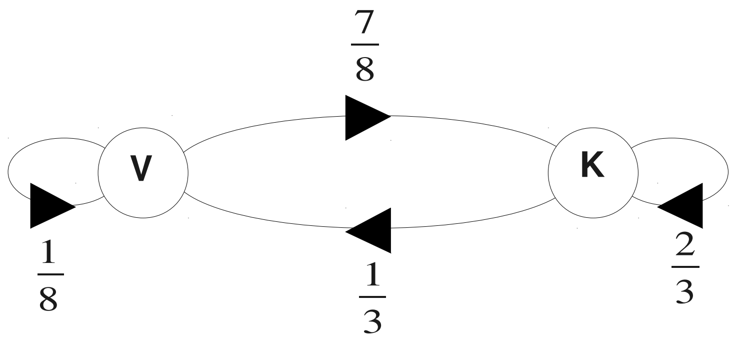
LANGZEITVERHALTEN VON MARKOW-KETTEN

NORA LOOSE

1. Buchstabensalat und Definition

Andrei Andrejewitsch Markow berechnete Anfang des 20. Jahrhunderts die Buchstabensequenzen in russischer Literatur. 1913 untersuchte er in einem Artikel [1] eine Folge von 20 000 Buchstaben aus dem Gedicht „Eugen Onegin“ von Puschkin und stellte fest, dass Konsonanten (K) und Vokale (V) so aufeinander folgen:



Aus diesem Ansatz entwickelten sich die Markow-Ketten, Modelle für Zufallsprozesse, die in diskreter Zeit eine Folge von Zuständen durchlaufen. Dabei hängt das Verhalten des Systems nur vom aktuellen Zustand und nicht von der gesamten Vorgeschichte ab. Diese Gedächtnislosigkeit nennt man die Markow-Eigenschaft.
Eine Möglichkeit, eine Markow-Kette zu beschreiben, ist die folgende:

Definition 1. Sei $S = \{1 \dots N\}$ die Menge aller möglichen Zustände.
Eine Folge von Zufallsversuchen $\{X_n\}_{n=0,1,2...}$ mit Werten in S heißt (zeit-) **homogene Markow-Kette** mit diskretem Zeitparameter n und endlichem Zustandraum S, falls reelle Zahlen $\{p_{ij}\}_{i,j=1...N}$ existieren mit

$$0 \leq p_{ij} \leq 1, \qquad \sum_{i=1}^{N} p_{ij} = 1,$$

so dass folgende Eigenschaft erfüllt ist:
Ist der Prozess im j-ten Zustand, springt er mit dem nächsten Schritt mit der Wahrscheinlichkeit p_{ij} in den i-ten Zustand.

Im Folgenden wollen wir homogene Markow-Ketten mit diskretem Zeitparameter und endlichem Zustandsraum untersuchen. Wir nennen solche Prozesse der Einfachheit halber kurz Markow-Ketten.
Eine Markow-Kette ist vollständig bestimmt, wenn man außer den Übergangswahrscheinlichkeiten $\{p_{ij}\}_{i,j=1...N}$ noch eine **Anfangsverteilung** $\{\mu_i\}_{i=1...N}$

angibt, wobei μ_i die Wahrscheinlichkeit ist, dass die Kette im Zustand i startet.

Kommen wir nun zu unserer Markow-Kette vom Anfang zurück.
Die Menge der möglichen Zustände ist $S = \{V, K\}$ und wir haben die Übergangswahrscheinlichkeiten p_{11}, p_{12}, p_{21} und p_{22} in einen gerichteten Graphen eingetragen.
Um wie Markow die Prozentsätze der Konsonanten und Vokale im Buch „Eugen Onegin“ schätzen zu können, stellen wir uns also die Frage nach dem Langzeitverhalten dieser Markow-Kette.

Wir übersetzen das Problem zunächst in die Sprache der Linearen Algebra und bedienen uns an Werkzeugen der Matrizentheorie.

2. Die Lineare Algebra als Hilfsmittel

Wir tragen die Übergangswahrscheinlichkeiten $\{p_{ij}\}_{i,j=1\ldots N}$ einer Markow-Kette in eine $N \times N$-Matrix ein, die Übergangsmatrix P.

