Markov-Ketten

Seminar Stochastik vom 4.-5.2.2010
2. Version Oktober 2010
Markus Penz

1. Vorbemerkungen zu bedingten Wahrscheinlichkeiten

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $X : \Omega \to \mathbb{R}$ eine $\mathcal{F}$-messbare sowie integrierbare Zufallsvariable. Der bedingte Erwartungswert $\mathbb{E}(X|\mathcal{G})$ unter einer Sigma-Algebra $\mathcal{G} \subseteq \mathcal{F}$ als $\mathcal{G}$-messbare Zufallsvariable ist implizit durch

$$\forall B \in \mathcal{G} : \int_B X \mathrm{d}\mathbb{P} = \int_B \mathbb{E}(X|\mathcal{G}) \mathrm{d}\mathbb{P} \tag{1.1}$$

definiert. Es besteht folgender einfacher Zusammenhang zu bedingten Wahrscheinlichkeiten unter einer gegebenen Sigma-Algebra:

$$\mathbb{P}(A|\mathcal{G}) := \mathbb{E}(\mathbb{1}_A|\mathcal{G}). \tag{1.2}$$

Aus der Definition des bedingten Erwartungswertes folgt für alle $B \in \mathcal{G}$

$$\int_B \mathbb{P}(A|\mathcal{G}) \mathrm{d}\mathbb{P} = \int_B \mathbb{E}(\mathbb{1}_A|\mathcal{G}) \mathrm{d}\mathbb{P} = \int_B \mathbb{1}_A \mathrm{d}\mathbb{P} = \mathbb{P}(A \cap B). \tag{1.3}$$

$\mathbb{P}(A|\mathcal{G})(\cdot)$ ist damit wieder eine $\mathcal{G}$-messbare Zufallsvariable.

Mit $\sigma(X)$ sei die durch die Zufallsvariable X erzeugte Sigma-Algbra bezeichnet. Damit definiert man bedingte Wahrscheinlichkeiten unter einer Zufallsvariable:

$$\mathbb{P}(A|X) := \mathbb{P}(A|\sigma(X)), \tag{1.4}$$

was wieder eine $\mathcal{G}$-messbare Zufallsvariable liefert.

2. Grundlegende Definition

Definition 1. Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum mit einer Filtration $(\mathcal{F}_n)$, $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}$ und die Zufallsvariablen $X_n : \Omega \to \mathbb{R}$ alle $\mathcal{F}_n$-messbar. Eine stochastische Folge $X = (X_n, \mathcal{F}_n)$ heißt **Markov-Kette** zu $\mathbb{P}$, falls sie für alle $m \geq n \geq 0$ und für alle $B \in \mathcal{B}(\mathbb{R})$ die sog. Markov-Eigenschaft erfüllt:

$$\mathbb{P}(X_m \in B|\mathcal{F}_n) = \mathbb{P}(X_m \in B|X_n) \qquad (\mathbb{P}\text{-f.s.}). \tag{2.5}$$

Im Spezialfall einer natürlichen Filtration $\mathcal{F}_n = \mathcal{F}_n^X := \sigma(\{X_0, \ldots, X_n\})$ spricht man kurz von (X_n) als Markov-Kette. Es gilt:

$$(X_n, \mathcal{F}_n) \text{ Markov-Kette} \Rightarrow (X_n, \mathcal{F}_n^X) \text{ Markov-Kette}. \tag{2.6}$$

In diesem Fall entspricht die Markov-Eigenschaft der Bedingung, dass für alle $B_i \in \mathcal{B}(\mathbb{R})$ mit $\mathbb{P}(X_i \in B_i|i = 1, \ldots, n) > 0$ gilt

$$\mathbb{P}(X_m \in B|X_i \in B_i, i = 1, \ldots, n) = \mathbb{P}(X_m \in B|X_n \in B_n). \tag{2.7}$$

Das Wissen um zusätzliche Ereignisse in der Vergangenheit $X_i \in B_i, i = 1, \ldots, n-1$ trägt also nichts zur Vorhersage für $X_m \in B$ bei, wenn man auch $X_n \in B_n$ weiß.

