Linear Algebra II

Nicolas Bartzsch
nbartzsch@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/~nbartzsch/`.

Vielen Dank und viel Erfolg mit Lin Alg II.

Nicolas Bartzsch

Version: 27. März 2025

Inhaltsverzeichnis

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:

- $rang(A) = n$
- 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:

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

Äussere Operation:

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

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

Schauen wir uns diese einmal genauer an.

Innere Operation:

$$\oplus : V \times V \to V$$

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:

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

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

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

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

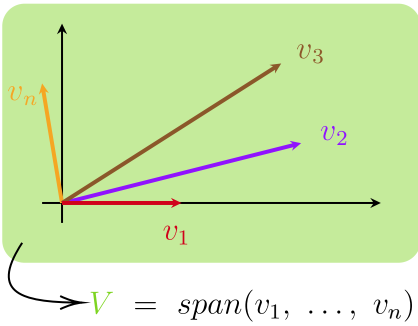
Ein Unterraum ist eine nichtleere Teilmenge U eines Vektorraums V, welche folgende Eigenschaften erfüllt:

$$\begin{aligned}
&\text{(i) } \forall u, v \in U : u \oplus v \in U \\
&\text{(ii) } \forall \alpha \in \mathbb{R},\ \forall u \in U : \alpha \odot u \in U
\end{aligned}$$

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 $v_1, \ldots, v_n$. Die Menge aller Linearkombinationen nennt sich lineare Hülle und wird mit $\text{span}(v_1, \ldots, v_n)$ abgekürzt. Diese Menge aller Linearkombinationen ist ein Vektorraum V. Die Vektoren $v_1, \ldots, v_n$ sind dann ein Erzeugendensystem von V. Es lassen sich alle Vektoren in V durch Linearkombinationen der Erzeugendenvektoren bilden.



Wenn alle Vektoren $v_1, \ldots, v_n$ 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

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

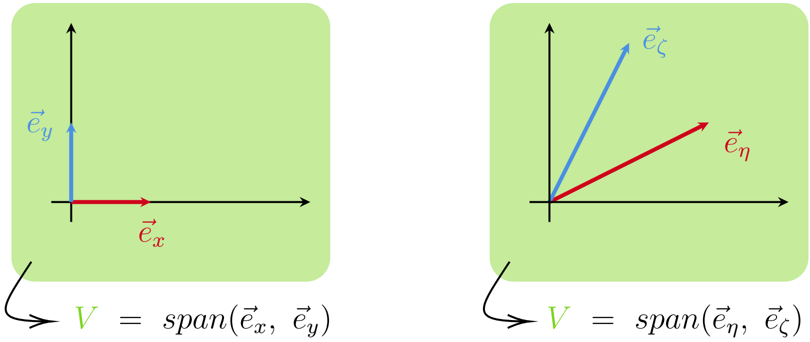
Jeder 2-D Vektor kann eindeutig als Linearkombination dieser beiden Basisvektoren dargestellt werden. Wir können jedoch auch zwei andere, linear unabhängige, Vektoren als Basisvektoren verwenden. Seien beispielsweise

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

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

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

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

$$\begin{aligned} i. \quad & \mathcal{F}(x+y) = \mathcal{F}(x) + \mathcal{F}(y) \\ ii. \quad & \mathcal{F}(\alpha x) = \alpha \mathcal{F}(x). \end{aligned}$$

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

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

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

$$\begin{aligned} i. \quad & \mathcal{F}(x+y) = 3(x+y) = 3x + 3y = \mathcal{F}(x) + \mathcal{F}(y) \\ ii. \quad & \mathcal{F}(\alpha x) = 3(\alpha x) = \alpha \cdot 3x = \alpha \cdot \mathcal{F}(x). \end{aligned}$$

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

Aus den Bedingungen i. und ii. foglen 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 $\mathcal{F} : V \to W$ durch eine Matrix $A \in \mathbb{R}^{m \times n}$ dargestellt werden. Allgemein kann man eine lineare Abbildung in der Form

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

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

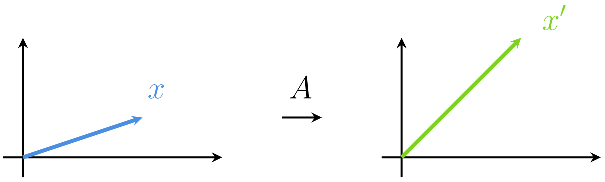
$$\mathcal{F} : \mathbb{R}^2 \to \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+y \\ x-2y \\ 3x \end{pmatrix}.$$

