

Kointegrationsmodelle

Eine neue Technik zur Lösung von Regressionsproblemen

Dipl.-Volkswirt Jürgen Jerger, Freiburg

In jüngster Zeit hat eine neue ökonometrische Methode in der empirischen Wirtschaftsforschung eine rasche und weite Verbreitung gefunden: Die Technik der Kointegrationsmodelle. Mit Hilfe dieses Ansatzes sollen Probleme bei der Schätzung von Zusammenhängen zwischen nichtstationären Zeitreihen gelöst werden. Für die angewandte Forschung scheint damit ein Werkzeug zur Verfügung zu stehen, das nicht — wie bisher — entweder dem Vorwurf der Verletzung statistischer Annahmen ausgesetzt ist oder die Zeitreihen in ersten Differenzen in die Regression aufnehmen muß. Das Ziel dieses Beitrags ist es, die Grundidee und die Anwendung der Kointegrationstechnik mit einem Minimum an formalem Aufwand darzustellen.

1. Probleme bei der Schätzung mit nichtstationären Zeitreihen

Eine wichtige Voraussetzung des Klassischen Modells der linearen Regression (KMLR) ist die Stationaritätseigenschaft der Regressoren und der abhängigen Variable. In einer für konkrete Zeitreihen operationalen Begriffsbestimmung verlangt dies die Unabhängigkeit des Mittelwertes und der Varianz vom Stichprobenumfang T . Diese Voraussetzung ist jedoch bei ökonomischen Zeitreihen typischerweise nicht erfüllt. Wenn man sich die zeitliche Entwicklung beispielsweise von Sozialprodukt, Konsumausgaben oder Geldmenge vor Augen hält, ist zumindest die Verletzung der Annahme der Mittelwertkonstanz evident. Das Problem bei der Schätzung mit solchen Zeitreihen ist nun, daß sehr leicht „spurious regressions“, d.h. Scheinkorrelationsbefunde auftreten. In Abschnitt 2 werden wir in der Lage sein, dieses Phänomen begründen zu können.

Eine naheliegende Strategie, mit der das Problem der Nichtstationarität gelöst werden kann, ist die Verwendung der ersten Differenzen der jeweiligen Zeitreihen: $\Delta x_t \equiv x_t - x_{t-1}$ kann in den meisten Fällen als stationär gelten. Als Beispiel soll hier eine einfache Trendgerade $x_t = b \cdot t$ dienen. Während x_t linear ansteigt, (d.h. sowohl der Mittelwert als auch — wie zu zeigen sein wird — die Varianz permanent wachsen) ist $\Delta x_t = b \cdot t - b \cdot (t - 1) = b$ konstant, d.h. unabhängig vom Stichprobenumfang.

Problematisch bei dieser Vorgehensweise ist der dabei auftretende Informationsverlust bezüglich des Niveaузusammenhanges zwischen den Zeitreihen. Während die

Ursprungsreihen Informationen sowohl über das Niveau als auch über die Veränderungen beinhalten, geben die ersten Differenzen nur noch über die Veränderungen Auskunft. In unserem Beispiel ist $\Delta x_t = b$ kompatibel mit einer Schar von Trendgeraden: $x_t^* = a + b \cdot t$, $a \in \mathbf{R}$.

Die Kointegrationstheorie hat nun — als Lösung dieser Probleme — nachgewiesen, daß unter bestimmten Umständen nichtstationäre Zeitreihen im Rahmen des KMLR verwendet werden können, ohne daß die oben genannte Kritik zutrifft. Bevor auf die Konzeption der Kointegrationsmodelle und auf die empirischen Testverfahren eingegangen werden kann, müssen zunächst noch zwei theoretische Prozesse näher untersucht werden. Auf der Unterscheidung von (nichtstationärem) random walk und einem (stationären) Markov-Prozeß erster Ordnung bauen die empirischen Kointegrationstests auf.

