

Wirtschaftsmathematik

Wirtschaftsstatistik

Ökonometrie

ARMA-Prozesse

Prof. Dr. Franz Seitz, Weiden / Dr. Benjamin R. Auer, Leipzig

Neben den formalen Grundlagen von ARMA-Prozessen (Autoregressive Moving Average Processes) wird hier diskutiert, wie sich diese Prozesse simulieren und damit für die Modellierung von Finanzmarktzeitreihen nutzen lassen.

I. Klassische Ökonometrie vs. Zeitreihenökonomie

In den Wirtschaftswissenschaften muss man häufig die Entwicklung einer ökonomischen Größe (z.B. eines Geldmarktzinssatzes) durch eine geeignete Funktion charakterisieren. Dabei kann ein **strukturelles Modell** verwendet werden, das die Größe in Abhängigkeit von anderen ökonomischen Größen (z.B. der Inflationsrate) beschreibt. Die Zeitreihenökonomie beschreibt einen anderen Weg. Hier geht man davon aus, dass sich die Entwicklung einer Variablen durch einen ihr zugrunde liegenden stochastischen Prozess beschreiben lässt. Wird unterstellt, dass sie einem **ARMA-Prozess** folgt, wird ihre Entwicklung durch ihre eigenen historischen Werte sowie aktuelle und vergangene zufällige Störgrößen bestimmt. Wird die Variable nur durch ihre Historie und eine kontemporäre Störgröße beschrieben, spricht man von einem reinen **AR-Prozess**. Erfolgt die Erklärung hingegen nur durch gegenwärtige und historische Störgrößen, liegt ein reiner **MA-Prozess** vor.

II. Grundbegriffe

Prozessbegriff

Der Grundgedanke der Zeitreihenökonomie besteht darin, eine Zeitreihe als Realisation einer Folge von Zufallsvariablen aufzufassen. Eine solche Folge heißt **stochastischer Prozess** und wird gewöhnlich mit $\{y_t\}_{t=1}^T = \{y_1, \dots, y_T\}$ abgekürzt. Jede einzelne Beobachtung einer Zeitreihe wird also als zufällig aus einer Wahrscheinlichkeitsverteilung gezogen betrachtet. Man kann auch sagen, dass eine Zeitreihe als Realisation aus einer gemeinsamen Verteilungsfunktion der T Prozessvariablen interpretiert wird.

Prozesscharakteristika

Ein stochastischer Prozess lässt sich allgemein durch seine **Prozessgleichung** und seine **Momentenfunktionen** konkretisieren. Erstere beschreibt genau die stochastische Natur des Prozesses, während Letztere daraus hervorgehen und so beschrieben werden können: Die Erwartungswertfunktion $E(y_t) = \mu_t$ eines Prozesses gibt als Funktion der Zeit an, mit welcher Realisation des stochastischen Prozesses zum Zeitpunkt t „im Mittel“ zu rechnen ist. Man kann sich $E(y_t)$ daher auch als durchschnittliche Realisation zum Zeitpunkt t über verschiedene mögliche Realisationen hinweg vorstellen. Analoges gilt für die Varianzfunktion $Var(y_t) = \gamma_{t,0}$ und die Autokovarianzfunktion $Cov(y_t, y_{t-k}) = \gamma_{t,k}$ mit $k = 1, 2, \dots$. Der Wert der Varianz wird hier mit $\gamma_{t,0}$ bezeichnet, da die Kovarianz einer Größe mit sich selbst ($k = 0$) gerade ihrer Varianz entspricht.

Stationarität

Zu den in der Statistik bedeutendsten Prozessen gehören die stationären stochastischen Prozesse. Ein stochastischer Prozess $\{y_t\}_{t=1}^T$ heißt allgemein stationär, wenn sich seine Eigenschaften im Zeitverlauf nicht ändern. Er ist **stark stationär**, wenn die gemeinsame Verteilungsfunktion einer Sequenz $y_t, y_{t+1}, \dots, y_{t+k}$ für alle t, k und s mit der der Sequenz $y_{t+s}, y_{t+1+s}, \dots, y_{t+k+s}$ identisch ist (vgl. Tsay, S. 25) bzw. die gemeinsame Verteilungsfunktion für beliebige Zeitpunkte nicht von einer beliebigen Verschiebung entlang der Zeitachse abhängig ist. Dies impliziert, dass die Verteilung von y_1 die gleiche ist wie für jedes andere y_t , und außerdem z.B. die Kovarianzen zwischen y_t und y_{t-k} für jedes k unabhängig von t sind. Da eine derart starke Bedingung empirisch schwer zu verifizieren ist, wird häufig angenommen, dass allein Mittelwert, Varianz und Autokovarianzen eines Prozesses zeitinvariant sind und nicht die gesamte Verteilung. Man spricht dann auch von schwacher Stationarität. Formal ist ein Prozess $\{y_t\}_{t=1}^T$ **schwach stationär**, wenn für alle t gilt:

- (1) $E(y_t) = \mu < \infty$,
- (2) $\text{Var}(y_t) = E[(y_t - \mu)^2] = \gamma_0 < \infty$,
- (3) $\text{Cov}(y_t, y_{t-k}) = E[(y_t - \mu)(y_{t-k} - \mu)] = \gamma_k$, für $k = 1, 2, \dots$

Die Bedingungen (1) und (2) fordern dabei, dass der Prozess einen konstanten und endlichen Erwartungswert und eine konstante und endliche Varianz besitzt. Die Bedingung (3), die von (1) Gebrauch macht, verlangt, dass die Autokovarianzen von y_t nur vom Abstand k der beiden betrachteten Zeitpunkte t und $t - k$ abhängen. Offenbar gilt bei schwacher Stationarität auch

$$\text{Cov}(y_t, y_{t-k}) = \text{Cov}(y_{t+s}, y_{t-k+s})$$

für alle s . Erwartungswert, Varianz und Autokovarianzen sind damit unabhängig vom Zeitpunkt, zu dem sie gemessen werden. Man sagt auch, sie sind **zeitinvariant** (vgl. Schmid/Trede, S. 114 f.).

