

Teil II

Lineare Algebra und Analytische Geometrie

Vorlesung

Sommersemester 2000

Volker Mehrmann

Übungen: Matthias Pester
Andreas Steinbrecher

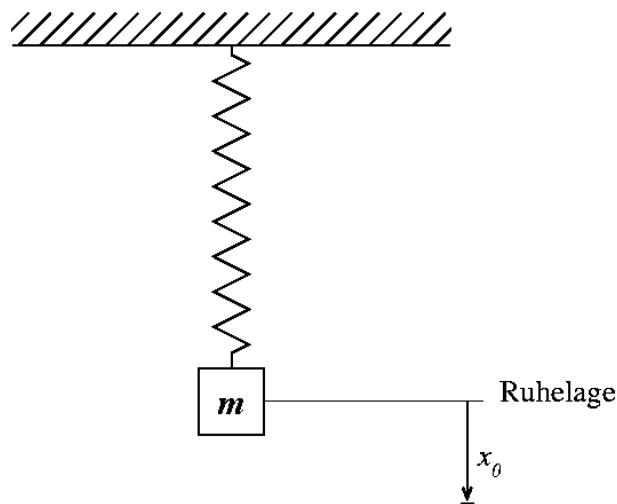
Kapitel 1

Lineare Differentialgleichungen

Im Kapitel 16 von Teil I haben wir die Jordan'sche Normalform einer Matrix kennengelernt. Wir wollen diese Ergebnisse nun anwenden, um damit die Lösungstheorie für lineare Differentialgleichungen zu bekommen.

Beispiel 1.1 *Schwingungsgleichung*

Ein Gewichtstück mit Masse m sei an einer Schraubenfeder mit Federkonstante μ aufgehängt. Zu Anfang wird das Gewicht um die Strecke x_0 ausgelenkt.



Außerdem erhalte es eine Anfangsgeschwindigkeit v_0 . Die Bewegungsgleichung lautet (mit dem Hook'schen Gesetz):

$$x'' = \frac{d^2x}{dt^2} = -\frac{\mu}{m}x$$

mit Anfangsbedingungen

$$x(0) = x_0, \quad x'(0) = v_0.$$

Wir führen als eine neue Variable die Geschwindigkeit ein:

$$v = x', \quad (v' = x'')$$

und erhalten das System von gewöhnlichen Differentialgleichungen 1. Ordnung

$$\begin{bmatrix} x \\ v \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ \frac{-\mu}{m} & 0 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix}. \quad (1.2)$$

Anfangsbedingungen : $\begin{bmatrix} x \\ v \end{bmatrix} (0) = \begin{bmatrix} x_0 \\ v_0 \end{bmatrix}.$

Wir möchten eine allgemeine Theorie zur Lösung von linearen Systemen gewöhnlicher Differentialgleichungen

$$y' = Ay + f, \quad y(0) = y_0, \quad A \in K^{n,n}, \quad y_0 \in K^n, \quad (1.3)$$

wobei

$$y, f : [0, a] \rightarrow K^n.$$

Satz 1.4 *Die Lösungen des homogenen Differentialgleichungssystems*

$$y' = Ay, \quad y : [0, a] \rightarrow \mathbb{C}^n, \quad A \in \mathbb{C}^{n,n}, \quad (1.5)$$

bilden einen Untervektorraum \mathcal{L}_A des (unendlich dimensionalen) Vektorraums der über dem Intervall $[0, a]$ stetig differenzierbaren Funktionen

$$\mathcal{C}^1([0, a], \mathbb{C}^n).$$

Beweis: Seien y_1, \dots, y_k Lösungen von (1.5), so folgt für alle $\alpha_1, \dots, \alpha_k \in \mathbb{C}$,

$$\left(\sum_{i=1}^k \alpha_i y_i \right)' = \sum_{i=1}^k \alpha_i y_i' = \sum_{i=1}^k \alpha_i A y_i = A \left(\sum_{i=1}^k \alpha_i y_i \right).$$

Also ist jede Linearkombination von Lösungen eine Lösung und die anderen Gesetze lassen sich leicht überprüfen, insbesondere ist die Funktion $y = 0$ das Nullelement. \square

Satz 1.6 *Jede Lösung des inhomogenen Systems*

$$y' = Ay + f \quad (1.7)$$

ist von der Form

$$y = y_1 + y_2, \quad (1.8)$$

wobei y_1 eine spezielle Lösung von (1.7) ist und y_2 eine Lösung des homogenen Systems (1.5).

Beweis: Seien y, y_1 Lösungen von (1.7), so gilt dass

$$\begin{aligned}(y - y_1)' &= Ay - Ay_1 + f - f \\ &= A(y - y_1),\end{aligned}$$

also ist $y - y_1$ Lösung des homogenen Systems. \square

Wie erhalten wir nun die Lösungen von (1.3)?

Da $A \in \mathbb{C}^{n,n}$, so gibt es nach Satz 16.1 bzw. Folgerung 16.6 von Teil I (S. 123, 127) eine nichtsinguläre Matrix T , so dass TAT^{-1} in Jordan'scher Normalform ist. Es gilt

$$Ty' = (Ty)' = TAT^{-1}Ty + Tf, \quad (1.9)$$

$$Ty(0) = (Ty)(0) = Ty_0 =: z_0. \quad (1.10)$$

Setze $Ty =: z$, $Tf = g$, so folgt

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{bmatrix}' = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{bmatrix} + \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_k \end{bmatrix}, \quad (1.11)$$

wobei $J_i = \lambda_i I + N_i$, $N_i = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$ und λ_i Eigenwert von A .

Wir erhalten lauter kleine Systeme

$$z_i' = J_i z_i + g_i, \quad i = 1, \dots, k. \quad (1.12)$$

Schauen wir uns so ein System im Detail an.

Sei $z_i = \begin{bmatrix} z_{i,1} \\ \vdots \\ z_{i,l_i} \end{bmatrix}$, $g_i = \begin{bmatrix} g_{i,1} \\ \vdots \\ g_{i,l_i} \end{bmatrix}$, so gilt

$$\begin{bmatrix} z_{i,1}'(x) \\ \vdots \\ z_{i,l_i}'(x) \end{bmatrix} = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} \begin{bmatrix} z_{i,1}(x) \\ \vdots \\ z_{i,l_i}(x) \end{bmatrix} + \begin{bmatrix} g_{i,1}(x) \\ \vdots \\ g_{i,l_i}(x) \end{bmatrix}, \quad (1.13)$$

$$\begin{bmatrix} z_{i,1} \\ \vdots \\ z_{i,l_i} \end{bmatrix} (0) = \begin{bmatrix} z_{i,1}^0 \\ \vdots \\ z_{i,l_i}^0 \end{bmatrix}.$$

Dieses System können wir durch Rückwärtseinsetzen lösen, sofern wir eine Gleichung der Form

$$w'(x) = \lambda w(x) + \gamma(x), \quad w(0) = w_0 \quad (1.14)$$

lösen können.

Multipliziere beide Seiten von (1.14) mit $e^{-\lambda x} \neq 0$, so gilt

$$e^{-\lambda x} w'(x) = \lambda e^{-\lambda x} w(x) + e^{-\lambda x} \gamma(x), \quad (1.15)$$

und da

$$\frac{d}{dx} (e^{-\lambda x} w(x)) = \frac{d}{dx} (e^{-\lambda x}) w(x) + e^{-\lambda x} w'(x) = -\lambda e^{-\lambda x} w(x) + e^{-\lambda x} w'(x),$$

so folgt äquivalent

$$\begin{aligned} e^{-\lambda x} (w' - \lambda w) &= \frac{d}{dx} (e^{-\lambda x} w) = e^{-\lambda x} \gamma \\ \iff e^{-\lambda x} w &= \int_0^x (e^{-\lambda t} \gamma) dt + c \\ \iff w &= e^{\lambda x} \int_0^x (e^{-\lambda t} \gamma) dt + e^{\lambda x} c \end{aligned}$$

$$w(0) = w_0 \implies w_0 = e^0 \left(\int_0^0 e^{-\lambda t} \gamma dt \right) + ce^0 \implies \boxed{c = w_0}.$$

Satz 1.16 Die eindeutige Lösung von (1.14) ist

$$w = e^{\lambda x} \int_0^x (e^{-\lambda t} \gamma) dt + e^{\lambda x} w_0.$$

Beweis: Die Eindeutigkeit und Lösbarkeit folgt aus der obigen Konstruktion. \square

Wir erhalten also rekursiv aus (1.13)

$$z_{i,l_i} = e^{\lambda_i x} \int_0^x (e^{-\lambda_i t} g_{i,l_i}(t)) dt + z_{i,l_i}^0 e^{\lambda_i x} \quad (1.17)$$

und

$$\begin{aligned} z_{i,j} &= e^{\lambda_i x} \int_0^x e^{-\lambda_i t} (g_{i,j}(t) + z_{i,j+1}(t)) dt + z_{i,j}^0 e^{\lambda_i x}, \\ j &= l_i - 1, \dots, 1. \end{aligned} \quad (1.18)$$

Korollar 1.19 Das Differentialgleichungssystem (1.3) hat eine eindeutige Lösung.

Beweis: System (1.3) hat eine eindeutige Lösung \iff (1.9), (1.10) hat eindeutige Lösung \iff alle Systeme (1.13), $i = 1, \dots, k$, haben eindeutige Lösung. \square

Wir können also die Existenz und Eindeutigkeit nachweisen und auch die Lösung mit Hilfe der Jordan'schen Normalform von A bestimmen.

Um die Lösungsdarstellung zu vereinfachen, führen wir jetzt für $A \in \mathbb{C}^{n,n}$ die Exponentialfunktion für Matrizen ein.

Definition 1.20 Die Exponentialfunktion von $A \in \mathbb{C}^{n,n}$ ist definiert durch

$$e^A = \sum_{i=0}^{\infty} \frac{1}{i!} A^i = \lim_{k \rightarrow \infty} \sum_{i=0}^k \frac{1}{i!} A^i, \quad (e^0 = I).$$

Dass diese Definition Sinn macht, werden wir später sehen. Es muss natürlich gezeigt werden, dass der Grenzwert existiert.

Lemma 1.21 Es gilt für $A, B \in \mathbb{C}^{n,n}$:

$$(a) \quad \frac{d}{dx}(e^{Ax}) = Ae^{Ax}.$$

$$(b) \quad (e^A)^{-1} = e^{-A}.$$

$$(c) \quad \text{Falls } AB = BA, \text{ so ist } e^{A+B} = e^A \cdot e^B.$$

Beweis: Mit formaler Potenzreihe. \square

Satz 1.22 Die Lösung des Systems von linearen Differentialgleichungen

$$y' = Ay + f, \quad y(0) = y^0$$

ist gegeben durch

$$y = e^{Ax} \int_0^x e^{-At} f dt + e^{Ax} y^0. \tag{1.23}$$

Beweis: Da wir die Eindeutigkeit schon nachgewiesen haben, brauchen wir nur zu zeigen, dass (1.23) eine Lösung darstellt.

$$\begin{aligned} y' &= \frac{d}{dx} \left(e^{Ax} \int_0^x e^{-At} f dt + e^{Ax} y^0 \right) \\ &= \left(\frac{d}{dx} e^{Ax} \right) \int_0^x e^{-At} f dt + e^{Ax} \left(\frac{d}{dx} \int_0^x e^{-At} f dt \right) + \left(\frac{d}{dx} e^{Ax} \right) y^0 \\ &= A \underbrace{\left[e^{Ax} \int_0^x e^{-At} f dt + e^{Ax} y^0 \right]}_y + f. \end{aligned}$$

Weiter gilt

$$\begin{aligned} y(0) &= e^0 \int_0^0 e^{-At} f dt + e^0 y^0 \\ &= y^0. \end{aligned}$$

Das heißt, die Anfangsbedingung ist auch erfüllt. \square

Satz 1.24 Sei $A \in \mathbb{C}^{n,n}$, so hat die Lösungsmenge des homogenen Systems $y' = Ay$ die Dimension n .

Beweis: Die gleiche Konstruktion wie bei der Lösung von $w' = \lambda w + f$ ergibt

$$y = \underbrace{e^{Ax} \int_0^x e^{-At} 0 dt}_{0} + e^{Ax} c,$$

wobei c ein beliebiger Vektor aus \mathbb{C}^n ist. \square

Beispiel 1.25 Betrachte das System aus (1.2):

$$\begin{bmatrix} x \\ v \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ -\frac{\mu}{m} & 0 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix}, \quad \begin{bmatrix} x \\ v \end{bmatrix}(0) = \begin{bmatrix} x_0 \\ v_0 \end{bmatrix}.$$

Lösung:

$$\begin{bmatrix} x \\ v \end{bmatrix} = e^{\begin{bmatrix} 0 & 1 \\ -\frac{\mu}{m} & 0 \end{bmatrix} t} \begin{bmatrix} x_0 \\ v_0 \end{bmatrix}.$$

Da $\begin{bmatrix} 0 & 1 \\ -\frac{\mu}{m} & 0 \end{bmatrix}$ die Eigenwerte $\lambda_{1,2} = \pm i\sqrt{\frac{\mu}{m}}$ hat und die Eigenvektoren dazu sich aus

$$\begin{bmatrix} \mp i\sqrt{\frac{\mu}{m}} & 1 \\ -\frac{\mu}{m} & \mp i\sqrt{\frac{\mu}{m}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \quad \text{ergeben,}$$

so ist $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -i\sqrt{\frac{\mu}{m}} \end{bmatrix}$ Eigenvektor zu $-i\sqrt{\frac{\mu}{m}}$

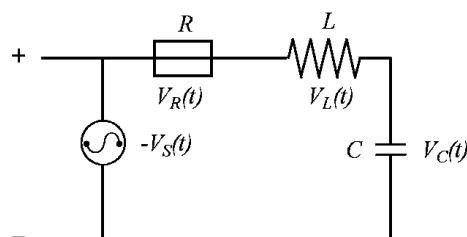
und $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 \\ +i\sqrt{\frac{\mu}{m}} \end{bmatrix}$ Eigenvektor zu $+i\sqrt{\frac{\mu}{m}}$.

Kapitel 2

Lineare differentiell–algebraische Gleichungen

Heutzutage spielen bei der Modellierung von technischen Systemen neben Differentialgleichungen, sogenannte differentiell–algebraische Gleichungen eine besondere Rolle.

Beispiel 2.1 Betrachte den einfachen Schaltkreis



V_S ist die Spannung, R Widerstand, L Induktivität und C Kapazität. Die Spannungsabfälle an Widerstand, Spule und Kondensator sind V_R, V_L, V_C . $I(t)$ ist die Stromstärke. Die Anwendung der Kirchhoff'schen Gesetze ergibt das System

$$\begin{bmatrix} L & 0 & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{I}(t) \\ \dot{V}_L(t) \\ \dot{V}_C(t) \\ \dot{V}_R(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -R & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} I(t) \\ V_L(t) \\ V_C(t) \\ V_R(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -V_S(t) \end{bmatrix}.$$

Dies ist ein lineares System von differentiellen und algebraischen Gleichungen.

Ein allgemeines System dieser Art hat die Form

$$E\dot{x} = Ax + f, \quad E, A \in K^{n,m}, \quad x \in \mathcal{C}^1([t_0, t_1], K^m), \quad f \in \mathcal{C}([t_0, t_1], K^n). \quad (2.2)$$

Wie können wir nun so ein System lösen? Ähnlich wie im Fall der linearen Differentialgleichungen können wir das System von links mit nichtsingulären Matrizen $P \in K^{n,n}$ multiplizieren und außerdem $x = Qy$ setzen, wobei $Q \in K^{m,m}$ nichtsingulär.

Definition 2.3 Zwei Paare von Matrizen $(E_1, A_1), (E_2, A_2)$, $E_i, A_i \in K^{n,m}$, $i = 1, 2$, heißen äquivalent, wenn es nichtsinguläre Matrizen $P \in K^{n,n}, Q \in K^{m,m}$ gibt, so dass

$$(E_1, A_1) = (PE_2Q, PA_2Q) \quad (2.4)$$

Beispiel 2.5 $m = n, E_1 = E_2 = I_n, P = Q^{-1}$ ergibt die Ähnlichkeit von A_1 und A_2 . Wir erwarten also eine Verallgemeinerung der Jordan'schen Normalform.

Wir wollen hier nur einen Spezialfall betrachten, nämlich sogenannte reguläre Paare.

Definition 2.6 Ein komplexes Matrizenpaar (E, A) , $E, A \in \mathbb{C}^{n,m}$, heißt regulär, wenn $n = m$ und ein $\lambda \in \mathbb{C}$ existiert, so dass

$$\det(\lambda E - A) \neq 0.$$

Falls (E, A) regulär, so heißen die Nullstellen von $\det(\lambda E - A)$ verallgemeinerte Eigenwerte von (E, A) .

Beispiel 2.7 (I, A) ist immer regulär, denn für alle λ , die nicht Eigenwert von A sind, gilt $\det(\lambda I - A) \neq 0$. Verallgemeinerte Eigenwerte von (I, A) sind die Eigenwerte von A .

Beispiel 2.8 Für den Schaltkreis aus Beispiel 2.1 gilt

$$\begin{aligned} & \det \left(\lambda \begin{bmatrix} L & 0 & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -R & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \right) \\ &= \det \begin{bmatrix} \lambda L & -1 & 0 & 0 \\ -1 & 0 & \lambda C & 0 \\ R & 0 & 0 & -1 \\ 0 & -1 & -1 & -1 \end{bmatrix} = \lambda L \det \begin{bmatrix} 0 & \lambda C & 0 \\ 0 & 0 & -1 \\ -1 & -1 & -1 \end{bmatrix} + \det \begin{bmatrix} -1 & \lambda C & 0 \\ R & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix} \\ &= \lambda^2 LC + (-1) \det \begin{bmatrix} 0 & -1 \\ -1 & -1 \end{bmatrix} - (\lambda C) \det \begin{bmatrix} R & -1 \\ 0 & -1 \end{bmatrix} = \lambda^2 LC + \lambda RC + 1. \end{aligned}$$

Es gibt also nur zwei verallgemeinerte Eigenwerte $\lambda_{1,2} = -\frac{1}{2L} \left(R \pm \sqrt{R^2 - \frac{4L}{C}} \right)$, obwohl die Matrizen 4×4 sind.

Beispiel 2.9

$$E = \begin{bmatrix} 0 & 1 \\ & 0 & 1 \\ & & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & & \\ & 0 & \\ & & 1 \end{bmatrix}$$

so folgt

$$\det(\lambda E - A) = \begin{bmatrix} -1 & \lambda & 0 \\ 0 & 0 & \lambda \\ 0 & 0 & -1 \end{bmatrix} = 0 \quad \text{für alle } \lambda,$$

also ist das Paar (E, A) nicht regulär.

Lemma 2.10 *Jedes Matrizenpaar, welches äquivalent zu einem regulären Paar ist, ist ebenfalls regulär. Äquivalente Paare haben dieselben verallgemeinerten Eigenwerte.*

Beweis: $E_2 = PE_1Q$, $A_2 = PA_1Q$, mit P, Q nicht singulär.

$$\begin{aligned} \det(\lambda E_2 - A_2) &= \det(\lambda PE_1Q - PA_1Q) \\ &= \det P \cdot \det(\lambda E_1 - A_1) \cdot \det Q. \\ \implies \det(\lambda E_2 - A_2) = 0 &\iff \det(\lambda E_1 - A_1) = 0. \end{aligned}$$

□

Satz 2.11 Weierstraß-Normalform

Seien $E, A \in \mathbb{C}^{n,n}$ und sei (E, A) regulär. Dann gibt es nichtsinguläre Matrizen $P, Q \in \mathbb{C}^{n,n}$, so dass

$$PEQ = \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, \quad PAQ = \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix},$$

wobei J eine Matrix in Jordan'scher Normalform ist und N eine nilpotente Matrix in Jordan'scher Normalform. Dabei sind N, J (bis auf Permutation der Jordanblöcke zu verschiedenen Eigenwerten) eindeutig bestimmt.

Beweis: Da (E, A) regulär, so gibt es $\lambda_0 \in \mathbb{C}$ mit $\det(\lambda_0 E - A) \neq 0$. Dann ist (E, A) äquivalent zu

$$\begin{aligned} &((\lambda_0 E - A)^{-1}E, (\lambda_0 E - A)^{-1}A) \\ &= ((\lambda_0 E - A)^{-1}E, (\lambda_0 E - A)^{-1}(-\lambda_0 E + A + \lambda_0 E)) \\ &= ((\lambda_0 E - A)^{-1}E, -I + \lambda_0(\lambda_0 E - A)^{-1}E) \\ &=: (\tilde{E}, -I + \lambda_0 \tilde{E}) \end{aligned}$$

Sei P_1 nichtsingulär, so daß

$$P_1^{-1} \tilde{E} P_1 = P_1^{-1} \left((\lambda_0 E - A)^{-1} E \right) P_1 = \tilde{J}$$

in Jordan'scher Normalform ist (existiert nach I./Folgerung 16.6).

$$\implies (\tilde{E}, -I + \lambda_0 \tilde{E}) \text{ ist äquivalent zu } (P_1^{-1} \tilde{E} P_1, P_1^{-1} (-I + \lambda_0 \tilde{E}) P_1) = (\tilde{J}, -I + \lambda_0 \tilde{J}).$$

Teile \tilde{J} auf als

$$\begin{bmatrix} \tilde{J}_1 & 0 \\ 0 & \tilde{N} \end{bmatrix}, \quad \text{mit } \tilde{J}_1 \text{ nichtsingulär und } \tilde{N} \text{ nilpotent.}$$

Wir erhalten also

$$(\tilde{J}, -I + \lambda_0 \tilde{J}) = \left(\left[\begin{array}{cc|c} \tilde{J}_1 & 0 & \\ \hline 0 & \tilde{N} & \\ \hline \end{array} \right], \left[\begin{array}{cc|c} -I + \lambda_0 \tilde{J}_1 & & 0 \\ \hline & & -I + \lambda_0 \tilde{N} \\ \hline \end{array} \right] \right).$$

$-I + \lambda_0 \tilde{N}$ ist nichtsingulär, da \tilde{N} strikte obere Dreiecksmatrix ist, also folgt, dass $(\tilde{J}, -I + \lambda_0 \tilde{J})$ äquivalent zu

$$\begin{aligned} & \left(\left[\begin{array}{cc|c} \tilde{J}_1^{-1} & 0 & \\ \hline 0 & (-I + \lambda_0 \tilde{N})^{-1} & \\ \hline \end{array} \right] \left[\begin{array}{cc} \tilde{J}_1 & 0 \\ 0 & \tilde{N} \end{array} \right], \left[\begin{array}{cc|c} \tilde{J}_1^{-1} & 0 & \\ \hline 0 & (-I + \lambda_0 \tilde{N})^{-1} & \\ \hline \end{array} \right] \left[\begin{array}{cc|c} -I + \lambda_0 \tilde{J}_1 & 0 & \\ \hline 0 & & -I + \lambda_0 \tilde{N} \\ \hline \end{array} \right] \right) \\ & = \left(\left[\begin{array}{cc} I & 0 \\ 0 & \hat{N} \end{array} \right], \left[\begin{array}{cc} \hat{J} & 0 \\ 0 & I \end{array} \right] \right) \end{aligned}$$

Seien P_2, P_3 nicht singuläre Matrizen, so dass $J = P_2 \hat{J} P_2^{-1}$ in Jordan'scher Normalform und $N = P_3 \hat{N} P_3^{-1}$ nilpotent in Jordan'scher Normalform. So folgt, dass

$$\begin{aligned} & \left[\begin{array}{cc} P_2 & 0 \\ 0 & P_3 \end{array} \right] \left[\begin{array}{cc} I & 0 \\ 0 & \hat{N} \end{array} \right] \left[\begin{array}{cc|c} P_2^{-1} & 0 & \\ \hline 0 & P_3^{-1} & \\ \hline \end{array} \right] = \left[\begin{array}{cc} I & 0 \\ 0 & N \end{array} \right] \\ \text{und} & \left[\begin{array}{cc} P_2 & 0 \\ 0 & P_3 \end{array} \right] \left[\begin{array}{cc} \hat{J} & 0 \\ 0 & I \end{array} \right] \left[\begin{array}{cc|c} P_2^{-1} & 0 & \\ \hline 0 & P_3^{-1} & \\ \hline \end{array} \right] = \left[\begin{array}{cc} J & 0 \\ 0 & I \end{array} \right]. \end{aligned}$$

Wir haben gezeigt, dass jedes reguläre Paar äquivalent zu einem Paar der Form

$$\left(\left[\begin{array}{cc} I & 0 \\ 0 & N \end{array} \right], \left[\begin{array}{cc} J & 0 \\ 0 & I \end{array} \right] \right)$$

ist, wir müssen noch die Eindeutigkeit zeigen.

Seien

$$\left(\left[\begin{array}{cc} I & 0 \\ 0 & N_1 \end{array} \right], \left[\begin{array}{cc} J_1 & 0 \\ 0 & I \end{array} \right] \right), \quad \left(\left[\begin{array}{cc} I & 0 \\ 0 & N_2 \end{array} \right], \left[\begin{array}{cc} J_2 & 0 \\ 0 & I \end{array} \right] \right)$$

zwei Normalformen von (E, A) , d.h.

$$P_1EQ_1 = \begin{bmatrix} I & 0 \\ 0 & N_1 \end{bmatrix}, \quad P_1AQ_1 = \begin{bmatrix} J_1 & 0 \\ 0 & I \end{bmatrix}$$

$$P_2EQ_2 = \begin{bmatrix} I & 0 \\ 0 & N_2 \end{bmatrix}, \quad P_2AQ_2 = \begin{bmatrix} J_2 & 0 \\ 0 & I \end{bmatrix}$$

\implies

$$P_1P_2^{-1} \begin{bmatrix} I & 0 \\ 0 & N_2 \end{bmatrix} Q_2^{-1}Q_1 = \begin{bmatrix} I & 0 \\ 0 & N_1 \end{bmatrix}$$

$$P_1P_2^{-1} \begin{bmatrix} J_2 & 0 \\ 0 & I \end{bmatrix} Q_2^{-1}Q_1 = \begin{bmatrix} J_1 & 0 \\ 0 & I \end{bmatrix}$$

Setze $P_1P_2^{-1} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$, $Q_1^{-1}Q_2 = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$,

so folgt

$$P_{11} = Q_{11}, \quad P_{12}N_2 = Q_{12}, \quad P_{21} = N_1Q_{21}, \quad P_{22}N_2 = N_1Q_{22}$$

und

$$P_{11}J_2 = J_1Q_{11}, \quad P_{12} = J_1Q_{12}, \quad P_{21}J_2 = Q_{21}, \quad P_{22} = Q_{22}.$$

$$\implies P_{21} = N_1P_{21}J_2 = N_1^2P_{21}J_2^2 = \dots = N_1^lP_{21}J_2^l = 0, \text{ f\"ur ein } l.$$

$$\implies P_{11} = Q_{11}, \quad P_{22} = Q_{22} \text{ nichtsingul\"ar.}$$

$$\implies P_{11}J_2P_{11}^{-1} = J_1 \implies J_1 = J_2, \text{ bis auf Permutation verschiedener Jordanbl\"ocke}$$

und

$$P_{22}N_2P_{22}^{-1} = N_1 \implies N_1 = N_2,$$

denn die Jordan'sche Normalform ist (bis auf Permutationen) eindeutig. \square

Beispiel 2.12

$$\begin{aligned} & \left(\begin{bmatrix} L & 0 & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -R & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \right) \\ & \sim \left(\begin{bmatrix} L & 0 & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -R & -1 & -1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \right) \\ & \sim \left(\begin{bmatrix} L & 0 & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -R & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right) \end{aligned}$$

$$\begin{aligned}
&\sim \left(\left[\begin{array}{ccc|c} L & 0 & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right], \left[\begin{array}{ccc|c} -R & 1 & -1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array} \right] \right) \\
&\sim \left(\left[\begin{array}{ccc|c} L & 0 & 0 & 0 \\ 0 & C & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right], \left[\begin{array}{ccc|c} -R & -1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array} \right] \right) \\
&\sim \left(\left[\begin{array}{ccc|c|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right], \left[\begin{array}{cc|c|c} -\frac{R}{L} & -\frac{1}{L} & 0 & 0 \\ \frac{1}{C} & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right] \right) \\
&\sim \left(\left[\begin{array}{ccc|c|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right], \left[\begin{array}{cc|c|c} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right] \right)
\end{aligned}$$

$$N = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & 0 \end{array} \right], \quad J = \left[\begin{array}{cc} \lambda_1 & 0 \\ 0 & \lambda_2 \end{array} \right].$$

Definition 2.13 Sei (E, A) ein reguläres Paar mit Weierstraß-Normalform

$$\left(\left[\begin{array}{cc} I & 0 \\ 0 & N \end{array} \right], \left[\begin{array}{cc} J & 0 \\ 0 & I \end{array} \right] \right),$$

So heißt der Nilpotenzindex von N der Index von (E, A) . Wir schreiben $\text{Ind}(E, A)$.

Die verallgemeinerten Eigenwerte sind die Eigenwerte von J und dazu kommt noch der Eigenwert ∞ von $(\lambda N - I)$, denn

$$\det(N - \mu I) = 0$$

hat nur die Nullstelle 0 und das charakteristische Polynom $\mu^{\text{Ind}(N, I)}$.

$$\det(N - \mu I) = \mu \det\left(\frac{1}{\mu}N - I\right).$$

Für „ $\lambda = \frac{1}{\mu} = \infty$ “, d.h. $(\mu = \frac{1}{\lambda} = 0)$, ergibt sich die entsprechende Jordanstruktur von N .

Nun können wir auch die differentiell-algebraische Gleichung

$$E\dot{x} = Ax + f, \quad x(t_0) = x^0,$$

mit (E, A) regulär, lösen.

Seien P, Q die Matrizen, die (E, A) auf Weierstraß-Normalform bringen, so gilt mit $x = Qy$

$$PEQ\dot{y} = PAQy + Pf. \quad y(t_0) = Q^{-1}x^0 =: y^0.$$

Teile $g := Pf$ und y entsprechend der Aufteilung von (PEQ, PAQ) auf, d.h.

$$\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} \begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \end{bmatrix} = \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \quad (2.14)$$

$$y_1(t_0) = y_1^0, \quad y_2(t_0) = y_2^0.$$

So ist y_1 die Lösung von

$$\dot{y}_1 = Jy_1 + g_1, \quad y_1(t_0) = y_1^0. \quad (2.15)$$

Existenz, Eindeutigkeit und Lösungsdarstellung siehe letztes Kapitel.
 y_2 ist die Lösung von

$$Ny_2 = y_2 + g_2 \quad (2.16)$$

mit Anfangsbedingung

$$y_2(t_0) = y_2^0. \quad (2.17)$$

Satz 2.18 Sei $g_2 \in C^\nu([t_0, t_1], \mathbb{C}^l)$, wobei ν der Nilpotenzindex der nilpotenten Matrix $N \in \mathbb{C}^{l,l}$ ist, so hat die Lösung von (2.16) die Form

$$y_2 = - \sum_{i=0}^{\nu-1} N^i g_2^{(i)}. \quad (2.19)$$

Die Lösung ist eindeutig, falls keine Anfangsbedingung (2.17) gegeben ist. Ist $y_2(t_0) = y_2^0$ als Anfangsbedingung gegeben, so existiert die Lösung und ist eindeutig genau dann, wenn

$$y_2^0 = - \sum_{i=0}^{\nu-1} N^i g_2^{(i)}(t_0). \quad (2.20)$$

Beweis: Dass jede Funktion der Form (2.19) die Gleichung (2.16) löst, folgt durch Einsetzen.

$$\begin{aligned} Ny_2 &= N \cdot \left(- \sum_{i=0}^{\nu-1} N^i g_2^{(i+1)} \right) \\ &= - \sum_{i=0}^{\nu-1} N^{i+1} g_2^{(i+1)} \\ &= - \sum_{i=1}^{\nu} N^i g_2^{(i)} \\ &= - \sum_{i=1}^{\nu-1} N^i g_2^{(i)} \quad (N^\nu = 0) \\ &= - \underbrace{\sum_{i=0}^{\nu-1} N^i g_2^{(i)}}_{y_2} + \underbrace{N^0 g_2^{(0)}}_{g_2}. \end{aligned}$$

Dass jede Lösung die Form (2.19) hat, folgt durch Lösung der einzelnen Teilblöcke der Jordan-Form von N .

$$N = \begin{bmatrix} N_1 & & \\ & \ddots & \\ & & N_k \end{bmatrix}, \quad N_i = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} \in \mathbb{C}^{l_i, l_i}$$

$$N_i \dot{z}_i = z_i + h_i.$$

Der Rest ist klar. Da (2.19) schon eindeutig ist, muss die Anfangsbedingung erfüllt sein oder die Lösung existiert nicht. \square

Bemerkung 2.21 *Im Gegensatz zu gewöhnlichen Differentialgleichungen können wir Anfangsbedingungen nicht beliebig wählen.*

Außerdem muss die Inhomogenität genügend oft ($\text{Ind}(E, A)$ -mal) differenzierbar sein.

Kapitel 3

Bilinearform und Hermite'sche Form

Ein wichtiges Hilfsmittel in vielen Teilen der Mathematik ist der Begriff der Bilinearform, den wir nun einführen wollen.

Definition 3.1 Sei V ein Vektorraum über einem Körper K . Eine Abbildung

$$\beta : V \times V \rightarrow K$$

heißt Bilinearform, falls folgende Bedingungen gelten:

$$(i) \quad \beta(v_1 + v_2, w) = \beta(v_1, w) + \beta(v_2, w)$$

$$(ii) \quad \beta(\lambda v, w) = \lambda \beta(v, w)$$

$$(iii) \quad \beta(v, w_1 + w_2) = \beta(v, w_1) + \beta(v, w_2)$$

$$(iv) \quad \beta(v, \lambda w) = \lambda \beta(v, w).$$

für alle $v, v_1, v_2, w, w_1, w_2 \in V, \lambda \in K$.

Die Bilinearform heißt symmetrisch, falls

$$\beta(v, w) = \beta(w, v), \quad \forall v, w \in V.$$

Die Abbildung β heißt Hermite'sche Form, falls (i), (ii) und (v) erfüllt sind.

$$(v) \quad \beta(v, w) = \overline{\beta(w, v)}, \quad \forall v, w \in V.$$

Beispiel 3.2

(a) Das in Teil I Kap. 9 eingeführte Skalarprodukt in \mathbb{R}^n

$$\langle u, v \rangle = v^\top u$$

ist eine symmetrische Bilinearform. Das analoge komplexe Skalarprodukt in \mathbb{C}^n ist eine Hermite'sche Form.

(b) Sei $V = \mathcal{C}([a, b], \mathbb{R})$. So bildet die Abbildung

$$\beta(f, g) = \int_a^b f \cdot g dx$$

eine symmetrische Bilinearform.

Lemma 3.3

- (i) Sei V ein Vektorraum über \mathbb{R} und $\beta : V \times V \rightarrow \mathbb{R}$ eine symmetrische Bilinearform. β definiert ein Skalarprodukt in V genau dann, wenn $\beta(v, v) > 0$ für $v \neq 0$.
- (ii) Sei V ein Vektorraum über \mathbb{C} und $\beta : V \times V \rightarrow \mathbb{C}$ eine Hermite'sche Form. β definiert ein komplexes Skalarprodukt in V genau dann, wenn $\beta(v, v) > 0$ für $v \neq 0$.

Beweis: Aus den Eigenschaften der symmetrischen Bilinearform (Hermite'schen Form) folgen sofort Bedingungen (1) und (2) des Skalarprodukts und die dritte Bedingung „positive Definitheit“ ist eben genau

$$\beta(v, v) > 0, \quad v \neq 0.$$

□

Es gibt eine Matrixdarstellung jeder Bilinearform. Sei V ein endlichdimensionaler Vektorraum über K und $\beta : V \times V \rightarrow K$ eine Bilinearform. Sei $\{v_1, \dots, v_n\}$ eine Basis von V . Dann heißt die Matrix $B \in K^{n,n}$, mit den Einträgen

$$b_{ij} = \beta(v_j, v_i), \quad i, j = 1, \dots, n,$$

die Matrixdarstellung von β bezüglich der Basis $\{v_1, \dots, v_n\}$.

Falls β eine symmetrische Bilinearform ist, so ist die Matrixdarstellung B bezüglich jeder Basis symmetrisch, d.h., $B = B^\top$. Falls β eine Hermite'sche Form ist, so ist die Matrixdarstellung B bezüglich jeder Basis Hermite'sch, d.h., $B = B^H = \overline{B}^\top$.

Beachte, auch im komplexen Fall gibt es symmetrische Bilinearformen, die nicht Hermite'sch sind (komplex symmetrische Matrizen).

Beispiel 3.4

$$V = \mathbb{C}^3, D = \begin{bmatrix} 1 & & \\ & 1 & \\ & & -1 \end{bmatrix}, \quad \beta_1(v, w) = w^\top Dv = v_1 w_1 + v_2 w_2 - v_3 w_3$$

ist eine symmetrische Bilinearform, denn

$$\beta_1(w, v) = w_1 v_1 + w_2 v_2 - w_3 v_3 = \beta_1(v, w).$$

Dagegen ist $\beta_2(v, w) = w^H Dv$ eine Hermite'sche Form.

Für die Basis $\{e_1, e_2, e_3\}$ ergibt sich die Matrixdarstellung $B = D$ und für

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ c \\ -\bar{s} \end{bmatrix}, \begin{bmatrix} 0 \\ s \\ \bar{c} \end{bmatrix} \right\} \quad \text{mit} \quad |c|^2 + |s|^2 = 1$$

ergibt sich

$$\begin{aligned} B &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & s \\ 0 & -\bar{s} & \bar{c} \end{bmatrix}^H D \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & s \\ 0 & -\bar{s} & \bar{c} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \bar{c} & -s \\ 0 & \bar{s} & c \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & s \\ 0 & \bar{s} & -\bar{c} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & |c|^2 - |s|^2 & 2\bar{c}s \\ 0 & 2\bar{s}c & |s|^2 - |c|^2 \end{bmatrix}. \end{aligned}$$

Lemma 3.5 Seien $\{v_1, \dots, v_n\}, \{\tilde{v}_1, \dots, \tilde{v}_n\}$ Basen des Vektorraumes V über \mathbb{R} und sei P die Basisübergangsmatrix, d.h., falls $v = \sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \tilde{\lambda}_i \tilde{v}_i$, so gilt

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = P \begin{bmatrix} \tilde{\lambda}_1 \\ \vdots \\ \tilde{\lambda}_n \end{bmatrix}.$$

Ist $\beta : V \times V \rightarrow \mathbb{R}$ eine Bilinearform und B die Matrixdarstellung von β bezüglich $\{v_1, \dots, v_n\}$, so ist

$$\tilde{B} = P^\top B P$$

die Matrixdarstellung von β bezüglich $\{\tilde{v}_1, \dots, \tilde{v}_n\}$.

Beweis:

$$v = \sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \tilde{\lambda}_i \tilde{v}_i,$$

$$w = \sum_{i=1}^n \delta_i v_i = \sum_{i=1}^n \tilde{\delta}_i \tilde{v}_i.$$

$$\Lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}, \quad \tilde{\Lambda} = \begin{bmatrix} \tilde{\lambda}_1 \\ \vdots \\ \tilde{\lambda}_n \end{bmatrix}, \quad \Delta = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_n \end{bmatrix}, \quad \tilde{\Delta} = \begin{bmatrix} \tilde{\delta}_1 \\ \vdots \\ \tilde{\delta}_n \end{bmatrix}$$

So gilt $\Lambda = P\tilde{\Lambda}$, $\Delta = P\tilde{\Delta}$ und

$$\begin{aligned}
 \beta(v, w) &= \beta\left(\sum_{i=1}^n \lambda_i v_i, \sum_{j=1}^n \delta_j v_j\right) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \delta_j \beta(v_i, v_j) \\
 &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \delta_j b_{ij} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}^\top B \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_n \end{bmatrix} = (P\tilde{\Lambda})^\top B(P\tilde{\Delta}) \\
 &= \tilde{\Lambda}^\top \underbrace{P^\top B P}_{\tilde{B}} \tilde{\Delta} = \sum_{ij=1}^n \tilde{\lambda}_i \tilde{\delta}_j \tilde{b}_{ij} = \sum_{i=1}^n \tilde{\lambda}_i \tilde{\delta}_j \beta(\tilde{v}_i, \tilde{v}_j) \\
 &= \beta(\tilde{v}, \tilde{w}).
 \end{aligned}$$

□

Definition 3.6 Seien $A, B \in \mathbb{R}^{n,n}$, so heißen A und B kongruent, falls es eine nichtsinguläre Matrix P gibt, so dass $B = P^\top A P$.

