

■ SEKUNDARSTUFE II
11.–13. Schuljahr

Der Palio, das Pferde- rennen von Siena

Bei jedem „Palio“ müssen 10 Teilnehmer aus 17 Kandidaten ausgewählt werden.

Wir modellieren Auswahlprozesse und erschließen so das Thema „Markoff-Ketten“.

Hans Humenberger

geb. 1963, Privatdozent an der Universität Dortmund
Institut für Entwicklung und Erforschung des Mathe-
matikunterrichts
E-Mail: hans.humenberger@mathematik.uni-
dortmund.de



Der Palio di Siena

Foto: dpa

In Siena (Toskana, Norditalien) findet jährlich ein sehr berühmtes Pferderennen mit einer schon jahrhundertelangen Tradition, der so genannte „Palio di Siena“ (kurz: Palio) statt: Ein waghalsiges Rennen auf ungesattelten Pferden rund um die *Piazza del Campo*; es dauert jeweils nur etwa $1\frac{1}{2}$ Minuten, wobei allerdings dieses Rennen von zahlreichen traditionellen Zeremonien tagelang umrahmt wird.

Vorformen des Palio reichen bis ins 11. Jh. zurück; etwa seit Mitte des 17. Jh. findet er in einer zur heutigen ähnlichen Form statt (jeweils am 2. Juli), seit 1802 sogar zweimal pro Jahr (2. Juli und 16. August). Dabei stellen die Juli-Rennen und die August-Rennen zwei eigenständige „Schienen“ dar. Das Juli- und das August-Rennen des selben Jahres hängen nicht zusammen, wohl aber die einzelnen Juli-Rennen verschiedener Jahre sowie die August-Rennen verschiedener Jahre.

Seit 1729 gibt es in Siena 17 Stadtbezirke (Contraden), und jeder Stadtbezirk wird durch ein Gespann Reiter-Pferd vertreten. Doch dabei ergibt sich eine gewisse „Misere“: Die Platzverhältnisse auf der genau festgelegten „Renn-Strecke“ im alten Siena sind so eng, dass es zu gefährlich wäre, das Rennen mit 17 Pferden zu bestreiten. Man hat sich schon seit sehr langer Zeit darauf geeinigt, bei jedem Palio nur 10 Reitergespanne gegeneinander antreten zu lassen (und damit nur 10 Contraden).

Damit erhebt sich natürlich die Frage, welche 10 der 17 Contraden je ein Reiter-Pferd-Gespann zum Palio

entsenden dürfen. Es ist also ein gerechtes Auswahlverfahren nötig, das festlegt, welche Contraden die Chance bekommen, die beliebte Trophäe zu erobern. Diese Trophäe ist ein Banner, das früher aus wertvollem roten Samt gemacht wurde und „Pallium“ hieß, woraus der Name „Palio“ für das Rennen resultiert. Genaueres zur Geschichte und zum ganzen Fest (Rennen) findet man zum Beispiel im Internet unter <http://www.comune.siena.it>

Modelle für die Auswahl

Schülerinnen und Schüler werden hier schnell an ihre Kinderzeit erinnert. Dort gab es sicher oft Situationen, in denen eine bestimmte Anzahl von Kindern vorhanden war, aber ein geplantes Spiel nur weniger Teilnehmer zuließ: Man brauchte ein geeignetes Auswahl- bzw. Auslosungsverfahren, wer denn nun zu den (glücklichen) Teilnehmern und wer zunächst zum Zuschauer verurteilt war.

Mache konkrete Vorschläge für ein Auswahlverfahren und berechne die Wahrscheinlichkeit, mit der eine bestimmte Contrade am Palio teilnimmt.

