

Mathe für Informatiker II

Mitschrift der Vorlesung von Christian Bender Sommersemester 17

Lukas Koschorke

10. Juli 2017

Vorwort

Hallo zusammen, ich versuche mein Skript immer aktuell zu halten und wenn ich es eh schon abtippe, kann ich es auch noch zusätzlich auf StudyDrive hochladen. Ich kann leider keine Garantie geben, dass es immer den kompletten Stoff aller Vorlesungen beinhaltet, auch wenn ich mein Bestes dafür gebe. Wenn euch Fehler auffallen, haut sie einfach in die Kommentare von StudyDrive.

Inhaltsverzeichnis

| | | |
|----------|---|-----------|
| 1 | Lineare Gleichungssysteme und der Gaußalgorithmus | 1 |
| 1.1 | Beispiel (PageRank) | 1 |
| 1.2 | Definition Zeilenstufenform | 4 |
| 1.3 | Algorithmus (Rückwärtseinsetzen) | 6 |
| 1.4 | Satz | 8 |
| 1.5 | Satz | 8 |
| 1.6 | Algorithmus | 8 |
| 1.7 | Beispiel | 10 |
| 1.8 | Algorithmus (Gaußsche Eliminationsverfahren zur Lösung von $Ax = b$) | 11 |
| 1.9 | Folgerung | 11 |
| 1.10 | Beispiel (Rundungsfehler im Gauß-Algorithmus) | 12 |
| 2 | Vektorräume | 15 |
| 2.1 | Definition | 15 |
| 2.2 | Beispiel | 15 |
| 2.3 | Satz | 15 |
| 2.4 | Definition | 16 |
| 2.5 | Beispiel | 17 |
| 2.6 | Definiton | 17 |
| 2.7 | Beispiel | 18 |
| 2.8 | Beispiel | 18 |
| 2.9 | Definition | 18 |
| 2.10 | Bemerkung | 18 |
| 2.11 | Beispiel | 19 |
| 2.12 | Beispiel | 19 |
| 3 | Lineare Unabhängigkeit und Basen | 21 |
| 3.1 | Definition | 21 |
| 3.2 | Bemerkung | 21 |
| 3.3 | Beispiel | 22 |
| 3.4 | Definiton | 22 |
| 3.5 | Bemerkung | 23 |
| 3.6 | Beispiele | 23 |
| 3.7 | Satz | 24 |
| 3.8 | Folgerung | 25 |
| 3.9 | Lemma (Austauschlemma) | 26 |
| 3.10 | Satz | 26 |

| | | |
|----------|--|-----------|
| 3.11 | Folgerung | 27 |
| 3.12 | Folgerung | 27 |
| 3.13 | Definition | 27 |
| 3.14 | Beispiel | 28 |
| 3.15 | Beispiel (und Erinnerung) | 28 |
| 3.16 | Folgerung(Basisergänzungssatz) | 28 |
| 3.17 | Folgerung | 29 |
| 4 | Matrizen | 31 |
| 4.1 | Definition | 31 |
| 4.2 | Algorithmus | 31 |
| 4.3 | Bemerkung | 32 |
| 4.4 | Bemerkung | 33 |
| 4.5 | Beispiel | 33 |
| 4.6 | Definition | 33 |
| 4.7 | Bemerkung | 34 |
| 4.8 | Satz | 34 |
| 4.9 | Bemerkung | 34 |
| 4.10 | Satz | 35 |
| 4.11 | Satz | 35 |
| 4.12 | Definition | 35 |
| 4.13 | Satz | 36 |
| 4.14 | Definition | 36 |
| 4.15 | Bemerkung | 36 |
| 4.16 | Satz | 38 |
| 4.17 | Bemerkung | 38 |
| 4.18 | Beispiel | 39 |
| 4.19 | Bemerkung | 40 |
| 4.20 | Bemerkung | 40 |
| 5 | Lineare Abbildungen | 41 |
| 5.1 | Definition | 41 |
| 5.2 | Beispiel | 41 |
| 5.3 | Beispiel | 41 |
| 5.4 | sehr einfacher Satz | 42 |
| 5.5 | Definition (Endomorphismus und Automorphismus) | 42 |
| 5.6 | Folgerung | 42 |
| 5.7 | Definition (Bild, Kern, Rang und Defekt) | 43 |
| 5.8 | Satz | 43 |
| 5.9 | Satz | 43 |
| 5.10 | Folgerung | 44 |
| 5.11 | Satz | 45 |
| 5.12 | Lemma | 45 |
| 5.13 | Definition (affiner Unterraum) | 45 |

| | | |
|----------|--|-----------|
| 5.14 | Bemerkung | 46 |
| 5.15 | Satz | 46 |
| 5.16 | Bemerkung | 46 |
| 5.17 | Beispiel | 46 |
| 6 | Koordinatentransformationen | 47 |
| 6.1 | Satz Koordinatensystem | 47 |
| 6.2 | Bemerkung | 47 |
| 6.3 | Beispiel (Diskrete Kosinustransformation zur Datenkomprimierung) | 48 |
| 6.4 | Bemerkung | 48 |
| 6.5 | sehr einfaches Beispiel | 49 |
| 6.6 | Satz | 49 |
| 6.7 | Satz | 50 |
| 6.8 | Bemerkung | 50 |
| 6.9 | Beispiel | 50 |
| 7 | Eigenwertproblem | 53 |
| 7.1 | Definition | 53 |
| 7.2 | Bemerkung | 53 |
| 7.3 | Definition | 53 |
| 7.4 | Satz | 53 |
| 7.5 | Beispiel | 54 |
| 7.6 | Bemerkung | 55 |
| 7.7 | Satz | 55 |
| 7.8 | Folgerung | 55 |
| 7.9 | Bemerkung | 56 |
| 7.10 | Satz | 56 |
| 7.11 | Satz (Jordansche Normalform) | 56 |
| 8 | Eigenwertprobleme für stochastische Matrizen | 59 |
| 8.1 | Lemma | 59 |
| 8.2 | Definition | 60 |
| 8.3 | Lemma | 61 |
| 8.4 | Satz | 61 |
| 8.5 | Satz (Spezialfall des Satzes von Frobenius-Perron) | 62 |
| 8.6 | Bemerkung | 62 |
| 8.7 | Folgerung | 63 |
| 8.8 | Bemerkung | 65 |
| 8.9 | Bemerkung | 65 |
| 9 | Determinante und charakteristisches Polynom | 67 |
| 9.1 | Beispiel | 67 |
| 9.2 | Beispiel | 67 |
| 9.3 | Definition | 68 |

| | | |
|-----------|---|-----------|
| 9.4 | Bemerkung | 69 |
| 9.5 | Satz | 69 |
| 9.6 | Bemerkung | 69 |
| 9.7 | Satz | 70 |
| 9.8 | Algorithmus | 71 |
| 9.9 | Folgerung | 72 |
| 9.10 | Beispiel | 72 |
| 9.11 | Satz (Leibniz Formel) | 72 |
| 9.12 | Satz (Laplacescher Entwicklungssatz) | 73 |
| 9.13 | Beispiel | 73 |
| 9.14 | Satz | 73 |
| 9.15 | Beispiel | 74 |
| 10 | Prä- Hilberträume und Orthogonalprojektionen | 77 |
| 10.1 | Beispiel | 77 |
| 10.2 | Definition | 78 |
| 10.3 | Bemerkung | 79 |
| 10.4 | Beispiel | 79 |
| 10.5 | Satz (Cauchy-Schwarz-Ungleichung) | 79 |
| 10.6 | Bemerkung | 80 |
| 10.7 | Definition | 80 |
| 10.8 | Satz | 81 |
| 10.9 | Definition | 82 |
| 10.10 | Beispiel | 82 |
| 10.11 | Satz | 82 |
| 10.12 | Definition | 82 |
| 10.13 | Bemerkung | 83 |
| 10.14 | Satz | 83 |
| 10.15 | Beispiel | 85 |
| 10.16 | Satz | 86 |
| 10.17 | Satz | 88 |
| 11 | Fourierreihen | 89 |
| 11.1 | Satz | 89 |
| 11.2 | Bemerkung | 90 |
| 11.3 | Beispiel | 91 |
| 11.4 | Definition (Fourierreihe) | 92 |
| 11.5 | Satz | 93 |
| 11.6 | Bemerkung | 93 |
| 11.7 | Bemerkung | 93 |
| 12 | Orthogonale Matrizen | 95 |
| 12.1 | Definition | 95 |
| 12.2 | Bemerkung | 95 |

| | |
|-------------------------|----|
| 12.3 Satz | 95 |
| 12.4 Beispiel | 96 |

1 Lineare Gleichungssysteme und der Gaußalgorithmus

1.1 Beispiel (PageRank)

Ein Nutzer surft auf den vorhandenen Seiten des Internets S_1, \dots, S_N . Er beginnt auf irgendeiner Seite und folgt typischerweise einem der Links. Er kann aber auch auf eine beliebige Seite springen.

Zur Modellierung sei $0 < d < 1$ (Damping-Faktor, typischerweise $d = 0,85$). Auf einer Seite angekommen, folgt der Nutzer mit Wahrscheinlichkeit d einem rein zufällig ausgewählten Link, mit Wahrscheinlichkeit $1 - d$ springt er auf eine rein zufällig gewählte Seite. (Konvention: Falls die Seite keinen Outlink hat, so wählt man rein zufällig eine Seite aus)

„Auf lange Sicht“ (d.h. wenn die Anzahl der Surfschritte gegen unendlich strebt, ist die Wahrscheinlichkeit p_i , sich auf der Seite S_i zu befinden, beschrieben durch

$$p_i = \frac{1-d}{N} + d \sum_{j=1}^N a_{ij} p_j, i = 1, \dots, N \quad (1.1)$$

wobei:

$$a_{ij} = \begin{cases} \frac{1}{c_j} & , \text{ falls Seite } S_j \text{ auf Seite } S_i \text{ verlinkt} \\ \frac{1}{N} & , \text{ falls } S_j \text{ keinen Outlink hat} \\ 0 & , \text{ sonst} \end{cases}$$

und c_j die Anzahl an Seiten angibt, auf die S_j verlinkt (siehe dazu das „Das Kapital der Markovketten“ in der MFI3).

p_i wird PageRank der Seite S_i genannt.

1.1 ist ein System von N linearen Gleichungen zu N Unbekannten.

In einem Netz mit 3 Seiten sei die Verlinkung schematisch dargestellt durch:

$$S_1 \rightleftharpoons S_3, S_2 \rightarrow S_1, S_3 \rightarrow S_2$$

1 Lineare Gleichungssysteme und der Gaußalgorithmus

Hier ist also $c_3 = 2, c_1 = c_2 = 1$, d.h. 1.1 wird im Spezialfall zu:

$$p_1 - dp_2 - \frac{d}{2}p_3 = \frac{1-d}{3} \quad (I)$$

$$p_2 - \frac{d}{2}p_3 = \frac{1-d}{3} \quad (II)$$

$$-dp_1 + p_3 = \frac{1-d}{3} \quad (III)$$

Indem man (I) durch $(\tilde{I}) = (I) + d(II) + \frac{1}{d}(III)$ ersetzt, erhält man die äquivalente Form:

$$-dp_1 + p_3 = \frac{1-d}{3} \quad (III)$$

$$p_2 - \frac{d}{2}p_3 = \frac{1-d}{3} \quad (II)$$

$$\left(\frac{1}{d} - \frac{d^2}{2} - \frac{d}{2}\right)p_3 = \frac{1-d}{3}\left(1 + \frac{1}{d} + d\right) \quad (\tilde{I})$$

Nun haben wir das System in „Zeilenstufenform“ und man kann die Lösung direkt ablesen ($0 < d < 1$):

$$p_3 = \frac{1-d}{3} \frac{d^2 + d + 1}{1 - \frac{d^3}{2} - \frac{d^2}{2}}$$

$$p_2 = \frac{1-d}{3} + \frac{d}{2}p_3$$

$$p_1 = \frac{1}{d}\left(p_3 - \frac{1-d}{3}\right)$$

Da $p_3 > \frac{1-d}{3}$ folgt $p_1 > 0$ und $p_2 > 0$.

Da ferner $p_1 + p_2 + p_3 = 1$ (Nachrechnen), kann der PageRank tatsächlich als Wahrscheinlichkeit interpretiert werden.

Für $d = 0,85$ ist:

$$p_1 \approx 0,397$$

$$p_2 \approx 0,215$$

$$p_3 \approx 0,388$$

Betrachten wir den „Grenzfall“ $d = 1$, in dem nur Links zum Surfen verwendet werden,

dann:

$$p_1 - p_2 - \frac{1}{2}p_3 = 0 \quad (I)$$

$$p_2 - \frac{1}{2}p_3 = 0 \quad (II)$$

$$-p_1 + p_3 = 0 \quad (III)$$

und wie zuvor:

$$-p_1 + p_3 = 0 \quad (III)$$

$$p_2 - \frac{1}{2}p_3 = 0 \quad (II)$$

$$0 = 0 \quad (\tilde{I})$$

Hier lösen alle p_1, p_2, p_3 der Form

$$p_1 = p_3, p_2 = \frac{1}{2}p_3, p_3 \in \mathbb{R}$$

das System 1.1. Fordert man zusätzlich die Wahrscheinlichkeitsbedingung $p_1 + p_2 + p_3 = 1$, so ist die eindeutige Lösung:

$$p_3 = p_1 = 0,4, p_2 = 0,2$$

Zur Bedeutung des Damping-Faktor:

- a)** Die Wahl $0 \leq d < 1$ sichert, das 1.1 genau eine Lösung hat und, dass diese Lösung positive Einträge hat.
Dies werden wir im Zusammenhang mit „Eigenwertproblemen“ besser verstehen.
- b)** Im realistischen Problem ist N sehr groß ($N > 20\text{ Mio.}$). Dann müssen numerische Verfahren zur Approximation des PageRank herangezogen werden. Wir werden sehen, dass die Wahl von d eine wichtige Rolle bei der Konvergenzgeschwindigkeit derartiger Verfahren spielt.

Problem:

Gegeben sei ein System von $m \in \mathbb{N}$ Gleichungen mit $n \in \mathbb{N}$ Unbekannten:

$$\left. \begin{array}{cccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \right\} (L)$$

Hierbei seien die Koeffizienten $a_{ij} \in \mathbb{R}$ (mit doppelter Indexschreibweise) für $i = 1, \dots, m$, $j = 1, \dots, n$, sowie die rechte Seite $b_i \in \mathbb{R}$, $i = 1, \dots, m$, gegeben.

Gesucht sind alle n -Tupel (x_1, \dots, x_n) von reellen Zahlen, die das System (L) lösen.

Zur übersichtlichen Notation schreibt man die Koeffizienten als Rechteckschema, das $m \times n$ -Matrix genannt wird

$$A := (a_{ij})_{i=1, \dots, m, j=1, \dots, n} := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

und die b_i 's und x_j 's als **Spaltenvektor**

$$x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, b := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Definiert man nun ein Produkt zwischen $m \times n$ -Matrix und Spaltenvektor der Höhe n , das einen Spaltenvektor der Höhe m ergibt, durch:

$$Ax := \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}$$

dann kann (L) umgeschrieben werden in die Form

$$Ax = b \text{ (L')}$$

als Gleichung von Spaltenvektoren der Höhe m .

Wir bezeichnen die Menge aller Spaltenvektoren der Höhe m mit reellen Einträgen als \mathbb{R}^m und die Menge aller $m \times n$ -Matrizen mit reellen Einträgen als $\mathbb{R}^{m \times n}$. Dann ist die Lösungsmenge von (L) gegeben durch:

$$\text{Lös}(A, b) := \{x \in \mathbb{R}^n | Ax = b\}$$

1.2 Definition Zeilenstufenform

Wir sagen, eine $m \times n$ -Matrix $A = (a_{ij})_{i=1, \dots, m, j=1, \dots, n}$ hat **Zeilenstufenform**, falls:

1. Es gibt ein $0 \leq r \leq m$, sodass in den Zeilen mit Index 1 bis r nicht alle Einträge gleich 0 sind und in der Zeile $(r+1)$ bis m alle Einträge gleich 0 sind und
2. es gelte

$$j_1 < \dots < j_r$$

wobei für jedes $1 \leq i \leq r$

$$j_i = \min\{j | a_{ij} \neq 0\}$$

der niedrigste Spaltenindex angibt, in dem in der i -ten Zeile ein von 0 verschiedener Eintrag steht.

Die Zahl r wird **Zeilenrang** von A genannt.

Die Einträge a_{i,j_i} , $i = 1, \dots, r$, werden **Pivots** genannt.

Offenbar gilt: $r \leq \min\{m, n\}$

$$A = \left(\begin{array}{ccc|c} (*) & & & \\ 0 & (*) & & \\ 0 & 0 & \ddots & \\ 0 & 0 & 0 & (*) \end{array} \right)$$

Matrix in Zeilenstufenform: Die mit $(*)$ gekennzeichneten Einträge sind die von 0 verschiedenen Pivots, die Einträge unterhalb der „Stufenlinie“ sind gleich 0, die übrigen Einträge sind beliebig.

Im Beispiel 1.1 hat die Koeffizientenmatrix zum System (III) , (II) , (\tilde{I}) Zeilenstufenform, nämlich:

$$\left(\begin{array}{ccc|c} -d & 0 & 1 & \\ 0 & 1 & -\frac{d}{2} & \\ 0 & 0 & (\frac{1}{2} - \frac{d^2}{2} - \frac{d}{2}) & \end{array} \right)$$

Fall 1:

Löse $Ax = b$, wobei A Zeilenstufenform hat.

- a)** Durch Umordnung der Spalten und damit Umnummerierung der Unbekannten, kann man stets voraussetzen, dass die Pivots in den ersten r Spalten stehen, d.h. $j_i = i$ für $1 \leq i \leq r$.

Beispiel:

$$\underbrace{\left(\begin{array}{cccccc} 0 & 2 & 0 & 4 & 0 & 5 \\ 0 & 0 & 1 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)}_{j_1=2, j_2=3, j_3=5} \rightsquigarrow \left(\begin{array}{cccccc} 2 & 0 & 0 & 5 & 0 & 4 \\ 0 & 1 & 1 & 0 & 0 & 3 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

- b)** Wir betrachten die **erweiterte Koeffizientenmatrix** zu $Ax = b$ (A beliebige $m \times n$ -Matrix).

$$(A, b) = \left(\begin{array}{cccc} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right)$$

, bei der die rechte Seite b als $(n+1)$ -te Spalte an die Koeffizientenmatrix angehängt wird. Ist A in Zeilenstufenform mit Pivots in den ersten r Spalten, so

$$(A, b) = \left(\begin{array}{ccc|ccc} a_{11} & & & & & b_1 \\ & \ddots & & & & \vdots \\ & & a_{rr} & & & b_r \\ & & & & & b_{r+1} \\ & & & & & \vdots \\ & & & & & b_m \end{array} \right)$$

1.3 Algorithmus (Rückwärtseinsetzen)

Input: $m \times n$ -Matrix A in Zeilenstufenform mit Zeilenrang r ($0 \leq r \leq m$) und Pivots in den ersten r Spalten, $b \in \mathbb{R}^m$.

Output: „es existiert keine Lösung“ oder eine bijektive Abbildung:

$$\Phi_{A,b} : \mathbb{R}^{n-r} \rightarrow \text{Lös}(A, b)$$

1. Falls $b_i \neq 0$ für ein $i = r+1, \dots, m$, so gebe „keine Lösung“ als Output aus.

2.1 Andernfalls setze in Abhängigkeit von $\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{n-r} \end{pmatrix}$

$$x_{r+1}(\lambda) := \lambda_1, \dots, x_n(\lambda) := \lambda_{n-r}$$

und für $i = r, \dots, 1$ rekursiv

$$x_i := \frac{1}{a_{ij}} \left(b_i - \sum_{j=i+1}^n a_{ij} x_j(\lambda) \right)$$

2.2 Gebe $\Phi_{A,b} : \mathbb{R}^{n-r} \rightarrow \text{Lös}(A, b), \lambda \mapsto x(\lambda) := \begin{pmatrix} x_1(\lambda) \\ \vdots \\ x_n(\lambda) \end{pmatrix}$ aus.

Bemerkung:

Ist $r = n$, so ist $\mathbb{R}^{n-r} = \mathbb{R}^0 := \{0\}$ eindeutig und das Rückwärtseinsetzen liefert genau eine Lösung für $Ax = b$

Beweis der Korrektheit des Algorithmus:

Ist $b_i \neq 0$ für ein $i = r + 1, \dots, m$, so lautet die i-te Gleichung

$$0 * x_1 + \dots + 0 * x_n = b_i \neq 0$$

und kann also für kein $x \in \mathbb{R}^n$ erfüllt sein. Andernfalls sind die Gleichungen r+1 bis m immer erfüllt ($0=0$).

Für $i = 1, \dots, r$ lautet die i-te Gleichung

$$a_{ii}x_i + a_{i,i+1}x_{i+1} + \dots + a_{in}x_n = b_i$$

und ist wegen $a_{ii} \neq 0$ äquivalent zu:

$$x_i = \frac{1}{a_{ii}}(b_i - \sum_{j=i+1}^n a_{ij}x_j)$$

Also: $x(\lambda) \in \text{Lös}(A, b)$ für alle $\lambda \in \mathbb{R}^{n-r}$

Zur Bijektivität:

Injektiv ist klar, da $x(\lambda) = \begin{pmatrix} x_1(\lambda) \\ \vdots \\ x_r(\lambda) \\ \lambda_1 \\ \vdots \\ \lambda_{n-r} \end{pmatrix}$

und also $x(\lambda) \neq x(\tilde{\lambda})$ für $\lambda \neq \tilde{\lambda}$.

Surjektiv: Sei $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \text{Lös}(A, b)$. Setze $\lambda = \begin{pmatrix} x_{i+1} \\ \vdots \\ x_n \end{pmatrix}$. Dann:

$$x_{r+i}(\lambda) = \lambda_i = x_{r+i} \forall i = 1, \dots, n - r$$

Für $i = r, \dots, 1$ folgt induktiv (Induktionsannahme: $x_j(\lambda) = x_j \forall j \geq i + 1$):

$$x_1(\lambda) = \frac{1}{a_{ii}}(b_i - \sum_{j=i+1}^n a_{ij}x_j(\lambda)) = \frac{1}{a_{ij}}(b_i - \sum_{j=i+1}^n a_{ij}x_j) = x_i \square$$

Fall 2:

Löse $Ax = b$, wobei A in allgemeiner Form.

Strategie: Finde ein Äquivalentes System $\tilde{A}x = \tilde{b}$ (also mit $\text{Lös}(A, b) = \text{Lös}(\tilde{A}, \tilde{b})$), wobei \tilde{A} Zeilenstufenform hat.

Dazu: Unter **elementaren Zeilenumformungen** einer Matrix verstehen wir (vorläufig):

- 1) Das Vertauschen von zwei Zeilen
- 2) Die Addition des λ -fachen der i-ten Zeile zur j-ten Zeile ($i \neq j, \lambda \in \mathbb{R} \setminus \{0\}$)

1.4 Satz

Sei (A, b) die erweiterte Koeffizientenmatrix zu $Ax = b$ und (\tilde{A}, \tilde{b}) gehe aus (A, b) durch endlich viele Zeilenumformungen hervor. Dann:

$$\text{Lös}(A, b) = \text{Lös}(\tilde{A}, \tilde{b})$$

Beweis:

Es reicht zu zeigen, dass sich die Lösungsmenge bei einer elementaren Zeilenumformung nicht ändert.

Typ1: Da alle Gleichungen in (L) simultan erfüllt sein müssen, ist die Reihenfolge der Gleichungen egal.

Typ2: Da nur die Zahlen i und j betroffen sind, reicht es zu zeigen, dass die Systeme

$$\left. \begin{array}{ccccccc} a_{i1}x_1 & + & \cdots & + & a_{in}x_n & = & b_i \\ a_{j1}x_1 & + & \cdots & + & a_{jn}x_n & = & b_j \end{array} \right\} (*)$$

und

$$\left. \begin{array}{ccccccc} a_{i1}x_1 & + & \cdots & + & a_{in}x_n & = & b_i \\ (a_{j1} + \lambda a_{i1})x_1 & + & \cdots & + & (a_{jn} + \lambda a_{in})x_n & = & b_j + \lambda b_i \end{array} \right\} (**)$$

Löst $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ das System $(*)$, so auch die erste Gleichung in $(**)$ und

durch Addition des λ -fachen der ersten Gleichung in $(*)$ zur zweiten auch die zweite Gleichung in $(**)$.

Umgekehrt schließt man analog durch Subtraktion des λ -fachen der 1. Gleichung in $(**)$ von der zweiten. \square

1.5 Satz

Jede $m \times n$ -Matrix kann durch endlich viele elementare Zeilenumformungen in Zeilenstufenform überführt werden. Zum Beweis geben wir einen Algorithmus an, der das Geforderte leistet:

1.6 Algorithmus

Input: $m \times n$ -Matrix A

Output: $m \times n$ -Matrix in Zeilenstufenform

1. setze $A_1 = A$, $p_1 = m$, $q_1 = n$
2. Wende so lange die Funktion „Zeilenstufenschritt“ an, d.h. setze
 $(A_k, p_k, q_k) := \text{Zeilenstufenschritt}(A_{k-1}, p_{k-1}, q_{k-1})$ bis $\min(p_k, q_k) = 0$
3. Nenne den Abbruchindex k_0 und gebe A_{k_0} aus.

