Prognose makroökonomischer Zeitreihen: Ein Vergleich linearer Modelle mit neuronalen Netzen

WOLFGANG KOLLER

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

PL ACADEMIC RESEARCH

1 Einleitung

Seit der Wiederentdeckung der neuronalen Netze (NN) etwa Anfang der 80er Jahre des vorigen Jahrhunderts erlebte dieses Forschungsgebiet ein spektakuläres Anwachsen von Forschungsaktivitäten und -ergebnissen. Inspiriert durch die ursprüngliche Idee, mit Hilfe einfacher rechnerischer Einheiten die Funktionsweise von Neuronen und des Gehirns nachzubilden, wurde die Theorie immer weiter ausgebaut und breite Anwendungsmöglichkeiten gefunden, zu denen neben Mustererkennung, Signalverarbeitung und Prozesssteuerung bald auch die Zeitreihenanalyse und -prognose gehörte. Für dieses Anwendungsgebiet kommen meist die auch in der vorliegenden Arbeit untersuchten Feedforward-Netze zum Einsatz, für die besonders viele und gut abgesicherte Ergebnisse vorliegen. Wenn in dieser Arbeit von NN die Rede ist, bezieht sich dies ab nun auf die Feedforward-Netze.

NN-Modelle werden heute als semi-parametrische oder parametrische nicht-lineare Modelle interpretiert, für die die entsprechende statistische Theorie zur Anwendung kommen sollte. Viele der ursprünglichen Innovationen der NN-Forschung haben eine statistisch-theoretische Absicherung erfahren bzw. wurden in die bestehende statistische Theorie integriert.

Frühe Arbeiten im Bereich der Prognose von Zeitreihen ließen die Zeitdimension der Daten oft außer Acht und bauten kaum auf den verfügbaren Grundlagen der Zeitreihenökonometrie und angewandten Statistik auf. Inzwischen hat sich die Sichtweise durchgesetzt, dass in diesem Anwendungsgebiet NN als nicht-lineare Zeitreihenmodelle eingesetzt werden können und im Vergleich mit etablierten linearen und nicht-linearen Zeitreihenmodellen bewertet werden sollten. Insbesondere für die autoregressiven Neuronale-Netz-Modelle (ARNN), die als Generalisierung der autoregressiven Modelle (AR) gelten, wurden seit Mitte der 90er Jahre wichtige theoretische und praktische Ergebnisse erzielt. Diese Entwicklungen erfolgten parallel und in wachsender Verzahnung mit der Forschung zu anderen nicht-linearen Zeitreihenmodellen und zu anderen benachbarten Gebieten der Zeitreihenanalyse wie der Testung auf Nicht-Linearität von Zeitreihen und der Messung und Testung der Prognosegüte.

Der Einsatz von NN für die Modellierung und Prognose ökonomischer Zeitreihen stellt eine Herausforderung dar, da diese eine Reihe von Eigenschaften aufweisen, die besondere Aufmerksamkeit und teilweise eine Erweiterung des Instrumentariums der NN erforderlich machen. Ökonomische Zeitreihen zeichnen sich durch vergleichsweise geringe Länge und einen hohen Anteil des nicht-prognostizierbaren Fehlers aus. Beides sind Faktoren, die gerade den Einsatz von NN erschweren. Zudem spielen nicht-stationäre Komponenten (stochastischer bzw. deterministischer Trend und Saisonalität) eine wichtige Rolle, deren richtige Handhabung auf die Prognosequalität große Auswirkungen haben kann. Schließlich kann a priori nicht mit Sicherheit gesagt werden, ob in einer ökonomischen Zeitreihe Nicht-Linearität jener Art vorhanden ist, die mit neuronalen Netzen modelliert werden kann.

