Mathe-LK Abitur 2002

BW

5. Juli 2003

Zusammenfassung

Hier ist gesammelt, was ich meinen Schülern aufgeschrieb, die im Jahre 2002 ihr Abitur machten. Es sind sicher Fehler drin, und alles ist aus konkreten Situationen entstanden, dennoch mag es nützlich sein.

Inhaltsverzeichnis

Teil I
12.1: Analysis

1 Orientierung

Was ist bisher erreicht? Du hast vertraute Dinge, vor allem die Ableitung einer Funktion, von einem etwas anderen Standpunkt aus betrachtet. Eine (abhängige) Größe y sei durch eine Gleichung

$$y = f(x) \tag{1}$$

an eine veränderliche Größe x gebunden. Eine Änderung Δx der einen Größe zieht eine Änderung Δy der anderen Größe nach sich. Das Verhältnis $\frac{\Delta y}{\Delta x}$ dieser Änderungen ist die mittlere Änderungsrate. Für $\Delta x \to 0$ geht diese gegen die lokale Änderungsrate, das ist die Ableitung der Funktion f an der Stelle x.

$$\frac{\Delta y}{\Delta x} \quad \stackrel{\Delta x \to 0}{\longrightarrow} \quad f'(x) \tag{2}$$

Diesen Sachverhalt kannst du am Graphen verdeutlichen, und du weißt, dass es da um Sekanten- und Tangentensteigungen geht.

Für "kleines" Δx haben wir praktisch gut brauchbare Näherungen, nämlich

$$\frac{\Delta y}{\Delta x} \approx f'(x) \quad \text{bzw} \quad \Delta y \approx f'(x)\Delta x \tag{3}$$

Mit Hilfe dieser Beziehungen können wir bequem einschätzen, wie sich ein relativer Fehler $\frac{\Delta x}{x}$ der Variablen x auf die Variable y auswirkt. Dazu berechnen wir den relativen Fehler $\frac{\Delta y}{y}$ von y.

Mehrfach waren wir auch in der Lage, für eine Größe y überhaupt erst einen Term zu finden, indem wir mit Hilfe von Δy und Δx die Ableitung von y berechneten und dann "aufleiteten".

Auch unsere Berechnung der Ableitungen von cos und sin haben wir über solche Änderungen bestimmt, und sie waren uns bei unserer Herleitung der Ableitungsregeln nützlich. Ableiten kann man, darüber spricht man nicht groß.

So, was machen wir nun mit unseren technischen Hilfsmitteln - außer Anwendungsaufgaben aus dem Buch? Wir sehen uns Kurven an, die durch Parameterdarstellungen gegeben sind: Ein Punkt P bewegt sich in der Ebene und befindet sich zur "Zeit" t an der Stelle $P(t) = (x(t), y(t))$. Das gibt Kreise, Parabeln und was nicht alles - eine herrliche Spielwiese. Die altbekannten Funktionsgraphen sind Spezialfälle dieser Kurven, da ist $x(t) = t$. Wir können viel untersuchen:

Wie "schnell" ist der Punkt zur Zeit t?

Wie lang ist ein Stück der Bahnkurve?

Hat die Bahnkurve besondere Punkte?

Wie groß ist der Inhalt eines Flächenstücks, das von einer oder mehreren Kurven begrenzt wird?

Eine Parabel ist offensichtlich in der Nähe des Scheitelpunkts stärker gekrümmt als weiter weg. Gibt es ein exaktes Maß für die Krümmung?

...

2 Anmerkungen zur Klausur

Dies ist ein Beispiel, wie ich mir eine Klausur so vorstelle. Die Aufgabe a) ist Routine, b) und d) und e) sind Anwendungsaufgaben verschiedenen Schwierigkeitsgrads. In c) wird überprüft, ob du den Mittelwertsatz verstanden hast und ob du dein Handwerkszeug anwenden kannst. Teil f) dann ist aus meiner Sicht ziemlich happig, du sollst unsere Methoden in einem neuen Kontext anwenden.

In der Regel läufst du in die Irre, wenn du blind drauf los rechnest. Ich empfehle dir dringend, Handskizzen zu zeichnen, die dir helfen, dir den Hintergrund klar zu machen. Vermeide, einfach Variable einzuführen ohne genau festzulegen, was sie bedeuten. Schreibe nicht zu eng und füge ruhig etwas erläuternden Text ein, damit deine Lösung lesbar wird. Und bleibe cool!

2.1 Klausuraufgabe

Es sei $y = f(x) = x^2(4 - x)$.

a) Bestimme die Extrema der Funktion und skizziere den Graphen.

b) Die Kurve schließt mit der x-Achse für $0 \leq x \leq 4$ ein Flächenstück ein. Bestimme seinen Inhalt.

c) Es sei $x > 0$. Dann garantiert uns der Mittelwertsatz, dass es ein z zwischen 0 und x mit einer bestimmten Eigenschaft gibt. Schreibe auf, um welche Eigenschaft es sich handelt, und berechne dieses $z = z(x)$.

d) Berechne das größte x mit $f(x) = 2$. Das wird nur näherungsweise gehen.

e) Es sei $P(x|y)$ ein Punkt der Kurve, und es sei $0 \leq x \leq 4$. Bestimme das Rechteck mit dem größten Flächeninhalt, dessen Seiten parallel zu den Achsen sind und das die Strecke vom Nullpunkt des Koordinatensystems bis P als Diagonale hat.

f) Es sei $l(x)$ die Länge der Kurve vom Nullpunkt bis zum Punkt $P(x|y)$ und $x \geq 0$. Bestimme $l'(x)$.

3 Klausur Nr. 1

1) Leite ab.

$$3x^2, \quad (2x+\pi)^4, \quad \frac{\sin x}{\cos x}, \quad (xf(x))^n, \quad a^3 + \sin\frac{1}{x}, \quad \sqrt{ax^2 + \sqrt{2}}, \quad \sqrt{5 - x^2}.$$

2) Die Strecke mit den Endpunkten $A(0|1)$ und $B(100|0)$ hat bestimmt Schnittpunkte mit der Sinuskurve $y = \sin x$.

a) Wie viele sind es? Begründe deine Antwort.

b) Berechne den x-Wert des ersten Schnittpunkts (zweimal Newton).

3) Der relative Fehler $\Delta x/x$ betrage 10%. Wie groß ist dann der relative Fehler $\Delta y/y$ der Größe $y = \sqrt{x}$?

4) Hummel Hermine fliegt so, dass sie sich zur Zeit t im Punkt $P(t) = (t|4 - t^2)$ befindet. Startzeit ist $t = -2$, Ankunftszeit $t = 2$.

a) Zeichne Hermines Flugbahn. (Im Laufe der Aufgabe musst du vielleicht mehrere Skizzen der Bahn zeichnen, aber das geht ja schnell)

b) Die Sonne steht so, dass der gleiche Sonnenstrahl durch die Punkte $P(-1)$ und $P(2)$ von Hermines Bahn geht. Zu welchem Zeitpunkt fliegt Hermine genau in Richtung des Sonnenlichts? (Anmerkung: Sonnenlicht besteht aus parallelen Strahlen)

c) Warum weiß ich von vornherein, dass Hermine auf jeden Fall irgendwann in der Zeit von -1 bis 2 in Richtung des Sonnenlichts fliegt? Stichwort genügt.

d) Wie groß ist der Inhalt des Flächenstücks, das Hermine in der Zeit von -1 bis $-\frac{1}{2}$ überfliegt?

e) Die Tangente an Hermines Bahn im Punkt $P(t)$ schließt für jedes t zwischen 0 und 2 mit den Koordinatenachsen ein Dreieck ein. Welchen Inhalt hat dieses Dreieck?

f) Zeige, dass Hermines Entfernung vom Nullpunkt $(0|0)$ zur Zeit t durch

$$\sqrt{t^4 - 7t^2 + 16}$$

gegeben ist.

g) Das Ergebnis von f) ist für Hermine deshalb wichtig, weil sie im Nullpunkt ihren schlimmsten Fressfeind vermutet. Dem will sie natürlich nicht zu nahe kommen. Zu welcher Zeit ist sie dem Nullpunkt am nächsten?

5) Du siehst eine steigende Kurve $y = f(x)$ und eine fallende Kurve $y = g(x)$. Zeige, dass für den Flächeninhalt $A(x)$ gilt: $A'(x) = f(x) - g(x)$.

4 Nachschreibklausur 1

1) Leite ab.

$$\text{a) } x^3 - \sqrt{2} \qquad \text{b) } (x^2 - 2\pi)^3 x \qquad \text{c) } \frac{\cos x}{\sin x} \qquad \text{d) } x f(x^n) \qquad \text{e) } (\cos f(x))^2$$

2) Es sei $y = f(x) = \sqrt{4 - x^2}$.

a) Skizziere grob die Graphen von f und von f'. Beschreibe den Graphen von f genauer: ist er symmetrisch, hat er Extrempunkte?

b) Zeige, dass in jedem Punkt P des Graphen von f die Tangente senkrecht zur Ursprungsgeraden durch P ist.

3) Es sei $y = f(x) = x^2 + \sin x$. Wieviele Extrema hat der Graph und wie viele Wendepunkte? Wo liegen diese Punkte? (Wenn genaue Werte nur mit großem Aufwand zu gewinnen sind, bin ich auch mit groben Angaben zufrieden)

4) Beschreibe, wie das Newton-Verfahren funktioniert, so dass man die Idee erkennt, und demonstriere das Rezept am Beispiel $y = f(x) = x^3 - x + 5$.

5) Es sei $y = f(x) = x^2 + \frac{1}{x}$.

a) Skizziere den Graphen von f. Bestimme dazu Extrem- und Wendepunkte.

b) Gibt es eine Tangente an den Graphen von f in einem Punkt zwischen $x = -1$ und $x = 1$, die parallel ist zur Sekante durch die Punkte $P(-1|*)$ und $Q(1|*)$ des Graphen?

c) Erkläre die Aussage des Mittelwertsatzes und beurteile, ob er mit dem Ergebnis von b) vereinbar ist.

6) Die Punkte $A(-1|0)$ und $B(x|y)$ des Einheitskreises sind Endpnkte einer Diagonalen eines Rechtecks mit achsenparallelen Seiten. Zeichne den Einheitskreis und zeichne ein solches Rechteck ein. Wo muss B liegen, damit der Inhalt des Rechtecks möglichst groß wird?

5 Riemannsche Summen

Da hatte unser Henrik schon den richtigen Riecher: Eine Funktion f ist über einem Intervall $[a, b]$ gegeben. Wir teilen das Intervall durch eine Zerlegung Z

$$Z: \quad a = z_0 < z_1 < \cdots < z_{i-1} < z_i < \cdots < z_n = b$$

in n Teilstücke. Das i-te Teilstück geht von z_{i-1} bis z_i. Darin wählen wir einen x-Wert x_i und bilden das Produkt

$$f(x_i)(z_i - z_{i-1}) \quad . \tag{4}$$

All diese Produkte addieren wir auf, das gibt

$$\sum_{i=1}^{n} f(x_i)(z_i - z_{i-1}) \quad , \tag{5}$$

und sowas nennt man eine Riemannsche Summe. Das ist eine unglaublich flexible und nützliche Begriffsbildung. Wenn wir sie ins Auge fassen und darüber intensiv meditieren, fällt uns alles in den Schoß. Du wirst sehen!

Henrik hatte die Vorstellung, dass diese $\Delta x_i = z_i - z_{i-1}$ irgendwie "unendlich klein" sein sollten. Dann würde die Summe aber unendlich viele Summanden haben, und das wäre eben keine sauber handhabare Summe mehr, sondern etwas Neues, nämlich ein Grenzwert Riemannscher Summen: das Leibnizsche **Integral** [1]

$$\int_a^b f(x)\,dx \quad .$$

Was soll das alles bedeuten? Nun, $f(x_i)(z_i - z_{i-1})$ ist für positives $f(x_i)$ der Inhalt eines Kästchens mit der Breite $\Delta x_i = z_i - z_{i-1}$ und der Höhe $f(x_i)$, und die Summe der Kästcheninhalte sollte ungefähr den Inhalt der Fläche zwischen Kurve und x-Achse in den Grenzen von $x = a$ und $x = b$ angeben. Falls $f(x) \geq 0$ ist in $[a, b]$, bedeutet das Integral also den Flächeninhalt.

Zweitens. Nimm an, wir haben eine Funktion F mit $F' = f$, also eine Stammfunktion (oder, wie wir sagten, eine "Aufleitung") von f. Dann können wir nach dem Mittelwertsatz das x_i im i-ten Teilstück so wählen, dass

$$f(x_i)(z_i - z_{i-1}) = F(z_i) - F(z_{i-1})$$

ist, und das geht natürlich für jedes der n Teilstücke. Der Wert der Riemannschen Summe ist dann

$$F(z_1) - F(z_0) + F(z_2) - F(z_1) + \cdots + F(z_n) - F(z_{n-1}) = F(z_n) - F(z_0) = F(b) - F(a) \ .$$

Da das für jede beliebige Zerlegung richtig ist, muss $F(b) - F(a)$ auch der Grenzwert der Riemannschen Summen sein:

$$\int_a^b f(x)\,dx = F(b) - F(a) \quad \text{für eine beliebige Funktion } F \text{ mit } F' = f \tag{6}$$

In Worten: Die Änderung einer Größe in der Zeit von a bis b ist das Integral über die Ableitung dieser Größe von a bis b. Konkretes Beispiel: Der in der Zeit von a bis b zurückgelegte Weg ist das Integral von a bis b über die Geschwindigkeit. Und Riemannsche Summen liefern Näherungswerte, wenn man das Integral nicht ausrechnen kann.

6 Taylorpolynome

Gegeben sei eine Funktion f. Wir suchen ein Polynom

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = \sum_{i=0}^{n} a_i x^i \quad , \tag{7}$$

so dass die durch $y = p(x)$ gegebene Funktion an der Stelle 0 mit f im Funktionswert und in den Werten der ersten n Ableitungen übereinstimmt. Für jedes n gibt es genau ein solches Polynom,

[1] Du erkennst unschwer die Herkunft dieses Symbols. Das Integralzeichen $\int$ ist ein stilisiertes Summenzeichen, darin steht der Summand $f(x)\,dx$. Dieses Integral ist also der Grenzwert, den die Riemannschen Summen annehmen, wenn man die Zerlegung immer feiner macht. Dass das mit dem Grenzwert funktioniert, ist noch zu zeigen.

sein Grad ist höchstens n. Denn für die i-te Ableitung von $p(x)$, die wir mit $p^{(i)}(x)$ bezeichnen wollen, gilt

$$p^{(i)}(0) = i(i-1)(i-2) \cdot \ \ldots \ \cdot 1a_i \quad , \tag{8}$$

- die Summanden, bei denen der Exponent von x kleiner als i ist, überstehen das Ableiten nicht; bei den Summanden, bei denen der Exponent von x größer als i ist, bleiben Faktoren x übrig, deshalb verschwinden sie, wenn man $x = 0$ einsetzt. Aus $p^{(i)}(0) = f^{(i)}(0)$ ergibt sich folglich mit der Abkürzung $i! := 1 \cdot 2 \cdot \ \ldots \ \cdot i$

$$a_i = \frac{f^{(i)}(0)}{i!} \tag{9}$$

und somit

$$p(x) = f(0) + f'(0)x + \frac{f^{(2)}(0)}{2}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \cdots + \frac{f^{(n)}(0)}{n!}x^n \quad . \tag{10}$$

Wenn wir $0! := 1$ und $1! := 1$ und $f^{(0)}(x) := f(x)$ setzen, können wir

$$p(x) = \sum_{i=0}^{n} \frac{f^{(i)}(0)}{i!}x^i \tag{11}$$

schreiben. Dies ist ein Taylor-Polynom. Erhöht man den Grad n, kommen nur weitere Summanden hinzu.

Einige Taylor-Polynome haben wir berechnet:

$$e^x \approx \sum_{i=0}^{n} \frac{1}{i!}x^i \tag{12}$$

$$\sin x \approx \sum_{i=0}^{n} (-1)^i \frac{x^{2i+1}}{(2i+1)!} \tag{13}$$

$$\cos x \approx \sum_{i=0}^{n} (-1)^i \frac{x^{2i}}{(2i)!} \tag{14}$$

Diese Reihen eröffnen überhaupt erst die Möglichkeit, Werte der Exponentialfunktion und der trigonometrischen Funktionen zu berechnen! Das ist schon eine wesentliche Anwendung der Taylor-Polynome.

Eine zweite Anwendungsmöglichkeit will ich jetzt aufzeigen. Wie du weißt, kennen wir keine Stammfunktion von

$$\varphi(x) := \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2} \quad , \tag{15}$$

und es gibt auch keine, die man in üblicher Form hinschreiben könnte. Hier kommen wir so weiter: Wir nehmen ein Taylor-Polynom $p(x)$ der Exponentialfunktion. Dann ist

$$\varphi(x) \approx \frac{1}{\sqrt{2\pi}}p(-\frac{1}{2}x^2) \quad , \tag{16}$$

und statt über φ integrieren wir einfach über diese Näherung. Man muss sich nur Gedanken über den Grad des Polynoms machen, um den Fehler in akzeptablen Grenzen zu halten. Die Funktion φ beschreibt Gauß´ Standardnormalverteilung, sie ist in der Praxis sehr wichtig. Und man braucht häufig Integrale über diese Funktion.

7 Über den Fehler des Taylor-Polynoms

Die Aufgabe, eine Polynomfunktion p vom Grade n zu finden, die an der Stelle 0 im Funktionswert und in den Werten der ersten n Ableitungen mit einer gegebenen Funktion f übereinstimmt, führte zum **Taylor-Polynom**

$$p(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \frac{f^{(3)}(0)}{6}x^3 + \cdots + \frac{f^{(n)}(0)}{n!}x^n = \sum_{i=0}^{n} \frac{f^{(i)}(0)}{i!}x^i \quad . \tag{17}$$

Wir haben uns an Beispielen angesehen, dass sich die Graphen der Taylor-Polynome dem Graphen der gegebenen Funktion mit steigendem Grad in der Tat immer besser anpassen. Somit kann man $p(x)$ als Näherungswert für $f(x)$ verwenden. Das ist von entscheidender Bedeutung: Du erinnerst dich, dass erst Taylorpolynome der e-Funktion uns ermöglichen, Funktionswerte dieser Funktion zu berechnen. Entsprechend sieht es mit der Sinus- und der Kosinusfunktion aus.

Ein Näherungswert ist eine heikle Sache, wenn man keine Vorstellung hat, wie groß der Fehler ist, also der Unterschied

$$r(x) := f(x) - p(x) \quad . \tag{18}$$

Den wollen wir uns nun ansehen.

Aus der Konstruktion des Taylor-Polynoms ergibt sich, dass

$$r(0) = r'(0) = r^{(2)}(0) = \cdots = r^{(n)}(0) = 0 \tag{19}$$

ist. Ferner ist

$$r^{(n+1)}(x) = (f(x) - p(x))^{(n+1)} = f^{(n+1)}(x) \quad , \tag{20}$$

denn die $(n+1)$-te Ableitung von $p(x)$ ist 0.

Wir können den Fehler abschätzen, wenn wir eine zusätzliche Voraussetzung über $f(x)$ machen, nämlich dass es eine Zahl M gibt mit

$$f^{(n+1)}(t) \leq M \text{ für alle } t \text{ mit } 0 \leq t \leq x \quad . \tag{21}$$

Damit geht das folgendermaßen (beachte (3)!):

$$\begin{aligned}
r^{(n)}(t) &= \int_0^t r^{(n+1)}(u)\,du \leq \int_0^t M\,du &&= Mt \\
r^{(n-1)}(t) &= \int_0^t r^{(n)}(u)\,du \leq \int_0^t Mu\,du &&= \frac{1}{2}Mt^2 \\
r^{(n-2)}(t) &= \int_0^t r^{(n-1)}(u)\,du \leq \int_0^t \frac{1}{2}Mu^2\,du &&= \frac{1}{3!}Mt^3 \\
&\cdots \\
r(t) &= \int_0^t r^{(1)}(u)\,du \leq \int_0^t \frac{1}{n!}Mu^n\,du &&= \frac{1}{(n+1)!}Mt^{n+1}
\end{aligned}$$

Damit haben wir folgendes Resultat bewiesen:

1 Satz *Es sei $p(x)$ das Taylor-Polynom vom Grade n an der Stelle 0 der Funktion f und es sei $|f^{(n+1)}(t)| \leq M$ für alle t zwischen 0 und x für eine feste Zahl M. Dann gilt*

$$|f(x) - p(x)| \leq \frac{Mx^{n+1}}{(n+1)!} \quad .$$

Eine Anwendung: Wenn wir $\sin x$ kennen für $0 \leq x \leq \frac{\pi}{4}$, können wir daraus die Sinuswerte beliebiger x berechnen, wie du dir am Graphen der Sinusfunktion klarmachen kannst. Welchen Grad muss ein Taylor-Polynom der Sinusfunktion haben, wenn wir damit Sinuswerte auf vier Stellen berechnen wollen?

8 Rückblick auf 12.1

In diesem Halbjahr solltest du deine Kenntnisse im Teilgebiet Analysis ausbauen. Du hast Funktionen in diesem Zusammenhang neu kennengelernt (Exponentialfunktion $y = e^x$, Logarithmusfunktion $y = \ln x$, trigonometrische Funktionen sin und cos), du kennst ihre Graphen, ihre Ableitungen und einige Anwendungen.

Du kennst die Ableitungsregeln (Summen-, Produkt-, Quotienten- und Kettenregel), und du weisst, wie man Potenzfunktionen $y = x^r$ für $r \in \mathbb{R}$ ableitet.

Du kannst mit Änderungen rechnen: du kennst die wichtige Beziehung $\Delta y \approx f'(x)\Delta x$ und den Mittelwertsatz $\Delta y = f'(z)\Delta x$ für ein z zwischen x und $x + \Delta x$ und seine geometrische Bedeutung.

Der wichtigste neue Begriff ist der Integralbegriff. Du kennst seine Bedeutung, den Zusammenhang mit Riemannschen Summen, für ordentliches f seine Berechnung über den Hauptsatz

$$\int_a^b f(x)\,dx = F(b) - F(a) \text{ für eine Stammfunktion } F \text{ von } f$$

sowie die typischen Ansätze

$$\int_a^b f(x)\,dx = (b-a)f(z) \text{ für ein } z \in [a,b]$$

und

$$f(x+\Delta x) = f(x) + \int_x^{x+\Delta x} f'(x)\,dx \quad .$$

Du kannst Flächen- und Rauminhalte mit Hilfe eines einfachen oder eines mehrfachen Integrals berechnen, dabei mag auch ein uneigentliches Integral

$$\int_a^\infty f(x)\,dx$$

auftreten. Ebenso kennst du Integrationsverfahren (Partielle Integration, Subsitutionsregel) und einige Regeln für Integrale (z.B. $\int(f+g) = \int f + \int g$ oder $\int(cf) = c(\int f)$). Und du weißt, wie du im Notfall mit Hilfe Riemannscher Summen einen Näherungswert berechnest.

Du kennst Beispiele für Kurven in Parameterdarstellung als Beschreibung der Bewegung eines Punktes in der Ebene oder im Raum, du kannst die Momentangeschwindigkeit $v(t)$ des Punktes, die Länge s des zurückgelegten Weges und den Inhalt A einer von einer Kurve eingeschlossenen Fläche berechnen:

$$v(t) = \sqrt{x'(t)^2 + y'(t)^2} \quad , \quad s = \int_a^b v(t)\,dt \quad , \quad A = \int_a^b x(t)y'(t)\,dt \quad .$$

Schließlich kennst du Taylor-Polynome

$$\sum_{i=0}^n \frac{f^{(i)}(0)}{i!}x^i \quad ,$$

die eine Funktion f approximieren, und du kennst eine Abschätzung des Fehlers.

