

Lineare Algebra II

Nicolas Bartsch
nbarzsch@ethz.ch

Frühling 2025

Vorwort

Dieses Skript basiert auf den Unterlagen meiner Übungen der Linearen Algebra II aus dem Semester FS25. Trotz Revisionen kann ich **weder die Vollständigkeit noch die Korrektheit garantieren**. Es ist möglich, dass kleine Fehler enthalten sind. Falls dir ein solcher Fehler auffällt, wäre ich dir dankbar, wenn du mich per E-Mail darüber informierst, damit das Skript korrigiert werden kann.

Eine aktuelle Version dieses Skripts sowie weitere Unterlagen findest du auf meiner Website: n.ethz.ch/~nbartzs/.

Vielen Dank und viel Erfolg mit Lin Alg II.

Nicolas Bartsch

Version: 1. Juli 2025

Inhaltsverzeichnis

0	Wiederholung: Vektorräume und Basis	1
0.1	Vektorräume	1
0.2	Unterräume	2
0.3	Erzeugendensystem und Basis	3
1	Lineare Abbildungen	5
2	Eigenwertproblem	6
2.1	Eigenwerte	6
2.2	Eigenvektoren	7
3	Normen und Skalarprodukte	10
3.1	Normen	10
3.2	Skalarprodukte	11
3.3	Vom Skalarprodukt induzierte Norm	12
4	Orthonormalbasis	13
4.1	Orthogonalprojektion	13
4.2	Orthonormalbasis	13
4.3	Gram-Schmidt-Orthogonalisierungsverfahren	14
5	Bild und Kern	18
5.1	Bild	18
5.2	Kern	19
5.3	Beziehung zwischen Kern und Bild	21
6	Basiswechsel und Diagonalisierung	25
6.1	Basiswechsel für Vektoren	25
6.2	Basiswechsel für Matrizen	27
6.3	Diagonalisieren	28
6.4	Diagonalisierbarkeit	30
7	Quadratische Formen	33
7.1	Quadratische Formen	33
7.2	Definitheit	34
7.3	Rein quadratische Formen	35
8	Kegelschnitte und Quadriken	37
8.1	Kegelschnitte	37
8.2	Quadriken	38
8.3	Hauptachsentransformation und Normalformen	39
9	Methode der Kleinsten Quadrate und QR-Zerlegung	43
9.1	Methode der kleinsten Quadrate	43
9.2	QR-Zerlegung	46
9.3	Kleinste Quadrate mit QR-Zerlegung	48
10	Lineare Differentialgleichungssysteme	49

10.1	Homogene lineare Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten	49
10.2	Systeme homogener linearer Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten	49
10.3	Lineare Differentialgleichungen höherer Ordnung	53
10.4	Spezialfall: Differentialgleichungssysteme mit komplexen Eigenwerten . .	54

0 Wiederholung: Vektorräume und Basis

Bevor wir uns mit dem neuen Stoff beschäftigen, blicken wir zunächst zurück in die Lineare Algebra I. (Für eine komplette Wiederholung siehe Lineare Algebra I, Woche 10). Folgende grundlegende Zusammenhänge für LGS und Matrizen sind weiterhin wichtig und werden folgend erweitert:

Wichtige Zusammenhänge

Folgende Aussagen sind für $A \in \mathbb{R}^{n \times n}$ äquivalent:

- $\text{rang}(A) = n$
- Das LGS $Ax = b$ ist für beliebiges b lösbar
- Das LGS $Ax = b$ hat genau eine Lösung
- Das homogene LGS $Ax = 0$ hat nur die triviale Lösung $x = 0$
- Die Zeilen und Spalten von A sind linear unabhängig
- A ist invertierbar (regulär, nicht singulär)
- $\det(A) \neq 0$
- Die Spalten von A bilden eine Basis von \mathbb{R}^n

0.1 Vektorräume

Wir wollen unsere Vorstellung von Vektoraddition und Multiplikation, von den uns bekannten Vektorpfeilen im Raum und Ebene, auf alle möglichen Vektoren abstrahieren (verallgemeinern). Damit können wir später die Eigenschaften von Vektoren auf z.B. Funktionen (welche sich als Vektoren darstellen lassen) anwenden. Formell können wir Vektorräume wie folgt definieren:

Vektorräume

Sei V eine Menge von Objekten. V heisst Vektorraum, wenn eine **innere Operation** (Kombination von zwei Objekten) und eine **äussere Operation** (Kombination eines Objekts mit einem Skalar) definiert sind:

Innere Operation:

$$\begin{aligned}\oplus : V \times V &\rightarrow V \\ (a, b) &\mapsto a \oplus b\end{aligned}$$

Äussere Operation:

$$\begin{aligned}\odot : \mathbb{K} \times V &\rightarrow V \\ (\alpha, a) &\mapsto \alpha \odot a\end{aligned}$$

Es müssen nun sowohl eine innere als auch eine äussere Operation definiert werden.

Schauen wir uns diese einmal genauer an.

Innere Operation:

$$\oplus : V \times V \rightarrow V$$

Hier nimmt \oplus zwei Vektoren a, b aus einer Menge V und ordnet dem Paar einen anderen Vektor in V zu.

$$(a, b) \mapsto a \oplus b$$

zeigt uns dann genau wie diese Operation definiert ist.

Äussere Operation:

$$\odot : \mathbb{K} \times V \rightarrow V$$

Hier nimmt \odot einen Vektor a aus einer Menge V und einen Skalar α und produziert einen neuen Vektor in V .

$$(\alpha, a) \mapsto \alpha \odot a$$

zeigt uns dann genau wie diese Operation definiert ist.

Basierend auf diesen Operationen müssen nun folgende Axiome gelten:

Axiome für Vektorräume

(A1) $\forall u, v \in V :$	$u \oplus v = v \oplus u$
(A2) $\forall u, v, w \in V :$	$(u \oplus v) \oplus w = u \oplus (v \oplus w)$
(A3) $\exists 0 \in V, \forall u \in V :$	$u \oplus 0 = u$
(A4) $\forall u \in V, \exists -u \in V :$	$u \oplus (-u) = 0$
(M1) $\forall \alpha, \beta \in \mathbb{R}, \forall u \in V :$	$\alpha \odot (\beta \odot u) = (\alpha \cdot \beta) \odot u$
(M2) $\forall \alpha, \beta \in \mathbb{R}, \forall u, v \in V :$	$(\alpha + \beta) \odot u = (\alpha \odot u) \oplus (\beta \odot u)$
(M3) $\forall u \in V :$	$1 \odot u = u$

Nun kennen wir alle Regeln und können überprüfen, ob eine Menge ein Vektorraum beschreibt. Dafür müssen wir **alle** Axiome überprüfen. Ausschlaggebend sind hier die Operationen nicht die Objekte.

0.2 Unterräume

Ein Unterraum ist eine nichtleere Teilmenge U eines Vektorraums V , welche folgende Eigenschaften erfüllt:

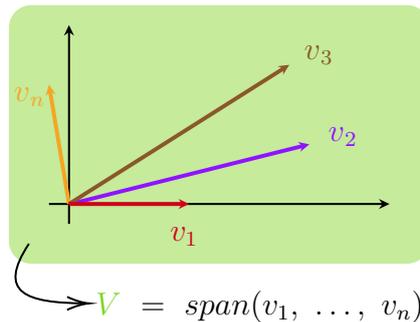
- (i) $\forall u, v \in U : u \oplus v \in U$
- (ii) $\forall \alpha \in \mathbb{R}, \forall u \in U : \alpha \odot u \in U$

Dabei sagt uns (i), dass die Summe von zwei Elementen aus U weiterhin ein Teil von U ist. (ii) sagt uns, dass wenn ein Element aus U mit einem Skalar multipliziert wird, so ist

das skalierte Element weiterhin ein Teil von U . Aus diesen Eigenschaften folgt, dass ein Unterraum auch ein Vektorraum ist und immer den Nullvektor enthält.

0.3 Erzeugendensystem und Basis

Sei $v := \sum_{i=1}^n x_i v_i$, dann ist v eine Linearkombination von v_1, \dots, v_n . Die Menge aller Linearkombinationen nennt sich lineare Hülle und wird mit $\text{span}(v_1, \dots, v_n)$ abgekürzt. Diese Menge aller Linearkombinationen ist ein Vektorraum V . Die Vektoren v_1, \dots, v_n sind dann ein Erzeugendensystem von V . Es lassen sich alle Vektoren in V durch Linearkombinationen der Erzeugendenvektoren bilden.



Wenn alle Vektoren v_1, \dots, v_n eines Erzeugendensystems linear unabhängig sind, dann bilden sie eine Basis. Die Vektoren heißen dann Basisvektoren.

In 2-D wird oft die Standardbasis verwendet. Diese besteht aus den Vektoren

$$\vec{e}_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ und } \vec{e}_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Jeder 2-D Vektor kann eindeutig als Linearkombination dieser beiden Basisvektoren dargestellt werden. Wir können jedoch auch zwei andere, linear unabhängige, Vektoren als Basisvektoren verwenden. Seien beispielsweise

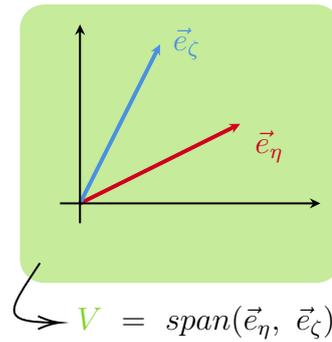
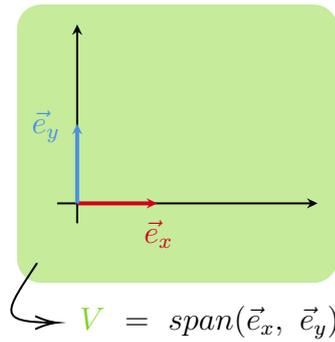
$$\vec{e}_\eta = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ und } \vec{e}_\zeta = \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

auch hier können wir jeden 2-D Vektor eindeutig als Linearkombination dieser beiden Vektoren darstellen. Beide Basen beschreiben hier also denselben Vektorraum. Wir sehen, dass die Anzahl von Basisvektoren erhalten bleibt. Die Anzahl an Basisvektoren nennt man Dimension. In unserem Fall mit $V = \text{span}(\vec{e}_x, \vec{e}_y) = \text{span}(\vec{e}_\eta, \vec{e}_\zeta)$ hat der Vektorraum V die Dimension 2. Es ist also möglich mehrere Basen für einen endlichdimensionalen Vektorraum zu finden.

Versuchen wir nun einen spezifischen Vektor in der Standardbasis \vec{e}_x, \vec{e}_y und der Basis $\vec{e}_\eta, \vec{e}_\zeta$ auszudrücken. Betrachten wir hierfür zunächst den Vektor

$$\begin{pmatrix} 3 \\ 3 \end{pmatrix} = \sum_{i=1}^n x_i v_i.$$

In der Standardbasis suchen wir also x_1, x_2 sodass



$$\begin{pmatrix} 3 \\ 3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Wir erkennen sofort, dass $x_1 = x_2 = 3$ sein muss. Somit ist der Koordinatenvektor bezüglich der Standardbasis \vec{e}_x, \vec{e}_y :

$$\begin{pmatrix} 3 \\ 3 \end{pmatrix}_{x,y}.$$

In der Basis $\vec{e}_\eta, \vec{e}_\zeta$ suchen wir nun x_1, x_2 sodass

$$\begin{pmatrix} 3 \\ 3 \end{pmatrix} = x_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Hier erkennen wir sofort, dass $x_1 = x_2 = 1$ sein muss. Somit ist der Koordinatenvektor bezüglich der Basis $\vec{e}_\eta, \vec{e}_\zeta$:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\eta,\zeta}.$$

Der Koordinatenvektor hängt also immer von der gewählten Basis ab!

1 Lineare Abbildungen

Lineare Abbildung

Seien V, W reelle Vektorräume. Dann heisst

$$\mathcal{F} : V \rightarrow W, \quad x \mapsto \mathcal{F}(x)$$

lineare Abbildung, falls $\forall x, y \in V, \forall \alpha \in \mathbb{R}$ gilt:

$$i. \quad \mathcal{F}(x + y) = \mathcal{F}(x) + \mathcal{F}(y)$$

$$ii. \quad \mathcal{F}(\alpha x) = \alpha \mathcal{F}(x).$$

Wenn wir nun prüfen wollen, ob eine Abbildung linear ist, dann müssen wir prüfen ob i. und ii. erfüllt sind. Als Beispiel betrachten wir die Abbildung

$$\mathcal{F} : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto 3x.$$

Wir prüfen nun i. und ii. für diese Abbildung:

$$i. \quad \mathcal{F}(x + y) = 3(x + y) = 3x + 3y = \mathcal{F}(x) + \mathcal{F}(y)$$

$$ii. \quad \mathcal{F}(\alpha x) = 3(\alpha x) = \alpha \cdot 3x = \alpha \cdot \mathcal{F}(x).$$

In diesem Fall ist die Abbildung \mathcal{F} linear.

Aus den Bedingungen i. und ii. folgen bestimmte Eigenschaften die jede lineare Abbildung besitzt. So wird z.B. der Nullvektor immer auf den Nullvektor abgebildet. Wenn ausserdem V, W endlichdimensionale Vektorräume sind, z.B. $V = \mathbb{R}^n, W = \mathbb{R}^m$, dann kann jede lineare Abbildung $\mathcal{F} : V \rightarrow W$ durch eine Matrix $A \in \mathbb{R}^{m \times n}$ dargestellt werden. Allgemein kann man eine lineare Abbildung in der Form

$$\mathcal{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad x \mapsto Ax$$

schreiben, wobei $A \in \mathbb{R}^{m \times n}$ die Abbildungsmatrix von \mathcal{F} ist. Als Beispiel betrachten wir die Abbildung \mathcal{F} gegeben durch

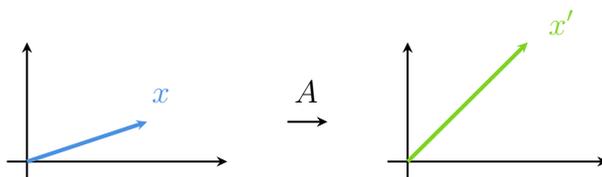
$$\mathcal{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x + y \\ x - 2y \\ 3x \end{pmatrix}.$$

Wir können die Abbildung in die Form $x \mapsto Ax$ bringen, indem wir die Abbildungsmatrix A bestimmen. In diesem Fall ist

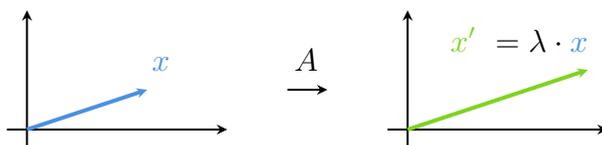
$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x + y \\ x - 2y \\ 3x \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & -2 \\ 3 & 0 \end{pmatrix}}_{:=A} \begin{pmatrix} x \\ y \end{pmatrix}.$$

2 Eigenwertproblem

Betrachten wir zunächst eine lineare Abbildung gegeben durch $x \mapsto Ax$ mit $A \in \mathbb{C}^{n \times n}$. Wie wir bereits in der Linearen Algebra I sahen, kann diese Abbildung analog zu einer Funktion interpretiert werden. Die Abbildung nimmt einen Vektor x und bildet ihn auf den Vektor $x' = Ax$ ab.



Wir fragen uns nun, ob es bestimmte Eingabevektoren x gibt, welche durch die Abbildung $x \mapsto Ax$ nur um einen Faktor λ gestreckt bzw. gestaucht werden.



Wir wissen also, dass in diesem Fall für eine Abbildung, gegeben durch $x \mapsto Ax$, folgendes gelten muss: $Ax = x' = \lambda \cdot x$. Daraus können wir die allgemeine Eigenwertgleichung ableiten:

$$Ax = \lambda x. \quad (1)$$

Der Vektor x welcher durch die Abbildung nur gestreckt bzw. gestaucht wird, nennt sich Eigenvektor. Der Faktor λ , um welchen der Vektor gestreckt bzw. gestaucht wird, nennen wir Eigenwert.

2.1 Eigenwerte

Wir suchen nun ein λ , sodass $Ax = \lambda x$ gilt. (Wir suchen nur nichttriviale Lösungen $x \neq 0$). Durch Umstellen von (1) erhalten wir ein homogenes lineares Gleichungssystem der Form

$$(A - \lambda I)x = 0. \quad (2)$$

Wann hat dieses HLGS nur nichttriviale Lösungen?

Recall von Lineare Algebra I

Folgende Aussagen sind für $A^{n \times n}$ äquivalent:

- Das homogene LGS $Ax = 0$ besitzt nur die triviale Lösung
- $\det(A) \neq 0$

Demnach hat das HLGS nichttriviale Lösungen genau dann wenn

$$\det(A - \lambda I) = 0$$

Mit dieser Gleichung können wir nun λ bestimmen.

