8 Markov-Ketten

8.1 Grundlegendes zu Markov-Ketten

Eine Markov-Kette ist ein stochastischer Prozess, der in diskreten Zeitschritten abläuft. Dabei wird jeweils von einem Zustand in einen nächsten übergegangen.

In diesem Abschnitt sind nur einige grundlegende Definitionen und Tatsachen (meist ohne Beweise) zusammengestellt. Ausführlicheres findet man zum Beispiel in den Büchern von Behrends (2000) und in dem auch Online unter `http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/book.html` verfügbaren Werk von Grinstead und Snell (1997), sowie etwa bei Feller (1968) und Kemeny und Snell (1976).

8.1 Definition Eine endliche *Markov-Kette* ist durch ein Paar $M = (S, \mathbf{P})$ festgelegt. Dabei bezeichnet S eine endliche Menge von *Zuständen* und $\mathbf{P}$ eine zeilenstochastische $S \times S$-Matrix, die die *Übergangswahrscheinlichkeiten* enthält. (Für alle $i, j \in S$ ist also $0 \leqslant P_{ij} \leqslant 1$ und $\sum_j P_{ij} = 1$.) ◇

Für alle $i, j \in S$ ist P_{ij} als Wahrscheinlichkeit zu interpretieren, dass M vom Zustand i in den Zustand j übergeht. Diese Wahrscheinlichkeit hängt also nur von i ab, und nicht etwa von vorher durchlaufenen Zuständen oder davon, der wievielte Schritt ausgeführt wird.

Im allgemeinen erlaubt man bei Markov-Ketten auch abzählbar unendlich große Zustandsmengen. *Dieser Fall ist im Folgenden stets ausgeschlossen.*

8.2 Im folgenden bezeichnet X_t stets die Zufallsvariable, die den Zustand einer Markov-Kette zum Zeitpunkt t angibt. Es ist also $\mathbf{Pr}[X_{t+1} = j \mid X_t = i] = P_{ij}$. Auch der Anfangszustand X_0 ist im allgemeinen nicht deterministisch, sondern gemäß einer Verteilung festgelegt.

8.3 Ist $\mathbf{q}$ ein Vektor, der zu einem Zeitpunkt t angibt, mit welcher Wahrscheinlichkeit $\mathbf{q}_i$ eine Markov-Kette in Zustand i ist, dann ist $\mathbf{qP}$ der entsprechende Vektor für Zeitpunkt $t+1$, denn es gilt:

$$\mathbf{Pr}[X_{t+1} = j] = \sum_i \mathbf{Pr}[X_t = i]\, \mathbf{Pr}[X_{t+1} = j \mid X_t = i] = \sum_i \mathbf{q}_i P_{ij} = (\mathbf{qP})_j \;.$$

Analog ergibt sich $\mathbf{qP}^t$ für die Verteilung nach t Schritten.

Wir schreiben $P_{ij}^{(t)}$ für die Wahrscheinlichkeit, dass die Markov-Kette in t Schritten von Zustand i in Zustand j übergeht. Es ist also $P_{ij}^{(t)} = (\mathbf{P}^t)_{ij}$.

8.4 Definition

- Eine nichtleere Teilmenge $C \subseteq S$ von Zuständen einer endlichen Markov-Kette heißt *abgeschlossen*, falls für alle $i \in C$ und alle $j \in S \setminus C$ gilt: $P_{ij} = 0$.
- Eine abgeschlossene Teilmenge C heißt *irreduzibel*, falls keine echte (nichtleere) Teilmenge von C auch abgeschlossen ist.
- Eine Markov-Kette heiße *irreduzibel*, falls S als abgeschossene Teilmenge irreduzibel ist. ◇

Jede Markov-Kette besitzt mindestens eine abgeschlossene Teilmenge, nämlich $C = S$.

