Vorlesung: Numerik 1 für Ingenieure

Version 30.10.23

Michael Karow

4. Vorlesung

Thema: Matrixgeometrie und Singulärwertzerlegung

Ziel der nächsten beiden Vorlesungen:

Verständnis der Fehlerformel für lineare Gleichungssysteme.

Gegeben: Invertierbare Matrizen $A, \tilde{A} \in \mathbb{R}^{n\times n}$ und Vektoren $b, \tilde{b}, x, \tilde{x} \in \mathbb{R}^n$, $b \neq 0$, so dass

$$Ax = b, \qquad \tilde{A}\tilde{x} = \tilde{b}. \quad \text{(gestörtes Gleichungssystem)}$$

Dann gilt für die **relativen Fehler**,

$$\boxed{\frac{\|\tilde{x}-x\|}{\|x\|} \le \frac{\text{cond}(A)}{1-\text{cond}(A)\frac{\|\tilde{A}-A\|}{\|A\|}} \left(\frac{\|\tilde{A}-A\|}{\|A\|} + \frac{\|\tilde{b}-b\|}{\|b\|}\right)}$$

sofern der Nenner > 0 ist. Dabei ist $\|x\|$ eine beliebige **Norm**,

$$\|A\| = \max_{x\neq 0} \frac{\|Ax\|}{\|x\|}$$

die zugehörige **Matrixnorm** und

$$\text{cond}(A) = \|A\|\,\|A^{-1}\|. \qquad \textbf{(Konditionszahl)}$$

Eine lineare Abbildung $x \longmapsto Ax$ bildet Geraden auf Geraden oder Punkte ab.

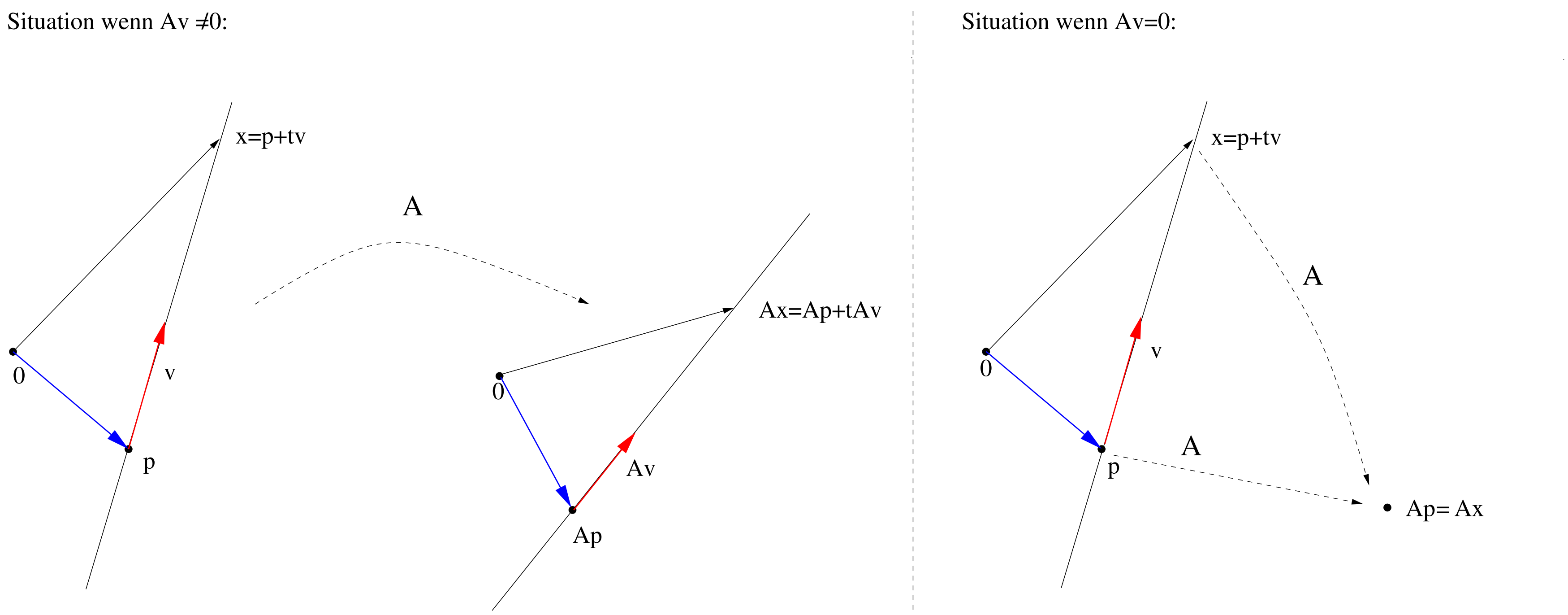
Erläuterung: Die Gerade durch den Punkt $p \in \mathbb{R}^n$ in Richtung $v \in \mathbb{R}^n, v \neq 0$, ist die Menge aller $x \in \mathbb{R}^n$ mit

$$x = p + t\,v, \qquad t \in \mathbb{R}.$$

Multipliziert man alle x auf der Geraden mit $A \in \mathbb{R}^{m\times n}$, dann bekommt man die Menge aller

$$Ax = Ap + t\,Av, \qquad t \in \mathbb{R}. \qquad (**)$$

Dies ist die Gerade durch den Punkt Ap mit Richtungsvektor Av, es sei denn $Av = 0$. In diesem Fall ist die Bildmenge $(**)$ nur der Punkt Ap.



Bemerkung: Der Fall $v \neq 0$ und $Av = 0$ kommt nicht vor, wenn A invertierbar ist.

Eine lineare Abbildung $x \longmapsto Ax$ bildet zueinander parallele Geraden auf auf zu einander parallele Geraden oder Punkte ab.

Erläuterung: Zwei Geraden sind zu einander parallel, wenn sie denselben Richtungsvektor haben (oder die Richtungsvektoren Vielfache von einander sind), z.B.

$$\text{Gerade 1:} \qquad x_1 = p_1 + t\,v, \qquad t \in \mathbb{R}.$$

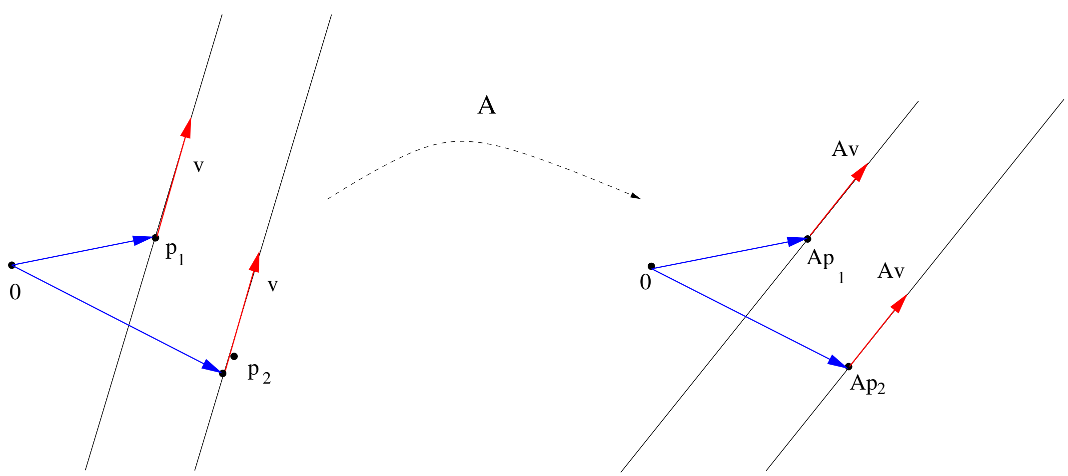
$$\text{Gerade 2:} \qquad x_2 = p_2 + t\,v, \qquad t \in \mathbb{R}.$$

Die Bildgeraden sind

$$\text{Bildgerade 1:} \qquad Ax_1 = Ap_1 + t\,Av, \qquad t \in \mathbb{R}.$$

$$\text{Bildgerade 2:} \qquad Ax_2 = Ap_2 + t\,Av, \qquad t \in \mathbb{R}.$$

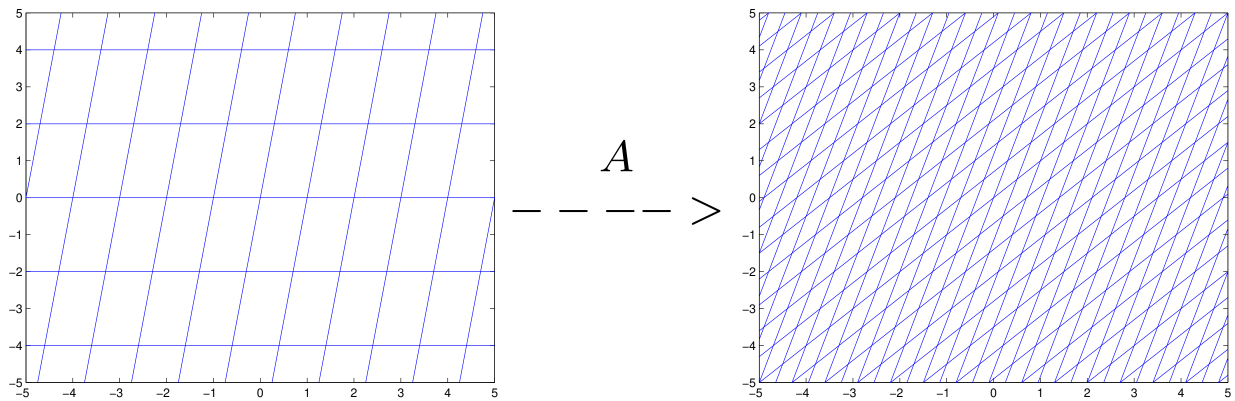
Die Richtungsvektoren der Bildgeraden sind wieder gleich. Die Bildgeraden sind also auch parallel.



Weitere Tatsache: Geht eine Gerade durch 0, dann geht auch die Bildgerade durch 0.

Folgerung aus den bisher genannten Tatsachen:

Eine invertierbare lineare Abbildung $x \mapsto Ax, x \in \mathbb{R}^n$,
macht aus einem Gitter paralleler Geraden wieder ein Gitter paralleler Geraden.



Das lineare Bild des Einheitsgitters

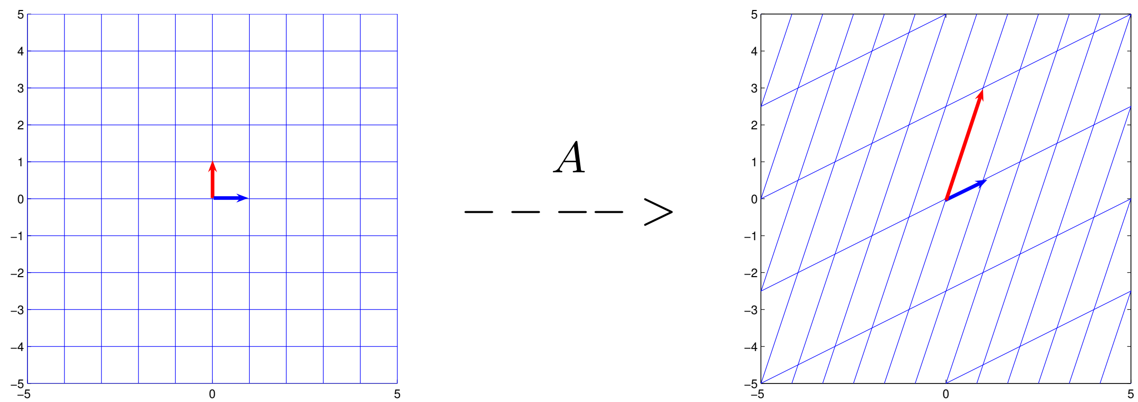
Das Bild des Einheitsgitters kann man direkt an der Matrix ablesen.

