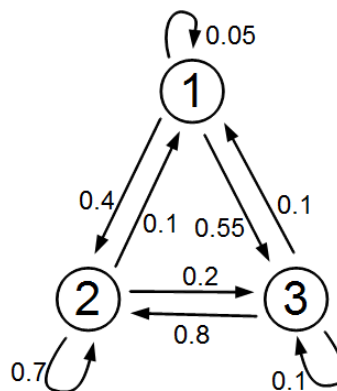


Lineare Algebra und analytische Geometrie II

Vorlesung 54

Stochastische Matrizen



DEFINITION 54.1. Eine reelle quadratische Matrix

$$M = (a_{ij})_{1 \leq i, j \leq n}$$

heißt *spaltenstochastisch*, wenn alle Einträge

$$a_{ij} \geq 0$$

sind und für jede Spaltensumme (also jedes j)

$$\sum_{i=1}^n a_{ij} = 1$$

gilt.

Die zugrunde liegende Interpretation für eine spaltenstochastische Matrix ist folgendermaßen: Man hat eine Menge von n möglichen Plätzen, Positionen, Netzwerkknoten, Internetseiten oder ähnliches, in denen man sich mit einer gewissen Wahrscheinlichkeit (einer Verteilung, einer Gewichtung)

aufhalten kann. Eine solche Verteilung wird durch ein n -Tupel $\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ mit

reellen nichtnegativen Zahlen v_i mit $\sum_{i=1}^n v_i = 1$ beschrieben. Man spricht

von einem *Verteilungsvektor*. Eine spaltenstochastische Matrix beschreibt die Übergangswahrscheinlichkeiten des gegebenen Netzwerks in einem bestimmten Zeitabschnitt. Der Eintrag a_{ij} ist die Wahrscheinlichkeit, dass ein im Knoten j befindliches Objekt (ein Besucher der Netzseite j) zur Position i hinüberwechselt (sich zur Netzseite i weiterklickt). Der j -te Standardvektor entspricht der Verteilung, in der alles im j -ten Knotenpunkt konzentriert ist. Die j -te Spalte der Matrix beschreibt das Bild dieses Standardvektors unter der Matrix. Generell wird zu einer Verteilung v durch Anwenden der Matrix die Bildverteilung Mv ausgerechnet, siehe Aufgabe 54.1. Naheliegende Fragen sind, ob es Verteilungen gibt, die stationär sind (*Stationäre Verteilung* oder *Fixverteilung* oder *Eigenverteilung*), also in sich selbst überführt werden, ob es periodische Verteilungen gibt, ob es bei „unendlich vielen“ Iterationen der Matrix Grenzverteilungen gibt, und wie man diese ausrechnen kann.

BEISPIEL 54.2. Eine spaltenstochastische 2×2 -Matrix hat die Form

$$\begin{pmatrix} p_1 & p_2 \\ 1 - p_1 & 1 - p_2 \end{pmatrix}$$

mit

$$0 \leq p_1, p_2 \leq 1.$$

Das charakteristische Polynom ist

$$\begin{aligned} & (X - p_1)(X - 1 + p_2) - (1 - p_1)p_2 \\ &= X^2 + (p_2 - p_1 - 1)X + p_1(1 - p_2) - p_2(1 - p_1) \\ &= X^2 + (p_2 - p_1 - 1)X + p_1 - p_2 \\ &= (X - 1)(X + p_2 - p_1). \end{aligned}$$

Eigenwerte sind also 1 und $p_1 - p_2$. Eine stationäre Verteilung ist (der Fall $p_1 = 1$ und $p_2 = 0$ ist für die folgende Rechnung auszuschließen) durch

