

Mathematik für Volkswirte

Mathematical Methods for Economists

Josef Leydold

Institute for Statistics and Mathematics · WU Wien

Wintersemester 2017/18



© 2009–2017 Josef Leydold

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Austria License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/at/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

Einleitung

Literatur

- ▶ KNUT SYDSÆTER, PETER HAMMOND
Essential Mathematics for Economics Analysis
Prentice Hall, 3rd ed., 2008
- ▶ KNUT SYDSÆTER, PETER HAMMOND, ATLE SEIERSTAD, ARNE STRØM
Further Mathematics for Economics Analysis
Prentice Hall, 2005
- ▶ ALPHA C. CHIANG, KEVIN WAINWRIGHT
Fundamental Methods of Mathematical Economics
McGraw-Hill, 2005
- ▶ JOSEF LEYDOLD
Mathematik für Ökonomen
3. Auflage, Oldenbourg Verlag, München, 2003

Weitere Übungsbeispiele

Die Bücher aus der Reihe *Schaum's Outline Series* (McGraw Hill) bieten umfangreiche Sammlungen von Musteraufgaben und Übungsbeispielen mit zum Teil ausführlichen Lösungen. Insbesondere seien die folgenden Bücher erwähnt:

- ▶ SEYMOUR LIPSCHUTZ, MARC LIPSON
Linear Algebra, 4th ed., McGraw Hill, 2009
- ▶ RICHARD BRONSON
Matrix Operations, 2nd ed., McGraw Hill, 2011
- ▶ ELLIOT MENDELSON
Beginning Calculus, 3rd ed., McGraw Hill, 2003
- ▶ ROBERT WREDE, MURRAY R. SPIEGEL
Advanced Calculus, 3rd ed., McGraw Hill, 2010
- ▶ ELLIOTT MENDELSON
3,000 Solved Problems in Calculus, McGraw Hill, 1988

Über die mathematische Methode

Man kann also gar nicht prinzipieller Gegner der mathematischen Denkformen sein, sonst müsste man das Denken auf diesem Gebiete überhaupt aufgeben. Was man meint, wenn man die mathematische Methode ablehnt, ist vielmehr die höhere Mathematik. Man hilft sich, wo es absolut nötig ist, lieber mit schematischen Darstellungen und ähnlichen primitiven Behelfen, als mit der angemessenen Methode.

Das ist nun aber natürlich unzulässig.

Joseph Schumpeter (1906)

Statische (Gleichgewichts-) Analyse

- ▶ Welcher Preis herrscht in Marktgleichgewicht?
Finde den Preis bei dem Angebots- und Nachfragefunktion übereinstimmen.
- ▶ Welche Gütermengen müssen in einer Volkswirtschaft produziert werden, damit Konsum und Exporte befriedigt werden können?
Finde Inverse einer Matrix in einem Leontief Input-Output Modell.
- ▶ Wie verhält sich ein Konsument, der seinen Nutzen optimiert?
Finde des absolute Maximum der Nutzenfunktion.
- ▶ Wie lautet das optimale Produktionsprogramm einer Firma?
Finde das absolute Maximum der Erlösfunktion.

Komparativ-statische Analyse

- ▶ In welche Richtung bewegen sich die Preise, wenn das Marktgleichgewicht gestört wird?
Bestimme die Ableitung des Preises als Funktion der Zeit.
- ▶ Wie lautet der marginale Produktionsvektor, wenn sich die Nachfrage in einem Leontief-Modell ändert?
Bestimme die Ableitung einer vektorwertigen Funktion.
- ▶ Wie ändert sich der optimale Nutzen eines Konsumenten, wenn sich Einkommen oder Preise ändern?
Bestimme die Ableitung des maximalen Nutzens nach den Modellparametern.

Dynamische Analyse

- ▶ Wir kennen die Änderungsrate eines Preises nach der Zeit.
Welchen Verlauf nimmt die Preisentwicklung?
Löse eine Differential- oder Differenzgleichung.
- ▶ Welche Investitionspolitik eines Staates optimiert das
Wirtschaftswachstum?
Bestimme die Parameter einer Differentialgleichung, sodass der
Endpunkt der Lösungsfunktion maximal wird.
- ▶ Wie lautet die Anlagestrategie eines Konsumenten, die seinen
intertemporalen Nutzen maximiert.
Bestimme die Sparrate (als Funktion der Zeit), die die Summe des
diskontierten Konsums optimiert.

Lernziele – Grundlagen

- ▶ Lineare Algebra:
Matrix und Vektor · Matrixalgebra · Determinante · Eigenwerte
- ▶ Univariate Analysis:
Funktion · Graph · injektiv und surjektiv · Limes · Stetigkeit ·
Differentialquotient und Ableitung · Monotonie · konvex und
konkav
- ▶ Multivariate Analysis:
partielle Ableitung · Gradient und Jacobische Matrix · totales
Differential · implizite und inverse Funktion · Hessematrix und
quadratische Form · Taylorreihe

Lernziele – Optimierung

- ▶ Statische Optimierung:
lokale und globale Extrema · Lagrange-Funktion und Kuhn-Tucker Bedingung · Umhüllungssatz
- ▶ Dynamische Analyse:
Integration · (Systeme von) Differentialgleichung · stabiler und instabiler Fixpunkt · Sattelpunkt · Transversalitätsbedingung · Kontrolltheorie und Hamiltonfunktion

Ablauf der Lehrveranstaltung

- ▶ Eigenständiges Vorbereiten eines neuen Kapitels (Handouts).
- ▶ Präsentationen des neuen Lehrstoffes mit Beispielen.
- ▶ Hausübungen.
- ▶ Besprechung der Übungsaufgaben (mittwochs).
- ▶ Endtest.

Voraussetzungen*

Mathematische Grundkenntnisse gehören zu den Voraussetzungen zum erfolgreichen Abschluß dieser Lehrveranstaltung und sollten bereits in der Schule oder in den Einführungslehrveranstaltungen Ihres Bakkalaureatsstudiums erworben sein.

Auf der Webseite dieser Lehrveranstaltung finden Sie daher das Skriptum *Mathematik – Grundlagen*. Es enthält eine Zusammenfassung dieser Grundkenntnisse und bietet die Möglichkeit, eventuell vorhandene Wissenslücken zu beheben. Dieser Stoff ist daher auch prüfungsrelevant.

Einige der Folien behandeln trotzdem diese Grundlagen. Sie sind durch ein * im Folientitel gekennzeichnet. Diese Folien werden aber nur bei Bedarf erklärt.

Voraussetzungen – Probleme*

Folgende Aufgaben bereiten erfahrungsgemäß besondere Probleme:

- ▶ das Zeichnen (oder Skizzieren) von Funktionsgraphen,
- ▶ Äquivalenzumformungen von Gleichungen,
- ▶ das Arbeiten mit Ungleichungen,
- ▶ die korrekte Handhabung von Bruchtermen,
- ▶ das Rechnen mit Exponenten und Logarithmen,
- ▶ das unnötige Ausmultiplizieren von Produkten,
- ▶ das Verwenden der mathematischen Notation.

Die präsentierten „Lösungen“ derartiger (Teil-) Aufgaben sind überraschend oft falsch.

Inhaltsverzeichnis – I – Propädeutik

Logik, Mengen und Abbildungen

Aussagenlogik

Mengen

Abbildungen

Zusammenfassung

Inhaltsverzeichnis – II – Lineare Algebra

Matrixalgebra

Prolog

Matrix

Rechnen mit Matrizen

Vektoren

Lineare Gleichungssysteme

Das Gaußsche Eliminationsverfahren

Das Gauß-Jordansche Verfahren

Epilog

Zusammenfassung

Vektorräume

Der Vektorraum

Rang einer Matrix

Basis und Dimension

Lineare Abbildung

Inhaltsverzeichnis – II – Lineare Algebra / 2

Zusammenfassung

Determinante

Definition und Eigenschaften

Berechnung

Cramersche Regel

Zusammenfassung

Eigenwerte

Eigenwerte und Eigenvektoren

Diagonalisieren

Quadratische Form

Hauptkomponentenanalyse

Zusammenfassung

Inhaltsverzeichnis – III – Analysis

Funktionen

Reelle Funktionen

Spezielle Funktionen

Elementare Funktionen

Grenzwert

Stetigkeit

Funktionen in mehreren Variablen

Wege

Allgemeine reelle Funktionen

Zusammenfassung

Differentialrechnung

Differentialquotient

Differential

Ableitung

Monotonie

Inhaltsverzeichnis – III – Analysis / 2

Krümmung

Elastizität

Partielle Ableitung

Partielle Elastizitäten

Gradient

Totales Differential

Jacobische Matrix

Zusammenfassung

Inverse und implizite Funktionen

Inverse Funktionen

Implizite Funktionen

Zusammenfassung

Taylorreihen

Taylorreihen

Konvergenz

Inhaltsverzeichnis – III – Analysis / 3

Rechnen mit Taylorreihen
Funktionen in mehreren Variablen
Zusammenfassung

Integration

Riemann-Integral
Stammfunktion
Hauptsatz der Integral- und Differentialrechnung
Uneigentliches Integral
Differenzieren unter dem Integral
Doppelintegrale
Zusammenfassung

Inhaltsverzeichnis – IV – Statische Optimierung

Extrema

Konvexe Mengen

Konvex und konkav

Extrema

Lokale Extrema

Quasi-konvex und quasi-konkav

Umhüllungssatz

Zusammenfassung

Lagrange-Funktion

Optimierung unter Nebenbedingungen

Lagrange-Ansatz

Viele Variablen und Gleichungen

Globale Extrema

Umhüllungssatz

Zusammenfassung

Kuhn-Tucker Bedingung

Graphisches Verfahren

Optimierung unter Nebenbedingungen

Die Kuhn-Tucker Bedingung

Der Satz von Kuhn-Tucker

Zusammenfassung

Inhaltsverzeichnis – V – Dynamische Optimierung

Differentialgleichungen

Was ist eine Differentialgleichung?

Lösungstechniken

Spezielle Differentialgleichungen

Lineare Differentialgleichung 2. Ordnung

Qualitative Analyse

Zusammenfassung

Kontrolltheorie

Das Standardproblem

Zusammenfassung

Mathematischer Zweig

- ▶ Courses hold in the *international scientific language*, i.e, broken English (© Robert Trapp).
- ▶ Discuss basics of mathematical reasoning.
- ▶ Extend our tool box of mathematical methods for static optimization and dynamic optimization.
- ▶ For more information see the corresponding web pages for the courses *Mathematics I* and *Mathematics II*.

Viel Erfolg!

Kapitel 1

Logik, Mengen und Abbildungen

Aussage

Um Mathematik betreiben zu können, sind ein paar Grundkenntnisse der **mathematischen Logik** erforderlich. Im Zentrum steht dabei die Aussage.

Eine **Aussage** ist ein Satz der entweder **wahr** (W) oder **falsch** (F) ist.

- ▶ „*Wien liegt an der Donau*“ ist eine wahre Aussage.
- ▶ „*Bill Clinton war Präsident der Republik Österreich*“ ist eine falsche Aussage.
- ▶ „*19 ist eine Primzahl*“ ist eine wahre Aussage.
- ▶ „*Dieser Satz ist falsch*“ ist keine Aussage.

Elementare Aussageverbindungen

Die Aussagenlogik verknüpft einfache zu komplexeren Aussagen und gibt deren Wahrheitswert an.

Dies geschieht durch die aus der Alltagssprache bekannten Wörter „und“, „oder“, „nicht“, „wenn ... dann“, und „genau dann ... wenn“.

Aussageverbindung	Symbol	Name
nicht P	$\neg P$	Negation
P und Q	$P \wedge Q$	Konjunktion
P oder Q	$P \vee Q$	Disjunktion
wenn P dann Q	$P \Rightarrow Q$	Implikation
P genau dann, wenn Q	$P \Leftrightarrow Q$	Äquivalenz

Wahrheitswerte

Wahrheitswerte elementarer Aussageverbindungen.

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
W	W	F	W	W	W	W
W	F	F	F	W	F	F
F	W	W	F	W	W	F
F	F	W	F	F	W	W

Aussagen $P =$ „ x ist durch 2 teilbar“ und $Q =$ „ x ist durch 3 teilbar“.
Die Aussage $P \wedge Q$ ist genau dann wahr, wenn x durch 2 und 3 teilbar ist.

Negation und Disjunktion

- ▶ Die *Negation* (*Verneinung*) $\neg P$ ist nicht das „Gegenteil“ der Aussage P .

Die Verneinung von $P =$ „Alle Katzen sind grau“

ist $\neg P =$ „Nicht alle Katze sind grau“

(Und keinesfalls „Alle Katzen sind nicht grau“!)

- ▶ Die *Disjunktion* $P \vee Q$ ist im *nicht-ausschließenden* Sinn gemeint:

$P \vee Q$ ist genau dann wahr, wenn P wahr ist, oder wenn Q wahr ist, oder wenn P und Q wahr sind.

Implikation

Die Wahrheitswerte der *Implikation* $P \Rightarrow Q$ erscheinen etwas mysteriös.

Beachte aber, dass $P \Rightarrow Q$ keine Aussage über den Wahrheitswert von P oder Q macht!

Welche der beiden Aussagen ist wahr?

- ▶ „*Wenn* Bill Clinton österreichischer Staatsbürger ist, *dann* kann er zum Präsidenten der Republik Österreich gewählt werden.“
- ▶ „*Wenn* Karl österreichischer Staatsbürger ist, *dann* kann er zum Präsidenten der Republik Österreich gewählt werden.“

Die Implikation $P \Rightarrow Q$ ist äquivalent zur Aussage $\neg P \vee Q$.

Symbolisch:

$$(P \Rightarrow Q) \Leftrightarrow (\neg P \vee Q)$$

Ein einfacher logischer Beweis

Wir können den Wahrheitswert der Aussage $(P \Rightarrow Q) \Leftrightarrow (\neg P \vee Q)$ mittels Wahrheitstabellen herleiten:

P	Q	$\neg P$	$(\neg P \vee Q)$	$(P \Rightarrow Q)$	$(P \Rightarrow Q) \Leftrightarrow (\neg P \vee Q)$
W	W	F	W	W	W
W	F	F	F	F	W
F	W	W	W	W	W
F	F	W	W	W	W

Die Aussage $(P \Rightarrow Q) \Leftrightarrow (\neg P \vee Q)$ ist also immer wahr, unabhängig von den Wahrheitswerten für P und Q .

Wir sagen daher, dass die beiden Aussagen $P \Rightarrow Q$ und $\neg P \vee Q$ äquivalent sind.

Theoreme

*Mathematics consists of propositions of the form: P implies Q ,
but you never ask whether P is true. (Bertrand Russell)*

Ein **mathematischer Satz** (*Theorem, Proposition, Lemma, Korollar*) ist eine Aussage der Form $P \Rightarrow Q$.

P heißt dann eine **hinreichende** Bedingung für Q .

Eine *hinreichende* Bedingung P garantiert, dass die Aussage Q wahr ist. Q kann aber auch dann wahr sein, wenn P falsch ist.

Q heißt dann eine **notwendige** Bedingung für P , $Q \Leftarrow P$.

Eine *notwendige* Bedingung Q muss wahr sein, damit die Aussage P wahr sein kann. Sie garantiert nicht, dass P wahr ist.

Quantoren

Mathematische Texte verwenden öfters die Ausdrücke „für alle“ bzw. „es existiert ein“.

In formaler Notation werden dafür folgende Symbole verwendet:

Quantor	Symbol
für alle	\forall
es existiert ein	\exists
es existiert genau ein	$\exists!$

Mengen*

Der Begriff der *Menge* ist fundamental für die moderne Mathematik.
Wir begnügen uns mit einer höchst einfachen Definition.

Eine **Menge** ist eine Sammlung von unterscheidbaren Objekten.

Ein Objekt a einer Menge A heißt **Element** der Menge:

$$a \in A$$

Mengen werden durch *Aufzählung* oder *Beschreibung* ihrer Elemente in *geschwungenen Klammern* $\{ \dots \}$ definiert.

$$A = \{1,2,3,4,5,6\}$$

$$B = \{x \mid x \text{ ist eine natürliche Zahl und durch 2 teilbar}\}$$

Wichtige Mengen*

Symbol	Beschreibung
--------	--------------

\emptyset	leere Menge (nur in der Schule: $\{\}$)
-------------	--

\mathbb{N}	natürliche Zahlen $\{1,2,3,\dots\}$
--------------	-------------------------------------

\mathbb{Z}	ganze Zahlen $\{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$
--------------	---

\mathbb{Q}	rationale Zahlen, Bruchzahlen $\{\frac{k}{n} \mid k, n \in \mathbb{Z}, n \neq 0\}$
--------------	--

\mathbb{R}	reelle Zahlen
--------------	---------------

$[a, b]$	abgeschlossenes Intervall $\{x \in \mathbb{R} \mid a \leq x \leq b\}$
----------	---

(a, b)	offenes Intervall $\{x \in \mathbb{R} \mid a < x < b\}$
----------	---

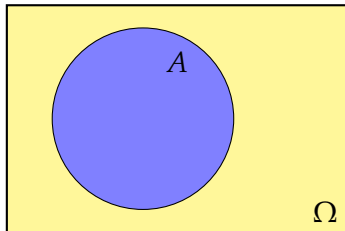
$[a, b)$	halboffenes Intervall $\{x \in \mathbb{R} \mid a \leq x < b\}$
----------	--

\mathbb{C}	komplexe Zahlen
--------------	-----------------

Venn-Diagramme*

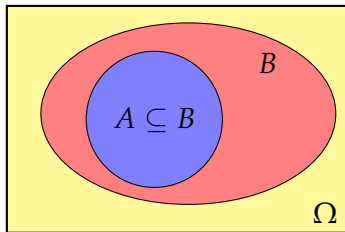
Beim Arbeiten mit Mengen nimmt man meist an, dass alle betrachteten Mengen Teilmengen einer vorgegebenen **Obermenge** Ω sind.

Mengen können durch sogenannte **Venn-Diagramme** dargestellt werden. Die Obermenge wird durch ein Rechteck, die einzelnen Mengen durch Kreise oder Ovale dargestellt.



Teilmenge*

Eine Menge A heißt **Teilmenge** von B , $A \subseteq B$, falls jedes Element von A auch Element von B ist, formal: $x \in A \Rightarrow x \in B$.



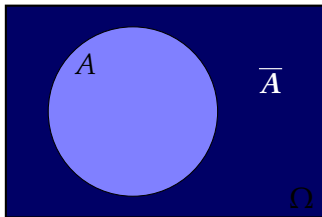
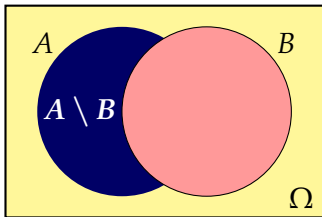
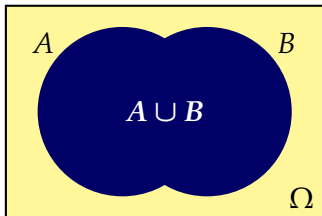
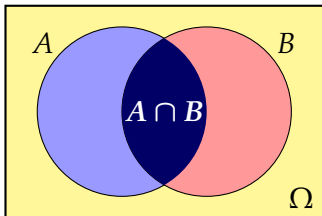
Eine Menge A heißt **echte Teilmenge** von B , $A \subset B$, falls $A \subseteq B$ und $A \neq B$.

Mengenverknüpfungen*

Symbol	Definition	Bezeichnung
$A \cap B$	$\{x x \in A \wedge x \in B\}$	Durchschnitt
$A \cup B$	$\{x x \in A \vee x \in B\}$	Vereinigung
$A \setminus B$	$\{x x \in A \wedge x \notin B\}$	Mengendifferenz
\overline{A}	$\Omega \setminus A$	Komplement
$A \times B$	$\{(x, y) x \in A, y \in B\}$	Cartesisches Produkt

Zwei Mengen A und B heißen **disjunkt** falls $A \cap B = \emptyset$.

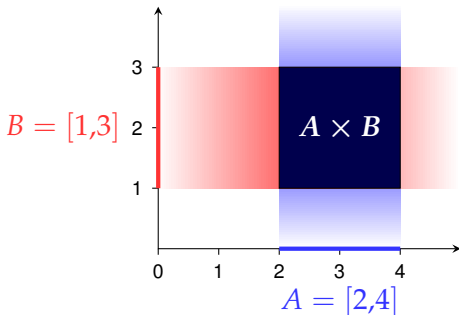
Mengenverknüpfungen*



Cartesisches Produkt*

Das Cartesische Produkt aus $A = \{0,1\}$ und $B = \{2,3,4\}$ ist $A \times B = \{(0,2), (0,3), (0,4), (1,2), (1,3), (1,4)\}$.

Das Cartesische Produkt aus $A = [2,4]$ und $B = [1,3]$ ist $A \times B = \{(x,y) \mid x \in [2,4] \text{ und } y \in [1,3]\}$.



Rechenregeln für Mengenverknüpfungen*

Regel

Bezeichnung

$$A \cup A = A \cap A = A$$

Idempotenz

$$A \cup \emptyset = A \quad \text{und} \quad A \cap \emptyset = \emptyset$$

Identität

$$(A \cup B) \cup C = A \cup (B \cup C) \quad \text{und} \\ (A \cap B) \cap C = A \cap (B \cap C)$$

Assoziativität

$$A \cup B = B \cup A \quad \text{und} \quad A \cap B = B \cap A$$

Kommutativität

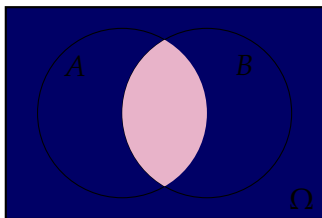
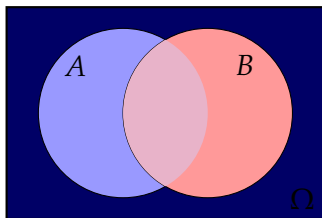
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \quad \text{und} \\ A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Distributivität

$$\bar{A} \cup A = \Omega \quad \text{und} \quad \bar{A} \cap A = \emptyset$$

Gesetz von De Morgan*

$$\overline{(A \cup B)} = \bar{A} \cap \bar{B} \quad \text{und} \quad \overline{(A \cap B)} = \bar{A} \cup \bar{B}$$



Abbildung

Eine **Abbildung** f ist definiert durch

- (i) eine **Definitionsmenge** D ,
- (ii) eine **Wertemenge** W und
- (iii) eine **Zuordnungsvorschrift**,
die jedem Element von D_f *genau ein* Element von W_f zuordnet.

$$f: D_f \rightarrow W_f, \quad x \mapsto y = f(x)$$

- ▶ x heißt **unabhängige** Variable, y heißt **abhängige** Variable.
- ▶ y ist das **Bild** von x , x ist das **Urbild** von y .
- ▶ $f(x)$ heißt **Funktionswert**, x heißt **Argument** der Abbildung.

Andere Bezeichnungen: *Funktion, Transformation*

Injektiv · surjektiv · bijektiv*

Jedes Argument besitzt immer genau ein Bild. Die Anzahl der Urbilder eines Elementes $y \in W$ kann jedoch beliebig sein. Wir können daher Funktionen nach der Anzahl der Urbilder einteilen.

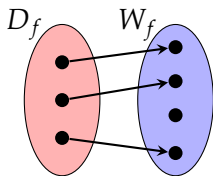
- ▶ Eine Abbildung f heißt **injektiv**, wenn jedes Element aus der Wertemenge *höchstens* ein Urbild besitzt.
- ▶ Sie heißt **surjektiv**, wenn jedes Element aus der Wertemenge *mindestens* ein Urbild besitzt.
- ▶ Sie heißt **bijektiv**, wenn sie sowohl injektiv als auch surjektiv ist.

Injektive Abbildungen haben die folgende wichtige Eigenschaft:

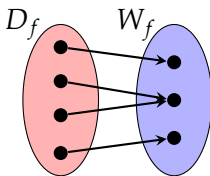
$$f(x) \neq f(y) \Leftrightarrow x \neq y$$

Injektiv · surjektiv · bijektiv*

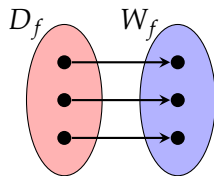
Abbildungen können durch „Pfeildiagramme“ veranschaulicht werden.



injektiv
(nicht surjektiv)



surjektiv
(nicht injektiv)



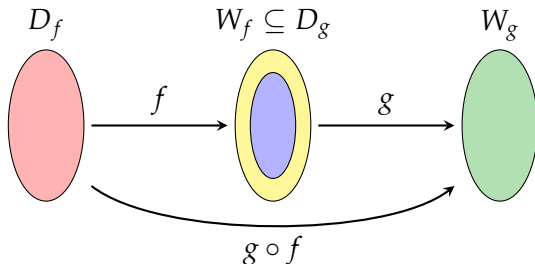
bijektiv

Zusammengesetzte Funktion*

Seien $f: D_f \rightarrow W_f$ und $g: D_g \rightarrow W_g$ Funktionen mit $W_f \subseteq D_g$.
Dann heißt die Funktion

$$g \circ f: D_f \rightarrow W_g, x \mapsto (g \circ f)(x) = g(f(x))$$

zusammengesetzte Funktion („ g zusammengesetzt f “).



Inverse Abbildung*

Bei einer **bijektiven** Abbildung $f: D_f \rightarrow W_f$ können wir jedem $y \in W_f$ sein Urbild $x \in D_f$ zuordnen.

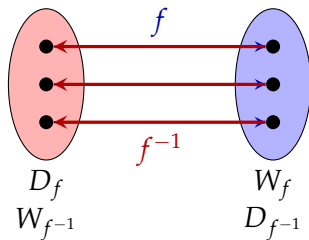
Wir erhalten dadurch wieder eine Abbildung f^{-1} mit der Definitionsmenge W_f und der Wertemenge D_f :

$$f^{-1}: W_f \rightarrow D_f, y \mapsto x = f^{-1}(y)$$

Diese Abbildung heißt **Umkehrfunktion** oder **inverse Abbildung**. Sie hat die Eigenschaft, dass für alle Elemente $x \in D_f$ und $y \in W_f$ gilt:

$$f^{-1}(f(x)) = f^{-1}(y) = x \quad \text{und} \quad f(f^{-1}(y)) = f(x) = y$$

Inverse Abbildung*



Identische Abbildung*

Die einfachste Funktion ist die **Einheitsfunktion** (oder **identische Abbildung** id), die das Argument auf sich selbst abbildet, d.h.

$$\text{id}: D \rightarrow W = D, x \mapsto x$$

Die Einheitsfunktion bei zusammengesetzten Abbildungen die Rolle der Zahl 1 bei der Multiplikation von Zahlen.

$$f \circ \text{id} = f \quad \text{und} \quad \text{id} \circ f = f$$

Insbesondere gilt:

$$f^{-1} \circ f = \text{id}: D_f \rightarrow D_f \quad \text{und} \quad f \circ f^{-1} = \text{id}: W_f \rightarrow W_f$$

Zusammenfassung

- ▶ Aussagenlogik
- ▶ Theorem
- ▶ Notwendige und hinreichende Bedingung
- ▶ Mengen
- ▶ Mengenverknüpfungen
- ▶ Abbildung
- ▶ Zusammengesetzte Funktion
- ▶ Inverse Abbildung

Kapitel 2

Matrixalgebra

Ein sehr einfaches Leontief-Modell

Eine Stadt betreibt die Unternehmen ÖFFENTLICHER VERKEHR, ELEKTRIZITÄT und GAS.

Technologiematrix und wöchentliche Nachfrage (in Werteeinheiten):

Verbrauch an für	Verkehr	Elektrizität	Gas	Konsum
Verkehr	0,0	0,2	0,2	7,0
Elektrizität	0,4	0,2	0,1	12,5
Gas	0,0	0,5	0,1	16,5

Wie groß muss die wöchentliche Produktion sein, damit die Nachfrage befriedigt werden kann?

Ein sehr einfaches Leontief-Modell

Wir bezeichnen die unbekannte Produktion von VERKEHR, ELEKTRIZITÄT und GAS mit x_1 , x_2 bzw. x_3 . Für die Produktion muss dann gelten:

Nachfrage = Produktion – interner Verbrauch

$$7,0 = x_1 - (0,0 x_1 + 0,2 x_2 + 0,2 x_3)$$

$$12,5 = x_2 - (0,4 x_1 + 0,2 x_2 + 0,1 x_3)$$

$$16,5 = x_3 - (0,0 x_1 + 0,5 x_2 + 0,1 x_3)$$

Durch Umformen erhalten wir das lineare Gleichungssystem:

$$1,0 x_1 - 0,2 x_2 - 0,2 x_3 = 7,0$$

$$-0,4 x_1 + 0,8 x_2 - 0,1 x_3 = 12,5$$

$$0,0 x_1 - 0,5 x_2 + 0,9 x_3 = 16,5$$

Wie müssen wir x_1 , x_2 und x_3 wählen?

Matrix

Eine $m \times n$ -**Matrix** ist ein rechteckiges Schema bestehend aus m Zeilen und n Spalten.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = (a_{ij})$$

Die Zahlen a_{ij} heißen **Elemente** oder **Koeffizienten** der Matrix \mathbf{A} , die Zahl i der **Zeilenindex**, die Zahl j der **Spaltenindex**.

Matrizen werden mit lateinischen Großbuchstaben bezeichnet, deren Koeffizienten mit den entsprechenden Kleinbuchstaben.

In der Literatur werden auch eckige Klammern $[a_{ij}]$ verwendet.

Vektor

- ▶ Ein (Spalten-) **Vektor** ist eine $n \times 1$ -Matrix:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

- ▶ Ein **Zeilenvektor** ist eine $1 \times n$ -Matrix:

$$\mathbf{x}^t = (x_1, \dots, x_n)$$

- ▶ Der i -te **Einheitsvektor** \mathbf{e}_i ist der Vektor, in dem die i -te Komponente gleich 1 und alle anderen gleich 0 sind.

Vektoren werden mit *kleinen* lateinischen Buchstaben bezeichnet.

Wir schreiben $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ für eine Matrix mit den Spalten(vektoren) $\mathbf{a}_1, \dots, \mathbf{a}_n$.

Spezielle Matrizen I

- ▶ Eine $n \times n$ -Matrix heißt **quadratische Matrix**.
- ▶ Eine **obere Dreiecksmatrix** ist eine quadratische Matrix, deren Elemente *unterhalb* der Hauptdiagonale alle Null sind.

$$\mathbf{U} = \begin{pmatrix} -1 & -3 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -2 \end{pmatrix}$$

- ▶ Eine **untere Dreiecksmatrix** ist eine quadratische Matrix, deren Elemente *oberhalb* der Hauptdiagonale alle Null sind.
- ▶ Eine **Diagonalmatrix** ist eine quadratische Matrix, bei der alle Elemente außerhalb der Hauptdiagonale gleich Null sind.

Spezielle Matrizen II

- ▶ Eine Matrix, in der alle Koeffizienten gleich Null sind, heißt **Nullmatrix** und wird mit $\mathbf{O}_{n,m}$ oder kurz $\mathbf{0}$ bezeichnet.
- ▶ Die **Einheitsmatrix** ist eine Diagonalmatrix, bei der die Hauptdiagonalelemente gleich 1 sind. Sie wird mit \mathbf{I}_n oder kurz \mathbf{I} bezeichnet. (In der deutschsprachigen Literatur auch mit \mathbf{E} .)

$$\mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Hinweis: Sowohl die Einheitsmatrix \mathbf{I}_n als auch die symmetrische Nullmatrix $\mathbf{O}_{n,n}$ sind ebenfalls Beispiele für Diagonalmatrizen, obere und untere Dreiecksmatrizen.

Transponierte Matrix

Die **Transponierte** A^t (oder A') einer Matrix A erhalten wir, wenn wir aus Zeilen Spalten machen und umgekehrt:

$$(a_{ij}^t) = (a_{ji})$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

Multiplikation mit einer Konstanten

- ▶ Zwei Matrizen heißen **gleich**, $\mathbf{A} = \mathbf{B}$, wenn die Anzahl der *Zeilen* und *Spalten* übereinstimmen und die Matrizen *koeffizientenweise* gleich sind, d.h. $a_{ij} = b_{ij}$.
- ▶ Eine Matrix \mathbf{A} wird mit einer Konstanten $\alpha \in \mathbb{R}$ *komponentenweise* multipliziert:

$$\alpha \cdot \mathbf{A} = (\alpha \cdot a_{ij})$$

$$3 \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 9 & 12 \end{pmatrix}$$

Addition zweier Matrizen

Zwei $m \times n$ -Matrizen **A** und **B** werden *komponentenweise* addiert:

$$\mathbf{A} + \mathbf{B} = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

Die Addition zweier Matrizen ist nur möglich, wenn die Anzahl der Zeilen und Spalten der beiden Matrizen übereinstimmen!

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}$$

Multiplikation zweier Matrizen

Das Produkt zweier Matrizen \mathbf{A} und \mathbf{B} ist nur dann definiert, wenn die Anzahl der Spalten der ersten Matrix gleich der Anzahl der Zeilen der zweiten Matrix ist.

D.h., wenn \mathbf{A} eine $m \times n$ -Matrix ist, so muss \mathbf{B} eine $n \times k$ -Matrix sein. Die Produktmatrix $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ ist dann eine $m \times k$ -Matrix.

Zur Berechnung des Elements c_{ij} der Produktmatrix wird die i -te Zeile der ersten Matrix mit der j -ten Spalte der zweiten Matrix „multipliziert“ (im Sinne eines Skalarprodukts):

$$c_{ij} = \sum_{s=1}^n a_{is} \cdot b_{sj}$$

Die Matrizenmultiplikation ist **nicht kommutativ!**

Falksches Schema

$\mathbf{A} \cdot \mathbf{B} \rightarrow$				1	2
\downarrow				3	4
				5	6
1	2	3		c_{11}	c_{12}
4	5	6		c_{21}	c_{22}
7	8	9		c_{31}	c_{32}

$$c_{21} = 1 \cdot 4 + 5 \cdot 3 + 6 \cdot 5 = 49$$

$$\mathbf{A} \cdot \mathbf{B} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 22 & 28 \\ 49 & 64 \\ 76 & 100 \end{pmatrix}$$

Nicht-Kommutativität

Achtung!

Die Matrizenmultiplikation ist **nicht kommutativ!**

Im Allgemeinen gilt:

$$\mathbf{A \cdot B \neq B \cdot A}$$

Potenz einer Matrix

$$\begin{aligned} \mathbf{A}^2 &= \mathbf{A} \cdot \mathbf{A} \\ \mathbf{A}^3 &= \mathbf{A} \cdot \mathbf{A} \cdot \mathbf{A} \\ &\vdots \\ \mathbf{A}^n &= \underbrace{\mathbf{A} \cdot \dots \cdot \mathbf{A}}_{n \text{ mal}} \end{aligned}$$

Inverse Matrix

Falls für eine quadratische Matrix \mathbf{A} eine Matrix \mathbf{A}^{-1} mit der Eigenschaft

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I}$$

existiert, dann heißt \mathbf{A}^{-1} die **inverse Matrix** von \mathbf{A} .

Die Matrix \mathbf{A} heißt **invertierbar** falls sie eine Inverse besitzt, andernfalls heißt sie **singulär**.

Achtung!

Die inverse Matrix ist nur für *quadratische* Matrizen definiert.

Rechengesetze für Matrizen

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$$

$$\mathbf{A} + \mathbf{0} = \mathbf{A}$$

$$(\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C})$$

$$\mathbf{I} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{I} = \mathbf{A}$$

$$(\alpha \mathbf{A}) \cdot \mathbf{B} = \alpha(\mathbf{A} \cdot \mathbf{B})$$

$$\mathbf{A} \cdot (\alpha \mathbf{B}) = \alpha(\mathbf{A} \cdot \mathbf{B})$$

$$\mathbf{C} \cdot (\mathbf{A} + \mathbf{B}) = \mathbf{C} \cdot \mathbf{A} + \mathbf{C} \cdot \mathbf{B}$$

$$(\mathbf{A} + \mathbf{B}) \cdot \mathbf{D} = \mathbf{A} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{D}$$

A und **B** invertierbar

$\Rightarrow \mathbf{A} \cdot \mathbf{B}$ invertierbar

$$(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}$$

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$$

$$(\mathbf{A} \cdot \mathbf{B})^t = \mathbf{B}^t \cdot \mathbf{A}^t$$

$$(\mathbf{A}^t)^t = \mathbf{A}$$

$$(\mathbf{A}^t)^{-1} = (\mathbf{A}^{-1})^t$$

Achtung!