2. Stationaritätseigenschaften von random walk und Markov (1)-Prozeß

Ein Markov (1)-Prozeß (siehe *Assenmacher*, 1984, S. 143 ff.) wird geschrieben als

$$x_t = a \cdot x_{t-1} + \varepsilon_t \quad |a| < 1. \quad (2.1)$$

Dabei ist ε_t ein white noise-Prozeß, mit einem Erwartungswert von Null und einer in allen Zeitpunkten gleichen Varianz.

Durch sukzessives Einsetzen höherer lags in (2.1) (z.B.

$$x_{t-1} = a \cdot x_{t-2} + \varepsilon_{t-1} \Rightarrow x_t = a^2 \cdot x_{t-2} + a \cdot \varepsilon_{t-1} + \varepsilon_t \Rightarrow E(x_t) = a^2 \cdot E(x_{t-2}))$$

erhält man folgende — mit (2.1) äquivalente — Schreibweise:

$$x_t = a^T \cdot x_0 + \sum_{s=0}^{T-1} a^s \varepsilon_{t-s}, \quad (2.2)$$

wobei x_0 der Startwert des Prozesses ist. T ist der zugrundeliegende Stichprobenumfang. Bilden wir den Erwartungswert von (2.2) und betrachten dessen Entwicklung für $T \rightarrow \infty$, so ergibt sich:

$$\lim_{T \rightarrow \infty} E(x_t) = \underbrace{\lim_{T \rightarrow \infty} a^T \cdot E(x_0)}_{= 0 \text{ wg. } \lim_{T \rightarrow \infty} a^T = 0} + \underbrace{E\left(\sum_{s=0}^{T-1} a^s \varepsilon_{t-s}\right)}_{= 0 \text{ wg. } E(\varepsilon_t) = 0 \forall t} = 0. \quad (2.3)$$

Damit ist gezeigt, daß der Erwartungswert eines Markov (1)-Prozesses – unabhängig vom Stichprobenumfang – asymptotisch konstant und gleich Null ist. Die Varianz

$$\text{Var}(x_t) \equiv E[x_t - E(x_t)]^2 = E(x_t)^2 = E[(\sum a^s \varepsilon_{t-s})^2] \quad (2.4)$$

kann man unter Berücksichtigung der angenommenen Homoskedastieeigenschaft von ε_t (d.h. $\text{Var}(\varepsilon_t) = \text{Var}(\varepsilon_{t-s}) \equiv \sigma_\varepsilon^2 \forall s$) und der Summenformel für eine unendliche geometrische Reihe schreiben als:

$$\text{Var}(x_t) = \frac{1}{1-a^2} \sigma_\varepsilon^2. \quad (2.5)$$

σ_ε^2 ist dabei die Varianz des white noise-Störprozesses. Aus (2.5) kann abgelesen werden, daß die Varianz eines Markov (1)-Prozesses nicht vom Stichprobenumfang, sondern ausschließlich von der Störvarianz sowie dem (konstanten) Parameter a abhängt. Damit ist nachgewiesen, daß ein stationärer Prozeß vorliegt.

Für einen random walk (siehe Schlittgen/Streitberg, 1989, S. 75), der durch

$$x_t = x_{t-1} + \varepsilon_t \quad (2.6)$$

gegeben ist, ergeben die Berechnung von Mittelwert und Varianz folgende Ergebnisse:

Durch sukzessives Einsetzen höherer lags in (2.6) erhält man:

$$E(x_t) = E\left(x_0 + \sum_{i=1}^t \varepsilon_i\right) = E(x_0) = x_0. \quad (2.7)$$

Der Erwartungswert ist also in jeder Periode gleich dem Startwert des Prozesses. Damit ist auch für einen random walk die Eigenschaft der Mittelwertstationarität nachgewiesen. Bildet man jedoch die Varianz eines random walk, so ergibt sich:

$$\text{Var}(x_t) = \sigma_\varepsilon^2 \cdot t. \quad (2.8)$$

Gleichung (2.8) impliziert, daß die Varianz proportional zum Stichprobenumfang wächst, d.h. ein random walk ist ein nichtstationärer Prozeß.