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{pmatrix}$$

P hat nichtnegative Elemente und Spaltensummen 1.
Eine Matrix mit diesen Eigenschaften heißt **stochastische Matrix**.
Eine gegebene Anfangsverteilung $\{\mu_i\}_{i=1\ldots N}$ stellen wir als stochastischen Vektor

$$\overrightarrow{p}(0) = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{pmatrix}$$

dar.
Wir machen folgende Beobachtung:
Sei P die Übergangsmatrix einer Markow-Kette. Der ij-te Eintrag $p_{ij}^{(n)}$ der Matrix P^n gibt die Wahrscheinlichkeit an, dass die Markow-Kette nach n Schritten im Zustand i ist, wenn sie im Zustand j gestartet ist.
Wenn man eine Anfangsverteilung (Verteilung zum Zeitpunkt 0)

$$\overrightarrow{p}(0) = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{pmatrix}$$

gegeben hat, lässt sich die Verteilung zum Zeitpunkt n berechnen durch:

$$\overrightarrow{p}(n) = P^n \overrightarrow{p}(0) \quad n = 0, 1, 2, \ldots$$

Um etwas über das Langzeitverhalten von Markow-Ketten zu erfahren, müssen wir also die Potenzen der Übergangsmatrix P untersuchen.

Aufgabe 1. (Unterrichtspraxis und/oder GeoGebra) Benutzen Sie einen GTR oder GeoGebra, um folgende Berechnungen/graphische Darstellungen zu realisieren:

(1) Untersuchen Sie für Markows Buchstabensequenzen für verschiedene Anfangsverteilungen die Entwicklung des Vektors $\overrightarrow{p}(n)$.
Man kann beobachten, dass sich, unabhängig von der Anfangsverteilung, der Vektor $\overrightarrow{p}(n)$ einer festen Verteilung annähert.
(2) Berechnen Sie die ersten paar Potenzen von P. Welche Vermutung liegt nahe?
Die Matrix konvergiert und in den Spalten steht der gleiche Vektor.
Für Matrizenrechnungen mit GeoGebra geben Sie die Matrixelemente in die Tabelle der Tabellenansicht ein. Unter dem Icon „Liste erzeugen“ finden Sie auch den Icon „Matrix erzeugen“. Die Matrix wird dann in der Algebra-Ansicht dargestellt. In der Eingabeleiste kann man nun Berechnungen mit Matrizen eingeben, z.B. Matrix1ˆ3 usw.

3. Der Hauptsatz für reguläre Markow-Ketten

Definition 2. Eine Markow-Kette mit Übergangsmatrix P heißt **regulär**, falls ein $k \in \mathbb{N}$ existiert, so dass P^k nur strikt positive Einträge hat.

Es gilt der **Hauptsatz für reguläre Markow-Ketten**:

Satz 1. *Sei P die Übergangsmatrix einer regulären Markow-Kette.*

i) *Dann existiert $\lim_{n\to\infty} P^n = W$ und in der Grenzmatrix W steht in allen Spalten der gleiche Vektor $\overrightarrow{w}$.*
ii) *$\overrightarrow{w}$ ist ein stochastischer Vektor mit strikt positiven Einträgen.*

Aufgabe 2. (Rationales Argumentieren)
Der Hauptsatz für reguläre Markow-Ketten lässt sich mithilfe von elementaren Techniken aus der Folgentheorie beweisen.
Ergänzen Sie im Beweis die Lücken!
Beweis von Satz 1: Sei P die $N \times N$-Übergangsmatrix einer regulären Markow-Kette, $N \geq 2$, und sei $d \geq 0$ das kleinste Element von P.

Wir wollen zeigen, dass die Potenzen P^n gegen eine Matrix W konvergieren, die lauter gleiche Spalten hat. Äquivalent können wir zeigen, dass W konstante Zeilen hat.
Wählen wir also eine feste Zeile i in der Matrix $P^n = (p_{ij}^{(n)})$. Wir nennen ihr maximales und minimales Element

$$M_n = \max_j p_{ij}^{(n)}, \qquad m_n = \min_j p_{ij}^{(n)} = p_{ia}^{(n)}.$$

Wir rechnen:

$$
\begin{aligned}
M_{n+1} = \max_j p_{ij}^{(n+1)} &= \max_j \sum_{k=1}^{N} p_{ik}^{(n)} \cdot p_{kj} = \max_j \left(\sum_{k \neq a} p_{ik}^{(n)} \cdot p_{kj} + m_n \cdot p_{aj} \right) \\
&\leq \max_j \left(M_n \sum_{k \neq a} p_{kj} + m_n \cdot p_{aj} \right) = M_n + \max_j (m_n - M_n) p_{aj} \\
&= M_n - (M_n - m_n) \min_j p_{aj} \leq M_n - d(M_n - m_n).
\end{aligned}
$$