Im Allgemeinen gilt

$$\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$$

Rechnen mit Matrizen

Für *geeignet* dimensionierte Matrizen gelten ähnliche Rechengesetze wie für reelle Zahlen. Wir müssen dabei aber beachten:

- ▶ Die Nullmatrix $\mathbf{0}$ spielt dabei die Rolle der Zahl 0.
- ▶ Die Einheitsmatrix \mathbf{I} entspricht dabei der Zahl 1.
- ▶ Die Matrizenmultiplikation ist nicht kommutativ!
Im Allgemeinen gilt $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$.

$$(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + \mathbf{A} \cdot \mathbf{B} + \mathbf{B} \cdot \mathbf{A} + \mathbf{B}^2$$

$$\begin{aligned} \mathbf{A}^{-1} \cdot (\mathbf{A} + \mathbf{B}) \cdot \mathbf{B}^{-1} \mathbf{x} &= (\mathbf{A}^{-1} \cdot \mathbf{A} + \mathbf{A}^{-1} \mathbf{B}) \cdot \mathbf{B}^{-1} \mathbf{x} \\ &= (\mathbf{I} + \mathbf{A}^{-1} \mathbf{B}) \cdot \mathbf{B}^{-1} \mathbf{x} = \\ &= (\mathbf{B}^{-1} + \mathbf{A}^{-1} \cdot \mathbf{B} \mathbf{B}^{-1}) \mathbf{x} \\ &= (\mathbf{B}^{-1} + \mathbf{A}^{-1}) \mathbf{x} \\ &= \mathbf{B}^{-1} \mathbf{x} + \mathbf{A}^{-1} \mathbf{x} \end{aligned}$$

Matrixgleichungen

Wird eine Matrixgleichung mit einer Matrix multipliziert, so muss dies auf beiden Seiten des Gleichheitszeichens von **derselben** Seite (entweder „von links“ oder „von rechts“) erfolgen!

Sei $\mathbf{B} + \mathbf{A} \mathbf{X} = 2\mathbf{A}$, wobei \mathbf{A} und \mathbf{B} bekannte Matrizen sind.
Wie lautet \mathbf{X} ?

$$\begin{array}{rcl} \mathbf{B} + \mathbf{A} \mathbf{X} = 2 \mathbf{A} & | & \mathbf{A}^{-1} \cdot \\ \mathbf{A}^{-1} \cdot \mathbf{B} + \mathbf{A}^{-1} \cdot \mathbf{A} \mathbf{X} = 2 \mathbf{A}^{-1} \cdot \mathbf{A} & & \\ \mathbf{A}^{-1} \cdot \mathbf{B} + \mathbf{I} \cdot \mathbf{X} = 2 \mathbf{I} & | & - \mathbf{A}^{-1} \cdot \mathbf{B} \\ \mathbf{X} = 2 \mathbf{I} - \mathbf{A}^{-1} \cdot \mathbf{B} & & \end{array}$$

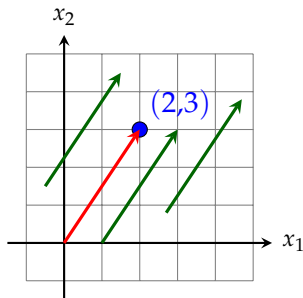
In dieser Gleichung ist natürlich darauf zu achten, dass die Matrizenoperationen tatsächlich definiert sind.

Geometrische Interpretation I

Wir haben Vektoren als Spezialfälle von Matrizen kennen gelernt.
Wir können aber Vektoren auch geometrisch interpretieren.

Wir können uns den Vektor $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ denken als

- ▶ **Punkt** (x_1, x_2) in der xy -Ebene.
- ▶ Pfeil vom Ursprung zum Punkt (x_1, x_2) (**Ortsvektor**).
- ▶ Pfeil mit gleicher Länge, Richtung und Orientierung wie dieser Ortsvektor. („**Klasse von Pfeilen**“).



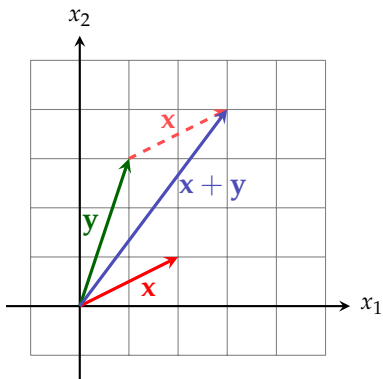
Wir wählen uns immer die Interpretation aus, die uns am besten passt.

Mit diesen Bildern können wir denken („Denkkrücke“).

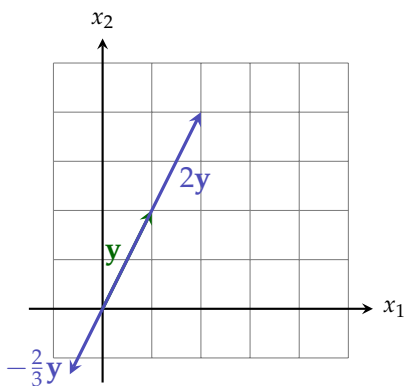
Rechnen müssen wir aber mit den Formeln!

Geometrische Interpretation II

Vektoraddition



Multiplikation mit Skalar



Skalarprodukt

Das **innere Produkt** (oder **Skalarprodukt**) zweier Vektoren \mathbf{x} und \mathbf{y} ist

$$\mathbf{x}^t \mathbf{y} = \sum_{i=1}^n x_i y_i$$

Zwei Vektoren heißen **orthogonal**, wenn $\mathbf{x}^t \mathbf{y} = 0$.

Sie stehen dann *normal* (*senkrecht, im rechten Winkel*) aufeinander.

Das innere Produkt von $\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und $\mathbf{y} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ lautet

$$\mathbf{x}^t \cdot \mathbf{y} = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 32$$

Norm

Die **Norm** $\|\mathbf{x}\|$ eines Vektors \mathbf{x} ist

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^t \mathbf{x}} = \sqrt{\sum_{i=1}^n x_i^2}$$

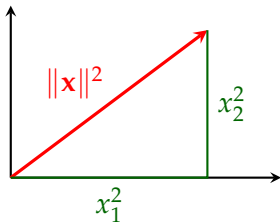
Ein Vektor \mathbf{x} heißt **normiert**, falls $\|\mathbf{x}\| = 1$.

Die Norm von $\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ lautet

$$\|\mathbf{x}\| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$$

Geometrische Interpretation

Die Norm eines Vektors kann als Länge interpretiert werden:



Pythagoräischer Lehrsatz:

$$\|\mathbf{x}\|^2 = x_1^2 + x_2^2$$

Das innere Produkt misst den Winkel zwischen zwei Vektoren.

$$\cos \angle(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x}^t \mathbf{y}}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$$

Eigenschaften der Norm

(i) $\|\mathbf{x}\| \geq 0$.

(ii) $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$.

(iii) $\|\alpha\mathbf{x}\| = |\alpha| \cdot \|\mathbf{x}\|$ für alle $\alpha \in \mathbb{R}$.

(iv) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$. (Dreiecksungleichung)

Un/gleichungen

- ▶ **Cauchy-Schwarzsche Ungleichung**

$$|\mathbf{x}^t \mathbf{y}| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$$

- ▶ **Minkowski Ungleichung** (Dreiecksungleichung)

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

- ▶ Satz von **Pythagoras**

Für orthogonale Vektoren \mathbf{x} und \mathbf{y} gilt

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

Lineares Gleichungssystem

Lineares Gleichungssystem aus m Gleichungen und n Unbekannten.

$$a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n = b_1$$

$$a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n = b_2$$

$$\vdots \quad \quad \quad \ddots \quad \quad \quad \vdots \quad \quad \quad \vdots$$

$$a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n = b_m$$

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}}_{\text{Koeffizientenmatrix}} \cdot \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{\text{Variablen}} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}_{\text{Konstantenvektor}}$$

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

Matrixdarstellung

Vorteile der Matrixdarstellung:

- ▶ Abgekürzte, kompakte Schreibweise.
- ▶ Die Anzahl der Variablen geht in dieser Darstellung nicht mehr ein.
- ▶ Die Lösungen können mit Hilfe der Matrizenrechnung berechnet und interpretiert werden.
- ▶ Wir können die einzelnen Bestandteile mit Namen versehen, etwa PRODUKTIONSVEKTOR, NACHFRAGEVEKTOR, TECHNOLOGIEMATRIX, etc. im Falle des Leontief-Modells.

Leontief Modell

Input-Output Modell mit

A ... Technologiematrix

x ... Produktionsvektor

b ... Nachfragevektor

Dann gilt: $\mathbf{x} = \mathbf{Ax} + \mathbf{b}$

Für eine vorgegebene Nachfrage **b** erhalten wir die notwendige Produktion durch

$$\begin{array}{l} \mathbf{x} = \mathbf{Ax} + \mathbf{b} \quad | \quad - \mathbf{Ax} \\ \mathbf{x} - \mathbf{Ax} = \mathbf{b} \\ (\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{b} \quad | \quad (\mathbf{I} - \mathbf{A})^{-1} \cdot \\ \mathbf{x} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} \end{array}$$

Lösung eines linearen Gleichungssystem

Es gibt drei Lösungsmöglichkeiten:

- ▶ Das Gleichungssystem hat *genau eine* Lösung.
- ▶ Das Gleichungssystem ist *inkonsistent* (nicht lösbar).
- ▶ Das Gleichungssystem hat *unendlich viele* Lösungen.

Aus der Anzahl der Gleichungen und Unbekannten kann noch nicht geschlossen werden, wie viele Lösungen ein Gleichungssystem besitzt.

Beim **Gaußschen Eliminationsverfahren** wird die erweiterte Koeffizientenmatrix (\mathbf{A}, \mathbf{b}) in die Stufenform umgeformt.

In der **Stufenform** nimmt die Anzahl der Elemente gleich 0 auf der linken Seite von Zeile zu Zeile um mindestens eins zu.

Durch **Rücksubstitution** lässt sich die Lösung bestimmen.

Gaußsches Eliminationsverfahren

Es sind (nur) die folgenden Operationen erlaubt:

- ▶ Multiplikation einer Zeile mit einer Konstanten ($\neq 0$).
- ▶ Addition des Vielfachen einer Zeile zu einer anderen Zeile.
- ▶ Vertauschen zweier Zeilen.

Diese Operationen lassen die Lösung des Gleichungssystems unverändert. (*Äquivalenzumformungen*)

Gaußsches Eliminationsverfahren

$$\begin{array}{ccc|c} 1,0 & -0,2 & -0,2 & 7,0 \\ -0,4 & 0,8 & -0,1 & 12,5 \\ 0,0 & -0,5 & 0,9 & 16,5 \end{array}$$

Wir addieren zunächst das 0,4-fache der ersten Zeile zur zweiten Zeile.
Wir schreiben dafür kurz:

$$Z2 \leftarrow Z2 + 0,4 \times Z1$$

$$\begin{array}{ccc|c} 1 & -0,20 & -0,20 & 7,0 \\ 0 & 0,72 & -0,18 & 15,3 \\ 0 & -0,50 & 0,90 & 16,5 \end{array}$$

Gaußsches Eliminationsverfahren

$$Z3 \leftarrow Z3 + \frac{0,5}{0,72} \times Z2$$

$$\begin{array}{ccc|c} 1 & -0,20 & -0,20 & 7,0 \\ 0 & 0,72 & -0,18 & 15,3 \\ 0 & 0 & 0,775 & 27,125 \end{array}$$

Rücksubstitution

$$\begin{array}{ccc|c} 1 & -0,20 & -0,20 & 7,0 \\ 0 & 0,72 & -0,18 & 15,3 \\ 0 & 0 & 0,775 & 27,125 \end{array}$$

Aus der dritten Zeile erhalten wir direkt:

$$0,775 \cdot x_3 = 27,125 \quad \Rightarrow \quad x_3 = 35$$

Restlichen Variablen x_2 und x_1 durch **Rücksubstitution**:

$$0,72 \cdot x_2 - 0,18 \cdot 35 = 15,3 \quad \Rightarrow \quad x_2 = 30$$

$$x_1 - 0,2 \cdot 30 - 0,2 \cdot 35 = 7 \quad \Rightarrow \quad x_1 = 20$$

Lösung ist eindeutig: $\mathbf{x} = (20,30,35)^t$

Beispiel 2

Suche die Lösung des Gleichungssystems:

$$3x_1 + 4x_2 + 5x_3 = 1$$

$$x_1 + x_2 - x_3 = 2$$

$$5x_1 + 6x_2 + 3x_3 = 4$$

$$\begin{array}{ccc|c} 3 & 4 & 5 & 1 \\ 1 & 1 & -1 & 2 \\ 5 & 6 & 3 & 4 \end{array}$$

$$Z2 \leftarrow 3 \times Z2 - Z1, \quad Z3 \leftarrow 3 \times Z3 - 5 \times Z1$$

$$\begin{array}{ccc|c} 3 & 4 & 5 & 1 \\ 0 & -1 & -8 & 5 \\ 0 & -2 & -16 & 7 \end{array}$$

Beispiel 2

$$Z3 \leftarrow Z3 - 2 \times Z2$$

$$\begin{array}{ccc|c} 3 & 4 & 5 & 1 \\ 0 & -1 & -8 & 5 \\ 0 & 0 & 0 & -3 \end{array}$$

Aus der dritten Zeile erhalten wir $0 = -3$, ein *Widerspruch*.

Das Gleichungssystem ist **inkonsistent**.

Beispiel 3

Suche die Lösung des Gleichungssystems:

$$\begin{aligned}2x_1 + 8x_2 + 10x_3 + 10x_4 &= 0 \\x_1 + 5x_2 + 2x_3 + 9x_4 &= 1 \\-3x_1 - 10x_2 - 21x_3 - 6x_4 &= -4\end{aligned}$$

$$\begin{array}{cccc|c}2 & 8 & 10 & 10 & 0 \\1 & 5 & 2 & 9 & 1 \\-3 & -10 & -21 & -6 & -4\end{array}$$

$$Z2 \leftarrow 2 \times Z2 - Z1, \quad Z3 \leftarrow 2 \times Z3 + 3 \times Z1$$

$$\begin{array}{cccc|c}2 & 8 & 10 & 10 & 0 \\0 & 2 & -6 & 8 & 2 \\0 & 4 & -12 & 18 & -8\end{array}$$

Beispiel 3

$$Z3 \leftarrow Z3 - 2 \times Z2$$
$$\begin{array}{cccc|c} 2 & 8 & 10 & 10 & 0 \\ 0 & 2 & -6 & 8 & 2 \\ 0 & 0 & 0 & 2 & -12 \end{array}$$

Dieses Gleichungssystem hat **unendlich** viele Lösungen. Das können wir daran erkennen, dass **nach** Erreichen der Stufenform mehr Variablen als Gleichungen übrig bleiben.

Beispiel 3

Aus der dritten Zeile erhalten wir direkt:

$$2 \cdot x_4 = -12 \quad \Rightarrow \quad x_4 = -6$$

Durch Rücksubstitution erhalten wir

$$2 \cdot x_2 - 6 \cdot x_3 + 8 \cdot (-6) = 2$$

Wir setzen x_3 gleich einer *Pseudolösung* $\alpha \in \mathbb{R}$, $x_3 = \alpha$, und erhalten

$$x_2 - 3 \cdot \alpha + 4 \cdot (-6) = 1 \quad \Rightarrow \quad x_2 = 25 + 3\alpha$$

$$2 \cdot x_1 + 8 \cdot (25 + 3 \cdot \alpha) + 10 \cdot \alpha + 10 \cdot (-6) = 0$$

$$\Rightarrow \quad x_1 = -70 - 17 \cdot \alpha$$

Beispiel 3

Jede Belegung der Pseudolösung α liefert eine gültige Lösung:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -70 - 17 \cdot \alpha \\ 25 + 3\alpha \\ \alpha \\ -6 \end{pmatrix} = \begin{pmatrix} -70 \\ 25 \\ 0 \\ -6 \end{pmatrix} + \alpha \begin{pmatrix} -17 \\ 3 \\ 1 \\ 0 \end{pmatrix}, \quad \alpha \in \mathbb{R}$$

Äquivalente Lösungen

Wir hätten in Beispiel 3 genauso $x_2 = \alpha'$ setzen können, und daraus das x_3 ausgerechnet:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -\frac{215}{3} \\ 0 \\ -\frac{25}{3} \\ -6 \end{pmatrix} + \alpha' \begin{pmatrix} -\frac{17}{3} \\ 1 \\ \frac{1}{3} \\ 0 \end{pmatrix}, \quad \alpha' \in \mathbb{R}$$

Die beiden Lösungsmengen sind aber gleich!

Es handelt sich dabei nur zwei verschiedene – aber äquivalente – Parameterdarstellungen derselben Gerade.

Die Lösungsmenge ist immer eindeutig bestimmt, die Darstellung der Lösung hingegen nicht.

Das Gauß-Jordansche Verfahren

Berechnung der inversen Matrix:

- (1) Stelle die erweiterte Matrix auf, die *links* die zu invertierende Matrix und *rechts* die (entsprechend dimensionierte) *Einheitsmatrix* enthält.
- (2) Formen die erweiterte Matrix mit den *Umformungsschritten* des Gaußschen Eliminationsverfahrens so um, dass die linke Seite zur Einheitsmatrix wird.
- (3) Entweder ist das Verfahren erfolgreich, dann erhalten wir auf der rechten Seite die *inverse Matrix*.
- (4) Oder das Verfahren bricht ab (wir erhalten auf der linken Seite eine Zeile aus Nullen). Dann ist die Matrix *singulär*.

Beispiel 1

Wir suchen die inverse Matrix zu

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 6 \\ 1 & 1 & 3 \\ -3 & -2 & -5 \end{pmatrix}$$

(1) Stelle die erweiterte Matrix auf:

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

Beispiel 1

(2) Umformen:

$$Z1 \leftarrow \frac{1}{3} \times Z1, \quad Z2 \leftarrow 3 \times Z2 - Z1, \quad Z3 \leftarrow Z3 + Z1$$

$$\left(\begin{array}{ccc|ccc} 1 & \frac{2}{3} & 2 & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 3 & -1 & 3 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{array} \right)$$

$$Z1 \leftarrow Z1 - \frac{2}{3} \times Z2$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & 0 \\ 0 & 1 & 3 & -1 & 3 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{array} \right)$$

Beispiel 1

$$Z2 \leftarrow Z2 - 3 \times Z3$$

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

(3) Die Matrix ist daher invertierbar und ihre Inverse lautet

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

Beispiel 2

Wir suchen die inverse Matrix zu

$$\mathbf{A} = \begin{pmatrix} 3 & 1 & 3 \\ 2 & 4 & 1 \\ 5 & 5 & 4 \end{pmatrix}$$

(1) Stelle die erweiterte Matrix auf:

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

Beispiel 2

(2) Umformen:

$$Z1 \leftarrow \frac{1}{3} \times Z1, \quad Z2 \leftarrow 3 \times Z2 - 2 \times Z1, \quad Z3 \leftarrow 3 \times Z3 - 5 \times Z1$$

$$\left(\begin{array}{ccc|ccc} 1 & \frac{1}{3} & 1 & \frac{1}{3} & 0 & 0 \\ 0 & 10 & -3 & -2 & 3 & 0 \\ 0 & 10 & -3 & -5 & 0 & 5 \end{array} \right)$$

$$Z1 \leftarrow Z1 - \frac{1}{30} \times Z2, \quad Z2 \leftarrow \frac{1}{10} \times Z2, \quad Z3 \leftarrow Z3 - Z2$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & \frac{11}{10} & \frac{4}{10} & -\frac{1}{10} & 0 \\ 0 & 1 & -\frac{3}{10} & -\frac{2}{10} & \frac{3}{10} & 0 \\ 0 & 0 & 0 & -3 & -3 & 5 \end{array} \right)$$

(4) Die Matrix **A** ist nicht invertierbar.

Leontief Modell

A ... Technologiematrix **p** ... Güterpreise
x ... Produktionsvektor **w** ... Arbeitslöhne
b ... Nachfragevektor

Kosten der Produktion müssen durch Preise gedeckt sein:

$$p_j = \sum_{i=1}^n a_{ij} p_i + w_j = a_{1j} p_1 + a_{2j} p_2 + \cdots + a_{nj} p_n + w_j$$

$$\mathbf{p} = \mathbf{A}^t \mathbf{p} + \mathbf{w}$$

Bei fixen Löhnen muss daher gelten:

$$\mathbf{p} = (\mathbf{I} - \mathbf{A}^t)^{-1} \mathbf{w}$$

Für das Input-Output Modell gilt weiters:

$$\mathbf{x} = \mathbf{A} \mathbf{x} + \mathbf{b}$$

Leontief Modell

Nachfrage gegeben durch Löhne für die produzierten Gütermengen:

$$\text{Nachfrage} = w_1x_1 + w_2x_2 + \cdots + w_nx_n = \mathbf{w}^t\mathbf{x}$$

Angebot gegeben durch Preise der nachgefragten Gütermenge:

$$\text{Angebot} = p_1b_1 + p_2b_2 + \cdots + p_nb_n = \mathbf{p}^t\mathbf{b}$$

Falls in einem Input-Output Modell die Gleichungen

$$\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{b} \quad \text{und} \quad \mathbf{p} = \mathbf{A}^t\mathbf{p} + \mathbf{w}$$

gelten, dann herrscht Markgleichgewicht, d.h. $\mathbf{w}^t\mathbf{x} = \mathbf{p}^t\mathbf{b}$.

Beweis:

$$\mathbf{w}^t\mathbf{x} = (\mathbf{p}^t - \mathbf{p}^t\mathbf{A})\mathbf{x} = \mathbf{p}^t(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{p}^t(\mathbf{x} - \mathbf{A}\mathbf{x}) = \mathbf{p}^t\mathbf{b}$$

Zusammenfassung

- ▶ Matrix und Vektor
- ▶ Dreiecks- und Diagonalmatrizen
- ▶ Nullmatrix und Einheitsmatrix
- ▶ Transponierte Matrix
- ▶ Inverse Matrix
- ▶ Matrizenrechnung (Matrixalgebra)
- ▶ Matrixgleichung
- ▶ Norm und inneres Produkt von Vektoren
- ▶ Lineare Gleichungssysteme
- ▶ Gaußsches Eliminationsverfahren
- ▶ Gauß-Jordansches Verfahren

Kapitel 3

Vektorräume

Reeller Vektorraum

Die Menge aller Vektoren x mit n Komponenten bezeichnen wir mit

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_i \in \mathbb{R}, 1 \leq i \leq n \right\}$$

und wird als **n -dimensionaler (reeller) Vektorraum** bezeichnet.

Definition

Ein **Vektorraum** \mathcal{V} ist eine Menge, deren Elemente sich addieren und mit einer Zahl multiplizieren lassen, wobei Summen und Vielfache von Elementen wieder Elemente der Menge sind. Die Elemente so eines Vektorraumes heißen **Vektoren**.

Teilraum

Ein **Unterraum** (oder **Teilraum**) eines Vektorraums ist eine Teilmenge, die selbst wieder einen Vektorraum bildet.

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} : x_i \in \mathbb{R}, 1 \leq i \leq 3 \right\} \subset \mathbb{R}^3 \text{ ist ein Teilraum des } \mathbb{R}^3.$$

$$\left\{ \mathbf{x} = \alpha \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} : \alpha \in \mathbb{R} \right\} \subset \mathbb{R}^3 \text{ ist ein Teilraum des } \mathbb{R}^3.$$

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_i \geq 0, 1 \leq i \leq 3 \right\} \subset \mathbb{R}^3 \text{ ist **kein** Teilraum des } \mathbb{R}^3.$$

Homogenes lineares Gleichungssystem

Sei \mathbf{A} eine $m \times n$ -Matrix.

Die Lösungsmenge \mathcal{L} des *homogenen* linearen Gleichungssystems

$$\mathbf{Ax} = 0$$

bildet einen Teilraum des \mathbb{R}^n :

Seien $\mathbf{x}, \mathbf{y} \in \mathcal{L} \subseteq \mathbb{R}^n$, i.e., $\mathbf{Ax} = 0$ und $\mathbf{Ay} = 0$.

Dann ist auch die Summe $\mathbf{x} + \mathbf{y} \in \mathcal{L}$,

$$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{Ax} + \mathbf{Ay} = 0 + 0 = 0$$

und jedes Vielfache von \mathbf{x} liegt in \mathcal{L} ,

$$\mathbf{A}(\alpha\mathbf{x}) = \alpha\mathbf{Ax} = \alpha 0 = 0$$

Linearkombination

Seien $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ Vektoren und $c_1, \dots, c_k \in \mathbb{R}$ beliebige Zahlen.
Dann erhalten wir durch **Linearkombination** einen neuen Vektor:

$$\mathbf{x} = c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k = \sum_{i=1}^k c_i \mathbf{v}_i$$

Seien $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} -2 \\ -2 \\ -2 \end{pmatrix}$, $\mathbf{v}_4 = \begin{pmatrix} -1 \\ 0 \\ -3 \end{pmatrix}$. Dann sind

$$\mathbf{x} = 1 \mathbf{v}_1 + 0 \mathbf{v}_2 + 3 \mathbf{v}_3 - 2 \mathbf{v}_4 = (-3, -4, 3)^t,$$

$$\mathbf{y} = -\mathbf{v}_1 + \mathbf{v}_2 - 2 \mathbf{v}_3 + 3 \mathbf{v}_4 = (4, 7, -2)^t, \quad \text{und}$$

$$\mathbf{z} = 2 \mathbf{v}_1 - 2 \mathbf{v}_2 - 3 \mathbf{v}_3 + 0 \mathbf{v}_4 = (0, 0, 0)^t = \mathbf{0}$$

Linearkombinationen der Vektoren \mathbf{v}_1 , \mathbf{v}_2 , \mathbf{v}_3 und \mathbf{v}_4 .

Aufgespannter Unterraum

Die Menge aller *Linearkombinationen* der Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$

$$\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = \{c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k : c_i \in \mathbb{R}\}$$

heißt der von $\mathbf{v}_1, \dots, \mathbf{v}_k$ **aufgespannte Unterraum** des \mathbb{R}^n .

$$\text{Seien } \mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} -2 \\ -2 \\ -2 \end{pmatrix}, \mathbf{v}_4 = \begin{pmatrix} -1 \\ 0 \\ -3 \end{pmatrix}.$$

$\text{span}(\mathbf{v}_1) = \{c \mathbf{v}_1 : c \in \mathbb{R}\}$ ist eine Gerade durch den Ursprung im \mathbb{R}^3 .

$\text{span}(\mathbf{v}_1, \mathbf{v}_2)$ ist Ebene durch den Ursprung im \mathbb{R}^3 .

$\text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$

$\text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4) = \mathbb{R}^3$.

Lineare Unabhängigkeit

Ein Vektor $\mathbf{x} \in \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ lässt sich immer als Linearkombination von $\mathbf{v}_1, \dots, \mathbf{v}_k$ darstellen.

$$\text{Seien } \mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} -2 \\ -2 \\ -2 \end{pmatrix}, \mathbf{v}_4 = \begin{pmatrix} -1 \\ 0 \\ -3 \end{pmatrix}.$$

$$\mathbf{x} = \begin{pmatrix} -3 \\ -4 \\ 3 \end{pmatrix} = 1 \mathbf{v}_1 + 0 \mathbf{v}_2 + 3 \mathbf{v}_3 - 2 \mathbf{v}_4 = -1 \mathbf{v}_1 + 2 \mathbf{v}_2 + 6 \mathbf{v}_3 - 2 \mathbf{v}_4$$

Diese Darstellung ist aber nicht immer eindeutig!

Grund: $2 \mathbf{v}_1 - 2 \mathbf{v}_2 - 3 \mathbf{v}_3 + 0 \mathbf{v}_4 = \mathbf{0}$

Lineare Unabhängigkeit

Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ heißen **linear unabhängig** falls das Gleichungssystem

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k = 0$$

nur die Lösung $c_1 = c_2 = \dots = c_k = 0$ besitzt. Sie heißen **linear abhängig**, wenn das Gleichungssystem andere Lösungen besitzt.

Sind Vektoren linear abhängig, dann lässt sich *ein* Vektor (aber nicht notwendigerweise *jeder*!) als Linearkombination der anderen Vektoren darstellen.

$$2 \mathbf{v}_1 - 2 \mathbf{v}_2 - 3 \mathbf{v}_3 + 0 \mathbf{v}_4 = 0 \quad \Leftrightarrow \quad \mathbf{v}_1 = \mathbf{v}_2 + \frac{3}{2} \mathbf{v}_3$$

Daher ist $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \text{span}(\mathbf{v}_2, \mathbf{v}_3)$.

Lineare Unabhängigkeit

Bestimmung der linearen Unabhängigkeit

- (1) Fasse die Vektoren als Spaltenvektoren einer Matrix \mathbf{V} auf.
- (2) Bringe Matrix \mathbf{V} mit den Umformungsschritten des Gaußschen Eliminationsverfahrens in die Stufenform.
- (3) Zähle die Zeilen, die ungleich dem Nullvektor sind.
- (4) Ist diese Anzahl gleich der Anzahl der Vektoren, so sind diese Vektoren linear unabhängig.
Ist sie kleiner, so sind die Vektoren linear abhängig.

In diesem Verfahren wird festgestellt ob das lineare Gleichungssystem $\mathbf{V} \cdot \mathbf{c} = 0$ eindeutig lösbar ist.

Beispiel – linear unabhängig

Sind die Vektoren

$$\mathbf{v}_1 = \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}$$

linear unabhängig?

(1) Wir bringen diese drei Vektoren in Matrixform:

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

Beispiel – linear unabhängig

(2) Durch Umformung erhalten wir

$$\begin{pmatrix} 3 & 1 & 3 \\ 2 & 4 & 1 \\ 2 & 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 3 & 1 & 3 \\ 0 & 10 & -3 \\ 0 & 1 & -3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 3 & 1 & 3 \\ 0 & 10 & -3 \\ 0 & 0 & -27 \end{pmatrix}$$

(3) Es gibt 3 von Null verschiedene Zeilen.

(4) Diese Anzahl stimmt mit der Anzahl der Vektoren (= 3) überein.
Die drei Vektoren \mathbf{v}_1 , \mathbf{v}_2 und \mathbf{v}_3 sind daher linear unabhängig.

Beispiel – linear abhängig

Sind die Vektoren $\mathbf{v}_1 = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}$

linear unabhängig?

(1) Wir bringen diese Vektoren in Matrixform ...

(2) und formen um:

$$\begin{pmatrix} 3 & 1 & 3 \\ 2 & 4 & 1 \\ 5 & 5 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 3 & 1 & 3 \\ 0 & 10 & -3 \\ 0 & 10 & -3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 3 & 1 & 3 \\ 0 & 10 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

(3) Es gibt 2 von Null verschiedene Zeilen.

(4) Diese Anzahl ist kleiner als die Anzahl der Vektoren (= 3).

Die drei Vektoren \mathbf{v}_1 , \mathbf{v}_2 und \mathbf{v}_3 sind daher linear abhängig.

Rang einer Matrix

Der **Rang** $\text{rank}(\mathbf{A})$ einer Matrix \mathbf{A} ist die maximale Anzahl an linear unabhängigen Spalten.

Es gilt:

$$\text{rank}(\mathbf{A}^t) = \text{rank}(\mathbf{A})$$

Der Rang einer $n \times k$ -Matrix ist immer $\leq \min(n, k)$.

Eine $n \times n$ -Matrix heißt **regulär**, falls sie **vollen Rang** hat, d.h. falls $\text{rank}(\mathbf{A}) = n$.

Rang einer Matrix

Berechnung des Ranges:

- (1) Bringen die Matrix mit den Umformungsschritten des Gaußschen Eliminationsverfahrens in die Stufenform.
- (2) Der Rang der Matrix ergibt sich dann aus der Anzahl der von Null verschiedenen Zeilen.

$$\begin{pmatrix} 3 & 1 & 3 \\ 2 & 4 & 1 \\ 2 & 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 3 & 1 & 3 \\ 0 & 10 & -3 \\ 0 & 0 & -27 \end{pmatrix} \Rightarrow \text{rank}(\mathbf{A}) = 3$$

$$\begin{pmatrix} 3 & 1 & 3 \\ 2 & 4 & 1 \\ 5 & 5 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 3 & 1 & 3 \\ 0 & 10 & -3 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \text{rank}(\mathbf{A}) = 2$$

Invertierbar und regulär

Eine $n \times n$ -Matrix A ist genau dann *invertierbar*, wenn sie *regulär* ist, also *vollen Rang* hat.

Die 3×3 -Matrix $\begin{pmatrix} 3 & 1 & 3 \\ 2 & 4 & 1 \\ 2 & 1 & 1 \end{pmatrix}$ hat vollen Rang (3).

Sie ist daher regulär und damit invertierbar.

Die 3×3 -Matrix $\begin{pmatrix} 3 & 1 & 3 \\ 2 & 4 & 1 \\ 5 & 5 & 4 \end{pmatrix}$ hat nur Rang 2.

Sie ist daher nicht regulär und damit singulär (i.e., nicht invertierbar).

Basis

Eine Menge von Vektoren $\{\mathbf{v}_1, \dots, \mathbf{v}_d\}$ **erzeugt** einen Vektorraum \mathcal{V} , falls

$$\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_d) = \mathcal{V}$$

Diese Vektoren heißen ein **Erzeugendensystem** für den Vektorraum.

Sind diese Vektoren *linear unabhängig*, so heißt diese Menge eine **Basis** des Vektorraumes.

Die Basis eines Vektorraumes ist nicht eindeutig bestimmt!

Die Anzahl an Vektoren in einer Basis ist hingegen eindeutig bestimmt und heißt die **Dimension** des Vektorraumes.

$$\dim(\mathcal{V}) = d$$

Beispiel – Basis

Die **kanonische Basis** des \mathbb{R}^n besteht aus den n Einheitsvektoren:

$$\mathcal{B}_0 = \{\mathbf{e}_1, \dots, \mathbf{e}_n\} \subset \mathbb{R}^n$$

Andere Basis des \mathbb{R}^3 :

$$\left\{ \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \right\}$$

Keine Basen des \mathbb{R}^3 sind (linear abhängig bzw. $\text{span}(\mathbf{v}_1, \mathbf{v}_2) \neq \mathbb{R}^3$)

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} -2 \\ -2 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -3 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix} \right\}$$

Koordinaten eines Vektors

Die Koordinaten \mathbf{c} eines Vektors \mathbf{x} bezüglich einer Basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ erhalten wir durch Lösen des Gleichungssystems

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n = \mathbf{x}$$

bzw. in Matrixschreibweise mit $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$:

$$\mathbf{V} \cdot \mathbf{c} = \mathbf{x} \quad \Rightarrow$$

$$\mathbf{c} = \mathbf{V}^{-1} \mathbf{x}$$

\mathbf{V} hat per Konstruktion vollen Rang.

Genau genommen sind x_1, \dots, x_n nur die Koordinaten des Vektors \mathbf{x} bezüglich der kanonischen Basis.

Jeder n -dimensionale Vektorraum \mathcal{V} ist daher *isomorph* (d.h., sieht so aus wie) der \mathbb{R}^n .

Beispiel

Wir suchen die Koordinaten \mathbf{c} von $\mathbf{x} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$

bezüglich der Basis $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 6 \end{pmatrix} \right\}$

Wir lösen das Gleichungssystem $\mathbf{V}\mathbf{c} = \mathbf{x}$:

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 3 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 3 & 3 & -1 \\ 3 & 5 & 6 & 2 \end{array} \right)$$

Beispiel

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

Durch Rücksubstitution erhalten wir $c_1 = 4$, $c_2 = -8$ und $c_3 = 5$.

Der Koordinatenvektor von \mathbf{x} bezüglich der Basis \mathcal{B} lautet daher

$$\mathbf{c} = \begin{pmatrix} 4 \\ -8 \\ 5 \end{pmatrix}$$

Alternative könnten wir auch \mathbf{V}^{-1} berechnen und erhalten $\mathbf{c} = \mathbf{V}^{-1}\mathbf{x}$.

