

Zeitdiskrete vs. zeitstetige Modellierung von Preismechanismen zur Regulierung von Angebots- und Nachfragemengen

Dipl.-Kfm. Dipl.-Volksw. Thomas Mazzoni*

2004

Zusammenfassung

Dieser Beitrag vergleicht drei Modellvarianten für einen Preismechanismus, die sich im Hinblick auf die Behandlung der Variablen „Zeit“ unterscheiden. Zunächst wird ausgehend von einem originären Zeitreihenmodell ein Grenzübergang für unendlich kleine Zeitschritte durchgeführt, um zu zeigen, dass ein so gewonnenes zeitstetiges Modell völlig andere Eigenschaften aufweist, und auch ein anschließend extrahiertes exaktes diskretes Modell strukturell nicht mehr kompatibel zur Ausgangsform ist.

Anschließend wird mit Hilfe eines kontinuierlich/diskreten Zustandsraummodells eine Variante diskutiert, die die zu diskreten Zeitpunkten beobachteten Preise als Momentaufnahmen eines latenten zeitstetigen Prozesses (zum Beispiel einem Prozess der Meinungsbildung) interpretiert. Es wird dabei gezeigt, dass diese Vorgehensweise es nicht nur ermöglicht Werte zwischen den Beobachtungszeitpunkten zu inter- bzw. extrapolieren, sondern auch Aussagen über deren Genauigkeit zulässt.

Schlüsselbegriffe: Differenzengleichung; *Kalman*-Filter; Maximum Likelihood Schätzung; Stochastische Differentialgleichung; Zustandsraummodell.

1. Einleitung

Wirtschaftliche Transaktionen auf Märkten werden in der volkswirtschaftlichen Analyse traditionell als in der Zeit diskrete Handlungen betrachtet, die in vorzugsweise zyklischen Zeitabständen vorgenommen werden. Die Annahme von stark lokalisierten Märkten stützt diese Betrachtungsweise, da sie impliziert, dass ein Ausweichen auf andere Märkte in Perioden, in denen kein Handel erfolgt, ausgeschlossen ist.

Diese Modellvorstellung muss vor dem Hintergrund wachsender Markttransparenz und Transaktionsmobilität durch neue Medien, wie Internet und effiziente Logistiknetze, deren globaler Ausbau stetig voranschreitet, neu überdacht werden. Dabei ist der Übergang von einer zeitdiskreten Form der Modellierung zu einer zeitstetigen keineswegs so unproblematisch wie man vermuten könnte. In der Tat zeigt sich, dass die beiden Theorieebenen nicht beliebig gegeneinander ausgetauscht werden können, sondern dass grundlegende Überlegungen zur Struktur des untersuchten Modells in dessen Formulierung einfließen müssen.

*Wissenschaftlicher Mitarbeiter am Lehrstuhl für angewandte Statistik und Methoden der empirischen Sozialforschung (Univ.-Prof. Dr. Hermann Singer) der Fernuniversität in Hagen, Universitätsstraße 41, 58084 Hagen, Tel.: 02331/9872106, E-Mail: Thomas.Mazzoni@Fernuni-Hagen.de

2. Das Ausgangsmodell

Ein oft zitiertes Beispiel für die Nachfrage- und Angebotsmengenregulierung durch einen freien Preismechanismus zu diskreten Zeitpunkten ist das Agrarmarktbeispiel für Hühnereier (*Smith* [23] S. 4), bzw. Schweine, Weizen, etc. (*Woll* [24] S. 100 ff.). Die Nachfrage wird dabei als lineare, streng monoton fallende Funktion in Abhängigkeit des aktuellen Preises notiert. Die indizierte Zeitvariable t_1, \dots, t_T stellt dabei die Zeitpunkte, zu denen Transaktionen stattfinden dar, wobei die handelsfreien Zeitintervalle Δt als konstant betrachtet werden können. Im Weiteren wird die Notation $p(t_i) \equiv p_i$ verwendet. Für die Nachfragesituation gilt damit

$$q_i^d = \alpha - \beta p_i. \quad (1)$$

Die Anbieter sind hinsichtlich ihrer Disposition von Vorleistungen, im Beispiel Hühnerfutter, Stroh etc. und Produktionsmengen auf die Kalkulation mit einem Preisniveau angewiesen, das von der Nachfrage abhängt, die sich zum nächsten Handelstermin manifestiert. Darum stellt sich aus ihrer Perspektive die Preisbildung als Zufallsprozess dar, der idealisierter Weise lediglich von der letzten bekannten Ausprägung der Zufallsvariable „Preis“ abhängt.

Da die Produzenten, im Gegensatz zu den Konsumenten, grosse Mengen eines Produktes umsetzen, können sie sich Größensparnisse zu Nutze machen. Ähnliche Strategien auf der Verbraucherseite, durch Bündelung von Abnahmekantitäten, werden nicht betrachtet. Vernachlässigt man ferner die Fixkosten, was im Hinblick auf das zugrundeliegende Beispiel nicht zu unrealistischen Verzerrungen führt, lässt sich die Angebotsfunktion in der Form

$$q_i^s = -ap_{i-1} + bp_{i-1}^2 \quad (2)$$

formulieren¹. Der lineare Term in Gleichung (2) drückt die variablen Kosten aus, da über den Preis auch unmittelbar die Quantitäten bestimmt werden. Der parabolische Anteil bildet die „economies of scale“² ab.

Aufgrund der nicht vorhandenen Lagerfähigkeit der gehandelten Güter, muss die gesamte angebotene Menge auch abgesetzt werden, d.h. $q_i^d = q_i^s$, wodurch sich ein unreglementierter Preis einstellt, zu dem der Markt gerade geräumt wird. Gleichsetzen von (1) und (2) und Auflösen ergibt

$$p_i = \frac{\alpha}{\beta} + \frac{b}{\beta} p_{i-1} \left(\frac{a}{b} - p_{i-1} \right). \quad (3)$$

Mit Gleichung (3) liegt nun eine nichtlineare Differenzgleichung vor, deren abhängiger Teil als logistische Funktion identifiziert werden kann, die beispielsweise zur Charakterisierung von Populationswachstum bei begrenztem Lebensraum eingesetzt wird. Der Faktor $\frac{b}{\beta}$ stellt dabei den Wachstumsfaktor dar, während $\frac{a}{b}$ die maximale Population ausdrückt.

In diesem Fall stellt (3) eine dynamische Vorschrift der Preisentwicklung dar, die unter bestimmten Voraussetzungen gegen einen Gleichgewichtspreis konvergiert. Ein Gleichgewicht liegt genau dann vor, wenn die Preise der aufeinanderfolgenden Handelszeitpunkte identisch sind, d.h. wenn $p_i = p_{i-1}$ gilt. Bezeichne p^* den periodenunabhängigen gleichgewichtigen Preis, dann erhält man aus (3) durch Umstellen

$$p^* = \frac{a - \beta + \sqrt{(a - \beta)^2 + 4\alpha b}}{2b}. \quad (4)$$

¹Diese einfache Form gilt nur näherungsweise für einen bestimmten Bereich einer „normalen“ Produktionsfunktion, ist aber völlig ausreichend für den diskutierten Sachverhalt.