Da die Autokovarianz von der Einheit abhängig ist, in der Variablen gemessen werden, greift man gewöhnlich auf die **Autokorrelation** zurück. Sie ist bei einem stationären Prozess durch

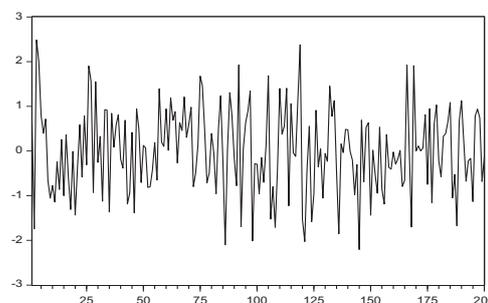
$$\text{Corr}(y_t, y_{t-k}) = \gamma_k / \gamma_0 = \rho_k \text{ mit } -1 \leq \rho_k \leq 1$$

gegeben, da Stationarität garantiert, dass $E(y_t) = \dots = E(y_{t-k})$ und $\text{Var}(y_t) = \dots = \text{Var}(y_{t-k})$, und stellt eine Funktion von k dar. Diese Funktion spielt eine zentrale Rolle bei der Veranschaulichung der durch einen konkreten Prozess entstehenden linearen Zusammenhänge zwischen zwei Zeitpunkten, die k Perioden voneinander getrennt sind.

Ist ein Prozess stark stationär (und sind seine ersten beiden Momente endlich), ist er auch schwach stationär. Umgekehrt gilt dies nicht generell. Unterstellt man jedoch als gemeinsame Verteilung eine Normalverteilung, so sind schwache und starke Stationarität äquivalent, da eine multivariate Normalverteilung schon durch die ersten beiden Momente eindeutig festgelegt ist (vgl. Verbeek, S. 273). Da es kaum vorkommt, dass ein Prozess schwach und nicht gleichzeitig auch stark stationär ist, ist eine Unterscheidung beider Formen von Stationarität empirisch kaum relevant (vgl. Greene, S. 636 ff.). Hier ist bei Verwendung des Begriffs Stationarität stets die schwache Stationarität gemeint. Es sei außerdem angemerkt, dass ein Prozess nur stationär sein kann, wenn die Parameter seiner Prozessgleichung nicht zeitvariabel sind (vgl. Pindyck/Rubinfeld, S. 493).

White-Noise

Der wohl wichtigste stationäre Prozess ist der **White-Noise-Prozess** $\{\varepsilon_t\}_{t=1}^T$, der eine Sequenz unabhängig und identisch verteilter („independent and identically distributed“, i.i.d.) Zufallsvariablen darstellt. Konkret wird für ihn unterstellt, dass jede Variable ε_t mit $E(\varepsilon_t) = 0$ und $\text{Var}(\varepsilon_t) = \sigma_\varepsilon^2$ verteilt ist und $E(\varepsilon_t \varepsilon_{t-k}) = 0$ für $k \neq 0$ gilt. Die Autokorrelationsfunktion des Prozesses ist damit für alle $k \geq 1$ gleich null. Sind die ε_t normalverteilte Zufallsvariablen, spricht man auch von einem **Gaußschen White-Noise-Prozess**. Die Unabhängigkeit der Momentenfunktionen von der Zeit ist beim White-Noise-Prozess unmittelbar zu erkennen. Im Gegensatz dazu wäre z.B. bei einem typischen nichtstationären Prozess wie dem **deterministischen Trend**, der $y_t = \alpha + \beta t + \varepsilon_t$ entspricht, klar ein zeitvariabler Erwartungswert $E(y_t) = \alpha + \beta t$ mit konstanter Varianz $\text{Var}(y_t) = \sigma_\varepsilon^2$ festzustellen.



EViews-Quellcode:
 genr yt = nrnd
 plot yt

Abb. 1: Simulierter Gaußscher White-Noise-Prozess

Abb. 1 zeigt eine typische Realisation eines Gaußschen White-Noise-Prozesses. Jeder der 200 Zeitreihenwerte wurde durch zufälliges Ziehen aus einer standardnormalverteilten Grundgesamtheit (Erwartungswert null, Varianz eins) bestimmt, wobei die einzelnen Züge unabhängig voneinander sind. Wie bei jedem stationären Prozesses ist deutlich zu

erkennen, dass die simulierten Beobachtungen in einer konstanten Bandbreite (Varianz) um ein festes Niveau (Erwartungswert) schwanken.

Frage 1: Warum ist ein White-Noise-Prozess als stationär zu betrachten?

III. Moving-Average-Prozesse

**Allgemeine
Prozessgleichung**

Folgt eine Variable y_t einem Moving-Average-Prozess der Ordnung $q \neq \infty$, kurz **MA(q)-Prozess**, entsteht jede Beobachtung y_t als gewichtete Summe zufälliger Störgrößen, die um bis zu q Perioden verzögert sind. Formal bedeutet dies

$$(4) \quad y_t = \delta + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q},$$

wobei δ und $\theta_1, \theta_2, \dots, \theta_q$ Parameter sind, die positive oder negative Werte annehmen können. Es wird angenommen, dass die ε_t einem White-Noise-Prozess folgen. Der MA(q)-Prozess wird damit genau durch die $q + 2$ Parameter $\theta_1, \theta_2, \dots, \theta_q, \delta, \sigma_\varepsilon^2$ beschrieben.

