

10 Markov-Ketten

10.1 Die Kain und Abel-Aufgabe (nach A. Engel)

Abel schlägt seinem Bruder Kain folgendes Spiel vor. Sie werfen abwechselnd eine faire Münze, bis erstmals eine der Ziffernfolgen

- a) 1111 oder
- b) 0011 auftritt.

Kain gewinnt bei (a) und Abel bei (b). Wie groß ist die Wahrscheinlichkeit, daß Kain gewinnt? Ein möglicher Pfad im Verlauf des Spiels ist 01100101110011.

Wir stellen nun die Struktur der Aufgabe mit Hilfe eines Graphen dar, wobei die “Zustände” die möglichen Ergebnismuster sind und die Pfeile die möglichen Übergänge angeben. Alle Übergänge geschehen mit Wahrscheinlichkeit $\frac{1}{2}$.

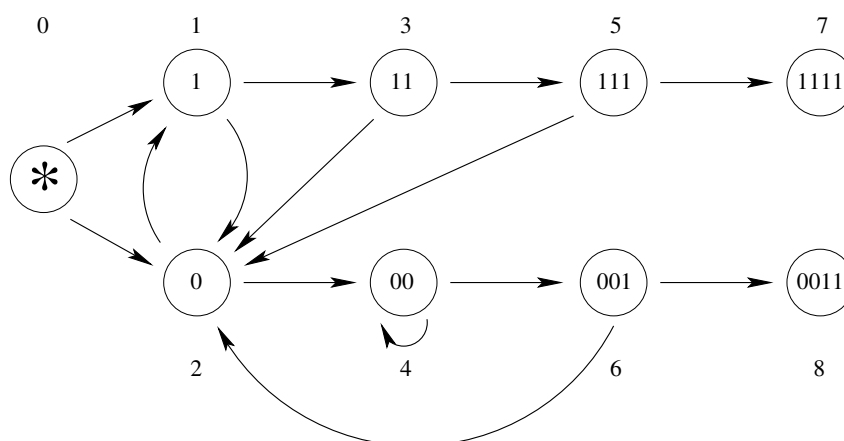


Abbildung 10.1: Der Graph der möglichen Übergänge

Die Zustände sind mit den Ziffern 0 bis 8 bezeichnet, wobei “0” den Zustand benennt, bei dem noch kein Ergebnis vorliegt. Der angegebene Spielverlauf 01100101110011 übersetzt sich nun (in eindeutiger Zuordnung) in 21324621352468. Wie läßt sich nun die Aufgabe lösen? Wir bezeichnen mit p_i die Wahrscheinlichkeit bei Start im Zustand “ i ” den Zustand “7” zu erreichen bevor man den Zustand “8” erreicht hat. Dann gilt natürlich sofort $p_7 = 1$ und $p_8 = 0$. Außerdem sind die folgenden Gleichungen intuitiv plausibel.

$$p_1 = \frac{1}{2}p_2 + \frac{1}{2}p_3$$

$$p_2 = \frac{1}{2}p_1 + \frac{1}{2}p_4$$

$$p_3 = \frac{1}{2}p_2 + \frac{1}{2}p_5$$

$$p_4 = \frac{1}{2}p_4 + \frac{1}{2}p_6$$

$$p_5 = \frac{1}{2} + \frac{1}{2}p_2$$

$$p_6 = \frac{1}{2}p_2.$$

Das Schema ist ein lineares Gleichungssystem mit sechs Unbekannten und Gleichungen. Durch Einsetzen erhält man sofort

$$p_4 = p_6 = \frac{1}{2}p_2$$

$$p_1 = \frac{3}{2}p_2$$

$$p_3 = \frac{1}{4} + \frac{3}{4}p_2$$

$$\text{sowie } p_1 = \frac{1}{8} + \frac{7}{8}p_2.$$

Damit ist $p_1 = \frac{3}{10}$, $p_2 = \frac{1}{5}$ und folglich $p_0 = \frac{1}{4}$. Das heißt Kain gewinnt nur mit Wahrscheinlichkeit $\frac{1}{4}$.

10.2 Definition von Markov-Ketten und erste Folgerungen

10.2.1 Definition (Markov-Kette)

(Ω, P) sei Wahrscheinlichkeitsraum, E sei endliche Menge. Seien $X_i : \Omega \rightarrow E$, $i = 0, 1, 2, 3, \dots$ Zufallsvariablen. Die Menge der Zufallsvariablen $\mathcal{X} = \{X_0, X_1, X_2, \dots, X_n, \dots\}$ heißt Markov-Kette (der Länge n), falls

$$P(X_i = x_i \mid X_0 = x_0, \dots, X_{i-1} = x_{i-1}) = P(X_i = x_i \mid X_{i-1} = x_{i-1})$$

für $i = 1, 2, \dots, n, \dots$ gilt. Eine Markov-Kette \mathcal{X} heißt stationär, falls

$$P(X_i = x_1 \mid X_{i-1} = x_0) = P(X_1 = x_1 \mid X_0 = x_0)$$

für $i = 1, 2, \dots$ gilt.

