# Endliche Markov-Ketten

Michael Krühler

24. Oktober 2013

## Inhaltsverzeichnis

# 1 Einführung

Dieser Vortrag stellt einen speziellen stochastischen Prozess, genannt Markov-Kette, vor. Dabei beschränken wir uns auf einen endlichen Zustandsraum. Grundlegende Eigenschaften einer solchen endlichen Markov-Kette sollen eingeführt und anhand von einigen Beispielen erläutert werden. Zum Ende hin beschäftigen wir uns zudem mit einigen Konvergenzaussagen für spezielle endliche Markov-Ketten.

## 1.1 Mathematische Einführung

Gegeben sei eine stochastische Folge $(M_n)_{n\geq 0}$ auf einem W-Raum $(\Omega, \mathscr{A}, \mathbb{P})$, die Werte in einem beliebigen endlichen messbaren Raum $(\mathscr{S}, \mathscr{G})$ annehme. Diesen Raum nennen wir im Folgenden **Zustandsraum** und seine Elemente $s \in \mathscr{S}$ **Zustände**.

**Definition 1 (*Markov-Kette*)**

*$(M_n)_{n\geq 0}$ wie oben heißt **endliche Markov-Kette**, wenn sie die folgenden zwei Eigenschaften besitzt:*

1. ***Markov-Eigenschaft (ME):** Für alle $n \in \mathbb{N}_0$ gilt*

$$\mathbb{P}(M_{n+1} \in A | M_0, \ldots, M_n) = \mathbb{P}(M_{n+1} \in A | M_n) \qquad \forall A \in \mathscr{G}$$

2. ***Zeithomogenität (ZH):** Für alle $n \in \mathbb{N}_0$ gilt:*

$$\mathbb{P}(M_{n+1} = j | M_n = i) = p_{ij} \qquad \forall i, j \in \mathscr{S}$$

*mit $0 \leq p_{ij} \leq 1$.*

*Die Zahlen $p_{ij}$ heißen **(Ein-Schritt-)Übergangswahrscheinlichkeiten**.*

Mit anderen Worten ausgedrückt, erfüllt eine stochastische Folge $M = (M_n)_{n\geq 0}$ die *Markov-Eigenschaft*, wenn die bedingte Verteilung des $(n+1)$-ten Folgegliedes, gegeben die Vergangenheit der Markov-Kette $M_0, \ldots, M_n$, lediglich vom Zustand des $n$-ten Folgegliedes abhängt. *Zeitlich homogen* ist diese Folge, wenn die Wahrscheinlichkeit für einen Übergang von einem Zustand $i \in \mathscr{S}$ in einen Zustand $j \in \mathscr{S}$ nicht davon abhängt, zu welchem Zeitpunkt $n$ dieser Übergang stattfindet.

Eindeutig festgelegt ist eine endliche Markov-Kette bereits durch die Angabe einer Anfangsverteilung $\mathbb{P}^{M_0}$ für den Startzustand $M_0$ und der Übergangswahrscheinlichkeiten $p_{ij}$, denn:

$$\mathbb{P}^M \overset{\text{WT}}{=} \mathbb{P}^{M_0} \otimes \bigotimes_{n\geq 1} \mathbb{P}^{M_{n+1}|M_0,\ldots M_n} \overset{\text{ME}}{=} \mathbb{P}^{M_0} \otimes \bigotimes_{n\geq 1} \mathbb{P}^{M_{n+1}|M_n}$$

## 1.2 Interpretation als 'Brettspiel mit Marker'

Die nun mathematisch eingeführten Markov-Ketten können in Form eines Brettspieles interpretiert werden und auf sehr anschauliche Weise durch ein Übergangsdiagramm dargestellt werden. Dabei symbolisieren Knotenpunkte die Zustände des Zustandsraumes $(\mathscr{S},\mathscr{G})$, häufig beschriftet mit $0,1,2,\ldots$ Alle positiven Ein-Schritt-Übergangswahrscheinlichkeiten $p_{ij} > 0$ werden durch einen Pfeil von Zustand $i \in \mathscr{S}$ nach Zustand $j \in \mathscr{S}$ gekennzeichnet und zumeist mit der Zahl $p_{ij}$ beschriftet ($i = j$ nicht ausgeschlossen). Ein Marker wandert Schritt für Schritt entlang der Pfeile über das Spielbrett. Ist er nach n Schritten auf dem Knoten k, entspricht dies $M_n = k \in \mathscr{S}$.

Es folgen nun einige einfache Beispiele für Markov-Ketten:

**Beispiel 1. 1** ***(Das Warte-Spiel)***
Der Marker startet bei Knoten 0, d.h. $M_0 = 0$. Nun wird eine Münze mit Wahrscheinlichkeit $0 < p < 1$ für Kopf und $q = 1 - p$ für Zahl wiederholt und unabhängig voneinander geworfen. Sobald die Münze Kopf zeigt, wechselt der Marker zu Knoten 1 und das Spiel ist beendet (vgl. Abb. 1). In diesem Beispiel gilt:

$$\mathscr{S} = \{0,1\}, \quad \mathscr{G} = \mathscr{P}(\mathscr{S}) = \{\emptyset,\{0\},\{1\},\{0,1\}\},$$
$$p_{00} = q, \quad p_{01} = p, \quad p_{10} = 0, \quad p_{11} = 1$$

Die Unabhängigkeit der Münzwürfe garantiert die Gültigkeit der Markov-Eigenschaft, während die Zeithomogenität offensichtlich gegeben ist, weil immer dieselbe Münze geworfen wird, p und q also konstant bleiben.

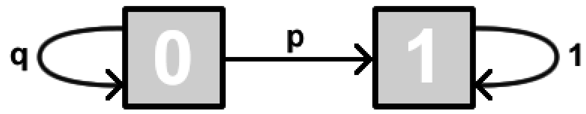


Abbildung 1: Übergangsdiagramm des Warte-Spiels

**Beispiel 2. 1** ***(Getränkeautomat)***
Ein Getränkeautomat wird täglich einmal überprüft und entweder für intakt (Zustand 0) oder für defekt (Zustand 1) befunden. Der Zeitindex $n$ gibt also die Tage ab Aufstellung des Gerätes an.