1.1 Zielsetzung und Ausrichtung der Arbeit

Das Ziel der Arbeit ist es, die Eignung des Instrumentariums der NN, im konkreten der ARNN-Modelle, zur Modellierung und Prognose von ökonomischen Zeitreihen zu untersuchen und mit jener der AR- und ARMA-Modelle (autoregressive Moving-Average-Modelle) als den wichtigsten Vertretern der univariaten linearen Zeitreihenmodelle zu vergleichen. Als beispielhaftes Anwendungsgebiet werden die beiden monatlichen Zeitreihen der österreichischen Arbeitslosenrate und des österreichischen Industrieproduktionsindex herangezogen. Die Arbeit beinhaltet die Enwicklung einer Reihe von Erweiterungen am Instrumentarium der ARNN-Modelle, die durch die besonderen Anforderungen des Anwendungsgebiets motiviert sind. Somit umfasst die Zielsetzung der Arbeit sowohl die Gewinnung von Aussagen über das eingesetzte Instrumentarium als auch über die untersuchten Zeitreihen:

Erkenntnisse zum Instrumentarium Wie wird die Familie der ARNN-Modelle definiert und welche Erweiterungen an der Spezifikation sind vorzusehen? Welche Techniken und Vorgangsweisen, die für die Prognose mit ARNN-Modellen, aber auch mit linearen Modellen entwickelt werden, haben eine gute Prognosegüte hinsichtlich Effizienz und Verlässlichkeit?

Erkenntnisse zu den Zeitreihen Was sind die besonderen Eigenschaften der beiden Beispielzeitreihen, die ihre Vorhersagbarkeit durch lineare und nicht-lineare Methoden möglich machen? Dies sind z.B. Trend, Saisonalität, Strukturbrüche, vor allem aber die Frage nach nicht-linearen Strukturen, die den Einsatz von nicht-linearen Modellen erst aussichtsreich erscheinen lassen.

Es werden sowohl die theoretischen Grundlagen als auch die angewandten Aspekte des Einsatzes von NN und linearen Modelle erläutert. Die hauptsächliche Ausrichtung ist eine angewandte und kann mit jener von wichtigen Referenzarbeiten wie Swanson und White (1997b), Teräsvirta et al. (2005), und Medeiros et al. (2006) verglichen werden, die ARNN-Modelle in ähnlicher Weise wie die vorliegende Arbeit spezifizieren und ebenso auf einer Auswahl von Zeitreihen deren Eignung im Vergleich zu alternativen Modellen untersuchen. Während in diesen Arbeiten die Auswahl der Beispielzeitreihen breiter ist, wird in der vorliegenden Arbeit auf nur zwei realweltliche Zeitreihen eingeschränkt, dafür jedoch die Anzahl der verschiedenen Modellierungsoptionen vertieft betrachtet. Ein systematischer, umfangreicher Prognosevergleich streng nach dem „Out-of-Sample“-Prinzip erlaubt einen fairen Vergleich der Prognosegüte der verschiedenen Ansätze und Modellierungsstrategien, wobei auch statistische Tests auf Prognosegüte zum Einsatz kommen.

Die in dieser Arbeit eingeführten Neuheiten am Instrumentarium der ARNN-Modellierung sind meist naheliegende Erweiterungen bestehender Methoden und nicht von großer theoretischer Reichweite, haben jedoch, wie sich zeigen wird, großen potentiellen praktischen Nutzen auf dem Anwendungsgebiet der ökonomischen Zeitreihen. Besonders zentral ist die Erweiterung der Spezifikation der ARNN-Modelle zur Berücksichtigung von sparsamer Modellspezifikation, Einschluss von deterministischer Saisonalität und Trendkomponente und schließlich zur Ermöglichung von direkten Mehr-Schritt-Prognosen.

Weiters wird ein Repertoire von verschiedenen konkreten Vorgangsweisen für die Modellierung und Prognose von ökonomischen Zeitreihen entwickelt, das auf verschiedenen Ansätzen beruht: dem statistisch-parametrischen Ansatz, dem klassischen Ansatz mit Early-Stopping, dem Ansatz mit Regulierung und dem Bayesianischen Evidenzansatz. Für die Auswahl dieser Methoden sowie deren Weiterentwicklung im Rahmen der vorliegenden Arbeit sind als besonders wichtige Ausgangspunkte die Monographien von Bishop (1995, 2006) und Anders (1997) hervorzuheben, letztere insbesondere für die zentrale Idee von Modellierungsstrategien.