Es gilt nämlich: **Die Spalten von A sind die Bilder der kanonischen Basisvektoren.**

Im 2×2 Fall:

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{A} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \qquad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix}$$

Illustration:

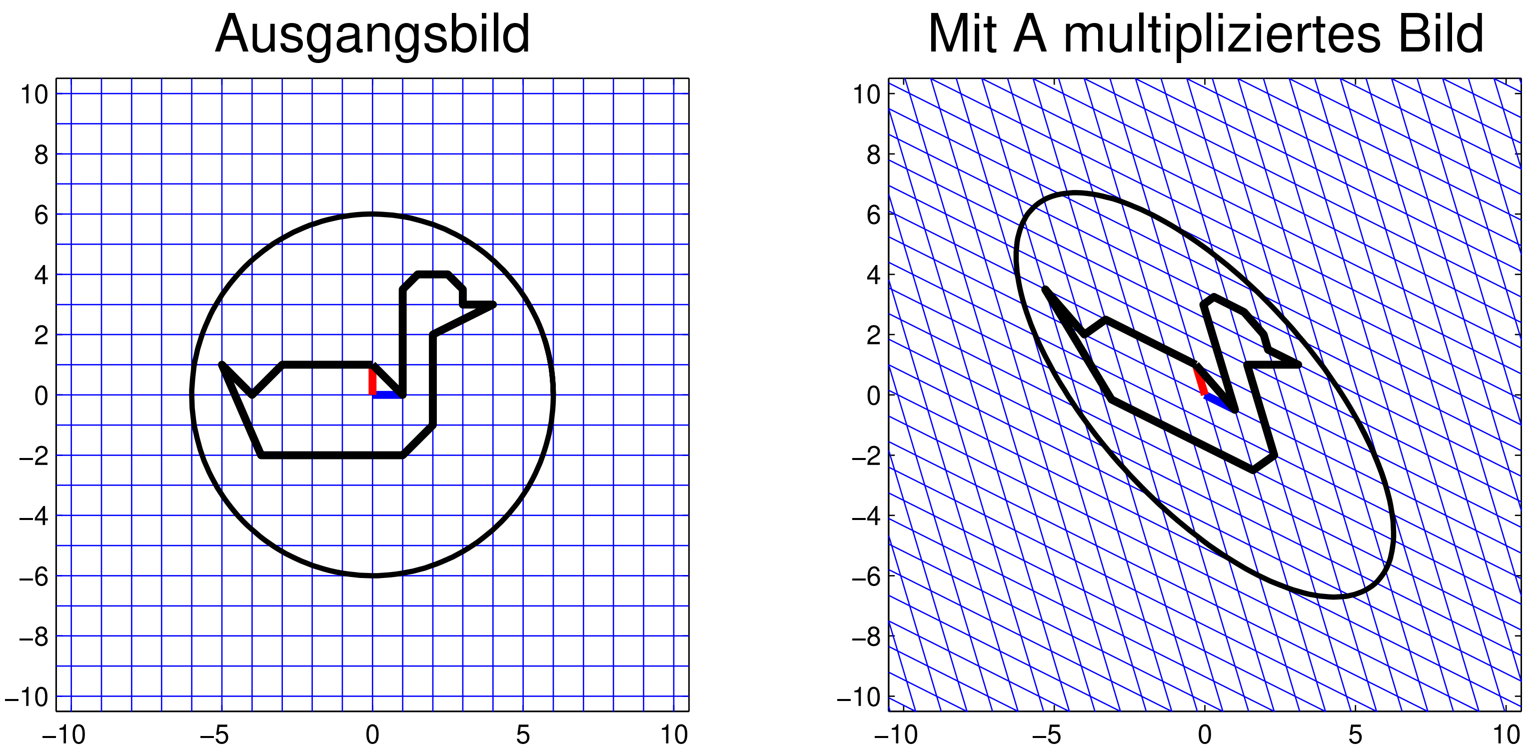


Linkes Bild: Einheitsgitter mit den kanonischen Basisvektoren

Rechtes Bild: Das A-Bild des Einheitsgitters mit den A-Bildern der Basisvektoren.

Die Matrix in diesem Beispiel ist $A = \begin{bmatrix} 1 & 1 \\ 0.5 & 3 \end{bmatrix}$.

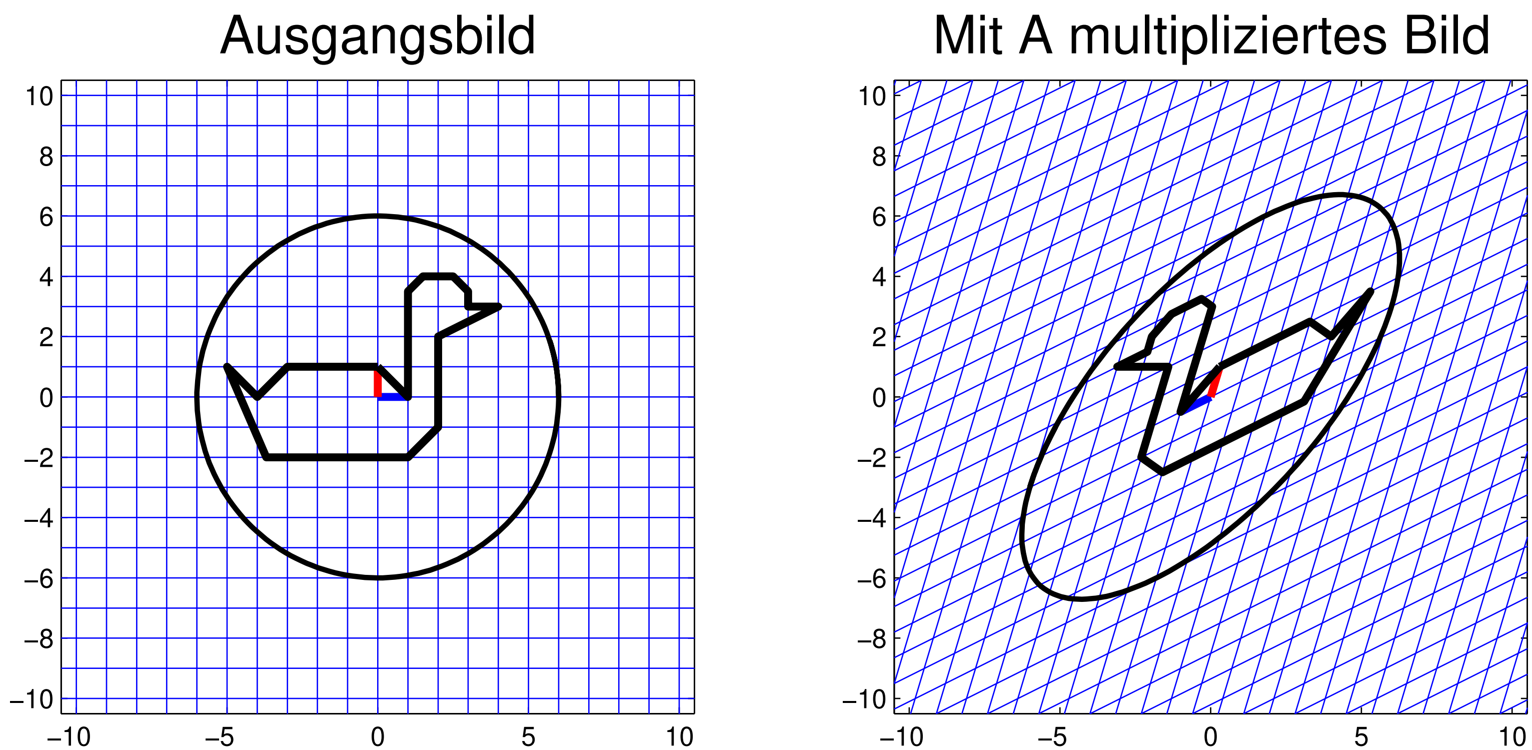
Beispiel: Lineare Transformation von Figuren I



In diesem Beispiel ist $A = \begin{bmatrix} 1 & -0.5 \\ -0.3 & 1 \end{bmatrix}$.

Bemerkung: Der Flächeninhalt einer beliebigen Figur multipliziert sich beim Übergang $x \mapsto Ax$ mit dem Faktor $|det(A)|$.

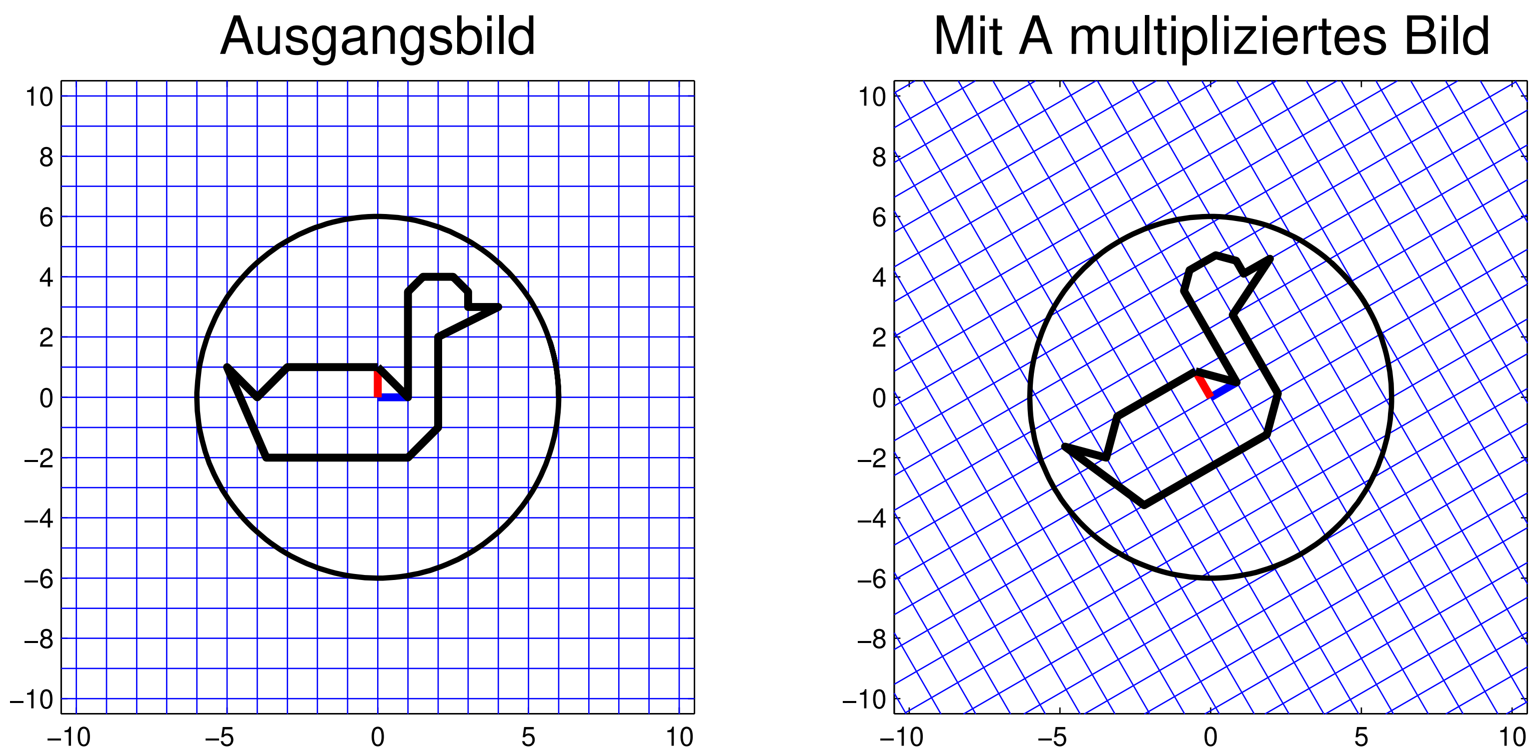
Beispiel: Lineare Transformation von Figuren II



In diesem Beispiel ist $A = \begin{bmatrix} -1 & 1 \\ -0.5 & 3 \end{bmatrix}$. Es ist $det(A) < 0$.

