

Einführung in die mehrdimensionale Differentialrechnung

Angewandte Mathematik für Informatiker*innen

Roman Wetenkamp

30. Oktober 2021



Inhaltsverzeichnis

I. Grundlagen der Linearen Algebra	4
1. Vektoren	4
2. Vektorräume	5
3. Skalarprodukt	5
4. Normen	6
 II. Einführung in die Analysis mehrerer Veränderlicher	 8
5. Folgen, Reihen und Konvergenz	8
6. Stetigkeit	9
7. Differenzierbarkeit	9
8. Approximation	11

Vorwort

Liebe Leser*innen,

schwere Brocken schocken trocken, wenn man nicht auf sie gefasst ist. So passt es, weiter zu lesen und zu schreiben, damit offene Fragen nicht mehr offen bleiben. Als wir zu Beginn dieses Semesters den Sprung von der eindimensionalen Analysis hin zur auf einmal so stark mit der Linearen Algebra verknüpften mehrdimensionalen Analysis schaffen sollten, überforderte mich das. Diese Lücke zwischen Analysis 1 und Analysis 2 möchte ich mit diesem Skript adressieren und – falls möglich – schließen. Im Gegensatz zu hochprofessionellen Lehrbüchern für Mathematikstudierende versuche ich eher, verständlich und ausführlich genug zu bleiben. Mit Beispielen und Aufgaben versuche ich, euch einen Mehrwert zu bieten und nebenbei wieder uns alle erfolgreich ins nächste Semester aufrücken zu lassen.

Dieses Skript baue ich wie folgt auf:

1. Wiederholung wichtiger Grundlagen der linearen Algebra
2. Einführung in die mehrdimensionale Analysis (mit Wiederholung der eindimensionalen Analysis)
3. Anwendungsszenarien

Dort, wo ich es für sinnvoll erachte, sind Hinweise auf weiterführende Materialien wie Bücher oder Videos gegeben.

Viel Erfolg!

Roman Wetenkamp
Mannheim, den 30. Oktober 2021

Warnung Das Studium an einer Dualen Hochschule unterscheidet sich von dem Studium an Universitäten oder regulären Fachhochschulen insbesondere dadurch, dass aufgrund der Dualität von Theorie und Praxis meist nur die Hälfte der Zeit zur Vermittlung des Stoffes zur Verfügung steht (wenn dann auch intensiver). Daher gehen Sie bitte nicht davon aus, dass Sie dieses Skript ausreichend auf Klausuren in regulären Vollzeitstudiengängen vorbereitet!

Hinweis Dieses Dokument ist kein Vorlesungsmaterial, hat nicht den Anspruch auf Vollständigkeit und enthält mit Sicherheit Fehler. Desweiteren ist es noch lange nicht vollendet (es ist infrage zu stellen, ob es das je sein wird), und doch möchte ich Sie ermutigen, beizutragen! Jegliche Fehler, Probleme oder Anmerkungen können Sie mir gerne über das dazugehörige GitHub-Repository unter der URL <https://github.com/RWetenkamp/analysis2> zukommen lassen. Danke!

Teil I.

Grundlagen der Linearen Algebra

Erinnern wir uns zurück an das erste Semester. Nach einer Einführung in grundlegendste Themen wie Mengen, Relationen und Gruppen widmeten wir uns den eigentlichen Inhalten der linearen Algebra, Vektoren, Vektorräumen, linearen Abbildungen und Matrizen. Dieses Kapitel soll einen ganz rudimentären Überblick geben, um womöglich Vergessenes zurück ins Bewusstsein zu rufen, damit wir daran anschließen können.

1. Vektoren

Beginnen wir zunächst mit Vektoren. Ein Vektor wird häufig als ein Pfeil dargestellt, weil er physikalische Größen wie Geschwindigkeiten, Kräfte oder Ausdehnungen beschreibt. Grundsätzlich wird in der Physik unterschieden zwischen **vektoriellen** Größen, also den obigen beispielsweise, und **skalaren** Größen. Skalare Größen sind diejenigen, die durch eine einzige reelle Zahl angegeben werden, zum Beispiel Temperaturen oder Massen.

