1 Lineare Algebra

1.1 Kern, Bild und Rang

Ein **Kern** ($Ker(A)$) existiert, wenn $\det(A) = 0$.
Der Kern einer Matrix A ist die Lösungsmenge von $A \cdot \vec{v} = \vec{0}$
$\rightarrow$ LGS=0 durch elem. Zeilenoperationen lösen.

Das **Bild** ($Im(A)$) einer Matrix gibt an, welche Menge an Vektoren als Lösungen auftreten können (vgl. Wertebereich bei Funktionen).
Das Bild einer Matrix A ist die Lösungsmenge von $A \cdot \vec{v} = \vec{b}$

Der **Rang** ($rank(A)$) einer Matrix A ist die Anzahl der linear unabhängigen Zeilen- bzw. Spaltenvektoren.
Rang = Anzahl der Nichtnullzeilen der Matrix in Zeilenstufenform.
$\rightarrow$ A durch elem. Zeilenoperationen umformen.

1.2 Determinante

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

2x2 Matrix

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + bfg + cdh - gec - hfa - ibd$$

3x3 Matrix

Ergänzung: Laplace'scher Entwicklungssatz bei höherrangigen Matrizen

Siehe: https://www.mathebibel.de/laplace-entwicklungssatz

1.3 Eigenwerte, Eigenvektoren und Eigenraum

Eine Zahl λ heißt Eigenwert der Matrix A, wenn es einen Vektor $\vec{v}$ gibt, der nicht der Nullvektor ist, so dass gilt:

$$\begin{aligned} Av &= \lambda v \\ Av - \lambda v &= 0 \\ (A - \lambda I)v &= 0 \end{aligned}$$

1.3.1 Charakteristisches Polynom berechnen

Anstatt o.g. Gleichungzu lösen: Bestimmung der Nullstellen des charakteristischen Polynoms $p_A(\lambda)$ der Matrix A.

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda E) \\ &= \begin{vmatrix} a_{11} - \lambda & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} - \lambda \end{vmatrix} \stackrel{!}{=} 0 \end{aligned}$$

1.3.2 Eigenvektoren berechnen

Der zu einem Eigenwert λ_i gehörende Eigenvektor $\vec{v_i}$ ist die Lösung der Gleichung:

$$\begin{aligned} A\vec{v_i} &= \lambda_i \vec{v_i} \\ (A - \lambda_i I) \cdot \vec{v_i} &= \vec{0} \end{aligned}$$

Rechenweg:

1. λ_i für λ in die Matrix $(A - \lambda E)$ einsetzen (siehe charakterisches Polynom)
2. Das folgende LGS durch elementare Zeilenoperationen lösen:
$$\left(\begin{array}{ccc|c} a_{11} - \lambda & \cdots & a_{1n} & 0 \\ \vdots & \ddots & \vdots & 0 \\ a_{n1} & \cdots & a_{nn} - \lambda & 0 \end{array}\right)$$
3. Für Nullzeilen ergeben sich beliebige Lösungen, die gleich 1 gesetzt werden können.

1.3.3 Eigenraum berechnen

Der Eigenraum $E_A(\lambda_i)$ einer Matrix A zu einem Eigenwert λ_i ist die Menge aller Eigenvektoren $\vec{v_i}$ zu λ_i.

Lösung: Vielfaches der Eigenvektoren in Mengenschreibweise festhalten:

$$E_A(\lambda_i) = \{k \cdot \vec{v_i} | k \in \mathbb{R}\}$$

1.3.4 algebraische vs. geometrische Vielfachheit von λ

- **algebraische Vielfachheit**: Anzahl gleicher Eigenwerte im charakteristischen Polynom
- **geometrische Vielfachheit**: Dimension (Anzahl der Vektoren) des Eigenraums $E(\lambda)$; $\leq$ algebraische Vielfachheit

1.4 Orthogonale Matrizen

Zwei Vektoren sind orthogonal, wenn ihr Skalarprodukt

$$\langle a, b\rangle = a_1 b_1 + \ldots + a_i b_i \ = 0$$

Äquivalente Aussagen:

- Matrix B ist orthogonal
- $B^T B = I$, d.h. B ist invertierbar mit $B^{-1} = B^T$.
- Die Spaltenvektoren von B definieren eine Orthonomalbasis von $\mathbb{R}^n$

1.4.1 Orthogonalen Vektor mit dem Kreuzprodukt finden

Für $\vec{a} \perp \vec{b}$ ergibt sich $\vec{c}$ mit $\vec{c} \perp \vec{a}, \vec{c} \perp \vec{b}$ aus:

$$\vec{c} = \vec{a} \times \vec{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$$

1.4.2 Gram-Schmidt-Verfahren

Ziel: Orthonormalbasis (ONB) zu einem Vektorraum $B = \{b_1, b_2, \ldots b_n\}$ finden.

1. Ersten Basisvektor normieren: $\vec{q_1} = \frac{\vec{q_1}}{||\vec{q_1}||}$
2. Fälle das Lot von b_2 auf die von q_1 erzeugte Gerade: $l_2 = b_2 - \langle b_2, q_1\rangle q_1$
3. Normiere das Lot: $\vec{q_2} = \frac{\vec{l_2}}{||\vec{l_2}||}$
4. Wiederhole Schritte 2 und 3 für alle Basisvektoren:
   $l_i = b_i - \langle b_i, q_1\rangle q_1 - \langle b_i, q_2\rangle q_2 - \ldots - \langle b_i, q_{i-1}\rangle q_{i-1}$ und $\vec{q_i} = \frac{\vec{l_i}}{||\vec{l_i}||}$

1.5 Diagonalisierbarkeit

1.5.1 Diagonalisierbarkeit

A ist diagonalisierbar, wenn

- für jeden Eigenwert von A die algebraische Vielfachheit gleich der geometrischen Vielfachheit ist, oder
- wenn alle Eigenwerte (λ_i) von A unterschiedlich sind.