Basiswechsel

Seien \mathbf{c}_1 und \mathbf{c}_2 die Koordinatenvektoren eines Vektors \mathbf{x} bezüglich der Basis $\mathcal{B}_1 = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ bzw. $\mathcal{B}_2 = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\}$.

Es gilt daher $\mathbf{c}_2 = \mathbf{W}^{-1}\mathbf{x} = \mathbf{W}^{-1}\mathbf{V}\mathbf{c}_1$.

Dieses „Umrechnen“ wird als **Basiswechsel** oder **Basistransformation** bezeichnet.

Die Matrix

$$\mathbf{U} = \mathbf{W}^{-1}\mathbf{V}$$

heißt **Transformationsmatrix** zum Basiswechsel von \mathcal{B}_2 nach \mathcal{B}_1 .
(Man beachte die Umkehrung der Reihenfolge, da $\mathbf{V} = \mathbf{W}\mathbf{U}$.)

Beispiel – Basiswechsel

Seien

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \\ 6 \end{pmatrix} \right\} \text{ und } \mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 6 \end{pmatrix} \right\}$$

zwei Basen des \mathbb{R}^3 .

Transformationsmatrix für den Basiswechsel von \mathcal{B}_2 nach \mathcal{B}_1 :

$$\mathbf{U} = \mathbf{W}^{-1} \cdot \mathbf{V}.$$

$$\mathbf{W} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 3 & 5 & 6 \end{pmatrix} \quad \Rightarrow \quad \mathbf{W}^{-1} = \begin{pmatrix} 3 & -1 & 0 \\ -3 & 3 & -1 \\ 1 & -2 & 1 \end{pmatrix}$$

$$\mathbf{V} = \begin{pmatrix} 1 & -2 & 3 \\ 1 & 1 & 5 \\ 1 & 1 & 6 \end{pmatrix}$$

Beispiel – Basiswechsel

Transformationsmatrix für den Basiswechsel von \mathcal{B}_2 nach \mathcal{B}_1 :

$$\mathbf{U} = \mathbf{W}^{-1} \cdot \mathbf{V} = \begin{pmatrix} 3 & -1 & 0 \\ -3 & 3 & -1 \\ 1 & -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 & 3 \\ 1 & 1 & 5 \\ 1 & 1 & 6 \end{pmatrix} = \begin{pmatrix} 2 & -7 & 4 \\ -1 & 8 & 0 \\ 0 & -3 & -1 \end{pmatrix}$$

Sei $\mathbf{c}_1 = (3, 2, 1)^t$ der Koordinatenvektor von \mathbf{x} bezüglich Basis \mathcal{B}_1 .

Dann lautet der Koordinatenvektor \mathbf{c}_2 bezüglich Basis \mathcal{B}_2

$$\mathbf{c}_2 = \mathbf{U}\mathbf{c}_1 = \begin{pmatrix} 2 & -7 & 4 \\ -1 & 8 & 0 \\ 0 & -3 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -4 \\ 13 \\ -7 \end{pmatrix}$$

Lineare Abbildung

Eine Abbildung φ zwischen Vektorräumen \mathcal{V} und \mathcal{W}

$$\varphi: \mathcal{V} \rightarrow \mathcal{W}, \mathbf{x} \mapsto \mathbf{y} = \varphi(\mathbf{x})$$

heißt **linear**, falls für alle $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ und $\alpha \in \mathbb{R}$ gilt

(i) $\varphi(\mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}) + \varphi(\mathbf{y})$

(ii) $\varphi(\alpha \mathbf{x}) = \alpha \varphi(\mathbf{x})$

Lineare Abbildung

Sei \mathbf{A} eine $m \times n$ -Matrix. Dann ist die Abbildung

$\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{x} \mapsto \varphi_{\mathbf{A}}(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x}$ linear:

$$\varphi_{\mathbf{A}}(\mathbf{x} + \mathbf{y}) = \mathbf{A} \cdot (\mathbf{x} + \mathbf{y}) = \mathbf{A} \cdot \mathbf{x} + \mathbf{A} \cdot \mathbf{y} = \varphi_{\mathbf{A}}(\mathbf{x}) + \varphi_{\mathbf{A}}(\mathbf{y})$$

$$\varphi_{\mathbf{A}}(\alpha \mathbf{x}) = \mathbf{A} \cdot (\alpha \mathbf{x}) = \alpha (\mathbf{A} \cdot \mathbf{x}) = \alpha \varphi_{\mathbf{A}}(\mathbf{x})$$

Umgekehrt können wir jede lineare Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ durch eine geeignete $m \times n$ -Matrix \mathbf{A} darstellen: $\varphi(\mathbf{x}) = \mathbf{A}_{\varphi} \mathbf{x}$.

Matrizen beschreiben somit alle denkbaren linearen Abbildungen zwischen Vektorräumen.

Lineare Abbildungen sind so einfach, dass man noch viel darüber aussagen und ausrechnen kann.

Geometrische Interpretation linearer Abbildungen

Man kann folgende „elementare“ Abbildungen unterscheiden:

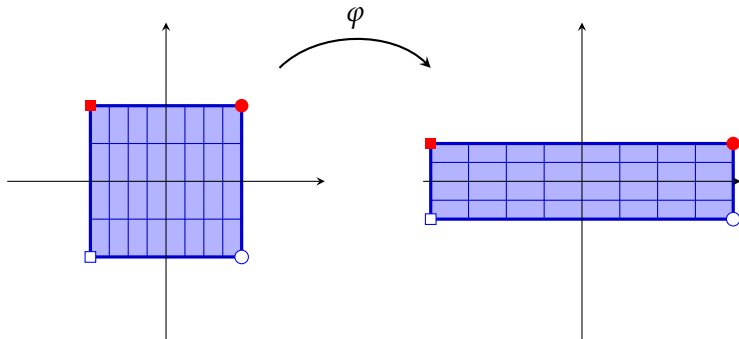
- ▶ *Streckung / Stauchung* in eine Richtung
- ▶ *Projektion* in einen Unterraum
- ▶ *Drehung*
- ▶ *Spiegelung* an einem Unterraum

Diese einfachen Abbildungen können zu komplexeren zusammengesetzt werden, z.B., Streckdrehungen.

Streckung / Stauchung

Die Abbildung $\varphi: \mathbf{x} \mapsto \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \mathbf{x}$

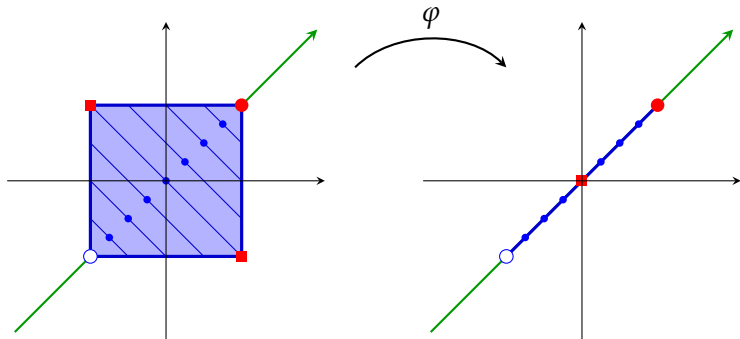
streckt die x -Koordinate um den Faktor 2 und
staucht die y -Koordinate um den Faktor $\frac{1}{2}$.



Projektion

Die Abbildung $\varphi: \mathbf{x} \mapsto \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \mathbf{x}$

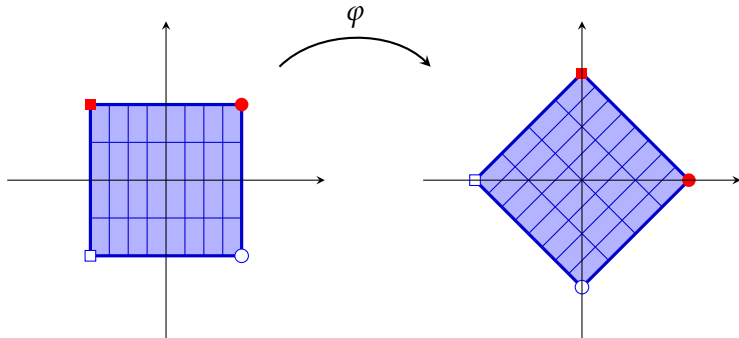
projiziert den Punkt \mathbf{x} orthogonal auf den von $(1,1)^t$ aufgespannten Unterraum.



Drehung

Die Abbildung $\varphi: \mathbf{x} \mapsto \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \mathbf{x}$

dreht den Punkt \mathbf{x} um 45° im Uhrzeigersinn um den Ursprung.



Spiegelung

Die Abbildung $\varphi: \mathbf{x} \mapsto \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x}$

spiegelt den Punkt \mathbf{x} an der y -Achse.

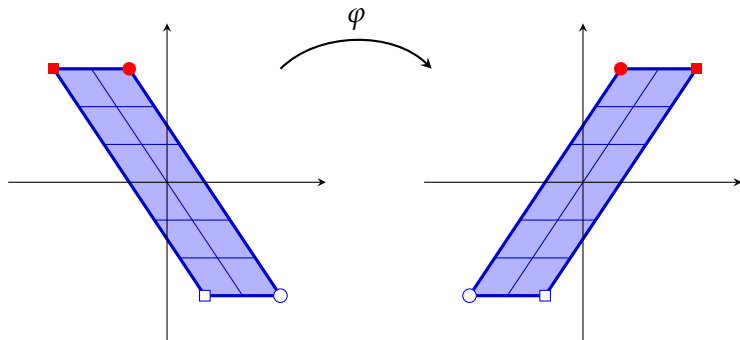


Image und Kern

Sei $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{x} \mapsto \varphi(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x}$ eine lineare Abbildung.

Das **Bild (Image)** von φ ist ein Teilraum des \mathbb{R}^m .

$$\text{Im}(\varphi) = \{\varphi(\mathbf{v}) : \mathbf{v} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$$

Der **Kern (oder Nullraum)** von φ ist ein Teilraum des \mathbb{R}^n .

$$\text{Ker}(\varphi) = \{\mathbf{v} \in \mathbb{R}^n : \varphi(\mathbf{v}) = \mathbf{0}\} \subseteq \mathbb{R}^n$$

Der Kern ist das Urbild von $\mathbf{0}$.

Der *Kern* von \mathbf{A} , $\text{Ker}(\mathbf{A})$, ist der Kern der entsprechenden linearen Abbildung.

Erzeugendensystem des Bildraumes

Sei $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ und $\mathbf{x} \in \mathbb{R}^n$ ein beliebige Vektor.

Wir können \mathbf{x} als Linearkombination der kanonischen Basis darstellen:

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

Weiters ist $\mathbf{A}\mathbf{e}_i = \mathbf{a}_i$, da für die k -te Komponente gilt:

$$(\mathbf{A}\mathbf{e}_i)_k = \sum_{j=1}^n a_{kj} (\mathbf{e}_i)_j = a_{ki}$$

Daher ist das Bild von \mathbf{x} eine Linearkombination der Spalten von \mathbf{A} :

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{A} \cdot \sum_{i=1}^n x_i \mathbf{e}_i = \sum_{i=1}^n x_i \mathbf{A}\mathbf{e}_i = \sum_{i=1}^n x_i \mathbf{a}_i$$

Die Spaltenvektoren \mathbf{a}_i spannen den Bildraum $\text{Im}(\varphi)$ auf.

Dimension von Image und Kern

Seien $\mathbf{v}_1, \mathbf{v}_2 \in \text{Ker}(\varphi)$.

Dann ist auch jede Linearkombination von $\mathbf{v}_1, \mathbf{v}_2 \in \text{Ker}(\varphi)$:

$$\varphi(\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2) = \alpha_1 \varphi(\mathbf{v}_1) + \alpha_2 \varphi(\mathbf{v}_2) = \alpha_1 \mathbf{0} + \alpha_2 \mathbf{0} = \mathbf{0}$$

Wir erhalten eine Basis von $\text{Ker}(\varphi)$ durch Lösen des linearen Gleichungssystems $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$.

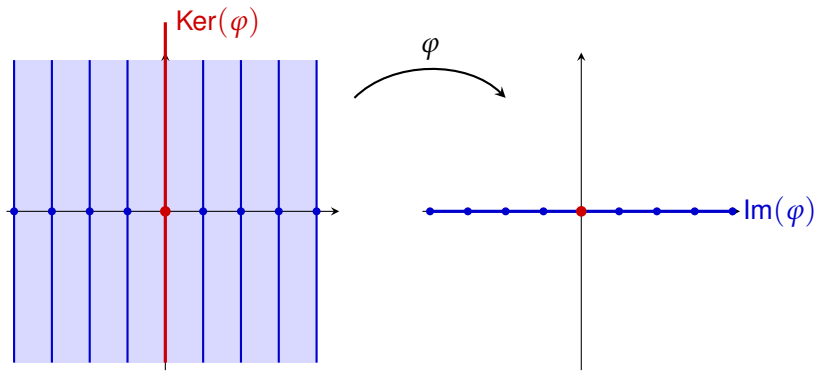
Zusammenhang zwischen diesen Vektorräumen:

$$\dim \mathcal{V} = \dim \text{Im}(\varphi) + \dim \text{Ker}(\varphi)$$

Dimension von Image und Kern

Die Abbildung $\varphi: \mathbf{x} \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{x}$

projiziert einen Punkt \mathbf{x} orthogonal auf die x -Achse.



Lineare Abbildung und Rang

Der Rang einer Matrix $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ ist (per definitionem) die Dimension von $\text{span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$.

Er gibt daher die Dimension des Bildes der korrespondierenden linearen Abbildung an.

$$\dim \text{Im}(\varphi_{\mathbf{A}}) = \text{rank}(\mathbf{A})$$

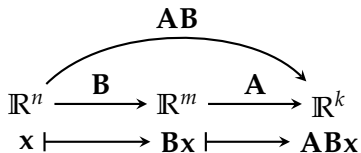
Die Dimension der Lösungsmenge \mathcal{L} eines homogenen linearen Gleichungssystems $\mathbf{A} \mathbf{x} = 0$ erhalten wir durch den Kern dieser linearen Abbildung.

$$\dim \mathcal{L} = \dim \text{Ker}(\varphi_{\mathbf{A}}) = \dim \mathbb{R}^n - \dim \text{Im}(\varphi_{\mathbf{A}}) = n - \text{rank}(\mathbf{A})$$

Matrixmultiplikation

Durch *Multiplizieren* zweier Matrizen **A** und **B** erhalten wir eine **zusammengesetzte** Abbildung:

$$(\varphi_{\mathbf{A}} \circ \varphi_{\mathbf{B}})(\mathbf{x}) = \varphi_{\mathbf{A}}(\varphi_{\mathbf{B}}(\mathbf{x})) = \mathbf{A}(\mathbf{B}\mathbf{x}) = (\mathbf{A} \cdot \mathbf{B})\mathbf{x}$$



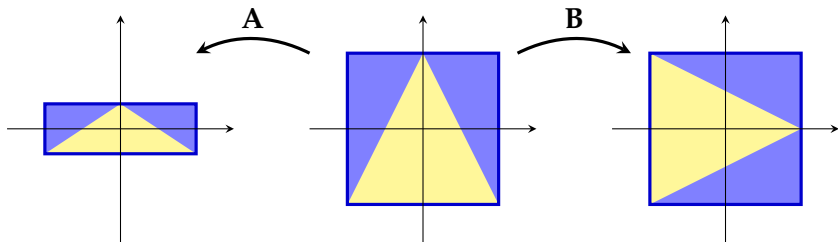
Aus dieser Sichtweise folgt:

$$\text{rank}(\mathbf{A} \cdot \mathbf{B}) \leq \min \{ \text{rank}(\mathbf{A}), \text{rank}(\mathbf{B}) \}$$

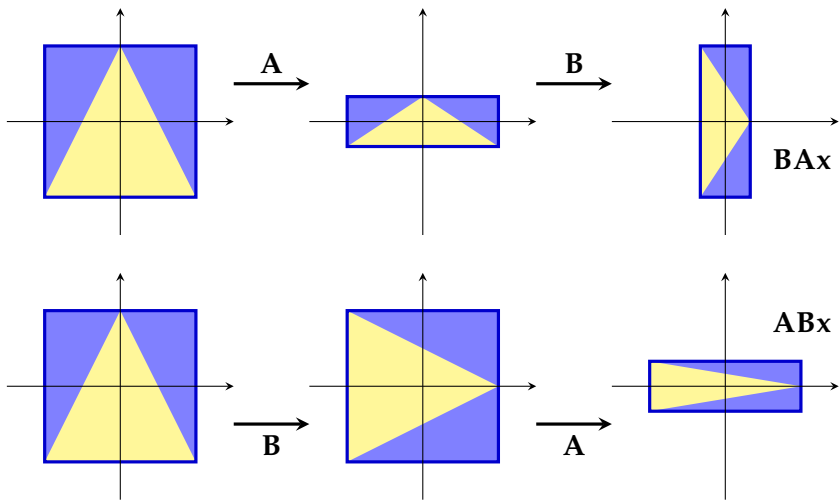
Nicht-kommutative Matrizenmultiplikation

$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}$ beschreibt eine Stauchung der y -Koordinate.

$\mathbf{B} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ beschreibt eine Drehung im Uhrzeigersinn um 90° .



Nicht-kommutative Matrizenmultiplikation



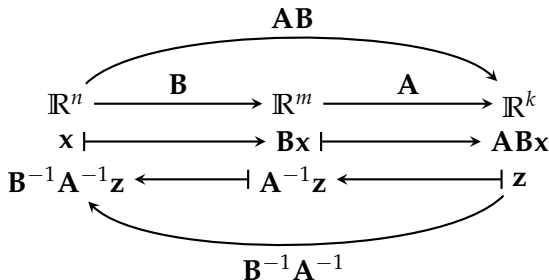
Inverse Matrix

Die *inverse Matrix* \mathbf{A}^{-1} von \mathbf{A} existiert genau dann, wenn die Abbildung $\varphi_{\mathbf{A}}(\mathbf{x}) = \mathbf{A} \mathbf{x}$ bijektiv ist, wenn also

$$\varphi_{\mathbf{A}}(\mathbf{x}) = x_1 \mathbf{a}_1 + \cdots + x_n \mathbf{a}_n = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{0}$$

d.h., wenn \mathbf{A} *regulär* ist.

Aus dieser Sichtweise wird klar, warum $(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}$



Ähnliche Matrizen

Die Basis eines Vektorraumes und damit die Koordinatendarstellung eines Vektors ist nicht eindeutig. Die Matrix \mathbf{A}_φ einer linearen Abbildung φ hängt ebenfalls von der verwendeten Basis ab.

Sei nun \mathbf{A} die Matrix bezüglich der Basis \mathcal{B}_1 .

Wie sieht nun die entsprechende Matrix \mathbf{C} bezüglich der Basis \mathcal{B}_2 aus?

$$\begin{array}{ccc} \text{Basis } \mathcal{B}_1 & \mathbf{U} \mathbf{x} \xrightarrow{\mathbf{A}} \mathbf{A} \mathbf{U} \mathbf{x} & \\ & \mathbf{U} \uparrow \qquad \qquad \downarrow \mathbf{U}^{-1} & \text{also } \mathbf{C} \mathbf{x} = \mathbf{U}^{-1} \mathbf{A} \mathbf{U} \mathbf{x} \\ \text{Basis } \mathcal{B}_2 & \mathbf{x} \xrightarrow{\mathbf{C}} \mathbf{U}^{-1} \mathbf{A} \mathbf{U} \mathbf{x} & \end{array}$$

Zwei $n \times n$ -Matrizen \mathbf{A} und \mathbf{C} heißen **ähnlich**, falls es eine invertierbare Matrix \mathbf{U} gibt, mit

$$\mathbf{C} = \mathbf{U}^{-1} \mathbf{A} \mathbf{U}$$

Zusammenfassung

- ▶ Vektorraum
- ▶ Lineare Unabhängigkeit und Rang
- ▶ Basis und Dimension
- ▶ Koordinatenvektor
- ▶ Basiswechsel
- ▶ Lineare Abbildungen

Kapitel 4

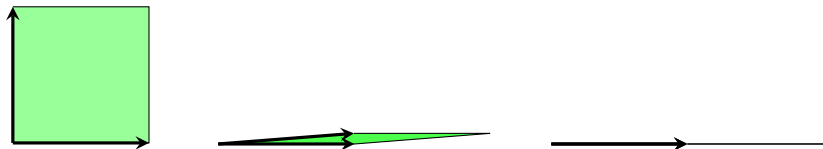
Determinante

Was ist eine Determinante?

Wir wollen „messen“, ob n Vektoren im \mathbb{R}^n linear abhängig sind bzw. wie weit sie davon entfernt sind.

Die Idee:

Zwei Vektoren im \mathbb{R}^2 spannen ein Parallelogramm auf.



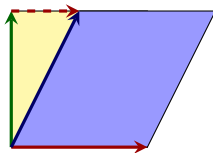
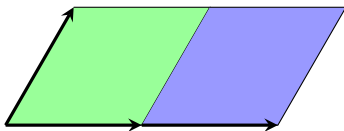
Vektoren sind linear *abhängig* \Leftrightarrow Flächeninhalt ist Null

Wir verwenden das n -dimensionale Volumen für unsere Funktion zum „Messen“ der linearen Abhängigkeit.

Eigenschaften des Volumens

Definieren diese Funktion indirekt durch Eigenschaften des Volumens.

- ▶ Multiplizieren wir einen Vektor mit einer Zahl α erhalten wir das α -fache Volumen.
- ▶ Addieren wir zu einem Vektor das Vielfache eines anderen Vektors, so bleibt das Volumen konstant.
- ▶ Sind zwei Vektoren gleich, so ist das Volumen gleich Null.
- ▶ Das Volumen eines Würfels mit Seitenlänge eins ist gleich eins.



Determinante

Die **Determinante** ist eine Funktion, die jeder $n \times n$ -Matrix $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ eine reelle Zahl $\det(\mathbf{A})$ zuordnet und folgende Eigenschaften besitzt:

(1) Die Determinante ist **linear** in jeder Spalte:

$$\det(\dots, \mathbf{a}_i + \mathbf{b}_i, \dots) = \det(\dots, \mathbf{a}_i, \dots) + \det(\dots, \mathbf{b}_i, \dots)$$

$$\det(\dots, \alpha \mathbf{a}_i, \dots) = \alpha \det(\dots, \mathbf{a}_i, \dots)$$

(2) Die Determinante ist Null, falls zwei Spaltenvektoren gleich sind:

$$\det(\dots, \mathbf{a}_i, \dots, \mathbf{a}_i, \dots) = 0$$

(3) Die Determinante ist **normiert**:

$$\det(\mathbf{I}) = 1$$

Schreibweisen:

$$\det(\mathbf{A}) = |\mathbf{A}|$$

Beispiel

(1)

$$\begin{vmatrix} 1 & 2 + 10 & 3 \\ 4 & 5 + 11 & 6 \\ 7 & 8 + 12 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} + \begin{vmatrix} 1 & 10 & 3 \\ 4 & 11 & 6 \\ 7 & 12 & 9 \end{vmatrix}$$

$$\begin{vmatrix} 1 & 3 \cdot 2 & 3 \\ 4 & 3 \cdot 5 & 6 \\ 7 & 3 \cdot 8 & 9 \end{vmatrix} = 3 \cdot \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$$

(2)

$$\begin{vmatrix} 1 & 2 & 1 \\ 4 & 5 & 4 \\ 7 & 8 & 7 \end{vmatrix} = 0$$

Weitere Eigenschaften

Die Eigenschaften (1) – (3) definieren eindeutig eine Funktion.

Diese Funktion besitzt noch eine Reihe weiterer Eigenschaften, die sich aus diesen drei Grundeigenschaften herleiten lassen.

(4) Die Determinante ist **alternierend**:

$$\det(\dots, \mathbf{a}_i, \dots, \mathbf{a}_k, \dots) = -\det(\dots, \mathbf{a}_k, \dots, \mathbf{a}_i, \dots)$$

(5) Der Wert der Determinante ändert sich nicht, wenn zu einer Spalte das Vielfache einer anderen Spalte addiert wird:

$$\det(\dots, \mathbf{a}_i + \alpha \mathbf{a}_k, \dots, \mathbf{a}_k, \dots) = \det(\dots, \mathbf{a}_i, \dots, \mathbf{a}_k, \dots)$$

(6) Beim Transponieren ändert sich der Wert der Determinante nicht, d.h. die Aussagen über Spalten gelten analog für Zeilen:

$$\det(\mathbf{A}^t) = \det(\mathbf{A})$$

Beispiel

(4)

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = - \begin{vmatrix} 1 & 3 & 2 \\ 4 & 6 & 5 \\ 7 & 9 & 8 \end{vmatrix}$$

(5)

$$\begin{vmatrix} 1 & 2 + 2 \cdot 1 & 3 \\ 4 & 5 + 2 \cdot 4 & 6 \\ 7 & 8 + 2 \cdot 7 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$$

(6)

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

Weitere Eigenschaften

- (7) $\det(\mathbf{A}) \neq 0 \Leftrightarrow$ Spalten von \mathbf{A} sind linear unabhängig
 $\Leftrightarrow \mathbf{A}$ ist regulär
 $\Leftrightarrow \mathbf{A}$ ist invertierbar

- (8) Die Determinante des Produktes zweier Matrizen ist gleich dem Produkt ihrer Determinanten:

$$\det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$$

- (9) Die Determinante der inversen Matrix ist gleich dem Kehrwert der Determinante:

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$$

Weitere Eigenschaften

- (10) Die Determinante einer Dreiecksmatrix ist das Produkt ihrer Diagonalelemente:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix} = a_{11} \cdot a_{22} \cdot a_{33} \cdot \dots \cdot a_{nn}$$

- (11) Der Absolutbetrag der Determinante $|\det(\mathbf{a}_1, \dots, \mathbf{a}_n)|$ ist das Volumen des von den Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ aufgespannten Parallelepipeds.

Regel von Sarrus

- ▶ 2×2 -Matrix:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$$

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 3 = -2$$

- ▶ 3×3 -Matrix: Regel von Sarrus

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{vmatrix} = 1 \cdot 5 \cdot 9 + 2 \cdot 6 \cdot 7 + 3 \cdot 4 \cdot 8 - 7 \cdot 5 \cdot 3 - 8 \cdot 6 \cdot 1 - 9 \cdot 4 \cdot 2 = 0$$

Umformen in Dreiecksmatrix

(1) Forme Matrix in obere Dreiecksmatrix um.

Das Verfahren ähnelt dem Gaußschen Eliminationsverfahren:

- ▶ Addiere Vielfaches einer Zeile zu einer anderen Zeile. (5)
- ▶ Multipliziere eine Zeile mit einer Konstanten $\alpha \neq 0$ und die Determinante mit dem Kehrwert $\frac{1}{\alpha}$. (1)
- ▶ Vertausche zwei Zeilen und ändere Vorzeichen der Determinante. (4)

(2) Berechne Determinante als Produkt der Diagonalelemente. (10)

Umformen in Dreiecksmatrix

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{vmatrix} \\ = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{vmatrix} = 1 \cdot (-3) \cdot 0 = 0$$

$$\begin{vmatrix} 0 & 2 & 4 \\ 1 & 2 & 3 \\ 2 & 5 & 6 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 2 & 5 & 6 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 1 & 0 \end{vmatrix} \\ = -\frac{1}{2} \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & -4 \end{vmatrix} = -\frac{1}{2} \cdot 1 \cdot 2 \cdot (-4) = 4$$

Der Laplacesche Entwicklungssatz

Entwicklung nach der k -ten Spalte bzw. i -ten Zeile:

$$\det(\mathbf{A}) = \sum_{i=1}^n a_{ik} \cdot (-1)^{i+k} |\mathbf{S}_{ik}| = \sum_{k=1}^n a_{ik} \cdot (-1)^{i+k} |\mathbf{S}_{ik}|$$

\mathbf{S}_{ik} ist die $(n-1) \times (n-1)$ -Matrix, die man erhält, wenn die i -te Zeile und k -te Spalte gestrichen wird („Streichungsmatrix“).

Die Determinante $|\mathbf{S}_{ik}|$ ist ein sogenannter **Minor** von \mathbf{A} .

Es ist dabei völlig egal, nach welcher Zeile oder Spalte entwickelt wird.

Die Vorzeichen $(-1)^{i+k}$ erhält man auch über ein Schachbrettmuster:

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$$

Entwicklung nach der ersten Zeile

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1 \cdot (-1)^{1+1} \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} \\ + 2 \cdot (-1)^{1+2} \cdot \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} \\ + 3 \cdot (-1)^{1+3} \cdot \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} \\ = 1 \cdot (-3) - 2 \cdot (-6) + 3 \cdot (-3) \\ = 0$$

Entwicklung nach der zweiten Spalte

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 2 \cdot (-1)^{1+2} \cdot \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} \\ + 5 \cdot (-1)^{2+2} \cdot \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} \\ + 8 \cdot (-1)^{3+2} \cdot \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} \\ = -2 \cdot (-6) + 5 \cdot (-12) - 8 \cdot (-6) \\ = 0$$

Adjungierte Matrix

Die Faktoren $A_{ik} = (-1)^{i+k} |\mathbf{S}_{ik}|$ aus dem Laplaceschen Entwicklungssatz heißen die **Kofaktor** von a_{ik} .

$$\det(\mathbf{A}) = \sum_{i=1}^n a_{ik} \cdot A_{ik}$$

Wir können diese zur **Kofaktorenmatrix** \mathbf{A}^* zusammenfassen.
Durch Transponieren erhalten wir die **adjungierte Matrix** \mathbf{A}^{*t} von \mathbf{A} .

$$\mathbf{A}^{*t} = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

Produkt $\mathbf{A} \cdot \mathbf{A}^{*t}$

$$\begin{aligned} \mathbf{A} \cdot \mathbf{A}^{*t} &= \begin{pmatrix} 0 & 2 & 4 \\ 1 & 2 & 3 \\ 2 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} & -\begin{vmatrix} 2 & 4 \\ 5 & 6 \end{vmatrix} & \begin{vmatrix} 2 & 4 \\ 2 & 3 \end{vmatrix} \\ -\begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} & \begin{vmatrix} 0 & 4 \\ 2 & 6 \end{vmatrix} & -\begin{vmatrix} 0 & 4 \\ 1 & 3 \end{vmatrix} \\ \begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix} & -\begin{vmatrix} 0 & 2 \\ 2 & 5 \end{vmatrix} & \begin{vmatrix} 0 & 2 \\ 1 & 2 \end{vmatrix} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 2 & 4 \\ 1 & 2 & 3 \\ 2 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} -3 & 8 & -2 \\ 0 & -8 & 4 \\ 1 & 4 & -2 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} = |\mathbf{A}| \cdot \mathbf{I} \end{aligned}$$

Produkt $\mathbf{A}^{*t} \cdot \mathbf{A}$

Produkt aus k -ter Zeile von \mathbf{A}^{*t} und j -ter Spalte von \mathbf{A} :

$$\begin{aligned} \left(\mathbf{A}^{*t} \cdot \mathbf{A}\right)_{kj} &= \sum_{i=1}^n A_{ik} \cdot a_{ij} = \sum_{i=1}^n a_{ij} \cdot (-1)^{i+k} |\mathbf{S}_{ik}| \\ \text{[Entwicklung nach } k\text{-ter Spalte]} &= \det(\mathbf{a}_1, \dots, \underbrace{\mathbf{a}_j}_{k\text{-te Spalte}, \dots, \mathbf{a}_n) \end{aligned}$$

[\mathbf{a}_i sind die einzelnen Spaltenvektoren von \mathbf{A}]

$$= \begin{cases} |\mathbf{A}| & \text{falls } j = k \\ 0 & \text{falls } j \neq k \end{cases}$$

$$\mathbf{A}^{*t} \cdot \mathbf{A} = |\mathbf{A}| \cdot \mathbf{I} \quad \Rightarrow \quad \mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \cdot \mathbf{A}^{*t}$$

Inverse Matrix

Für die Inverse einer Matrix \mathbf{A} erhalten wir

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \cdot \mathbf{A}^{*t}$$

Für eine 2×2 -Matrix \mathbf{A} gilt

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{|\mathbf{A}|} \cdot \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Beispiel

$$\begin{pmatrix} 0 & 2 & 4 \\ 1 & 2 & 3 \\ 2 & 5 & 6 \end{pmatrix}^{-1} = \frac{1}{4} \cdot \begin{pmatrix} -3 & 8 & -2 \\ 0 & -8 & 4 \\ 1 & 4 & -2 \end{pmatrix} = \begin{pmatrix} -\frac{3}{4} & 2 & -\frac{1}{2} \\ 0 & -2 & 1 \\ \frac{1}{4} & 1 & -\frac{1}{2} \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^{-1} = \frac{1}{-2} \cdot \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

Cramersche Regel

Die **Cramersche Regel** ist eine Methode zur Lösung des Gleichungssystem

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

Falls \mathbf{A} invertierbar ist (d.h., $|\mathbf{A}| \neq 0$), dann gilt

$$\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b} = \frac{1}{|\mathbf{A}|} \mathbf{A}^{*t} \cdot \mathbf{b}$$

Wir erhalten daher für x_k

$$\begin{aligned} x_k &= \frac{1}{|\mathbf{A}|} \sum_{i=1}^n A_{ik} \cdot b_i = \frac{1}{|\mathbf{A}|} \sum_{i=1}^n b_i \cdot (-1)^{i+k} |\mathbf{S}_{ik}| \\ &= \frac{1}{|\mathbf{A}|} \det(\mathbf{a}_1, \dots, \underbrace{\mathbf{b}}_{k\text{-te Spalte}}, \dots, \mathbf{a}_n) \end{aligned}$$

Cramersche Regel

Sei \mathbf{A}_k die Matrix, die wir aus \mathbf{A} erhalten, wenn wir die k -te Spalte durch den Vektor \mathbf{b} ersetzen.

Falls \mathbf{A} eine invertierbare $n \times n$ -Matrix ist, dann erhalten wir die Lösung von

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

durch

$$\mathbf{x} = \frac{1}{|\mathbf{A}|} \begin{pmatrix} |\mathbf{A}_1| \\ \vdots \\ |\mathbf{A}_n| \end{pmatrix}$$

Dieses Verfahren eignet sich nur, wenn \mathbf{A} eine reguläre quadratische Matrix ist.

Beispiel

Gesucht ist die Lösung von

$$\begin{pmatrix} 9 & 11 & 3 \\ 9 & 13 & 4 \\ 2 & 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Lösung:

$$\mathbf{x} = \frac{1}{|\mathbf{A}|} \begin{pmatrix} |\mathbf{A}_1| \\ |\mathbf{A}_2| \\ |\mathbf{A}_3| \end{pmatrix} = \begin{pmatrix} 12 \\ -22 \\ 45 \end{pmatrix}$$

$$|\mathbf{A}| = \begin{vmatrix} 9 & 11 & 3 \\ 9 & 13 & 4 \\ 2 & 3 & 1 \end{vmatrix} = 1$$

$$|\mathbf{A}_1| = \begin{vmatrix} 1 & 11 & 3 \\ 2 & 13 & 4 \\ 3 & 3 & 1 \end{vmatrix} = 12$$

$$|\mathbf{A}_2| = \begin{vmatrix} 9 & 1 & 3 \\ 9 & 2 & 4 \\ 2 & 3 & 1 \end{vmatrix} = -22$$

$$|\mathbf{A}_3| = \begin{vmatrix} 9 & 11 & 1 \\ 9 & 13 & 2 \\ 2 & 3 & 3 \end{vmatrix} = 45$$

Zusammenfassung

- ▶ Definition der Determinante
- ▶ Eigenschaften
- ▶ Zusammenhang zu Rang und Invertierbarkeit
- ▶ Volumen eines Parallelepipeds
- ▶ Berechnung der Determinante
(Regel von Sarrus, Umformen in Dreiecksmatrix)
- ▶ Laplacescher Entwicklungssatz
- ▶ Cramersche Regel

Kapitel 5

Eigenwerte

Geschlossenes Leontief-Modell

Ein Leontief-Modell für eine Volkswirtschaft heißt geschlossen, wenn der Konsum gleich der Produktion ist, d.h. wenn

$$V \cdot x = x$$

Es handelt sich dabei um einen Spezialfall eines **Eigenwertproblems**.