Ist die Maschine ordnungsgemäß in Betrieb, ist sie am nächsten Tag mit Wahrscheinlichkeit $0 < \delta < 1$ defekt. Im Falle eines technisches Fehlers findet die Reparatur mit Wahrscheinlichkeit $0 < \gamma < 1$ bis zum nächsten Tag statt (vgl. Abb. 2). Es ist also:

$$\mathscr{S} = \{0,1\}, \quad \mathscr{G} = \mathscr{P}(\mathscr{S}) = \{\emptyset,\{0\},\{1\},\{0,1\}\},$$
$$p_{00} = 1 - \delta, \quad p_{01} = \delta, \quad p_{10} = \gamma, \quad p_{11} = 1 - \gamma$$

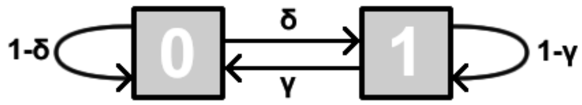


Abbildung 2: Getränkeautomat

**Beispiel 3. 1** ***(Gambler's Ruin)***
Der Spieler einer Wette startet mit einem Guthaben von $0 < k < 5$ € und setzt in jeder Spielrunde 1€ ein. Ein Münzwurf mit Wahrscheinlichkeit $0 < p < 1$ für Kopf entscheidet darüber, ob der Spieler seinen Einsatz verliert (Zahl) oder verdoppelt (Kopf). Das Spiel endet, sobald der Spieler sein im Vorfeld gestecktes Ziel von 5€ erreicht hat oder wenn er kein Geld mehr zur Verfügung hat, also mit dem "*Gambler's Ruin*" (vgl. Abb. 3). Es gilt:

$$\begin{aligned}
&\mathscr{S} = \{0,1,2,3,4,5\}, \quad \mathscr{G} = \mathscr{P}(\mathscr{S}) \\
&p_{12} = p_{23} = p_{34} = p_{45} = p \\
&p_{43} = p_{32} = p_{21} = p_{10} = 1 - p = q \\
&p_{00} = p_{55} = 1
\end{aligned}$$

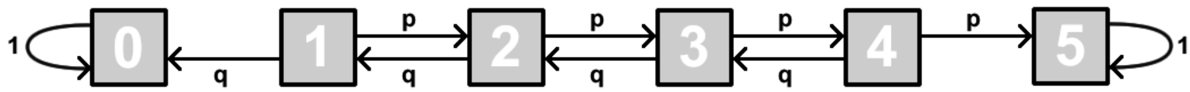


Abbildung 3: Gambler's Ruin

**Beispiel 4. 1** ***(Alles oder nichts)***
Genau wie im vorherigen Beispiel startet ein Spieler mit einem Guthaben von $0 < k < 5$ €. Die Wette ist beendet, wenn der Spieler entweder sein Ziel von 5€ erreicht hat oder wenn er kein Geld mehr zur Verfügung hat. Der Wurf einer $p$-Münze entscheidet in jeder Runde, ob der Spieler seinen Einsatz verliert (Zahl) oder verdoppelt (Kopf, Wahrscheinlichkeit $0 < p < 1$).

Dabei setzt der Spieler in jeder Runde sein ganzes Guthaben ein, mit Ausnahme solcher Spielrunden, in denen die Verdoppelung eines Anteils $m$ von seinem Guthaben bereits den Sieg bedeuten würde. In diesem Fall setzt er $m$€ (vgl. Abb. 4). Es gilt:

$$\begin{aligned}
&\mathscr{S} = \{0,1,2,3,4,5\}, \quad \mathscr{G} = \mathscr{P}(\mathscr{S}) \\
&p_{12} = p_{24} = p_{45} = p_{35} = p \\
&p_{10} = p_{20} = p_{31} = p_{43} = 1 - p = q \\
&p_{00} = p_{55} = 1
\end{aligned}$$

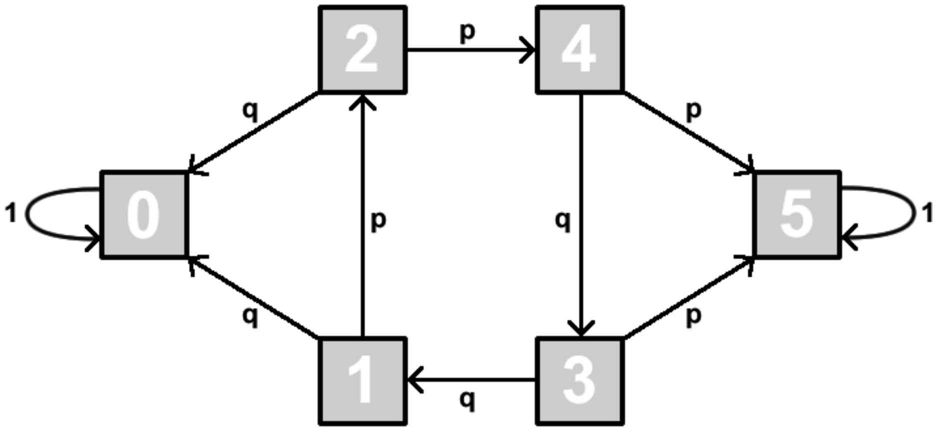


Abbildung 4: Alles oder nichts

## 2 Übergangswahrscheinlichkeiten

Um den Axiomen der Wahrscheinlichkeitsrechnung gerecht zu werden, muss für die Ein-Schritt-Übergangswahrscheinlichkeiten stets

$$0 \le p_{ij} \le 1 \quad \forall i, j \in \mathscr{S} \qquad \text{und} \quad \sum_{j \in \mathscr{S}} p_{ij} = 1 \quad \forall i \in \mathscr{S}$$

gelten. Dies bedeutet, dass sich alle Wahrscheinlichkeiten für die Pfade, die von einem Knoten $i$ wegführen (inklusive des Pfades von Konten $i$ zu sich selbst), zu 1 aufaddieren.

