Modul Algebra
(Lineare Algebra II und Einführung in die Algebra)

SoSe 2005
Wiland Schmale

Teil A:

Ergänzungen zur linearen Algebra der Vektorräume mit Skalarprodukt

Nur teilweise ausgearbeitete Notizen[1], nicht alles wird in der Vorlesung behandelt.

---

[1] Hinweise auf Fehler bitte an wiland.schmale@uni-oldenburg.de

Inhaltsverzeichnis

1 Projektionen und Abstände

In diesem Abschnitt geht es in erster Linie um ein paar Wiederholung und etwas Hintergrundwissen u.A. für die ersten Übungsaufgaben.

Sei im Folgenden stets V ein euklidischer Vektorraum mit Skalarprodukt $(\!(\cdot\,,\cdot)\!)$. Wie üblich wird dann für $v, v' \in V$ der Abstand von v, v' mit $d(v, v')$ und die Länge von v mit $\| \, v \, \|$ bezeichnet. Dabei ist $\| \, v \, \| = \sqrt{(\!(v\,,v)\!)}$ und $d(v, v') = \sqrt{(\!(v - v'\,, v - v')\!)} = \| \, v - v' \, \|$.
Eine Orthogonalbasis von V ist eine Basis bestehend aus paarweise orthogonalen Vektoren. Sie brauchen nicht die Länge 1 zu haben. Ist Letzteres der Fall, so spricht man von einer Orthonormalbasis.

(A) Orthogonale Projektion auf Untervektorräume: Sei W ein endlichdimensionaler Untervektorraum von V und sei $(w_1, \ldots, w_r)$ eine Orthogonalbasis von W. Für $v \in V$ ist dann

$$P_W(v) := \sum_{i=1}^{r} \frac{(\!(v\,, w_i)\!)}{(\!(w_i\,, w_i)\!)} w_i$$

die orthogonale Projektion von v auf W. Man kann zeigen, dass gilt:

$$d(v, P_W(v)) = \| \, v - P_W(v) \, \| \leq \| \, v - w \, \| = d(v, w) \quad \text{für alle} \quad w \in W$$
$$\text{und ‚=‘ nur für } w = P_W(v).$$

(B) Abstände von Unterräumen: Seien jetzt W und W' endlichdimensionale Untervektorräume von V, $x, x' \in V$ und $\Gamma = x + W, \Gamma' = x' + W'$ affine Unterräume. Ihr Abstand ist definiert als

$$d(\Gamma, \Gamma') := \mathsf{inf}\{d(y, y') : \ y \in \Gamma, y' \in \Gamma'\} \ .$$

Man kann nun zeigen, dass gilt:

$$d(\Gamma, \Gamma') = d(x - x', W + W') \ .$$

$d(x - x', W + W') = d(\{x - x'\}, W + W')$ ist der Abstand des Punktes (spezieller affiner Unterraum) $x - x'$ von $W + W'$. Wegen (A) gilt

$$d(x - x', W + W') = d(x - x', P_{W+W'}(x - x')) = \| \, (x - x') - P_{W+W'}(x - x') \, \| \ .$$

Damit kann der Abstand zweier endlichdimensionaler affiner Unterräume mit Hilfe von (A) berechnet werden, indem man zuerst eine Orthogonalbasis von $W + W'$ bestimmt, dann $y^* := P_{W+W'}(x - x')$ berechnet und dann

$$d(\Gamma, \Gamma') = d(x - x', y^*) = \| \, x - x' - y^* \, \| \ .$$

2 Lineare Abbildungen euklidischer Vektorräume

Lineare Abbildungen sind die "strukturerhaltenden" Abbildungen von Vektorräumen, die sogenannten "Morphismen". Euklidische Vektorräume haben zusätzliche Struktur. Die zugehörigen "Morphismen" sind die metrischen Abbildungen oder Isometrien.

Definition 2.1: Seien V, W euklidische Vektorräume mit Skalarprodukten $(\!(\ ,\ )\!)_1$, $(\!(\ ,\ )\!)_2$ und sei $F : V \to W$ eine Abbildung. F heißt **metrisch** bzw. **mit dem Skalarprodukt verträglich,** wenn für alle $v, v' \in V$ gilt

$$(\!(F(v)\,, F(v'))\!)_2 = (\!(v\,, v')\!)_1\ .$$

Man lässt dabei i. A. bei den Skalarprodukten die Indizes weg, da ja stets klar ist "wo" man sich befindet.
Die Forderung nach der Verträglichkeit mit dem Skalarprodukt allein ist schon so stark, dass sie die Linearität erzwingt.

Beobachtung 2.2: Metrische Abbildungen sind linear.

Beweis: Sei $U =< \{F(v) : v \in V\} >_{\mathbb{R}}$ der von $F(V)$ erzeugte Untervektorraum von W. Seien $v, v' \in V$ und $w := F(v+v') - F(v) - F(v')$. Es ist w aus U, und wir müssen zeigen, dass $w = 0$. Für alle $v'' \in V$ gilt:

$$\begin{aligned}((w, F(v''))) &= (\!(F(v+v')\,, F(v''))\!) - (\!(F(v)\,, F(v''))\!) - (\!(F(v')\,, F(v''))\!)\\ &\underset{(1)}{=} (\!(v+v'\,, v'')\!) - (\!(v\,, v'')\!) - (\!(v'\,, v'')\!) \underset{((\ ,\ ))\text{ bilinear}}{=} 0\end{aligned}$$

Dann gilt aber auch $w \perp U$ und insbesondere $w \perp w$ bzw. $(\!(w\,, w)\!) = 0$, denn $w \in U$. Es folgt $w = 0$. Analog zeigt man für $\lambda \in \mathbb{R}$ und $w = F(\lambda v) - \lambda F(v)$, dass gilt: $w \perp U$ ☐

Definition 2.3: Seien V, W euklidische Vektorräume und $F : V \to W$ linear.

(a) F heißt **Isometrie** oder **abstandstreu**, wenn für alle $v, v' \in V$ gilt: $d(F(v), F(v')) = d(v, v')$
dabei ist (Wdhlg.) $d(v, v') = \sqrt{(\!(v - v'\,, v - v')\!)} = \parallel v - v' \parallel$

(b) F heißt **längentreu**, wenn für alle $v \in V$ gilt: $\parallel F(v) \parallel = \parallel v \parallel$

(c) F heißt **winkeltreu**, wenn für alle $v, v' \in V$ gilt:

$$(\!(F(v)\,, F(v'))\!) \cdot \parallel v \parallel \cdot \parallel v' \parallel = (\!(v\,, v')\!) \cdot \parallel F(v) \parallel \cdot \parallel F(v') \parallel$$

Bei (c) erinnere man sich z.B. an [F], S. 276.

Beobachtung 2.4: Seien V, W euklidisch, $F : V \to W$ linear. Es gilt:

$$\begin{array}{llclcl} F & \text{metrisch} & \Leftrightarrow & F \text{ Isometrie} & \Leftrightarrow & F \text{ längentreu}\\ & \quad\Downarrow & & & & \\ F & \text{winkeltreu} & & & & \end{array}$$

Beweis: F längentreu $\Rightarrow$ F metrisch:
Seien $v, v' \in V$. Dann ist zunächst $(\!(v+v'\,, v+v')\!) = (\!(v\,, v)\!) + 2(\!(v\,, v')\!) + (\!(v'\,, v')\!)$.

M.a.W. $(\!(v\,,v')\!) = \frac{1}{2}\left((\!(v+v'\,,v+v')\!) - (\!(v\,,v)\!) - (\!(v'\,,v')\!)\right)$. Dies gilt auch für $F(v), F(v')$ statt v, v' und da F längentreu ist, folgt:

$$\begin{aligned}(\!(F(v)\,,F(v'))\!) &= \frac{1}{2}\Big((\!(\underbrace{F(v)+F(v')}_{F(v+v')}\,,\underbrace{F(v)+F(v')}_{F(v+v')})\!) - (\!(F(v)\,,F(v))\!) - (\!(F(v')\,,F(v'))\!)\Big)\\ &= \frac{1}{2}\left((\!(v+v'\,,v+v')\!) - (\!(v\,,v)\!) - (\!(v'\,,v')\!)\right) = (\!(v\,,v')\!)\end{aligned}$$

Alle übrigen Implikationen ergeben sich direkt aus den Definitionen. ☐

Bemerkung 2.5: **(a)** Aus der Definition 2.3 ergibt sich im Falle $V = W$ unmittelbar, dass ein längentreuer Endomorphismus von V nur die Eigenwerte $+1$ oder -1 haben kann.

(b) Wenn F längentreu, $\lambda \in \mathbb{R}\backslash\{0, 1, -1\}$, dann ist $\lambda \cdot F$ nicht mehr längentreu, denn $\parallel (\lambda \cdot F)(v) \parallel = |\lambda| \parallel F(v) \parallel = |\lambda| \parallel v \parallel$. Allerdings ist $\lambda \cdot F$ immer noch winkeltreu!

(c) Verbreitete Bezeichnung: **"orthogonale" Abbildung** für Isometrie eines euklidischen Vektorraumes.

(d) Längentreue Abbildungen sind relevant in der angewandten Mathematik, Numerik, bei Optimierungsaufgaben wie z. B. bei der Ausgleichsrechnung.

Beobachtung 2.6: Längentreue Abbildungen sind injektiv.

Beweis: $\forall v \in V : F(v) = 0 \Rightarrow (\!(F(v)\,,F(v))\!) = 0 \underset{\text{F metrisch}}{\Rightarrow} (\!(v\,,v)\!) = 0 \Rightarrow v = 0$ ☐

$F(V)$ und V sind demnach bei einer metrischen (also insbesondere linearen) Abbildung isomorph. Wir konzentrieren uns im Folgenden auf metrische Endomorphismen, dann ist $V = W$, und $(\!(\,,\,)\!)_1 = (\!(\,,\,)\!)_2$.

Standardbeispiel 2.7: Sei $V = \mathbb{R}^n$ versehen mit dem **Standardskalarprodukt**. D.h. es ist $\boldsymbol{(\!(v\,,v')\!) = {}^{\mathsf{t}}vv}$ für $v, v' \in V$. Sei außerdem $\boldsymbol{F = L_A}$ mit $A \in \mathbb{R}^{n\times n}$. F eine Isometrie. Dann muss für $1 \leq i,\ j \leq n$ gelten:

$$(\!(e_i\,,e_j)\!) = \delta_{ij} = (\!(F(e_i)\,,F(e_j))\!) = (\!(A_{\bullet i}\,,A_{\bullet j})\!) = {}^{\mathsf{t}}A_{\bullet i} \cdot A_{\bullet j}$$

M. a. W: Es gilt dann ${}^{\mathsf{t}}AA = E_n$

Definition 2.8: $A \in \mathbb{R}^{n\times n}$ heißt **orthogonal**, wenn gilt:

$${}^{\mathsf{t}}AA = E_n \tag{1}$$

Konsequenter wäre die Bezeichnung: "orthonormiert".

Beobachtung 2.9: Sei $A \in \mathbb{R}^{n\times n}$ orthogonal. Sowohl die Spalten als auch die Zeilen von A bilden eine Orthonormalbasis. Außerdem gilt: $\mathsf{det}\,A = \pm 1$, $A^{-1} = {}^{\mathsf{t}}A$, $A\,{}^{\mathsf{t}}A = E_n$, $A \in Gl(n, \mathbb{R})$ und $\sigma_{\mathbb{R}}(A) \subseteq \{1, -1\}$.

Beweis: Wegen (1) bilden die Spalten von A eine Orthonormalbasis. Da $E_n = {}^{\mathsf{t}}AA$, gilt auch $A^{-1} = {}^{\mathsf{t}}AAA^{-1} = {}^{\mathsf{t}}A$ und somit $E_n = AA^{-1} = A\,{}^{\mathsf{t}}A$. Also bilden auch die Zeilen von A eine Orthonormalbasis. Da $1 = \mathsf{det}\,E_n = \mathsf{det}\,A\,{}^{\mathsf{t}}A = \mathsf{det}\,A \cdot \mathsf{det}\,A = (\mathsf{det}\,A)^2$, folgt: $\mathsf{det}\,A \pm 1$. ☐

Satz 2.10: $\mathcal{O}(n) = \{A \in \mathbb{R}^{n\times n} : A \text{ orthogonal}\}$ ist eine Untergruppe von $Gl(n,\mathbb{R})$, $\mathcal{O}_1(n) = \{A \in \mathcal{O}(n) : \mathsf{det}\, A = 1\}$ ist Untergruppe von $\mathcal{O}(n)$.

Beweis: mit A, B sind auch $A \cdot B$ und $A^{-1} = {}^{\mathsf{t}}A \in \mathcal{O}(n)$.
Bemerkung: $\mathcal{O}_1(n)$ ist sogar ein sogenannter Normalteiler in $\mathcal{O}(n)$ d.h.: für alle $A \in \mathcal{O}_1(n)$, $B \in \mathcal{O}(n)$ ist $BAB^{-1} \in \mathcal{O}_1(n)$. □

Beispiel 2.11: (a) $\mathcal{O}(1) = \{1, -1\}$
(b) Behauptung

$$A \in \mathcal{O}(2) \Leftrightarrow \exists a, b \in \mathbb{R} : [A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \text{ oder } A = \begin{bmatrix} a & b \\ a & -a \end{bmatrix}] \textbf{ und } a^2 + b^2 = 1.$$

Beweis: Sei $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{O}(2)$.

$A^{\mathsf{t}}A = E_2$ bedeutet $\boxed{a^2 + b^2 = 1,\ ac + bd = 0,\ c^2 + d^2 = 1}$

${}^{\mathsf{t}}AA = E_2$ bedeutet $\boxed{a^2 + c^2 = 1,\ ab + cd = 0,\ b^2 + d^2 = 1}$

Es folgt: $b = \pm c,\ a = \pm d$ also: $d = \pm a$. Wenn $d = a$: dann $c = -b$. Wenn $d = -a$: dann $c = b$.
Umgekehrt rechnet man nach, dass solche Matrizen orthogonal sind.

(c) Geometrische Interpretation von $\boldsymbol{A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}}$ und $\boldsymbol{a^2 + b^2 = 1}$:

Dann ist $\mathsf{det}\, A = -1$ und $\sigma(A) = \{1, -1\}$, denn

$$\mathsf{det}\,(A - tE) = t^2 - a^2 - b^2 = t^2 - 1 = (t+1)(t-1).$$

A ist demnach diagonalisierbar. Man berechnet für $b \neq 0$:

$$W := \mathsf{Eig}\,(A, 1) =< \begin{bmatrix} b \\ 1 - a \end{bmatrix} >_{\mathbb{R}} \text{ und } \mathsf{Eig}\,(A, -1) =< \begin{bmatrix} b \\ -1 - a \end{bmatrix} >_{\mathbb{R}} .$$

L_A ist die **Spiegelung an $\boldsymbol{W}$.** Für $b = 0$ ist $W =< e_1 >_{\mathbb{R}}$.
Sonderfall: $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Hier ist $A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$ und $W =< \begin{bmatrix} 1 \\ 1 \end{bmatrix} >_{\mathbb{R}}$.

Beachte: A ist im Sonderfall eine 2×2-Permutationsmatrix.

