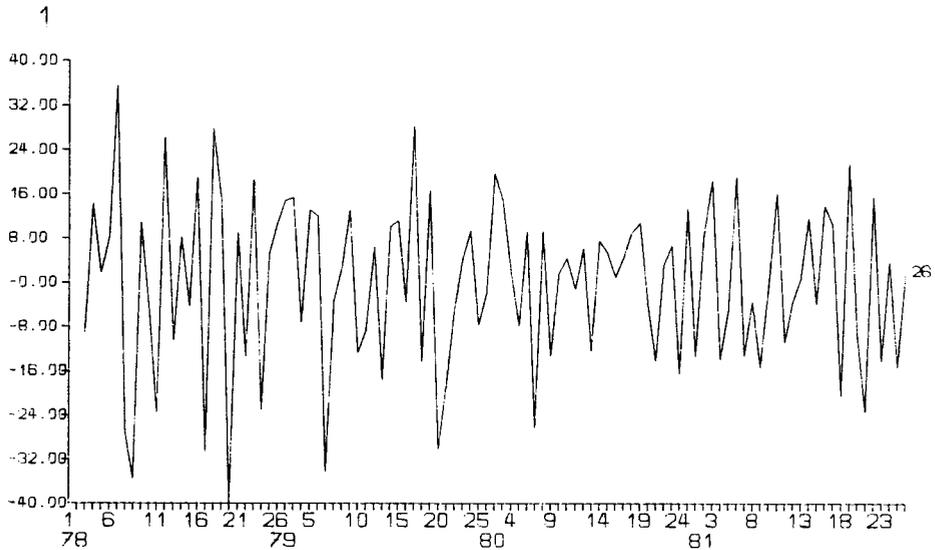


Abbildung 6: Erste Differenzen der Auflage einer Publikumszeitschrift

Die Einzelheiten des Spektrums treten am deutlichsten hervor, wenn man $m = 36$ Autokovarianzen schätzt, die in Abbildung 7 wiedergegeben sind.

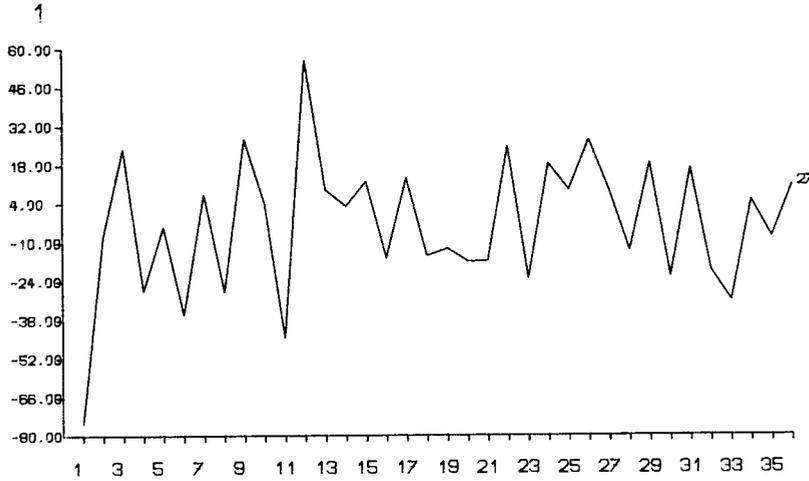


Abbildung 7: Autokovarianzfunktion einer Publikumszeitschrift

Mit der Gewichtsfunktion von Parzen erhält man aus Gleichung (47) das geschätzte Spektrum $\hat{f}(\lambda)$ dieser Zeitreihe, wie Abbildung 8 zeigt.

$$\hat{f}(\lambda)$$

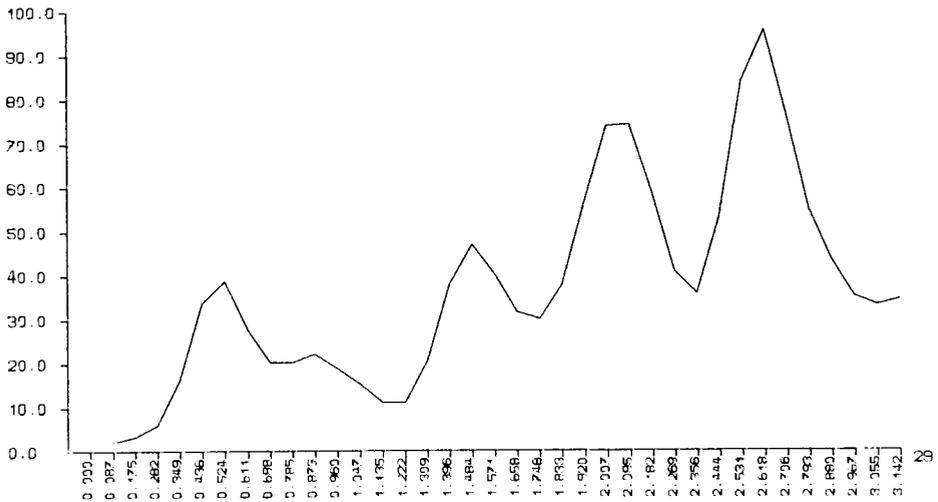


Abbildung 8: Spektrum der Auflage einer Publikumszeitschrift

Abbildung 8 läßt erkennen, daß an vier Stellen große Varianzanteile konzentriert sind. Infolge der Unsicherheit der Schätzung können die zugehörigen Frequenzen natürlich nur angenähert angegeben werden. Sie liegen etwa bei $\lambda_1 = 0,5$, $\lambda_2 = 1,5$, $\lambda_3 = 2$ und $\lambda_4 = 2,6$. Mit Gleichung (45) errechnet man die Periodenlänge der zugehörigen Zyklen zu

$P_1 \approx 12$ bis 13	Perioden
$P_2 \approx 4$	Perioden
$P_3 \approx 3$	Perioden
$P_4 \approx 2$ bis 3	Perioden

Der erste Zyklus stellt eindeutig den Saisonzklus dieser Zeitschrift dar, der auch in der Autokovarianzfunktion gefunden wurde. Die kurzen Zyklen zwischen 2 und 4 Perioden werden von der unterschiedlichen Attraktion der Titelseite (Text, Bild und Aufmachung) dieser Zeitschrift verursacht, die ja ein grobes Spiegelbild des Inhalts der Zeitschrift ist, und das Kaufverhalten der Konsumenten erheblich beeinflußt. Offensichtlich gelingt es der Redaktion nur in Abständen von 3 bis 4 Erscheinungszeitpunkten, überdurchschnittliches Interesse an der Zeitschrift zu wecken.

Die Spektralanalyse vermag zwar nur Zyklen „aufzudecken“ und keine Einflußfaktoren, doch führt häufig die inhaltliche Analyse eines gefundenen Zyklus auf die Spur eines periodisch wirkenden Einflußfaktors, den man anschließend in ein multiples Regressionsmodell aufnehmen kann. Ein Beispiel für eine solche fruchtbare Kombination von Spektral- und Regressionsanalyse ist in (Hansmann 1980c, S. 231 ff.) beschrieben.

5. Kritische Zusammenfassung

Die Spektralanalyse ermöglicht die Identifizierung von verschiedenen, sich gegenseitig überlappenden Saisonzyklen, sowie die Bestimmung der Zykluslängen. Mit diesen Informationen kann man die Saisonverläufe in der Form von geeigneten Sinus- bzw. Kosinusfunktionen darstellen und bei Geltung der Zeitstabilitätshypothese zur Prognose künftiger Zeitreihenwerte verwenden.

Die Anwendung der Spektralanalyse ist jedoch *nicht unproblematisch*. Folgende Gesichtspunkte sind dabei besonders hervorzuheben:

- Die konkret vorliegende Zeitreihe muß als Realisation eines *stationären* stochastischen Prozesses aufgefaßt werden können.
- Der Zeitreihenumfang muß genügend groß sein (möglichst mehr als 100 Zeitreihenwerte).
- Die Struktur der Zeitreihe, insbesondere die Zykluslänge, darf sich im Zeitablauf nicht verändern.

- Die Schätzung des Spektrums ist nicht einfach (hoher mathematischer Aufwand, Wahl der geeigneten Gewichtungsfunktion, Bestimmung des maximalen Lags m , Signifikanzprüfung).

Ist eine der ersten drei Bedingungen nicht erfüllt, sollte die Spektralanalyse möglichst nicht angewendet werden. Aufgrund dieser eingeschränkten Verwendungsmöglichkeit ist die Spektralanalyse kein „eigentliches“ Prognoseverfahren, sondern ein Hilfsinstrument zur Aufdeckung von Saisonzyklen, deren Kenntnis bei einem darauf aufbauenden Prognoseverfahren, z. B. der Regressionsanalyse (Kapitel G.), ausgenutzt werden kann (Hansmann 1980c).

E. Prognosen mit autoregressiven Methoden

I. Grundgedanken der Zeitreihenanalyse

Es war das Ziel der exponentiellen Glättung, den zukünftigen Verlauf einer Zeitreihe aus ihren Vergangenheitswerten zu prognostizieren, wie sich aus Gleichung (7) erkennen läßt.

$$(7) \quad \hat{x}_{t+1} = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i x_{t-i} + (1-\alpha)^t \hat{x}_1$$

Die Gewichtung der Zeitreihenwerte x_{t-i} erfolgte jedoch nach einem *starren* Schema: $\alpha(1-\alpha)^i$, das auf die Eigenart einer konkreten Zeitreihe nicht abgestimmt werden konnte.

In diesem Kapitel werden wir Verfahren kennenlernen, die die Gewichte der Vergangenheitswerte individuell für jede Zeitreihe zu optimieren versuchen. Sie werden üblicherweise unter dem Begriff „Moderne Zeitreihenanalyseverfahren“ (time series analysis methods) zusammengefaßt und sind – nach Vorarbeiten von Yule (1926), Walker (1931), Slutsky (1937), Wold (1954), Quenouille (1957) und Hannan (1960) – wesentlich von Box/Jenkins (1976) und Makridakis/Wheelwright (1978a) zur Konstruktion von Prognosesystemen weiterentwickelt worden.

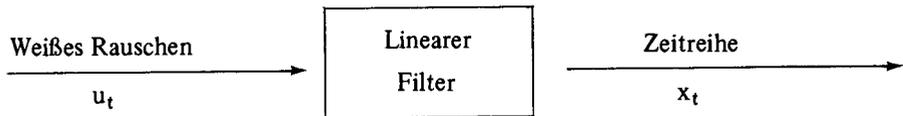
Diese Verfahren stehen auf wesentlich höherem mathematischen Niveau und sind so aufgebaut, daß sie z. B. die exponentielle Glättung als Spezialfall enthalten (vgl. S. 70). Zur Zeit dürften sie die flexibelsten und numerisch anspruchvollsten Verfahren der univariaten Prognosemodelle sein.

Bevor das Box-Jenkins-Verfahren und das „adaptive Filtern“ von Makridakis/Wheelwright näher erläutert werden, ist zunächst eine kurze Einführung in die gemeinsamen Grundgedanken angebracht.

Ausgangspunkt der Überlegungen sind wie im vorigen Kapitel die *stationären stochastischen Prozesse*, wobei schon hier darauf hingewiesen sei, daß später auch nicht-stationäre Prozesse behandelt werden. Nach einer Idee von Yule (1927) wird unterstellt, daß ein stationärer stochastischer Prozeß von einer Störvariablen erzeugt wird, deren Realisationen $u_t, u_{t-1}, u_{t-2} \dots$ in bestimmter Weise gefiltert werden.

Die Störvariable u besitze – wie auf S. 30 dargelegt – den Erwartungswert Null, die Varianz σ_u^2 und sei über die Zeit unkorreliert. Eine Störvariable mit diesen Eigenschaften wird im technischen Bereich als „weißes Rauschen“ bezeichnet.

Aus dem weißen Rauschen wird nun durch Anwendung eines linearen Filters die konkrete Zeitreihe als Realisation des zugrunde liegenden stochastischen Prozesses herausgefiltert.



Als Filter verwenden Yule (1927) und Box-Jenkins (1976) die gewichtete Summe der Störvariablen-Realisationen $u_t, u_{t-1} \dots$, so daß sich folgende Darstellung ergibt:

$$(49) \quad x_t = \mu + u_t + a_1 u_{t-1} + a_2 u_{t-2} + a_3 u_{t-3} + \dots$$

μ ist hier der Mittelwert des stationären Prozesses und a_1, a_2, \dots sind die (zu schätzenden) Gewichte der Störvariablen u . Damit der Prozeß stationär ist, muß die Anzahl der Glieder endlich sein oder die Reihe bei unendlicher Anzahl konvergieren.

Den weiten Anwendungsbereich des Prozesses (49) erkennt man leicht, wenn man die Reihe für $t-1, t-2$ aufschreibt

$$(49a) \quad x_{t-1} = \mu + u_{t-1} + a_1 u_{t-2} + a_2 u_{t-3} + a_3 u_{t-4} + \dots$$

$$(49b) \quad x_{t-2} = \mu + u_{t-2} + a_1 u_{t-3} + a_2 u_{t-4} + a_3 u_{t-5} + \dots$$

sukzessive nach u_{t-1}, u_{t-2} auflöst und in (49) einsetzt

$$(50) \quad x_t = (1 - a_1 - a_2 + a_1^2) \mu + u_t + a_1 x_{t-1} + (a_2 - a_1^2) x_{t-2} \\ + (-a_1(a_2 - a_1^2) - a_1 a_2 + a_3) u_{t-3} \\ + (-a_2(a_2 - a_1^2) - a_1 a_3 + a_4) u_{t-4} \\ + (-a_3(a_2 - a_1^2) - a_1 a_4 + a_5) u_{t-5}$$

Schon nach diesen beiden Umformungen ist zu erkennen, daß x_t auch durch eine gewichtete Summe von x_{t-1} und x_{t-2} ausgedrückt werden kann. Führt man diesen Prozeß für $t-2, t-3, \dots$ aus, ergibt sich die Gleichung

$$(51) \quad x_t = \mu' + b_1 x_{t-1} + b_2 x_{t-2} + b_3 x_{t-3} + \dots + u_t$$

wobei zur Abkürzung $a_1 = b_1, a_2 - a_1^2 = b_2$ und $(1 - a_1 - a_2 + a_1^2) \mu = \mu'$ gesetzt wurde.

$b_3, b_4 \dots$ sind Abkürzungen für die komplizierteren Ausdrücke, die sich als Koeffizienten von $x_{t-3}, x_{t-4} \dots$ nach den Umformungen ergeben.

Die Gleichungen (49) und (51) sind die Grundlagen des Box-Jenkins-Verfahrens und aller ähnlichen Methoden der modernen Zeitreihenanalyse. Sie besagen, daß eine konkrete Zeitreihe als Realisation eines stationären stochastischen Prozesses ausgedrückt werden kann durch die gewogene Summe der

1. Realisation einer zufallsbestimmten Störvariablen oder der
2. Vergangenheitswerte der Zeitreihe.

Um die Konstanten μ bzw. μ' zu sparen, werden häufig die Zeitreihenwerte als Abweichungen von ihrem Mittelwert ausgedrückt, so daß man in (49) und (51) μ und μ' gleich Null setzen kann.

Im folgenden sind zwei grundlegende Probleme zu lösen:

1. Wieviele Glieder der Gleichungen (49) bzw. (51) sollen zur Anpassung des Modells an eine gegebene Zeitreihe berücksichtigt werden?
2. Wie können geeignete Gewichte $a_1, a_2 \dots$ und $b_1, b_2 \dots$ gefunden werden?

Diese Fragen werden in den nächsten beiden Abschnitten beantwortet.

II. Das Box-Jenkins-Verfahren (ARIMA-Modelle)

1. Autoregressive Prozesse (AR-Modelle)

Zur Definition der AR-Modelle ist es zweckmäßig, von Gleichung (51) auszugehen, wobei im folgenden stets vorausgesetzt sei, daß $\mu = \mu' = 0$ gilt, d. h. die Zeitreihenwerte als Abweichungen vom Mittelwert ausgedrückt werden.

Falls die ersten p Gewichte in (51) ungleich Null und alle übrigen gleich Null sind, sprechen Box/Jenkins (1976, S. 51) von einem *autoregressiven Prozeß p -ter Ordnung* oder kurz von einem *AR(p)-Modell* zum Gegenwartszeitpunkt t

$$(52) \quad x_t = b_1 x_{t-1} + b_2 x_{t-2} + \dots + b_p x_{t-p} + u_t$$

Der Name autoregressiv wurde gewählt, weil (52) einer *Regressionsgleichung* ähnlich ist, mit der Eigenschaft, daß die Variablen der rechten Seite die zeitverschobenen Werte der abhängigen Variablen selbst enthalten.

Zur Beantwortung der Frage, ob eine konkret vorliegende Zeitreihe einen autoregressiven Prozeß bestimmten Grades widerspiegelt, analysieren Box/Jenkins zunächst die *Autokorrelationsfunktion* r_τ ($\tau = 1, 2 \dots$), wobei τ der time-lag ist. Man erhält sie auf folgende Weise: (52) wird mit $x_{t-\tau}$ multipliziert

$$(53) \quad x_{t-\tau} x_t = b_1 x_{t-\tau} x_{t-1} + b_2 x_{t-\tau} x_{t-2} + \dots + b_p x_{t-\tau} x_{t-p} + x_{t-\tau} u_t$$

Durch Bildung des Erwartungswertes $E[x_{t-\tau} x_t]$ ergibt sich die *Autokovarianz* R_τ und nach Division durch die Varianz R_0 der Autokorrelationskoeffizient r_τ

$$(54) \quad r_\tau = b_1 r_{\tau-1} + b_2 r_{\tau-2} + \dots + b_p r_{\tau-p} + E[x_{t-\tau} u_t] / R_0$$

Der letzte Summand verschwindet, weil $x_{t-\tau}$ gemäß (49) nur von den Störvariablen $u_{t-\tau}, u_{t-\tau-1} \dots$ beeinflusst wird, die mit u_t unkorreliert sind.

Es kann nun bewiesen werden (Box/Jenkins 1976, S. 55), daß die Autokorrelationsfunktion bei AR-Modellen mit zunehmendem time-lag τ entweder exponentiell abfällt oder eine gedämpfte Sinusschwingung vollzieht.

Wir wollen diesen Sachverhalt am Beispiel des AR(1) Modells veranschaulichen. Ein autoregressiver Prozeß erster Ordnung läßt sich gemäß (52) darstellen als

$$(55) \quad x_t = b_1 x_{t-1} + u_t \quad (-1 < b_1 < 1)$$

wobei die Beschränkung für b_1 aus der Stationarität folgt. Die Autokorrelationsfunktion ergibt sich aus (54) zu

$$(56) \quad r_\tau = b_1^\tau r_{\tau-1} \quad \text{und da } r_0 = 1$$

$$(57) \quad r_\tau = b_1^\tau \quad (\tau \geq 0; -1 < b_1 < 1)$$

r_τ ist also eine geometrische Folge, die für $b_1 > 0$ exponentiell und monoton gegen Null und für $b_1 < 0$ in der Form einer gedämpften Sinusschwingung gegen Null strebt. In der folgenden Abbildung 9 sind die Verläufe für $b_1 = 0,9$ und $b_1 = -0,9$ dargestellt.

Mit der Autokorrelationsfunktion kann man nun entscheiden, ob eine gegebene Zeitreihe einem AR-Prozeß entspricht. Sie sollte einen ähnlichen Verlauf aufweisen wie Abbildung 9 zeigt.

Hat man in einer Zeitreihe einen AR-Prozeß erkannt, ist als nächstes die Ordnung des Prozesses zu bestimmen. Dazu stellen Box/Jenkins (1976) folgende Überlegung an:

Ist der Prozeß tatsächlich von der Ordnung p , dann muß in Gleichung (52) der letzte Koeffizient b , der ungleich Null ist, b_p sein, da $b_{p+1}, b_{p+2} \dots$ ex definitione gleich Null sind.

Man fängt zweckmäßigerweise bei $p = 1$ an und prüft, ob in der Gleichung

$$(58a) \quad x_t = b_1 x_{t-1} + u_t$$

$b_1 \neq \text{Null}$ ausfällt. Wenn nein, ist der Prozeß null-ter Ordnung, im anderen Fall muß nun geprüft werden, ob b_1 der letzte Koeffizient ungleich Null ist. Das geschieht mit Hilfe der Gleichung

$$(58b) \quad x_t = b_1 x_{t-1} + b_2 x_{t-2} + u_t$$

$b_2 = 0$ würde hier auf einen Prozeß erster Ordnung hinweisen. Falls $b_2 \neq 0$, muß die Prüfung fortgesetzt werden mit

$$(58c) \quad x_t = b_1 x_{t-1} + b_2 x_{t-2} + b_3 x_{t-3} + u_t$$

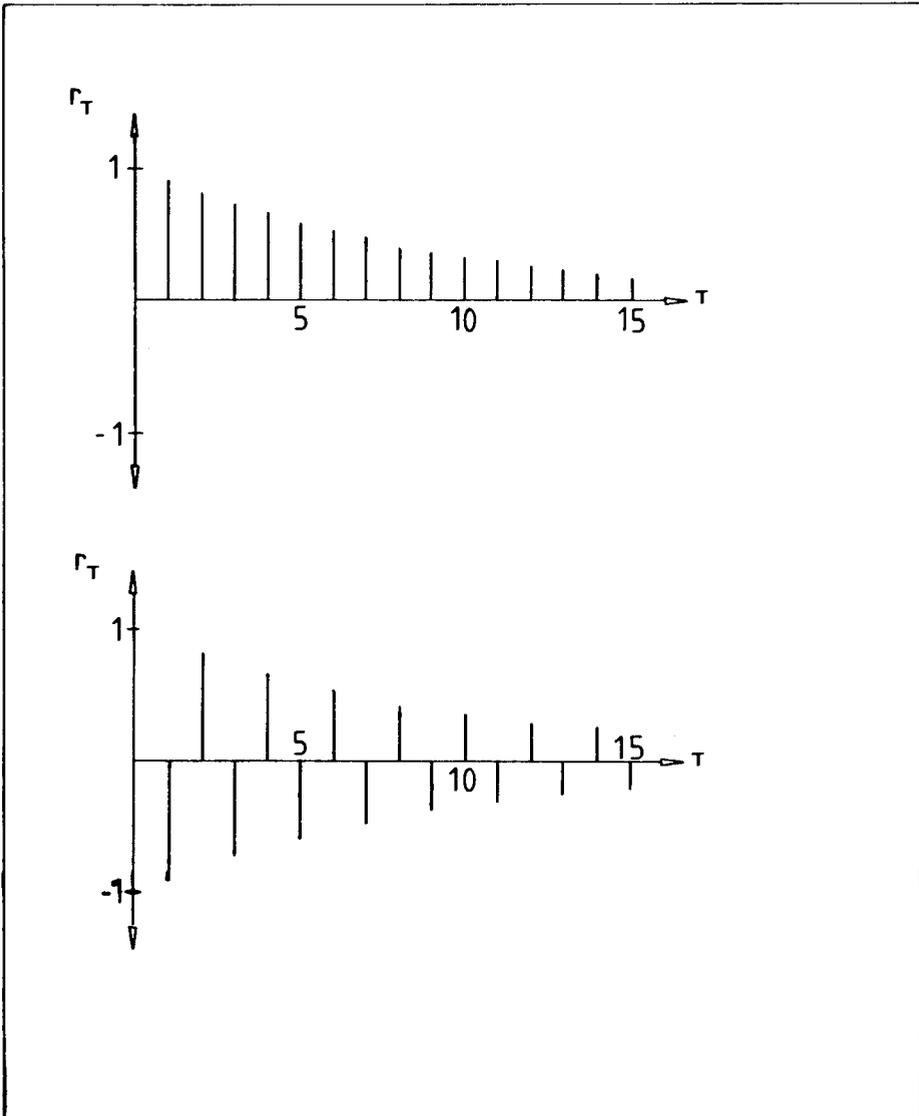


Abbildung 9
 Autokorrelationsfunktion von (57) mit $b_1 = 0,9$ und $b_1 = -0,9$

Dieses Verfahren ist fortzusetzen bis der letzte Koeffizient $\neq 0$ gefunden wurde. Er bestimmt die Ordnung des Prozesses.

Das oben beschriebene Vorgehen ist numerisch aufwendig und langwierig, so daß man

vorteilhafter von Gleichung (54) ausgeht. Für $\tau = 1 \dots p$ erhält man die sog. *Yule-Walker-Gleichungen* (Yule 1926, Walker 1931)

$$(59) \quad \begin{array}{rcccc} r_1 & = & b_1 & + & b_2 r_1 & + & b_3 r_2 & + & \dots & + & b_p r_{p-1} \\ r_2 & = & b_1 r_1 & + & b_2 & + & b_3 r_1 & + & \dots & + & b_p r_{p-2} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ r_p & = & b_1 r_{p-1} & + & b_2 r_{p-2} & + & b_3 r_{p-3} & + & \dots & + & b_p \end{array}$$

Hier ist zu beachten, daß $r_0 = 1$ gilt und infolge der Symmetrie des time-lags $r_{-i} = r_i$ angenommen werden kann. Dieses lineare Gleichungssystem läßt sich für beliebiges p lösen und ergibt die gesuchten Werte der Koeffizienten $b_1, b_2 \dots b_p$ für die Bestimmung der Ordnung des AR-Prozesses.

Setzt man zunächst $p = 1$, ergibt sich sofort $\hat{b}_1 = r_1$. Bei $p = 2$ erhält man \hat{b}_2 aus dem Gleichungssystem

$$\begin{aligned} r_1 &= b_1 + b_2 r_1 \quad \text{und} \quad r_2 = b_1 r_1 + b_2 \\ \hat{b}_2 &= \frac{r_2 - r_1^2}{1 - r_1^2} \end{aligned}$$

Ist $\hat{b}_2 \neq 0$, so prüft man nun, ob der Prozeß dritter Ordnung ist ($p=3$), indem man aus dem Gleichungssystem (59) für $p=3$ den Koeffizienten \hat{b}_3 ermittelt.

Dies Verfahren wird nun stufenweise ($p = 4, 5 \dots$) fortgesetzt. Der jeweils letzte Koeffizient in der Stufe p , nämlich \hat{b}_p , heißt partieller Autokorrelationskoeffizient, die Folge $\hat{b}_1, \hat{b}_2 \dots$ *partielle Autokorrelationsfunktion*. Diese etwas mißverständlichen Begriffe wurden von Box/Jenkins (1976, S. 64) gewählt, um ihren Zusammenhang mit den (echten) Autokorrelationskoeffizienten gemäß Gleichung (59) zu verdeutlichen und um darauf hinzuweisen, daß z. B. der Koeffizient \hat{b}_1 nur den isolierten, d. h. partiellen Einfluß von x_{t-1} auf x_t wiedergibt und dabei die Einflüsse von $x_{t-2}, x_{t-3} \dots$ vernachlässigt werden.

Die partiellen Autokorrelationskoeffizienten dienen also dazu, den Grad eines AR-Prozesses festzustellen:

Der AR-Prozeß ist von p -ter Ordnung, wenn die ersten p partiellen Autokorrelationskoeffizienten ungleich Null und alle weiteren gleich Null sind.

Mit Hilfe der echten und der partiellen Autokorrelationsfunktion kann man also die Existenz und den Grad, d. h. die Ordnung eines autoregressiven Prozesses bestimmen. In der Praxis gehört dazu jedoch eine gewisse Geschicklichkeit, da wir die Autokorrelationsfunktion aus einer Zeitreihe schätzen und damit Schätzfehler auftreten können. Diesem Problem werden wir in II. 4. auf S. 74 ff. nachgehen.

2. Moving-average-Prozesse (MA-Modelle)

Zur Definition der MA-Modelle wollen wir uns in Erinnerung zurückrufen, daß die beiden Grundgleichungen (49) und (51) auseinander hervorgehen und damit eine gewisse *Dualität* aufzeigen. Während die autoregressiven Prozesse auf Gleichung (51) und damit auf den Vergangenheitswerten der Zeitreihe basieren, gehen wir nun von Gleichung (49), d. h. von der Analyse der Störvariablen aus, wobei wieder $\mu = 0$ vorausgesetzt sei.

Wir sprechen von einem *Moving-average-Prozeß* q -ter Ordnung oder kurz von einem MA(q)-Modell zum Gegenwartszeitpunkt t , wenn folgende Gleichung vorliegt¹:

$$(60) \quad x_t = u_t - a_1 u_{t-1} - a_2 u_{t-2} - \dots - a_q u_{t-q}$$

Leider ist diese Benennung, die auf Wold (1938) zurückgeht, mißverständlich, da es sich nicht um gleitende Durchschnitte herkömmlicher Art handelt. Welcher Zusammenhang überhaupt besteht, wird klar, wenn wir folgende Umformungen für den Spezialfall $q = 1$ (MA-Prozeß erster Ordnung) vornehmen:

$$(61) \quad x_t = u_t - a_1 u_{t-1}$$

Wir setzen in (61) für den Index t sukzessive die Indizes $t-1, t-2 \dots$ ein

$$(61a) \quad x_{t-1} = u_{t-1} - a_1 u_{t-2}$$

$$(61b) \quad x_{t-2} = u_{t-2} - a_1 u_{t-3}$$

$$\begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{array}$$

lösen die Gleichungen nach $u_{t-1}, u_{t-2} \dots$ auf und substituieren in (61). Daraus resultiert die Gleichung

$$(62) \quad x_t = u_t - a_1 x_{t-1} - a_1^2 x_{t-2} - a_1^3 x_{t-3} - \dots$$

Aus dieser Gleichung können wir drei Erkenntnisse gewinnen:

- a) Ein *endlicher* MA-Prozeß (hier: erster Ordnung) kann durch eine *unendliche* Reihe gewogener Vergangenheitswerte der Zeitreihe, also *autoregressiv*, approximiert werden.
- b) Da x_t als gewogener *Durchschnitt* der Vergangenheitswerte $x_{t-1}, x_{t-2} \dots$ ausgedrückt wird, kann man diesen Sachverhalt als Erklärung für den Namen „Moving average“ ansehen.

¹ Die Minuszeichen sind hier nur wegen der Konformität mit der Literatur gewählt. Man könnte auch wie in (49) die Reihe mit positiven Vorzeichen versehen.

c) Der *Moving-average-Prozeß erster Ordnung* entspricht formal dem konstanten Modell der *exponentiellen Glättung*, die sich somit als Spezialfall des Box-Jenkins-Verfahrens erweist. Die exponentiell abnehmenden Gewichte müssen dabei die Bedingung $|a_1| < 1$ erfüllen.

Wie bestimmt man nun Existenz und Grad (Ordnung) eines MA-Prozesses? Im Hinblick auf die Dualität von AR- und MA-Prozessen liegt es nahe, zur Beantwortung dieser Frage wieder die echte und die partielle Autokorrelationsfunktion heranzuziehen.

Box und Jenkins (1976, S. 68 ff.) haben gezeigt, daß für einen MA(q)-Prozeß folgendes gilt:

- Die Autokorrelationskoeffizienten sind bis zum time-lag q (einschließlich) *ungleich Null*, nach q *gleich* Null.
- Die partiellen Autokorrelationskoeffizienten streben exponentiell oder in Form einer gedämpften Sinusschwingung gegen Null.

Wir finden also ein genau umgekehrtes Ergebnis im Vergleich zu AR-Prozessen. Die partielle Autokorrelationsfunktion entscheidet darüber, ob eine vorliegende Zeitreihe ein MA-Prozeß *ist*, die Autokorrelationsfunktion legt die Ordnung q des Prozesses fest.

Wir wollen dieses Ergebnis für den wichtigsten Spezialfall $q = 1$ herleiten, um die Zusammenhänge noch einmal klar hervortreten zu lassen.

Die Varianz R_0 des MA(1)-Prozesses

$$(63) \quad x_t = u_t - a_1 u_{t-1} \quad \text{lautet}$$

$$(64) \quad E[x_t^2] = R_0 = E[u_t^2 - 2a_1 u_t u_{t-1} + a_1^2 u_{t-1}^2]$$

Da die Realisation der Störvariablen in verschiedenen Perioden unkorreliert sind und wir für $E[u_t^2] = \sigma_u^2$ gesetzt hatten (vgl. S. 30 und 63) ergibt sich

$$(64a) \quad R_0 = \sigma_u^2 (1 + a_1^2)$$

Die Autokovarianz mit dem time-lag $\tau = q = 1$ lautet, da die Störvariablen verschiedener Zeitpunkte unkorreliert sind,

$$(64b) \quad R_1 = E[x_t x_{t-1}] = E[u_t u_{t-1} - a_1 u_t u_{t-2} - a_1 u_{t-1}^2 + a_1^2 u_{t-1} u_{t-2}] = -a_1 \sigma_u^2$$

Der *Autokorrelationskoeffizient* $r_1 = R_1/R_0$ ergibt sich zu

$$(64c) \quad r_1 = \frac{-a_1}{1 + a_1^2}$$

Der Autokorrelationskoeffizient für $\tau = 2 > q = 1$ muß Null sein, da

$$(65) \quad R_2 = E[x_t x_{t-2}] = E[u_t u_{t-2} - a_1 u_t u_{t-3} - a_1 u_{t-1} u_{t-2} + a_1^2 u_{t-1} u_{t-3}] = 0$$

Dies gilt auch für alle weiteren r_τ mit $\tau > q$.

Für die partielle Autokorrelationsfunktion kann man zunächst aus Gleichung (62) ableiten, daß sie *unendlich* viele Koeffizienten aufweist, die *nicht* plötzlich bei einem bestimmten time-lag auf Null abfallen.

Setzt man in den Yule-Walker-Gleichungen (S. 68) z. B. $p = 2$ und für r_1 gemäß (64c)

$\frac{-a_1}{1+a_1^2}$ sowie für r_2 den Wert Null (wegen (65)) ein, so ergibt sich

$$(66a) \quad \hat{b}_1 = -a_1 \cdot \frac{1}{1+a_1^2} \quad \text{und}$$

$$(66b) \quad \hat{b}_2 = -a_1^2 \cdot \frac{1}{1+a_1^2+a_1^4}$$

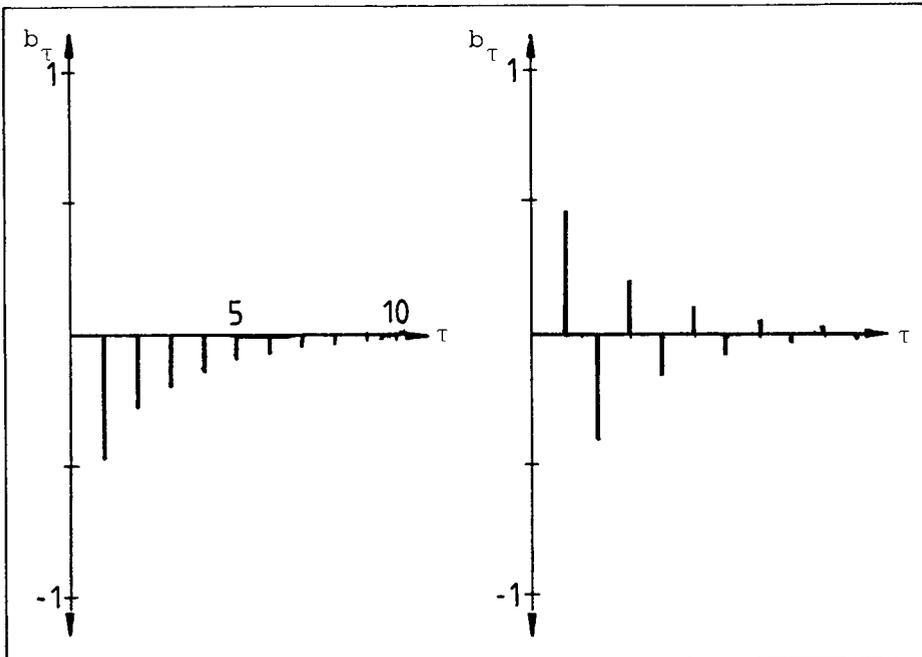


Abbildung 10
Partielle Autokorrelationskoeffizienten von (66a) und (66b)
 $|a_1| = 0,7$

und allgemein

$$(66c) \quad \hat{b}_\tau = -a_1^\tau \cdot \frac{1}{\sum_{k=0}^{\tau-1} a_1^{2k}} \quad (|a_1| < 1)$$

Die partiellen Autokorrelationskoeffizienten streben also mit zunehmendem time-lag τ *exponentiell* gegen Null, und zwar monoton, falls $a_1 > 0$ und mit wechselndem Vorzeichen (gedämpfter Sinusschwingung) falls $a_1 < 0$ ausfällt. Abbildung 10 zeigt dies für $|a_1| = 0,7$.