Hierbei ist:

Zeilenstufenschritt

$$(\mathbb{R}^{m \times n} \times \{0, \dots, m\} \times \{0, \dots, n\} \rightarrow \mathbb{R}^{m \times n} \times \{0, \dots, m\} \times \{0, \dots, n\})$$

Wie folgt definiert:

$\text{Zeilenstufenschritt}(A, p, q) := (\tilde{A}, \tilde{p}, \tilde{q})$, wobei:

1. Sei $B = (b_{ij})_{i=1, \dots, p, j=1, \dots, q}$ mit $b_{ij} := a_{m-p+i, n-q+j}$
 Veranschaulichung:

$$A = \left(\begin{array}{ccccc} a_{11} & \cdots & a_{1, n-p} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m-q, 1} & \cdots & a_{m-q, n-p} & \cdots & a_{m-q, n} \\ \vdots & \ddots & \vdots & \boxed{B} & \\ a_{m1} & \cdots & a_{m, n-p} & & \end{array} \right)$$

2. Ist $b_{ij} = 0$ für alle $i = 1, \dots, p$ und $j = 1, \dots, q$, setze $\tilde{p} = \tilde{q} = 0$ und $\tilde{A} = A$
- 3.1 Andernfalls seien: $j_1 = \min\{j | b_{ij} \neq 0 \text{ für ein } i = 1, \dots, p\}$ (die Spalte von B mit minimalem Index, in der ein von 0 verschiedener Eintrag ist.)
 und $i_1 = \min\{i | b_{ij} \neq 0\}$ (die Zeile mit minimalem Index, in der j_1 -ten Spalte ein von 0 verschiedener Eintrag steht)
- 3.2 Setze $\tilde{B} = (\tilde{b}_{ij})_{i=1, \dots, p, j=1, \dots, q}$, wobei $\tilde{b}_{1j} := b_{i_1, j} \forall j = 1, \dots, q$ und für $i = 2, \dots, p$ und $j = 1, \dots, q$:

$$\tilde{b}_{ij} = \begin{cases} -\frac{b_{ij_1}}{\tilde{b}_{1j_1}} \tilde{b}_{1j} + b_{ij} & i \neq i_1 \\ -\frac{b_{1j_1}}{\tilde{b}_{1j_1}} \tilde{b}_{1j} + b_{1j} & i = i_1 \end{cases}$$

(Hier wird erst die i_1 -te mit der ersten Zeile vertauscht und dann ein geeignetes Vielfaches der neuen ersten Zeile zu den anderen Zeilen addiert. \tilde{B} geht also durch elementare Zeilenumformungen aus B hervor.)

3.3 Setze $\tilde{p} = p - 1, q = \tilde{q} - j_1$

$$\tilde{A} = \left(\begin{array}{cccc|c} a_{11} & \cdots & a_{1,n-p} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m-q,1} & \cdots & a_{m-q,n-p} & \cdots & a_{m-q,n} \\ \vdots & \ddots & \vdots & & \\ a_{m1} & \cdots & a_{m,n-p} & & \end{array} \right) \begin{array}{c} \\ \\ \tilde{B} \\ \end{array}$$

Beachte, die Matrix \tilde{B} hat die Form

$$\tilde{B} = \left(\begin{array}{cccc|ccc} 0 & \cdots & 0 & b_{i_1,j_1} & * & \cdots & * \\ \vdots & \ddots & \vdots & 0 & \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{array} \right) \quad * : \text{beliebige Einträge, } b_{i_1,j_1} \neq 0$$

1.7 Beispiel

$$\begin{aligned} A &= \left(\begin{array}{ccccc} 0 & 0 & 1 & 2 & 9 \\ 0 & 3 & 4 & 5 & 9 \\ 0 & 6 & 7 & 8 & 9 \\ 0 & 9 & 9 & 9 & 9 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccccc} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 6 & 7 & 8 & 9 \\ 0 & 9 & 9 & 9 & 9 \end{array} \right) \rightsquigarrow \left(\begin{array}{c|ccc} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 0 & -1 & -2 & -9 \\ 0 & 0 & -3 & -6 & -18 \end{array} \right) = A_2 \\ &\rightsquigarrow \left(\begin{array}{c|ccc} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 9 \end{array} \right) = A_3 \rightsquigarrow \left(\begin{array}{c|ccc} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) = A_4 \end{aligned}$$

Wir zeigen: Algorithmus 1.6 führt jede $m \times n$ -Matrix A mit endlich vielen elementaren Zeilenumformungen in Zeilenstufenform über.

Beweis: Da $\min(p_k, q_k) < \min(p_{k-1}, q_{k-1})$, bricht der Algorithmus nach endlich vielen Schritten ab. Ist $A_1 = A$ die Nullmatrix, so ist A bereits in Zeilenstufenform und der Algorithmus liefert A als Output. Andernfalls entsteht A_2 aus A_1 durch endlich viele elementare Zeilenumformungen und hat die Form

$$A_2 = \left(\begin{array}{cccc|c} 0 & \cdots & 0 & (*) & * & \cdots & * \\ \vdots & \ddots & \vdots & 0 & & & \\ \vdots & \ddots & \vdots & \vdots & & & \\ 0 & \cdots & 0 & 0 & & & \end{array} \right) \begin{array}{c} \\ \\ B_2 \\ \end{array}$$

Der Eintrag $(*) \neq 0$ steht in der $m - p_2 = 1$ -ten Zeile und $n - q_2$ -ten Spalte. Also ist B_2 eine $p_2 \times q_2$ -Matrix.

Ist B_2 die Nullmatrix, so ist A_2 in Zeilenstufenform und der Algorithmus gibt A_2 aus.

1.8 Algorithmus (Gaußsche Eliminationsverfahren zur Lösung von $Ax = b$)

Andernfalls wird B_2 im nächsten Schritt mit endlich vielen Elementaren Zeilenumformungen auf die Form (1.1) gebracht. Dehnt man diese Umformungen auf die Zeilen 2 bis m von A_2 aus, ändern sich die ersten $n - q_2$ Spalten von A_2 nicht, da dort nur Nullen stehen.

A_2 wird also in diesem Schritt mit endlich vielen elementaren Zeilenumformungen auf die Form

$$A_3 = \begin{pmatrix} 0 & \cdots & 0 & (*) & * & \cdots & \cdots & \cdots & \cdots & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & 0 & \cdots & 0 & (*) & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots & 0 & \boxed{} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \boxed{B_3} \end{pmatrix}$$

gebracht. Induktiv überzeugt man sich leicht, dass der Algorithmus das Geforderte leistet. \square

Diese Überlegungen führen zu:

1.8 Algorithmus (Gaußsche Eliminationsverfahren zur Lösung von $Ax = b$)

Input: $m \times n$ -Matrix A , $b \in \mathbb{R}^m$

Output: „es ex. keine Lösung“ od. $(0 \leq r \leq n$ und eine Bijektion $\Phi_{A,b} : \mathbb{R}^{n-r} \rightarrow \text{Lös}(A, b)$)

1. Bringe die erweiterte Koeffizientenmatrix (A, b) in Zeilenstufenform (\tilde{A}, \tilde{b}) mit Algorithmus 1.6
2. Bestimme den Zeilenrang r von \tilde{A} (m minus die Anzahl der Nullstellen von \tilde{A})
3. Tausche die Spalten von \tilde{A} (falls nötig) so, dass die Pivots in den ersten r Spalten stehen. (Wir bezeichnen die entstehende Matrix weiterhin mit \tilde{A})
4. Verwende Algorithmus 1.3 mit Input \tilde{A}, \tilde{b}, r und setzen im Existenzfall $\Phi_{(A,b)} := \Phi_{(\tilde{A}, \tilde{b})}$ (wobei ggf. die Umnummerierung der Unbekannten aus Schritt 3 zu beachten ist).
Die Korrektheit des Algorithmus folgt aus Satz 1.4.

1.9 Folgerung

Die $\text{Lös}(A, b)$ für $Ax = b$ ist entweder leer oder einelementig oder überabzählbar unendlich.

1.10 Beispiel (Rundungsfehler im Gauß-Algorithmus)

Zu lösen ist:

$$\begin{pmatrix} 10^{-4} & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Der Gaußalgorithmus liefert:

$$\left(\begin{array}{cc|c} 10^{-4} & 1 & 1 \\ 1 & 1 & 2 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|c} 10^{-4} & 1 & 1 \\ 0 & -9999 & -9998 \end{array} \right)$$

Also:

$$x_2 = \frac{9998}{9999} \approx 0,9999$$

$$x_1 = 10^4(1 - 1 * \frac{9998}{9999}) \approx 1,0001$$

Führen wir nun den Algorithmus mit der Rechengenauigkeit von 3 Dezimalstellen aus, so:

$$\left(\begin{array}{cc|c} 1,00 * 10^{-4} & 1,00 & 1,00 \\ 1,00 & 1,00 & 2,00 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|c} 1,00 * 10^{-4} & 1,00 & 1,00 \\ 0 & 1,00 * 10^4 & 1,00 * 10^4 \end{array} \right)$$

Also: $x_2 = 1, x_1 = 10^4(1 - 1) = 0$, was stark von der exakten Lösung abweicht.

Problem:

Der erste Pivot 10^{-4} , durch den zu Beginn des Algorithmus geteilt wird, ist „sehr klein“ und führt zu einem „großen“ Rundungsfehler, durch im Lauf des Algorithmus fortsetzen kann.

Ausweg:

Vertausche die Zeilen im Verlauf des Gauß-Algorithmus so, dass in der jeweiligen Spalte immer ein betragsmäßig möglichst großer Pivot entsteht (Gauß-Algorithmus mit Teilpivotierung)

In dieser Form ist der Gauß-Algorithmus „stabiler“, also weniger anfällig gegen Rundungsfehler.

Im Beispiel mit Rundungen auf 3 Dezimalstellen:

$$\left(\begin{array}{cc|c} 1,00 * 10^{-4} & 1,00 & 1,00 \\ 1,00 & 1,00 & 2,00 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|c} 1,00 & 1,00 & 2,00 \\ 1,00 * 10^{-4} & 1,00 & 1,00 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|c} 1,00 & 1,00 & 2,00 \\ 0 & 1,00 & 1,00 \end{array} \right)$$

da $1 - 10^{-4} = 0,9999$ und $1 - 2 * 10^{-4} = 0,9998$ auf 3 Stellen gerundet gleich 1 ist, also $x_2 = 1,00$ und $x_1 = 1,00$, was der Rundung der korrekten Lösung auf 3 Stellen entspricht.

Formal bedeutet die Teilpivotierung:

Verwende in „Zeilenstufenschritt“ (Alg. 1.6) in Schritt 3.1:

$$i_1 := \min\{i \mid |b_{i,j_1}| \geq |b_{k,j_1}| \forall k = 1, \dots, p\}$$

Rechenaufwand des Gauß-Algorithmus

Wir bestimmen die Anzahl der Multiplikationen, die im Verlauf der Gauß-Algorithmus durchgeführt werden, im Fall $m = n = r$.

- Umformen in Zeilenstufenform:

Die erste Zeile von (A, b) auf Zeilenstufenform zu bringen, benötigt

$$\begin{array}{ccccccc} (n-1) & * & 1 & + & (n-1) & * & n & = & (n-1) * (n+1) \\ \uparrow & & \uparrow & & \uparrow & & \uparrow & & \\ \text{Zeilen} & & \text{Faktor} & & \text{Zeilen} & & \text{Spalten} & & \end{array}$$

Multiplikationen.

Iteration führt zu:

$$\sum_{k=1}^n (k-1)(k+1) = \sum_{k=1}^n (k^2 - 1) = \frac{1}{3}n^3 + \frac{1}{2}n^2 - \frac{5}{6}n$$

Multiplikationen.

- Rückwärtseinsetzen:

Da der Zeilenrang von \tilde{A} gleich n ist, so werden zur Berechnung von x_k

$$(n - k + 1)$$

Multiplikationen benötigt, also insgesamt:

$$\sum_{k=1}^n (n - k + 1) = \sum_{k=1}^n k = \frac{(n+1)n}{2}$$

Multiplikationen.

- Gesamtaufwand: $\frac{1}{3}n^3 + O(n^2)$ Multiplikationen

Ist n sehr groß (s. Beispiel 1.1), so müssen approximative Lösungsverfahren mit niedrigerem Rechenaufwand zum Einsatz gebracht werden.

2 Vektorräume

Sei $A \in \mathbb{R}^m \times n$ und $b \in \mathbb{R}^m$. Um das System $Ax = b$ weiter zu studieren, bietet es sich an, die Funktion

$$f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m, x \mapsto Ax$$

zu untersuchen. Dies werden im allgemeineren Rahmen von „Linearen Abbildungen“ zwischen „Vektorräumen“ vornehmen.

Zunächst einige Begriffe:

2.1 Definition

Eine Menge G zusammen mit einer Verknüpfung $*$, also einer Abbildung $*$: $G \times G \rightarrow G$ heißt **Gruppe**, falls

(G1) $\forall a, b, c \in G : (a * b) * c = a * (b * c)$ (assoziativ)

(G2) Es gibt:

- ein $e \in G$ mit: $\forall a \in G : e * a = a$
- für alle $a \in G$ ein $a^{-1} \in G$ mit: $a^{-1} * a = e$ (e wird **neutrales Element** genannt, a^{-1} **das inverse Element zu a**)

Gilt zudem $a * b = b * a$ für alle $a, b \in G$, so wird G **abelsch** (oder **kommutativ**) genannt.

2.2 Beispiel

- $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ sind jeweils mit der Addition eine abelsche Gruppe. $\mathbb{Q} \setminus \{0\}, \mathbb{R} \setminus \{0\}, \mathbb{C} \setminus \{0\}$ bilden mit der Multiplikation als Verknüpfung eine abelsche Gruppe.
- Die Menge aller bijektiven Funktionen von \mathbb{R} nach \mathbb{R} mit der Hintereinanderausführung \circ als Verknüpfung ist eine Gruppe, aber nicht kommutativ (Erinnerung: $(f \circ g)(x) = f(g(x))$)

2.3 Satz

Sei $(G, *)$ eine Gruppe. Dann:

- Das neutrale Element e ist eindeutig und $a * e = a \forall a \in G$

2 Vektorräume

ii) Das inverse Element a^{-1} zu $a \in G$ ist eindeutig und $a * a^{-1} = e \forall a \in G$

iii) Für alle $a, b \in G$ gilt: $(a^{-1})^{-1} = a$ und $(a * b)^{-1} = b^{-1} * a^{-1}$

iv) Für $a, b, \tilde{b} \in G$ gelten die „Kürzungsregeln“:

$$a * b = a * \tilde{b} \Rightarrow b = \tilde{b}$$

$$b * a = \tilde{b} * a \Rightarrow b = \tilde{b}$$

Beweis:

Es seien $a \in G, a^{-1}$ ein inverses Element zu a und $(a^{-1})^{-1}$ ein inverses Element zu a^{-1} .
Dann:

$$a * a^{-1} = e * (a * a^{-1}) = ((a^{-1})^{-1} * a^{-1}) * (a * a^{-1}) = (a^{-1})^{-1} * (a^{-1} * a) * a^{-1} = (a^{-1})^{-1} * (e * a^{-1})$$

$$(a^{-1})^{-1} * a^{-1} = e$$

Somit:

$$a * e = a * (a^{-1} * a) = (a * a^{-1}) * a = e * a = a$$

Sei \tilde{e} ein weiteres neutrales Element. Dann:

$$\tilde{e} * e = \tilde{e} \text{ und } \tilde{e} * e = e$$

Also $e = \tilde{e}$

Ist a' ein weiteres neutrales Element zu a . Dann:

$$a' = a' * e = a' * (a * a^{-1}) = (a' * a) * a^{-1} = e * a^{-1} = a^{-1}$$

Dies zeigt i) und ii). iii) und iv): Übung \square

Erinnerung:

2.4 Definition

Eine Menge K mit Verknüpfungen $+: K \times K \rightarrow K, *: K \times K \rightarrow K$ heißt **Körper**, falls

K1) $(K, +)$ ist eine abelsche Gruppe. Man bezeichnet das neutrale Element mit 0 und das inverse Element zu a mit $-a$.

K2) $(K \setminus \{0\}, *)$ ist eine abelsche Gruppe. Man bezeichnet das neutrale Element mit 1 und das inverse Element zu a mit a^{-1} oder $\frac{1}{a}$.

K3) Für alle $a, b, c \in K$ gilt:

$$(a + b) * c = a * c + b * c \text{ Distributivgesetz}$$

2.5 Beispiel

- i) $(\mathbb{Q}, +, *)$, $(\mathbb{R}, +, *)$, $(\mathbb{C}, +, *)$ sind Körper, $(\mathbb{Z}, +, *)$ nicht.
- ii) Ist p eine Primzahl, so bilden die Restklassen modulo p , versehen mit der modularen Addition und Multiplikation einen Körper, der mit \mathbb{Z}_p oder \mathbb{F}_p bezeichnet wird (siehe MFI1).

Erinnerung: $\mathbb{F}_2 = \{0, 1\}$ mit Addition und Multiplikation erklärt durch:

| + | 0 | 1 | * | 0 | 1 |
|---|---|---|---|---|---|
| 0 | 0 | 1 | 0 | 0 | 0 |
| 1 | 1 | 0 | 1 | 0 | 1 |

Notation: Sei K ein Körper und $n \in \mathbb{N}_0$. Dann schreiben wir $a^n := a * a * \cdots * a$,
 $a^{-n} := \left(\frac{p}{a}\right)$, $a^0 := 1$

2.6 Definition

Sei K ein Körper. Eine Menge V zusammen mit Abbildungen:

$$+ : V \times V \rightarrow V \text{ (Addition)}$$

$$* : V \times V \rightarrow V \text{ (skalare Multiplikation)}$$

wird **K-Vektorraum** (oder **Vektorraum** über K) genannt, wenn:

- V1)** $(V, +)$ ist eine abelsche Gruppe (Das neutrale Element heißt Nullvektor und wird mit 0 bezeichnet, das inverse Element zu $v \in V$ mit $-v$).
- V2)** Die skalare Multiplikation ist verträglich mit den Übrigen Verknüpfungen, d.h. für alle $\lambda, \mu \in K$ und $u, v \in V$ gilt:

$$(\lambda + \mu) * v = \lambda * v + \mu * v, \lambda * (v + w) = \lambda * v + \lambda * w$$

$$(\lambda * \mu) * v = \lambda * (\mu * v), 1 * v = v (1 \in K)$$

Achtung

$+$ bezeichnet hier sowohl die Addition im Körper, als auch die Addition im Vektorraum. Aus dem Kontext wird klar, was gemeint ist, **zum Beispiel:**

$$\begin{array}{ccccc}
 (\lambda & + & \mu) * v = \lambda * v & + & \mu * v \\
 & \uparrow & & \uparrow & \\
 & \text{Addition} & & \text{Addition} & \\
 & \text{im Körper} & & \text{im Vektorraum} &
 \end{array}$$

Analog wird $*$ für die Multiplikation im Körper und für die skalare Multiplikation verwendet und 0 kann für die neutrale Elemente bzgl. der Addition im Körper und im Vektorraum stehen.

2.7 Beispiel

Sei K ein Körper.

$$\text{i) } K^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in K \forall i = 1, \dots, n \right\}$$

Also die Menge aller Spalten der Höhe n mit Einträgen in K mit

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

,

$$\lambda * \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda * x_1 \\ \vdots \\ \lambda * x_n \end{pmatrix}$$

ist ein **K-Vektorraum**.

- ii) Für eine nichtleere Menge A sei $K^A := \{f : A \rightarrow K\}$ die Menge aller Funktionen von A nach K . Dann wird K^A zu einem K -Vektorraum mit $(f + g)(a) := f(a) + g(a)$, $(\lambda * f)(a) := \lambda * f(a)$, $a \in A$

2.8 Beispiel

Ein Datenblock aus n Bits kann als Vektor in \mathbb{F}_2 notiert werden, also Vektor der Länge n mit Einträgen $\{0, 1\}$. Aufgrund der Additionstabelle in (Beispiel 2.5) folgt:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \quad \text{wobei } z_i = \begin{cases} 0 & , v_i = w_i \\ 1 & , v_i \neq w_i \end{cases}$$

2.9 Definition

Sei V ein K -Vektorraum und $U \subset V$ nicht leer. Dann wird U ein **Untervektorraum (UVR)** von V genannt, falls

1. $\forall u, v \in U : u + v \in U$
2. $\forall \lambda \in K, u \in U : \lambda * u \in U$

2.10 Bemerkung

Sei U ein UVR von V . Schränkt man die Addition und die skalare Multiplikation auf $U \times U$ bzw. $K \times U$ ein, so ist $(U, +, *)$ ein Vektorraum. Denn : Für jeden Vektor $u \in U$ gilt: $-u = (-1) * u \in U$ und $0 = u + (-u) \in U$.

2.11 Beispiel

i) Sei $[a, b] \subset \mathbb{R}$ ein Intervall. Dann sind

$$\zeta([a, b]) = \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ stetig}\}$$

$$\zeta^1([a, b]) = \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ stetig, differenzierbar}\}$$

$$\zeta^\infty([a, b]) = \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ beliebig oft differenzierbar}\}$$

ii) Sei K ein Körper. Die Menge der **Polynomfunktionen**

$$Pol(K) = \{f : K \rightarrow K \mid f(\lambda) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_0 \lambda^0, \text{ für ein } n \in \mathbb{N}_0 \text{ und } a_0, \dots, a_n \in K\}$$

ist ein UVR des K^K .

Achtung:

Ist etwa $K = \mathbb{F}_2$, so gilt für die Polynomfunktion

$$f : \mathbb{F}_2 \rightarrow \mathbb{F}_2, \lambda \mapsto \lambda + \lambda^2 :$$

$$f(0) = 0 + 0 = 0$$

$$f(1) = 1 + 1 = 0 \text{ in } \mathbb{F}_2$$

Also ist $f : \mathbb{F}_2 \rightarrow \mathbb{F}_2$ die Nullfunktion. Polynome mit verschiedenen Koeffizienten können in endlichen Körpern die gleiche Funktion darstellen.

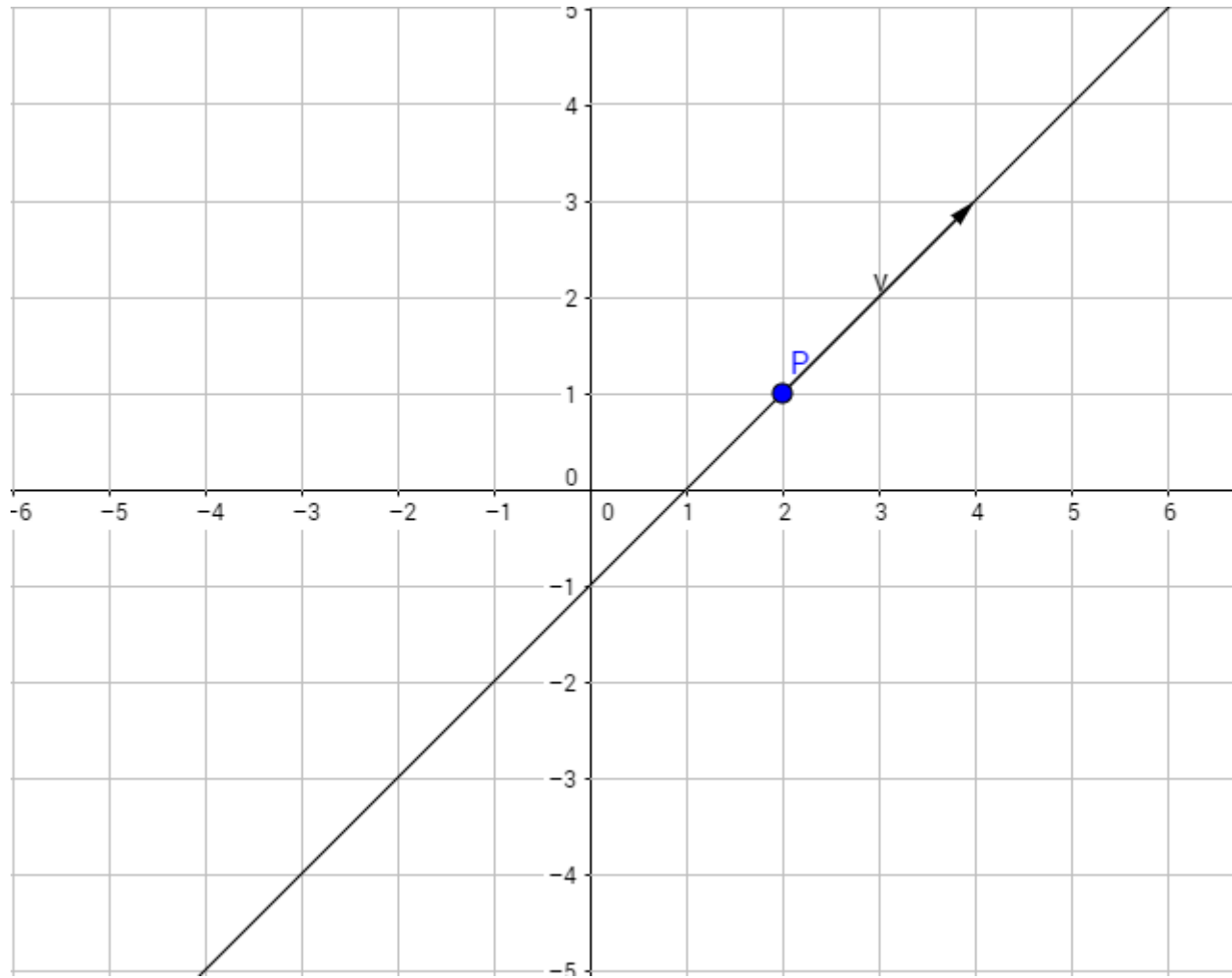
2.12 Beispiel

Im \mathbb{R}^n seien Vektoren p und $v \neq 0$ gegeben. Eine Menge der Form

$$L(p, v) := \{p + \lambda * v \mid \lambda \in \mathbb{R}\}$$

heißt **Gerade** mit **Aufpunkt** p und **Richtungsvektor** v .