Wir haben also

$$M_{n+1} \leq M_n - d(M_n - m_n). \tag{1}$$

Zeigen Sie analog zu oben

$$m_{n+1} \geq m_n + d(M_n - m_n). \tag{2}$$

Sei

$$M_n = \max_j p_{ij}^{(n)} = p_{ib}^{(n)}$$

.

$$
\begin{aligned}
m_{n+1} = \min_j p_{ij}^{(n+1)} &= \min_j \sum_{k=1}^{N} p_{ik}^{(n)} \cdot p_{kj} = \min_j \left(\sum_{k \neq b} p_{ik}^{(n)} \cdot p_{kj} + M_n \cdot p_{bj} \right) \\
&\geq \min_j \left(m_n \sum_{k \neq b} p_{kj} + M_n \cdot p_{bj} \right) = m_n + (M_n - m_n) \min_j p_{bj} \\
&\geq m_n + d(M_n - m_n).
\end{aligned}
$$

Argumentieren Sie, dass aus (1)

$$M_1 \geq M_2 \geq M_3 \geq \dots$$

und aus (2)

$$m_1 \leq m_2 \leq m_3 \leq \dots$$

folgt.
Folgt, da $M_n - m_n \geq 0$ und $d \geq 0$.
Also sind beide Folgen monoton und beschränkt:

$$m_1 \leq m_n \leq \dots \leq M_n \leq M_1$$

Es folgt, dass die Grenzwerte $m := \lim_{n\to\infty} m_n$ und $M := \lim_{n\to\infty} M_n$ existieren. Wir wissen, dass $m \leq M$.
Wir sind mit Teil i) des Beweises fertig, wenn wir zeigen können $m = M$. (Wieso?)
Weil dann jeder Eintrag jeder Zeile von P^n konvergiert (und damit die ganze Matrix P^n) und die Grenzmatrix hat konstante Zeilen.
Das ist der Fall, wenn die Folge $M_n - m_n$ eine Nullfolge ist.
Um dies zu zeigen, verwenden wir nun, dass P eine reguläre Matrix ist, d.h. ein $k \in \mathbb{N}$ existiert, so dass P^k nur strikt positive Einträge hat.

a) Zunächst nehmen wir an, dass schon P nur strikt positive Einträge hat (also $k = 1$ ist). Begründen Sie:

$$0 < d \leq \frac{1}{2}$$
$$\iff 0 \leq 1 - 2d < 1$$

Nach Voraussetzung sind alle Einträge von P strikt positiv, also $d > 0$. $d \leq \frac{1}{2}$ gilt, da wir angenommen haben, dass $N \geq 2$ und die Spaltensummen von P jeweils 1 sind.

Subtraktion der Ungleichungen (1) und (2) liefert:

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

Daraus folgt:

$$M_n - m_n \leq (1 - 2d)^{n-1}(M_1 - m_1).$$

Argumentieren Sie, dass die Folge $M_n - m_n$ eine Nullfolge ist.
Gilt, da $0 \leq 1 - 2d < 1$.

b) Falls P^k, $k > 1$, strikt positive Einträge hat, wenden Sie die Argumentation aus a) auf die Matrix $\tilde{P} := P^k$ an.
Folgern Sie, dass die Folge $M_{nk} - m_{nk}$ eine Nullfolge ist.
Sei

$$\tilde{M_n} = \max_j \tilde{p}_{ij}^{(n)}, \qquad \tilde{m_n} = \min_j \tilde{p}_{ij}^{(n)}.$$