Analog, falls $A, B \in \mathbb{C}^{n,n}$, so heißen A und B kongruent, falls es eine nichtsinguläre Matrix P gibt, so dass $B = P^H A P$.

Zur Wiederholung: Ein reeller Vektorraum V mit Skalarprodukt (positiv definiten symmetrischer Bilinearform) heißt Euklidischer Raum (V, β) oder $(V, \langle \cdot, \cdot \rangle)$.

Analog nennen wir einen komplexen Vektorraum mit Skalarprodukt (positiv definiten Hermitescher Form) unitären Raum (V, β) oder $(V, \langle \cdot, \cdot \rangle)$.

Falls wir einen euklidischen (unitären) Raum haben, verwenden wir typischerweise $(V, \langle \cdot, \cdot \rangle)$ und nicht (V, β) .

Definition 3.7 Sei V ein euklidischer (unitärer) Raum. Ein Endomorphismus $f : V \rightarrow V$ heißt orthogonal (unitär), falls

$$\langle f(v), f(w) \rangle = \langle v, w \rangle$$

für alle $v, w \in V$.

Korollar 3.8 Sei V ein euklidischer (unitärer) Raum mit Orthonormalbasis $\{v_1, \dots, v_n\}$. Sei $f : V \rightarrow V$ ein orthogonaler (unitärer) Endomorphismus. Dann gelten folgende Aussagen:

- (i) Die Matrixdarstellung von f bezüglich der Basis $\{v_1, \dots, v_n\}$ ist eine orthogonale (unitäre) Matrix. (Die Umkehrung, dass jede orthogonale (unitäre) Matrix einen orthogonalen (unitären) Endomorphismus definiert, gilt natürlich auch.)
- (ii) f ist injektiv und f^{-1} ist ebenfalls orthogonal (unitär).
- (iii) Ist λ ein Eigenwert von f , so ist $|\lambda| = 1$.
- (iv) Ist $\|\cdot\|$ die durch $\langle \cdot, \cdot \rangle$ induzierte Norm, d.h., $\|v\| = \sqrt{\langle v, v \rangle}$ so gilt $\|f(v)\| = \|v\|$.

Beweis:

- (i) Sei $F = [f_{ij}]$ die Matrixdarstellung von f , d.h., $f(v) = Fv$. Wir erhalten wie üblich die Elemente f_{ij} über die Basisdarstellung der Bilder von f , d.h.,

$$f(v_j) = \sum_{i=1}^n f_{ij}v_i \quad j = 1, \dots, n.$$

Im reellen Fall ist mit $v = \sum_{i=1}^n \lambda_i v_i$, $w = \sum_{i=1}^n \gamma_i v_i$

$$\begin{aligned} \langle v, w \rangle &= \sum_{i,j=1}^n \lambda_i \gamma_j \langle v_i, v_j \rangle = \sum_{i=1}^n \lambda_i \gamma_i \\ \langle f(v), f(w) \rangle &= \langle Fv, Fw \rangle \\ &= \sum_{i,j=1}^n \lambda_i \gamma_j \langle Fv_i, Fv_j \rangle \\ &= \sum_{i,j=1}^n \lambda_i \gamma_j v_j^\top F^\top F v_i = \langle v, F^\top F w \rangle \end{aligned}$$

f ist orthogonal $\Rightarrow \langle v, w \rangle = \langle f(v), f(w) \rangle \quad \forall v, w \in V$, und damit:

$$\langle v, w \rangle = \langle v, F^\top F w \rangle \quad \forall v, w \in V,$$

insbesondere damit

$$w = F^\top F w, \quad \forall w \in V \quad \Longrightarrow \quad F^\top F = I.$$

Die Umkehrung ist trivialerweise erfüllt, denn im reellen Fall

$$\langle Fv, Fw \rangle = \langle v, F^\top F w \rangle = \langle v, w \rangle.$$

und im komplexen

$$\langle Fv, Fw \rangle = \langle v, F^H F w \rangle = \langle v, w \rangle.$$

- (iii) Sei v ein Eigenvektor von f , d.h., $f(v) = \lambda v$.

Dann folgt, da f (als invertierbare Matrix F) nicht den Eigenwert 0 hat,

$$\begin{aligned} 0 \neq \langle v, v \rangle &= \langle f(v), f(v) \rangle = \langle \lambda v, \lambda v \rangle \\ &= \begin{cases} \lambda^2 \langle v, v \rangle, & \text{falls } K = \mathbb{R} \text{ und } \lambda \in \mathbb{R} \\ \lambda \bar{\lambda} \langle v, v \rangle, & \text{falls } K = \mathbb{C} \text{ oder } \lambda \notin \mathbb{R}, \end{cases} \\ \implies 1 &= \begin{cases} \lambda^2, & \text{falls } K = \mathbb{R} \text{ und } \lambda \in \mathbb{R} \\ |\lambda|^2, & \text{falls } K = \mathbb{C} \text{ oder } \lambda \notin \mathbb{R}. \end{cases} \end{aligned}$$

(Falls $K = \mathbb{R}$ und $\lambda \notin \mathbb{R}$, muss das komplexe Skalarprodukt verwendet werden.)

Also $\lambda \in \{+1, -1\}$, falls $K = \mathbb{R}$ und $\lambda \in \mathbb{R}$, und $|\lambda| = 1$, falls $K = \mathbb{C}$ oder $\lambda \notin \mathbb{R}$.

(iv)

$$\begin{aligned}\|v\| &= \sqrt{\langle v, v \rangle} = \sqrt{\langle f(v), f(v) \rangle} \\ &= \|f(v)\|\end{aligned}$$

□

Die Umkehrung von Teil (iv) des Korollar 3.8 gilt ebenfalls.

Lemma 3.9 Sei V ein euklidischer Vektorraum und $f : V \rightarrow V$ eine beliebige Abbildung mit $f(0) = 0$ und

$$\|v - w\| = \|f(v) - f(w)\| = \|f(v - w)\| \quad \text{für alle } v, w \in V,$$

so ist f ein orthogonaler Endomorphismus.

Beweis: Es gilt für alle $v, w \in V$

$$\langle v, w \rangle = \frac{1}{2} (\langle v, v \rangle + \langle w, w \rangle - \langle v - w, v - w \rangle)$$

Analog

$$\begin{aligned}\langle f(v), f(w) \rangle &= \frac{1}{2} (\langle f(v), f(v) \rangle + \langle f(w), f(w) \rangle - \langle f(v - w), f(v - w) \rangle) \\ &= \frac{1}{2} (\langle f(v), f(v) \rangle + \langle f(w), f(w) \rangle - \langle v - w, v - w \rangle)\end{aligned}$$

mit $w = 0$ ergibt sich

$$\begin{aligned}0 &= \langle f(v), 0 \rangle = \frac{1}{2} (\langle f(v), f(v) \rangle + \langle 0, 0 \rangle - \langle v, v \rangle) \\ \implies \langle f(v), f(w) \rangle &= \frac{1}{2} (\langle v, v \rangle + \langle w, w \rangle - \langle v - w, v - w \rangle) \\ &= \langle v, w \rangle.\end{aligned}$$

Wir müssen nur noch die Linearität von f zeigen, d.h.,

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2), \quad \forall v_1, v_2 \in V, \lambda_1, \lambda_2 \in K.$$

$$\begin{aligned}&\|f(\lambda_1 v_1 + \lambda_2 v_2) - \lambda_1 f(v_1) - \lambda_2 f(v_2)\|^2 \\ &= \langle f(\lambda_1 v_1 + \lambda_2 v_2), f(\lambda_1 v_1 + \lambda_2 v_2) \rangle + \lambda_1^2 \langle f(v_1), f(v_1) \rangle + \lambda_2^2 \langle f(v_2), f(v_2) \rangle \\ &\quad - 2\lambda_1 \langle f(v_1), f(\lambda_1 v_1 + \lambda_2 v_2) \rangle - 2\lambda_2 \langle f(v_2), f(\lambda_1 v_1 + \lambda_2 v_2) \rangle \\ &\quad + 2\lambda_1 \lambda_2 \langle f(v_1), f(v_2) \rangle \\ &= \langle \lambda_1 v_1 + \lambda_2 v_2, \lambda_1 v_1 + \lambda_2 v_2 \rangle + \lambda_1^2 \langle v_1, v_1 \rangle + \lambda_2^2 \langle v_2, v_2 \rangle \\ &\quad - 2\langle \lambda_1 v_1 + \lambda_2 v_2, \lambda_1 v_1 + \lambda_2 v_2 \rangle \\ &\quad + 2\lambda_1 \lambda_2 \langle v_1, v_2 \rangle \\ &= 0 \implies \text{Behauptung.}\end{aligned}$$

□

Wir haben bereits gesehen, dass alle Eigenwerte von orthogonalen bzw. unitären Matrizen den Betrag 1 haben.

Satz 3.10 Reelle Schur-Form (Reelle Version von Teil I, Satz 13.15)

Sei $A \in \mathbb{R}^{n,n}$, dann gibt es $U \in \mathcal{O}(n) = \{A \in \mathbb{R}^{n,n} \mid A \text{ orthogonal}\}$, so dass

$$U^{-1}AU = \begin{bmatrix} R_{11} & \cdots & R_{1m} \\ & \ddots & \vdots \\ & & R_{mm} \end{bmatrix}$$

wobei die Diagonalblöcke R_{ii} entweder 1×1 oder 2×2 sind, und falls sie 2×2 sind, so haben sie ein Paar komplex konjugierter Eigenwerte.

Beweis: Da A reell ist, treten alle nicht reellen Eigenwerte in komplex konjugierten Paaren auf. Falls alle Eigenwerte reell sind, folgt der Beweis genau wie bei Teil I, Satz 13.15, nur reell. Ansonsten mit Induktion:

I.A.: $n = 1, n = 2 \Rightarrow$ klar.

I.V.: Die Behauptung sei richtig für $A \in \mathbb{R}^{n-1, n-1}$

I.S.: Falls $\lambda = a + ib$, $a, b \in \mathbb{R}$, Eigenwert von A mit Eigenvektor $x + iy$, so gilt

$$A(x + iy) = (a + ib)(x + iy)$$

und

$$A(x - iy) = (a - ib)(x - iy), \quad (x + iy)^\top (x - iy) = 1.$$

Es gilt dann durch Addition und Subtraktion dieser Gleichungen

$$A[x, y] = [x, y] \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

und damit ist $[x, y]$ ein reeller invarianter Unterraum.

Mit der QR -Zerlegung (Gram-Schmidt, Teil I, Folgerung 10.13) gibt es U_1 orthogonal, so dass

$$U_1^\top [x, y] = \begin{bmatrix} x_{11} & y_{11} \\ 0 & y_{21} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}, \quad \text{mit} \quad \begin{bmatrix} x_{11} & y_{11} \\ 0 & y_{21} \end{bmatrix} \text{ invertierbar,}$$

und damit

$$U_1^\top AU_1 U_1^\top [x, y] = U_1^\top [x, y] \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

und es folgt, dass

$$U_1^\top A U_1 = \left[\begin{array}{cc|c} a_{11} & a_{12} & * \\ a_{21} & a_{12} & \\ \hline 0 & 0 & \\ \vdots & \vdots & A_2 \\ 0 & 0 & \end{array} \right]$$

und daß $R_{11} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{12} \end{bmatrix}$ ähnlich ist zu $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$

Per Induktionsvoraussetzung gibt es \tilde{U}_2 , so dass

$$U_2^\top A_2 U_2 = \begin{bmatrix} R_{22} & & R_{2m} \\ & \ddots & \\ & & R_{mm} \end{bmatrix}$$

von der gewünschten Form ist.

Mit $U = U_1 \begin{bmatrix} I_2 & 0 \\ 0 & U_2 \end{bmatrix}$ folgt die Behauptung.

□

Wir kommen nun zum zentralen Satz für orthogonale Matrizen.

Satz 3.11

(i) Sei $A \in \mathcal{U}(n) = \{A \in \mathbb{C}^{n,n} \mid A \text{ unitär}\}$. Dann gibt es eine unitäre Matrix P , so dass

$$P^{-1} A P = P^H A P = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

(ii) Sei $A \in \mathcal{O}(n) := \{A \in \mathbb{R}^{n,n} \mid A \text{ orthogonal}\}$. Dann gibt es eine orthogonale Matrix P , so dass

$$P^{-1} A P = P^\top A P = \begin{bmatrix} I_p & & & \\ & -I_q & & \\ & & D_1 & \\ & & & \ddots \\ & & & & D_s \end{bmatrix}$$

mit $D_i = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \in \mathbb{R}^{2,2}$, $c^2 + s^2 = 1$.

Beweis:

- (i) Nach dem Satz von Schur (Teil I, Satz 13.15) gibt es eine unitäre Matrix $P \in \mathbb{C}^{n,n}$, so dass

$$P^{-1}AP = P^H AP = R = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ & & r_{nn} \end{bmatrix}.$$

Da P^{-1} , A , P unitär sind, folgt mit der Übung (da $\mathcal{U}(n)$ eine Gruppe ist), dass R unitär.

$$\Rightarrow R^H R = I_n, \quad \text{d.h.,} \quad \delta_{ik} = \sum_{j=1}^n \bar{r}_{ji} r_{jk}.$$

Es gilt

$$|r_{11}|^2 = 1 \quad \Rightarrow \quad r_{11} \neq 0$$

und

$$\bar{r}_{11} r_{1j} = 0, \quad j = 2, \dots, n, \quad \Rightarrow \quad r_{1j} = 0, \quad j = 2, \dots, n.$$

Per Induktion folgt dann die Behauptung.

Um Teil (ii) zu beweisen, verwenden wir analog die reelle Schurform aus 3.10. □

Kapitel 4

Hermite'sche, symmetrische Matrizen, adjungierte Endomorphismen

Im folgenden sei V ein euklidischer oder ein unitärer Raum.

Definition 4.1 Seien $f, g : V \rightarrow V$ Endomorphismen. g heißt adjungiert zu f , falls

$$\langle g(v), w \rangle = \langle v, f(w) \rangle \quad \forall v, w \in V.$$

Lemma 4.2 Sei V ein endlichdimensionaler euklidischer (unitärer) Raum und $f : V \rightarrow V$ ein Endomorphismus. Dann gibt es einen eindeutigen adjungierten Endomorphismus f^{ad} zu f , und es gilt $(f^{ad})^{ad} = f$.

Beweis: Wir zeigen nur den reellen Fall, der Beweis im komplexen Fall ist Übungsaufgabe.

Sei $V^* = L(V, \mathbb{R})$ der Dualraum zu V , (siehe Teil I, Kapitel 11) und sei:

$$\begin{aligned} \varphi : V &\rightarrow V^* \\ v &\mapsto \langle v, \cdot \rangle \end{aligned}$$

φ ist ein Vektorraumisomorphismus (s. Teil I, Satz 11.7).

Sei $f^* : V^* \rightarrow V^*$ die zu f duale Abbildung

$$f^*(g) = gf, \quad \text{d.h.,} \quad (f^*(g))(v) = g(f(v)), \quad v \in V, g \in V^*,$$

und definiere

$$f^{ad} : V \rightarrow V \quad \text{durch} \quad f^{ad}(v) = (\varphi^{-1} f^* \varphi)(v).$$

f^{ad} ist natürlich ein Endomorphismus und es gilt

$$\begin{aligned} \langle f^{ad}(v), w \rangle &= \varphi(f^{ad}(v))(w) \\ &= (f^*(\varphi(v)))(w) \\ &= \varphi(v)(f(w)) = \langle v, f(w) \rangle \end{aligned}$$

Wir müssen noch die Eindeutigkeit zeigen. Sei $g : V \rightarrow V$ ein Endomorphismus mit

$$\langle g(v), w \rangle = \langle v, f(w) \rangle \implies \varphi(g(v)) = \varphi(f^{ad}(v))$$

$$\varphi \text{ injektiv} \implies g(v) = f^{ad}(v)$$

$$\begin{aligned} \langle (f^{ad})^{ad}(v), w \rangle &= \langle v, f^{ad}(w) \rangle = \langle w, f(v) \rangle = \langle f(v), w \rangle \\ \implies f(v) &= (f^{ad})^{ad}(v), \forall v \in V, \\ \implies f &= (f^{ad})^{ad} \end{aligned}$$

□

Beispiel 4.3

$$(a) \quad V = \mathbb{R}^3$$

$$f : \begin{array}{l} V \rightarrow V \\ v \mapsto Fv \end{array} \quad F = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 3 & 0 \end{bmatrix}$$

$$\langle g(v), w \rangle = \langle v, Fw \rangle = (Fw)^\top v = w^\top F^\top v = \langle F^\top v, w \rangle$$

\implies

$$g = f^{ad} : \begin{array}{l} V \rightarrow V \\ v \mapsto F^\top v \end{array}$$

$$(b) \quad V = \{f \in C^\infty([a, b], \mathbb{R}), f(a) = f(b) = 0\}$$

$$\langle f, g \rangle = \int_a^b f \cdot g \, dx, \quad D : \begin{array}{l} V \rightarrow V \\ f \mapsto f' \end{array}$$

$$\langle z(f), g \rangle = \langle f, D(g) \rangle$$

$$\int_a^b z(f) \cdot g \, dx = \int_a^b f \cdot g' \, dx = \underbrace{f \cdot g}_a^b - \int_a^b f' g \, dx = - \int_a^b f' g \, dx$$

$$\implies z(f) = -f' = D^{ad}(f).$$

Definition 4.4 Ein Endomorphismus $f : V \rightarrow V$ heißt selbstadjungiert, falls $\forall v, w \in V$

$$\langle f(v), w \rangle = \langle v, f(w) \rangle.$$

Lemma 4.5 Sei V ein endlichdimensionaler euklidischer (unitärer) Raum mit Orthonormalbasis $\{v_1, \dots, v_n\}$. Sei $f : V \rightarrow V$ ein Endomorphismus und F die Matrixdarstellung von f bzgl. $\{v_1, \dots, v_n\}$.

i) Falls V euklidisch, so ist f selbstadjungiert genau dann, wenn F symmetrisch.

ii) Falls V unitär, so ist f selbstadjungiert genau dann, wenn F hermite'sch.

Beweis:

$$i) \text{ Sei } v = \sum_{i=1}^n \lambda_i v_i, \quad w = \sum_{i=1}^n \gamma_i v_i, \quad \Lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}, \quad \Gamma = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix}.$$

$$\begin{aligned} \langle f(v), w \rangle &= \Gamma^\top F \Lambda \\ \langle v, f(w) \rangle &= (F\Gamma)^\top \Lambda = \Gamma^\top F^\top \Lambda \end{aligned}$$

$$\begin{aligned} \implies \Gamma^\top F^\top \Lambda &= \Gamma^\top F \Lambda, \quad \forall \Lambda, \Gamma \in \mathbb{R}^n. \\ \implies F^\top &= F. \end{aligned}$$

ii) Übungsaufgabe

□

Wir haben also den Zusammenhang

Euklidischer Raum

- symmetrische Bilinearform
- selbstadjungierter Endomorphismus
- symmetrische Matrix

Unitärer Raum

- Hermite'sche Form
- selbstadjungierter Endomorphismus
- Hermite'sche Matrix

Satz 4.6

(i) Sei $A \in \mathbb{R}^{n,n}$ symmetrisch, so gibt es $Q \in \mathcal{O}(n)$, so dass

$$Q^{-1} A Q = Q^\top A Q = D \in \mathbb{R}^{n,n}$$

diagonal ist.

Insbesondere ist damit A diagonalisierbar und alle Eigenwerte sind reell.

(ii) Sei $A \in \mathbb{C}^{n,n}$ Hermite'sch, so gibt es $Q \in \mathcal{U}(n)$, so dass

$$Q^{-1} A Q = Q^H A Q = D \in \mathbb{R}^{n,n}$$

diagonal ist.

Insbesondere ist damit A diagonalisierbar und alle Eigenwerte sind reell.

Beweis:

(i) Nach Satz 3.10 (reelle Schur-Form) gibt es $Q \in \mathcal{O}(n)$, so dass

$$Q^{-1}AQ = Q^{\top}AQ = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1m} \\ & \ddots & & \\ & & \ddots & \\ & & & R_{m,m} \end{bmatrix} = R$$

$$\text{Aber } R^{\top} = (Q^{\top}AQ)^{\top} = Q^{\top}A^{\top}Q = Q^{\top}AQ = R.$$

$$\text{Wegen } R = R^{\top} \implies R_{i,j} = 0, j > i.$$

Nach Satz 3.10 sind die R_{ii} (1×1) oder (2×2) reell symmetrisch.

Da aber im (2×2)-Fall $R_{ii} = \begin{bmatrix} a_i & b_i \\ b_i & c_i \end{bmatrix}$ das charakteristische Polynom

$$\lambda^2 - (a_i + c_i)\lambda + (a_i c_i) - b_i^2 = 0$$

die Nullstellen $\frac{a_i + c_i \pm \sqrt{(a_i - c_i)^2 + b_i^2}}{2}$ hat und diese beide reell sind, gibt es keine (2×2)-Blöcke.

\implies Behauptung.

(ii) Nach dem Satz von Schur, Teil I Satz 13.15 gibt es $Q \in \mathcal{U}(n)$, so dass

$$Q^{-1}AQ = Q^H A Q = \begin{bmatrix} r_{11} & & r_{1n} \\ & \ddots & \\ & & r_{nn} \end{bmatrix}$$

$$\text{Aber } (Q^H A Q)^H = Q^H A^H Q = Q^H A Q.$$

$$\implies r_{ij} = 0, j > i \text{ und } r_{ii} = \overline{r_{ii}}, \implies r_{ii} \text{ reell.}$$

□

Beispiel 4.7

$$A = \begin{bmatrix} 2 & -1 & \frac{3}{2}\sqrt{2}i \\ -1 & 2 & -\frac{3}{2}\sqrt{2}i \\ -\frac{3}{2}\sqrt{2}i & \frac{3}{2}\sqrt{2}i & 3 \end{bmatrix}$$

$$\begin{aligned} P_A(\lambda) &= \lambda^3 - 7\lambda^2 + \lambda(6 + 6 + 4 - \frac{9}{2} - \frac{9}{2} - 1) + (-12 - \frac{9}{2} - \frac{9}{2} + 9 + 3 + 9) \\ &= \lambda^3 - 7\lambda^2 + 6\lambda = (\lambda - 1)\lambda(\lambda - 6) \end{aligned}$$

Eigenwerte sind $\lambda_{1,2,3} = 1, 0, 6$.

Berechnung von Q :

Bilde

$$(\lambda_1 I - A) = \begin{bmatrix} -1 & 1 & -\frac{3}{2}\sqrt{2}i \\ 1 & -1 & \frac{3}{2}\sqrt{2}i \\ \frac{3}{2}\sqrt{2}i & -\frac{3}{2}\sqrt{2}i & -2 \end{bmatrix}$$

Eliminiere

$$Q_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad Q_1^H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$Q_1^H (\lambda_1 I - A) Q_1 = \begin{bmatrix} 0 & 0 & 0 \\ \frac{2}{\sqrt{2}} & -\frac{2}{\sqrt{2}} & 3\sqrt{2}i \\ \frac{3}{2}\sqrt{2}i & -\frac{3}{2}\sqrt{2}i & -2 \end{bmatrix} Q_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & 3i \\ 0 & -3i & -2 \end{bmatrix}$$

$$\implies Q_1^H A Q_1 = \lambda_1 I - \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & 3i \\ 0 & -3i & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -3i \\ 0 & 3i & 3 \end{bmatrix}$$

usw.

Satz 4.8 Sei V endlichdimensionaler euklidischer oder unitärer Raum. Seien f_1, f_2 selbstadjungierte Endomorphismen von $V \rightarrow V$. Dann sind die folgenden Aussagen äquivalent.

(i) Es gibt eine Orthonormalbasis $\{v_1, \dots, v_n\}$ von V , so dass $\{v_1, \dots, v_n\}$ Eigenvektoren für f_1 und f_2 sind, d.h. f_1 und f_2 sind simultan diagonalisierbar.

(ii) $f_1 \circ f_2 = f_2 \circ f_1$

(In Matrixterminologie: Zwei reelle symmetrische (Hermite'sche) Matrizen F_1, F_2 sind genau dann simultan diagonalisierbar, wenn $F_1 F_2 = F_2 F_1$.)

Beweis:

(i) \implies (ii)

Sei v_i Eigenvektor von f_1 zum Eigenwert λ_i^1 und von f_2 zum Eigenwert λ_i^2 .

$$\implies (f_1 \circ f_2)(v_i) = f_1(f_2(v_i)) = \lambda_i^1 \lambda_i^2 v_i = \lambda_i^2 \lambda_i^1 v_i = (f_2 \circ f_1)(v_i) \quad \forall i = 1, \dots, n.$$

(ii) \implies (i)

Da f_1, f_2 selbstadjungierte Endomorphismen sind, so gibt es Matrixdarstellungen F_1, F_2 und eine orthogonale (unitäre) Matrix P_1

$$\begin{aligned} P_1^H F_1 P_1 &= D_1, & \text{mit: } D_1 &= \text{diag}(\lambda_1 I_{j_1}, \dots, \lambda_k I_{j_k}) \text{ mit } \lambda_i \neq \lambda_j \text{ für } i \neq j, \\ P_1^H F_1 F_2 P_1 &= P_1^H F_1 P_1 P_1^H F_2 P_1 = D_1 P_1^H F_2 P_1 \\ &= P_1^H F_2 F_1 P_1 = P_1^H F_2 P_1 D_1. \end{aligned}$$

Setze $P_1^H F_2 P_1 = \begin{bmatrix} F_{11} & \cdots & F_{1k} \\ \vdots & \ddots & \vdots \\ F_{k1} & \cdots & F_{kk} \end{bmatrix}$, analog zu D_1 partitioniert. Es folgt

$$\lambda_i F_{ij} = F_{ij} \lambda_j \quad \forall i, j.$$

Wegen $\lambda_i \neq \lambda_j$ für $i \neq j$ gilt $F_{ij} = 0$ für $i \neq j$,

$$\implies P_1^H F_2 P_1 = \begin{bmatrix} F_{11} & & \\ & \ddots & \\ & & F_{kk} \end{bmatrix}.$$

Setze $\tilde{P}_1 = \text{diag} \begin{bmatrix} \Pi_1 & & \\ & \ddots & \\ & & \Pi_k \end{bmatrix}$, Π_i orthogonal (unitär), so dass $\Pi_i^H F_{ii} \Pi_i = \Delta_i$ diagonal,

so folgt, dass

$$\begin{aligned} \tilde{P}_1^H P_1^H F_1 P_1 \tilde{P}_1 &= \begin{bmatrix} \lambda_1 I_{j_1} & & \\ & \ddots & \\ & & \lambda_k I_{j_k} \end{bmatrix} & \text{und} \\ \tilde{P}_1^H P_1^H F_2 P_1 \tilde{P}_1 &= \begin{bmatrix} \Delta_1 & & \\ & \ddots & \\ & & \Delta_k \end{bmatrix} \end{aligned}$$

diagonal sind. □

Definition 4.9 Sei V ein endlichdimensionaler Vektorraum und $f : V \rightarrow V$ ein Endomorphismus.

f heißt normal, falls

$$f^{ad} \circ f = f \circ f^{ad}$$

In Matrixterminologie in $\mathbb{C}^{n,n}$ (oder $\mathbb{R}^{n,n}$): F heißt normal, falls

$$F^H F = F F^H \quad (F^\top F = F F^\top).$$

Beispiel 4.10

$$\begin{aligned} A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & \text{ so ist } A^\top A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ & \text{ und } AA^\top = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \\ A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} & \text{ so ist } A^\top A = A^2 = AA^\top. \end{aligned}$$

Satz 4.11 Sei V ein endlichdimensionaler Vektorraum, so gilt

- (i) Ist $f : V \rightarrow V$ ein orthogonaler (unitärer) Endomorphismus, so ist f normal.
- (ii) Ist $f : V \rightarrow V$ selbstadjungiert, so ist f normal.

Beweis:

- (i) $\forall v, w \in V$ gilt

$$\langle f(v), f(w) \rangle = \langle v, w \rangle \quad \text{und} \quad \langle f^{ad}(v), w \rangle = \langle v, f(w) \rangle.$$

Da V endlichdimensional ist, ist f invertierbar und f^{-1} orthogonal (unitär).

Also gilt $\forall v, w \in V$

$$\langle v, f(w) \rangle = \langle f^{-1}(v), w \rangle = \langle f^{ad}(v), w \rangle \implies f^{-1} = f^{ad} \quad \text{und somit}$$

$$f \circ f^{ad} = f^{ad} \circ f.$$

- (ii) $f = f^{ad} \implies f \circ f^{ad} = f \circ f = f^{ad} \circ f.$

□

Normale Matrizen (Endomorphismen) bilden also eine Obermenge von orthogonalen und symmetrischen (bzw. unitären und Hermite'schen) Matrizen (Endomorphismen). Die Eigenschaft der Diagonalisierbarkeit wird damit auch vererbt.

Lemma 4.12 Sei V ein endlichdimensionaler Vektorraum und $f : V \rightarrow V$ ein normaler Endomorphismus. Dann gilt

$$(i) \text{ Kern}(f) = \text{Kern}(f^{ad}).$$

$$(ii) \text{ Für alle Eigenwerte } \lambda \text{ von } f \text{ gilt } E(\lambda, f) = E(\bar{\lambda}, f^{ad}).$$

(Zur Erinnerung: $E(\lambda, f)$ ist der Eigenraum von f zum Eigenwert λ).

Beweis:

$$(i) \forall v \in V \text{ gilt}$$

$$\begin{aligned} \langle f(v), f(v) \rangle &= \langle (f^{ad} \circ f)(v), v \rangle \\ &= \langle (f \circ f^{ad})v, v \rangle = \langle f^{ad}(v), f^{ad}(v) \rangle \end{aligned}$$

Aus der Definitheit von $\langle \cdot, \cdot \rangle$ folgt die Behauptung.

$$(ii) \text{ Sei } h = f - \lambda \cdot id_V, \text{ so folgt } h^{ad} = f^{ad} - \bar{\lambda} \cdot id_V.$$

(Beachte, dass bei komplexen Eigenwerten auch das komplexe Skalarprodukt verwendet werden muss.)

$$\begin{aligned} \langle h^{ad}(v), w \rangle &= \langle v, h(w) \rangle \\ &= \langle v, f(w) \rangle - \langle v, \lambda w \rangle \\ &= \langle f^{ad}(v), w \rangle - \bar{\lambda} \langle v, w \rangle \\ &= \langle (f^{ad} - \bar{\lambda} \cdot id_V)(v), w \rangle. \end{aligned}$$

$$\begin{aligned} E(\lambda, f) &= \text{Kern}(f - \lambda \cdot id_V) \stackrel{(i)}{=} \text{Kern}(f^{ad} - \bar{\lambda} \cdot id_V) \\ &= E(\bar{\lambda}, f^{ad}). \end{aligned}$$

□

Damit können wir nun zeigen, dass die durch unitäre Ähnlichkeit diagonalisierbaren Endomorphismen genau die normalen sind.

Satz 4.13 Sei V ein endlichdimensionaler Vektorraum und $f : V \rightarrow V$ ein Endomorphismus. Dann sind die folgenden Aussagen äquivalent.

$$(i) \text{ Es gibt eine Orthonormalbasis von } V \text{ aus Eigenvektoren von } f.$$

$$(ii) f \text{ ist normal.}$$

(In Matrixterminologie: A ist unitär diagonalisierbar genau dann, wenn A normal ist.)

Beweis:

(i) \implies (ii) Sei $\{v_1, \dots, v_n\}$ ONB von V aus Eigenvektoren von f .

$$\langle f^{ad}(v_i), v_j \rangle = \langle v_i, f(v_j) \rangle = \langle v_i, \lambda v_j \rangle = \bar{\lambda} \langle v_i, v_j \rangle = \begin{cases} \bar{\lambda}, & i = j \\ 0, & i \neq j \end{cases},$$

$$\implies f^{ad}(v_i) = \sum_{j=1}^n \langle f^{ad}(v_i), v_j \rangle v_j = \bar{\lambda} v_i$$

$$\implies (f \circ f^{ad})(v_i) = \lambda \bar{\lambda} v_i = (f^{ad} \circ f)(v_i) \quad \forall i = 1, \dots, n$$

$$\implies f \circ f^{ad} = f^{ad} \circ f.$$

(ii) \implies (i) Mit Induktion:

I.A.: $n = 1$ ist klar.

I.V.: Behauptung sei richtig für $n - 1$.

I.S.: Sei v_1 mit $\|v_1\|_2 = 1$ ein Eigenvektor von f zum Eigenwert λ und $W = \mathcal{L}(v_1)^\perp$.

Sei $w \in W$, so gilt

$$\langle v_1, f(w) \rangle = \langle f^{ad}(v_1), w \rangle \stackrel{\text{Lemma 4.12}}{=} \langle \bar{\lambda} v_1, w \rangle = \bar{\lambda} \langle v_1, w \rangle = 0$$

$$\implies f(W) \subset W.$$

Weiter gilt: $f|_W$ ist natürlich auch normal.

Da $\dim(W) = \dim(V) - 1$, so gibt es ONB $\{v_2, \dots, v_n\}$ aus Eigenvektoren von $f|_W$. Dann ist $\{v_1, \dots, v_n\}$ ONB von V aus Eigenvektoren von f .

In Matrixterminologie ist der Beweis eine Folgerung vom Satz von Schur und Satz 4.8.

(i) \implies (ii) Falls A unitär diagonalisierbar, so gilt $P^*AP = D$,

$$\begin{aligned} P^*A^*AP &= P^*A^*PP^*AP \\ &= D^*D = DD^* = P^*AA^*P \\ \implies A^*A &= AA^*. \end{aligned}$$

(ii) \implies (i) Sei $A^*A = AA^*$.

$$\text{Nach Schur gibt es } P^*AP = \Delta = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ 0 & & r_{nn} \end{bmatrix} \quad (\text{obere Dreiecksmatrix}).$$

$$(P^*AP)^* = P^*A^*P = \Delta^*$$

$$\begin{array}{ccc} \begin{array}{|c|} \hline \triangle \\ \hline \end{array} \cdot \begin{array}{|c|} \hline \triangle \\ \hline \end{array} & = & \begin{array}{|c|} \hline \triangle \\ \hline \end{array} \cdot \begin{array}{|c|} \hline \triangle \\ \hline \end{array} \end{array}$$

$$\Delta^* \cdot \Delta = \Delta \cdot \Delta^*$$

$$|r_{11}|^2 = |r_{11}|^2 + \sum_{j=2}^n |r_{1j}|^2 \implies \sum_{j=2}^n |r_{1j}|^2 = 0 \implies \Delta = \left[\begin{array}{c|ccc} r_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \ddots & \\ 0 & & & \end{array} \right]$$

Rest mit Induktion.

□

Für normale Matrizen liefert also der Satz von Schur die Diagonalisierung.

Zusammenfassung:

A Hermite'sch, $A = A^H$
ONB aus Eigenvektoren
Eigenwerte reell

A reell symmetrisch, $A = A^\top$
ONB aus reellen Eigenvektoren
Eigenwerte reell

A unitär, $A^H A = I$
ONB aus Eigenvektoren
Eigenwerte $|\lambda| = 1$

A orthogonal, $A^\top A = I$
ONB aus i.a. komplexen Eigenvektoren
Eigenwerte $|\lambda| = 1$

A schiefhermite'sch, $A = -A^H$
ONB aus Eigenvektoren
Eigenwerte rein imaginär

A reell schiefsymmetrisch, $A = -A^\top$
ONB aus komplexen Eigenvektoren
Eigenwerte rein imaginär

Kapitel 5

Die Singulärwertzerlegung

Wir haben bisher verschiedene Transformationen kennengelernt, $A \in K^{n,m}$,

	Transformation		Normalform
$n = m$	$A \rightarrow PAQ$	P, Q nichtsingulär	$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$
$n = m$	$A \rightarrow Q^{-1}AQ$	Q nichtsingulär	Jordan'sche NF.
$n = m$	$A \rightarrow Q^H A Q$	Q unitär	Schur-Form

Es stellt sich die Frage, was sich ergibt, wenn wir $A \rightarrow PAQ$ betrachten, P, Q unitär (orthogonal).

Um die Betrachtungen elegant zu machen, führen wir zuerst mal ein paar neue Normen ein:

Sei $A = [a_{i,j}] \in \mathbb{C}^{n,m}(\mathbb{R}^{n,m})$.

$$\|A\|_F := \left[\sum_{i=1}^n \sum_{j=1}^m |a_{ij}|^2 \right]^{\frac{1}{2}} = (\operatorname{tr}(A^H A))^{\frac{1}{2}} \quad (5.1)$$

heißt Frobeniusnorm von A (Euklidische Norm in $\mathbb{C}^{n \cdot m}$ ($\mathbb{R}^{n \cdot m}$)).

$$\|A\|_2 := \sup_{\substack{x \neq 0 \\ x \in \mathbb{C}^m}} \frac{\|Ax\|_2}{\|x\|_2} \quad (5.2)$$

heißt Spektralnorm von A .

Lemma 5.3 *Es gilt für $A \in \mathbb{C}^{n,m}$, $P \in \mathbb{C}^{n,n}$, $Q \in \mathbb{C}^{m,m}$, P, Q unitär:*

$$\begin{aligned} \|A\|_F &= \|PAQ\|_F, \\ \|A\|_2 &= \|PAQ\|_2. \end{aligned}$$

(Analog für $A \in \mathbb{R}^{n,m}$, $P \in \mathbb{R}^{n,n}$, $Q \in \mathbb{R}^{m,m}$, P, Q orthogonal).

Beweis:

$$\|\cdot\|_F : \left(\operatorname{tr} \left((PAQ)^H (PAQ) \right) \right)^{\frac{1}{2}} = \left(\operatorname{tr} (Q^H A^H A Q) \right)^{\frac{1}{2}} = \left(\operatorname{tr} (A^H A) \right)^{\frac{1}{2}}.$$

$$\begin{aligned} \|\cdot\|_2 : \sup_{\substack{x \neq 0 \\ x \in \mathbb{C}^n}} \frac{\|Ax\|_2}{\|x\|_2} &= \sup_{x \neq 0} \frac{\sqrt{\langle Ax, Ax \rangle}}{\sqrt{\langle x, x \rangle}}, \\ \sup_{x \neq 0} \frac{\|PAQx\|_2}{\|x\|_2} &= \sup_{x \neq 0} \frac{\langle PAQx, PAQx \rangle^{\frac{1}{2}}}{\langle x, x \rangle^{\frac{1}{2}}} = \sup_{x \neq 0} \frac{\langle AQx, AQx \rangle^{\frac{1}{2}}}{\langle Qx, Qx \rangle^{\frac{1}{2}}} \\ &= \sup_{y \neq 0} \frac{\|Ay\|_2}{\|y\|_2}, \quad \text{mit } y = Qx. \end{aligned}$$

□

Satz 5.4 Sei $A \in \mathbb{C}^{n,m}$ ($\mathbb{R}^{n,m}$), $n \geq m$, so gibt es P, Q unitär (orthogonal), $P \in \mathbb{C}^{n,n}$, $Q \in \mathbb{C}^{m,m}$ ($P \in \mathbb{R}^{n,n}$, $Q \in \mathbb{R}^{m,m}$), so dass

$$P^H A Q = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_m & \\ & & & 0 \end{bmatrix} \quad \text{mit } \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m \geq 0.$$

Beweis: Seien $x \in \mathbb{C}^m$, $y \in \mathbb{C}^n$ ($x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$), mit $\|x\|_2 = \|y\|_2 = 1$ und $Ax = \sigma y$, wobei $\sigma = \|A\|_2$. (Das kann man natürlich immer erreichen.)

Sei $P = [y \ P_1] \in \mathbb{C}^{n,n}$, $Q = [x \ Q_1] \in \mathbb{C}^{m,m}$ unitär (Ergänzung zu ONB), so folgt

$$A_1 = P^H A Q = \begin{bmatrix} \sigma & w^H \\ 0 & A_2 \end{bmatrix} \begin{matrix} 1 \\ n-1 \end{matrix} \begin{matrix} \\ 1 \\ m-1 \end{matrix}.$$

Da

$$\left\| A_1 \begin{bmatrix} \sigma \\ w \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \sigma^2 + w^H w \\ A_2 w \end{bmatrix} \right\|_2^2 \geq \sigma^2 + w^H w,$$

so folgt

$$\|A_1\|_2 = \sup_{\substack{z \neq 0 \\ z \in \mathbb{C}^m}} \frac{\|A_1 z\|_2}{\|z\|_2} \implies \|A_1\|_2^2 \geq \sigma^2 + w^H w.$$

Aber

$$\sigma^2 = \|A\|_2^2 = \|A_1\|_2^2 \implies w^H w = 0 \implies w = 0.$$

Der Rest folgt mit Induktion.