Eine Gerade ist genau dann ein UVR des \mathbb{R}^n falls $0 \in l(p, v)$.



3 Lineare Unabhängigkeit und Basen

3.1 Definition

Sei V ein K -Vektorraum und $(v_i)_{i \in I}$ eine Familie von Vektoren aus V . (d.h. $I \neq \emptyset$ ist eine Menge und für alle $i \in I$ gilt: $v_i \in V$)

i) v heißt **Linearkombination** von $(v_i)_{i \in I}$, falls es ein $r \in \mathbb{N}$, $i_1, \dots, i_r \in I$ und $\lambda_1, \dots, \lambda_r \in K$ gibt mit

$$v = \sum_{k=1}^r \lambda_k v_{i_k}.$$

Die Menge aller Linearkombinationen von $(v_i)_{i \in I}$ wird mit $\text{span}((v_i)_{i \in I})$ notiert und heißt **Spann** von $(v_i)_{i \in I}$.

Konvention: $\text{span}(\emptyset) = \{0\}$

ii) $(v_i)_{i \in I}$ wird **Erzeugendensystem** von V genannt, falls $\text{span}((v_i)_{i \in I}) = V$

iii) V heißt **endlich erzeugt**, falls es ein $r \in \mathbb{N}_0$ und $v_1, \dots, v_r \in V$ so, dass (v_1, \dots, v_r) ein Erzeugendensystem von V ist.

3.2 Bemerkung

Sei $U := \text{span}((v_i)_{i \in I})$. Dann ist U ein UVR von V . Denn sind $u_1, u_2 \in U$, so existiert $r \in \mathbb{N}$, $i_1, \dots, i_r \in I$ und $\lambda_k^{(1)}, \lambda_k^{(2)} \in K$ ($k = 1, \dots, r$) mit:

$$u_1 = \sum_{k=1}^r \lambda_k^{(1)} v_{i_k}, u_2 = \sum_{k=1}^r \lambda_k^{(2)} v_{i_k}$$

Somit folgt für alle $\mu_1, \mu_2 \in K$:

$$\mu_1 * u_1 + \mu_2 * u_2 = \sum_{k=1}^r \underbrace{(\mu_1 \lambda_k^{(1)} + \mu_2 \lambda_k^{(2)})}_{\in K} * v_{i_k} \in U$$

3.3 Beispiel

i) Für einen Körper K ist K^n endlich erzeugt. Sei:

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{-te Position,}$$

i-te Einheitsvektor im K^n ($i=1, \dots, n$). Dann ist (e_1, \dots, e_n) ein Erzeugendensystem des K^n , denn für alle $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ gilt:

$$x = \sum_{i=1}^n x_i * e_i$$

ii) Der \mathbb{R} -VR der Polynomfunktion $\text{Pol}(\mathbb{R})$ ist nicht endlich erzeugt. Denn seien $p_1, \dots, p_n \in \text{Pol}(\mathbb{R})$ und $\lambda_1, \dots, \lambda_n$ reelle Zahlen. Sei k der maximale Grad dieser n -Polynome. Dann hat

$$p := \sum_{i=1}^n \lambda_i p_i$$

höchstens Grad k . Also ist (p_1, \dots, p_n) kein Erzeugendensystem von $\text{Pol}(\mathbb{R})$.

3.4 Definition

Sei V ein K -VR.

i) Eine endliche Familie von Vektoren (v_1, \dots, v_n) wird **Linear unabhängig** genannt, falls für alle $\lambda_1, \dots, \lambda_n \in K$ gilt:

$$\sum_{i=1}^n \lambda_i v_i = 0 \in V \Rightarrow \forall_{i=1, \dots, n} \lambda_i = 0 \in K$$

ii) Eine beliebige Familie von Vektoren $(v_i)_{i \in I}$ wird **linear unabhängig** genannt, falls jede endliche Teilfamilie linear unabhängig ist (d.h. für alle $J \subset I$ endlich gilt: $(v_j)_{j \in J}$ ist linear unabhängig). Andernfalls heißt $(v_i)_{i \in I}$ **linear abhängig**.

iii) Unter einer **Basis** von V verstehen wir ein linear unabhängiges Erzeugendensystem von V .

3.5 Bemerkung

Lineare Abhängigkeit bedeutet also, dass man aus der Familie $(v_i)_{i \in I}$ den Nullvektor durch eine nicht-triviale Linearkombination darstellen kann (d.h. durch eine Linearkombination, bei der nicht alle Koeffizienten gleich 0 sind).

Die Familie

$$\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} \right) \text{ im } \mathbb{R}^3$$

ist also linear abhängig, denn

$$-2 * \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-2) * \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 1 * \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

3.6 Beispiele

i) Die Familie $(e_i)_{i=1, \dots, n}$ im K^n ist eine Basis. Dann seien $\lambda_1, \dots, \lambda_n \in K$ mit

$$\sum_{i=1}^n \lambda_i e_i = 0 \in K^n$$

Da $\sum_{i=1}^n \lambda_i e_i = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$, folgt $\lambda_i = 0 \in K$ für alle $i=1, \dots, n$. Also ist (e_1, \dots, e_n)

linear unabhängig. Nach Beispiel 3.3 ist (e_1, \dots, e_n) auch ein Erzeugendensystem, also eine Basis.

(e_1, \dots, e_n) wird **kanonische Basis** in K^n genannt.

ii) Die Familie $\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right)$ ist eine Basis im \mathbb{R}^3

Lineare Unabhängigkeit:

$$\lambda_1 * \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 * \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \lambda_3 * \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow 3 * \lambda_3 = 0 \Rightarrow \lambda_3 = 0$$

$$\Rightarrow \lambda_1 * \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 * \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \lambda_2 = 0 \Rightarrow \lambda_1 = 0.$$

3 Lineare Unabhängigkeit und Basen

Erzeugendensystem: Sei $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$. Dann $\lambda_1 * \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 * \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \lambda_3 * \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ für $\lambda_3 = \frac{x_3}{3}, \lambda_2 = \frac{x_2}{2} - \frac{x_3}{3}, \lambda_1 = x_1 - \frac{x_2}{2}$.

iii) Seien $q_n(t) = t^n, n \in \mathbb{N}_0$. Dann ist $(q_n)_{n \in \mathbb{N}_0}$ eine Basis von $\text{Pol}(\mathbb{R})$.

Lineare Unabhängigkeit: Sei $n \in \mathbb{N}$ beliebig. Ist $\lambda_i \neq 0$ für ein $i = 0, \dots, n$, dann hat

$$\left(\sum_{i=0}^n q_i \right) (t) = \sum_{i=0}^n \lambda_i t^i$$

höchstens n Nullstellen. Also ist $\sum_{i=0}^n \lambda_i q_i \neq 0 \in \text{Pol}(\mathbb{R})$.

Erzeugendensystem: Sei $p \in \text{Pol}(\mathbb{R})$. Dann

$$p(t) = \sum_{k=0}^n \lambda_k t^k \text{ für ein } n \in \mathbb{N}_0, \lambda_1, \dots, \lambda_n \in \mathbb{R}.$$

$$\text{Also } p = \sum_{k=0}^n \lambda_k q_k.$$

Zur Charakterisierung von Basen:

3.7 Satz

Sei $B = (v_1, \dots, v_n)$ eine Familie von Vektoren eines K -VR V . Dann sind äquivalent:

- i) B ist Basis
- ii) B ist ein unverkürzbares Erzeugendensystem, dass heißt für jedes $r \in \{1, \dots, n\}$ ist $(v_1, \dots, v_{r-1}, v_{r+1}, \dots, v_n)$ kein Erzeugendensystem, aber B ist Erzeugendensystem.
- iii) Zu jedem Vektor $v \in V$ ex. eindeutig bestimmte $\lambda_1, \dots, \lambda_n \in K$ mit: $v = \sum_{i=1}^n \lambda_i v_i$
- iv) B ist unverlängerbar linear unabhängig, d.h. B ist linear unabhängig und für alle $v \in V$ ist (v_1, \dots, v_n, v) linear abhängig.

Beweis:

i) \Rightarrow ii): Sei B ein Erzeugendensystem. Ist B verkürzbar, so gilt:

$$v_r = \lambda_1 v_1 + \dots + \lambda_{r-1} v_{r-1} + \lambda_{r+1} v_{r+1} + \dots + \lambda_n v_n$$

Also:

$$\lambda_1 v_1 + \dots + \lambda_{r-1} v_{r-1} + \lambda_{r+1} v_{r+1} + \dots + \lambda_n v_n + (-1) * v_r = 0,$$

d.h. B ist nicht linear unabhängig und also keine Basis.

ii) \Rightarrow iii) : Sei B ein Erzeugendensystem. Ist die Eindeutigkeitsbedingung nicht erfüllt, so gibt es ein $v \in V$ mit

$$v = \lambda_1 v_1, \dots, \lambda_n v_n = \mu_1 v_1 + \dots + \mu_n v_n$$

($\lambda_r \neq \mu_r$ für mindestens ein r).

Für ein solches r gilt:

$$v_r = \frac{\mu_1 - \lambda_1}{\lambda_r - \mu_r} v_1 + \dots + \frac{\mu_{r-1} - \lambda_{r-1}}{\lambda_r - \mu_r} v_{r-1} + \frac{\mu_{r+1} - \lambda_{r+1}}{\lambda_r - \mu_r} v_{r+1} + \dots + \frac{\mu_n - \lambda_n}{\lambda_r - \mu_r} v_n$$

Also ist B verkürzbar.

iii) \Rightarrow iv) : Aus iii) mit $v = 0 \in V$ folgt, direkt die lineare Unabhängigkeit. Sei $v \in V$. Dann hat v die Darstellung:

$$v = \sum_{i=1}^n \lambda_i v_i$$

Also

$$\sum_{i=1}^n \lambda_i v_i + (-1) * v = 0,$$

d.h. (v_1, \dots, v_n, v) ist linear abhängig.

iv) \Rightarrow i) : Aus iv) folgt: Für jedes $v \in V$ existieren $\lambda_1, \dots, \lambda_n, \lambda \in K$ mit

$$\sum_{i=1}^n \lambda_i \lambda v = 0,$$

wobei nicht alle $\lambda_1, \dots, \lambda_n, \lambda$ gleich 0 sind. Da B linear unabhängig ist, muss gelten $\lambda \neq 0$. Also:

$$v = \sum_{i=1}^n \frac{-\lambda_i}{\lambda} v_i,$$

d.h. B ist auch ein Erzeugendensystem. \square

3.8 Folgerung

Jeder endlich erzeugte K-VR hat eine Basis.

Beweis:

Von einem endlichen Erzeugendensystem nehme man so lange Vektoren weg, bis es unverkürzbar ist. \square

Konvention:

Die leere Familie von Vektoren ist linear unabhängig, also ist sie Basis von $V = \{0\}$.

Frage:

Kann ein endlich erzeugter Vektorraum Basen unterschiedlicher Länge (also mit unterschiedlich vielen Vektoren) haben?

3.9 Lemma (Austauschlemma)

Sei V ein K -VR mit Basis $B = (v_1, \dots, v_n)$ und $w = \sum_{k=1}^n \lambda_k v_k \in V$.

Falls $\lambda_k \neq 0$, so ist $(v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_n) =: B'$ ebenfalls eine Basis von V .

Beweis:

Durch Umnummerierung kann man erreichen, dass $k = 1$.

- $B' = (w, v_2, \dots, v_n)$ ist Erzeugendensystem:
Sei $v \in V$. Dann: $v = \sum_{i=1}^n \mu_i v_i, \mu_1, \dots, \mu_n \in K$. Da $\lambda_1 \neq 0$, folgt:

$$v_1 = \frac{1}{\lambda_1} w - \sum_{i=2}^n \frac{\lambda_i}{\lambda_1} v_i$$

Somit:

$$v = \frac{\mu_1}{\lambda_1} w + \sum_{i=2}^n \left(\mu_i - \frac{\mu_1}{\lambda_1} \lambda_i \right) v_i$$

- B' ist linear unabhängig: Sei

$$0 = \mu * w + \sum_{i=2}^n \mu_i v_i$$

Da $w = \sum_{i=1}^n \lambda_i v_i$, folgt:

$$0 = \mu \lambda_1 v_1 + \sum_{i=2}^n (\mu_i + \mu \lambda_i) v_i$$

Da B linear unabhängig ist, folgt: $\mu \lambda_1 = 0$ und $(\mu_i + \mu \lambda_i) = 0$ für $i = 2, \dots, n$.

Da $\lambda_1 \neq 0$, gilt $\mu = 0$ und $\mu_i = 0$ für $i = 2, \dots, n$. \square

3.10 Satz

Sei V ein K -VR mit Basis $B = (v_1, \dots, v_r)$ und (w_1, \dots, w_n) eine linear unabhängige Familie in V . Dann:

i) $n \leq r$

ii) Es gibt Indizes $i_1, \dots, i_n \in \{1, \dots, r\}$, sodass die Familie B' , die entsteht, wenn man in B v_{i_1}, \dots, v_{i_n} durch w_1, \dots, w_n ersetzt weiterhin eine Basis von V ist.

Beweis:

Induktion über $n \in \mathbb{N}_0$

- $n = 0$: Hier ist nichts zu zeigen
- $(n - 1) \rightsquigarrow n$: Da auch (w_1, \dots, w_{n-1}) linear unabhängig ist, ergibt die Induktionsannahme, dass $(n - 1) \leq r$ und dass nach geeigneter Umnummerierung der Basis B $(w_1, \dots, w_{n-1}, v_n, \dots, v_r)$ Basis von V ist (Umnummerierung mit $i_1 = 1, \dots, i_{n-1} = n - 1$).

Um $n \leq r$ zu zeigen, genügt es, $n - 1 = r$ zu einem Widerspruch zu führen. In diesem Fall wäre (w_1, \dots, w_{n-1}) eine Basis von V im Widerspruch zu Satz 3.7 iv).

Ferner hat w_n die Darstellung

$$w_n = \lambda_1 w_1 + \dots + \lambda_{n-1} w_{n-1} + \lambda_n v_n + \dots + \lambda_r v_r$$

Da (w_1, \dots, w_n) linear unabhängig ist, folgt $\lambda_k \neq 0$ für ein $k = n, n + 1, \dots, r$. Also kann nach dem Austauschslemma v_k durch w_n ersetzt werden, ohne dass die Basiseigenschaft verloren geht. \square

3.11 Folgerung

Sei V ein K-VR mit Basen (v_1, \dots, v_n) und (w_1, \dots, w_r) . Dann: $n = r$.

Beweis

Der Austauschsatz liefert: $n \leq r$ und $r \leq n$. \square

3.12 Folgerung

Hat ein K-VR eine endliche Basis, so sind alle seine Basen endlich.

Beweis

Sei (v_1, \dots, v_r) eine endliche Basis und $(w_i)_{i \in I}$ eine unendliche Basis. Dann existiert $i_1, \dots, i_{r+1} \in I$ mit $(w_{i_1}, \dots, w_{i_{r+1}})$ linear unabhängig ist. Dies widerspricht dem Austauschsatz. \square

Also ist der folgende Dimensionsbegriff wohldefiniert.

3.13 Definition

Ist V ein K-VR, so definieren wir die **Dimension** von V durch:

$$\dim V = \begin{cases} \infty & , \text{ falls } V \text{ nicht endlich erzeugt} \\ r & , \text{ falls } V \text{ Basis der Länge } r \in \mathbb{N}_0 \text{ hat.} \end{cases}$$

Hierbei sei die **Länge** einer endlichen Basis die Anzahl von Vektoren in der Basis. Man schreibt auch $\dim_K V$ statt $\dim V$, falls der Körper aus dem Kontext nicht klar ist.

3.14 Beispiel

i) $\dim K^n = n$

ii) $\dim \text{Pol}(\mathbb{R}) = \infty$

3.15 Beispiel (und Erinnerung)

Auf dem \mathbb{R}^2 führen wir folgende Multiplikation ein:

$$* : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2 :$$

$$\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) \mapsto \begin{pmatrix} x_1 y_1 & - & x_2 y_2 \\ x_1 y_2 & + & x_2 y_1 \end{pmatrix}$$

Mit diesem Beispiel wird $(\mathbb{R}, +, *)$ zu einem Körper, dem Körper der **komplexen Zahlen**, \mathbb{C} . Das neutrale Element bzgl. $*$ ist $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Man notiert daher:

$$1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, i := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Jedes $z \in \mathbb{C}$ hat dann eine eindeutige Darstellung der Form:

$$z = \lambda_1 * 1 + \lambda_2 * i, \lambda_1, \lambda_2 \in \mathbb{R}$$

λ_1 heißt **Realteil** von z ($\text{Re}(z)$), λ_2 **Imaginärteil** von z ($\text{Im}(z)$). i wird **imaginäre Einheit** genannt. Es gilt:

$$i^2 = i * i = -1 \in \mathbb{C}$$

Die reellen Zahlen können als Teilmenge von \mathbb{C} aufgefasst werden, indem man $\lambda \in \mathbb{R}$ mit der komplexen Zahl $\lambda * 1 \in \mathbb{C}$ identifiziert. (Die reellen Zahlen findet man also auf der x-Achse). Man schreibt dann: $z = \lambda_1 + \lambda_2 * i$ als Kurzform für $z = \lambda_1 * 1 + \lambda_2 * i$ ($\lambda_1, \lambda_2 \in \mathbb{R}$). Nach Beispiel 3.14 gilt:

$$\dim_{\mathbb{C}} \mathbb{C} = 1$$

$$\dim_{\mathbb{R}} \mathbb{C} = \dim_{\mathbb{R}} \mathbb{R}^2 = 2$$

Einige weitere Folgerungen des Austauschsatzes:

3.16 Folgerung(Basisergänzungssatz)

Sei V ein endlich erzeugter K -VR und (v_1, \dots, v_n) linear unabhängig. Dann existiert $r \in \mathbb{N}_0$ und $v_{n+1}, \dots, v_{n+r} \in V$, sodass (v_1, \dots, v_{n+r}) eine Basis von V ist.

3.17 Folgerung

Ist U ein UVR eines endlich erzeugten K -VR V . Dann:

i) $\dim U \leq \dim V$

ii) $\dim U = \dim V \Rightarrow U = V$

4 Matrizen

Sei K ein Körper. Wir bezeichnen mit $K^{m \times n}$ die Menge aller $m \times n$ -Matrizen mit Einträgen in K , also der Rechteckschemata der Form:

$$A = (a_{ij})_{i=1, \dots, m, j=1, \dots, n} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, a_{ij} \in K.$$

Mit Addition:

$$(a_{ij})_{i=1, \dots, m, j=1, \dots, n} + (b_{ij})_{i=1, \dots, m, j=1, \dots, n} := (a_{ij} + b_{ij})_{i=1, \dots, m, j=1, \dots, n}$$

und der skalaren Multiplikation:

$$\lambda * (a_{ij}) = (\lambda * a_{ij}), \lambda \in K$$

wird $K^{m \times n}$ zu einem K -VR.

4.1 Definition

Sei $A \in K^{m \times n}$ mit Zeilen:

$$a_1 := \begin{pmatrix} a_{11} \\ \vdots \\ a_{1n} \end{pmatrix}, \dots, a_m := \begin{pmatrix} a_{m1} \\ \vdots \\ a_{mn} \end{pmatrix}$$

(als Spaltenvektor geschrieben). Dann wird $Z(A) := \text{span}(a_1, \dots, a_m) \subset K^n$ der **Zeilenraum** von A genannt.

4.2 Algorithmus

Input: $A \in K^{m \times n}$

Output: Basis B von $Z(A)$

- 1.: Bringe A auf Zeilenstufenform mit Algorithmus 1.6 und bezeichne die resultierende Matrix mit $\tilde{A} = (\tilde{a}_{ij})_{i=1, \dots, m, j=1, \dots, n}$
- 2.: Bestimme den Zeilenrang r von \tilde{A} (also die Anzahl von Zeilen von \tilde{A} , die nicht nur Nulleinträge haben).

4 Matrizen

3.: Setze für $i = 1, \dots, r$

$$b_i = \begin{pmatrix} \tilde{a}_{i1} \\ \vdots \\ \tilde{a}_{in} \end{pmatrix} \text{ (i-te Zeile von } \tilde{A} \text{ als Spaltenvektor)}$$

und gebe $B = (b_1, \dots, b_r)$ aus.

Beweis der Korrektheit von Algorithmus 4.2

i) Elementare Zeilenumformungen ändern den Zeilenraum nicht, d.h. $Z(A) = Z(\tilde{A})$. Ist etwa $v \in Z(A)$ und A_1 gehe aus A hervor, indem das λ -fache der j -ten Zeile zur i -ten Zeile addiert wird. Dann:

$$v = \sum_{k=1}^m \mu_k a_k = \left(\sum_{k \in \{1, \dots, m\} \setminus \{i, j\}} \mu_k a_k \right) + \mu_i(a_i + \lambda a_j) + (\mu_i - \lambda \mu_i) a_j$$

Also $v \in Z(A_1)$, d.h. $Z(A) \subset Z(A_1)$

Analog: $Z(A_1) \subset Z(A)$

klar: Zeilenvertauschungen ändern den Zeilenraum nicht.

ii) B ist Erzeugendensystem von $Z(A)$:

Dies folgt aus i), da B ein Erzeugendensystem von $Z(\tilde{A})$ (Weglassen der Nullzeilen ändert den Spann nicht).

iii) B ist linear unabhängig:

Sei \tilde{a}_{i, j_i} der i -te Pivot von \tilde{A} . Dann:

$$\tilde{a}_{k, j_1} = 0 \forall k > i.$$

Seien $\lambda_1, \dots, \lambda_r \in K$ mit

$$\lambda_1 b_1 + \dots + \lambda_r b_r = 0 \in K^n$$

Betrachtung des j_1 -ten Eintrags führt zu

$$\lambda_1 \underbrace{\tilde{a}_{1, j_1}}_{\neq 0} + \lambda_2 \underbrace{\tilde{a}_{2, j_1}}_{=0} + \dots + \lambda_r \underbrace{\tilde{a}_{r, j_1}}_{=0} = 0 \in K$$

Also: $\lambda_1 = 0$. Iteriert man dies, folgt leicht: $\lambda_k = 0 \forall k = 1, \dots, r$. \square

4.3 Bemerkung

Die Algorithmen 1.3, 1.6, 1.8 behalten ihre Gültigkeit, falls man \mathbb{R} durch einen beliebigen Körper K ersetzt. Spezielle Eigenschaften wurden dort nicht verwendet.

4.4 Bemerkung

Man nennt die Dimension des Zeilenraums einer Matrix $A \in K^{m \times n}$ den **Zeilenrang** von A. Für Matrizen in Zeilenstufenform stimmt dies mit der Begriffsbildung aus Definition 1.2 überein. Die Zahl r im Output von Algorithmus 1.8 ist also der Zeilenrang von A, da elementare Zeilenumformungen den Zeilenrang nicht ändern.