Satz 1. *Sei $A_i := \{X_i \in B_i\}$ mit $B_i \in \mathcal{B}(\mathbb{R})$ so, dass $\mathbb{P}(A_i) > 0$. Dann ist die Markov-Eigenschaft für (X_n) äquivalent zu*

$$\mathbb{P}(A_{k-1}, A_{k+1}|A_k) = \mathbb{P}(A_{k-1}|A_k) \cdot \mathbb{P}(A_{k+1}|A_k). \tag{2.8}$$

Beweis. „$\Rightarrow$“: Mit dem Satz von Bayes und der Markov-Eigenschaft für $n = 2$:

$$\begin{aligned}\mathbb{P}(A_{k-1}, A_{k+1}|A_k) &= \mathbb{P}(A_{k-1}|A_k) \cdot \mathbb{P}(A_{k+1}|A_{k-1}, A_k) \\ &= \mathbb{P}(A_{k-1}|A_k) \cdot \mathbb{P}(A_{k+1}|A_k).\end{aligned} \tag{2.9}$$

„$\Leftarrow$“: Wir starten mit $k = 2$. Im ersten Schritt wendet man wieder den Satz von Bayes und (2.8) an:

$$\begin{aligned}&\mathbb{P}(A_3|A_1, A_2) \cdot \mathbb{P}(A_1|A_2) = \mathbb{P}(A_1, A_3|A_2) = \mathbb{P}(A_1|A_2) \cdot \mathbb{P}(A_3|A_2) \\ &\Rightarrow \mathbb{P}(A_3|A_1, A_2) = \mathbb{P}(A_3|A_2) \\ &\Rightarrow \text{(Induktion) } \mathbb{P}(A_n|A_1, \ldots, A_{n-1}) = \mathbb{P}(A_n|A_{n-1}).\end{aligned} \tag{2.10}$$

Dies entspricht genau der Markov-Eigenschaft und beschließt damit den Beweis. □

Die Markov-Eigenschaft kann in Worten also auch aufgefasst werden als: „Bei fixierter Gegenwart sind Vergangenheit und Zukunft stochastisch unabhängig.“

3. Diskrete Markov-Ketten

Definition 2. Ein weiterer wichtiger Spezialfall sind sogenannte **diskrete Markov-Ketten**, d.h. $X = (X_n), X_n : \Omega \to E = \{1, 2, 3, \ldots\}$ (E ist die Menge der **Zustände**) mit fester **Übergangsfunktion** $P(i; B) := \mathbb{P}(X_{n+1} \in B|X_n = i)$ für alle $n \geq 0$. Man schreibt kurz $p_{ij} := P(i; \{j\})$, also für die Wahrscheinlichkeit, dass es einen Übergang vom Zustand i in den Zustand j gibt.

Definition 3. Als **Anfangsverteilung** einer diskreten Markov-Kette (X_n) wird das Maß $\pi(B) := \mathbb{P}(X_0 \in B)$ mit $B \subseteq E$ bezeichnet.

Satz 2. *(π, P) bestimmt vollständig die wahrscheinlichkeitstheoretischen Eigenschaften der Folge X.*

Beweis. Sei $A \subseteq E^{n+1}$ für ein $n \geq 0$ beliebig. Dann ist die Wahrscheinlichkeit des Ereignisses $\{(X_0, \ldots, X_n) \in A\}$ gegeben durch:

$$\begin{aligned}\mathbb{P}((X_0, \ldots, X_n) \in A) &= \sum_{(x_0,\ldots,x_n)\in A} \mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) \\ &= \sum_{(x_0,\ldots,x_n)\in A} \underbrace{\mathbb{P}(X_n = x_n|X_{n-1} = x_{n-1})}_{P(x_{n-1};\{x_n\})} \cdot \ldots \cdot \underbrace{\mathbb{P}(X_0 = x_0)}_{\pi(\{x_0\})}.\end{aligned} \tag{3.11}$$

Das Wissen um (π, P) legt also alle Wahrscheinlichkeiten fest. □

Bemerkung 1. Ein Paar (π, P) mit einem Wahrscheinlichkeitsmaß π und $P(\cdot, B)$ messbar sowie $P(x; \cdot)$ ein Wahrscheinlichkeitsmaß erzeugt also eindeutig eine Markov-Kette.
Im Fall $\pi(\{x\}) = 1$ für ein $x \in E$ nennt man diese eine „im Punkt x beginnende Kette“. Mit einem P ist also eine ganze Familie von Markov-Ketten mit verschiedenen Anfangsverteilungen verknüpft.