Um die Diagonalmatrix $D = S^{-1} A S$ bzw. $A = S D S^{-1}$ zu bestimmen:

1. Eigenwerte λ_i von A bestimmen $\rightarrow$ *Nullstellen char. Polynom*
2. Eigenvektoren $\vec{v_i}$ zu λ_i bestimmen $\rightarrow$ *Spalten der Matrix S*
3. Diagonalmatrix $D = diag(\lambda_1, \lambda_2, \ldots \lambda_i)$ bestimmen

1.5.2 Orthogonale Diagonalisierbarkeit

Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt orthogonal diagonalisierbar, falls es eine orthogonale Matrix $S \in \mathbb{R}^{n \times n}$ gibt, so dass $D = S^T AS = S^{-1}AS$ eine Diagonalmatrix ist ($S^T S = I \Rightarrow$ Orthogonalität $S^{-1} = S^T$).

Dies ist genau dann der Fall, wenn A symmetrisch ist:

$$\boldsymbol{A^T} = (SDS^T)^T = (S^T)^T D^T S^T = SDS^T = \boldsymbol{A}$$

Vorgehensweise analog zur Diagonalisierbarkeit; zusätzlich müssen die Eigenvektoren $\vec{v_i}$ zu λ_i noch normiert werden ($\tilde{v_i} = \frac{v_i}{||v_i||}$)

1.6 Pseudo-Inverse A^+

Approximation einer inversen Matrix Für nicht-quadratische Matrizen mit Hilfe der Singulärwertzerlegung (siehe 1.7).

$$A^+ = V \cdot \Sigma^{-1} \cdot U^T$$

wobei $\Sigma^{-1} = diag(\sigma_1^{-1}, \dots \sigma_r^{-1})$

Eigenschaften:

- $AA^+A = A$
- $A^+AA^+ = A^+$
- $(AA^+)^T = AA^+ \rightarrow AA^+$ ist symmetrisch
- $(A^+A)^T = A^+A \rightarrow A^+A$ ist symmetrisch
- $A^+ = A^{-1}$, wenn A invertierbar ist
- $A = U\Sigma V^T \Leftrightarrow A^T = V\Sigma U^T$

1.7 Singulärwertzerlegung

$$\underbrace{A}_{\mathbb{R}^{m \times n}} = \underbrace{U}_{\mathbb{R}^{m \times m}} \overbrace{\Sigma}^{\mathbb{R}^{m \times n}} \underbrace{V^T}_{\mathbb{R}^{n \times n}}$$

- U und V sind orthogonale/unitäre Matrizen
- Σ mit $\sigma_1 \geq \sigma_2 \geq \ldots \geq 0$ (sortiert) auf der Hauptdiagonalen

1.7.1 Verfahren

1. Form prüfen: ist A „hochkant“?
$\rightarrow$ sonst aufwendiger zu lösen
Umstellen zu A^T ist möglich, da $(A^T)^T = A$; d.h. $A^T = V\Sigma^T U^T$

$$A = \begin{pmatrix} 1 & 0 \\ 2 & 2 \\ 0 & 1 \end{pmatrix}$$

2. Eigenwerte von $A^T A$ bestimmen
Eigenwerte (≥ 0) über *Nullstellen char. Polynom* bestimmen, absteigend sortieren!

$$A^T A = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$$

$$p_A(\lambda) = \det(A^T A - \lambda I) \overset{!}{=} 0$$

$$(5-\lambda)^2 - 16 = 0$$

$$\lambda_1 = 9, \lambda_2 = 1$$

3. Σ aufstellen
Diagonalmatrix mit $\sigma_i = \sqrt{\lambda_i}$, Rest $= 0$

$$\Sigma = \begin{pmatrix} 3 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

4. Spaltenvektoren für V ermitteln
Eigenvektoren zu λ aus 2. bestimmen, normieren und in Matrix V eintragen
Für SVD: V^T bilden

<u>für $\lambda_1 = 9$:</u>

$$\left(\begin{array}{cc|c} 5-9 & 4 & 0 \\ 4 & 5-9 & 0 \end{array}\right) \Leftrightarrow$$

$$\left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right) \Rightarrow v_1^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$v_1 = \frac{v_1^*}{||v_1^*||} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

<u>für $\lambda_2 = 1$:</u>

...

$$v_2 = \frac{v_2^*}{||v_2^*||} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

5. U aufstellen
a) für vorhandene Singulärwerte:
$u_i = \frac{1}{\sigma_i} A v_i$
b) sonst: u_i so finden, dass u_i ONB sind
$\rightarrow$ Kreuzprodukt ($\mathbb{R}^3$)
$\rightarrow$ Gram-Schmidt

$$u_1 = \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}, u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

b) $u_3 = u_1 \times u_2 = \frac{1}{6} \begin{pmatrix} 4 \\ -2 \\ 4 \end{pmatrix}$

2 Optimierung

2.1 Konvexe Funktionen/Mengen

2.2 Optimierung ohne Nebenbedingungen

2.2.1 Gradientenabstiegsverfahren

Inkl. stochastisches Gradientenabstiegsverfahren (3.3.4)

2.2.2 Newton-Verfahren

2.3 Optimierung unter Nebenbedingungen

Lagrangemultiplikatoren
Dualität
KKT-Bedingungen