**Eigenschaften eines
MA(1)-Prozesses**

Zur Veranschaulichung der Eigenschaften von MA(q)-Prozessen empfiehlt sich ein Blick auf einen einfachen **MA(1)-Prozess**. Dieser hat die Prozessgleichung

$$(5) \quad y_t = \delta + \varepsilon_t + \theta_1 \varepsilon_{t-1}.$$

Der Erwartungswert und die Varianz dieses Prozesses lauten

$$(6) \quad E(y_t) = \delta,$$

$$(7) \quad \text{Var}(y_t) = \sigma_\varepsilon^2 (1 + \theta_1^2).$$

Dies kann relativ einfach bewiesen werden: Aufgrund der für alle t geltenden White-Noise-Eigenschaften $E(\varepsilon_t) = 0$, $\text{Var}(\varepsilon_t) = \sigma_\varepsilon^2$ und $E(\varepsilon_t \varepsilon_{t-k}) = 0$ für $k \neq 0$ ergibt sich nämlich

$$E(y_t) = E(\delta + \varepsilon_t + \theta_1 \varepsilon_{t-1}) = \delta + E(\varepsilon_t) + \theta_1 E(\varepsilon_{t-1}) = \delta$$

und

$$\begin{aligned} \text{Var}(y_t) &= E[(y_t - \delta)^2] = E[(\varepsilon_t + \theta_1 \varepsilon_{t-1})^2] = E(\varepsilon_t^2 + 2\theta_1 \varepsilon_t \varepsilon_{t-1} + \theta_1^2 \varepsilon_{t-1}^2) \\ &= E(\varepsilon_t^2) + 2\theta_1 E(\varepsilon_t \varepsilon_{t-1}) + \theta_1^2 E(\varepsilon_{t-1}^2) = \sigma_\varepsilon^2 (1 + \theta_1^2). \end{aligned}$$

Die Autokovarianz und -korrelation des Prozesses erhält man als

$$(8) \quad \text{Cov}(y_t, y_{t-k}) = \begin{cases} \theta_1 \sigma_\varepsilon^2 & \text{für } k = 1 \\ 0 & \text{für } k > 1 \end{cases},$$

$$(9) \quad \text{Corr}(y_t, y_{t-k}) = \begin{cases} \frac{\theta_1}{1 + \theta_1^2} & \text{für } k = 1 \\ 0 & \text{für } k > 1 \end{cases}.$$

Es gilt nämlich für $k = 1$ stets

$$\begin{aligned} \text{Cov}(y_t, y_{t-1}) &= E[(y_t - \delta)(y_{t-1} - \delta)] = E[(\varepsilon_t + \theta_1 \varepsilon_{t-1})(\varepsilon_{t-1} + \theta_1 \varepsilon_{t-2})] \\ &= E(\varepsilon_t \varepsilon_{t-1}) + \theta_1 E(\varepsilon_t \varepsilon_{t-2}) + \theta_1 E(\varepsilon_{t-1} \varepsilon_{t-1}) + \theta_1^2 E(\varepsilon_{t-1} \varepsilon_{t-2}) \\ &= \theta_1 \sigma_\varepsilon^2. \end{aligned}$$

Für $k > 1$ erhält man stattdessen bereits

$$\begin{aligned} \text{Cov}(y_t, y_{t-k}) &= E[(y_t - \delta)(y_{t-k} - \delta)] = E[(\varepsilon_t + \theta_1 \varepsilon_{t-1})(\varepsilon_{t-k} + \theta_1 \varepsilon_{t-k-1})] \\ &= E(\varepsilon_t \varepsilon_{t-k}) + \theta_1 E(\varepsilon_t \varepsilon_{t-k-1}) + \theta_1 E(\varepsilon_{t-1} \varepsilon_{t-k}) + \theta_1^2 E(\varepsilon_{t-1} \varepsilon_{t-k-1}) \\ &= 0. \end{aligned}$$

Die Autokorrelationskoeffizienten ergeben sich daraus über die in Abschnitt 2 beschriebene Autokorrelationsformel. Dies gilt auch für die Prozesse der Abschnitte 4 und 5.

Aus den Resultaten (6) bis (9) lässt sich zunächst schließen, dass es sich bei einem MA(1)-Prozess stets um einen stationären Prozess handelt, da sein Erwartungswert und seine Varianz — bedingt durch die Stationarität des White-Noise-Prozesses — konstant und auch seine Autokovarianzen zeitinvariant sind. Die Autokorrelationsfunktion zeigt bekanntlich, wie stark ein Prozesswert mit vorhergehenden Prozesswerten korreliert ist bzw. wie lange (und wie stark) ein zufälliger Schock (ε_t) im Prozess die Werte von y_t beeinflusst. Für einen MA(1)-Prozess gilt, dass seine Autokorrelationskoeffizienten für Verzögerungen von $k > 1$ den Wert null annehmen. Beobachtungen, die zwei oder mehr Perioden voneinander entfernt sind, sind also nicht mehr korreliert. Man sagt in diesem Zusammenhang auch, dass der Prozess ein Gedächtnis von nur einer Periode hat. Ein Schock beeinflusst y_t insgesamt nur in zwei Perioden, sodass sich der Prozess bei Fehlen neuer Schocks bereits nach zwei Perioden auf seinen Erwartungswert zurückbewegt.

Frage 2: Welchen Autokorrelationskoeffizienten erster Ordnung besitzt eine Größe y_t , die einem MA(1)-Prozess der Form $y_t = \varepsilon_t + 0,1\varepsilon_{t-1}$ folgt?

