I. Markov-Ketten mit diskretem Zeitparameter

1. Elementare Eigenschaften von Markov-Ketten

Gegeben sei eine Folge von Zufallsvariablen (X_n) auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $X_n : \Omega \to S$ wobei S nicht leer, und endlich oder abzählbar unendlich ist.

Definition
Eine $S \times S$-Matrix $P = (p_{ij})$ heißt *stochastische Matrix*, falls $p_{ij} \geq 0$ ist und für alle $i \in S$ die Zeilensumme $\sum_{j \in S} p_{ij} = 1$ ist.

Definition
Sei P eine stochastische Matrix. Eine (endliche oder unendliche) Folge $X_0, X_1, X_2, \ldots$ von S-wertigen Zufallsvariablen heißt (homogene[1]) *Markov-Kette* mit Übergangsmatrix P, falls für alle $n \in \mathbb{N}$[2] und für alle Zustände $i_k \in S$ mit

$$P(X_0 = i_0, \ldots, X_n = i_n) > 0$$

gilt

$$P(X_{n+1} = i_{n+1} \mid X_0 = i_0, \ldots, X_n = i_n) = P(X_{n+1} = i_{n+1} \mid X_n = i_n) \coloneqq p_{i_n i_{n+1}}.$$

Die p_{ij} heißen Übergangswahrscheinlichkeiten und die *Startverteilung* ν der Kette ist definiert durch $\nu(i) \coloneqq P(X_0 = i)$ für $i \in S$.

Bemerkung: Jede Folge von unabhängigen Zufallsvariablen ist eine Markov-Kette.

Satz 1.1 (Eigenschaften von Markov-Ketten)
(X_n) ist genau dann eine Markov-Kette mit Übergangsmatrix P, falls gilt:

$$P(X_k = i_k,\, 0 \leq k \leq n) = P(X_0 = i_0) \prod_{k=0}^{n-1} p_{i_k i_{k+1}} \quad \forall n \in \mathbb{N}_0 \ \forall i_k \in S$$

genau dann wenn gilt:

$$P(X_k = i_k,\, 1 \leq k \leq n \mid X_0 = i_0) = \prod_{k=0}^{n-1} p_{i_k i_{k+1}} \quad \forall n \in \mathbb{N}_0 \ \forall i_k \in S \text{ mit } P(X_0 = i_0) > 0$$

[1] kurz für zeit-homogen. Die Übergangswahrscheinlichkeiten hängen nicht vom aktuellen Zeitpunkt ab.
[2] Hier ist $\mathbb{N} = 1, 2, \ldots$

genau dann wenn gilt:

$$P(X_k = i_k,\, m \le k \le m+n) = P(X_m = i_m) \prod_{k=m}^{m+n-1} p_{i_k i_{k+1}} \quad \forall m,n \in \mathbb{N}_0 \; \forall i_k \in S$$

Beweis
Zur ersten Äquivalenz. Sei $A_k \coloneqq [X_k = i_k]$, $k \in \mathbb{N}_0$.

„$\Longrightarrow$“ Induktion über n: $n = 0$ ✓, $n \curvearrowright n+1$:

$$\begin{aligned} P(A_0 A_1 \dots A_n A_{n+1}) &= P(A_0 \dots A_n) \cdot P(A_{n+1} \mid A_0 \dots A_n) \\ &= P(A_0 \dots A_n) \cdot p_{i_n i_{n+1}} && \text{(Markov-Eigenschaft)} \\ &= P(X_0 = i_0) \prod_{k=0}^{n} p_{i_k i_{k+1}} && \text{(I.V.)} \end{aligned}$$

„$\Longleftarrow$“

$$\begin{aligned} P(A_{n+1} \mid A_0 \dots A_n) &= \frac{P(A_0 \dots A_n A_{n+1})}{P(A_0 \dots A_n)} \\ &= p_{i_n i_{n+1}} && \text{(Vor.)} \end{aligned}$$

Die weiteren Äquivalenzen sind ähnlich zu beweisen. ■

Konstruktion einer Markov-Kette. Seien (Y_n) Zufallsvariablen, unabhängig und identisch verteilt (u.i.v.), in Z. Weiter ist $g : S \times Z \to S$ eine messbare Abbildung. Definiere die Folge (X_n) mit

$$X_0 = c \in S, \quad X_n = g(X_{n-1}, Y_n).$$

Die so konstruierte Folge (X_n) ist eine Markov-Kette mit Werten in S und Übergangsmatrix $P = (p_{ij})$ mit $p_{ij} = P(g(i, Y_n) = j)$.

Beweis
Die Variablen $X_0, \dots, X_n$ hängen nur von $X_0, Y_1, \dots, Y_n$ ab, sind also unabhängig von Y_{n+1}.

$$\begin{aligned} P(X_{n+1} = i_{n+1} \mid X_k = i_k, 0 \le k \le n) &= \frac{P(X_k = i_k,\, 0 \le k \le n+1)}{P(X_k = i_k,\, 0 \le k \le n)} \\ &= \frac{P(X_k = i_k,\, 0 \le k \le n,\, g(i_n, Y_{n+1}) = i_{n+1})}{P(X_k = i_k,\, 0 \le k \le n)} \\ &= \frac{P(X_k = i_k,\, 0 \le k \le n) \cdot P(g(i_n, Y_{n+1}) = i_{n+1})}{P(X_k = i_k,\, 0 \le k \le n)} \\ &= P(g(i_n, Y_{n+1}) = i_{n+1}) \\ &= \frac{P(g(i_n, Y_{n+1}) = i_{n+1}) \cdot P(X_n = i_n)}{P(X_n = i_n)} \\ &= \frac{P(g(i_n, Y_{n+1}) = i_{n+1}, X_n = i_n)}{P(X_n = i_n)} \\ &= P(g(i_n, Y_{n+1}) = i_{n+1} \mid X_n = i_n) \\ &= P(X_{n+1} = i_{n+1} \mid X_n = i_n) \end{aligned}$$

■

Bemerkung: Umgekehrt kann zu jeder stochastischen Matrix P eine Markov-Kette (X_n) konstruiert werden mit $X_n = g(X_{n-1}, Y_n)$, wobei (Y_n) u.i.v. und o.B.d.A. $Y_n \sim U[0,1]$.

Beispiel 1.1 (Lagerhaltung)
Sei Y_n die Nachfrage nach einem gelagerten Produkt im Zeitintervall $(n-1, n]$. (Y_n) sei u.i.v. und $Y_n \in \mathbb{N}_0$. Die Auffüll-Politik sei eine (z, Z)-Politik mit $z \leq Z$, $z, Z \in \mathbb{N}$, die wie folgt funktioniert: Falls der Lagerbestand zur Zeit $n \leq z$ ist, dann fülle auf Z auf, sonst tue nichts.

Sei X_n der Lagerbestand zum Zeitpunkt n, $S = \mathbb{N}_0$. Es gilt

$$X_n = \begin{cases} (Z - Y_n)^+, & X_{n-1} \leq z \\ (X_{n-1} - Y_n)^+, & X_{n-1} > z \end{cases}$$

Also ist (X_n) eine Markov-Kette mit Übergangsmatrix $P = (p_{ij})$ und

$$p_{ij} = \begin{cases} P((Z - Y_n)^+ = j), & i \leq z \\ P((i - Y_n)^+ = j), & i > z \end{cases}$$

Beispiel 1.2 (Ruinspiel)
Zwei Spieler mit Startkapital $B \in \mathbb{N}$ Euro spielen in Runden um jeweils einen Euro, etwa mit einem Münzwurf. Spieler I gewinnt dabei mit Wahrscheinlichkeit p. Sei $Y_n = 1$, falls Spieler I die n-te Runde gewinnt, und $Y_n = -1$, falls er die n-Runde verliert. Wir nehmen an, dass Y_n u.i.v. ist.

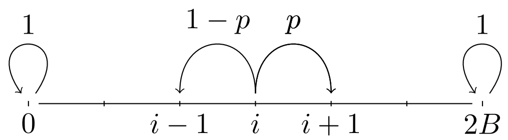
Wir interessieren uns für das Kapital X_n von Spieler I nach der n-ten Runde. Damit ist der Zustandsraum $S = \{0, 1, \ldots, 2B\}$.

Es gilt $X_0 = B$ und

$$X_n = \begin{cases} 2B, & X_{n-1} = 2B \\ X_{n-1} + Y_n, & 0 < X_{n-1} < 2B \\ 0, & X_{n-1} = 0. \end{cases}$$

Es folgt aus der Konstruktion direkt dass (X_n) eine Markov-Kette ist mit Übergangsmatrix $P = (p_{ij})$, und $p_{00} = p_{2B,2B} = 1$ sowie

$$p_{ij} = \begin{cases} p, & j = i+1 \\ 1-p, & j = i-1 \end{cases} \quad \text{für } 0 < i < 2B.$$



Beispiel 1.3 (Wartesystem)
Zu jedem Zeitpunkt $n = 0, 1, \ldots$ können maximal m Kunden bedient werden. Y_n sei die Anzahl der zufällig im Zeitintervall $(n-1, n]$ eintreffenden Kunden und sei u.i.v.