Beispiel Sei $A = \begin{pmatrix} 3 & 0 & 0 \\ 5 & -6 & 0 \\ 0 & 1 & 3 \end{pmatrix}$, bestimme alle Eigenwerte λ_i von A .

$$\det(A - \lambda I) = \det \left[\begin{pmatrix} 3 & 0 & 0 \\ 5 & -6 & 0 \\ 0 & 1 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right] = \det \begin{pmatrix} 3 - \lambda & 0 & 0 \\ 5 & -6 - \lambda & 0 \\ 0 & 1 & 3 - \lambda \end{pmatrix} = 0$$

$$(3 - \lambda)(-6 - \lambda)(3 - \lambda) = 0$$

$$\lambda_1 = 3, \lambda_2 = 3, \lambda_3 = -6$$

Das Polynom $p_A := \det(A - \lambda I)$ heisst charakteristisches Polynom. Wenn $A \in \mathbb{C}^{n \times n}$, dann ist $p_A(\lambda)$ ein Polynom vom Grad n . In unserem Beispiel ist der Eigenwert 3 eine doppelte Nullstelle des charakteristischen Polynoms. Wir sagen dann, dass die algebraische Vielfachheit des Eigenwertes 3 gleich 2 ist. Für die Matrix von oben bedeutet das

$\lambda_{1,2} = 3$ hat die algebraische Vielfachheit 2

$\lambda_3 = -6$ hat die algebraische Vielfachheit 1

Merkmale und Definitionen von Eigenvektoren einer Matrix $A \in \mathbb{C}^{n \times n}$:

- A hat mindestens einen Eigenwert $\lambda \in \mathbb{C}$
- A hat höchstens n Eigenwerte $\lambda \in \mathbb{C}$
- A hat genau n Eigenwerte $\lambda \in \mathbb{C}$, wenn man die Vielfachheit zählt
- die Menge aller Eigenwerte von A heisst Spektrum von A
- Zwei Matrizen $A, B \in \mathbb{C}^{n \times n}$ heissen ähnlich, wenn für eine reguläre Matrix $T \in \mathbb{C}^{n \times n}$ gilt: $B = T^{-1}AT$. Ferner haben A und B dann:
 - dasselbe charakteristische Polynom
 - dieselben Eigenwerte mit denselben algebraischen Vielfachheiten
 - dasselbe Spektrum
 - dieselbe Determinante

2.2 Eigenvektoren

Wir wissen nun wie wir die Eigenwerte λ erhalten, nun müssen wir noch die dazugehörigen Eigenvektoren berechnen. Das heisst den Vektor x bestimmen, für welchen gilt:

$$Ax = \lambda x \quad (x \neq 0)$$

Das Problem kann wieder zu einem homogenen LGS der Form (??) umgeschrieben werden. Nun setzen wir einen der Eigenwerte ein und lösen das HLGS. Demnach sind die Eigenvektoren immer mit einem Eigenwert verknüpft.

Beispiel Sei $A = \begin{pmatrix} 3 & 0 & 0 \\ 5 & -6 & 0 \\ 0 & 1 & 3 \end{pmatrix}$ mit den Eigenwerten $\lambda_{1,2} = 3, \lambda_3 = -6$. Bestimme die Eigenvektoren $v_{\lambda_1}, v_{\lambda_2}, v_{\lambda_3}$ von A .

Für den ersten Eigenvektor müssen wir folgendes HLGS lösen:

$$v_{\lambda_1} : (A - \lambda_1 I)x = (A - 3I)x = 0.$$

Mit der gegebenen Matrix ergibt sich

$$\begin{pmatrix} 0 & 0 & 0 \\ 5 & -9 & 0 \\ 0 & 1 & 0 \end{pmatrix} x = 0.$$

Welches die Lösung

$$x_3 = t \in \mathbb{R}; \quad x_2 = 0; \quad x_1 = 0$$

besitzt und somit in den folgenden Eigenvektor resultiert:

$$v_{\lambda_1} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Da die Eigenwerte $\lambda_1 = \lambda_2$ sind, ist der Eigenvektor v_{λ_2} identisch zu v_{λ_1} . Für den dritten Eigenvektor müssen wir folgendes HLGS lösen:

$$v_{\lambda_3} : (A - \lambda_3 I)x = (A + 6I)x = 0.$$

Mit der gegebenen Matrix ergibt sich

$$\begin{pmatrix} 9 & 0 & 0 \\ 5 & 0 & 0 \\ 0 & 1 & 9 \end{pmatrix} x = 0.$$

Welches die Lösung

$$x_1 = 0; \quad x_2 = -9x_3$$

besitzt und somit in den folgenden Eigenvektor resultiert:

$$v_{\lambda_3} = \begin{pmatrix} 0 \\ -9 \\ 1 \end{pmatrix}.$$

Da wir bei der Berechnung für Eigenvektoren ein HLGS mit $\det = 0$ lösen, wird es immer unendlich viele Lösungen für $(A - \lambda I)x = 0$ geben. D.h. jede mögliche Linearkombination von Eigenvektoren von einem Eigenwert ist auch wieder ein Eigenvektor zum gleichen Eigenwert. Der dadurch aufgespannte Vektorraum ist dann ein Unterraum von \mathbb{C}^n und nennt sich Eigenraum. Eigenräume werden mit E_{λ_i} bezeichnet.

Beispiel Sei $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ mit den Eigenwerten $\lambda_{1,2} = 1, \lambda_3 = 4$. Bestimme die Eigenvektoren von A .

Genau wie oben müssen wir ein HLGS lösen.

$$E_{\lambda_1} = E_{\lambda_2} = E_1 : (A - \lambda_1 I)x = (A - 1I)x = 0.$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} x = 0 \xrightarrow{\text{Gauss}} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x = 0.$$

Nun haben wir zwei freie Parameter:

$$x_3 = t; \quad x_2 = s; \quad x_1 = -s - t \quad s, t \in \mathbb{R}.$$

Der resultierende Eigenraum zum Eigenwert 1 ist somit

$$E_1 = \text{span} \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

E_4 kann auf die gleiche Weise berechnet werden.

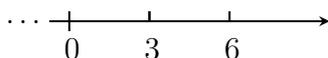
In diesem Beispiel ist also der Eigenraum E_1 zum Eigenwert 1 ein zweidimensionaler Unterraum. Die Dimension des Eigenraums $\dim(E_{\lambda_i})$ heisst geometrische Vielfachheit des Eigenwertes λ_i . In unserem Beispiel ist also die geometrische Vielfachheit von $\lambda_1 = 2$. Die geometrische Vielfachheit ist immer gleich der Anzahl freier Parameter des HLGS $(A - \lambda I)x = 0$. Allgemein ist immer zu beachten, dass für einen Eigenwert λ der Matrix $A \in \mathbb{C}^{n \times n}$ gelten muss:

$$1 \leq \text{geom. Vielfachheit}(\lambda) \leq \text{alg. Vielfachheit}(\lambda) \leq n.$$

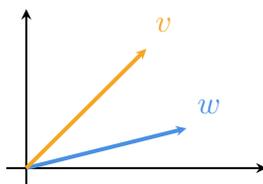
3 Normen und Skalarprodukte

3.1 Normen

Bisher haben wir mit den Vektorräumen die Vektoroperationen verallgemeinert. Wir wissen also wie wir, unabhängig von den Objekten in einem Vektorraum rechnen können. Nun wollen wir einen weiteren Schritt machen und die Grösse von Vektoren aus einem Vektorraum vergleichen. Dafür ordnen wir jedem Vektor v , in einem Vektorraum, eine reelle positive Zahl zu, da wir deren Grösse gut vergleichen können, z.B. können wir ganz klar sagen, dass die Zahl 6 grösser als die Zahl 3 ist.



Wie sieht es jedoch mit Vektoren aus? Hier ist die Antwort nicht so offensichtlich. Welcher dieser Vektoren ist grösser?



$$v = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \text{ oder } w = \begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$

Nun kommt es darauf an was grösser in diesem Kontext bedeutet z.B. könnte man argumentieren, dass:

- v ist grösser, da die geometrische Länge grösser ist.
- w ist grösser, da der Vektor eine grössere Zahl enthält.

Normen werden uns nun helfen diese Frage eindeutig zu beantworten. Denn eine Norm ist eine fixe Regel, mit welcher man jedem Vektor eine reelle positive Zahl zuordnen kann. Dadurch kann man Sie dann vergleichen. Mathematisch lässt sich das durch eine Abbildung ausdrücken:

$$\|\cdot\| : V \rightarrow \mathbb{R}, \quad v \mapsto \|v\|.$$

Wie im Beispiel von oben, gibt es viele Möglichkeiten diese Zuordnung zu machen.

- Geometrische Länge (Euklidische Norm)

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2} \quad \rightarrow \quad \|v\|_2 = \sqrt{18}$$
$$\|w\|_2 = \sqrt{17}$$

- grösster Eintrag (Maximumsnorm)

$$\|x\|_\infty = \max\{|x_1|, |x_2|\} \quad \rightarrow \quad \begin{aligned} \|v\|_\infty &= 3 \\ \|w\|_\infty &= 4 \end{aligned}$$

Achtung! Nicht alles ist eine Norm, es müssen bestimmte Bedingungen erfüllt sein damit eine Norm vorhanden ist.

Vektornormen

Eine Norm ordnet jedem Vektor eine reelle Zahl zu und kann so als eine Art Mass verstanden werden. $\forall v, w \in V, \alpha \in \mathbb{R}$ muss gelten:

$$\|v\| \geq 0 \text{ und } \|v\| = 0 \Leftrightarrow v = 0$$

$$\|\alpha \cdot v\| = |\alpha| \cdot \|v\|$$

$$\|v + w\| \leq \|v\| + \|w\|$$

Beispiele solcher Normen sind:

$$\text{Auf } \mathbb{R}^n \left\{ \begin{array}{ll} \|v\|_2 := \sqrt{v_1^2 + \dots + v_n^2} & \text{(Euklidische Norm)} \\ \|v\|_\infty := \max\{|v_1|, \dots, |v_n|\} & \text{(Maximumsnorm)} \\ \|v\|_p := (|v_1|^p + \dots + |v_n|^p)^{\frac{1}{p}} & \text{(p-Norm, } 1 \leq p < \infty) \end{array} \right.$$

$$\text{Auf } \mathbb{R}^{n \times m} \left\{ \begin{array}{ll} \|A\|_2 := \sqrt{\sum_{i,j} |a_{ij}|^2} & \text{(Hilbert-Schmidt-Norm)} \\ \|A\|_{SM} := \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}| & \text{(Spaltenmaximumsnorm)} \\ \|A\|_{ZM} := \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}| & \text{(Zeilenmaximumsnorm)} \end{array} \right.$$

3.2 Skalarprodukte

Wir haben nun einen Weg um die Grösse von Vektoren in einem Vektorraum zu vergleichen. Im nächsten Schritt wollen wir die Beziehung zwischen zwei Vektoren mit einer reellen Zahl beschreiben. Wieder lässt sich dies durch eine Abbildung beschreiben.

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}, \quad (x, y) \mapsto \langle x, y \rangle.$$

Auch hier müssen bestimmte Bedingungen erfüllt sein.

Skalarprodukte

Ein Skalarprodukt ordnet jedem Vektorpaar eine reelle Zahl zu, diese beschreibt die Beziehung der beiden Vektoren zueinander. $\forall x, y, z \in V, \alpha \in \mathbb{R}$ muss gelten:

$$\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$$

$$\langle x, \alpha \cdot y \rangle = \alpha \cdot \langle x, y \rangle$$

$$\langle x, y \rangle = \langle y, x \rangle$$

$$\langle x, x \rangle \geq 0 \text{ und } \langle x, x \rangle = 0 \Leftrightarrow x = 0$$

Wenn $\langle x, y \rangle = 0$, dann sind x, y orthogonal zueinander ($x \perp y$).

3.3 Vom Skalarprodukt induzierte Norm

Das Skalarprodukt kann auch benutzt werden, um eine Norm zu induzieren. Wenn auf einem Vektorraum ein Skalarprodukt gegeben ist, kann daraus direkt eine Norm der Form

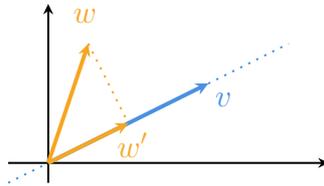
$$\|x\| := \sqrt{\langle x, x \rangle}$$

abgeleitet werden. Diese Norm erfüllt automatisch alle Bedingungen einer Norm, egal welches Skalarprodukt benutzt wird.

4 Orthonormalbasis

4.1 Orthogonalprojektion

Um die Orthonormalbasis besser zu verstehen, schauen wir uns nochmals die Orthogonalprojektion an. Grundlegend gibt uns die Orthogonalprojektion den Anteil eines Vektors der in dieselbe Richtung eines anderen Vektors zeigt. Oder bildlich:



Hier projizieren wir den Vektor w orthogonal auf den Vektor v . Das Resultat ist der Vektor w' . Mathematisch lässt sich das mit dem Skalarprodukt ausrechnen

$$w' = \frac{\langle w, v \rangle}{\langle v, v \rangle} v.$$

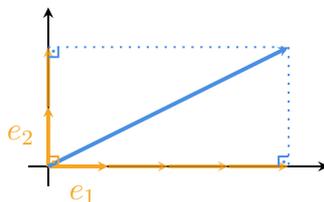
4.2 Orthonormalbasis

Wenn wir eine Basis für einen Vektorraum wählen, ist es vorteilhaft, wenn diese bestimmte Eigenschaften erfüllt. Eine dieser wünschenswerten Eigenschaften ist das die Basis orthonormal ist. Aber was bedeutet es, wenn eine Basis orthonormal ist, und warum ist das wünschenswert?

Bei einer Orthonormalbasis stehen die Basisvektoren senkrecht aufeinander (orthogonal) und haben die Länge 1 (normal). Zum Beispiel ist die Standardbasis

$$\mathcal{E} = \left\{ e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\},$$

eine Orthonormalbasis. Wenn wir in einer solchen Basis die Koordinaten von einem Vektor v finden wollen, können wir einfach den Vektor orthogonal auf die Basisvektoren projizieren. Nehmen wir Beispielsweise den blauen Vektor v in der folgenden Abbildung:



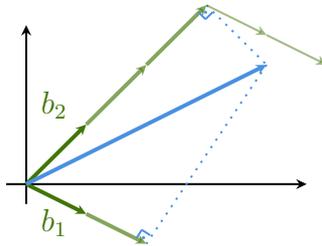
Durch die Orthogonalprojektionen auf die Basisvektoren erhalten wir die Koordinaten von v in der Basis \mathcal{E} . In diesem Fall sind die Koordinaten gegeben durch

$$v = \left[\begin{pmatrix} 4 \\ 2 \end{pmatrix} \right]_{\mathcal{E}}.$$

Bei einer nicht orthonormalen Basis müssen die Basisvektoren weder orthogonal noch normiert sein. Zum Beispiel

$$\mathcal{B} = \left\{ b_1 = \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix} \text{ und } b_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}.$$

Versuchen wir nun die Koordinaten desselben Vektors v in dieser Basis \mathcal{B} zu finden. Wieder projizieren wir den Vektor v orthogonal auf die Basisvektoren.



Durch die Orthogonalprojektion erhalten wir die Koordinaten $[2 \ 3]^T$. Diese entsprechen jedoch nicht dem Vektor v .

$$v \neq \begin{bmatrix} 2 \\ 3 \end{bmatrix}_{\mathcal{B}}.$$

Wenn wir einen Vektor in einer solchen nicht-orthogonalen Basis darstellen wollen, dann können wir nicht einfach orthogonal projizieren. Um die entsprechenden Koordinaten zu finden, müssen wir eine Linearkombination der Basisvektoren finden (LGS lösen).

Für unsere Orthonormalbasis bedeutet das nun folgendes. Sei $\mathcal{E} = \{e_1, e_2, \dots, e_n\}$ eine Orthonormalbasis des Vektorraums V , dann gilt $\forall x \in V$

$$x = \sum_{k=1}^n \langle x, e_k \rangle e_k.$$

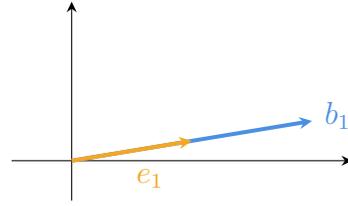
In anderen Worten, können wir jeden Vektor als Linearkombination der Basisvektoren schreiben, wobei die Koeffizienten durch die Orthogonalprojektionen gegeben sind. Um die Koordinaten zu finden, müssen wir kein LGS lösen. Beachte das der Term $\langle e_k, e_k \rangle$ wegfällt, da dieser aufgrund der Normierung immer 1 ist.

4.3 Gram-Schmidt-Orthogonalisierungsverfahren

Aufgrund der oben genannten Eigenschaften wollen wir oft mit Orthonormalbasen arbeiten. Um eine solche Basis zu finden, benutzen wir das Gram Schmidt Orthogonalisierungsverfahren. Dieses Verfahren erlaubt uns aus einer jeder Basis eine Orthonormalbasis zu erstellen. Dafür benötigen wir eine beliebige Basis $\mathcal{B} = \{b_1, \dots, b_k\}$ und ein Skalarprodukt, mit welchen wir dann eine Orthonormalbasis $\mathcal{E} = \{e_1, \dots, e_k\}$ erstellen. Im Folgenden gehen wir das Verfahren Schritt für Schritt durch.

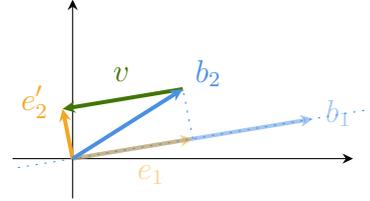
1. Wähle einen beliebigen Basisvektor b_1 und normiere ihn.

$$e_1 = \frac{b_1}{\|b_1\|} = \frac{b_1}{\sqrt{\langle b_1, b_1 \rangle}}$$



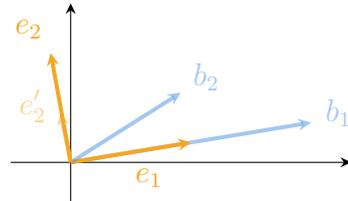
2. Wähle einen zweiten Basisvektor b_2 , ziehe zuerst den zu e_1 parallelen Teil ab

$$e'_2 = b_2 - \underbrace{\langle b_2, e_1 \rangle}_{v} e_1$$



und normiere ihn dann

$$e_2 = \frac{e'_2}{\|e'_2\|} = \frac{e'_2}{\sqrt{\langle e'_2, e'_2 \rangle}}$$



3. Wiederhole für jeden Basisvektor b_i

$$e'_i = b_i - \langle b_i, e_1 \rangle e_1 - \langle b_i, e_2 \rangle e_2 - \dots - \langle b_i, e_{i-1} \rangle e_{i-1}$$

$$e_i = \frac{e'_i}{\|e'_i\|} = \frac{e'_i}{\sqrt{\langle e'_i, e'_i \rangle}}$$

Beispiel Sei auf \mathcal{P}_4 das folgende Skalarprodukt gegeben:

$$\langle p, q \rangle = \int_0^1 p(x)q(x) dx,$$

finde eine Orthonormalbasis für den Vektorraum $\text{span}\{1, 3x^4\}$.