²Es wird an dieser Stelle keine Differenzierung zwischen Grössen- und Verbundersparnissen vorgenommen, sondern alle Vorteile einer Produktionsausweitung unter dem Begriff „economies of scale“ summiert.

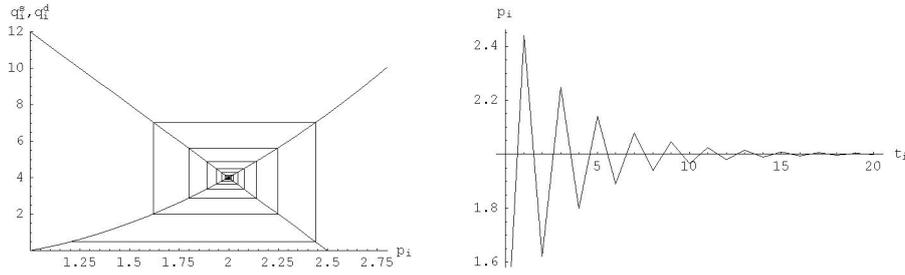


Abbildung 1: Trajektorie durch $p_0 = 1.2$ mit $a = 2$, $b = 2$, $\alpha = 20$, $\beta = 8$

Liegen bestimmte Parameterkonstellationen vor, kristallisiert sich ein stabiles Marktgleichgewicht heraus. Aus (3) und (4) erhält man die notwendige Bedingung

$$\beta^2 - \alpha b > \frac{(a - \beta)^2}{4} \quad (5)$$

für die Existenz eines stabilen Gleichgewichts. Abbildung 1 zeigt die Trajektorie durch p_0 als Projektion in der pq -Ebene, wobei die Charakteristik des 'Cobweb-Zyklus' deutlich wird und die durch Linien verbundenen Preise zu den Zeitpunkten t_i . Der stabile Fall wird in der Literatur als der Normalfall angenommen (Smith [23] S. 5, Woll [24] S. 102)³. An dieser Stelle besitzen aber auch alle anderen Konstellationen Relevanz, da beurteilt werden muss, zu welchem Grad ein Substitutionsmodell äquivalent ist.

Ist (5) beispielsweise nicht mehr erfüllt, bildet p^* keinen stabilen Gleichgewichtspreis mehr und es kann sogar zur Ausbildung von Grenzyklen kommen. Abbildung 2 zeigt verschiedene Trajektorien in der Umgebung des Schnittpunktes der Angebots- und Nachfragekurve. Offensichtlich stößt p^* jetzt alle Trajektorien ab, jedoch nur in einem lokalen Gebiet. Es bildet sich ein sog. periodischer Orbit.

Ökonomisch betrachtet würde ein solches Phänomen im Kontext der zugrundeliegenden Marktbeschreibung implizieren, dass die Produzenten sich ständig um einen identischen Angebotsbetrag verkalkulieren. Eine solche Situation stellt die gesamte Modellkonzeption in Frage, da die Anbieter zweifellos eine so einfache Systematik durchschauen und den theoretischen Handlungsrahmen durch Intervention durchbrechen würden.

Ferner ist für logistische Differenzgleichungen bekannt, dass sie unter bestimmten Parameterrestriktionen auch chaotische Funktionsverläufe generieren können, wobei streng zwischen chaotischen und zufälligen Systempfaden unterschieden werden

³Die Autoren gehen von einem linearen Modell des Anbieterverhaltens aus, das im Fall eines stabilen Gleichgewichts auch qualitativ gleiche Ergebnisse liefert.

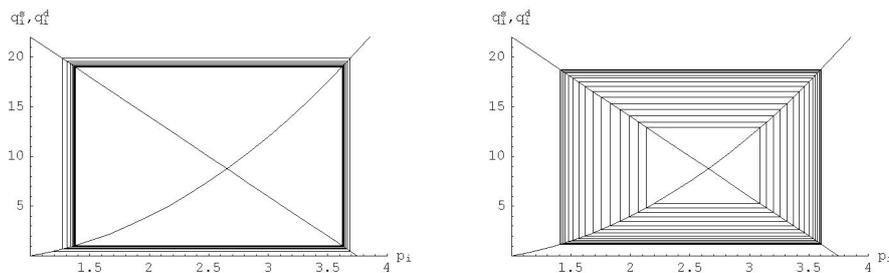


Abbildung 2: Trajektorien durch $p_0 = 1.2$ und $p_0 = 2.2$ für $a = 2$, $b = 2$, $\alpha = 30$, $\beta = 8$

muss. Diese Eigenschaften sind weder wünschenswert für die Modellbildung, noch werden sie in der Realität beobachtet. Offenbar gilt das Differenzenmodell nur in ganz bestimmten Bereichen.

3. Übergang zu stetiger Zeit

Die Differenz zwischen dem gegenwärtigen Preis und seiner Abbildung in die Zukunft kann mit Hilfe von (3) nach Umstellen in der Form

$$p_{i+1} - p_i = \frac{\alpha}{\beta} + \left(\frac{a - \beta}{\beta} \right) p_i - \frac{b}{\beta} p_i^2 \quad (6)$$

notiert werden. Bezeichne η_k für $k = 0, 1, 2$ die Koeffizienten von (6) und verhalten die Koeffizienten sich bei Veränderung von Δt proportional zum Zeitintervall, d.h. $\eta_k = \theta_k \Delta t$ für $\Delta t \rightarrow 0$, gewinnt man durch einen Grenzübergang die zur Differenzgleichung (3) korrespondierende Differentialgleichung

$$\dot{p} = \lim_{\Delta t \rightarrow 0} \frac{p_{t+\Delta t} - p_t}{\Delta t} = \theta_0 + \theta_1 p_t - \theta_2 p_t^2. \quad (7)$$

Gleichung (7) wird als autonome *Ricattische* Differentialgleichung bezeichnet (*Bronstein, Semendjajew, Musiol und Mühlig* [4] S. 508), die im Allgemeinen nicht durch Quadraturen gelöst werden kann. Dennoch lässt sich durch geschachtelte Substitution eine Vorschrift für $p(t)$ gewinnen⁴

$$p(t) = \tilde{p}^* + \frac{1}{\left(\frac{1}{p_0 - \tilde{p}^*} + \frac{\theta_2}{\tilde{\delta}} \right) \cdot e^{\tilde{\delta}t} - \frac{\theta_2}{\tilde{\delta}}} \quad (8a)$$

mit

$$\tilde{p}^* = \frac{\theta_1 + \tilde{\delta}}{2\theta_2}, \quad \tilde{\delta} = \sqrt{\theta_1^2 + 4\theta_0\theta_2}. \quad (8b)$$

Abbildung 3 auf der nächsten Seite zeigt das Verhalten der Trajektorien in stetiger Zeit. Im Gegensatz zum zeitdiskreten Fall findet kein Einschwingvorgang statt, sondern die Trajektorien relaxieren sofort zum Gleichgewichtspreis. Das ist auch unmittelbar einsichtig, da (8a) einen stetigen Produktions- und Absatzfluss beschreibt, in dem die Quantitäten instantan an den gegenwärtigen Preis angepasst werden können. Ein zeitliches „Lag“ wie in Gleichung (3) ist nicht vorhanden, bzw. ist von unendlich kurzer Dauer.

