§ 30 Stochastische Matrizen

Def: Eine Matrix $M = (p_{ij}) \in \mathbb{R}^{k \times k}$ mit

$p_{ij} \in [0,1]$ und $\sum_{i=1}^{n} p_{ij} = 1$ für $j = 1, \dots n$

heißt stochastische Matrix.

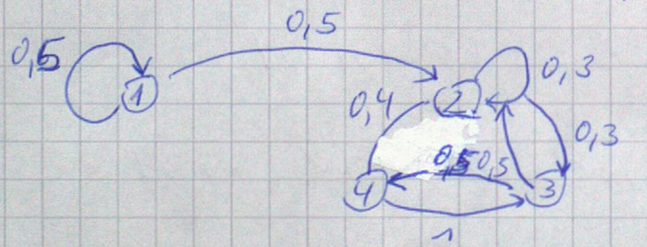
Eine stochastische Matrix beschreibt die Übergangswahrscheinlichkeit eines zeitunabhängigen Prozesses mit k Zuständen.
(eine sogenannte Markow-Kette)

Anwendung in der Informatik

- maschinelles Lernen
- Modellierung von Netzwerken
- Google Page Ranking

Bsp.:

a) Automat mit 4 Zuständen



p_{ij}: Wahrscheinlichkeit von Zustand j in den Zustand i zu kommen

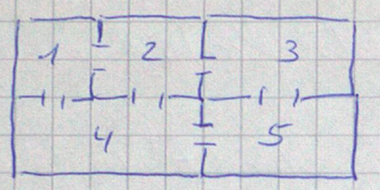
lten-
mme
1

von 1 zu 1

$$M = \begin{pmatrix} 0{,}5 & 0 & 0 & 0 \\ 0{,}5 & 0{,}3 & 0{,}5 & 0 \\ 0 & 0{,}3 & 0 & 1 \\ 0 & 0{,}4 & 0{,}5 & 0 \end{pmatrix}$$

von 1 zu 4

b) Maus im Labyrinth geht mit gleicher Wahrscheinlichkeit in einen der angrenzenden Räume



$$\mathfrak{P} = \begin{pmatrix} 0 & 0{,}\overline{3} & 0 & 0{,}\overline{3} \\ 0{,}5 & 0 & 0{,}5 & 0{,}\overline{3} \\ 0 & 0{,}\overline{3} & 0{,}5 & 0 \\ 0 & 0 & 0{,}5\;\;0 & 0{,}5 \\ 0{,}5 & 0{,}\overline{3} & 0 & 0 \end{pmatrix}$$

Diagonale 0

$$\begin{matrix} 0 \\ 0 \\ 0{,}5 \\ 0 \\ 0{,}5 \end{matrix}$$

man kann wieder zurück in den Raum

Bei $\mathfrak{P}^2$ sind die Diagonaleinträge > 0.

$$\mathfrak{P}^2 = \Big(\tfrac{1}{2}\cdot\tfrac{1}{3} + \tfrac{1}{2}\cdot\tfrac{1}{3}$$

c) Random surfen für Google Page Ranking: $k \sim 10^9$

Eigenschaften von stochastischen Matrizen

Satz: Sei $\mathfrak{P}$ eine stochastische Matrix. Dann gilt:

1) $\lambda = 1$ ist ein EW von $\mathfrak{P}$

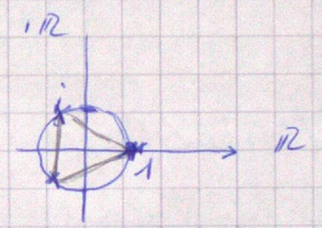
2) $|\lambda| \leq 1$ für jeden EW von $\mathfrak{P}$

Minimum der Diagonaleinträge

3) Ist $d = \min(p_{ii}) > 0$, dann sind alle EW von $\mathfrak{P}$ in der Kreisscheibe $\{z \in \mathbb{C} \mid |z-d| \leq 1-d\}$ enthalten. Insbesondere ist $\lambda = 1$ der einzige EW mit $|\lambda| = 1$.

4) Ist λ ein EW mit $|\lambda| = 1$, dann hat die Summe der Jordanblöcke von $\mathfrak{P}$ zu λ die Gestalt $\lambda \cdot E_s$, wobei s = Vielfachheit der Nullstelle λ im $\chi_{\mathfrak{P}}(t) = \det(\mathfrak{P} - tE)$.