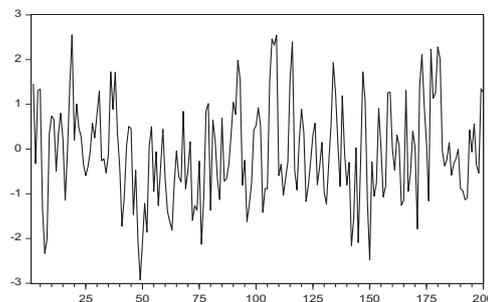
Eigenschaften von MA(q)-Prozessen

MA(q)-Prozesse höherer Ordnung ($q > 1$) besitzen mit MA(1)-Prozessen vergleichbare Eigenschaften. So sind auch hier Erwartungswert, Varianz und Autokovarianzen zeitinvariant. Zudem sind die Autokorrelationen für Verzögerungen bis q von null verschieden und für $k > q$ gleich null (vgl. Greene, S. 718 ff.; Pindyck/Rubinfeld, S. 522 ff.).

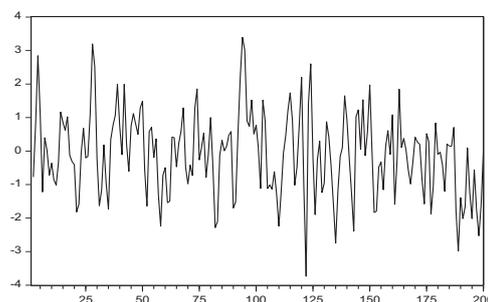
MA(1)-Prozesse grafisch

Abb. 2 zeigt zwei simulierte MA(1)-Prozesse für $\delta = 0$ und $\theta_1 = 0,5$ bzw. $\theta_1 = 0,9$. Die Simulation eines MA(1)-Prozesses ist dabei denkbar einfach (vgl. Davidson/MacKinnon, S. 573 ff.). Wie der Quelltext zeigt, muss dazu nur eine Zeitreihe von White-Noise-Realisationen (hier aus einer Standardnormalverteilung) erzeugt werden, die anschließend mittels der Prozessgleichung verbunden werden.

Beim optischen Vergleich beider Prozessbilder lässt sich kein starker Unterschied erkennen. Beide schwanken um null, wobei die Streuung aufgrund von (7) für $\theta_1 = 0,9$ leicht größer ausfällt als für $\theta_1 = 0,5$. Im Vergleich zum White-Noise-Prozess in Abb. 1 lässt sich feststellen, dass die simulierten MA(1)-Prozesse breiter streuen, was auf ihre höhere Varianz (7) zurückzuführen ist.



EViews-Quellcode:
 genr wn = nrnd
 genr yt = wn + 0.5*wn(-1)
 plot yt



EViews-Quellcode:
 genr wn = nrnd
 genr yt = wn + 0.9*wn(-1)
 plot yt

Abb. 2: Einige simulierte MA(1)-Prozesse

MA(1)-Prozesse in der Praxis

Die Prozessbilder in Abb. 2 geben Hinweise darauf, dass sich MA(1)-Prozesse potenziell zur Modellierung von **Aktienrenditen in ruhigen Marktphasen** eignen, da sich diese dann (bei ausreichend liquiden Aktien) häufig in gleichmäßiger Schwankungsbreite um ihren Mittelwert bewegen. Auch lässt sich insbesondere bei Tagesrenditen wertgewichteter Aktienindizes häufig beobachten, dass ihre Autokorrelationsfunktion für $k > 1$ nahe null liegt (vgl. Campbell/Lo/MacKinlay, S. 66 ff.). In Krisenphasen (wie etwa bei der jüngsten Finanzmarktkrise) ist eine solche Modellierung hingegen nicht geeignet, da sich hier sog. Volatilitätscluster bilden, d.h. Phasen mit äußerst hohen und/oder niedrigen Rendi-

ten im Vergleich zur sonstigen Renditeentwicklung auftreten. Eine solche Zeitvariabilität in der Varianz der Renditen kann durch einen stationären Prozess (mit zeitkonstanter Varianz) nicht abgebildet werden.

IV. Autoregressive Prozesse

Allgemeine Prozessgleichung

Folgt eine Variable y_t einem autoregressiven Prozess der Ordnung $p \neq \infty$, kurz **AR(p)-Prozess**, so entsteht der gegenwärtige Prozesswert y_t als gewichtete Summe historischer Prozesswerte aus bis zu p Perioden in der Vergangenheit und einer Störgröße ε_t der gegenwärtigen Periode. Dies bedeutet

$$(10) \quad y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t,$$

wobei δ und $\phi_1, \phi_2, \dots, \phi_p$ Parameter sind. Es wird angenommen, dass ε_t einem White-Noise-Prozess folgt und von den vergangenen y_{t-1}, y_{t-2}, \dots unabhängig ist. Anders als ein MA(q)-Prozess ist ein AR(p)-Prozess nicht generell stationär. Er ist dies nur, wenn die Nullstellen z_1, z_2, \dots, z_p seines charakteristischen Polynoms

$$f(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p$$

betragsmäßig größer als eins sind bzw. außerhalb des Einheitskreises liegen, d.h. $|z_i| > 1$ für $i = 1, \dots, p$ (vgl. Greene, S. 718 ff.).

Eigenschaften eines AR(1)-Prozesses

Um dies und weitere Eigenschaften von AR(p)-Prozessen zu veranschaulichen, empfiehlt sich ein Blick auf einen einfachen **AR(1)-Prozess**:

$$(11) \quad y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t.$$

Das dazugehörige charakteristische Polynom lautet $f(z) = 1 - \phi_1 z$ mit der Nullstelle $z = 1 / \phi_1$. Diese Nullstelle ist betragsmäßig nur dann größer als eins, wenn ϕ_1 betragsmäßig kleiner als eins ist. Ein AR(1)-Prozess ist also nur dann stationär, wenn $|\phi_1| < 1$ gilt. Der bekannteste nichtstationäre AR(1)-Prozess ist der **Random Walk mit Drift**. Er ist durch $\phi_1 = 1$, d.h. $y_t = \delta + y_{t-1} + \varepsilon_t$ charakterisiert.