Damit haben wir die theoretischen Grundlagen für die Anpassung konkreter Box-Jenkins-Modelle an stationäre Zeitreihen gelegt und wenden uns nun dem nicht-stationären Fall zu.

3. Nicht-stationäre Prozesse

Um das Box-Jenkins-Verfahren auch auf nicht-stationäre stochastische Prozesse, die in der Ökonomie besonders häufig auftreten, anwenden zu können, müssen nicht-stationäre Zeitreihen zunächst in stationäre umgeformt werden. Dies geschieht mit Hilfe der *Differenzenbildung*.

Zur Erläuterung des Verfahrens nehmen wir einen Prozeß mit *linearem* Trend an

$$(67) \quad x_t = a + bt + u_t \quad (t = 0, 1, \dots, T)$$

Zur Eliminierung des Trends werden die *ersten* Differenzen gebildet

$$(68) \quad \Delta x_t = x_t - x_{t-1} = a + bt + u_t - a - b(t-1) - u_{t-1} = b + u_t - u_{t-1}$$

Dieser Differenzenprozeß ist *stationär*, da der Erwartungswert der Störvariablen u gleich Null und b konstant ist. Δx_t ($t = 1, 2, \dots, T$) ist nun die Input-Reihe für das Box-Jenkins-Verfahren.

Liegt ein *quadratischer* Trend vor, z. B.

$$(69) \quad x_t = a + bt + ct^2 + u_t \quad (t = 0, 1, \dots, T)$$

so sind die ersten Differenzen noch nicht stationär

$$(70) \quad \begin{aligned} \Delta x_t = x_t - x_{t-1} &= a + bt + ct^2 + u_t - a - b(t-1) - c(t-1)^2 - u_{t-1} \\ &= b + 2ct - c + u_t - u_{t-1} \end{aligned}$$

Erst die *zweiten Differenzen* erzeugen einen stationären Prozeß, der in das Box-Jenkins-Verfahren eingeht:

$$(70a) \quad \Delta^2 x_t = \Delta x_t - \Delta x_{t-1} = 2c + u_t - 2u_{t-1} + u_{t-2}$$

Entsprechend kann man Polynome n -ten Grades durch n -fache Differenzenbildung stationär machen. Dieselben Ergebnisse lassen sich auch durch *Differentiation* gewinnen, wenn wir gedanklich von diskreten Zeitreihen x_t ($t = 0, 1, \dots, T$) zu kontinuierlichen Prozessen $x(t) \in \mathbb{R}^+$ übergehen.

Dieses von Box und Jenkins vorgeschlagene Verfahren der Differenzenbildung verfehlt jedoch sein Ziel der Trendausschaltung, wenn *nicht-polynomiale* Trends wie z. B. logarithmische, exponentielle oder unregelmäßig strukturierte Trends vorliegen.

Bei folgendem logarithmischen Trend

$$(71) \quad x(t) = \ln t + u_t \quad (t > 0)$$

eliminiert die 1. Ableitung, die der Differenzenbildung im diskreten Fall entspricht, nicht den Trend, sondern transformiert den positiven Trend gemäß

$$(72) \quad \frac{dx}{dt} = \frac{1}{t} \quad (t > 0)$$

in einen *negativen* Trend. Stationarität wird auch bei weiteren Differentiationen nicht erreicht. Infolgedessen wird vom Verfasser ein anderes Verfahren, eine *modifizierte* Differenzenbildung vorgeschlagen.

Anstatt der einfachen Differenzen $x_t - x_{t-1}$ wird die Differenz

$$(73) \quad x_t - \rho x_{t-1} = \Delta x_t \quad (0 \leq \rho \leq 1)$$

zugrunde gelegt. Durch Variation von ρ läßt sich das Verfahren sehr flexibel an viele unterschiedliche – auch unregelmäßige – Trends anpassen. Bei polynomialen Trends wird ρ gleich eins gesetzt, in den übrigen Fällen sollte der Autokorrelationskoeffizient erster Ordnung (r_1) als Initialwert für ρ verwendet werden, da z. B. bei einem exponentiellen Trend der Form

$$(74) \quad x_t = c\rho^t \quad (|\rho| < 1)$$

die modifizierte Differenzenbildung

$$(74a) \quad x_t - \rho x_{t-1} = c\rho^t - \rho c\rho^{t-1} = 0$$

unmittelbar zur Stationarität führt. Solche Trends, die sich durch den Autokorrelationskoeffizienten erster Ordnung approximativ beschreiben lassen, treten in der ökonomischen Praxis relativ häufig auf. Im computergestützten Prognosesystem des Verfassers ist daher die modifizierte Differenzenbildung implementiert.

Wir sind nunmehr auch in der Lage, *saisonbehaftete* Zeitreihen mit dem Box-Jenkins-Verfahren zu behandeln. Im letzten Kapitel (S. 53 ff.) wurde gezeigt, daß sich Saison-

zyklen entweder durch hohe Autokorrelationskoeffizienten im Korrelogramm oder durch hohe Werte des Spektrums der Zeitreihe sichtbar machen lassen. Hat man auf diese Weise die *Saisonalänge* ermittelt, kann der Saisoneinfluß durch Bildung der Differenzen

$$(75) \quad x_t - \rho x_{t-L} \quad (L = \text{Länge des Saisonzyklus})$$

eliminiert werden. Diese Saisonausschaltung wird üblicherweise mit der Trendausschaltung kombiniert, indem von (75) die ersten Differenzen zur Beseitigung des Trends errechnet werden.

Erst wenn man mit den beschriebenen Verfahren – eventuell nach mehreren Versuchen – einen stationären Prozeß erzeugt hat, darf der nächste Schritt, die Auswahl des für diesen Prozeß geeignetsten Modells erfolgen.

4. Identifizierung des geeigneten ARIMA-Modells

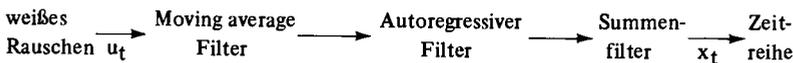
a) Kennzeichnung von ARIMA-Modellen

Autoregressive-, Moving-average- und Differenzenprozesse sind die Grundbausteine aller Box-Jenkins-Modelle. Durch Anwendung dieser drei Filtermethoden auf „weißes Rauschen“ versucht man, die vorgelegte Zeitreihe zu erzeugen, d. h. „herauszufiltern“. Dabei ist zu beachten, daß der Differenzenfilter umzukehren ist, wenn aus einem stationären Prozeß der ehemalige nicht-stationäre Prozeß wiedergewonnen werden soll.

Diesen inversen Filter nennt man „Summenfilter“

$$(76) \quad x_t = \Delta x_t + \Delta x_{t-1} + \Delta x_{t-2} + \dots = \sum_{i=0}^{\infty} \Delta x_{t-i}$$

Graphisch lassen sich die drei Filterprozesse folgendermaßen veranschaulichen.



Das allgemeine Box-Jenkins-Modell verwendet alle drei Filter, ist also eine Kombination aus den in den Abschnitten 1. bis 3. dieses Kapitels besprochenen Verfahren. Box und Jenkins haben diesem allgemeinen Modell den Namen

„autoregressive integrated moving average model“

abgekürzt ARIMA-Modell gegeben, wobei das Wort „integrated“ den Summenfilter bezeichnen soll. Nicht-stationäre Zeitreihen werden daher in einem ARIMA-Modell, sta-

tionäre in einem ARMA-Modell abgebildet. Bei der Analyse der echten und der partiellen Autokorrelationsfunktion muß sich herausstellen, ob ein AR-, ein MA- oder ein gemischtes Modell problemadäquat ist.

b) Identifizierung des passenden Modelltyps

Zur Ermittlung des adäquaten Modells untersuchen wir zunächst die vorgelegte Zeitreihe hinsichtlich ihrer Stationarität. Eine nicht-stationäre, trendbehaftete Zeitreihe ist z. B. die Umlaufrendite festverzinslicher Wertpapiere in der BRD, die in Abbildung 11 wiedergegeben ist².

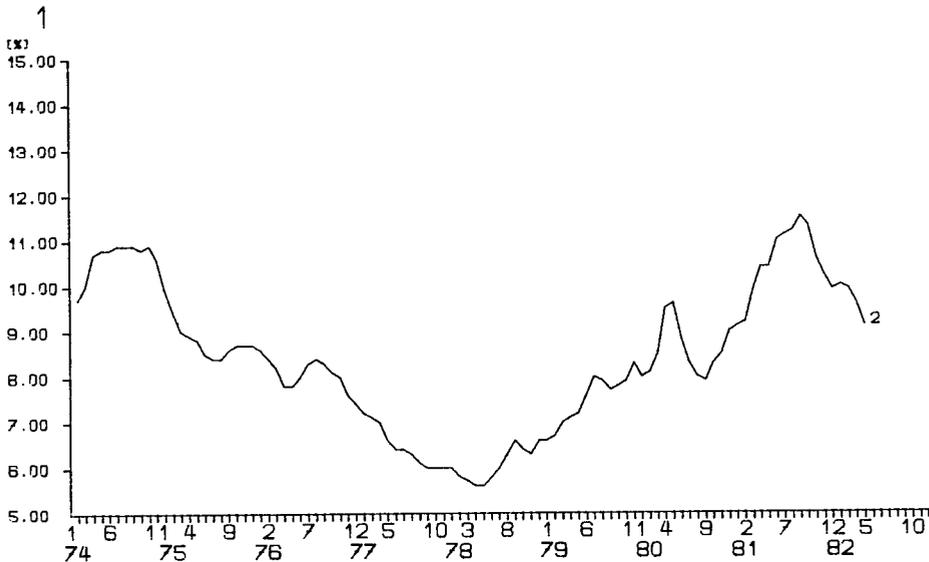


Abbildung 11
Umlaufrendite festverzinslicher Wertpapiere in der BRD

Diese Reihe zeigt ein typisches Verhalten ihrer Autokorrelationsfunktion, die in der folgenden Abbildung 12 wiedergegeben ist.

² Stat. Beihefte zu den Monatsberichten der Deutschen Bundesbank, Reihe 2, Ziffer 8b, Werte von Januar 1974 bis April 1982.

LAG	-1	0																		+1	WERT			
	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	
1																							*	.978
2																							*	.947
3																							*	.906
4																							*	.861
5																							*	.815
6																							*	.763
7																							*	.706
8																							*	.649
9																							*	.592
10																							*	.538
11																							*	.484
12																							*	.435
13																							*	.391
14																							*	.351
15																							*	.315
16																							*	.280
17																							*	.242
18																							*	.200
19																							*	.159
20																							*	.119
21																							*	.079
22																							*	.034
23																							*	-.018
24																							*	-.074
25																							*	-.132

STANDARDABWEICHUNG-UNTERGRENZE : -.196
STANDARDABWEICHUNG-OBERGRENZE : .196
VERTRAUENSBEREICH : 95% (ANNAHME: NORMALVERTEILUNG)

Abbildung 12
Autokorrelationsfunktion einer nicht-stationären Zeitreihe
(Umlaufrendite festverzinslicher Wertpapiere in der BRD)

Die Graphik zeigt ein beinahe lineares Abnehmen der Autokorrelationskoeffizienten mit wachsendem time lag, das durch den Trend erklärt werden kann. Bei *geringem* time lag haben die Abweichungen der trendbehafteten Zeitreihenwerte von ihrem Mittelwert sehr häufig das gleiche Vorzeichen, was gemäß Gleichung (46)

$$(46) \quad \hat{R}(\tau) = \frac{1}{T} \sum_{t=1}^{T-\tau} (x_t - \bar{x}) (x_{t+\tau} - \bar{x}) \quad \tau = 1, 2, \dots$$

zu hohen Autokovarianzen und ebenfalls zu hohen Autokorrelationskoeffizienten führt. Mit zunehmender zeitlicher Verschiebung weisen $(x_t - \bar{x})$ und $(x_{t+\tau} - \bar{x})$ immer

ist und welche Werte für die Parameter p und q gewählt werden sollten. Zur Beantwortung dieser Fragen dienen uns die in den Unterabschnitten 1. und 2. angestellten Überlegungen, die hier noch einmal kurz zusammengefaßt werden sollen.

1. Die Autokorrelationsfunktion eines AR-Prozesses fällt mit zunehmendem time lag τ exponentiell oder in Form einer gedämpften Sinusschwingung gegen Null ab.
2. Ein AR-Prozeß ist von p -ter Ordnung, wenn die ersten p *partiellen* Autokorrelationskoeffizienten ungleich Null und die restlichen gleich Null sind.
3. Die *partielle* Autokorrelationsfunktion eines MA-Prozesses strebt mit zunehmendem time lag τ exponentiell oder in Form einer gedämpften Sinusschwingung gegen Null.
4. Ein MA-Prozeß ist von q -ter Ordnung, wenn die ersten q Autokorrelationskoeffizienten ungleich Null und die übrigen gleich Null sind.
5. Die Autokorrelationsfunktion eines gemischten ARMA-Prozesses strebt nach einem time lag von $(q - p)$ Perioden wie die eines AR-Prozesses gegen Null, während sich die partielle Autokorrelationsfunktion des ARMA-Prozesses nach $(p - q)$ Perioden wie diejenige eines MA-Prozesses verhält. Falls die Differenzen $(q - p)$ bzw. $(p - q)$ negativ sind, werden sie gleich Null gesetzt.

Diese auf den ersten Blick eindeutig erscheinenden Kriterien zur Bestimmung des geeigneten ARMA-Modells verlieren aus zwei Gründen in der Praxis an Unterscheidungskraft:

- Eine vorgelegte Zeitreihe ist nur eine *Realisation* des zugrunde liegenden stochastischen Prozesses und nicht der Prozeß selbst.
- Die echten und partiellen Autokorrelationskoeffizienten müssen aus der Zeitreihe *geschätzt* werden und sind daher mit Unsicherheit behaftet. Insbesondere die Frage, ob sie ungleich Null sind, kann nur mit einem statistischen Test beantwortet werden.

Diese Ausführungen machen deutlich, daß die Identifizierung eines ARMA-Prozesses aufgrund einer Zeitreihe ein äußerst schwieriges und nicht immer zufriedenstellend zu lösendes Problem darstellt. Ehe wir auf diese Problematik weiter eingehen, soll jedoch zunächst entschieden werden, unter welcher Bedingung wir einen Autokorrelationskoeffizienten aufgrund einer empirischen Schätzung als von Null verschieden ansehen wollen.

Zur Konstruktion eines aussagefähigen Tests benötigen wir die Standardabweichung der geschätzten Autokorrelationskoeffizienten. Bartlett (1946) hat hierfür einen Ausdruck entwickelt:

$$(77) \quad \hat{\sigma}(\hat{r}_k) = \frac{1}{\sqrt{n}} \sqrt{1 + 2(\hat{r}_1^2 + \hat{r}_2^2 + \dots + \hat{r}_q^2)} \quad (k > q)$$

n bedeutet hier die Anzahl der Zeitreihenwerte und q einen bestimmten time lag. Diejenigen \hat{r}_k , deren theoretischer Wert Null ist, sind dabei, wie Anderson (1942) gezeigt hat, annähernd normalverteilt.

Ein ähnliches Ergebnis hat Quenouille (1949) für die partiellen Autokorrelationskoeffizienten abgeleitet. Sie sind ebenfalls annähernd normalverteilt und ihre Standardabweichung lautet bei einem AR(p)-Prozeß

$$(78) \quad \hat{\sigma}(\hat{r}_k \text{ partiell}) = \frac{1}{\sqrt{n}} \quad (k \geq p + 1)$$

In der Literatur wird (77) meistens vereinfacht (vgl. Makridakis/Wheelwright 1978a, S. 337) zu

$$\hat{\sigma}(\hat{r}_k) = \frac{1}{\sqrt{n}}$$

und damit formal (78) angeglichen. Wir schließen uns diesem Vorgehen an und formulieren folgenden Signifikanztest:

Ein echter oder partieller Autokorrelationskoeffizient ist mit 95 % Sicherheitswahrscheinlichkeit – signifikant – von Null verschieden, wenn sein Schätzwert \hat{r}_k außerhalb des Bereichs $\pm 1,96 / \sqrt{n}$ liegt.

Für $-1,96 / \sqrt{n} < \hat{r}_k < 1,96 / \sqrt{n}$ nehmen wir die Null-Hypothese: $r_k = 0$ an.

Dieses Konfidenzintervall ist in den Abbildungen 12 und 13 mit senkrechten gepunkteten Linien gekennzeichnet. Es findet sich auch in jeder weiteren Graphik von Autokorrelationsfunktionen. Die Abbildung 14 zeigt die partielle Autokorrelationsfunktion unserer stationären Beispiel-Zeitreihe.

Mit Hilfe der Abbildungen 13 und 14 versuchen wir nun, das geeignete Modell für diese Zeitreihe zu identifizieren.

Dies gestaltet sich nicht ganz einfach, da nur jeweils der erste echte und partielle Autokorrelationskoeffizient signifikant von Null verschieden ist und der Verlauf der Funktionen nicht zweifelsfrei als exponentiell abnehmend bzw. gedämpft sinusförmig identifiziert werden kann. Aber auch die gegenteilige Annahme, daß die Koeffizienten nach dem time lag von einer Periode abrupt auf Null fallen, läßt sich durch die Gestalt der Funktionen nicht stützen.

Wir können versuchsweise anhand der Graphiken folgende Vermutung anstellen: Die wahre Autokorrelationsfunktion des Prozesses strebt exponentiell gegen Null, die partielle Autokorrelationsfunktion strebt in gedämpften Sinusschwingungen gegen Null. Dann liegt nach Regel 5. auf S. 78 ein ARMA-Prozeß vor. Nach Regel 2. ist der AR-Teil von erster Ordnung und nach Regel 4. der MA-Teil ebenfalls von erster Ordnung. Da wir die ersten Differenzen ($d = 1$) gebildet hatten, identifizieren wir die *Umlaufrendite* als einen ARIMA (p,d,q)-Prozeß mit $p = 1$, $d = 1$, $q = 1$ oder kurz: ARIMA (1,1,1)-Prozeß mit

$$\Delta x_t = b_1 \Delta x_{t-1} + u_t - a_1 u_{t-1}$$

LAG	-1	0	+1	WERT
1	I	I	*	.501
2	.	* I	.	-.159
3	.	I *	.	.086
4	.	*	.	.007
5	.	I *	.	.105
6	.	* I	.	-.032
7	.	* I	.	-.136
8	.	I *	.	.083
9	.	* I	.	-.108
10	.	I *	.	.098
11	.	* I	.	-.106
12	.	I *	.	.062
13	.	* I	.	-.118
14	.	I *	.	.091
15	.	*	.	.009
16	.	I *	.	.141
17	.	* I	.	-.044
18	.	* I	.	-.037
19	.	* I	.	-.059
20	.	I *	.	.085
21	.	*	.	.016
22	.	I *	.	.060
23	.	*	.	.022
24	.	* I	.	-.142
25	.	*	.	.012

STANDARDABWEICHUNG-UNTERGRENZE : -.197
STANDARDABWEICHUNG-OBERGRENZE : .197
VERTRAUENSBEREICH : 95% (ANNAHME:NORMALVERTEILUNG)

Abbildung 14
Partielle Autokorrelationsfunktion einer stationären Zeitreihe
(Differenzenreihe der Umlaufrendite)

Die Eignung dieses (von uns vermuteten) Modells läßt sich nun testen, indem wir a_1 und b_1 so schätzen, daß die Summe der quadrierten Residuen (Abweichungen der wahren von den geschätzten Reihenwerten) minimal wird, und anschließend prüfen, ob die Residuen *unkorreliert* sind. Diese Prüfung, die im Anschluß an den nächsten Abschnitt vorgenommen wird (vgl. S. 88), fällt positiv aus. Zuvor muß aber das Problem der Schätzung von a_1 und b_1 gelöst werden.

5. Schätzung der Parameter

a) Anfangsschätzungen

Die Schätzung der Parameter eines ARMA-Modells ist mathematisch aufwendig, so daß man sie zweckmäßigerweise in mehrere Stufen unterteilt. Zunächst werden Anfangswerte für die Parameter geschätzt, die in einem zweiten Schritt mit einem nicht-linearen Verfahren der Kleinsten Quadrate optimiert werden.

Am leichtesten finden wir die Anfangswerte für einen *AR(p)-Prozeß*

$$(52) \quad x_t = b_1 x_{t-1} + b_2 x_{t-2} + \dots + b_p x_{t-p} + u_t$$

durch Anwendung der Yule-Walker-Gleichungen (59). Mit den aus der Zeitreihe berechneten Werten für die p Autokorrelationskoeffizienten $\hat{r}_1, \hat{r}_2, \dots, \hat{r}_p$ läßt sich dieses lineare Gleichungssystem einfach lösen. Entsprechende Prozeduren sind an den meisten EDV-Anlagen implementiert.

Hat man z. B. einen *AR(2)-Prozeß* identifiziert, so lautet das Gleichungssystem (59)

$$\begin{aligned} \hat{r}_1 &= \hat{b}_1 + \hat{b}_2 \hat{r}_1 \\ \hat{r}_2 &= \hat{b}_1 \hat{r}_1 + \hat{b}_2 \end{aligned}$$

Daraus ergeben sich die Anfangsschätzungen

$$\hat{b}_1 = \frac{\hat{r}_1(1 - \hat{r}_2)}{1 - \hat{r}_1^2} \quad \text{und} \quad \hat{b}_2 = \frac{\hat{r}_2 - \hat{r}_1^2}{1 - \hat{r}_1^2}$$

In ähnlicher Weise werden die \hat{b}_i ($i = 1, 2, \dots, p$) für höhere AR-Prozesse geschätzt.

Schwieriger gestaltet sich die Anfangsschätzung für *MA-Prozesse*. Man geht dazu von den geschätzten Autokorrelationskoeffizienten \hat{r}_k des MA-Prozesses

$$(60) \quad x_t = u_t - a_1 u_{t-1} - a_2 u_{t-2} - \dots - a_q u_{t-q}$$

aus, die man durch Bildung des Erwartungswertes $E[x_t x_{t-k}]$ und anschließende Division durch die Varianz erhält.

$$(79) \quad E[x_t x_{t-k}] = \hat{R}_k = E[(u_t - a_1 u_{t-1} - \dots - a_q u_{t-q})(u_{t-k} - a_1 u_{t-k-1} - \dots - a_q u_{t-k-q})]$$

$$(79a) \quad \begin{aligned} \hat{R}_k &= E[u_t u_{t-k} - a_1 u_t u_{t-k-1} - \dots - a_q u_t u_{t-k-q} \\ &\quad - a_1 u_{t-1} u_{t-k} + a_1^2 u_{t-1} u_{t-k-1} + \dots + a_1 a_q u_{t-1} u_{t-k-q} - \dots \\ &\quad - a_q u_{t-q} u_{t-k} + a_1 a_q u_{t-q} u_{t-k-1} + \dots + a_q^2 u_{t-q} u_{t-k-q}] \end{aligned}$$

Dabei gilt $E[u_i u_j] = \sigma_u^2$, falls $i=j$ und $E[u_i u_j] = 0$ für $i \neq j$.

Setzen wir $k = 0$, ergibt sich die Varianz der Zeitreihe

$$(80) \quad \hat{R}_0 = \sigma_u^2 (1 + a_1^2 + a_2^2 + \dots + a_q^2)$$

Die Autokovarianz mit einem time lag von *einer* Periode ($k = 1$) hat folgende Gestalt:

$$(81) \quad \hat{R}_1 = \sigma_u^2 (-a_1 + a_1 a_2 + a_2 a_3 + \dots + a_{q-1} a_q)$$

Daraus folgt der Autokorrelationskoeffizient \hat{r}_1 zu

$$(82) \quad \hat{r}_1 = \frac{\hat{R}_1}{\hat{R}_0} = \frac{-a_1 + a_1 a_2 + a_2 a_3 + \dots + a_{q-1} a_q}{1 + a_1^2 + a_2^2 + \dots + a_q^2}$$

Nach dieser Methode kann man alle \hat{r}_k ermitteln. Aus (82) ist bereits ersichtlich, daß die \hat{r}_k *nicht-lineare* Funktionen der a_i sind. Dies bestätigt der allgemeinen Fall:

$$(83) \quad \hat{r}_k = \frac{-a_k + a_1 a_{k+1} + a_2 a_{k+2} + \dots + a_{q-k} a_q}{1 + a_1^2 + a_2^2 + \dots + a_q^2} \quad k = 1, 2, \dots, q$$

Man erhält auf diese Weise ein *nicht-lineares Gleichungssystem* von q Gleichungen und q Unbekannten a_1, a_2, \dots, a_q , das nur *iterativ* gelöst werden kann.

Für einen MA(2)-Prozeß sei der *Iterationsvorgang beispielhaft* durchgeführt. Wir gehen von Gleichung (79a) aus und setzen $\hat{R}_0 = 2,6$, $\hat{R}_1 = 1,3$ und $\hat{R}_2 = 0,5$ voraus. Explizit lautet das Gleichungssystem

$$(84) \quad \begin{aligned} \hat{R}_0 &= \sigma_u^2 (1 + a_1^2 + a_2^2) \\ \hat{R}_2 &= \sigma_u^2 (-a_2) \\ \hat{R}_1 &= \sigma_u^2 (-a_1 + a_1 a_2) \end{aligned}$$

das folgendermaßen umgeschrieben wird

$$(84a) \quad \begin{aligned} \sigma_u^2 &= \frac{\hat{R}_0}{1 + a_1^2 + a_2^2} \\ a_2 &= -\frac{\hat{R}_2}{\sigma_u^2} \\ a_1 &= -\frac{\hat{R}_1}{\sigma_u^2} + a_1 a_2 \end{aligned}$$

Zu Beginn des Iterationsprozesses setzen wir $a_1 = a_2 = 0$ und erhalten mit den oben angegebenen $\hat{R}_0 = 2,6$, $\hat{R}_1 = 1,3$ und $\hat{R}_2 = 0,5$ im ersten Schritt $\sigma_u^2 = 2,6$, $a_2 = -0,19$ und $a_1 = -0,5$.

Diese Werte werden sukzessive in die rechten Seiten von (84a) eingesetzt und ergeben in der zweiten Iteration $\sigma_u^2 = 2,02$, $a_2 = -0,248$ und $a_1 = -0,52$. Nach 10 Iterationen mit dem Taschenrechner haben sich die Werte auf drei Stellen hinter dem Komma stabilisiert:

$$\sigma_u^2 = 1,914 \quad a_1 = -0,538 \quad a_2 = -0,261$$

Mit diesem Iterationsverfahren lassen sich Anfangsschätzwerte für beliebige MA(q)-Prozesse ermitteln. Auf den von Box-Jenkins (1976, S. 203) vorgeschlagenen schneller konvergierenden, aber nicht so robusten Newton-Raphson-Algorithmus sei an dieser Stelle nur hingewiesen. In dem am Lehrstuhl des Verfassers entwickelten EDV-Programm zum Box-Jenkins-Verfahren sind beide Iterationsverfahren enthalten³.

Liegt ein (gemischtes) ARMA (p,q)-Modell

$$(85) \quad x_t = b_1 x_{t-1} + b_2 x_{t-2} + \dots + b_p x_{t-p} + u_t - a_1 u_{t-1} - a_2 u_{t-2} - \dots - a_q u_{t-q}$$

vor, so schätzt man die Anfangswerte $b_1, b_2, \dots, b_p, a_1, a_2, \dots, a_q$ in drei Schritten:

1. Schätzung der b_1, b_2, \dots, b_p mit Hilfe der Yule-Walker-Gleichungen;
2. Bildung der Reihe $w_t = x_t - \hat{b}_1 x_{t-1} - \hat{b}_2 x_{t-2} - \dots - \hat{b}_p x_{t-p}$;
3. Schätzung der a_1, a_2, \dots, a_q aus der Reihe $w_t = u_t - a_1 u_{t-1} - a_2 u_{t-2} - \dots - a_q u_{t-q}$ mit dem oben angegebenen Iterationsverfahren.

Mit diesen Anfangswerten beginnt man das eigentliche Schätzverfahren, das im nächsten Unterabschnitt beschrieben wird.

b) Das nicht-lineare Hauptschätzverfahren

Es ist unser Ziel, die Parameter a_1, a_2, \dots, a_q und b_1, b_2, \dots, b_p des allgemeinen ARMA(p, q)-Modells

$$(85) \quad x_t = b_1 x_{t-1} + b_2 x_{t-2} + \dots + b_p x_{t-p} + u_t - a_1 u_{t-1} - a_2 u_{t-2} - \dots - a_q u_{t-q}$$

so festzulegen, daß die Summe der quadrierten Schätzfehler (Residuen) minimal wird.

$$(86) \quad Z = \sum_{t=1}^T u_t^2 \Rightarrow \text{Min!}$$

Dieses Verfahren der sog. *Kleinste-Quadrate-Schätzung* (KQ), das ausführlich im Kapitel G. beschrieben wird, liefert für die Praxis ausreichende Approximationen für Maximum-likelihood-Schätzungen (vgl. Box-Jenkins 1976, S. 213, 243 und 279).

³ Unveröffentlichte Diplomarbeit von K. Sabathil, *Ausgewählte Prognoseverfahren der Unternehmensplanung und ihre EDV-technische Realisierung*, Hamburg 1981

Liegt nur ein AR(p)-Prozeß vor, so hängt die Störvariable u_t linear von den Parametern b_1, b_2, \dots, b_p ab. Für die Anwendung des KQ-Verfahrens ist somit nur das lineare Gleichungssystem

$$\frac{\partial Z}{\partial b_i} = 0 \quad (i = 1, 2, \dots, p) \text{ zu lösen.}$$

Sobald aber ein MA-Teil im Modell enthalten ist, treten nicht-lineare Beziehungen zwischen u_t und den Parametern a_1, a_2, \dots, a_q auf. Das sieht man sofort, wenn wir uns die Gleichungen (60) bis (62) vergegenwärtigen, die dort für einen MA(1)-Prozeß abgeleitet wurden. Schon für diesen einfachen Spezialfall erhielten wir die nicht-lineare Beziehung

$$(62) \quad u_t = x_t + a_1 x_{t-1} + a_1^2 x_{t-2} + a_1^3 x_{t-3} + \dots$$

Bei höheren MA-Prozessen, d.h. $q \geq 2$, ist u_t auch mit den Parametern a_2, a_3, \dots, a_q nicht-linear verknüpft. Wir werden uns deshalb bei der weiteren Analyse des Hauptschätzverfahrens auf MA-Prozesse beschränken, um die Ableitungen etwas durchsichtiger zu gestalten. Selbstverständlich können auch gemischte ARMA-Modelle mit dem Verfahren gelöst werden.

Ausgangspunkt sind die Anfangswerte $a_1^0, a_2^0, \dots, a_q^0$ aus Unterabschnitt a) sowie vor dem Schätzzeitraum liegende Werte für $u_0, u_{-1}, u_{-2}, \dots, u_{1-q}$, die man gleich Null setzt oder aus früheren Schätzungen übernimmt. Hieraus läßt sich rekursiv

$$(87) \quad u_t^0 = x_t + a_1^0 u_{t-1} + a_2^0 u_{t-2} + \dots + a_q^0 u_{t-q} \quad (t = 1, \dots, T)$$

berechnen.

Um die Anfangswerte zu verbessern, entwickeln wir die Funktion $u_t = f(a_1, a_2, \dots, a_q)$ in eine Taylor-Reihe⁴, die wir nach dem linearen Term abbrechen.

$$(88) \quad u_t = u_t^0 + \frac{\partial u_t}{\partial a_1}(a_1^0) [a_1 - a_1^0] + \frac{\partial u_t}{\partial a_2}(a_2^0) [a_2 - a_2^0] + \dots + \frac{\partial u_t}{\partial a_q}(a_q^0) [a_q - a_q^0]$$

Die partiellen Ableitungen werden für EDV-Programme zweckmäßigerweise *numerisch* ausgewertet nach der Vorschrift

$$(89) \quad \frac{\partial u_t}{\partial a_i} = [u_t(a_1, a_2, \dots, a_i + \delta_i, \dots, a_q) - u_t(a_1, a_2, \dots, a_i, \dots, a_q)] / \delta_i$$

⁴ Zum Begriff und zur Anwendung von Taylor-Reihen vgl. z. B. H. Körth et al. (Hrsg.), Lehrbuch der Mathematik für Wirtschaftswissenschaften, 3. Aufl., Opladen 1975, S. 448 ff.

Setzen wir zur Vereinfachung der Schreibweise

$$\frac{\partial u_t}{\partial a_i} = G_{ti}$$

und beachten (88), können wir folgende Zielfunktion formulieren:

$$(90) \quad Z = \sum_{t=1}^T u_t^2 = \sum_{t=1}^T [u_t^0 + G_{t1}^0 (a_1 - a_1^0) + G_{t2}^0 (a_2 - a_2^0) + \dots + G_{tq}^0 (a_q - a_q^0)]^2 \Rightarrow \text{Min!}$$

Dies Problem kann mit der einfachen Methode der kleinsten Quadrate gelöst werden, da die partiellen Ableitungen $\partial Z/\partial a_i$ linear in den Parametern a_i sind. Die aus der Lösung des linearen Gleichungssystems $\partial Z/\partial a_i = 0$ ($i = 1, \dots, q$) resultierenden Werte für $\hat{a}_1^1, \hat{a}_2^1, \dots, \hat{a}_q^1$ treten nun an die Stelle der Anfangswerte $a_1^0, a_2^0, \dots, a_q^0$. Damit ist die erste Iteration des nicht-linearen Optimierungsprozesses abgeschlossen.

In der zweiten Iteration werden die verbesserten Parameterwerte $\hat{a}_1^1, \dots, \hat{a}_q^1$ in die Gleichungen (87), (88) und (89) eingesetzt, um anschließend die Zielfunktion (90) wieder zu minimieren. Das Iterationsverfahren wird so lange fortgesetzt, bis sich die Parameterwerte $\hat{a}_1^j, \hat{a}_2^j, \dots, \hat{a}_q^j$ von ihren Vorgängern $\hat{a}_1^{j-1}, \hat{a}_2^{j-1}, \dots, \hat{a}_q^{j-1}$ nur noch um einen vorgegebenen Betrag ϵ unterscheiden.

Die Konvergenz des Verfahrens hängt sehr stark von der geschickten Wahl der Anfangswerte $a_1^0, a_2^0, \dots, a_q^0$ ab. Wenn diese mit dem im Unterabschnitt a) beschriebenen Iterationsverfahren berechnet wurden, konvergierte das Hauptschätzverfahren in allen bisher vom Verfasser analysierten empirischen Zeitreihen mit einer Maximalanzahl von 16 Iterationen.

