

# Repetitorium 5

Rückblende: In den letzten Wochen haben wir grundlegende Konzepte und Modelle der akademischen Elektrotechnik kennengelernt.

Heute unternehmen wir einen Seitensprung und wiederholen Konzepte und Methoden der Mathematik, die vor allem in der Hauptvorlesung benötigt werden. Leider ist Alzheimer nicht nur bei uns Alten, sondern auch bei den jungen StudentInnen ein allgegenwärtiges Problem, daher die Wiederholung. Wie ich immer betont habe, habe ich die Mathematik „nicht erfunden“ und so manche Tipps und Tricks sind so gestaltet, dass die praktische Anwendung im Vordergrund gegenüber der mathematischen Korrektheit ist.

## Zentralthema „Ausführbarkeit von Rechenoperationen“

Eine zentrale Grundlage der Mathematik ist die Sicherstellung, dass die notwendigen Rechenoperationen bedenkenlos ausgeführt werden können.

Der Ausgangspunkt mathematischer Tätigkeit ist ein mathematischer Raum. Darunter versteht man eine (beispielweise über Ordnungsrelationen) strukturierte Menge mathematischer Elemente wie Zahlen, Vektoren oder Funktionen. Je nach den speziellen Eigenschaften eines solchen mathematischen Raums, oft auch zusammen mit Verknüpfungen, werden Zahlenräume, (lineare) Vektorräume oder Funktionenräume definiert.

Zahlenmengen sind die Grundlage der Vektorräume und Funktionenräume. Zusammen mit einer oder mehreren mathematischen Operationen entsteht ein Verknüpfungsgebilde. Die Gruppentheorie gibt Auskunft, ob die Zusammenfügung von Zahlenmenge und mathematischen Operationen Sinn ergibt oder nicht, das heißt, ob diese Operationen in der gegebenen Zahlenmenge tatsächlich bedenkenlos ausgeführt werden können.

Von einer Gruppe spricht man, falls für eine Menge zusammen mit einer Verknüpfung je zweier Elemente dieser Menge, hier geschrieben als  $a * b$ , die folgenden Anforderungen erfüllt sind:

- Abgeschlossenheit: Die Verknüpfung zweier Elemente der Menge ergibt wiederum ein Element derselben Menge.
- Assoziativgesetz: Für die Verknüpfung ist die Klammerung unerheblich, das heißt, es gilt  $(a * b) * c = a * (b * c)$  für alle  $a, b, c$ .
- Existenz eines neutralen Elementes: Es gibt ein Element  $e$  in der Menge, das bezüglich der Verknüpfung nichts bewirkt, also ein  $a * e = e * a = a$  für alle  $a$ .
- Existenz der inversen Elemente: Zu jedem Element  $a$  gibt es bezüglich der Verknüpfung ein Umkehr – Element, also ein  $a^*$ . Dieses hat die Eigenschaft, beim Verknüpfen mit  $a$  das neutrale Element zu ergeben:  $a * a^* = a^* * a = e$ .

Am Anfang stand die intuitiv scheinende Konstruktion der natürlichen Zahlen zusammen mit der Operation der Addition: Das elementare Verknüpfungsgebilde  $(\mathbb{N}, +)$ . Dieses Verknüpfungsgebilde weist die elementaren Eigenschaften einer Gruppe nicht auf: Weder das neutrale noch die inversen Elemente sind Elemente der natürlichen Zahlen. Die Erweiterung auf die ganzen Zahlen ist nötig. Diese bilden zusammen mit der Addition eine Gruppe.

Um auch Multiplizieren zu können, ist die Erweiterung der Gruppeneigenschaften zum Körper erforderlich. Ein Körper ist eine ausgezeichnete algebraische Struktur, in der die Addition,

Subtraktion, Multiplikation und Division auf eine bestimmte Weise durchgeführt werden können. Die Bezeichnung „Körper“ wurde im 19. Jahrhundert von Richard Dedekind eingeführt.

Die wichtigsten Körper, die in fast allen Gebieten der Mathematik benutzt werden, sind der Körper  $\mathbb{Q}$  der rationalen Zahlen, der Körper  $\mathbb{R}$  der reellen Zahlen und der Körper  $\mathbb{C}$  der komplexen Zahlen. Will man sich nicht explizit festlegen, welche Zahlenmenge gemeint ist, schreibt man allgemein  $K$  und nennt diese (skalare) Menge einen Skalarkörper.

Ein Körper muss neben der vorausgesetzten Abgeschlossenheit also folgende Einzelaxiome erfüllen:

➤ Additive Eigenschaften:

- Assoziativgesetz:  $a + (b + c) = (a + b) + c$
- Kommutativgesetz:  $a + b = b + a$
- Es gibt ein neutrales Element:  $0 + a = a$
- Zu jedem Element in  $K$  existiert das additive Inverse:  $(-a) + a = 0$

➤ Multiplikative Eigenschaften:

- Assoziativgesetz:  $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
- Kommutativgesetz:  $a \cdot b = b \cdot a$
- Es gibt ein neutrales Element:  $1 \cdot a = a$
- Zu jedem Element in  $K$  ausgenommen dem Nullelement existiert das multiplikative Inverse  $a^{-1}$ :  $a^{-1} \cdot a = 1$

➤ Zusammenspiel von additiver und multiplikativer Struktur:

- Links – Distributivgesetz:  $a \cdot (b + c) = a \cdot b + a \cdot c$
- Rechts – Distributivgesetz:  $(b + c) \cdot a = b \cdot a + c \cdot a$

Will man auf einem Raum eine Metrik einführen, beispielsweise die alte euklidische Metrik, die einem Vektor eine nichtnegative Zahl zuordnet, die die Länge des Vektors abbildet, stößt man unweigerlich auf das Problem des vollständigen Raumes. So war bereits in der Antike bekannt, dass die Länge der Diagonale eines Quadrates mit der Seitenlänge 1 keine rationale Zahl sein kann. Bis heute ist kein zwingender Beweis der Existenz der irrationalen und damit der reellen Zahlen bekannt, sie werden axiomatisch eingeführt. (Hinweis: Das schon in der Sekundarstufe oft gelehrt „Cantors zweites Diagonalargument“ ist bis heute nicht unwidersprochen.)