(d) Geometrische Interpretation von $\boldsymbol{A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}}$ mit $\boldsymbol{a^2 + b^2 = 1}$:

Sei $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$ ein Vektor der Länge 1, also mit $x^2 + y^2 = 1$. Auch $A \begin{bmatrix} x \\ y \end{bmatrix}$ hat die Länge 1 und es ist

$$(\!(\begin{bmatrix} y \\ x \end{bmatrix} , A \begin{bmatrix} x \\ y \end{bmatrix})\!) = (\!(\begin{bmatrix} x \\ y \end{bmatrix} , \begin{bmatrix} ax + by \\ -bx + ay \end{bmatrix})\!) = a(x^2 + y^2) = a = \cos \underbrace{\measuredangle(\begin{bmatrix} x \\ y \end{bmatrix} ,\ A \begin{bmatrix} x \\ y \end{bmatrix})}_{=:\varphi}$$

Für Letzteres siehe z.B. [F] S. 276. Offensichtlich ist φ unabhängig von $\begin{bmatrix} x \\ y \end{bmatrix}$.
L_A ist die **Drehung um φ gegen den Uhrzeigersinn** (vgl. [F], S. 106-107).
Beachte: Wenn $a = \cos\varphi$ und $a^2 + b^2 = 1$, dann ist $b = \pm\sin\varphi$. Man berechnet noch: $\sigma_{\mathbb{C}}(A) = \{a + ib, a - ib\} = \{\text{Eigenwerte von } A \text{ in } \mathbb{C}\}$. Fasst man die reelle Matrix A als komplexe Matrix auf, so erhält man für $b \neq 0$ in $\mathbb{C}$ die Eigenräume $\mathsf{Eig}\,(a + ib) =< \begin{bmatrix} 1 \\ i \end{bmatrix} >_{\mathbb{C}}$, $\mathsf{Eig}\,(a - ib) =< \begin{bmatrix} 1 \\ -i \end{bmatrix} >_{\mathbb{C}}$.

(e) Elementare orthogonale Matrizen für $n > 2$:

$$\begin{bmatrix}
1 & & & & & & & & 0 \\
 & \ddots & & & & & & & \\
 & & a & & \cdots & & b & & \\
 & & & 1 & & & & & \\
 & & \vdots & & \ddots & & \vdots & & \\
 & & & & & 1 & & & \\
 & & c & & \cdots & & d & & \\
 & & & & & & & \ddots & \\
0 & & & & & & & & 1
\end{bmatrix}$$

Dabei ist $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ wie in (b)(c)(d) eine Drehung oder eine Spiegelung
Beachte: die Elementarmatrizen P^i_j sind elementare orthogonale Matrizen mit Determinante -1.

(f) Permutationsmatrizen sind orthogonal als endliche Produkte von Elementarmatrizen (Vertauschungsmatrizen).

Matrixdarstellungen von Isometrien, strukturgerechte Basiswahl und Koordinatisierung.

Beobachtung 2.12: Seien V ein euklidischer Vektorraum, $\mathsf{dim}_V = n$, $\mathfrak{A} = (v_1, \ldots, v_n)$ eine Orthonormalbasis. Dann ist die Koordinatenabbildung $\Phi_{\mathfrak{A}} : \mathbb{R}^{n\times n} \to V$ metrisch (Definition 2.1), wobei in $\mathbb{R}^{n\times n}$ das Standardskalarprodukt benutzt wird.

Beweis: Seien $u = \sum\limits_{i=1}^{n} x_i v_i$, $w = \sum\limits_{i=1}^{n} y_i v_i \in V$, $x = {}^t[x_1, \ldots, x_n]$, $y = {}^t[y_1, \ldots, y_n]$.
Dann gilt:

$$((\Phi_{\mathfrak{A}}(x)\,, \Phi_{\mathfrak{A}}(y))) = ((u\,, w)) \underset{\mathfrak{A}\text{ ist Orthonormalbasis}}{=} \sum_{i=1}^{n} x_i y_i = {}^t x y = ((x\,, y)) \qquad \square$$

Für "strukturgerechte" Matrixdarstellungen erhalten wir:

Satz 2.13: Sei V ein euklidischer Vektorraum, $\mathsf{dim}_V = n$, $F : V \to V$ linear, $\mathfrak{A}$ Orthonormalbasis und $A := M^{\mathfrak{A}}_{\mathfrak{A}}(F)$. Dann gilt

$$F \text{ Isometrie } \Leftrightarrow L_A \text{ Isometrie } \Leftrightarrow A \text{ orthogonal}$$

Wir benutzen beim Beweis folgenden

Hilfssatz 2.14: (a) die Hintereinanderausführung metrischer (linearer) Abbildungen ist wieder metrisch.

(b) die Umkehrabbildung einer surjektiven metrischen (linearen) Abbildung ist wieder metrisch.

Beweis. Nachrechnen als Übung.
Beweis von Satz 2.13:

$$\begin{tikzcd} V \arrow[r, "F"] & V \\ \mathbb{R}^{n\times 1} \arrow[u, "\Phi_{\mathfrak{A}}"] \arrow[r, "L_A"'] & \mathbb{R}^{n\times 1} \arrow[u] \end{tikzcd} \qquad A = M_{\mathfrak{A}}^{\mathfrak{A}}(F)$$

Es gilt (LA1): $L_A = \Phi_{\mathfrak{A}}^{-1} \circ F \circ \Phi_{\mathfrak{A}}$ und $F = \Phi_{\mathfrak{A}} \circ L_A \circ \Phi_{\mathfrak{A}}^{-1}$. Nach Hilfssatz 2.14 und Beobachtung 2.12 ist also die erste ”$\Leftrightarrow$” richtig.
$\underline{L_A \text{ Isometrie} \Rightarrow A \text{ orthogonal:}}$ siehe Beispiel 2.7.
$\underline{A \text{ orthogonal} \Rightarrow L_A \text{ Isometrie:}}$ A orthogonal heißt nach Beobachtung 2.9:
$\mathfrak{A} = (v_1, \ldots, v_n)$ mit $v_i = A_{\bullet i}$ ist eine Orthonormalbasis von $R^{n\times n}$. Daher ist $L_A = \Phi_{\mathfrak{A}}$ metrisch nach Beobachtung 2.12 und somit Isometrie nach Beobachtung 2.4. ☐
Kurz: ”wenn man Orthonormalbasen zugrundegelegt, passt alles”, alle Abbildungen im Standarddiagramm (s.o.) sind metrisch.

Zwei Fragen entstehen hier:

(I) lässt sich die Orthonormalbasis $\mathfrak{A}$ so wählen, dass $M_{\mathfrak{A}}^{\mathfrak{A}}(F) = A$ ”schön” ist? Mit anderen Worten: gibt es $P \in \mathcal{O}(n)$ derart, dass $P^{-1}AP = {}^tPAP$ ”schön” ist; ”orthogonale Ähnlichkeit”

(II) Was bewirken ”orthogonale Zeilenumformungen” und speziell: lässt sich eine orthogonale Matrix P als Produkt gewisser ”elementarer orthogonaler” Matrizen darstellen?

Zuerst (II) (wichtiger für Anwendungen , (I) ist schwieriger)
Wir brauchen genauere Bezeichnungen:

Definition 2.15: Seien $n \geq 2, \varepsilon_1, \varepsilon_2 \in \{1, -1\}$ und sei

$$P = \begin{bmatrix} 1 &&&&&&&& 0 \\ & \ddots &&&&&& \iddots & \\ && a && 0 && b && \\ && \vdots & \ddots && \iddots & \vdots && \\ && 0 && 1 && 0 && \\ && \vdots & \iddots && \ddots & \vdots && \\ && \varepsilon_2 \cdot b & \cdots & 0 & \cdots & \varepsilon_1 \cdot a && \\ & \iddots &&&&&& \ddots & \\ 0 &&&&&&&& 1 \end{bmatrix}$$

eine elementare orthogonale Matrix in $\mathbb{R}^{n\times n}$. Es gilt dann $\varepsilon_1\varepsilon_2 = -1,\ a^2 + b^2 = 1$. Die Nummern der von der Einheitsmatrix abweichenden Zeilen/Spalten seien i, k mit $i < k$. P heißt:

(a) elementare Drehung(smatrix), Bezeichnung $D(i, k; a, b)$, wenn $\varepsilon_1 = 1$ (und dann $\varepsilon_2 = -1$)

(b) elementare Spiegelung(smatrix), Bezeichnung $S(i, k; a, b)$, wenn $\varepsilon_1 = -1$ (und dann $\varepsilon_2 = 1$).

Beispiel bei $\boldsymbol{n = 3}$: Wir betrachten $D(1, 3; a, b) = \begin{bmatrix} a & 0 & b \\ 0 & 1 & 0 \\ -b & 0 & a \end{bmatrix}$, und $S(1, 3; a, b) = \begin{bmatrix} a & 0 & b \\ 0 & 1 & 0 \\ b & 0 & -a \end{bmatrix}$. In beiden Fällen lässt die entsprechende lineare Abbildung den Vektor e_2 und damit auch die ganze "Achse" $< e_2 >_{\mathbb{R}}$ punktweise fest. Im ersten Fall findet in der Ebene $< e_1, e_3 >_{\mathbb{R}}$ eine Drehung, im zweiten Fall eine Spiegelung statt.

Beachte: $\begin{bmatrix} a \\ -b \end{bmatrix}$ mit $a^2 + b^2 = 1$ ist ein wohldefinierter Punkt auf dem Einheitskreis und $\underbrace{\begin{bmatrix} a & b \\ -b & a \end{bmatrix}}_{\text{Drehmatrix}} e_1 = \begin{bmatrix} a \\ -b \end{bmatrix} \cdot$ Es wird also nach $\begin{bmatrix} a \\ -b \end{bmatrix}$ und analog e_2 nach $\begin{bmatrix} b \\ a \end{bmatrix}$ gedreht.

Der "Drehwinkel" kann demnach mit Hilfe von $\begin{bmatrix} a \\ -b \end{bmatrix}$ (etwa) festgelegt werden. Dies geschieht meist durch folgende Bijektion:

$$[0, 2\pi) \to C_1 = \left\{ \begin{bmatrix} a \\ -b \end{bmatrix} : a, b \in \mathbb{R},\ a^2 + b^2 = 1 \right\} \quad , \varphi \mapsto \begin{bmatrix} \cos\varphi \\ -\sin\varphi \end{bmatrix}$$

Beobachtung 2.16: Es gilt: $D(i, k; a, b)^{-1} = D(i, k; a, -b)$ und $S(i, k; a, b)^{-1} = S(i, k; a, b)$

Satz 2.17: Orthogonale Zeilenstufenform. Sei $n \geq 2$. Zu $A \in \mathbb{R}^{n \times m}$ gibt es $r \in \mathbb{N}_+$ und elementare orthogonale Matrizen $Q_1, \ldots, Q_r$ so dass $Q_r \cdots Q_1 A$ in Zeilenstufenform ist mit positiven Eckeinträgen bis zur mindestens vorletzten Zeile $\neq 0$ und mit $1 \leq r \leq \frac{n(n-1)}{2}$. Jedes $Q_i,\ 1 \leq i \leq r$ kann entweder als elementare Spiegelung oder als elementare Drehung gewählt werden. Insbesondere sind die beiden folgenden Fälle stets realisierbar:

'**alle** Q_i el. Drehung' und '**alle** Q_i el. Spiegelung'

Beweis: Konstruktiv, Induktion nach n. Wenn man elementare orthogonale Matrizen für $n = 1$ als $+1, -1$ festlegt, dann ist $n = 1$ möglich, es lohnt aber den Fall $n = 2$ zusätzlich zu betrachten.

$\boldsymbol{n = 2}$: Sei $A = \begin{bmatrix} a_1 & * \\ a_2 & * \end{bmatrix} \in \mathbb{R}^{2 \times m}$. Wenn $a_1 = a_2 = 0$, dann ist $A = \underbrace{\begin{bmatrix} 0 & a_1' & * \\ 0 & a_2' & * \end{bmatrix}}_{=A'}$, falls $m \geq 2$ und man betrachtet A' statt A. Wenn die Behauptung für A' gilt, dann auch für A.

Sei also o. E: $\boldsymbol{a_1 \neq 0}$ **oder** $\boldsymbol{a_2 \neq 0}$.

Dann ist $\mathrm{D}\big(1, 2; \frac{a_1}{\delta}, \frac{a_2}{\delta}\big) \cdot A = \begin{bmatrix} \frac{a_1}{\delta} & \frac{a_2}{\delta} \\ \frac{-a_2}{\delta} & \frac{a_1}{\delta} \end{bmatrix} \cdot \begin{bmatrix} a_1 & * \\ a_2 & * \end{bmatrix} = \begin{bmatrix} \delta & * \\ 0 & * \end{bmatrix}$

und ebenso $S\left(1, 2, \frac{a_1}{\delta}, \frac{a_2}{\delta}\right) \cdot A = \begin{bmatrix} \frac{a_1}{\delta} & \frac{a_2}{\delta} \\ \frac{a_2}{\delta} & \frac{-a_1}{\delta} \end{bmatrix} \cdot \begin{bmatrix} a_1 & * \\ a_2 & * \end{bmatrix} = \begin{bmatrix} \delta & * \\ 0 & * \end{bmatrix}$, wobei $\delta = \sqrt{a_1^2 + a_2^2} > 0$.

Beidesmal liegt eine Zeilenstufenform vor mit positivem Eckeintrag in der ersten Zeile und es ist $r \leq 1$.

$n \geq 2$: Die Behauptung sei richtig für $A \in \mathbb{R}^{n\times m}$, und $A \in R^{(n+1)\times m}$, und sei o. E. die erste Spalte von A nicht 0. Sei etwa $A = \begin{bmatrix} a_1 & & \\ \vdots & & \\ a_k & * & * \\ \vdots & & \\ a_{n+1} & & \end{bmatrix}$.