Durch den Abbruch der Taylor-Entwicklung nach dem linearen Glied haben wir also den nicht-linearen Optimierungsprozeß durch eine Folge von linearen Optimierungsprozessen ersetzt, die mit der KQ-Methode gelöst werden. Zur Verdeutlichung soll dieser Prozeß noch einmal an einem kleinen numerischen Beispiel nachvollzogen werden.

Es sei ein MA(1)-Modell von folgender stationärer Zeitreihe identifiziert worden⁵:

$x_1 = -0,82 \quad x_2 = -1,22 \quad x_3 = 1,11 \quad x_4 = -0,36$. Der geschätzte Autokorrelationskoeffizient \hat{r}_1 betrage $-0,31$.

Nach Unterabschnitt a) berechnen wir zunächst den Anfangsschätzwert \hat{a}_1^0 aus Gleichung (83)

$$-0,31 = \frac{-\hat{a}_1}{1 + \hat{a}_1^2} \quad \text{wobei wir beachten, daß}$$

⁵ 4 Zeitreihenwerte sind natürlich zur Anwendung von ARMA-Modellen zu wenig und wurden nur aus rechenökonomischen Gründen gewählt.

$-1 < \hat{a}_1 < 1$ gelten muß. Es ergibt sich

$$\hat{a}_1^0 = 0,34$$

Das Versuchsmodell lautet nun $u_t = x_t + 0,34 u_{t-1}$

Um die *erste Iteration* starten zu können, benötigen wir den Wert $u_0 = 0,5$ (aus vorangegangener Schätzung). Die partiellen Ableitungen werden mit der Schrittweite $\delta_1 = 0,01$ errechnet, d. h. an den Stellen $\hat{a}_1^0 = 0,34$ und $\hat{a}_1^0 + \delta_1 = 0,35$.

Die Werte für x_t , u_t^0 und G_{t1}^0 enthält folgende Tabelle:

Tabelle 9

t	x_t	$u_t^0 (\hat{a}_1 = 0,34)$	$u_t (\hat{a}_1 = 0,35)$	G_{t1}^0
0	—	0,50	0,500	—
1	-0,82	-0,65	-0,645	0,5
2	-1,22	-1,44	-1,446	-0,6
3	1,11	0,62	0,604	-1,6
4	-0,36	-0,15	-0,149	0,1

Die Zielfunktion (90) bekommt für dieses Beispiel folgendes Aussehen:

$$Z = \sum_{t=1}^4 u_t^2 = \sum_{t=1}^4 [u_t^0 + G_{t1}^0 (a_1 - \hat{a}_1^0)] \Rightarrow \text{Min!}$$

Notwendige Bedingung für ein Minimum ist⁶

$$\frac{\partial Z}{\partial a_1} = 2 \sum_{t=1}^4 [u_t^0 + G_{t1}^0 (a_1 - \hat{a}_1^0)] \cdot G_{t1}^0 = 0$$

Daraus ergibt sich

$$(91) \quad \hat{a}_1 = \hat{a}_1^0 - \frac{\sum_{t=1}^4 u_t^0 G_{t1}^0}{\sum_{t=1}^4 (G_{t1}^0)^2}$$

⁶ Die hinreichende Bedingung $\partial^2 Z / \partial a_1^2 = 2 \sum_{t=1}^4 (G_{t1}^0)^2 > 0$ ist stets erfüllt.

Aus den Werten der Tabelle 9 lassen sich die benötigten Summen berechnen

$$\hat{a}_1 = 0,34 + 0,147 = 0,487$$

Wir schneiden die dritte Dezimalstelle ab und starten mit dem Schätzwert $\hat{a}_1^1 = 0,48$ die *zweite* Iteration mit dem Modell

$$u_t = x_t + 0,48 u_{t-1}$$

Die charakteristischen Werte enthält Tabelle 10

Tabelle 10

t	x_t	$u_t^1 (\hat{a}_1 = 0,48)$	$u_t (\hat{a}_1 = 0,49)$	G_{t1}^1
0	—	0,5000	0,5000	—
1	-0,82	-0,5800	-0,5750	0,50
2	-1,22	-1,4984	-1,5018	-0,34
3	1,11	0,3900	0,3740	-1,60
4	-0,36	-0,1724	-0,1767	-0,43

Nach Auswertung der Summen von Gleichung (91) ergibt sich

$$\hat{a}_1 = 0,48 + 0,106 = 0,586$$

Die *dritte Iteration* wird mit dem Schätzwert

$$\hat{a}_1 = 0,58$$

und dem Modell $u_t = x_t + 0,58 u_{t-1}$ gestartet.

Nach vier Iterationen stabilisiert sich der Parameterwert, so daß der Iterationsprozeß abgebrochen werden kann und als *endgültiges Prognosemodell*

$$x_t = u_t - 0,614 u_{t-1}$$

erhalten wird. Tabelle 11 faßt den Iterationsprozeß zusammen.

Damit ist die Darstellung des Hauptschätzverfahrens, das – in etwas modifizierter Form – dem in diesem Buch benutzten EDV-Programm zugrunde liegt, abgeschlossen. Wie oben schon erwähnt, dient es nicht nur zur Schätzung der MA-Parameter, sondern wird generell bei ARMA(p,q)-Modellen angewandt.

Tabelle 11

Iteration	Schätzwert \hat{a}_1
0	0,340
1	0,480
2	0,580
3	0,609
4	0,614

6. Prüfung des Prognosemodells

Nach erfolgter Schätzung der Modellparameter kann bei stationären Zeitreihen die Gleichung (85)

$$(85) \quad x_t = b_1 x_{t-1} + b_2 x_{t-2} + \dots + b_p x_{t-p} + u_t - a_1 u_{t-1} - a_2 u_{t-2} - \dots - a_q u_{t-q}$$

direkt als *Prognosegleichung* benutzt werden. Nicht-stationäre Reihen, deren Differenzen mit ARMA-Modellen analysiert wurden, müssen mit dem Summenfilter (vgl. S. 74) rücktransformiert werden, bevor die Prognose errechnet wird. Außerdem ist der Mittelwert μ hinzuzufügen, falls die x_t – wie in diesem Kapitel vorausgesetzt – als Abweichungen der Reihenwerte vom Mittelwert definiert sind.

Mit der Gleichung (85) läßt sich nun für jeden Zeitpunkt t ($t = 1, \dots, T$) eine *Ex-post-Prognose* \hat{x}_t erstellen und mit dem wahren Zeitreihenwert x_t vergleichen. Der Schätzfehler $e_t = (x_t - \hat{x}_t)$ ist zugleich eine Schätzung der Störvariablen u und wird als *Residuum* \hat{u}_t bezeichnet. Da wir für die Störvariable vorausgesetzt haben (vgl. S. 30 und 81)

$$E[u_t] = 0 \quad E[u_t u_{t+k}] = \begin{cases} \sigma_u^2 & \text{für } k = 0 \\ 0 & \text{für } k \neq 0 \end{cases}$$

kann man nun anhand des Verhaltens der Residuen \hat{u}_t prüfen, ob diese Annahme empirisch bestätigt wird und damit das ARMA-Modell *problemadäquat* ist. Diese Prüfung vollzieht sich in zwei Schritten.

Zunächst werden die *Autokorrelationskoeffizienten* der Residuen berechnet und graphisch dargestellt. Wir knüpfen an Abbildung 14 an und zeigen die Autokorrelationsfunktion der Residuen der Umlaufrendite für festverzinsliche Wertpapiere in der folgenden Abbildung.

($k = 1, \dots, K$) heran und weist die Zeitreihe T Werte auf, dann ist bei einem adäquaten ARIMA(p, d, q)-Modell die Größe

$$Q = n \sum_{k=1}^K r_k^2(\hat{u}) \quad (n = T - d)$$

annähernd χ^2 -verteilt mit $K - p - q$ Freiheitsgraden. Falls der Mittelwert μ im Modell mitgeschätzt wurde, verringert sich die Zahl der Freiheitsgrade um einen.

Für die Umlaufrendite festverzinslicher Wertpapiere ergibt sich bei $K = 26$ $T = 96$
 $d = 1$ $p = 1$ $q = 1$ (μ im Modell geschätzt)

$Q = 13,683$ mit 22 Freiheitsgraden.

Dieser Wert wird nun mit dem theoretischen Wert aus der χ^2 -Verteilung mit 22 Freiheitsgraden und 95 % Sicherheitswahrscheinlichkeit $\chi_{krit}^2(0,95) = 33,924$ verglichen. Danach kann die Hypothese: „Das Modell ist adäquat“ auf dem 95 %-Signifikanzniveau nicht abgelehnt werden. Mit diesem Test, der ebenfalls Bestandteil des in diesem Buch benutzten EDV-Programms ist, wollen wir die Prüfung des Prognosemodells abschließen. Im nächsten Abschnitt wird die Prognose des Aktienkursindex des Statistischen Bundesamtes mit dem Box-Jenkins-Verfahren durchgeführt.

7. Aktienkursprognose mit einem ARIMA-Modell

In der Abbildung 20 ist der Aktienkursindex des Statistischen Bundesamtes für den Zeitraum von Januar 1974 bis April 1982 dargestellt. Die Reihe besteht aus 100 Werten.

Die folgende Abbildung 16 zeigt die Autokorrelationsfunktion des Aktienindex.

Aus dem Verlauf des Index und aus der Autokorrelationsfunktion ist deutlich ablesbar, daß es sich um eine nicht-stationäre Reihe handelt, die durch Differenzenbildung zunächst in eine stationäre Reihe umgeformt werden muß. Dies geschieht mit Gleichung (73), $\Delta_1 x_t = x_t - \rho_1 x_{t-1}$, wobei $\rho_1 = 0,94$ gesetzt wird (Autokorrelationskoeffizient 1. Ordnung).

In den Abbildungen 17 und 18 sind die echten und partiellen Autokorrelationsfunktionen dieser Differenzenreihe wiedergegeben.

Der Verlauf der beiden Funktionen ist nicht ganz eindeutig, da sich signifikante Koeffizienten nicht nur bei einem time-lag von 1, sondern auch bei 6, 11 und 15 befinden. Zur Identifizierung des Modells benutzen wir die jeweils ersten signifikanten Koeffizienten und identifizieren ein gemischtes ARIMA(1,1,1)-Modell mit $p = q = 1$.

LAG	-1	0											+1	WERT					
1	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	.942
2874
3806
4740
5671
6608
7558
8515
9462
10411
11350
12274
13200
14136
15085
16060
17039
18014
19	-.008
20	-.012
21	-.019
22	-.026
23	-.021
24	-.005
25010

STANDARDABWEICHUNG-UNTERGRENZE : -.196
 STANDARDABWEICHUNG-OBERGRENZE : .196
 VERTRAUENSBEREICH : 95% (ANNAHME:NORMALVERTEILUNG)

Abbildung 16
 Autokorrelationsfunktion des Aktienindex

LAG	-1	0	+1	WERT
1	I	I	*	.240
2	.	*I	.	-.046
3	.	I*	.	.049
4	.	*	.	.018
5	.	*I	.	-.155
6	*	I	.	-.233
7	.	*I	.	-.104
8	.	*	.	-.022
9	.	*I	.	-.109
10	.	I*	.	.122
11	.	I	*	.232
12	.	I*	.	.026
13	.	*	.	-.006
14	.	*	.	-.006
15	*	I	.	-.237
16	.	*I	.	-.162
17	.	*	.	-.068
18	.	*	.	-.017
19	*	I	.	-.177
20	.	*I	.	-.089
21	.	I*	.	.063
22	.	*	.	-.003
23	.	*I	.	-.031
24	.	I*	.	.110
25	.	*	.	.011

STANDARDABWEICHUNG-UNTERGRENZE : -.197
STANDARDABWEICHUNG-OBERGRENZE : .197
VERTRAUENSBEREICH : 95% (ANNAHME:NORMALVERTEILUNG)

EMPIRISCHER CHI-QUADRAT-WERT = 37.366

Abbildung 17
Autokorrelationsfunktion der Aktienindex-Differenzenreihe

LAG	-1	0	+1	WERT
1	I.I.I.I.I.I.I.I.I.I.I.I.I.I.I.I.I.I.I	. I .*		.240
2		. * I .		-.110
3		. I * .		.094
4		. * I .		-.026
5		. * I .		-.154
6		. * I .		-.172
7		. * I .		-.034
8		. * .		-.006
9		. * I .		-.098
10		. I * .		.189
11		. I * .		.111
12		. * I .		-.085
13		. * .		-.002
14		. * I .		-.078
15	* .	. I .		-.292
16	. I*	. .		.028
17	. *019
18	. *011
19	. * I .	. .		-.165
20	. * I .	. .		-.039
21	. * I .	. .		-.137
22	. * I .	. .		-.137
23	. *009
24	. I*	. .		.050
25	. * I .	. .		-.067

STANDARDABWEICHUNG-UNTERGRENZE : -.197
STANDARDABWEICHUNG-OBERGRENZE : .197
VERTRAUENSBEREICH : 95% (ANNAHME:NORMALVERTEILUNG)

EMPIRISCHER CHI-QUADRAT-WERT = 36.049

Abbildung 18
Partielle Autokorrelationsfunktion der Aktienindex-Differenzenreihe

Die signifikant von Null verschiedenen echten und partiellen Autokorrelationskoeffizienten mit dem time-lag von 15 Perioden deuten auf einen schwachen Saisoneinfluß hin, den wir durch Differenzbildung nach Gleichung (75)

$$\Delta_2 x_t = \Delta_1 x_t - \rho_2 \Delta_1 x_{t-15}$$

mit dem Wert $\rho_2 = -0,24$ auszuschalten versuchen. Gemäß unserem ARIMA(1,1,1)-Modell ist also folgende Gleichung dem Schätzverfahren zugrunde zu legen:

$$(*) \quad \Delta_2 x_t = \mu + b_1 \Delta_2 x_{t-1} + u_t - a_1 u_{t-1}$$

Folgende Schätzwerte können nach einer Iteration mit dem Hauptschätzverfahren ermittelt werden:

$$\hat{\mu} = 9,0682 \quad \hat{b}_1 = -0,1912 \quad \hat{a}_1 = -0,4643$$

Für die Prognose ist die Schätzgleichung in die Normalform (85), die die absoluten Zeitreihenwerte enthält, zurückzutransformieren.

Aus (*) erhält man in Verbindung mit (75)

$$\Delta_1 x_t - \rho_2 \Delta_1 x_{t-15} = \hat{\mu} + \hat{b}_1 (\Delta_1 x_{t-1} - \rho_2 \Delta_1 x_{t-16}) + \hat{u}_t - \hat{a}_1 \hat{u}_{t-1}$$

und weiter mit (73)

$$x_t - \rho_1 x_{t-1} - \rho_2 (x_{t-15} - \rho_1 x_{t-16}) = \hat{\mu} + \hat{b}_1 (x_{t-1} - \rho_1 x_{t-2} - \rho_2 (x_{t-16} - \rho_1 x_{t-17})) + \hat{u}_t - \hat{a}_1 \hat{u}_{t-1}$$

Daraus folgt nach Ordnen die Prognosegleichung

$$(**) \quad \hat{x}_t = \hat{\mu} + (\rho_1 + \hat{b}_1) x_{t-1} - \rho_1 \hat{b}_1 x_{t-2} + \rho_2 x_{t-15} - \rho_2 (\rho_1 + \hat{b}_1) x_{t-16} + \rho_1 \rho_2 \hat{b}_1 x_{t-17} + \hat{u}_t - \hat{a}_1 \hat{u}_{t-1}$$

oder nach Einsetzen der Schätzwerte (incl. ρ_1 und ρ_2)

$$(***) \quad \hat{x}_t = 9,0682 + 0,749x_{t-1} + 0,18x_{t-2} - 0,240x_{t-15} + 0,180x_{t-16} + 0,043x_{t-17} + \hat{u}_t - 0,46\hat{u}_{t-1}$$

Da die Zeitreihe 100 Werte umfaßt, prognostizieren wir für die Periode $t = 101$, wobei wir beachten, daß das Residuum \hat{u}_{101} noch nicht bekannt ist und gemäß $E[\hat{u}_t] = 0$ gleich Null gesetzt wird, während $\hat{u}_{100} = x_{100} - \hat{x}_{100} = 0,742$ bereits vorliegt. Die benötigten Zeitreihenwerte lauten

$$x_{84} = 98,3 \quad x_{85} = 98,2 \quad x_{86} = 97,5 \quad x_{99} = 103,1 \quad x_{100} = 104,1$$

Somit ergibt sich ein Prognosewert von $\hat{x}_{101} = 104,8$.

Aus der Autokorrelationsfunktion der Residuen läßt sich ablesen, ob dieses ARIMA (1,1,1)-Modell problemangemessen ist.

LAG	-1	0	+1	WERT
1	I	*I		-.051
2		*		-.005
3		*		.023
4		*		.015
5	*	I		-.169
6	*	I		-.203
7		*I		-.123
8		*I		-.026
9	*	I		-.155
10		I *		.119
11		I *		.206
12		I*		.034
13		I*		.025
14		I*		.067
15		*		-.003
16		*I		-.094
17		*I		-.061
18		I*		.071
19	*	I		-.190
20		*I		-.098
21		I*		.063
22		*I		-.054
23		*I		-.034
24		I*		.067
25		*		.010

STANDARDABWEICHUNG-UNTERGRENZE : -.197
 STANDARDABWEICHUNG-OBERGRENZE : .197
 VERTRAUENSBEREICH : 95% (ANNAHME: NORMALVERTEILUNG)

EMPIRISCHER CHI-QUADRAT-WERT = 24.931
 THEORETISCHER CHI-QUADRAT-WERT = 32.671
 FREIHEITSGRADE = 21
 WAHRSCHEINLICHKEITSBEREICH = .950%

Abbildung 19
 Autokorrelationsfunktion der Residuen

Man erkennt, daß die Residuen im wesentlichen nicht signifikant von Null verschieden sind und auch der χ^2 -Test von Box und Pierce den Schluß zuläßt, daß das Modell adäquat ist ($Q = 24,9 < \chi^2(0,95) = 32,7$).

Trotz der Schwierigkeiten bei der Identifizierung des Modells zeigen auch die Prognose-Gütemaße

$$MAA = 1,80 \quad WMQA = 2,27 \quad U = 0,97$$

$$\text{mittlerer prozentualer Prognosefehler} = 1,77\%$$

daß ein einigermaßen befriedigendes Prognosemodell erarbeitet werden konnte. Der Theil'sche Ungleichheitskoeffizient ist jedoch nicht sehr gut und macht deutlich, daß die Prognosewerte den wahren Werten oft eine Periode hinterherlaufen. Diesen Eindruck vermittelt auch Abb. 20, die die Prognosewerte (Prognosehorizont: eine Periode) den wahren Werten gegenüberstellt.

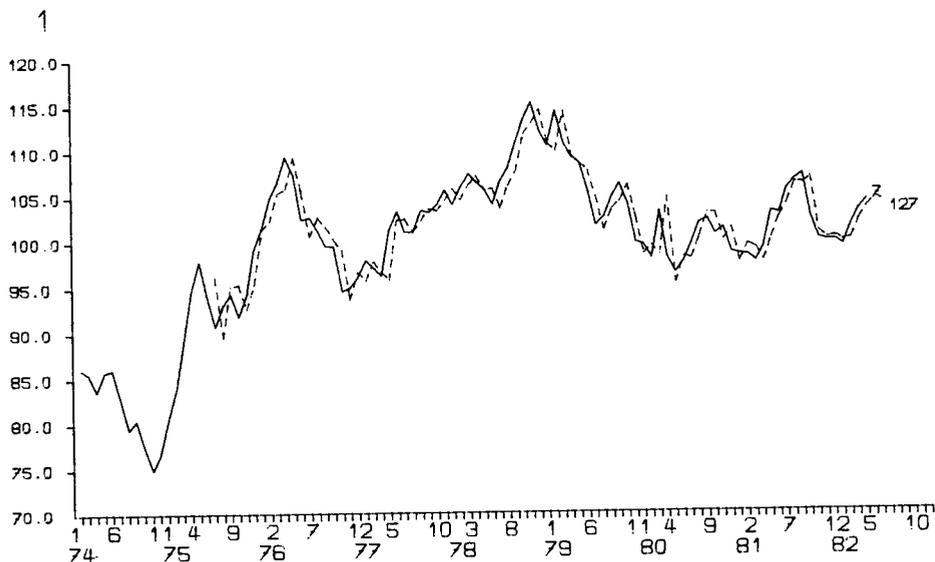


Abbildung 20
Aktienindex und Box-Jenkins-Prognose (gestrichelt)

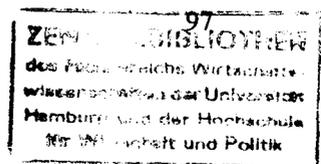
8. Kritik am Box-Jenkins-Verfahren

Das Box-Jenkins-Verfahren ist ohne Zweifel das anspruchvollste und detaillierteste univariate Prognoseverfahren, das heute in der betrieblichen Praxis angewendet wird. Seine Vorzüge können wie folgt beschrieben werden:

- Durch die konsequente Anwendung der Theorie *stationärer Prozesse* ist das Verfahren theoretisch gut fundiert. Die Erkenntnisse über die echte und partielle Autokorrelationsfunktion sind statistisch abgesichert und bilden die Grundlage eines wirksamen theoretischen Instrumentariums.
- Das allgemeine ARIMA-Modell ist *umfassend* angelegt, so daß stationäre und nicht-stationäre Zeitreihen, ggf. mit Saisonkomponente, behandelt werden können. Die exponentielle Glättung ist als Spezialfall im allgemeinen Modell enthalten.
- Die Anpassung des Modells an eine gegebene Zeitreihe ist sehr *flexibel* durch die Vielzahl möglicher Kombinationen der Parameter p , d und q . Darüber hinaus sind Prüfkriterien zur Identifikation des geeigneten Modells vorhanden.
- Für die Prognoseergebnisse sind statistisch ermittelte *Konfidenzintervalle* verfügbar, die eine objektive Beurteilung der Prognose ermöglichen.

Neben diesen unbestreitbaren Vorzügen weist das Verfahren natürlich auch einige Unzulänglichkeiten und Anwendungsprobleme auf, die im folgenden zusammengestellt sind.

- Das Verfahren ist *mathematisch* sehr anspruchsvoll und wird in der Praxis als außerordentlich kompliziert angesehen. Dies gilt sowohl für das Verständnis der Algorithmen des Anfangs- und des Hauptschätzverfahrens als auch für die EDV-technische Realisierung. Die (noch) relativ geringe Verbreitung in der praktischen Unternehmensprognose beruht hauptsächlich auf der Schwierigkeit der Benutzer, den *Prognose-Mechanismus* zu durchschauen.
- Der *Programmieraufwand* und die *Rechenzeiten* sind infolge der benutzten Schätzverfahren recht hoch. Die hierfür anfallenden Kosten stehen nicht immer im richtigen Verhältnis zur Verbesserung der Prognose gegenüber einfachen Verfahren.
- Das Verfahren erfordert vom Benutzer sehr viel *Erfahrung* und auch Fingerspitzengefühl, z. B. bei der Identifizierung des geeigneten Modells. Die Gefahr der Wahl eines falschen Modells auf Grund nicht eindeutiger Autokorrelationsfunktionen wird bei mechanistischer Anwendung des Verfahrens sehr groß.
- Der von Box und Jenkins verwendete *Differenzenfilter* ist für nicht-polynomiale Trends weniger geeignet. Eine Verbesserung der Ergebnisse konnte der Verfasser mit einer Modifikation des Verfahrens erzielen (vgl. S. 73 ff.).
- Das Verfahren ist am besten für *kurzfristige* Prognosen geeignet. Ist der Prognosezeitraum länger als eine Periode, werden die Prognosen schlechter, weil in der Prognosegleichung benötigte Werte noch nicht zur Verfügung stehen. Diese Verschlechterung läßt sich allerdings an der Verbreiterung der Konfidenzintervalle im voraus erkennen.
- Das Verfahren ist zunächst nur für die univariate Zeitreihenanalyse geeignet. Kausale Einflußfaktoren auf die Zeitreihenwerte können so nicht erfaßt werden. Schritte zur Erweiterung der Modellstruktur zur multivariaten Analyse, auf die wir hier nicht eingehen können, sind jedoch inzwischen eingeleitet worden (vgl. MARMA-Modelle, z. B. bei Makridakis/Wheelwright 1978a).



Das wichtigste Hindernis für eine weitere Verbreitung des Box-Jenkins-Verfahrens in der Praxis ist der mathematische Komplexitätsgrad. Um diesen herabzusetzen, haben Makridakis/Wheelwright (1978a) eine *vereinfachte Variante* entwickelt, die zusätzlich noch eine dynamische Anpassung eines Parameters analog zur exponentiellen Glättung mit α -Anpassung erlaubt. Dieses unter dem Namen adaptives Filtern bekannt gewordene Verfahren ist Gegenstand des nächsten Abschnitts.

III. Das Verfahren des adaptiven Filterns

1. Beziehungen zum Box-Jenkins-Verfahren

Dem Verfahren des adaptiven Filterns liegt dieselbe Prognosegleichung

$$(85) \quad x_t = b_1 x_{t-1} + b_2 x_{t-2} + \dots + b_p x_{t-p} + u_t - a_1 u_{t-1} - a_2 u_{t-2} - \dots - a_q u_{t-q}$$

zugrunde, die wir im vorigen Abschnitt benutzt haben. Darüber hinaus ist die Theorie der stationären Prozesse genau wie im Box-Jenkins-Verfahren theoretisches Fundament. Weitere wichtige Analogien ergeben sich bei der Identifizierung des geeigneten ARIMA-Modells mit Hilfe der echten und partiellen Autokorrelationsfunktionen und bei der Bestimmung der Anfangsschätzwerte für die Modellparameter. Bis hierher sind Box-Jenkins-Verfahren und adaptives Filtern praktisch identisch.

Der entscheidende Unterschied beider Ansätze liegt im *Hauptschätzverfahren* der Modellparameter und bezieht sich auf zwei Sachverhalte (Makridakis/Wheelwright 1978a, S. 286 ff.):

- Makridakis/Wheelwright passen die Modellparameter *dynamisch* an, d. h. zu jedem Zeitpunkt wird mit den aktuellen Zeitreihenwerten und den Schätzfehlern eine neue Anpassung vorgenommen.
- Sie verwenden eine *Lernkonstante* zur Steuerung der Schrittweite des Anpassungsprozesses.

Diese beiden Sachverhalte, die das adaptive Filtern kennzeichnen, werden im folgenden im Detail erläutert.

2. Die dynamischen Anpassungsgleichungen

Ausgangspunkt der Betrachtungen ist das Ziel, den mittleren quadrierten Schätzfehler

$$\frac{1}{T} \sum_{t=1}^T u_t^2$$

zu minimieren. Dieses Ziel ist adäquat zu

$$\sum_{t=1}^T u_t^2 \Rightarrow \text{Min!},$$

da die Division durch die Konstante T die Lage des Minimums im Raum der Parameter nicht verändert.

Mit den Anfangsschätzwerten $a_1^0, a_2^0, \dots, a_q^0, b_1^0, b_2^0, \dots, b_p^0$ haben wir einen Wert der Funktion

$$(92) \quad Z^0 = \sum_{t=1}^T u_t^2 = \sum_{t=1}^T (x_t - b_1^0 x_{t-1} - b_2^0 x_{t-2} - \dots - b_p^0 x_{t-p} + a_1^0 u_{t-1} + a_2^0 u_{t-2} + \dots + a_q^0 u_{t-q})^2$$

gewonnen, der aber nicht notwendigerweise das Minimum darstellt. Man bewegt sich von hier aus schnellstens in Richtung auf das Minimum, wenn man der Richtung des *negativen Gradienten* $-(\partial Z/\partial b_1, \dots, \partial Z/\partial b_p, \partial Z/\partial a_1, \dots, \partial Z/\partial a_q)$ folgt, da diese den steilsten Abstieg der Funktion anzeigt. Z ist nun allerdings eine Summenfunktion und damit hängt der Gradient von allen T Zeitreihenwerten ab, was eine umständliche Handhabung bedeutet.

Widrow (1966) hat deshalb eine Approximation des Gradienten vorgeschlagen, die nur die letzten p bzw. q Zeitreihenwerte benötigt und sich damit schneller an Veränderungen der Zeitreihe anpaßt.

Wir ersetzen dazu die Elemente des obigen Gradientenvektors durch die partiellen Ableitungen $\partial u_t^2 / \partial b_1, \partial u_t^2 / \partial b_2, \dots, \partial u_t^2 / \partial a_q$, was bedeutet, daß wir uns mit einem „durchschnittlichen“ Gradienten (über T Zeitpunkte) zufrieden geben und erhalten für den AR-Teil von (92)

$$(93) \quad \frac{\partial u_t^2}{\partial b_i} = 2 u_t \frac{\partial u_t}{\partial b_i} = 2 u_t (-x_{t-i}) \quad (i = 1, \dots, p)$$

und für den MA-Teil

$$(94) \quad \frac{\partial u_t^2}{\partial a_i} = 2 u_t \frac{\partial u_t}{\partial a_i} = 2 u_t u_{t-i} \quad (i = 1, \dots, q)$$

Diese partiellen Ableitungen hängen nicht mehr von der gesamten Zeitreihe, sondern nur noch von den letzten p bzw. q Werten ab.

Beachtet man das Minuszeichen des negativen Gradienten (Richtung des steilsten *Abstiegs*) und führt noch eine *Schrittweite* K ein, deren Bestimmung im nächsten Abschnitt behandelt wird, so ergeben sich folgende neuen Parameterwerte aus den Anfangsschätzungen

$$(95) \quad b_i = b_i^0 + 2 K u_t x_{t-i} \quad (i = 1, \dots, p) \quad \text{AR-Teil}$$

$$(96) \quad a_i = a_i^0 - 2 K u_t u_{t-i} \quad (i = 1, \dots, q) \quad \text{MA-Teil}$$

Bei geeigneter Wahl der Schrittweite K (vgl. nächsten Abschnitt) *konvergiert* der Iterationsprozeß, der die Parameter zu jedem Zeitpunkt t gemäß der folgenden Vorschrift anpaßt, in endlich vielen Schritten (Widrow 1966, S. 25):

$$(95a) \quad b_{it} = b_{i,t-1} + 2Ku_t x_{t-i} \quad (i = 1, \dots, p) \quad \text{AR-Teil}$$

$$(96a) \quad a_{it} = a_{i,t-1} - 2Ku_t u_{t-i} \quad (i = 1, \dots, q) \quad \text{MA-Teil}$$

Man hat also durch die Approximation des Gradienten mit (95a) und (96a) zwei leicht zu programmierende und auszuwertende Formeln gewonnen, die darüber hinaus den Vorzug besitzen, die Parameter in Abhängigkeit von K und dem aktuellen Schätzfehler u_t dynamisch anzupassen. Diese Gleichungen sind Bestandteil des am Lehrstuhl des Verfassers von Herrn Gunter Paetow entwickelten und implementierten EDV-Programms.

Die besten Werte von b_{iT} und a_{iT} erhält man mit der letzten Anpassung b_{iT} und a_{iT} am Ende der Zeitreihe. Diese Werte kann man nun benutzen, um am Anfang der Zeitreihe neu zu starten (2. Iteration). Im allgemeinen konvergieren die Parameterwerte bei geeigneter Wahl von K nach wenigen Iterationen.

3. Steuerung der Schrittweite durch eine Lernkonstante

Die hohe Flexibilität des Adaptiven Filterverfahrens wird maßgeblich durch die Lernkonstante K erreicht. Die Konvergenz des Verfahrens ist jedoch nicht bei jeder beliebig großen Schrittweite gewährleistet, sondern nur, wenn gilt (Widrow 1966, S. 25)

$$(97) \quad 0 < K < \frac{1}{\sum_{i=t-p}^{t-1} x_i^2 + \sum_{i=t-q}^{t-1} u_i^2}$$

Mit der Kenntnis der p letzten Zeitreihenwerte und der q letzten Schätzfehler kann somit die Obergrenze von K angegeben werden.

Standardisiert man alle Zeitreihenwerte und alle Schätzfehler, indem sie z. B. jeweils durch ihr Maximum dividiert werden⁷, so läßt sich eine „vorsichtige“ Obergrenze von K durch folgende Überlegung gewinnen:

Der größte standardisierte Wert ist gleich Eins, alle anderen liegen zwischen Null und Eins. Daher gilt

$$\sum_{i=t-p}^{t-1} x_i^2 \leq p \quad \text{und} \quad \sum_{i=t-q}^{t-1} u_i^2 \leq q$$

⁷ Evtl. sollten die Werte vorher in den positiven Bereich (\mathbb{R}^+) verschoben werden.

und wir erhalten die einfach auszuwertende Ungleichung

$$(98) \quad 0 < K < \frac{1}{p+q}$$

Innerhalb dieses Bereiches läßt sich K variieren, ohne die Konvergenz des Verfahrens zu gefährden. Der Anpassungsprozeß läuft im allgemeinen um so rascher ab, d. h. die Zahl der Iterationen ist um so kleiner, je näher K an der Obergrenze gewählt wird.

Als Anschauungsbeispiel für die dynamische Anpassung nehmen wir die künstliche Zeitreihe

$$x_1 = 128 \quad x_2 = 64 \quad x_3 = 32 \quad x_4 = 16 \quad x_5 = 8 \quad x_6 = 4 \quad x_7 = 2 \quad x_8 = 1$$

und behandeln sie mit einem AR(1)-Modell

$$x_t = b_1 x_{t-1} + u_t$$

Man sieht sofort, daß b_1 den Wert 0,5 erhalten müßte. Wir wählen $b_1^0 = 1,0$ und $K = 0,1$, um den Anpassungsprozeß zu verfolgen. Tabelle 12 zeigt, daß nach zwei Iterationen b_1 gegen den Wert 0,5 konvergiert, so daß das Verfahren abgebrochen werden kann.

Für die dynamische Anpassung der Parameter sollte nicht direkt auf Gleichung (95a) zurückgegriffen werden. Eine Normalisierung des Satzes der Zeitreihenwerte (x_{t-i} ; $i = 1, \dots, p$) und des Schätzfehlers (u_t) führt in der Anpassungsphase zu einer schnelleren Konvergenz der Parameter gegen den Satz der „optimalen“ Parameter.

Die normalisierten Zeitreihenwerte (x_{t-i}^* ; $i = 1, \dots, p$) und den normalisierten Schätzfehler (u_t^*) erhalten wir, indem wir die Zeitreihenwerte und den Schätzfehler durch

$$NK = \sqrt{\sum_{i=1}^p x_{t-i}^2}$$

teilen.

Die Verwendung von NK als Normalisierungskonstante führt zu einer um 5 % schnelleren Anpassung der Parameter an die „optimalen“ Parameter als die Verwendung des größten Zeitreihenwertes aus jedem Satz von p Zeitreihenwerten als Normalisierungskonstante (vgl. Makridakis/Reschke/Wheelwright 1980, S. 94 f.).