Matrizen mit negativer Determinante bewirken eine Umkehrung der Orientierung.

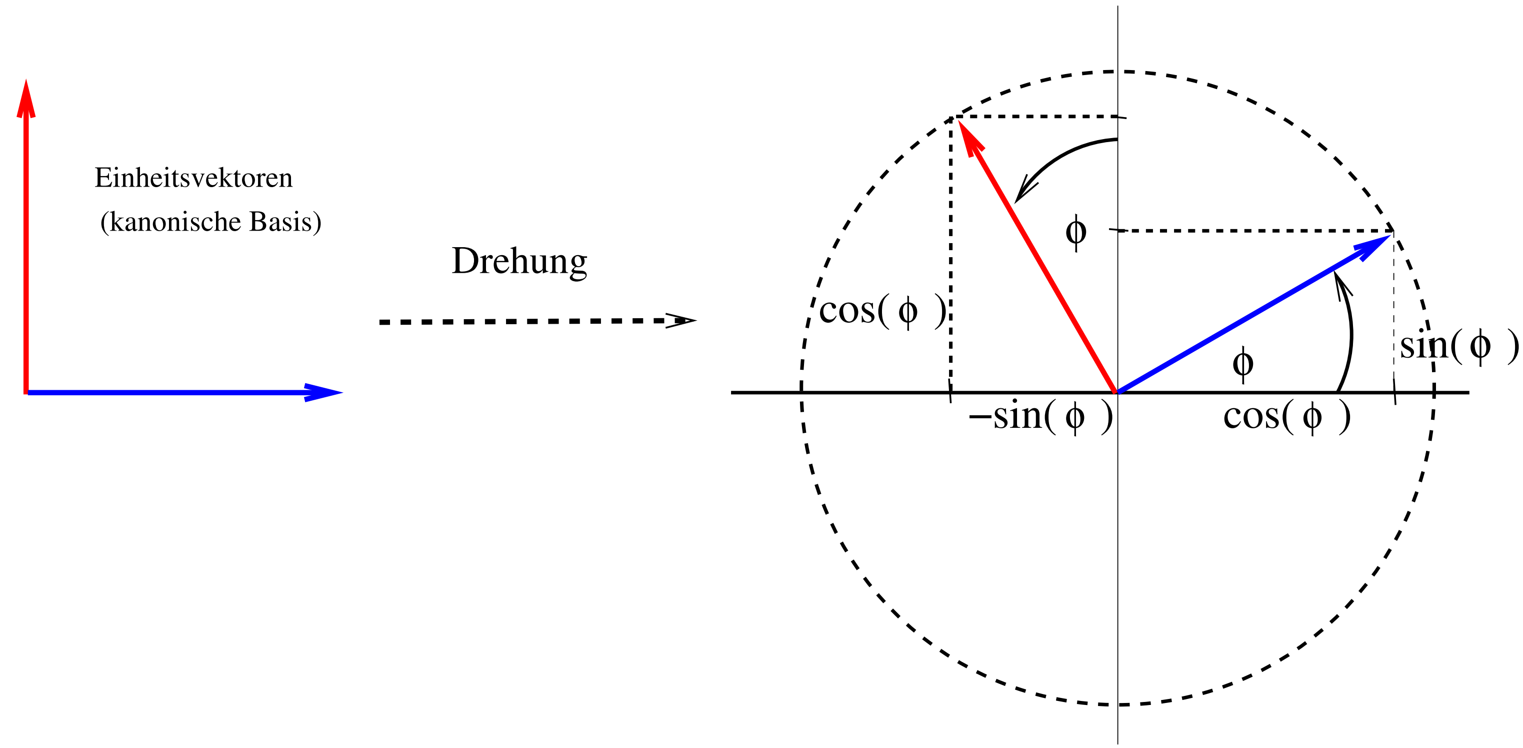
Beispiel: Lineare Transformation von Figuren III



In diesem Beispiel ist $A = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}, \ \ \phi = \pi/6.$

A ist eine **Drehmatrix**.

Konstruktion von Drehmatrizen im $\mathbb{R}^2$



Die Bilder der kanonischen Basisvektoren bei Drehung um den Winkel ϕ gegen den Uhrzeigersinn sind

$$\begin{bmatrix} \cos(\phi) \\ \sin(\phi) \end{bmatrix}, \quad \begin{bmatrix} -\sin(\phi) \\ \cos(\phi) \end{bmatrix}.$$

In den Spalten der zugehörigen Drehmatrix stehen die Bilder der Basisvektoren.

Die Drehmatrix ist also

$$A = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}.$$

Das Standard-Skalarprodukt

Das Skalarprodukt zweier Vektoren $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \;\; y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ ist bekanntlich

$$x \cdot y = x^T y = x_1\, y_1 + \ldots + x_n\, y_n \; = \sum_{j=1}^{n} x_j\, y_j.$$

Euklidische Länge eines Vektors:

$$\|x\|_2 \;=\; |x| \;=\; \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = \sqrt{x^T x}. \qquad (1)$$

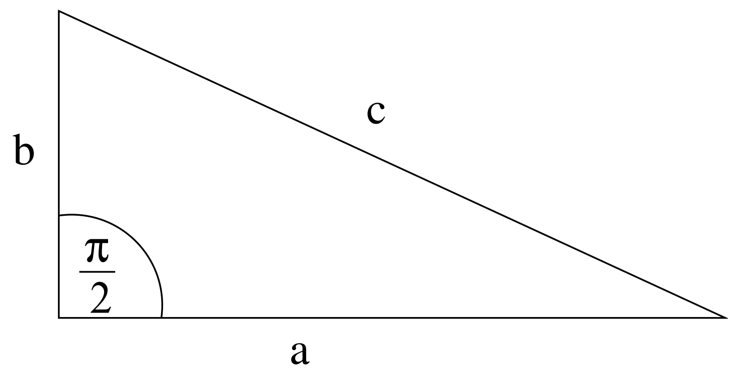
Winkel zwischen x und y:

$$\phi = \text{arccos}\left(\frac{x^T y}{\|x\|_2\, \|y\|_2}\right). \qquad (2)$$

Die Gleichungen (1) und (2) sind Definitionen, die durch die Elementargeometrie motiviert sind. (1) ist der Satz des Pythagoras. (2) folgt für Vektoren aus dem Anschauungsraum aus dem Cosinussatz (siehe die nächsten Seiten).

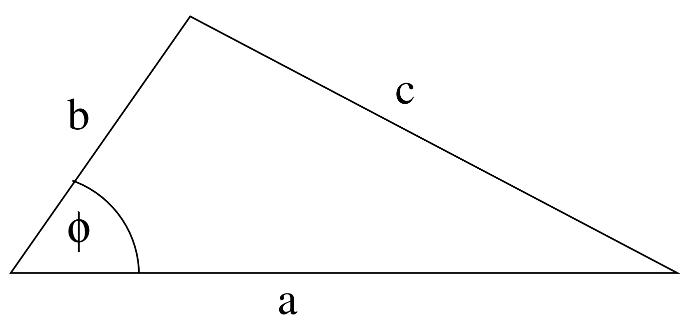
Der Cosinussatz ist eine Verallgemeinerung des Pythagoras auf beliebige Winkel.

Pythagoras:



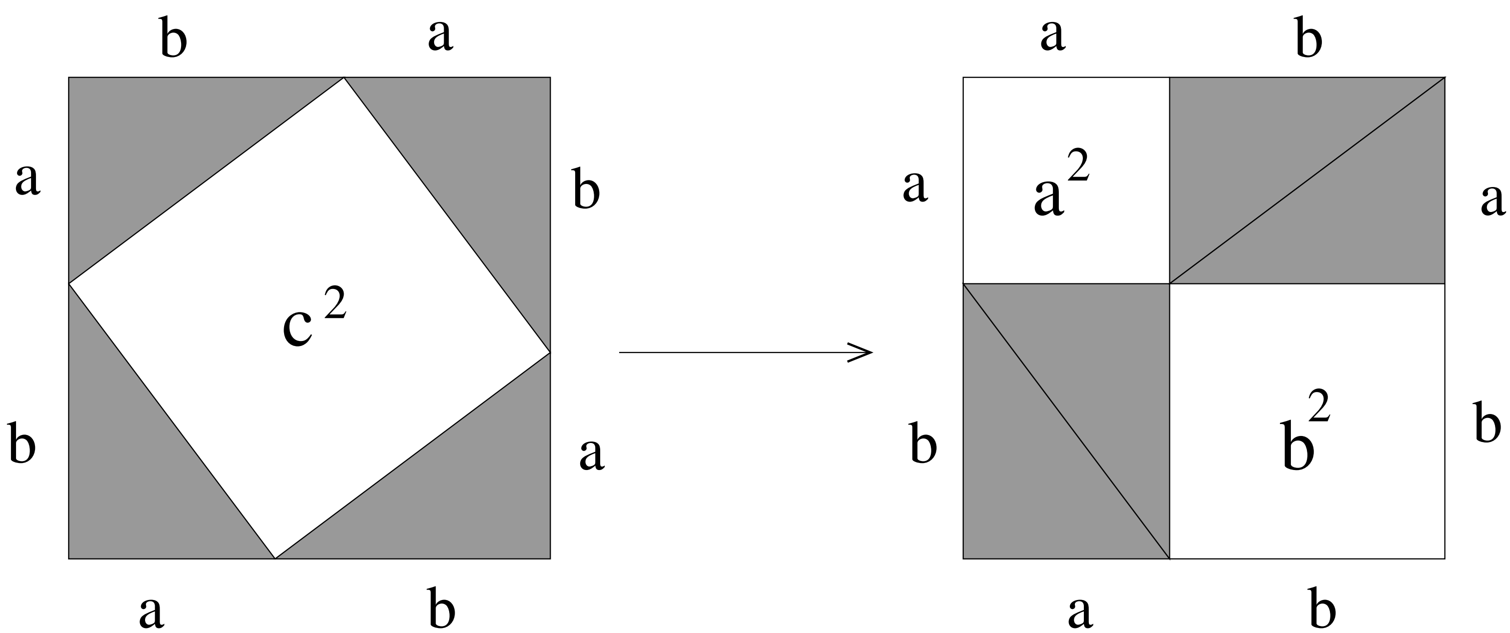
$$c^2 = a^2 + b^2$$

Cosinussatz:



$$c^2 = a^2 + b^2 - 2\,ab\,\cos(\phi)$$

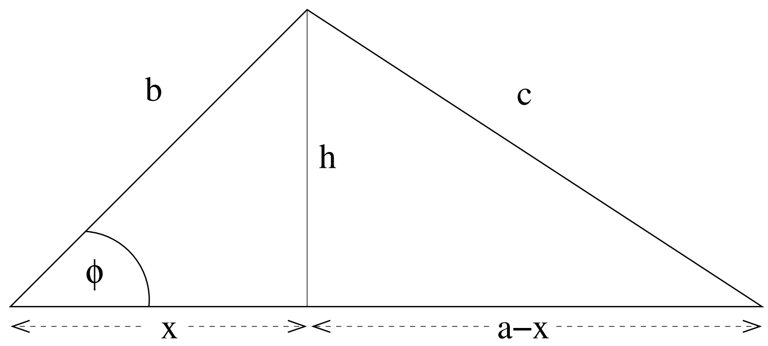
Beweis des Satzes von Pythagoras



Bemerkung.
Alle natürlichen Zahlentripel (a, b, c) so dass $a^2 + b^2 = c^2$ können geschrieben werden als

$$a = (u^2 - v^2)^2, \qquad b = (2uv)^2, \qquad c = (u^2 + v^2)^2, \qquad u, v = 0, 1, 2, \ldots.$$

Zum Beweis des Cosinussatzes (hier nur für spitze Winkel)



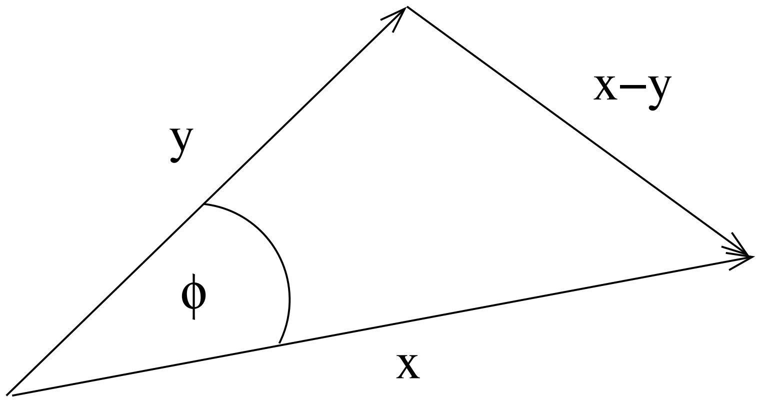
$$\begin{aligned} c^2 &= (a-x)^2 + h^2 \\ &= (a-x)^2 + b^2 - x^2 \\ &= a^2 - 2ax + x^2 + b^2 - x^2 \\ &= a^2 + b^2 - 2ax \end{aligned} \qquad b^2 = h^2 + x^2$$

Es ist

$$x = b \cos\phi.$$

$\Rightarrow$ $$c^2 = a^2 + b^2 - 2ab\cos\phi$$

Winkel zwischen zwei Vektoren und der Cosinussatz.