Ist ein AR(1)-Prozess stationär, gilt für seinen Erwartungswert und seine Varianz

$$(12) \quad E(y_t) = \frac{\delta}{1 - \phi_1},$$

$$(13) \quad \text{Var}(y_t) = \frac{\sigma_\varepsilon^2}{1 - \phi_1^2}.$$

Die Herleitung dieser Ausdrücke ist relativ einfach. Wenn die Stationaritätsbedingung $|\phi_1| < 1$ erfüllt ist, ist der Erwartungswert im Zeitverlauf konstant. Dies impliziert $E(y_t) = E(y_{t-1})$ und es gilt

$$E(y_t) = E(\delta + \phi_1 y_{t-1} + \varepsilon_t) = \delta + \phi_1 E(y_{t-1}) = \delta + \phi_1 E(y_t)$$

bzw. äquivalent

$$E(y_t) = \delta / (1 - \phi_1).$$

Bei Stationarität folgt zudem eine konstante Varianz, d.h. $\text{Var}(y_t) = \text{Var}(y_{t-1})$. Wird zur Vereinfachung $\delta = 0$ angenommen und berücksichtigt man $E(y_{t-1} \varepsilon_t) = 0$, ergibt sich damit für die Varianz der Ausdruck

$$\begin{aligned} \text{Var}(y_t) &= E(y_t^2) = E[(\phi_1 y_{t-1} + \varepsilon_t)^2] = E(\phi_1^2 y_{t-1}^2 + 2\phi_1 y_{t-1} \varepsilon_t + \varepsilon_t^2) \\ &= \phi_1^2 \text{Var}(y_t) + \sigma_\varepsilon^2 \end{aligned}$$

bzw. äquivalent

$$\text{Var}(y_t) = \sigma_\varepsilon^2 / (1 - \phi_1^2).$$

Dasselbe Resultat gilt auch für $\delta \neq 0$, es ist dann allerdings mit höherem Rechenaufwand verbunden. Analoges gilt auch für die noch folgenden Beweise, in denen vereinzelt wieder auf diese Vereinfachung zurückgegriffen wird. Autokovarianz und -korrelation des Prozesses ergeben sich bei Stationarität aus

$$(14) \quad \text{Cov}(y_t, y_{t-k}) = \phi_1^k \text{Var}(y_t),$$

$$(15) \quad \text{Corr}(y_t, y_{t-k}) = \phi_1^k.$$

Bei erfüllter Stationaritätsbedingung und $\delta = 0$ ergibt sich nämlich für die Autokovarianz mit $k = 1$ bei Beachtung von $E(\varepsilon_t y_{t-1}) = 0$ zunächst

$$\begin{aligned} \text{Cov}(y_t, y_{t-1}) &= E(y_t y_{t-1}) = E[(\phi_1 y_{t-1} + \varepsilon_t) y_{t-1}] \\ &= \phi_1 E(y_{t-1}^2) + E(\varepsilon_t y_{t-1}) = \phi_1 \text{Var}(y_t). \end{aligned}$$

Bei der Berechnung der Autokovarianz mit $k = 2$ ist zu berücksichtigen, dass aus

$$y_t = \phi_1 y_{t-1} + \varepsilon_t \text{ und } y_{t-1} = \phi_1 y_{t-2} + \varepsilon_{t-1}$$

durch Substitution

$$y_t = \phi_1 (\phi_1 y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \phi_1^2 y_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t$$

folgt. Es gilt daher unter Beachtung von $E(\varepsilon_{t-1} y_{t-2}) = 0$ und $E(\varepsilon_t y_{t-2}) = 0$ allgemein

$$\text{Cov}(y_t, y_{t-2}) = E(y_t y_{t-2}) = E[(\phi_1^2 y_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t) y_{t-2}] = \phi_1^2 \text{Var}(y_t).$$

Analog dazu folgt durch weitere Substitution

$$\text{Cov}(y_t, y_{t-k}) = \phi_1^k \text{Var}(y_t).$$

Die Prozesseigenschaften (12) bis (15) lassen zunächst erkennen, dass die Varianz nur dann endlich und positiv ist, wenn $|\phi_1| < 1$ gilt. Darüber hinaus zeigt sich, dass die Autokorrelationsfunktion für einen AR(1)-Prozess eine relativ einfache Funktion von k ist, die bei positivem ϕ_1 geometrisch mit steigendem k sinkt. Bei negativem ϕ_1 ist hingegen ein oszillierendes geometrisches Abfallen zu beobachten. Der Prozess besitzt ein unbegrenztes Gedächtnis, da der gegenwärtige Wert mit allen vergangenen Werten korreliert ist, auch wenn das Ausmaß dieser Korrelation für weiter in der Vergangenheit liegende Werte (bzw. mit zunehmendem Abstand k der Werte) schwächer wird. Ein zufälliger Schock (ε_t) beeinflusst also alle künftigen Perioden, allerdings mit abnehmendem Effekt.

Eigenschaften von AR(p)-Prozessen

Für AR(p)-Prozesse höherer Ordnung ($p > 1$) lassen sich bei Stationarität vergleichsweise komplexere Ausdrücke für Erwartungswert, Varianz, Autokovarianz und -korrelation herleiten. Letztere nimmt auch hier mit steigendem k ab (vgl. Pindyck/Rubinfeld, S. 527 ff.; Tsay, S. 33 ff.).