Um zu untersuchen, ob durch den Übergang von diskreter zu stetiger Zeit äquivalente Modellklassen entstanden sind, kann mit Hilfe des Ansatzes

$$p(t - \Delta t) = \tilde{p}^* + \frac{1}{\left(\frac{1}{p_0 - \tilde{p}^*} + \frac{\theta_2}{\tilde{\delta}} \right) \cdot e^{\tilde{\delta}(t - \Delta t)} - \frac{\theta_2}{\tilde{\delta}}} \quad (9)$$

⁴Durch die Substitution $p = \frac{z}{\theta_2} + \pi$ erhält man für den Koeffizienten der ersten Potenz von z den Ausdruck $\theta_1 - 2\theta_2\pi$, der verschwindet wenn $\pi = \frac{\theta_1}{2\theta_2}$ gesetzt wird. Somit ergibt sich die Normalform $\dot{z} = -z^2 + \frac{\theta_1^2 + 4\theta_0\theta_2}{4}$. Durch Umstellen und Auflösen erhält man für z zwei partikuläre Lösungen, $z_{1/2} = \pm \frac{\tilde{\delta}}{2}$. Durch eine weitere Substitution der Form $z = \frac{1}{u} + z_1$ erhält man die lineare DGL $\dot{u} - \tilde{\delta}u = 1$, die durch Einsetzen der partikulären Lösung $u_1 = \frac{1}{z_2 - z_1} = -\frac{1}{\tilde{\delta}}$ auf die allgemeine Lösung $u = -\frac{1}{\tilde{\delta}} + c_0 e^{\tilde{\delta}t}$ führt. Rücksubstitution ergibt die gesuchte Vorschrift $p(t) = \tilde{p}^* + \frac{1}{\theta_2 \left(c_0 e^{\tilde{\delta}t} - \frac{1}{\tilde{\delta}} \right)}$. Mit Hilfe der Anfangswertbedingung $p(0) = p_0$ erhält man $c_0 = \frac{1}{\theta_2(p_0 - \tilde{p}^*)} + \frac{1}{\tilde{\delta}}$ als Lösung für die Konstante und somit die allgemeine Lösung der Differentialgleichung.

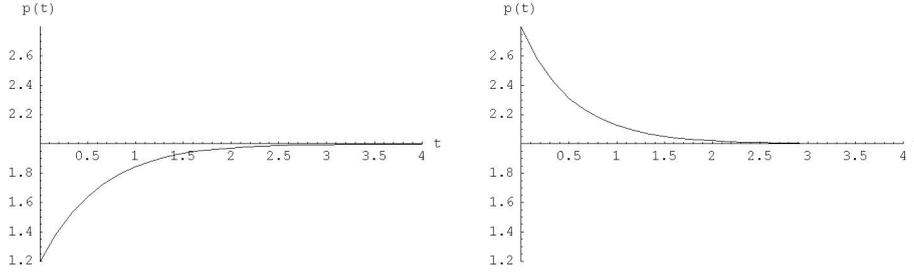


Abbildung 3: Trajektorien durch $p_0 = 1.2$ und $p_0 = 2.8$ für $\theta = (2.5, -0.75, 0.25)'$

durch Umformen und Einsetzen wieder ein exaktes diskretes Modell aus (8a) extrahiert werden. Wird die ursprüngliche Skalierung des beobachtungsfreien Zeitintervalls Δt gewählt, können die Koeffizienten θ_k rücksubstituiert werden, und man erhält die exakte zeitdiskrete Modellgleichung

$$p_i = p^* - \frac{1}{\left(\frac{1}{p^* - p_{i-1}} + \frac{b}{\beta\delta}\right) \cdot e^{\delta\Delta t} - \frac{b}{\beta\delta}}, \quad (10)$$

mit $\delta = \frac{\sqrt{(a-\beta)^2 + 4\alpha b}}{\beta}$. Gleichungen des Typs (8a) bzw. (10) werden ebenfalls als logistische Funktion (Schwarze [18] S. 130) oder logistische Sättigungskurve (Fahrmeir, Künstler, Pigeot und Tutz [6] S. 554) bezeichnet. Sie konvergieren exponentiell gegen einen Gleichgewichtswert, in diesem Fall p^* , und werden beispielsweise zur Berechnung der Halbwertszeit radioaktiver Isotope oder zur Beschreibung des Entladevorgangs von Kondensatoren eingesetzt.

Im Gegensatz zum ursprünglichen zeitdiskreten Modell (3), kann (10) keine oszillatorischen Verläufe erzeugen. Die durch beide Modelle generierten Abbildungen stimmen lediglich im Startwert p_0 überein und konvergieren gegen denselben Grenzwert p^* , wobei das ursprüngliche Modell alternierend von beiden Seiten konvergiert, das logistische jeweils nur von einer. Ferner scheint das abgeleitete exakte diskrete Modell schneller zu konvergieren als das originäre.

Überhaupt wirken sich alle Parameteränderungen in (10) lediglich auf die Konvergenzgeschwindigkeit aus. Um diese Eigenschaft einer ökonomischen Interpretation zuzuführen, ist es sinnvoll, die Angebots- und Nachfrageelastizitäten $\varepsilon(q, p) = \frac{\partial q}{\partial p} \cdot \frac{p}{q}$ aus (1) und (2) unter Vernachlässigung des temporalen Lags auszuwerten.

$$\varepsilon(q^d, p) = 1 - \frac{\alpha}{\alpha - \beta p} \quad (11a)$$

$$\varepsilon(q^s, p) = 2 + \frac{a}{bp - a} \quad (11b)$$

Vergleicht man diese Lösungen mit der Struktur von δ , sieht man leicht, dass die Konvergenzgeschwindigkeit steigt, wenn der Betrag der Preiselastizität von Angebot oder Nachfrage sinkt. Diese Bedingung ist ökonomisch einsichtig, da reziprok die Mengeneelastizität des Preises steigt und der Preis somit empfindlicher auf Mengenänderungen reagiert, bzw. sich schneller an gegebene Mengen anpasst. Dieser Schluss gilt jedoch nur für Bereiche positiver Angebotsmengen, da formal die Injektivität der Funktionen q^d und q^s gefordert werden muss, die für q^s jedoch nicht im gesamten Intervall $[0; \infty]$ gegeben ist.

Zusammenfassend kann gesagt werden, dass die beiden Modelle, die durch den Grenzübergang $\Delta t \rightarrow 0$ entstanden sind, signifikante strukturelle Unterschiede aufweisen. Im zeitstetigen Modell, resp. im daraus wieder abgeleiteten exakten diskreten Modell, treten nicht dieselben Effekte auf, wie im ursprünglichen zeitdiskreten Fall. Diese Diskrepanz ist nicht auf einen Modell- oder Spezifikationsfehler

zurückzuführen, sondern auf die Tatsache, dass ein völlig anderer Zusammenhang beschrieben wird. Bei vollkommener Markttransparenz und unendlich hoher Reaktionsgeschwindigkeit kommt es nicht zu Fehlbeurteilungen und den daraus abzuleitenden Unregelmäßigkeiten⁵. Diese Argumentation führt zu dem Schluss, dass ein zum Zeitreihenmodell analoges, zeitstetiges Pendant nicht durch beliebiges Verkleinern des Zeitintervalls zwischen den Beobachtungspunkten gewonnen werden kann. Des Weiteren scheint der perfekte Markt als Explorationsinstrument im Hinblick auf die zugrundeliegende Problematik völlig zu versagen, da seine Betrachtung nicht nur zu falschen Ergebnissen führt, sondern kategorisch ganze Ergebnisklassen ausschließt, die zwar beobachtet, aber durch die Theorie nicht erklärt werden können.