Mit diesem Wissen über die Entwicklung der Varianz eines random walk kann nun mit Hilfe der Gütestatistik R^2 sehr einfach gezeigt werden, wie sich eine Verletzung der Stationaritätsannahme im KMLR auswirkt.

Für eine Regression

$$x_t = \text{constant} + a \cdot y_t + z_t \quad (2.9)$$

(y_t ist ein Regressor zur Erklärung von x_t , z_t ist die Reststreuung)

ist das R^2 definiert als (siehe Assenmacher, 1984, S. 116)

$$R^2 = 1 - \frac{s_z^2}{s_x^2} \quad (2.10)$$

wobei s_z^2 bzw. s_x^2 die empirischen Varianzen von geschätztem Störprozeß und abhängiger Variable angeben. Folgt x_t einem random walk, so wird aus (2.10) unmittelbar deutlich, daß das R^2 gegen Eins gehen muß, weil der Nenner des zu subtrahierenden Bruches gegen Unendlich strebt. An dieser Stelle wird auch deutlich, warum die Mittel-

wertstationarität für eine korrekte Schätzung im KMLR erforderlich ist: **Jede** trendbehaftete Zeitreihe (z.B. auch die deterministische Trendgerade $x_t = 2 \cdot t$) hat eine im Zeitablauf steigende empirische Varianz:

$$s_x^2 = \frac{1}{T} \sum (x_t - \mu)^2, \quad \mu: \text{Mittelwert.}$$

D.h. selbst wenn von „Streuung“ im üblichen Sinne gar nicht die Rede sein kann, steigt die empirische Varianz. Dabei gilt, daß die Varianz mit dem Stichprobenumfang und der Steigung der Reihe wächst. Für das Beispiel zweier deterministischer Trendgeraden mit unterschiedlichen Steigungen soll das Gesagte in *Abb. 1* graphisch verdeutlicht werden. Die empirische Varianz ist die Summe der quadrierten Abweichungen vom Mittelwert μ , dividiert durch den Stichprobenumfang. Es ist evident, daß bei gleichem Stichprobenumfang die Varianz der linken Trendgeraden größer ist als die der rechten.

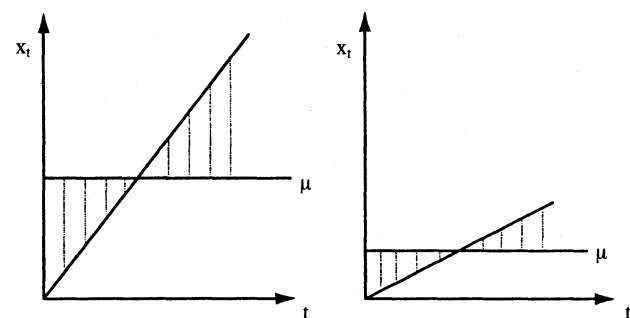


Abb. 1

Am Beispiel eines random walk läßt sich auch erläutern, wie ein nichtstationärer Prozeß durch die Bildung der ersten Differenz stationarisiert werden kann. Aus (2.6) ergibt sich sofort

$$\Delta x_t \equiv x_t - x_{t-1} = \varepsilon_t. \quad (2.11)$$

Das heißt, die erste Differenz eines random walk ist der als white noise angenommene und damit stationäre Störprozeß ε_t .

3. Das Konzept der Kointegration

Kointegrationsmodelle setzen nun genau an der besprochenen Eigenschaft der Nichtstationarität der meisten ökonomischen Zeitreihen an. Bestandteile eines Kointegrationsmodelles sind Zeitreihen mit jeweils identischen Integrationsordnungen. Die Integrationsordnung einer Zeitreihe ergibt sich dabei aus der für die Stationarisierung notwendigen Anzahl der Differenzenbildung. (Ist d-fache Differenzenbildung notwendig, so heißt die Reihe integriert von der Ordnung d: I (d).) Bei dem Beispiel des random walk ist — wie bei sehr vielen ökonomischen Zeitreihen — die Integrationsordnung gleich Eins (vgl. Gleichung (2.11)). Dieser — für die Praxis überaus bedeutsame — Spezialfall ist auch im folgenden Gegenstand der Betrachtung. Um die Darstellung noch weiter zu vereinfachen, wird der Fall betrachtet, daß genau zwei