AR(1)-Prozesse grafisch

Abb. 3 zeigt einige typische Realisationen von AR(1)-Prozessen mit $\delta = 0$ und $\phi_1 = 0,5$, $\phi_1 = 0,9$, $\phi_1 = 1$ bzw. $\phi_1 = 1,02$. Ihre Simulation wird im Vergleich zu MA(1)-Prozessen durch die Tatsache erschwert, dass y_0 bzw. ein y_t -Startwert des Prozesses festgelegt werden muss (vgl. Davidson/MacKinnon, S. 573 ff.). Dieser kann bei einem stationären AR(1)-Prozess mit dem Erwartungswert des Prozesses belegt werden. So wurde auch in Abb. 3 vorgegangen, d.h. konkret

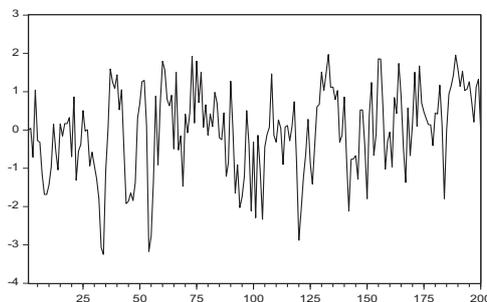
$$y_0 = E(y_t) = 0/(1 - \phi_1) = 0$$

verwendet. Bei nichtstationären Prozessen ist y_0 hingegen benutzerdefiniert festzulegen (hier: $y_0 = 0$). Die verbleibenden Beobachtungen werden dann rekursiv unter Hinzuziehung von White-Noise-Realisationen erzeugt. Soll bei einem stationären AR(1)-Prozess in wiederholten Simulationen ein stochastischer Startwert modelliert werden, so ist es sinnvoll, diesen aus der Verteilung der Prozessvariablen mit Erwartungswert (12) und Varianz (13) zufällig zu ziehen.

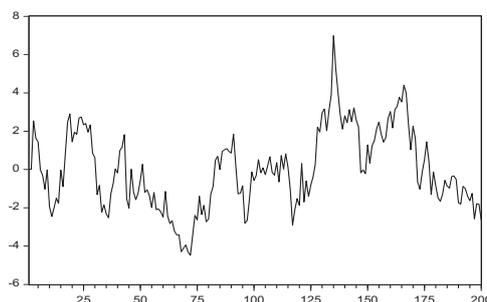
Frage 3: Nehmen Sie an, Sie möchten einen AR(1)-Prozess simulieren, der einen Erwartungswert $E(y_t) = 10$ besitzt und einen Parameter $\phi_1 = 0,5$ aufweisen soll. Auf welchen Wert müssen Sie δ festlegen?

Betrachtet man zunächst die stationären AR(1)-Prozesse ($\phi_1 = 0,5$, $\phi_1 = 0,9$), so unterscheiden sich diese dadurch, dass sich beim Prozess mit höherem ϕ_1 in stärkerem Maße „Tendenzen in eine Richtung“ erkennen lassen. Dies ist darauf zurückzuführen, dass bei höherem ϕ_1 aufgrund von (15) jede Beobachtung hochgradiger mit den umliegenden Beobachtungen korreliert ist. Da die Autokorrelationsfunktion bei höherem ϕ_1 erst für vergleichsweise höhere k gegen null geht, dauert es außerdem nach einem Schock länger, bis der Prozess auf seinen Erwartungswert zurückfindet. Der Prozess weist also gewisse

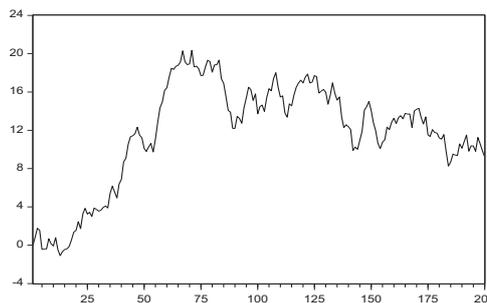
Persistenzen auf. Vergleicht man die stationären AR(1)-Prozesse mit den nichtstationären ($\phi_1 = 1, \phi_1 = 1,02$), lässt sich feststellen, dass die stationären Prozesse mit konstanter Varianz um ihren konstanten Mittelwert schwanken, d.h. die Tendenz besitzen, sich auf diesen zurückzubewegen (**Mean Reverting**). Ein solches Verhalten zeigt sich bei den nichtstationären Prozessen nicht, da z.B. für einen Random Walk mit Drift die Beziehungen $E(y_t) = y_0 + t\delta$ und $Var(y_t) = t\sigma_\varepsilon^2$ gelten (vgl. Schmid/Trede, S. 121 ff.).



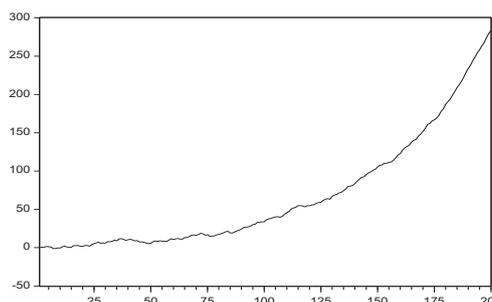
EViews-Quellcode:
 smpl 1 1
 genr yt = 0
 smpl 2 200
 genr yt = 0.5*yt(-1) + nrnd
 smpl 1 200
 plot yt



EViews-Quellcode:
 smpl 1 1
 genr yt = 0
 smpl 2 200
 genr yt = 0.9*yt(-1) + nrnd
 smpl 1 200
 plot yt



EViews-Quellcode:
 smpl 1 1
 genr yt = 0
 smpl 2 200
 genr yt = yt(-1) + nrnd
 smpl 1 200
 plot yt



EViews-Quellcode:
 smpl 1 1
 genr yt = 0
 smpl 2 200
 genr yt = 1.02*yt(-1) + nrnd
 smpl 1 200
 plot yt

Abb. 3: Einige simulierte AR(1)-Prozesse

**AR(1)-Prozesse
in der Praxis**

Die Prozessbilder, die sich in Abb. 3 für $\phi_1 = 0,9$ und $\phi_1 = 1$ zeigen, erinnern an typische Kursverläufe, wie man sie von Börsennachrichten kennt. In der Tat werden in der Forschung häufig AR(1)-Prozesse zur **Modellierung von Aktienkursen** herangezogen. Hier steht im einfachsten Fall häufig der nichtstationäre Random Walk ohne Drift $y_t = y_{t-1} + \varepsilon_t$, d.h. ein AR(1)-Modell mit $\delta = 0$ und $\phi_1 = 1$ (und zur Vermeidung negativer Kurse geeignet festgelegtem Startwert y_0), im Mittelpunkt des Interesses.