Der einzige konsistente Ausweg aus diesem Dilemma besteht in der Aufspaltung der formalen Betrachtung in ein mikroskopisches, auf zeitstetiger Ebene spezifiziertes Prozessmodell und ein makroskopisches, an diskreten Zeitpunkten definiertes Beobachtungsmodell. Diese Vorgehensweise eröffnet eine Reihe von weiteren Rahmenbedingungen, die vor allem zu einer allgemeineren und geschlossenen Theorie führen (für eine ausführliche Diskussion siehe *Bergstrom* [3] S. 1 ff.).

4. Der latente zeitstetige Prozess

Fasst man das ursprüngliche Differenzenmodell (3) als kontinuierlichen Prozess $p(t)$ auf, der nur zu den Zeitpunkten t_i , quasi in Form von Momentaufnahmen, beobachtet wird, $p_i = p(t_i)$, ergibt sich die Problematik, den nicht beobachtbaren Systemzustand zwischen den Messzeitpunkten modellieren zu müssen. Zu diesem Zweck werden einige Vorüberlegungen bezüglich der Preis- und Mengenreaktionen angestellt.

Steigt die gehandelte Gütermenge über die gleichgewichtige an, liegt ein Überangebot vor, das nur durch Preissenkung beseitigt werden kann. Liegt hingegen ein Preis vor, der den gleichgewichtigen Preis übersteigt, herrscht offenbar Übernachfrage, worauf die Produzenten die angebotene Gütermenge ausweiten werden.

Ein System von Mengen- und Preisänderungen kann somit in folgender Form formuliert werden

$$\dot{p} = -k_0^p(q - q^*) \quad (12a)$$

$$\dot{q} = k_0^q(p - p^*), \quad (12b)$$

wobei die Variable t nicht notiert wurde. Wird (12a) noch einmal nach der Zeit abgeleitet und (12b) eingesetzt ergibt sich

$$\ddot{p} = -k_0^p k_0^q (p - p^*), \quad (13)$$

was man als harmonischen Oszillator identifiziert. Abbildung 4 zeigt die Ellipsen im Phasenraum, der für (13) durch „Orts-“ und „Geschwindigkeits“-Variablen aufgespannt wird. Offenbar wird ein konservatives System beschrieben,

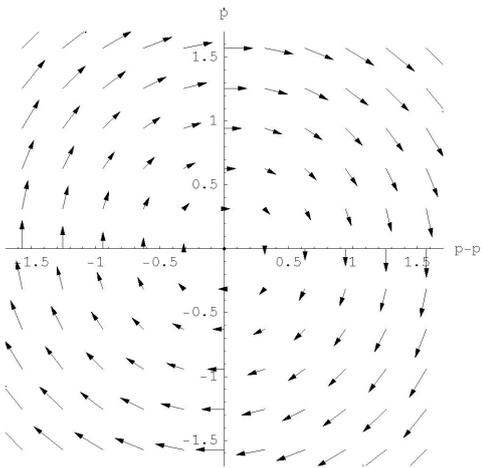


Abbildung 4: Phasenportrait

das unendlich lange mit der selben Amplitude schwingt, wenn der Startwert von p^* verschieden ist. Dasselbe gilt analog auch für die korrespondierenden Gütermengen. Diese Konstellation ist, wie weiter oben bereits diskutiert wurde, nicht realistisch. Die Marktteilnehmer werden entlang der zyklischen Schwankungen intervenieren, mit dem Ziel, die Differenz zwischen zukünftigen und gleichgewichtigen

⁵Diese Überlegung kann als latenter Faktor hinter dem Argument der Arbitragefreiheit interpretiert werden.

Werten zu minimieren. Aus diesem Grund ist es sinnvoll einen Dämpfungsterm zu modellieren, der analog zur physikalischen Reibung in das System (12a) und (12b) eingefügt wird. Ein solches System wird als phasenraumschrumpfend oder dissipativ bezeichnet, vgl. Abbildung 1 auf Seite 3, da die Abweichungen vom Gleichgewichtspreis, resp. der gleichgewichtigen Menge, im Zeitablauf kleiner werden. Da die physikalische Reibung sich proportional zur Geschwindigkeit verhält, wird auch der Dissipationsterm in Abhängigkeit von der ersten Zeitableitung, also der Phasengeschwindigkeit modelliert. Betrachtet man beispielsweise die inhomogene Differentialgleichung $\dot{p} = \lambda^p(p - p^*)$, die durch den Ansatz $p^* + (p_0 - p^*)e^{\lambda^p t}$ gelöst wird (Arnold [2] S. 51), erkennt man, dass für $\lambda^p < 0$ eine exponentielle Dämpfung beschrieben wird. Nimmt man λ^p als Funktion von p und q an, die durch lineare Parameterkombination in einen konstanten und einen abhängigen Teil zerlegt wird, kann der Dissipationsterm in der Form

$$\lambda^p(p, q) = - (k_1^p + k_2^p ((p - p^*)^2 + (q - q^*)^2)) \quad (14)$$

modelliert werden. Dasselbe gilt analog für den Mengenfall. Die quadratische Summation von Preis und Gütermenge hat nicht nur den Vorteil vorzeichenrestringierend zu wirken, sondern führt beim Übergang zu Polarkoordinaten zu einer sehr einfachen Form, wie noch gezeigt wird.

Wird weiterhin eine symmetrische Wirkung der Variablen aufeinander unterstellt, d.h. $k_j^p = k_j^q$ für $j = 0, 1, 2$, kann ein realistischeres Preis- Mengenänderungssystem als Summe aus Zuwachs und Dämpfung formuliert werden.

$$\dot{p} = -k_0(q - q^*) + \lambda(p, q) \cdot (p - p^*) \quad (15a)$$

$$\dot{q} = k_0(p - p^*) + \lambda(p, q) \cdot (q - q^*) \quad (15b)$$

Übergang zu Polarkoordinaten mit $p = p^* + r \cos \phi$ und $q = q^* + r \sin \phi$ ergibt das transformierte und entkoppelte Gleichungssystem

$$\dot{r} = -k_1 r - k_2 r^3 \quad (16a)$$

$$\dot{\phi} = -k_0. \quad (16b)$$

Diese Form ist erheblich kompakter als die Darstellung in kartesischen Koordinaten. Ferner ist die Winkelgeschwindigkeit aus (16b) bereits bekannt. Da das ursprüngliche Modell im Zeitraum Δt alternierende Werte generiert hat, gilt $\dot{\phi} = \frac{\pi}{\Delta t}$.