Wir benutzen das Gram-Schmidt-Verfahren mit den Basisvektoren $b_1 = 1$ und $b_2 = 3x^4$.

1.

$$e_1 = \frac{b_1}{\|b_1\|} = \frac{b_1}{\sqrt{\langle b_1, b_1 \rangle}}$$

$$\rightarrow \langle b_1, b_1 \rangle = \int_0^1 1 \cdot 1 dx = [x]_0^1 = 1$$

$$e_1 = 1$$

2.

$$e'_2 = b_2 - \langle b_2, e_1 \rangle e_1$$

$$\rightarrow \langle b_2, e_1 \rangle = \int_0^1 3x^4 \cdot 1 \, dx = \left[\frac{3}{5}x^5 \right]_0^1 = \frac{3}{5}$$

$$e'_2 = b_2 - \frac{3}{5} = 3x^4 - \frac{3}{5}$$

$$e_2 = \frac{e'_2}{\|e'_2\|} = \frac{e'_2}{\sqrt{\langle e'_2, e'_2 \rangle}}$$

$$\rightarrow \langle e'_2, e'_2 \rangle = \int_0^1 \left(3x^4 - \frac{3}{5} \right)^2 dx = \int_0^1 9x^8 - \frac{18}{5}x^4 + \frac{9}{25} dx$$

$$= \left[x^9 - \frac{18}{25}x^5 + \frac{9}{25}x \right]_0^1 = \frac{16}{25}$$

$$\rightarrow \sqrt{\langle e'_2, e'_2 \rangle} = \frac{4}{5}$$

$$e_2 = \frac{5}{4} \left(3x^4 - \frac{3}{5} \right) = \frac{15}{4}x^4 - \frac{3}{4}$$

Die Orthonormalbasis ist gegeben durch

$$\mathcal{E} = \left\{ 1, \frac{15}{4}x^4 - \frac{3}{4} \right\}.$$

Zusammengefasst kann das Gram-Schmidt-Verfahren in einem Kochrezept zusammengefasst werden.

Gram-Schmidt-Verfahren

Für das Gram-Schmidt-Verfahren benötigt man eine beliebige Basis $\mathcal{B} = \{b_1, \dots, b_k\}$, sowie ein beliebiges Skalarprodukt. Daraus lässt sich dann eine Orthonormalbasis $\mathcal{E} = \{e_1, \dots, e_k\}$ erstellen.

1. Wähle einen beliebigen Basisvektor b_1 und normiere ihn mit der vom Skalarprodukt induzierten Norm.

$$e_1 = \frac{b_1}{\|b_1\|} = \frac{b_1}{\sqrt{\langle b_1, b_1 \rangle}}$$

2. Wähle einen zweiten Basisvektor b_2 , ziehe den zu b_1 parallelen Teil ab und normiere ihn.

$$e'_2 = b_2 - \langle b_2, e_1 \rangle e_1$$

$$e_2 = \frac{e'_2}{\|e'_2\|} = \frac{e'_2}{\sqrt{\langle e'_2, e'_2 \rangle}}$$

3. Wiederhole für jeden Basisvektor b_i

$$e'_i = b_i - \langle b_i, e_1 \rangle e_1 - \langle b_i, e_2 \rangle e_2 - \dots - \langle b_i, e_{i-1} \rangle e_{i-1}$$

$$e_i = \frac{e'_i}{\|e'_i\|} = \frac{e'_i}{\sqrt{\langle e'_i, e'_i \rangle}}$$

5 Bild und Kern

5.1 Bild

Ganz am Anfang der Linearen Algebra I sahen wir, wie die Matrix-Vektor-Multiplikation ausgeführt wird. Wir können diese Multiplikation, einer Matrix A mit einem Vektor x , auch als Linearkombination der Spaltenvektoren von A ausdrücken

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} x_2 + \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} x_3.$$

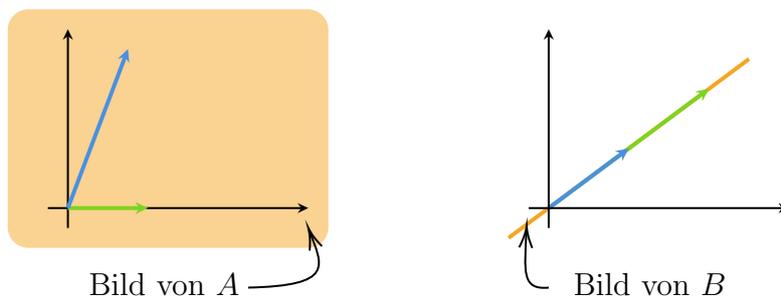
Das heisst, dass alle möglichen Ergebnisse einer Matrix-Vektor-Multiplikation Ax durch die Linearkombination der Spaltenvektoren von A dargestellt werden können. Die Menge aller dieser Ergebnisse nennt sich Bild. Mathematisch ist das Bild einer Matrix $A \in \mathbb{R}^{m \times n}$ durch

$$\text{Bild}(A) = \{y \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n \text{ sodass } y = Ax\}$$

definiert. Grafisch können wir uns das Bild gut im zweidimensionalen Raum vorstellen. Seien beispielsweise die Matrizen

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

gegeben. Wir können nun grafisch die Bilder von A und B darstellen.



Das Bild von A spannt den gesamten zweidimensionalen Raum auf, da die Spaltenvektoren linear unabhängig sind (zeigen in unterschiedliche Richtungen). Das Bild von B , im Gegensatz, ist nur eine Linie da die Spaltenvektoren linear abhängig sind (zeigen in dieselbe Richtung). Dies stimmt auch mit der obigen Definition überein, denn

$$Ax = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 0 \end{pmatrix},$$

linear unabhängig

$$Bx = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$

linear abhängig

Wir können hier erkennen, dass die Dimension des Bildes gleich der Anzahl linear unabhängiger Spalten bzw. dem Rang der Matrix ist. Allgemein gilt:

$$\dim(\text{Bild}(A)) = \text{Rang}(A).$$

Beispiel Oft muss ein Erzeugendensystem oder Basis des Bildes einer Matrix gefunden werden. Sei beispielsweise eine Matrix A gegeben durch:

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 5 \\ 4 & 3 & 9 \end{pmatrix}.$$

- **Aufgabe:** Finden Sie ein Erzeugendensystem für das Bild von A .
Hier können wir einfach die Spaltenvektoren ablesen und erhalten

$$\text{Bild}(A) = \left\{ \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 9 \end{pmatrix} \right\}.$$

- **Aufgabe:** Finden Sie eine Basis für das Bild von A .

Wir suchen nun alle linear unabhängigen Spalten von A . Dafür transponieren wir die originale Matrix A , wodurch die Spalten zu den Zeilen werden und umgekehrt. Nun wenden wir den Gauss-Algorithmus an und nehmen die Zeilen, die nicht Null sind, als Basis des Bildes. Wir erhalten:

$$A^T = \begin{pmatrix} 1 & 3 & 4 \\ 2 & 1 & 3 \\ 4 & 5 & 9 \end{pmatrix} \xrightarrow{\text{Gauss}} \begin{pmatrix} 1 & 3 & 4 \\ 0 & -5 & -5 \\ 0 & 0 & 0 \end{pmatrix}.$$

Somit ist eine mögliche Basis gegeben durch

$$\text{Bild}(A) = \left\{ \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ -5 \\ -5 \end{pmatrix} \right\}.$$

(Falls man direkt erkennt wie viele linear unabhängige Spalten eine Matrix hat, kann man ohne Rechnung die entsprechende Anzahl an linear unabhängigen Spaltenvektoren nehmen. Bei grossen Matrizen kann es aber schwierig sein, lineare Abhängigkeit zwischen Spalten zu erkennen.)

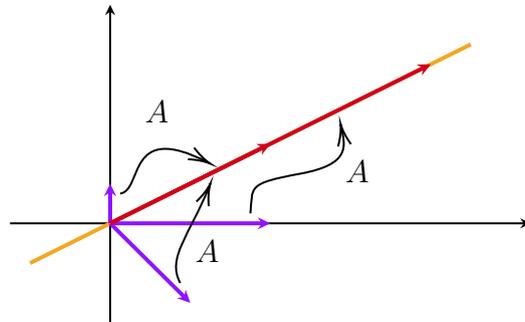
5.2 Kern

Um eine Intuition für den Kern zu bekommen, blicken wir nochmals auf das Bild einer Matrix. Das Bild beschreibt den Raum auf welchen alle beliebigen Vektoren x durch A abgebildet werden. Für die Matrix

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

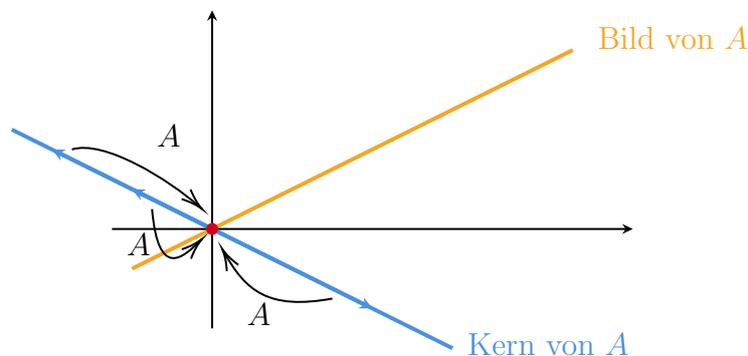
können wir uns das Bild grafisch als eine Linie vorstellen da A zwei linear abhängige Spalten hat. Betrachten wir nun was passiert, wenn einige Vektoren v durch A abgebildet werden, d.h. $Av = u$

$$\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} \quad \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$



Man erkennt, dass alle Vektoren nach der Multiplikation mit A wie erwartet im **Bild** liegen (in diesem fall auf der orangen Linie). Wir können uns nun fragen welche Vektoren x durch A auf null abgebildet werden. D.h. wir suchen all x , sodass $Ax = 0$. Zum Beispiel erfüllen diese Vektoren die Bedingung:

$$\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$



Wir erkennen, dass alle Vektoren auf der blauen Linie diese Bedingung erfüllen. Diese resultierende Menge, aller Vektoren, die auf null abgebildet werden, wird Kern genannt. Diese Menge ist äquivalent zur Lösung des HLGS $Ax = 0$. Mathematisch ausgedrückt ist der Kern einer Matrix A definiert durch

$$\text{Kern}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}.$$

Mit der Definition von Kern können wir auch die im ersten Abschnitt eingeführte Zusammenhangs-Liste ergänzen.

Wichtige Zusammenhänge

Folgende Aussagen sind für $A \in \mathbb{R}^{n \times n}$ äquivalent:

- $\text{rang}(A) = n$
- Das LGS $Ax = b$ ist für beliebiges b lösbar
- Das LGS $Ax = b$ hat genau eine Lösung
- Das homogene LGS $Ax = 0$ hat nur die triviale Lösung $x = 0$
- Die Zeilen und Spalten von A sind linear unabhängig
- A ist invertierbar (regulär, nicht singular)
- $\det(A) \neq 0$
- Die Spalten von A bilden eine Basis von \mathbb{R}^n
- **Der Kern besteht nur aus dem Nullvektor**
- **Kein Eigenwert von A ist Null**

Um den Zusammenhang mit Eigenwerten zu verstehen, erinnern wir uns nochmals was Eigenwerte und Eigenvektoren sind. Ein Eigenvektor ist ein Vektor, der durch eine Multiplikation mit einer Matrix nur um einen Faktor λ , dem Eigenwert, skaliert wird. Wenn nun ein Eigenwert Null ist, dann gibt es Vektoren (die nicht der Nullvektor sind) welche auf null abgebildet werden. D.h. das HLGS $Ax = 0$ hat nicht triviale Lösungen $x \neq 0$ und dadurch ist der Rang der Matrix nicht voll.

Beispiel

- Bestimme den Kern der Matrix

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

Löse das HLGS $Ax = 0$:

$$\begin{array}{c|c} 2 & 4 & 0 \\ 1 & 2 & 0 \end{array} \xrightarrow{\text{Gauss}} \begin{array}{c|c} 2 & 4 & 0 \\ 0 & 0 & 0 \end{array} \quad \begin{array}{l} x_2 = t \\ x_1 = -2t \end{array} \quad \rightarrow \quad \text{Kern}(A) = \text{span} \left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$$

5.3 Beziehung zwischen Kern und Bild

Im Folgenden wollen wir Zusammenhänge zwischen Kern und Bild genauer betrachten. Spezifischer wollen wir zwei wichtige Eigenschaften beweisen. Die erste Eigenschaft

$$i. \quad \dim(\text{Kern}) + \dim(\text{Bild}) = n,$$

lässt sich sowohl grafisch als auch mathematisch Beweisen.

Grafische Intuition

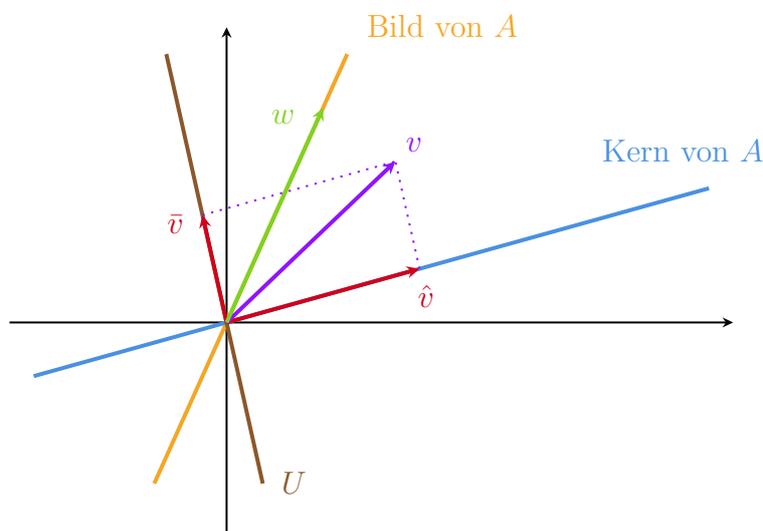
Beginnen wir mit dem grafischen Beweis in 2 Dimensionen. Sei $A \in \mathbb{R}^{m \times n}$ und $v \in \mathbb{R}^n$ ein beliebiger Vektor. Ausserdem sei $Av = w$. Wir können nun v in zwei Komponenten zerlegen:

$$v = \hat{v} + \bar{v} \quad \begin{cases} \hat{v} \in \text{Kern}(A) \\ \bar{v} \text{ orthogonal zu } \hat{v} \end{cases}$$

Dadurch ist:

$$Av = A(\hat{v} + \bar{v}) = \overset{0}{A\hat{v}} + A\bar{v} = A\bar{v} = w.$$

Der Term $A\hat{v}$ ist Null, da \hat{v} im Kern von A liegt. Letztlich definieren wir noch einen zum Kern orthogonalen Unterraum U . Per Definition ist dadurch $\bar{v} \in U$. Zusammengefasst können wir das ganze grafisch darstellen:



Folgende Beobachtungen können nun gemacht werden:

- Der **Kern** und der zum Kern orthogonale Unterraum U haben zusammen die Dimension n , da wir durch die Summe $v = \hat{v} + \bar{v}$ alle Vektoren v erzeugen können.

$$\dim(\text{Kern}) + \dim(U) = n.$$

- Jeder Vektor \bar{v} in U korrespondiert zu einem Vektor in **Bild** und umgekehrt, da $A\bar{v} = w$. Die Dimension von U ist also gleich der Dimension des **Bildes**.

$$\dim(U) = \dim(\text{Bild}).$$

Zusammen gilt also

$$\dim(\text{Kern}) + \dim(\text{Bild}) = n.$$

Mathematisch

Sei $A \in \mathbb{R}^{m \times n}$ und der Kern per Definition die Lösungsmenge des HLGS $Ax = 0$. Diese Lösungsmenge hat $n - r$ freie Parameter. Dabei ist r der Rang von A (Anzahl Pivots). Betrachten folgendes Beispiel:

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 5 \\ 4 & 3 & 9 \end{pmatrix} \xrightarrow{\text{HLGS}} \begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 3 & 1 & 5 & 0 \\ 4 & 3 & 9 & 0 \end{array} \xrightarrow{\text{Gauss}} \begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 0 & -5 & -7 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

wodurch:

$$x_3 = t, \quad x_2 = -\frac{7}{5}t, \quad x_1 = -\frac{6}{5}t;$$

und

$$r = 2, \quad n = 3, \quad n - r = 1.$$

Der Lösungsraum von $Ax = 0$ hat also die Dimension 1. Damit hat auch der Kern die Dimension 1.

$$\dim(\text{Kern}) = n - r.$$

Durch das Gauss-Verfahren haben wir bereits die Zeilenstufenform R von A gefunden. Die Bilder von A und R lassen sich durch die jeweiligen Spalten beschreiben

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 5 \\ 4 & 3 & 9 \end{pmatrix} \qquad R = \begin{pmatrix} 1 & 2 & 4 \\ 0 & -5 & -7 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\text{Bild}(A) = \text{span} \{a^{(1)}, a^{(2)}, a^{(3)}\} \stackrel{!}{=} \text{Bild}(R) = \text{span} \{r^{(1)}, r^{(2)}, r^{(3)}\}$$