Variablen in das Kointegrationsmodell eingehen. Für zwei Zeitreihen, x_t und y_t , die das Modell bilden, gilt dann:

x_t und y_t heißen kointegriert, wenn eine Linearkombination

$$x_t - a \cdot y_t = z_t \quad (3.1)$$

existiert, die gewährleistet, daß z_t stationär ist; a ist dabei der Kointegrationsparameter.

D.h. für die Kointegrationseigenschaft wird verlangt, daß die beiden $I(1)$ -Reihen x_t und y_t in der Lage sein müssen, eine $I(0)$ -Reihe zu erklären. Die Abweichungen der beiden Integrationsmuster müssen durch den Parameter a hinlänglich abgebildet werden können.

Für die ökonomische Verwertung relevant ist die Feststellung, daß bei Vorliegen von Kointegration (d.h. wenn z_t stationär ist) die Beziehung zwischen den Niveauwerten der Variablen x_t und y_t langfristig stabil ist. Diese Interpretation wird evident, wenn man sich verdeutlicht, daß die Abweichungen von x_t und y_t eine Größenordnung geringer sind als die Niveauentwicklungen selbst.

Der Reiz der Technik wird schon an dieser Stelle deutlich: Durch eine einfache OLS-Schätzung von (3.1) in der Form

$$x_t = a \cdot y_t + z_t \quad (3.2)$$

kann nun eine zuverlässige Aussage über die langfristige Beziehung von x_t und y_t gemacht werden. Bei Vorliegen von Kointegration trifft auch der Vorwurf der Verletzung statistischer Annahmen nicht mehr zu. *Stock* (1987) hat nachgewiesen, daß in kointegrierten Systemen die Schätzung von a — trotz der Nichtstationarität von x_t und y_t — konsistent ist. Bevor dies näher ausgeführt wird, sollen die Merkmale konsistenter Schätzer kurz in Erinnerung gerufen werden. Ein Schätzer ist konsistent, wenn

- (a) asymptotische Erwartungstreue und
- (b) mit wachsendem Stichprobenumfang verschwindende Varianz vorliegen. (Dies sind die hinreichenden Bedingungen für Konsistenz.)

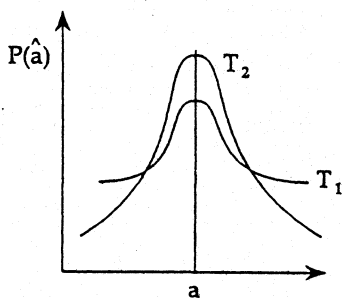


Abb. 2

In Abb. 2 ist die Verteilung eines (konsistenten) Schätzers \hat{a} um den wahren Wert a in Abhängigkeit vom Stichprobenumfang ($T_1 < T_2$) eingetragen. Konsistenz bedeutet hier, daß mit wachsendem Stichprobenumfang die Wahrscheinlichkeitsverteilung $P(\hat{a})$ über dem wahren Wert a

Soeben erschienen:
Neue PC-Software

Brink/Damhorst
Kramer/von Zwehl

Lineare und ganzzahlige Optimierung mit impac

Handbuch und Programmdiskette

Von Dr. Alfred Brink, Dipl.-Kfm. Hubert Damhorst, Dipl.-Kfm. Dominik Kramer und Prof. Dr. Wolfgang von Zwehl