Wenn $a_2 = \cdots a_{n+1} = 0$, dann ist $a_1 \neq 0$ und

$$D(1,2;\frac{a_1}{|a_1|},0)\;A = \begin{bmatrix} |a_1| & * \\ 0 & \\ \vdots & A' \\ 0 & \end{bmatrix} \qquad \text{(Drehung um } 180^\circ \text{ oder } \pi).$$

Dieselbe erste Spalte erreicht man aber auch mit $S(1,2;\frac{a_1}{|a_1|},0)$ statt $D(1,2;\frac{a_1}{|a_1},0)$.
Wenn etwa $a_k \neq 0$ für ein $k > 1$, dann ist

$$D(1,k;\frac{a_1}{\delta},\frac{a_k}{\delta})A = \begin{bmatrix} \delta & \\ a_2 & \\ \vdots & \\ a_{k-1} & * \\ 0 & \\ a_{k+1} & \\ \vdots & \\ a_{n+1} & \end{bmatrix} \qquad \text{mit} \quad \delta = \sqrt{a_1^2 + a_k^2} > 0$$

Wichtig: Die Einträge $a_2, \ldots, a_{k-1}, a_{k+1}, \ldots, a_{n+1}$ bleiben unberührt, und das Gleiche funktioniert auch mit”S” statt ”D”! Ist nun noch $a_l \neq 0,\ l \neq 1, l \neq k$ dann ist

$$D\left(1,l;\frac{\delta}{\delta'},\frac{a_l}{\delta'}\right)\quad D\left(1,k,\frac{a_1}{\delta},\frac{a_k}{\delta}\right)A = \begin{bmatrix} \delta & & \\ * & & \\ 0 & * & * \\ * & & \\ 0 & & \\ * & & \\ a_{n+1} & & \end{bmatrix}$$

und $\delta' = \sqrt{\delta^2 + a_\ell^2} > 0$; analog mit ”$S$” statt ”$D$”. Nach höchstens n Schritten ist erreicht, dass

$$Q_{\nu_1} \cdots Q_1 A = \begin{bmatrix} \delta_1 & * \\ 0 & \\ \vdots & A' \\ 0 & \end{bmatrix} \qquad \nu_1 \leq n$$

mit elementaren orthogonalen Matrizen $Q_1, \ldots, Q_{\nu_1}$ jeweils ”S” oder ”D”, wie wir wollen. Wenn $m = 1$, ist man fertig. Wenn $m > 1$ ist $A' \in \mathbb{R}^{n\times(m-1)}$. Es gibt dann nach Induktionsannahme $Q'_1, \cdots, Q'_{r'}$ derart, dass $Q'_{r'}, \cdots, Q'_1 A'$ in Zeilenstufenform ist mit positiven Eckeinträgen bis mindestens zur vorletzten Zeile und mit $r' \leq \frac{n(n-1)}{2}$. Insgesamt hat dann

$$\underbrace{\begin{bmatrix}1 & 0\\ 0 & Q'_{r'}\end{bmatrix}\cdots\begin{bmatrix}1 & 0\\ 0 & Q'_{1}\end{bmatrix} Q_{\nu_1}\cdots Q_1}_{r\text{ Faktoren}}\cdot A$$

die behauptete Form und es ist $r = \nu_1 + r' \leq n + \frac{n(n-1)}{2} = \frac{(n+1)((n+1)-1)}{2}$. □

Bemerkung: Ein stärkeres Ergebnis erhält man für Spiegelungen mit Householder-Matrizen. Man braucht höchstens $n-1$ solcher Matrizen (siehe Übungsaufgaben).

Eine Anwendung: Zu lösen sei ein lineares Gleichungssystem ”$Ax = b$” mit $A \in \mathbb{R}^{n\times m}$, $b \in \mathbb{R}^{n\times 1}$. Aus LA1 wissen wir

$$\mathsf{Lös}\,(A,b) = \varnothing \Leftrightarrow b \notin SR(A)$$

und wenn $b \in SR(A)$, dann gilt:

$$\mathsf{Lös}\,(A,b) = v + \underbrace{\mathsf{Kern}\,L_A}_{=\mathsf{Lös}\,(A,0)} \text{ für alle } v \in \mathsf{Lös}\,(A,b)$$

Wir versehen $\mathbb{R}^{n\times 1}$, $\mathbb{R}^{m\times 1}$ mit dem Standardskalarprodukt. Eine verallgemeinerte Aufgabe ist dann:

Suche den kürzesten Vektor $v^* \in \mathbb{R}^{n\times 1}$ so, dass zugleich $Av^* - b$ am kürzesten ist unter allen Vektoren $Av - b$, $v \in \mathbb{R}^{n\times 1}$

Es besteht ein direkter Zusammenhang zur sogenannten Ausgleichsrechnung, siehe etwa [MV], S. 219 ff. für eine detaillierte Darstellung im Fall $m = 2$, und $\mathsf{Rang}\,A = 2$. Die Anwendung von Satz 2.17 ermöglicht Transformation der Aufgabenstellung: ”$Ax = b$”→ ”$PAx = Pb$” mit längentreuem L_P und PA in Zeilenstufenform.

Rechen-Beispiel 2.18: $A = \begin{bmatrix}1 & * & *\\ 4 & * & *\\ 3 & * & *\end{bmatrix}$: $Q_1\begin{bmatrix}\frac{1}{\sqrt{17}} & \frac{4}{\sqrt{17}} & 0\\ -\frac{4}{\sqrt{17}} & \frac{1}{\sqrt{17}} & 0\\ 0 & 0 & 1\end{bmatrix} = D\left(1,2;\frac{1}{\sqrt{17}},\frac{4}{\sqrt{17}}\right)$

$$Q_1A = \begin{bmatrix}\sqrt{17} & \\ 0 & *\\ 3 & \end{bmatrix},\quad Q_2 = \begin{bmatrix}\frac{\sqrt{17}}{\sqrt{26}} & 0 & \frac{3}{\sqrt{26}}\\ 0 & 1 & 0\\ -\frac{3}{\sqrt{26}} & 0 & \frac{\sqrt{17}}{\sqrt{26}}\end{bmatrix},\ Q_2Q_1A = \begin{bmatrix}\sqrt{26} & \\ 0 & *\\ 0 & \end{bmatrix}.$$

Beachte: Die 1 in A verleitet, mit elementaren Umformungen zu arbeiten, diese sind aber nicht alle orthogonal/längentreu, insbesondere nicht die von Typ 2: $Q^i_j(\lambda)$, vgl. etwa [F], S. 164.
Als Folgerung erhalten wir:

Satz 2.19: **(a)** Sei $A \in \mathcal{O}_1(n)$. Dann ist $A = D_1\cdots D_r$ mit elementaren Drehungen $D_1,\dots,D_r$ und $1 \leq r \leq \frac{n(n-1)}{2}$.

(b) Sei $A \in \mathcal{O}(n)\backslash\mathcal{O}_1(n)$: Dann ist: $A = S_1\cdots S_r$ mit el. Spiegelungen $S_1,\dots,S_r$ und $1 \leq r \leq \frac{n(n-1)}{2} + 1$ und r ungerade.
Es ist aber auch:
$A = D_1\cdots D_r\cdot S(1,n;1,0)$ mit elementaren Drehungen, $D_1,\dots,D_r$ und $1 \leq r \leq \frac{n(n-1)}{2}$.

Beweis: Nach Satz 2.17 ist mit geeigneten elementaren orthogonalen Matrizen $Q_1, \dots, Q_r$ die Matrix $B := Q_r \cdots Q_1 A$ in Zeilenstufenform mit positiven Eckeinträgen bis zur vorletzten Zeile. A ist in (a) und (b) orthogonal also auch B. Insbesondere ist B invertierbar, also obere Dreiecksmatrix mit von Null verschiedenen Diagonaleinträgen. und - da orthogonal - eine Diagonalmatrix. Nur der letzte Diagonaleintrag ist u. U. nicht positiv, also ist: $B = \operatorname{diag}(1, \dots, 1, \varepsilon)$ mit $\varepsilon \in \{1, -1\}$.
(a) Jetzt ist $\det A = 1$. wir wählen die Q_i als el. Drehungen, dann ist $\det Q_1 = 1$, für $1 \le i \le r$ und $\det B = 1, \varepsilon = 1$ also: $B = E_n$ und $A = Q_1^{-1} \dots Q_r^{-1}$. Nach Beobachtung 2.16 sind auch die Q_i^{-1} elementare Drehungen. Der Rest der Behauptung folgt nun mit Satz 2.17.
(b) Jetzt ist $\varepsilon = -1,\ B = \operatorname{diag}(1, \dots, 1, -1) = S(1, n; 1, 0)$. Die Behauptungen ergeben sich jetzt analog zu (a). ☐

Bemerkungen: Die zu den Drehungen in Satz 2.19 (a) gehörenden Drehwinkel heißen **"Eulersche Winkel" von A oder L_A.** Allerdings muss dann ein ganz bestimmtes Trigonalisierungsvorgehen bei Satz 17 festgelegt werden, damit Eindeutigkeit besteht. Der Fall $n = 3$ ist dabei ein Sonderfall, denn nur hier ist $\frac{n(n-1)}{2} = n$. Man überlege sich, dass im Fall $n = 3$ bei (a) in Satz 2.19 stets $r \le 2$ ausreicht! Anwendungen der Eulerschen Winkel gibt es u. A. in Mechanik und Robotik.

Nun zu Frage (I) vor Definition 2.15:
"Block-Diagonalisierung durch orthogonalen Basiswechsel"

Definition 2.20: $A \in \mathbb{R}^{n \times n}$ heißt **normal,** wenn ${}^{\mathsf{t}}AA = A{}^{\mathsf{t}}A$. (vgl. [F], S. 344)

Beispiele:

(a) Symmetrische Matrizen, orthogonale Matrizen, schiefsymmetrische Matrizen

(b) $n = 2 : A \in \mathbb{R}^{2\times 2}$ normal, dann ist $A = \begin{bmatrix} a & b \\ \varepsilon_2 b & \varepsilon_1 a \end{bmatrix}$ mit $\varepsilon_1\varepsilon_2 \in \{1, -1\}$ **und** $\varepsilon_1\varepsilon_2 = -1$ (siehe Satz 21 und Beweis).

Satz 2.21: Orthogonale Diagonalisierung. Sei $A \in \mathbb{R}^{n\times n}$ normal, dann gibt es $P \in \mathcal{O}(n)$ derart, dass

$$P^{-1}AP = \left[\begin{array}{c|c} \operatorname{diag}(\lambda_1, \dots, \lambda_s) & 0 \\ \hline 0 & \operatorname{diag}(B_1, \dots, B_r) \end{array}\right]$$

mit $\lambda_1, \dots, \lambda_s \in \sigma_{\mathbb{R}}(A) = \{\text{reelle Eigenwerte von } A\}$ und normalen 2×2-Matrizen $B_1, \dots, B_r$ ohne reelle Eigenwerte. Letztere haben die Form $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ mit $b \neq 0$.
Dabei ist $s = 0$, falls $\sigma_{\mathbb{R}}(A) = \varnothing$ und $r = 0$, falls A diagonalisierbar ist.
Für die **Sonderfälle** "symmetrisch" und "orthogonal" s. u. Satz 2.22.

Da es im Satz 2.21 um eine Ähnlichkeitstransformation geht, ist die Verwendung elementarer orthogonaler Matrizen schwierig, denn mit den Zeilen müssen immer zugleich auch die Spalten entsprechend umgeformt werden. Wir gehen vor wie in LA 1 und benutzen unsere Kenntnisse über Eigenwerte und Eigenvektoren.
Beweis: Induktion nach n: $\boldsymbol{n = 1 :}$ wähle $P = 1$.
$\boldsymbol{n \ge 1 :}$ Die Behauptung gelte für normale Matrizen aus $\mathbb{R}^{n\times n}$. Sei jetzt $A \in \mathbb{R}^{(n+1)\times(n+1)}$.

Fall 1: $\boldsymbol{\sigma_{\mathbb{R}}(A) \neq \varnothing}$: Sei v_1 Eigenvektor zu einem $\lambda \in \sigma_{\mathbb{R}}(A)$ und $||v_1|| = 1$.
Ergänze v_1 zu einer Orthonormalbasis $(v_1, \ldots, v_{n+1})$ von $\mathbb{R}^{n+1}$ und setze $P_1 = [v_1, \ldots, v_{n+1}]$.
Dann gilt:

$$P_1^{-1} = {}^tP_1 \text{ und } {}^tP_1 A P_1 = \begin{bmatrix} \lambda & u \\ 0 & A_1 \end{bmatrix} =: B$$

Auch B ist normal. Der (1,1)-Eintrag von $B{}^{\mathsf{t}}B - {}^{\mathsf{t}}BB$ ist $u{}^{\mathsf{t}}u = \parallel u \parallel^2$. Daher muss $u = 0$ sein und auch A_1 ist normal. Mit Hilfe der Induktionsannahme ergibt sich nun die Behauptung für $n + 1$.

Fall 2: $\sigma_{\mathbb{R}}(A) = \varnothing$. $n + 1$ ist dann gerade!

$\underline{n + 1 = 2 :}$ Sei etwa $A : \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. A ist normal, also gilt

$${}^{\mathsf{t}}AA = \begin{bmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{bmatrix} = \begin{bmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix} = A \cdot {}^{\mathsf{t}}A$$

Es folgt: $b^2 = c^2$ und $(a - d)(b - c) = 0$. Wenn $b = c$: dann ist A symmetrisch und hat $\lambda = \frac{1}{2}\left(a + d + \sqrt{(a - d)^2 + 4b^2}\right)$ als reellen Eigenwert. W! Also ist $b = -c$ und $a = d$ und entsprechend $A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$. Da $\sigma_A = \varnothing$ sein soll, ist $b \neq 0$. (A hat die komplexen Eigenwerte $a \pm ib$.)

$\underline{n + 1 \geq 4 :}$ Vorbemerkung: χ_A hat stets Nullstellen in $\mathbb{C}$. Da $\chi_A \in \mathbb{R}[t]$, ist mit λ stets auch $\overline{\lambda}$ Nullstelle von χ_A. Sei $\lambda \in \sigma_{\mathbb{C}}(A)$ und $v \in \mathbb{C}^{n+1}$ ein Eigenvektor. $\overline{v}$ ist dann Eigenvektor zum Eigenwert $\overline{\lambda}$. Seien $v_1 = \frac{1}{2}(v + \overline{v})$ und $v_2 = \frac{i}{2}(v - \overline{v})$, dann sind v_1, v_2 reelle Vektoren. Andersherum ist $\overline{v} = v_1 + iv_2$ und $v = v_1 - iv_2$. Sei $W := < v_1, v_2 >_{\mathbb{R}}$.
Dann gilt:

$$\text{(i)} \quad \mathsf{dim}_{\mathbb{R}} W = 2 \qquad \text{(ii)} \quad AW \subseteq W \qquad \text{(iii)} \quad {}^{\mathsf{t}}AW \subseteq W$$

Beweis:

(i) Nach Voraussetzung ist $\lambda \neq \overline{\lambda}$, also ist das Paar der Eigenvektoren $(v, \overline{v})$ $\mathbb{C}$-linear unabhängig. Da $< v, \overline{v} >_{\mathbb{C}} = < v_1, v_2 >_{\mathbb{C}}$, folgt (v_1, v_2) ist $\mathbb{C}$-linear unabhängig und umsomehr $\mathbb{R}$-linear unabhängig.

(ii)
$$\begin{aligned} Av_1 &= A(\tfrac{1}{2}v + \overline{v}) = \tfrac{1}{2}(\lambda v + \overline{\lambda}\overline{v}) = \tfrac{1}{2}((\lambda_1 + i\lambda_2)v + (\lambda_1 - i\lambda_2)\overline{v}) \\ &= \tfrac{1}{2}\lambda_1(v + \overline{v}) + \tfrac{1}{2}\lambda_2 i(v - \overline{v}) = \lambda_1 v_1 + \lambda_2 v_2 \in W, \end{aligned}$$
analog für Av_2.

(iii) Es genügt zu zeigen, dass $v, \overline{v}$ auch $\mathbb{C}$-Eigenvektoren von ${}^{\mathsf{t}}A$ sind. Dies geht mit demselben Trick, wie im Fall 1, allerdings jetzt komplex gerechnet:

Zu $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C}^n$ ist dann $||x|| = \sum\limits_{i=1}^{n} \underbrace{\overline{x}_i x_i}_{\text{reell}}$ reell $= {}^{\mathsf{t}}\overline{x}x$

In unserem Fall ist:

$$\begin{aligned} 0 &= ||(A - \lambda E_{n+1})v|| = {}^{\mathsf{t}}\overline{v}\,{}^{\mathsf{t}}(A - \overline{\lambda}E_{n+1})(A - \lambda E_{n+1})v \\ &= {}^{\mathsf{t}}\overline{v}(A - \lambda E_{n+1}){}^{\mathsf{t}}(A - \overline{\lambda}E_{n+1})v \\ &= ||({}^{\mathsf{t}}A - \overline{\lambda}E_{n+1})v|| \,. \end{aligned}$$

Analog für $\overline{v}$. Da ${}^{\mathsf{t}}A$ auch reell ist folgt ${}^{\mathsf{t}}AW \subseteq W$.