Die Analysen und Modelle werden mit Hilfe der frei verfügbaren Statistik-Software R implementiert, wobei die selbst entwickelten Techniken und Methoden in Form von zwei R-Stoftwarepaketen zusammengefasst und dokumentiert werden. Es ist dies einerseits das Paket `NNUTS`, welches Funktionen und Algorithmen im Zusammenhang mit ARNN-Modellen enthält und andererseits `seasonal`, das verschiedene Funktionen im Zusammenhang mit der linearen Modellierung von saisonalen Zeitreihen bereitstellt. Diese Software und ihre Dokumentation sind vom Autor bzw. vom Comprehensive R Archive Network (CRAN) unter `http://CRAN.R-project.org` erhältlich. Es ist zu hoffen, dass auch auf dem Weg der Bereitstellung von freier wissenschaftlicher Software ein Beitrag zur weiteren Forschung auf dem Gebiet der Modellierung und Prognose von ökonomischen Zeitreihen mit linearen und Neuronale-Netz-Methoden geleistet werden kann.

Aufgrund der Arbeit soll ein interessierter Praktiker der Zeitreihenanalyse eine umfassende Beurteilung der Eignung von neuronalen Netzen zur Modellierung und Prognose von ökonomischen Zeitreihen zur Verfügung haben. Dies soll die Grundlage bieten können sowohl für die Entscheidung, ob man überhaupt das Instrumentarium der neuronalen Netze einsetzt, als auch für die Auswahl eines bestimmten Modellierungszugangs.

1.2 Vorschau

Den Zielsetzungen der Arbeit folgend, sowohl Aussagen über das lineare und NN-Instrumentarium als auch über die Zeitreihen zu generieren, weist jedes Kapitel sowohl theoretische bzw. methodische Abschnitte als auch Anwendungen anhand der beiden Zeitreihen auf. Die Arbeit ist in vier Hauptkapitel (Kap. 2–4) gegliedert. Kapitel 1 führt in die Grundlagen der Zeitreihenanalyse und der linearen Modellierung ein und hat in diesem Sinne eher vorbereitenden Charakter. Doch auch hier werden sowohl wichtige Aussagen zu den beiden untersuchten Zeitreihen als auch zu den verschiedenen Techniken der Modellierung gewonnen.

Kapitel 3 bringt Grundlagen zu nicht-linearen Zeitreihenmodellen und stellt die Frage nach dem Vorhandensein von Nicht-Linearität aus der Perspektive von statistischen Hypothesentests. Nur wenn in einer Zeitreihe ausreichend Nicht-Linearität von jenem Typ vorhanden ist, der mit NN behandelt werden kann, darf man erwarten, mit NN die Prognosegüte von linearen Modellen zu übertreffen. Eine Abfolge von Nicht-Linearitätstests, die in Hinblick auf die Erfordernisse von ökonomischen Zeitreihen zusammengestellt und teilweise angepasst wird, wird diskutiert und auf die beiden Beispielzeitreihen angewendet.

In Kapitel 4 werden die ARNN-Modelle entwickelt und Methoden für ihre Schätzung und Spezifikation, d.h. für die Modellierung insgesamt, vorgestellt und weiterentwickelt. Insbesondere wurde der Schwerpunkt auf vier verschiedene Modellierungsansätze gelegt, mit denen das für NN besonders prekäre Problem des Overfitting bzw. der Generalisierungsfähigkeit gelöst werden kann. Die Funktionsweise dieser Techniken wird sowohl anhand einer einfachen simulierten nicht-linearen Zeitreihe als auch anhand der beiden Beispielzeitreihen demonstriert. Dieses Kapitel richtet zwar das Hauptinteresse auf Fragen zum Instrumentarium, aber auch hier können aus den Problemen, die bei der Anwendung des Instrumentariums auf die realweltlichen Zeitreihen auftreten, wichtige Schlüsse über die Eigenschaften derselben gezogen werden.