In der mathematischen Praxis wird man daher häufig neben den Körpereigenschaften auch die Vollständigkeit des zugrundeliegenden Zahlenkörpers fordern oder gegebenenfalls herstellen müssen. Vollständigkeit bedeutet, dass sämtliche Cauchy – Folgen zu einem Grenzwert konvergieren müssen, der Element des zugrundeliegenden Skalarkörpers ist. Sollte das einmal nicht funktionieren, muss der entsprechende Grenzwert ergänzt werden, man spricht dann von einer Vervollständigung des vorher unvollständigen metrischen Raumes.

Definition: Eine Folge  $\langle a_i \rangle$  heißt Cauchy – Folge, wenn es zu jedem  $\varepsilon > 0$  einen Index  $N$  gibt, sodass ab diesem Index alle Folgenglieder (nicht nur die benachbarten!) weniger als  $\varepsilon$  voneinander entfernt sind.

Der wesentliche Vorteil des Cauchy'schen Konvergenzbegriffes besteht darin, dass der Grenzwert der Folge weder bekannt sein muss noch überhaupt Element der Menge sein muss, über der die Cauchy – Folge konstruiert ist!

## Polynome

Ein Polynom ist die Summe von Potenzen einer Variablen mit Konstanten. Man schreibt beispielsweise

$$p(x) = \sum_0^n a_n \cdot x^n$$

und nennt eine solche Konstruktion Polynom vom Grad  $n$ . Vor allem im Sekundarbereich wird dafür auch immer wieder die Bezeichnung „Ganzrationale Funktion“ verwendet. Wenn  $n$  unendlich ist, nennt man solche Polynome auch „Potenzreihe“.

In der linearen Algebra identifiziert man Polynome bzw. Potenzreihen auch mit Vektoren, indem man die Folge der Konstanten zu einem Vektor zusammenfasst, beispielsweise

$$p(x) \triangleq (a_0, a_1, \dots, a_{n-2}, a_{n-1}, a_n)^T$$

Jedes Polynom kann (bis auf die Reihenfolge) eindeutig als Produktterm dargestellt werden:

$$p(x) = a_n \cdot \prod_0^n (x - x_n)$$

Die „lästige“ Konstante  $a_n$  kann durch Dividieren des gesamten Polynoms vorher entfernt werden, was die Darstellung vereinfacht.

Aus dieser Darstellung ist unmittelbar einsichtig, dass jedes Polynom  $n$ . Grades  $n$  Nullstellen hat. Diese sind allerdings häufig einerseits komplex, andererseits mehrfach. Die Umrechnung der beiden Darstellungen ineinander ist in der Praxis unterschiedlich schwierig: Von der Produktdarstellung zum Polynom ist eine elementare Rechenaufgabe. Der umgekehrte Weg erfordert die Berechnung der Nullstellen des Polynoms. Für Grad 1 ist das eine Aufgabe der Sekundarstufe 1.

Die Lösungsformel für quadratische Polynome lässt sich wie folgt ableiten:

$$x^2 + px + q = 0$$

$$x^2 + px = -q$$

$$x^2 + px + \left(\frac{p}{2}\right)^2 = \left(\frac{p}{2}\right)^2 - q$$

$$\left(x + \frac{p}{2}\right)^2 = \frac{p^2}{4} - q$$

$$x + \frac{p}{2} = \pm \sqrt{\frac{p^2}{4} - q}$$

$$x = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$$

Diese Formel ist dermaßen gebräuchlich, dass sie vor allem in Deutschland auch als „Mitternachtsformel“ bezeichnet wird.

Polynome dritten und vierten Grades sind an sich mittels der cardanischen Formeln direkt lösbar, der hohe Aufwand ist gegenüber verschiedenen numerischen Verfahren aber kaum mehr zu rechtfertigen. Umgekehrt gibt es Online – Rechner, die gemäß den cardanischen Formeln arbeiten.

Polynome ab dem fünften Grad sind grundsätzlich nicht mittels geschlossener Formeln lösbar (Satz von Abel – Ruffini).

Interessiert man sich nicht direkt für die Nullstellen eines Polynoms, sondern für dessen Funktionswert

$$y = p(x)$$

so ist die Funktionsauswertung durch Auswertung jedes einzelnen Terms und danach Summenbildung nicht effektiv. Klüger ist die Anwendung des Horner – Schemas, bei dem lediglich Multiplikation und Addition erforderlich sind. Es gilt allgemein: Zu einem Polynom vom Grade n

$$p(x) = a_n x^n + \dots + a_2 x^2 + a_1 x + a_0$$

ist das Horner-Schema definiert als:

$$p(x) = (\dots (a_n x + a_{n-1}) x + \dots) x + a_0$$

Beispiel:

$$p(x) = x^3 + 2x^2 + 3x + 4$$

$$p(x) = (x^3 + 2x^2 + 3x) + 4$$

$$p(x) = (x^2 + 2x + 3)x + 4$$

$$p(x) = ((x^2 + 2x) + 3)x + 4$$

$$p(x) = ((x + 2)x + 3)x + 4$$

Dadurch wird die Anzahl der elementaren Rechenschritte deutlich reduziert: Im obigen Beispiel von 5 Multiplikationen und 3 Additionen auf 2 Multiplikationen und 3 Additionen.

# Die Darstellung von Geraden

Bei der formelmäßigen Darstellung von Geraden ist zwischen der Geraden in der  $x - y$  - Ebene und einer Geraden in höheren Dimensionen zu unterscheiden.

Gerade in der Ebene werden wie folgt dargestellt:

$$\text{Implizit} \quad a x + b y + c = 0$$

Diese Darstellung ist besonders hilfreich, wenn der Normalenvektor bekannt ist, denn  $(a,b)^T$  ist bereits der Normalenvektor.

$$\text{Explizit} \quad y = k x + d$$

Diese Darstellung ist besonders hilfreich, wenn Steigung  $k$  (natürlich gleich  $y'$ ) und Achsenabschnitt auf der  $y$  - Achse ( $y = d$  für  $x = 0$ ) bekannt sind.