Die Aufgabe fordert dazu auf, den Wahlvorgang zu *modellieren*, und eignet sich sehr gut für eine Gruppenarbeit. Zwei Lösungsvorschläge:

■ Die 17 Contraden „liegen“ in einer Urne – 10 einzelne Ziehungen ohne Zurücklegen oder eine Ziehung von 10 Contraden mit einem Griff:

17 Kugeln liegen in einer Urne; sie enthalten zusammengerollte Zettel als „Innenleben“, auf denen die Namen der 17 Contraden stehen, darunter unsere Heimat-Contrade C. Mit welcher Wahrscheinlichkeit nehmen wir am Palio teil? Man erhält:

$$P(C \text{ wird gezogen}) = P(C \text{ nimmt teil}) = \frac{10}{17}.$$

■ **Urne enthält „Lose“:**

Eine andere Möglichkeit der Auslosung wäre: In einer Urne sind wieder 17 Kugeln. 10 davon enthalten einen Zettel mit der Aufschrift „Teilnahme“ (T) und 7 einen Zettel mit „Nicht-Teilnahme“ (-T), also 10 Gewinnlose und 7 Nieten. Jede Contrade zieht eine Kugel (T oder -T) ohne Zurücklegen aus der Urne. Eine kurze stochastische Analyse zeigt, dass die Reihenfolge der Ziehungen keine Rolle spielt, dass bei jeder Ziehung die Chance (die a-priori-Wahrscheinlichkeit), ein T zu erwischen, jeweils $\frac{10}{17}$ beträgt:

$$P(C \text{ zieht } T) = P(C \text{ nimmt teil}) = \frac{10}{17}.$$

Die beiden bis jetzt beschriebenen Modelle des Auswählens haben gemeinsam:

1. Für jede Contrade beträgt die Teilnahmewahrscheinlichkeit $\frac{10}{17}$ – das Wahlverfahren ist in diesem Sinn also gerecht, keine Contrade wird bevorzugt.
2. Die Auslosung geschieht jedes Mal „neu“ unter den selben Bedingungen: Die Konstellationen des Vorjahres – ob eine Contrade am Palio teilgenommen hat oder nicht; ob sie gewonnen oder verloren hat – beeinflussen die Teilnahmewahrscheinlichkeit einer Contrade am diesjährigen Palio nicht. Jede Contrade stellt sich jedes Jahr der Wahl „neu“ (Unabhängigkeit der einzelnen Wahlvorgänge; Bernoulli-Kette).

Markoff-Ketten

Die beschriebene Unabhängigkeit der jährlichen Auswahlvorgänge ist ein entscheidender Nachteil in den obigen Modellen: Wenn jedes Jahr nach dem Prinzip „neues Jahr, neues Rennen, neues Glück“ verfahren würde, bestünde durchaus die Möglichkeit, dass eine bestimmte Contrade über mehrere Jahre am Palio nicht teilnehmen könnte. Was natür-

lich zu einer sehr hohen Frustration führen würde.

Welche Möglichkeiten gibt es, diese für alle beteiligten Contraden unerwünschte Frustrationsmöglichkeit zu beseitigen?

Hier werden die Schülerinnen und Schüler schnell nahe liegende Modelle vorschlagen: Man könnte die 7 Nichtteilnehmer eines Rennens zu Fixstärtern für das nächste Rennen erklären. Wie könnten dann die restlichen 3 Startplätze für den nächsten Palio gerecht vergeben werden? Auf diese Frage könnten verschiedene Vorschläge kommen:

- Die Contraden auf den *ersten* 3 Plätzen eines Palio sollen auch im nächsten starten dürfen: *Belohnung* für gute Leistung. Oder: Die Contraden auf den *letzten* 3 Plätzen eines Palio sollen auch im nächsten starten dürfen: *Trost* für verpatzte Rennen.
- Aus den 10 Teilnehmern eines Palio sollen 3 gewählt werden, die auch am nächsten teilnehmen dürfen. Dies kann nach einem der vorhin beschriebenen Verfahren geschehen: jede an einem Palio teilnehmende Contrade hätte dann eine 3/10-Chance, auch beim nächsten Mal dabei zu sein.

Dieses letzte Modell wird in Siena gewählt! In dieser Weise hängen die Juli-Rennen untereinander und die August-Rennen untereinander zusammen. Die Juli- und die August-Rennen sind aber zwei voneinander unabhängige „Stränge“ (für eine eher hochschulmathematische Analyse siehe Götz/Grosser 1999).

Dadurch wird erreicht, dass keine Contrade zweimal hintereinander pausieren muss: nach einer Nicht-Teilnahme folgt mit Sicherheit eine

Teilnahme; nach einer Teilnahme folgt mit Wahrscheinlichkeit 3/10 eine erneute Teilnahme.