Definition 1.1. Ein n -Tupel $(a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ bezeichnen wir als **Vektor**. Die reellen Zahlen a_1, \dots, a_n heißen **Koordinaten** oder **Komponenten** eines Vektors. Wir stellen einen Vektor a als \vec{a} dar.

Vermutlich werden Sie Vektoren bereits in Ihrer Schulzeit behandelt haben. Sie haben Geraden mit Orts- und Richtungsvektoren beschrieben, mit Geraden Ebenen aufgespannt und darin Abstände zwischen Punkt, Gerade und Ebene berechnet. Dabei werden Sie nicht um die Rechenregeln bei Vektoren herum gekommen sein:

Definition 1.2. Seien $\vec{a} = (a_1, a_2, \dots, a_n), \vec{b} = (b_1, b_2, \dots, b_n)$ zwei Vektoren. Wir definieren die **Summe** beider Vektoren als

$$\vec{a} + \vec{b} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \dots \\ a_n + b_n \end{pmatrix}.$$

Ebenso definieren wir die **Multiplikation** eines Vektors und eines Skalars als

$$k \cdot \vec{a} = \begin{pmatrix} k \cdot a_1 \\ k \cdot a_2 \\ \dots \\ k \cdot a_n \end{pmatrix}.$$

Diese beiden Rechenoperationen auf Vektoren benötigen wir, um im Folgenden Vektorräume definieren zu können. Doch zunächst betrachten wir eine weitere Eigenschaft von Vektoren:

Definition 1.3. Die **Länge** oder der **Betrag** eines Vektors $\vec{a} \in \mathbb{R}^n$ wird definiert durch

$$\|\vec{a}\| = \sqrt{\sum_{j=1}^n a_j^2} = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}$$

Bemerkung 1. Der Betrag eines Vektors ist immer eine nichtnegative reelle Zahl und $\|\vec{a}\| = 0 \Leftrightarrow \vec{a} = 0$

2. Vektorräume

Definition 2.1. Sei \mathbb{K} ein beliebiger Körper und V eine Menge mit zwei Verknüpfungen:

1. **Vektoraddition:** Je zwei Elementen $\vec{a}, \vec{b} \in V$ wird ein Element $\vec{a} + \vec{b} \in V$ zugeordnet, sodass $(V, +)$ eine kommutative („abelsche“) Gruppe wird. Neutrales Element ist dabei der Nullvektor und das zu \vec{a} inverse Element ist $-\vec{a}$.
2. **Multiplikation mit einem Skalar:** Je einem $\vec{a} \in V$ und einem $k \in \mathbb{K}$ wird ein Element $k \cdot \vec{a} \in V$ (wie oben) zugeordnet.

Dann wird V ein **\mathbb{K} -Vektorraum** genannt.

Da wir nun Vektoren auch aus anderen Körpern als \mathbb{R} betrachten, genügt unsere Definition der Länge eines Vektors nicht mehr. Vorstellbar wären nun nämlich auch Vektoren eines Vektorraums über einem Restklassenkörper, wo auch andere Vektoren als der Nullvektor die Länge 0 hätten. Dem begegnen wir, indem wir eine Funktion dafür definieren.

Definition 2.2. Sei V ein reeller Vektorraum mit einer Funktion $\|\cdot\| : V \mapsto [0, \infty)$. Wenn

$$\begin{aligned} \|\vec{a}\| &> 0 \text{ für } \vec{a} \neq 0 && \text{(Positivität)} \\ \|k \cdot \vec{a}\| &= |k| \cdot \|\vec{a}\| && \text{(Homogenität)} \\ \|\vec{a} + \vec{b}\| &\leq \|\vec{a}\| + \|\vec{b}\| && \text{(Dreiecksungleichung)} \end{aligned}$$

für alle $\vec{a}, \vec{b} \in V$ und $k \in \mathbb{R}$, so nennen wir diese Funktion eine **Norm** und den Vektorraum V einen **normierten Raum**.