1991. XII, 202 Seiten. Gebunden.
Mit einer 5¼-Zoll-Diskette nach dem Industriestandard unter MS-DOS.

DM 64,-
ISBN 3-8006-1586-X

Die lineare und die ganzzahlige Optimierung gehören zu den wichtigsten verfahrenstechnischen Instrumenten der Unternehmensplanung. Ein Einsatz dieser Planungsinstrumente ohne EDV-Unterstützung ist praktisch nicht denkbar. Allerdings ist die am Markt verfügbare PC-Software bislang nicht in der Lage, Problemstellungen mit Ganzzahligkeitsbedingungen zu lösen. Mit dem Programmpaket impac (integrated mathematical programming package) steht nun eine Software zur Lösung linearer und ganzzahliger Optimierungsprobleme für Personalcomputer bereit. Der praktische Einsatz des Programms wird durch eine menügesteuerte Benutzeroberfläche erleichtert, die dem Anwender eine schnelle und einfache Einarbeitung erlaubt. Das Handbuch enthält eine Einführung in die lineare und in die ganzzahlige Optimierung sowie eine ausführliche Beschreibung des Programms impac. Zahlreiche Anwendungsbeispiele verdeutlichen die Einsatzmöglichkeiten von impac. Handbuch und Programm richten sich nicht nur an Fachleute auf dem Gebiet der Unternehmensforschung, sie sind vielmehr so konzipiert, daß Interessenten anderer Fachrichtungen in die Anwendung der mathematischen Optimierung eingeführt werden.

Verlag Vahlen München

kollabiert, d.h. immer enger wird. Im KMLR konvergiert ein Schätzer \hat{a} mit der Rate $T^{-1/2}$ gegen den wahren Wert (Judge et al., 1988, S. 86). In kointegrierten Systemen gilt sogar, daß die Schätzungen mit der schnelleren Rate T^{-1} gegen die wahren Parameterwerte konvergieren („Superkonsistenz“). Wenn jedoch keine Kointegration zwischen x_t und y_t gefunden werden kann, muß bedacht werden, daß die Kritik bezüglich der Verwendung nichtstationärer Schätzer zum Tragen kommt. In diesem Fall sind die Parameterschätzungen inkonsistent, d.h. die errechneten Koeffizientenwerte und deren Varianzen können großen Verzerrungen unterliegen. Das *Stocks*che Ergebnis ist intuitiv zu begreifen, wenn man sich vor Augen hält, daß der (im Falle der Kointegration) stationäre Residuenterm z_t mit wachsendem Stichprobenumfang gegenüber den nichtstationären Reihen x_t und y_t immer unbedeutender wird. Daraus wird umgekehrt auch klar, warum die Superkonsistenz nur in kointegrierten Systemen vorliegt.

Die beschriebenen Anforderungen an die langfristige gemeinsame Entwicklung der beiden Zeitreihen werden in der Literatur oft als „Gleichgewichtsbeziehung“ bezeichnet, wobei die Abweichungen z_t dann „Gleichgewichtsfehler“ sind. Diese Begriffsbildung ist insofern etwas unglücklich, als „Gleichgewicht“ ein stark theoretisch vorbelasteter Terminus ist. Der Gehalt eines Kointegrationsgleichgewichts beschränkt sich auf die Aussage, daß die Abweichungen z_t der empirischen Zeitreihen von der langfristigen „Gleichgewichtsbeziehung“ $x_t = a \cdot y_t$ eine kleinere Integrationsordnung aufweisen als die Reihen selbst. In einem statistischen Sinn kann gesagt werden, daß die Linkhandvariable durch die in die Regression (3.2) aufgenommene Rechthandvariable „gut“ erklärt werden kann. Die Abweichungen sind — verglichen mit den Niveauwerten — gering.

Vorsicht ist jedoch auch im Rahmen dieser Methode bei der kausalen Interpretation, die ja eigentlicher Gegenstand des ökonomischen Interesses ist, geboten. Schon im Fall der Kointegration zweier Variablen sind drei Kausalinterpretationen möglich; bei der Aufnahme mehrerer Variablen in den Schätzansatz steigt die Anzahl der möglichen Kausalinterpretationen sehr rasch.