Man kann auch die Grundregel nützen, dass Richtungsvektor und Normalenvektor stets aufeinander normal stehen (hehe...) und daher ihr Skalarprodukt Null sein muss:

$$\vec{n} \cdot \vec{r} = 0$$

Jeder Vektor, der den „Ausgangspunkt“ der Geraden  $\vec{x}_0$  mit einem anderen Punkt der Geraden verbindet, ist ein Richtungsvektor:

$$\vec{n} \cdot (\vec{x} - \vec{x}_0) = 0$$

Durch Ausmultiplizieren erhält man die

$$\text{Normalenvektorform} \quad \vec{n} \cdot \vec{x} = \vec{n} \cdot \vec{x}_0$$

Die Umrechnung auf die implizite Darstellung ist unmittelbar einsichtig:

$$\vec{n} \cdot \vec{x} = ax + by$$

$$\vec{n} \cdot \vec{x}_0 = -c$$

Vor allem für Messaufgaben (beispielsweise die Bestimmung des Abstandes eines Punktes von einer Geraden) hilfreich ist es, wenn man statt eines beliebig langen Normalenvektors einen mit Länge 1 verwendet. Diesen nennen wir beispielsweise  $\vec{n}_e$ . So erhält man die

$$\text{Hessesche Normalform} \quad \vec{n}_e \cdot \vec{x} - d = 0$$

Für die Darstellung von Geraden in höheren Dimensionen sind all diese Formeln nicht geeignet. Man bildet die Gerade entweder als Schnittmenge zweier nichtparalleler Ebenen im Raum oder verwendet gleich die allgemeine Vektorgleichung mit Bezugspunkt und Richtungsvektor mit dem Parameter  $t$ . (Das ist übrigens meine persönliche Lieblingsdarstellung, nur am Rande.)

$$\text{Richtungsvektor(Parameter)form} \quad \vec{x} = \vec{x}_0 + t \cdot \vec{r}$$

Die Umrechnung der Parameterdarstellung einer Geraden in der Ebene auf ihre explizite Darstellung gelingt mit der offensichtlichen Identität

$$\vec{r} = \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} \frac{dx}{dy} \\ \frac{dy}{dx} \end{pmatrix} = \begin{pmatrix} 1 \\ y' \end{pmatrix}$$

leicht:

$$\vec{x} = \vec{x}_0 + t \cdot \vec{r}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + t \cdot \begin{pmatrix} 1 \\ y' \end{pmatrix}$$

Als Gleichungssystem schreiben

$$x = x_0 + t$$

$$y = y_0 + t \cdot y'$$

Parameter eliminieren

$$y = y_0 + (x - x_0) \cdot y'$$

Alternativ formulieren wir mit den Funktionswerten

$$f(x) = f(x_0) + (x - x_0) \cdot f'(x)$$

**Wie wir gleich sehen werden, ist diese Gleichung extrem bedeutend! Wer schlau ist, lernt sie bis zum sicheren und schnellen Können!**

## Die Darstellung von Ebenen

Im dreidimensionalen Raum  $\mathbb{R}^3$  ist eine Ebene direkt über die implizite Normalvektorform darstellbar:

$$a x + b y + c z + d = 0$$

Diese funktioniert ausschließlich in  $\mathbb{R}^3$ . Ähnlich wie bei der Geraden ist

$$\vec{n} = (a, b, c)^T$$

Will man Ebenen auch in höheren Räumen darstellen, ist wiederum die Parameterform hilfreich:

$$\vec{x} = \vec{x}_0 + s \cdot \vec{p} + t \cdot \vec{q}$$

Die Ebene ist also durch einen Bezugspunkt  $\vec{x}_0$  und die beiden Richtungsvektoren  $\vec{p}$  und  $\vec{q}$  definiert. In Worten: Die Menge aller Ortsvektoren  $\vec{x}$  führt zu allen Punkten der Ebene, daher können wir die Menge dieser Ortsvektoren mit der Ebene identifizieren. Um zu einem Punkt der Ebene zu gelangen, gehen wir zuerst zum Bezugspunkt, dann die Länge s den Richtungsvektor  $\vec{p}$  entlang, und dann die Länge t den Richtungsvektor  $\vec{q}$  entlang. Schon haben wir einen Punkt der Ebene erreicht.

Die Umrechnung zwischen Parameterform und implizite Normalvektorform geschieht in  $\mathbb{R}^3$  über das Kreuzprodukt:

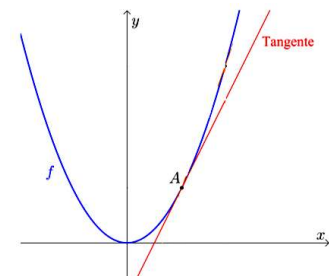
$$\vec{n} = \vec{p} \times \vec{q}$$

Damit sind a, b, c festgelegt. Dann  $x_0$  einsetzen und man erhält d.

## Die Linearisierung von (stetigen und glatten) Funktionen

Sowohl in der Mathematik wie in der Technik steht man immer wieder von dem Problem nichtlinearer Funktionen, deren Funktionsauswertung oder mathematische Weiterverarbeitung schwierig oder gar nur numerisch möglich sind. Daher wurde bereits früh versucht, solche nichtlineare Funktionen zumindest in einem begrenzten Bereich durch die leichter handhabbaren linearen Funktionen zu approximieren. Voraussetzung für die hier beschriebene Methode ist, dass die Funktion stetig und glatt, also mindestens ein Mal differenzierbar ist.

Einsichtig scheint, dass in der Umgebung eines Punktes einer Funktion die Tangente an diesem Punkt die optimale Linearisierung darstellt. Das heißt, wir kennen die Koordinaten der Anschlussstelle  $A(x_0, f(x_0))$  und die Steigung der Tangente im Punkt A, nämlich  $f'(x)$ , ausgewertet an der Stelle  $x_0$ . Dafür schreiben wir beispielsweise  $f'(x_0)$ . (Auch andere Symbole sind üblich, das müsst Ihr dann aus dem Kontext erkennen.) Diese bekannten bzw. leicht zu berechnenden Werte werden in die Geradengleichung eingesetzt.



