

Lineare Algebra II
Hubert Grassmann
HU Berlin, Inst. f. Mathematik
22. Februar 2012

$$c_A(z) = \sum b_i z^{n-i}$$

$$b_i = -\frac{1}{i} \sum_{j=0}^{i-1} b_j \operatorname{Spur} (A^{i-j})$$

Inhaltsverzeichnis

0	Vorwort	5
0.1	Kleinigkeiten	5
1	Polynome	7
2	Normalformen von Matrizen	31
2.1	Invariante Unterräume	31
2.2	Nilpotente Endomorphismen	32
2.3	Jordansche Normalform	35
2.4	Rekursive Folgen	38
2.5	Lineare Differentialgleichungssysteme	40
3	Euklidische Vektorräume	43
3.1	Skalarprodukt, Orthonormalbasis	43
3.2	Orthogonale Abbildungen und Matrizen	48
3.3	Die adjungierte Abbildung	52
3.4	Pseudoinverse Matrizen	57
3.5	Unlösbare und unterbestimmte Gleichungssysteme	60
3.6	Überbestimmte Gleichungssysteme: Die Methode der kleinsten Quadrate (Gauß)	62
3.7	Householder-Transformationen	63
3.8	QR-Zerlegung	64
3.9	Hessenberg-Zerlegung	66
3.10	Singularwertzerlegung	69
3.11	Vektor- und Matrixnormen	70
3.12	Positiv definite Matrizen	73
4	Euklidische und projektive Geometrie	77
4.1	Euklidische Geometrie	77
4.2	Sphärische Geometrie	82
4.3	Konvexe Mengen und lineare Ungleichungssysteme	84
4.4	Projektive Geometrie	89
5	Polynommatrizen	97
5.1	Smithsche Normalform	99
5.2	Die rationale Normalform	105

5.3	Lokale Minimalpolynome eines Endomorphismus	106
6	Universelle Konstruktionen; abstract nonsense	109
7	Tensorprodukte	115
8	Halbeinfache Algebren und Moduln	121
8.1	Grundlagen	121
8.2	Darstellungen endlicher Gruppen	131
8.3	Charaktere	136
8.4	Die diskrete Fourier-Transformation	140
9	Zerlegung endlichdimensionaler Algebren	143
10	Das Jacobson-Radikal	153
11	Primzahltest und Faktorisierung ganzer Zahlen	159
12	Boolesche Algebren und Boolesche Funktionen	167
Index		171

Kapitel 0

Vorwort

Zu Beginn des zweiten Semesters werden Polynome behandelt (Kapitel 1): größter gemeinsamer Teiler, Newtonsche Formeln für symmetrische Polynome und als Anwendung eine Rekursionsformel zur Berechnung der Koeffizienten des charakteristischen Polynoms einer Matrix.

Dann haben wir eine Folge von langen Beweisen, als deren Ergebnis die Jordansche Normalform erscheint (Kapitel 2). Zu bemerken ist, daß konsequent auf den Begriff des Faktorraums verzichtet wird, der in der Vektorraumtheorie ja eigentlich auch überflüssig ist. Als Anwendung werden rekursive Folgen behandelt. Es folgt ein umfangreiches Kapitel 3 über Euklidische Vektorräume. Hier wird neben dem Üblichen auf einige für numerische Anwendungen relevante Verfahren eingegangen. Kapitel 4 behandelt einige Fragen der Euklidischen Geometrie und führt in die projektive Geometrie ein. Danach werden Polynommatrizen und deren Normalformen behandelt, ein Thema, das nicht zum Standardumfang der linearen Algebra gehört, aber einen abrundenden Rückblick gestattet (Kapitel 5).

0.1 Kleinigkeiten

Was ist schwieriger: Lesen, Schreiben oder Rechnen ¹

Ich habe mal erfahren, daß $p = 2^{20996011} - 1$ die größte bekannte Primzahl war; sie hat $\log_{10}(p) = 6.3$ Millionen Stellen, das sind 1500 Druckseiten.

1. Berechnung (schnelle Potenzierung) braucht $\log n = \log \log p$ Rechenoperationen mit großen Zahlen. Zeit: 89 min.
2. Schreiben (Umwandlung in einen String aus 0,1,...,9) braucht $\log_{10}(p)$ Divisionen langer Zahlen durch 10 (es nützt nichts, daß 10 „klein“ ist). Zeit: 919 min = 15 h
3. Lesen (Umwandlung String in Bytefolge) braucht mit dem Horner Schema $n = \log_{10}(p)$ Operationen, aber im Wesentlichen Multiplikationen mit der kleinen Konstanten 10. Zeit: 63 min

¹Grundschüler werden gefragt: „Was macht mehr Spaß?“

Am 9.6.2004 stand in der FAZ, daß $2^{24036583} - 1$ die größte bekannte Primzahl ist, sie hat 7 235 733 Stellen; ich habe sie in 190 Minuten berechnet.

Quadratzahlen und Quadratwurzeln in \mathbb{N}

Es sei $n \in \mathbb{N}$ gegeben, wir suchen die natürliche Zahl x mit $x^2 \leq n < (x+1)^2$, sie kann mit dem folgenden Algorithmus bestimmt werden:

```
x = n;
do
  y = (x+n/x)/2;
  z = x;
  x = y;
while (y < z);
return z;
```

Zum Beweis der Korrektheit zeigen wir zunächst, daß

$$y = \left(x + \frac{n}{x}\right)/2 \geq \sqrt{n}$$

für *alle* x gilt. Dies ist genau dann der Fall, wenn die folgenden äquivalenten Ungleichungen erfüllt sind:

$$\begin{aligned} x^2 + 2n + \frac{n^2}{x^2} &\geq 4n \\ x^2 + \frac{n^2}{x^2} &\geq 2n \\ x^4 + n^2 &\geq 2nx^2 \\ x^4 - 2nx^2 + n^2 &= (x^2 - n)^2 \geq 0 \end{aligned}$$

Die letzte Ungleichung ist klar. Also ist das Ergebnis des Algorithmus nicht kleiner als \sqrt{n} .

Sein nun $q = \sqrt{n}$. Wenn $x > q$ ist, so ist $x \geq q + 1$. Wir betrachten das y von oben:

$$y - x = \left\lfloor \frac{x + n/x}{2} \right\rfloor - x = \left\lfloor \frac{n/x - x}{2} \right\rfloor = \left\lfloor \frac{n - x^2}{2x} \right\rfloor$$

Wegen $x \geq q + 1 > \sqrt{n}$ gilt $n - x^2 < 0$, also $y - x < 0$, ein Widerspruch.

Nun stellen wir uns die Frage, ob eine natürliche Zahl n eine Quadratzahl ist. Wenn das der Fall ist, so ist auch die Restklasse von n in jedem Ring $\mathbb{Z}/m\mathbb{Z}$ ein Quadrat. Die Zahl der Quadrate modulo $m = 64, 63, 65, 11$ ist 12, 16, 21, 6. Die Wahrscheinlichkeit, daß ein Quadrat nicht entdeckt wird, ist also

$$\frac{12}{64} \cdot \frac{16}{63} \cdot \frac{21}{65} \cdot \frac{11}{6} = \frac{6}{715} < 0.01.$$

Wir merken uns, welche Restklassen Quadrate sind, indem wir in vier Feldern

q11, q63, q64, q65 jeweils an den Stellen $k^2 \bmod m, k = 0 \dots m/2$ Einsen eintragen.

Num geht es schnell: Wir setzen $t = n \bmod 2882880$ (das ist das Produkt unserer Moduln) und prüfen nacheinander, ob im Feld qm an der Stelle $t \bmod m$ eine Null steht. Wenn dies der Fall ist, wissen wir, daß n kein Quadrat ist und brechen ab. Sonst berechnen wir $q = \lfloor \sqrt{n} \rfloor$; wenn $q^2 = n$ ist, so ist n ein Quadrat, sonst nicht.

Kapitel 1

Polynome

Wir verlassen für ein paar Augenblicke die lineare Algebra und stellen uns einige Hilfsmittel zusammen, die wir später verwenden werden.

Ein Term der Form $f(z) = a_0z^n + a_1z^{n-1} + \dots + a_n$ heißt ein Polynom in z , die Zahlen a_0, \dots, a_n heißen die Koeffizienten des Polynoms. Die Summe $f + g$ zweier Polynome f und g ist wieder ein Polynom in z , ebenso ihr Produkt fg . Wenn in der obigen Darstellung der Koeffizient a_0 von Null verschieden ist, so sagt man, das Polynom f habe den Grad n , als Abkürzung schreiben wir $\deg(f) = n$. Die Menge aller Polynome in z mit Koeffizienten aus R bezeichnet man mit $R[z]$. Sei $\deg(f) = n, \deg(g) = m$. Bitte überlegen Sie sich, daß $\deg(f + g) \leq \max(m, n)$ und $\deg(fg) = n + m$ ist.

Werte berechnen

Für das Polynom $f(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n$ soll der Wert $f(b)$ berechnet werden.

Die naive Methode geht so:

$$\begin{aligned} w &= a_n \\ \text{für } i &= 1, \dots, n : w = w + a_{n-i} \cdot b^i \end{aligned}$$

Aufwand: Es sind $1 + 2 + \dots + n = \frac{n(n+1)}{2} \sim n^2$ Multiplikationen nötig.

Das Schema von Horner (1819) benötigt nur n Multiplikationen, denn wegen

$$f(b) = (\dots((a_0b + a_1)b + a_2)b + \dots + a_{n-1})b + a_n$$

kann man so rechnen:

$$\begin{aligned} w &= a_0 \\ \text{für } i &= 1, \dots, n : \\ w &= w \cdot b + a_i \end{aligned}$$

Potenzieren

Dieselbe Idee kann man verwenden, um (große) Potenzen einer Zahl zu berechnen. Um x^n zu berechnen, stellen wir den Exponenten im Binärsystem dar:

$$n = (b_k \dots b_0)_2 = \sum b_i 2^i, \quad b_i \in \{0, 1\}.$$

Wir sehen also, daß n der Wert eines gewissen Polynoms an der Stelle 2 ist. Ein Summand (1) verlangt eine Multiplikation, ein Faktor (2) verlangt eine Quadrierung.

$$x^n = x^{\sum b_i 2^i} = \prod x^{b_i 2^i} = \prod_{b_i \neq 0} x^{2^i}$$

```

y = 1
w = x
solange n >= 0:
    wenn n ungerade:
        y = y * w
    wenn n > 1:
        w = w * w
    n = n / 2

```

Das Ergebnis ist y

Dies entspricht der Auswertung des Exponenten mittels des Horner-Schemas: Wir stellen uns die binär dargestellte Zahl als Polynom vor, dessen Koeffizienten Nullen und Einsen sind, in das die Zahl 2 eingesetzt ist.

Definition: Das Polynom g heißt Teiler des Polynoms f , wenn ein Polynom h existiert, so daß $f = gh$ ist.

Wenn $r \neq 0$ eine Zahl (also ein Polynom vom Grade 0) ist, so gibt es immer ein Polynom h mit $f = rh$. Gerade diesen trivialen Fall wollen wir immer ausschließen, wenn wir von Teilbarkeit sprechen.

Zum Beispiel hat das Polynom $f(z) = z^2 + pz + q$ die (nichtkonstanten) Teiler

$$z + p/2 \pm (p^2/4 - q)^{1/2}.$$

Nicht zu unterschätzen ist das folgende Lemma über die Division mit Rest:

Lemma 1.0.1 *Seien f, g Polynome, $g \neq 0$. Dann gibt es Polynome q und r , so daß $f = gq + r$ gilt und entweder $r = 0$ oder $\deg(r) < \deg(g)$ ist.*

Das Polynom $q(x)$ heißt der Quotient, $r(x)$ der Rest der Polynomdivision.

Beweis: Wenn $\deg(g) > \deg(f)$ ist, so setzen wir $q = 0$ und $r = f$. Weiterhin sei $\deg(f) \geq \deg(g)$. Wir führen die Induktion über $n = \deg(f)$.

Ohne Beschränkung der Allgemeinheit können wir annehmen, daß die höchsten Koeffizienten von f und g gleich 1 sind (solche Polynome heißen „normiert“).

Sei also $n = 1$, d.h. $f(z) = z + a$. Dann ist $g(z) = z + b$ oder $g(z) = 1$, im ersten Fall wählen wir $q(z) = 1, r(z) = a - b$ und im zweiten Fall $q(z) = f(z), r(z) = 0$.

Wir setzen nun voraus, daß das Lemma für Polynome von einem Grad, der kleiner als n ist, bewiesen ist. Sei also

$$f(z) = z^n + a_1 z^{n-1} + \dots, \quad g(z) = z^m + b_1 z^{m-1} + \dots,$$

dann setzen wir $q_1(z) = z^{n-m}$, es ist

$$q_1(z)g(z) = z^n + b_1z^{n-1} + \dots$$

und das Polynom

$$f_1(z) = f(z) - q_1(z)g(z) = (a_1 - b_1)z^{n-1} + \dots$$

hat einen Grad, der kleiner als n ist, also gibt es Polynome $q_2(z)$ und $r(z)$ mit $f_1 = q_2g + r$ und wir wissen, daß $r = 0$ oder $\deg(r) < \deg(g)$ gilt. Dann haben wir mit

$$f = f_1 + q_1g = (q_2 + q_1)g + r$$

die gewünschte Zerlegung gefunden. □

Falls nun $g(x) = x - b$ ein lineares Polynom ist, dann ist der Divisionsrest eine Konstante, und zwar gleich $f(b)$, wie man aus

$$f(x) = (x - b)q(x) + r$$

durch Einsetzen von $x = b$ sieht.

Das wäre eine neue Methode der Wertberechnung. Wie effektiv ist sie?

Wir werden gleich sehen, daß man mittels einer kleinen Modifizierung des Horner-Schemas den Quotienten und den Rest der Polynomdivision berechnen kann.

Sei wieder $f(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n$. In unserem Fall hat der Quotient den Grad $n - 1$; wir machen den Ansatz

$$q(x) = b_0x^{n-1} + \dots + b_{n-k-1}x^k + \dots + b_kx^{n-k-1} + \dots + b_{n-1}$$

und vergleichen in

$$f(x) = q(x)(x - b) + b_n$$

die Koeffizienten der x -Potenzen (den Rest nennen wir b_n ; gelegentlich ist eine geschickte Notation hilfreich):

$$\begin{aligned} a_0 &= b_0 \\ x^{n-k} : \quad a_k &= b_k - b_{k-1}b \end{aligned}$$

Damit erhalten wir schrittweise die Koeffizienten von $q(x)$:

$$b_0 = a_0$$

$$b_k = a_k + b_{k-1}b, \quad k = 1, \dots, n$$

und $b_n = f(b)$ ist der Divisionsrest. Der Rechenaufwand ist derselbe wie beim Horner-Schema.

Wenn man eine reelle Nullstelle b des Polynoms $f(x)$ bestimmt hat, so kann man dieses Verfahren nutzen, um den Faktor $x - b$ aus $f(x)$ zu entfernen und kann nach Nullstellen des Quotienten suchen.

Wenn $a + bi$ eine komplexe Nullstelle von $f(x)$ ist, so ist auch $a - bi$ eine Nullstelle von $f(x)$, d.h. $f(x)$ ist durch das quadratische (reelle) Polynom $(x - a - bi)(x - a + bi) = x^2 - 2ax + a^2 + b^2$ teilbar. Um diesen Faktor aus $f(x)$ zu entfernen, verwenden wir eine Verallgemeinerung des Hornerchemas (Collatz 1940):

Wir berechnen also den Quotienten und den Rest bei der Division von $f(x)$ durch $x^2 + px + t$. Der Quotient sei

$$q(x) = b_0x^{n-2} + b_1x^{n-3} + \dots + b_{n-2},$$

der Rest sei

$$r(x) = b_{n-1}x + b_n.$$

Der Koeffizientenvergleich in $f(x) = (x^2 + px + t)q(x) + r(x)$ liefert

$$\begin{aligned} a_0 &= b_0 \\ a_1 &= b_1 + pb_0 \\ &\dots \\ a_k &= b_k + pb_{k-1} + tb_{k-2}, \quad k = 2, \dots, n-1 \\ &\dots \\ a_n &= tb_{n-2} + b_n \end{aligned}$$

also

$$\begin{aligned} b_0 &= a_0 \\ b_1 &= a_1 - pb_0 \\ &\dots \\ b_k &= a_k - pb_{k-1} - tb_{k-2}, \quad k = 2, \dots, n-1 \\ &\dots \\ b_n &= a_n - tb_{n-2} \end{aligned}$$

Um die vorgestellten Verfahren zur Nullstellenabtrennung nutzen zu können, muß man zuerst eine Nullstelle kennen. Im Buch von *Beresin und Shidkov, Numerische Methoden 2, VEB Deutscher Verlag der Wissenschaften, Berlin 1971* wird auf S. 169 ein Verfahren von Muller (1956) vorgestellt, mit dessen Hilfe komplexe Nullstellen von Polynomem berechnet werden können. Der Witz besteht darin, daß (anders als beim Newton-Verfahren) keine Anfangsnäherung bekannt sein muß; das Verfahren ist zwar theoretisch nicht streng abgesichert, liefert aber in der Praxis gute Ergebnisse.

Für das Verfahren von Graeffe zur Bestimmung reeller Nullstellen habe ich z.Z. nur die Referenz auf *Zurmühl, Praktische Mathematik für Ingenieure und Physiker, Springer 1957, S. 62*). Zur Bestimmung komplexer Nullstellen gibt es Verfahren von Nickel bzw. Laguerre, die z.B. bei *Gander, Computermathematik, Birkhäuser 1992, S. 110 ff* vorgestellt werden.

Wir sahen: Wenn $f(a) = 0$ ist, so ist $z - a$ ein Teiler von $f(z)$.

Definition: Seien f_1 und f_2 Polynome. Ein Polynom d heißt größter gemeinsamer Teiler von f_1 und f_2 , wenn gilt:

1. d ist ein Teiler von f_1 und von f_2 (also ein gemeinsamer Teiler),
 2. wenn h irgendein gemeinsamer Teiler von f_1 und f_2 ist, so ist h ein Teiler von d .
- Als Abkürzung schreiben wir $d = \text{ggT}(f_1, f_2)$.

Trivial ist das

Lemma 1.0.2 *Der größte gemeinsame Teiler zweier Polynome ist bis auf einen konstanten Faktor eindeutig bestimmt.*

Beweis: Die Polynome d_1 und d_2 mögen die Bedingungen der Definition erfüllen. Aus 2. folgt, daß Polynome p und q existieren, so daß $d_1 = pd_2$ und $d_2 = qd_1$. Daraus folgt $\deg(p) = \deg(q) = 0$. \square

Zur Berechnung des größten gemeinsamen Teilers zweier Polynome benutzen wir den Euklidischen Algorithmus:

Seien f_1, f_2 gegeben, wir dividieren fortlaufend mit Rest, bis die Division aufgeht:

$$f_1 = q_1 f_2 + f_3$$

$$f_2 = q_2 f_3 + f_4$$

$$f_3 = q_3 f_4 + f_5$$

...

$$f_{m-3} = q_{m-3} f_{m-2} + f_{m-1}$$

$$f_{m-2} = q_{m-2} f_{m-1}$$

Wegen $\deg(f_2) > \deg(f_3) > \deg(f_4) > \dots$ muß nach endlich vielen Schritten ein Rest gleich Null sein, hier ist es f_m .

Behauptung: $\text{ggT}(f_1, f_2) = f_{m-1}$.

Beweis:

1. Klar ist, daß f_{m-2} von f_{m-1} geteilt wird. Weiter ist

$$f_{m-3} = (q_{m-3} q_{m-2} + 1) f_{m-1}$$

durch f_{m-1} teilbar. Jetzt haben wir den Anfang in der Hand: Schauen Sie sich die obigen Gleichungen von der letzten bis zur ersten an! Das Polynom f_{m-1} teilt die beiden f 's auf der rechten Seite, damit aber auch das f mit kleinerem Index auf der linken Seite. Am Ende sehen wir, daß f_{m-1} sowohl f_1 als auch f_2 teilt.

2. Sei h ein gemeinsamer Teiler von f_1 und f_2 . Es ist

$$f_3 = f_1 - q_1 f_2,$$

also ist h ein Teiler von f_3 . So schrauben wir uns an den Indizes nach oben und erhalten zum Schluß, daß h das Polynom f_{m-1} teilt. \square

Beispiel:

```

+2x^6+2x^5-3x^4+1x^3+1x^2-4x+2
+1x^5-1x^4-4x^3-2x^2-2x+3
quot +1x+2
rest +9/2x^4+21/2x^3+13/2x^2-1x-5
Rest +1x^4+7/3x^3+13/9x^2-2/9x-10/9
quot +1x-10/3
rest +7/3x^3+82/27x^2-44/27x-19/27
Rest +1x^3+82/63x^2-44/63x-19/63
quot +1x+65/63
rest +3175/3969x^2+3175/3969x-3175/3969
Rest +1x^2+1x-1
quot +1x+19/63
rest NULL
ggT +1x^2+1x-1

```

Lemma 1.0.3 Sei $d = ggT(f_1, f_2)$, dann gibt es Polynome g_1, g_2 mit $f_1g_1 + f_2g_2 = d$.

Beweis: Wir lesen die obigen Gleichungen von rechts nach links und von unten nach oben und sehen: Das Polynom f_i läßt sich aus f_{i-1} und f_{i-2} kombinieren. Also läßt sich f_{m-1} aus f_1 und f_2 mit gewissen Faktoren kombinieren. \square

Interessanter ist das

Lemma 1.0.4 Der größte gemeinsame Teiler von f_1 und f_2 ist das (normierte) Polynom d von minimalem Grad, für das Polynome g_1 und g_2 existieren, so daß $f_1g_1 + f_2g_2 = d$ ist.

Beweis: Sei $d = f_1g_1 + f_2g_2$ und $\deg(d)$ minimal.

1. Wir dividieren mit Rest:

$$f_1 = q_1d + r_1 = q_1f_1g_1 + q_1f_2g_2 + r_1,$$

also

$$r_1 = f_1(1 - q_1g_1) - f_2q_1g_2,$$

aber wegen $\deg(r_1) < \deg(d)$ ist dies ein Widerspruch zur Minimalität des Grades von d , also ist $r = 0$. d.h d teilt f_1 .

2. Sei h ein gemeinsamer Teiler von f_1 und f_2 , dann ist h auch ein Teiler von $f_1g_1 + f_2g_2 = d$. \square

Wir wollen uns nun mit einem klassischen Verfahren der näherungsweise Nullstellenberechnung befassen.

Das einfachste Näherungsverfahren ist die Newton-Interpolation: Wenn für eine reelle Zahl a der Wert von $f(a)$ beinahe Null ist, so ist $a - f(a)/f'(a)$ eine bessere Näherung. Eine Voraussetzung dafür ist aber, daß f'' in der Nähe des Startpunkts der Iteration sein Vorzeichen nicht ändert. Man muß also erst einmal in die Nähe einer Nullstelle geraten.

Als erstes kann man ein Intervall angeben, in dem alle Nullstellen liegen müssen.

Satz 1.0.1 (Cauchysche Ungleichung)

Sei $f(x) = \sum a_i x^{n-i} \in \mathbb{C}[x]$, dann gilt für jede Nullstelle z von $f(x)$ die Ungleichung

$$|z| < 1 + \frac{\max(|a_1|, \dots, |a_n|)}{|a_0|}.$$

Beweis: Sei $h = \max(|a_1|, \dots, |a_n|)$ und $f(z) = 0$, dann ist

$$a_0 z^n = -a_1 z^{n-1} - \dots - a_n,$$

$$|a_0| |z|^n \leq h \cdot (|z|^{n-1} + \dots + 1) < \frac{h \cdot |z|^n}{|z| - 1},$$

also $|a_0| \cdot (|z| - 1) < h$. □

Als nächstes behandeln wir ein Verfahren, das es gestattet, die Anzahl der Nullstellen eines Polynoms in einem gegebenen Intervall zu berechnen. Durch fortgesetzte Intervallteilung kann man jede einzelne Nullstelle dann beliebig genau einschachteln.

Der Sturmsche Lehrsatz

Sei $f(x) \in \mathbb{R}[x]$ und $a \in \mathbb{R}$. Wie oben haben wir

$$f(x) = (x - a)q(x) + f(a),$$

also

$$q(x) = \frac{f(x) - f(a)}{x - a}.$$

Wenn a eine einfache Nullstelle von $f(x)$ ist, so ist

$$q(a) = f'(a) \neq 0$$

und dies gilt in einer Umgebung von a . Es gibt zwei Fälle:

- 1) $f'(a) > 0$, dann ist $f(x)$ bei a monoton wachsend; wenn x von links durch a hindurch läuft, ist $f(x)$ zuerst negativ, dann positiv.
- 2) $f'(a) < 0$, dann ist $f(x)$ bei a monoton fallend; wenn x von links durch a hindurch läuft, ist $f(x)$ zuerst positiv, dann negativ.

Beiden Fällen ist folgendes gemeinsam: Wenn x wachsend durch a läuft, gehen die Funktionen $f(x)$ und $f'(x)$ von verschiedenen zu gleichen Vorzeichen über.

Wir konstruieren nun eine „Sturmsche Kette“, wir setzen $f_1(x) = f'(x)$ und dividieren fortlaufend mit Rest:

$$\begin{aligned} f &= q_1 \cdot f_1 - f_2 \\ f_1 &= q_2 \cdot f_2 - f_3 \\ &\dots \\ f_{i-1} &= q_i \cdot f_i - f_{i+1} \\ &\dots \\ f_m &= c \text{ konstant. Beachten Sie das Minuszeichen.} \end{aligned}$$

Wenn $f(x)$ nur einfache Nullstellen besitzt, ist $\text{ggT}(f, f_1) = 1$, also $f_m \neq 0$.

Nun gilt: Für keine Zahl $r \in \mathbb{R}$ gibt es ein i mit $f_i(r) = 0 = f_{i+1}(r)$, denn sonst wäre $f(r) = 0 = f'(r)$, also r eine mehrfache Nullstelle. Sei r eine reelle Zahl, sei $w(r)$ die Zahl der Indizes i mit $\text{sgn}(f_i(r)) \neq \text{sgn}(f_{i+1}(r))$, dies ist die Zahl der Vorzeichenwechsel in der Sturmschen Kette $f(r), f_1(r), \dots, f_m(r)$.

Satz 1.0.2 Die Zahl der im Intervall (a, b) gelegenen Nullstellen von $f(x)$ ist gleich $w(a) - w(b)$.

Beweis: Wenn r die x -Achse entlangläuft, so kann sich $w(r)$ nur dann ändern, wenn für ein i gilt $f_i(r) = 0$. Dann ist aber

$$f_{i-1}(r) = -f_{i+1}(r).$$

Schauen wir uns die Werte von f_{i-1}, f_i und f_{i+1} in der Nähe von r an:

	f_{i-1}	f_i	f_{i+1}
$r - \epsilon$	$+t$	p	$-t$
r	$+t$	0	$-t$
$r + \epsilon$	$+t$	$-p$	$-t$

Dabei sei $p, t \in \{\pm 1\}$; egal, welchen Wert p hat, p oder $-p$ stimmt mit einem ihrer Nachbarn überein. Beim Durchgang durch r findet also hier gar keine Änderung der Wechselzahl statt. Eine Änderung der Wechselzahl sich kann also nur beim Durchgang durch eine Nullstelle von f ereignen, und oben haben wir gesehen, das dort ein Vorzeichenwechsel zwischen f und f' verlorenght.

Schauen wir uns das in einem Beispiel an:

$$\begin{aligned} f(x) &= x^5 - 4x - 2 \\ f_1(x) &= 5x^4 - 4 \\ x^5 - 4x - 2 &= (5x^4 - 4) \cdot x/5 - 16/5x + 2, \\ \text{wir setzen } f_2(x) &= 8x + 5 \\ f_3(x) &> 0 \end{aligned}$$

x	f	f_1	f_2	f_3	$w(x)$
-2	-	+	-	+	3
-1	+	+	-	+	2
0	-	-	+	+	1
1	-	+	+	+	1
2	+	+	+	+	0

Also liegen zwischen -2 und -1, zwischen -1 und 0 sowie zwischen 1 und 2 je eine Nullstelle.

Die folgende Konstruktion erlaubt es festzustellen, ob zwei Polynome gemeinsame Nullstellen besitzen.

Resultante und Diskriminante

Folgerung 1.0.1 Die Polynome $f(x)$ und $g(x)$ besitzen genau dann gemeinsame Nullstellen, wenn $\text{Res}(f, g) = 0$ ist.

Beispiel: $f(x) = x^2 - 1$, $g(x) = x^2 + 2x + 1$

$$\begin{vmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{vmatrix} = 0$$

Wir wollen uns nun genauer mit dem Zusammenhang zwischen den Koeffizienten und den Nullstellen eines Polynoms befassen. Sei

$$f(x) = x^n + b_1x^{n-1} + \dots + b_{n-1}x + b_n = (x - x_1) \dots (x - x_n)$$

ein Polynom mit den Koeffizienten b_i und den Nullstellen x_i .

Die Formeln von Viète lauten

$$\begin{aligned} b_1 &= -(x_1 + \dots + x_n) \\ b_2 &= x_1x_2 + \dots + x_{n-1}x_n \\ b_k &= \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} x_{i_1} \dots x_{i_k} \\ b_n &= (-1)^n x_1 \dots x_n \end{aligned}$$

Diese Ausdrücke heißen die elementarsymmetrischen Funktionen der x_i . Ich habe in einem Algebra-Buch aus dem 18. Jahrhundert gelesen, daß man mit Hilfe dieser Formeln „leicht“ die Nullstellen aus den Koeffizienten berechnen könnte; das ist aber nicht ausgeführt worden: es geht ja auch nicht.

Seien $f(x) = \prod(x - x_i)$, $g(x) = \prod(x - y_j)$ normierte Polynome, dann sind die Koeffizienten von f bzw. von g die elementarsymmetrischen Funktionen der Nullstellen x_i bzw. y_j , also ist auch die Resultante ein Polynom in den x_i und y_j . Da die Resultante stets verschwindet, wenn ein x_i gleich einem y_j ist, ist sie durch alle Differenzen $(x_i - y_j)$ teilbar, und wenn wir uns Grad und höchsten Koeffizienten genauer anschauen, finden wir

$$\text{Res}(f, g) = \prod (x_i - y_j).$$

Die Resultante von f und f' wird als Diskriminante von f bezeichnet, sie verschwindet genau dann, wenn f mehrfache Nullstellen besitzt.

Wir wissen, daß $f(x)$ durch die Linearfaktoren $x - x_i$ teilbar ist. Wie sehen die Koeffizienten von $\frac{f(x)}{x - x_i}$ aus?

Satz 1.0.4
$$\frac{\sum_{i=0}^n b_{n-i}x^i}{x - x_1} = \sum_{i=0}^{n-1} x^{n-1-i} \sum_{j=0}^i b_j x_1^{i-j}.$$

Beweis: Wir multiplizieren $\sum_{i=0}^{n-1} x^{n-1-i} \sum_{j=0}^i b_j x_1^{i-j}$ und $x - x_1$ miteinander und stellen fest, ob $f(x)$ herauskommt.

$$\begin{aligned}
& \left(\sum_{i=0}^{n-1} x^{n-1-i} \sum_{j=0}^i b_j x_1^{i-j} \right) \cdot (x - x_1) \\
&= \sum_{i=0}^{n-1} x^{n-i} \sum_{j=0}^i b_j x_1^{i-j} - \sum_{i=0}^{n-1} x^{n-1-i} \sum_{j=0}^i b_j x_1^{i-j+1} \\
&= \sum_{i=0}^{n-1} x^{n-i} \sum_{j=0}^i b_j x_1^{i-j} - \sum_{k=1}^n x^{n-k} \sum_{j=0}^{k-1} b_j x_1^{k-j} \quad (k := i + 1) \\
&= \sum_{i=1}^{n-1} x^{n-i} \left(\underbrace{\sum_{j=0}^i b_j x_1^{i-j} - \sum_{j=0}^{i-1} b_j x_1^{i-j}}_{= b_i} \right) + x^n + b_n - \sum_{j=0}^{n-1} b_j x_1^{n-j} - b_n \\
&= \sum_{i=0}^n x^{n-i} b_i + \sum_{i=0}^n x_1^{n-i} b_i \\
&= f(x) - f(x_1) = f(x). \quad \square
\end{aligned}$$

Abkürzung: Sei $f(x)$ ein Polynom vom Grade n mit den Nullstellen x_1, \dots, x_n . Wir setzen

$$\begin{aligned}
s_0 &= n, \\
s_1 &= x_1 + \dots + x_n \\
s_2 &= x_1^2 + \dots + x_n^2, \\
&\dots \\
s_i &= x_1^i + \dots + x_n^i.
\end{aligned}$$

Die Zahl s_i heißt die i -te Potenzsumme der x_j .

Wir stellen nun eine Beziehung zwischen den Potenzsummen der Wurzeln und den Koeffizienten des Polynoms auf, dies sind die sogenannten Newtonschen Formeln.

Satz 1.0.5 $b_i = -\frac{1}{i} \sum_{j=0}^{i-1} b_j s_{i-j}$

Beweis: Es ist $f(x) = \prod (x - x_i)$. Wir betrachten die Ableitung $f'(x)$ von $f(x)$:

$$\begin{aligned}
f'(x) &= \sum_k \prod_{i \neq k} (x - x_i) = \sum \frac{f(x)}{x - x_k} \\
&= \sum_{k=1}^n \sum_{i=0}^{n-1} x^{n-i-1} \sum_{j=0}^i b_j x_k^{i-j}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^{n-1} \sum_{j=0}^i x^{n-i-1} b_j \sum_k x_k^{i-j} \\
&= \sum_{i=0}^{n-1} \sum_{j=0}^i x^{n-i-1} b_j s_{i-j}.
\end{aligned}$$

Andererseits gilt

$$f'(x) = \sum_i x^{n-i-1} (n-i) b_i$$

und durch Koeffizientenvergleich erhalten wir

$$(n-i)b_i = \sum_{j=0}^i b_j s_{i-j} = \sum_{j=0}^{i-1} b_j s_{i-j} + n b_i,$$

und daraus ergibt sich die Behauptung. \square

Wir kehren nun doch nach diesem Seitensprung wieder zu unseren lieben Matrizen zurück. Aber wir wenden das Gelernte an:

Lemma 1.0.5 *Seien z_1, \dots, z_n die Eigenwerte der Matrix A , dann ist $s_i = Sp(A^i)$.*

Beweis: A^i hat die Eigenwerte z_1^i, \dots, z_n^i und die Spur einer Matrix ist die Summe ihrer Eigenwerte. \square

Nun können wir die Newtonschen Formeln verwenden, um die Koeffizienten des charakteristischen Polynoms einer Matrix zu bestimmen, ohne irgendwelche Determinanten ausrechnen zu müssen.

Folgerung 1.0.2 *Sei A eine Matrix und $c_A(z) = \sum b_i z^{n-i}$ ihr charakteristisches Polynom. Dann ist $b_i = -\frac{1}{i} \sum_{j=0}^{i-1} b_j Sp(A^{i-j})$.* \square

Beispiel:

$$A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}, \quad A^2 = \begin{pmatrix} -2 & 0 & -6 \\ 3 & 4 & 3 \\ 3 & 0 & 7 \end{pmatrix}, \quad A^3 = \begin{pmatrix} -6 & & \\ & 8 & \\ & & 15 \end{pmatrix}$$

$$b_1 = -Sp(A) = -5, \quad b_2 = -\frac{1}{2}(b_0 Sp(A^2) + b_1 Sp(A)) = 8,$$

$$b_3 = -\frac{1}{3}(b_0 Sp(A^3) + b_1 Sp(A^2) + b_2 Sp(A)) = -4 (= \det(A))$$

Wenn wir beachten, daß die Zuordnung $A \mapsto Sp(A)$ eine lineare Abbildung ist, so können wir die obige Formel auch so schreiben:

$$b_i = -\frac{1}{i} Sp\left(\sum_{j=1}^{i-1} b_j A^{i-j}\right),$$

oder schrittweise:

$$\begin{aligned} b_1 &= -Sp(A) \\ b_2 &= -\frac{1}{2}Sp((A + b_1E)A) \\ b_3 &= -\frac{1}{3}Sp(((A + b_1E)A + b_2E)A) \\ b_4 &= -\frac{1}{4}Sp((((A + b_1E)A + b_2E)A + b_3E)A) \end{aligned}$$

Das funktioniert also wie beim Horner-Schema.

Wir führen das mit der Matrix

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

durch; bei ersten Faktor wurde das entsprechende Vielfache der Einheitsmatrix jeweils addiert:

$$b_1 = -Sp(A) = -4$$

$$b_2 = -\frac{1}{2}Sp\left(\begin{pmatrix} -3 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}\right) = -\frac{1}{2}\begin{pmatrix} -2 & -2 & 1 & 0 \\ -2 & -1 & -2 & 1 \\ 1 & -2 & 2 & -2 \\ 0 & 1 & -2 & 1 \end{pmatrix} = 3$$

$$b_3 = -\frac{1}{3}Sp\left(\begin{pmatrix} 1 & -2 & 1 & 0 \\ -2 & -1 & -2 & 1 \\ 1 & -2 & 2 & -2 \\ 0 & 1 & -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}\right) = -\frac{1}{3}\begin{pmatrix} -1 & 0 & -1 & 1 \\ 0 & -2 & 1 & -1 \\ -1 & 1 & -2 & 0 \\ 1 & -1 & 0 & -1 \end{pmatrix} = 2$$

$$b_4 = -\frac{1}{4}Sp\left(\begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}\right) = -\frac{1}{4}Sp\left(\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\right) = -1 = \det(A)$$

Wir sehen, daß wir nebenbei noch (fast) die Inverse von A berechnet haben.

Sei $f(z) = \sum a_i z^{n-i}$ ein normiertes Polynom (also $a_0 = 1$) und sei A eine Matrix, dann setzen wir $f(A) = \sum a_i A^{n-i}$, dies ist wieder eine Matrix.

Wenn $f(A) = 0$ die Nullmatrix ist, so heißt f ein die Matrix A annullierendes Polynom. Wie wir im Satz von Hamilton-Cayley gesehen haben, ist das charakteristische Polynom ein annullierendes Polynom.

Definition: Ein (normiertes) Polynom f mit $f(A) = 0$, das den kleinstmöglichen Grad hat, heißt Minimalpolynom von A .

Lemma 1.0.6 Sei $m(z)$ ein Minimalpolynom der Matrix A und $f(z)$ irgendein annullierendes Polynom. Dann ist m ein Teiler von f .

Beweis: Wir dividieren mit Rest:

$$f(z) = q(z)m(z) + r(z),$$

es ist $r = 0$ oder $\deg(r) < \deg(m)$. Wenn $r = 0$ ist, so folgt die Teilbarkeit. Sonst setzen wir A ein:

$$0 = f(A) = q(A)m(A) + r(A),$$

wegen $m(A) = 0$ folgt $r(A) = 0$, d.h. $r(z)$ wäre ein A annullierendes Polynom mit einem Grad, der kleiner als der von m ist, ein Widerspruch. \square

Folgerung 1.0.3 *Das Minimalpolynom von A ist eindeutig bestimmt.*

Beweis: Wir nehmen an, wir hätten zwei Minimalpolynome. Dann teilen sie sich gegenseitig, und da sie normiert sein sollten, sind sie gleich. \square

Folgerung 1.0.4 *Die Nullstellen des Minimalpolynoms der Matrix A sind Eigenwerte von A .*

Beweis: Das Minimalpolynom teilt das charakteristische Polynom. \square

Wir bemerken, daß auch die Umkehrung gilt: Jeder Eigenwert von A ist Nullstelle des Minimalpolynoms, dazu benötigen wir noch ein paar Hilfsmittel.

Satz 1.0.6 *Sei $f(x) = \sum a_i x^i \in R[x]$ ein Polynom, P eine invertierbare Matrix und A eine beliebige Matrix. Dann gilt $f(P^{-1}AP) = P^{-1}f(A)P$.*

Beweis: Die Idee ist $(P^{-1}AP)^2 = P^{-1}APP^{-1}AP = P^{-1}A^2P$, also $f(P^{-1}AP) = \sum a_i (P^{-1}AP)^i = \sum a_i P^{-1}A^i P = P^{-1}(\sum a_i A^i)P$. \square

Folgerung 1.0.5 *Wenn $f(A) = 0$ ist, so ist auch $f(P^{-1}AP) = 0$.* \square

Satz 1.0.7 *Wenn $f(A) = 0$ ist, so gilt $f(z_i) = 0$ für alle Eigenwerte z_i von A .*

Beweis: Wir wissen, daß es eine invertierbare Matrix P gibt, so daß $P^{-1}AP = D$ eine obere Dreiecksmatrix ist und daß auf der Diagonalen von D die Eigenwerte z_i von A stehen. Wie wir eben gesehen haben, gilt $f(D) = 0$, man rechnet nun leicht aus, daß auf der Diagonalen von $f(D)$ gerade die Ausdrücke $f(z_i)$ stehen. \square

Folgerung 1.0.6 *Jeder Eigenwert von A ist Nullstelle des Minimalpolynoms von A .*

Wir können dies anwenden:

Wenn A idempotent ist, so ist sein Minimalpolynom ein Teiler von $z^2 - z$, also hat A nur die Eigenwerte 1 und 0.

Wenn A nilpotent ist, so hat sein Minimalpolynom die Form z^k , also ist 0 der einzige Eigenwert.

Wenn A involutiv ist (d.h. $A^2 = E$), so ist sein Minimalpolynom von der Form $z^2 - 1$, also kommen nur die Eigenwerte 1 und -1 in Frage.

Wir haben bisher recht naiv mit Polynomen $f(x)$ gerechnet; was ist eigentlich das x ? Manchmal nennt man es eine Unbestimmte, aber kann man mit unbestimmten Objekten rechnen?

Wir machen folgende Konstruktion:

$$F = \{(a_0, a_1, \dots, a_n, 0, \dots) \mid a_i \in R\}$$

sei die Menge aller endlichen Folgen, d.h. nur endlich viele Glieder dürfen von Null verschieden sein. Wir führen eine Addition und eine Multiplikation in F ein:

$$(a_i) + (b_i) = (a_i + b_i),$$

$$(a_i) \cdot (b_i) = (c_k) = \left(\sum_{i+j=k} a_i b_j \right),$$

man sieht leicht, daß die rechten Seiten ebenfalls endliche Folgen sind.

Bezüglich der Addition bildet F eine kommutative Gruppe, auch die Multiplikation ist kommutativ. Wir zeigen die Gültigkeit des Assoziativgesetzes:

$$\begin{aligned} ((a_i)(b_j))(c_l) &= \left(\sum_{i+j=k} a_i b_j \right) (c_l) \\ &= \sum_{k+l=p} \sum_{i+j=k} a_i b_j c_l \\ &= \sum_{i+j+l=p} a_i b_j c_l \end{aligned}$$

und diesen symmetrischen Ausdruck können wir in die gewünschte Form überführen. Bei der Multiplikation ist die Folge $(1, 0, 0, \dots)$ das neutrale Element und die Elemente der Form $(a, 0, 0, \dots)$ verhalten sich bei Addition und Multiplikation wie Elemente von R , wir werden also $(a, 0, 0, \dots)$ mit $a \in R$ identifizieren.

Wir setzen nun

$$x = (0, 1, 0, \dots),$$

dann ist

$$x \cdot x = (0, 0, 1, 0, \dots)$$

und

$$x^i = (0, \dots, 0, 1, 0, \dots),$$

wo die 1 an der $(i+1)$ -ten Position steht. Dann hat jedes Element von F die Form

$$(a_0, a_1, \dots, a_n, 0, \dots) = \sum_{i=0}^n a_i x^i.$$

Wir können also sagen: Die Unbestimmte x ist die obengenannte Folge und ein Polynom ist ein Element von F .

Für die Festlegung der Rechenoperationen in F war die Forderung nach der Endlichkeit der Folgen eigentlich nicht wesentlich. Wir setzen

$$P = \{(a_0, a_1, \dots) \mid a_i \in R\}$$

und vereinbaren eine Addition und eine Multiplikation wie soeben. Dann hat jedes Element von P die Gestalt

$$f = (a_i) = \sum_{i=0}^{\infty} a_i x^i,$$

dies nennen wir eine formale Potenzreihe (wir kümmern uns nicht um Konvergenzfragen).

Aufgabe: Betrachten Sie die (formalen) Potenzreihen, die die Sinus- und die Cosinusfunktion darstellen, und zeigen Sie mithilfe der obigen Rechenregeln, daß $\sin^2(x) + \cos^2(x) = 1$ gilt.

Sei $f(x) = \sum a_i x^i$ eine Potenzreihe mit $a_0 \neq 0$; wir zeigen, daß eine Potenzreihe $g(x)$ mit $f(x) \cdot g(x) = 1$ existiert, d.h. $f(x)$ ist invertierbar.

Wir machen einen Ansatz $g(x) = \sum b_j x^j$, dann soll

$$\sum_{i+j=k} a_i b_j = \begin{cases} 0 & \text{für } k > 0 \\ 1 & \text{für } k = 0 \end{cases}$$

gelten. Also muß der Reihe nach $a_0 b_0 = 1$, $a_0 b_1 + a_1 b_0 = 0$, ... gelten, und diese Gleichungen sind offenbar lösbar.

Aufgabe: Berechnen Sie $\frac{1}{\cos(x)}$ und daraus $\tan(x) = \frac{\sin(x)}{\cos(x)}$.

Zum Schluß behandeln wir noch das Problem der Interpolation: Zu n gegebenen, verschiedenen Zahlen x_1, \dots, x_n und gegebenen y_1, \dots, y_n ist ein Polynom $f(x)$ vom Grad $n - 1$ gesucht, so daß $f(x_i) = y_i$ ist.

Dazu sind die folgenden, von Lagrange gefundenen Polynome hilfreich:

$$L_i = \frac{(x - x_1) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_n)}{(x_i - x_1) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n)}.$$

Offensichtlich gilt $L_i(x_j) = \delta_{ij}$ und damit ist $f(x) = \sum y_i L_i(x)$ das gesuchte Polynom.

Die Lagrange-Polynome für die Zahlen 1, 2, 3 lauten beispielsweise

$$L_1 = \frac{(z - 2)(z - 3)}{(1 - 2)(1 - 3)} = \frac{1}{2}(z^2 - 5z + 6)$$

$$L_2 = -z^2 + 4z - 3$$

$$L_3 = \frac{1}{2}(z^2 - 3z + 2)$$

Wir suchen eine Formel für $s_n = 1 + 2 + \dots + n$ und vermuten, daß dies ein Polynom in n vom Grad 2 ist. Mit $s_1 = 1$, $s_2 = 3$, $s_3 = 6$ erhalten wir $s_z = \frac{1}{2}z(z + 1)$.

Satz 1.0.8 $\sum L_i(x) = 1$.

Beweis: Wir wählen $y_i = 1$, $i = 1, \dots, n$. □

Wir werden dies anwenden, um die sogenannte Spektralzerlegung eines Endomorphismus zu berechnen.

Projektor auf Eigenraum zu 3

$$\begin{pmatrix} 2 & -2 & -2 & 2 \\ 0 & 0 & 0 & 0 \\ 2 & -5 & -3 & 4 \\ 1 & -4 & -2 & 3 \end{pmatrix}$$

$2 \cdot P_1 + 3 \cdot P_2$

$$\begin{pmatrix} 4 & -2 & -2 & 2 \\ 0 & 2 & 0 & 0 \\ 2 & -5 & -1 & 4 \\ 1 & -4 & -2 & 5 \end{pmatrix}$$

Lemma 1.0.7 Seien p_1, p_2 orthogonale Idempotente und $g(x) = \sum a_i x^i$ ein Polynom. Dann ist $g(ap_1 + bp_2) = g(a)p_1 + g(b)p_2$.

Beweis: $(ap_1 + bp_2)^n = \sum \binom{n}{k} a^k b^{n-k} p_1^k p_2^{n-k} = a^n p_1 + b^n p_2$. □

Der folgende Satz gibt eine Konstruktion für die Projektionen in der Spektralzerlegung an.

Satz 1.0.10 Seien $p_i : V \rightarrow V, i = 1, \dots, m$ Projektionen und $V = \text{Imp}_1 \oplus \dots \oplus \text{Imp}_m$. Weiter seien z_1, \dots, z_m paarweise verschiedene Zahlen und $f = \sum z_i p_i$. Dann sind die z_i Eigenwerte von f , die Eigenräume sind die Imp_i und f ist diagonalisierbar. Seien die L_i die Lagrangeschen Interpolationspolynome für z_1, \dots, z_m . Dann gilt $p_i = L_i(f)$. Da sich jeder diagonalisierbare Endomorphismus in der angegebenen Form darstellen läßt, erhält man so die Projektionen auf die Eigenräume.

Beweis: Sei $v_j \in \text{Imp}_j$, dann ist $f(v_j) = \sum z_i p_i(v_j) = z_j v_j$. Wir wählen in jedem Eigenraum eine Basis und erhalten so eine Basis von V aus Eigenvektoren von f , also ist f diagonalisierbar. Schließlich gilt

$$L_j(f) = L_j\left(\sum z_i p_i\right) = \sum L_j(z_i) p_i = p_j.$$

Als Beispiel betrachten wir die folgende diagonalisierbare Matrix

$$A = \begin{pmatrix} 4 & -2 & -2 & 2 \\ 0 & 2 & 0 & 0 \\ 2 & -5 & -1 & 4 \\ 1 & -4 & -2 & 5 \end{pmatrix}$$

ihre Eigenwerte sind 2 und 3. die entsprechenden Lagrangepolynome sind $-x + 3$ und $x - 2$. Der Projektor P_2 auf den Eigenraum zu 2 ist

$$\begin{pmatrix} -1 & 2 & 2 & -2 \\ 0 & 1 & 0 & 0 \\ -2 & 5 & 4 & -4 \\ -1 & 4 & 2 & -2 \end{pmatrix}$$

Der Projektor P_3 auf den Eigenraum zu 3 ist

$$\begin{pmatrix} 2 & -2 & -2 & 2 \\ 0 & 0 & 0 & 0 \\ 2 & -5 & -3 & 4 \\ 1 & -4 & -2 & 3 \end{pmatrix}$$

Es ist $2P_2 + 3P_3 = A$, die Eigenräume sind die Bilder der Projektoren; jeweils die erste und zweite Spalte bilden eine Basis des jeweiligen Eigenraums. \square

Nun wollen wir noch überlegen, wie sich die Eigenwerte bei „algebraischen“ Operationen verhalten.

Satz 1.0.11 *Seien $z_1, \dots, z_n \in \mathbb{C}$ die Eigenwerte von A und sei $p(z) \in \mathbb{C}[z]$ ein Polynom; dann sind die Zahlen $p(z_1), \dots, p(z_n)$ die Eigenwerte von $p(A)$.*

Beweis: Es sei

$$g(z) = \prod_{j=1}^m (y_j - z)$$

ein normiertes Polynom mit den Nullstellen y_1, \dots, y_m . Dann ist

$$g(A) = \prod_{j=1}^m (y_j E - A).$$

Wenn

$$c_A(z) = \prod_{i=1}^n (z - z_i) = \det(zE - A)$$

das charakteristische Polynom ist, so folgt aus dem Determinantenmultiplikationssatz

$$\det(g(A)) = \prod_j \det(y_j E - A) = \prod_j c_A(y_j) =$$

$$\prod_j \prod_i (y_j - z_i) = \prod_i \prod_j (y_j - z_i) = g(z_1) \cdots g(z_n).$$

Um die Eigenwerte von $p(A)$ zu bestimmen, setzen wir $g(z) = w - p(z)$, dann ist $g(A) = wE - p(A)$ und damit

$$\det(wE - p(A)) = (w - p(z_1)) \cdots (w - p(z_n)),$$

dies ist eine Zerlegung von $c_{p(A)}(w)$ in Linearfaktoren, also hat $p(A)$ die Eigenwerte $p(z_1), \dots, p(z_n)$. \square

Algebraische Körpererweiterungen

Eine (komplexe) Zahl α heißt algebraische Zahl, wenn es ein Polynom $f(x) \in \mathbb{Q}[x]$ gibt, so daß $f(\alpha) = 0$ ist. Das normierte Polynom $m(x)$ minimalen Grades mit $m(\alpha) = 0$ heißt das Minimalpolynom von α .

Satz 1.0.12 *Das Minimalpolynom $m(x)$ einer algebraischen Zahl α ist irreduzibel, d.h. es gibt keine Polynome $f(x), g(x) \in \mathbb{Q}[x]$ mit $m(x) = f(x) \cdot g(x)$.*

Beweis: Wenn $m(x) = f(x) \cdot g(x)$ gilt, so gilt auch $m(\alpha) = f(\alpha) \cdot g(\alpha)$, also muß wegen $m(\alpha) = 0$ auch $f(\alpha) = 0$ oder $g(\alpha) = 0$ gelten. Dies widerspricht der Minimalität des Grades von $m(x)$. \square

Von nun an wollen wir mit griechischen Buchstaben stets algebraische Zahlen bezeichnen.

Der kleinste Körper, die sowohl alle rationalen Zahlen als auch die Zahl α enthält, besteht aus allen Brüchen der Form $\frac{f(\alpha)}{g(\alpha)}$ mit $g(\alpha) \neq 0$, er wird mit $\mathbb{Q}(\alpha)$ bezeichnet.

Solch ein Körper wird als einfache Erweiterung des Körpers \mathbb{Q} bezeichnet. Wir werden sehen, daß wir seine Elemente etwas einfacher darstellen können, so daß sie auch ein Computer versteht. Denken Sie an $\alpha = \sqrt{2}$.

Zunächst überlegen wir uns, daß es ein Polynom $u(x)$ gibt, so das $u(\alpha) = \frac{1}{g(\alpha)}$ gilt,

d.h. jedes Element von $\mathbb{Q}(\alpha)$ ist ein Polynom in α , nicht ein Quotient zweier Polynome. In der Tat: Da das Minimalpolynom $m(x)$ von α irreduzibel ist, gibt es nur zwei Möglichkeiten: Entweder ist $m(x)$ ein Teiler von $g(x)$, dann wäre aber $g(\alpha) = 0$, oder der größte gemeinsame Teiler von $g(x)$ und $m(x)$ ist gleich 1. Dann gibt es aber Polynome $u(x), v(x)$ mit $g(x) \cdot u(x) + m(x) \cdot v(x) = 1$, nun setzen wir $x = \alpha$ und sehen $g(\alpha) \cdot u(\alpha) = 1$, wie gewünscht.

Weiter: Der Grad von $m(x)$ sei gleich n . Dann gibt es für jedes Polynom $f(x)$ ein Polynom $r(x)$ von einem Grade, der kleiner als n ist, mit $f(\alpha) = r(\alpha)$. Wenn der Grad von $f(x)$ größer oder gleich n ist, so dividieren wir $f(x)$ durch $m(x)$:

$$f(x) = q(x) \cdot m(x) + r(x), \deg(r) < n.$$

Wir setzen wieder $x = \alpha$ und erhalten $f(\alpha) = r(\alpha)$.

Folgerung 1.0.7 $\mathbb{Q}(\alpha)$ ist ein \mathbb{Q} -Vektorraum der Dimension n .

Beweis: Die Menge $\{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\}$ bildet eine \mathbb{Q} -Basis von $\mathbb{Q}(\alpha)$. \square

Da das Minimalpolynom $m(x)$ irreduzibel ist, ist das von $m(x)$ erzeugte Ideal $(m(x))$ des Polynomrings $\mathbb{Q}[x]$ ein maximales Ideal, also ist der Faktorring $\mathbb{Q}[x]/(m(x))$ ein Körper. Durch die Zuordnung

$$x^i \bmod m(x) \rightarrow \alpha^i$$

wird ein Isomorphismus $\mathbb{Q}[x]/(m(x)) \rightarrow \mathbb{Q}(\alpha)$ gegeben. Also können wir die Elemente $\beta \in \mathbb{Q}(\alpha)$ folgendermaßen im Rechner darstellen: $\beta = \sum b_i x^i$ ist ein Polynom vom Grade $\leq n - 1$, wir rechnen mit diesen Elementen modulo $m(x)$. Wenn der Grad eines Produkts zweier Polynome die Zahl $n - 1$ übersteigt, ersetzen wir es durch seine Restklasse mod $m(x)$. Wie wir oben gesehen haben, ist das Inverse eines Polynoms mit Hilfe des Euklidischen Algorithmus berechenbar.

Es wäre aber auch zu überlegen, ob man die zugegebenermaßen zeitaufwendige Inversenberechnung nicht so lange wie möglich vor sich herschieben sollte. Wir fügen einfach

jedem Element von $\mathbb{Q}(\alpha)$ einen Nenner hinzu, der anfangs gleich 1 gesetzt wird. Nun wird wie mit gewöhnlichen Brüchen gerechnet, bei einer Division wird einfach mit dem reziproken Bruch multipliziert, und Zähler und Nenner werden stets modulo $m(x)$ reduziert. Erst, wenn eine Ausgabe notwendig ist, werden die Nenner bereinigt: der Zähler wird mit dem Inversen des Nenners multipliziert.

Damit haben wir die Arithmetik in endlichen algebraischen Erweiterungskörpern von \mathbb{Q} im Griff.

Schwieriger wird es, wenn wir zwei beliebige algebraische Zahlen α, β addieren oder multiplizieren wollen. Beide Zahlen sind als Nullstellen ihrer jeweiligen Minimalpolynome zu verstehen, wir müßten also das Minimalpolynom von $\alpha + \beta$ bestimmen. Alle drei Zahlen liegen im Körper $\mathbb{Q}(\alpha, \beta)$, der folgende Satz sagt aus, daß dies ebenfalls eine einfache Erweiterung von \mathbb{Q} ist.