Damit hätten wir den Grundstein für weitere Betrachtungen und Definitionen gelegt – es liegt noch einiges vor uns.

3. Skalarprodukt

Im vorherigen Abschnitt haben wir bereits die Multiplikation eines Vektors mit einem Skalar betrachtet. Nun führen wir die Multiplikation zweier Vektoren ein. Nur diese Multiplikation nennen wir **Skalarmultiplikation**, weil wir als Ergebnis einen Skalar erhalten.

Definition 3.1. Das Produkt zweier Vektoren $\vec{a} = (a_1, a_2, \dots, a_n)$ und $\vec{b} = (b_1, b_2, \dots, b_n)$ ist definiert durch

$$\langle \vec{a}, \vec{b} \rangle = \sum_{j=1}^n a_j \cdot b_j = a_1 \cdot b_1 + a_2 \cdot b_2 + \dots + a_n \cdot b_n$$

und wird als **Skalarprodukt** oder **inneres Produkt** der beiden Vektoren bezeichnet.

Das Skalarprodukt verwenden wir nun, um folgenden Zusammenhang festzuhalten:

Satz 3.1.

$$\cos(\phi) = \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{v}\| \cdot \|\vec{w}\|}$$

Der Beweis des Satzes folgt aus der Cauchy-Schwarzschen Ungleichung:

Satz 3.2 (Cauchy-Schwarzsche Ungleichung). Sei V ein \mathbb{R} -Vektorraum mit einem Skalarprodukt. Dann gilt für $\vec{v}, \vec{w} \in V$:

$$|\langle \vec{v}, \vec{w} \rangle|^2 \leq \langle \vec{v}, \vec{v} \rangle \cdot \langle \vec{w}, \vec{w} \rangle$$

4. Normen

Nun müssen wir uns näher mit Normen beschäftigen. Wir haben gesehen, dass die Norm eine Abbildung ist, die die drei Eigenschaften Definitheit, Homogenität und die Dreiecksungleichung erfüllen muss. Wir begannen mit der Länge eines Vektors und abstrahierten darauf basierend die Norm. Die Formel für die Länge eines Vektors ist eine spezielle Norm, die sogenannte **euklidische Norm**. Nun ist aber offensichtlich, dass es weitere Abbildungen auf Vektorräumen geben kann, die die drei Axiome erfüllen - folglich gibt es eine Vielzahl verschiedener Normen, von denen wir einige im Folgenden betrachten wollen.

Definition 4.1. Wir betrachten einen Vektor $\vec{v} \in \mathbb{R}^n$.

1. Die **Summennorm** oder **1-Norm** des Vektors ist definiert durch:

$$\|\vec{v}\|_1 = \sum_{i=1}^n |v_i|$$

2. Die **euklidische Norm** oder **2-Norm** des Vektors ist definiert durch:

$$\|\vec{v}\|_2 = \sqrt{\sum_{i=1}^n |v_i|^2}$$

3. Die **Maximumsnorm** oder ∞ -**Norm** des Vektors ist definiert durch:

$$\|\vec{v}\|_\infty = \max_{i=1, \dots, n} |v_i|$$

Darauf basierend existiert folgender Satz:

Satz 4.1. Die euklidische Norm $\|\vec{v}\|_2 = \sqrt{\sum_{i=1}^n v_i^2}$ und die Maximumsnorm $\|\vec{v}\|_\infty = \max_{i=1, \dots, n} |v_i|$ sind äquivalent. Sie lassen sich gegeneinander mit Konstanten abschätzen.

$$k_1 \cdot \|\vec{v}\|_2 < \|\vec{v}\|_\infty < k_2 \cdot \|\vec{v}\|_2$$

Nun greifen wir den Vorlesungsinhalten etwas vor und führen einen weiteren Begriff ein, den wir später gebrauchen werden.

