

Zur Existenz von Markov-Prozessen höherer Ordnung

Henze, Ernst

Veröffentlicht in:
Abhandlungen der Braunschweigischen
Wissenschaftlichen Gesellschaft Band 36, 1984,
S.33-39



Verlag Erich Goltze KG, Göttingen

Zur Existenz von Markov-Prozessen höherer Ordnung

Von **Ernst Henze**, Braunschweig

(Eingegangen am 6.4.1984)

In früheren Arbeiten ([3], [4], [5]) haben der Verfasser und andere Autoren schon Ergebnisse zur Beantwortung der Frage nach der Existenz Markovscher Prozesse – speziell Markovscher Ketten – angegeben.

In der vorliegenden Abhandlung sollen nun diese Resultate, die die Nichtexistenz solcher Prozesse unter sehr allgemeinen Bedingungen zeigten, gestraft werden. Dabei ist die primäre Idee, daß die im allgemeinen in der gesamten Wahrscheinlichkeitstheorie gemachte Annahme, daß nämlich Zufallsvariable meßbare, injektive Abbildungen $X: \Omega \rightarrow \mathfrak{X}$ sind, auch hier konsequent verfolgt wird. Das bedeutet mit anderen Worten, daß man sich auf die nach CHUNG [1] oder DYNKIN [2] sog. Standardversionen – oder entsprechende einseitige Forderungen bei der Parameterannäherung – beschränkt, was aber vernünftig und zumindest für alle Anwendungen keine Einschränkung ist.

Zum allgemeinen Verständnis der Problemstellungen ist auch eine kurze Wiederholung von Ergebnissen der früheren Arbeiten notwendig.

Dabei werden wir uns auch hier, wie in den früheren Arbeiten, auf Markov-Prozesse zweiter Ordnung beschränken, weil die Ergebnisse ganz offensichtlich sofort auf solche Prozesse beliebiger, endlicher Ordnung zu übertragen sind.

1. Markovsche Ketten zweiter Ordnung mit diskretem Parameter

Gegeben sei ein Wahrscheinlichkeitsraum $(\Omega, \mathfrak{B}, P)$ und eine unendliche Familie $\{X_n; n \geq 0\}$ von Zufallsvariablen auf diesem Raum mit Werten in einer abzählbaren Menge I , dem Zustandsraum.

Definition 1: Eine doppelte Markovkette mit diskretem Parameter ist eine Folge von Zufallsvariablen $\{X_n; n \geq 0\}$ mit den folgenden Eigenschaften: Für jedes $n \geq 2$ und $i_v \in I$;

$$(1) \quad \begin{aligned} & 0 \leq v \leq n \text{ gilt } P(X_n(\omega) = i_n | X_v(\omega) = i_v; 0 \leq v \leq n-1) = \\ & = P(X_n(\omega) = i_n | X_{n-2}(\omega) = i_{n-2}, X_{n-1}(\omega) = i_{n-1}), \\ & \text{wenn die linke Seite definiert ist.} \end{aligned}$$

Wir bezeichnen mit $(i, j, k \in I)$

$$(2) \quad \begin{aligned} p_i(0) &= P(X_0(\omega) = i) \\ p_{ij}(0,1) &= P(X_1(\omega) = j | X_0(\omega) = i) \\ p_{ijk}(r, r+1, r+2) &= P(X_{r+2}(\omega) = k | X_r(\omega) = i, X_{r+1}(\omega) = j), r \geq 0 \end{aligned}$$

die Anfangsverteilung $\{p_i(0), p_{ij}(0,1); i, j \in I\}$ und die (elementaren) Übergangswahrscheinlichkeiten $p_{ijk}(r, r+1, r+2); r \geq 0$. Offensichtlich ist

$$(3) \quad P(X_v(\omega) = i_v; 0 \leq v \leq n) = p_{i_0}(0) p_{i_0 i_1}(0,1) \prod_{q=0}^{n-2} p_{i_q i_{q+1} i_{q+2}}(q, q+1, q+2)$$

und die Wahrscheinlichkeit für den allgemeinen Zylinder

$$P(X_{t_v}(\omega) = i_v; 0 \leq v \leq n); 0 \leq t_0 < t_1 < \dots < t_n$$

kann aus (3) durch Summation über die Indizes $i_\mu \in I$ gewonnen werden, die hier nicht fixiert sind.