Tabelle 12

Iteration	t	b_{1t-1}	x_t	\hat{x}_t	x_{t-1}^*	u_t	u_t^*
1	1	—	128,00	—	—	—	—
	2	1,00	64,00	128,00	1,00	-64,00	-0,50
	3	0,90	32,00	57,60	1,00	-25,60	-0,40
	4	0,82	16,00	26,24	1,00	-10,24	-0,32
	5	0,76	8,00	12,16	1,00	- 4,16	-0,26
	6	0,71	4,00	5,68	1,00	- 1,68	-0,21
	7	0,67	2,00	2,68	1,00	- 0,68	-0,17
	8	0,64	1,00	1,28	1,00	- 0,28	-0,14
2	1	—	128,00	—	—	—	—
	2	0,61	64,00	78,08	1,00	-14,08	-0,11
	3	0,59	32,00	37,76	1,00	- 5,76	-0,09
	4	0,57	16,00	18,24	1,00	- 2,24	-0,07
	5	0,56	8,00	8,96	1,00	- 0,96	-0,06
	6	0,55	4,00	4,40	1,00	- 0,40	-0,05
	7	0,54	2,00	2,16	1,00	- 0,16	-0,04
	8	0,53	1,00	1,06	1,00	- 0,06	-0,03

4. Aktienkursprognose durch adaptives Filtern

Bei der Analyse des Box-Jenkins-Verfahrens hatten wir die Zeitreihe des Aktienindex als ARIMA (1,1,1)-Modell identifiziert. Wir wollen an dieser Stelle zeigen, daß der MA-Teil durch einen AR-Teil höherer Ordnung (vgl. auch dazu Gleichung (62)) approximiert werden kann. Nach vielen Experimenten mit verschiedenen AR-Modellen erwies sich ein AR(12)-Modell ohne Differenzenbildung als adäquat, das nunmehr vorgestellt werden soll.

Mit $p = 12$ ergibt sich für K die obere Schranke $\frac{1}{12} = 0,08\bar{3}$. Eine bessere Anpassung brachte jedoch der Wert $K = 0,005$. Nach 10 Iterationen betrug die mittlere quadrierte Abweichung $MQA = 6,07$ und die Wurzel aus $MQA = 2,46$, was einem prozentualen mittleren Fehler von 2,5 % entspricht.

Aufgrund der geschätzten Parameter konnte folgende Prognosegleichung für das AR(12)-Modell aufgestellt werden:

$$\begin{aligned}\hat{x}_t = & 1,107x_{t-1} - 0,158x_{t-2} + 0,004x_{t-3} + 0,045x_{t-4} - 0,019x_{t-5} \\ & - 0,063x_{t-6} + 0,0003x_{t-7} + 0,163x_{t-8} - 0,14x_{t-9} + 0,146x_{t-10} \\ & + 0,039x_{t-11} - 0,122x_{t-12}\end{aligned}$$

Im Rahmen der Box-Jenkins-Theorie ist ein Wert von $b_1 > 1$ nicht optimal, doch ist dies in unserem Fall darauf zurückzuführen, daß die nicht-stationäre Zeitreihe selbst verwendet wird, ohne eine Differenzenreihe zu bilden. In dem Koeffizienten $b_1 =$

1,107 kommt daher zusätzlich der Trend der Zeitreihe zum Ausdruck. Gerade in einem solchen Fall der nicht-adäquaten Modellspezifizierung (Nicht-Stationarität, Vernachlässigung des MA-Teils) zeigt sich die *Robustheit* des Verfahrens, das trotzdem gute Prognoseergebnisse liefert.

Weitere Einzelheiten zum adaptiven Filtern mit einer umfangreichen Fallstudie findet man in (Hüttner 1982, S. 142 ff.).

5. Kritische Zusammenfassung

Das adaptive Filtern zeichnet sich durch vier Hauptvorzüge aus:

- Selbstanpassung der Modellparameter in einem dynamischen Prozeß
- Einfache Herleitung und Handhabung der Anpassungsgleichungen (95a) und (96a)
- Universelle Anwendbarkeit auf alle ARIMA-Modelle
- Robustheit bei falscher Modellspezifikation (vgl. Abschnitt 4.)

Diese Vorzüge haben dem Verfahren im angloamerikanischen Bereich eine gewisse Verbreitung in der Praxis verschafft und damit dem Box-Jenkins-Verfahren eine ernstzunehmende Konkurrenzmethode an die Seite gestellt, die die Vorteile der exponentiellen Glättung (Einfachheit, Selbstanpassung) mit der Flexibilität von Box-Jenkins gut kombiniert.

Hauptnachteil des Verfahrens ist der fehlende statistisch-theoretische Unterbau, da die statistischen Eigenschaften des Filters noch weitgehend ungeklärt sind und keine Konfidenzintervalle für die Prognosen abgeleitet werden können. Darüber hinaus ist die *optimale* Bestimmung der Lernkonstanten K ein Problem, das noch nicht gelöst werden konnte, dessen Lösung aber dem Verfahren in Zukunft noch größere Anwendungsmöglichkeiten erschließen könnte.

F. Prognosen mit Wachstums- und Sättigungsmodellen

I. Langfristige Trendextrapolation

In den zurückliegenden Kapiteln haben wir uns überwiegend mit Prognoseverfahren befaßt, die kurz- bis höchstens mittelfristige Prognosen erlauben, weil sie auf die Zeitreihenwerte und die Schätzfehler der Vergangenheit zurückgreifen. Umfassen die Prognosen einen Zeitraum, der länger als eine Periode ist, so fehlen z. B. in der Prognosegleichung

$$x_{t+2} = b_1 x_{t+1} + b_2 x_t + b_3 x_{t-1} + u_{t+2} - a_1 u_{t+1} - a_2 u_t - a_3 u_{t-1}$$

der Zeitreihenwert x_{t+1} sowie der Schätzfehler u_{t+1} , da sie zum Zeitpunkt der Prognose noch nicht bekannt sind. Solche mehrperiodigen Prognosen werden also mit zunehmendem Zeithorizont immer unsicherer und verlieren bei etwa 4 bis 6 Prognoseperioden jede praktische Bedeutung.

Will man *langfristige* Prognosen abgeben, so ist man gezwungen, das ehrgeizige Ziel einer möglichst genauen Voraussage der zukünftigen Zeitreihenstruktur aufzugeben und sich statt dessen mit der Prognose einer „allgemeinen Richtung“ der Zeitreihe, d. h. ihres *Trends*, ungeachtet etwaiger kurzfristiger Schwankungen, zu bescheiden. Mit den Verfahren, die dieser Aufgabe der „Trendprognose“ gewidmet sind, beschäftigen wir uns in diesem Kapitel.

Neben der einfachen Trendextrapolation mit polynomialen Funktionen werden die exponentielle Wachstumsfunktion sowie zwei Ansätze behandelt, die ein Sättigungsniveau in die Analyse einbeziehen und damit besonders zur Prognose der Verbreitung langlebiger Konsumgüter in der Bevölkerung (z. B. Pkw-Bestand) herangezogen werden.

Die einfache *Trendextrapolation* paßt an die Vergangenheitswerte der Zeitreihe ein Polynom (möglichst) niedrigen Grades an, das die Summe der quadrierten Abweichungen zwischen den wahren Zeitreihenwerten und den Trendwerten minimiert.

Das so gewonnene Trendpolynom kann dann beliebig weit in die Zukunft extrapoliert werden, solange die Zeitstabilitätshypothese gilt.

Vor der quantitativen Darstellung der Trendextrapolation ist noch zu klären, warum ein niedriger Polynomgrad anzustreben ist. Es läßt sich zwar durch jede Zeitreihe von T Werten ein Polynom $(T-1)$ -ten Grades legen, das alle Zeitreihenwerte enthält, doch ist damit für die *Prognose* überhaupt nichts gewonnen, da die Extrapolation eines Polynoms von höherem als zweitem Grade durch die starke Steigung der Funktion schnell zu sinnlosen Entwicklungen führt. Dies sei an folgendem Beispiel verdeutlicht.

Eine Zeitreihe bestehe aus den vier Beobachtungswerten

$$x_1 = 0 \quad x_2 = 2 \quad x_3 = 1 \quad x_4 = 3$$

Wir können hierfür ein Polynom 3. Grades in der Zeitvariablen t finden, das diese Werte reproduziert, d. h. sich „optimal“ an die Zeitreihe anpaßt.

$$x_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3$$

Die Parameter a_0, a_1, a_2, a_3 bestimmt man durch Einsetzen der Zeitreihenwerte und der Zeitvariablen t in die obige Gleichung:

$$0 = a_0 + a_1 + a_2 + a_3$$

$$2 = a_0 + 2a_1 + 4a_2 + 8a_3$$

$$1 = a_0 + 3a_1 + 9a_2 + 27a_3$$

$$3 = a_0 + 4a_1 + 16a_2 + 64a_3$$

Aus dem so entstehenden linearen Gleichungssystem von vier Gleichungen ergeben sich die „optimal“ angepaßten Parameterwerte, so daß der Ansatz für diese Zeitreihe lautet:

$$x_t = -11 + \frac{35}{2}t - \frac{15}{2}t^2 + t^3 + u_t$$

Obwohl bei dieser Anpassung keine Abweichungen der Zeitreihenwerte von den errechneten Trendwerten auftreten, ist dieses Polynom für die Prognose *ungeeignet*, wie die extrapolierten Werte x_5, x_6, x_7 in der folgenden Tabelle zeigen:

Tabelle 13

Trendextrapolation mit einem Polynom dritten Grades

t	x_t	\hat{x}_t	\hat{u}_t
1	0	0	0
2	2	2	0
3	1	1	0
4	3	3	0
5	—	14	—
6	—	40	—
7	—	87	—

Die extrapolierten Werte entsprechen sicherlich in keiner Weise der Struktur der Zeitreihe x_1 bis x_4 . Der steile Anstieg des Polynoms ist für Prognosezwecke ungeeignet.

Aus diesem Grund versucht man, mit linearen oder höchstens quadratischen Funktionen auszukommen. Ein Ausschalten der Zufallskomponente u_t „mit Gewalt“, wie dies durch die obige Anpassung geschah, ist sicher für die *stochastischen* Prozesse der Realität unangemessen.

Die Annahme eines linearen Trends sollte immer am Anfang einer Trendanalyse stehen. Man unterstellt in diesem Fall eine Prognosegleichung der Form

$$(99) \quad x_t = a_0 + a_1 t + u_t$$

und schätzt die Parameter a_0 und a_1 mit dem Ziel, die Summe der Abweichungsquadrate

$$\sum_{t=1}^T u_t^2$$

während der Beobachtungsperiode $t = 1, \dots, T$ zu minimieren. Hier tritt die Ähnlichkeit zum Trendmodell der exponentiellen Glättung zutage (vgl. S. 33 ff.), bei dem wir die Summe der *gewogenen* Abweichungsquadrate minimiert haben.

Wir minimieren also folgende Zielfunktion (Kleinste-Quadrate-Methode)

$$(100) \quad Z = \sum_{t=1}^T u_t^2 = \sum_{t=1}^T (x_t - a_0 - a_1 t)^2 \Rightarrow \text{Min!}$$

indem wir die ersten partiellen Ableitungen nach a_0 und a_1 gleich Null setzen.

$$(101a) \quad \frac{\partial Z}{\partial a_0} = -2 \sum_{t=1}^T (x_t - a_0 - a_1 t) = 0$$

$$(101b) \quad \frac{\partial Z}{\partial a_1} = -2 \sum_{t=1}^T (x_t - a_0 - a_1 t) \cdot t = 0$$

Für nicht-negative quadratische Funktionen sind auch die hinreichenden Bedingungen für ein Minimum erfüllt, so daß wir mit (101) a_0 und a_1 bestimmen können und damit die sog. *Normalgleichungen* erhalten

$$(102a) \quad \hat{a}_0 T + \hat{a}_1 \sum_{t=1}^T t = \sum_{t=1}^T x_t$$

$$(102b) \quad \hat{a}_0 \sum_{t=1}^T t + \hat{a}_1 \sum_{t=1}^T t^2 = \sum_{t=1}^T x_t \cdot t$$

Für unser Beispiel in Tabelle 13 ergeben sich folgende Ausdrücke:

$$\sum_{t=1}^4 t = 10 \quad \sum_{t=1}^4 x_t = 6 \quad \sum_{t=1}^4 t^2 = 30 \quad \sum_{t=1}^4 x_t \cdot t = 19$$

Die Prognosegleichung wird aufgestellt durch Auflösen von (102) nach a_0 und a_1 sowie Einsetzen in (99)

$$(103) \quad x_t = -\frac{1}{2} + \frac{4}{5}t + u_t \quad (t = 1, 2, \dots, 7)$$

Tabelle 14 enthält die Anpassungs- und Prognosewerte.

Tabelle 14

Trendextrapolation mit linearer Funktion

t	x_t	\hat{x}_t	\hat{u}_t	\hat{u}_t^2
1	0	0,3	-0,3	0,09
2	2	1,1	0,9	0,81
3	1	1,9	-0,9	0,81
4	3	2,7	0,3	0,09
5	-	3,5	-	-
6	-	4,3	-	-
7	-	5,1	-	-

Obwohl die Anpassung mit $\sum \hat{u}_t^2 = 1,8$ schlechter als mit dem Polynom dritten Grades ausfällt, scheint sich die Prognose doch besser im Rahmen des zu erwartenden Trends zu bewegen.

Stellt man bei der Betrachtung des Graphen einer konkreten Zeitreihe eine tatsächlich *nicht-lineare* Struktur fest, so sollte man zunächst prüfen, ob eine quadratische Funktion problemangemessen ist. Dies geschieht durch Minimierung der Zielfunktion

$$(104) \quad Z = \sum_{t=1}^T u_t^2 = \sum_{t=1}^T (x_t - a_0 - a_1 t - a_2 t^2)^2 \Rightarrow \text{Min!}$$

und Bestimmung der Parameter \hat{a}_0 , \hat{a}_1 und \hat{a}_2 in genau derselben Weise wie oben beschrieben. Ist die quadratische Trendanpassung unbefriedigend, so lassen sich vielleicht nicht-polynomiale Trendfunktionen anwenden, die in den nächsten Abschnitten dieses Kapitels besprochen werden.

II. Die Exponentialfunktion

1. Prognosegleichung und Parameterschätzung

Eine Reihe von ökonomischen Prozessen zeigt einen Verlauf, der durch eine nur in gewissen Grenzen schwankende, im Trend aber weitgehend *konstante Wachstumsrate* gekennzeichnet ist. Die Zeitreihe entwickelt sich dann in exponentieller Weise, wie unten im einzelnen gezeigt wird. Beispiele für solche Prozesse sind Bruttosozialprodukt und Volkseinkommen sowie die davon abhängigen Zeitreihen, aber auch zuweilen das Bevölkerungswachstum und Kapitalakkumulationsprozesse.

Fundament der exponentiellen Wachstumsfunktion ist folgende Annahme:

Das Wachstum der Zeitreihe $\{x_t\}$, d. h. ihre zeitliche Veränderung Δx_t ist *proportional* zum bisher erreichten Niveau x_t .

Diese Annahme, deren Zutreffen natürlich für jede konkrete Zeitreihe geprüft werden muß, kleiden wir nun in eine Gleichung, wobei wir aus mathematischen Gründen (Anwendung der Infinitesimalrechnung) statt der diskreten Zeitreihe $\{x_t\}$ einen kontinuierlichen Prozeß $x(t)$ zugrunde legen.

$$(105) \quad x'(t) = \frac{dx}{dt} = ax(t) \quad \text{oder} \quad (105a) \quad \frac{x'(t)}{x(t)} = a$$

x' gibt dabei die marginale zeitliche Veränderung von x an, und x'/x die relative Veränderung, d. h. die *Wachstumsrate*, die gemäß Annahme konstant sein soll.

(105) ist eine *Differentialgleichung*, die gelöst werden muß, um die Funktion $x(t)$ explizit anschreiben zu können. Zu diesem Zweck formen wir (105) um in

$$(106) \quad \frac{1}{x} dx = a \cdot dt \quad \text{und integrieren}$$

$$(107) \quad \int \frac{1}{x} dx = \int a dt + \ln C$$

wobei $\ln C$ die Integrationskonstante bedeutet. Die Integration liefert die Funktion

$$(108) \quad \ln x = at + \ln C$$

die nach Entlogarithmieren in die *Exponentialfunktion* übergeht.

$$(109) \quad x(t) = Ce^{at} \quad (e = 2,71828 \dots ; t \in \mathbb{R})$$

Will man (109) als Prognosegleichung benutzen, so müssen die Parameter C und a aus der Zeitreihe geschätzt werden. Dafür steht uns das Verfahren der kleinsten Quadrate

(KQ) aus dem letzten Abschnitt (S. 106) zur Verfügung. Wir können es jedoch nicht direkt auf (109) anwenden, da die Schätzgleichung *linear* in den Parametern sein muß. Gleichung (108) ist jedoch dafür geeignet und wir erhalten für eine diskrete Zeitreihe den Ansatz

$$(108a) \quad \ln x_t = \ln C + at + u_t \quad (t = 1, \dots, T)$$

und die Zielfunktion

$$(110) \quad Z = \sum_{t=1}^T u_t^2 = \sum_{t=1}^T (\ln x_t - \ln C - at)^2 \Rightarrow \text{Min!}$$

In genau gleicher Weise wie in Abschnitt I ergeben sich nun die geschätzten Parameter $\ln \hat{C}$ und \hat{a} , wenn man in den Gleichungen (102a) und (102b) $\hat{a}_0 = \ln \hat{C}$ und $\hat{a}_1 = \hat{a}$ setzt. Nach Enthlogarithmieren der Schätzgleichung erhalten wir die endgültige Exponentialfunktion

$$(111) \quad \hat{x}_t = \hat{C}e^{\hat{a}t + \hat{u}_t} \quad (t = 1, \dots, T)$$

mit der *Prognosegleichung*

$$(112) \quad \hat{x}_t = \hat{C}e^{\hat{a}t} \quad (t = T + 1, T + 2, \dots)$$

Zur Illustration sei nun ein Beispiel für das exponentielle Modell durchgerechnet.

2. Anwendungsbeispiel: Prognose der Weltbevölkerung

Die seit 1920 in 10-Jahres-Abständen vorliegenden Werte der Zeitreihe Entwicklung der Weltbevölkerung (Quelle: G. Bruckmann (1978), S. 51) sollen durch eine exponentielle Funktion reproduziert werden. Die Zeitreihenwerte sind in Tabelle 15 dargestellt.

Tabelle 15

Entwicklung der Weltbevölkerung in Mrd.

Jahr	t	x_t	Jahr	t	x_t
1920	1	1,794	1960	5	2,982
1930	2	1,996	1970	6	3,632
1940	3	2,228	1980	7	4,330
1950	4	2,486	1990	8	—

Nach Anwendung der KQ-Methode auf Gleichung (108a) ergeben sich für die Parameter Werte von

$$\hat{C} = 1,4713$$

und $\hat{a} = 0,1476$,

so daß die Prognosegleichung (112) für unser Beispiel lautet:

$$\hat{x}_t = 1,4713 \cdot e^{0,1476 \cdot t}$$

Die Zeitreihen- und Prognosewerte sind in Tabelle 16 wertmäßig und in Abbildung 21 graphisch dargestellt.

Tabelle 16

Jahr	t	x_t	\hat{x}_t
1920	1	1,794	1,705
1930	2	1,996	1,977
1940	3	2,228	2,291
1950	4	2,486	2,655
1960	5	2,982	3,078
1970	6	3,632	3,567
1980	7	4,330	4,134
1990	8	—	4,792
	MAA = 0,1	WMQA = 0,1	U = 0,2

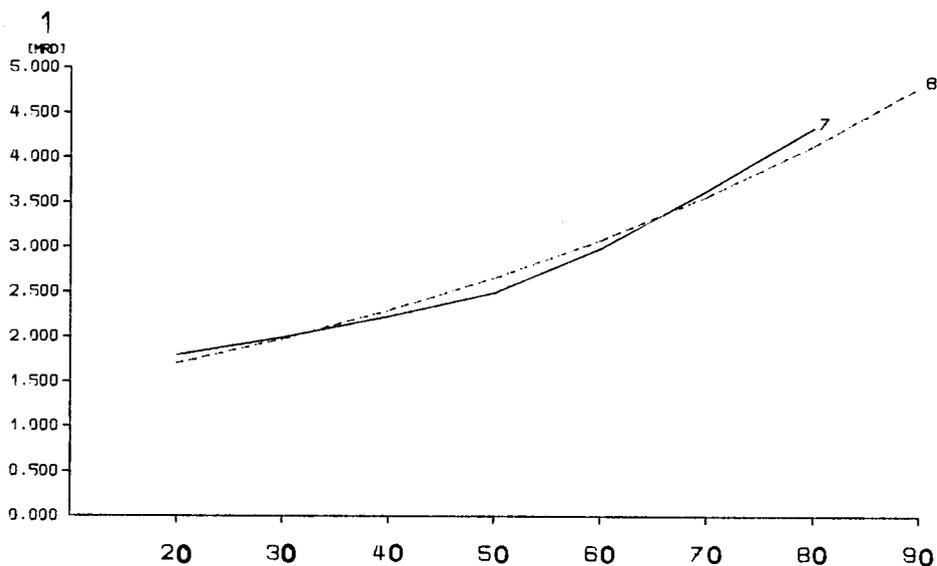


Abbildung 21
Entwicklung der Weltbevölkerung (Prognose gestrichelt)

Die gute Übereinstimmung der Ist- und Prognosewerte wird durch die Gütemaße sowie durch die Abbildung der Zeitreihen bestätigt.

III. Die logistische Funktion

Die logistische Funktion und ihre Varianten sind in zahlreichen Veröffentlichungen der Prognoseliteratur beschrieben worden (vgl. vor allem Lewandowski 1974, S. 265 ff., Lewandowski 1980, S. 156 ff. und Mertens 1981, S. 192 ff.). Hieraus geht schon ihre Bedeutung für die betriebliche Praxis hervor. So geht z. B. der Absatz vieler Produkte nach einer Wachstumsphase in eine ruhigere Entwicklung über, die schließlich in der Stagnation endet, weil der Markt mit diesem Produkt *gesättigt* ist.

Beispiele für ein solches Wachstum mit Tendenz zur Sättigung lassen sich auf dem Markt für langlebige Konsumgüter leicht finden. Es ist für die Produzenten außerordentlich wichtig,

- den Trend der Absatzentwicklung und
- das vermutliche Sättigungsniveau

zuverlässig zu prognostizieren, um sich rechtzeitig auf andere, wachstumsträchtigere Produkte umstellen zu können. Wir wollen dieses Problem nun quantitativ zu lösen versuchen.

1. Ableitung der logistischen Anpassungsgleichung

Auch die logistische Funktion beruht wie die exponentielle auf einer *Fundamentalannahme*, die wie folgt formuliert werden kann:

Das Wachstum $\frac{dx}{dt}$ der Zeitreihe $x(t)$ ist *proportional*

- a) dem jeweils erreichten Niveau $x(t)$ und
- b) dem Abstand zwischen dem erreichten Niveau $x(t)$ und dem absoluten Sättigungsniveau S .

Es gibt also eine fördernde Wachstumskomponente ($x(t)$) wie im exponentiellen Modell, darüber hinaus aber auch eine hemmende Komponente, da der Abstand ($S - x(t)$) mit zunehmendem Niveau x abnimmt.

Die Fundamentalannahme lautet quantitativ

$$(113) \quad x' = \frac{dx}{dt} = ax(S - x)$$

Daraus folgt eine abnehmende Wachstumsrate der Zeitreihe

$$(114) \quad \frac{x'}{x} = a(S - x)$$

Die Differentialgleichung (113) kann nicht ohne weitere Umformungen integriert werden, so daß wir die Ableitung kurz nachvollziehen wollen. Wir gehen von (113) aus und setzen aus mathematischen Gründen $x \neq 0$ und $x \neq S$ voraus.

$$(113a) \quad dx = ax(S - x)dt \quad \text{bzw.} \quad \frac{1}{x(S-x)} dx = a dt$$

Dies wird umgeformt zu

$$\frac{S - x + x}{Sx(S - x)} dx = \frac{S - x}{Sx(S - x)} dx + \frac{x}{Sx(S - x)} dx = a dt$$

Hieraus ergibt sich

$$\frac{1}{x} dx + \frac{1}{S - x} dx = aS dt$$

Setzen wir nun $z = S - x$ und damit $dz/dx = -1$, so können wir die Integration

$$\int \frac{1}{x} dx + \int -\frac{1}{z} dz = aS \int dt + C$$

ausführen und erhalten (C ist die Integrationskonstante)

$$\ln x - \ln z = aS t + C$$

Hier ist z wieder durch $S - x$ zu ersetzen und die Logarithmenregel II anzuwenden ($\ln a - \ln b = \ln(\frac{a}{b})$)

$$\ln\left(\frac{x}{S - x}\right) = aS t + C$$

Entlogarithmieren ergibt $x/(S - x) = e^{aS t + C}$ und schließlich durch Auflösen nach x die Gleichung der

logistischen Funktion

$$(115) \quad x(t) = \frac{S}{1 + e^{-aS t - C}}$$

Auf den ersten Blick müssen diesmal drei Parameter geschätzt werden: a , S und C . Dies trifft jedoch nur scheinbar zu, da wir C durch S schätzen können. Formel (115) lautet nämlich für $t = 1$

$$x(1) = \frac{S}{1 + e^{-aS-C}} \quad \text{und daraus ergibt sich}$$

$$e^{-C} = \frac{S - x(1)}{x(1) e^{-aS}} \quad (x(1) \neq 0)$$

Wir haben also nur die Parameter S und a zu schätzen.

2. Schätzung der Parameter

Gleichung (115) ist zur Schätzung von S und a nicht unmittelbar geeignet, weil die Parameter nicht linear verknüpft sind. Wir transformieren daher die Zeitreihe $\{x_t\}$ nach folgender Vorschrift und führen dabei neue Variable y_t und z_t ein:

$$y_t = \frac{1}{x_t} \quad z_t = \frac{1}{x_{t+1}}$$

Eingesetzt in (115) führt dies bei kontinuierlicher Betrachtung zu

$$y = \frac{1 + e^{-aS t - C}}{S} \quad \text{und}$$

$$z = \frac{1 + e^{-aS(t+1) - C}}{S} = \frac{(1 + e^{-aS t - C}) e^{-aS} + 1 - e^{-aS}}{S}$$

Hiermit gewinnt man eine lineare Funktion in y und z

$$z = e^{-aS} \cdot y + \frac{1 - e^{-aS}}{S}$$

oder, wenn $e^{-aS} = b_1$ und $(1 - e^{-aS})/S = b_0$ gesetzt wird

$$(116) \quad z_t = b_0 + b_1 y_t + u_t \quad (t = 1, \dots, T)$$

deren Parameter b_0 und b_1 mit der Methode der Kleinsten Quadrate geschätzt werden können.

Aus den Schätzwerten \hat{b}_0 und \hat{b}_1 errechnen wir nun \hat{S} und \hat{a} :

$$\hat{S} \cdot \hat{b}_0 = 1 - e^{-\hat{a}\hat{S}} = 1 - \hat{b}_1 \quad \text{Daraus folgt}$$

$$(117) \quad \hat{S} = \frac{1 - \hat{b}_1}{\hat{b}_0}$$

und aus $-\hat{a}\hat{S} = \ln \hat{b}_1$ ergibt sich

$$(118) \quad \hat{a} = \frac{\ln \hat{b}_1 \cdot \hat{b}_0}{1 - \hat{b}_1}$$

Damit haben wir formal das Schätzproblem gelöst. An empirischen Beispielen zeigt sich jedoch, daß insbesondere die Schätzung des Sättigungsniveaus S sehr unsicher ist und die Prognose sehr beeinträchtigen kann.

Es sollte grundsätzlich angestrebt werden, das Sättigungsniveau nicht aus der Zeitreihe selbst zu schätzen, sondern exogen, d. h. aus externen Quellen und zusätzlichen Überlegungen zu bestimmen. So versucht man z. B. das Sättigungsniveau des Pkw-Bestandes in der BRD aus der Bevölkerungsentwicklung abzuleiten. Wenn eine solche externe Bestimmung von S möglich ist, vereinfacht sich das Schätzproblem wie folgt:

Aus (115) ergibt sich $S/x = 1 + e^{-aS t - C}$ und durch Logarithmieren

$$(119) \quad \ln \left(\frac{S}{x} - 1 \right) = -C - a S t$$

Auf diese Gleichung kann man direkt die KQ-Methode anwenden und bei bekanntem S den Parameter a schätzen.

Häufig haben wir jedoch – insbesondere bei jungen Produkten – keine exogenen Vorstellungen über das Sättigungsniveau und müssen dann die ausführliche Schätzung von a und S mit (116) vornehmen.

3. Anwendungsbeispiel: Prognose der Ausstattung deutscher Privathaushalte mit Farbfernsehern

Seit 1968 liegt eine Zeitreihe vor, die angibt, welcher Anteil der deutschen Privathaushalte ein Farbfernsehgerät besitzt (Quelle: Statistische Jahrbücher, 20.7 Ausstattung ausgewählter privater Haushalte mit ausgewählten langlebigen Gebrauchsgütern, Haushaltstyp 2). Die Entwicklung dieser Zeitreihe (vgl. Abbildung 22) folgt einer logistischen Funktion. Wir bestimmen das Sättigungsniveau exogen, indem wir annehmen, daß maximal 90 % der Haushalte ein Farbfernsehgerät haben werden und die restlichen 10 % entweder kein Fernsehen mögen oder nicht auf Farbgeräte umsteigen wollen.

Die in Tabelle 17 dargestellte Zeitreihe gibt den Anteil deutscher Privathaushalte mit Farbfernsehgeräten wieder.

Tabelle 17

Anteil der Privathaushalte mit Farbfernsehgerät (in %) an allen Privathaushalten

Jahr	t	x_t	Jahr	t	x_t
1968	1	2,0	1975	8	29,3
1969	2	3,4	1976	9	42,2
1970	3	3,5	1977	10	50,1
1971	4	3,8	1978	11	60,9
1972	5	9,1	1979	12	69,2
1973	6	10,9	1980	13	73,8
1974	7	21,1			

Tabelle 18

Prognose des Anteils der Farbfernsehgeräte bis 1985

Jahr	t	x_t	\hat{x}_t
1968	1	2,000	1,571
1969	2	3,400	2,485
1970	3	3,500	3,904
1971	4	3,800	6,078
1972	5	9,100	9,331
1973	6	10,900	14,033
1974	7	21,100	20,503
1975	8	29,300	28,834
1976	9	42,200	38,644
1977	10	50,100	49,123
1978	11	60,900	59,171
1979	12	69,200	67,861
1980	13	73,800	74,734
1981	14	—	79,794
1982	15	—	83,327
1983	16	—	85,702
1984	17	—	87,260
1985	18	—	88,265
	MAA = 1,3	WMQA = 1,7	U = 0,2

Durch die Vorgabe des Sättigungsniveaus können wir die Parameter a und C nach Gleichung (119) schätzen. Es ergeben sich Werte von

$$\hat{a} = 0,0052 \quad \text{und}$$

$$\hat{C} = -4,4979.$$

Die Prognosegleichung (115) für unser Beispiel lautet

$$\hat{x}_t = \frac{90}{1 + e^{-0,468 \cdot t - (-4,4979)}}$$

Prognose- und Zeitreihenwerte sind in Tabelle 18 wiedergegeben.

Die Gütemaße zeigen, daß sich mit Hilfe der logistischen Funktion das Verhalten der Zeitreihe relativ gut wiedergeben läßt. 1985 ist die Sättigung praktisch erreicht. Abbildung 22 vereinigt Zeitreihe und Prognosereihe (bis 1985) in einem Bild.

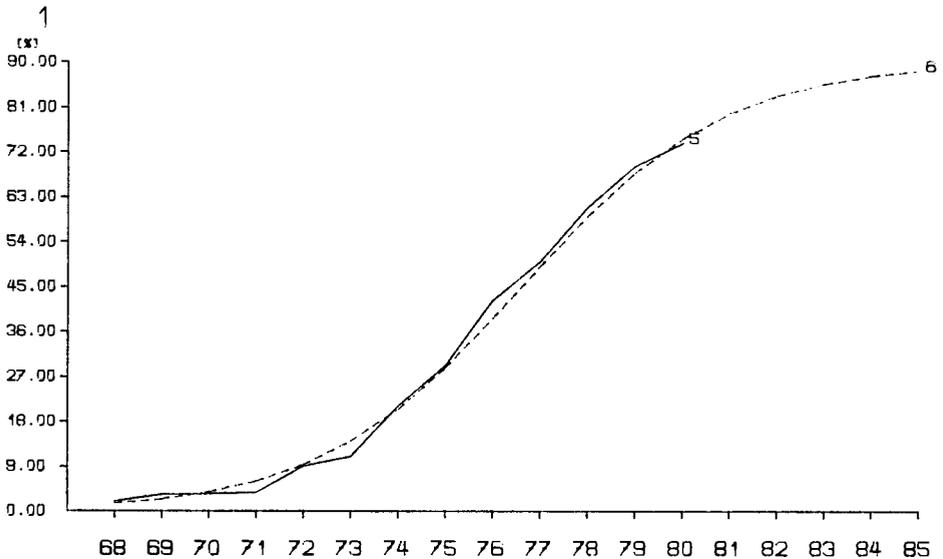


Abbildung 22

Ausstattung privater Haushalte mit Farbfernsehgeräten (Prognose gestrichelt)

IV. Die Gompertz-Funktion

1. Ableitung der Anpassungsgleichung

Die Gompertz-Funktion – benannt nach dem österreichischen Statistiker Gompertz – ist im Gegensatz zur logistischen Funktion asymmetrisch aufgebaut, da die Wachstumsrate nicht so schnell abnimmt.

Fundamentalannahme der Gompertz-Funktion:

Das Wachstum $\frac{dx}{dt}$ der Zeitreihe $x(t)$ ist *proportional*

- dem jeweils erreichten Niveau $x(t)$
- der *logarithmischen* Differenz zwischen Sättigungsniveau S und bisher erreichtem Niveau $x(t)$.

Quantitativ lautet diese Annahme

$$(120) \quad x' = \frac{dx}{dt} = ax (\ln S - \ln x) = ax \ln \left(\frac{S}{x} \right)$$

Wir substituieren $z = \ln \left(\frac{S}{x} \right)$ mit $\frac{dz}{dx} = -\frac{1}{x}$ und erhalten

$$\frac{dx}{x} \cdot \frac{1}{z} = a dt \quad \text{oder} \quad -\frac{dz}{z} = a dt$$

Das Integral

$$-\int \frac{1}{z} dz = \int a dt + C \quad \text{ergibt}$$

$$-\ln z = at + C \quad \text{und nach zweimaligem Entlogarithmieren}$$

$$x = S \cdot e^{-e^{-at-C}}$$

Setzt man hier $e^{-a} = A$ und $e^{-C} = B$ ein, so folgt der bekannte Ausdruck der Gompertz-Funktion

$$(121) \quad x(t) = S e^{-BA^t}$$

wofür die drei Parameter S , A und B aus den Zeitreihenwerten zu schätzen sind.

2. Schätzung der Parameter

Falls das Sättigungsniveau S exogen bestimmt wurde und damit als bekannt vorausgesetzt werden kann, gestaltet sich die Schätzung von A und B relativ einfach.

Aus (121) entwickeln wir durch zweimaliges Logarithmieren

$$(122) \quad \ln \left[\ln \left(\frac{S}{x_t} \right) \right] = \ln B + \ln A \cdot t$$

und schätzen hieraus mit der KQ-Methode $\ln \hat{B}$ und $\ln \hat{A}$ sowie \hat{B} und \hat{A}^1 .

Soll das Sättigungsniveau S endogen mitgeschätzt werden, so setzen wir $z_t = x_{t+1}$ und erhalten

$$z_t = S e^{-BA^{(t+1)}} = S(e^{-BA^t})^A = S^{1-A} (S e^{-BA^t})^A = S^{1-A} x_t^A$$

und daraus die Schätzgleichung für die KQ-Methode

$$(123) \quad \ln z_t = (1 - A) \ln S + A \ln x_t$$

Hieraus werden A und S ermittelt, so daß man S in (122) einsetzen kann, um nun B zu schätzen. Die (nahe) Übereinstimmung der Schätzwerte für A nach (122) und (123) ist dabei ein Maß für die Güte der Schätzung.