Da in der Realität ferner nie derart glatte Funktionsverläufe auftreten, ist es sinnvoll, das abgeleitete System durch Einfügen einer Zufallsfluktuation zu ergänzen. An dieser Stelle erweisen sich auch frühere Annahmen als notwendig. Ginge man beispielsweise von einem konservativen *Hamiltonschen* System aus, würde also kein Reibungsterm modelliert, würde die ständige stochastische „Energiezufuhr“ das System quasi sprengen, d.h. die Systemvariablen würden über alle Grenzen wachsen können. Erst die Formulierung eines dissipativen Modells ermöglicht stationäre Lösungen trotz stochastischer Systemgleichungen.

Die Verbesserung des Modells durch die Ergänzung eines stochastischen Terms hat aber auch ihren Preis. Zum Einen ist die stochastische Differentialgleichung nur über ihre korrespondierende Integralgleichung lösbar, da ein zufälliger Prozess, wie beispielsweise der von *Wiener* beschriebene (vgl. z.B. *Singer* [21] S. 22), nicht differenzierbar ist, und sich diese Eigenschaft auf die gesamte Gleichung ausdehnt. Zum Anderen erhält man als Lösungen Funktionswerte $r(t, \omega)$, die im Gegensatz zum deterministischen Modell zusätzlich von Zufallsfluktuationen abhängen.

5. Das kontinuierlich/diskrete Zustandsraummodell

Zustandsraummodelle werden in den Ingenieurwissenschaften vielfach eingesetzt, da sie auf die Problematik eines latenten nicht beobachtbaren Prozesses, der nur durch diskrete Messungen abgebildet werden kann zugeschnitten sind. Das Zustandsraum- oder State Space Modell wird durch zwei Gleichungen gegeben, die Zustandsgleichung, die den latenten Prozess beschreibt, und die Messgleichung, die den funktionalen Zusammenhang mit den Beobachtungswerten herstellt. Auf diese Weise kann das ursprüngliche Differenzgleichungsmodell als Hybrid aus zeitdiskreten Beobachtungswerten und dem im letzten Abschnitt hergeleiteten, latenten zeitstetigen Prozess formuliert werden

$$dr(t) = f(r(t))dt + g(t)dW(t) \quad t \in [t_0, t_T] \quad (17a)$$

$$p(t_i) = p^* + r(t_i) \cdot \cos(\pi i) \quad i = 0, \dots, T \quad (17b)$$

mit $f(r(t)) = -k_1 r(t) - k_2 r(t)^3$. Die Mess- oder Beobachtungsgleichung (17b) stellt lediglich die Rücktransformation der latenten Zustandswerte in kartesische Koordinaten dar, da keine Messfehler berücksichtigt werden. Die Schreibweisen $dr(t)$, dt und $dW(t)$ sind als Kurzform für die korrespondierende Integralgleichung aufzufassen. Insbesondere ist $r(t)$ nicht differenzierbar, da $W(t)$ den *Wiener*-Prozess darstellt, der durch zufällige, unabhängige Zuwächse charakterisiert ist⁶.

Die Aufgabe der Parameterschätzung kann mit Hilfe des Maximum-Likelihood Verfahrens (*Singer* [20] S. 113 ff.) gelöst werden. Die ML-Schätzung ist asymptotisch erwartungstreu, konsistent und effizient (*Greene* [10] S. 476 ff., *Singer* [20] S. 89), benötigt jedoch die Verteilungsfunktion der Zufallsvariablen.

Es kann gezeigt werden, dass im Falle der Existenz einer eindeutigen Lösung der stochastischen Differentialgleichung die Lösung durch einen *Markov*-Prozess gegeben ist (für einen ausführlichen Beweis siehe *Arnold* [1] S. 118 ff., 158 ff.). Weiterhin kann gezeigt werden, dass ein Diffusionsprozess⁷ existiert, der die stochastische Differentialgleichung (17a) löst. Die zugehörige Diffusionsgleichung für die Wahrscheinlichkeitsdichte $\pi(r, t)$ ist mit der *Fokker-Planck*-Gleichung gegeben (*Arnold* [1] S. 56 ff., *Friedrich* und *Lange* [8] S. 92 ff., *Haken* [11] S. 171 ff., *Risken* [17] S. 63 ff., *Singer* [20] S. 44 f., [21] S. 48 ff.)

$$\dot{\pi}(r, t) = -\frac{\partial}{\partial r} [f(r)\pi(r, t)] + \frac{1}{2} \frac{\partial^2}{\partial r^2} [\Omega\pi(r, t)], \quad (18)$$

wobei $\Omega = g^2$ gilt⁸. Die gesuchte Verteilung kann durch Lösung der *Fokker-Planck*-Gleichung gewonnen werden, was jedoch nur im Ausnahmefall analytisch gelingt (*Risken* [17] S. 7). Im Fall einer linearen Driftfunktion und eines konstanten Diffusionsterms erhält man unabhängig von der Stationarität des Prozesses eine *Gauß*-Verteilung.

⁶Handelt es sich bei dem Diffusionsterm $g(t)$ um eine Konstante oder eine nicht stochastische Funktion, kann das stochastische Integral $\int_{t_0}^t g(s)dW(s)$ als gewöhnliches *Riemann-Stieltjes*-Integral aufgefasst und im herkömmlichen Sinne als Summengrenzwert ausgewertet werden (*Singer* [20] S. 34). Stellt g eine stochastische Funktion dar, hängt der Grenzwert der Summe von der Lokalisation der Zwischenpunkte ab, wodurch das stochastische Integral nur im Rahmen der *Itô*- oder *Stratonovich*-Interpretation einer Lösung zuzuführen ist (*Arnold* [1] S. 72 ff., *Gard* [9] S. 35 ff., *Singer* [21] S. 27 ff.)

⁷Diffusionsprozesse stellen spezielle *Markov*-Prozesse mit stetigen Realisierungen dar.

⁸Handelt es sich bei g um eine stochastische Funktion, gilt die Äquivalenz der Terme nur für die *Itô*-Lösung des stochastischen Integrals. Die *Stratonovich*-Lösung kann jedoch generell mit Hilfe der *Wong-Zakai*-Korrektur (*Arnold* [1] S. 180, *Singer* [20] S. 41) in die *Itô*-Lösung überführt werden. Für nicht stochastische g stimmen beide Interpretationen überein.

Mit Hilfe der *Fokker-Planck*-Gleichung gelingt es ebenfalls, Gleichungen für die zeitabhängigen Momente abzuleiten. Für das erste Moment gilt

$$\begin{aligned}\dot{\mu}(t) &= \int r \dot{\pi}(r, t) dr \\ &= \int r \left(-\frac{\partial}{\partial r} [f(r(t))\pi(r, t)] + \frac{1}{2} \frac{\partial^2}{\partial r^2} [\Omega\pi(r, t)] \right) dr \\ &= \int f(r(t))\pi(r, t) dr = E[f(r(t))].\end{aligned}\quad (19a)$$

Die Ableitungen wurden dabei durch partielle Integration auf r abgewälzt, wobei ausgenutzt wurde, dass alle Randterme für $|r| \rightarrow \infty$ gegen Null gehen. Für das 2. Moment erhält man auf analogem Weg

$$\dot{\sigma}^2(t) = 2 \operatorname{cov}[f(r(t)), r(t)] + E[\Omega]. \quad (19b)$$

Es handelt sich bei (19a) und (19b) formal nicht um Differentialgleichungen, da auf der rechten Seite Erwartungswerte stehen. Trotzdem liefern sie zwei wichtige Zusammenhänge. Zum einen zeigt (19a), dass der Erwartungswert der Lösung nicht zwingend die deterministische Gleichung $dr = f(r(t))dt$ erfüllt, die durch Weglassen des Fluktuationsterms gebildet wird, zum anderen wird durch eine *Taylor*-Entwicklung um den Erwartungswert klar, dass nur für lineare Driftfunktionen geschlossene Gleichungen entstehen. Nichtlineare Konvektionsterme erzeugen somit keine *Gauß*-schen Übergangswahrscheinlichkeiten.