Da A und R dieselbe Lösungsmenge beschreiben, spannen sie auch dasselbe Bild auf und haben auch dieselbe Dimension. Die Dimension von $\text{Bild}(R)$ ist einfach zu bestimmen da sie gleich dem Rang r ist. Es gilt also

$$\dim(\text{Bild}(A)) = \dim(\text{Bild}(R)) = r.$$

Dadurch ist

$$\dim(\text{Kern}) + \dim(\text{Bild}) = n - r + r = n,$$

und somit

$$\dim(\text{Kern}) + \dim(\text{Bild}) = n.$$

Die zweite Eigenschaft, die wir beweisen wollen ist, gegeben durch:

$$ii. \quad \text{Kern}(A^\top) \perp \text{Bild}(A).$$

Den Beweis dazu können wir durch folgende Überlegung erlangen. Sei A definiert durch:

$$A = \begin{pmatrix} \vdots & \vdots & \vdots \\ a^{(1)} & a^{(2)} & a^{(3)} \\ \vdots & \vdots & \vdots \end{pmatrix}$$

wobei $a^{(n)}$ die Spalten von A sind. Dadurch gilt das $\text{Bild}(A) = \{a^{(1)}, a^{(2)}, a^{(3)}\}$. Durch das Transponieren werden die Spalten $a^{(n)}$ zu den Zeilen $a^{[n]}$.

$$A^\top = \begin{pmatrix} \dots & a^{[1]} & \dots \\ \dots & a^{[2]} & \dots \\ \dots & a^{[3]} & \dots \end{pmatrix}$$

Betrachten wir nun das HLGS $A^\top y = 0$. Allgemein ist dies definiert durch

$$A^\top y = \begin{pmatrix} \dots & a^{[1]} & \dots \\ \dots & a^{[2]} & \dots \\ \dots & a^{[3]} & \dots \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wir sehen, dass der Vektor y immer im Kern von A^\top liegen muss, um diese Gleichung zu erfüllen. Nun können wir das Produkt von A^\top mit y auch als das Skalarprodukt von den Spalten $a^{(n)}$ mit dem Vektor y interpretieren. Dieses Skalarprodukt $\langle a^{(n)}, y \rangle$ ist immer Null, woraus folgt:

$$\text{Kern}(A^\top) \perp \text{Bild}(A).$$

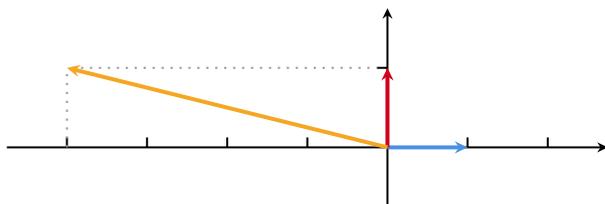
Aus dieser Bedingung folgt die Fredholm Alternative welche besagt, dass $Ax = b$ lösbar ist (b liegt im Bild) genau dann, wenn b senkrecht auf allen Lösungen des adjungierten LGS $A^\top y = 0$ steht.

6 Basiswechsel und Diagonalisierung

6.1 Basiswechsel für Vektoren

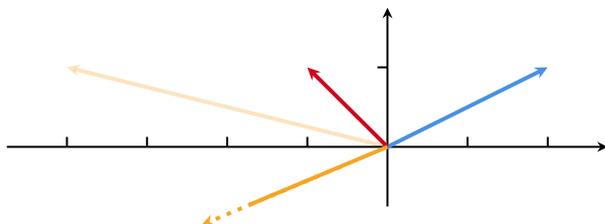
Was bedeutet es nochmal, wenn wir einen Vektor mit den Koordinaten $[-4 \ 1]^T$ beschreiben? Grundlegend beschreiben Koordinaten um wie viel die Basisvektoren des assoziierten Vektorraums skaliert werden. In der Standardbasis ist dies bereits recht intuitiv.

$$-4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$$



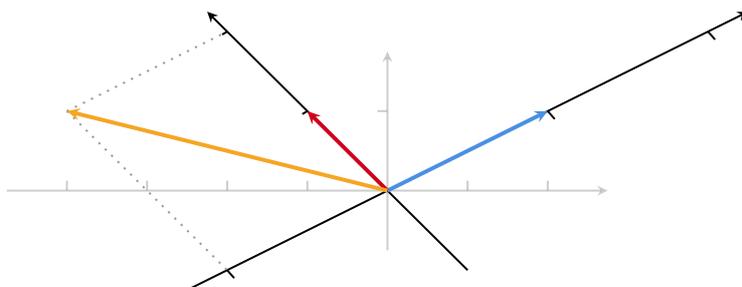
Wir können aber auch eine andere Basis wählen, z.B. gegeben durch die Vektoren $[2 \ 1]^T$ und $[-1 \ 1]^T$. Wenn wir nun dieselben Koordinaten wie in der Standardbasis benutzen, bekommen wir einen anderen Vektor.

$$-4 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -7 \\ -3 \end{bmatrix} \neq \begin{bmatrix} -4 \\ 1 \end{bmatrix}$$



Mit den neuen Basisvektoren wären die richtigen Koordinaten für denselben Vektor gegeben durch

$$-1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$$



In der neuen Basis wird derselbe Vektor durch die Koordinaten $[-1 \ 2]^T$ beschrieben. D.h. die Koordinaten hängen immer von den Basisvektoren ab. Wenn wir von einer Basis

in eine andere wechseln, müssen demnach auch die Koordinaten geändert werden. Dafür führen wir die Übergangsmatrix bzw. Transformationsmatrix ein. Um dieses Konzept besser verstehen zu können betrachten wir zunächst das Beispiel von oben. Seien die zwei Basen gegeben durch:

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \quad \mathcal{B}' = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}.$$

Wir wollen nun von der Basis \mathcal{B} in die Basis \mathcal{B}' transformieren. Gesucht sind also die Koordinaten eines Vektors $[v]_{\mathcal{B}}$ in der neuen Basis \mathcal{B}' . Mathematisch können wir das so ausdrücken:

$$x_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\mathcal{B}} + x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}_{\mathcal{B}} = \underbrace{\begin{bmatrix} -4 \\ 1 \end{bmatrix}_{\mathcal{B}}}_{\text{Beschreiben denselben Vektor}} \Leftrightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{\mathcal{B}'}$$

Dies können wir in ein LGS in Matrixschreibweise umformen:

$$\underbrace{\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}}_{:= T_{\mathcal{B}' \rightarrow \mathcal{B}}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{\mathcal{B}'} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}_{\mathcal{B}} \Leftrightarrow T_{\mathcal{B}' \rightarrow \mathcal{B}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{\mathcal{B}'} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}_{\mathcal{B}}.$$

Wir erkennen hier, dass wenn ein Vektor in der Basis \mathcal{B}' mit der Matrix $T_{\mathcal{B}' \rightarrow \mathcal{B}}$ multipliziert wird, derselbe Vektor in der Basis \mathcal{B} resultiert. Die Matrix $T_{\mathcal{B}' \rightarrow \mathcal{B}}$ transformiert also einen Vektor aus der Basis \mathcal{B}' in die Basis \mathcal{B} . Wir suchen jedoch eine Matrix die einen Vektor aus der Basis \mathcal{B} in die Basis \mathcal{B}' transformiert. Dafür können wir einfach die Inverse nehmen:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{\mathcal{B}'} = T_{\mathcal{B}' \rightarrow \mathcal{B}}^{-1} \begin{bmatrix} -4 \\ 1 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}_{\mathcal{B}'}$$

Die gewünschte Übergangsmatrix ist also gegeben durch:

$$T_{\mathcal{B}' \rightarrow \mathcal{B}}^{-1} = T_{\mathcal{B} \rightarrow \mathcal{B}'} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}$$

Basiswechsel für Vektoren

Für beliebige Basen $\mathcal{Q} = \{q_1, q_2, \dots, q_n\}$ und $\mathcal{W} = \{w_1, w_2, \dots, w_n\}$ gilt:

$$T_{\mathcal{Q} \rightarrow \mathcal{W}} = ([q_1]_{\mathcal{W}}, [q_2]_{\mathcal{W}}, \dots, [q_n]_{\mathcal{W}})$$

$$[v]_{\mathcal{W}} = T_{\mathcal{Q} \rightarrow \mathcal{W}} [v]_{\mathcal{Q}}$$

$$T_{\mathcal{Q} \rightarrow \mathcal{W}} = T_{\mathcal{W} \rightarrow \mathcal{Q}}^{-1}$$

Beispiel Eine Basis sei gegeben mit $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}_{\mathcal{S}}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{\mathcal{S}} \right\}$. Transformiere den Vektor $v = \begin{bmatrix} 3 \\ 4 \end{bmatrix}_{\mathcal{S}}$ aus der Standardbasis \mathcal{S} in die Basis \mathcal{B} .

Wir suchen also die Übergangsmatrix $T_{\mathcal{S} \rightarrow \mathcal{B}} = ([s_1]_{\mathcal{B}}, [s_2]_{\mathcal{B}})$. Die Vektoren $[s_i]_{\mathcal{B}}$ kennen wir nicht, nur die Vektoren $[b_i]_{\mathcal{S}}$ sind bekannt. Damit ist auch die Matrix $T_{\mathcal{B} \rightarrow \mathcal{S}}$ gegeben.

$$T_{\mathcal{B} \rightarrow \mathcal{S}} = ([b_1]_{\mathcal{S}}, [b_2]_{\mathcal{S}}) = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

Wir wissen nun das $T_{\mathcal{S} \rightarrow \mathcal{B}} = T_{\mathcal{B} \rightarrow \mathcal{S}}^{-1}$. Damit können wir den Vektor v transformieren:

$$[v]_{\mathcal{B}} = T_{\mathcal{S} \rightarrow \mathcal{B}} [v]_{\mathcal{S}} = T_{\mathcal{B} \rightarrow \mathcal{S}}^{-1} [v]_{\mathcal{S}} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}.$$

6.2 Basiswechsel für Matrizen

Wir können nun Vektoren von einer Basis in eine andere transformieren. Nun stellt sich die Frage, ob das auch für Matrizen gilt. Denn lineare Abbildungen und deren Darstellungsmatrizen sind auch an eine Basis gebunden. Sei A eine Darstellungsmatrix einer linearen Abbildung in der Basis \mathcal{Q} . Dann gilt:

$$[A]_{\mathcal{Q}} [x]_{\mathcal{Q}} = [y]_{\mathcal{Q}}.$$

Wir suchen nun die Darstellungsmatrix der gleichen Abbildung in einer Basis \mathcal{W} . Mathematisch ausgedrückt

$$[A]_{\mathcal{W}} [x]_{\mathcal{W}} = [y]_{\mathcal{W}}.$$

Durch Umformen erhalten wir:

$$\begin{aligned} [A]_{\mathcal{Q}} [x]_{\mathcal{Q}} &= [y]_{\mathcal{Q}} && | \cdot T_{\mathcal{Q} \rightarrow \mathcal{W}} \\ T_{\mathcal{Q} \rightarrow \mathcal{W}} [A]_{\mathcal{Q}} [x]_{\mathcal{Q}} &= T_{\mathcal{Q} \rightarrow \mathcal{W}} [y]_{\mathcal{Q}} && | T_{\mathcal{Q} \rightarrow \mathcal{W}} [y]_{\mathcal{Q}} = [y]_{\mathcal{W}} \\ T_{\mathcal{Q} \rightarrow \mathcal{W}} [A]_{\mathcal{Q}} [x]_{\mathcal{Q}} &= [y]_{\mathcal{W}} && | I = T_{\mathcal{Q} \rightarrow \mathcal{W}}^{-1} T_{\mathcal{Q} \rightarrow \mathcal{W}} \\ T_{\mathcal{Q} \rightarrow \mathcal{W}} [A]_{\mathcal{Q}} T_{\mathcal{Q} \rightarrow \mathcal{W}}^{-1} T_{\mathcal{Q} \rightarrow \mathcal{W}} [x]_{\mathcal{Q}} &= [y]_{\mathcal{W}} && | T_{\mathcal{Q} \rightarrow \mathcal{W}} [x]_{\mathcal{Q}} = [x]_{\mathcal{W}} \\ \underbrace{T_{\mathcal{Q} \rightarrow \mathcal{W}} [A]_{\mathcal{Q}} T_{\mathcal{Q} \rightarrow \mathcal{W}}^{-1}}_{[A]_{\mathcal{W}}} [x]_{\mathcal{W}} &= [y]_{\mathcal{W}} \end{aligned}$$

Basiswechsel für Matrizen

Für beliebige Basen $\mathcal{Q} = \{q_1, q_2, \dots, q_n\}$ und $\mathcal{W} = \{w_1, w_2, \dots, w_n\}$ gilt:

$$[A]_{\mathcal{W}} = T_{\mathcal{Q} \rightarrow \mathcal{W}} [A]_{\mathcal{Q}} T_{\mathcal{Q} \rightarrow \mathcal{W}}^{-1}$$

Doch warum müssen wir hier zweimal mit einer Transformationsmatrix multiplizieren? Betrachten wir dafür die einzelnen Terme etwas genauer.

$$[A]_{\mathcal{W}}[x]_{\mathcal{W}} = [y]_{\mathcal{W}}$$

$$T_{\mathcal{Q} \rightarrow \mathcal{W}} [A]_{\mathcal{Q}} T_{\mathcal{Q} \rightarrow \mathcal{W}}^{-1} [x]_{\mathcal{W}} = [y]_{\mathcal{W}}$$

$$\left| \begin{array}{l} \left| \begin{array}{l} \left| \begin{array}{l} \text{Transformation von } [x]_{\mathcal{W}} \text{ zu Basis } \mathcal{Q} \\ \text{Abbildung von } [x]_{\mathcal{Q}} \text{ in Basis } \mathcal{Q} \\ \text{Rücktransformation von } [y]_{\mathcal{Q}} \text{ zu Basis } \mathcal{W} \end{array} \right. \end{array} \right. \end{array} \right.$$

Beispiel Eine Basis sei gegeben mit $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}_{\mathcal{S}}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{\mathcal{S}} \right\}$. Transformiere die Matrix $[A]_{\mathcal{S}} = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$ aus der Standardbasis \mathcal{S} in die Basis \mathcal{B} .

Wir benutzen hier die oben eingeführte Transformationsmatrix in diesem Fall bekommen wir

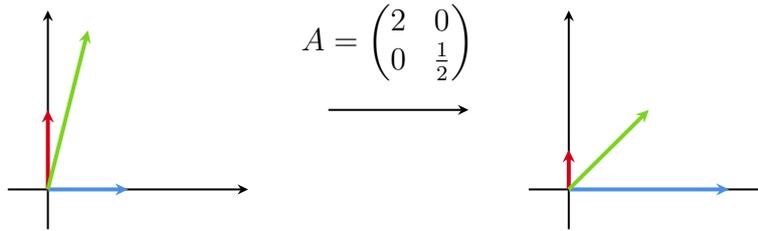
$$\begin{aligned} [A]_{\mathcal{B}} &= T_{\mathcal{S} \rightarrow \mathcal{B}} [A]_{\mathcal{S}} T_{\mathcal{S} \rightarrow \mathcal{B}}^{-1} \\ &= \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 5 & 1 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ -6 & 0 \end{pmatrix} \end{aligned}$$

6.3 Diagonalisieren

Wir wissen nun wie wir von einer Basis zur anderen wechseln können, indem wir Vektoren und Matrizen transformieren. Da lineare Abbildungen durch Matrizen beschrieben werden können, sind wir auch in der Lage lineare Abbildungen von einer Basis in die andere zu transformieren. Für solche Abbildungen ist es vorteilhaft, wenn die dazugehörige Abbildungsmatrix diagonal ist. Denn Diagonalmatrizen sind leicht invertierbar und Matrixmultiplikation ist bedeutend unkomplizierter. Beim sogenannten Diagonalisieren wollen wir einen Basiswechsel durchführen, sodass eine gewünschte Matrix in der neuen Basis diagonal ist. Konkreter wollen wir für eine quadratische Matrix $A \in \mathbb{C}^{n \times n}$ eine Matrix $T \in \mathbb{C}^{n \times n}$ finden, sodass

$$T^{-1} A T = D = \text{diag}(d_1, d_2, \dots, d_n),$$

wobei D die Matrix in der neuen Basis ist. Um die Diagonalisierung durchzuführen, müssen wir demnach die Basis finden in der A diagonal ist. Aber wann ist das der Fall? Nehmen wir nochmals einen Schritt zurück und überlegen uns was der Effekt von diagonalen Matrizen und deren assoziierter Abbildung im zweidimensionalen Raum ist. Im folgenden Beispiel betrachten wir wie eine Diagonalmatrix verschiedene Vektoren im Raum abbildet.



$$\begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix}$$

Wir erkennen, dass die Basisvektoren (hier die Standardbasis in Blau und Rot) nur skaliert werden, aber nicht Ihre Richtung ändern, genau wie Eigenvektoren. In anderen Worten sind bei Diagonalmatrizen die Eigenvektoren auch die Basisvektoren. Indem wir also die Eigenvektoren der ursprünglichen Matrix als neue Basisvektoren wählen, garantieren wir, dass die Matrix in der neuen Basis diagonal ist. Weiterhin sehen wir, dass die Diagonaleinträge der Matrix genau die Faktoren sind mit denen die Basisvektoren skaliert werden. D.h. die Diagonaleinträge der diagonalisierten Matrix sind die Eigenwerte der ursprünglichen Matrix.