Sei X_n die Anzahl der zur Zeit n wartenden Kunden, $S = \mathbb{N}_0$. Es gilt $X_0 = c$ und $X_n = (X_{n-1} - m)^+ + Y_n$. Also ist (X_n) eine Markov-Kette mit Übergangsmatrix $P = (p_{ij})$ und $p_{ij} = P(Y_n = j - (i-m)^+)$, $i, j \in \mathbb{N}_0$.

Definition
Sei P eine stochastische $S \times S$-Matrix. Dann heißen die Elemente $p_{ij}^{(n)}$ von P^n die n-Schritt-Übergangswahrscheinlichkeiten zu P. Wir definieren $P^0 = E$, also $p_{ij}^{(0)} = \delta_{ij}$.

Satz 1.2
Sei (X_n) eine Markov-Kette mit Übergangsmatrix P. Dann gilt:

a) $P(X_{n+m} = j \mid X_m = i) = p_{ij}^{(n)}$ für alle $i, j \in S$, $m, n \in \mathbb{N}_0$ mit $P(X_m = i) > 0$.

b) $P(X_n = j) = \sum_{i \in S} P(X_0 = i) p_{ij}^{(n)}$, $j \in S$, $n \in \mathbb{N}$.

Beweis
a)

$$\begin{aligned} P(X_{n+m} = i_{n+m}, X_m = i_m) &= \sum_{i_{m+1},\ldots,i_{n+m-1} \in S} P(X_m = i_m) \prod_{k=m}^{m+n-1} p_{i_k i_{k+1}} \\ &= P(X_m = i_m) p_{i_m i_{m+n}}^{(n)} \end{aligned}$$

b)

$$\begin{aligned} P(X_n = j) &= \sum_{i \in S} P(X_n = j, X_0 = i) \\ &= \sum_{i \in S} P(X_n = j \mid X_0 = i) \cdot P(X_0 = i) \\ &= \sum_{i \in S} P(X_0 = i) p_{ij}^{(n)} \end{aligned}$$

■

Bemerkung: i) Wegen $P^{n+m} = P^n \cdot P^m$ gilt:

$$p_{ij}^{(n+m)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj}^{(m)} \text{ für } i, j \in S$$

Dies ist die „Chapman-Kolmogorov-Gleichung“.

ii) Ist $X_0 \sim \nu$, so gilt $X_n \sim \nu \cdot P^n$.

Satz 1.3 (Existenzsatz für Markov-Ketten)
Sei ν ein Wahrscheinlichkeitsmaß auf S und P eine stochastische $S \times S$-Matrix. Sei X_n die n-te Projektion auf $\Omega := S^{\mathbb{N}_0}$, also $X_n : \Omega \to S$, $n \in \mathbb{N}_0$ mit $X_n(\omega) = X_n((i_0, i_1, \ldots)) = i_n$.

Dann existiert ein Wahrscheinlichkeitsmaß P auf $\mathcal{F} = \oplus_{n=0}^{\infty} \mathcal{P}(S)$, sodass (X_n) eine Markov-Kette mit Übergangsmatrix P und Startverteilung ν ist, d.h:

- $P(X_0 = i_0) = \nu(i_0)$, $i_0 \in S$
- $P(X_{n+1} = j \mid X_n = i) = p_{ij}$, $i, j \in S$, $P(X_n = i) > 0$.

Beweis
Satz von Ionescu-Tulcea über die Fortsetzung von Maßen und die Existenz zufälliger Folgen. ∎

2. Klassifikation von Zuständen, Rekurrenz und Transienz

In diesem Paragraphen widmen wir uns Fragestellungen wie diesen: Welche Zustände in S werden von der Markov-Kette mit Sicherheit besucht und welche nicht? Wenn sie besucht werden, wie oft?

Definition
Sei (X_n) eine Markov-Kette mit Übergangsmatrix $P = (p_{ij})$.

a) $i \in S$ *führt nach* $j \in S$ (kurz $i \rightsquigarrow j$), falls es ein $n \in \mathbb{N}$ gibt mit $p_{ij}^{(n)} > 0$.

b) $i \in S$ *kommuniziert mit* $j \in S$ (kurz $i \leftrightarrow j$) falls sowohl $i \rightsquigarrow j$ als auch $j \rightsquigarrow i$ gilt.

Bemerkung: Für $i, j \in S$ sei $i \sim j$ definiert als $(i \leftrightarrow j) \vee (i = j)$. Diese Relation ist eine Äquivalenzrelation auf S, da sie reflexiv, symmetrisch und transitiv ist.

Dies liefert uns eine Partition von S mit den Äquivalenzklassen $K(i) := \{j \in S \mid i \sim j\}$. Die Äquivalenzklasse $K(i)$ enthält i selbst und die mit i kommunizierenden Zustände.

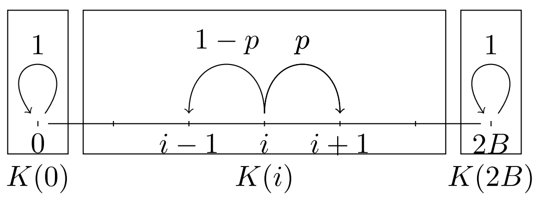
Definition
Sei (X_n) eine Markov-Kette mit Übergangsmatrix $P = (p_{ij})$.

a) $J \subset S$ heißt *abgeschlossen*, wenn es keine zwei Zustände $j \in J$ und $i \in S \setminus J$ gibt mit $j \rightsquigarrow i$.

b) Die Markov-Kette (X_n) beziehungsweise die Übergangsmatrix P heißen *irreduzibel*, falls S nur aus einer Klasse besteht, also für alle $i, j \in S$, $i \neq j$, gilt $i \leftrightarrow j$.

Beispiel 2.1
Skizze, hier ausgelassen

Beispiel 2.2 (Ruinspiel)



Lemma 2.1
$J \subset S$ ist genau dann abgeschlossen, wenn $(p_{ij},\ i, j \in J)$ stochastisch ist.

Beweis
„$\Longrightarrow$": Klar. „$\Longleftarrow$": Es gilt: $(p_{ij},\ i, j \in J)$ stochastisch $\iff$ $(p_{ij}^{(n)},\ i, j \in J)$ stochastisch für alle $n \in \mathbb{N}$. ∎

Sei (X_n) eine Markov-Kette mit Übergangsmatrix $P = (p_{ij})$. Es sei

$$T_i := \inf\{n \in \mathbb{N} \mid X_n = i\}$$

die (zufällige) Ersteintrittszeit der Markov-Kette in den Zustand i.

Wir setzen dabei $\inf \emptyset := \infty$. Weiter sei für $i, j \in S$, $n \in \mathbb{N}$:

$$\begin{aligned} f_{ij}^{(n)} &:= P(T_j = n \mid X_0 = i) = P_i(T_j = n) \\ &= P(X_n = j,\, X_\nu \neq j \text{ für } 1 \leq \nu < n \mid X_0 = i) \\ f_{ij}^{(0)} &:= 0 \end{aligned}$$

Offenbar ist $f_{ij}^{(1)} = p_{ij}$. Weiter definieren wir

$$f_{ij}^* := \sum_{n=0}^{\infty} f_{ij}^{(n)} = \sum_{n=0}^{\infty} P_i(T_j = n) = P_i(T_j < \infty) = P_i(\exists n \in \mathbb{N} : X_n = j) \in [0, 1]$$

Definition
Ein Zustand $i \in S$ heißt *rekurrent*, falls $f_{ii}^* = 1$ und *transient* sonst.

Lemma 2.2
Für alle $n \in \mathbb{N}$, $i, j \in S$ gilt:

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

Beweis (Methode des ersten Besuches)
Unter Verwendung der Formel $P(AB \mid C) = P(B \mid C) \cdot P(A \mid BC)$ für Ereignisse A, B, C zeigen wir:

$$\begin{aligned} p_{ij}^{(n)} &= P_i(X_n = j) = \sum_{k=1}^{n} P(X_n = j,\, X_\mu \neq j,\, 1 \leq \mu < k,\, X_k = j \mid X_0 = i) \\ &= \sum_{k=1}^{n} P_i(X_\mu \neq j,\, 1 \leq \mu < k,\, X_k = j) \cdot \\ &\qquad \underbrace{P(X_n = j \mid X_0 = i,\, X_\mu \neq j,\, 1 \leq \mu < k,\, X_k = j)}_{= P(X_n = j \mid X_k = j)} \\ &= \sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)} \end{aligned}$$

■

Satz 2.3
$i \in S$ ist rekurrent genau dann, wenn gilt:

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$$

Beweis
Für $s \in (0,1)$ erhalten wir aus Lemma 2.2:

$$\begin{aligned}\sum_{n=0}^{\infty} p_{ii}^{(n)} s^n = 1 + \sum_{n=1}^{\infty} p_{ii}^{(n)} s^n &= \\ &= 1 + \sum_{n=1}^{\infty} s^n \sum_{k=1}^{n} f_{ii}^{(k)} p_{ii}^{(n-k)} \\ &= 1 + \sum_{k=1}^{\infty} f_{ii}^{(k)} s^k \sum_{n=k}^{\infty} p_{ii}^{(n-k)} s^{n-k} \\ &= 1 + \sum_{k=1}^{\infty} f_{ii}^{(k)} s^k \sum_{n=0}^{\infty} p_{ii}^{(n)} s^n\end{aligned}$$

Abkürzend schreiben wir $F(s) := \sum_{k=1}^{\infty} f_{ii}^{(k)} s^k$ und $P(s) = \sum_{n=0}^{\infty} p_{ii}^{(n)} s^n$, also gilt

$$P(s) = 1 + F(s) \cdot P(s).$$

Nun sei $s \to 1$ (monotone Konvergenz!), und wir erhalten

$$P(1) = 1 + f_{ii}^* \cdot P(1).$$

Es folgt: Ist $f_{ii}^* = 1$, so gilt $P(1) = 1 + P(1)$, also ist $P(1) = \sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$. Ist ansonsten $f_{ii}^* < 1$, so gilt $P(1) = \frac{1}{1-f_{ii}^*} < \infty$. ■

Bemerkung: Die im Satz 2.3 auftretende Reihe kann wie folgt interpretiert werden:

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \sum_{n=0}^{\infty} E_i[1_{[X_n=i]}] = E_i(\sum_{n=0}^{\infty} 1_{[X_n=i]})$$

Sie bezeichnet also die erwartete Anzahl der Besuche des Zustandes $i \in S$.

Satz 2.4 (Solidaritätsprinzip)
Ist ein Zustand $i \in S$ rekurrent (bzw. transient), so ist jeder Zustand in $K(i)$ rekurrent (bzw. transient).

Beweis
Sei i rekurrent und $j \in K(i)$, $j \neq i$, das heißt es gibt $n, m \in \mathbb{N}$, sodass $p_{ij}^{(m)} \cdot p_{ji}^{(n)} > 0$. Mit der Abschätzung

$$\sum_{k=0}^{\infty} p_{jj}^{(k)} \geq \sum_{k=0}^{\infty} p_{jj}^{(m+n+k)} \geq \sum_{k=0}^{\infty} p_{ji}^{(n)} p_{ii}^{(k)} p_{ij}^{(m)} = p_{ij}^{(m)} p_{ji}^{(n)} \sum_{k=0}^{\infty} p_{ii}^{(k)}$$

und Satz 2.3 ist $\sum_{k=0}^{\infty} p_{jj}^{(k)} = \infty$ und j rekurrent. ■

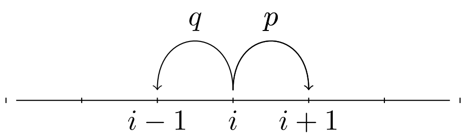
Bemerkung: Ist $i \in S$ rekurrent (bzw. transient), so sagen wir $K(i)$ ist rekurrent (bzw. transient).

Ist (X_n) irreduzibel und ein $i \in S$ ist rekurrent (bzw. transient), so sagen wir (X_n) ist rekurrent (bzw. transient).

Beispiel 2.3 (Irrfahrt auf den ganzen Zahlen, „Random Walk“)
Es sei $X_n = \sum_{k=1}^{n} Y_k = X_{n-1} + Y_n$ und $X_0 = 0$, wobei (Y_n) u.i.v. mit

$$P(Y_n = 1) = p = 1 - P(Y_n = -1) = 1 - q, \quad p \in (0,1)$$

ist $(S = \mathbb{Z})$.



(X_n) ist nach Konstruktion eine irreduzible Markov-Kette. Ist (X_n) rekurrent oder transient?

Wir wenden Satz 2.3 an und untersuchen o.B.d.A. $i = 0$. Es gilt für alle $n \in \mathbb{N}_0$:

$$\begin{aligned} p_{00}^{(2n+1)} &= 0 \\ p_{00}^{(2n)} &= p^n q^n \binom{2n}{n} = p^n q^n \frac{(2n)!}{(n!)^2} \end{aligned}$$

Mit der Stirling-Formel $(n! \simeq (\frac{n}{e})^n \cdot \sqrt{2\pi n})$ erhält man dann

$$p_{00}^{(2n)} \approx (pq)^n \cdot \frac{(\frac{2n}{e})^{2n}\sqrt{4\pi n}}{(\frac{n}{e})^{2n} 2\pi n} = \frac{(4pq)^n}{\sqrt{\pi n}}.$$

Fall 1 Ist $p = q = \frac{1}{2}$, so ist $p_{00}^{(2n)} \approx \frac{1}{\sqrt{\pi n}}$, also ist $\sum_{n=0}^{\infty} p_{00}^{(2n)} = \infty$ und die Markov-Kette ist rekurrent.

Fall 2 Ist dagegen $p \neq q$, also $pq < \frac{1}{4}$, so ist $\sum_{n=0}^{\infty} p_{00}^{(2n)} \leq c \cdot \sum_{n=0}^{\infty} (4pq)^n < \infty$, also ist die Markov-Kette transient.

Bemerkung: Betrachtet man die symmetrische Irrfahrt auf $\mathbb{Z}^d$, also $p_{ij} = \frac{1}{2d}$ für $\|i - j\| = 1$, mit $\|\cdot\|$ der l^1-Norm und $i, j \in \mathbb{Z}^d$, so ist die Irrfahrt rekurrent für $d = 1, 2$ und transient sonst.

Lemma 2.5
Liegen i und j in der selben rekurrenten Klasse, so gilt $f_{ij}^* = f_{ji}^* = 1$.

Lemma 2.6
Für alle $i, j \in S$ gilt: Wenn j transient ist, dann gilt

$$\sum_{n=0}^{\infty} p_{ij}^{(n)} < \infty$$

Insbesondere ist

$$\lim_{n \to \infty} p_{ij}^{(n)} = 0.$$

Beweis
Summiere die Gleichung in Lemma 2.2 über alle n:

$$\begin{aligned}\sum_{n=0}^{\infty} p_{ij}^{(n)} &= \delta_{ij} + \sum_{n=1}^{\infty}\sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)} \\ &= \delta_{ij} + \sum_{k=1}^{\infty} f_{ij}^{(k)} \sum_{n=k}^{\infty} p_{jj}^{(n-k)} \\ &= \delta_{ij} + f_{ij}^{*} \underbrace{\sum_{n=0}^{\infty} p_{jj}^{(n)}}_{<\infty \text{ da } j \text{ transient}} < \infty\end{aligned}$$

■

Satz 2.7
Ist eine Klasse $K \subseteq S$ rekurrent, so ist K abgeschlossen bzw. $(p_{ij},\, i, j \in K)$ ist stochastisch.

Beweis
Wir zeigen: Ist $i \in K$ rekurrent und $i \rightsquigarrow j$, $i \neq j$, dann gilt $j \rightsquigarrow i$ und damit $j \in K$.

Angenommen, $j \rightsquigarrow i$ gelte nicht, also $p_{ji}^{(n)} = 0$ für alle $n \in \mathbb{N}$, und sei $N \in \mathbb{N}$ die kleinste Zahl mit $p_{ij}^{(N)} > 0$. Es gilt nun für alle $n \in \mathbb{N}$

$$P_i(X_N = j, X_n = i) = 0.$$

Denn für $n > N$ gilt: $P_i(X_N = j, X_n = i) = p_{ij}^{(N)} p_{ji}^{(n-N)} = 0$ und für $n < N$ gilt: $P_i(X_N = j, X_n = i) = p_{ii}^{(n)} p_{ij}^{(N-n)} = 0$, da $N - n < N$.

Also ist $P_i(T_i \leq m, X_N = j) = \sum_{n=1}^{m} P_i(T_i = n, X_N = j) \leq \sum_{n=1}^{m} P_i(X_n = i, X_N = j) = 0$ und damit

$$\begin{aligned}\sum_{n=1}^{m} f_{ii}^{(n)} &= P_i(T_i \leq m) \\ &= P_i(T_i \leq m, X_N \neq j) \\ &\leq P_i(X_N \neq j) \\ &= 1 - P_i(X_N = j) = 1 - p_{ij}^{(N)}.\end{aligned}$$

Für $m \to \infty$ folgt dann

$$1 = f_{ii}^{*} = \sum_{n=1}^{\infty} f_{ii}^{(n)} \leq 1 - p_{ij}^{(N)} < 1,$$

was ein Widerspruch ist. ■

Satz 2.8
Ist die Klasse K endlich und abgeschlossen, so ist K rekurrent.

Beweis
Da $(p_{ij},\, i,j \in K)$ stochastisch ist, folgt induktiv, dass $(p_{ij}^{(n)},\, i,j \in K)$ stochastisch für alle $n \in \mathbb{N}$ ist. Angenommen, K wäre transient. Sei dann $j \in K$, dann ist nach Lemma 2.6 $p_{ij}^{(n)} \to 0$ für $n \to \infty$ und alle $i \in S$. Für $i \in K$ folgt also: $1 = \sum_{j \in K} p_{ij}^{(n)} \to 0$ für $n \to \infty$. Widerspruch. ■

Bemerkung: Insbesondere gilt: Ist S endlich und P irreduzibel, so ist die Markov-Kette rekurrent.

