Beispiele • $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$

Gesucht: Diagonalisierung von A

$$\chi_A = \det\begin{pmatrix} x-2 & -1 & -1 \\ -1 & x-2 & -1 \\ -1 & -1 & x-2 \end{pmatrix} = (x-2)^3 - 1 - 1 - 3(x-2) =$$

$$= x^3 - 6x^2 + 9x - 4 = (x-1)(x^2 - 5x + 4) = (x-1)^2(x-4)$$

$\Rightarrow$ A ist ähnlich zu $\operatorname{diag}(1,1,4) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$

Gesucht: orthogonale Transformationsmatrix

dazu: Bestimmung der Eigenräume

E_1: Lösungsraum des homogenen LGS mit Matrix $A - I_3$:

$$E_1 = \left\langle \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\rangle$$

Hierauf wenden wir das Gram-Schmidtsche Orthogonalisierungsverfahren an:

$$u_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\end{pmatrix} \quad \text{und} \quad w_2 = \begin{pmatrix}1\\-1\\0\end{pmatrix} - \underbrace{\frac{1}{\sqrt{2}}}_{\text{Ergebnis Skalarprodukt}} u_1 = \frac{1}{2}\begin{pmatrix}1\\-2\\1\end{pmatrix}$$

$$\Rightarrow u_2 = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\-2\\1\end{pmatrix}$$

E_4: Lösungsraum des homogenen LGS mit Matrix $A - 4 \cdot I_3$:

$$E_4 = \left\langle \begin{pmatrix}1\\1\\1\end{pmatrix} \right\rangle$$

$$\Rightarrow u_3 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}$$

$$\Rightarrow \quad S = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \in O_3(\mathbb{R}) \quad \text{und} \quad S^{-1}AS = \begin{pmatrix}1&0&0\\0&1&0\\0&0&4\end{pmatrix}$$

- Versuch, den Hauptachsentransformationssatz auf $\mathbb{C}$ anzuwenden statt $\mathbb{R}$

$$A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} \in \mathbb{C}^{2\times 2}$$

$$\Rightarrow \chi_A = \det\begin{pmatrix} x-1 & -i \\ -i & x+1 \end{pmatrix} = (x-1)(x+1) - (-i)^2 =$$

$$= x^2 \cancel{-x} \cancel{+x} - 1 + 1 = x^2$$

$\Rightarrow$ einziger Eigenwert 0 mit $m_a(0) = 2$, aber: geometrische Vielfachheit ist $m_g(0) = 1$.

$\Rightarrow$ Satz gilt nicht für $\mathbb{C}$ statt $\mathbb{R}$

Matrix hat rang 1

$$A \cdot x = 0 \cdot x = 0$$

$$\begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}\begin{pmatrix} -i \\ 1 \end{pmatrix} = 0$$

<u>Definition</u> $A \in \mathbb{R}^{n \times n}$ symmetrisch heißt

- <u>positiv definit</u>, falls alle Eigenwerte positiv sind
- <u>positiv semidefinit</u>, falls alle Eigenwerte ≥ 0 sind
- <u>negativ definit</u>, falls alle Eigenwerte negativ sind
- <u>negativ semidefinit</u>, falls alle Eigenwerte ≤ 0 sind
- <u>indefinit</u>, falls es positive und negative Eigenwerte gibt

<u>Satz</u> $A \in \mathbb{R}^{n \times n}$ symmetrisch ist genau dann positiv definit, wenn für alle $v \in \mathbb{R}^n \setminus \{0\}$ gilt $\langle v, Av \rangle > 0$.
A ist positiv semidefinit, wenn $\langle v, Av \rangle \geq 0$ gilt.
Entsprechendes gilt für negativ (semi)definit.

Beweis Es gibt $S \in O_n(\mathbb{R})$ mit

$SDS^T = A$

$$\underline{S^T A S} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \underline{D},$$ wobei die λ_i die Eigenwerte von A sind.

Wegen $S^T \in GL_n(\mathbb{R})$ ist für jeden Vektor $v \in \mathbb{R}^n \setminus \{0\}$ auch $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = S^T v \neq 0$ und jeder Vektor aus $\mathbb{R}^n \setminus \{0\}$ tritt als ein $S^T v$ mit $v \in \mathbb{R}^n \setminus \{0\}$ auf.

$(S^T)^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = v$

Es gilt $\langle v, Av \rangle = v^T A v = v^T \underline{(SDS^T)} v = (x_1, \ldots, x_n) \cdot D \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n \underline{\lambda_i} x_i^2.$

$\underbrace{v^T S}_{= (S^T v)^T} \cdot D \cdot S^T v$

$\square$

wenn alle λ_i positiv sind, dann ist die Summe nur dann $= 0$, wenn alle $x_i = 0$

Beispiel Sei $A = \begin{pmatrix} a & 0 & -a & 0 \\ 0 & b & 0 & -b \\ -a & 0 & a & 0 \\ 0 & -b & 0 & b \end{pmatrix}$ mit $a, b \in \mathbb{R}$