Definition 4 (Übergangswahrscheinlichkeit in n Schritten)**.**

$$p_{ij}^{(n)} := \sum_{i_2 \in E, \ldots, i_n \in E} p_{ii_2} p_{i_2 i_3} \ldots p_{i_n j}. \tag{3.12}$$

Satz 3 (Chapman-Kolmogorov)**.**

$$p_{ij}^{(n+m)} = \sum_{k \in E} p_{ik}^{(n)} p_{kj}^{(m)} \tag{3.13}$$

Fragen:

(i) Existiert ein Grenzwert $\pi_j = \lim_n p_{ij}^{(n)}$ unabhängig von i?

(ii) Bilden diese $(\pi_1, \pi_2, \ldots)$ eine Wahrscheinlichkeitsverteilung auf E?

(iii) Ist die Kette **ergodisch**? D.h. gilt $\pi_j \gneqq 0$, also jeder Zustand kann im Limes erreicht werden.

(iv) Existiert eine eindeutige, stationäre Verteilung $(q_1, q_2, \ldots)$ mit $q_j = \sum_{i \in E} q_i p_{ij}$?

Zur Beantwortung dieser Fragen empfiehlt es sich, eine Klassifikation von Zuständen einer Markov-Kette zu entwerfen, abhängig von den arithmetischen und asymptotischen Eigenschaften von $p_{ij}^{(n)}$.

4. Klassifikation nach arithmetischen Eigenschaften

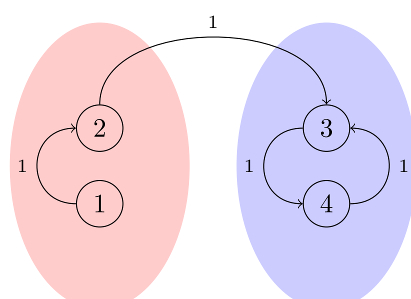
Definition 5. Ein Zustand $i \in E$ heißt **unwesentlich**, falls er nach endlich vielen Schritten verlassen wird und man nie zu ihm zurückkehrt, d.h.

$$\exists\, j \neq i\ \exists\, m > 0 : p_{ij}^{(m)} > 0 \quad \text{und} \quad \forall n > 0 : p_{ji}^{(n)} = 0. \tag{4.14}$$

Ansonsten heißt er **wesentlich**.

Beispiel 1. In der Markov-Kette mit der Übergangsmatrix

$$(p_{ij}) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{4.15}$$



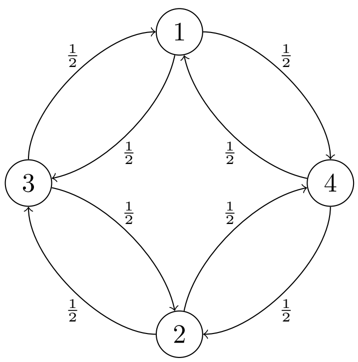
sind die Zustände 1 und 2 unwesentlich, 3 und 4 dagegen wesentlich. Sobald die wesentlichen Zustände erreicht sind, werden diese nicht mehr verlassen. Bei der Frage nach dem asymptotischen Verhalten kommt es also nur auf die wesentlichen Zustände an.

Definition 6. Ein wesentlicher Zustand $j \in E$ heißt **erreichbar** von $i \in E$, wenn es ein $m \geq 0$ gibt, so dass $p_{ij}^{(m)} > 0$ gilt. Für $m = 0$ wird die Konvention $p_{ij}^{(0)} = \delta_{ij}$ getroffen. Ist j von i aus erreichbar, notiert man kurz $i \to j$.
Zustände i und j heißen **verbunden**, wenn $i \to j \wedge j \to i$ gilt, man notiert $i \leftrightarrow j$. Diese Beziehung zwischen Zuständen definiert eine Äquivalenzrelation. Daraus folgt, dass alle wesentlichen Zustände in Äquivalenzklassen – sog. **unzerlegbare Klassen** – zerfallen, die nicht wechselseitig verbunden sind:

$$E = U \mathbin{\dot\cup} E_1 \mathbin{\dot\cup} E_2 \mathbin{\dot\cup} \ldots. \tag{4.16}$$

(Hier wurden die unwesentlichen Zustände in U zusammengefasst und die unzerlegbaren Klassen als $E_1, E_2, \dots$ bezeichnet.)

Im nächsten Schritt wird eine weitere Zerlegung dieser Klassen angestrebt, auch wenn diese als „unzerlegbar“ bezeichnet werden. Diese fußt auf dem Begriff der „Periode“ für Zustände bzw. unzerlegbare Klassen. Dazu sei zuerst ein Beispiel angegeben.



