Jonas Gloor
Gymnasium Oberwil BL, jonas.gloor@sbl.ch

Markow-Ketten als Anwendung der linearen Algebra im Unterricht

Markow-Ketten bieten eine Reihe für den Unterricht interessante Zugänge und erlauben es, verschiedene im Schwerpunktfach behandelte Themen (lineare Algebra, Graphentheorie, Wahrscheinlichkeit oder geometrische Reihen) zu vernetzen. Für die sonst als eher trocken geltende Matrizenrechnung ergeben sich praxisnahe Anwendungen. Auch im Bulletin waren Markow-Ketten und die in diesem Artikel diskutierte *Erste Mittelwertsregel* bereits Thema: *Hans Ulrich Keller: Markow-Ketten*, [1]. (Für weitere Anwendungen und Grundlagen zur linearen Algebra für die Sekundarstufe II: *Themenheft Matrizen*, [2].)

Beispiele für Markow-Ketten sind schnell so kompliziert, dass die vorkommenden Matrizen mühsam gross sind – oder dann so einfach, dass die Prozesse elementar beschreibbar sind (vgl. Abschnitt 2). In diesem Artikel werden im Unterricht erprobte Zugänge anhand eines Beispiels beschrieben, welches zwar nicht zu kompliziert ist, gleichzeitig aber den Einsatz fortgeschrittener Methoden rechtfertigt. Die diskutierten Methoden werden alle in gängigen Lehrmitteln behandelt (z.B. *Lambacher Schweizer, Mathematik – Stochastik*, [3]), das Beispiel ist dem Autor in der Literatur nicht begegnet.

1 Fragestellung

Eine Spielfigur startet in der Mitte eines 5×5-Schachfelds wie in Abbildung 1 zu einer Irrfahrt. In jedem Zeitschritt bewegt sich die Figur zufällig in ein Nachbarfeld, wobei auch diagonal benachbarte Felder zulässig sind. Für jedes Nachbarfeld betrage die Wahrscheinlichkeit $\frac{1}{8}$. Sobald die Figur auf einem Randfeld angekommen ist, endet die Irrfahrt. Mit welcher Wahrscheinlichkeit p_{schwarz} landet die Figur auf einem schwarzen Randfeld?

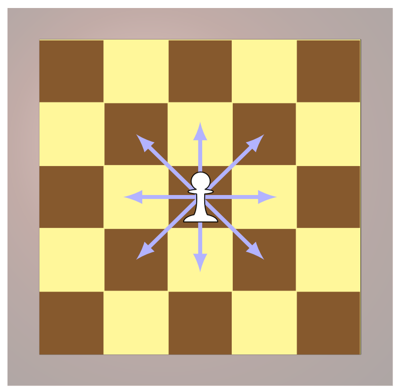

Abbildung 1: Irrfahrt auf einem 5×5-Schachfeld

Es wird sich herausstellen, dass die Wahrscheinlichkeit für ein schwarzes Randfeld 'nur' rund 0.4857 beträgt, obschon es mehr schwarze als weisse Randfelder gibt.

Alleine anhand von Plausibilitätsbetrachtungen kommt man kaum zu diesem Ergebnis. Im Unterricht bietet es sich an, die Schüler:innen mit einem Zufallsgenerator viele Irrfahrten durchführen zu lassen und so schnell zu einer Stichprobe vom Umfang $n \approx 200$ zu kommen. Stärkere Aussagen, als dass $p_{\text{schwarz}} \approx 0.5$ ist, werden sich allerdings aus dieser Stichprobe wohl kaum ableiten lassen.

Ein systematischer Zugang ist daher unabdingbar.