Definition 4.2. Sei E ein reeller Vektorraum. Unter einer **Metrik** verstehen wir eine Abbildung $d : E \times E \rightarrow \mathbb{R}$ mit folgenden Eigenschaften:

1. $d(\vec{x}, \vec{y}) \geq 0$ und $d(\vec{x}, \vec{y}) = 0 \Leftrightarrow \vec{x} = \vec{y}$
2. $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$
3. $d(\vec{x}, \vec{y}) \leq d(\vec{x}, \vec{z}) + d(\vec{z}, \vec{y})$ (Dreiecksungleichung)

Bemerkung 2. Die Wahl des Buchstabens d für die Metrik $d(\vec{v}, \vec{w})$ erklärt, was die Metrik angibt: Den Abstand - die Distanz - zwischen zwei Punkten bzw. Vektoren.

Damit können wir folgenden Satz formulieren:

Satz 4.2. Aus jeder Norm N kann eine Metrik d_N wie folgt abgeleitet werden:

$$d_N(\vec{v}, \vec{w}) := N(\vec{v} - \vec{w})$$

Aufgaben

Aufgabe 1 Berechnen Sie für die angegebenen Vektoren jeweils die drei definierten Normen.

1. $\vec{w}_1 = \begin{pmatrix} 17 \\ 12 \\ 24 \\ 7 \end{pmatrix}$

2. $\vec{w}_2 = \begin{pmatrix} x+3 \\ 12-z \\ \frac{x}{2} \end{pmatrix}$, mit $5 < x < z$

3. $\vec{w}_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$

Lösung (folgt)

Aufgabe 2 Zeigen Sie, dass die aus der Norm abgeleitete (vermeintliche) Metrik tatsächlich alle drei geforderten Axiome einer Metrik erfüllt.

Lösung (folgt)

Aufgabe 3 Beurteilen Sie, inwiefern Skalarprodukt, Norm und Metrik in einem Zusammenhang stehen.

Lösung (folgt)

Teil II.

Einführung in die Analysis mehrerer Veränderlicher

Nun ist der Koffer gepackt, alle wichtigen Hilfsmittel der linearen Algebra liegen vor, um sich ganz der Analysis zu verschreiben. Ich verzichte auf eine separate Wiederholung der Konzepte der Analysis I und beginne direkt damit, beides ineinander zu verweben, auch, um deutlich zu machen, dass Analysis II die Analysis I im Wesentlichen nur ergänzt.

5. Folgen, Reihen und Konvergenz

Steigen wir ein, indem wir uns die Definition von Folgen zurück ins Gedächtnis rufen.

Definition 5.1. Eine Folge ist eine endliche oder nicht-endliche Auflistung von Zahlen, die in aufsteigender Reihenfolge lückenlos durchnummeriert sind. Wir schreiben eine Folge a_n als $(a_n) = a_1, a_2, \dots, a_n, a_n + 1, \dots, a_m$.

In der Analysis I haben wir Folgen auf der Menge der reellen Zahlen betrachtet. Nun lassen sich jedoch auch Folgen von Vektoren betrachten.

Beispiel 5.1.

$$(a_n) = \left(\frac{1}{n} \right), n \in \mathbb{N}$$

Über Folgen können wir eine Reihe von Aussagen treffen. Eine davon ist eine Beschreibung des Verhaltens sehr großer Folgenglieder, das Grenzwertverhalten.

Definition 5.2. Sei eine Folge (a_n) und außerdem eine Zahl $g \in \mathbb{R}$ gegeben. g heißt **Grenzwert** von (a_n) , falls $\forall \epsilon > 0 \exists N \in \mathbb{N}$ mit:

$$|a_n - g| < \epsilon \quad \forall n \geq N$$

Eine Folge hat höchstens einen Grenzwert. Existiert ein Grenzwert, so **konvergiert** die Folge.

Mit unseren Definitionen aus der linearen Algebra können wir die Definition der Konvergenz auf Folgen von Vektoren ausweiten.

Definition 5.3. Gegeben sei eine Folge von Vektoren $(a_n) \in \mathbb{R}^n$ und eine Metrik d . Dann konvergiert (a_n) gegen einen Grenzwert $g \in \mathbb{R}^n$, falls gilt:

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \quad \forall n > N : d(g, a_n) < \epsilon$$

Nun werden wir noch einen Schritt komplizierter und betrachten die Konvergenz von Funktionen.