$\begin{pmatrix} \frac{p_2}{p_2 - p_1 + 1} \\ \frac{1 - p_1}{p_2 - p_1 + 1} \end{pmatrix}$ gegeben, es ist ja

$$\begin{aligned} \begin{pmatrix} p_1 & p_2 \\ 1 - p_1 & 1 - p_2 \end{pmatrix} \begin{pmatrix} \frac{p_2}{p_2 - p_1 + 1} \\ \frac{1 - p_1}{p_2 - p_1 + 1} \end{pmatrix} &= \begin{pmatrix} p_1 \frac{p_2}{p_2 - p_1 + 1} + p_2 \frac{1 - p_1}{p_2 - p_1 + 1} \\ (1 - p_1) \frac{p_2}{p_2 - p_1 + 1} + (1 - p_2) \frac{1 - p_1}{p_2 - p_1 + 1} \end{pmatrix} \\ &= \begin{pmatrix} \frac{p_1 p_2 + p_2 (1 - p_1)}{p_2 - p_1 + 1} \\ \frac{(1 - p_1) p_2 + (1 - p_2)(1 - p_1)}{p_2 - p_1 + 1} \end{pmatrix} \\ &= \begin{pmatrix} \frac{p_2}{p_2 - p_1 + 1} \\ \frac{1 - p_1}{p_2 - p_1 + 1} \end{pmatrix}. \end{aligned}$$

BEISPIEL 54.3. Die spaltenstochastische 2×2 -Matrix

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

führt die Verteilung $\begin{pmatrix} p \\ 1-p \end{pmatrix}$ in die Verteilung $\begin{pmatrix} 1-p \\ p \end{pmatrix}$ über. Die Verteilung $\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ wird in sich selbst überführt, ist also eine stationäre Verteilung. Die Verteilung $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ wird in $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ überführt und umgekehrt, es handelt sich also um periodische Verteilungen der Periodenlänge 2.

BEISPIEL 54.4. Die spaltenstochastische $n \times n$ -Matrix

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

führt die Verteilung $\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ in die Verteilung

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n v_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

über. Der erste Standardvektor ist ein Eigenvektor zum Eigenwert 1, die weiteren Standardvektoren werden, wie jeder Verteilungsvektor, in den ersten Standardvektor überführt. Der Kern wird von den Vektoren $e_1 - e_j$, $j \geq 2$, erzeugt und enthält keine Verteilungsvektoren.

BEISPIEL 54.5. Es sei N ein Netzwerk (oder ein „gerichteter Graph“), bestehend aus einer Menge K aus Knotenpunkten und einer Menge an gerichteten Verbindungen, die zwischen Knotenpunkten bestehen können. Beispielsweise ist K die Menge aller Seiten im Internet und von der Seite $j \in K$ besteht ein Verbindungspfeil nach $i \in K$, falls es auf der Internetseite j einen Link auf die Seite i gibt. Die Verlinkungsstruktur kann man durch die *Adjazenzmatrix*

$$A = (a_{ij})$$

ausdrücken, wobei

$$a_{ij} := \begin{cases} 1, & \text{falls es einen Link von } j \text{ auf } i \text{ gibt,} \\ 0, & \text{sonst,} \end{cases}$$

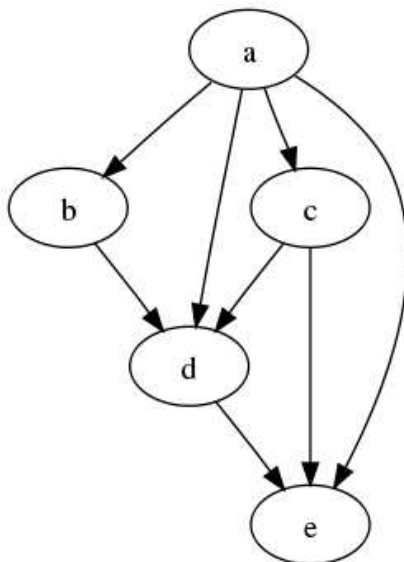
festgelegt ist (in der j -ten Spalte sind die von j ausgehenden Links ablesbar), oder aber durch die spaltenstochastische Matrix

$$B = (b_{ij}),$$

wobei

$$b_{ij} = \frac{a_{ij}}{d_j}$$

und d_j die Anzahl der Links angibt, die vom j -ten Knoten überhaupt ausgehen. Diese Division sichert, dass die Spaltensummennorm gleich 1 wird (es sei vorausgesetzt, dass von jedem Knoten mindestens ein Link ausgeht).