**Achtung: In der Geradengleichung auf der letzten Seite ist die Ableitung allgemein angegeben. Das ist bei einer Geraden natürlich nebensächlich, da dort die Ableitung an allen Stellen gleich ist! Hier müssen wir den Wert der Ableitung an der Stelle  $x_0$  verwenden!**

$$f(x) = f(x_0) + (x - x_0) \cdot f'(x_0)$$

Damit ist die Approximation schon fertig, denn in der Umgebung der Anschlussstelle wird der Funktionswert der Originalfunktion wohl „so ungefähr“ gleich dem Funktionswert der Tangente an der entsprechenden Stelle sein:

$$f(x) \approx f(x_0) + (x - x_0) \cdot f'(x_0)$$

Von dieser einfachen Approximation gehen viele andere Verfahren aus, von denen wir einige wenige an dieser Stelle wiederholen:

## Das Newton'sche Näherungsverfahren

Die Frage nach den Nullstellen einer Funktion ist in der Analysis grundlegend. Doch schon bei Polynomen ab Grad 3 ist das Auffinden nicht mehr so einfach, für andere Typen von Funktionen kann das mittels geschlossener Formeln unmöglich oder zumindest überaus mühsam sein. Intuitiv hilft man sich mit grafischen Methoden oder Intervallschachtelungen. Grafische Methoden gelangen schnell an ihre Genauigkeitsgrenzen und die Intervallschachtelungen konvergieren langsam. Das Newton'sche Näherungsverfahren kann (!) eine Hilfe sein, da – wenn es funktioniert – quadratische Konvergenz vorliegt. Praktisch bedeutet das, dass pro Schritt zwei signifikante Stellen Genauigkeit gewonnen werden können.

Das Funktionsprinzip: Wir nehmen an, dass der Schnittpunkt der Tangente an die Anschlussstelle mit der x – Achse eine genauere Nullstelle ist, als die Anschlussstelle selbst. Daher schneiden wir die Tangente

$$f(x) = f(x_0) + (x - x_0) \cdot f'(x_0)$$

mit der x – Achse

$$f(x) = 0$$

Einsetzen. Dabei beachten wir, dass eine rekursive Funktion entstehen wird und ändern gleich die Indizes:  $x_0 \rightarrow x_n$  und  $x \rightarrow x_{n+1}$

$$f(x_n) + (x_{n+1} - x_n) \cdot f'(x_n) = 0$$

Elementarumformung

$$\frac{f(x_n)}{f'(x_n)} + x_{n+1} - x_n = 0$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Diese Gleichung ist das Newton'sche Näherungsverfahren.

Dieses Verfahren hat zwei Probleme: Formulierung und Konvergenz.

Zum Konvergenzproblem siehe die ausführliche Darstellung in <https://de.wikipedia.org/wiki/Newtonverfahren#Konvergenzbetrachtungen>.

**Gerade Ungeübte vergessen leicht, dass die auszuwertende Funktion die allgemeine Form  $f(x)=0$  haben muss.**

Beispiel: Die Quadratwurzel von 2 ist zu berechnen. Es gilt daher  $x^2 = 2$ . Als Startwert wählen wir  $x_0 = 1$ . Wir stellen die Gleichung auf

$$x_{n+1} = 1 - \frac{x^2| (1)}{2x| (1)} = 1 - \frac{1}{2} = \frac{1}{2}$$

Und das ist schon falsch!

Also nochmal: Richtig ist  $x^2 - 2 = 0$ , Startwert  $x_0 = 1$

$$x_1 = 1 - \frac{(x^2 - 2)| (1)}{2x| (1)} = 1 - \frac{(1 - 2)}{2} = 1 - \frac{-1}{2} = \frac{3}{2}$$

6% Fehler.

$$x_2 = \left(\frac{3}{2}\right) - \frac{(x^2 - 2)| \left(\frac{3}{2}\right)}{2x| \left(\frac{3}{2}\right)} = \left(\frac{3}{2}\right) - \frac{\left(\frac{9}{4}\right) - \left(\frac{8}{4}\right)}{\left(\frac{6}{2}\right)} = \left(\frac{18}{12}\right) - \frac{\left(\frac{1}{4}\right)}{\left(\frac{12}{4}\right)} = \left(\frac{18}{12}\right) - \left(\frac{1}{12}\right) = \frac{17}{12}$$

Und schon haben wir nur mehr 1,8‰ Fehler!

Dieses Verfahren zum Ziehen der Quadratwurzel war in Mesopotamien übrigens schon bereits zur Zeit von Hammurapi I. (ca. 1750 v. Chr.), eines Königs von Babylon, bekannt. Daher heißt es bis heute Heron-Verfahren, Heronsches Näherungsverfahren oder babylonisches Wurzelziehen. Die allgemeine Formel zur Berechnung der Quadratwurzel von  $a$  lautet

$$x_{n+1} = x_n - \frac{(x_n)^2 - a}{2 \cdot x_n} = \frac{2 \cdot (x_n)^2 - (x_n)^2 - a}{2 \cdot x_n} = \frac{2 \cdot (x_n)^2 - (x_n)^2 + a}{2 \cdot x_n} = \frac{(x_n)^2 + a}{2 \cdot x_n}$$

$$x_{n+1} = \frac{1}{2} \left( x_n + \frac{a}{x_n} \right)$$



Das schöne an dem Verfahren ist, dass es für alle Startwerte außer 0 konvergiert. Für negative Startwerte konvergiert es zur negativen Wurzel.

Eine andere wichtige Anwendung des Newton'schen Näherungsverfahrens ist die Division. Während Addition, Subtraktion und Multiplikation geradlinig zu rechnen sind, ist die Division mit ihrer grundlegenden Frage „Wie oft geht sich der Divisor im restlichen Dividenden aus?“ nicht nur für Volksschüler tückisch, auch elektronische Rechenwerke haben damit ihre Probleme. Bereits Ende des 17. Jahrhunderts fanden Newton und Raphson das heute nach ihnen benannte „Newton-Raphson-Divisions-Verfahren“. Auch dieses basiert auf dem Newton'schen Näherungsverfahren, zeigt aber wie wichtig eine kluge mathematische Darstellung der gesuchten Funktion ist.