Dann ist mit Teil a) die Folge $\tilde{M_n} - \tilde{m_n}$ eine Nullfolge. Das ist aber gerade die Folge $M_{nk} - m_{nk}$.
Argumentieren Sie mit Teilfolgen, um auch für diesen Fall die Aussage zu beweisen.
m_n und M_n konvergieren, also auch $M_n - m_n$. Eine Teilfolge von $M_n - m_n$ konvergiert gegen 0. Also konvergiert auch schon die Folge $M_n - m_n$ gegen Null.
Warum stimmt Teil ii) des Satzes?
Jeder Eintrag von $\overrightarrow{w}$ ist nach dem eben gezeigten strikt größer Null, da er der Grenzwert der monoton wachsenden Folge m_n ist und für irgendein k ist m_k ein Eintrag einer Matrix P^k, die nur strikt positive Einträge hat.
Der Vektor $\overrightarrow{w}$ ist stochastisch, da er Spaltenvektor der Matrix W ist, die der Grenzwert von stochastischen Matrizen ist.

4. Langzeitverhalten von regulären Markow-Ketten

Aufgabe 3. (Rationales Argumentieren)
Sei P die Übergangsmatrix einer regulären Markow-Kette und $W = \lim_{n\to\infty} P^n$ die Grenzmatrix mit Spaltenvektoren $\overrightarrow{w}$.

(1) Zeigen Sie, dass $P\overrightarrow{w} = \overrightarrow{w}$ gilt, d.h. $\overrightarrow{w}$ ist ein Fixvektor für P.
Es gilt $W\overrightarrow{w} = W$: der i-te Eintrag von $W\overrightarrow{w}$ ist gleich

$$w_i \cdot w_1 + w_i \cdot w_2 + \ldots + w_i \cdot w_n = w_i(w_1 + w_2 + \ldots\ w_n) = w_i,$$

da $\overrightarrow{w}$ ein stochastischer Vektor ist. Aus $P^{n+1} = PP^n$ folgt für $n \to \infty$ $W = PW$. Multiplikation von rechts mit $\overrightarrow{w}$ ergibt das Gewünschte.

(2) Falls man $P\overrightarrow{v} = \overrightarrow{v}$ für einen Vektor $\overrightarrow{v}$ hat, gilt: $\overrightarrow{v} = c\overrightarrow{w}$ für ein $c \in \mathbb{R}$.
D.h. $\overrightarrow{w}$ ist der einzige stochastische Fixvektor für P.
Aus $P\overrightarrow{v} = \overrightarrow{v}$ folgt $P^n\overrightarrow{v} = \overrightarrow{v}$ für alle n und für $n \to \infty$ erhält man $W\overrightarrow{v} = \overrightarrow{v}$. Andererseits kann man $W\overrightarrow{v}$ wie eben ausrechnen: Der i-te Eintrag von $W\overrightarrow{v}$ ist gleich
$$w_i \cdot v_1 + w_i \cdot v_2 + \ldots + w_i \cdot v_n = w_i(v_1 + v_2 + \ldots v_n) = c \cdot w_i,$$
wobei c die Summe der Komponenten von $\overrightarrow{v}$ ist.
Also $\overrightarrow{v} = W\overrightarrow{v} = c \cdot \overrightarrow{w}$

(3) Sei $\overrightarrow{p}(0)$ eine beliebige Anfangsverteilung.
Was gilt für $W\overrightarrow{p}(0)$?
Da $\overrightarrow{p}(0)$ ein stochastischer Vektor ist, gilt wie eben, dass der i-te Eintrag von $W\overrightarrow{p}(0)$ gleich
$$w_i \cdot \mu_1 + w_i \cdot \mu_2 + \ldots + w_i \cdot \mu_n = w_i(\mu_1 + \mu_2 + \ldots \mu_n) = w_i$$
ist. Also gilt $W\overrightarrow{p}(0) = \overrightarrow{w}$
Interpretieren Sie das Ergebnis.
Unabhängig vom Startzustand konvergiert die Markow-Kette für $n \to \infty$ gegen die Grenzverteilung $\overrightarrow{w}$.

Die Ergebnisse aus Übung 3 legen eine Erklärung nahe, wie Markow die Prozentsätze von Konsonanten und Vokalen in „Eugen Onegin“ approximieren konnte.