8.5 **Definition** Es seien $C_1, \dots, C_r$ alle irreduziblen Teilmengen einer Markov-Kette S und $T = S \setminus (C_1 \cup \dots \cup C_r)$. Die Zustände in T (sofern welche existieren) heißen *transient* und die Zustände in den C_k *rekurrent* (oder auch *persistent*). ◇

8.6 **Definition**

- Die Wahrscheinlichkeit, ausgehend von Zustand i nach t Schritten *erstmals* in Zustand j überzugehen, werde mit $f_{ij}^{(t)} = \mathbf{Pr}[X_t = j \wedge \forall\, 1 \leqslant s \leqslant t-1 : X_s \neq j \mid X_0 = i]$ bezeichnet.
- Die Wahrscheinlichkeit f_{ij}^*, ausgehend von Zustand i irgendwann Zustand j zu erreichen, ist $f_{ij}^* = \sum_{t>0} f_{ij}^{(t)}$.
- Der Erwartungswert m_{ij} für die benötigte Anzahl Schritte, um ausgehend von Zustand i irgendwann zum ersten Mal Zustand j zu erreichen, kann definiert werden als

$$m_{ij} = \begin{cases} \sum_{t\geqslant 1} t \cdot f_{ij}^{(t)} & \text{falls } f_{ij}^* = 1 \\ \infty & \text{sonst} \end{cases}$$

Man beachte, dass auch im Fall $f_{ij}^* = 1$ immer noch $m_{ij} = \infty$ sein kann. ◇

8.7 Jedem Random Walk auf einem Graphen G, wie wir ihn im vorangegangenen Kapitel betrachtet haben, entspricht offensichtlich eine endliche Markov-Kette M_G, indem man $P_{ij} = 0$ setzt, falls keine Kante von i nach j existiert, und andernfalls $P_{ij} = 1/d(i)$, wobei $d(i)$ der Ausgangsgrad des Knotens i ist.

Umgekehrt entspricht jeder Markov-Kette M ein Graph G_M, dessen Knoten die Zustände $i \in S$ sind und in dem eine Kante von i nach j genau dann vorhanden ist, wenn $P_{ij} > 0$ ist. In diesem Fall kann man sich auch noch die Kante mit P_{ij} gewichtet vorstellen.

Man kann nun zeigen:

8.8 **Lemma.** Es sei eine endliche Markov-Kette gegeben. Dann gilt: Ein Zustand i ist genau dann transient, wenn eine der folgenden (äquivalenten) Bedingungen erfüllt ist:

- $f_{ii}^* < 1$.
- $\sum_{t\geqslant 0} P_{ii}^{(t)} < \infty$.
- Ein Random Walk, der in i startet, kehrt mit Wahrscheinlichkeit 0 unendlich oft nach i zurück.

Analog ergibt sich:

8.9 **Lemma.** Es sei eine endliche Markov-Kette gegeben. Dann gilt: Ein Zustand i ist genau dann rekurrent, wenn eine der folgenden (äquivalenten) Bedingungen erfüllt ist:

- $f_{ii}^* = 1$.
- $\sum_{t\geqslant 0} P_{ii}^{(t)} = \infty$.
- Ein Random Walk, der in i startet, kehrt mit Wahrscheinlichkeit 1 unendlich oft nach i zurück.

8.10 **Definition** Ein rekurrenter Zustand i einer Markov-Kette heißt

- *positiv persistent*, falls $m_{ii} < \infty$ ist, und
- *null-persistent*, falls $m_{ii} = \infty$ ist. ◇

8.2 Irreduzible und ergodische Markov-Ketten

8.11 Für die Anwendungen in dieser Vorlesung sind vor allem irreduzible Markov-Ketten interessant. In diesem Fall ist die gesamte Kette die einzige irreduzible Teilmenge von Zuständen und es gibt also gar keine transienten Zustände.

8.12 DEFINITION Die *Periode* d_i eines Zustandes i ist der größte gemeinsame Teiler aller Zahlen der Menge $N_i = \{t \mid P_{ii}^{(t)} > 0 \wedge t \in \mathbb{N}_+\}$. Ein Zustand mit Periode 1 heißt auch *aperiodisch*.
Ein Zustand, der aperiodisch und positiv persistent ist, heißt auch *ergodisch*. ◇

8.13 DEFINITION Eine Markov-Kette ist *aperiodisch*, wenn alle ihre Zustände aperiodisch sind. Eine irreduzible und aperiodische Markov-Kette heißt auch *ergodisch*. ◇

Man beachte, dass für aperiodische Zustände *nicht* gilt, dass für alle t automatisch $P_{ii}^{(t)} > 0$ ist. Man kann aber zeigen:

8.14 LEMMA. Es sei $M \subseteq \mathbb{N}$ eine Menge natürlicher Zahlen mit der Eigenschaft, dass $M + M = \{k + \ell \mid k, \ell \in M\} \subseteq M$ und $\gcd M = 1$. Dann gibt es ein $k_0 \in \mathbb{N}$ mit $\{k_0, k_0 + 1, k_0 + 2, \dots\} \subseteq M$, d. h. M enthält ab irgendeinem k_0 *alle* natürlichen Zahlen.