Die Verwendung eines derartigen Kursmodells lässt sich damit begründen, dass nach der Theorie informationseffizienter Finanzmärkte der aktuelle Kurs eines Wertpapiers stets alle auf dem Markt verfügbaren Informationen enthält, aufgrund derer die Aktionäre handeln und so den Kurs beeinflussen. Eine Kursänderung kann damit nur durch neue, unerwartete Informationen entstehen, die in ihrer Natur stochastisch sind. Der heutige Kurs y_t ergibt sich damit aus dem gestrigen y_{t-1} plus einer unerwarteten Kursänderung

ε_t . Die Tatsache, dass in einem solchen Modell implizit stationäre Kursänderungen bzw. Renditen unterstellt werden, machen es jedoch nicht für jede Marktphase realistisch.

Wie sich in der Empirie zeigt (vgl. Schmid/Trede, S. 147 ff.), eignen sich stationäre AR(1)-Prozesse besonders zur Modellierung der Entwicklung von (deutschen) **Zinsspreads**, d.h. der Differenz zwischen lang- und kurzfristigen Zinssätzen.

V. ARMA-Prozesse

Allgemeine Prozessgleichung

Viele stationäre Zufallsprozesse können nicht als reine MA- oder reine AR-Prozesse modelliert werden, da sie Eigenheiten beider Prozessgattungen aufweisen. Die logische Erweiterung der in den Abschnitten 3 und 4 präsentierten Prozesse ist daher ein gemischter Prozess, der üblicherweise als **ARMA(p, q)-Prozess** abgekürzt wird. Er besitzt die Form

$$(16) \quad y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

ARMA(p,q)-Prozesse haben die gleichen Stationaritätsbedingungen wie AR(p)-Prozesse, d.h. alle Nullstellen des charakteristischen Polynoms des AR-Teils müssen betragsmäßig größer als eins sein. Die $\theta_1, \theta_2, \dots, \theta_q$ unterliegen keinen Beschränkungen, was die Stationarität des Prozesses anbelangt.

Eigenschaften eines ARMA(1,1)-Prozesses

Zur Veranschaulichung der Prozesseigenschaften wird ein einfacher **ARMA(1,1)-Prozess** betrachtet:

$$(17) \quad y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

Wie sich zeigen wird, sind seine Eigenschaften Verallgemeinerungen des AR(1)-Prozesses mit einigen kleinen Modifikationen zur Berücksichtigung der MA(1)-Komponente.

Der Erwartungswert und die Varianz des Prozesses (17) sind bei Stationarität

$$(18) \quad E(y_t) = \frac{\delta}{1 - \phi_1},$$

$$(19) \quad \text{Var}(y_t) = (1 + \theta_1^2 + 2\phi_1\theta_1) \frac{\sigma_\varepsilon^2}{1 - \phi_1^2}$$

Im stationären Fall ($|\phi_1| < 1$) ergibt sich nämlich bei Berücksichtigung der Eigenschaften von White-Noise-Prozessen der Ausdruck

$$E(y_t) = E(\delta + \phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}) = \delta + \phi_1 E(y_{t-1})$$

bzw. äquivalent

$$E(y_t) = \delta / (1 - \phi_1)$$

Setzt man zur Vereinfachung $\delta = 0$, ergibt sich zudem

$$\begin{aligned} \text{Var}(y_t) &= E[(\phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1})^2] \\ &= \phi_1^2 \text{Var}(y_{t-1}) + \sigma_\varepsilon^2 + \theta_1^2 \sigma_\varepsilon^2 + 2\phi_1\theta_1 E(y_{t-1}\varepsilon_{t-1}) \end{aligned}$$

Daraus liefert ein stationaritätsbegründetes Einsetzen des Erwartungswertes der mit ε_t multiplizierten Prozessgleichung, d.h.

$$E(y_t \varepsilon_t) = E(\varepsilon_t^2) + \theta_1 E(\varepsilon_t \varepsilon_{t-1}) = E(\varepsilon_t^2) = \sigma_\varepsilon^2,$$

die Äquivalenzumformung

$$\begin{aligned} \text{Var}(y_t) - \phi_1^2 \text{Var}(y_{t-1}) &= (1 + \theta_1^2 + 2\phi_1\theta_1) \sigma_\varepsilon^2 \\ \Leftrightarrow \text{Var}(y_t) &= (1 + \theta_1^2 + 2\phi_1\theta_1) \sigma_\varepsilon^2 / (1 - \phi_1^2) \end{aligned}$$

Für Autokovarianz und -korrelation des Prozesses gilt bei Stationarität

$$(20) \quad \text{Cov}(y_t, y_{t-k}) = \begin{cases} \phi_1 \text{Var}(y_t) + \theta_1 \sigma_\varepsilon^2 & \text{für } k = 1 \\ \phi_1 \text{Cov}(y_t, y_{t-k+1}) & \text{für } k > 1 \end{cases}$$

$$(21) \quad \text{Corr}(y_t, y_{t-k}) = \begin{cases} \frac{(1 + \phi_1 \theta_1)(\phi_1 + \theta_1)}{1 + \theta_1^2 + 2\phi_1 \theta_1} & \text{für } k = 1 \\ \phi_1 \text{Corr}(y_t, y_{t-k+1}) & \text{für } k > 1 \end{cases}$$