Dieses Auswahlverfahren ist keine Bernoulli-Kette mehr, sondern eine so genannte *Markoff-Kette*. Die Teilnahmewahrscheinlichkeit einer Contrade C am Rennen $n + 1$ hängt (nur) davon ab, ob C beim Rennen n dabei war oder nicht.

Ein Großteil des Folgenden dient nun der Beantwortung der Frage:

Ist die Palio-Auswahl „gerecht“? Mit welcher Wahrscheinlichkeit nehmen wir (die Contrade C) am Palio nach 1 (3, 10, 100, ...) Jahren teil?

Dabei liegt es nahe, unter „gerecht“ zu verstehen, dass sich auf lange Sicht, das heißt für große n , für jede Contrade die selbe Teilnahmewahrscheinlichkeit ergibt, und zwar unabhängig davon, ob die Contrade beim ersten Rennen teilgenommen hat oder nicht. Andernfalls bedeutete eine (Nicht-) Teilnahme von C beim ersten Rennen eine Verbesserung (bzw. Verschlechterung) ihrer Teilnahmewahrscheinlichkeit „auf ewig“, was nicht wünschenswert wäre. Von so einer gerechten Teilnahmewahrscheinlichkeit auf lange Sicht erwarten wir natürlich, dass sie den Wert $\frac{10}{17}$ hat. Dies folgt auch aus der Überlegung über die relative Teilnahmehäufigkeit bei den ersten n Rennen: Bei n Rennen (n groß) gibt es insgesamt $10 \cdot n$ Rennteilnahmen, so dass im „gerechten“ Fall bei 17 Konkurrenten jedem ein Anteil von $\frac{10}{17} \cdot n$ zusteht.

Zunächst betrachten wir Zufallsgrößen

$$X_n = \begin{cases} T, & \text{falls } C \text{ am Rennen } n \\ & \text{teilnimmt} \\ -T, & \text{falls } C \text{ am Rennen } n \\ & \text{nicht teilnimmt} \end{cases}$$

Das Rennen auf ungesattelten Pferden ist nicht ungefährlich



Foto: dpa

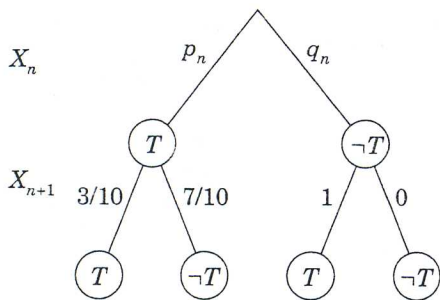


Abb. 1: 2-stufiges Baumdiagramm für die Zufallsgrößen X_n bzw. X_{n+1}

und interessieren uns für die Wahrscheinlichkeitsverteilungen von X_n . Mit der Abkürzung

$$p_n := P(X_n = T),$$

$$q_n := 1 - p_n = P(X_n = -T)$$

interessiert uns die Folge $(p_n)_{n \in \mathbb{N}}$. Damit ist auch die Folge $(q_n)_{n \in \mathbb{N}}$ festgelegt.

Für große n sollte $p_n \approx \frac{10}{17}$ sein, und zwar unabhängig davon, ob C beim Anfangsrennen dabei war oder nicht.

So ohne weiteres können wir über die Werte p_n und q_n keine direkten Aussagen machen, hängen sie doch davon ab, ob C beim Rennen davor teilgenommen hat oder nicht. Der „Mechanismus“ dieser Abhängigkeit von einem Rennen n zum nächsten Rennen $n + 1$ ist uns bekannt und für alle n der selbe; er ist im Baumdiagramm in **Abbildung 1** dargestellt.

Die Pfadregeln ergeben:

$$p_{n+1} = \frac{3}{10} \cdot p_n + 1 \cdot q_n \quad (1)$$

$$q_{n+1} = \frac{7}{10} \cdot p_n + 0 \cdot q_n \quad (2)$$

Mithilfe dieser beiden Iterationsformeln können wir nun ausgehend von einem bestimmten Anfangsrennen, an dem wir teilgenommen haben (oder auch nicht: je nachdem ist dann $p_0 = 1$ oder $p_0 = 0$) schrittweise die Teilnahmewahrscheinlichkeit p_i und q_i konkret berechnen:

$$(p_0, q_0) \rightarrow (p_1, q_1) \rightarrow (p_2, q_2) \rightarrow \dots$$

Dies sollten die Schülerinnen und Schüler (noch ohne Computer) bis zum Index 3 oder 4 wirklich ausführen. Sie merken schnell, dass hierbei wiederholt das selbe zu tun ist. Das ist immer ein Anzeichen dafür, dass Computer gut einsetzbar sein müssten.