Satz 1.0.13 (Satz vom primitiven Element) *Seien α, β algebraische Zahlen, dann gibt es eine algebraische Zahl δ mit $\mathbb{Q}(\alpha, \beta) = \mathbb{Q}(\delta)$, d.h. es gibt Polynome $p(x, y)$, $q(x)$, $r(x)$ mit $\delta = p(\alpha, \beta)$, $\alpha = q(\delta)$, $\beta = r(\delta)$.*

Beweis: Sei $f(x)$ das Minimalpolynom von α , seine Nullstellen seien $\alpha = \alpha_1, \alpha_2, \dots, \alpha_n$. Sei $g(x)$ das Minimalpolynom von β , seine Nullstellen seien $\beta = \beta_1, \beta_2, \dots, \beta_m$. Dann gibt es eine rationale Zahl c mit

$$\alpha_i + c\beta_k \neq \alpha_1 + c\beta_1 \text{ für alle } i, k,$$

denn jede der Gleichungen

$$\alpha_i + x\beta_k = \alpha_1 + x\beta_1$$

hat höchstens eine Lösung $x \in \mathbb{Q}$. Wir setzen $\delta = \alpha + c\beta \in \mathbb{Q}(\alpha, \beta)$, also ist $\mathbb{Q}(\delta) \subseteq \mathbb{Q}(\alpha, \beta)$.

Wir betrachten nun $g(x)$ und $f(\delta - cx)$ als Polynome in $\mathbb{Q}(\delta)[x]$. Es gilt $g(\beta) = 0$ und $f(\delta - c\beta) = f(\alpha) = 0$ sowie

$$\delta - c\beta_k = \alpha_1 + c\beta - c\beta_k \neq \alpha_i \text{ für alle } k,$$

also $f(\delta - c\beta_k) \neq 0$ für $k > 1$. Also ist β die einzige gemeinsame Nullstelle der Polynome $g(x)$ und $f(\delta - cx)$, demnach ist ihr größter gemeinsamer Teiler, der im Polynomring $\mathbb{Q}(\delta)[x]$ zu berechnen ist, gleich $x - \beta$. Das heißt aber, daß β im Koeffizientenkörper $\mathbb{Q}(\delta)$ liegt, damit ist auch $\alpha = \delta - c\beta \in \mathbb{Q}(\delta)$, also $\mathbb{Q}(\delta) = \mathbb{Q}(\alpha, \beta)$.

Beispiel: $\alpha = \sqrt{2}$, $\beta = \sqrt{3}$, $\delta = \sqrt{2} + \sqrt{3}$, das Minimalpolynom von δ ist $x^4 - 10x^2 + 1$, es gilt $\sqrt{2} = (\delta^3 - 9\delta)/2$, $\sqrt{3} = (11\delta - \delta^3)/2$.

Resultante und Diskriminante

Seien zwei Polynome

$$f(x) = a_0x^m + a_1x^{m-1} + \dots + a_m$$

$$g(x) = b_0x^n + b_1x^{n-1} + \dots + b_n$$

gegeben.

Beispiel: $f(x) = x^2 - 1$, $g(x) = x^2 + 2x + 1$

$$\begin{vmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{vmatrix} = 0$$

Sei $f(x) = \prod(x - x_i)$, $g(x) = \prod(x - y_j)$ normierte Polynome, dann sind die Koeffizienten von f bzw. von g die elementarsymmetrischen Funktionen der Nullstellen x_i bzw. y_j , also ist auch die Resultante ein Polynom in den x_i und y_j . Da die Resultante stets verschwindet, wenn ein x_i gleich einem y_j ist, ist sie durch alle Differenzen $(x_i - y_j)$ teilbar, und wenn wir uns Grad und höchsten Koeffizienten genauer anschauen, finden wir

$$\text{Res}(f, g) = \prod (x_i - y_j).$$

Die Resultante von f und f' wird als Diskriminante von f bezeichnet, sie verschwindet genau dann, wenn f mehrfache Nullstellen besitzt.

Man kann nun direkt ein Polynom angeben, daß als Nullstelle die Zahl $\alpha + \beta$ besitzt: Wir wählen eine neue Unbestimmte y und betrachten das Polynom $f(y - x)$ als Polynom in x , die Resultante

$$r(y) = \text{Res}(f(y - x), g(x))$$

verschwindet genau dann, wenn $f(y - x)$ und $g(x)$ eine gemeinsame Nullstelle besitzen, dies ist für $y = \alpha + \beta$ der Fall, die Nullstelle ist gleich $x = \beta$.

In unserem Beispiel sieht das folgendermaßen aus:

$$f(x) = x^2 - 2, g(x) = x^2 - 3, f(y - x) = x^2 - 2xy + y^2 - 2,$$

$$r(y) = \begin{vmatrix} 1 & -2y & y^2 - 2 & 0 \\ 0 & 1 & -2y & y^2 - 2 \\ 1 & 0 & -3 & 0 \\ 0 & 1 & 0 & -3 \end{vmatrix} = y^4 - 10y + 1.$$

Folgerung 1.0.9 Die Menge der algebraischen Zahlen ist ein Körper.

Kapitel 2

Normalformen von Matrizen

2.1 Invariante Unterräume

Während dieses gesamten Kapitels sei V ein fixierter Vektorraum und $f : V \rightarrow V$ ein Endomorphismus. Alle einzuführenden Begriffe beziehen sich auf diese Situation, auch wenn es später nicht ausdrücklich erwähnt wird.

Definition: Ein Unterraum $U \subseteq V$ heißt invariant (bezüglich f), wenn $f(U) \subseteq U$ gilt.

Sei U_1 ein invarianter Unterraum von V . Wir wählen eine Basis B_1 von U_1 und ergänzen sie durch eine Menge B_2 zu einer Basis B von V . Wie sieht die Darstellungsmatrix $A_{BB}(f)$ aus?

Nun, das Einzige, was wir wissen, ist, das für $b_i \in B_1$ das Bild $f(b_i)$ in $L(B_1)$ liegt, also

$$A_{BB}(f) = \begin{pmatrix} \star & \dots & \star & ? & \dots & ? \\ & \dots & & & & \\ \star & \dots & \star & ? & \dots & ? \\ & \dots & & & & \\ 0 & \dots & 0 & ? & \dots & ? \\ & \dots & & & & \\ 0 & \dots & 0 & ? & \dots & ? \end{pmatrix}$$

Die Darstellungsmatrix besitzt links unten einen Null-Block.

In besonderen Fällen kann es vorkommen, daß die Darstellungsmatrix auch noch rechts oben einen Null-Block besitzt, etwa so:

$$\begin{pmatrix} \star & \dots & \star & 0 & \dots & 0 \\ & \dots & & & & \\ \star & \dots & \star & 0 & \dots & 0 \\ & \dots & & & & \\ 0 & \dots & 0 & ? & \dots & ? \\ & \dots & & & & \\ 0 & \dots & 0 & ? & \dots & ? \end{pmatrix}$$

Das ist dann der Fall, wenn $f(B_2) \subseteq L(B_2)$ ist, d.h. der Unterraum $U_2 = L(B_2)$ ist ebenfalls invariant. Wir bemerken nebenbei, daß V die direkte Summe von U_1 und U_2 ist: $V = U_1 \oplus U_2$.

Definition: Der Vektorraum V heißt (bezüglich f) zerlegbar, wenn invariante Unterräume U_1, U_2 von V existieren, so daß $V = U_1 \oplus U_2$ ist, anderenfalls heißt V unzerlegbar.

Lemma 2.1.1 *Es gibt unzerlegbare invariante Unterräume U_1, \dots, U_r mit*

$$V = U_1 \oplus \dots \oplus U_r.$$

Beweis: Wenn V unzerlegbar ist, so setzen wir $r = 1$ und $U_1 = V$. Wenn V zerlegbar ist, so zerlegen wir V in invariante Unterräume: $V = U_1 \oplus U_2$. Wenn U_1 und U_2 unzerlegbar sind, sind wir fertig, wenn nicht, so zerlegen wir diese Unterräume weiter, bis wir lauter unzerlegbare invariante Summanden erhalten. \square

2.2 Nilpotente Endomorphismen

Definition: Ein Endomorphismus $f : V \rightarrow V$ heißt nilpotent vom Grade n , wenn $f^{n-1} \neq o$ und $f^n = o$ ist.

Das folgende Lemma ist trivial.

Lemma 2.2.1 *Sei $f : V \rightarrow V$ nilpotent vom Grade n und $V = V_1 \oplus \dots \oplus V_r$ eine direkte Summe invarianter Unterräume. Es sei $f_i = f | V_i$ die Einschränkung von f auf den Unterraum V_i . Dann ist f_i nilpotent vom Grade $\leq n$. \square*

Satz 2.2.1 *Sei $f : V \rightarrow V$ nilpotent vom Grade n , dann ist*

$$\{o\} \subset \text{Ker}(f) \subset \text{Ker}(f^2) \subset \dots \subset \text{Ker}(f^{n-1}) \subset V$$

und alle Inklusionen sind echt.

Beweis: Wenn $f^i(v) = o$ ist, so ist auch $f^{i+1}(v) = o$, also ist $\text{Ker}(f^i)$ in $\text{Ker}(f^{i+1})$ enthalten.

Wir nehmen an, daß $\text{Ker}(f^i) = \text{Ker}(f^{i+1})$ für ein i gelte. Das heißt, wenn ein Vektor in $\text{Ker}(f^{i+1})$ liegt, so liegt er auch in $\text{Ker}(f^i)$. Nun existiert ein v mit $f^{n-1}(v) \neq o$, dann ist

$$o = f^n(v) = f^{i+1}(f^{n-i-1}(v))$$

und damit

$$f^i(f^{n-i-1}(v)) = f^{n-1}(v) = o,$$

ein Widerspruch. \square

Satz 2.2.2 *Sei f nilpotent vom Grade n und $f^{n-1}(v) \neq o$, dann sind $v, f(v), \dots, f^{n-1}(v)$ linear unabhängig.*

Beweis: Es ist $f^i(v) \in \text{Ker}(f^{n-i})$, denn $f^{n-i}(f^i(v)) = f^n(v) = o$, aber $f^i(v)$ liegt nicht in $\text{Ker}(f^{n-i-1})$, wie man schnell nachrechnet. Sei nun

$$r_0v + r_1f(v) + \dots + r_{n-1}f^{n-1}(v) = o$$

und es sei i die kleinste Zahl, so daß $r_i \neq 0$ ist (also $r_0 = \dots = r_{i-1} = 0$). Dann ist

$$-r_0f^i(v) = r_{i-1}f^{i+1}(v) + \dots + r_{n-1}f^{n-1}(v),$$

die Summanden auf der rechten Seite liegen aber alle in $\text{Ker}(f^{n-i-1})$, ein Widerspruch. \square

Wir betrachten nochmals die Inklusion der Kerne, um eine besonders schöne Basis von V zu konstruieren.

Sei zunächst $v \in \text{Ker}(f^i)$, also $f^i(v) = o = f^{i-1}(f(v))$, d.h. $f(v) \in \text{Ker}(f^{i-1})$.

Wir wollen schrittweise Unterräume $T_{i-1} \subset \text{Ker}(f^i)$ konstruieren, so daß

$$\text{Ker}(f^{i-1}) \oplus T_{i-1} = \text{Ker}(f^i)$$

gilt. Als T_{n-1} wählen wir einen beliebigen zu $\text{Ker}(f^{n-1})$ komplementären Unterraum von $\text{Ker}(f^n)$.

Sei $\{v_1, \dots, v_k\}$ eine Basis von T_{i-1} . Dann liegen $f(v_1), \dots, f(v_k)$ in $\text{Ker}(f^{i-1})$. Wir nehmen an, eine Linearkombination dieser Vektoren läge in $\text{Ker}(f^{i-2})$:

$$f^{i-2}\left(\sum r_i f(v_i)\right) = o = f^{i-1}\left(\sum r_i v_i\right),$$

also

$$\sum r_i v_i \in \text{Ker}(f^{i-1}) \cap T_{i-1} = \{o\},$$

d.h. $\sum r_i v_i = o$, also $r_1 = \dots = r_k = 0$.

Unsere Vektoren sind sogar linear unabhängig, wir ergänzen sie zu einer Basis eines zu $\text{Ker}(f^{i-2})$ komplementären Teilraums T_{i-2} von $\text{Ker}(f^{i-1})$.

So fahren wir fort und erhalten den

Satz 2.2.3 Sei $f : V \longrightarrow V$ nilpotent, dann gibt es Vektoren $v_1, \dots, v_k \in V$, so daß

$$\{v_1, f(v_1), \dots, f^{n_1}(v_1), \dots, v_k, f(v_k), \dots, f^{n_k}(v_k)\}$$

eine Basis von V ist.

Man kann auch wie folgt vorgehen: Wir wählen eine Basis von $\text{Ker}(f)$ und bestimmen Urbilder der Basisvektoren (in $\text{Ker}(f^2)$), diese Urbilder sind linear unabhängig, wir ergänzen sie zu einer Basis von $\text{Ker}(f^2)$ und verfahren mit dieser Basis analog.

Nun übertragen wir dies auf Matrizen, indem wir sie als Darstellungsmatrizen von Endomorphismen auffassen:

Folgerung 2.2.1 Sei A eine nilpotente Matrix, dann gibt es eine reguläre Matrix X , so daß $X^{-1}AX$ eine Block-Diagonalmatrix ist:

$$X^{-1}AX = \begin{pmatrix} A_1 & & 0 \\ & \dots & \\ 0 & & A_k \end{pmatrix},$$

und die A_i sind n_i -reihige Matrizen der Form

$$\begin{pmatrix} 0 & & \dots & 0 \\ 1 & 0 & & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ & & & \dots & \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}. \square$$

Beispiel: $A = \begin{pmatrix} -4 & 2 & 3 \\ -6 & 3 & 5 \\ -2 & 1 & 1 \end{pmatrix}$, $A^2 = \begin{pmatrix} -2 & 1 & 1 \\ -4 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, $A^3 = 0$. Der Vektor $v = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ liegt nicht in $\text{Ker}(A^2)$, es ist $Av = \begin{pmatrix} -4 \\ -6 \\ -2 \end{pmatrix}$, $A^2v = \begin{pmatrix} -2 \\ -4 \\ 0 \end{pmatrix}$, und mit $X = \begin{pmatrix} 1 & -4 & -2 \\ 0 & -6 & -4 \\ 0 & -2 & 0 \end{pmatrix}$ haben wir $X^{-1}AX = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

Noch ein Beispiel:

$$A = \begin{pmatrix} 7 & -7 & 1 & 2 & 0 \\ 15 & -19 & 6 & 10 & -5 \\ 17 & -19 & 5 & 9 & -4 \\ 20 & -32 & 14 & 22 & -14 \\ 24 & -36 & 15 & 24 & -15 \end{pmatrix} \quad A^2 = \begin{pmatrix} 1 & 1 & -2 & -3 & 3 \\ 2 & 2 & -4 & -6 & 6 \\ 3 & 3 & -6 & -9 & 9 \\ 2 & 2 & -4 & -6 & 6 \\ 3 & 3 & -6 & -9 & 9 \end{pmatrix} \quad A^3 = 0$$

$$\text{Ker}A^2 = \text{Spaltenraum von } \begin{pmatrix} 1 & -2 & -3 & 3 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \notin \text{Ker}A^2 \quad v_2 = Av_1 = \begin{pmatrix} 7 \\ 15 \\ 17 \\ 20 \\ 24 \end{pmatrix} \quad v_3 = Av_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 2 \\ 3 \end{pmatrix}$$

$$v_4 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \in \text{Ker}A^2, \quad v_4 \notin \text{Ker}A, \quad v_5 = Av_4 = \begin{pmatrix} 14 \\ 34 \\ 36 \\ 52 \\ 60 \end{pmatrix}$$

und mit $X = (v_1, \dots, v_5)$ erhalten wir zwei Blöcke:

$$X^{-1}AX = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

2.3 Jordansche Normalform

In diesem Abschnitt setzen wir voraus, daß das charakteristische Polynom zerfällt, wir wählen am Besten den Grundkörper \mathbb{C} .

Definition: Eine Matrix der Form

$$J(z) = \begin{pmatrix} z & & \dots & 0 \\ 1 & z & & 0 \\ 0 & 1 & z & \dots & 0 \\ & & & \dots & \\ 0 & \dots & 0 & 1 & z \end{pmatrix}$$

heißt Jordankästchen. Wir sagen, eine Matrix liegt in Jordanscher Normalform vor, wenn sie eine Blockdiagonalmatrix ist, deren Diagonalblöcke Jordankästchen sind:

$$\begin{pmatrix} J(z_1) & \dots & 0 \\ & \dots & \\ 0 & \dots & J(z_k) \end{pmatrix}$$

Die Eigenwerte einer Matrix in Jordanscher Normalform sind offenbar die z_1, \dots, z_k , die Eigenwerte in verschiedenen Jordankästchen müssen nicht voneinander verschieden sein, z.B. sind für dreireihige Matrizen folgende Jordanschen Normalformen möglich:

$$\begin{pmatrix} x & & \\ & y & \\ & & z \end{pmatrix}, \begin{pmatrix} x & & \\ & x & \\ & & y \end{pmatrix}, \begin{pmatrix} x & & \\ 1 & x & \\ & & y \end{pmatrix}, \begin{pmatrix} x & & \\ & x & \\ & & x \end{pmatrix}, \begin{pmatrix} x & & \\ 1 & x & \\ & & x \end{pmatrix}, \begin{pmatrix} x & & \\ 1 & x & \\ & 1 & x \end{pmatrix}.$$

Wir werden sehen, daß zu jeder Matrix A eine reguläre Matrix X existiert, so daß $X^{-1}AX$ Jordansche Normalform besitzt. Dies folgt aus dem

Satz 2.3.1 Sei $f : V \rightarrow V$ ein Endomorphismus, dann gibt es eine Basis B von V , so daß $A_{BB}(f)$ Jordansche Normalform besitzt.

Beweis: Sei z ein Eigenwert von f , dann ist $(f - z \text{id})$ nicht injektiv, wir setzen $g = f - z \text{id}$. Es gilt $\text{Ker}(g) \neq \{o\}$ und es sei

$$\{o\} \subset \text{Ker}(g) \subset \text{Ker}(g^2) \subset \text{Ker}(g^3) \subset \dots \subset \text{Ker}(g^m) = \text{Ker}(g^{m+1}),$$

die ersten Inklusionen seien alle echt, wir überlegen uns, daß die Kerne der noch höheren Potenzen von g alle übereinstimmen: Sei $g^{m+2}(v) = o$, dann ist $g^{m+1}(g(v)) = o$, also auch $g^m(g(v)) = g^{m+1}(v) = o$, usw.

Wir setzen nun $U_1 = \text{Ker}(g^m)$ und $U_2 = \text{Im}(g^m)$.

Behauptung: $V = U_1 \oplus U_2$. In der Tat: Sei $v \in U_1 \cap U_2$, also $v = g^m(w)$, dann ist $o = g^m(v) = g^{2m}(w)$, also liegt w in $\text{Ker}(g^{2m}) = \text{Ker}(g^m)$, also ist $v = g^m(w) = o$.

Andererseits gilt $\dim V = \dim \text{Im}(g^m) + \dim \text{Ker}(g^m)$, also ist V die direkte Summe von U_1 und U_2 .

Man sieht leicht, daß U_1 und U_2 g -invariante Unterräume sind: Sei $v \in \text{Ker}(g^m)$, dann ist $g^m(g(v)) = g(g^m(v)) = g(o) = o$, also $g(v) \in \text{Ker}(g^m)$. Sei $v \in \text{Im}(g^m)$, $v = g^m(w)$, dann ist $g(v) = g(g^m(w)) = g^m(g(w)) \in \text{Im } g^m$. Weiter ist die Einschränkung von g auf U_1 nilpotent vom Grade m .

Wir wenden nun unsere Kenntnisse über nilpotente Abbildung an: Wir zerlegen U_1 in eine direkte Summe unzerlegbarer invarianter Unterräume, oBdA können wir annehmen, daß U_1 selbst schon unzerlegbar ist. Also gibt es eine Basis B von U_1 , die folgende Gestalt hat

$$B = \{v, g(v), \dots, g^{m-1}(v)\}.$$

Wie wirkt nun f auf diese Basis? Es ist $f = g + z \text{id}$.

$$\begin{aligned} f(v) &= g(v) + zv \\ f(g(v)) &= g^2(v) + zg(v) \\ &\dots \\ f(g^{m-2}(v)) &= g^{m-1}(v) + zg^{m-2}(v) \\ f(g^{m-1}(v)) &= o + zg^{m-1}(v). \end{aligned}$$

Die Darstellungsmatrix der Einschränkung von f auf U_1 ist also ein Jordankästchen.

Nun schränken wir f auf U_2 ein und beginnen von vorn. Wir suchen einen Eigenwert, bilden ein neues g , stellen fest, wo sich die aufsteigende Folge der Kerne der Potenzen von g stabilisiert usw. Damit ist der Satz bewiesen. \square

Wir betrachten das folgende Beispiel aus der Webseite www.prenhall.com/Leon von Steven Leon:

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

Der einzige Eigenwert ist gleich 1 und der zugehörige Eigenraum ist 2-dimensional, es gibt also zwei Jordankästchen, deren Größen 2 und 3 oder 1 und 4 sein könnten. Um dies zu entscheiden, berechnen wir den Nilpotenzgrad der Matrix $A - E_5$: Es ist $(A - E_5)^2 \neq 0$, $(A - E_5)^3 \neq 0$, aber $(A - E_5)^4 = 0$. Also tritt der zweite Fall ein.

Noch ein Beispiel:

$$A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}, \quad c_A(z) = -(z - 2)^3,$$

$$B = A - 2E = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -2 \\ 1 & 2 & 3 \end{pmatrix} \quad B^2 = \begin{pmatrix} 0 & 2 & 2 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \end{pmatrix}$$

$$\text{Ker}B = L\left(\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\right), \text{Ker}B^2 = L\left(\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}\right)$$

Wir wählen $v_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \notin \text{Ker}B^2$, dann ist $v_2 = Bv_1 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$, $v_3 = B^2v_1 = \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix}$,
dann ist

$$X = \begin{pmatrix} 0 & 3 & 2 \\ 0 & -1 & 2 \\ 1 & 1 & 2 \end{pmatrix}, X^{-1} = \frac{1}{4} \begin{pmatrix} 0 & -3 & -1 \\ 0 & 2 & 2 \\ 4 & 4 & 0 \end{pmatrix}, X^{-1}AX = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}$$

Sei nun A eine Matrix und $J = X^{-1}AX$ ihre Jordansche Normalform. Sei $m(z)$ das Minimalpolynom von A , dann ist $m(z)$ auch das Minimalpolynom von J , wie man sich schnell überlegt.

Wir betrachten ein Beispiel:

$$J = \begin{pmatrix} J(z_1) & \\ & J(z_2) \end{pmatrix}$$

Das Kästchen zum Eigenwert z_1 habe p Reihen, das zum Eigenwert z_2 habe q Reihen. Das Minimalpolynom $m(z)$ kann nur die Nullstellen z_1 und z_2 haben, wie wir gesehen haben, wir müssen noch ihre Vielfachheit erraten. Wir wollen es nicht zu spannend machen, das Minimalpolynom ist in unserem Beispiel

$$m(z) = (z - z_1)^p(z - z_2)^q,$$

rechnen Sie es nach! Dann haben Sie die Beweisidee der

Folgerung 2.3.1 Die Matrix A habe die verschiedenen Eigenwerte z_1, \dots, z_l , das größte Jordankästchen zum Eigenwert z_i habe p_i Reihen. Dann ist $\prod (z - z_i)^{p_i}$ das Minimalpolynom von A . \square

Folgerung 2.3.2 Jeder Eigenwert von A ist Nullstelle des Minimalpolynoms von A .

Beweis: Die Eigenwerte von A und $X^{-1}AX$ stimmen überein, also muß in der Jordanschen Normalform von A zu jedem Eigenwert mindestens ein Kästchen vorhanden sein. \square

Wir können nun die Frage beantworten, wann es zu einer Matrix A eine reguläre Matrix X gibt, so daß $X^{-1}AX$ eine Diagonalmatrix ist, oder, was dasselbe heißt, ob der Vektorraum R^n eine Basis aus Eigenvektoren von A besitzt.

Satz 2.3.2 Die Matrix A ist genau dann diagonalisierbar, wenn ihr Minimalpolynom nur einfache Nullstellen besitzt, d.h.

$$m(z) = \prod (z - z_i), \quad z_i \neq z_j \text{ für } i \neq j.$$

Beweis: In diesem Fall haben alle Jordankästchen der Normalform von A die Größe 1, d.h. die Jordansche Normalform ist eine Diagonalmatrix. \square

Beispiel:

Sei A eine idempotente Matrix, also $A^2 = A$, das Minimalpolynom von A ist $m(z) = (z - 1)z$, hat also einfache Nullstellen, also ist A diagonalisierbar.

Folgerung 2.3.3 Sei A eine idempotente Matrix, dann ist $\text{rg}(A) = \text{Sp}(A)$.

Beweis: Der Rang einer diagonalisierbaren Matrix ist gleich der Differenz der Reihenzahl und der Vielfachheit des Eigenwerts 0. Da alle Eigenwerte von A gleich 0 oder 1 sind, ist der Rang gleich der Spur. \square

2.4 Rekursive Folgen

Wir betrachten eine Folge $(x_n)_{n \geq 0}$, deren Glieder einer Rekursionsformel genügen:

$$x_n = a_1 x_{n-1} + \dots + a_k x_{n-k}.$$

Alle Glieder sind eindeutig bestimmt, wenn die Anfangsglieder x_0, x_1, \dots, x_{k-1} gegeben sind. Das Problem besteht darin, eine explizite Formel für x_n zu finden, so daß man etwa x_{1000} sofort ausrechnen kann und nicht vorher x_{999}, x_{998} usw. kennen muß. Wir schreiben die Formel als Matrixprodukt:

$$\begin{pmatrix} x_n \\ x_{n-1} \\ \dots \\ \dots \\ x_{n-k+1} \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & \dots & a_k \\ 1 & 0 & & 0 \\ & 1 & 0 & \\ & & \dots & \\ & & & 1 & 0 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_{n-2} \\ \dots \\ \dots \\ x_{n-k} \end{pmatrix}$$

Den Vektor auf der linken Seite nennen wir X_n , auf der rechten Seite steht dann X_{n-1} , multipliziert mit der Matrix A . Der Vektor X_{k-1} enthält die Anfangswerte x_0, \dots, x_{k-1} . Dann gilt also

$$X_n = AX_{n-1} = A^i X_{n-i} = A^{n-k+1} X_{k-1}.$$

Die erste Zeile dieser Formel berechnet unser x_n , jedoch ist dies eigentlich keine explizite Formel, da für A^i keine explizite Formel bekannt ist, man muß eine Potenz nach der anderen ausrechnen (wenn man genauer hinsieht, stellt man fest, daß man zur Berechnung der i -ten Potenz nicht i Rechenoperationen, sondern nur $\log(i)$ durchführen muß). Nichtsdestoweniger ist es überhaupt nicht trivial, z.B. für $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n$ eine explizite Formel zu finden.

Bei Jordankästchen jedoch ist das Potenzieren ein Kinderspiel (Mit $\binom{n}{i}$ bezeichnen wir den Binomialkoeffizienten „ n über i “):

Lemma 2.4.1

$$\begin{pmatrix} z & & \dots & 0 \\ 1 & z & & 0 \\ 0 & 1 & z & \dots & 0 \\ & & & \dots & \\ 0 & \dots & 0 & 1 & z \end{pmatrix}^n = \begin{pmatrix} z^n & & & & \dots & 0 \\ \binom{n}{1} z^{n-1} & z^n & & & \dots & 0 \\ \binom{n}{2} z^{n-2} & \binom{n}{1} z^{n-1} & z^n & & \dots & 0 \\ & & & \dots & & \\ & & & & & z^n \end{pmatrix}$$

Den Induktionsbeweis überlassen wir dem Leser. \square

Wir kehren zu unserer Folge zurück. Sei J die Jordansche Normalform von A und $A = YJY^{-1}$, dann ist

$$X_n = A^{n-k+1} X_{k-1} = (YJY^{-1})^{n-k+1} X_{k-1} = YJ^{n-k+1} Y^{-1} X_{k-1},$$

also ist X_n ein Vektor, in dem Linearkombinationen der Potenzen der Eigenwerte von A stehen, damit ist eine explizite Formel für x_n gegeben.

Wir müssen uns aber nicht die Mühe machen, die Jordansche Normalform von A zu bestimmen, sondern wir können für x_n einen Ansatz $x_n = \sum \binom{n}{j} b_{ij} z_i^j$ machen, dabei sind die z_i die Eigenwerte von A und die b_{ij} bestimmt man aus den Anfangswerten der Folge. Wenn alle Eigenwerte von A paarweise verschieden sind, so ist J eine Diagonalmatrix und es reicht der Ansatz $x^n = \sum b_i z_i^n$.

Nun ist aber A nicht irgendeine Matrix, es ist nicht schwierig, ihr charakteristisches Polynom zu bestimmen:

Lemma 2.4.2

$$c_A(z) = (-1)^k \sum (-a_i) z^{k-i} \quad (a_0 = 1).$$

Also einfacher gehts wirklich nicht. Den Beweis überlassen wir allerdings wieder dem Leser. \square

Beispiele:

$$x_n = 2x_{n-1} - x_{n-2}$$

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, \quad c_A(z) = z^2 - 2z + 1, \quad \text{Eigenwerte } 1, 1;$$

Transformation in Jordansche Normalform:

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

$$\begin{pmatrix} x_n \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_{n-2} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_0 \end{pmatrix}$$

$$= \begin{pmatrix} n(x_1 - x_0) + x_0 \\ (n-1)(x_1 - x_0) + x_0 \end{pmatrix}$$

also ist $x_n = n(x_1 - x_0) + x_0$.

Fibonacci-Zahlen

$x_n = x_{n-1} + x_{n-2}$, Anfangswerte $x_0 = 0, x_1 = 1$. Das charakteristische Polynom ist $z^2 - z - 1$ und hat die einfachen Nullstellen $z_i = \frac{1 \pm \sqrt{5}}{2}$.

Wir machen den Ansatz $x_n = az_1^n + bz_2^n$ und bestimmen a und b aus den Anfangswerten zu $a = \frac{1}{\sqrt{5}} = -b$.

2.5 Lineare Differentialgleichungssysteme

Wir setzen hier ein paar Vorkenntnisse aus der Analysis voraus. Diesen Abschnitt behandeln wir nicht in der Vorlesung, er ist für späteres Nachschlagen gedacht.

Sei $y(x)$ eine differenzierbare Funktion und $a(y)$ eine gegebene Funktion, dann heißt eine Gleichung der Form $y'(x) = a(y(x))$ eine Differentialgleichung.

Wenn n^2 Funktionen $a_{ij}(y)$ gegeben und n Funktionen $y_1(x), \dots, y_n(x)$ gesucht sind, so daß

$$y_1' = a_{11}(y_1) + \dots + a_{1n}(y_n)$$

...

$$y_n' = a_{n1}(y_1) + \dots + a_{nn}(y_n)$$

gilt, so nennen wir diese Bedingungen ein „lineares homogenes Differentialgleichungssystem 1. Ordnung“. Wir schreiben es auch kurz in der Form $y' = Ay$.

Lemma 2.5.1 *Die Menge aller Lösungen von $y' = Ay$ bildet einen Vektorraum.* \square

Ein „Anfangswertproblem“ besteht darin, eine Lösung zu finden, so daß $y_i(0) = c_i$, ($i = 1, \dots, n$) für ein gegebenes n -tupel $c \in \mathbf{R}^n$ gilt.

Im allereinfachsten Fall ist $n = 1$, $a(y) = y$, die Differentialgleichung $y' = y$ hat die Lösung $y(x) = ce^x$.

Der nächste einfachste Spezialfall ist $y' = ay$, $a \in \mathbf{R}$, hier haben wir die Lösung $y = e^{ax}$. Wir werden im folgenden versuchen, eine Exponentialfunktion für Matrizen einzuführen. Dazu überlegen wir zuerst, wie man die Matrix einer Differentialgleichung transformieren kann. Wir beschränken uns auf den Spezialfall, daß A eine konstante Matrix ist.

Sei $y = (y_1, \dots, y_n)$, dann setzen wir $y' = (y_1', \dots, y_n')$. Weiter sei M eine reguläre Matrix, dann besteht $w = My$ aus Linearkombinationen der y_i , und da die Ableitung eine lineare Abbildung ist, gilt $w_i' = (\sum m_{ij}y_j)' = \sum m_{ij}y_j'$, also $(My)' = My'$. Also gilt

$$y' = Ay \quad \text{gdw.} \quad M^{-1}w' = AM^{-1}w \quad \text{gdw.} \quad w' = MAM^{-1}w,$$

Wir können also ohne Beschränkung der Allgemeinheit annehmen, daß die Koeffizientenmatrix A in Jordanscher Normalform vorliegt.

Das Differentialgleichungssystem

$$y' = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_k \end{pmatrix} y$$

zerfällt nun in k voneinander unabhängigen Differentialgleichungen, so daß wir nur noch den Fall zu untersuchen haben, wo A ein Jordan-Block ist.

Die gewöhnliche Exponentialfunktion ist durch

$$e^x = \sum \frac{1}{i!} x^i$$

gegeben, diese Reihe konvergiert für alle $x \in \mathbf{R}$. Wir definieren: Eine Matrixfolge $C^{(k)} = (c_{ij}^{(k)})$ konvergiert, wenn alle Folgen $c_{ij}^{(k)}$ konvergieren, und der Grenzwert der Matrixfolge sei die Matrix der Grenzwerte.

Wir definieren nun

$$e^C = \sum \frac{1}{i!} C^i$$

und setzen zunächst voraus, daß diese Reihe konvergiert.

Dann gilt

$$e^{M^{-1}CM} = \sum \frac{1}{i!} (M^{-1}CM)^i = M^{-1} \left(\sum \frac{1}{i!} C^i \right) M = M^{-1} e^C M,$$

es genügt also, die Exponentialfunktion für Jordan-Blöcke zu berechnen. Sei also nun

$$C = \begin{pmatrix} z & & & \\ & 1 & z & \\ & & \ddots & \\ & & & 1 & z \end{pmatrix}$$

Lemma 2.5.2 Die Matrixreihe e^C konvergiert.

Beweis: In der Matrixreihe steht an der Stelle $(k+l+1, l)$ die Summe

$$\sum \frac{1}{i!} \binom{i}{k} z^{i-k} = \sum \frac{1}{i!} \frac{i!}{k!(i-k)!} z^{i-k} = \frac{1}{k!} \sum \frac{1}{(i-k)!} z^{i-k} = \frac{1}{k!} e^z,$$

und diese Summe existiert. □

In unserem Fall ist $C = Ax$ und C^i hat die Komponenten $\binom{i}{k} z^{i-k} x^i$, also hat e^{Ax} die Komponenten

$$\sum \frac{1}{i!} \binom{i}{k} z^{i-k} x^i = \frac{1}{k!} \sum \frac{1}{(i-k)!} z^{i-k} x^i = \frac{1}{k!} x^k \sum \frac{1}{(i-k)!} (zx)^{i-k} = \frac{1}{k!} x^k e^{zx}.$$

Satz 2.5.1 Für jedes n -tupel $c \in \mathbf{R}^n$ ist $y(x) = e^{Ax}c$ eine Lösung von $y' = Ay$.

Beweis: Wir berechnen $(e^{Ax}c)'$:

$$\left(\sum \frac{1}{k!} x^k e^{zx} c_l \right)' = \sum \frac{1}{k!} (kx^{k-1} e^{zx} c_l + zx^k e^{zx} c_l) = \sum \left(\frac{1}{(k-1)!} x^{k-1} e^{zx} + \frac{1}{k!} zx^k e^{zx} \right) c_l,$$

dies sind die Komponenten von

$$\begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & & \\ & & & 1 & 0 \end{pmatrix} e^{Ax}c + \begin{pmatrix} z & & & \\ & z & & \\ & & \ddots & \\ & & & z \end{pmatrix} e^{Ax}c = Ae^{Ax}c.$$

□

Kapitel 3

Euklidische Vektorräume

3.1 Skalarprodukt, Orthonormalbasis

Definition: Sei $b : V \times V \rightarrow \mathbb{R}$ eine Bilinearform, b heißt symmetrisch, wenn $b(v, w) = b(w, v)$ gilt, b heißt positiv definit, wenn $b(v, v) \geq 0$ für alle $v \in V$ und $b(v, v) = 0$ nur für $v = o$ gilt. Eine positiv definite symmetrische Bilinearform heißt Skalarprodukt. Zur Abkürzung schreiben wir bei Skalarprodukten $b(v, w) = \langle v, w \rangle$. Ein Vektorraum, in dem ein Skalarprodukt ausgezeichnet ist, heißt Euklidischer Vektorraum. Die Zahl $|v| = \sqrt{\langle v, v \rangle}$ heißt der Betrag des Vektors v .

Wenn wir zum Beispiel im Vektorraum \mathbb{R}^2 eine Basis $\{v_1, v_2\}$ wählen, und zwei Vektoren $v = r_1v_1 + r_2v_2$, $w = s_1v_1 + s_2v_2$ gegeben sind, so ist durch $\langle v, w \rangle = r_1s_1 + r_2s_2$ ein Skalarprodukt gegeben.

Eigenschaften des Betrags:

1. $|v| \geq 0$, wenn $|v| = 0$ ist, so ist $v = o$.
2. $|rv| = |r| |v|$ für $r \in \mathbb{R}$.
3. $|\langle v, w \rangle| \leq |v| |w|$ (Cauchy-Schwarzsche Ungleichung)

Beweis: Sei $r \in \mathbb{R}$, wir betrachten $u = v + rw$. Es gilt

$$0 \leq |u|^2 = \langle v + rw, v + rw \rangle = \langle v, v \rangle + 2r\langle v, w \rangle + r^2\langle w, w \rangle.$$

Wenn $w = o$ ist, so ist die Behauptung richtig. Nun sei $w \neq o$, wir setzen $r = -\frac{\langle v, w \rangle}{|w|^2}$ ein:

$$0 \leq \langle v, v \rangle - 2\frac{\langle v, w \rangle^2}{|w|^2} + \frac{\langle v, w \rangle^2}{|w|^2},$$

also

$$0 \leq |v|^2 |w|^2 - \langle v, w \rangle^2.$$

4. $|v + w| \leq |v| + |w|$ (Dreiecksungleichung)

Beweis: $|v + w|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle \leq |v|^2 + |w|^2 + 2|v| |w| = (|v| + |w|)^2$

5. Die Zahl $c = \frac{\langle v, w \rangle}{|v||w|}$ liegt zwischen -1 und 1 , wir setzen $\cos x = c$ und definieren x als den „Winkel“ zwischen den Vektoren v und w . Dann bedeutet $\langle v, w \rangle = 0$, daß v und w senkrecht aufeinander stehen.

Definition: Eine Menge $\{v_1, \dots, v_n\}$ heißt ein Orthonormalsystem, wenn $\langle v_i, v_j \rangle = \delta_{ij}$ gilt.

Lemma 3.1.1 *Orthonormalsysteme sind linear unabhängig.*

Beweis: Sei $\sum s_i v_i = o$, wir multiplizieren skalar mit v_j und erhalten $0 = \sum s_i \langle v_i, v_j \rangle = s_j$. \square

Definition: Eine Basis von V , die ein Orthonormalsystem ist, heißt Orthonormalbasis.

Satz 3.1.1 (Orthonormierungsverfahren von E. Schmidt) *Sei $\{v_1, \dots, v_n\}$ eine Basis von V , dann gibt es eine Orthonormalbasis $\{e_1, \dots, e_n\}$ von V , so daß*

$$\begin{aligned} \mathcal{L}(e_1) &= \mathcal{L}(v_1), \\ \mathcal{L}(e_1, e_2) &= \mathcal{L}(v_1, v_2), \\ &\dots \\ \mathcal{L}(e_1, \dots, e_i) &= \mathcal{L}(v_1, \dots, v_i), \quad (i = 1, \dots, n). \end{aligned}$$

Beweis: Wir setzen $e_1 = \frac{v_1}{|v_1|}$, dann ist $|e_1| = 1$ und $\mathcal{L}(e_1) = \mathcal{L}(v_1)$. Sei e_1, \dots, e_{i-1} schon konstruiert. Wir machen den Ansatz

$$e_i = r_1 e_1 + \dots + r_{i-1} e_{i-1} + v_i.$$

Die Bedingungen $\langle e_j, e_i \rangle = 0$ für $j = 1, \dots, i-1$ dienen zur Berechnung der r_j , indem wir e_i skalar mit e_j multiplizieren:

$$\begin{aligned} 0 &= \langle e_j, e_i \rangle \\ &= \langle e_j, r_1 e_1 \rangle + \dots + \langle e_j, r_{i-1} e_{i-1} \rangle + \langle e_j, v_i \rangle \\ &= r_j + \langle e_j, v_i \rangle, \end{aligned}$$

da e_j schon senkrecht auf e_1, \dots, e_{i-1} steht. Damit kann r_j berechnet werden. Falls nun $|e_i| \neq 1$ ist, so ersetzen wir e_i durch $\frac{e_i}{|e_i|}$ (e_i ist keinesfalls der Nullvektor, denn sonst

läge v_i in $\mathcal{L}(e_1, \dots, e_{i-1}) = \mathcal{L}(v_1, \dots, v_{i-1})$, was unmöglich ist).

Schließlich ist $\mathcal{L}(e_1, \dots, e_i) \subseteq \mathcal{L}(v_1, \dots, v_i) = \mathcal{L}(e_1, \dots, e_{i-1}, v_i)$ und umgekehrt $v_i \in \mathcal{L}(e_1, \dots, e_i)$, also stimmen beide Mengen überein. \square

Im folgenden Beispiel verwenden wir das „kanonische“ Skalarprodukt im \mathbb{R}^n :

$$\langle v, w \rangle = \sum v_i w_i.$$

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 3 \\ -3 \\ 3 \end{pmatrix}$$

$$w_1 = v_1, \quad w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} - \frac{2}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

$$w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 = \begin{pmatrix} 3 \\ -3 \\ 3 \end{pmatrix} - \frac{6}{2} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} - \frac{-6}{6} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Wenn in einem Euklidischen Vektorraum eine Orthonormalbasis gewählt wird, vereinfachen sich die Rechnungen:

Lemma 3.1.2 Sei $\{e_1, \dots, e_n\}$ eine Orthonormalbasis von V und $v = \sum v_i e_i$, $w = \sum w_i e_i$, dann gilt $v_i = \langle v, e_i \rangle$, also $v = \sum \langle v, e_i \rangle e_i$, $\langle v, w \rangle = \sum v_i w_i$, $|v|^2 = \sum v_i^2$. \square

Lemma 3.1.3 (Besselsche Ungleichung) Sei $\{e_1, \dots, e_k\}$ ein Orthonormalsystem und $v \in V$, dann gilt $|v|^2 \geq \sum \langle v, e_i \rangle^2$.

Beweis: Wir ergänzen $\{e_1, \dots, e_k\}$ zu einer Orthonormalbasis (das geht!) und haben

$$|v|^2 = \sum_{i=1}^k \langle v, e_i \rangle^2 + \sum_{i=k+1}^n \langle v, e_i \rangle^2,$$

die zweite Summe ist nichtnegativ. \square

Definition: Sei $U \subseteq V$ ein Unterraum, dann sei

$$U^\perp = \{v \in V \mid \langle v, u \rangle = 0 \text{ für alle } u \in U\}.$$

U^\perp heißt das orthogonale Komplement von U .

Lemma 3.1.4 U^\perp ist ein Unterraum von V , $\dim U^\perp = \dim V - \dim U$; es gilt $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$, $U^{\perp\perp} = U$ und $U \cap U^\perp = \{o\}$. Wenn $U \subset W$ ist, so ist $W^\perp \subset U^\perp$.

Beweis: Seien $v_1, v_2 \in U^\perp$, also $\langle v_1, u \rangle = \langle v_2, u \rangle = 0$ für alle $u \in U$; sei $r \in \mathbb{R}$, dann gilt $\langle v_1 + r v_2, u \rangle = 0$, also ist U^\perp ein Unterraum.

Die Dimensionsbeziehung sehen wir, wenn wir eine Orthonormalbasis von U wählen und diese zu einer Orthonormalbasis von V ergänzen: die ergänzten Vektoren liegen in U^\perp .

Sei $v \in (U_1 + U_2)^\perp$, also $\langle v, u_1 + u_2 \rangle = 0$ für alle $u_i \in U_i$, dann gilt speziell $\langle v, u_1 \rangle = \langle v, u_2 \rangle = 0$, also $v \in U_1^\perp \cap U_2^\perp$.

Sei umgekehrt $v \in U_1^\perp \cap U_2^\perp$, also $\langle v, u_1 \rangle = \langle v, u_2 \rangle = 0$, dann folgt $\langle v, u_1 + u_2 \rangle = 0$, also $v \in (U_1 + U_2)^\perp$.

Es gilt $v \in U^{\perp\perp}$ genau dann, wenn $\langle v, w \rangle = 0$ für alle $w \in U^\perp$ gilt, also für alle w mit $\langle w, u \rangle = 0$ für alle $u \in U$.

Sei nun $u \in U$, dann ist $\langle u, v \rangle = 0$ für alle $v \in U^\perp$, also $u \in U^{\perp\perp}$, d.h. $U \subseteq U^{\perp\perp}$ und die Gleichheit ergibt sich aus der Dimensionsformel.

Sei schließlich $U \subset W, v \in W^\perp$, dann ist $\langle v, w \rangle = 0$ für alle $w \in W$, speziell ist $\langle v, u \rangle = 0$ für alle $u \in U$, also $v \in U^\perp$. \square

Wir wollen unsere Kenntnisse für die Elementargeometrie nutzen:

Satz 3.1.2 Die Mittelsenkrechten eines Dreiecks schneiden sich in einen Punkt.

Beweis: Die Mittelsenkrechte der Strecke AB ist die Menge allen Punkte P mit

$$\left| \overrightarrow{AP} \right| = \left| \overrightarrow{PB} \right|.$$

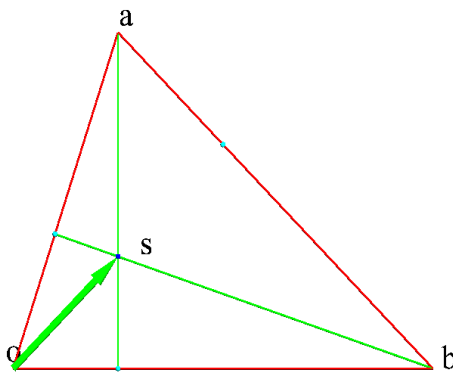
Für den Schnittpunkt S zweier Mittelsenkrechten gilt also

$$\left| \overrightarrow{AS} \right| = \left| \overrightarrow{SB} \right| = \left| \overrightarrow{SC} \right|,$$

also liegt S auch auf der dritten Mittelsenkrechten. □

Satz 3.1.3 Die Höhen im Dreieck schneiden sich in einen Punkt.

1. Beweis:



Es seien $p = \frac{\langle a, b \rangle}{\langle b, b \rangle} b$, $q = \frac{\langle a, b \rangle}{\langle a, a \rangle} a$ die Projektionen zweier Dreiecksseiten aufeinander, dann sind $l = \frac{\langle a, b \rangle}{\langle b, b \rangle} b - a$ und $m = \frac{\langle a, b \rangle}{\langle a, a \rangle} a - b$ zwei Höhen im Dreieck. Für deren Schnittpunkt $s = a + rl = b + tm$ gilt also

$$a + r \left(\frac{\langle a, b \rangle}{\langle b, b \rangle} b - a \right) = b + t \left(\frac{\langle a, b \rangle}{\langle b, b \rangle} a - b \right)$$

$$a \left(1 - r - t \frac{\langle a, b \rangle}{\langle a, a \rangle} \right) = b \left(1 - t - r \frac{\langle a, b \rangle}{\langle b, b \rangle} \right)$$

Die Koeffizienten von a und b sind Null, wir erweitern mit dem Nennern und erhalten

$$\begin{aligned} \langle a, a \rangle r + \langle a, b \rangle t &= \langle a, a \rangle \\ \langle a, b \rangle r + \langle b, b \rangle t &= \langle b, b \rangle \end{aligned}$$

Wir verwenden die Cramersche Regel, mit $n = \langle a, a \rangle \langle b, b \rangle - \langle a, b \rangle^2$ ist

$$r = (\langle a, a \rangle \langle b, b \rangle - \langle a, b \rangle \langle b, b \rangle) / n.$$

Somit ist

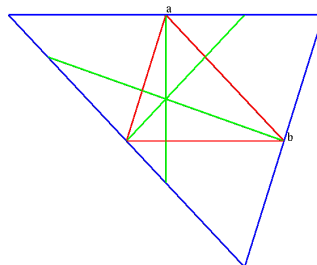
$$s = a + (\langle a, a \rangle \langle b, b \rangle - \langle a, b \rangle \langle b, b \rangle) / n \cdot \left(\frac{\langle a, b \rangle}{\langle b, b \rangle} b - a \right)$$

der Schnittpunkt beider Höhen. Dann ist

$$\begin{aligned} ns &= (\langle a, a \rangle \langle b, b \rangle - \langle a, b \rangle^2) a + (\langle a, a \rangle - \langle a, b \rangle) \langle b, b \rangle \cdot \left(\frac{\langle a, b \rangle}{\langle b, b \rangle} b - a \right) \\ &= (\langle a, a \rangle \langle b, b \rangle - \langle a, b \rangle^2) a + (-\langle a, a \rangle + \langle a, b \rangle) \langle b, b \rangle a + (\langle a, a \rangle - \langle a, b \rangle) \cdot \langle a, b \rangle b \\ &= (-\langle a, b \rangle^2 + \langle a, b \rangle \langle b, b \rangle) a + (\langle a, a \rangle - \langle a, b \rangle) \cdot \langle a, b \rangle b \end{aligned}$$

Nun ist es eine Schönschreibaufgabe, $\langle ns, a - b \rangle = 0$ nachzuweisen. Damit liegt s auch auf der dritten Höhe.

2. Beweis (Gauß 1810)



Wir ziehen durch die Eckpunkte Parallelen zu den Gegenseiten, die Mittelsenkrechten des großen Dreiecks sind dann die Höhen des gegebenen Dreiecks.

3. Beweis (Koecher/Krieg, Ebene Geometrie) mit $[a, b, c] = \det(a, b) + \det(b, c) + \det(c, a)$ ist der Höhenschnittpunkt

$$s = \frac{1}{[a, b, c]} (\langle a, b - c \rangle a^\perp + \langle b, c - a \rangle b^\perp + \langle c, a - b \rangle c^\perp).$$

□

3.2 Orthogonale Abbildungen und Matrizen

V und W seien Euklidische Vektorräume und $f : V \rightarrow W$ sei eine lineare Abbildung, f heißt orthogonale Abbildung, wenn

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

für alle $v, w \in V$ gilt.

Lemma 3.2.1 *Sei $f : V \rightarrow W$ eine orthogonale Abbildung, dann gilt $|f(v)| = |v|$. Wenn v auf w senkrecht steht, so steht $f(v)$ auf $f(w)$ senkrecht und der Winkel zwischen $f(v)$ und $f(w)$ ist gleich dem Winkel zwischen v und w . \square*

Lemma 3.2.2 *Wenn $|f(v)| = |v|$ für alle $v \in V$ gilt, so ist f eine orthogonale Abbildung.*

Beweis: $\langle v + w, v + w \rangle = \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle$, also ist

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

und damit

$$\langle f(v), f(w) \rangle = \frac{1}{2}(|f(v + w)|^2 - |f(v)|^2 - |f(w)|^2). \quad \square$$

Lemma 3.2.3 *Orthogonale Abbildungen sind injektiv.*

Beweis: Sei $f(v) = o$, dann ist $0 = \langle f(v), f(v) \rangle = \langle v, v \rangle$, also ist $v = o$. \square

Satz 3.2.1 *Sei $f : V \rightarrow V$ ein orthogonaler Endomorphismus und $B = \{e_1, \dots, e_n\}$ eine Orthonormalbasis von V . Dann ist $A_{BB}(f)^T = A_{BB}(f)^{-1}$.*

Beweis: Sei $f(e_i) = \sum f_{ji}e_j$, dann ist $A_{BB}(f) = (f_{ji}) = F$ und $\langle f(e_i), f(e_k) \rangle = \langle \sum f_{ji}e_j, \sum f_{lk}e_l \rangle = \sum f_{ji}f_{lk} = \langle e_i, e_k \rangle = \delta_{ik}$, d.h. $F^T F = E$. \square

Definition: Eine Matrix A mit $A^T = A^{-1}$ heißt orthogonal.

Also gehören zu orthogonalen Abbildungen bezüglich Orthonormalbasen orthogonale Darstellungsmatrizen.

Man kann die Orthogonalität einer Matrix so veranschaulichen: Bezüglich des Skalarprodukts $\langle v, w \rangle = \sum v_i w_i$ im \mathbb{R}^n sind die Beträge der Zeilen und der Spalten gleich 1, das Skalarprodukt verschiedener Zeilen oder Spalten ist null.

Lemma 3.2.4 *Das Produkt und die Inverse von orthogonalen Matrizen sind orthogonal. Die n -reihigen orthogonalen Matrizen bilden eine Gruppe $O(n)$, die „orthogonale“ Gruppe.*

Beweis: Sei $A^T A = E, B^T B = E$, dann ist $(AB)^T AB = B^T A^T AB = B^T B = E$. \square

Die Determinante einer orthogonalen Matrix hat den Wert -1 oder 1, eine Matrix mit Determinante 1 heißt speziell und die speziellen orthogonalen Matrizen bilden die „spezielle orthogonale“ Gruppe $SO(n)$.

Wir wollen uns eine Übersicht über die orthogonalen Matrizen verschaffen. Für kleine Matrizen ist dies trivial: $O(1) = \{1, -1\}$.

Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O(2)$, dann ist $E = A^T A = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix}$, wir setzen $a = \cos w, b = \sin w$ und aus den anderen Relationen folgt $c = \pm b, d = \mp a$, also treten zwei Fälle auf:

$$A = \begin{pmatrix} \cos w & \sin w \\ -\sin w & \cos w \end{pmatrix},$$

dies ist eine Drehung, oder

$$A = \begin{pmatrix} \cos w & \sin w \\ \sin w & -\cos w \end{pmatrix},$$

dies ist das Produkt einer Drehung und einer Spiegelung.

Satz 3.2.2 Jede Matrix $A \in SO(n)$ läßt sich als Produkt von $\frac{n(n-1)}{2}$ Drehmatrizen darstellen.

Beweis: Die Matrix, die in der Ebene der i -ten und der j -ten Koordinatenachse eine Drehung um den Winkel w veranstaltet, bezeichnen wir mit $D_{ij}(w)$.

Wir multiplizieren die Matrix A der Reihe nach mit D_{12}, D_{13}, \dots und bestimmen w so, daß im Produkt an den Stellen $(1, 2), (1, 3), \dots$ Nullen stehen und an der Stelle $(1, 1)$ eine positive Zahl steht. Das diese beiden Forderungen erfüllbar sind, möge sich der Leser bitte klarmachen (eine analoge Rechnung haben wir im Zusammenhang mit der Jacobi-Diagonalisierung durchgeführt). Nach $n - 1$ Schritten haben wir

$$AD = \begin{pmatrix} a & 0 & \dots & 0 \\ & & \dots & \\ * & & \dots & * \end{pmatrix},$$

diese Matrix ist orthogonal, also ist der Betrag der ersten Zeile gleich 1, also ist $a = 1$. Auch der Betrag der ersten Spalte ist gleich 1, also stehen in der ersten Spalte unter der 1 lauter Nullen. So fahren wir fort, bis wir als Produkt die Einheitsmatrix erhalten.

\square

Als nächstes wollen wir eine besondere Klasse orthogonaler Abbildungen untersuchen: Sei V ein Vektorraum und $a \neq o$ ein Vektor aus V . Wir definieren eine Abbildung $s_a : V \rightarrow V$ durch

$$s_a(w) = w - 2 \frac{\langle a, w \rangle}{|a|^2} a.$$

Diese Abbildung ist orthogonal, denn

$$\langle s_a(v), s_a(w) \rangle = \langle v - 2 \frac{\langle a, v \rangle}{|a|^2} a, w - 2 \frac{\langle a, w \rangle}{|a|^2} a \rangle = \langle v, w \rangle.$$

Es gilt $s_{ra} = s_a$ für $r \in \mathbb{R}$ und $s_a(a) = -a$. Wenn aber $\langle a, w \rangle = 0$ ist, so gilt $s_a(w) = w$, wie man leicht nachrechnet. Also: Die Elemente von $\mathcal{L}(a)^\perp$ werden bei s_a nicht verändert, d.h. s_a ist die Spiegelung an $\mathcal{L}(a)^\perp$.

Daß s_a eine lineare Abbildung ist, folgt aus dem

Satz 3.2.3 Sei $f : V \rightarrow V$ eine Abbildung mit $\langle f(v), f(w) \rangle = \langle v, w \rangle$ für alle $v, w \in V$, dann ist f linear.

Beweis:

$$\begin{aligned} & \langle f(v+w) - f(v) - f(w), f(v+w) - f(v) - f(w) \rangle \\ &= \langle f(v+w), f(v+w) \rangle - \langle f(v+w), f(v) \rangle - \dots \\ &= \langle v+w, v+w \rangle - \langle v+w, v \rangle - \dots = 0. \square \end{aligned}$$

Sei nun $U \subseteq V$ ein Unterraum und $a \in U$, dann ist $s_a : U \rightarrow U$ eine Spiegelung in U . Wenn aber v ein beliebiger Vektor aus V ist, so ist auch $s_a(v)$ definiert, wir haben also eine Spiegelung des gesamten Raums V , die die Spiegelung von U fortsetzt und U^\perp festhält (es gilt ja $U^\perp \subseteq \mathcal{L}(v)^\perp$).