Wir probieren wieder einmal intuitiv: Gesucht ist der Kehrwert von N, also sollte die Formel lauten

$$f(x) = \frac{1}{N}$$

Das ist natürlich falsch, weil ja die Form  $f(x)=0$  notwendig ist. Also noch einmal

$$f(x) = x - \frac{1}{N}$$

Sieht schon besser aus, da der Funktionswert 0 wird, wenn x der Kehrwert von N ist. Leider ist auch diese Formel unbrauchbar, da die Ableitung von x gleich 1 ist, was bei Newton nicht hilfreich ist. Also macht man einen Trick und bildet den Kehrwert der rechten Seite:

$$f(x) = \frac{1}{x} - N$$

Stimmt offensichtlich noch immer, denn wenn x der Kehrwert von N ist, wird der Funktionswert 0. Und in dieser Form ist das Verfahren endlich brauchbar!

Newton Basisformel

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Raphson – Ansatz einsetzen

$$x_{n+1} = x_n - \frac{x_n^{-1} - N}{-x_n^{-2}} = x_n + x_n^2(x_n^{-1} - N) = x_n + x_n - N \cdot x_n^2 = 2 \cdot x_n - N \cdot x_n^2$$

$$x_{n+1} = x_n \cdot (2 - N \cdot x_n)$$

Und das ist die Formel für das Newton-Raphson-Divisions-Verfahren. Es benötigt pro Schritt lediglich eine Subtraktion und zwei Multiplikationen und konvergiert schnell und vorhersehbar, falls ein geeigneter Startwert verwendet wird. Falls! Der Startwert wird entweder mit einer Lookup – Table oder mittels des Alternantensatzes von Tschebyscheff bestimmt. Mit dem Startwert darf man nicht sorglos umgehen, denn das Verfahren läuft schnell davon.

## Mehrdimensionale Approximation

Wir haben nun einige Verfahren zur linearen Approximation (Linearisierung) in  $\mathbb{R}^2$  kennengelernt. Gleichartig funktioniert das auch in  $\mathbb{R}^3$ . Doch selbst für die im Folgenden vorgestellte einfache Abstraktion brauchen wir ein paar Begriffe aus der Analysis:

Ein Feld ist eine Funktion, deren Argument ein Vektor ist. Es gibt zwei Arten von Feldern je nach Art des Funktionswertes: Entweder der Funktionswert ist ein Skalar (z.B. die Höhe eines Punktes  $h$  über der Ebene als Funktion der Position des Punktes bezüglich der Ebene), dann spricht man von einem Skalarfeld

$$h = f(x, y)$$

Oder aber der Funktionswert ist ein Vektor (z.B. die Kraft an einer Stelle des elektrostatischen Feldes)

$$\vec{F} = f(x, y, z)$$

Dann spricht man von einem Vektorfeld.

Man beachte, wie leicht hier die Identifikation des mathematischen Feld – Begriffes einer Art von Funktion mit der physikalischen Bedeutung als messbare Eigenschaft eines physikalischen Raumes von der Hand geht.

Die partielle Ableitung einer Funktion ist die Ableitung einer Funktion mit mehreren Argumenten nach einem dieser Argumente. Alle anderen Argumente werden als Konstante interpretiert. Zur Unterscheidung von der „normalen“ Ableitung wird ein anderes Symbol verwendet:

$$\frac{\partial f}{\partial x}$$

Beispiel: Gegeben sei die Funktion

$$z = f(x, y) = x^2 + y^2$$

Ich will wissen, wie sich die Funktionswerte verändern, wenn ich kleine Schritte entlang der  $x$  – Achse gehe:

$$\frac{\partial f}{\partial x} = 2x$$

Der geordnete Vektor aus allen partiellen Ableitungen einer Funktion  $f$  heißt Gradient von  $f$ . Man schreibt (unter anderem, die Symbolik ist wie so oft leider nicht einheitlich) Nabla  $f$ :

$$\nabla f := \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$$

Statt der für  $\mathbb{R}^2$  bzw.  $\mathbb{R}^3$  reservierten Bezeichnung  $x, y, z$  findet man auch allgemein  $x_1, x_2, x_3$ , etc.

Definition: Der Gradient ist ein mathematischer Operator, genauer ein Differentialoperator, der auf ein Skalarfeld angewandt werden kann und in diesem Fall ein Vektorfeld liefert, das Gradientenfeld genannt wird. Der Gradient steht dabei senkrecht auf der Niveaulfläche des Skalarfeldes in einem Punkt  $P$  und der Betrag des Gradienten gibt die größte Änderungsrate des Skalarfeldes im Punkt  $P$  an. Interpretiert man beispielsweise die Reliefkarte einer Landschaft als

eine Funktion  $h(x,y)$ , die jedem Ort die Höhe an dieser Stelle zuordnet, dann ist der Gradient von  $h$  an der Stelle  $(x,y)$  ein Vektor in der  $xy$ -Ebene, der in die Richtung des steilsten Anstiegs von  $h$  an dieser Stelle zeigt und dessen Länge ein Maß für die Steilheit (Steigung) ist.

Beispiel: Gegeben sei wieder die Funktion

$$z = f(x, y) = x^2 + y^2$$

Da  $z$  kein Argument ist, ist der Gradient zweidimensional

$$\nabla f = (2x, 2y)$$

Anders beispielsweise bei einem Potentialfeld: Das Potential eines Gravitationsfeldes hat die Form

$$\overrightarrow{F(\vec{r})} = m \cdot \overrightarrow{a(\vec{r})} = -m \cdot \nabla \Phi(\vec{r})$$

In Worten: Die Gravitationskraft an der Stelle des Ortsvektors  $\vec{r}$  ist erstens gleich der Masse des Probekörpers mal der Beschleunigung an dieser Stelle. Die Beschleunigung ist wiederum ein Vektor, der sich als Gradient des Potentials an der Stelle des Ortsvektors  $\vec{r}$  ergibt.