Das obige Gleichungssystem kann man auch mittels einer Matrix schreiben:

$$\begin{pmatrix} \frac{3}{10} & 1 \\ \frac{7}{10} & 0 \end{pmatrix} \cdot \begin{pmatrix} p_n \\ q_n \end{pmatrix} = \begin{pmatrix} p_{n+1} \\ q_{n+1} \end{pmatrix}$$

$:= A \quad := \vec{\pi}_n \quad \vec{\pi}_{n+1}$

Nun werden die Verbindungen zur Linearen Algebra schon deutlich. Die Matrizenrechnung wird ab nun vorausgesetzt, obwohl sie nicht an jeder Oberstufe unterrichtet wird; es wäre auch möglich, sie an dieser Stelle einzuführen.

$$\text{Der (Spalten-)Vektor } \vec{\pi}_n := \begin{pmatrix} p_n \\ q_n \end{pmatrix}$$

beschreibt offenbar die Teilnahmewahrscheinlichkeit der Contrade beim n -ten Rennen: Er beschreibt die Wahrscheinlichkeitsverteilung der Zufallsgröße X_n (die ja nur zwei mögliche Ausprägungen hat). Der Vektor $\vec{\pi}_n$ heißt deshalb auch *Wahrscheinlichkeitsvektor zum Zeitpunkt n* oder *Verteilung zum Zeitpunkt n* .

Multipliziert man ihn von links mit der Matrix A , so erhält man den Vektor $\vec{\pi}_{n+1}$ der die Teilnahmewahrscheinlichkeit am $(n + 1)$ -ten Rennen beschreibt.

Die Matrix A sorgt für den Vektorübergang $\vec{\pi}_n \rightarrow \vec{\pi}_{n+1}$ und heißt deshalb „Übergangsmatrix“.

Was hat man von dieser Erkenntnis, dass das Gleichungssystem in Matrixform geschrieben werden kann? Es eröffnet eine Möglichkeit,

ausgehend von $\vec{\pi}_0 = \begin{pmatrix} p_0 \\ q_0 \end{pmatrix}$ direkt zu

$$\vec{\pi}_n = \begin{pmatrix} p_n \\ q_n \end{pmatrix} \text{ zu kommen, ohne } n\text{-mal}$$

die mühselige Iteration durchführen zu müssen:

$$\vec{\pi}_1 = A \cdot \vec{\pi}_0$$

$$\vec{\pi}_2 = A \cdot \vec{\pi}_1 = A \cdot (A \cdot \vec{\pi}_0) = A^2 \cdot \vec{\pi}_0$$

$$\vec{\pi}_3 = A \cdot \vec{\pi}_2 = A \cdot (A^2 \cdot \vec{\pi}_0) = A^3 \cdot \vec{\pi}_0$$

:

$$\vec{\pi}_n = A \cdot \vec{\pi}_{n-1} = \dots = A^n \cdot \vec{\pi}_0$$

Um $\vec{\pi}_n$ zu erhalten, muss man also „nur“ A^n bilden und mit $\vec{\pi}_0$ multiplizieren.

Computereinsatz

Diese Arbeit wird vom Computer vermutlich iterativ (schrittweise) erledigt. Durch die Matrixdarstellung wird also nicht notwendig Rechenarbeit erspart, sondern nur eine CAS-taugliche Form gewählt, so dass das Mühselige an den Computer übertragen werden kann, ohne selbst ein Programm dafür schreiben zu müssen.

Man erlebt die Sinnhaftigkeit und Kompaktheit der Darstellungsform

„Matrix“, die Bedeutung der Matrizenmultiplikation, des Potenzierens von Matrizen, der Multiplikation einer Matrix mit einem Vektor und der Assoziativität. Ohne Computer kann man natürlich keine hohen Matrixpotenzen berechnen, sondern nur solche von sehr niedriger Ordnung; mit einem CAS und der Möglichkeit, Matrizen in Sekundenbruchteilen zu potenzieren, steht hingegen einer weitgehend selbstständigen Entdeckung einiger Phänomene nichts im Wege.