Beispiel 2.4 (Irrfahrt mit reflektierenden Grenzen)
Die Irrfahrt ist irreduzibel und rekurrent nach Satz 2.8.

Absorbtionswahrscheinlichkeiten

Sei $(X_n)_{n\in\mathbb{N}}$ eine Markov-Kette mit Zustandsraum S und Übergangsmatrix $P = (p_{ij})$. Aufgrund der bisherigen Ergebnisse können wir S zerlegen in rekurrente Klassen $K_1, K_2, \ldots$ und eine Menge von transienten Zuständen T, also $S = T \cup K_1 \cup K_2 \cup \ldots$.

Es sei $\tau := \inf\{n \geq 0 \mid X_n \notin T\}$ die Austrittszeit aus der Menge der transienten Zustände. Für $i \in T$, $k \in T^c$ interessiert uns $u_{ik} = P_i(X_\tau = k)$, vorausgesetzt $P_i(\tau < \infty) = 1$.

Wir unterteilen $P = (p_{ij})$ in

$$P = \begin{pmatrix} Q & R \\ 0 & \tilde{P} \end{pmatrix}$$

wobei Q die Einschränkung von P auf die transienten Zustände ist, also $Q = (q_{ij}) = (p_{ij}, i, j \in T)$.

Satz 2.9
Für $i \in T$, $j \in T^c$ gilt:

$$u_{ij} = \sum_{k \in T} q_{ik} u_{kj} + p_{ij}.$$

Beweis
Sei $i \in T$, $j \in T^c$.

$$\begin{aligned}
u_{ij} &= P_i(X_\tau = j) \\
&= \sum_{k \in S} P_i(X_\tau = j, X_1 = k) \\
&= \underbrace{\sum_{k \in T} P_i(X_\tau = j, X_1 = k)}_{=: A} + \underbrace{\sum_{k \in T^c} P_i(X_\tau = j, X_1 = k)}_{= p_{ij}} \\
A &= \sum_{k \in T} \sum_{n \geq 2} P_i(\tau = n, X_n = j, X_1 = k) \\
&= \sum_{k \in T} \sum_{n \geq 2} P_i(X_2 \in T, \ldots, X_{n-1} \in T, X_n = j, X_1 = k) \\
&= \sum_{k \in T} \sum_{n \geq 2} P_i(X_2 \in T, \ldots, X_{n-1} \in T, X_n = j \mid X_1 = k) \cdot P_i(X_1 = k) \\
&= \sum_{k \in T} \sum_{n \geq 2} p_{ik} P_k(X_1 \in T, \ldots, X_{n-2} \in T, X_{n-1} = j) \\
&= \sum_{k \in T} \sum_{n \geq 2} p_{ik} P_k(\tau = n - 1, X_\tau = j) \\
&= \sum_{k \in T} p_{ik} P_k(X_\tau = j) \\
&= \sum_{k \in T} p_{ik} u_{kj}
\end{aligned}$$

Da für $i, k \in T$ gilt: $p_{ik} = q_{ik}$, folgt die Behauptung. ■

Bemerkung: Es sei $U = (u_{ij})_{i \in T, j \in T^c}$. Dann lässt sich Satz 2.9 schrieben als $U = QU + R$ bzw. $U - QU = R$, also $(I - Q)U = R$. Falls $I - Q$ invertierbar ist, ist $U = (I - Q)^{-1} R$

3. Stationäre Verteilungen

Sei (X_n) eine Markov-Kette mit Übergangsmatrix P und Startverteilung ν.

Dann ist $X_n \sim \nu \cdot P^n$. Im Allgemeinen hängt diese Verteilung von n ab. Es gibt aber spezielle Verteilungen ν, sodass die mit dieser Verteilung gestartete Kette zu jedem Zeitpunkt n die selbe Verteilung ν besitzt. Man sagt dann, die Kette ist *im Gleichgewicht* bzw. *stationär*.

Definition
Eine Abbildung $\nu : S \to \mathbb{R}_{\geq 0}$ heißt *Maß*.

NB: Ein Maß ν definiert ein Maß $\mathcal{P}(S) \to \mathbb{R}_{\geq 0}, A \mapsto \sum_{a \in A} \nu(a)$ im gewöhnlichen Sinne. Falls $\sum_{a \in S} \nu(a) = 1$, so definiert es sogar eine Verteilung.

Definition
Sei (X_n) eine Markov-Kette mit Übergangsmatrix $P = (p_{ij})$. Ein Maß ν heißt *invariant* für P, falls $\nu \cdot P = \nu$, d.h. falls gilt:

$$\sum_{i \in S} \nu(i) \cdot p_{ij} = \nu(j).$$

Ist ν eine Verteilung und invariant, so nennt man ν auch *stationäre Verteilung* oder *Gleichgewichtsverteilung.*

Bemerkung: a) Ist S endlich, so kann jedes (nichtdegenerierte) invariante Maß zu einer stationären Verteilung normiert werden.

b) Ist ν invariant, so gilt $\nu \cdot P^n = \nu$ $(n \in \mathbb{N}_0)$.

Ist ν eine stationäre Verteilung, so gilt

$$P_\nu(X_n = j) = \sum_{i \in S} \nu(i) \cdot p_{ij}^{(n)} = \nu(j),$$

d.h. die mit ν gestartete Kette hat zu jedem Zeitpunkt die Verteilung ν.

c) Ist P irreduzibel und $\nu \neq 0$ ein invariantes Maß, so ist $\nu(j) > 0$ für jedes $j \in S$.

Denn: $\nu \neq 0$, also existiert $i_0 \in S$ mit $\nu(i_0) > 0$. Wegen der Irreduzibilität gibt es ferner für jedes $j \in S$ ein $n \in \mathbb{N}_0$ mit $p_{i_0 j}^{(n)} > 0$. Zusammen:

$$\nu(j) = \sum_{i \in S} \nu(i) \cdot p_{ij}^{(n)} \geq \nu(i_0) p_{i_0 j}^{(n)} > 0$$

Gibt es immer ein invariantes Maß bzw. eine stationäre Verteilung? Ist es eindeutig?

Im Folgenden sei P irreduzibel. Wir definieren für ein beliebiges $k \in S$ das Maß γ_k wie folgt:

$$\gamma_k(i) := E_k[\sum_{n=1}^{T_k} 1_{[X_n = i]}]$$

Satz 3.1
Sei (X_n) irreduzibel und rekurrent, $k \in S$. Dann gilt:

a) γ_k ist ein invariantes Maß

b) $0 < \gamma_k < \infty$

c) γ_k ist das einzige invariante Maß mit $\gamma_k(k) = 1$ (d.h. γ_k ist eindeutig bis auf Vielfache).

Beweis
a) Zunächst gilt:

$$\begin{aligned}\gamma_k(i) &= E_k[\sum_{n=1}^{T_k} 1_{[X_n = i]}] = E_k[\sum_{n=1}^{\infty} 1_{[X_n = i, n \leq T_k]}] = \sum_{n=1}^{\infty} P_k(X_n = i, n \leq T_k) \\ &= \sum_{n=1}^{\infty} \sum_{j \in S, j \neq k} P_k(X_n = i, X_{n-1} = j, n \leq T_k)\end{aligned}$$

Für $j \neq k$ erhält man

$$\begin{aligned}
&P_k(X_n = i, X_{n-1} = j, n \leq T_k)\\
=&P(X_n = i, X_{n-1} = j, X_{n-2} \neq k, \ldots, X_1 \neq k, X_0 = k)/P(X_0 = k)\\
=&P(X_n = i \mid X_{n-1} = j, X_{n-2} \neq k, \ldots, X_1 \neq k, X_0 = k) \cdot P_k(X_{n-1} = j, X_{n-2} \neq k, \ldots, X_1 \neq k)\\
=&p_{ji} \cdot P_k(X_{n-1} = j, n \leq T_k)
\end{aligned}$$

Die so erhaltene Identität gilt auch für $j = k$ und es folgt

$$\begin{aligned}
\gamma_k(i) &= \sum_{n=1}^{\infty} \sum_{j \in S} P_k(X_n = i, X_{n-1} = j, n \leq T_k)\\
&= \sum_{n=1}^{\infty} \sum_{j \in S} p_{ji} \cdot P_k(X_{n-1} = j, n \leq T_k)\\
&= \sum_{j \in S} p_{ji} \sum_{n=0}^{\infty} P_k(X_n = j, n+1 \leq T_k)\\
&= \sum_{j \in S} p_{ji} E_k[\sum_{n=0}^{T_k - 1} 1_{[X_n = j]}]\\
&= \sum_{j \in S} p_{ji} E_k[\sum_{n=1}^{T_k} 1_{[X_n = j]}] = \sum_{j \in S} \gamma_k(j) p_{ji}
\end{aligned}$$

b) Nach Bemerkung c) oben gilt $\gamma_k > 0$, denn $\gamma_k(k) = 1$. Wegen $\gamma_k = \gamma_k \cdot P^n$ folgt

$$1 = \gamma_k(k) \geq \gamma_k(j) \cdot p_{jk}^{(n)}$$

für jedes $j \in S$. Es gibt allerdings mindestens ein $p_{jk}^{(n)} > 0$, denn die Markov-Kette ist irreduzibel; daran erkennt man $\gamma_k < \infty$.