Ganz analog zum Beweis der Existenz von einfachen Markovschen Ketten, wie er z. B. im bekannten Buch von K.L. CHUNG [1] ausgeführt ist, kann man auch die Existenz von Markovschen Ketten 2. Ordnung zeigen: Durch Vorgabe des Zustandsraumes, der Startverteilung und der Übergangswahrscheinlichkeiten ist die Existenz eines Wahrscheinlichkeitsraumes $(\Omega, \mathfrak{B}, P)$ und einer doppelten Markovkette $\{X_n; n \geq 0\}$ darauf mit diesen Wahrscheinlichkeiten gesichert.

Der Beweis folgt dem von CHUNG so analog, daß er hier nicht vorgeführt werden soll.

Aus den Start- und Übergangsverteilungen kann man nun allgemein die Größen

$$(4) \quad \begin{aligned} p_i(r) &= P(X_r(\omega) = i), r \geq 0 \\ p_{ij}(r,s) &= P(X_s(\omega) = j | X_r(\omega) = i), 0 \leq r < s \\ p_{ijk}(r,s,t) &= P(X_t(\omega) = k | X_r(\omega) = i, X_s(\omega) = j); 0 \leq r < s < t \end{aligned}$$

berechnen, für die die Funktionalgleichungen

$$(5) \quad \sum_{q \in I} p_q(r) p_{qi}(r,s) p_{qij}(r,s,t) = p_i(s) p_{ij}(s,t), 0 \leq r < s < t$$

$$(6) \quad \sum_{q \in I} p_{je}(s,u) p_{jek}(s,u,u+1) p_{okl}(u,u+1,t) = p_{jk}(s,u+1) p_{jkl}(s,u+1,t)$$

und $0 \leq s < u < u+1 < t$

$$(7) \quad \sum_{q \in I} p_{ije}(r,s,s+1) p_{jek}(s,s+1,v) = p_{ijk}(r,s,v); 0 \leq r < s < s+1 < v$$

gelten, natürliche Verallgemeinerungen der Chapman-Kolmogorov-Gleichungen.

Die Markov-Eigenschaft (1) ist doch nun offensichtlich in der folgenden, allgemeineren Bedingung enthalten: Für jedes $n \geq 2; 0 \leq t_1 < t_2 < \dots < t_n$ und alle $i_v \in I; 0 \leq v \leq n$ gilt

$$(8) \quad \begin{aligned} P(X_{t_n}(\omega) = i_n | X_{t_v}(\omega) = i_v; 0 \leq v \leq n-1) = \\ P(X_{t_n}(\omega) = i_n | X_{t_{n-2}}(\omega) = i_{n-2}, X_{t_{n-1}}(\omega) = i_{n-1}), \end{aligned}$$

wenn die linke Seite definiert ist.

Nun ist wohlbekannt, daß die (8) entsprechende Relation für einfache Markovsche Ketten mit der Markov-Eigenschaft (1) äquivalent ist. Hier gilt das nicht, man findet durch elementare, aber sehr langwierige Rechnungen, wenn man die „Lücken“ zwi-

schen den Parameterwerten $\{t, 1 \leq v \leq n$, „auffüllt“ und dann in den (3) entsprechenden Zylindern wieder heraussummiert, daß (8) aus der Definition 1 genau dann folgt, wenn noch die folgende Funktionalgleichung gilt, die nicht einfach wie die Gleichungen (5) bis (7) eine direkte wahrscheinlichkeitstheoretische Folgerung der Definition 1 ist:

$$\sum_{q \in I} p_{ije}(r, r+1, r+2) p_{jek}(r+1, r+2, v) p_{eki}(r+2, v, t) = p_{ijk}(r, r+1, v) p_{jki}(r+1, v, t)$$

$$(9) \quad 0 \leq r < r+1 < r+2 < v < t.$$

Aus der Gültigkeit der Definitionsgleichung (8) folgen nun die Funktionalgleichungen

$$(10) \quad \sum_{q \in I} p_{je}(s, u) p_{jek}(s, u, v) p_{eki}(u, v, t) = p_{jk}(s, v) p_{jki}(s, v, t), \quad 0 \leq s < u < v < t$$

und

$$(11) \quad \sum_{q \in I} p_{ije}(r, s, u) p_{jek}(s, u, v) p_{eki}(u, v, t) = p_{ijk}(r, s, v) p_{jki}(s, v, t), \quad 0 \leq r < s < u < v < t.$$