5) Ist λ EW mit $|\lambda| = 1$, dann existiert ein $m \in \mathbb{N}$, so dass $\lambda^m = 1$, d.h. λ ist eine <u>m-te Einheitswurzel</u>.



$\lambda = x + iy$

$|\lambda| = \sqrt{x^2+y^2} = 1$

$\Rightarrow x^2 + y^2 = 1$ (Kreis!)

$m = 3$

$1,\ 1^{\frac{\pi}{3}},\ 1^{\frac{2\pi}{3}}$

$m = 4$

$1,\ i,\ -1,\ -i$

<u>Beweis:</u>

1) Sei $v = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^k$. Dann ist $\mathfrak{P}^t v = \begin{pmatrix} \sum_{j=1}^{k} p_{1j} \\ \vdots \\ \sum_{j=1}^{k} p_{kj} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = v$

$\Rightarrow \lambda = 1$ EW von $\mathfrak{P}^t$

$\Rightarrow \lambda = 1$ EW von $\mathfrak{P}$

$$0 = \det(\mathfrak{P}^t - E) = \det((\mathfrak{P}^t - E)^t) = \det(\mathfrak{P} - E)$$

2) Wir betrachten die Spaltensummen

$$\|\mathfrak{P}\|_s = \max_j \Big(\sum_{i=1}^{k} |p_{ij}| = 1$$

Also $|\lambda| \leq \|\mathfrak{P}\|_s = 1$ für jeden EW nach Satz über Eigenwertabschätzung bzgl. Matrixnormen.

3) Nach Satz von Gerschgorin liegen die EW ④
in der Vereinigung der Kreisscheiben

$$\{ z \in \mathbb{C} \mid |z - p_{ii}| \leq \sum_{i \neq j} p_{ij} = 1 - p_{ii} \}$$

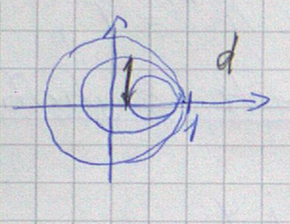
Alle Kreisscheiben sind in

$K = \{ z \in \mathbb{C} \mid |z - d| \leq 1 - d \}$ enthalten

und $\lambda = 1$ ist der einzige Punkt von K

mit $|\lambda| = 1$ im Fall also:

1 ist immer der rechte Punkt aller Kreise



4) Sei J die Jordansche Normalform von M
und $T \in GL(n, \mathbb{C})$ die Transformationsmatrix

$$TMT^{-1} = J = \begin{pmatrix} J_{r_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{r_m}(\lambda_m) \end{pmatrix}$$

Jordanblock / Kästchen

wobei $J_r(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$ $r \times r$ Matrix

Es gilt:

$$J^N = (TMT^{-1})^N = TM^NT^{-1}$$

Sei B ein Jordankästchen von J zum EW λ

$$B = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \Rightarrow B^2 = \begin{pmatrix} \lambda & \lambda\lambda & 1 & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & \ddots & \lambda\lambda \\ 0 & & & & \lambda \end{pmatrix}$$

$$\overset{\text{induktiv}}{\Longrightarrow} \begin{pmatrix} \lambda^N & N\lambda^{N-1} & & * \\ & \ddots & \ddots & \\ & & \ddots & N\lambda^{N-1} \\ 0 & & & \lambda^N \end{pmatrix}$$

⑤

Für $|\lambda| = 1$ gilt

$$N + 1 = |\lambda^N| + N|\lambda^{N-1}| \leq \|B^N\|_s \leq \|J^N\|_s = \|TM^NT^{-1}\|_s$$

$$\leq \|T\|_s \cdot \|M^N\|_s \cdot \|T^{-1}\|_s$$

$$\leq \|T\|_s \underbrace{\|M\|_s^N}_{=1} \cdot \|T^{-1}\|_s$$

$$= \|T\|_s \cdot \|T^{-1}\|_s$$

Dies ist ein Widerspruch, wenn nicht $B = (\lambda)$ eine
1×1 Matrix und damit $B^N = (\lambda^N)$ ist.
Alle Jordanblöcke haben also Größe 1 und
es gilt $\mathrm{Eig}(M, \lambda) = s$. Dimension
für alle λ mit $|\lambda| = 1$

5) Für den letzten Fall brauchen wir etwas
Notation.