2 Markow-Ketten im Mathematikunterricht

Markow-Ketten (hier synonym für *diskrete Markow-Prozesse*) sind eine spezielle Klasse von stochastischen Prozessen. Erste stochastische Prozesse werden im gymnasialen Mathematikunterricht in der Regel im Rahmen der elementaren Wahrscheinlichkeitsrechnung behandelt – ohne dass diese als solche bezeichnet werden. Eine typische Aufgabenstellung lautet: Ariane und Bernd werfen abwechselnd Steinchen auf eine Blechdose. Es gewinnt, wer es als Erste(r) schafft, die Blechdose umzuwerfen. Ariane wirft die Dose bei jedem Wurf mit einer Wahrscheinlichkeit von 0.45 um, Bernd mit einer Wahrscheinlichkeit von 0.6. Ariane beginnt. Gesucht ist die Wahrscheinlichkeit p, dass Ariane gewinnt.

Üblicherweise werden solche Aufgaben mit einem Baumdiagramm untersucht. Die gesuchte Wahrscheinlichkeit ergibt sich als unendliche geometrische Reihe. In unserem Fall etwa ist

$$p = 0.45 \cdot \sum_{k=0}^{\infty} (0.55 \cdot 0.4)^k = 0.45 \cdot \frac{1}{1 - 0.55 \cdot 0.4} = 0.577.$$

Das Spiel lässt sich auch als Markow-Kette mit vier Zuständen auffassen.

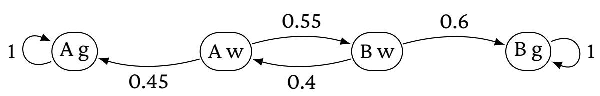

Abbildung 2: Prozessdiagramm

Der Prozess hat die beiden *transienten Zustände* 'Ariane wirft als Nächste' (A w) und Bernd wirft als Nächster' (B w) und die beiden *absorbierenden Zustände* 'Ariane hat gewonnen' (A g) und 'Bernd hat gewonnen' (B g). Transient bedeutet, dass die Zustände mit Wahrscheinlichkeit 1 nur endlich oft besucht werden. Das Prozessdiagramm (Abbildung 2) ist ein endlicher gerichteter und gewichteter Graph. Die Kantengewichte entsprechen den Übergangswahrscheinlichkeiten. Die Adjazenzmatrix des Graphen ist die *Übergangsmatrix* des Prozesses:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.45 & 0 & 0.55 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Die Übergangsmatrix ist eine *stochastische Matrix*, d.h. die Matrix ist quadratisch, die Einträge sind allesamt nichtnegative reelle Zahlen und die Zeilensummen sind alle 1. Diverse Fragestellungen lassen sich mit Hilfe der Übergangsmatrix diskutieren. Allerdings zeigt die Berechnung der Wahrscheinlichkeit p mit Hilfe einer geometrischen Reihe, dass das Beispiel ohne Theorie über Markow-Ketten behandelt werden kann.

Unser Schachbrettbeispiel ist zwar ähnlich einfach zugänglich, lässt sich aber nicht mehr mittels elementarer Techniken wie Baumdiagrammen oder geometrischen Reihen behandeln, wie im folgenden Abschnitt klar werden wird.

3 Modellierung des Problems mit einer Markow-Kette

Wir modellieren die Irrfahrt auf dem 5×5-Schachbrett mit einem diskreten stochastischen Prozess. Aus Symmetriegründen können die 25 Felder auf 5 Zustände reduziert werden (vgl. Abbildung 3).

Bei den Zuständen 4 und 5 endet die Irrfahrt. Solche Zustände heissen *absorbierend*. Die Zustände 1 bis 3 sind *transiente Zustände*. Gesucht ist die Absorptionswahrscheinlichkeit im Zustand 5 beim Start im Zustand 1.

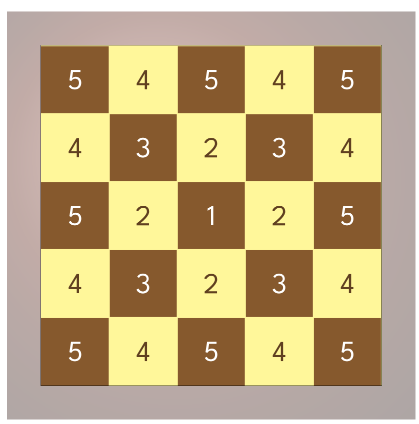

Abbildung 3: Zustände der Markow-Kette

3.1 Vom Prozessdiagramm zur Lösung des Problems mit einem Gleichungssystem

Abbildung 4 zeigt das Prozessdiagramm der Markow-Kette. Die fünf Zustände der Markow-Kette bilden die fünf Knoten des Prozessdiagramms. Die Gewichte der gerichteten Kanten sind durch die Übergangswahrscheinlichkeiten gegeben:

Befindet sich die Spielfigur beispielsweise im Zustand 3, so wird sie im folgenden Schritt mit Wahrscheinlichkeit $\frac{3}{8} = 0.375$ im Zustand 5 landen – daher hat die Kante $3 \to 5$ das Gewicht 0.375.

Der Ansatz, aus dem Prozessdiagramm ein Baumdiagramm zu erstellen und die gesuchte Wahrscheinlichkeit mit Hilfe einer unendlichen Reihe zu berechnen, ist nicht mehr zielführend. Davon kann man sich leicht selbst überzeugen.

Bezeichnen wir die Wahrscheinlichkeit, dass wir ausgehend vom inneren Zustand i im Zustand 5 absorbiert werden, mit a_i, so erhalten wir die Gleichungen

$$
\begin{aligned}
a_1 &= 0.5a_2 + 0.5a_3 \\
a_2 &= 0.125a_1 + 0.25a_2 + 0.25a_3 + 0.125 \\
a_3 &= 0.125a_1 + 0.25a_2 + 0.375
\end{aligned}
$$

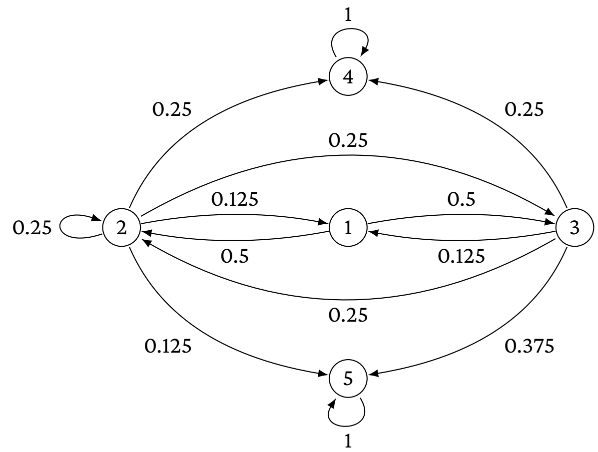

Abbildung 4: Prozessdiagramm

Die dritte Gleichung z.B. lässt sich wie folgt begründen: Sind wir im Zustand 3, so gelangen wir mit Wahrscheinlichkeit 0.375 im nächsten Schritt in den Zustand 5 (daher der Summand 0.375). Mit Wahrscheinlichkeit 0.125 aber gelangen wir in den Zustand 1 und werden anschliessend von dort mit a_1 im Zustand 5 absorbiert (daher der Summand $0.125a_1$). Schliesslich gelangen wir mit Wahrscheinlichkeit 0.25 in den Zustand 2, daher der Summand $0.25a_2$.

Nach kurzer Umformung erhalten wir die Matrix-Vektor-Gleichung

$$
\begin{pmatrix} 1 & -0.5 & -0.5 \\ -0.125 & 0.75 & -0.25 \\ -0.125 & -0.25 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0.125 \\ 0.375 \end{pmatrix} \tag{1}
$$

mit der Lösung $a_1 = 0.4857, a_2 = 0.4286, a_3 = 0.5429$. Starten wir also in der Mitte bzw. im Zustand 1, so endet die Irrfahrt mit einer Wahrscheinlichkeit von 48.57% auf einem schwarzen Feld (und mit Wahrscheinlichkeit 51.43% auf einem weissen Feld).

3.2 Lösung des Problems mit der Fundamentalmatrix

Die in 3.1 diskutierte Lösung erlaubt Querbezüge zur Graphentheorie und greift auf lineare Gleichungssysteme zurück. Mit etwas fortgeschritteren Methoden der linearen Algebra lässt sich das Problem systematischer angehen und dient als gute Grundlage für weitere Fragestellungen.