Satz 3.2.4 Sei $f : V \rightarrow V$ eine orthogonale Abbildung, $f \neq id$, $\dim V = n$, dann ist f ein Produkt von höchstens n Spiegelungen.

Beweis: Wir führen die Induktion über n .

Wenn $n = 1$ und $f \neq id$ ist, so ist $f(v) = -v$ und dies ist eine Spiegelung.

Sei der Satz für die Dimension $n - 1$ bewiesen.

Wir treffen noch eine zusätzliche Voraussetzung: Es soll ein Vektor $v \neq 0$ existieren, für den $f(v) = v$ gilt.

Wenn nun $\langle w, v \rangle = 0$ ist, so ist auch $\langle f(w), f(v) \rangle = \langle f(w), v \rangle = 0$, also ist $H = \mathcal{L}(v)^\perp$ ein invarianter Unterraum. Also ist die Einschränkung $f|_H$ von f auf H ein Produkt von höchstens $n - 1$ Spiegelungen. Dies sind auch Spiegelungen von V , die $H^\perp = \mathcal{L}(v)$ festlassen, also ist ihr Produkt gleich f .

Wir betrachten nun den allgemeinen Fall. Da $f \neq id$ ist, gibt es einen Vektor v mit $f(v) \neq v$. Wir setzen $w = f(v) - v$, $H = \mathcal{L}(w)^\perp$, sei s_w die Spiegelung an H . Wir zeigen $s_w(f(v)) = v$:

Es ist $s_w(w) = -w$, also

$$s_w(f(v) - v) = -f(v) + v = s_w(f(v)) - s_w(v)$$

und es gilt

$$\langle f(v) + v, f(v) - v \rangle = \langle f(v), f(v) \rangle - \langle v, v \rangle = 0,$$

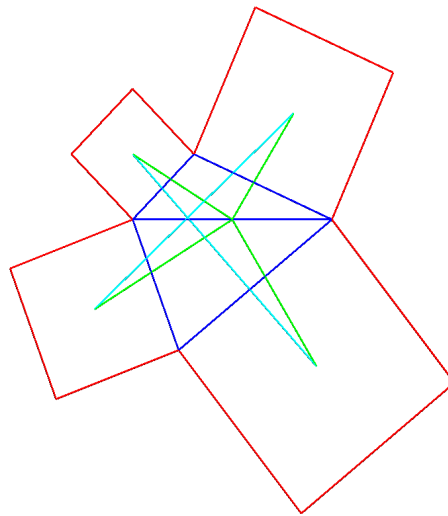
also liegt $f(v) + v$ in H und damit ist

$$s_w(f(v) + v) = f(v) + v = s_w(f(v)) + s_w(v)$$

Beide Gleichungen addiert ergeben $2s_w(f(v)) = 2v$. Damit erfüllt die Abbildung $s_w \circ f$ die obige spezielle Voraussetzung, ist also Produkt von höchstens $n - 1$ Spiegelungen. Damit ist f ein Produkt von höchstens n Spiegelungen. \square

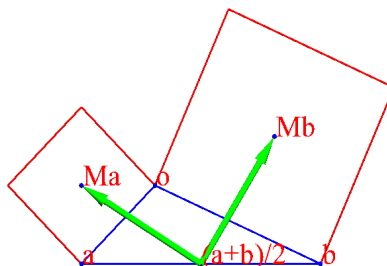
Beispiel: Wir betrachten $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$; es ist $Ae_1 \neq e_1$, wir wählen $v = e_1$. Es ist $w = Ae_1 - e_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $s_w(Av) = s_w\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - 2\frac{1}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $s_w(Ae_2) = s_w\left(\begin{pmatrix} -1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} - 1\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, also ist $s_w \circ A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ die Spiegelung an e_2 .

Der Satz von van Aubel (1878)



Satz 3.2.5 *In einem Viereck sind die Strecken zwischen den Mittelpunkten gegenüberliegender auf den Seiten errichteter Quadrate gleichlang und zueinander orthogonal.*

Beweis: Wir zerlegen die Figur entlang einer Diagonale:



Die Dreiecksseitenvektoren seien $a = (a_1, a_2)$ und $b = (b_1, b_2)$, die dazu senkrechten Quadratseiten sind dann $d = (-a_2, a_1)$ und $c = (b_2, -b_1)$. Wir bezeichnen die Vektoren vom Mittelpunkt der dritten Dreiecksseite zu den Mittelpunkten der Quadrate mit v und w ; es ist $v = \frac{1}{2}(a + d - a - b) = \frac{1}{2}(d - b)$, $w = \frac{1}{2}(c - a)$. Nun gilt

$$\langle v, w \rangle = \langle d - b, c - a \rangle = \langle d, c \rangle + \langle b, a \rangle = -a_2 b_2 - a_1 b_1 + a_1 b_1 + a_2 b_2 = 0.$$

Durch eine analoge Rechnung erhält man $|v| = |w|$. Der Vektor w geht aus v durch eine Drehung um 90 Grad hervor.

Genauso geht es für die untere Hälfte: die Strecken vom Diagonalenmittelpunkt zu den Quadratmittelpunkten sind gleichlang und orthogonal. Also sind die beiden im oberen Bild eingezeichneten Dreiecke kongruent, also sind die dritten Dreiecksseiten (die Verbindungen der Quadratmittelpunkte) gleichlang und orthogonal. \square

3.3 Die adjungierte Abbildung

Das Skalarprodukt auf V ist eine bilineare Abbildung, dazu gehört also eine Abbildung $t: V \rightarrow V^*$, $t(v)(w) = \langle v, w \rangle$.

Lemma 3.3.1 *Die Abbildung t ist bijektiv.*

Beweis: Sei $t(v) = o$, d.h. $t(v)(w) = \langle v, w \rangle = 0$ für alle $w \in V$, speziell ist $t(v)(v) = \langle v, v \rangle = 0$, also $v = o$. Also ist t injektiv und wegen der Dimensionsgleichheit von V und V^* ist t bijektiv. \square

Folgerung 3.3.1 *Sei $l: V \rightarrow \mathbb{R}$ eine Linearform, dann gibt es einen eindeutig bestimmten Vektor $w \in V$, so daß $l(v) = \langle v, w \rangle$ für alle $v \in V$ gilt.*

Beweis: Sei $w = t^{-1}(l)$, dann ist $t(w) = l$, also $t(w, v) = l(v) = \langle v, w \rangle$. \square

Satz 3.3.1 *Seien V, W Euklidische Vektorräume und $f: V \rightarrow W$ eine lineare Abbildung. Dann gibt es eine eindeutig bestimmte lineare Abbildung $f^*: W \rightarrow V$ mit $\langle f(v), w \rangle = \langle v, f^*(w) \rangle$.*

Beweis: Für festes $w \in W$ ist die Zuordnung $v \rightarrow \langle f(v), w \rangle$ eine Linearform, also gibt es einen Vektor $u \in V$ mit $\langle f(v), w \rangle = \langle v, u \rangle$. Wir setzen dann $f^*(w) = u$. Die Linearität von f^* ergibt sich aus der Linearität von $\langle v, \cdot \rangle$. \square

Definition: f^* heißt die zu f adjungierte Abbildung.

Lemma 3.3.2 $(f + g)^* = f^* + g^*$, $id^* = id$, $(f \circ g)^* = g^* \circ f^*$, $f^{**} = f$. \square

Beweis: Zunächst zeigen wir, daß aus $\langle v, w \rangle = \langle v, u \rangle$ für alle v folgt, daß $w = u$:

$\langle v, w - u \rangle = 0$, wir setzen $v = w - u$: $\langle w - u, w - u \rangle = 0$, also $w - u = o$.

$$\langle (f + g)(v), w \rangle = \langle f(v), w \rangle + \langle g(v), w \rangle = \langle v, f^*(w) \rangle + \langle v, g^*(w) \rangle = \langle v, (f + g)^*(w) \rangle$$

$$\langle f(g(v)), w \rangle = \langle v, (fg)^*(w) \rangle = \langle g(v), f^*(w) \rangle = \langle v, g^*(f^*(w)) \rangle$$

$$\langle f(v), w \rangle = \langle v, f^*(w) \rangle = \langle f^*(w), v \rangle = \langle w, f^{**}(v) \rangle$$

\square

Lemma 3.3.3 Sei $B = \{e_1, \dots, e_n\}$ eine Orthonormalbasis von V und $f : V \rightarrow V$ eine lineare Abbildung. Dann ist $A_{BB}(f^*) = A_{BB}(f)^T$.

Beweis: Sei $f(e_i) = \sum f_{ji}e_j$, dann gilt

$$\langle f(e_i), e_k \rangle = \langle e_i, f^*(e_k) \rangle = \langle \sum f_{ji}e_j, e_k \rangle = \sum f_{ji} \langle e_j, e_k \rangle = f_{ki},$$

also

$$f^*(e_k) = \sum f_{ki}e_i \quad \square$$

Lemma 3.3.4 Wenn f eine orthogonale Abbildung ist, so ist $f^* = f^{-1}$.

Beweis: $\langle f(v), f(w) \rangle = \langle v, f^*(f(w)) \rangle = \langle v, w \rangle$ für alle v , also ist $f^* \circ f = id$. □

Definition: Wenn $f^* = f$ ist, so heißt f selbstadjungiert.

Lemma 3.3.5 Sei B eine Orthonormalbasis, f ist genau dann selbstadjungiert, wenn $A_{BB}(f)$ eine symmetrische Matrix ist. □

Wir wollen für kurze Zeit als Grundkörper den Körper \mathbb{C} der komplexen Zahlen wählen. Hier ist die durch $\langle \sum v_i e_i, \sum w_j e_j \rangle = \sum v_i w_i$ gegebene Bilinearform nicht mehr positiv definit. Abhilfe schaffen hier Hermitesche Formen:

Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ heißt Hermitesch, wenn sie linear im ersten Faktor ist und $\langle v, w \rangle = \overline{\langle w, v \rangle}$ gilt (der Strich bedeutet die konjugiert komplexe Zahl). Für Hermitesche Formen gelten die bisher bewiesenen Resultate ebenfalls, blättern Sie zurück und beweisen Sie es.

Eine Abbildung, für die $\langle f(v), f(w) \rangle = \langle v, w \rangle$ gilt, heißt hier unitär. Die Darstellungsmatrix B zur adjungierten Abbildung ist die komplex konjugierte der Transponierten der Darstellungsmatrix A der gegebenen Abbildung, die Matrix B wird dann als die zu A adjungierte Matrix bezeichnet: $B = A^*$.

Satz 3.3.2 Die Eigenwerte einer selbstadjungierten Abbildung sind reell.

Beweis: Sei $f(v) = zv$, dann ist

$$\langle f(v), v \rangle = \langle zv, v \rangle = z \langle v, v \rangle = \langle v, f(v) \rangle = \langle v, zv \rangle = \bar{z} \langle v, v \rangle,$$

also ist $z = \bar{z}$ reell. □

Wir wollen nun die folgende Frage beantworten:

V sei ein Euklidischer Vektorraum und $f : V \rightarrow V$ ein Endomorphismus. Gibt es dann eine Orthonormalbasis von V , die aus Eigenvektoren von f besteht?

Oder anders ausgedrückt: Sei eine n -reihige Matrix A gegeben. Gibt es dann eine orthogonale (bzw. unitäre) Matrix X , so daß X^*AX eine Diagonalmatrix ist?

Wir werden sehen, daß dies für eine spezielle Klasse von Endomorphismen bzw. Matrizen der Fall ist.

Der folgende Satz stammt von I. Schur.

Satz 3.3.3 (Schur-Zerlegung) *Zu jeder Matrix A gibt es eine unitäre Matrix U , so daß U^*AU eine obere Dreiecksmatrix ist.*

Beweis: Wir führen die Induktion über n . Für $n = 1$ ist nichts zu zeigen. Sei der Satz für $(n - 1)$ -reihige Matrizen bewiesen.

Sei z ein Eigenwert und u_1 ein entsprechender Eigenvektor von A , also $Au_1 = zu_1$. Wir ergänzen u_1 zu einer Orthonormalbasis $\{u_1, \dots, u_n\}$ von \mathbb{C}^n , also gilt $u_i^*u_j = \delta_{ij}$, d.h. die Matrix U_1 mit den Spalten u_1, \dots, u_n ist unitär. Es gilt

$$AU_1 = U_1 \begin{pmatrix} z & \star \\ 0 & \\ \dots & \\ 0 & A_1 \end{pmatrix}$$

also

$$U_1^*AU_1 = \begin{pmatrix} z & \star \\ 0 & \\ \dots & \\ 0 & A_1 \end{pmatrix}$$

Nun sei U_2 eine unitäre Matrix, so daß $U_2^*A_1U_2 = D_1$ eine Dreiecksmatrix ist. Dann ist

$$\begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}^* U_1^*AU_1 \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix} = \begin{pmatrix} z & \star \\ 0 & D_1 \end{pmatrix}$$

eine Dreiecksmatrix. □

Satz 3.3.4 (reelle Schur-Zerlegung) *Zu einer reellen Matrix A gibt es eine orthogonale Matrix Q , so daß $Q^T A Q$ eine Block-Dreiecksgestalt hat, deren Diagonalblöcke \mathbb{R}_k die folgende Form haben: $\mathbb{R}_k = (d_k) \in M_{11}$, wenn d_k ein reeller Eigenwert von A ist, und $\mathbb{R}_k \in M_{22}$, wenn $c \pm is$ ein Paar komplexer Eigenwerte von A ist.*

Beweis: Reelle Eigenwerte werden wie oben bearbeitet.

Sei $z = c + is$ ein komplexer Eigenwert von A , dann gibt es $x, y \in \mathbb{R}^n$ mit

$$A(x + iy) = (c + is)(x + iy),$$

also

$$Ax = cx - sy, \quad Ay = sx + cy.$$

Die Zahl $c - is$ ist ebenfalls ein Eigenwert von A , der zugehörige Eigenvektor ist $x - iy$, denn

$$A(x - iy) = cx - sy - isx - icy = (c - is)(x - iy).$$

Da $c + is \neq c - is$ ist, ist $\{x + iy, x - iy\}$ eine linear unabhängige Menge. Wir zeigen, daß auch $\{x, y\}$ linear unabhängig ist.

Sei $rx + ty = o$, wir betrachten $(r - it)(x + iy)$ und dem dazu komplex konjugieren Vektor:

$$(r - it)(x + iy) + (r + it)(x - iy) = 2(rx + ty) = o,$$

wegen der linearen Unabhängigkeit von $\{x + iy, x - iy\}$ muß $r + it = 0$ sein, d.h. $r = t = 0$.

Aus den obigen Beziehungen sehen wir, daß $\mathcal{L}(x, y)$ ein A -invarianter Unterraum ist, in dem wir eine Orthonormalbasis $\{u, v\}$ wählen. Nun schreiben wir die Komponenten von u und v in die ersten beiden Spalten von Q und verfahren weiter wie bei der komplexen Schur-Zerlegung. \square

Definition: Ein Endomorphismus $f : V \rightarrow V$ heißt normal, wenn $f \circ f^* = f^* \circ f$ gilt. Eine Matrix A heißt normal, wenn $AA^* = A^*A$ gilt.

Lemma 3.3.6 Sei A normal und U unitär, dann ist U^*AU normal.

Beweis: $(U^*AU)(U^*AU)^* = U^*AUU^*A^*U = U^*AA^*U = U^*A^*AU = U^*A^*UU^*AU = (U^*AU)^*(U^*AU)$. \square

Lemma 3.3.7 Eine normale Dreiecksmatrix hat Diagonalgestalt.

Beweis: Sei $A = (a_{ij})$, $a_{ij} = 0$ für $i > j$, dann ist das j -te Diagonalelement der Matrix AA^* gleich

$$\sum_{i=1}^n a_{ji}\bar{a}_{ji}$$

und das j -te Diagonalelement von A^*A ist gleich

$$\sum_{i=1}^n \bar{a}_{ij}a_{ij},$$

aus der Gleichheit folgt der Reihe nach $a_{12} = \dots = a_{1n} = 0$, $a_{23} = \dots = a_{2n} = 0$ usw.

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ & \ddots & \vdots \\ & & a_{nn} \end{pmatrix} \begin{pmatrix} \bar{a}_{11} & & \\ \vdots & \ddots & \\ \bar{a}_{1n} & \dots & \bar{a}_{nn} \end{pmatrix} = \begin{pmatrix} a_{11}\bar{a}_{11} + a_{12}\bar{a}_{12} + \dots & & \\ & \ddots & \\ & & \bar{a}_{nn} \end{pmatrix}$$

$$\begin{pmatrix} \bar{a}_{11} & & \\ \vdots & \ddots & \\ \bar{a}_{1n} & \dots & \bar{a}_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1n} \\ & \ddots & \vdots \\ & & a_{nn} \end{pmatrix} = \begin{pmatrix} \bar{a}_{11}a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{pmatrix}$$

\square

Folgerung 3.3.2 (Spektralsatz) 1. Wenn A eine normale Matrix ist, so gibt es eine unitäre Matrix U , so daß U^*AU eine Diagonalmatrix ist.

2. Wenn $f : V \rightarrow V$ ein normaler Endomorphismus ist, so besitzt V eine Orthonormalbasis $\{e_1, \dots, e_n\}$ aus Eigenvektoren von f .

3. Wenn die e_i eine Orthonormalbasis von Eigenvektoren von f bilden, dann sind die e_i sind auch Eigenvektoren von f^* .

Beweis: 1. Es gibt eine unitäre Matrix U , so daß U^*AU eine normale Dreiecksmatrix ist.

2. ist äquivalent zu 1.

3. Sei $f(e_i) = z_i e_i$, dann ist

$$\langle f(e_i), e_j \rangle = \langle z_i e_i, e_j \rangle = \langle e_i, \bar{z}_i e_j \rangle = z_i \delta_{ij} = \langle e_i, f^*(e_j) \rangle,$$

dies ist gleich Null für $i \neq j$, also liegt $f^*(e_j)$ in $\mathcal{L}(e_j)$, für $j = i$ erhalten wir $\langle e_i, f^*(e_i) \rangle = \bar{z}_i$, also $f^*(e_i) = \bar{z}_i e_i$. \square

Es gilt aber auch die Umkehrung:

Satz 3.3.5 *Der Vektorraum V besitze eine Orthonormalbasis aus Eigenvektoren von $f : V \rightarrow V$. Dann ist f normal.*

Beweis: Wie oben folgt $f^*(e_i) = \bar{z}_i e_i$, dann ist $f \circ f^*(e_i) = z_i \bar{z}_i e_i = f^* \circ f(e_i)$ für alle i , also $ff^* = f^*f$. \square

Analog beweist man den

Satz 3.3.6 *Sei $f : V \rightarrow V$ normal. Wenn alle Eigenwerte von f reell sind, so ist f selbstadjungiert. Wenn alle Eigenwerte von f den Betrag 1 haben, so ist f unitär.* \square

Wir wenden dies an: Sei A eine symmetrische Matrix mit den Eigenwerten $z_1 \geq z_2 \geq z_3 \geq \dots \geq z_n$ und $Q = (q_1 \ \dots \ q_n)$ eine orthogonale Matrix mit

$$A = QDQ^T = (q_1 \ \dots \ q_n) \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{pmatrix} \begin{pmatrix} q_1^T \\ \vdots \\ q_n^T \end{pmatrix} = z_1 q_1 q_1^T + \dots + z_n q_n q_n^T.$$

Die z_i tragen Informationen über A , aber die größeren mehr als die kleineren ($z_i = 0$ trägt gar nichts bei). Wenn man nun A ungefähr, aber mit geringerem Aufwand beschreiben will, so setzt man eine Grenze für die zu berücksichtigenden Eigenwerte und betrachtet $z_1 q_1 q_1^T + \dots + z_r q_r q_r^T$.

Satz 3.3.7 *Sei $f : V \rightarrow V$ ein normaler Endomorphismus, dann gilt $\langle f(v), f(w) \rangle = \langle f^*(v), f^*(w) \rangle$.*

Beweis: $\langle f(v), f(w) \rangle = \langle v, f^* f(w) \rangle \langle v, f f^*(w) \rangle = \langle v, f^* f^* f^*(w) \rangle = \langle f^*(v), f^*(w) \rangle$ \square

Satz 3.3.8 *Wenn $\langle f(v), f(w) \rangle = \langle f^*(v), f^*(w) \rangle$ für alle $v, w \in V$ gilt, dann ist f normal.*

Beweis:

$$\langle f(v), f(w) \rangle + \langle f(w), f(v) \rangle = \langle f(v), f(w) \rangle + \overline{\langle f(v), f(w) \rangle} = 2\operatorname{Re}\langle f(v), f(w) \rangle$$

$$\langle f^*(v), f^*(w) \rangle + \langle f^*(w), f^*(v) \rangle = \langle f^*(v), f^*(w) \rangle + \overline{\langle f^*(v), f^*(w) \rangle} = 2\operatorname{Re}\langle f^*(v), f^*(w) \rangle,$$

also $\operatorname{Re}\langle f(v), f(w) \rangle = \operatorname{Re}\langle f^*(v), f^*(w) \rangle$.

$$\langle f(v), if(w) \rangle + \langle if(w), f(v) \rangle = -i\langle f(v), f(w) \rangle + i\overline{\langle f(v), f(w) \rangle} = 2\operatorname{Im}\langle f(v), f(w) \rangle,$$

also $\operatorname{Im}\langle f(v), f(w) \rangle = \operatorname{Im}\langle f^*(v), f^*(w) \rangle$. Nun ist

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

für alle $w \in V$, also $f^* f = f f^*$. \square

Den folgenden Zusammenhang werden wir später benötigen.

Satz 3.3.9 Sei $f : V \rightarrow W$ eine lineare Abbildung, dann ist $(\text{Ker}(f^*))^\perp = \text{Im}(f)$.

Beweis: Sei $f^*(v) = 0$, dann ist für alle Vektoren w die Gleichung

$$\langle f^*(v), w \rangle = 0 = \langle v, f(w) \rangle,$$

erfüllt, also liegt v in $(\text{Im}(f))^\perp$, also ist $\text{Im}(f) \subseteq (\text{Ker}(f^*))^\perp$. Sei F die Darstellungsmatrix von f bezüglich irgendwelcher Basen, sei $\dim V = n$, $\dim W = m$, dann ist

$$\dim \text{Im}(f) = \text{rg}(F) = \text{rg}(F^*) = \dim \text{Im}(f^*) = r$$

und

$$\dim \text{Ker}(f^*) = m - r,$$

also

$$\dim(\text{Ker}(f^*))^\perp = r,$$

daraus folgt die Behauptung. □

Anwendung (Fredholmsche Alternative): Das Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn $b \in \text{Ker}A^\perp$, also wenn $A^\perp b = 0$.

3.4 Pseudoinverse Matrizen

Wir wollen den Begriff der inversen Matrix auf nichtreguläre und nichtquadratische Matrizen verallgemeinern.

Wir betrachten Matrizen von fixierter Größe, und zwar seien sie stets entweder aus M_{nm} oder aus M_{mn} , wenn Produkte $A_1 A_2 A_3 \dots$ gebildet werden, so wollen wir immer voraussetzen, daß diese Matrizen sich auch wirklich multiplizieren lassen, also abwechselnd aus M_{mn} und aus M_{nm} sind.

Definition: Eine Matrix A heißt pseudo-regulär, wenn eine Matrix X mit $AXA = A$ existiert. Die Matrizen A und X heißen zueinander pseudoinvers, wenn $AXA = A$ und $XAX = X$ gilt.

Lemma 3.4.1 Wenn die Matrix A regulär ist, so ist sie auch pseudoregulär und die einzige zu A pseudo-inverse Matrix ist A^{-1} . □

Lemma 3.4.2 Wenn A pseudo-regulär ist, so besitzt es eine pseudo-inverse Matrix.

Beweis: Sei $AXA = A$ und $Y = XAX$, dann ist $AYA = AXAXA = AXA = A$, $YAY = XAXAXAX = XAXAX = XAX = Y$, also sind A und Y zueinander pseudo-invers. □

Satz 3.4.1 Sei $M = \{X \mid AXA = A\}$, dann ist jede Matrix $Y = X_1 A X_2$, wo $X_1, X_2 \in M$ sind, zu A pseudo-invers und jede zu A pseudoinverse Matrix hat diese Form. Seien A und X zueinander pseudoinvers, dann sind AX und XA idempotente Matrizen.

Beweis: 1. $AYA = AX_1AX_2A = AX_2A = A$, $YAY = X_1AX_2AX_1AX_2 = X_1AX_1AX_2 = X_1AX_2 = Y$.

2. Sei A zu Y pseudo-invers, dann liegt $Y = YAY$ in M .

3. $(AX)^2 = AXAX = AX$, analog für XA . □

Folgerung 3.4.1 *Seien A und B zueinander pseudo-invers, dann gilt $\text{rg}(A) = \text{rg}(B)$ und $\text{Sp}(AB) = \text{Sp}(BA) = \text{rg}(A)$.*

Beweis: Aus $ABA = A$ folgt $\text{rg}(A) \leq \text{rg}(B)$, aus Symmetriegründen folgt die Gleichheit. die zweite Aussage folgt aus der Idempotenz von AB . □

Satz 3.4.2 *Sei B pseudo-invers zu A . Das Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn $ABb = b$ ist und die allgemeine Lösung hat die Form $x = y - BAy + Bb$, wobei $y \in \mathbb{R}^n$ beliebig ist.*

Beweis: Sei u eine Lösung, also $Au = b$, dann ist $ABb = ABAu = Au = b$. Wenn umgekehrt $ABb = b$ gilt, so ist Bb eine Lösung.

Wegen $Ay - ABAy = 0$ ist $x = y - BAy$ eine Lösung des homogenen Systems $Ax = 0$; wenn umgekehrt $Ax = 0$ ist, so hat x die Form $x = x - BAx$. □

Wir betrachten ein Beispiel:

Sei $A = (1, 1)$, dann ist $B = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ zu A pseudo-invers. Wir betrachten das Gleichungssystem $Ax = 5$. Es ist $AB = (1)$, also $ABb = b$ und wegen $BA = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ ist die allgemeine Lösung gleich

$$\begin{pmatrix} a \\ b \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} \frac{5}{2} \\ \frac{5}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}a - \frac{1}{2}b + \frac{5}{2} \\ -\frac{1}{2}a + \frac{1}{2}b + \frac{5}{2} \end{pmatrix}$$

Definition: Die Matrix X heißt Moore-Penrose-Inverse der Matrix A , wenn

1. $AXA = A$, 2. $XAX = X$, 3. $(AX)^* = AX$, 4. $(XA)^* = XA$
gilt, d.h. A und X sind zueinander pseudo-invers und AX und XA sind selbstadjungiert.

Lemma 3.4.3 *Eine Matrix A besitzt höchstens eine Moore-Penrose-Inverse.*

Beweis: Seien X und Y pseudo-invers zu A , dann gilt

$$X = XAX = XX^*A^* = XX^*A^*Y^*A^* = XX^*A^*AY = XAXAY = XAY,$$

analog zeigt man $Y = XAY$, damit ist $X = Y$. □

Wir wollen nun zeigen, dass jede Matrix eine Moore-Penrose-Inverse besitzt. Dazu brauchen wir ein paar Hilfsbetrachtungen.

Lemma 3.4.4 *Wenn $AA^* = 0$ ist, so ist $A = 0$. Wenn $BA^*A = 0$ ist, so ist $BA^* = 0$.*

Beweis: Sei $AA^* = 0$, dann ist $\text{Sp}(AA^*) = 0$, diese Spur ist aber gleich $\sum a_{ij}\bar{a}_{ij}$, folglich sind alle $a_{ij} = 0$.

Wenn $BA^*A = 0$ ist, so ist auch $BA^*AB^* = (BA^*)(BA^*)^* = 0$ und daraus folgt die Behauptung. \square

Satz 3.4.3 *Sei A eine beliebige (rechteckige) Matrix. Das Minimalpolynom $m(z)$ von A^*A hat die Null höchstens als einfache Nullstelle.*

Beweis: Wir nehmen an, $m(z)$ hätte die Null als k -fache Nullstelle, $k > 1$, also $m(z) = g(z)z^k$. Dann gilt

$$g(A^*A)(A^*A)^k = 0,$$

daraus folgt

$$g(A^*A)(A^*A)^{k-1}A^* = 0$$

und

$$g(A^*A)(A^*A)^{k-1} = 0$$

im Widerspruch dazu, daß $m(z)$ das Minimalpolynom ist. \square

Wir bestimmen die Moore-Penrose-Inverse A^- von $A \in M_{mn}$ im Spezialfall, wenn A^*A regulär ist: $A^- = (A^*A)^{-1}A^*$. Wir weisen die 4 Eigenschaften nach:

$$AA^-A = A(A^*A)^{-1}A^*A = A,$$

$$A^-AA^- = (A^*A)^{-1}A^*A(A^*A)^{-1}A^* = (A^*A)^{-1}A^*,$$

$$(A^-A)^* = ((A^*A)^{-1}A^*A)^* = E^* = A^-A,$$

$$(AA^-)^* = (A(A^*A)^{-1}A^*)^* = A((A^*A)^{-1})^*A^* = A((A^*A)^*)^{-1}A^* = A(A^*A)^{-1}A^* = AA^-.$$

Der folgende Satz wurde von Penrose 1956 veröffentlicht.

Satz 3.4.4 *Jede Matrix A besitzt eine Moore-Penrose-Inverse A^- .*

Beweis: Wenn 0 kein Eigenwert von A^*A ist, so ist A^*A regulär, diesen Fall haben wir soeben behandelt. Andernfalls hat das Minimalpolynom $m(z)$ von A^*A (bis auf einen konstanten Faktor) die Form

$$m(z) = g(z)z^2 - z,$$

wir setzen

$$X = g(A^*A)A^*.$$

Es gilt

$$XAA^*A = g(A^*A)A^*AA^*A = g(A^*A)(A^*A)^2 = A^*A,$$

also $(XA - E)A^*A = 0$, daraus folgt $(XA - E)A^* = 0$, d.h.

$$XAA^* = A^*.$$

Nun folgt

$$(XA)^* = A^*X^* = XAA^*X^* = X(XAA^*)^* = XA,$$

$$AXA = A(XA)^* = AA^*X^* = A.$$

Die Matrix A^*A ist selbstadjungiert, damit ist auch die Matrix $g(A^*A)$ selbstadjungiert, daraus folgt

$$\begin{aligned} AX &= Ag(A^*A)A^* = Ag(A^*A)^*A^* = (Ag(A^*A)A^*)^* = (AX)^*, \\ XAX &= X(AX)^* = XX^*A^* = g(A^*A)A^*X^*A^* = g(A^*A)(AXA)^* = \\ &g(A^*A)A^* = X. \end{aligned}$$

Damit sie die vier Eigenschaften bewiesen. \square

Der Beweis liefert auch gleich ein Konstruktionsverfahren für die Moore-Penrose-Inverse, das aber in der Praxis nicht angewandt wird.

Wir wollen aber ein Beispiel betrachten. Sei $A = \begin{pmatrix} 3 & 2 & 1 \end{pmatrix}$, wir wollen die Moore-Penrose-Inverse von A bestimmen. Es ist

$$A^*A = \begin{pmatrix} 9 & 6 & 3 \\ 6 & 4 & 2 \\ 3 & 2 & 1 \end{pmatrix},$$

$rgA = rg(A^*A) = 1$, also hat das charakteristische Polynom die Form $z^3 - 14z^2$ und das Minimalpolynom ist $\frac{1}{14}z^2 - z$, also ist $X = \frac{1}{14}A^*$.

3.5 Unlösbare und unterbestimmte Gleichungssysteme

In praktischen Beispielen kommen lineare Gleichungssysteme vor, die zu wenige Gleichungen enthalten, um eine eindeutig bestimmte Lösung zu besitzen. Manchmal ist es sinnvoll, aus der Lösungsmannigfaltigkeit eine Lösung mit minimalem Betrag auszuwählen. Wie das zu geschehen hat, wollen wir uns ansehen.

Sei also $Ax = b$ ein lineares Gleichungssystem, $x \in \mathbb{R}^m, b \in \mathbb{R}^n$. Wir wissen vom Ende des letzten Kapitels, daß

$$\mathbb{R}^m = \text{Im}(A^*) \oplus \text{Ker}(A)$$

gilt. Sei x eine Lösung des Systems, wir zerlegen $x = x_1 + x_2$, wo x_1 in $\text{Im}(A^*)$ und x_2 in $\text{Ker}(A)$ liegt, dann ist

$$Ax = Ax_1 + Ax_2 = Ax_1,$$

also ist x_1 auch eine Lösung, weiter ist

$$|x|^2 = |x_1|^2 + |x_2|^2 \geq |x_1|^2$$

und die untere Grenze wird für $x \in \text{Im}(A^*)$ angenommen.

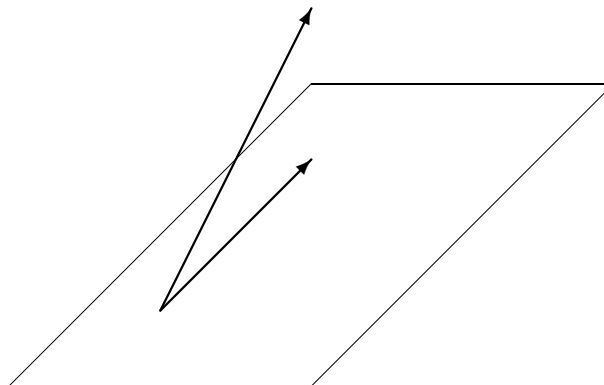
Weiter kann es vorkommen, daß ein Gleichungssystem nicht lösbar ist, die Ursache kann in kleinen (vernachlässigbaren) Ungenauigkeiten der Eingangsdaten liegen, die sich aber fatal auswirken: Der Gaußsche Algorithmus liefert nicht etwa eine Näherungslösung,

sondern gar keine. Wir wollen eine „Lösung“ x suchen, wo $|Ax - b|$ minimal ist, sicher ist das ein vernünftiger Kompromiß.

Der Wert von $|Ax - b|$ ist dann minimal, wenn Ax die orthogonale Projektion von b auf $\text{Im}(A)$ ist. Es ist

$$\mathbb{R}^n = \text{Im}(A) \oplus \text{Ker}(A^*)$$

wir zerlegen b entsprechend: $b = b_1 + b_2$. Da b_1 in $\text{Im}(A)$ liegt, ist $Ax = b_1$ lösbar und $|Ax - b|$ ist minimal. Dadurch ist x noch nicht notwendigerweise eindeutig bestimmt, evtl. ist noch eine „Lösung“ minimalen Betrags auszuwählen.



Der folgende Satz wurde von Penrose 1957 veröffentlicht.

Satz 3.5.1 Für $x = A^-b$ sind $|Ax - b|$ und $|x|$ minimal.

Beweis: Zu zeigen ist, daß x in $\text{Im}(A^*)$ liegt und daß $AA^-b = b_1$ ist. Die erste Aussage folgt aus

$$A^-b = g(A^*A)A^*b$$

(mit den Bezeichnungen des ersten Satzes von Penrose).

Weiter ist AA^-b die Orthogonalprojektion von b auf $\text{Im}(A)$, denn $AA^-b \in \text{Im}(A)$ und

$$\langle AA^-b - b, AA^-b \rangle = \langle AA^-b, AA^-b \rangle - \langle b, AA^-b \rangle,$$

AA^- ist selbstadjungiert, also

$$\begin{aligned} &= \langle b, AA^-AA^-b \rangle - \langle b, AA^-b \rangle \\ &= \langle b, AA^-b \rangle - \langle b, AA^-b \rangle = 0. \quad \square \end{aligned}$$

Beispiel:

Wir betrachten das unlösbare Gleichungssystem

$$\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Wir erhalten $\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{50} \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{50} \begin{pmatrix} -2 \\ -4 \end{pmatrix}$, die Probe ergibt $\begin{pmatrix} -\frac{1}{5} \\ \frac{3}{5} \end{pmatrix}$,

doch eine gehörige Abweichung.

Moore-Penrose-Inverse von normalen Matrizen lassen sich leicht berechnen. Zunächst rechnen Sie bitte nach, daß man die Moore-Penrose-Inverse einer Diagonalmatrix dadurch erhält, daß man die von Null verschiedenen Diagonalelemente durch ihre Inversen ersetzt.

Satz 3.5.2 *Sei A eine (quadratische) normale Matrix, U eine unitäre Matrix und D eine Diagonalmatrix, so daß $A = UDU^*$ ist. Dann ist $A^- = UD^-U^*$.*

Beweis: Man rechnet es eben aus:

$$AA^-A = UDU^*UD^-U^*UDU^* = UDD^-DU^* = UDU^* = A \text{ usw.} \quad \square$$

3.6 Überbestimmte Gleichungssysteme: Die Methode der kleinsten Quadrate (Gauß)

Der folgende Text ist angeregt von Leon, Linear Algebra with applications.

Sei $A \in M_{mn}$, $m > n$. Wir betrachten das Gleichungssystem $Ax = b$, wir setzen $r(x) = b - Ax$, das ist das „Residuum“ von x . Wir suchen ein x , wo $|b - Ax| = |r(x)|$ minimal ist, dann ist auch $|r(x)|^2$ minimal (daher der Name). Ein solcher Vektor existiert und ist eindeutig bestimmt.

Satz 3.6.1 *Sei $W \subset \mathbb{R}^m$ ein Unterraum. Für jedes $b \in \mathbb{R}^m$ existiert ein eindeutig bestimmter Vektor p mit $|b - y| > |b - p|$ für alle $y \in W, y \neq p$, und zwar gilt $b - p \in W^\perp$.*

Beweis: Wir zerlegen $\mathbb{R}^m = W \oplus W^\perp$, entsprechend sei $b = p + z$. Sei nun $y \in W$ beliebig, dann ist $|b - y|^2 = |(b - p) + (p - y)|^2$, wegen $p - y \in W$ und $p - b = z \in W^\perp$ gilt der Satz des Pythagoras, also

$$|b - y|^2 = |b - p|^2 + |p - y|^2,$$

also $|b - y|^2 > |b - p|^2$. □

Wenn speziell $b \in W$ gilt, so ist $p = b$, denn $b = b + o$ und $o \in W^\perp$.

Sei x_0 Lösung des kleinste-Quadrat-Problems $Ax = b$. Wenn $p = Ax_0 \in \text{Im}A$ minimalen Abstand zu b hat, so ist dies die Projektion von b auf $\text{Im}A$, d.h. $b - p = b - Ax_0 = r(x_0) \in (\text{Im}A)^\perp$.

Wie findet man so ein x_0 ?

Es ist $(\text{Im}A)^\perp = \text{Ker}(A^T)$, also $o = A^T r(x_0) = A^T(b - Ax_0)$, d.h. $A^T Ax_0 = A^T b$.

Satz 3.6.2 *Wenn $A \in M_{mn}$ den (vollen) Rang n hat, so ist die eindeutig bestimmte Lösung von $A^T Ax = A^T b$ der Vektor $x_0 = (A^T A)^{-1} A^T b$.*

Wir werden später diese Lösung mithilfe der QR-Zerlegung bestimmen.

Beispiel: Wir suchen die lineare Funktion $y = a + bx$ mit der „Wertetabelle“ $\begin{array}{c|c|c|c} x & 0 & 3 & 6 \\ \hline y & 1 & 4 & 5 \end{array}$.

Es soll also

$$\begin{pmatrix} 1 & 0 \\ 1 & 3 \\ 1 & 6 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix}$$

gelten. Es ist

$$A^T A = \begin{pmatrix} 3 & 9 \\ 9 & 45 \end{pmatrix}, \quad A^T \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 10 \\ 42 \end{pmatrix}, \quad \begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 4 \\ 2 \end{pmatrix},$$

also $y = \frac{4}{3} + \frac{2}{3}x$, die richtige Wertetabelle ist also $\begin{array}{c|c|c|c} x & 0 & 3 & 6 \\ \hline y & \frac{4}{3} & \frac{10}{3} & \frac{16}{3} \end{array}$.

3.7 Householder-Transformationen

Seien

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad e = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Spaltenvektoren aus M_{n1} , wir setzen

$$P = E - 2 \frac{vv^T}{v^T v},$$

Es ist $Px = x - 2 \frac{v^T x}{v^T v} v$, also $Pv = -v$ und wenn $\langle x, v \rangle = 0$ ist, so folgt $Px = x$. Demnach ist P die Spiegelung am Unterraum $\mathcal{L}(v)^\perp$. Derartige Abbildungen werden als Householder-Transformationen bezeichnet.

Lemma 3.7.1 *Sei x gegeben, dann kann v so gewählt werden, daß Px in $\mathcal{L}(e_1)$ liegt.*

Beweis: Wenn Px in $\mathcal{L}(e_1)$ liegen soll, so muß v in $\mathcal{L}(e_1, x)$ liegen. Wir setzen $v = x + re_1$, dann ist

$$Px = x - 2 \frac{x^T x + rx_1}{x^T x + 2rx_1 + r^2} x - 2r \frac{x^T x + rx_1}{v^T v} e_1,$$

der Koeffizient von x ist gleich Null, wenn $r = |x|$ ist, also ist $v = x \pm |x|e_1$ zu wählen. (Das ist auch anschaulich klar: v ist die Winkelhalbierende oder die Normale der Winkelhalbierenden zwischen x und e_1 .) \square

3.8 QR-Zerlegung

Sei eine Matrix A gegeben, gesucht wird eine orthogonale Matrix Q , so daß $Q^T A = R$ eine obere Dreiecksmatrix ist. Dann gilt also $A = QR$, man nennt dies die QR -Zerlegung von A , sie hat viele Anwendungen in der numerischen Mathematik.

Wir werden diese Zerlegung mit Hilfe von Householder-Transformationen herstellen. Die Spalten von A seien a_1, \dots, a_n . Es sei P_1 eine Householder-Matrix, so daß

$$P_1 a_1 = \begin{pmatrix} \star \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

ist, wir bilden

$$P_1 A = \begin{pmatrix} \star & \star & \dots & \star \\ 0 & \clubsuit & & \\ \vdots & & & \\ 0 & \clubsuit & \dots & \star \end{pmatrix}$$

(Beachten Sie, daß wir die Elemente in der zweiten Spalte hervorgehoben haben.)

Nun sei P'_2 die Householder-Matrix (mit einer Reihe weniger als P_1), für die

$$P'_2 \begin{pmatrix} \clubsuit \\ \vdots \\ \clubsuit \end{pmatrix} = \begin{pmatrix} \star \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

ist, wir setzen

$$P_2 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & P'_2 & \end{pmatrix}$$

Dann sind in $P_2 P_1 A$ schon die ersten beiden Spalten richtig eingerichtet. So fahren wir fort, nach n Schritten haben wir $Q = P_1 \dots P_n$ gefunden. \square

Ein anderes Verfahren der QR-Zerlegung:

Die Spaltenvektoren der Matrix A seien a_1, \dots, a_n . Wir orthonormieren diese Vektoren und erhalten q_1, \dots, q_n . Dann gilt

$$a_i = \langle a_i, q_1 \rangle q_1 + \dots + \langle a_i, q_n \rangle q_n.$$

Wir bilden aus den Spaltenvektoren q_i die orthogonale Matrix Q ; die obige Gleichung lautet dann in Matrixform

$$(a_1, \dots, a_n) = (q_1, \dots, q_n) \begin{pmatrix} \langle a_1, q_1 \rangle & & \langle a_n, q_1 \rangle \\ & \dots & \\ \langle a_1, q_n \rangle & & \langle a_n, q_n \rangle \end{pmatrix}.$$

Beim Orthonormierungsverfahren wird $q_i \in \mathcal{L}(a_1, \dots, a_i)$ und $q_i \perp \mathcal{L}(a_1, \dots, a_{i-1})$, also ist

$$R = \begin{pmatrix} \langle a_1, q_1 \rangle & & \langle a_n, q_1 \rangle \\ 0 & \dots & \\ 0 & 0 & \langle a_n, q_n \rangle \end{pmatrix}.$$

eine obere Dreiecksmatrix.

Aufgabe: Sei die Matrix A gegeben, sei

$$A = Q_0 R_0$$

ihre QR -Zerlegung. Wir setzen

$$A_1 = R_0 Q_0$$

und bilden wieder die QR -Zerlegung:

$$A_1 = Q_1 R_1.$$

Nun sei wieder

$$A_2 = R_1 Q_1.$$

So fahren wir fort. Die Folge der A_i konvergiert gegen eine Matrix B (das kann man zeigen), probieren Sie aus, welche Matrix das ist! Ein Computerprogramm dafür ist schnell geschrieben. In der Numerik-Vorlesung lernen Sie, daß dieses Verfahren hervorragend zur Eigenwertberechnung geeignet ist: Wenn A verschiedene Eigenwerte hat, so konvergiert dieses Verfahren gegen eine Diagonalmatrix, die die Eigenwerte enthält. Wir betrachten ein Beispiel:

```

2 2 2 2 2 2
2 4 7 12 21 38
2 7 18 43 106 279
2 12 43 128 381 1240
2 21 106 381 1250 4421
2 38 279 1240 4421 15552

```

Das charakteristische Polynom dieser Matrix ist gleich

$$x^6 - 16954x^5 + 676066x^4 + 29218948x^3 - 78775416x^2 - 26167472x + 5263104.$$

Die Potenzsummen sind

```

6
16954
2.86085984E8
4.838752092928E12
8.1842294913699632E16
1.3842745943986054E21
2.341351931977203E25

```

Wenn wir das charakteristische Polynom mit Hilfe der Newtonschen Formeln mit Gleitkommaarithmetik berechnen, erhalten wir

$$x^6 - 1,695E04x^5 + 6,761E05x^4 + 2,922E07x^3 - 7,878E07x^2 - 2,619E07x + 5,263E08,$$

die Fehler entstehen, weil die Potenzsummen sehr unterschiedliche Größenordnungen haben.

Nach 50 Iterationen erhalten wir

1,691E04	0.0	0.0	0.0	0.0	0.0
0.0	6,539E01	0.0	0.0	0.0	0.0
0.0	0.0	-2,786E01	0.0	0.0	0.0
0.0	0.0	0.0	2,820E00	0.0	0.0
0.0	0.0	0.0	0.0	-4,238E-01	0.0
0.0	0.0	0.0	0.0	0.0	1,429E-01

Das Produkt der Diagonalglieder ist 5263104.000000306, man kann also davon ausgehen, daß die Eigenwerte gut angenähert wurden.

Wie versprochen wollen wir die Gleichung $A^T Ax = A^T b$ lösen, wenn $A \in M_{mn}$ den Rang n hat: Sei $A = QR$, dann ist $Q \in M_{mn}$ und hat orthogonale Spalten und R ist eine invertierbare obere Dreiecksmatrix. Dann gilt $R^T Q^T QRx = R^T Q^T b$, Der Faktor R^T kann gekürzt werden und $Q^T Q = E$, also bleibt $Rx = Q^T b$.

Beispiel:

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 4 \\ -2 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \quad Q = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} \end{pmatrix} \quad R = \begin{pmatrix} 3 & -6 \\ 0 & 9 \end{pmatrix},$$

also

$$\begin{pmatrix} 3 & -6 \\ 0 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Q^T b = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad x = \begin{pmatrix} \frac{8}{9} \\ \frac{1}{9} \end{pmatrix}.$$

3.9 Hessenberg-Zerlegung

Sei A eine Matrix, dann gibt es eine orthogonale Matrix Q , so daß

$$Q^T A Q = \begin{pmatrix} * & & \dots & * \\ * & * & & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ & & \dots & \\ 0 & \dots & 0 & * & * \end{pmatrix}$$

eine „Hessenberg“-Matrix ist. (Dies ist eine Dreiecksmatrix, wo unter der Diagonalen noch eine Reihe besetzt ist.)

Die Existenz ist klar: Die reelle Schur-Zerlegung hat eine solche Gestalt. Wir werden sehen, daß man hier Householder-Transformationen nutzen kann.

Sei $A = (a_{ij})$, es sei P'_1 eine Householder-Matrix, die den Vektor $\begin{pmatrix} a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{pmatrix}$ in $\begin{pmatrix} * \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

überführt. Wir rändern die Matrix P'_1 mit Nullen in der ersten Zeile und Spalte und einer Eins links oben:

$$P_1 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ & & \dots & \\ 0 & & & P'_1 \end{pmatrix}$$

dann ist

$$P_1 A = \begin{pmatrix} * & & \dots & * \\ * & * & & * \\ 0 & * & * & * \\ & & \dots & \\ 0 & * & \dots & * & * \end{pmatrix}$$

wenn wir nun noch, wie gefordert, von rechts mit P_1^T heranzumultiplizieren, bleiben die Nullen in der ersten Spalte erhalten. Nun suchen wir eine $(n-2)$ -reihige Householdermatrix, die einen Teil der zweiten Spalte von $P_1 A P_1^T$ in Nullen überführt, rändern wieder zu einer n -reihigen Matrix usw. \square

Lemma 3.9.1 *Wenn die Matrix A symmetrisch und $H = Q^T A Q$ eine Hessenbergmatrix ist, so ist H tridiagonal, d.h. in H sind nur die Diagonale und die beiden Reihen über und unter der Diagonalen besetzt.*

Beweis: $H^T = (Q^T A Q)^T = Q^T A^T Q = Q^T A Q = H$. \square

Wir wissen zwar, daß eine symmetrische Matrix sogar diagonalisierbar ist, haben dafür z. B. Iterationsverfahren zur Verfügung. Die tridiagonale Form kann man jedoch in wenigen Schritten erzeugen.

Satz 3.9.1 *Sei*

$$T_r = \begin{pmatrix} a_1 & b_2 & & & \\ b_2 & a_2 & b_3 & & \\ & b_3 & \dots & & \\ & & & \dots & \\ & & & & b_r \\ & & & & b_r & a_r \end{pmatrix}$$

eine tridiagonale symmetrische r -reihige Matrix und $c_r(z)$ ihr charakteristisches Polynom, dann gilt

$$c_r(z) = (a_r - z)c_{r-1}(z) - b_r^2 c_{r-2}(z).$$

Beweis: Es ist

$$\begin{aligned} & \det \begin{pmatrix} a_1 - z & b_2 & & & \\ b_2 & a_2 - z & b_3 & & \\ & b_3 & \dots & & \\ & & & \dots & \\ & & & b_{r-1} & a_{r-1} - z & b_r \\ & & & & b_r & a_r - z \end{pmatrix} \\ &= (a_r - z)c_{r-1}(z) - b_r \det \begin{pmatrix} a_1 - z & b_2 & & & \\ b_2 & a_2 - z & b_3 & & \\ & b_3 & \dots & & \\ & & & \dots & \\ & & & & a_{r-2} - z & 0 \\ & & & & b_{r-1} & b_r \end{pmatrix} \end{aligned}$$

und durch Entwicklung des zweiten Summanden nach der letzten Spalte erhalten wir die Behauptung. \square

Weil wir ihn als ein Hilfsmittel für die anschließende Herleitung der Singulärwertzerlegung verwenden können, beweisen wir hier den verallgemeinerten Determinantenmultiplikationssatz:

Satz 3.9.2 *Seien $A \in M_{mn}$ und $B \in M_{nm}$ Matrizen und es sei $m \leq n$. Dann gilt $\det(AB) = \sum_K \det A_{IK} \det B_{KI}$, wobei $I = \{1, \dots, m\}$ ist und K alle Indexsysteme $\{k_1, \dots, k_m\}$ mit $1 \leq k_1 < k_2 < \dots < k_m \leq n$ durchläuft, A_{IK} ist die Untermatrix von A , die nur die Zeilen aus I und die Spalten aus K enthält.*

Beweis: Es gilt

$$AB = \begin{pmatrix} \sum_{i_1} a_{1i_1} b_{i_1 1} & \dots & \sum_{i_m} a_{1i_m} b_{i_m m} \\ \dots & \dots & \dots \\ \sum_{i_1} a_{mi_1} b_{i_1 1} & \dots & \sum_{i_m} a_{mi_m} b_{i_m m} \end{pmatrix}$$

Die Determinantenfunktion ist linear in jeder Spalte, also gilt

$$\det(AB) = \sum_{i_1} \dots \sum_{i_m} \det \begin{pmatrix} a_{1i_1} & \dots & a_{1i_m} \\ \dots & \dots & \dots \\ a_{mi_1} & \dots & a_{mi_m} \end{pmatrix} b_{i_1 1} \dots b_{i_m m}$$

Die Determinante auf der rechten Seite ist null, wenn einige Indizes übereinstimmen, wir betrachten also nur Summanden, wo alle i_l paarweise verschieden sind. Sei $\{i_1, \dots, i_m\} = \{k_1, \dots, k_m\}$, wo $k_1 < \dots < k_m$ ist. Sei p die Permutation mit $p(i_l) = k_l$. Dann ist die in Frage stehende Determinante gleich

$$\operatorname{sgn}(p) \det A_{IK}$$

und damit

$$\det(AB) = \sum_K \det A_{IK} \sum_p \operatorname{sgn}(p) b_{i_1 1} \dots b_{i_m m},$$

wegen $i_l = p^{-1}(k_l)$ ist die rechte Summe gleich $\det(B_{KI})$. \square

Folgerung 3.9.1 *Sei $C = AB$, sei C_{JK} eine Untermatrix von C , dann ist*

$$C_{JK} = A_{JI} B_{IK} \quad (I = \{1, \dots, n\})$$

und damit

$$\det C_{JK} = \sum \det A_{JL} \det B_{LK}. \square$$

Wir wenden dies für den Fall $B = A^T$ an und betrachten den Hauptminor $\det C_{JJ}$ von AA^T : $\det C_{JJ} = \sum \det A_{JK} \det A_{JK}$.

Folgerung 3.9.2 *Die charakteristischen Polynome von AA^T und $A^T A$ unterscheiden sich um den Faktor z^{n-m} , d.h. die von Null verschiedenen Eigenwerte von AA^T und $A^T A$ stimmen überein.*

Beweis: Der Koeffizient von z^{n-k} im charakteristischen Polynom von AA^T ist die Summe der k -Hauptminoren, also gleich

$$\sum_I \sum_J \det A_{IJ} \det A_{IJ},$$

bei $A^T A$ ist der Koeffizient von z^{m-k} gleich

$$\sum_I \sum_J \det A_{JI} \det A_{JI},$$

also stimmen sie überein. □

3.10 Singularwertzerlegung

Sei $f : V \rightarrow W$ eine lineare Abbildung, dabei sei $\dim V = n$, $\dim W = m$, $\dim \operatorname{Im}(f) = r$. Wir wissen, daß die von Null verschiedenen Eigenwerte von $f^* f$ und von $f f^*$ übereinstimmen und daß sie reell sind. Wir überlegen, daß die Eigenwert nicht negativ sein können: Sei $A^T A v = z v$, dann ist $v^T A^T A v = |Av|^2 = z v^T v = z |v|^2 \geq 0$. Wir können also die gemeinsamen Eigenwerte von $f f^*$ und $f^* f$ mit a_1^2, \dots, a_s^2 bezeichnen. Nun gibt es eine Orthonormalbasis $B = \{v_1, \dots, v_n\}$ von V aus Eigenvektoren von $f^* f$ und eine Orthonormalbasis $C = \{w_1, \dots, w_m\}$ aus Eigenvektoren von $f f^*$. Also ist

$$f^* f(v_i) = a_i^2 v_i,$$

also

$$f(f^* f(v_i)) = (f f^*)(f(v_i)) = a_i^2 f(v_i),$$

d.h. $f(v_i)$ ist ein Eigenvektor von $f f^*$ zum Eigenwert a_i^2 , also gilt

$$f(v_i) = r w_i.$$

Analog erhalten wir

$$f^*(w_i) = p v_i.$$

Nun gilt

$$\langle f(v_i), w_i \rangle = \langle r w_i, w_i \rangle = r = \langle v_i, f^*(w_i) \rangle = \langle v_i, p v_i \rangle = p.$$

Wir setzen oben $v_i = f^*\left(\frac{1}{r} w_i\right)$ ein:

$$f f^* f f^*\left(\frac{1}{r} w_i\right) = \frac{1}{r} a_i^4 w_i = a_i^2 f(v_i) = a_i^2 r w_i,$$

folglich ist $|r| = a_i$ und wir können (evtl. nach Ersetzung von w_i durch $-w_i$) $r = a_i$ annehmen, also

$$f(v_i) = a_i w_i.$$

Die Zahlen a_i heißen die Singulärwerte von f und die Basen B und C ein Paar singulärer Basen.

Folgerung 3.10.1 1. Es gibt Orthogonalbasen B, C von V bzw. W mit

$$A_{BC}(f) = \begin{pmatrix} a_1 & & 0 \\ & \dots & \\ 0 & & a_s \end{pmatrix}.$$

2. Zur Matrix A gibt es orthogonale Matrizen U, V , so daß

$$A = U^T \begin{pmatrix} a_1 & & 0 \\ & \dots & \\ 0 & & a_s \end{pmatrix} V. \square$$

Beweis: Sei $A \in M_{mn}$. Die Eigenwerte der $n \times n$ -Matrix $A^T A$ seien $z_1 = a_1^2, \dots, z_r = a_r^2, z_{r+1} = \dots, z_n = 0$. Die linear unabhängigen orthogonalen Eigenvektoren v_1, \dots, v_n bilden die Spalten der Matrix V .