Bei den uns in dieser Vorlesung interessierenden Anwendungen von Markov-Ketten kann man sich mitunter auf aperiodische beschränken. Wir skizzieren im folgenden zunächst, warum. Insofern ist es für den weiteren Verlauf der Vorlesung „in Ordnung“, wenn man vor allem an aperiodische Markov-Ketten denkt.

8.15 Ist eine Markov-Kette M mit Matrix $\mathbf{P}$ nicht aperiodisch, dann kann man daraus wie folgt eine neue, aperiodische Markov-Kette M' konstruieren: In M werden alle Wahrscheinlichkeiten mit $1/2$ multipliziert und für jeden Zustand i die Wahrscheinlichkeit P_{ii} um $1/2$ erhöht. Mit anderen Worten ist $\mathbf{P}' = \frac{1}{2}(\mathbf{I} + \mathbf{P})$. ($\mathbf{I}$ bezeichne die Einheitsmatrix.)
Bei dieser Vorgehensweise bleiben einige „Dinge“ erhalten. Zum Beispiel gilt: Ist $\mathbf{wP} = \mathbf{w}$, dann ist auch $\mathbf{wP}' = \mathbf{w}$ und umgekehrt. Allgemeiner haben sogar die beiden Matrizen die gleichen Eigenvektoren. Und aus einem Eigenwert λ von $\mathbf{P}$ wird ein Eigenwert $1/2 + \lambda/2$ von $\mathbf{P}'$. Wir werden im Folgenden sehen, warum gerade das interessant ist.

8.16 SATZ. *Es sei* $\mathbf{P}$ *die Matrix einer ergodischen Markov-Kette. Dann gilt:*

- $\mathbf{W} = \lim_{t\to\infty} \mathbf{P}^t$ *existiert.*
- $\mathbf{W}$ *besteht aus identischen Zeilen* $\mathbf{w}$.
- *Alle Einträge von* $\mathbf{w} = (w_1, \dots, w_n)$ *sind echt größer* 0 *und* $\sum_{i=1}^n w_i = 1$.

8.17 BEWEIS. Da die Markov-Kette ergodisch ist, folgt aus Lemma 8.14, dass es eine Potenz $\mathbf{P}^k$ gibt, deren Einträge *alle echt* größer Null sind. Um die Notation zu vereinfachen nehmen wir im folgenden einfach an, dass schon $\mathbf{P}$ diese Eigenschaft habe. Ansonsten müsste man im Folgenden statt dessen immer mit $\mathbf{P}^k$ arbeiten.
Sei nun zunächst $\mathbf{y}$ ein beliebiger Vektor.

1. Wir zeigen zunächst: Ist $d > 0$ der kleinste in $\mathbf{P}$ vorkommende Eintrag und sind m_0 und M_0 der kleinste resp. der größte Wert eines Vektors $\mathbf{y}$ und m_1 und M_1 der kleinste resp. der größte Wert von $\mathbf{Py}$, dann gilt: $M_1 - m_1 \leqslant (1 - 2d)(M_0 - m_0)$.

   Die Einträge jeder Zeile von P addieren sich zu 1. Für jedes i ist $(\mathbf{Py})_i = \sum_j P_{ij} y_j$. Offensichtlich ist

- $m_1 = \min_i \sum_j P_{ij} y_j \geqslant dM_0 + (1-d)m_0$
- $M_1 = \max_i \sum_j P_{ij} y_j \leqslant dm_0 + (1-d)M_0$

Also ist $m_0 \leqslant m_1 \leqslant M_1 \leqslant M_0$.

Außerdem ist $M_1 - m_1 \leqslant (dm_0 + (1-d)M_0) - (dM_0 + (1-d)m_0) = (1-2d)(M_0 - m_0)$.