Eigenwert und Eigenvektor

Ein Vektor $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq 0$, heißt **Eigenvektor** einer $n \times n$ -Matrix \mathbf{A} zum **Eigenwert** λ , falls

$$\mathbf{A} \cdot \mathbf{x} = \lambda \cdot \mathbf{x}$$

Die Eigenwerte der Matrix \mathbf{A} sind alle Zahlen λ für die ein Eigenvektor existiert.

Für eine 3×3 -Diagonalmatrix gilt

$$\mathbf{A} \cdot \mathbf{e}_1 = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ 0 \\ 0 \end{pmatrix} = a_{11} \cdot \mathbf{e}_1$$

\mathbf{e}_1 ist also Eigenvektor zum Eigenwert $\lambda = a_{11}$.

Berechnung der Eigenwerte

Für eine $n \times n$ -Matrix \mathbf{A} müssen wir die Gleichung

$$\mathbf{A} \mathbf{x} = \lambda \mathbf{x} = \lambda \mathbf{I} \mathbf{x} \quad \Leftrightarrow \quad (\mathbf{A} - \lambda \mathbf{I}) \mathbf{x} = \mathbf{0}$$

lösen. Falls $(\mathbf{A} - \lambda \mathbf{I})$ invertierbar ist, dann ist

$$\mathbf{x} = (\mathbf{A} - \lambda \mathbf{I})^{-1} \mathbf{0} = \mathbf{0}$$

Dann ist aber λ kein Eigenwert.

λ ist genau dann *Eigenwert* von \mathbf{A} , wenn $(\mathbf{A} - \lambda \mathbf{I})$ *nicht invertierbar* ist, d.h. wenn

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

Charakteristisches Polynom

$\det(\mathbf{A} - \lambda\mathbf{I})$ ist ein Polynom n -ten Grades in λ und heißt das **charakteristische Polynom** der Matrix \mathbf{A} .

Die Eigenwerte sind die Nullstellen dieses Polynoms.

Bemerkung:

Es kann sein, dass (alle oder einzelne) Nullstellen des charakteristischen Polynoms komplex sind.

Man spricht dann von komplexen Eigenwerten.

Beispiel

Wir suchen die Eigenwerte von $\mathbf{A} = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}$.

Dazu bilden wir das charakteristische Polynom und berechnen deren Nullstellen.

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 1 - \lambda & -2 \\ 1 & 4 - \lambda \end{vmatrix} = \lambda^2 - 5\lambda + 6 = 0$$

Die Lösungen dieser quadratischen Gleichung sind

$$\lambda_1 = 2 \quad \text{und} \quad \lambda_2 = 3.$$

\mathbf{A} besitzt daher die Eigenwerte 2 und 3.

Berechnung der Eigenvektoren

Wir können die Eigenvektoren \mathbf{x} zum *bekanntem* Eigenwert λ durch Einsetzen in $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$ berechnen.

Eigenvektoren von $\mathbf{A} = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}$ zum Eigenwert $\lambda_1 = 2$:

$$(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} = \begin{pmatrix} -1 & -2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Lösung mittels Gauß-Elimination: $x_2 = \alpha$ und $x_1 = -2\alpha$

$$\mathbf{v}_1 = \alpha \begin{pmatrix} -2 \\ 1 \end{pmatrix} \quad \text{für ein } \alpha \in \mathbb{R} \setminus \{0\}$$

Eigenraum

Falls \mathbf{x} ein Eigenvektor zum Eigenwert λ ist, dann auch jedes Vielfache $\alpha\mathbf{x}$:

$$\mathbf{A} \cdot (\alpha\mathbf{x}) = \alpha(\mathbf{A} \cdot \mathbf{x}) = \alpha\lambda \cdot \mathbf{x} = \lambda \cdot (\alpha\mathbf{x})$$

Falls \mathbf{x} und \mathbf{y} Eigenvektoren zum gleichen Eigenwert λ sind, dann ist auch $\mathbf{x} + \mathbf{y}$ ein Eigenvektor:

$$\mathbf{A} \cdot (\mathbf{x} + \mathbf{y}) = \mathbf{A} \cdot \mathbf{x} + \mathbf{A} \cdot \mathbf{y} = \lambda \cdot \mathbf{x} + \lambda \cdot \mathbf{y} = \lambda \cdot (\mathbf{x} + \mathbf{y})$$

Die Menge aller Eigenvektoren zum Eigenwert λ (vereinigt mit 0) ist daher ein *Unterraum* des \mathbb{R}^n und wird als **Eigenraum** bezeichnet.

Man gibt daher immer nur eine *Basis des Eigenraumes* an.

Die Dimension des Eigenraumes wird auch als (geometrische) *Vielfachheit des Eigenwertes* bezeichnet.

Beispiel

$$\text{Sei } \mathbf{A} = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}.$$

$$\text{Eigenvektor zum Eigenwert } \lambda_1 = 2: \quad \mathbf{v}_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

$$\text{Eigenvektor zum Eigenwert } \lambda_2 = 3: \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Die Eigenvektoren zum Eigenwert λ_i sind dann alle nichtverschwindenden Vielfache (d.h., $\neq 0$) von \mathbf{v}_i .

Computerprogramme geben meist normierte Eigenvektoren aus:

$$\mathbf{v}_1 = \begin{pmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix} \quad \text{und} \quad \mathbf{v}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

Beispiel

Eigenwerte und Eigenvektoren von $\mathbf{A} = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 1 \\ 0 & 6 & 2 \end{pmatrix}$.

Wir bilden das charakteristische Polynom und berechnen dessen Nullstellen.

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 2 - \lambda & 0 & 1 \\ 0 & 3 - \lambda & 1 \\ 0 & 6 & 2 - \lambda \end{vmatrix} = (2 - \lambda) \cdot \lambda \cdot (\lambda - 5) = 0$$

Wir erhalten die Eigenwerte

$$\lambda_1 = 2, \lambda_2 = 0 \text{ und } \lambda_3 = 5.$$

Beispiel

Eigenvektor(en) zum Eigenwert $\lambda_3 = 5$ durch Lösen der Gleichung

$$(\mathbf{A} - \lambda_3 \mathbf{I})\mathbf{x} = \begin{pmatrix} (2-5) & 0 & 1 \\ 0 & (3-5) & 1 \\ 0 & 6 & (2-5) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

Durch Gauß-Elimination erhalten wir

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

Somit $x_3 = \alpha$, $x_2 = \frac{1}{2}\alpha$ und $x_1 = \frac{1}{3}\alpha$ für ein beliebiges $\alpha \in \mathbb{R} \setminus \{0\}$.

Eigenvektor $\mathbf{v}_3 = (2, 3, 6)^t$

Beispiel

Eigenvektor zum Eigenwert $\lambda_1 = 2$: $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

Eigenvektor zum Eigenwert $\lambda_2 = 0$: $\mathbf{v}_2 = \begin{pmatrix} -3 \\ -2 \\ 6 \end{pmatrix}$

Eigenvektor zum Eigenwert $\lambda_3 = 5$: $\mathbf{v}_3 = \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix}$

Die Eigenvektoren zum Eigenwert λ_i sind dann alle nichtverschwindenden Vielfache (d.h., $\neq 0$) von \mathbf{v}_i .

Eigenschaften von Eigenwerten

1. \mathbf{A} und \mathbf{A}^t besitzen dieselben Eigenwerte.
2. Seien \mathbf{A} und \mathbf{B} zwei $n \times n$ -Matrizen.
Dann besitzen die Matrizen $\mathbf{A} \cdot \mathbf{B}$ und $\mathbf{B} \cdot \mathbf{A}$ dieselben Eigenwerte.
3. Ist \mathbf{x} ein Eigenvektor von \mathbf{A} zum Eigenwert λ ,
dann ist \mathbf{x} ein Eigenvektor von \mathbf{A}^k zum Eigenwert λ^k .
4. Ist \mathbf{x} ein Eigenvektor einer regulären Matrix \mathbf{A} zum Eigenwert λ ,
dann ist \mathbf{x} ein Eigenvektor von \mathbf{A}^{-1} zum Eigenwert $\frac{1}{\lambda}$.
5. Die Determinante einer $n \times n$ -Matrix \mathbf{A} ist gleich dem Produkt der Eigenwerte λ_i von \mathbf{A} :

$$\det(\mathbf{A}) = \prod_{i=1}^n \lambda_i$$

6. Die Summe der Eigenwerte λ_i einer Matrix \mathbf{A} ist gleich der Summe der Diagonalelemente von \mathbf{A} (der Spur von \mathbf{A}).

Eigenwerte ähnlicher Matrizen

Sei \mathbf{U} eine Transformationsmatrix und $\mathbf{C} = \mathbf{U}^{-1} \mathbf{A} \mathbf{U}$.

Ist \mathbf{x} ein Eigenvektor von \mathbf{A} zum Eigenwert λ ,

dann ist $\mathbf{U}^{-1}\mathbf{x}$ Eigenvektor von \mathbf{C} zum gleichen Eigenwert:

$$\mathbf{C} \cdot (\mathbf{U}^{-1}\mathbf{x}) = (\mathbf{U}^{-1}\mathbf{A}\mathbf{U})\mathbf{U}^{-1}\mathbf{x} = \mathbf{U}^{-1}\mathbf{A}\mathbf{x} = \mathbf{U}^{-1}\lambda\mathbf{x} = \lambda \cdot (\mathbf{U}^{-1}\mathbf{x})$$

Ähnliche Matrizen \mathbf{A} und \mathbf{C} haben die gleichen Eigenwerte und besitzen (unter Berücksichtigung des Basiswechsels) die gleichen Eigenvektoren.

Wir wählen für unsere Abbildung eine Basis, sodass die entsprechende Matrix \mathbf{A} möglichst einfach wird.

Die einfachsten Matrizen sind *Diagonalmatrizen*.

Können wir immer eine Darstellung durch eine Diagonalmatrix finden?

Leider nicht immer.

Symmetrische Matrix

Eine $n \times n$ -Matrix \mathbf{A} heißt **symmetrisch**, falls $\mathbf{A}^t = \mathbf{A}$.

Für eine *symmetrische* Matrix \mathbf{A} gilt:

- ▶ Alle n Eigenwerte sind reell.
- ▶ Die Eigenvektoren \mathbf{v}_i zu verschiedenen Eigenwerten λ_i sind *orthogonal*, d.h. $\mathbf{v}_i^t \cdot \mathbf{v}_j = 0$ falls $i \neq j$.
- ▶ Es gibt eine **Orthonormalbasis** $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ aus Eigenvektoren für den \mathbb{R}^n , d.h. die \mathbf{v}_i sind normiert und paarweise orthogonal.

Die Transformationsmatrix $\mathbf{U} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ ist dann eine **Orthogonalmatrix**:

$$\mathbf{U}^t \cdot \mathbf{U} = \mathbf{I} \quad \Leftrightarrow \quad \mathbf{U}^{-1} = \mathbf{U}^t$$

Diagonalisieren

Für den i -ten Einheitsvektor \mathbf{e}_i gilt

$$\mathbf{U}^t \mathbf{A} \mathbf{U} \cdot \mathbf{e}_i = \mathbf{U}^t \mathbf{A} \mathbf{v}_i = \mathbf{U}^t \lambda_i \mathbf{v}_i = \lambda_i \mathbf{U}^t \mathbf{v}_i = \lambda_i \cdot \mathbf{e}_i$$

Also

$$\mathbf{U}^t \mathbf{A} \mathbf{U} = \mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Jede symmetrische Matrix wird bezüglich einer Orthonormalbasis aus Eigenvektoren zu einer Diagonalmatrix.

Deren Diagonalelemente sind die Eigenwerte von \mathbf{A} .

Man nennt diesen Vorgang **Diagonalisieren** der Matrix \mathbf{A} .

Beispiel

Wir wollen $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ diagonalisieren.

Eigenwerte sind

$$\lambda_1 = -1 \text{ und } \lambda_2 = 3$$

mit den normierten Eigenvektoren

$$\mathbf{v}_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \text{ bzw. } \mathbf{v}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

Bezüglich dieser Basis wird \mathbf{A} zur Diagonalmatrix

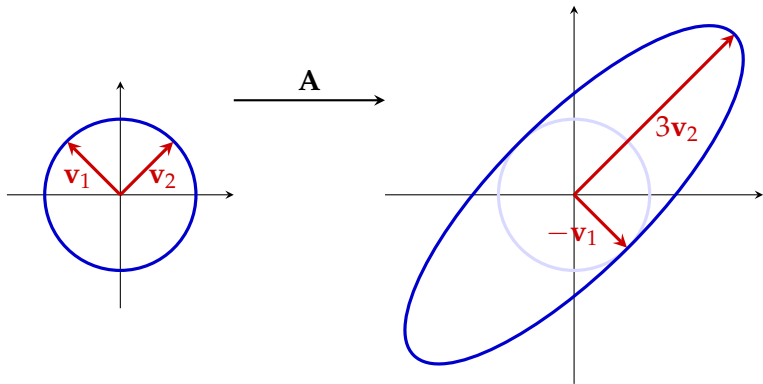
$$\begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$$

Ein geometrische Interpretation I

Die Abbildung $x \mapsto \mathbf{A}x = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} x$ bildet den

Einheitskreis im \mathbb{R}^2 in eine Ellipse ab.

Die beiden Halbachsen werden durch $\lambda_1 \mathbf{v}_1$ bzw. $\lambda_2 \mathbf{v}_2$ gebildet.



Quadratische Form

Sei A eine *symmetrische Matrix*. Dann heißt die Funktion

$$q_A : \mathbb{R}^n \rightarrow \mathbb{R}, \mathbf{x} \mapsto q_A(\mathbf{x}) = \mathbf{x}^t \cdot A \cdot \mathbf{x}$$

eine **quadratische Form**.

Sei $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. Dann ist

$$q_A(\mathbf{x}) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^t \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1^2 + 2x_2^2 + 3x_3^2$$

Beispiel

Allgemein gilt für eine $n \times n$ -Matrix $\mathbf{A} = (a_{ij})$:

$$q_{\mathbf{A}}(\mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

$$q_{\mathbf{A}}(\mathbf{x}) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^t \cdot \begin{pmatrix} 1 & 1 & -2 \\ 1 & 2 & 3 \\ -2 & 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^t \cdot \begin{pmatrix} x_1 + x_2 - 2x_3 \\ x_1 + 2x_2 + 3x_3 \\ -2x_1 + 3x_2 + x_3 \end{pmatrix}$$

$$= x_1^2 + 2x_1x_2 - 4x_1x_3 + 2x_2^2 + 6x_2x_3 + x_3^2$$

Definitheit

Eine quadratische Form $q_{\mathbf{A}}$ heißt

- ▶ **positiv definit**, wenn für alle $\mathbf{x} \neq 0$, $q_{\mathbf{A}}(\mathbf{x}) > 0$.
- ▶ **positiv semidefinit**, wenn für alle \mathbf{x} , $q_{\mathbf{A}}(\mathbf{x}) \geq 0$.
- ▶ **negativ definit**, wenn für alle $\mathbf{x} \neq 0$, $q_{\mathbf{A}}(\mathbf{x}) < 0$.
- ▶ **negativ semidefinit**, wenn für alle \mathbf{x} , $q_{\mathbf{A}}(\mathbf{x}) \leq 0$.
- ▶ **indefinit** in allen anderen Fällen.

Eine Matrix \mathbf{A} heißt *positiv* (negativ) *definit* (semidefinit), wenn die entsprechende quadratische Form *positiv* (negativ) *definit* (semidefinit) ist.

Definitheit

Jede symmetrische Matrix ist *diagonalisierbar*. Sei $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ eine Orthonormalbasis Eigenvektoren von \mathbf{A} . Dann gilt für jeden Vektor \mathbf{x} :

$$\mathbf{x} = c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n = \sum_{i=1}^n c_i \mathbf{v}_i = \mathbf{U} \mathbf{c}$$

$\mathbf{U} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ ist die Transformationsmatrix für die Orthonormalbasis, \mathbf{c} der entsprechende Koeffizientenvektor. Für die quadratische Form erhalten wir dann

$$q_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^t \cdot \mathbf{A} \cdot \mathbf{x} = (\mathbf{U} \mathbf{c})^t \cdot \mathbf{A} \cdot \mathbf{U} \mathbf{c} = \mathbf{c}^t \cdot \mathbf{U}^t \mathbf{A} \mathbf{U} \cdot \mathbf{c} = \mathbf{c}^t \cdot \mathbf{D} \cdot \mathbf{c}$$

wobei \mathbf{D} die Diagonalmatrix aus den Eigenwerten λ_i von \mathbf{A} ist. Daher

$$q_{\mathbf{A}}(\mathbf{x}) = \sum_{i=1}^n c_i^2 \lambda_i$$

Definitheit und Eigenwerte

Aus $q_{\mathbf{A}}(\mathbf{x}) = \sum_{i=1}^n c_i^2 \lambda_i$ folgt sofort:

Seien λ_i die Eigenwerte der Matrix \mathbf{A} . Dann gilt:

Die quadratische Form $q_{\mathbf{A}}$ ist genau dann

- ▶ *positiv definit*, wenn alle $\lambda_i > 0$ sind.
- ▶ *positiv semidefinit*, wenn alle $\lambda_i \geq 0$ sind.
- ▶ *negativ definit*, wenn alle $\lambda_i < 0$ sind.
- ▶ *negativ semidefinit*, wenn alle $\lambda_i \leq 0$ sind.
- ▶ *indefinit*, wenn es positive und negative Eigenwerte gibt.

Beispiel

- ▶ Die Eigenwerte von $\begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$ lauten $\lambda_1 = 6$ und $\lambda_2 = 1$.

Die Matrix ist daher positiv definit.

- ▶ Die Eigenwerte von $\begin{pmatrix} 5 & -1 & 4 \\ -1 & 2 & 1 \\ 4 & 1 & 5 \end{pmatrix}$ lauten $\lambda_1 = 0$, $\lambda_2 = 3$ und

$\lambda_3 = 9$. Die Matrix ist daher positiv semidefinit.

- ▶ Die Eigenwerte von $\begin{pmatrix} 7 & -5 & 4 \\ -5 & 7 & 4 \\ 4 & 4 & -2 \end{pmatrix}$ lauten

$\lambda_1 = -6$, $\lambda_2 = 6$ und $\lambda_3 = 12$. Die Matrix ist daher indefinit.

Hauptminoren

Die Definitheit einer Matrix kann auch mit Hilfe der sogenannten Hauptminoren bestimmt werden.

Sei $\mathbf{A} = (a_{ij})$ eine symmetrische $n \times n$ -Matrix. Dann heißt die Determinanten der Untermatrix

$$H_k = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{vmatrix}$$

der k -te (führende) **Hauptminor** von \mathbf{A} .

Hauptminoren und Definitheit

Eine quadratische Matrix \mathbf{A} ist genau dann

- ▶ *positiv definit*, wenn alle $H_k > 0$ sind.
- ▶ *negativ definit*, wenn $(-1)^k H_k > 0$ für alle k .
- ▶ *indefinit*, wenn $|\mathbf{A}| \neq 0$ und keiner der beiden Fälle zutrifft.

$(-1)^k H_k > 0$ bedeutet, dass

- ▶ $H_1, H_3, H_5, \dots < 0$, und
- ▶ $H_2, H_4, H_6, \dots > 0$.

Beispiel

Gesucht ist die Definitheit der Matrix

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

$$H_1 = \det(a_{11}) = a_{11} = 2 > 0$$

$$H_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = 5 > 0$$

$$H_3 = |\mathbf{A}| = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 2 \end{vmatrix} = 8 > 0$$

\mathbf{A} und $q_{\mathbf{A}}$ sind positiv definit.

Beispiel

Gesucht ist die Definitheit der Matrix

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

$$H_1 = \det(a_{11}) = a_{11} = 1 > 0$$

$$H_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 > 0$$

$$H_3 = |\mathbf{A}| = \begin{vmatrix} 1 & 1 & -2 \\ 1 & 2 & 3 \\ -2 & 3 & 1 \end{vmatrix} = -28 < 0$$

\mathbf{A} und $q_{\mathbf{A}}$ sind indefinit.

Allgemeine Hauptminoren

Die Bedingung für Semidefinitheit ist leider nicht so einfach.

Die k -ten **allgemeinen Hauptminoren** sind die Unterdeterminanten

$$\tilde{H}_{i_1, \dots, i_k} = \begin{vmatrix} a_{i_1, i_1} & \dots & a_{i_1, i_k} \\ \vdots & \ddots & \vdots \\ a_{i_k, i_1} & \dots & a_{i_k, i_k} \end{vmatrix} \quad 1 \leq i_1 < \dots < i_k \leq n.$$

Allgemeine Hauptminoren und Semidefinitheit

Eine quadratische Matrix \mathbf{A} ist genau dann

- ▶ *positiv semidefinit*, wenn alle $\tilde{H}_{i_1, \dots, i_k} \geq 0$ sind.
- ▶ *negativ semidefinit*, wenn $(-1)^k \tilde{H}_{i_1, \dots, i_k} \geq 0$ für alle k .
- ▶ *indefinit* in allen anderen Fällen.

$(-1)^k \tilde{H}_{i_1, \dots, i_k} \geq 0$ bedeutet, dass

- ▶ $\tilde{H}_{i_1, \dots, i_k} \geq 0$, falls k gerade ist, und
- ▶ $\tilde{H}_{i_1, \dots, i_k} \leq 0$, falls k ungerade ist.

Beispiel

Gesucht ist die Definitheit der Matrix

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

Die Matrix ist daher positiv semidefinit.

(Aber nicht positiv definit!)

1-te Hauptminoren:

$$\tilde{H}_1 = 5 \geq 0 \quad \tilde{H}_2 = 2 \geq 0$$

$$\tilde{H}_3 = 5 \geq 0$$

2-te Hauptminoren:

$$\tilde{H}_{1,2} = \begin{vmatrix} 5 & -1 \\ -1 & 2 \end{vmatrix} = 9 \geq 0$$

$$\tilde{H}_{1,3} = \begin{vmatrix} 5 & 4 \\ 4 & 5 \end{vmatrix} = 9 \geq 0$$

$$\tilde{H}_{2,3} = \begin{vmatrix} 2 & 1 \\ 1 & 5 \end{vmatrix} = 9 \geq 0$$

3-ter Hauptminor:

$$\tilde{H}_{1,2,3} = |\mathbf{A}| = 0 \geq 0$$

Beispiel

Gesucht ist die Definitheit der Matrix

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

Die Matrix ist negativ semidefinit.

(Aber nicht negativ definit!)

1-te Hauptminoren:

$$\tilde{H}_1 = -5 \leq 0 \quad \tilde{H}_2 = -2 \leq 0$$

$$\tilde{H}_3 = -5 \leq 0$$

2-te Hauptminoren:

$$\tilde{H}_{1,2} = \begin{vmatrix} -5 & 1 \\ 1 & -2 \end{vmatrix} = 9 \geq 0$$

$$\tilde{H}_{1,3} = \begin{vmatrix} -5 & -4 \\ -4 & -5 \end{vmatrix} = 9 \geq 0$$

$$\tilde{H}_{2,3} = \begin{vmatrix} -2 & -1 \\ -1 & -5 \end{vmatrix} = 9 \geq 0$$

3-ter Hauptminor:

$$\tilde{H}_{1,2,3} = |\mathbf{A}| = 0 \leq 0$$

Beispiel

Jede positiv definite Matrix ist auch positiv semidefinit
(aber nicht notwendigerweise umgekehrt).

Die Matrix

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

ist positiv definit da alle führenden Hauptminoren > 0 sind
(siehe oben).

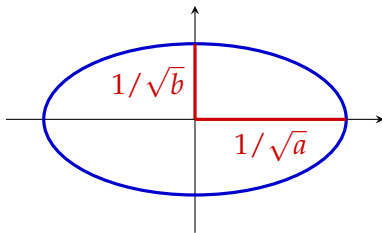
\mathbf{A} ist daher auch positiv semidefinit.

Ellipse

Die Gleichung

$$ax^2 + by^2 = 1, \quad a, b > 0$$

beschreibt eine *Ellipse* in **Hauptlage**.



Die Halbachsen haben die Längen $\frac{1}{\sqrt{a}}$ bzw. $\frac{1}{\sqrt{b}}$.

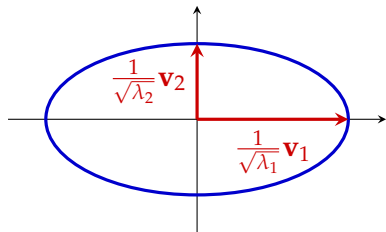
Eine geometrische Interpretation II

Der Term $ax^2 + by^2$ ist aber eine quadratische Form mit Matrix

$$\mathbf{A} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

Diese hat Eigenwerte und Eigenvektoren

$$\lambda_1 = a \text{ mit } \mathbf{v}_1 = \mathbf{e}_1 \quad \text{und} \quad \lambda_2 = b \text{ mit } \mathbf{v}_2 = \mathbf{e}_2$$



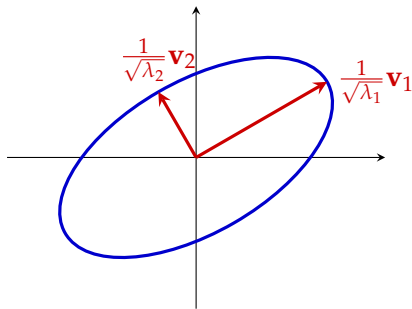
Die Eigenvektoren spannen die Hauptachsen der Ellipse auf.

Eine geometrische Interpretation II

Sei \mathbf{A} eine symmetrische 2×2 -Matrix mit *positiven* Eigenwerten.
Die Gleichung

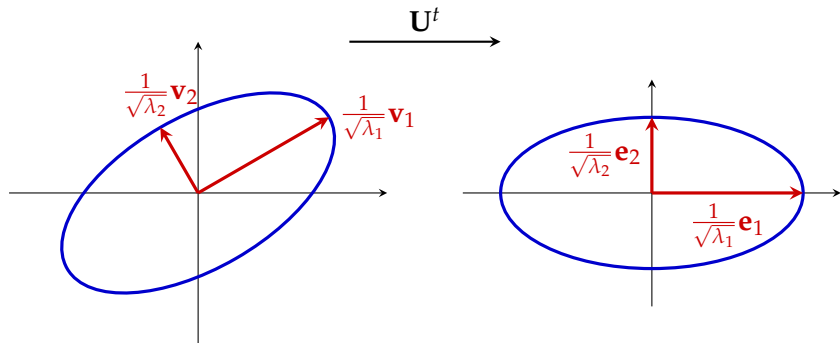
$$\mathbf{x}^t \mathbf{A} \mathbf{x} = 1$$

beschreibt eine *Ellipse*, deren Hauptachsen durch die normierten Eigenvektoren \mathbf{v}_1 und \mathbf{v}_2 von \mathbf{A} aufgespannt werden.



Eine geometrische Interpretation II

Durch den Basiswechsel $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2)$ von $\{\mathbf{e}_1, \mathbf{e}_2\}$ zu $\{\mathbf{v}_1, \mathbf{v}_2\}$ wird die Ellipse in die Hauptlage gedreht.



Man nennt daher diesen Vorgang (*Diagonalisieren*) auch **Hauptachsentransformation**.

Eine statistische Anwendung

Wir haben n Beobachtungen von k metrischen Merkmalen X_1, \dots, X_k , die wir zu Vektoren zusammenfassen können:

$$\mathbf{x}_i = (x_{i1}, \dots, x_{ik}) \in \mathbb{R}^k$$

Das arithmetischen Mittel ist (wie bei univariaten Daten)

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i = (\bar{x}_1, \dots, \bar{x}_k)$$

Die *Summe der Abweichungsquadrate* ist ein Maß für die Streuung

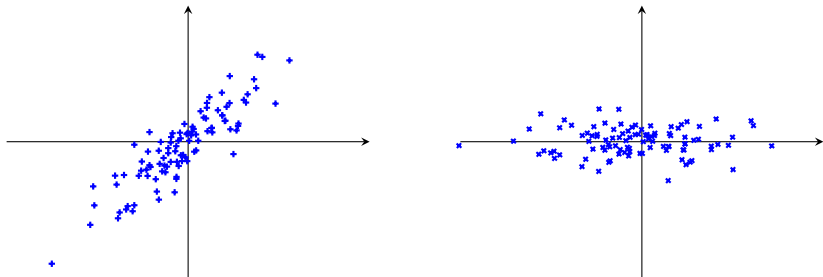
$$\text{TSS} = \sum_{i=1}^n \|\mathbf{x}_i - \bar{\mathbf{x}}\|^2 = \sum_{j=1}^k \left(\sum_{i=1}^n |x_{ij} - \bar{x}_j|^2 \right) = \sum_{j=1}^k \text{TSS}_j$$

und kann komponentenweise berechnet werden.

Eine statistische Anwendung

Ein orthogonaler Basiswechsel verändert die Summe der Abweichungsquadrate TSS nicht.

Er verändert aber die Aufteilung auf die einzelnen Komponenten.



Kann man die Basis so wählen, dass sich ein großer Beitrag für die TSS auf wenige Komponenten konzentriert?

Hauptkomponentenanalyse (PCA)

Annahme:

- ▶ Die Daten sind näherungsweise *multinormal* verteilt.

Vorgangsweise:

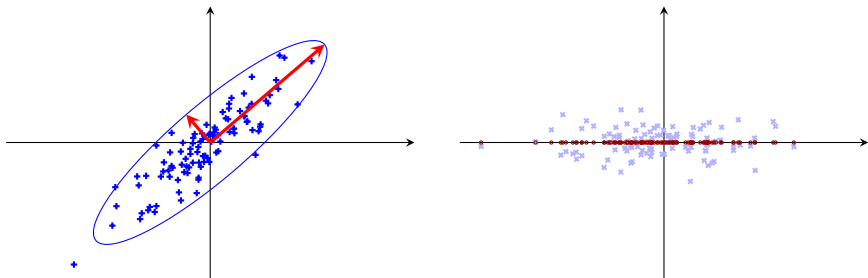
1. Berechne Kovarianzmatrix Σ
2. Berechne Eigenwerte und Eigenvektoren von Σ
3. Ordne Eigenwerte (und -vektoren) so, dass

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$$

4. Verwende Eigenvektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ als neue Basis.
5. Der Beitrag der m Komponenten in dieser Basis zum TSS ist näherungsweise

$$\frac{\sum_{j=1}^m \text{TSS}_j}{\sum_{j=1}^k \text{TSS}_j} \approx \frac{\sum_{j=1}^m \lambda_j}{\sum_{j=1}^k \lambda_j}$$

Hauptkomponentenanalyse (PCA)



Mit Hilfe der PCA ist es möglich die Anzahl der Dimensionen so zu reduzieren, dass die Gesamtstreuung näherungsweise erhalten bleibt.

Zusammenfassung

- ▶ Eigenwerte und Eigenvektoren
- ▶ Eigenraum
- ▶ Symmetrische Matrizen
- ▶ Diagonalisieren
- ▶ Quadratische Formen
- ▶ Definitheit
- ▶ Hauptminoren
- ▶ Hauptkomponentenanalyse

Kapitel 6

Funktionen

Reelle Funktion*

Reelle Funktionen sind Abbildungen, in denen sowohl die Definitionsmenge als auch die Wertemenge Teilmengen von \mathbb{R} (üblicherweise Intervalle) sind.

Bei reellen Funktionen wird meist weder Definitionsmenge noch Wertemenge angegeben. In diesem Fall gilt:

- ▶ Die *Definitionsmenge* ist die größtmögliche *sinnvolle* Teilmenge von \mathbb{R} , in der die Zuordnungsvorschrift definiert ist.
- ▶ Die *Wertemenge* ist die **Bildmenge**

$$f(D) = \{y \mid y = f(x) \text{ für ein } x \in D_f\} .$$

Beispiel*

Die Produktionsfunktion $f(x) = \sqrt{x}$ ist eine Abkürzung für

$$f: [0, \infty) \rightarrow [0, \infty), x \mapsto f(x) = \sqrt{x}$$

(Es gibt keine „negativen“ Produktionsmengen.
Ausserdem ist \sqrt{x} für $x < 0$ nicht reell!)

Die abgeleitete Funktion $f'(x) = \frac{1}{2\sqrt{x}}$ ist eine Abkürzung für

$$f': (0, \infty) \rightarrow (0, \infty), x \mapsto f'(x) = \frac{1}{2\sqrt{x}}$$

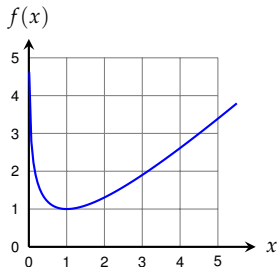
(Beachten sie das offene Intervall $(0, \infty)$.
 $\frac{1}{2\sqrt{x}}$ ist für $x = 0$ nicht definiert!)

Graph einer Funktion*

Jedem Paar $(x, f(x))$ entspricht ein Punkt in der xy -Ebene. Die Menge aller dieser Punkte bildet eine Kurve und heißt **Graph** der Funktion.

$$\mathcal{G}_f = \{(x, y) \mid y = f(x) \text{ für ein } x \in D_f\}$$

Wir können Funktionen mit Hilfe des Graphen veranschaulichen. Viele Eigenschaften von Funktionen lassen sich bereits aus deren Graphen herauslesen.

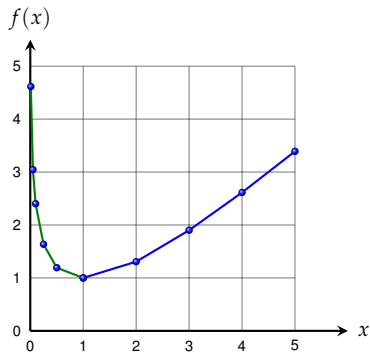


$$f(x) = x - \ln(x)$$

Zeichnen eines Graphen*

1. Wir überlegen uns zuerst wie der Graph wahrscheinlich aussehen wird. Graphen von elementaren Funktionen sollten bereits aus dem Gedächtnis skizziert werden können.
2. Wir wählen einen geeigneten Bereich auf der x -Achse aus. (Er sollte einen charakteristischen Ausschnitt zeigen.)
3. Wir erstellen eine Wertetabelle und zeichnen die entsprechenden Zahlenpaare in der xy -Ebene ein.
Charakteristische Punkte wie etwa lokale Extrema oder Wendepunkte sollten verwendet werden.
4. Wir überprüfen, ob aus den gezeichneten Punkten der Verlauf der Kurve ersichtlich ist. Andernfalls verlängern wir die Wertetabelle um einige geeignete Werte.
5. Die eingezeichneten Punkte werden in geeigneter Weise miteinander verbunden.

Beispiel*



Graph der Funktion

$$f(x) = x - \ln x$$

Wertetabelle:

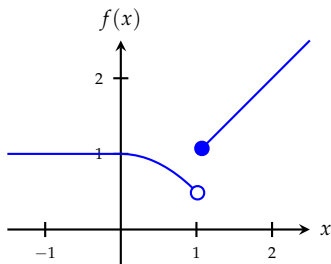
x	$f(x)$
0	ERROR
1	1
2	1,307
3	1,901
4	2,614
5	3,391
0,5	1,193
0,25	1,636
0,1	2,403
0,05	3,046

Stückweise definierte Funktionen*

Die Zuordnungsvorschrift einer Funktion kann auch in verschiedenen Intervallen des Definitionsbereichs verschieden definiert sein.

An den Intervallgrenzen müssen wir dann kennzeichnen, welche Punkte Bestandteil des Graphen sind:

- (Bestandteil) und \circ (nicht Bestandteil).



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

Bijektivität*

Jedes Argument besitzt immer genau ein Bild. Die Anzahl der Urbilder eines Elementes $y \in W$ kann jedoch beliebig sein. Wir können daher Funktionen nach der Anzahl der Urbilder einteilen.

- ▶ Eine Abbildung f heißt **injektiv**, wenn jedes Element aus der Wertemenge *höchstens* ein Urbild besitzt.
- ▶ Sie heißt **surjektiv**, wenn jedes Element aus der Wertemenge *mindestens* ein Urbild besitzt.
- ▶ Sie heißt **bijektiv**, wenn sie sowohl injektiv als auch surjektiv ist.