Bezeichnungsweisen

1. E heißt Zustandsraum,
2. $q(x, y), x, y \in E$ heißt stochastische Matrix, falls $q(x, y) \geq 0$ und $\sum_{y \in E} q(x, y) = 1$,
 $\forall x \in E$ gilt,
3. $\pi(x_0) = P(X_0 = x_0)$ mit $x_0 \in E$ heißt Startverteilung.

Konstruktion von stationären Markov-Ketten

Gegeben seien:

1. E endlich oder abzählbar,
2. $q(x, y), x, y \in E$ eine stochastische Matrix,
3. $\pi(x), x \in E$ eine Wahrscheinlichkeitsfunktion auf E , d.h. $\pi(x) \geq 0$ für alle $x \in E$ und $\sum_{x \in E} \pi(x) = 1$.

Zur Existenz:

Sei $\Omega_{n+1} := \{\omega = (x_0, \dots, x_n) \mid x_i \in E, i = 0, \dots, n\}$ und $X_i(\omega) := x_i$, für $i = 0, 1, \dots, n$.

10.2.2 Satz:

Sei $n \in \mathbb{N}$. Durch $p(\omega) := \pi(x_0)q(x_0, x_1) \dots q(x_{n-1}, x_n)$ wird eine Wahrscheinlichkeitsfunktion auf Ω_{n+1} gegeben, so dass gilt:

1. $P(X_0 = x_0) = \pi(x_0)$,
2. $P(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = p(\omega)$,
3. $P(X_{i+1} = x_{i+1} \mid X_0 = x_0, \dots, X_i = x_i) = q(x_i, x_{i+1}) = P(X_{i+1} = x_{i+1} \mid X_i = x_i)$
für $i = 0, \dots, n-1$.

Beweis:**Zu (2):**

Sei $\omega = (x_0, \dots, x_n)$. Mit $p(\omega) = \pi(x_0)q(x_0, x_1) \dots q(x_{n-1}, x_n)$ gilt: $p(\omega) \geq 0$ und $\sum_{\omega \in \Omega_{n+1}} p(\omega) = 1$. Da $X_i(\omega) = x_i$ ist, folgt $P(\{\omega \mid X_i(\omega) = x_i, i = 0, \dots, n\}) = P(\{\omega\}) = p(\omega)$.

Zu (1):

$$\begin{aligned} (+) P(\{X_0 = x_0, \dots, X_i = x_i\}) &= \sum_{x_{i+1}, \dots, x_n} \pi(x_0)q(x_0, x_1)q(x_{i-1}, x_i) \dots q(x_{n-1}, x_n) \\ &= \pi(x_0)q(x_0, x_1) \dots q(x_{i-1}, x_i) \end{aligned}$$

für $i = 0, 1, 2, \dots, n-1$.

Zu (3):

Es gilt

$$\begin{aligned} P(X_{i+1} = x_{i+1} \mid X_0 = x_0, \dots, X_i = x_i) &= \frac{\pi(x_0)q(x_0, x_1) \dots q(x_{i-1}, x_i)q(x_i, x_{i+1})}{\pi(x_0)q(x_0, x_1) \dots q(x_{i-1}, x_i)} \\ &= q(x_i, x_{i+1}) \end{aligned}$$

sowie

$$\begin{aligned} P(X_{i+1} = x_{i+1} \mid X_i = x_i) &= \frac{P(X_{i+1} = x_{i+1}, X_i = x_i)}{P(X_i = x_i)} \\ &= \frac{\sum_{x_0, \dots, x_{i-1}} \pi(x_0)q(x_0, x_1) \dots q(x_{i-1}, x_i)q(x_i, x_{i+1})}{\sum_{x_0, \dots, x_{i-1}} \pi(x_0)q(x_0, x_1) \dots q(x_{i-1}, x_i)} \\ &= q(x_i, x_{i+1}). \end{aligned}$$

Bemerkungen:

$$1. P(X_i = x_i) = \sum_{x_0, \dots, x_{i-1}} \pi(x_0)q(x_0, x_1) \dots q(x_{i-1}, x_i) = \sum_{x_0} \pi(x_0)q^i(x_0, x_i).$$

Dabei ist $q^{n+1}(x, y) = \sum_{z \in E} q^n(x, z)q(z, y)$ das $n+1$ -fache Matrixprodukt.

2. Mit Satz 10.2.2 ist eine Markov-Kette der Länge n konstruiert!
3. Wegen (+) im Beweis von Satz 10.2.2 ist die Markov-Kette der Länge i eingebettet in eine Markov-Kette der Länge n für $i \leq n$.
4. Mit der Bedingung (+) lässt sich eine Markov-Kette beliebiger Länge konstruieren.