c) Sei λ ein weiteres, invariantes Maß mit $\lambda(k) = 1$. Es gilt also:

$$\begin{aligned}
\lambda(j) &= \sum_{i \in S \setminus \{k\}} \lambda(i) \cdot p_{ij} + 1 \cdot p_{kj}\\
&= \sum_{i \in S \setminus \{k\}} \left(\sum_{l \in S \setminus \{k\}} \lambda(l) \cdot p_{li} + p_{ki} \right) p_{ij} + p_{kj}\\
&= \sum_{i,l \in S \setminus \{k\}} \lambda(l) \cdot p_{li} \cdot p_{ij} + \sum_{i \in S \setminus \{k\}} p_{ki} \cdot p_{ij} + p_{kj}\\
&= \sum_{i,l \in S \setminus \{k\}} \lambda(l) \cdot p_{li} \cdot p_{ij} + P_k(X_2 = j, T_k \geq 2) + P_k(X_1 = j, T_k \geq 1)
\end{aligned}$$

Iterativ erhalten wir für $n \in \mathbb{N}$:

$$\begin{aligned}
\lambda(j) &= \sum_{i_0, i_1, \ldots, i_n \in S \setminus \{k\}} \lambda(i_n) \prod_{r=1}^{n} p_{i_r i_{r-1}} p_{i_0 j} + \sum_{r=1}^{n+1} P_k(X_r = j, T_k \geq r)\\
&\geq \sum_{r=1}^{n+1} P_k(X_r = j, T_k \geq r) = E_k[\sum_{r=1}^{\min(n+1, T_k)} 1_{[X_r = j]}] \xrightarrow{n \to \infty} \gamma_k(j).
\end{aligned}$$

Es ist also $\lambda - \gamma_k$ ebenfalls ein invariantes Maß mit $(\lambda - \gamma_k)(k) = 0$; nach Bemerkung c) folgt $\lambda - \gamma_k = 0$, d.h. $\lambda = \gamma_k$. ■

Bemerkung: a) Ist S endlich und P irreduzibel, so folgt aus Satz 3.1, dass eine stationäre Verteilung existiert.

b) Ist (X_n) irreduzibel und transient, so kann keine stationäre Verteilung existieren, denn:

Angenommen es existiert eine stationäre Verteilung π. Dann ist

$$\pi(j) = \sum_{i \in S} \pi(i) p_{ij}^{(n)}.$$

Für $n \to \infty$ wird daraus (mit majorisierter Konvergenz und der bekannten Eigenschaft transienter Übergangswahrscheinlichkeiten):

$$\begin{aligned}\pi(j) &= \sum_{i \in S} \pi(i) \lim_{n \to \infty} p_{ij}^{(n)} \\ &= \sum_{i \in S} \pi(i) \cdot 0 \\ &= 0\end{aligned}$$

Definition
Für $i \in S$ sei

$$\begin{aligned}m_i \coloneqq E_i[T_i] &= \sum_{n=1}^{\infty} n \cdot P_i(T_i = n) + \infty \cdot (1 - f_{ii}^*) \\ &= \sum_{n=1}^{\infty} n \cdot f_{ii}^{(n)} + \infty \cdot (1 - f_{ii}^*)\end{aligned}$$

die *mittlere Rückkehrzeit* des Zustands i.

Bemerkung: Ist j transient, so folgt $m_j = \infty$.

Definition
Ein Zustand $i \in S$ heißt *positiv rekurrent*, falls $m_i < \infty$ und *nullrekurrent*, falls i rekurrent und $m_i = \infty$ ist.

Bemerkung: Jeder positiv rekurrente Zustand ist auch rekurrent.

Satz 3.2
Sei (X_n) eine irreduzible Markov-Kette. Die folgenden Aussagen sind äquivalent:

i) Es existiert eine stationäre Verteilung.

ii) Es gibt einen positiv rekurrenten Zustand $i \in S$.

iii) Alle Zustände in S sind positiv rekurrent.

Sind diese Bedingungen erfüllt, so ist die stationäre Verteilung eindeutig und durch

$$\pi(i) = \frac{1}{m_i}$$

gegeben.

Beweis

(iii) $\implies$ (ii) ✓

(ii) $\implies$ (i) Sei $k \in S$ mit $m_k < \infty$, dann ist (X_n) rekurrent und mit Satz 3.1 ist γ_k ein invariantes Maß. Dann gilt:

$$\begin{aligned}\sum_{j\in S}\gamma_k(j) &= \sum_{j\in S} E_k[\sum_{n=1}^{T_k} 1_{[X_n=j]}] \\ &= E_k[\sum_{n=1}^{T_k}\sum_{j\in S} 1_{[X_n=j]}] \\ &= E_k[\sum_{n=1}^{T_k} 1] \\ &= E_k[T_k] = m_k < \infty\end{aligned}$$

also ist γ_k normierbar.

(i) $\implies$ (iii) Sei π eine stationäre Verteilung und $k \in S$. Insbesondere sind $\pi(j) > 0$ für alle $j \in S$. Dann ist $\gamma := \frac{\pi}{\pi(k)}$ ein invariantes Maß mit $\gamma(k) = 1$. Nach Satz 3.1 c) folgt $\gamma = \gamma_k$. Beachte dass im Beweis von 3.1 c) die Voraussetzung (X_n) rekurrent nicht verwendet wurde. Wie oben ist

$$m_k = \sum_{j\in S}\gamma_k(j) = \frac{1}{\pi(k)}\sum_{j\in S}\pi(j) = \frac{1}{\pi(k)} < \infty.$$

Da $k \in S$ beliebig ist, folgt die Behauptung (iii).

Außerdem ist gezeigt, dass $\pi(k) = \frac{1}{m_k}$. ■

Bemerkung: i) Es gilt also die folgende Trichotomie für irreduzible Markov-Ketten, das heißt eine irreduzible Markov-Kette gehört immer zu genau einem der folgenden Fälle:

- Die Markov-Kette ist transient, es gibt keine stationäre Verteilung.
- Die Markov-Kette ist nullrekurrent, insbesondere gilt für alle $i, j \in S$:

  $$P_i(T_j < \infty) = 1 \text{ und } E_i[T_j] = \infty$$

  und es gibt ein (bis auf Vielfache) eindeutiges invariantes Maß, aber keine stationäre Verteilung.
- Die Markov-Kette ist positiv rekurrent, für alle $i, j \in S$ ist $E_i[T_j] < \infty$ und es gibt eine stationäre Verteilung.

ii) Ist S endlich und die Markov-Kette irreduzibel, so ist sie automatisch positiv rekurrent.

iii) Ist π eine stationäre Verteilung, so gilt:

$$\pi(i) = \frac{\gamma_k(i)}{\sum_{j\in S}\gamma_k(j)} = \frac{E_k[\sum_{n=1}^{T_k} 1_{[X_n=i]}]}{E_k[T_k]}.$$

$\pi(i)$ ist also der durchschnittliche Bruchteil der Zeit, den die Markov-Kette im Zustand i verbringt, während sie einen Zyklus durchläuft.

Beispiel 3.1
Sei $S = \{1, 2\}$. Die Übergangsmatrix P sei gegeben durch

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \quad \alpha, \beta \in (0, 1].$$

Also ist die Markov-Kette irreduzibel und positiv rekurrent. Die stationäre Verteilung ist die Lösung des linearen Gleichungssystems

$$\pi P = \pi$$

unter Berücksichtigung von $\pi \geq 0$ und $\pi(1) + \pi(2) = 1$, also

$$\pi(1) = \frac{\beta}{\alpha + \beta}, \quad \pi(2) = \frac{\alpha}{\alpha + \beta}.$$

Beispiel 3.2 (Irrfahrt)
Siehe Beispiel 2.3: Für $p \neq q$ ist die Markov-Kette transient. Existiert ein invariantes Maß? Ansatz:

$$\gamma(j) = \sum_{i \in \mathbb{Z}} \gamma(i) p_{ij} = \gamma(j+1) \cdot q + \gamma(j-1) \cdot p$$

$$\Longrightarrow \gamma(j+1) - \gamma(j) = \frac{p}{q}(\gamma(j) - \gamma(j-1))$$

Also: $\gamma_1(j) = 1$ und $\gamma_2(j) = (\frac{p}{q})^j$, $j \in \mathbb{Z}$, sind verschiedene invariante Maße.