Definition 3. Eine Wahrscheinlichkeitsverteilung $\overrightarrow{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{pmatrix}$ heißt **stationär**, falls $P\overrightarrow{\mu} = \overrightarrow{\mu}$

Übung 3 zeigt also, dass die Grenzverteilung $\overrightarrow{w}$ gegeben ist durch die eindeutige stationäre Verteilung von P.
In der Sprache der Linearen Algebra ausgedrückt, entspricht diese gerade dem eindeutigen stochastischen Eigenvektor von P zum Eigenwert 1.
Als nächstes wollen wir aus schulpraktischer Perspektive verschiedene Methoden diskutieren, die stationäre Verteilung einer regulären Markow-Kette zu berechnen.

5. Berechnung der stationären Verteilung

Aufgabe 4. (Unterrichtspraxis)

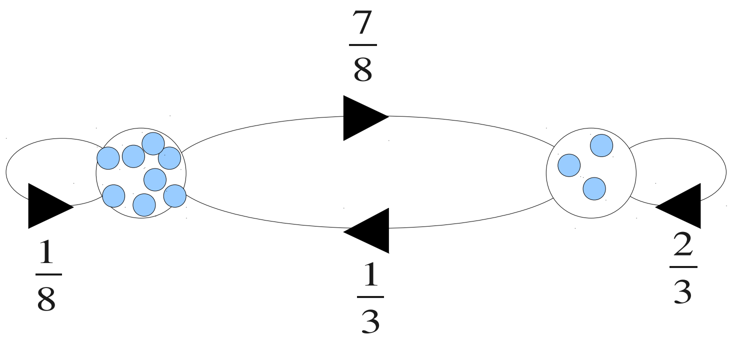
(1) Stellen Sie verschiedene Ansatzmöglichkeiten von Schülerlösungen vor, die stationäre Verteilung von Markows Buchstabenkette zu berechnen.

(2) In seinem Buch „Wahrscheinlichkeitsrechnung und Statistik“ [2] präsentiert A. Engel einen Algorithmus zur Bestimmung der stationären Verteilung einer regulären Markow-Kette, wenn die Übergangswahrscheinlichkeiten rationale Zahlen sind:

Der Übergangsgraph einer Markow-Kette wird als Spielbrett aufgefasst, auf dem Steine durch den Graphen bewegt werden, wobei die Übergangswahrscheinlichkeiten respektiert werden.

Folgen Sie der Spielanleitung, um die stationäre Verteilung für das Buchstabenbeispiel zu finden!

Für jeden Zustand j sei a_j das kleinste gemeinsame Vielfache der Nenner der nichtverschwindenen Einträge in der j-ten Spalte der Übergangsmatrix.
Starten Sie, indem Sie für jedes j a_j Spielsteine in den j-ten Zustand legen. Als nächstes ziehen Sie von jedem Zustand j $a_j \cdot p_{ij}$ Steine in den i-ten Zustand. Die Anzahl der Steine im Zustand j nach dem Zug muss kein Vielfaches von a_j sein. Laden Sie also jeden Zustand j mit der kleinsten Anzahl von Steinen nach, so dass ein Vielfaches von a_j erreicht wird. Dann führen Sie an allen Zuständen gleichzeitig den nächsten Zug durch usw.
Dieser Prozess wird letztlich einen Punkt erreichen, an dem die Anzahl von Spielsteinen in jedem Zustand nach einem Zug dieselbe ist wie vor dem Zug. Jetzt haben Sie den Fixvektor gefunden und durch Normierung auf die Komponentensumme 1 erhalten Sie die stationäre Verteilung.



$\xrightarrow{\text{laden}} (8,3) \xrightarrow{\text{ziehen}} (2,9) \xrightarrow{\text{nachladen}} (8,9) \xrightarrow{\text{ziehen}} (4,13) \xrightarrow{\text{nachladen}} (8,15) \xrightarrow{\text{ziehen}}$
$(6,17) \xrightarrow{\text{nachladen}} (8,18) \xrightarrow{\text{ziehen}} (7,19) \xrightarrow{\text{nachladen}} (8,21) \xrightarrow{\text{ziehen}} (8,21)$

Normierung ergibt die stationäre Verteilung $\overrightarrow{\mu} = \begin{pmatrix} 8/29 \\ 21/29 \end{pmatrix}$

6. Zusatz: Auch Google funktioniert nicht ohne Markow-Ketten (Bearbeitung optional)

Larry Page ist ein US-amerikanischer Informatiker und Mitbegründer der Suchmaschine Google. Page entwickelte 1998 an der Stanford University den nach ihm benannten PageRank-Algorithmus, der einen wichtigen Bestandteil des Sortieralgorithmus der Suchmaschine Google darstellt.
Der PageRank-Algorithmus ist eine spezielle Methode, die Linkpopularität einer Webseite festzulegen. Google hat ein Bewertungssystem eingeführt, in

dem Webseiten der „PageRank“, ein Wert von 0 bis 10 zugewiesen wird, wobei 10 den höchsten Bekanntheitsgrad darstellt.