□

Definition 5.5 Sei $A \in \mathbb{C}^{n,m}(\mathbb{R}^{n,m})$, $n \geq m$. Die Zerlegung in

$$A = P \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_m & \\ & & 0 & \end{bmatrix} Q^H,$$

mit P, Q unitär (orthogonal), heißt Singulärwertzerlegung von A , die σ_i heißen Singulärwerte und die Spalten von P (Q) Linkssingulärvektoren (Rechtssingulärvektoren). (Matlab: SVD)
(Für $m \geq n$ gilt alles mit der Singulärwertzerlegung von A^H).

Satz 5.6 Sei $A \in \mathbb{C}^{n,m}(\mathbb{R}^{n,m})$, $n \geq m$, und sei mit $P = [p_1, \dots, p_n]$, $Q = [q_1, \dots, q_m]$

$$A = P \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_m & \\ & & 0 & \end{bmatrix} Q^H$$

mit $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_m = 0$ die Singulärwertzerlegung von A , dann gilt

- (i) $r = \text{Rang}(A)$,
- (ii) $\text{Kern}(A) = \text{span}\{q_{r+1}, \dots, q_m\}$,
- (iii) $\text{Bild}(A) = \text{span}\{p_1, \dots, p_r\}$,
- (iv) $\|A\|_2 = \sigma_1$,
- (v) $\|A\|_F^2 = \sigma_1^2 + \dots + \sigma_m^2$.

Beweis:

- (i) Die Behauptung folgt aus den Sätzen 11.5, 4.2 Teil I.
- (ii) Die Spalten von Q bilden ONB von \mathbb{C}^m ,

$$A[q_{r+1}, \dots, q_m] = 0 \implies \{q_{r+1}, \dots, q_m\} = \text{Kern}(A)$$

- (iii) Für das Bild gilt die Betrachtung analog.
- (iv) Nach Konstruktion

$$\|P^H A Q\|_2 = \left\| \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_m & \\ & & 0 & \end{bmatrix} \right\|_2 = \sup_{\substack{x \in \mathbb{C}^n \\ x \neq 0}} \frac{\|[\sigma_i x_i]\|_2}{\|x\|_2} = \sigma_1 \quad \text{mit } x = e_1.$$

$$(v) \|A\|_F^2 = \|P^H A Q\|_F^2 = \sigma_1^2 + \dots + \sigma_m^2.$$

□

Bem.: Die Singulärwerte sind die Wurzeln der Eigenwerte von $A^H A$ bzw. AA^H .

Mit der Singulärwertzerlegung erhalten wir das Analogon zur Normalform unter Äquivalenz. Allerdings bietet die Singulärwertzerlegung neben der Tatsache, dass man sie numerisch sehr gut ausrechnen kann, auch noch weitere wichtige Eigenschaften.

Wir kommen nochmal auf das allgemeine lineare Gleichungssystem

$$Ax = b \quad \text{mit} \quad A \in \mathbb{C}^{n,m} \quad \text{und} \quad b \in \mathbb{C}^n, x \in \mathbb{C}^m \quad (A \in \mathbb{R}^{n,m}, b \in \mathbb{R}^n, x \in \mathbb{R}^m). \quad (5.7)$$

Wir setzen jetzt aber nicht voraus, dass $b \in \text{Bild}(A)$. Damit ist $Ax = b$ aber nicht mehr unbedingt lösbar. Wir lösen statt dessen

$$\|Ax - b\| = \min, \quad (5.8)$$

und wie schon vorher im Fall der lösbaren Gleichungssysteme fordern wir

$$\|x\| = \min, \quad (5.9)$$

damit die Lösung eindeutig wird.

Satz 5.10 *Betrachte die Minimierungsaufgabe $\|Ax - b\|_2 = \min$ unter der Zusatzforderung $\|x\|_2 = \min$. Sei $\text{Rang } A = r$. Diese Aufgabe hat eine eindeutige Lösung, die mittels der Singulärwertzerlegung von $A = P\Sigma Q^H = P \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} Q^H$, mit Σ_r nichtsingulär, wie folgt angegeben werden kann:*

$$\hat{x} = Q \begin{bmatrix} \Sigma_r^{-1} [I_r \ 0] P^H b \\ 0 \end{bmatrix}. \quad (5.11)$$

Das Residuum ist gegeben durch

$$\|A\hat{x} - b\|_2 = \|[0 \ I_{n-r}] P^H b\|_2.$$

Beweis:

$$\|Ax - b\|_2 = \min \quad \text{und} \quad \|x\|_2 = \min \Leftrightarrow \|P^H A Q Q^H x - P^H b\|_2 = \min \quad \text{und} \quad \|Q^H x\|_2 = \min.$$

Das Gleichungssystem $P^H A Q Q^H x = P^H b$ hat die Form

$$\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} [I_r \ 0] P^H b \\ [0 \ I_{n-r}] P^H b \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} [I_r \ 0] Q^H x \\ [0 \ I_{m-r}] Q^H x \end{bmatrix} \quad (5.12)$$

Dessen Minimallösung lautet

$$\begin{aligned} y_1 &= \Sigma_r^{-1} b_1 = \Sigma_r^{-1} [I_r \ 0] P^H b, \\ y_2 &= 0, \\ \Rightarrow \hat{x} &= Q \begin{bmatrix} \Sigma_r^{-1} [I_r \ 0] P^H b \\ 0 \end{bmatrix}. \end{aligned}$$

Damit haben wir eine Lösung. Und da die Minimallösung von (5.12) eindeutig ist, ist auch die Lösung (5.11) eindeutig. Für das Residuum gilt:

$$\|A\hat{x} - b\|_2 = \|P^H A Q Q^H \hat{x} - P^H b\|_2 = \|b\|_2 = \|[0 \ I_{n-r}] P^H b\|_2$$

□

Beispiel 5.13

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \text{Singularwertzerlegung von } A = P \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} Q^H.$$

$$P^H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P^H A = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad P^H b = \begin{bmatrix} \frac{3}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 3 \end{bmatrix},$$

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \Sigma = P^H A Q = \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad y = Q^H x,$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 3 \end{bmatrix} \implies y_1 = \frac{3}{2\sqrt{2}}, \quad y_2 = 0,$$

$$x = Qy = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{3}{2\sqrt{2}} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} \\ \frac{3}{4} \end{bmatrix}.$$

Residuum:

$$\|Ax - b\|_2 = \left\| \begin{bmatrix} \frac{3}{2} \\ \frac{3}{2} \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -3 \end{bmatrix} \right\|_2 = \sqrt{\frac{1}{4} + \frac{1}{4} + 9} = \sqrt{\frac{19}{2}}.$$

Definition 5.14 Sei $A \in \mathbb{C}^{n,m}$ ($\mathbb{R}^{n,m}$) und sei $A = P \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} Q^H = P \Sigma Q^H$ mit $\Sigma_r = \text{diag}(\sigma_1, \dots, \sigma_r)$ die Singularwertzerlegung von A . Die Matrix

$$A^+ = Q \Sigma^+ P^H \quad \text{mit} \quad \Sigma^+ = \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{m,n}$$

heißt Moore–Penrose–Inverse von A .

Satz 5.15 Sei $A \in \mathbb{C}^{n,m}$ ($\mathbb{R}^{n,m}$).

- (i) Falls $\text{Rang}(A) = m$, so gilt $A^+ = (A^H A)^{-1} A^H$.
- (ii) Falls $n = m$ und A nichtsingulär, so gilt $A^+ = A^{-1}$.

(iii) A^+ ist die eindeutige Matrix X , welche die vier Moore–Penrose-Gleichungen erfüllt:

- (a) $AXA = A$,
- (b) $XAX = X$,
- (c) $(AX)^H = AX$,
- (d) $(XA)^H = XA$.

Beweis: $A = P\Sigma Q^H$, $\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$.

(i) Da $\text{Rang}(A) = m$, d.h., $A = \begin{bmatrix} \Sigma_m \\ 0 \end{bmatrix}$, so folgt $A^+ = Q\Sigma^+P^H = Q[\Sigma_m^{-1} \ 0]P^H$.

$$\begin{aligned} (A^H A)^{-1} A^H &= \left(Q[\Sigma_m \ 0]P^H P \begin{bmatrix} \Sigma_m \\ 0 \end{bmatrix} Q^H \right)^{-1} \cdot Q[\Sigma_m \ 0]P^H \\ &= Q\Sigma_m^{-2}Q^H Q[\Sigma_m \ 0]P^H = Q[\Sigma_m^{-1} \ 0]P^H. \end{aligned}$$

(ii) Klar.

(iii) $Y = Q^H X P$

(a) $P^H A Q \quad Q^H X P \quad P^H A Q = P^H A Q$,

$$\begin{aligned} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \\ \implies \Sigma_r Y_{11} \Sigma_r = \Sigma_r &\implies Y_{11} = \Sigma_r^{-1}. \end{aligned}$$

(b) $\begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$,

$$\begin{aligned} \implies Y_{11} \Sigma_r Y_{11} &= Y_{11}, \\ Y_{21} \Sigma_r Y_{11} &= Y_{21}, \\ Y_{11} \Sigma_r Y_{12} &= Y_{12}, \\ Y_{21} \Sigma_r Y_{12} &= Y_{22}. \end{aligned}$$

(c) $\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$ Hermite'sch $\implies \Sigma_r Y_{12} = 0$, $\Sigma_r Y_{11} = Y_{11}^H \Sigma_r^H$.

(d) $\begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$ Hermite'sch $\implies Y_{21} \Sigma_r = 0$, $(Y_{11} \Sigma_r)^H = Y_{11} \Sigma_r$.

Alle 4 Gleichungen zusammen ergeben

- (c) $\implies Y_{12} = 0$,
- (d) $\implies Y_{21} = 0$,
- (b) $\implies Y_{22} = 0$,
- (a) $\implies Y_{11} = \Sigma_r^{-1}$.

Lösung ist eindeutig \implies Behauptung. □

Kapitel 6

Die Hauptachsentransformation

Nachdem wir mit der Singulärwertzerlegung bereits eines der wichtigsten praktischen Werkzeuge der linearen Algebra kennengelernt haben, welches uns wichtige Eigenschaften wie Rang, Kern, Bild einer Matrix erzeugt, kommen wir nun zu weiteren wichtigen Transformationen. Bisher haben wir Äquivalenztransformationen PAQ mit P, Q nichtsingulär, orthogonal oder unitär betrachtet und $P^{-1}AP$ mit P nichtsingulär, orthogonal oder unitär. Im Fall P orthogonal ist das P^TAP .

Was passiert nun aber, wenn P nicht orthogonal ist und wir P^TAP betrachten? Wir haben schon gesehen, dass dies die Form ist, die ein Basiswechsel bei einer Bilinearform hat, siehe Lemma 3.5. Im Komplexen betrachten wir natürlich analog P^HAP .

Definition 6.1 Sei K ein Körper. Seien $A, B \in K^{n,n}$. A und B heißen kongruent, falls es eine nichtsinguläre Matrix $P \in K^{n,n}$ gibt, so dass

$$A = P^TBP.$$

Zu jeder Matrix können wir sofort eine Bilinearform konstruieren und umgekehrt.

Sei K ein beliebiger Körper, V ein Vektorraum und $\{v_1, \dots, v_n\}$ eine Basis von V . Ist $\alpha : V \times V \rightarrow K$ eine Bilinearform, so setze $A = [a_{ij}]$ mit $a_{ij} = \alpha(v_j, v_i)$, $i, j = 1, \dots, n$. So ist A die Matrixdarstellung von α bezüglich der Basis $\{v_1, \dots, v_n\}$,

Umgekehrt sei $A = [a_{ij}] \in K^{n,n}$, so definieren wir auf V eine Bilinearform $\alpha : V \times V \rightarrow K$ wie folgt: Seien $v, v' \in V$ und sei

$$v = \sum_{i=1}^n b_i v_i, \quad v' = \sum_{i=1}^n c_i v_i \quad \text{und} \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}, \quad c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix},$$

so setze $\alpha(v, v') := c^T A b = \sum_{i,j=1}^n c_i a_{ij} b_j$.

Beachte, diese Bijektion zwischen Matrizen und Bilinearformen hängt natürlich von der Basis ab. Wir hatten schon in Lemma 3.5 gesehen, dass für zwei Basen $\{v_1, \dots, v_n\}, \{v'_1, \dots, v'_n\}$ von V mit Basisübergangsmatrix P , die Matrixdarstellung von einer Bilinearform $\alpha : V \times V \rightarrow K$ bezüglich dieser beiden Basen durch Kongruenztransformation gegeben ist.

Falls $v, w \in V$, $v = \sum_{i=1}^n b_i v_i = \sum_{i=1}^n b'_i v'_i$ und $w = \sum_{i=1}^n c_i v_i = \sum_{i=1}^n c'_i v'_i$, so ist

$$b := \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = P \begin{bmatrix} b'_1 \\ \vdots \\ b'_n \end{bmatrix} =: Pb'$$

$$c := \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = P \begin{bmatrix} c'_1 \\ \vdots \\ c'_n \end{bmatrix} =: Pc'$$

und es ergibt sich

$$\alpha(v, w) = c^\top Ab = (Pc')^\top APb' = c'^\top P^\top APb'.$$

Beachte, wir betrachten Bilinearformen (auch falls $K = \mathbb{C}$.)

Wir hatten auch bereits gesehen, dass die Matrixdarstellung von α symmetrisch ist genau dann, wenn α symmetrisch ist, und dass es in diesem Fall eine Orthonormalbasis aus Eigenvektoren gibt.

Beispiel 6.2 $V = \mathbb{R}^2, \{v_1, v_2\} = \{e_1, e_2\}$,

$$A = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix},$$

$$\alpha\left(\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}\right) = \alpha(b_1 e_1 + b_2 e_2, c_1 e_1 + c_2 e_2) = c^\top Ab$$

$$= 4b_1 c_1 + b_2 c_1 + b_1 c_2 + 2b_2 c_2.$$

A hat das charakteristische Polynom $P_A(\lambda) = \lambda^2 - 6\lambda + 7$.

Eigenwerte:

$$\lambda_1, \lambda_2 = 3 \pm \sqrt{9-7} = 3 \pm \sqrt{2}.$$

Eigenvektoren:

$$A - \lambda_1 I = \begin{bmatrix} 4 - 3 + \sqrt{2} & 1 \\ 1 & 2 - 3 + \sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 + \sqrt{2} & 1 \\ 1 & -1 + \sqrt{2} \end{bmatrix}.$$

$$Q^\top = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}, \quad c = \frac{1 + \sqrt{2}}{\sqrt{(1 + \sqrt{2})^2 + 1^2}}, \quad s = \frac{1}{\sqrt{(1 + \sqrt{2})^2 + 1^2}},$$

$$Q^\top (A - \lambda_1 I) Q = \begin{bmatrix} c(1 + \sqrt{2}) + s & c + s(-1 + \sqrt{2}) \\ -s(1 + \sqrt{2}) + c & -s + c(-1 + \sqrt{2}) \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

$$\begin{aligned}
&= \begin{bmatrix} \frac{(1+\sqrt{2})^2+1}{\sqrt{(1+\sqrt{2})^2+1}} & \frac{1+\sqrt{2}-1+\sqrt{2}}{\sqrt{(1+\sqrt{2})^2+1}} \\ 0 & \frac{-1+2-1}{\sqrt{(1+\sqrt{2})^2+1}} \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \\
&= \frac{1}{\sqrt{(1+\sqrt{2})^2+1}} \begin{bmatrix} (1+\sqrt{2})^2+1 & 2\sqrt{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \\
&= \frac{1}{\left(\sqrt{(1+\sqrt{2})^2+1}\right)^2} \begin{bmatrix} ((1+\sqrt{2})^2+1)(1+\sqrt{2})+\sqrt{2} & (1+\sqrt{2})2\sqrt{2}-1-(1+\sqrt{2})^2 \\ 0 & 0 \end{bmatrix} \\
&= \frac{1}{(1+\sqrt{2})^2+1} \begin{bmatrix} (4+2\sqrt{2})(1+\sqrt{2})+2\sqrt{2} & 2\sqrt{2}+4-1-1-2\sqrt{2}-2 \\ 0 & 0 \end{bmatrix} \\
&= \frac{1}{(1+\sqrt{2})^2+1} \begin{bmatrix} 8+8\sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} \\
\implies Q^T A Q &= \begin{bmatrix} 3+\sqrt{2} & 0 \\ 0 & 3-\sqrt{2} \end{bmatrix}.
\end{aligned}$$

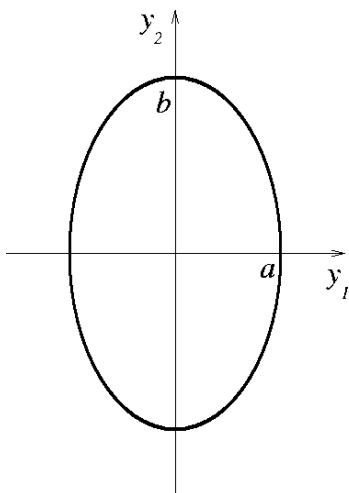
Was für eine Menge stellt

$$E = \left\{ x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \mid x^T A x - 1 = 0 \right\} = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \mid 4x_1^2 + 2x_1x_2 + 2x_2^2 = 1 \right\}$$

dar? — Dies ist eine Ellipse.

Wenn wir zu dem Koordinatensystem übergehen, welches durch die ONB $\left\{ \begin{bmatrix} c \\ s \end{bmatrix}, \begin{bmatrix} -s \\ c \end{bmatrix} \right\}$ gegeben ist, so ist

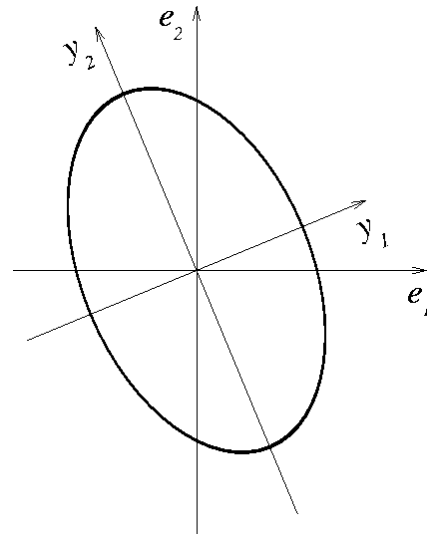
$$\begin{aligned}
E &= \left\{ y = Q^T x \in \mathbb{R}^2 \mid y^T \begin{bmatrix} 3+\sqrt{2} & 0 \\ 0 & 3-\sqrt{2} \end{bmatrix} y = 1 \right\} \\
&= \left\{ y \in \mathbb{R}^2 \mid y_1^2(3+\sqrt{2}) + y_2^2(3-\sqrt{2}) = 1 \right\}, \quad \text{d.h., die Ellipse}
\end{aligned}$$



$$E = \left\{ y \in \mathbb{R}^2 \mid \frac{y_1^2}{a^2} + \frac{y_2^2}{b^2} = 1 \right\}, \quad \text{mit}$$

$$a = \sqrt{\frac{1}{3+\sqrt{2}}}, \quad b = \sqrt{\frac{1}{3-\sqrt{2}}}.$$

Im kanonischen Koordinatensystem mit der Basis $\{e_1, e_2\}$ ist das also die Ellipse



Definition 6.3 Sei V ein euklidischer Raum und $\alpha : V \times V \rightarrow \mathbb{R}$ eine symmetrische Bilinearform. Dann heißt α positiv definit, falls $\alpha(v, v) > 0, \forall v \in V, v \neq 0$, und positiv semidefinit, falls $\alpha(v, v) \geq 0 \forall v \in V$.

Satz 6.4 Sei V ein endlichdimensionaler euklidischer Raum und $\alpha : V \times V \rightarrow \mathbb{R}$ eine symmetrische Bilinearform.

- (i) α ist positiv definit genau dann, wenn es eine Basis $\{v_1, \dots, v_n\}$ von V gibt, so dass die Matrixdarstellung von α bezüglich $\{v_1, \dots, v_n\}$ nur positive reelle Eigenwerte hat.
- (ii) α ist positiv semidefinit genau dann, wenn es eine Basis $\{v_1, \dots, v_n\}$ von V gibt, so dass die Matrixdarstellung von α bezüglich $\{v_1, \dots, v_n\}$ nur nichtnegative reelle Eigenwerte hat.
- (iii) Es gibt eine Orthonormalbasis $\{v_1, \dots, v_n\}$ von V , so dass die Matrixdarstellung von α bezüglich dieser Basis diagonal ist.

Beweis: Wir beweisen (i) und (ii) zusammen.

Sei α positiv (semi)definit und $A = [a_{ij}]$ die Matrixdarstellung von α bezüglich $\{v_1, \dots, v_n\}$, d.h., $a_{ij} = \alpha(v_j, v_i)$. A ist reell symmetrisch $\implies \exists$ Orthonormalbasis aus Eigenvektoren von A , $P = [p_1, \dots, p_n]$, und bezüglich der Basis $\{v_1, \dots, v_n\} \cdot P$ hat α die Matrixdarstellung

$$P^\top AP = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Ist w der Koordinatenvektor von $z \in V$, in der Basis $\{v_1, \dots, v_n\} \cdot P$, dann gilt

$$\alpha(z, z) = w^\top P^\top AP w = \sum_{i=1}^n w_i^2 \lambda_i.$$

Falls es $\lambda_i \leq 0$ ($\lambda_i < 0$) gibt, so betrachte $w = e_i$ und es gilt $\alpha(z, z) = \lambda_i \leq 0$ (< 0) und das ist ein Widerspruch zur positiven Definitheit (Semidefinitheit).

Für die andere Richtung: Ist x der Koordinatenvektor von $z \in V$ in der Basis $\{v_1, \dots, v_n\}$, dann ergibt sich aus

$$\alpha(z, z) = x^\top A x = x^\top P \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} P^\top x = \sum_{i=1}^n x_i^2 \lambda_i > 0 \quad (\geq 0)$$

sofort, dass α positiv definit (semidefinit) ist.

(iii) ist klar. Man nimmt einfach eine Matrixdarstellung bezüglich irgendeiner Basis und erhält A . Mit der Orthonormalbasis aus Eigenvektoren ergibt sich die Behauptung. \square

Beachte: Wir haben sogar bewiesen, dass die Matrixdarstellung einer positiv (semi-)definiten, symmetrischen Bilinearform bezüglich jeder Basis von V nur positive (nicht-negative) reelle Eigenwerte hat.

Wir kommen nun zu der Frage, welches die Invarianten unter der Transformation $P^\top A P$, angewandt auf symmetrische Matrizen A , sind. Es ist klar, dass für P orthogonal dies die Eigenwerte sind, aber für P nichtorthogonal sind dies nur die Anzahlen der Eigenwerte $< 0, = 0, > 0$.

Satz 6.5 (Trägheitssatz von Sylvester)

Sei V ein endlichdimensionaler euklidischer Raum und $\alpha : V \times V \rightarrow \mathbb{R}$ eine symmetrische Bilinearform. Seien $\{v_1, \dots, v_n\}, \{v'_1, \dots, v'_n\}$ Basen von V und A, A' Matrixdarstellungen von α bezüglich dieser beiden Basen.

Seien (π, ν, ω) und (π', ν', ω') die Anzahl der Eigenwerte ($> 0, < 0, = 0$) von A bzw. A' . Dann gilt $\pi = \pi', \nu = \nu'$ und $\omega = \omega'$ und $\text{Rang}(A) = \text{Rang}(A')$.

Beweis: Es gibt Q, Q' orthogonal, so dass

$$Q^\top A Q = D, \quad Q'^\top A' Q' = D'$$

und die Diagonalelemente von D, D' seien geordnet als

$$D = \begin{bmatrix} D_1 & & \\ & D_2 & \\ & & 0 \end{bmatrix}, \quad D' = \begin{bmatrix} D'_1 & & \\ & D'_2 & \\ & & 0 \end{bmatrix},$$

wobei D_1 und D'_1 die positiven Eigenwerte von A und A' enthalten und D_2 und D'_2 die negativen.

Da dies mit einer Permutation immer möglich ist, können wir gleich Q, Q' so wählen, dass das gilt. Wir wissen, dass $A' = P^\top A P$, also gilt

$$D' = \tilde{P}^\top D \tilde{P} \quad \text{mit} \quad \tilde{P} = Q^\top P Q'$$

und außerdem $\text{Rang}(A) = \text{Rang}(D) = \text{Rang}(D') = \text{Rang}(A')$.

Da A' und A symmetrisch sind, ist (mit der Singulärwertzerlegung) $\text{Rang}(A) = \pi + \nu$ und $\text{Rang}(A') = \pi' + \nu'$. Es reicht also zu zeigen, dass $\pi = \pi'$, dann gilt auch $\nu = \nu'$.

Wir erhalten aus der Diagonalisierung Räume V_π, V_ν, V_ω bzw. $V'_{\pi'}, V'_{\nu'}, V'_{\omega'}$, die aus den ONB Q, Q' gebildet werden. Wir zeigen nun, dass

$$V_\pi \cap (V'_{\nu'} \oplus V'_{\omega'}) = \{0\}.$$

Dazu sei v aus $V_\pi \cap (V'_{\nu'} \oplus V'_{\omega'})$, so gilt $\alpha(v, v) \geq 0$ und $\alpha(v, v) \leq 0$, also $v = 0$.

Mit dem Dimensionssatz (Teil I, Satz 8.26) folgt

$$\dim V_\pi + \dim V'_{\nu'} + \dim V'_{\omega'} \leq n.$$

Aber da

$$\dim V'_{\pi'} + \dim V'_{\nu'} + \dim V'_{\omega'} = n,$$

folgt:

$$\dim V_\pi \leq \dim V'_{\pi'} \implies \pi \leq \pi'.$$

Mit $V'_{\pi'} \cap [V_\nu \oplus V_\omega]$ folgt $\pi' \leq \pi$.

Also $\pi' = \pi$ und dann $\nu' = \nu$. □

Definition 6.6 *Das zu einer symmetrischen Bilinearform α gehörende Tripel*

$$(\pi, \nu, \omega)$$

heißt Trägheitsindex von α .

Dieser Begriff stammt ursprünglich aus der Mechanik (Beschreibung von Trägheitsmomenten). Ein direktes Korollar aus Satz 6.5 ist die Normalform unter Kongruenz.

Korollar 6.7 *Sei V ein endlichdimensionaler euklidischer Raum der Dimension n und α eine symmetrische Bilinearform von V , so gibt es eine Basis $\{v_1, \dots, v_n\}$ von V , so dass die Matrixdarstellung von α bezüglich dieser Basis die Form hat:*

$$\begin{bmatrix} I_\pi & & \\ & -I_\nu & \\ & & 0 \end{bmatrix} \in \mathbb{R}^{n,n}.$$

Beweis: Wir wissen schon, dass π, ν, ω Invarianten unter Kongruenz sind und dass es P orthogonal gibt, so dass

$$P^\top AP = \begin{bmatrix} D_1 & & \\ & D_2 & \\ & & 0 \end{bmatrix} \begin{matrix} \} \pi \\ \} \nu \\ \} \omega \end{matrix}. \quad \text{Setze } \tilde{P} = P \cdot \begin{bmatrix} D_1^{-\frac{1}{2}} & & \\ & (-D_2)^{-\frac{1}{2}} & \\ & & I_\omega \end{bmatrix},$$

so folgt

$$\tilde{P}^\top A \tilde{P} = \begin{bmatrix} D_1^{-\frac{1}{2}} D_1 D_1^{-\frac{1}{2}} & & \\ & (-D_2)^{-\frac{1}{2}} D_2 (-D_2)^{-\frac{1}{2}} & \\ & & 0 \end{bmatrix} = \begin{bmatrix} I_\pi & & \\ & -I_\nu & \\ & & 0 \end{bmatrix}.$$

□

Man beachte, dass die Basis, d.h. die Transformationsmatrix, nicht eindeutig ist, aber die Form $\begin{bmatrix} I_\pi & & \\ & -I_\nu & \\ & & 0 \end{bmatrix}$.

Definition 6.8 Sei $A \in \mathbb{R}^{n,n}$ symmetrisch, so heißt A positiv definit (positiv semidefinit), falls

$$x^\top A x > 0 \ (\geq 0), \quad \forall x \in \mathbb{R}^n \setminus \{0\}.$$

A heißt negativ definit (negativ semidefinit), falls

$$x^\top A x < 0 \ (\leq 0), \quad \forall x \in \mathbb{R}^n \setminus \{0\}.$$

Korollar 6.9 Sei $A \in \mathbb{R}^{n,n}$ symmetrisch.

(a) Die folgenden Aussagen sind äquivalent.

- (i) A ist positiv definit.
- (ii) Alle Eigenwerte von A sind positiv.
- (iii) Der Trägheitsindex von A ist $(n, 0, 0)$.

(b) Die folgenden Aussagen sind äquivalent.

- (i) A ist positiv semidefinit.
- (ii) Alle Eigenwerte von A sind nichtnegativ.
- (iii) Der Trägheitsindex von A ist $(\pi, 0, \omega)$.

Beweis: Der Beweis ist Übungsaufgabe.

□

Kapitel 7

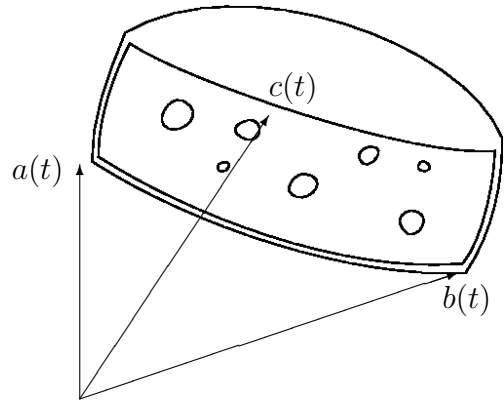
Bewegung starrer Körper

Als Beispiel wollen wir nun die Bewegung eines starren Körpers betrachten.

Markiere einen seiner Punkte und beschreibe dessen Bahn durch die Funktion $a(t) (t \geq 0)$. Da der Körper starr ist, ändern die anderen Punkte ihren Abstand zu a nicht, so dass für alle Punkte $x(t)$ gilt

$$x(t) - a(t) = Q(t) (x(0) - a(0))$$

mit einer t -abhängigen orthogonalen Matrix Q .



Mit der Kenntnis der Bahnbewegung zweier weiterer Punkte $b(t), c(t)$ (die zu $a(t)$ in allgemeiner Lage liegen) kann man dann $Q(t)$ bestimmen.

$$\begin{aligned} x_1(t) &= b(t) - a(t) = Q(t) (b(0) - a(0)) \\ x_2(t) &= c(t) - a(t) = Q(t) (c(0) - a(0)) \\ x_3(t) &= (b(t) - a(t)) \times (c(t) - a(t)) \\ &= (Q(t) (b(0) - a(0))) \times (Q(t) (c(0) - a(0))). \end{aligned}$$

Die Vektoren sind linear unabhängig.

Da Q orthogonal und $x_3(0) \perp x_1(0), x_2(0)$, gilt auch

$$Q(t)x_3(0) \perp Q(t)x_1(0), Q(t)x_2(0) \implies Q(t)x_3(0) = \lambda(Q(t)x_1(0) \times Q(t)x_2(0)).$$

Da sich die Orientierung des Körpers nicht ändern kann, gilt $\lambda > 0$ und (siehe Teil I)

$$\begin{aligned} \|(Qx_1) \times (Qx_2)\|_2^2 &= \|Qx_1\|_2^2 \|Qx_2\|_2^2 - \langle Qx_1, Qx_2 \rangle^2 \\ &= \|x_1\|_2^2 \|x_2\|_2^2 - \langle x_1, x_2 \rangle^2 = \|x_3\|_2^2 \\ &= \|Qx_3\|_2^2 \implies \lambda = 1. \end{aligned}$$

Damit haben wir gezeigt, dass auch für x_3 die Bewegung durch $x_3(t) = Q(t)x_3(0)$ beschrieben wird.

Da $\det[Qx_1, Qx_2, Qx_3] = \det Q \cdot \det[x_1, x_2, x_3]$, aber $\det Q \in \{+1, -1\}$, bedeutet die Erhaltung der Orientierung also: $\det Q = +1$.

$$Q(t) \in \mathcal{SO}(3) = \{Q \in \mathcal{O}(3) \mid \det Q = 1\}.$$

Man beachte, dass $a(0), a(t)$ nicht eindeutig bestimmt sind, aber $Q(t)$. Für die Geschwindigkeit $\dot{x} = \frac{d}{dt}x(t)$ eines Punktes $x(t)$ gilt

$$\dot{x}(t) = \dot{a}(t) + \dot{Q}(t)(x(0) - a(0)).$$

Was wissen wir über die Ableitung einer orthogonalen Matrix?

$$\begin{aligned} Q^\top Q = I &\implies \dot{Q}^\top Q + Q^\top \dot{Q} = 0 \implies Q^\top \dot{Q} \text{ ist schiefsymmetrisch,} \\ QQ^\top = I &\implies \dot{Q}Q^\top + Q\dot{Q}^\top = 0 \implies \dot{Q}Q^\top \text{ ist schiefsymmetrisch.} \end{aligned}$$

Nun gilt aber

$$\begin{aligned} \dot{Q}(x(0) - a(0)) &= \dot{Q}Q^\top Q(x(0) - a(0)), \\ \dot{x} &= \dot{a} + \dot{Q}Q^\top(Q(x(0) - a(0))) = \dot{a} + \dot{Q}Q^\top(x(t) - a(t)). \end{aligned}$$

$\dot{Q}Q^\top$ ist schiefsymmetrisch. Sei also

$$\dot{Q}Q^\top = \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix}, \quad \text{so ist}$$

$$\dot{Q}Q^\top(x(t) - a(t)) = \underbrace{\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}}_{\omega(t)} \times (x(t) - a(t)) \quad (\text{Vektorprodukt in } \mathbb{R}^3).$$

$$\implies \dot{x} = \dot{a} + \omega(t) \times (x(t) - a(t)).$$

Die Momentanbewegung eines starren Körpers ist zusammengesetzt aus der Translation des Punktes $a(t)$ mit der Geschwindigkeit $\dot{a}(t)$ und einer Rotation um die durch $\omega(t)$ definierte Achse durch $a(t)$ mit der Winkelgeschwindigkeit $\|\omega(t)\|_2$.

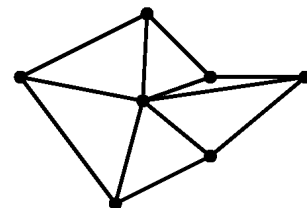
$\omega(t)$ ist eindeutig, \dot{a} und die Zerlegung von \dot{x} in \dot{a} und ω sind nicht eindeutig.

Wir wählen nun als $a(t)$ die Bahn des Schwerpunktes des Körpers

$$a(t) = \sum_{i=1}^N \frac{m_i}{m} y_i(t)$$

wobei y_i die Orte einzelner Massenpunkte sind, m_i ihre Massen und $m = \sum_{i=1}^N m_i$ die Gesamtmasse.

(Im nichtdiskreten Fall sind die Summen durch Volumenintegrale zu ersetzen).



Setze $v = \|\dot{a}\|_2$, $v_i = \|\dot{y}_i\|_2$, $r_i = y_i - a$, $\varrho_i = \|r_i\|_2$.

⇒ Kinetische Energie des Körpers:

$$\begin{aligned} T &= \sum_{i=1}^N \frac{1}{2} m_i v_i^2 = \sum_{i=1}^N \frac{1}{2} m_i \|\dot{a} + \omega \times r_i\|_2^2 \\ &= \sum_{i=1}^N \left(\frac{m_i}{2} v^2 + m_i \langle \dot{a}, \omega \times r_i \rangle + \frac{1}{2} m_i \|\omega \times r_i\|_2^2 \right) \\ &= \frac{m}{2} v^2 + \underbrace{\langle \dot{a}, \omega \times \sum_{i=1}^N m_i r_i \rangle}_0 + \sum_{i=1}^N \frac{1}{2} m_i \left(\|\omega\|_2^2 \varrho_i^2 - \langle \omega, r_i \rangle^2 \right) \end{aligned}$$

Nun ist aber $\sum_{i=1}^N m_i r_i = 0$ und mit $r_i = \begin{bmatrix} r_{i1} \\ r_{i2} \\ r_{i3} \end{bmatrix}$, $\omega = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}$ folgt

$$\begin{aligned} \|\omega\|_2^2 \varrho_i^2 - \langle \omega, r_i \rangle^2 &= (\omega_1^2 + \omega_2^2 + \omega_3^2) \varrho_i^2 - (\omega_1 r_{i1} + \omega_2 r_{i2} + \omega_3 r_{i3})^2 \\ &= \sum_{k,l=1}^3 \omega_k \omega_l (\varrho_i^2 \delta_{kl} - r_{ik} r_{il}). \end{aligned}$$

Setze $\Theta = [\Theta_{kl}] \in \mathbb{R}^{3,3}$ mit

$$\Theta_{k,l} = \sum_{i=1}^N m_i \{ \varrho_i^2 \delta_{kl} - r_{ik} r_{il} \},$$

so folgt

$$T = \frac{m}{2} v^2 + \frac{1}{2} \omega^\top \Theta \omega.$$

$\frac{m}{2} v^2$ heißt Translationsenergie,

$\frac{1}{2} \omega^\top \Theta \omega$ heißt Rotationsenergie.

Die Normierung der Bilinearform Θ als

$$\Theta_e = \frac{\omega^\top \Theta \omega}{\|\omega\|_2^2}$$

heißt das auf die Achse $\frac{\omega}{\|\omega\|_2}$ bezogene Trägheitsmoment des Körpers.

Θ ist symmetrisch und (falls nicht alle Massenpunkte auf einer Geraden liegen) positiv definit. Die Gleichung $\omega^\top \Theta \omega = 1$ beschreibt ein Ellipsoid, das Trägheitsellipsoid, das die bei fester Rotationsenergie zu den verschiedenen Richtungen gehörenden Winkelgeschwindigkeiten veranschaulicht. Die Hauptachsen dieses Ellipsoids sind Achsen, um die der Körper reine Rotationsbewegungen ausführen kann.

Dazu kommt dann noch die Kreisbewegung, die behandeln wir hier nicht.

Kapitel 8

Quadriken

Wir wollen nun einen Bezug zur Geometrie herstellen und damit die Klassifikation von geometrischen Objekten im \mathbb{R}^2 und \mathbb{R}^3 vornehmen. Sei K ein Körper. Ein quadratisches Polynom in n Variablen ist ein Ausdruck der Form

$$P(x_1, \dots, x_n) = \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j + \sum_{1 \leq i \leq n} b_i x_i + c, \quad (8.1)$$

in dem nicht alle a_{ij} verschwinden. P ist eine nichtlineare Abbildung

$$P : K^n \rightarrow K.$$

Definition 8.2 *Eine Teilmenge $Q \subset K^n$ heißt Quadrik oder Hyperfläche zweiter Ordnung, falls es ein quadratisches Polynom gibt, so dass*

$$Q = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in K^n \mid P(x_1, \dots, x_n) = 0 \right\}.$$

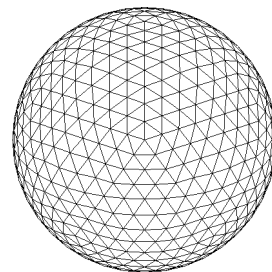
Beispiel 8.3

(1) $K = \mathbb{R}$, $n = 3$

$$P(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 1,$$

$$Q = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 - 1 = 0 \right\}$$

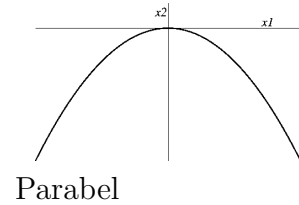
ist gerade die Oberfläche einer Kugel mit Mittelpunkt $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ und Radius 1.



(2) $K = \mathbb{R}, n = 2$

$$P(x_1, x_2) = x_1^2 + 2x_2$$

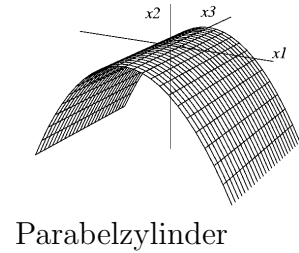
$$Q = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \mid x_1^2 + 2x_2 = 0 \right\}$$



(3) $K = \mathbb{R}, n = 3$

$$P(x_1, x_2, x_3) = x_1^2 + 2x_2$$

$$Q = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_1^2 + 2x_2 = 0 \right\}$$



Im folgenden betrachten wir nur Körper, in denen $1 + 1 \neq 0$ ist.