Wir setzen $u_i = \frac{1}{a_i} A v_i$ für $i = 1, \dots, r$ und u_{r+1}, \dots, u_m sei eine ON-Basis von $\text{Ker}(A^T)$, diese Vektoren bilden die Spalten der Matrix $U \in M_{mm}$. \square

Beispiel: $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}$, $A^T A = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$, $z_1 = 4$, $a_1 = 2$, $z_2 = 0$ Die normierten Eigenvektoren ergeben $V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 0 \end{pmatrix}$ Dann ist $u_1 = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$.

Eine Basis von $\text{Ker}(A^T)$ ist $u_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, die Vektoren sind bereits orthogonal, werden normiert, das Ergebnis ist

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

Man kann die Singulärwertzerlegung nutzen, um die Moore-Penrose-Inverse A^- einer Matrix A zu bestimmen: Sei $A = USV^T$, dann ist

$$S^- = \begin{pmatrix} \frac{1}{a_1} & & & & \\ & \dots & & & \\ & & \frac{1}{a_r} & & \\ & & & 0 & \\ & & & & \dots \\ & & & & & 0 \end{pmatrix}$$

und $A^- = VS^-U^T$, denn $AA^-A = USV^T VS^-U^T USV^T = USV^T = A$ usw.

3.11 Vektor- und Matrixnormen

Wir wählen in diesem Abschnitt den Körper \mathbb{R} der reellen Zahlen als Grundkörper. Eine Zuordnung $V \rightarrow \mathbb{R}$, $x \mapsto \|x\|$ mit den Eigenschaften

1. $\|x\| \geq 0$, $\|x\| = 0$ genau für $x = 0$,
2. $\|rx\| = |r| \|x\|$ ($r \in \mathbb{R}$),
3. $\|x + y\| \leq \|x\| + \|y\|$

heißt eine Vektornorm.

Zum Beispiel ist der Betrag $\|x\|_2 = \sqrt{\langle x, x \rangle}$ eine Vektornorm, sie heißt die euklidische Norm. Weitere Beispiele sind die Summennorm $\|x\|_1 = \sum |x_i|$ und die Maximumnorm $\|x\|_\infty = \max |x_i|$.

Man rechnet die obigen Eigenschaften leicht nach. Veranschaulichen Sie sich durch eine Skizze, wie die jeweiligen „Einheitskreise“ aussehen.

Lemma 3.11.1 $\|x\|_\infty \leq \|x\|_1 \leq n \|x\|_\infty$,
 $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty$,
 $\|x\|_2 \leq \|x\|_1 \leq \sqrt{n} \|x\|_2$.

Definition: Sei $\|\cdot\|_p$ eine Vektornorm, $p \in \{1, 2, \infty\}$; wir setzen für eine Matrix A

$$\|A\|_p = \max \left\{ \frac{\|Ax\|_p}{\|x\|_p} \mid x \neq 0 \right\} = \max_{\|x\|_p=1} \|Ax\|_p,$$

dies nennen wir die durch $\|\cdot\|_p$ induzierte Matrixnorm.

Also gilt $\|Ax\|_p \leq \|A\|_p \cdot \|x\|_p$.

Weitere Matrixnormen sind die folgenden:

$$\|A\|_z = \max_k \sum_j |a_{kj}| \quad \text{Zeilensummennorm}$$

$$\|A\|_s = \max_j \sum_k |a_{kj}| \quad \text{Spaltensummennorm}$$

$$\|A\|_F = \sqrt{\sum a_{ij}^2} \quad \text{Frobeniusnorm}$$

Lemma 3.11.2 Die Spaltensummennorm wird durch die Maximumnorm, die Zeilensummennorm wird durch die Summennorm induziert.

Beweis: Zunächst gilt

$$\|A\|_\infty = \max_i \left| \sum_j a_{ij} x_j \right| \leq \max_i \sum_j |a_{ij}| \cdot |x_j| \leq \|x\|_\infty \max_i \sum_j |a_{ij}| = \|x\|_\infty \cdot \|A\|_z,$$

wir zeigen, daß die obere Schranke tatsächlich angenommen wird. Sei die k -te die Zeilensumme, wo das Maximum angenommen wird:

$$\max_i \sum_j |a_{ij}| = \sum_j |a_{kj}|,$$

dann setzen wir

$$x_i = \begin{cases} \frac{|a_{ki}|}{a_{ki}} & \text{für } a_{ki} \neq 0 \\ 0 & \text{sonst} \end{cases},$$

dann ist $\|x\|_\infty = 1$ und $\max |\sum a_{ij}x_j| = \sum_j |a_{kj}|$.

Die andere Aussage wird analog bewiesen:

$$\|A\|_1 = \max \|Ax\|_1 = \max \sum_i |a_{ij}x_j| = \max_j \sum_i |a_{ij}|.$$

□

Die verschiedenen Matrixnormen haben folgende Eigenschaften:

1. Für jede orthogonale Matrix Q gilt $\|x\|_2^2 = x^T x = x^T Q^T Q x = \|Qx\|_2^2$.
2. Wenn Q und R orthogonale Matrizen sind, so gilt $\|QAR\|_2 = \|A\|_2$, denn $\max \|QARx\|_2 = \max \|ARx\|_2 = \max \|Ax\|_2$.
3. $\|ABx\|_p = \|A(Bx)\|_p \leq \|A\|_p \|Bx\|_p \leq \|A\|_p \|B\|_p \|x\|_p$, also $\|AB\|_p \leq \|A\|_p \|B\|_p$.

4. Sei

$$D = \begin{pmatrix} d_1 & & \\ & \dots & \\ & & d_n \end{pmatrix}$$

eine Diagonalmatrix, dann ist $\|D\|_1 = \|D\|_2 = \|D\|_\infty = \max |d_i|$.

5. Sei

$$QAR = \begin{pmatrix} a_1 & & \\ & \dots & \\ & & a_n \end{pmatrix}$$

die Singulärwertzerlegung von A und $a_1 = \max a_i$, dann ist $\|A\|_2 = a_1$.

6. Sei z ein Eigenwert von A , dann $|z| \leq \|A\|_p$, denn für einen zugehörigen Eigenvektor v gilt $|z| \cdot \|v\| = \|zv\| = \|Av\| \leq \|A\| \cdot \|v\|$.
7. Sei A eine quadratische Matrix, für ihre Eigenwerte gelte $|z_1| \geq \dots \geq |z_n|$ und ihre Singulärwerte seien $a_1 \geq \dots \geq a_n$. Dann gilt $\prod a_i = |\det(A)|$, da orthogonale Matrizen die Determinante ± 1 haben, und $a_n \leq |z_i| \leq a_1$. Die rechte Ungleichung folgt aus dem oben Gesagten, die linke Ungleichung ist für $a_n = 0$ trivial, wenn aber $a_n \neq 0$ ist, so ist A regulär und A^{-1} hat die Eigenwerte $\frac{1}{z_i}$ und die Singulärwerte $\frac{1}{a_i}$, die linke Ungleichung folgt dann aus der rechten.
8. Wenn A eine symmetrische Matrix ist, so gilt $A^T Av_i = a_i^2 v_i = A^2 v_i = z_i^2 v_i$, also $a_i = |z_i|$.

Durch einige Rechnung kann man die folgende Hölder-Ungleichung beweisen:

$$\sum u_k v_k \leq \left(\sum u_k^p\right)^{\frac{1}{p}} \left(\sum u_k^q\right)^{\frac{1}{q}}, \quad \text{wobei } \frac{1}{p} + \frac{1}{q} = 1.$$

Hieraus folgt für $p > 1$ die Minkowski-Ungleichung

$$\left(\sum |x_k + y_k|^p\right)^{\frac{1}{p}} \leq \left(\sum |x_k|^p\right)^{\frac{1}{p}} + \left(\sum |y_k|^p\right)^{\frac{1}{p}},$$

dies ist für $p = 2$ die Dreiecksungleichung und allgemein bedeutet es, daß für $\|x\|_p = \left(\sum |x_k|^p\right)^{\frac{1}{p}}$ die oben geforderte dritte Normeigenschaft gilt (die beiden anderen sind ebenfalls erfüllt), damit haben wir eine ganze Serie von Vektornormen und durch sie induzierter Matrixnormen erhalten.

3.12 Positiv definite Matrizen

Wir beginnen mit ein paar Vorüberlegungen. Quadratische Gleichungen löst man durch quadratische Ergänzung:

Wegen

$$ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2 = 0$$

gilt

$$x = \left(-\frac{b}{a} \pm \sqrt{\frac{1}{a}\left(c - \frac{b^2}{a}\right)}\right)y.$$

Man kann dies in Matrixform schreiben:

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{b}{a} & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \begin{pmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{pmatrix},$$

wobei $a \neq 0$ und $d = c - \frac{b^2}{a}$ ist.

Dies kann in zwei verschiedenen Weisen auf symmetrische $n \times n$ -Matrizen verallgemeinert werden. Sei

$$S = \begin{pmatrix} T & w \\ w^T & a \end{pmatrix},$$

wo T eine $(n-1)$ -reihige Matrix und $a \neq 0$ ist. Dann gilt

$$S = \begin{pmatrix} E & \frac{1}{a}w \\ 0 & 1 \end{pmatrix} \begin{pmatrix} T - \frac{1}{a}ww^T & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} E & 0 \\ \frac{1}{a}w^T & 1 \end{pmatrix}.$$

Wenn T eine invertierbare Matrix ist, so gilt

$$S = \begin{pmatrix} E & 0 \\ w^T(T^{-1}) & 1 \end{pmatrix} \begin{pmatrix} T & 0 \\ 0 & a - w^T T^{-1} w \end{pmatrix} \begin{pmatrix} E & T^{-1}w \\ 0 & 1 \end{pmatrix}.$$

Rechnen Sie es nach!

Wir setzen $d = a - w^T T^{-1} w$, es ist $d = \frac{\det(S)}{\det(T)}$. Die jeweiligen rechten und linken Faktoren sind zueinander transponierte Dreiecksmatrizen, auf deren Diagonalen Einsen stehen, solche Matrizen heißen unipotent.

Definition: Die Determinante der Untermatrix einer Matrix A , die durch Streichen der letzten $n - i$ Zeilen und Spalten entstehen, nennen wir den i -Anfangsminor von A .

In der Literatur werden diese oft als „Hauptminoren“ bezeichnet, wir haben diesen Begriff aber schon vergeben.

Satz 3.12.1 (Jacobi) *Die Anfangsminoren d_1, \dots, d_n der symmetrischen Matrix S seien von Null verschieden, dann gibt es eine unipotente Matrix W mit*

$$S = W^T \begin{pmatrix} d_1 & & & & \\ & \frac{d_2}{d_1} & & & \\ & & \frac{d_3}{d_2} & & \\ & & & \dots & \\ & & & & \frac{d_n}{d_{n-1}} \end{pmatrix} W.$$

Beweis: Den Anfang haben wir soeben gemacht: Die oben eingeführte Zahl d ist gleich $\frac{d_n}{d_{n-1}}$ und da $\det(T) = d_{n-1} \neq 0$ ist, kann das ganze Verfahren auf die Untermatrix T angewandt werden, usw. \square

Definition: Eine symmetrische Matrix S heißt positiv definit, falls die Bilinearform $b_S(y, x) = x^T S y$ positiv definit ist.

Zum Beispiel sei $S = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, dann ist

$$a \cdot x^T S x = a^2 x_1^2 + 2abx_1x_2 + acx_2^2 = (ax_1 + bx_2)^2 + (ac - b^2)x_2^2 > 0$$

genau dann, wenn $a > 0$ und $ac - b^2 > 0$.

Wir bemerken, daß Diagonalmatrizen mit positiven Diagonalelementen positiv definit sind.

Wenn W invertierbar und S positiv definit sind, so ist $W^T S W$ auch positiv definit, denn $(Wx)^T S W x = x^T W^T S W x = 0$ genau dann, wenn $Wx = 0$, also wenn $x = 0$.

Schließlich ist mit

$$S = \begin{pmatrix} T & \star \\ \star & \star \end{pmatrix},$$

auch die $m \times m$ -Matrix T positiv definit, dies sehen wir, wenn wir b_T auf Vektoren anwenden, deren letzte $n - m$ Komponenten gleich Null sind.

Satz 3.12.2 *Sei S eine symmetrische Matrix, dann sind folgende Bedingungen äquivalent:*

1. S ist positiv definit,
2. es gibt eine invertierbare Matrix W mit $S = W^T W$,
3. alle Anfangsminoren von S sind positiv.

Beweis: (2 \Rightarrow 1) Es ist $S = W^T E W$ und E ist positiv definit.

(1 \Rightarrow 3) Sei T eine Anfangs-Teilmatrix von S :

$$S = \begin{pmatrix} T & \star \\ \star & \star \end{pmatrix},$$

dann ist T positiv definit, also ist b_T eine nichtausgeartete Bilinearform, also ist $\det(T) \neq 0$, alle Anfangsminoren sind von Null verschieden. Nach dem Satz von Jacobi gibt es eine unipotente (also invertierbare) Matrix U , so daß $S = U^T D U$ ist, wobei D eine Diagonalmatrix mit den Diagonaleinträgen $\frac{d_i}{d_{i-1}}$ ist. Mit S ist auch die Diagonalmatrix D positiv definit, also gilt $d_i > 0$.

(3 \Rightarrow 2) Wenn die Anfangsminoren positiv sind, so ist $S = U^T D U$ und es gibt reelle Zahlen a_i mit $\frac{d_i}{d_{i-1}} = a_i^2$. Wir setzen

$$A = \begin{pmatrix} a_1 & & \\ & \dots & \\ & & a_n \end{pmatrix},$$

dann ist $D = A^2$; wenn wir $W = A U$ setzen, erhalten wir $S = U^T D U = U^T A^2 U = W^T W$. \square

Kapitel 4

Euklidische und projektive Geometrie

4.1 Euklidische Geometrie

Sei V ein euklidischer Vektorraum mit dem Skalarprodukt $\langle \cdot, \cdot \rangle$. Sei $\{v_1, \dots, v_n\} \subset V$ eine Basis und $b_1, \dots, b_n \in \mathbb{R}$ gegeben. Wir suchen den Vektor x mit

$$\langle v_i, x \rangle = b_i, \quad i = 1, \dots, n,$$

wir setzen dazu $x = \sum y_j v_j$ ein:

$$\langle v_i, \sum y_j v_j \rangle = \sum_j \langle v_i, v_j \rangle y_j = b_i,$$

d.h. die erste Bedingung ist äquivalent zum Gleichungssystem für die Koordinaten y_i von x ; dessen Koeffizientenmatrix wird als Gram-Matrix bezeichnet:

$$G(v_1, \dots, v_n) = (\langle v_i, v_j \rangle).$$

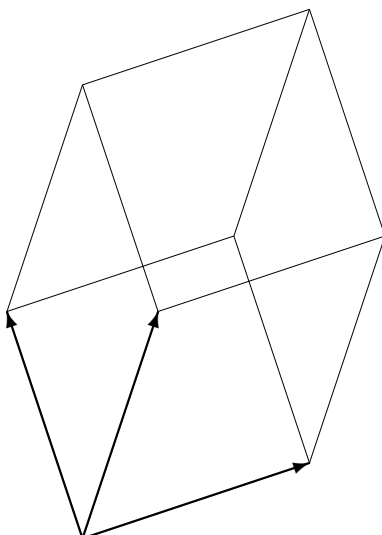
Es ist $\det(G(v_1, \dots, v_n)) \neq 0$. Solch eine Matrix kann auch für eine beliebige Anzahl auch linear abhängiger Vektoren eingeführt werden:

$$G(v_1, \dots, v_m) = (\langle v_i, v_j \rangle).$$

Lemma 4.1.1 *Wenn $\{v_1, \dots, v_m\} \subset V$ linear unabhängig ist, so ist $G(v_1, \dots, v_m)$ positiv definit.*

Seien $r_1, \dots, r_m \in \mathbb{R}$ nicht alle null, dann ist $x = \sum r_i v_i \neq 0$ und $0 < \langle x, x \rangle = \sum r_i r_j \langle v_i, v_j \rangle$, d.h. die Matrix $G(v_1, \dots, v_m)$ ist positiv definit. \square

Wir können die Gram-Matrix zur Volumenberechnung verwenden:



Wir stellen uns vor, daß die Vektoren a_1, \dots, a_m einen von Parallelogrammen begrenzten Körper aufspannen, so etwas nennt man ein Parallelepiped. Dann setzen wir

$$\text{vol}(a_1, \dots, a_m) = \sqrt{\det(G(a_1, \dots, a_m))}$$

(wir wissen, daß der Radikand nicht negativ ist).

Zur Rechtfertigung dieser Definition betrachten wir Spezialfälle:

$$m = 1: \text{vol}(a) = \sqrt{\det(\langle a, a \rangle)} = |a|.$$

$$m = 2: \text{vol}(a, b) = \sqrt{|a|^2 |b|^2 - \langle a, b \rangle^2} = |a| |b| \sin(\alpha).$$

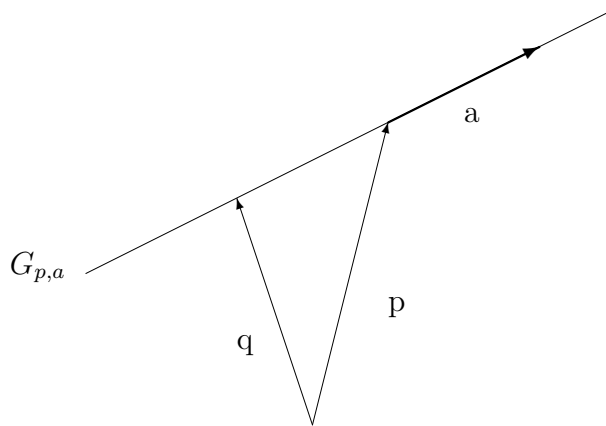
Wir überlegen, wie sich das Volumen ändert, wenn die Vektoren einer linearen Abbildung unterworfen werden.

Lemma 4.1.2 Seien $b_1, \dots, b_m \in \mathcal{L}(a_1, \dots, a_m)$ und $a_j = \sum a_{jk} b_k$, dann ist $\text{vol}(a_1, \dots, a_m) = |\det(A)| \cdot \text{vol}(b_1, \dots, b_m)$ mit $A = (a_{ij})$.

Beweis: Es ist $\langle a_i, a_j \rangle = \langle \sum a_{il} b_l, \sum a_{jk} b_k \rangle = \sum a_{il} a_{jk} \langle b_l, b_k \rangle$, also $G(a_1, \dots, a_m) = A \cdot G(b_1, \dots, b_m) \cdot A^T$. \square

Wir wollen nun Abstände und Winkel zwischen Geraden und Hyperebenen berechnen. Wir identifizieren dabei Punkte mit ihren „Ortsvektoren“, d.h. wir fassen den Vektorraum V als affinen Raum auf.

Eine Gerade ist durch einen ihrer Punkte p und einen Richtungsvektor a bestimmt:
 $G_{p,a} = \{q \mid q = p + ra\}$



Wenn $M \subset V$ eine Teilmenge und $q \in V$ ein Punkt ist, so definieren wir deren Abstand als

$$d(q, M) = \min\{|q - m| \mid m \in M\}.$$

Lemma 4.1.3 *Der Abstand des Punkts q von der Geraden $G_{p,a}$ ist*

$$d(q, G_{p,a}) = \frac{1}{|a|} \sqrt{|a|^2 |p - q|^2 - \langle p - q, a \rangle^2},$$

der Fußpunkt des Lotes von q auf $G_{p,a}$ ist

$$f = p - \frac{\langle p - q, a \rangle}{|a|^2} a.$$

Beweis: Für $r \in \mathbb{R}$ gilt

$$\begin{aligned} |(p + ra) - q|^2 &= \langle (p - q) + ra, (p - q) + ra \rangle \\ &= |p - q|^2 + 2r \langle p - q, a \rangle + r^2 |a|^2 \\ &= (r |a| + \frac{\langle p - q, a \rangle}{|a|})^2 + |p - q|^2 - \frac{\langle p - q, a \rangle^2}{|a|^2} \end{aligned}$$

und das Minimum wird für $r |a|^2 = -\langle p - q, a \rangle$ angenommen. \square

Sei $U \subset V$ ein Unterraum mit $\dim U = \dim V - 1$ und $p \in V$, dann heißt $H = p + U$ eine Hyperebene. Es gilt $v \in H$ genau dann, wenn die Koordinaten von v eine Gleichung erfüllen, etwa

$$H = H_{c,r} = \{v \in V \mid \langle c, v \rangle = r\},$$

dann ist $U = \mathcal{L}(c)^\perp$, denn es ist $u \in U$ gdw. $u = v_1 - v_2$ mit $v_i \in H$, also $\langle c, u \rangle = \langle c, v_1 \rangle - \langle c, v_2 \rangle = r - r = 0$.

Die Unterräume $p_1 + U_1$ und $p_2 + U_2$ sind parallel, wenn $U_1 \subset U_2$ oder $U_2 \subset U_1$ gilt, also sind zwei Geraden $G_{p,a}$ und $G_{q,b}$ genau dann parallel, wenn a und b linear abhängig sind, und zwei Hyperebenen $H_{c,r}$ und $H_{d,s}$ sind genau dann parallel, wenn c und d linear abhängig sind. Schließlich ist die Gerade $G_{p,a}$ genau dann parallel zur Hyperebene $H_{c,r}$, wenn $a \perp c$ ist.

Satz 4.1.1 Sei $G = G_{p,a}$ eine Gerade und $H = H_{c,r}$ eine Hyperebene. Dann sind folgende Bedingungen äquivalent:

1. G und H schneiden sich in einem Punkt.
2. G und H sind nicht parallel.

In diesem Fall ist der Schnittpunkt gleich $p + \frac{r - \langle p, c \rangle}{\langle a, c \rangle} a$.

Beweis: Wir suchen ein $x \in \mathbb{R}$ mit $\langle c, p + xa \rangle = r$, also $x \langle c, a \rangle = r - \langle c, p \rangle$, solch ein x existiert, wenn $\langle c, a \rangle \neq 0$ ist, und damit ergibt sich der obige Parameter für den Schnittpunkt. Wenn $\langle c, a \rangle = 0$ ist, so sind G und H parallel. \square

Folgerung 4.1.1 Das Lot vom Punkt p auf die Hyperebene $H_{c,r}$ ist die Gerade $G_{p,c}$, der Fußpunkt des Lots ist $f = p + \frac{r - \langle p, c \rangle}{|c|^2} c$ und $d(p, H) = \frac{|r - \langle p, c \rangle|}{|c|}$. \square

Satz 4.1.2 Sei $U \subset V$ ein Unterraum und $\{b_1, \dots, b_m\}$ eine Orthonormalbasis von U , dann ist für $x \in V$

$$p_U(x) = \sum_{i=1}^m \langle b_i, x \rangle b_i$$

die orthogonale Projektion von x auf U und

$$d(x, U) = \sqrt{|x|^2 - |p_U(x)|^2}.$$

Beweis: Man ergänzt die Basis von U zu einer Orthonormalbasis von V . \square

Damit können wir den Abstand beliebiger affiner Teilräume bestimmen. Sei $X_1 = p_1 + U_1$, $X_2 = p_2 + U_2$, dann ist

$$\begin{aligned} d(X_1, X_2) &= \min(|x_1 - x_2|) \\ &= \min(|p_1 + u_1 - p_2 - u_2|) \\ &= \min(|p_1 - p_2 - (u_2 - u_1)|) \\ &= d(p_1 - p_2, U_1 + U_2). \end{aligned}$$

Als nächstes betrachten wir Sphären: Sei M ein Punkt und r eine reelle Zahl,

$$\begin{aligned} S_{M,r} &= \{x \mid d(x, m) = r\} \\ &= \{x \mid |x - M| = r\} \\ &= \{x \mid \sum (x_i - m_i)^2 = r^2\} \end{aligned}$$

ist die Sphäre mit dem Mittelpunkt M und dem Radius r . Wir betrachten eine Gerade und eine spezielle Sphäre:

$$\begin{aligned} x &= p + ta, \quad |a| = 1, \\ |x|^2 &= r^2. \end{aligned}$$

Ein Punkt x ist ein Schnittpunkt, wenn

$$\langle p + ta, p + ta \rangle = r^2$$

d.h.

$$|p|^2 + 2\langle p, a \rangle t + t^2 - r^2 = 0,$$

also

$$(t + \langle p, a \rangle)^2 = \langle p, a \rangle^2 - (|p|^2 - r^2).$$

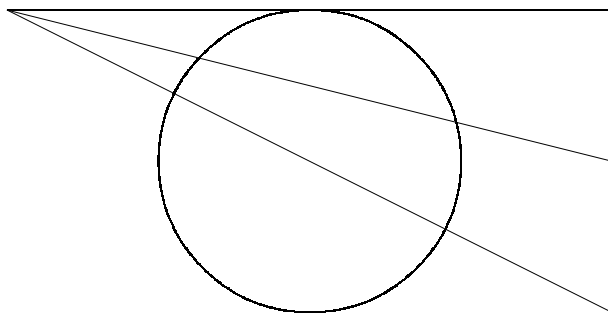
In Abhängigkeit vom Vorzeichen der rechten Seite gibt es zwei, einen oder keinen Schnittpunkt. Die Abstände von p zu den Schnittpunkten sind gleich den Lösungen der quadratischen Gleichung (1):

$$d(p, x_1) = |t_1 a| = |t_1|,$$

$$d(p, x_2) = |t_2 a| = |t_2|,$$

ihr Produkt ist gleich dem konstanten Term in (1), also

$$d(p, x_1)d(p, x_2) = |t_1 t_2| = |p|^2 - r^2.$$



Diese Konstante hängt nicht von der Richtung der Geraden ab, sondern nur vom Punkt p . Diese Aussage ist als Sekanten-Tangentensatz bekannt.

Für die folgenden Untersuchungen benötigen wir als Hilfsmittel das Vektorprodukt von Vektoren im \mathbb{R}^3 . Sei $\{i, j, k\}$ eine Orthonormalbasis des \mathbb{R}^3 , dies war für $a, b \in \mathbb{R}^3$ als

$$a \times b = \det \begin{pmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

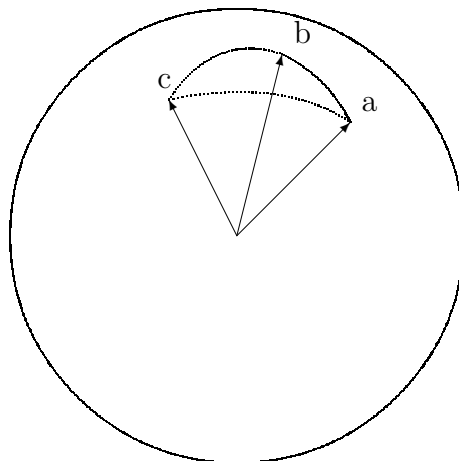
definiert. Wir werden nun einige Identitäten für Vektorprodukte herleiten.

Wir erinnern daran, daß das Produkt $a \times b$ linear in a und in b ist, daher genügt es, wenn die zu beweisenden Gleichungen nachgewiesen werden, wenn die einzelnen Faktoren Basiselemente sind. Somit erhalten wir

1. $\langle a \times b, c \rangle = \det(a \ b \ c)$
2. $a \times (b \times c) = \langle a, c \rangle b - \langle a, b \rangle c$ Graßmann-Identität
3. $a \times (b \times c) + b \times (c \times a) + c \times (a \times b) = 0$ Jacobi-Identität
4. $(a \times b) \times (c \times d) = \langle a \times b, d \rangle c - \langle a \times b, c \rangle d = \det(a \ b \ d)c - \det(a \ b \ c)d$
5. $\langle a \times b, c \times d \rangle = \langle a, c \rangle \langle b, d \rangle - \langle b, c \rangle \langle a, d \rangle$
6. $|v \times w| = |v| |w| \sin \alpha$

4.2 Sphärische Geometrie

Als nächstes wollen wir uns mit der Geometrie auf der Oberfläche der Einheitskugel beschäftigen.



Ein sphärisches Dreieck ist durch drei Einheitsvektoren a, b, c gegeben. Wir wählen die Reihenfolge so, daß $\det(a \ b \ c) > 0$ ist. Die „Seiten“ des Dreiecks sind die Winkel A, B, C zwischen den Vektoren a, b, c , die „Winkel“ α, β, γ des Dreiecks sind die Winkel zwischen den Ebenen $\mathcal{L}(a, b), \mathcal{L}(a, c), \mathcal{L}(b, c)$, also die Winkel zwischen deren Normalenvektoren. Da die Vektoren den Betrag 1 haben, gelten folgende Beziehungen:

$$\cos A = \langle b, c \rangle \quad \cos B = \langle a, c \rangle \quad \cos C = \langle b, a \rangle$$

Weiter ist

$$|b \times c|^2 = \langle b \times c, b \times c \rangle = \langle b, b \rangle \langle c, c \rangle - \langle c, b \rangle \langle b, c \rangle = 1 - \langle c, b \rangle^2 = 1 - \cos^2 A = \sin^2 A,$$

also

$$\begin{aligned} \sin A &= |a \times c| & \sin B &= |a \times b| & \sin C &= |b \times a| \\ \cos \alpha &= \frac{\langle a \times c, a \times b \rangle}{|a \times c| |a \times b|} & \cos \beta &= \frac{\langle b \times a, b \times c \rangle}{|b \times a| |b \times c|} & \cos \gamma &= \frac{\langle c \times b, c \times a \rangle}{|c \times b| |c \times a|} \end{aligned}$$

Aus den obigen Formeln für das Vektorprodukt folgt

$$\begin{aligned} |(a \times b) \times (a \times c)| &= |\det(a \ b \ c)| |a| \\ &= \det(a \ b \ c) \\ &= |a \times b| |a \times c| \sin \alpha \\ &= \sin C \cdot \sin B \cdot \sin \alpha. \end{aligned}$$

Daraus folgt der sphärische Sinussatz:

Satz 4.2.1

$$\frac{\sin \alpha}{\sin A} = \frac{\det(a \ b \ c)}{\sin A \sin B \sin C} = \frac{\sin \beta}{\sin B} = \frac{\sin \gamma}{\sin C}. \quad \square$$

Wenn die Seiten klein sind, so können wir sie die Sinuswerte durch die Argumente ersetzen und erhalten den ebenen Sinussatz.

Wir erhalten zwei Cosinussätze:

- Satz 4.2.2** 1. $\cos A = \cos B \cos C + \sin B \sin C \cos \alpha$,
 2. $\sin C \cos B = \sin B \cos C \cos \alpha + \sin A \cos \beta$.

Aus der ersten Formel geht hervor, daß man die Winkel aus den Seiten berechnen kann. Beweis: 1. $\sin B \sin C \cos \alpha = |a \times c| |a \times b| \cos \alpha = \langle a \times c, a \times b \rangle = \langle a, a \rangle \langle b, c \rangle - \langle a, b \rangle \langle a, c \rangle = \cos A - \cos B \cos C$.

2. Die Behauptung ist genau dann richtig, wenn

$$|a \times b| \langle a, c \rangle = |a \times c| \langle a, b \rangle \frac{\langle a \times c, a \times b \rangle}{|a \times c| |a \times b|} + |b \times c| \frac{\langle b \times a, b \times c \rangle}{|b \times a| |b \times c|}$$

gdw.

$$|a \times b|^2 \langle a, c \rangle = \langle a, b \rangle \langle a \times c, a \times b \rangle + \langle b \times a, b \times c \rangle,$$

die linke Seite ist gleich $(1 - \langle a, b \rangle^2) \langle a, c \rangle$, die rechte Seite ist gleich

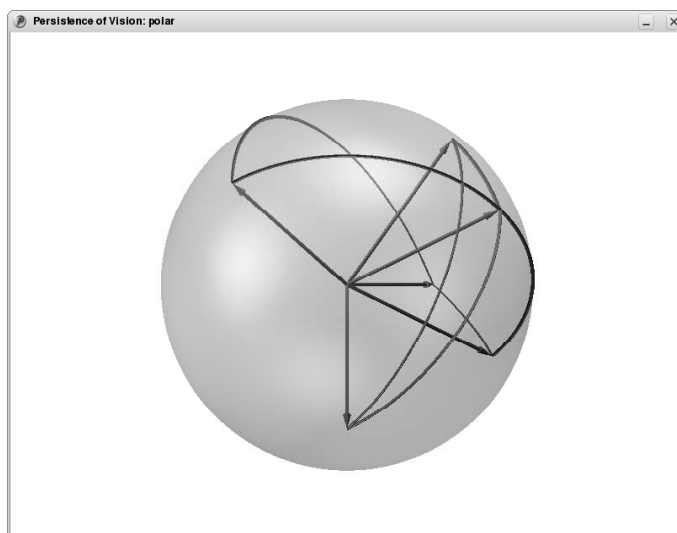
$$\langle a, b \rangle (\langle a, a \rangle \langle c, b \rangle - \langle c, a \rangle \langle a, b \rangle) + \langle b, b \rangle \langle a, c \rangle - \langle a, b \rangle \langle b, c \rangle,$$

dies stimmt mit der linken Seite überein. □

Das Dreieck, dessen definierenden Vektoren senkrecht auf den Seiten des durch a, b, c gegebenen Dreiecks stehen, wird das Polardreieck genannt, es hat die Ecken

$$a' = \frac{b \times c}{|b \times c|}, \quad b' = \frac{c \times a}{|c \times a|}, \quad c' = \frac{a \times b}{|a \times b|}$$

die Seiten A', B', C' und die Winkel α', β', γ' .



Satz 4.2.3 (Viètà-Formeln)

$$\cos A' = -\cos \alpha, \quad \cos \alpha' = -\cos A.$$

Beweis:

$$\cos A' = \langle b', c' \rangle = \frac{\langle c \times a, a \times b \rangle}{|c \times a| |a \times b|}. \square$$

Als Folgerung erhalten wir den polaren Cosinussatz:

Folgerung 4.2.1

$$-\cos \alpha = \cos \beta \cos \gamma + \sin \beta \sin \gamma \cos A. \square$$

Das bedeutet, daß die Seiten des Dreiecks aus den Winkeln berechnet werden können. Es gibt also keine ähnlichen Dreiecke, die nicht kongruent sind.

Wenn man die geografischen Längen L_i und Breiten B_i zweier Orte kennt, so kann man mit dem ersten Cosinussatz deren Entfernung berechnen. Man betrachtet das Dreieck, das von beiden Orten und dem Nordpol gebildet wird, dessen zwei Seiten sind gleich $\pi/2 - B_i$ und der der gesuchten Seite gegenüberliegende Winkel ist gleich $L_1 - L_2$.

Beispiel: Paris: $(2, 3^\circ; 48, 8^\circ)$, Berlin: $(13, 4^\circ; 52, 5^\circ)$, damit berechnet man $\cos A = 0,99\dots$, also $A = 7,939^\circ$, bei einem Erdradius von 6378 km ergibt dies eine Entfernung von 884 km.

4.3 Konvexe Mengen und lineare Ungleichungssysteme

Sei V ein Vektorraum und $U \subset V$ ein Unterraum, dann ist die Menge $v + U$ ein affiner Unterraum des affinen Raums (V, V) , wir nennen sie eine affine Teilmenge. Wenn $M \subset V$ eine beliebige Teilmenge ist, so sei I_M eine derartige Indexmenge, daß

$$\{A_i \mid i \in I_M\} = \{A \supseteq M \mid A \text{ affiner Unterraum}\}$$

die Menge aller affinen Unterräume von V ist, die M umfassen. Den Durchschnitt all dieser Mengen

$$A(M) = \bigcap_{i \in I_M} A_i$$

bezeichnen wir als die affine Hülle von M .

Lemma 4.3.1 $A(M) = m + \bigcap_{i \in I_M} U_i$ mit $m \in M$.

Beweis: Sei $m \in M$ beliebig, dann ist $m + \bigcap U_i \subset m + U_j$ für alle $j \in I_M$, also $m + \bigcap U_i \subset A(M)$.

Sei andererseits $x \in A(M)$, dann ist $x \in m + U_i$, also $x - m \in U_i$ für alle i , folglich ist $x - m \in \bigcap U_i$, also $x \in m + \bigcap U_i$ und damit $A(M) \subset m + \bigcap U_i$. \square

Wir bezeichnen den Unterraum $\bigcap U_i$ mit U_M .

Folgerung 4.3.1 Sei $M = \{m_1, \dots, m_k\}$. Es ist $U_M = \mathcal{L}(\{m_i - m_1\}) = \{\sum r_i m_i \mid \sum r_i = 0\}$.

Beweis: Der zu einem affinen Unterraum gehörige Vektorraum besteht aus den Verbindungsvektoren der Punkte. Jeder Vektor $u \in U_M$ hat die Gestalt $u = \sum r_i(m_i - m_1) = \sum r_i m_i - (\sum r_i)m$ und die Summe aller Koeffizienten ist Null. \square

Folgerung 4.3.2 $A(M) = \{\sum r_i m_i \mid \sum r_i = 1\}$. \square

Definition: Eine Teilmenge $K \subset V$ heißt konvex, wenn mit $u, v \in K$, $r, s \in \mathbb{R}$, $r, s \geq 0$, $r + s = 1$ auch $ru + sv \in K$ ist.

Eine konvexe Teilmenge enthält also mit je zwei Punkten auch deren Verbindungsstrecke.

Beispiele:

1. Jeder Unterraum ist konvex.

2. Die Menge $\{v \mid |v| \leq 1\}$ ist konvex, denn sei $|u| \leq 1, |v| \leq 1$, dann ist $|ru + sv| \leq |r||u| + |s||v| \leq r + s = 1$.

Definition: Seien $v_1, \dots, v_n \in V$, $r_1, \dots, r_n \in \mathbb{R}$, dann heißt ein Vektor $v = \sum r_i v_i$ mit $r_i \geq 0$, $\sum r_i = 1$ eine konvexe Linearkombination der v_i .

Lemma 4.3.2 Eine konvexe Menge enthält alle konvexen Linearkombinationen seiner Elemente.

Beweis: Sei K eine konvexe Menge. Nach Definition enthält K jede konvexe Linearkombination zweier seiner Elemente. Es sei schon gezeigt, daß K jede konvexe Linearkombination von $m - 1$ seiner Elemente enthält, und seien $v_1, \dots, v_m \in K$. Sei $r_i \geq 0, \sum r_i = 1$, wir betrachten den Vektor $v = \sum r_i v_i$. Falls $r_m = 1$ ist, so ist $v = v_m \in K$. Andernfalls ist $r_m < 1$, wir setzen $r = \sum_{i=1}^{m-1} r_i = 1 - r_m$, dann ist $v = r \sum \frac{r_i}{r} v_i + r_m v_m$, wobei $\sum \frac{r_i}{r} = \frac{\sum r_i}{r} = 1$ ist, der erste Summand liegt nach Voraussetzung in K , also ist v als konvexe Linearkombination zweier Elemente von K auch in K enthalten. \square

Definition: $K(M) = \bigcap \{K \mid M \subset K, K \text{ konvex}\}$ heißt die konvexe Hülle von K .

Satz 4.3.1 $K(M)$ ist die Menge aller (endlichen) konvexen Linearkombinationen von Elementen aus M .

Beweis: Die Menge

$$K_0 = \{v = \sum r_i v_i \mid v_i \in M, \sum r_i = 1\}$$

umfaßt M und ist konvex, denn seien $v = \sum r_i v_i$, $w = \sum s_i w_i \in M$, $r + s = 1$, dann ist $rv + sw = \sum rr_i v_i + \sum ss_i w_i$ und $\sum rr_i + \sum ss_i = r \sum r_i + s \sum s_i = r + s = 1$. \square

Definition: Eine Linearkombination $\sum r_i v_i$ heißt positiv, wenn $r_i \geq 0$ ist.

Wenn $M = \{v_1, \dots, v_m\}$ eine endliche Menge ist, so heißt $K(M)$ ein konvexes Polyeder.

Ein Element eines konvexen Polyeders $K(M)$ heißt Ecke, wenn es nicht als konvexe Linearkombination anderer Elemente von $K(M)$ dargestellt werden kann.

Satz 4.3.2 Die Ecken von $K(v_1, \dots, v_m)$ sind genau diejenigen v_i , die sich nicht als konvexe Linearkombination der restlichen v_j darstellen lassen.

Beweis: Es sei v_m keine konvexe Linearkombination von v_1, \dots, v_{m-1} , wir zeigen, daß v_m auch keine konvexe Linearkombination anderer Elemente ist.

Angenommen, es gilt $v_m = \sum_{i=1}^k a_i w_i$, $\sum a_i = 1$, $a_i > 0$, es sei $w_i = \sum_{j=1}^m a_{ij} v_j$, $\sum_j a_{ij} = 1$. Dann gilt $v_m = \sum_j \sum_i a_i a_{ij} v_j$, wir setzen $r_j = \sum_i a_i a_{ij}$, es gilt $\sum r_j = 1$. Falls $r_m = \sum a_i a_{im} = 1$ gilt, so muß $a_{im} = 1$ für ein i und $a_{ij} = 0$ für $j < m$ gelten, also wäre $v_m = w_i$ als konvexe Linearkombination dargestellt.

Also gilt $r_m < 1$ und wir haben v_m als Linearkombination von v_1, \dots, v_{m-1} dargestellt:

$$(1 - r_m)v_m = \sum_{i=1}^{m-1} r_i v_i.$$

Dies ist eine konvexe Linearkombination, wie man leicht nachrechnet. Damit erhalten wir einen Widerspruch. \square

Definition: Sei $M = \{m_0, \dots, m_k\}$; wenn $\dim A(M) = s$ ist, so heißt $K(M)$ ein s -dimensionales Polyeder. Wenn $\dim A(M) = k$ ist, so heißt es ein k -Simplex. Sei $S = K(x_0, \dots, x_k)$ ein k -Simplex, dann ist $K(x_{i_0}, \dots, x_{i_r})$ ein r -Simplex, es heißt Seite von S .

Satz 4.3.3 Der Durchschnitt eines Simplex und eines Untervektorraums ist ein konvexes Polyeder oder leer.

Beweis: Sei $P = K(x_0, \dots, x_m)$ ein Simplex und $U \subset V$ ein Unterraum. Jeder Punkt $x \in P \cap U$ liegt auf gewissen Seiten von P , sei S_x die Seite kleinster Dimension, die x enthält. Wir nennen einen Punkt $x \in P \cap U$ markant, wenn $S_x \cap U = \{x\}$ ist, also wenn $S_x \not\subset U$ ist. Die Zahl der markanten Punkte ist endlich, da die Zahl der Seiten endlich ist. Wir zeigen

$$P \cap U = K(\{x \mid x \text{ markant}\}).$$

Wenn $x \in P \cap U$ nicht markant ist, so enthält S_x noch einen Punkt z aus $P \cap U$. Dann gilt $x, z \in U$, also ist $y = z - x \in U$ nicht der Nullvektor. Es sei $S_x = K(x_0, \dots, x_r)$ und $x = \sum a_k x_k$ mit $\sum a_k = 1$, dann ist $a_k > 0$ für alle k wegen der Minimalität von S_x . Sei $z = \sum b_k x_k$, $b_k \geq 0$, $\sum b_k = 1$. Dann ist $y = \sum (b_k - a_k) x_k$, wir setzen $c_k = b_k - a_k$, dann gilt $\sum c_k = 0$, aber nicht alle c_k sind null. Also ist mindestens eins der c_k positiv und mindestens eins ist negativ. Sei

$$a = \max_{c_k < 0} \left(-\frac{a_k}{c_k} \right),$$

diese Zahl ist positiv. Dann gilt

$$a_k + a c_k \geq 0, \quad k = 0, \dots, r$$

und das Gleichheitszeichen kommt für ein k vor. Analog gibt es ein $b > 0$ mit

$$a_k - b c_k \geq 0, \quad k = 0, \dots, r$$

und auch hier kommt das Gleichheitszeichen vor. Wir betrachten

$$u = x + ay = \sum (a_k + ac_k)x_k,$$

hier sind die Koeffizienten nicht negativ und die Koeffizientensumme ist gleich 1. Analog sei

$$v = x - by = \sum (a_k - bc_k)x_k,$$

auch dies ist eine konvexe Linearkombination der x_k . Nun liegen aber u und v auf einer echten Seite von S_x , denn jeweils ein Koeffizient ist null. Es gilt

$$bu + av = bx + bay + ax - aby = (a + b)x,$$

folglich ist

$$x = \frac{1}{a+b}(bu + av)$$

eine konvexe Linearkombination von Elementen, die auf niedrigerdimensionalen Seiten als x liegen. Wenn u, v markant sind, so sind wir fertig. Wenn nicht, so zerlegen wir, wie soeben x , nun u und v . \square

Definition: Eine Menge der Form $P = P(v_0, \dots, v_m) = \{\sum r_i v_i \mid r_i \geq 0\}$ heißt konvexe Pyramide. Zwei Vektoren u, v heißen positiv parallel, wenn ein $r > 0$ existiert, so daß $ru = v$ ist. Eine Pyramide P heißt spitz, wenn $P \cap (-P) = \{0\}$ ist.

OBdA seien unter den v_i keine positiv parallelen Vektoren; dann heißen diese die Kanten von P . Keine Kante ist eine positive Linearkombination der restlichen.

Seien $H_1, \dots, H_k \subset \mathbb{R}^n$ Hyperebenen, die jeweils durch eine homogene Gleichung $f_i(x) = 0$ beschrieben seien. Dann ist $H = \bigcap H_i$ ein Unterraum, wir betrachten $H^+ = \{x \in H \mid x_i \geq 0\}$. Die Elemente von H^+ können also durch Gleichungen und Ungleichungen beschreiben werden.

Satz 4.3.4 H^+ ist eine spitze konvexe Pyramide.

Beweis: Sei $\{e_1, \dots, e_n\}$ die kanonische Basis, sie erzeugen das $(n-1)$ -Simplex $P = K(e_1, \dots, e_n)$ und der Durchschnitt $P \cap H$ ist ein konvexes Polyeder, also von der Form $K(x_0, \dots, x_m)$. Nun ist aber H^+ die positive Hülle der x_i , H^+ ist spitz, da in H^+ nur Vektoren mit nichtnegativen Komponenten, in $-H^+$ aber nur Vektoren mit nichtpositiven Komponenten vorkommen. \square

Wir betrachten jetzt ein lineares Ungleichungssystem, dies ist eine Folge von Bedingungen der Form

$$f(x) \geq a, \quad g(x) \leq b, \quad h(x) = c,$$

wo f, g, h Linearformen sind.

Durch Einföhrung neuer Unbekannter und neuer Bedingungen können wir das Bild vereinheitlichen:

$f(x) \geq a$ ist äquivalent zu $f(x) - y = a, \quad y \geq 0,$

$g(x) \leq b$ ist äquivalent zu $g(x) + z = b, \quad z \geq 0,$

wir können die Ungleichungen also durch Gleichungen und Positivitätsforderungen an die Unbekannten ersetzen. Wenn an eine Unbekannte z keine Positivitätsforderung gestellt ist, so ergänzen wir $z = z' - z''$, $z' \geq 0$, $z'' \geq 0$.

Somit können wir eine einfache Form des Ungleichungssystems annehmen:

$$Ax = b, \quad x_i \geq 0, \quad (i = 1, \dots, n).$$

Satz 4.3.5 *Die Menge der Lösungen des homogenen Ungleichungssystems $Ax = 0$, $x_i \geq 0$ ist eine spitze konvexe Pyramide.*

Beweis: Die Lösungsmenge hat die Form H^+ (s.o.). □

Zum inhomogenen Ungleichungssystem

$$Ax = b, \quad x_i \geq 0$$

betrachten wir das Ungleichungssystem

$$AZ - bz_0 = 0, \quad z_i \geq 0,$$

hierbei sei $Z = (z_1, \dots, z_n)^T$, wir betrachten die Lösungen von (2) mit $z_0 > 0$, aus diesen erhalten wir Lösungen von (1): $x_i = \frac{z_i}{z_0}$.

Satz 4.3.6 *Die Lösungsmenge eines inhomogenen linearen Ungleichungssystems ist die Summe eines konvexen Polyeders und einer konvexen Pyramide.*

Beweis: Die Menge aller Lösungen $Z = (z_1, \dots, z_n, z_0)$ von (2) bilden eine konvexe Pyramide, diese habe die Kanten

$$Z_k = (z_{1k}, \dots, z_{nk}, z_{0k}), \quad k = 1, \dots, s,$$

also hat jede Lösung die Form $Z = \sum r_i Z_i$, $r_i \geq 0$.

Falls für $k = 1, \dots, s$ stets $z_{0k} = 0$ ist, so ist $z_0 = 0$ für jede Lösung von (2), d.h. (1) besitzt keine Lösung.

Sei also oBdA $z_{01} > 0, \dots, z_{0r} > 0$, $z_{0,r+1} = \dots = 0$, es ist $z_j = \sum r_i z_{ji}$ und damit

$$(x_j) = \left(\frac{z_j}{z_0} \right) = \sum_{k=1}^r r_k \frac{z_{0k}}{z_0} \left(\frac{z_{jk}}{z_{0k}} \right) + \sum_{k=r+1}^s \frac{r_k}{z_0} (z_{jk}),$$

die zweite Summe beschreibt ein Element einer Pyramide; wir betrachten die Koeffizientensumme der ersten Summe:

$$\sum_{j=1}^r r_k \frac{z_{0k}}{z_0} = \frac{1}{z_0} \sum r_k z_{0k} = \frac{z_0}{z_0} = 1,$$

also beschreibt die erste Summe ein Element eines konvexen Polyeders. □

4.4 Projektive Geometrie

Definition: Sei V ein Vektorraum, $P(V)$ die Menge aller 1-dimensionalen Unterräume von V , $G(V)$ sei die Menge aller 2-dimensionalen Unterräume von V und ϵ sei die Inklusionsrelation auf $P(V) \times G(V)$, also $p \in g \iff p \subset g$. Dann heißt $(P(V), G(V), \epsilon)$ der projektive Raum über V . Die Elemente von $P(V)$ heißen Punkte, die Elemente von $G(V)$ heißen Geraden. Der Punkt p liegt auf der Geraden g , wenn $p \in g$, d.h. $p \subset g$ gilt. Sei $U \subset V$ ein Unterraum, dann gilt $P(U) \subset P(V)$ und $G(U) \subset G(V)$ und $(P(U), G(U), \epsilon|_{P(U) \times G(U)})$ heißt projektiver Unterraum. Wir setzen $\dim(P(V)) = \dim V - 1$.

Lemma 4.4.1 $P(U)$ und $P(W)$ seien Unterräume von $P(V)$. Wenn $P(U) = P(W)$ ist, dann gilt $U = W$.

Beweis: Sei $P(U) \subset P(W)$ und $o \neq u \in U$, dann hat $p = \mathbb{R}u$ die Dimension 1, also gilt $p \in P(U)$, also $p \in P(W)$ und damit $U \subset W$. \square

Satz 4.4.1 $P(V)$ sei ein projektiver Raum, dann gilt:

- (1) Zu zwei Punkten p, q gibt es genau eine Gerade $g = (p, q)$ mit $p \in g$, $q \in g$.
- (2) Auf jeder Geraden liegen mindestens drei Punkte.
- (3) Seien p, q, r, s verschiedene Punkte. Wenn die Geraden (p, q) und (r, s) einen Schnittpunkt haben, so schneiden sich auch (p, r) und (q, s) .

Beweis: 1. Seien $p, q \in P(V)$, $p \neq q$, dann ist der Unterraum $p + q \subset V$ 2-dimensional und es gilt $p \subset p + q$, $q \subset p + q$, also ist $p + q \in G(V)$ und $p \in p + q$, $q \in p + q$ und $(p, q) = p + q$ ist der einzige zweidimensionale Unterraum von V mit dieser Eigenschaft.

2. Sei $g = \mathbb{R}x + \mathbb{R}y$ eine Gerade, dann sind x, y linear unabhängig, also sind $\mathbb{R}x$, $\mathbb{R}y$, $\mathbb{R}(x + y)$ drei verschiedene Punkte auf g .

3. Sei $p = \mathbb{R}u$, $q = \mathbb{R}v$, $r = \mathbb{R}x$, $s = \mathbb{R}y$, nach Voraussetzung sind jeweils zwei der Vektoren u, v, x, y linear unabhängig. Es ist $(p, q) = \mathbb{R}u + \mathbb{R}v$, $(r, s) = \mathbb{R}x + \mathbb{R}y$. Wenn (p, q) und (r, s) sich schneiden, so ist $(\mathbb{R}u + \mathbb{R}v) \cap (\mathbb{R}x + \mathbb{R}y)$ 1-dimensional, enthält also einen vom Nullvektor verschiedenen Vektor

$$z = au + bv = cx + dy, \quad a, b, c, d \in \mathbb{R}.$$

Dann ist

$$z' = au - cx = -bv + dy \in (p, r) \cap (q, s),$$

denn falls $z' = o$ wäre, so wäre wegen $a = b = c = d = 0$ auch $z = o$. \square

Alle im Folgenden zu beweisenden Aussagen können aus den drei Aussagen dieses Satzes hergeleitet werden, diese Aussagen könnten also als Axiome der projektiven Geometrie dienen.

Wir beweisen einige elementare geometrische Eigenschaften.

Satz 4.4.2 (4) Zwei verschiedene Geraden g, h schneiden sich höchstens in einem Punkt.

(5) Sei $\dim(P(V)) = 2$ (d.h. $P(V)$ ist eine projektive Ebene), dann schneiden sich zwei verschiedene Geraden aus $P(V)$ genau in einem Punkt.

Beweis: 4. Seien $p, q \in P(V)$ mit $p \in h$, $q \in h$, $p \in g$, $q \in g$, $p \neq q$, also $h = (p, q)$, $g = (p, q)$, wegen (1) folgt $g = h$.

5. Seien $g, h \in G(V)$. Wenn $g \cap h = \{o\}$, dann ist

$$\dim g + \dim h = 4 > \dim V = 3,$$

aus dem Widerspruch folgt $\dim(g \cap h) = 1$ und der Schnittpunkt ist wegen (4) eindeutig bestimmt. \square

Satz 4.4.3 Seien p, q, r drei Punkte, dann ist $p \in (q, r)$ genau dann, wenn $q \in (p, r)$.

Beweis: Sei $p \in (q, r)$, also $\mathbb{R}p \subset \mathbb{R}q + \mathbb{R}r$ dann ist auch $p \in (p, r)$, beide Geraden enthalten p und r , also ist $(q, r) = (p, r)$ und damit $q \in (p, r)$. \square

Lemma 4.4.2 $P(U)$ sei ein Unterraum von $P(V)$, dann gilt: (6) Wenn $p, q \in P(U)$, $p \neq q$, dann ist $(p, q) \in G(U)$.

(7) Sei $g \in G(U)$, $p \in P(V)$, wenn $p \in g$, so ist $p \in P(U)$.

Beweis: 6. Wir haben $p, q \subset U$, also $p + q \subset U$, also ist $(p, q) = p + q \in G(U)$.

7. Wegen $p \in g$ ist $p \subset g \subset U$, also $p \subset U$. \square

Wir können Unterräume eines projektiven Raums wie folgt charakterisieren:

Satz 4.4.4 Seien $P \subset P(V)$, $G \subset G(V)$ Teilmengen mit folgenden Eigenschaften:

(a) wenn $p, q \in P$ und $p \neq q$ ist, so ist $(p, q) \in G$,

(b) sei $g \in G$ und $p \in P(V)$, wenn $p \in g$ ist, so ist $p \in P$.

Dann gibt es einen Unterraum $U \subset V$ mit $P = P(U)$, $G = G(U)$.

Beweis: Wir setzen $U = \{u \in V \mid u = o \text{ oder } \mathbb{R}u \in P\}$ und zeigen, daß dies ein Unterraum ist.

Seien $u, v \in U$, also $\mathbb{R}u, \mathbb{R}v \in P$. Wenn u, v linear abhängig sind, $\mathbb{R}(u + v) = \mathbb{R}u \in P$, also $u + v \in U$. Seien nun u, v linear unabhängig, dann ist $\mathbb{R}u + \mathbb{R}v \in G$ nach (a). Weiter ist $\mathbb{R}(u + v) \in \mathbb{R}u + \mathbb{R}v$, wegen (b) ist also $\mathbb{R}(u + v) \in P$, d.h. $u + v \in U$.

Für $r \in \mathbb{R}$ ist $ru \in U$, da $\mathbb{R}ru = \mathbb{R}u$.

Nach Konstruktion ist $P(U) = P$ und aus (a) und (7) gilt $G(U) \subset G$; aus (6) und (b) folgt $G \subset G(U)$. \square

Definition: Seien $X, Y \in P(V)$ Teilmengen, wir setzen

$$X + Y = \{p \in P(V) \mid \text{es gibt } x \in X, y \in Y \text{ mit } p \in (x, y)\} \cup X \cup Y,$$

$$2X = X + X, \dots, nX = (n - 1)X + X,$$

$H(X) = \bigcup_n nX$ heißt die lineare Hülle von X .