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

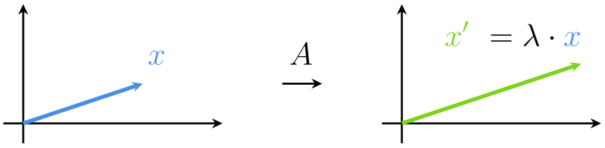
$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+y \\ x-2y \\ 3x \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & -2 \\ 3 & 0 \end{pmatrix}}_{:=A} \begin{pmatrix} x \\ y \end{pmatrix}.$$

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:

$$Ax = \lambda x. \tag{1}$$

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

$$(A - \lambda I)x = 0. \tag{2}$$

Wann hat dieses HLGS nur nichttriviale Lösungen?

Erinnerung aus der Linearen 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

$$\begin{aligned} \lambda_{1,2} = 3 \quad & \text{hat die algebraische Vielfachheit 2} \\ \lambda_3 = -6 \quad & \text{hat die algebraische Vielfachheit 1} \end{aligned}$$

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 Eingenwerte $\lambda \in \mathbb{C}$, 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 (2) umgeschrieben werden. Nun setzten 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

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

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

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

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:

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

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.

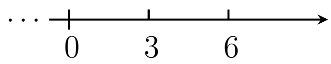
In diesem Beispiel ist also der Eigenraum E_1 zum Eigenwert 1 ein zweidimensionaler Unterraum. Die Dimension des Eigenraums $\dim(E_{\lambda_i})$ heisst geometrische Vielfachheit des Eigenwertes λ_i. In unserem Beispiel ist also die geometrische Vielfachheit von $\lambda_1 = 2$. Die geometrische Vielfachheit 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:

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

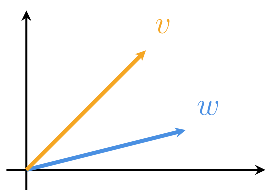
3 Normen und Skalarprodukte

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:

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

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

- Geometrische Länge (Euklidische Norm)

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2} \quad \to \quad \begin{aligned} \|v\|_2 &= \sqrt{18} \\ \|w\|_2 &= \sqrt{17} \end{aligned}$$

- 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:

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

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

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 beschrieben. Wieder lässt sich dies durch eine Abbildung beschreiben.

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

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

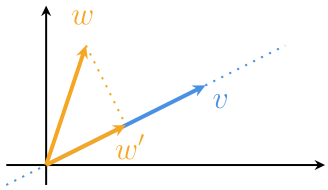
$$\|x\| := \sqrt{\langle x, x \rangle}$$

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

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

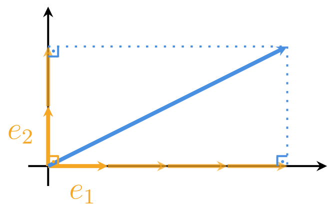
4.2 Orthonormalbasis

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:



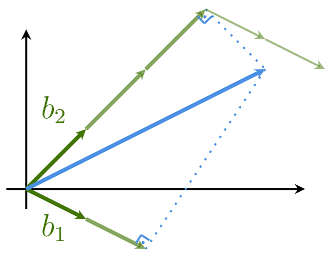
Durch die Orthogonalprojektionen auf die Basisvektoren erhalten wir die Koordinaten von v in der Basis $\mathcal{E}$. In diesem Fall sind die Koordinaten gegeben durch

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

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 $\begin{bmatrix} 2 & 3 \end{bmatrix}^\intercal$. Diese entsprechen jedoch nicht dem Vektor v.

$$v \neq \left[\begin{pmatrix} 2 \\ 3 \end{pmatrix} \right]_{\mathcal{B}}.$$

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 $\forall x \in V$

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

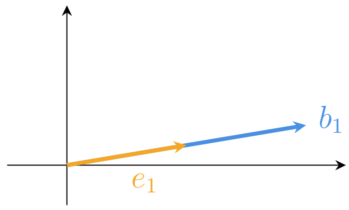
In anderen Worten, können wir jeden Vektor als Linearkombination der Basisvektoren schreiben, wobei die Koeffizienten durch die Orthogonalprojektionen gegeben sind. Um die Koordinaten zu finden, müssen wir kein LGS lösen. Beachte das der Term $\langle e_k, e_k \rangle$ wegfällt, da dieser aufgrund der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

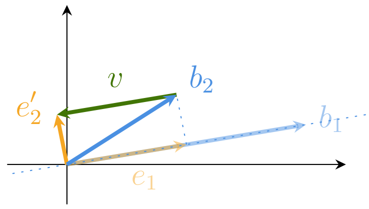
1. Wähle einen beliebigen Basisvektor b_1 und normiere ihn.