Das letzte Hauptkapitel (Kap. 5) ist schließlich einer systematisch angelegten Prognose- und Evaluierungsanordnung gewidmet, in der auf einem Evaluierungsset („Out-of-Sample"), d.h. einem zur Seite gelegten Abschnitt der Zeitreihe, die Prognosegüte der linearen und NN-Methoden getestet wird. In diesem Kapitel liegt der Schwerpunkt auf Erkenntnissen zu den Daten. Sind die Zeitreihen mit einer der verwendeten Varianten der ARNN-Modelle besser, d.h. deutlich bzw. in statistisch signifikanter Weise besser, als mit linearen Modellen zu prognostizieren, so ist dies ein Nachweis von nicht-linearen Strukturen in diesen beiden konkreten Zeitreihen. In diesem Sinne liefert der systematische Prognosevergleich ein aussagekräftigeres Ergebnis zum Vorhandensein nicht-linearer Strukturen als die Nicht-Linearitätstests des Kapitels 3. Außerdem erfordert die Zusammenstellung der Prognose- und Evaluierungsanordnung die Entwicklung neuen und zusätzlichen Instrumentariums, insbesondere zur Ermöglichung von Mehr-Schritt-Prognosen.

Zwei Anhänge enthalten die Herleitung der Ableitungen der Fehlerfunktion eines ARNN-Modells nach den einzelnen Koeffizienten des Modells, die für die numerische Umsetzung der verschiedenen Lernverfahren notwendig sind (Anhang A), sowie zusätzliche Tabellen mit Ergebnissen zu verschiedenen Sensitivitätsanalysen im Zusammenhang mit dem systemantischen Vergleich der Prognosegüte (Anhang B).

1.3 Mathematischen Notation und Sprachgebrauch

In dieser Arbeit wird versucht, so weit wie möglich eine einheitliche mathematische Notation zugrunde zu legen. Aufgrund der großen Anzahl und Vielfalt der formelmäßig darzustellenden Konzepte, die zudem aus unterschiedlichen Fachrichtungen stammen, ist dies nicht immer möglich bzw. sinnvoll, da neben das Ziel der einheitlichen Notation auch jenes der Verwendung der aus der jeweiligen Fachliteratur vertrauten Notation tritt. Es müssen daher in einzelnen Fällen mathematische Symbole in verschiedener Bedeutung verwendet werden, wofür der Leser um sein Verständnis gebeten wird. In jedem dieser Fälle ist die gewählte Definition des Symbols im unmittelbaren Zusammenhang angegeben und aus dem Kontext unmissverständlich. Für die Arbeit zentrale Konzepte werden weitgehend in einheitlicher Notation dargestellt. Manche Symbole, die fest im ökonometrischen Sprachgebrauch verankert sind, wie t-Wert, F-Statistik, Q-Statistik u.ä. werden ohne besonderen Hinweis verwendet, auch wenn diese Symbole an anderer Stelle in anderer Definition eingesetzt werden.

Mathematische Symbole werden im allgemeinen kursiv gesetzt. Davon ausgenommen sind verschiedene besondere statistische und mathematische Funktionen wie Erwartungswert, Varianz, Kovarianz, Indikatorfunktion usw., die aufrecht gesetzt werden. Ebenso aufrecht gesetzt werden Vektoren und Matrizen, die zusätzlich fettgeschrieben werden. Für Mengen wird kalligraphischer Schriftsatz verwendet. Griechische Symbole werden vorwiegend für Koeffizienten in Modellen sowie für Hyperparameter und Heuristiken im Rahmen von Algorithmen und Modellierungsstrategien verwendet. Durch die über einem Symbol gesetzten Superskripte $\tilde{\cdot}$, $\bar{\cdot}$, $\hat{\cdot}$, $\dot{\cdot}$ wird eine Modifikation des ursprünglichen Terms, Durchschnitts- oder Erwartungswertbildung, Schätzung sowie Vorläufigkeit oder Ziehung aus einer Simulation zum Ausdruck gebracht (in dieser Reihenfolge). Durch ein hochgestelltes Sternchen bei einer Variable, $\cdot^*$, wird angezeigt, dass diese einen besonderen Wert annimmt, z.B. im Sinne einer Restriktion oder Optimierung. Ein hochgestelltes $\cdot^{(\tau)}$ bezeichnet die Abhängigkeit vom Iterationszeitpunkt. Für die Transposition eines Vektors oder einer Matrix, ebenso wie für die Ableitung einer Funktion wird das Apostroph verwendet, $\cdot'$. Bei der Schreibweise von Dezimalzahlen wird, der englischsprachigen Konvention folgend, nicht das Komma sondern der Punkt als Dezimaltrennzeichen verwendet. Eine umfassende Aufstellung der verwendeten mathematischen Symbole befindet sich am Ende der Arbeit vor dem Literaturverzeichnis.