Da (9) ein Spezialfall von (11) ist und (11) aus (8) folgt, ist (9) mit (11) äquivalent, genauer gesagt für Prozesse, für die (1) gilt.

Weiterhin gibt es noch einen Zusammenhang zwischen (10) und (11): Setzt man noch konsequenterweise, weil die Zufallsvariablen $X_n(\omega)$ des Prozesses – wie schon erwähnt – meßbare und injektive Abbildungen sind

$$(12) \quad p_{ijk}(r, s, s) = \delta_{jk} = \begin{cases} 1 \\ 0 \end{cases} \text{ für } \begin{cases} j = k \\ j \neq k \end{cases}$$

und

$$(13) \quad p_{ijk}(r, r, t) = \begin{cases} p_{jk}(r, t) & \text{für } i = j \\ 0 & \text{sonst} \end{cases},$$

so kann man leicht sehen, daß (10) aus (11) folgt, wenn man (13) annimmt.

Als Hauptergebnis dieses Abschnittes zeigen wir nun, daß die Äquivalenz von (8) mit der klassischen Definition (1) die Markovkette 2. Ordnung zu einer erster Ordnung, d. h. einer einfachen Markovschen Kette degenerieren läßt: Wir beweisen dazu den

Satz 1: Sei $\{X_n; n \geq 0\}$ eine Markovsche Kette 2. Ordnung mit diskretem Parameter, für die die Markov-Eigenschaft (8) gilt, dann reduziert sich diese Kette zu einer einfachen Markovschen Kette genau dann, wenn die Funktionalgleichung (11) für $u=s$ und die Anfangsbedingung (12) gilt.

Beweis: Aus (8) folgt wie oben gesagt (11), für $u=s$ soll diese Relation ebenfalls gelten. Summiert man nun (11) in bezug auf $i \in I$ auf, so folgt (für $u=s$) mit (12)

$$\sum_{q \in I} p_{ije}(r, s, s) p_{jek}(s, s, v) = p_{ijk}(s, s, v) = p_{ijk}(r, s, v)$$

und damit hängt die linke Seite nicht von $i \in I$ und $r \geq 0$ ab, also kann auch die rechte Seite nicht davon abhängen. Die Umkehrung ist offensichtlich.

Wir gehen nun zu kontinuierlichen Parametern über, weisen aber schon darauf hin, daß die übliche Definition der Ketten hier notwendigerweise an die Gleichung (8) anschließt und daß daher – und nun allgemein – ebenfalls eine Degeneration des Konzeptes zu erwarten ist.

2. Markovsche Ketten zweiter Ordnung mit kontinuierlichem Parameter

Wir bezeichnen mit $T = [0, \infty)$ und $T^0 = (0, \infty)$. In einer Arbeit aus dem Jahre 1971 hat IOSIFESCU [6] ein Beispiel einer Markovkette 2. Ordnung mit kontinuierlichem Parameter gegeben, von dem er zeigen konnte, daß dieses, übrigens stationäre, Beispiel sich bei der Forderung einer Standard-Anfangsbedingung $\lim_{t \downarrow s} p_{ijk}(r, s, t) = \delta_{jk}$ auf eine Kette 1. Ordnung reduzierte. Wir wollen wieder notwendige und hinreichende Bedingungen dafür finden, daß das allgemeine der Fall ist.