- a) $y_t \rightarrow x_t$ (y_t verursacht x_t)
- b) $x_t \rightarrow y_t$ (x_t verursacht y_t)
- c) $x_t \quad y_t$ (x_t und y_t hängen nicht voneinander ab, die
 $\uparrow \quad \uparrow$ beobachtete Kovariation entsteht durch
 $e_t \quad e_t$ eine unbekannt gemeinsame Ursache e_t .)

Die Ergebnisse von Kointegrationstests können nicht zwischen den drei Alternativen diskriminieren. Für diese Beurteilung müssen ergänzende ökonometrische Tests bzw. die ökonomische Theorie herangezogen werden.

Es klang bereits an, daß das Kointegrationskonzept nicht auf die Betrachtung von zwei Variablen beschränkt sein muß. Es ist — analog zu der gegebenen Darstellung des bivariaten Falles — auch möglich, multivariate Beziehungen zu analysieren. Ebenso muß die Integrationsordnung

der involvierten Zeitreihen nicht auf Eins beschränkt sein. Für die Kointegration eines solchen Systems werden zwei Anforderungen gestellt:

- (a) Alle Zeitreihen des betrachteten Systems müssen integriert von der gleichen Ordnung sein.
- (b) Die Integrationsordnung der Regressionsresiduen muß mindestens um Eins geringer sein als die Integrationsordnung der in den Ansatz aufgenommenen Variablen.

4. Liegt Kointegration vor? Empirische Testverfahren

An den beiden Anforderungen für das Vorliegen von Kointegration kann die Vorgehensweise der empirischen Testverfahren bereits abgelesen werden. Zunächst ist es erforderlich, die Integrationsordnung der in die sogenannte Kointegrationsregression (3.2) aufgenommenen Zeitreihen festzustellen. In einem zweiten Schritt wird dann geprüft, ob die Integrationsordnung der Residuen z_t tatsächlich kleiner ist als die der Variablen. Die bisher für diesen Zweck vorgestellten regressionsanalytischen Testverfahren versuchen nun, die Nullhypothese eines random walk gegen die Alternativhypothese eines Markov (1)-Prozesses zu testen. Damit kann überprüft werden, ob eine Zeitreihe $I(1)$ oder $I(0)$ ist. Die Literatur verwendet für diese Aufgabe bisher den *Sargan-Bhargava-Test* (Sargan/Bhargava, 1983), der die *Durbin-Watson-Statistik* heranzieht, sowie den *Dickey-Fuller-Test* (Dickey/Fuller, 1979, 1981) in einer einfachen und erweiterten Form. Diese Tests basieren jeweils auf einer gewöhnlichen t-Statistik.

Dargestellt wird hier die Vorgehensweise des *Sargan/Bhargava-Tests*. Der Grundgedanke ist bei den beiden anderen Tests jeweils derselbe.

Zunächst soll der Integrationsgrad einer Zeitreihe x_t festgestellt werden. Dazu wird die sog. Niveaugression

$$x_t = c + \xi_t \quad (c: \text{Konstante}) \quad (4.1)$$

geschätzt.

Analysiert werden dann die „Residuen“ ξ_t aus (4.1). Dies geschieht ausschließlich aus Gründen der Praktikabilität, weil in dieser Form eine *Durbin-Watson-Statistik* für ξ_t (bzw. x_t) mit jedem Regressionsprogramm leicht berechenbar ist; die Subtraktion einer Konstanten c ändert selbstverständlich nichts an der Stationarität/Nichtstationarität einer Zeitreihe. Die Untersuchung von ξ_t ist identisch mit der Untersuchung von x_t .

Die Nullhypothese der random walk-Eigenschaft von ξ_t (resp. x_t), wird gegen die Alternativhypothese, daß die Zeitreihe einem Markov(1)-Prozeß folgt, getestet. Relevante Teststatistik ist der *Durbin-Watson-Koeffizient* der Niveaugression (4.1). Um die Logik des Tests darzustellen, ist es nützlich, eine approximative Schreibweise zu wählen, die z.B. bei Assenmacher (1984), S. 147, abgeleitet wird:

$$DW \approx 2(1 - \hat{\rho}), \quad (4.2)$$