4.5 Beispiel

Sei

$$U = \text{span} \left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 9 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 4 \\ 5 \\ 9 \end{pmatrix}, \begin{pmatrix} 0 \\ 6 \\ 7 \\ 8 \\ 9 \end{pmatrix}, \begin{pmatrix} 0 \\ 9 \\ 9 \\ 9 \\ 9 \end{pmatrix} \right) \subset \mathbb{R}^5$$

Wir betrachten:

$$A = \begin{pmatrix} 0 & 0 & 1 & 2 & 9 \\ 0 & 3 & 4 & 5 & 9 \\ 0 & 6 & 7 & 8 & 9 \\ 0 & 9 & 9 & 9 & 9 \end{pmatrix}$$

Algorithmus 1.6 liefert(s. Beispiel 1.7):

$$\tilde{A} = \begin{pmatrix} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Also ist $\dim(Z(A)) = 3$ und

$$U = \text{span} \left(\begin{pmatrix} 0 \\ 3 \\ 4 \\ 5 \\ 9 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 9 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 9 \end{pmatrix} \right) \text{ ist Basis von } Z(A).$$

4.6 Definition

Seien $A = (a_{ij}) \in K^{m \times n}$ und $B = (b_{jk}) \in K^{n \times r}$ (d.h. A hat so viele Spalten, wie B Zeilen). Dann ist das Produkt von A und B definiert durch

$$A \cdot B := C \in K^{m \times r}, C = (c_{ik})_{i=1, \dots, m, k=1, \dots, r}$$

$$c_{ik} := \sum_{j=1}^n a_{ij} b_{jk}$$

4 Matrizen

$A \cdot B$ hat also so viele Zeilen wie A und so viele Spalten wie B.

$$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{i1} & a_{i2} & \cdot & \cdot & a_{in} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot & b_{1k} & \cdot \\ \cdot & b_{2k} & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & b_{rk} & \cdot \end{pmatrix} = \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & a_{ik} & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$$

4.7 Bemerkung

Die Matrixmultiplikation ist nicht kommutativ, selbst dann nicht, wenn AB und BA definiert werden können (d.h. $A \in K^{m \times n}, b \in K^{n \times m}$).

Gegenbeispiel

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$
$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Wir bezeichnen mit E_n die n-reihige Einheitsmatrix, d.h.:

$$E_n := \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in K^{n \times n}$$

Die folgenden Rechenregeln können elementar überprüft werden:

4.8 Satz

Seien $A, A' \in K^{m \times n}, B, B' \in K^{n \times r}, C \in K^{r \times s}$ und $\lambda \in K$, dann:

- i) $A \cdot (B + B') = A \cdot B + A \cdot B'$ und $(A + A') \cdot B = A \cdot B + A' \cdot B$ (Distributivgesetze)
- ii) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ (Assoziativgesetz)
- iii) $E_m \cdot A = A \cdot E_n$
- iv) $A \cdot (\lambda B) = (\lambda A) \cdot B = \lambda(A \cdot B)$

4.9 Bemerkung

- i) Man schreibt auch AB , statt $A \cdot B$

- ii) Eine Matrix mit n Zeilen und 1 Spalte ist ein Spaltenvektor der Höhe n . Man schreibt K^n , statt $K^n \times 1$. Als Spezialfall von Definition 4.6 erhält man:

$$A * x = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}, x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$$

Analog werden Elemente von $K^{1 \times n}$ werden **Zeilenvektoren** genannt

4.10 Satz

Sei $A \in K^{m \times n}$ und $b \in K^m$. Dann sind äquivalent:

- i) $\text{Lös}(A, b) \neq \emptyset$, d.h. es existiert eine Lösung zu $Ax = b$
- ii) $\text{Zeilenrang}(A) = \text{Zeilenrang}((A, b))$

Beweis

Wir verwenden den Algorithmus 1.8 und Bemerkung 4.3 und 4.4.

Sei $r := \text{Zeilenrang}(\tilde{A})$. Da $\text{Zeilenrang}(A) = \text{Zeilenrang}((A, b)) \Leftrightarrow \text{Zeilenrang}(\tilde{A}) = \text{Zeilenrang}((\tilde{A}, \tilde{b})) \Leftrightarrow$ Die Gleichungen $r + 1, \dots, m$ von $\tilde{A}x = \tilde{b}$ lauten $0 = 0 \Leftrightarrow$ Algorithmus 1.8 liefert mindestens eine Lösung zu $Ax = b$.

4.11 Satz

Sei $A \in K^{m \times n}$, $b \in K^m$. Dann sind äquivalent:

- i) $\text{Lös}(A, b)$ hat höchstens ein Element, d.h. es gibt höchstens eine Lösung $x \in K^n$ zu $Ax = b$
- ii) $\text{Zeilenrang}(A) = n$

Beweis

Nach Bemerkung 4.4 folgt: $\text{Zeilenrang}(A) = n \Leftrightarrow$ Algorithmus 1.8 liefert „keine Lösung“ oder Bijektion $\Phi_{A, B} : \mathbb{R}^{n-r} \rightarrow \text{Lös}(A, b) \Rightarrow \text{Lös}(A, b)$ höchstens 1 Element.

4.12 Definition

Eine Matrix $A \in K^{n \times n}$ wird **invertierbar** genannt, falls ein $B \in K^{n \times m}$ existiert mit $B \cdot A = E_n$. Wir bezeichnen mit $GL(n, k)$ die Menge aller invertierbaren Matrizen in $k^{m \times n}$ ($GL \Leftrightarrow$ „general linear“).

4.13 Satz

Sei $A \in K^{n \times n}$. Dann sind äquivalent:

- i) $\forall b \in K^n Ax = b$ hat genau eine Lösung $x \in K^n$
- ii) $\text{Lös}(A, 0) = \{0\}$
- iii) $\text{Zeilenrang}(A) = n$
- iv) Die Spalten von A bilden eine Basis in K^n
- v) A ist invertierbar, d.h. ex. $B \in K^{n \times n}$ mit $B \cdot A = E_n$
 In diesem Fall gilt: B ist eindeutig bestimmt und $B \in GL(n, K)$. Man nennt B die **Inverse** von A und schreibt A^{-1} statt B.

Zur Vorbereitung:

4.14 Definition

Sei $A = (a_{ij}) \in K^{m \times n}$. Dann wird die durch :

$$A^T = (a'_{ji})_{j=1, \dots, n, i=1, \dots, m}, (a_{ij}) =: (a'_{ji})$$

definierte Matrix $\in K^{n \times m}$ die **Transponierte** von A genannt. Hier wird also die i-te Zeile von A zur i-ten Spalte von A^T .

Man prüft leicht: $\forall A \in K^{m \times n}, B \in K^{n \times r}$

$$(A^T)^T = A \text{ und } (AB)^T = B^T A^T$$

4.15 Bemerkung

Unter dem Spaltenraum von $A \in K^{m \times n}$ versteht man den Spalten einer Matrix aufgespannten U-VR von K^m . Dies ist nach Definition identisch mit dem Zeilenraum von A^T . Seine Dimension wird **Spaltenrang** von A genannt.

Bedingung iv) in Satz 4.13 ist äquivalent zu

iv') $\text{Spaltenrang}(A) = n$

Später sehen wir: Es gilt stets: $\text{Zeilenrang}(A) = \text{Spaltenrang}(A)$.

Beweis von Satz 4.13: Wir zeigen $iii) \Rightarrow ii) \Rightarrow iv) \Rightarrow v) \Rightarrow iii)$ und $i) \Leftrightarrow iv)$

iii) \Rightarrow ii) Nach Satz 4.11 hat $\text{Lös}(A, 0)$ höchstens ein Element. Da stets $A \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, folgt: $\text{Lös}(A, 0) = \{0\}$

ii) \Rightarrow iv) Sei a_i die i-te Spalte von A, $x \in K^n$.

Dann $Ax = \sum_{i=1}^n x_i a_i$

ii)

$\text{Lös}(A, b) = \{0\} \Leftrightarrow Ax = 0$ gilt nur für $x_1 = \dots = x_n = 0$

$\Leftrightarrow x_1 a_1 + \dots + x_n a_n = 0$ nur für $x_1 = \dots = x_n = 0$

$\Leftrightarrow a_1, \dots, a_n$ sind linear unabhängig.

Da alle Basen des K^n die Länge n haben (Folgerung 3.11) ist (a_1, \dots, a_n) nach Satz 3.10 unverlängerbar linear unabhängig, also eine Basis von K^n (Satz 3.7)

iv) \Leftrightarrow i) Da $Ax = b \Leftrightarrow x_1 a_1 + \dots + x_n a_n = b$, liefert Satz 3.7 die Äquivalenz von iv) und i) (vgl. Folien).

iv) \Rightarrow v) : Nach Bem 4.13 ist dann $\text{Zeilenrang}(A^T) = n$ und so ist auch $\text{Zeilenrang}(A^T|b) = n \forall b \in K^n$.

Satz 4.10 liefert: $\forall b \in K^n$ existiert eine Lösung $x \in K^n$ zu $A^T x = b$. Für $i = 1, \dots, n$

wähle $b := e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{-te Zeile. Dann existieren } c_i \in K^n \text{ mit } A^T c_i = e_i. \text{ Die}$

Matrix C mit Spalten c_i erfüllt dann

$$A^T \cdot C = E_n$$

Sei $B := C^T$, dann gilt: $B \cdot A = ((B \cdot A)^T)^T = ((C^T \cdot A)^T)^T = (A^T \cdot C)^T = (E_n)^T = E_n \Rightarrow A$ ist invertierbar.

v) \Rightarrow iii) Sei C wie oben, d.h. $C = B^T$, dann: $A^T C = A^T B^T = (BA)^T = E_n^T = E_n$

$$(A^T C)x = E_n x$$

$$\begin{array}{ccc} \parallel & & \parallel \\ A^T(Cx) & = & x \quad \forall x \in K^n \text{ ist darstellbar als lineare Kombination von Spalten} \\ \downarrow & & \end{array}$$

der A^T (mit Koeffizienten (Cx)). Somit erzeugen die Spalten von A^T und also die Zeilen von A den K^n . Da $\dim K^n = n$ und A n Zeilen hat, bilden die Zeilen von A ein unverkürzbares Erzeugendensystem, also eine Basis von K^n (Satz 3.7). Somit ist $\text{Zeilenrang}(A) = n$.

Zusatz: Sei $A \in K^{n \times n}$ invertierbar, d.h. $\exists B \in K^{n \times n}$ mit $B \cdot A = E_n$. Wir zeigen, dass B eindeutig bestimmt und invertierbar ist. Sei $B \cdot A = E_n$. Dann gilt für alle $x \in K^n$: $(BA)x = E_n x \Leftrightarrow B(Ax) = x$. Also erzeugen die Spalten von B den K^n . Da $\dim K^n = n$ und B n Spalten hat, bilden die Spalten von B ein unverkürzbares Erzeugendensystem, und also eine Basis von K^n . Nach Satz 4.13(iv) \Leftrightarrow v)) folgt, dass B invertierbar ist. Die Eindeutigkeit von B ergibt sich aus Satz 2.3 und dem folgenden Satz:

4.16 Satz

$GL(n, K)$ bildet versehen mit der Matrixmultiplikation eine Gruppe.

Beweis

- Seien $A, B \in GL(n, K)$. Dann existieren $A', B' \in GL(n, K)$ mit $A'A = E_n, B'B = E_n$. Ferner, $(B'A')AB = B' \underbrace{(A'A)}_{=E_n} B = B'B = E_n$. Also: $AB \in GL(n, K)$.
- Assoziativität: Folgt aus Satz 4.8.
- Neutrales Element: E_n
- Inverses Element: $A \in GL(n, K) \Rightarrow \exists A' \in GL(n, K)$ mit $A'A = E_n \Rightarrow A'$ ist Inverse zu A . \square

4.17 Bemerkung

- i) Ist A invertierbar, so gilt nach Satz 2.3: $A^{-1}A = AA^{-1} = E_n$
- ii) Ist A invertierbar, so ist auch A^T invertierbar. Und es gilt: $(A^T)^{-1} = (A^{-1})^T$. Denn: $(A^{-1})^T A^T = (AA^{-1})^T = (E_n)^T = E_n$

Zur Berechnung der Inversen:

Nach Bem 4.17 i) und Satz 4.13 gilt: Ist A invertierbar, so $\exists A^{-1}$ mit $AA^{-1} = E_n$.

Sei $A^{-1} = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}$, dann für jedes $i = 1, \dots, n$ ist $x_{\cdot i} := \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix}$ die eindeutige Lösung zu $Ax = e_i$.

Man geht also wie folgt vor:

- i) Betrachte die erweiterte $n \times (2n)$ Matrix $(A | \underbrace{E_n}_{e_1 \dots e_n})$ und bringe diese mit Algorithmus 1.6 auf Zeilenstufenform: $(A | E_n) \rightsquigarrow (\tilde{A} | \tilde{B})$

ii) Ist Zeilenrang von $\tilde{A} < n$ (d.h. hat \tilde{A} eine Nullzeile), so ist A nicht invertierbar.

iii) Andernfalls hat $(\tilde{A}|\tilde{B})$ die Form:

$$\left(\begin{array}{ccc|c} \tilde{a}_{11} & \cdots & \cdots & \tilde{b}_1 \\ \tilde{a}_{22} & \cdots & \cdots & \tilde{b}_2 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \tilde{a}_{nn} & \tilde{b}_n \end{array} \right) \quad \text{mit } \tilde{a}_{ii} \neq 0 \forall i = 1, \dots, n.$$

iv) Eliminiere sukzessive für $i = n, \dots, 1$ durch Addition eines geeigneten Vielfaches der i -ten Zeile zu den darüberliegenden Zeilen die Einträge oberhalb von \tilde{a}_{ii} . Dies führt zu:

$$\left(\begin{array}{ccc|c} \tilde{a}_{11} & \cdots & \cdots & \tilde{b}_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \tilde{a}_{nn} & \tilde{b}_n \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} \tilde{a}_{11} & \cdots & 0 & \tilde{b}_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \tilde{a}_{nn} & \tilde{b}_n \end{array} \right)$$

iv) Dann erhält man A^{-1} indem man die i -te Zeile von \tilde{B} mit \tilde{a}_{ii}^{-1} multipliziert:

$$\left(\begin{array}{ccc|c} \tilde{a}_{11} & \cdots & 0 & \tilde{b}_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \tilde{a}_{nn} & \tilde{b}_n \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & \cdots & 0 & \tilde{b}_1 \tilde{a}_{11}^{-1} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & \tilde{b}_n \tilde{a}_{nn}^{-1} \end{array} \right)$$

Fazit: $(A|E_n) \rightsquigarrow (E_n|A^{-1})$
elementare Zeilenumformung

4.18 Beispiel

Zu berechnen ist (im Fall der Existenz) A^{-1} zur Matrix

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

$$\left(\begin{array}{ccc|ccc} 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array} \right)$$

\rightsquigarrow

$$\begin{array}{c}
\left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array} \right) \\
\rightsquigarrow \\
\left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 & -1 & 1 \end{array} \right) \\
\rightsquigarrow \\
\left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & -1 & 1 \end{array} \right) \\
\rightsquigarrow \\
\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & -1 & 2 & -1 \\ 0 & 1 & 0 & -3 & 4 & -4 \\ 0 & 0 & -1 & 1 & -1 & 1 \end{array} \right) \\
\rightsquigarrow \\
\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & -6 & 7 \\ 0 & 1 & 0 & -3 & 4 & -4 \\ 0 & 0 & -1 & 1 & -1 & 1 \end{array} \right)
\end{array}$$

Also:

$$A^{-1} = \begin{pmatrix} 5 & -6 & 7 \\ -3 & 4 & -4 \\ -1 & 1 & -1 \end{pmatrix}$$

4.19 Bemerkung

Üblicherweise betrachtet man drei Typen von elementaren Zeilenumformungen einer Matrix $A \in K^{m \times n}$:

Typ I: Multiplikation einer Zeile mit $\lambda \in K \setminus \{0\}$

Typ II: Addition des λ -fachen der i -ten Zeile zur j -ten Zeile ($i \neq j, \lambda \in K \setminus \{0\}$)

Typ III: Vertauschen der i -ten und j -ten Zeile ($i \neq j$)

Das Verfahren zur Invertierung von Matrizen zeigt:

Jede Matrix $A \in GL(n, K)$ kann durch endlich viele elementare Zeilenumformungen vom Typ I-III in die n -reihige Einheitsmatrix E_n überführt werden.

4.20 Bemerkung

Ist $A \in GL(n, K)$ und ist A^{-1} bereits berechnet, so kann für jede rechte Seite $b \in K^n$ die eindeutige Lösung x zu $A * x = b$ durch Matrixmultiplikation berechnet werden:

$$x = (A^{-1} * A)x = A^{-1}(Ax) = (A^{-1}b)$$

5 Lineare Abbildungen

5.1 Definition

Eine Abbildung $F : V \rightarrow W$ zwischen K -VR V und W wird **linear** oder (Vektorraum-) **Homomorphismus** genannt, falls

- i) $\forall v \in V, \lambda \in K : F(\lambda v) = \lambda * F(v)$
- ii) $\forall v_1, v_2 \in V : F(v_1 + v_2) = F(v_1) + F(v_2)$

Ist F bijektiv, so wird F ein **Isomorphismus** genannt.

5.2 Beispiel

- i) Ist $A \in K^{m \times n}$, so ist $F_A : \mathbb{R}^n \rightarrow \mathbb{R}^m, x \mapsto A * x$ linear. Es gilt: $F_A(e_j) = A * e_j =$

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}, j = 1, \dots, n, \text{ d.h.}$$

die Spalten von A sind die Bilder der Einheitsvektoren unter F_A .

Falls $m = n$, so folgt leicht aus Satz 4.13:

F_A bijektiv $\Leftrightarrow F_A$ injektiv $\Leftrightarrow F_A$ surjektiv.

- ii) Sei $[a, b] \subset \mathbb{R}$ ein Intervall $(-\infty < a < b < \infty)$. Dann sind:

I: $C([a, b]) \rightarrow C^1([a, b]) : f \mapsto (x \mapsto \int_a^x f(y) dy)$

D: $C^1([a, b]) \rightarrow C([a, b]) : f \mapsto f'$

linear. Nach dem Hauptsatz der Differential- und Integralrechnung gilt: $D \circ I = id_{C([a, b])}$, sodass I injektiv und D surjektiv ist. Da für konstante Funktionen $G \neq 0$ gilt: $(I \circ D)(g) = 0$ sind weder I noch D Isomorphismen.

5.3 Beispiel

Ein Datenblock bestehend aus n Bits soll übertragen werden. Zur Fehlererkennung bei der Datenübertragung wird bei der Paritätsprüfung ein weiteres Bit, das sogenannte Paritätsbit an den Datenblock angehängt (vor der Datenübertragung). Es ist:

1 , falls der Datenblock eine gerade Zahl von Einsen hat,

0 , sonst.

5 Lineare Abbildungen

Dieses Vorgehen führt zu einer linearen Abbildung:

$$F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^{n+1} : x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_1 + x_2 + \dots + x_n \end{pmatrix}$$

Dann hat $F(x)$ stets eine gerade Zahl von Einsen, sodass auf jeden Fall ein Übertragungsfehler vorliegt, wenn der empfangene Datenblock eine ungerade Zahl von Einsen hat.

In Matrixschreibweise gilt: $F = F_A$ mit $A = \begin{pmatrix} E_n & \\ 1 & \dots & 1 \end{pmatrix}$

5.4 sehr einfacher Satz

Sei $F : V \rightarrow W$ linear. Dann:

- i) $F(0) = 0$
- ii) $F(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i F(v_i)$ ($\lambda_1, \dots, \lambda_n \in K, v_1, \dots, v_n \in V$)
- iii) Sind $V' \subset V, W' \subset W$ UVR, so sind $F(V') \subset W$ und $F^{-1}(W') \subset V$ UVR.
- iv) Ist $(v_i)_{i \in I}$ linear abhängig in V , so ist $(F(v_i))_{i \in I}$ linear abhängig in W .
- v) Ist F ein Isomorphismus, so ist F^{-1} linear.
- vi) Ist $G : W \rightarrow U$ linear, so ist $G \circ F$ linear.

Beweis

i)-iv),vi) sind eine gute Übung.

- v) Seien $w, w' \in W, \lambda, \mu \in K$. Seien $v := F^{-1}(w), v' := F^{-1}(w')$. Dann: $\lambda w + \mu w' = \lambda F(v) + \mu F(v') = F(\lambda v + \mu v')$
Anwenden der Umkehrfunktion liefert:
 $F^{-1}(\lambda w + \mu w') = \lambda v + \mu v' = \lambda F^{-1}(w) + \mu F^{-1}(w') \square$

5.5 Definition (Endomorphismus und Automorphismus)

Ist $F : V \rightarrow V$ linear, so wird F ein **Endomorphismus** genannt. Bijektive Endomorphismen heißen **Automorphismen**.

5.6 Folgerung

Sei V ein K-VR und $\text{Aut}(V)$ die Menge alle Automorphismen auf dem V . Dann wird $\text{Aut}(V)$ mit der Hintereinanderausführung zu einer Gruppe.

5.7 Definition (Bild, Kern, Rang und Defekt)

Sei $F : V \rightarrow W$ linear. Dann heißt $Im(F) := F(V)$ das **Bild** von F und $Ker(F) := F^{-1}(\{0\})$ der **Kern** von F .

Beides sind UVR nach Satz 5.4. Man nennt $dim(Im(F))$ den **Rang** von F und $dim(Ker(F))$ den **Defekt** von F .

5.8 Satz

Sei $F : V \rightarrow W$ linear. Dann:

$$F \text{ ist injektiv} \Leftrightarrow Ker(F) = \{0\}$$

In diesem Fall gilt: Ist $(v_i)_{i \in I}$ linear unabhängig in V , so ist $(F(v_i))_{i \in I}$ linear unabhängig in W .

Beweis

„ \Rightarrow “ klar

„ \Leftarrow “ Seien $v, v' \in V$ mit $F(v) = F(v')$. Dann $F(v - v') = 0$, d.h. $v - v' \in Ker(F)$.
Somit: $v = v'$.

Zusatz:

Sei F injektiv und $(v_i)_{i \in I}$ linear unabhängig. Falls für eine Linearkombination

$$\sum_{j=1}^n \lambda_j F(v_{i_j}) = 0$$

, so

$$F\left(\sum_{j=1}^n \lambda_j v_{i_j}\right) = 0$$

Also:

$$\sum_{j=1}^n \lambda_j v_{i_j} = 0,$$

da $Ker(F) = \{0\}$. Somit $\lambda_j = 0$ für alle $j = 1, \dots, n$, da $(v_i)_{i \in I}$ linear unabhängig. \square

5.9 Satz

Sei $F : V \rightarrow W$ linear und $dim(V) < \infty$. Sind dann Basen (v_1, \dots, v_k) von $Ker(F)$ und (w_1, \dots, w_r) in $Im(F)$ gegeben. Ferner seien $u_1 \in F^{-1}(\{w_1\}), \dots, u_r \in F^{-1}(\{w_r\})$. Dann ist

$$A = (u_1, \dots, u_r, v_1, \dots, v_k)$$

5 Lineare Abbildungen

eine Basis von V . Insbesondere:

$$\dim(V) = \dim(\operatorname{Im}(F)) + \dim(\operatorname{Ker}(F)) \quad (\text{Dimensionsformel})$$

Beweis

Für $v \in V$ seien $\mu_1, \dots, \mu_r \in K$ gegeben mit

$$F(v) = \mu_1 w_1 + \dots + \mu_r w_r.$$

Sei $v' := \mu_1 u_1 + \dots + \mu_r u_r$. Da $F(v') = F(v)$, folgt: $v - v' \in \operatorname{Ker}(F)$

Also: $v - v' = \lambda_1 v_1 + \dots + \lambda_k v_k$ ($\lambda_1, \dots, \lambda_k \in K$).

Zusammen: $v = (v - v') + v' = \lambda_1 v_1 + \dots + \lambda_k v_k + \mu_1 u_1 + \dots + \mu_r u_r$.

Also ist A erzeugend.

Ist $\mu_1 u_1 + \dots + \mu_r u_r + \lambda_1 v_1 + \dots + \lambda_k v_k = 0$ für $\mu_1, \dots, \mu_r, \lambda_1, \dots, \lambda_k \in K$, so folgt durch Anwenden von F : $\mu_1 w_1 + \dots + \mu_r w_r = 0$.

Da (w_1, \dots, w_r) linear unabhängig, folgt $\mu_1 = \mu_2 = \dots = \mu_r = 0$.

Also: $\lambda_1 v_1 + \dots + \lambda_k v_k = 0$.

Da (v_1, \dots, v_k) linear unabhängig, folgt: $\lambda_1 = \dots = \lambda_k = 0$. Also ist A linear unabhängig. \square

5.10 Folgerung

Seien V und W endlich-dimensionale K -VR und $F : V \rightarrow W$ ein Isomorphismus. Dann:

$$\dim(V) = \dim(W)$$

.

Anwendung auf lineare Gleichungssysteme:

Sei $A \in K^{m \times n}$. Wir betrachten das sogenannte **homogene Problem**

$$Ax = 0 \in K^m.$$

Da $\operatorname{Lös}(A, 0) = \operatorname{Ker}(F_A)$, ist $\operatorname{Lös}(A, 0)$ ein UVR des \mathbb{R}^n .

Sei $\Phi_{A,0}$ die Abbildung aus Algorithmus 1.8. Induktiv überzeugt man sich, dass $c_{ik} \in K$ existieren, sodass mit $r := \operatorname{Zeilenrang}(\tilde{A})$

$$x_i(\lambda) = \frac{1}{\tilde{a}_{ii}} \left(- \sum_{j=i+1}^n \tilde{a}_{ij} x_j(\lambda) \right) = \sum_{k=1}^{n-r} c_{ik} \lambda_k$$

Also ist $\Phi_{A,0} : \mathbb{R}^{n-r} \rightarrow \operatorname{Lös}(A, 0)$ in Algorithmus 1.8 linear und nach dem dort Gezeigten bijektiv. Folgerung 5.10 impliziert:

$$\dim(\operatorname{Lös}(A, 0)) = n - \operatorname{Zeilenrang}(A)$$

Andererseits impliziert die Dimensionsformel

$$n = \text{Rang}(F_A) + \dim(\text{Lös}(A, 0))$$

Da für alle $\lambda_1, \dots, \lambda_n \in K$ gilt:

$$F_A\left(\sum_{j=1}^n \lambda_j e_j\right) = \sum_{j=1}^n \lambda_j A * e_j = \sum_{j=1}^n \lambda_j \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix},$$

folgt: Das Bild von F_A ist der Spaltenraum von A .