Die Antwort auf die Frage, wie die Wahrscheinlichkeiten eines Pfades, der über eine Knotenfolge $i_0, i_1, \ldots, i_k \in \mathscr{S}$ verläuft, berechnet werden kann, liefert der folgende

**Satz 2 (*Multiplikationsregel für Pfad-Wahrscheinlichkeiten*)**

*Gegeben ein Übergangsdiagramm mit Ein-Schritt-Übergangswahrscheinlichkeiten $p_{ij}$ und eine Anfangsverteilung $\pi$ für $M_0$, so gilt*

$$\mathbb{P}(M_0 = i_0, M_1 = i_1, \ldots, M_k = i_k) = \pi_{i_0} p_{i_0,i_1} p_{i_1,i_2} \cdots p_{i_{k-1},i_k},$$

*mit* $i_0, i_1, \ldots, i_k \in \mathscr{S}$

**Beweis :** Gegeben eine Folge von Zuständen $i_0, i_1, \ldots$, definieren wir das Ereignis

$$A_k := \{M_k = i_k\}$$

(zum Zeitpunkt $k$ zeigt der Marker auf Knoten $i_k$ im Übergangsdiagramm) und das Ereignis

$$B_k := \{M_0 = i_0, M_1 = i_1, \ldots, M_k = i_k\}$$

(der Marker durchläuft den Pfad $i_0 \to i_1 \to \ldots \to i_k$)
Mit Hilfe der Markov-Eigenschaft gilt:

$$\mathbb{P}(A_{n+1}|B_k) \overset{\text{(ME)}}{=} \mathbb{P}(A_{n+1}|A_n) = p_{i_n,i_{n+1}} \quad \forall n \ge 0 \tag{1}$$

Des Weiteren stellen wir fest, dass $A_k \cap B_{k-1} = B_k$ gilt und folgern aus der definierenden Gleichung für bedingte Wahrscheinlichkeiten $\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B)$, dass

$$\mathbb{P}(B_k) = \mathbb{P}(A_k \cap B_{k-1}) = \mathbb{P}(A_k|B_{k-1})\mathbb{P}(B_{k-1})$$

Diesen Vorgang iterieren wir bis $B_0$ und erhalten

$$\begin{aligned}\mathbb{P}(B_k) =& \mathbb{P}(A_k|B_{k-1})\mathbb{P}(A_{k-1}|B_{k-2}) \cdots \mathbb{P}(A_1|B_0)\mathbb{P}(B_0) \\ \underset{(1)}{\overset{B_0=A_0}{=}}& p_{i_{k-1},i_k} p_{i_{k-2},i_{k-1}} \cdots p_{i_0,i_1} \pi_{i_0}\end{aligned}$$

■

**Definition 3 (*n-Schritt-Übergangswahrscheinlichkeiten*)**

*Gegeben eine Markov Kette $(M_n)_{n\geq 0}$ mit Ein-Schritt-Übergangswahrscheinlichkeiten $p_{ij}$, so definiere für alle $i, j \in \mathscr{S}$ die* ***n-Schritt-Übergangswahrscheinlichkeiten*** *als*

$$p_{ij}^{(n)} := \mathbb{P}(M_n = j | M_0 = i), \quad n \geq 0$$

**Bemerkung**

1. Insbesondere gelten:

$$p_{ij}^{(1)} = p_{ij}, \quad p_{ii}^{(0)} = 1 \quad \text{und} \quad p_{ij}^{(0)} = 0 \text{ für alle } i \neq j$$

2. Durch die Zeithomogenität erhalten wir ebenfalls

$$p_{ij}^{(n)} = \mathbb{P}(M_{k+n} = j | M_k = i) \quad \forall k \geq 0$$

**Definition 4 (*Übergangsmatrizen*)**

*Definiere die* ***Ein-Schritt-Übergangsmatrix*** *$P$ und für alle $n \in \mathbb{N}_0$ die* ***n-Schritt-Übergangsmatrix*** *$P^{(n)}$ wie folgt:*

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots \\ p_{21} & p_{22} & \\ \vdots & & \ddots \end{pmatrix} \qquad P^{(n)} = \begin{pmatrix} p_{11}^{(n)} & p_{12}^{(n)} & \cdots \\ p_{21}^{(n)} & p_{22}^{(n)} & \\ \vdots & & \ddots \end{pmatrix}$$

**Bemerkung**

1. Es kann vorkommen, dass die Indizes der in einer Übergangsmatrix eingetragenen Übergangswahrscheinlichkeiten nicht mit der Position in der Matrix übereinstimmen (beispielsweise wenn die Zustände mit $0, 1, 2, \ldots$ gekennzeichnet sind).

2. Die $i$-te Zeile von $P^{(n)}$ ist die bedingte Verteilung von $M_n$, gegeben $M_0 = i$. Daher gilt f.a. $n \in \mathbb{N}$:

$$0 \leq p_{ij}^{(n)} \leq 1 \quad \forall i, j \in \mathscr{S} \qquad \text{und} \qquad \sum_{j \in \mathscr{S}} p_{ij}^{(n)} = 1 \quad \forall i \in \mathscr{S}$$

Jede Matrix mit diesen beiden Eigenschaften wird **stochastische Matrix** genannt.

**Satz 5 (*Chapman-Kolmogorov-Gleichung*)**

*Für alle Zustände* $i, j \in \mathscr{S}$ *und alle* $n \in \mathbb{N}_0, m \in \mathbb{N}$ *gilt:*

$$p_{ij}^{(n+m)} = \sum_k p_{ik}^{(n)} p_{kj}^{(m)}$$

**Beweis :** Konditioniere mit $M_n = k$ über den Satz von der totalen Wahrscheinlichkeit:

$$\begin{aligned}
p_{ij}^{(n+m)} &\overset{\text{Def}}{=} \mathbb{P}(M_{n+m} = j | M_0 = i) \\
&\underset{\text{tot. W'keit}}{\overset{\text{Satz v. d.}}{=}} \sum_{k \in \mathscr{S}} \mathbb{P}(M_{n+m} = j | M_n = k, M_0 = i) \mathbb{P}(M_n = k | M_0 = i) \\
&\overset{\text{(ME)}}{=} \sum_{k \in \mathscr{S}} \mathbb{P}(M_{n+m} = j | M_n = k) \mathbb{P}(M_n = k | M_0 = i) \\
&\overset{\text{(ZH)}}{=} \sum_{k \in \mathscr{S}} \mathbb{P}(M_m = j | M_0 = k) \mathbb{P}(M_n = k | M_0 = i) \\
&\overset{\text{Def}}{=} \sum_{k \in \mathscr{S}} p_{kj}^{(m)} p_{ik}^{(n)}
\end{aligned}$$