Auch zum Sprachgebrauch sind einige Vorbemerkungen angebracht. Englischsprachige Fachausdrücke werden soweit als möglich ins Deutsche übertragen. Meist gibt es eindeutige deutsche Übersetzungen. Doch gelegentlich ist eine Übersetzung nur schwer möglich oder verwirrend und wird in diesen Fällen nicht vorgenommen. Die Einführung von neuer deutschsprachiger Fachsprache ist keine wesentliche Aufgabe dieser Arbeit. Wird in Einzelfällen dennoch der Versuch unternommen, so wird immer auf den englischen Ausdruck hingewiesen. Teilweise kann das Problem durch die Verwendung von Abkürzungen umgangen werden. Abkürzungen werden bei ihrer erstmaligen Verwendung definiert und sind in einem Abkürzungsverzeichnis am Ende der Arbeit aufgelistet.

2 Lineare Modellierung von Zeitreihen

Lineare univariate Zeitreihenanalyse wird in der Ökonomie so oft und in so vielen verschiedenen Gebieten angewendet, dass eine Rechtfertigung der Verwendung dieses Instrumentariums nicht notwendig erscheint. Die Alternativen zu linearen Modellen bestehen einerseits im gänzlichen bzw. teilweisen Verzicht auf quantitative Methoden (d.h. im Rückgriff auf reine Heuristiken, Expertenurteile und sogenannte „Judgement-Methoden“) andererseits in der Anwendung nicht-linearer Modelle, die jedoch als Weiterentwicklungen und Verallgemeinerungen der linearen Modelle gesehen werden können und daher auf den Begriffen und der Theorie der linearen Zeitreihenanalyse aufbauen.

In diesem Sinne liefert dieses Kapitel einen Überblick nicht nur über die theoretischen und praktischen Aspekte der linearen Zeitreihenanalyse, sondern greift auch die Grundlagen der Analyse und insbesondere der Prognose von Zeitreihen insgesamt auf. Diese Fragen nehmen mitunter einen ausgesprochen wissenschaftstheoretischen Charakter an („Was ist überhaupt Prognose?“). Andererseits werden die in diesem Kapitel gelegten theoretischen Grundlagen und praktischen Aspekte auch in den darauffolgenden Kapiteln aufgegriffen und vorausgesetzt.

Abschnitt 2.1 führt in die statistischen Grundlagen der Zeitreihenanalyse ein und stellt jene linearen Zeitreihenmodelle vor, die im Rahmen dieser Arbeit untersucht und mit nicht-linearen Modellen verglichen werden sollen. Abschnitt 2.2 bringt einen Überblick über die Anwendung linearer Zeitreihenanalyse für ökonomische Zeitreihen und stellt jene beiden Zeitreihen vor, die in dieser Arbeit durchgehend als Anwendungsbeispiele herangezogen werden. Abschnitt 2.3 widmet sich der Frage, wie die Stationarität einer Zeitreihe festgestellt werden kann und was geeignete Wege zur Herbeiführung von Stationarität sind. Abschnitt 2.4 dient der Entwicklung einer konkreten Vorgangsweise bei der linearen Modellierung und wendet diese auf die beiden Beispielzeitreihen an. Abschnitt 2.5 fasst einige wichtige Erkenntnisse zusammen, die sich aus der linearen Modellierung der beiden Beispielzeitreihen ergeben.