Zu unserer Contrade

Wie groß ist die Wahrscheinlichkeit, dass sie in 3 Jahren teilnehmen/nicht teilnehmen wird? Wir brauchen

$$\text{dafür } A^3 \approx \begin{pmatrix} 0.45 & 0.79 \\ 0.55 & 0.21 \end{pmatrix}$$



Sie habe im Palio dieses Jahres teilgenommen, das heißt

$$\vec{\pi}_0 = \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \text{ Dann ist}$$

$$\vec{\pi}_3 = \begin{pmatrix} p_3 \\ q_3 \end{pmatrix} = A^3 \cdot \vec{\pi}_0$$

$$\approx \begin{pmatrix} 0.45 & 0.79 \\ 0.55 & 0.21 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.45 \\ 0.55 \end{pmatrix}$$

Die gesuchte Teilnahmewahrscheinlichkeit beträgt also $p_3 \approx 0.45$.

Welche Werte haben p_3 und q_3 , wenn die Contrade in diesem Jahr nicht teilgenommen hat, also bei

$$\vec{\pi}_0 = \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} ? \text{ Es ist}$$

$$\vec{\pi}_3 = \begin{pmatrix} p_3 \\ q_3 \end{pmatrix} \approx \begin{pmatrix} 0.45 & 0.79 \\ 0.55 & 0.21 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.79 \\ 0.21 \end{pmatrix}$$

Die Teilnahmewahrscheinlichkeit beträgt hier $p_3 \approx 0.79$ und unterscheidet sich von obiger doch beträchtlich. Die diesjährige Teilnahme oder Nichtteilnahme wirkt sich also noch stark auf die Teilnahmechancen in drei Jahren aus.

Die selben Fragen für 10 Jahre sind nur noch mit CAS sinnvoll zu bearbeiten:

$$A^{10} \approx \begin{pmatrix} 0.60 & 0.57 \\ 0.40 & 0.43 \end{pmatrix}.$$

Bei Teilnahme in diesem Jahr folgt

$$\begin{aligned} \bar{\pi}_{10} &= \begin{pmatrix} p_{10} \\ q_{10} \end{pmatrix} \approx \begin{pmatrix} 0.60 & 0.57 \\ 0.40 & 0.43 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0.60 \\ 0.40 \end{pmatrix}. \end{aligned}$$



Foto: dpa

Festliche Umzüge umrahmen das Rennen

Keine Teilnahme in diesem Jahr liefert

$$\begin{aligned} \bar{\pi}_{10} &= \begin{pmatrix} p_{10} \\ q_{10} \end{pmatrix} \approx \begin{pmatrix} 0.60 & 0.57 \\ 0.40 & 0.43 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 0.57 \\ 0.43 \end{pmatrix}. \end{aligned}$$

Nach 10 Jahren unterscheiden sich die beiden Teilnahmewahrscheinlichkeiten p_{10} und die beiden Wahrscheinlichkeitsvektoren $\bar{\pi}_{10}$ in Bezug auf diesjährige Teilnahme/Nicht-Teilnahme nur mehr wenig: $p_{10} \approx 0.60$ und $p_{10} \approx 0.57$.

Wird dieser Unterschied mit wachsendem n tatsächlich immer geringer? Um diese Vermutung experimentell zu bestätigen, wählen wir bei größeren n vorsichtshalber eine höhere Genauigkeit, zum Beispiel die

CAS-StandardEinstellung mit 6 signifikanten Ziffern.

Für $n = 100$ Jahre folgt:

$$\begin{aligned} A^{100} &\approx \begin{pmatrix} 0.588235 & 0.588235 \\ 0.411765 & 0.411765 \end{pmatrix} \\ \bar{\pi}_{100} &= \begin{pmatrix} p_{100} \\ q_{100} \end{pmatrix} = A^{100} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \approx \begin{pmatrix} 0.588235 \\ 0.411765 \end{pmatrix}. \end{aligned}$$

Nach 100 Jahren unterscheiden sich die beiden Wahrscheinlichkeitsvektoren $\bar{\pi}_{100}$ praktisch nicht mehr!