Beispiel 2. Betrachtet wird eine unzerlegbare Markov-Kette mit Übergangsmatrix

$$(p_{ij}) = \begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{pmatrix}. \tag{4.17}$$

Man kann die Zustände in die Mengen $C_0 = \{1, 2\}$ und $C_1 = \{3, 4\}$ $(E = C_0 \dot\cup C_1)$ mit „zyklischer“ Eigenschaft aufteilen, d.h. ein Zustand aus C_0 geht immer in einen in C_1 über und umgekehrt. Da es genau zwei solche Mengen gibt, weisen wir der unzerlegbaren Klasse die Periode 2 zu. Zuerst sei die Periode aber für Zustände definiert.

Definition 7. $j \in E$ hat **Periode** $d = d(j)$, wenn d der größte gemeinsame Teiler aller n mit $p_{jj}^{(n)} > 0$ ist.

Satz 4. *Alle Zustände einer unzerlegbaren Klasse E haben die gleiche Periode $d = d(E)$.*

Beweis. Für alle $i, j \in E$ findet man positive Übergangswahrscheinlichkeiten in beide Richtungen, also existieren k, l so, dass $p_{ij}^{(k)} > 0$ und $p_{ji}^{(l)} > 0$. Mittels Satz (3) folgern wir $p_{ii}^{(k+l)} \geq p_{ij}^{(k)} p_{ji}^{(l)} > 0$ und damit $d(i) \,|\, k + l$.
Sei $n > 0$ mit $p_{jj}^{(n)} > 0$ ($d(j)$ ist ggT aller dieser n) dann gilt $p_{ii}^{(k+l+n)} \geq p_{ij}^{(k)} p_{jj}^{(n)} p_{ji}^{(l)} > 0$ und damit wiederum $d(i) \,|\, k + l + n$. Zusammen mit dem oberen Ergebnis zeigt dies $d(i) \,|\, n$ für alle solchen n und deshalb $d(i) \leq d(j)$.
Aus Symmetriegründen folgt auch $d(i) \geq d(j)$ und damit $d(i) = d(j)$. □

Korollar 1. *Startet man in irgendeinem $j \in E$ mit E einer unzerlegbaren Klasse, so ist man in frühestens $d = d(E)$ Schritten wieder dort.*

Definition 8. Über die Wahl eines beliebigen Zustandes $i_0 \in E$ einer unzerlegbaren Klasse E mit $d = d(E)$ definiert man nun **zyklische Unterklassen**:

$$C_p = \{j \in E \,|\, \exists n \in \mathbb{N}_0 : p_{i_0 j}^{(n \cdot d + p)} > 0\} \quad \text{mit} \quad p \in \{0, \dots, d-1\}. \tag{4.18}$$

Ein solches j ist von i_0 aus also in $p \bmod d$ Schritten zu erreichen.

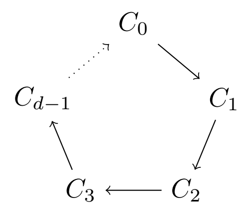
Damit diese Definition sinnvoll ist, ist noch zu zeigen:

$$p_{i_0 j}^{(n \cdot d + p)} > 0 \Rightarrow p_{i_0 j}^{(n' \cdot d + p')} = 0 \quad \text{für} \quad p \neq p' \bmod d. \tag{4.19}$$

Reductio ad absurdum: Sei $p_{i_0 j}^{(n' \cdot d + p')} > 0$. Da alle Zustände in E verbunden sind gibt es ein m mit $p_{j i_0}^{(m)} > 0$. Zusammen mit den oberen Annahmen und Satz (3) gilt also $p_{i_0 i_0}^{(n \cdot d + p + m)} > 0$ sowie $p_{i_0 i_0}^{(n' \cdot d + p' + m)} > 0$. Daraus folgern wir $d \,|\, n \cdot d + p + m$ und $d \,|\, n' \cdot d + p' + m$. Es folgt $d \,|\, p + m$ sowie $d \,|\, p' + m$ und damit $d \,|\, p - p'$, was einen Widerspruch darstellt.