Erwähnung verdient auch, dass du Erfahrungen mit Maple gesammelt hast - sicher nicht ihr alle in gleichem Maße, aber im Mittel ganz ordenlich. Insgesamt liest sich das ganz gut, wir können mit dem Erreichten zufrieden sein. Ich hätte gern die Kurven intensiver behandelt, und wir entwickeln keine geschlossene Theorie, sondern bearbeiten Probleme mit adhoc-Methoden. Manche von euch hätten sich mehr Anwendungsaufgaben gewünscht. Ich denke, ich bringe beides noch unter. Erstmal greifen wir entschlossen mitten hinein in die Stochastik, und zwar gleich nächste Woche.

Teil II

12.2: Stochastik

9 Einstieg in die Wahrscheinlichkeitstheorie

9.1 Vorbemerkung

Da das lange Wort Wahrscheinlichkeit sehr oft vorkommt, verabreden wir hiermit, es mit W. abzukürzen. Dann ist es also W-theorie, womit wir uns in der nächsten Zeit beschäftigen, und der Name ist Programm: es handelt sich um eine mathematische Theorie, eine sehr schöne übrigens. Inspiriert ist sie von den Alltagsphänomenen Zufall und Ungewissheit, und sie hilft durchaus, mit diesen Phänomenen umzugehen, denn die Alltagsphänomene verhalten sich in erstaunlichem Maße so wie geeignet gebildete w-theoretische Modelle. Aber diese Übereinstimmung ist ein empirischer Befund, W-theorie selbst ist reine Mathematik.

Am Anfang einer ausformulierten mathematischen Theorie stehen Axiome, die zentrale Eigenschaften der Gegenstände der Theorie erfassen. Das übliche Axiomensystem der W-theorie wurde von Andrej Kolmogorow in den dreißiger Jahren des zwanzigsten Jahrhunderts formuliert, also erst nach einigen hundert Jahren Beschäftigung mit Fragen, die zur W-theorie gehören. Es bekäme dir nicht, wenn ich das Axiomensystem an den Anfang unserer Arbeit stellte und daraus Sätze deduzierte. Ich will aber auch nicht alles auf dein intuitives Bild von "Wahrscheinlichkeit" stützen und einfach praktischen Fragen nachgehen. Deshalb gebe ich ein Modell vor, das unmittelbar einleuchtende Eigenschaften hat[2]. Damit und mit den daran aufgehängten Begriffen arbeiten wir dann. Auf diese Weise sind wir gleich mitten drin, ohne monatelange Durststrecke durch das Dickicht der Kombinatorik, das auf den ersten hundert Seiten unseres Lehrbuches wuchert.

9.2 Das Standardmodell

Ich denke mir eine Maschine, die auf Befehl eines der Elemente ω einer vorgegebenen Menge Ω so auswählt, dass alle Elemente von Ω gleichberechtigt sind. Das war´s schon! Die Elemente $\omega \in \Omega$ heißen **Ergebnisse**, die Menge Ω selbst heißt **Stichprobenraum**. Man muss von der Menge Ω selbst und von vernünftigen Teilmengen $A \subseteq \Omega$ die Größe messen können, wir bezeichnen die Größe von A mit $|A|$. Dann wird die **Wahrscheinlichkeit** $P(A)$ auf

$$P(A) := \frac{|A|}{|\Omega|} \tag{22}$$

festgesetzt,[3] und diese vernünftigen Teilmengen heißen **Ereignisse**.

9.3 Beispiele

1. Dies ist der für uns wichtigste Typ: Es sei Ω die Menge der Punkte des Einheitsquadrats, und für eine Teilmenge A von Ω sei $|A|$ einfach der Flächeninhalt von A. Vernünftige A´s sind solche, die einen Flächeninhalt haben! Dann ist
$P(\text{Dreieck mit den Eckpunkten } (0|0),\ (1/2|0) \text{ und } (1|1/3)) = \frac{1}{12}/1 = 1/12$
$P(\text{Fläche unter der Normalparabel}) = \int_0^1 x^2\,dx/1 = \frac{1}{3}$
$P(\text{Diagonale des Quadrats}) = 0$ und
$P(\{(0|1)\}) = 0.$

[2] Was noch lange nicht sagt, dass ein solcher Gegenstand real existiert, nicht einmal, dass die postulierten Eigenschaften des Modells logisch widerspruchsfrei sind.

[3] Aufgepasst: Nach langem fruchtlosem Grübeln, was "W." denn sei, wurde der gordische Knoten durchgeschlagen, indem man eine rechnerisch fassbare Größe bildete und festlegte, dass diese Größe eben W. heißt. Natürlich verhält sich die Größe so, wie man es von der W. erwartet, sonst wäre nichts gewonnen.

2. Es sei Ω das Intervall $[0;2]$ der Zahlengeraden. Als vernünftige Teilmengen lassen wir solche zu, die sich aus Teilintervallen zusammensetzen, ihr Maß ist die Summe der Längen der Teilintervalle. Dann ist $P([0;1]) = 1/2$.

3. Es sei $\Omega = \{1, 2, 3, \ldots, 20\}$, und für eine Teilmenge A von Ω sei $|A|$ einfach die Anzahl der Elemente von A. Wenn nun A alle durch drei teilbaren Zahlen von Ω enthält, ist $P(A) = 6/20$.

9.4 Zufallsgrößen

Mathematisch ist eine Zufallsgröße X einfach eine Abbildung $X : \Omega \to \mathbb{R}$. Am besten stellst du dir das so vor, dass jedem $\omega \in \Omega$ ein Zettel angeheftet ist, auf dem eine reelle Zahl steht. Die Zahl auf dem Zettel an ω ist der Wert $X(\omega)$ von ω unter X. Gleich zwei Beispiele, die sich auf das erste Ω des vorigen Abschnitts beziehen: Wir heften jedem Punkt $(x|y)$ des Einheitsquadrats einen Zettel an und schreiben darauf schlicht die x-Koordinate des Punkts. Dann haben wir eine Zufallsgröße X definiert, und es ist $X((x|y)) = x$. Trotz ihrer Einfachheit ist dies eine sehr brauchbare Zufallsgröße. Man könnte auf den Zettel an $\omega = (x|y)$ auch den Abstand des Punkts vom Nullpunkt schreiben, dann hätte man eine Zufallsgröße, meinetwegen mit dem Namen Y, und es wäre $Y((x|y)) = \sqrt{x^2 + y^2}$.

Der Zufall kommt bei der Zufallsgröße so ins Spiel. Wenn die Maschine auf Knopfdruck ihr ω ausspuckt, sehen wir es uns nicht selbst an, sondern nur seinen Wert $X(\omega)$ unter X. Unter $P(X = 1/2)$ oder unter $P(1/3 < X < 1/2)$ oder unter $P(Y \leq 1)$ versteht man die Wahrscheinlichkeit der Teilmenge der $\omega \in \Omega$, deren X-Wert $1/2$ ist oder deren X-Wert zwischen $1/3$ und $1/2$ liegt oder deren Y-Wert höchstens 1 ist. Hier ist übrigens

$$\begin{aligned} P(X = 1/2) &= 0 \\ P(1/3 < X < 1/2) &= 1/6 \\ P(Y \leq 1) &= \pi/4 \end{aligned}$$

Mache dir das an einer Zeichnung klar!

9.5 Vokabelheft

Ich gebe dir hier eine kurze Liste mit den wichtigsten Begriffen, den dafür üblichen Symbolen und entsprechenden englischen Fachausdrücken.

Begriff	Symbol	englische Bezeichnung	Anmerkung
Stichprobenraum	Ω	sample space	eine Menge
Ergebnis	ω	outcome	$\omega \in \Omega$
Ereignis	$A, B, \ldots$	event	Teilmenge von Ω
Wahrscheinlichkeit	P	probability	$P(A)$ für $A \subseteq \Omega$
Zufallsgröße	X, Y, Z	random variable	Abb. $\Omega \to \mathbb{R}$
Verteilungsfunktion von X	F	distribution function	$F(x) := P(X \leq x)$
Dichtefunktion von X	f, φ	probability density function	$F(x) = \int_{-\infty}^{x} f(t)\, dt$

10 Übungsaufgaben zu stetig verteilten Zufallsgrößen

Diese Aufgaben stammen aus dem Buch “A First Course in Probability” von Sheldon Ross, erschienen 1976 in New York.

10.1 Example 1a, p. 122

Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

What is the value of C? Find $P(X > 1)$.

10.2 Example 1b, p. 122

The amount of time, in hours, that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

What is the probability that a computer will function between 50 and 150 hours before breaking down? What is the probability that it will function less than 100 hours?

10.3 Example 1c, p. 123

The lifetime in hours of a certain kind of radio tube is a random variable having a probability density function given by

$$f(x) = \begin{cases} 0 & x \leq 100 \\ \frac{100}{x^2} & x > 100 \end{cases}$$

What is the probability that exactly 2 of 5 such tubes in a radio set will have to be replaced within the first 150 hours of operation? [4]

10.4 Example 2a, p. 125

If X is uniformly distributed over $(0; 10)$ [5], calculate the probability that (1) $X < 3$, (2) $X > 6$, and (3) $3 < X < 8$.

10.5 Example 2b, p. 126

Buses arrive at a specified stop at 15-minute intervals starting at 7 a.m. That is, they arrive at 7, 7:15, 7:30, 7:45, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7:30, find the probability that he waits (1) less than 5 minutes for a bus? (2) More than 10 minutes for a bus?

10.6 Example 2c, p. 126

Consider a "random chord" of a circle. What is the probability that the length of the chord will be greater than the side of the equilateral triangle inscribed that circle?[6]

[4] Diese Aufgabe ist nicht, wie die vorigen Aufgaben, nur eine technische Übung. Da müssen wir uns eine ganze Reihe von Gedanken machen und einige Voraussetzungen formulieren.

[5] gleich-verteilt, das heißt, dass die Dichtefunktion zwischen 0 und 10 konstant ist und sonst 0 ist.

[6] Diese Aufgabe ist schwierig!

10.7 Example 3c, p. 133

An expert witness in a paternity suit testifies that the length (in days) of pregnancy (that is, the time from impregnation to the delivery of the child) is approximately normally distributed with parameters $\mu = 270$ and $\sigma^2 = 100$. The defendant in the suit is able to prove that he was out of the country during a period that began 290 days before the birth of the child and ended 240 days before the birth. If the defendant was, in fact, the father of the child, what is the probability that the mother could have had the very long or very short pregnancy indicated by the testimony?

10.8 Text: Exponential Random Variables, p. 136

A continous random variable whose probability density function is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

[...] The exponential distribution often arises, in practice, as being the distribution of the amount of time until some specific event occurs. For instance, the amount of time (starting from now) until an earthquake occurs, or until a new war breaks out, or until a telephone call you receive turns out to be a wrong number are all random variables that tend in practice to have exponential distributions.

10.9 Example 4a, p. 136

Suppose that the length of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{10}$. If someone arrives immidiately ahead of you at a public telephone booth, find the probability that you will have to wait (1) more than 10 minutes, and (2) between 10 and 20 minutes.

10.10 Example 4b, p. 137

Consider a post office that is manned by two clerks. Suppose that when Mr. Smith enters the system he discovers that Mr. Jones is being served by one of the clerks and Mr. Brown by the other. Suppose also that Mr. Smith is told that his service will begin as soon as either Jones or Brown leaves. If the amount of time that a clerk spends with a customer is exponentially distributed with parameter λ, what is the probability that, of the three customers, Mr. Smith is the last to leave the office?[7]

10.11 Zerfall

Ein Teilchen einer radioaktiven Substanz, das wir jetzt vor uns sehen, zerfällt zu einem zufälligen Zeitpunkt. Die auf diese Weise gegebene Zufallsgröße ist exponentialverteilt mit Parameter $\lambda = 8$.

Wir sperren zwei Teilchen dieser Substanz in ein Gefäß, sie heißen Hans und Fritz. Berechne die Wahrscheinlichkeiten der folgenden Ereignisse.
a) Hans lebt mindestens doppelt so lange wie Fritz.
b) Nach 20 (Tagen) gibt es beide Teilchen nicht mehr.
c) Nach vier (Tagen) ist mindestens eins der Teilchen zerfallen.
d) Hans zerfällt im Zeitraum $[4, 6]$.
e) Nimm an dass Fritz nach sechs (Tagen) noch da ist und dass Hans nicht mehr zu sehen ist. Wie groß ist die W, dass Hans im Zeitraum $[4, 6]$ zerfallen ist, wenn er mit Sicherheit nicht entkommen konnte (er vielleicht entkommen ist)?

Bedenke, dass du zwei exponentialverteilte Zufallsgrößen brauchst, dass du für Ω einen geeigneten Körper wählen musst und für die Ereignisse die richtigen Gebiete der xy-Ebene erkennen musst.

[7] Diese Aufgabe ist schwierig!

10.12 Rechnen mit Verteilungen ohne realistischen Bezug

a) Die Zufallsgröße X habe die Dichtefunktion

$$f(x) = \begin{cases} Cx(4-x) & \text{für } 0 \le x \le 4 \\ 0 & \text{sonst} \end{cases}$$

Wie groß ist C?
b) Die Zufallsgröße Y habe die Dichtefunktion

$$g(y) = \begin{cases} Cy & \text{für } 0 \le y \le 4 \\ 0 & \text{sonst} \end{cases}$$

Wie groß ist C?
c) Durch

$$z = \begin{cases} f(x)g(y) & \text{für } 0 \le x, y \le 4 \\ 0 & \text{sonst} \end{cases}$$

ist eine Funktion $\mathbb{R}^2 \to \mathbb{R}$ gegeben. Unser Ω ist der Körper zwischen dem Graphen dieser Funktion und der xy-Ebene. Ein Punkt ω dieses Körpers werde zufällig gewählt. Seine x-Koordinate liefert die Zufallsgröße X, seine y-Koordinate die Zufallsgröße Y. Weise die folgenden Tatsachen nach.
c(i) Es ist $P(a \le X \le b) = \int_a^b f(x)\,dx$, das heißt, dass unser f aus a) Dichtefunktion von X ist.[8]
c(ii) Desgleichen für Y und g.
c(iv) Berechne $P(1 \le X, Y \le 2)$ und $P(X + Y \le 4)$.[9]

Hinweis: Lass dir den Körper und die Teile, die du betrachtest, durch Maple darstellen. Auch die Integrale kannst du mit Maple nachrechnen.

11 Zum Begriff der Wahrscheinlichkeit

In einem bahnbrechenden Buch[10] schafft A. Kolmogoroff der Wahrscheinlichkeitstheorie eine axiomatische Grundlage und macht sie dadurch zu einer sauberen und leistungsfähigen Theorie. Wichtige philosophische Fragen um die Wahrscheinlichkeit können offen bleiben; Kolmogoroff betrachtet eine Funktion P von einem geeigneten System von Teilmengen einer Menge Ω - den Ereignissen - in das Intervall $[0; 1]$ und verlangt von ihr einige einfache Eigenschaften - das genügt. Darauf steht der ganze Bau.

Natürlich fallen die Axiome mit den Eigenschaften nicht vom Himmel, sondern sind von den Vorstellungen von Wahrscheinlichkeit inspiriert, die den Leuten vorschwebten. Komogoroff beschreibt dies ganz deutlich in einem eigenen Abschnitt mit der Überschrift "Das Verhältnis zu Erfahrungswelt"[11]. Ich gebe einige Zeilen daraus wieder.

A Man kann praktisch sicher sein, daß, wenn man [...] eine große Anzahl von n Malen wiederholt und dabei durch m die Anzahl der Fälle bezeichnet, bei denen das Ereignis A stattgefunden hat, das Verhältnis m/n sich von $P(A)$ nur wenig unterscheidet.

B Ist $P(A)$ sehr klein, so kann man praktisch sicher sein, daß bei einer einmaligen Realisation [...] das Ereignis A nicht stattfindet.

---

[8] a und b schränkst du selbst sinnvoll ein
[9] Versäume nicht nachzurechnen, dass

$$\int_0^4 \int_0^4 z\,dx\,dy = 1$$

[10] Grundbegriffe der Wahrscheinlichkeitsrechnung. Berlin 1933
[11] Seite 3ff

Bemerkung I Aus der praktischen Sicherheit zweier Behauptungen folgt die praktische Sicherheit der Behauptung ihrer gleichzeitigen Richtigkeit, obwohl der Sicherheitsgrad sich dabei ein wenig erniedrigt. Ist jedoch die Anzahl der Behauptungen sehr groß, so lassen sich aus der praktischen Sicherheit jeder einzelnen Behauptung überhaupt keine Schlüsse ziehen. Deshalb folgt aus dem Prinzip **A** noch keineswegs, daß bei einer sehr großen Anzahl von Serien von Versuchen, von denen jede Serie aus n Versuchen besteht, in *jeder* Serie der Quotient m/n sich von $P(A)$ wenig unterscheiden wird.

Bemerkung II Dem unmöglichen Ereignis (der leeren Menge) entspricht kraft unserer Axiome die Wahrscheinlichkeit $P(\emptyset) = 0$, während umgekehrt aus $P(A) = 0$ die Unmöglichkeit des Ereignisses A durchaus nicht zu folgen braucht; nach dem Prinzip **B** folgt aus dem Nullwerden der Wahrscheinlichkeit nur, daß bei einer einmaligen Realisation [...] das Ereignis A praktisch unmöglich ist. Das bedeutet keineswegs, daß auch bei einer genügend langen Reihe von Versuchen das Ereignis A nicht auftreten wird. [...]

12 Klausur 3

1) In der dunklen Werkstatt hat Herr Karl eine Schublade mit Arbeitshandschuhen. Zwei linke und ein rechter Handschuh liegen darin. Herr Karl nimmt so lange einzeln Handschuhe aus der Schublade, bis er ein Paar hat.

a) Nenne die Anzahl der erforderlichen Ziehungen X und berechne $E(X)$.

b) Was bedeuten die Aussagen $E(X) = 2,1$ und $P(X = 2) = 0,7$ in diesem Zusammenhang praktisch?

2) a) Es sei $f(x) = c$ für $0 \leq x \leq \frac{1}{2}$ und 0 sonst. Bestimme c so, dass f Dichtefunktion einer Zufallsgröße X ist, bestimme die Verteilungsfunktion F von X und zeichne die beiden Graphen.

b) Gib ohne Rechnung den Erwartungswert von X an. Begründe deine Antwort kurz.

3) Zeichne in ein Koordinatensystem den Halbkreis mit Mittelpunkt $M(0|0)$ und Radius 5, der oberhalb der x-Achse liegt, und verbinde den Punkt $A(3|0)$ mit dem Kreispunkt $B(3|4)$.

a) Ein Punkt P des Halbkreises[12] werde zufällig gewählt. Mit welcher Wahrscheinlichkeit schneiden sich die Strecken $\overline{MP}$ und $\overline{AB}$

(i) überhaupt

(ii) im Mittelpunkt von $\overline{AB}$?

b) Nimm an, es sei schon bekannt, dass sich die beiden Strecken schneiden. Mit welcher Wahrscheinlichkeit liegt der Schnittpunkt dann unterhalb des Mittelpunkts von $\overline{AB}$?

4) Das Flächenstück unter der Normalparabel zwischen $x = 0$ und $x = 1$ rotiere um die x-Achse. Es entsteht ein Rotationskörper. Ein Punkt ω des Körpers werde zufällig gewählt, die Zufallsgröße X bezeichne seine x-Koordinate.

a) Berechne $P(X \leq \frac{1}{2})$.

b) Bilde die Verteilungsfunktion F und die Dichtefunktion $f = F'$ von X.

c) Berechne $E(X)$.

d) Welche physikalische Bedeutung mag der Wert aus c) haben?

5) Der Zerfallszeitpunkt des radioaktiven Teilchens Fritz sei eine Zufallsgröße X mit der Dichtefunktion $f(x) = e^{-x}$ für $x \geq 0$.

a) Berechne $E(X)$. Was bedeutet der Wert?

b) Mit welcher Wahrscheinlichkeit zerfällt Fritz in der ersten Zeiteinheit?

[12] Hier meine ich wirklich nur die Kreislinie, nicht die Kreisscheibe!

c) Nimm an, Fritz erlebt den Zeitpunkt 4. Mit welcher Wahrscheinlichkeit zerfällt er dann in der nächsten Zeiteinheit? Das sollst du hier wirklich ausrechnen und dann sagen, warum du es eigentlich nicht rechnen musstest.

d) Wie lange dauert es, bis Fritz (i) mit Sicherheit (ii) mit $99,9\%$-iger Wahrscheinlichkeit zerfallen ist?

e) Wir sperren wieder Fritz zusammen mit dem Teilchen Klara gleicher Sorte in eine Bleikammer. In einer Zeiteinheit wird der Strahlenschutzbeauftragte erwartet. Mit welcher Wahrscheinlichkeit werden sie dann beide zerfallen sein?

f) Mit welcher Wahrscheinlichkeit lebt Klara mindestens doppelt so lange wie Fritz, aber höchstens zwei Zeiteinheiten?

13 Übung zum Rechnen mit Erwartungswerten

Wir würfeln zweimal mit einem Tetraeder - seine Seitenflächen zeigen die Zahlen 1 bis 4. Gewürfelt ist die Zahl auf der Fläche, auf der das Tetraeder zu liegen kommt.

1) Es sei X_i die im i-ten Wurf gewürfelte Augenzahl. Berechne $E(X_i)$ und $V(X_i)$. Gib Erwartungswert und Varianz von $X_1 + X_2$ und von $X_1 \cdot X_2$ an.

2) Es sei G die größte gewürfelte Augenzahl und K die kleinste.

a) Berechne die folgenden Wahrscheinlichkeiten:

$$\begin{array}{ccc} P(G=3) & P(K=2) & P(G=3, K=2) \\ P_{G=3}(K=2) & P_{G=3}(K=4) & P(G \leq 2) \end{array}$$

Am besten zeichnest du dir Ω als Rechteck hin und teilst es in Spalten und Zeilen ein, so dass in jeder Zeile die Elemente ω mit dem gleichen $G(\omega)$ und in jeder Spalte die Elemente mit dem gleichen $K(\omega)$ stehen. Sind G und K unabhängig?

b) Zeichne den Graphen der Verteilungsfunktion von G.

c) Berechne Varianz und Erwartungswert von G und von K, von $G + K$ und von $G \cdot K$. Vergleiche deine Ergebnisse mit dem, was die Regeln vorhersagen.

3) Wir werfen den Würfel n-mal, bezeichnen die im i-ten Wurf erzielte Augenzahl mit X_i und bilden den Mittelwert

$$S_n = \frac{1}{n} \sum_{i=1}^{n} X_i \quad .$$

a) Gib Erwartungswert und Varianz von S_n an.

b) Was sagt die Ungleichung von Tschebyschew über

$$P(|S_n - E(S_n)| \geq 0,1)$$

und was bedeutet das für große n? Bleibt das Ergebnis stabil, wenn man $0,1$ durch eine beliebig kleine positive Zahl ϵ ersetzt? Formuliere das Ergebnis präzise und würdige es: Es ist von großer Bedeutung für die Wahrscheinlichkeitstheorie.

14 Statistik

Damit du besser durchblickst, schreibe ich hier ein paar Dinge zusammen; die Überschrift passt nicht ganz.