Wir beginnen mit der

Definition 2: Eine Markovsche Kette 2. Ordnung mit kontinuierlichem Parameter $\{X_t; t \in T\}$ hat die folgende Eigenschaft: Für jedes $n \geq 2$; $0 \leq t_0 < t_1 < \dots < t_n$ und beliebige $i_v \in I$; $0 \leq v \leq n$ gilt:

$$(14) \quad \begin{aligned} P(X_{t_n}(\omega) = i_n | X_{t_v}(\omega) = i_v; 0 \leq v \leq n-1) = \\ = P(X_{t_n}(\omega) = i_n | X_{t_{n-2}}(\omega) = i_{n-2}, X_{t_{n-1}}(\omega) = i_{n-1}), \\ \text{wenn die linke Seite definiert ist.} \end{aligned}$$

Es ist klar, daß die Funktionalgleichungen (5)–(7) sowie (10) und (11) – mit den entsprechenden Bezeichnungen – hier ebenfalls gelten, wobei die Parameterwerte jetzt aus T sind. Wieder gelingt der Beweis des Existenzsatzes analog zum Fall diskreter Parameter, der der Kürze wegen auch hier fortgelassen werden soll.

Wie im diskreten Fall kann man Grenz- oder Anfangsbedingungen fordern, analog zu (12) und (13) also

$$(15) \quad \lim_{t \downarrow s} p_{ijk}(r, s, t) = \delta_{jk}$$

$$(16) \quad \lim_{s \uparrow r} p_{ijk}(r, s, t) = \begin{cases} p_{jk}(r, t) & \text{für } i = j \\ 0 & \text{sonst} \end{cases}.$$

Wir formulieren nun den folgenden Satz, dessen Beweis offensichtlich ist.

Satz 2: Sei $\{X_t; t \in T\}$ eine Markovsche Kette 2. Ordnung mit kontinuierlichem Parameter für die (a) die Konvergenz in (11) gleichmäßig in u ($s < u < v$) ist und für die (b) $\lim_{t \downarrow s} p_{jk}(s, t)$ für alle $j, k \in I$ und $\lim_{s \uparrow r} p_{ijk}(r, s, t)$ für alle $i, j, k \in I$ existiert. Dann reduziert sich der Prozeß auf eine einfache Markovsche Kette genau dann, wenn (c) $\lim_{u \downarrow s} p_{ijk}(r, s, u)$ für alle $i, j, k \in I$ existiert.

Man beachte, daß der Satz 2 in der Tat ein vollständiges Analogon des Satzes 1 ist, der Gültigkeit der Gleichung (11) auch für $u = s$ entspricht hier die Existenz der beiden Grenzwerte.

Beide führen unter der ersten Bedingung des Satzes 2 zur Degeneration einer Markovschen Kette 2. Ordnung mit kontinuierlichem Parameter zu einer einfachen, sie sind ja offensichtlich Sonderfälle der Bedingung des Satzes 2; bei (15) sieht man das sofort, (16) ist aber äquivalent zu (15), wie man auch (11) – nach Summation über 1 – leicht folgert. Die näheren Umstände gehen aus dem Abschnitt 4 hervor.

3. Markov-Prozesse zweiter Ordnung mit diskretem Parameter

Wir lassen jetzt die Voraussetzung fallen, daß der Zustandsraum I diskret ist und betrachten auf dem zugrunde liegenden Wahrscheinlichkeitsraum $(\Omega, \mathfrak{B}, P)$ einen Markov-Prozeß $\{X_n, n \geq 0\}$ mit Werten in dem beliebigen polnischen Raum \mathfrak{X} , $(\mathfrak{X}, \mathfrak{B}_{\mathfrak{X}})$ sei der zugehörige meßbare Raum, $\mathfrak{B}_{\mathfrak{X}}$ die σ -Algebra der Borelmengen von \mathfrak{X} . Dann lautet die

Definition 3: *Ein Markov-Prozeß zweiter Ordnung mit diskretem Parameter ist eine Folge von Zufallsvariablen $\{X_n, n \geq 0\}$ mit den folgenden Eigenschaften: Für jedes $n \geq 2$ und beliebiges $A \in \mathfrak{B}_{\mathfrak{X}}$ gilt*

$$(17) \quad P(X_n(\omega) \in A | X_v, 0 \leq v \leq n-1) = P(X_n(\omega) \in A | X_{n-2}, X_{n-1}) \quad P\text{-f. s. } \dots$$