Dazu beschreiben wir die Markow-Kette mit ihrer Übergangsmatrix:

$$
P = \begin{pmatrix} 0 & 0.5 & 0.5 & 0 & 0 \\ 0.125 & 0.25 & 0.25 & 0.25 & 0.125 \\ 0.125 & 0.25 & 0 & 0.25 & 0.375 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} Q & R \\ 0 & E \end{pmatrix} \tag{2}
$$

Die blaue Teilmatrix Q beschreibt die Übergangsmatrix des transienten Teils der Kette, die rote Teilmatrix R die Absorbtionswahrscheinlichkeiten in den absorbierenden Zuständen. Unten rechts steht (grün) die Einheitsmatrix, unten links die Nullmatrix. Der Eintrag P_{ij} ist gleich der Wahrscheinlichkeit, in einem Zeitschritt vom Zustand i in den Zustand j zu gelangen.

Wir sehen, dass die 3×3-Matrix A in der Matrix-Vektor-Gleichung (1) entsteht, wenn wir die Matrix Q von der Einheitsmatrix E subtrahieren: $A = E - Q$ bzw.

$$\begin{pmatrix} 1 & -0.5 & -0.5 \\ -0.125 & 0.75 & -0.25 \\ -0.125 & -0.25 & 1 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} - \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.125 & 0.25 & 0.25 \\ 0.125 & 0.25 & 0 \end{pmatrix}$$

Die Inverse dieser Matrix, $F = A^{-1} = (E - Q)^{-1}$ heisst *Fundamentalmatrix* der Markow-Kette. Im Beispiel ist

$$F = \begin{pmatrix} 1.2571 & 1.1429 & 0.9143 \\ 0.2857 & 1.7143 & 0.5714 \\ 0.2286 & 0.5714 & 1.2571 \end{pmatrix}.$$

Es kann gezeigt werden, dass die formal richtige Identität

$$F = (E - Q)^{-1} = E + Q + Q^2 + Q^3 + \ldots$$

tatsächlich gilt. (Die Einträge der Matrizen Q^k streben für $k \to \infty$ gegen 0, weil die Matrix Q nur transiente Zustände beschreibt. Dass die Reihe tatsächlich konvergiert, ist damit natürlich noch nicht gezeigt.) Der Eintrag F_{ij} ist gleich dem Erwartungswert für die Anzahl Besuche im Zustand j bis zur Absorption bei Start im Zustand i.

Auch der im Gleichungssystem (1) vorkommende Vektor $\begin{pmatrix} 0 & 0.125 & 0.375 \end{pmatrix}^T$ erscheint in der Übergangsmatrix P. Es ist die zum absorbierenden Zustand 5 gehörende Spalte der Matrix R. Bezeichnen wir ihn mit $\vec{b}$, so erhalten wir die gesuchten Absorbtionswahrscheinlichkeiten $\vec{a}$ als Lösung der Gleichung

$$A \cdot \vec{a} = (E - Q) \cdot \vec{a} = \vec{b}.$$

Umgeformt:

$$\vec{a} = (E - Q)^{-1} \cdot \vec{b} = F \cdot \vec{b}.$$

Diese Tatsache ist als *Erste Mittelwertsregel* bekannt:

Erste Mittelpunktsregel: Eine Markow-Kette habe die Fundamentalmatrix F. Weiter beschreibe der Vektor $\vec{b}$ die Wahrscheinlichkeiten, mit welchen man von einem inneren Zustand direkt in den absorbierenden Zustand i übergeht. Dann stehen im Vektor

$$\vec{a} = F \cdot \vec{b}$$

die Wahrscheinlichkeiten, dass man von den inneren Zuständen im Zustand i absorbiert wird. In unserem Beispiel erhalten wir

$$\vec{a} = \begin{pmatrix} 0.4857 \\ 0.4286 \\ 0.5429 \end{pmatrix}.$$

Der erste Eintrag des Vektors $\vec{a}$ ist die gesuchte Wahrscheinlichkeit, dass die Irrfahrt ausgehend von Feld 1 auf einem schwarzen Randfeld endet.