10.2.3 Beispiele für Markov-Ketten

1) Kain und Abel-Aufgabe:

Der Zustandsraum ist $E = \{0, 1, 2, \dots, 8\}$. Für π gilt $\pi(0) = 1$. Weiter ist

$$(q(x, y))_{x, y \in E} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

2) Sei $E = \{1, 2, 3\}$,

$$q(x, y) = \begin{pmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix} \text{ und } q^2(x, y) = \begin{pmatrix} \frac{5}{8} & 0 & \frac{3}{8} \\ \frac{1}{2} & \frac{3}{8} & \frac{1}{8} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{pmatrix}.$$

Für jede Startverteilung π mit $\pi(x) > 0$ für $x \in E$ gilt $\lim_{n \rightarrow \infty} \sum_{x \in E} \pi(x) q^n(x, y) = \alpha(y)$ mit $\alpha(1) = \frac{8}{19}$, $\alpha(2) = \frac{6}{19}$, $\alpha(3) = \frac{5}{19}$.

3) Seien Y_1, Y_2, \dots unabhängig, $X_i = \sum_{j=1}^i Y_j$ und $X_0 = x_0$. Dann ist $\{X_0, X_1, X_2, \dots\}$ eine Markov-Kette.

Beweis:

Sei $m \in \mathbb{N}$ und sei $y_i = x_i - x_{i-1}$ für $i \geq 1$ und $y_0 = x_0$. Dann ist:

$$\begin{aligned} & P(X_{m+1} = x | X_0 = x_0, \dots, X_m = x_m) \\ &= P(X_{m+1} = x | X_0 = y_0, X_1 = y_0 + y_1, X_2 = y_0 + y_1 + y_2, \dots, X_m = y_0 + \dots + y_m) \\ &= P(Y_{m+1} = x - (y_0 + y_1 + \dots + y_m) | Y_0 = y_0, Y_1 = y_1, \dots, Y_m = y_m) \\ &= \frac{P(Y_{m+1} = x - (y_0 + y_1 + \dots + y_m)) P(Y_0 = y_0) P(Y_1 = y_1) \dots P(Y_m = y_m)}{P(Y_0 = y_0) P(Y_1 = y_1) \dots P(Y_m = y_m)} \\ &= P(Y_{m+1} = x - (y_0 + y_1 + \dots + y_m)) \\ &= P(X_{m+1} = x | X_m = x_m). \end{aligned}$$

4) Ehrenfest'sches Urnenmodell

N Teilchen befinden sich in den Behältern 1 und 2.

i Teilchen seien in Behälter 1, $N - i$ Teilchen in Behälter 2.

In jeder Zeiteinheit springt ein Teilchen entweder von $1 \rightarrow 2$ oder von $2 \rightarrow 1$.

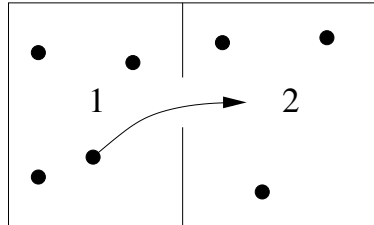


Abbildung 10.2: Ehrenfest'sches Urnenmodell

Die stochastische Matrix wird gegeben durch $q(i, i + 1) = \frac{N-i}{N}$, $q(i, i - 1) = \frac{i}{N}$ und $q(i, j) = 0$ für $j \neq i \pm 1$.

10.3 Absorbierende Zustände

Wir können eine Markov-Kette als Bewegung eines Teilchens durch die Zustände auffassen. In manchen Markov-Ketten gibt es Zustände, die nicht verlassen werden können. Solche Zustände heißen absorbierend.

10.3.1 Definition (absorbierender Zustand, absorbierender Rand)

Sei $\mathcal{X} = \{X_0, X_1, \dots\}$ eine stationäre Markov-Kette mit Zustandsraum E .

Ein Zustand $x \in E$ heißt absorbierend, wenn gilt $q(x, x) = 1$.

Die Menge $A \subset E$ aller absorbierender Zustände heißt absorbierender Rand von \mathcal{X} .

Bei der "Kain und Abel"-Aufgabe ist $\{7, 8\}$ der absorbierende Rand. Der folgende Satz liefert eine Begründung für die Kain und Abel-Aufgabe.

10.3.2 Satz

Sei \mathcal{X} eine stationäre Markov-Kette und $A \subset E$ ihr absorbierender Rand. Wir nehmen an, dass für jedes $x \in E$ die Wahrscheinlichkeit nach A zu kommen gleich 1 ist.

Sei $A = A_1 \cup A_2$ und $A_1 \cap A_2 = \emptyset$. Sei $P_{A_1}(x)$ die Wahrscheinlichkeit bei Start in x beim ersten Eintritt in A nach A_1 zu gelangen. Dann gilt:

(a) $P_{A_1}(x) = 1$ für $x \in A_1$ und $P_{A_1}(x) = 0$ für $x \in A_2$,

(b) Für $x \notin A$ ist $P_{A_1}(x) = \sum_{z \in A_1} q(x, z) + \sum_{n \geq 2} \sum_{\substack{x_i \notin A \text{ für} \\ i=1, \dots, n-1, \\ z \in A_1}} q(x, x_1)q(x_1, x_2) \dots q(x_{n-1}, z)$,

(c) Es gilt stets $P_{A_1}(x) = \sum_{y \in E} q(x, y)P_{A_1}(y)$.