14.1 Lage- und Streuungsparameter einer Zufallsgröße

Es sei X eine Zufallsgröße auf einem Ergebnisraum Ω. Zunächst können wir Wahrscheinlichkeiten ausrechnen, dass X einen Wert annimmt, der in einem bestimmten Bereich liegt. Dann können wir fragen, was denn typischerweise herauskommt, wenn man einen Wert von X auslost. Eine Antwort auf diese Frage ist der Erwartungswert $E(X)$ der Zufallsgröße.

14.1.1 Der Erwartungswert $E(X)$ der Zufallsgröße X

Die Definition ist

$$E(X) := \sum_{x \in X(\Omega)} xP(X = x) \tag{23}$$

für eine diskrete Zufallsgröße und

$$E(X) := \int_{X(\Omega)} xf(x)\, dx \tag{24}$$

für stetig verteiltes X mit Dichtefunktion f.[13]

Die Bedeutung des Erwartungswerts muss dir klar sein: Wenn du viele Werte von X auslost und ihren Durchschnitt bildest, sollte ungefähr der Erwartungswert herauskommen. Das heißt keineswegs, dass alle Werte von X in der Nähe des Erwartungswerts lägen, vielleicht gibt es in der Nähe des Erwartungswerts überhaupt keine Werte von X.

Damit sind wir beim nächsten Punkt: Die Werte von X streuen. Für diese Streuung hätte man gern ein Maß, und das gängigste Streuungsmaß ist die Varianz $V(X)$ der Zufallsgröße bzw. die daraus berechnete Standardabweichung $\sigma := \sqrt{V(X)}$.

14.1.2 Varianz $V(X)$ und Standardabweichung σ der Zufallsgröße X

Ich schreibe eine Definition der Varianz hin:

$$V(X) := E\left((X - E(X))^2\right) \tag{25}$$

Die Varianz heißt demnach auch mittlere quadratische Abweichung.

Zwei wichtige Aussagen über die Varianz:

1) Die Zufallsgröße X besitzt Werte sowohl im Bereich $[E(X) - \sigma, E(X) + \sigma]$ als auch in dem Bereich, der aus den beiden Teilen $(-\infty, E(X) - \sigma]$ und $[E(X) + \sigma, \infty)$ besteht.[14]

2) Die Wahrscheinlichkeit, dass die Zufallsgröße Werte liefert, die von $E(X)$ mindestens a entfernt sind, kann nicht beliebig groß sein. Eine obere Schranke für diese Wahrscheinlichkeit liefert die berühmte **Ungleichung von Tschebyschew**

$$P\left(|X - E(X)| \geq a\right) \leq \frac{V(X)}{a^2} \tag{26}$$

14.1.3 Gesetze für E und V

Erwartungswert und Varianz gehorchen den folgenden Gesetzen für Zahlen $r \in \mathbb{R}$ und Zufallsgrößen X und Y auf dem gleichen Ω:

$$E(X + Y) = E(X) + E(Y) \qquad E(rX) = rE(X) \tag{27}$$

$$V(rX) = r^2V(X) \qquad V(X + r) = V(X) \tag{28}$$

Falls X und Y unabhängig von einander sind, gilt sogar noch

$$E(XY) = E(X)E(Y) \qquad V(X + Y) = V(X) + V(Y) \tag{29}$$

[13] Dir sollten die Symbole inzwischen so vertraut sein, dass du sie verstehst und selbst sachgerecht benutzen kannst. Lies da bloß nicht so drüber weg!

[14] Zeichne dir bei solchen Angaben immer eine Skizze der Merkmalsachse und suche die Bereiche auf!

14.1.4 Bernoullis Schwaches Gesetz der großen Zahl

Nun haben wir genug beisammen, ein wichtiges Ergebnis der W-Theorie zu formulieren und zu beweisen, das oben im Abschnitt Erwartungswert schon angedeutet wurde. Dazu losen wir n Werte der Zufallsgröße X aus und bilden den Mittelwert. Wenn man n Werte von X auslost, erhält man eine Liste $x_1, x_2, \ldots, x_n$ von Werten von XM; ich spreche manchmal auch von Worten, das Alphabet ist die Menge der Werte von X. Der i-te Buchstabe x_i des Worts ist selbst Wert einer Zufallsgröße X_i, und X_i ist praktisch eine Kopie von X selbst. Insbesondere hat X_i den gleichen Erwartungswert und die gleiche Varianz wie X. Nun können wir Erwartungswert und Varianz des Mittelwerts M ausrechnen:

$$E(M) = E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X) = \frac{1}{n} nE(X) = E(X) \tag{30}$$

Analog:

$$V(M) = V(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2} nV(X) = \frac{1}{n} V(X) \tag{31}$$

Der Mittelwert M streut also weniger als X selbst! Die Tschebyschew-Ungleichung liefert nun das Ergebnis. Für jede (noch so kleine) Zahl $\epsilon > 0$ gilt nämlich

$$P\left(|M - E(M)| \geq \epsilon\right) \leq \frac{V(M)}{\epsilon^2} = \frac{V(X)}{n\epsilon^2} \tag{32}$$

Die rechte Seite der Ungleichung strebt aber gegen 0 für n gegen ∞! Das heißt: Die Wahrscheinlichkeit, dass der Mittelwert von n Werten von X den Erwartungswert von X um mindestens $\epsilon > 0$ verfehlt, strebt gegen 0 für n gegen ∞. Dies ist Bernoullis Schwaches Gesetz der großen Zahl. Bei Mittelwerten realer Messwerte stellt man eben dieses Verhalten fest, das wir gerade für die Zufallsgrößen der W-Theorie bewiesen haben - ein Hinweis auf eine Anwendbarkeit der W-Theorie auf die Wirklichkeit.

14.2 Bernoulli-Versuche und binomialverteilte Zufallsgrößen

14.2.1 Definition und Eigenschaften

Wenn Ω nur zwei Elemente hat, nennt man das eine Erfolg, das andere Misserfolg, und bezeichnet die zugehörigen Wahrscheinlichkeiten mit p und mit $q = 1-p$. Den Zufallsversuch selbst nennt man dann Bernoulli-Versuch. Wiederholt man den Bernoulli-Versuch n-mal unabhängig, nennt man dies eine Bernoulli-Kette der Länge n, und man interessiert sich für die Anzahl X der Erfolge, die dabei auftreten. Für die Werte $k = 0, 1, \ldots, n$, die X annehmen kann, hat man die Wahrscheinlichkeiten

$$B(n, p, k) := P(X = k) = \binom{n}{k} p^k q^{n-k} \tag{33}$$

Die Zufallsgröße X heißt binomialverteilt mit den Parametern n und p, kurz $B(n,p)$-verteilt.

Der Erwartungswert eines Bernoulli-Versuchs ist p, die Varianz pq. Erwartungswert und Varianz von X selbst berechnet man bequem mit Hilfe der Regeln (wie oben im Abschnitt über Bernoullis Schwaches Gesetz der großen Zahl) zu

$$E(X) = np \quad \text{und} \quad V(X) = npq \tag{34}$$

14.2.2 Die integrale Näherungsformel von de Moivre und Laplace

Für große n und k ist die Berechnung von $P(X = k)$ und besonders für $P(X \leq k)$ sehr aufwendig (obwohl man mit einem Rechner recht weit kommt). Hier verwendet man eine Näherungsformel, die man traditionell anwendet, wenn $npq \geq 9$ ist. Beweise laufen mit viel harter Analysis; wir übernehmen hier nur das Ergebnis, und das sieht so aus:

$$P(X \leq k) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{k+0,5-E(X)}{\sigma}} e^{-\frac{1}{2}x^2}\, dx =: \Phi(\frac{k+0,5-E(X)}{\sigma}) \tag{35}$$

Der Hintergrund ist, dass das Histogramm der zu X gehörigen standardisierten Zufallsgröße $Z := \frac{X-E(X)}{\sigma}$ durch die Gauß-Funktion φ angenähert wird, deshalb bekommt man die gewünschte Wahrscheinlichkeit näherungsweise durch ein Integral über die Gauß-Funktion. Dessen Werte muss man freilich wieder einer Tabelle entnehmen. Die Gauß-Funktion φ kennst du aus der Analysis und als Dichtefunktion einer stetig-verteilten Zufallsgröße.

14.2.3 Anwendungen

Binomialverteilte Zufallsgrößen sind für uns so wichtig, weil wir damit praktische Vorgänge modellieren können. Die Anzahl der defekten Stücke in einer Lieferung aus einer Massenfertigung mit konstantem Ausschussanteil kann man als binomialverteilte Zufallsgröße ansehen, das ist das eine wichtige Beispiel. Zweites Beispiel sind Umfragen, bei denen nur mit ja oder nein geantwortet werden kann. Einzelheiten folgen im nächsten Abschnitt.

14.3 Beispiele für Schätz- und Testaufgaben

Damit sind die Grundaufgaben der beurteilenden Statistik genannt. Ich gebe ein Beispiel. Der Bürgermeister lässt n Leute befragen, ob sie mit ihm zufrieden sind. Welche Schlüsse auf den Anteil der Zufriedenen lässt ein solches Umfrageergebnis zu? Beispiele für solche Schätzaufgaben, auch welche mit großer praktischer Bedeutung, gibt es in Fülle.

Zweites Beispiel: Lieferant und Kunde verabreden, dass eine Stichprobe aus einer Lieferung gezogen und untersucht wird. Finden sich darin höchstens k fehlerhafte Stücke, passiert dies, sonst das. Eine Lieferung mit Ausschussanteil p wird dann mit der Wahrscheinlichkeit $P(X_p \leq k)$ angenommen und mit der Wahrscheinlichkeit $P(X_p > k)$ abgelehnt. Diese Wahrscheinlichkeiten sind Funktionen von p, und wir können ausnutzen, dass wir mit Funktionen umzugehen verstehen.

14.4 Parameterschätzung: Das Konfidenzintervall

Sagen wir, 110 von 200 Leuten geben bei einer Umfrage an, dass sie mit der Arbeit ihres Bürgermeisters zufrieden sind. Was kann man daraus schließen? Man kann dem Bürgermeister natürlich sagen, dass eine Umfrage ergeben habe, dass 55% der Bevölkerung mit ihm einverstanden sind, aber es ist klar, dass diese Angabe mit einer Unsicherheit behaftet ist und dass die Unsicherheit kleiner wäre, hätten sich 330 von 600 Leuten für ihn ausgesprochen.

Zäumen wir einmal das Pferd von hinten auf: Wenn der Anteil der Leute in der Bevölkerung, die mit der Politik des Bürgermeisters einverstanden sind, p_0 beträgt und wir n zufällig gewählte Bürger nach ihrer Meinung befragen, können wir so tun, als hätten wir es mit einer Bernoulli-Kette der Länge n mit Erfolgswahrscheinlichkeit p_0 zu tun. Die Anzahl X der Ja-Stimmen ist dann eine $B(n, p_0)$-verteilte Zufallsgröße. Theoretisch kann, wenn p_0 nicht 0 oder 1 ist, jeder der Werte von 0 bis n herauskommen, aber der Löwenanteil der Wahrscheinlichkeit konzentriert sich auf die Werte, die nicht zu weit vom Erwartungswert weg liegen. Man fasst sich nun ein Herz und trifft die folgende

Annahme: Der erhaltene Wert k von X liegt in der 95%-Umgebung des Erwartungswerts

und ist sich dabei der Tatsache bewusst, dass diese Annahme mit der Wahrscheinlichkeit 5% falsch ist. Diese 5% heißen Irrtumswahrscheinlichkeit.[15] Dieser Schritt mag dir gegen den Strich gehen, er ist aber unumgänglich. Wer kein Irrtumsrisiko in Kauf nehmen will, kann nicht mit Hilfe von Wahrscheinlichkeitstheorie versuchen, aus unvollständigen Datenlagen das Beste zu machen.

[15] Man hätte statt 95% natürlich auch 99% oder sonstwas nehmen können.

Man bestimmt nun alle Wahrscheinlichkeiten p mit der Eigenschaft, dass der beobachtete Wert k von X in der 95%-Umgebung des Erwartungswerts np der $B(n,p)$-verteilten Zufallsgröße X_p liegt. Diese p bilden ein Intervall, das sogenannte **Konfidenzintervall** zum Umfragergebnis k und zur Irrtumswahrscheinlichkeit 5%.

Am einfachsten geht die Berechnung der Grenzen des Konfidenzintervalls so, wie Bernd es vorgeschlagen hat: Wir berechnen

$$r = \Phi^{-1}(0,975) \approx 1,956 \tag{36}$$

- das ist der Radius der 95%-Umgebung des Erwartungswerts der zu X_p gehörigen standardisierten Zufallsgröße, wie du dir am Graphen der Gaußfunktion φ klarmachst. Der gesuchte Radius der 95%-Umgebung des Erwartungswerts von X_p selbst ist dann

$$r\sigma = r\sqrt{np(1-p)} \tag{37}$$

Das Umfrageergebnis k liegt in der $r\sigma$-Umgebung des Erwartungswerts np von X_p, wenn

$$|k - np| \leq r\sigma \tag{38}$$

ist. Quadriert man dies und ersetzt das $\leq$ durch $=$, erhält man eine quadratische Gleichung, deren beide Lösungen das größte und das kleinste zulässige p sind. Die Rechnung ist überaus lästig. Maple hat das für mich erledigt, für $k = 110$ und $n = 200$ ist das Konfidenzintervall $[0,481; 0,617]$.

Die Aussage, laut Umfrage läge der Anteil der Anhänger des Bürgermeisters mit der W. 95% zwischen 48% und 62% ist so suggestiv und verbreitet wie blödsinnig. Wahrscheinlichkeitsaussagen sind grundsätzlich nur über Ergebnisse von Zufallsversuchen möglich, der Anteil der Anhänger des Bürgermeisters ist aber nicht Ergebnis eines Zufallsversuchs. Die korrekte Auskunft lautete so: Das gelieferte Konfidenzintervall ist mit Hilfe eines Verfahrens bestimmt worden, das mit der W. 95% ein Intervall ergibt, das das wahre p_0 überdeckt. Kaum einem Anwender kann man eine so filigrane Expertise zumuten, das gebe ich zu, aber billiger geht es nicht, wenn man redlich bleiben will.

Die Folgen liegen auf der Hand: von den 95%-Konfidenzintervallen, die eine Firma am laufenden Band liefert, sind im Mittel 5% falsch. Die Fehlerquote bei den 10.000 Behauptungen aufgrund statistischer Untersuchungen, die täglich die Nachrichtenagenturen überfluten, halte ich für deutlich höher. Mehr als ein gewisser Unterhaltungswert ist da nicht drin, aber darin liegt ja auch ein gewisser Reiz.

15 Klausur Nr. 4

1) Ein Hersteller von Tennisbällen trifft mit einem großen Kunden die Verabredung, dass eine Lieferung zurückgewiesen wird, wenn unter hundert geprüften Bällen mehr als zwölf Mängel aufweisen.

a) Wie groß ist die Wahrscheinlichkeit, dass eine Lieferung mit einem Anteil $p = 0,1$ nicht einwandfreier Bälle zurückgewiesen wird?

b) Wie groß ist die Wahrscheinlichkeit, dass eine Lieferung mit einem Anteil $p = 0,2$ nicht einwandfreier Bälle unbeanstandet bleibt?

c) Zeichne aus der Hand, wie du dir den Graphen der Funktion

$$f(p) = P(X_p \leq 12)$$

vorstellst (dabei ist X_p natürlich $B(100,p)$-verteilt).

2) a) Die Zufallsgröße X nimmt die Werte 5 und -5 jeweils mit der Wahrscheinlichkeit $\frac{1}{2}$ an. Bestimme Erwartungswert und Varianz von X und zeichne die Graphen der Funktionen $f(a) = V(X)/a^2$ und $w(a) = P(|X - E(X)| \geq a)$ in ein Schaubild. Was soll dieser Auftrag?

b) Es sei X eine Zufallsgröße, die nur Werte ≥ 0 annimmt. Schreibe den Rechenausdruck für den Erwartungswert $E(X)$ hin, wende die Beweisidee der Tschebyschew-Ungleichung an und schließe, dass

$$P(X \geq a) \leq \frac{E(X)}{a}$$

sein muss.

c) Die Zufallsgröße X sei stetig verteilt mit Dichtefunktion

$$f(x) = \begin{cases} \frac{1}{x^2} & \text{für } x \geq 1 \\ 0 & \text{sonst} \end{cases}$$

Versuche, den Erwartungswert $E(X)$ zu berechnen.

3) Eine Urne enthält fünf rote und 10 weiße Kugeln. Es wird n-mal eine Kugel gezogen, die gezogene Kugel wird **nicht** zurückgelegt.

a) Berechne die Wahrscheinlichkeit (bei (iv) und (v) genügen Rechenausdrücke), dass
(i) die erste gezogene Kugel rot ist;
(ii) die zweite gezogene Kugel rot ist;
(iii) die dritte gezogene Kugel rot ist;
(iv) $rrwr$ gezogen wird;
(v) von den ersten vier gezogenen Kugeln genau drei rot sind.

b) Es sei nun X die Anzahl der weißen unter den n gezogenen Kugeln. Wir suchen $E(X)$. Bezeichne dazu die Anzahl der weißen Kugeln bei der i-ten Ziehung mit X_i und begründe, dass $E(X_i) = p = \frac{2}{3}$ ist. Dann erhältst du $E(X)$ über die Summenregel, wie üblich.

c) Jetzt denkst du vielleicht, auch $V(X)$ fiele dir für das X aus b) in den Schoß. Aber $V(X)$ ist **nicht** $nV(X_i)$. Ist unsere Summenregel für die Varianz falsch?

4) Klaus hat mal wieder in der Schule gefehlt, und nun hat er Mühe, den Stoff nachzuarbeiten. "Das Histogramm der zur Zufallsgröße X gehörigen Zufallsgröße Z sieht etwa so aus wie der Graph der Gauß-Funktion φ, und deshalb kann man mit Hilfe des Integrals über die Gauß-Funktion Wahrscheinlichkeiten $P(a \leq X \leq b)$ ausrechnen", sagte man ihm. Hm, erstmal ein Beispiel. Klaus nimmt für X eine $B(2, \frac{1}{2})$-verteilte Zufallsgröße und zeichnet ihr Histogramm und das der zugehörigen standardisierten Zufallsgröße Z. Dann berechnet er $P(0 \leq X \leq 2)$ mit Hilfe von ϕ. Natürlich ist er mit dem Ergebnis nicht sonderlich zufrieden. Er beschwert sich bei seinem Mitschüler Rudi: "Wie soll das denn überhaupt funktionieren? Das Z nimmt nur Werte in einem begrenzten Bereich an, aber $\varphi(x)$ ist für jedes $x \in \mathbb{R}$ positiv." Zum Glück weiß Rudi Bescheid. "Stimmt", antwortet er. "Bleibe meinetwegen bei $p = \frac{1}{2}$. Dann musst du das n aber so groß wählen, dass die Werte von Z wenigstens den Bereich von -3 bis 3 abdecken. Sonst wird der Fehler zu groß."

a) Welchen Satz hatte der Kurs in Klaus´ Fehlstunde gelernt?

b) Was steht auf Klaus´ Zettel? Zeichne die Histogramme seines X und seines Z und berechne die Wahrscheinlichkeit $P(0 \leq X \leq 2)$ auf Klaus´ Art.

c) Nimm zu Klaus´ Einwand Stellung!

d) Wie groß ist die Wahrscheinlichkeit, die Rudi vernachlässigt, wenn er φ nur im Bereich von -3 bis 3 betrachtet?

e) Wie groß soll Klaus das n nach Rudis Rat wählen? Berechne den Wert.

f) Wie groß ist die Varianz des X zu $p = \frac{1}{2}$ und dem in e) berechneten n? Sagt dir das Ergebnis etwas?

16 Nachschreibklausur Nr. 4

1) a) Es sei X eine $B(3, \frac{1}{3})$-verteilte Zufallsgröße. Zeichne das Histogramm von X und das der zugehörigen standardisierten Zufallsgröße.

b) Nenne ein Beispiel für eine $B(10, \frac{1}{6})$-verteilte Zufallsgröße X, gib die Formel für $P(X = k)$ an und begründe, dass sie richtig ist. Gib einen Beweis für $E(X) = \frac{10}{6}$.

2) Die Zufallsgröße X sei stetig verteilt mit der Dichtefunktion

$$f(x) = \begin{cases} \frac{2}{x^3} & \text{für } x \geq 1 \\ 0 & \text{sonst} \end{cases}$$

Berechne $E(X)$ (Ergebnis: 2) und versuche, $V(X)$ zu berechnen.

3) "Jedes dritte Los ein Treffer", lautet die Werbung einer Lotteriegesellschaft. Der Verein "Kampf dem unlauteren Wettbewerb" lässt 100 Lose untersuchen.

a) Wieviel Gewinnlose kann man vernünftigerweise unter den 100 Losen erwarten?

b) Wenn weniger als zwanzig Gewinnlose gefunden werden sollten, will der Verein eine Pressekampagne gegen die Lotteriegesellschaft starten. Wie groß ist die Wahrscheinlichkeit, dass der Lotteriebetreiber Ziel einer Kampagne wird, obwohl seine Werbung wahr ist?

c) In der Stichprobe werden 42 Gewinnlose gefunden. Was kann man über den Anteil der Gewinnlose aussagen?

Hinweis: Die Zahl der Lose soll so riesig sein, dass die Entnahme der Stichprobe keinen Einfluss auf den Anteil der Gewinnlose hat.

4) In einem Kursraum mit $100\,m^3$ Volumen befinden sich, sagen wir, 10^{26} Luftteilchen. Diese sind in ständiger regelloser Bewegung; es ist unmöglich vorherzusagen, an welcher Stelle des Kursraums sich ein Luftteilchen in einer Minute befinden wird, wenn man weiß, wo es jetzt gerade ist. Deshalb kann man gut sagen, dass der Ort des Teilchens in einer Minute nur vom Zufall abhängt.

a) Wir grenzen in Gedanken einen Bereich von $2m \times 1m \times 1m$ des Kursraums ein. Da sitzt du. Mit welcher Wahrscheinlichkeit befindet sich ein bestimmtes Luftteilchen in einer Minute in deinem Bereich?

b) Wieviele Teilchen sollten im Mittel in deinem Bereich sein?

c) Mit welcher Wahrscheinlichkeit befinden sich in einer Minute zufällig **alle** Luftteilchen in deinem Bereich? (schlecht für deine Kollegen, die haben keine Luft zum Atmen)

d) Wie groß ist die Wahrscheinlichkeit, dass die Anzahl der Teilchen in deinem Bereich vom Mittelwert aus b) um mindestens 1% abweicht

(i) näherungsweise

(ii) maximal nach Tschebyschew?

d) Was sagt das Schwache Gesetz der großen Zahl aus?

Teil III

13.1: Lineare Algebra

17 Einstieg ins Halbjahrsthema

Liebe Leute, ich wünsche euch, dass ihr schöne Ferien hattet, und ich wünsche euch ein gutes letztes Schuljahr! Da ich gleich eine ganze Woche weg bin, seid ihr einige Tage auf euch gestellt: ihr sollt euch einen Gegenstand selbst erarbeiten.

Unser Gebiet heißt "Lineare Algebra". Algebra handelte ursprünglich vom Lösen von Gleichungen, und die Gleichungen der Linearen Algebra sind linear, das heißt, es kommen keine Quadrate oder Wurzeln der Unbekannten oder noch Schlimmeres vor. Was dann noch bleibt, ist so schlicht, dass du vielleicht schon befürchtest, man käme dir in der 13 mit trivialen Inhalten. Gemach, Systeme linearer Gleichungen haben eigentümliche Eigenschaften, die reichlich Stoff für Forschungen geben.