Sei zum Beispiel eine Matrix A in der Standardbasis \mathcal{S} gegeben durch $A = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}$. Weiter seien die Eigenwerte und Eigenvektoren von A gegeben durch

$$\lambda_1 = 2, \quad v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \lambda_2 = 1, \quad v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Um A zu diagonalisieren suchen wir eine Basis \mathcal{B} , in der die Basisvektoren durch die Eigenvektoren v_1, v_2 gegeben sind. Anschliessend benötigen wir nur noch eine Transformationsmatrix, um in die entsprechende Basis zu wechseln. Mit den gegebenen Eigenvektoren können wir bereits die Übergangsmatrix $T_{\mathcal{B} \rightarrow \mathcal{S}}$ bestimmen, indem wir die Eigenvektoren als Spalten nehmen:

$$T_{\mathcal{B} \rightarrow \mathcal{S}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Die diagonalisierte Matrix D ist dann gegeben durch

$$D = T_{\mathcal{B} \rightarrow \mathcal{S}}^{-1} A T_{\mathcal{B} \rightarrow \mathcal{S}} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Die Eigenwerte der Matrix A sind nun auch die Diagonaleinträge der Matrix D . Gleichermassen gilt auch

$$A = T_{\mathcal{B} \rightarrow \mathcal{S}} D T_{\mathcal{B} \rightarrow \mathcal{S}}^{-1}.$$

Zusammengefasst können wir dies in einem ‘‘Kochrezept’’ aufschreiben.

Diagonalisieren

Um eine Matrix $A \in \mathbb{C}^{n \times n}$ zu diagonalisieren ($D = T^{-1}AT$), gehe wie folgt vor:

1. Bestimme die Eigenwerte λ_i und Eigenvektoren v_i von A .
2. Die Matrix $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ ist die diagonalisierte Matrix.
3. Setze die Eigenvektoren als Spalten in die Matrix T ein (**Gleiche Reihenfolge wie bei D!**).
4. Bestimme T^{-1} . Falls Eigenvektoren orthonormal sind, gilt: $T^{-1} = T^\top$.

6.4 Diagonalisierbarkeit

Nicht alle Matrizen sind Diagonalisierbar. Es kann also sein das keine Transformationsmatrix T existiert, sodass

$$D = T^{-1}AT.$$

Ein Kriterium für die Diagonalisierbarkeit einer Matrix A ist dadurch, dass die Transformationsmatrix T regulär (invertierbar) sein muss. Ein Weg dies zu Überprüfen ist die lineare Abhängigkeit der Spalten/ Zeilen zu betrachten. Nur wenn alle Spalten linear unabhängig sind, ist die Matrix auch invertierbar. In diesem Fall sind die Spalten der Matrix T auch die Eigenvektoren von A . Das bedeutet, dass wir prüfen müssen, ob die Eigenvektoren von A linear unabhängig sind. Dieses Kriterium können wir umschreiben denn, wenn alle Eigenvektoren linear unabhängig sind dann bilden sie eine Basis. Eine solche Matrix nennen wir dann halbeinfach.

Halbeinfache Matrizen

Eine Matrix $A \in \mathbb{C}^{n \times n}$ ist halbeinfach, wenn Ihre Eigenvektoren eine Basis für \mathbb{C}^n bilden.

Ein Kriterium zur Diagonalisierbarkeit einer Matrix ist dadurch, das sie halbeinfach ist. Wir können prüfen, ob eine Matrix halbeinfach ist, indem wir uns die Vielfachheiten der Eigenwerte und Eigenvektoren anschauen. Denn wir wissen, dass Eigenvektoren von unterschiedlichen Eigenwerten jeweils linear unabhängig sind und die Eigenvektoren desselben Eigenwerts auch linear unabhängig sein müssen. Das bedeutet dass, nur wenn die algebraische Vielfachheit jedes Eigenwertes gleich der geometrischen Vielfachheit ist, ist die Matrix halbeinfach. D.h. für jeden Eigenwert λ muss gelten:

$$\text{alg. Vfh. } \lambda = \text{geom. Vfh. } \lambda.$$

Somit können wir das Kriterium für die Diagonalisierbarkeit einer Matrix zusammenfassen.

Halbeinfache Matrizen

Eine quadratische Matrix $A \in \mathbb{C}^{n \times n}$ ist diagonalisierbar, falls eine reguläre Matrix T existiert, sodass $D = T^{-1}AT$ eine Diagonalmatrix ist.

$$A \text{ halbeinfach} \Leftrightarrow A \text{ diagonalisierbar}$$

Beispiel

Halbeinfache Matrix: $A = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}$. Mit den Eigenwerten und Eigenvektoren

$$\lambda_1 = 2, \lambda_2 = 1, \quad v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

In diesem Fall sind alle geometrischen und algebraischen Vielfachheiten gleich 1 damit bilden die Eigenvektoren eine Basis. Die Matrix ist also halbeinfach und damit auch diagonalisierbar.

Nicht halbeinfache Matrix: $A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$. Mit den Eigenwerten und Eigenvektoren

$$\lambda_{1,2} = 2, \quad v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Hier ist die algebraische Vielfachheit des Eigenwertes 2 gleich 2, aber die geometrische Vielfachheit nur gleich 1 damit bilden die Eigenvektoren keine Basis. Demnach ist die Matrix nicht halbeinfach und damit auch nicht diagonalisierbar.

6.4.1 Spezialfall: Symmetrische Matrizen

Wenn $A \in \mathbb{R}^{n \times n}$ symmetrisch ist, gilt:

- A ist halbeinfach, also auch diagonalisierbar.
- A besitzt eine orthonormale Eigenbasis.
- Es existiert eine orthogonale Matrix T , sodass $T^{-1}AT = T^T AT$ diagonal ist.

6.4.2 Anwendung: Potenzen von Matrizen

Sei $A \in \mathbb{C}^{n \times n}$ diagonalisierbar. Berechne A^3 .

Da A diagonalisierbar ist, gilt $A = TDT^{-1}$. Damit gilt für A^3 :

$$A^3 = (TDT^{-1})^3 = \underbrace{(TDT^{-1})}_I \underbrace{(TDT^{-1})}_I (TDT^{-1}) = TD^3 T^{-1}.$$

und weil D eine Diagonalmatrix ist, gilt

$$A^3 = T \underbrace{\text{diag}(\lambda_1^3, \lambda_2^3, \dots, \lambda_n^3)}_{D^3} T^{-1}.$$

Allgemein gilt für jede diagonalisierbare Matrix A

$$A^k = T \operatorname{diag}(\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k) T^{-1}.$$

Insbesondere vereinfacht sich dadurch auch das Matrixexponential

$$e^{At} = T \operatorname{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_n t}) T^{-1}.$$
$$e^A = \sum_{m=0}^{\infty} \frac{A^m}{m!}$$

7 Quadratische Formen

7.0.1 Notation. In den folgenden zwei Kapiteln zu quadratischen Formen und Quadriken werden wir für Vektoren immer **fett gedruckte** Zeichen verwenden. Zum Beispiel ist der Koordinatenvektor $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top$. Analog zu anderen Vektoren $\mathbf{y}, \mathbf{z}, \mathbf{a}$ etc.

7.1 Quadratische Formen

Quadratische Gleichungen sind bereits aus der Schulzeit bekannt. In der Regel haben solche Gleichungen in einer Variabel die allgemeine Form:

$$f(x) = ax^2 + bx + c.$$

Wie bereits aus der Analysis bekannt, können Gleichungen auch mehr als eine Variable umfassen. Wir werden oft alle Variablen in einem Vektor $\mathbf{x} = (x_1, \dots, x_n)$ zusammenfassen. Nun können sich auch quadratische Funktionen in mehreren Variablen ergeben. Ein Beispiel hierfür ist eine quadratische Funktion in zwei Variablen:

$$q(x_1, x_2) = x_1^2 + 4x_1x_2 + x_2^2.$$

Bei genauerer Betrachtung sieht man, dass diese Funktion auch durch eine Matrix-Vektor Multiplikation beschrieben werden kann.

$$q(x_1, x_2) = (x_1 \ x_2) \underbrace{\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}}_A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{x}^\top A \mathbf{x}.$$

In diesem Fall sagen wir das $q(\mathbf{x}) = x_1^2 + 4x_1x_2 + x_2^2$ die quadratische Form der Matrix A ist. Allgemein können wir für jede symmetrische Matrix $A \in \mathbb{R}^{n \times n}$, die dazugehörige quadratische Form finden. Sie ist allgemein, wie folgt definiert:

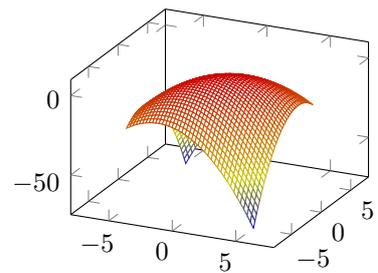
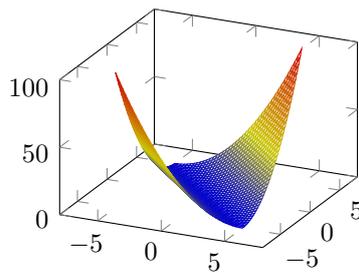
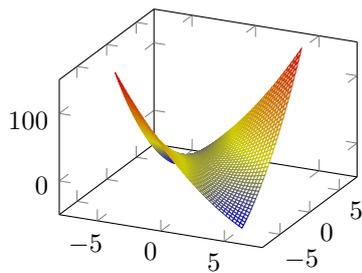
$$\begin{aligned} q_A : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \mathbf{x} &\longmapsto \mathbf{x}^\top A \mathbf{x} = \sum_{i,j=1}^n a_{ij} x_i x_j. \end{aligned}$$

Für 2 und 3 Dimensionen können wir eine allgemeine Form explizit aufschreiben.

$$\mathbb{R}^2 : q_A(\mathbf{x}) = \mathbf{x}^\top \begin{pmatrix} a & b \\ b & c \end{pmatrix} \mathbf{x} = ax_1^2 + 2bx_1x_2 + cx_2^2,$$

$$\mathbb{R}^3 : q_A(\mathbf{x}) = \mathbf{x}^\top \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \mathbf{x} = ax_1^2 + 2bx_1x_2 + 2cx_1x_3 + dx_2^2 + 2ex_2x_3 + fx_3^2.$$

Je nach Matrix kann die quadratische Form grafisch anders aussehen. Beispiele für 2 Variablen sind zum Beispiel.



7.2 Definitheit

Die unterschiedlichen quadratischen Formen und die assoziierten symmetrischen Matrizen lassen sich wie folgt nach Definitheit klassifizieren.

Definitheit quadratischer Formen

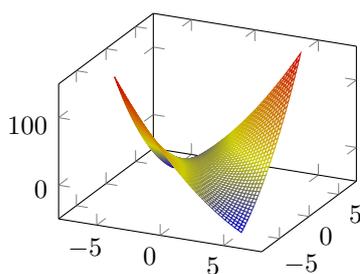
Eine quadratische Form, gegeben durch $q(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x}$, heisst:

- positiv definit: $q(\mathbf{x}) > 0, \forall \mathbf{x} \neq 0$
- negativ definit: $q(\mathbf{x}) < 0, \forall \mathbf{x} \neq 0$
- positiv semidefinit: $q(\mathbf{x}) \geq 0, \forall \mathbf{x} \neq 0$
- negativ semidefinit: $q(\mathbf{x}) \leq 0, \forall \mathbf{x} \neq 0$
- indefinit: sonst

Die Definitheit kann auch über die Eigenwerte von A bestimmt werden. Somit ist die Matrix A

- positiv definit: Alle $\lambda > 0$
- negativ definit: Alle $\lambda < 0$
- positiv semidefinit: Alle $\lambda \geq 0$
- negativ semidefinit: Alle $\lambda \leq 0$
- indefinit: sonst

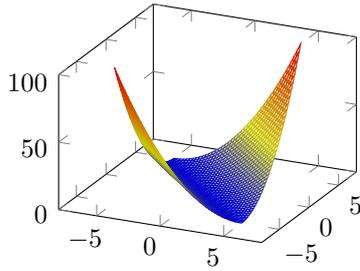
Betrachten wir nochmals die oben gezeigten Beispiele und klassifizieren diese anhand ihrer Definitheit.



$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \rightarrow x_1^2 + 4x_1x_2 + x_2^2$$

$$\text{Eigenwerte: } \lambda_1 = 3, \quad \lambda_2 = -1$$

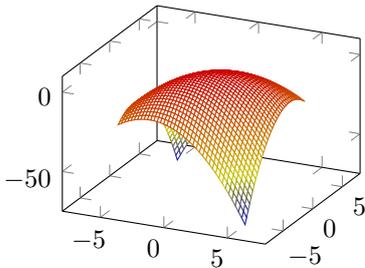
Indefinit



$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \rightarrow x_1^2 + 2x_1x_2 + x_2^2$$

$$\text{Eigenwerte: } \lambda_1 = 2, \quad \lambda_2 = 0$$

Positiv semidefinit



$$A = \begin{pmatrix} -1 & \frac{1}{2} \\ \frac{1}{2} & -1 \end{pmatrix} \rightarrow -x_1^2 + x_1x_2 - x_2^2$$

$$\text{Eigenwerte: } \lambda_1 = -\frac{3}{2}, \quad \lambda_2 = -\frac{1}{2}$$

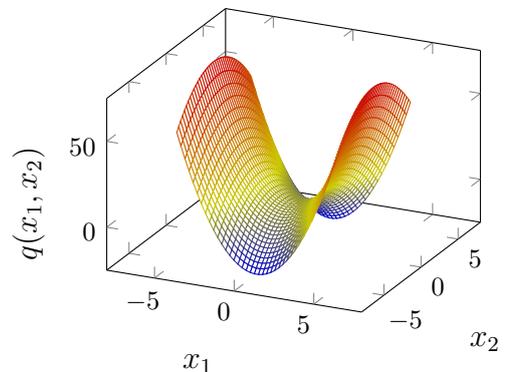
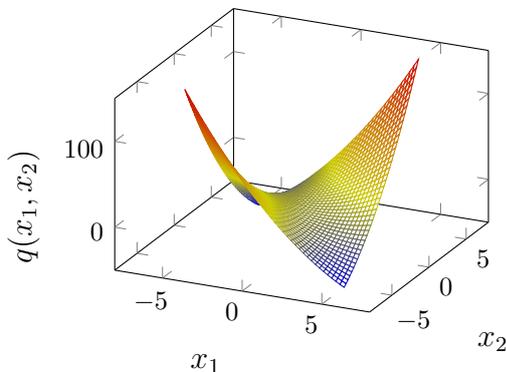
Negativ definit

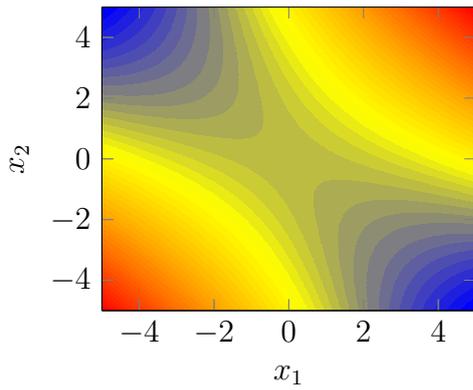
7.3 Rein quadratische Formen

Eine rein quadratische Form hat keine Mischterme. D.h. die Matrix A hat nur Diagonaleinträge. Allgemein können wir rein quadratische Formen in dieser Form schreiben

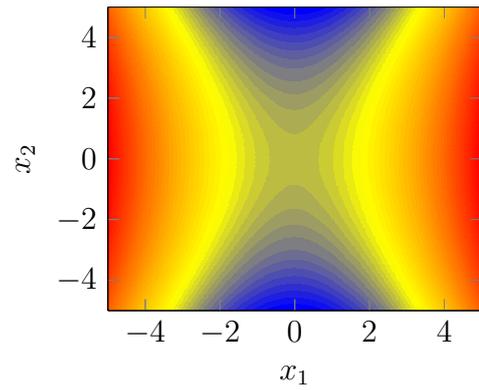
$$q_A(\mathbf{x}) = (x_1 \quad x_2 \quad \dots \quad x_n) \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_1x_1^2 + a_2x_2^2 + \dots + a_nx_n^2.$$

Wie unterscheiden sich rein quadratische Formen von allgemeinen quadratischen Formen? Betrachten wir zwei Beispiele. In einem Fall gibt es Mischterme welche durch Elemente neben der Diagonalen beschrieben werden. Im anderen Fall ist die Matrix diagonal und es gibt keine Mischterme.





$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$



$$A = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$$

Es scheint hier als ob die beiden Formen dieselbe Fläche beschreiben mit dem einzigen Unterschied, dass sie rotiert sind. In folgenden Kapiteln werden wir sehen wie wir durch eine Basistransformation eine nicht reine Form in eine rein quadratische Form umwandeln können. Dafür werden wir das Konzept der Diagonalisierung verwenden.

8 Kegelschnitte und Quadriken

8.1 Kegelschnitte

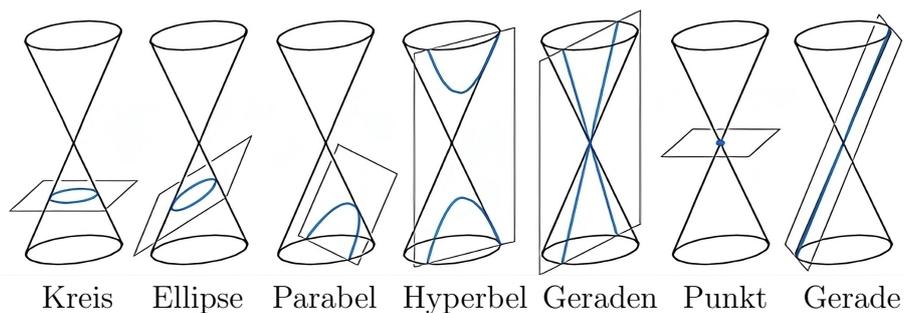
Wie wir im vorherigen Kapitel gesehen haben, können wir die quadratische Form einer symmetrischen Matrix $A \in \mathbb{R}^{2 \times 2}$ wie folgt schreiben

$$q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^2.$$

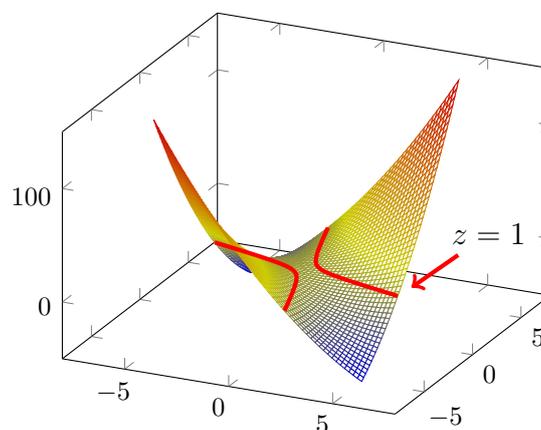
Für eine solche quadratische Form können wir nun eine Niveaumenge definieren. Eine solche Niveaumenge kann man sich wie einen Schnitt durch unsere quadratische Funktion vorstellen. Mathematisch definieren wir eine solche Niveaumenge mit

$$\{x \in \mathbb{R}^2 : q_A(\mathbf{x}) = 1\}.$$

Die resultierenden Niveaumengen einer quadratischen Funktion, mit zwei Variablen, ergeben dann Kegelschnitte. Abhängig von der Matrix A erhalten wir einen der folgenden Kegelschnitte.