Auch bei jedem anderen „Wahrscheinlichkeits-Startvektor“

$$\bar{\pi}_0 = \begin{pmatrix} p_0 \\ q_0 = 1 - p_0 \end{pmatrix}, \text{ der die Anfangs-}$$

Teilnahmewahrscheinlichkeit beschreibt, hätte sich der selbe Vektor $\bar{\pi}_{100}$ ergeben, weil jede Zeile in A^{100} aus fast identischen Werten besteht und $q_0 = 1 - p_0$ ist:

$$\begin{aligned} \bar{\pi}_{100} &= \begin{pmatrix} p_{100} \\ q_{100} \end{pmatrix} = A^{100} \cdot \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} \\ &\approx \begin{pmatrix} 0.588235 \\ 0.411765 \end{pmatrix}. \end{aligned}$$

Berechnet man auch A^{101} und $\bar{\pi}_{101}$, so sieht man, dass sich A^{100} von A^{101} und $\bar{\pi}_{100}$ von $\bar{\pi}_{101}$ bei 6 signifikanten Ziffern nicht mehr unterscheiden.

Der experimentelle Befund lässt vermuten, dass die Folge der Matrixpotenzen (A^n) gegen eine Grenzmatrix G konvergiert, deren Eigenschaft man in $A^{100} \approx G$ schon ablesen kann:

Die Zeilen von G haben jeweils gleiche Werte beziehungsweise G hat identische Spalten.

Wenn dies richtig ist, dann konvergieren auch die Wahrscheinlichkeitsvektoren $\bar{\pi}_n$ gegen einen eindeutig bestimmten Grenzvektor $\bar{\pi} \approx \begin{pmatrix} 0.588235 \\ 0.411765 \end{pmatrix}$, der nichts anderes

ist als eine Spalte von G . Diese Konvergenz $\bar{\pi}_n \rightarrow \bar{\pi}$ ist dann sogar unabhängig vom Startvektor $\bar{\pi}_0$.

Diese spezielle Konvergenz können Schülerinnen und Schüler einfach durch Probieren am Computer selbstständig entdecken. Es bleibt die Frage, was es mit den Werten 0.588235 und 0.411765 auf sich hat.

Auf lange Sicht hatten wir doch, bei Gerechtigkeit des Auswahlverfahrens, eine Teilnahmewahrschein-

lichkeit von $\frac{10}{17}$ erwartet; tatsächlich gilt:

$$\frac{10}{17} \approx 0.588235 \text{ und } \frac{7}{17} \approx 0.411765.$$

Dadurch drängt sich natürlich sofort die Vermutung auf, dass diese Werte des Grenzvektors $\bar{\pi} \approx \bar{\pi}_{100}$ vielleicht auch einfacher zu bekommen sein müssten, ohne die Grenzmatrix $G \approx A^{100}$ zu berechnen.

Allein die gesicherte Existenz einer Grenzmatrix $G = \lim_{n \rightarrow \infty} A^n$ mit

Zeilen aus jeweils gleichen Werten, ohne sie konkret zu kennen, hat entscheidende Konsequenzen. Deswegen drängt sich die Frage nach Bedingungen dafür ja geradezu auf:

Wenn es nämlich eine solche Grenzmatrix G gibt, so gibt es klarerweise auch den Grenzvektor $\bar{\pi}$ (eindeutig und unabhängig von $\bar{\pi}_0$). In

Analogie zu den Grenzwert-Sätzen im Eindimensionalen erhalten wir ($\bar{\pi}_0$ ist ein konstanter Vektor!):

$$\begin{aligned} \bar{\pi} &= \lim_{n \rightarrow \infty} (A^n \cdot \bar{\pi}_0) = \left(\lim_{n \rightarrow \infty} A^n \right) \cdot \bar{\pi}_0 \\ &= G \cdot \bar{\pi}_0. \end{aligned}$$

Dieser Grenzvektor $\bar{\pi}$ darf sich durch Multiplikation mit A natürlich nicht mehr verändern, es muss gelten: $A \cdot \bar{\pi} = \bar{\pi}$; $\bar{\pi}$ muss also Fixvektor von A sein.