Nach dem Cosinussatz:

$$\|x-y\|_2^2 = \|x\|_2^2 + \|y\|_2^2 - 2\,\|x\|_2\,\|y\|_2\,\cos\phi. \qquad (1)$$

Mit direkter Rechnung folgt:

$$\begin{aligned}
\|x-y\|_2^2 &= (x-y)^T(x-y) \\
&= (x^T-y^T)(x-y) \\
&= x^Tx - x^Ty - y^Tx + y^Ty \qquad \text{(Es ist } x^Ty = y^Tx\text{)} \\
&= \|x\|_2^2 + \|y\|_2^2 - 2\,x^Ty \qquad (2)
\end{aligned}$$

Vergleich von (1) und (2) ergibt

$$\|x\|_2\,\|y\|_2\cos\phi = x^Ty.$$

Also

$$\phi = \arccos\left(\frac{x^Ty}{\|x\|_2\,\|y\|_2}\right).$$

Senkrechte Projektion eines Vektors auf eine Achse

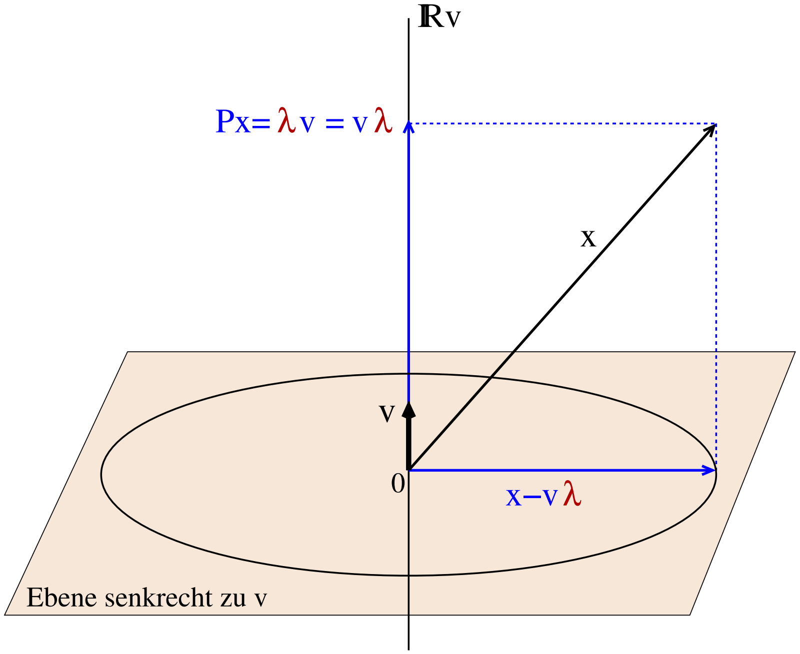
Sei $v \in \mathbb{R}^n$ ein Einheitsvektor ($\|v\|_2 = 1$). Sei $x \in \mathbb{R}^n$ beliebig.
Gesucht ist die senkrechte Projektion Px von x auf die Gerade $\mathbb{R}v$. Genauer formuliert:
Es ist $Px = \lambda x = x\lambda$ mit $\lambda \in \mathbb{R}$ so dass $x - Px = x - v\lambda$ senkrecht auf v steht, also

$$0 = v^\top(x - Px) = v^\top(x - v\lambda) = v^\top x - \underbrace{v^\top v}_{=1}\lambda \qquad \Rightarrow \qquad \lambda = v^\top x.$$

Folglich:

$$Px = v(v^\top x) = (vv^\top)x = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \ddots & v_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} v_1^2 & v_1 v_2 & \dots & v_1 v_n \\ v_2 v_1 & v_2^2 & \dots & v_2 v_n \\ \vdots & & & \vdots \\ v_n v_1 & v_n v_2 & \dots & v_n^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

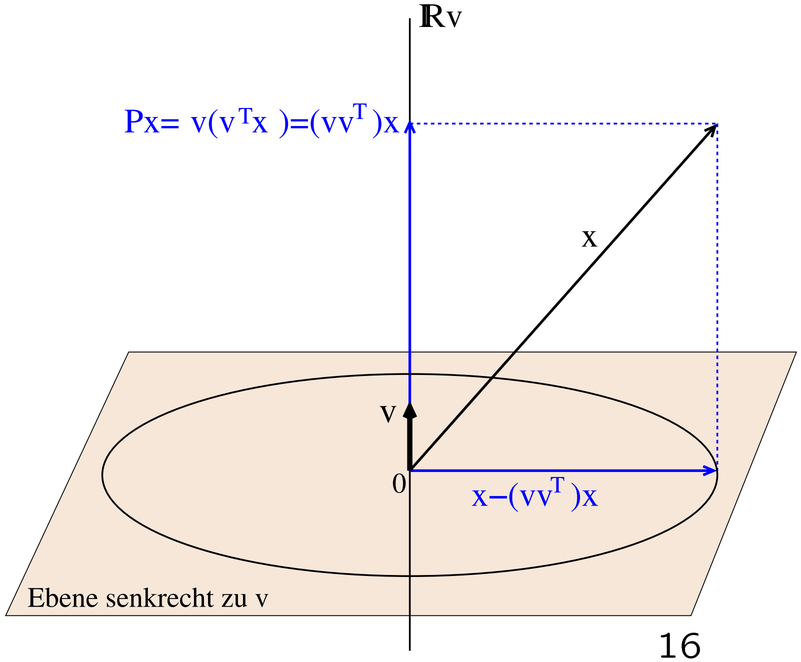
Ansatz:



$$\lambda = v^\top x$$

$$- - - - - >$$

Lösung:



Senkrechte Projektion eines Vektors auf eine Achse

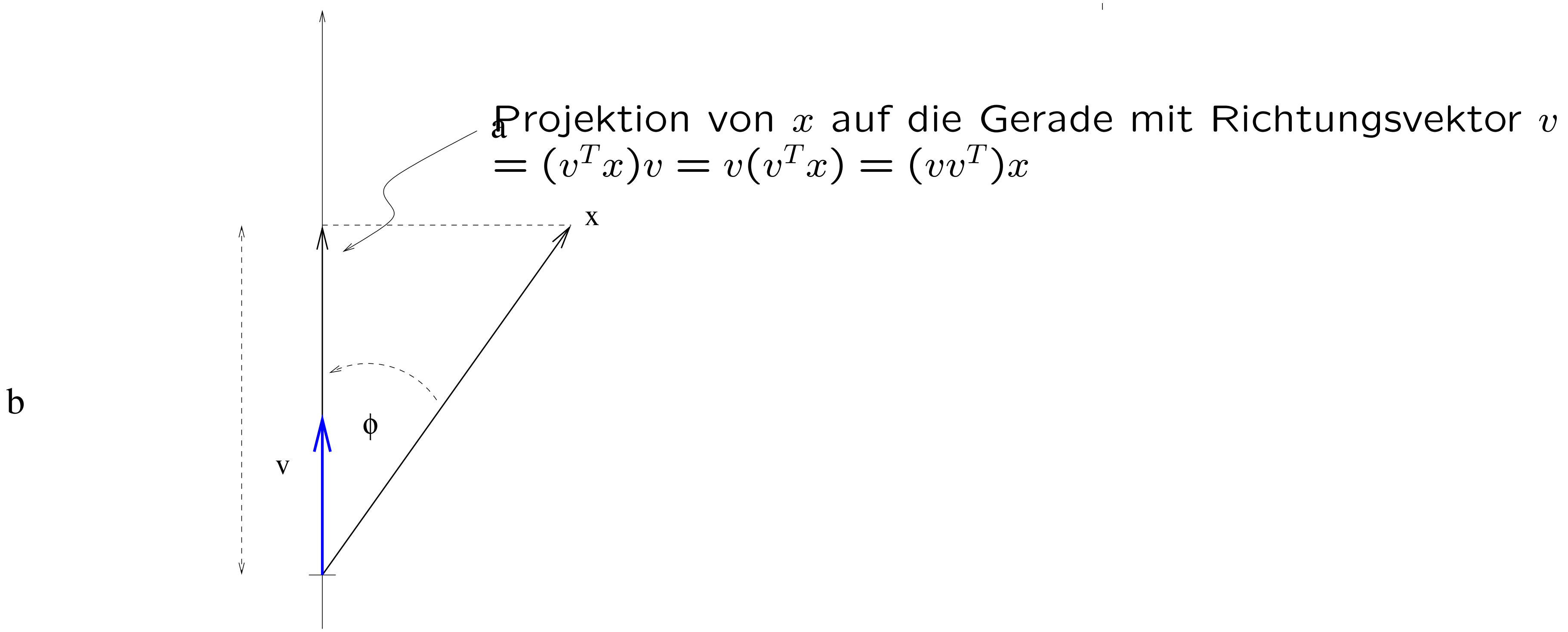
Erinnerung: Für den Winkel ϕ zwischen 2 Vektoren x, v gilt $\cos\phi = \frac{v^T x}{\|v\|_2\,\|x\|_2}$.

Wenn v die Länge 1 hat, dann folgt $v^T x = \|x\|_2 \cos\phi$.

Die senkrechte Projektion von x auf die Gerade durch 0 mit Richtungsvektor v ist dann

$$x \longmapsto (v^T x)\,v = v\,(v^T x) = (v\,v^T)\,x.$$

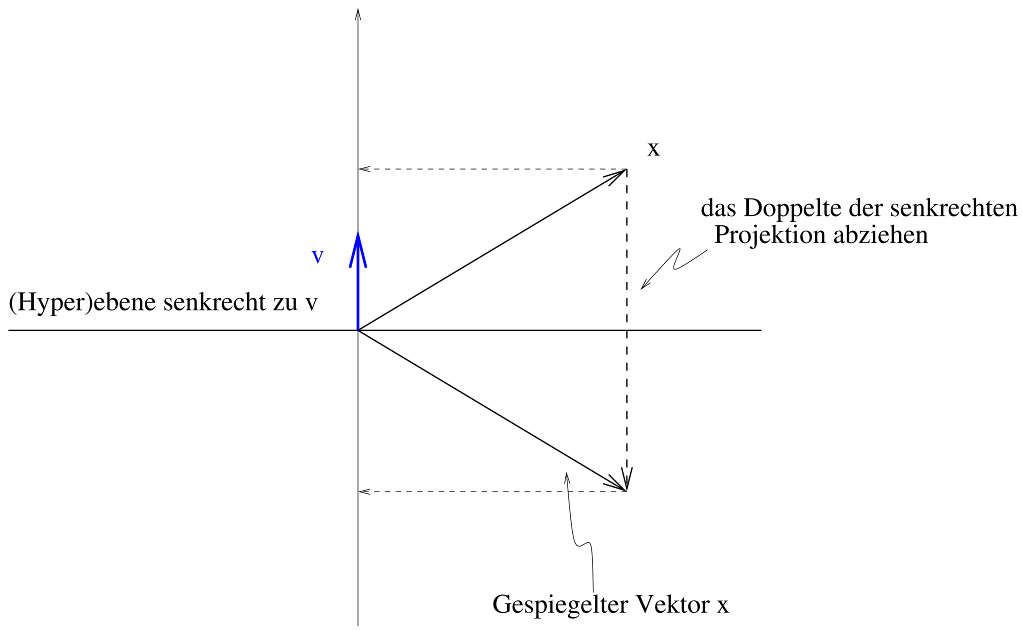
Folgerung: Die Projektion ist eine (nicht invertierbare) lineare Abbildung $x \mapsto Ax$ mit der Matrix $A = v\,v^T$.