Die Adjazenzmatrix und die spaltenstochastisch gemachte Adjazenzmatrix zum Graphen rechts (wobei wir durchgängig Selbstlinks hinzunehmen) sind

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{3} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{2} & 1 \end{pmatrix}.$$

Potenzien von stochastischen Matrizen

Wir untersuchen nun die Potenzen von stochastischen Matrizen mit Hilfe der Summennorm und den Ergebnissen der letzten Vorlesung.

KOROLLAR 54.6. *Eine spaltenstochastische Matrix ist stabil.*

Beweis. Für einen beliebigen Vektor $v \in V$ ist Wegen

$$\begin{aligned} \|Mv\|_{\text{sum}} &= \sum_{i=1}^n |(Mv)_i| \\ &= \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} v_j \right| \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} |v_j| \right) \\
&= \sum_{j=1}^n |v_j| \left(\sum_{i=1}^n a_{ij} \right) \\
&= \sum_{j=1}^n |v_j| \\
&= \|v\|_{\text{sum}} .
\end{aligned}$$

Iterative Anwendung dieser Beobachtung zeigt, dass Satz 53.10 (2) erfüllt ist. \square

LEMMA 54.7. *Es sei*

$$M = (a_{ij})_{1 \leq i, j \leq n}$$

eine reelle quadratische Matrix mit nichtnegativen Einträgen. Dann ist M genau dann spaltenstochastisch, wenn M für Vektoren mit nichtnegativen Einträgen isometrisch bezüglich der Summennorm ist, wenn also

$$\|Mv\|_{\text{sum}} = \|v\|_{\text{sum}}$$

für alle $v \in \mathbb{R}_{\geq 0}^n$ gilt.

Beweis. Es sei M eine spaltenstochastische Matrix und

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

ein Vektor mit nichtnegativen Einträgen. Dann ist

$$\begin{aligned}
\|Mv\|_{\text{sum}} &= \sum_{i=1}^n (Mv)_i \\
&= \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} v_j \right) \\
&= \sum_{j=1}^n v_j \left(\sum_{i=1}^n a_{ij} \right) \\
&= \sum_{j=1}^n v_j \\
&= \|v\|_{\text{sum}} .
\end{aligned}$$

Wenn umgekehrt die angegebene isometrische Eigenschaft gilt, so gilt insbesondere für die Bilder der Standardvektoren, dass ihre Summennorm gleich 1 sein muss. Diese Bilder stehen in der entsprechenden Spalte der Matrix, alle Spaltensummen haben also den Wert 1. \square

LEMMA 54.8. *Es sei M eine spaltenstochastische Matrix. Dann gelten folgende Aussagen.*

- (1) *Es gibt Eigenvektoren zum Eigenwert 1.*
- (2) *Wenn es eine Zeile gibt, in der alle Einträge positiv sind, so gilt für jeden Vektor $v \in V$, der sowohl positive als auch negative Einträge besitzt, die Abschätzung*

$$\|Mv\|_{\text{sum}} < \|v\|_{\text{sum}} .$$

- (3) *Wenn es eine Zeile gibt, in der alle Einträge positiv sind, so ist der Eigenraum zum Eigenwert 1 eindimensional. Es gibt dann einen Eigenvektor, der nur nichtnegative Einträge hat und insbesondere eine eindeutig bestimmte stationäre Verteilung.*

Beweis. (1). Die transponierte Matrix ist zeilenstochastisch und besitzt daher

den Eigenvektor $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ zum Eigenwert 1. Daher besitzt nach Satz 23.2 das

charakteristische Polynom der transponierten Matrix eine Nullstelle an der Stelle 1 und dies gilt nach Aufgabe 23.19 dann auch für die ursprüngliche Matrix. Daher besitzt M einen Eigenvektor zum Eigenwert 1.