Die Berechnung des Potentials ist wahlweise eine zähe oder eine unsaubere Angelegenheit, die für gegenständliches Thema keinen Beitrag leistet. Daher nenne ich hier lediglich das Ergebnis

$$\Phi(r) = \frac{G M}{r}$$

$G$  ist die Gravitationskonstante und  $M$  die Masse des Gravitationszentrums. Zur Frage des Vorzeichens verweise ich auf die einschlägige Literatur. Man beachte, dass hier der Abstand zum Gravitationszentrum bereits skalar dargestellt wird! Das ist genau die wesentliche Vereinfachung, für die der ganze Aufwand für die Einführung des Potential – Begriffes sinnvoll war.

Mit

$$r = \sqrt{x^2 + y^2 + z^2}$$

$$\overrightarrow{F(x, y, z)} = -m \cdot \nabla \Phi(x, y, z) = -m \cdot \nabla \left( \frac{G M}{\sqrt{x^2 + y^2 + z^2}} \right)$$

$$\overrightarrow{F(x, y, z)} = -G \cdot M \cdot m \cdot \nabla \left( (x^2 + y^2 + z^2)^{-\frac{1}{2}} \right)$$

$$\overrightarrow{F(x, y, z)} = -G \cdot M \cdot m \cdot \left( \frac{\partial \left( (x^2 + y^2 + z^2)^{-\frac{1}{2}} \right)}{\partial x}, \frac{\partial \left( (x^2 + y^2 + z^2)^{-\frac{1}{2}} \right)}{\partial y}, \frac{\partial \left( (x^2 + y^2 + z^2)^{-\frac{1}{2}} \right)}{\partial z} \right)$$

$$\overrightarrow{F(x, y, z)} = -G \cdot M \cdot m \cdot \left( -\frac{1}{2} \cdot (x^2 + y^2 + z^2)^{-\frac{3}{2}} \cdot (2x), -\frac{1}{2} \cdot (x^2 + y^2 + z^2)^{-\frac{3}{2}} \cdot (2y), -\frac{1}{2} \cdot (x^2 + y^2 + z^2)^{-\frac{3}{2}} \cdot (2z) \right)$$

$$\overrightarrow{F(x, y, z)} = -G \cdot M \cdot m \cdot \left( -\frac{x}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, -\frac{y}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, -\frac{z}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \right)$$

$$\overrightarrow{F(x,y,z)} = G \cdot M \cdot m \cdot \left( \frac{x}{(x^2 + y^2 + z^2) \cdot \sqrt{(x^2 + y^2 + z^2)}}, \frac{y}{(x^2 + y^2 + z^2) \cdot \sqrt{(x^2 + y^2 + z^2)}}, \frac{z}{(x^2 + y^2 + z^2) \cdot \sqrt{(x^2 + y^2 + z^2)}} \right)$$

$$\overrightarrow{F(x,y,z)} = \frac{G \cdot M \cdot m}{(x^2 + y^2 + z^2)} \cdot \left( \frac{x}{\sqrt{(x^2 + y^2 + z^2)}}, \frac{y}{\sqrt{(x^2 + y^2 + z^2)}}, \frac{z}{\sqrt{(x^2 + y^2 + z^2)}} \right)$$

$$\overrightarrow{F(\vec{r})} = \frac{G \cdot M \cdot m}{r^2} \cdot \frac{\vec{r}}{|\vec{r}|}$$

Und das ist genau das Gravitationsgesetz in vektorieller Form.

Nach dieser Wiederholung der Grundlagen nun die Verallgemeinerung der Methode der Linearisierung auf den mehrdimensionalen Fall. Da dies ein Repetitorium der Elektrotechnik ist und keine Analysis, muss ich in Kauf nehmen, dass die richtigen Mathematiker ein großes Frottee – Tuch nehmen, richtig nass machen, zusammen winden und dann – na Ihr wisst schon. Wir verallgemeinern nämlich rein formal:

Aus einem Skalar  $x$  wird ein Vektor  $\vec{x}$

Aus einer Funktion  $f(x)$  wird ein Feld  $f(\vec{x})$

Aus der Ableitung  $f'(x)$  wird der Gradient  $\nabla f$

Und einfach einsetzen. Aus

$$f(x) \approx f(x_0) + (x - x_0) \cdot f'(x_0)$$

wird

$$f(\vec{x}) \approx f(\vec{x}_0) + (\vec{x} - \vec{x}_0) \cdot \nabla f(\vec{x}_0)$$

Und fertig ist die mehrdimensionale Approximation.

Beispiel: Gegeben sei ein Rotationsparaboloid

$$f(x,y) = x^2 + y^2$$

Gesucht ist die Tangentialebene an der Stelle (2,3)

$$f(\vec{x}) = f(\vec{x}_0) + (\vec{x} - \vec{x}_0) \cdot \nabla f(\vec{x}_0)$$

$$z = (2^2 + 3^2) + \left( \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right) \cdot \begin{pmatrix} 2x \\ 2y \end{pmatrix} \Big| \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$z = 13 + \begin{pmatrix} x - 2 \\ y - 3 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

$$z = 13 + (4x - 8 + 6y - 18)$$

$$z = 4x + 6y - 13$$

Jetzt interessiert mich noch, wie groß der Fehler zwischen der Originalfunktion und der Approximation an der Stelle (2.5,3.5) ist:

Der korrekte Wert ist

$$f(2.5, 3.5) = 2.5^2 + 3.5^2 = 18.5$$

Die Approximation liefert

$$z = 4 \cdot 2.5 + 6 \cdot 3.5 - 13 = 18$$

War doch nicht so schlecht!