2.1 Grundlagen der linearen Zeitreihenanalyse

Dieser Abschnitt enthält eine Einführung in Denkweisen, Grundbegriffe und die wichtigsten Modelle der linearen Zeitreihenanalyse. Da insbesondere die Details in einer Vielzahl von ausgezeichneten Lehrbüchern der Zeitreihenanalyse nachgelesen werden können (z.B. Mills, 1990; Brockwell und Davis, 1991; Harvey, 1993; Enders, 1995; Schlittgen und Streitberg, 1997), werden einige Aspekte nur knapp behandelt, während andere im Zusammenhang mit dem Ziel der Arbeit stehende Aspekte hingegen stärker hervorgehoben werden. Zudem wird die Behandlung einiger wichtiger Konzepte, die auch für die lineare Zeitreihenanalyse wichtig sind, auf spätere Kapitel verschoben, da sie, wie etwa das Generalisierungsproblem (vgl. Abschnitt 4.2) oder Modellselektionskriterien (vgl. Abschnitt 4.6) im Zusammenhang mit der Verwendung von neuronalen Netzen eine besonders ausgeprägte Bedeutung haben.

Denkweisen und Grundbegriffe

Eine Zeitreihe (x_t) ist eine zeitlich geordnete Folge von Beobachtungen einer interessierenden Größe x_t, wobei $t \in \mathbb{N}$ der Zeitindex ist.[1] Kennt man von einer Zeitreihe nur vergangene Beobachtungen $x_t, t \leq t^*$, wobei t^* der gegenwärtige Zeitpunkt ist, so ist in vielen Anwendungsgebieten die Prognose zukünftiger $x_t, t > t^*$ eine interessante Aufgabenstellung.

Prognose beruht auf der Annahme, dass die Zukunft – zumindest teilweise – wie die Vergangenheit ist. In der Sprache der Zeitreihenanalyse manifestiert sich diese Denkweise im Begriff des „datengenerierenden Prozesses“ (DGP): es wird angenommen, dass eine im Zeitverlauf unveränderliche Gesetzmäßigkeit existiert, die der beobachteten Zeitreihe zugrundeliegt. Auf der Basis dieser Annahme kann man es wagen, von der Vergangenheit auf die Zukunft zu schließen.[2] Im Rahmen der Zeitreihenanalyse wird versucht, den DGP durch ein Modell abzubilden und dieses gemeinsam mit den bekannten Werten der Zeitreihe zur Prognose der zukünftigen Ausprägungen der Zeitreihe zu verwenden. Die vorliegende Arbeit beschränkt sich auf die univariate Zeitreihenanalyse, die zur Prognose zukünftiger $x_t, t > t^*$ ausschließlich die Informationen verwendet, die in den bisherigen $x_t, t \leq t^*$ enthalten ist.

Für ökonomische Zeitreihen ist es eine weithin akzeptierte Hypothese, dass im Rahmen des DGP auch der Zufall eine wesentliche Rolle spielt, man spricht daher von „stochastischen Prozessen“.[3] Jede einzelne Beobachtung x_t ist hierbei die Realisation einer Zufallsvariable. Die gegenteilige Annahme, ökonomische Zeitreihen könnten das Ergebnis eines rein deterministischen Prozesses sein, wird in Abschnitt 3.1 aufgegriffen. Die Unterscheidung zwischen der Zeitreihe und dem sie generierenden stochastischen Prozess ist essentiell. Die beobachtete Zeitreihe ist nur eine der möglichen Realisationen des stochastischen Prozesses, genauso wie eine einzelne Beobachtung im Rahmen eines Zufallsexperiments nur eine Realisation einer Zufallsvariablen ist. Da in der Wirtschaft wie auch in den meisten anderen

[1] Da in dieser Arbeit immer aus dem Zusammenhang klar ist, wann von der Zeitreihe und wann von der einzelnen Beobachtung die Rede ist, wird diese Unterscheidung in der Notation im weiteren fallen gelassen.