2. Durch Induktion ergibt sich hieraus für die kleinsten und größten Einträge m_k und M_k von $\mathbf{P}^k\mathbf{y}$: $m_0 \leqslant m_1 \leqslant \cdots m_k \leqslant M_k \leqslant \cdots \leqslant M_1 \leqslant M_0$ und $M_k - m_k \leqslant (1-2d)^k(M_0 - m_0)$.

   Die Folgen der m_k und der M_k sind also beschränkt und monoton, d. h. sie besitzen jeweils einen Grenzwert $m = \lim_{k\to\infty} m_k$ bzw. $M = \lim_{k\to\infty} M_k$.

3. O. B. d. A. nehmen wir nun an, dass $\mathbf{P}$ mindestens 2 Zeilen und Spalten hat (im Fall 1 ist die zu beweisende Aussage trivialerweise richtig). Folglich ist $0 < d \leqslant 1/2$ und damit $0 \leqslant 1 - 2d < 1$. Dann ist aber $\lim_{k\to\infty} M_k - m_k = 0$ und daher $M = m$.

4. Es sei $u = M = m$. Da alle Einträge in $\mathbf{P}^k\mathbf{y}$ zwischen m_k und M_k liegen, ist $\lim_{k\to\infty} \mathbf{P}^k\mathbf{y} = \mathbf{u}$, wobei $\mathbf{u}$ der konstante Vektor ist, dessen Einträge alle gleich u sind.

Wir betrachten nun den Fall, dass $\mathbf{y} = \mathbf{e}_j$ der Einheitsvektor ist, dessen j-te Komponente 1 ist und alle anderen gleich 0.

5. Dann ist $\mathbf{P}^k\mathbf{e}_j$ die j-te Spalte von $\mathbf{P}^k$. Für jedes j konvergiert also die Folge der j-ten Spalten von $\mathbf{P}^k$ gegen einen konstanten Vektor. Also existiert $\lim_{k\to\infty} \mathbf{P}^k = \mathbf{W}$ und besteht aus lauter konstanten Spalten, d. h. mit anderen Worten aus lauter gleichen Zeilen $\mathbf{w}$.

6. Um zu zeigen, dass alle Einträge in $\mathbf{w}$ echt größer 0 sind, benutzen wir die Voraussetzung, dass $\mathbf{P}$ keine Nulleinträge hat. Dann gilt für jedes j: $\mathbf{P}\mathbf{e}_j$ enthält nur echt positive Werte, d. h. in diesem Fall ist $m_1 > 0$ und daher auch $m > 0$. Dieses m ist aber gerade die j-te Komponente von $\mathbf{w}$.

7. Die Tatsache, dass $\sum_{i=1}^n w_i = 1$ ist, ergibt sich daraus, dass für alle k die Potenzen $\mathbf{P}^k$ stochastische Matrizen sind, d. h. Zeilensumme 1 haben.

∎

Im Folgenden habe $\mathbf{w}$ stets die gleiche Bedeutung wie im obigen Beweis.

8.18 Satz. *Für jede ergodische Markov-Kette mit Matrix* $\mathbf{P}$ *gilt:*

1. $\mathbf{wP} = \mathbf{w}$.
2. *Falls* $\mathbf{vP} = \mathbf{v}$ *ist, ist* $\mathbf{v} = (\sum_j v_j)\mathbf{w}$.
3. *Es gibt genau eine Verteilung* $\mathbf{w}$ *(i. e. Summe der Einträge gleich 1) mit* $\mathbf{wP} = \mathbf{w}$.

8.19 Beweis.

1. $\mathbf{W} = \lim_{k\to\infty} \mathbf{P}^k = \lim_{k\to\infty} \mathbf{P}^{k+1} = (\lim_{k\to\infty} \mathbf{P}^k) \cdot \mathbf{P} = \mathbf{WP}$. Insbesondere gilt also für jede Zeile $\mathbf{w}$ von $\mathbf{W}$: $\mathbf{wP} = \mathbf{w}$.

2. Wenn $\mathbf{vP} = \mathbf{v}$ ist, dann auch $\mathbf{vP}^k = \mathbf{v}$ für jedes k und folglich $\mathbf{vW} = \mathbf{v}$. Ist $r = \sum_j v_j$ die Summe der Komponenten von $\mathbf{v}$, dann ist andererseits $\mathbf{vW} = r\mathbf{w}$, also $\mathbf{v} = r\mathbf{w}$.

