Mathematik I

Vorlesung 14 - Eigenwerte und Eigenvektoren

Prof. Dr. Sandra Eisenreich

08./11. Januar 2023

Hochschule Landshut

Motivation + Anwendung in der Informatik

- Eine Drehung im R^3 ist eine lineare Abbildung. Gibt es einen Vektor, der durch die Drehung fest bleibt? Genau, die Achse, um die man dreht!
- Eine Spiegelung an einer Ebene lässt auch manche Vektoren konstant: alle in der Spiegelebene.
- Man kann zeigen: für alle linearen Abbildungen gibt es Vektoren (sog. **Eigenvektoren**) oder ganze Unterräume (sog. **Eigenräume**), die konstant bleiben oder nur gestreckt werden. Der Streckungsfaktor heißt **Eigenwert**.

- z.B. zum Ausrechnen von großen Potenzen von Matrizen
- Datenkomprimierung
- Bildverarbeitung
- Statistik
- Datenanalyse (Hauptkomponentenanalyse)
- Praktische Informatik (z.B. PageRank - Algorithmus)

Eigenwert und Eigenvektor

Definition

Sei $f : K^n \to K^n$ (K Körper) eine lineare Abbildung. Ein $\lambda \in K$ heißt **Eigenwert von** f, wenn es einen Vektor $v \in K^n$ gibt, so dass $v \neq 0$ und

$$f(v) = \lambda \cdot v.$$

v heißt dann **Eigenvektor zum Eigenwert** λ (das heißt er wird unter f nur um λ gestreckt).

Definition (Eigenraum)

Ist $\lambda \in K$ ein Eigenwert der linearen Abbildung $f \cdot K^n \to K^n$, dann ist der **Eigenraum zu** λ definiert als die Menge aller Vektoren, die unter f nur um λ gestreckt werden:

$$T_\lambda := \{v \in K^n | f(v) = \lambda \cdot v\}$$

Satz

Jeder Eigenraum ist ein linearer Unterraum.

Beispiele → Mitschrift

Herleitung: Berechnung von Eigenvektoren → Mitschrift

Berechnung von Eigenvektoren

Das, was wir eben hergeleitet haben, gilt allgemein:

Satz

Sei $f\colon K^n \longrightarrow K^n$ eine lineare Abbildung, dargestellt durch $A =$ Mat $f \in K^{n \times n}$. Außerdem sei E_n die $n \times n$-Einheitsmatrix. Dann gilt:

- *v ist genau dann Eigenvektor zum Eigenwert λ, wenn gilt:*

$$(A - \lambda \cdot E_n) \cdot v = 0 \ \textit{(LGS!)}, \ \textit{d.h. wenn } v \in \ker(A - \lambda \cdot E_n)$$

- *Ist λ ein Eigenwert von f bzw. A, so gilt:*

$$T_\lambda = \ker(A - \lambda \cdot E_n) = \ \textit{Lösungsmenge des LGS: } (A - \lambda \cdot E_n) \cdot v = 0$$

- *λ ist genau dann ein Eigenwert von f bzw. A, wenn*

$$\ker(A - \lambda \cdot E_n) \neq \{0\}, \ \textit{d.h. wenn } \dim \ker(A - \lambda \cdot E_n) > 0.$$

Herleitung: Eigenwertberechnung

Falls wir wissen, dass λ ein Eigenwert ist, bekommen wir den Eigenraum T_λ als die Lösungsmenge (mit dem Gauß-Algorithmus leicht zu bestimmen!) des LGS

$$(A - \lambda \cdot E_n) \cdot v = 0.$$

Also bleibt die Frage: Wie finden wir die Eigenwerte λ, d.h. die mit $\dim \ker(A - \lambda \cdot E_n) > 0$?

Erinnerung: Es gilt: $\dim \ker(A - \lambda \cdot E_n) + \operatorname{rang}(A - \lambda \cdot E_n) = n$

Antwort: Aus obigem folgt:

$$\dim \ker(A - \lambda \cdot E_n) > 0 \Leftrightarrow \operatorname{rang}(A - \lambda \cdot E_n) < n \Leftrightarrow \det(A - \lambda \cdot E_n) = 0$$

$\Rightarrow$ Vorgehen: Berechne die Determinante von $A - \lambda \cdot E_n$ mit λ als Variable. Dies ist ein Polynom in λ (man nennt es das **charakteristische Polynom**). Die Nullstellen sind die Eigenwerte!