3. Anwendungsbeispiele

a) Ausstattung privater Haushalte mit Waschmaschinen

Die in Tabelle 19 dargestellte Zeitreihe, Ausstattung privater Haushalte mit Waschmaschinen (Quelle: Statistische Jahrbücher, 20.7 Ausstattung ausgewählter privater Haushalte mit ausgewählten langlebigen Gebrauchsgütern, Haushaltstyp 2), soll mit Hilfe der Gompertz-Funktion prognostiziert werden.

Wir gehen von einem Sättigungsniveau von 100 % aus, d. h. jeder Haushalt wird maximal eine Waschmaschine haben. Bei exogen gegebenem Sättigungsniveau lassen sich die Parameter der Funktion nach Gleichung (122) unter Anwendung der KQ-Methode schätzen. Wir erhalten für die Parameter Werte von

$$\hat{A} = 0,8925$$

und $\hat{B} = 1,8941$,

so daß sich die Prognosefunktion nach Gleichung (121) für unser Beispiel ergibt

$$\hat{x}_t = 100 \cdot e^{-1,8941 \cdot 0,8925^t}$$

Die Zeitreihen- und Prognosewerte sind in Tabelle 20 wiedergegeben.

1 Die entlogarithmierten Schätzwerte sind nicht mehr erwartungstreu, dürften aber in der Nähe der wahren Werte liegen. Zur Erwartungstreue vgl. Kapitel G., S. 129 f.

Tabelle 19

Ausstattung privater Haushalte mit Waschmaschinen (in %) von allen Privathaushalten

Jahr	t	x_t	Jahr	t	x_t
1965	1	19,8	1973	9	49,1
1966	2	18,9	1974	10	52,0
1967	3	27,4	1975	11	57,6
1968	4	32,0	1976	12	63,2
1969	5	35,4	1977	13	62,1
1970	6	37,5	1978	14	65,0
1971	7	45,6	1979	15	75,6
1972	8	45,2	1980	16	78,8

Tabelle 20

Zeitreihen- und Prognosewerte des Beispiels von Tab. 19

Jahr	t	x_t	\hat{x}_t
1965	1	19,800	18,443
1966	2	18,900	22,120
1967	3	27,400	26,016
1968	4	32,000	30,069
1969	5	35,400	34,216
1970	6	37,500	38,398
1971	7	45,600	42,561
1972	8	45,200	46,655
1973	9	49,100	50,641
1974	10	52,000	54,485
1975	11	57,600	58,161
1976	12	63,200	61,651
1977	13	62,100	64,942
1978	14	65,000	68,028
1979	15	75,600	70,905
1980	16	78,800	73,575
1981	17	—	76,043
1982	18	—	78,316
1983	19	—	80,402
1984	20	—	82,310
1985	21	—	84,051
	MAA = 2,3	WMQA = 2,6	U = 0,5

Eine gute Reproduktion der Zeitreihe durch die Gompertz-Funktion weisen die obigen Werte der Gütemaße nach, was optisch durch Abbildung 23 bestätigt wird.

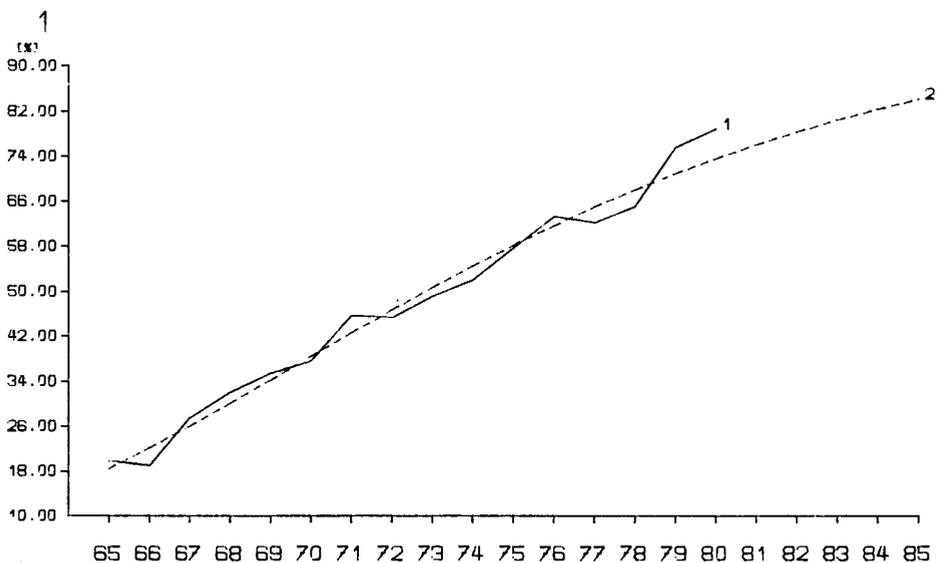


Abbildung 23

Ausstattung privater Haushalte mit Waschmaschinen (Prognose gestrichelt)

Nachdem im obigen Beispiel das Sättigungsniveau exogen bestimmt wurde, soll nun neben den Parametern A und B auch das Sättigungsniveau endogen geschätzt werden.

b) Pkw je 1000 Erwachsene Wohnbevölkerung

Die Entwicklung des Pkw-Bestandes ist schon seit 1955 ein bevorzugter Gegenstand der Langfristprognose. Die Werte der Zeitreihe, Pkw je 1000 Erwachsene Wohnbevölkerung (Quelle: Deutsche Shell AG, Aktuelle Wirtschaftsanalysen 13, Aufschwung nach der Talfahrt, Shell-Prognose des Pkw-Bestandes bis zum Jahr 2000, Hamburg 1981, S. 20), für unser Beispiel liegen seit 1952 vor und sind in Tabelle 21 wiedergegeben.

Tabelle 21

Pkw je 1 000 Erwachsene Wohnbevölkerung

Jahr	t	x_t	Jahr	t	x_t
1952	1	28,237	1967	16	258,534
1953	2	34,499	1968	17	273,691
1954	3	41,009	1969	18	295,398
1955	4	48,600	1970	19	323,342
1956	5	58,979	1971	20	344,765
1957	6	69,582	1972	21	360,773
1958	7	82,910	1973	22	373,144
1959	8	97,758	1974	23	379,473
1960	9	116,344	1975	24	396,886
1961	10	137,432	1976	25	418,440
1962	11	156,048	1977	26	441,908
1963	12	177,888	1978	27	465,526
1964	13	197,946	1979	28	482,140
1965	14	219,606	1980	29	489,689
1966	15	243,649			

Das Sättigungsniveau soll endogen geschätzt werden. Nach Anwendung der KQ-Methode auf Gleichung (123) erhalten wir für die Parameter die Schätzwerte

$$\hat{S} = 848,5238$$

und $\hat{A} = 0,9356$.

Mit Hilfe dieses Sättigungsniveaus lassen sich durch Anwendung der KQ-Methode auf Gleichung (122) die Parameter A und B schätzen. Es ergeben sich Werte von

$$\hat{A} = 0,9349$$

und $\hat{B} = 3,6222$,

so daß die Prognosegleichung (121) für dieses Beispiel lautet

$$\hat{x}_t = 848,5238 \cdot e^{-3,6222 \cdot 0,9349^t}$$

Beim Vergleich der Schätzwerte für A nach Gleichung (122) und Gleichung (123) stellen wir fest, daß eine annähernde Übereinstimmung besteht. Dies läßt eine hohe Güte der Schätzung vermuten.

Tabelle 22

Prognose Pkw je 1000 Erwachsene Wohnbevölkerung

Jahr	t	x_t	\hat{x}_t
1952	1	28,237	26,702
1953	2	34,499	35,776
1954	3	41,009	43,959
1955	4	48,600	53,295
1956	5	58,979	63,810
1957	6	69,582	75,509
1958	7	82,910	88,380
1959	8	97,758	102,391
1960	9	116,344	117,492
1961	10	137,432	133,620
1962	11	156,048	150,695
1963	12	177,888	168,628
1964	13	197,946	187,320
1965	14	219,606	206,665
1966	15	243,649	226,555
1967	16	258,534	246,879
1968	17	273,691	267,526
1969	18	295,398	288,390
1970	19	323,342	309,365
1971	20	344,765	330,354
1972	21	360,773	351,263
1973	22	373,144	372,008
1974	23	379,473	392,509
1975	24	396,886	412,698
1976	25	418,440	432,511
1977	26	441,908	451,895
1978	27	465,526	470,803
1979	28	482,140	489,196
1980	29	489,689	507,042
1981	30	—	524,314
1982	31	—	540,996
1983	32	—	557,071
1984	33	—	572,533
1985	34	—	587,376
	MAA = 8,2	WMQA = 9,6	U = 0,5

Aufgrund der Gütemaße und der Abbildung 24, die die Entwicklung der Prognose- und Zeitreihenwerte im Zeitablauf wiedergibt, können wir feststellen, daß die Gompertz-Funktion die Entwicklung des Pkw-Bestandes pro 1000 Erwachsene Wohnbevölkerung gut reproduziert.

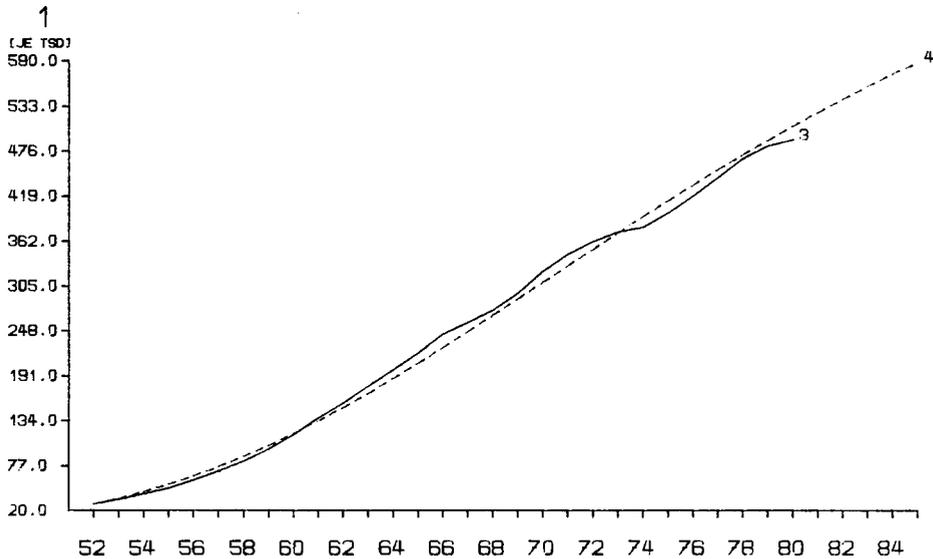


Abbildung 24

Pkw je 1000 Erwachsene Wohnbevölkerung (Prognose gestrichelt)

V. Kritische Zusammenfassung

Die Wachstums- und Sättigungsmodelle eignen sich zur Trendextrapolation von Zeitreihen, deren Struktur sich nicht mit einem einfachen linearen oder quadratischen Trend angemessen beschreiben läßt. Sie sind zur langfristigen Prognose von makroökonomischen Variablen (z. B. Bruttosozialprodukt, Rohstoffvorkommen, Energieversorgung) ebenso geeignet wie zur langfristigen Absatzprognose von Konsum- und Investitionsgütern, die einem Lebenszyklusprozeß unterliegen.

Der mathematische und programmierungstechnische Aufwand der Verfahren ist nicht so hoch, wie es auf den ersten Blick scheint, wenn ein Regressionsprogramm (z. B. SPSS) für Kleinste-Quadrate-Schätzungen verfügbar ist.

Der größte Nachteil der Verfahren ist ihre mangelnde Robustheit gegenüber Verletzungen der jeweiligen Fundamentalannahmen. Sind diese – etwas einschränkenden und mechanistischen – Annahmen nicht erfüllt, zeigen die Verfahren unangemessene Er-

gebnisse. Ihre Annahmen müssen dann verfeinert werden, woraus neue Prognosemodelle resultieren.

Ein weiterer Nachteil ist die große Schätzunsicherheit bei der Bestimmung des Sättigungsniveaus. Beim praktischen Einsatz neigen die logistischen und Gompertz-Modelle dazu, das Sättigungsniveau zu unterschätzen.

Es sind jedoch Anstrengungen unternommen worden, diese Unzulänglichkeiten durch Verfeinerung der Verfahren zu beseitigen. So arbeitet Bonus (1968) mit einem vom Volkseinkommen abhängigen variablen Sättigungsniveau und Lewandowski (1974) hat mit seinen generalisierten logistischen Funktionen die etwas mechanistischen Fundamentalannahmen zu sehr flexiblen Annahmen umgestaltet. Einen sehr guten Überblick zum weiteren Einarbeiten in die verfeinerten Wachstums- und Sättigungsmodelle mit Literaturangaben geben Lewandowski (1974 und 1980) und Mertens (1981).

G. Prognosen mit der multiplen Regressionsanalyse

I. Grundgedanken kausaler Prognosemethoden

In den vorangehenden Kapiteln wurde das zukünftige Verhalten einer Zeitreihe aus ihren eigenen Vergangenheitswerten prognostiziert. Alle angesprochenen Prognoseverfahren gehörten in den Bereich der Zeitreihenanalyse. In diesem Kapitel wird versucht, das Verhalten einer Zeitreihe auf das Verhalten anderer, sog. *exogener* Zeitreihen zurückzuführen und eine solche Kausalbeziehung zur Prognose auszunutzen.

Es muß ausdrücklich betont werden, daß eine Kausalbeziehung zwischen zwei Zeitreihen niemals mit statistischen Mitteln aufgedeckt oder gar bewiesen werden kann, sondern nur mit Hilfe einer fachwissenschaftlichen Theorie. Die Statistik stellt nur Korrelationen fest, die Theorie bestimmt die Kausalrichtung.

In vielen Fällen läßt sich die Kausalrichtung leicht feststellen, z. B. beeinflußt die Zeitreihe der Baugenehmigungen die Zeitreihe des Badewannenabsatzes und nicht umgekehrt. Ähnlich verhält es sich mit dem Auftragseingang der Industrie und ihrem Produktionsvolumen.

Schwieriger wird es mit Zeitreihen, die sich gegenseitig beeinflussen wie privater Konsum und Volkseinkommen. Diese sog. interdependenten Systeme (vgl. z. B. Schneeweiß 1978 S. 242 ff.) verlangen besondere Verfahren der Ökonometrie für ihre Analyse, auf die wir in diesem Buch nicht eingehen können.

Im betriebswirtschaftlichen Bereich ist die Kausalrichtung bei absatzpolitischen Maßnahmen, wie Werbekampagnen oder Preisveränderungen, von vornherein feststellbar. Auch der Einfluß makroökonomischer Variablen, wie z. B. des Volkseinkommens oder des langfristigen Kapitalmarktzinses auf den Absatz einzelner Produkte (Pkw, Eigentumswohnungen, Unterhaltungselektronik u. a.) ist in seiner Richtung eindeutig, so daß sich die nachfolgend beschriebenen Verfahren besonders für diese Gebiete eignen.

Hat man für ein konkretes Prognoseproblem die Kausalrichtung in den Beziehungen zwischen den Zeitreihen theoretisch geklärt, so kann nun mit den Mitteln der Statistik und Ökonometrie das *Ausmaß* dieser Beziehungen quantitativ erforscht werden. Bezeichnen wir die zu prognostizierende Zeitreihe, auch *abhängige oder endogene Variable* genannt, mit $\{y_t\}$ und die n beeinflussenden Zeitreihen, auch *unabhängige oder exogene Variablen* genannt, mit $\{x_{ti}\}$ ($i = 1, \dots, n$), so gilt, da alle Zeitreihen Realisationen stochastischer Prozesse sind,

$$(124) \quad y_t = f(x_{1t}, x_{2t}, \dots, x_{nt}) + u_t \quad (t = 1, \dots, T)$$

u_t ist hier wieder die uns schon bekannte Störvariable mit den auf S. 30 angegebenen Eigenschaften.

Gleichung (124) nennt man eine *Regressionsfunktion*, die die Variation von $\{y_t\}$ auf die Variation der exogenen Variablen zurückführt.

Zur Prognose eignet sich die Regressionsfunktion, wenn es gelingt,

- den Funktionsverlauf aus den Daten richtig zu spezifizieren und
- die (zukünftigen) Werte der unabhängigen Variablen $\{x_{it}\}$ rechtzeitig für eine Prognose von $\{y_t\}$ zu ermitteln.

Die zweite Bedingung ist leicht zu erfüllen, wenn abhängige Variable und unabhängige Variable *zeitlich verzögert* miteinander korreliert sind, z. B.

$$y_t = f(x_{1t-2}, x_{2t-2}, \dots, x_{nt-2}) + u_t$$

da in diesem Fall die gemessenen Werte der exogenen Variablen eine Prognose von y_t für zwei Perioden im voraus erlauben. Korrelieren alle Variablen ohne time-lag miteinander, so müssen zunächst die exogenen Variablen für die Prognoseperiode prognostiziert und dann in die Regressionsfunktion eingesetzt werden. Eine solche Regressionsprognose lohnt sich nur dann, wenn die exogenen Variablen leichter zu prognostizieren sind als $\{y_t\}$, sonst sollte man die *Zeitreihenanalyse* der vorigen Kapitel anwenden.

Eine Schätzung von $\{y_t\}$ mit Hilfe der Regressionsfunktion (124) wird *multiple Regressionsanalyse* genannt. Ihre Voraussetzungen, Schätzmethoden und Anwendungsprobleme sind Gegenstand des nächsten Abschnitts.

II. Das multiple Regressionsmodell

1. Grundannahmen

Das allgemeine Regressionsmodell geht von der Regressionsfunktion (124) aus. Grundsätzlich könnte man versuchen, den Funktionsverlauf (z. B. Polynomgrad oder Exponentialform) aus den Daten der unabhängigen und abhängigen Variablen zu schätzen. Da es aber sehr viele unterschiedliche mögliche Funktionsformen gibt, reicht bei ökonomischen Problemen die *Zeitreihenlänge* in den meisten Fällen nicht aus, um die Funktionsform hinreichend genau schätzen zu können.

Man spezifiziert daher von vornherein die Regressionsfunktion zu einer *linearen Funktion*

$$(125) \quad y_t = b_0 + b_1 x_{1t} + b_2 x_{2t} + \dots + b_n x_{nt} + u_t \quad (t = 1, \dots, T)$$

und schätzt nur noch die Parameter b_i ($i = 0, 1, \dots, n$).

Diese Beschränkung der Funktionsform scheint auf den ersten Blick recht willkürlich, ist jedoch aus zwei Gründen vorteilhaft:

- Das Schätzproblem vereinfacht sich ungemein, da bei Anwendung der KQ-Methode nur ein lineares Gleichungssystem zu lösen ist.

– Nicht-lineare Funktionen können innerhalb eines gewissen Variationsbereiches der Variablen linear approximiert werden (vgl. die MA-Schätzungen im Box-Jenkins-Verfahren, S. 84 ff.).

Insbesondere bei den Wachstums- und Sättigungsmodellen konnten wir auch nicht-lineare Funktionen durch Logarithmieren linearisieren, so daß es generell gerechtfertigt erscheint, im folgenden nur noch lineare Regressionsfunktionen zu analysieren.

Außer dieser Linearitätsannahme benötigen wir noch weitere Grundannahmen über die Störvariable u und die Lösbarkeit der Normalgleichungen, die im folgenden zusammengestellt sind.

Zur kürzeren und übersichtlicheren Schreibweise bedienen wir uns von hier ab der Matrixschreibweise, wie sie auch in den einschlägigen Lehrbüchern der Statistik und Ökonometrie verwandt wird (vgl. Schneeweiß 1978, Johnston 1963, Menges 1961, Hochstädter/Uebe 1970).

Durch Einführen einer Scheinvariablen x_{0t} , die für alle Perioden $t = 1, \dots, T$ den Wert Eins erhält, können wir unsere Regressionsfunktionen

$$(125) \quad y_t = b_0 x_{0t} + b_1 x_{1t} + \dots + b_n x_{nt} + u_t \quad (t = 1, \dots, T)$$

in Matrizenform schreiben

$$(125a) \quad y = Xb + u$$

mit

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_T \end{pmatrix} \quad X = \begin{pmatrix} x_{01} & x_{11} & \dots & x_{n1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{0T} & x_{1T} & \dots & x_{nT} \end{pmatrix} \quad u = \begin{pmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ u_T \end{pmatrix} \quad b = \begin{pmatrix} b_0 \\ b_1 \\ \cdot \\ \cdot \\ b_n \end{pmatrix}$$

Für diese Regressionsfunktion setzen wir nun folgende Prämissen (vgl. Schneeweiß 1978, S. 92 ff.)¹:

$$P 1 : E[u|X] = 0$$

Der Erwartungswert der Störvariablen bei einer gegebenen Datenkonstellation X soll Null sein. Diese Annahme haben wir in ähnlicher Form in den vorangehenden Kapiteln gemacht.

$$P 2 : E[u u^T | X] = \sigma^2 E \quad (E \triangleq \text{Einheitsmatrix})$$

¹ Der hochgestellte Index T gibt die Transponierung der Matrix an.

Diese Annahme, die wir wiederholt benutzt haben, beinhaltet einmal die zeitliche Konstanz der Störvariablen (Homoskedastizität) und zum anderen die Unkorreliertheit der u_t für verschiedene t : $E[u_{t+i}] = 0$ ($i \neq 0$)

$$P 3 : X^T X \text{ ist invertierbar, d. h. } \det(X^T X) \neq 0$$

Zwischen den exogenen Variablen darf also keine *lineare Abhängigkeit* bestehen, da die Normalgleichungen der KQ-Methode in diesem Fall nicht eindeutig lösbar sind.

$$P 4 : \lim_{T \rightarrow \infty} \frac{1}{T} X^T X \text{ ist endlich und regulär.}$$

Erwartungswerte und Varianzen der exogenen Variablen müssen bei $T \rightarrow \infty$ endliche Werte annehmen. Für trendbehaftete Variable ist diese Annahme *nicht* erfüllt. Annahme P4 beinhaltet aber ebenfalls die Forderung, daß die Varianzen der exogenen Variablen nicht Null sein dürfen, da sonst eine Schätzung unmöglich ist.

$$P 5 : u \text{ ist (bei gegebenem } X) \text{ normalverteilt.}$$

Diese Annahme wird zur Konstruktion von Konfidenzintervallen, z. B. für die Regressionsparameter, benötigt. Sie erscheint auch plausibel, da die Störvariable den gemeinsamen Einfluß vieler unbedeutender Faktoren widerspiegeln soll und der zentrale Grenzwertsatz eine Normalverteilung dieser Gesamtheit von Störungen nahelegt.

Mit diesen fünf Prämissen und der Regressionsfunktion (125a) ist das multiple Regressionsmodell vollständig beschrieben und wir können uns der Parameterschätzung zuwenden.

2. Schätzung der Regressionsparameter

Am verbreitetsten in der Ökonometrie ist das Maximum-Likelihood-Prinzip (ML) zur Schätzung der Regressionsparameter. Dabei werden die b_i ($i = 0, \dots, n$) so festgesetzt, daß die y_t die größte Wahrscheinlichkeitsdichte besitzen, wenn die b_i -Schätzungen die tatsächlichen Parameterwerte sind.

Dieses Schätzprinzip führt, wenn keine interdependenten Regressionsvariablen vorhanden sind, zu denselben Schätzformeln wie die uns bereits bekannte Methode der Kleinsten Quadrate (KQ) (Schneeweiß 1978, S. 166). Da wir die KQ-Methode bisher nur mit einer unabhängigen Variablen (t) angewandt haben, leiten wir nun die Normalgleichungen für ein multiples Modell in Matrizenform hier ab.

Die Summe der Abweichungsquadrate $\sum_{t=1}^T u_t^2$ wird bei der KQ-Methode minimiert, wobei (125a) zu beachten ist

$$(126) \quad u^T u = (y - Xb)^T (y - Xb) = y^T y - 2y^T Xb + b^T X^T Xb \Rightarrow \text{Min}$$

Die erste Ableitung im Matrizenkalkül lautet gleich Null gesetzt

$$(127) \quad \frac{d(\mathbf{u}^T \mathbf{u})}{d\mathbf{b}} = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \mathbf{b} = 0$$

Daraus ergeben sich die Normalgleichungen

$$(128) \quad \mathbf{X}^T \mathbf{X} \hat{\mathbf{b}} = \mathbf{X}^T \mathbf{y}$$

und nach Inversion von $\mathbf{X}^T \mathbf{X}$ (nach P 3 ist dies möglich!)

$$(129) \quad \hat{\mathbf{b}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

Bei mehreren unabhängigen Variablen und normaler Zeitreihenlänge ist diese Inversion nur mit EDV-Unterstützung zu bewältigen. Hierfür stehen aber benutzerfreundliche Programmpakete (z. B. SPSS) zur Verfügung.

Nach Durchführung der Matrizeninversion enthält der Vektor $\hat{\mathbf{b}}$ die KQ-Schätzwerte der Regressionsparameter, die den Maximum-Likelihood-Schätzungen in den hier beschriebenen Anwendungsbeispielen entsprechen.

Die Schätzung der Varianz der Störvariablen σ^2 (vgl. Grundannahme P 2) erfolgt mit Hilfe der Residual-Varianz

$$\hat{\sigma}^2 = \frac{1}{T-n-1} \hat{\mathbf{u}}^T \hat{\mathbf{u}} = \frac{\sum_{t=1}^T \hat{u}_t^2}{T-n-1}$$

$T-n-1$ ist die Anzahl der Freiheitsgrade, da die Residuen den $n+1$ linearen Restriktionen

$$(130) \quad \mathbf{X}^T \hat{\mathbf{u}} = \mathbf{0}$$

unterliegen, die man erhält, wenn $\hat{\mathbf{u}} = \mathbf{y} - \hat{\mathbf{y}}$ und (125a) in (128) eingesetzt wird.

Wir wollen uns nun mit den *Eigenschaften* unserer KQ-Schätzfunktion² beschäftigen.

Zunächst ist die Schätzfunktion (129) linear in bezug auf die abhängige Variable y .

Die zweite Eigenschaft von $\hat{\mathbf{b}}$ ist die *Erwartungstreue* (Unverzerrtheit). Sie läßt sich aus (129) und (125a) ableiten:

$$(131) \quad \hat{\mathbf{b}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{X} \mathbf{b} + \mathbf{u}) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} \mathbf{b} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u}$$

$$(131a) \quad E[\hat{\mathbf{b}}] = \mathbf{b} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T E[\mathbf{u}]$$

$$(131b) \quad E[\hat{\mathbf{b}}] = \mathbf{b}$$

² (129) ist eine *Schätzfunktion*. Man erhält die *Schätzwerte* durch Einsetzen der Realisationen in die Schätzfunktion.

Unter allen linearen erwartungstreuen Schätzfunktionen ist die KQ-Schätzfunktion $\hat{\mathbf{b}}$ insofern die *beste*, als sie die *kleinste Varianz* aufweist. Wegen des formalen Aufwands sei dieses Ergebnis hier nicht abgeleitet. Der Beweis findet sich in den Lehrbüchern der Ökonometrie (vgl. z. B. Schneeweiß 1978, S. 108–109).

Diese drei Eigenschaften bilden zusammen die sog. BLUE-Eigenschaft (= *best linear unbiased estimator*), die den KQ-Schätzvektor $\hat{\mathbf{b}}$ auszeichnet.

Eine weitere wichtige Eigenschaft von $\hat{\mathbf{b}}$ ist die *Konsistenz*, die besagt, daß die Varianz der Regressionsschätzungen $\hat{\mathbf{b}}_i$ ($i = 0, 1, \dots, n$) mit zunehmendem Zeitreihenumfang gegen Null strebt.

Um dies einzusehen, gehen wir von der Kovarianzmatrix von $\hat{\mathbf{b}}$ aus

$$\text{VAR}(\hat{\mathbf{b}}) = E[(\hat{\mathbf{b}} - \mathbf{b})(\hat{\mathbf{b}} - \mathbf{b})^T | \mathbf{X}]$$

Durch Einsetzen von (131) ergibt sich nach einigen Umformungen

$$(132) \quad \text{VAR}(\hat{\mathbf{b}}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} = \frac{\sigma^2}{T} \left[\frac{1}{T} (\mathbf{X}^T \mathbf{X}) \right]^{-1}$$

Wächst der Zeitreihenumfang über alle Grenzen, so gilt

$$(133) \quad \lim_{T \rightarrow \infty} \text{VAR}(\hat{\mathbf{b}}) = \lim_{T \rightarrow \infty} \frac{\sigma^2}{T} \cdot \lim_{T \rightarrow \infty} \left[\frac{1}{T} (\mathbf{X}^T \mathbf{X}) \right]^{-1}$$

Da der erste Faktor gegen Null strebt und der zweite Faktor gemäß Annahme P 4 endlich ist, folgt

$$(134) \quad \lim_{T \rightarrow \infty} \text{VAR}(\hat{\mathbf{b}}) = \mathbf{0}$$

Die Schätzfunktion $\hat{\mathbf{b}}$ ist also konsistent.

Bevor wir uns der Prognose mit dem multiplen Regressionsmodell zuwenden, seien zunächst Prüfkriterien für die Güte der Regressionsschätzung entwickelt.

3. Prüfkriterien für die Regressionsschätzung

a) Das Bestimmtheitsmaß

Das in der Praxis am häufigsten verwendete Maß für die Güte einer Regressionsschätzung beruht auf folgendem Gedankengang: Die Variation der abhängigen Variablen y wird bestimmt durch

- die Variation der unabhängigen Variablen \mathbf{X}
- die Schwankungen des Residuums $\hat{\mathbf{u}}$

Die Regressionsschätzung kann als um so besser angesehen werden, je größer der Variationsanteil von y ausfällt, der auf die unabhängigen Variablen zurückzuführen ist. Diese Überlegung läßt sich quantifizieren.

Aus $y = X\hat{b} + \hat{u} = \hat{y} + \hat{u}$ folgt

$$(135) \quad y^T y = (\hat{y} + \hat{u})^T (\hat{y} + \hat{u}) = \hat{y}^T \hat{y} + 2\hat{y}^T \hat{u} + \hat{u}^T \hat{u} = \hat{y}^T \hat{y} + \hat{u}^T \hat{u}$$

da nach (130) $\hat{y}^T \hat{u} = (X\hat{b})^T \hat{u} = \hat{b}^T X^T \hat{u} = 0$ gilt.

Division von (135) durch T und beiderseitige Subtraktion von $\bar{y}^2 = (\bar{y} - \bar{\hat{u}})^2 = \bar{y}^2$ ergibt³

$$(135a) \quad \frac{1}{T} y^T y - \bar{y}^2 = \frac{1}{T} \hat{y}^T \hat{y} - \bar{y}^2 = \frac{1}{T} \hat{u}^T \hat{u}$$

d. h. eine Zerlegung der Varianzen

$$(135b) \quad s_y^2 = s_{\hat{y}}^2 + s_{\hat{u}}^2$$

Man nennt $s_{\hat{y}}^2$ die „erklärte“ Varianz von y , da sie auf die unabhängigen Variablen zurückgeführt werden kann und $s_{\hat{u}}^2$ die „nicht erklärte“ Varianz. Der Ausdruck

$$(136) \quad R^2 = \frac{s_{\hat{y}}^2}{s_y^2} = 1 - \frac{s_{\hat{u}}^2}{s_y^2}$$

heißt *Bestimmtheitsmaß* und gibt den Anteil der erklärten Varianz an der Gesamtvarianz der abhängigen Variablen an.

R^2 variiert zwischen Null und Eins. Man bezeichnet häufig ein Regressionsmodell als geeignet, wenn R^2 nahe bei Eins liegt, doch läßt sich keine exakte Grenze angeben.

Darüber hinaus sollte man aus drei Gründen das Bestimmtheitsmaß vorsichtig interpretieren:

- Das nach (136) berechnete R^2 ist eine *Schätzfunktion* für das *wahre* Bestimmtheitsmaß und kann, insbesondere bei kleinem Zeitreihenumfang, mehr oder weniger stark von diesem abweichen.
- R^2 wächst monoton mit der Anzahl der unabhängigen Variablen; dabei nimmt aber auch die Gefahr zu, daß P 3 durch lineare Abhängigkeiten in der Matrix X verletzt wird.

3 \bar{y} bedeutet das arithmetische Mittel von y : $\frac{1}{T} \sum_{t=1}^T y_t$

Da $E[u] = 0$ gilt, nimmt man an, daß $\bar{\hat{u}} = \frac{1}{T} \sum_{t=1}^T \hat{u}_t$ bei großem Zeitreihenumfang nahe bei Null liegt.

- Enthalten die Variablen einen versteckten Trend, so wächst R^2 mit *zunehmendem* Zeitreihenumfang, da dann s_y^2 größer wird, während s_λ^2 möglicherweise konstant bleibt. Die „verbesserte Eignung des Modells ist hier nur scheinbar.

Berücksichtigt man diese Tatbestände, so läßt sich das Bestimmtheitsmaß durchaus als erster Anhaltspunkt für die Güte einer Regressions-schätzung verwenden. Darüber hinaus sind jedoch zusätzliche Prüfkriterien nötig, die nun entwickelt werden.

b) Parameter-tests für den Schätzvektor

Wir wollen mit diesen Tests zwei Fragen beantworten:

1. Ist der Einfluß *aller* unabhängigen Variablen auf die abhängige Variable *signifikant*?
2. Ist der Einfluß einer bestimmten Variablen i ($i = 1, \dots, n$) auf die abhängige Variable *signifikant*?

Zur Beantwortung der ersten Frage testen wir die Null-Hypothese

$$H_0 : b_1 = b_2 = b_3 = \dots = b_n = 0$$

gegen die Alternativhypothese H_1 , daß die Regressionskoeffizienten b_i ($i = 1, \dots, n$) ungleich Null sind.

Die Null-Hypothese zur Beantwortung der zweiten Frage lautet entsprechend

$$H_0 : b_i = 0 \quad (\text{ein einzelner Parameter})$$

Zur Konstruktion einer geeigneten Testgröße benötigen wir die Grundannahme P 5: „ u ist normalverteilt“, auf die deshalb noch einmal hingewiesen sei.

Den Vektor b , der sich aus der Null-Hypothese ergibt, bezeichnen wir mit b_0 sowie entsprechend

$$y_0 = Xb_0 \quad \text{und} \quad u_0 = y - y_0$$

Für u_0 ergibt sich dann folgender Ausdruck

$$u_0 = y - \hat{y} + \hat{y} - y_0 = \hat{u} + X\hat{b} - Xb_0 = \hat{u} + X(\hat{b} - b_0)$$

Daraus läßt sich die Abweichungsquadratsumme $u_0^T u_0$ entwickeln

$$(137) \quad u_0^T u_0 = \hat{u}^T \hat{u} + (\hat{b} - b_0)^T X^T X (\hat{b} - b_0)$$

da das Produkt $2\hat{u}^T X(\hat{b} - b_0)$ nach (130) Null wird.

$\hat{u}^T \hat{u}$ ist χ^2 -verteilt mit $T-n-1$ Freiheitsgraden (vgl. S. 129), und $(\hat{b} - b_0)^T X^T X (\hat{b} - b_0)$

ist χ^2 -verteilt mit $n + 1$ Freiheitsgraden, die aus den $n + 1$ Werten $(\hat{b}_i - b_{0i})$ ($i = 0, \dots, n$) resultieren.