Berechne nun eine Orthonormalbasis (w_1, w_2) von W und ergänze sie zu einer Orthonormalbasis $(w_1, \dots w_{n+1})$ von $\mathbb{R}^{n+1}$. Setze $P = [w_1, \dots, w_{n+1}]$ und erhalte:

$$P^{-1}AP = \begin{bmatrix} B_1 & 0 \\ 0 & A_1 \end{bmatrix} \text{ mit } P^{-1} = {}^{\mathsf{t}}P.$$

Auch $P^{-1}AP$ ist normal und somit auch A_1. Mit der Induktionsannahme für A_1 ergibt sich jetzt die Behauptung für A. ☐

Sonderfälle von Satz 2.21:
Wenn $\boldsymbol{A \in \mathcal{O}(n)}$, dann sind die B_i, falls sie vorkommen, alle orthogonal und ohne reelle Eigenwerte. Die B_i sind daher Drehungsmatrizen mit Drehwinkel $\neq 0, \pi$.
Wenn A symmetrisch ist, dann kommen keine B_i vor $(r = 0)$ denn die B_i müssten dann symmetrisch sein und symmetrische 2×2-Matrizen haben stets einen reellen Eigenwert.
Man erhält:

Satz 2.22: Spektralsatz. Symmetrische Matrizen sind orthogonal diagonalisierbar. Insbesondere gibt es eine Orthonormalbasis aus Eigenvektoren, und es ist: $\sigma_{\mathbb{C}}(A) = \sigma_{\mathbb{R}}(A)$.

Beispiel 2.23: Sei $A = \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$. A ist schiefsymmetrisch. Man berechnet der Reihe nach: $\sigma(A) = \{0\}$, $\mathsf{Eig}\,(A, 0) = \mathsf{Kern}\, L_A =< \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} >_{\mathbb{R}} =< \underbrace{\frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}}_{\substack{\text{Länge1} \\ =: v_1}} >_{\mathbb{R}}$,

$$P := \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{\sqrt{2}}{\sqrt{3}} \end{bmatrix} \text{ und } P^{-1}AP = \left[\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & 0 & \sqrt{3} \\ 0 & -\sqrt{3} & 0 \end{array}\right]. \text{ Der Block } \begin{bmatrix} 0 & \sqrt{3} \\ -\sqrt{3} & 0 \end{bmatrix}$$

entspricht einer (winkeltreuen aber nicht orthogonalen) Drehstreckung.

Beispiel 2.24: Sei $A := \begin{bmatrix} \frac{9}{25} & \frac{12}{125} & \frac{116}{125} \\ -\frac{12}{25} & \frac{109}{125} & \frac{12}{125} \\ \frac{4}{5} & \frac{12}{25} & -\frac{9}{25} \end{bmatrix}$. Man rechnet nach, dass gilt:
$A = Q_1 \cdot Q_2 \cdot Q_3$ mit den elementaren orthogonalen Matrizen,

$$Q_1 := \begin{bmatrix} \frac{3}{5} & \frac{4}{5} & 0 \\ \frac{-4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad Q_2 := \begin{bmatrix} \frac{3}{5} & 0 & \frac{4}{5} \\ 0 & 1 & 0 \\ \frac{4}{5} & 0 & \frac{-3}{5} \end{bmatrix}, \quad Q_3 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{5} & \frac{4}{5} \\ 0 & \frac{-4}{5} & \frac{3}{5} \end{bmatrix}.$$

A hat den Eigenwert -1 mit normiertem Eigenvektor $w_1 := \left[\frac{6}{11}, \frac{2}{11}, -\frac{9}{11}\right]$.
A hat das konjugiert komplexe Eigenwertpaar $\frac{117}{125} + i\frac{44}{125}$, $\frac{117}{125} - i\frac{44}{125}$.
Der dazugehörende A-invariante reelle Untervektorraum W hat als orthonormale Basis:

$$w_2 := \left[-\frac{4\sqrt{13}}{143},\ \frac{3\sqrt{13}}{11},\ \frac{6\sqrt{13}}{143}\right],\ w_3 := \left[\frac{3\sqrt{13}}{13},\ 0, \frac{2\sqrt{13}}{13}\right].$$

Die Matrix P mit w_1, w_2, w_3 als Spalten führt nun zu

$$P^{(-1)}AP = \begin{bmatrix} -1 & 0 & 0 \\ 0 & \frac{117}{125} & \frac{-44}{125} \\ 0 & \frac{44}{125} & \frac{117}{125} \end{bmatrix}.$$

Bemerkung:
Dieses Beispiel ist natürlich extra so konstruiert, dass "schöne" Zahlen vorkommen. Dies ist aber untypisch, da normalerweise die komplexen Eigenwerte nur näherungsweise berechnet werden können und auch die notwendigen Normierungen (Wurzelberechnungen) zwangsläufig zu "unschönen" Zahlen führen.

Beispiel 2.25: Sonderfall $\boldsymbol{n = 3}$, $A \in \mathcal{O}(3)$. χ_A hat Grad 3 und somit stets mindestens eine reelle Nullstelle $\lambda \in \{1, -1\}$. Wie im Beweis von Satz 2.21, Fall 1 erhält man dann mit einem $P \in \mathcal{O}(3)$:

$$D := P^{-1}AP = \left[\begin{array}{c|cc} \lambda & 0 & 0 \\ \hline 0 & a & b \\ 0 & c & d \end{array}\right]$$

mit $\underbrace{\begin{matrix} a & b \\ c & d \end{matrix}}_{A'}$ als unterem rechten Block.

Da auch $D \in \mathcal{O}(3)$ ist $A' \in \mathcal{O}(2)$ also (siehe Beispiel 11(b)) $a = d,\ c = -b$ und $a^2 + b^2 = 1$ oder $a = -d,\ b = c$ und $a^2 + b^2 = 1$.
Mehr Details zu $\boldsymbol{n = 3}$ bei [F], S. 306/307 → "Satz vom Fußball".

Bemerkung 2.26: Umkehrung zu Satz 2.21.

- **(a)** Wenn $A \in \mathbb{R}^{n\times n}$ orthogonal diagonalisierbar ist, dann ist A normal.
- **(b)** Seien $A \in \mathbb{R}^{n\times n},\ P \in \mathcal{O}(n),\ B := {}^{\mathsf{t}}PAP$. Hat B die Gestalt aus Satz 2.21 (mit normalen B_i), dann ist A normal.

Definition 2.27: Selbstadjungierte lineare Abbildungen. Sei V ein euklidischer Vektorraum. Eine lineare Abbildung $F : V \to V$ mit der Eigenschaft

$$(\!(v\,, F(w))\!) = (\!(F(v)\,, w)\!) \text{ für alle } v, w \in V$$

heißen **selbstadjungiert.**

Beobachtung 2.28: Ist V endlichdimensional, dann ist F genau dann selbstadjungiert, wenn die Matrix von F bezüglich einer Orthonormalbasis von V symmetrisch ist.

Beweis: Seien $\Phi : \mathbb{R}^{n\times 1} \to V$ die Koordinatenabbildung zur Orthonormalbasis $\mathcal{A} = (v_1, \ldots, v_n)$ und A die Matrix von F bezüglich $\mathcal{A}$. Dann stellt man u.A: mit Hilfe von Beobachtung 2.12 fest

$${}^{\mathsf{t}}e_iAe_j = (\!(e_i\,, Ae_j)\!) = (\!(\Phi(e_i)\,, \Phi(Ae_j))\!) = (\!(v_i\,, F(v_j))\!)$$

$$= \langle\!\langle F(v_i)\,, v_j \rangle\!\rangle = \langle\!\langle \Phi(Ae_i)\,, e_j \rangle\!\rangle = \langle\!\langle Ae_i\,, e_j \rangle\!\rangle = {}^{\mathsf{t}}e_i{}^{\mathsf{t}}Ae_j$$

Also ist A symmetrisch, wenn F selbstadjungiert ist. Die Umkehrung bleibt als Übung. □

Zusammenfassung §12:

- Metrische lineare Abbildungen
- Isometrien, Standardfall: $\mathbb{R}^{m+1} \to \mathbb{R}^{m+1}$ mit Standardskalarprodukt
- $\mathcal{O}(n)$, $\mathcal{O}(2)$ und geometrische Interpretation
- elementaren orthogonalen Matrizen (sogenannte "Givens"-Drehungen im D-Fall)
- Matrixdarstellungen von Isometrien
- Zeilenstufenform durch elementare orthogonale Zeilenumformungen
- Anwendung: Ausgleichsrechnung
- Erzeugung von $\mathcal{O}(n)$, $\mathcal{O}_1(n)$ durch elementare orthogonale Matrizen. (Satz 2.19)
- Eulersche Winkel
- "orthogonale Diagonalisierbarkeit" normaler Matrizen (Satz 2.21)
- Spezialfall: symmetrische Matrizen

Offen bleibt u. A. "orthogonale Äquivalenz", dazu mehr im nächsten Abschnitt.

3 SVD und Rayleigh-Quotient

Definition 3.1: **(a)** $A, B \in \mathbb{R}^{n\times n}$ heißen **orthogonal ähnlich,**
wenn mit $P \in \mathcal{O}(n)$ gilt: ${}^{\mathsf{t}}PAP = B$.
(b) $A, B \in \mathbb{R}^{m\times n}$ heißen **orthogonal äquivalent,**
wenn mit $P \in \mathcal{O}(m)$ und $Q \in \mathcal{O}(n)$ gilt $PAQ = B$.
In beiden Fällen handelt es sich um Äquivalenzrelationen.

Die Suche nach entsprechenden Normalformen gehört zu den Grundaufgaben der linearen Algebra. Normalformen für die orthogonale Äquivalenz und die orthogonale Ähnlichkeit sind von besonderem Interesse für Anwendungen,[2] da orthogonale Koordinatenwechsel i.A. die Problemparameter nicht "verzerren". Im vorigen Abschnitt ergaben sich bereits wichtige Informationen für die orthogonale Ähnlichkeit bei speziellen Matrizentypen. Jetzt untersuchen wir die orthogonale Äquivalenz für eine beliebige Matrix $A \in \mathbb{R}^{m\times n}$.

Satz 3.2: Singulärwertzerlegung (SVD). Sei $A \in \mathbb{R}^{m\times n}\backslash\{0\}$.
Dann gibt es $P \in \mathcal{O}(m)$, $Q \in \mathcal{O}(n)$ und $s_1, \ldots, s_r \in \mathbb{R}$, derart, dass

$$0 < s_1 \leq \cdots \leq s_r$$

---

[2] z.B. bei Simulation, Numerik, Matrizenarithmetik, angewandte Stochastik, Optimierung, Bildkompression, ...

und

$$PAQ = \begin{bmatrix} s_r & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & s_1 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \qquad (m \times n\text{-Diagonalmatrix}) \ . \qquad (2)$$

$s_1, \ldots, s_r$ sind durch A **eindeutig** festgelegt und heißen **Singulärwerte von A**. Dabei gilt

$$\{s_1, \ldots, s_r\} = \{\sqrt{\mu} : \ \mu \in \sigma({}^{\mathsf{t}}A\,A) \setminus \{0\}\} \ .$$

Die Zerlegung $A = {}^{\mathsf{t}}P\text{diag}(s_r, \ldots s_1, \ 0, \ldots){}^{\mathsf{t}}Q$ heißt **Singulärwertzerlegung von A, Singular Value Decomposition "SVD"**

In dem Buch von [K], S. 199-200, findet man eine Version von Satz 3.2, die dort für invertierbare A dem französischen Mathematiker E. Cartan (1859 - 1951) zugeschrieben wird. Demnach wäre es angebrachter, von der **Cartan-Normalform** zu sprechen statt den schwer motivierbaren Namen "Singulärwertzerlegung" zu benutzen. Zum historischen Hintergrund, einigen Anwendungen vor 1992 und der "seltsamen" Namensgebung, siehe etwa [St]. Eine gründliche Darstellung der SVD und ihrer Rolle in der angewandten Mathematik ist als Kapitel 4.3 in dem Buch [HP] enthalten. Hinweise zu Anwendungen der SVD sind u. a. auch zu finden in: [M], [MV].

Beweis in mehreren Schritten:
Existenz:

(a) Behauptung: Man kann o. E annehmen: $m = n$ und $A \in GL_n(\mathbb{R})$.
Bemerkung: Ein direkter Beweis für beliebige m, n ist unter Umständen "schöner", aber meist schwerer zu verstehen (siehe z.B. [Se] S. 128 ff.). Unser Vorgehen zeigt zugleich, wie man rechnen kann.
Beweis: Wir benutzen die orthogonale Zeilenstufenform von Satz 2.17:
Sei $A \neq 0$.

$$\begin{aligned} &\exists P \in \mathcal{O}(n) : \quad PA = \begin{bmatrix} Z \\ 0 \end{bmatrix} \ , \ Z \text{ mit linear unabhängigen Zeilen.} \\ &\exists Q \in \mathcal{O}(m) : \quad {}^{\mathsf{t}}Q\ {}^{\mathsf{t}}Z \text{ in Zeilenstufenform bzw.} ZQ = [S, 0] \\ &\qquad\qquad\qquad\quad S \text{ mit linear unabhängigen Spalten.} \end{aligned}$$

Insgesamt ergibt dies: $PAQ = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}$ mit $S \in \mathbb{R}^{r \times r}$ und $r =$ Rang $A =$ Rang S, insbesondere $S \in Gl(r, \mathbb{R})$. Gelingt es, die Existenz einer SVD für S nachzuweisen, so ergibt sich demnach die Existenz einer SVD auch für A.
Sei ab jetzt: $A \in Gl_n(\mathbb{R})$.

(b) Da $A \in Gl_n(\mathbb{R})$, gilt für $x \in \mathbb{R}^{n \times 1} \setminus \{0\}$: ${}^{\mathsf{t}}x{}^{\mathsf{t}}AAx = \| Ax \|^2 > 0$. Damit folgt: $\sigma({}^{\mathsf{t}}AA) \subset \mathbb{R}_+$, denn ist x ein Eigenvektor zum Eigenwert λ von ${}^{\mathsf{t}}AA$, dann gilt:

$$0 < \| Ax \|^2 = {}^{\mathsf{t}}x \underbrace{({}^{\mathsf{t}}AAx)}_{=\lambda x} = \lambda{}^{\mathsf{t}}xx = \lambda \| x \|^2.$$

Es folgt $\lambda > 0$.