Aber zunächst nur Syteme linearer Gleichungen, abgekürzt LGS. Nehmt das Buch, lest die Einleitung auf der ersten Seite bis (ausschließlich) A, werft nur einen Blick auf die Beispiele in A (seht daran, dass LGS vielerorts vorkommen) und geht zu B auf Seite 10. Lest euch den Text von B durch, arbeitet dann die Paragraphen 1.2 und 1.3 durch. Löst dann die Aufgabe 1.2 a) auf S. 17 und das System mit den Büffelpreisen auf S. 11. Das ist schon alles.

Wenn du fertig bist, kennst du Beispiele für LGS und ein Lösungsverfahren dafür, nämlich den Gaußschen Algorithmus. Das ist das Standardverfahren, und das wird zur Lösung von LGS benutzt. Kraut- und Rüben-Methoden sind nicht zulässig. Halte dich auch strikt an die Schreibweise, schreibe die Variablen schön untereinander, wie das im Buch in Aufgabe 1.2 gemacht ist. Und notiere hinter jeder Gleichung eines Systems beim Lösen, wie du sie bekommst. Wenn du die dritte Gleichung des bearbeiteten Systems als Summe der dritten Gleichung und des 7-fachen der zweiten Gleichung des vorigen systems bekommst, schreibe dahinter

$$|III + 7II \quad .$$

Und es wird immer das ganze System geschrieben, Biotope von Gleichungen werden nicht geduldet.

Vielleicht wundert dich, dass ich so auf einem Formalismus beharre. An den LGS hängen die Begriffe Vektor und Matrix, Matrixabbildung, Vektorraum, Basis und Dimension und viele andere: eben die ganze Lineare Algebra. Es ist eine im Vergleich zur Analysis recht junge Theorie, und sie ist sehr erfolgreich. Ob man ein Ziel erreicht oder nicht, ob man zu einem wichtigen Begriff gelangt, ihn sieht und versteht oder nicht, hängt in der Linearen Algebra oft daran, ob man eine angemessene Schreibweise benutzt oder nicht.

Damit habe ich das Programm genannt bis auf einen wichtigen Punkt: Es ist ungeheuer fruchtbar, die Begriffe der Linearen Algebra geometrisch zu deuten. Manches kann man ohne diese geometrische Deutung überhaupt nicht angemessen verstehen. Die Geometrie ihrerseits erfährt duch die Lineare Algebra eine Ausdehnung auf höherdimensionale Räume. Das ist eine spannende Sache, du wirst sehen.

18 Vektoren und Matrizen

18.1 Die Vektor-Matrix-Schreibweise eines LGS

Sagen wir, nach der Anwendung des Gaußschen Algorithmus sehe unser Gleichungssystem so aus:

$$\begin{array}{ccccccc} x & + & y & - & z & = & 4 \\ & & y & + & 2z & = & -3 \end{array}$$

Dies ist ein 2×3-LGS: es hat zwei Gleichungen und drei Variable. Eine Lösung finden wir leicht: wir können z frei wählen, z.B. $z = 0$. Dann ist $y = -3$ und schließlich $x = 7$. Das ist aber nur **eine** Lösung, deshalb erfindet man neue Objekte, um die Lösung angemessen schreiben zu können. Man fasst die drei Werte in einer Dreierspalte zusammen. Die Stelle, an der ein Wert steht, zeigt an, zu welcher Variablen er gehört. Diese Spalten heißen **Vektoren**, man setzt zur Kennzeichnung einen Pfeil über den Namen des Vektors. Unsere Lösung schreiben wir dann so:

$$\vec{a} := \begin{pmatrix} 7 \\ -3 \\ 0 \end{pmatrix} \in \mathbb{R}^3$$

Man erfasst nun das Koeffizientenschema der linken Seite in Form einer 2×3-Matrix A. Die sieht so aus:

$$A := \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 2 \end{pmatrix} \tag{39}$$

Eine $m \times n$-Matrix ist einfach ein rechteckiges Zahlenschema mit m Zeilen und n Spalten. Die Zahlen darin heißen die Einträge der Matrix, der Eintrag in der i-ten Zeile und der j-ten Spalte heißt a_{ij}.

Mit der rechten Seite

$$\vec{b} := \begin{pmatrix} 4 \\ -3 \end{pmatrix} \in \mathbb{R}^2$$

können wir unser LGS kurz so schreiben:

$$A\vec{x} = \vec{b} \quad . \tag{40}$$

Dies ist die Vektor-Matrix-Schreibweise des LGS.

18.2 Rechnen mit Vektoren und Matrizen

Unsere neuen Objekte, nämlich Vektoren und Matrizen, gestatten eine Kurzschreibweise für LGS. Das ist zwar schön, aber entscheidend ist etwas ganz anderes: Mit den neuen Objekten **rechnet** man. Beginnen wir mit den Vektoren[16]. Die Menge aller Vektoren mit n Einträgen heißt der Spaltenraum $\mathbb{R}^n$. Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ und $r \in \mathbb{R}$ definiert man Summe und Vielfaches auf diese Weise:

$$\vec{x} + \vec{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \qquad r\vec{x} = r \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} rx_1 \\ rx_2 \\ \vdots \\ rx_n \end{pmatrix} \tag{41}$$

Eine $m \times n$ Matrix $A = (a_{ij})$ kann man mit einem n-Vektor $\vec{x}$ multiplizieren. Das Ergebnis ist ein m-Vektor $\vec{b}$:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{42}$$

Dabei sind die Einträge von $\vec{b}$ auf folgende Weise festgelegt:

$$b_i := \sum_{j=1}^{n} a_{ij} x_j \tag{43}$$

18.3 Die Struktur der Lösungsmenge unseres LGS

Du magst dich vielleicht darauf einlassen, dass man den $\mathbb{R}^n$ einführt, damit eine Lösung eines LGS auch nur ein Ding ist. Aber warum rechnet man mit diesen neuen Objekten? Der Grund ist der folgende. Eine Menge, mit deren Elementen man rechnen kann, ist nicht mehr nur ein Sammelsurium von Gegenständen, sondern sie trägt eine Struktur. Den Nutzen sollst du sogleich sehen.

Sämtliche Lösungen unseres LGS gewinnen wir nach Markus′ Methode. Wir wählen $z \in \mathbb{R}$ beliebig. Dann muss $y = -3 - 2z$ sein, damit die zweite Gleichung erfüllt ist, und anschließend

$$x = 4 - y + z = 4 - (-3 - 2z) + z = 7 + 3z$$

sein, damit auch die erste Gleichung erfüllt ist. Somit erhalten wir als Lösungsmenge des LGS die Menge

$$M = \left\{ \begin{pmatrix} 7 + 3z \\ -3 - 2z \\ z \end{pmatrix} \;\middle|\; z \in \mathbb{R} \right\}$$

[16] Einen Vektor kannst du als Spezialfall einer Matrix auffassen, nämlich als einspaltige Matrix.

An dem von z abhängigen Vektor in M schrauben wir etwas herum, und zwar zerlegen wir ihn in einen festen und in einen von z abhängigen Teil und ziehen aus dem zweiten Teil das z heraus.

$$\begin{pmatrix} 7+3z \\ -3-2z \\ z \end{pmatrix} = \begin{pmatrix} 7 \\ -3 \\ 0 \end{pmatrix} + \begin{pmatrix} 3z \\ -2z \\ z \end{pmatrix} = \begin{pmatrix} 7 \\ -3 \\ 0 \end{pmatrix} + z \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$$

Was haben wir erreicht? Die Lösungsmenge des LGS lässt sich so darstellen: Man addiert zu einem festen Vektor - dem wir übrigens oben schon begegnet waren - sämtliche Vielfache eines zweiten Vektors. Na und? Nun: Die Vektoren des $\mathbb{R}^3$ können wir uns als Pfeile vorstellen, die vom Nullpunkt des Koordinatensystems zu den Raumpunkten mit den entsprechenden Koordinaten zeigen[17]. Die Menge aller Vielfachen eines Vektors steht dann für die Gerade durch den Nullpunkt und den Punkt des Vektors. Dadurch, dass man einen festen Vektor hinzuaddiert, hängt man die Gerade nur in einem anderen Raumpunkt auf, man verschiebt sie im Raum so, dass sie durch den Punkt des festen Vektors geht.

Fazit: Die Lösungsmenge eines LGS ist nicht irgendwelcher Müll. Es kommen nur Mengen heraus, die für klar definierte geometrische Objekte stehen, nämlich einzelne Punkte, Geraden, Ebenen, ... Dazwischen gibt es nichts. Durch dieses Zusammenspiel zwischen Geometrie und Algebra versteht man die Sache erst richtig. - Jetzt lasst uns das alles in Ruhe betrachten und einüben.

19 Algebra und Geometrie

Wir haben den Spaltenraum $\mathbb{R}^n$ eingeführt. Das ist zunächst die Menge der n-Spalten. Die Menge wird dadurch zu einem "Raum", dass man in der Menge rechnet. Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ und $r \in \mathbb{R}$ haben wir $\vec{x} + \vec{y}$ und $r\vec{x}$ erklärt. Dieser Raum ist der harte Boden der Tatsachen unserer Arbeit. Beweise werden letztlich immer hier geführt: wir treiben abstrakte Algebra.

Dennoch: wie im letzten Abschnitt schon angesprochen wurde, lieferte die Geometrie die Vorlage für die algebraischen Strukturen. Wer mit den Begriffen der Algebra vertraut werden will, tut gut daran, sich das geometrische Modell anzusehen, das die Begriffsbildungen angeregt hat. Die Algebra gewährleistet Verlässlichkeit (und bietet der Geometrie sogar ein festes Fundament), die Geometrie speist die Intuition.

Damit dir dieses Wechselspiel so deutlich wie möglich wird, stelle ich die algebraischen Begriffe und ihre Entsprechungen in der Geometrie in einer Tabelle gegenüber. Dabei seien die Vektoren zunächst aus dem $\mathbb{R}^3$, und wir benutzen ein räumliches Koordinatensystem mit Ursprung O.

[17] Das sind die Ortsvektoren der Physik.

algebraischer Begriff	geometrische Entsprechung
$\vec{x}$	1) Punkt X mit den durch $\vec{x}$ gegebenen Koordinaten 2) Pfeil $\overrightarrow{OX}$ vom Nullpunkt zu X
$\vec{z} = \vec{x} + \vec{y}$	Pfeil $\overrightarrow{OZ}$ im Parallelogramm $OXZY$
$\vec{z} = r\vec{x}$	Pfeil $\overrightarrow{OZ}$ ist $\overrightarrow{OX}$, auf die $\lvert r\rvert$-fache Länge gestreckt. Für $r < 0$ wird die Richtung umgekehrt.
$\langle\vec{x}\rangle := \{r\vec{x} \,\vert\, r \in \mathbb{R}\}$	Gerade durch O und X (für $\vec{x} \neq \vec{0}$)
$\langle\vec{x}, \vec{y}\rangle := \{r\vec{x} + s\vec{y} \,\vert\, r, s \in \mathbb{R}\}$	Ebene durch O, X und Y - falls die drei Punkte nicht auf einer Geraden liegen.
$\vec{a} + \langle\vec{b}\rangle$	Gerade durch A parallel zu OB (falls $\vec{b} \neq \vec{0}$)
$\vec{a} + \langle\vec{b}, \vec{c}\rangle$	Ebene durch A, B und C - falls $\vec{b}$ und $\vec{c}$ eine Ebene erzeugen
?	Länge einer Strecke $\overline{AB}$
?	senkrechtstehen (Orthogonalität)
?	Winkel
?	sauberer Dimensionsbegriff
?	Kugel
?	Tangentialebene
?	Abbildungen: Drehung, Streckung, Projektion, Spiegelung

Was erst einmal für den $\mathbb{R}^3$ etabliert ist, schieben wir auf den $\mathbb{R}^n$ hoch und erhalten so eine n-dimensionale Geometrie.

20 Das Skalarprodukt

Das Skalarprodukt öffnet das Tor zu neuen Möglichkeiten: Man kann mit seiner Hilfe Längen von Vektoren und Winkel zwischen Vektoren im $\mathbb{R}^n$ bilden und erhält so eine Geometrie in einem n-dimensionalen Raum. Und eine Reihe schwieriger Dinge werden ganz leicht, wenn man das Skalarprodukt verwendet, als habe man sie mit einem Zauberstab berührt.

20.1 Definition und algebraische Eigenschaften

In unserem Buch steht in §4.2 eine schöne Motivation, wie Fragen der räumlichen Geometrie zum Skalarprodukt führen. Ich kann hier also mit der Definition beginnen, und ich formuliere sie gleich für den $\mathbb{R}^n$.

1 Definition *Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ definieren wir das Skalarprodukt $\vec{x} * \vec{y}$ durch*

$$\vec{x} * \vec{y} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} * \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := x_1 y_1 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i$$

Beachte, dass das Skalarprodukt zweier Vektoren stets eine reelle Zahl ist - es handelt sich also keineswegs um eine Multiplikation innerhalb des $\mathbb{R}^n$.

Wir notieren sofort einige algebraische Eigenschaften des Skalarprodukts. Wenn du die Produkte ausschreibst, siehst du sofort ein, dass sie gelten.

2 Lemma *Für $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ und $r \in \mathbb{R}$ gilt*

$$\vec{x} * (\vec{y} + \vec{z}) = \vec{x} * \vec{y} + \vec{x} * \vec{z} \qquad \textit{und} \qquad (r\vec{x}) * \vec{y} = \vec{x} * (r\vec{y}) = r(\vec{x} * \vec{y})$$

Das heißt, mit dem Skalarprodukt kann man in gewissen Grenzen vernünftig rechnen.

20.2 Skalarprodukt und räumliche Geometrie

Das Skalarprodukt im $\mathbb{R}^3$ passt in der folgenden Weise zur räumlichen Geometrie.

1. Die Länge $|\vec{x}|$ des Pfeils, der für den Vektor $\vec{x} \in \mathbb{R}^3$ steht, ist gegeben durch $\sqrt{\vec{x} * \vec{x}}$.
2. Die Pfeile, die für $\vec{x}, \vec{y} \in \mathbb{R}^3$ stehen, bilden genau dann einen rechten Winkel, wenn $\vec{x} * \vec{y} = 0$ ist.[18]
3. Für $\vec{a} \in \mathbb{R}^3$, $\vec{a} \neq \vec{0}$ bildet die Menge aller $\vec{x}$ mit $\vec{a} * \vec{x} = 0$ eine Ebene durch den Nullpunkt.[19]

Ein Beispiel zum letzten Punkt: Für

$$\vec{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad \text{ist} \qquad \vec{a} * \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \iff x + 2y + 3z = 0 \quad .$$

Die Lösungsmenge dieses 1×3-LGS ist

$$\left\{ \begin{pmatrix} -2y - 3z \\ y \\ z \end{pmatrix} \,\middle|\, y, z \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \right\rangle$$

20.3 Normalenform einer Ebene im Raum

Zeichne dir unbedingt auf, was ich jetzt beschreibe!

Nimm einen Vektor $\vec{n} \in \mathbb{R}^3$, $\vec{n} \neq \vec{0}$. Es sei V die Menge aller Vektoren $\vec{x}$ mit $\vec{n} * \vec{x} = 0$. Dann steht V für eine Ebene durch den Nullpunkt, wie wir gerade gesehen haben. Nimm nun einen weiteren Vektor $\vec{a}$, der aber nicht in V liegt. Dann ist $\vec{a} * \vec{n}$ eine Zahl $c \neq 0$. Wenn du zu $\vec{a}$ einen Vektor $\vec{b} \in V$ addierst und die Summe mit $\vec{n}$ skalar multiplizierst, erhältst du wieder c:

$$(\vec{a} + \vec{b}) * \vec{n} = \vec{a} * \vec{n} + \vec{b} * \vec{n} = \vec{a} * \vec{n} + 0 = \vec{a} * \vec{n} = c$$

Und umgekehrt: Ist für einen Vektor $\vec{d}$ das Skalarprodukt $\vec{d} * \vec{n}$ gleich $\vec{a} * \vec{n}$, dann ist $\vec{d} \in \vec{a} + V$:

$$\vec{d} * \vec{n} = \vec{a} * \vec{n} \iff \vec{d} * \vec{n} - \vec{a} * \vec{n} = 0 \iff (\vec{d} - \vec{a}) * \vec{n} = 0 \iff \vec{d} - \vec{a} \in V \iff \vec{d} \in \vec{a} + V$$

Es gilt also der folgende Satz.

[18] Um dieses "genau dann" zu haben, nimmt man in Kauf, dass der "Nullpfeil" zu allen Pfeilen senkrecht steht. Dies erspart auch lästige Sonderbetrachtungen beim Rechnen.

[19] Nähme man den "Nullpfeil" aus, hätte die Ebene ein Loch.

3 Lemma *Für $\vec{n} \in \mathbb{R}^3$, $\vec{n} \neq \vec{0}$ und $c \in \mathbb{R}$ ist die Lösungsmenge der Gleichung $\vec{n} * \vec{x} = c$ eine Ebene im Raum. Die von $\vec{n}$ erzeugte Gerade $\langle\vec{n}\rangle$ steht auf dieser Ebene senkrecht.*

Den Vektor $\vec{n}$ des Lemmas nennt man einen **Normalenvektor** der Ebene, die Gleichung $\vec{x} * \vec{n} = c$ heißt eine Normalenform der Ebene.

20.4 Anwendungsbeispiel

Wir betrachten die Ebene im Raum, die durch die Punkte $A(4|0|0)$, $B(0|2|0)$ und $C(0|0|6)$ geht, und wir suchen den Fußpunkt des Lots vom Koordinatenursprung auf diese Ebene.

Ein Normalenvektor $\vec{n}$ der Ebene muss auf den Vektoren $\vec{b} - \vec{a}$ und $\vec{c} - \vec{a}$ senkrecht stehen. Dies ist gleichbedeutend mit

$$(\vec{b} - \vec{a}) * \vec{n} = 0 \qquad \text{und} \qquad (\vec{c} - \vec{a}) * \vec{n} = 0$$

Das ist schlicht das 2×3-LGS

$$\begin{array}{rrrl} -4 & 2 & 0 & |0 \\ -4 & 0 & 6 & |0 \end{array}$$

und dessen Lösungsmenge ist

$$\left\{ n_3 \begin{pmatrix} \frac{3}{2} \\ 3 \\ 1 \end{pmatrix} \,\middle|\, n_3 \in \mathbb{R} \right\}$$

Als Normalenvektor nehmen wir bequemlichkeitshalber $\vec{n} = \begin{pmatrix} 3 \\ 6 \\ 2 \end{pmatrix}$. Und in der Tat ist nun

$$\vec{a} * \vec{n} = \vec{b} * \vec{n} = \vec{c} * \vec{n} = 12$$

Unsere Ebene ist genau die Lösungsmenge der Gleichung $\vec{x} * \vec{n} = 3x + 6y + 2z = 12$.

Wie kommen wir nun an den Fußpunkt F des Lotes vom Nullpunkt auf die Ebene? Dieser Punkt muss zwei Bedingungen genügen. Er muss auf der von $\vec{n}$ erzeugten Geraden liegen und er muss in der Ebene liegen. Also muss $\vec{f} = r\vec{n}$ sein, und $\vec{f}$ muss der Gleichung der Ebene genügen. Ich setze einfach $r\vec{n}$ für $\vec{x}$ in die Ebenengleichung ein. Dann ist

$$(r\vec{n}) * \vec{n} = 12, \qquad \text{also} \qquad r(\vec{n} * \vec{n}) = 12, \qquad \text{also} \qquad r = \frac{12}{49}$$

Wir erhalten $\vec{f} = \frac{12}{49}\vec{n}$. Daraus können wir auch sofort den Abstand der Ebene vom Nullpunkt berechnen, er ist $|\vec{f}|$.

21 Klausur Nr.1

Hier ist eine Liste der Vektoren bzw. Punkte, die du in den folgenden Aufgaben brauchst.

$$A = \begin{pmatrix} -1 \\ 4 \\ 2 \end{pmatrix} \quad B = \begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix} \quad C = \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix} \quad D = \begin{pmatrix} 11 \\ 0 \\ 0 \end{pmatrix} \quad P = \begin{pmatrix} 50 \\ 20 \\ -20 \end{pmatrix}$$

1) Beschreibe die folgenden Punktmengen im Koordinatensystem in Worten oder durch eine

Skizze. Eine exakte Zeichnung wird nicht erwartet.

$$M_1 = \left\{ \begin{pmatrix} 1+r \\ -r \\ 2+3r \end{pmatrix} \;\middle|\; r \in \mathbb{R} \right\} \quad M_2 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x^2+y^2+z^2 \leq 1 \right\}$$

$$M_3 = \left\{ \begin{pmatrix} 2\cos t \\ 2\sin t \\ z \end{pmatrix} \;\middle|\; 0 \leq t < 2\pi, -5 \leq z \leq 5 \right\} \quad M_4 = \{ r\vec{a} \mid 0 \leq r \leq 1 \}$$

$$M_5 = \{ r\vec{a} + s\vec{b} \mid 0 \leq r, s \leq 1 \} \quad M_6 = \{ r\vec{a} + s\vec{b} + t\vec{c} \mid 0 \leq r, s, t \leq 1 \}$$

2) Geometrie im Raum.
a) Zeichne das Dreieck ABC und verbinde seine Eckpunkte mit dem Koordinatenursprung O.
b) Berechne die Längen der Strecken $\overline{OA}$ und $\overline{BC}$, den Winkel zwischen den Strecken $\overline{OA}, \overline{OC}$ und den Winkel α des Dreiecks ABC.
c) Die Ebene durch die Punkte A, B und C nennen wir E. Berechne einen Normalenvektor dieser Ebene.
d) Stelle die Ebene E durch A, B und C auf zwei Arten dar.
e) Berechne den Abstand von A zur Geraden durch B und C und den Abstand des Koordinatenursprungs von der Ebene E durch A, B und C. Erläutere deine Ansätze angemessen!
f) Schreibe eine Gleichung für die Ebene durch D hin, die zur Ebene aus e) parallel ist.
g) Entscheide, ob der Punkt P auf der gleichen Seite der Ebene aus e) liegt wie der Koordinatenursprung.
h) Bestimme Punkte auf der Geraden durch A und B, die vom Koordinatenursprung die Entfernung 81 haben. Wie viele müsste es geben?
i) Schneidet die Gerade durch D und P die Ursprungsgerade durch A?

3) Lineare Gleichungsysteme.
a) Löse das durch

$$\begin{aligned} \vec{a} * \vec{x} &= 5 \\ \vec{b} * \vec{x} &= 5 \\ \vec{c} * \vec{x} &= 5 \end{aligned}$$

gegebene LGS. Kannst du das Ergebnis geometrisch interpretieren?
b) Schreibe die Lösungsmengen der folgenden LGS hin und gib eine geometrische Interpretation. Für $\vec{x}$ musst du jeweils einen Vektor der richtigen Größe ansetzen.