Zum Beispiel gibt die quadratische Form $q(\mathbf{x}) = x_1^2 + 4xy + y^2$, aus dem vorherigen Kapitel, einen hyperbolischen Kegelschnitt, wenn wir die quadratische Form bei $z = 1$ anschauen. Dies kann einfach in einer Abbildung gesehen werden, ist aber schwer direkt aus der quadratischen Form abzulesen.



Ein besonderer Fall ergibt sich, wenn die Matrix A diagonal ist, dann sind die Hauptachsen des Kegelschnitts parallel zu den Koordinatenachsen. Ein Kegelschnitt im Hauptachsensystem kann in Normalform geschrieben werden. Die Normalformen aller möglichen Kegelschnitte sind in der folgenden Liste gegeben.

Rang(A) = 2:

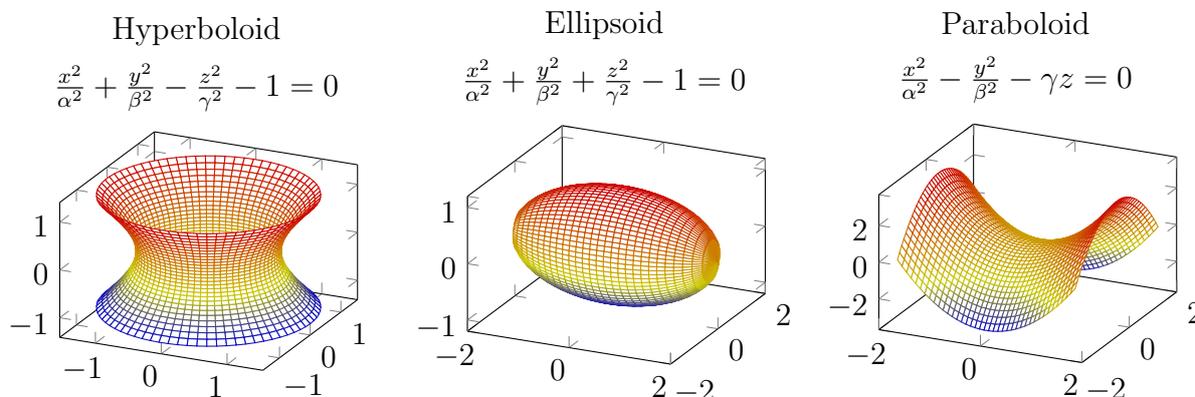
$$\begin{aligned} \frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - 1 = 0 & \quad \text{Ellipse} \\ \frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} - 1 = 0 & \quad \text{Hyperbel} \\ \frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + 1 = 0 & \quad \text{leere Menge} \\ x^2 + \beta^2 y^2 = 0 & \quad \text{Punkt} \\ x^2 - \beta^2 y^2 = 0 & \quad \text{Geradenpaar} \end{aligned}$$

Rang(A) = 1:

$$\begin{aligned} x^2 - \gamma y = 0 & \quad \text{Parabel} \\ x^2 - \alpha^2 = 0 & \quad \text{paralleles Geradenpaar} \\ x^2 + \alpha^2 = 0 & \quad \text{leere Menge} \\ x^2 = 0 & \quad \text{Gerade} \end{aligned}$$

8.2 Quadriken

Im allgemeinen Fall, wenn $A \in \mathbb{R}^{n \times n}$ und $n > 2$ ist, kann die Lösungsmenge der quadratischen Form nicht mehr durch Kegelschnitte beschrieben werden. Im Allgemeinen nennen wir dann die Lösungsmenge einer quadratischen Form eine Quadrik. Für den Fall $n = 3$ können wir die resultierenden Quadriken noch grafisch als Flächen 2. Grades darstellen. Auch hier können wir die Normalformen der Quadriken in einem Hauptachsensystem finden. Einige Beispiele sind:



Aber wie genau bringen wir quadratische Formen in die entsprechende Normalform?

8.3 Hauptachsentransformation und Normalformen

Bei einer Hauptachsentransformation wird eine quadratische Form in ihr Hauptachsensystem gebracht, wobei A dann diagonal ist. Zusätzlich, können wir Kegelschnitte bzw. Quadriken in Normalform schreiben. Wie das genau funktioniert, schauen wir uns in einem Beispiel an.

Aufgabe aus der Prüfung Winter 2018.

- Gegeben sei die quadratische Form

$$q : \mathbb{R}^2 \rightarrow \mathbb{R}, \mapsto q(\mathbf{x}) = -x_1^2 + 4x_1x_2 - x_2^2.$$

Ein Kegelschnitt Q ist gegeben durch

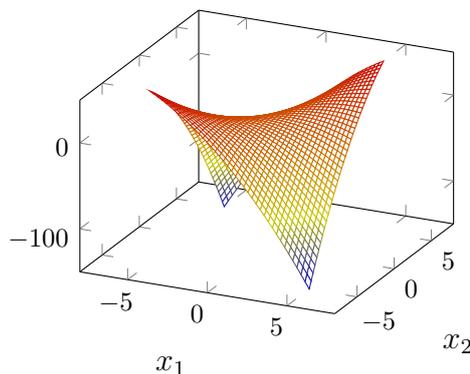
$$q(\mathbf{x}) + \mathbf{a}^\top \mathbf{x} = 0, \text{ wobei } \mathbf{a}^\top = (6 \quad -6).$$

Bringen Sie den Kegelschnitt durch eine Hauptachsentransformation $x = Ty$ und eine Translation auf Normalform, und geben Sie dabei auch T explizit an.

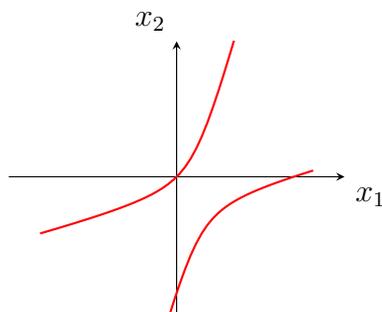
Betrachten wir zunächst einmal die gegebene quadratische Form $q(\mathbf{x})$. Diese lässt sich im 3D Raum darstellen und kann durch eine symmetrische Matrix $A \in \mathbb{R}^{2 \times 2}$ beschreiben werden.

$$A = \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix}$$

$$q(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x}$$



Der Kegelschnitt Q beschreibt die Schnittmenge von $q(\mathbf{x})$ mit einer Ebene. In 2 Dimensionen lässt sich die Schnittmenge wie folgt darstellen.



$$Q(\mathbf{x}) : -x_1^2 + 4x_1x_2 - x_2^2 + 6x_1 - 6x_2 = 0$$

Zuerst rotieren wir den Kegelschnitt damit die Hauptachsen mit den Koordinatenachsen übereinstimmen. Dazu führen wir einen Basiswechsel $\mathbf{x} = T\mathbf{y}$ durch. In der neuen Basis

muss die Matrix A eine Diagonalmatrix sein. Demnach muss A diagonalisiert werden. Die dafür benötigten Eigenwerte und Eigenvektoren sind mit

$$\lambda_1 = -3, \quad v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \lambda_2 = 1, \quad v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

gegeben. Da wir den Kegelschnitt nur rotieren wollen, ohne seine Form zu verändern, müssen wir unsere Transformationsmatrix T orthogonal wählen. Dadurch ergibt sich

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Nun können wir die Koordinatentransformation durchführen und den Kegelschnitt in der neuen Basis ausdrücken. Per Definition ist $q(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x}$ wodurch wir Q umschreiben können.

$$Q : q(\mathbf{x}) + \mathbf{a}^\top \mathbf{x} = \mathbf{x}^\top A \mathbf{x} + \mathbf{a}^\top \mathbf{x} = 0$$

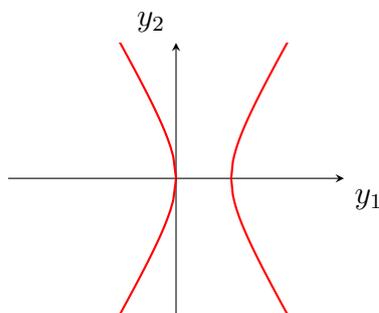
Nun setzen wir $\mathbf{x} = T\mathbf{y}$ ein und erhalten

$$\mathbf{x}^\top A \mathbf{x} + \mathbf{a}^\top \mathbf{x} = \underbrace{(T\mathbf{y})^\top A (T\mathbf{y})}_{\mathbf{y}^\top T^\top A T \mathbf{y}} + \mathbf{a}^\top T\mathbf{y} = 0$$

Da wir die Matrix A diagonalisieren und T orthogonal ist, gilt $T^\top A T = D$. Der Kegelschnitt vereinfacht sich weiterhin zu

$$\mathbf{y}^\top D \mathbf{y} + \mathbf{a}^\top T\mathbf{y} = -3y_1^2 + y_2^2 + 6\sqrt{2}y_1 = 0.$$

Was grafisch, wie erwartet, eine Rotation des Kegelschnitts darstellt.



$$Q(\mathbf{y}) : -3y_1^2 + y_2^2 + 6\sqrt{2}y_1 = 0$$

Um die Normalform zu erhalten, müssen wir nur noch, durch eine Translation, den y_1 -Term loswerden. Dafür ergänzen wir quadratisch und erhalten

$$\begin{aligned} 0 &= -3y_1^2 + y_2^2 + 6\sqrt{2}y_1 \\ &= -3 \left(y_1^2 - 2\sqrt{2}y_1 \right) + y_2^2 \\ &= -3 \left((y_1 - \sqrt{2})^2 + 2 \right) + y_2^2 \\ &= -3 \left(y_1 - \sqrt{2} \right)^2 + y_2^2 + 6. \end{aligned}$$

Nun wird eine Translation $\mathbf{z} = \mathbf{y} + \mathbf{c}$ durchgeführt, wobei der Vektor \mathbf{c} die Verschiebung in y_1 Richtung korrigiert. In diesem Fall gilt

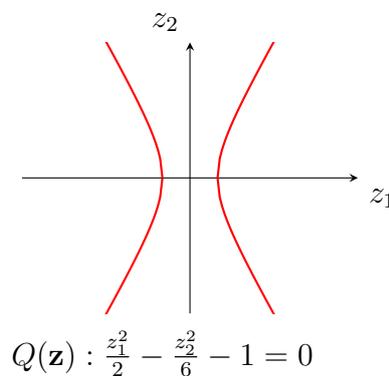
$$\mathbf{z} = \mathbf{y} + \mathbf{c} = \mathbf{y} + \begin{pmatrix} -\sqrt{2} \\ 0 \end{pmatrix}.$$

Durch einsetzen erhalten wir die Normalform

$$-3z_1^2 + z_2^2 + 6 = 0$$

$$\frac{z_1^2}{2} - \frac{z_2^2}{6} = 1.$$

In der Normalform sehen wir nun eindeutig, dass es sich bei dem Kegelschnitt um eine Hyperbel handelt. Grafisch sieht die Normalform wie folgt aus.



Wenn wir nun alle Formen vergleichen, ist zu erkennen, dass die Normalform am einfachsten zu interpretieren ist.

$$Q : -x_1^2 + 4x_1x_2 - x_2^2 + 6x_1 - 6x_2 = 0$$

$$-3y_1^2 + y_2^2 + 6\sqrt{2}y_1 = 0$$

$$\frac{z_1^2}{2} - \frac{z_2^2}{6} - 1 = 0.$$

Das allgemeine Vorgehen können wir in einem Kochrezept zusammenfassen.

Quadriken in Normalform bringen

Durch zwei Koordinatentransformationen (Drehung $\mathbf{y} = T\mathbf{x}$ und Verschiebung $\mathbf{z} = \mathbf{y} + \mathbf{c}$) können wir jede Quadrik $Q : \mathbf{x}^\top A \mathbf{x} + \mathbf{a}^\top \mathbf{x} + b = 1$ in eine Normalform bringen. Falls nur nach der Hauptachstransformation der quadratischen Form gefragt ist, kann die Verschiebung weglassen werden.

1. Bestimme die Matrix $A \in \mathbb{R}^{n \times n}$ sodass $q(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x}$.
2. Diagonalisiere A und wähle T orthogonal.
3. Setze $\mathbf{x} = T\mathbf{y}$ ein und vereinfache so weit wie möglich. (Hauptachsentransformation ist hier fertig)
4. Bestimme den Vektor $\mathbf{a} \in \mathbb{R}^n$ und die Konstante $b \in \mathbb{R}$ und multipliziere alles aus um $\mathbf{y}^\top D \mathbf{y} + \mathbf{a}^\top T\mathbf{y} + b = 0$ zu erhalten.
5. Ergänze quadratisch, um die Verschiebung \mathbf{c} zu bestimmen.
6. Setze $\mathbf{z} = \mathbf{y} + \mathbf{c}$ ein und vereinfache so weit wie möglich.

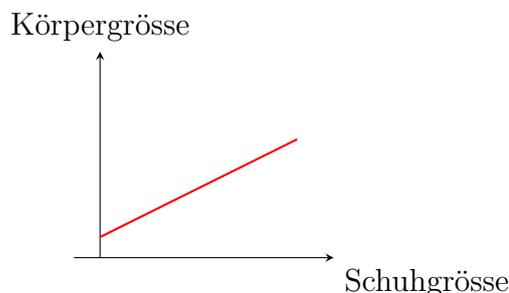
9 Methode der Kleinsten Quadrate und QR-Zerlegung

9.1 Methode der kleinsten Quadrate

Die Methode der kleinsten Quadrate wird immer dann angewendet, wenn wir eine möglichst *gute* Lösung für ein überbestimmtes Gleichungssystem $Ax = c$ finden wollen, für welches es keine eindeutige Lösung gibt. Ein Gleichungssystem ist überbestimmt, wenn es mehr Gleichungen als Unbekannte hat. Für die Matrix $A \in \mathbb{R}^{m \times n}$ gilt dann $m > n$. In realen Anwendungen sind Gleichungssysteme sehr oft Überbestimmt, weswegen es wichtig ist für solche Systeme möglichst *gute* Lösungen zu finden. Was dabei eine *gute* Lösung ausmacht, betrachten wir im folgenden Beispiel.

Bei einer Umfrage werden Menschen bezüglich Körpergröße und Schuhgröße befragt. Wir wollen einen Zusammenhang zwischen Körpergröße und Schuhgröße finden damit wir nachher basierend auf der Schuhgröße eine Schätzung für die Körpergröße machen können. Wir suchen also eine Gleichung der Form

$$\text{Körpergröße} = x_1 \cdot \text{Schuhgröße} + x_2, \quad (\ddagger)$$



wobei die Koeffizienten x_1, x_2 unbekannt sind und experimentell durch eine Umfrage ermittelt werden. In einer Umfrage werden nun 8 Menschen befragt, wodurch wir für zwei Unbekannte x_1 und x_2 , 8 Gleichungen erhalten. Wir werden also keine eindeutige Lösung für x_1 und x_2 finden können. Wie wählen wir nun die besten Werte für x_1 und x_2 damit die Gleichung (\ddagger) möglichst gut die Realität repräsentiert? Betrachten wir zunächst die Werte der Umfrage. Aus einer Tabelle mit allen Umfragewerten können wir das überstimmte LGS $Ax = c$ aufstellen.