Sei $P(x_1, \dots, x_n)$ quadratisches Polynom wie in (8.1). Wir wollen nun alle Quadriken mit Hilfe von Matrizen beschreiben. Dazu konstruieren wir aus $A = [a_{ij}] \in K^{n,n}, b \in K^n, c \in K$ die folgende Matrix

$$\hat{A} = \begin{bmatrix} c & \frac{b_1}{2} & \frac{b_2}{2} & \dots & \frac{b_n}{2} \\ \frac{b_1}{2} & a_{11} & \frac{a_{12}}{2} & \dots & \frac{a_{1n}}{2} \\ \frac{b_2}{2} & \frac{a_{12}}{2} & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \frac{a_{n-1,n}}{2} \\ \frac{b_n}{2} & \frac{a_{1n}}{2} & \dots & \frac{a_{n-1,n}}{2} & a_{nn} \end{bmatrix} = [\hat{a}_{ij}]. \tag{8.4}$$

\hat{A} ist symmetrisch und enthält alle Koeffizienten des Polynoms. Weiter gilt

$$x \in Q \iff \hat{x}^\top \hat{A} \hat{x} = 0 \quad \text{für } \hat{x} = \begin{bmatrix} 1 \\ x \end{bmatrix} = \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Durch diese Erweiterung gehören also zu den Punkten aus Q diejenigen erweiterten Vektoren, für die die über \hat{A} definierte Bilinearform

$$\hat{\alpha} : K^{n+1} \times K^{n+1} \rightarrow K$$

ergibt $\hat{\alpha}(\hat{x}, \hat{x}) = \hat{x}^\top \hat{A} \hat{x} = 0$.

Wir wollen nun spezielle Abbildungen betrachten, die Abstände erhalten.

Sei V ein endlichdimensionaler euklidischer Vektorraum mit einer Abstandsfunktion

$$d(v, w) = \|v - w\|_2.$$

Eine Abbildung $f : V \rightarrow V$, für die gilt $d(v, w) = d(f(v), f(w)) \quad \forall v, w \in V$ heißt abstandserhaltend (wird manchmal auch Kongruenzabbildung genannt).

Betrachte nun die Abbildung

$$g : V \rightarrow V \\ v \mapsto g(v) = f(v) - f(0).$$

Es gilt natürlich, dass g wieder abstandserhaltend ist und $g(0) = 0$.

Also folgt aus Lemma 3.9, dass g ein orthogonaler Endomorphismus ist. Es gibt also zu jeder abstandserhaltenden Funktion $f : V \rightarrow V$ einen orthogonalen Endomorphismus g , so dass

$$f(v) = a + g(v) \quad \forall v \in V, \quad (a = f(0)).$$

Umgekehrt gilt natürlich sofort, dass alle Abbildungen $v \mapsto a + g(v)$, mit $a \in V$ und g orthogonal, abstandserhaltend sind.

Was ist die Matrixdarstellung von f bzw. g ?

Von g ist das natürlich eine orthogonale Matrix und von f eine Matrix der Form

$$\hat{G} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline a_1 & & & \\ \vdots & & G & \\ a_n & & & \end{array} \right] \quad (8.5)$$

wobei G die (orthogonale) Matrixdarstellung von g ist und $a = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$.

Lemma 8.6 *Ist Q eine Quadrik in \mathbb{R}^n , beschrieben durch die Matrix \hat{A} und $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine abstandserhaltende Abbildung mit der Matrixdarstellung*

$$\hat{G} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline a_1 & & & \\ \vdots & & G & \\ a_n & & & \end{array} \right],$$

so ist $f(Q)$ eine Quadrik, beschrieben durch die Matrix

$$\hat{G}^{-\top} \hat{A} \hat{G}^{-1}. \quad (8.7)$$

Wir sehen, dass dies eine (spezielle) Kongruenztransformation mit \hat{G}^{-1} ist.

Beweis: Sei $y = f(x)$, $\hat{y} = \hat{G}\hat{x}$ mit

$$\hat{y} = \begin{bmatrix} 1 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}, \hat{x} = \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} \implies \hat{x} = \hat{G}^{-1}\hat{y}.$$

(Aus der Gruppeneigenschaft folgt, dass \hat{G} invertierbar ist.)

$$y \in f(Q) \iff x \in Q \iff \hat{x}^\top \hat{A} \hat{x} = 0 \iff \hat{y}^\top \hat{G}^{-\top} \hat{A} \hat{G}^{-1} \hat{y} = 0.$$

Also wird $f(Q)$ gerade durch die Matrix $\hat{G}^{-\top} \hat{A} \hat{G}^{-1}$ beschrieben und ist damit eine Quadrik, denn \hat{G}^{-1} hat die gleiche Form wie \hat{G} . \square

Wir wollen noch einmal anschauen, was

$$\hat{y}^\top \hat{G}^{-\top} \hat{A} \hat{G}^{-1} \hat{y} = 0 \tag{8.8}$$

ist:

$$\hat{y} = \begin{bmatrix} 1 \\ y \end{bmatrix}, \quad \hat{G} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline a_1 & & & \\ \vdots & & & \\ a_n & & & G \end{array} \right] =: \begin{bmatrix} 1 & 0 \\ a & G \end{bmatrix}, \quad \hat{G}^{-1} = \begin{bmatrix} 1 & 0 \\ -G^{-1}a & G^{-1} \end{bmatrix},$$

$$\hat{A} = \begin{bmatrix} c & \frac{b_1}{2} & \cdots & \frac{b_n}{2} \\ \frac{b_1}{2} & a_{11} & & \frac{a_{1n}}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{b_n}{2} & \frac{a_{n1}}{2} & \cdots & a_{nn} \end{bmatrix} =: \begin{bmatrix} c & \frac{b^\top}{2} \\ \frac{b}{2} & \tilde{A} \end{bmatrix}$$

$$\hat{G}^{-1} \begin{bmatrix} 1 \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ -G^{-1}a + G^{-1}y \end{bmatrix} = \begin{bmatrix} 1 \\ G^{-1}(y - a) \end{bmatrix}$$

$$\hat{y}^\top \hat{G}^{-\top} \hat{A} \hat{G}^{-1} \hat{y} = \begin{bmatrix} 1 \\ G^{-1}(y - a) \end{bmatrix}^\top \begin{bmatrix} c & \frac{b^\top}{2} \\ \frac{b}{2} & \tilde{A} \end{bmatrix} \begin{bmatrix} 1 \\ G^{-1}(y - a) \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ G^{-1}(y - a) \end{bmatrix}^\top \begin{bmatrix} c + \frac{b^\top}{2} G^{-1}(y - a) \\ \frac{b}{2} + \tilde{A} G^{-1}(y - a) \end{bmatrix}$$

$$= c + \frac{b^\top}{2} (G^{-1}(y - a)) + (G^{-1}(y - a))^\top \frac{b}{2} + (G^{-1}(y - a))^\top \tilde{A} (G^{-1}(y - a))$$

$$= c + \left(\frac{(y - a)^\top}{2} G^{-\top} b \right) + \left(\frac{(y - a)^\top}{2} G^{-\top} b \right)^\top + (y - a)^\top (G^{-\top} \tilde{A} G^{-1})(y - a).$$

Wir erhalten also, dass \tilde{A} durch eine orthogonale Kongruenztransformation mit G^{-1} transformiert wird.

Deswegen sprechen wir von Abbildung von Quadriken unter Kongruenz.

Satz 8.9 (Klassifikation der Quadriken in \mathbb{R}^n unter Kongruenz)

Sei Q eine Quadrik in \mathbb{R}^n . Dann gibt es eine Kongruenzabbildung f , natürliche Zahlen $\tilde{\pi}, \tilde{\nu}$ und reelle Zahlen $\beta_i > 0, 1 \leq i \leq \tilde{\pi} + \tilde{\nu}$, so dass die Quadrik $f(Q)$ durch eine der folgenden Gleichungen beschrieben wird:

$$\sum_{i=1}^{\tilde{\pi}} \frac{x_i^2}{\beta_i^2} - \sum_{i=\tilde{\pi}+1}^{\tilde{\pi}+\tilde{\nu}} \frac{x_i^2}{\beta_i^2} = 0, \quad \tilde{\pi} \geq \tilde{\nu}, \quad (8.10)$$

$$\sum_{i=1}^{\tilde{\pi}} \frac{x_i^2}{\beta_i^2} - \sum_{i=\tilde{\pi}+1}^{\tilde{\pi}+\tilde{\nu}} \frac{x_i^2}{\beta_i^2} = 1, \quad (8.11)$$

$$\sum_{i=1}^{\tilde{\pi}} \frac{x_i^2}{\beta_i^2} - \sum_{i=\tilde{\pi}+1}^{\tilde{\pi}+\tilde{\nu}} \frac{x_i^2}{\beta_i^2} = x_{\tilde{\pi}+\tilde{\nu}+1}, \quad \tilde{\pi} + \tilde{\nu} < n. \quad (8.12)$$

Beweis: Sei $Q = \{x \in \mathbb{R}^n \mid P(x_1, \dots, x_n) = \hat{x}^\top \hat{A} \hat{x} = 0\}$, wobei \hat{x}, \hat{A} wie in (8.4) gebildet sind,

$$\hat{A} = \begin{bmatrix} c & \frac{b^\top}{2} \\ \frac{b}{2} & \tilde{A} \end{bmatrix}, \quad \text{und} \quad \tilde{A} = \tilde{A}^\top.$$

1. Schritt: Diagonalisierung von \tilde{A} . Da \tilde{A} reell symmetrisch ist, so gibt es nach dem Trägheitssatz von Sylvester eine orthogonale Matrix P , so dass

$$P^\top \tilde{A} P = \begin{bmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_\pi & \\ & & & \lambda_{\pi+1} & & \\ & & & & \ddots & \\ & & & & & \lambda_{\pi+\nu} & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & & 0 \end{bmatrix}$$

mit $\lambda_1, \dots, \lambda_\pi > 0, \lambda_{\pi+1} \dots \lambda_{\pi+\nu} < 0$. Setze $\hat{P} = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & P \end{array} \right]$, so gilt

$$A_1 = \hat{P}^\top \hat{A} \hat{P} = \left[\begin{array}{c|c} c & \frac{b^\top}{2} P \\ \hline P^\top \frac{b}{2} & P^\top \tilde{A} P \end{array} \right] =: \left[\begin{array}{cccc} c & \gamma_1 & \cdots & \gamma_n \\ \gamma_1 & \lambda_1 & & \\ \vdots & & \ddots & \\ \gamma_n & & & \lambda_n \end{array} \right]$$

2. Schritt: Verschiebung des Nullpunktes

$$\text{Sei } T = \begin{bmatrix} 1 & & & & & & \\ -\frac{\gamma_1}{\lambda_1} & 1 & & & & & \\ \vdots & & \ddots & & & & \\ -\frac{\gamma_{\pi+\nu}}{\lambda_{\pi+\nu}} & & & 1 & & & \\ 0 & & & & 1 & & \\ \vdots & & & & & \ddots & \\ 0 & & & & & & 1 \end{bmatrix},$$

$$A_2 = T^\top A_1 T = \left[\begin{array}{c|ccc|c} c & 0 & \cdots & 0 & \gamma_{\pi+\nu+1} \cdots \gamma_n \\ \hline 0 & \lambda_1 & & & 0 \\ \vdots & & \ddots & & \\ 0 & & & \lambda_{\pi+\nu} & \\ \hline \gamma_{\pi+\nu+1} & & & & 0 \\ \vdots & 0 & & & \\ \gamma_n & & & & \end{array} \right].$$

3. Schritt: Jetzt unterscheiden wir 3 Fälle.

(a) $c = 0$ und $\gamma_{\pi+\nu+1} = \dots = \gamma_n = 0$.

Mit der Setzung $\beta_i = \frac{1}{\sqrt{|\lambda_i|}}$ erhalten wir dann $\text{Rang } \tilde{A} = \text{Rang } \hat{A}$ und Teil (8.10).

(b) $c \neq 0$ und $\gamma_{\pi+\nu+1} = \dots = \gamma_n = 0$, also $\text{Rang } \hat{A} = \text{Rang } \tilde{A} + 1$.

Mit $\beta_i = \sqrt{\frac{|c|}{|\lambda_i|}}$ erhalten wir (8.11).

(c) $c \neq 0$ und es gibt $\gamma_j \neq 0$, $\pi + \nu < j \leq n \implies \text{Rang } \hat{A} = \text{Rang } \tilde{A} + 2$.

$$\text{Setze } \gamma = \begin{bmatrix} \gamma_{\pi+\nu+1} \\ \vdots \\ \gamma_n \end{bmatrix} \text{ und } \varphi_1 = \frac{\gamma}{\|\gamma\|_2}.$$

Ergänze φ_1 durch $\varphi_2, \dots, \varphi_{n-\pi-\nu}$ zu einer Orthonormalbasis von $\mathbb{R}^{n-\pi-\nu}$, setze

$$\tilde{c} = -\frac{c}{2\|\gamma\|_2} \quad \text{und} \quad V = \left[\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & I_{\pi+\nu} & 0 \\ \hline \tilde{c}\varphi_1 & 0 & \tilde{V} \end{array} \right] \text{ mit } \tilde{V} = [\varphi_1 \varphi_2 \cdots \varphi_{n-\pi-\nu}].$$

Dann gilt

$$V^T A_2 V = \left[\begin{array}{ccc|ccc} 0 & 0 & \cdots & 0 & \|\gamma\|_2 & 0 & \cdots & 0 \\ \hline 0 & \lambda_1 & & & & & & \\ \vdots & & \ddots & & & & & 0 \\ 0 & & & \lambda_{\pi+\nu} & & & & \\ \hline \|\gamma\|_2 & & & & & & & \\ 0 & & & & & & & \\ \vdots & & & 0 & & & & 0 \\ 0 & & & & & & & \end{array} \right]$$

und mit $\beta_i = \sqrt{\frac{2\|\gamma\|_2}{|\lambda_i|}}$, $i = 1, \dots, \pi + \nu$ erhalten wir (8.12).

Beachte, je nach Vorzeichen von c wechseln die Rollen von π und ν . □

Beispiel 8.13

$$p(x) = x_1^2 + 9x_2^2 - 6x_1x_2 + 20x_1 - 4x_2 - 10$$

$$\Rightarrow \tilde{A} = \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix}, \quad \hat{A} = \begin{bmatrix} -10 & \frac{20}{2} & -\frac{4}{2} \\ \frac{20}{2} & 1 & -3 \\ -\frac{4}{2} & -3 & 9 \end{bmatrix} = \begin{bmatrix} -10 & 10 & -2 \\ 10 & 1 & -3 \\ -2 & -3 & 9 \end{bmatrix}$$

Rang $\tilde{A} = 1$, Rang $\hat{A} = 3$, da invertierbar. Also haben wir den Fall (8.12).

(1) Diagonalisierung von \tilde{A} :

$$P_{\tilde{A}}(\lambda) = \lambda^2 - 10\lambda, \quad \Rightarrow \quad \text{Eigenwerte } 0, 10$$

$$P = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}, \quad \text{so dass } P^T(\tilde{A} - 0 \cdot I)P = \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix}$$

$$s = \frac{-3}{\sqrt{3^2 + 1}} = -\frac{3}{\sqrt{10}}, \quad c = \frac{1}{\sqrt{10}}$$

$$P^T \tilde{A} P = \begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & P^T \end{bmatrix} \hat{A} \begin{bmatrix} 1 & 0 \\ 0 & P \end{bmatrix} = \begin{bmatrix} -10 & \frac{8}{5}\sqrt{10} & \frac{14}{5}\sqrt{10} \\ \frac{8}{5}\sqrt{10} & 10 & 0 \\ \frac{14}{5}\sqrt{10} & 0 & 0 \end{bmatrix}$$

(2) Verschiebung des Nullpunktes

$$T = \left[\begin{array}{c|cc} 1 & 0 & 0 \\ -\frac{8\sqrt{10}}{5 \cdot 10} & 1 & 0 \\ 0 & 0 & 1 \end{array} \right], \quad A_2 = T^T A_1 T = \begin{bmatrix} -\frac{314}{25} & 0 & \frac{14}{5}\sqrt{10} \\ 0 & 10 & 0 \\ \frac{14}{5}\sqrt{10} & 0 & 0 \end{bmatrix}$$

$$\gamma = \left[\frac{14}{5}\sqrt{10} \right] \in \mathbb{R}^1, \quad \varphi_1 = [1], \quad \tilde{c} = \frac{314}{25 \cdot 2 \cdot \|\gamma\|_2} = \frac{157}{700}\sqrt{10} \quad .$$

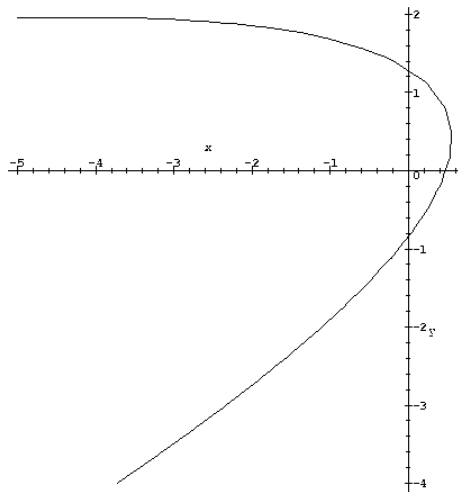
$$V = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{157}{700}\sqrt{10} & 0 & 1 \end{bmatrix}, \quad V^T A_2 V = \begin{bmatrix} 0 & 0 & \frac{14}{5}\sqrt{10} \\ 0 & 10 & 0 \\ \frac{14}{5}\sqrt{10} & 0 & 0 \end{bmatrix}$$

Setze $\beta_1 = \sqrt{\frac{2 \cdot 14\sqrt{10}}{5 \cdot 10}}$. Dann erhalten wir die transformierte Gleichung

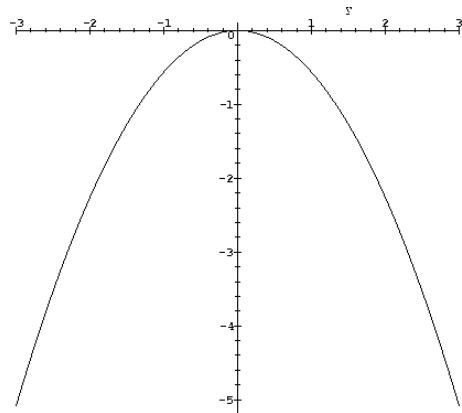
$$\frac{y_1^2}{\beta_1^2} = y_2 \quad \text{oder} \quad \frac{5}{28}\sqrt{10}y_1^2 - y_2 = 0, \quad \text{wobei}$$

$$\hat{x} = P \cdot T \cdot V \hat{y} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{359}{700} & \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \\ \frac{493}{700} & -\frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} 1 \\ y_1 \\ y_2 \end{bmatrix}, \quad \text{bzw.}$$

$$\hat{y} = V^{-1} \cdot T^{-1} \cdot \hat{P}^{-1} \hat{x} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{4}{25}\sqrt{10} & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ -\frac{157}{700}\sqrt{10} & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} .$$



$$p(x) = x_1^2 + 9x_2^2 - 6x_1x_2 + 20x_1 - 4x_2 - 10 = 0$$

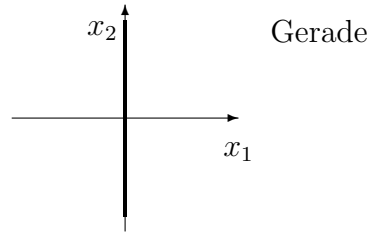


$$q(y) = \frac{5}{28}\sqrt{10}y_1^2 - y_2 = 0$$

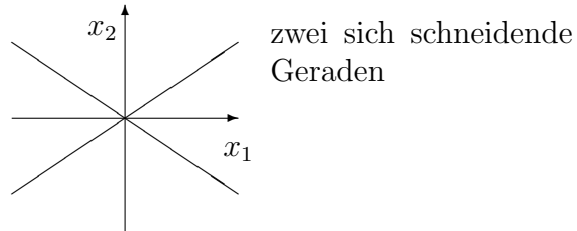
Wir können damit alle Quadriken klassifizieren. Im \mathbb{R}^2 erhalten wir die folgenden Möglichkeiten:

Tabelle 8.14 Quadriken in \mathbb{R}^2

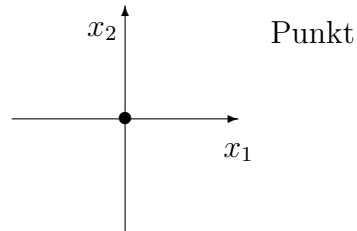
(8.10) $\nu = 0, \quad \pi = 1, \quad x_1^2 = 0$



$\nu = 1, \quad \pi = 1, \quad \frac{x_1^2}{\beta_1^2} - x_2^2 = 0$
 $x_1 = \pm \beta_1 x_2$



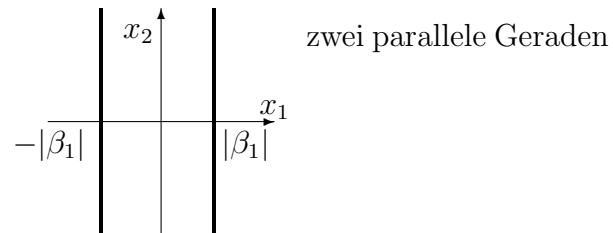
$\nu = 0, \quad \pi = 2, \quad \frac{x_1^2}{\beta_1^2} + x_2^2 = 0$



(8.11) $\nu = 1, \quad \pi = 0, \quad \frac{-x_1^2}{\beta_1^2} = 1$

\emptyset

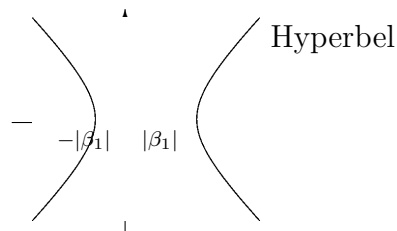
$\nu = 0, \quad \pi = 1, \quad \frac{x_1^2}{\beta_1^2} = 1$



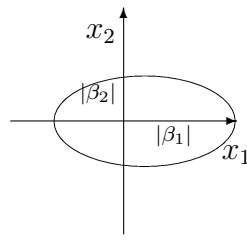
$\nu = 2, \quad \pi = 0, \quad \frac{-x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} = 1$

\emptyset

$\nu = 1, \quad \pi = 1, \quad \frac{x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} = 1$

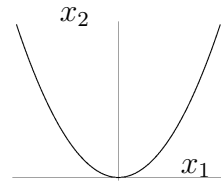


$$\nu = 0, \quad \pi = 2, \quad \frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} = 1$$



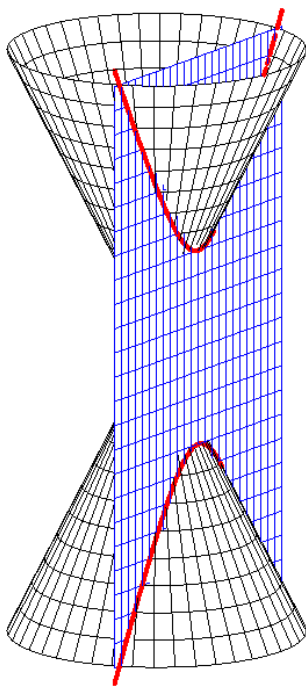
Ellipse

(8.12) $\nu = 0, \quad \pi = 1, \quad \frac{x_1^2}{\beta_1^2} = x_2$

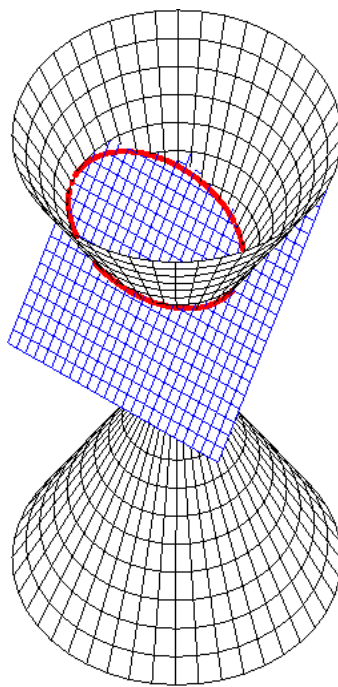


Parabel

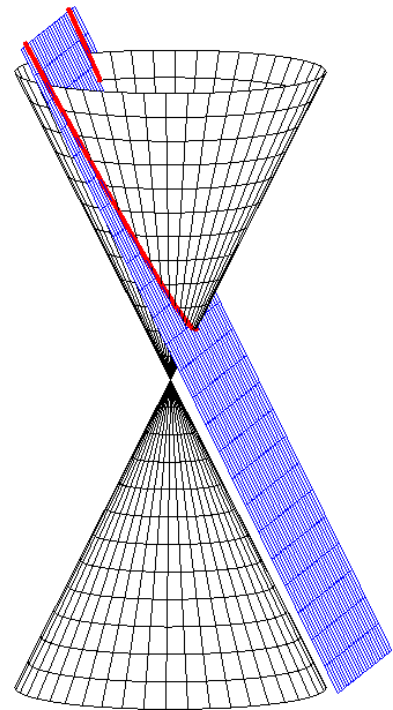
Quadriken sind Schnitte von Ebenen mit einem doppelten Kreiskegel, sie werden daher auch Kegelschnitte genannt.



Hyperbel



Ellipse

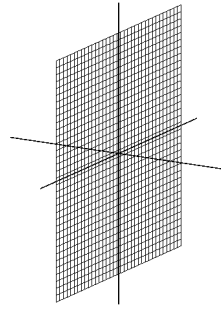


Parabel

Tabelle 8.15 Quadriken in \mathbb{R}^3

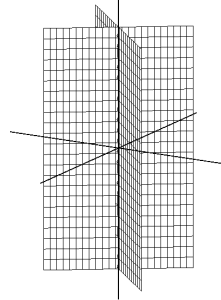
$$(8.10) \quad \nu = 0, \quad \pi = 1, \quad x_1^2 = 0$$

eine Ebene



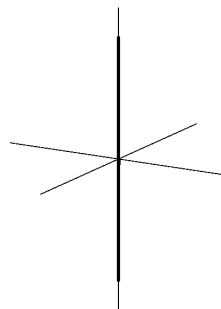
$$\nu = 1, \quad \pi = 1, \quad \frac{x_1^2}{\beta_1^2} - x_2^2 = 0$$

zwei sich schneidende Ebenen



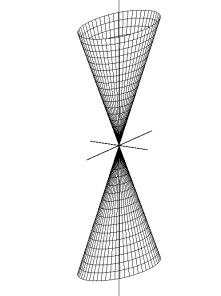
$$\nu = 0, \quad \pi = 2, \quad \frac{x_1^2}{\beta_1^2} + x_2^2 = 0$$

eine Gerade



$$\nu = 1, \quad \pi = 2, \quad \frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} - x_3^2 = 0$$

Ellipsenkegel



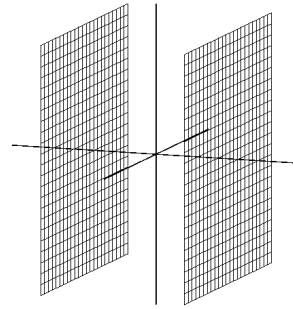
$$\nu = 0, \quad \pi = 3, \quad \frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} + x_3^2 = 0$$

Punkt

$$(8.11) \quad \nu = 1, \quad \pi = 0, \quad \frac{-x_1^2}{\beta_1^2} = 1$$

 \emptyset

$$\nu = 0, \quad \pi = 1, \quad \frac{x_1^2}{\beta_1^2} = 1$$

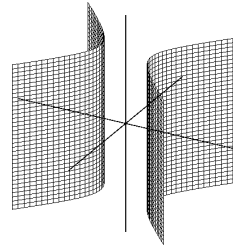


zwei parallele Ebenen

$$\nu = 2, \quad \pi = 0, \quad \frac{-x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} = 1$$

\emptyset

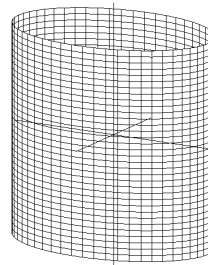
$$\nu = 1, \quad \pi = 1, \quad \frac{x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} = 1$$



Hyperbelzylinder

$$\nu = 0, \quad \pi = 2, \quad \frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_1^2} = 1$$

Ellipsenzylinder

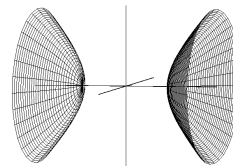


$$\nu = 3, \quad \pi = 0, \quad \frac{-x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} - \frac{x_3^2}{\beta_3^2} = 1$$

\emptyset

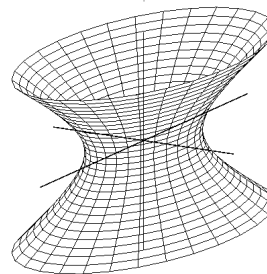
$$\nu = 2, \quad \pi = 1, \quad \frac{x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} - \frac{x_3^2}{\beta_3^2} = 1$$

zweischaliges Hyperboloid



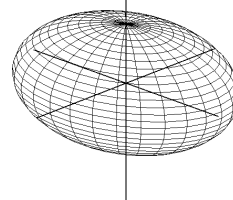
$$\nu = 1, \quad \pi = 2, \quad \frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} - \frac{x_3^2}{\beta_3^2} = 1$$

einschaliges Hyperboloid

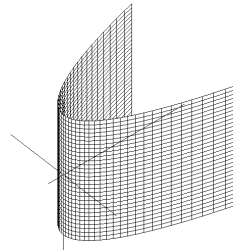


$$\nu = 0, \quad \pi = 3, \quad \frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} + \frac{x_3^2}{\beta_3^2} = 1$$

Ellipsoid

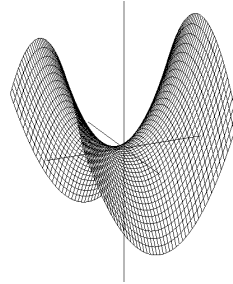


$$(8.12) \quad \nu = 0, \quad \pi = 1, \quad \frac{x_1^2}{\beta_1^2} = x_2$$



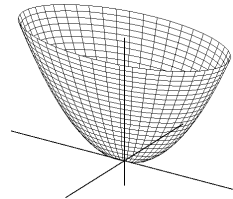
parabolischer
Zylinder

$$\nu = 1, \quad \pi = 1, \quad \frac{x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} = x_3$$



hyperbolisches Pa-
raboloid

$$\nu = 0, \quad \pi = 2, \quad \frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} = x_3$$



elliptisches Parabo-
loid

Definition 8.16 Sei V ein Vektorraum über K . Eine Abbildung $f : V \rightarrow V$ heißt affin linear, falls es $a \in V$ und lineare Abbildung $g : V \rightarrow V$ gibt, so dass $f(v) = a + g(v)$. Zwei Quadriken Q_1, Q_2 heißen affin äquivalent, wenn es eine bijektive affine Abbildung $f : V \rightarrow V$ gibt mit $f(Q_1) = Q_2$.

Korollar 8.17 Jede Quadrik in \mathbb{R}^n ist affin äquivalent zu einer Quadrik in \mathbb{R}^n , die durch eine der folgenden Gleichungen gegeben ist:

$$\sum_{i=1}^{\pi} x_i^2 - \sum_{i=\pi+1}^{\pi+\nu} x_i^2 = 0 \tag{8.18}$$

$$\sum_{i=1}^{\pi} x_i^2 - \sum_{i=\pi+1}^{\pi+\nu} x_i^2 = 1 \tag{8.19}$$

$$\sum_{i=1}^{\pi} x_i^2 - \sum_{i=\pi+1}^{\pi+\nu} x_i^2 = x_{\pi+\nu+1} \tag{8.20}$$

Beweis: Wir können natürlich annehmen, dass wir schon eine Beschreibung der Quadrik Q in der Form (8.10 - 8.12) haben. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ die lineare Abbildung mit der Matrixdar-

stellung

$$\left[\begin{array}{ccccccc} \beta_1^{-1} & & & & & & \\ & \ddots & & & & & \\ & & \beta_{\pi+\nu}^{-1} & & & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & & 1 \end{array} \right] \text{ bzgl. } \{e_1, \dots, e_n\},$$

so ist f natürlich bijektiv und wir haben sofort, dass $f(Q)$ einer der Gleichungen (8.18 - 8.20) genügt, denn sei zum Beispiel (8.10) die Form von Q , so gilt

$$y = f(x) \quad \text{mit} \quad \begin{array}{ll} y_i = \frac{1}{\beta_i} x_i & i = 1, \dots, \pi + \nu \\ y_i = x_i & i = \pi + \nu + 1, \dots, n \end{array}$$

$$\begin{aligned} y \in f(Q) &\iff x \in Q \iff \sum_{i=1}^{\pi} \frac{x_i^2}{\beta_i^2} - \sum_{i=\pi+1}^{\pi+\nu} \frac{x_i^2}{\beta_i^2} = 0 \\ &\iff \sum_{i=1}^{\pi} y_i^2 - \sum_{i=\pi+1}^{\pi+\nu} y_i^2 = 0. \end{aligned}$$

Die anderen Fälle sind analog.

□

Kapitel 9

Positiv definite Matrizen (Bilinearformen)

Wir haben bereits gesehen, dass symmetrisch positiv definite Matrizen (Bilinearformen) eine wichtige Rolle spielen (Skalarprodukte, Quadriken usw.), und dass diese dadurch charakterisiert sind, dass alle Eigenwerte reell und positiv sind. Jetzt wollen wir noch einige weitere wichtige Eigenschaften kennenlernen.

Zuerst noch ein paar Bemerkungen zum komplexen Fall.

Definition 9.1 Sei V ein unitärer Raum und $\beta : V \times V \rightarrow \mathbb{C}$ eine Hermite'sche Form, so heißt β positiv definit (positiv semidefinit) genau dann, wenn $\beta(v, v) > 0$ ($\beta(v, v) \geq 0$) für alle $v \in V \setminus \{0\}$.

Beachte, dass damit auch $\beta(v, v) \in \mathbb{R}$ gilt.

Korollar 9.2 Sei V ein unitärer Raum und $\beta : V \times V \rightarrow \mathbb{C}$ eine Hermite'sche Form. Dann sind folgende Aussagen äquivalent:

- (i) β ist positiv definit.
- (ii) Jede Matrixdarstellung B von β (bezüglich einer Basis $\{v_1, \dots, v_n\}$ von V) hat nur positive reelle Eigenwerte.
- (iii) Für jede Matrixdarstellung B von β gibt es eine nichtsinguläre Matrix T , so dass

$$T^{-H} B T^{-1} = I.$$

- (iv) Für jede Matrixdarstellung B von β gibt es eine nichtsinguläre untere Dreiecksmatrix L , so dass $B = LL^H$ (Cholesky-Zerlegung).
- (v) Für jede Matrixdarstellung B von β gibt es eine Hermite'sche Matrix A mit beliebigem Trägheitsindex $(\pi, \nu, 0)$, so dass $A^2 = B$.

Beweis: Der Beweis, dass (i) \iff (ii), gilt analog wie im reellen Fall.

„(ii) \implies (iii)“: Nach dem Satz von Schur gibt es eine unitäre Matrix Q , so dass $B = Q^H D Q$ mit $D = \text{diag}(d_1, \dots, d_n)$, $d_i > 0$. Setze

$$D^{\frac{1}{2}} = \begin{bmatrix} d_1^{\frac{1}{2}} & & \\ & \ddots & \\ & & d_n^{\frac{1}{2}} \end{bmatrix},$$

so folgt mit $T = D^{\frac{1}{2}} Q$, dass T invertierbar und

$$B = T^H T \quad \implies \quad T^{-H} B T^{-1} = I.$$

„(iii) \implies (iv)“: Nach Satz 10.11 mit Folgerung 10.13 aus Teil I gibt es Q unitär, so dass $T = QR$ mit einer oberen Dreiecksmatrix R . Es folgt

$$B = T^H T = R^H Q^H Q R = R^H R \quad \text{mit} \quad L = R^H.$$

„(iv) \implies (iii)“: trivial

„(iii) \implies (ii)“: Übungsaufgabe

„(ii) \implies (v)“: $B = Q^H D Q$, $D = \text{diag}(d_1, \dots, d_n)$ mit $d_i > 0$, $i = 1, \dots, n$.

Setze $A = Q^H D^{\frac{1}{2}} \begin{bmatrix} I_\pi & \\ & -I_\nu \end{bmatrix} Q$, so folgt, dass A Hermite'sch ist und

$$A^2 = Q^H D^{\frac{1}{2}} \begin{bmatrix} I_\pi & \\ & -I_\nu \end{bmatrix} Q Q^H D^{\frac{1}{2}} \begin{bmatrix} I_\pi & \\ & -I_\nu \end{bmatrix} Q = Q^H D^{\frac{1}{2}} D^{\frac{1}{2}} Q = Q^H D Q.$$

„(v) \implies (ii)“: Bilde $A = Q^H D Q$ mit $D = \text{diag}(d_1, \dots, d_n)$ und $d_i \neq 0$, reell, $i = 1, \dots, n$, so folgt $B = A^2 = Q^H D^2 Q$, und damit hat B nur positive Eigenwerte. \square

Um die Cholesky-Zerlegung einer Hermite'schen, positiv definiten Matrix $A \in \mathbb{C}^{n,n}$ zu berechnen, betrachten wir die Gleichung

$$A = LL^H = \begin{bmatrix} l_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ l_{n1} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} \overline{l_{11}} & \cdots & \overline{l_{n1}} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \overline{l_{nn}} \end{bmatrix}, \quad L = [l_{ij}], \quad l_{ij} = 0 \quad \text{für} \quad j > i.$$

So gilt (für Zeile 1 von A)

$$a_{11} = l_{11} \overline{l_{11}} \implies l_{11} = \sqrt{a_{11}} = \overline{l_{11}}, \quad (\text{wir können } l_{11} \text{ so wählen})$$

$$a_{1j} = l_{11} \overline{l_{j1}} \implies l_{j1} = \frac{\overline{a_{1j}}}{l_{11}}, \quad j = 2, \dots, n. \tag{9.3}$$

Analog (für Zeile i)

$$\begin{aligned}
 a_{ii} &= \sum_{j=1}^i l_{ij} \overline{l_{ij}} \implies l_{ii} = \sqrt{a_{ii} - \sum_{j=1}^{i-1} |l_{ij}|^2} = \overline{l_{ii}}, \quad (\text{wir können } l_{ii} \text{ so wählen}) \\
 a_{ij} &= \sum_{k=1}^n l_{ik} \overline{l_{jk}} = \sum_{k=1}^i l_{ik} \overline{l_{jk}} + \sum_{k=1}^{i-1} l_{ik} \overline{l_{jk}} + l_{ii} \overline{l_{ji}} \\
 \implies l_{ji} &= \frac{1}{l_{ii}} \left(a_{ij} - \sum_{k=1}^{i-1} l_{ik} \overline{l_{jk}} \right) \quad \text{für } j > i.
 \end{aligned} \tag{9.4}$$

Beispiel 9.5

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

$$l_{11} = \sqrt{a_{11}} = \sqrt{2}, \quad l_{j1} = \frac{\overline{a_{1j}}}{l_{11}} \implies l_{21} = \frac{\overline{a_{12}}}{l_{11}} = \frac{1}{\sqrt{2}}, \quad l_{31} = 0,$$

$$l_{22} = \sqrt{a_{22} - \sum_{j=1}^1 |l_{2j}|^2} = \sqrt{a_{22} - |l_{21}|^2} = \sqrt{2 - \frac{1}{2}} = \sqrt{\frac{3}{2}},$$

$$l_{32} = \frac{1}{l_{22}} \left(a_{23} - \sum_{k=1}^1 l_{2k} \overline{l_{3k}} \right) = \frac{a_{23} - l_{21} \overline{l_{31}}}{l_{22}} = \sqrt{\frac{2}{3}} (1 - 0) = \sqrt{\frac{2}{3}},$$

$$l_{33} = \sqrt{a_{33} - |l_{31}|^2 - |l_{32}|^2} = \sqrt{2 - \frac{2}{3}} = \sqrt{\frac{4}{3}},$$

$$L = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{\frac{3}{2}} & 0 \\ 0 & \sqrt{\frac{2}{3}} & \sqrt{\frac{4}{3}} \end{bmatrix}, \quad L^H = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \sqrt{\frac{3}{2}} & \sqrt{\frac{2}{3}} \\ 0 & 0 & \sqrt{\frac{4}{3}} \end{bmatrix}.$$

Probe:

$$LL^H = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

Es gibt noch viele weitere wichtige Eigenschaften Hermite'scher (positiv definiten) Matrizen. Um diese wichtigen Sätze zu betrachten, ordnen wir die Eigenwerte im folgenden.