Sei $K = \{1, \ldots k\}$; Für $j \in K$ sei

Übergangswahrscheinlichkeit

$$M_1(j) = \{ i \mid p_{ij} > 0 \}$$

die Zustände, die vom Zustand j in
<u>einem</u> Schritt erreicht werden können.

Für $K_0 \subset K$ eine Teilmenge der Zustände definieren wir

$$M_1(K_0) = \bigcup_{j \in K_0} M_1(j)$$

Rekursiv dafür

$$M_t(K_0) = M_1(M_{t-1}(K_0))$$

die Menge, der in t Schritten zu erreichende Zustand

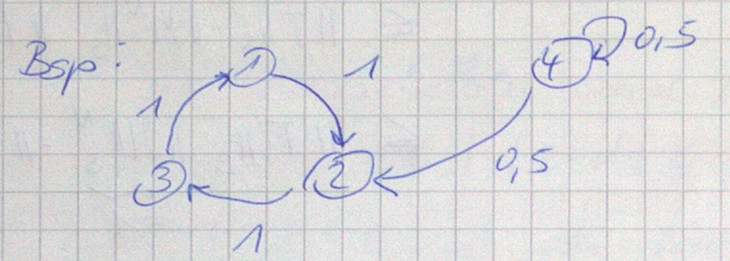
wobei $M_0(K_0) = K_0$

(6)

$$\mathcal{N}(K_0) = \bigcup_{t \geq 0} \mathcal{N}_t(K_0)$$

die aus K_0 überhaupt erreichbaren Zustände.

Bsp:



$K = \{1, \dots 4\}$

$K_0 = \{1\}$: $\mathcal{N}_1(K_0) = \{2\}$ // $\mathcal{N}_1(\mathcal{N}_1(K_0))$ von 1 kann man nur zur 2

$\mathcal{N}_2(K_0) = \{3\} = \mathcal{N}_1(\{2\}) = \{3\}$

$\mathcal{N}_3(K_0) = \{1\}$

$\rightarrow \mathcal{N}(K_0) = \{1, 2, 3\}$

$K_0 = 4$ $\quad \mathcal{N}_1(K_0) = \{2, 4\}$

$\mathcal{N}_2(K_0) = \mathcal{N}(\{2,4\}) = \{2, 3, 4\}$

$\mathcal{N}_3(K_0) = \{1, 2, 3, 4\} = K$

Sei jetzt λ ein EW mit $|\lambda| = 1$
und $x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$ ein EW von $\mathcal{N}^t$

d.h. $\quad \lambda x_i = \sum_{i=1}^{k} x_i p_{ij} \qquad \forall j = 1, \dots k$

Sei $y = \max |x_j|$ und $K_0 = \{j \mid |x_j| = y\} \neq 0$

Dann gilt für $j = K_0$:

$$y = |x_j| = |\lambda x_j| = \left| \sum_{i=1}^{k} x_i p_{ij} \right| \overset{\triangle\text{-Ungl.}}{\leq} \sum_{i=1}^{k} |x_i| p_{ij}$$

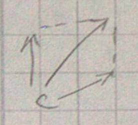
$$\leq \sum_{i=1}^{k} y p_{ij} = y \left| \sum_{i=1}^{k} p_{ij} \right| = y$$

so dass überall Gleichheit gelten muss.

(7)

Somit ist für $i \in \mathcal{N}_1(j)$ $\quad (p_{ij} \neq 0)$

$\underline{|x_i| = y = |x_j|}$ also $i \in K_0$



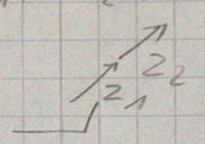
Außerdem ist

$$\left| \sum_{i=1}^{k} x_i p_{ij} \right| = \sum_{i=1}^{k} |x_i| p_{ij}$$

$|z_1 + z_2| = |z_1| + |z_2|$

$z_1, z_2 \in \mathbb{C}$

$\Rightarrow z_1 = c z_2$



Die komplexen Zahlen x_i, x_h zeigen in die gleiche Richtung $\forall i, h \in \mathcal{N}_1(j)$