Beachte: nicht alle symmetrischen Matrizen haben positive Eigenwerte; z. B. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

(c) Nach Satz 2.22 ist ${}^{\mathsf{t}}AA$ orthogonal ähnlich zu $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Zwangsläufig gilt nach (b): $0 < \lambda_i$ für $1 \le i \le n$. Mit elementaren Spiegelungen $P_1, \ldots, P_s$ erreicht man noch:

$$P_s \cdots P_1 \mathrm{diag}(\lambda_1, \ldots, \lambda_n) P_1 \cdots P_s = \mathrm{diag}(\mu_n, \cdots, \mu_1) \text{ mit } 0 < \mu_1 \le \cdots \le \mu_n.$$

Setze $s_i := {}_{+}\sqrt{\mu_i}$ und $D := \mathrm{diag}(s_n, \ldots, s_1)$. Zur Erklärung des Vorstehenden sei daran erinnert, dass $P_i^{-1} = P_i$ und das $S(i, j, 0, 1) = P_j^i$.

(d) Sei (vgl. Satz 2.22) $(v_n, \ldots, v_1)$ eine Orthonormalbasis von Eigenvektoren von ${}^{\mathsf{t}}AA$. Dann ist v_i Eigenvektor zum Eigenwert $s_i^2 = \mu_i$.
Setze $Q := [v_n, \ldots, v_1]$ und $P := D^{-1}\,{}^{\mathsf{t}}Q\,{}^{\mathsf{t}}A$. Dann gilt $({}^{\mathsf{t}}A \cdot A)Q = QD^2$ und

$$PAQ = (D^{-1}\,{}^{\mathsf{t}}Q\,{}^{\mathsf{t}}A)AQ = D^{-1}\,{}^{\mathsf{t}}Q({}^{\mathsf{t}}AAQ) = D^{-1}\,{}^{\mathsf{t}}QQD^2 = D$$

Die letzte Gleichung ergibt sich, weil nach Konstruktion ${}^{\mathsf{t}}QQ = E_n$. Dabei ist $Q \in \mathcal{O}(n)$ und, da außerdem

$$P^{\mathsf{t}}P = D^{-1}\,{}^{\mathsf{t}}Q^{\mathsf{t}}\underbrace{AAQ}\,D^{-1} = D^{-1}\,\underbrace{{}^{\mathsf{t}}QQ}_{=E_n}\,D^2D^{-1} = E_n,$$

also auch $P \in \mathcal{O}(n)$.

Eindeutigkeit: Wenn $PAQ = D$ bzw. $A = {}^{\mathsf{t}}PD^{\mathsf{t}}Q$, dann ist ${}^{\mathsf{t}}AA = QD\underbrace{P^{\mathsf{t}}P}_{E_n}D^{\mathsf{t}}Q = QD^2\,{}^{\mathsf{t}}Q$. Es folgt $\chi_{{}^{\mathsf{t}}AA} = \prod_{i=1}^{n}(t - \mu_i) = \chi_{D^2}$ □

Ein völlig anderer Beweis ergibt sich mit Hilfe des sogenannten Rayleigh-Quotienten (J. W. Rayleigh 1842 - 1919) und Matrix-Normen. Da die Begriffe sowieso wichtig sind, werden Sie jetzt eingeführt und dann Satz 3.2 nochmal bewiesen.

Definition 3.3: **(a)** $A \in \mathbb{R}^{m \times n}$
$\| A \|_2 := \max\limits_{x \ne o} \dfrac{\| Ax \|}{\| x \|}$ heißt (2-)Norm von A. Dies ist die Standardmatrixnorm zum Standardskalarprodukt.
Vorsicht! Dass das Maximum angenommen wird, wird unten erst gezeigt, kann aber auch direkt mit analytischen Überlegungen erkannt werden.

(b) Sei $A \in \mathbb{R}^{n \times n}$**symmetrisch.** Die Abbildung $R_A : \mathbb{R}^{n \times 1} \backslash \{0\} \to \mathbb{R}$ mit $R_A(x) = \frac{{}^{\mathsf{t}}xAx}{{}^{\mathsf{t}}xx}$ heißt **Rayleighquotient** zur Matrix A.

(c) Die **Frobenius-Norm** einer Matrix $A \in \mathbb{R}^{m \times n}$ ist die Länge des 'Vektors' $A \in \mathbb{R}^{m \times n}$ bzgl. dem Standardskalarprodukt:

$$\| A \|_F = \sqrt{\sum_{\substack{1 \le i \le m \\ 1 \le j \le n}} a_{ij}^2} \ .$$

Beispiel 3.4: **(a)** $\| E_n \|_2 = 1, \| E_n \|_F = \sqrt{n},\ R_{E_n}(x) = 1$.

(b) Sei $A = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$. dann ist $\| A \|_F = \sqrt{2 + s^2}$. Was ist $\| A \|_2$? Behauptung: $\| A \|_2 = 1 + \frac{s^2}{2} + \frac{1}{2}\sqrt{4s^2 + s^4}$. Dies ergibt sich mit dem folgenden Satz 3.6.

(c) Normberechnung in Maple: z. B.: so:

`with(linalg): A:=matrix(.......);norm(A,2);`

(d) Sei A symmetrisch und $Ax = \lambda x$ für ein $x \neq 0$, dann ist $R_A(x) = \lambda$.

Beobachtung 3.5: $\| A \|_F$ und $\| A \|_2$ sind sogenannte Matrixnormen, d.h. es gilt (mit $\| \ \|$ statt $\| \ \|_F$ oder $\| \ \|_2$) :
(i) $\forall A \in \mathbb{R}^{m\times n} : \| A \| = 0 \Leftrightarrow A = 0$.
(ii) $\forall \lambda \in \mathbb{R},\ A,\ B \in \mathbb{R}^{m\times n} : \| \lambda A \| = |\lambda| \| A \|$ und $\| A + B \| \leq \| A \| + \| B \|$.
(iii) $\forall A \in \mathbb{R}^{m\times n},\ B \in \mathbb{R}^{n\times p},\ \lambda \in \mathbb{R}$:

$$\| Ax \| \leq \| A \| \| x \| \text{ und } \| AB \| \leq \| A \| \cdot \| B \|$$

Bei (iii) wird angenommen, dass "gleichartige" Normen in $\mathbb{R}^{m\times n}$ und $\mathbb{R}^n$ vorliegen. Bezüglich (i) und (ii) sei auch auf [F], Seite 275 verwiesen.

Beweis: Übungsaufgabe. ((iii) für $\| \ \|_2$ trivial; für $\| \ \|_F$ in Beweis von Satz 3.6 (b) enthalten.)

Satz 3.6: (a) sei $A \in \mathbb{R}^{m\times n}$ **symmetrisch.** Sei $\lambda_1 = \min \sigma(A)$ und $\lambda_n = \max \sigma(A)$.
(i) $\lambda_1 = \min\limits_{\times\neq 0} R_A(x)$ und $\lambda_n = \max\limits_{x\neq 0} R_A(x)$.
(ii) Für alle $x \in \mathbb{R}^{n\times 1}\backslash\{0\}$: $\begin{array}{ccc} R_A(x) = \lambda_1 & \Leftrightarrow & x \in \mathsf{Eig}\,(A, \lambda_1) \\ R_A(x) = \lambda_n & \Leftrightarrow & x \in \mathsf{Eig}\,(A, \lambda_n) \end{array}$ beachte: λ_1, λ_n speziell!

(b) Sei $A \in \mathbb{R}^{m\times n}$, dann ist

$$\| A \|_2 = \max\left\{\sqrt{\lambda} : \lambda \in \sigma({}^{\mathsf{t}}AA)\right\} = \sqrt{\max\{R_{{}^{\mathsf{t}}AA}(x) : x \in \mathbb{R}^{n\times 1}\backslash\{0\}\}} \leq \| A \|_F$$

Man beachte den Zusammenhang mit den Singulärwerten: $\| A \|_2$ = größter Singulärwert.

Beweis:

(a) Da A symmetrisch ist, gibt es eine Orthonormalbasis $(w_1, \ldots, w_n)$ von Eigenvektoren (Satz 2.22, die wir so so ordnen, dass w_1 zum kleinsten Eigenwert λ_1 und w_n zum größten Eigenwert λ_n gehört: $\lambda_1 \leq \cdots \leq \lambda_n$, u. U. $\lambda_1 = \lambda_n$!

Mit $x = \sum\limits_{i=1}^{n} x_i w_i$ gilt:

$$\begin{aligned} {}^{\mathsf{t}}xAx &= {}^{\mathsf{t}}\Big(\sum_{i=1}^{n} x_i w_i\Big) A \Big(\sum_{j=1}^{n} x_j w_j\Big) &= {}^{\mathsf{t}}\Big(\sum_{i=1}^{n} x_i w_i\Big)\Big(\sum_{j=1}^{n} x_j A w_j\Big) \\ &= {}^{\mathsf{t}}\Big(\sum_{i=1}^{n} x_i w_i\Big)\Big(\sum_{j=1}^{n} \lambda_i x_i w_i\Big) &= \sum_{i=1}^{n} \lambda_i x_i^2 \end{aligned}$$

und ${}^{\mathsf{t}}xx = \| x \|^2$. Daher ist

$$R_A(x) = \sum_{i=1}^{n} \lambda_i \underbrace{\frac{x_i^2}{\| x \|^2}}_{\geq 0} \tag{3}$$

Es folgt $\lambda_1 \leq \lambda_1 \dfrac{\| x \|^2}{\| x \|^2} = \sum\limits_{i=1}^{n} \lambda_1 \dfrac{x_i^2}{\| x \|^2} \underset{\overset{\uparrow}{\lambda_1 \leq \lambda_i}}{\leq} \underbrace{\sum_{i=1}^{n} \lambda_i \frac{x_i^2}{\| x \|^2}}_{=R_A(x)} \underset{\overset{\uparrow}{\lambda_i \leq \lambda_n}}{\leq} \sum\limits_{i=1}^{n} \lambda_n \dfrac{x_i^2}{\| x \|^2} \leq$

λ_n .

Außerdem ergibt (1) für $x = w_1 : R_A(w_1) = \lambda_1$ und für $x = x_n : R_A(w_n) = \lambda_n$. Insgesamt ergibt sich (i).
zu (ii): Wenn $R_A(x) = \lambda_1$, dann muss in (1) gelten:

$$x_k = 0 \text{ falls } \lambda_k \neq \lambda_1 \, ,$$

denn dann ist $\lambda_1 < \lambda_k$. Es folgt $x = \sum\limits_{\substack{w_i \text{Eigenvektor} \\ \text{zum Eigenwert} \lambda_1}} x_i w_i$ und $Ax = \lambda_1 x$, analog, wenn $R_A(x) = \lambda_n$. Die Umkehrung ist jeweils trivial (vgl. Beispiel 3.4(d)).

(b) Aufgrund von (a) ergibt sich: $\| A \|_2^2 = \max\limits_{x \neq 0} \frac{{}^t x {}^t A A x}{{}^t x x} = \max\limits_{x \neq 0} R_{{}^t AA}(x) = \max \sigma({}^t AA)$.
Zu zeigen ist nun noch $\| A \|_2 \leq \| A \|_F$.
Für $x \in \mathbb{R}^{n \times 1}$ gilt, wenn etwa $A = (a_{ij})$:

$$\| Ax \|^2 = \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} a_{ij} x_i \Big)^2 \leq \sum_{i=1}^{n} \Big(\sum_{j=1}^{\infty} a_{ij}^2 \Big) \Big(\sum_{j=1}^{n} x_j^2 \Big) = \| A \|_F^2 \| x \|^2$$

Dabei ergibt sich die Ungleichung aufgrund der Cauchy-Schwarz-schen Ungleichung für das Standardskalarprodukt der Vektoren ${}^t A_{i\bullet}$ und x . Es folgt für alle $x \neq 0$;

$$\frac{\| Ax \|}{\| x \|} \leq \| A \|_F \quad \text{bzw.} \quad \| A \|_2 \leq \| A \|_F \qquad \square$$

Zurück zu Beispiel 3.4(b):
Dort ist ${}^t AA = \begin{bmatrix} 1 & 0 \\ s & 1 \end{bmatrix} \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & s \\ s & s^2 + 1 \end{bmatrix}$, $\chi_{{}^t AA} = (1 + t)((s^2 + 1) - t) - s^2 = t^2 - (2 + s^2)t + 1$ und

$$\sigma({}^t AA) = \Big\{ \underbrace{1 + \frac{s^2}{2} - \frac{1}{2}\sqrt{4s^2 + s^4}}_{=:\lambda_1}, \underbrace{1 + \frac{s^2}{2} + \frac{1}{2}\sqrt{4s^2 + s^4}}_{\lambda_2} \Big\} \; .$$

Bemerkung: Der Rayleigh-Quotient und die Matrizennormen spielen eine wichtige Rolle bei der sogenannten **"Lokalisierung von Eigenwerten".** Es gilt nämlich (erste grobe Information)

Satz 3.7: Sei $A \in \mathbb{R}^{n \times n}$ und $\lambda \in \sigma_{\mathbb{C}}(A)$. Dann ist

$$\Big(\sqrt{\lambda \overline{\lambda}} \Big) = |\lambda| \leq \| A \|_2$$

Der kurze Beweis erfordert die Ausweitung der Theorie ins Komplexe und dann gilt der Satz auch für $A \in \mathbb{C}^{n \times n}$.
Mehr Information dazu (z. B. Min Max-Prinzip für Eigenwerte) meist in Büchern über allgemeine lineare Algebra und Matrizenrechnung, Vorlesungen zur Ingenieurmathematik und der Numerik.
Um die Methoden zu erproben, folgt nun ein völlig andersartiger Beweis von Satz 3.2 mit Hilfe von Beobachtung 3.5 und Satz 3.6:
Sei $0 \neq A \in \mathbb{R}^{m \times n}$ und sei v ein Vektor derart, dass $s := \| A \|_2 = \frac{\| Av \|}{\| v \|} = \max\limits_{x \neq 0} \frac{\| Ax \|}{\| x \|} > 0$.

Wir setzen $v_1 = \frac{1}{\| v \|} v$, dann ist $\| Av_1 \| = \| A \|_2 = s$. Sei weiter $u_1 = s^{-1}Av_1$. Dann gilt:

$$Av_1 = s \cdot u_1 \quad \text{und} \quad \| v_1 \| = 1, \ \| u_1 \| = 1$$

Seien $(v_1, \ldots, v_n)$ eine Orthonormalbasis von $\mathbb{R}^{n\times 1}$, $(u_1, \ldots, u_m)$ eine Orthonormalbasis von $\mathbb{R}^{m\times 1}$, und seien $Q = [v_1, \ldots, v_n],\ P = [u_1, \ldots, u_m]$. Dann ist

$$B := {}^{\mathsf{t}}PAQ = \begin{bmatrix} {}^{\mathsf{t}}u_1 \\ \vdots \\ {}^{\mathsf{t}}u_n \end{bmatrix} [Av_1, \ldots, Av_n] = \begin{bmatrix} {}^{\mathsf{t}}u_1 \\ \vdots \\ {}^{\mathsf{t}}u_n \end{bmatrix} [su_1, Av_2, \ldots, Av_n] = \begin{bmatrix} s & {}^{\mathsf{t}}w \\ 0 & \\ \vdots & A_1 \\ 0 & \end{bmatrix} = B\ .$$

Dabei ist ${}^{\mathsf{t}}w = [{}^{\mathsf{t}}u_1 Av_2, \ldots, {}^{\mathsf{t}}u_1 Av_n]$.
Automatisch ist $w = 0$, denn einerseits ist $B \cdot \begin{bmatrix} s \\ w \end{bmatrix} = \begin{bmatrix} s^2 + {}^{\mathsf{t}}ww \\ A_1 w \end{bmatrix}$ und somit

$$\| B \begin{bmatrix} s \\ w \end{bmatrix} \|^2 = (s^2 + {}^{\mathsf{t}}ww)^2 + \| A_1 w \|^2 \geq (s^2 + {}^{\mathsf{t}}ww)^2 = \| \begin{bmatrix} s \\ w \end{bmatrix} \|^4$$

und daher

$$\| B \|_2^2 \geq s^2 + {}^{\mathsf{t}}ww.$$

Andererseits ist aber auch :

$$\begin{aligned} s^2 = \| A \|_2^2 &= \max \sigma({}^{\mathsf{t}}AA) = \max \sigma({}^{\mathsf{t}}(PB{}^{\mathsf{t}}Q)(PB{}^{\mathsf{t}}Q) \\ &= \max \sigma(Q{}^{\mathsf{t}}BB{}^{\mathsf{t}}Q) = \max \sigma({}^{\mathsf{t}}BB) = \| B \|_2^2 \end{aligned}$$

Insgesamt folgt: $w = 0$.