Falls $\lambda_1(A), \dots, \lambda_n(A) \in \mathbb{R}$ die Eigenwerte von A sind (gezählt mit ihrer Multiplizität), so gelte

$$\lambda_1(A) \geq \lambda_2(A) \geq \dots \geq \lambda_n(A).$$

Dann gilt der

Satz 9.6 (Satz von Courant–Fischer)

Sei $A \in \mathbb{C}^{n,n}$ Hermite'sch, so gilt für $k = 0, \dots, n - 1$

$$\lambda_{k+1}(A) = \min_{\substack{\dim(S)=n-k \\ S \subset \mathbb{C}^n}} \max_{0 \neq x \in S} \frac{x^H Ax}{x^H x}.$$

Beweis: Betrachte

$$\max x^H Ax \quad \text{mit den Nebenbedingungen} \quad \|x\|_2 = 1 \quad \text{und} \quad p_i^H x = 0, \quad i = 1, \dots, k,$$

wobei $p_i \neq 0$ irgendwelche linear unabhängige Vektoren sind. Nach Satz 4.6 gibt es eine unitäre Matrix Q , so dass

$$Q^H A Q = D, \quad \text{mit} \quad D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Wir setzen $y = Q^H x$, so folgt $x^H Ax = x^H Q Q^H A Q Q^H x = y^H D y$. Setzen wir noch $q_i = Q^H p_i$, können wir also stattdessen betrachten:

$$\max \sum_{i=1}^n \bar{y}_i \lambda_i y_i \quad \text{mit den Nebenbedingungen} \quad y^H y = 1 \quad \text{und} \quad q_i^H y = 0, \quad i = 1, \dots, k.$$

Für beliebige Vektoren p_i und damit q_i erhalten wir damit $k + 1 \leq n$ Gleichungen für die Komponenten von y .

Für jede Wahl y der Form

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_{k+1} \\ 0 \end{bmatrix} \in \mathbb{R}^n, \quad \text{so dass} \quad \begin{bmatrix} q_1^H \\ \vdots \\ q_k^H \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_k \\ y_{k+1} \\ 0 \end{bmatrix} = 0,$$

können wir y so normieren, dass $\|y\|_2 = 1$, und erhalten die zu maximierende Funktion

$$\sum_{i=1}^{k+1} |y_i|^2 \lambda_i \geq \lambda_{k+1} \sum_{i=1}^{k+1} |y_i|^2 = \lambda_{k+1}.$$

Also $\max x^H Ax \geq \lambda_{k+1}$ für alle Wahlen p_i und damit auch

$$\min_{\substack{\dim(S)=n-k \\ S \subset \mathbb{C}^n}} \max_{0 \neq x \in S} \frac{x^H Ax}{x^H x} \geq \lambda_{k+1}.$$

Für die spezielle Wahl $p_i = Q e_i$, $i = 1, \dots, k$, ergibt sich $q_i = e_i$ und damit $y_1, \dots, y_k = 0$.

Also

$$\begin{aligned} x^H Ax &= y^H \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} y = \sum_{i=k+1}^n \lambda_i |y_i|^2 \leq \lambda_{k+1} \sum_{i=k+1}^n |y_i|^2 = \lambda_{k+1} \\ \implies \max \frac{x^H Ax}{x^H x} &\leq \lambda_{k+1} \implies \min \max \frac{x^H Ax}{x^H x} \leq \lambda_{k+1}. \end{aligned}$$

Also folgt die Behauptung. □

Kapitel 10

Eigenwert– und Eigenvektorberechnung

Wir wollen in diesem Kapitel einige Methoden zur Eigenwert- und Eigenvektorberechnung betrachten. In Teil I (Seminar) haben wir schon die Nullstellenberechnung für Polynome behandelt. Man berechnet allerdings fast immer besser die Eigenwerte einer Matrix, ohne über das charakteristische Polynom zu gehen. Wir definieren zunächst den Begriff des orthogonalen Projektors.

Definition 10.1 Sei $S \in K^n$ ein Unterraum.
 $P \in K^{n,n}$ heißt orthogonaler Projektor auf S , falls

$$\text{Bild}(P) = S, \quad P^2 = P \quad \text{und} \quad P^\top = P \quad \text{für} \quad K = \mathbb{R}, \quad (\text{bzw.} \quad P^H = P \quad \text{für} \quad K = \mathbb{C}).$$

Die orthogonalen Projektoren kann man z. B. aus der Singulärwertzerlegung erhalten. Sei $A = U\Sigma V^H$ die Singulärwertzerlegung von A und $r = \text{Rang}(A)$. Setze

$$\begin{aligned} U &= [U_r \quad \tilde{U}_r] \\ V &= [V_r \quad \tilde{V}_r], \end{aligned}$$

so ist

$$\begin{aligned} V_r V_r^H & \text{ die orthogonale Projektion auf Bild}(A^H), \\ \tilde{V}_r \tilde{V}_r^H & \text{ die orthogonale Projektion auf Kern}(A), \\ U_r U_r^H & \text{ die orthogonale Projektion auf Bild}(A), \\ \tilde{U}_r \tilde{U}_r^H & \text{ die orthogonale Projektion auf Kern}(A^H). \end{aligned}$$

Für zwei Unterräume $S_1, S_2 \in K^n$ lässt sich damit ein Abstand definieren:

$$\text{dist}(S_1, S_2) := \|P_1 - P_2\|_2, \quad P_i \text{ orthogonaler Projektor auf } S_i, \quad i = 1, 2.$$

Die erste Methode, die wir betrachten wollen, ist die Potenzmethode.

Algorithmus 10.2 Gegeben $A \in \mathbb{C}^{n,n}$ und ein beliebiger Vektor $x^{(0)} \in \mathbb{C}^n$ mit $\|x^{(0)}\|_2 = 1$.

for $k = 1, 2, \dots$

(1) Setze $y^{(k)} = Ax^{(k-1)}$.

(2) Normalisiere $x^{(k)} = \frac{y^{(k)}}{\|y^{(k)}\|_2}$.

(3) Teste, ob $\|x^{(k)} - x^{(k-1)}\|_2$ klein.

Seien $\lambda_1, \lambda_2, \dots, \lambda_n$ mit $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$ die Eigenwerte einer diagonalisierbaren Matrix A mit den Eigenvektoren z_k , d.h.

$$Az_k = \lambda_k z_k \quad \text{mit} \quad \|z_k\|_2 = 1. \quad (10.3)$$

Wann konvergiert nun die Potenzmethode?

Satz 10.4 Sei $A \in \mathbb{C}^{n,n}$ diagonalisierbar und es gelte $|\lambda_1| > |\lambda_2|$, so konvergiert die Potenzmethode mit der Konvergenzgeschwindigkeit $\left| \frac{\lambda_2}{\lambda_1} \right|$ gegen z_1 , falls $\alpha_1 \neq 0$ in der Darstellung

$$x^{(0)} = \sum_{i=1}^n \alpha_i z_i.$$

Beweis: $x^{(0)} = \sum_{i=1}^n \alpha_i z_i$, also

$$\begin{aligned} A^k x^{(0)} &= A^k \sum_{i=1}^n \alpha_i z_i \\ &= \sum_{i=1}^n \alpha_i \lambda_i^k z_i = \alpha_1 \lambda_1^k z_1 + \sum_{i=2}^n \alpha_i \lambda_i^k z_i \\ &= \alpha_1 \lambda_1^k \left[z_1 + \sum_{j=2}^n \frac{\alpha_j}{\alpha_1} \left(\frac{\lambda_j}{\lambda_1} \right)^k z_j \right]. \end{aligned}$$

Für $k \rightarrow \infty$ geht nach Voraussetzung $\left(\frac{\lambda_j}{\lambda_1} \right)^k \rightarrow 0$, also konvergiert die Folge gegen z_1 .

Da $x^{(k)} \in \text{span} \{A^k x^{(0)}\}$, folgt

$$\text{dist} \left(\mathcal{L}(x^{(k)}), \mathcal{L}(z_1) \right) \rightarrow 0 \quad \text{und} \quad |\lambda_1 - \lambda^{(k)}| \rightarrow 0 \quad \text{für} \quad k \rightarrow \infty.$$

Dabei ist $\lambda^{(k)} = \frac{y_i^{(k)}}{x_i^{(k-1)}} = \frac{(Ax^{(k-1)})_i}{x_i^{(k-1)}}$ (für eine Komponente $x_i^{(k-1)} \neq 0$)

oder $|\lambda^{(k)}| = \|y^{(k)}\|_2$ oder $\lambda^{(k)} = (x^{(k)})^H A x^{(k)}$. □

Beispiel 10.5

$$A = \begin{bmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{bmatrix} \quad \text{hat die Eigenwerte } \{3, 4, 10\}.$$

Verwende $x^{(0)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, dann ergibt sich bei der Potenzmethode für $\lambda^{(k)}$:

k	$\frac{y_1^{(k)}}{x_1^{(k-1)}}$	$\ y^{(k)}\ _2$	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\ Ax^{(k)} - \lambda^{(k)}x^{(k)}\ _2$
1	-261.00000,	994.495349,	-.26244467,	-.53293361,	-.80442810,	274.06048
2	13.2145592,	13.0605581,	-.26553923,	-.53354216,	-.80300790,	2.4954564
3	10.7599327,	10.7191503,	-.26654951,	-.53401840,	-.80235635,	.55266728
4	10.2230953,	10.2072991,	-.26696201,	-.53428429,	-.80204214,	.15983316
5	10.0698519,	10.0632843,	-.26713624,	-.53441568,	-.80189657,	.05006162
6	10.0225526,	10.0198037,	-.26720953,	-.53447621,	-.80183181,	.01626419
7	10.0074380,	10.0062959,	-.26724003,	-.53450290,	-.80180385,	.00540406
8	10.0025084,	10.0020379,	-.26725260,	-.53451432,	-.80179205,	.00184831
9	10.0008544,	10.0006622,	-.26725774,	-.53451912,	-.80178714,	.00063392
10	10.0002994,	10.0002214,	-.26725982,	-.53452111,	-.80178512,	.00021929

$$\frac{x^{(10)}}{x_1^{(10)}} = \begin{bmatrix} 1.000000000 \\ 2.000005497 \\ 3.000021151 \end{bmatrix} \quad \text{Eigenvektor zu } \lambda_1 = 10 : \quad z_1 = \frac{1}{\sqrt{14}} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Der Faktor $\frac{\lambda_2}{\lambda_1}$ ist 0.4, also $\left(\frac{\lambda_2}{\lambda_1}\right)^{10} \approx 0.0001$.

Die direkte Verallgemeinerung der Potenzmethode ist die orthogonale Iteration.

Sei $Q_0 \in \mathbb{C}^{n,p}$ eine Matrix mit orthonormalen Spalten.

Algorithmus 10.6

for $k = 1, 2, \dots$

$$\begin{aligned} Z_k &= A Q_{k-1} \\ Q_k R_k &= Z_k \quad \text{QR-Zerlegung.} \end{aligned}$$

Sei

$$P^H A P = \Gamma = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} + N$$

die Schur-Zerlegung von A mit $|\lambda_1| \geq |\lambda_2| \dots \geq |\lambda_p| > |\lambda_{p+1}| \geq \dots \geq |\lambda_n|$ und

$$P = \begin{bmatrix} P_1 & P_2 \\ p & n-p \end{bmatrix}, \quad \Gamma = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} & & \\ 0 & \Gamma_{22} & & \\ & & & \\ & & & \end{bmatrix} \begin{matrix} p \\ n-p \\ \\ \end{matrix}.$$

So konvergiert die Folge $\{\text{Bild}(Q_k)\}$ gegen $\text{Bild}(P_1)$ mit der Rate $\left| \frac{\lambda_{p+1}}{\lambda_p} \right|^k$.

Beispiel 10.7 A wie in Bsp. 10.5

$$Q_0 = [e_1, e_2], \quad \text{Faktor } \frac{\lambda_2}{\lambda_3} = \frac{3}{4} = 0.75$$

Wir betrachten den Abstand $d_k = \text{dist}(\text{Bild}(P_1), \text{Bild}(Q_k)) = \|W_1 - V_k\|_2$, wobei W_1 orthogonaler Projektor auf $\text{Bild}(P_1)$ und V_k orthogonaler Projektor auf $\text{Bild}(Q_k)$ ist, sowie das Verhältnis $q_k = \frac{d_k}{d_{k-1}}$, ($d_0 = 1$).

k	d_k	q_k	r_{11}	r_{22}
1	0.0052170	0.9187	994.4953494	-4.1139003
2	0.0047929	0.8217	13.0605581	3.9728619
4	0.0030776	0.7814	10.2072916	3.9806330
5	0.0023527	0.7645	10.0632840	3.9862537
6	0.0017808	0.7569	10.0198036	3.9901713
7	0.0013417	0.7535	10.0063037	3.9928825
8	0.0010088	0.7519	10.0020396	3.9947836
9	0.0007577	0.7511	10.0006716	3.9961424
10	0.0005688	0.7507	10.0002254	3.9971303

Für $p = n$ ergibt sich der QR-Algorithmus.

Algorithmus 10.8

$A_0 = A$.
for $k = 1, 2, \dots$

$$A_{k-1} = Q_k R_k \quad (\text{QR-Zerlegung})$$

$$A_k = R_k Q_k = Q_k^H A_{k-1} Q_k = Q_k^H \dots Q_1^H A_0 Q_1 \dots Q_k.$$

Falls $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$, so konvergieren die Elemente unter der Diagonalen mit den Raten $\left| \frac{\lambda_i}{\lambda_j} \right|^k$ gegen 0.

Beispiel 10.9 Wir betrachten die Matrix A wie vorher.

$\frac{\lambda_i}{\lambda_j} =$	$\frac{4}{10}$	$\frac{3}{10}$	$\frac{3}{4}$
k	$\mathcal{O}(a_{21})$	$\mathcal{O}(a_{31})$	$\mathcal{O}(a_{32})$
1	10^{-1}	10^{-2}	10^{-2}
2	10^{-2}	10^{-2}	10^{-4}
3	10^{-2}	10^{-3}	10^{-3}
4	10^{-3}	10^{-3}	10^{-3}
5	10^{-3}	10^{-3}	10^{-3}
6	10^{-3}	10^{-5}	10^{-3}
7	10^{-4}	10^{-5}	10^{-3}
8	10^{-4}	10^{-6}	10^{-3}
9	10^{-4}	10^{-6}	10^{-3}
10	10^{-5}	10^{-7}	10^{-3}

In der Praxis bringt man die Matrix A durch Drehungen und Spiegelungen erst mal auf Hessenberggestalt:

$$Q^H A Q = \begin{bmatrix} x & x & \cdots & \cdots & x \\ x & x & \cdots & \cdots & x \\ 0 & x & x & \cdots & x \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \cdots & 0 & x & x \end{bmatrix} = A_0 = [a_{ij}^{(0)}].$$

Dann kostet jede Iteration anstatt $\mathcal{O}(n^3)$ nur noch $\mathcal{O}(n^2)$ „flops“.

Außerdem beschleunigt man die Konvergenz, indem man mit $(A - \lambda I)$ rechnet, wobei λ als „Eigenwertnäherung“ ein Eigenwert von

$$\begin{bmatrix} a_{n-1,n-1}^{(i)} & a_{n-1,n}^{(i)} \\ a_{n,n-1}^{(i)} & a_{n,n}^{(i)} \end{bmatrix}$$

ist.

Für symmetrische Matrizen ist die Hessenberggestalt tridiagonal.

$$A_0 = Q^H A Q = \begin{bmatrix} x & x & & & \\ x & x & x & & \\ & x & x & \ddots & \\ & & \ddots & \ddots & x \\ & & & x & x \end{bmatrix}.$$

Dann kostet die Iteration noch $\mathcal{O}(n)$ pro Schritt.

Für reell symmetrische Matrizen (oder auch Hermite'sche) gibt es als ganz andere Methode noch das **Jacobi-Verfahren**.

Also folgt aus

$$\begin{aligned} b_{pp}^2 + b_{qq}^2 + 2b_{pq}^2 &= a_{pp}^2 + a_{qq}^2 + 2a_{pq}^2 \\ b_{pp}^2 + b_{qq}^2 &= a_{pp}^2 + a_{qq}^2 + 2a_{pq}^2 - 2b_{pq}^2 \end{aligned}$$

$$\implies \|B\|_F^2 - \sum_{i=1}^n b_{ii}^2 = \|A\|_F^2 - \sum_{i=1}^n a_{ii}^2 - 2a_{pq}^2 + 2b_{pq}^2.$$

Die Elimination wird so ausgeführt, daß $b_{pq} = 0$. So wird die Summe der Quadrate der Außenelemente um $2a_{pq}^2$ reduziert.

Mit $m = \frac{n(n-1)}{2}$ folgt, da $2[a_{pq}^{(k)}]^2 \geq \frac{1}{m} \text{off}(A^{(k)})$, daß

$$\text{off}(A^{(k+1)}) \leq \left(1 - \frac{1}{m}\right) \text{off}(A^{(k)})$$

und per Induktion

$$\text{off}(A^{(k)}) \leq \left(1 - \frac{1}{m}\right)^k \text{off}(A^{(0)}).$$

Beispiel 10.14

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{bmatrix}$$

k	$\mathcal{O}(\sqrt{\text{off}(A^{(k)})})$
0	10^1
m	10^0
$2m$	10^{-3}
$3m$	10^{-12}
$4m$	$< 10^{-20}$

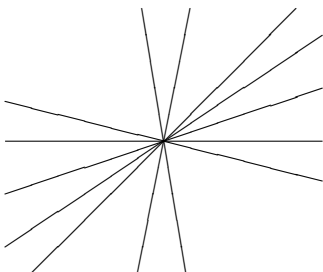
Es gibt viele weitere Varianten – siehe Vorlesungen über numerische Mathematik.

Kapitel 11

Projektive Räume und projektive Geometrie

Wir wollen nun eine weitere wichtige Struktur der linearen Algebra kennenlernen, den projektiven Raum.

Definition 11.1 Sei V ein Vektorraum über einem Körper K und sei $\mathbb{P}(V)$ die Menge der eindimensionalen linearen Unterräume von V . $\mathbb{P}(V)$ heißt der zu V gehörende projektive Raum. Die Elemente von $\mathbb{P}(V)$ nennt man Punkte des projektiven Raums.



Die eindimensionalen Unterräume sind geometrisch die Geraden durch 0.

Man kann nun auch für $\mathbb{P}(V)$ wiederum Unterräume definieren und Dimensionen.

Definition 11.2 Sei V Vektorraum über K .

Ist $U \subset V$ Untervektorraum von V so nennen wir $\mathbb{P}(U) \subset \mathbb{P}(V)$ einen projektiven Unterraum von $\mathbb{P}(V)$.

Falls V endlichdimensional ist, so setzen wir $\dim_K(\mathbb{P}(V)) = \dim_K(V) - 1$ und nennen diese Zahl die projektive Dimension von $\mathbb{P}(V)$.

Beispiel 11.3 $V = \mathbb{R}^n$.

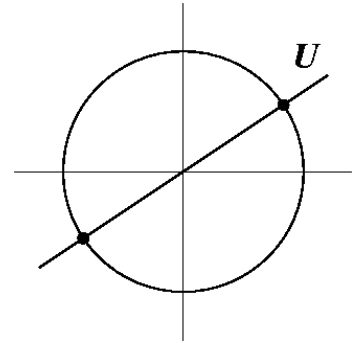
$n = 0$	$\dim \mathbb{P}(V) = -1$	$\mathbb{P}(V) = \phi$
$n = 1$	$\dim \mathbb{P}(V) = 0$	$\mathbb{P}(V)$ hat ein Element
$n = 2$	$\dim \mathbb{P}(V) = 1$	$\mathbb{P}(V)$ heißt <u>projektive Gerade</u>
$n = 3$	$\dim \mathbb{P}(V) = 2$	$\mathbb{P}(V)$ heißt <u>projektive Ebene</u>

Ist $V = K^{n+1}$, so schreiben wir $\mathbb{P}_n(K)$ anstatt $\mathbb{P}(K^{n+1})$ und nennen diesen Raum den kanonischen n -dimensionalen projektiven Raum.

Beispiel 11.4 (1) $K = \mathbb{R}$

Sei

$$S^n = \left\{ \begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n+1} \mid \sum_{i=0}^n x_i^2 = 1 \right\}.$$



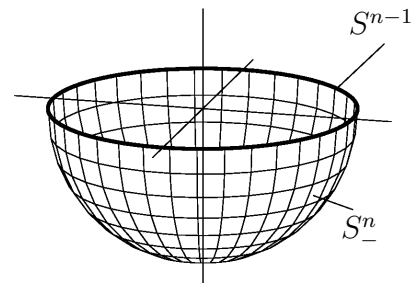
Dies ist die n -Sphäre in \mathbb{R}^{n+1} . Für $n = 2$ ist das gerade die Oberfläche der Einheitskugel (Beachte: $S^n \subset \mathbb{R}^{n+1}$). Sei $U \subset \mathbb{R}^{n+1}$ ein eindimensionaler Unterraum, so ist die Menge $S^n \cap U$ gerade die Menge der beiden Schnittpunkte der Geraden mit der Kugel.

Falls $\begin{bmatrix} z_0 \\ \vdots \\ z_n \end{bmatrix} \in S^n \cap U$, so ist auch $\begin{bmatrix} -z_0 \\ \vdots \\ -z_n \end{bmatrix} \in S^n \cap U$.

Sei $S_-^n = \left\{ \begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n+1} \mid \sum_{i=0}^n x_i^2 = 1, x_n \leq 0 \right\}$

die Oberfläche einer Halbkugel. Es folgt, dass der Rand von S_-^n eine $(n - 1)$ -Sphäre $S^{n-1} \subset S_-^n$ ist.

$\mathbb{P}_n(\mathbb{R})$ entsteht aus S_-^n durch Identifizieren der Antipodalpunkte aus der Randsphäre S^{n-1} . Wir können uns also $\mathbb{P}_1(\mathbb{R})$ als S^1 vorstellen.



Betrachte nun $\mathbb{P}_n(K)$. Führe Koordinaten ein:

Ist $x = \begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix} \in K^{n+1} \setminus \{0\}$, so ist $Kx = \{\lambda x \mid \lambda \in K\}$ ein eindimensionaler Unterraum.

Wir schreiben $(x_0; x_1; \dots; x_n) = K \begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix}$ und nennen dies homogene Koordinaten.
 Offensichtlich gilt

$$(x_0; x_1; \dots; x_n) = (y_0; y_1; \dots; y_n) \iff \exists \lambda \in K \setminus \{0\} \text{ mit } y_i = \lambda x_i \quad \forall i = 0, \dots, n.$$

Homogene Koordinaten sind also nur bis auf einen Faktor $\neq 0$ bestimmt.

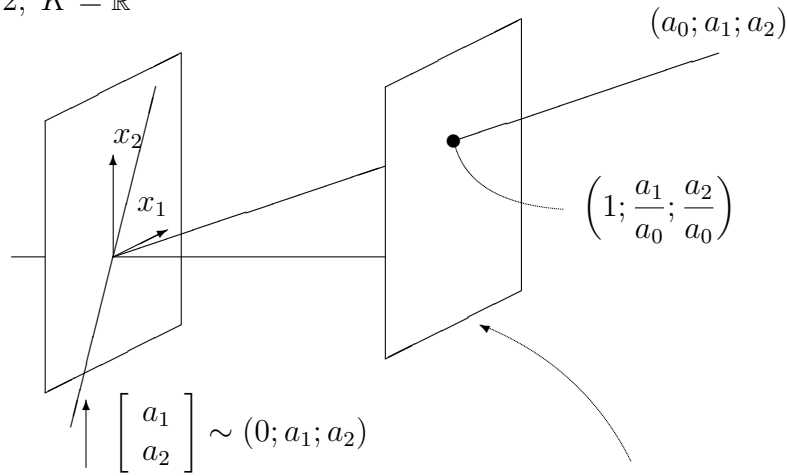
Wir definieren nun eine Einbettung $K^n \rightarrow \mathbb{P}_n(K)$ durch

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \rightarrow (1; x_1; \dots; x_n).$$

Als Bild erhält man gerade die Elemente $(x_0; \dots; x_n) \in \mathbb{P}_n(K)$ mit $x_0 \neq 0$.

Das Komplement des Bildes besteht gerade aus den Punkten $(0; y_1; \dots; y_n)$. Dies ist ein projektiver Unterraum der Dimension $n - 1$ und heißt „uneigentliche Hyperebene“ oder „unendlich ferne Hyperebene“.

Beispiel 11.5 $n = 2, K = \mathbb{R}$



die Punkte der Form $(0; *; *)$,
sie entsprechen den Geradenrichtungen in \mathbb{R}^2 .

Die Punkte in dieser Ebene
entsprechen bijektiv den Punkten von \mathbb{R}^2 .

Wir führen nun einen Abbildungsbegriff ein.

Definition 11.6 Seien V, W K -Vektorräume.

Eine Abbildung $f : \mathbb{P}(V) \rightarrow \mathbb{P}(W)$ heißt projektiv, wenn es eine injektive lineare Abbildung

$F : V \rightarrow W$ gibt, mit

$$f(Kv) = KF(v) \quad \forall v \in V \setminus \{0\}. \quad (11.7)$$

Man schreibt dann $f = \mathbb{P}(F)$.

Eine bijektive projektive Abbildung heißt Projektivität.

Sei $f : \mathbb{P}(V) \rightarrow \mathbb{P}(W)$ eine projektive Abbildung, so ist f injektiv und es gilt

$$\dim \mathbb{P}(W) \geq \dim \mathbb{P}(V),$$

falls V, W endlichdimensional sind.

Eine Menge $\{p_0, \dots, p_r\}$ von Punkten eines projektiven Raums $\mathbb{P}(V)$ heißt projektiv unabhängig, wenn es linear unabhängige Vektoren

$$v_0, \dots, v_r \in V \quad \text{gibt mit} \quad p_i = Kv_i \quad \text{für } i = 0, \dots, r.$$

Eine Menge $\{p_0, \dots, p_{n+1}\}$ von Punkten aus $\mathbb{P}(V)$ heißt projektive Basis von $\mathbb{P}(V)$, wenn je $n + 1$ Punkte projektiv unabhängig sind, wobei $n = \dim \mathbb{P}$.

Beispiel 11.8 $\mathbb{P}_n(K) = \mathbb{P}(K^{n+1})$

Projektive Basis: $p_0 = (1; 0; \dots; 0), \dots, p_n = (0; \dots; 0; 1), p_{n+1} = (1; \dots; 1; 1)$

Lemma 11.9 Sei V ein Vektorraum mit $\dim V = n + 1$. Ist $\{p_0, \dots, p_{n+1}\}$ eine projektive Basis von $\mathbb{P}(V)$, so gibt es eine Basis $\{v_0, \dots, v_n\}$ von V mit $p_i = Kv_i, 0 \leq i \leq n$, und

$$p_{n+1} = K(v_0 + \dots + v_n).$$

Beweis: Da $\{p_0, \dots, p_n\}$ projektiv unabhängig sind, gibt es $\{x_0, \dots, x_n\} \in V$ linear unabhängig mit

$$p_i = Kx_i \quad i = 0, \dots, n.$$

Da diese linear unabhängig sind und $\dim V = n + 1$, bilden die x_i eine Basis von V . Also gibt es

$$\lambda_0, \dots, \lambda_n \in K \text{ mit } p_{n+1} = K(\lambda_0 x_0 + \dots + \lambda_n x_n).$$

Da $p_0, \dots, p_{i-1}, p_{i+1}, \dots, p_{n+1}$ projektiv unabhängig sind, folgt, dass $\lambda_i \neq 0, i = 0, \dots, n$. Mit $v_i = \lambda_i x_i, 0 \leq i \leq n$, erhalten wir, dass

$$p_i = Kv_i, \quad 0 \leq i \leq n \quad \text{und} \quad p_{n+1} = K(v_0 + \dots + v_n).$$

□

Korollar 11.10 Seien $\mathbb{P}(V)$ und $\mathbb{P}(W)$ projektive Räume gleicher Dimension mit projektiven Basen

$$\{p_0, \dots, p_{n+1}\}, \quad \{q_0, \dots, q_{n+1}\}.$$

Dann gibt es genau eine Projektivität

$$f : \mathbb{P}(V) \rightarrow \mathbb{P}(W)$$

mit $f(p_i) = q_i$.

Beweis: Wähle Basen $\{v_0, \dots, v_n\}$ von V und $\{w_0, \dots, w_n\}$ von W wie in Lemma 11.9 und die lineare Abbildung $F : V \rightarrow W$ durch $F(v_i) = w_i$ für $0 \leq i \leq n$.

Sei $v \in V \setminus \{0\}$. $v = \sum_{i=0}^n \lambda_i v_i$, so setze

$$f(Kv) = KFv = K \cdot \sum_{i=0}^n \lambda_i F(v_i) = K \cdot \sum_{i=0}^n \lambda_i w_i.$$

So ist f eine Projektivität.

Seien f, g Projektivitäten mit $f(p_i) = q_i, g(p_i) = q_i$ und dazu $F, G : V \rightarrow W$ mit

$$\begin{aligned} f(Kv) &= KF(v) \\ g(Kv) &= KG(v) \end{aligned} \quad \text{für } v \in V \setminus \{0\}.$$

So gibt es $\lambda_0, \dots, \lambda_n \in K \setminus \{0\}$, $\mu_0, \dots, \mu_n \in K \setminus \{0\}$ mit

$$F(v_i) = \lambda_i w_i, \quad G(v_i) = \mu_i w_i$$

dann

$$\begin{aligned} q_i &= Kw_i = f(p_i) = f(Kv_i) = KF(v_i) \quad \text{und} \\ q_i &= Kw_i = g(p_i) = g(Kv_i) = KG(v_i). \end{aligned}$$

Da

$$\begin{aligned} f(p_{n+1}) &= q_{n+1}, \\ g(p_{n+1}) &= q_{n+1}, \end{aligned}$$

so gibt es $\lambda, \mu \in K \setminus \{0\}$ mit

$$\begin{aligned} F(v_0 + \dots + v_n) &= \lambda(w_0 + \dots + w_n) \\ G(v_0 + \dots + v_n) &= \mu(w_0 + \dots + w_n) \\ \implies \lambda &= \lambda_0 = \dots = \lambda_n \\ \mu &= \mu_0 = \dots = \mu_n \end{aligned}$$

Sei $v \in V \setminus \{0\}$, $v = \sum_{i=0}^n \mu_i v_i$

$$\begin{aligned} f(Kv) &= KF(v) = K \sum_{i=0}^n \lambda_i \mu_i w_i \\ &= K \sum_{i=0}^n \lambda \mu_i w_i = K \sum_{i=0}^n \mu_i w_i = K \sum_{i=0}^n \mu_i \mu w_i \\ &= KG(v) = g(Kv). \end{aligned}$$

$\implies f = g.$

□

Definition 11.11 Sei K ein Körper und V ein Vektorraum über K .

Ein projektives Koordinatensystem in $\mathbb{P}(V)$ ist eine Projektivität

$$k : \mathbb{P}_n(K) \rightarrow \mathbb{P}(V).$$

Ist $p = K(x_0; \dots; x_n) \in \mathbb{P}(V)$, so heißt $(x_0; \dots; x_n)$ homogener Koordinatenvektor des Punktes p .

Nach dem Korollar 11.10 folgt, dass zu jeder projektiven Basis $\{p_0, \dots, p_{n+1}\}$ von $\mathbb{P}(V)$ ein projektives Koordinatensystem

$$k : \mathbb{P}_n(K) \rightarrow \mathbb{P}(V)$$

durch

$$p_0 = k(1; 0; \dots; 0), \dots, p_n = k(0; \dots; 0; 1), p_{n+1} = k(1; \dots; 1)$$

gegeben ist.

Definition 11.12 Sei V ein Vektorraum. Seien X_1, \dots, X_r projektive Unterräume des projektiven Raums $\mathbb{P}(V)$, wobei z.B. $X_i = \mathbb{P}(Z_i), Z_i \subset V$.

(a) Dann setzen wir $X_1 \vee \dots \vee X_r = \mathbb{P}(\sum_{i=1}^r Z_i)$. Dieser Raum heißt Verbindungsraum und ist der kleinste projektive Unterraum von $\mathbb{P}(V)$, der die Vereinigung $\bigcup_{i=1}^r X_i$ enthält.

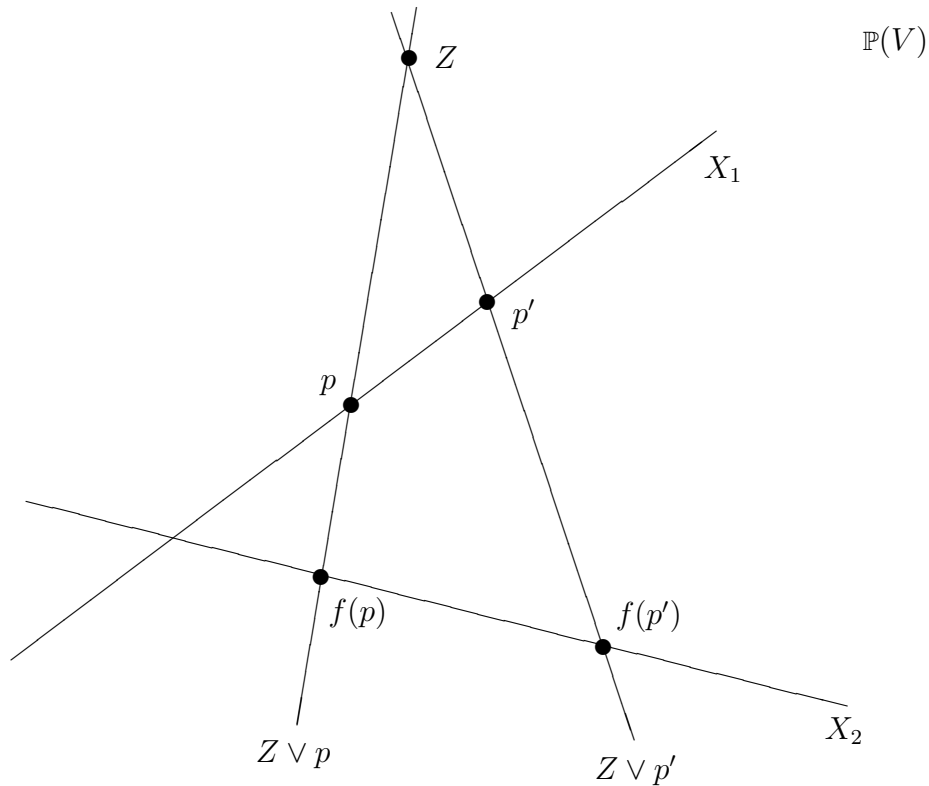
(b) Seien $X_1, X_2 \subset \mathbb{P}(V)$ projektive Unterräume gleicher Dimension. Eine Abbildung $f : X_1 \rightarrow X_2$

heißt Zentralprojektion, wenn es einen projektiven Unterraum $Z \subset \mathbb{P}(V)$ (genannt das Zentrum) gibt, so dass

- (i) $Z \vee X_1 = Z \vee X_2 = \mathbb{P}(V)$
- (ii) $Z \cap X_1 = Z \cap X_2 = \emptyset$
- (iii) Für alle $p \in X_1$ ist $f(p) = (Z \vee p) \cap X_2$.

Beispiel 11.13 Sind $p, q \in \mathbb{P}(V)$, so ist $p \vee q$ eine projektive Gerade, die p und q verbindet. Beachte: eine projektive Gerade ist ein zweidimensionaler Unterraum von V .

Beispiel 11.14



Bedingungen (i) und (ii) in Definition 11.12 besagen, dass Z außerhalb von $X_1 \cup X_2$.

Lemma 11.15 Sei V ein Vektorraum, $X_1, X_2 \subset \mathbb{P}(V)$ projektive Unterräume gleicher Dimension. Sei $f : X_1 \rightarrow X_2$ eine Zentralprojektion. So ist f eine Projektivität.

Beweis: Sei $X_i = \mathbb{P}(W_i)$, $i = 1, 2$. Wir müssen $F : W_1 \rightarrow W_2$ finden, so dass

$$f(Kv) = KF(v) \quad \forall v \in V \setminus \{0\}.$$

Aus Definition 11.12 i) und ii) folgt für das Zentrum $Z = \mathbb{P}(W)$, dass

$$V = W \oplus W_1 = W \oplus W_2.$$

Sei

$$P : V = W \oplus W_2 \rightarrow W_2, \quad \text{mit } P(w + w_2) = w_2, \quad w \in W, w_2 \in W_2$$

und sei F die Einschränkung von P auf W_1 . Zeige, dass F ein Isomorphismus ist.

F ist offensichtlich linear.

Sei nun $w_1 \in W_1$ mit $P(w_1) = 0$. Da $V = W \oplus W_2$, so gibt es $w \in W, w_2 \in W_2$ mit $w_1 = w + w_2$, als

$$0 = P(w_1) = P(w + w_2) = w_2 \quad \text{und damit } w = w_1 \in W.$$

Da $V = W \oplus W_1$, folgt $w_1 = 0$. Also ist F injektiv.

Da $\dim W_1 = \dim W_2$ ist, folgt sofort, dass F ein Isomorphismus ist. Wir müssen nun noch zeigen, dass

$$f(Kv) = KF(v).$$

Sei $p = Kw_1 \in Z_1$ und sei $U \subset V$ der Unterraum, definiert durch $U = W \oplus Kw_1$. Dann ist $\mathbb{P}(U) = Z \vee p$. Nach Definition von F ist

$$F(Kw_1) = U \cap W_2, \quad \text{also}$$

$$f(p) = \mathbb{P}(U \cap W_2) = \mathbb{P}(U) \cap \mathbb{P}(W_2) = (Z \vee p) \cap Z_2.$$

□

Lemma 11.16 *Seien $Z, \tilde{Z} \subset \mathbb{P}_2(K)$ projektive Geraden. Dann gilt $Z \cap \tilde{Z} \neq \emptyset$. (D.h., zwei projektive Geraden in $\mathbb{P}_2(K)$ schneiden sich immer.)*

Beweis: Seien $Z = \mathbb{P}(U)$, $\tilde{Z} = \mathbb{P}(\tilde{U})$, wobei U, \tilde{U} zweidimensionale Unterräume des K^3 sind. Es folgt, dass $Z \cap \tilde{Z} = \mathbb{P}(U \cap \tilde{U})$. Da aber

$$\begin{aligned} \dim U \cap \tilde{U} &= \dim U + \dim \tilde{U} - \dim(U + \tilde{U}) \\ &= 4 - \underbrace{\dim(U + \tilde{U})}_{\leq 3} \geq 1 \end{aligned}$$

ist, folgt $\dim \mathbb{P}(U \cap \tilde{U}) \geq 0$ und somit $Z \cap \tilde{Z} \neq \emptyset$.

□

Wir hatten bei der Einführung der homogenen Koordinaten eine Einbettung

$$K^2 \rightarrow \mathbb{P}_2(K)$$

gemacht und dabei als Komplement die unendlich ferne Gerade eingeführt. Diese enthält gerade die Schnittpunkte von parallelen Geraden in K^2 .

Seien G, \tilde{G} parallele Geraden in K^2

$$G = v + U, \quad \tilde{G} = \tilde{v} + U, \quad \text{wobei } U \subset K^2 \text{ eindimensional und } v \neq \tilde{v}.$$

Seien $X_1, X_2 \in G, \tilde{X}_1, \tilde{X}_2 \in \tilde{G}$ jeweils verschiedene Punkte. Setze

$$\begin{aligned} Z &= X_1 \vee X_2 \subset \mathbb{P}_2(K) \\ \tilde{Z} &= \tilde{X}_1 \vee \tilde{X}_2 \subset \mathbb{P}_2(K) \end{aligned}$$

und damit gilt $G \subset Z, \tilde{G} \subset \tilde{Z}$.

Aus Lemma 11.16 folgt, dass $Z \cap \tilde{Z} \neq \emptyset$. Da $G \cap \tilde{G} = \emptyset$ ist, so muss der Schnittpunkt im Komplement liegen, der unendlich fernen Geraden.

Man kann sich jedoch jede projektive Gerade als unendlich ferne Gerade vorstellen, wie das folgende Lemma zeigt:

Lemma 11.17 *Sei $E \subset K^3$ Unterraum der Dimension 2. Dann ist $\mathbb{P}_2(K) \setminus \mathbb{P}(E)$ bijektiv zu K^2 , d.h. wir können uns $\mathbb{P}(E)$ als unendlich ferne Gerade vorstellen.*

Beweis:

$$E = \left\{ \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} \in K^3 \mid a_0x_0 + a_1x_1 + a_2x_2 = 0 \right\}$$

und mindestens eins von $a_0, a_1, a_2 \in K$ ist ungleich Null. O.B.d.A. sei dies a_0 . Definiere

$$\varphi : \mathbb{P}_2(K) \setminus \mathbb{P}(E) \rightarrow K^2 \quad \text{durch} \quad \varphi(x_0; x_1; x_2) = \begin{bmatrix} \frac{x_1}{a_0x_0 + a_1x_1 + a_2x_2} \\ \frac{x_2}{a_0x_0 + a_1x_1 + a_2x_2} \end{bmatrix}$$

und zeige dann, dass φ bijektiv ist.