Der Quotient dieser beiden unabhängigen Größen, dividiert durch die Anzahl der Freiheitsgrade, ergibt unsere Testgröße F

$$(138) \quad \frac{(\hat{\mathbf{b}} - \mathbf{b}_0)^T \mathbf{X}^T \mathbf{X} (\hat{\mathbf{b}} - \mathbf{b}_0)}{\hat{\mathbf{u}}^T \hat{\mathbf{u}}} \cdot \frac{T-n-1}{n+1} = F_{T-n-1}^{n+1}$$

Sie ist F -verteilt mit $n+1$ und $T-n-1$ Freiheitsgraden. Eine interpretationsfähige Form der Testgröße ergibt sich, wenn wir (137) in (138) einsetzen

$$(138a) \quad F_{T-n-1}^{n+1} = \frac{\mathbf{u}_0^T \mathbf{u}_0 - \hat{\mathbf{u}}^T \hat{\mathbf{u}}}{\hat{\mathbf{u}}^T \hat{\mathbf{u}}} \cdot \frac{T-n-1}{n+1}$$

Die Null-Hypothese kann um so eher abgelehnt werden, je größer ihre Abweichungsquadratsumme $\mathbf{u}_0^T \mathbf{u}_0$ gegenüber derjenigen der KQ-Schätzung ist, d. h. je größer F ausfällt. Aus der theoretischen F -Verteilung kann man für ein bestimmtes Signifikanzniveau (z. B. 95 %) einen kritischen Wert entnehmen, bei dessen Überschreiten die Null-Hypothese abgelehnt wird.

Nun wollen wir dieses allgemeine Ergebnis auf unsere erste Null-Hypothese anwenden. Wenn keine Variable einen Einfluß auf y ausübt, sind die b_i ($i = 1, \dots, n$) gleich Null, nur die Konstante b_0 nicht. Sie ist in diesem Fall (entsprechend der KQ-Methode) gleich dem arithmetischen Mittel von y , d. h. es gilt

$$\begin{aligned} \mathbf{u}_0 &= \mathbf{y} - \mathbf{X}\mathbf{b}_0 = \mathbf{y} - \bar{y} & (b_0^T &= (b_0, 0, \dots, 0)) \\ & & (\bar{y}^T &= (\bar{y}, \bar{y}, \dots, \bar{y})) \end{aligned}$$

Daraus folgt für (138a)

$$\begin{aligned} (138b)^4 \quad F_{T-n-1}^n &= \frac{(\mathbf{y} - \bar{y})^T (\mathbf{y} - \bar{y}) - \hat{\mathbf{u}}^T \mathbf{u}}{\hat{\mathbf{u}}^T \hat{\mathbf{u}}} \cdot \frac{T-n-1}{n} \\ &= \frac{s_y^2 - s_{\hat{\mathbf{u}}}^2}{s_{\hat{\mathbf{u}}}^2} \cdot \frac{T-n-1}{n} \end{aligned}$$

Berücksichtigen wir die Definition des Bestimmtheitsmaßes in Gleichung (136), kann die Testgröße F durch R^2 ausgedrückt werden

$$(139) \quad F_{T-n-1}^n = \frac{R^2}{1 - R^2} \cdot \frac{T-n-1}{n}$$

⁴ Die Anzahl der Zählerfreiheitsgrade vermindert sich um einen, da wir nur n Parameter (ohne b_0) testen.

Überschreitet diese empirisch ermittelte Testgröße den kritischen Wert der F-Verteilung mit n und $T-n-1$ Freiheitsgraden, so wird die Nullhypothese abgelehnt und statt dessen angenommen, daß *alle* unabhängigen Variablen *zusammen* einen signifikanten Einfluß auf die abhängige Variable ausüben und auch das Bestimmtheitsmaß signifikant ist.

Zum Test der zweiten Null-Hypothese: „Der i -te Regressionsparameter ist gleich Null“ gehen wir von (138) aus, sowie von der Definition

$$\hat{\sigma}^2 = \frac{1}{T-n-1} \hat{\mathbf{u}}^T \hat{\mathbf{u}} \text{ von Seite 129.}$$

Daraus ergibt sich

$$(140) \quad F_{T-n-1}^{n+1} = \frac{1}{\hat{\sigma}^2} \cdot \frac{1}{n+1} (\hat{\mathbf{b}} - \mathbf{b}_0)^T \mathbf{X}^T \mathbf{X} (\hat{\mathbf{b}} - \mathbf{b}_0)$$

Vergegenwärtigen wir uns die Kovarianzmatrix des Schätzvektors in Gleichung (132) bzw. deren Schätzung

$$(141) \quad \text{VAR}(\hat{\mathbf{b}}) = \hat{\sigma}^2 (\mathbf{X}^T \mathbf{X})^{-1}$$

so können wir die invertierte Matrix $[\text{VAR}(\hat{\mathbf{b}})]^{-1} = \frac{1}{\hat{\sigma}^2} (\mathbf{X}^T \mathbf{X})$ in (140) einsetzen und erhalten

$$(142) \quad F_{T-n-1}^{n+1} = \frac{1}{n+1} (\hat{\mathbf{b}} - \mathbf{b}_0)^T [\text{VAR}(\hat{\mathbf{b}})]^{-1} (\hat{\mathbf{b}} - \mathbf{b}_0)$$

Für einen einzelnen Regressionsparameter i vereinfacht sich diese Testgröße zu

$$(143) \quad F_{T-n-1}^1 = \frac{(\hat{b}_i - b_{0i})^2}{\hat{\sigma}^2 (b_i)}$$

Wir berechnen diese Testgröße für jeden Regressionsparameter und nehmen die Alternativhypothese $b_i \neq 0$ an, wenn F größer als der kritische Wert der F-Verteilung ausfällt. In diesem Fall sagen wir, daß die unabhängige Variable i einen signifikanten Einfluß auf die abhängige Variable ausübe.

Mit den in diesem Abschnitt entwickelten Prüfkriterien sind wir in der Lage, die Eignung eines Regressionsmodells zu beurteilen. Bevor wir zur Prognose übergehen, sind aber noch zwei Sonderprobleme zu betrachten, die bei der Regression ökonomischer Zeitreihen sehr häufig auftreten.

4. Autokorrelation und Multikollinearität

Bei vielen ökonomischen Zeitreihen, insbesondere solchen, die einen versteckten Trend aufweisen, ist die Grundannahme P 2 verletzt, weil die u_t miteinander korreliert sind. Diesen Fall bezeichnet man als *Autokorrelation* der Störvariablen.

Die KQ-Schätzfunktion ist zwar noch erwartungstreu, hat aber nicht mehr minimale Varianz. Darüber hinaus unterschätzen die $\hat{\sigma}^2(b_i)$ die wahren Varianzen.

Wir benötigen daher

- ein Prüfmaß zur Feststellung einer evtl. Autokorrelation und
- ein Verfahren zur Beseitigung der Auswirkungen der Autokorrelation.

Zur Prüfung der Autokorrelation wird in der ökonometrischen Praxis meistens der Durbin-Watson-Test (Durbin/Watson 1950) durchgeführt, der folgende Prüfgröße benutzt

$$(144) \quad d = \frac{\sum_{t=2}^T (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^T \hat{u}_t^2}$$

Durch Ausmultiplizieren erkennt man die Verwandtschaft mit dem Autokorrelationskoeffizienten 1. Ordnung des Residuums ($\hat{\rho}$)

$$(144a) \quad d = \frac{\sum_{t=2}^T \hat{u}_t^2}{\sum_{t=1}^T \hat{u}_t^2} - 2 \frac{\sum_{t=2}^T \hat{u}_t \hat{u}_{t-1}}{\sum_{t=1}^T \hat{u}_t^2} + \frac{\sum_{t=2}^T \hat{u}_{t-1}^2}{\sum_{t=1}^T \hat{u}_t^2} \approx 2 - 2\hat{\rho}$$

Bei positiver Autokorrelation ($\hat{\rho} \rightarrow 1$) strebt d gegen Null, bei negativer Autokorrelation ($\hat{\rho} \rightarrow -1$) gegen 4, so daß der Definitionsbereich der Testgröße $0 \leq d \leq 4$ ist.

Ist keine Autokorrelation vorhanden ($\hat{\rho} \rightarrow 0$), so nähert sich d dem Wert 2.

Durbin und Watson (1951) haben Tabellen entwickelt, die es gestatten, in Abhängigkeit von T und n Signifikanzgrenzen für d anzugeben. Werden z. B. $n = 4$ unabhängige Variable mit jeweils $T = 60$ Zeitreihenwerten im Regressionsmodell verwendet, so liegt mit 95 % Sicherheitswahrscheinlichkeit *keine* Autokorrelation vor, wenn d in das Intervall

$$1,73 < d < 2,27$$

fällt. Wegen weiterer Werte sei auf die ökonometrische Fachliteratur verwiesen (z. B. Schneeweiß 1978, Anhang E).

Hat man eine Autokorrelation im Regressionsmodell festgestellt, so kann man diese in vielen Fällen durch eine Variablentransformation beseitigen (Hansmann 1980 d, S. 115 ff.).

Wir nehmen dazu an, daß für die Störvariable ein autoregressiver Prozeß 1. Ordnung besteht, d. h. daß nur unmittelbar benachbarte u_t korreliert sind (vgl. die Analogie zum MA(1)-Prozeß bei der Zeitreihenanalyse). Dies trifft in der Realität häufig zu. Der Prozeß läßt sich dann wie folgt schreiben

$$(145) \quad u_t = \rho u_{t-1} + \epsilon_t$$

$$\text{mit} \quad E[\epsilon_t] = 0 \quad E[\epsilon_t^2] = \sigma_\epsilon^2 \quad E[\epsilon_t \epsilon_{t+i}] = 0 \quad \text{für} \quad i \neq 0$$

Die neue Störvariable ϵ_t erfüllt unsere Grundannahme P 2.

Gleichung (145) wird nun benutzt, um die „autokorrelierte“ Regressionsgleichung

$$(146) \quad y_t = b_0 + \sum_{i=1}^n b_i x_{it} + u_t$$

zu transformieren

$$(147) \quad \epsilon_t = (y_t - b_0 - \sum_{i=1}^n b_i x_{it}) - \rho (y_{t-1} - b_0 - \sum_{i=1}^n b_i x_{it-1})$$

Man erhält durch Umstellung die neue Regressionsgleichung

$$(148) \quad y_t - \rho y_{t-1} = b_0(1 - \rho) + \sum_{i=1}^n b_i (x_{it} - \rho x_{it-1}) + \epsilon_t$$

deren Störvariable nicht autokorreliert ist. Wenn die Regression dieser „modifizierten Differenzen“ zu ähnlichen Schätzwerten \hat{b}_i führt wie die Regression der ursprünglichen Gleichung, so hat man damit für die praktische Anwendung das Problem der Autokorrelation im Griff (verfeinerte Verfahren finden sich in Schneeweiß 1978)⁵.

Ein zweites Problem, das bei ökonomischen Regressionsmodellen gelegentlich auftritt, ist die *Multikollinearität*. Man versteht darunter eine hohe Korrelation zwischen den exogenen Variablen. Ist ein Korrelationskoeffizient zwischen ihnen gleich Eins, liegt lineare Abhängigkeit vor und $X^T X$ ist nicht invertierbar, d. h. Grundannahme P 3 ist verletzt und die Normalgleichungen sind nicht eindeutig lösbar.

Aber auch bei stochastischer linearer Abhängigkeit ($|r| < 1$) ist die Multikollinearität sehr unangenehm, da die Einflüsse der korrelierten exogenen Variablen auf y nicht richtig getrennt werden können. Die Parameter b_i sind nur schwach identifizierbar. Außerdem treten große Schätzfehler auf.

⁵ Für ρ kann man die Schätzung aus Gleichung (144 a) verwenden.

Zu erkennen ist die Multikollinearität an der Korrelationsmatrix der exogenen Variablen und an der Determinante dieser Matrix. Hat die Determinante einen Wert nahe bei Null, so kann man auf Multikollinearität schließen. Die betroffenen Variablen sind in der Regel durch hohe Korrelationskoeffizienten gekennzeichnet.

Ein einfaches Verfahren zur Verminderung der Multikollinearität besteht darin, einige der hoch korrelierenden Variablen aus dem Modell zu entfernen. Dies ist aber nicht ganz ungefährlich, da systematische Verzerrungen die Folge sein können (Schneeweiß 1978, S. 140). Andere Möglichkeiten sind erstens die Bildung von Differenzen der Zeitreihen, um Trends auszuschalten, die häufig die Multikollinearität verursachen oder zweitens die Substitution der korrelierenden Variablen durch geeignete Linearkombinationen.

Bleibt die Multikollinearität auch während des Prognosezeitraums erhalten, so sind die schädlichen Auswirkungen für die *Prognose* nicht so gravierend, da die Güte der Prognose nicht von der Identifizierbarkeit der Einflüsse einzelner Variablen abhängt.

5. Die Prognosegleichungen

Zum Abschluß der Darstellung des multiplen Regressionsmodells sollen hier die Prognosegleichungen zusammengestellt werden, die wir im nächsten Abschnitt praktisch anwenden werden.

Unter der Voraussetzung, daß die Werte der exogenen Variablen für den Prognosezeitraum m bekannt sind, lautet die Prognosegleichung

$$(149) \quad \hat{y}_{t+m} = \hat{b}_0 + \hat{b}_1 x_{1t+m} + \hat{b}_2 x_{2t+m} + \dots + \hat{b}_n x_{nt+m}$$

Diese KQ-Prognoseformel weist die kleinste Fehlervarianz auf.

Liegt im Regressionsmodell *Autokorrelation* 1. Ordnung vor, so kann man Gleichung (145) benutzen, um die autokorrelierte Störvariable der Prognosegleichung hinzuzufügen, z. B.:

$$(150) \quad \hat{y}_{t+1} = \hat{b}_0 + \sum_{i=1}^n \hat{b}_i x_{it+1} + \rho \hat{u}_t$$

Dadurch verbessern sich – nach den Erfahrungen des Verfassers – die Prognosen (vgl. Hansmann 1980 d, S. 118).

III. Prognose des Aktienkursindex mit einem multiplen Regressionsmodell

Die Zeitreihe des Aktienkursindex besteht aus den 100 monatlichen Indexwerten vom Januar 1974 bis April 1982, die wir schon bei der Prognose mit dem Box-Jenkins-Verfahren zugrunde gelegt hatten (vgl. S. 90).

Die zusätzliche Aufgabe bei der Prognose mit einem Regressionsmodell besteht darin, geeignete Zeitreihen zu finden, die den Aktienkursindex beeinflussen und damit als unabhängige Variable (Regressoren) in Frage kommen.

Nach eingehender theoretischer und empirischer Prüfung haben sich folgende sechs Zeitreihen am besten bewährt:

1. *Zuwachsrates des Mindestreservesolls auf Inlandsverbindlichkeiten*
Diese Rate spiegelt die Veränderung der Geldmenge wider, die ihrerseits die Nachfrage nach Kapitalanlagemöglichkeiten (z. B. Aktien) beeinflusst.
2. *Rendite festverzinslicher Wertpapiere*
Sie sollte mit dem Aktienkursindex negativ korreliert sein, weil festverzinsliche Wertpapiere eine Anlagealternative zu Aktien sind.
3. *Geschäftsklimaindex des Ifo-Instituts für Vorprodukte*
Dieser Index kann als vorlaufender Indikator für die Konjunkturschätzung der Unternehmen angesehen werden.
4. *Dollarkurs (in DM ausgedrückt)*
Der Dollarkurs wird z. T. durch den grenzüberschreitenden Kapitalverkehr bestimmt, der seinerseits Ergebnis unterschiedlicher Renditen der in- und ausländischen Kapitalanlagemöglichkeiten ist.
5. *Dow-Jones-Index*
Durch diesen Index wird der Einfluß des amerikanischen Aktienmarktes auf die deutschen Aktienkurse berücksichtigt.
6. *Politisches Umfeld*
Diese Variable, deren Werte auf einer Rating-Skala subjektiv geschätzt werden müssen, soll innen- und außenpolitische Krisen (Regierungswechsel, Afghanistankonflikt usw.), die die Börse zeitweise beeinflussen können, berücksichtigen.

Das Ergebnis der Regression mit diesem Modell ist in Tab. 23 wiedergegeben.

Mit einem Bestimmtheitsmaß von $R^2 = 0,882$ konnten also 88,2 % der Varianz erklärt werden. Der theoretische F-Wert $F_{95\%} = 4$ zeigt, daß alle Regressoren und das Modell insgesamt signifikant sind.

Der Durbin-Watson Koeffizient beträgt 0,90 und deutet daher auf einen versteckten Trend bzw. auf eine Saison hin, die wir bei Anwendung des Box-Jenkins-Verfahrens mit 15 Monaten spezifiziert hatten.

Der mittlere absolute Fehler beträgt $MAA = 1,80$, der $WMQA = 2,27$ und der Theil'sche Ungleichheitskoeffizient $U = 0,96$.

Tabelle 23

Reihe	Koeffizient \hat{b}	$\hat{\sigma}(\hat{b})$	F-Wert
Zuwachsrate Mindestreservesoll	0,5371	0,15397	12,171
Wertpapierrendite	-1,4129	0,29057	23,644
Geschäfts- klimaindex	0,1488	0,02333	40,694
Dollarkurs	-15,8297	1,22222	167,742
Dow-Jones-Index	0,0243	0,00377	41,558
Politisches Umfeld	2,3722	0,32532	53,169
Konstante	122,9441	—	—
$R^2 = 0,882$			$F_{95\%} = 4$
			DW = 0,90

Die Abbildung 25 zeigt die wahren und die prognostizierten Indizes des Aktienkurses.

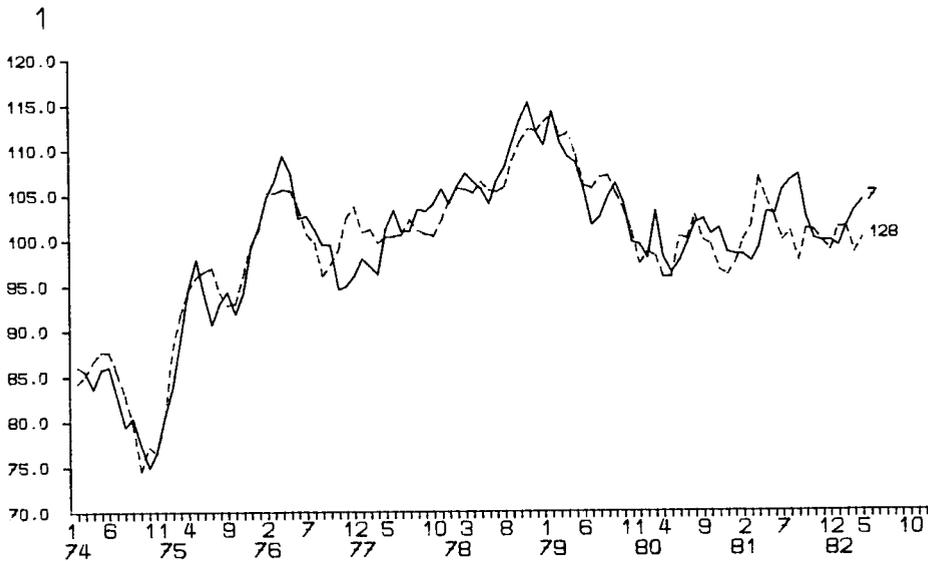


Abbildung 25
Aktienindex und Regressionsprognose (gestrichelt)

IV. Kritische Zusammenfassung

Die Regressionsanalyse zeichnet sich im Hinblick auf Prognosen durch folgende *Vorzüge* aus:

- Sie ist das leistungsfähigste Instrument zur Messung von kausalen Beziehungen zwischen Variablen und erlaubt dadurch, die Prognose einer Zeitreihe auf die Prognose von anderen – evtl. leichter zu prognostizierenden – Zeitreihen zurückzuführen.
- Die theoretischen und statistischen Implikationen der Methode sind weitgehend geklärt, so daß sie ein solides theoretisches Fundament besitzt.
- Die Probleme, die durch Verletzung der Grundannahmen auftreten (Autokorrelation, Multikollinearität), sind durch zusätzliche Überlegungen prinzipiell lösbar oder für die Prognose nicht so gravierend.
- Für die Prognosewerte können Konfidenzintervalle angegeben werden.
- Es liegen benutzerfreundliche Software-Systeme an den meisten Hochschulen und auch zunehmend in der betrieblichen Praxis vor.

Diesen Vorzügen stehen im wesentlichen folgende *Nachteile* bzw. Anwendungsprobleme gegenüber:

- Die Methode stellt hohe Anforderungen an das mathematische und statistische Verständnis der Benutzer und ist darüber hinaus rechenzeit- und speicheraufwendig.
- Die Auswahl des richtig spezifizierten Modells, insbesondere der Funktionsform und der exogenen Variablen, bedarf subtiler Überlegungen des Fachwissenschaftlers. Eine mechanistische Anwendung des Instrumentariums kann zu erheblichen Prognosefehlern führen.
- Eine Überlegenheit der Regressionsanalyse gegenüber den Methoden der Zeitreihenanalyse ist nur gegeben, wenn die exogenen Variablen der abhängigen Variablen vorauseilen, da dann ihre Werte für den Prognosezeitraum von y_t schon bekannt sind, oder wenn sie leichter zu prognostizieren sind als die abhängige Variable selbst.

Da diese beiden letzten Bedingungen in der Praxis nicht immer zutreffen, empfiehlt es sich, die Regressionsanalyse mit autoregressiven Methoden (Box-Jenkins-Verfahren) zu kombinieren, um die Prognoseergebnisse zu verbessern (Hansmann 1980 c).

H. Kriterien zur Auswahl geeigneter Prognoseverfahren

Nach Darstellung und kritischer Würdigung der wichtigsten Prognoseverfahren bleibt noch die Frage zu beantworten, nach welchen Kriterien das geeignete Prognoseverfahren für ein konkretes Prognoseproblem zu bestimmen ist.

In der Literatur existieren bereits verschiedene Kriterienkataloge zur Auswahl von Prognoseverfahren, die sich z. T. überschneiden (Scheer 1980, S. 41 ff., Hüttner 1982, S. 158 ff. und Makridakis, Reschke, Wheelwright 1980, S. 223 ff.).

Im folgenden wird versucht, auf diesen Kriterienkatalogen aufbauend, Hinweise zur Auswahl des geeigneten Prognoseverfahrens aus der Teilmenge der in diesem Buch beschriebenen Verfahren zu geben. Dabei werden wir uns auf die folgenden fünf wichtigen Kriterien beschränken:

1. Prognosegenauigkeit
2. Prognosekosten
3. Komplexität und Benutzerfreundlichkeit des Verfahrens
4. Prognosezeitraum
5. Prognosedatenbasis
 - a) Quantitative Daten oder subjektive Einschätzungen
 - b) Eine Zeitreihe oder zusätzliche erklärende Reihen
 - c) Umfang der Zeitreihen
 - d) Zeitreihenmuster (Trend, Saison).

Die drei ersten Kriterien – Genauigkeit, Kosten und Komplexität – sind untereinander abhängig, so daß sie gemeinsam erörtert werden sollen. Es liegt auf der Hand, daß die Prognosekosten mit zunehmender Komplexität der Verfahren und mit höheren Ansprüchen an die Prognosegenauigkeit steigen. Im allgemeinen verbessert sich auch die Genauigkeit der Prognose bei Anwendung komplizierterer Verfahren, doch kann man auch Gegenbeispiele finden.

Bei der Auswahl des geeigneten Prognoseverfahrens muß also sorgfältig abgewogen werden, ob die zusätzlichen Kosten eines komplizierteren Prognoseverfahrens durch erhöhte Prognosegenauigkeit gerechtfertigt sind. Das hängt einmal von der Bedeutung der Prognose für die Entscheidungen des Unternehmens ab, zum anderen aber auch von der Verfügbarkeit qualifizierten Personals, das die komplizierteren Prognosemodelle durchschaut und richtig einzusetzen versteht.

Die *Prognosegenauigkeit* mißt man mit den statistischen Fehlermaßen, die in Kapitel A. IV. (S. 14 ff.) ausführlich dargestellt wurden und daher hier nicht noch einmal behandelt werden müssen.

Die *Prognosekosten* umfassen vor allem die fixen Kosten für Entwicklung (bzw. Beschaffung) und Implementierung der Software sowie die variablen Kosten für Rechenzeit, Datenbeschaffung und -aufbereitung.

Die *Komplexität* eines Prognoseverfahrens ist ein Kriterium, das nicht nur unter Kostengesichtspunkten betrachtet werden darf. Je höher der mathematisch-statistische Aufwand eines Verfahrens ist, um so weniger sind die Prognoseergebnisse für Nicht-Experten nachprüfbar und um so zurückhaltender reagieren die Entscheidungsträger, die sich auf die Prognoseergebnisse stützen müssen. Es kann also durchaus günstiger sein, ein weniger komplexes Prognoseverfahren zu wählen und damit eine geringere Prognosegenauigkeit in Kauf zu nehmen, wenn auf diese Weise der Widerstand von Entscheidungsinstanzen gegen quantitative Methoden eher abgebaut werden kann.

Spielen die Kriterien Kosten, Komplexität und Genauigkeit für ein konkretes Prognoseproblem keine dominierende Rolle oder ist aufgrund dieser Kriterien noch keine Entscheidung für ein bestimmtes Verfahren möglich, dienen die Kriterien *Prognosezeitraum* und *Prognosedatenbasis* als Anhaltspunkte für die endgültige Auswahl eines Verfahrens.

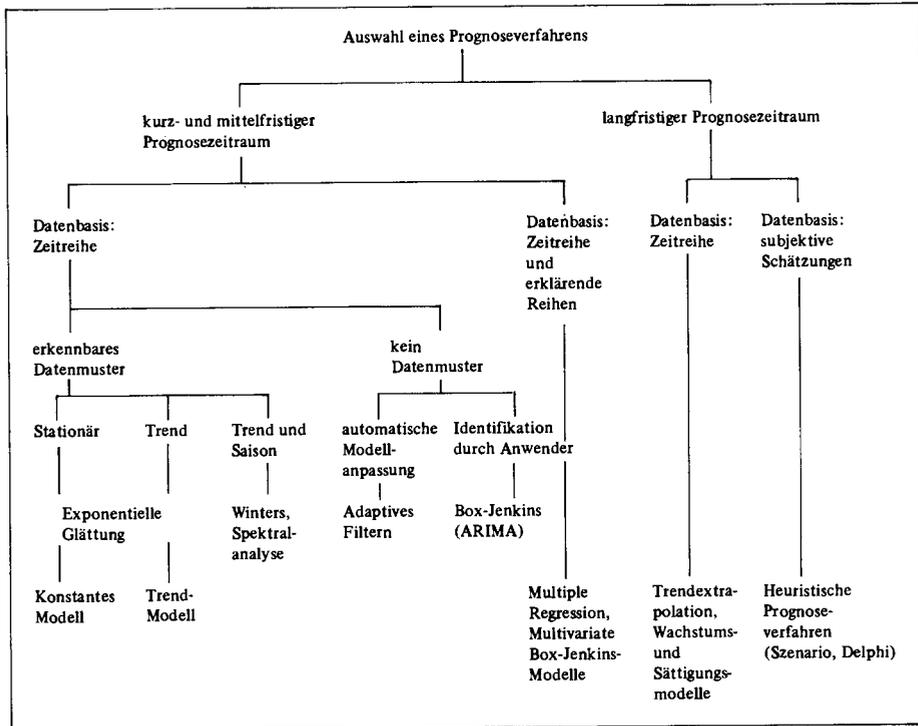
In Anlehnung an Scheer (1980, S. 43) wird im folgenden ein Entscheidungsbaum für die Verfahrensauswahl entwickelt, der es erlaubt, die in diesem Buch verwendeten Prognoseverfahren nach den Eigenschaften der Datenbasis und des Prognosezeitraums so zu klassifizieren, daß das Auswahlproblem computergestützt im Dialog gelöst werden kann (vgl. Abbildung S. 143).

Bei der Anwendung des Entscheidungsbaums ist jedoch zu beachten, daß die Prognoseverfahren nicht ganz überschneidungsfrei zu klassifizieren sind. So kann natürlich auch bei einer Zeitreihe mit erkennbarem Datenmuster z. B. das Box-Jenkins-Verfahren angewendet werden.

Im Interesse der Übersichtlichkeit gibt der Entscheidungsbaum nur die typischen Zuordnungen wieder und sollte daher als eine erste und noch grobe Orientierungsmöglichkeit aufgefaßt werden, die zusammen mit den Kriterien Genauigkeit, Kosten und Komplexität für die Auswahl eines geeigneten Prognoseverfahrens eine Entscheidungsgrundlage bietet.

Mit der Erörterung der Auswahlkriterien für ein Prognoseverfahren sind nunmehr alle wesentlichen Bausteine für Prognosen vorhanden. Der nachfolgende Übungsteil soll dazu dienen, die Kenntnis der Verfahren und ihrer theoretischen Grundlagen zu überprüfen und darüber hinaus die Fertigkeit zur Anwendung auf konkrete Prognoseprobleme einzuüben.

Entscheidungsbaum für die Verfahrensauswahl



Aufgaben und Lösungen

Zu Kapitel A.: Grundlagen der Prognoserechnung

Aufgabe 1:

Definieren und charakterisieren Sie den Begriff ‚Prognose‘.

Lösung:

Eine Prognose ist eine Aussage über ein oder mehrere zukünftige Ereignisse, die auf Beobachtungen der Vergangenheit und einer Theorie beruht. Eine Prognose ist durch das Bemühen charakterisiert, mit Hilfe von Methoden Gesetzmäßigkeiten in der Entwicklung vergangener Ereignisse zu erkennen, die einen Schluß auf zukünftige Ereignisse ermöglichen.

Aufgabe 2:

Verbinden Sie inhaltlich die Begriffe ‚Zeitreihe‘ und ‚Zeitstabilitätshypothese‘.

Lösung:

Eine Zeitreihe ist eine Menge von Beobachtungswerten, die im gleichen zeitlichen Abstand aufeinander folgen. Die Zeitstabilitätshypothese fordert, daß in der Entwicklung der Zeitreihenwerte eine zeitstabile Gesetzmäßigkeit (Trend, Saison) zum Ausdruck kommt, von der angenommen werden kann, daß sie auch in der Zukunft gültig sein wird.

Aufgabe 3:

Welche Überlegungen sind hinsichtlich der Auswahl eines Prognosemodells für einen konkreten Prognosegegenstand anzustellen?

Lösung:

Die Auswahl eines Prognosemodells ist formal an der Beschaffenheit der Datenbasis orientiert, mit der die Entwicklung des Prognosegegenstandes im Zeitablauf beschrieben ist. Prognosegegenstand, Art und Qualität (Art: quantitativ, qualitativ; Qualität:

exakte Messung, Schätzung) der Daten sowie deren Fristigkeit (z. B. Jahres-, Monats-, Tageswerte) determinieren die Prognosemodellart im Sinne einer Vorauswahl. Eine feinere Analyse des Datenmaterials in bezug auf erkennbare Gesetzmäßigkeiten führt zur Identifikation spezieller Modelltypen, insbesondere für den Bereich des kurz- und mittelfristen Prognosezeitraumes.

Da für einen Prognosegegenstand mehrere Prognosemodelle geeignet sein können, kann eine endgültige Entscheidung abhängig gemacht werden von der Prognosegenauigkeit, den Prognosekosten sowie der Komplexität und Benutzerfreundlichkeit des Verfahrens.

Aufgabe 4:

Zur Beurteilung der Prognosequalität werden statistische Fehlermaße herangezogen. Welche Eigenschaften sollte ein Fehlermaß erfüllen?

Lösung:

Zwei wesentliche Eigenschaften sind zu unterscheiden:

- a) Die Dimension des Maßes sollte artverwandt mit der Dimension der Zeitreihenwerte sein (z. B. dimensionsgleich, prozentuale Abweichungen).
- b) Der Wert des Maßes sollte im Sinne einer Normierung zwischen guten und schlechten Prognosen diskriminieren können.

Zu Kapitel B.: Heuristische Prognoseverfahren

Aufgabe 5:

Skizzieren Sie inhaltlich die Szenario-Technik.

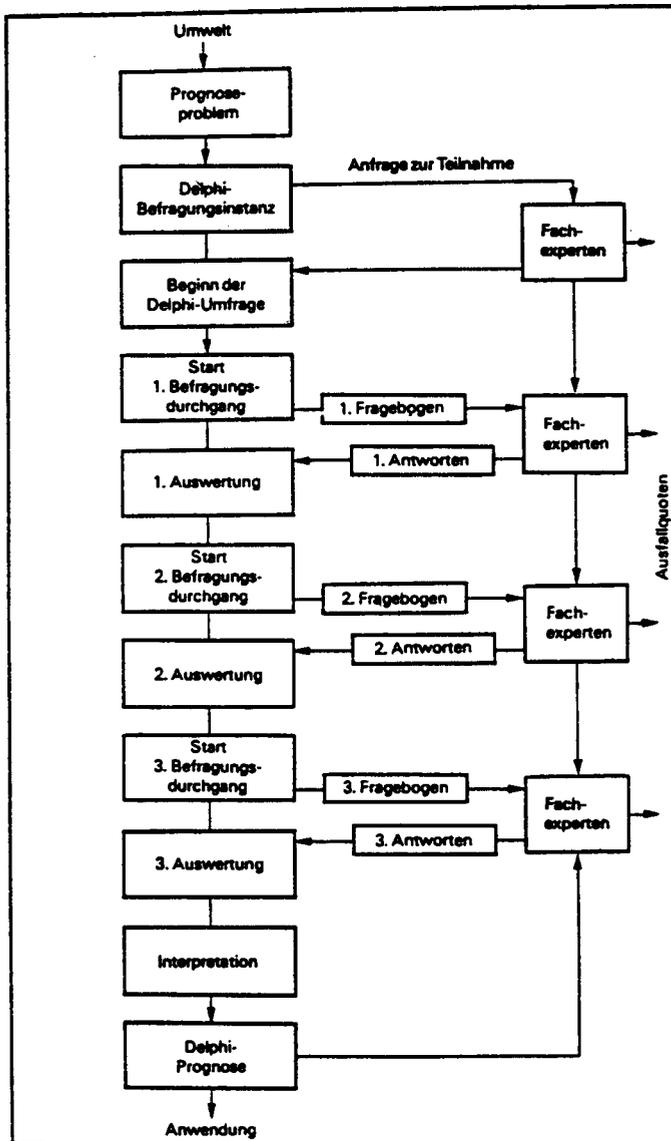
Lösung:

Die Szenario-Technik orientiert sich an der Fragestellung: „Wie könnte oder wie sollte die Zukunft aussehen?“ Als Ziel folgt daraus die Entwicklung eines möglichst umfassenden und widerspruchsfreien Bildes denk- und wünschbarer zukünftiger Ereignisse. Es ist charakteristisch für Szenarien, daß die zugrunde liegende Theorie nicht ausschließlich auf Erfahrungen und Strukturen der Vergangenheit beruht.

Aufgabe 6:

Entwickeln Sie ein Ablaufdiagramm für eine Delphi-Umfrage.

Lösung:



Quelle: Hüttner, Manfred: Markt- und Absatzprognosen. Stuttgart-Berlin-Köln-Mainz 1982, S. 30.

Aufgabe 7:

Eine Gruppe von acht problemvertrauten Personen hat in den drei Runden einer Delphi-Umfrage folgende Werte für den durchschnittlichen Sollzins privater Girokonten für das Jahr 1983 geschätzt:

1. Runde

8,0 % 10,5 % 11,0 % 12,5 % 14,0 % 15,8 % 17,0 % 18,0 %

2. Runde

8,0 % 11,0 % 11,5 % 11,5 % 13,0 % 14,5 % 16,0 % 17,0 %

3. Runde

9,0 % 11,5 % 11,5 % 12,0 % 12,5 % 14,0 % 15,5 % 16,4 %

- Errechnen Sie die statistischen Kennzahlen für die einzelnen Runden.
- Mit welchen Argumenten könnten die Schätzwerte außerhalb der Quartilspanne begründet werden?
- Interpretieren Sie die Entwicklung der statistischen Kennzahlen über alle drei Runden und geben Sie das Ergebnis der Gruppenprognose an.