Das Verfahren kann mithilfe der Markow-Ketten-Theorie interpretiert werden[3]:
Die Menge aller Webseiten wird als riesiger Zustandsraum $S = \{1 \dots N\}$ einer regulären Markow-Kette modelliert. Sei $0 < d < 1$ fest.
Wir stellen uns einen zufälligen Surfer vor, der sich folgendermaßen durch das Netz bewegt: Angenommen der Surfer befinde sich auf der Webseite j und auf der Webseite j gibt es Links auf insgesamt C_j verschiedene Webseiten. Dann wählt der Surfer mit der Wahrscheinlichkeit d eine der C_j auf j verlinkten Webseiten und jede dieser mit gleicher Wahrscheinlichkeit $d \cdot \frac{1}{C_j}$.
Mit der Wahrscheinlichkeit $1 - d$ wählt er eine beliebige neue Seite, ohne sich an den Links zu orientieren.

Aufgabe 5. (Unterrichtspraxis)
Geben Sie die Einträge in der Übergangsmatrix an!

$$p_{ij} = \begin{cases} (d-1)/N & \text{falls } i \text{ nicht von } j \text{ aus verlinkt} \\ (d-1)/N + d/C_j & \text{falls } i \text{ von } j \text{ aus verlinkt} \end{cases}$$

Angenommen die Markow-Kette ist regulär, dann gibt es eine eindeutige stationäre Verteilung $\overrightarrow{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{pmatrix}$, die unabhängig von der Anfangsverteilung die Wahrscheinlichkeiten angibt, mit der die Seiten von dem Surfer nach einer langen Zeit besucht werden.
Der PageRank der Webseite j wird proportional zu der entsprechenden Wahrscheinlichkeit μ_j aus der stationären Verteilung gewählt.

7. Ausblick: Langzeitverhalten irreduzibler Markow-Ketten (Bearbeitung optional)

Wir haben uns die ganze Zeit mit regulären Markow-Ketten beschäftigt. Erinnern wir uns, dass das bedeutet, dass ein $k \in \mathbb{N}$ existiert, so dass man von irgendeinem Zustand j zu jedem Zustand i in genau k Schritten gelangen kann. Diese Voraussetzung ist nicht sehr natürlich. Es gibt eine zu „regulär“ äquivalente, natürlichere Bedingung. Dazu benötigen wir folgende Definitionen:

Definition 4. Eine Markow-Kette heißt **irreduzibel**, wenn es für zwei beliebige Zustände j und i ein $k \in \mathbb{N}$ gibt, so dass $p_{ij}^{(k)} > 0$. D.h. man kann von einem beliebigen Zustand j zu jedem Zustand i gelangen, aber nicht notwendigerweise in derselben Schrittzahl.

Diese Eigenschaft kann man oft einem Graphen ansehen.

Definition 5. Für eine Markow-Kette heißt $d = \text{ggT}\{n | p_{ii}^{(n)} > 0\}$ die **Periode** des Zustands i.
Eine Markow-Kette heißt **aperiodisch**, wenn die Periode jedes Zustands 1 ist.

Wenn z.B. $p_{ii} > 0$, dann ist die Periode von i gleich 1.
Es gilt der folgende Satz, den wir nicht beweisen wollen:

Lemma 1. *Eine Markow-Kette ist genau dann regulär, wenn sie irreduzibel und aperiodisch ist.*

Für irreduzible Markow-Ketten gilt, dass alle Zustände dieselbe Periode haben.
Wenn wir also schon wissen, dass eine Markow-Kette irreduzibel ist, reicht es z.B. einen Zustand i zu finden mit $p_{ii} > 0$, um zu folgern, dass die Kette auch aperiodisch ist und damit regulär.