Beweis:

Zu (a): Dies folgt direkt aus der Definition.

Zu (b): Sei $x \notin A$. Dann ist

$$\begin{aligned} P_{A_1}(x) &= P(X_0 = x, X_1 \in A_1) + P\left(\bigcup_{n \geq 2} \{X_0 = x, X_i \notin A \text{ für } i < n, X_n \in A_1\}\right) \\ &= \sum_{z \in A_1} q(x, z) + \sum_{n \geq 2} P(\{X_0 = x, X_i \notin A \text{ für } i < n, X_n \in A_1\}) \\ &= \sum_{z \in A_1} q(x, z) + \sum_{n \geq 2} \sum_{x_1, x_2, \dots, x_{n-1} \notin A, z \in A_1} q(x, x_1) \dots q(x_{n-1}, z). \end{aligned}$$

Zu (c): Sei $x \notin A$. So gilt

$$\begin{aligned} P_{A_1}(x) &= P(X_0 = x, X_1 \in A_1) + P(X_0 = x, X_1 \notin A, X_2 \in A_1) \\ &\quad + P(X_0 = x, \exists n \geq 3 \text{ mit } X_n \in A_1 \text{ und } X_i \notin A \text{ für } i < n) \\ &= \sum_{y \in A_1} q(x, y) + \sum_{y \notin A, z \in A_1} q(x, y)q(y, z) \\ &\quad + \sum_{n \geq 3} \sum_{y \notin A} \sum_{\substack{x_i \notin A, \\ i=2, \dots, n-1, \\ z \in A_1}} q(x, y)q(y, x_2) \dots q(x_{n-1}, z) \\ &\stackrel{(*)}{=} \sum_{y \in A} q(x, y)P_{A_1}(y) + \sum_{y \notin A} q(x, y) \left(\sum_{z \in A_1} q(y, z) \right) \\ &\quad + \sum_{y \notin A} q(x, y) \left(\sum_{n \geq 3} \sum_{\substack{x_i \notin A, \\ i=2, \dots, n-1, \\ z \in A_1}} q(y, x_2) \dots q(x_{n-1}, z) \right) \\ &= \sum_{y \in A} q(x, y)P_{A_1}(y) + \sum_{y \notin A} q(x, y)P_{A_1}(y) \\ &= \sum_y q(x, y)P_{A_1}(y). \end{aligned}$$

(*) $P_{A_1}(x) = 1$ für $x \in A_1$ und $P_{A_1}(x) = 0$ für $x \in A \setminus A_1$.

Bemerkung: Die Gleichung c) aus Satz 10.3.2 lautet in Vektorschreibweise $P_{A_1} = qP_{A_1}$. Dies bedeutet P_{A_1} ist rechter Eigenvektor von q zum Eigenwert 1. Man sagt auch, P_{A_1} ist harmonisch oder P_{A_1} erfüllt die Mittelwerteigenschaft.

10.3.3 Berechnung von Ruin-Wahrscheinlichkeiten

Hans und Rudolf spielen ein Spiel. In jeder Runde gewinnt Hans mit der Wahrscheinlichkeit p und Rudolf gewinnt mit der Wahrscheinlichkeit $q = 1 - p$. Der Gewinner einer Runde erhält von seinem Gegner einen Euro. Es wird so lange gespielt bis einer der Spieler kein Geld mehr hat. Wie hoch ist die Ruin-Wahrscheinlichkeit $P(x)$ von Hans, wenn Hans zu Beginn x Euro hat und Hans und Rudolf zusammen b Euro haben?

Seien $x, b \in \mathbb{N}$ mit $0 < x < b$. Weiter seien X_1, X_2, \dots unabhängige Zufallsvariablen mit $P(X_i = 1) = p = 1 - P(X_i = -1)$ für alle i und es sei $S_0 = x$ und $S_n = x + \sum_{i=1}^n X_i$ für $n \geq 1$. Damit ist S_0, S_1, S_2, \dots eine Markov-Kette und S_n ist das Kapital, das Hans nach n Runden besitzt. Wir frage nach der Ruin-Wahrscheinlichkeit:

$$P(x) = P(\exists n \text{ mit } S_n = 0 \text{ und } 0 < S_i < b \text{ für } i < n \mid S_0 = x).$$

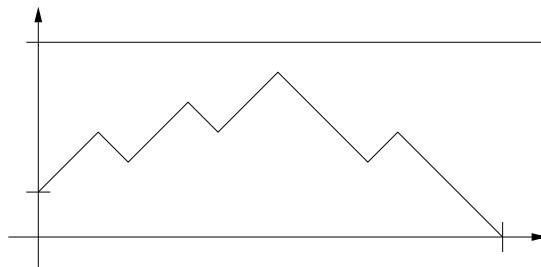


Abbildung 10.3: Ruin-Wahrscheinlichkeiten

10.3.4 Satz

Sei $q = 1 - p$.