Zusammen ergibt sich:

5.11 Satz

Sei $A \in K^{m \times n}$. Dann:

$$\text{Rang}(F_A) = \text{Spaltenrang}(A) = \text{Zeilenrang}(A)$$

und $\text{Lös}(A, 0)$ ist ein UVR des \mathbb{R}^n der Dimension $n - \text{Rang}(F_A)$

Wir betrachten nun das **inhomogene Problem** $Ax = b, b \in K^m \setminus \{0\}$.

5.12 Lemma

Sei $x \in \text{Lös}(A, b)$. Dann:

$$y \in \text{Lös}(A, b) \Leftrightarrow \exists z \in \text{Lös}(A, 0) : y = x + z$$

Beweis

$$„\Rightarrow“ \quad A(x - y) = Ax - Ay = b - b = 0$$

$$„\Leftarrow“ \quad Ay = A(x + z) = Ax + Az = b + 0 = b \quad \square$$

5.13 Definition (affiner Unterraum)

Sei U ein UVR eines K -VR V und $v \in V$. Eine Menge der Form

$$v + U := \{v + u | u \in U\}$$

wird ein **affiner Unterraum** von V genannt.

5.14 Bemerkung

Ist $v+U = v'+U'$, so gilt $U = U'$ und $v' \in v+U$. Insbesondere ist $\dim(v+U) := \dim(U)$ wohldefiniert.

5.15 Satz

Seien $A \in K^{m \times n}$ und $b \in K^m \setminus \{0\}$. Dann ist $\text{Lös}(A, b)$ entweder leer oder ein affiner Unterraum der Dimension $n - \text{Rang}(F_A)$, nämlich

$$\text{Lös}(A, b) = x + \text{Lös}(A, 0),$$

wobei $x \in \text{Lös}(A, b)$ beliebig.

Beweis

Lemma 5.12 und Satz 5.11. \square

5.16 Bemerkung

Die Gesamtheit der Lösung des inhomogenen Problems findet man, indem man zunächst eine Lösung zu $Ax = b$ bestimmt (spezielle Lösung) und dann alle Lösungen des homogenen Problems bestimmt.

5.17 Beispiel

Sei $a \in K^n \setminus \{0\}$ und $b \in K$. Eine Menge der Form

$$H(a, b) = \{x \in K^n \mid a^T * x = b\}$$

wird **Hyperebene** genannt.

Da $\text{Rang}(F_{a^T}) = 1$, ist $H(a, b)$ ein affiner Unterraum der Dimension $n - 1$ im K^n . Ist $K^n = \mathbb{R}^3$, so sind die Hyperebenen gerade die Ebenen im „üblichen“ Sinn.

6 Koordinatentransformationen

6.1 Satz Koordinatensystem

Sei V ein K -VR mit Basis $B = (v_1, \dots, v_n)$. Dann existiert genau ein Isomorphismus $\Phi_B : K^n \rightarrow V$ mit $\Phi_B(e_j) = v_j$ für $j = 1, \dots, n$. Man nennt Φ_B das durch B bestimmte **Koordinatensystem** in V und $\Phi_B^{-1}(v)$ die **Koordinaten** von v .

Beweis

Sei $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$. Dann impliziert die Linearität:

$$\Phi_B(x) = \Phi_B\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n x_j \Phi_B(e_j) = \sum_{j=1}^n x_j v_j$$

Also ist Φ_B eindeutig bestimmt.

Da (v_1, \dots, v_n) eine Basis von V ist, existiert zu jedem $v \in V$ genau ein $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$

mit $v = \sum_{j=1}^n x_j v_j = \Phi_B(x)$. Also ist Φ_B bijektiv. \square

6.2 Bemerkung

Zwei K -VR U und V werden **isomorph** genannt, falls ein Isomorphismus $F : U \rightarrow V$ existiert.

Satz 6.1 und Folgerung 5.10 zeigen: Zwei endlichdimensionale K -VR sind isomorph genau dann, wenn sie die gleiche Dimension haben.

6.3 Beispiel (Diskrete Kosinustransformation zur Datenkomprimierung)

Ein Vektor $v \in \mathbb{R}^n$ soll in komprimierter Form gespeichert werden.
Eine Möglichkeit: Transformation in den Frequenzbereich. Dazu sei

$$v_k := \begin{pmatrix} \cos\left(\frac{0,5}{n}(k-1)\pi\right) \\ \cos\left(\frac{1,5}{n}(k-1)\pi\right) \\ \vdots \\ \cos\left(\left(1 - \frac{0,5}{n}\right)(k-1)\pi\right) \end{pmatrix}$$

Man kann zeigen, dass $B = (v_1, \dots, v_n)$ eine Basis des \mathbb{R}^n bilden. Sie besteht aus Kosinusschwingungen zu unterschiedlichen Frequenzen, die an den Punkten $\frac{j+0,5}{n}$ abgetastet

werden. Seien $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ die Koordinaten von $v \in \mathbb{R}^n$ im durch B bestimmten Koordinatensystem, das heißt $v = \sum_{k=1}^n x_k v_k$. In einem Quantisierungsschritt setzt man für ein vorgegebenes $\sigma > 0$:

$$\tilde{x}_k = \begin{cases} x_k & , |x_k| \geq \sigma \\ 0 & , |x_k| < \sigma \end{cases}$$

und speichert nur $\tilde{x}_1, \dots, \tilde{x}_{i_{\max}}$, wobei i_{\max} der größte Index ist, für den $\tilde{x}_{i_{\max}} \neq 0$. Die Rekonstruktion

$$\tilde{v} = \sum_{k=1}^{i_{\max}} \tilde{x}_k v_k$$

ist besser, je kleiner σ ist. Die Einsparung bei Speicherplatz ist höher, je größer σ ist. Ein ähnliches Verfahren für Matrizen statt Vektoren kommt bei JPEG zum Einsatz.

6.4 Bemerkung

Sei $B = (b_1, \dots, b_n)$ eine Basis im K^n . Dann hat das Koordinatensystem Φ_B die Darstellung:

$$\Phi_B(x) = T_B * x$$

wobei die Vektoren b_1, \dots, b_n die Spalten von $T_B \in K^{n \times n}$ sind. Die Matrix T_B ist dann invertierbar nach Satz 4.13 und es gilt:

$$\Phi_B^{-1}(v) = T_B^{-1} * v, v \in K^n.$$

Man erhält also die Koordinaten durch Multiplikation mit T_B^{-1} .

6.5 sehr einfaches Beispiel

Sei $b_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, b_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$. Dann ist $B = (b_1, b_2)$ eine Basis im \mathbb{R}^2 . Mit dem Verfahren zu Matrixinvertierung aus Kapitel 4 berechnet man leicht:

$$T_B^{-1} = \begin{pmatrix} \frac{3}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{pmatrix}$$

Also gilt für jeden Vektor $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$:

$$v = \left(\frac{3}{5}v_1 - \frac{1}{5}v_2 \right) * b_1 + \left(-\frac{1}{5}v_1 + \frac{2}{5}v_2 \right) * b_2$$

6.6 Satz

Gegeben sei K-VRe V mit Basis $A = (v_1, \dots, v_n)$ und W mit Basis $B = (w_1, \dots, w_m)$. Dann gibt es zu jeder linearen Abbildung $F : V \rightarrow W$ genau eine Matrix $M_B^A(F) = (a_{ij}) \in K^{m \times n}$ mit

$$F(v_j) = \sum_{i=1}^m a_{ij} w_i, j = 1, \dots, n(*)$$

Ferner gilt:

$$(\Phi_B^{-1} \circ F \circ \Phi_A)(x) = M_B^A(F) * x, x \in K^n(**)$$

Man nennt $M_B^A(F)$ die **darstellende Matrix** von F bzgl. A und B .

Beweis

Da B eine Basis von W ist, sind die Spalten von $M_B^A(F)$ durch $(*)$ eindeutig bestimmt. Aufgrund der Linearität reicht es, $(**)$ auf den kanonischen Einheitsvektor zu prüfen. Es gilt:

$$\begin{aligned} \Phi_B^{-1} \circ F \circ \Phi_A(e_j) &= \Phi_B^{-1}(F(v_j)) = \Phi_B^{-1}\left(\sum_{i=1}^m a_{ij} w_i\right) \\ &= \sum_{i=1}^m a_{ij} \Phi_B^{-1}(w_i) = \sum_{i=1}^m a_{ij} e_i \\ &= \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = M_B^A(F) * e_j. \square \end{aligned}$$

Frage:

Wie wählt man die Basen A, B so, dass die darstellende Matrix besonders „einfach“ ist.

6.7 Satz

Sei $F : V \rightarrow W$ linear, dann $\dim(V) = n$, $\dim(W) = m$ und r der Rang von F . Dann gibt es Basen A von V und B von W , sodass

$$M_B^A(F) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

Beweis:

Eine Basis (w_1, \dots, w_r) von $\text{Im}(F)$ werde auf irgendeine Weise zu einer Basis $B = (w_1, \dots, w_m)$ von W ergänzt. Ferner seien $v_i \in F^{-1}(\{w_i\})$ für $i = 1, \dots, r$ und (v_{r+1}, \dots, v_n) eine Basis von $\text{Ker}(F)$. Nach Satz 5.9 ist $A = (v_1, \dots, v_n)$ eine Basis von V . Es gilt:

$$F(v_j) = \begin{cases} w_j & , j = 1, \dots, r \\ 0 & , j = r+1, \dots, n \end{cases}$$

Also $M_B^A(F) = (a_{ij})$ mit

$$a_{ij} = \begin{cases} 1 & , 0 \leq i = j \leq r \\ 0 & , \text{sonst} \end{cases} \quad \square$$

6.8 Bemerkung

Seien $A, b \in K^{m \times n}$. Dann wird A **äquivalent** zu B genannt, falls $T \in \text{GL}(n, k), S \in \text{GL}(m, k)$ existieren mit

$$B = S * A * T^{-1}.$$

Wendet man Satz 6.7 auf $F_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ an und setzt man $T = T_{\mathcal{A}}^{-1}, S = T_{\mathcal{B}}^{-1}$ so gilt mit Satz 6.6 und Bemerkung 6.4:

$$SAT^{-1} = T_{\mathcal{B}}^{-1} A T_{\mathcal{A}}^{-1} = M_{\mathcal{B}}^{\mathcal{A}}(F_A) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

wobei r der Rang von A ist. Also ist A äquivalent zu $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$. Man nennt die die Normalform von A (bzgl. Äquivalenz).

6.9 Beispiel

Sei $A = \begin{pmatrix} 1 & 2 \\ 2 & 2 \\ 0 & 1 \end{pmatrix}$. Gesucht sind Basen \mathcal{A} im \mathbb{R}^2 und \mathcal{B} im \mathbb{R}^3 , sodass $M_{\mathcal{B}}^{\mathcal{A}}(F_A)$ Normalform (also die Form aus Satz 6.7) hat.

Da $\text{Im}(F_A) = ZR(A^T)$, kann man mit Algorithmus 4.2 eine Basis von $\text{Im}(F_A)$ berechnen.

$$A^T = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{pmatrix}$$

Also ist $\left(\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right)$ eine Basis vom $\text{Im}(F_A)$

Diese ergänzen wir zu einer Basis des \mathbb{R}^3

$$\mathcal{B} = \left(\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

Da $\text{Rang}(F_A) = 2$, ist $\text{Ker}(F_A) = \{0\}$. Die Basis \mathcal{A} im \mathbb{R}^2 besteht also aus Lösungen

$$v_1, v_2 \text{ zu } A * v_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, A * v_2 = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$$

Der Gaußalgorithmus liefert:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & 2 & 2 & -2 \\ 0 & 1 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & 0 & -2 \\ 0 & 1 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$$\text{Also: } \mathcal{A} = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right)$$

7 Eigenwertproblem

Frage:

Sei V ein n -dim K -VR und $F : V \rightarrow V$ ein Endomorphismus. Wie findet man eine Basis \mathcal{B} von V , sodass $M_{\mathcal{B}}^{\mathcal{B}}(F)$ eine „möglichst einfache“ Darstellung hat?

Diese (schwierige) Frage führt zu:

7.1 Definition

Sei $F : V \rightarrow V$ ein Endomorphismus. $\lambda \in K$ wird **Eigenwert** zu F genannt, falls ein $v \in V \setminus \{0\}$ existiert mit

$$F(v) = \lambda * v$$

Jedes solche v wird ein **Eigenvektor** von F zum Eigenwert λ genannt.

Mit $\text{Eig}(F, \lambda) := \{v \in V | F(v) = \lambda v\} = \text{Ker}(F - \lambda \text{id}_V)$ bezeichnen wir den **Eigenraum** von F zu $\lambda \in K$. Hier ist $\text{id}_V : V \rightarrow V, v \mapsto v$.

7.2 Bemerkung

λ Eigenwert zu $F \Leftrightarrow \text{Eig}(F, \lambda) \neq \{0\}$. In diesem Fall ist $\text{Eig}(F, \lambda) \setminus \{0\}$ die Menge aller Eigenvektoren zum Eigenwert λ .

7.3 Definition

Ein Endomorphismus $F : V \rightarrow V$ wird **diagonalisierbar** genannt, falls eine Basis

$\mathcal{B} = (v_1, \dots, v_n)$ und $\lambda_1, \dots, \lambda_n \in K$ existieren mit $M_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$

7.4 Satz

Sei V ein n -dim K -VR und $F : V \rightarrow V$ linear. Dann:

$$F \text{ ist diagonalisierbar} \Leftrightarrow \text{Es ex. eine Basis aus Eigenvektoren zu } F \text{ in } V$$

Beweis

Sei F diagonalisierbar und für eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ und $\lambda_1, \dots, \lambda_n \in K$ gelte:

$$M_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Dann:

$$F(v_j) = F(\Phi_{\mathcal{B}}(e_j)) = \Phi_{\mathcal{B}}(\underbrace{M_{\mathcal{B}}^{\mathcal{B}}(F) * e_j}_{=\lambda_j e_j}) = \lambda_j \Phi_{\mathcal{B}}(e_j) = \lambda_j v_j$$

Also ist v_j ein Eigenvektor zum Eigenwert λ_j .

Ist andererseits $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis. Eigenvektoren zu (nicht notwendigerweise verschiedenen) Eigenwerten $\lambda_1, \dots, \lambda_n$.

Dann:

$$M_{\mathcal{B}}^{\mathcal{B}}(F) * e_j = (\Phi_{\mathcal{B}}^{-1} \circ F \circ \Phi_{\mathcal{B}})(e_j) = \Phi_{\mathcal{B}}^{-1}(\underbrace{(F(v_j))}_{=\lambda_j v_j}) = \lambda_j \Phi_{\mathcal{B}}^{-1}(v_j) = \lambda_j e_j.$$

Also hat $M_{\mathcal{B}}^{\mathcal{B}}(F)$ Diagonalform \square

7.5 Beispiel

Für $\alpha \in \{-1, 0, 1\}$ sei $A = \begin{pmatrix} 0 & \alpha \\ 1 & 0 \end{pmatrix}$.

Frage: Ist $F_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2, x \mapsto Ax$ diagonalisierbar?

λ ist Eigenwert zu $F_A \Leftrightarrow \text{Lös}(A - \lambda E_2, 0) \neq \{0\}$

Mit dem Gauß-Algorithmus berechnen wir:

$$\left(\begin{array}{cc|c} -\lambda & \alpha & 0 \\ 1 & -\lambda & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & -\lambda & 0 \\ -\lambda & \alpha & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & -\lambda & 0 \\ 0 & -\lambda^2 + \alpha & 0 \end{array} \right)$$

Also Lös: $(A - \lambda E_2, 0) \neq \{0\} \Leftrightarrow \lambda^2 = \alpha$.

- $\alpha = 1$: $\lambda_1 = 1$ und $\lambda_2 = -1$ sind Eigenwerte mit Eigenvektoren $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Mit $\mathcal{B} = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$ folgt: $M_{\mathcal{B}}^{\mathcal{B}}(F_A) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

- $\alpha = -1$: $\lambda^2 = -1$ hat keine Lösung in \mathbb{R} . Also ist $F_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ nicht diagonalisierbar. Wir betrachten: $\mathbb{C}^2 \rightarrow \mathbb{C}^2$. Dann lösen $\lambda_1 = i$ und $\lambda_2 = -i$ die Gleichung $\lambda^2 = -1$. Zugehörige Eigenvektoren sind: $v_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$. Also folgt mit $\mathcal{B} = \left(\begin{pmatrix} i \\ 1 \end{pmatrix}, \begin{pmatrix} -i \\ 1 \end{pmatrix} \right)$, dass $M_{\mathcal{B}}^{\mathcal{B}}(F_A) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. Also ist F_A über \mathbb{C} diagonalisierbar.

- $\alpha = 0$: Hier ist $\lambda = 0$ einziger Eigenwert und $\text{Eig}(F_A, 0) = \text{span}(e_2)$ hat Dimension 1. Also existiert weder über \mathbb{R} noch über \mathbb{C} eine Basis aus Eigenvektoren, dass heißt F_A ist nicht diagonalisierbar.

7.6 Bemerkung

Eigenwerte kann man als Nullstellen des sog charakteristischen Polynoms finden. Dazu später mehr.

7.7 Satz

Sei $F : V \rightarrow V$ linear und seien v_1, \dots, v_n Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_n$. Dann ist (v_1, \dots, v_n) linear unabhängig. Insbesondere gilt: $n \leq \dim(V)$.

Beweis

Induktion über n:

n=1: klar, da $v_1 \neq 0$

n-1 \rightsquigarrow n: Sei

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$$

für $\alpha_1, \dots, \alpha_n \in K$. Dann folgt durch Anwenden von F bzw Multiplikation λ_n :

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 \dots + \alpha_n \lambda_n v_n = 0$$

$$\alpha_1 \lambda_n v_1 + \alpha_2 \lambda_n v_2 \dots + \alpha_n \lambda_n v_n = 0$$

Also:

$$\alpha_1 (\lambda_1 - \lambda_n) v_1 + \alpha_2 (\lambda_2 - \lambda_n) v_2 + \dots + \alpha_{n-1} (\lambda_{n-1} - \lambda_n) v_{n-1}$$

Nach Induktionsannahme ist (v_1, \dots, v_{n-1}) linear unabhängig. Da $\lambda_j \neq \lambda_n$ für alle $j = 1, \dots, n-1$, folgt: $\alpha_1 = \alpha_2 = \dots = \alpha_{n-1} = 0$. Da $v_n \neq 0$ folgt $\alpha_n = 0$. \square

7.8 Folgerung

$F : V \rightarrow V$ linear, habe n paarweise verschiedene Eigenwerte, wobei $n = \dim(V)$. Dann ist F diagonalisierbar.

Beweis

Satz 7.7 und Satz 7.4 \square

7.9 Bemerkung

- i) Ist $A \in K^{n \times n}$, so nennt man die Eigenwerte (bzw. -vektoren) von F_A auch Eigenwerte (bzw. -vektoren) von A .
- ii) $A, C \in K^{n \times n}$ heißen **ähnlich**, falls ein $B \in \text{GL}(n, K)$ existiert mit $A = BCB^{-1}$. A wird diagonalisierbar genannt, falls A ähnlich zu einer Diagonalmatrix

$$C = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

ist (oder anders formuliert, wenn F_A diagonalisierbar ist).

7.10 Satz

Sei $A \in K^{n \times n}$ und C sei ähnlich zu A . Dann haben A, A^T, C die selben Eigenwerte.

Man beachte:

λ ist Eigenwert zu $A \Leftrightarrow A - \lambda E_n$ nicht invertierbar.

Beweis

$$\begin{aligned} A - \lambda E_n \text{ invertierbar} &\Leftrightarrow (A - \lambda E_n)^T \text{ invertierbar} \\ &\Leftrightarrow A^T - (\lambda E_n)^T \text{ invertierbar} \\ &\Leftrightarrow A^T - \lambda E_n \text{ invertierbar} \end{aligned}$$

Also: λ Eigenwert zu $A \Leftrightarrow \lambda$ Eigenwert zu A^T .

Sei nun $A = BCB^{-1}$ für ein $B \in \text{GL}(n, k)$.

Dann:

$$A - \lambda E_n = BCB^{-1} - \lambda E_n = BCB^{-1} - B\lambda E_n B^{-1} = B(C - \lambda E_n)B^{-1}$$

Also:

$(A - \lambda E_n)v = 0 \Leftrightarrow (C - \lambda E_n)(B^{-1}v) = 0$, d.h. v ist Eigenvektor von A zum Eigenwert λ genau dann, wenn $B^{-1}v$ Eigenvektor von C zum Eigenwert λ ist. \square

Im Fall $K = \mathbb{C}$ kann man zeigen:

7.11 Satz (Jordansche Normalform)

Jede Matrix $A \in \mathbb{C}^{n \times n}$ ist ähnlich zu einer Matrix der Form:

$$\begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_p \end{pmatrix}$$

wobei jeder „Jordanblock“ J_i eine quadratische Matrix der Form:

$$J_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}$$

ist, wobei λ_i (nicht notwendigerweise verschiedene) Eigenwerte von A sind.

8 Eigenwertprobleme für stochastische Matrizen

Zurück zum Page-Rank-Beispiel 1.1:

Gegeben seien Internetseiten. Wir definieren Matrizen $L = (l_{ij})_{i,j=1,\dots,n}$, $S = (s_{ij})_{i,j=1,\dots,n}$, $G = (g_{ij})_{i,j=1,\dots,n}$ durch

$$l_{ij} = \begin{cases} \frac{1}{c_j} & , \text{ falls Seite } j \text{ auf Seite } i \text{ verlinkt} \\ c_j & , \text{ sonst} \end{cases}$$

c_j = Anzahl der Seiten, auf die j verlinkt.

$$a_j = \begin{cases} 1 & , \text{ Seite } j \text{ verlinkt auf keine Seite} \\ 0 & , \text{sonst} \end{cases}$$

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$$

$$S := L + \frac{1}{n}ea^T$$

$$G := d * S + (1 - d)\frac{1}{n}ee^T$$

für einen Damping-Faktor $0 < d < 1$.

Gleichungssystem aus Bsp: 1.1 (*) ist dann gerade $(d * S - E_n)x = -\frac{1-d}{n}e$

8.1 Lemma

$x \in \mathbb{R}^n$ löst (P) genau dann, wenn x Eigenvektor von G zum Eigenwert 1 ist und $\sum_{i=1}^n x_i = 1$

Beweis:

x löse (P). Dann:

$$\sum_{i=1}^n x_i = e^T x = e^T E_n x = e^T (d * S x + \frac{1-d}{n}e)$$

8 Eigenwertprobleme für stochastische Matrizen

$$= d * e^T S + \frac{1-d}{n} e^T e$$

$$\text{Da } E e^T e = \sum_{i=1}^n 1 * 1 = n$$

$$e^T S = (\sigma_1, \dots, \sigma_n) \in \mathbb{R}^{1 \times n},$$

wobei

$$\sigma_j = \sum_{i=1}^n 1 * s_{ij}$$

die Summe der Einträge der j-ten Spalte von S ist. Da

$$\sigma_j = \sum_{i=1}^n l_{ij} + a_j = 1,$$

folgt:

$$e^T S = e^T$$

und also:

$$\sum_{i=1}^n x_i = d \left(\sum_{i=1}^n x_i \right) + (1-d)$$

Da $d \neq 1$ folgt:

$$e^T x = \sum_{i=1}^n x_i = 1$$

Aus (P) folgt weiter:

$$Gx = d * Sx + (1-d) \underbrace{\frac{1}{n} e^T x}_{=1} = E_n x = x$$

Also ist x Eigenvektor von G zum Eigenwert 1.

Sei nun x Eigenvektor von G zum Eigenwert 1 mit $\sum_{i=1}^n x_i = 1$. Dann:

$$(d * S - E_n)x = \underbrace{Gx - x}_{=0} - (1-d) \underbrace{\frac{1}{n} e^T x}_{=1} = -\frac{1-d}{n} e$$

Also löst x (P). \square

8.2 Definition

Eine Matrix $A \in \mathbb{R}^{n \times n}$ mit nicht negativen Einträgen (also $a_{ij} \geq 0 \forall i, j = 1, \dots, n$) heißt:

- **Zeilenstochastisch**, falls $A * e = e$
- **Spaltenstochastisch**, falls $e^T A = e^T$
(Hier sind also im ersten Fall alle Zeilensummen = 1, im zweiten Fall alle Spaltensummen = 1).

8.3 Lemma

S und G sind spaltenstochastisch.

Beweis:

Die Nichtnegativität der Einträge ist klar. $e^T S = e^T$ wurde in Lemma 8.1 mitgezeigt.