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



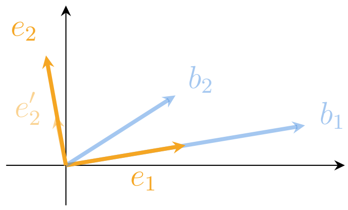
2. Wähle einen zweiten Basisvektor b_2, ziehe zuerst den zu e_1 parallelen Teil ab

$$e_2' = b_2 - \underbrace{\langle b_2, e_1\rangle e_1}_{v}$$



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

$$e_i' = b_i - \langle b_i, e_1\rangle e_1 - \langle b_i, e_2\rangle e_2 - \ldots - \langle b_i, e_{i-1}\rangle e_{i-1}$$

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

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 $span\{1, 3x^4\}$.

Wir benutzten 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.

$$e_2' = b_2 - \langle b_2, e_1 \rangle e_1$$

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

$$e_2' = b_2 - \frac{3}{5} = 3x^4 - \frac{3}{5}$$

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

$$\to \langle e_2', e_2' \rangle = \int_0^1 \left(3x^4 - \frac{3}{5} \right)^2 dx = \int_0^1 9x^8 - \frac{18}{5} x^4 + \frac{9}{25} \, dx$$

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

$$\to \sqrt{\langle e_2', e_2' \rangle} = \frac{4}{5}$$

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

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 $\mathcal{B} = \{b_1, \ldots, b_k\}$, sowie ein beliebeiges Skalarprodukt. Daraus lässt sich dann eine Orthonormalbasis $\mathcal{E} = \{e_1, \ldots, e_k\}$ erstellen.

1. Wähle einen beliebeigen 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 zwieten Basisvektor b_2, ziehe den zu b_1 parallelen Teil ab un 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

$$e_i' = b_i - \langle b_i, e_1 \rangle e_1 - \langle b_i, e_2 \rangle e_2 - \ldots - \langle b_i, e_{i-1} \rangle e_{i-1}$$

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

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

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

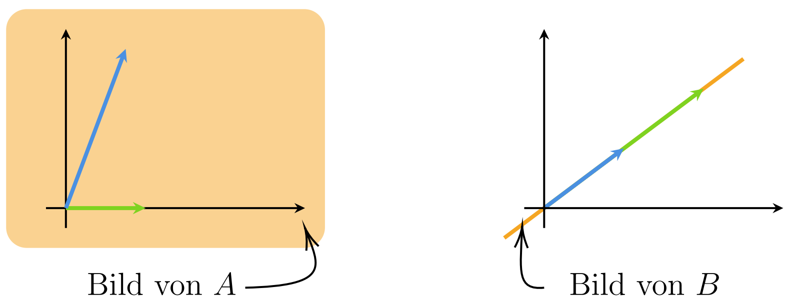
Das heisst, dass alle möglichen Ergebnisse einer Matrix-Vektor-Multiplikation Ax durch die Linearkombination der Spaltenvektoren von A dargestellt werden können. Die Menge aller dieser Ergebnisse nennt sich Bild. Mathematisch ist das Bild einer Matrix $A \in \mathbb{R}^{m \times n}$ durch

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

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

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

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

$$Ax = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{x_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 0 \end{pmatrix}}_{\text{linear unabhängig}},$$

$$Bx = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{x_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 4 \end{pmatrix}}_{\text{linear abhängig}}.$$

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

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

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

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

- **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:

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

  Somit ist eine mögliche Basis gegeben durch

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

  (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.)