Man gewinnt aber dennoch approximativ geschlossene Gleichungen, wenn die *Taylor*-Reihe nur linear entwickelt wird. Hängt Ω nicht vom Zufall ab, ergibt sich mit $\frac{\partial}{\partial r} f(r) = A(r)$

$$\dot{\mu} \simeq f(\mu) \quad (20a)$$

$$\dot{\sigma}^2 \simeq 2A(\mu)\sigma^2 + \Omega, \quad (20b)$$

wobei die Zeitvariable zur Vereinfachung nicht notiert wurde. Durch diese Näherung lassen sich auch die Übergangswahrscheinlichkeiten für nichtlineare Driftfunktionen approximativ als *Gauß*-sch behandeln, was speziell für die Parameterschätzung vorteilhaft ist. Dabei wird ausgenutzt, dass alle endlich dimensionalen Verteilungen des Systemzustands mit Hilfe der *Bayes*-Formel $P(r_0, r_1, r_2) = P(r_2|r_1, r_0)P(r_1|r_0)P(r_0)$ und der *Markov*-Eigenschaft $P(r_i|r_{i-1}, \dots, r_0) = P(r_i|r_{i-1})$ als Produkt von Übergangswahrscheinlichkeiten geschrieben werden können

$$P(r_0, \dots, r_T) = \prod_{i=1}^T P(r_i|r_{i-1})P(r_0). \quad (21)$$

Sei $\boldsymbol{\theta}$ der Parametervektor, $p_{t|i} = p^* + \mu_{t|i} \cdot \cos(\pi t)$ der prognostizierte Beobachtungswert zum Zeitpunkt t , gegeben die Beobachtung zum Zeitpunkt t_i und $\sigma_{t|i}^2$ der mittlere quadratische Prognosefehler zum Zeitpunkt t , gegeben t_i , dann können die Beobachtungswerte p_i ebenfalls approximativ als bedingt normalverteilte Zufallsvariablen mit $E[p_i|p_{i-1}] = p_{i|i-1}$ und $E[(p_i - p_{i|i-1})^2|p_{i-1}] = \sigma_{i|i-1}^2$ aufgefasst werden. Für bekanntes p_0 kann dann die Likelihood-Funktion in der Form

$$L(\boldsymbol{\theta}; p_0, \dots, p_T) = \prod_{i=1}^T \frac{1}{\sqrt{2\pi\sigma_{i|i-1}^2}} \cdot e^{-\frac{1}{2} \frac{(p_i - p_{i|i-1})^2}{\sigma_{i|i-1}^2}} \quad (22)$$

geschrieben werden. Die Schätzwerte für den Parametervektor können durch Maximierung der Likelihood-Funktion gefunden werden, $\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} L(\boldsymbol{\theta}; \mathbf{p})$. Die für

die Berechnung benötigten bedingten Momente werden mit Hilfe von (20a) und (20b) gewonnen.

$$\mu_{t|i} = \mu_{i|i} + \int_{t_i}^t f(\mu_{\tau|i}) d\tau \quad (23a)$$

$$\sigma_{t|i}^2 = \int_{t_i}^t e^{2 \int_{\tau}^t A(\mu_{s|i}) ds} g^2 d\tau \quad (23b)$$

Gleichung (23b) ergibt sich als allgemeine Lösung der linear-inhomogenen Differentialgleichung (20b) unter der Anfangsbedingung $\sigma_{i|i}^2 = 0$ (Arnold [2] S. 51). Damit lässt sich der Parametervektor θ unter Verwendung der Maximum-Likelihood Methode schätzen.

Diese sequentielle Berechnung der Likelihood-Funktion entspricht der Arbeitsweise des *Kalman*-Filters (Kalman [14]). Es wäre ebenso möglich, die Dichtefunktion der Beobachtungswerte $\pi(p_0, \dots, p_T) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, basierend auf dem Satz über die Normalkorrelation (Fahrmeir [7] S. 27, Singer [20] S. 136 f.) explizit anzugeben, jedoch muss in diesem Fall mit einem Erwartungswertvektor $\boldsymbol{\mu}$ der Länge $T \times 1$ und einer Kovarianzmatrix $\boldsymbol{\Sigma}$ der Dimension $T \times T$ gerechnet werden. Aus diesem Grund ist man für alle praktischen Anwendungen daran interessiert ein Verfahren einzusetzen, das die benötigten Likelihood-Terme sequentiell berechnet und gleichzeitig eine Echtzeitprognose ermöglicht.

Ein solcher Algorithmus liegt für lineare Differentialgleichung mit dem *Kalman*-Filter vor, bzw. für nichtlineare Zustandsgleichungen mit dem Extended *Kalman*-Filter (EKF), für den die approximativen Momente (20a) und (20b) als Ausgangsgleichungen fungieren (Harvey [12] S. 104 ff., Jazwinski [13] S. 162 ff., S. 340 ff., Kalman [14], Singer [20] S. 143 ff., [21] S. 86 ff., S. 169 ff., [22]). Die Likelihood-Funktion kann dabei als Prognose-Fehler-Zerlegung berechnet werden (Schweppe [19], Singer [20] S. 113 f.). Der *Kalman*-Filter kann für zukünftige Werte eine Prognose liefern, Echtzeitdaten filtern und für Werte vor dem letzten Messupdate als Glätter fungieren (Rauch, Tung und Striebel [16], Singer [20] S. 142, [21] S. 91).