■

Die Berechnung der n-Schritt-Übergangswahrscheinlichkeiten $p_{ij}^{(n)}$ einer Markov-Kette $(M_n)_{n \geq 0}$ kann auf zwei verschiedene Weisen verwirklicht werden:

1. Finde alle Pfade der Länge $n$, die bei Knoten $i$ starten und bei Knoten $j$ enden. Mit der Länge ist ausdrücklich die Anzahl der Zeitschritte gemeint; ein Schritt $k \to k$ von Knoten $k$ zu sich selbst wird also berücksichtigt. Fasse all diese Pfade in $T_{ij}^{(n)}$ zusammen. Somit gilt:

$$p_{ij}^{(n)} = \sum_{\gamma \in T_{ij}^{(n)}} \mathbb{P}(\gamma)$$

2. Anhand von Matrizenrechnung mit dem folgenden

**Satz 6**

*Für alle* $n \in \mathbb{N}$ *gilt:*

$$P^{(n)} = P^n$$

**Beweis :** über Indukion nach $n$

Ind.anfang: $n = 1$: Nach Definition klar

IA: Behauptung gelte für festes $n \in \mathbb{N}$

IS: $n \to n + 1$:

$$p_{ij}^{(n+1)} \overset{\text{Satz 5}}{=} \sum_k p_{ik}^{(n)} p_{kj}^{(1)} \overset{\text{IA}}{=} \sum_k (P^n)_{ik} (P)_{kj} = (P^{n+1})_{ij}$$

■

Beide Rechenwege werden wir nun anhand des nächsten Beispieles näher studieren:

**Beispiel 1.2** ***(Das Warte-Spiel)***
Der Marker startet bei Knoten 0. Wie groß ist die Wahrscheinlichkeit, dass er nach $k$ Spielrunden bei Knoten 1 ist?

1. Im Fall $k=2$ sind die gesuchten Pfade

$$\gamma_1 : 0 \to 1 \to 1 \quad \text{und} \quad \gamma_2 : 0 \to 0 \to 1$$

mit Pfadwahrscheinlichkeiten $\mathbb{P}(\gamma_1) \overset{\text{Satz 2}}{=} p$ und $\mathbb{P}(\gamma_2) \overset{\text{Satz 2}}{=} p + qp$. Es gilt also

$$\mathbb{P}(M_2 = 1 | M_0 = 0) = p + qp$$

Allgemein betrachtet stellt man fest, dass jeder gesuchte Pfad genau ein Mal den Wechsel $0 \to 1$ beinhaltet. Findet dieser im $m$-ten Schritt statt, hat der Pfad die Gestalt

$$\gamma_m : \underbrace{0 \to 0 \to \ldots \to 0}_{m} \to \underbrace{1 \to 1 \to \ldots \to 1}_{k+1-m} \tag{2}$$

und die Wahrscheinlichkeit $\mathbb{P}(\gamma_m) = q^{m-1}p$
Da alle gesuchten Pfade genau die Gestalt (2) mit $1 \leq m \leq k$ haben, gilt:

$$\begin{aligned} \mathbb{P}(M_k = 1 | M_0 = 0) &= \sum_{m=1}^{k} q^{m-1} p \underset{\text{shift}}{\overset{\text{Index-}}{=}} p \sum_{m=0}^{k-1} q^m \\ &\underset{\text{Reihe}}{\overset{\text{geom.}}{=}} p \frac{q^k - 1}{q-1} = 1 - q \frac{q^k - 1}{q-1} = -(q^k - 1) = 1 - q^k \end{aligned}$$

2. Die Ein-Schritt-Übergangsmatrix ist in diesem Fall eine obere Dreiecksmatrix:

$$P = \begin{pmatrix} q & p \\ 0 & 1 \end{pmatrix}$$

Für die Matrixmultiplikation zweier oberer $N \times N$-Dreiecksmatrizen $PQ = R$ gilt

- $R$ ist wieder eine obere Dreiecksmatrix der Größe $N \times N$
- $R_{ii} = P_{ii} Q_{ii} \quad \forall 1 \leq i \leq N$

Da sich in $P^{(n)}$ die Einträge einer Zeile zu 1 aufaddieren müssen, gilt

$$P^{(k)} \overset{\text{S 6}}{=} P^k = \begin{pmatrix} q^k & 1 - q^k \\ 0 & 1 \end{pmatrix}$$
$$\Rightarrow \quad \mathbb{P}(M_k = 1 | M_0 = 0) = p_{01}^{(k)} = P_{01}^k = 1 - q^k$$

# 3 Eigenschaften und Klassifikation von Zuständen

In diesem Abschnitt widmen wir uns den Zuständen eine Markov-Kette und suchen nach charakterisierenden Eigenschaften wie Kommunikation mit anderen Zuständen, Absorption und Transienz. Anhand dessen können wir die Zustände klassifizieren und vergrößern so unser Verständnis über die Prozesse in einer Markov-Kette.

**Definition 7 (*Kommunikation*)**

*Sei $(M_n)_{n\geq 0}$ eine endliche Markov-Kette mit Zustandsraum $(\mathscr{S},\mathscr{G})$. Definiere für zwei Zustände $i,j \in \mathscr{S}$*

1. *$j$ heißt **erreichbar** von $i$, kurz $i \to j$, falls ein Pfad von $i$ nach $j$ existiert, wenn also:*
$$\exists k \in \mathbb{N}: \quad p_{ij}^{(k)} > 0$$
2. *$i$ und $j$ heißen **kommunizierend** oder **verbunden**, kurz $i \leftrightarrow j$, falls $i \to j$ und $j \to i$.*

**Bemerkung** Es gilt stets $i \to i$ für alle $i \in \mathscr{S}$ durch einen Pfad der Länge 0, also gilt auch $i \leftrightarrow i$ für alle $i \in \mathscr{S}$.