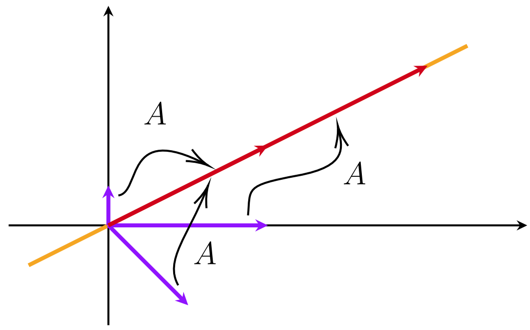
5.2 Kern

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

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

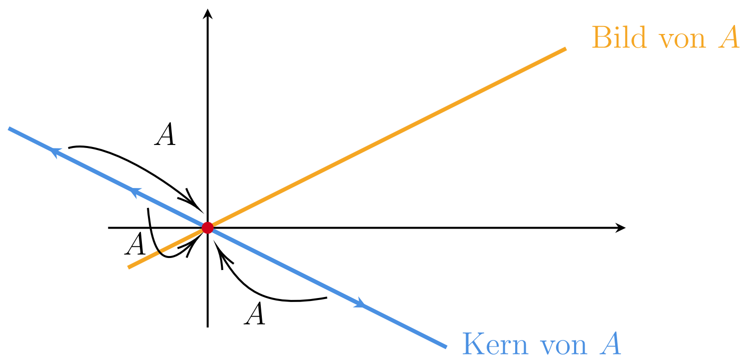
können wir uns das Bild grafisch als eine Linie vorstellen da A zwei linear abhängige Spalten hat. Betrachten wir nun was passiert, wenn einige Vektoren v durch A abgebildet werden, d.h. $Av = u$

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



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:

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



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:

- $rang(A) = n$
- 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 $\mathbb{R}^n$
- 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 $Ax = 0$:

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

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

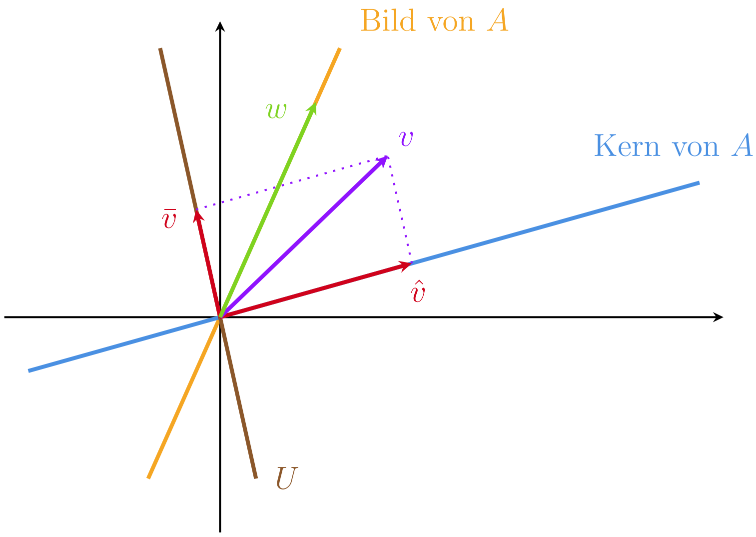
Beginnen wir mit dem grafischen Beweis in 2 Dimensionen. Sei $A \in \mathbb{R}^{m \times n}$ und $v \in \mathbb{R}^n$ ein beliebiger Vektor. Ausserdem sei $Av = w$. Wir können nun v in zwei Komponenten zerlegen:

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

Dadurch ist:

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

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:

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 5 \\ 4 & 3 & 9 \end{pmatrix} \quad \xrightarrow{\text{HLGS}} \quad \left.\begin{matrix} 1 & 2 & 4 \\ 3 & 1 & 5 \\ 4 & 3 & 9 \end{matrix}\,\right|\,\begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \quad \xrightarrow{\text{Gauss}} \quad \left.\begin{matrix} 1 & 2 & 4 \\ 0 & -5 & -7 \\ 0 & 0 & 0 \end{matrix}\,\right|\,\begin{matrix} 0 \\ 0 \\ 0 \end{matrix}$$

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

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

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

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:

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

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 $\mathrm{Bild}(A) = \left\{a^{(1)}, a^{(2)}, a^{(3)}\right\}$. Durch das Transponieren werden die Spalten $a^{(n)}$ zu den Zeilen $a^{[n]}$.

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

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

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

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:

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

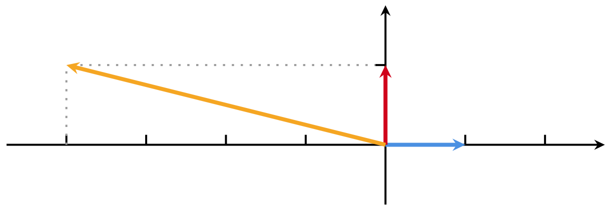
Aus dieser Bedingung folgt die Fredholm Alternative welche besagt, dass $Ax = b$ lösbar ist (b liegt im Bild) genau dann, wenn b senkrecht auf allen Lösungen des adjungierten LGS $A^\top y = 0$ steht.