Eine Funktion besitzt genau dann eine *Umkehrfunktion* wenn sie *bijektiv* ist.

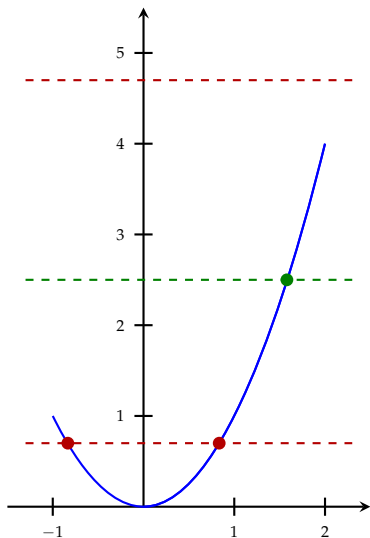
Horizontalen-Test*

Wie kann man feststellen, ob eine reelle Funktion injektiv / surjektiv ist?

Wie viele Urbilder kann ein $y \in W_f$ besitzen?

- (1) Wir zeichnen den Graphen der zu untersuchenden Funktion.
- (2) Wir zeichnen ein $y \in W$ auf der y -Achse ein und legen eine Gerade parallel zur x -Achse (Horizontale) durch diesen y -Wert.
- (3) Die Anzahl der Schnittpunkte von Horizontale und Graph ist die Anzahl der Urbilder von y .
- (4) Wir wiederholen (2) und (3) für eine *repräsentative* Auswahl von y -Werten.
- (5) Interpretation: Schneidet jede Horizontale den Graphen in
 - (a) *höchstens* einem Punkt, so ist f **injektiv**;
 - (b) *mindestens* einem Punkt, so ist f **surjektiv**;
 - (c) *genau* einem Punkt, so ist f **bijektiv**.

Beispiel*



$$f: [-1,2] \rightarrow \mathbb{R}, x \mapsto x^2$$

- ▶ ist nicht injektiv;
- ▶ ist nicht surjektiv;

$$f: [0,2] \rightarrow \mathbb{R}, x \mapsto x^2$$

- ▶ ist injektiv;
- ▶ ist nicht surjektiv;

$$f: [0,2] \rightarrow [0,4], x \mapsto x^2$$

- ▶ ist bijektiv;

Definitions- und Wertemenge sind Bestandteil der Funktion!

Die zusammengesetzte Funktion*

Seien $f: D_f \rightarrow W_f$ und $g: D_g \rightarrow W_g$ Funktionen mit $W_f \subseteq D_g$.
Dann heißt die Funktion

$$g \circ f: D_f \rightarrow W_g, x \mapsto (g \circ f)(x) = g(f(x))$$

zusammengesetzte Funktion („ g zusammengesetzt f “).

Seien

$$g: \mathbb{R} \rightarrow [0, \infty), x \mapsto g(x) = x^2,$$
$$f: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) = 3x - 2.$$

Dann ist

$$(g \circ f): \mathbb{R} \rightarrow [0, \infty),$$
$$x \mapsto (g \circ f)(x) = g(f(x)) = g(3x - 2) = (3x - 2)^2$$

und

$$(f \circ g): \mathbb{R} \rightarrow \mathbb{R},$$
$$x \mapsto (f \circ g)(x) = f(g(x)) = f(x^2) = 3x^2 - 2$$

Inverse Funktion*

Eine **bijektive** Funktion $f: D_f \rightarrow W_f$ besitzt eine **Umkehrfunktion** $f^{-1}: W_f \rightarrow D_f$ mit der Eigenschaft $f^{-1} \circ f = \text{id}$ und $f \circ f^{-1} = \text{id}$, i.e.,

$$f^{-1}(f(x)) = f^{-1}(y) = x \quad \text{und} \quad f(f^{-1}(y)) = f(x) = y$$

Die Zuordnungsvorschrift der inversen Abbildung einer reellen Funktion erhalten wir durch Vertauschen der Rollen von Argument x und Bild y .

Beispiel*

Zur Berechnung der inversen Funktion drücken wir x als Funktion von y aus.

Wir suchen die Umkehrfunktion von

$$y = f(x) = 2x - 1$$

Durch Umformung erhalten wir:

$$y = 2x - 1 \quad \Leftrightarrow \quad y + 1 = 2x \quad \Leftrightarrow \quad \frac{1}{2}(y + 1) = x$$

Die Umkehrfunktion lautet daher $f^{-1}(y) = \frac{1}{2}(y + 1)$.

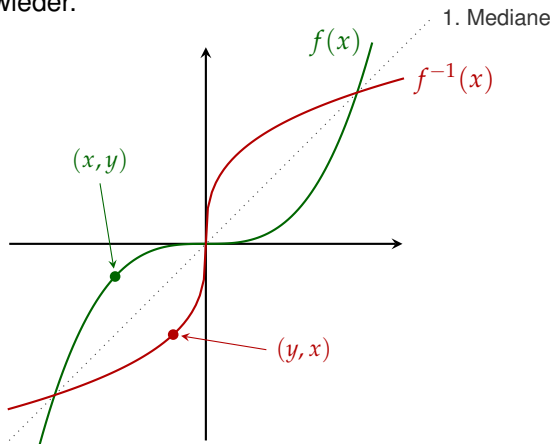
Da es üblich ist, das Argument mit x zu bezeichnen, schreiben wir

$$f^{-1}(x) = \frac{1}{2}(x + 1)$$

Die Umkehrfunktion von $f(x) = x^3$ ist $f^{-1}(x) = \sqrt[3]{x}$.

Geometrische Interpretation*

Das Vertauschen von x und y spiegelt sich auch im Graphen der Umkehrfunktion wieder.



(Graph der Funktion $f(x) = x^3$ und ihrer Inversen.)

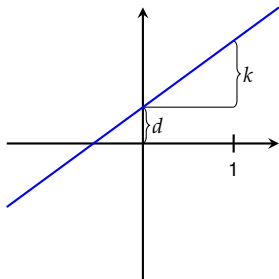
Lineare Funktion und Absolutbetrag*

► lineare Funktion

$$f(x) = kx + d$$

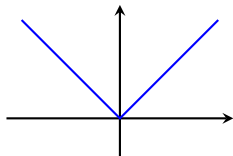
k ... **Steigung**

d ... **konstantes Glied**



► Betragsfunktion

$$f(x) = |x| = \begin{cases} x & \text{für } x \geq 0 \\ -x & \text{für } x < 0 \end{cases}$$

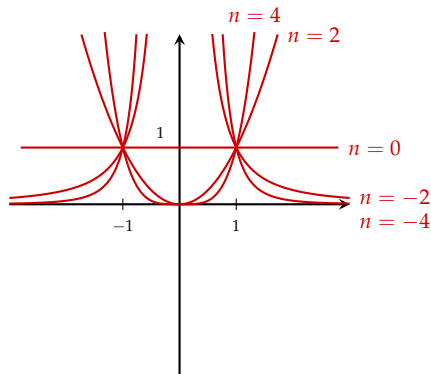
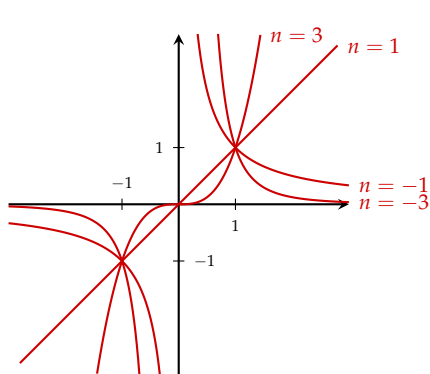


Potenzfunktion*

Potenzfunktion mit ganzzahligen Exponenten:

$$f: x \mapsto x^n \quad n \in \mathbb{Z}$$

$$D = \begin{cases} \mathbb{R} & \text{falls } n \geq 0 \\ \mathbb{R} \setminus \{0\} & \text{falls } n < 0 \end{cases}$$



Rechenregeln für Potenzen und Wurzeln*

$$x^{-n} = \frac{1}{x^n}$$

$$x^0 = 1 \quad (x \neq 0)$$

$$x^{n+m} = x^n \cdot x^m$$

$$x^{\frac{1}{m}} = \sqrt[m]{x} \quad (x \geq 0)$$

$$x^{n-m} = \frac{x^n}{x^m}$$

$$x^{\frac{n}{m}} = \sqrt[m]{x^n} \quad (x \geq 0)$$

$$(x \cdot y)^n = x^n \cdot y^n$$

$$x^{-\frac{n}{m}} = \frac{1}{\sqrt[m]{x^n}} \quad (x \geq 0)$$

$$(x^n)^m = x^{n \cdot m}$$

Achtung!

$-x^2$ ist **nicht** gleich $(-x)^2$

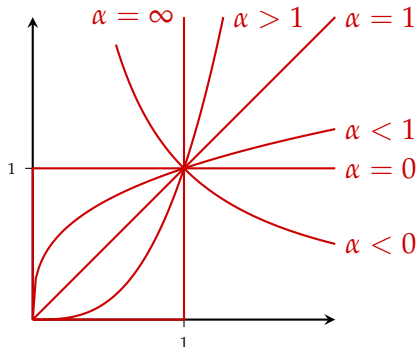
$(x + y)^n$ ist **nicht** gleich $x^n + y^n$

$x^n + y^n$ kann (im Allgemeinen) **nicht** vereinfacht werden!

Potenzfunktion II*

Potenzfunktion mit reellen Exponenten:

$$f: x \mapsto x^\alpha \quad \alpha \in \mathbb{R}$$



Polynome und rationale Funktionen*

- ▶ **Polynome** von Grad n :

$$f(x) = \sum_{k=0}^n a_k x^k$$

$a_i \in \mathbb{R}$, für $i = 1, \dots, n$

$a_n \neq 0$

- ▶ **Rationale Funktionen:**

$$D \rightarrow \mathbb{R}, \quad x \mapsto \frac{p(x)}{q(x)}$$

$p(x)$ und $q(x)$ sind Polynome

$D = \mathbb{R} \setminus \{\text{Nullstellen von } q\}$

Rechenregeln für Brüche und Bruchterme*

Seien $b, c, e \neq 0$

$$\frac{c \cdot a}{c \cdot b} = \frac{a}{b}$$

$$\frac{a}{b} = \frac{c \cdot a}{c \cdot b}$$

$$\frac{a}{b} \cdot \frac{d}{c} = \frac{a \cdot d}{b \cdot c}$$

$$\frac{a}{b} : \frac{e}{c} = \frac{a}{b} \cdot \frac{c}{e}$$

$$\frac{a}{b} + \frac{d}{b} = \frac{a + d}{b}$$

$$\frac{a}{b} + \frac{d}{c} = \frac{a \cdot c + d \cdot b}{b \cdot c}$$

$$\frac{\frac{a}{b}}{\frac{e}{c}} = \frac{a \cdot c}{b \cdot e}$$

Bei der Addition zuerst auf **gemeinsamen Nenner** bringen!

Rechenregeln für Brüche und Bruchterme*

Achtung!

$$\frac{a+c}{b+c} \text{ ist nicht gleich } \frac{a}{b}$$

$$\frac{x}{a} + \frac{y}{b} \text{ ist nicht gleich } \frac{x+y}{a+b}$$

$$\frac{a}{b+c} \text{ ist nicht gleich } \frac{a}{b} + \frac{a}{c}$$

$$\frac{x+2}{y+2} \neq \frac{x}{y}$$

$$\frac{1}{2} + \frac{1}{3} \neq \frac{1}{5}$$

$$\frac{1}{x^2 + y^2} \neq \frac{1}{x^2} + \frac{1}{y^2}$$

Exponentialfunktion*

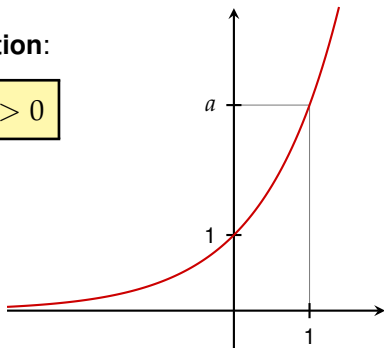
► **Exponentialfunktion:**

$$\mathbb{R} \rightarrow \mathbb{R}^+, \quad x \mapsto \exp(x) = e^x$$

$e = 2,7182818\dots$ Eulersche Zahl

► **Allgemeine Exponentialfunktion:**

$$\mathbb{R} \rightarrow \mathbb{R}^+, \quad x \mapsto a^x \quad a > 0$$



Logarithmusfunktion*

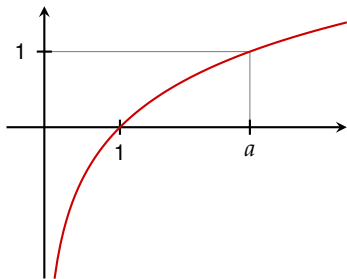
► **Logarithmusfunktion:**

Inverse Funktion zur *Exponentialfunktion*.

$$\mathbb{R}^+ \rightarrow \mathbb{R}, \quad x \mapsto \log(x) = \ln(x)$$

► **Allgemeine Logarithmusfunktion** zur Basis a

$$\mathbb{R}^+ \rightarrow \mathbb{R}, \quad x \mapsto \log_a(x)$$



Rechnen mit Exponenten und Logarithmus*

Eine Zahl y heißt Logarithmus von x zur Basis a , falls $a^y = x$. Der Logarithmus ist der Exponent einer Zahl bezüglich einer Basis a . Wir schreiben dafür

$$y = \log_a(x) \quad [\Leftrightarrow \quad x = a^y]$$

Wichtige Logarithmen:

- ▶ **natürlicher Logarithmus** $\ln(x)$ zur Basis $e = 2,7182818\dots$ (*Eulersche Zahl*)
- ▶ **dekadischer Logarithmus** $\lg(x)$ zur Basis 10

Rechnen mit Exponenten und Logarithmus*

Umrechnungsformel:

$$a^x = e^{x \ln(a)} \quad \log_a(x) = \frac{\ln(x)}{\ln(a)}$$

Achtung:

Oft schreibt man nur $\log(x)$ ohne Basisangabe.

In diesem Fall ist (sollte) die verwendete Basis aus dem Zusammenhang oder einer Konvention ersichtlich sein.

- ▶ Im mathematischen Bereich: natürlicher Logarithmus
Finanzmathematik, Programme wie R, *Mathematica*, *Maxima*, ...
- ▶ Im technischen Bereich: dekadischer Logarithmus
Wirtschaftswissenschaften, Taschenrechner, Excel, ...

Rechenregeln für Exponenten und Logarithmus*

$$a^{x+y} = a^x \cdot a^y$$

$$a^{\log_a(x)} = x$$

$$\log_a(x \cdot y) = \log_a(x) + \log_a(y)$$

$$\log_a(a^x) = x$$

$$\log_a\left(\frac{x}{y}\right) = \log_a(x) - \log_a(y)$$

$$\log_a(1) = 0$$

$$\log_a(x^\beta) = \beta \cdot \log_a(x)$$

$$\log_a(a) = 1$$

Achtung:

$\log_a(x)$ ist nur für $x > 0$ definiert!

$\log_a(x + y)$ ist **nicht** gleich $\log_a(x) + \log_a(y)$.

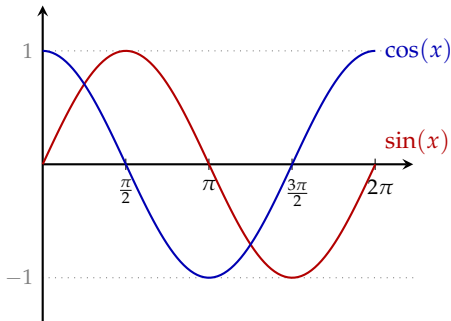
Winkelfunktionen*

► **Sinusfunktion:**

$$\mathbb{R} \rightarrow [-1,1], x \mapsto \sin(x)$$

► **Cosinusfunktion:**

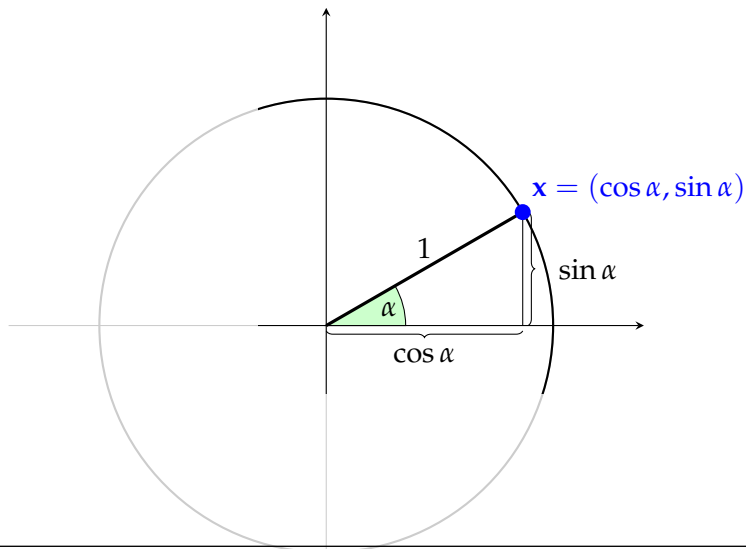
$$\mathbb{R} \rightarrow [-1,1], x \mapsto \cos(x)$$



Achtung!

x wird im *Bogenmaß (Radiant)* angegeben, d.h., ein rechter Winkel entspricht $x = \pi/2$.

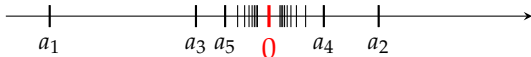
Sinus und Cosinus*



Grenzwert einer Folge*

Betrachten wir die Folge von Zahlen

$$(a_n)_{n=1}^{\infty} = \left((-1)^n \frac{1}{n} \right)_{n=1}^{\infty} = \left(-1, \frac{1}{2}, -\frac{1}{3}, \frac{1}{4}, -\frac{1}{5}, \frac{1}{6}, \dots \right)$$



Die Folgenglieder „*streben*“ mit wachsendem n gegen 0.

Wir sagen, die Folge (a_n) **konvergiert** gegen 0.

Wir schreiben dafür

$$(a_n) \rightarrow 0 \quad \text{oder} \quad \lim_{n \rightarrow \infty} a_n = 0$$

Grenzwert einer Folge / Definition*

Definition:

Eine Zahl $a \in \mathbb{R}$ heißt **Grenzwert** (Limes) einer Folge (a_n) , wenn es für jedes noch so kleine Intervall $(a - \varepsilon, a + \varepsilon)$ ein N gibt, sodass $a_n \in (a - \varepsilon, a + \varepsilon)$ für alle $n \geq N$.

M.a.W.: alle Folgenglieder ab a_N liegen im Intervall.

Äquivalente Formulierung:

Eine Folge (a_n) konvergiert gegen den Grenzwert $a \in \mathbb{R}$, wenn für jedes $\varepsilon > 0$ ein N existiert, sodass $|a_n - a| < \varepsilon$ für alle $n \geq N$.

[Mathematiker verwenden gerne ε für eine ganz kleine positive Zahl.]

Eine Folge, die einen Grenzwert besitzt, heißt **konvergent**.

Sie **konvergiert** gegen ihren Grenzwert.

Nicht jede Folge besitzt einen Grenzwert.

So eine Folge heißt **divergent**.

Grenzwert / Beispiele*

Die Folge $(a_n)_{n=1}^{\infty} = \left(\frac{1}{2^n}\right)_{n=1}^{\infty} = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots\right)$ konvergiert gegen 0:

$$\lim_{n \rightarrow \infty} a_n = 0$$

Die Folge $(b_n)_{n=1}^{\infty} = \left(\frac{n-1}{n+1}\right)_{n=1}^{\infty} = \left(0, \frac{1}{3}, \frac{2}{4}, \frac{3}{5}, \frac{4}{6}, \frac{5}{7}, \dots\right)$ ist konvergent:

$$\lim_{n \rightarrow \infty} b_n = 1$$

Die Folge $(c_n)_{n=1}^{\infty} = \left((-1)^n\right)_{n=1}^{\infty} = (-1, 1, -1, 1, -1, 1, \dots)$ ist divergent.

Die Folge $(d_n)_{n=1}^{\infty} = \left(2^n\right)_{n=1}^{\infty} = (2, 4, 8, 16, 32, \dots)$ ist divergent, strebt aber gegen ∞ . Man schreibt daher (nicht ganz korrekt):

$$\lim_{n \rightarrow \infty} d_n = \infty$$

Grenzwerte wichtiger Folgen*

$$\lim_{n \rightarrow \infty} n^a = \begin{cases} 0 & \text{für } a < 0 \\ 1 & \text{für } a = 0 \\ \infty & \text{für } a > 0 \end{cases}$$

$$\lim_{n \rightarrow \infty} q^n = \begin{cases} 0 & \text{für } |q| < 1 \\ 1 & \text{für } q = 1 \\ \infty & \text{für } q > 1 \\ \nexists & \text{für } q \leq -1 \end{cases}$$

$$\lim_{n \rightarrow \infty} \frac{n^a}{q^n} = \begin{cases} 0 & \text{für } |q| > 1 \\ \infty & \text{für } 0 < q < 1 \\ \nexists & \text{für } -1 < q < 0 \end{cases} \quad (|q| \notin \{0,1\})$$

Rechenregeln*

Seien $(a_n)_{n=1}^{\infty}$ und $(b_n)_{n=1}^{\infty}$ konvergente Folgen mit $\lim_{n \rightarrow \infty} a_n = a$ und $\lim_{n \rightarrow \infty} b_n = b$; und $(c_n)_{n=1}^{\infty}$ eine beschränkte Folge.

$$(1) \quad \lim_{n \rightarrow \infty} (k \cdot a_n + d) = k \cdot a + d$$

$$(2) \quad \lim_{n \rightarrow \infty} (a_n + b_n) = a + b$$

$$(3) \quad \lim_{n \rightarrow \infty} (a_n \cdot b_n) = a \cdot b$$

$$(4) \quad \lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{a}{b} \quad \text{für } b \neq 0$$

$$(5) \quad \lim_{n \rightarrow \infty} (a_n \cdot c_n) = 0 \quad \text{falls } a = 0$$

$$(6) \quad \lim_{n \rightarrow \infty} a_n^k = a^k$$

Rechenregeln / Beispiele*

$$\lim_{n \rightarrow \infty} \left(2 + \frac{3}{n^2} \right) = 2 + 3 \underbrace{\lim_{n \rightarrow \infty} n^{-2}}_{=0} = 2 + 3 \cdot 0 = 2$$

$$\lim_{n \rightarrow \infty} (2^{-n} \cdot n^{-1}) = \lim_{n \rightarrow \infty} \frac{n^{-1}}{2^n} = 0$$

$$\lim_{n \rightarrow \infty} \frac{1 + \frac{1}{n}}{2 - \frac{3}{n^2}} = \frac{\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} \right)}{\lim_{n \rightarrow \infty} \left(2 - \frac{3}{n^2} \right)} = \frac{1}{2}$$

$$\lim_{n \rightarrow \infty} \underbrace{\sin(n)}_{\text{beschränkt}} \cdot \underbrace{\frac{1}{n^2}}_{\rightarrow 0} = 0$$

Die Eulersche Zahl*

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e = 2,7182818284590 \dots$$

Dieser Grenzwert ist in der Finanzmathematik wichtig (stetige Verzinsung).

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n &= \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n/x}\right)^n \\ &= \lim_{m \rightarrow \infty} \left(1 + \frac{1}{m}\right)^{mx} && \left(m = \frac{n}{x}\right) \\ &= \left(\lim_{m \rightarrow \infty} \left(1 + \frac{1}{m}\right)^m\right)^x = e^x \end{aligned}$$

Dreiecksungleichung

Für zwei beliebige Zahlen $a, b \in \mathbb{R}$ gilt

$$|a + b| \leq |a| + |b|$$

Beweis:

Wir verwenden

$$|x| = \begin{cases} x & \text{falls } x \geq 0 \\ -x & \text{falls } x < 0 \end{cases} \quad \text{und} \quad x \leq |x|$$

Zwei Fälle:

1. $(a + b) \geq 0$: $|a + b| = a + b \leq |a| + |b|$.
2. $(a + b) < 0$: $|a + b| = -(a + b) = (-a) + (-b) \leq |a| + |b|$.

Dreiecksungleichung / Anwendung

Seien $(a_n) \rightarrow a$ und $(b_n) \rightarrow b$ zwei konvergente Folgen. Dann gilt

$$\lim_{n \rightarrow \infty} (a_n + b_n) = a + b \quad (\text{Regel 2})$$

Beweis:

Sei $\varepsilon > 0$ beliebig.

$(a_n) \rightarrow a$ heißt: Es gibt ein N_a sodass $|a_n - a| < \frac{\varepsilon}{2}$ für alle $n \geq N_a$.

$(b_n) \rightarrow b$ heißt: Es gibt ein N_b sodass $|b_n - b| < \frac{\varepsilon}{2}$ für alle $n \geq N_b$.

Daher gilt für alle $n \geq N = \max(N_a, N_b)$:

$$\begin{aligned} |(a_n + b_n) - (a + b)| &= |(a_n - a) + (b_n - b)| \\ &\stackrel{[\text{Dreiecksungleichung}]}{\leq} \underbrace{|a_n - a|}_{< \frac{\varepsilon}{2}} + \underbrace{|b_n - b|}_{< \frac{\varepsilon}{2}} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

I.e.: $(a_n + b_n) \rightarrow (a + b)$ q.e.d.

Grenzwert einer Funktion*

Was passiert mit dem Funktionswert einer Funktion f , wenn das Argument x gegen einen bestimmten Wert x_0 strebt?

Wenn für jede konvergente Folge von Argumenten $(x_n) \rightarrow x_0$ die Folge der Funktionswerte $(f(x_n))$ gegen eine Zahl a konvergiert, so heißt a der **Grenzwert** (oder **Limes**) der Funktion f *an der Stelle* x_0 .

Wir schreiben dafür

$$\lim_{x \rightarrow x_0} f(x) = a \quad \text{oder} \quad f(x) \rightarrow a \text{ für } x \rightarrow x_0$$

x_0 muss nicht in der Definitionsmenge liegen und kann daher auch ∞ sein. Genauso muss a nicht in der Wertemenge der Funktion liegen.

Für Limiten von Funktionen gelten analoge Rechenregeln wie für Grenzwerte von Folgen.

Bestimmen eines Grenzwertes*

Für einfache Funktionen eignet sich folgende Vorgangsweise:

1. Wir zeichnen den Graphen der Funktion.
2. Wir zeichnen den Wert x_0 auf der x -Achse ein.
3. Wir setzen den Bleistift auf dem Graphen und führen ihn auf dem Graphen von *rechts* bis zum x_0 -Wert.
4. Wir lesen den y -Wert dieses Punktes von y -Achse ab. Dieser Wert heißt der **rechtsseitige Grenzwert** von f an der Stelle x_0 :

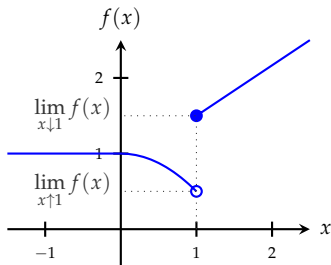
$$\lim_{x \downarrow x_0} f(x).$$

5. Analog erhalten wir von der *linken* Seite den **linksseitige Grenzwert** von f an der Stelle x_0 : $\lim_{x \uparrow x_0} f(x)$.

6. Wenn beide Limiten *gleich* sind, so existiert der Grenzwert und es gilt

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

Beispiel*



$$f(x) = \begin{cases} 1, & \text{für } x < 0 \\ 1 - \frac{x^2}{2}, & \text{für } 0 \leq x < 1 \\ \frac{x}{2} + 1, & \text{für } x \geq 1 \end{cases}$$

$$0,5 = \lim_{x \uparrow 1} f(x) \neq \lim_{x \downarrow 1} f(x) = 1,5$$

d.h., der Grenzwert an der Stelle $x_0 = 1$ existiert nicht.

Der Grenzwert an anderen Stellen existiert hingegen,

z.B. $\lim_{x \rightarrow 0} f(x) = 1.$

Stetigkeit*

Beim Zeichnen von Graphen fällt auf, dass es Funktionen gibt, die sich *ohne Absetzen des Bleistifts* zeichnen lassen.

Solche Funktionen heißen **stetig**.

Andere Funktionen besitzen *Sprungstellen* und man muss beim Zeichnen den Bleistift vom Papier heben. Solche Stellen heißen *Unstetigkeitsstellen* der Funktion.

Formal lässt sich das so ausdrücken:

Definition:

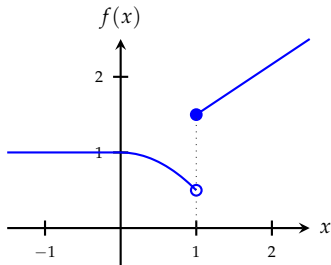
Eine Funktion f heißt **stetig** an der Stelle $x_0 \in D$, falls $\lim_{x \rightarrow x_0} f(x)$ existiert und gleich dem Funktionswert $f(x_0)$ ist.

Die Funktion heißt *stetig*, falls sie in allen Punkten des Definitionsbereichs stetig ist.

Stetigkeit*

Vorgangsweise für einfache Funktionen:

- (1) Wir zeichnen den Graphen der Funktion.
- (2) In allen Punkten des *Definitionsbereichs*, in denen wir beim Zeichnen *nicht* den Bleistift absetzen müssen, ist die Funktion stetig.
- (3) In allen Punkten des *Definitionsbereichs*, in denen wir absetzen müssen ist, die Funktion nicht stetig.



$$f(x) = \begin{cases} 1, & \text{für } x < 0 \\ 1 - \frac{x^2}{2}, & \text{für } 0 \leq x < 1 \\ \frac{x}{2} + 1, & \text{für } x \geq 1 \end{cases}$$

f ist überall stetig
außer im Punkt $x = 1$.

Funktionen in mehreren Variablen*

Eine **reelle Funktion in mehreren Variablen** ist eine Abbildung, die jedem Vektor \mathbf{x} eine reelle Zahl zuordnet.

$$f: \mathbb{R}^n \rightarrow \mathbb{R}, \mathbf{x} \mapsto f(\mathbf{x}) = f(x_1, x_2, \dots, x_n)$$

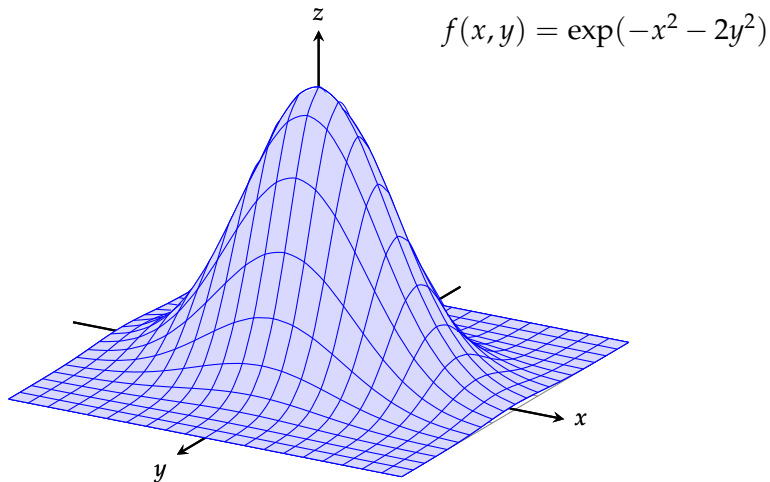
Die Komponenten x_i des Vektors \mathbf{x} heißen die **Variablen** der Funktion f .

Funktionen in *zwei* Variablen lassen sich durch den **Graphen** (Funktionengebirge) veranschaulichen:

$$\mathcal{G}_f = \{(\mathbf{x}, y) \mid y = f(\mathbf{x}) \text{ für ein } \mathbf{x} \in D_f\}$$

(Der Graph einer Funktion mit vielen Variablen ist analog definiert, er dient aber nicht mehr zur Veranschaulichung.)

Graph einer bivariaten Funktion*



Niveaulinien einer bivariaten Funktion*

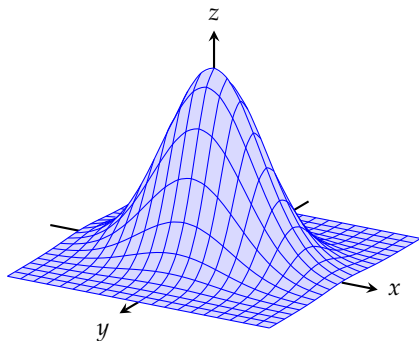
Die Menge aller Punkte (x, y) mit $f(x, y) = c$ für ein $c \in \mathbb{R}$ wird als **Niveaulinie** der Funktion f bezeichnet.

Die Funktion f hat daher auf einer Höhenlinie den gleichen Funktionswert.

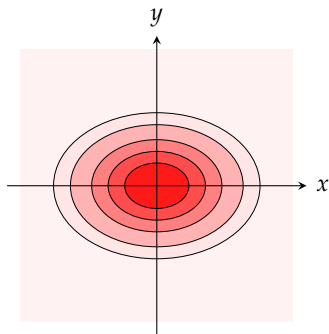
Andere Bezeichnungen:

- ▶ *Indifferenzkurve*
- ▶ *Isoquante* (Isonutzenlinie)
- ▶ *Höhenlinie*
- ▶ *Contourlinie*

Niveaulinien einer bivariaten Funktion*



Graph



Niveaulinien

$$f(x, y) = \exp(-x^2 - 2y^2)$$

Weg*

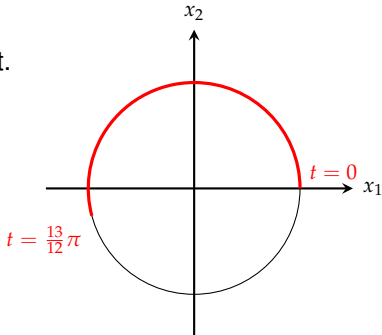
Eine Funktion

$$s: \mathbb{R} \rightarrow \mathbb{R}^n, t \mapsto s(t) = \begin{pmatrix} s_1(t) \\ \vdots \\ s_n(t) \end{pmatrix}$$

heißt ein **Weg** (oder *Pfad*) im \mathbb{R}^n .

Die Variable t wird oft als *Zeit* interpretiert.

$$[0, \infty) \rightarrow \mathbb{R}^2, t \mapsto \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$



Vektorwertige Funktion

Allgemeine vektorwertige Funktionen:

$$\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{x} \mapsto \mathbf{y} = \mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}$$

- ▶ Univariate Funktionen:

$$\mathbb{R} \rightarrow \mathbb{R}, x \mapsto y = x^2$$

- ▶ Multivariate Funktionen:

$$\mathbb{R}^2 \rightarrow \mathbb{R}, \mathbf{x} \mapsto y = x_1^2 + x_2^2$$

- ▶ Wege:

$$[0,1) \rightarrow \mathbb{R}^n, s \mapsto (s, s^2)^t$$

- ▶ Lineare Abbildungen:

$$\mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{x} \mapsto \mathbf{y} = \mathbf{A}\mathbf{x}$$

$\mathbf{A} \dots m \times n$ -Matrix

Zusammenfassung

- ▶ Abbildung
- ▶ Reelle Funktionen
- ▶ Graph einer Funktion
- ▶ Injektive, surjektiv, bijektiv
- ▶ Zusammengesetzte und inverse Funktion
- ▶ Potenzfunktion, Polynome und rationale Funktionen
- ▶ Exponentialfunktion und Logarithmus
- ▶ Winkelfunktionen
- ▶ Grenzwert
- ▶ Stetigkeit
- ▶ Funktionen in mehreren Variablen
- ▶ Wege und vektorwertige Funktionen

Kapitel 7

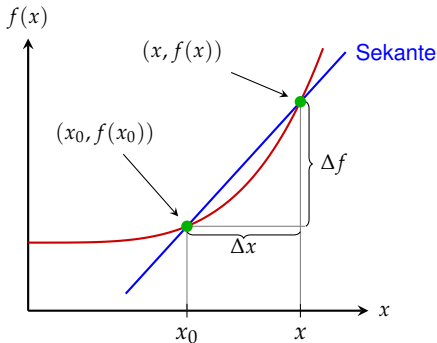
Differentialrechnung

Differenzenquotient*

Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion. Der Quotient

$$\frac{\Delta f}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \frac{f(x) - f(x_0)}{x - x_0}$$

heißt **Differenzenquotient** an der Stelle x_0 .