**Lemma 8**

*Die Beziehung $\leftrightarrow$ ist eine Äquivalenzrelation.*

**Beweis :**
1. Nach obiger Bemerkung ist die Beziehung $\leftrightarrow$ reflexiv
2. $\leftrightarrow$ ist nach Definition symmetrisch
3. Seien $i,j,k \in \mathscr{S}$ mit $i \leftrightarrow j$ und $j \leftrightarrow k$
$$\Rightarrow \exists n,m \in \mathbb{N}: \quad p_{ij}^{(m)} > 0, \quad p_{jk}^{(n)} > 0$$
$$p_{ik}^{(m+n)} \overset{\text{S5}}{=} \sum_l p_{il}^{(m)} p_{lk}^{(n)} \geq p_{ij}^{(m)} p_{jk}^{(n)} > 0 \quad \Rightarrow \quad i \to k$$
Analog zeigt man nun $k \to i$, sodass $i \leftrightarrow j$ gilt. $\quad \Rightarrow \quad \leftrightarrow$ ist transitiv ■

Da $\leftrightarrow$ eine Äquivalenzrelation ist, erhalten wir über folgende Definition eine Partition der Menge aller Zustände $\mathscr{S}$:

**Definition 9 (*Klasse*)**

*Die **Klasse** eines Zustandes $i \in \mathscr{S}$ ist die Menge aller Zustände $j \in \mathscr{S}$, die mit $i$ verbunden sind.*
$$C(i) := \{j \in \mathscr{S} | i \leftrightarrow j\}$$

Des Weiteren definieren wir:

**Definition 10 (*irreduzibel, absorbierend, abgeschlossen, transient*)**

1. *Eine Markov-Kette heißt* ***irreduzibel****, wenn alle Zustände einer einzigen Klasse angehören, d.h. $\mathscr{S} = C(i)$ für ein bel. $i \in \mathscr{S}$.*

2. *Ein Zustand $i \in \mathscr{S}$ heißt* ***absorbierend****, falls $p_{ii} = 1$.*

3. *Eine Klasse C heißt* ***abgeschlossen****, wenn jeder in C startende Pfad in C verbleibt, wenn also*

$$\mathbb{P}(M_n \in C \quad \forall n \geq 1 | M_0 = i) = 1 \quad \forall i \in C$$

*bzw.*

$$\sum_{j \in C} p_{ij} = 1 \quad \forall i \in C$$

4. *Eine Klasse C heißt* ***transient****, wenn es einen Pfad gibt, der aus C heraus führt:*

$$\sum_{j \in C} p_{ij} < 1 \quad \text{für ein } i \in C$$

*Zustände in transienten Klassen heißen ebenfalls* ***transient****.*

Übertragen wir auch diese Begriffe auf die Beispiele:

**Beispiel 1.3** ***(Das Warte-Spiel)***
Werfen wir zunächst einen Blick auf die Ein-Schritt-Übergangsmatrix $P$. Aus dieser lassen sich viele Eigenschaften direkt ablesen.

$$P = \begin{pmatrix} q & 1-q \\ 0 & 1 \end{pmatrix}$$

Es gilt $C(0) = \{0\}$ und $C(1) = \{1\}$, da zwar 1 von 0 erreichbar ist, 0 aber nicht von 1 ($p_{10} = 0$). Diese Markov-Kette ist also nicht irreduzibel.

Die Klasse $C(1)$ ist abgeschlossen. Denn der einzige Zustand dieser Klasse ist absorbierend ($p_{11} = 1$), sodass kein Pfad aus dieser Klasse hinaus führt. Im Gegensatz dazu ist $C(0)$ transient, weil $p_{01} > 0$ gilt.

**Beispiel 2.2** ***(Getränkeautomat)***
Auch hier notieren wir zunächst die Übergangsmatrix $P$:

$$P = \begin{pmatrix} 1-\delta & \delta \\ \gamma & 1-\gamma \end{pmatrix}$$

Offensichtlich gilt $0 \leftrightarrow 1 \quad \Rightarrow \quad C(0) = \{0,1\} = \mathscr{S} \quad \Rightarrow$ Die Kette ist irreduzibel
Absorbierende Zustände gibt es nicht.

Eine weitere spannende Eigenschaft eines Zustandes ist die Periodizität. Verlässt man einen Knoten und gelangt über einen Pfad wieder zu ihm, ist dann die Pfadlänge immer ein Vielfaches einer speziellen Zahl $d > 0$? Wir definieren:

**Definition 11 (*periodisch, aperiodisch*)**

*Ein Zustand $i \in \mathscr{S}$ ist **periodisch** mit **Periode**$d$ (kurz $d$-periodisch), wenn*

$$d = ggT\{n \geq 1 : p_{ii}^{(n)} > 0\}$$

*Ein Zustand mit Periode 1 heißt auch **aperiodisch**. Eine irreduzible Markov-Kette heißt **aperiodisch**, wenn jeder Zustand aperiodisch ist.*

**Lemma 12**

*Ist $i \in \mathscr{S}$ $d$-periodisch und gilt $i \leftrightarrow j$, so folgt dass $j$ ebenfalls $d$-periodisch ist.*

**Beweis :** Da $i \leftrightarrow j$ gilt, existieren folgende Pfade

$$\alpha\colon i \to \ldots \to j$$
$$\beta\colon j \to \ldots \to i$$

Ferner sei $\gamma$ ein beliebiger Pfad von $j$ nach $j$. Setze:

$$\alpha \cdot \beta \colon \underbrace{i \to \ldots \to}_{\text{Pfad } \alpha} j \underbrace{\to \ldots \to i}_{\text{Pfad } \beta}$$

$$\alpha \cdot \gamma \cdot \beta \colon \underbrace{i \to \ldots \to}_{\text{Pfad } \alpha} j \underbrace{\to \ldots \to}_{\text{Pfad } \gamma} j \underbrace{\to \ldots \to i}_{\text{Pfad } \beta}$$

$$\begin{aligned} \text{Länge}(\alpha \cdot \gamma \cdot \beta) &= \text{Länge}(\alpha) + \text{Länge}(\gamma) + \text{Länge}(\beta) \\ &= \text{Länge}(\alpha \cdot \beta) + \text{Länge}(\gamma) \end{aligned}$$

$$i \text{ periodisch mit Periode } d \Rightarrow d \text{ teilt Länge}(\alpha \cdot \gamma \cdot \beta) \text{ und } d \text{ teilt Länge}(\alpha \cdot \beta)$$

$$\Rightarrow d \text{ teilt Länge}(\gamma) \overset{\gamma \text{ bel.}}{\Rightarrow} j \text{ periodisch mit Periode } d$$