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} * \vec{x} = 0 \qquad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} * \vec{x} = 0$$

4) Neue Abenteuer von Hinz und Kunz.
a) Hinz und Kunz betrachten irgend zwei Geraden $g : \vec{x}(r) = \vec{a} + r\vec{b}$ und $h : \vec{y}(s) = \vec{c} + s\vec{d}$ (allgemeine Vektoren, **nicht** die von oben). Hinz hat ausgerechnet, dass sich die Geraden schneiden und dass die Richtungsvektoren zu einander senkrecht sind. Deshalb meint er, dass g und h auf einander senkrecht stehen. Kunz brummelt, dass aber zum Beispiel $\vec{x}(1) * \vec{y}(1)$ keineswegs Null ergebe. Deshalb seien die Geraden auch nicht senkrecht zu einander. Was sagst du dazu?
b) Hinz hat sich etwas Tolles ausgedacht. Er nimmt den Vektor $\vec{a}$ von oben. Kunz könne sich einen Vektor $\vec{x} \in \mathbb{R}^3$ beliebig aussuchen. Hinz will dann eine Zahl r und einen Vektor $\vec{y}$ senkrecht zu $\vec{a}$ finden, so dass

$$\vec{x} = r\vec{a} + \vec{y}$$

ist. Das r könne man ganz leicht berechnen, man müsse nur beide Seiten der Gleichung skalar mit $\vec{a}$ multiplizieren.
Begründe, dass Hinz recht hat, und führe die Sache für $\vec{x} = \vec{b}$ (von oben) durch.

22 Zwei Aufgaben aus dem Abitur 1999

22.1 Aufgabe 1

Sperling Alfred, vom Größenwahn gepackt, fliegt Loopings. Sieh dir nur seine neueste Heldentat an, die der Flugschreiber aufgezeichnet hat. Zur Zeit t befand sich Alfred im Punkt

$$P(t) = (x(t), y(t)) \quad \text{mit} \quad x(t) = 4 - t^2 \quad \text{und} \quad y(t) = 4t - t^3 \quad .$$

a) In welcher Richtung hat Alfred die Bahn durchflogen?
b) Wann war er im Nullpunkt des Systems, wann in dem Punkt mit dem größten x-Wert?
c) Wie groß war Alfreds Bahngeschwindigkeit beim Durchfliegen des Nullpunkts?
d) Wann befand sich Alfred im höchsten, wann im tiefsten Punkt der Schleife, und wie sind die Koordinaten dieser Punkte?
e) In welchem Punkt hat Alfreds Flugbahn die Steigung 1?
f) Bestimme die Extremwerte von Alfreds Bahngeschwindigkeit.
g) Berechne den Flächeninhalt des von Alfred umflogenen Flächenstücks.

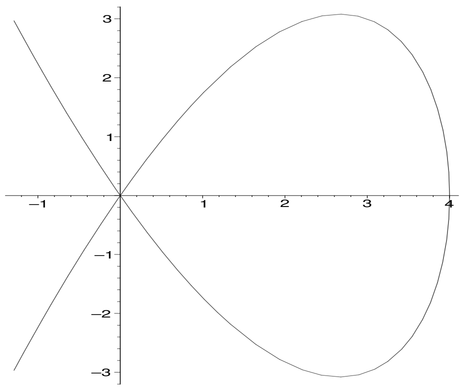


Abbildung 1: Alfreds Flugbahn für $t \in [-2,3; 2,3]$

22.2 Aufgabe 2

Es sei

$$\vec{a} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ -1 \\ -3 \end{pmatrix}, \quad \vec{c} = \begin{pmatrix} 2 \\ -4 \\ 4 \end{pmatrix}$$

und $\epsilon = \langle \vec{a}, \vec{b} \rangle$.
a) Gib eine Gleichung für ϵ an.
b) Rechne nach, daß das LGS $r\vec{a} + s\vec{b} = \vec{c}$ keine Lösung hat. Was bedeutet das?
c) Nun soll die Länge des Vektors

$$\vec{c} - (r\vec{a} + s\vec{b})$$

so klein wie möglich gemacht werden. Interpretiere den Ansatz geometrisch und bestimme die gesuchte minimale Länge.
d) Bestimme die Matrix P der Projektion des Raums auf ϵ, die $\vec{c}$ auf den Nullpunkt abbildet.
e) Zeichne das Parallelepiped, das von den Vektoren $\vec{a}, \vec{b}$ und $\vec{c}$ aufgespannt wird.

23 Matrixabbildungen

23.1 Definition und wichtigste Eigenschaften

Wir kennen LGS in Vektormatrixschreibweise, den allgemeinen Ausdruck findest du in 1.2.4 auf Seite 2. Eigentlich suchte man eine Lösung eines LGS $A\vec{x} = \vec{b}$ für eine $m \times n$-Matrix A. Nun nimmt man einen Standpunktwechsel vor: Zu einem $\vec{x} \in \mathbb{R}^n$ rechnet man einfach $\vec{y} := A\vec{x}$ aus. Dadurch hat man eine sogenannte Matrixabbildung

$$\varphi : \vec{x} \mapsto A\vec{x} \quad \text{oder} \quad \vec{y} = \varphi(\vec{x}) := A\vec{x} \tag{44}$$

gegeben. Warum man das macht, ist in unserem Buch auf Seite 177ff gut beschrieben. Dort sind auch wichtige Eigenschaften von Matrixabbildungen aufgeschrieben und plausibel gemacht. Ich liste hier auf, was mir am wichtigsten ist.

- In den Spalten der Matrix stehen die Bilder der Einheitsvektoren, das heißt $A = (\varphi(\vec{e}_1), \varphi(\vec{e}_2), \ldots, \varphi(\vec{e}_n))$. Dies hilft uns, Matrixabbildungen ganz gezielt zu konstruieren.
- Matrixabbildungen sind linear, das heißt es gilt $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}$ und $A(r\vec{x}) = rA\vec{x}$ für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$ und alle $r \in \mathbb{R}$. Daraus ergibt sich, dass Matrixabbildungen geometrische Objekte vernünftig abbilden. Aus einer Geraden wird wieder eine Gerade (oder ein Punkt), aus einer Strecke $\overline{PQ}$ die Strecke $\overline{P'Q'}$ der Bildpunkte $A\vec{p}$ und $A\vec{q}$ (die allerdings zu einem Punkt entarten kann).

23.2 Drehungen im Raum als Matrixabbildungen

Wir haben uns im Unterricht überlegt, dass Drehungen des Raums um die Achsen des Systems als Matrixabbildungen geschrieben werden können, und wir haben die entsprechenden Matrizen aufgestellt. Die Matrix $M_z(t)$ gehört zu einer Drehung um die z-Achse um den Winkel t (im Bogenmaß). Entsprechend sind $M_y(t)$ und $M_x(t)$ gebildet.

$$M_z(t) = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{45}$$

$$M_y(t) = \begin{pmatrix} \cos t & 0 & \sin t \\ 0 & 1 & 0 \\ -\sin t & 0 & \cos t \end{pmatrix} \tag{46}$$

$$M_x(t) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix} \tag{47}$$

Wir werden diese (und einige andere) Standardmatrizen in Maple vorbereiten, damit wir damit bequem arbeiten können.

Was kann man denn nun damit tun? Eine geometrische Figur, einen Quader zum Beispiel, können wir als Liste von Raumpunkten kodieren, die nach einem bestimmten Schema durch Strecken verbunden werden. Wenn wir ein Bild des gedrehten Körpers haben wollen, multiplizieren wir alle Punkte der Liste (von links!) mit der entsprechenden Drehmatrix. Die Bildpunkte sind dann wieder nach dem Schema durch Strecken zu verbinden. Kurz gesagt: Bildbearbeitungsprogramme des Rechners verwenden Matrizen.

23.3 Abbildungsverkettung und Matrizenmultiplikation

Wenn wir zu einem Vektor $\vec{x}$ den Bildvektor $\vec{y} = A\vec{x}$ unter einer Matrixabbildung berechnen und dann auf dieses $\vec{y}$ eine weitere Matrixabbildung mit Matrix B anwenden, erhalten wir einen Vektor

$\vec{z} := B\vec{y}$. [20] Es gibt dann eine Matrix C, die erlaubt, $\vec{z}$ gleich aus $\vec{x}$ zu berechnen. Sie heißt Produktmatrix von B und A, man schreibt $C = BA$. Wie wir nachgerechnet haben, bekommt man die Matrix C, indem man die Bilder $B\vec{a}_j$ der Spaltenvektoren von A zu einer Matrix zusammenfasst:

$$BA := (B\vec{a}_1, B\vec{a}_2, \ldots, B\vec{a}_n) \tag{48}$$

Man kann auch sagen

$$(b_{ki})(a_{ij}) = (c_{kj}) \quad \text{mit} \quad c_{kj} = \sum_{i=1}^{m} b_{ki}a_{ij} \tag{49}$$

Dabei ist a_{ij} das Element von A, das in der i-ten Zeile an der j-ten Stelle steht.

23.4 Übungen

23.4.1 Verkettung von Drehungen

Rechne nach, dass $M_x(t) = M_y(\pi/2)M_z(t)M_y(-\pi/2)$ ist, und suche den geometrischen Kern dieser algebraischen Identität.

23.4.2 Drehung um eine Würfeldiagonale

Unser Würfel W, den wir verschiedentlich benutzt haben, hat eine Diagonale, die von dem Vektor $\vec{d}$ mit $d_1 = d_2 = d_3 = 1$ erzeugt wird.
a) Finde eine Drehmatrix $M_z(t_1)$, die $\vec{d}$ in einen Vektor $\vec{d'}$ der xz-Ebene dreht, und anschließend eine Drehmatrix $M_y(t_2)$, die $\vec{d'}$ auf die z-Achse dreht.
b) Begründe, dass $M_z(-t_1)M_y(-t_2)M_z(t)M_y(t_2)M_z(t_1)$ eine Drehung um den Winkel t um die Achse $\langle\vec{d}\rangle$ ist.
c) Welche Matrix ergibt sich, wenn du für das t in b) den Wert $t = 2\pi/3$ setzt? Wie wirkt diese Drehung auf die Punkte des Würfels? (Siehe dazu die folgende Aufgabe)

23.4.3 Ein Zeichenprogramm für den Würfel W

Stelle eine Liste der Punkte des Würfels W auf und ein Schema, welche Punkte zu verbinden sind, und schreibe damit eine plot-Anweisung für Maple. Drehe dann den Würfel und lasse auch das Bild des gedrehten Würfels zeichnen. (Tip: Im Paket geom3d von Maple gibt es einen fertigen Befehl zum Zeichnen eines Würfels. Man kann auch anschauen, wie Maple den Würfel kodiert. Probiere faces(W) und vertices von W, wenn du den Würfel als cube(W,sqrt(3)) eingeführt hast. Einen Überblick bekommst du bei ?geom3d.

23.4.4 2D-Geometrie

a) Finde eine Matrix $M(t)$ für die Drehung des $\mathbb{R}^2$ um den Nullpunkt um den Winkel t gegen den Uhrzeigersinn.
b) Finde eine Matrix S_x für die Spiegelung der Ebene an der x-Achse.
c) Was wird durch die Matrix $M = M(\pi/4)S_xM(-\pi/4)$ bewirkt?
d) Berechne MS_x und S_xM.

24 Übungszettel zu Matrixabbildungen

Es sei A eine 3×3-Matrix und $\varphi : \vec{x} \mapsto A\vec{x}$ die dadurch gegebene Matrixabbildung des $\mathbb{R}^3$ in sich.

24.1 Parallelität

Es seien g und h parallele Geraden. Bilden die Mengen $\varphi(g) := \{\varphi(\vec{x}) \,|\, \vec{x} \in g\}$ und $\varphi(h)$ wieder parallele Geraden? Beweise deine Behauptung sauber.

[20] Natürlich müssen die Matrizen die passenden Größen haben, damit das funktioniert.

24.2 Konstruktion einer Matrixabbildung mit einigen vorgegebenen Bildpunkten

Du weißt, dass du drei beliebige Vektoren $\vec{b}_1, \vec{b}_2, \vec{b}_3$ als Bilder der Einheitsvektoren festlegen kannst und problemlos eine Matrix A hinschreiben kannst, deren Matrixabbildung die Einheitsvektoren genau auf diese Vektoren bringt.
a) Führe dies für die folgenden drei Vektoren durch.

$$\vec{b}_1 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}, \quad \vec{b}_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

b) Wir wählen die Vektoren $\vec{a}_1 = \vec{e}_1$, $\vec{a}_2 = \vec{e}_2 + \vec{e}_3$ und $\vec{a}_3 = \vec{e}_3$. Suche eine Matrixabbildung, die $\vec{a}_1$ und $\vec{a}_2$ fest lässt und $\vec{a}_3$ auf $\vec{a}_1$ abbildet.

24.3 Kern und Bild

Die Mengen $Kern(\varphi)$ und $Bild(\varphi)$ einer Matrixabbildung φ sind folgendermaßen definiert:

$$Kern(\varphi) := \{\vec{x} \,|\, \varphi(\vec{x}) = \vec{0}\} \tag{50}$$

$$Bild(\varphi) := \{\varphi(\vec{x}) \,|\, \vec{x} \in \mathbb{R}^3\} \tag{51}$$

(Natürlich muss man die $\vec{x}$ in der Definition von $Bild(\varphi)$ aus dem $\mathbb{R}^n$ nehmen, auf dem φ definiert ist; es ist nicht immer der $\mathbb{R}^3$)
Aufgabe: Berechne Kern und Bild der Abbildungen der vorigen Aufgabe.
Aufgabe: Definitionen fallen nicht vom Himmel, sondern werden so geprägt, dass sie nützlich sind. Warum interessiert man sich für Kern und Bild einer Matrixabbildung?

24.4 Spiegelung des Raums an einer Ebene

Es sei E eine Ebene, nimm meinetwegen erst einmal die von den Vektoren $\vec{a}_1$ und $\vec{a}_2$ der vorigen Aufgabe erzeugte Ebene. Wie findet man zu jedem Vektor den Bildvektor?

25 Kern und Bild einer Matrixabbildung

25.1 Definitionen und Eigenschaften von Kern und Bild

Was hier geschieht, geht allgemein für beliebige Matrizen, aber wir sehen uns die Begriffe für 3×3-Matrizen an, damit dir das Umfeld vertrauter ist. Zunächst gibt es kurz und trocken die Definitionen.

4 Definition *Es sei $\varphi : \vec{x} \mapsto A\vec{x}$ eine Matrixabbildung. Dann definieren wir den Kern und das Bild von φ wie folgt.*

$$Kern(\varphi) := \{\vec{x} \,|\, \varphi(\vec{x}) = \vec{0}\} \quad \textit{und} \quad Bild(\varphi) := \{\varphi(\vec{x}) \,|\, \vec{x} \in \mathbb{R}^3\}$$

Natürlich muss man in der Definition von $Bild(\varphi)$ statt $\mathbb{R}^3$ den $\mathbb{R}^n$ passender Größe nehmen, wenn die Matrix der Abbildung nicht eine 3×3-Matrix ist. - Zunächst einige Eigenschaften der Vektormengen $Kern(\varphi)$ und $Bild(\varphi)$.

5 Lemma *Die Vektormenge $Kern(\varphi)$ hat die folgenden Eigenschaften.*

- $\vec{0} \in Kern(\varphi)$
- *Wenn $\vec{x} \in Kern(\varphi)$ ist, dann ist auch $r\vec{x} \in Kern(\varphi)$ für jede Zahl $r \in \mathbb{R}$.*

- *Wenn $\vec{x}, \vec{y} \in Kern(\varphi)$ ist, dann ist auch $\vec{x} + \vec{y} \in Kern(\varphi)$.*
- *Es ist $\varphi(\vec{x} + \vec{y}) = \varphi(\vec{x})$ für jeden Vektor $\vec{y} \in Kern(\varphi)$.*
- *Es ist $\varphi(\vec{x}) = \varphi(\vec{y})$ genau dann, wenn $\vec{x} - \vec{y} \in Kern(\varphi)$ ist.*

Ich beweise die dritte Aussage: Es seien $\vec{x}, \vec{y} \in Kern(\varphi)$. Dann ist $\varphi(\vec{x}) = \vec{0}$ und $\varphi(\vec{y}) = \vec{0}$. Damit ist dann auch $\varphi(\vec{x} + \vec{y}) = \varphi(\vec{x}) + \varphi(\vec{y}) = \vec{0} + \vec{0} = \vec{0}$, also liegt auch $\vec{x} + \vec{y}$ in $Kern(\varphi)$.

Ähnliche Aussagen lassen sich auch über $Bild(\varphi)$ machen.

6 Lemma *Für jede Matrixabbildung φ gelten die folgenden Aussagen.*

- $\vec{0} \in Bild(\varphi)$
- *$r\vec{x} \in Bild(\varphi)$ für alle $\vec{x} \in Bild(\varphi)$ und alle $r \in \mathbb{R}$.*
- *$\vec{x} + \vec{y} \in Bild(\varphi)$ für alle $\vec{x}, \vec{y} \in Bild(\varphi)$.*

Was soll das nun alles? Zunächst stellen die im zweiten Lemma aufgeführten Eigenschaften sicher, dass $Bild(\varphi)$ und $Kern(\varphi)$ nicht irgendwelche wilden Vektormengen sein können, sondern nur $\{\vec{0}\}$, eine Ursprungsgerade, eine Ebene durch den Nullpunkt oder der ganze Raum $\mathbb{R}^3$. Wir nehmen für den Augenblick einmal an, dass wir es mit einer Abbildung φ zu tun haben, deren Kern eine Ursprungsgerade ist, es sei also $Kern(\varphi) = \langle\vec{b}\rangle$ für einen Vektor $\vec{b} \neq \vec{0}$. Alle Vektoren im Kern gehen unter φ auf den Nullvektor. Nehmen wir ein $\vec{a}$, das nicht im Kern liegt. Dann ist $\varphi(\vec{a}) \neq \vec{0}$, und für alle $\vec{x} \in Kern(\varphi)$ ist $\varphi(\vec{a} + \vec{x}) = \varphi(\vec{a})$. Das heißt: alle Punkte der zu $Kern(\varphi)$ parallelen Geraden mit Stützvektor $\vec{a}$ gehen auf den selben Bildpunkt $\varphi(\vec{a})$. Gibt es noch einen Vektor $\vec{y}$ mit $\varphi(\vec{y}) = \varphi(\vec{a})$, dann muss $\vec{a} - \vec{y} \in Kern(\varphi)$ liegen, und somit gehört dann $\vec{y}$ zu einem Punkt der Geraden mit dem allgemeinen Vektor $\vec{a} + r\vec{b}$ - und das ist wieder die Gerade von oben. Wenn du dir den ganzen Raum zerlegt denkst in zu $Kern(\varphi)$ parallele Geraden, so wie ein Baumstamm aus Fasern besteht, dann haben alle Punkte einer Faser (und nur die!) den gleichen Bildpunkt. Die Menge aller Bildpunkte, das sei hier schon verraten, bildet dann eine Ebene im Raum.

25.2 Folgerungen für Lineare Gleichungssysteme

Nun können wir über das LGS $A\vec{x} = \vec{c}$ folgende Aussage machen. Für $\vec{c} = \vec{0}$ ist die Lösungsmenge gerade der Kern der Matrixabbildung $\varphi : \vec{x} \mapsto A\vec{x}$, also eine der Vektormengen, die oben aufgezählt wurden, sagen wir mal eine Ursprungsgerade. Falls die rechte Seite $\vec{c}$ des LGS zu $Bild(\varphi)$ gehört, ist die Lösungsmenge eine zu $Kern(\varphi)$ parallele Gerade, und sonst ist sie leer. Das ist eine sehr wichtige Einsicht in die Struktur der Lösungsmengen linearer Gleichungssysteme. Und wie du dich erinnerst, hatten wir Matrixabbildungen eingeführt, um Erkenntnisse über LGS zu gewinnen. Genau dies ist nun gelungen.

25.3 Berechnung des Bildraums einer Matrixabbildung

Nehmen wir der Einfachheit halber wieder eine Matrixabbildung $\varphi : \vec{x} \mapsto A\vec{x}$ für eine 3×3-Matrix A. Die folgenden Überlegungen gelten für beliebige Matrixabbildungen ganz entsprechend.

Wenn man den Kern $Kern(\varphi)$ berechnen will, muss man nur das LGS $A\vec{x} = \vec{0}$ lösen.[21] Das ist weiter kein Problem.

Aber wie bekommt man den Bildraum $Bild(\varphi)$? Man müsste ja $A\vec{x}$ bilden und $\vec{x}$ den ganzen $\mathbb{R}^3$ durchlaufen lassen?! Nun wissen wir ja, dass

$$A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + x_3\vec{a}_3$$

ist, wenn die x_i die Komponenenten des Vektors $\vec{x}$ und die $\vec{a}_i$ die Spaltenvektoren der Matrix A sind. Es gilt also folgendes Lemma.

[21] Ein LGS, dessen rechte Seite der Nullvektor ist, heißt **homogenes** LGS.

7 Lemma *Es sei $A = (\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n)$ eine Matrix und $\varphi : \vec{x} \mapsto A\vec{x}$ die zugehörige Matrixabbildung. Dann ist $Bild(\varphi)$ das Erzeugnis der Spaltenvektoren von A:*

$$Bild(\varphi) = \langle \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n \rangle$$

Ein durchgerechnetes Beispiel für die Bestimmung des Bildraums einer Matrixabbildung findest du im Buch auf Seite 168 (Beispiel 6.3). Freilich fällt die Einsicht, dass $A\vec{e}_1 = A\vec{e}_2 + A\vec{e}_3$ ist, vom Himmel. Und was ist, wenn das Warten auf einen Einfall keinen Erfolg hat? Nun, zunächst einmal liegt ja $\vec{a}_1 = \varphi(\vec{e}_1) = A\vec{e}_1$ in $Bild(\varphi)$, also auch $\langle \vec{a}_1 \rangle$, und das ist eine Gerade durch den Nullpunkt. Der Vektor $\vec{a}_2 = \varphi(\vec{e}_2)$ liegt offensichtlich nicht auf dieser Geraden, also erzeugen $\vec{a}_1$ und $\vec{a}_2$ eine Ebene durch den Nullpunkt: $Bild(\varphi)$ enthält jedenfalls eine Ebene. Wenn $\vec{a}_3$ aus dieser Ebene "herausragt", wenn er also nicht in dieser Ebene liegt, dann ist $Bild(\varphi)$ der ganze Raum $\mathbb{R}^3$. Rechne nach, dass $\vec{a}_3 \in \langle \vec{a}_1, \vec{a}_2 \rangle$ ist - dazu musst du ein LGS lösen. Dann weißt du, dass der Bildraum $Bild(\varphi)$ die von $\vec{a}_1$ und $\vec{a}_2$ erzeugte Ebene ist.

Wir vertiefen diese Überlegungen im nächsten Kapitel, sie führen auf die Begriffe Basis und Dimension.

26 Teilräume des $\mathbb{R}^n$, Basis und Dimension

Die Vektormengen $Kern(\varphi)$ und $Bild(\varphi)$ waren Teilmengen des $\mathbb{R}^n$ mit einigen bemerkenswerten Eigenschaften. Geometrisch gesprochen handelte sich um Ursprungsgeraden, Ebenen durch den Nullpunkt, den ganzen Raum oder die Menge, die nur den Nullpunkt enthält. Dies folgt aus der algebraischen Eigenschaft, dass die Mengen abgeschlossen sind gegenüber Addition und Multiplikation mit Zahlen: diese Operationen führen nicht aus der Menge heraus. All dies ist wichtig genug, dass man einen neuen Begriff prägt.[22]

8 Definition *Eine Teilmenge M des $\mathbb{R}^n$ heißt Teilraum des $\mathbb{R}^n$, in Zeichen: $M \leq \mathbb{R}^n$, wenn sie die folgenden drei Eigenschaften hat.*

- $\vec{0} \in M$
- $r\vec{x} \in M \ \forall \vec{x} \in M \ \forall r \in \mathbb{R}$
- $\vec{x} + \vec{y} \in M \ \forall \vec{x}, \vec{y} \in M$

Für eine Matrixabbildung φ sind also $Kern(\varphi)$ und $Bild(\varphi)$ stets Teilräume des $\mathbb{R}^n$.