(2). Es seien nun zusätzlich alle Einträge der k -ten Zeile positiv und $v \in V$ sei ein Vektor mit (mindestens) einem positiven und einem negativen Eintrag. Dann ist

$$\begin{aligned} \|Mv\|_{\text{sum}} &= \sum_{i=1}^n |(Mv)_i| \\ &= \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} v_j \right| \\ &= \sum_{i \neq k} \left| \sum_{j=1}^n a_{ij} v_j \right| + \left| \sum_{j=1}^n a_{kj} v_j \right| \\ &< \sum_{i \neq k} \sum_{j=1}^n a_{ij} |v_j| + \sum_{j=1}^n |a_{kj} v_j| \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} |v_j| \\ &= \sum_{j=1}^n |v_j| \left(\sum_{i=1}^n a_{ij} \right) \\ &= \sum_{j=1}^n |v_j| \\ &= \|v\|_{\text{sum}} . \end{aligned}$$

(3). Wie im Beweis zu (2) seien alle Einträge der k -ten Zeile positiv. Für einen jeden Eigenvektor v zum Eigenwert 1 sind nach (2) entweder alle Einträge nichtnegativ oder nichtpositiv. Somit ist für einen solchen Vektor wegen $Mv = v$ der k -te Eintrag ungleich 0. Seien v, w solche Eigenvektoren. Dann gehört auch $\frac{w_k}{v_k}v - w$ zum Fixraum. Allerdings ist die k -te Komponente davon gleich 0 und daher ist es der Nullvektor. Das bedeutet, dass v und w linear abhängig sind. Somit ist dieser Eigenraum eindimensional. Wegen (2) gibt es einen Eigenvektor zum Eigenwert 1 mit nichtnegativen Einträgen. Durch Normieren sieht man, dass es auch eine stationäre Verteilung gibt. \square

BEISPIEL 54.9. Wir betrachten die spaltenstochastische 3×3 -Matrix

$$\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{2}{3} & 0 \\ \frac{1}{6} & 0 & \frac{2}{3} \end{pmatrix},$$

bei der alle Einträge der ersten Zeile positiv sind. Nach Lemma 54.8 gibt es eine eindeutige Eigenverteilung. Um diese zu bestimmen, berechnet man den Kern von

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{2}{3} & 0 \\ \frac{1}{6} & 0 & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{2} & \frac{1}{3} & 0 \\ -\frac{1}{6} & 0 & \frac{1}{3} \end{pmatrix}.$$

Dieser wird von $\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$ erzeugt und die stationäre Verteilung ist

$$\begin{pmatrix} \frac{2}{6} \\ \frac{3}{6} \\ \frac{1}{6} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{2} \\ \frac{1}{6} \end{pmatrix}.$$

BEISPIEL 54.10. Für die spaltenstochastische 3×3 -Matrix

$$\begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & \frac{2}{3} \end{pmatrix}$$

ist der Eigenraum zum Eigenwert 1 gleich $\langle e_1, e_2 \rangle$, also zweidimensional. Die Aussage Lemma 54.8 gilt also nicht, wenn es eine Spalte (aber keine Zeile) mit ausschließlich positiven Einträgen gibt.

SATZ 54.11. *Es sei M eine spaltenstochastische Matrix mit der Eigenschaft, dass es eine Zeile gibt, in der alle Einträge positiv sind. Dann konvergiert zu jedem Verteilungsvektor $v \in \mathbb{R}_{\geq 0}^n$ mit $\sum_{i=1}^n v_i = 1$ die Folge $M^n v$ gegen die eindeutig bestimmte stationäre Verteilung von M .*

Beweis. Es sei $w \in \mathbb{R}^n$ die nach Lemma 54.8 (3) eindeutig bestimmte stationäre Verteilung und

$$U = \left\{ \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \mid \sum_{i=1}^n u_i = 0 \right\} \subseteq \mathbb{R}^n.$$

Dies ist ein Untervektorraum von \mathbb{R}^n der Dimension $n-1$. Nach Lemma 54.8 (2) hat w ausschließlich nichtnegative Einträge und gehört damit nicht zu U . Wegen

$$\begin{aligned} \sum_{i=1}^n (Mu)_i &= \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} u_j \right) \\ &= \sum_{j=1}^n u_j \left(\sum_{i=1}^n a_{ij} \right) \\ &= \sum_{j=1}^n u_j \end{aligned}$$

ist U invariant unter der Matrix M . Somit ist

$$V = U \oplus \mathbb{R}w$$

eine direkte Summenzerlegung in invariante Untervektorräume. Für jedes $u \in U$ mit $\|u\|_{\text{sum}} = 1$ ist

$$\|Mu\|_{\text{sum}} < 1$$

nach Lemma 54.8 (2). Da die Sphäre zum Radius 1 bezüglich jeder Norm kompakt ist, ist die induzierte Maximumsnorm von $M|_U$ kleiner als 1. Nach Lemma 53.8 und Satz 53.6 konvergiert daher die Folge $M^n u$ für jedes $u \in U$ gegen den Nullvektor.