Differentialquotient*

Falls der Limes

$$\lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert, so heißt die Funktion f **differenzierbar** an der Stelle x_0 und dieser Grenzwert **Differentialquotient** oder (**erste**) **Ableitung** der Funktion an der Stelle x_0 .

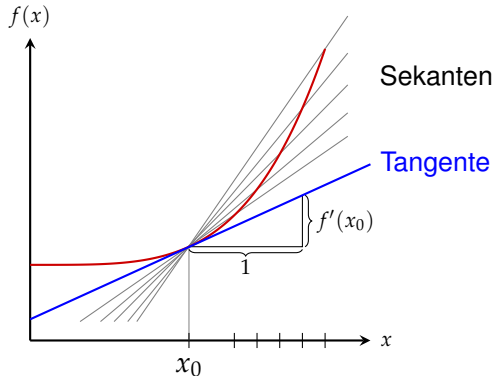
Eine Funktion f heißt *differenzierbar*, wenn sie in jedem Punkt des Definitionsbereichs differenzierbar ist.

Schreibweisen:

$$f'(x_0) = \left. \frac{df}{dx} \right|_{x=x_0}$$

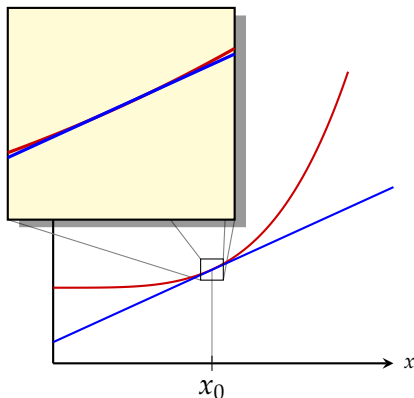
Graphische Interpretation des Differentialquotienten*

- ▶ Anstieg der Tangente an den Graphen der Funktion $f(x)$ an der Stelle x_0 .



Interpretation als „Grenzfunktion“*

- ▶ Marginalquote, oder „Grenzfunktion“ einer Wirkungsgröße $y = f(x)$ bezüglich einer Faktorgröße x .



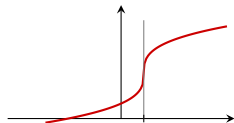
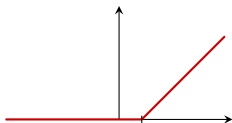
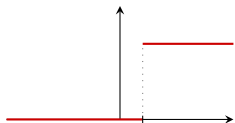
Existenz des Differentialquotienten*

Eine Funktion f ist differenzierbar in allen Punkten, in denen sich eine Tangente mit endlicher Steigung an den Graphen legen lässt.

In allen Punkten in denen das nicht möglich ist, ist die Funktion *nicht* differenzierbar.

Das sind vor allem

- ▶ Unstetigkeitsstellen („Sprungstellen“)
- ▶ „Knicke“ im Graph der Funktion
- ▶ Senkrechte Tangenten



Berechnung des Differentialquotienten*

Der Differentialquotient kann durch Bestimmen des Grenzwertes berechnet werden.

Sei $f(x) = x^2$. Dann ist

$$\begin{aligned} f'(x_0) &= \lim_{h \rightarrow 0} \frac{(x_0 + h)^2 - x_0^2}{h} \\ &= \lim_{h \rightarrow 0} \frac{x_0^2 + 2x_0h + h^2 - x_0^2}{h} \\ &= \lim_{h \rightarrow 0} \frac{2x_0h + h^2}{h} = \lim_{h \rightarrow 0} (2x_0 + h) \\ &= 2x_0 \end{aligned}$$

Marginale Änderung*

Für kleine Werte von Δx können wir die Ableitung $f'(x_0)$ durch den Differenzenquotienten mit kleinem Δx abschätzen:

$$f'(x_0) = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \approx \frac{\Delta f}{\Delta x}$$

Daher können wir umgekehrt die Änderung Δf von f in der Nähe von x_0 für *kleine* Änderungen Δx abschätzen:

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

Beachte:

- ▶ $f'(x_0) \cdot \Delta x$ ist eine *lineare Funktion* von Δx .
- ▶ Diese Funktion ist die *bestmögliche Approximation* von f durch eine lineare Funktion in der *Nähe* von x_0 .
- ▶ Diese Approximation ist nur für „kleine“ Werte von Δx brauchbar.

Differential*

Der Approximationsfehler geht dabei schneller gegen 0 als Δx :

$$\lim_{\Delta x \rightarrow 0} \frac{|f(x + \Delta x) - f(h)) - f'(x_0) \cdot \Delta x|}{|\Delta x|} = 0$$

Wenn wir die Differenzen Δf und Δx durch *infinitesimale* („unendlich kleine“) Größen df und dx ersetzen, erhalten wir das **Differential** der Funktion f an der Stelle x_0 :

$$df = f'(x_0) dx$$

df und dx heißen die **Differentiale** der Funktion f bzw. der unabhängigen Variable x .

Differential*

Wir können das Differential von f als lineare Funktion in dx auffassen und damit die Funktion f näherungsweise berechnen.

$$f(x_0 + dx) \approx f(x_0) + df$$

Sei $f(x) = e^x$.

Differential von f an der Stelle 1:

$$df = f'(1) dx = e^1 dx$$

Approximation von $f(1,1)$ mit Hilfe dieses Differentials:

$$\Delta x = (x_0 + dx) - x_0 = 1,1 - 1 = 0,1$$

$$f(1,1) \approx f(1) + df = e + e \cdot 0,1 \approx 2,99$$

Zum Vergleich: $f(1,1) = 3,004166 \dots$

Ableitung einer Funktion*

Die Funktion

$$f' : D \rightarrow \mathbb{R}, x \mapsto f'(x) = \left. \frac{df}{dx} \right|_x$$

heißt die **erste Ableitung** der Funktion f . Die Definitionsmenge D ist die Menge aller Punkte, in denen der Differentialquotient existiert.

Die Berechnung der Ableitung wird als **Ableiten** oder **Differenzieren** der Funktion bezeichnet.

Ableitung elementarer Funktionen*

$f(x)$	$f'(x)$
c	0
x^α	$\alpha \cdot x^{\alpha-1}$
e^x	e^x
$\ln(x)$	$\frac{1}{x}$
$\sin(x)$	$\cos(x)$
$\cos(x)$	$-\sin(x)$

Differentiationsregeln*

▶ $(c \cdot f(x))' = c \cdot f'(x)$

▶ $(f(x) + g(x))' = f'(x) + g'(x)$

Summenregel

▶ $(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x)$

Produktregel

▶ $(f(g(x)))' = f'(g(x)) \cdot g'(x)$

Kettenregel

▶ $\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{(g(x))^2}$

Quotientenregel

Beispiel*

$$(3x^3 + 2x - 4)' = 3 \cdot 3 \cdot x^2 + 2 \cdot 1 - 0 = 9x^2 + 2$$

$$(e^x \cdot x^2)' = (e^x)' \cdot x^2 + e^x \cdot (x^2)' = e^x \cdot x^2 + e^x \cdot 2x$$

$$((3x^2 + 1)^2)' = 2(3x^2 + 1) \cdot 6x$$

$$(\sqrt{x})' = \left(x^{\frac{1}{2}}\right)' = \frac{1}{2} \cdot x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}}$$

$$(a^x)' = \left(e^{\ln(a) \cdot x}\right)' = e^{\ln(a) \cdot x} \cdot \ln(a) = a^x \ln(a)$$

$$\left(\frac{1+x^2}{1-x^3}\right)' = \frac{2x \cdot (1-x^3) - (1+x^2) \cdot 3x^2}{(1-x^3)^2}$$

Höhere Ableitungen*

Die Ableitung einer Funktion kann wiederum differenziert werden.

Dadurch erhalten wir die

- ▶ **zweite Ableitung** $f''(x)$ der Funktion f ,
- ▶ **dritte Ableitung** $f'''(x)$, usw.
- ▶ **n -te Ableitung** $f^{(n)}(x)$.

Die ersten 5 Ableitungen der Funktion $f(x) = x^4 + 2x^2 + 5x - 3$ sind

$$f'(x) = (x^4 + 2x^2 + 5x - 3)' = 4x^3 + 4x + 5$$

$$f''(x) = (4x^3 + 4x + 5)' = 12x^2 + 4$$

$$f'''(x) = (12x^2 + 4)' = 24x$$

$$f^{IV}(x) = (24x)' = 24$$

$$f^V(x) = 0$$

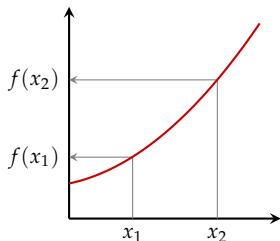
Monotonie*

Eine Funktion f heißt **monoton steigend**, falls

$$x_1 \leq x_2 \Leftrightarrow f(x_1) \leq f(x_2)$$

Sie heißt *streng monoton steigend*, falls

$$x_1 < x_2 \Leftrightarrow f(x_1) < f(x_2)$$

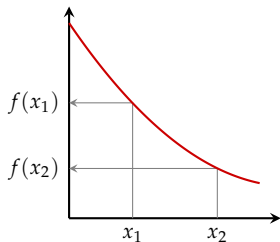


Eine Funktion f heißt **monoton fallend**, falls

$$x_1 \leq x_2 \Leftrightarrow f(x_1) \geq f(x_2)$$

Sie heißt *streng monoton fallend*, falls

$$x_1 < x_2 \Leftrightarrow f(x_1) > f(x_2)$$



Monotonie*

Für differenzierbare Funktionen gilt:

$$f \text{ monoton steigend} \Leftrightarrow f'(x) \geq 0 \quad \text{für alle } x \in D_f$$

$$f \text{ monoton fallend} \Leftrightarrow f'(x) \leq 0 \quad \text{für alle } x \in D_f$$

$$f \text{ streng monoton steigend} \Leftarrow f'(x) > 0 \quad \text{für alle } x \in D_f$$

$$f \text{ streng monoton fallend} \Leftarrow f'(x) < 0 \quad \text{für alle } x \in D_f$$

Die Funktion $f: (0, \infty), x \mapsto \ln(x)$ ist streng monoton steigend, da

$$f'(x) = (\ln(x))' = \frac{1}{x} > 0 \quad \text{für alle } x > 0$$

Lokale Monotonie*

Eine Funktion kann in einem bestimmten Intervall monoton steigend, in einem anderen monoton fallend sein. Sie wird dann als **lokal monoton**, auf dem entsprechenden Abschnitt bezeichnet.

Für *stetig* differenzierbare Funktionen eignet sich folgende Vorgangsweise:

1. Berechne erste Ableitung $f'(x)$.
2. Bestimme Nullstellen von $f'(x)$.
3. Erhalte Intervalle, in denen $f'(x)$ das Vorzeichen nicht wechselt.
4. Wähle "geeigneten" Punkt in jedem Intervall und bestimme dort das Vorzeichen von $f'(x)$.
5. $f'(x)$ kann innerhalb eines Intervalls das Vorzeichen nicht ändern.

Beispiel – Lokale Monotonie*

In welchen Bereichen ist $f(x) = 2x^3 - 12x^2 + 18x - 1$ monoton?

Suchen den Bereich, wo $f'(x) \geq 0$:

1. $f'(x) = 6x^2 - 24x + 18$

2. Nullstellen: $x^2 - 4x + 3 = 0 \Rightarrow x_1 = 1, x_2 = 3$

3. Erhalte 3 Intervalle: $(-\infty, 1]$, $[1, 3]$ und $[3, \infty)$

4. Vorzeichen von $f'(x)$ an “geeigneten” Punkt in jedem Intervall:
 $f'(0) = 3 > 0$, $f'(2) = -1 < 0$ und $f'(4) = 3 > 0$.

5. $f'(x)$ kann innerhalb eines Intervalls das Vorzeichen nicht ändern:
 $f'(x) \geq 0$ in $(-\infty, 1]$ und $[3, \infty)$.

Die Funktion $f(x)$ ist monoton steigend in $(-\infty, 1] \cup [3, \infty)$.

Monotonie und inverse Funktion*

Beachte:

Falls f *streng monoton* steigend ist, dann folgt aus

$$x_1 < x_2 \Leftrightarrow f(x_1) < f(x_2)$$

sofort auch

$$x_1 \neq x_2 \Leftrightarrow f(x_1) \neq f(x_2)$$

M.a.W., f ist injektiv. Falls f auch surjektiv ist, ist f somit *invertierbar*.

Falls eine surjektive Funktion f streng monoton steigend oder fallend ist, dann ist sie auch invertierbar.

(Die Umkehrung gilt jedoch nicht.)

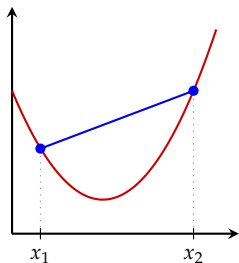
Konvexität und Konkavität*

Eine Funktion f heißt **konvex**, wenn

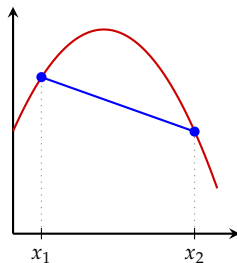
$$f((1-h)x_1 + hx_2) \leq (1-h)f(x_1) + hf(x_2)$$

für alle $x_1, x_2 \in D_f$ und alle $h \in [0,1]$. Sie heißt **konkav**, wenn

$$f((1-h)x_1 + hx_2) \geq (1-h)f(x_1) + hf(x_2)$$



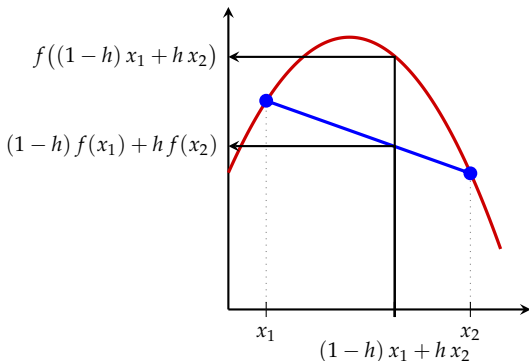
konvex



konkav

Konkave Funktion*

$$f((1-h)x_1 + hx_2) \geq (1-h)f(x_1) + hf(x_2)$$



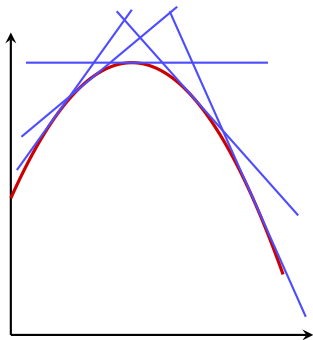
Sehne unterhalb des Funktionsgraphen.

Konvexität und Konkavität*

Für differenzierbare Funktionen gilt:

$$f \text{ konvex} \Leftrightarrow f''(x) \geq 0 \quad \text{für alle } x \in D_f$$

$$f \text{ konkav} \Leftrightarrow f''(x) \leq 0 \quad \text{für alle } x \in D_f$$



$f'(x)$ ist monoton fallend,
daher $f''(x) \leq 0$

Streng konvex / konkav*

Eine Funktion f heißt **streng konvex**, wenn

$$f((1-h)x_1 + hx_2) < (1-h)f(x_1) + hf(x_2)$$

für alle $x_1, x_2 \in D_f$, $x_1 \neq x_2$ und alle $h \in (0,1)$.

Sie heißt **streng konkav**, wenn

$$f((1-h)x_1 + hx_2) > (1-h)f(x_1) + hf(x_2)$$

Für differenzierbare Funktionen gilt:

$$\begin{aligned} f \text{ streng konvex} &\Leftrightarrow f''(x) > 0 && \text{für alle } x \in D_f \\ f \text{ streng konkav} &\Leftrightarrow f''(x) < 0 && \text{für alle } x \in D_f \end{aligned}$$

Beispiel – konvex*

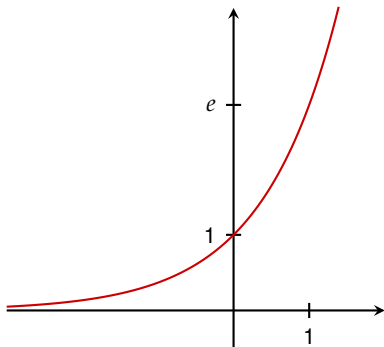
Exponentialfunktion:

$$f(x) = e^x$$

$$f'(x) = e^x$$

$$f''(x) = e^x > 0 \quad \text{für alle } x \in \mathbb{R}$$

$\exp(x)$ ist (streng) konvex.



Beispiel – konkav*

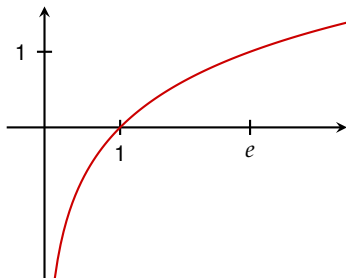
Logarithmusfunktion: $(x > 0)$

$$f(x) = \ln(x)$$

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

$$f''(x) = -\frac{1}{x^2} < 0 \quad \text{für alle } x > 0$$

$\ln(x)$ ist (streng) konkav.



Lokale Konkavität*

Eine Funktion kann auch auf einem bestimmten Intervall konkav, auf einem anderen konvex sein. Sie wird dann als **lokal konkav**, bzw. **lokal konvex** auf dem entsprechenden Abschnitt bezeichnet.

Für *stetig* differenzierbare Funktionen eignet sich folgende Vorgangsweise:

1. Berechne zweite Ableitung $f''(x)$.
2. Bestimme Nullstellen von $f''(x)$.
3. Erhalte Intervalle, in denen $f''(x)$ das Vorzeichen nicht wechselt.
4. Wähle "geeigneten" Punkt in jedem Intervall und bestimme dort das Vorzeichen von $f''(x)$.
5. $f''(x)$ kann innerhalb eines Intervalls das Vorzeichen nicht ändern.

Beispiel – Lokale Konkavität*

In welchem Bereich ist $f(x) = 2x^3 - 12x^2 + 18x - 1$ konkav?

Suchen den Bereich, wo $f''(x) \leq 0$.

1. $f''(x) = 12x - 24$

2. Nullstellen: $12x - 24 = 0 \Rightarrow x = 2$

3. Erhalte 2 Intervalle: $(-\infty, 2]$ und $[2, \infty)$

4. Vorzeichen von $f''(x)$ an "geeigneten" Punkt in jedem Intervall:
 $f''(0) = -24 < 0$ und $f''(4) = 24 > 0$.

5. $f''(x)$ kann innerhalb eines Intervalls das Vorzeichen nicht ändern: $f''(x) \leq 0$ in $(-\infty, 2]$

Die Funktion $f(x)$ ist konkav in $(-\infty, 2]$.

Elastizität*

Die erste Ableitung einer Funktion gibt die *Änderungsrate* einer Funktion f an der Stelle x_0 in *absoluten* Zahlen an. Sie ist somit abhängig von der *Skalierung* von Argument und Funktionswert.

Wir sind aber in vielen Fällen an *relativen* Änderungsraten interessiert.

Skaleninvarianz und *relative* Änderungsraten erhalten wir durch

$$\frac{\text{Änderung des Funktionswertes in \% des Funktionswertes}}{\text{Änderung des Arguments in \% des Argumentes}}$$

bzw. für die marginale Änderungsrate

$$\lim_{\Delta x \rightarrow 0} \frac{\frac{f(x+\Delta x) - f(x)}{f(x)}}{\frac{\Delta x}{x}} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \cdot \frac{x}{f(x)} = f'(x) \cdot \frac{x}{f(x)}$$

Elastizität*

Der Ausdruck

$$\varepsilon_f(x) = x \cdot \frac{f'(x)}{f(x)}$$

heißt die **Elastizität** von f an der Stelle x .

Sei $f(x) = 3e^{2x}$. Dann ist

$$\varepsilon_f(x) = x \cdot \frac{f'(x)}{f(x)} = x \cdot \frac{6e^{2x}}{3e^{2x}} = 2x$$

Sei $f(x) = \beta x^\alpha$. Dann ist

$$\varepsilon_f(x) = x \cdot \frac{f'(x)}{f(x)} = x \cdot \frac{\beta \alpha x^{\alpha-1}}{\beta x^\alpha} = \alpha$$

Elastische Funktionen*

Eine Funktion f heißt in x

- ▶ **elastisch**, falls $|\varepsilon_f(x)| > 1$
- ▶ **1-elastisch**, falls $|\varepsilon_f(x)| = 1$
- ▶ **unelastisch**, falls $|\varepsilon_f(x)| < 1$

Für eine elastische Funktion gilt daher:

Der Funktionswert ändert sich *relative* stärker als das Argument.

Die Funktion $f(x) = 3e^{2x}$ ist

$$[\varepsilon_f(x) = 2x]$$

- ▶ 1-elastisch, für $x = -\frac{1}{2}$ und $x = \frac{1}{2}$;
- ▶ unelastisch, für $-\frac{1}{2} < x < \frac{1}{2}$;
- ▶ elastisch, für $x < -\frac{1}{2}$ oder $x > \frac{1}{2}$.

elastische Nachfragefunktion*

Sei $q(p)$ eine *elastische* Nachfragefunktion, p der Preis.

Es gilt: $p > 0$, $q > 0$, und $q' < 0$ (q ist monoton fallend). Also gilt

$$\varepsilon_q(p) = p \cdot \frac{q'(p)}{q(p)} < -1$$

Was passiert mit dem Umsatz (= Preis \times Absatz)?

$$\begin{aligned} u'(p) &= (p \cdot q(p))' = 1 \cdot q(p) + p \cdot q'(p) \\ &= q(p) \cdot \underbrace{\left(1 + p \cdot \frac{q'(p)}{q(p)}\right)}_{=\varepsilon_q < -1} \\ &< 0 \end{aligned}$$

Das heißt, der Umsatz nimmt ab, falls wir den Preis erhöhen.

Elastizität II*

Wir können die relative Änderungsrate von f ausdrücken als Ableitung

$$\ln(f(x))' = \frac{f'(x)}{f(x)}$$

Was passiert, wenn wir $\ln(f(x))$ nach $\ln(x)$ ableiten?

$$\text{Sei } v = \ln(x) \Leftrightarrow x = e^v$$

Ableiten mittels Kettenregel ergibt:

$$\frac{d(\ln(f(x)))}{d(\ln(x))} = \frac{d(\ln(f(e^v)))}{dv} = \frac{f'(e^v)}{f(e^v)} e^v = \frac{f'(x)}{f(x)} x = \varepsilon_f(x)$$

$$\varepsilon_f(x) = \frac{d(\ln(f(x)))}{d(\ln(x))}$$

Elastizität II*

Wir können die Kettenregel *formal* auch so schreiben:

Sei

- ▶ $u = \ln(y)$,
- ▶ $y = f(x)$,
- ▶ $x = e^v \Leftrightarrow v = \ln(x)$

Dann erhalten wir

$$\frac{d(\ln f)}{d(\ln x)} = \frac{du}{dv} = \frac{du}{dy} \cdot \frac{dy}{dx} \cdot \frac{dx}{dv} = \frac{1}{y} \cdot f'(x) \cdot e^v = \frac{f'(x)}{f(x)} x$$

Partielle Ableitung*

Wir untersuchen die Änderung einer Funktion $f(x_1, \dots, x_n)$, wenn wir eine Variable x_i variieren und alle anderen konstant lassen.

$$\frac{\partial f}{\partial x_i} = \lim_{\Delta x_i \rightarrow 0} \frac{f(\dots, x_i + \Delta x_i, \dots) - f(\dots, x_i, \dots)}{\Delta x_i}$$

heißt die (erste) **partielle Ableitung** von f nach x_i .

Es haben sich eine Reihe von weiteren Symbolen für die partielle Ableitung $\frac{\partial f}{\partial x_i}$ eingebürgert:

- ▶ $f_{x_i}(\mathbf{x})$ (Ableitung nach der Variable x_i)
- ▶ $f_i(\mathbf{x})$ (Ableitung nach der i -ten Variable)
- ▶ $f'_i(\mathbf{x})$ (i -te Komponente des Gradienten)

Berechnung der partiellen Ableitung*

Wir erhalten die partielle Ableitung nach x_i , wenn wir alle anderen Variablen als Konstante auffassen und f nach den bekannten Regeln für Funktionen in einer Variable nach x_i ableiten.

Erste partiellen Ableitungen von

$$f(x_1, x_2) = \sin(2x_1) \cdot \cos(x_2)$$

$$f_{x_1} = 2 \cdot \cos(2x_1) \cdot \underbrace{\cos(x_2)}_{\text{als Konstante betrachtet}}$$

$$f_{x_2} = \underbrace{\sin(2x_1)}_{\text{als Konstante betrachtet}} \cdot (-\sin(x_2))$$

Höhere partielle Ableitungen*

Analog zu den Funktionen in einer Variablen können wir partielle Ableitungen nochmals ableiten und erhalten so **höhere partielle Ableitungen**.

$$f_{x_i x_k}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_k \partial x_i}(\mathbf{x}) \quad \text{bzw.} \quad f_{x_i x_i}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_i^2}(\mathbf{x})$$

Falls alle zweiten partiellen Ableitungen existieren und *stetig* sind, dann kommt auf die Reihenfolge beim Differenzieren nicht an.

$$\frac{\partial^2 f}{\partial x_k \partial x_i}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_i \partial x_k}(\mathbf{x})$$

Beispiel*

Gesucht sind alle ersten und zweiten partiellen Ableitungen von

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

Erste partielle Ableitungen:

$$f_x = 2x + 3y \quad f_y = 0 + 3x$$

Zweite partielle Ableitungen:

$$\begin{aligned} f_{xx} &= 2 & f_{xy} &= 3 \\ f_{yx} &= 3 & f_{yy} &= 0 \end{aligned}$$

Partielle Elastizitäten*

Die **partiellen Elastizitäten** geben *relative* Änderungsraten bezüglich den einzelnen Variablen an.

$$\varepsilon_{f,i}(\mathbf{x}) = x_i \cdot \frac{f_{x_i}(\mathbf{x})}{f(\mathbf{x})}$$

Gesucht sind die partiellen Elastizitäten von

$$f(x_1, x_2) = x_1^2 + x_2^3$$

$$\varepsilon_{f,1}(\mathbf{x}) = x_1 \cdot \frac{f_{x_1}(\mathbf{x})}{f(\mathbf{x})} = x_1 \cdot \frac{2x_1}{x_1^2 + x_2^3} = \frac{2x_1^2}{x_1^2 + x_2^3}$$

$$\varepsilon_{f,2}(\mathbf{x}) = x_2 \cdot \frac{f_{x_2}(\mathbf{x})}{f(\mathbf{x})} = x_2 \cdot \frac{3x_2^2}{x_1^2 + x_2^3} = \frac{3x_2^3}{x_1^2 + x_2^3}$$

Kreuzpreiselastizität*

Zwei Güter werden zu den Preisen p_1 bzw. p_2 angeboten. Die Nachfrage q_1 nach Gut 1 hängt nicht nur vom Preis für Gut 1 ab, sondern auch vom Preis des anderen Gutes:

$$\mathbf{q}(\mathbf{p}) = \begin{pmatrix} q_1(p_1, p_2) \\ q_2(p_1, p_2) \end{pmatrix}$$

Die partiellen Elastizitäten $\varepsilon_{q_1,2}(\mathbf{p})$ und $\varepsilon_{q_2,1}(\mathbf{p})$ heißen die **Kreuzpreiselastizitäten**.

- ▶ $\varepsilon_{q_i,j}(\mathbf{p}) > 0 \Rightarrow$ Güter sind *Substitute*
- ▶ $\varepsilon_{q_i,j}(\mathbf{p}) < 0 \Rightarrow$ Güter sind *komplementär*
- ▶ $\varepsilon_{q_i,j}(\mathbf{p}) = 0 \Rightarrow$ Güter ohne Beziehung

Im Allgemeinen ist $\varepsilon_{q_1,2}(\mathbf{x}) \neq \varepsilon_{q_2,1}(\mathbf{x})$.

Der Gradient

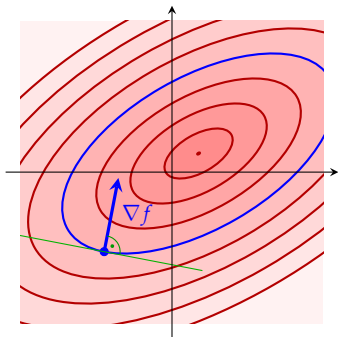
Wir fassen die partiellen Ableitungen erster Ordnung zu einem Zeilenvektor, dem **Gradienten** an der Stelle \mathbf{x} , zusammen.

$$\nabla f(\mathbf{x}) = (f_{x_1}(\mathbf{x}), \dots, f_{x_n}(\mathbf{x}))$$

- ▶ ∇f heißt auch „Nabla f “.
- ▶ Der Gradient wird oft auch als Spaltenvektor geschrieben.
- ▶ Andere Notationen: $f'(\mathbf{x})$
- ▶ Der Gradient „spielt“ die gleiche Rolle wie die erste Ableitung bei Funktionen in einer Variablen.

Eigenschaften des Gradienten

- ▶ Der Gradient einer Funktion f zeigt in die Richtung des steilsten Anstiegs von f .
- ▶ Seine Länge gibt diese Steigung an.
- ▶ Der Gradient steht immer normal auf die entsprechende Niveaulinie.



Beispiel

Gesucht ist der Gradient von

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

an der Stelle $\mathbf{x} = (3, 2)$.

$$f_x = 2x + 3y$$

$$f_y = 0 + 3x$$

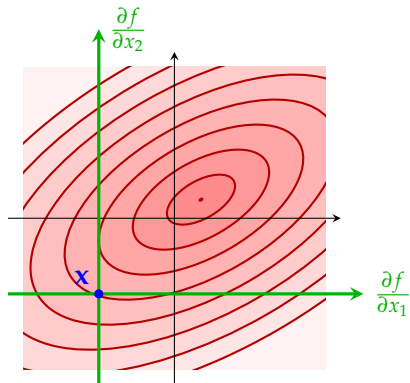
$$\nabla f(\mathbf{x}) = (2x + 3y, 3x)$$

$$\nabla f(3, 2) = (12, 9)$$

Die Richtungsableitung

Wir erhalten die partielle Ableitung $\frac{\partial f}{\partial x_i}$ durch Ableiten der univariaten Funktion $g(t) = f(x_1, \dots, x_i + t, \dots, x_n) = f(\mathbf{x} + t \cdot \mathbf{h})$ mit $\mathbf{h} = \mathbf{e}_i$ an der Stelle $t = 0$:

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \left. \frac{dg}{dt} \right|_{t=0} = \left. \frac{d}{dt} f(\mathbf{x} + t \cdot \mathbf{h}) \right|_{t=0}$$

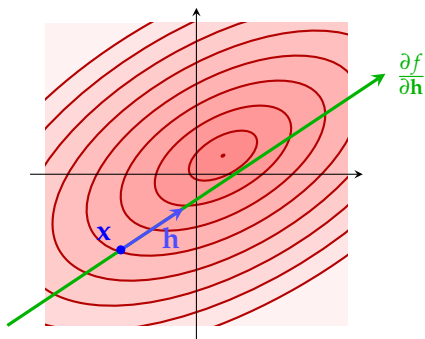


Die Richtungsableitung

Verallgemeinerung:

Wir erhalten die **Richtungsableitung** $\frac{\partial f}{\partial \mathbf{h}}$ in Richtung \mathbf{h} mit Länge 1 durch Ableiten der univariaten Funktion $g(t) = f(\mathbf{x} + t \cdot \mathbf{h})$ an der Stelle $t = 0$:

$$\frac{\partial f}{\partial \mathbf{h}}(\mathbf{x}) = \left. \frac{dg}{dt} \right|_{t=0} = \left. \frac{d}{dt} f(\mathbf{x} + t \cdot \mathbf{h}) \right|_{t=0}$$



Die Richtungsableitung gibt die Änderung von f an, wenn wir \mathbf{x} in Richtung \mathbf{h} verschieben.

Die Richtungsableitung

Es gilt (für $\|\mathbf{h}\| = 1$):

$$\frac{\partial f}{\partial \mathbf{h}}(\mathbf{x}) = f_{x_1}(\mathbf{x}) \cdot h_1 + \cdots + f_{x_n}(\mathbf{x}) \cdot h_n = \nabla f(\mathbf{x}) \cdot \mathbf{h}$$

Falls \mathbf{h} nicht Norm 1 hat, muss zuerst normiert werden:

$$\frac{\partial f}{\partial \mathbf{h}}(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \frac{\mathbf{h}}{\|\mathbf{h}\|}$$

Beispiel

Wir suchen die Richtungsableitung von

$$f(x_1, x_2) = x_1^2 + 3 x_1 x_2$$

nach $\mathbf{h} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ an der Stelle $\mathbf{x} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$.

Norm von \mathbf{h} :

$$\|\mathbf{h}\| = \sqrt{\mathbf{h}^t \mathbf{h}} = \sqrt{1^2 + (-2)^2} = \sqrt{5}$$

Die Richtungsableitung lautet daher

$$\frac{\partial f}{\partial \mathbf{h}}(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \frac{\mathbf{h}}{\|\mathbf{h}\|} = \frac{1}{\sqrt{5}} (12, 9) \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} = -\frac{6}{\sqrt{5}}$$

Das totale Differential

Wir wollen eine Funktion f durch eine lineare Funktion so approximieren, dass der Fehler möglichst klein ist.

Den Funktionswert an einer Stelle $\mathbf{x} + \mathbf{h}$ können wir näherungsweise analog zur Richtungsableitung berechnen.

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) \approx f_{x_1}(\mathbf{x}) h_1 + \dots + f_{x_n}(\mathbf{x}) h_n$$

Das totale Differential erhalten wir, wenn wir die h_i durch „unendlich kleine“ Differentiale dx_i ersetzen.

Die *lineare Funktion*

$$df = f_{x_1}(\mathbf{x}) dx_1 + \dots + f_{x_n}(\mathbf{x}) dx_n = \sum_{i=1}^n f_{x_i} dx_i$$

heißt das **totale Differential** von f an der Stelle \mathbf{x} .

Beispiel

Wir suchen das totale Differential von

$$f(x_1, x_2) = x_1^2 + 3 x_1 x_2$$

an der Stelle $\mathbf{x} = (3,2)$.

$$df = f_{x_1}(3,2) dx_1 + f_{x_2}(3,2) dx_2 = 12 dx_1 + 9 dx_2$$

Approximation von $f(3,1;1,8)$ mit Hilfe des totalen Differentials:

$$\begin{aligned} f(3,1;1,8) &\approx f(3;2) + df \\ &= 27 + 12 \cdot 0,1 + 9 \cdot (-0,2) = 26,40 \end{aligned}$$

Zum Vergleich: $f(3,1;1,8) = 26,35$

$$\mathbf{h} = (\mathbf{x} + \mathbf{h}) - \mathbf{x} = \begin{pmatrix} 3,1 \\ 1,8 \end{pmatrix} - \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 0,1 \\ -0,2 \end{pmatrix}$$

Differenzierbarkeit

Satz:

Eine Abbildung $f: \mathbb{R} \rightarrow \mathbb{R}$ ist differenzierbar in x_0 genau dann, wenn es eine lineare Abbildung ℓ gibt, die f in x_0 bestmöglich approximiert:

$$\lim_{h \rightarrow 0} \frac{|(f(x_0 + h) - f(x_0)) - \ell(h)|}{|h|} = 0$$

Offensichtlich: $\ell(h) = f'(x_0) \cdot h$ ist das Differential von f .

Definition:

Eine Abbildung $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt differenzierbar in \mathbf{x}_0 , wenn es eine lineare Abbildung ℓ gibt, die \mathbf{f} in \mathbf{x}_0 bestmöglich approximiert:

$$\lim_{\mathbf{h} \rightarrow 0} \frac{\|(\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0)) - \ell(\mathbf{h})\|}{\|\mathbf{h}\|} = 0$$

$\ell(\mathbf{h}) = \mathbf{J} \cdot \mathbf{h}$ heißt das Differential von \mathbf{f} .