Definition 5.4. Sei $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine Funktion und $a \in X$. Wir nennen $L_a \in \mathbb{R}^m$ Grenzwert von f bezüglich der Annäherung von x an a , falls für jede konvergente Folge $x_n \rightarrow a$ die Folge $f(x_n)$ gegen L_a konvergiert. Dies lässt sich schreiben als

$$\lim_{x \rightarrow a} f(x) = L_a$$

Bemerkung 3. Wir betrachten hier bereits eine mehrdimensionale Funktion. Zunächst konstruieren wir eine Folge (x_n) aus dem Definitionsbereich dieser Funktion und betrachten ihr Grenzverhalten. Konvergiert jede erdenkliche Folge gegen a , so konvergiert auch die Folge nach Anwendung der Funktion f , also $f(x_n)$. Dieser Grenzwert ist selbstverständlich im Allgemeinen nicht a , sondern wird von uns mit L_a für „Limes“ bezeichnet.

6. Stetigkeit

Besinnen wir uns zurück auf das Konzept der Stetigkeit. Als ganz profane und rudimentäre Erklärung sagten wir, dass eine Funktion genau dann stetig ist, wenn wir beim Zeichnen ihres Graphen den Stift nicht absetzen müssen, also eine durchgezogene Linie ohne Sprünge zu Papier bringen können. Wir ergänzten und formalisierten diese Definition zu der folgenden:

Definition 6.1 (Epsilon-Delta-Kriterium). Eine Funktion f ist genau dann stetig in x_0 , wenn

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in D_f : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon$$

gilt.

Eine weitere, andere Definition nutzt die Grenzwertbildung aus.

Definition 6.2. Eine Funktion f heißt stetig in x_0 , falls

$$\lim_{x \rightarrow x_0} f(x) = f(x_0)$$

gilt.

Bemerkung 4. Das heißt: Existiert für jedes x , dass ich gegen x_0 laufen lasse, ein Grenzwert und stimmt dieser mit dem Funktionswert von x_0 , also $f(x_0)$ überein, dann ist diese Funktion stetig in x_0 .

Eine Funktion insgesamt ist dann stetig, wenn sie für alle Werte ihres Definitionsbereiches stetig ist.

7. Differenzierbarkeit

Stetig wird es schwerer, ebenso aber natürlich auch spannender. Nun sind wir bereit, Funktionen abzuleiten und die fundamentalen Konzepte der Analysis anzuwenden. Betrachten wir zunächst den Differentialquotienten aus der Analysis I:

Definition 7.1. Die lokale Änderungsrate einer Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ wird durch den Differentialquotienten

$$f'(x) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

angegeben.

Nun wollen wir diese Definition auf mehrdimensionale Funktionen erweitern. Dafür werden wir uns zunächst bewusst, dass wir einen weiteren Parameter benötigen: Einen Vektor, der angibt, in welche Richtung wir die Ableitung bestimmen wollen.

Definition 7.2. Sei $f : U \rightarrow \mathbb{R}$ eine Funktion. Für einen Vektor $\vec{h} \in \mathbb{R}^n$ und einen Punkt $a \in U$ heißt der Grenzwert

$$\delta_n f(a) := \lim_{t \rightarrow 0} \frac{f(a + t\vec{h}) - f(a)}{t}$$

die **Richtungsableitung** von f am Punkt a in Richtung \vec{h} .

Ein Spezialfall dessen ist die **partielle Ableitung** mit einem Einheitsvektoren als Richtungsangabe.

Definition 7.3. Für die Standardbasisvektoren \vec{e}_i nennen wir

$$\frac{\delta f(a)}{\delta x_i} = \delta_{e_i} f(a) = \lim_{t \rightarrow 0} \frac{f(a + t\vec{e}_i) - f(a)}{t}$$

die **partielle Ableitung** von f in a nach x_i .