Dies führt zur Fixvektorgleichung

$$\begin{pmatrix} \frac{3}{10} & 1 \\ \frac{7}{10} & 0 \end{pmatrix} \cdot \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix},$$

wobei für die Wahrscheinlichkeiten $p, q \geq 0$ noch $p + q = 1$ gelten muss. Die leicht verifizierbare Lösung ist

$$\bar{\pi} = \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \frac{10}{17} \\ \frac{7}{17} \end{pmatrix}.$$

In diesem Zusammenhang geben wir die folgende

Definition:
Ein Vektor mit nichtnegativen Einträgen und Spaltensumme 1 heißt „stochastischer Vektor“ (z. B. $\bar{\pi}_0, \bar{\pi}_n, \bar{\pi}$). Eine $n \times n$ -Matrix mit nichtnegativen Einträgen und alle Spaltensummen = 1 heißt „stochastische Matrix“ (z. B. A, A^n, G).

Eine entscheidende Frage ist nun: kann es andere stochastische Fixvektoren als den gesuchten Grenzvektor $\vec{\pi}$ geben? Wenn nicht, so könnten wir $\vec{\pi}$ einfach als den stochastischen Fixvektor berechnen.

Jeder stochastische Fixvektor \vec{s} ist klarerweise auch Grenzvektor (nämlich bei dem Startvektor $\vec{\pi}_0 = \vec{s}$) und der Grenzvektor $\vec{\pi}$ ist ja eindeutig und unabhängig vom stochastischen Startvektor $\vec{\pi}_0$.

Wir haben dadurch die Absicherung: der unbekannte stochastische Grenzvektor $\vec{\pi}$ ist einfach zu bestimmen, nämlich als der eindeutig existierende, stochastische Fixvektor von A .

Für den Palio bedeutet dies: Wie auch immer der Startvektor und die Startverteilung

$$\vec{\pi}_0 = \begin{pmatrix} p_0 \\ 1 - p_0 = q_0 \end{pmatrix}$$

„aussieht“, beim Palio-Auswahlmodus ist die Teilnahmewahrscheinlichkeit nach „vielen“ Jahren davon nahezu unabhängig und nähert sich dem Wert $\frac{10}{17}$. Die Frage nach der Gerechtigkeit des Palio-Auswahlverfahrens kann also positiv beantwortet werden.

Dies alles erhielten wir unter der Voraussetzung, dass eine wie oben beschriebene Grenzmatrix G existiert, so dass die Frage auf der Hand liegt: Wie kann man es einer stochastischen Matrix A ansehen, ob die Matrixpotenzen A^n gegen eine solche Grenzmatrix G konvergieren?

Bemerkung: Hier eröffnet sich mit CAS ein weites Probierfeld mit stochastischen Matrizen (2×2 , 3×3 , 4×4 , ...):

Finde stochastische Matrizen, so dass A^n gegen eine solche Grenzmatrix G konvergiert (bzw. nicht konvergiert).

Besonders für den Fall von 2×2 -Matrizen ist eine Vermutung leicht zu erhalten: Bis auf die zwei leicht zu

findenden Ausnahmen $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

und $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ konvergiert (A^n) für alle

stochastischen 2×2 -Matrizen A gegen eine Grenzmatrix G mit den beschriebenen Eigenschaften.

Nun folgt der „entscheidende“ Satz von Markoff.

Satz von Markoff:

Sei A eine stochastische Matrix. Wenn es ein $n \geq 1$ gibt, so dass A^n (mindestens) eine Zeile mit nur positiven Elementen hat, so konvergiert die Folge (A^n) gegen eine Grenzmatrix G mit Zeilen aus jeweils gleichen Werten.

In unserem Fall hatte A selbst ($n = 1$) bereits eine Zeile mit nur positiven Elementen: die erste Zeile.

Der allgemeine Beweis dieses Satzes ist im Schulunterricht nicht sinnvoll; die Formulierung, die Diskussion seiner Bedeutung und seine Verwendung jedoch sehr wohl.

Für den Spezialfall von 2×2 -Matrizen ist der Beweis sehr einfach (z. B. „Intervallschachtelungen“ – siehe Humenberger 2002; auch im allgemeinen Fall kann mit der selben Beweisidee gearbeitet werden).