3. Unter allen Vektoren $r\mathbf{w}$ gibt es offensichtlich genau einen, für den die Summe aller Einträge gleich 1 ist.

■

8.20 Definition Eine Verteilung $\mathbf{w}$ heißt *stationär*, falls $\mathbf{w} = \mathbf{wP}$ ist. ◇

8.21 Als erstes Beispiel wollen wir die stationäre Verteilung von Markov-Ketten M_G berechnen, die durch Graphen induziert werden. Dazu sei $G = (V, E)$ ein endlicher, zusammenhängender, ungerichteter Graphen, der nicht bipartit ist. Dann gehört jeder Knoten in G zu einem Zyklus ungerader Länge; außerdem gehört jeder Knoten zu einem Zyklus der Länge 2 (zu einem Nachbarn und zurück). Also hat in M_G jeder Zustand Periode 1 und die Markov-Kette ist aperiodisch. Da der Graph als ungerichtet und zusammenhängend angenommen wurde, ist die Kette auch irreduzibel, also auch ergodisch.

In diesem Fall kann man die eindeutige stationäre Verteilung $(w_1, \ldots, w_n)$ leicht angeben:

8.22 Lemma. Für alle $v \in V$ ist $w_v = d(v)/2m$.

8.23 Beweis. Da die stationäre Verteilung gegebenenfalls eindeutig ist, genügt es nachzuweisen, dass $\mathbf{q}$ mit $q_v = d(v)/2m$ eine Verteilung und stationär ist.

$$\begin{aligned}
\sum_{v\in V} q_v &= \sum_{v\in V} d(v)/2m = 1/2m \sum_{v\in V} d(v) = 1\,. \\
(\mathbf{qP})_v &= \sum_{u\in V} q_u P_{uv} = \sum_{(u,v)\in E} q_u P_{uv} = \sum_{(u,v)\in E} \frac{d(u)}{2m}\cdot\frac{1}{d(u)} = \sum_{(v,u)\in E} \frac{1}{2m} = \frac{d(v)}{2m}\,.
\end{aligned}$$

■

8.24 Wegen der Bemerkung in Punkt 8.15 gilt der dritte Teil der Aussage aus Satz 8.18 für irreduzible Markov-Ketten, auch wenn sie nicht aperiodisch sind: *Jede irreduzible Markov-Kette* $\mathbf{P}$ *besitzt genau eine stationäre Verteilung* $\mathbf{w}$.

Man beachte aber, dass $\lim_{t\to\infty} \mathbf{P}^t$ für irreduzible Markov-Ketten im allgemeinen nicht existiert! Für $\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ und alle k ist z. B. $\mathbf{P}^{2k} = \mathbf{I}$ und $\mathbf{P}^{2k+1} = \mathbf{P}$.

8.25 Für ergodische Markov-Ketten existiert $\lim_{t\to\infty} \mathbf{P}^t = \mathbf{W}$. Folglich existiert auch der sogenannte Cesàro-Grenzwert $\lim_{t\to\infty} \mathbf{A}_t$, wobei $\mathbf{A}_t = \frac{1}{t+1}\sum_{k=0}^{t} \mathbf{P}^k$ ist und es ist $\lim_{t\to\infty} \mathbf{A}_t = \mathbf{W}$.

Da jeder Eintrag $P_{ij}^{(k)}$ die Wahrscheinlichkeit angibt, in k Schritten von i nach j zu gelangen, ist jeder Eintrag an Stelle (i, j) von $\mathbf{A}_t$ der erwartete Anteil (als Zahl zwischen 0 und 1) von Zeitpunkten zwischen 0 und t, zu denen man in Zustand j ist, wenn man in Zustand i startet. Diese Interpretation ist natürlich immer richtig (nicht nur für ergodische Markov-Ketten).

Der interessante Punkt ist, dass man auch für periodische irreduzible Markov-Ketten noch beweisen kann (wir unterlassen das hier):

8.26 Satz. *Es sei* $\mathbf{P}$ *die Übergangsmatrix einer irreduziblen Markov-Kette* M. *Dann gilt:*

- $\lim_{t\to\infty} \mathbf{A}_t = \mathbf{W}$ *existiert.*
- *Alle Zeilen von* $\mathbf{W}$ *sind gleich.*
- *Die Zeile* $\mathbf{w}$ *ist die eindeutig bestimmte stationäre Verteilung von* M.