Interessant ist auch die Feststellung, dass

$$\vec{a} = F \cdot \vec{b} = (E + Q + Q^2 + \ldots) \cdot \vec{b}.$$

4 Ausblick

Etwas alltagsnäher ist dieselbe Fragestellung auf einem Schachbrett der ‘richtigen’ Grösse 8×8. Das oben beschriebene Vorgehen lässt sich übertragen. Es braucht hier allerdings 14 Zustände (Abbildung 5).

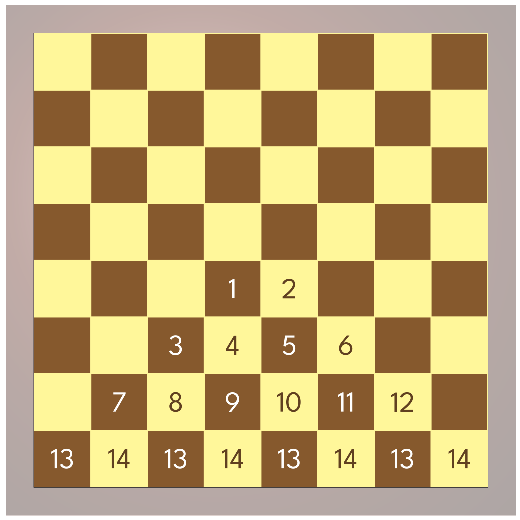

Abbildung 5: Zustände der Markow-Kette beim richtigen Schachbrett

Das Aufstellen der Fundamentalmatrix ist eine Konzentrations- und Fleissaufgabe, ist aber noch machbar. Man findet beispielsweise, dass die Wahrscheinlichkeit, auf einem schwarzen Randfeld absorbiert zu werden, bei einem Start auf Feld 1 gleich 0.5002 ist. Mit einem Wert von 0.5582 am grössten ist diese Wahrscheinlichkeit, wenn auf Feld 11 gestartet wird.

Die hier diskutierten Beispiele lassen sich beliebig modifizieren: Es können z.B. frei nach Leonard Euler Rösselsprünge auf dem Schachbrett untersucht werden. Eine weitere mögliche Anwendung ist das Bestimmen *mittlerer Wartezeiten* (Zeitschritte bis zur Absorption auf einem Randfeld).

Zusammenfassend halten wir fest:

- Einfach verständliche Beispiele bieten verschiedene Zugänge und laden zum entdeckenden Lernen ein.
- Querbezüge zur Wahrscheinlichkeitsrechnung, zu geometrischen Reihen oder zur Graphentheorie ergeben sich auf natürliche Weise.
- Die vorkommenden Gleichungssysteme sind allesamt linear. Das Beschreiben der Gleichungen mit Matrizen erweist sich als vorteilhaft: Erst ein systematischer Zugang mit Matrizen führt zur Möglichkeit, Beispiele ohne aufwändigen Umweg über das Prozessdiagramm zu erfassen.
- Weitere Anwendungen der linearen Algebra ergeben sich (z.B. Adjazenzmatrizen von Graphen oder Eigenwerte stochastischer Matrizen).
- Markow-Ketten lassen sich mit Computereinsatz programmiertechnisch problemlos mit numerischen Simulationen untersuchen. Entweder kann dazu die Übergangsmatrix wiederholt auf den Ausgangszustand angewendet werden (in vielen Lehrbüchern als Iterationsmethode beschrieben) oder es lassen sich zahlreiche Irrfahrten simulieren.

In diesem Sinn möchte der Artikel als Plädoyer für den Einsatz von Markow-Ketten im Schwerpunktfachunterricht dienen.

Links

[1] *Hans Ulrich Keller*, Markow-Ketten, Bulletin des VSMP, No 121, Januar 20213, S. 23 – 24

[2] *Themenheft Matrizen* (wird aufs Schuljahr 2023/24 erscheinen, dmk.vsmp.ch)

[3] *Lambacher Schweizer*, Mathematik – Stochastik, Ernst Klett Verlag GmbH, Stuttgart 2012