Wie kommt es nun, dass diese Teilräume nur die oben aufgezählten besonderen Gebilde sein können? Nun, es sei M ein solcher Teilraum des $\mathbb{R}^n$, es sei also $M \leq \mathbb{R}^n$. Wenn M dann überhaupt einen Vektor $\vec{x}_1 \neq \vec{0}$ enthält, enthält es gleich die ganze von $\vec{x}_1$ erzeugte Ursprungsgerade. Wenn diese Gerade noch nicht ganz M ausmacht, gibt es einen Vektor $\vec{x}_2 \in M$, der nicht zu $\langle \vec{x}_1 \rangle$ gehört. Dann aber muss M gleich die ganze von $\vec{x}_1$ und $\vec{x}_2$ erzeugte Ebene enthalten. Wenn diese Ebene $\langle \vec{x}_1, \vec{x}_2 \rangle$ noch nicht ganz M ausmacht, gibt es einen weiteren Vektor $\vec{x}_3 \in M$, der aus der Ebene herausragt. Dann enthält M aber das ganze Erzeugnis der drei Vektoren, und dies ist bereits der ganze Raum $\mathbb{R}^3$. So ungefähr läuft die Argumentation, sie lässt sich im Prinzip auch für den allgemeinen Fall des $\mathbb{R}^n$ führen.[23]

Nebenbei macht die Argumentation plausibel, dass folgendes Lemma gilt:

9 Lemma *Es sei $M \leq \mathbb{R}^n$, und M enthalte einen Vektor $\neq \vec{0}$. Dann kann man M als Erzeugnis geeigneter Vektoren $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m$ aus M schreiben:*

$$M = \langle \vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m \rangle$$

für geeignete Vektoren $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m$ aus M.

[22] Das Kürzel $\forall$ in der folgenden Definition bedeutet schlicht "für alle".

[23] Eine Schwierigkeit dabei ist, dass man sicherstellen muss, dass dieser Prozess zum Ende kommt und dass man nicht etwa immer neue Vektoren hinzunehmen kann, ohne jemals ganz M auszuschöpfen.

Einige Begriffe: Die Vektoren aus dem Lemma bilden ein **Erzeugendensystem** für M. Man wird natürlich alle überflüssigen Vektoren weglassen, dann kommt man zu einem **minimalen Erzeugendensystem** oder **Basis** von M. Für solche Basen gilt der folgende Satz, den Beweis gebe ich auf Wunsch.

2 Satz *Es sei $M \leq \mathbb{R}^n$, und M enthalte einen Vektor $\neq \vec{0}$. Dann haben alle minimalen Erzeugendensysteme von M die gleiche Anzahl von Elementen.*

Diese Anzahl von Elementen einer (und damit jeder) Basis von M heißt die **Dimension** von M. Sie wird bezeichnet mit dem Symbol $dim(M)$.

Übungen

1) Es seien $\vec{x}_1, \ldots, \vec{x}_m \in \mathbb{R}^n$. Zeige, dass dann $\langle \vec{x}_1, \ldots, \vec{x}_m \rangle \leq \mathbb{R}^n$ gilt.
2) Es sei $M = \{\vec{x} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$. Zeige, dass $M \leq \mathbb{R}^3$ ist.
3) Es sei $M = \vec{e}_3 + \langle \vec{e}_1, \vec{e}_2 \rangle$. Zeige, dass M kein Teilraum des $\mathbb{R}^3$ ist.
4) Es sei $\vec{a} \in \mathbb{R}^3$ und $M = \{\vec{x} \in \mathbb{R}^3 \mid \vec{x} * \vec{a} = 0\}$. Zeige, dass $M \leq \mathbb{R}^3$ gilt.

27 Basen von Kern und Bild

Ich werde nun an einem Beispiel vorführen, wie man Basen von Kern und Bild einer gegebenen Matrixabbildung gewinnt.

Es sei A meinetwegen eine 6×8 -Matrix. Sie gehört dann zu einer Matrixabbildung $\varphi : \mathbb{R}^8 \to \mathbb{R}^6$, die Spalten der Länge 8 frisst und Spalten der Länge 6 ausgibt. Wir versuchen nun, den Kern der Abbildung zu berechnen, daraus ergibt sich dann alles Weitere.

27.1 Berechnung des Kerns

Ein Vektor $\vec{x} \in Kern(\varphi)$ erfüllt die Bedingung $A\vec{x} = \vec{0}$. Der Kern ist also die Lösungsmenge des homogenen LGS

$$A\vec{x} = \vec{0} \tag{52}$$

Dieses System bearbeiten wir - wie üblich - mit dem Gaußschen Algorithmus. Als Ergebnis erhalten wir eine Koeffizientenmatrix, die so aussieht wie die folgende Matrix B.[24]

$$B = \begin{pmatrix} 1 & * & * & . & . & . & . & . \\ 0 & 0 & 0 & 1 & . & . & . & . \\ 0 & 0 & 0 & 0 & 1 & . & . & . \\ 0 & 0 & 0 & 0 & 0 & 1 & * & . \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{53}$$

Sowohl die Punkte als auch die Sternchen stehen für irgendwelche Zahlen. Wie sie genau aussehen, spielt keinerlei Rolle.

Schreiben wir nun den allgemeinen Vektor $\vec{x}$ von $Kern(\varphi)$ hin. Die sechste Gleichung stellt keinerlei Forderungen. Die fünfte verlangt aber, dass $x_8 = 0$ ist, dann ist sie zufrieden. Den Eintrag x_7 von $\vec{x}$ können wir frei wählen und x_6 dann so berechnen, dass auch die vierte Gleichung erfüllt ist. Die dritte und die zweite Gleichung legen dann x_5 und x_4 fest. Schließlich bleiben x_3 und x_2 frei wählbar. Der Eintrag x_1 wird dann so berechnet, dass auch die erste Gleichung erfüllt ist.

[24] Die "Treppenform" erhält man, indem man nötigenfalls Gleichungen - also Zeilen des Systems! - vertauscht. Dass die erste Zahl $\neq 0$ in jeder Zeile $= 1$ ist, kriegt man durch Multiplikation der Gleichung mit einer Zahl $\neq 0$ hin.

Insgesamt erhalten wir für den allgemeinen Vektor $\vec{x} \in Kern(\varphi)$ die folgende Darstellung:

$$\vec{x} = \begin{pmatrix} \diamond \\ x_2 \\ x_3 \\ \diamond \\ \diamond \\ \diamond \\ x_7 \\ 0 \end{pmatrix} = x_2 \begin{pmatrix} \diamond \\ 1 \\ 0 \\ \diamond \\ \diamond \\ \diamond \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} \diamond \\ 0 \\ 1 \\ \diamond \\ \diamond \\ \diamond \\ 0 \\ 0 \end{pmatrix} + x_7 \begin{pmatrix} \diamond \\ 0 \\ 0 \\ \diamond \\ \diamond \\ \diamond \\ 1 \\ 0 \end{pmatrix} =: x_2\vec{y}_1 + x_3\vec{y}_2 + x_7\vec{y}_3 \tag{54}$$

An den mit $\diamond$ besetzten Stellen stehen dabei irgendwelche Linearkombinationen von x_2, x_3 und x_7.

Aus $\vec{x} = \vec{0}$ folgt sofort $x_2 = x_3 = x_7 = 0$, also sind die drei Vektoren $\vec{y}_1, \vec{y}_2, \vec{y}_3$, als deren Linearkombination $\vec{x}$ in (54) dargestellt ist, linear unabhängig. Sie bilden eine Basis von $Kern(\varphi)$, die Dimension des Kerns ist 3. Das ist natürlich die Anzahl der Sternchen in der Matrix B in (53). Unser Ergebnis:

$$Kern(\varphi) = \langle \vec{y}_1, \vec{y}_2, \vec{y}_3 \rangle \quad \text{und} \quad \dim(Kern(\varphi)) = 3 \tag{55}$$

27.2 Berechnung des Bilds

Wir wissen von vornherein, dass $Bild(\varphi)$ von den Spaltenvektoren $\vec{a}_1, \ldots, \vec{a}_8$ der ursprünglichen Matrix A erzeugt wird. An $Kern(\varphi)$ erkennen wir nun, welche der $\vec{a}_i$ wir weglassen können, so dass eine Basis von $Bild(\varphi)$ übrig bleibt.

Jeder der Vektoren $\vec{y}_1, \vec{y}_2, \vec{y}_3$ aus (54) liefert uns eine Darstellung des Nullvektors als Linearkombination der $\vec{a}_i$. Die Darstellung, die $\vec{y}_1$ liefert, gestattet, $\vec{a}_2$ als Linearkombination der übrigen zu schreiben. Wegen der Nullen in $\vec{y}_1$ werden dabei $\vec{a}_3$ und $\vec{a}_7$ nicht benötigt[25]. Nutzen wir nun noch $\vec{y}_2$ und $\vec{y}_3$ aus, sehen wir, dass wir auch $\vec{a}_3$ und $\vec{a}_7$ als Linearkombination von $\vec{a}_1, \vec{a}_4, \vec{a}_5, \vec{a}_6$ und $\vec{a}_8$ schreiben können. Somit gilt schon einmal

$$Bild(\varphi) = \langle \vec{a}_1, \vec{a}_4, \vec{a}_5, \vec{a}_6, \vec{a}_8 \rangle \quad , \tag{56}$$

die Dimension von $Bild(\varphi)$ ist höchstens 5. Schauen wir genauer hin: wenn wir in der Darstellung von $\vec{x} \in Kern(\varphi)$ in (54) die frei wählbaren Parameter x_2, x_3, x_7 alle auf Null setzen, werden auch alle $\diamond = 0$. Es gibt also nur die Darstellung des Nullvektors als Linearkombination von $\vec{a}_1, \vec{a}_4, \vec{a}_5, \vec{a}_6, \vec{a}_8$, in der alle Koeffizienten $= 0$ sind. Somit bilden diese Vektoren eine Basis von $Bild(\varphi)$, es ist

$$\dim(Bild(\varphi)) = 5 \tag{57}$$

27.3 Die Dimensionsformel

Was ich hier an einer "konkreten" Matrix A vorgeführt habe, läuft im allgemeinen Fall genau so ab. Es gilt stets für die durch $\varphi : \vec{x} \mapsto A\vec{x}$ definierte Matrixabbildung

$$\dim(Kern(\varphi)) + \dim(Bild(\varphi)) = \text{ Anzahl der Spalten von } A \tag{58}$$

28 Klausur Nr. 2

1) Es sei

$$A = \begin{pmatrix} 2 & -3 \\ -4 & 6 \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad .$$

a) Berechne Kern und Bild der Matrixabbildung $\varphi : \vec{x} \mapsto A\vec{x}$ und zeichne die zugehörigen Punktmengen der Ebene.

[25] $\vec{a}_8$ auch nicht, aber das ist egal.

b) Bestimme die Menge $M = \{\vec{x} \mid \varphi(\vec{x}) = \varphi(\vec{b})\}$ und zeichne die zugehörige Punktmenge der Ebene noch in deine Zeichnung zu a) ein.
c) Erkläre anhand deiner Zeichnung zu b), welche allgemeine Einsicht in die Struktur der Lösungsmengen von LGS wir mit Hilfe der Matrixabbildungen gewonnen haben.
d) Berechne A^2.

2) a) Welche geometrische Bedeutung haben die Abbildungen zu den folgenden Matrizen?

$$A = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

b) Jede der drei Matrizen hat eine inverse Matrix. Schreibe die inversen Matrizen hin.

3) Zu der Matrix

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

gibt es Vektoren $\vec{a}, \vec{b} \neq \vec{0}$ mit folgenden Eigenschaften: Es ist $A\vec{a} = \vec{a}$ und $A\vec{b} = -\vec{b}$ und $\vec{a} * \vec{b} = 0$.
a) Berechne solche Vektoren $\vec{a}$ und $\vec{b}$.
b) Was bewirkt die zu A gehörige Matrixabbildung geometrisch? Anmerkung: das kannst du mit Hilfe der angegebenen Eigenschaften auch beantworten, wenn du keine konkreten Vektoren berechnet hast.

4) Es seien $\vec{a}$ und $\vec{b}$ zwei feste Vektoren aus dem gleichen $\mathbb{R}^n$. Wir setzen

$$M := \{\vec{x} \in \mathbb{R}^n \mid \vec{x} * \vec{a} = 0 \ \text{ und } \ \vec{x} * \vec{b} = 0\} \quad .$$

Zeige, dass M ein Teilraum des $\mathbb{R}^n$ ist, und zwar
a) direkt über die Definition
b) indem du eine Matrixabbildung angibst, deren Kern gerade M ist.
c) Berechne M für

$$\vec{a} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad , \quad \vec{b} = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \quad .$$

Welche Dimension hat dieses M? Interpretiere die Angelegenheit geometrisch.

5) a) Kunz meint: "Wenn ich drei Vektoren des $\mathbb{R}^3$ habe, von denen keine zwei in dieselbe Richtung zeigen, dann bilden die drei Vektoren eine Basis des $\mathbb{R}^3$." Stimmt das?
b) Was heißt denn genau, dass $\vec{a}$ und $\vec{b}$ in dieselbe Richtung zeigen? Gib eine präzise Formulierung.

6) Hinz will Kern und Bild einer Matrixabbildung $\varphi : \vec{x} \mapsto A\vec{x}$ berechnen. Er bearbeitet die Matrix sachgerecht und erhält als Ergebnis die folgende Matrix $B =$

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

a) Gib Basen des Kerns und des Bildes der Abbildung an. Beschreibe dabei, wie du vorgehst, und schreibe nicht nur ein Ergebnis hin.
b) Was heißt überhaupt, dass Hinz die Matrix "sachgerecht bearbeitet"? - Ein halber Satz reicht hier als Antwort.

7) Der Vektor $\vec{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ soll durch Drehungen auf den Teil der z-Achse mit $z \geq 0$ gebracht werden, und zwar soll er zuerst um die z-Achse in die yz-Ebene und dann um die x-Achse auf die

z-Achse gedreht werden.
a) Berechne die beiden Drehwinkel t_z und t_x.
b) Gesucht ist die Matrix D einer Drehung um die von $\vec{a}$ erzeugte Achse gegen den Uhrzeigersinn um den Winkel u. Schreibe D als ein Produkt aus unseren Drehmatrizen $M_x(t)$, $M_y(t)$ und $M_z(t)$. Vorsicht: du sollst nur das Produkt solcher Symbole hinschreiben, aber nicht die konkreten 3×3-Matrizen ausschreiben und schon gar nicht die konkrete Produktmatrix ausrechnen!
c) Ohne die Drehmatrix zu kennen, kann man schon den Bildvektor $D\vec{a}$ angeben. Welcher ist es?

29 Klausur Nr. 2: Sonderausgabe für Viktor

1) Es sei

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \quad \text{und} \quad \vec{v} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad .$$

Wir betrachten die drei Abbildungen $\varphi : \vec{x} \mapsto A\vec{x}$, $\psi : \vec{x} \mapsto B\vec{x}$ und $\tau : \vec{x} \mapsto \vec{x} + \vec{v}$ des $\mathbb{R}^2$ in sich.
a) Wende die Abbildungen auf das Einheitsquadrat an und deute die Abbildungen geometrisch.
b) Gib zu jeder der drei Abbildungen die inverse Abbildung an, die sie wieder rückgängig macht.
c) Berechne die Matrix BAB^{-1} und wende die zugehörige Matrixabbildung auf das Einheitsquadrat an. Um was handelt es sich geometrisch?
d) Ist τ eine lineare Abbildung?

2) Es sei

$$\vec{a} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{c} = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} \quad .$$

a) Weise nach, dass $\vec{a}, \vec{b}, \vec{c}$ eine Basis des $\mathbb{R}^3$ ist.
b) Mache aus dieser Basis eine Basis aus paarweise orthogonalen Vektoren.
c) Der Vektor $\vec{c}$ soll in die xy-Ebene gedreht werden. Gib eine Drehachse und einen Drehwinkel an.
d) Bestimme eine Matrixabbildung, deren Kern die xy-Ebene und deren Bild die von $\vec{c}$ erzeugte Gerade ist. Gibt es mehrere Möglichkeiten?
e) Gibt es auch eine Matrixabbildung, deren Kern die xy-Ebene und deren Bild das Erzeugnis von $\vec{a}$ und $\vec{c}$ ist?

3) Es sei

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

a) Bestimme die Menge aller Vektoren $\vec{x}$ mit $A\vec{x} = \vec{x}$.
b) Bestimme die Menge aller Vektoren $\vec{y}$ mit $A\vec{y} = -\vec{y}$.
c) Was bewirkt die zu A gehörige Matrixabbildung geometrisch?
d) Findet man zu jeder Matrixabbildung Vektoren $\neq \vec{0}$, die die Bedingungen in a) bzw. in b) erfüllen?

4) Es sei

$$\vec{a} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \quad \text{und} \quad \vec{c} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

Bestimme die Menge aller Vektoren $\vec{x} \in \mathbb{R}^4$, die auf allen drei Vektoren (gleichzeitig) senkrecht stehen. Deute die Aufgabenstellung und dein Ergebnis geometrisch.

30 Orthonormalbasen

Ein linear unabhängiges System $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ von Vektoren des $\mathbb{R}^n$ erzeugt einen m-dimensionalen Teilraum $V := \langle \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m \rangle$ des $\mathbb{R}^n$, das heißt es ist eine Basis von V.

Wie wir wissen, lässt sich jeder Vektor $\vec{x} \in V$ eindeutig als Linearkombination der Basisvektoren schreiben:

$$\vec{x} \in V \Longrightarrow \vec{x} = r_1\vec{a}_1 + r_2\vec{a}_2 + \cdots + r_m\vec{a}_m = \sum_{i=1}^{m} r_i\vec{a}_i \tag{59}$$

Will man die Koeffizienten r_i für ein konkretes $\vec{x} \in V$ berechnen, muss man ein $n \times m$-LGS lösen. Das ist im Prinzip kein Problem, aber doch mit der Hand nur für ganz kleine m zu machen. Hier gibt es einen interessanten Ansatz.

10 Lemma *Die Vektoren $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ des $\mathbb{R}^n$ seien alle $\neq \vec{0}$ und paarweise orthogonal. Dann sind sie linear unabhängig.*

Der Beweis ist ganz leicht. Wenn man eine Linearkombination

$$\vec{x} = r_1\vec{a}_1 + r_2\vec{a}_2 + \cdots + r_m\vec{a}_m = \sum_{i=1}^{m} r_i\vec{a}_i \tag{60}$$

skalar mit $\vec{a}_i$ multipliziert, fallen wegen der Orthogonalität alle Summanden ausser $r_i\vec{a}_i * \vec{a}_i$ weg. Aus $\vec{x} = \vec{0}$ folgt dann sofort $r_i\vec{a}_i * \vec{a}_i = \vec{0} * \vec{a}_i = 0$, also $r_i = 0$. Damit ist die lineare Unabhängigkeit schon gezeigt. Und wenn man genau hinschaut, sieht man, dass sogar das folgende Lemma mit bewiesen ist.

11 Lemma *Es sei $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ ein System paarweiser orthogonaler Vektoren $\neq \vec{0}$ des $\mathbb{R}^n$. Dann gilt für jedes $\vec{x} = \sum_{i=1}^{m} r_i\vec{a}_i$ des von den $\vec{a}_i$ erzeugten Teilraums des $\mathbb{R}^n$*

$$r_i = \frac{\vec{x} * \vec{a}_i}{\vec{a}_i * \vec{a}_i} \quad \text{für } i = 1, \ldots, m$$

Haben alle $\vec{a}_i$ die Länge 1, wird das Ergebnis noch einfacher, dann ist nämlich schlicht $r_i = \vec{x}_i * \vec{a}_i$ für $i = 1, \ldots, m$. Eine solche Basis heißt Orthonormalbasis.

12 Beispiel *Die Standardbasis $\vec{e}_1, \ldots, \vec{e}_n$ ist eine Orthonormalbasis des $\mathbb{R}^n$.*

Wie aber bekommt man Orthonormalbasen? Das ist einfacher, als du denkst. Gehen wir also von einer Basis $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ mit Erzeugnis V aus. Wir beschaffen uns nun ein System paarweise orthogonaler Vektoren $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m$, die den gleichen Teilraum V erzeugen. Wenn du willst, kannst du anschließend die Längen der Vektoren auf 1 bringen.

Erster Schritt:

$$\vec{b}_1 := \vec{a}_1 \tag{61}$$

Warum auch nicht, ein Vektor ist wie der andere.

Zweiter Schritt:

$$\vec{b}_2 := \vec{a}_2 - \frac{\vec{a}_2 * \vec{b}_1}{\vec{b}_1 * \vec{b}_1}\vec{b}_1 \tag{62}$$

Die neuen Vektoren $\vec{b}_1, \vec{b}_2$ haben das gleiche Erzeugnis wie $\vec{a}_1, \vec{a}_2$, und sie sind orthogonal zu einander. Du kannst das nachrechnen, oder du kannst dir überlegen, was das geometrisch bedeutet, was da abläuft: Man subtrahiert von $\vec{a}_2$ den Anteil parallel zu $\vec{b}_1$, so dass das, was übrigbleibt, orthogonal zu $\vec{b}_1$ ist. Die dadurch gegebene Abbildung $\vec{a}_2 \mapsto \vec{b}_2$ ist eine Scherung mit Scherrichtung $\vec{b}_1$.

Dritter Schritt:

$$\vec{b}_3 := \vec{a}_3 - \frac{\vec{a}_3 * \vec{b}_1}{\vec{b}_1 * \vec{b}_1}\vec{b}_1 - \frac{\vec{a}_3 * \vec{b}_2}{\vec{b}_2 * \vec{b}_2}\vec{b}_2 = \vec{a}_3 - \sum_{i=1}^{2} \frac{\vec{a}_3 * \vec{b}_i}{\vec{b}_i * \vec{b}_i}\vec{b}_i \tag{63}$$

Und so weiter. Hat man schon $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_k$ auf diese Weise konstruiert und ist man noch nicht fertig, gewinnt man den nächsten Vektor $\vec{b}_{k+1}$ durch

$$\vec{b}_{k+1} := \vec{a}_{k+1} - \sum_{i=1}^{k} \frac{\vec{a}_{k+1} * \vec{b}_i}{\vec{b}_i * \vec{b}_i} \vec{b}_i \tag{64}$$

Dieses Verfahren ist als Gram-Schmidtsches Orthogonalisierungsverfahren bekannt. Wir werden davon noch Gebrauch machen. Mache dich damit gründlich vertraut.

31 Beweisübungen

31.1 Lineare Abbildungen

Es sei $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ eine Abbildung mit folgenden Eigenschaften:

$$\varphi(\vec{x} + \vec{y}) = \varphi(\vec{x}) + \varphi(\vec{y}) \qquad \forall\, \vec{x}, \vec{y} \in \mathbb{R}^n \tag{65}$$

$$\varphi(r\vec{x}) = r\varphi(\vec{x}) \qquad \forall\, \vec{x} \in \mathbb{R}^n, r \in \mathbb{R} \tag{66}$$

Eine solche Abbildung heißt lineare Abbildung.

Beweise, dass jede lineare Abbildung eine Matrixabbildung ist.