[2] An der Annahme eines DGP ist vor allem in der Ökonomie vielfach Kritik geäußert worden. Oberflächliche Kritik argumentiert, dass es in ökonomischen Zeitreihen häufig zu Strukturbrüchen kommt bzw. dass die Gesetzmäßigkeiten einem stetigen Wandel unterworfen sind. Diese Kritik kann leicht entkräftet werden, indem man zu einem entsprechend allgemeiner formulierten Modell übergeht (vgl. Abschnitt 3.4). Hingegen stellt Keuzenkamp (1995) das Konzept eines DGP als eine „Reifikation“ eines Denkkonstrukts grundsätzlich in Frage.

[3] Für eine grundlegende und aktuelle Einführung in stochastische Prozesse bieten sich unter anderem an: Brockwell und Davis (1991), Mills (1990), Harvey (1993) und Schlittgen und Streitberg (1997)

Anwendungsgebieten wiederholbare Experimente selten anzutreffen sind, steht die Zeitreihenanalyse vor dem Problem, aus nur *einer* Realisation auf den stochastischen Prozess schließen zu müssen. Dieses Problem ist nur mit Hilfe bestimmter Annahmen bezüglich des stochastischen Prozesses lösbar. Es sind dies die beiden Annahmen der Ergodizität und der Stationariät.

Ergodizität eines stochastischen Prozesses ist, grob gesprochen, dann gegeben, wenn die Stichprobenmomente für endlich lange Realisationen des Prozesses gegen die wahren Momente konvergieren, lässt man die Länge der Realisationen gegen unendlich gehen. Da diese Annahme naturgemäß nicht getestet werden kann, wird sie bei der Anwendung zeitreihenanalytischer Instrumente implizit vorausgesetzt.

Die Annahme der *Stationarität* wird in zwei verschieden starken Ausprägungen verwendet. Streng stationär ist ein Prozess, wenn seine Eigenschaften unabhängig von einer Verschiebung des Zeit-Ursprungs $t = 0$ sind. Das heißt, die gemeinsame Verteilung für jede beliebige Indexmenge $\{t_1, t_2, \ldots, t_m\}$ muss identisch mit der gemeinsamen Verteilung für $\{t_1 + k, t_2 + k, \ldots, t_m + k\}$ sein, wobei k eine beliebige Verschiebung entlang der Zeitachse ist. Die Annahme der schwachen Stationarität hingegen verlangt nur, dass die ersten und zweiten Momente unabhängig von einer Verschiebung entlang der Zeitachse sind. Es muss also für alle t und k gelten:

(i) $\mathrm{E}(x_1) = \mathrm{E}(x_2) = \ldots = \mathrm{E}(x_t) = \mu$,

(ii) $\mathrm{Cov}(x_1, x_{1+k}) = \mathrm{Cov}(x_2, x_{2+k}) = \ldots = \mathrm{Cov}(x_t, x_{t+k})$.

Es ist klar, dass der zweite Teil dieser Bedingung auch die Unabhängigkeit der Varianz von Verschiebungen entlang der Zeitachse einschließt: $\mathrm{Var}(x_1) = \mathrm{Var}(x_2) = \ldots = \mathrm{Var}(x_t) = \sigma^2$. Strenge Stationarität bringt immer schwache Stationarität mit sich. Hingegen kann von schwacher Stationarität nur dann auf starke Stationarität geschlossen werden, wenn zusätzlich der Prozess ein Gauß'scher Prozess ist, da eine Normalverteilung durch die ersten beiden Momente bereits vollständig definiert ist.

Aus dem in der Definition des Stationaritätsbegriffes verwendeten Begriff der Autokovarianz

$$\mathrm{Cov}(x_t, x_{t-k}) = \mathrm{E}((x_t - \mu)(x_{t-k} - \mu)) \tag{2.1}$$

leitet sich der Begriff der Autokorrelation ab:

$$\rho_k = \frac{\mathrm{Cov}(x_t, x_{t-k})}{(\mathrm{Var}(x_t)\mathrm{Var}(x_{t-k}))^{0.5}}. \tag{2.2}$$