Zusammengefasst:

$$x_h = x_h \sum_{i=1}^{k} p_{ij} = \sum_{i=1}^{k} x_i p_{ij} = \lambda x_i \qquad \forall h \in \mathcal{N}_1(j)$$

Sei nun $j \in K_0$ <u>fest</u> mit $|x_j| = y$

Wir betrachten

$K_1 = \{ l \mid x_l = x_j \}$

Dann gilt $\mathcal{N}_1(K_1) \neq \emptyset$ und für $i \in \mathcal{N}_1(K_1)$

$x_i = \lambda x_j$

Induktiv erhalten wir

$\mathcal{N}_t(K_0) \neq 0$ und für $i \in \mathcal{N}_t(K_1)$

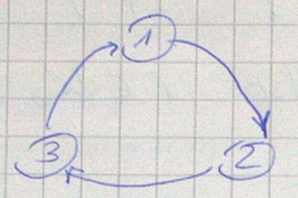
$x_i = \lambda^t x_j$

Die Mengen $\mathcal{N}_0(K_1), \mathcal{N}_1(K_1), \dots, \mathcal{N}_t(K_1)$ können nicht alle disjunkt sein, da deren Vereinigung endlich ist.

$\Rightarrow \exists_{s<t} s, t : \mathcal{N}_t(K_1) \cap \mathcal{N}_s(K_1) \neq \emptyset$

$\Rightarrow \lambda^t x_j = x_i = \lambda^s x_i \Rightarrow \lambda^{t-s} = 1$ d.h. nach Def eine Einheitswurzel $\square$

Bsp.:



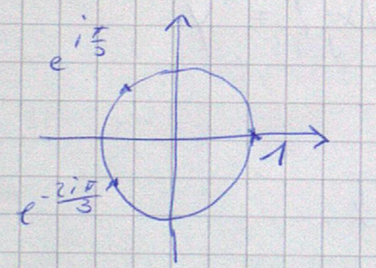
$$P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\chi_P(t) = -t^3 + 1 \quad \Rightarrow \quad t^3 = 1$$

$$\rightarrow \lambda_1 = 1, \ \lambda_2 = e^{i\frac{\pi}{3}}, \ \lambda_3 = e^{\frac{2\pi}{3}}$$

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

$$v_2 = \operatorname{Ke} \begin{pmatrix} -e^{i\frac{\pi}{3}} & 0 & 1 \\ 1 & -e^{i\frac{\pi}{3}} & 0 \\ 0 & 1 & -e^{i\frac{\pi}{3}} \end{pmatrix} = \begin{pmatrix} 1 \\ e^{\frac{2i\pi}{3}} \\ e^{\frac{i\pi}{3}} \end{pmatrix} x_j$$



$$\lambda x_j = e^{\frac{2\pi}{3}}$$

$$\lambda^2 x_j = 1$$

Satz: Ergodensatz

Sei $P \in \mathbb{R}^{k \times k}$ eine stochastische Matrix.

a) Dann existiert stets

$$Q = \lim_{n \to \infty} \frac{1}{m} \left(\sum_{l=0}^{m-1} P^l \right)$$

und es gilt

$$Q^2 = Q = QP = PQ$$

b) Rang $Q = s$ ist die Vielfachheit des EWs $\lambda = 1$ von P und Q beschreibt eine Projektion auf $\operatorname{Eig}(P, 1)$

c) Ist $\lambda = 1$ der einzige EW zu P mit $|\lambda| = 1$ $\quad (\exists \, p_{ij} \neq 0)$

dann gilt:

$$\lim_{n \to \infty} P^n = Q$$

d) Ist $\lambda = 1$ der einzige EW von P mit $|\lambda| = 1$ und $\dim \operatorname{Eig}(P, 1) = 1$ nur 1 Vektor

dann ist

$$\lim_{n \to \infty} P^n = \begin{pmatrix} z_1 & \cdots & z_1 \\ \vdots & & \vdots \\ z_k & \cdots & z_k \end{pmatrix}$$

mit $z = [0,1]^k \subset \mathbb{R}^k$; $\sum_{i=1}^{k} z_i = 1$ der eindeutig bestimmte Gleichgewichtszustand von P, d.h. $Pz = z$.