Nun kann mit vollständiger Induktion der Beweis abgeschlossen werden. □

4 Bilinearformen und Quadriken

Motivation: Geometrie, CAD, NW, Verallgemeinerung des Begriffs Skalarprodukt.
Wiederholung: Bilinearformen, symmetrische Bilinearformen, positiv definiertes Skalarprodukt.
Wir konzentrieren uns auf reelle Vektorräume.
Sei im Folgenden stets V ein $\mathbb{R}$-Vektorraum und $s : V \times V \to \mathbb{R}$ eine Bilinearform.

Definition 4.1: Sei $\mathfrak{A} = (v_1, \ldots, v_n)$ eine Basis von V. $M_{\mathfrak{A}}(s) := (s(v_i,\ v_j))_{1\leq i,j\leq n}$ heißt **Gram'sche Matrix von *s*.** (vgl. [F], S. 289 ff.)
Mit Hilfe von $M_{\mathfrak{A}}(s)$ kann die Berechnung von $s(v, w)$ über Koordinaten erfolgen: Sei $v = \Phi_{\mathfrak{A}}(x),\ \ w = \Phi_{\mathfrak{A}}(y)$, dann ist

$$\boxed{s(v, w) = {}^{\mathsf{t}}xM_{\mathfrak{A}}(s)\, y.}$$

Es entstehen folgende **Fragen:**
Wie hängt $M_{\mathfrak{A}}(s)$ von $\mathfrak{A}$ ab ?
Lassen sich Eigenschaften von s an $M_{\mathfrak{A}}(s)$ ablesen?
Wie sehen die "Flächen" $\{v : s(v, v) = const\}$ aus ?

Beobachtung 4.2: s symmetrische Bilinearform $\Leftrightarrow M_{\mathfrak{A}}(s)$ symmetrische Matrix.

Die Symmetrie von s lässt sich demnach bezüglich jeder Basis an $M_{\mathfrak{A}}(s)$ erkennen.

Satz 4.3: Seien $\mathfrak{A} = (v_1, \ldots, v_n)$, $\mathfrak{B} = (w_1, \ldots, w_n)$ Basen von V. Sei $P := M_{\mathfrak{B}}^{\mathfrak{A}}(id_V) = (p_{ik})$ und damit für $i \leq i \leq n :\ v_i = \sum\limits_{k=1}^{n} p_{ki}\, w_k$. Dann ist

$$M_{\mathfrak{A}}(s) = {}^{\mathsf{t}}PM_{\mathfrak{B}}(s)P$$

Beweis: Seien $A := M_{\mathfrak{A}}(s),\ B := M_{\mathfrak{B}}(s),\ A = (a_{ij}),\ B = (b_{ij})$. Nach Definition 4.1 ist dann $a_{ij} = s(v_i,\ v_j),\ b_{ij} = s(w_i, w_j)$ für $1 \leq i,\ j \leq n$ und man berechnet:
$a_{ij} = s\Big(\sum\limits_{k=1}^{n} p_{ki}w_k, v_j\Big) = \sum\limits_{k=1}^{n} p_{ki}s(w_k, v_j) = \sum\limits_{k=1}^{n}\sum\limits_{l=1}^{n} p_{ki}s(w_k, w_l)p_{l_j} = ({}^{\mathsf{t}}PBP)_{ij}$ ☐

Beachte: i. A. ist **nicht** ${}^{\mathsf{t}}P = P^{-1}$. Dies ist nur der Fall, wenn $P \in \mathcal{O}(n)$.

Definition 4.4: Seien $A,\ B \in \mathbb{R}^{n\times n}$. A, B heißen **kongruent ($A \equiv B$)**, wenn mit $P \in GL(n, \mathbb{R})$ gilt:

$$A = {}^{\mathsf{t}}PBP$$

$\equiv$ ist eine Äquivalenzrelation.

Beispiel 4.5: **(a)** Seien $V = \mathbb{R}^{3\times 1},\ A \in \mathbb{R}^{3\times 3}$ und für $x, y \in \mathbb{R}^{3\times 1} : s_A(x, y) := {}^{\mathsf{t}}xAy$. Offensichtlich ist s genau dann symmetrisch, wenn A symmetrisch ist. Dies ergibt sich auch direkt aus Beobachtung 4.2, denn mit der kanonischen Basis $\mathfrak{K}$ ist $M_{\mathfrak{K}}(s_A) = A$. Wann ist s_A ein Skalarprodukt bzw. s positiv definit? Dazu ergeben sich im Verlauf dieses Paragraphen Kriterien. s_A ist die **Standardbilinearform** zur Matrix A bzgl. der kanonischen Basis.

(b) Auf dem reellen Vektorraum $C_1([a, b])$, wird durch die Festlegung $s(f, g) = \int\limits_a^b f(t)g(t)dt$ für $f, g \in C_1([A, B])$ ein Skalarprodukt erklärt (Wdhlg). Sei $V =< 1, f, \ldots, f^{n-1} >_{\mathbb{R}}$, mit $f = id_{[a,b]}$. $\mathfrak{A}= \{1, \ldots, f^{n-1}\}$ ist Basis von V. Wie äußert sich hier die Positivdefinitheit von s in der Matrix $M_{\mathfrak{A}}(s)$?

Wir interessieren uns besonders für die einer Bilinearform zugeordnete **quadratische Form**

$$q_s : V \to \mathbb{R},\ v \mapsto s(v, v).$$

Den Zusammenhang zwischen s und q_s beschreibt die folgende

Beobachtung 4.6: Sei s eine Bilinearform auf V.

(a) Sei $s^* : V \times V \to \mathbb{R}$ die Bilinearform mit $s^*(v, w) = s(w, v)$ für $v, w_i \in V$. Dann gilt

$$s = \frac{1}{2}\underbrace{(s + s^*)}_{s^+} + \frac{1}{2}\underbrace{(s - s^*)}_{s^-}$$

Dabei ist s^+ symmetrisch und s^- schiefsymmetrisch. Letzteres besagt: $s^-(v, w) = -s^-(w, v)$ für alle $v, w \in V$.

(b) $q_s = q_{s^+}$ und für alle $v, w \in V$ gilt: $s^+(v, w) = \frac{1}{2}(q_s(v + w) - q_s(v) - q_s(w))$

Beweis: Nachrechnen (vgl. [F] S. 291) ☐

Wir konzentrieren uns auf symmetrische Bilinearformen, u. A. da wir insbesondere an quadratischen Flächen interessiert sind: "$\{v : q_s(v) = \text{const. }\}$". Dafür genügt es symmetrische Matrizen "modulo" Kongruenz zu betrachten (Satz 2).
Eine **Normalform für symmetrische Matrizen unter Kongruenz** ist leicht zu erreichen.

Satz 4.7: $\forall A \in \mathbb{R}^{n\times n}, \boldsymbol{A}$ **symmetrisch,** $\exists P \in Gl(n, \mathbb{R})$:

$$
\begin{aligned}
{}^{\mathsf{t}}PAP &= \operatorname{diag}\Big(\underbrace{1, \dots, 1}_{p}, \underbrace{-1, \dots, -1}_{q}, \underbrace{0, \dots, 0}_{n-r}\Big) \\
&= \begin{bmatrix} \boxed{E_p} & & \\ & \boxed{-E_q} & \\ & & \boxed{0} \end{bmatrix} = \operatorname{diag}(E_p, -E_q, 0)
\end{aligned}
$$

($p = 0$ und/oder $q = 0$ und/oder $v = 0$ erlaubt). Dabei ist r der Rang von A.

Beweis:

(a) Wir benutzen elementare Umformungen der Art $A \mapsto {}^{\mathsf{t}}PAP$ mit Elementarmatrizen P. Dies genügt, denn $Gl(n, \mathbb{R})$ wird von den Elementarmatrizen erzeugt. Im Einzelnen gilt:

(i) $${}^{\mathsf{t}}P_k^1 A P_k^1 = \begin{bmatrix}
\boldsymbol{a_{kk}} & a_{k2} & \dots & a_{k,k+1} & a_{k1} & a_{k,k+1} & \dots & a_{kn} \\
a_{2k} & & & & a_{21} & & & \\
\vdots & * & & & \vdots & * & & \\
a_{k-1,k} & & & & a_{k-1,1} & & & \\
a_{1k} & a_{12} & \dots & a_{1k-1} & \boldsymbol{a_{11}} & a_{1,k+1} & \dots & a_{1n} \\
a_{k+1,k} & & & & a_{k+1,1} & & & \\
\vdots & * & & & \vdots & * & & \vdots \\
a_{n,k} & & & & a_{n1} & & \dots & a_{nn}
\end{bmatrix}$$

dabei ist ${}^{\mathsf{t}}P_k^1 = P_k^1 = P_1^k$. Alle Einträge in den mit $*$ gekennzeichneten Bereichen werden unverändert von A übernommen.
Die Diagonaleinträge können also vertauscht werden, außerdem können so von 0 verschiedene Einträge in die erste Zeile und erste Spalte gebracht werden.

(ii) ${}^{\mathsf{t}}Q_k^1(\lambda) A Q_k^1(\lambda) = \begin{bmatrix} a_{11} + 2\lambda a_{1k} + \lambda^2 a_{kk} & * \\ * & * \end{bmatrix}$, dabei ist ${}^{\mathsf{t}}Q_k^1(\lambda) = Q_1^k(\lambda)$.
Sollte $a_{11} = 0$ sein und a_{1k} oder $a_{kk} \neq 0$, dann kann so erreicht werden, dass die $(1, 1)$-Position $\neq 0$ wird.

(iii) $${}^{\mathsf{t}}Q_1^k(\mu) A Q_1^k(\mu) = \begin{bmatrix}
a_{11} & \dots & \boldsymbol{a_{1k} + \mu a_{11}} & * \\
\vdots & & \vdots & \vdots \\
\boldsymbol{a_{k1} + \mu a_{11}} & \dots & * & * \\
* & & * & *
\end{bmatrix}$$

Wenn $a_{11} \neq 0$ kann - da ja $a_{k1} = a_{1k}$ - simultan in der ersten Zeile und der ersten Spalte ausgeräumt werden.

(b) Nun lässt sich A mit vollständiger Induktion wie folgt diagonalisieren:
$\boldsymbol{n = 1}$ **:** $A = a$, wenn $a \neq 0$, dann ist ${}^{\mathsf{t}}PAP = \mathrm{sgn(a)}$ mit $P = \frac{1}{\sqrt{|a|}}$.
$\boldsymbol{n \geq 1}$ **:** Induktionsannahme: Für alle symmetrischen $A' \in \mathbb{R}^{n \times n}$ existiert ein $P \in GL(n, \mathbb{R})$, derart, dass ${}^{\mathsf{t}}PA'P$ eine Diagonalmatrix ist.
Sei nun $A \in R^{(n+1)\times(n+1)}$.
Wenn $\boldsymbol{A = 0}$, wähle $P = E_{n+1}$.
Wenn $\boldsymbol{A \neq 0}$, dann sei etwa $a_{ik} = a_{ki} \neq 0$. Falls $i \neq 0$, kann durch die Umformung $A \mapsto {}^{\mathsf{t}}P^1_k A P^1_k$ erreicht werden, dass in der ersten Zeile ein Eintrag $\neq 0$ wird. Sei also o. E. $i = 1$, und gelte $a_{1k} = a_{k1} \neq 0$. Falls dann $k \neq 1$, dann kann durch die Umformung $A \mapsto {}^{\mathsf{t}}Q^1_k(\lambda) A Q^1_k(\lambda)$ mit einem $\lambda \in \mathbb{R}$ derart, dass $a_{11} + 2\lambda a_{1k} + \lambda^2 a^2_{kk} \neq 0$, erreicht werden das der $(1,1)$-Eintrag $\neq 0$ wird (siehe (ii)): Sei daher o. E. auch $k = 1$ und $a_{11} \neq 0$. Dann berechnet man

$$
{}^{\mathsf{t}}Q^{n+1}_1(\mu_{n+1}) \dots {}^{\mathsf{t}}Q^2_1(\mu_2) A Q^2_1(\mu_2) \dots Q^{n+1}_1(\mu_{n+1}) = \left[\begin{array}{c|c} a_{11} & 0 \quad \dots \\ \hline 0 & \\ \vdots & \qquad A' \end{array}\right]
$$

mit $\mu_k = -a_{1k}/a_{11}$ für $2 \leq k \leq n+1$. Die Induktionsannahme für A' führt jetzt zum Ziel.

(c) Sei nun $A = \mathrm{diag}(a_1, \dots, a_n)$. Durch Transformationen des Typs $A \mapsto {}^{\mathsf{t}}P^1_k A P^1_k$ können die Diagonaleinträge sortiert werden, derart, dass

$$
\underbrace{a_1 > 0, \dots, a_p > 0}_{\substack{\text{falls ein } a_k > 0, \\ \text{sonst } p=0}} \;\; \underbrace{a_{p+1} < 0, \dots, a_{p+q} < 0}_{\substack{\text{falls ein } a_k < 0, \\ \text{sonst } q=0}}, \;\; \underbrace{a_{p+q+1} = \dots = a_n = 0}_{\substack{\text{falls ein } a_k = 0, \\ \text{sonst } p+q=n}}
$$

Wenn nun $a_k \neq 0$, dann hat
${}^{\mathsf{t}}S_k\left(\frac{1}{\sqrt{|a_k|}}\right) \mathrm{diag}(a_1, \dots, a_n) S_k\left(\frac{1}{\sqrt{|a_k|}}\right)$ den Eintrag $\frac{a_k}{|a_k|}$statt a_k □

Beispiel 4.8: Sei $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 3}$. Wir berechnen:

$$
A \equiv\!\!\longrightarrow \begin{bmatrix} 1 & 0 & 3 \\ 0 & -3 & -4 \\ 3 & -4 & 1 \end{bmatrix} \equiv\!\!\longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & -4 \\ 0 & -4 & -8 \end{bmatrix} \equiv\!\!\longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -\frac{8}{3} \end{bmatrix} \equiv\!\!\longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \; \square
$$

Beobachtung 4.9: Da die Werte einer Bilinearform 's' auf einen endlichdimensionalen $\mathbb{R}$-Vektorraum bei gegebener Basis $\mathfrak{A}$ mit Hilfe von $M_{\mathfrak{A}}(s)$ berechnet werden können und wegen Satz 4.3 und Satz 4.7, gibt es, wenn s symmetrisch ist, eine Basis $\mathfrak{A}$ von V derart, dass

$$
M_{\mathfrak{A}}(s) = \mathrm{diag}(E_p, -E_q, Q_{n-r}) =: \Lambda
$$

und es ist dann

$$
s(\Phi_{\mathfrak{A}}(x), \Phi_{\mathfrak{A}}(y)) = {}^{\mathsf{t}}x \Lambda y = \sum_{i=1}^{p} x_i y_i - \sum_{i=p+1}^{p+q} x_i y_i = s_\Lambda(x, y)
$$

und $q_s(\Phi_{\mathscr{a}}(x)) = {}^{\mathsf{t}}x\Lambda x = \sum\limits_{i=1}^{p} x_i^2 - \sum\limits_{i=p+1}^{p+q} x_i^2$

Satz 4.10: Sylvesterscher Trägheitssatz. (James Joseph Sylvester, 1814 - 1897, klassische Version von Sylvester 1852, zitiert nach [McD])

Die Diagonalmatrix in Satz 4.7 ist eindeutig durch *A* bestimmt.

