

Eigenwerte und positiv semidefinite Matrizen

Definition 11.1 Sei $A \in \mathbb{R}^{n \times n}$. Eine Zahl $\lambda \in \mathbb{R}$ heißt Eigenwert von A falls es ein $x \in \mathbb{R}^n \setminus \{0\}$ gibt mit $Ax = \lambda x$

$$\Leftrightarrow Ax - \lambda x = 0 \Leftrightarrow (A - \lambda I) \cdot x = 0 \Leftrightarrow \det(A - \lambda I) = 0$$

(Regel der 6. Vorlesung, Seite ③)

Die Abb. $\lambda \mapsto \det(A - \lambda I)$ heißt **charakteristisches Polynom**.

Vektoren $x \neq 0$ mit $A \cdot x = \lambda x$ heißen Eigenvektoren von A zum Eigenwert λ .

Bsp:

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 6 & 2 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix} = 6 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Eigenvektor Eigenwert

Falls λ ein Eigenwert von A ist, ist $\{x \mid Ax = \lambda x\}$ der Eigenraum zum Eigenwert λ .

(Auch wenn $x = 0$ kein Eigenvektor ist, liegt $x = 0$ trotzdem im Eigenraum.)

Erinnerung: $a, b \in \mathbb{R}^n$ stehen senkrecht zueinander, wenn $a^T b = 0$ gilt. ②

- Seien nun $u^{(1)}, \dots, u^{(n)} \in \mathbb{R}^n$ Vektoren der Länge 1

$$\left(\|u^{(i)}\|_2 = \sqrt{(u^{(i)})^T u^{(i)}} = \sqrt{(u_1^{(i)})^2 + (u_2^{(i)})^2 + \dots + (u_n^{(i)})^2} = 1 \right)$$

die paarweise senkrecht zueinander stehen.

$$\text{also } (u^{(i)})^T u^{(j)} = 0 \text{ für } i \neq j$$

$$(u^{(i)})^T u^{(i)} = 1$$

⊗

$$\text{Bsp. } \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{e_1}, \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}}_{e_2}, \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}}_{e_3}, \underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}}_{e_4} \in \mathbb{R}^4 \text{ oder } \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \in \mathbb{R}^2 \dots$$

Bilde die Matrix $U = (u^{(1)}, \dots, u^{(n)})$

Dann gilt

$$\left(U^T U \right)_{ij} = \left(\begin{pmatrix} (u^{(1)})^T \\ \vdots \\ (u^{(n)})^T \end{pmatrix} \begin{pmatrix} u^{(1)} & \dots & u^{(n)} \end{pmatrix} \right)_{ij} = (u^{(i)})^T u^{(j)} \quad \otimes$$

$$\text{bzw. } U^T U = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = I$$

U^T ist die Inverse Matrix von U , $U^{-1} = U^T$ (3)

d.h. $U^T U = I = U U^T$.

Solche Matrizen heißen orthogonale Matrizen.

- Falls nun eine orthogonale Matrix U und eine Diagonalmatrix D gegeben sind, so ist

$A := U D U^T$ eine symmetrische Matrix, denn

$$A^T = (U D U^T)^T = U^T D^T U = U D U^T = A.$$

Reihenfolge vertauschen und
die einzelnen Matrizen transponieren

$$\text{Ferner: } A \cdot u^{(i)} = U D U^T u^{(i)} = U D \underbrace{\begin{pmatrix} (u^{(1)})^T \\ \vdots \\ (u^{(n)})^T \end{pmatrix}}_{e_i} u^{(i)} \\ \underbrace{D_{ii} \cdot e_i}_{D_{ii} \cdot u^{(i)}}$$

Also ist $u^{(i)}$ ein Eigenvektor zum Eigenwert D_{ii}

A hat also n Eigenvektoren, die paarweise senkrecht aufeinander stehen.

Man nennt die Zerlegung $A = UDU^T$ 4
mit Diagonalmatrix D und orthogonaler Matrix U
daher Eigenwertzerlegung von A .

Diese ist nicht eindeutig, aber D ist bis auf
Vertauschung von Diagonalelementen eindeutig.

Satz 11.2 Falls $A = A^T \in \mathbb{R}^{n \times n}$ beliebig gegeben ist,
so existieren eine orthogonale Matrix U und eine
Diagonalmatrix D mit $A = UDU^T$.

Die Berechnung von U, D ist etwas "technisch" aber
Softwareprogramme können das "sehr gut" (stabil und
mit moderatem Aufwand) berechnen.