■

**Beispiel 4.2 *(Alles oder nichts)***
Die Zustände dieser Markov-Kette gliedern sich in zwei abgeschlossen Klassen, die je einen absorbierenden Zustand beinhalten,

$$C(0) = \{0\} \quad \text{und} \quad C(5) = \{5\}$$

und eine transiente Klasse

$$C(1) = \{1,2,3,4\}$$

Verlässt man den Zustand 1 und kehrt wieder in diesen zurück, kann das nur über den Pfad

$$1 \to 2 \to 4 \to 3 \to 1$$

geschehen sein.

$$\Rightarrow \quad \text{Zustand 1 ist periodisch mit Periode 4}$$

$$\xRightarrow{\text{Lemma 12}} \quad \text{Die Zustände } 1,2,3,4 \text{ sind periodische mit Periode 4}$$

**Beispiel 2.3 *(Getränkeautomat)***
Da es einen Pfad der Länge 1 von Zustand 0 zu sich selbst gibt, ist dieser *aperiodisch*. Aus der *Irreduzibilität* der Kette und Lemma 12 folgt dirket, dass Zustand 1 ebenfalls aperiodisch ist.

**Definition 13 (*absorbierende Markov-Kette*)**

*Eine Markov-Kette heißt* ***absorbierend****, wenn jeder Zustand entweder absorbierend oder transient ist.*

Diesen absorbierenden Markov-Ketten soll ein eigenständiger Vortrag gewidmet werden, sodass wir an dieser Stelle nicht weiter auf sie eingehen.

# 4 Reguläre Ketten

Während absorbierende Markov-Ketten zur Beschreibung von terminierenden Prozessen dienen, liegt der Fokus in diesem Abschnitt auf solchen Vorgängen, die beständig fortlaufen.

**Definition 14 (*reguläre Markov-Kette*)**

*Eine endliche Markov-Kette bzw. ihre Übergangsmatrix heißt **regulär**, wenn sie irreduzibel und aperiodisch ist.*

Ein Kriterium für diese Eigenschaft liefert uns

**Lemma 15**

*Sei $P$ die Übergangsmatrix einer endlichen Markov-Kette $(M_n)_{n\geq 0}$. Dann gilt:*

$$\exists n \in \mathbb{N}: \quad \text{Jeder Eintrag von } P^n \text{ ist positiv} \quad \Leftrightarrow \quad (M_n)_{n\geq 0} \text{ ist regulär}$$

**Beweis :** '$\Leftarrow$': Die Gültigkeit dieser Aussage ist direkt ersichtlich. Deshalb verzichten wir an dieser Stelle auf den formalen Beweis und verweisen auf Goodman, S. 161 und S. 154f.
'$\Rightarrow$': Sei $n \in \mathbb{N}$ so gewählt, dass jeder Eintrag von $P^n$ positiv ist. Aus der Chapman-Kolmogorov-Gleichung (Satz 5) folgt für die Einträge von $P^{n+1}$

$$p_{ij}^{(n+1)} = \sum_k p_{ik} p_{kj}^{(n)} \tag{3}$$

Da $P^n$ eine stochastische Matrix ist und somit die Zeilensummen 1 ergeben, gibt es für jedes $i \in \mathscr{S}$ ein $k \in \mathscr{S}$, sodass $p_{ik} > 0$. Da alle Einträge von $P^n$ nach Voraussetzung positiv sind, ist die rechte Seite in Gleichung (3) größer als Null. Iteriert man diese Vorgehensweise, erhält man:

$$p_{ij}^{(m)} > 0 \quad \forall m \geq n$$
$$\Rightarrow \begin{cases} i \leftrightarrow j \quad \forall i, j \in \mathscr{S} & \Rightarrow \quad \text{irreduzibel} \\ \text{Es existieren Pfade jeder Länge } m \geq n \text{ von } i \text{ nach } j & \Rightarrow \quad \text{aperiodisch} \end{cases}$$
$$\Rightarrow \quad \text{regulär}$$

■

In vielen Anwendungen ist es sicherlich interessant, das Langzeitverhalten eines Prozesses studieren und Antworten auf Konvergenzfragen geben zu können. Eine Konvergenz fast sicher $(M_n)_{n\geq 0} \xrightarrow[n\to\infty]{} s^* \in \mathscr{S}$ wird bei regulären Ketten nicht vorliegen, weil es keinen absorbierenden Zustand gibt.

Jedoch könnte man im Beispiel *Getränkeautomat* auch Fragen folgender Art stellen: Gesetzt meine Maschine ist intakt, wie wahrscheinlich ist es, dass sie es in 10, in 100 oder in 250 Tagen ist? Gibt es dabei ein Konvergenzverhalten?

Zur Beantwortung dieser Fragen betrachten wir abermals die Übergangsmatrix $P$ und widmen uns zunächst dem

**Beispiel 2.4** ***(Getränkeautomat)***
Gehen wir davon aus, dass die Wahrscheinlichkeiten für den Wechsel der Maschine von intakt zu defekt $\delta = 0.2$ und von defekt zu intakt $\gamma = 0.9$ betragen. Somit ist

$$P = \begin{pmatrix} 0.8 & 0.2 \\ 0.9 & 0.1 \end{pmatrix}$$

und weiter durch Matrizenmultiplikation:

$$P^2 = \begin{pmatrix} 0.82 & 0.18 \\ 0.81 & 0.19 \end{pmatrix} \qquad P^3 = \begin{pmatrix} 0.818 & 0.182 \\ 0.819 & 0.181 \end{pmatrix}$$

$$P^4 = \begin{pmatrix} 0.8182 & 0.1818 \\ 0.8181 & 0.1819 \end{pmatrix} \approx P^5$$

Betrachten wir zunächst nur die Entwicklung der ersten Zeile, das heißt die bedingte Verteilung, gegeben dass der Getränkeautomat zu Beginn intakt ist

$$\mathbb{P}^{M_n|M_0=0}, \quad 1 \leq n \leq 5, \quad j \in \{0,1\}$$

Es fällt auf, dass die Differenzen

$$|p_{0j}^{(n+1)} - p_{0j}^{(n)}|, \quad 1 \leq n \leq 4, \quad j \in \{0,1\},$$

mit steigendem $n$ kleiner werden. Ein ähnliches Verhalten in Zeile zwei begründet den Verdacht, dass die Zeilen der Matrizen $(P^n)_{n\geq 1}$ konvergieren.