7.1. **Irreduzible Markow-Ketten.** Betrachten wir folgende Übergangsmatrix einer Markow-Kette:

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

Aufgabe 6. (Unterrichtspraxis)
Prüfen Sie, ob die Matrix irreduzibel und/oder aperiodisch ist.
Die Matrix ist irreduzibel, hat jedoch Periode 2 (und ist somit nicht aperiodisch).
Existiert eine Grenzmatrix?
Nein, $P^{2k-1} = P$ und $P^{2k} = I$ für alle $k \geq 1$.
Existiert eine stationäre Verteilung $\overrightarrow{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$?
Ja, $\overrightarrow{\mu} = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$.
Falls ja, konvergiert $P^n \overrightarrow{p}(0)$ wieder für jede Anfangsverteilung $\overrightarrow{p}(0)$?
Nein, z.B. für die Anfangsverteilung $\overrightarrow{p}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist für $k \geq 1$ $P^{2k-1}\overrightarrow{p}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und $P^{2k}\overrightarrow{p}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Satz 2. *Für eine irreduzible Markow-Kette mit Übergangsmatrix P ist es immer noch wahr, dass es einen eindeutigen stochastischen Fixvektor $P\overrightarrow{w} = \overrightarrow{w}$ gibt und $\overrightarrow{w}$ ist strikt positiv.*

Aufgabe 7. (Rationales Argumentieren)
Beweisen Sie diese Behauptung, indem Sie für die irreduzible Matrix P Folgendes zeigen:

(1) $\tilde{P} = (1/2)I + (1/2)P$ ist regulär.
Die Diagonaleinträge von $\tilde{P}$ sind alle strikt größer als Null. Mit Lemma 1 und der nachfolgenden Bemerkung gilt also, dass $\tilde{P}$ regulär ist.

(2) $\tilde{P}$ hat dieselben Fixvektoren wie P.

$$\tilde{P}v = v \Leftrightarrow 1/2(I + P)v = v \Leftrightarrow 1/2Pv = 1/2v \Leftrightarrow Pv = v$$

Warum folgt jetzt die Behauptung?
Nach dem Hauptsatz für reguläre Markow-Ketten konvergiert $\tilde{P}$ gegen eine Grenzmatrix W, die in allen Spalten denselben Vektor $\overrightarrow{w}$ stehen hat, dessen Einträge strikt positiv sind. Dieser Vektor ist nach Aufgabe 3 der eindeutige stochastische Fixvektor der Matrix $\tilde{P}$. Da die Fixvektoren von $\tilde{P}$ und P übereinstimmen, ist $\overrightarrow{w}$ der eindeutige stochastische Fixvektor von P.

Für irreduzible Markow-Ketten hat die stationäre Verteilung auch eine interessante Interpretation:

Satz 3 (**Gesetz der Großen Zahlen für Irreduzible Markow-Ketten**). *Gegeben sei eine irreduzible Markow-Kette mit stationärer Verteilung* $\overrightarrow{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix}$. *Sei* $H_j^{(n)}$ *die Anzahl der Zeitpunkte bis zum* n*-ten Schritt, in denen die Kette im Zustand* j *ist. Dann gilt für beliebiges* $\epsilon > 0$

$$P\left[|H_j^{(n)} - w_j| > \epsilon\right] \longrightarrow 0,$$

unabhängig vom Startzustand.

Reguläre Markow-Ketten sind insbesondere irreduzibel. Das gibt uns eine neue Interpretation der errechneten stationären Verteilung für die Buchstaben-Markow-Kette. Satz 3 sagt, dass auf Dauer gesehen, die stationäre Verteilung auch wirklich die auftretenden Prozentsätze von Vokalen und Konsonanten angibt.

Literatur

[1] A.A. Markow. An Example of Statistical Analysis of the Text of Eugene Onegin Illustrating the Association of Trials into a Chain. Bulletin de lAcadamie Imperiale des Sciences de St. Petersburg, ser. 6, Band 7 (1913), pp. 153162.

[2] A. Engel. Wahrscheinlichkeitsrechnungn und Statistik. Band 2, Stuttgart: Klett Verlag, 1976.

[3] http://de.wikipedia.org/wiki/PageRank