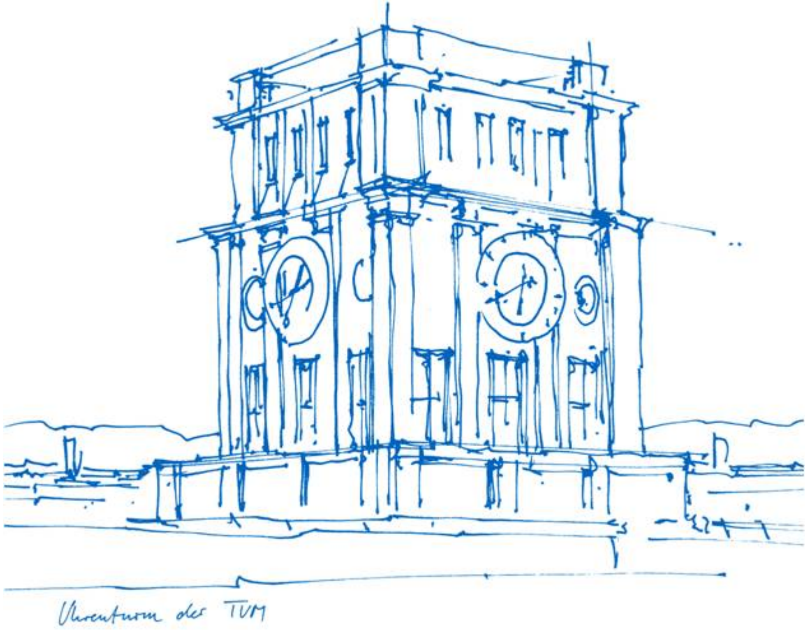
Lineare Algebra für Informatiker – Tutorium KW 27

Jeremias Bohn

Technische Universität München

Fakultät für Mathematik

Garching, 01./02. Juli 2021



Recap: Diagonalisierung

Für die Diagonalisierung einer Matrix $A \in K^{n \times n}$ benötigen wir folgende Schritte:

- Eigenwerte berechnen
  - Charakteristisches Polynom $\chi_A(\lambda) = \det(\lambda \cdot I_n - A)$ aufstellen
  - Nullstellen des charakteristischen Polynoms berechnen (diese sind die Eigenwerte)
- Für jeden konkreten Eigenwert λ_i den Eigenraum mit $\mathrm{Kern}(\lambda_i \cdot I_n - A)$ berechnen
- Diagonalisierbarkeit überprüfen (für jeden Eigenwert gilt $m_g(\lambda_i) = m_a(\lambda_i)$)
- Matrix D aufstellen (Eigenwerte auf einer Diagonalen jeweils mit Häufigkeit $m_a(\lambda_i)$)
- Matrix S aufstellen (Basisvektoren der Eigenräume als Spalten, jeweils in der Spalte, in der in D der zugehörige Eigenwert steht, jeder Basisvektor exakt einmal!)
- Es gilt $A = SDS^{-1}$ bzw. $D = S^{-1}AS$ und A und D sind ähnlich

PageRank-Algorithmus (Random Surfer)

Der PageRank-Algorithmus ist der ursprünglich von Google genutzte Algorithmus zur Bestimmung der Relevanz von Webpages. Der Algorithmus funktioniert für n Seiten wie folgt:

- Bilde die **Weblink-Matrix** $W = \left(w_{i,j}\right)$ mit $w_{i,j} = \begin{cases} 1, & \text{falls Link von Seite } i \text{ zu Seite } j \text{ existiert} \\ 0, & \text{sonst} \end{cases}$
- Bilde (mithilfe der Matrix W) die **Übergangsmatrix** $S = \left(s_{i,j}\right)$, wobei jeder Link von einer Seite mit gleicher Wahrscheinlichkeit aufgerufen wird oder, wenn die Seite keine Links besitzt, jede Seite mit gleicher Wahrscheinlichkeit

  als nächste aufgerufen wird, d.h. $s_{i,j} = \begin{cases} \frac{1}{\sum_{j=1}^{n} w_{i,j}}, & \text{falls Link von Seite } i \text{ zu Seite } j \text{ existiert} \\ \frac{1}{n}, & \text{falls } \sum_{j=1}^{n} w_{i,j} = 0 \\ 0, & \text{sonst} \end{cases}$
- Bilde die **Google-Matrix** $G \coloneqq (1-\alpha) \cdot S + \frac{\alpha}{n} \cdot (\mathbf{1} \cdot \mathbf{1}^T)$ mit Dämpfungsfaktor α

PageRank-Algorithmus (ff.)

Der **PageRank-Vektor** entspricht der Relevanz einer Website, wobei eine Website wichtiger ist, wenn viele Websites auf sie verlinken, sie selbst aber auf wenige andere Websites verlinkt. Der PageRank-Wert ist dabei die Wahrscheinlichkeit, auf dieser Seite zu landen, sodass sich die Wahrscheinlichkeitsverteilung nach einem erneuten Übergang nicht mehr ändert (eine sog. stationäre Verteilung)

- Dieser Vektor ist ein Zeilenvektor und es gilt $p \cdot G = p$ bzw. $G^T \cdot p^T = p^T$
- Dies entspricht dem Eigenvektor der transponierten Google-Matrix zum Eigenwert 1 (normiert auf Länge 1)
- Alternativ lässt sich auch $\lim_{n \to \infty} G^n$ für die Google-Matrix berechnen, jede Zeile dieser Matrix entspricht dann dem PageRank-Vektor p