Bemerkung 2. Im Fall der obigen Definition gilt für $i, j \in E$:

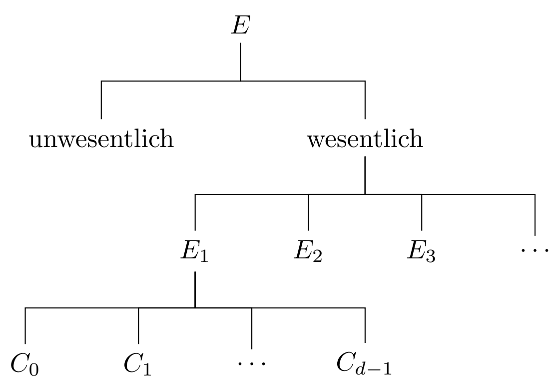
$$i \in C_p \ \wedge \ p_{ij} > 0 \Rightarrow j \in C_{p+1 (\mathrm{mod}\, d)}. \tag{4.20}$$



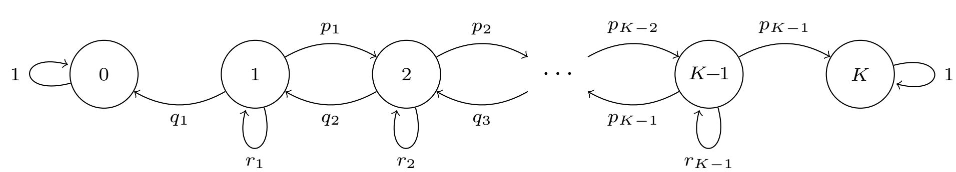
Definition 9. Eine unzerlegbare Klasse E ist **aperiodisch**, wenn $d(E) = 1$.

Korollar 2. *Jeder zyklischen Unterklasse C_p kann man eine neue Markov-Kette mit Übergangsmatrix $\left(p_{ij}^{(d)}\right)_{i,j \in C_p}$ zuordnen, die unzerlegbar und aperiodisch ist. Es genügt also, das Grenzverhalten für solche Markov-Ketten zu untersuchen.*

Aus diesen Betrachtungen folgt also folgendes Klassifizierungsschema:



Beispiel 3 (Strahlungsschaden-Modell nach Reid und Landau). Der Zustand eines Organismus, der radioaktiver Strahlung ausgesetzt ist, soll mit den Stufen 0 (gesund) bis K (irreparabler Strahlungsschaden) gleichgesetzt werden. Der Zustandsraum ist also $E = \{0, 1, \ldots, K\}$. In jedem Zeitschritt wird ein Organismus im Zustand i mit einer Wahrscheinlichkeit von q_i um eine Stufe „gesünder", verbleibt mit Wahrscheinlichkeit r_i statisch oder die Strahlenkrankheit schreitet mit Wahrscheinlichkeit p_i fort. Es muss also $q_i + r_i + p_i = 1$ gelten.



$$(p_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ q_1 & r_1 & p_1 & 0 & \cdots & 0 & 0 \\ 0 & q_2 & r_2 & p_2 & \cdots & 0 & 0 \\ \vdots & & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & r_{K-1} & p_{K-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \tag{4.21}$$

Reid und Landau wählten das noch einfachere Modell $r_i = 0, p_i = \frac{i}{K}$ und $q_i = 1 - \frac{i}{K}$.

$$(p_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \frac{K-1}{K} & 0 & \frac{1}{K} & 0 & \cdots & 0 & 0 \\ 0 & \frac{K-2}{K} & 0 & \frac{2}{K} & \cdots & 0 & 0 \\ \vdots & & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & \frac{K-1}{K} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \tag{4.22}$$

Hier zerfällt der Zustandsraum in drei unzerlegbare Klassen $E_1 = \{0\}$, $E_2 = \{1, \ldots, K-1\}$ und $E_3 = \{K\}$ mit den Perioden $d(E_1) = d(E_3) = 1$, $d(E_2) = 2$. Als „Heilungswahrscheinlichkeit“ λ_0 sei die Wahrscheinlichkeit bezeichnet, dass man ausgehend vom Zustand 1 den Zustand 0 erreicht.

$$\lambda_0 = p_{1,0} \cdot \sum_{n=0}^{\infty} p_{1,1}^{(n)} \tag{4.23}$$

5. Klassifikation nach asymptotischen Eigenschaften

Definition 10. Mit $f_{ii}^{(k)}$ bezeichnet man die Wahrscheinlichkeit für die erste Rückkehr in einen Zustand i innerhalb von k Schritten und analog dazu mit $f_{ij}^{(k)}$ $(i \neq j)$ das erste Erreichen eine Zustands j von i ausgehend in k Schritten. Beide Definitionen lassen sich so zusammenfassen:

$$f_{ij}^{(k)} := \mathbb{P}(X_k = j, X_l \neq j, 1 \leq l \leq k-1 | X_0 = i). \tag{5.24}$$