6. Anwendungsbeispiel

Die im ursprünglichen Modell als Normalfall bezeichnete Konstellation, vgl. Abbildung 1 auf Seite 3, kann mit Hilfe eines State Space Modells dargestellt werden, indem die zeitdiskret berechneten Preise als Beobachtungswerte des Messmodells aufgefasst werden. Da offensichtlich kein Grenzyklus auftritt, kann der Parameter k_2 gleich Null gesetzt werden, was die analytische Handhabung erheblich vereinfacht, da nun ein lineares Problem vorliegt. Man erhält

$$\mu_{i|i-1} = \mu_{i-1|i-1} \cdot e^{-k_1 \Delta t} \quad (24a)$$

$$\sigma_{i|i-1}^2 = \frac{g^2}{2k_1} \cdot (1 - e^{-2k_1 \Delta t}) \quad (24b)$$

für die Momente, und unter der Bedingung $\Delta t = 1$ ergibt sich in Verbindung mit (3), (4) und (22) die Likelihood-Funktion

$$L(\theta; \mathbf{p}) = \prod_{i=1}^T \sqrt{\frac{k_1}{\pi g^2 \cdot (1 - e^{-2k_1})}} \cdot e^{-k_1 \cdot \frac{(p_i - p^* - (p^* - p_{i-1}) \cdot e^{-k_1})^2}{g^2 \cdot (1 - e^{-2k_1})}}. \quad (25)$$

Die Maximierung von (25) kann mit Hilfe eines numerischen Optimierungsverfahrens wie dem *Newton-Raphson*-Algorithmus oder ähnlichen durchgeführt werden

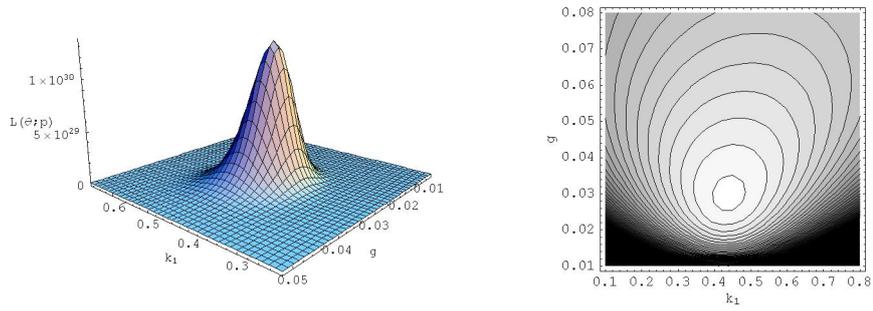


Abbildung 5: Likelihoodfunktion und Log-Likelihoodisohypsen

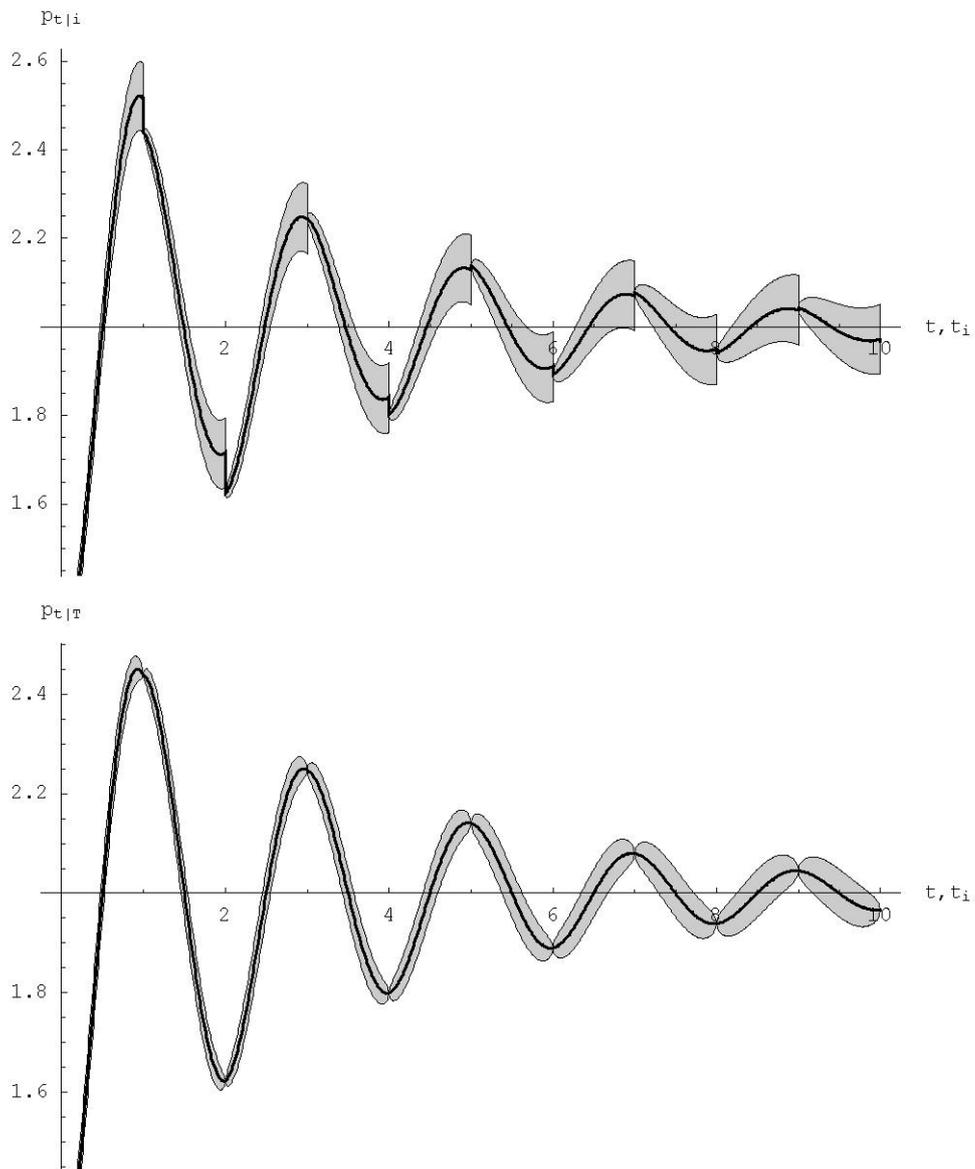


Abbildung 6: Gefilterte und geglättete Trajektorie durch $p_0 = 1.2$

(für eine ausführliche Diskussion siehe *Dennis* und *Schnabel* [5] S. 86 ff.). Als Ergebnis im Maximum der Likelihoodfunktion findet man den geschätzten Parametervektor

$$\hat{\theta} = \begin{pmatrix} \hat{k}_1 \\ \hat{g} \end{pmatrix} = \begin{pmatrix} 0.4383 \\ 0.0294 \end{pmatrix} \quad (26)$$

bei Betrachtung der ersten 30 Beobachtungswerte. Abbildung 5 auf der vorherigen Seite zeigt die scharf lokalisierte Likelihood und eine topografische Darstellung der logarithmierten Likelihood. Die zugehörige gefilterte bzw. geglättete Trajektorie ist in Abbildung 6 dargestellt. Die grau unterlegten Gebiete bilden einen 99.9%-Konfidenzbereich. Es ist zu erkennen, dass die gefilterte Trajektorie an den Beobachtungspunkten Unstetigkeiten aufweist. Das ist nicht verwunderlich, da an diesen Punkten neue Informationen in Form von Beobachtungsdaten hinzukommen. Bei Glättung ist der Beobachtungswert vorher schon bekannt, weshalb die Funktion an jeder Stelle stetig ist. Formal gilt für die Eigenschaften von Filter und Glätter

$$p_F(t) = E[p(t)|p_0, \dots, p_i; t_i \leq t] \quad (27a)$$

$$p_G(t) = E[p(t)|p_0, \dots, p_T]. \quad (27b)$$

Weiterhin fällt in Abbildung 6 auf, dass die Konfidenzbereiche unterschiedliche Formen annehmen. In der gefilterten Lösung ergibt sich ein Trichter, was zu erwarten war, da der zukünftige Wert nur prognostiziert wird. Da in der geglätteten Variante der zukünftige Wert schon feststeht, herrscht genau in der Mitte des beobachtungsfreien Intervalls die größte Unsicherheit. Aus diesem Grund nimmt der Konfidenzbereich eine elliptische Form an. Es kommt ebenfalls nicht zu einer plötzlichen Abnahme der Entropie an den Beobachtungspunkten, da keine neue Information hinzukommt.