Lemma 4.4.3 $(X + Y) + Z = X + (Y + Z)$ \square

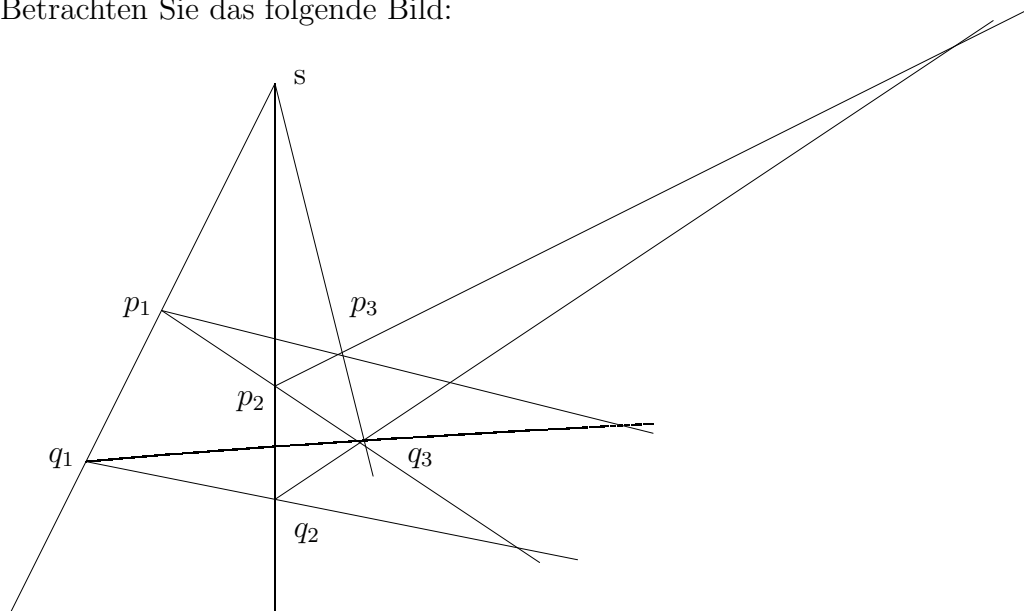
Satz 4.4.5 Seien $p_0, \dots, p_n \in P(V)$, $p_i = \mathbb{R}x_i$, $\mathcal{L}(x_0, \dots, x_n) = U$, dann gilt

$$H(p_0, \dots, p_n) = P(U), \dim(H(p_0, \dots, p_n)) \leq n. \quad \square$$

Folgerung 4.4.1 $\dim P + \dim Q = \dim(P + Q) + \dim(P \cap Q)$. \square

Sei nun $H \subset P(V)$ eine Hyperebene und g eine Gerade, dann ist $\dim(g \cap H) = \dim H + \dim g - \dim(H + g) \geq n - 1 + 1 - n = 0$, also schneiden sie sich in jedem Fall. In einem projektiven Raum gibt es keine Parallelität.

Betrachten Sie das folgende Bild:



Satz 4.4.6 (Desargues(1591-1661)) Seien $p_1, p_2, p_3, q_1, q_2, q_3, s$ paarweise verschiedene Punkte und $s \in (p_i, q_i)$, $i = 1, 2, 3$. Dann gibt es kollineare (d.h. auf einer Geraden gelegene Punkte) b_{12}, b_{23}, b_{13} mit $b_{ij} \in (p_i, p_j)$ und $b_{ij} \in (q_i, q_j)$.

Beweis: Sei $p_i = \mathbb{R}x_i$, $q_i = \mathbb{R}y_i$, $s = \mathbb{R}z$, dann sind jeweils x_i, y_i linear unabhängig. Es ist aber $z \in \mathcal{L}(x_i, y_i)$, oBdA können wir $z = x_i - y_i$, ($i = 1, 2, 3$) annehmen, also

$$x_i - x_j = y_j - y_i.$$

Die gesuchten Schnittpunkte sind dann

$$b_{ij} = \mathbb{R}(x_i - x_j) \in \mathbb{R}x_i + \mathbb{R}x_j = (p_i, p_j)$$

und

$$\mathbb{R}(y_i - y_j) \in \mathbb{R}y_i + \mathbb{R}y_j = (q_i, q_j).$$

Weiter ist $x_1 - x_3 = (x_1 - x_2) + (x_2 - x_3)$, also

$$\mathbb{R}(x_1 - x_3) \subset \mathbb{R}(x_1 - x_2) + \mathbb{R}(x_2 - x_3),$$

also $b_{13} \in (b_{12}, b_{23})$. \square

Wir wollen nun einen Zusammenhang mit den von früher bekannten affinen Räumen herstellen.

Satz 4.4.7 Sei $V = W \oplus \mathbb{R}a$ und $A = P(V) \setminus P(W)$, dann gibt es zu jedem Punkt $p \in A$ genau einen Vektor $f(p) \in W$ mit $p = \mathbb{R}(f(p) + a)$.

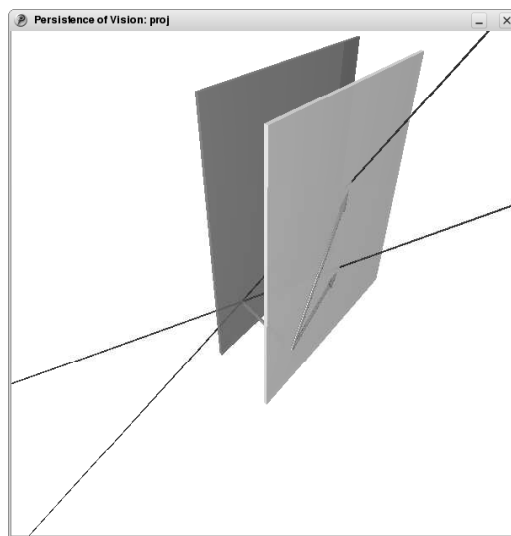
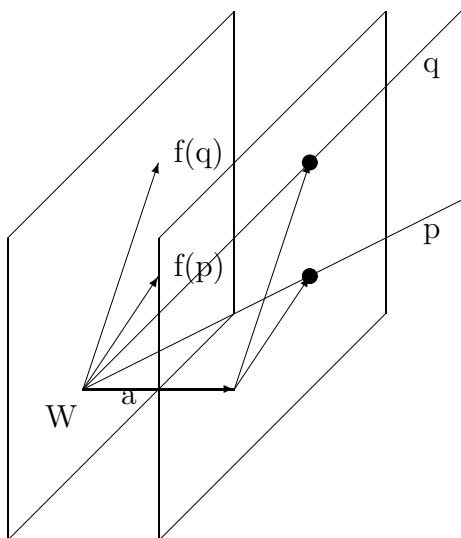
Beweis: Sei $p = \mathbb{R}x \in A$, dann ist $\mathbb{R}x \not\subset W$, also gibt es ein eindeutig bestimmtes $w \in W$ und $t \in \mathbb{R}$ mit $x = w + ta$, wir setzen $f(p) = \frac{1}{t}w$. Dann gilt $\frac{1}{t}x = f(p) + a$.

Wenn $y = sx$, ($s \neq 0$) ein anderer Repräsentant von p ist, so gilt $y = sw + sta$, also ist $f(p) = \frac{1}{st}sw = \frac{1}{t}w$ wohldefiniert. \square

Satz 4.4.8 Sei $P(W)$ eine projektive Hyperebene in $P(V)$. Dann ist $A = P(V) \setminus P(W)$ ein affiner Raum mit dem Translationsraum W .

Wir nennen $P(W)$ die bezüglich A uneigentliche Hyperebene.

Beweis: Wegen $\dim W = \dim V - 1$ ist $V = W \oplus \mathbb{R}a$ für ein $a \in V \setminus W$, hierzu haben wir die obige Abbildung $f: A \rightarrow W$. Sei $p = \mathbb{R}x$, $w \in W$, wir setzen $q = \mathbb{R}(w + tx)$ mit $tx = f(p) + a$, dann ist $w + tx = (w + f(p)) + a = f(q) + a$ (nach Konstruktion von f), also setzen wir $\overrightarrow{pq} = w = f(q) - f(p)$. Man rechnet leicht nach, daß $\overrightarrow{pq} = \overrightarrow{pr} + \overrightarrow{rq}$ gilt. \square



Folgerung 4.4.2 Sei $g \in G(V)$, $g \not\subset P(W)$, dann ist $g \setminus P(W) = g \cap A$ eine affine Gerade und $g \cap P(W)$ besteht aus genau einem Punkt.

Beweis: Es ist $\dim W = \dim V - 1$, $\dim g = 2$, also $\dim(g \cap W) = 1$. Sei $g = \mathbb{R}a + \mathbb{R}w$ mit $a \notin W$, wir können $w \in W$ wählen. Sei $p_0 = \mathbb{R}a \in g \cap A$, dann ist ein beliebiger Punkt $p \in g \cap A$ von der Form $p + \mathbb{R}(tw + a)$ (oBdA kann man den Faktor von a gleich 1 wählen), d.h. $f(p) = tw$ ist der Verbindungsvektor von p_0 zu p , also gilt (im affinen Sinn) $p = p_0 + tw$, diese Punkte bilden eine affine Gerade. Weiter ist $g \cap P(W) \neq \emptyset$, und wenn im Durchschnitt zwei Punkte lägen, so läge g in $P(W)$. \square

Satz 4.4.9 Wenn $h \subset A$ eine affine Gerade ist, dann gibt es genau eine projektive Gerade \bar{h} mit $h = \bar{h} \cap A$.

Beweis: Sei $h = \mathbb{R}w + p_0 \subset A$, $p_0 = \mathbb{R}a \in A$, dann ist $\bar{h} = \mathbb{R}a + \mathbb{R}w$ eine projektive Gerade. Jeder Punkt $p = \mathbb{R}(tw + a)$ von h liegt in $\bar{h} \cap A$, also ist $h \subset \bar{h} \cap A$. Wenn umgekehrt $p \in \bar{h} \cap A$ ist, so ist $p = \mathbb{R}(tw + a) \in h$. \square

Folgerung 4.4.3 Die Zuordnung $h \rightarrow \bar{h}$ ist eine Bijektion zwischen den Geraden von A und den projektiven Geraden, die nicht in $P(W)$ enthalten sind.

Folgerung 4.4.4 Die affinen Geraden h_1, h_2 sind genau dann parallel, wenn $\bar{h}_1 \cap P(W) = \bar{h}_2 \cap P(W)$.

Beweis: Seien (affin) $h_i = p_i + \mathbb{R}w_i$ mit $p_i = \mathbb{R}a_i$, dann sind die zugehörigen projektiven Geraden gerade $\bar{h}_i = \mathbb{R}w_i + \mathbb{R}a_i$, deren Schnittpunkt mit $P(W)$ sind die Punkte $\mathbb{R}w_i$. Affin gilt aber $h_1 \parallel h_2$ genau dann, wenn $\mathbb{R}w_1 = \mathbb{R}w_2$. \square

Der affine Raum (A, W) wird also zu einem projektiven Raum „vervollständigt“, indem zu jeder Parallelschar in A ein „uneigentlicher“ Punkt hinzugefügt wird. Alle Punkte bilden eine „uneigentliche Hyperebene“. Jede Gerade wird zu einer projektiven Geraden verlängert.

Wir wollen nun die Punkte eines projektiven Raums durch Koordinaten beschreiben.

Definition: Sei $\dim P(V) = n$. Die Punkte p_0, \dots, p_n bilden eine projektive Basis, wenn sie nicht in einer Hyperebene enthalten sind.

Satz 4.4.10 Die Punkte $p_0 = \mathbb{R}x_0, \dots, p_n = \mathbb{R}x_n$ bilden genau dann eine projektive Basis von $P(V)$, wenn $\{x_0, \dots, x_n\}$ eine Basis von V bilden.

Beweis: $\{x_0, \dots, x_n\}$ sind genau dann linear abhängig, wenn sie in einem n -dimensionalen Unterraum W von V enthalten sind, dann liegen die entsprechenden Punkte aber in der Hyperebene $P(W)$. \square

Definition: Die Punkte p_0, \dots, p_{n+1} bilden ein projektives Koordinatensystem, wenn je $n + 1$ dieser Punkte eine projektive Basis bilden.

Sei $\{p_0, \dots, p_{n+1}\}$ ein projektives Koordinatensystem und $p_i = \mathbb{R}x_i$. Dann sind die Vektoren x_0, \dots, x_{n+1} linear abhängig, aber je $n + 1$ von ihnen sind linear unabhängig. Es gibt also eine Linearkombination

$$t_0x_0 + \dots + t_{n+1}x_{n+1} = 0$$

wo alle Koeffizienten t_i ungleich Null sind. Wir können also ohne Beschränkung der Allgemeinheit annehmen, daß

$$x_0 + \dots + x_{n+1} = 0$$

ist. Die Punkte p_0, \dots, p_n heißen dann Grundpunkte, p_{n+1} heißt Einheitspunkt.

Wenn nun $p \in P(V)$, $p = \mathbb{R}x$ irgendein Punkt ist, so ist $x = a_0x_0 + \dots + a_nx_n$ und die Zahlen a_0, \dots, a_n heißen die homogenen Koordinaten von p . Wegen des folgenden Satzes wird das Koordinatentupel mit $(a_0 : \dots : a_n)$ bezeichnet.

Satz 4.4.11 *Die homogenen Koordinaten sind bis auf ein Vielfaches mit einem Faktor $\neq 0$ eindeutig bestimmt. Zu jedem Tupel $(a_0 : \dots : a_n) \neq (0 : \dots : 0)$ gibt es einen Punkt mit diesen homogenen Koordinaten.*

Beweis: Seien $p_i = \mathbb{R}x_i = \mathbb{R}x'_i$ mit $x_0 + \dots + x_{n+1} = 0 = x'_0 + \dots + x'_{n+1}$, also $x'_i = t_i x_i$. Dann ist $\{x_0, \dots, x_n\}$ eine Basis von V und es gilt

$$-t_{n+1}x_{n+1} = t_0x_0 + \dots + t_nx_n,$$

aber auch

$$-t_{n+1}x_{n+1} = t_{n+1}x_0 + \dots + t_{n+1}x_n,$$

also $t_0 = \dots = t_{n+1} = t$. Wenn nun $p = \mathbb{R}x$ ist, so haben wir $x = a_0x'_0 + \dots + a_nx'_n = ta_0x_0 + \dots + ta_nx_n$, also unterscheiden sich die homogenen Koordinaten bei verschiedenen Repräsentanten der Elemente des projektiven Koordinatensystems nur um die Konstante t .

Zum Tupel $(a_0 : \dots : a_n)$ haben wir $p = \mathbb{R}x$ mit $x = a_0x_0 + \dots + a_nx_n$. \square

Folgerung 4.4.5 *Die homogenen Koordinaten der Grundpunkte sind $(0 : \dots : 1 : \dots : 0)$ und die des Einheitspunkts sind $(1 : \dots : 1)$.* \square

Folgerung 4.4.6 *Sei $P(W)$ die Hyperebene durch p_1, \dots, p_n und p habe die homogenen Koordinaten $(a_0 : \dots : a_n)$, dann ist p ein eigentlicher Punkt, wenn $a_0 \neq 0$ ist, sonst ein uneigentlicher Punkt. Die affinen Koordinaten eines eigentlichen Punkts sind $(1 : \frac{a_1}{a_0} : \dots : \frac{a_n}{a_0})$*

Beweis: Es ist $x = a_0x_0 + \dots + a_nx_n \in W$ genau dann, wenn $a_0 = 0$ gilt. \square

Folgerung 4.4.7 *Bezüglich des affinen Koordinatensystems $\{a = \mathbb{R}x_0, x_1, \dots, x_n\}$ möge p die Koordinaten (p_1, \dots, p_n) haben, dann sind seine homogenen Koordinaten $(1 : p_1 : \dots : p_n)$.* \square

Sei eine affine Gerade durch eine Gleichung

$$ax_1 + bx_2 = c$$

gegeben, wir wollen ihre uneigentlichen Punkte finden. Wir homogenisieren die Gleichung, indem wir x_i durch $\frac{x_i}{x_0}$ ersetzen und die Nenner „hochmultiplizieren“:

$$ax_1 + bx_2 = cx_0. \quad \star$$

Ihre eigentlichen Punkte haben die homogenen Koordinaten $(1 : x_1 : x_2)$, die (\star) erfüllen. Es gibt genau einen uneigentlichen Punkt auf der Geraden, er hat die homogenen Koordinaten $(0 : -b : a)$, man sieht die Verwandtschaft zum Richtungsvektor der Geraden.

Analog besitzt eine Ebene genau eine uneigentliche Gerade: wenn die Ebene durch die Gleichung $ax_1 + bx_2 + cx_3 = d$ gegeben ist, so ist $\mathbb{R} \begin{pmatrix} -b \\ a \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} -c \\ 0 \\ a \end{pmatrix}$ diese Gerade.

Ein Kreis ist durch eine Gleichung

$$(x_1 - a)^2 + (x_2 - b)^2 = r^2$$

gegeben, wir homogenisieren:

$$x_1^2 - 2ax_0 + x_0^2 + x_2^2 - 2bx_0 + x_0^2 = r^2x_0^2.$$

Um uneigentliche Punkte zu finden, setzen wir $x_0 = 0$:

$$x_1^2 + x_2^2 = 0,$$

die Lösungen sind $(0 : x_1 : x_2) = (0 : 1 : x_2)$ mit $x_2^2 = -1$, also $(0 : 1 : i)$ und $(0 : 1 : -i)$. In diesen beiden unendlich fernen imaginären Punkten schneiden sich also alle Kreise, denn die Parameter a, b, r sind herausgefallen. Diese Punkte werden die „Kreispunkte“ genannt.

Überprüfen Sie, daß sich alle zueinander ähnlichen Ellipsen ebenfalls jeweils in zwei imaginären uneigentlichen Punkten schneiden.

Betrachten wir nun die durch

$$ax_1^2 + x_2 = 0$$

gegebene Parabel. Ihre homogene Gleichung lautet

$$ax_1^2 + x_0x_2 = 0$$

und ihr uneigentlicher Punkt ist $(0 : 0 : 1)$.

Betrachten wir schließlich eine Hyperbel, ihre homogene Gleichung lautet

$$ax_1^2 - bx_2^2 = cx_0^2,$$

sie hat die uneigentlichen Punkte $(0 : 1 : \pm\sqrt{\frac{a}{b}})$, sie entsprechen den Richtungen der Asymptoten.

Kapitel 5

Polynommatrizen

Definition: $M_n(R[x])$ sei die Menge aller $n \times n$ -Matrizen $A(x) = (a_{ij}(x))$, wo die $a_{ij}(x)$ Polynome sind. Solche Matrizen heißen Polynommatrizen.

Sei $A(x)$ eine Polynommatrix, k sei das Maximum der Grade der $a_{ij}(x)$, dann heißt k der Grad von $A(x)$, $k = \deg(A(x))$. Dann können wir jedes $a_{ij}(x)$ als

$$a_{ij}(x) = a_{ij}^{(0)}x^k + a_{ij}^{(1)}x^{k-1} + \dots + a_{ij}^{(k)}$$

schreiben und mindestens ein $a_{ij}^{(0)}$ ist von Null verschieden.

Wir betrachten nun die Matrizen

$$A_l = (a_{ij}^{(l)}) \in M_{nn},$$

dann ist

$$A(x) = A_0x^k + A_1x^{k-1} + \dots + A_k$$

und A_0 ist nicht die Nullmatrix.

Zum Beispiel:

$$\begin{pmatrix} x^2 + x + 1 & x^3 - x + 2 \\ 2x & x - 3x - 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x^3 + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} x^2 + \begin{pmatrix} 1 & -1 \\ 2 & -3 \end{pmatrix} x + \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$$

Definition: Die Polynommatrix $a(x)$ heißt regulär, wenn A_0 regulär ist.

Polynommatrizen werden wie üblich addiert und multipliziert.

Lemma 5.0.4 $\deg(A(x) + B(x)) \leq \max(\deg(A(x)), \deg(B(x)))$,
 $\deg(A(x)B(x)) \leq \deg(A(x)) + \deg(B(x))$, wenn $A(x)$ oder $B(x)$ regulär ist, so gilt im zweiten Fall Gleichheit.

Beweis: Sei $A(x) = A_0x^n + \dots$, $B(x) = B_0x^n + \dots$, wo $A_0 \neq 0$ oder $B_0 \neq 0$ ist, dann ist

$$A(x) + B(x) = (A_0 + B_0)x^n + \dots,$$

also ist der Grad höchstens gleich n . Weiter sei $A_0 \neq 0$ und $B_0 \neq 0$, dann ist

$$A(x)B(x) = A_0B_0x^{n+m} + \dots,$$

also ist der Grad höchstens gleich $n + m$. Wenn z.B. A_0 regulär ist, so ist $A_0B_0 \neq 0$, der Grad also gleich $n + m$. \square

Satz 5.0.12 (Division mit Rest) Seien $A(x)$, $B(x)$ Polynommatrizen, $B(x)$ sei regulär. Dann gibt es eindeutig bestimmte Polynommatrizen $Q(x)$, $R(x)$ mit $A(x) = Q(x)B(x) + R(x)$, wobei $R(x) = 0$ oder $\deg R(x) < \deg B(x)$ gilt. $Q(x)$ heißt der rechte Quotient, $R(x)$ der rechte Rest von $A(x)$ bzgl. $B(x)$.

Beweis: Sei $\deg(A(x)) = l$, $\deg(B(x)) = m$. Falls $l < m$ ist, setzen wir $Q(x) = 0$, $R(x) = A(x)$.

Sei also $l \geq m$, wir führen die Induktion über l . Es gilt

$$B_0^{-1}B(x)x^{l-m} = Ex^l + B_0^{-1}B_1x^{l-1} + \dots + B_0^{-1}B_mx^{l-m},$$

die Matrix

$$A_0B_0^{-1}B(x)x^{l-m} = A_0x^l + \dots$$

hat denselben höchsten Koeffizienten wie $A(x)$, also hat

$$A(x) - A_0B_0^{-1}B(x)x^{l-m}$$

höchstens den Grad $l - 1$. Nach Induktionsvoraussetzung gibt es also Matrizen $P(x)$ und $R(x)$, wo $\deg(R) < \deg(B)$ oder $R = 0$ gilt, so daß

$$A(x) - A_0B_0^{-1}B(x)x^{l-m} = P(x)B(x) + R(x),$$

d.h.

$$A(x) = (P(x) + A_0B_0^{-1}x^{l-m})B(x) + R(x).$$

Den Faktor vor $B(x)$ nennen wir $Q(x)$.

Die Matrizen Q und R sind eindeutig bestimmt, denn sonst gäbe es P und S mit

$$A = QB + R = PB + S$$

also

$$(Q - P)B = P - R,$$

da $Q - P \neq 0$ sein sollte, steht links eine Matrix vom Grad $\geq m$, rechts aber eine Matrix vom Grad $< m$, also ist $P = Q$ und $R = S$. \square

Folgerung 5.0.8 Dasselbe gilt mit vertauschten Faktoren: $A = BP + S$, $S = 0$ oder $\deg(S) < \deg(B)$, P heißt linker Quotient und S linker Rest. \square

Es ist sicher nicht verwunderlich, das bei linker und rechter Division unterschiedliche Quotienten und Reste auftreten, es kann sogar vorkommen, daß eine Matrix A bei rechter Division durch B einen von Null verschiedenen Rest läßt, aber von links durch B teilbar ist. Hier ist ein Beispiel:

$$\underbrace{\left(\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} x + \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \right)}_B \cdot \underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}}_P = \underbrace{\left(\begin{pmatrix} 5 & 6 \\ 11 & 16 \end{pmatrix} x + \begin{pmatrix} 3 & 2 \\ 11 & 16 \end{pmatrix} \right)}_A$$

Wir bestimmen nun Q, R mit $A = QB + R$:

$$B_0^{-1}B = Ex + \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} = Ex + \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix},$$

dann ist

$$A_0B_0^{-1}B = A_0x + \begin{pmatrix} 5 & 6 \\ 11 & 16 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix} = A_0x + \begin{pmatrix} -7 & -1 \\ -21 & -5 \end{pmatrix}$$

und

$$A - A_0B_0^{-1}B = \begin{pmatrix} -4 & 1 \\ -4 & -3 \end{pmatrix} = R, \quad Q = A_0B_0^{-1} = \begin{pmatrix} -15 & 7 \\ -47 & 21 \end{pmatrix}$$

Früher haben wir in ein gegebenes Polynom eine (skalare) Matrix eingesetzt. Nun wollen wir in eine Polynommatrix $A(x) \in M_n(R[x])$ eine Matrix $B \in M_{nn}$ einsetzen, dabei müssen wir aber aufpassen: Sei

$$A(x) = A_0x^k + A_1x^{k-1} + \dots + A_k,$$

dann definieren wir

$$A(B) = A_0B^k + A_1B^{k-1} + \dots + A_k.$$

Satz 5.0.13 *Es sei $B \in M_{nn}$ und $A(x) = Q(x)(Ex - B) + R$, dann hat R den Grad 0, ist also eine skalare Matrix und es gilt $A(B) = R$.*

Beweis: Wie Sie selbst bitte überprüfen, gilt

$$Ex^i - B^i = (Ex^{i-1} + Bx^{i-2} + \dots + B^{i-2}x + B^{i-1})(Ex - B),$$

wir multiplizieren von links mit A_{k-i} und summieren:

$$\begin{aligned} & A_0Ex^k - A_0B^k + A_1Ex^{k-1} - A_1B^{k-1} + \dots + A_k - A_k \\ &= A(x) - A(B) = \sum A_{k-i}(Ex^{i-1} + \dots + B^{i-1})(Ex - B), \end{aligned}$$

den Faktor vor $(Ex - B)$ bezeichnen wir mit $Q(x)$ und erhalten

$$A(x) = Q(x)(Ex - B) + A(B),$$

also $A(B) = R$. □

5.1 Smithsche Normalform

Die folgenden Darlegungen stammen aus dem Buch *Theory of matrices* von P. Lancaster, das mir in der russischen Raubübersetzung von 1978 vorlag.

Wir wollen Polynommatrizen Operationen folgenden Typs unterwerfen:

1. Vertauschen von Zeilen bzw. Spalten,

2. Multiplikation einer Reihe mit einer Zahl $r \neq 0$,
3. Addition des $f(x)$ -fachen einer Zeile zu einer anderen, dasselbe auch für Spalten, dabei sei $f(x)$ ein Polynom.

Definition: Zwei Polynommatrizen heißen äquivalent, wenn sie durch eine Folge von elementaren Operationen auseinander hervorgehen.

Zum Beispiel gehen die folgenden Matrizen durch elementare Operationen auseinander hervor:

$$\begin{pmatrix} x & x+1 \\ x^2-x & x^2-1 \end{pmatrix} \begin{pmatrix} x & x+1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Satz 5.1.1 Jede Polynommatrix ist zu einer Polynommatrix der Form

$$\begin{pmatrix} i_1 & & & 0 \\ & \dots & & \\ & & i_r & \\ 0 & & & 0 \end{pmatrix}$$

äquivalent, wobei jeweils i_k ein Teiler von $i_{k+1}(x)$ ist.

Beweis: Durch Zeilen- und Spaltenvertauschungen wird erreicht, daß $\deg(a_{11}(x))$ minimal ist. Das Polynom a_{1k} aus der ersten Zeile wird mit Rest durch a_{11} dividiert:

$$a_{1k} = qa_{11} + r, \deg(r) < \deg(a_{11}) \text{ oder } r = 0.$$

Nun subtrahieren wir das q -fache der ersten Spalte von der k -ten Spalte, dann bleibt an der Stelle $(1, k)$ das r stehen. Wenn $r = 0$ ist, ist es gut, sonst bringen wir es an die Stelle $(1,1)$ und beginnen von vorn. Nach endlich vielen Schritten sind alle Elemente der ersten Zeile (außer dem ersten) gleich Null. Dasselbe veranstalten wir mit der ersten Spalte. Also ist $A(x)$ äquivalent zur Matrix

$$\begin{pmatrix} a_{11}(x) & 0 & \dots & 0 \\ \dots & & & \\ 0 & & A_1(x) & \end{pmatrix}$$

Wenn a_{11} alle Komponenten von $A_1(x)$ teilt, so bleibt das auch bei allen Operationen, die wir künftig mit $A_1(x)$ ausführen, erhalten. Wenn etwa $a_{ij}(x)$ nicht von a_{11} geteilt wird, so addieren wir die i -te Zeile zur ersten und beginnen von vorn. Dabei wird sich der Grad von a_{11} weiter verkleinern. Wenn wir erreicht haben, daß a_{11} alle Komponenten von $A_1(x)$ teilt, widmen wir uns $A_1(x)$ und bringen es in Diagonalgestalt. Irgendwann sind wir fertig. \square

Wir fragen uns nun, ob die Polynome i_1, i_2, \dots von den gewählten elementaren Operationen oder nur von der Matrix $A(x)$ abhängen. Die Antwort können wir aber erst etwas später geben. Zuerst überlegen wir uns, daß die Wirkung dieser Operationen durch Multiplikation mit Matrizen folgender Art realisiert werden kann:

$$\begin{pmatrix} 1 & & 0 \\ & \dots & \\ & r & \\ & \dots & \\ & & 1 \end{pmatrix}, \begin{pmatrix} 1 & & 0 \\ & \dots & \\ & f(x) & \\ & \dots & \\ & & 1 \end{pmatrix}, \begin{pmatrix} 1 & & 0 \\ & \dots & 1 \\ & 1 & \dots \\ & & & 1 \end{pmatrix}.$$

Dies sind Polynommatrizen, deren Determinante nicht von x abhängt, die also im Bereich der Polynommatrizen eine Inverse besitzen.

Definition: Sei $A(x) = (a_{ij}(x))$ eine Polynommatrix.

Sei $d_1(x)$ der größte gemeinsame Teiler aller $a_{ij}(x)$,

$d_2(x)$ der größte gemeinsame Teiler aller 2-Minoren von $A(x)$,

...

$d_i(x)$ der größte gemeinsame Teiler aller i -Minoren von $A(x)$,

...

$d_n(x) = \det A(x)$. Alle d_i seien normiert. Die d_i heißen die Determinantenteiler von $A(x)$.

Lemma 5.1.1 Für alle i gilt: $d_i(x)$ teilt $d_{i+1}(x)$.

Beweis: Nach dem Entwicklungssatz ist jeder $(i+1)$ -Minor von $A(x)$ eine Linearkombination von i -Minoren, also teilt d_i jeden $(i+1)$ -Minor und damit auch d_{i+1} . \square

Definition: Wir setzen $i_0(x) = 1$, $i_k(x) = \frac{d_k(x)}{d_{k-1}(x)}$, die $i_k(x)$ heißen die Invariantenteiler von $A(x)$.

Satz 5.1.2 Die Determinantenteiler einer Matrix ändern sich bei elementaren Operationen nicht. Äquivalente Matrizen haben dieselben Determinantenteiler.

Beweis: Wir betrachten die äquivalenten Matrizen $A(x)$ und $P(x)A(x)Q(x)$, wo $P(x)$ und $Q(x)$ Produkte von Elementarmatrizen sind, ihre Inversen sind also auch Polynommatrizen. Sei $b_j(x)$ ein l -Minor von $P(x)A(x)Q(x)$, nach dem verallgemeinerten Determinantenmultiplikationssatz gilt

$$b_j = \sum p_i a_i q_i,$$

wo die Polynome p_i, a_i, q_i jeweils gewisse l -Minoren von $P(x), A(x)$ bzw. $Q(x)$ sind. Nun sei d_l der l -te Determinantenteiler von $A(x)$. Dann teilt d_l jedes a_i , also teilt es auch jeden l -Minor von PAQ und damit auch den l -ten Determinantenteiler von PAQ . Da durch Multiplikation von PAQ mit P^{-1} und Q^{-1} wieder A erhalten wird, stimmen die Determinantenteiler überein. \square

Satz 5.1.3 Sei $A(x)$ zu $\begin{pmatrix} a_1(x) & & \\ & \ddots & \\ & & a_n(x) \end{pmatrix}$ äquivalent, weiter möge jedes a_k ein Teiler von a_{k+1} sein, dann sind die $a_k(x)$ die Invariantenteiler von $A(x)$.

Beweis: Beide Matrizen haben dieselben Determinantenteiler d_k , da sie äquivalent sind. Das Polynom a_1 teilt alle Elemente der zweiten Matrix, also ist $d_1 = a_1$. Die 2-Minoren haben die Form $a_i a_j$, sie werden alle von $a_1 a_2$ geteilt, also ist $d_2 = a_1 a_2$. Analog sieht man $d_k = a_1 \dots a_k$.

Nun ist $i_1 = d_1 = a_1$, allgemeiner

$$i_k = \frac{d_k}{d_{k-1}} = \frac{a_1 \dots a_k}{a_1 \dots a_{k-1}} = a_k. \square$$

Damit können wir unsere obige Frage beantworten: Die oben verbliebenen Diagonalelemente sind die Invariantenteiler der Matrix.

Folgerung 5.1.1 *Zwei Polynommatrizen sind genau dann äquivalent, wenn sie dieselben Invariantenteiler besitzen.* \square

Definition: Zwei Matrizen $A, B \in M_n$ heißen ähnlich, wenn eine reguläre Matrix $X \in M_{nn}$ existiert, so daß $X^{-1}AX = B$ ist.

Im Kapitel über Normalformen haben wir uns ständig mit ähnlichen Matrizen befaßt (ohne es zu wissen).

Satz 5.1.4 *Die Matrizen A und B sind genau dann ähnlich, wenn die Polynommatrizen $A - Ex$ und $B - Ex$ äquivalent sind, also dieselben Invariantenteiler besitzen.*

Beweis: Sei $X^{-1}AX = B$, dann ist

$$X^{-1}(A - Ex)X = X^{-1}AX - Ex = B - Ex,$$

also sind $A - Ex$ und $B - Ex$ äquivalent.

Seien umgekehrt $A - Ex$ und $B - Ex$ äquivalent, dann gibt es invertierbare Polynommatrizen $P(x), Q(x)$, so daß

$$P(x)(A - Ex)Q(x) = B - Ex$$

gilt. Wir setzen

$$R(x) = P(x)^{-1},$$

dann gilt

$$(A - Ex)Q(x) = R(x)(B - Ex).$$

Wir dividieren nun $Q(x)$ von rechts mit Rest durch $B - Ex$ und $R(x)$ von links durch $A - Ex$:

$$\begin{aligned} Q(x) &= T(x)(B - Ex) + Q_0, \\ R(x) &= (A - Ex)S(x) + R_0, \end{aligned}$$

dabei sind Q_0 und R_0 skalare Matrizen (sie haben den Grad 0). Also ist

$$(A - Ex)(T(x)(B - Ex) + Q_0) = ((A - Ex)S(x) + R_0)(B - Ex)$$

$$(A - Ex)(T(x) - S(x))(B - Ex) = -(A - Ex)Q_0 + R_0(B - Ex)$$

Falls $S \neq T$ ist, hat die linke Matrix einen Grad ≥ 2 , die rechte Matrix hat aber höchstens den Grad 1, also ist

$$S(x) = T(x)$$

und damit

$$(A - Ex)Q_0 = R_0(B - Ex) = AQ_0 - Q_0x = R_0B - R_0x,$$

also

$$R_0 = Q_0 \text{ und } AQ_0 = AR_0 = R_0B.$$

Um die Ähnlichkeit von A und B zu beweisen, müssen wir noch zeigen, daß R_0 regulär ist. Dazu dividieren wir $P(x) = R(x)^{-1}$ mit Rest durch $(B - Ex)$:

$$P(x) = (B - Ex)U(x) + P_0,$$

dann ist

$$\begin{aligned} E &= R(x)P(x) = ((A - Ex)S(x) + R_0)((B - Ex)U(x) + P_0) \\ &= (A - Ex)S(x)(B - Ex)U(x) + R_0(B - Ex)U(x) + (A - Ex)S(x)P_0 + R_0P_0, \end{aligned}$$

Es ist

$$R_0(B - Ex) = (A - Ex)Q_0,$$

also ist

$$E = (A - Ex)(Q(x)U(x) + S(x)P_0) + R_0P_0,$$

dies ist eine Darstellung der Restdivision von E durch $(A - Ex)$, die sieht aber so aus:

$$E = (A - Ex)0 + E,$$

also ist $R_0P_0 = E$ und R_0 eine reguläre Matrix. □ □

Beispiel:

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

daraus wird

$$\begin{aligned} \begin{pmatrix} 1 & 2+x & 1 \\ x & 0 & -2 \\ 1 & 0 & 3+x \end{pmatrix} &\rightarrow \begin{pmatrix} 1 & 2+x & 1 \\ 0 & -2x-x^2 & -2-x \\ 0 & -2-x & 2+x \end{pmatrix} \rightarrow \\ &\begin{pmatrix} 1 & 0 & 0 \\ 0 & -x^2-3x-2 & -2-x \\ 0 & 0 & 2+x \end{pmatrix} \rightarrow \begin{pmatrix} 1 & & \\ & 2+x & \\ & & x^2+3x+2 \end{pmatrix} \end{aligned}$$

und der letzte Term ist gleich $(x+1)(x+2)$.

Die in einem Invariantenteiler einer Matrix auftretenden Potenzen eines irreduziblen Teilers des Invariantenteilers heißen deren Weierstrasssche Elementarteiler. Diese entsprechen (über dem Körper der komplexen Zahlen) den Jordanblöcken zum entsprechenden Eigenwert. Zwei Matrizen sind genau dann ähnlich, wenn ihre Elementarteiler übereinstimmen.

Wir wenden uns nun dem Minimalpolynom der Matrix $A \in M_{nn}$ zu. Dazu betrachten wir die Polynommatrix $A - Ex$. Die Matrix

$$B(x) = (b_{ij}(x))$$

sei die Adjunkte (die aus den $(n-1)$ -Minoren bestehende von $A - Ex$ gebildete) Matrix, sie hat den Grad $n-1$, es sei $d_1(x)$ der erste Determinantenteiler von $B(x)$, also der

größte gemeinsame Teiler der $b_{ij}(x)$. Wir teilen alle $b_{ij}(x)$ durch $d_1(x)$, es gibt also eine Polynommatrix $C(x)$, deren erster Determinantenteiler gleich 1 ist, mit

$$B(x) = d_1(x)C(x).$$

Aus der Formel für die Inverse einer Matrix erkennen wir

$$(A - Ex)B(x) = \det(A - Ex)E = c_A(x)E,$$

dabei ist $c_A(x)$ das charakteristische Polynom von A . Also gilt

$$c_A(x)E = d_1(x)(A - Ex)C(x),$$

also ist $c_A(x)$ durch $d_1(x)$ teilbar:

$$c_A(x) = d_1(x)m(x),$$

$$m(x)E = (A - Ex)C(x),$$

d.h. die Polynommatrix $m(x)E$ ist ohne Rest durch $A - Ex$ teilbar, also gilt

$$m(A)E = m(A) = 0,$$

also ist $m(x)$ ein annullierendes Polynom für A .

Satz 5.1.5 $m(x)$ ist das Minimalpolynom von A , es gilt $m(x)d_1(x) = c_A(x)$.

Beweis: Sei $n(x)$ das Minimalpolynom von A , dann ist $m(x) = f(x)n(x)$ und $n(x)E$ ist durch $A - Ex$ teilbar:

$$n(x)E = (A - Ex)D(x),$$

also

$$m(x)E = (A - Ex)D(x)f(x) = (A - Ex)C(x),$$

folglich ist $C(x) = D(x)f(x)$, d.h. $f(x)$ ist ein gemeinsamer Teiler der Komponenten von $C(x)$, also ist $f(x) = 1$ und $m(x) = n(x)$. \square

Folgerung 5.1.2 (Hamilton-Cayley) $c_A(A) = 0$. \square

Folgerung 5.1.3 Das Minimalpolynom von A ist gleich dem höchsten Invariantenteiler von $A - Ex$.

Beweis: Sei

$$P(x)(A - Ex)Q(x) = \begin{pmatrix} i_1 & & \\ & \dots & \\ & & i_n \end{pmatrix}.$$

Wir wissen, daß $c_A(x) = i_1 \dots i_n$ ist. Sei wieder $B(x)$ die Matrix der Adjunkten von $A - Ex$, dann ist

$$\begin{aligned} (A - Ex)B(x) &= c_A(x)E \\ &= P(x)c_A(x)P(x)^{-1} \end{aligned}$$

$$\begin{aligned}
&= P(x)(A - Ex)Q(x)Q(x)^{-1}B(x)P^{-1} \\
&= \begin{pmatrix} i_1 & & \\ & \cdots & \\ & & i_n \end{pmatrix} \begin{pmatrix} b_n & & ? \\ & \cdots & \\ ? & & b_1 \end{pmatrix} = \begin{pmatrix} i_1 \cdots i_n & & \\ & \cdots & \\ & & i_1 \cdots i_n \end{pmatrix}
\end{aligned}$$

da die $i_k \neq 0$ sind, ist auch die zweite Matrix eine Diagonalmatrix und es gilt

$$\begin{aligned}
b_n &= i_2 \cdots i_n, \\
b_{n-1} &= i_1 i_3 \cdots i_n, \\
&\dots \\
b_2 &= i_1 \cdots i_{n-2} i_n, \\
b_1 &= i_1 \cdots i_{n-1}.
\end{aligned}$$

Nun teilt b_1 das Polynom b_2 , b_2 teilt b_3 usw., also sind die b_k die Invariantenteiler von $B(x)$, es ist $c_A(x) = b_1 m(x)$, also ist $m(x) = i_n(x)$. \square

5.2 Die rationale Normalform

Zum Schluß wollen wir noch eine weitere Normalform einer skalaren Matrix finden.

Lemma 5.2.1 Sei $f(x) = x^n + a_1 x^{n-1} + \dots + a_n$ und

$$M(f) = \begin{pmatrix} 0 & & & -a_n \\ 1 & 0 & & -a_{n-1} \\ 0 & 1 & 0 & \\ & & \cdots & \\ & & & 1 & -a_1 \end{pmatrix}$$

dann ist $\det(M(f) - xE) = f(x)$ das Minimalpolynom von $M(f)$.

Beweis: Das charakteristische Polynom von $M(f)$ gleich $f(x)$ ist, sollten Sie schon wissen. Die zweite Behauptung folgt aus der ersten. Wir bestimmen die Determinantenteiler von $M(f) - xE$: Es gibt Minoren der Ordnung $1, \dots, n-1$, die gleich 1 sind, damit ist $d_1 = \dots = d_{n-1} = 1$ und $d_n = f$. \square

Satz 5.2.1 (rationale Normalform) Sei A eine skalare Matrix und i_r, \dots, i_n seien die nichtkonstanten Invariantenteiler von $A - xE$. Dann gibt es eine reguläre Matrix X , so daß

$$X^{-1}AX = \begin{pmatrix} M(i_r) & & \\ & \cdots & \\ & & M(i_n) \end{pmatrix}$$

eine Blockdiagonalmatrix ist.

Beweis: Nach dem Lemma stimmen die Invariantenteiler der zugehörigen Polynommatrizen überein. \square

5.3 Lokale Minimalpolynome eines Endomorphismus

Wir hatten früher gesehen, daß man am Minimalpolynom einer Matrix erkennen kann, ob es eine Basis aus Eigenvektoren gibt oder nicht: Dies ist genau dann der Fall, wenn alle Nullstellen des Minimalpolynoms einfach sind.

Ob mehrfache Nullstellen vorhanden sind, kann man erkennen, ohne diese berechnen zu müssen:

Lemma 5.3.1 *Wenn $f(x) = (x-x_0)^k g(x)$, $g(x_0) \neq 0$, $k > 1$ eine mehrfache Nullstelle x_0 besitzt, so ist x_0 auch eine Nullstelle von $f'(x)$, und umgekehrt.*

Beweis: Es ist $f'(x) = k(x-x_0)^{k-1}g(x) + (x-x_0)^k g'(x)$ und wegen $k > 1$ ist $f'(x_0) = 0$; wenn $k = 1$ gilt, so ist $f'(x_0) = g(x_0) \neq 0$. \square

Folgerung 5.3.1 *Das Polynom $f(x)$ hat genau dann mehrfache Nullstellen, wenn $ggT(f, f') \neq 1$ ist.* \square

Wenn A eine „zufällige“ Matrix ist, so sind deren Eigenwerte auch zufällig, also „oft“ voneinander verschieden. Demnach ist „fast jede“ Matrix diagonalisierbar. Schwieriger zu behandeln, aber mathematisch interessant sind die Sonderfälle.

Wir wollen uns nun näher mit Minimalpolynomen beschäftigen. Es sei im Folgenden $f : V \rightarrow V$ ein fixierter Endomorphismus und $m_f(x)$ sein Minimalpolynom.

Satz 5.3.1 *Sei $m_f(x) = g_1(x)g_2(x)$ mit teilerfremden Polynomen g_1, g_2 , dann gibt es invariante Unterräume $U_1, U_2 \subset V$ und das Minimalpolynom der Einschränkung $f|_{U_i}$ ist gleich $g_i(x)$.*

Beweis: Wir setzen $U_i = \{v \in V \mid g_i(f)(v) = o\}$. Wegen der Teilerfremdheit gibt es Polynome $h_i(x)$ mit

$$g_1 h_1 + g_2 h_2 = 1,$$

also

$$g_1(f) \circ h_1(f) + g_2(f) \circ h_2(f) = id_V.$$

Sei $v \in V$ beliebig, dann gilt

$$v = id_V(v) = g_1(f) \circ h_1(f)(v) + g_2(f) \circ h_2(f)(v)$$

und der erste Summand liegt in U_2 und der zweite in U_1 , denn (z.B. $i = 1$)

$$g_1(f) \circ g_2(f) \circ h_2(f)(v) = m_f(f) \circ h_2(f)(v) = o.$$

Somit ist $V = U_1 + U_2$. Sei nun $v \in U_1 \cap U_2$, also

$$g_1(f)(v) = g_2(f)(v) = o,$$

dann ist

$$v = id_V(v) = g_1(f) \circ h_1(f)(v) + g_2(f) \circ h_2(f)(v) = o,$$

also $U_1 \cap U_2 = \{o\}$.

Nach Konstruktion ist g_i ein annullierendes Polynom für $f|_{U_i}$. Wenn ein echter Teiler $h(x)$ von g_1 schon ein annullierendes Polynom für $f|_{U_1}$ wäre, so wäre $h \circ g_2$ ein annullierendes Polynom für f im Widerspruch zur Minimalität von $m_f(x)$. \square

Beispiel: $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$, $A^2 = \begin{pmatrix} 14 & 28 & 42 \\ 28 & 56 & 84 \\ 42 & 84 & 126 \end{pmatrix}$, $m_A(x) = x^2 - 14x$, denn A hat den Rang 1. Wir setzen $g_1(x) = x - 14$, $g_2(x) = x$, $U_1 = \{v \in R^3 \mid (A - 14E)v = o\} = L \begin{pmatrix} 16 \\ -2 \\ 3 \end{pmatrix}$, $U_2 = \{v \mid Av = o\} = L \left\{ \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \right\}$.

Definition: Sei $f : V \rightarrow V$ ein Endomorphismus und $v \in V$. Das normierte Polynom $m_{f,v}(x) \in R[x]$ heißt Minimalpolynom von f für v , wenn es das Polynom kleinsten Grades ist, so daß $m_{f,v}(f)(v) = o$ ist.

Beispiel: Sei $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, dann ist $A^2 = \begin{pmatrix} 4 & 4 \\ 0 & 4 \end{pmatrix}$ also $Ae_1 = 2e_1$, somit $(A - 2E)e_1 = o$, also $m_{A,e_1}(x) = x - 2$; $Ae_2 = e_1 + 2e_2$, $A^2e_2 = 4e_1 + 4e_2 = 4Ae_2 - 4Ee_2$, da $4Ae_2 = 4e_1 + 8e_2$ ist. Also gilt $(A^2 - 4A + 4E)e_2 = o$, demnach ist $m_{A,e_2} = x^2 - 4x + 4 = (x - 2)^2$.

Wenn $m_f(x)$ das Minimalpolynom von f ist, also $m_f(f) = 0$ ist, dann ist $m_f(f)(v) = 0$ für alle $v \in V$, also ist $m_{f,v}(x)$ ein Teiler von $m_f(x)$.

Welche Beziehungen bestehen nun zwischen verschiedenen Minimalpolynomen? Wir halten den Endomorphismus f fest und schreiben m_v anstelle von $m_{f,v}$.

Satz 5.3.2 Seien $v, w \in V$; wenn m_v und m_w teilerfremd sind, so ist $m_{v+w} = m_v m_w$.

Beweis: Sei $h(x)$ ein annullierendes Polynom für $v + w$, d.h. $h(f)(v + w) = o$. Dann ist

$$m_w(f)h(f)(v) = m_w(f)h(f)(v + w) - \underbrace{h(f)m_w(f)(w)}_{= o} = o$$

(Polynome in f kommutieren), also gilt $m_v \mid m_w h$, wegen der Teilerfremdheit von m_v und m_w folgt $m_v \mid h$ und analog $m_w \mid h$. Also wird jedes $v + w$ annullierende Polynom von $m_v m_w$ geteilt, also ist dies das Minimalpolynom. \square

Lemma 5.3.2 Sei $\{v_1, \dots, v_n\} \subset V$ eine Basis, dann ist $m_f(x)$ das kleinste gemeinsame Vielfache der m_{v_i} .

Beweis: Sei $g(x)$ ein gemeinsames Vielfaches der m_{v_i} , also $g(x) = h_i(x)m_{v_i}(x)$ und sei $v = \sum r_i v_i \in V$, dann gilt

$$g(f)(v) = \sum r_i h_i(f) m_{v_i}(f)(v_i) = o.$$

Wenn umgekehrt $g(f)$ alle Vektoren in V annulliert, so annulliert es auch die v_i , also ist $g(x)$ durch m_{v_i} teilbar, also ein gemeinsames Vielfaches der m_{v_i} . Das Polynom minimalen Grades mit dieser Eigenschaft ist das kleinste gemeinsame Vielfache. \square

Satz 5.3.3 *Es gibt einen Vektor $v \in V$ mit $m_{f,v} = m_f$.*

Beweis: Wir betrachten zunächst einen Spezialfall:

Sei $m_f(x) = g(x)^k$ die Potenz eines irreduziblen Polynoms und sei $\{v_1, \dots, v_n\}$ eine Basis von V . Die Minimalpolynome der Basisvektoren sind dann Teiler von $g(x)^k$, also $m_{v_i}(x) = g(x)^{k_i}$. Sei nun $m = \max k_i$, dann ist $g(x)^m = kgV(g(x)^{k_i}) = m_f(x) = g(x)^k$, also $l = k$ und ein Basisvektor v_i , wo das Maximum angenommen wird, leistet das Verlangte.

Sei nun

$$m_f(x) = \prod_{i=1}^m g_i(x)^{k_i}, \quad ggT(g_i, g_j) = 1 \text{ für } i \neq j,$$

dann ist $V = U_1 \oplus \dots \oplus U_m$ mit zu den $g(x)^{k_i}$ gehörigen invarianten Unterräumen, diese Polynome sind paarweise teilerfremd. Wir wählen Vektoren $u_i \in U_i$ mit den „richtigen“ Minimalpolynomen, das Minimalpolynom von $u_1 + \dots + u_m$ ist dann gleich $\prod_{i=1}^m g_i(x)^{k_i} = m_f(x)$. \square

Also

Folgerung 5.3.2 (rationale Normalform) *Sei V bezüglich f unzerlegbar, dann gibt es einen Vektor v , so daß $m_{f,v}(x) = m_f(x) = x^n + a_1x^{n-1} + \dots + a_n$ ist, die Vektoren $f^i(v)$, $i = 0, \dots, n-1$ sind linear unabhängig, die Darstellungsmatrix von f hat bezüglich der Basis $\{v, f(v), \dots, f^{n-1}(v)\}$ dann die Form*

$$M(f) = \begin{pmatrix} 0 & & & & -a_n \\ 1 & 0 & & & -a_{n-1} \\ 0 & 1 & 0 & & \\ & & & \dots & \\ & & & 1 & -a_1 \end{pmatrix}.$$

\square

Kapitel 6

Universelle Konstruktionen; abstract nonsense

Der Begriff der *Klasse* wird nicht definiert; stellen Sie sich eine Ansammlung von *Objekten* vor. Wenn M ein Objekt der Klasse C ist, so schreiben wir $M \in Ob(C)$.

Man kann also die Klasse aller Mengen betrachten, während der Begriff „Menge aller Mengen“ zur Verwirrung führt:

Sei \mathbf{M} die Menge aller Mengen, dann gilt natürlich $\mathbf{M} \in \mathbf{M}$. Sei nun \mathbf{N} die Menge der Mengen, die sich nicht als Element enthalten, also

$$M \in \mathbf{N} \Leftrightarrow M \notin M.$$

Gilt dann $\mathbf{N} \in \mathbf{N}$?

Wenn die Antwort „ja“ lautet, so ist $\mathbf{N} \notin \mathbf{N}$, bei „nein“ gälte $\mathbf{N} \in \mathbf{N}$

Definition

Eine Kategorie C besteht aus einer Klasse $Ob(C)$ von Objekten, wobei zu jedem Paar $A, B \in Ob(C)$ eine Menge $Mor(A, B)$ von Morphismen existiert. Weiter soll zu je drei Objekten A, B, C eine Kompositionsabbildung

$$Mor(B, C) \times Mor(A, B) \longrightarrow Mor(A, C)$$

$$(f, g) \mapsto f \circ g$$

existieren, so daß folgende Eigenschaften erfüllt sind:

1. $Mor(A, B) \cap Mor(A', B') = \emptyset$, wenn $A \neq A'$ oder $B \neq B'$ ist.
2. Zu jedem $A \in Ob(C)$ gibt es einen Morphismus $id_A \in Mor(A, A)$, so daß für alle $f \in Mor(A, B)$, $g \in Mor(B, A)$ gilt

$$f \circ id_A = f, id_A \circ g = g.$$

3. Wenn $f \in Mor(A, B)$, $g \in Mor(B, C)$, $h \in Mor(C, D)$ gilt, so folgt

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Anstelle von $f \in \text{Mor}(A, B)$ schreiben wir auch einfach $f : A \longrightarrow B$.

Beispiele für Kategorien sind die Kategorie der Vektorräume über einem Körper K mit den linearen Abbildungen als Morphismen, die Kategorie der Gruppen (der endlichen Gruppen, der abelschen Gruppen) mit den Gruppenhomomorphismen als Morphismen, die Kategorie der linken Moduln über einem Ring R mit den R -Modulhomomorphismen als Morphismen sowie die Kategorie der Mengen mit den Abbildungen als Morphismen. In all diesen Beispielen ist der Morphismus id_A die identische Abbildung.

Definition

Ein Morphismus $f : A \longrightarrow B$ heißt Isomorphismus, wenn ein Morphismus $g : B \longrightarrow A$ mit $f \circ g = id_B$, $g \circ f = id_A$ existiert. Ein Morphismus f heißt Monomorphismus, wenn aus $f \circ g_1 = f \circ g_2$ folgt $g_1 = g_2$. Ein Morphismus f heißt Epimorphismus, wenn aus $g_1 \circ f = g_2 \circ f$ folgt $g_1 = g_2$.

Die Begriffe „Monomorphismus“ und „Epimorphismus“ sind *duale* Begriffe; sie vertauschen sich, wenn man die Pfeilrichtungen vertauscht.

Satz 6.0.4 *In den obengenannten Kategorien sind die Monomorphismen genau die injektiven Abbildungen und Epimorphismen sind genau die surjektiven Abbildungen.*

Beweis: (Wir behandeln der Einfachheit halber die Kategorie der Mengen.) Es sei $f : Y \longrightarrow Z$ eine injektive Abbildung, $g_1, g_2 : X \longrightarrow Y$ und $f g_1 = f g_2$, dann ist $f(g_1(x)) = f(g_2(x))$ für alle x ; wegen der Injektivität folgt $g_1(x) = g_2(x)$ für alle x , also $g_1 = g_2$.

Sei f nicht injektiv, also $f(y_1) = f(y_2)$ mit $y_1 \neq y_2$. Wir wählen $X = \{x\}$ und setzen $g_1(x) = y_1$, $g_2(x) = y_2$. Dann ist $g_1 \neq g_2$, aber $f g_1 = f g_2$.

Nun sei $f : X \longrightarrow Y$ surjektiv und $g_1, g_2 : Y \longrightarrow Z$ mit $g_1 f = g_2 f$ gegeben. Dann gilt $g_1(f(x)) = g_2(f(x))$ für alle x . Jedes $y \in Y$ hat wegen der Surjektivität die Form $y = f(x)$, also gilt $g_1(y) = g_2(y)$ für alle y , also $g_1 = g_2$.

Sei f nicht surjektiv, dann gibt es ein $y \notin \text{Im}(f)$, wir wählen ein $y_1 \neq y$ und setzen $Z = Y$, $g_1 = id_Y$, $g_2(y') = y'$ für $y' \neq y$, $g_2(y) = y_1$. Dann gilt $g_1 f = g_2 f$. \square

Satz 6.0.5 *Wenn f, g Monomorphismen sind, so ist $f \circ g$ auch ein Monomorphismus. Wenn f, g Epimorphismen sind, so ist $f \circ g$ auch ein Epimorphismus. Wenn $f \circ g$ ein Monomorphismus ist, so ist g ein Monomorphismus. Wenn $f \circ g$ ein Epimorphismus ist, so ist f ein Epimorphismus.*

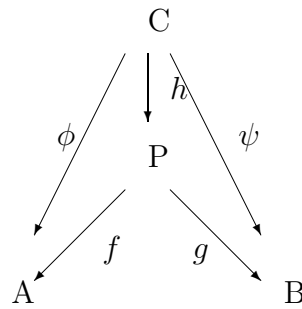
Beweis: Wir beweisen jeweils die Aussagen für Monomorphismen, der Rest geschieht durch Dualisierung.

Sei $f \circ g \circ h_1 = f \circ g \circ h_2$, dann ist wegen der Monomorphie von f auch $g \circ h_1 = g \circ h_2$ und aus der Monomorphie von g folgt $h_1 = h_2$.

Sei $g \circ h_1 = g \circ h_2$, dann gilt $(f \circ g) \circ h_1 = (f \circ g) \circ h_2$, also $h_1 = h_2$. \square

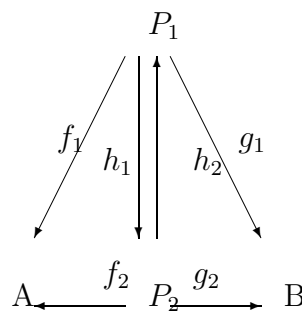
Definition

Seien A, B Objekte; ein Produkt von A, B ist ein Tripel $(P, f : P \longrightarrow A, g : P \longrightarrow B)$, so daß für alle $\phi : C \longrightarrow A$, $\psi : C \longrightarrow B$ ein eindeutig bestimmter Morphismus $h : C \longrightarrow P$ existiert, so daß $\phi = f \circ h$, $\psi = g \circ h$ gilt, d.h. das folgende Diagramm ist kommutativ:



Satz 6.0.6 Wenn ein Produkt von A, B existiert, dann ist es bis auf Isomorphie eindeutig bestimmt; man bezeichnet es mit $A \times B$.