Wieder kann man zeigen, daß aus der Vorgabe von Startverteilungen und der Übergangswahrscheinlichkeiten die Existenz eines solchen Markov-Prozesses bewiesen werden kann, der Kürze wegen soll darauf wieder nicht eingegangen werden, auch nicht auf den Beweis der Degeneration solcher Prozesse zu einfachen Markov-Prozessen, wenn man die allgemeinere Bedingung

$$(18) \quad \begin{aligned} & P(X_{t_n}(\omega) \in A | X_{t_v}; 0 \leq v \leq n-1) \\ & = P(X_{t_n}(\omega) \in A | X_{t_{n-2}}, X_{t_{n-1}}) \quad P\text{-f. s.} \\ & 0 \leq t_0 < t_1 < \dots < t_n \text{ fordert.} \end{aligned}$$

Wir gehen der Kürze wegen gleich zum allgemeineren Fall über.

4. Markov-Prozesse zweiter Ordnung mit kontinuierlichem Parameter

Hier sei wieder $T = [0, \infty)$ die Parametermenge, dann lautet die – auf demselben Zustandsraum wie in 3. gegeben –

Definition 4: *Ein Markov-Prozeß zweiter Ordnung mit kontinuierlichem Parameter ist eine Familie von Zufallsvariablen $\{X_t; t \in T\}$ mit den folgenden Eigenschaften: Für jedes $n \geq 3$, Parameterwerte $0 \leq t_1 < t_2 < \dots < t_n$ und beliebiges $A \in \mathfrak{B}_{\mathfrak{X}}$ gilt*

$$(19) \quad \begin{aligned} & P(X_{t_n}(\omega) \in A | X_{t_v}; 1 \leq v \leq n-1) = \\ & = P(X_{t_n}(\omega) \in A | X_{t_{n-2}}, X_{t_{n-1}}) \quad P\text{-f. s.} \end{aligned}$$

Mit $p^r(A)$, ${}^r p^s(x, A)$ und ${}^{r,s} p^t(x, y, A)$ seien die Wahrscheinlichkeiten der Start- und Übergangsverteilungen bezeichnet, wobei jetzt $r, s, t \in T$ beliebige Parameterwerte sind. Dann gilt offensichtlich die Funktionalgleichung für die Übergangswahrscheinlichkeiten, die Verallgemeinerung der Chapman-Kolmogorov-Gleichung:

$$(20) \quad {}^{r,s} p^t(x, y, A) = \int_{\mathfrak{X}} {}^{r,s} p^u(x, y, dz) {}^{s,u} p^t(y, z, A),$$

$$x, y, z \in \mathfrak{X}, 0 \leq r \leq s \leq u \leq t.$$

Auch hier kann – wie auch im Falle diskreter Parameter – ein Existenzsatz analog zu dem für Markovsche Ketten bewiesen werden, der Zustandsraum ist deshalb als polnisch vorausgesetzt.

Wir beweisen nun die Degeneration dieser Prozesse zweiter Ordnung zu solchen erster Ordnung:

Satz 3: *Es sei $\{X_t, t \in T\}$ ein Markov-Prozeß zweiter Ordnung mit kontinuierlichem Parameter, der von rechts stochastisch stetig ist und für den (a) ${}^s p^t(x, A)$ für alle $x \in \mathfrak{X}$ und $A \in \mathfrak{B}_{\mathfrak{X}}$ für $t|s$ schwach konvergent ist, (b) $\lim_{u|s} {}^{s,u} p^t(y, z, A)$ für alle $y \in \mathfrak{X}$ und $A \in \mathfrak{B}_{\mathfrak{X}}$ gleichmäßig in bezug auf $z \in \mathfrak{X}$ existiert. Dann reduziert sich der Prozeß auf einen einfachen Markov-Prozeß genau dann, wenn*

$$(c) \quad {}^{r,s} p^t(x, y, B), B \in \mathfrak{B}_{\mathfrak{X}} \text{ für } t|s \text{ schwach konvergent ist.}$$