Ist $p = q = \frac{1}{2}$, so ist die Markov-Kette rekurrent. Ist sie nullrekurrent oder positiv rekurrent? Es gibt keine stationäre Verteilung (siehe oben), die Markov-Kette ist also nullrekurrent.

Beispiel 3.3 (Geburts- und Todesprozess in diskreter Zeit)
Es sei (X_n) eine Markov-Kette mit $S = \mathbb{N}_0$ und Übergangsmatrix

$$P = \begin{pmatrix} p_{00} & p_{01} & 0 & \cdots & \\ p_{10} & p_{11} & p_{12} & 0 & \cdots \\ 0 & p_{21} & p_{22} & p_{23} & \ddots \\ \vdots & 0 & p_{32} & p_{33} & \ddots \\ & \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

mit $p_{01} > 0$, $p_{i,i+1} > 0$, $p_{i,i-1} > 0$ für alle $i \geq 1$, also ist (X_n) irreduzibel. Wann ist (X_n) positiv rekurrent?

Der Ansatz $\pi P = \pi$ liefert:

$$\begin{aligned} \pi(0) &= p_{00} \cdot \pi(0) + p_{10} \cdot \pi(1) \\ \pi(i) &= p_{i-1,i} \cdot \pi(i-1) + p_{ii} \cdot \pi(i) + p_{i+1,i} \cdot \pi(i+1) \\ \iff p_{i,i-1} \cdot \pi(i) + p_{i,i+1} \pi(i) &= p_{i-1,i} \cdot \pi(i-1) + p_{i+1,i} \cdot \pi(i+1) \\ \iff p_{i+1,i} \cdot \pi(i+1) - p_{i,i+1} \cdot \pi(i) &= p_{i,i-1} \cdot \pi(i) - p_{i-1,i} \cdot \pi(i-1) \\ &= \ldots = p_{10} \cdot \pi(1) - p_{01} \cdot \pi(0) \end{aligned}$$

Aus der ersten Gleichung folgt $p_{01} \cdot \pi(0) = p_{10} \cdot \pi(1)$ und damit:

$$\begin{aligned} p_{i+1,i} \cdot \pi(i+1) - p_{i,i+1} \cdot \pi(i) = 0 \\ \implies \pi(i+1) &= \frac{p_{i,i+1}}{p_{i+1,i}} \pi(i) \\ &= \ldots = \pi(0) \cdot \prod_{k=0}^{i} \frac{p_{k,k+1}}{p_{k+1,k}}. \end{aligned}$$

Für $\pi(0) > 0$ erhält man ein invariantes Maß. (X_n) ist positiv rekurrent, falls

$$\sum_{i=1}^{\infty} \prod_{k=0}^{i} \frac{p_{k,k+1}}{p_{k+1,k}} < \infty.$$

Im Spezialfall $p_{k,k+1} = p$, $p_{k,k-1} = q = 1 - p$, $k \geq 1$ und $p_{01} = p$, $p_{00} = 1 - p$ gilt

$$(X_n) \text{ ist positiv rekurrent } \iff \sum_{i=1}^{\infty} \left(\frac{p}{q}\right)^{i+1} < \infty \iff p < q.$$

4. Konvergenz gegen die stationäre Verteilung

Sei (X_n) eine Markov-Kette mit Übergangsmatrix P. Wir nehmen an, dass (X_n) bzw. P aperiodisch ist, das heißt: Für alle Zustände $i \in S$ gilt:

$$d_i := \mathrm{ggT}\{n \in \mathbb{N} \mid p_{ii}^{(n)} > 0\} = 1$$

Lemma 4.1
P ist genau dann irreduzibel und aperiodisch, wenn für alle $i, j \in S$ eine Zahl $n_0 \in \mathbb{N}$ existiert, sodass für alle $n \geq n_0$ gilt: $p_{ij}^{(n)} > 0$.

(ohne Beweis)

Satz 4.2 (Konvergenzsatz)
Es sei (X_n) irreduzibel, aperiodisch und positiv rekurrent mit stationärer Verteilung π. Dann gilt für alle $i, j \in S$:

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi(j) = \frac{1}{m_j}$$

Beweis
Wir benutzen ein sogenanntes "Kopplungsargument".

(1) Sei (Y_n) eine weitere Markov-Kette, unabhängig von (X_n) mit gleicher Übergangsmatrix und Startverteilung π, also $Y_n \sim \pi$ für alle $n \in \mathbb{N}$. Es sei

$$T \coloneqq \inf\{n \in \mathbb{N} \mid X_n = Y_n\}$$

die Treffzeit der Markov-Ketten. Wir zeigen zunächst: $P(T < \infty) = 1$.

Offenbar ist $(X_n, Y_n)_{n\in\mathbb{N}_0}$ eine Markov-Kette auf S^2 mit Übergangsmatrix $\hat{P} = (\hat{p}_{(ij)(kl)})$, wobei $\hat{p}_{(ij)(kl)} = p_{ik} \cdot p_{jl}$. Weiter ist $\hat{p}^{(n)}_{(ij)(kl)} = p^{(n)}_{ik} \cdot p^{(n)}_{jl}$ und mit Lemma 4.1: Die Kette (X_n, Y_n) ist irreduzibel und aperiodisch. Man kann nachrechnen:

$$\hat{\pi}(i,j) \coloneqq \pi(i) \cdot \pi(j)$$

ist eine stationäre Verteilung für (X_n, Y_n), also ist sie nach Satz 3.2 positiv rekurrent.

Sei $X_0 = i$ und die Startverteilung $\hat{\nu}$ von (X_n, Y_n) gegeben durch $\hat{\nu}(k,l) = \delta_i(k) \cdot \pi(l)$. Für $b \in S$ sei

$$T_{(b,b)} \coloneqq \inf\{n \in \mathbb{N} \mid (X_n, Y_n) = (b,b)\}.$$

Offenbar ist $T \leq T_{(b,b)}$ und $P_{\hat{\nu}}(T_{(b,b)} < \infty) = 1$. Daraus folgt, dass $P_{\hat{\nu}}(T < \infty) = 1$.

(2) Betrachte $(Z_n)_{n\in\mathbb{N}_0}$ definiert durch

$$Z_n = \begin{cases} X_n, & \text{für } n \leq T \\ Y_n, & \text{für } n > T. \end{cases}$$

Es ist (Z_n) eine Markov-Kette mit Übergangsmatrix P und $Z_0 = i$, denn:

Nach Satz 1.1 genügt es zu zeigen, dass

$$\forall n \in \mathbb{N}, i_k \in S : P_{\hat{\nu}}(Z_k = i_k,\, 0 \leq k \leq n) = \delta_i(i_0) \prod_{k=0}^{n-1} p_{i_k, i_{k+1}}$$

Es gilt

$$\begin{aligned}
&P_{\hat{\nu}}(Z_k = i_k,\, 0 \leq k \leq n) \\
&= \sum_{r=0}^{n} P_{\hat{\nu}}(Z_k = i_k,\, 0 \leq k \leq n,\, T = r) \\
&\quad + P_{\hat{\nu}}(Z_k = i_k,\, 0 \leq k \leq n,\, T > n) \\
&= \sum_{r=0}^{n} \underbrace{P_{\hat{\nu}}(X_k = i_k,\, 0 \leq k \leq r,\, Y_k = i_k,\, r+1 \leq k \leq n,\, Y_0 \neq i_0, \ldots, Y_{r-1} \neq i_{r-1},\, Y_r = i_r)}_{=:\mathrm{I}} \\
&\quad + \underbrace{P_{\hat{\nu}}(X_k = i_k,\, 0 \leq k \leq n,\, Y_0 \neq i_0, \ldots, Y_n \neq i_n)}_{=:\mathrm{II}}
\end{aligned}$$

mit

$$\begin{aligned}
\mathrm{I} &= P_{\hat\nu}(X_k = i_k, 0 \le k \le r) \cdot P_{\hat\nu}(Y_k = i_k, r+1 \le k \le n | Y_0 \ne i_0, \ldots, Y_{r-1} \ne i_{r-1}, Y_r = i_r) \cdot \\
&\quad P_{\hat\nu}(Y_0 \ne i_0, \ldots, Y_{r-1} \ne i_{r-1}, Y_r = i_r) \\
&= \delta_i(i_0) \cdot \prod_{k=0}^{r-1} p_{i_k, i_{k+1}} \cdot \prod_{k=r}^{n-1} p_{i_k, i_{k+1}} \cdot P_{\hat\nu}(Y_0 \ne i_0, \ldots, Y_{r-1} \ne i_{r-1}, Y_r = i_r) \\
&= \delta_i(i_0) \cdot \prod_{k=0}^{n-1} p_{i_k, i_{k+1}} \cdot P_{\hat\nu}(Y_0 \ne i_0, \ldots, Y_{r-1} \ne i_{r-1}, Y_r = i_r), \\
\mathrm{II} &= \ldots = \delta_i(i_0) \cdot \prod_{k=0}^{n-1} p_{i_k, i_{k+1}} \cdot P_{\hat\nu}(Y_0 \ne i_0, \ldots, Y_{n-1} \ne i_{n-1}, Y_n \ne i_n)
\end{aligned}$$