Darauf aufbauend definieren wir nun die partielle Differenzierbarkeit.

Definition 7.4. Eine Funktion $f : U \rightarrow \mathbb{R}$ heißt partiell differenzierbar im Punkt $a \in U$, falls die partiellen Ableitungen

$$\frac{\delta f(a)}{\delta x_1}, \dots, \frac{\delta f(a)}{\delta x_n}$$

existieren und stetig sind. In diesem Fall nennen wir die Matrix

$$df(a) := \begin{pmatrix} \frac{\delta f(a)}{\delta x_1} & \dots & \frac{\delta f(a)}{\delta x_n} \end{pmatrix}$$

das **Differential** von f im Punkt a .

Dasselbe auf dem Kopf ist die Definition des Gradienten:

Definition 7.5. Der Vektor

$$\nabla f(a) := \begin{pmatrix} \frac{\delta f(a)}{\delta x_1} \\ \dots \\ \frac{\delta f(a)}{\delta x_n} \end{pmatrix}$$

wird als **Gradient** bezeichnet. Ferner gilt:

$$df(a) \cdot h = \langle \nabla f(a), h \rangle$$

Für das Differential gelten zudem eine Reihe von Eigenschaften:

Satz 7.1. Für das Differential einer differenzierbaren Funktion $f : U \rightarrow \mathbb{R}$ gilt für alle $a \in U$:

- $df(a)(h) := df(a) \cdot h$ - Dies ist eine lineare Abbildung von $\mathbb{R}^n \rightarrow \mathbb{R}$.
- $df(a) \cdot h = \delta_n f(a)$
- $d(f \cdot g) = g(a)df + f(a)dg$
- $d(f + g) = df + dg$

Dies ist nun genug Handwerkszeug, um die erste Ableitung mehrdimensionaler Funktionen zu berechnen. Für bestimmte Fragen, wie beispielsweise nach Extrema, wird ebenfalls die zweite Ableitung verlangt. Oben haben wir eine Funktion in jede Dimension (also nach jeder Variable) abgeleitet und einen Vektor (den Gradienten) erhalten. Wollen wir nun die zweite Ableitung bestimmen, müssen wir diesen Vektor ableiten.

Definition 7.6. Sei $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Unter der **Hesseschen Matrix** von f im Punkt $x \in U$ versteht man die $n \times n$ -Matrix

$$(\text{Hess}f)(x) = \begin{pmatrix} \frac{\delta}{\delta x_1} \frac{\delta}{\delta x_1} & \cdots & \frac{\delta}{\delta x_1} \frac{\delta}{\delta x_n} \\ \vdots & \ddots & \vdots \\ \frac{\delta}{\delta x_n} \frac{\delta}{\delta x_1} & \cdots & \frac{\delta}{\delta x_n} \frac{\delta}{\delta x_n} \end{pmatrix}$$

Diese Matrix ist nach dem Satz von Schwarz symmetrisch.

Bemerkung 5. Die Reihenfolge, nach welcher Variable zuerst abgeleitet wird, spielt folglich keine Rolle. Die Annahme, man können hier die Ableitungen miteinander multiplizieren, ist falsch. Viel mehr handelt es sich bei den Delta-Brüchen um eine Kennzeichnung, dass hier eine Ableitung gebildet wurde.

8. Approximation

In der Analysis I haben wir uns bereits mit Approximationen beschäftigt. Damit bezeichnen wir das Verfahren, eine komplizierte Funktion durch eine einfachere – häufig lineare oder polynomiale – anzunähern. Diese Annäherungen sind immer bis auf einen bestimmten Fehler genau und genügen so vielen Fragestellungen aus der Anwendung. Ein prominentes Beispiel ist die **Taylor-Approximation**, die wir an dieser Stelle kurz wiederholen wollen.

Definition 8.1. Das n -te **Taylorpolynom** von f an der Entwicklungsstelle a ist definiert durch:

$$T_n f(x, a) := \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} \cdot (x - a)^k$$

Der Genauigkeit halber wird ein Restglied $R_n f(x, a)$ hinzu addiert.