Für alle $\lambda \in K \setminus \{0\}$ gilt per Definition von φ , dass $\varphi(x_0; x_1; x_2) = \varphi(\lambda x_0; \lambda x_1; \lambda x_2)$, damit ist φ wohldefiniert.

$$\text{Sei nun } \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in K^2, \text{ so setze } \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{a_0} - \frac{1}{a_0}(a_1y_1 + a_2y_2) \\ y_1 \\ y_2 \end{bmatrix}.$$

$$\text{Dann ist } \varphi(x_0; x_1; x_2) = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Damit folgt, dass jedes Element von K^2 als Bild vorkommt, also ist φ surjektiv.

Für die Injektivität seien

$$(x_0; x_1; x_2), (y_0; y_1; y_2) \in \mathbb{P}_2(K) \quad \text{mit } \varphi(x_0; x_1; x_2) = \varphi(y_0; y_1; y_2), \quad \text{d.h.}$$

$$\begin{bmatrix} \frac{x_1}{a_0x_0 + a_1x_1 + a_2x_2} \\ \frac{x_2}{a_0x_0 + a_1x_1 + a_2x_2} \end{bmatrix} = \begin{bmatrix} \frac{y_1}{a_0y_0 + a_1y_1 + a_2y_2} \\ \frac{y_2}{a_0y_0 + a_1y_1 + a_2y_2} \end{bmatrix}$$

Mit $\lambda = \frac{a_0x_0 + a_1x_1 + a_2x_2}{a_0y_0 + a_1y_1 + a_2y_2} \neq 0$ folgt $x_1 = \lambda y_1, x_2 = \lambda y_2$.

Wir müssen also nur noch $x_0 = \lambda y_0$ zeigen.

Aus $x_1 = \lambda y_1, x_2 = \lambda y_2$ folgt

$$\begin{aligned} a_1 \cdot | \quad a_0x_1y_0 + \underline{a_2x_1y_2} &= a_0y_1x_0 + \underline{a_2y_1x_2} && \text{und} \\ a_2 \cdot | \quad a_0x_2y_0 + \underline{a_1x_2y_1} &= a_0y_2x_0 + \underline{a_1y_2x_1} \end{aligned}$$

Summe:

$$\begin{aligned} a_0a_1x_1y_0 + a_2a_0x_2y_0 &= a_0a_1y_1x_0 + a_0a_2y_2x_0 \\ &\quad + a_0^2x_0y_0 \quad \quad + a_0^2x_0y_0 \\ a_0y_0(a_0x_0 + a_1x_1 + a_2x_2) &= a_0x_0(a_1y_1 + a_0y_0 + a_2y_2) \\ \implies x_0 &= \lambda y_0. \end{aligned}$$

□

Punkte eines projektiven Raumes heißen *kollinear*, wenn sie auf einer projektiven Gerade liegen.

Seien p_0, p_1, p_2, p_3 Punkte in $\mathbb{P}(V)$ über einem Körper K , wobei p_0, p_1, p_2 paarweise verschieden sind und p_0, p_1, p_2, p_3 kollinear.

Dann ist p_0, p_1, p_2 eine Basis der gemeinsamen projektiven Gerade G . Es gibt also in G ein projektives Koordinatensystem

$$k : \mathbb{P}_1(K) \rightarrow G$$

mit

$$k(1; 0) = p_0, \quad k(0; 1) = p_1, \quad k(1, 1) = p_2. \quad (11.18)$$

Seien $(\lambda; \mu) = k^{-1}(p_3)$ die homogenen Koordinaten von p_3 . Dann definieren wir als Doppelverhältnis

$$DV(p_0; p_1; p_2; p_3) = \begin{cases} \frac{\lambda}{\mu} & \mu \neq 0, \text{ d.h. } p_3 \neq p_0 \\ \infty & \mu = 0 \end{cases}$$

Lemma 11.19 *Das Doppelverhältnis bleibt unter Projektivität erhalten.*

Beweis: Seien $p_0, p_1, p_2, p_3 \in \mathbb{P}(V)$ kollinear und p_0, p_1, p_2 paarweise verschieden, und sei $f : \mathbb{P}(V) \rightarrow \mathbb{P}(W)$ eine Projektivität.

Dann sind $f(p_0), f(p_1), f(p_2), f(p_3)$ wieder kollinear. Denn sei G die Gerade, auf der p_0, p_1, p_2, p_3 liegen, so liegen die Bilder auf $f(G)$ und $f(G)$ ist wieder eine projektive Gerade.

Zeige nun

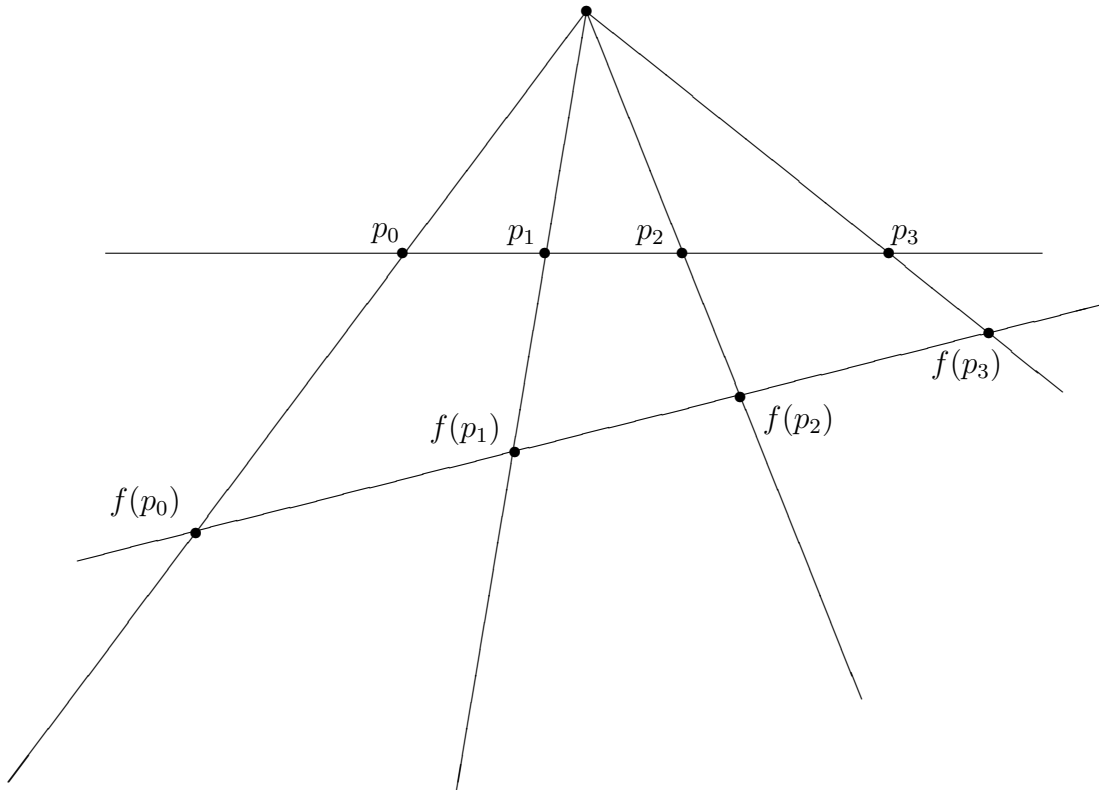
$$DV(p_0, p_1, p_2, p_3) = DV(f(p_0), f(p_1), f(p_2), f(p_3))$$

Sei $k : \mathbb{P}_1(K) \rightarrow G$ ein projektives Koordinatensystem wie in (11.18) und $\tilde{k} : \mathbb{P}_1(K) \rightarrow f(G)$ ein projektives Koordinatensystem mit

$$\tilde{k}(1; 0) = f(p_0), \quad \tilde{k}(0; 1) = f(p_1), \quad \tilde{k}(1; 1) = f(p_2).$$

Dann ist $f|_G \circ k = \tilde{k}$, denn es hat die gleichen Eigenschaften wie \tilde{k} und damit folgt die Behauptung wie in Korollar 11.10.

Da $k^{-1}(p) = \tilde{k}^{-1}(f(p))$ folgt dann auch, dass die Doppelverhältnisse gleich sind. □



$$\begin{aligned} DV(p_0, p_1, p_2, p_3) &= \frac{\overline{p_0 p_2}}{\overline{p_2 p_1}} : \frac{\overline{p_0 p_3}}{\overline{p_3 p_1}} \\ &= \frac{\overline{f(p_0) f(p_2)}}{\overline{f(p_2) f(p_1)}} : \frac{\overline{f(p_0) f(p_3)}}{\overline{f(p_3) f(p_1)}} \end{aligned}$$

Wir können das Doppelverhältnis mit Hilfe von Determinanten beschreiben.

Lemma 11.20 *Gegeben sind vier Punkte*

$p_i = (\lambda_i; \mu_i) \in \mathbb{P}_1(K)$, $i = 0, 1, 2, 3$ so dass p_0, p_1, p_2 paarweise verschieden sind. Dann ist

$$DV(p_0, p_1, p_2, p_3) = \frac{\det \begin{bmatrix} \lambda_3 & \lambda_1 \\ \mu_3 & \mu_1 \end{bmatrix} \cdot \det \begin{bmatrix} \lambda_0 & \lambda_2 \\ \mu_0 & \mu_2 \end{bmatrix}}{\det \begin{bmatrix} \lambda_0 & \lambda_3 \\ \mu_0 & \mu_3 \end{bmatrix} \cdot \det \begin{bmatrix} \lambda_2 & \lambda_1 \\ \mu_2 & \mu_1 \end{bmatrix}}$$

Beweis: Sei $k : \mathbb{P}_1(K) \rightarrow \mathbb{P}_1(K)$ ein projektives Koordinatensystem mit

$$k(1; 0) = (\lambda_0; \mu_0), \quad k(0; 1) = (\lambda_1; \mu_1), \quad k(1; 1) = (\lambda_2; \mu_2).$$

Sei $F : K^2 \rightarrow K^2$ mit $k = \mathbb{P}(F)$. Dann wird F bezüglich der kanonischen Basis durch folgende Matrix A beschrieben.

Man rechnet dann sofort nach, dass

$$A = \begin{bmatrix} \lambda_0 \det \begin{bmatrix} \lambda_2 & \lambda_1 \\ \mu_2 & \mu_1 \end{bmatrix} & \lambda_1 \det \begin{bmatrix} \lambda_0 & \lambda_2 \\ \mu_0 & \mu_2 \end{bmatrix} \\ \mu_0 \det \begin{bmatrix} \lambda_2 & \lambda_1 \\ \mu_2 & \mu_1 \end{bmatrix} & \mu_1 \det \begin{bmatrix} \lambda_0 & \lambda_2 \\ \mu_0 & \mu_2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix},$$

$$A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \alpha \begin{bmatrix} \lambda_0 \\ \mu_0 \end{bmatrix}, \quad A \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \beta \begin{bmatrix} \lambda_1 \\ \mu_1 \end{bmatrix}, \quad A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \gamma \begin{bmatrix} \lambda_2 \\ \mu_2 \end{bmatrix}.$$

Das Doppelverhältnis erhält man dann über

$$A^{-1} \begin{bmatrix} \lambda_3 \\ \mu_3 \end{bmatrix} = \frac{1}{\det A} \cdot \begin{bmatrix} a_4 & -a_2 \\ -a_3 & a_1 \end{bmatrix} \begin{bmatrix} \lambda_3 \\ \mu_3 \end{bmatrix}$$

Den konstanten Faktor können wir weglassen und erhalten

$$\begin{bmatrix} a_4 & -a_2 \\ -a_3 & a_1 \end{bmatrix} \begin{bmatrix} \lambda_3 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} (\lambda_3 \mu_1 - \mu_3 \lambda_1) (\lambda_0 \mu_2 - \mu_0 \lambda_2) \\ (\lambda_0 \mu_3 - \mu_0 \lambda_3) (\lambda_2 \mu_1 - \mu_2 \lambda_1) \end{bmatrix}$$

Das ist das gesuchte Verhältnis. □

Definition 11.21 *Seien p_0, p_1, p_2, p_3 vier verschiedene kollineare Punkte in $\mathbb{P}(V)$, so liegen (p_0, p_1) und (p_2, p_3) harmonisch, falls*

$$DV(p_0, p_1, p_2, p_3) = -1.$$

Lemma 11.22 *Seien p_0, p_1, p_2, p_3 vier verschiedene kollineare Punkte in $\mathbb{P}(V)$ so sind (p_0, p_1) und (p_2, p_3) harmonisch, wenn*

$$DV(p_0, p_1, p_2, p_3) = DV(p_1, p_0, p_2, p_3)$$

Beweis: Sei $k : \mathbb{P}_1(K) \rightarrow \mathbb{P}(V)$ projektives Koordinatensystem der gemeinsamen Gerade G und seien $q_i = (\lambda_i; \mu_i) \in \mathbb{P}_1(K)$ die Urbilder von p_i , $i = 0, 1, 2, 3$.

Wegen der Invarianz des Doppelverhältnisses reicht es, die Behauptung für q_i , $i = 0, 1, 2, 3$, zu zeigen.

Aus Lemma 11.20 folgt, falls $DV(q_0, q_1, q_2, q_3) = \lambda$, so ist

$$DV(q_1, q_0, q_2, q_3) = \frac{\det \begin{bmatrix} \lambda_3 & \lambda_0 \\ \mu_3 & \mu_0 \end{bmatrix} \cdot \det \begin{bmatrix} \lambda_1 & \lambda_2 \\ \mu_1 & \mu_2 \end{bmatrix}}{\det \begin{bmatrix} \lambda_1 & \lambda_3 \\ \mu_1 & \mu_3 \end{bmatrix} \cdot \det \begin{bmatrix} \lambda_2 & \lambda_0 \\ \mu_2 & \mu_0 \end{bmatrix}} = \lambda^{-1}$$

Ist $DV(q_0, q_1, q_2, q_3) = -1$, so ist $DV(q_1, q_0, q_2, q_3) = (-1)^{-1} = -1$.

Umgekehrt

$$DV(q_0, q_1, q_2, q_3) = DV(q_1, q_0, q_2, q_3) \implies DV(q_0, q_1, q_2, q_3)^2 = 1.$$

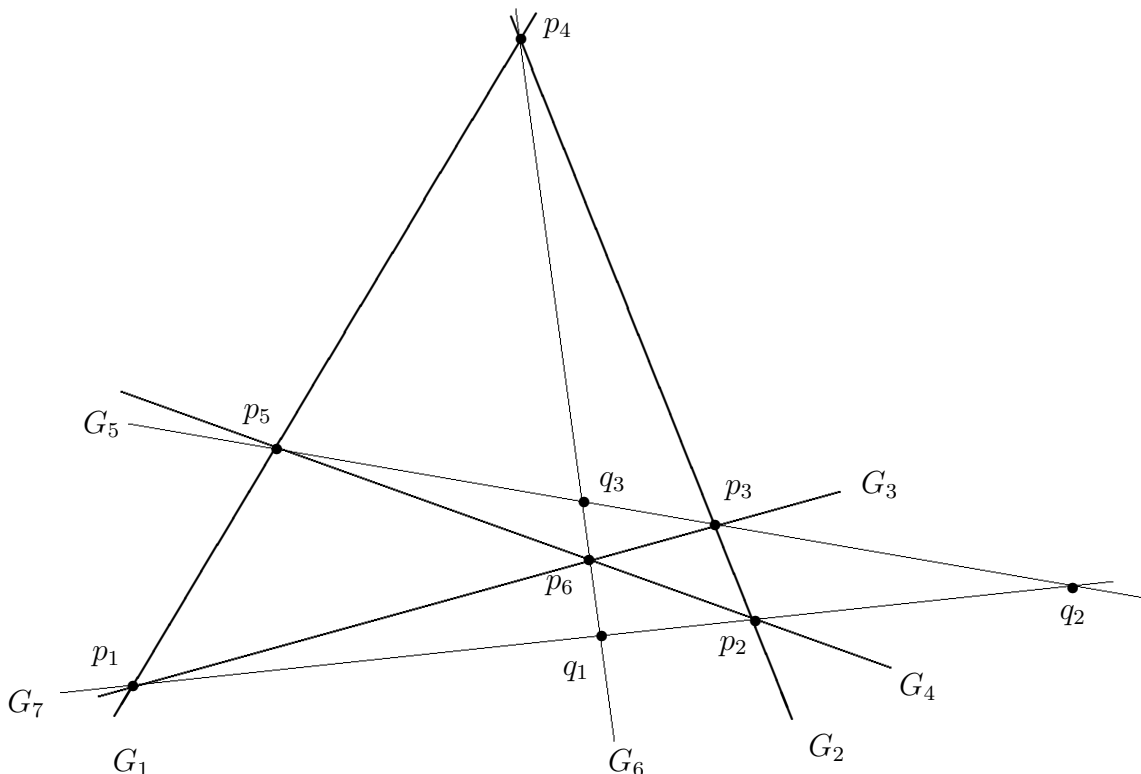
Sei $q_3 = (\lambda_3; \mu_3)$.

Da $(\lambda_3; \mu_3) \neq (1, 1) = (\lambda_2, \mu_2)$, folgt $DV(q_0, q_1, q_2, q_3) = -1$. □

Eine klassische Aufgabe der Geometrie ist es nun, zu 3 kollinearen Punkten einen vierten Punkt zu finden, der harmonisch dazu liegt (*und dies nur mit Hilfe eines Lineals*).

Betrachte dazu 4 Geraden in allgemeiner Lage in einer projektiven Ebene, d.h. keine 3 Geraden haben einen gemeinsamen Schnittpunkt. So etwas heißt *vollständiges Vierseit*.

Seien p_1, \dots, p_6 die Schnittpunkte. Bilde dann die Verbindungsgeraden G_5, G_6, G_7 zwischen den Ecken. Die Schnittpunkte der Diagonalen seien q_1, q_2, q_3 .



Satz 11.23 Auf jeder Diagonalen eines vollständigen Vierseits liegen die Punktepaare $(p_i, p_j), (q_k, q_l)$, bestehend aus Ecken und Diagonalschnittpunkten harmonisch.

Beweis: Zeige exemplarisch $DV(p_1, p_2, q_1, q_2) = -1$.

Wir verwenden dazu die Zentralprojektion mit dem Zentrum p_4 . Nach dem Satz über das Doppelverhältnis 11.19 gilt

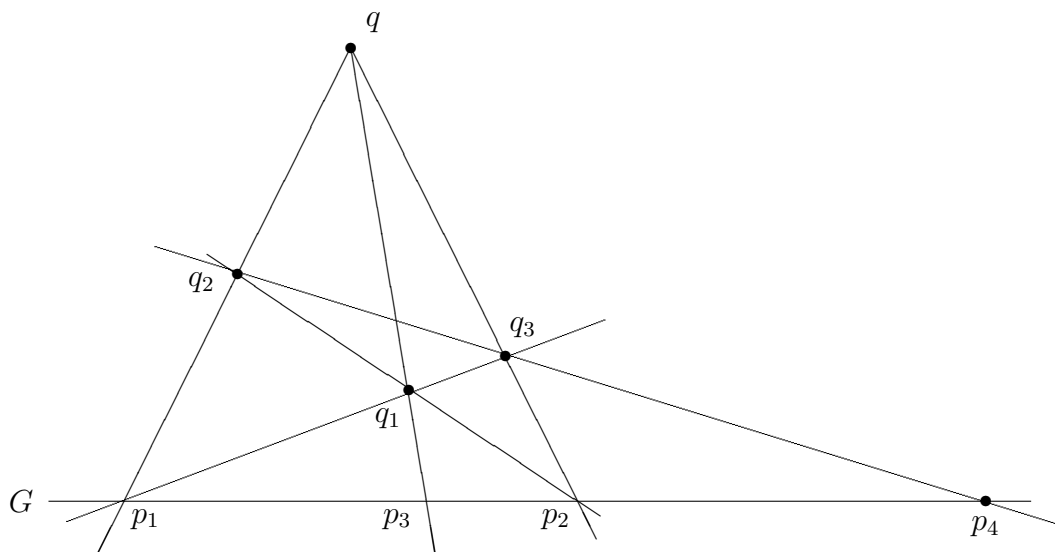
$$DV(p_1, p_2, q_1, q_2) = DV(p_5, p_3, q_3, q_2).$$

Die Zentralprojektion mit dem Zentrum P_6 ergibt

$$DV(p_5, p_3, q_3, q_2) = DV(p_2, p_1, q_1, q_2).$$

Damit folgt die Behauptung aus Lemma 11.22. □

Damit können wir die gewünschte Konstruktionsaufgabe lösen. Seien $p_1, p_2, p_3 \in G$ Punkte auf einer Geraden G . Wähle $q \notin G$.



Betrachte Geraden: $q \vee p_1, q \vee p_2, q \vee p_3$.

Wähle Punkt $q_1 \in q \vee p_3, q_1 \neq q, p_3$ und bilde $p_1 \vee q_1, p_2 \vee q_1$.

Sei $q_2 = (p_2 \vee q_1) \cap (q \vee p_1), q_3 = (p_1 \vee q_1) \cap (q \vee p_2)$ und $p_4 = (q_2 \vee q_3) \cap G$.

Dann sind $(p_1, p_2), (p_3, p_4)$ harmonisch.

Korollar 11.24 Seien p_0, p_1, p_2, p_3, p_4 Punkte einer projektiven Geraden, so dass p_0, p_1, p_2 paarweise verschieden.

Ist $DV(p_0, p_1, p_2, p_3) = DV(p_0, p_1, p_2, p_4)$, so folgt $p_3 = p_4$.

Beweis: (Übung) □

Zwei weitere wichtige Sätze der ebenen Geometrie erhalten wir jetzt als Folgerungen.

Satz 11.25 (Satz von Desarque)

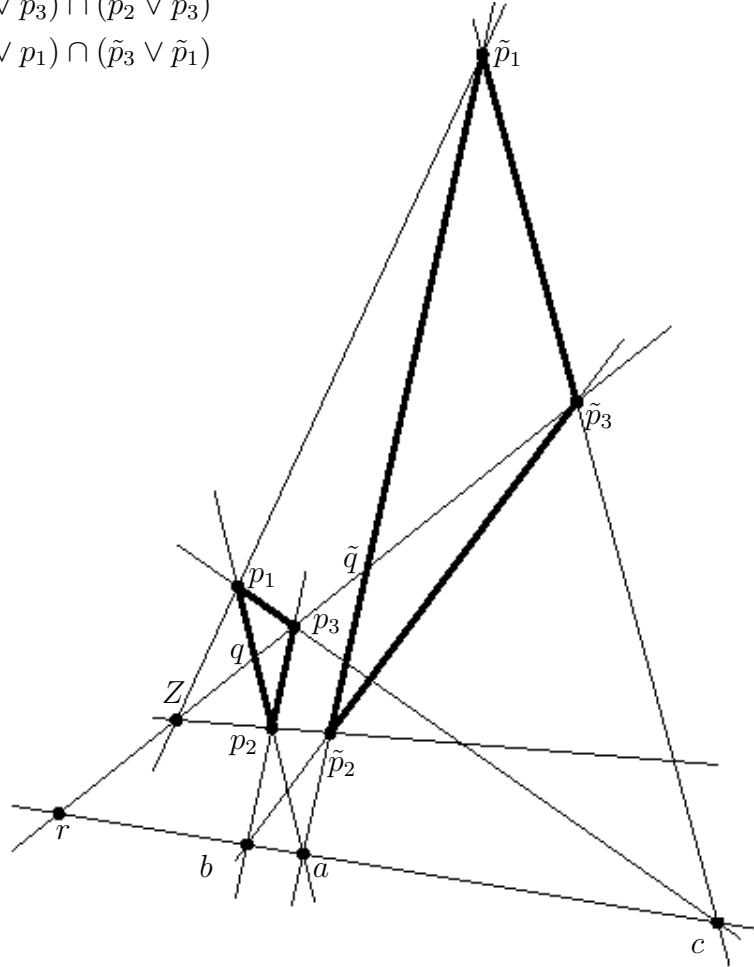
Sei K ein Körper und seien zwei Dreiecke in $\mathbb{P}_2(K)$ gegeben, die in perspektivischer Lage zueinander liegen, d.h. es gibt Punkte $p_1, p_2, p_3, \tilde{p}_1, \tilde{p}_2, \tilde{p}_3$, so dass sich $p_1 \vee \tilde{p}_1, p_2 \vee \tilde{p}_2, p_3 \vee \tilde{p}_3$ in einem Punkt Z schneiden. Dann sind die Schnittpunkte

$$a = (p_1 \vee p_2) \cap (\tilde{p}_1 \vee \tilde{p}_2)$$

$$b = (p_2 \vee p_3) \cap (\tilde{p}_2 \vee \tilde{p}_3)$$

$$c = (p_3 \vee p_1) \cap (\tilde{p}_3 \vee \tilde{p}_1)$$

kollinear.



Beweis: Sei

$$q = (p_1 \vee p_2) \cap (p_3 \vee \tilde{p}_3),$$

$$\tilde{q} = (\tilde{p}_1 \vee \tilde{p}_2) \cap (p_3 \vee \tilde{p}_3),$$

$$r = (a \vee c) \cap (p_3 \vee \tilde{p}_3),$$

$$b' = (a \vee c) \cap (p_2 \vee p_3),$$

$$b'' = (a \vee c) \cap (\tilde{p}_2 \vee \tilde{p}_3).$$

Zu zeigen ist $b = b' = b''$.

Zeige zuerst, dass $b' = b''$. Dafür zeigen wir, dass $DV(a, c, r, b') = DV(a, c, r, b'')$.

Für die Zentralprojektion mit dem Zentrum p_3 gilt

$$DV(a, c, r, b') = DV(a, p_1, q, p_2).$$

Analog folgt für die Zentralprojektion mit dem Zentrum Z , dass

$$DV(a, p_1, q, p_2) = DV(a, \tilde{p}_1, \tilde{q}, \tilde{p}_2)$$

und für die Zentralprojektion mit dem Zentrum \tilde{p}_3

$$DV(a, \tilde{p}_1, \tilde{q}, \tilde{p}_2) = DV(a, c, r, b'').$$

Also folgt aus Korollar 11.24, dass $b' = b''$. Analog folgt $b = b'$. \square

Satz 11.26 (Satz von Pappos)

Sei K ein Körper und seien in $\mathbb{P}_2(K)$ zwei Geraden G, \tilde{G} gegeben und darauf paarweise verschiedene Punkte

$$p_1, p_2, p_3 \in G, \quad \tilde{p}_1, \tilde{p}_2, \tilde{p}_3 \in \tilde{G}.$$

Dann sind

$$a = (p_1 \vee \tilde{p}_2) \cap (\tilde{p}_1 \vee p_2)$$

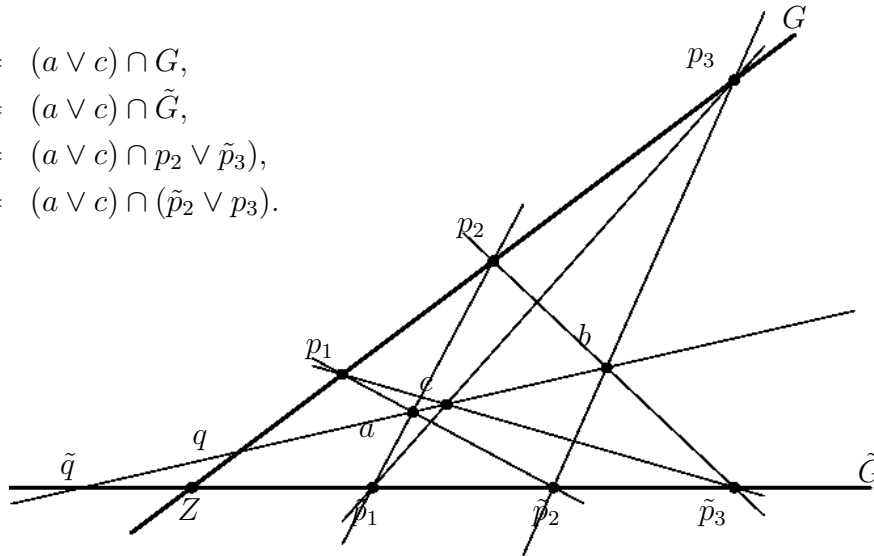
$$b = (p_2 \vee \tilde{p}_3) \cap (\tilde{p}_2 \vee p_3)$$

$$c = (p_3 \vee \tilde{p}_1) \cap (\tilde{p}_3 \vee p_1)$$

kollinear.

Beweis: Sei $Z = G \cap \tilde{G}$ (existiert nach Lemma 11.16),

$$\begin{aligned} q &= (a \vee c) \cap G, \\ \tilde{q} &= (a \vee c) \cap \tilde{G}, \\ b' &= (a \vee c) \cap p_2 \vee \tilde{p}_3, \\ b'' &= (a \vee c) \cap (\tilde{p}_2 \vee p_3). \end{aligned}$$



Zeige $b' = b''$ und dazu $DV(q, c, \tilde{q}, b') = DV(q, c, \tilde{q}, b'')$.

Verwende dazu Zentralprojektionen

$$f : a \vee c \rightarrow G \quad \text{mit Zentrum } \tilde{p}_3$$

$$g : G \rightarrow \tilde{G} \quad \text{mit Zentrum } a$$

$$h : \tilde{G} \rightarrow a \vee c \quad \text{mit Zentrum } p_3$$

Es folgt

$$\begin{aligned} DV(q, c, \tilde{q}, b') &\stackrel{f}{=} DV(q, p_1, Z, p_2) \stackrel{g}{=} DV(\tilde{q}, \tilde{p}_2, Z, \tilde{p}_1) \stackrel{h}{=} DV(\tilde{q}, b'', q, c) \\ &= DV(q, c, \tilde{q}, b'') \end{aligned}$$

\square

Kapitel 12

Das Kroneckerprodukt

Wir wollen nun ein weiteres wichtiges Hilfsmittel der linearen Algebra kennenlernen, welches, wie wir sehen werden, in der Stabilitätstheorie dynamischer Systeme eine große Bedeutung hat.

Definition 12.1 Sei K ein Körper und seien

$$A = [a_{ij}] \in K^{m,m}, B = [b_{ij}] \in K^{n,n}$$

so definieren wir durch

$$\otimes : (K^{m,m}, K^{n,n}) \rightarrow K^{m \cdot n, m \cdot n}$$
$$(A, B) \mapsto A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mm}B \end{bmatrix}.$$

das Tensor- oder Kroneckerprodukt.

Beispiel 12.2

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}, \quad \implies \quad A \otimes B \in K^{6,6},$$

$$A \otimes B = \left[\begin{array}{ccc|ccc} a_{11}b_{11} & a_{11}b_{12} & a_{11}b_{13} & a_{12}b_{11} & a_{12}b_{12} & a_{12}b_{13} \\ a_{11}b_{21} & a_{11}b_{22} & a_{11}b_{23} & a_{12}b_{21} & a_{12}b_{22} & a_{12}b_{23} \\ a_{11}b_{31} & a_{11}b_{32} & a_{11}b_{33} & a_{12}b_{31} & a_{12}b_{32} & a_{12}b_{33} \\ \hline a_{21}b_{11} & a_{21}b_{12} & a_{21}b_{13} & a_{22}b_{11} & a_{22}b_{12} & a_{22}b_{13} \\ a_{21}b_{21} & a_{21}b_{22} & a_{21}b_{23} & a_{22}b_{21} & a_{22}b_{22} & a_{22}b_{23} \\ a_{21}b_{31} & a_{21}b_{32} & a_{21}b_{33} & a_{22}b_{31} & a_{22}b_{32} & a_{22}b_{33} \end{array} \right].$$

Wir wollen nun einige Rechenregeln kennenlernen.

Lemma 12.3 *Seien im folgenden die Matrixdimensionen so, dass die Summen und Produkte definiert sind, und sei $\mu \in K$. Dann gilt*

$$\begin{aligned} (i) \quad & (\mu A) \otimes B = A \otimes (\mu B) = \mu(A \otimes B), \\ (ii) \quad & (A + B) \otimes C = (A \otimes C) + (B \otimes C), \\ (iii) \quad & A \otimes (B + C) = (A \otimes B) + (A \otimes C), \\ (iv) \quad & A \otimes (B \otimes C) = (A \otimes B) \otimes C, \\ (v) \quad & (A \otimes B)^\top = A^\top \otimes B^\top. \end{aligned}$$

Beweis: Übungsaufgabe. □

Beachte, dass sich in (v) bei diesem Produkt die Reihenfolge beim Transponieren nicht umdreht.

Den Zusammenhang zwischen Matrixmultiplikation und Kroneckerprodukt gibt das folgende Lemma.

Lemma 12.4 *Seien $A, C \in K^{m,m}$, $B, D \in K^{n,n}$. So gilt*

$$(A \otimes B) \cdot (C \otimes D) = (AC) \otimes (BD). \quad (12.5)$$

Beweis:

$$\begin{aligned} A \otimes B &= [a_{ij} B] \\ C \otimes D &= [c_{ij} D] \end{aligned}$$

Dann ist der Block (i, j) in der Blockmatrix $[F_{ij}] = (A \otimes B) \cdot (C \otimes D)$ mit $F_{ij} \in K^{n,n}$ gegeben durch

$$F_{ij} = \sum_{k=1}^m (a_{ik} B)(c_{kj} D) = \sum_{k=1}^m a_{ik} c_{kj} \cdot B \cdot D = \left(\sum_{k=1}^m a_{ik} c_{kj} \right) BD.$$

Für die Blockmatrix $[G_{ij}] = (A \cdot C) \otimes (B \cdot D)$ mit $G_{ij} \in K^{n,n}$ erhalten wir

$$G_{ij} = g_{ij} BD \quad \text{mit} \quad g_{ij} = \sum_{k=1}^m a_{ik} c_{kj}.$$

Daraus folgt die Behauptung. □

Aus diesem Lemma folgen einige wichtige Konsequenzen:

Korollar 12.6 *Seien $A \in K^{m,m}$, $B \in K^{n,n}$ so gilt:*

$$\begin{aligned} (i) \quad A \otimes B &= (A \otimes I_n) \cdot (I_m \otimes B) \\ &= (I_m \otimes B) \cdot (A \otimes I_n) \end{aligned}$$

$$(ii) \quad (A \otimes B)^{-1} = A^{-1} \otimes B^{-1}, \quad \text{falls } A^{-1}, B^{-1} \text{ existieren.}$$

D.h., $I_m \otimes B$ und $A \otimes I_n$ sind vertauschbar und die Inverse des Kroneckerprodukts ist das Kroneckerprodukt (in gleicher Reihenfolge) der Inversen.

Korollar 12.7 Seien $A_1, \dots, A_r \in K^{m,m}$, $B_1, \dots, B_r \in K^{n,n}$. So gilt

$$(A_1 \otimes B_1)(A_2 \otimes B_2) \dots (A_r \otimes B_r) = (A_1 \cdot \dots \cdot A_r) \otimes (B_1 \cdot \dots \cdot B_r).$$

Im allgemeinen ist das Kroneckerprodukt nicht kommutativ, aber es gilt:

Lemma 12.8 Seien $A \in K^{m,m}$, $B \in K^{n,n}$, so gibt es eine Permutationsmatrix $P \in K^{mn,mn}$, so dass

$$P^\top(A \otimes B)P = B \otimes A.$$

Beweis: Sei P die Permutationsmatrix, so dass $P^\top(A \otimes I_n)P = I_n \otimes A$.

P erhält man über die Permutation

$$\begin{aligned} \sigma(1, \dots, m \cdot n) = & (1, \quad n+1, \quad 2n+1, \quad \dots, \quad (m-1) \cdot n+1, \\ & 2, \quad n+2, \quad 2n+2, \quad \dots, \quad (m-1) \cdot n+2, \\ & \quad \quad \quad \dots, \\ & n, \quad 2n, \quad 3n, \quad \dots, \quad m \cdot n). \end{aligned}$$

Es gilt außerdem $P^\top(I_m \otimes B)P = B \otimes I_m$.

Aus Korollar 12.6 (i) folgt dann

$$\begin{aligned} P^\top(A \otimes B)P &= P^\top(A \otimes I_n) \cdot (I_m \otimes B)P \\ &= P^\top(A \otimes I_n)PP^\top(I_m \otimes B)P \\ &= (I_n \otimes A) \cdot (B \otimes I_m) \stackrel{12.6(i)}{=} (B \otimes I_m) \cdot (I_n \otimes A) \\ &= B \otimes A. \end{aligned}$$

□

Korollar 12.9 Seien $A \in K^{m,m}$, $B \in K^{n,n}$. So gilt

$$\det(A \otimes B) = (\det A)^n \cdot (\det B)^m$$

Beweis: Aus Korollar 12.6 (i) folgt

$$\det(A \otimes B) = \det((A \otimes I_n) \cdot (I_m \otimes B)) = \det(A \otimes I_n) \cdot \det(I_m \otimes B).$$

Nun gilt aber

$$\det(A \otimes I_n) \stackrel{(12.8)}{=} \det(P(I_n \otimes A)P^\top) = \det(I_n \otimes A) = (\det A)^n$$

und $\det(I_m \otimes B) = (\det B)^m$.

□

Die Abbildung \otimes geht von $(K^{m,m}, K^{n,n})$ nach $K^{m \cdot n, m \cdot n}$. Wir brauchen jetzt auch noch eine Abbildung von $K^{m,n}$ in $K^{m \cdot n}$. Dies ist die vec -Operation, gegeben durch

$$\begin{aligned} \text{vec} : K^{m,n} &\rightarrow K^{m \cdot n} \\ A = [a_{ij}] &\mapsto [\tilde{a}_i] \\ \text{vec}(A) &:= \begin{bmatrix} \tilde{a}_1 \\ \vdots \\ \tilde{a}_{m \cdot n} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{12} \\ \vdots \\ a_{m2} \\ \vdots \\ a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} \end{aligned} \quad (12.10)$$

vec ist eine lineare Abbildung, denn

$$\text{vec}(\alpha A + \beta B) = \alpha \text{vec}(A) + \beta \text{vec}(B) \quad (12.11)$$

für alle $A, B \in K^{m,n}$, $\alpha, \beta \in K$.

Lemma 12.12 *Die Matrizen*

$$A_1, \dots, A_k \in K^{m,n}$$

sind linear unabhängig als Elemente des Vektorraums $K^{m \cdot n}$ genau dann, wenn

$$\text{vec}(A_1), \dots, \text{vec}(A_k)$$

linear unabhängig als Elemente des Vektorraums $K^{m \cdot n}$ sind.

Beweis: Übungsaufgabe. □

Den Zusammenhang zwischen \otimes und vec ergibt der folgende Satz.

Satz 12.13 *Sei $A = [a_{ij}] \in K^{m,m}$, $B = [b_{kj}] \in K^{n,n}$, $X \in K^{m,n}$ so gilt*

$$\text{vec}(AXB) = (B^\top \otimes A) \text{vec}(X).$$

Beweis: Für $j = 1, \dots, n$ ist die j -te Spalte von AXB gegeben durch

$$\begin{aligned} (AXB)e_j &= (AX)(Be_j) = \sum_{k=1}^n b_{kj} (AX)e_k \\ &= \sum_{k=1}^n (b_{kj}A)(Xe_k) \end{aligned}$$

$$\implies (AXB)e_j = [b_{1j}A, b_{2j}A, \dots, b_{nj}A] \text{vec}(X).$$

□

Korollar 12.14 Seien $A \in K^{m,m}$, $B \in K^{n,n}$, $X \in K^{m,n}$. So gilt:

- (i) $\text{vec}(AX) = (I_n \otimes A) \text{vec}(X)$,
- (ii) $\text{vec}(XB) = (B^\top \otimes I_m) \text{vec}(X)$,
- (iii) $\text{vec}(AX + XB) = [(I_n \otimes A) + (B^\top \otimes I_m)] \text{vec}(X)$,
- (iv) $\text{tr}(A \otimes B) = \text{tr}(A) \cdot \text{tr}(B)$,
- (v) $\text{Rang}(A \otimes B) = \text{Rang}(A) \cdot \text{Rang}(B)$.