Für $v = \begin{pmatrix} x_1 \\ \vdots \\ x_4 \end{pmatrix} \in \mathbb{R}^4$ gilt $\langle v, Av \rangle = v^T A v = (x_1, \ldots, x_4) \cdot \begin{pmatrix} a(x_1 - x_3) \\ b(x_2 - x_4) \\ -a(x_1 - x_3) \\ -b(x_2 - x_4) \end{pmatrix}$

$= a(x_1 - x_3)^2 + b(x_2 - x_4)^2$

$x_1 \cdot a(x_1 - x_3) + x_2 \cdot b(x_2 - x_4) + x_3 \cdot (-a)(x_1 - x_3) + x_4 \cdot (-b)(x_2 - x_4)$

$\Rightarrow$ A ist positiv semidefinit, falls $a, b \geq 0$, negativ semidefinit, falls $a, b \leq 0$ und sonst indefinit

Test auf pos./neg. (Semi) definitheit <u>kann</u> sehr viel weniger Arbeit benötigen als die direkte Bestimmung der Eigenwerte

13 Anwendung: Hauptkomponentenanalyse

Ziel: Identifikation einer geeigneten Basis zur Darstellung einer gegebenen Punktmenge

Datenanalyse: (1) Feature Representation
(2) Distance Metric
(3) Clustering / Classification

wichtig, schwer

gute Ergebnisse von einfachen Standardalgorithmen, falls (1) & (2) qualitativ hochwertig

Die Dimension der Datenmenge kann sehr hoch sein.

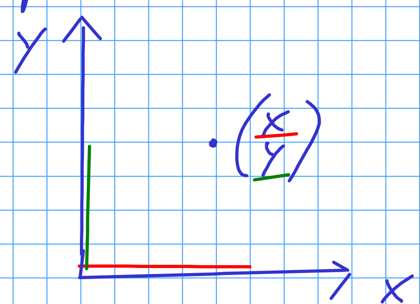
Ziel dann: Dimensionsreduktion unter möglichst wenig Informationsverlust

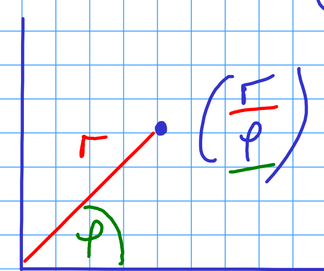
Input: Datenpunkte $x_1,\dots,x_n \in \mathbb{R}^m$, $X=(x_1,\dots,x_n)\in\mathbb{R}^{m\times n}$,
diese sind dargestellt über Standardbasis B
(typischerweise da die Punkte so gemessen wurden)

Output: Basis C für den $\mathbb{R}^m$ (später: niedrigere Dimension),
die die Daten „besser" darstellt.
C soll durch lineare Transformation aus B bestehen,
so wie das bei unseren Basiswechseln der Fall war

Einschränkung

Gegenbeispiel: Verschiedene Arten der Darstellung von 2D-Vektoren





$X \in \mathbb{R}^{m \times n}$ alte Datenpunkte

$Y \in \mathbb{R}^{m \times n}$ neue Darstellung der Datenpunkte und

$A \cdot X = Y$

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \qquad Y = (y_1, \dots, y_n)$$

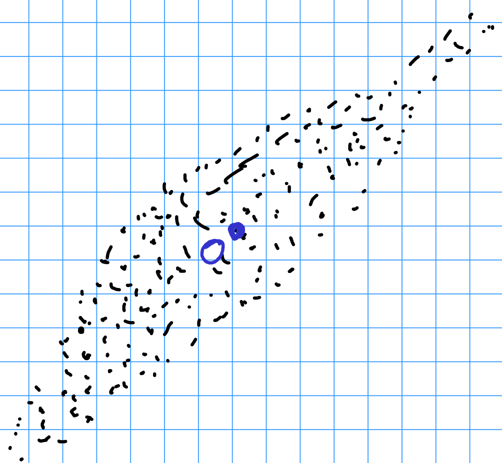
$\Rightarrow$

- A transformiert X zu Y
- Geometrisch interpretiert ist A eine Rotation und Streckung von X zu Y
- Die Zeilen von A, also $\{a_1, \dots, a_m\}$ sind eine Menge von Basisvektoren zur Darstellung von Y:

$$A \cdot X = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} (x_1, \dots, x_n) = Y = \begin{pmatrix} a_1^T x_1 & \dots & a_1^T x_n \\ \vdots & & \\ a_m^T x_1 & \dots & a_m^T x_n \end{pmatrix}$$

$$\Rightarrow \quad Y_i = \begin{pmatrix} a_1^T x_i \\ \vdots \\ a_m^T x_i \end{pmatrix}$$

Wie wählt man die neue Basis? Beste Antworten: Statistik



Wahl des Nullpunkts als arithmetisches Mittel aller Punkte

meist: Interpretation einer Datenwolke als eine Menge von <u>Zufallsvariablen</u> zu einer zu Grunde liegenden <u>Verteilung</u>

Fachbegriffe aus der Statistik

→ Rechtfertigung für diese Wahl des Nullpunkts, denn dann hat die zugehörige Verteilung einen Erwartungswert im Nullpunkt