Charakteristisches Polynom

Definition

Für unbestimmtes λ ist $\det(A - \lambda \cdot E_n)$ ein Polynom von Grad n, das sogenannte **Charakteristische Polynom** $\chi_A(\lambda)$.

Satz

λ ist ein Eigenwert von A genau dann, wenn $\det(A - \lambda \cdot E_n) = 0$, das heißt die Eigenwerte sind die Nullstellen des charakteristischen Polynoms $\chi_A(\lambda)$.

Rechenregel

Berechnung von Eigenwerten und Eigenräumen:

- 1. Schritt: Berechne das charakteristische Polynom $\chi_A(\lambda) = \det(A - \lambda \cdot E_n)$
- 2. Schritt: Berechne die Nullstellen von $\chi_A(\lambda)$. Das sind die Eigenwerte.
- 3. Schritt: Für jeden in Schritt 2 berechneten Eigenwert, löse das LGS $(A - \lambda \cdot E_n)v = 0$ durch Anwenden des Gauß-Algorithmus. Der Lösungsraum ist T_λ.

Beispiele → Mitschrift

Eigenvektoren und Basen

Beispiel in $\mathbb{R}^3$ von vorhin: die erzeugenden Vektoren für alle Eigenräume waren:

$$\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

Was fällt auf? - Die Vektoren sind linear unabhängig, also eine Basis des $\mathbb{R}^3$!

Dies gilt allgemein:

Satz

- *Sind $v_1, \dots, v_k$ Eigenvektoren zu unterschiedlichen Eigenwerten $\lambda_1, \dots, \lambda_k$, dann sind die Vektoren $v_1, \dots, v_k$ linear unabhängig.*
- *Sind die Eigenwerte $\lambda_1, \dots, \lambda_n$ einer Matrix $A \in K^{n \times n}$ alle verschieden und in K, dann bilden die Eigenvektoren $v_1, \dots, v_n$ zu den jeweiligen Eigenwerten $\lambda_1, \dots, \lambda_n$ eine Basis von K^n.*

reelle symmetrische Matrizen

Bisher hatten wir "schöne" Beispiele, in denen χ_A für reelle Vektorräume immer in Linearfaktoren zerfallen ist. Aber: das muss nicht sein, z.B. wenn $\chi_A(\lambda) = x^2 + 1$ (keine Nullstelle in $\mathbb{R}$, also keine Eigenwerte!)

Frage: Wann sind wir in einem "schönen" Fall mit reellen Nullstellen?

Satz

- *Falls n ungerade, dann besitzt χ_A immer mindestens eine reelle Nullstelle.*
- *Falls A eine reelle symmetrische Matrix ist (d.h. $A = A^T$), dann besitzt χ_A nur reelle Nullstellen.*
- *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch, d.h. $A = A^T$. Dann sind alle Eigenwerte reell und Eigenvektoren v, v' zu unterschiedlichen Eigenwerten stehen senkrecht aufeinander, d.h. $v \cdot v' = 0$.*

Symmetrische Matrizen sind also gut! Glücklicherweise kommen in Anwendungen (z.B. in der KI) oft symmetrische Matrizen vor.

Diagonalisierbare Matrizen

Dass es wie bei reellen symmetrischen Matrizen eine Basis aus Eigenvektoren gibt ist so eine tolle Eigenschaft, dass man ihr einen eigenen Namen gibt.

Definition

Eine Matrix A, für die n Eigenvektoren $v_1, \dots, v_n$ zu den Eigenwerten $\lambda_1, \dots, \lambda_n$ (möglicherweise mit Vielfachheit) bekannt sind und eine Basis des $\mathbb{R}^n$ bilden, nennt man **diagonalisierbar**.