Tatsächlich gilt also

$$P_{\hat\nu}(Z_k = i_k, 0 \le k \le n) = \delta_i(i_0) \cdot \prod_{k=0}^{n-1} p_{i_k, i_{k+1}}$$

(3) Es gilt nun

$$\begin{aligned}
p_{i,j}^{(n)} &= P_{\hat\nu}(Z_n = j) = P_{\hat\nu}(Z_n = j, T \le n) + P_{\hat\nu}(Z_n = j, T > n) \\
\pi(j) &= P_{\hat\nu}(Y_n = j) = \underbrace{P_{\hat\nu}(Y_n = j, T \le n)}_{= P_{\hat\nu}(Z_n = j, T \le n)} + P_{\hat\nu}(Y_n = j, T > n) \\
\Rightarrow \quad & |p_{i,j}^{(n)} - \pi(j)| \le 2 \cdot P_{\hat\nu}(\underbrace{T > n}_{\downarrow \{T = \infty\}}) \longrightarrow 0
\end{aligned}$$

■

Satz 4.3
Seien i, $j \in S$ Zustände sowie d_j die Periode von j. Dann gilt:

$$\lim_{n\to\infty} p_{i,j}^{(n \cdot d_j + r)} = \frac{d_j}{m_j} \sum_{k=0}^{\infty} f_{i,j}^{(k \cdot d_j + r)} \qquad r = 1, \ldots, d_j$$

Speziell:

a) Ist j transient oder nullrekurrent (d.h. $m_j = \infty$), so gilt $p_{i,j}^{(n)} \longrightarrow 0$

b) Ist j aperiodisch (d.h. $d_j = 1$), so gilt $p_{i,j}^{(n)} \longrightarrow \frac{1}{m_j} f_{i,j}^*$

5. Markov-Ketten und Martingale

Erinnerung: (X_n) heißt *Martingal*, falls

a) $E|X_n| < \infty$

b) $E[X_{n+1}|X_1, \dots, X_n] = X_n$, bzw.

b') $E[X_{n+1}|\mathcal{F}_n] = X_n$, wobei $\mathcal{F}_n := \sigma(X_1, \dots, X_n)$ die *natürliche Filtration* bezeichnet.

Erinnerung: Sei $(\Omega, \mathcal{F}, P)$ W-Raum, X Zufallsvariable mit $E|X| < \infty$, $\mathcal{G} \subseteq \mathcal{F}$ Unter-σ-Algebra. $Z \coloneqq E[X|\mathcal{G}]$ heißt bedingter Erwartungswert von X bzgl. $\mathcal{G}$, falls

a) Z ist $\mathcal{G}$-messbar

b) $\int_A Z \cdot dP = \int_A X \cdot dP \quad \forall A \in \mathcal{G}$

Sei (X_n) eine Markov-Kette auf dem Zustandsraum S mit Übergangsmatrix P sowie $\mathcal{F}_n := \sigma(X_1, \dots, X_n)$ die natürliche Filtration. Weiter sei $h : S \to \mathbb{R}$ und $P : (S \to \mathbb{R}) \to (S \to \mathbb{R})$ definiert durch

$$(Ph)(i) := \sum_{j \in S} p_{i,j} \cdot h(j) \qquad \forall i \in S$$

NB: "Ph" macht Sinn im Sinne einer "Matrix·Vektor"-Multiplikation.

Lemma 5.1
Sei $h : S \to \mathbb{R}$ eine Funktion mit $\|P\,|h|\|_\infty < \infty$. Dann gilt:

$$(Ph)(X_n) = E[h(X_{n+1})|\mathcal{F}_n]$$

Beweis
Wir prüfen nach, dass $(Ph)(X_n)$ ein bedingter Erwartungswert von $h(X_{n+1})$ bzgl. $\mathcal{F}_n$ ist.

a) $(Ph)(X_n)$ ist $\mathcal{F}_n$-messbar, denn X_n ist $\mathcal{F}_n$-messbar.

b) Sei $A := \{X_0 = i_0, \dots, X_n = i_n\} \in \mathcal{F}_n$. Dann:

$$\begin{aligned}
\int_A (Ph)(X_n) \cdot dP &= \int 1_A \cdot (Ph)(X_n) \cdot dP = \int 1_{\{X_0=i_0,\dots,X_n=i_n\}}(Ph)(i_n) \cdot dP \\
&= (Ph)(i_n) \cdot P(A) = \sum_{j \in S} p_{i_n,j} \cdot h(j) \cdot P(X_0 = i_0) \cdot \prod_{k=0}^{n-1} p_{i_k,i_{k+1}} \\
&= \sum_{j \in S} h(j) \cdot P(X_0 = i_0, \dots, X_n = i_n, X_{n+1} = j) = \sum_{j \in S} \int_{A \cap \{X_{n+1}=j\}} h(X_{n+1}) \cdot dP \\
&= \sum_{j \in S} \int_A 1_{\{X_{n+1}=j\}} h(X_{n+1}) \cdot dP \overset{LDC}{=} \int_A h(X_{n+1}) \cdot dP
\end{aligned}$$

Da jedes $A \in \mathcal{F}_n$ abzählbare Vereinigung solcher "Elementarereignisse" ist, folgt die Behauptung. ∎

Satz und Definition 5.2
Seien P eine Übergangsmatrix auf S und $h : S \to \mathbb{R}$ eine Funktion mit $\|P\,|h|\|_\infty < \infty$. Gilt dann $Ph = h$, so nennen wir h *harmonisch*. Im Fall $Ph \geq h$ bzw. $Ph \leq h$ heißt h *sub-* bzw. *superharmonisch*.

Ist (X_n) eine Markov-Kette mit Übergangsmatrix P und h [sub-/super-]harmonisch, so ist $(h(X_n))$ ein $(\mathcal{F}_n)$-[Sub-/Super-]Martingal.

Beweis
Nach Lemma 5.1 gilt

$$E[h(X_{n+1})|\mathcal{F}_n] = Ph(X_n) = h(X_n) + (Ph - h)(X_n)$$

$\rightsquigarrow$ Behauptung. ∎

Bemerkung: Ist $h : S \to \mathbb{R}$ eine beschränkte Funktion, so wird durch

$$\begin{aligned} Z_n^h &:= h(X_n) - \sum_{k=1}^{n} (E[h(X_k) \mid \mathcal{F}_{k-1}] - h(X_{k-1})) \\ &= h(X_n) - \sum_{k=1}^{n} (Ph - h)(X_{k-1}) \end{aligned}$$

ein Martingal definiert, genannt Levi-Martingal zu (X_n).

Die Markoveigenschaft lässt sich über Levi-Martingale charakterisieren:

Satz 5.3
Es sei $(X_n)_{n\in\mathbb{N}_0}$ ein stochastischer Prozess mit Werten in S und natürlicher Filtration $(\mathcal{F}_n)$, d.h. $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$, und P sei eine stochastische Matrix auf S. Ist dann für alle beschränkten $h : S \to \mathbb{R}$ der Prozess

$$Z_n^h := h(X_n) - \sum_{k=1}^{n} (Ph - h)(X_{k-1})$$

ein $(\mathcal{F}_n)$-Martingal, so ist X_n eine homogene Markov-Kette mit Übergangsmatrix P.

Beweis
Aus $E[Z_{n+1}^h \mid \mathcal{F}_n] = Z_n^h$ erhält man

$$E[h(X_{n+1}) \mid \mathcal{F}_n] = (Ph)(X_n)$$

bzw.

$$\int_A h(X_{n+1})dP = \int_A (Ph)(X_n)dP \text{ für alle } A \in \mathcal{F}_n$$

Es sei $i_0, i_1, \ldots, i_{n+1} \in S$ und $A := \{X_0 = i_0, \ldots, X_n = i_n\}$ und wir setzen $h := 1_{\{i_{n+1}\}}$. Die linke Seite der letzten Gleichung ergibt dann

$$\int_A h(X_{n+1})dP = P(X_0 = i_0, \ldots, X_n = i_n, X_{n+1} = i_{n+1})$$

und die rechte Seite ergibt

$$\begin{aligned}\int_A (Ph)(X_n)dP &= \int_{\{X_0=i_0,\ldots,X_n=i_n\}} (Ph)(i_n)dP \\ &= \int_{\{X_0=i_0,\ldots,X_n=i_n\}} p_{i_n i_{n+1}} dP \\ &= p_{i_n i_{n+1}} \cdot P(X_0 = i_0, \ldots, X_n = i_n).\end{aligned}$$