Für $p = \frac{1}{2}$ gilt $P(x) = 1 - \frac{x}{b}$ für $0 \leq x \leq b$.

Für $p \neq \frac{1}{2}$ gilt $P(x) = \frac{(\frac{q}{p})^b - (\frac{q}{p})^x}{(\frac{q}{p})^b - 1}$ für $0 \leq x \leq b$.

Beweis:

Fall 1: Sei $p = \frac{1}{2}$.

Satz 10.3.2 liefert: $P(x) = \frac{1}{2}P(x-1) + \frac{1}{2}P(x+1)$ für $0 < x < b$, $P(0) = 1$, $P(b) = 0$.

Durch Lösen dieses linearen Gleichungssystems ergibt sich: $P(x) = 1 - \frac{x}{b}$.

Fall 2: Sei $p \neq \frac{1}{2}$.

Satz 10.3.2 liefert: $P(x) = pP(x+1) + qP(x-1)$ für $0 < x < b$, $P(b) = 0$, $P(0) = 1$.

Dann ist

$$\begin{aligned} P(x) &= pP(x+1) + qP(x-1) \\ pP(x) + qP(x) &= pP(x+1) + qP(x-1) \\ P(x+1) - P(x) &= \frac{q}{p}(P(x) - P(x-1)). \end{aligned}$$

Wiederholtes Anwenden liefert

$$P(x+1) - P(x) = \left(\frac{q}{p}\right)^x (P(1) - P(0)).$$

Sei $r := \frac{q}{p}$.

Aufaddieren liefert $P(x) - P(0) = \frac{r^x - 1}{r - 1} (P(1) - P(0))$.

Wegen $P(0) = 1$ gilt $P(x) = 1 + \frac{r^x - 1}{r - 1} (P(1) - 1)$.

Wegen $P(b) = 0$ gilt $P(1) - 1 = -\frac{r-1}{r^b-1}$.

Einsetzen in die vorangegangene Gleichung liefert die Behauptung.

10.4 Rekurrente und transiente Zustände

10.4.1 Bezeichnungen

Sei $f^n(x, y) := P(X_n = y, X_i \neq y \text{ für } 1 \leq i < n | X_0 = x)$. Dann bezeichnet $f^n(x, y)$ die Wahrscheinlichkeit bei Start in x nach n Schritten erstmals den Zustand y zu erreichen.

$f^*(x, y) := P(X_n = y \text{ für ein } n \geq 1 | X_0 = x) = \sum_{n \geq 1} f^n(x, y)$ ist die Wahrscheinlichkeit bei Start in x irgendwann nach y zu gelangen.

$q^*(x, y) := \sum_{n=0}^{\infty} q^n(x, y)$ ist die erwartete Anzahl der Besuche in y bei Start in x , denn:

$$\begin{aligned} q^*(x, y) &= \sum_{n=0}^{\infty} q^n(x, y) = \sum_{n=0}^{\infty} P(X_n = y | X_0 = x) \\ &= \sum_{n=0}^{\infty} E(1_{\{X_n=y\}} | X_0 = x) = E\left(\sum_{n=0}^{\infty} 1_{\{X_n=y\}} \middle| X_0 = x\right). \end{aligned}$$

10.4.2 Definition (rekurrenter Zustand)

$x \in E$ heißt rekurrent, falls $f^*(x, x) = 1$,

$x \in E$ heißt transient, falls $f^*(x, x) < 1$.

10.4.3 Satz

1. $x \in E$ ist rekurrent $\Leftrightarrow q^*(x, x) = \infty$,
2. $x \in E$ ist transient $\Leftrightarrow q^*(x, x) < \infty$.

Beweis: Dieser folgt aus nachfolgendem Lemma.

10.4.4 Lemma

Es gilt

a) $q^*(x, y) = f^*(x, y) q^*(y, y) + \delta(x, y)$,

b) $f^*(x, x) = \frac{q^*(x, x) - 1}{q^*(x, x)}$,

c) $q^*(x, x) = \frac{1}{1 - f^*(x, x)}$.

Beweis:

Für alle $x, y \in E$ gilt: $q^n(x, y) = \sum_{i=1}^n f^i(x, y) q^{n-i}(y, y)$. Denn es gilt mit

$$T_y = \min\{m \geq 1 | X_m = y\}$$

$$q^n(x, y) = P(X_n = y | X_0 = x)$$

$$\begin{aligned}
&= \sum_{i=1}^n P(T_y = i | X_0 = x) P(X_n = y | T_y = i, X_0 = x) \\
&= \sum_{i=1}^n f^i(x, y) P(X_n = y | X_i = y) \\
&= \sum_{i=1}^n f^i(x, y) P(X_{n-i} = y | X_0 = y) \\
&= \sum_{i=1}^n f^i(x, y) q^{n-i}(y, y).
\end{aligned}$$