Lösung:

a) Auswertung 1. Runde

Z	Q _u	Q _o	Q _m	\bar{x}	s	R	
13,25	10,5	17	6,5	13,35	3,47	10	[%]

Auswertung 2. Runde

Z	Q _u	Q _o	Q _m	\bar{x}	s	R	
12,25	11	16	5	12,81	2,94	9	[%]

Auswertung 3. Runde

Z	Q _u	Q _o	Q _m	\bar{x}	s	R	
12,25	11,5	15,5	4	12,8	2,40	7,4	[%]

- Die Schätzwerte im oberen Quartil beruhen auf pessimistischen Erwartungen bezüglich der volkswirtschaftlichen Entwicklung. Als Gründe wurden eine fortgesetzte restriktive Geldmengenpolitik der Bundesbank infolge einer erwarteten Politik des „knappen Geldes“ der Amerikanischen Notenbank angegeben.

Die Schätzwerte im unteren Quartil beruhen auf optimistischen Erwartungen bezüglich der volkswirtschaftlichen Entwicklung. Als Gründe wurden erwartete Verände-

rungen wirtschaftspolitischer Rahmenbedingungen (Diskontsatz- und Lombardsatzsenkungen) angegeben.

- c) Die statistischen Kennzahlen Q_m (Quartilspanne), s (Standardabweichung) und R (Range) signalisieren insgesamt eine Konvergenz der Meinungen. Die „Experten-Gruppe“ erwartet einen durchschnittlichen Sollzinssatz (Median Z) für das Jahr 1983 von 12,25 %, also einen Rückgang um 2,75 Prozentpunkte von aktuell 15 %.

Zu Kapitel C.: Prognosen auf der Grundlagen der exponentiellen Glättung

Aufgabe 8:

Welche Gemeinsamkeiten und Unterschiede bestehen zwischen dem Verfahren der gleitenden Durchschnitte und den methodischen Varianten der exponentiellen Glättung?

Lösung:

Die Methode der gleitenden Durchschnitte und die Varianten der exponentiellen Glättung haben gemeinsam, daß die Beobachtungswerte einer Zeitreihe durch die Prognosegleichung gewichtet werden. Mit Ausnahme der Variante dynamischer α -Anpassung wird bei den Methoden die Gewichtung exogen vorgegeben. Unterschiedlich ist die Höhe der Gewichtung. Während die Methode der gleitenden Durchschnitte alle Zeitreihenwerte gleich mit dem Faktor $\frac{1}{n}$ gewichtet, berücksichtigen die Varianten der exponentiellen Glättung die Aktualität eines Zeitreihenwertes durch stärkere Gewichtung. Darüber hinaus geht der Prognosefehler der Gegenwart (e_t) in die folgenden Prognosen ein.

Aufgabe 9:

Welche Eigenschaft einer Zeitreihe ist entscheidend für die Wahl einer methodischen Variante der exponentiellen Glättung?

Lösung:

Entscheidend ist die Gesetzmäßigkeit, die in der zeitlichen Entwicklung der Zeitreihenwerte zum Ausdruck kommt. Für Entwicklungen ohne klaren Trend und Saisonschwankungen ist das konstante Modell angemessen; liegt keine annähernd konstante Entwicklung vor, ist das Trendmodell der exponentiellen Glättung angezeigt, wobei eine Verbesserung der Prognosequalität in der Regel durch die endogene Anpassung des Glättungsparameters α möglich ist.

Aufgabe 10:

Welche Konsequenzen sind mit der Wahl eines Wertes für den Glättungsparameter α verbunden? Verdeutlichen Sie sich die Wirkung alternativer Werte für α (0,2; 0,4; 0,6; 0,8; 1,0) anhand der fiktiven Zeitreihen

t :	1	2	3	4	5	6	7
I : x_t :	4	8	4	4	4	4	4
II : x_t :	2	4	2	4	2	4	2

und dem konstanten Modell der exponentiellen Glättung.

Lösung:

Mit dem Wert von α wird festgelegt, wie stark das Prognosemodell auf Wertänderungen der Zeitreihenelemente reagieren soll.

Große Werte für α führen zu einer schnellen Reaktion des Systems, was bei einer grundlegenden Niveauänderung der Zeitreihenwerte (Strukturbruch) sinnvoll ist.

Kleine Werte für α bewirken eine stärkere Gewichtung der Werte der Vergangenheit, was zu einer verzögerten Reaktion des Systems führt; einmalige „Ausrutscher“ beeinträchtigen die Prognoseergebnisse weniger stark.

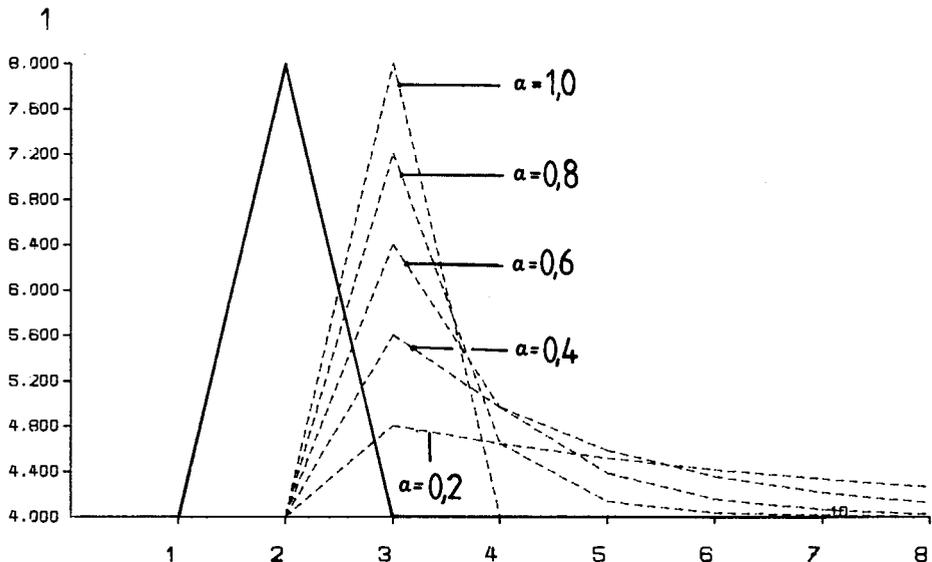


Abbildung 26

Zeitreihe I und alternative Werte für α

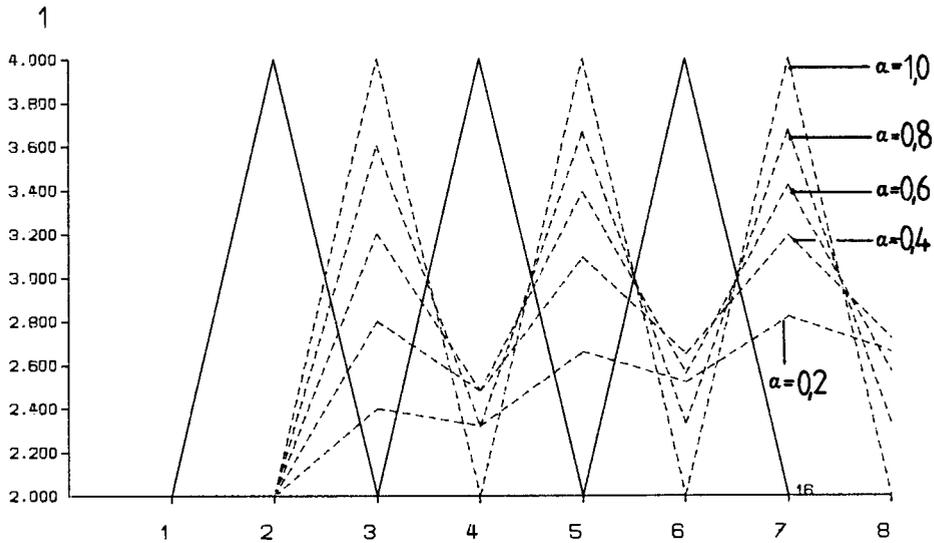


Abbildung 27

Zeitreihe II und alternative Werte für α

Aufgabe 11:

Welche Überlegungen können zur Bestimmung eines vergleichsweise ‚guten‘ Wertes für den Glättungsparameter α führen?

Lösung:

Eine wesentliche Hilfestellung bietet die graphische Darstellung der Zeitreihe. Neben dem Erkennen von Gesetzmäßigkeiten (Trend, Saison) können sowohl singuläre „Ausreißer“ als auch generelle Niveauverschiebungen (Strukturbrüche) optisch identifiziert werden. Dadurch wird einerseits die Wahl des Prognosemodells erleichtert, andererseits ergeben sich im Zusammenhang mit der Kenntnis der Wirkung eines Wertes für α (vgl. Aufgabe 10) Anhaltspunkte für die Wahl eines konkreten Wertes. Weitere Verbesserungen sind durch alternative Rechnungen und Vergleich der Fehlermaße möglich.

Aufgabe 12:

Tabelle 24

I: Flugzeugbestand (Startgewicht bis 20 t) deutscher Fluggesellschaften
(Quelle: Verkehr in Zahlen 1981, S. 97)

1970	375	1975	345
1971	370	1976	320
1972	303	1977	316
1973	335	1978	400
1974	334	1979	378

Tabelle 25

II: Stromverbrauch der Deutschen Bundesbahn in (MRD-kwh/p.a.)
(Quelle: Verkehr in Zahlen 1981, S. 230 f.)

1960	1,743	1970	5,335
1961	1,923	1971	5,466
1962	2,292	1972	5,780
1963	2,689	1973	5,958
1964	2,967	1974	6,111
1965	3,344	1975	5,691
1966	3,527	1976	5,840
1967	3,667	1977	5,837
1968	4,106	1978	6,085
1969	4,884	1979	6,725
		1980	6,935

Für die vorstehenden zwei Zeitreihen (Tabellen 24 und 25) soll jeweils eine Prognose für ein Jahr im voraus erstellt werden.

- a) Stellen Sie die Zeitreihen graphisch dar und entscheiden Sie, welche Variante der exponentiellen Glättung für die einzelnen Zeitreihen zur Anwendung kommen sollte.
- b) Führen Sie die Prognoserechnung für Zeitreihe I mit Hilfe des konstanten Modells der exponentiellen Glättung und einem Wert des Glättungsparameters $\alpha = 0,6$ durch.
- ba) Welchen prozentualen und absoluten Einfluß übt der Schätzwert \hat{x}_1 auf den Prognosewert \hat{x}_{t+1} aus?
- bb) Aufgrund der Modellkonzeption gilt:

$$\hat{x}_{t+m} = \hat{x}_{t+1} \quad (m = 2, 3, \dots)$$

Beurteilen Sie die Sinnhaftigkeit der Prognoseergebnisse für $m \geq 2$.

bc) Beurteilen Sie die unter b) erzielten Ergebnisse mit Hilfe der Fehlermaße

- MAA (mittlere absolute Abweichung)
- WMQA (Wurzel aus mittlerer quadrierter Abweichung)
- U (Theil'scher Ungleichheitskoeffizient).

c) Führen Sie die Prognoserechnung für Zeitreihe II mit Hilfe des Trendmodells der exponentiellen Glättung und dynamischer Anpassung des Glättungsparameters α durch. Wählen Sie als Startwerte:

$$\bar{\alpha}_0 = 0,9 \quad \beta = 0,4 \quad \gamma = 0,05 \quad MA_0 = 0 \quad MAA_0 = 0,4$$

$$e_0 = e_1 = e_2 = 0$$

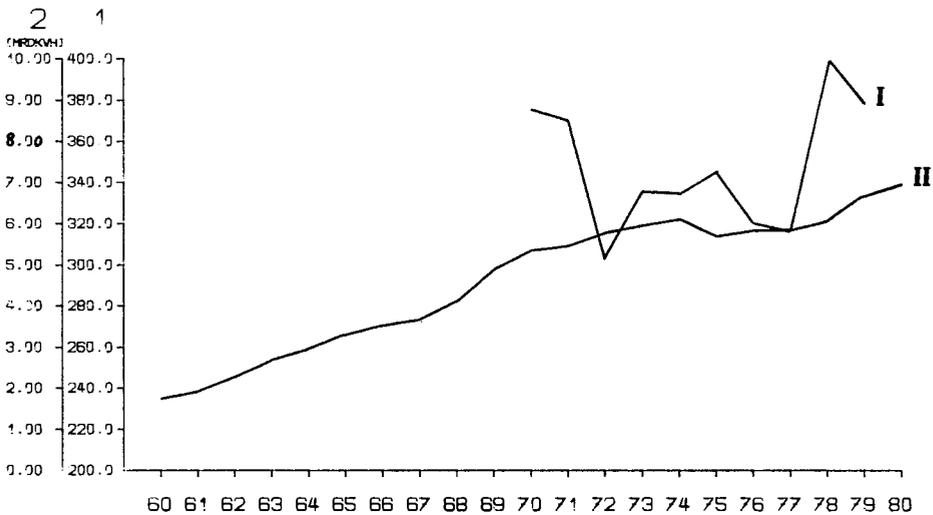


Abbildung 28

Darstellung der Zeitreihen I und II

a) Der Abbildung 28 ist zu entnehmen, daß für Zeitreihe I das konstante Modell der exponentiellen Glättung angemessen ist, da weder ein klarer Trend noch eine Saison erkennbar sind. Demgegenüber läßt Zeitreihe II einen klaren Trend erkennen.

b) *Initialisierung des Verfahrens*

– nach (5) bzw. (7): $\hat{x}_1 = x_1 = 375$

– *Prognose*

$$\text{nach (5) : } \hat{x}_2 = 0,6 \cdot 375 + 0,4 \cdot 375 = 375$$

$$\text{nach (5) : } \hat{x}_3 = 0,6 \cdot 370 + 0,4 \cdot 375 = 372$$

$$\text{nach (5) : } \hat{x}_4 = 0,6 \cdot 303 + 0,4 \cdot 372 = 330,6$$

Die vollständigen Ergebnisse sind in Tabelle 26 wiedergegeben.

Tabelle 26

Jahr	t-1	x_{t-1}	\hat{x}_{t-1}	e_{t-1}
1970	1	375	375,0	0,0
1971	2	370	375,0	- 5,0
1972	3	303	372,0	-69,0
1973	4	335	330,6	4,4
1974	5	334	333,2	0,8
1975	6	345	333,7	11,3
1976	7	320	340,5	-20,5
1977	8	316	328,2	-12,2
1978	9	400	320,9	79,1
1979	10	378	368,4	9,6
1980	11	-	374,1	-

Für das Jahr 1980 ergibt sich ein Prognosewert von $\hat{x}_{1980} = 374,1$ d. h. ein Flugzeugbestand von 374 Maschinen mit einer Tragfähigkeit von 20 t pro Maschine.

ba) nach (7)

$$(1-0,6)^{10} = 0,0001048$$

Der prozentuale Einfluß beträgt 0,0105 %

Der absolute Einfluß beträgt

$$\frac{375 \cdot 0,0105}{100} = 0,04, \text{ d. h.}$$

die Forderung ‚genügend großer Zeitreihenumfang‘ ist praktisch erfüllt.

bb) Mit zunehmendem Prognosehorizont entbehren die Prognosewerte einer inhaltlichen Fundierung, da keine aktuellen Zeitreihenwerte mehr in die Prognose eingehen.

bc) MAA = 23,54
WMQA = 36,29
U = 0,93

Die Werte von MAA und WMQA signalisieren in bezug auf den Durchschnittsbestand von ca. 347 Flugzeugen eine wenig effiziente Prognose. Nach dem

Theil'schen Ungleichheitskoeffizient sind die Prognoseergebnisse jedoch besser als die einer Naivprognose.

c) Bestimmung der Glättungsoperatoren \hat{S}_0^1 und \hat{S}_0^2 (Initialisierung, $t=0$)

$$\hat{b}_0 = 0,1800$$

$$\hat{a}_0 = 1,5630$$

$$\text{nach (25a) : } 1,5630 = 2 \hat{S}_0^1 - \hat{S}_0^2$$

$$\text{nach (25b) : } 0,1800 = \frac{0,9}{0,1} (\hat{S}_0^1 - \hat{S}_0^2)$$

$$\text{aus (25a) und (25b) : } \hat{S}_0^1 = 1,5430 ; \hat{S}_0^2 = 1,5230$$

– Fortschreibung der Glättungsoperatoren, α -Anpassung und Prognose ($t = 1$)

$$\text{nach (19) : } \hat{S}_1^1 = 1,7230$$

$$\text{nach (23) : } \hat{S}_1^2 = 1,7030$$

$$\text{nach (25a) : } \hat{a}_1 = 1,7430$$

$$\text{nach (25b) : } \hat{b}_1 = 0,1800$$

– – α -Anpassung

$$\text{nach (27) : } MA_1 = 0$$

$$\text{nach (28) : } MAA_1 = 0,24$$

$$\text{nach (29) : } \alpha_1 = 0$$

$$\text{nach (30) : } \bar{\alpha}_1 = 0,8550$$

– – Prognose

Da zur Initialisierung des Verfahrens die Zeitreihenwerte $x_1 = 1,7430$ und $x_2 = 1,9230$ verwendet werden, erfolgt die erste Prognose für 1962 (Periode $t = 3$). Dafür müssen \hat{S}_2^1 , \hat{S}_2^2 und $\bar{\alpha}_2$ ermittelt werden.

– Fortschreibung der Glättungsoperatoren, α -Anpassung und Prognose ($t = 2$)

$$\text{nach (19) : } \hat{S}_2^1 = 1,8940$$

$$\text{nach (23) : } \hat{S}_2^2 = 1,8663$$

$$\text{nach (25a) : } \hat{a}_2 = 1,9217$$

$$\text{nach (25b) : } \hat{b}_2 = 0,1633$$

– – α -Anpassung

$$\text{nach (27) : } MA_2 = 0$$

$$\text{nach (28) : } MAA_2 = 0,144$$

$$\begin{aligned} \text{nach (29)} & : \alpha_2 = 0 \\ \text{nach (30)} & : \bar{\alpha}_2 = 0,81225 \end{aligned}$$

-- Prognose

$$\begin{aligned} \text{nach (26)} & : \hat{x}_{2+1} = \hat{x}_3 = 2,0805 \\ \text{Prognosefehler:} & \quad e_3 = 0,2070 \end{aligned}$$

– Fortschreibung der Glättungsoperatoren, α -Anpassung und Prognose ($t = 3$)

$$\begin{aligned} \text{nach (19)} & : \hat{S}_3^1 = 2,2173 \\ \text{nach (23)} & : \hat{S}_3^2 = 2,1514 \\ \text{nach (25a)} & : \hat{a}_3 = 2,2832 \\ \text{nach (25b)} & : \hat{b}_3 = 0,2851 \end{aligned}$$

-- α -Anpassung

$$\begin{aligned} \text{nach (27)} & : MA_3 = 0,0828 \\ \text{nach (28)} & : MAA_3 = 0,1692 \\ \text{nach (29)} & : \alpha_3 = 0,4894 \\ \text{nach (30)} & : \bar{\alpha}_3 = 0,7961 \end{aligned} \quad \begin{array}{l} \text{(Abweichung vom Tabellenwert} \\ \text{(Tabelle 27) aufgrund von} \\ \text{Rundungsfehlern)} \end{array}$$

-- Prognose

$$\begin{aligned} \text{nach (26)} & : \hat{x}_{3+1} = \hat{x}_4 = 2,5683 \\ & \quad e_4 = 0,1208 \end{aligned}$$

Unter Verwendung des jeweils folgenden Zeitreihenwertes x_t und fortlaufender α -Anpassung, führt die jährlich fortschreitende Rechnung zu den Ergebnissen der Tabelle 27.

Tabelle 27

Jahr	t	x_t	\hat{x}_t	e_t	\hat{a}_t	\hat{b}_t	$\bar{\alpha}_t$
1960	1	1,7340	—	—	1,7430	0,1800	0,85
1961	2	1,9230	1,9230	0,0000	1,9217	0,1633	0,81
1962	3	2,2920	2,0850	0,2070	2,2832	0,2851	0,80
1963	4	2,6890	2,5683	0,1208	2,6828	0,3514	0,79
1964	5	2,9670	3,0343	-0,0673	2,9693	0,3040	0,76
1965	6	3,3440	3,2733	0,0707	3,3377	0,3276	0,75
1966	7	3,5270	3,6653	-0,1383	3,5342	0,2394	0,72
1967	8	3,6670	3,7736	-0,1066	3,6729	0,1695	0,71
1968	9	4,1060	3,8424	0,2636	4,0837	0,2998	0,70
1969	10	4,8840	4,3836	0,5004	4,8365	0,5330	0,70
1970	11	5,3695	5,4040	-0,0345	5,3392	0,5224	0,70
1971	12	5,4660	5,8616	-0,3956	5,5008	0,3256	0,68
1972	13	5,7800	5,8264	-0,0464	5,7817	0,2879	0,66
1973	14	5,9580	6,0696	-0,1116	5,9682	0,2272	0,66
1974	15	6,1110	6,1955	-0,0845	6,1201	0,1871	0,66
1975	16	5,6910	6,3073	-0,6163	5,7626	-0,0805	0,67
1976	17	5,8400	5,6820	0,1580	5,8225	-0,0118	0,66
1977	18	5,8370	5,8107	0,0263	5,8340	0,0000	0,65
1978	19	6,0850	5,8341	0,2509	6,0535	0,1047	0,63
1979	20	6,7250	6,1582	0,5668	6,6464	0,3254	0,64
1980	21	6,9350	6,9718	-0,0368	6,9412	0,3163	0,64
1981	22	—	7,2575	—	—	—	—

Damit ergibt sich ein Prognosewert $\hat{x}_{1981} = 7,2575$, d. h. der prognostizierte Stromverbrauch der Deutschen Bundesbahn beträgt für das Jahr 1981 7,2575 Mrd-kWh.

Zu Kapitel D.: Prognosen bei saisonbehafteten Zeitreihen

Aufgabe 13:

Für die Anwendung des Verfahrens von Winters ist die Zeitreihe der monatlichen Bierproduktion in der Bundesrepublik Deutschland (Quelle: Statistisches Bundesamt, Fachserie 14, Reihe 9.2.1) von Januar 1977 bis Dezember 1981 ($t = 1, \dots, 60$) gegeben. Die Reihe umfaßt 60 Werte, die mit der Einheit 1000 Hektoliter (hl) gemessen sind. Es kann eine Saison von 12 Monaten unterstellt werden.

Der letzte Zeitreihenwert lautet: $x_{60} = 7881$, die Glättungsparameter haben folgende Werte:

$$\alpha = 0,5 \quad \beta = 0,1 \quad \gamma = 0,5$$

In der Vergangenheit sind die Saisonfaktoren

$$\hat{s}_{48,0} = 0,7837 \quad \hat{s}_{49,1} = 0,675 \quad \hat{s}_{50,2} = 0,6844$$

sowie der Grundwert $\hat{a}_{60} = 9989,3624$ und der Trendwert $\hat{b}_{60} = 5,754$ geschätzt worden.

Berechnen Sie den Saisonfaktor für Dezember 1981.

Lösung:

Nach Gleichung (37) ergibt sich dieser Saisonfaktor zu

$$\hat{s}_{60,0} = \hat{s}_{48,0} + \gamma \left(\frac{x_{60}}{\hat{a}_{60}} - \hat{s}_{48,0} \right) = 0,7837 + 0,5 \left(\frac{7881}{9989,3624} - 0,7837 \right)$$

$$\hat{s}_{60,0} = 0,7863$$

Aufgabe 14:

Prognostizieren Sie die Bierproduktion für die Monate Januar und Februar 1982.

Lösung:

Zur Prognose wird Gleichung (34) zugrunde gelegt. Da für den Monat Januar 1982 kein effektiver Wert bekannt ist, können Grund- und Trendwert vom Dezember 1981 nicht fortgeschrieben werden. Für die Februar-Prognose beträgt somit der Zeithorizont zwei Monate ($m=2$).

$$\hat{x}_{61} = (\hat{a}_{60} + \hat{b}_{60} \cdot 1) \cdot \hat{s}_{49,1} = (9989,3624 + 5,754) \cdot 0,675$$

$$\hat{x}_{62} = (\hat{a}_{60} + \hat{b}_{60} \cdot 2) \cdot \hat{s}_{50,2} = (9989,3624 + 5,754 \cdot 2) \cdot 0,6844$$

$$\hat{x}_{61} = 6746,7036 \quad \hat{x}_{62} = 6844,5957$$

Aufgabe 15:

Nennen Sie den Haupteinwand gegen das Verfahren von Winters.

Lösung:

Der von Winters benutzte Saisonfaktor ist nicht aktuell, sondern liegt eine Saisonzykluslänge (L Perioden) zurück. Da sich die Zeitreihenstruktur während dieser L Perioden geändert haben kann, droht die Gefahr eines systematischen Prognosefehlers, wenn der Saisonfaktor nicht auf dem neuesten Stand ist. Die Konstruktion des Prognosesystems von Winters erlaubt es nicht, diesen Mangel an Aktualität abzustellen.

Aufgabe 16:

Die stochastischen Prozesse sind für die Spektralanalyse grundlegend.

- a) Definieren Sie einen stochastischen Prozeß.
- b) Verknüpfen Sie diesen Begriff mit dem Begriff ‚Zeitreihe‘.
- c) Gewinnt man durch Übergang vom Zeitbereich (Autokorrelationsfunktion) in den Frequenzbereich (Spektrum) mehr Information aus einer Zeitreihe?

Lösung:

- a) Ein stochastischer Prozeß ist eine Familie von Zufallsvariablen, die auf der Menge der Elementarereignisse und der Menge der Zeitpunkte definiert ist.
- b) Eine Zeitreihe kann als Realisation eines stochastischen Prozesses aufgefaßt werden.
- c) Die Information, die in der Zeitreihe steckt, wird durch diese Transformation nicht verändert. Durch die Darstellung im Frequenzbereich gelingt es jedoch, die Varianz der Zeitreihe additiv den einzelnen Frequenzen zuzuordnen und damit evtl. mehrere sich überlappende Saisonzyklen zu identifizieren.

Zu Kapitel E.: Prognosen mit autoregressiven Methoden

Aufgabe 17:

Wodurch ist ein autoregressiver Prozeß p -ter Ordnung gekennzeichnet?

Lösung:

Ein autoregressiver Prozeß p -ter Ordnung zeichnet sich dadurch aus, daß die partiellen Autokorrelationskoeffizienten b_i ($i = 1, 2, \dots, p$) ungleich Null und b_i ($i > p$) gleich Null sind.

Aufgabe 18:

Leiten Sie die Autokorrelationsfunktion $r(\tau)$ ab.

Lösung:

Gleichung (52) wird mit $x_{t-\tau}$ multipliziert und anschließend der Erwartungswert gebildet. Dies ergibt die Autokovarianz R_τ . Sie wird durch die Varianz R_0 dividiert und als Ergebnis erhält man die Autokorrelationsfunktion r_τ (Gleichung 54).

Aufgabe 19:

Zeigen Sie,

- daß ohne die Beschränkung $-1 \leq b_1 \leq 1$ (vgl. (55)) die Stationarität verletzt wäre,
- daß $r_0 = 1$ gilt.

Lösung:

- Gleichung (55) wird mit $b_1 = 2$ fortgeschrieben. Man erkennt, daß die Reihe explodiert.

t	0	1	2	3	4
x_t	1	2	4	8	16

$$b) r_0 = \frac{E[(x_t - E[x_t]) \cdot (x_t - E[x_t])]}{\sqrt{E[(x_t - E[x_t])^2] \cdot E[(x_t - E[x_t])^2]}} = \frac{E[(x_t - E[x_t])^2]}{E[(x_t - E[x_t])^2]} = 1$$

Aufgabe 20:

Bestimmen Sie die partiellen Autokorrelationskoeffizienten \hat{b}_1, \hat{b}_2 (mit $p = 2$) des folgenden Zeitreihenausschnitts

t	1	2	3	4	5
x_t	21,019	20,125	19,901	19,677	19,007

aus den Yule-Walker-Gleichungen (59), wenn die Varianz der Reihe $\sigma^2 = 2,627$ und der Mittelwert $\mu = 18,747$ beträgt.

Lösung:

Für $p = 1$ ergibt sich aus (59)

$$\hat{b}_1 = r_1 = \frac{\frac{1}{T} \sum_{t=2}^5 (x_{t-1} - \mu)(x_t - \mu)}{\sigma^2} = \frac{1,20662}{2,627}$$

also $\hat{b}_1 = 0,4593$.

Setzen wir $p = 2$, so folgt

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

Zunächst errechnen wir

$$r_2 = \frac{\frac{1}{T} \sum_{t=3}^5 (x_{t-2} - \mu)(x_t - \mu)}{\sigma^2} = \frac{0,8404}{2,627} = 0,3199.$$

Daraus folgt für \hat{b}_2 :

$$\hat{b}_2 = \frac{0,3199 - 0,4593^2}{1 - 0,4593^2} = 0,138.$$

Aufgabe 21:

- Formulieren Sie die Bestimmungsgleichung für ein ARMA(2,2)-Modell.
- Bilden Sie die Anfangsschätzwerte für die in Aufgabe 20 genannten Daten (3 Iterationen).
- Prognostizieren Sie den Zeitreihenwert \hat{x}_6 , wenn $u_5 = -0,2169$ und $u_4 = -0,0685$ gegeben ist.

Lösung:

a) Das ARMA(2,2)-Modell lautet

$$x_t = u_t + \sum_{i=1}^2 b_i x_{t-i} - \sum_{i=1}^2 a_i u_{t-i}.$$

b) Die Anfangsschätzwerte für den AR-Teil ergeben sich aus den Yule-Walker-Gleichungen (59) für $p = 2$: $r_1 = b_1 + b_2$ und $r_2 = b_1 r_1 + b_2$. In Aufgabe 20 sind r_1 und r_2 errechnet worden, so daß sich folgende Werte für b_1 und b_2 ergeben:

$$b_1 = 0,3959$$

$$b_2 = 0,138$$

Die für die Iterationen erforderliche Varianz R_0 und die Kovarianzen R_1 und R_2 sind ebenfalls aus Aufgabe 20 bekannt¹

$$R_0 = \sigma^2 = 2,627$$

$$R_1 = r_1 \sigma^2 = 1,2066$$

$$R_2 = r_2 \sigma^2 = 0,8404 .$$

Die erste Iteration mit $a_1 = a_2 = 0$ führt zu (vgl. (84a)):

$$\sigma_u^2 = \frac{R_0}{1 + a_1^2 + a_2^2} = 2,627$$

$$a_2 = -\frac{R_2}{\sigma_u^2} = -\frac{0,8404}{2,627} = -0,3199$$

$$a_1 = -\frac{R_1}{\sigma_u^2} + a_1 a_2 = -\frac{1,2066}{2,627} = -0,4593$$

Die zweite Iteration führt zu:

$$\sigma_u^2 = \frac{2,627}{1 + 0,211 + 0,1023} = 2,000$$

$$a_2 = -\frac{0,8404}{2,000} = -0,4202$$

$$a_1 = -\frac{1,2066}{2,000} + (-0,4593)(-0,3199) = -0,4564$$

Nach der dritten Iteration ergibt sich schließlich

$$\sigma_u^2 = \frac{2,627}{1,385} = 1,897$$

¹ Da sie in dieser Aufgabe als Schätzgröße aufgefaßt werden, müßten sie eigentlich ein „^“ tragen. Wegen der Kontinuität zu Aufgabe 20 sei jedoch darauf verzichtet.

$$a_2 = -\frac{0,8404}{1,897} = -0,4430$$

$$a_1 = -\frac{1,2066}{1,897} + (-0,4564)(-0,4202) = -0,4443$$

c) Wir definieren den Zeitpunkt 6 als t+1 und erhalten

$$\begin{aligned}\hat{x}_{t+1} &= 0,3959x_t + 0,138x_{t-1} + 0,4443u_t + 0,443u_{t-1} \\ \hat{x}_6 &= 4,45\end{aligned}$$

Zu Kapitel E. III.: Prognose mit dem adaptiven Filtern

Aufgabe 22:

- Stellen Sie die Prognosegleichung für ein AR(2)-Modell auf.
- Leiten Sie, ausgehend von der unter (a) aufgestellten Prognosegleichung, die Vorschrift (Gleichung) zur dynamischen Anpassung der Modellparameter ab.
- Welche Bedeutung hat, in der unter (b) ermittelten Vorschrift, die Lernkonstante K und in welchen Grenzen darf K variiert werden, ohne die Konvergenz des Verfahrens zu gefährden?

Lösung:

a) Prognosegleichung:

$$x_t = b_1x_{t-1} + b_2x_{t-2} + u_t$$

b) Ableiten der Vorschrift für die dynamische Anpassung der Modellparameter:

$$\begin{aligned}u_t &= x_t - (b_1x_{t-1} + b_2x_{t-2}) \\ u_t^2 &= [x_t - b_1x_{t-1} - b_2x_{t-2}]^2 \\ \frac{\partial u_t^2}{\partial b_1} &= 2u_t(-x_{t-1}) \\ \frac{\partial u_t^2}{\partial b_2} &= 2u_t(-x_{t-2})\end{aligned}$$

oder allgemein

$$\frac{\partial u_t^2}{\partial b_i} = 2u_t(-x_{t-i}) \quad (i = 1, 2)$$

Beachtet man das Minuszeichen des negativen Gradienten und führt die Lernkonstante K ein, so ergeben sich die neuen Parameterwerte aus den alten durch Addition des Produktes aus negativem Gradient und Lernkonstante.

$$b_{it} = b_{it-1} + 2Ku_t x_{t-i} \quad (i = 1, 2)$$

- c) Beginnend mit beliebigen Ausgangswerten der Modellparameter stellt man fest, in welcher Richtung des Parameterraumes der quadratische Prognosefehler am meisten verringert wird (negativer Gradient) und macht einen Schritt in diese Richtung. Die Schrittweite entlang des negativen Gradienten wird durch den Wert der Lernkonstanten K bestimmt.

Die Konvergenz des Verfahrens ist für ein AR(2)-Modell nicht gefährdet, wenn man folgende Restriktion beachtet

$$0 < K < \frac{1}{2}.$$

Aufgabe 23:

Tabelle 28 gibt die Zeitreihenwerte der Erteilungen der allgemeinen Fahrerlaubnis in Führerscheinklasse 1 (Quelle: Bundesminister für Verkehr, Verkehr in Zahlen 1981, Bonn 1981, S. 122) wieder.

Tabelle 28

Jahr	t	x_t (in Mio)	Jahr	t	x_t (in Mio)
1975	1	1,701	1978	4	1,964
1976	2	1,857	1979	5	2,073
1977	3	1,866	1980	6	2,110

- a) Passen Sie die Modellparameter für ein AR(2)-Modell über 2 *Iterationen* dynamisch an. Gehen Sie dabei von $b_{1t-1} = 0,25$, $b_{2t-2} = 0,25$ und $K = 0,25$ aus.
- b) Prognostizieren Sie mit Hilfe der unter (a) gefundenen „optimalen“ Modellparameter die Anzahl der Erteilungen der allgemeinen Fahrerlaubnis in Führerscheinklasse 1 für das Jahr 1981.

Lösung:

a) Wir verwenden die Normalisierungskonstante

$$NK = \sqrt{\sum_{i=1}^2 x_{t-i}^2}$$

Tabelle 29 zeigt den Verlauf der dynamischen Anpassung der Modellparameter.