31.2 Summe und Schnitt von Teilräumen

Es seien V und W Teilräume des $\mathbb{R}^n$. Beweise, dass dann auch

$$V \cap W := \{\vec{x} \,|\, \vec{x} \in V \text{ und } \vec{x} \in W\}$$

und

$$V + W := \{\vec{v} + \vec{w} \,|\, \vec{v} \in V, \vec{w} \in W\}$$

Teilräume des $\mathbb{R}^n$ sind.

32 Fourierreihen

In diesem Kapitel sollst du ein Beispiel für die Anwendungsmöglichkeiten unserer Theorie sehen. Um die Sachen einfach zu halten, betrachten wir aber nur einen Spezialfall. Schaffen wir zunächst die technischen Voraussetzungen. Du bist ja kein Anfänger mehr, da können wir gleich zur Sache kommen.

32.1 Theorie

13 Lemma *Für $n, m \in \mathbb{N}$ ist*

$$\int_{-\pi}^{\pi} \sin(nx)\sin(mx)\,dx = \begin{cases} \pi & \text{für } m = n \\ 0 & \text{für } m \neq n \end{cases} \tag{67}$$

Dass das Lemma gilt, rechnest du einfach nach. Für $m \neq n$ solltest du zweimal partiell integrieren, dann hast du das Ausgangsintegral mit einem Vorfaktor wieder auf der rechten Seite. Du kannst die Gleichung dann nach dem Integral auflösen. Für $m = n$ integrierst du einmal partiell und ersetzt $\cos^2(nx)$ durch $1 - \sin^2(nx)$. Dann kannst du die Gleichung wieder nach dem gesuchten Integral auflösen.

Aus dem Lemma ergibt sich, dass die Vektoren $\vec{v}_n := (x \mapsto \sin(nx), x \in [-\pi;\pi])$ alle $\neq \vec{0}$ und paarweise orthogonal und folglich linear unabhängig sind. Sie erzeugen einen unendlich-dimensionalen Teilraum S des Vektorraums aller Funktionen auf $[-\pi;\pi]$. Dabei benutzen wir das übliche Skalarprodukt unseres Funktionenraums

$$(x \mapsto f(x)) * (x \mapsto g(x)) := \int_{-\pi}^{\pi} f(x)g(x)\,dx \tag{68}$$

Wozu ist das gut? Nehmen wir eine $\vec{h} := (x \mapsto h(x))$. Für jedes $n \in \mathbb{N}$ können wir $\vec{h}$ als Summe eines Vektors $\vec{u}_n \in \langle \vec{v}_1, \ldots, \vec{v}_n \rangle$ und eines Vektors $\vec{r}_n$ darstellen, der auf jedem der $\vec{v}_1, \ldots, \vec{v}_n$ senkrecht steht. Und $\vec{u}_n$ kann man sogar gleich hinschreiben:

$$\vec{u}_n = \sum_{i=1}^{n} a_i \vec{v}_i \quad \text{mit} \quad a_i = \frac{\vec{h} * \vec{v}_i}{\vec{v}_i * \vec{v}_i} = \frac{1}{\pi}\int_{-\pi}^{\pi} h(x)\sin(ix)\,dx \;\text{ für } i = 1, \ldots, n. \tag{69}$$

Du erkennst die Konstruktion: $\vec{u}_n$ ist die orthogonale Projektion von $\vec{h}$ in den von $\vec{v}_1, \ldots, \vec{v}_n$ erzeugten Teilraum des Funktionenraums. Die Koeffizienten sind genau die aus unserem Gram-Schmidt-Verfahren.

Die Frage ist nun, wie gut die Annäherung für $\vec{h}$ ist, die man auf diese Weise bekommt. Jedenfalls gilt Folgendes: Es ist

$$\vec{h} * \vec{h} = (\vec{u}_n + \vec{r}_n)(\vec{u}_n + \vec{r}_n) = \vec{u}_n * \vec{u}_n + \vec{r}_n * \vec{r}_n \quad , \tag{70}$$

und

$$\vec{u}_n * \vec{u}_n = \sum_{i=1}^{n} a_i^2 \vec{v}_i * \vec{v}_i = \sum_{i=1}^{n} a_i^2 \pi \tag{71}$$

Der Betrag der Näherung $\vec{u}_n$ wird mit wachsendem n also tendenziell größer und der Betrag des Rests $\vec{r}_n$ entsprechend tendenziell kleiner. Mit unseren $\sin(nx)$, auf die ich mich der Einfachheit halber beschränkt habe, kommt man nicht aus, aber wenn man noch die Funktionen $\vec{w}_n = (x \mapsto \cos(nx))$ für $n \in \mathbb{N}$ und $\vec{w}_0 = (x \mapsto 1)$ hinzunimmt, kann man auch ziemlich wilde Funktionen beliebig gut annähern[26].

32.2 Anwendungsbeispiel 1

Im Unterricht haben wir $\vec{h} = (x \mapsto x)$ behandelt. Die Koeffizienten sind

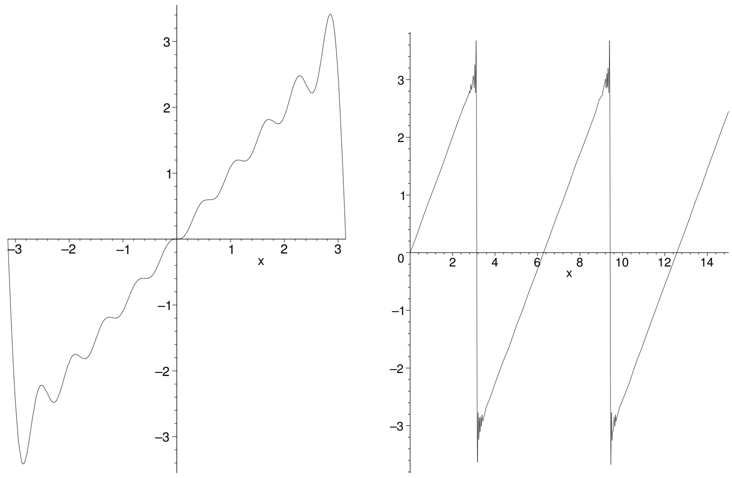
$$\begin{aligned}
a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} x\sin(nx)\,dx && (72)\\
&= \frac{1}{\pi}\left(\left(x(-\frac{1}{n}\cos(nx)\right)\Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} -\frac{1}{n}\cos(nx)\,dx\right) && (73)\\
&= \frac{1}{\pi}\left(\frac{2}{n}\pi - 0\right) && (74)\\
&= \frac{2}{n} \quad . && (75)
\end{aligned}$$

Folglich erhalten wir die Näherung

$$x \approx \sum_{i=1}^{n} \frac{2}{n}\sin(nx) \quad . \tag{76}$$

Wie das für $n = 10$ aussieht, zeigt Abbildung 1(a). Dass man an dieser komplizierten Darstellung

[26] Das Fachwort für annähern ist approximieren.



(a) $n = 10, \ -\pi \leq x \leq \pi$ (b) $n = 100, \ 0 \leq x \leq 15$

Abbildung 2: Endliche Fourierreihen für $h(x) = x$

der harmlosen Funktion $h(x) = x$ überhaupt Interesse hat, liegt daran, dass die trigonometrischen Funktionen periodisch sind. Abbildung 1(b) zeigt, wie der Graph der Fourierreihe für $n = 100$ im Bereich $0 \leq x \leq 15$ aussieht.

Wir haben also einen gut handhabbaren Rechenausdruck für eine Sägezahnkurve gefunden. Man kann die Fourierreihe ableiten und integrieren, sie ist nicht etwa abschnittweise gegeben. Joseph Fourier, der die Theorie entwickelt hat, trieb nicht l'art pour l'art, sondern suchte brauchbares Werkzeug für seine Physik. Er wollte Wärmeleitung untersuchen und rechnerisch in den Griff bekommen. Im letzten Abschnitt deute ich an, in welcher Weise Fourierreihen dabei hilfreich sind.

32.3 Anwendungsbeispiel 2

Wir nehmen uns folgende Funktion vor.

$$h(x) = \begin{cases} x + \pi & \text{für } -\pi \leq x - \frac{\pi}{2} \\ -x & \text{für } -\frac{\pi}{2} < x \leq \frac{\pi}{2} \\ -x + \pi & \text{für } \frac{\pi}{2} < x \leq \pi \end{cases} \tag{77}$$

Zeichne selbst den Graphen von h im Intervall $[-\pi; \pi]$ und berechne die Koeffizienten a_n. Dazu musst du den Integrationsbereich in drei Abschnitte unterteilen. Der Graph der Fourierreihe mit $n = 10$ für unser h ist in Abbildung 2 gezeichnet.

32.4 Ausblick für Unerschrockene: Wozu?

Man schreibt eine gegebene periodische Funktion als Fourierreihe, damit man sie physikalisch realisieren kann. Mit Hilfe von geeigneten Frequenzgeneratoren könnte man eine Rechteckspannung oder eine Sägezahnspannung erzeugen (ob man das tatsächlich tut, weiß ich aber nicht). Mathematisch benutzt man Fourierreihen in folgender Weise. Man muss eine Differentialgleichung lösen, man sucht also zum Beispiel eine Funktion y mit

$$y'' + 3y' = g(x) \quad .$$

Wenn man $g(x)$ als Fourierreihe schreiben kann, setzt man für y eine Fourierreihe mit unbestimmten Koeffizienten an, berechnet y' und y'', setzt alles in die Diffehrentialgleichung ein und kann

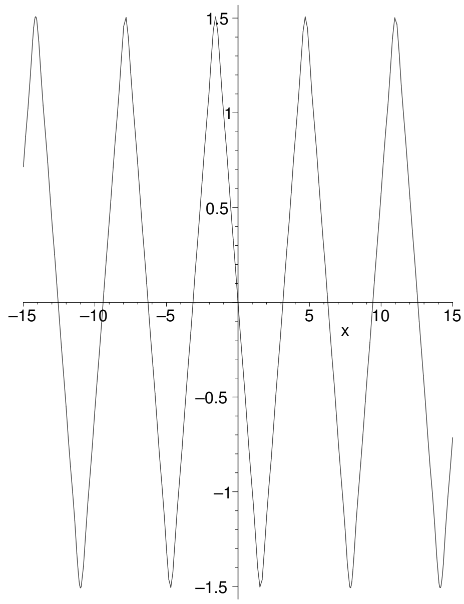


Abbildung 3: Näherung für h, $n = 10$, $-15 \leq x \leq 15$

dann vielleicht durch Koeffizientenvergleich die Koeffizienten der Fourierreihe für y bestimmen - Alltagspraxis des Physikers.

33 Gauß' Methode der kleinsten Quadrate

Die Klasse 8b soll das Hookesche Gesetz bestätigen. Dazu hängen sie Gewichtstücke an eine Schraubenfeder und messen jeweils die Verlängerung der Schraubenfeder. Denken wir uns Messwerte aus. Die Anzahl der Gewichtstücke bezeichnen wir dabei mit x, die Verlängerung mit y.

x	0	1	2	3
y	0	1	3	4

Eigentlich sollten die Messpunkte auf einer Geraden liegen[27], also sollten für geeignete Zahlen m, b die Gleichungen

$$y_i = mx_i + b \quad \text{für } i = 1, \ldots, 4 \tag{78}$$

gelten. Diesen Satz von Gleichungen kann man bequem vektoriell schreiben:

$$\vec{y} = m\vec{x} + b\vec{e} \quad \text{mit } \vec{e} = \vec{e}_1 + \ldots \vec{e}_4 \tag{79}$$

Der Vektor $\vec{e}$ ist also dabei der, dessen Einträge sämtlich $= 1$ sind.

Was bringt das? Nun, wir erkennen, dass alle Wertesätze $\vec{y}$, die sich aus dem gegebenen Wertesatz $\vec{x}$ durch eine echte Geradenvorschrift ergeben, im von $\vec{x}$ und $\vec{e}$ erzeugten Teilraum

$$V := \langle \vec{x}, \vec{e} \rangle \tag{80}$$

des $\mathbb{R}^4$ liegen, und zu jedem dieser Wertesätze $\vec{y} \in V$ gehört genau eine Geradenvorschrift. Der durch eine Messung ermittelte Wertesatz $\vec{y}$ gehört in aller Regel nicht zu V. Was nun? Klar, man bildet die orthogonale Projektion[28] des gemessenen $\vec{y}$ in V, und so erhält man den Vektor von V, der $\vec{y}$ am nächsten liegt. Die Gerade zu diesem Vektor ist die sogenannte Ausgleichsgerade

[27] In der Praxis ist das natürlich nie der Fall. Entsprechend habe ich y nach der Vorschrift $y = \sqrt{2}x$ berechnet und auf Einer gerundet.

[28] Die Richtung der orthogonalen Projektion ist durch die Normale der Ebene gegeben.

zur Wolke der Messpunkte nach der Methode der kleinsten Quadrate. Berechne sie zu den Beispielwerten und schreibe eine Handreichung für den praktisch tätigen Naturforscher, wie er seine Messwertesätze vernünftig behandeln sollte.

Meine 8-er wissen all das natürlich nicht. Sie zeichnen einfach die Punkte in ein Koordinatensystem und legen per Augenmaß eine Ausgleichsgerade hindurch. Was sollen sie auch sonst machen?

Abbildung 1 gibt einen Hinweis, wie die Methode zu ihrem Namen kommt. Du siehst die Messpunkte und die Ausgleichsgerade mit der Gleichung $y = g(x) = mx + b$. Bei jedem Messpunkt ist ein kleines Quadrat eingezeichnet, seine Kantenlänge ist der Unterschied des y-Werts des Messpunkts und des y-Werts des Geradenpunkts zum gleichen x-Wert. Die Summe der Flächeninhalte ist

$$\sum_{i=1}^{4}(y_i - g(x_i))^2 \quad ,$$

und das ist nichts anderes als

$$(\vec{y} - \vec{v}) * (\vec{y} - \vec{v}) \quad \text{mit} \quad \vec{v} = \begin{pmatrix} g(0) \\ g(1) \\ g(2) \\ g(3) \end{pmatrix} \quad .$$

Mit anderen Worten: zu jedem $\vec{v} \in V$ gehört eine Gerade in der Zeichenebene. Das Quadrat der Länge der Differenz $\vec{y} - \vec{v}$ ist gleich der Summe der Flächeninhalte der kleinen Quadrate in Abbildung 1 für die zu $\vec{v}$ gehörige Gerade. Für $\vec{v} \in V$ nimmt die Länge der Differenz $\vec{y} - \vec{v}$ den kleinstmöglichen Wert an, wenn $\vec{v}$ eben die orthogonale Projektion von $\vec{y}$ in V ist. Die Gaußsche Ausgleichsgerade hat also die kleinstmögliche Quadratsumme.

34 Stochastische Matrizen: Arbeitsaufträge für Gruppen am 25.1.2002

Ein Vektor heißt Wahrscheinlichkeitsvektor (W-Vektor), wenn seine Einträge sämtlich ≥ 0 sind und die Summe der Einträge $= 1$ ist. Eine Matrix heißt stochastisch, wenn ihre Spaltenvektoren sämtlich W-Vektoren sind.

Die folgenden Themen sollen in Gruppen bearbeitet werden.

34.1 Janusz' Performance im Unterricht und weitere Beispiele

Überprüfe mit dem Rechner, ob sich $P^n\vec{x}$ tatsächlich dem berechneten Fixvektor nähert, und sieh dir auch P^n an. Hast du eine Vermutung, wie die Grenzmatrix aussehen könnte?

Nimm dir nun einmal Viktor vor. Zustände: träumt, passt auf, ist nicht da. Wähle die Übergangswahrscheinlichkeiten selbst und rechne die Sache durch.

Weiteres Beispiel: Ein Käfer läuft auf einem Tetraeder über die Kanten von Ecke zu Ecke. In jeder Ecke wählt er irgendwie zufällig, auf welcher Kante er weiterlaufen will. In der Spitze des Tetraeders sitzt ein Spinne. Wenn er die Kante zur Spinne wählt, ist es um ihn geschehen. Kläre Käfers Schicksal.

34.2 Bergtrinker

Ein Betrunkener wankt von der Berghütte nach Hause. Der Pfad ist vier Schritte breit. Jeden Vorwärtsschritt führt er in einer von drei Weisen aus: einen Schritt geradeaus, einen Schritt nach rechts und einen geradeaus, einen Schritt nach links und einen geradeaus.

Rechne zwei Versionen: Kammweg mit Abgründen rechts und links, Hangweg mit Abgrund links und Steilhang rechts (der ihn quasi reflektiert, wenn er da gegen läuft). - Vergiss auch nicht, aus den Ergebnissen der Rechnung Lehren fürs Leben zu ziehen.

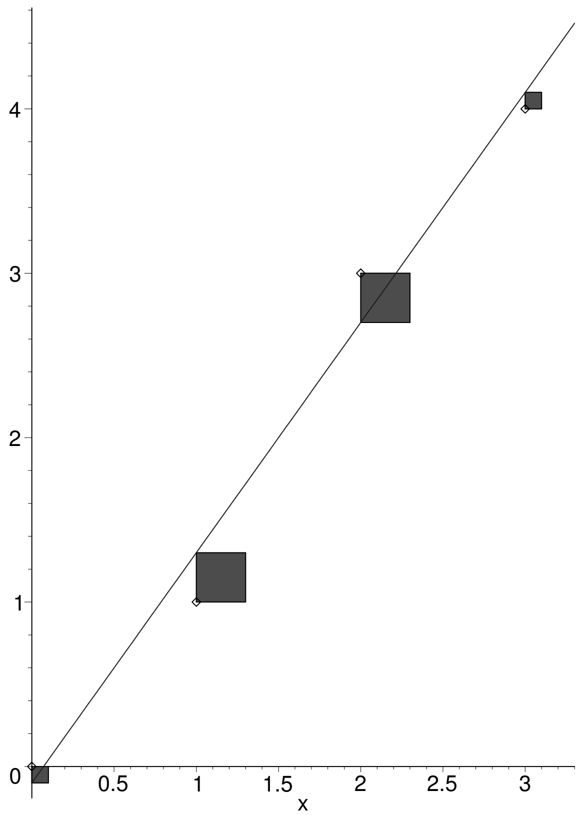


Abbildung 4: Wie die Methode zu ihrem Namen kommt

34.3 Geometrische Interpretation stochastischer Matrizen

Schreibe eine stochastische 3×3-Matrix P hin, meinetwegen die Übergangsmatrix aus der Wetteraufgabe aus dem Buch. Welche Punktmenge bilden die W-Vektoren? Stelle sicher, dass für jeden W-Vektor $\vec{x}$ der Vektor $P\vec{x}$ wieder W-Vektor ist. Berechne einen W-Vektor, der unter $\varphi : \vec{x} \mapsto P\vec{x}$ fest bleibt (heißt: Fixvektor von P). Nimm einen W-Vektor $\vec{y}$ und vergleiche $|\vec{x} - \vec{y}|$ mit $|P\vec{x} - P\vec{y}|$ (das ist übrigens $|\vec{x} - P\vec{y}|$). Wenn du zeigen kannst, dass der Unterschied durch die Anwendung von P kleiner geworden ist, kannst du vielleicht das Rätsel der Grenzmatrix lösen.

35 Differentialgeometrie

Der Titel klingt interessant, aber, ihr Leute, wir werden nicht weit kommen. Die Messung von Krümmungen und von Flächeninhalten zum Beispiel muss ich euch leider schuldig bleiben. Die Frau Ministerin will nicht, dass ihr allzuviel lernt; vielleicht schadete das eurer Gesundheit. - Nun will ich die knappen Ressourcen nicht mit Geschwätz vergeuden. Worum geht es also? Um Kurven und Flächen und deren Eigenschaften. Gewisse Kenntnisse hast du bereits, du kennst Parameterdarstellungen von Kurven aus der Analysis und Geraden und Ebenen, Längen und Winkel aus der linearen Algebra. Daran knüpfen wir an.

35.1 Tangentialvektoren einer Kurve

Wenn ich zur Beschreibung einer Kurve, die du aus 12.1 kennst, die Sprache von 13.1 heranziehe, ist eine Kurve gegeben durch eine Darstellung der Form

$$\vec{p}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \tag{81}$$

mit Funktionen $x(t), y(t)$ und $z(t)$, die auf einem Intervall $[a; b]$ definiert sind. Du kannst an einen beweglichen Punkt denken, der sich zur Zeit t im Raumpunkt mit dem Ortsvektor $\vec{p}(t)$ befindet.[29]

In der Zeitspanne Δt ändert sich der Ort des Punktes um $\vec{p}(t+\Delta t) - \vec{p}(t)$. Dividierst du diesen Vektor durch Δt und bildest anschließend den Grenzwert $\Delta t \to 0$, erhältst du den Tangentialvektor $\vec{p}\,'(t)$. Dieser Vektor gibt praktisch die Geschwindigkeit des beweglichen Punkts zur Zeit t an, die ja eine gerichtete Größe ist. Seine Länge ist die Bahngeschwindigkeit.

$$\frac{1}{\Delta t}\left(\vec{p}(t+\Delta t) - \vec{p}(t)\right) \xrightarrow[\Delta t \to 0]{} \vec{p}\,'(t) \tag{82}$$

Es ist klar, dass

$$\vec{p}\,'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix} \tag{83}$$

ist, und dies macht man sich bei der konkreten Berechnung des Tangentialvektors zunutze.

Beispiel Durch

$$\vec{p}(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ t \end{pmatrix}$$

ist eine Schraubenlinie gegeben. Der Tangentialvektor

$$\vec{p}\,'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \\ 1 \end{pmatrix}$$

hat die konstante Länge $\sqrt{2}$. Natürlich ist er am Nullpunkt des Systems anzutragen, aber oft ist hilfreich, sich ihn bei $\vec{p}(t)$ angeheftet zu denken.

35.2 Flächen

Wir wollen uns nur mit Flächen beschäftigen, die uns quasi als Graphen von Funktionen

$$z = f(x, y) \tag{84}$$

begegnen. Die Punkte einer solchen Fläche sind dann die mit den Ortsvektoren

$$\vec{p}(x, y) = \begin{pmatrix} x \\ y \\ f(x, y) \end{pmatrix} \quad ,$$

wobei die (x, y) ein vernünftiges „Gebiet" der xy-Ebene füllen müssen. Du hast eine Reihe von Beispielen gesehen, die (nach den ebenen Flächen) einfachsten sind das Rotationsparaboloid mit $z = 9 - x^2 - y^2$ und die Sattelfläche mit $z = xy$.

[29] Wenn $x(t), y(t), z(t)$ von der Form $mt + n$ sind, hast du es mit einer Geraden zu tun. Und die Zahl der Koordinaten muss nicht drei sein. Man kann zwei Koordinaten nehmen, dann handelt es sich um eine ebene Kurve, aber du kannst auch mehr als drei Koordinaten nehmen.

Wenn du den Punkt (x, y) der xy-Ebene eine Kurve durchlaufen lässt, durchläuft der zugehörige Punkt

$$\begin{pmatrix} x \\ y \\ f(x,y) \end{pmatrix}$$

eine Raumkurve, die in der Fläche liegt. Die wichtigsten Beispiele sind die Geraden (x, b) bzw. (a, y) der xy-Ebene, die parallel zu den Koordinatenachsen verlaufen. Dazu gehören die Kästchenlinien deines Heftes. Die zugehörigen Raumkurven bilden ein Netz in der Fläche, sie heißen Koordinatenlinien.