Damit ist folgendes gemeint: Wenn

$$\begin{aligned} A &\equiv \operatorname{diag}(E_p, -E_q, 0_{n-r}) =: \Lambda \quad \text{und} \\ A &\equiv \operatorname{diag}(E_{p'}, -E_{q'}, 0_{n-r'}) =: \Lambda' \,, \end{aligned}$$

dann ist $p = p',\ q = q'$ und dann automatisch $r = r'$.

Beweis: Wenn $A \equiv \Lambda$ und $A \equiv \Lambda'$, dann folgt $\Lambda \equiv \Lambda'$. Gelte mit $P \in GL(n, \mathbb{R})$: ${}^{\mathsf{t}}P\Lambda P = \Lambda'$. Dann gilt für alle $x \in \mathbb{R}^{n\times 1}$: $\underbrace{{}^{\mathsf{t}}x{}^{\mathsf{t}}P}_{={}^{\mathsf{t}}y} \Lambda \underbrace{Px}_{=:y} = {}^{\mathsf{t}}x\Lambda' x$ bzw.

$$\sum_{i=1}^{p} y_i^2 - \sum_{i=p+1}^{p+q} y_i^2 = \sum_{i=1}^{p'} x_i^2 - \sum_{i=p'+1}^{p'+q'} x_i^2 \qquad (*)$$

Beachte: $p + q = p' + q' =$ Rang A liegt fest.
Annahme $\boldsymbol{p < p'}$: Es gibt $x \in \mathbb{R}^{n\times 1}$ derart, dass $x_{p'+1} = \ldots = x_n = 0$ (falls $p' < n$) und $x \neq 0$ und $\underbrace{\begin{bmatrix} P_{1\bullet} \\ \vdots \\ P_{p\bullet} \end{bmatrix}}_{\in \mathbb{R}^{p\times n}} x = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} = 0$, denn $p < p'$!

In $(*)$ folgt damit

$$-\sum_{i=p+1}^{p+q} y_i^2 = \sum_{r=1}^{p'} x_i^2 > 0 \ \ , W!$$

Aus "Symmetriegründen" folgt $p = p'$ und dann $q = q'$. □

Definition 4.11: Sei $A \in \mathbb{R}^{n+n}$ symmetrisch. Die nach Satz 4.10 und eindeutig durch A bestimmten Zahlen p, q heißen **Trägheitsindizes** von A bzw. von s_A. Die **Signatur von *A* bzw. von $\boldsymbol{s_A}$** ist $\sigma = p - q$.

Bemerkungen:

(a) Die Paare $(\sigma, \mathsf{Rang}\, A)$ und (p, q) bestimmen sich gegenseitig (bei gegebenen n.) Es ist $\mathsf{Rang}\, A = p + q$ und $\sigma = p - q$.

(b) Häufig werden die Trägheitsindizes auch anders definiert. Unsere Definition folgt dem Buch [G], in [MV] wird z.B. anders vorgegangen.

(c) Zur Bestimmung von p, q genügt es $P \in Gl(n, \mathbb{R})$ zu bestimmen, derart, dass ${}^{\mathsf{t}}PAP$ Diagonalmatrix ist und dann die positiven Einträge abzuzählen, denn nach Normieren der Beträge der Diagonalmatrix und nach Umsortieren entsteht die Form in Satz 4.7.

Als Folgerung eine etwas abstraktere Version des Trägheitssatzes: (vgl. [F], S. 322)

Satz 4.12: Sei $s : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform, V endlichdimensional und sei
$\mathscr{A} = (v_1, \ldots, v_n)$ eine Basis von V, derart, dass

$$\begin{array}{lcll} s(v_i, v_i) & > & 0 & \text{für } 1 \leq i \leq p, \\ s(v_i, v_i) & < & 0 & \text{für } p+1 \leq i \leq p+q, \\ s(v_i, v_i) & = & 0 & \text{für } p+q+1 \leq i \leq n. \end{array}$$

Dabei ist $p = 0, q = 0$ und $p + q = n$ erlaubt.
Dann sind (p, q) die Trägheitsindizes von $M_{\mathscr{A}}(s)$ und diese durch s bereits eindeutig festgelegt.

Beweis: Sätze 4.3, 4.7, 4.10 □

Bemerkung 4.13: Sei in Satz 4.12:
$V^+ =< v_1, \ldots, v_p >_{\mathbb{R}}$, $V^- =< v_{p+1}, \ldots, v_{p+q} >$, $V_0 =< v_{p+q+1}, \ldots, v_n >_{\mathbb{R}}$
dann gilt:

$$V_0 \oplus V^+ \oplus V^-.$$

Mit anderen Worten: es gilt: $V = V_0 + V^+ + V^-$ und $V_0 \cap (V^+ + V^-) = \{0\} = V^+ \cap (V_0 + V^-) = V^- \cap (V_0 + V^+)$.
Außerdem sind die drei Untervektorräume paarweise "orthogonal" bzgl. s und wenn etwa $V = \mathbb{R}^n$ auch bezüglich des Standardskalarprodukts (vgl. [F], S. 322).

Bemerkung 4.14: Wir wissen aus Satz 2.22, dass es zu einer symmetrischen Matrix $A \in \mathbb{R}^{n \times n}$ ein $P \in \mathcal{O}(n)$ gibt, derart, dass ${}^{\mathsf{t}}PAP$ eine Diagonalmatrix ist. Verfügt man über dieses P, so braucht nur noch normiert zu werden, um die Form aus Satz 4.7 zu erhalten. **Aber:** solch ein P ist i.A. schwer zu berechnen.

Ablesen der Positivdefinitheit von s an $M_{\mathscr{A}}(s)$ bezüglich einer beliebigen Basis $\mathscr{A}$.

Beobachtung 4.15: Sei s symmetrische Bilinearform.

(a) s nicht ausgeartet $\Leftrightarrow \forall v \in V \backslash \{0\} \exists w \in V : s(v, w) \neq 0 \Leftrightarrow p + q = n$

(b) [s positiv definit $\Leftrightarrow p = n$] und [s negativ definit $\Leftrightarrow q = n$]

Eine Bilinearform s ist dabei negativ definit, wenn $-s$ positiv definit ist. Wir nennen eine **symmetrische Matrix positiv bzw. negativ definit,** wenn die zugehörige Bilinearform $(x, y) \mapsto {}^{\mathsf{t}}xAy$ positiv bzw. negativ definit ist.
Beweis von Beobachtung 4.15: klar nach Beobachtung 9. □

Beispiel 4.16: **(a)** Seien $n = 2, V = \mathbb{R}^{2 \times 1}$, $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$, $s_A(x, y) = {}^{\mathsf{t}}xAy$ für $x, y \in \mathbb{R}^{n \times 1}$. Dann ist $M_{\mathscr{K}}(s_A) = A$. Hier berechnet man

$$A \equiv \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Es folgt: $p = q = 1$. s_A bzw. A ist weder positiv, noch negativ definit.

(b) Gegeben sei die reelle Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) = \alpha x^2 + \beta xy + \gamma y^2 + \delta x + \varepsilon y +$ const. f ist zweimal stetig differenzierbar. Das folgende hinreichende Kriterium für die Existenz lokaler Extrema ist Spezialfall entsprechender Kriterien aus der Analysis. Wenn an der Stelle $(a,b) \in \mathbb{R}^2$ gilt:

$$f'(a,b) = 0 \text{ und } f''(a,b) \text{ positiv oder negativ definit}$$

dann liegt bei (a,b) ein Extremum von f vor. Man berechnet $f' = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (2\alpha x + \beta y + \delta, \beta x + 2\gamma y + \varepsilon)$ und

$$f'' = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2\alpha & \beta \\ \beta & 2\gamma \end{bmatrix} .$$

Zusätzlich sei vorausgesetzt: $\Delta := 4\alpha\gamma - \beta^2 \neq 0$. Dann hat das lineare Gleichungssystem $f' = 0$ die Lösung

$$(a,b) = \frac{1}{\Delta}(-2\gamma\delta + \varepsilon\beta, 2\alpha\varepsilon - \delta\beta) .$$

Weiter berechnet man unter der Voraussetzung $\alpha \neq 0$:

$$f'' \equiv \begin{bmatrix} 2\alpha & 0 \\ 0 & \frac{\Delta}{2\alpha} \end{bmatrix} \equiv sgn(\alpha) \begin{bmatrix} 1 & 0 \\ 0 & sgn(\Delta) \end{bmatrix} .$$

Mit Hilfe des Kriteriums erkennt man jetzt, dass bei (a,b) ein Extremum von f vorliegt, wenn $\Delta > 0$.

Bemerkung: Das Verfahren aus dem Beweis von Satz 4.7 ist i. A. das effizienteste, um die Trägheitsindizes und insbesondere die Positivdefinitheit zu ermitteln. Es genügt, Diagonalform zu erreichen, da durch die Normierung der Beträge der Diagonaleinträge die Vorzeichen nicht mehr geändert werden. Kennt man das charakteristische Polynom von A, dann führen die Vorzeichenregeln von Descartes zum Ziel, siehe etwa [F], S. 69 ff.

Weitere Methoden/Kriterien zur Entscheidung, ob eine symmetrische Matrix positiv definiert ist.

Orthogonale Basiswechsel $A \to {}^tPAP$ (siehe Bemerkung 4.14)

Abschätzung von $\frac{\| Ax \|}{\| x \|}$**:** Siehe dazu auch Abschnitt 3 (Matrixnormen, Rayleigh-Quotient)
Wenn für alle $x \in \mathbb{R}^{n \times 1}$ und mit einem $\alpha > 0$ gilt

$$\| Ax \| \leq \alpha \| x \|$$

dann ist A, bzw. s_A positiv definit.

Hauptminoren-Kriterium oder Hurwitz-Kriterium: Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch, $A = (a_{ij})$. Dann gilt:

$$A \text{ positiv definit} \iff \text{für } 1 \leq k \leq n : \quad \det\left((a_{ij})_{1 \leq i,j \leq k}\right) > 0.$$

Die dabei auftretenden Determinanten heißen Hauptminoren. Der Beweis wird Gegenstand einer Übungsaufgabe. Für größere Matrixformate ist dieses Kriterium wegen des zu großen Rechenaufwandes praktisch nicht anwendbar.
Für das allgemeine **Hurwitz-Kriterium** zum Ablesen der Signatur an obiger Determinantenfolge sei auf Seite 398 im Teil 2 von [SchSt] verwiesen.

Quadriken und Hauptachsentransformation [3]

Definition 4.17: Ein Polynom(ausdruck) f oder $f(x)$ vom Grad 2 mit reellen Koeffizienten hat die Form $f(x) = \sum\limits_{1\leq i,\, j\leq n} a_{ij}x_ix_j + \sum\limits_{i=1}^{n} b_ix_i + c$ mit $x = {}^{\mathsf{t}}[x_1, \dots, x_n]$.
Mit anderen Worten: $f(x) \in \mathbb{R}[x_1, \dots, x_n]$ ist ein Polynom mit höchstens quadratischen Termen. Dazu gehört immer eine Abbildung $f : \mathbb{R}^{n\times 1} \to \mathbb{R},\ v \mapsto f(v)$.
Die in einem Polynom 2. Grades auftretenden Koeffizienten lassen sich zusammenfassen als

$$A = (a_{ij}) \in \mathbb{R}^{n\times n},\ b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \in \mathbb{R}^{n\times 1}, c \in \mathbb{R}.$$

Die Nullstellenmenge $N_f := \{v \in \mathbb{R}^{n\times 1} : f(v) = 0\}$ heißt Fläche 2. Ordnung oder **Quadrik**. Folgende kompakte Schreibweise wird benutzt

$$\boxed{f(x) = {}^{\mathsf{t}}xAx + {}^{\mathsf{t}}bx + c} \qquad (4)$$

Wenn $A = 0$, dann ist f linear und N_f ein affiner Unterraum von $\mathbb{R}^{n\times 1}$. Ab jetzt sei $A \neq 0$ vorausgesetzt.
Wenn A nicht symmetrisch ist, dann kann $A' := \frac{1}{2}(A + {}^{\mathsf{t}}A)$ gesetzt werden. Das entsprechende veränderte Polynom sei $f' = {}^{\mathsf{t}}xA'x + {}^{\mathsf{t}}bx + c$ und man stellt fest: $f' = f$.
Im Folgenden kann also o. E. vorausgesetzt werden $\boldsymbol{A \neq 0}$ **und symmetrisch.**

Satz 4.18: Normalform für Polynome vom Grad 2
Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch und $\neq 0$. Seien $b \in \mathbb{R}^{n\times 1},\ c \in \mathbb{R}$ und f wie in (4), dann gibt es $P \in \mathcal{O}_1(n),\ v \in \mathbb{R}^{n\times 1}$ und $\alpha, \beta \in \mathbb{R}$ derart, dass mit $y \in \mathbb{R}^{n\times 1}$ gilt:

$$f(Py + v) = \sum_{i=1}^{r} \lambda_i y_i^2 + \alpha \qquad r \leq n$$

oder

$$f(Py + v) = \sum_{i=1}^{r} \lambda_i y_i^2 + \beta y_n \qquad r < n$$

Dabei ist $r = \mathsf{Rang}\, A,\ \beta > 0$, und $\lambda_1, \dots, \lambda_r$ sind die von 0 verschiedenen, u. U. mehrfachen, Eigenwerte von A.
Beweis: Für einen vollständigen Beweis sei auf das Buch [MV] von Mackens und Voß verwiesen. Hier wird nur der relativ kurz und einfach beweisbare Fall behandelt, wo zusätzlich vorausgesetzt ist, dass $\boldsymbol{b \in \mathsf{Bild}\, L_A}$.
Nach Satz 2.22 gibt es eine Orthonormalbasis $(v_1, \dots, v_n)$ von Eigenvektoren von A. Mit

---

[3] orientiert an [MV], Seiten 301 ff.