## Nichtlineare Approximation: Taylor – Reihe

Für viele Zwecke ist die lineare Approximation doch nicht das optimale Verfahren, man versucht die Krümmung einer Funktion ebenso in der Approximation nachzubilden. Eine altherwürdige, aber bis heute zu diesem Zweck häufig verwendete Potenzreihe ist die Taylor – Reihe. Wir führen sie als Verallgemeinerung der linearen Approximation ein:

Sehen wir

$$f(x) \approx f(x_0) + (x - x_0) \cdot f'(x_0)$$

nur als kleinen Ausschnitt einer umfassenderen Reihe, so beginnen wir formal mit der Einfügung vorerst nicht wirksamer Koeffizienten und Potenzen

$$f(x) \approx \frac{f(x_0)}{0!} \cdot (x - x_0)^0 + \frac{f'(x_0)}{1!} \cdot (x - x_0)^1$$

Bis jetzt ist nichts passiert! Aber in dieser Gestalt können wir die Reihe weiterführen

$$f(x) = \frac{f(x_0)}{0!} \cdot (x - x_0)^0 + \frac{f'(x_0)}{1!} \cdot (x - x_0)^1 + \frac{f''(x_0)}{2!} \cdot (x - x_0)^2 + \dots + \frac{f^n(x_0)}{n!} \cdot (x - x_0)^n + \dots$$

Oder in Summenschreibweise

$$f(x) = \sum_{n=0}^{\infty} \frac{f^n(x_0)}{n!} \cdot (x - x_0)^n$$

Wobei  $f^n$  die  $n$  – te Ableitung von  $f$  ist.  $x_0$  nennen wir die Anschlussstelle, die Entwicklungsstelle oder den Entwicklungspunkt der Taylor – Reihe.

Diese Potenzreihe ist die Taylor – Reihe. Damit dreht sich die Formulierung um: Man kann die ausführlich besprochene lineare Approximation als Taylor – Reihe betrachten, die nach dem linearen Glied abgebrochen wird.

Von zentraler Bedeutung für die Nutzbarkeit der Taylor – Reihe ist die Frage der Konvergenz. Wie das bei den unendlichen Reihen so ist, kann das leicht passieren, dass die Werte über alle Grenzen gehen und die Approximation nicht funktioniert. Intuitiv ist, dass die Gefahr der Divergenz umso größer wird, je weiter  $x$  und  $x_0$  auseinander sind. Man bestimmt daher den Konvergenzradius (die Bezeichnung Radius macht natürlich nur über  $\mathbb{C}$  Sinn, über  $\mathbb{R}$  wäre vielleicht Konvergenzbereich besser, aber so ist es halt üblich), also den maximalen Abstand zwischen  $x$  und  $x_0$ , für den die Reihe noch konvergiert. Details bitte der einschlägigen Literatur zu entnehmen.

Aufgrund des binomischen Lehrsatzes wird die Funktionsauswertung von  $(x - x_0)^n$  mit steigendem  $n$  immer aufwändiger. Wenn möglich versucht man daher die Anschlussstelle 0 zu verwenden. Die Bedeutung dieser vereinfachten Reihe ist dermaßen groß, dass sie einen eigenen Namen hat: MacLaurinsche Reihe. Sie lautet in Summenschreibweise

$$f(x) = \sum_{n=0}^{\infty} \frac{f^n(0)}{n!} \cdot x^n$$

Funktionen, die an der Stelle 0 nicht definiert oder nicht differenzierbar sind, können nicht durch eine MacLaurinsche Reihe dargestellt werden.

## Die zweidimensionale Taylor – Reihe

Die mehrdimensionale Taylor – Reihe führt ganz schnell in die Abgründe der Analysis, daher zeige ich an dieser Stelle lediglich die Ergänzung der Linearisierung mit dem quadratischen Glied in zwei Dimensionen. Wir hatten die mehrdimensionale Approximation rein formal zu

$$f(\vec{x}) \approx f(\vec{x}_0) + (\vec{x} - \vec{x}_0) \cdot \nabla f(\vec{x}_0)$$

bestimmt. Dabei war der Gradient die erste mehrdimensionale Ableitung. Die zweite mehrdimensionale Ableitung ist die Hesse – Matrix. Diese ist im zweidimensionalen Fall wie folgt definiert, falls  $f(\vec{x})$  eine zweimal stetig differenzierbare Funktion ist:

$$H_f(\vec{x}) := \begin{pmatrix} \left( \frac{\partial^2 f(\vec{x})}{\partial x \partial x} \right) & \left( \frac{\partial^2 f(\vec{x})}{\partial x \partial y} \right) \\ \left( \frac{\partial^2 f(\vec{x})}{\partial y \partial x} \right) & \left( \frac{\partial^2 f(\vec{x})}{\partial y \partial y} \right) \end{pmatrix}$$

Damit bilden wir die zweidimensionale Taylor – Reihe aus

$$f(x) \approx \frac{f(x_0)}{0!} \cdot (x - x_0)^0 + \frac{f'(x_0)}{1!} \cdot (x - x_0)^1 + \frac{f''(x_0)}{2!} \cdot (x - x_0)^2$$

Formal einsetzen:

$$f(\vec{x}) \approx f(\vec{x}_0) + (\vec{x} - \vec{x}_0) \cdot \nabla f(\vec{x}_0) + \frac{H_f(\vec{x}_0)}{2!} \cdot (\vec{x} - \vec{x}_0)^2 \quad (\text{leider falsch})$$

Ganz klappt dieser rein formale Trick aber leider nicht, da die Auswertung von  $(\vec{x} - \vec{x}_0)^2$  einen Skalar ergibt, was multipliziert mit einer Matrix eine Matrix ergibt, wir aber für die Konsistenz der Reihe aber einen Skalar benötigen. Daher müssen wir wie folgt korrigieren:

$$f(\vec{x}) \approx f(\vec{x}_0) + (\vec{x} - \vec{x}_0) \cdot \nabla f(\vec{x}_0) + \frac{1}{2} \cdot (\vec{x} - \vec{x}_0)^T \cdot H_f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$$

Die Notation ist hier leider missverständlich, aber man kommt in der Praxis schnell darauf, wie das mit dem Zeilenvektor und dem Spaltenvektor rund um die Hesse – Matrix gemeint ist: Links steht der Zeilenvektor, rechts der Spaltenvektor, sonst ist die Voraussetzung für die Multiplikation nicht erfüllt.