Es gilt – indem man erstes Erreichen und die restlichen Schritte trennt –

$$p_{ij}^{(n)} = \sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)}. \tag{5.25}$$

Die Wahrscheinlichkeit, dass man in i startend in beliebig vielen Zeitschritten zurückkehrt, ist

$$f_{ii} := \sum_{n=1}^{\infty} f_{ii}^{(n)}. \tag{5.26}$$

Definition 11. Mit der obigen Definition kann man nun zwischen rekurrenten und transienten Zuständen unterscheiden. Ein $i \in E$ mit $f_{ii} = 1$ heißt **rekurrent**, ansonsten bei $f_{ii} < 1$ **transient**.

Ein rekurrenter Zustand i – d.h. sichere Rückkehr zu diesem Zustand – kann zusätzlich danach klassifiziert werden, ob die mittlere Rückkehrzeit

$$\mu_i := \sum_{n=1}^{\infty} n f_{ii}^{(n)} \tag{5.27}$$

endlich oder unendlich ist. Im Fall $\mu_i^{-1} > 0$ heißt er **positiv-rekurrent**, ansonsten bei $\mu_i^{-1} = 0$ **null-rekurrent**.

Lemma 1 (ohne Beweis).

(i) Aus $j \in E$ transient folgt für alle $i \in E$, dass $p_{ij}^{(n)} \overset{n\to\infty}{\longrightarrow} 0$.

(ii) Aus $j \in E$ rekurrent und $i \leftrightarrow j$ folgt i rekurrent.

(iii) Aus $j \in E$ rekurrent, $i \leftrightarrow j$ und $d(j) = 1$ (also aperiodisch) folgt

$$p_{ij}^{(n)} \overset{n\to\infty}{\longrightarrow} \mu_j^{-1} \quad \begin{cases} = 0 & \text{falls } j \text{ null-rekurrent} \\ > 0 & \text{falls } j \text{ positiv-rekurrent.} \end{cases} \tag{5.28}$$

Korollar 3. *Eine Markov-Kette sei unzerlegbar und aperiodisch.*

(i) Falls alle Zustände null-rekurrent oder transient sind, gilt:

$$\forall i, j \in E : p_{ij}^{(n)} \overset{n\to\infty}{\longrightarrow} 0. \tag{5.29}$$

(ii) Falls alle Zustände positiv-rekurrent sind, gilt:

$$\forall i, j \in E : p_{ij}^{(n)} \overset{n\to\infty}{\longrightarrow} \mu_j^{-1} > 0. \tag{5.30}$$

Satz 5. *In einer unzerlegbaren, aperiodischen Markov-Kette mit endlich vielen Zuständen $E = \{1, \ldots, r\}$ sind alle Zustände positiv-rekurrent.*

Beweis. Reductio ad absurdum: Wir nehmen an, alle Zustände seien transient. Nach Lemma (1-i) gilt für alle $i, j \in E$, dass $p_{ij}^{(n)} \overset{n\to\infty}{\longrightarrow} 0$. Natürlich gilt

$$\sum_{j=1}^{r} p_{ij}^{(n)} = 1. \tag{5.31}$$

Lässt man nun $n \to \infty$ gehen und vertauscht den Limes mit der Summe folgt der Widerspruch

$$\sum_{j=1}^{r} \lim_{n\to\infty} p_{ij}^{(n)} = 0 \neq 1. \tag{5.32}$$

Demzufolge gibt es mindestens einen rekurrenten Zustand, nach Lemma (1-ii) sind damit aber alle Zustände rekurrent, da wir eine unzerlegbare Kette betrachten, in der alle Zustände wechselseitig verbunden sind. Falls man nun annimmt, dass alle Zustände null-rekurrent sind, folgt genau derselbe Widerspruch. Es existiert also auch mindestens ein positiv-rekurrenter Zustand i_0 und für alle $i \in E$ mit $i \leftrightarrow i_0$ gilt

$$\exists s, t : p_{i_0 i}^{(s)} > 0, p_{i i_0}^{(t)} > 0 \tag{5.33}$$

so dass

$$p_{ii}^{(n+s+t)} \geq p_{ii_0}^{(t)} p_{i_0 i_0}^{(n)} p_{i_0 i}^{(s)} \overset{n\to\infty}{\longrightarrow} p_{ii_0}^{(t)} \frac{1}{\mu_{i_0}} p_{i_0 i}^{(s)} > 0. \tag{5.34}$$