Ferner kann man beobachten, dass auch die Differenzen

$$|p_{0j}^{(n)} - p_{1j}^{(n)}|, \quad 1 \leq n \leq 4, \quad j \in \{0,1\},$$

bei steigendem $n$ fallen.

Die Zeilen der Matrizen $(P^n)_{n\geq 1}$ aus dem Beispiel scheinen jeweils gegen ein und denselben Zeilenvektor $(a_1 \dots a_d)$ zu konvergieren. Diesen Verdacht bestätigt der folgende Satz. Es gilt also

$$\sum_{j\in A} p_{ij}^{(n)} = \mathbb{P}^{M_n|M_0=i}(A) \xrightarrow[n\to\infty]{} \mathbb{P}^*(A) := \sum_{j\in A} a_j, \quad \forall i \in \mathscr{S},$$

d.h. es liegt Verteilungskonvergenz vor.

**Satz 16 (*Grenzwerttheorem*)**

*Sei P eine reguläre Übergangsmatrix. Dann gilt:*

$$P^{(n)} \quad \xrightarrow[n\to\infty]{} \quad A = \begin{pmatrix} a_1 & a_2 & \dots & a_d \\ a_1 & a_2 & \dots & a_d \\ \vdots & & \ddots & \vdots \\ a_1 & a_2 & \dots & a_d \end{pmatrix},$$

*mit* $0 < a_j \leq 1$ *f.a.* $j \in \mathscr{S}$ *und* $a_1 + \ldots + a_d = 1$.
*A heißt* ***Grenzmatrix****, der Vektor* $(a_1 \ldots a_d)$ ***Grenzvektor****.*

Zum Beweis des Grenzwerttheorems benötigen wir das folgende Mittlungslemma. Dieses bleibt hier ohne Beweis.

**Lemma 17 (*Mittlungslemma*)**

*Für die Gewichte* $w_1, \ldots, w_d \geq \beta > 0$ *mit* $\sum_{k=1}^{d} w_k = 1$ *und die Zahlen* $v_1, \ldots, v_d \geq 0$ *mit* $m := \min\limits_{1\leq i\leq d} v_i$ *und* $M := \max\limits_{1\leq i\leq d} v_i$ *gilt:*

$$(1-\beta)m + \beta M \quad \leq \quad \sum_{i=1}^{d} v_i w_i \quad \leq \quad \beta m + (1-\beta)M$$

**Beweis von Satz 16 :** Die Markov-Kette mit Übergangsmatrix P habe mindestens zwei Zustände (sonst ist nichts zu zeigen). Ferner dürfen wir o. E.

$$p_{ij} \geq \beta > 0 \quad \forall i, j \in \mathscr{S}$$

annehmen. Sonst existiert mit Lemma 15 ein Folgenindex $n$, für den alle Einträge in $P^{(n)}$ positiv sind. Argumentiere dann ab diesem Index.
Da P eine mindestens $2 \times 2$ große stochastische Matrix ist, gilt $\beta \leq \frac{1}{2}$
Betrachte nun eine feste aber beliebige Spalte $j$ von $P$ und definiere

$$m_n := \min_{1\leq k\leq d} p_{kj}^{(n)} \qquad M_n := \max_{1\leq k\leq d} p_{kj}^{(n)}$$

$p_{ij}^{(n+1)}$ berechnet sich nach Matrizenrechnung durch $\sum_{k=1}^{d} p_{ik} p_{kj}^{(n)}$. Da $p_{kj}^{(n)} \geq 0$, f.a. $k \in \mathscr{S}$, $p_{i1} + \ldots + p_{id} = 1$ (stochastische Matrix) und $p_{ik} > 0$, f.a. $k, i \in \mathscr{S}$, nach Voraussetzung gelten, können wir das Mittlungslemma anwenden:

$$\overset{\text{ML}}{\Longrightarrow} \quad (1-\beta)m_n + \beta M_n \quad \leq \quad p_{ij}^{(n+1)} \quad \leq \quad \beta m_n + (1-\beta)M_n, \qquad (4)$$
$$\forall n \in \mathbb{N} \text{ und } \forall i \in \mathscr{S}$$

Da die Ungleichung (4) für alle $i \in \mathscr{S}$ gilt, folgt:

$$\begin{aligned} M_{n+1} &\leq \beta m_n + (1-\beta)M_n &&\leq \beta M_n + (1-\beta)M_n = M_n \\ m_{n+1} &\geq (1-\beta)m_n + \beta M_n &&\geq (1-\beta)m_n + \beta m_n = m_n \end{aligned} \tag{5}$$

$$\Rightarrow \quad M_{n+1} - m_{n+1} \leq (1-2\beta)(M_n - m_n)$$

Iteriere diesen Prozess:

$$M_{n+1} - m_{n+1} \leq \underbrace{(1-2\beta)}_{\in[0,1)}{}^n \underbrace{(M_1 - m_1)}_{<1} \quad \xrightarrow[n\to\infty]{} \quad 0$$

$$\Rightarrow \quad p_{ij}^{(n)} \quad \xrightarrow[n\to\infty]{} \quad a_j \in \mathbb{R} \quad \forall i \in \mathscr{S}$$

Alle Einträge der j-ten Spalte konvergieren also gegen dieselbe Zahl $a_j$ Des Weiteren stellen wir fest:

$$\overset{(5)}{\Longrightarrow} \quad m_n \underset{n\to\infty}{\uparrow} a_j \quad \Rightarrow \quad a_j > 0$$

Da alle Zeilensummen von $P^n$ für alle $n \in \mathbb{N}$ 1 ergeben, gilt dies auch im Grenzübergang $P^n \xrightarrow[n\to\infty]{} A$ für die Matrix $A$. ■

**Bemerkung** Das Grenzwerttheorem trifft weitere folgende interessante Aussagen für eine reguläre Markov-Kette, interpretiert als Brettspiel:

1. Jeder Knoten kann bei genügend vielen Schritten mit positiver Wahrscheinlichkeit erreicht werden, egal wo der Marker startet.