Es sei nun $v \in V$ ein Verteilungsvektor, den wir wegen

$$\sum_{i=1}^n v_i = 1 = \sum_{i=1}^n w_i$$

als

$$v = w + u$$

mit $u \in U$ schreiben können. Wegen

$$M^n v = M^n(w + u) = M^n w + M^n u = w + M^n u$$

und der Vorüberlegung konvergiert diese Folge gegen w . \square

BEMERKUNG 54.12. In der Situation von Lemma 54.8 kann man die Eigenverteilung dadurch finden, dass man ein lineares Gleichungssystem löst. Wenn es sich um eine sehr große Matrix (man denke an $\geq 10^9$ Knoten) handelt, ist eine solche Rechnung sehr aufwändig. Häufig muss man die Eigenverteilung aber gar nicht genau kennen, sondern es reicht eine gute Approximation aus.

Dann kann man zu einer beliebigen Startverteilung endlich viele Iterationen ausrechnen und weiß aufgrund von Satz 54.11, dass dieses Verfahren die Eigenverteilung beliebig gut approximiert. Eine Suchmaschine im Internet erstellt beispielsweise zu einem Suchbegriff eine Reihenfolge von Seiten, die diesen Begriff enthalten. Wie kommt diese Reihenfolge zustande? Die wahre Antwort ist, zumindest für die ersten Einträge, dass dies davon abhängt, wer am meisten dafür zahlt. Ansonsten ist ein natürlicher Ansatz, der auch Grundlage des Page ranks ist, die numerische Ordnung in der Eigenverteilung als ausschlaggebend anzusehen. Der oberste Eintrag ist derjenige, bei dem die meisten Leute „schließlich“ landen, wenn sie mit gleicher Wahrscheinlichkeit den möglichen Links folgen. Diese Wanderungsbewegung wird eben durch die stochastische Matrix, die man im Sinne von Beispiel 54.5 erhält, modelliert.¹

Den numerischen Unterschied zwischen dem exakten Lösen eines linearen Gleichungssystems zur Bestimmung eines Eigenvektors und der *Potenzmethode* kann man folgendermaßen erfassen. Sei eine $n \times n$ -Matrix gegeben. Das Gaussche Eliminationsverfahren braucht, um die erste Variable in $n - 1$ Gleichungen zu eliminieren, $n(n - 1) \sim n^2$ Multiplikationen (Additionen sind vom Rechenaufwand her einfacher und werden hier nicht berücksichtigt), die Größenordnung der Multiplikationen der Gesamtelimination ist somit

$$n^2 + (n - 1)^2 + (n - 2)^2 + \dots + 1 \sim \frac{1}{6}n^3.$$

Dagegen sind für die Auswertung der Matrix auf einen Vektor n^2 Multiplikationen nötig. Wenn man k Iterationen berechnen möchte, braucht man also kn^2 Operationen. Wenn also k deutlich kleiner als n gewählt werden kann, so ist der Gesamtrechenaufwand deutlich kleiner.

1

Unter *Modellierung* versteht man in der (insbesondere angewandten) Mathematik den Vorgang, realweltliche Phänomene mathematisch zu erfassen, zu verstehen und zu beeinflussen. Mathematisch modelliert werden physikalische Prozesse, Wetterphänomene, Finanzaktionen, etc.

Abbildungsverzeichnis

Quelle = Stochmatgraph.png , Autor = Benutzer NikelsenH auf Commons, Lizenz = CC-by-sa 3.0	1
Quelle = Tred-G.svg , Autor = Benutzer Dmitry Dzhus auf Commons, Lizenz = gemeinfrei	4