Tangentialebene und Normalenvektor

Betrachte einen Punkt einer vernünftigen Fläche, sie sei durch $z = f(x, y)$ gegeben. Die Tangentialvektoren aller Raumkurven durch den Punkt, die in der Fläche verlaufen, bilden eine Ebene, die sogenannte **Tangentialebene**. Zur Berechnung verwenden wir Koordinatenlinien: Der betrachtete Punkt P gehöre zum Punkt (a, b) der xy-Ebene, seine z-Koordinate sei also $f(a, b)$. Die Koordinatenlinien sind dann die beiden Kurven mit den Parameterdarstellungen

$$\vec{p}_x(t) = \begin{pmatrix} t \\ b \\ f(t,b) \end{pmatrix} \quad \text{und} \quad \vec{p}_y(t) = \begin{pmatrix} a \\ t \\ f(a,t) \end{pmatrix} \quad . \tag{85}$$

Wenn wir die Tangentialvektoren berechnen wollen, müssen wir im Funktionsterm $f(x, y)$ eine Koordinate festhalten und nach der anderen ableiten. Diese Ableitungen heißen **partielle** Ableitungen von f, sie werden mit

$$\frac{\partial f}{\partial x}(x,y) \quad \text{bzw.} \quad \frac{\partial f}{\partial y}(x,y)$$

bezeichnet. Wenn wir diese Symbolik verwenden, erhalten wir für die Tangentialvektoren der Koordinatenlinien in P bzw. den Normalenvektor $\vec{n}$ der Tangentialebene die Vektoren

$$\vec{p}_x\,'(a,b) = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(a,b) \end{pmatrix}, \quad \vec{p}_y\,'(a,b) = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(a,b) \end{pmatrix} \quad \text{bzw.} \quad \vec{n} = \begin{pmatrix} \frac{\partial f}{\partial x}(a,b) \\ \frac{\partial f}{\partial y}(a,b) \\ -1 \end{pmatrix} \quad . \tag{86}$$

Es ist klar, dass die beiden Tangentialvektoren die Tangentialebene erzeugen. Deshalb ist es nicht schwer, für die (im Flächenpunkt P befestigte) Tangentialebene eine Punkt-Richtungs-Form

$$\vec{q}(x,y) = \begin{pmatrix} a \\ b \\ f(a,b) \end{pmatrix} + (x-a)\begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(a,b) \end{pmatrix} + (y-b)\begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(a,b) \end{pmatrix} \tag{87}$$

und eine Gleichung

$$x\frac{\partial f}{\partial x}(a,b) + y\frac{\partial f}{\partial y}(a,b) - z = a\frac{\partial f}{\partial x}(a,b) + b\frac{\partial f}{\partial y}(a,b) - f(a,b) \tag{88}$$

hinzuschreiben (die letzte Gleichung lautet nur $\vec{n} * \vec{x} = \vec{n} * \vec{p}$).

In der Nähe des Berührpunkts kann man näherungsweise die Fläche durch die (im Berührpunkt befestigte) Tangentialebene ersetzen. Entfernt man sich vom Punkt (a, b) der xy-Ebene um Δx in x-Richtung und um Δy in y-Richtung, ändert sich der Funktionswert von f folglich etwa um

$$\Delta z \approx \frac{\partial f}{\partial x}(a,b)\Delta x + \frac{\partial f}{\partial y}(a,b)\Delta y \quad . \tag{89}$$

Das ist völlig analog zum eindimensionalen Fall, und wir werden es auch genau so benutzen, um ein zweidimensionales Newton-Verfahren zu gewinnen.

Dass alle Tangentialvektoren dieser Kurven im betrachteten Punkt überhaupt eine Ebene bilden, ist eigentlich höchst erstaunlich. Stelle dir vor, eine Platte berührte ein Modell des Paraboloids in einem festen Punkt. Das ist, anschaulich gesehen, nur auf eine Weise möglich, nämlich so, dass die Platte mit der (in den Punkt verschobenen) Tangentialebene zusammenfällt.

Beispiel Für das durch $z = f(x,y) = 9 - x^2 - y^2$ gegebene Rotationsparaboloid erhalten wir

$$\vec{p_x}\,'(x) = \begin{pmatrix} 1 \\ 0 \\ -2x \end{pmatrix}, \quad \vec{p_y}\,'(y) = \begin{pmatrix} 0 \\ 1 \\ -2y \end{pmatrix} \quad \text{bzw.} \quad \vec{n} = \begin{pmatrix} -2x \\ -2y \\ -1 \end{pmatrix} \quad .$$

Wir haben Tangentialebene und Normalenvektor benutzt, um die Richtung des stärksten Anstiegs in einem Punkt der Fläche zu finden und um sagen zu können, in welche Richtung ein Strahl reflektiert wird, der die Fläche in einem Punkt trifft.

35.3 Übungen

1) Die Punkte $A(1|0)$, $B(3|0)$ und $C(1|2)$ sind Eckpunkte eines Dreiecks der xy-Ebene. Berechne die Innenwinkel des zugehörigen Dreiecks in der durch $z = 9 - x^2 - y^2$ bzw. der durch $z = xy$ gegebenen Fläche.
2) Was ist mit den Seitenlängen der Dreiecke in Aufgabe 1? Kannst du sie berechnen?
3) Berechne für beide Flächen aus Aufgabe 1 jeweils das Volumen des Körpers zwischen der xy-Ebene und der Fläche, dessen Grundfläche das Dreieck ABC ist.
4) Es sei $z = f(x,y) = x^2y^3$. Bestimme eine Gleichung der Schnittgeraden der Tangentialebene an die Fläche im Punkt P mit der x-Koordinate 2 und der y-Koordinate 3 und der xy-Ebene.
5) Zur Strecke mit den Endpunkten $A(4|-1)$ und $B(4|1)$ der xy-Ebene gehört eine Kurve in der durch $z = \sqrt{25 - x^2 - y^2}$ gegebenen Fläche. Berechne ihre Länge.
Übrigens: diese Kurve ist **nicht** die kürzeste Verbindung der Flächenpunkte zu A und B, die innerhalb der Fläche verläuft.
6) Markus' Problem: Finde möglichst viel über die Schnittkurve einer Ebene und eines Rotationsparaboloids heraus! Nimm dazu konkret unser Paraboloid mit $z = 9 - x^2 - y^2$ und die Ebene mit der Gleichung $3x - z = 0$.

36 Klausur 13.2

1) Es sei $z = f(x,y) = x^3 - 3xy^2$. Wir betrachten die Fläche mit der Gleichung $z = f(x,y)$ und den Punkt $P(2,1,z)$ der Fläche.
a) Gib eine Punkt-Richtungsform und eine Gleichung der Tangentialebene an die Fläche im Punkt P an.
b) Stelle dir die Fläche als Oberfläche eines Körpers aus poliertem Stahl vor. Ein Lichtstrahl geht vom Nullpunkt des Systems aus und trifft im Punkt P auf die Fläche. Wie verläuft der Strahl nach der Reflektion?
c) Im Punkt P sitze ein kleines Krabbeltier. In welche Richtung muss es laufen, wenn es möglichst steil bergan will? In welche Richtung muss es laufen, wenn es auf gleicher Höhe bleiben will?
d) Nun lassen wir das Krabbeltier längs der durch

$$\vec{p}(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ z(t) \end{pmatrix}$$

gegebenen Kurve über die Fläche krabbeln. Welches sind die höchsten, welches die tiefsten Punkte seiner Wanderung?
e) Unter welchem Winkel schneiden sich die Koordinatenlinien, die in der Fläche durch den Punkt $Q(2,1,z)$ verlaufen?

f) Wie groß ist das Volumen des Körpers zwischen Fläche und xy-Ebene über dem Quadrat der xy-Ebene mit den Eckpunkten $(1,0),(2,0),(2,1)$ und $(1,1)$?
g) Kann die Fläche Mulden oder Gipfel haben? (Hinweis: Dort müssten beide partiellen Ableitungen von f den Wert 0 haben)

2) Die Funktionen $\vec{a} = (x \mapsto \sin(x))$, $\vec{b} = (x \mapsto \cos(x))$ und $\vec{c} = (x \mapsto 1)$ erzeugen einen Teilraum V des Vektorraums aller auf $\mathbb{R}$ definierten Funktionen.
a) Beweise ordentlich, dass $\vec{a}, \vec{b}, \vec{c}$ ein linear unabhängiges System bilden, die Dimension von V also drei ist.
b) Die Abbildung φ ordne jeder Funktion

$$\vec{v} = r\vec{a} + s\vec{b} + t\vec{c} = (x \mapsto r\sin(x) + s\cos(x) + t)$$

ihre Ableitung zu. Schreibe $\varphi(\vec{v})$ wieder als Linearkombination von $\vec{a}, \vec{b}$ und $\vec{c}$ und gib eine Matrix A an, so dass im Spaltenvektor

$$A \begin{pmatrix} r \\ s \\ t \end{pmatrix}$$

gerade die Vorfaktoren dieser Linearkombination stehen.
c) Auf V verwenden wir das übliche Skalarprodukt

$$(x \mapsto f(x)) * (x \mapsto g(x)) = \int_0^\pi f(x)g(x)\,dx \quad .$$

Berechne die Längen von $\vec{a}$ und von $\vec{c}$ und den Winkel zwischen $\vec{a}$ und $\vec{c}$, die durch dieses Skalarprodukt gegeben sind.
d) Integriere einmal partiell, dann solltest du sehen, dass $\vec{a}$ und $\vec{b}$ bezüglich unseres Skalarprodukts orthogonal sind. Schreibe den Vektor $\vec{d}$ hin, durch den du $\vec{c}$ ersetzen müsstest, damit du eine orthogonale Basis $\vec{a}, \vec{b}, \vec{d}$ von V erhältst. Berechnen musst du ihn nicht.

3) Wenn Jonas, der Partylöwe, heute ausgeht, geht er morgen mit der Wahrscheinlichkeit $\frac{1}{2}$ wieder aus. Wenn er heute zu Hause bleibt, geht er morgen mit der Wahrscheinlichkeit $\frac{3}{4}$ aus. Zeichne unseren Graphen mit den Knoten „bleibt zu Hause“ und „geht aus“, schreibe die Übergangswahrscheinlichkeiten ordentlich an die Pfeile und stelle die Übergangsmatrix auf. Schätze die Wahrscheinlichkeit, dass Jonas heute in drei Wochen ausgeht.

4) Nun die lange erwartete Hinz-und-Kunz-Geschichte zur Belohnung für die, die zur Originalklausur kommen. Heute haben sich die beiden ein heikles Thema vorgenommen, sie reden nämlich über Fourierreihen:

„Kenne ich“, beginnt Kunz. „Man hat paarweise orthogonale Vektoren $\vec{v}_1, \vec{v}_2, \ldots$ aus einem Funktionenraum und stellt irgendeine Funktion $\vec{w}$ des Raums näherungsweise als Linearkombination der $\vec{v}_i$ dar. Je mehr man von den $\vec{v}_i$ benutzt, desto besser ist die Näherung. Für die $\vec{v}_i$ nimmt man bei den Fourierreihen zum Beispiel

$$\vec{v}_i = (x \mapsto \sin(ix))$$

und als Skalarprodukt das Integral von 0 bis π.“

Hinz schaut nachdenklich in seinen Bierkrug. „Ja, so rechnet man das, und die Näherung wird auch immer besser. Aber es ist doch geradezu ein Wunder, dass sie beliebig gut wird. Deine $\vec{v}_i$ erzeugen einen unendlich-dimensionalen Teilraum des Funktionenraums, aber der ganze Funktionenraum ist noch riesig viel größer...“

Hinz ist ein kluger Kopf, seine scharfsinnigen Erwägungen kannst du nur zur Kenntnis nehmen. Berechne immerhin den Koeffizenten r_i von $\vec{v}_i$, den man bekommt, wenn man Kunzens Fourierreihe von $\vec{w} = (x \mapsto 1)$ bildet. Du darfst dabei neben Kunzens Aussagen noch benutzen, dass

$$\int_0^\pi (\sin(nx))^2\,dx = \frac{\pi}{2}$$

ist. Skizziere dann, wie du dir den Graphen der Funktion vorstellst, die durch die Fourierreihe gegeben ist (beachte, dass du nur Sinusfunktionen verwandt hast!) und begründe anschaulich-geometrisch, dass die Näherung immer besser werden sollte, je mehr $\vec{v}_i$ man verwendet. (Jedenfalls wird die Näherung garantiert nicht schlechter...)

37 Zugabe: Das Newton-Verfahren

Dies hier bekommst du hinzu, weil es schöne Mathematik ist und weil du Einsicht gewinnen kannst. Prüfungsrelevant ist es nicht.

37.1 Das alte Newton-Verfahren

Wenn man eine Gleichung $f(x) = 0$ nicht exakt lösen kann, kann man zu der Funktion f mit der Gleichung $y = f(x)$ übergehen und versuchen, mit dem Newton-Verfahren näherungsweise eine Nullstelle der Funktion zu berechnen. Es ist klar, dass die Nullstellen der Funktion genau die Lösungen der Gleichung sind. Du weißt noch gut, wie das Newton-Verfahren funktioniert. Man beginnt mit einem Schätzwert x_0 und berechnet $f(x_0)$. Das ist natürlich in aller Regel von Null verschieden. Deshalb verändert man x_0 um eine Korrektur und erhält den verbesserten Schätzwert

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \quad . \tag{90}$$

Wie kommt man auf den Korrekturterm? Man ersetzt den Graphen der Funktion durch die Tangente an den Graphen im Punkt $(x_0, f(x_0))$. Der neue Näherungswert x_1 ist die Nullstelle der Tangente. Da müssen wir jetzt nicht groß was rechnen: Die Tangente hat an der Stelle x_0 die Höhe $f(x_0)$. Um welche Abweichung Δx müssen wir x_0 verändern, so dass wir auf die Höhe 0 kommen? Aus $-f(x_0) = \Delta y = f'(x_0)\Delta x$ folgt sofort für die Korrektur

$$\Delta x = -\frac{f(x_0)}{f'(x_0)} \quad . \tag{91}$$

37.2 Verallgemeinerung auf Funktionen von zwei Variablen

Wir fragen nun nach Lösungen einer Gleichung

$$f(x, y) = 0 \quad . \tag{92}$$

Wenn wir einfach nachbauen, was wir eben gemacht haben, beginnen wir mit einem Startpunkt (x_0, y_0) der xy-Ebene und berechnen seinen Funktionswert $f(x_0, y_0)$. In aller Regel wird das nicht Null sein. Wir ersetzen den Graphen von f, also die Fläche mit der Gleichung $z = f(x, y)$, durch die Tangentialebene im Punkt $(x_0, y_0, f(x_0, y_0))$ und fragen nach einer Korrektur des Startwerts (x_0, y_0), der uns in einen Schnittpunkt der Tangentialebene mit der xy-Ebene bringt. Diese Korrektur $(\Delta x, \Delta y)$ soll eine Änderung $\Delta z = -f(x_0, y_0)$ der Höhe des zugehörigen Tangentialebenenpunkts herbeiführen. Es muss also gelten (erinnere dich an unsere Gleichung der Tangentialebene):

$$\frac{\partial f}{\partial x}(x_0, y_0)\Delta x + \frac{\partial f}{\partial y}(x, y)\Delta y = -f(x_0, y_0) \quad . \tag{93}$$

Na schön. Die Tangentialebene schneidet die xy-Ebene in der Regel in einer ganzen Geraden, und schon die Schnittmenge der Fläche $z = f(x, y)$ mit der xy-Ebene ist in der Regel eine Kurve oder sogar mehrere. Wenn unser Problem so gestellt sein soll, dass es mit numerischen Mitteln vernünftig behandelt werden kann, brauchen wir ausser der Gleichung (92) eine zweite Gleichung

$$g(x, y) = 0 \quad , \tag{94}$$

die gleichzeitig erfüllt sein soll. Diese liefert eine zweite Bedingung

$$\frac{\partial g}{\partial x}(x_0, y_0)\Delta x + \frac{\partial g}{\partial y}(x, y)\Delta y = -g(x_0, y_0) \quad . \tag{95}$$

Unsere Korrektur $(\Delta x, \Delta y)$ ist dann Lösung eines LGS

$$\begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \\ \frac{\partial g}{\partial x}(x_0, y_0) & \frac{\partial g}{\partial y}(x_0, y_0) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = -\begin{pmatrix} f(x_0, y_0) \\ g(x_0, y_0) \end{pmatrix} \quad . \tag{96}$$

Aus dieser Gleichung berechnen wir die Korrektur $(\Delta x, \Delta y)$ zu

$$\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = -\begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \\ \frac{\partial g}{\partial x}(x_0, y_0) & \frac{\partial g}{\partial y}(x_0, y_0) \end{pmatrix}^{-1} \begin{pmatrix} f(x_0, y_0) \\ g(x_0, y_0) \end{pmatrix} \quad . \tag{97}$$

Die Koeffizientenmatrix in (96) heißt die **Jacobi-Matrix** der Funktion, die das Paar (x, y) auf $(f(x, y), g(x, y))$ abbildet, an der Stelle (x_0, y_0). Ich verwende für diese Matrix das Symbol $J(x_0, y_0)$.

Die Rekursionsvorschrift für das Newton-Verfahren lautet dann so:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = J(x_0, y_0) \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \tag{98}$$

Ein mit Maple gerechnetes Beispiel findest du als Anlage.

38 Abiturklausur

Aufgabe 1

a) Es sei $y = f(x) = xe^{-x}$. Berechne die ersten drei Ableitungen von f und gib einen Term für die n-te Ableitung an. Ein strenger Beweis für dessen Richtigkeit ist nicht verlangt.

b) Betrachten wir die durch $y = f_t(x) = (x - t)e^{-x}$ gegebene Kurvenschar. Skizziere, wie der Graph von f_t aussieht, so dass man die wesentlichen Eigenschaften erkennt.

c) Auf welcher Kurve liegen die Hochpunkte der Scharkurven?

d) Wie groß ist der Inhalt des Flächenstücks zwischen der x-Achse und dem Kurvenstück, das oberhalb der x-Achse liegt? (Zwischenergebnis: der Inhalt ist e^{-t})

e) Begründe, dass man f_t nur für $t = 0$ für die Dichtefunktion einer stetig verteilten Zufallsgröße X verwenden könnte (das ist übrigens das f aus Teil a). Schreibe einen Rechenausdruck für den Erwartungswert von X hin, rechne ihn aber nicht aus.

f) Schreibe die Taylorreihe für das f aus a) hin (Entwicklungspunkt ist 0, wie bei uns üblich).

Aufgabe 2

Unsere Freunde Hinz und Kunz unterhalten sich angeregt beim Bier. Hören wir eine Weile zu.

Kunz: Eine Basis für eine Vektorraum zu wählen, ist doch eine mühsame Sache. Ich frage mich, ob es riskant ist, einfach Vektoren in der richtigen Anzahl zufällig zu wählen. Bilden solche Vektoren nicht in der Regel eine Basis? Lass uns das doch einmal für den $\mathbb{R}^3$ überlegen.

Hinz: Hm, einen Vektor aus dem $\mathbb{R}^3$ zufällig wählen kann man nicht. Aber man könnte sich ja auf eine Kugel um den Nullpunkt vom Radius 4 meinetwegen beschränken und daraus drei Vektoren zufällig wählen. Die Wahrscheinlichkeit, dass die drei dann eine Basis bilden, kann man schon angeben. Aber sehr realistisch ist das Modell nicht, denke ich.

Kunz: Ja, wenn ich Vektoren wähle, nehme ich natürlich welche mit ganzzahligen Koordinaten. Wählen wir also zufällig drei Vektoren $\vec{a}, \vec{b}, \vec{c}$, deren Koordinaten ganzzahlig, mindestens 0 und höchstens 4 sind, und suchen wir die Wahrscheinlichkeit p, dass sie linear unabhängig sind.

Hinz: Ich denke, diese Wahrscheinlichkeit p ist mindestens

$$\frac{124}{125} \cdot \frac{120}{125} \cdot \frac{100}{125} \quad .$$

Machen wir doch eine Simulation auf dem Rechner.

Gesagt, getan. Die beiden schreiben ein kleines Programm, das drei Vektoren, wie sie Kunz in seinem zweiten Beitrag beschrieben hat, zufällig wählt und gleich prüft, ob sie ein linear unabhängiges System bilden. Bei 3000 Versuchen erhalten sie für den Anteil der Basen den Wert $0,885$.

a) Mit welcher Wahrscheinlichkeit erhält man eine Basis des $\mathbb{R}^3$, wenn man so vorgeht, wie Hinz in seinem ersten Beitrag vorschlägt?

b) Verträgt sich das Ergebnis der Simulation mit dem Wert, den Hinz in seinem zweiten Beitrag für p nennt? Gib eine qualifizierte Antwort.

c) Wie kommt Hinz auf den Ausdruck für p in seinem zweiten Beitrag? Und hast du eine Ahnung, wieso er vorsichtig sagt, p sei mindestens so groß wie der genannte Wert?

d) Ich habe den Rechner einmal mit dem Programm der beiden drei konkrete Vektoren wählen lassen:

$$\vec{a} = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix}, \quad \vec{c} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

Bilden sie eine Basis des $\mathbb{R}^3$?

e) Wende auf die Vektoren aus e) unser Orthogonalisierungsverfahren nach Gram-Schmidt an: Gib die ersten beiden Vektoren der neuen Basis konkret an und erläutere dabei den geometrischen Hintergrund des Verfahrens. Für den dritten Basisvektor brauchst du nur einen Rechenausdruck hinzuschreiben.

Aufgabe 3

Aus der xy-Ebene erhebt sich ein gläserner Berg. Seine Oberfläche ist gegeben durch die Vorschrift

$$z = f(x, y) = 9 - x^2 - 4y^2 \quad .$$

a) Beschreibe den Berg: Wie sieht er aus?
b) Wir betrachten das Teilstück des Bergkörpers, das auf dem Rechteck der Punkte der xy-Ebene mit $0 \le x \le 2$ und $0 \le y \le 1$ steht. Wie groß ist sein Volumen?
c) Gib einen Rechenausdruck für das gesamte Volumen des Berges an, werte ihn aber nicht aus.
d) Es sei nun P der Punkt der Bergoberfläche mit den Koordinaten $x = 2$, $y = 1$ und $z = f(2; 1)$. Wir denken uns drei Wege auf der Bergoberfläche, die durch P laufen, und zwar

$$\vec{p}_1(t) = \begin{pmatrix} t \\ 1 \\ f(t,1) \end{pmatrix}, \quad \vec{p}_2(t) = \begin{pmatrix} 2 \\ t \\ f(2,t) \end{pmatrix} \quad \text{und} \quad \vec{p}_3(t) = \begin{pmatrix} 2t \\ t \\ f(2t,t) \end{pmatrix} \quad .$$

Die Vektoren

$$\vec{b}_1 = \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix}, \ \vec{b}_2 = \begin{pmatrix} 0 \\ 1 \\ -8 \end{pmatrix} \quad \text{und} \quad \vec{b}_3 = \begin{pmatrix} 2 \\ 1 \\ -16 \end{pmatrix}$$

sind jeweils Tangentialvektoren der Kurven im Punkt P. Rechne dies für $\vec{b}_3$ nach.
e) Bestätige rechnerisch, dass $\vec{b}_3 \in \langle \vec{b}_1, \vec{b}_2 \rangle$ ist, und kommentiere.
f) Gib eine Punkt-Richtungsform und eine Gleichung für die Ebene E durch P mit den Richtungsvektoren $\vec{b}_1$ und $\vec{b}_2$ an.
g) Ein Laserstrahl geht vom Punkt $Q(\frac{1}{2}; 12; 0)$ der xy-Ebene aus, trifft im Punkt P auf den Glasberg und wird dort reflektiert. Gib je einen Richtungsvektor für den Strahl und für den reflektierten Strahl an.