Also:

$$e^T G = de^T s + \frac{1-d}{n} \underbrace{e^T e}_{=n} e^T = de^T + (1-d)e^T = e^T$$

□

8.4 Satz

$A \in \mathbb{R}^{n \times n}$ sei zeilen, oder spaltenstochastisch. Dann:

- i) 1 ist Eigenwert zu A .
- ii) Betrachtet man A als Element von $\mathbb{C}^{n \times n}$ und ist $\lambda \in \mathbb{C}$ Eigenwert von A , so: $|\lambda| \leq 1$.
Kurz: 1 ist der betragsgrößte Eigenwert zu A .

Erinnerung: Ist $z = a + ib \in \mathbb{C}$ ($a, b \in \mathbb{R}$), so $|z| := \sqrt{a^2 + b^2}$

Beweis:

Ohne Einschränkung sei A zeilenstochastisch (Satz 7.10).

- i) Nach Def. 8.2 ist dann e Eigenvektor von A zum Eigenwert 1.
- ii) Sei $v \in \mathbb{C}^n \setminus \{0\}$ ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{C}$.
Dann gilt nach den Rechenregeln für den komplexen Betrag aus MfI 1:

$$\begin{aligned} |\lambda| * \max_{i=1, \dots, n} |v_i| &= \max_{i=1, \dots, n} (|\lambda| |v_i|) \\ \max |\lambda v_i| &=_{Av=\lambda v} \max \left| \sum_{j=1}^n a_{ij} v_j \right| \\ &\leq \max \sum_{j=1}^n \underbrace{a_{ij}}_{\geq 0} |v_j| \leq \max_{k=1, \dots, n} |v_k| \max_{i=1, \dots, n} \underbrace{\sum_{j=1}^n a_{ij}}_1 = \max_{k=1, \dots, n} |v_k| \end{aligned}$$

Da $v \neq 0$, folgt $\max_k |v_k| \neq 0$. Also $|\lambda| \leq 1$. □

Das folgende als **Potenzmethode** oder **einfache Vektoiteration** bekannte Verfahren kann verwendet werden, um unter geeigneten Voraussetzungen einen Eigenvektor zum betragsgrößten Eigenwert einer Matrix $A \in \mathbb{C}^{n \times n}$ zu bestimmen:

Sei λ der betragsgrößte Eigenwert von $v^0 \in \mathbb{C}^n$ ein „geeignet“ gewählter Startvektor. Setze:

$$v^{(k)} := Av^{(k-1)}, k \in \mathbb{N}$$

$$= A^k v^{(0)}, a^k = \underbrace{A \cdots A}_{k\text{-mal}}$$

Dann gilt typischerweise:

Für alle $i = 1, \dots, n$ konvergiert $(\lambda^{-k} v_i^{(k)})_{k \in \mathbb{N}}$ gegen ein $v_i \in \mathbb{C}$, wobei $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ ein

Eigenvektor zum Eigenwert λ ist.

„(s. Stoer/Bulirsch, *Einführung in die Numerische Mathematik II*, Springer)“

Für spaltenstochastische Matrizen gilt:

8.5 Satz (Spezialfall des Satzes von Frobenius-Perron)

$A \in \mathbb{R}^{n \times n}$ sei spaltenstochastisch mit strikt positiven Einträgen.

i) Es gibt genau einen Eigenvektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ zum Eigenwert 1 mit $\sum_{i=1}^n x_i = 1$.
Dieser erfüllt $x_i > 0$ für alle $i = 1, \dots, n$.

ii) Gegeben ein Startvektor $x^{(0)} \in \mathbb{R}^n$ sein $x^{(k)} = Ax^{(k-1)}$. Dann gilt für alle $i = 1, \dots, n$

$$\lim_{k \rightarrow \infty} x_i^{(k)} = x_i * \left(\sum_{j=1}^n x_j^{(0)} \right)$$

8.6 Bemerkung

Die einfache Vektoriteration konvergiert hier also gegen einen Eigenvektor zum Eigenwert 1, sofern

$$x^{(0)} \notin H(e, 0) = \{y \in \mathbb{R}^n \mid \sum_{i=1}^n y_i = 0\}.$$

8.7 Folgerung

Für $0 < d < 1$ hat das Page-Rank-Gleichungssystem (P) genau eine Lösung $x \in \mathbb{R}^n$. Diese kann mittels der einfachen Vektoriteration konstruiert werden, sofern:

$$x^{(0)} \in H(e, 1) = \{y \mid \sum_{i=1}^n y_i = 1\}$$

Beweis von Satz 8.5

Für $k \geq 1$ seien $a_{ij}^{(k)}$ die Einträge A^k . Sei

$$m_i^{(k)} = \min_{j=1, \dots, n} a_{ij}^{(k)}$$

$$M_i^{(k)} = \max_{j=1, \dots, n} a_{ij}^{(k)}$$

Schritt 1: Es reicht zu zeigen, dass für alle $i = 1, \dots, n$ ein $x_i > 0$ existiert, mit:

$$\lim_{k \rightarrow \infty} = \lim_{k \rightarrow \infty} M_i^{(k)} = x_i$$

Da $x^{(k)} = A^k x^{(0)}$, folgt für ein beliebiges $x^{(0)} \in \mathbb{R}^n$:

$$\begin{aligned} x_i^{(k)} &= e_i^T A^k \left(\sum_{j=1}^n x_j^{(0)} e_j \right) = \sum_{j=1}^n x_j^{(0)} e_j^T A^k e_i \\ &= \sum_{j=1}^n x_j^{(0)} a_{ji}^{(k)} \end{aligned}$$

Da $m_i^{(k)} \leq a_{ij}^{(k)} \leq M_i^{(k)}$ für alle $j = 1, \dots, n$, folgt aus dem Sandwich Theorem (MFI1), dass

$$\lim_{k \rightarrow \infty} a_{ij}^{(k)} = x_i \text{ für alle } j = 1, \dots, n$$

Somit:

$$\lim_{k \rightarrow \infty} x_i^{(k)} = \lim_{k \rightarrow \infty} \sum_{j=1}^n x_j^{(0)} a_{ji}^{(k)} = x_i \sum_{j=1}^n x_j^{(0)}$$

Sei $x^{(0)}$ ein Eigenvektor zum Eigenwert 1 (s. Satz 8.4). Dann folgt induktiv:

$$x^{(k)} = A^k x^{(0)} = A^{k-1} \underbrace{A x^{(0)}}_{=x^{(0)}} = A^{k-1} x^{(0)} = \dots = x^{(0)}$$

Also folgt:

$$x_i^{(0)} = \lim_{k \rightarrow \infty} x_i^{(k)} = x_i \left(\sum_{j=1}^n x_j^{(0)} \right)$$

8 Eigenwertprobleme für stochastische Matrizen

Da $x^{(0)} \neq 0 \in \mathbb{R}^n$, folgt $\sum_{j=1}^n x_j^{(0)} \neq 0$. Sei $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Dann ist $x = \frac{1}{\sum_{j=1}^n x_j^{(0)}} * x^{(0)}$ ein Eigenvektor zum Eigenwert 1 mit $\sum_{i=1}^n x_i = 1$.

Für jeden Eigenvektor $x^{(0)}$ zum Eigenwert 1 mit $\sum_{i=1}^n x_i^{(0)} = 1$, folgt:

$$x^{(0)} = \left(\sum_{j=1}^n x_j^{(0)} \right) x = x$$

Schritt 2: Für alle $i = 1, \dots, n$ existiert ein $x_i > 0$ mit $\lim_{k \rightarrow \infty} M_i^{(k)} = \lim_{k \rightarrow \infty} m_i^{(k)} = x_i$.
Da die Einträge von A strikt positiv sind, gilt:

$$M_i^{(1)} \geq m_i^{(1)} \geq \epsilon := \min_{i=1, \dots, n, j=1, \dots, n} a_{ij} > 0$$

Da

$$\begin{aligned} m_i^{(k+1)} &= \min_j a_{ij}^{(k+1)} = \min \left(\sum_{l=1}^n a_{il}^{(k)} a_{lj} \right) \\ &\geq \min_{v=1, \dots, n} a_{iv}^{(k)} \underbrace{\min_j \sum_{l=1}^n a_{lj}}_{=1} = m_i^{(k)} \end{aligned}$$

und analog $M_i^{k+1} \leq M_i^{(k)}$, ist $(m_i^{(k)})_{k \in \mathbb{N}}$ monoton wachsend und nach oben beschränkt durch $M_i^{(1)}$ und $(M_i^{(k)})_{k \in \mathbb{N}}$ monoton fallend und nach unten beschränkt durch $m_i^{(1)} \geq \epsilon$.

Beide Folgen haben also Grenzwerte \underline{x}_i (unterstrichen) bzw. \overline{x}_i (überstrichen) die jeweils größer oder gleich ϵ sind. Es bleibt zu zeigen, dass \underline{x}_i (unterstrichen) = \overline{x}_i (überstrichen), also:

$$\lim_{k \rightarrow \infty} (M_i^{(k)} - m_i^{(k)}) = 0$$

Dazu beachte man zunächst, dass A^k splatenstochastisch ist, da induktiv:

$$e^T A^k = e^T A A^{k-1} = e^T A^{k-1} = \dots = e^T A = e^T$$

Somit:

$$\begin{aligned} a_{ij}^{k+1} &= \sum_{l=1}^n a_{il}^k a_{lj} = \sum_{l=1}^n a_{il}^k (a_{lj} - \epsilon a_{li}^{(k)}) + \epsilon \sum_{l=1}^n a_{il}^{(k)} a_{li}^{(k)} \\ &\stackrel{A^{2k} = A^k * A^k}{=} \sum_{l=1}^n a_{il}^k \underbrace{(a_{lj} - \epsilon a_{li}^{(k)})}_{\geq 0} + \epsilon a_{ii}^{(2k)} \end{aligned}$$

$$\begin{aligned} &\geq m_i^{(k)} \sum_{l=1}^n (a_{lj} - \epsilon a_{li}^{(k)}) + \epsilon a_{ii}^{(2k)} \\ &\quad m_i^{(k)} (1 - \epsilon) + \epsilon a_{ii}^{(2k)} \end{aligned}$$

Also:

$$\begin{aligned} M_i^{k+1} - m_i^{(k+1)} &= \max_j a_{ij}^{(k)} - \min_j a_j^{(k)} \\ &\leq (1 - \epsilon)(M_i^{(k)} - m_i^{(k)}) \end{aligned}$$

Also:

$$0 \leq M_i^{(k)} - m_i^{(k)} \leq (1 - \epsilon)^{k-1} (M_i^{(1)} - m_i^{(1)})$$

Da $\lim_{k \rightarrow \infty} (1 - \epsilon)^k = 0$, folgt:

$$\lim_{k \rightarrow \infty} M_i^{(k)} - m_i^{(k)} = 0. \square$$

8.8 Bemerkung

Im Kontext von Satz 8.5 gelte für den Startvektor $x_i^{(0)}$ und $\sum_{i=1}^n x_i^{(0)} = 1$. Dann folgt leicht aus dem Beweis von Satz 8.5, dass für $\epsilon := \min_{i,j} a_{ij}$

$$\max_{i=1,\dots,n} |x_i^{(k)} - x_i| \leq \sum_{j=1}^n x_j^{(0)} \underbrace{\max_i |a_{ij}^{(k)} - x_i|}_{\leq M_i^{(k)} - m_i^{(k)}} = (1 - \epsilon)^k$$

Mit mehr Aufwand kann man zeigen, dass für

$$\Lambda := \max \{ |\lambda| \mid \lambda \in \mathbb{C} \text{ Eigenwert zu } A \in \mathbb{C}^{n \times n}, \lambda \neq 1 \} < 1$$

gilt: Es gibt Konstanten $C \geq 0$ und $0 \leq m \leq n - 2$, sodass:

$$\max_{i=1,\dots,n} |x_i^{(k)} - x_i| \leq C k^m \Lambda^k$$

Im Page-Rank-Beispiel $A = G$ kann man zeigen, dass $\Lambda \leq d$, während

$$1 - \epsilon \geq \frac{n - 1 + d}{n}$$

8.9 Bemerkung

Löst man das Page-Rank-Gleichungssystem mittels einfacher Vektoriteration, so gilt wegen $e^T x^{(k)} = e^T A^k x^{(0)} = e^T x^{(0)} = 1$:

$$x^{(k+1)} = Gx^{(k)} = d * \left(Lx^{(k)} \right) + e \left(\frac{d}{n} a^T x^{(k)} + \frac{1-d}{n} \right)$$

8 Eigenwertprobleme für stochastische Matrizen

Die Matrix L ist dünn besetzt (viele Nullen). Für das Matrix-Vektor-Produkt $Lx^{(k)}$ müssen nur so viele Multiplikationen in \mathbb{R} durchgeführt werden, wie das Netz Links hat (ca. 10 Links pro Seite im Durchschnitt). Es werden also in jedem Iterationsschritt $O(n)$ Multiplikationen durchgeführt. Da die Iteration mit einer exponentiellen Rate konvergiert, erwartet man, dass typischerweise $\log(n)$ Iterationen ausreichen, um eine hinreichend genaue Approximation zu erhalten. Der Gesamtaufwand beträgt dann $O(n * \log(n))$ Multiplikationen.

9 Determinante und charakteristisches Polynom

Sei $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in K^{2 \times 2}$. Wir definieren die Determinante von A durch:

$$\det(A) := a * d - b * c$$

9.1 Beispiel

Sei $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ Vektoren im \mathbb{R}^2 mit positiven Einträgen. Wir betrachten das von v und w aufgespannte Parallelogramm:

Sei F der Flächeninhalt des Parallelogramms und die weiteren Großbuchstaben bezeichnen die Fläche der Recht- bzw. Dreieckflächen. Dann:

$$\begin{aligned} F &= (v_1 + w_1) * (v_2 + w_2) - (A + B + C_1 + C_2 + D_1 + D_2) \\ &= v_1 v_2 + w_1 v_2 + v_2 w_2 + w_1 w_2 - (2v_2 * w_1 + w_1 w_2 + v_1 v_2) \\ &= v_1 w_2 - v_2 w_1 = \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \end{aligned}$$

9.2 Beispiel

Wir betrachten das Lin-Gleichungssystem $Ax = b$, $A = (a_{ij})_{i,j=1,2} \in K^{2 \times 2}, b \in K^2$. A sei invertierbar. Zeilenumformungen (1. Zeile \rightsquigarrow $a_{22} * 1.$ Zeile - $a_{12} * 2.$ Zeile; 2. Zeile \rightsquigarrow $a_{11} * 2.$

Zeile - $a_{21} * 1.$ Zeile) zeigen: die eindeutige Lösung $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ löst auch:

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{12}b_2$$

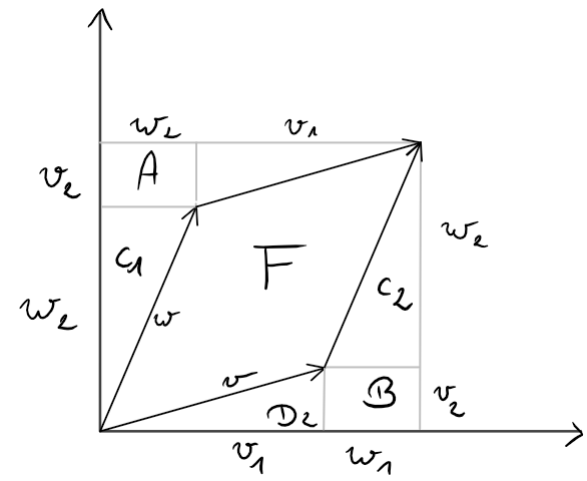
$$(a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - a_{21}b_1$$

Man überprüft leicht, dass A invertierbar ist, genau dann, wenn $\det(A) \neq 0$ (s. Folgerung 9.9).

Also:

$$x_1 = \frac{1}{\det(A)} \det \begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}, x_2 = \frac{1}{\det(A)} \det \begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}$$

Dies ist der einfachste Fall der **Cramerschen Regel**.



9.3 Definition

Sei K ein Körper und $n \in \mathbb{N}$. Eine Abbildung

$$\det : K^{n \times n} \rightarrow K$$

wird **Determinante** genannt, falls

(D1) \det ist linear in jeder Zeile, d.h. für jedes $i = 1, \dots, n$ gilt:

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i + b_i \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ b_i \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix}$$

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ \lambda a_i \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix} = \lambda \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix}$$

Hierbei sind $a_1, \dots, a_n, b_i \in K^{1 \times n}$ Zeilenvektoren und $\lambda \in K$.

(D2) \det ist alternierend, d.h. hat A zwei gleiche Zeilen, so ist $\det(A) = 0$

(D3) \det ist normiert, d.h. $\det(E_n) = 1$

9.4 Bemerkung

Für $n = 2$ prüft man leicht nach, dass

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a * d - b * c$$

eine Determinante definiert.

Für $n = 3$ definiert die **Sarrus-Regel**

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$:= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$$

Man kann zeigen:

9.5 Satz

Zu jedem Körper K und jedem $n \in \mathbb{N}$ existiert eine eindeutig bestimmte Determinante.

9.6 Bemerkung

- i) Seien $I^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \mid 0 \leq x_i \leq 1 \right\}$ der Einheitswürfel im \mathbb{R}^n und $A \in \mathbb{R}^{n \times n}$

eine Menge der Form $P = \{Ax \mid x \in I^n\}$ wird **Parallelotop** mit Abbildungsmatrix A genannt. Im Fall $n = 2$ sind dies gerade Parallelogramme. In Verallgemeinerung zu Beispiel 9.1 gilt für einen geeigneten Volumenbegriff (s. MFI 3): $|\det(A)|$ ist das Volumen von P .

- ii) Die Verallgemeinerung der Cramerschen Regel für $n > 2$ lautet:

Ist $A = (a_{ij})_{i,j} \in GL(n, K)$ und $b \in K^n$, so ist die eindeutige Lösung $x \in K^n$ zu $Ax = b$ gegeben durch

$$x_i = \frac{\det(a_{\cdot 1}, \dots, a_{\cdot (i-1)}, b, a_{\cdot (i+1)}, \dots, a_{\cdot n})}{\det(A)}$$

wobei $a_{\cdot j}$ die j -te Spalte von A ist.

Beispiel:

Das Gleichungssystem:

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} * x = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

9 Determinante und charakteristisches Polynom

soll gelöst werden. Mit der Sarrus-Regel berechnet man:

$$\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} = 2, \det \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} = -2$$

$$\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 0 & 1 \end{pmatrix} = 4, \det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 0 \end{pmatrix} = -2$$

Also:

$$x = \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix}$$

9.7 Satz

Für $\det : K^{n \times n} \rightarrow K$ gilt:

- i) Entsteht B aus A durch eine Zeilenvertauschung, so gilt: $\det(B) = -\det(A)$
- ii) Entsteht B aus A durch Addition des λ -fachen der j -ten Zeile zur i -ten Zeile ($i \neq j, \lambda \in K$), so: $\det(B) = \det(A)$
- iii) Ist A eine obere Dreiecksmatrix, also:

$$A = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \ddots & & \vdots \\ & & \ddots & * \\ 0 & & & \lambda_n \end{pmatrix}$$

Dann gilt: $\det(A) = \lambda_1 \dots \lambda_n$

Beweis

- i) Es werde die i -te mit der j -ten Zeile vertauscht ($i \neq j$). Zur Vereinfachung der Notation schreiben wir nur die relevanten Zeilen:

$$\begin{aligned} \det(A) + \det(B) &= \det \begin{pmatrix} a_i \\ a_j \end{pmatrix} + \det \begin{pmatrix} a_j \\ a_i \end{pmatrix} \\ &\stackrel{D2}{=} \det \begin{pmatrix} a_i \\ a_j \end{pmatrix} + \det \begin{pmatrix} a_i \\ a_i \end{pmatrix} + \det \begin{pmatrix} a_j \\ a_i \end{pmatrix} + \det \begin{pmatrix} a_j \\ a_j \end{pmatrix} \\ &\stackrel{D1}{=} \det \begin{pmatrix} a_i \\ a_j + a_i \end{pmatrix} + \det \begin{pmatrix} a_j \\ a_i + a_j \end{pmatrix} \end{aligned}$$

$$\underbrace{=}_{D1} \det \begin{pmatrix} a_i + a_j \\ a_i + a_j \end{pmatrix}$$

$$\underbrace{=}_{D2} 0$$

ii)

$$\det(B) = \det \begin{pmatrix} a_i + \lambda a_j \\ a_j \end{pmatrix}$$

$$\underbrace{=}_{D1} \det(A) + \lambda \det \begin{pmatrix} a_j \\ a_j \end{pmatrix}$$

$$\underbrace{=}_{D2} \det(A)$$

iii) Sind alle $\lambda_i \neq 0$, so folgt durch wiederholtes Anwenden von ii):

$$\det(A) = \det \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \underbrace{=}_{D1} \lambda_1 \dots \lambda_n \det(E_n) \underbrace{=}_{D3} \lambda_1 \dots \lambda_n$$

Andernfalls wählen wir i maximal mit $\lambda_i = 0$ (d.h. $\lambda_{i+1}, \dots, \lambda_n \neq 0$). Wiederholtes Anwenden von ii) führt zu:

$$\det(A) = \det \begin{pmatrix} \lambda_1 & & & & & * \\ & \ddots & & & & \\ 0 & & \lambda_{i-1} & & & \\ 0 & \dots & \dots & 0 & \dots & 0 \\ & & & \lambda_{i+1} & & * \\ & & & & \ddots & \\ 0 & & & & & \lambda_n \end{pmatrix} = \underbrace{=}_{D1} = 0$$

$$= \lambda_1 \dots \lambda_n \cdot \square$$

Satz 9.7 rechtfertigt folgenden Algorithmus zur Berechnung der Determinanten:

9.8 Algorithmus

Input: $A \in K^{n \times n}$

Output: $\det(A)$

1.) Bringe A mit Algorithmus 1.6 auf Zeilenstufenform $\tilde{A} = (\tilde{a}_{ij})_{i,j=1,\dots,n}$ und merke dabei die Anzahl k der Zeilenumformungen.

2.) Gebe $(-1)^k \tilde{a}_{11} * \dots * \tilde{a}_{nn}$ als Output aus.

Insbesondere folgt die Eindeutigkeit der Determinante in Satz 9.5.

Der Algorithmus zur Berechnung der Inversen aus Kapitel 4 zeigt:

$$A \text{ invertierbar} \Leftrightarrow \forall_{i=1, \dots, n} : \tilde{a}_{ii} \neq 0.$$

Also:

9.9 Folgerung

$A \in K^{n \times n}$ ist invertierbar, genau dann, wenn $\det(A) \neq 0$.

9.10 Beispiel

Mit Algorithmus 9.8 berechnen wir:

$$\begin{aligned} \det \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix} &= -\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix} = -\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & -1 & 1 \end{pmatrix} \\ &= -\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix} = -3 \end{aligned}$$

Zum Existenzbeweis der Determinante kann man eine explizite Berechnungsformel, die sogenannte Leibniz-Formel angeben. Dazu sei:

$$S_n = \{\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\} \mid \sigma \text{ bijektiv}\}$$

die sogenannte **symmetrische Gruppe** vom Grad n . Ihre Elemente heißen **Permutationen**. Es sei:

$$\text{sign}(\sigma) := \begin{cases} 1 & , \text{ falls } \{(i, j) \in \{1, \dots, n\}^2 \mid i < j \text{ und } \sigma(i) > \sigma(j)\} \\ & \text{eine gerade Anzahl von Elementen hat} \\ -1 & , \text{ sonst} \end{cases}$$

das Vorzeichen von σ .

9.11 Satz (Leibniz Formel)

Ist K ein Körper und $n \geq 1$, so definiert:

$$\det : K^{n \times n} \rightarrow K : (a_{ij})_{i,j=1, \dots, n} \mapsto \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)}$$

eine und (also die) Determinante.

Beweis:

Fischer, Lineare Algebra.

9.12 Satz (Laplacescher Entwicklungssatz)

Sei $n \geq 2$ und $A = (a_{ij})_{i,j=1,\dots,n} \in K^{n \times n}$. Es sei für $i, j \in \{1, \dots, n\}$ A'_{ij} die $(n-1) \times (n-1)$ Matrix, die entsteht, wenn man aus A die i -te Zeile und j -te Spalte entfernt. Dann gilt:
Für beliebiges $i = 1, \dots, n$

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A'_{ij})$$

(Entwicklung nach der i -ten Zeile) und für beliebiges $j = 1, \dots, n$

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A'_{ij})$$

(Entwicklung nach der j -ten Spalte).