Weiter gilt nach Vertauschung der Summation

$$\begin{aligned}
q^*(x, y) &= \sum_{i=1}^{\infty} q^n(x, y) + \delta(x, y) \\
&= \sum_{n=1}^{\infty} \sum_{i=1}^n f^i(x, y) q^{n-i}(y, y) + \delta(x, y) \\
&= \sum_{i=1}^{\infty} \sum_{n=i}^{\infty} f^i(x, y) q^{n-i}(y, y) + \delta(x, y) \\
&= \sum_{i=1}^{\infty} f^i(x, y) \sum_{m=0}^{\infty} q^m(y, y) + \delta(x, y) \\
&= f^*(x, y) \cdot q^*(y, y) + \delta(x, y).
\end{aligned}$$

Damit ist a) gezeigt. Für $x = y$ folgt $q^*(x, x) - f^*(x, x)q^*(x, x) = 1$.

Auflösen nach $q^*(x, x)$ liefert $q^*(x, x)(1 - f^*(x, x)) = 1$ und damit $q^*(x, x) = \frac{1}{1 - f^*(x, x)}$.

Auflösen nach $f^*(x, x)$ liefert: $f^*(x, x) = \frac{q^*(x, x) - 1}{q^*(x, x)}$.

10.4.5 Beispiel auf \mathbb{Z} (Irrfahrt auf \mathbb{Z})

Seien Y_1, Y_2, \dots unabhängige Zufallsvariable mit $P(Y_i = 1) = p$, $P(Y_i = -1) = q$ für alle i . Weiter sei $X_0 = 0$ und $X_n = Y_1 + Y_2 + \dots + Y_n$ für $n \geq 1$.

Behauptung: Für $p = q = \frac{1}{2}$ ist 0 rekurrent, für $p \neq q$ ist 0 transient.

Beweis:

Wende Satz 10.4.3 an! Sei $n \in \mathbb{N}$. Nach $2n + 1$ Schritten liegt ein ungerader Zustand vor, also kann X_{2n+1} nicht 0 sein und damit gilt: $q^{2n+1}(0, 0) = 0$. Fällt in $2n$ Experimenten insgesamt n -mal die 1 und n -mal die -1 , so ist $X_{2n} = 0$ und es gilt: $q^{2n}(0, 0) = \binom{2n}{n} p^n q^n$. Für große n können wir die Stirling'schen Formel anwenden:

$$q^{2n}(0, 0) \sim \frac{1}{\sqrt{\pi n}} (4pq)^n, \text{ da } \binom{2n}{n} \sim \frac{1}{\sqrt{\pi n}} 2^{2n}.$$

Fall 1: $p \neq q \Rightarrow (4pq) < 1 \Rightarrow \sum_{n \geq 1} q^{2n}(0, 0) = \sum_{n \geq 1} \frac{1}{\sqrt{\pi n}} (4pq)^n (1 + o(1)) < \infty \Rightarrow$ „0“ ist transient.

Fall 2: $p = q = \frac{1}{2} \Rightarrow (4pq) = 1$

$$\begin{aligned} \Rightarrow q^{2n}(0,0) &\sim \frac{1}{\sqrt{\pi n}} \\ \Rightarrow \sum_n q^{2n}(0,0) &\sim \sum_n \frac{1}{\sqrt{\pi n}} = \infty \Rightarrow \text{„0“ ist rekurrent.} \end{aligned}$$

10.4.6 Definition (irreduzibel)

Die stochastische Matrix q heißt irreduzibel, falls für alle $x, y \in E$ ein $m \in \mathbb{N}$ existiert mit $q^m(x, y) > 0$.

10.4.7 Satz

Sei q irreduzibel. Dann gilt: Existiert ein $z \in E$, das rekurrent ist, so sind alle $x \in E$ rekurrent.

Beweis:

Sei $x \in E$ beliebig. Dann existieren $k, l \in \mathbb{N}$ mit $q^k(x, z) > 0$ und $q^l(z, x) > 0$.

Dann ist

$$q^n(x, x) \geq q^k(x, z)q^m(z, z)q^l(z, x),$$

falls $n = k + m + l$ ist. Es folgt

$$q^*(x, x) = \sum_{n \geq 0} q^n(x, x) = q^k(x, z) \sum_{m \geq 0} q^m(z, z)q^l(z, x).$$

Da die rechte Seite nach Voraussetzung gleich unendlich ist, ist auch $q^*(x, x) = \infty$ und damit $f^*(x, x) = 1$.

10.5 Stationäre Verteilungen

10.5.1 Definition (stationäre Verteilung)

Eine Wahrscheinlichkeitsfunktion π auf E heißt stationär bezüglich q , falls für alle $y \in E$ gilt:

$$\sum_x \pi(x)q(x, y) = \pi(y).$$

In Vektorschreibweise: $\pi^T q = \pi^T$, d. h. π ist linker Eigenvektor von q mit Eigenwert 1.