Spiegelung an einer (Hyper)ebene

Sei $v \in \mathbb{R}^n$ ein Einheitsvektor (d.h. $\|v\|_2 = 1$).

Wenn man von einem Vektor x das doppelte der senkrechten Projektion von x auf die Gerade durch v abzieht, dann hat man x dadurch an der (Hyper)ebene senkrecht zu v gespiegelt.



Dies ist wieder eine lineare Abbildung: $\quad x \longmapsto x - 2\,vv^T x = \underbrace{(I - 2vv^T)}_{A}\,x.$

Die Spiegelungsmatrix A nennt man auch **Householdermatrix**.

Solche Matrizen sind in der Numerik sehr wichtig. Dazu später mehr.

Spiegelungen in $\mathbb{R}^2$

Wir haben gesehen, dass

$$\text{die Matrizen } R_t = \begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix} \text{ Drehmatrizen sind.}$$

Wir zeigen, dass

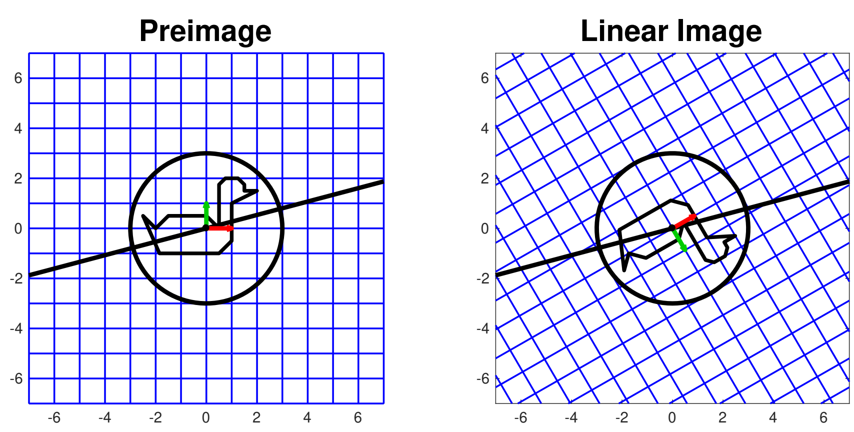
$$\text{die Matrizen } \widetilde{R}_t = \begin{bmatrix} \cos(t) & \sin(t) \\ \sin(t) & -\cos(t) \end{bmatrix} \text{ Spiegelungsmatrizen sind.}$$

Beweis. Sei $c = \cos(t/2), s = \sin(t/2)$.
Wegen der Additionstheoreme für sin und cos und Pythagoras ($c^2 + s^2 = 1$) hat man,

$$\underbrace{\begin{bmatrix} c^2 - s^2 & 2\,sc \\ 2\,sc & s^2 - c^2 \end{bmatrix}}_{\widetilde{R}_t} \underbrace{\begin{bmatrix} c \\ s \end{bmatrix}}_{u} = \underbrace{\begin{bmatrix} c \\ s \end{bmatrix}}_{u}, \qquad \underbrace{\begin{bmatrix} c^2 - s^2 & 2\,sc \\ 2\,sc & s^2 - c^2 \end{bmatrix}}_{\widetilde{R}_t} \underbrace{\begin{bmatrix} -s \\ c \end{bmatrix}}_{v} = -\underbrace{\begin{bmatrix} -s \\ c \end{bmatrix}}_{v}.$$

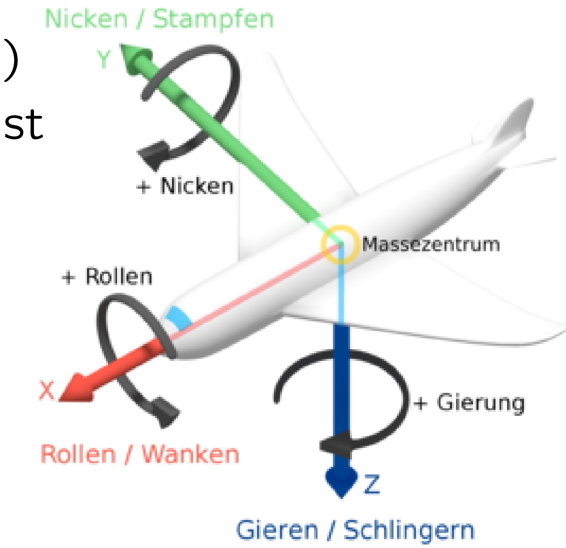
Folgerung: Multiplikation mit $\widetilde{R}_t$ läßt alle Vektoren der gerade $\mathbb{R}u$ fest und multipliziert alle Vektoren auf $\mathbb{R}v$ mit -1. Man rechnet leicht nach, dass $\widetilde{R}_t = I - 2\,vv^\top$.

Beispiel: $\tilde{R}$ for $t = \pi/6$:



Drehungen (Rotationen) im $\mathbb{R}^3$, I

Alle Drehungen im $\mathbb{R}^3$ sind Produkte (Hinereinanderausführung) von Drehungen um x, y und z-Achse. Die Drehmatrix $A \in \mathbb{R}^{3,3}$ ist



$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix}}_{\substack{\text{Rot. um } x\text{-Achse} \\ \text{Wanken} \\ \text{(eng.: roll)}}} \underbrace{\begin{bmatrix} \cos(\beta) & 0 & -\sin(\beta) \\ 0 & 1 & 0 \\ \sin(\beta) & 0 & \cos(\beta) \end{bmatrix}}_{\substack{\text{Rot. um } y\text{-Achse} \\ \text{Nicken} \\ \text{(engl.: pitch)}}} \underbrace{\begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\substack{\text{Rot. um } z\text{-Achse} \\ \text{Gieren} \\ \text{(engl.: yaw)}}}$$

Dies ist die **Tait-Bryan-Darstellung** einer Drehung.

Euler verwendete drei Drehungen um nur zwei Achsen:

$$A = \underbrace{\begin{bmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Rot. um } z\text{-Achse}} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi) & -\sin(\psi) \\ 0 & \sin(\psi) & \cos(\psi) \end{bmatrix}}_{\text{Rot. um } x\text{-Achse}} \underbrace{\begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Rot. um } z\text{-Achse}}$$

Bem.: Die Matrizen kommutieren nicht. Änderung der Reihenfolge ändert die Drehung.

Alle Drehungen in $\mathbb{R}^3$ können aber auch als Drehung um eine einzige Achse dargestellt werden. Siehe nächste Seite.

Das Bild oben ist von Georg Eckert (Wikipedia).

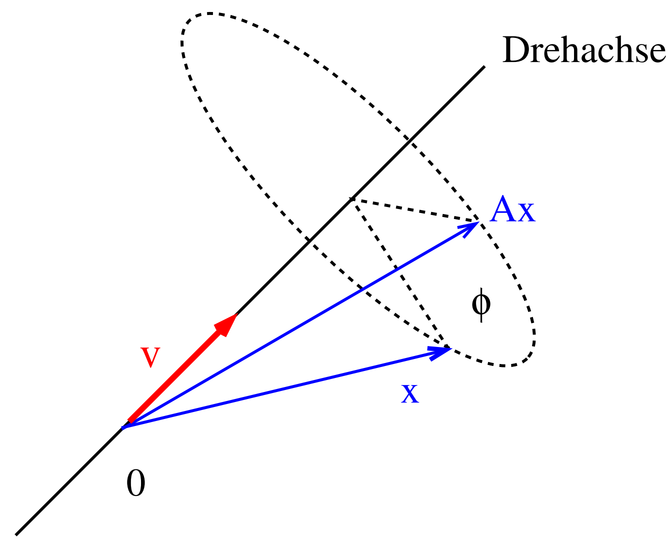
Drehungen im $\mathbb{R}^3$, II

Sei $v \in \mathbb{R}^3$ ein Einheitsvektor, d.h. $\|v\|_2 = 1$. Die Drehung im $\mathbb{R}^3$ um die Achse mit Richtungsvektor v und um den Winkel ϕ ist folgende lineare Abbildung:

$$x \longmapsto Ax = 2 \begin{bmatrix} p_0^2 + p_1^2 - \frac{1}{2} & p_1 p_2 - p_0 p_3 & p_1 p_3 + p_0 p_2 \\ p_1 p_2 + p_0 p_3 & p_0^2 + p_2^2 - \frac{1}{2} & p_2 p_3 - p_0 p_1 \\ p_1 p_3 - p_0 p_2 & p_2 p_3 + p_0 p_1 & p_0^2 + p_3^2 - \frac{1}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

wobei

$$p_0 = \cos(\phi), \qquad \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \sin(\phi)\, v.$$



Bemerkung: Es ist

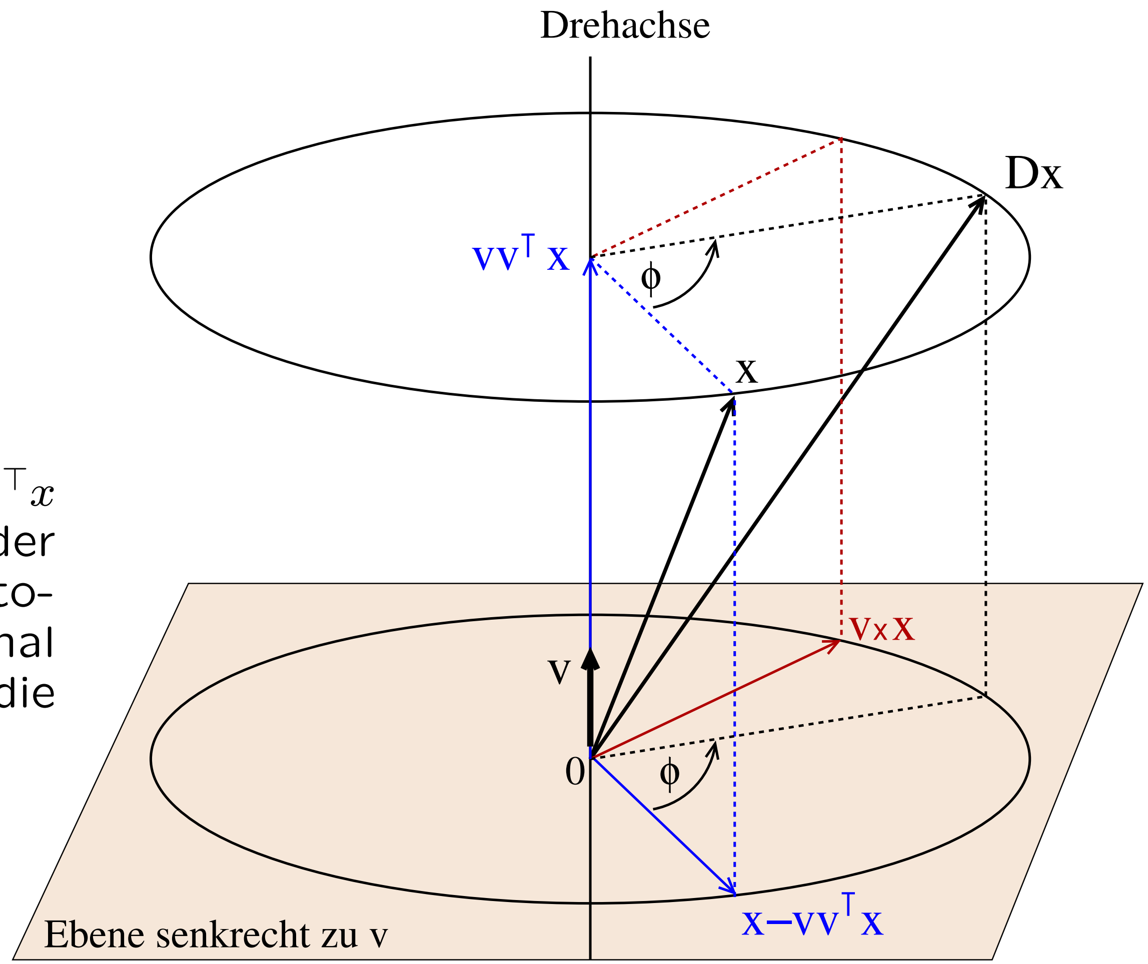
$$\left\| \begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{bmatrix} \right\|_2 = 1.$$

D.h. der p-Vektor ist ein Einheitsvektor im 4-dimensionalen Raum.