Mit anderen Worten: Bezeichnet $N_i(j, t)$ *die Anzahl der Besuche von Zustand* j *während der ersten* t *Schritte bei Start in* i, *so ist* $\lim_{t\to\infty} N_i(j, t)/t = w_j$ *unabhängig vom Anfangszustand* i.

Man mache sich an einem Gegenbeispiel klar, dass bei der folgenden Aussage die Forderung nach Aperiodizität unverzichtbar ist.

8.27 Satz. *Für jede ergodische Markov-Kette* $\mathbf{P}$ *und jede Verteilung* $\mathbf{v}$ *gilt:* $\lim_{t\to\infty} \mathbf{vP}^t = \mathbf{w}$.

8.28 Beweis. Es gilt $\lim_{k\to\infty} \mathbf{vP}^k = \mathbf{vW}$. Da sich die Einträge in $\mathbf{v}$ zu 1 summieren und alle Zeilen von $\mathbf{W}$ gleich $\mathbf{w}$ sind, ist $\mathbf{vW} = \mathbf{w}$. ∎

8.29 Satz. *Für jede irreduzible Markov-Kette mit stationärer Verteilung* $\mathbf{w} = (w_1, \dots, w_n)$ *gilt für alle* i*:* $w_i = 1/m_{ii}$.

8.30 Beweis. Wir beginnen mit einfachen Überlegungen zu den m_{ij}.

1. Falls $i \neq j$ ist, ist

$$m_{ij} = P_{ij} \cdot 1 + \sum_{k\neq j} P_{ik}(m_{kj} + 1) = 1 + \sum_{k\neq j} P_{ik} m_{kj}$$

2. Falls $i = j$ ist, ist

$$m_{ii} = P_{ii} \cdot 1 + \sum_{k\neq i} P_{ik}(m_{ki} + 1) = 1 + \sum_{k\neq i} P_{ik} m_{ki}$$

3. Es bezeichne nun $\mathbf{E}$ die Matrix, deren Einträge alle gleich 1 seien, $\mathbf{M}$ die Matrix mit

$$\mathbf{M}_{ij} = \begin{cases} m_{ij} & \text{falls } i \neq j \\ 0 & \text{falls } i = j \end{cases}$$

und $\mathbf{D}$ die Matrix mit

$$\mathbf{D}_{ij} = \begin{cases} 0 & \text{falls } i \neq j \\ m_{ii} & \text{falls } i = j \end{cases}$$

Dann lassen sich die eben genannten Gleichungen ausdrücken als Matrixgleichung

$$\mathbf{M} + \mathbf{D} = \mathbf{E} + \mathbf{PM} .$$

Dazu äquivalent ist

$$(\mathbf{I} - \mathbf{P})\mathbf{M} = \mathbf{E} - \mathbf{D} .$$

Da $\mathbf{wP} = \mathbf{w}$ ist, ist $\mathbf{w}(\mathbf{I} - \mathbf{P}) = \mathbf{0}$ und folglich $\mathbf{wE} = \mathbf{wD}$. Das bedeutet aber ausgeschrieben nichts anderes als

$$(1, 1, \dots, 1) = (w_1 m_{11}, w_2 m_{22}, \dots, w_n m_{nn})$$

∎

Zusammenfassung

1. Ergodische Markov-Ketten besitzen eine eindeutig bestimmte stationäre Verteilung.
2. Für irreduzible Markov-Ketten gilt das auch, aber man kann die stationäre Verteilung im allgemeinen nicht mehr durch Grenzwertübergang der P^k erhalten.

Literatur

Behrends, Ehrhard (2000). *Introduction to Markov Chains*. Advanced Lectures in Mathematics. Vieweg (siehe S. 71).

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*. third. Bd. I. John Wiley & Sons (siehe S. 71).

Grinstead, Charles M. und J. Laurie Snell (1997). *Introduction to Probablity: Second Revised Edition*. American Mathematical Society. ISBN: 0-8218-0749-8 (siehe S. 71).

Kemeny, John G. und J. Laurie Snell (1976). *Finite Markov Chains*. Springer-Verlag (siehe S. 71).