10.5.2 Bemerkung

Ist π stationär und Startverteilung von \mathcal{X} , d.h. $P(X_0 = x) = \pi(x)$, so gilt:

$$P(X_n = y) = \pi(y).$$

Beweis: Es gilt

$$P(X_n = y) = \sum_x P(X_n = y \mid X_0 = x)\pi(x)$$

$$\begin{aligned}
&= \sum_x \pi(x) q^n(x, y) \\
&= \sum_x \pi(x) \sum_z q(x, z) q^{n-1}(z, y) \\
&= \sum_z \sum_x \pi(x) q(x, z) q^{n-1}(z, y) \\
&= \sum_z q^{n-1}(z, y) \pi(z) \\
&\quad \vdots \\
&= \pi(y).
\end{aligned}$$

10.5.3 Beispiel (Ehrenfestsches Urnenmodell)

Behauptung:

$\pi(i) = \binom{N}{i} 2^{-N}$ ist stationäre Verteilung des Ehrenfestschen Urnenmodells.

Beweis:

Zu zeigen: $\sum_i \pi(i) q(i, j) = \pi(j)$.

$$\begin{aligned}
&\sum_i \pi(i) q(i, j) \\
&= \sum_i \binom{N}{i} 2^{-N} q(i, j) \\
&= \binom{N}{j-1} 2^{-N} \frac{N-(j-1)}{N} + \binom{N}{j+1} 2^{-N} \frac{j+1}{N} \\
&= \left[\frac{N!}{(j-1)!(N-(j-1))!} \frac{N-(j-1)}{N} + \frac{N!}{(j+1)!(N-(j+1))!} \frac{j+1}{N} \right] 2^{-N} \\
&= \left[\frac{(N-1)!}{(j-1)!(N-j)!} \frac{j}{j} + \frac{(N-1)!}{j!(N-j-1)!} \frac{N-j}{N-j} \right] 2^{-N} \\
&= \frac{(N-1)!j + (N-1)!(N-j)}{j!(N-j)!} 2^{-N} \\
&= \binom{N}{j} 2^{-N} = \pi(j).
\end{aligned}$$

Es gibt noch einen einfacheren Weg die Stationarität von π zu zeigen. Man verwendet dazu das folgende Lemma.

10.5.4 Lemma

Gilt für alle $x, y \in E$ $\pi(x)q(x, y) = \pi(y)q(y, x)$, so ist π stationär.

Beweis:

$$\begin{aligned} \sum_x \pi(x)q(x, y) &= \sum_x \pi(y)q(y, x) \\ &= \pi(y) \sum_x q(y, x) \\ &= \pi(y). \end{aligned}$$

Bemerkung: Für das Ehrenfestsche Urnenmodell lautet die in 10.5.4 vorausgesetzte Gleichung $\binom{N}{i} \frac{N-i}{N} = \binom{N}{i+1} \frac{i+1}{N}$. Es ist offensichtlich, daß diese gilt.

Der folgende Satz liefert die Existenz einer stationären Verteilung.

10.5.5 Satz

Sei q irreduzibel. Dann sind folgende Aussagen äquivalent:

- Es existiert ein $z \in E$ mit $E(T_z | X_0 = z) < \infty$,
- Für alle $x, y \in E$ gilt $E(T_y | X_0 = x) < \infty$,
- Es existiert eine stationäre Wahrscheinlichkeitsfunktion π bezüglich q .

Bemerkung: Bedingung (a) ist stärker als Rekurrenz, denn es gibt Markov-Ketten mit einem rekurrenten Zustand x für den $E(T_x | X_0 = x) = \infty$ gilt. Siehe z.B. die Irrfahrt von oben.

Beweisskizze: Zeige (a) \Rightarrow (c) Setze: $P_z(\cdot) = P(\cdot | X_0 = z)$. Dann ist P_z das Wahrscheinlichkeitsmaß bei Start in z .

Sei $H(x) := \sum_{n=1}^{\infty} P_z(X_n = x, T_z \geq n)$. Dann gilt $H(z) = 1$ und weiter

$$\begin{aligned} \sum_{x \in E} H(x) &= \sum_{n=1}^{\infty} \sum_{x \in E} P_z(X_n = x, T_z \geq n) \\ &= \sum_{n=1}^{\infty} P_z(T_z \geq n) = E_z T_z. \end{aligned}$$

Wir behaupten (+): $H(y) = \sum_{x \in E} H(x)q(x, y)$.

Setzt man nun $\pi(x) = \frac{H(x)}{E_z T_z}$, so ist $\pi(y) = \sum_x \pi(x)q(x, y)$ und $\sum_y \pi(y) = 1$, d.h. π ist stationäre Verteilung.