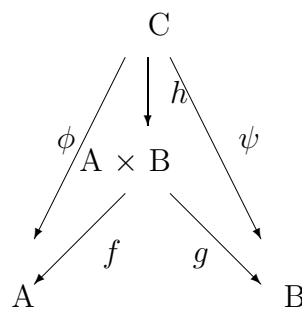
Beweis: Seien (P_1, f_1, g_1) und (P_2, f_2, g_2) zwei Produkte von A und B . Dann existieren eindeutig bestimmte h_1, h_2 mit $f_1 = f_2 \circ h_1$, $g_1 = g_2 \circ h_1$ und $f_2 = f_1 \circ h_2$, $g_2 = g_1 \circ h_2$.



Es folgt $f_1 = f_1 \circ h_2 \circ h_1$ und $g_1 = g_1 \circ h_2 \circ h_1$. Wegen der Einzigkeit von h folgt $h_2 \circ h_1 = id_{P_1}$ und analog $h_1 \circ h_2 = id_{P_2}$, also sind P_1 und P_2 isomorph. \square

Satz 6.0.7 In der Kategorie der Mengen (Gruppen, R -Moduln,...) existiert das Produkt je zweier Objekte, und zwar ist es das kartesische Produkt zusammen mit den beiden Projektionen auf die Faktoren.

Beweis:



Die Abbildung h ist durch $h(c) = (\phi(c), \psi(c))$ gegeben. \square

Wir dualisieren den Begriff des Produkts.

Definition: Seien A, B Objekte; ein Koprodukt von A, B ist ein Tripel $(P, f : A \rightarrow P, g : B \rightarrow P)$, so daß für alle $\phi : A \rightarrow C, \psi : B \rightarrow C$ ein eindeutig bestimmter Morphismus $h : P \rightarrow C$ existiert, so daß $\phi = h \circ f, \psi = h \circ g$ gilt.

Satz 6.0.8 Wenn ein Koproduct von A, B existiert, dann ist es bis auf Isomorphie eindeutig bestimmt; man bezeichnet es mit $A \oplus B$ oder $A \sqcup B$.

Der Beweis ist dual zum obigen. □

Satz 6.0.9 In der Kategorie der Mengen existiert das Koproduct beliebiger Mengen A, B , es besteht aus der disjunkten Vereinigung $A \sqcup B$ zusammen mit den beiden Einbettungen von A und B in $A \sqcup B$.

Beweis: Die gesuchte Abbildung $h : A \sqcup B \rightarrow C$ ist durch $h(a) = \phi(a)$, $h(b) = \psi(b)$ gegeben. □

Satz 6.0.10 In der Kategorie der R -Moduln existiert das Koproduct beliebiger Moduln A, B , es besteht aus der direkten Summe $A \oplus B$ zusammen mit den beiden Einbettungen von A und B in $A \oplus B$.

Beweis: Die gesuchte Abbildung $h : A \oplus B \rightarrow C$ ist durch $h(a + b) = \phi(a) + \psi(b)$ gegeben. □

Definition

Ein Objekt U heißt universell, wenn es zu jedem Objekt A genau einen Morphismus $f_A : U \rightarrow A$ gibt.

Wir fragen uns nach der Existenz. Offenbar sind alle universellen Objekte einer Kategorie isomorph, falls welche existieren.

In der Kategorie der Mengen ist die leere Menge universell, in der Kategorie der Gruppen leistet dies eine einelementige Gruppe, in der Kategorie der R -Moduln ist der Nullmodul universell.

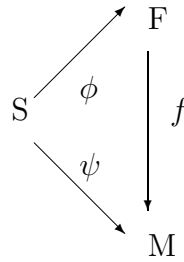
Wir betrachten nun eine Kategorie, deren Objekte keine Mengen (mit einer algebraischen Struktur) sind:

Sei S eine Menge und R ein Ring. Die Objekte der Kategorie R_S seien Abbildungen $\phi : S \rightarrow M$, wobei M ein R -Modul ist. Ein Morphismus von $\phi : S \rightarrow M$ in $\psi : S \rightarrow N$ ist ein R -Modulhomomorphismus $f : M \rightarrow N$, so daß das Diagramm

$$\begin{array}{ccc}
 & & M \\
 & \nearrow \phi & \downarrow f \\
 S & & \\
 & \searrow \psi & \downarrow \\
 & & N
 \end{array}$$

kommutativ wird, d.h. $f \circ \phi = \psi$.

Ein universelles Objekt der Kategorie R_S ist also eine Abbildung $\phi : S \rightarrow F$ in einen R -Modul F , so daß zu jeder Abbildung $\psi : S \rightarrow M$ in einen R -Modul M ein eindeutig bestimmter R -Modulhomomorphismus $f : F \rightarrow M$ mit kommutativem Diagramm



existiert. Wir werden gleich sehen, daß ein derartiges universelles Objekt existiert; der Modul F heißt der von S erzeugte freie R -Modul.

Satz 6.0.11 Sei $F = \{\sum r_i s_i \in r_i \in R, s_i \in S\}$ die Menge aller formaler Linearkombinationen und $\phi : S \hookrightarrow F$ die Einbettung; dies ist ein universelles Objekt in R_S .

Beweis: Offenbar ist F ein R -Modul. Wenn $\psi : S \rightarrow M$ eine beliebige Abbildung ist, so ist durch $f(\sum r_i s_i) = \sum r_i \psi(s_i)$ ein R -Modulhomomorphismus gegeben und es gilt $f \circ \phi = \psi$. \square

Definition

Ein Objekt O heißt Nullobjekt, wenn es sowohl universell als auch kouniversell ist, d.h. zu jedem Objekt A gibt es genau einen Morphismus $O \rightarrow A$ und genau einen Morphismus $A \rightarrow O$.

Ein Morphismus $o : A \rightarrow B$ heißt Nullmorphismus, wenn er gleich einem Produkt $A \rightarrow O \rightarrow B$ ist.

Sei $f : A \rightarrow B$ ein Morphismus; ein Morphismus $k : K \rightarrow A$ heißt Kern von f , wenn $f \circ k = o$ ist und wenn es zu jedem $g : G \rightarrow A$ mit $f \circ g = o$ einen eindeutig bestimmten Morphismus $h : G \rightarrow K$ mit $g = k \circ h$ gibt.

$$K \rightarrow A \rightarrow B$$

$$\begin{array}{c} \uparrow \nearrow \\ G \end{array}$$

Lemma 6.0.3 Jeder Kern ist ein Monomorphismus.

Seien $g_1, g_2 : X \rightarrow K$ Morphismen mit $kg_1 = kg_2$. Wir setzen $g = kg_1 : X \rightarrow A$, dann ist $fg = fkg_1 = o$, also existiert ein eindeutig bestimmtes $h : X \rightarrow K$ mit $g = kh = kg_1 = kg_2$, also ist $h = g_1 = g_2$. \square

Lemma 6.0.4 Der Kern eines Monomorphismus ist der Nullmorphismus.

Beweis: Aus $fk = o = fo$ und der Monomorphie von f folgt $k = o$. \square

Nun nehmen Sie ein Buch über Kategorientheorie und beweisen Sie als Übungsaufgaben alle Sätze.

Kapitel 7

Tensorprodukte

Sei im folgenden sei stets R ein Körper und M und N seien R -Vektorräume.

Wir machen eine Vorbemerkung:

Sei $f : M \rightarrow N$ eine R -lineare Abbildung und $U \subset M$ ein Unterraum mit $f(U) = \{0\}$. Dann wird durch f eine Abbildung $f' : M/U \rightarrow N$ induziert, wir setzen $f'(m+U) = f(m)$, dies ist wohldefiniert, wie man schnell nachrechnet.

Wir wollen einen neuen R -Vektorraum $M \otimes_R N$ konstruieren, der von Elementen der Form $m \otimes n$, $m \in M, n \in N$ erzeugt wird, wobei folgende Rechenregeln gelten sollen:

$$\begin{aligned}(m + m') \otimes n &= m \otimes n + m' \otimes n \\ m \otimes (n + n') &= m \otimes n + m \otimes n' \\ (rm) \otimes n &= m \otimes (rn), \quad (r \in R)\end{aligned}$$

Wir konstruieren diesen Vektorraum folgendermaßen: Sei $F(M \times N)$ der von der Menge $M \times N$ erzeugte R -Vektorraum, sei S der Unterraum von $F(M \times N)$, der von allen Elementen der Form

$$\begin{aligned}(m + m', n) - (m, n) - (m', n) \\ (m, n + n') - (m, n) - (m, n') \\ (rm, n) - (m, rn) \\ (m, m' \in M, n, n' \in N, a \in R)\end{aligned}$$

erzeugt wird.

Wir setzen $M \otimes_R N = F(M \times N)/S$, die Äquivalenzklasse von (m, n) bezeichnen wir mit $m \otimes n$, dann sind die obigen Rechenregeln erfüllt. Der Vektorraum $M \otimes_R N$ wird das Tensorprodukt von M und N genannt. Die Elemente von $M \otimes N$ haben die Gestalt $\sum_{i \in I} m_i \otimes n_i$, $m_i \in M, n_i \in N$.

Ein Element der Form $m \otimes n$ heißt zerfallender Tensor, die zerfallenden Tensoren bilden ein Erzeugendensystem. Zum Beispiel ist

$$\begin{aligned}m_1 \otimes n_1 + 2m_1 \otimes n_2 + 2m_2 \otimes n_1 + 4m_2 \otimes n_2 \\ = m_1 \otimes (n_1 + 2n_2) + 2m_2(n_1 + 2n_2) = (m_1 + 2m_2) \otimes (n_1 + 2n_2)\end{aligned}$$

ein zerfallender Tensor.

Definition: Seien M, N, P Vektorräume; eine Abbildung $f : M \times N \longrightarrow P$ mit

$$f(m + m', n) = f(m, n) + f(m, n')$$

$$f(m, n + n') = f(m, n) + f(m, n')$$

$$f(rm, n) = rf(m, n) = f(m, rn)$$

heißt R -bilinear.

Beispiel: Die kanonische Abbildung $k : M \times N \longrightarrow M \otimes_R N$, $(m, n) \mapsto m \otimes n$, ist bilinear.

Satz 7.0.12 Sei $f : M \times N \longrightarrow P$ eine bilineare Abbildung. Dann gibt es eine eindeutig bestimmte R -lineare Abbildung $g : M \otimes_R N \longrightarrow P$, so daß das Diagramm

$$\begin{array}{ccc} M \times N & \longrightarrow & P \\ k \downarrow & \nearrow g & \\ M \otimes_R N & & \end{array}$$

kommutativ ist. („Die bilineare Abbildung f induziert die lineare Abbildung g “.)

Beweis: Wir setzen zunächst f linear auf $F(M \times N)$ fort:

$$f' : F(M \times N) \longrightarrow P, \quad f'(\sum r_i(m_i, n_i)) = \sum r_i f(m_i, n_i).$$

Wir überlegen, daß $f'(s) = 0$ für $s \in S$ gilt. Sei etwa $s = (m + m', n) - (m, n) - (m', n)$, dann ist

$$\begin{aligned} f'(s) &= f'((m + m', n) - (m, n) - (m', n)) \\ &= f'(m + m', n) - f'(m, n) - f'(m', n) \\ &= f(m + m', n) - f(m, n) - f(m', n) = 0 \end{aligned}$$

wegen der Bilinearität von f . Also induziert f' eine lineare Abbildung $g : M \otimes_R N \longrightarrow P$ mit $g(m \otimes n) = g((m, n) + S) = f'(m, n) = f(m, n)$.

Da die Elemente der Form $m \otimes n$ den Vektorraum $M \otimes_R N$ erzeugen, ist die Abbildung g eindeutig bestimmt. \square

Satz 7.0.13 (Isomorphieeigenschaften) 1. $M \otimes_R N \cong N \otimes_R M$,

2. $(M \otimes_R N) \otimes_R P \cong M \otimes_R (N \otimes_R P)$,

3. $R \otimes_R M \cong M \cong M \otimes_R R$.

Beweis: 1. Die Abbildung

$$f : N \times M \longrightarrow M \otimes_R N, \quad f(n, m) = m \otimes n,$$

ist bilinear und induziert eine lineare Abbildung $g : N \otimes_R M \longrightarrow M \otimes_R N$, diese besitzt offensichtlich eine Inverse.

2. Wir fixieren ein $p \in P$ und betrachten

$$h_p : M \times N \longrightarrow M \otimes_R (N \otimes_R P),$$

$$h_p(m, n) = m \otimes (n \otimes p),$$

sie ist bilinear und induziert eine lineare Abbildung

$$g_p : M \otimes_R N \longrightarrow M \otimes_R (N \otimes_R P).$$

Wir betrachten nun

$$g : (M \otimes_R N) \times P \longrightarrow M \otimes_R (N \otimes_R P),$$

$$(m \otimes n, p) \mapsto m \otimes (n \otimes p),$$

diese Abbildung ist bilinear und induziert eine lineare Abbildung

$$(M \otimes_R N) \otimes_R P \longrightarrow M \otimes_R (N \otimes_R P),$$

die wiederum offensichtlich eine Inverse besitzt.

3. Die Abbildung $f : R \times M, f(r, m) = rm$, ist bilinear, usw. □

Seien $f : M \longrightarrow P, g : N \longrightarrow Q$ R -lineare Abbildungen, dann ist die Abbildung

$$h : M \times N \longrightarrow P \otimes_R Q$$

$$(m, n) \mapsto f(m) \otimes g(n)$$

bilinear, also wird eine lineare Abbildung

$$f \otimes g : M \otimes_R N \longrightarrow P \otimes_R Q$$

induziert, sie heißt das Tensorprodukt der Abbildungen f und g .

Lemma 7.0.5

$$f \otimes g(m \otimes n) = f(m) \otimes g(n)$$

$$(f + f') \otimes g = f \otimes g + f' \otimes g$$

$$f \otimes (g + g') = f \otimes g + f \otimes g'$$

$$(af) \otimes g = f \otimes (ag) = a f \otimes g$$

$$(f_2 \circ f_1) \otimes (g_2 \circ g_1) = (f_2 \otimes g_2) \circ (f_1 \otimes g_1) \quad \square$$

Lemma 7.0.6 $(M \oplus M') \otimes_R N \cong M \otimes_R N \oplus M' \otimes_R N$.

Beweis: Sei $P = M \oplus M'$, wir können die direkte Summe durch idempotente Endomorphismen charakterisieren: Wir haben die Projektionen

$$e_1, e_2 : P \longrightarrow P, \quad e_1 + e_2 = id, \quad e_i \circ e_j = \delta_{ij} e_i, \quad M = \text{Im } e_1, \quad M' = \text{Im } e_2.$$

Wir tensorieren mit $id : N \longrightarrow N$, also $f_i = e_i \otimes id : P \otimes_R N \longrightarrow P \otimes_R N$ und erhalten für die f_i analoge Relationen. \square

Sei nun $M = R^m$, dann folgt

$$M \otimes_R N = \left(\bigoplus_{i=1}^m R \right) \otimes_R N = \bigoplus_{i=1}^m (R \otimes_R N) \cong \bigoplus_{i=1}^m N = N^m.$$

Wenn auch noch beachten, daß $N = R^n$ ist, so folgt

$$M \otimes_R N = (R^m)^n = R^{mn}.$$

Folgerung 7.0.3 Sei K ein Körper und seien V, W K -Vektorräume, dann gilt $\dim V \otimes_K W = \dim V \cdot \dim W$. Seien $\{v_1, \dots, v_n\}$, $\{w_1, \dots, w_m\}$ Basen von V bzw. W , dann ist $\{v_i \otimes w_j \mid i = 1, \dots, n; j = 1, \dots, m\}$ eine Basis von $V \otimes_K W$ und jedes Element $t \in V \otimes_K W$ läßt sich mit $t = \sum t_{ij} v_i \otimes w_j$ durch einen „Tensor“ (t_{ij}) beschreiben.

Seien nun $f_1 : V_1 \longrightarrow W_1$, $f_2 : V_2 \longrightarrow W_2$ lineare Abbildungen, dazu erhalten wir die lineare Abbildung

$$f_1 \otimes f_2 : V_1 \otimes V_2 \longrightarrow W_1 \otimes W_2,$$

wir wollen deren Darstellungsmatrix bestimmen. Dazu wählen wir Basen $\{b_i\}$ von V_1 , $\{c_j\}$ von V_2 , $\{d_k\}$ von W_1 , $\{e_l\}$ von W_2 und ordnen die Basiselemente von $V_1 \otimes V_2$ bzw. $W_1 \otimes W_2$ folgendermaßen an:

$$G = \{b_1 \otimes c_1, \dots, b_n \otimes c_1, b_1 \otimes c_2, \dots, b_n \otimes c_2, \dots, b_1 \otimes c_m, \dots, b_n \otimes c_m\},$$

$$H = \{d_1 \otimes e_1, \dots, d_p \otimes e_1, d_1 \otimes e_2, \dots, d_p \otimes e_2, \dots, d_1 \otimes e_q, \dots, d_p \otimes e_q\}.$$

Sei $X = A_{BD}(f_1)$, $Y = A_{CE}(f_2)$, dann hat A_{GH} die folgende Block-Gestalt:

$$\begin{pmatrix} Xy_{11} & \dots & Xy_{1m} \\ & \dots & \\ Xy_{p1} & \dots & Xy_{pq} \end{pmatrix}.$$

Diese Matrix heißt das Kroneckerprodukt (oder auch Tensorprodukt) der Matrizen X und Y .

Wir betrachten die folgende (offenbar lineare) Abbildung

$$f : V \otimes W^* \longrightarrow \text{Hom}(W, V), \quad f(v \otimes l) = f_{v \otimes l},$$

$$f_{v \otimes l}(w) = l(w) v.$$

(W^* ist der zu W duale Vektorraum.)

Wir zeigen, daß f ein Isomorphismus ist. Da

$$\dim(V \otimes W^*) = \dim V \cdot \dim W^* = \dim V \cdot \dim W = \dim \operatorname{Hom}(W, V)$$

ist, genügt es, die Surjektivität nachzuweisen.

Seien $B = \{b_i\}$, $\{c_j\}$ Basen von W bzw. V . Dann bilden die Abbildungen $G_{ij} : W \rightarrow V$ mit $g_{ij}(b_k) = \delta_{ik} c_j$ eine Basis von $\operatorname{Hom}(W, V)$. Sei $\{b_i^*\}$ die B duale Basis von W^* , dann ist $f(c_j \otimes b_i^*) = g_{ij}$, folglich ist f surjektiv.

Folgerung 7.0.4

$$\operatorname{Hom}(V, W) \cong V^* \otimes W. \square$$

Wir kehren wieder zu beliebigen Grundringen zurück.

Satz 7.0.14 *Seien $U \subset M$, $T \subset N$ Untermoduln, dann ist*

$$(M/U) \otimes_R (N/T) \cong M \otimes_R N / (U \otimes_R N + M \otimes_R T).$$

Beweis: Seien $f : M \rightarrow M/U$, $g : N \rightarrow N/T$ die kanonischen Abbildungen, wir betrachten die Abbildung

$$f \otimes g : M \otimes_R N \rightarrow M/U \otimes_R N/T$$

$$m \otimes n \mapsto \bar{m} \otimes \bar{n} = f(m) \otimes g(n).$$

Sei $m \in U$ oder $t \in T$, dann ist $f \otimes g(m \otimes n) = 0$, also

$$f \otimes g|_{U \otimes_R N + M \otimes_R T} = 0,$$

also wird eine Abbildung

$$h : M \otimes N / (U \otimes_R N + M \otimes_R T) \rightarrow M/U \otimes_R N/T$$

induziert, wobei $\overline{m \otimes n} \mapsto \bar{m} \otimes \bar{n}$. Wir zeigen, daß die Umkehrabbildung

$$k : M/U \otimes_R N/T \rightarrow M \otimes N / (U \otimes_R N + M \otimes_R T)$$

$$\bar{m} \otimes \bar{n} \mapsto \overline{m \otimes n}$$

wohldefiniert ist:

$$\overline{(m+u) \otimes (n+t)} = \bar{m} \otimes \bar{n} + \bar{u} \otimes \bar{n} + \bar{m} \otimes \bar{t} + \bar{u} \otimes \bar{t},$$

der erste Summand wird auf $\overline{m \otimes n}$, die restlichen auf Null abgebildet.

Folgerung 7.0.5 *Seien I und J Ideale von R , dann gilt*

$$R/I \otimes_R R/J \cong R/(I+J). \square$$

Beispiel: $\mathbf{Z}/2\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/3\mathbf{Z} = 0$, $\mathbf{Z}/2\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/2\mathbf{Z} = \mathbf{Z}/2\mathbf{Z}$.

Kapitel 8

Halbeinfache Algebren und Moduln

8.1 Grundlagen

Eine Algebra ist ein spezielles algebraisches Objekt.

Definition: Sei K ein Körper, A ein K -Vektorraum und gleichzeitig ein Ring; vermöge der Abbildung $K \rightarrow A$ mit $k \mapsto k \cdot 1$ wird K in A eingebettet, wir identifizieren K mit $K \cdot 1$. Es gelte

$$k \cdot a = a \cdot k \text{ für alle } k \in K, a \in A.$$

Dann heißt A eine K -Algebra.

Sei M ein linker A -Modul, wegen $K \subseteq A$ operiert auch K auf M , d.h. M ist auch ein K -Vektorraum.

Beispiele:

$K[x]$, $K[x_1, \dots, x_n]$, $K[x_1, \dots, x_n]/I$ für ein Ideal I von $K[x_1, \dots, x_n]$, $M_{nn}(K)$, die Menge $T_n(K)$ der oberen Dreiecksmatrizen, die Menge D_n der Diagonalmatrizen, $K \times \dots \times K$ mit komponentenweiser Addition und Multiplikation.

Wir vereinbaren, daß alle in diesem Abschnitt betrachteten Vektorräume endlichdimensional sind (dann fallen zwei der obigen Beispiele aus dem Rahmen).

Sei $\dim A = n$, dann wählen wir eine Basis $\{e_1, \dots, e_n\}$ des Vektorraums A , dann lassen sich die Produkte $e_i e_j$ als Linearkombination der e_k darstellen:

$$e_i e_j = \sum a_{ijk} e_k \text{ mit } a_{ijk} \in K.$$

Die n^3 Zahlen a_{ijk} heißen die Strukturkonstanten der Algebra (bezüglich der gewählten Basis), durch sie ist die Multiplikation in A eindeutig bestimmt:

$$\sum x_i e_i \cdot \sum y_j e_j = \sum x_i y_j a_{ijk} e_k.$$

Die Strukturkonstanten sind zur Konstruktion einer Algebra nicht willkürlich wählbar, denn die Multiplikation soll assoziativ sein, also muß gelten:

$$(e_i e_j) e_l = \sum a_{ijk} e_k e_l = \sum a_{ijk} a_{klm} e_m,$$

$$e_i(e_j e_l) = e \sum_i a_{jlk} e_k = \sum a_{jlk} a_{ikm} e_m,$$

dafür ist notwendig und hinreichend, daß

$$\sum a_{ijk} a_{klm} = \sum a_{jlk} a_{ikm} \text{ für alle } i, j, l, m.$$

Eine Menge S heißt Halbgruppe, wenn eine Multiplikation $\cdot : S \times S \rightarrow S$ gegeben ist, die das Assoziativgesetz erfüllt und für die es ein neutrales Element 1 gibt.

Sei S eine endliche Halbgruppe, wir ordnen ihr die folgende „Halbgruppenalgebra“ K^S zu. Wir setzen

$$K^S = \{f : S \rightarrow K\}$$

und definieren Addition, K -Multiplikation und Ringmultiplikation wie folgt:

$$(f_1 + f_2)(s) = f_1(s) + f_2(s),$$

$$(k \cdot f)(s) = k \cdot f(s),$$

$$(f_1 \cdot f_2)(s) = \sum_{rt=s} f_1(r) \cdot f_2(t),$$

Es ist nicht schwierig nachzurechnen, daß K^S eine K -Algebra der Dimension $|S|$ ist.

Definition: Sei A ein Ring und M ein linker A -Modul, M heißt einfach, wenn M keine echten Untermoduln besitzt. Ein Linksideal $L \subseteq A$ heißt minimal, wenn $\{0\}$ das einzige echt in L enthaltene Linksideal ist. Ein Linksideal L heißt maximal, wenn A das einzige L echt enthaltende Linksideal ist.

Satz 8.1.1 1. Jedes minimale Linksideal ist ein einfacher A -Modul.

2. Für jedes maximale Linksideal $L \subseteq A$ ist A/L ein einfacher A -Modul.

3. Jeder einfache A -Modul ist isomorph zu A/L für ein geeignetes maximales Linksideal $L \subseteq A$.

Beweis: Die beiden ersten Aussagen sind trivial, wir beweisen nur die dritte: Sei M einfach und $0 \neq m \in M$, dann ist $\{0\} \neq Am \subseteq M$ ein Untermodul, also $Am = M$. Wir betrachten den folgenden Modulhomomorphismus

$$f : A \rightarrow M, a \mapsto am,$$

dieser ist offenbar surjektiv und $L = \text{Ker}(f)$ ist ein Linksideal von A . Nach dem Homomorphiesatz gilt $A/L \cong M$ und wegen der Einfachheit von M muß L maximal sein.

□

Beispiele:

Sei $A = K$, die einfachen K -Moduln sind 1-dimensionale Vektorräume, also isomorph zu K , K besitzt das minimale Ideal K und das maximale Ideal $\{0\}$.

Wir betrachten die Matrixalgebra $A = M_{22}(K)$ und deren Unterräume, wir bezeichnen mit $\begin{pmatrix} * & 0 \\ * & 0 \end{pmatrix}$ die Menge aller Matrizen, deren zweite Spalte Null ist. Die Menge $\begin{pmatrix} * & 0 \\ * & 0 \end{pmatrix}$ ist ein minimales und ein maximales Linksideal von A , der Vektorraum $K^2 = \begin{pmatrix} * \\ * \end{pmatrix}$ der Spaltenvektoren ist ein einfacher A -Modul.

Der Original-Beweis des folgenden Satzes ist etwa eine Druckseite lang.

Lemma 8.1.1 (Schur) 1. Sei M ein einfacher A -Modul und $f : M \rightarrow M$ eine A -lineare Abbildung, dann ist f ein Isomorphismus oder $f = 0$.

2. Sei A eine \mathbb{C} -Algebra, M ein einfacher A -Modul und $f : M \rightarrow M$ eine A -lineare Abbildung, dann gilt $f = z \cdot id_M$ für ein $z \in \mathbb{C}$.

Beweis: 1. $\text{Ker}(f)$, $\text{Im}(f) \subseteq M$ sind Untermoduln, also müssen sie gleich $\{0\}$ oder gleich M sein.

2. Sei $z \in \mathbb{C}$ ein Eigenwert der linearen Abbildung f , dann ist $f - z \cdot id_M$ kein Isomorphismus, also $f - z \cdot id = 0$. \square

Eine A -lineare Abbildung eines A -Moduls M in sich nennt man einen A -Endomorphismus, die Menge aller A -Endomorphismen von M wird mit $\text{End}_A(M)$ bezeichnet. Mit der Nacheinanderausführung von Endomorphismen als Multiplikation wird $\text{End}_A(M)$ eine K -Algebra.

Folgerung 8.1.1 Sei A eine \mathbb{C} -Algebra und M ein einfacher A -Modul. Dann gilt $\text{End}_A(M) \cong \mathbb{C}$.

Beweis: Wir ordnen dem Endomorphismus f seinen Eigenwert zu. \square

Definition: Eine K -Algebra A heißt einfache Algebra, wenn A genau zwei Ideale besitzt, nämlich $\{0\}$ und A .

Beispiele:

1. Ein Körper K ist eine einfache K -Algebra.
2. Sei R ein kommutativer Ring ohne echte Ideale. Sei $0 \neq r \in R$, dann ist das von r erzeugte Ideal $rR \neq \{0\}$, also $rR = R \ni 1$, also gibt es ein $s \in R$ mit $rs = 1$. Jedes von Null verschiedene Element von R besitzt ein Inverses, also ist R ein Körper.
3. Wir wollen nachweisen, daß die Matrixalgebra $A = M_{nn}(K)$ eine einfache K -Algebra ist. Sei also $\{0\} \neq I \subseteq A$ ein Ideal. Dann muß für jede Matrix $M = (m_{ij}) \in I$ und alle Matrizen $X, Y \in A$ auch das Produkt XYM in I liegen. Mit E_{ij} bezeichnen wir die Matrix, die nur an der Stelle (i, j) eine 1 und sonst Nullen enthält. Sei $m_{ij} \neq 0$, dann gilt

$$\frac{1}{m_{ij}}(E_{1i}ME_{j1} + E_{2i}ME_{j2} + \dots + E_{ni}ME_{jn}) = E \in I,$$

also $I = A$.

Satz 8.1.2 (Wedderburn) Jede (endlichdimensionale) einfache \mathbb{C} -Algebra A ist isomorph zu einer Matrixalgebra $M_{nn}(\mathbb{C})$.

Beweis: Sei $L \subseteq A$ ein minimales Linksideal. Zum Element $x \in L$ betrachten wir die Abbildung

$$f_x : L \rightarrow L, f_x(l) = l \cdot x,$$

sie ist A -linear, denn

$$f_x(al_1 + l_2) = (al_1 + l_2)x = al_1x + l_2x = af_x(l_1) + f_x(l_2).$$

Also gilt $f_x \in \text{End}_A(L) \cong \mathbb{C}$, also $f_x = z_x \cdot \text{id}$ mit $z_x \in \mathbb{C}$. Sei nun $a \in A$ beliebig, wir betrachten die Abbildung

$$g_a : L \rightarrow L, g_a(l) = al.$$

Die Abbildung g_a ist offenbar \mathbb{C} -linear. Es gilt

$$g_1 = \text{id}_L, g_{a+b} = g_a + g_b, g_{ab} = g_a \cdot g_b,$$

also ist

$$g : A \rightarrow \text{End}_{\mathbb{C}}(L), g(a) = g_a,$$

ein Ringhomomorphismus, es ist $g \neq 0$, somit ist $\text{Ker}(g) \neq A$ ein Ideal von A , also $\text{Ker}(g) = \{0\}$, somit ist g injektiv. Wir wollen noch beweisen, daß g surjektiv ist.

Die Menge $L \cdot A$ ist ein Ideal von A , also gilt $L \cdot A = A$, also

$$g(A) = g(LA) = g(L) \cdot g(A).$$

Wir zeigen nun, daß $g(L) \subseteq \text{End}_{\mathbb{C}}(L)$ ein Linksideal ist.

Sei also $h \in \text{End}_{\mathbb{C}}(L)$ und $x, l \in L$. Dann gilt $h \circ g_l(x) = h(lx) = h(f_x(l)) = h(z_x \cdot l) = z_x \cdot h(l) = f_x \circ h(l) = h(l) \cdot x = g_{h(l)}(x)$, also $h \circ g_l = g_{h(l)} \in g(L)$. Folglich ist $\text{End}_{\mathbb{C}}(L) \cdot g(L) = g(L)$ und wegen $1 \in g(A)$ gilt $\text{End}_{\mathbb{C}}(L) \cdot g(A) = \text{End}_{\mathbb{C}}(L)$.

Nun folgt $g(A) = g(L) \cdot g(A) = \text{End}_{\mathbb{C}}(L) \cdot g(L) \cdot g(A) = \text{End}_{\mathbb{C}}(L) \cdot g(A) = \text{End}_{\mathbb{C}}(L)$, also ist g surjektiv, d.h. $A \cong \text{End}_{\mathbb{C}}(L)$.

Wenn nun $\dim_{\mathbb{C}}(L) = n$ ist, so ist $\text{End}_{\mathbb{C}}(L) \cong M_{nn}(\mathbb{C})$. □

Definition: Sei A eine beliebige K -Algebra, ein A -Modul M heißt halbeinfach, wenn $M = M_1 \oplus \dots \oplus M_k$ eine direkte Summe einfacher A -Moduln ist.

Satz 8.1.3 Sei M ein halbeinfacher A -Modul und $U \subseteq M$ ein Untermodul. Dann gibt es einen halbeinfachen Untermodul $V \subseteq M$ mit $U \oplus V = M$.

Beweis: Sei $M = \bigoplus_{i \in I} M_i$ mit einfachen Moduln M_i . Wir wählen eine maximale Teilmenge $J \subseteq I$, so daß die Summe $U + \bigoplus_{j \in J} M_j$ direkt ist. Dann ist

$$M_i \subseteq U + \bigoplus_{j \in J} M_j \text{ für alle } i \in I,$$

andernfalls wäre

$$M_i \cap (U + \bigoplus_{j \in J} M_j) \subset M_i, \text{ also}$$

$$M_i \cap (U + \bigoplus_{j \in J} M_j) = \{0\},$$

daß heißt, die Menge J wäre nicht maximal. Damit gilt

$$M = \bigoplus M_i = U + \bigoplus_{j \in J} M_j$$

und $V = \bigoplus_{j \in J} M_j$ ist ein halbeinfacher A -Modul. \square

Wir können auch die Umkehrung beweisen:

Satz 8.1.4 *Sei M ein A -Modul, $\dim_K(M) < \infty$, jeder Untermodul von M sei ein direkter Summand. Dann ist M halbeinfach.*

Beweis: Wir wählen einen minimalen Untermodul $M_1 \subseteq M$, dieser ist ein einfacher Modul und es gilt

$$M = M_1 \oplus U_1$$

für einen Untermodul U_1 . Sei $M_2 \subseteq U_1$ ein minimaler (also einfacher) Untermodul, dann ist $M_1 \cap M_2 = \{0\}$, also ist deren Summe direkt und es gibt einen Untermodul U_2 mit $M = M_1 \oplus M_2 \oplus U_2$, usw.

Nach endlich vielen Schritten haben wir M in einfache Untermoduln zerlegt. \square

Satz 8.1.5 *Untermoduln und Faktormoduln halbeinfacher Moduln sind halbeinfach.*

Beweis: Sei M halbeinfach und $U \subseteq M$ ein Untermodul. Dann existiert ein halbeinfacher Modul $V \subseteq M$ mit $U \oplus V = M$, und zu V existiert ein halbeinfacher Modul $W \subseteq M$ mit $V \oplus W = M$. Nun ist

$$M/V = (U \oplus V)/V \cong U \cong (V \oplus W)/V \cong W,$$

also ist U halbeinfach, und

$$M/U \cong V$$

ist ebenfalls halbeinfach. \square

Satz 8.1.6 *Sei $M = M_1 \oplus \dots \oplus M_n$ mit einfachen Untermoduln M_i und $U \subseteq M$ sei ein weiterer einfacher Untermodul. Es gelte $U \cong M_i$ für $i = 1, \dots, r$ und $M_j \not\cong U$ für $j > r$. Dann gilt $U \subseteq M_1 \oplus \dots \oplus M_r$.*

Beweis: Es sei $p_i : M \rightarrow M_i$ die Projektion; für $u \in U$ gilt dann

$$u = \sum p_i(u)$$

mit $p_i(u) \in M_i$. Die Einschränkung $p_i | U : U \rightarrow M_i$ ist A -linear, wenn also $p_i(u) \neq 0$ ist, so ist $p_i | U$ ein Isomorphismus zwischen U und M_i , also $p_j | U = 0$ für $j > r$. \square

Wir betrachten nun eine spezielle Klasse von K -Algebren, die sich als direkte Summen von Linksidealen darstellen lassen. Dazu erweitern wir unsere Kenntnisse über Modulhomomorphismen. Die Menge aller A -linearen Abbildungen $f : M \rightarrow N$ zwischen zwei R -Moduln bezeichnen wir mit $\text{Hom}_A(M, N)$.

Lemma 8.1.2 $\text{Hom}_A(A, M) \cong M$.

Beweis: Sei $f : A \rightarrow M$ eine A -lineare Abbildung, dann gilt

$$f(r) = f(r \cdot 1) = rf(1) \text{ für alle } r \in A,$$

also ist f durch $f(1)$ eindeutig bestimmt. \square

Folgerung 8.1.2 $\text{End}_A(A) \cong A$. \square

Wir hatten früher gesehen, daß zu direkten Zerlegungen

$$M = M_1 \oplus \dots \oplus M_n$$

Endomorphismen $p_1, \dots, p_n \in \text{End}(M)$ gehören, für die $p_i \circ p_j = p_i \delta_{ij}$ galt (man nennt solche Elemente „orthogonale Idempotenten“) und es galt $M_i = p_i(M)$.

Wenn wir nun eine Zerlegung einer Algebra $A = L_1 \oplus \dots \oplus L_n$ in Linksideale vornehmen, so entspricht dem die Existenz orthogonaler Idempotenter e_i in $\text{End}_R(R) = R$ und es gilt $L_i = Ae_i$.

Definition: Eine K -Algebra A heißt halbeinfach, wenn $A = L_1 \oplus \dots \oplus L_n$ eine direkte Summe minimaler Linksideale ist.

Eine Algebra ist also genau dann halbeinfach, wenn sie als linker Modul über sich selbst halbeinfach ist. Wir zeigen, daß dies keine besondere Bevorzugung der linken Seite bedeutet.

Satz 8.1.7 *Eine Algebra A ist genau dann eine direkte Summe minimaler Linksideale, wenn sie eine direkte Summe minimaler Rechtsideale ist.*

Beweis: Sei $A = \bigoplus L_i = \bigoplus Ae_i$ mit $e_i e_j = e_i \delta_{ij}$ und $\sum e_i = 1$ eine Zerlegung von A in eine direkte Summe minimaler Linksideale. Dann ist

$$A = \bigoplus e_i A$$

eine direkte Zerlegung in Rechtsideale. Wir zeigen, daß diese minimal sind.

Sei $a \in e_i Ae_i \subseteq e_i A$, also $a = e_i x e_i$ für ein x , dann ist $e_i a = e_i^2 x e_i = e_i x e_i = a$, also $aA \subseteq e_i A$. Analog gilt $Aa \subseteq Ae_i$, also $Aa = Ae_i$. Wir betrachten die Abbildung $f : Ae_i \rightarrow Aa$ mit $f(xe_i) = xe_i a = xa$, dies ist ein Homomorphismus linker A -Moduln, also ein Isomorphismus. Weiter sei $A = Aa \oplus U$, wir definieren $g : A \rightarrow A$ durch $g(xa + u) = f^{-1}(xa) = xe_i$, dies ist ein Homomorphismus linker A -Moduln, also $g(y) = y \cdot g(1)$, wir setzen $g(1) = b$. Dann ist $e_i = g(b) = a \cdot b$, also $e_i \in aA \subseteq e_i A$, also $aA = e_i A$. \square

Die Moduln halbeinfacher Algebren sind besonders einfach:

Satz 8.1.8 *Sei M ein (endlich erzeugter) Modul über der halbeinfachen Algebra A . Dann ist M halbeinfach.*

Beweis: Es gibt einen freien A -Modul F , so daß $M \cong F/U$ mit $U \subseteq F$ gilt. Mit A ist auch F halbeinfach, also auch dessen Faktormodul M . \square

Satz 8.1.9 *Jeder einfache Modul einer halbeinfachen Algebra A ist isomorph zu einem minimalen Linksideal von A .*

Beweis: Sei M ein einfacher A -Modul, dann gilt $M \cong A/L$ für ein maximales Linksideal L . Es gibt ein Linksideal $H \subseteq A$ mit $A = L \oplus H$ und da L maximal ist, kann H nur minimal sein. Nun gilt aber $M \cong A/L \cong H$. \square

Wie sehen eigentlich die zweiseitigen Ideale einer halbeinfachen Algebra aus?

Satz 8.1.10 *Sei $A = \bigoplus L_i$ eine halbeinfache Algebra, die L_i seien minimale Linksideale und $L_1 \cong \dots \cong L_r$ und $L_1 \not\cong L_i$ für $i > r$. Dann ist $I = L_1 \oplus \dots \oplus L_r$ ein minimales zweiseitiges Ideal von A .*

Beweis: Sei $a \in A$; zu $i \leq r$ betrachten wir die Abbildung $f_a : L_i \rightarrow A$ mit $f_a(l) = la$, die Abbildung f_a ist A -linear, also ist $\text{Im}(f_a) = \{0\}$ oder $\text{Im}(f_a) \cong L_i$. Folglich ist $\text{Im}(f_a) = L_i a \subseteq L_1 \oplus \dots \oplus L_r$, also ist I ein zweiseitiges Ideal.

Wir zeigen noch: jedes in I enthaltene minimale Linksideal (etwa L_1) erzeugt I als Ideal.

Sei $p_1 : A \rightarrow L_1$ die Projektion und $f : L_1 \rightarrow L_j$ ein A -Isomorphismus. Wir betrachten $f \circ p_1 : A \rightarrow L_j$, es ist $f \circ p_1(a) = af \circ p_1(1)$ für $a \in A$. Sei $l \in L_1$, dann gilt $p_1(l) = l$, also $f \circ p_1(l) = f(l) = lf \circ p_1(1)$, also ist $L_j = f(L_1) = L_1 \cdot f \circ p_1(1) \subseteq L_1 A$, also ist I ein minimales Ideal. \square

Folgerung 8.1.3 *Sei A eine halbeinfache Algebra, dann ist A eine direkte Summe minimaler Ideale: $A = I_1 \oplus \dots \oplus I_s$, jedes I_i ist eine direkte Summe paarweise isomorpher minimaler Linksideale.* \square

Satz 8.1.11 *Sei $A = I_1 \oplus \dots \oplus I_s$ mit minimalen Idealen I_i . Dann gilt $I_i \cdot I_j = \{0\}$ für $i \neq j$ und jedes I_i ist eine einfache Algebra.*

Beweis: Für $i \neq j$ gilt $I_i \cdot I_j \subseteq I_i \cap I_j = \{0\}$. Sei $1 = e_1 + \dots + e_s$ mit $e_i \in I_i$, dann ist e_i das neutrale Element von I_i . Sei $J \subseteq I_1$ ein Ideal von I_1 , also $I_1 J I_1 = J$, dann ist $A J A = A I_1 J I_1 A = I_1 J I_1 = J$, da $A I_1 = I_1 A = I_1$ ist. Also ist J ein Ideal von A , also $J = \{0\}$ oder $J = I_1$. \square

Sei nun $G = \{g_1, \dots, g_n\}$ eine endliche Gruppe. Mit

$$KG = \left\{ \sum r_i g_i \mid r_i \in K \right\}$$

bezeichnen wir die Menge aller formaler Linearkombinationen der Elemente der Gruppe G , dies ist ein n -dimensionaler Vektorraum. Wir führen eine Multiplikation ein, die durch die Multiplikation in G induziert wird:

$$\left(\sum r_i g_i \right) \left(\sum s_j g_j \right) = \sum r_i s_j (g_i g_j),$$

diese erfüllt offenbar das Assoziativgesetz, die Körperelemente kommutieren mit allen Elementen und das Einselement von G ist das neutrale Element. Die Menge KG ist also eine K -Algebra, sie heißt die Gruppenalgebra der Gruppe G .

Wir bemerken, daß KG isomorph zu Halbgruppenalgebra K^G ist.

Satz 8.1.12 (Maschke) *Wenn die Zahl $|G|$ in K invertierbar ist, so ist KG eine halbeinfache Algebra.*

Beweis: Wir zeigen viel mehr: Jeder Untermodul eines KG -Moduls ist ein direkter Summand.

Sei M ein KG -Modul und $U \subseteq M$ ein Untermodul. Speziell ist U ein Unterraum von M und es gibt einen Unterraum V von M mit $M = U \oplus V$, also gibt es eine K -lineare Abbildung $p: M \rightarrow M$ mit $p^2 = p$ und $p(M) = U$, nämlich die Projektion auf U . Wir konstruieren nun eine KG -lineare Abbildung:

$$q(x) = \frac{1}{|G|} \sum g_i^{-1} p(g_i x) \text{ für } x \in M.$$

Zunächst gilt $\text{Im}(q) \subseteq \text{Im}(p)$, wenn $u \in U$ ist, so gilt

$$q(u) = \frac{1}{|G|} \sum g_i^{-1} p(g_i u) = \frac{1}{|G|} \sum g_i^{-1} g_i u = u,$$

denn wegen $g_i u \in U$ ist $p(g_i u) = g_i u$, also gilt $\text{Im}(q) = \text{Im}(p)$. Sei nun $h \in G$, dann ist

$$h^{-1} q(hx) = \frac{1}{|G|} \sum h^{-1} g_i^{-1} p(g_i h x) = q(x),$$

denn mit g_i durchläuft auch $g_i h$ die Gruppe G , d.h.

$$q(hx) = h q(x),$$

also ist q ein KG -Homomorphismus. Schließlich ist

$$q \circ q(x) = q\left(\frac{1}{|G|} \sum g_i^{-1} p(g_i x)\right) = \frac{1}{|G|} \sum g_i^{-1} q \circ p(g_i x),$$

wegen $p(g_i x) \in U$ ist dies gleich $\frac{1}{|G|} \sum g_i^{-1} p(g_i x) = q(x)$, also ist q idempotent und damit $M = \text{Im}(q) \oplus \text{Ker}(q)$, somit ist U ein direkter Summand von M . \square

Wir wollen das an einem einfachen, aber nichttrivialen Beispiel vorführen:

Sei $G = \{e, a, a^2\}$, $a^3 = e$ die zyklische Gruppe der Ordnung 3, dann wird $M = L(e_1, e_2, e_3)$ mittels

$$a \cdot e_1 = e_2$$

$$a \cdot e_2 = e_3$$

$$a \cdot e_3 = e_1$$

ein $\mathbb{R}G$ -Modul: Wir können die Operation von a^k auf e_i durch

$$a^k \cdot e_i = e_{(i+k-1) \pmod{3}+1}$$

darstellen, dann ist

$$a^k(a^l \cdot e_i) = a^k \cdot e_{(i+l-1) \pmod{3}+1} = e_{((i+l+k-1) \pmod{3})+1}$$

und

$$a^{k+l} \cdot e_i = e_{(i+k+l-1) \pmod{3}+1},$$

die Operation ist also assoziativ.

Die Darstellungsmatrizen von a bzw. A^2 sind

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, A^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Der Vektorraum $U_1 = L(e_1 + e_2 + e_3)$ ist offenbar ein $\mathbb{R}G$ -Untermodul, durch

$$P_1 = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

wird eine Projektion auf U_1 gegeben. Die obige Konstruktion

$$Q_1 = \frac{1}{3}(P_1 + A^{-1}P_1A + A^{-2}P_1A^2)$$

ergibt aber nichts Neues, da $P_1 - 1$ sehr homogen aussieht: Eine Zeilen- und die inverse Spaltenoperation an P_1 prallen ab.

Betrachten wir einen anderen $\mathbb{R}G$ -Untermodul $U_2 = \text{Ker}P_1 = L(e_1 - e_2, e_1 - e_3)$: eine Projektion auf U_2 wird z.B. durch

$$P_2 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

gegeben. Die Konstruktion

$$Q_2 = \frac{1}{3}(P_2 + A^{-1}P_2A + A^{-2}P_2A^2)$$

liefert

$$Q_2 = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

Wird dadurch tatsächlich ein $\mathbb{R}G$ -Homomorphismus gegeben? Ja: $Q_2 = E_3 - P_1$!

Berechnung

Ein Rechtsideal R in einer halbeinfachen Algebra A ist ein direkter Summand, also ist $R = eA$ für ein idempotentes Element $e \in R$. Wir wollen ein solches Element berechnen: Es sei x_1, \dots, x_n eine K -Basis von A und a_{ijk} seien die zugehörigen Strukturkonstanten. Weiter sei c_1, \dots, c_k eine K -Basis von R , etwa $c_j = \sum_l k_{jl}x_l$. Wir suchen ein Idempotent $e = \sum_i e_i c_i \in R$, dies ist dann ein linkes Einselement in R , also gilt $e \cdot c_j = c_j$ für alle j , d.h.

$$e \cdot c_j = \sum_i \sum_l e_i k_{il} x_l \cdot c_j = \sum_i \sum_l e_i k_{il} x_l \sum_m k_{jm} x_m$$

$$= \sum_i \sum_l e_i k_{il} \sum_m k_{jm} \sum_p a_{lmp} x_p = c_j = \sum_p k_{jp} x_p,$$

durch Koeffizientenvergleich ergibt sich ein lineares Gleichungssystem mit kn Gleichungen für die Unbekannten e_i, \dots, e_k .

Beispiel:

Die Matrix $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ erzeugt in M_3 ein 6-dimensionales Rechtsideal. Ein entsprechendes Idempotent ist $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 2 & 0 \end{pmatrix}$. Es wurde mit der folgenden Methode berechnet:

```

/** die Spalten von K erzeugen das Rechtsideal R = (c_1, ..., c_k);
    gesucht ist ein Idempotent e = \sum e_i c_i mit R = eA;
    die letzte Spalte des Ergebnisses ist eine spezielle Loesung */
public static QM idpt(QA a, QM K)
{
    int i, j, k = K.n, l, m, n = a.dim, p, z = 0; Q g;
    QM b = new QM(k*n, k+1), h;
    for (j = 1; j <= k; j++)
        for (p = 1; p <= n; p++)
            {
                z++; // aktuelle Zeile
                for (i = 1; i <= k; i++)
                    {
                        g = new Q(0);
                        for (l = 1; l <= n; l++)
                            {
                                h = new QM(n, 1);
                                for (m = 1; m <= n; m++)
                                    h.mat[l][1] =
                                        Q.add(h.mat[l][1], Q.mult(K.mat[m][j], a.st[l][m][p]));
                                g = Q.add(g, Q.mult(K.mat[l][i], h.mat[l][1]));
                            }
                        b.mat[z][i] = g;
                    }
                b.mat[z][k+1] = K.mat[p][j];
            }
    QM.GAUSS(b);
    return e = QM.loesung(b);
}

public static void probe_idpt()
{
    int d, i, j, k, n; QA a = vMatrix(3); k = 1; n = a.dim;

```

```

Q[] e = new Q[n+1];
for (i = 1; i <= n; i++) e[i] = new Q(i);
QM s = rechtsIdeal(a, e); d = s.n; // Dim des Rechtsideals
QM b = idpt(a, s);
QM v = new QM(d, 1);
for (i = 1; i <= d; i++) v.mat[i][1] = b.mat[i][b.n];
QM.write(v); QM p = QM.mult(s, v); QM.write(p);
}

```

8.2 Darstellungen endlicher Gruppen

In diesem Abschnitt werden wir uns mit den Elementen der Darstellungstheorie beschäftigen. Das Ziel kann grob so umrissen werden, daß „abstrakte“ Gruppen als „konkrete“ Gruppen von Matrizen beschrieben werden sollen.

Definition: Sei G eine Gruppe und V ein Vektorraum über dem Körper K , dann heißt ein Gruppenhomomorphismus

$$\rho : G \longrightarrow GL(V)$$

eine Darstellung von G in V ; ein Gruppenhomomorphismus

$$R : G \longrightarrow GL(n, K)$$

heißt Matrixdarstellung von G über K vom Grade n .

Für eine Darstellung ρ von G gilt also für $g, h \in G$: $\rho(gh) = \rho(g) \circ \rho(h)$ sowie $\rho(g^{-1}) = \rho(g)^{-1}$, $\rho(1) = id$.

Sei nun $t : V \xrightarrow{\sim} W$ ein Isomorphismus von Vektorräumen und ρ eine Darstellung von G , dann haben wir also Automorphismen $\rho(g) : V \longrightarrow V$. Wir konstruieren nun eine neue Darstellung $\rho' : G \longrightarrow GL(W)$ wie folgt: $\rho'(g) : W \longrightarrow W$ sei durch

$$\rho'(g) = t \circ \rho(g) \circ t^{-1}$$

gegeben, d.h. wir haben ein kommutatives Diagramm

$$\begin{array}{ccc}
 V & \xrightarrow{\rho(g)} & V \\
 t \downarrow & & \downarrow t \\
 W & \xrightarrow{\rho'(g)} & W
 \end{array}$$

und es gilt $\rho'(gh) = t\rho(g)t^{-1}t\rho(h)t^{-1} = \rho'(g)\rho'(h)$, also ist ρ' eine Darstellung von G in W .

Definition: Zwei Darstellungen $\rho : G \longrightarrow GL(V)$, $\rho' : G \longrightarrow GL(W)$ heißen äquivalent, wenn ein Isomorphismus $t : V \xrightarrow{\sim} W$ existiert, so daß

$$\rho'(g) = t \circ \rho(g) \circ t^{-1} \text{ für alle } g \in G$$

gilt.

Zu einer Darstellung ρ von G erhalten wir eine Matrixdarstellung, indem wir eine Basis B in V wählen und jedem Element $g \in G$ die Darstellungsmatrix des Automorphismus $\rho(g)$ zuordnen:

$$R(g) = A_{BB}(\rho(g)).$$

Wenn B' eine andere Basis und X die Basiswechselmatrix ist, so erhalten wir eine neue Matrixdarstellung R' , die aber wegen $R'(g) = X^{-1}R(g)X$ für alle $g \in G$ zu R äquivalent ist.

Beispiele:

1. $\rho_1 : G \longrightarrow GL(K) = K^*$, $\rho_1(g) = 1$ heißt 1-Darstellung von G .
2. X sei eine Menge, auf der die Gruppe G operiere, d.h. es gibt eine Abbildung $G \times X \longrightarrow X$, $(g, x) \mapsto g \cdot x$ mit $(gh)x = g(hx)$, $1x = x$. Dann ist $V = \{f : X \longrightarrow K\}$ ein Vektorraum und die Abbildung $\rho : G \longrightarrow GL(V)$ mit

$$(\rho(g)(f))(x) = f(g^{-1}x)$$

ist eine Darstellung, denn

$$(\rho(gh)(f))(x) = f(h^{-1}g^{-1}x) = (\rho(h)(f))(g^{-1}x) = \rho(g)(\rho(h)(f))(x).$$

3. Sei $G = \{1 = g_1, g_2, \dots, g_n\}$, $V = L(b_1, \dots, b_n)$ ein n -dimensionaler Vektorraum, wir setzen

$$\rho(g_i)(b_j) = b_m, \text{ falls } g_i g_j = g_m$$

ist. Wegen $g_i G = G$ ist $\rho(g_i)$ invertierbar, daß ρ ein Homomorphismus ist, rechnet man leicht nach. Diese Darstellung heißt die reguläre Darstellung von G .

4. Sei $K_4 = \{1, g, h, gh\}$ die Kleinsche Vierergruppe. Sei R ihre reguläre Matrixdarstellung, dann haben wir

$$\begin{aligned} R(g)(b_1) &= b_2, \text{ da } g \cdot 1 = g \\ R(g)(b_2) &= b_1, \text{ da } g^2 = 1 \\ R(g)(b_3) &= b_4, \\ R(g)(b_4) &= b_3. \end{aligned}$$

also

$$R(g) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

und analog für die anderen Elemente.

5. Sei $G = C_n = \langle a \rangle$, $a^n = 1$ die zyklische Gruppe der Ordnung n , für deren reguläre Darstellung gilt

$$\begin{aligned} \rho(a)(a^i) &= a^{i+1}, \quad i = 1, \dots, n-1 \\ \rho(a)(a^{n-1}) &= 1 \end{aligned}$$

also

$$R(a) = \begin{pmatrix} 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & \\ & & \dots & \\ 0 & \dots & 1 & 0 \end{pmatrix}.$$

6. Sei wieder $G = C_n$ und $A \in GL(V)$ mit $A^n = E$, wir setzen $\rho(a^i) = A^i$, dies ist eine Darstellung. Wenn speziell $z \in \mathbb{C}$ eine n -te Einheitswurzel ist, so erhalten wir mit $\rho(a^i) = z^i$ eine Darstellung vom Grad 1. Wenn z_1, \dots, z_m n -te Einheitswurzeln sind, so ist durch

$$T(a) = \begin{pmatrix} z_1 & 0 & \dots & 0 \\ & & \dots & \\ 0 & & \dots & z_m \end{pmatrix}$$

eine Darstellung vom Grad m gegeben.

7. Wenn $m = n$ und die z_i paarweise verschieden sind, so sind die Darstellungen unter 5. und 6. äquivalent: Sei

$$S = \begin{pmatrix} z_1 & \dots & z_n \\ z_1^2 & \dots & z_n^2 \\ \dots & \dots & \dots \\ 1 & \dots & 1 \end{pmatrix}$$

die Vandermondsche Matrix, dann gilt

$$R(A)^T S = S T(A).$$

8. Sei $G \subset S_n$ eine Untergruppe, $p \in G$ eine Permutation von $\{1, \dots, n\}$ und $\{x_1, \dots, x_n\}$ eine Basis von V . Wir definieren $\rho(p)(x_i) = x_{p(i)}$, dies ist eine Darstellung von G . Zum Beispiel kann die Kleinsche Vierergruppe als Gruppe von Permutationen dargestellt werden:

$$K_4 = \{1 = (1), g = (12)(34), h = (3)(24), gh = (14)(23)\}.$$

Dann ist

$$R(1), R(g) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, R(h) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, R(gh) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

eine zu K_4 isomorphe Matrixgruppe.

Als nächstes wollen wir uns einen Überblick über alle 1-dimensionalen komplexen Darstellungen einer Gruppe verschaffen. Dies sind also Homomorphismen $\rho : G \rightarrow \mathbb{C}^*$ in die multiplikative Gruppe des Körpers \mathbb{C} .