$P = [v_1, \dots, v_n] \in \mathcal{O}_1(n)(!)$ gilt dann ${}^{\mathsf{t}}PAP = \operatorname{diag}(\lambda_1, \dots, \lambda_n) =: \Lambda$.
Nun ist

$$\begin{aligned} f(Py+v) &= ({}^{\mathsf{t}}y{}^{\mathsf{t}}P + {}^{\mathsf{t}}v)A(Py+v) + {}^{\mathsf{t}}b(Py+v) + c \\ &= {}^{\mathsf{t}}y\Lambda y + \underline{{}^{\mathsf{t}}y{}^{\mathsf{t}}PAv} + \underline{{}^{\mathsf{t}}vAPy} + {}^{\mathsf{t}}vAv + \underline{{}^{\mathsf{t}}bPy} + {}^{\mathsf{t}}bv + c \\ &= ty\Lambda y + \underbrace{{}^{\mathsf{t}}(2Av+b)Py}_{=0\ ??} + \underbrace{{}^{\mathsf{t}}vAv + {}^{\mathsf{t}}bv + c}_{=:\lambda} \end{aligned}$$

wenn $b \in$ Bild L_A, also $\mathsf{L\ddot{o}s}\,(A, -\frac{1}{2}b) \neq \varnothing$, dann kann v so gewählt werden, dass $2Av + b = 0$. ☐

Definition 4.19: Die Abbildung $F : \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$ mit $F(y) = Py + v$ aus Satz 4.18 heißt **Hauptachsentransformation.** I. A. ist F eine sogenannte **Bewegung**, dies ist eine Hintereinanderausführung einer Translation und einer Isometrie. Die Bewegungen sind genau (zu beweisender Satz) die abstandstreuen Abbildungen. F heißt **abstandstreu,** wenn für alle $v, w \in V$ gilt: $\parallel v - w \parallel = \parallel F(v) - F(w) \parallel$. Die Hauptachsentransformation fördert die geometrischen "optimalen" (wenig Koeffizienten $\neq 0$) Koordinatenachsen zu Tage.

Beispiel 4.20: nach [MV], S. 299.
Sei $f = 8x_1^2 - 12x_1x_2 + 17x_2^2 - 20x_1 - 10x_2 + 5$.
Hier ist $A = \begin{bmatrix} 8 & -6 \\ -6 & 17 \end{bmatrix}$, $b = \begin{bmatrix} -20 \\ -10 \end{bmatrix}$ und $c = 5$.
Man berechnet: $\chi_A = t^2 - 25t + 100 = (t-5)(t-20)$ und dann $P = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$.
Probe: ${}^{\mathsf{t}}PAP = \begin{bmatrix} 5 & 0 \\ 0 & 20 \end{bmatrix}$. Hier ist $\mathsf{det}\, A \neq o$ und $\mathsf{L\ddot{o}s}\,(A, b) = \begin{bmatrix} -4 \\ -2 \end{bmatrix}$, bzw. $b \in \mathsf{Bild}\, L_A$.
Setze nun wie im Beweis $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, dann erhält man:

$$f(Py+v) = 5y_1^2 + 20y_2^2 - 20 \ .$$

Die **Klassifizierung der Quadriken** setzt nach der Hauptachsentransformation ein. Für die Fälle $n = 2, n = 3$ ist sie in dem Buch von Mackens und Voß vollständig durchgeführt mit Abbildungen typischer Fälle. Siehe dort auf den Seiten 304 - 306. Z. B. mit Hilfe von Maple kann man sich Quadriken für $n = 3$ mit der Anweisung

$$\mathsf{plots[implicitplot]}\big(f(x,y,z), x = \alpha \dots \beta, y = \gamma \dots \delta,\ z = \kappa \dots \lambda,\ \text{Optionen}\big).$$

Die Untersuchung von Quadriken ist Teil der sogenannten "analytischen Geometrie". Zur Hauptachsentransformation im Rahmen der analytischen Geometrie siehe: Fischer, Analytische Geometrie, Vieweg. Dort wir u.A. eine schöne historische Einführung gegeben und ein tieferes Verständnis der Quadriken im Rahmen der projektiven Geometrie ermöglicht.
Quadriken sind "trotz alledem" recht einfach gebaute Flächen. Komplexere algebraische definierbare Flächen werden in der **algebraischen Geometrie** untersucht. Ein paar Beispiele zum Anschauen finden Sie unter

www.mathematik.uni-kl.de/~keilen/Flaechen.html

Faszinierend finde ich z. B. die sogenannte Barth-Dezik, die Nullstellenmenge eines Polynoms in 3 Variablen vom Grad 10.

5 Ein paar Ergänzungen für unitäre Vektorräume

Bisher konzentrierten wir uns auf reelle Vektorräume. Die naheliegende Ausdehnung auf komplexe Vektorräume mit Skalarprodukt führt recht direkt zu einer ganzen Reihe analoger Ergebnisse, die im Folgenden wiedergegeben sind. Nur einige werden in der Vorlesung und in Übungsaufgaben behandelt.
Zu erinnern bzw. wiederholen (siehe z.B. [F],[V]) sind die Begriffe **Hermitesche Form und unitärer Raum**, und die darauf aufbauenden Verallgemeinerungen von $\| \;\ \|$, $(\!(\, , \,)\!)$, Orthogonalität, Orthonormalbasis, usf.
Entlang der Abschnitte 1 und 2 lässt nun sich fast alles unmittelbar auf endlichdimensionale unitäre Räume ausdehnen.

Definition 5.1: **(a)** Die **konjugiert transponierte Matrix** ${}^{t}\overline{A}$ zu einer Matrix $A \in \mathbb{C}^{m\times n}$ bezeichnen wir mit $\boldsymbol{{}^{\star}\!A}$.

(b) Eine Matrix $A \in \mathbb{C}^{n\times n}$ heißt **Hermitesch**, wenn ${}^{\star}\!A = A$.

(c) Eine Matrix $A \in \mathbb{C}^{n\times n}$ heißt **unitär**, wenn gilt: $A^{-1} = {}^{\star}\!A$.

Symmetrische reelle Matrizen sind Hermitesch und orthogonale Matrizen sind unitär.
Die Diagonaleinträge einer Hermiteschen Matrix sind stets reell.
Die metrischen Abbildungen oder Isometrien endlichdimensionaler unitärer Vektorräume führen jetzt zu den unitären Matrizen. Die unitären $(n \times n)$-Matrizen übernehmen die Rolle der orthogonalen $(n \times n)$-Matrizen.

Beobachtung 5.2: **(a)** Sei $A \in \mathbb{C}^{n\times n}$ eine Hermitesche Matrix, dann ist $\sigma(A) \subseteq \mathbb{R}$.

(b) Sei $A \in \mathbb{C}^{n\times n}$ eine unitäre Matrix. Dann ist $|\mathsf{det}\, A| = 1$. Die Determinante bildet $\mathcal{U}(n)$ ab auf den komplexen Einheitskreis. Außerdem liegen alle Eigenwerte von A auf dem komplexen Einheitskreis.

(c) Die unitären $(n \times n)$-Matrizen bilden die Untergruppe $\mathcal{U}(n)$ von $Gl(n, \mathbb{C})$. Sie enthält den Normalteiler $\mathcal{U}_1(n) = \{U \in \mathcal{U}(n) :\ \mathsf{det}\, U = 1\}$. Desweiteren enthält sie auch den Normalteiler $\mathcal{U}_{\pm 1}(n) = \{U \in \mathcal{U}(n) :\ \mathsf{det}\, U \in \{1, -1\}\}$.

(d) $\mathcal{U}_{\pm 1}(2) = \{ \begin{bmatrix} a & b \\ \varepsilon_1 \overline{b} & \varepsilon_2 \overline{a} \end{bmatrix} : a, b \in \mathbb{C}, a\overline{a} + b\overline{b} = 1\}$ mit $\varepsilon_1, \varepsilon_2 \in \{1, -1\}, \varepsilon_1\varepsilon_2 = -1$
und $\mathcal{U}_1(2) = \{ \begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix} :\ a, b \in \mathbb{C}, a\overline{a} + b\overline{b} = 1\}\}$

Ü: Wie lautet die analoge Aussage für $\mathcal{U}(n)$?

Definition 5.3: Elementare unitäre Matrizen werden ganz analog zu Definition 2.15 erklärt:

$$P = \begin{bmatrix} 1 &&&&&&&& 0 \\ & \ddots &&&&&& \iddots & \\ && a && 0 && b && \\ && \vdots & \ddots && \iddots & \vdots && \\ && 0 && 1 && 0 && \\ && \vdots & \iddots && \ddots & \vdots && \\ && \varepsilon_2 \cdot \overline{b} & \cdots & 0 & \cdots & \varepsilon_1 \cdot \overline{a} && \\ & \iddots &&&&&& \ddots & \\ 0 &&&&&&&& 1 \end{bmatrix}, \varepsilon_1, \varepsilon_2 \in \{1, -1\}, \varepsilon_1\varepsilon_2 = -1,\ a\overline{a}+b\overline{b} = 1.$$

Offensichtlich ist $\mathsf{det}\, P \in \{1, -1\}$.

Als **Analoga zu den Sätzen 2.13, 2.17, 2.19, 3.2** ergeben sich

Satz 5.4: Sei V ein unitärer Vektorraum, $\mathsf{dim}_V = n,\ F : V \to V$ linear, $\mathcal{A}$ Orthonormalbasis und $A := M^{\mathcal{A}}_{\mathcal{A}}(F)$. Dann gilt

$$F \text{ Isometrie } \Leftrightarrow L_A \text{ Isometrie } \Leftrightarrow A \text{ unitär}$$

Satz 5.5: Unitäre Zeilenstufenform. Sei $n \geq 2$. Zu $A \in \mathbb{R}^{n\times m}$ gibt es $r \in \mathbb{N}_+$ und elementare unitäre Matrizen $Q_1, \ldots, Q_r$ so dass $Q_r \cdots Q_1 A$ in Zeilenstufenform ist mit positivem Eckeinträgen bis zur mindestens vorletzten Zeile $\neq 0$ und mit $1 \leq r \leq \frac{n(n-1)}{2}$. Jedes einzelne $Q_i,\ 1 \leq i \leq r$ kann entweder mit $\mathsf{det}\, Q_i = 1$ oder mit $\mathsf{det}\, Q_i = -1$ gewählt werden. Insbesondere sind die beiden folgenden Fälle realisierbar:

$$\mathsf{det}\, Q_i = 1 \text{ für } 1 \leq i \leq r \text{ oder } \mathsf{det}\, Q_i = -1 \text{ für } 1 \leq i \leq r.$$

Satz 5.6: **(a)** Sei $A \in \mathcal{U}_1(n)$. Dann ist $A = P_1 \cdots P_r$ mit elementaren unitären Matrizen $P_1, \ldots, P_r$, mit $\mathsf{det}\, P_i = 1$ für $1 \leq i \leq r$ und mit $1 \leq r \leq \frac{n(n-1)}{2}$.

(b) Sei $A \in \mathcal{U}_{\pm 1}(n) \backslash \mathcal{U}_1(n)$: Dann ist: $A = Q_1 \cdots Q_r$ mit elementaren unitären Matrizen $Q_1, \ldots, Q_r$, mit $\mathsf{det}\, Q_i = -1$ und mit $1 \leq r \leq \frac{n(n-1)}{2} + 1$ und r ungerade.
Es ist aber auch:
$A = P_1 \cdots P_r \cdot Q$ mit elementaren unitären Matrizen $P_1, \ldots, P_r, Q$, mit $\mathsf{det}\, P_1 = \ldots = \mathsf{det}\, P_r = 1, \mathsf{det}\, Q = -1$ und mit $1 \leq r \leq \frac{n(n-1)}{2}$.

Definition 5.7: $A \in \mathbb{C}^{n\times n}$ heißt normal, wenn $A\,{}^\star\!A = {}^\star\!A\,A$.

Hermitesche und unitäre Matrizen sind normal und natürlich auch alle reellen normalen Matrizen.

Satz 5.8: Unitäre Diagonalisierung. Sei $A \in \mathbb{C}^{n\times n}$ normal, dann gibt es $P \in \mathcal{U}(n)$ derart, dass

$$P^{-1}AP = \mathrm{diag}(\lambda_1, \ldots, \lambda_s)$$

mit $\lambda_1, \ldots, \lambda_s \in \sigma_{\mathbb{C}}(A)$. Ist A unitär, dann gilt (s.o.) $|\lambda_i| = 1$ für $1 \leq i \leq s$, ist A Hermitesch, dann gilt (s.o.) $\lambda_i \in \mathbb{R}$ für $1 \leq i \leq s$. Insbesondere gibt es eine Orthonormalbasis aus Eigenvektoren.

Die vorstehenden Aussagen beinhalten u.A. den **Spektralsatz** für Hermitesche Matrizen.
Man erhält auch eine Singulärwertzerlegung für komplexe Matrizen:

Satz 5.9: Singulärwertzerlegung. Sei $A \in \mathbb{C}^{m\times n} \backslash \{0\}$.
Dann gibt es $P \in \mathcal{U}(m),\ Q \in \mathcal{U}(n)$ und $s_1, \ldots, s_r \in \mathbb{R}$, derart, dass

$$0 < s_1 \leq \cdots \leq s_r$$

und

$$PAQ = \begin{bmatrix} s_r & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & s_1 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \qquad (m \times n\text{-Diagonalmatrix})$$

$s_1, \dots, s_r$ sind durch A **eindeutig** festgelegt und heißen **Singulärwerte von A.** Dabei gilt

$$\{s_1, \dots, s_r\} = \{\sqrt{\mu} : \ \mu \in \sigma({}^{\star}A\,A) \setminus \{0\}\} \ .$$

Die Zerlegung $A = {}^{\mathsf{t}}P\mathrm{diag}(s_r, \dots s_1,\ 0, \dots){}^{\mathsf{t}}Q$ heißt **Singulärwertzerlegung von A**.

Auch die Sätz 4.7 und 4.10 haben unmittelbare Verallgemeinerungen.

Literatur

[B] Theodor Bröcker: *Lineare Algebra und analytische Geometrie*, Birkhäuser, 2004.

[F] Gerd Fischer: *Lineare Algebra*, vieweg, 13. Auflage, 2002.

[G] Felix Gantmacher: *Matrizenrechnung*, Springer, 1986.

[HP] Diederich Hinrichsen und Anthony Pritchard: *Mathematical System Theory I, Modelling, State Space Analysis, Stability and Robustness*, Springer, 2005.

[K] Max Köcher: *Lineare Algebra und analytische Geometrie*, Springer, 1983.

[McD] Cyrus MacDuffee: *The Theory of Matrices.* Chelsea, 1946.

[MV] Wolfgang Mackens und Heinrich Voss: *Mathematik I, Lehrbuch für Studierende der Ingenieurwissenschaften* , Heco-Verlag, 1993.

[M] Herbert Möller: *Algorithmische Mathematik*, Vieweg, 1997.

[SchSt] Günter Scheja, Uwe Storch: *Lehrbuch der Algebra*, Teubner, 1980 und 1988.

[Se] Denis Serre: *Matrices*, Springer, 2002.

[St] G.W. Stewart: *Early history of SVD*, University Maryland (Report) (z.Zt. unter http://www.math.ucdavis.edu/∼saito/courses/167/stewart-svd.pdf erhältlich) und in SIAM Review, 1993.

[V] Udo Vetter: *Lineare Algebra*, Skript zur Vorlesung im Wintersemester 2004/2005, Institut für Mathematik, Oldenburg.