Drehungen im $\mathbb{R}^3$, II

Sei $v \in \mathbb{R}^3$ mit $\|v\|_2 = 1$.
Die Drehung im $\mathbb{R}^3$ um die Achse mit mit Richtungsvektor v und um den Winkel ϕ ist die unten angegebene lineare Abbildung $x \mapsto Ax$.

Grund: Die orthogonale Projektion $vv^\top x$ von x auf die Drehachse bleibt bei der Drehung unverändert. Die beiden Vektoren $x - vv^\top x$ und $v \times x$ sind orthogonal zueinander, gleich lang und spannen die Ebene senkrecht zur Drehachse auf.



$$
\begin{aligned}
Ax &= vv^\top x + \overbrace{\cos(\phi)\,(x - vv^\top x) + \sin(\phi)\, v \times x}^{\text{Drehung in der Ebene}} \\
&= \underbrace{\begin{bmatrix} (1-c)\,v_1^2 + c & (1-c)\,v_1 v_2 - s\,v_3 & (1-c)\,v_1 v_3 + s\,v_2 \\ (1-c)\,v_2 v_1 + s\,v_3 & (1-c)\,v_2^2 + c & (1-c)\,v_2 v_3 - s\,v_1 \\ (1-c)\,v_3 v_1 - s\,v_2 & (1-c)\,v_3 v_2 + s\,v_1 & (1-c)\,v_3^2 + c \end{bmatrix}}_{\text{Drehmatrix}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}
\end{aligned}
$$

Drehungen im $\mathbb{R}^3$, III

Drehmatrix von der vorherigen Seite:

$$A \;=\; \begin{bmatrix} (1-c)\,v_1^2+c & (1-c)\,v_1v_2-s\,v_3 & (1-c)\,v_1v_3+s\,v_2 \\ (1-c)\,v_2v_1+s\,v_3 & (1-c)\,v_2^2+c & (1-c)\,v_2v_3-s\,v_1 \\ (1-c)\,v_3v_1-s\,v_2 & (1-c)\,v_3v_2+s\,v_1 & (1-c)\,v_3^2+c \end{bmatrix}, \quad \text{wobei} \quad \begin{array}{l} c=\cos(\phi), \\ s=\sin(\phi). \end{array}$$

Eine andere Darstellung der Drehmatrix bekommt man mit den trigonometrischen Identitäten

$$\cos(\phi)=1-2\,\sin^2(\phi/2), \qquad \sin(\phi)=2\,\sin(\phi/2)\,\cos(\phi/2)$$

und den Definitionen

$$\boxed{q_0 := \cos(\phi/2), \quad q_k := \sin(\phi/2)\,v_k, \quad k=1,2,3.}$$

Es ist

$$q_0^2+q_1^2+q_2^2+q_3^2 = v_1^2+v_2^2+v_3^2 = 1$$

und

$$\boxed{\begin{aligned} A \;&=\; \begin{bmatrix} 1-2(q_2^2+q_3^2) & 2(q_1\,q_2-q_0\,q_3) & 2(q_1\,q_3+q_2\,q_0) \\ 2(q_2\,q_1+q_0\,q_3) & 1-2(q_1^2+q_3^2) & 2(q_2\,q_3-q_1\,q_0) \\ 2(q_3\,q_1-q_2\,q_0) & 2(q_3\,q_2+q_1\,q_0) & 1-2(q_1^2+q_2^2) \end{bmatrix} \\ &=\; \begin{bmatrix} -q_1 & q_0 & -q_3 & q_2 \\ -q_2 & q_3 & q_0 & -q_1 \\ -q_3 & -q_2 & q_1 & q_0 \end{bmatrix} \begin{bmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{bmatrix} \end{aligned}}$$

Drehungen und Spiegelungen sind längen- und winkelerhaltend.

Solche linearen Abbildungen nennt man **orthogonale Transformationen**.

Die zugehörigen Matrizen heißen **orthogonale Matrizen**.

Orthogonale Matrizen

Satz: Sei $A \in \mathbb{R}^{n\times n}$. Dann sind die folgenden Aussagen äquivalent:

1. $\|Ax\|_2 = \|x\|_2$ für alle $x \in \mathbb{R}^n$.

2. $(Ax)^T(Ay) = x^Ty$ für alle $x, y \in \mathbb{R}^n$.

3. $A^TA = I$ ($I=$ Einheitsmatrix)

4. $A^T = A^{-1}$.

5. Die Spalten von A stehen senkrecht aufeinander und haben alle die Länge 1. Sie bilden also eine **Orthonormalbasis** von $\mathbb{R}^n$. Formal:
$$A = [a_1 \;\; a_2 \;\; \dots \; a_n] \quad \Rightarrow \quad a_j^T a_k = \delta_{jk} = \begin{cases} 1 & \text{falls } k = j \\ 0 & \text{sonst} \end{cases}$$

6. Die Zeilen von A stehen senkrecht aufeinander und haben die Länge 1.

7. $A\,A^T = I$.

Wenn eine dieser Bedingungen erfüllt ist, dann sind auch alle anderen Bedingungen erfüllt. In diesem Fall nennt man die Matrix A **orthogonal**.

Für orthogonale Matrizen gilt stets $1 = \det(I) = \det(A^TA) = \det(A^T)\det(A) = \det(A)^2$, also
$$\det(A) = 1 \qquad \text{oder} \qquad \det(A) = -1.$$

Gramsche Matrizen

Sei $A = [a_1 \;\; a_2 \;\; \ldots \;\; a_n] \in \mathbb{R}^{m\times n}$ eine beliebige Matrix mit Spalten a_k. Dann gilt

$$A^T A = \begin{bmatrix} a_1^T a_1 & a_1^T a_2 & \ldots & a_1^T a_n \\ a_2^T a_1 & a_2^T a_2 & \ldots & a_2^T a_n \\ \vdots & & & \vdots \\ a_n^T a_1 & a_n^T a_2 & \ldots & a_n^T a_n \end{bmatrix}.$$

Die symmetrische Matrix $A^T A$ enthält also sämtliche Skalarprodukte der Spalten von A.

Insbesondere gilt

$A^T A = I \quad (I =$Einheitmatrix$) \quad \Leftrightarrow \quad$ Die Spalten von A sind senkrecht zueinander und haben Länge 1 .

Definition: $A^T A$ heißt **Gramsche Matrix** von A.

Jørgen Pedersen Gram (1850-1916), Dänischer Versicherungsmathematiker.

Erinnerung an die Lineare-Algebra-Vorlesung

Seien $v_1, v_2, \ldots, v_p$ Eigenvektoren von $A \in \mathbb{C}^{n\times n}$, so dass

$$Av_k = \lambda_k v_k, \qquad \lambda_k \in \mathbb{C}, \qquad k = 1, \ldots, p. \qquad (*)$$

Dann gilt

$$AV = V\Lambda, \qquad (**)$$

wobei

$$V = [v_1\ v_2\ \ldots v_p], \qquad \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{bmatrix}.$$

Umgekehrt folgt auch $(*)$ aus $(**)$ (siehe unten).

Wenn $p = n$ und die Eigenvektoren linear unabhängig sind, also eine Basis des $\mathbb{C}^n$ bilden, dann ist die Matrix V invertierbar und aus $(**)$ folgt

$$A = V\Lambda V^{-1}. \qquad (***)$$

Diese Produktdarstellung nennt man eine **Diagonalisierung** von A.

Nicht jede Matrix besitzt eine Basis von Eigenvektoren (Stichwort: Hauptvektoren, Jordansche Normalform).

Herleitung von $(**)$ aus $(*)$: Durch direktes Nachrechnen zeigt man

$$A[v_1\ v_2\ \ldots v_p] = [Av_1\ Av_2\ \ldots Av_p] = [\lambda_1 v_1\ \lambda_2 v_2 \ldots \lambda_p v_p] = [v_1\ v_2\ \ldots v_p]\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{bmatrix}.$$

Im allgemeinen können bei einer Matrix $A \in \mathbb{R}^{n\times n}$ die folgenden Fälle auftreten:

- Die Matrix A besitzt keine Basis von Eigenvektoren.

  Beispiel: Es ist

$$\underbrace{\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix}}_{J} \underbrace{\begin{bmatrix} \alpha \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}}_{v} = \lambda \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}, \qquad \alpha \in \mathbb{C}$$

  Man kann zeigen, dass die Jordanmatrix J außer den Vektoren v keine weiteren Eigenvektoren hat. Es gibt also keine Basis von Eigenvektoren.

- Die Eigenwerte und -Vektoren können komplex sein.

  Beispiel Drehmatrizen:

$$\begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \begin{bmatrix} 1 \\ \pm i \end{bmatrix} = e^{\mp i\phi} \begin{bmatrix} 1 \\ \pm i \end{bmatrix}.$$

- Die Eigenvektoren können in jedem beliebigen, insbesondere auch sehr spitzem Winkel zueinander stehen.

  Beispiel:

$$\begin{bmatrix} 1 & x^2-1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \pm 1 \\ \pm 1 \end{bmatrix} = (\pm x) \begin{bmatrix} x \pm 1 \\ \pm 1 \end{bmatrix}$$

  Für den Cosinus des Winkels ϕ_x zwischen den Eigenvektoren (Skalarprodukt durch Produkt der Längen) hat man $\lim_{x\to\infty} \cos(\phi_x) = 1$. Der Winkel zwischen den Eigenvektoren konvergiert also gegen 0.

Satz von der Hauptachsentransformation einer symmetrischen Matrix (Spektralsatz):

Sei $A \in \mathbb{R}^{n \times n}$ eine **symmetrische** Matrix. Dann gibt es eine Diagonalmatrix $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ und eine orthogonale Matrix $V = [v_1, \ldots, v_n] \in \mathbb{R}^{n \times n}$, so dass

$$A = V \Lambda V^T.$$

Die Spalten von V bilden eine Orthonormalbasis von Eigenvektoren von A zu den Eigenwerten λ_k. Genauer gilt:

$$A v_k = \lambda_k v_k, \qquad \textit{und} \qquad v_j^T v_k = \delta_{jk} = \begin{cases} 1 & \text{falls } j = k \\ 0 & \text{sonst} \end{cases}$$

Bemerkung:

In $(*)$ kann man statt V^T auch V^{-1} schreiben, denn weil V orthogonal ist, gilt $V^T = V^{-1}$.