Ein Gruppenelement der Form $g^{-1}h^{-1}gh$ wird als Kommutator bezeichnet. In einer kommutativen Gruppe sind alle Kommutatoren gleich 1. Die von allen Kommutatoren erzeugte sogenannte Kommutatorgruppe

$$G' = \langle g^{-1}h^{-1}gh \mid g, h \in G \rangle$$

stellt also ein Maß für die Abweichung von der Kommutativität dar.

Da \mathbb{C}^* kommutativ ist, gilt $G' \subset \text{Ker}(\rho)$, also wird ein Homomorphismus

$$\bar{\rho} : G/G' \longrightarrow \mathbb{C}^*$$

induziert; ρ ist durch $\bar{\rho}$ eindeutig bestimmt: $\rho(g) = \bar{\rho}(gG')$.

Da die Gruppe G/G' abelsch ist, müssen nur die eindimensionalen Darstellungen abelscher Gruppen behandelt werden. Nach dem Hauptsatz über endliche abelsche Gruppen ist

$$G \cong \mathbb{Z}/p_1^{n_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_m^{n_m}\mathbb{Z},$$

wir wählen Erzeugende a_1, \dots, a_m mit $a_i^{p_i^{n_i}} = 1$ und $p_i^{n_i}$ -te Einheitswurzeln z_i , dann liefert $\rho(a_i) = z_i$ eine 1-dimensionale Darstellung.

Satz 8.2.1 *G sei eine abelsche Gruppe der Ordnung $p_1^{n_1} \cdots p_m^{n_m}$, dann gibt es genau $|G|$ verschiedene 1-dimensionale Darstellungen von G über \mathbb{C} .*

Beweis: Jede Darstellung ist durch die Bilder der Erzeugenden eindeutig bestimmt, es gibt $p_i^{n_i}$ verschiedene $p_i^{n_i}$ -te Einheitswurzeln, insgesamt also $p_1^{n_1} \cdots p_m^{n_m} = |G|$ Wahlmöglichkeiten. \square

Folgerung 8.2.1 *Eine endliche Gruppe G besitzt genau $|G/G'|$ verschiedene 1-dimensionale Darstellungen über \mathbb{C} .*

Seien nun ρ_1, ρ_2 Darstellungen von G in V_1, V_2 , dann können wir die Darstellung

$$\rho_1 \oplus \rho_2 = \rho : G \longrightarrow \text{GL}(V_1 \oplus V_2)$$

mit

$$\rho(g)(v_1 + v_2) = \rho_1(g)(v_1) + \rho_2(g)(v_2)$$

betrachten, sie heißt die direkte Summe der Darstellungen ρ_1, ρ_2 .

Analog definiert man die direkte Summe von Matrixdarstellungen:

$$R_1 \oplus R_2(g) = \begin{pmatrix} R_1(g) & 0 \\ 0 & R_2(g) \end{pmatrix}.$$

Das Tensorprodukt

$$\rho_1 \otimes \rho_2 = \rho : G \longrightarrow \text{GL}(V_1 \otimes V_2)$$

mit

$$\rho_1 \otimes \rho_2(g) = \rho_1(g) \otimes \rho_2(g)$$

ist ebenfalls eine Darstellung, die entsprechenden Darstellungsmatrizen sind die Kroneckerprodukte der gegebenen Darstellungsmatrizen.

Wenn wieder $\rho : G \longrightarrow \text{GL}(V)$ eine Darstellung und KG die Gruppenalgebra ist, so wird V durch $(\sum a_i g_i) \cdot v := \sum a_i \rho(g_i)(v)$ ein linker KG -Modul. Umgekehrt: Wenn V ein KG -Modul ist, so definiert $\rho(g)(v) = g \cdot v$ eine Darstellung von G in V . Wenn

wir KG als linken Modul über sich selbst auffassen, so entspricht dem die reguläre Darstellung von G .

Definition: Sei $\rho : G \rightarrow GL(V)$ eine Darstellung. Ein Unterraum $U \subset V$ heißt ρ -invariant, wenn $\rho(g)(U) \subset U$ für alle $g \in G$ gilt. (In diesem Fall ist U ein KG -Untermodul von V .) Die Darstellung ρ heißt irreduzibel, wenn es außer $\{0\}$ und V keine invarianten Unterräume gibt, sonst reduzibel. (Der KG -Modul V ist im ersten Fall einfach, sonst nicht.)

Aus dem Satz von Maschke folgt:

Folgerung 8.2.2 *Jede endlichdimensionale Darstellung ist eine direkte Summe irreduzibler Darstellungen.* \square

Der folgende Satz erlaubt schon einmal eine Zerlegung eine Darstellung in eine direkte Summe von Unterdarstellungen, denn idempotenten Endomorphismen entsprechen direkte Summanden:

Satz 8.2.2 *Für jede Darstellung $\rho : G \rightarrow GL(V)$ ist*

$$p = \frac{1}{|G|} \sum_{g \in G} \rho(g) : V \rightarrow V$$

ein idempotenter Endomorphismus.

Beweis: Für $h \in G$ ist

$$\rho(h) \circ p = \frac{1}{|G|} \sum_{g \in G} \rho(h) \circ \rho(g) = \frac{1}{|G|} \sum_{g \in G} \rho(hg) = p,$$

also

$$p \circ p = \frac{1}{|G|} \sum_{g \in G} \rho(h) \circ p = \frac{1}{|G|} \sum_{g \in G} p = \frac{|G|}{|G|} p = p.$$

Die Verträglichkeit mit der KG -Operation zeigt man analog. \square

Das Lemma von Schur liest sich im Darstellungszusammenhang folgendermaßen:

Satz 8.2.3 *1. Seien $\rho_i : G \rightarrow GL(V_i)$, $i = 1, 2$ irreduzible Darstellungen und $f : V_1 \rightarrow V_2$ eine lineare Abbildung mit $f \circ \rho_1(g) = \rho_2(g) \circ f$ für alle $g \in G$, dann ist $f = 0$ oder f ist ein Isomorphismus, d.h. ρ_1 und ρ_2 sind äquivalent.*

2. Wenn $V_1 = V_2$ und zusätzlich $K = \mathbb{C}$ ist, so ist $f = z \cdot id$. \square

Mit demselben Trick wie beim Satz von Maschke oder im obigen Satz erhält man Vertauschungsrelationen, für die man das Schursche Lemma anwenden kann:

Sei $h : V_1 \rightarrow V_2$ beliebig und $a \in G$, wir setzen

$$h_a = \sum_g \rho_2(g) h \rho_1(a g^{-1}), \quad (*)$$

dies ist eine lineare Abbildung und es gilt

$$\rho_2(u)h_a\rho_1(u^{-1}) = \sum_g \rho_2(ug)h\rho_1(a(ug)^{-1}) = h_a,$$

also

$$\rho_1(u)h_a = h_a\rho_1(u).$$

Folgerung 8.2.3 Wenn ρ_1, ρ_2 irreduzibel und nicht äquivalent sind, so ist $h_a = 0$. Wenn ρ_1, ρ_2 äquivalent sind, so gibt es ein $z_a \in \mathbb{C}$ mit $h_a = z_a \text{id}$. Dabei gilt $z_e = \frac{|G|}{\dim V} Sp(h)$.

Beweis: Wir bilden in (*) die Spur. □

8.3 Charaktere

Definition: Sei $\rho : G \rightarrow GL(V)$ eine Darstellung, wir konstruieren dazu die Funktion $\chi : G \rightarrow K$ mit $\chi(g) = Sp(\rho(g))$, sie heißt der zu ρ gehörige Charakter der Gruppe G . Wenn ρ eine irreduzible Darstellung ist, so nennt man χ einen irreduziblen Charakter.

Beispiel: Sei ρ die reguläre Darstellung, der entspricht als Modul die Gruppenalgebra. Wie operiert $\rho(g_i)$ auf $\{g_1 = 1, g_2, \dots, g_n\}$?

$$\rho(g_i)(g_j) = \begin{cases} g_i g_j \neq g_j & \text{für } i \neq 1 \\ g_i & \text{für } i = 1 \end{cases},$$

d.h. $\rho(g_1)$ hat als Darstellungsmatrix die Einheitsmatrix, also $\chi(1) = Sp(\rho(1)) = n$; für $i \neq 1$ hat die Darstellungsmatrix von $\rho(g_i)$ nur Nullen auf der Diagonalen, also $\chi(g_i) = Sp(\rho(g_i)) = 0$.

Satz 8.3.1 1. Die Charaktere zu äquivalenten Darstellungen sind gleich.

2. $\chi(gh) = \chi(hg)$ für $g, h \in G$.

3. Ein Charakter ist konstant auf den Klassen konjugierter Elemente.

4. Seien χ_i die zu ρ_i gehörigen Charaktere, dann gehört zur Darstellung $\rho_1 \oplus \rho_2$ der Charakter $\chi_1 + \chi_2$.

Beweis: 1. Es gilt $\rho_1(g) = t^{-1}\rho_2(g)t$ für einen Isomorphismus t , also sind die Spuren von $\rho_1(g)$ und $\rho_2(g)$ gleich.

2. gilt wegen $Sp(AB) = Sp(BA)$.

3. $\chi(h^{-1}gh) = Sp(\rho(h^{-1}gh)) = Sp(\rho(h)^{-1}\rho(g)\rho(h)) = Sp(\rho(g)) = \chi(g)$.

4. Die zugehörigen Darstellungsmatrizen sind $\begin{pmatrix} R_1(g) & 0 \\ 0 & R_2(g) \end{pmatrix}$, deren Spur ist gleich $Sp(R_1(g)) + Sp(R_2(g))$. □

Satz 8.3.2 Jeder Charakter von G ist eine Summe irreduzibler Charaktere.

Beweis: Jede Darstellung ist eine direkte Summe irreduzibler Darstellungen. \square

Definition: Seien $\phi, \psi : G \longrightarrow K$ beliebige Abbildungen; wir definieren ein Skalarprodukt

$$\langle \phi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \phi(g)\psi(g^{-1}).$$

Satz 8.3.3 *Das Skalarprodukt $\langle \cdot, \cdot \rangle$ ist symmetrisch und nicht ausgeartet.*

Beweis: $\langle \phi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \phi(g)\psi(g^{-1}) = \frac{1}{|G|} \sum_{h^{-1} \in G} \phi(h^{-1})\psi(h) = \langle \psi, \phi \rangle$.

Sei $\langle \phi, \psi \rangle = 0$ für alle Abbildungen $\psi : G \longrightarrow K$. Für ein beliebiges $h \in G$ definieren wir

$$\psi_h(g) = \begin{cases} 1, & g = h^{-1} \\ 0 & \text{sonst} \end{cases},$$

dann gilt

$$0 = \langle \phi, \psi_h \rangle = \frac{1}{|G|} \phi(h),$$

also $\phi = 0$. \square

Beispiel: Es sei χ_{reg} der Charakter zur regulären Darstellung von G , also

$$\chi_{reg}(g) = \begin{cases} |G|, & g = 1 \\ 0 & \text{sonst} \end{cases},$$

χ sei der Charakter zu $\rho : G \longrightarrow GL(V)$, dann ist

$$\langle \chi, \chi_{reg} \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_{reg}(g)\chi(g^{-1}) = \frac{1}{|G|} |G| \chi(1) = \dim(V).$$

Da die Spur des Kroneckerprodukts $A \otimes B$ zweier Matrizen gleich $Sp(A)Sp(B)$ ist, gilt entsprechendes für den Charakter zum Tensorprodukt: Seien $\rho_i : G \longrightarrow GL(V_i)$ Darstellungen und χ_i die zugehörigen Charaktere, weiter sei χ der Charakter zu $\rho_1 \otimes \rho_2$, dann gilt

$$\chi(g) = \chi_1(g)\chi_2(g).$$

Wir wollen nun eine Darstellung von G im $\text{Hom}(V_1, V_2)$ konstruieren:

$$\rho(g) : V_1^* \otimes V_2 \longrightarrow V_1^* \otimes V_2,$$

$$\rho(g) = \rho_1(g)^{-1} \otimes \rho_2(g),$$

führt zu

$$\rho(g) : \text{Hom}(V_1, V_2) \longrightarrow \text{Hom}(V_1, V_2),$$

$$\rho(g)(f) = \rho_2(g) \circ f \circ \rho_1(g).$$

Wegen $Sp(A^T) = Sp(A)$ gehört zu ρ der Charakter χ mit

$$\chi(g) = \chi_2(g)\chi_1(g^{-1}).$$

Zu dieser Darstellung bilden wir den idempotenten Endomorphismus

$$p = \frac{1}{|G|} \sum_g \rho(g) = \frac{1}{|G|} \sum_g \rho_2(g) \otimes \rho_1(g^{-1})^*,$$

es gilt

$$rg(p) = Sp(p) = \frac{1}{|G|} \sum_g \chi_2(g) \chi_1(g^{-1}) = \langle \chi_2, \chi_1 \rangle.$$

Nun können wir die nützlichen Orthogonalitätsrelationen für irreduzible Charaktere beweisen:

Satz 8.3.4 *Seien $\rho_i : G \rightarrow GL(V)$ irreduzible Darstellungen über \mathbb{C} und χ_1, χ_2 die zugehörigen Charaktere, dann gilt*

$$\langle \chi_1, \chi_2 \rangle = \begin{cases} 1, & \text{wenn } \rho_1, \rho_2 \text{ äquivalent,} \\ 0, & \text{wenn } \rho_1, \rho_2 \text{ nicht äquivalent.} \end{cases}$$

Beweis: Wir betrachten die obige Abbildung p , es gilt $\langle \chi_1, \chi_2 \rangle = rg(p) = \dim(\text{Im}(p))$. Sei $h \in \text{Hom}(V_1, V_2)$, dann ist (vergleiche die obige Konstruktion (*))

$$p(h) = \frac{1}{|G|} \sum_g \rho_2(g) h \rho_1(g^{-1}) = \frac{1}{|G|} h_e = 0,$$

falls ρ_1 und ρ_2 nicht äquivalent sind. Wenn aber ρ_1, ρ_2 äquivalent sind, so sei oBdA $\rho_1 = \rho_2$, $V_1 = V_2$, dann ist

$$p(h) = \frac{1}{|G|} h_e = z_e id$$

für alle h , d.h. $\text{Im}(p) = L(id)$ ist eindimensional, also $\langle \chi_1, \chi_1 \rangle = 1$. □

Satz 8.3.5 *Sei $\rho : G \rightarrow GL(V)$ eine Darstellung mit dem Charakter χ , $\phi : G \rightarrow GL(U)$ sei eine irreduzible Darstellung mit dem Charakter ψ , dann ist $\langle \chi, \psi \rangle$ gleich der Anzahl der irreduziblen Summanden von ρ , die äquivalent zu ϕ sind.*

Beweis: Wir zerlegen den KG -Modul V in einfache Untermoduln und fassen zueinander isomorphe zusammen:

$$V = \bigoplus n_i V_i,$$

die zugehörige Darstellung ist $\rho = \sum n_i \rho_i$ mit dem Charakter $\chi = \sum n_i \chi_i$. OBdA sei ψ äquivalent zu ρ_1 , dann folgt aus den Orthogonalitätsrelationen

$$\langle \chi, \psi \rangle = \sum n_i \langle \chi_i, \psi \rangle = n_1. \square$$

Wesentlich einfacher als im Satz von Krull-Schmidt erhalten wir die

Folgerung 8.3.1 *Eine Zerlegung einer Darstellung in irreduzible Darstellungen ist bis auf die Reihenfolge der Summanden eindeutig bestimmt.*

Beweis: Die Vielfachheiten ergeben sich als Skalarprodukte. □

Folgerung 8.3.2 Seien χ_1, \dots, χ_k alle irreduziblen Charaktere von G und n_1, \dots, n_k die Dimensionen der zugehörigen Darstellungsräume V_i , dann gilt $\chi_{reg} = \sum n_i \chi_i$ und $|G| = \sum n_i^2$.

Beweis: $n_i = \langle \chi_{reg}, \chi_i \rangle = \dim(V_i)$, $\chi_{reg}(1) = |G| = \sum n_i \chi_i(1) = \sum n_i^2$, da $\chi(1) = \dim(V)$. \square

Satz 8.3.6 Seien $\rho, \rho' : G \rightarrow GL(V_i)$ Darstellungen mit Charakteren χ, χ' . Die Darstellungen ρ, ρ' sind genau dann äquivalent, wenn $\chi = \chi'$.

Beweis: Sei $\rho = \sum n_i \rho_i$, $\rho' = \sum l_i \rho_i$ mit irreduziblen ρ_i . Dann folgt aus $\chi = \chi'$ sofort $n_i = \langle \chi, \chi_i \rangle = \langle \chi', \chi_i \rangle = l_i$. \square

Satz 8.3.7 Eine Darstellung $\rho : G \rightarrow GL(V)$ mit dem Charakter χ ist genau dann irreduzibel, wenn $\langle \chi, \chi \rangle = 1$ ist.

Beweis: Wenn ρ irreduzibel ist, so folgt $\langle \chi, \chi \rangle$ aus den Orthogonalitätsrelationen. Sei umgekehrt $\rho = \sum m_i \rho_i$, dann gilt $\langle \chi, \chi \rangle = \sum m_i^2 = 1$ genau dann, wenn ein m_i gleich 1 und die restlichen gleich Null sind, also wenn ρ irreduzibel ist. \square

Definition: Eine Abbildung $f : G \rightarrow K$ heißt Klassenfunktion, wenn $f(gh) = f(hg)$ für alle $g, h \in G$ gilt, d.h. f ist auf den Klassen konjugierter Elemente konstant.

Charaktere sind Klassenfunktionen.

Sei $f : G \rightarrow K$ eine Klassenfunktion und $\rho : G \rightarrow GL(V)$ eine irreduzible Darstellung, wir betrachten

$$T(f, \rho) = \sum_g f(g^{-1}) \rho(g) \in \text{End}(V).$$

Es gilt

$$\begin{aligned} \rho(h) T(f, \rho) \rho(h^{-1}) &= \sum f(g^{-1}) \rho(hgh^{-1}) \\ &= \sum f(h^{-1}g^{-1}h) \rho(hgh^{-1}) \\ &= \sum f(u^{-1}) \rho(u) \\ &= T(f, \rho), \end{aligned}$$

also ist $T(f, \rho)$ mit allen $\rho(h)$ vertauschbar, aus dem Lemma von Schur folgt also

$$T(f, \rho) = z_{f, \rho} \text{id}, \quad z_{f, \rho} \in \mathbb{C}.$$

Wir bilden die Spur:

$$Sp(T(f, \rho)) = \dim(V) z_{f, \rho}$$

oder

$$\frac{\dim(V)}{|G|} z_{f, \rho} = \frac{1}{|G|} Sp(T(f, \rho)) = \frac{1}{|G|} \sum f(g^{-1}) \chi(g) = \langle f, \chi \rangle,$$

also

$$z_{f, \rho} = \frac{|G|}{\dim(V)} \langle f, \chi \rangle.$$

Satz 8.3.8 *Jede Klassenfunktion ist eine Linearkombination irreduzibler Charaktere.*

Beweis: Andernfalls gibt es eine Klassenfunktion f mit $\langle f, \chi_i \rangle = 0$ für alle irreduziblen Charaktere χ_i , dann ist die obige Konstante z_{f, ρ_i} für alle i gleich Null. Also ist $T(f, \rho_i) = 0$. Sei $\rho_{reg} = \sum n_i \rho_i$, dann ist

$$T(f, \rho_{reg}) = \sum_i n_i \sum_g f(g^{-1} \rho_i(g)) = \sum_i n_i T(f, \rho_i) = 0.$$

Wir wenden dies auf 1 an und beachten $\rho_{reg}(g)(1) = g$:

$$\sum f(g^{-1}) \rho_{reg}(1) = \sum f(g^{-1}) g = 0,$$

also $f(g^{-1}) = 0$, d.h. $f = 0$. □

Satz 8.3.9 *Die irreduziblen Charaktere bilden eine Orthonormalbasis des Vektorraums der Klassenfunktionen.* □

Folgerung 8.3.3 *Die Anzahl der irreduziblen Charaktere ist gleich der Anzahl der Konjugationsklassen von G .*

Beweis: Die Dimension des Raums der Klassenfunktionen ist gleich der Zahl der Konjugationsklassen. □

Satz 8.3.10 *Sei G eine abelsche Gruppe, dann ist jede irreduzible Darstellung eindimensional, d.h. jeder irreduzible Charakter $\chi_i : G \rightarrow \mathbb{C}^*$ ist ein Homomorphismus.*

Beweis: Sei $|G| = k$, alle Konjugationsklassen sind einelementig, also gibt es k Stück. Weiter ist $k = |G| = \sum_{i=1}^k n_i^2$, also $n_i = 1$ für $i = 1, \dots, k$. □

Ohne Beweis teilen wir abschließend mit, daß die Dimensionen der irreduziblen Darstellungen Teiler der Gruppenordnung sind.

8.4 Die diskrete Fourier-Transformation

Die Gruppe G sei kommutativ, dann ist KG kommutativ. Wir betrachten den Spezialfall $K = \mathbb{C}$. Nach dem Satz von Maschke ist $\mathbb{C}G$ eine halbeinfache Algebra, wir haben gesehen, das halbeinfache Algebren sich als eine direkte Summe einfacher Algebren darstellen lassen:

$$\mathbb{C}G = A_1 \oplus \dots \oplus A_m.$$

Nach dem Satz von Wedderburn ist jede einfache \mathbb{C} -Algebra isomorph zu einer Matrixalgebra:

$$A_i \cong M_{n_i n_i}(\mathbb{C}).$$

Für $n_i > 1$ ist diese Algebra nichtkommutativ, also müssen alle $n_i = 1$ sein, also $A_i \cong \mathbb{C}$. Insgesamt erhalten wir

$$\mathbb{C}G \cong \mathbb{C} \times \dots \times \mathbb{C}.$$

Es sei $C_n = \{1, g, g^2, \dots, g^{n-1}\}$ mit $g^n = 1$ die zyklische Gruppe mit n Elementen. Dann ist $\mathbb{C}C_n$ isomorph zur Faktor algebra $\mathbb{C}[x]/(x^n - 1)$ des Polynomrings $\mathbb{C}[x]$. Die Multiplikation in $\mathbb{C}[x]/(x^n - 1)$ ist relativ komplex, wenn das Produkt zweier Polynome zu berechnen ist, so sind etwa n^2 Multiplikationen von Körperelementen durchzuführen. Nach den obigen Resultaten ist aber

$$\mathbb{C}C_n \cong \mathbb{C} \times \dots \times \mathbb{C}$$

und die Multiplikation in dieser Algebra geschieht komponentenweise, für eine Multiplikation von Algebra-Elementen benötigt man also nur n Multiplikationen von Körperelementen. Es wäre schön, wenn wir den obigen Isomorphismus explizit kennen würden.

Dies ist möglich. Wir bestimmen nämlich die Idempotenten e_i mit $A_i = \mathbb{C}C_n e_i$.

Lemma 8.4.1 *Sei G eine kommutative Gruppe und $f : G \rightarrow \mathbb{C} \setminus \{0\}$ ein Homomorphismus von multiplikativen Gruppen, dann ist*

$$\frac{1}{|G|} \sum f(g_i) g_i \in \mathbb{C}G$$

idempotent.

Beweis:

$$\begin{aligned} \frac{1}{|G|} \sum f(g_i) g_i \cdot \frac{1}{|G|} \sum f(g_j) g_j &= \frac{1}{|G| |G|} \sum f(g_i g_j) g_i g_j \\ &= \frac{1}{|G| |G|} \sum f(g_i) g_i \cdot |G| = \frac{1}{|G|} \sum f(g_i) g_i. \square \end{aligned}$$

Wenn speziell $G = C_n = \langle g \rangle$ ist, so ist jeder Homomorphismus $f : C_n \rightarrow \mathbb{C} \setminus \{0\}$ durch $f(g) \in \mathbb{C}$ bestimmt, wegen $g^n = 1$ muß $f(g)^n = 1$ sein, d.h. $f(g)$ ist eine n -te Einheitswurzel. Sei also ω eine primitive n -te Einheitswurzel, dann gibt es die folgenden n Homomorphismen $f_i : C_n \rightarrow \mathbb{C} \setminus \{0\}$ mit

$$f_i(g) = \omega^i.$$

Also haben wir n Idempotenten in der Gruppenalgebra:

$$e_i = \frac{1}{n} (1 + \omega^i g + \omega^{2i} g^2 + \dots + \omega^{(n-1)i} g^{n-1}), \quad i = 0, \dots, n-1.$$

Also hat der Isomorphismus

$$F^{-1} : \bigoplus \mathbb{C}C_n e_i \rightarrow \mathbb{C}C_n$$

bezüglich der Basen $\{e_1, \dots, e_n\}$ und $\{1, g, \dots, g^{n-1}\}$ die Darstellungsmatrix

$$\frac{1}{n} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \omega & \dots & \omega^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{n-1} & \dots & \omega^{(n-1)^2} \end{bmatrix}$$

der (eigentlich interessante) inverse Isomorphismus F hat die zu dieser inverse Darstellungsmatrix, diese hat die Gestalt

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \xi & \dots & \xi^{n-1} \\ \dots & \dots & \dots & \dots \\ 1 & \xi^{n-1} & \dots & \xi^{(n-1)^2} \end{bmatrix}$$

mit $\xi = \omega^{-1}$.

Kapitel 9

Zerlegung endlichdimensionaler Algebren

Von nun an sei A eine beliebige endlichdimensionale K -Algebra.

Definition: Ein Element $a \in A$ heißt nilpotent, wenn ein $n \in \mathbb{N}$ mit $a^n = 0$ existiert. Ein Linksideal $L \subseteq A$ heißt nilpotent, wenn $L^n = \{0\}$ für ein $n \in \mathbb{N}$, d.h. wenn alle Produkte $l_1 \dots l_n$ von Elementen aus L Null sind.

Beispiele: In $K[x]/(x^n)$ ist das von \bar{x} erzeugte Ideal nilpotent.

In der Algebra $\begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$ der Dreiecksmatrizen ist das Linksideal $\begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix}$ nilpotent.

Lemma 9.0.2 *In einer halbeinfachen Algebra A gibt es keine von $\{0\}$ verschiedenen nilpotenten Linksideale.*

Beweis: Ein Linksideal $L \subseteq A$ ist ein direkter Summand, wird also von einem Idempotent e erzeugt, von dem keine Potenz verschwindet. \square

Deshalb traten bisher keine nilpotenten Linksideale auf.

Lemma 9.0.3 *Die Summe zweier nilpotenter Linksideale ist nilpotent.*

Beweis: Seien $N, L \subseteq A$ Linksideale und $N^p = L^q = \{0\}$. Wir betrachten ein Produkt von $p + q + 1$ Faktoren aus $N + L$:

$$(n_1 + l_1) \dots (n_{p+q+1} + l_{p+q+1}),$$

wenn dies ausmultipliziert wird, so enthalte ein Summand r Faktoren aus N und $p + q + 1 - r$ Faktoren aus L . Wenn $r \geq p + 1$ ist, so liegt er in $N^p(N + L) = \{0\}$, wenn $r < p + 1$ ist, so ist $p + q + 1 - r > q$, also liegt der Summand in $L^q(N + L) = \{0\}$. \square

In einer endlichdimensionalen Algebra ist die Summe unendlich vieler Linksideale stets gleich der Summe von nur endlich vielen dieser Linksideale, denn sonst könnte man eine unendliche echt aufsteigende Kette von Linksidealen konstruieren. Somit erhalten wir den

Satz 9.0.1 *Die Summe aller nilpotenter Linksideale von A ist nilpotent.* \square

Satz 9.0.2 Sei $a \in A$, dann sind die folgenden Bedingungen äquivalent:

1. Es ist $aM = \{0\}$ für alle einfachen linken A -Moduln M .
2. Das Element a liegt im Durchschnitt aller maximalen Linksideale von A .
3. Für alle $b \in A$ besitzt $1 - ba$ ein Inverses.
4. Für alle $b \in A$ ist ba nilpotent.
5. Das Element a liegt in einem nilpotenten Linksideal.
6. Es ist $Ma = \{0\}$ für alle einfachen rechten A -Moduln M .
7. Das Element a liegt im Durchschnitt aller maximalen Rechtsideale von A .
8. Für alle $b \in A$ besitzt $1 - ab$ ein Inverses.
9. Für alle $b \in A$ ist ab nilpotent.
10. Das Element a liegt in einem nilpotenten Rechtsideal.

Beweis: Wir zeigen $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 5 \Rightarrow 1$ und $4 \Leftrightarrow 9$, daraus folgt der Rest.

$1 \Rightarrow 2$: Sei L ein maximales Linksideal und $a \notin L$, dann ist der A -Modul A/L einfach und $a \cdot (1 + L) = a + L \neq 0 + L$, also ist $a \cdot (A/L) \neq \{0\}$, ein Widerspruch.

$2 \Rightarrow 3$: Wir zeigen zuerst, daß $1 - ba$ ein Linksinverses besitzt. Das Element ba liegt in jedem maximalen Linksideal, also liegt $1 - ba$ in keinem maximalen Linksideal, denn aus $ba \in L$ und $1 - ba \in L$ folgt $1 \in L$, also $L = A$. Folglich ist $A(1 - ba) = A$, also gibt es ein $c \in A$ mit $c(1 - ba) = 1$. Aus dem soeben bewiesenen folgt, daß auch $1 + cba = c$ ein Linksinverses d besitzt: $dc = 1$, also hat c ein Links- und ein Rechtsinverses, somit stimmen beide überein: $d = 1 - ba$.

$3 \Rightarrow 4$: Sei $L = Aa$, es ist

$$L \supset L^2 \supset L^3 \supset \dots \supset L^n = L^{n+1}$$

und wir nehmen an, das $L^n \neq \{0\}$ wäre. Sei dann $N \subseteq L$ ein Linksideal, das minimal mit der Eigenschaft $L^n \cdot N \neq \{0\}$ ist, also gibt es ein $x \in N$ mit $L^n x \neq \{0\}$. Wir betrachten das Linksideal $L^n x$: es ist

$$L^n L^n x = L^n x \neq \{0\},$$

also $L^n x = N$. Also gibt es ein $y \in L^n$ mit $yx = x$ und es gilt $y \in L = Aa$, also $y = ba$ für ein $b \in A$. Folglich besitzt $1 - y$ ein Inverses c . Nun folgt $x = 1 \cdot x = c(1 - y)x = c(x - yx) = 0$, ein Widerspruch, also ist das Linksideal Aa nilpotent.

$4 \Rightarrow 5$: Nach Voraussetzung besteht $L = Aa$ aus nilpotenten Elementen, sei wieder $L^n = L^{n+1} \neq \{0\}$, wie oben erhalten wir Elemente $0 \neq x, y \in L$ mit $yx = x$, daraus folgt $y^n x = x$ für alle n , aber für großes n ist $y^n = 0$, ein Widerspruch.

$5 \Rightarrow 1$: Sei M ein einfacher A -Modul, dann gilt entweder $AaM = \{0\}$ oder $AaM = M$, denn AaM ist ein Untermodul von M . Im ersten Fall folgt $aM = \{0\}$, im zweiten $(Aa)^n M = M$ für alle n , wegen der Nilpotenz von Aa ist dies ein Widerspruch.

$4 \Leftrightarrow 9$: Wenn $(ba)^n = 0$ ist, so ist ebenfalls $(ab)^{n+1} = a(ba)^n b = 0$. □

Satz 9.0.3 Die Menge J aller Elemente $a \in A$, die den Bedingungen des vorigen Satzes genügen, ist ein Ideal von A .

Beweis: Nach 2. ist J der Durchschnitt aller maximaler Linksideale, also selbst ein Linksideal, nach 7. ist J auch ein Rechtsideal. □

Dieses Ideal J wird als Jacobson-Radikal bezeichnet.

Folgerung 9.0.1 J ist das eindeutig bestimmte maximale nilpotente (Links-, Rechts-) Ideal von A . Wenn A halbeinfach ist, so ist $J = \{0\}$. \square

Satz 9.0.4 Das Radikal von A/J ist Null.

Beweis: Sei $N/J \subseteq A/J$ ein nilpotentes Linksideal, dabei ist $J \subseteq N \subseteq A$. Aus $(N/J)^n = J/J = \{\bar{0}\}$ folgt $N^n \subseteq J$ und aus $J^m = \{0\}$ folgt $N^{nm} = \{0\}$, also ist N nilpotent, also $N \subseteq J$, d.h. N/J ist Null. \square

Die Umkehrung des folgenden Satzes haben wir oben gesehen.

Satz 9.0.5 Wenn $J = \{0\}$ ist, so ist A halbeinfach.

Beweis: Wir zeigen zuerst, daß jedes minimale Linksideal $L \subseteq A$ ein idempotentes Element enthält.

Es ist $L^2 \subseteq L$, also $L^2 = \{0\}$ oder $L^2 = L$, wobei der erste Fall nicht eintreten kann, weil es wegen $J = \{0\}$ keine nilpotenten Linksideale gibt. Also gibt es ein $a \in L$ mit $La \neq \{0\}$, also $La = L$. Wir betrachten $N = \{b \in L \mid ba = 0\}$, dies ist ein Linksideal, das in L enthalten und von L verschieden ist, also muß $N = \{0\}$ sein, d.h. aus $ba = 0$ folgt $b = 0$. Wegen $La = L$ gibt es ein $e \in L$ mit $ea = a$, dann ist auch $e^2a = a$, d.h. $(e^2 - e)a = 0$, also $e^2 - e = 0$, also ist e idempotent.

Nun zerlegen wir A schrittweise in eine direkte Summe minimaler Linksideale.

Das Linksideal Ae ist nicht Null und in L enthalten, also $L = Ae$. Sei nun $a \in A$ beliebig, dann ist $a = ae + (a - ae)$, dabei ist der erste Summand ein Element von L und die Elemente der Form $a - ae$ bilden ein Linksideal N , folglich ist $A = L + N$.

Sei $b \in L \cap N$, dann ist einerseits $b = b_1e \in L = Ae$, also $be = b_1e^2 = b_1e = b$, andererseits ist $b = b_2 - b_2e \in N$, also $be = b_2e - b_2e^2 = b_2e - b_2e = 0$, also ist $L \cap N = \{0\}$, d.h. $A = L \oplus N$. Nun wählen wir ein minimales Linksideal, das in N enthalten ist und spalten es als direkten Summanden ab. Nach endlich vielen Schritten sind wir fertig. \square

Folgerung 9.0.2 A/J ist halbeinfach. \square

Satz 9.0.6 Ein A -Modul M ist genau dann halbeinfach, wenn $J \cdot M = \{0\}$ ist.

Beweis: Sei $M = \bigoplus M_i$ eine direkte Summe einfacher Moduln, dann gilt $J \cdot M_i = \{0\}$ nach Definition des Radikals.

Sei umgekehrt $J \cdot M = \{0\}$, dann wird M wie folgt ein A/J -Modul:

$$(a + J) \cdot m = am,$$

dies ist wegen $Jm = \{0\}$ wohldefiniert. Also ist M als A/J -Modul halbeinfach, es gibt einfache A/J -Moduln M_i mit $M = \bigoplus M_i$. Jeder A/J -Modul ist aber auch ein A -Modul, also ist M halbeinfach als A -Modul. \square

Wenn die Algebra A durch ihre Strukturkonstanten gegeben ist, also als Matrixalgebra dargestellt ist, so kann ihr Radikal J leicht berechnet werden:

Satz 9.0.7 (Satz von Dickson) *Das Radikal der Matrixalgebra A ist gleich*

$$J = \{X \in A \mid \text{Spur}(XY) = 0 \text{ für alle } Y \in A\}.$$

Beweis: Sei $X \in J$ und $Y \in A$, dann ist XY nilpotent, also sind alle Eigenwerte von XY gleich 0, also $\text{Spur}(XY) = 0$.

Sei umgekehrt $X \in A$ und $\text{Spur}(XY) = 0$ für alle $Y \in A$. Sei $f(x) = x^n + a_1x^{n-1} + \dots + a_n$ das charakteristische Polynom von XY und z_1, \dots, z_n dessen Nullstellen. Dann sind die Eigenwerte von $(XY)^r$ die Zahlen z_i^r und die Spur von $(XY)^r$ ist die r -te Potenzsumme s_r der z_i , also auch gleich 0. Nach den Newtonschen Formeln lassen sich die Koeffizienten a_j aus den Potenzsummen berechnen, also sind alle a_j gleich 0, d.h. $f(x) = x^n$ und nach Hamilton-Cayley ist somit XY nilpotent, also $X \in J$. \square

```
import HUMath.Algebra.*;
/** Algebren */
public class QA
{
/** Strukturkonstanten */
    public Q[] [] [] st;
/** Dimension */
    public int dim;

/** {x | Spur(xy) = 0 fuer alle y} steht in den Spalten
der Ergebnismatrix*/
public static QM radikal(QA a)
{
    int dim = a.dim;
    QM am = new QM(dim, dim);
    int i,j,k,l;
    Q aji;
    for (j = 1; j <= dim; j++)
        for (i = 1; i <= dim; i++)
            {
                aji = new Q(0);
                for (l = 1; l <= dim; l++)
                    for (k = 1; k <= dim; k++)
                        aji = Q.add(aji, Q.mult(a.st[i][k][l], a.st[j][l][k]));
                am.mat[j][i] = aji;
            }
    QM.GAUSS(am);
    QM r = QM.nullraum(am);
    return r;
}
```

Da wir über A keine weiteren Voraussetzungen machen, können wir nicht erwarten, daß sich jeder Modul in eine direkte Summe einfacher Untermoduln zerlegen läßt. Aber Zerlegungen wird es schon geben.

Definition: Sei M ein A -Modul. Wenn nichttriviale Untermoduln $U, V \subseteq M$ existieren, so daß $M = U \oplus V$ gilt, so heißt M zerlegbar, andernfalls heißt M unzerlegbar.

Satz 9.0.8 *Jeder (endlichdimensionale) A -Modul M ist eine direkte Summe unzerlegbarer Untermoduln.*

Beweis: Entweder ist M unzerlegbar oder eine direkte Summe von Untermoduln. Diese Summanden sind entweder unzerlegbar oder lassen sich in Summen zerlegen. Und so weiter. \square

Definition: Ein idempotentes Element $e \in A$ heißt primitiv, wenn es keine orthogonalen Idempotenten s, t mit $e = s + t$ gibt.

Da Idempotente zu direkten Summen führen, gilt der

Satz 9.0.9 *Sei $M = \oplus M_i$ und $e_i : M \rightarrow M$ sei die Projektion auf M_i . Die Moduln M_i sind genau dann unzerlegbar, wenn die e_i primitive orthogonale Idempotente in $\text{End}_A(M)$ sind und $\sum e_i = \text{id}_M$ ist.* \square

Folgerung 9.0.3 *Seien $e_1, \dots, e_k \in A$ idempotente Elemente, dann ist äquivalent:*

1. $A = Ae_1 \oplus \dots \oplus Ae_k$ und die Ae_i sind unzerlegbare Linksideale.
2. $A = e_1A \oplus \dots \oplus e_kA$ und die e_iA sind unzerlegbare Rechtsideale.
3. $\sum e_i = 1$ und die e_i sind primitive orthogonale Idempotente. \square

Beispiel: $A = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$, primitive Idempotente sind $e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ und $e_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, sie erzeugen die Linksideale $Ae_1 = \begin{pmatrix} * & 0 \\ 0 & 0 \end{pmatrix}$ und $Ae_2 = \begin{pmatrix} 0 & * \\ 0 & * \end{pmatrix}$, letzteres ist nicht minimal, denn es enthält das Linksideal $\begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix}$, dieses wiederum ist kein direkter Summand von A .

Wir werden nun einen Zusammenhang zwischen den Zerlegungen von A/J in eine direkte Summe minimaler Linksideale und von A in eine Summe unzerlegbarer Linksideale herstellen.

Satz 9.0.10 (Liften von Idempotenten) *Seien $f_i + J \in A/J$ orthogonale Idempotente, dann gibt es orthogonale Idempotente $e_i \in A$ mit $e_i + J = f_i + J$.*

Beweis: Wir setzen $f_1 = a$, dann gilt $a^2 - a \in J$. Wir machen einen Ansatz

$$e = a + x(1 - 2a),$$

wobei $x \in J$ sein wird, also $a + J = e + J$. Dann ist

$$\begin{aligned} e^2 &= a^2 + x^2(1 + 4a^2 - 4a) + 2ax(1 - 2a), \\ e^2 - e &= a^2 + x^2 + 4a^2x^2 - 4ax^2 + 2ax - 4a^2x - a - x + 2ax \\ &= (x^2 - x)(1 + 4(a^2 - a)) + a^2 - a \end{aligned}$$

und dies ist jedenfalls dann gleich Null, wenn

$$x = \frac{1}{2} \left(1 - \frac{1}{\sqrt{1+4n}} \right) = \frac{1}{2} (2n - \binom{4}{2}n^2 + \binom{6}{3}n^3 - \dots),$$

dabei haben wir zur Abkürzung $n = a^2 - a$ gesetzt. Nebenrechnung:

$$\sqrt{\frac{1}{4} - \frac{n}{1+4n}} = \frac{1}{2\sqrt{1+4m}}.$$

Wegen $n = a^2 - a \in J$ ist n nilpotent und die Potenzreihe bricht ab. Also haben wir ein Idempotent e_1 in der Klasse von f_1 gefunden.

Analog sei e_2 eine Liftung von f_2 . Diese Idempotenten sind eventuell nicht orthogonal. Aber es ist

$$e_1 e_2 \equiv f_1 f_2 \equiv 0 \pmod{J},$$

also $e_1 e_2 \in J$, folglich besitzt $1 - e_1 e_2$ ein Inverses (nämlich $1 + e_1 e_2 + (e_1 e_2)^2 + \dots$), wir betrachten nun

$$e_2^* = (1 - e_1 e_2) e_2 (1 - e_1 e_2)^{-1},$$

dann ist $e_2^{*2} = e_2^*$ und $e_2^* \equiv f_2 \pmod{J}$, denn $e_2^*(1 - e_1 e_2) = e_2 - e_1 e_2 = e_2^* - e_2^* e_1 e_2$, also ist $e_2^* - e_2 = e_2^* e_1 e_2 - e_1 e_2 \in J$. Weiterhin gilt

$$e_1 e_2^* = (e_1 e_2 - e_1 e_2) (1 - e_1 e_2)^{-1} = 0.$$

Wir setzen nun $e_2^\# = e_2^*(1 - e_1)$, dann ist $e_1 e_2^\# = e_2^\# e_1 = 0$ und

$$e_2^{\#2} = e_2^*(1 - e_1) e_2^*(1 - e_1) = e_2^*(e_2^* - e_1 e_2^*)(1 - e_1) = e_2^*(1 - e_1) = e_2^\#.$$

Seien nun $e_1, e_2, e_3 \in A$ idempotente Elemente und $e_1 e_2 = e_2 e_1 = 0$ und $e_1 e_3 = e_3 e_1 = 0$. Nun wird e_3 zu $e_3^\#$ geändert, so daß $e_2 e_3^\# = e_3^\# e_2 = 0$ gilt. Wir rechnen nach, ob dann noch $e_1 e_3^\# = e_3^\# e_1 = 0$ gilt: Es ist

$$e_3^\# = (1 - e_2 e_3) e_3 (1 + e_2 e_3 + (e_2 e_3)^2 + \dots + (e_2 e_3)^k) (1 - e_2),$$

also gilt

$$e_1 e_3^\# = e_3^\# e_1 = 0.$$

So kann jede Menge orthogonaler Idempotenten von A/J zu A geliftet werden. \square

Folgerung 9.0.4 Jede direkte Zerlegung von A/J in unzerlegbare (d.h. minimale) Linksideale läßt sich zu einer Zerlegung von A in unzerlegbare Linksideale liften.

Beweis: Sei $A/J = \bigoplus (A/J)(f_i + J)$, wobei die $f_i + J$ primitive orthogonale Idempotenten mit $\sum f_i + J = 1 + J$ sind. Seien e_i orthogonale Liftungen dieser Idempotenten, wir zeigen, daß die e_i primitiv sind und daß $\sum e_i = 1$ gilt.

Andernfalls wäre $e_i = p_i + q_i$ mit $p_i q_i = q_i p_i = 0$, dann ist aber auch $e_i + J = p_i + J + q_i + J$ und $(p_i + J) \cdot (q_i + J) = J$ im Widerspruch zur Primitivität der $f_i + J$.

Falls $\sum e_i \neq 1$ wäre, so ist $1 - \sum e_i$ ein weiteres Idempotent, aber $(1 - \sum e_i) + J = (1 + J) - (1 + J) = J$, also $1 - \sum e_i \in J$, ein Widerspruch. \square

Wir wollen den Zusammenhang zwischen den minimalen Linksidealen von A/J und deren Liftungen noch genauer untersuchen. Dazu sind einige Hilfsmittel nötig.

Satz 9.0.11 *Jedes nichtnilpotente Linksideal $L \subseteq A$ enthält ein Idempotent.*

Beweis: Sei $N \subseteq L$ ein Linksideal, das in der Menge der in L enthaltenen nichtnilpotenten Linksidealen minimal ist. Dann ist $N^2 = N$, also gibt es ein $a \in N$ mit $Na = N$, denn falls $Na \neq N$ für alle $a \in N$ gälte, so wäre N nilpotent). Folglich ist

$$Q = \{n \in N \mid na = 0\} \subset N$$

ein echtes Unterideal, also nilpotent. Nun gibt es ein $c \in N$ mit $ca = a$, damit $c^2a = a$, also $(c^2 - c)a = 0$, d.h. $c^2 - c \in Q$ und wie vorhin können wir ein idempotentes Element $e = c + x(1 - 2c)$ finden. \square

Satz 9.0.12 *Ein unzerlegbares Linksideal $L \subseteq A$ ist genau dann ein direkter Summand, wenn L nicht nilpotent ist.*

Beweis: Wenn L ein direkter Summand ist, so ist es von der Form $L = Ae$ mit idempotentem e , also ist L nicht nilpotent.

Sei umgekehrt L nicht nilpotent, dann enthält es ein Idempotent e . Nun ist Le das Bild der Multiplikation von L mit e , deren Kern ist $Q = \{l \in L \mid le = 0\}$ und aus der Idempotenz von e folgt

$$L = Le \oplus Q,$$

wegen der Unzerlegbarkeit von L muß also $Q = 0$ und $L = Le$ sein. Schließlich ist $Le \subseteq Ae \subseteq L$, also ist $L = Ae$ ein direkter Summand.

Satz 9.0.13 *Ein nichtnilpotentes Linksideal L ist genau dann unzerlegbar, wenn alle echt in L enthaltenen Linksideale nilpotent sind.*

Beweis: Seien Alle $N \subset L$ nilpotent, dann kann L nicht die Summe von Untermoduln sei, es wäre sonst selbst nilpotent.

Sei umgekehrt L unzerlegbar und $N \subseteq L$ minimal unter den nichtnilpotenten Linksidealen. N kann nicht Summe von Untermoduln sei (sonst wäre es nilpotent), also ist N unzerlegbar, folglich ein direkter Summand von A und damit auch ein direkter Summand von L . Also muß $N = L$ gelten, also jeder Untermodul von L ist nilpotent. \square

Satz 9.0.14 *Sei $L = Ae, e^2 = e$, ein unzerlegbares Linksideal, dann ist $Je \subseteq Ae$ das eindeutig bestimmte maximale Unterlinksideal, somit ist Ae/Je ein einfacher A -Modul.*

Beweis: Jeder Linksmodul $N \subset L$ ist nilpotent, also $N \subseteq J \cap L = Je$, also ist Je maximal in L . \square

Wir können nun die einfachen Moduln einer Algebra genauer beschreiben:

Satz 9.0.15 *Sei $A = \bigoplus Ae_i$ mit orthogonalen primitiven Idempotenten e_i . Sei M ein einfacher A -Modul, dann gibt es ein i mit $M \cong Ae_i/Je_i$.*

1. Beweis: Es iat $AM \neq \{0\}$, also $Ae_iM \neq \{0\}$ für ein i , also $Ae_im = M$ für ein $m \in M$. Damit haben wir einen surjektiven Modulhomomorphismus $f : Ae_i \rightarrow M$ mit $f(ae_i) = ae_im$, dessen Kern ist maximal in Ae_i , also gleich Je_i .

2. Beweis: M ist einfach, also $JM = 0$, d.h. M ist ein A/J -Modul und als einfacher Modul isomorph zu einem minimalen Ideal der halbeinfachen Algebra A/J , etwa zu Ae_i/Je_i . \square

Lemma 9.0.4 (Fitting) *Sei M ein unzerlegbarer A -Modul und $f : M \rightarrow M$ ein A -Endomorphismus, dann ist f ein Isomorphismus oder nilpotent.*

Beweis: Wir haben eine absteigende Folge

$$M \supset f(M) \supset f^2(M) \supset \dots \supset f^n(M) = f^{n+1}(M),$$

die sich stabilisiert, also ist die Einschränkung

$$f^n | f^n(M) : f^n(M) \rightarrow f^n(M)$$

surjektiv und folglich auch injektiv (alle Vektorräume sind endlichdimensional). Also gilt $f^n(M) \cap \text{Ker}(f^n) = \{0\}$. Sei nun $m \in M$, dann gilt $f^n(m) = f^{2n}(b)$ für ein $b \in M$ und

$$m = f^n(b) + (m - f^n(b)),$$

der erste Summand liegt in $f^n(M)$, der zweite in $\text{Ker}(f^n)$, also ist $M = f^n(M) \oplus \text{Ker}(f^n)$ eine direkte Summe, daraus folgt $M = f^n(M)$ oder $f^n = 0$. \square

Folgerung 9.0.5 *Wenn M ein unzerlegbarer A -Modul ist, so ist dessen Radikal $J(\text{End}_A(M))$ das eindeutig bestimmte maximale Ideal (d.h. $\text{End}_A(M)$) ist ein lokaler Ring). \square*

Satz 9.0.16 *Seien $e_1, e_2 \in A$ primitive Idempotente, dann gilt $Ae_1/Je_1 \cong Ae_2/Je_2$ genau dann, wenn $Ae_1 \cong Ae_2$.*

Beweis: Sei $f : Ae_1/Je_1 \rightarrow Ae_2/Je_2$ ein Isomorphismus und $f(e_1 + J) = r + J$ mit $r \in Ae_2$. Dann ist $r + J = f(e_1 + J) = f(e_1^2 + J) = e_1 f(e_1 + J) = e_1 r + J$. Es gilt $Ae_1 r \subseteq Ae_2$ und $e_1 r$ ist nicht nilpotent, also ist $Ae_2 = Ae_1 r$, d.h. $g : Ae_1 \rightarrow Ae_2$ mit $g(x) = xr$ ist ein surjektiver Homomorphismus linker A -Moduln. Analog erhalten wir einen Homomorphismus $h : Ae_2 \rightarrow Ae_1$ mit $h(y) = ys$, also ist $g \circ h : Ae_1 \rightarrow Ae_1$ ein surjektiver Endomorphismus, also nicht nilpotent. Nach dem Lemma von Fitting ist $g \circ h$ ein Isomorphismus, also ist h auch ein Isomorphismus.

Wenn umgekehrt $h : Ae_1 \rightarrow Ae_2$ ein Isomorphismus ist, so betrachten wir die Komposition

$$Ae_1 \rightarrow Ae_2 \rightarrow Ae_2/Je_2,$$

deren Kern ist gleich Je_1 . Nach dem Homomorphiesatz folgt $Ae_1/Je_1 \cong Ae_2/Je_2$. \square

Wir hatten gesehen, daß sich jeder Modul in eine direkte Summe unzerlegbarer Untermoduln zerlegen läßt. Wir stellen uns die Frage, ob dies auf mehrere verschiedene Weisen möglich sein kann.

Lemma 9.0.5 *Sei M ein unzerlegbarer A -Modul und für $i = 1, \dots, n$ seien $f_i : M \rightarrow M$ A -lineare Abbildungen, so daß $\sum f_i$ ein Isomorphismus ist. Dann ist zumindest eines der f_i ein Isomorphismus.*

Beweis: Andernfalls sind alle f_i nilpotent, liegen also im Radikal von $\text{End}_A(M)$, dort liegt auch deren Summe, kann also kein Isomorphismus sein. \square

Lemma 9.0.6 Sei $f : M_1 \oplus M_2 \rightarrow N_1 \oplus N_2$ ein Isomorphismus, es sei $f(m_1, 0) = (h(m_1), g(m_1))$ und $h : M_1 \rightarrow N_1$ sei ein Isomorphismus, dann ist $M_2 \cong N_2$.

Beweis: Sei $f(m_1, m_2) = (n_1, n_2)$, wir setzen

$$p(m_1, m_2) = (n_1, n_2 - gh^{-1}(n_1)).$$

Die Abbildung p ist injektiv, denn aus $(n_1, n_2 - gh^{-1}(n_1)) = (0, 0)$ folgt $n_1 = 0 = n_2$ und aus der Injektivität von f folgt $m_1 = m_2 = 0$. Wegen $\dim(M_1 \oplus M_2) = \dim(N_1 \oplus N_2)$ ist p auch ein Isomorphismus. Weiter gilt

$$p(m_1, 0) = (h(m_1), g(m_1) - gh^{-1}(h(m_1))) = (h(m_1), 0),$$

also $M_2 \cong M_1 \oplus M_2 / M_1 \oplus 0 \xrightarrow{p} (M_1 \oplus M_2) / h(M_1) \oplus 0 = N_1 \oplus N_2 / N_1 \oplus 0 \cong N_2$. \square

Nun können wir zeigen, daß es keine wesentlich verschiedenen Zerlegungsmöglichkeiten eines Moduls gibt:

Satz 9.0.17 (Krull/Schmidt/Remak/Wedderburn) Wenn $M_1 \oplus \dots \oplus M_m \cong N_1 \oplus \dots \oplus N_n$ und die M_i und N_j unzerlegbare A -Moduln sind, so ist $m = n$ und bei geeigneter Numerierung $M_i \cong N_i$.

Beweis: Sei $f : M_1 \oplus \dots \oplus M_m \rightarrow N_1 \oplus \dots \oplus N_n$ ein Isomorphismus, wir verknüpfen ihn und sein Inverses mit Einbettungen und Projektionen:

$$g_k : M_k \xrightarrow{i_k} M_1 \oplus \dots \oplus M_m \xrightarrow{f} N_1 \oplus \dots \oplus N_n \xrightarrow{p_1} N_1,$$

$$h_k : N_1 \xrightarrow{j_1} N_1 \oplus \dots \oplus N_n \xrightarrow{f^{-1}} M_1 \oplus \dots \oplus M_m \xrightarrow{q_k} M_k,$$

dann ist

$$\sum g_k h_k = \sum p_1 f i_k q_k f^{-1} j_1 = p_1 f \sum i_k q_k f^{-1} j_1 = p_1 j_1 = id_{N_1},$$

also ist einer der Summanden, etwa $g_1 h_1$, ein Isomorphismus. Also ist g_1 surjektiv, weiter ist $h_1 g_1$ nicht nilpotent, also ein Isomorphismus und damit ist g_1 injektiv, also $M_1 \cong N_1$.

Nun ist $f(m_1, 0, \dots, 0) = (g_1(m_1), \dots)$, also folgt aus dem obigen Lemma, daß $M_2 \oplus \dots \oplus M_m \cong N_2 \oplus \dots \oplus N_n$ gilt. \square

Kapitel 10

Das Jacobson-Radikal

In diesem Abschnitt bezeichne R stets einen Ring, der nicht notwendigerweise eine 1 besitzen muß. Dann können wir eine Eins „adjungieren“: In der additiven Gruppe $R^1 = \mathbb{Z} \oplus R$ führen wir auf natürliche Weise eine Multiplikation ein:

$$(n + r)(m + q) = nm + mr + nq + rq,$$

der erste Summand liegt in \mathbb{Z} , die anderen in R .

Für $x \in R$ betrachten wir folgende Abbildung

$$L(x) : R \longrightarrow R, \quad L(x)(y) = xy.$$

Aus dem Assoziativgesetz $x(yz) = (xy)z$ folgt

$$L(x)(L(y)(z)) = L(xy)(z),$$

also

$$L(x) \circ L(y) = L(xy),$$

und aus dem Distributivgesetz $(x + y)z = xy + xz$ folgt

$$L(x + y)(z) = L(x)(z) + L(y)(z) = (L(x) + L(y))(z),$$

also

$$L(x + y) = L(x) + L(y),$$

somit ist $L : R \longrightarrow \text{End}_R(R)$ ein Ringhomomorphismus.

Wenn R eine 1 besitzt, so ist $\text{End}(R) \cong R$, dann entspricht L der identischen Abbildung.

Definition: Ein Element $x \in R$ heißt quasi-invertierbar mit dem Quasi-Inversen y , wenn $1 - x \in R^1$ invertierbar mit dem Inversen $1 + y$ ist.

Bemerkung: Sei $1 - x \in R^1$ invertierbar mit einem Inversen $m + z$, dann gilt $(1 - x)(m + z) = m - mx + z - xz = 1$, also muß $m = 1$ sein, d.h. das Inverse hat stets die Form $1 + y$.

Lemma 10.0.7 *Wenn R eine Eins besitzt, dann sind x und $L(x)$ gleichzeitig invertierbar.*

Satz 10.0.18 Für $x \in R$ ist äquivalent:

- a) x ist quasi-invertierbar,
- b) es gibt ein $y \in R$ mit $y - x = xy = yx$,
- c) $id - L(x) \in \text{End}R$ ist invertierbar.

Das Quasi-Inverse von x ist in diesem Fall gleich $y = (id - L(x))^{-1}(x)$.