6 Basiswechsel und Diagonalisierung

6.1 Basiswechsel für Vektoren

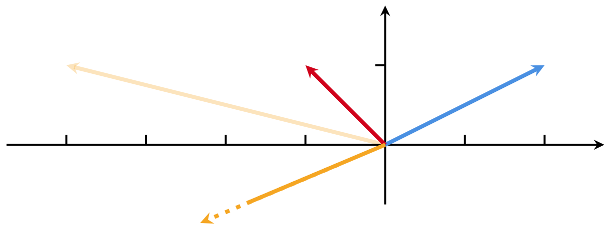
Was bedeutet es nochmal, wenn wir einen Vektor mit den Koordinaten $\begin{bmatrix}-4 & 1\end{bmatrix}^\top$ beschreiben? Grundlegend beschreiben Koordinaten um wie viel die Basisvektoren des assoziierten Vektorraums skaliert werden. In der Standardbasis ist dies bereits recht intuitiv.

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



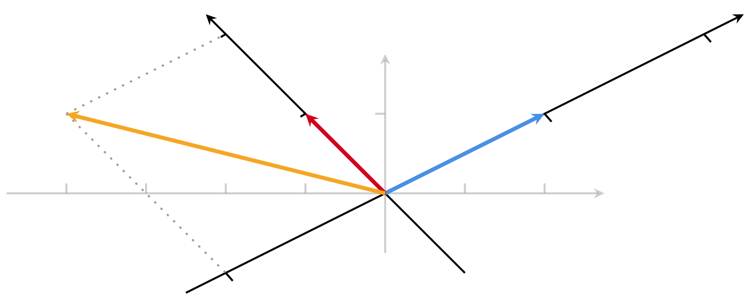
Wir können aber auch eine andere Basis wählen, z.B. gegeben durch die Vektoren $\begin{bmatrix}2 & 1\end{bmatrix}^\top$ und $\begin{bmatrix}-1 & 1\end{bmatrix}^\top$. Wenn wir nun dieselben Koordinaten wie in der Standardbasis benutzten, bekommen wir einen anderen Vektor.

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



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 $\begin{bmatrix}-1 & 2\end{bmatrix}^\top$ 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:

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

Dies können wir in ein LGS in Matrixschreibweise umformen:

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

Wir erkennen hier, dass wenn ein Vektor in der Basis $\mathcal{B}'$ mit der Matrix $T_{\mathcal{B}' \mapsto \mathcal{B}}$ multipliziert wird, derselbe Vektor in der Basis $\mathcal{B}$ resultiert. Die Matrix $T_{\mathcal{B}' \mapsto \mathcal{B}}$ 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:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{\mathcal{B}'} = T^{-1}_{\mathcal{B}' \mapsto \mathcal{B}} \begin{bmatrix} -4 \\ 1 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}_{\mathcal{B}'}.$$

Die gewünschte Übergangsmatrix ist also gegeben durch:

$$T^{-1}_{\mathcal{B}' \mapsto \mathcal{B}} = T_{\mathcal{B} \mapsto \mathcal{B}'} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}$$

Basiswechsel für Vektoren

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

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

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

$$T_{\mathcal{Q} \mapsto \mathcal{W}} = T^{-1}_{\mathcal{W} \mapsto \mathcal{Q}}$$

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

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

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

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

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:

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

Basiswechsel für Matrizen

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

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

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

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

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

Transformation von $[x]_{\mathcal{W}}$ zu Basis $\mathcal{Q}$

Abbildung von $[x]_{\mathcal{Q}}$ in Basis $\mathcal{Q}$

Rücktransformation von $[y]_{\mathcal{Q}}$ zu Basis $\mathcal{W}$

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 benutzten hier die oben eingeführte Transformationsmatrix in diesem Fall bekommen wir

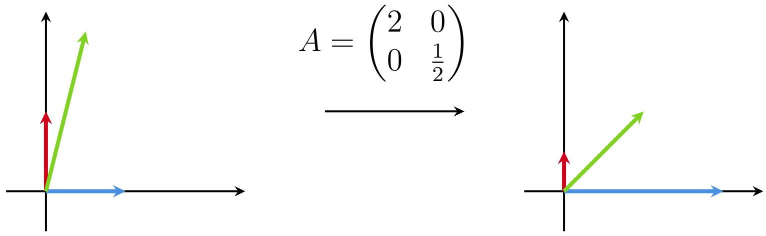
$$\begin{aligned} [A]_{\mathcal{B}} &= T_{\mathcal{S}\mapsto\mathcal{B}}\ [A]_{\mathcal{S}}\ T^{-1}_{\mathcal{S}\mapsto\mathcal{B}} \\ &= \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\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

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

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

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 seinen 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 $T_{\mathcal{B}\mapsto\mathcal{S}}$ bestimmen, indem wir die Eigenvektoren als Spalten nehmen:

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

Die diagonalisierte Matrix D ist dann gegeben durch

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

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

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

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: $T^{-1} = T^\top$.