Jacobische Matrix

Die $m \times n$ -Matrix

$$D\mathbf{f}(\mathbf{x}_0) = \mathbf{f}'(\mathbf{x}_0) = \mathbf{J} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

heißt die **Jacobische Matrix** von \mathbf{f} and der Stelle \mathbf{x}_0 .

Für $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ist $Df(\mathbf{x}_0) = \nabla f(\mathbf{x}_0)$.

Wir können $D\mathbf{f}(\mathbf{x}_0)$ auch schreiben als

$$D\mathbf{f}(\mathbf{x}_0) = \begin{pmatrix} \nabla f_1(\mathbf{x}_0) \\ \vdots \\ \nabla f_m(\mathbf{x}_0) \end{pmatrix}$$

Beispiel

$$\blacktriangleright f(\mathbf{x}) = f(x_1, x_2) = \exp(-x_1^2 - x_2^2)$$

$$\begin{aligned} Df(\mathbf{x}) &= \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right) = \nabla f(\mathbf{x}) \\ &= (-2x_1 \exp(-x_1^2 - x_2^2), -2x_2 \exp(-x_1^2 - x_2^2)) \end{aligned}$$

$$\blacktriangleright \mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, x_2) = \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} x_1^2 + x_2^2 \\ x_1^2 - x_2^2 \end{pmatrix}$$

$$D\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1 & 2x_2 \\ 2x_1 & -2x_2 \end{pmatrix}$$

$$\blacktriangleright \mathbf{s}(t) = \begin{pmatrix} s_1(t) \\ s_2(t) \end{pmatrix} = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

$$D\mathbf{s}(t) = \begin{pmatrix} \frac{ds_1}{dt} \\ \frac{ds_2}{dt} \end{pmatrix} = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}$$

Kettenregel

Seien $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ und $g: \mathbb{R}^m \rightarrow \mathbb{R}^k$. Dann gilt

$$D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = D\mathbf{g}(\mathbf{f}(\mathbf{x})) \cdot D\mathbf{f}(\mathbf{x})$$

$$\mathbf{f}(x, y) = \begin{pmatrix} e^x \\ e^y \end{pmatrix} \quad \mathbf{g}(x, y) = \begin{pmatrix} x^2 + y^2 \\ x^2 - y^2 \end{pmatrix}$$

$$D\mathbf{f}(x, y) = \begin{pmatrix} e^x & 0 \\ 0 & e^y \end{pmatrix} \quad D\mathbf{g}(x, y) = \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix}$$

$$\begin{aligned} D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) &= D\mathbf{g}(\mathbf{f}(\mathbf{x})) \cdot D\mathbf{f}(\mathbf{x}) = \begin{pmatrix} 2e^x & 2e^y \\ 2e^x & -2e^y \end{pmatrix} \cdot \begin{pmatrix} e^x & 0 \\ 0 & e^y \end{pmatrix} \\ &= \begin{pmatrix} 2e^{2x} & 2e^{2y} \\ 2e^{2x} & -2e^{2y} \end{pmatrix} \end{aligned}$$

Beispiel – Richtungsableitung

Wir können die Richtungsableitung von $f: \mathbb{R}^n \rightarrow \mathbb{R}$ in Richtung \mathbf{h} (mit $\|\mathbf{h}\| = 1$) an der Stelle \mathbf{x}_0 auch so herleiten:

Sei $\mathbf{s}(t)$ ein Weg in Richtung \mathbf{h} :

$$\mathbf{s}: \mathbb{R} \rightarrow \mathbb{R}^n, t \mapsto \mathbf{x}_0 + t\mathbf{h}$$

Dann ist

$$\begin{aligned} Df(\mathbf{s}(0)) &= Df(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \\ D\mathbf{s}(0) &= \mathbf{h} \end{aligned}$$

und somit

$$\frac{\partial f}{\partial \mathbf{h}} = D(f \circ \mathbf{s})(0) = Df(\mathbf{s}(0)) \cdot D\mathbf{s}(0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{h}$$

Beispiel – Indirekte Abhängigkeit

Sei $f(x_1, x_2, t)$ wobei $x_1(t)$ und $x_2(t)$ ebenfalls von t abhängen.
Wie ändert sich f mit t ?

Kettenregel:

$$\text{Sei } \mathbf{x}: \mathbb{R} \rightarrow \mathbb{R}^3, t \mapsto \begin{pmatrix} x_1(t) \\ x_2(t) \\ t \end{pmatrix}$$

$$\begin{aligned} \frac{df}{dt} &= D(f \circ \mathbf{x})(t) = Df(\mathbf{x}(t)) \cdot D\mathbf{x}(t) \\ &= \nabla f(\mathbf{x}(t)) \cdot \begin{pmatrix} x_1'(t) \\ x_2'(t) \\ 1 \end{pmatrix} = (f_{x_1}(\mathbf{x}(t)), f_{x_2}(\mathbf{x}(t)), f_t(\mathbf{x}(t))) \cdot \begin{pmatrix} x_1'(t) \\ x_2'(t) \\ 1 \end{pmatrix} \\ &= f_{x_1}(\mathbf{x}(t)) \cdot x_1'(t) + f_{x_2}(\mathbf{x}(t)) \cdot x_2'(t) + f_t(\mathbf{x}(t)) \\ &= f_{x_1}(x_1, x_2, t) \cdot x_1'(t) + f_{x_2}(x_1, x_2, t) \cdot x_2'(t) + f_t(x_1, x_2, t) \end{aligned}$$

Zusammenfassung

- ▶ Differenzen- und Differentialquotient
- ▶ Differential
- ▶ Ableitung und höhere Ableitungen
- ▶ Monotonie
- ▶ Konkav und konvex
- ▶ Elastizität
- ▶ Partielle Ableitungen und Elastizität
- ▶ Gradient
- ▶ Richtungsableitung
- ▶ Totales Differential
- ▶ Jacobische Matrix
- ▶ Kettenregel

Kapitel 8

Inverse und implizite Funktionen

Inverse Funktion

Sei $f: D_f \subseteq \mathbb{R}^n \rightarrow W_f \subseteq \mathbb{R}^m$, $x \mapsto y = f(x)$. Eine Funktion

$$f^{-1}: W_f \rightarrow D_f, y \mapsto x = f^{-1}(y)$$

heißt **inverse Funktion** zu f , falls

$$f^{-1} \circ f = f \circ f^{-1} = \text{id}$$

id ist die **Einheitsfunktion** oder **identische Funktion**: $\text{id}(x) = x$

$$f^{-1}(f(x)) = f^{-1}(y) = x \quad \text{und} \quad f(f^{-1}(y)) = f(x) = y$$

f^{-1} existiert genau dann, wenn f bijektiv ist.

Wir erhalten $f^{-1}(y)$ als **eindeutige** Lösung x der Gleichung $y = f(x)$.

Lineare Funktion

Sei $f: \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto y = f(x) = ax + b$.

$$y = ax + b \quad \Leftrightarrow \quad ax = y - b \quad \Leftrightarrow \quad x = \frac{1}{a}y - \frac{b}{a}$$

Also:

$$f^{-1}(y) = a^{-1}y - a^{-1}b$$

Vorausgesetzt: $a \neq 0$ $[a = f'(x)]$

Beachte:

$$(f^{-1})'(y) = a^{-1} = \frac{1}{a} = \frac{1}{f'(x)}$$

Lineare Funktion

Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{x} \mapsto \mathbf{y} = f(\mathbf{x}) = \mathbf{A} \mathbf{x} + \mathbf{b}$ mit $m \times n$ -Matrix \mathbf{A} .

$$\mathbf{y} = \mathbf{A} \mathbf{x} + \mathbf{b} \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{A}^{-1} \mathbf{y} - \mathbf{A}^{-1} \mathbf{b}$$

Also:

$$f^{-1}(\mathbf{y}) = \mathbf{A}^{-1} \mathbf{y} - \mathbf{A}^{-1} \mathbf{b}$$

Vorausgesetzt: \mathbf{A} ist invertierbar. $[\mathbf{A} = Df(\mathbf{x})]$

(Insbesondere: $n = m$)

Beachte:

$$D(f^{-1})(\mathbf{y}) = \mathbf{A}^{-1} = (Df(\mathbf{x}))^{-1}$$

Lokal invertierbare Funktion

Die Funktion ist nicht bijektiv:

$$f: \mathbb{R} \rightarrow [0, \infty), x \mapsto f(x) = x^2$$

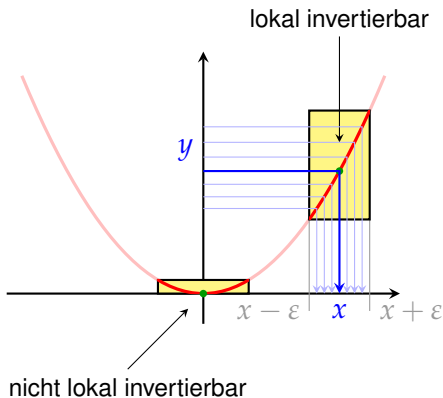
f^{-1} existiert daher nicht *global*.

Für manche x_0 gibt es ein *offenes* Intervall $(x_0 - \varepsilon, x_0 + \varepsilon)$ über dem $y = f(x)$ eindeutig nach x auflösbar ist.

Wir sagen:

f ist um x_0 **lokal invertierbar**.

Für manche x_0 gibt kein derartiges (noch so kleines) Intervall.



Existenz und Ableitung

1. Für welche x_0 ist f lokal invertierbar?
2. Wie lautet die Ableitung von f^{-1} an der Stelle $y_0 = f(x_0)$.

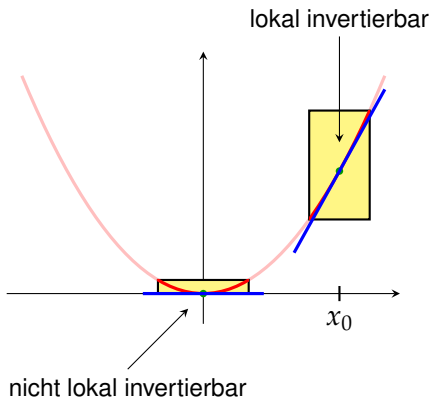
Idee:

Ersetze f durch das Differential:

$$f(x_0 + h) \approx f(x_0) + Df(x_0) \cdot h$$

Daher:

1. $Df(x_0)$ muss invertierbar sein.
2. $D(f^{-1})(y_0) = (Df(x_0))^{-1}$



Satz über inverse Funktionen

Sei $f: D_f \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion und x_0 ein Punkt mit $f'(x_0) \neq 0$.
Dann existiert ein Intervall \mathcal{U} um x_0 , sodass $f|_{\mathcal{U}}$ bijektiv in ein Intervall \mathcal{V} um $y_0 = f(x_0)$ abbildet.

Es existiert daher die inverse Funktion $f^{-1}: \mathcal{V} \rightarrow \mathcal{U}$.

Für die Ableitung gilt:

$$(f^{-1})'(y_0) = (f'(x_0))^{-1}$$

Seien $f: \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto y = f(x) = x^2$ und $x_0 = 3$, $y_0 = f(x_0) = 9$.
Da $f'(x_0) = 6 \neq 0$, ist f um $x_0 = 3$ lokal invertierbar und es gilt

$$(f^{-1})'(9) = \frac{1}{f'(3)} = \frac{1}{6}$$

Für $x_0 = 0$ liefert dieser Satz keine Aussage, da $f'(0) = 0$.

Satz über inverse Funktionen II

Seien $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ und \mathbf{x}_0 ein Punkt mit $|Df(\mathbf{x}_0)| \neq 0$.

Dann existiert ein Rechteck \mathcal{U} um \mathbf{x}_0 , sodass f \mathcal{U} bijektiv in ein Rechteck \mathcal{V} um $\mathbf{y}_0 = f(\mathbf{x}_0)$ abbildet.

Es existiert daher in \mathcal{V} die inverse Funktion $f^{-1}: \mathcal{V} \rightarrow \mathcal{U}$.

Für die Ableitung gilt:

$$D(f^{-1})(\mathbf{y}_0) = (Df(\mathbf{x}_0))^{-1}$$

Die Determinante der Jacobischen Matrix wird auch als **Funktionaldeterminante** bezeichnet. Notation:

$$\frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)} = |Df(\mathbf{x}_0)|$$

Beispiel

Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $\mathbf{x} \mapsto f(\mathbf{x}) = \begin{pmatrix} x_1^2 - x_2^2 \\ x_1 x_2 \end{pmatrix}$.

Dann gilt:

$$Df(\mathbf{x}) = \begin{pmatrix} 2x_1 & -2x_2 \\ x_2 & x_1 \end{pmatrix}$$

$$\frac{\partial(f_1, f_2)}{\partial(x_1, x_2)} = \begin{vmatrix} 2x_1 & -2x_2 \\ x_2 & x_1 \end{vmatrix} = 2x_1^2 + 2x_2^2 \neq 0 \quad \text{für alle } \mathbf{x} \neq 0.$$

f ist um alle $\mathbf{x}_0 \neq 0$ lokal invertierbar.

$$D(f^{-1})(f(1,1)) = (Df(1,1))^{-1} = \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{4} & \frac{2}{4} \\ -\frac{1}{4} & \frac{2}{4} \end{pmatrix}$$

f ist jedoch nicht bijektiv: $f(1,1) = f(-1,-1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Explizite und implizite Funktion

Der Zusammenhang zwischen zwei Variablen x und y kann gegeben werden durch eine

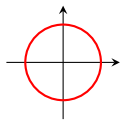
explizite Funktion:

$$y = f(x)$$

Beispiel:

$$y = x^2$$

existiert nicht



implizite Funktion:

$$F(x, y) = 0$$

Beispiel:

$$y - x^2 = 0$$

$$x^2 + y^2 - 1 = 0$$

Fragen:

- ▶ Wann kann man eine implizite Funktion (**lokal**) auch als explizite Funktion darstellen?
- ▶ Wie lautet die Ableitung von y nach der Variable x ?

Fall: Lineare Funktion

Im Falle einer linearen Funktion

$$F(x, y) = ax + by$$

sind beide Fragen leicht zu beantworten:

$$ax + by = 0 \quad \Rightarrow \quad y = -\frac{a}{b}x \quad (\text{falls } F_y = b \neq 0)$$

$$\frac{dy}{dx} = -\frac{a}{b} = -\frac{F_x}{F_y}$$

Fall: Allgemeine Funktion

Sei $F(x, y)$ eine Funktion und (x_0, y_0) ein Punkt mit $F(x_0, y_0) = 0$.

Wenn F nicht linear ist, dann können wir die Ableitung $\frac{dy}{dx}$ im Punkt x_0 berechnen, indem wir die Funktion *lokal* durch das totale Differential von F ersetzen.

$$dF = F_x dx + F_y dy = d0 = 0$$

Daraus erhalten wir

$$\frac{dy}{dx} = -\frac{F_x}{F_y}$$

Beispiel

Gesucht ist die implizite Ableitung $\frac{dy}{dx}$ von

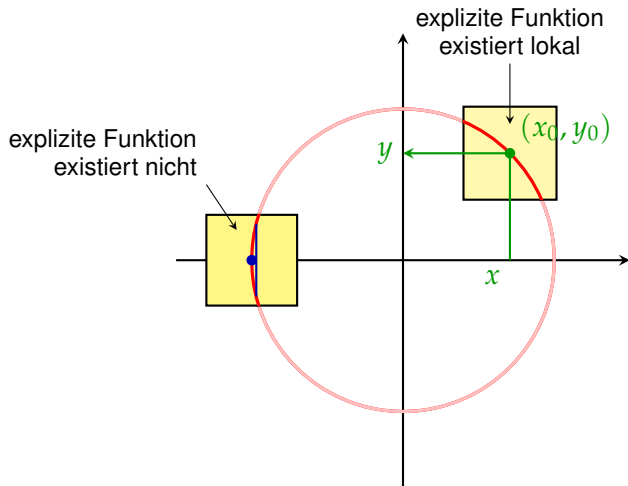
$$F(x, y) = x^2 + y^2 - 1 = 0$$

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{2x}{2y} = -\frac{x}{y}$$

Wir können auch die Ableitung von x nach der Variable y ausrechnen:

$$\frac{dx}{dy} = -\frac{F_y}{F_x} = -\frac{2y}{2x} = -\frac{y}{x}$$

Lokale Existenz einer expliziten Funktion



$y = f(x)$ existiert lokal, wenn $F_y \neq 0$.

Satz über implizite Funktionen

Sei $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ und sei (x_0, y_0) ein Punkt mit

$$F(x_0, y_0) = 0 \quad \text{und} \quad F_y(x_0, y_0) \neq 0.$$

Dann existiert ein Rechteck um (x_0, y_0) , sodass gilt:

- ▶ $F(x, y) = 0$ hat dort eine eindeutige Lösung $y = f(x)$, und

- ▶
$$\frac{dy}{dx} = -\frac{F_x}{F_y}$$

Sei $F(x, y) = x^2 + y^2 - 8$ und $(x_0, y_0) = (2, 2)$.

Da $F(x_0, y_0) = 0$ und $F_y(x_0, y_0) = 2y_0 = 4 \neq 0$, lässt sich y lokal als

Funktion von x darstellen und $\frac{dy}{dx}(x_0) = -\frac{F_x(x_0, y_0)}{F_y(x_0, y_0)} = -\frac{2x_0}{2y_0} = -1$.

Satz über implizite Funktionen II

Sei $F: \mathbb{R}^{n+1} \rightarrow \mathbb{R}$, $(\mathbf{x}, y) \mapsto F(\mathbf{x}, y) = F(x_1, \dots, x_n, y)$, und sei (\mathbf{x}_0, y_0) ein Punkt mit

$$F(\mathbf{x}_0, y_0) = 0 \quad \text{und} \quad F_y(\mathbf{x}_0, y_0) \neq 0.$$

Dann existiert ein Rechteck um (\mathbf{x}_0, y_0) , sodass gilt:

- ▶ $F(\mathbf{x}, y) = 0$ hat dort eine eindeutige Lösung $y = f(\mathbf{x})$, wobei $f: \mathbb{R}^n \rightarrow \mathbb{R}$, und

- ▶
$$\frac{\partial y}{\partial x_i} = - \frac{F_{x_i}}{F_y}$$

y ist die von uns ausgewählte Variable und muss nicht notwendigerweise an letzter Position stehen.

Beispiel

Gesucht ist $\frac{\partial x_2}{\partial x_3}$ der impliziten Funktion

$$F(x_1, x_2, x_3, x_4) = x_1^2 + x_2 x_3 + x_3^2 - x_3 x_4 - 1 = 0$$

an der Stelle $(x_1, x_2, x_3, x_4) = (1, 0, 1, 1)$.

Da $F(1, 0, 1, 1) = 0$ und $F_{x_2}(1, 0, 1, 1) = 1 \neq 0$ können wir x_2 lokal als Funktion von (x_1, x_3, x_4) darstellen: $x_2 = f(x_1, x_3, x_4)$.

Für die partielle Ableitung nach x_3 erhalten wir

$$\frac{\partial x_2}{\partial x_3} = -\frac{F_{x_3}}{F_{x_2}} = -\frac{x_2 + 2x_3 - x_4}{x_3} = -1$$

An den Stellen $(1, 1, 1, 1)$ und $(1, 1, 0, 1)$ kann der Satz über implizite Funktionen nicht für x_2 angewendet werden:

$$F(1, 1, 1, 1) \neq 0 \text{ und } F_{x_2}(1, 1, 0, 1) = 0.$$

Jacobische Matrix

Sei

$$\mathbf{F}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} F_1(x_1, \dots, x_n, y_1, \dots, y_m) \\ \vdots \\ F_m(x_1, \dots, x_n, y_1, \dots, y_m) \end{pmatrix} = 0$$

dann heißt

$$\frac{\partial \mathbf{F}(\mathbf{x}, \mathbf{y})}{\partial \mathbf{y}} = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}$$

Jacobische Matrix von $\mathbf{F}(\mathbf{x}, \mathbf{y})$ bezüglich \mathbf{y} .

Analog: $\frac{\partial \mathbf{F}(\mathbf{x}, \mathbf{y})}{\partial \mathbf{x}}$

Satz über implizite Funktionen III

Sei $F: \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$,

$$(\mathbf{x}, \mathbf{y}) \mapsto F(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} F_1(x_1, \dots, x_n, y_1, \dots, y_m) \\ \vdots \\ F_m(x_1, \dots, x_n, y_1, \dots, y_m) \end{pmatrix}$$

und sei $(\mathbf{x}_0, \mathbf{y}_0)$ ein Punkt mit

$$F(\mathbf{x}_0, \mathbf{y}_0) = 0 \quad \text{und} \quad \left| \frac{\partial F(\mathbf{x}, \mathbf{y})}{\partial \mathbf{y}} \right| \neq 0 \quad \text{für} \quad (\mathbf{x}, \mathbf{y}) = (\mathbf{x}_0, \mathbf{y}_0).$$

Dann existiert ein Rechteck um $(\mathbf{x}_0, \mathbf{y}_0)$, sodass gilt:

- ▶ $F(\mathbf{x}, \mathbf{y}) = 0$ hat dort eine eindeutige Lösung $\mathbf{y} = \mathbf{f}(\mathbf{x})$, wobei $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, und

- ▶
$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = - \left(\frac{\partial F}{\partial \mathbf{y}} \right)^{-1} \cdot \left(\frac{\partial F}{\partial \mathbf{x}} \right)$$

Beispiel

$$\text{Sei } \mathbf{F}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} F_1(x_1, x_2, y_1, y_2) \\ F_2(x_1, x_2, y_1, y_2) \end{pmatrix} = \begin{pmatrix} x_1^2 + x_2^2 - y_1^2 - y_2^2 + 3 \\ x_1^3 + x_2^3 + y_1^3 + y_2^3 - 11 \end{pmatrix}$$

und $(\mathbf{x}_0, \mathbf{y}_0) = (1, 1, 1, 2)$.

$$\frac{\partial \mathbf{F}}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1 & 2x_2 \\ 3x_1^2 & 3x_2^2 \end{pmatrix} \quad \frac{\partial \mathbf{F}}{\partial \mathbf{x}}(1, 1, 1, 2) = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}$$

$$\frac{\partial \mathbf{F}}{\partial \mathbf{y}} = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2} \\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2} \end{pmatrix} = \begin{pmatrix} -2y_1 & -2y_2 \\ 3y_1^2 & 3y_2^2 \end{pmatrix} \quad \frac{\partial \mathbf{F}}{\partial \mathbf{y}}(1, 1, 1, 2) = \begin{pmatrix} -2 & -4 \\ 3 & 12 \end{pmatrix}$$

Da $\mathbf{F}(1, 1, 1, 2) = 0$ und $\left| \frac{\partial \mathbf{F}(\mathbf{x}, \mathbf{y})}{\partial \mathbf{y}} \right| = -12 \neq 0$ können wir den Satz über implizite Funktionen anwenden und es gilt

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = - \left(\frac{\partial \mathbf{F}}{\partial \mathbf{y}} \right)^{-1} \cdot \left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}} \right) = -\frac{1}{-12} \begin{pmatrix} 12 & 4 \\ -3 & -2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 0 & 0 \end{pmatrix}$$

Zusammenfassung

- ▶ Lokale Existenz einer inversen Funktion
- ▶ Ableitung einer inversen Funktion
- ▶ Explizite und implizite Funktionen
- ▶ Lokale explizite Darstellung einer impliziten Funktion
- ▶ Ableitung einer impliziten Funktion

Kapitel 9

Taylorreihen

Näherung erster Ordnung

Wir wollen eine Funktion f durch möglichst einfache Funktionen approximieren.

Approximation durch eine **lineare Funktion**:

$$f(x) \doteq f(x_0) + f'(x_0)(x - x_0)$$

\doteq bedeutet „**in erster Näherung**“.

Wenn wir die Näherung verwenden, rechnen wir mit der Tangente der Funktion an der Stelle x_0 anstatt mit f .

Polynome

Besser Approximationen erhalten wir durch Verwendung eines **Polynoms** $P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$.

Ansatz:

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n + R_n(x)$$

Dabei wird $R_n(x)$ als **Restglied** bezeichnet. Es gibt den Fehler an, den wir beim Ersetzen von $f(x)$ durch $P_n(x)$ machen.

Wählen dabei die Koeffizienten a_i so, dass die ersten n Ableitungen von f und P_n an der Stelle $x_0 = 0$ übereinstimmen.

Ableitungen

$$f(x) = a_0 + a_1x + \cdots + a_nx^n = P_n(x)$$

$$\Rightarrow f(0) = a_0$$

$$f'(x) = a_1 + 2 \cdot a_2x + \cdots + n \cdot a_nx^{n-1} = P'_n(x)$$

$$\Rightarrow f'(0) = a_1$$

$$f''(x) = 2 \cdot a_2 + 3 \cdot 2 \cdot a_3x + \cdots + n \cdot (n-1) \cdot a_nx^{n-2} = P''_n(x)$$

$$\Rightarrow f''(0) = 2a_2$$

$$f'''(x) = 3 \cdot 2 \cdot a_3 + \cdots + n \cdot (n-1) \cdot (n-2) \cdot a_nx^{n-3} = P'''_n(x)$$

$$\Rightarrow f'''(0) = 3! a_3$$

⋮

$$f^{(n)}(x) = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 1 \cdot a_n = P_n^{(n)}(x)$$

$$\Rightarrow f^{(n)}(0) = n! a_n$$

MacLaurinpolynom

Die Koeffizienten des Polynoms lauten daher

$$a_k = \frac{f^{(k)}(0)}{k!}$$

$f^{(k)}(x_0)$ bezeichnet dabei die k -te Ableitung von f an der Stelle x_0 .

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k + R_n(x)$$

Das Polynom

$$f(0) + f'(0) x + \frac{f''(0)}{2!} x^2 + \dots + \frac{f^{(n)}(0)}{n!} x^n$$

heißt das **MacLaurinpolynom** n -ter Ordnung von f .

Taylorpolynom

Wenn wir die Ableitungen an einer beliebigen Stelle x_0 betrachten erhalten wir das **Taylorpolynom** n -ter Ordnung von f im Punkt x_0 :

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + R_n(x)$$

Die unendliche Reihe ($n \rightarrow \infty$) heißt die **Taylorreihe** von f .

Wenn $\lim_{n \rightarrow \infty} R_n(x) = 0$, dann konvergiert die Taylorreihe gegen $f(x)$.

Wir sprechen dann von der **Taylorreihenentwicklung** von f mit **Entwicklungspunkt** x_0 .

Beispiel – Exponentialfunktion

Taylorreihenentwicklung von $f(x) = e^x$ an der Stelle $x_0 = 0$:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + R_n(x)$$

$$f(x) = e^x \Rightarrow f(0) = 1$$

$$f'(x) = e^x \Rightarrow f'(0) = 1$$

$$f''(x) = e^x \Rightarrow f''(0) = 1$$

$$f'''(x) = e^x \Rightarrow f'''(0) = 1$$

\vdots

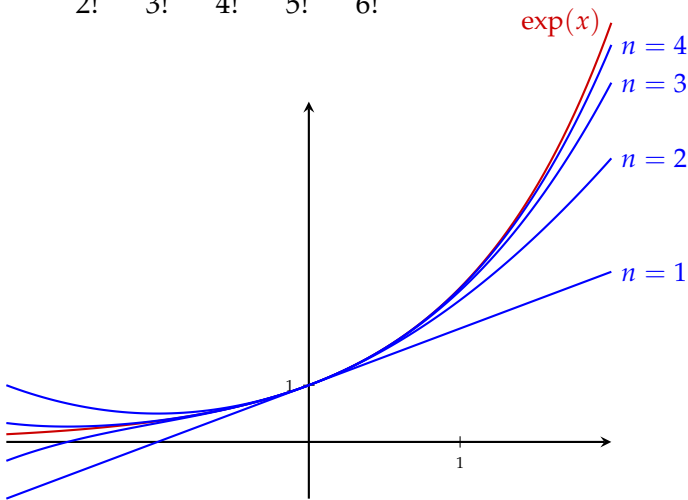
$$f^{(n)}(x) = e^x \Rightarrow f^{(n)}(0) = 1$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots$$

Diese Taylorreihe konvergiert für alle $x \in \mathbb{R}$.

Beispiel – Exponentialfunktion

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \dots$$



Beispiel – Logarithmus

Taylorreihenentwicklung von $f(x) = \ln(1+x)$ an der Stelle $x_0 = 0$:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \dots + \frac{f^{(n)}(0)}{n!}x^n + R_n(x)$$

$$f(x) = \ln(1+x) \quad \Rightarrow \quad f(0) = 0$$

$$f'(x) = (1+x)^{-1} \quad \Rightarrow \quad f'(0) = 1$$

$$f''(x) = -1 \cdot (1+x)^{-2} \quad \Rightarrow \quad f''(0) = -1$$

$$f'''(x) = 2 \cdot 1 \cdot (1+x)^{-3} \quad \Rightarrow \quad f'''(0) = 2!$$

\vdots

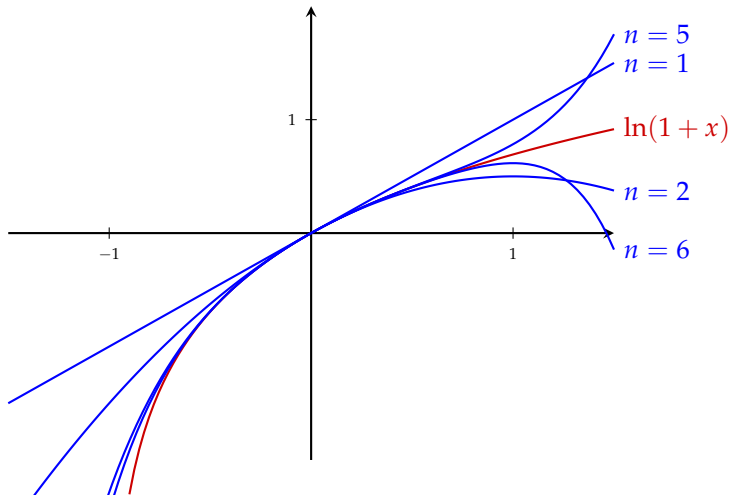
$$f^{(n)}(x) = (-1)^{n-1}(n-1)!(1+x)^{-n+1} \Rightarrow f^{(n)}(0) = (-1)^{n-1}(n-1)!$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots + (-1)^{n-1} \frac{x^n}{n} + \dots$$

Diese Taylorreihe konvergiert für alle $x \in (-1,1)$.

Beispiel – Logarithmus

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \frac{x^6}{6} + \dots$$



Konvergenzradius

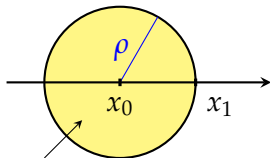
Es gibt Taylorreihen, die nicht für alle $x \in \mathbb{R}$ konvergieren.

z.B.: $\ln(1 + x)$

Es gilt jedoch:

Falls eine Taylorreihe für ein x_1 mit $|x_1 - x_0| = \rho$ konvergiert, so konvergiert sie für alle x mit $|x - x_0| < \rho$.

Das größtmögliche derartige ρ heißt der **Konvergenzradius** der Taylorreihe.



Taylorreihe konvergiert

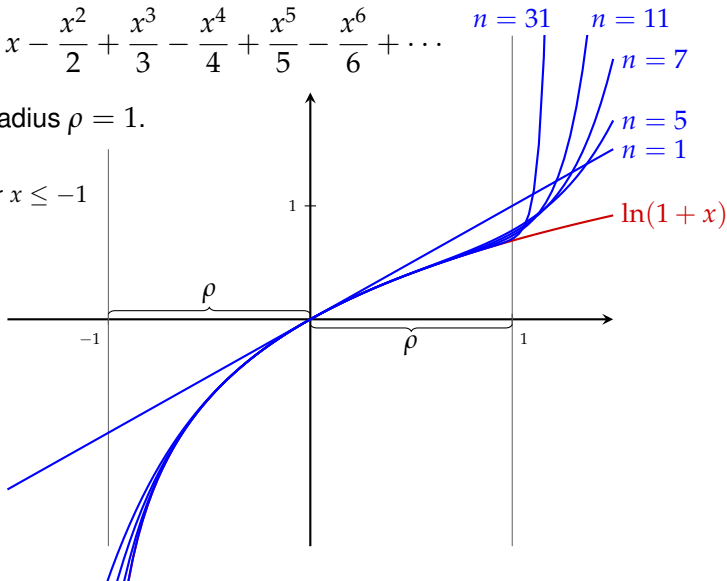
Beispiel – Logarithmus

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \frac{x^6}{6} + \dots$$

Konvergenzradius $\rho = 1$.

Indiz:

$\ln(1+x)$ ist für $x \leq -1$
nicht definiert.



Approximationsfehler

Das Restglied gibt den Fehler bei der Approximation durch die Taylorreihe an.

$$\text{Fehler} = |R_n(x)| = \left| f(x) - \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \right|$$

Der Approximationsfehler $|R_n(x)|$ ist umso kleiner

- ▶ je näher x am Entwicklungspunkt x_0 ist;
- ▶ je größer die Ordnung n ist.

Falls die Taylorreihe konvergiert, dann gilt

$$R_n(x) = O((x - x_0)^{n+1}) \quad \text{für } x \rightarrow x_0$$

Wir sagen: der Fehler geht mit groß O von x^{n+1} gegen 0.

Landau-Symbol

Seien $f(x)$ und $g(x)$ zwei Funktionen.

Wir schreiben

$$f(x) = O(g(x)) \quad \text{für } x \rightarrow x_0$$

falls eine Konstante C existiert, sodass

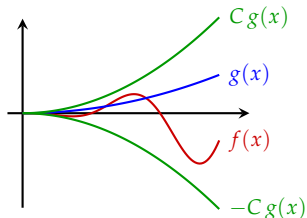
$$|f(x)| < C \cdot |g(x)|$$

für alle x mit $|x - x_0| < \varepsilon$.

$O(\cdot)$ heißt **Landau-Symbol** („groß O von ...“).

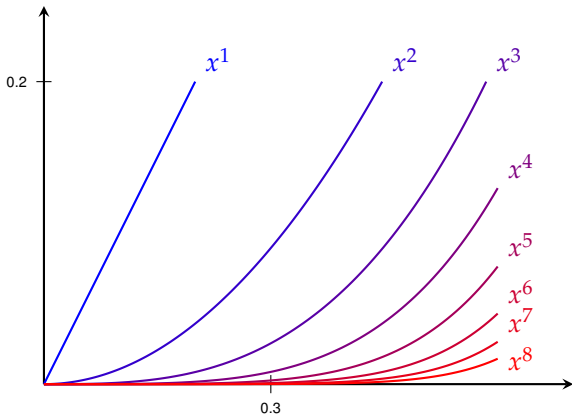
Man schreibt daher auch

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + O((x - x_0)^{n+1})$$



Einfluss der Ordnung der Potenzen

Je höher die Ordnung einer Potenz wird, desto kleiner wird ihr Beitrag in der Taylorreihenentwicklung in der Nähe des Entwicklungspunkt.



Wichtige Taylorreihen

$f(x)$	MacLaurinreihe	ρ
$\exp(x)$	$= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots$	∞
$\ln(1+x)$	$= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$	1
$\sin(x)$	$= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$	∞
$\cos(x)$	$= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$	∞
$\frac{1}{1-x}$	$= 1 + x + x^2 + x^3 + x^4 + \dots$	1

Rechnen mit Taylorreihen

Wir können Taylorreihen bequem

- ▶ addieren (gliedweise)
- ▶ differenzieren (gliedweise)
- ▶ integrieren (gliedweise)
- ▶ multiplizieren
- ▶ dividieren
- ▶ substituieren

Man verwendet daher oft Taylorreihen zur *Definition* einer Funktion.