Beweis: a) \Leftrightarrow b) folgt aus $(1 - x)(1 + y) = 1 - x + y - xy = 1$.

a) \Leftrightarrow c): $1 - x \in R^1$ ist genau dann invertierbar, wenn $L(1 - x) \in R^1$ invertierbar ist. Es gilt $L(1 - x)(n + r) = n - nx + r - xr$, dies ist gleich 0 genau dann, wenn $n = 0$ und $r - xy = 0$ ist, also wenn $r \in \text{Ker}(id - L(x))$, also ist $L(1 - x)$ genau dann bijektiv, wenn $id - L(x) \in \text{End}(R)$ bijektiv ist. \square

Beispiele: 1. Wann ist $x \in \mathbf{Z}$ invertierbar? Es muß also $y - x = xy$ sein, d.h. $y(1 - x) = x$. Dies ist einerseits für $x = y = 0$ der Fall, sonst ist $x - 1$ ein Teiler von x , also $x - 1 \leq \sqrt{x}$, hieraus folgt $x^2 - 1 \leq 3x$ und damit $x = 2, y = -2$.

2. Wenn x nilpotent vom Grade n ist, dann gilt

$$(1 - x)(1 + x + x^2 + \cdots + x^{n-1}) = 1,$$

also ist x quasi-invertierbar.

Wir fixieren ein Element $u \in R$ und führen in R eine neue Multiplikation ein:

$$x \circ y = x u y.$$

Wenn wir die Abhängigkeit von u hervorheben wollen, schreiben wir \circ_u . Die Ringaxiome für die neue Multiplikation überprüft man schnell, z.B.

$$(x \circ y) \circ z = x u y u z = x \circ (y + z),$$

$$x \circ (y + z) = x u (y + z) = x u y + x u z = x \circ y + x \circ z.$$

(Das neutrale Element wäre u^{-1} , falls dies existiert.)

Wir bezeichnen diesen Ring mit R_u , er heißt das u -Homotop von R . Die Linksmultiplikation in R_u bezeichnen wir mit $L_u : L_u(x) = x \circ y = x u y$, also

$$L_u(x) = L(x) L(u) = L(xu).$$

Folgerung 10.0.6 Für $x \in R$ sind äquivalent:

- a) x ist in R_u quasi-invertierbar,
- es gibt ein y mit $y - x = x u y = y u x$,
- c) $id - L(xu)$ ist invertierbar.

Das Quasi-Inverse von x ist in diesem Fall gleich $(id - L(xu))^{-1}(x)$. \square

Wir bezeichnen in folgenden das Quasi-Inverse von x in R_u mit $q(x, u)$, falls es existiert, und setzen $B(x, u) = id - L(xu) = id - L_u(x)$.

Satz 10.0.19 (Symmetriesatz) Für $x, y \in R$ sind äquivalent:

- a) $q(x, u)$ existiert,
- b) $q(xu, 1)$ existiert,
- c) $q(u, x)$ existiert (d.h. u ist quasi-invertierbar in R_x),
- d) $q(ux, 1)$ existiert,
- e) $B(x, u)$ ist invertierbar,
- f) $B(u, x)$ ist invertierbar.

Dann gilt $q(x, u) = B(x, u)^{-1}(x)$ und $q(u, x) = uq(x, u)u + u$.

Beweis: a) \Leftrightarrow e) folgt aus der Folgerung, b) \Leftrightarrow e) folgt aus dem obigen Satz.

a) \Leftrightarrow c): Sei $w = q(x, u)$, also $w - x = x \circ_u w = xuw = wux$, dann setzen wir $z = wu + u$ und erhalten

$$z - u = (wu + u) - u = wu = uxu + uxuwu = u \circ_x (u + wu) = u \circ_x z,$$

also $z - u = u \circ_x z$, d.h. $z = q(u, x)$.

Der Rest folgt aus Symmetriegründen. \square

Bemerkung: Seien $a, b \in R^1, x \in R$, dann ist

$$axb = (m + a')x(n + b') = \dots \in R (!).$$

Satz 10.0.20 (Verschiebungssatz) Seien $a, b \in R^1, x, u \in R$, dann existiert $q(axb, u)$ genau dann, wenn $q(x, bua)$ existiert, und dann gilt $q(axb, u) = aq(x, bua)b$.

Beweis: Sei $w = q(x, bua)$, also $w - x = wbuax = xbuaw$, also $awb - axb = (awb)u(axb) = (axb)u(awb)$, d.h. es existiert $q(axb, u) = awb$.

Wenn umgekehrt $q(axb, u)$ existiert, dann existiert wegen der Symmetrie auch $q(u, axb)$, wegen des soeben Bewiesenen folgt die Existenz von $q(bua, x)$ und wegen der Symmetrie folgt wieder die Existenz von $q(x, bua)$. \square

Satz 10.0.21 (Additionssatz) Es existiere $q(x, u)$, dann gilt

$$B(x, u)B(q(x, u), z) = B(x, u + z)$$

und $q(q(x, u), z)$ existiert genau dann, wenn $q(x, u + z)$ existiert; in diesem Fall gilt $q(q(x, u), z) = q(x, u + z)$.

Beweis: Sei $w = q(x, u)$, also $w - x = wux = xuw$, daraus folgt $xuwz = wz - xz$. Weiter gilt

$$\begin{aligned} B(x, u)B(q(x, u), z) &= (id - L(xu))(id - L(xz)) \\ &= id - L(xu) - L(wz) + L(xuwz) \\ &= id - L(x(u + z)) \\ &= B(x, u + z). \end{aligned}$$

$B(x, u)$ ist invertierbar, also ist $B(q(x, u), z)$ genau dann invertierbar, wenn $B(x, u + z)$ invertierbar ist, und dann gilt

$$q(x, u + z) = B(x, u + z)^{-1}(x) = B(w, z)^{-1}B(x, u)^{-1}(x) = B(w, z)^{-1}(w) = q(w, z). \square$$

Satz 10.0.22 $J = \{x \in R \mid x \text{ ist quasi-invertierbar allen Ringen } R_u\}$ ist ein Ideal von R .

Beweis: Seien $a, b \in R^1$, $x \in J$. Dann existiert $q(x, u)$, also existiert auch $q(x, aub)$ für jedes u . Durch Verschiebung erhalten wir: $q(bxa, u)$ existiert für alle u , d.h. $bxa \in J$. Seien $x, y \in J$, wegen der Symmetrie existieren dann $q(u, x)$ und $q(v, y)$ für alle u, v . Wir setzen speziell $v = q(u, x)$, dann existiert $q(q(u, x), y) = q(u, x + y)$ und damit existiert $q(x + y, u)$ für alle u , also ist $x + y \in J$. \square

Definition: $J = \text{Rad}(R)$ heißt das Jacobson-Radikal von R . Wenn $\text{Rad}(R) = 0$ ist, so heißt R semi-primitiv.

Satz 10.0.23

$$\text{Rad}(R/\text{Rad}(R)) = \{0\}$$

Beweis: Wir setzen $\bar{R} = R/\text{Rad}(R)$, sei $\bar{x} \in \text{Rad}(\bar{R})$, also gibt es zu jedem $\bar{u} \in \bar{R}$ ein $\bar{w} \in \bar{R}$ mit $\bar{w} - \bar{x} = \bar{w}\bar{u}\bar{x} = \bar{x}\bar{u}\bar{w}$, also folgt für die Repräsentanten

$$w - x - wux \in \text{Rad}(R).$$

Damit ist $B(w - x - wux, -u)$ invertierbar und

$$\begin{aligned} B(w - x - wux, -u) &= id - L(-wu + xu - xwu) \\ &= (id - L(xu))(id + L(wu)) \\ &= B(x, u)B(w, u), \end{aligned}$$

also ist auch $B(x, u)$ invertierbar und damit $x \in \text{Rad}(R)$, folglich $\bar{x} = \bar{0}$. \square

Definition: Ein (Links-, Rechts-) Ideal von R heißt quasi-invertierbar, wenn jedes Element quasi-invertierbar ist. Es heißt nil, wenn jedes Element nilpotent ist.

Satz 10.0.24 L sei ein quasi-invertierbares Linksideal, dann ist $L \subset \text{Rad}(R)$.

Beweis: Seien $x \in L$, $u \in R$, dann ist $ux \in L$, also existiert $q(ux, 1)$, folglich existiert $q(x, u)$ für alle u , also ist $x \in \text{Rad}(R)$. \square

Folgerung 10.0.7 Jedes nil-Linksideal ist in $\text{Rad}(R)$ enthalten.

Folgerung 10.0.8 $\text{Rad}(R)$ ist das maximale quasi-invertierbare Linksideal von R .

Beweis: Sei $x \in \text{Rad}(R)$, wir zeigen, daß x quasi-invertierbar ist. Zunächst existiert $q(x, x)$, also auch $q(x^2, 1)$, d.h. $1 - x^2 \in R^1$ ist invertierbar. Wegen $1 - x^2 = (1 - x)(1 + x)$ ist auch $(1 - x) \in R^1$ invertierbar, also ist x quasi-invertierbar. \square

Folgerung 10.0.9 $\text{Rad}(R) = \{x \in R \mid xu \text{ ist quasi-invertierbar für alle } u\}$.

Der Beweis folgt aus dem Symmetriesatz. \square

Bemerkung: Sei I ein Links- oder Rechtsideal und $x \in I$ quasi-invertierbar in R_u , d.h. $q(x, u)$ und $q(u, x)$ existieren. Dann liegt $q(x, u) = xq(u, x)x + x$ ebenfalls in I .

Satz 10.0.25 *Sei $I \subset R$ ein Ideal, dann ist $\text{Rad}(I) = I \cap \text{Rad}(R)$.*

Beweis: Sei $x \in I \cap \text{Rad}(R)$, dann ist $w = q(x, u) \in I$, dann gilt $w - x = wux = xuw$ speziell für alle $u \in I$, also ist $x \in \text{Rad}(I)$.

Sei umgekehrt $x \in \text{Rad}(I)$, d.h. zu jedem $y \in I$ gibt es ein $v \in I$ mit $v - x = v y x = x y v$, d.h. $q(x, y)$ existiert für alle $y \in I$. Folglich existiert $q(x, uxu)$ für alle $u \in R$, also ist die folgende Abbildung invertierbar:

$$\begin{aligned} B(x, uxu) &= id - L(xuxu) \\ &= (id - L(xu))(id + L(xu)) \\ &= B(x, u)B(x, -u), \end{aligned}$$

damit ist auch $B(x, u)$ invertierbar, d.h. $q(x, u)$ existiert, also ist $x \in \text{Rad}(R)$. Es folgt $\text{Rad}(I) \subset I \cap \text{Rad}(R)$. \square

Folgerung 10.0.10 *Jedes Ideal eines semi-primitiven Rings ist semi-primitiv.* \square

Satz 10.0.26

$$\text{Rad}(R_u) = \{x \in R \mid uxu \in \text{Rad}(R)\}$$

$$\text{Rad}(R) = \bigcap_{u \in R} \text{Rad}(R_u)$$

Beweis: 1. Für die Homotope von R_u gilt $(R_u)_v = R_{uvu}$, denn $x \circ_{uvu} y = xuvvy = x \circ_u v \circ_u y$.

2.

$$\begin{aligned} x \in \text{Rad}(R_u) &\Leftrightarrow x \text{ ist quasi-invertierbar in } (R_u)_v \text{ für alle } v \\ &\Leftrightarrow x \text{ ist quasi-invertierbar in } R_{uvu} \text{ für alle } v \\ &\Leftrightarrow q(x, uvu) \text{ existiert für alle } v \\ &\Leftrightarrow q(uxu, v) \text{ existiert für alle } v \text{ (Verschiebung)} \\ &\Leftrightarrow uxu \in \text{Rad}(R). \end{aligned}$$

3. $\text{Rad}(R)$ ist ein Ideal, also gilt $u\text{Rad}(R)u \subset \text{Rad}(R)$ für alle u , also gilt $\text{Rad}(R) \subset \text{Rad}(R_u)$ und damit $\text{Rad}(R) \subset \bigcap \text{Rad}(R_u)$. Sei umgekehrt $x \in \text{Rad}(R_u)$ für alle u , dann existiert $q(uxu, y)$ für alle $u, y \in R$, also ist die Abbildung $B(uxu, x) = B(u, x)B(u, -x)$ invertierbar, damit ist auch $B(u, x)$ invertierbar, also ist $x \in \text{Rad}(R)$.

\square

Wir berechnen zum Abschluß das Radikal eines Matrixrings.

Satz 10.0.27

$$\text{Rad}(M_{nn}(R)) = M_{nn}(\text{Rad}(R)).$$

Kapitel 11

Primzahltest und Faktorisierung ganzer Zahlen

Eine Zahl p ist genau dann eine Primzahl, wenn sie keinen Primteiler $\leq \sqrt{p}$ besitzt. Um also große Zahlen auf Primalität zu untersuchen, sind Tabellen kleiner Primzahlen nützlich.

Wenn wir zum Beispiel alle Primzahlen zwischen 100 und 200 suchen, so müssen wir aus diesen Zahlen alle Vielfachen von 2, 3, 5, 7, 11 und 13 wegstreichen, was ganz einfach ist; es bleiben 21 Zahlen übrig.

Lehmer hat 1909 eine Tabelle aller Primzahlen unter 10 Millionen als Buch veröffentlicht, und in einem zweiten Buch sind die kleinsten Primteiler aller Zahlen unter 10 Millionen enthalten, die nicht durch 2, 3, 5 oder 7 teilbar sind. Damit ist eine vollständige Faktorisierung dieser Zahlen möglich. Man sollte aber bedenken, daß ein Computer eine solche Zahl schneller zerlegen kann, als man in einer Tabelle nachschlagen kann.

Für ein Sieb-Programm braucht man nicht viel Speicherplatz: man repräsentiert jede Zahl durch ein Bit (die i -te Zahl durch das i -te Bit), setzt alle Bits auf 0, geht für alle relevanten Primzahlen p in p er Schritten über die Bits und setzt Einsen. Die restlichen Null-Bits repräsentieren Primzahlen.

Wenn aus irgendeinem Grund alle Primzahlen nacheinander durchforstet werden müssen, so sollte man $(p_i - p_{i-1})/2$ tabellieren, hierfür reichen 6 Bit für $p < 10^6$, also braucht man insgesamt 59 KByte, die Primzahldifferenzen für $p < 10^9$ passen in 8 Bit, dafür braucht man insgesamt 51 MByte, und für $p < 10^{12}$ benötigt man 12 Bit (72 GByte).

Die Anzahl $\pi(x)$ der Primzahlen unterhalb einer gegebenen Schranke x berechnet sich nach Lagrange (1830) wie folgt (alle Brüche sind als ihre ganzen Teile zu verstehen):

$$1 + \pi(x) = \pi(\sqrt{x}) + x - \sum_{p_i \leq \sqrt{x}} \frac{x}{p_i} + \sum_{p_i < p_j \leq \sqrt{x}} \frac{x}{p_i p_j} - \dots$$

Hier ist x gleich der Zahlen $\leq x$, $\sum \frac{x}{p_i}$ ist die Zahl der Zahlen $\leq x$ mit dem Primteiler p_i , $\sum \frac{x}{p_i p_j}$ ist die Zahl der durch zwei Primzahlen teilbaren Zahlen usw.

Nach dem Ein- und Ausschlußprinzip wird die Zahl der Vielfachen kleiner Primzahlen weggenommen, die der Vielfachen zweier kleiner Primzahlen hinzugefügt usw.

$$\pi(100) = \pi(10) + 100 - 50 - 33 - 20 - 14 + 16 + 10 + 7 - 0 + 1 = 25$$

Lagrange berechnete so die Zahl der Primzahlen unter 10^6 zu 78.526; eine damals aktuelle Primzahlentabelle enthielt 78.492 Einträge, der richtige Wert ist 78.489.

Meisel hat 1885 $\pi(10^9) = 50.847.478$ berechnet, dies sind 56 Zahlen zu wenig, wie von Lehmer erst 1958 bemerkt wurde.

Ein Primzahltest ist eine Bedingung $P(n) \in \{0, 1\}$ mit $P(n) = 1$ genau dann, wenn n eine Primzahl ist. Demgegenüber ist ein Kompositionstest eine Bedingung K , so daß aus $K(n) = 1$ folgt, daß n zusammengesetzt ist, d.h. wenn n eine Primzahl ist, so gilt stets $K(n) = 0$, es ist aber auch $K(xy) = 0$ möglich.

Satz 11.0.28 (Wilson) *Die natürliche Zahl p ist genau dann eine Primzahl, wenn*

$$(p-1)! \equiv -1 \pmod{p}.$$

Beweis: Sei p prim; dann ist $\mathbb{Z}/p\mathbb{Z}$ ein Körper, also hat die Gleichung $x^2 = 1$ nur die Lösungen ± 1 . Im Produkt $(p-1)!$ kommt mit jeder Restklasse auch ihre Inverse vor, daraus folgt die Behauptung.

Wenn p einen echten Teiler q besitzt, so ist dies ein Teiler von $(p-1)!$, also ein gemeinsamer Teiler von $(p-1)!$ und p , während -1 teilerfremd zu p ist. \square

Dies ist Primzahltest völlig ungeeignet: Die Berechnung von $n!$ für eine 10-stellige Zahl dauert etwa 100 Sekunden.

Der kleine Satz von Fermat besagt: Ist p eine Primzahl und $\text{ggT}(a, p) = 1$, so gilt

$$a^{p-1} \equiv 1 \pmod{p}.$$

Wenn also

$$a^{N-1} \not\equiv 1 \pmod{N}$$

gilt, so ist N zusammengesetzt.

Dieser Test ist schnell: Für eine 10-stellige Zahl hat man nach knapp 1 Sekunde das Ergebnis.

Allerdings kann dieser Test versagen: $341 = 11 \cdot 31$, aber $2^{340} \equiv 1 \pmod{341}$, jedoch wird die Zahl 341 entlarvt, wenn wir $a = 3$ wählen: $3^{340} \equiv 56 \pmod{341}$. Man sagt: 341 ist eine Fermatsche Pseudoprimzahl zur Basis 2.

Von Lehmer (1936) stammt eine Liste aller Pseudoprimzahlen zur Basis 2 zwischen 10^7 (dem Ende damaliger Primzahllisten) und 10^8 , die keinen Faktor ≤ 313 haben. Pomerance, Selfridge und Wagstaff veröffentlichten 1980 eine Liste von 1770 Zahlen unter $25 \cdot 10^9$, die pseudoprim zu allen Basen 2, 3, 5, 7 sind. In diesem Bereich genügen also vier Fermat-Tests als Primzahltest. Man bedenke, daß zur Berechnung von a^n , $n \approx 10^9 = 2^{30}$ jeweils nur 60 Multiplikationen und MOD-Operationen nötig sind.

Leider gibt es Zahlen, die pseudoprim zu jeder Basis sind, diese heißen Carmichael-Zahlen ; die kleinste ist $561 = 3 \cdot 11 \cdot 17$, es gibt etwa 100000 unter 10^{15} .

Der Fermat-Test kann durch weitere Tests ergänzt werden, für die wir im folgenden die theoretischen Grundlagen legen.

Definition: Es sei $ggT(a, n) = 1$, dann heißt a ein quadratischer Rest modulo n , wenn $x^2 \equiv a \pmod{n}$ für ein x gilt, anderenfalls heißt a quadratischer Nichtrest. Wenn p eine ungerade Primzahl ist, so ist das folgende Legendre-Symbol definiert:

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & a \text{ quadratischer Rest} \pmod{p} \\ -1, & a \text{ Nichtrest} \end{cases}$$

Lemma 11.0.8 1. Die Menge der quadratischen Reste modulo n ist eine Untergruppe von $(\mathbf{Z}/n\mathbf{Z})^*$.

2. Wenn $(\mathbf{Z}/n\mathbf{Z})^*$ zyklisch ist (z.B. wenn n eine Primzahl ist), dann gibt es gleichviele Reste wie Nichtreste und das Produkt zweier Nichtreste ist ein Rest.

3. Wenn $a \equiv b \pmod{m}$, so ist $\left(\frac{a}{m}\right) = \left(\frac{b}{m}\right)$.

Beweis: 1. Sei $x^2 \equiv a$, $y^2 \equiv b$, dann ist $(xy)^2 \equiv ab$.

2. Wenn $(\mathbf{Z}/n\mathbf{Z})^* = \langle g \rangle$ ist, dann sind $1, g^2, g^4, \dots$ quadratische Reste und g, g^3, \dots sind quadratische Nichtreste. Schließlich ist das Produkt zweier ungerader Potenzen eine gerade Potenz. \square

Folgerung 11.0.11 $\left(\frac{a}{p}\right) \left(\frac{b}{p}\right) = \left(\frac{ab}{p}\right)$.

Satz 11.0.29 (Euler) Wenn $ggT(a, p) = 1, p \neq 2$ ist, so gilt $\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}$.

Beweis: Sei $\mathbf{Z}/p\mathbf{Z} = \langle g \rangle$. Es sei $g^s \equiv -1$. Dann ist $g^{2s} = 1$, also ist $2s$ ein Vielfaches von $p-1$, denn $ord(g) = p-1$. Also sind für s nur die Werte $0, (p-1)/2, p-1$ möglich. Im ersten und dritten Fall ist aber $g^s = +1$, also folgt $g^{(p-1)/2} \equiv -1$. Sei nun $a = g^t$, dann ist $a^{(p-1)/2} = g^{t(p-1)/2} \equiv (-1)^t$, also ist a genau dann quadratischer Rest, wenn t gerade ist, und genau dann gilt $a^{(p-1)/2} \equiv 1$. \square

Ohne Beweis geben wir das folgende „quadratische Reziprozitätsgesetz“ an:

Satz 11.0.30 Seien p, q ungerade Primzahlen, dann gilt

$$\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right) (-1)^{\frac{p-1}{2} \frac{q-1}{2}}.$$

Wir können dies anwenden, um das Legendre-Symbol schnell zu berechnen: 4567 ist eine Primzahl.

$$\left(\frac{123}{4567}\right) = \left(\frac{3}{4567}\right) \left(\frac{41}{4567}\right) = \left(\frac{4567}{3}\right) (-1)^{2283} \left(\frac{4567}{41}\right) (-1)^{20 \cdot 2283} = \left(\frac{1}{3}\right) (-1) \left(\frac{16}{41}\right) = -1.$$

Das folgende Jacobi-Symbol verallgemeinert das Legendre-Symbol:

Definition: Sei $n = \prod p_i^{a_i}$, wir setzen $\left(\frac{a}{n}\right) = \prod \left(\frac{a}{p_i}\right)^{a_i}$.

Satz 11.0.31 Wenn a, b, n ungerade sind und $\text{ggT}(a, n) = \text{ggT}(b, n) = 1$ ist, so gilt $\left(\frac{a}{n}\right) \left(\frac{b}{n}\right) = \left(\frac{ab}{n}\right)$. \square

Wenn also a ein quadratischer Rest ist, so gilt $\left(\frac{a}{n}\right) = 1$, aber die Umkehrung gilt nicht.

Wir kommen zurück zum Primzahltest.

Wenn N ungerade und $\text{ggT}(a, N) = 1$, aber $a^{\frac{N-1}{2}} \not\equiv \pm 1 \pmod{N}$ ist, so ist N zusammengesetzt, denn nach dem Eulerschen Satz ist dieser Term im Primzahlfall gleich dem Legendre-Symbol.

Falls aber $a^{\frac{N-1}{2}} \equiv \pm 1$ ist, so berechne man das Jacobi-Symbol $\left(\frac{a}{N}\right)$ mit Hilfe des quadratischen Reziprozitätsgesetzes. Wenn nun $a^{\frac{N-1}{2}} \not\equiv \left(\frac{a}{N}\right)$ ist, so ist N zusammengesetzt, denn wenn N eine Primzahl wäre, so wären Jacobi- und Legendre-Symbole gleich und man wendet den Satz von Euler an.

Eine Zahl N mit $a^{\frac{N-1}{2}} \equiv \left(\frac{a}{N}\right)$ heißt Euler-pseudoprim zur Basis a .

Einige Carmichael-Zahlen werden durch Euler entlarvt, z.B. 1729, denn $11^{864} \equiv 1 \pmod{1729}$, aber $\left(\frac{11}{1729}\right) = -1$.

Pinch (1993) hat festgestellt, daß so bekannte Computeralgebrasysteme wie Mathematica, Maple V und Axiom einige bekannte Carmichael-Zahlen als Primzahlen passieren lassen.

Der folgende Satz von Lucas und Lehmer kehrt den Satz von Fermat um:

Satz 11.0.32 Sei $N - 1 = \prod q_i^{a_i}$ die Primzahlzerlegung. Wenn ein a existiert, so daß

$$a^{(N-1)/q_i} \not\equiv 1 \pmod{N} \text{ für alle } i,$$

aber

$$a^{N-1} \equiv 1 \pmod{N},$$

dann ist N eine Primzahl.

Beweis: Wenn N prim ist, so existiert in $\mathbf{Z}/N\mathbf{Z}^*$ ein Element der Ordnung $N - 1$, und sonst nicht, da dann $\varphi(N) < N$ ist. Nach Voraussetzung ist die Ordnung von a ein Teiler von $N - 1$. Jeder echte Teiler von $N - 1$ ist ein Teiler von $(N - 1)/q_i$, diese sind aber nicht gleich der Ordnung von a , also ist die Ordnung von a gleich $N - 1$. \square

Für die Auswahl dieser Zahl a sollte man keine mit $\left(\frac{a}{N}\right) = 1$ nehmen, denn diese können keine Erzeugenden sein.

Schließlich erwähnen wir den Lucas-Primzahltest für die Mersenne-Zahlen $M_n = 2^n - 1$. Diese Zahl ist genau dann prim, wenn für die rekursive Folge

$$v_0 = 4, v_i \equiv v_{i-1}^2 - 2 \pmod{M_n}$$

gilt:

$$v_{n-2} \equiv 0 \pmod{M_n}$$

gilt.

Wir wollen uns nun mit der Faktorisierung von Zahlen beschäftigen. Als Beispiel für die Schwierigkeiten, die hier zu erwarten sind, beginnen wir mit den Fermatzahlen

$$F_n = 2^{2^n} + 1,$$

diese Zahlen sind für kleine Werte von n Primzahlen: 3, 5, 17, 129; $F_4 = 65537$ ist die größte bekannte Primzahl unter den Fermatzahlen. Pépin hat gezeigt: F_n ist genau dann eine Primzahl, wenn

$$3^{\frac{F_n-1}{2}} \equiv -1 \pmod{F_n}$$

ist. Die kleinste Fermatzahl, wo die Zerlegbarkeit unklar ist, ist $F_{24} = 10^5 \text{ Mio}$; die Zahl F_{14} ist seit 1963 als zusammengesetzt bekannt, aber man kennt keinen Faktor.

Die Feststellung, ob eine Zahl eine Primzahl ist, ist relativ schnell zu machen. Faktorisierungen sind schwierig und langwierig. Deshalb sollte einen Faktorisierungsalgorithmus nur auf eine Zahl anwenden, von der man weiß, daß sie zerlegbar ist.

Ein erstes Beispiel der Zerlegung einer wirklich großen Zahl wurde durch Cole (1903) gegeben:

$$2^{67} - 1 \approx 10^9 \cdot 10^{13}.$$

1. Als erstes sollte man Probedivisionen durch kleine Primzahlen machen, die als Tabelle vorliegen oder on-line erzeugt werden.

Besser ist es, nicht nur durch Primzahlen zu teilen, sondern durch alle Zahlen der Form $6k \pm 1$ (sowie durch 2 und 3), dann braucht man in keiner Tabelle nachzuschlagen. Die Zahlen dieser Form erzeugt man sich, indem man $z = 5$ und $d = 2$ als Anfangswerte wählt und dann $z := z + d$, $d := 6 - d$ iteriert.

Die Grenze der Probedivision liegt bei $z > \sqrt{N}$, also wenn $z > N/z$ ist, man braucht also keine Wurzel zu berechnen, während der Quotient sowieso berechnet werden muß. Man achte darauf, daß jeder gefundene Faktor sofort wegdividiert werden muß und daß Primfaktoren auch mehrfach auftreten können.

2. Wir fassen Produkte von Primzahlen blockweise zusammen:

$$p_0 = \prod_{2 \leq p \leq 97} P \approx 10^{37}$$

$$p_1 = \prod_{101 \leq p \leq 199} P \approx 10^{46}$$

...

$$p_9 = \prod_{907 \leq p \leq 997} P \approx 10^{42}$$

Der mit dem Euklidischen Algorithmus berechenbare $ggT(N, p_i)$ teilt N .

3. Fermatsche Methode

Wir versuchen, die Zahl $N = a \cdot b$ in der Form $N = x^2 - y^2 = (x+y)(x-y)$ darzustellen. Dabei ist $x > \sqrt{N}$, wir beginnen also mit $m = \lfloor \sqrt{N} \rfloor + 1$, setzen $z = m^2 - N$ und

überprüfen, ob dies eine Quadratzahl ist. Wenn nicht, so erhöhen wir m um 1, d.h. die nächste zu testende Zahl ist $z + 2m + 1$.

Es sei nebenbei bemerkt, daß man Quadratwurzeln wie folgt berechnen kann (eine Zahl ist dann eine Quadratzahl, wenn sie gleich dem Quadrat ihrer Wurzel ist): zunächst gibt es ein Iterationsverfahren

$$x_n = \frac{x_{n-1} + \frac{a}{x_{n-1}}}{2},$$

oder man verwendet die binomische Formel

$$(1+x)^{\frac{1}{2}} = 1 + \binom{\frac{1}{2}}{1}x + \binom{\frac{1}{2}}{2}x^2 + \dots$$

Schließlich kann man es einer Zahl schnell ansehen, daß sie keine Quadratzahl ist, denn die letzten beiden Dezimalstellen einer Quadratzahl haben die Form

$$uu, g1, g4, 25, u6, g9,$$

wobei u eine ungerade und g eine gerade Zahl bedeutet.

4. Legendre

Wir suchen $x \not\equiv \pm y$ mit $x^2 \equiv y^2 \pmod{N}$. Dann ist

$$x^2 - y^2 \equiv 0 \equiv (x+y)(x-y) \pmod{N},$$

also $(x+y)(x-y) = tN$. Also sind die Primteiler von N auch in $x \pm y$ vorhanden, man berechne also $\text{ggT}(N, x-y)$.

Aber wie soll man solche x, y finden? Die Antwort gibt das quadratisch Sieb von Pomerance (1981), es funktioniert für Zahlen mit bis zu 110 Dezimalstellen.

Wir suchen für fixierte „kleine“ Primzahlen p_i folgende Kongruenzen zu lösen (es sei $n = \sqrt{N}$ und $x_k = (n + k^2) - N$, k klein):

$$x_k^2 \equiv (-1)^{e_{0k}} p_1^{e_{1k}} \dots p_m^{e_{mk}} \pmod{N}.$$

Wenn wir genügend viele gefunden haben, so sind die Vektoren (e_{0k}, \dots, e_{mk}) modulo 2 linear abhängig, also gibt es a_k mit

$$\sum_{k=1}^n a_k (e_{0k}, \dots, e_{mk}) \equiv (0, \dots, 0),$$

also

$$\sum_{k=1}^n a_k (e_{0k}, \dots, e_{mk}) = 2(v_0, \dots, v_m).$$

Wir setzen dann $x = \prod x_k^{a_k}$, $y = (-1)^{v_0} p_1^{v_1} \dots p_m^{v_m}$, es gilt

$$\begin{aligned} x^2 &= \prod (x_k^2)^{a_k} = \left(\prod (-1)^{e_{0k}} p_1^{e_{1k}} \dots p_m^{e_{mk}} \right)^{a_k} \\ &= (-1)^{\sum a_k e_{0k}} p_1^{\sum a_k e_{1k}} \dots p_m^{\sum a_k e_{mk}} \\ &= (-1)^{2v_0} p_1^{2v_1} \dots p_m^{2v_m} \\ &= y^2. \end{aligned}$$

Beispiel: $n = 1909$, $\sqrt{n} \approx 43$, $44^2 = 1936$, also probieren wir:

$$44^2 \equiv 3 \cdot 3 \cdot 3 \pmod{n} \quad (11.1)$$

$$45^2 \equiv 2 \cdot 2 \cdot 29 \pmod{n} \quad (11.2)$$

$$46^2 \equiv 3 \cdot 3 \cdot 23 \pmod{n} \quad (11.3)$$

$$47^2 \equiv 2 \cdot 2 \cdot 3 \cdot 5 \cdot 5 \pmod{n} \quad (11.4)$$

Wenn wir nur die Primzahlen 2, 3, 5 betrachten, so sind die Exponentenvektoren der ersten und der letzten Zahl (0, 3, 0) und (2, 1, 2), also modulo 2 gleich. Damit ist

$$(44 \cdot 47)^2 \equiv 90^2$$

und $\text{ggT}(\underbrace{44 \cdot 47 - 90}_{1978}, 1909) = 23$ ist ein Teiler.

Moderne Faktorisierungsmethoden wie die von Pollard, auf die wir aber hier nicht eingehen, haben ihre Leistungsgrenze bei 120 bis 150 Dezimalstellen.

Daß eine Zahl N durchschnittlich $\ln \ln N$ verschiedene Primfaktoren hat, wollen wir einmal hinnehmen. Wie groß aber sind diese?

Wir ordnen die Primteiler von N der Größe nach:

$$N = P_s(N)P_{s-1}(N) \cdots P_2(N)P_1(N), \quad P_s \leq P_{s-1} \leq \dots \leq P_2 \leq P_1.$$

Dann hat $N/P(N)$ noch $s - 1$ verschiedene Primteiler, also

$$\begin{aligned} s - 1 &\approx \ln \ln \frac{N}{P(N)} = \ln(\ln N - \ln P(N)) \\ &= \ln \ln N + \ln\left(1 - \frac{\ln P(N)}{\ln N}\right) \\ &= s + \ln\left(1 - \frac{\ln P(N)}{\ln N}\right), \end{aligned}$$

also $\ln\left(1 - \frac{\ln P(N)}{\ln N}\right) \approx -1$, d.h. $1 - \frac{\ln P(N)}{\ln N} \approx \frac{1}{e}$, also $\ln P(N) \approx \left(1 - \frac{1}{e}\right) \ln N = 0,632 \ln N$ und somit

$$P(N) \approx N^{0,632}.$$

Die Stellenzahl des größten Primteilers von N ist also etwa 63 % der Stellenzahl von N , die Stellenzahl des zweitkleinsten beträgt 63 % des Rests, also 23 %.

Primzahlen und Kryptographie

Bei der Kryptographie geht es darum, zunächst einer Zeichenkette eine Zahl T zuzuordnen, die mittels einer Verschlüsselungsfunktion f zu einem Code $C = f(t)$ chiffriert wird. Der Code C wird an den Empfänger gesandt, der über die Inverse f^{-1} der Verschlüsselungsfunktion verfügen muß, um den Originaltext lesen zu können.

Die Funktionen f und f^{-1} sind im einfachsten Fall Tabellen. Besser sind Funktionen, die von einem Parameter abhängen (einem Schlüssel), diesen Schlüssel sollte man häufig wechseln.

Ein erfahrener Verschlüßler geht davon aus, daß dem unberechtigten Mithörer seiner Nachricht der Verschlüßlungsalgorithmus bekannt geworden ist, (hoffentlich) nicht aber der jeweilige Schlüssel.

Ein einfacher Schlüssel, der schon von Caesar benutzt wurde, besteht im Verschieben der Buchstaben um eine konstante Zahl: $C = T + k$. Durch Rivest, Shamir, Adleman ist nun Codierungsverfahren mit öffentlichem Schlüssel entwickelt worden, das RSA-Verfahren. Dabei ist

$$C = T^k \pmod{N}, \text{ also } T = \sqrt[k]{N} \pmod{N},$$

die erste Operation ist schnell durchgeführt, die zweite ist schwierig, falls N eine zusammengesetzte Zahl ist.

Wir verfahren wie folgt: Wir suchen Zahlen N, k , so daß ein k' mit $T = C^{k'} \pmod{N}$ existiert, d.h. $a^{kk'} \equiv a \pmod{N}$ für alle a , und es soll schwer sein, eine Zahl s mit $a^{ks} \equiv a \pmod{N}$ zu bestimmen.

Für die Eulersche Funktion φ gilt $a^{\varphi(N)} \equiv 1 \pmod{N}$, falls $\text{ggT}(a, N) = 1$ ist.

Die Carmichael-Funktion λ ist wie folgt definiert:

$$\lambda(1) = \lambda(2) = 1, \lambda(4) = 2$$

$$\lambda(2^r) = 2^{r-2} \text{ für } r > 2$$

$$\lambda(p^r) = p^{r-1} \cdot (p-1) \text{ für } p > 2, r > 0$$

$$\lambda(p_1^{r_1} \cdot \dots \cdot p_n^{r_n}) = \text{kgV}(\lambda(p_1^{r_1}), \dots, \lambda(p_n^{r_n}))$$

$\lambda(n)$ ist das kleinste m mit $a^m \equiv 1 \pmod{n}$ für alle a mit $\text{ggT}(a, n) = 1$.

Der Funktionswert $\lambda(N)$ ist ein Teiler von $\varphi(N)$, und es gilt

$$a^{\lambda(N)} \equiv 1 \pmod{N},$$

für alle a , falls N ein Produkt verschiedener Primzahlen ist. Wenn $p \neq q$ Primzahlen sind, dann gilt

$$\lambda(pq) = \frac{(p-1)(q-1)}{\text{ggT}(p-1, q-1)} = \text{kgV}(p-1, q-1).$$

Dann gilt $a^{kk'} \equiv a \pmod{N}$ genau dann, wenn $kk' \equiv 1 \pmod{\lambda(N)}$. Wir setzen also $N = pq$ und suchen ein m mit

$$m\lambda(N) + 1 = k \cdot k'$$

indem wir am Besten k vorgeben und die Kongruenz $m\lambda(N) + 1 \equiv 0 \pmod{k}$ lösen. Wenn $\lambda(N)$ nicht bekannt ist, kann ein Unbefugter die Zahl k' nicht bestimmen.

Wir bestimmen dann $C = f(T) = T^k \pmod{N}$ und versenden diesen Code. Dabei sollte k nicht zu klein sein, damit $T^k > N$ wird und tatsächlich eine \pmod{N} -Reduktion vorgenommen wird.

Der Empfänger berechnet

$$C^{k'} \equiv T^{kk'} \equiv T^{m\lambda(N)+1} \equiv T \pmod{N}.$$

Die Schlüssel N und k kann man als Zeitungsinserat veröffentlichen. Der unbefugte Entschlüßler muß die Zahl $\lambda(N)$ berechnen, dazu muß er die Zahl N zerlegen, und das schafft er nicht in einer vernünftigen Zeit, wenn p und q groß sind.

Kapitel 12

Boolesche Algebren und Boolesche Funktionen

Definition:

Eine Menge B mit drei Operationen $+ : B \times B \longrightarrow B$, $\cdot : B \times B \longrightarrow B$ und $\bar{} : B \longrightarrow B$ sowie zwei ausgezeichneten Elementen $0, 1 \in B$ heißt Boolesche Algebra, wenn für $a, b, c \in B$ folgende Rechenregeln erfüllt sind:

$$a + (b + c) = (a + b) + c, \quad a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad (\text{Assoziativität})$$

$$a + b = b + a, \quad a \cdot b = b \cdot a \quad (\text{Kommutativität})$$

$$a + a = a, \quad a \cdot a = a \quad (\text{Idempotenz})$$

$$a + (b \cdot c) = (a + b) \cdot (a + c), \quad a \cdot (b + c) = a \cdot b + a \cdot c \quad (\text{Distributivität})$$

$$a + (a \cdot b) = a, \quad a \cdot (a + b) = a \quad (\text{Absorbtion})$$

$$0 + a = a, \quad 0 \cdot a = 0$$

$$1 + a = 1, \quad 1 \cdot a = a$$

$$a + \bar{a} = 1, \quad a \cdot \bar{a} = 0.$$

Manchmal schreibt man anstelle von $+$ auch \vee oder \cup und nennt diese Operation Disjunktion, Vereinigung oder Supremum; für \cdot schreibt man dann \wedge oder \cap und nennt es Konjunktion, Durchschnitt oder Infimum. Die Operation $\bar{}$ heißt Komplementierung oder Negation.

Das einfachste Beispiel einer Booleschen Algebra ist die Algebra $\mathbb{B} = \{0, 1\}$, wo sich die Definition der Rechenoperationen schon aus den obigen Regeln ergibt.

Ein weiteres Beispiel ist die Potenzmenge $P(M) = \{U \mid U \subseteq M\}$ einer Menge M mit Durchschnitt und Vereinigung sowie Komplementärtmengenbildung als Rechenoperationen, da, wie man oben sieht, für die Addition und die Multiplikation genau dieselben Rechenregeln gelten, ist es egal, ob wir den Durchschnitt als Addition oder als Multiplikation auffassen.

Wenn B und C Boolesche Algebren sind, so ist $B \times C$ mit komponentenweise definierten Rechenoperationen ebenfalls eine Boolesche Algebra, insbesondere also auch jede kartesische Potenz B^n von B .

Von nun an bezeichne B stets eine Boolesche Algebra.

Für die Komplementierung gelten die folgenden DeMorganschen Regeln:

Satz 12.0.33 $\overline{x \cdot y} = \bar{x} + \bar{y}$, $\overline{x + y} = \bar{x} \cdot \bar{y}$.

Beweis: Wenn a das Komplement von $x \cdot y$ bezeichnet, so haben wir $a + (x \cdot y) = 1$ und $a \cdot (x \cdot y) = 0$ nachzuweisen:

$$(x \cdot y) + (\bar{x} + \bar{y}) = (x + \bar{x} + \bar{y}) \cdot (y + \bar{x} + \bar{y}) = 1 \cdot 1 = 1,$$

$(x \cdot y) \cdot (\bar{x} + \bar{y}) = (x \cdot y \cdot \bar{x}) + (x \cdot y \cdot \bar{y}) = 0 + 0 = 0$. Der Beweis der anderen Regel verläuft analog. \square

Die soeben benutzte Beweismethode („analog“) ist typisch für die Arbeit mit Booleschen Algebren: Man vertauscht die Rechenoperationen miteinander und wendet die analogen Regeln an; dies nennt man „Dualisierung“.

Lemma 12.0.9 (Kürzungsregel) Für $x, y, z \in B$ gelte (1) $x \cdot y = x \cdot z$ und (2) $x + y = x + z$. Dann folgt $y = z$.

Beweis: Zur ersten Gleichung wird sowohl y als auch z addiert:

$$\begin{aligned} (x \cdot y) + y &= (x + y) \cdot (y + y) = (x + y) \cdot y = y \\ &= (x \cdot z) + y = (x + y) \cdot (z + y), \\ (x \cdot z) + z &= (x + z) \cdot (z + z) = (x + y) \cdot z = z \\ &= (x \cdot y) + z = (x + z) \cdot (y + z) \end{aligned}$$

und die beiden letzten Terme jeder Gleichung stimmen wegen (2) überein. \square

Wir können in B wie folgt eine Ordnung einführen: $a \leq b$ genau dann, wenn $a \cdot b = a$ gilt.

Lemma 12.0.10 $a \leq b$ gdw. $a + b = b$.

Beweis: $b = (a + b) \cdot b = a \cdot b + b \cdot b = a + b$. \square

Definition:

Seien $a \leq b \in B$, dann heißt die Menge $\{x \mid a \leq x \leq b\} = [a, b]$ das durch a und b bestimmte Intervall von B .

Wir bemerken, daß $[a, b]$ bezüglich der Addition und Multiplikation abgeschlossen sind. Wenn wir a als Nullelement und b als Einselement auffassen und die Komplementierung in $[a, b]$ relativ zu diesen durchführt (was auch immer das heißen mag), so wird $[a, b]$ wieder eine Boolesche Algebra.

Eine Abbildung zwischen Booleschen Algebren, die mit den jeweiligen drei Rechenoperationen verträglich ist, heißt Homomorphismus Boolescher Algebren. Ein bijektiver Homomorphismus heißt Isomorphismus.

Nun beweisen wir einen Struktursatz, der eine Übersicht über alle endlichen Booleschen Algebren ergibt.

Satz 12.0.34 Wenn B eine endliche Boolesche Algebra ist, so gilt $B \cong \mathbb{B}^n$ für eine natürliche Zahl n .

Beweis: Wir führen die Induktion über $|B|$. Wenn $|B| = 2$ ist, so ist nichts zu zeigen. Sei also die Behauptung für „kleine“ Boolesche Algebren schon bewiesen. Wir wählen ein Element $a \in B$, $a \neq 0, 1$. Wir setzen

$$X_a = \{(a \cdot b, a + b) \mid b \in B\},$$

dies ist eine Teilmenge von $[0, a] \times [a, 1]$.

Weiter sei $f : B \longrightarrow X_a$ folgende Abbildung: $f(b) = (a \cdot b, a + b)$. Nach der obigen Kürzungsregel ist f injektiv. Wir zeigen die Verträglichkeit mit den Rechenoperationen:

$$\begin{aligned} f(b \cdot c) &= (a \cdot (b \cdot c), a + (b \cdot c)), \\ f(b) \cdot f(c) &= (a \cdot b, a + b) \cdot (a \cdot c, a + c) \\ &= (a \cdot a \cdot b \cdot c, (a + b) \cdot (a + c)) \\ &= (a \cdot b \cdot c, a + (b \cdot c)), \\ f(b + c) &= f(b) + f(c) \end{aligned}$$

analog. Beim Komplement müssen wir aufpassen: Wir zeigen zunächst, daß $a \cdot \bar{b}$ das Komplement von $a \cdot b$ in $[0, a]$ ist.

$a \cdot \bar{b} + a \cdot b = a \cdot (b + \bar{b}) = a \cdot 1 = a$ ist das größte Element und $(a \cdot \bar{b}) \cdot (a \cdot b) = a \cdot 0 = 0$ ist das kleinste.

Analog: $a + \bar{b}$ ist das Komplement von $a + b$ in $[a, 1]$, da $a + \bar{b} + a + b = 1$ und $(a + \bar{b}) \cdot (a + b) = a + (\bar{b} \cdot b) = a + 0 = a$ ist das kleinste Element.

Nun folgt $f(\bar{b}) = (a \cdot \bar{b}, a + \bar{b}) = \overline{(a \cdot b, a + b)} = \overline{f(b)}$.

Nun ist f auch noch surjektiv, denn für $(x, y) \in [0, a] \times [a, 1]$ setzen wir $b = y \cdot (\bar{a} + x)$, dann ist $f(b) = (a \cdot y \cdot (\bar{a} + x), a + y \cdot (\bar{a} + x)) = (a \cdot y \cdot \bar{a} + a \cdot y \cdot x, (a + y)(a + \bar{a} + x))$; der erste Term ist Null, der zweite wegen $x \leq a \leq y$ gleich x , der dritte Term ist gleich $(a + y) \cdot (a + \bar{a} + x) = (a + y) \cdot 1 = y$.

Also ist f ein Isomorphismus Boolescher Algebren.

Da nun sowohl $[0, a]$ als auch $[a, 1]$ weniger Elemente als B haben, gilt für sie die Induktionsvoraussetzung: $[0, a] \cong \mathbb{B}^k$, $[a, 1] \cong \mathbb{B}^m$, also $B \cong \mathbb{B}^{k+m}$. \square

Die Menge \mathbb{B}^n ist isomorph zur Potenzmenge der Menge $\{1, \dots, n\}$, wir ordnen dem Tupel (i_1, \dots, i_n) die Menge der k mit $i_k \neq 0$ zu. Dies ist mit den Operationen verträglich.

Folgerung 12.0.12 (Stonescher Darstellungssatz) $B \cong P(M)$ für eine endliche Menge M . \square

Folgerung 12.0.13 Zwei gleichmächtige endliche Boolesche Algebren (mit 2^n Elementen) sind isomorph (zu \mathbb{B}^n). \square

Wir betrachten nun n -stellige Abbildungen der Form $f : B^n \longrightarrow B$. Wenn f, g zwei solcher Abbildungen sind, so können wir $(f \cdot g)(x) = f(x) \cdot g(x)$, $(f + g)(x) = f(x) + g(x)$ und $\overline{(f)}(x) = \overline{f(x)}$ setzen und es ist nicht schwer nachzuweisen, daß die Menge $F_n(B) = \{f : B^n \longrightarrow B\}$ so eine Boolesche Algebra wird.

Definition:

Ein Boolesches Polynom in x_1, \dots, x_n ist folgendes:

- (1) $x_1, \dots, x_n, 0, 1$ sind Boolesche Polynome,
- (2) wenn p und q Boolesche Polynome sind, so sind auch $(p) + (q)$, $(p) \cdot (q)$ und $\overline{(p)}$ Boolesche Polynome.

Ein Boolesches Polynom ist also einfach eine Zeichenkette, es gilt $x_1 + x_2 \neq x_2 + x_1$. Wenn aber $f(x_1, \dots, x_n)$ ein Boolesches Polynom und B eine Boolesche Algebra ist, so können wir eine Funktion $f^* : B^n \rightarrow B$ durch $f^*(b_1, \dots, b_n) = f(b_1, \dots, b_n)$ konstruieren, indem wir die b_i einfach in f einsetzen und den Wert ausrechnen. Dann gilt natürlich $(x_1 + x_2)^* = (x_2 + x_1)^*$.

Definition:

Zwei Boolesche Polynome f, g heißen äquivalent ($f \sim g$), wenn die zugehörigen Funktionen auf der Algebra \mathbb{B} gleich sind.

Zur Vereinfachung führen wir folgende Schreibweisen ein: $x^1 = x, x^{-1} = \bar{x}$.

Satz 12.0.35 Jedes Boolesche Polynom ist äquivalent zu einer „disjunktiven Normalform“

$$f_d(x_1, \dots, x_n) = \sum_{i_1, \dots, i_n \in \{0,1\}} d_{i_1 \dots i_n} \cdot x_1^{i_1} \cdot \dots \cdot x_n^{i_n}, \quad d_{i_1 \dots i_n} \in \{0,1\}.$$

Jedes Boolesche Polynom ist äquivalent zu einer „konjunktiven Normalform“

$$f_k(x_1, \dots, x_n) = \prod (k_{i_1 \dots i_n} + x_1^{i_1} + \dots + x_n^{i_n}), \quad k_{i_1 \dots i_n} \in \{0,1\}.$$

Beweis: Es ist $f^*(1^{j_1}, \dots, 1^{j_n}) = \sum d_{i_1 \dots i_n} 1^{i_1 j_1} \dots 1^{i_n j_n}$ und ein Summand ist genau dann gleich 1, wenn $i_1 = j_1, \dots, i_n = j_n$ und $d_{i_1 \dots i_n} = 1$ ist, das heißt, die f_d mit verschiedenen d sind jeweils inäquivalent. Nun ist aber die Anzahl der disjunktiven Normalformen gleich 2^{2^n} , also gleich der Zahl aller Funktionen $\mathbb{B}^n \rightarrow \mathbb{B}$.

Die zweite Aussage ergibt sich durch Dualisierung. □

Folgerung 12.0.14 In der obigen Darstellung ist $d_{i_1 \dots i_n} = f^*(1^{i_1}, \dots, 1^{i_n})$.

Beispiel: $f = ((x_1 + x_2) \cdot \bar{x}_1) + (x_2 \cdot (x_1 + \bar{x}_2))$, dann ist $f(0,0) = f(1,0) = 0$ und $f(0,1) = f(1,1) = 1$, die disjunktive Normalform von f erhalten wir, indem wir in der Wertetabelle die Stellen aufsuchen, wo der Wert 1 angenommen wird. Wenn hier ein Argument gleich 0 ist, so ist die entsprechende Variable zu komplementieren, sonst nicht. Also $f \sim \bar{x}_1 x_2 + x_1 x_2$. Dies kann weiter vereinfacht werden: $f \sim (\bar{x}_1 + x_1) \cdot x_2 = 1 \cdot x_2 = x_2$.

Wir überlegen nun, wie man eine Darstellung von Polynomen vereinfachen kann.

Definition Es seien p und q Boolesche Polynome; wir sagen, daß p das Polynom q impliziert, wenn aus $p^*(b_1, \dots, b_n) = 1$ folgt, daß auch $q^*(b_1, \dots, b_n) = 1$ gilt (dabei ist $b_i \in \{0,1\}$).

Wir bezeichnen ein Polynom als „Produkt“, wenn es kein $+$ -Zeichen enthält.

Das Polynom p heißt Primimplikant von q , wenn gilt

- 1) p ist ein Produkt,

- 2) p impliziert q ,
 3) kein Teilprodukt von p impliziert q .

Sei zum Beispiel $q = x_1x_2x_3 + x_1\bar{x}_2x_3 + \bar{x}_1\bar{x}_2\bar{x}_3$ und $p = x_1x_2$, dann wird q von p impliziert, denn $p^* = 1$ gilt nur für $x_1 = x_2 = 1$ und es ist $q^*(1, x_2, 1) = (x_2 + \bar{x}_2 + \bar{x}_2)^* = 1$, aber z.B. x_1 impliziert q nicht, da $q^*(1, x_2, x_3) = (x_2x_3 + \bar{x}_2x_3 + 0)^* = (x_2 + \bar{x}_2)x_3 = x_3 \neq 1$ ist.

Wir bemerken, daß ein Produkt genau dann 1 ist, wenn alle nichtkomplementierten Variablen gleich 1 und alle komplementierten Variablen gleich 0 gesetzt werden. Alle Summanden einer disjunktiven Normalform sind Implikanten.

Satz 12.0.36 *Jedes Polynom ist äquivalent zur Summe seiner Primimplikanten.*

Beweis: Seien p_1, \dots, p_m die Primimplikanten von q , wir setzen $p = p_1 + \dots + p_m$. Sei nun $p^*(b_1, \dots, b_n) = 1$, dann gibt es ein p_i mit $p_i(b_1, \dots, b_n) = 1$ und da p_i das Polynom q impliziert, gilt auch $q^*(b_1, \dots, b_n) = 1$.

Sei umgekehrt $q^*(b_1, \dots, b_n) = 1$, wir setzen $s = x_1^{i_1} \cdots x_n^{i_n}$ mit $i_k = 1$, falls $b_k = 1$ und $i_k = -1$ für $b_k = 0$, dann ist s ein Implikant von q . Wir lassen nun aus dem Term s alle die x_i weg, für die $q^*(b_1, \dots, b_{i-1}, \bar{b}_i, \dots) = 1$ ist; das Ergebnis sei r . Dann gilt: r impliziert q , aber kein Teilwort von r impliziert q , folglich ist r als Primimplikant gleich einem der p_j , also folgt $p^*(b_1, \dots, b_n) = 1$, d.h. $p \sim q$. \square

Von der disjunktiven Normalform eines Polynoms ausgehend kann man eine Darstellung als Summe von Primimplikanten erhalten, indem man für alle Paare von Summanden, wo dies möglich ist, die Regel $px + p\bar{x} \sim p$ anwendet.

Index

- ähnliche Matrizen, 102
- äquivalente Darstellungen, 132
- äquivalente Polynommatrizen, 100

- adjungierte Abbildung, 52
- affine Hülle, 84
- Algebra, 121
- Algebraische Körpererweiterungen, 25
- Anfangsminoren, 74

- Carmichaelzahlen, 161
- Charakter, 136

- Darstellung, 131
- Determinantenteiler, 101

- einfache Algebra, 123
- einfacher Modul, 122
- Einsdarstellung, 132

- Fermatsche Methode, 163

- größter gemeinsamer Teiler, 11
- Gram-Matrix, 77

- halbeinfache Algebra, 126
- halbeinfacher Modul, 124
- Horner-Schema, 7

- Interpolation, 22
- Invariantenteiler, 101
- invarianter Unterraum, 31, 135
- irreduzibel, 135

- Jacobisymbol, 161
- Jacobson-Radikal, 156
- Jordankästchen, 35
- Jordansche Normalform, 35

- Klassenfunktion, 139
- Kommutator, 133

- konvex, 85
- konvexe Hülle, 85
- konvexe Pyramide, 87
- konvexes Polyeder, 85

- Legendre, 164
- Legendresymbol, 161

- Matrixdarstellung, 131
- Matrixnorm, 71
- minimales Linksideal, 122
- Minimalpolynom, 19
- Moore-Penrose-Inverse, 58

- Newtonsche Formeln, 17
- nilpotent, 32, 143
- normale Matrix, 55
- normaler Endomorphismus, 55

- orthogonale Matrix, 48
- orthogonales Komplement, 45
- Orthogonalitätsrelationen, 138
- Orthonormalbasis, 44
- Orthonormalsystem, 44

- Polynom, 7
- positiv definit, 74
- positive Linearkombination, 85
- Potenzieren, 7
- Potenzreihe, 22
- Potenzsummen, 17
- primitiv, 147
- pseudoreguläre Matrix, 57

- quadratischer Rest, 161
- quadratisches Reziprozitätsgesetz, 161
- quasi-invertierbar, 153

- Radikal, 144
- reduzibel, 135

- reguläre Darstellung, 132
- reguläre Polynommatrix, 97
- Resultante, 14

- Satz von Euler, 161
- Satz von Hamilton-Cayley, 104
- Seite eines Simplex, 86
- selbstadjungierte Abbildung, 53
- semi-primitiv, 156
- Simplex, 86
- Skalarprodukt, 43
- Smithsche Normalform, 99
- Spektralzerlegung, 22
- sphärische Geometrie, 82

- Tensorprodukt von linearen Abbildungen,
117
- Tensorprodukt von Moduln, 115

- unipotent, 74
- unzerlegbarer Vektorraum, 32
- unzerlegbarer Modul, 147

- Vektornorm, 71
- Vektorprodukt, 81

- zerlegbarer Modul, 147
- zerlegbarer Vektorraum, 32