Der Rayleigh-Quotient

Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Laut dem Satz von der Hauptachsentransformation gibt es eine orthonormale Eigenvektorbasis $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$,

$$Av_j = \lambda_j v_j, \quad \lambda_k \in \mathbb{R}, \qquad v_j^T v_k = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}$$

Jeder Vektor $x \in \mathbb{R}^n$ kann geschrieben werden als Linearkombination von Eigenvektoren:

$$x = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n, \quad x_j \in \mathbb{R}.$$

Direkte Rechnung ergibt:

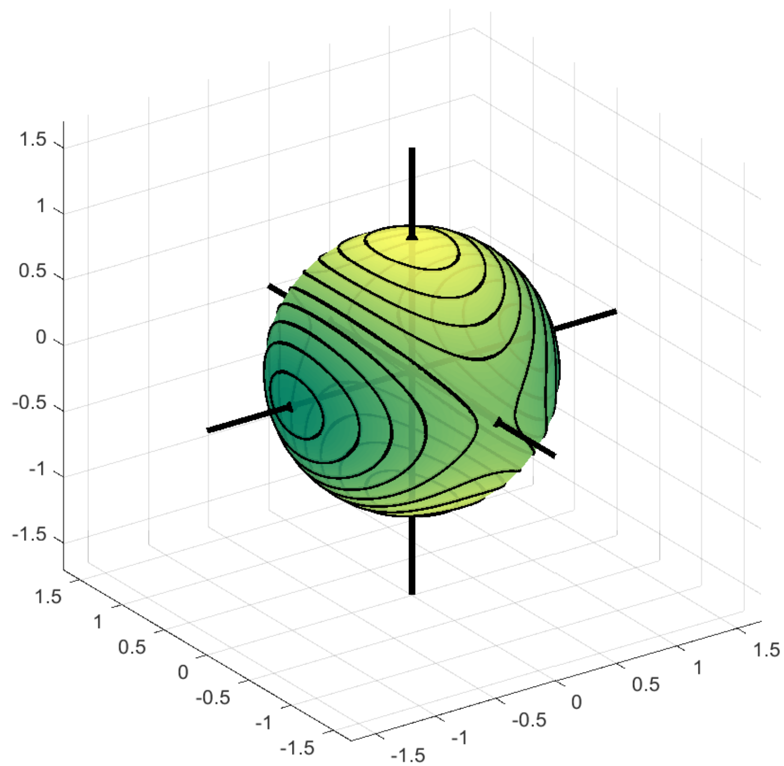
$$\begin{aligned} x^T Ax &= (x_1 v_1 + x_2 v_2 + \ldots + x_n v_n)^T A(x_1 v_1 + x_2 v_2 + \ldots + x_n v_n) \\ &= (x_1 v_1 + x_2 v_2 + \ldots + x_n v_n)^T (x_1 \lambda_1 v_1 + x_2 \lambda_2 v_2 + \ldots + x_n \lambda_n v_n) \\ &= \lambda_1 x_1^2 \underbrace{v_1^T v_1}_{=1} + \lambda_2 x_1 x_2 \underbrace{v_1^T v_2}_{=0} + \ldots \lambda_2 x_2^2 \underbrace{v_2^T v_2}_{=1} + \ldots \\ &= \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2 \end{aligned}$$

Wir definieren den **Rayleigh-Quotient** von x als

$$R(x) := \frac{x^T Ax}{\|x\|_2^2} = \frac{\lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2}{x_1^2 + x_2^2 + \ldots + x_n^2} \begin{cases} \leq \lambda_1 & \text{wenn } \lambda_1 \text{ größter Eigenwert} \\ \geq \lambda_n & \text{wenn } \lambda_n \text{ kleinster Eigenwert} \end{cases}$$

Extrema werden angenommen bei $x = v_1$ bzw. $x = v_n$.

Rayleigh-Quotient: Anschauung



Das Bild zeigt die Niveau-Mengen des Rayleigh-Quotienten

$$R(v) = \frac{v^\top A v}{\|v\|_2^2}, \quad A = \mathrm{diag}\{1, 2, 3\},$$

auf der Einheitssphäre im $\mathbb{R}^3$.

Das Bild zeigt auch die Eigenräume von A

Das Eigenwertkriterium für positive Definitheit

Satz: Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Dann gilt:

A ist genau dann positiv definit, wenn alle Eigenwerte von A positiv sind.

Beweis des Eigenwertkriteriums: Nach dem Spektralsatz existiert eine Orthonormalbasis $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$ von Eigenvektoren:

$$A v_j = \lambda_j v_j, \qquad v_j^T v_k = \begin{cases} 1 & \text{falls } j = k \\ 0 & \text{sonst} \end{cases}$$

Jeder Vektor $x \in \mathbb{R}^n$ lässt sich dann als Linearkombination der Eigenvektoren schreiben:

$$x = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n, \qquad x_j \in \mathbb{R}.$$

Eine direkte Rechnung ergibt:

$$\begin{aligned} x^T A x &= (x_1 v_1 + x_2 v_2 + \ldots + x_n v_n)^T A (x_1 v_1 + x_2 v_2 + \ldots + x_n v_n) \\ &= (x_1 v_1 + x_2 v_2 + \ldots + x_n v_n)^T (x_1 \lambda_1 v_1 + x_2 \lambda_2 v_2 + \ldots + x_n \lambda_n v_n) \\ &= \lambda_1 x_1^2 \underbrace{v_1^T v_1}_{=1} + \lambda_2 x_1 x_2 \underbrace{v_1^T v_2}_{=0} + \ldots \lambda_2 x_2^2 \underbrace{v_2^T v_2}_{=1} + \ldots \\ &= \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2 \end{aligned}$$

Wenn alle Eigenwerte positiv sind, folgt $x^T A x > 0$ für alle $x \neq 0$.

Wenn ein Eigenwert λ_k negativ ist, dann folgt $v_k^T A v_k = \lambda_k < 0$.

Gramsche Matrizen sind positiv semidefinit

Sei $A \in \mathbb{R}^{m \times n}$ eine beliebige reelle Matrix. Dann ist für alle $x \in \mathbb{R}^n$,

$$x^T A^T A\, x = (Ax)^T (Ax) = \|Ax\|_2^2 \geq 0.$$

$\Rightarrow$ $A^T A$ ist positiv semidefinit und hat daher keine negativen Eigenwerte.

Wenn A linear unabhängige Spalten hat, dann ist $Ax = 0$ nur für $x = 0$.
Also ist $x^T A^T A x = 0$ nur für $x = 0$.
Also ist $A^T A$ positiv definit und hat daher nur positive Eigenwerte.

Rayleigh-Quotient einer Gramschen Matrix, Streckfaktor und Spektralnorm

Sei $A \in \mathbb{R}^{m\times n}$ (muss nicht quadratisch sein).

Der Quotient $\|Ax\|_2/\|x\|_2$ ist der **Streckfaktor** von x bei Multiplikation mit A.

Der **maximale Streckfaktor** heißt **Spektralnorm** von A,

$$\|A\|_2 := \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}.$$

Das Quadrat des Streckfaktors ist ein Rayleigh-Quotient:

$$\frac{\|Ax\|_2^2}{\|x\|_2^2} = \frac{(Ax)^T(Ax)}{\|x\|_2^2} = \frac{x^T A^T A x}{\|x\|_2^2} = \frac{\lambda_1\, x_1{}^2 + \lambda_2\, x_2{}^2 + \ldots + \lambda_n\, x_n{}^2}{x_1{}^2 + x_2{}^2 + \ldots + x_n{}^2}$$

Dabei ist

$$x = x_1\, v_1 + x_2\, v_2 + \ldots + x_n\, v_n, \qquad x_j \in \mathbb{R}$$

dargestellt als eine Linearkombination der Eigenvektoren von $A^T A$ mit zugehörigen Eigenwerten $\lambda_k = \lambda_k(A^T A)$. Es folgt, dass

$$\sqrt{\lambda_{min}(A^T A)} \leq \frac{\|Ax\|_2}{\|x\|_2} \leq \sqrt{\lambda_{max}(A^T A)} = \|A\|_2$$

<u>Übungsaufgabe:</u> Zeige, dass für invertierbares A,

$$\sqrt{\lambda_{min}(A^T A)} = \|A^{-1}\|_2^{-1}.$$

Definition. Die Quadtatwurzeln $\sigma_k := \sqrt{\lambda_k(A^T A)}$ heißen **Singulärwerte** von A.

Singulärwertzerlegung (hier nur für quadratische Matrizen)

Den **Satz von der Singulärwertzerlegung** (singular value decomposition, kurz: SVD) kann man auf 2 verschiedene Weisen formulieren:

Sei $A \in \mathbb{R}^{n\times n}$ eine beliebige Matrix.

Formulierung 1: Es gibt Orthonormalbasen $u_1, u_2, \dots, u_n \in \mathbb{R}^n$ und $v_1, v_2, \dots, v_n \in \mathbb{R}^n$ und nicht negative Zahlen $\sigma_1, \sigma_2, \dots, \sigma_n \in \mathbb{R}$, so dass gilt

$$A v_k = \sigma_k u_k, \qquad k = 1, \dots, n \qquad (*)$$

Terminologie: Die Vektoren u_k, v_k heißen singuläre Vektoren.
Die Zahlen σ_k heißen Singulärwerte.

Formulierung 2: Es gibt orthogonale Matrizen $U, V \in \mathbb{R}^{n\times n}$ und eine Diagonalmatrix Σ mit nicht negativen Diagonalelementen, so dass

$$A = U\,\Sigma\,V^T. \qquad (**)$$

Zusammenhang zwischen beiden Formulierungen: $(*) \;\Rightarrow$

$$A\,\underbrace{[v_1\ v_2\ \dots\ v_n]}_{V} = [Av_1\ Av_2\ \dots\ Av_n] = [\sigma_1 u_1\ \sigma_2 u_2\ \dots\ \sigma_n u_n] = \underbrace{[u_1\ \ u_2\ \ \dots\ \ u_n]}_{U}\underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}}_{\Sigma}.$$

D.h.: $AV = U\Sigma$. Multiplikation mit $V^T = V^{-1}$ ergibt: $A = U\Sigma V^T$.

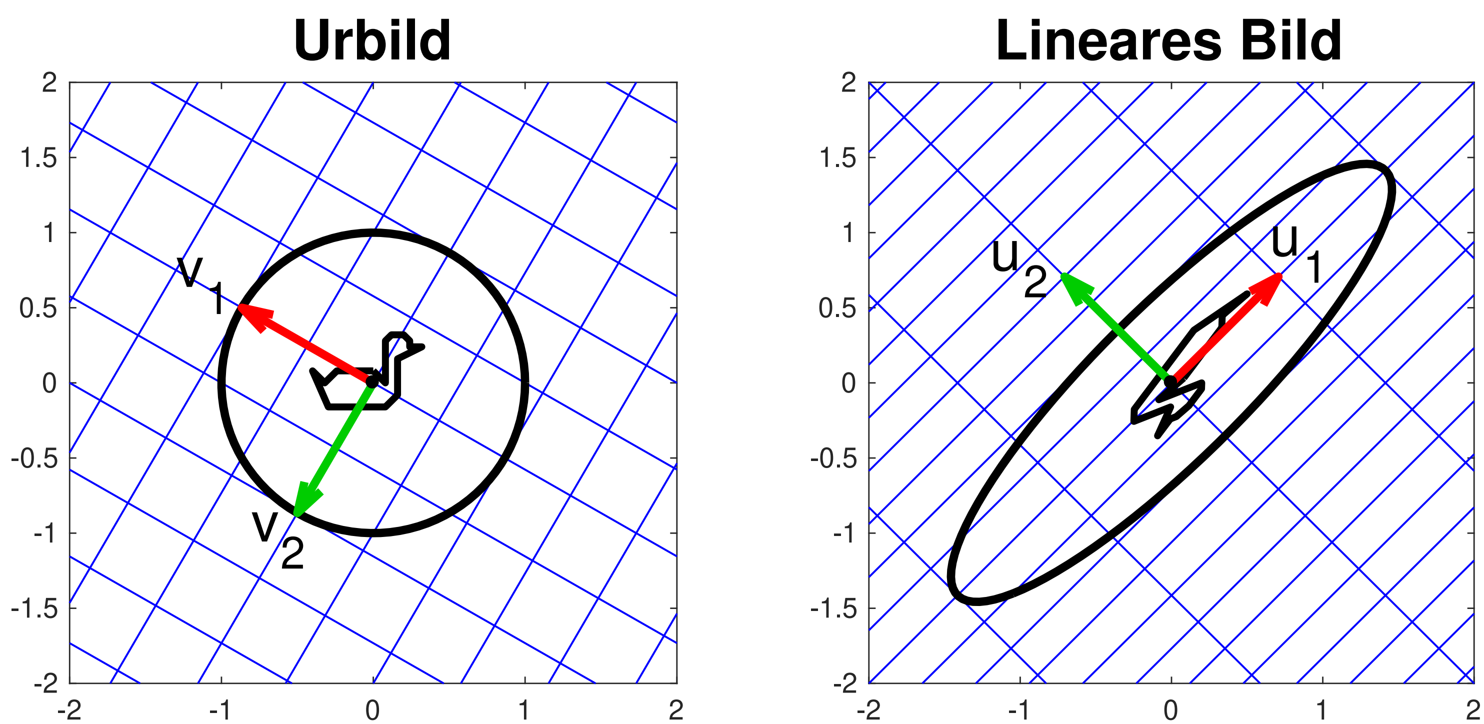
Anschauung zu Formulierung 1 des Satzes von der Singulärwertzerlegung.