Satz

- *Reelle symmetrische Matrizen sind diagonalisierbar.*
- *Ist A eine diagonalisierbare Matrix mit Eigenvektoren $v_1, \dots, v_n$ zu den Eigenwerten $\lambda_1, \dots, \lambda_n$, dann gilt für die Matrix B mit den Eigenvektoren als Spalten ($B = (v_1 \dots v_n)$):*

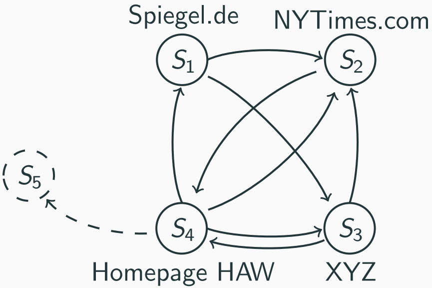
$$B^{-1} \cdot A \cdot B = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

Anwendung: Effektives Potenzieren von Matrizen

Für diagonalisierbare Matrizen wie oben kann man schnell Potenzen ausrechenen! Es gilt nämlich (mit der Bezeichnung von vorhin):

$$
\begin{aligned}
A^k &= B \cdot \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \cdot B^{-1} \cdot B \cdot \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \cdot B^{-1} \cdot \ldots \\
&= B \cdot \begin{pmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{pmatrix} \cdot B^{-1}.
\end{aligned}
$$

Beispiele → Mitschrift



Gegeben: Das Internet bestehend aus Webseiten $S_1, \ldots, S_N$ und Verlinkungen von Webseiten S_i zu Webseiten S_j.

Idee: Eine Webseite ist wichtig, wenn andere wichtige Webseites auf diese verlinken.

Problem: Implizite Definition

Um die Wichtigkeit einer Webseite zu bestimmen betrachtet man das sog. **Zufallssurfer / Random Surfer -Modell**.

Regel: Befindet sich der Zufallssurfer zum Zeitpunkt t auf der Seite S_i und existieren l_i Links von S_i zu anderen Seiten S_j, dann wird jede dieser Seiten zum Zeitpunkt $t+1$ mit Wahrscheinlichkeit $\frac{1}{l_i}$ angewählt. Wir bezeichnen mit $w_i(t)$ die Wahrscheinlichkeit, dass man sich zum Zeitpunkt t auf der Seite S_i befindet,

Beispiel → Mitschrift

Also falls die Wahrscheinlichkeitsverteilung stationär wird (d.h. $w(t)$ verändert sich nicht mehr/oder kaum ab einem gewissen t_0), so gilt

$$w \approx w(t) = M \cdot w(t-1) = M \cdot w$$

Also im Grenzwert gilt $w = M \cdot w$, und somit ist die Wahrscheinlichkeitsverteilung w ein Eigenvektor von M.

Problem: Gibt es Seiten, auf denen keine Links zu andern Seiten existieren, die aber selbst verlinkt sind, dann kann der Zufallssurfer auf solchen Seiten gefangen sein.

Ausweg: In jeden Schritt wählt der Zufallssurfer einen der bestehenden Links mit Wahrscheinlichkeit $1-d$, und geht auf eine beliebige Seite mit Wahrscheinlichkeit $\frac{d}{N}$. Hierbei ist $d \in (0,1)$ der sogenannte Dämpfungsfaktor. Dann ergibt sich folgende Wahrscheinlichkeitsverteilung zum Zeitpunkt t:

$$w(t) \quad = \quad (1-d) \cdot M \cdot w(t-1) + d \cdot \begin{pmatrix} 1/N \\ 1/N \\ \vdots \\ 1/N \end{pmatrix} = (1-d) \cdot M \cdot w(t-1) + \frac{d}{N} \cdot \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Falls die Verteilung $w(t)$ stationär wird $(= w)$ dann gilt:

$$
\begin{aligned}
w &= (1-d)\cdot M \cdot w + \frac{d}{N}\cdot \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \\
\Rightarrow \quad w &= (1-d)\cdot M \cdot w + \frac{d}{N}\cdot \underbrace{\begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \dots & \dots & 1 \end{pmatrix} w}_{\text{da } w_1+w_2+\ldots+w_N=1} \\
\Rightarrow \quad w &= \left((1-d)\cdot M + \frac{d}{N}\cdot \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \dots & \dots & 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}
\end{aligned}
$$

$$
\text{Also: } w = \underbrace{\left((1-d)\cdot M + \frac{d}{N}\cdot E\right)}_{M_d} \cdot w, \text{ wobei } \mathbb{E} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \dots & \dots & 1 \end{pmatrix}
$$

Satz (Perron Frobenius)

Die Matrix M_d hat die Eigenschaft, dass es genau einen reell-wertigen Eigenvektor mit ausschließlich positiven Koordinaten gibt. Weiter gilt, dass der zugehörige Eigenwert reell ist und den höchsten Betrag aller Eigenwerte hat.