Schuhgröße	Körpergröße
41	172
45	190
42	180
43	183
43	178
38	163
44	180
40	178

$$172 = x_1 \cdot 41 + x_2$$

$$190 = x_1 \cdot 45 + x_2$$

$$180 = x_1 \cdot 42 + x_2$$

$$183 = x_1 \cdot 43 + x_2$$

$$178 = x_1 \cdot 43 + x_2$$

$$163 = x_1 \cdot 38 + x_2$$

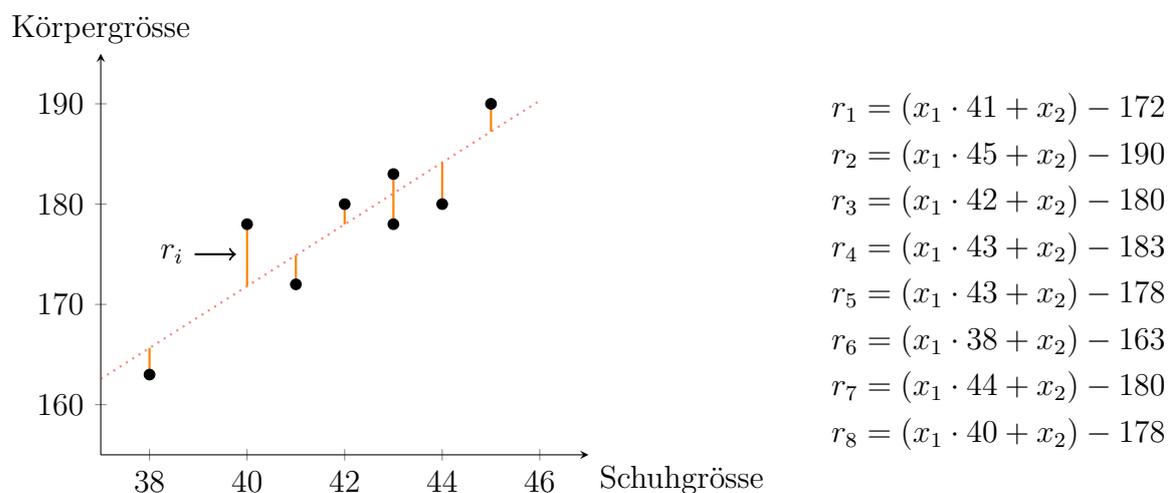
$$180 = x_1 \cdot 44 + x_2$$

$$178 = x_1 \cdot 40 + x_2$$

Die Matrix $A \in \mathbb{R}^{8 \times 2}$ und der Vektor $c \in \mathbb{R}^{8 \times 1}$ sind dann wie folgt definiert:

$$A = \begin{pmatrix} 41 & 1 \\ 45 & 1 \\ \vdots & \vdots \\ 40 & 1 \end{pmatrix}, \quad c = \begin{pmatrix} 172 \\ 190 \\ \vdots \\ 178 \end{pmatrix}.$$

Wir können nun alle Messwerte in einem Koordinatensystem auftragen um zu sehen, wie der Zusammenhang zwischen Schuhgröße und Körpergröße aussieht. Als Nächstes wollen wir eine Gerade der Form (§) finden, die möglichst nah an allen Messpunkten ist. Ein Weg dies zu erreichen ist es alle Abstände r_i von der Geraden zu den Messwerten zu minimieren.



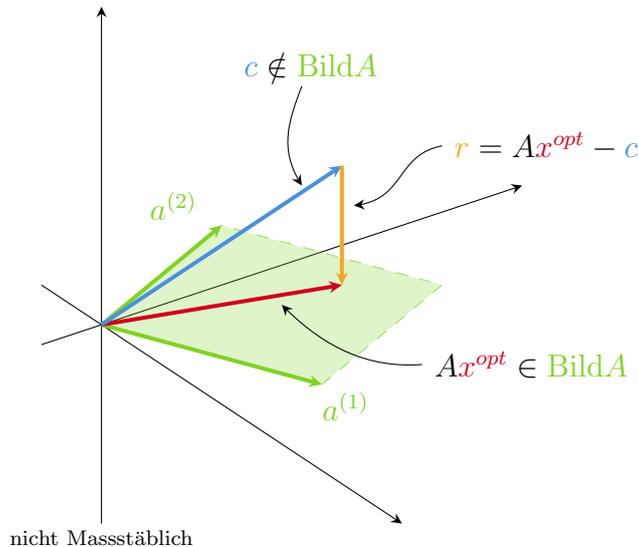
Da alle Abstände gleichzeitig minimiert werden sollen, fassen wir alle r_i im sogenannten Residuenvektor $r = (r_1, \dots, r_m)^\top$ zusammen und minimieren dann die gesamte Länge von r , das entspricht der quadratischen Minimierung des Fehlers r . Wir suchen also das x sodass

$$\|r\|_2 = \|Ax - c\|_2,$$

minimal wird. In unserem Fall suchen wir $x^{opt} = (x_1^{opt} \quad x_2^{opt})^\top$ welches die Länge von r minimiert. Aber wie können wir dieses Problem lösen? Betrachten wir dafür ein reduziertes Gleichungssystem mit nur drei Gleichungen und zwei Unbekannten, dafür nehmen wir einfach die ersten drei Messwerte von oben.

$$\begin{array}{l} x_1 \cdot 41 + x_2 = 172 \\ x_1 \cdot 45 + x_2 = 190 \\ x_1 \cdot 42 + x_2 = 180 \end{array} \quad \rightarrow \quad \underbrace{\begin{pmatrix} 41 & 1 \\ 45 & 1 \\ 42 & 1 \end{pmatrix}}_A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{\begin{pmatrix} 172 \\ 190 \\ 180 \end{pmatrix}}_c$$

Genau wie oben ist dieses LGS überbestimmt und in dieser Form nicht exakt lösbar. Grafisch können wir uns das Problem wie folgt vorstellen.



Der Vektor x^{opt} repräsentiert hier die optimale Lösung, welche den kürzesten Residuenvektor r hat und im Bild von A liegt. Grafisch können wir hier auch sehen, dass der Vektor r minimal ist, wenn er orthogonal zum Bild A ist. Für die optimale Lösung muss also gelten:

$$\underbrace{Ay}_{\text{Bild } A} \perp r, \quad \forall y \in \mathbb{R}^n.$$

Mit dem Skalarprodukt können wir das umformulieren zu

$$\langle Ay, r \rangle = 0,$$

$$(Ay)^\top r = 0.$$

Nun können wir unsere Definition für r einsetzen

$$y^\top A^\top (Ax^{opt} - c) = 0,$$

und alles ausmultiplizieren

$$y^\top A^\top Ax^{opt} - y^\top A^\top c = 0.$$

Schliesslich können wir noch umstellen und nach x^{opt} auflösen

$$A^\top c = A^\top Ax^{opt},$$

$$(A^\top A)^{-1} A^\top c = x^{opt}.$$

Dieser Ausdruck gibt uns nun die optimale Lösung, indem der quadratische Fehler (Länge des Residuenvektors) minimiert wird. Oft wollen wir jedoch nicht die inverse von $A^\top A$ berechnen und können alternativ das folgende LGS lösen

$$A^\top Ax^{opt} = A^\top c.$$

Kleinste Quadrate

Wir wollen die optimale Lösung x^{opt} finden welche den quadratischen Fehler $\|r\|_2 = \|Ax - c\|_2$ minimiert. Oft ist das LGS schon in der Form $Ax - c = r$ gegeben.

1. Bestimme A und c
2. Berechne $A^T A$ und $A^T c$
3. Löse das LGS $A^T A x^{opt} = A^T c$

9.2 QR-Zerlegung

Bei der QR-Zerlegung werden wir eine Matrix $A \in \mathbb{R}^{m \times n}$ in das Produkt einer orthogonalen Matrix $Q \in \mathbb{R}^{m \times m}$ und einer oberen Dreiecksmatrix $R \in \mathbb{R}^{m \times n}$ zerlegen. Es gilt dann

$$\begin{array}{ccc} A & = & Q \cdot R \\ \begin{array}{c} m \\ \square \\ n \end{array} & = & \begin{array}{c} m \\ \square \\ m \end{array} \cdot \begin{array}{c} m \\ \begin{array}{c} \diagdown \\ 0 \quad \diagdown \\ 0 \quad 0 \end{array} \\ n \end{array} \end{array}$$

Später können wir die QR-Zerlegung als ein alternatives Lösungsverfahren der kleinsten Quadrate verwenden. Um die QR-Zerlegung zu berechnen, müssen einige Schritte befolgt werden. Betrachten wir das Ganze an einem Beispiel.

Wir suchen nun Q und R sodass für die Matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix},$$

gilt $A = QR$. Dafür eliminieren wir schrittweise alle Elemente a_{ij} unterhalb der Hauptdiagonalen von A . In unserem fall beginnen wir mit dem Eintrag a_{31} . Zuerst notieren wir die Werte für i, j und die Einträge a_{jj} , a_{ij}

$$i = 3, j = 1.$$

$$a_{jj} = 1, a_{ij} = 1.$$

Anschliessend berechnen wir den Wert einer Zwischenvariable ω gegeben durch

$$\omega = \sqrt{a_{jj}^2 + a_{ij}^2} = \sqrt{1^2 + 1^2} = \sqrt{2}.$$

Als Nächstes suchen wir eine Rotationsmatrix Q'^T . Dafür fangen wir mit einer Identitätsmatrix $I \in \mathbb{R}^{m \times m}$ an und füllen die Einträge wie folgt ein

$$i_{ii} = \cos(\alpha), \quad i_{ij} = -\sin(\alpha), \quad i_{ji} = \sin(\alpha), \quad i_{jj} = \cos(\alpha).$$

Wobei die spezifischen Werte für \cos und \sin durch das ω gegeben sind

$$\sin(\alpha) = \frac{a_{ij}}{\omega} = \frac{1}{\sqrt{2}}, \quad \cos(\alpha) = \frac{a_{jj}}{\omega} = \frac{1}{\sqrt{2}}.$$

In unserem Fall ist Q'^T gegeben durch

$$Q'^T = \begin{pmatrix} \cos(\alpha) & 0 & \sin(\alpha) \\ 0 & 1 & 0 \\ -\sin(\alpha) & 0 & \cos(\alpha) \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Nun Berechnen wir die Matrix $A' = Q'^T A$ und überprüfen ob A' eine obere Dreiecksmatrix ist.

$$A' = Q'^T A = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & 1 \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Falls, wie in unserem Fall, A' keine obere Dreiecksmatrix ist, wiederholen wir den Prozess und nehmen den nächsten Eintrag unterhalb der Hauptdiagonalen von A' . Hier wäre das der Eintrag a'_{32} dadurch werden $i = 3, j = 2$. Wir lesen wieder die Einträge a'_{ij}, a'_{jj} und ω ab

$$a'_{jj} = 1, \quad a'_{ij} = \frac{1}{\sqrt{2}}, \quad \omega = \sqrt{\frac{3}{2}}.$$

Weiter konstruieren wir wie oben eine Rotationsmatrix Q''^T welche die folgende Form annimmt

$$Q''^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) \\ 0 & -\sin(\alpha) & \cos(\alpha) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix},$$

wodurch wir die Matrix $A'' = Q''^T A'$ erhalten

$$A'' = \begin{pmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{3}}{\sqrt{2}} \\ 0 & 0 \end{pmatrix}.$$

Nun ist A'' eine obere Dreiecksmatrix und es gilt $A'' = R$. Um schliesslich noch die Matrix Q zu erhalten, multiplizieren wir alle Rotationsmatrizen zusammen

$$Q^T = Q''^T Q'^T.$$

Somit ist die QR-Zerlegung vollendet. Zusammengefasst müssen also folgende Schritte befolgt werden.

QR-Zerlegung

Bei der QR-Zerlegung zerlegen wir eine Matrix $A \in \mathbb{R}^{m \times n}$ in das Produkt einer orthogonalen Matrix $Q \in \mathbb{R}^{m \times m}$ und einer oberen Dreiecksmatrix $R \in \mathbb{R}^{m \times n}$. Nach und nach eliminieren wir die Einträge unterhalb der Hauptdiagonalen von A .

1. Wähle das zu eliminierende Element a_{ij} unter der Hauptdiagonalen
2. Notiere die Werte i, j, a_{ij}, a_{jj}
3. Berechne $\omega = \sqrt{a_{ij}^2 + a_{jj}^2}$
4. Finde die Rotationsmatrix Q'^T basierend auf einer Identitätsmatrix $I \in \mathbb{R}^{m \times m}$ und setze $i_{ii} = \cos(\alpha), i_{ij} = -\sin(\alpha), i_{ji} = \sin(\alpha), i_{jj} = \cos(\alpha)$
5. Für die Werte von $\cos(\alpha)$ und $\sin(\alpha)$ setze $\sin(\alpha) = \frac{a_{ij}}{\omega}, \cos(\alpha) = \frac{a_{jj}}{\omega}$
6. Berechne $A' = Q'^T A$
7. Überprüfe ob A' eine obere Dreiecksmatrix ist, falls nicht, wiederhole den Prozess mit dem nächsten Eintrag a'_{ij} von A' solange, bis alle Einträge unterhalb der Hauptdiagonalen eliminiert sind
8. Wenn nach k Wiederholungen $A^{(k)} = R$ ist, berechne $Q^T = Q^{(k)T} \cdot \dots \cdot Q''^T \cdot Q'^T$
9. Nun ist $A = QR$

9.3 Kleinste Quadrate mit QR-Zerlegung

In gewissen Fällen kann es sein das die oben gezeigte Methode zur Lösung von überbestimmten LGS ungenaue Ergebnisse liefert. Vor allem, wenn Sie numerisch mit dem Computer berechnet werden. Oft kann man für eine numerisch stabilere Lösung die QR-Zerlegung verwenden.

Kleinste Quadrate mit QR-Zerlegung

Ein alternativer Lösungsweg um überbestimmte LGS $Ax = c$ durch kleinste Quadrate zu lösen.

1. Bestimme A und c
2. Berechne die QR-Zerlegung $A = QR$
3. Berechne $d = Q^T c$
4. Löse das LGS $R_0 x = d_0$, wobei R_0 die extrahierte Dreiecksmatrix von R ist und d_0 die dazugehörigen oberen Einträge von d
5. Die Lösung x ist die optimale Lösung für $Ax = c$

10 Lineare Differentialgleichungssysteme

10.1 Homogene lineare Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten

Solche Gleichungen sind bereits aus der Analysis bekannt. Allgemein ist eine solche Differentialgleichung gegeben durch

$$y'(t) = ay(t), \quad a \in \mathbb{R},$$

und der Lösung

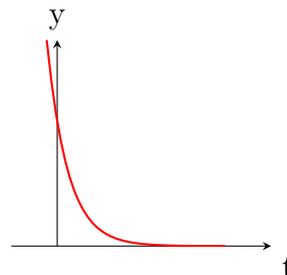
$$y(t) = y(0) e^{at}.$$

Wobei $y(0)$ die Anfangsbedingung ist. Ein konkretes Beispiel hierfür ist die Differentialgleichung

$$y'(t) = -2y(t), \quad y(0) = 3.$$

Die Lösung ist dann gegeben durch

$$y(t) = y(0)e^{-2t} = 3e^{-2t}.$$



Die Lösungsmenge dieser Differentialgleichung $\{y \in C^1(\mathbb{R}) : y' = ay\}$ ist ein 1-D Unterraum von den 1-mal stetig differenzierbaren Funktionen $C^1(\mathbb{R})$.

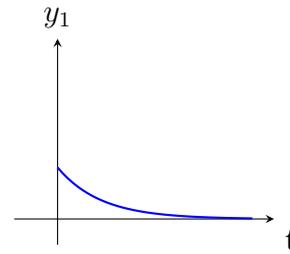
10.2 Systeme homogener linearer Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten

In der linearen Algebra I haben wir bereits gesehen wie wir Systeme von Gleichungen lösen und manipulieren können. Das geht auch mit Differentialgleichungen. Ein Beispiel dafür wäre zum Beispiel das Gleichungssystem gegeben durch

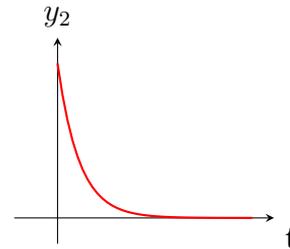
$$\begin{cases} y_1'(t) = -2y_1(t), & y_1(0) = 1 \\ y_2'(t) = -4y_2(t), & y_2(0) = 3. \end{cases}$$

Da beide Gleichungen von unterschiedlichen Variablen abhängen sind sie unabhängig voneinander, solche Systeme nennen wir entkoppelt. In einem solchen entkoppelten System können wir die Differentialgleichungen separat voneinander lösen.

$$y_1(t) = y_1(0)e^{-2t} = e^{-2t}$$



$$y_2(t) = y_2(0)e^{-4t} = 3e^{-4t}$$



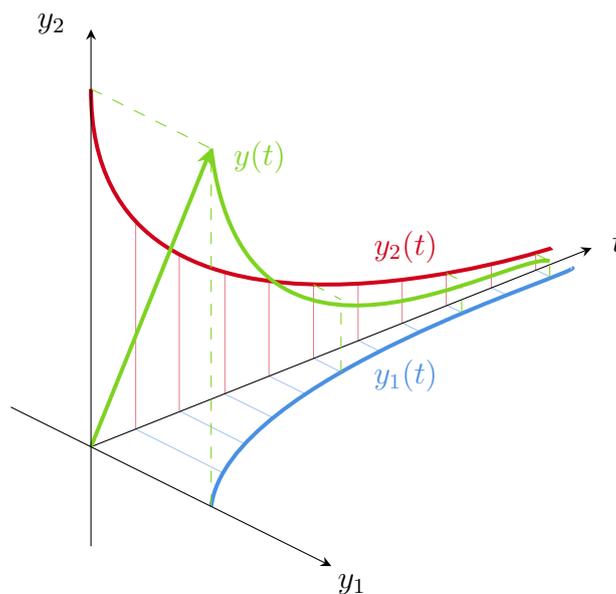
Wie bei den herkömmlichen LGS der Form $Ax = b$, können wir auch dieses System von linearen Differentialgleichungen in Matrixform schreiben, bloss, dass die allgemeine Form nun $y' = Ay$ ist. In unserem Fall ergibt sich

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & -4 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}, \quad \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} = y(0) = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Die Lösung dieses Systems lässt sich dann als Vektor schreiben:

$$y(t) = \begin{pmatrix} e^{-2t} \\ 3e^{-4t} \end{pmatrix} = 1 \cdot e^{-2t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \cdot e^{-4t} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Grafisch können wir uns die zusammengesetzte Lösung y wie eine Linearkombination von y_1 und y_2 vorstellen, bloss das jetzt alles noch von t abhängt.