Die Matrix $A \in \mathbb{R}^{2\times 2}$ im Beispiel unten hat die Singulärwerte

$$\sigma_1 = 2, \qquad \sigma_2 = 1/2.$$

Die singulären Vektoren von A sind v_1, v_2 und u_1, u_2. Sie bilden Orthonormalbasen. Es gilt

$$Av_1 = \sigma_1\, u_1, \qquad Av_2 = \sigma_2\, u_2.$$



Deutung der Singulärwertzerlegung $A = U\Sigma V^T$:

Jede lineare Abbildung $x \longmapsto Ax$ ist eine Hintereinanderausführung

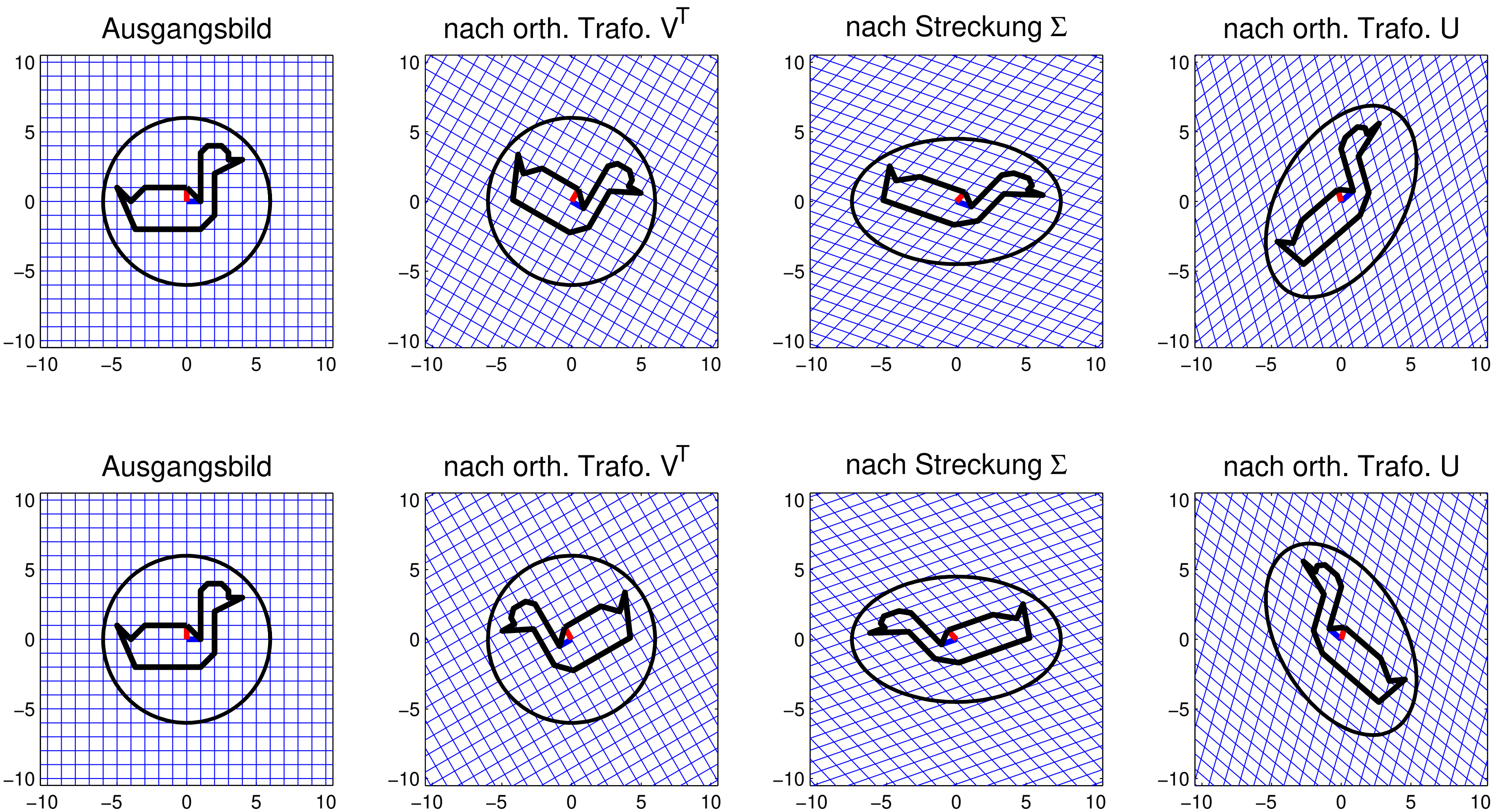
- einer orthogonalen (d.h. längen- und winkelerhaltenden) Abbildung V^T,
- Einer Streckung/Stauchung Σ entlang der Achsen des Standard-Koordinatensystems,
- einer orthogonalen (d.h. längen- und winkelerhaltenden) Abbildung U.

$$x \xmapsto{V^T} V^T x \xmapsto{\Sigma} \Sigma V^T x \xmapsto{U} U \ \Sigma V^T x = Ax$$

Siehe die Beispiele auf der nächsten Seite.

Beispiel zur Singulärwertzerlegung, Formulierung 2

Singulärwertzerlegung: $\mathbf{A} = \mathbf{U}\,\mathbf{\Sigma}\mathbf{V}^{\mathbf{T}}$ mit $U^TU = I, \;\; V^TV = I, \;\; \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}.$



Bilder: die Trafos V^T, Σ, U für $A = \begin{bmatrix} 0.87 & -0.25 \\ 0.75 & 0.87 \end{bmatrix}$ (oben) und $A = \begin{bmatrix} -0.87 & 0.25 \\ 0.75 & 0.87 \end{bmatrix}$ (unten).

In beiden Fällen ist $\sigma_1(A) \approx 1.25$, $\sigma_2(A) \approx 0.75$.

Folgerungen aus der Singulärwertzerlegung

Sei $A = U\Sigma V^T$ eine Singulärwertzerlegung von A.

Dann ist

$A^T = V\Sigma U^T$ eine Singulärwertzerlegung von A^T,

$A^{-1} = V\Sigma^{-1}U^T$ eine Singulärwertzerlegung von A^{-1} (falls A^{-1} existiert)

Außerdem folgt

$$\begin{aligned} A^T A &= (V\Sigma \underbrace{U^T)(U}_{I} \Sigma V^T) = V\Sigma^2 V^T, && (*) \\ AA^T &= (U\Sigma \underbrace{V^T)(V}_{I} \Sigma V^T) = U\Sigma^2 U^T. && (**) \end{aligned}$$

Dies sind Diagonalisierungen der symmetrischen und positiv (semi)definiten Matrizen $A^T A$, AA^T. Somit sind die Einträge von

$$\Sigma^2 = diag(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$$

die Eigenwerte von $A^T A$ und AA^T. Anders formuliert:

Die Singulärwerte von A sind die Wurzeln aus den Eigenwerten von $A^T A$ und von AA^T,

$$\sigma_k(A) = \sqrt{\lambda_k(A^T A)} = \sqrt{\lambda_k(AA^T)}.$$

Aus $(*)$ und $(**)$ folgt weiter:

Die singulären Vektoren V sind die Eigenvektoren von $A^T A$, die singulären Vektoren U sind die Eigenvektoren von AA^T.

Berechnung der Singulärwertzerlegung einer invertierbaren Matrix

Aus den Folgerungen auf der vorigen Seite ergibt sich folgendes Verfahren zur Berechnung einer Singulärwertzerlegung einer invertierbaren Matrix.

- Berechne eine Orthonormalbasis $V = [v_1 \;\; v_2 \;\; \dots \;\; v_n]$ von Eigenvektoren von $A^T A$, so dass $A^T A\, v_k = \lambda_k\, v_k$.

- Die Eigenwerte λ_k der positiv definiten symmetrischen Matrix $A^T A$ sind positiv. Setze $\sigma_k := \sqrt{\lambda_k}$. Dann ist

$$(Av_j)^T (Av_k) = v_j^T A^T A v_k = v_j(\sigma_k^2\, v_k) = \sigma_k^2 v_j^T v_k = \sigma_k^2\, \delta_{jk}, \qquad (*)$$

  Insbesondere $\|Av_k\|_2 = \sqrt{(Av_k)^T (Av_k)} = \sigma_k$.

- Setze $u_k := Av_k / \sigma_k$ und bilde $U = [u_1 \;\; u_2 \;\; \dots \;\; u_n]$. Dann gilt

$$Av_k = \sigma_k\, u_k \qquad (**)$$

  und aus $(*)$ folgt, dass die u_k eine Orthonormalbasis bilden.
  Wegen $(**)$ hat man $AV = U\Sigma$ mit $\Sigma := diag(\sigma_1, \dots, \sigma_n)$, also auch

$$A = U \Sigma V^T.$$

Singulärwertzerlegung (singular value decomposition) mit Matlab

s = svd(A) gibt einen Vektor s zurück, in dem die Singulärwerte von A stehen.

[U,S,V] = svd(A) gibt die volle Singulärwertzerlegung zurück: $A = USV^T$.

Bei symmetrischen Matrizen sind Diagonalisierung und Singulärwertzerlegung bis aufs Vorzeichen dasselbe

Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch und sei

$$A = V \Lambda V^T, \qquad V^T V = I, \qquad \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

eine orthogonale Diagonalisierung. Sei

$$\Sigma = |\Lambda| = \mathrm{diag}(|\lambda_1|, \ldots, |\lambda_n|), \qquad D = \mathrm{diag}(\mathrm{sign}(\lambda_1), \ldots, \mathrm{sign}(\lambda_n)),$$

wobei $\mathrm{sign}(\cdot) =$ Vorzeichen. Dann ist $\Lambda = D\Sigma$ und

$$A = \underbrace{V D}_{=:U} \Sigma V^T. \qquad (*)$$

U ist orthogonal: $U^T U = (VD)^T (VD) = \underbrace{D^T}_{=D} \underbrace{V^T V}_{=I} D = D^2 = I$.

Daher ist $(*)$ eine Singulärwertzerlegung.

Wenn A positiv definit ist, dann ist $D = I$ und $\Lambda = \Sigma$.

$\Rightarrow$ Bei positiv definiten sym. Matrizen sind die Eigenwerte die Singulärwerte.

Singulärwertzerlegung und Konditionszahlen

Aus der Singulärwertzerlegung lässt sich leicht herleiten, dass

$$cond_2(A) = \|A\|_2 \, \|A^{-1}\|_2 = \frac{\sigma_{\mathsf{max}}(A)}{\sigma_{\mathsf{min}}(A)}.$$

Genaueres in der nächsten Vorlesung.