Damit gilt $\mu_i^{-1} > 0$. □

Korollar 4 (zu Fragen (i), (ii)). *Aus dem Beweis von Satz (5) folgt sofort, dass bei einer unzerlegbaren, aperiodischen Markov-Kette mit endlich vielen Zuständen für die Grenzwerte $\pi_j = \lim_n p_{ij}^{(n)}$ gilt, dass $\pi_1 + \ldots + \pi_r = 1$. Sie bilden also eine Wahrscheinlichkeitsverteilung auf E.*

Korollar 5 (zu Frage (iii)). *Eine unzerlegbare, aperiodische Markov-Kette mit endlich vielen Zuständen ist stets ergodisch.*

Beweis. Für die Ergodizität ist zu zeigen, dass für alle $j \in E$ gilt $\pi_j > 0$. Nach Satz (5) sind aber alle Zustände positiv-rekurrent und somit folgt mit Lemma (1-iii) $\pi_j = \mu_j^{-1} > 0$. □

Bemerkung 3. Tatsächlich gilt für endlich viele Zustände auch die Umkehrung, also die Äquivalenz zwischen Ergodizität auf der einen und Unzerlegbarkeit und Aperiodizität auf der anderen Seite.

Satz 6 (zu Frage (iv))**.** *Für eine endliche, ergodische Markov-Kette mit $E = \{1, \ldots, r\}$ existiert eine eindeutige, stationäre Verteilung $\mathbf{q} = (q_1, \ldots, q_r)$.*

Beweis. Zuerst soll die Existenz gezeigt werden. Wir definieren A als den Grenzwert

$$P^n = \left(p_{ij}^{(n)}\right) \stackrel{n\to\infty}{\longrightarrow} \begin{pmatrix} \pi_1 & \cdots & \pi_r \\ \vdots & \ddots & \vdots \\ \pi_1 & \cdots & \pi_r \end{pmatrix} =: A. \tag{5.35}$$

Damit folgt

$$A \cdot P = \lim_{n\to\infty} P^n \cdot P = A. \tag{5.36}$$

Die Zeilen von A liefern also genau die stationäre Verteilung $\mathbf{q} = (q_1, \ldots, q_r) := (\pi_1, \ldots, \pi_r)$ mit $\mathbf{q} \cdot P = \mathbf{q}$ bzw. wie gefordert

$$q_j = \sum_{i=1}^{r} q_i p_{ij}. \tag{5.37}$$

Zu zeigen bleibt noch die Eindeutigkeit. Sei dazu $\mathbf{q}'$ eine zweite stationäre Verteilung. Dann folgt aus $\mathbf{q}' \cdot P = \mathbf{q}'$, dass $\mathbf{q}' \cdot P^n = \mathbf{q}'$. Bilden wir nun den Grenzwert $n \to \infty$ folgt $\mathbf{q}' \cdot A = \mathbf{q}'$ und somit für alle l, dass

$$(\underbrace{q_1' + \ldots + q_r'}_{1})q_l = q_l = q_l'. \tag{5.38}$$

□

Bemerkung 4. Die Umkehrung dieses Satzes gilt allerdings nicht.

Mit diesen Aussagen zur Existenz von stationären Verteilungen in endlichen, ergodischen Markov-Ketten soll die kurze Einführung in das Gebiet der Markov-Ketten geschlossen werden.

Referenzen

[1] A.N. Širjaev, *Probability*, vol. 95 *Graduate Texts in Mathematics*, Springer (1984)
[2] P. Guttorp, *Stochastic Modeling of Scientific Data*, Chapman and Hall/CRC (1995)
[3] S. Geiß, *Stochastic processes in discrete time* (2009)
`http://www.uibk.ac.at/mathematik/personal/geiss/skripte/processes-discrete-time.pdf`
[4] Ch. Geiß, *Stochastic Modeling* (2009)
`http://www.uibk.ac.at/mathematik/personal/geiss/skripte/stochastic-models.pdf`