Beweis: (i) – (iii) sind triviale Folgerungen aus Satz 12.13 und der Definition von vec .
 (iv) $(A \otimes B) = [a_{ij}B]$

$$\implies \text{tr}(A \otimes B) = \sum_{i=1}^m \sum_{j=1}^n a_{ii} b_{jj} = \left(\sum_{i=1}^m a_{ii} \right) \left(\sum_{j=1}^n b_{jj} \right) = \text{tr}(A) \cdot \text{tr}(B).$$

(v) Seien P_A, Q_A, P_B, Q_B nichtsinguläre Matrizen, so dass

$$P_A A Q_A = \begin{bmatrix} I_{r_1} & 0 \\ 0 & 0 \end{bmatrix} = R_A$$

$$P_B B Q_B = \begin{bmatrix} I_{r_2} & 0 \\ 0 & 0 \end{bmatrix} = R_B$$

$$\text{Rang}(A) \cdot \text{Rang}(B) = \text{Rang}(P_A A Q_A) \cdot \text{Rang}(P_B B Q_B) = r_1 \cdot r_2,$$

$$\begin{aligned} \text{Rang}(A \otimes B) &= \text{Rang} \left[(P_A^{-1} R_A Q_A^{-1}) \otimes (P_B^{-1} R_B Q_B^{-1}) \right] \\ &\stackrel{12.5}{=} \text{Rang} \left[(P_A^{-1} \otimes P_B^{-1}) \cdot (R_A \otimes R_B) \cdot (Q_A^{-1} \otimes Q_B^{-1}) \right] \\ &= \text{Rang} \left[(P_A \otimes P_B)^{-1} \cdot (R_A \otimes R_B) \cdot (Q_A \otimes Q_B)^{-1} \right] \\ &= \text{Rang}(R_A \otimes R_B) = r_1 \cdot r_2. \end{aligned}$$

□

Als nächstes werden wir nun den Zusammenhang zwischen den Eigenwerten von A, B und $A \otimes B$ studieren.

Sei $p(x, y) = \sum_{i,j=0}^l c_{ij} x^i y^j$ ein Polynom in zwei Variablen und seien $A \in \mathbb{C}^{m,m}$, $B \in \mathbb{C}^{n,n}$.

Setze

$$p(A; B) := \sum_{i,j=0}^l c_{ij} A^i \otimes B^j. \quad (12.15)$$

Beispiel 12.16 Ist $p(x, y) = 2x + 3xy^2 = 2x^1 y^0 + 3x^1 y^2$

$$p(A; B) = 2A \otimes I_n + 3A \otimes B^2.$$

Dann haben wir den folgenden fundamentalen Satz.

Satz 12.17 (Satz von Stephanos)

Sei $p(A; B)$ wie in (12.15) und seien $\lambda_1, \dots, \lambda_m$ die Eigenwerte von $A \in \mathbb{C}^{m,m}$ und μ_1, \dots, μ_n die Eigenwerte von $B \in \mathbb{C}^{n,n}$, so sind die $m \cdot n$ Zahlen

$$p(\lambda_r, \mu_s), \quad r = 1, \dots, m, \quad s = 1, \dots, n$$

die Eigenwerte von $p(A; B)$.

Beweis: Seien $P \in \mathbb{C}^{m,m}$, $Q \in \mathbb{C}^{n,n}$, so dass $P^{-1}AP = J_A$ und $Q^{-1}BQ = J_B$ in Jordan'scher Normalform sind.

Da J_A und J_B obere Dreiecksmatrizen sind, so sind auch J_A^i, J_B^j und $J_A^i \otimes J_B^j$ obere Dreiecksmatrizen. Die Eigenwerte von J_A^i sind λ_r^i und die von J_B^j sind μ_s^j .

Damit sind die Diagonalelemente von $p(J_A; J_B)$ gerade $p(\lambda_r, \mu_s)$, $r = 1, \dots, m$, $s = 1, \dots, n$.

$$\begin{aligned} p(A; B) &= \sum_{i,j=0}^l c_{ij} (P J_A P^{-1})^i \otimes (Q J_B Q^{-1})^j \\ &\stackrel{12.3/12.5}{=} \sum_{i,j=0}^l c_{ij} (P \otimes Q) \cdot (J_A^i \otimes J_B^j) (P^{-1} \otimes Q^{-1}) \\ &= (P \otimes Q) \left(\sum_{i,j=0}^l c_{ij} (J_A^i \otimes J_B^j) \right) (P \otimes Q)^{-1} \end{aligned}$$

Damit folgt die Behauptung. □

Korollar 12.18 Seien $A \in \mathbb{C}^{m,m}$, $B \in \mathbb{C}^{n,n}$ und $\lambda_1, \dots, \lambda_m$ die Eigenwerte von A und μ_1, \dots, μ_n die Eigenwerte von B .

- (i) Die Eigenwerte von $A \otimes B$ sind die $m \cdot n$ Produkte $\lambda_r \cdot \mu_s$, $r = 1, \dots, m$, $s = 1, \dots, n$.
- (ii) Die Eigenwerte von $A \otimes I_n + I_m \otimes B$ sind die $m \cdot n$ Summen $\lambda_r + \mu_s$, $r = 1, \dots, m$, $s = 1, \dots, n$.
- (iii) Sei $C = (A \otimes I_n) + (I_m \otimes B)$, so gilt $e^C = e^A \otimes e^B$.

Beweis: (i) und (ii) folgen direkt aus Satz 12.17.

Für (iii) folgt aus Korollar 12.6, dass $A \otimes I_n$ und $I_m \otimes B$ vertauschbar sind, also ist

$$\begin{aligned} e^C &= e^{A \otimes I_n + I_m \otimes B} = \sum_{j=0}^{\infty} \frac{1}{j!} [(A \otimes I_n) + (I_m \otimes B)]^j \\ &= e^{A \otimes I_n} \cdot e^{I_m \otimes B} \\ &= \left(\sum_{j=0}^{\infty} \frac{1}{j!} (A \otimes I_n)^j \right) \left(\sum_{i=0}^{\infty} \frac{1}{i!} (I_m \otimes B)^i \right) \\ &= \sum_{j=0}^{\infty} \frac{1}{j!} \sum_{i=0}^{\infty} \frac{1}{i!} (A \otimes I_n)^j \cdot (I_m \otimes B)^i \end{aligned}$$

$$\begin{aligned} &= \sum_{j=0}^{\infty} \frac{1}{j!} \sum_{i=0}^{\infty} \frac{1}{i!} (A^j \otimes B^i) \\ &= e^A \otimes e^B. \end{aligned}$$

□

Wir werden diese Ergebnisse nun anwenden, um die Stabilität von dynamischen Systemen zu überprüfen. Dies ist die berühmte Stabilitätstheorie von Lyapunov.

Kapitel 13

Lineare Matrixgleichungen

Wir werden nun das Kroneckerprodukt verwenden, um lineare Matrixgleichungen zu behandeln. Betrachte die allgemeine lineare Matrixgleichung

$$A_1XB_1 + A_2XB_2 + \dots + A_pXB_p = C \quad (13.1)$$

mit $A_j \in \mathbb{C}^{m,m}$, $B_j \in \mathbb{C}^{n,n}$, $j = 1, \dots, p$, $X, C \in \mathbb{C}^{m,n}$.

Satz 13.2 *Eine Matrix $X \in \mathbb{C}^{m,n}$ löst (13.1) genau dann, wenn der Vektor $x = \text{vec}(X)$ eine Lösung des linearen Gleichungssystems*

$$Gx = c \quad (13.3)$$

mit $c = \text{vec}(C)$ und

$$G = \sum_{j=1}^p B_j^\top \otimes A_j. \quad (13.4)$$

Beweis: Der Beweis folgt direkt aus Satz 12.13. □

Korollar 13.5

- a) Gleichung (13.1) hat eine Lösung genau dann, wenn $\text{Rang } [G, c] = \text{Rang } G$.
- b) Gleichung (13.1) hat eine eindeutige Lösung, falls G nichtsingulär.

Beweis: Siehe Teil I, Lemma 4.12. □

Wir betrachten nun den Spezialfall der Sylvester-Gleichung

$$AX + XB = C, \quad (13.6)$$

mit $A \in \mathbb{C}^{m,m}$, $B \in \mathbb{C}^{n,n}$, $X, C \in \mathbb{C}^{m,n}$.

Satz 13.7 Die Sylvester-Gleichung (13.6) mit $A \in \mathbb{C}^{m,m}$, $B \in \mathbb{C}^{n,n}$, $C \in \mathbb{C}^{m,n}$ hat eine eindeutige Lösung genau dann, wenn für die Spektren gilt $\sigma(A) \cap \sigma(-B) = \emptyset$, d.h., A und $-B$ haben keine gemeinsamen Eigenwerte.

Beweis: Wir können (13.6) schreiben als

$$(I_n \otimes A + B^\top \otimes I_m) \text{vec}(X) = \text{vec}(C).$$

Die Eigenwerte von $G = I_n \otimes A + B^\top \otimes I_m$ sind nach dem Satz von Stephanos (12.17) $\lambda_r + \mu_s$, wobei $\sigma(A) = \{\lambda_1, \dots, \lambda_m\}$, $\sigma(B) = \{\mu_1, \dots, \mu_n\}$. Also ist G nichtsingulär genau dann, wenn $\lambda_r + \mu_s \neq 0 \forall r = 1, \dots, m, s = 1, \dots, n$. \square

Für die Sylvester-Gleichung (13.6) können wir eine Lösungsformel angeben.

Satz 13.8 Falls alle Eigenwerte von $A \in \mathbb{C}^{m,m}$ und $B \in \mathbb{C}^{n,n}$ negativen Realteil haben, so ist die eindeutige Lösung von (13.6) gegeben durch

$$X = - \int_0^{\infty} (e^{At} C e^{Bt}) dt. \quad (13.9)$$

(Integral wird komponentenweise berechnet.)

Beweis: Aus der Eigenwerteigenschaft folgt nach Satz 13.7, dass es eine eindeutige Lösung von 13.6 gibt.

Sei nun $Z(t) : (0, \infty) \rightarrow \mathbb{C}^{m,n}$ definiert durch die Differentialgleichung

$$\frac{dZ}{dt} = AZ + ZB, \quad Z(0) = C. \quad (13.10)$$

Integration dieser Gleichung von $t = 0$ bis $t = \infty$ ergibt

$$\lim_{t \rightarrow \infty} Z(t) - Z(0) = A \int_0^{\infty} Z(t) dt + \left(\int_0^{\infty} Z(t) dt \right) B.$$

Mit der Lösung

$$Z(t) = e^{At} C e^{Bt} \quad (13.11)$$

von (13.10) erhalten wir

$$\begin{aligned} \lim_{t \rightarrow \infty} Z(t) &= \lim_{t \rightarrow \infty} e^{At} C e^{Bt} \\ &\stackrel{\text{Vorausss.}}{=} \lim_{t \rightarrow \infty} \underbrace{P e^{J_A t} P^{-1}}_{\rightarrow 0} \underbrace{C Q}_{\text{konst.}} \underbrace{e^{J_B t} Q^{-1}}_{\rightarrow 0} = 0 \end{aligned}$$

und damit folgt dann, dass $X = - \int_0^{\infty} Z(t) dt = - \int_0^{\infty} e^{At} C e^{Bt} dt$ die eindeutige Lösung von (13.6) ist. \square

Korollar 13.12 $A \in \mathbb{C}^{n,n}$ habe nur Eigenwerte mit negativem Realteil, und $W \in \mathbb{C}^{n,n}$ sei Hermite'sch positiv definit. Dann hat die Lyapunov-Gleichung

$$AX + XA^H = -W \quad (13.13)$$

eine eindeutige Lösung X , die auch Hermite'sch positiv definit ist.

Beweis: Da die Realteile der Eigenwerte von A^H gleich den Realteilen der Eigenwerte von A sind, können wir Satz 13.8 anwenden und erhalten die eindeutige Lösbarkeit mit der Lösung

$$X = - \int_0^{\infty} e^{At} (-W) e^{A^H t} dt. \quad (13.14)$$

Für alle $x \in \mathbb{C}^n \setminus \{0\}$ gilt

$$x^H X x = x^H \left(\int_0^{\infty} e^{At} W e^{A^H t} dt \right) x = \int_0^{\infty} \underbrace{(x^H e^{At}) W (e^{A^H t} x)}_{>0} dt > 0.$$

□

Satz 13.15

(a) Die diskrete Sylvester-Gleichung

$$X = AXB + C \quad (13.16)$$

mit $A \in \mathbb{C}^{m,m}$, $B \in \mathbb{C}^{n,n}$ und $C \in \mathbb{C}^{m,n}$ hat eine eindeutige Lösung X genau dann, wenn $\lambda_r \mu_s \neq 1$ für alle $\lambda_r \in \sigma(A)$ und $\mu_s \in \sigma(B)$.

(b) Falls für die Spektralradien von A und B $\rho(A) \cdot \rho(B) < 1$ gilt, dann ist die eindeutige Lösung von (13.16) gegeben durch

$$X = \sum_{j=0}^{\infty} A^j C B^j. \quad (13.17)$$

Beweis:

(a) Wir können nach Satz 12.13 die Gleichung (13.16) schreiben als

$$\text{vec}(X) = (B^T \otimes A) \text{vec}(X) + \text{vec}(C) \quad (13.18)$$

oder

$$\underbrace{(I_{m \cdot n} - B^T \otimes A)}_G \text{vec}(X) = \text{vec}(C). \quad (13.19)$$

Nach dem Satz von Stephanos (12.17) sind die Eigenwerte von G von der Form $1 - \lambda_r \cdot \mu_s$. Also ist G invertierbar genau dann, wenn $\lambda_r \mu_s \neq 1$, $\forall r = 1, \dots, m$, $s = 1, \dots, n$.

(b) Setze nun $\{Z_k\}_{k=0}^\infty$ an als Lösung von $Z_{k+1} = AZ_kB + C$ mit $Z_0 = C$. So folgt

$$\begin{aligned} Z_{k+1} &= AZ_kB + C \\ &= A^2Z_{k-1}B^2 + ACB + C \\ &= \sum_{j=0}^{k+1} A^jCB^j. \end{aligned}$$

Da die eindeutige Lösung existiert und

$$\lim_{k \rightarrow \infty} Z_k = \lim_{k \rightarrow \infty} \sum_{j=0}^k A^jCB^j = \lim_{k \rightarrow \infty} P \left(\sum_{j=0}^k J_A^j \underbrace{(P^{-1}CQ)}_{\text{konst.}} J_B^j \right) Q^{-1}$$

konvergiert, folgt die Behauptung. □

Korollar 13.20 $A \in \mathbb{C}^{n,n}$ habe einen Spektralradius $\rho(A) < 1$ und $C = C^H \in \mathbb{C}^{n,n}$ sei positiv semidefinit. Dann hat die Stein-Gleichung

$$X = AXA^H + C \tag{13.21}$$

eine eindeutige Lösung X und es gilt, dass X und $X - C$ positiv semidefinit sind.

Beweis: $\rho(A^H) = \rho(A) \implies$ Wir können Satz 13.15 anwenden und erhalten die eindeutige Lösung

$$X = \sum_{j=0}^{\infty} A^j C (A^H)^j.$$

Damit folgt

$$\begin{aligned} x^H X x &= \sum_{j=0}^{\infty} \underbrace{(x^H A^j) C (A^H)^j x}_{\geq 0} \geq 0, \\ x^H (X - C) x &= \sum_{j=1}^{\infty} (x^H A^j) C (A^H)^j x \geq 0. \end{aligned}$$
□

Wir können die Lösung der Sylvester-Gleichung auch noch anders angeben. Dazu betrachten wir zuerst den Spezialfall der homogenen Gleichung mit $B = -A$

$$AX = XA, \quad A, X \in \mathbb{C}^{n,n}. \tag{13.22}$$

Die Lösungsmenge ist die Menge aller Matrizen, die mit A vertauschen. Sei $J_A = P^{-1}AP$ die Jordan'sche Normalform von A , so können wir mit $Y = P^{-1}XP$ auch die Gleichung

$$JY = YJ \tag{13.23}$$

betrachten.

Sei $J_A = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_p \end{bmatrix}$ mit $J_i = \lambda_i I_{n_i} + N_{n_i}$.

Setze $Y = \begin{bmatrix} Y_{11} & \cdots & Y_{1p} \\ \vdots & & \vdots \\ Y_{p1} & \cdots & Y_{pp} \end{bmatrix}$, partitioniert genau wie J_A , so erhalten wir das System

$$J_s Y_{st} = Y_{st} J_t, \quad s, t = 1, \dots, p,$$

oder äquivalent

$$(\lambda_s - \lambda_t) Y_{st} = Y_{st} N_{n_t} - N_{n_s} Y_{st}, \quad s, t = 1, \dots, p. \quad (13.24)$$

Fall 1: $\lambda_s \neq \lambda_t$.

Sukzessives Einsetzen und Multiplikation mit $(\lambda_s - \lambda_t)$ ergibt

$$(\lambda_s - \lambda_t)^r Y_{st} = \sum_{j=0}^r (-1)^j \binom{r}{j} N_{n_s}^j Y_{st} N_{n_t}^{r-j}, \quad (13.25)$$

wobei wir $N_{n_s}^0 = I_{n_s}$ und $N_{n_t}^0 = I_{n_t}$ verwendet haben.

Da $N_{n_j}^q = 0$ für genügend großes q , so folgt für genügend großes q , dass $(\lambda_s - \lambda_t)^q Y_{st} = 0$ und da $\lambda_s \neq \lambda_t$, so ist $Y_{st} = 0$.

Fall 2: $\lambda_s = \lambda_t$.

Dann hat (13.24) die Form $Y_{st} N_{n_t} = N_{n_s} Y_{st}$, $1 \leq s, t \leq p$.

Nachrechnen ergibt, dass für

$$\left. \begin{aligned} (a) \quad n_s = n_t &\implies Y_{st} = \begin{bmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_{n_s} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \gamma_2 \\ & & & \gamma_1 \end{bmatrix} =: T(\gamma_1, \dots, \gamma_{n_s}) \\ &\quad \text{(obere-Dreiecks-Toeplitz-Matrix)} \\ (b) \quad n_s < n_t &\implies Y_{st} = [0 \quad T(\gamma_1, \dots, \gamma_{n_s})] \\ (c) \quad n_s > n_t &\implies Y_{st} = \begin{bmatrix} T(\gamma_1, \dots, \gamma_{n_t}) \\ 0 \end{bmatrix}. \end{aligned} \right\} \quad (13.26)$$

Mit dieser Konstruktion haben wir den folgenden Satz bewiesen.

Satz 13.27 Sei $A \in \mathbb{C}^{n,n}$ und die Jordan'sche Normalform von A

$$J_A = P^{-1} A P = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_p \end{bmatrix}.$$

Eine Matrix X vertauscht mit A genau dann, wenn

$$X = PYP^{-1}, \quad \text{wobei } Y = \begin{bmatrix} Y_{11} & \cdots & Y_{1p} \\ \vdots & & \vdots \\ Y_{p1} & \cdots & Y_{pp} \end{bmatrix}$$

analog partitioniert ist wie J_A und $Y_{st} = 0$ für $\lambda_s \neq \lambda_t$ oder in einer der Formen (13.26) (a), (b), (c), falls $\lambda_s = \lambda_t$.

Eine Folgerung aus diesem Satz ist die Beantwortung der Frage, wann alle mit A vertauschbaren Matrizen Polynome in A sind.

Korollar 13.28 Sei $A \in \mathbb{C}^{n,n}$. Jede Matrix, die mit A vertauscht, ist ein Polynom in A genau dann, wenn es keine zwei Jordanblöcke zum gleichen Eigenwert gibt.

Beweis: Der Beweis ergibt sich durch sehr langwieriges Nachrechnen und durch Lösung einer Interpolationsaufgabe. Er soll hier nicht geführt werden. \square

Noch einige weitere Folgerungen.

Korollar 13.29

- (a) Diagonalisierbare Matrizen $A, X \in \mathbb{C}^{n,n}$ vertauschen genau dann, wenn sie eine gemeinsame Eigenvektorbasis besitzen.
- (b) Falls $A \in \mathbb{C}^{n,n}$ normal mit n verschiedenen Eigenwerten, so ist jede Matrix, die mit A vertauscht, auch normal.
- (c) Normale Matrizen $A, X \in \mathbb{C}^{n,n}$ vertauschen genau dann, wenn sie ein gemeinsames Orthonormalsystem aus Eigenvektoren besitzen.

Beweis:

- (a) Seien $\lambda_1, \dots, \lambda_s$ die paarweise verschiedenen Eigenwerte von A mit Vielfachheiten n_1, \dots, n_s und sei $P^{-1}AP = D$. Nach Satz 13.27 ist

$$X = PYP^{-1} = P \begin{bmatrix} Y_1 & & \\ & \ddots & \\ & & Y_s \end{bmatrix} P^{-1} \quad \text{mit } Y_i \in \mathbb{C}^{n_i, n_i}.$$

Jedes Y_i kann auf diagonale Jordanform \tilde{D}_i transformiert werden: $Y_i = Q_i \tilde{D}_i Q_i^{-1}$.

$$\text{Sei } Q = \begin{bmatrix} Q_1 & & \\ & \ddots & \\ & & Q_s \end{bmatrix}, \text{ so ist } X = PQ \begin{bmatrix} \tilde{D}_1 & & \\ & \ddots & \\ & & \tilde{D}_s \end{bmatrix} (PQ)^{-1} \text{ und}$$

$$A = (PQ)D(PQ)^{-1}.$$

Umgekehrt, falls A, X eine gemeinsame Eigenvektorbasis besitzen, dann sind A, X diagonalisierbar und $P^{-1}AP = D_1, P^{-1}XP = D_2$ mit D_1, D_2 diagonal und damit vertauschen A und X .

(b), (c) Matrizen sind normal, genau dann, wenn sie (unitär) diagonalisierbar sind. Folgt man dem Beweis von (a), ergeben sich die Behauptungen.

□

Wir kommen nun zur homogenen Sylvester-Gleichung

$$AX + XB = 0.$$

Man sieht sofort, dass $X \in \mathbb{C}^{m,n}$ genau dann eine Lösung ist, wenn die Matrizen

$$\begin{bmatrix} I_m & X \\ 0 & I_n \end{bmatrix}, \begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix} \quad (13.30)$$

vertauschen. Damit folgt

Satz 13.31 Sei $A \in \mathbb{C}^{m,m}$, $B \in \mathbb{C}^{n,n}$ und $A = PJ_AP^{-1}$, $B = QJ_BQ^{-1}$ mit

$$\left. \begin{aligned} J_A &= \text{diag}[\lambda_1 I_{m_1} + N_{m_1}, \dots, \lambda_p I_{m_p} + N_{m_p}] \\ J_B &= \text{diag}[\mu_1 I_{n_1} + N_{n_1}, \dots, \mu_q I_{n_q} + N_{n_q}] \end{aligned} \right\} \quad (13.32)$$

Jede Lösung X von

$$AX + XB = 0 \quad (13.33)$$

hat die Form $X = PYQ^{-1}$, wobei $Y = [Y_{st}]$, $s = 1, \dots, p$, $t = 1, \dots, q$, die allgemeine Lösung von $J_A Y + Y J_B = 0$ ist, mit $Y_{st} \in \mathbb{C}^{m_s, n_t}$, und $Y_{st} = 0$, falls $\lambda_s \neq -\mu_t$ bzw. von der Form (13.26), falls $\lambda_s = -\mu_t$.

Satz 13.34 Die Gleichung

$$AX + XB = C \quad \text{mit} \quad A \in \mathbb{C}^{m,m}, B \in \mathbb{C}^{n,n}, C \in \mathbb{C}^{m,n} \quad (13.35)$$

hat entweder keine Lösung, eine eindeutige Lösung (falls $\sigma(A) \cap \sigma(-B) = \emptyset$) oder unendlich viele Lösungen, gegeben durch

$$X = X_0 + \tilde{X},$$

wobei X_0 eine feste Partikulärlösung von (13.35) ist und \tilde{X} die allgemeine Lösung von (13.33).

Als letztes wollen wir noch eine Lösungstheorie in den Größen A, B, C aufstellen.

Satz 13.36 Gleichung (13.35) hat eine Lösung genau dann, wenn

$$\begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix} \text{ und } \begin{bmatrix} A & C \\ 0 & -B \end{bmatrix}$$

ähnlich sind.

Beweis: Falls (13.35) lösbar ist, so sind $\begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix}$ und $\begin{bmatrix} A & C \\ 0 & -B \end{bmatrix}$ ähnlich mit der Transformation $\begin{bmatrix} I & X \\ 0 & I \end{bmatrix}$ von links und $\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}$ von rechts.

Sei

$$\begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix} = P^{-1} \begin{bmatrix} A & C \\ 0 & -B \end{bmatrix} P \text{ mit } P \in \mathbb{C}^{m+n, m+n}. \quad (13.37)$$

Seien T_1, T_2 lineare Abbildungen, gegeben durch

$$T_1(Z) = \begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix} Z - Z \begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix},$$

$$T_2(Z) = \begin{bmatrix} A & C \\ 0 & -B \end{bmatrix} Z - Z \begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix}.$$

Wenn wir eine Matrix $Z = \begin{bmatrix} R & U \\ 0 & -I \end{bmatrix}$ in Kern (T_2) finden können, so ist U Lösung von (13.35), denn dann gilt

$$\begin{aligned} 0 &= \begin{bmatrix} A & C \\ 0 & -B \end{bmatrix} \begin{bmatrix} R & U \\ 0 & -I \end{bmatrix} - \begin{bmatrix} R & U \\ 0 & -I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix} \\ &= \begin{bmatrix} AR - RA & AU - C + UB \\ 0 & B - B \end{bmatrix}. \end{aligned}$$

Aus (13.37) folgt $T_1(Z) = P^{-1}T_2(PZ)$,

$$\begin{aligned} \implies \text{Kern}(T_2) &= \{PZ : Z \in \mathbb{C}^{m+n, m+n}, Z \in \text{Kern}(T_1)\}, \\ \implies \dim(\text{Kern}(T_1)) &= \dim(\text{Kern}(T_2)). \end{aligned}$$

Betrachte die Menge L der Paare (V, W) , $V \in \mathbb{C}^{n, m}, W \in \mathbb{C}^{n, n}$, so dass

$$BV + VA = 0, \quad BW = WB.$$

Dann ist L ein Vektorraum mit den Operationen

$$\begin{aligned} \alpha(V, W) &= (\alpha V, \alpha W), \\ (V_1, W_1) + (V_2, W_2) &= (V_1 + V_2, W_1 + W_2). \end{aligned}$$

Sei $\varphi_i : \text{Kern}(T_i) \rightarrow L$ linear mit

$$\varphi_i \left(\begin{bmatrix} R & U \\ V & W \end{bmatrix} \right) = (V, W),$$

so ist

$$\text{Kern}(\varphi_1) = \text{Kern}(\varphi_2) = \left\{ \begin{bmatrix} R & U \\ 0 & 0 \end{bmatrix} : AR = RA, AU + UB = 0 \right\} \quad (13.38)$$

und $\text{Bild}(\varphi_1) = L$, denn falls $(V, W) \in L$, so $BV + VA = 0$ und $BW = WB$. Also

$$\begin{aligned} \begin{bmatrix} 0 & 0 \\ V & W \end{bmatrix} &\in \text{Kern}(T_1) \text{ und} \\ \varphi_1 \left(\begin{bmatrix} 0 & 0 \\ V & W \end{bmatrix} \right) &= (V, W), \quad \text{also } L \subset \text{Bild}(\varphi_1). \end{aligned}$$

Die Umkehrung $\text{Bild}(\varphi_1) \subset L$ ist klar.

Nun gilt aber $\text{Bild}(\varphi_2) \subset L = \text{Bild}(\varphi_1)$ und

$$\dim(\text{Kern}(\varphi_i)) + \dim(\text{Bild}(\varphi_i)) = \dim(\text{Kern}(T_i)), \quad i = 1, 2.$$

Aus $\dim(\text{Kern}(T_1)) = \dim(\text{Kern}(T_2))$ und (13.38) folgt $\dim(\text{Bild}(\varphi_1)) = \dim(\text{Bild}(\varphi_2))$. Also

$$\text{Bild}(\varphi_1) = \text{Bild}(\varphi_2) = L. \quad (13.39)$$

Aber da $\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \in \text{Kern}(T_1)$ und $\varphi_1 \left(\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \right) = (0, -I)$, so folgt aus (13.39), dass es

$\begin{bmatrix} R_0 & U_0 \\ V_0 & W_0 \end{bmatrix} \in \text{Kern}(T_2)$ gibt, so dass

$$(0, -I) = \varphi_2 \left(\begin{bmatrix} R_0 & U_0 \\ V_0 & W_0 \end{bmatrix} \right).$$

Also ist $V_0 = 0$, $W_0 = -I$ und wir haben die gewünschte Matrix gefunden. \square

Im nächsten Kapitel werden wir diese Ergebnisse zur Stabilitätsanalyse verwenden.

Kapitel 14

Die Stabilitätstheorie von Lyapunov

Eine wichtige Theorie, die bei der Analyse und Lösung von Problemen der Mechanik, Steuerungstheorie und auch der Dynamik von wirtschaftswissenschaftlichen Modellen eine Rolle spielt, ist die Frage der Stabilität.

Definition 14.1

(a) Sei $y(t)$ die Lösung der linearen Differentialgleichung

$$y' = A_c y, \quad y(t_0) = y^0. \quad (14.2)$$

Die Lösung heißt asymptotisch stabil, falls $\lim_{t \rightarrow \infty} y(t) = 0$ für alle y^0 .

(b) Sei $\{y_k\}_{k=0}^{\infty}$ die Lösungsfolge der linearen Differenzgleichung

$$y_{k+1} = A_d y_k, \quad y_0 = y^0. \quad (14.3)$$

Die Lösung heißt asymptotisch stabil, falls für alle y^0

$$\lim_{k \rightarrow \infty} y_k = 0.$$

Satz 14.4

(a) Die Lösung von (14.2) ist asymptotisch stabil genau dann, wenn $\operatorname{Re}(\lambda) < 0$ für alle Eigenwerte von A_c .

(b) Die Lösung von (14.3) ist asymptotisch stabil genau dann, wenn $|\lambda| < 1$ für alle Eigenwerte λ von A_d .

Beweis:

(a) Die Lösung ist $y(t) = e^{A_c(t-t_0)}y^0$. Damit ist klar, dass die Bedingung an die Eigenwerte notwendig ist, denn ist $J_{A_c} = P A_c P^{-1}$ die Jordan'sche Normalform von A_c , so gilt

$$y(t) = P^{-1} e^{J_{A_c}(t-t_0)} P y^0.$$

Sei $z(t) = Py(t)$, $z^0 = Py^0$.

$$\lim_{t \rightarrow \infty} y(t) = 0 \iff \lim_{t \rightarrow \infty} z(t) = 0$$

$$z(t) = e^{J_{A_c}(t-t_0)} z^0.$$

Falls es $\tilde{\lambda}$ mit $\operatorname{Re}(\tilde{\lambda}) \geq 0$ gibt, so ordne die Eigenwerte von J_{A_c} so, dass $\tilde{\lambda}$ das letzte Diagonalelement ist, dann folgt

$$z_n(t) = e^{\tilde{\lambda}(t-t_0)} z_n^0$$

und $\lim_{t \rightarrow \infty} z_n(t) \neq 0$. Das ist ein Widerspruch.

Für die Umkehrung sind die Formeln für $e^{J_{A_c}} = \sum_{i=0}^{\infty} \frac{1}{i!} J_{A_c}^i$ anzuwenden.

- (b) Sei $J_{A_d} = PA_dP^{-1}$ die Jordan'sche Normalform von A_d . Die Lösung von (14.3) ist $y_k = A_d^k y^0$. Sei $z_k = Py_k$, $z^0 = Py_0$.

$$\lim_{k \rightarrow \infty} y_k = 0 \iff \lim_{k \rightarrow \infty} z_k = 0 \iff \lim_{k \rightarrow \infty} J_{A_d}^k z^0 = 0.$$

Damit dies für alle z^0 gilt, muss notwendigerweise gelten $|\lambda| < 1$ für alle Eigenwerte von A_d . Für die Umkehrung verwende die Formeln für $J_{A_d}^k$.

□

Definition 14.5 Sei $A \in \mathbb{C}^{n,n}$. Dann heißt A stabil (bzgl. der imaginären Achse), falls $\operatorname{Re}(\lambda) < 0$ für alle Eigenwerte λ von A , und es heißt A stabil (bzgl. des Einheitskreises), falls $|\lambda| < 1$ für alle Eigenwerte λ von A .

Die Grundlagen der Stabilitätstheorie von Lyapunov sind nun die folgenden.

Sei $y(t)$ Lösung von (14.2) und sei V eine positiv definite Matrix. Setze $v(y) = y^H V y$, so gilt mit $\dot{y} = Ay$

$$\dot{v} = \frac{dv}{dt} = \dot{y}^H V y + y^H V \dot{y} = y^H (A^H V + V A) y.$$

Wenn wir setzen

$$A^H V + V A = -W, \tag{14.6}$$

so ist W Hermite'sch und

$$\dot{v}(y) = -w(y) \quad \text{mit} \quad w(y) := y^H(t) W y(t).$$

Lyapunov bemerkte, dass für eine gegebene positiv definite Hermite'sche Form w , die Stabilität (bzgl. der imaginären Achse) durch die Existenz einer positiv definiten Lösung V von (14.6) charakterisiert werden kann, weil dann durch

$$\|y\|_V^2 = y^H V y \tag{14.7}$$

eine Norm in \mathbb{C}^n gegeben ist, und

$$\lim_{t \rightarrow \infty} y(t) = 0 \iff \lim_{t \rightarrow \infty} v(y(t)) = 0.$$

Denn wir erhalten, dass $\lim_{t \rightarrow \infty} y(t) = 0 \iff \exists$ positiv definite Hermite'sche Formen w, v , so dass

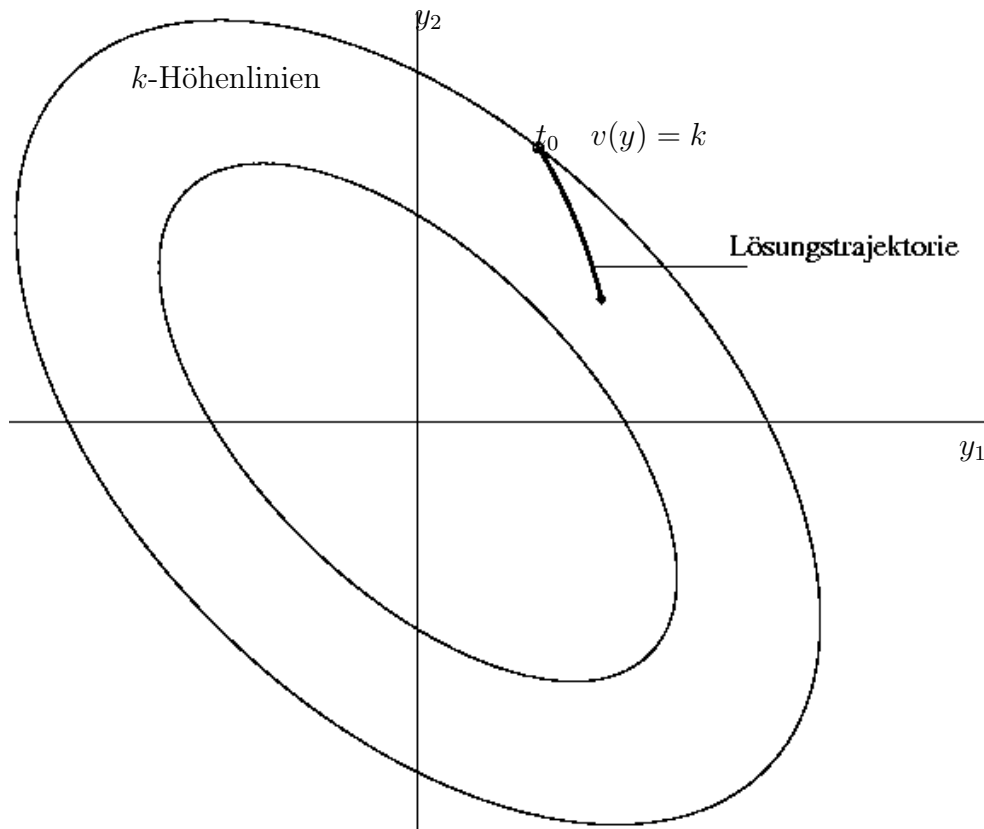
$$\dot{v}(y) = -w(y).$$

Die positiv definite Matrix W kann dabei beliebig gewählt werden.

Beispiel 14.8 Sei $A \in \mathbb{R}^{2,2}$ und $y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$ Lösung von

$$\dot{y}(t) = Ay(t).$$

Die Höhenlinien für eine positiv definite symmetrische Bilinearform $V(y)$ sind gerade Ellipsen in der durch y_1, y_2 aufgespannten Ebene.



Falls $y(t_0) = y^0$, so erhalten wir einen Punkt auf einer Ellipse (Höhenlinie $v(y) = k$). Für eine Lösungstrajektorie in der Umgebung von $\begin{bmatrix} y_1(t_0) \\ y_2(t_0) \end{bmatrix}$ gilt, falls $w > 0$ und $\dot{v} < 0$, so geht die Lösung in das Innere der Ellipse für $t > t_0$.

Satz 14.9 (Satz von Lyapunov) [Ляпунов]

Seien $A, W \in \mathbb{C}^{n,n}$, $W = W^H$ positiv definit.

(a) Falls A stabil ist (bzgl. der imaginären Achse), so hat die Gleichung

$$AH + HA^H = W \quad (14.10)$$

eine eindeutige Lösung H und H ist negativ definit.

(b) Falls es eine negativ definite Matrix H gibt, so dass (14.10) erfüllt ist, so ist A stabil (bzgl. der imaginären Achse).

Beweis:

(a) Sei A stabil, so haben A und $-A^H$ keine gemeinsamen Eigenwerte und wir haben nach Satz 13.8 die explizite Lösung

$$H = - \int_0^{\infty} e^{At} W e^{A^H t} dt.$$

Die Matrix e^{At} ist nichtsingulär für alle $t \in \mathbb{R}$ und W ist positiv definit,

$$\implies \forall x \in \mathbb{C}^n \setminus \{0\} : \quad x^H H x = - \int_0^{\infty} (x^H e^{At}) W (e^{A^H t} x) dt < 0.$$

(b) Sei λ ein Eigenwert von A^H mit Eigenvektor x , d.h.

$$A^H x = \lambda x.$$

Multipliziere (14.10) von links mit x^H und von rechts mit x , so ergibt sich

$$x^H A H x + x^H H A^H x = \bar{\lambda} x^H H x + \lambda x^H H x = x^H W x$$

$$\iff (\bar{\lambda} + \lambda) x^H H x = x^H W x.$$

Nach Voraussetzung ist $x^H H x < 0$ und $x^H W x > 0$,

$$\implies \bar{\lambda} + \lambda < 0 \implies \operatorname{Re}(\lambda) < 0.$$

□

Man kann diesen Satz noch wesentlich verschärfen, wenn man den Begriff des Trägheitsindex auf beliebige Matrizen erweitert.

Definition 14.11 Sei $A \in \mathbb{C}^{n,n}$. Der Trägheitsindex $(\pi(A), \nu(A), \omega(A))$ von A gibt die Anzahl der Eigenwerte mit $\operatorname{Re}(\lambda) > 0$, $\operatorname{Re}(\lambda) < 0$, $\operatorname{Re}(\lambda) = 0$ an.

Satz 14.12 (Ostrowski/Schneider)

Sei $A \in \mathbb{C}^{n,n}$. Falls $H = H^H \in \mathbb{C}^{n,n}$, so dass

$$AH + HA^H = W \quad (14.13)$$

mit W positiv definit, so ist H nichtsingulär und

$$\pi(A) = \pi(H), \quad \nu(A) = \nu(H), \quad \omega(A) = \omega(H). \quad (14.14)$$

Umgekehrt, falls $\omega(A) = 0$, so gibt es eine Hermite'sche Matrix H , so dass (14.13), (14.14) gelten.

Beweis: Angenommen, (14.13) ist erfüllt.

Sei λ Eigenwert von A^H und x zugehöriger Eigenvektor, so erhalten wir wie im Beweis von Satz 14.9

$$(\bar{\lambda} + \lambda)x^H H x = x^H W x.$$

Da W positiv definit ist, folgt $\bar{\lambda} + \lambda \neq 0$, also ist $\omega(A) = 0$.

Sei

$$P^{-1}AP = J = \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} \begin{matrix} p \\ n-p \\ p & n-p \end{matrix} \quad (14.15)$$

die Jordan'sche Normalform von A wobei J_1 nur Eigenwerte mit positivem Realteil hat und J_2 nur Eigenwerte mit negativem Realteil. Sei $R = P^{-1}HP^{-H}$, so können wir (14.13) schreiben als

$$JR + RJ^H = P^{-1}WP^{-H} =: W_0.$$

Die Aufteilung von W_0, R analog zu J als $R = \begin{bmatrix} R_{11} & R_{12} \\ R_{12}^H & R_{22} \end{bmatrix}$, $W_0 = \begin{bmatrix} W_{11} & W_{12} \\ W_{12}^H & W_{22} \end{bmatrix}$ ergibt

$$J_1 R_{11} + R_{11} J_1^H = W_{11}, \quad J_2 R_{22} + R_{22} J_2^H = W_{22}$$

mit W_{11}, W_{22} positiv definit.