Reelle symmetrische $n \times n$ -Matrizen haben also immer n
Eigenvektoren, die paarweise senkrecht zueinander stehen.

Bei nichtsymmetrischen Matrizen ist die Diskussion
komplizierter:

$$\begin{aligned} \text{Bsp } A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} &\Rightarrow \det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 1 \\ -1 & 1-\lambda \end{pmatrix} \\ &= (1-\lambda)^2 + 1 > 0 \quad \forall \lambda \in \mathbb{R} \end{aligned}$$

⇒ Dieses A hat keinen (reellen) Eigenwert 5
und daher auch keinen Eigenvektor!

(Ist auch nicht symmetrisch!)

Eine symmetrische Matrix $A = UDU^T$ mit U, D
wie oben heißt positiv **semi**-definit, wenn

$D_{ii} \geq 0$ für $1 \leq i \leq n$. Dies ist gleichbedeutend
mit $x^T A x \geq 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\}$.

• Wenn $D_{ii} > 0 \quad \forall i$ so folgt

$$x^T A x = (x^T U) D U^T x = \underbrace{(U^T x)^T}_{=: z} D \underbrace{U^T x}_z = z^T D z$$

$$= (z_1, \dots, z_n) \begin{pmatrix} D_{11} & & \\ & D_{22} & \\ & & \ddots \\ & & & D_{nn} \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = (z_1, \dots, z_n) \begin{pmatrix} D_{11} z_1 \\ \vdots \\ D_{nn} z_n \end{pmatrix}$$

$$= \underbrace{z_1^2}_{\geq 0} \underbrace{D_{11}}_{\geq 0} + \dots + \underbrace{z_n^2}_{\geq 0} \underbrace{D_{nn}}_{\geq 0} \geq 0$$

• Umkehrung : Übung.

Die Hesse-Matrix

⑥

Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$ stetig differenzierbar, d.h.

$$x \mapsto \nabla f(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} f(x) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x) \end{pmatrix} \text{ sei stetig.}$$

f heißt 2-mal stetig differenzierbar, wenn für jedes feste i die Abbildung $x \mapsto \frac{\partial}{\partial x_i} f(x)$

erneut stetig differenzierbar ist. Die Ableitung ist

$$\text{dann } \left[\frac{\partial}{\partial x_1} \left(\frac{\partial}{\partial x_i} f(x) \right), \dots, \frac{\partial}{\partial x_n} \left(\frac{\partial}{\partial x_i} f(x) \right) \right] \in \mathbb{R}^{1 \times n}$$

Diese Ableitungen fasst man in der Hesse-Matrix

$$\nabla^2 f(x) = D^2 f(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_1} f(x), \dots, \frac{\partial}{\partial x_n} \frac{\partial}{\partial x_1} f(x) \\ \vdots \\ \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_n} f(x), \dots, \frac{\partial}{\partial x_n} \frac{\partial}{\partial x_n} f(x) \end{pmatrix}$$

Zusammen.

Satz: Bei 2-maliger stetiger Differenzierbarkeit ist die Hessematrix symmetrisch.

7

Charakteristische Gleichung

Für kleine $\|\Delta x\|$ gilt

$$f(x+\Delta x) = f(x) + Df(x)\Delta x + \frac{1}{2}\Delta x^T D^2f(x)\Delta x + \underbrace{o(\|\Delta x\|^2)}$$

"noch viel kleiner
als $\|\Delta x\|^2$ "

Satz 11.3 (Lokale Minimalstellen)

- Wenn $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 2-mal stetig diff'bar ist und $\bar{x}$ eine lokale Minimalstelle von f ist so ist $\nabla f(\bar{x}) = 0$ und $D^2f(\bar{x})$ positiv semidefinit. Wenn umgekehrt ein Punkt $\tilde{x}$ gegeben ist mit $\nabla f(\tilde{x}) = 0$ und positiv definiten Matrix $D^2f(\tilde{x})$ so ist $\tilde{x}$ eine strikte lokale Minimalstelle von f .

$$\underbrace{f(\bar{x})}_{\text{min}} \leq \underbrace{f(\bar{x}+\Delta x)}_{\text{min}} = \underbrace{f(\bar{x})}_{\text{min}} + \underbrace{Df(\bar{x})\Delta x}_{=0} + \frac{1}{2}\Delta x^T D^2f(\bar{x})\Delta x + o(\|\Delta x\|^2)$$

$$0 \leq 0 + 0 + \underbrace{\frac{1}{2}\Delta x^T D^2f(\bar{x})\Delta x}_{\geq 0} + \underbrace{o(\|\Delta x\|^2)}_{\neq}$$