Und so passt das mit der zweidimensionalen Taylor – Reihe.

Probieren wir die noch aus, wieder mit dem schon bekannten Rotationsparaboloid

$$f(x, y) = x^2 + y^2$$

Gesucht ist die Approximation an der Entwicklungsstelle (2,3)

$$f(\vec{x}) = f(\vec{x}_0) + (\vec{x} - \vec{x}_0) \cdot \nabla f(\vec{x}_0) + \frac{1}{2} \cdot (\vec{x} - \vec{x}_0)^T \cdot H_f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$$

$$z = (2^2 + 3^2) + \left( \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right) \cdot \begin{pmatrix} 2x \\ 2y \end{pmatrix} \bigg|_{\begin{pmatrix} 2 \\ 3 \end{pmatrix}} + \frac{1}{2} \cdot (x - 2, y - 3) \cdot H_f(\vec{x}_0) \cdot \left( \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right)$$

Die Hesse - Matrix rechnen wir getrennt aus:

$$H_f(\vec{x}) = \begin{pmatrix} \left( \frac{\partial^2(x^2 + y^2)}{\partial x \partial x} \right) & \left( \frac{\partial^2(x^2 + y^2)}{\partial x \partial y} \right) \\ \left( \frac{\partial^2(x^2 + y^2)}{\partial y \partial x} \right) & \left( \frac{\partial^2(x^2 + y^2)}{\partial y \partial y} \right) \end{pmatrix}$$

$$H_f(\vec{x}) = \begin{pmatrix} \left( \frac{\partial(2x)}{\partial x} \right) & \left( \frac{\partial(2x)}{\partial y} \right) \\ \left( \frac{\partial(2y)}{\partial x} \right) & \left( \frac{\partial(2y)}{\partial y} \right) \end{pmatrix}$$

$$H_f(\vec{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Da diese Matrix konstant ist, entfällt die Funktionsauswertung.

$$z = 13 + \begin{pmatrix} x - 2 \\ y - 3 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 6 \end{pmatrix} + \frac{1}{2} \cdot (x - 2, y - 3) \cdot \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} x - 2 \\ y - 3 \end{pmatrix}$$

$$z = 13 + (4x - 8 + 6y - 18) + (x - 2, y - 3) \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x - 2 \\ y - 3 \end{pmatrix}$$

$$z = 4x + 6y - 13 + (x - 2, y - 3) \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x - 2 \\ y - 3 \end{pmatrix}$$

$$z = 4x + 6y - 13 + (x - 2, y - 3) \cdot \begin{pmatrix} x - 2 \\ y - 3 \end{pmatrix}$$

$$z = 4x + 6y - 13 + (x - 2)^2 + (y - 3)^2$$

$$z = 4x + 6y - 13 + (x^2 - 4x + 4) + (y^2 - 6y + 9)$$

$$z = 4x + 6y - 13 + x^2 - 4x + 4 + y^2 - 6y + 9$$

$$z = x^2 + y^2$$

Passt doch und sogar überall!

# Übungen

Ist das Verknüpfungsgebilde, bestehend aus der Menge  $G = \{g: g \in \mathbb{N} \wedge g < 10\}$  und der Operation der Addition eine Gruppe? Wenn nein, warum nicht?

Ist das Verknüpfungsgebilde  $(\mathbb{Z}, +, \cdot)$  ein Körper? Hinweis: Keep it simple! „Offensichtlich“ gültige Operationen brauchen nicht einzeln bewiesen zu werden. Gib gegebenenfalls die nicht möglichen Operationen an. Nur für Streber: Wie heißt das algebraische Gebilde, dessen Eigenschaften das genannte Verknüpfungsgebilde hat?

Beweise, dass  $\sqrt{2}$  keine rationale Zahl ist.

Gib eine rationale Cauchy – Folge an, deren Grenzwert  $\sqrt{2}$  ist. Beweise, dass es eine Cauchy – Folge ist. Hinweise: Die Folge darf auch rekursiv definiert sein. Keep it simple! Das ist keine Analysis – Vorlesung!

Erfinde ein Polynom 5. Grades und eine Stelle ungleich 0. Berechne den Funktionswert klassisch und mittels Horner – Schema. Vergleiche die Anzahl an Multiplikationen und Additionen bei den beiden Verfahren.

Bestimme die Gleichungen der Geraden durch die Punkte (3,1) und (4,5)

- Implizit
- Explizit
- Normalvektorform
- Hessesche Normalform
- Richtungsvektor(Parameter)form

Bestimme den Normalabstand des Punktes (11,7) von dieser Geraden.

Gegeben sei die Funktion  $y = x^2 + 3$ . Gib die lineare Approximation an der Stelle  $x = 1$  an.

Die Aufladung eines RC – Gliedes verhalte sich gemäß der Formel  $U = 1 - e^{-t}$ . Gib die lineare Approximation an der Stelle  $t = 0$  an. Zu welchem Zeitpunkt  $t$  weicht die Approximation mehr als 1% von der korrekten Funktion ab?

Gegeben sei die Funktion  $x^3 + 3,4x^2 - 1,67x - 4,62 = 0$

Bestimme die Nullstellen mittels des Newtonschen Näherungsverfahrens. Gib die Produktdarstellung an. Hinweis: Alle Nullstellen sind rational.

Bestimme den Kehrwert der Zahl 7 mittels des Newton–Raphson-Divisions-Verfahrens. Wie findest Du einen geeigneten Startwert?



Gegeben sei die Funktion  $z = f(x, y) = 2x^2 + 3y^2$ . Finde die besten linearen Approximationen an den Stellen  $(0,0)$  und  $(1,2)$ .

Gib die Taylor – Reihe für  $\sin(x)$  an der Entwicklungsstelle  $x = 0$  an. Wie heißt diese spezielle Reihe?

Gegeben sei die Funktion  $z = f(x, y) = 2x^2 + 3y^2$ . Finde die besten quadratischen Approximationen an den Stellen  $(0,0)$  und  $(1,2)$ .