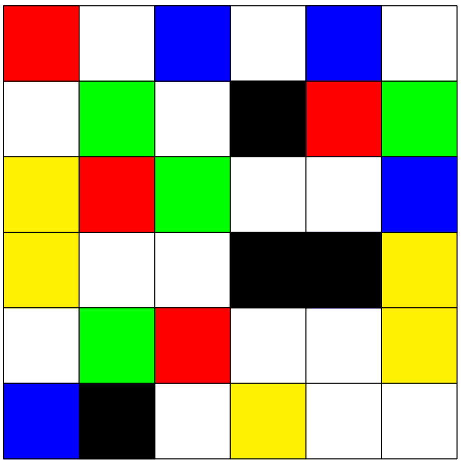
Durch Teilen der rechten Wahrscheinlichkeit erhält man

$$P(X_{n+1} = i_{n+1} \mid X_0 = i_0, \ldots, X_n = i_n) = p_{i_n i_{n+1}}. \qquad \blacksquare$$

Bemerkung: Es gilt: Ist $(X_n)_{n\in\mathbb{N}_0}$ ein nicht-negatives Supermartingal, so gibt es eine Zufallsvariable X_∞ mit

$$\lim_{n\to\infty} X_n = X_\infty \quad P-\text{f.s.}$$

Beispiel 5.1
Jedem Feld eines schachbrettartigen $N \times N$-Gitters wird eine von L möglichen Farben zugewiesen. Diese Färbung wird mittels Zufallsexperimenten modifiziert: Eine Zelle wird gleichverteilt gewählt, daraufhin wird einer der vier Nachbarn (modulo N) zufällig gewählt und dessen Farbe dem zuerst gewählten Feld zugewiesen. Offensichtlich ist dies eine Markov-Kette und die monochromen Zustände sind die absorbierenden.



Formal seien $L, N \in \mathbb{N}$, $L, N \geq 2$. $I := \{1, \ldots, N\}^2$, $S := \{1, \ldots, L\}^I = \{f : I \to \{1, \ldots, L\}\}$. Sei (X_n) die Markov-Kette in S, die die Zustandsfolge angibt.

Wie verhält sich die Folge für $n \to \infty$? Sei $l \in \{1, \ldots, L\}$ fest und Y_n sei die Anzahl der Felder mit Farbe l (zum Beispiel: blau) im Zustand X_n. Sei (A, B) ein Nachbarpaar im Gitter. Wäre dies die Wahl in einem Zustandsübergang, so gelte: Ist $X_n(A) = X_n(B)$ oder $X_n(A) \neq l, X_n(B) \neq l$, so gilt auch $Y_{n+1} = Y_n$. Ist dagegen $X_n(A) = l$ und $X_n(B) \neq l$, so ist $Y_{n+1} = Y_n - 1$. Ist letztlich $X_n(A) \neq l$ und $X_n(B) = l$, so ist $Y_{n+1} = Y_n + 1$.

Die Wahrscheinlichkeit, erst A, dann B zu wählen ist gleich der Wahrscheinlichkeit, erst B und dann A zu wählen. Damit ist

$$E[Y_{n+1} \mid X_1, \ldots, X_n] = Y_n.$$

Sei $\mathcal{F}_n \coloneqq \sigma(X_0, \ldots, X_n)$. Dann ist $(Y_n)_{n\in\mathbb{N}_0}$ ein $(\mathcal{F}_n)$-Martingal. Nach der Bemerkung folgt: $Y_n \to Y_\infty$ für ein $n \to \infty$ P-fast-sicher. Da (Y_n) ganzzahlig ist, ist $Y_n(\omega)$ konstant ab einem $n \geq n_0(\omega)$. Als Konstanten kommen nur 0 und N^2 in Frage, denn für $k \in \{1, \ldots, N^2 - 1\}$ gilt

$$\begin{aligned}
P(Y_{n+j} = k \mid Y_n = \ldots = Y_{n+j-1} = k) &\leq 1 - \frac{1}{N^2 4} \\
\implies P(Y_n = \ldots = Y_{n+j} = k) &\leq (1 - \frac{1}{N^2 4})^j \\
\implies P(\underbrace{Y_m = k, \forall m \geq n}_{=:A_n}) &= 0
\end{aligned}$$

Es gilt $\{\omega \in \Omega \mid \exists n \ \forall m \geq n : Y_m(\omega) = k\} = \bigcup_{n=0}^\infty A_n$. Damit ist

$$P(\lim_{n\to\infty} Y_n = k) = P(\exists n \ \forall m \geq n : Y_n = k) \leq \sum_{n=0}^\infty P(A_n) = 0$$

und wir folgern, dass $P(Y_\infty \in \{0, N^2\}) = 1$.

Außerdem gilt noch, da Y_n beschränkt ist:

$$EY_\infty = \lim_{n\to\infty} EY_n = EY_0.$$

Damit ist die Wahrscheinlichkeit, dass das Feld irgendwann komplett blau ist, gleich

$$P(Y_\infty = N^2) = \frac{1}{N^2} EY_\infty = \frac{1}{N^2} EY_0 = \frac{1}{N^2} \#\{A \in I \mid X_0(A) = l\}.$$

Anwendungen dieses Modells findet man in der Physik (Vielteilchensysteme), in der Biologie (Ausbreitung von Infektionen) oder in der Finanzmathematik (Kreditrisiken).

6. Die starke Markov-Eigenschaft

Gegeben sei eine Markov-Kette $(X_n)_{n\in\mathbb{N}_0}$ mit Zustandsraum S auf dem Wahrscheinlichkeitsraum $(\Omega = S_0^{\mathbb{N}_0}, \mathcal{F} \coloneqq \otimes_{n=0}^\infty \mathcal{P}(S), P)$. Beachte, dass die Mengen

$$Z(i_0, i_1, \ldots, i_n) \coloneqq \{i_0\} \times \{i_1\} \times \cdots \times \{i_n\} \times S \times S \times \ldots$$

für $n \in \mathbb{N}_0$, $i_0, \ldots i_n \in S$ ein durchschnittsstabiles Erzeugendensystem für $\mathcal{F}$ bilden. Weiter sei $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ mit $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$ die natürliche Filtration von (X_n).

$\tau : \Omega \to \mathbb{N}_0$ sei eine $(\mathcal{F}_n)$-Stoppzeit, das heißt $\{\tau \leq n\} \in \mathcal{F}_n$ und $P(\tau < \infty) = 1$. Die gestoppte Markov-Kette $X^\tau = (X_n^\tau)_{n\in\mathbb{N}_0}$ ist

$$X_n^\tau \coloneqq X_{\min(\tau,n)}$$

für $n \in \mathbb{N}_0$ und $Y = (Y_n)_{n\in\mathbb{N}_0}$ definiert durch

$$Y_n := X_{\tau+n}$$

heißt der Post-τ-Prozess.

Satz 6.1 (Starke Markov-Eigenschaft)
Mit den oben eingeführten Bezeichnungen gilt:

a) Y ist eine Markov-Kette mit Übergangsmatrix P und Startverteilung ν, wobei $X_\tau \backsim \nu$.

b) X^τ und Y sind unter X_τ bedingt unabhängig.

Beweis
a.) Es gilt:

$$\begin{aligned}
&P(Y_0 = i_0, \dots, Y_n = i_n, Y_{n+1} = i_{n+1}) \\
=&\sum_{k=0}^{\infty} P(\tau = k, X_k = i_0, \dots, X_{k+n+1} = i_{n+1}) \\
=&\sum_{k=0}^{\infty} P(X_{k+n+1} = i_{n+1} \mid X_k = i_0, \dots, X_{k+n} = i_n, \tau = k) \cdot P(X_k = i_0, \dots, X_{k+n} = i_n, \tau = k) \\
=&p_{i_n i_{n+1}} \sum_{k=0}^{\infty} P(X_k = i_0, \dots, X_{k+n} = i_n, \tau = k) \\
=&p_{i_n i_{n+1}} P(Y_0 = i_0, \dots, Y_n = i_n) \implies \text{ Behauptung.}
\end{aligned}$$

b.) Seien

$$\begin{aligned}
A &:= Z(i_0, \dots, i_m) = \{i_0\} \times \dots \times \{i_m\} \times S \times S \times \dots \\
B &:= Z(j_0, \dots, j_n),
\end{aligned}$$

dann gilt:

$$\begin{aligned}
&P(X^\tau \in A, Y \in B, X_\tau = j) \\
=&\sum_{k=0}^{\infty} P(\tau = k, X_{0\wedge k} = i_0, \dots, X_{m\wedge k} = i_m, X_k = j_0, \dots, X_{k+n} = j_n, X_k = j) \\
=&\sum_{k=0}^{\infty} P(X_k = j_0, \dots, X_{k+n} = j_n \mid X_{0\wedge k} = i_0, \dots, X_{m\wedge k} = i_m, X_k = j, \tau = k) \\
&\quad \cdot P(X_{0\wedge k} = i_0, \dots, X_{m\wedge k} = i_m, X_k = j, \tau = k) \\
=&P(X_0 = j_0, \dots, X_n = j_n \mid X_0 = j) \cdot P(X^\tau \in A, X_\tau = j) \\
=&P(Y \in B \mid X_\tau = j) \cdot P(X^\tau \in A \mid X_\tau = j) \cdot P(X_\tau = j)
\end{aligned}$$

Teilen durch $P(X_\tau = j) \implies$ Behauptung. ∎