2. Die Wahl des Startknotens zur Entscheidung der Frage, auf welchem Knoten der Marker nach $n$ Schritten ist, wird mit wachsendem $n$ immer unbedeutender, bis die Abhängigkeit im Grenzübergang $n \to \infty$ ganz verschwindet.

# 5 Stationäre Verteilung und Ergodensatz

Bislang haben wir unseren Marker zu Beginn auf einen beliebigen aber festen Knoten $i$ gesetzt und die Entwicklung der Markov-Kette $(M_n)_{n\geq 0}$, gegeben $M_0 = i$, betrachtet. Nun erlauben wir eine zufällige Auswahl des Startzustandes und ziehen uns dabei wiederum auf endliche Markov-Ketten mit Zustandsraum $\mathscr{S} = \{1, 2, \ldots, d\}$ zurück. Die Zähldichte der *Anfangsverteilung* schreiben wir als Zeilenvektor

$$\pi = (\pi_1 \ \pi_2 \ \ldots \ \pi_d), \quad \text{mit } \pi_i = \mathbb{P}(M_0 = i) \quad \forall i \in \mathscr{S}$$

sodass wir für die Anfangsverteilung $\pi$ folgende Zusammenhänge erhalten:

$$\mathbb{P}^{M_0}(A) = \mathbb{P}(M_0 \in A) \overset{\text{Def}}{=} \pi(A) = \sum_{i \in A} \pi_i, \quad \forall A \in \mathscr{G} = \mathscr{P}(\mathscr{S})$$

Um den Anforderungen an eine Zähldichte gerecht zu werden, muss gelten:

$$\sum_{i \in \mathscr{S}} \pi_i = 1$$

Mit dieser Anfangsverteilung können wir nun die Zähldichte für $M_n$ bestimmen. Konditioniere dafür $M_n$ mit $M_0$ über den Satz von der totalen Wahrscheinlichkeit:

$$\begin{aligned}
\mathbb{P}(M_n = j) &\underset{\text{tot. W'keit}}{\overset{\text{Satz v. d.}}{=}} \sum_{i \in \mathscr{S}} \mathbb{P}(M_0 = i)\mathbb{P}(M_n = j | M_0 = i) \\
&\overset{\text{Satz 6}}{=} \sum_{i \in \mathscr{S}} \pi_i (P^n)_{ij} \\
&= (\pi P^n)_j
\end{aligned}$$

Der Zeilenvektor $\pi P^n$ stellt also die Zähldichte für $M_n$ bereit.

**Definition 18 (*Stationäre Verteilung*)**

*Der Vektor $\pi = (\pi_1 \ \ldots \ \pi_d)$ mit $\sum_{i=1}^{d} \pi_i = 1$ heißt **stationäre Verteilung** der Markov-Kette mit Übergangsmatrix $P$, falls gilt*

$$\pi P = \pi$$

**Bemerkung** Ist $\pi$ eine stationäre Verteilung, gilt wegen

$$\pi P^2 = (\pi P) P = \pi P = \pi$$

auch

$$\pi P^n = \pi \quad \forall n \in \mathbb{N}$$

und somit die Konvergenz in Verteilung

$$\pi P^n \xrightarrow[n \to \infty]{} \pi$$

Weitere Konvergenzaussagen trifft der

**Satz 19** (***Ergodensatz für reguläre Markov-Ketten***)

*Sei $P$ die Übergangsmatrix einer regulären Markov-Kette und $\pi$ der zugehörige Grenzvektor. Dann gilt:*

1. *$\pi$ ist die einzige stationäre Verteilung dieser Kette.*
2. *Für jede Anfangsverteilung $\rho$ konvergiert die $n$-Schritt-Verteilung $\rho P^n$ für $n \to \infty$ gegen $\pi$*

**Beweis :** Seien $e_i$, $1 \leq i \leq d$, die $i$-ten Einheitsvektoren und A die Grenzmatrix der Markov-Kette, d.h.

$$A = \begin{pmatrix} \pi \\ \pi \\ \vdots \\ \pi \end{pmatrix}$$

wobei $\pi$ der Grenzvektor ist.

$$\begin{aligned} &\Rightarrow \quad e_i A = \pi \quad \forall 1 \leq i \leq d \\ &\Rightarrow \quad \lim_{n\to\infty} e_i P^n = \pi \quad \forall 1 \leq i \leq d \end{aligned} \tag{6}$$

Insbesondere gilt:

$$\begin{aligned} \pi P = \lim_{n\to\infty} e_1 P^{n+1} &= \lim_{n\to\infty} e_1 P^n = \pi \\ \Rightarrow \quad &\pi \text{ ist stationäre Verteilung} \end{aligned}$$

Ist $\rho$ eine beliebige Anfangsverteilung, so gilt

$$\rho = (\rho_1 \ \rho_2 \ \ldots \ \rho_d) = \sum_{i=1}^{d} r_i e_i \tag{7}$$

mit $r_i \geq 0$ und $r_1 + \ldots + r_d = 1$
Die Verteilung für $M_n$, $n \geq 1$, ist in diesem Fall gegeben durch

$$\rho P^n \stackrel{(7)}{=} \sum_{i=1}^{d} r_i e_i P^n \quad \forall n \in \mathbb{N}$$

Lasse nun $n \to \infty$ streben und nutze (6):

$$\lim_{n\to\infty} \rho P^n = \lim_{n\to\infty} \sum_{i=1}^{d} r_i e_i P^n \underset{\text{regeln}}{\stackrel{\text{Limes-Rechen-}}{=}} \sum_{i=1}^{d} r_i \lim_{n\to\infty} e_i P^n \stackrel{(6)}{=} \sum_{i=1}^{d} r_i \pi \stackrel{r_1+\ldots+r_d=1}{=} \pi \tag{8}$$

Somit ist 2. gezeigt.
Ist $\rho$ eine stationäre Verteilung für $P$, ist die linke Seite von Gleichung (8) gleich $\rho$

$$\begin{aligned} &\Rightarrow \quad \pi = \rho \\ &\Rightarrow \quad \pi \text{ ist die eindeutige stationäre Verteiltung} \quad \Rightarrow \quad 1. \end{aligned}$$

■