7. Ergebnisse

Mit dem kontinuierlich/diskreten Zustandsraummodell ist ein hybrides Konstrukt gegeben, das Prozesse, die lediglich in diskreter Zeit beobachtet werden können exakt abbilden kann. Weiterhin besteht nun die Möglichkeit, auch Werte zu Zeitpunkten, die nicht beobachtet werden können zu inter- bzw. extrapolieren. Vergleicht man den Verlauf der Trajektorien in Abbildung 1 auf Seite 3 und 6 auf der vorherigen Seite, erkennt man, dass die Funktion an den Messzeitpunkten genau übereinstimmt. Darüber hinaus lässt sich in diskreter Zeit keine Aussage formulieren, da keine Theorie vorhanden ist, die die Lücken zwischen den Messdaten füllen könnte. *Phillips* sagt über die stetige Behandlung der Zeit in ökonomischen Modellen (*Phillips* [15]):

„This form is intuitively appealing because the actual movements of economic variables, as distinct from the observations on them, are dependent on the continuous passage of time. If the variables of a model are assumed to be continuous and differentiable functions of time, it seems natural to represent the direction and strength of movements in the variables by derivatives. The dynamics of a system involving several variables of this type can then be described by a number of independent differential equations.“

Das State Space Modell stellt den Zusammenhang zwischen dem zeitstetigen Zustandssystem und den zeitdiskreten Beobachtungswerten her. Mit Hilfe des durch Maximierung von Funktion (25) gefunden Parametersatzes (26) können nun Aussagen über den Systemzustand zwischen den Beobachtungspunkten gemacht und ferner auch eine Abschätzung der Unsicherheit dieser Aussagen geliefert werden. Darüber hinaus ist das Zustandsraummodell sehr flexibel. Es können ohne weiteres

eventuelle Messfehler in den Filter- bzw. Glättalgorithmus aufgenommen werden. Von fundamentaler Wichtigkeit sind kontinuierlich/diskrete Modellierungsansätze in allen Bereichen, die nur zu diskreten Zeitpunkten Beobachtungsdaten liefern, deren stetiger Fortschritt aber unstrittig ist. Eine solche Situation ist in großen Teilen der Makroökonomie gegeben, was Ausbau und Anwendung der diskutierten Verfahren nahelegt.

Literatur

- [1] Arnold, L. (1973):
„*Stochastische Differentialgleichungen. Theorie und Anwendung*“, Oldenbourg, München, Wien.
- [2] Arnold, V.I. (2001):
„*Gewöhnliche Differentialgleichungen*“, 2. Auflage, Springer, Berlin, Heidelberg, New York.
- [3] Bergstrom, A.R. (1990):
„*Continuous Time Econometric Modelling*“, Oxford University Press, Oxford, New York, Toronto.
- [4] Bronstein, I.N./Semendjajew, K.A./Musiol, G./Mühlig, H. (2001):
„*Taschenbuch der Mathematik*“, Harri Deutsch, Thun, Frankfurt/Main.
- [5] Dennis, J.E./Schnabel, R.B. (1983):
„*Numerical Methods for Unconstrained Optimization and Nonlinear Equations*“, Prentice-Hall, New Jersey.
- [6] Fahrmeir, L./Künstler, R./Pigeot, I./Tutz, G. (2003):
„*Statistik. Der Weg zur Datenanalyse*“, 4. verbesserte Auflage, Springer, Berlin, Heidelberg, New York.
- [7] Fahrmeir, L./Hamerle, A./Tutz, G. (1996):
„*Multivariate statistische Verfahren*“, 2. überarbeitete Auflage, Walter de Gruyter, Berlin, New York.
- [8] Friedrich, H./Lange, C. (1999):
„*Stochastische Prozesse in Natur und Technik. Modellierung, Simulation, Zuverlässigkeit*“, Harri Deutsch, Thun, Frankfurt/Main.
- [9] Gard, T.C. (1988):
„*Introduction to Stochastic Differential Equations*“, Marcel Dekker, New York.
- [10] Greene, W.H. (2003):
„*Econometric Analysis*“, 5. Auflage, Prentice-Hall Inc., New Jersey.
- [11] Haken, H. (1990):
„*Synergetik. Eine Einführung. Nichtgleichgewichts-Phasenübergänge und Selbstorganisation in Physik, Chemie und Biologie*“, 3. erweiterte Auflage, Springer, Berlin, Heidelberg, New York.
- [12] Harvey, A.C. (1990):
„*Forecasting, structural time series models and the Kalman filter*“, Cambridge University Press, Cambridge, New York, Port Chester, Melbourne, Sydney.
- [13] Jazwinski, A.H. (1970):
„*Stochastic Processes and Filtering Theory*“, Academic Press, New York, London.
- [14] Kalman, R.E. (1960):
„*A New Approach to Linear Filtering and Prediction Problems*“, Transactions of the ASME-Journal of Basic Engineering, Band 82, Serie D, S. 35-45.
- [15] Phillips, P.C.B. (1972):
„*The Structural Estimation of a Stochastic Differential Equation System*“, Econometrica, Band 40, S. 1021-1041.
- [16] Rauch, H.E./Tung, F./Striebel, C.T. (1965):
„*Maximum Likelihood Estimates of Linear Dynamic Systems*“, AIAA Journal, Band 3, Nr. 8, S. 1445-1450.

- [17] Risken, H. (1989):
 „The Fokker-Planck Equation. Methods of Solution and Applications“, 2. Edition,
 Springer, Berlin, Heidelberg, New York.
- [18] Schwarze, J. (1994):
 „Grundlagen der Statistik I. Beschreibende Verfahren“, 7. Auflage, Neue Wirtschafts-
 Briefe, Herne, Berlin.
- [19] Schweppe, F. (1965):
 „Evaluation of Likelihood Functions for Gaussian Signals“, *IEEE Transactions on
 Information Theory*, Band 11, S. 61-70.
- [20] Singer, H. (1992):
 „Zeitkontinuierliche Dynamische Systeme“, Campus, Frankfurt/Main, New York.
- [21] Singer, H. (1999):
 „Finanzmarktökonomie. Zeitstetige Systeme und ihre Anwendung in Ökonometrie
 und empirischer Kapitalmarktforschung“, Physica, Heidelberg.
- [22] Singer, H. (2002):
 „Parameter Estimation of Nonlinear Stochastic Differential Equations: Simulated
 Maximum Likelihood versus Extended Kalman Filter and Itô-Taylor Expansion“,
Journal of Computational and Graphical Statistics, Band 11, Nr. 4, S. 972-995.
- [23] Smith, A. (1994):
 „Einführung in die Volkswirtschaftslehre“, 2. Auflage, Oldenbourg, München, Wien.
- [24] Woll, A. (1993):
 „Allgemeine Volkswirtschaftslehre“, 11. überarbeitete und ergänzte Auflage, Vahlen,
 München.