Die Annahme der Stationarität impliziert, dass die Autokorrelation nur vom *Lag* (d.h. der zeitlichen Verzögerung) k abhängt. Jene Funktion, die jedem Wert k die Autokorrelation ρ_k zuordnet, heißt Autokorrelationsfunktion (ACF). Ein eng mit der Autokorrelationsfunktion verbundenes Konzept ist die Partielle Autokorrelationsfunktion (PACF), die jedem k die partielle Autokorrelation ϕ_{kk} zuordnet. Diese wird mit Hilfe der Autoregression der Ordnung k definiert:

$$x_t = \phi_{k1} x_{t-1} + \phi_{k2} x_{t-2} + \ldots + \phi_{kk} x_{t-k} + u_t, \tag{2.3}$$

wobei u_t ein stationärer Residualprozess ist. Auf diese Weise misst ϕ_{kk} die verbleibende Korrelation zwischen x_t und x_{t-k}, nachdem um die Korrelation zwischen x_t und den dazwischen liegenden Variablen $x_{t-1}, \ldots, x_{t-k+1}$ bereinigt worden ist. ACF und PACF sind für die Identifikation linearer Modelle ein hilfreiches Instrument, spielen aber auch im Rahmen nicht-linearer Modellierung eine gewisse Rolle.

Ein grundlegender Baustein komplizierterer stochastischer Prozesse ist der als „Weißes Rauschen" oder als (unkorrelierter) Zufallsprozess bezeichnete stochastische Prozess ε_t mit den Eigenschaften

(i) ε_t identisch verteilt,

(ii) $\mathrm{E}(\varepsilon_t) = 0$

(iii) $\mathrm{Cov}(\varepsilon_t, \varepsilon_{t-k}) = 0$ für alle $k \neq 0$.

Ein Zufallsprozess, der zusätzlich die Bedingung erfüllt, dass ε_t für alle t unabhängig verteilt ist, ist ein Reiner Zufallsprozess („Reines Weißes Rauschen"). Ein normalverteilter Zufallsprozess heißt Normalprozess oder Gaußprozess. Realweltliche ökonomische Zeitreihen, die einem (Reinen) Weißen Rauschen entsprechen, sind selten. Üblicherweise wird jedoch angenommen, dass der Fehlerprozess (auch: Innovationsprozess) in einem linearen Modell ein weißes Rauschen ist. Die strengere Annahme eines Reinen Weißen Rauschens oder eines Gaußprozesses als Fehlerprozess ist für die theoretische Herleitung der Eigenschaften der linearen Modelle und der praktischen Schätzung im allgemeinen nicht notwendig.

Schließlich gehören der Backshift-Operator (auch: Lag-Operator) B^k, der eine Verzögerung einer Zeitreihe x_t um k Zeiteinheiten, $B^k x_t = x_{t-k}$, bewirkt, und der Differenzenoperator $\nabla_k^m = (1 - B^k)^m$ zu den Konzepten, von denen im folgenden ständig Gebrauch gemacht wird.

Lineare stochastische Prozesse

Generell wird ein linearer Prozess dadurch definiert, dass er als unendlicher linearer Filter eines unabhängig und identisch verteilten Zufallsprozesses ε_t angeschrieben werden kann:

$$x_t - \mu_t = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \ldots = \sum_{j=0}^{\infty} \phi_j \varepsilon_{t-j}, \ \phi_0 = 1. \tag{2.4}$$

Hierbei bezeichnet μ_t linear deterministische Komponenten, etwa einen konstanten Mittelwert, $\mu_t = \mu$, saisonale Mittelwerte oder Trendkomponenten. Die Annahme der Unabhängigkeit von ε_t ist in diesem Zusammenhang wesentlich. In der folgenden Diskussion der linearen Modelle reicht jedoch wieder die Annahme der Unkorreliertheit von ε_t aus.[4] Wold'sche Dekompositionstheorem besagt, dass jeder schwach

[4] Harvey (1993, S. 15) verwendet eine eigentümliche Trennung zwischen den Begriffen (linearer) Prozess und (lineares) Modell. Er nennt einen Prozess linear, wenn er nach Gleichung (2.4)