Zur Bestimmung der Kovarianzen wird zunächst die Prozessgleichung (bei unterstelltem $\delta = 0$) mit y_{t-k} multipliziert, was

$$y_t y_{t-k} = \phi_1 y_{t-1} y_{t-k} + \varepsilon_t y_{t-k} + \theta_1 \varepsilon_{t-1} y_{t-k}$$

liefert. Die Bildung des Erwartungswertes dieses Ausdrucks liefert für $k = 1$ den konkreten Ausdruck

$$\begin{aligned} \text{Cov}(y_t, y_{t-1}) &= \phi_1 E(y_{t-1} y_{t-1}) + E(\varepsilon_t y_{t-1}) + \theta_1 E(\varepsilon_{t-1} y_{t-1}) \\ &= \phi_1 \text{Var}(y_t) + \theta_1 \sigma_\varepsilon^2 \end{aligned}$$

und für $k > 1$ den Term

$$\begin{aligned} \text{Cov}(y_t, y_{t-k}) &= \phi_1 E(y_{t-1} y_{t-k}) + E(\varepsilon_t y_{t-k}) + \theta_1 E(\varepsilon_{t-1} y_{t-k}) \\ &= \phi_1 \text{Cov}(y_{t-1}, y_{t-k}) = \phi_1 \text{Cov}(y_t, y_{t-k+1}), \end{aligned}$$

wobei die letzte Umformung auf die Stationarität des Prozesses zurückzuführen ist.

Wie sich aus den Prozesseigenschaften (18) bis (21) zeigt, entspricht der Erwartungswert des ARMA(1,1)-Prozesses dem des korrespondierenden AR(1)-Prozesses. Ihre Varianzen unterscheiden sich um den Faktor $(\theta_1^2 + 2\phi_1 \theta_1)$. Im Fall von $\phi_1 = 0$ ergibt sich aus (18) und (19) gerade (6) und (7) und im Fall $\theta_1 = 0$ gerade (12) und (13). Zudem lässt sich erkennen, dass die Autokorrelationsfunktion bei einem Wert beginnt, der sowohl von ϕ_1 als auch von θ_1 abhängt, und aufgrund ihrer Rekursivität anschließend einen fallenden Verlauf zeigt. Im Fall von $\phi_1 = 0$ ergibt sich aus (20) und (21) gerade (8) und (9) und im Fall $\theta_1 = 0$ gerade (14) und (15).

Eigenschaften von ARMA(p,q)-Prozessen

Bei ARMA(p,q)-Prozessen höherer Ordnung ($p > 1, q > 1$) ergeben sich vergleichsweise unübersichtlichere Ausdrücke für den Erwartungswert und die Varianz. Die ersten q Autokorrelationskoeffizienten sind komplexe Funktionen der AR- und der MA-Parameter. Für $k > q$ besitzt die Autokorrelationsfunktion rekursiven Charakter (vgl. Greene, S. 721 ff.) und zeigt einen fallenden Verlauf (vgl. Asteriou/Hall, S. 241 ff.).

ARMA(1,1)-Prozesse grafisch und in der Praxis

Der grafische Verlauf eines ARMA(1,1)-Prozesses wird primär von seiner AR-Komponente bestimmt und ähnelt daher den Prozessbildern in Abb. 3. Es zeigt sich jedoch, dass man durch Hinzufügen einer MA-Komponente in ein AR(1)-Modell (abhängig vom Vorzeichen der Prozessparameter) die Schwankungsbreiten und die Autokorrelation von y_t erhöhen oder reduzieren kann. (Interessierten Lesern stellen die Autoren auf Anfrage ein selbst entwickeltes Tool zur Verfügung, mit dem man spielerisch ein Gefühl für den Verlauf von ARMA(1,1)-Prozessen erlangen kann.) ARMA-Prozesse mit sehr kleinen Werten von p und q haben sich bei optimaler Anpassung der Parameter an die historische Entwicklung von **Zinsvariablen** als äußerst nützlich für ihre Prognose erwiesen (vgl. Greene, S. 717 ff.). Vielfältige Verallgemeinerungen sind möglich und werden in der Praxis angewandt (vgl. Tsay; Verbeek, S. 287 f.).

Frage 4: Ist ein ARMA(1,1)-Prozess, dessen Parameter θ_1 einen Wert größer als eins annimmt, stationär?

Literatur:

Asteriou, D./Hall, S.G.: Applied Econometrics — A Modern Approach. New York 2007.
 Campbell, J.Y./Lo, A.W./MacKinlay, A.C.: The Econometrics of Financial Markets. Princeton 1997.
 Davidson, R./MacKinnon, J.G.: Econometric Theory and Methods. New York/Oxford 2004.
 Greene, W.H.: Econometric Analysis. 6. Aufl., Upper Saddle River 2008.
 Hamilton, J.D.: Time Series Analysis. Princeton 1994.
 Pindyck, R.S./Rubinfeld, D.L.: Econometric Models and Economic Forecasts. 4. Aufl., Singapore 1998.
 Schmid, F./Trede, M.: Finanzmarktstatistik. Berlin et al. 2006.
 Tsay, R.S.: Analysis of Financial Time Series. 2. Aufl., Hoboken 2005.
 Verbeek, M.: A Guide to Modern Econometrics. 3. Aufl., Chichester 2008.

Die Fragen werden im WISU-Repetitorium beantwortet.