Nach Satz 14.9 sind dann R_{11} und $-R_{22}$ positiv definit.

Sei nun

$$Q = \begin{bmatrix} I & -R_{11}^{-1}R_{12} \\ 0 & I \end{bmatrix},$$

so folgt

$$Q^H R Q = \begin{bmatrix} R_{11} & 0 \\ 0 & R_{22} - R_{12}^H R_{11}^{-1} R_{12} \end{bmatrix}$$

mit R_{11} positiv definit und $R_{22} - R_{12}^H R_{11}^{-1} R_{12}$ negativ definit, denn

$$x^H (R_{22} - R_{12}^H R_{11}^{-1} R_{12}) x = x^H R_{22} x - (x^H R_{12}^H) R_{11}^{-1} (R_{12} x) < 0 \quad \forall x \in \mathbb{C}^n \setminus \{0\},$$

also ist

$$\begin{aligned} \pi(H) &= \pi(R) = \pi(Q^H R Q) = \pi(A) \\ \nu(H) &= \nu(R) = \nu(Q^H R Q) = \nu(A). \end{aligned}$$

Für die Umkehrung nehmen wir an, dass $\omega(A) = 0$, dann hat A die Jordanform (14.15) und es gibt nach Satz 14.9

$$H_1 \in \mathbb{C}^{p,p}, \quad H_2 \in \mathbb{C}^{n-p,n-p},$$

so dass

$$\begin{aligned} -J_1 H_1 - H_1 J_1^H &= W_{11} \quad \text{positiv definit} \\ J_2 H_2 + H_2 J_2^H &= W_{22} \quad \text{positiv definit.} \end{aligned}$$

Dann ist $H = P \begin{bmatrix} -H_1 & 0 \\ 0 & H_2 \end{bmatrix} P^H$ Hermite'sch, erfüllt (14.13) und hat den gleichen Trägheitsindex wie \tilde{A} . \square

Wir erhalten auch analoge Sätze für Differenzgleichungen $x_{k+1} = Ax_k$, und Stabilität (bzgl. des Einheitskreises).

Satz 14.16 *Seien $A, V \in \mathbb{C}^{n,n}$ und V positiv definit.*

(a) *Falls A stabil (bzgl. des Einheitskreises), so hat die Stein-Gleichung*

$$H - A^H H A = V \tag{14.17}$$

eine eindeutige Lösung H und H ist positiv definit.

(b) *Falls es eine positiv definite Matrix H gibt, welche (14.17) erfüllt, so ist A stabil (bzgl. des Einheitskreises).*

Beweis:

(a) Da A stabil (bzgl. des Einheitskreises) ist, so ist $|\lambda| < 1$ für alle Eigenwerte $\implies A^H + I$ ist nichtsingulär und die Cayley-Transformation von A

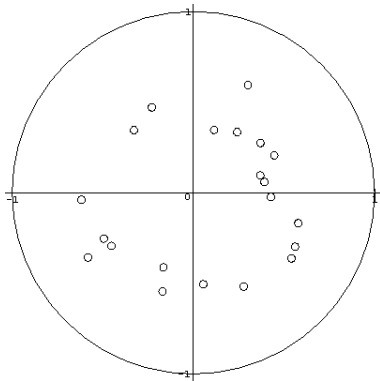
$$C = (A^H + I)^{-1} (A^H - I) \tag{14.18}$$

ist definiert. C hat die gleichen Eigen- und Hauptvektoren wie A^H , denn falls $A^H x = \lambda x$, so gilt

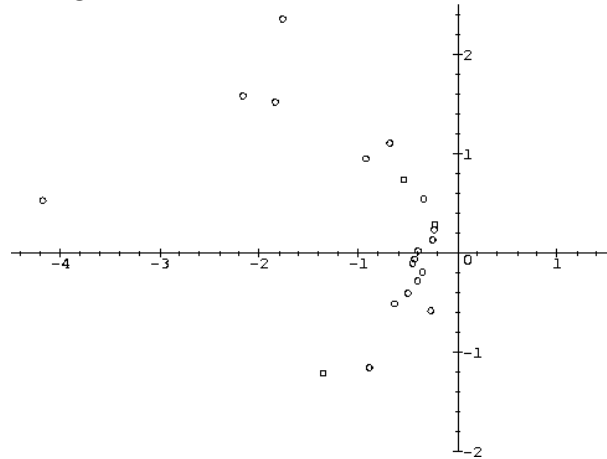
$$C x = (A^H + I)^{-1} (\lambda - 1) x = \frac{\lambda - 1}{\lambda + 1} x,$$

denn aus $A^H x = \lambda x$ folgt $(A^H + I)x = (\lambda + 1)x \implies \frac{1}{\lambda + 1}x = (A^H + I)^{-1}x$.

Eigenwerte von A



Eigenwerte von C



Die Cayley-Transformation entspricht einer Aufschneidung des Einheitskreises bei -1 und Abbildung des Inneren in die linke Halbebene, des Kreises auf die imaginäre Achse und des Äußeren in die rechte Halbebene.

$$\begin{aligned} CH + HC^H &= (A^H + I)^{-1} \left[(A^H - I)H(A + I) + (A^H + I)H(A - I) \right] \cdot (A + I)^{-1} \\ &= 2(A^H + I)^{-1}(A^H H A - H)(A + I)^{-1} \end{aligned}$$

$$\implies C(-H) + (-H)C^H = 2(A^H + I)^{-1}(H - A^H H A)(A + I)^{-1}.$$

Aus dem Satz von Lyapunov folgt für $H - A^H H A$ positiv definit, dass H positiv definit und eindeutig ist.

- (b) Für die Umkehrung verwenden wir dasselbe Argument wie im Satz für Lyapunov-Gleichungen. Angenommen, (14.17) gilt mit H positiv definit. Falls $Ax = \lambda x$, so gilt $x^H A^H = \bar{\lambda} x^H$ und damit

$$x^H V x = (1 - |\lambda|^2)x^H H x.$$

Da V und H positiv definit sind, so folgt

$$1 - |\lambda|^2 > 0.$$

□

Schließlich können wir auch noch einen analogen Satz zu 6.5 über den Trägheitsindex beweisen.

Satz 14.19 Sei $A \in \mathbb{C}^{n,n}$. Falls H Hermite'sche Lösung der Stein-Gleichung (14.17) ist mit V positiv definit, so ist H invertierbar und A hat $\pi(H)$ Eigenwerte im Einheitskreis, $\nu(H)$ Eigenwerte außerhalb des Einheitskreises und keinen Eigenwert vom Betrag 1.

Umgekehrt, falls A keine Eigenwerte vom Betrag 1 hat, so existiert eine Hermite'sche Matrix H , so dass (14.17) erfüllt ist und der Trägheitsindex von H $(\pi(H), \nu(H), 0)$ gibt die Anzahl von Eigenwerten von A innerhalb bzw. außerhalb und auf dem Einheitskreis an.

Beweis: Die Matrix A habe die Jordan'sche Normalform

$$J = \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} = P^{-1}AP, \quad (14.20)$$

mit $J_1 \in \mathbb{C}^{p,p}$, $J_2 \in \mathbb{C}^{n-p,n-p}$, und die Eigenwerte von J_1 und J_2 liegen innerhalb bzw. außerhalb des Einheitskreises.

Mit $R := P^H H P = \begin{bmatrix} R_{11} & R_{12} \\ R_{12}^H & R_{22} \end{bmatrix}$ und $V_0 = P^H V P = \begin{bmatrix} V_{11} & V_{12} \\ V_{12}^H & V_{22} \end{bmatrix}$ positiv definit ergibt sich

$$R - J^H R J = V_0$$

beziehungsweise

$$R_{11} - J_1^H R_{11} J_1 = V_{11}, \quad R_{22} - J_2^H R_{22} J_2 = V_{22}$$

mit positiv definiten V_{11}, V_{22} .

Aus Satz 14.16 folgt $R_{11}, -R_{22}$ positiv definit und damit

$$\pi(H) = \pi(R) = p, \quad \nu(H) = \nu(R) = n - p.$$

Für die Umkehrung sei $|\lambda| \neq 1$ für alle Eigenwerte von A , so hat A die Jordan'sche Normalform (14.20). Nach Satz 14.16 gibt es H_1, H_2 positiv definit, so dass

$$\begin{aligned} H_1 - J_1^H H_1 J_1 &= V_1, \\ H_2 - J_2^{-H} H_2 J_2^{-1} &= V_2, \end{aligned}$$

wobei V_1, V_2 positiv definit sind. Also gilt für

$$H = P^{-H} \begin{bmatrix} H_1 & 0 \\ 0 & -J_2^{-H} H_2 J_2^{-1} \end{bmatrix} P^{-1},$$

dass $H - A^H H A > 0$ und die Eigenwertbedingung folgt analog wie vorher. \square

Kapitel 15

Matrixgruppen

Wir hatten schon in Teil 1 gesehen, dass die invertierbaren Matrizen in $\mathbb{C}^{n,n}(\mathbb{R}^{n,n})$ eine Gruppe bilden und analog auch die unitären (orthogonalen) Matrizen $\mathcal{O}_n(\mathbb{C})$ bzw. $\mathcal{O}_n(\mathbb{R})$. Wir werden nun weitere solcher Matrixgruppen kennenlernen.

Definition 15.1 Sei $J = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix} \in \mathbb{R}^{n,n}$ und sei $\langle x, y \rangle_J = y^\top Jx$ eine Abbildung von $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, so heißt eine Matrix $Q \in \mathbb{R}^{n,n}$ J -orthogonal¹, falls

$$\langle Qx, Qy \rangle_J = \langle x, y \rangle_J, \quad (15.2)$$

und $S \in \mathbb{R}^{n,n}$ heißt J -symmetrisch, falls

$$\langle Sx, y \rangle_J = \langle x, Sy \rangle_J. \quad (15.3)$$

Die Abbildung $\langle \cdot, \cdot \rangle_J : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ definiert ein indefinites Skalarprodukt, bei dem die gleichen Bedingungen wie bei einem normalen Skalarprodukt gelten, bis auf die letzte Bedingung, denn

$$\langle x, x \rangle_J = 0 \not\Rightarrow x = 0.$$

Beispiel 15.4 $J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \neq 0$.

$$x^\top Jx = [1 \quad -1] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0.$$

Satz 15.5 Sei $\mathcal{S}_J = \{Q \mid Q \text{ ist } J\text{-orthogonal}\}$, so gilt für alle $Q \in \mathcal{S}_J$, dass

$$Q^\top JQ = J$$

und \mathcal{S}_J ist eine Gruppe bezüglich der normalen Matrixmultiplikation, und damit Untergruppe von $GL_n(\mathbb{R})$.

¹manchmal werden auch solche Matrizen symplektisch genannt (hier aber siehe Def. 15.27)

Beweis: Der Beweis der Gruppeneigenschaft ist Übungsaufgabe.

Sei $\{x_1, \dots, x_n\}$ eine Basis von \mathbb{R}^n , so gilt

$$\begin{aligned} \langle Qx_i, Qx_j \rangle &= x_j^\top Q^\top J Q x_i = x_j^\top J x_i \quad \forall i, j = 1, \dots, n \\ \implies \text{mit } X &= [x_1, \dots, x_n] \\ X^\top Q^\top J Q X &= X^\top J X \end{aligned}$$

und da die x_i eine Basis bilden, ist X nichtsingulär $\implies Q^\top J Q = J$. □

Wie sehen nun die Elemente von \mathcal{S}_J aus, und was für Eigenschaften hat \mathcal{S}_J ?

Falls $J = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix}$ und $p = 0$ oder $q = 0$ ist, so ist $J = -I$ oder $J = I$ und damit $\mathcal{S}_J = \mathcal{O}_n(\mathbb{R})$.

Falls jedoch $p, q \neq 0$, so ist die Situation komplizierter.

Lemma 15.6 Sei $J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, so haben alle J -orthogonalen Matrizen die Form

$$\begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix} \begin{bmatrix} \cosh(\varphi) & \sinh(\varphi) \\ \sinh(\varphi) & \cosh(\varphi) \end{bmatrix}, \quad \text{mit } \varepsilon_i^2 = 1, \quad i = 1, 2$$

wobei $\cosh(\varphi) = \frac{e^\varphi + e^{-\varphi}}{2}$, $\sinh(\varphi) = \frac{e^\varphi - e^{-\varphi}}{2}$

Beweis: Sei $Q = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ und $Q^\top \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} Q = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

$$\begin{aligned} \iff & \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ -c & -d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ \iff & a^2 - c^2 = 1, \quad ab - cd = 0, \quad b^2 - d^2 = -1 \\ & a^2 = 1 + c^2 \implies a = \pm\sqrt{1 + c^2} \\ & b^2 = -1 + d^2 \implies b = \pm\sqrt{d^2 - 1} \\ \iff & \pm\sqrt{1 + c^2}\sqrt{d^2 - 1} - cd = 0 \\ \implies & (1 + c^2)(d^2 - 1) = c^2 d^2 \\ & d^2 + c^2 d^2 - c^2 - 1 = c^2 d^2 \\ \implies & d^2 - c^2 = 1 \end{aligned}$$

Mit $a^2 - c^2 = 1 \implies d^2 - a^2 = 0 \implies d = \pm a$ und mit $ab = cd \implies c = \pm b$.

Also gilt $Q = \begin{bmatrix} a & b \\ \varepsilon b & \varepsilon a \end{bmatrix}$ mit $\varepsilon \in \{\pm 1\}$ und $a^2 = 1 + b^2$.

Die Gleichung $a = \pm\sqrt{1 + b^2}$ definiert eine Hyperbel.

Für die erste Zeile ergibt sich daraus

$$\hat{A} = \left[\begin{array}{c|cc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & * & \\ 0 & & & \end{array} \right] \quad \text{und} \quad \hat{B} = \left[\begin{array}{c|cc} 0 & \cdots & 0 \\ \hline & * & \end{array} \right].$$

Per Induktion folgt dann die Behauptung. Es ist dabei klar, dass wir jederzeit noch mit Diagonalmatrizen mit $\{\pm 1\}$ in der Diagonale multiplizieren können. \square

Beachte: Hyperbolische Rotationen sind keine orthogonalen Matrizen, denn i.a. gilt $H_{ij}^\top I H_{ij} \neq I$. Sie sind nur orthogonal im Sinne des Skalarprodukts $\langle \cdot, \cdot \rangle_J$.

Die Zerlegung von $G \in \mathcal{S}_J$ als Produkt von Rotationen und hyperbolischen Rotationen ist natürlich nicht eindeutig. Allerdings lassen sich die Elemente von \mathcal{S}_J über die Winkel φ_{ij} der Rotationen und die Vorzeichen parametrisieren. Man kann zeigen, dass $\frac{n(n-1)}{2}$ Winkel in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ für Rotationen bzw. Parameter φ_{ij} für hyperbolische Rotationen und n Vorzeichen (sind eigentlich auch Winkel) jedes Element charakterisieren.

Beispiel 15.8 $J = \left[\begin{array}{c|c} 1 & \\ \hline & -1 \end{array} \right]$

$$G = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{3\sqrt{2}}{2} & \frac{3\sqrt{2}}{2} & 2\sqrt{2} \\ -2 & 2 & 3 \end{bmatrix}, \quad H_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 2\sqrt{2} \\ 0 & 2\sqrt{2} & 3 \end{bmatrix},$$

$$H_{23}^{-1} = \frac{1}{9 - (2\sqrt{2})^2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -2\sqrt{2} \\ 0 & -2\sqrt{2} & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -2\sqrt{2} \\ 0 & -2\sqrt{2} & 3 \end{bmatrix},$$

$$\varphi_{23} = \operatorname{arcosh}(3),$$

$$H_{23}^{-1}G = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \varphi_{12} = \frac{\pi}{4}, \quad \varphi_{13} = 0.$$

Korollar 15.9 Sei $J = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$, so gilt $\det G = \pm 1$ für alle $G \in \mathcal{S}_J$.

Beweis: $\det(G^\top JG) = \det(J) \implies \det(G^2) \cdot \det(J) = \det(J) \implies (\det G)^2 = 1 \implies \det(G) = \pm 1$. \square

Eine weitere wichtige Matrixgruppe sind die Matrizen in $\mathbb{R}^{n,n}$ mit Determinante 1. Diese Gruppe wird mit $SL_n(\mathbb{R})$ bezeichnet.

Satz 15.10 Sei $J = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix}$, $p + q = n$. Die folgenden Klassen von Matrizen bilden Untergruppen von $GL_n(\mathbb{R})$.

- (a) $\mathcal{S}_J = \{G \in \mathbb{R}^{n,n} \mid G^\top JG = J\}$
- (b) $SL_n(\mathbb{R}) = \{A \in \mathbb{R}^{n,n} \mid \det A = 1\}$
- (c) $SL_n(\mathbb{R}) \cap \mathcal{S}_J$
- (d) $SL_n(\mathbb{R}) \cap \mathcal{O}_n(\mathbb{R})$
- (e) $\mathcal{S}_J \cap \mathcal{O}_n(\mathbb{R})$

Beweis: Übung □

Analoge Betrachtungen lassen sich auch im Komplexen durchführen.

Wir hatten in Teil 2 gesehen, dass orthogonale (unitäre Matrizen) alle Eigenwerte auf dem Einheitskreis haben. Kann man für J -orthogonale Matrizen etwas analoges zeigen?

Betrachte $Q \in \mathcal{S}_J$ mit $J = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$, $p + q = n$.

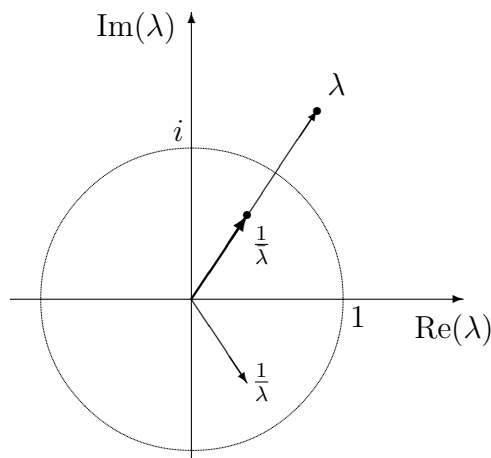
Sei $Qx = \lambda x$ mit $x \neq 0 \implies Jx = Q^\top JQx = \lambda Q^\top Jx$, also ist $y = (x^\top J^\top)$ Linkseigenvektor von Q und es gilt

$$y^\top = \lambda y^\top Q. \tag{15.11}$$

Da $Q \in \mathcal{S}_J$ nichtsingulär ist, gilt $\lambda \neq 0$ und es folgt

$$y^\top Q = \frac{1}{\lambda} y^\top \quad \text{oder} \quad Q^\top y = \frac{1}{\lambda} y.$$

Da aber Q und Q^\top die gleichen Eigenwerte haben, ergibt sich, dass mit λ auch immer $\frac{1}{\lambda}$ ein Eigenwert ist (im Komplexen $\frac{1}{\bar{\lambda}}$).



$$\begin{aligned} \lambda &= a + ib \\ \frac{1}{\bar{\lambda}} &= \frac{a + ib}{a^2 + b^2} \\ \frac{1}{\lambda} &= \frac{a - ib}{a^2 + b^2} \end{aligned}$$

Es folgt, dass für $p + q$ ungerade mindestens ein Eigenwert auf dem Einheitskreis liegen muss, und zwar im reellen Fall bei $+1$ oder -1 , denn sonst kann nicht $\lambda = \frac{1}{\lambda}$ sein.

Man kann für Matrizen aus \mathcal{S}_J auch eine Jordanform bestimmen unter Transformation mit Matrizen aus \mathcal{S}_J , d.h., man sucht für $Q \in \mathcal{S}_J$ ein $S \in \mathcal{S}_J$, so dass $S^{-1}QS$ in einer Normalform ist, die so etwas wie die Jordan'sche Normalform ist. Diese Form wurde erst kürzlich das erste Mal bestimmt und ist ziemlich schrecklich (siehe meine Webseite, bzw. <http://www.tu-chemnitz.de/sfb393/preprints.html>, Nr. 98-07 und 98-29)

Was kann man nun im J -symmetrischen oder J -schiefsymmetrischen Fall aussagen? Dazu führen wir zwei neue Matrixmultiplikationen ein.

Definition 15.12 Sei $J = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$, $p + q = n$ und

$$\mathcal{A}_J = \{A \in \mathbb{R}^{n,n} \mid (AJ)^\top = AJ\}$$

die Menge der J -symmetrischen Matrizen in $\mathbb{R}^{n,n}$ und

$$\mathcal{C}_J = \{A \in \mathbb{R}^{n,n} \mid (AJ)^\top = -AJ\}$$

die Menge der J -schiefsymmetrischen Matrizen in $\mathbb{R}^{n,n}$.
Definiere für $A, B \in \mathcal{A}_J$ die Multiplikation

$$(A, B) = AB + BA \tag{15.13}$$

und für $A, B \in \mathcal{C}_J$ die Multiplikation

$$[A, B] = AB - BA. \tag{15.14}$$

Lemma 15.15

- (i) Mit $A, B \in \mathcal{A}_J$ ist auch $(A, B) \in \mathcal{A}_J$.
- (ii) Mit $A, B \in \mathcal{C}_J$ ist auch $[A, B] \in \mathcal{C}_J$.
- (iii) Für $A \in \mathcal{C}_J$ ist $A^2 \in \mathcal{A}_J$.

Beweis: Beachte $J^\top = J$.

- (i) $A, B \in \mathcal{A}_J \implies (AJ)^\top = AJ, (BJ)^\top = BJ$
 $\implies (AB + BA)J = ABJ + BAJ$
 $= AJ^\top B^\top + BJ^\top A^\top = AJB^\top + BJA^\top$
 $= J^\top A^\top B^\top + J^\top B^\top A^\top = J(AB + BA)^\top.$
- (ii) $A, B \in \mathcal{C}_J \implies (AJ)^\top = -AJ, (BJ)^\top = -BJ$
 $\implies (AB - BA)J = ABJ - BAJ$
 $= -AJ^\top B^\top + BJ^\top A^\top = JA^\top B^\top - JB^\top A^\top$
 $= J(AB - BA)^\top$
- (iii) $A^2 J = -AJA^\top = J(A^\top)^2.$

□

Definition 15.16 Ein Vektorraum V mit einer bilinearen Operation „Multiplikation“

$$\begin{aligned} * & : V \times V \rightarrow V \\ (x, y) & \rightarrow x * y = xy \end{aligned}$$

heißt nichtassoziative Algebra. Falls die Operation $*$ assoziativ ist, so heißt die Algebra assoziativ, d.h., es gilt $(xy)z = x(yz)$.

Beispiel 15.17 Die Mengen $\mathbb{R}^{n,n}, \mathbb{C}^{n,n}, \mathbb{F}^{n,n}$ sind assoziative Algebren mit der normalen Matrixmultiplikation. Mit der Multiplikation $[A, B] = AB - BA$ sind sie nichtassoziative Algebren.

Definition 15.18 Ein Vektorraum L mit einem Produkt $[\cdot, \cdot]$, für das gilt

$$[x, y] = -[y, x], \quad \forall x, y \in L, \quad (15.19)$$

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0, \quad \forall x, y, z \in L \quad (\text{Jacobi-Identität}) \quad (15.20)$$

heißt Lie-Algebra.

Beispiel 15.21

(i) \mathbb{R}^3 mit dem \times -Produkt (Vektorprodukt) ist eine Lie-Algebra, denn es gilt

$$a \times (b \times c) = (a^\top c)b - (a^\top b)c.$$

(ii) Für $J = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix}$ ist $\mathcal{C}_J = \{A \in \mathbb{R}^{p+q, p+q} \mid AJ + J^\top A^\top = 0\}$ eine Lie-Algebra mit dem Lie-Produkt $[A, B] = AB - BA$.

Definition 15.22 Eine Derivation einer nichtassoziativen Algebra \mathcal{A} ist ein linearer Operator d auf \mathcal{A} , der die formale Leibniz-Regel für Ableitungen erfüllt.

$$d(xy) = (dx)y + x(dy) \quad \forall x, y \in \mathcal{A} \quad (15.23)$$

Beispiel 15.24 Für die Algebra der Polynome in x und ein festes Polynom $p(x)$ ist

$d = p(x) \frac{d}{dx}$ eine Derivation.

Die Menge der Derivationen einer nichtassoziativen Algebra \mathcal{A} ist ein Untervektorraum von der Menge aller linearen Abbildungen von \mathcal{A} in \mathcal{A} .

Das Produkt von Derivationen ist im allgemeinen keine Derivation jedoch das Lie-Produkt $[\cdot, \cdot]$. Denn für $[d_1, d_2]$ gilt:

$$\begin{aligned} (d_1 d_2 - d_2 d_1)(xy) &= d_1(d_2(xy)) - d_2(d_1(xy)) \\ &= d_1((d_2 x)y + x(d_2 y)) - d_2((d_1 x)y + x(d_1 y)) \\ &= (d_1 d_2 x)y + (d_2 x)(d_1 y) + (d_1 x)(d_2 y) + x(d_1 d_2 y) \\ &\quad - (d_2 d_1 x)y - (d_1 x)(d_2 y) - (d_2 x)(d_1 y) - x(d_2 d_1 y) \\ &= ((d_1 d_2 - d_2 d_1)x)y + x((d_1 d_2 - d_2 d_1)y). \end{aligned}$$

Korollar 15.25 Sei \mathcal{A} eine nichtassoziative Algebra und D die Menge der Derivationen von \mathcal{A} , so ist D eine Lie-Algebra. Man nennt D die Derivationsalgebra.

Bevor wir den Zusammenhang zwischen den Lie-Algebren und den Matrixgruppen studieren, wollen wir noch eine weitere klassische Matrixgruppe einführen.

Definition 15.26 Sei $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$. Die Menge der J -orthogonalen Matrizen in $\mathbb{R}^{2n,2n}$ heißt symplektische Gruppe.

$$Sp_{2n}(\mathbb{R}) = \{Q \in \mathbb{R}^{2n,2n} \mid Q^\top J Q = J\}.$$

Lemma 15.27

(a) Sei $K \in \mathbb{R}^{2n,2n}$ schiefssymmetrisch und nichtsingulär, so ist K kongruent zu J aus Definition 15.26.

(b) Sei $K \in \mathbb{R}^{n,n}$ symmetrisch und nichtsingulär, so ist K kongruent zu $\begin{bmatrix} I_p & \\ & -I_q \end{bmatrix}$.

Beweis:

(a) Da K reell schiefssymmetrisch ist, gibt es Q orthogonal, so dass

$$Q^\top K Q = \begin{bmatrix} K_{11} & & \\ & \ddots & \\ & & K_{ss} \end{bmatrix}$$

mit K_{ii} reell-schiefssymmetrisch 1×1 oder 2×2 (je nachdem, ob die Eigenwerte reell oder komplex konjugierte Paare sind). Da aber alle Eigenwerte von K auf der imaginären Achse liegen und K nichtsingulär ist, so folgt, dass alle K_{ii} die Form $\begin{bmatrix} 0 & a_i \\ -a_i & 0 \end{bmatrix}$ haben und Paare komplex konjugierter Eigenwerte haben, also $s = n$

Sei $D = \text{diag}(D_{11}, \dots, D_{nn})$ mit $D_{ii} = \begin{bmatrix} \frac{1}{\sqrt{a_i}} & 0 \\ 0 & \frac{\text{sgn}(a_i)}{\sqrt{a_i}} \end{bmatrix}$, so folgt

$$D^\top Q^\top K Q D = \begin{bmatrix} 0 & 1 & & \\ -1 & 0 & & \\ & & \ddots & \\ & & & 0 & 1 \\ & & & -1 & 0 \end{bmatrix}.$$

Mit $P = [e_1, e_3, \dots, e_{2n-1}, e_2, e_4, \dots, e_{2n}]$ ist dann $P^\top D^\top Q^\top K Q D P = J$.

(b) klar.

□

Damit können wir uns auf die Fälle

$$J = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix} \text{ oder } J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

beschränken.

Lemma 15.28 Sei $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$. Die Menge der J -symmetrischen Matrizen

$$\mathcal{A}_J(\mathbb{R}) = \{A \in \mathbb{R}^{n,n} \mid (AJ)^\top = AJ\}$$

ist eine Lie-Algebra mit dem Produkt $[\cdot, \cdot]$. Alle Matrizen in $\mathcal{A}_J(\mathbb{R})$ haben die Form

$$A = \begin{bmatrix} F & G \\ H & -F^\top \end{bmatrix} \text{ mit } H = H^\top, G = G^\top.$$

Diese Matrizen heißen Hamiltonische Matrizen.

Beweis:

$$A = \begin{bmatrix} F & G \\ H & L \end{bmatrix} \in \mathcal{A}_J(\mathbb{R})$$

$$\implies \begin{bmatrix} -G & F \\ -L & H \end{bmatrix} = (AJ)^\top = AJ = \begin{bmatrix} -G^\top & -L^\top \\ F^\top & H^\top \end{bmatrix}$$

$$\implies G = G^\top, H = H^\top, L = -F^\top.$$

$A, B \in \mathcal{A}_J(\mathbb{R})$

$$\begin{aligned} \implies (AB - BA)J &= ABJ - BAJ \\ &= AJ^\top B^\top - BJ^\top A^\top \\ &\stackrel{(J^\top = -J)}{=} -AJB^\top + BJA^\top \\ &= -J^\top A^\top B^\top + J^\top B^\top A^\top \\ &= J^\top (B^\top A^\top - A^\top B^\top) \end{aligned}$$

□

Lemma 15.29 Sei $X \in GL_n(\mathbb{R})$. Falls AX symmetrisch (schiefsymmetrisch) für alle symmetrischen (schiefsymmetrischen) Matrizen $A \in \mathbb{R}^{n,n}$, so gilt

$$X = \alpha I_n \quad \text{mit} \quad \alpha \in \mathbb{R} \setminus \{0\}.$$

Beweis: Es folgt mit $A = I_n$, dass X symmetrisch ist. Mit $A = e_k e_j^\top + e_j e_k^\top$ folgt

$$X(e_k e_j^\top + e_j e_k^\top) = (e_k e_j^\top + e_j e_k^\top)X, \quad \forall j, k \in \{1, \dots, n\}.$$

$$\begin{aligned} \implies x_{rk} e_j^\top e_l + x_{rj} e_k^\top e_l &= e_r^\top e_k x_{jl} + e_r^\top e_j x_{kl} \quad \forall j, k, r, l \in \{1, \dots, n\} \\ \implies \text{für } k \neq r, l = j = k &\text{ folgt } 2x_{rk} = 0 \text{ und} \\ \text{für } j = r, l = k &\text{ folgt } x_{kk} = x_{jj} \implies x = \alpha I_n \end{aligned}$$

Im schiefssymmetrischen Fall analoger Beweis (Übung). □

Satz 15.30 Sei $J = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$ oder $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ und sei $K = \mathcal{A}_J(\mathbb{R})$ oder $K = \mathcal{C}_J(\mathbb{R})$. Dann gilt $T^{-1}MT \in K$ für alle $M \in K$ genau dann, wenn

$$T^\top J T = \alpha J \quad \text{für } \alpha \in \mathbb{R} \setminus \{0\}$$

(wobei natürlich das passende J zu wählen ist).

Beweis: Falls $T^\top J T = \alpha J$ (oder $T^\top J = \alpha J T^{-1}$) ist, so gilt für $M \in K$, dass

$$\begin{aligned} J(T^{-1}MT) &= \frac{1}{\alpha} T^\top J M T = \pm \frac{1}{\alpha} T^\top M^\top J^\top T = \pm \frac{1}{\alpha} T^\top M^\top \alpha T^{-\top} J^\top \\ &= \pm (T^{-1}MT)^\top J^\top. \end{aligned}$$

Umgekehrt, falls für $M \in K$ auch $T^{-1}MT \in K$, so folgt

$$\begin{aligned} J(T^{-1}MT) &= \pm (T^{-1}MT)^\top J^\top = \pm T^\top M^\top T^{-\top} J^\top = \pm T^\top M^\top J^\top J T^{-\top} J^\top \\ &= T^\top J M J T^{-\top} J^\top, \quad \implies \\ JM(T J T^\top J^\top) &= (J T J^\top T^\top) J M \end{aligned}$$

und damit ist $JM(T J T^\top J^\top)$ symmetrisch.

Wenn also $T^{-1}MT \in K \quad \forall M \in K$, so ist $T J T^\top J^\top$ symmetrisch für alle symmetrischen (schiefsymmetrischen) $A = JM$, also folgt nach Lemma 15.29

$$J T J^\top T^\top = \alpha I \quad \implies \quad T J^\top T^\top = \alpha J^\top \text{ oder } T J T^\top = \alpha J \quad \implies \quad J T^\top = \alpha T^{-1} J$$

$$\begin{aligned} \implies J T^\top J^\top T &= \alpha T^{-1} J J^\top T = \alpha I \\ \implies T^\top J^\top T &= \alpha J^\top \\ \implies T^\top J T &= \alpha J. \end{aligned}$$

□

Wir erhalten also, dass Ähnlichkeitstransformation mit Elementen der Gruppe die entsprechende Algebra erhalten und eigentlich nur diese (bis auf eine Konstante). Man kann nun für Matrizen in $\mathcal{A}_J, \mathcal{C}_J$ mit $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ eine Schurform mit Matrizen in $\mathcal{O}_{2n}(\mathbb{R}) \cap Sp_{2n}(\mathbb{R})$ herleiten.

Satz 15.31 Sei $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$.

(i) Sei $H \in \mathcal{A}_J = \{A \in \mathbb{R}^{2n,2n} \mid (AJ)^\top = AJ\}$ und H habe keine Eigenwerte mit Realteil 0, so gibt es $S \in \mathcal{O}_{2n}(\mathbb{R}) \cap Sp_{2n}(\mathbb{R})$, so dass

$$S^{-1}HS = S^\top HS = \begin{bmatrix} F & G \\ 0 & -F^\top \end{bmatrix}$$

mit F in reeller Schurform, d.h. $F = \begin{bmatrix} F_{11} & \cdots & F_{1k} \\ & \ddots & \vdots \\ & & F_{kk} \end{bmatrix}$ mit F_{ii} 1×1 oder 2×2 .

(ii) Sei $H \in \mathcal{C}_J = \{A \in \mathbb{R}^{2n,2n} \mid (AJ)^\top = -AJ\}$, so gibt es $S \in \mathcal{O}_{2n}(\mathbb{R}) \cap Sp_{2n}(\mathbb{R})$, so dass

$$S^{-1}HS = S^\top HS = \begin{bmatrix} F & G \\ 0 & F^\top \end{bmatrix}$$

mit F in reeller Schurform.

Beweis: Wir haben bereits die Schurform für allgemeine Matrizen betrachtet. Der Beweis geht analog und soll hier nicht gemacht werden. \square

Bemerkung: Man kann diesen Satz als notwendige und hinreichende Bedingung formulieren und dabei Eigenwerte auf der imaginären Achse zulassen.

Siehe Webseite.

Inhaltsverzeichnis

1	Lineare Differentialgleichungen	1
2	Lineare differentiell–algebraische Gleichungen	7
3	Bilinearform und Hermite’sche Form	15
4	Hermite’sche, symmetrische Matrizen, adjungierte Endomorphismen	24
5	Die Singulärwertzerlegung	34
6	Die Hauptachsentransformation	40
7	Bewegung starrer Körper	47
8	Quadriken	50
9	Positiv definite Matrizen (Bilinearformen)	64
10	Eigenwert– und Eigenvektorberechnung	69
11	Projektive Räume und projektive Geometrie	76
12	Das Kroneckerprodukt	91
13	Lineare Matrixgleichungen	98
14	Die Stabilitätstheorie von Lyapunov	107
15	Matrixgruppen	115

Index

Numbers written in *italic* refer to the page where the corresponding entry is described, the ones underlined to the definition, the rest to the pages where the entry is used.

- A**
- abstandserhaltend 51
 - ähnlich 31
 - Ähnlichkeitstransformation 125
 - Äquivalenz 37, 40
 - affin äquivalent 62
 - affin linear 62
 - Algebra 125
 - (nicht-)assoziativ . . . 122
 - asymptotisch stabil 107
- B**
- Bilinearform . . 15, 40, 41, 64
- C**
- Cauchy'scher Trennungssatz 68
 - Cayley-Transformation 112, 113
 - Cholesky-Zerlegung . . 64, 65
 - Courant-Fischer 66
- D**
- Derivation 122
 - Derivationsalgebra 123
 - Desarque, Satz von 89
 - diagonalisierbar 31
 - simultan 29
 - Differentialgleichung 1, 99, 107
 - Differentialgleichungssystem 2, 5
 - Differenzgleichung . . . 107
 - Doppelverhältnis 84
- E**
- ebene Geometrie 88
 - Eigenvektor 69
 - Eigenwert 64, 69, 96, 107, 113
 - verallgemeinert 8
 - Ellipsenkegel 60
 - Ellipsenzylinder 61
 - euklidischer Raum . . 18, 24
 - Exponentialfunktion e^A . . 5
- F**
- Frobeniusnorm 34
- G**
- Gruppe 115, 119
- H**
- Hamiltonische Matrizen . 124
 - harmonisch 86
 - Hauptabschnittsmatrix . . 68
- Form 15, 64, 108
- Matrix 64
- Hessenberggestalt 73
- homogene Koordinaten . . 77
- Hyperbelzylinder 61
- Hyperboloid 61
- Hyperfläche 50
- I**
- Index 12
- J**
- Jacobi-Identität 122
 - Jacobi-Verfahren 73
 - Jordan'sche Normalform 1, 5
- K**
- Klassifikation, Quadriken 53, 58
 - kollinear 84, 86
 - Kongruenz 18, 40
 - Kongruenzabbildung 51
 - Kongruenztransformation 52
 - Kroneckerprodukt 91
- L**
- Leibniz-Regel 122
 - Lie-Algebra . . 122, 123, 124
 - Lie-Produkt 122
 - Lyapunov 97, 108
 - Satz von 109
 - Lyapunov-Gleichung . . . 100
- M**
- Matrix
 - Exponentialfunktion . . 5
 - Matrixdarstellung 40, 41, 52
 - Matrixgleichung 98
 - Matrixgruppen . . . 115, 119
 - Moore–Penrose–Inverse . . 38
- N**
- negativ definit 46
 - nichtlineare Abbildung . . 50
 - normal 30, 31, 33
 - Normalform
 - Kongruenz 45
 - Weierstraß 9
- O**
- Orientierung 47
 - orthogonal 30
 - Endomorphismus . 18, 30
 - orthogonale Iteration 71
 - orthogonaler Projektor . . 69
- P**
- Parabelzylinder 51
 - parabolischer Zylinder . . . 62
 - Paraboloid 62
 - positiv definit 43, 46, 64, 108, 111, 113
 - Potenzmethode 69, 71
 - projektive Abbildung . . . 78
 - projektive Basis 78
 - projektive Dimension . . . 76
 - projektive Ebene 76
 - projektive Gerade 76, 81–84, 88
 - projektiver Raum 76
 - projektives Koordinatensystem 80
 - Projektivität . . 78, 79–81, 85
- Q**
- QR-Algorithmus 72
 - quadratisches Polynom . . 50
 - Quadrik 50, 52
 - Klassifikation 53–62
- R**
- Rotationsenergie 49
- S**
- Satz von Schur 27, 65
 - Schur-Form 21
 - Schurform 125
 - Schwingungsgleichung . . . 1
 - selbstadjungiert 25
 - Endomorphismus . 29, 30
 - semidefinit 43, 46
 - Singularwert 36
 - Singularwertzerlegung 36, 40, 69
 - Spektralnorm 34
 - stabil 107
 - Stabilitätstheorie 97
 - Stein-Gleichung 101, 112, 113
 - Stephanos, Satz von 96
 - Störungstheorie 68
 - SVD 36
 - Sylvester, Trägheitssatz . . 44
 - Sylvester-Gleichung 98, 101, 104
 - diskrete 100
 - symmetrisch 30
 - Bilinearform 15
 - symplektisch 115
 - symplektische Gruppe . . 123
- T**
- Tensorprodukt 91

Trägheitssatz	<u>44</u>	unitär	30, 31	W	
Translationsenergie	49	Endomorphismus	18, 30	Weierstraß-Normalform	9
Trennungssatz	<u>68</u>	unitärer Raum	18, 24	Z	
	U	Untergruppe	116, 120	Zentralprojektion	81
unendlich ferne			V		
Gerade	83	vec-Operation	94		
Hyperebene	78	Verbindungsraum	81		