Bemerkung: In das Thema „Markoff-Ketten“ fließen die drei klassischen Gebiete der Oberstufenmathematik: Stochastik, Lineare Algebra, Analysis ein – „Vernetzung“. Für weitere Beispiele zum Thema siehe z. B. (Meyer 1998).

Unterrichtserfahrungen

Herr Florian Rösner hat in einem Leistungskurs der Klasse 12 diesen Zugang zu Markoff-Ketten bereits ausprobiert und dem Autor sehr Positives darüber berichtet (die Schülerinnen und Schüler waren im Rahmen einer Kultur-Woche in der Toskana sogar vor Ort in Siena und haben einen Palio gesehen). Ein anschließender Unterrichtsschwerpunkt lag vor allem bei selbstständigem Experimentieren mit CAS. Da die Lernenden mit CAS umgehen konnten, haben auch Leistungsschwächere herausgefunden, dass die Matrixpotenzen A^n gegen eine spezielle Grenzmatrix zu konvergieren scheinen, wodurch ein Begründungsbedürfnis entstanden ist. Dies ist zunächst in Form heuristischer Plausibilitätsbetrachtungen und anschließend in Form eines (anderen einfachen) Beweises für den Fall von 2×2 -Matrizen geschehen. Weitere Anwendungsaufgaben aus dem Bereich Wirtschaft-Marketing (auch nicht-homogene Markoff-Ketten) ergänzten den Lehrgang.

Es ist nicht denknotwendig, hier Matrizenrechnung ins Spiel zu brin-

gen. Bei Prozessen mit nur zwei möglichen Zuständen genügt die Lösung einer einfachen Rekursionsgleichung, um die Konvergenz $p_n \rightarrow \frac{10}{17}$

bzw. eine explizite Darstellung für p_n herauszufinden. Wenn man nur an der Lösung dieses oder anderer Prozesse mit nur zwei möglichen Zuständen interessiert ist, und nicht dieses Beispiel als Einstieg zum Thema „Markoff-Ketten mit auch mehr als 2 Zuständen“ in Matrixdarstellung benutzen will, so könnte man auch kürzer folgendermaßen vorgehen:

Aus der Gleichung (1) folgt mit $q_n = 1 - p_n$ die Beziehung

$$p_{n+1} = -\frac{7}{10} p_n + 1 \quad (*),$$

eine einfache Rekursionsgleichung. Der einzig mögliche (dadurch noch nicht nachgewiesene!) Grenzwert p ergibt sich aus der Gleichung

$$p = -\frac{7}{10} p + 1, \text{ also } p = \frac{10}{17}.$$

Daher liegt es besonders nahe, die Abweichungen $p_n - \frac{10}{17}$ zu untersuchen (und als Nullfolge zu entlarven): Aus (*) erhält man durch Subtraktion von $\frac{10}{17}$ auf beiden Seiten sofort

$$\left(p_{n+1} - \frac{10}{17}\right) = \left(-\frac{7}{10}\right) \cdot \left(p_n - \frac{10}{17}\right)$$

wodurch klar ist, dass die Differenzen $p_n - \frac{10}{17}$ eine geometrische Folge mit dem Faktor $-\frac{7}{10}$ bilden. Dadurch ist nicht

nur die Konvergenz $p_n \rightarrow \frac{10}{17}$, sondern auch die Konvergenzgeschwindigkeit klar, und dass die Werte p_n abwechselnd über bzw. unter dem Grenzwert $p = \frac{10}{17}$ liegen. Eine explizite Darstellung für p_n wäre durch

$$p_n = \left(-\frac{7}{10}\right)^n \left(p_0 - \frac{10}{17}\right) + \frac{10}{17}$$

gegeben. Analog könnte bei allen anderen Fällen der Art $p_{n+1} = a \cdot p_n + b$ vorgegangen werden.

Literatur

Götz, S./M. Grosser: Über das Pferderennen in Siena. – In: Mathematische Semesterberichte 46 (1999), 1, S. 77–92.
 Humenberger, H.: Der Palio, das Pferderennen von Siena – Ausgangspunkt für Modelle von Auswahlprozessen und Einstieg zum Thema Markoff-Ketten. – In: Stochastik in der Schule 22 (2002), 2.
 Meyer, D.: Markoff-Ketten. – In: Mathematik in der Schule 36 (1998), 12, S. 661–670 und 675–680.