9.13 Beispiel

$$\begin{aligned} \det \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix} &= 0 * \det \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} + (-1) * 1 * \det \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} + 2 * \det \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix} \\ &= 0 - 1 - 2 = -3 \end{aligned}$$

9.14 Satz

Seien $A, B \in K^{n \times n}$. Dann:

- i) $\det(A^T) = \det(A)$
- ii) $\det(A * B) = \det(A) * \det(B)$
Insbesondere $\det(A^{-1}) = (\det(A))^{-1}$ für $A \in GL(n, k)$

Beweisskizze

- ii) Ist $\text{Rang}(A) < n$, so ist auch $\text{Rang}(A * B) < n$. Nach Folgerung 9.9 sind dann beide Seiten der Gleichung gleich 0. Ist $\text{Rang}(A) = n$ und ist also invertierbar, so gibt es $\tilde{a}_{11}, \dots, \tilde{a}_{nn} \in K \setminus \{0\}$, so dass A aus:

$$\begin{pmatrix} \tilde{a}_{11} & & 0 \\ & \ddots & \\ 0 & & \tilde{a}_{nn} \end{pmatrix}$$

9 Determinante und charakteristisches Polynom

durch k-Zeilenumtauschen und mehrere Additionen von Vielfachen einer Zeile zu einer anderen hervorgeht (man kehre die Umformungen zur Matrixinvertierung von A einfach um). Die Matrix $A * B$ erhält man zuerst für alle $i = 1, \dots, n$, die i-te Zeile von B mit \tilde{a}_{ii} multipliziert und dann die obigen Zeilenumformungen auf die resultierende Matrix anwendet. Also folgt aus Satz 9.7:

$$\begin{aligned} \det(A * B) &= (-1)^k \det \begin{pmatrix} \tilde{a}_{11} * b_1 \\ \vdots \\ \tilde{a}_{nn} * b_n \end{pmatrix} \underbrace{=}_{D1} (-1)^k \tilde{a}_{11} \dots \tilde{a}_{nn} * \det(B) \\ &= \det(A) \det(B) \end{aligned}$$

- i) Hierzu überlegt man sich, dass bei der Berechnung von $\det(A^T)$ mittels der Leibnizformel der Summand zur Permutation σ gerade dem Summanden zur Permutation σ^{-1} bei der Berechnung von $\det(A)$ entspricht. \square

Fragestellung:

Wie kann man die Eigenwerte einer Matrix $A \in K^{n \times n}$ berechnen?

Dazu:

λ ist Eigenwert zu $A \Leftrightarrow (A - \lambda E_n)$ ist nicht invertierbar

$$\underbrace{\Leftrightarrow}_{\text{Folgerung 9.9}} \det(A - E_n) = 0$$

Wir betrachten die Funktion

$$P_A : K \rightarrow K, t \mapsto \det(A - tE_n),$$

deren Nullstellen gerade die Eigenwerte von A sind. P_A heißt **charakteristische Funktion** von A. Ist $K = \mathbb{R}$ oder $K = \mathbb{C}$, so ist P_A ein Polynom vom Grad n, das sogenannte **charakteristische Polynom**. Dies sieht man mittels der Leibniz-Formel (vgl. Übung).

9.15 Beispiel

Ist

$$A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix},$$

so folgt mit dem Laplaceschen Entwicklungssatz:

$$\begin{aligned} P_A(t) &= \det \begin{pmatrix} -t & -1 & 1 \\ -3 & -2-t & 3 \\ -2 & -2 & 3-t \end{pmatrix} \\ &= -t \det \begin{pmatrix} -2-t & 3 \\ -2 & 3-t \end{pmatrix} + 3 \det \begin{pmatrix} -1 & 1 \\ -2 & 3-t \end{pmatrix} - 2 \det \begin{pmatrix} -1 & 1 \\ -2-t & 3 \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
&= -t(-2-t)(3-t) - 6t - 3(3-t) + 6 + 6 + 2(-2-t) \\
&= -t^3 + t^2 + t - 1 = -(t-1)^2(t+1)
\end{aligned}$$

Also sind $\lambda_1 = 1$ und $\lambda_2 = -1$ die Eigenwerte von A . Mittels des Gauß-Algorithmus berechnet man die Eigenräume:

$$\begin{aligned}
\text{Eig}(A, 1) &= \text{Lös} \left(\begin{pmatrix} -1 & -1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right) \\
&= \{x \in \mathbb{R}^3 \mid -x_1 - x_2 + x_3 = 0\} \\
&= \text{span} \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right) \\
\text{Eig}(A, -1) &= \text{Lös} \left(\begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right) \\
&= \text{span} \left(\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right)
\end{aligned}$$

Somit ist

$$\mathcal{B} = \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right)$$

eine Basis aus Eigenvektoren. Es folgt, dass A diagonalisierbar ist, präziser ähnlich zu

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

10 Prä- Hilberträume und Orthogonalprojektionen

Motivation:

wir wollen Begriffe wie die Länge eines Vektors und den Winkel zwischen Vektoren einführen.

10.1 Beispiel

Für Vektoren $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ definieren wir das „kanonische Skalarprodukt“ durch

$$\langle x, y \rangle := x_1 y_1 + x_2 y_2$$

Es ist dann

$$\sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + x_2^2}$$

nach Satz von Pythagoras die Länge des Vektors x : Sei nun $x \neq 0$. Wir setzen

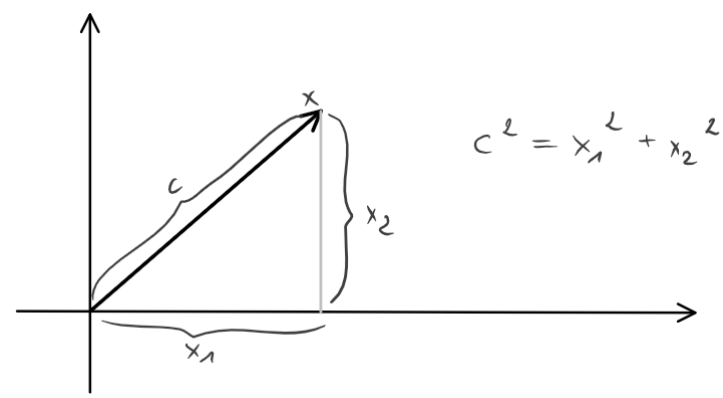
$$\tilde{x} = \langle x, x \rangle^{-\frac{1}{2}} * x$$

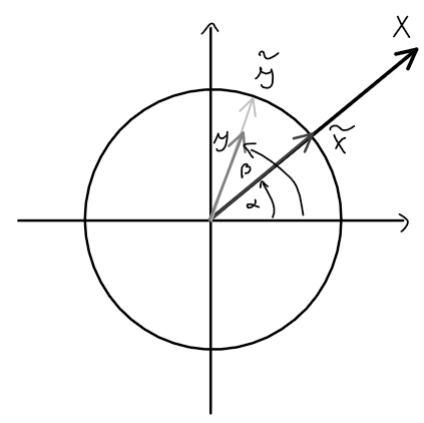
Dann:

$$\langle \tilde{x}, \tilde{x} \rangle = \langle x, x \rangle^{-1} * (x_1^2 + x_2^2) = 1$$

Analog sei für $y \neq 0$

$$\tilde{y} = \langle y, y \rangle^{-\frac{1}{2}} * y$$





Aus der MfI1 wissen wir: es existieren $\alpha, \beta \in [0, 2\pi)$ mit

$$\tilde{x} = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}$$

$$\tilde{y} = \begin{pmatrix} \cos(\beta) \\ \sin(\beta) \end{pmatrix}$$

Nach dem Additionstheorem für den Kosinus folgt:

$$\begin{aligned} \langle x, y \rangle &= \sqrt{\langle x, x \rangle} * \sqrt{\langle y, y \rangle} (\cos(\alpha) \cos(\beta) + \sin(\alpha) \sin(\beta)) \\ &= \sqrt{\langle x, x \rangle} * \sqrt{\langle y, y \rangle} \cos(\beta - \alpha) \end{aligned}$$

Wir betrachten das $\varphi \in [0, \pi]$ mit $\cos(\varphi) = \cos(\beta - \alpha)$ als den „unorientierten Winkel“ zwischen x und y . Für ihn gilt:

$$\varphi = \arccos \left(\frac{\langle x, y \rangle}{\sqrt{\langle x, x \rangle \langle y, y \rangle}} \right)$$

10.2 Definition

Sei V ein \mathbb{R} -VR. Unter einem Skalarprodukt verstehen wir eine Abbildung

$$\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$$

mit

(S1) Linearität in der ersten Variablen, d.h.

$$\langle \lambda v, w \rangle = \lambda \langle v, w \rangle \quad \text{und} \quad \langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$$

für alle $v, v_1, v_2, w \in V$ und $\lambda \in \mathbb{R}$

(S2) Symmetrie:

$$\langle v, w \rangle = \langle w, v \rangle \text{ für alle } v, w \in V$$

(S3) Positivität:

$$\langle v, v \rangle \geq 0 \text{ und } (\langle v, v \rangle = 0 \Rightarrow v = 0)$$

Ein \mathbb{R} -VR versehen mit einem Skalarprodukt wird **Euklidischer Raum** oder **Prä-Hilbertraum** genannt.

10.3 Bemerkung

Auf Grund der Symmetrie und der Linearität in der ersten Variable ist ein Skalarprodukt über einem \mathbb{R} -VR auch linear in der zweiten Variable, also eine „Bilinearform“.

10.4 Beispiel

i) In \mathbb{R}^n definiert

$$\langle x, y \rangle := x^T y = \sum_{i=1}^n x_i y_i$$

ein Skalarprodukt, das sog. kanonische Skalarprodukt oder Euklidische Skalarprodukt im \mathbb{R}^n . Zusammen mit diesem Skalarprodukt wird der \mathbb{R}^n auch der n -dimensionale Euklidische Raum genannt.

ii) Auf dem \mathbb{R} -VR $\varphi([a, b], \mathbb{R})$, $-\infty < a < b < \infty$, definiert

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx$$

ein Skalarprodukt.

10.5 Satz (Cauchy-Schwarz-Ungleichung)

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Prä-Hilbertraum und $\|v\| := \sqrt{\langle v, v \rangle}$. Dann:

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|, v, w \in V$$

Gleichheit gilt genau dann, wenn v und w linear abhängig sind.

Beweis

Sei $w \neq 0$. Dann gilt für alle $\lambda, \mu \in \mathbb{R}$

$$\begin{aligned} (*) \quad 0 &\leq \langle \lambda v + \mu w, \lambda v + \mu w \rangle \\ &= \lambda^2 \langle v, v \rangle + \mu^2 \langle w, w \rangle + 2\mu\lambda \langle v, w \rangle \end{aligned}$$

10 Prä-Hilberträume und Orthogonalprojektionen

Setzt man $\mu := \langle v, w \rangle$ und $\lambda := \langle w, w \rangle$ und teilt durch $\lambda > 0$, so

$$\begin{aligned} 0 &\leq \lambda \langle v, v \rangle + \langle v, w \rangle^2 - 2 \langle v, w \rangle^2 \\ &= \|w\|^2 \|v\|^2 - \langle v, w \rangle^2 \end{aligned}$$

Durch Wurzelziehen folgt die Ungleichung. Falls die Gleichheit gilt, muss das „ \leq “ in (*) ein „ $=$ “ sein, für die obige Wahl von λ und μ . Also:

$$\lambda v + \mu w = 0$$

Da $\lambda > 0$, sind v und w linear abhängig.

Existiert andererseits ein beliebiges $\begin{pmatrix} \lambda \\ \mu \end{pmatrix} \in \mathbb{R}^2 \setminus \{0\}$ mit $\lambda v + \mu w = 0$, so $v = -\frac{\mu}{\lambda} w$.

Also

$$|\langle v, w \rangle| = \left| \frac{-\mu}{\lambda} \right| \langle w, w \rangle = \|v\| \|w\|,$$

da

$$\|v\| = \sqrt{\langle \frac{-\mu}{\lambda} w, \frac{-\mu}{\lambda} w \rangle} = \sqrt{\left(\frac{\mu}{\lambda}\right)^2 \langle w, w \rangle} = \left| \frac{-\mu}{\lambda} \right| \sqrt{\langle w, w \rangle}$$

Ist $w = 0$, so sind v und w linear abhängig und beide Seiten der Ungleichung sind gleich Null. \square

10.6 Bemerkung

In einem Prä-Hilbertraum V gilt also:

$$\frac{\langle v, v \rangle}{\|v\| \|w\|} \in [-1, 1]$$

Somit können wir in Verallgemeinerung zu Bsp 10.1 den Winkel zwischen $v \neq 0$ und $w \neq 0$ definieren durch:

$$\arccos \left(\frac{\langle v, w \rangle}{\|v\| \|w\|} \right)$$

Hierbei ist wieder $\|v\| := \sqrt{\langle v, v \rangle}$ die „Länge“ des Vektors v .

Allgemein beschreibt das Konzept der Norm die Länge von Vektoren:

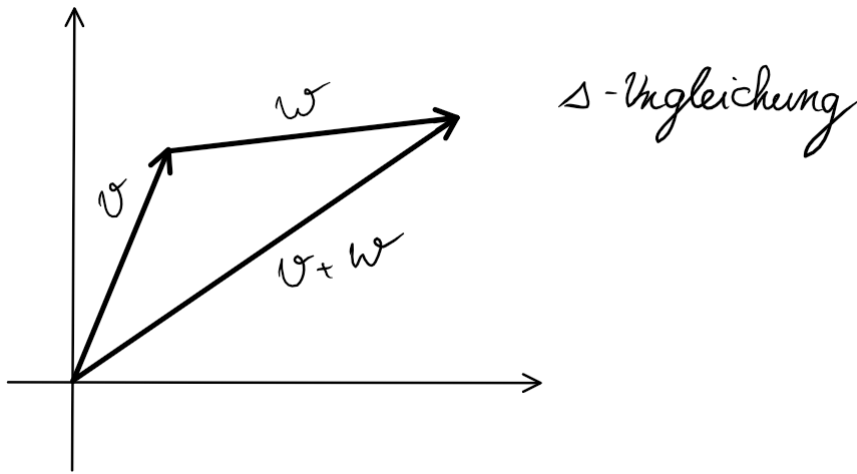
10.7 Definition

Sei V ein \mathbb{R} -VR. Eine Abbildung

$$\|\cdot\| : v \rightarrow \mathbb{R}$$

wird **Norm** genannt, falls:

(N1) $\|\lambda v\| = |\lambda| \|v\|$ für alle $\lambda \in \mathbb{R}, v \in V$ (positiv homogen)



(N2) Für alle v, w in V gilt

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

(Dreiecksungleichung)

(N3) $\|v\| \geq 0$ und $(\|v\| = 0 \Rightarrow v = 0)$ (Positivität)

10.8 Satz

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Prä-Hilbertraum. Dann definiert

$$\|v\| := \sqrt{\langle v, v \rangle}, v \in V,$$

eine Norm auf V , die von $\langle \cdot, \cdot \rangle$ **induzierte Norm**. Ferner gilt für alle $v, w \in V$:

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2 \langle v, w \rangle \quad (\mathbf{P})$$

Beweis

(N1) $\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{\lambda^2 \langle v, v \rangle} = |\lambda| \|v\|$

(N3) folgt direkt aus (S3).

(N2)+(P):

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + \langle w, w \rangle + 2 \langle v, w \rangle$$

$$\|v\|^2 + \|w\|^2 + 2 \langle v, w \rangle$$

$$\underbrace{\leq}_{\text{Cauchy-Schwarz}} \|v\|^2 + \|w\|^2 + 2\|v\|\|w\|$$

$$= (\|v\| + \|w\|)^2 \quad \square$$

10.9 Definition

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Prä-Hilbertraum. Dann heißen Vektoren v und w **orthogonal**, falls $\langle v, w \rangle = 0$. Nach Bemerkung 10.6 entspricht dies einem Winkel von $\frac{\pi}{2}$ im Bogenmaß, also 90° .

Eine Familie von Vektoren (v_1, \dots, v_n) heißt **orthogonal**, falls für alle $i \neq j$ v_i orthogonal zu v_j . Gilt zudem $\|v_i\| = 1$ für alle $i = 1, \dots, n$, so wird die Familie **orthonormal** genannt. Eine orthonormale Familie, die gleichzeitig Basis von V ist, heißt **Orthonormalbasis**.

10.10 Beispiel

Die kanonische Basis $\varphi_1, \dots, \varphi_n$ im \mathbb{R}^n ist eine Orthonormalbasis bzgl. des kanonischen Skalarprodukts.

10.11 Satz

Ist (v_1, \dots, v_n) orthogonal und $v_i \neq 0$ für alle $i = 1, \dots, n$, so ist (v_1, \dots, v_n) linear unabhängig.

Beweis

Sei $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ für $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Dann:

$$\begin{aligned} 0 &= \left\langle \sum_{i=1}^n \lambda_i v_i, \sum_{k=1}^n \lambda_k v_k \right\rangle \\ &= \sum_{i=1}^n \sum_{k=1}^n \lambda_i \lambda_k \underbrace{\langle v_i, v_k \rangle}_{=0 \text{ für } i \neq k} \\ &= \sum_{i=1}^n \lambda_i^2 \underbrace{\langle v_i, v_i \rangle}_{>0} \end{aligned}$$

Also $\lambda_i^2 = 0$ und somit $\lambda_i = 0$ für alle $i = 1, \dots, n$. \square

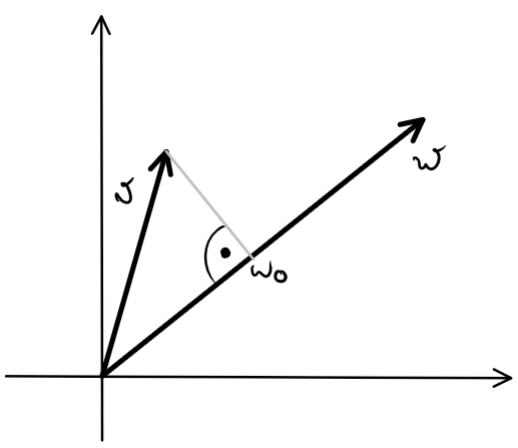
Frage:

Wie kann man Orthonormalbasen konstruieren?

10.12 Definition

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Prä-Hilbertraum, $W \subset V$ ein UVR und $v \in V$. Ein Vektor $w_0 \in W$ wird **Orthogonalprojektion** von v auf W genannt, falls

$$\langle v - w_0, w \rangle = 0 \text{ für alle } w \in W.$$



10.13 Bemerkung

Im Existenzfall ist w_0 eindeutig bestimmt. Denn ist \tilde{w}_0 eine weitere Orthogonalprojektion von v aus W , so gilt für $w = \tilde{w}_0 - w_0 \in W$

$$0 = \langle v - w_0, w \rangle = \langle v - \tilde{w}_0, w \rangle = \langle \tilde{w}_0 - w_0, \tilde{w}_0 - w_0 \rangle$$

Also: $\tilde{w}_0 - w_0 = 0$ nach (S3).

Wir beschreiben im Folgenden: $\Pi_W(v) = w_0$. offensichtlich gilt: $\Pi_W(w) = w$ für alle $w \in W$.

10.14 Satz

Sei $V, \langle \cdot, \cdot \rangle$ ein Prä-Hilbertraum und $W \subset V$ ein n -dimensionaler Unterraum ($n \in \mathbb{N}$). Dann:

- i) W hat eine Orthonormalbasis (w_1, \dots, w_n)
- ii) für jedes $v \in V$ und jede Orthonormalbasis (w_1, \dots, w_n) von W gilt:

$$\Pi_W(v) = \sum_{i=1}^n \langle v, w_i \rangle w_i$$

Insbesondere ist: $\Pi_W : V \rightarrow W$ linear.

Beweis

- ii) Sei (w_1, \dots, w_n) eine Orthonormalbasis von W . Sei zunächst $w \in W$. Dann:

$$w = \sum_{i=1}^n \lambda_i w_i$$

für gewisse $\lambda \in \mathbb{R}$. Es gilt für alle $k = 1, \dots, n$

$$\begin{aligned} \langle w, w_k \rangle &= \sum_{i=1}^n \lambda_i \underbrace{\langle w_i, w_k \rangle}_{= \begin{cases} 0, & i \neq k \\ 1, & i = k \end{cases}} = \lambda_k \end{aligned}$$

Sei nun $v \in V$. Dann folgt für alle $w \in W$

$$\begin{aligned} &\langle v - \sum_{i=1}^n \langle v, w_i \rangle w_i, w \rangle \\ &= \langle v - \sum_{i=1}^n \langle v, w_i \rangle w_i, \sum_{k=1}^n \langle w, w_k \rangle w_k \rangle \\ &= \sum_{k=1}^n \langle w, w_k \rangle \langle v, w_k \rangle - \sum_{i=1}^n \sum_{k=1}^n \langle v, w_i \rangle \langle w, w_k \rangle \underbrace{\langle w_i, w_k \rangle}_{= \begin{cases} 0, & i \neq k \\ 1, & i = k \end{cases}} \\ &= 0 \end{aligned}$$

Also:

$$\Pi_W(v) = \sum_{i=1}^n \langle v, w_i \rangle w_i$$

- i)** Das folgende konstruktive Verfahren heißt **Gram-Schmidt-Orthonormalisierung**:
 Sei $(\tilde{w}_1, \dots, \tilde{w}_n)$ irgendeine Basis von W . Sei $W_k = \text{span}(\tilde{w}_1, \dots, \tilde{w}_k)$, $k = 1, \dots, n$.
 Definiere:

$$\begin{aligned} w_1 &:= \frac{\tilde{w}_1}{\|\tilde{w}_1\|} \\ w_k &:= \frac{\tilde{w}_k - \Pi_{W_{k-1}}(\tilde{w}_k)}{\|\tilde{w}_k - \Pi_{W_{k-1}}(\tilde{w}_k)\|} \end{aligned}$$

Wir zeigen induktiv: Für alle $k = 1, \dots, n$ ist (w_1, \dots, w_k) eine Orthonormalbasis von W_k (und somit ist nach **ii)** w_{k+1} definiert).

k=1:

$$\text{klar, da } \|w_1\| = \frac{1}{\|\tilde{w}_1\|} \|\tilde{w}_1\| = 1$$

k-1 \rightsquigarrow k: Nach Induktionsannahme ist (w_1, \dots, w_{k-1}) eine orthonormale Familie und $w_l \in W_{k-1}$ für alle $l = 1, \dots, k-1$. Nach Definition der Orthogonalprojektion folgt für alle $l = 1, \dots, k-1$:

$$\langle w_k, w_l \rangle = \frac{1}{\|\tilde{w}_k - \Pi_{W_{k-1}}(\tilde{w}_k)\|} \langle \tilde{w}_k - \Pi_{W_{k-1}}(\tilde{w}_k), \underbrace{w_l}_{\in W_{k-1}} \rangle = 0$$

Da offensichtlich $\|w_k\| = 1$ und (wegen $W_{k-1} \subset W_k$) auch $w_k \in W_k$, ist (w_1, \dots, w_k) eine orthonormale Familie in W_k . Nach Satz 10.11 ist diese Familie linear unabhängig, also eine Basis, da $\dim(W_k) = k$. \square

10.15 Beispiel

$$W = \text{span}(\tilde{w}_1, \tilde{w}_2), \tilde{w}_3 \in \mathbb{R}^4$$

mit

$$\tilde{w}_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tilde{w}_2 = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \end{pmatrix}, \tilde{w}_3 = \begin{pmatrix} 1 \\ 4 \\ -2 \\ 3 \end{pmatrix}$$

Seien $W_1 = \text{span}(\tilde{w}_1)$, $W_2 = \text{span}(\tilde{w}_1, \tilde{w}_2)$. Gram-Schmidt-Orthonormalisierung liefert:

$$w_1 = \frac{\tilde{w}_1}{\|\tilde{w}_1\|} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

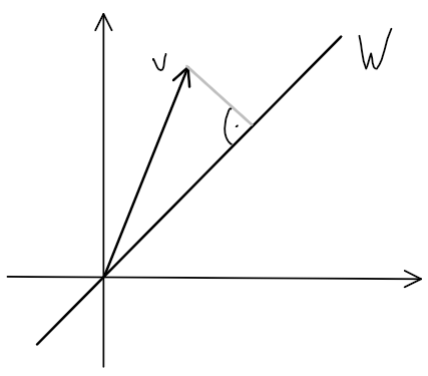
$$\begin{aligned} \tilde{w}_2 - \Pi_{W_1}(\tilde{w}_2) &= \tilde{w}_2 - \langle \tilde{w}_2, w_1 \rangle w_1 \\ &= \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\rangle * \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 3 \\ 0 \end{pmatrix} \end{aligned}$$

nach Satz 10.14(ii).

Also:

$$w_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{aligned} \tilde{w}_3 - \Pi_{W_2}(\tilde{w}_3) &= \tilde{w}_3 - \langle \tilde{w}_3, w_1 \rangle w_1 - \langle \tilde{w}_3, w_2 \rangle w_2 \\ &= \begin{pmatrix} 1 \\ 4 \\ -2 \\ 3 \end{pmatrix} - \underbrace{\left\langle \begin{pmatrix} 1 \\ 4 \\ -2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\rangle}_{=1} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \underbrace{\left\langle \begin{pmatrix} 1 \\ 4 \\ -2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle}_{=-2} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 4 \\ 0 \\ 3 \end{pmatrix}, \end{aligned}$$



da (w_1, w_2) eine Orthonormalbasis von W_2 ist.