Beweis: Aus (20) folgt für $u|s$ und unter den Voraussetzungen (b) und (c) des Satzes

$${}^{r,s} p^t(x, y, A) = \int_{\mathfrak{X}} \lim_{u|s} {}^{r,s} p^u(x, y, dz) {}^{s,u} p^t(y, z, A)$$

$$= {}^{r,s^+} p^t(x, y, dz) {}^{s,s^+} p^t(y, z, A).$$

Nun gilt für den ersten Faktor nach der Definition der allgemeinen bedingten Wahrscheinlichkeit mit einem beliebigen $B \in \mathfrak{B}_{\mathfrak{X}_s}$ und wenn wir mit $\mathfrak{X}_r, \mathfrak{X}_s$ die behandelten Basen des Zustandsraumes bezeichnen, da X_s als Zufallsvariable eine $\mathfrak{B}_{\mathfrak{X}_s}$ -meßbare Abbildung und der Prozeß nach Annahme von rechts stochastisch stetig ist.

$$P_{\mathfrak{X}_r} \otimes P_{\mathfrak{X}_s}((\mathfrak{X}_r \times B) \wedge (\mathfrak{X}_r \times \mathfrak{X}_s)) = P_{\mathfrak{X}_r \times \mathfrak{X}_s}(B \wedge (\mathfrak{X}_r \times \mathfrak{X}_s)) =$$

$$= P_{\mathfrak{X}_r \times \mathfrak{X}_s}(B \wedge (\mathfrak{X}_r \times B)) =$$

$$= \int_{\mathfrak{X}_r \times B} P(X_s \in B | X_s \in B, X_r) dP_{\mathfrak{X}_r \times \mathfrak{X}_s}.$$

Daher gilt bekanntlich

$$P(X_s \in B | X_s \in B, X_r) = 1 \quad \text{f. s.}$$

und ist damit von X_r unabhängig, mithin ${}^{r,s} p^{s^+}(x, y, B)$ von $r \in T$ und $x \in \mathfrak{X}_r$.

Umgekehrt folgt aus (a)

$${}^{r,s} p^t(x, y, A) = {}^s p^t(y, A),$$

daß $r_s p^t(x, y, A)$ schwach konvergiert für $t|s$ und natürlich gegen einen von $x \in \mathfrak{X}$ und $r \in T$ unabhängigen Grenzwert.

Man beachte, daß die Voraussetzungen (b) und (c) nur im ersten Teil des Satzes benötigt werden und (a) nur in der Umkehrung.

Für die hier behandelten Fälle Markovscher Prozesse – im allgemeinsten Falle von rechts stochastisch stetiger – zeigt man leicht, daß die früheren Bedingungen der entsprechenden Sätze mit den hier formulierten Bedingungen äquivalent sind, z. T. ist das sogar offensichtlich. Die stochastische Stetigkeit – auch die einseitige – ist also eine viel anschaulichere und natürlichere Bedingung als die früher erhobene Forderung, daß die entsprechenden Grenzwerte vom ersten Parameterwert und vom ersten Zustandswert unabhängig sein sollten.

Literatur

- [1] CHUNG, K.L.: Markov Chains with Stationary Transition Probabilities; 2nd Ed., Springer, New York 1967.
- [2] DYNKIN, E.B.: Die Grundlagen der Theorie der Markoffschen Prozesse; Springer, Heidelberg 1961.
- [3] HENZE, E., MASSÉ, J.C. und THEODORESCU, R.: On Multiple Markov Chains; Journal of Multivariate Analysis 7 (1977) 589–593.
- [4] HENZE, E. und WEHKING, A.: Räumlich homogene Markovketten höherer Ordnung; Abhandlungen der Braunschweigischen Wissenschaftlichen Gesellschaft XXX (1979), 50–53.
- [5] HENZE, E., THEODORESCU, R. und WEHKING, A.: On Multiple Markov Dependence; Rend. Sem. Mat. Univers. Politecn. Torino 38 (1980), 101–105.
- [6] IOSIFESCU, M.: On Multiple Markov Dependence; Proc. Fourth Conf. on Prob. Theory, Brasov (1971), 12–18.
- [7] WEHKING, A.: Zur Existenz Markovscher Prozesse höherer Ordnung; Diss. TU Braunschweig 1982.