Es kann aber auch sein, dass die Differentialgleichungen nicht entkoppelt sind. In dem Fall sind die Gleichungen voneinander abhängig und wir können sie nicht mehr separat lösen. Ein Beispiel dafür wäre das Gleichungssystem

$$\begin{cases} y_1'(t) = -4y_1(t) - y_2(t) \\ y_2'(t) = -2y_1(t) - 3y_2(t) \end{cases} \longrightarrow \begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} -4 & -1 \\ -2 & -3 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}, \quad y(0) = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Es stellt sich nun die Frage wie wir ein solches System dennoch lösen können. Betrachten wir dafür nochmals genauer die Matrizen A_1 und A_2 der beiden Systeme

$$A_1 = \begin{pmatrix} -2 & 0 \\ 0 & -4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -4 & -1 \\ -2 & -3 \end{pmatrix}.$$

Wir erkennen, dass die Matrix A_1 , des entkoppelten Systems, diagonal ist und die andere nicht. Um nun eine Lösung für das zweite System zu finden, können wir einen Basiswechsel $y = Tz$ durchführen und dadurch die Matrix A_2 diagonalisieren, wodurch das System entkoppelt gelöst werden kann. Wir transformieren also das gesamte Gleichungssystem in die Eigenbasis von A_2 . Dafür benötigen wir zunächst die Eigenwerte und Eigenvektoren von A_2 . Die Spalten der Transformationsmatrix T sind dann durch die Eigenvektoren gegeben.

$$\lambda_1 = -2, \quad \lambda_2 = -5, \quad T = \begin{pmatrix} -\frac{1}{2} & 1 \\ 1 & 1 \end{pmatrix}, \quad T^{-1} = \frac{1}{3} \begin{pmatrix} -2 & 2 \\ 2 & 1 \end{pmatrix}.$$

Nun können wir die Transformation durchführen.

$$\begin{aligned} \begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} &= \begin{pmatrix} -4 & -1 \\ -2 & -3 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} & \xrightarrow{z=T^{-1}y} & \begin{pmatrix} z_1'(t) \\ z_2'(t) \end{pmatrix} &= \underbrace{\begin{pmatrix} -2 & 0 \\ 0 & -5 \end{pmatrix}}_{T^{-1}AT} \begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} \\ y(0) &= \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} & & z(0) &= \begin{pmatrix} z_1(0) \\ z_2(0) \end{pmatrix} = \underbrace{\begin{pmatrix} 4 \\ 5 \\ 3 \end{pmatrix}}_{T^{-1}y(0)} \end{aligned}$$

In der Basis z können wir das Problem nun wie oben entkoppelt lösen

$$z(t) = \frac{4}{3} \cdot e^{-2t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{5}{3} \cdot e^{-5t} \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

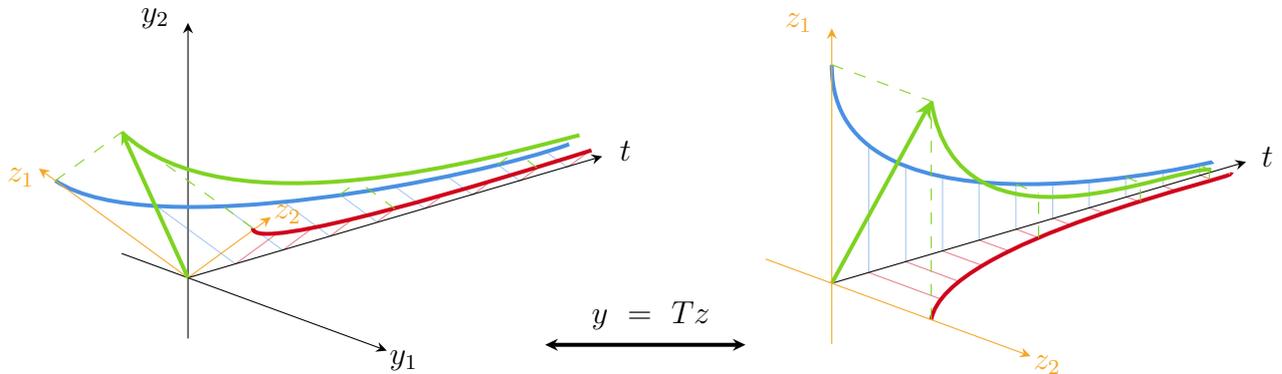
und danach wieder mit $y = Tz$ zurücktransformieren.

$$\begin{aligned} y(t) &= \frac{4}{3} \cdot e^{-2t} \begin{pmatrix} -\frac{1}{2} & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{5}{3} \cdot e^{-5t} \begin{pmatrix} -\frac{1}{2} & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \frac{4}{3} \cdot e^{-2t} \begin{pmatrix} -\frac{1}{2} \\ 1 \end{pmatrix} + \frac{5}{3} \cdot e^{-5t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (\dagger) \\ &= \frac{1}{3} \begin{pmatrix} -2e^{-2t} + 5e^{-5t} \\ 4e^{-2t} + 5e^{-5t} \end{pmatrix} \end{aligned}$$

In der Gleichung (\dagger) erkennen wir, dass die Lösung die Form $y(t) = z_1(0) \cdot e^{\lambda_1 t} v_1 + z_2(0) \cdot e^{\lambda_2 t} v_2$ annimmt, wobei v_1 und v_2 die Eigenvektoren von A_2 sind. Das ist bei Systemen, in denen A diagonalisierbar ist, immer der Fall weswegen wir die allgemeine Lösung immer direkt in der folgenden Form schreiben können:

$$y(t) = z_1(0)v_1e^{\lambda_1 t} + z_2(0)v_2e^{\lambda_2 t} + \dots + z_n(0)v_ne^{\lambda_n t} \quad (*)$$

Grafisch kann man sich diese Transformation wie folgt vorstellen.



10.2.1 Allgemeine Form Alle Differentialgleichungssysteme von homogenen linearen Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten lassen sich mithilfe von Matrizen schreiben. Die allgemeine Form und Lösung sind gegeben durch

$$\begin{cases} y_1' = a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n \\ y_2' = a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n \\ \dots = \dots \\ y_n' = a_{n1}y_1 + a_{n2}y_2 + \dots + a_{nn}y_n \end{cases} \longrightarrow y'(t) = Ay(t),$$

mit dem Variablen Vektor y und der Matrix A

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Weiterhin ist die Anfangsbedingung $y(0) = y_0 \in \mathbb{R}^n$ und damit die allgemeine Lösung gegeben durch

$$y(t) = e^{At}y_0.$$

Um ein solches System zu lösen, müssen wir das Matrixexponential e^{At} auswerten. Das ist im Allgemeinen über die Taylorreihe definiert und ist oft nicht einfach zu berechnen. Jedoch können wir, falls die Matrix A diagonalisierbar ist, das Matrixexponential vereinfachen und auf (*) rückschließen

$$y(t) = e^{At}y_0 = T \underbrace{\text{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_n t})}_{v_i \cdot e^{\lambda_i}} \underbrace{T^{-1}y_0}_{z_i(0)}.$$

Lösen von homogenen linearen Differentialgleichungssystemen

Für die Lösung eines homogenen linearen Differentialgleichungssystems 1. Ordnung in der Standardform $y' = Ay$ und den Anfangsbedingungen $y(0) = y_0$ gehen wir wie folgt vor: (A muss diagonalisierbar sein)

- Diagonalisiere die Matrix $A = TDT^{-1}$ und bestimme die Transformationsmatrix T , sodass $y = Tz$.
- Berechne die Anfangswerte $z(0)$ durch $z(0) = T^{-1}y_0$ oder mit dem LGS $Tz(0) = y_0$.
- Sei $t^{(i)}$ die i -te Spalte von T und d_{ii} der i -te Diagonaleintrag von D . Die Lösung des Systems lautet dann

$$y(t) = z_1(0)t^{(1)}e^{d_{11}t} + z_2(0)t^{(2)}e^{d_{22}t} + \dots + z_n(0)t^{(n)}e^{d_{nn}t}$$

10.3 Lineare Differentialgleichungen höherer Ordnung

Bis jetzt haben wir nur Differentialgleichungen 1. Ordnung betrachtet, also Gleichungen in denen maximal die erste Ableitung vorkommt. Wie lösen wir jedoch Differentialgleichungen höherer Ordnung? Betrachten wir hierfür die Differentialgleichung 3. Ordnung

$$y''' + 4y'' + 2y' - 3y = 0.$$

Durch eine Substitution können wir die Differentialgleichung in ein homogenes lineares Differentialgleichungssystem 1. Ordnung umwandeln. Dafür substituieren wir

$$\begin{array}{ll} y_0 := y & \\ y_1 := y' & \\ y_2 := y'' & \\ y_3 := y''' & \end{array} \quad \text{wodurch} \quad \begin{array}{l} y_1 = y'_0 \\ y_2 = y'_1 \\ y_3 = y'_2 \end{array}$$

und schreiben die originale Gleichung mit den neuen Variablen y_i , wobei die höchste Ableitung, hier y''' , durch eine erste Ableitung, hier y'_2 , ersetzt wird.

$$y'_2 + 4y_2 + 2y_1 - 3y_0 = 0.$$

Damit die Information über die Substitution nicht verloren geht, schreiben wir Sie als weitere Gleichungen in einem System. Daraus resultiert das folgende System

$$\begin{cases} y'_0 = y_1 \\ y'_1 = y_2 \\ y'_2 = 3y_0 - 2y_1 - 4y_2 \end{cases} \quad \text{oder} \quad \begin{pmatrix} y'_0 \\ y'_1 \\ y'_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 3 & -2 & -4 \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix}$$

Nun können wir dieses System wie oben beschrieben lösen. Das Resultat y der originalen Gleichung ist dann durch die Lösung für y_0 des Systems gegeben.

10.4 Spezialfall: Differentialgleichungssysteme mit komplexen Eigenwerten

Nehmen wir Beispielsweise die Gleichung

$$x''(t) = -8x(t) + 4x'(t), \quad (\star)$$

welche durch die Substitution

$$\begin{array}{l} y_0 := x \\ y_1 := x' \\ y_2 := x'' \end{array} \quad \longrightarrow \quad \begin{array}{l} y_1 = y_0' \\ y_2 = y_1' \end{array}$$

in das System

$$\begin{cases} y_0' = y_1 \\ y_1' = -8y_0 + 4y_1 \end{cases} \quad \text{bzw.} \quad \begin{pmatrix} y_0' \\ y_1' \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -8 & 4 \end{pmatrix}}_A \begin{pmatrix} y_0 \\ y_1 \end{pmatrix},$$

umgewandelt werden kann. Beachte, dass die Gleichung (\star) , 2. Ordnung ist, weswegen der Lösungsraum von (\star) die Dimension 2 haben muss (Intuitiv: Wir müssen zweimal integrieren und haben deswegen zwei Integrationskonstanten bzw. Freiheitsgrade). Um die allgemeine Lösung zu finden, können wir wie oben beschrieben vorgehen. Zuerst bestimmen wir die Eigenwerte

$$\det \begin{pmatrix} -\lambda & 1 \\ -8 & 4 - \lambda \end{pmatrix} = \lambda^2 - 4\lambda + 8 = 0 \quad \rightarrow \quad \lambda_{1,2} = 2 \pm 2i,$$

und Eigenvektoren

$$E_{2+2i} : \left(\begin{array}{cc|c} -2-2i & 1 & 0 \\ -8 & 2-2i & 0 \end{array} \right) \xrightarrow[\leftarrow_+]{\cdot \frac{-8}{2+2i}} \left(\begin{array}{cc|c} -2-2i & 1 & 0 \\ 0 & 0 & 0 \end{array} \right) \quad \begin{cases} y_0 = \frac{s}{2+2i} \\ y_1 = s \in \mathbb{R} \end{cases}$$

$$\text{für } s = 4 \quad \begin{cases} y_0 = \frac{4}{2+2i} \cdot \underbrace{\frac{2-2i}{2-2i}}_{=1} = \frac{4(2-2i)}{4+4} = 1-i \\ y_1 = 4 \end{cases}$$

$$E_{2+2i} : \text{span} \left\{ \begin{pmatrix} 1-i \\ 4 \end{pmatrix} \right\}, \quad E_{2-2i} : \text{span} \left\{ \begin{pmatrix} 1+i \\ 4 \end{pmatrix} \right\}.$$

Wodurch sich die folgende allgemeine Lösung des Systems ergibt

$$y(t) = e^{(2+2i)t} \begin{pmatrix} 1-i \\ 4 \end{pmatrix} + e^{(2-2i)t} \begin{pmatrix} 1+i \\ 4 \end{pmatrix}.$$

Diese Lösung ist jedoch komplex. Normalerweise wollen wir aber eine reelle Lösung. Um eine reelle Lösung zu erhalten, benutzen wir die Eulersche Formel $re^{i\phi} = r \cos(\phi) +$

$i r \sin(\phi)$. Um die Notation folgend zu vereinfachen betrachten, wir nur einen der beiden Eigenvektoren der Lösung (wenn man beide betrachtet kommt man auf das gleiche Ergebnis nur mit mehr Notation, am Ende findest du die Alternative mit beiden).

$$\begin{aligned}
 y(t) &= e^{(2+2i)t} \begin{pmatrix} 1-i \\ 4 \end{pmatrix} = e^{2t} e^{2it} \begin{pmatrix} 1-i \\ 4 \end{pmatrix} \\
 &= e^{2t} (\cos(2t) + i \sin(2t)) \begin{pmatrix} 1-i \\ 4 \end{pmatrix} \\
 &= e^{2t} \left\{ \begin{pmatrix} \cos(2t) + i \sin(2t) - i \cos(2t) - \sin(2t) \\ 4 \cos(2t) + 4i \sin(2t) \end{pmatrix} \right\} \\
 &= e^{2t} \left\{ \begin{pmatrix} \cos(2t) + \sin(2t) \\ 4 \cos(2t) \end{pmatrix} + i \begin{pmatrix} \sin(2t) - \cos(2t) \\ 4 \sin(2t) \end{pmatrix} \right\}
 \end{aligned}$$

Für den letzten Schritt benutzten wir ein Korollar aus der Vorlesung welches besagt, dass wenn y eine komplexe Lösung eines homogenen linearen Differentialgleichungssystems $y' = Ay$ ist, so sind $\operatorname{Re}(y)$ und $\operatorname{Im}(y)$ ebenfalls Lösungen des Systems. $\operatorname{Re}(y)$ und $\operatorname{Im}(y)$ sind hier weiterhin linear unabhängig und bilden dadurch eine Basis des Lösungsraums von y . So lässt sich eine allgemeine Lösung auch als Linearkombination von $\operatorname{Re}(y)$ und $\operatorname{Im}(y)$ schreiben

$$y(t) = e^{2t} \left\{ a \begin{pmatrix} \cos(2t) + \sin(2t) \\ 4 \cos(2t) \end{pmatrix} + b \begin{pmatrix} \sin(2t) - \cos(2t) \\ 4 \sin(2t) \end{pmatrix} \right\}.$$

Da wir eigentlich eine Lösung für (\star) suchen und durch die Substitution $x = y_0$ substituiert wurde, nehmen wir nur die erste Zeile des Vektors $y(t)$.

$$\begin{aligned}
 x(t) = y_0(t) &= e^{2t} \{a(\cos(2t) + \sin(2t)) + b(\sin(2t) - \cos(2t))\} \\
 &= e^{2t} \{(a-b) \cos(2t) + (a+b) \sin(2t)\} \\
 &= e^{2t} \{c_1 \cos(2t) + c_2 \sin(2t)\}.
 \end{aligned}$$

10.4.1 Alternative Lösung mit beiden Eigenvektoren: Anstatt nur einen der beiden Eigenvektoren zu benutzen, können wir auch beide benutzen. Das Resultat ist das gleiche, jedoch mit mehr Notationsaufwand.

$$\begin{aligned}
y(t) &= c_3 \cdot e^{(2+2i)t} \begin{pmatrix} 1-i \\ 4 \end{pmatrix} + c_4 \cdot e^{(2-2i)t} \begin{pmatrix} 1+i \\ 4 \end{pmatrix} \\
&= e^{2t} \left\{ c_3 (\cos(2t) + i \sin(2t)) \begin{pmatrix} 1-i \\ 4 \end{pmatrix} + c_4 (\cos(2t) - i \sin(2t)) \begin{pmatrix} 1+i \\ 4 \end{pmatrix} \right\} \\
&= e^{2t} \begin{pmatrix} c_3 \cos(2t) + c_3 i \sin(2t) - c_3 i \cos(2t) - c_3 \sin(2t) + c_4 \cos(2t) - c_4 i \sin(2t) + c_4 i \cos(2t) - c_4 \sin(2t) \\ 4c_3 \cos(2t) + 4c_3 i \sin(2t) + 4c_4 \cos(2t) - 4c_4 i \sin(2t) \end{pmatrix} \\
&= e^{2t} \begin{pmatrix} (c_3 + c_4) \cos(2t) + (c_3 - c_4) i \sin(2t) - (c_3 - c_4) i \cos(2t) + (c_3 + c_4) \sin(2t) \\ 4(c_3 + c_4) \cos(2t) + 4(c_3 - c_4) i \sin(2t) \end{pmatrix} \\
&= e^{2t} \left\{ \begin{pmatrix} (c_3 + c_4) \cos(2t) + (c_3 + c_4) \sin(2t) \\ 4(c_3 + c_4) \cos(2t) \end{pmatrix} + i \begin{pmatrix} (c_3 - c_4) \sin(2t) - (c_3 - c_4) \cos(2t) \\ 4(c_3 - c_4) \sin(2t) \end{pmatrix} \right\} \\
&= e^{2t} \left\{ a \begin{pmatrix} \cos(2t) + \sin(2t) \\ 4 \cos(2t) \end{pmatrix} + i b \begin{pmatrix} \sin(2t) - \cos(2t) \\ 4 \sin(2t) \end{pmatrix} \right\}
\end{aligned}$$

womit wir schliesslich durch dieselbe Argumentation von oben die Lösung für (\star) finden

$$\begin{aligned}
x(t) = y_0(t) &= e^{2t} \{a(\cos(2t) + \sin(2t)) + b(\sin(2t) - \cos(2t))\} \\
&= e^{2t} \{(a - b) \cos(2t) + (a + b) \sin(2t)\} \\
&= e^{2t} \{c_1 \cos(2t) + c_2 \sin(2t)\}.
\end{aligned}$$