Zeige Behauptung (+):

$$\begin{aligned}
H(y) &= \sum_{n=1}^{\infty} P_z(X_n = y, T_z \geq n) \\
&= P_z(X_1 = y) + \sum_{n \geq 2} P_z(X_n = y, X_j \neq z \text{ für } 1 \leq j \leq n-1) \\
&= q(z, y) + \sum_{n \geq 2} \sum_{x \in E \setminus \{z\}} P_z(X_{n-1} = x, X_n = y, X_j \neq z \text{ für } 1 \leq j \leq n-1) \\
&= q(z, y) + \sum_{x \in E \setminus \{z\}} \sum_{n \geq 2} P_z(X_{n-1} = x, X_n = y, X_j \neq z \text{ für } 1 \leq j \leq n-1) \\
&= q(z, y) + \sum_{x \in E \setminus \{z\}} \sum_{n \geq 2} P_z(X_{n-1} = x, X_j \neq z \text{ für } 1 \leq j \leq n-1) \cdot P(X_n = y | X_{n-1} = x) \\
&= q(z, y) + \sum_{x \in E \setminus \{z\}} \sum_{l \geq 1} P_z(X_l = x, X_j \neq z \text{ für } 1 \leq j \leq l) \cdot P(X_n = y | X_{n-1} = x) \\
&= q(z, y) + \sum_{x \in E \setminus \{z\}} H(x)q(x, y) \\
&= \sum_x H(x)q(x, y).
\end{aligned}$$

10.5.6 Korollar

Es gilt

$$\pi(x) = \frac{H(x)}{E_z T_z}$$

und insbesondere

$$\pi(z) = \frac{1}{E_z T_z},$$

wobei

$$H(x) = \sum_{n=1}^{\infty} P_z(X_n = x, T_z \geq n) \text{ ist für } z \in E.$$

Bemerkung: Beim Ehrenfestschen Urnenmodell sei $\sigma(i) = \min\{n > 0 \mid X_n = i\}$. Dann ist nach dem Korollar: $E_0 \sigma(0) = 2^{2k}$ und $E_k \sigma(k) = \binom{2k}{k}^{-1} 2^{2k}$, wobei $k = N/2$ ist. Setzt man $N = 6 \cdot 10^{23}$ und ist ein Zeitschritt 1 Sekunde lang, so ist $E_k \sigma(k) \cong 10^{12}$ Sekunden, aber $E_0 \sigma(0) \cong 10^{1,8 \cdot 10^{23}}$ Sekunden.

10.6 Konvergenz gegen die stationäre Verteilung

10.6.1 Definition (Periode, aperiodisch)

Sei q eine irreduzible stochastische Matrix. Weiter sei $N(x, y) := \{n \in \mathbb{N} \mid q^n(x, y) > 0\}$. Wir definieren die Periode von $x \in E$ durch $d(x) := \text{ggT}(N(x, x))$, wobei $\text{ggT}(B) := \sup\{k \in \mathbb{N} \mid B \subset k \cdot \mathbb{N}\}$ der größte gemeinsame Teiler der Menge B ist. q heißt aperiodisch, falls $d(x) = 1$ ist für alle $x \in E$. q hat Periode k , falls $d(x) = k$ ist für alle $x \in E$.

10.6.2 Beispiel

Für die stochastische Matrix beim Ehrenfest'schen Urnenmodell gilt $d = 2$.

10.6.3 Satz

Sei q irreduzibel und aperiodisch mit stationärer Verteilung π . Dann gilt für alle $x \in E$

$$\lim_{n \rightarrow \infty} \sum_{y \in E} |q^n(x, y) - \pi(x)| = 0.$$

Sei q irreduzibel mit Periode k und mit stationärer Verteilung π , so gilt

$$\lim_{n \rightarrow \infty} \sum_{y \in E} |q^{nk}(x, y) - \pi(x)| = 0.$$

Beweis: Zum Beweis siehe Dümbgen, S. 129.

10.6.4 Satz

Sei q irreduzibel und aperiodisch mit stationärer Verteilung π . Sei $f : E \rightarrow \mathbb{R}$, so daß $\sum_{x \in E} |f(x)|\pi(x) < \infty$ ist. Dann gilt für jede Startverteilung μ mit P_μ -Wahrscheinlichkeit 1:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N f(x_k) = \sum_{x \in E} f(x)\pi(x).$$

Bemerkung: Dies ist ein grundlegender Satz für Markov-Ketten, der dem Gesetz der Großen Zahlen für unabhängige Beobachtungen entspricht. Es ist ein Basisresultat für die Simulation bei Markov-Ketten, der sogenannten MCMC-Methode.

10.6.5 Bemerkungen zum Ehrenfest'schen Urnenmodell

Dieses Modell wurde entwickelt, um eine Streitfrage in der statistischen Physik zu klären. Boltzmann behauptete, ein großes Teilchensystem tendiere sehr schnell zu seinem Gleichgewicht, Zermelo entgegnete, auch dann müßten aus physikalischen Gründen sehr unwahrscheinliche Zustände angenommen werden können. Dies erschien paradox. Das Ehrenfest'sche Modell gestattet die Details zu berechnen.

Sei $N = 2k$. Sei $\tau(k) = \inf\{n \geq 0 \mid X_n = k\}$. Dann gilt:

a) $E_0(\tau(k)) = k \ln k + k + O(1)$

b) $E_k(\tau(0)) = \frac{1}{2k} 2^{2k} (1 + o(\frac{1}{k}))$ für $k \rightarrow \infty$.