Da

$$\left\| \begin{pmatrix} 0 \\ 4 \\ 0 \\ 3 \end{pmatrix} \right\| = \sqrt{0^2 + 4^2 + 0^2 + 3^2} = 5$$

folgt:

$$w_3 = \frac{1}{5} \begin{pmatrix} 0 \\ 4 \\ 0 \\ 3 \end{pmatrix}$$

Also ist

$$\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{5} \begin{pmatrix} 0 \\ 4 \\ 0 \\ 3 \end{pmatrix} \right)$$

eine Orthonormalbasis von W .

Die Orthogonalprojektion hat folgende wichtige Minimierungseigenschaft:

10.16 Satz

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Prä-Hilbertraum mit induzierter Norm $\| \cdot \|$. Sei $W \subset V$ ein endlich dimensionaler Unterraum. Dann existiert zu jedem $v \in V$ genau ein $w_0 \in W$ mit

$$\|v - w_0\| = \inf_{w \in W} \|v - w\|$$

und es gilt: $w_0 = \Pi_{W(v)}$

Beweis

$$\|v - w\|^2 = \|(v - \Pi_{W(v)}) + (\Pi_{W(v)} - w)\|^2$$

$$\begin{aligned}
& \underbrace{=}_P \|v - \Pi_{W(v)}\|^2 + \|\Pi_{W(v)} - w\|^2 + \underbrace{2 \langle v - \Pi_{W(v)}, \underbrace{\Pi_{W(v)} - w}_{\in W} \rangle}_{=0} \\
& = \|v - \Pi_{W(v)}\|^2 + \|\Pi_{W(v)} - w\|^2
\end{aligned}$$

Dieser Ausdruck wird minimal in w , genau dann, wenn $\|\Pi_{W(v)} - w\| = 0$, was gleichbedeutend mit $w = \Pi_{W(v)}$ ist. \square

Anwendung auf lineare Gleichungssysteme

Wir betrachten lineare Gleichungssysteme

$$Ax = b, \quad A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m > n.$$

Wir nehmen an, dass $\text{Rang}(A) = n$. Nach Satz 4.11 gilt:

Entweder hat $Ax = b$ genau eine Lösung $x = x(b)$ oder es existiert keine Lösung. Im letzteren Fall betrachten wir das Problem.

Finde $x_0 \in \mathbb{R}^n$ so, dass

$$\|Ax_0 - b\|$$

minimal wird ($\|\cdot\|$ die vom Euklidischen Skalarprodukt induzierte „Euklidische Norm“ auf \mathbb{R}^m). Es gilt nach Satz 10.16

$$\begin{aligned}
\inf_{x \in \mathbb{R}^n} \|Ax - b\| &= \inf_{y \in \text{Im}(F_A)} \|y - b\| \\
&= \|\Pi_{\text{Im}(F_A)}(b) - b\|
\end{aligned}$$

Sei $b_0 := \Pi_{\text{Im}(F_A)}(b)$. Dann: $b_0 = Ax(b_0)$.

Also:

$$\inf_{x \in \mathbb{R}^n} \|Ax - b\| = \|Ax(b_0) - b\|$$

Wir finden also die „beste Approximation“ x_0 zu $Ax = b$, indem wir zunächst b orthogonal auf $\text{Im}(F_A)$ projizieren und dann die Lösung x_0 zu

$$Ax = \Pi_{\text{Im}(F_A)}(b)$$

bestimmen. Wir geben eine explizite Form für x_0 an. Da

$$\begin{aligned}
A^T Ax = 0 &\Rightarrow x^T A^T Ax = 0 \\
&\Rightarrow (Ax)^T (Ax) = 0 \Rightarrow \|Ax\| = 0 \Rightarrow Ax = 0 \\
&\Rightarrow x = 0,
\end{aligned}$$

ist $A^T A$ invertierbar nach Satz 4.13. Sei $y \in \text{Im}(F_A)$, also $y = Ax$ für $x \in \mathbb{R}^n$. Dann gilt für alle $b \in \mathbb{R}^m$:

$$\langle b, y \rangle = \langle Ax, b \rangle = (Ax)^T b = x^T A^T b = x^T A^T A (A^T A)^{-1} A^T b$$

$$\underbrace{(Ax)^T}_{=y} A(A^T A)^{-1} A^T b = \langle A(A^T A)^{-1} A^T b, y \rangle.$$

Da $y \in \text{Im}(F_A)$ beliebig, folgt

$$\Pi_{\text{Im}(F_A)}(b) = A(A^T A)^{-1} A^T b$$

Man nennt $A^+ := (A^T A)^{-1} A^T$ die **Pseudoinverse** von A. Es ist dann also

$$x_0 = A^+ b$$

die Bestapproximation zu $Ax = b$.

Als Folgerung der obigen Überlegungen halten wir fest:

10.17 Satz

Sei $(\mathbb{R}^m, \langle \cdot, \cdot \rangle)$ der m -dimensionale Euklidische Raum und W ein n -dimensionaler Unterraum des \mathbb{R}^m mit Basis (w_1, \dots, w_n) . Dann gilt für alle $v \in \mathbb{R}^m$:

$$\Pi_W(v) = \sum_{i=1}^n \langle A^+ v, e_i \rangle w_i,$$

wobei $A \in \mathbb{R}^{m \times n}$ als Spalten w_1, \dots, w_n hat und e_i der i -te kanonische Einheitsvektor.

Beweis

Dann ist $W = \text{Im}(F_A)$. Also folgt:

$$\begin{aligned} \Pi_W(v) &= \Pi(\text{Im}(F_A))(v) = AA^+ v \\ &= A \left(\sum_{i=1}^n \langle A^+ v, e_i \rangle e_i \right) = \sum_{i=1}^n \langle A^+ v, e_i \rangle w_i \quad \square \end{aligned}$$

11 Fourierreihen

Ziel:

Approximation einer Funktion durch Überlagerung von Kosinus- und Sinusschwingungen. Dadurch kann ein Signal in die Anteile der einzelnen Frequenzen zerlegt werden. In diesem Kapitel bezeichnet $V = \phi([- \pi, \pi])$ der \mathbb{R} -Vr der stetigen Funktion $f : [- \pi, \pi] \rightarrow \mathbb{R}$ versehen mit dem Skalarprodukt

$$\langle f, g \rangle := \int_{-\pi}^{+\pi} f(x)g(x)dx$$

Für jedes $n \in \mathbb{N}$ sei

$$W_n := \text{span}(1, \cos(\cdot), \cos(2\cdot), \dots, \cos(n\cdot), \sin(\cdot), \sin(2\cdot), \dots, \sin(n\cdot)),$$

wobei etwa $\cos(k\cdot) : [- \pi, \pi] \rightarrow \mathbb{R}, x \mapsto \cos(kx)$. Die Elemente von W_n heißen **trigonometrische Polynome** vom Grad $\leq n$.

11.1 Satz

Für alle $n \in \mathbb{N}$ ist

$$(1, \cos(1\cdot), \dots, \cos(n\cdot), \sin(\cdot), \dots, \sin(n\cdot))$$

eine orthogonale Familie in V und also eine Basis von W_n .

Beweis

Da das System W_n erzeugt und die lineare Unabhängigkeit aus der Orthogonalität folgt (Satz 10.11) ist nur die Orthogonalität zu zeigen.

Man beachte für alle $l \in \mathbb{Z} \setminus \{0\}$

$$\int_{-\pi}^{+\pi} \cos(lx)dx = [\sin(lx) * \frac{1}{l}]_{-\pi}^{\pi} = 0$$

$$\int_{-\pi}^{+\pi} \sin(lx)dx = [\cos(lx) * \frac{-1}{l}]_{-\pi}^{\pi} = 0$$

Ferner folgt aus dem Additionstheoremen für Sinus und Kosinus:

$$\sin((l+m)x) + \sin((l-m)x)$$

$$\begin{aligned}
 &= \sin(lx) \cos(mx) + \cos(lx) \sin(mx) + \sin(lx) \cos(mx) \\
 &= -\sin(mx) \cos(lx) = 2 \sin(lx) \cos(mx)
 \end{aligned}$$

und analog:

$$\cos((l+m)x) + \cos((l-m)x) = 2 \cos(lx) \cos(mx)$$

Also gilt für $l, m > 0$:

$$\langle 1, \cos(m \cdot) \rangle = \int_{-\pi}^{\pi} \cos(mx) dx = 0$$

$$\langle 1, \sin(l \cdot) \rangle = \int_{-\pi}^{\pi} \sin(lx) dx = 0$$

$$\begin{aligned}
 \langle \sin(l \cdot), \cos(m \cdot) \rangle &= \frac{1}{2} \int_{-\pi}^{\pi} (\sin((l+m)x) + \sin((l-m)x)) dx \\
 &= 0
 \end{aligned}$$

(unter beachtung von $\sin(0) = 0$ im Fall $l = m$).

$$\begin{aligned}
 \langle \cos(l \cdot), \cos(m \cdot) \rangle &= \frac{1}{2} \int_{-\pi}^{\pi} (\cos((l+m)x) + \cos((l-m)x)) dx \\
 &= \begin{cases} 0 & l \neq m \\ \pi & l = m \end{cases}
 \end{aligned}$$

(unter Beachtung von $\cos(0) = 1$ im Fall $l = m$).

$$\begin{aligned}
 \langle \sin(l \cdot), \sin(m \cdot) \rangle &= \int_{-\pi}^{\pi} \sin(lx) \sin(mx) dx \\
 &= \int_{-\pi}^{\pi} \cos(lx) \cos(mx) dx - \underbrace{\int_{-\pi}^{\pi} \cos((l+m)x) dx}_{=0} \\
 &= \langle \cos(l \cdot), \cos(m \cdot) \rangle \\
 &= \begin{cases} 0 & l \neq m \\ \pi & l = m \end{cases}
 \end{aligned}$$

11.2 Bemerkung

Berechnet man noch

$$\langle 1, 1 \rangle = 2\pi,$$

so folgt, dass

$$\left(\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos(\cdot), \dots, \frac{1}{\sqrt{\pi}} \cos(n \cdot), \frac{1}{\sqrt{\pi}} \sin(\cdot), \dots, \frac{1}{\sqrt{\pi}} \sin(n \cdot) \right)$$

eine Orthonormalbasis von W_n ist.

Sei $f \in V$ beliebig. Dann gilt nach Satz 10.14 ii):

$$\Pi_{W_n}(f) = \langle f, \frac{1}{\sqrt{2\pi}} \rangle \frac{1}{\sqrt{2\pi}} + \sum_{l=1}^n \langle f, \cos(l \cdot) \cdot \frac{1}{\sqrt{\pi}} \rangle \frac{1}{\sqrt{\pi}} \cos(l \cdot) + \sum_{l=1}^n \langle f, \sin(l \cdot) \frac{1}{\sqrt{\pi}} \rangle \frac{1}{\sqrt{\pi}} \sin(l \cdot) = \frac{a_0(f)}{2} + \sum_{l=1}^n (a_l(f) \cos(l \cdot) + b_l(f) \sin(l \cdot))$$

mit

$$a_l(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(lx) dx, l = 0, \dots, n.$$

$$b_l(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(lx) dx, l = 1, \dots, n.$$

Die Koeffizienten $a_l(f), b_l(f)$ werden **Fourier-Koeffizienten** von f genannt.

Nach Satz 10.16 ist dies die Bestapproximation von f durch trigonometrische Polynome vom Grad $\leq n$ in dem Sinn, dass für alle $g \in W_n$

$$\begin{aligned} \int_{-\pi}^{\pi} \left| f(x) - \left(\frac{a_0(f)}{2} + \sum_{l=1}^n a_l(f) \cos(lx) + b_l(f) \sin(lx) \right) \right|^2 dx \\ \leq \int_{-\pi}^{\pi} |f(x) - g(x)|^2 dx \end{aligned}$$

11.3 Beispiel

Sei $f : [-\pi, \pi] \rightarrow \mathbb{R}, x \mapsto |x|$. Dann: $f \in V$ und f ist gerade, d.h. $f(x) = f(-x)$. Also:

$$\begin{aligned} \int_{-\pi}^0 f(x) \sin(lx) dx &= \int_0^{\pi} f(-x) \sin(-lx) dx \\ &= - \int_0^{\pi} f(x) \sin(lx) dx, \end{aligned}$$

d.h. $b_l(f) = 0$ für alle $l \geq 1$.

Ferner gilt für $l \geq 1$:

$$\begin{aligned} \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(lx) dx &= \frac{2}{\pi} \int_0^{\pi} f(x) \cos(lx) dx \\ &= \frac{2}{\pi} \int_0^{\pi} x \cos(lx) dx \\ &= \frac{2}{\pi} \left(\left[x \frac{\sin(lx)}{l} \right]_0^{\pi} - \int_0^{\pi} \frac{\sin(lx)}{l} dx \right) \\ &= \frac{2}{\pi} \left[\frac{\cos(lx)}{l^2} \right]_0^{\pi} = \frac{2}{\pi} ((-1)^l - 1) \frac{1}{l^2} \end{aligned}$$

11 Fourierreihen

Also:

$$a_l(f) = \begin{cases} \pi & , l = 0 \\ \frac{-4}{\pi l^2} & , l \text{ ungerade} \\ 0 & , l \text{ gerade} \end{cases}$$

Die Bestapproximation von $|x|$ durch ein trigonometrisches Polynom vom Grad n lautet also:

$$\frac{\pi}{2} + \sum_{l=1}^n \frac{2}{\pi l^2} ((-1)^l - 1) \cos(lx)$$

Da $\sum_{l=1}^{\infty} \frac{4}{\pi l^2} < \infty$, folgt aus dem Majorantenkriterium (Mf1), dass für alle $x \in [-\pi, \pi]$ die Reihe

$$\sum_{l=1}^{\infty} \frac{2}{\pi l^2} ((-1)^l - 1) \cos(lx) = - \sum_{l=0}^{\infty} \frac{4}{\pi(2l+1)^2} \cos((2l+1)x)$$

konvergiert. Nach Satz 11.5 und Bemerkung 11.6 (unten) gilt für alle $x \in [-\pi, \pi]$

$$|x| = \frac{\pi}{2} - \sum_{l=0}^{\infty} \frac{4}{\pi(2l+1)^2} \cos((2l+1)x)$$

Für $x = 0$ erhält man die Identität

$$\frac{\pi^2}{8} \sum_{l=0}^{\infty} \frac{1}{(2l+1)^2} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \dots$$

11.4 Definition (Fourierreihe)

Sei $f \in V$. Unter der **Fourierreihe** zu f verstehen wir die Folge von **Fourierpolynomen**

$$(S_n f)_{n \in \mathbb{N}}$$

mit

$$(S_n f)(x) := \frac{a_0(f)}{2} + \sum_{l=0}^n (a_l(f) \cos(lx) + b_l(f) \sin(lx))$$

und im Konvergenzfall auch den Wert der Reihe

$$(Sf)(x) := \lim_{n \rightarrow \infty} (S_n f)(x)$$

$$= \frac{a_0(f)}{2} + \sum_{l=1}^{\infty} (a_l(f) \cos(lx) + b_l(f) \sin(lx)).$$

Mit Techniken aus der Analysis kann man zeigen:

11.5 Satz

Sei $f \in V$. Dann:

- i) Falls für ein $x \in (-\pi, \pi)$ der Grenzwert $\lim_{n \rightarrow \infty} (S_n f)(x)$ existiert, so gilt: $(Sf)(x) = f(x)$, d.h. im Konvergenzfall stellt die Fourierreihe die Funktion dar.
- ii) Falls f an einer Stelle $x \in (-\pi, \pi)$ linksseitig und rechtsseitig dfb. ist, so ex $\lim_{n \rightarrow \infty} (S_n f)(x)$
- iii) $\sqrt{\int_{-\pi}^{\pi} (f(x) - (S_n f)(x))^2 dx} \rightarrow 0 (n \rightarrow \infty)$, d.h. der Abstand zwischen f und $(S_n f)$ in der von $\langle \cdot, \cdot \rangle$ induzierten Norm konvergiert gegen 0.

Beweis

Königsberger, Analysis I, Kapitel 16. \square

11.6 Bemerkung

An den Randpunkten $\pi, -\pi$ gilt: Im Konvergenzfall ist

$$\lim_{n \rightarrow \infty} (S_n f)(-\pi) = \lim_{n \rightarrow \infty} (S_n f)(\pi) = \frac{f(\pi)f(-\pi)}{2}$$

Und Konvergenz liegt vor, falls f linksseitig dfb. in π und rechtsseitig dfb. in $-\pi$ ist. Ist $f(\pi) \neq f(-\pi)$, so zeigt $(S_n f)(x)$ in der „Nähe“ von $-\pi$ und in der „Nähe“ von π ein ausgeprägtes Über- und Unterschwenken, das als Gibbsches Phänomen bekannt ist.

11.7 Bemerkung

Sei $f : [0, \pi] \rightarrow \mathbb{R}$ stetig dfb.. Wir ergänzen f zu einer gerade Funktion auf $[-\pi, \pi]$ durch $f(x) := f(-x), x \in [-\pi, 0)$. Sei $x_j := \frac{j-0,5}{n} * \pi, j = 1, \dots, n$. Dann bilden die Punkte $x_j, j = 1, \dots, n$, eine äquidistante Partitionierung des Intervalls $\left[\frac{0,5\pi}{n}, \pi - \frac{0,5\pi}{n}\right]$.

Sei $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ mit $y_i := f(x_j)$. Dann gilt für „großes“ n :

$$y_j = f(x_j) \approx (S_{n-1} f)(x_j) = \frac{a_0}{2} + \sum_{k=1}^{n-1} a_k \cos\left(k \frac{j-0,5}{n} \pi\right)$$

mit

$$\begin{aligned} a_k &= \frac{2}{\pi} \int_0^{\pi} f(x) \cos(kx) dx \approx \frac{2}{\pi} \sum_{l=1}^n (f(x_l) \cos(kx_l)) \frac{\pi}{n} \\ &= \frac{2}{n} \sum_{l=1}^n y_l \cos\left(k \left(l - 0,5\right) \frac{\pi}{n}\right) =: \hat{a}_k(y) \end{aligned}$$

11 Fourierreihen

Die Abbildung:

$$\mathbb{R}^n \rightarrow \mathbb{R}^n, y \mapsto \begin{pmatrix} \hat{a}_0(y) \\ \vdots \\ \hat{a}_{n-1}(y) \end{pmatrix}$$

wird **diskrete Kosinustransformation** genannt. Man überzeugt sich leicht, dass für alle $y \in \mathbb{R}^n$

$$y_j = \frac{\hat{a}_0(y)}{2} + \sum_{k=1}^n \hat{a}_k(y) \cos\left(\frac{\pi}{n}k\left(j - \frac{1}{2}\right)\right), j = 1, \dots, n.$$

vgl. Beispiel 6.3.

12 Orthogonale Matrizen

12.1 Definition

Eine Matrix $Q \in \mathbb{R}^{n \times n}$ wird **orthogonal** genannt, falls $Q^T Q = E_n$.

Wir bezeichnen die Menge der orthogonalen $n \times n$ -Matrizen mit $O(n)$. Sie heißen **orthogonale Gruppe**

12.2 Bemerkung

- i) Eine orthogonale Matrix Q ist also gerade eine invertierbare Matrix mit $Q^{-1} = Q^T$
- ii) $O(n)$ ist tatsächlich mit der Matrix-Multiplikation als Verknüpfung eine Gruppe (Übung).

12.3 Satz

Sei $Q \in \mathbb{R}^{n \times n}$ und $\langle \cdot, \cdot \rangle$ bezeichne das Euklidische Skalarprodukt auf \mathbb{R}^n . Dann sind äquivalent:

- a) $Q \in O(n)$
- b) Die Spalten von Q sind orthonormal bzgl. $\langle \cdot, \cdot \rangle$.
- c) Für $x, y \in \mathbb{R}^n$ gilt: $\langle Qx, Qy \rangle = \langle x, y \rangle$

Insbesondere erhält Multiplikation mit $Q \in O(n)$ die Länge eines Vektors, d.h. $\|Qx\| = \|x\|$ für alle $x \in \mathbb{R}^n$

Beweis

- a) \Leftrightarrow b) : Seien q_i die Spalten von Q . Sei $Q^T Q = (a_{ij})_{i,j=1,\dots,n}$. Dann: $a_{ij} = \langle q_i, q_j \rangle$, was die Äquivalenz von a) und b) zeigt.

- a) \Leftrightarrow c) Da

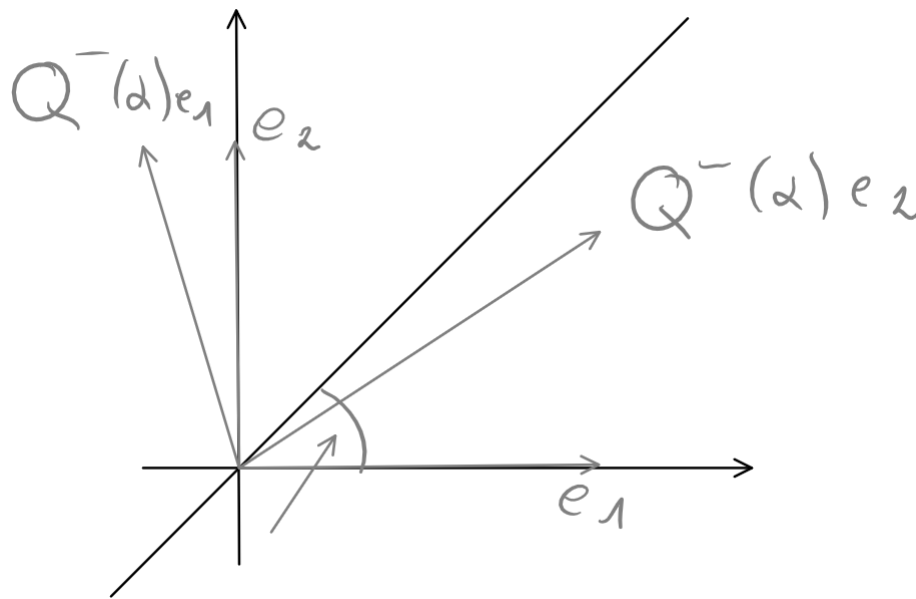
$$\langle Qx, Qy \rangle = (Qx)^T Qy = x^T (Q^T Q)y,$$

folgt c) aus a). Andererseits folgt mit $x = e_i, y = e_j$ aus c):

$$a_{ij} = \langle Qe_i, Qe_j \rangle \underset{c)}{=} \langle e_i, e_j \rangle = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad \square$$

$$Q^+(\alpha) := \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

$$Q^-(\alpha) := \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix}$$



12.4 Beispiel

Sei $n = 2$. Dann sind für alle $\alpha \in [0, 2\pi)$ orthogonal.

Die Abbildung $x \mapsto Q^+(\alpha) * x$ ist eine Drehung um den Winkel α , die Abbildung $x \mapsto Q^-(\alpha) * x$ ist eine Spiegelung an der Geraden mit Winkel $\frac{\alpha}{2}$ zur x_1 -Achse: Wir zeigen, dass dies alle Matrizen in $O(2)$ sind. Sei

$$Q = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in O(2)$$

Dann: 1.) $a^2 + b^2 = 1$, 2.) $c^2 + d^2 = 1$, 3.) $a * c + b * d = 0$

Wegen 1), 2) existiert $\alpha, \alpha' \in [0, 2\pi)$ mit $a = \cos(\alpha)$, $b = \sin(\alpha)$, $c = \cos(\alpha')$, $d = \sin(\alpha')$

Aus 3) folgt:

$$\cos(\alpha - \alpha') = \cos(\alpha) \cos(\alpha') + \sin(\alpha) \sin(\alpha')$$

$$= a * c + b * d = 0$$

Also: $\alpha - \alpha' = \frac{\pi}{2} + k\pi$ für ein $k \in \mathbb{Z}$.

Somit $c = \cos(\alpha') = \cos(\alpha - (\frac{\pi}{2} + k\pi)) = \sin(\alpha) \sin(\frac{\pi}{2} + k\pi)$

$d = \sin(\alpha') = \sin(\alpha - (\frac{\pi}{2} + kr)) = -\cos(\alpha) \sin(\frac{\pi}{2} + kr)$

Also:

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} \sin(\alpha) \\ -\cos(\alpha) \end{pmatrix} \text{ oder } \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}$$

da $\sin(\frac{\pi}{2} + k\pi) \in \{-1, 1\}$. Also: Die Matrizen in $O(2)$ beschreiben genau die Drehungen und Spiegelungen.

Sei $Q \in O(n)$. Dann folgt aus den Rechenregeln für die Determinante:

$$\det(Q)^2 = \det(Q) * \det(Q) = \det(Q^T) * \det(Q)$$

$$\det(Q^T Q) = \det(E_n) = 1$$

Also: $\det(Q) \in \{1, -1\}$. Man definiert daher:

$$SO(n) := O^+(n) := \{Q \in O(n) | \det(Q) = 1\}$$

$$O^-(n) = \{Q \in O(n) | \det(Q) = -1\}$$

$SO(n)$ wird **spezielle orthogonale Gruppe** genannt und ist eine Gruppe bzgl. der Matrixmultiplikation (Übung). Im Fall $n = 2$ entsprechen Elemente in $SO(2)$ gerade den Drehungen und sind also „orientierungserhaltend“. Elemente in $O^-(2)$ entsprechen den Spiegelungen und sind somit „orientierungsumkehrend“.