Tabelle 29

Iteration	t	b _{1t-1}	b _{2t-2}	x _t	\hat{x}_t	x _{t-1} *	x _{t-2} *	u _t	u _t *
1	1	—	—	1,701	—	—	—	—	—
	2	—	—	1,857	—	—	—	—	—
	3	0,250	0,250	1,866	0,890	0,737	0,675	0,976	0,388
	4	0,393	0,381	1,964	1,444	0,709	0,705	0,523	0,199
	5	0,464	0,451	2,073	1,753	0,725	0,689	0,320	0,188
	6	0,532	0,516	2,110	2,116	0,726	0,668	-0,006	-0,002
2	1	—	—	1,701	—	—	—	—	—
	2	—	—	1,857	—	—	—	—	—
	3	0,531	0,515	1,866	1,862	0,737	0,675	0,004	0,002
	4	0,532	0,516	1,964	1,951	0,709	0,705	0,013	0,005
	5	0,534	0,518	2,073	2,015	0,725	0,689	0,058	0,021
	6	0,542	0,525	2,110	2,155	0,726	0,668	-0,045	-0,016

b) Wir setzen die unter (a) ermittelten Modellparameter in die, für ein AR(2)-Modell modifizierte, Gleichung (85) ein und erhalten

$$\hat{x}_t = 0,542x_{t-1} + 0,525x_{t-2}$$

als Prognosegleichung.

Der Prognosewert der Zeitreihe für 1981 ergibt sich durch Einsetzen der Zeitreihenwerte der Jahre 1980 (x_{t-1}) und 1979 (x_{t-2}).

$$\hat{x}_{1981} = 0,542 \cdot 2,110 + 0,525 \cdot 2,073$$

$$\hat{x}_{1981} = 2,232$$

Zu Kapitel F.: Prognosen mit Wachstums- und Sättigungsmodellen

Aufgabe 24:

Eine genaue Untersuchung hat ergeben, daß die Preise für 1/2 kg Heringe ab 1974 einem *Trend* in Form der Exponentialfunktion folgen. Tabelle 30 zeigt diese Entwicklung für die Jahre 1974 bis 1978.

Tabelle 30

Jahr	t	x_t
1974	1	0,40
1975	2	0,70
1976	3	1,00
1977	4	1,50
1978	5	2,30

- a) Die Prognosegleichung lautet $\hat{x}_t = \hat{A} \cdot \hat{B}^t$.
 Schätzen Sie die Parameter der Prognosegleichung aus der obigen Zeitreihe.
- b) Prognostizieren Sie die Preise für die Jahre 1974 bis 1982.
- c) Berechnen Sie für den Zeitraum 1974 bis 1978 die mittlere quadratische Abweichung (MQA).

Lösung:

- a) Um die Parameter A und B der obigen Prognosegleichung mit Hilfe der KQ-Methode schätzen zu können, muß die Gleichung vorher linearisiert werden.

$$x_t = A \cdot B^t$$

$$\lg x_t = \lg A + t \cdot \lg B$$

Die Normalgleichungen (102a) und (102b) für diese Aufgabe lauten

$$(I) \quad T \cdot \lg \hat{A} + \lg \hat{B} \sum_{t=1}^5 t = \sum_{t=1}^5 \lg x_t$$

$$(II) \quad \lg \hat{A} \sum_{t=1}^5 t + \lg \hat{B} \sum_{t=1}^5 t^2 = \sum_{t=1}^5 \lg x_t \cdot t$$

Wir legen folgende Arbeitstabelle an:

Tabelle 31

Jahr	t	t ²	x _t	lg x _t	t · lg x _t
1974	1	1	0,40	-0,398	-0,398
1975	2	4	0,70	-0,155	-0,310
1976	3	9	1,00	0,000	0,000
1977	4	16	1,50	0,176	0,704
1978	5	25	2,30	0,362	1,810
Σ	15	55	5,90	-0,015	1,806

Durch Einsetzen in obige Normalgleichungen erhalten wir

$$(I') \quad 5 \lg \hat{A} + 15 \lg \hat{B} = -0,015$$

$$(II') \quad 15 \lg \hat{A} + 55 \lg \hat{B} = 1,806$$

$$(III=3I' - II') \quad -10 \lg \hat{B} = -1,851$$

$$\lg \hat{B} = 0,1851$$

$$\hat{B} = 10^{0,1851}$$

$$\hat{B} = 1,531$$

$\lg \hat{B} = 0,1851$ in I' eingesetzt.

$$5 \lg \hat{A} = -0,015 - 15 \cdot 0,1851$$

$$5 \lg \hat{A} = -2,792$$

$$\lg \hat{A} = -0,558$$

$$\hat{A} = 10^{-0,558}$$

$$\hat{A} = 0,277$$

b) Nach Einsetzen der geschätzten Parameter in die Prognosegleichung erhalten wird

$$\hat{x}_t = 0,277 \cdot 1,531^t.$$

Die Prognose- und Zeitreihenwerte sind in Tabelle 32 gegenübergestellt.

Tabelle 32

Jahr	t	x _t	\hat{x}_t
1974	1	0,40	0,42
1975	2	0,70	0,65
1976	3	1,00	0,99
1977	4	1,50	1,52
1978	5	2,30	2,33
1979	6	—	3,57
1980	7	—	5,46
1981	8	—	8,36
1982	9	—	12,80

$$\begin{aligned}
 \text{c) } \quad \text{MQA} &= \frac{1}{5} \sum_{t=1}^5 (x_t - \hat{x}_t)^2 \\
 &= 0,0009
 \end{aligned}$$

Aufgabe 25:

Ein großer Mineralölkonzern möchte von Ihnen, um den zukünftigen Absatz von Benzin, Super und Diesel besser planen zu können, den Pkw-Bestand je 1000 Erwachsene Wohnbevölkerung prognostizieren lassen. Aufgrund der in den USA gemachten Erfahrungen sollen Sie von einem Sättigungsniveau von $S = 700$ ausgehen.

- Schätzen Sie die Parameter der *logistischen Funktion* unter der Annahme, daß nur der Zeitraum von 1970 bis 1980 eine für die Zukunft relevante Entwicklung der Zeitreihe (Tabelle 21) beinhaltet.
- Stellen Sie die Prognosegleichung (vgl. (115)) auf und prognostizieren Sie den Pkw-Bestand je 1000 Erwachsene Wohnbevölkerung für das Jahr 2000.

Lösung:

- Weil das Sättigungsniveau exogen mit $S = 700$ gegeben ist, vereinfacht sich unser Schätzproblem. Wir wenden die KQ-Methode auf Gleichung (119) an. Zur Ermittlung der Normalgleichungen (vgl. (102a) und (102b)) legen wir eine Arbeitstabelle an.

Tabelle 33

Jahr	t	t ²	x _t	$\ln\left(\frac{S}{x_t} - 1\right)$	$t \cdot \ln\left(\frac{S}{x_t} - 1\right)$
1970	1	1	323	0,155	0,155
1971	2	4	345	0,029	0,058
1972	3	9	361	-0,063	-0,189
1973	4	16	373	-0,132	-0,528
1974	5	25	380	-0,172	-0,860
1975	6	36	397	-0,270	-1,620
1976	7	49	418	-0,394	-2,758
1977	8	64	442	-0,538	-4,304
1978	9	81	466	-0,689	-6,201
1979	10	100	482	-0,793	-7,930
1980	11	121	490	-0,847	-9,317
Σ	66	506		-3,714	-33,494

Die Normalgleichungen lauten:

$$(I) \quad 11 \cdot (-\hat{C}) - \hat{a}S \cdot 66 = -3,714$$

$$(II) \quad 66 \cdot (-\hat{C}) - \hat{a}S \cdot 506 = -33,494$$

$$(III = 6 I - II) \quad 110 \hat{a}S = 11,210$$

$$\hat{a}S = 0,1019$$

$$\hat{a} = 0,1019/700$$

$$\hat{a} = 0,00015$$

$\hat{a} \cdot \hat{S} = 0,1019$ in (I) eingesetzt:

$$(I) \quad -11 \hat{C} = -3,714 + 66 \cdot 0,1019$$

$$-11 \hat{C} = 3,0114$$

$$\hat{C} = -0,274$$

b) Nach Gleichung (115) erhalten wir als Prognosegleichung

$$\hat{x}_t = \frac{700}{1 + e^{-0,1019t + 0,274}} \cdot$$

Prognose für das Jahr 2000:

$$\hat{x}_{t=31} = \frac{700}{1 + e^{-2,8849}} = 662,97$$

Aufgabe 26:

Die in Tabelle 17 (vgl. Gliederungspunkt F. III. 3.) wiedergegebenen Prognose- und Zeitreihenwerte zeigen die tatsächliche (bis 1980) und die prognostizierte (bis 1985) Entwicklung des prozentualen Anteils der Privathaushalte mit Farbfernsehgerät an allen Privathaushalten. Prüfen Sie, ob eine Verringerung des Sättigungsniveaus von $S = 90\%$ auf $S = 80\%$ zu einer besseren Prognose führt (logistische Funktion).

a) Nach Anwendung der KQ-Methode auf Gleichung (119) erhalten wir folgende Regressionsgerade

$$\ln\left(\frac{S}{x_t} - 1\right) = 4,586 - 0,521 t.$$

Ermitteln Sie die Parameter der Prognosegleichung und stellen Sie diese auf.

b) Ermitteln Sie die Zeitreihen- und Prognosewerte für den Zeitraum von 1968 bis 1980, und stellen Sie beide Zeitreihen graphisch dar.

- c) Ermitteln Sie die mittlere absolute Abweichung (MAA) und die Wurzel aus der mittleren quadratischen Abweichung (MQA).
- d) Führt das Senken des Sättigungsniveaus zu einer besseren Prognose (ex post)?

Lösung:

- a) Ermittlung der Parameter:

$$\begin{aligned} -\hat{C} &= 4,586 \\ \hat{C} &= -4,586 \\ \hat{a}S &= 0,521 \\ \hat{a} &= 0,521/80 \\ \hat{a} &= 0,0065 \end{aligned}$$

Setzen wir die geschätzten Parameter in Gleichung (115) ein, so erhalten wir als Prognosegleichung

$$\hat{x}_t = \frac{80}{1 + e^{-0,52t + 4,586}}$$

- b) Die Prognose- und Zeitreihenwerte sind in Tabelle 34 gegenübergestellt.

Tabelle 34

Jahr	t	x_t	\hat{x}_t
1968	1	2,00	1,35
1969	2	3,40	2,24
1970	3	3,50	3,70
1971	4	3,80	6,04
1972	5	9,10	9,65
1973	6	10,90	15,00
1974	7	21,10	22,38
1975	8	29,30	31,61
1976	9	42,20	41,88
1977	10	50,10	51,91
1978	11	60,90	60,53
1979	12	69,20	67,15
1980	13	73,80	71,83

Die graphische Gegenüberstellung der Zeitreihen- und Prognosewerte erfolgt in Abbildung 29.

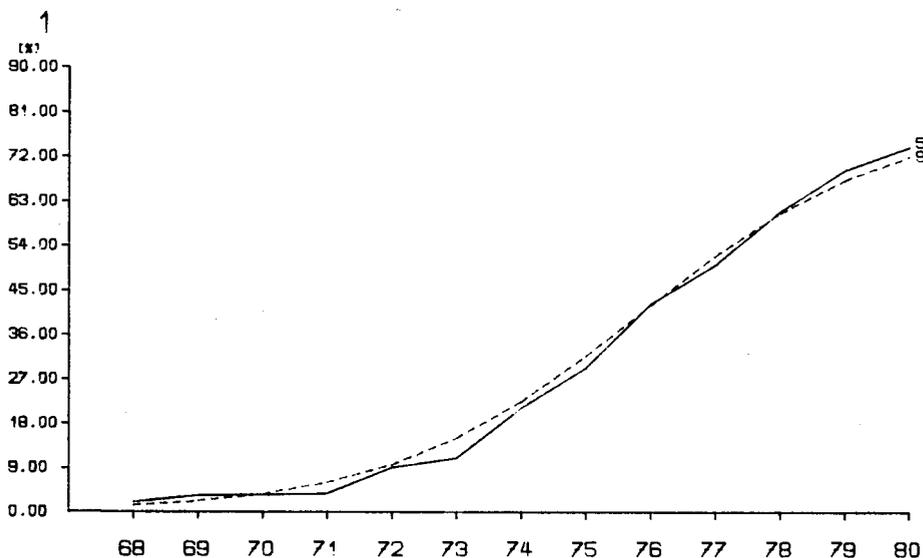


Abbildung 29

Ausstattung privater Haushalte mit Farbfernsehgeräten (Prognose gestrichelt)

$$c) \quad MAA = \frac{1}{13} \sum_{t=1}^{13} |(x_t - \hat{x}_t)| = \frac{18,97}{13} = 1,46$$

$$WMQA = \sqrt{\frac{1}{13} \sum_{t=1}^{13} (x_t - \hat{x}_t)^2} = \sqrt{\frac{42,51}{13}} = 1,81$$

d) Aufgrund der oben berechneten Gütemaße kann man festhalten, daß eine Senkung des Sättigungsniveaus von $S = 90\%$ auf $S = 80\%$ zu einer leichten Verschlechterung des Prognoseergebnisses (ex post) geführt hat.

Zu Kapitel G.: Prognosen mit der multiplen Regressionsanalyse

Aufgabe 27:

Bilden Sie für die Indexreihe y und die beiden vermuteten Regressoren x_1 und x_2 die Normalgleichungen und berechnen Sie die Koeffizienten b_i .

t	1	2	3
y_t	96,9	92,2	111
x_{1t}	97	105	120
x_{2t}	88	70	103

Lösung:

Gemäß (129) gilt

$$\hat{\mathbf{b}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

Mit $\mathbf{X} = \begin{pmatrix} 1 & 97 & 88 \\ 1 & 105 & 70 \\ 1 & 120 & 103 \end{pmatrix}$ und $\mathbf{y} = \begin{pmatrix} 96,9 \\ 92,2 \\ 111,0 \end{pmatrix}$

ergibt sich

$$\mathbf{X}^T = \begin{pmatrix} 1 & 1 & 1 \\ 97 & 105 & 120 \\ 88 & 70 & 103 \end{pmatrix}$$

$$\mathbf{X}^T \mathbf{y} = \begin{pmatrix} \sum y_t \\ \sum x_{1t} y_t \\ \sum x_{2t} y_t \end{pmatrix} = \begin{pmatrix} 300,1 \\ 32400,3 \\ 26414,2 \end{pmatrix}$$

$$\mathbf{X}^T \mathbf{X} = \begin{pmatrix} T & \sum x_{1t} & \sum x_{2t} \\ \sum x_{1t} & \sum x_{1t}^2 & \sum x_{1t} x_{2t} \\ \sum x_{2t} & \sum x_{2t} x_{1t} & \sum x_{2t}^2 \end{pmatrix} = \begin{pmatrix} 3 & 322 & 261 \\ 322 & 34834 & 28246 \\ 261 & 28246 & 23253 \end{pmatrix}$$

$$(\mathbf{X}^T \mathbf{X})^{-1} = \begin{pmatrix} 42,638017 & -0,404200 & 0,012407 \\ -0,404200 & 0,005744 & -0,002441 \\ 0,012407 & -0,002441 & 0,002869 \end{pmatrix}$$

Der Vektor der geschätzten Regressionskoeffizienten lautet also

$$\hat{\mathbf{b}} = \begin{pmatrix} 27,188 \\ 0,330 \\ 0,417 \end{pmatrix}$$

Aufgabe 28:

- a) Stellen Sie die Prognosegleichung für das Beispiel der Aufgabe 27 auf und prognostizieren Sie zwei Perioden in die Zukunft.

$$(x_{14} = 80, \quad x_{15} = 101, \quad x_{24} = 112, \quad x_{25} = 97)$$

- b) Bestimmen Sie den mittleren quadratischen Fehler über alle 5 Perioden, wenn die wahren Werte für den Index $y_4 = 93$, $y_5 = 100$ betragen.

Lösung:

- a) Die Prognosegleichung lautet

$$\hat{y}_t = 27,188 + 0,330 x_{1t} + 0,417 x_{2t}$$

Damit ergibt sich

$$\hat{y}_4 = 100,29$$

$$\hat{y}_5 = 100,97$$

- b) Mit der Formel für die mittlere quadratische Abweichung

$$\text{MQA} = \frac{1}{T} \sum_{t=1}^T (\hat{y}_t - y_t)^2$$

erhält man

$$\text{MQA} = \frac{1}{5} \cdot 58,06 = 11,61$$

$$\sqrt{\text{MQA}} = 3,41$$

Aufgabe 29:

Untersuchungen lassen vermuten, daß drei bestimmte Variable Einfluß auf die jährlichen Erträge von Zuckerrüben haben. Ein Regressionsmodell hat folgende Ergebnisse geliefert. Die Tabelle 35 gibt die tatsächlichen und geschätzten Rübenträge zur Zuckergewinnung für den Zeitraum 1969 bis 1981 wieder.

Tabelle 35

Jahr	x_t	\hat{x}_t
1969	42,3	41,79
1970	43,0	42,50
1971	42,2	44,29
1972	50,2	49,30
1973	49,5	48,57
1974	50,4	50,36
1975	50,5	51,31
1976	35,0	34,98
1977	52,1	52,11
1978	49,8	49,75
1979	48,5	50,37
1980	51,2	49,49
1981	59,9	59,79

Prüfen Sie,

a) ob der gemeinsame Einfluß aller Variablen auf die Rübenenerträge signifikant ist, ohne die Variablen im einzelnen zu kennen

und

b) ob die Prämisse P 2 verletzt ist.

Lösung:

a) Gemäß (138b) gilt:

$$F_{T-n-1}^n = \frac{s_y^2 - s_u^2}{s_u^2} \cdot \frac{T-n-1}{n}$$

Die Residuen \hat{u} betragen

Tabelle 36

t	Jahr	\hat{u}	\hat{u}^2	$(\hat{u}_t - \hat{u}_{t-1})$	$(\hat{u}_t - \hat{u}_{t-1})^2$
1	1969	0,51	0,260	—	—
2	1970	0,50	0,250	0,01	0,000
3	1971	-2,09	4,368	-2,59	6,708
4	1972	0,90	0,810	2,99	8,940
5	1973	0,93	0,865	0,03	0,001
6	1974	0,04	0,002	-0,89	0,792
7	1975	-0,81	0,656	-0,85	0,723
8	1976	0,02	0,000	0,83	0,689
9	1977	-0,01	0,000	-0,03	0,001
10	1978	0,05	0,003	0,06	0,004
11	1979	-1,87	3,497	-1,92	3,686
12	1980	1,71	2,924	3,58	12,816
13	1981	0,11	0,012	-1,60	2,560
	Σ	-0,01	1 13,647	-0,400	36,920

$$s_u^2 = \frac{1}{T-1} \left[\sum_{t=1}^T \hat{u}_t^2 - \frac{1}{T} \left(\sum_{t=1}^T \hat{u}_t \right)^2 \right]$$

$$= \frac{1}{12} \left[13,647 - \frac{1}{13} (-0,01)^2 \right] = 1,13725$$

$$s_y^2 = \frac{1}{T-1} \left[\sum_{t=1}^T y_t^2 - \frac{1}{T} \left(\sum_{t=1}^T y_t \right)^2 \right]$$

$$= \frac{1}{12} \left[30460,98 - \frac{1}{13} (624,6)^2 \right] = 37,61269$$

Der empirische F-Wert beträgt somit

$$F_{T-n-1}^n = \frac{37,61269 - 1,13725}{1,13725} \cdot \frac{13 - 3 - 1}{3} = 96,22011$$

Mit $n = 3$ Regressoren, $T = 13$ Werten und einer Irrtumswahrscheinlichkeit von $\alpha = 0,05$ ergibt sich aus der Tabelle ein theoretischer F-Wert von

$$F_{13-3-1}^3 = 3,86$$

Dieser ist kleiner als der empirische F-Wert, so daß die Regression als signifikant anzusehen ist.

b) Gemäß (144) ist der Durbin-Watson-Koeffizient definiert als

$$d = \frac{\sum_{t=2}^T (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^T \hat{u}_t^2} = \frac{36,920}{13,647} = 2,71$$

Dieser Wert liegt im Unschärfbereich des Koeffizienten:

$$4 - d_{4,13}^0 = 2,25 \leq 2,71 \leq 3,18 = 4 - d_{4,13}^u$$

und läßt keine Aussage darüber zu, ob Autokorrelation der Residuen vorliegt oder nicht.

Literaturverzeichnis

- Anderson, R. L. (1942): Distribution of the Serial Correlation Coefficient. In: Ann. Math.Stat., Vol. 13, No. 1, 1942.
- Bartlett, M. S. (1946): On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series. In: Journal Royal Stat.Soc., Bd. 8, No. 27, 1946.
- Box, G. E. P.; Jenkins, G. M. (1976): Time Series Analysis. San Francisco 1976.
- Brockhoff, K. (1977): Prognoseverfahren für die Unternehmensplanung. Wiesbaden 1977. SLB = X/12205=AR
12/1991
7/57190
- Brown, R. G. (1963): Smoothing, Forecasting and Prediction of Discrete Time Series. Englewood Cliffs 1963.
- Brown, R. G.; Meyer, R. F. (1961): The Fundamental Theorem of Exponential Smoothing. In: Operations Research, Vol. 9, 1961, S. 673–685.
- Bruckmann, G. (1978): Langfristige Prognosen. Möglichkeiten und Methoden der Langfristprognostik komplexer Systeme. 2. Aufl., Würzburg–Wien 1978.
- Chow, W. M. (1965): Adaptive Control of the Exponential Smoothing Constant. In: Journal of Industrial Engineering, 1965, S. 314 ff.
- Deutsche Bundesbank (1961): Erfahrungen mit der Anwendung der Regressionsrechnung auf die Saisonbereinigung statistischer Zeitreihen. In: Monatsberichte der Deutschen Bundesbank, 13. Jg., Nr. 8, 1961.
- Durbin, J.; Watson, G. S. (1950, 1951, 1971): Testing for the Serial Correlation in Least-Squares Regression I, II, III. In: Biometrika 37 (1950), S. 409–428; Biometrika 38 (1951), S. 159–178; Biometrika 58 (1971), S. 1–19.
- Geschka, H. (1978): Delphi. In: Bruckmann, G., Langfristige Prognosen, 2. Aufl., Würzburg–Wien 1978.
- Hannan, E. J. (1960): Time Series Analysis. London 1960.
- Hansmann, K.-W. (1979): Heuristische Prognoseverfahren. In: WISU, 8. Jg., 1979, H. 5, S. 229–233.
- Hansmann, K.-W. (1980a): Prognosen auf der Grundlage der exponentiellen Glättung. In: WISU, 9. Jg., 1980, H. 1, S. 17–21.
- Hansmann, K.-W. (1980b): Das Trendmodell der exponentiellen Glättung. In: WISU, 9. Jg., 1980, Teil I: H. 7, S. 323–325, Teil II: H. 8, S. 376–381.
- Hansmann, K.-W. (1980c): Die Anwendung der Spektral- und Regressionsanalyse bei der Absatzprognose von Zeitschriften. In: Angewandte Prognoseverfahren, Hrsg.: Schwarze, J., Herne–Berlin 1980, S. 231–244.
- Hansmann, K.-W. (1980d): Dynamische Aktienanlageplanung. Wiesbaden 1980.
- Harrison, R. M. (1965): Short-Term Sales Forecasting. In: Applied Statistics 1965, S. 102–139.
- Hochstädter, D.; Ube, G. (1970): Ökonometrische Methoden. Berlin–Heidelberg–New York 1970.
- Holt, C.; Modigliani, F.; Muth, J. F.; Simon, H. A. (1960): Planning Production, Inventories and Work Force. Englewood Cliffs 1960.
- Hüttner, M. (1982): Markt- und Absatzprognosen. Stuttgart–Berlin–Köln–Mainz 1982.

- Johnston, J. (1963): *Econometric Methods*. New York–Toronto–London 1963.
- Kahn, H. (1976): *The Next 200 Years*. New York 1976.
- König, H.; Wolters, J. (1972): *Einführung in die Spektralanalyse ökonomischer Zeitreihen*. Meisenheim am Glan 1972.
- Leiner, B. (1978): *Spektralanalyse ökonomischer Zeitreihen*. 2. Aufl., Wiesbaden 1978.
- Lewandowski, R. (1974): *Prognose- und Informationssysteme und ihre Anwendung*, Bd. 1, Berlin–New York 1974.
- Lewandowski, R. (1980): *Prognose- und Informationssysteme und ihre Anwendung*, Bd. 2, Berlin–New York 1980.
- Makridakis, S.; Wheelwright, S. C. (1978a): *Forecasting, Methods and Application*. New York 1978.
- Makridakis, S.; Wheelwright, S. C. (1978b): *Interactive Forecasting. Univariate and Multivariate Methods*. 2nd Edition, San Francisco 1978.
- Makridakis, S.; Reschke, H.; Wheelwright, S. C. (1980): *Prognosetechniken für Manager*. Wiesbaden 1980.
- Menges, G. (1961): *Ökonometrie*. Wiesbaden 1961.
- Mertens, P. (Hrsg.) (1981): *Prognoserechnung*. 4. Aufl., Würzburg–Wien 1981.
- Nullau, B. (1969): Darstellung des Verfahrens. In: *Das Berliner Verfahren. Beitrag zur Zeitreihenanalyse DIW-Beiträge zur Strukturforschung*, H. 7, 1969.
- Parzen, E. (1961): *Mathematical Considerations in the Estimation of Spectra*. In: *Technometrics*, Vol. 3, 1961.
- Quenouille, M. H. (1957): *Analysis of Multiple Time Series*. New York 1957.
- Quenouille, M. H. (1949): Approximate Tests of Correlation in Time Series. In: *Journal Royal Stat. Soc.*, B. 11, 1949, S. 68 ff.
- Scheer, A.-W. (1980): Absatzprognosen. In: *Rationelle Betriebswirtschaft*, Hrsg.: Müller, W.; Krink, J., Neuwied 1980, Kapitel XIV, S. 1–222.
- Schläger, W. (1981): Einführung in die Zeitreihenprognose bei saisonalen Bedarfsschwankungen und Vergleich der Verfahren von Winters und Harrison. In: *Prognoserechnung*, 4. Aufl., Hrsg.: Mertens, P., Würzburg–Wien 1981, S. 59–78.
- Schneeweiss, H. (1978): *Ökonometrie*. 3. Aufl., Würzburg–Wien 1978.
- Schröder, M. (1981): Einführung in die kurzfristige Zeitreihenprognose und Vergleich der einzelnen Verfahren. In: *Prognoserechnung*, 4. Aufl., Hrsg.: Mertens, P., Würzburg–Wien 1981, S. 23–58.
- Schwarze, J. (Hrsg.) (1980): *Angewandte Prognoseverfahren*. Herne–Berlin 1980.
- Shiskin, J.; Eisenpress, H. (1957): Seasonal Adjustments by Electronic Computer Methods. In: *Journal of the American Statistical Association*, Vol. 52, 1957, S. 415–449.
- Slutzky, E. (1937): The Summation of Random Causes as the Source of Cyclical Processes. In: *Econometrica*, Vol. 5, 1937.
- Smith, D. E. (1974): Adaptive Response for Exponential Smoothing. In: *Operational Research Quarterly*, Vol. 25, 1974, S. 421 ff.
- Theil, H. (1971): *Applied Economic Forecasting*. 2nd Edition, Amsterdam 1971.

- Trigg, D. W.; Leach, A. G. (1967): Exponential Smoothing with an Adaptive Response Rate. In: Operational Research Quarterly, Vol. 18, 1967, S. 53–59.
- Walker, G. (1931): On Periodicity in Series of Related Terms. In: Proc. Royal Soc., A131, 1931, S. 518 ff.
- Widrow, B. (1966): Adaptive Filters One: Fundamentals, Su-Sel-66-126, Systems Theory Laboratory. Stanford University, Stanford, Cal. 1966.
- Winters, P. R. (1960): Forecasting Sales by Exponentially Weighted Moving Averages. In: Management Science, Vol. 6, 1960, S. 324–342.
- Wold, H. O. (1954): A Study in the Analysis of Stationary Time Series. 2nd Edition, Uppsala 1954.
- Yule, G. U. (1927): On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Woelfer's Sunspot Numbers. In: Phil. Trans., A 226, 1927, S. 267 ff.
- Zwicky, F. (1971): Entdecken, Erfinden, Forschen im morphologischen Weltbild. München–Zürich 1971.

Quellen der Zeitreihenwerte

- Hrsg.: Der Bundesminister für Verkehr
Verkehr in Zahlen 1981, Bonn 1981
- Hrsg.: Deutsche Bundesbank
Monatsberichte, Beihefte, Reihe 2,
Frankfurt am Main, lfde. Jahrgänge
- Hrsg.: Deutsche Shell AG
Aktuelle Wirtschaftsanalysen.
Aufschwung nach der Talfahrt,
Hamburg 1981
- Hrsg.: IFO-Institut für Wirtschaftsforschung
Reihe: IFO-Schnelldienste, lfde. Jahrgänge
- Hrsg.: Statistisches Bundesamt
Statische Jahrbücher der Bundesrepublik
Deutschland. Wiesbaden, lfde. Jahrgänge
- Hrsg.: Statistisches Bundesamt
Wirtschaft und Statistik,
Wiesbaden, lfde. Jahrgänge

Stichwortverzeichnis

- Adaptives Filtern 98 ff., 163 ff.
- Anfangsschätzungen 81 ff., 162
- ARIMA-Modell 74, 89, 102
- ARMA-Modell 161
- AR-Modell 65, 101
- Autokorrelationsfunktion 54 f., 65, 71, 135, 137, 160
 - partielle 68 ff.
- Autokorrelationskoeffizienten 65
 - partielle 68, 70 f., 159 ff.
- Autokovarianzfunktion 54 f., 58, 65, 70

- Bestimmtheitsmaß 130, 138
- BLUE-Eigenschaft 130
- Box-Jenkins-Verfahren 63 ff., 65

- Delphi-Prognose 22 ff., 147 f.
- Differenzenbildung 72
- Dualität 69, 70
- × Durbin-Watson-Koeffizient 138, 176

- Exponentialfunktion 108 ff.
- × Exponentielle Glättung 27 ff., 40, 48, 149, 152 f.
 - höherer Ordnung 43 f.
 - konstantes Modell 28 ff., 152 ff.
 - Mehr-Parameter-Modelle 44
 - Saisonverfahren 47 ff.
 - Trendmodell 33 ff., 47, 153 ff.
- Exogen 125
- Endogen 125
- Erwartungstreue 129

- Fehlervarianz 137
- Filter 74
- Filtern
 - adaptives 98 ff., 163 ff.
 - lineares 63
- Fourier-Transformation 55 f.
- Frequenzbereich 55 f., 159

- Funktion
 - Exponential- 108 ff.
 - Gompertz- 117 ff.
 - logistische 111 ff.
 - nicht lineare 126

- Gewichte 64 f.
- Glättungsoperator 34 f., 37, 155 f.
- Glättungsparameter 28, 31, 35, 150 f.
 - dynamische Anpassung 39 ff., 155 f.
- Gleitende Durchschnitte 27 f., 46, 149
- Gompertz-Funktion 117 ff.
- Gradient 99, 164
- Grundwert 33, 38, 48 f., 155 f., 158
- Gütemaße 32, 39, 42, 51, 96, 102, 110, 119, 122, 138, 153, 168, 171, 173

- Hauptschätzverfahren 83 ff., 98 ff.
- Homoskedastizität 128

- Kausalbeziehung 125
- Konfidenzintervall 128
- Konsistenz 130
- Konvergenz 85 ff.
- Korrelation 136
- Korrelogramm 54
- KQ-Methode (siehe Methode der Kleinsten Quadrate)
- KQ-Schätzung 83 ff., 126

- Lernkonstante 98
- Lag Fenster 58
- Logistische Funktion 111 ff.

- MA-Modelle 69
- Maximum-Likelihood
 - Prinzip 128
 - Schätzung 83 ff.
- Methode der Kleinsten Quadrate 106, 108 f., 110, 113 f., 118, 121, 166 ff.

- Methode von Chow und Smith 40 f.
- Mittlere absolute Abweichung (MAA)
 - 15, 32, 39, 40, 42, 51, 96, 110, 115, 119, 122, 138, 153, 154, 171
- Mittlerer absoluter Prognosefehler 40
- Mittlerer Prognosefehler 40
- Mittlere quadratische Abweichung (MQA)
 - 15, 102, 168, 173
- Multikollinearität 135

- nicht-stationär 90
- Normalgleichungen 106, 128 f., 166 ff.
- Normalisierungskonstante 101, 165

- Parameterschätzung 81 ff., 128
- Parametertest 132 f.
- Prognose 11, 145, 171,
 - zeitlicher Ablauf 13 f.
 - Beurteilungsmaße 14 ff.
- Prognose (ex-post) 88
- Prognosefehler 15, 28
 - absoluter 15
 - mittlerer 40
 - mittlerer absoluter 40
 - quadratischer 15
 - relativer absoluter 15
- Prognosemodell 12 f.
 - Auswahl eines 141 f.
 - univariate 63
- Prognoseverfahren
 - heuristische 18 ff.
 - quantitative 27 ff.
- Prognosezeitraum 13
- Prozesse
 - autoregressiver 65, 136, 159
 - Grad des 68 ff.
 - Moving Average 69
 - nichtstationäre 72
 - Ordnung des 67 ff., 69 ff.
 - stationäre-stochastische 63 ff.
 - stochastische 125
- Prüfkriterien der Güte 130
- × Regressionsanalyse
 - multiple 125 ff., 171 ff.
- Regressionsfunktion 126
- Regressionskoeffizienten 125 ff., 172
- Regressionsprämissen 127
- Regressor 125 ff., 171
- Residual-Varianz 129
- Residuum 88, 130, 175

- Sättigungsniveau 111 ff., 117 ff., 168 f.
- Saisonfaktor 47 ff., 158
- Saisonverfahren 46 ff.
 - Winters 47 ff., 157 f.
- Saisonzyklus 47 ff., 55 f.
- Schätzfehler 88, 136
- Signifikant 79, 89, 94, 132 f., 174
- Signifikanz 135
- Spektralanalyse 53 ff.
- Spektrum 56 ff.
- Stationarität 66, 75, 160
- Stieltjes-Integral 55
- Stochastischer Prozeß 53 f., 58, 109
 - streng stationär 54
 - schwach stationär 54
- Störvariable 29 f., 33
- Strukturbruch 31, 151
- Summenfilter 74, 88
- System
 - interdependentes 125
- Szenario-Technik 18 ff., 146 f.

- Taylor-Reihe 84 f.
- Test 78
 - Bartlett 78
 - χ^2 89, 96
 - Durbin-Watson 135
- Theil'scher Ungleichheitskoeffizient
 - 15, 32, 39, 42, 51, 96, 110, 115, 119, 122, 138, 153, 154 f.
- Time lag 75, 94
- Trend 73, 132
- Tendextrapolation 104 ff.
- Trendwert 33, 38, 48 f., 155 f., 158

- Unverzerrtheit 129

- Varianz
 - erklärte 131
- Verteilung (F) 134
- Wachstums- und Sättigungsmodelle
 - 104 ff.
- Wendepunktfehlermaß 17
- × Winters Saisonverfahren 47 ff., 157 f.
- Wurzel aus der mittleren quadratischen
 - Abweichung (WMQA) 15, 32, 39, 42,
 - 51, 96, 102, 110, 115, 119, 122, 138,
 - 153, 154, 171, 173
- Yule-Walker-Gleichung 68, 71, 162
- Zeitreihe 11, 53, 125, 145
 - saisonbehaftete 73
- Zeitreihenanalyse 63, 125
- Zeitstabilitätshypothese 11, 28, 45, 145