Etwa

$$\exp(x) := 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Beispiel

Die erste Ableitung von $\exp(x)$ erhalten wir durch Ableiten der Taylorreihe:

$$\begin{aligned}(\exp(x))' &= \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots\right)' \\ &= 0 + 1 + \frac{2x}{2!} + \frac{3x^2}{3!} + \frac{4x^3}{4!} + \dots \\ &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \\ &= \exp(x)\end{aligned}$$

Beispiel

Wir erhalten die MacLaurinreihe von $f(x) = x^2 \cdot e^x$ durch Multiplizieren der MacLaurinreihe von x^2 mit der MacLaurinreihe von $\exp(x)$:

$$\begin{aligned}x^2 \cdot e^x &= x^2 \cdot \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots\right) \\ &= x^2 + x^3 + \frac{x^4}{2!} + \frac{x^5}{3!} + \frac{x^6}{4!} + \dots\end{aligned}$$

Wir erhalten die MacLaurinreihe von $f(x) = \exp(-x^2)$ durch Substituieren (Einsetzen) von $-x^2$ in die MacLaurinreihe von $\exp(x)$:

$$\begin{aligned}e^u &= 1 + u + \frac{u^2}{2!} + \frac{u^3}{3!} + \frac{u^4}{4!} + \dots \\ e^{-x^2} &= 1 + (-x^2) + \frac{(-x^2)^2}{2!} + \frac{(-x^2)^3}{3!} + \frac{(-x^2)^4}{4!} + \dots \\ &= 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!} - \dots\end{aligned}$$

Polynome

Die Idee der Taylorreihe kann auch für Funktionen in zwei oder mehreren Variablen verwirklicht werden.

Ein Polynom n -ten Grades in zwei Variablen hat die allgemeine Form

$$\begin{aligned} P_n(x_1, x_2) = & a_0 \\ & + a_{10} x_1 + a_{11} x_2 \\ & + a_{20} x_1^2 + a_{21} x_1 x_2 + a_{22} x_2^2 \\ & + a_{30} x_1^3 + a_{31} x_1^2 x_2 + a_{32} x_1 x_2^2 + a_{33} x_2^3 \\ & \vdots \\ & + a_{n0} x_1^n + a_{n1} x_1^{n-1} x_2 + a_{n2} x_1^{n-2} x_2^2 + \cdots + a_{nn} x_2^n \end{aligned}$$

Wählen die Koeffizienten a_{kj} so, dass alle partiellen Ableitungen von f und P_n bis zur Ordnung n am Entwicklungspunkt \mathbf{x}_0 übereinstimmen.

Taylorpolynom 2. Ordnung

Wir erhalten dadurch die Koeffizienten

$$a_{kj} = \frac{1}{k!} \binom{k}{j} \frac{\partial^k f(0)}{(\partial x_1)^{k-j} (\partial x_2)^j} \quad k \in \mathbb{N}, j = 0, \dots, k$$

Das Taylorpolynom zweiter Ordnung an der Stelle $\mathbf{x}_0 = 0$ lautet daher

$$\begin{aligned} f(\mathbf{x}) = & f(0) \\ & + f_{x_1}(0) x_1 + f_{x_2}(0) x_2 \\ & + \frac{1}{2} f_{x_1 x_1}(0) x_1^2 + f_{x_1 x_2}(0) x_1 x_2 + \frac{1}{2} f_{x_2 x_2}(0) x_2^2 + \dots \end{aligned}$$

Der lineare Term kann mittels Gradient dargestellt werden:

$$f_{x_1}(0) x_1 + f_{x_2}(0) x_2 = \nabla f(0) \cdot \mathbf{x}$$

Was ist mit dem quadratischen Term?

Hesse-Matrix

Wir fassen alle zweiten partiellen Ableitungen von f an der Stelle \mathbf{x}_0 zu einer 2×2 -Matrix zusammen.

$$\mathbf{H}_f(\mathbf{x}_0) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}_0) & f_{x_1x_2}(\mathbf{x}_0) \\ f_{x_2x_1}(\mathbf{x}_0) & f_{x_2x_2}(\mathbf{x}_0) \end{pmatrix}$$

Diese Matrix wird als **Hesse-Matrix** von f an der Stelle \mathbf{x}_0 bezeichnet.

Der quadratische Term kann mittels Hesse-Matrix dargestellt werden:

$$f_{x_1x_1}(0) x_1^2 + 2 f_{x_1x_2}(0) x_1 x_2 + f_{x_2x_2}(0) x_2^2 = \mathbf{x}^t \cdot \mathbf{H}_f(0) \cdot \mathbf{x}$$

Also

$$f(\mathbf{x}) = f(0) + \nabla f(0) \cdot \mathbf{x} + \frac{1}{2} \mathbf{x}^t \cdot \mathbf{H}_f(0) \cdot \mathbf{x} + O(\|\mathbf{x}\|^3)$$

Hesse-Matrix II

Allgemein fasst die **Hesse-Matrix** alle zweiten partiellen Ableitungen einer Funktion f in n Variablen an der Stelle \mathbf{x}_0 zu einer $n \times n$ -Matrix zusammen.

$$\mathbf{H}_f(\mathbf{x}_0) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}_0) & f_{x_1x_2}(\mathbf{x}_0) & \cdots & f_{x_1x_n}(\mathbf{x}_0) \\ f_{x_2x_1}(\mathbf{x}_0) & f_{x_2x_2}(\mathbf{x}_0) & \cdots & f_{x_2x_n}(\mathbf{x}_0) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{x}_0) & f_{x_nx_2}(\mathbf{x}_0) & \cdots & f_{x_nx_n}(\mathbf{x}_0) \end{pmatrix}$$

- ▶ Die Hesse-Matrix ist symmetrisch.
(Falls f zweimal stetig differenzierbar ist.)
- ▶ Die Hesse-Matrix “spielt“ die gleiche Rolle wie die zweite Ableitung bei Funktionen in einer Variablen.
- ▶ Andere Notation: $f''(\mathbf{x}_0)$

Taylorpolynom 2. Ordnung (II)

Taylorreihenentwicklung von f zweiter Ordnung an der Stelle \mathbf{x}_0 :

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2} \mathbf{h}^t \cdot \mathbf{H}_f(\mathbf{x}_0) \cdot \mathbf{h} + O(\|\mathbf{h}\|^3)$$

In anderer Notation sieht das ganze analog zur Taylorreihe einer Funktion in einer Variable aus:

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2} \mathbf{h}^t \cdot f''(\mathbf{x}_0) \cdot \mathbf{h} + O(\|\mathbf{h}\|^3)$$

Beispiel

Wir suchen das Taylorpolynom 2. Ordnung an der Stelle $\mathbf{x}_0 = 0$ von

$$f(x, y) = e^{x^2 - y^2} + x$$

$$\begin{aligned} f(x, y) &= e^{x^2 - y^2} + x && \Rightarrow f(0,0) = 1 \\ f_x(x, y) &= 2x e^{x^2 - y^2} + 1 && \Rightarrow f_x(0,0) = 1 \\ f_y(x, y) &= -2y e^{x^2 - y^2} && \Rightarrow f_y(0,0) = 0 \\ f_{xx}(x, y) &= 2 e^{x^2 - y^2} + 4x^2 e^{x^2 - y^2} && \Rightarrow f_{xx}(0,0) = 2 \\ f_{xy}(x, y) &= -4xy e^{x^2 - y^2} && \Rightarrow f_{xy}(0,0) = 0 \\ f_{yy}(x, y) &= -2 e^{x^2 - y^2} + 4y^2 e^{x^2 - y^2} && \Rightarrow f_{yy}(0,0) = -2 \end{aligned}$$

Gradient:

$$\nabla f(0) = (1, 0)$$

Hesse-Matrix:

$$\mathbf{H}_f(0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

Beispiel

Das Taylorpolynom lautet daher

$$\begin{aligned}f(x, y) &\approx f(0) + \nabla f(0) \cdot \mathbf{x} + \frac{1}{2} \mathbf{x}^t \cdot \mathbf{H}_f(0) \cdot \mathbf{x} \\&= 1 + (1, 0) \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2} (x, y) \cdot \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \\&= 1 + x + x^2 - y^2\end{aligned}$$

Zusammenfassung

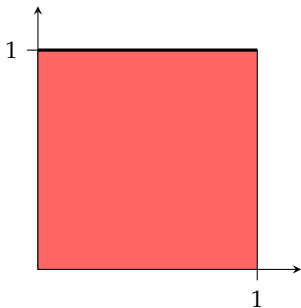
- ▶ MacLaurin- und Taylorpolynom
- ▶ Taylorreihenentwicklung
- ▶ Konvergenzradius
- ▶ Rechnen mit Taylorreihen
- ▶ Taylorreihen von Funktionen in mehreren Variablen
- ▶ Hesse-Matrix

Kapitel 10

Integration

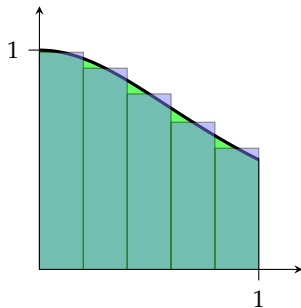
Flächeninhalt

Berechnen Sie die Inhalte der angegebenen Flächen!



$$f(x) = 1$$

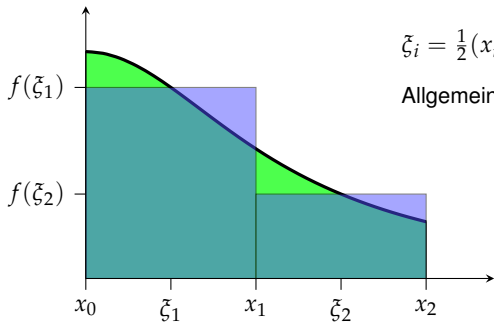
Fläche: $A = 1$



$$f(x) = \frac{1}{1+x^2}$$

Approximation
durch Treppenfunktion

Riemann-Summen

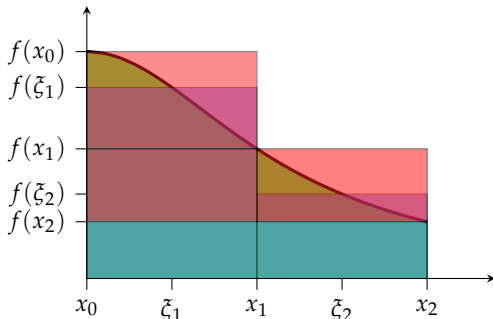


$$\xi_i = \frac{1}{2}(x_{i-1} + x_i)$$

Allgemeiner: $\xi_i \in (x_i, x_{i-1})$

$$A = \int_a^b f(x) dx \approx \sum_{i=1}^n f(\xi_i) \cdot (x_i - x_{i-1})$$

Approximationsfehler



$$\left| \int_a^b f(x) dx - \sum_{i=1}^n f(\xi_i) \cdot (x_i - x_{i-1}) \right| \leq (f_{\max} - f_{\min}) (b - a) \frac{1}{n} \rightarrow 0$$

Annahme: Funktion monoton; x_0, x_1, \dots, x_n äquidistant

Riemann-Integral

Falls die **Riemann-Summen**

$$I_n = \sum_{i=1}^n f(\xi_i) \cdot (x_i - x_{i-1})$$

eine konvergente Folge bilden, dann heißt deren Grenzwert das

Riemann-Integral von f . Notation: $\int_a^b f(x) dx$

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i) \cdot (x_i - x_{i-1})$$

Alle für uns relevanten Funktionen besitzen ein Riemann-Integral.

Riemann-Integral und Erwartungswert

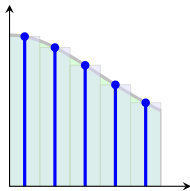
Angenommen eine stetige Zufallsvariable X mit Dichte f kann nur als diskrete Zufallsvariable D mit Wahrscheinlichkeitsfunktion w beobachtet werden. (**Klassifizieren** in n Klassen)

Sei $\xi_i \in (x_{i-1}, x_i)$ Klassenrepräsentant mit Wahrscheinlichkeit

$$w(\xi_i) = P(D = \xi_i) = P(x_{i-1} < X \leq x_i) \approx f(\xi_i) \cdot (x_i - x_{i-1})$$

Für den Erwartungswert von D gilt dann

$$E[D] \approx \sum_{i=1}^n \xi_i \cdot f(\xi_i) \cdot (x_i - x_{i-1}) \approx \int_a^b x \cdot f(x) dx = E[X]$$



Riemann-Integral – Eigenschaften

$$\int_a^b (\alpha f(x) + \beta g(x)) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx$$

$$\int_a^b f(x) dx = - \int_b^a f(x) dx$$

$$\int_a^a f(x) dx = 0$$

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx$$

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx \quad \text{falls } f(x) \leq g(x) \text{ für alle } x \in [a, b]$$

Stammfunktion

Eine Funktion $F(x)$ heißt **Stammfunktion** einer Funktion $f(x)$, falls

$$F'(x) = f(x)$$

Berechnung:

Vermuten und Verifizieren

Beispiel: Wir suchen die Stammfunktion von $f(x) = \ln(x)$.

Vermuten: $F(x) = x(\ln(x) - 1)$

Verifizieren: $F'(x) = (x(\ln(x) - 1))' =$
 $= 1 \cdot (\ln(x) - 1) + x \cdot \frac{1}{x} = \ln(x)$

Aber auch: $F(x) = x(\ln(x) - 1) + 5$

Stammfunktion

Die Stammfunktion wird mit dem Symbol

$$\int f(x) dx + c$$

bezeichnet und wird meist als das **unbestimmte Integral** der Funktion f bezeichnet. Die Zahl c heißt **Integrationskonstante**.

Für das Suchen von Stammfunktionen gibt es keine „Kochrezepte“ (sondern nur Werkzeuge, die man durchprobieren kann).

Es gibt Funktionen, deren Stammfunktionen sich nicht durch elementare Funktionen ausdrücken lassen.

E.g., die Stammfunktion von $\exp(-\frac{1}{2}x^2)$.

Grundintegrale

Zur Erleichterung gibt es Tabellen mit bekannten Stammfunktionen, sogenannten **Grundintegralen**:

$f(x)$	$\int f(x) dx$
0	c
x^a	$\frac{1}{a+1} \cdot x^{a+1} + c$
e^x	$e^x + c$
$\frac{1}{x}$	$\ln x + c$
$\cos(x)$	$\sin(x) + c$
$\sin(x)$	$-\cos(x) + c$

Integrationsverfahren

► Summenregel

$$\int \alpha f(x) + \beta g(x) dx = \alpha \int f(x) dx + \beta \int g(x) dx$$

► Partielles Integrieren

$$\int f \cdot g' dx = f \cdot g - \int f' \cdot g dx$$

► Substitution

$$\int f(g(x)) \cdot g'(x) dx = \int f(z) dz$$

mit $z = g(x)$ und $dz = g'(x) dx$

Beispiel – Summenregel

Stammfunktion von $f(x) = 4x^3 - x^2 + 3x - 5$.

$$\begin{aligned}\int f(x) dx &= \int 4x^3 - x^2 + 3x - 5 dx \\ &= 4 \int x^3 dx - \int x^2 dx + 3 \int x dx - 5 \int dx \\ &= 4 \frac{1}{4} x^4 - \frac{1}{3} x^3 + 3 \frac{1}{2} x^2 - 5x + c \\ &= x^4 - \frac{1}{3} x^3 + \frac{3}{2} x^2 - 5x + c\end{aligned}$$

Beispiel – Partielles Integrieren

Stammfunktion von $f(x) = x \cdot e^x$.

$$\int \underbrace{x}_f \cdot \underbrace{e^x}_{g'} dx = \underbrace{x}_f \cdot \underbrace{e^x}_g - \int \underbrace{1}_{f'} \cdot \underbrace{e^x}_g dx = x \cdot e^x - e^x + c$$

$$f = x \quad \Rightarrow \quad f' = 1$$

$$g' = e^x \quad \Rightarrow \quad g = e^x$$

Beispiel – Partielles Integrieren

Stammfunktion von $f(x) = x^2 \cos(x)$.

$$\int \underbrace{x^2}_f \cdot \underbrace{\cos(x)}_{g'} dx = \underbrace{x^2}_f \cdot \underbrace{\sin(x)}_g - \int \underbrace{2x}_{f'} \cdot \underbrace{\sin(x)}_g dx$$

Partielles Integrieren des zweiten Terms ergibt:

$$\begin{aligned} \int \underbrace{2x}_f \cdot \underbrace{\sin(x)}_{g'} dx &= \underbrace{2x}_f \cdot \underbrace{(-\cos(x))}_g - \int \underbrace{2}_{f'} \cdot \underbrace{(-\cos(x))}_g dx \\ &= -2x \cdot \cos(x) - 2 \cdot (-\sin(x)) + c \end{aligned}$$

Die Stammfunktion von f lautet daher:

$$\int x^2 \cos(x) dx = x^2 \sin(x) + 2x \cos(x) - 2 \sin(x) + c$$

Beispiel – Substitution

Stammfunktion von $f(x) = 2x \cdot e^{x^2}$.

$$\int \underbrace{\exp(x^2)}_{g(x)} \cdot \underbrace{2x}_{g'(x)} dx = \int \exp(z) dz = e^z + c = e^{x^2} + c$$

$$z = g(x) = x^2 \quad \Rightarrow \quad dz = g'(x) dx = 2x dx$$

Integrationsverfahren – Herleitung

Partielles Integrieren erhält man aus der Produktregel für das Differenzieren:

$$\begin{aligned} f(x) \cdot g(x) &= \int (f(x) \cdot g(x))' dx = \int (f'(x) g(x) + f(x) g'(x)) dx \\ &= \int f'(x) g(x) dx + \int f(x) g'(x) dx \end{aligned}$$

Substitution folgt aus der Kettenregel:

Sei F eine Stammfunktion von f und $z = g(x)$. Dann gilt

$$\begin{aligned} \int f(z) dz &= F(z) = F(g(x)) = \int (F(g(x)))' dx \\ &= \int F'(g(x)) g'(x) dx = \int f(g(x)) g'(x) dx \end{aligned}$$

Der Hauptsatz der Differential- und Integralrechnung

Sei $F(x)$ eine Stammfunktion einer *stetigen* Funktion $f(x)$,
dann gilt für das bestimmte **Integral**

$$\int_a^b f(x) dx = F(x) \Big|_a^b = F(b) - F(a)$$

Dieser Satz erlaubt es uns, Integrale einfach mittels Stammfunktionen auszurechnen! (**Bestimmtes Integral**)

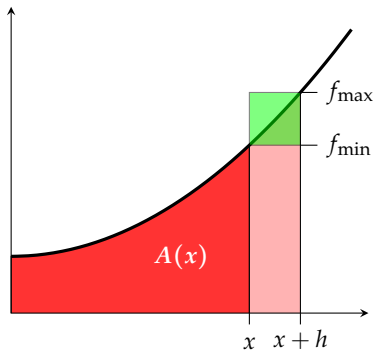
Beispiel:

Wir suchen das Integral der Funktion $f(x) = x^2$ im Intervall $[0,1]$.

$$\int_0^1 x^2 dx = \frac{1}{3} x^3 \Big|_0^1 = \frac{1}{3} \cdot 1^3 - \frac{1}{3} \cdot 0^3 = \frac{1}{3}$$

Der Hauptsatz / Beweisidee

Sei $A(x)$ die Fläche zwischen dem Graphen einer stetigen Funktion f und der x -Achse zwischen 0 und x .



$$f_{\min} \cdot h \leq A(x+h) - A(x) \leq f_{\max} \cdot h$$

$$f_{\min} \leq \frac{A(x+h) - A(x)}{h} \leq f_{\max}$$

Grenzübergang $h \rightarrow 0$: ($\lim_{h \rightarrow 0} f_{\min} = f(x)$)

$$f(x) \leq \underbrace{\lim_{h \rightarrow 0} \frac{A(x+h) - A(x)}{h}}_{=A'(x)} \leq f(x)$$

$$A'(x) = f(x)$$

d.h. $A(x)$ ist eine Stammfunktion von $f(x)$.

Integrationsverfahren / (2)

► Summenregel

$$\int_a^b \alpha f(x) + \beta g(x) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx$$

► Partielles Integrieren

$$\int_a^b f \cdot g' dx = f \cdot g \Big|_a^b - \int_a^b f' \cdot g dx$$

► Substitution

$$\int_a^b f(g(x)) \cdot g'(x) dx = \int_{g(a)}^{g(b)} f(z) dz$$

mit $z = g(x)$ und $dz = g'(x) dx$

Beispiel

Berechne das bestimmte Integral $\int_e^{10} \frac{1}{\ln(x)} \cdot \frac{1}{x} dx$.

$$\int_e^{10} \frac{1}{\ln(x)} \cdot \frac{1}{x} dx = \int_1^{\ln(10)} \frac{1}{z} dz =$$

$$z = \ln(x) \quad \Rightarrow \quad dz = \frac{1}{x} dx$$

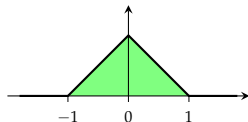
$$= \ln(z) \Big|_1^{\ln(10)} =$$

$$= \ln(\ln(10)) - \ln(1) \approx 0,834$$

Beispiel

Gesucht ist $\int_{-2}^2 f(x) dx$ für die Funktion

$$f(x) = \begin{cases} 1+x & \text{für } -1 \leq x < 0 \\ 1-x & \text{für } 0 \leq x < 1 \\ 0 & \text{für } x < -1 \text{ und } x \geq 1 \end{cases}$$



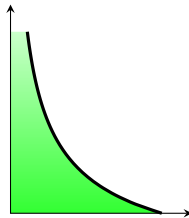
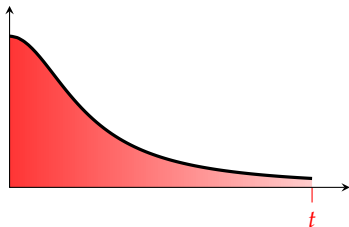
Es gilt

$$\begin{aligned} \int_{-2}^2 f(x) dx &= \int_{-2}^{-1} f(x) dx + \int_{-1}^0 f(x) dx + \int_0^1 f(x) dx + \int_1^2 f(x) dx \\ &= \int_{-2}^{-1} 0 dx + \int_{-1}^0 (1+x) dx + \int_0^1 (1-x) dx + \int_1^2 0 dx \\ &= \left(x + \frac{1}{2}x^2\right) \Big|_{-1}^0 + \left(x - \frac{1}{2}x^2\right) \Big|_0^1 \\ &= \frac{1}{2} + \frac{1}{2} = 1 \end{aligned}$$

Uneigentliches Integral

Uneigentliche Integrale sind Integrale, bei denen

- ▶ das Integrationsintervall unbeschränkt ist, oder
- ▶ die Funktion unbeschränkt ist.



$$\int_0^{\infty} f(x) dx = \lim_{t \rightarrow \infty} \int_0^t f(x) dx$$

$$\int_0^1 f(x) dx = \lim_{t \rightarrow 0} \int_t^1 f(x) dx$$

Beispiele

$$\int_0^1 \frac{1}{\sqrt{x}} dx = \lim_{t \rightarrow 0} \int_t^1 x^{-\frac{1}{2}} dx = \lim_{t \rightarrow 0} 2\sqrt{x} \Big|_t^1 = \lim_{t \rightarrow 0} (2 - 2\sqrt{t}) = 2$$

$$\int_1^{\infty} \frac{1}{x^2} dx = \lim_{t \rightarrow \infty} \int_1^t x^{-2} dx = \lim_{t \rightarrow \infty} -\frac{1}{x} \Big|_1^t = \lim_{t \rightarrow \infty} -\frac{1}{t} - (-1) = 1$$

$$\int_1^{\infty} \frac{1}{x} dx = \lim_{t \rightarrow \infty} \int_1^t \frac{1}{x} dx = \lim_{t \rightarrow \infty} \ln(x) \Big|_1^t = \lim_{t \rightarrow \infty} \ln(t) - \ln(1) = \infty$$

Das uneigentliche Integral existiert somit nicht.

Zwei Limiten

In der Wahrscheinlichkeitstheorie wird oft über einen Bereich integriert, in dem beide Grenzen unbeschränkt sind. Z.B.:

Der Erwartungswert einer Zufallsvariable X mit Dichte f ist definiert als

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

In diesem Fall müssen wir die beiden Grenzwerte getrennt auswerten:

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x \cdot f(x) dx \\ &= \lim_{t \rightarrow -\infty} \int_t^0 x \cdot f(x) dx + \lim_{s \rightarrow \infty} \int_0^s x \cdot f(x) dx \end{aligned}$$

Achtung!

Tritt dabei als Zwischenresultat $\infty - \infty$ auf,
so ist das Ergebnis **nicht** $\infty - \infty = 0$!

Leibniz Formel

Gegen sei eine *stetig differenzierbare* Funktion $f(x, t)$ und

$$F(x) = \int_{a(x)}^{b(x)} f(x, t) dt$$

Dann gilt für die Ableitung

$$F'(x) = f(x, b(x)) b'(x) - f(x, a(x)) a'(x) + \int_{a(x)}^{b(x)} \frac{\partial f(x, t)}{\partial x} dt$$

Falls $a(x) = a$ und $b(x) = b$ konstant sind, dann gilt insbesondere:

$$\frac{d}{dx} \int_a^b f(x, t) dt = \int_a^b \frac{\partial}{\partial x} f(x, t) dt$$

Beispiel

Sei $F(x) = \int_x^{2x} t x^2 dt$ für $x \geq 0$. Berechne $F'(x)$.

Wir setzen $f(x, t) = t x^2$, $a(x) = x$ und $b(x) = 2x$.

Dann gilt auf Grund der Leibniz Formel,

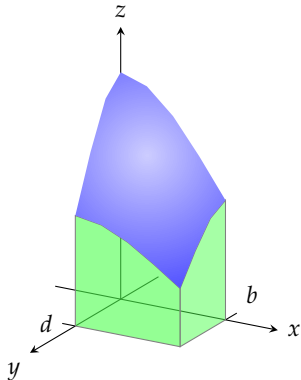
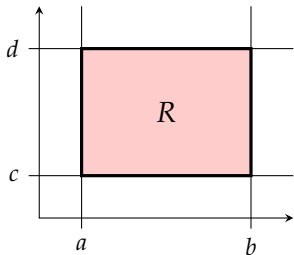
$$\begin{aligned} F'(x) &= f(x, b) \cdot b' - f(x, a) \cdot a' + \int_a^b f_x(x, t) dt \\ &= (2x) x^2 \cdot 2 - (x) x^2 \cdot 1 + \int_x^{2x} 2x t dt \\ &= 4x^3 - x^3 + \left(2x \frac{1}{2} t^2 \right) \Big|_x^{2x} \\ &= 4x^3 - x^3 + (4x^3 - x^3) \\ &= 6x^3 \end{aligned}$$

Volumen

Sei $f(x, y)$ eine Funktion definiert über dem Rechteck

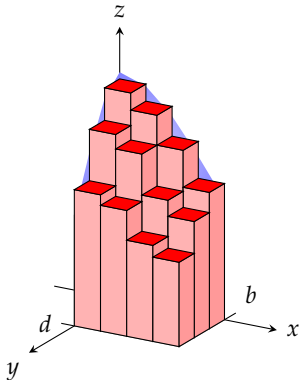
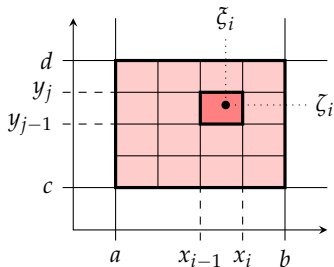
$$R = [a, b] \times [c, d] = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, c \leq y \leq d\}$$

Wie groß ist das Volumen V unter dem Graphen von f ?



Riemann-Summen

- ▶ Zerlege R in kleine Rechtecke $R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$
- ▶ Berechne $V \approx \sum_{i=1}^n \sum_{j=1}^k f(\xi_i, \zeta_j) (x_i - x_{i-1}) (y_j - y_{j-1})$



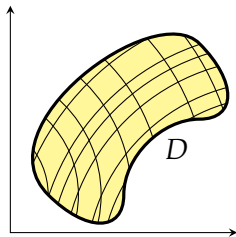
Riemann-Integral

Falls die **Riemann-Summe** für immer feinere Zerlegungen von R konvergiert, dann heißt dieser Grenzwert das **Riemann-Integral** von f über R :

$$\iint_R f(x, y) dx dy = \lim_{n, k \rightarrow \infty} \sum_{i=1}^n \sum_{j=1}^k f(\xi_i, \zeta_j) (x_i - x_{i-1}) (y_j - y_{j-1})$$

Das Riemann-Integral ist analog auf beliebigen Gebieten D definiert.

$$\iint_D f(x, y) dx dy$$



Satz von Fubini

Sei $f: R = [a, b] \times [c, d] \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ eine *stetige* Funktion. Dann gilt

$$\begin{aligned} \iint_R f(x, y) \, dx \, dy &= \int_a^b \left(\int_c^d f(x, y) \, dy \right) dx \\ &= \int_c^d \left(\int_a^b f(x, y) \, dx \right) dy . \end{aligned}$$

Der Satz von Fubini liefert uns ein Rezept zur Berechnung von Doppelintegralen:

1. Fasse x als Konstante auf und berechne das innere Integral $\int_c^d f(x, y) \, dy$.
2. Integriere das Ergebnis nach x .

Wir dürfen aber auch die Reihenfolge vertauschen und zuerst nach x und erst dann nach y integrieren.

Beispiel

Berechne $\int_{-1}^1 \int_0^1 (1 - x - y^2 + xy^2) dx dy$.

Wir müssen zweimal integrieren.

$$\begin{aligned} & \int_{-1}^1 \int_0^1 (1 - x - y^2 + xy^2) dx dy \\ &= \int_{-1}^1 \left(x - \frac{1}{2}x^2 - xy^2 + \frac{1}{2}x^2y^2 \Big|_0^1 \right) dy \\ &= \int_{-1}^1 \left(\frac{1}{2} - \frac{1}{2}y^2 \right) dy = \frac{1}{2}y - \frac{1}{6}y^3 \Big|_{-1}^1 \\ &= \frac{1}{2} - \frac{1}{6} - \left(-\frac{1}{2} + \frac{1}{6} \right) = \frac{2}{3} \end{aligned}$$

Integrationsgrenzen

Achtung, die Integrationsvariable und die dazu gehörigen Integrationsgrenzen sind von **innen** nach **außen** zu lesen.

Wenn wir die Reihenfolge beim Integrieren vertauschen, dann müssen wir das auch in der Reihenfolge der Integrationsgrenzen berücksichtigen:

$$\int_a^b \int_c^d f(x, y) dy dx = \int_c^d \int_a^b f(x, y) dx dy$$

Wir sehen das besser, wenn wir die (redundanten) Klammer einfügen:

$$\int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy .$$

Beispiel – Fortsetzung

Integration in umgekehrter Reihenfolge:

$$\begin{aligned} & \int_{-1}^1 \int_0^1 (1 - x - y^2 + xy^2) dx dy \\ &= \int_0^1 \left(\int_{-1}^1 (1 - x - y^2 + xy^2) dy \right) dx \\ &= \int_0^1 \left(y - xy - \frac{1}{3}y^3 + \frac{1}{3}xy^3 \Big|_{-1}^1 \right) dx \\ &= \int_0^1 \left(1 - x - \frac{1}{3} + \frac{1}{3}x - \left(-1 + x + \frac{1}{3} - \frac{1}{3}x \right) \right) dx \\ &= \int_0^1 \left(\frac{4}{3} - \frac{4}{3}x \right) dx = \frac{4}{3}x - \frac{4}{6}x^2 \Big|_0^1 = \frac{2}{3} \end{aligned}$$

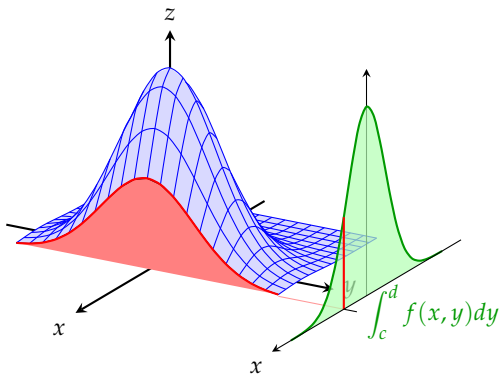
Satz von Fubini – Interpretation

$$\iint_{\mathbb{R}} f(x, y) dx dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_a^b A(x) dx$$

Wenn wir x festhalten,
dann ist

$$A(x) = \int_c^d f(x, y) dy$$

die Fläche unter
 $g(y) = f(x, y)$.



Zusammenfassung

- ▶ Flächeninhalt und Riemann-Integral
- ▶ Stammfunktion
- ▶ Hauptsatz der Differential- und Integralrechnung
- ▶ Integrationsverfahren
- ▶ Uneigentliches Integral
- ▶ Differenzieren unter dem Integral
- ▶ Volumen und Doppelintegral
- ▶ Satz von Fubini

Kapitel 11

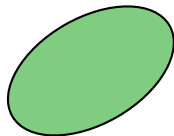
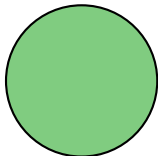
Extrema

Konvexe Menge

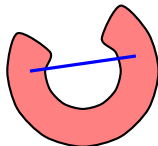
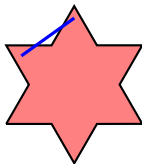
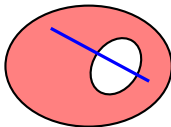
Eine Menge $D \subseteq \mathbb{R}^n$ heißt **konvex**, wenn für zwei beliebige Punkte $\mathbf{x}, \mathbf{y} \in D$ auch die Verbindungsstrecke dieser Punkte in D liegt, d.h.

$$(1 - h) \mathbf{x} + h \mathbf{y} \in D \quad \text{für alle } h \in [0, 1], \text{ und } \mathbf{x}, \mathbf{y} \in D$$

konvex:

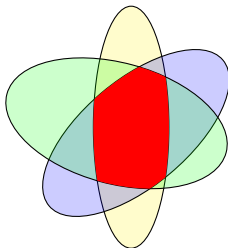


nicht konvex:



Durchschnitt konvexer Mengen

Seien S_1, \dots, S_k konvexe Teilmengen des \mathbb{R}^n , dann ist auch der Durchschnitt $S_1 \cap \dots \cap S_k$ konvex.



Die Vereinigung zweier konvexer Mengen muss nicht konvex sein.

Beispiel – Halbräume

Seien $\mathbf{p} \in \mathbb{R}^n$ und m fest, $\mathbf{p} \neq 0$. Dann beschreibt

$$H = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{p}^t \cdot \mathbf{x} = m\}$$

eine **Hyperebene**, die den \mathbb{R}^n in die zwei **Halbräume**

$$H_+ = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{p}^t \cdot \mathbf{x} \geq m\}$$

$$H_- = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{p}^t \cdot \mathbf{x} \leq m\}$$

teilt. Die Mengen H , H_+ und H_- sind konvex.

Sei \mathbf{x} ein Gütervektor, \mathbf{p} der entsprechende Preisvektor und m das verfügbare Einkommen. Dann ist die Budgetmenge konvex:

$$\begin{aligned} & \{\mathbf{x} \in \mathbb{R}^n : \mathbf{p}^t \cdot \mathbf{x} \leq m, \mathbf{x} \geq 0\} \\ &= \{\mathbf{x} : \mathbf{p}^t \cdot \mathbf{x} \leq m\} \cap \{\mathbf{x} : x_1 \geq 0\} \cap \dots \cap \{\mathbf{x} : x_n \geq 0\} \end{aligned}$$

Konvexe und konkave Funktionen

Eine Funktion f heißt **konvex** in $D \subseteq \mathbb{R}^n$, falls D konvex ist, und

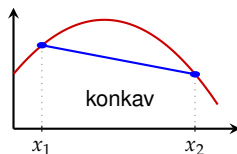
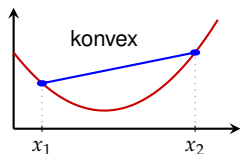
$$f((1-h)x_1 + hx_2) \leq (1-h)f(x_1) + hf(x_2)$$

für alle $x_1, x_2 \in D$ und alle $h \in [0, 1]$.

Die Funktion f heißt **konkav** in $D \subseteq \mathbb{R}^n$, falls D konvex ist, und

$$f((1-h)x_1 + hx_2) \geq (1-h)f(x_1) + hf(x_2)$$

für alle $x_1, x_2 \in D$ und alle $h \in [0, 1]$.



Streng konvexe und streng konkave Funktionen

Eine Funktion f heißt **streng konvex** in $D \subseteq \mathbb{R}^n$, falls D *konvex* ist, und

$$f((1-h)\mathbf{x}_1 + h\mathbf{x}_2) < (1-h)f(\mathbf{x}_1) + hf(\mathbf{x}_2)$$

für alle $\mathbf{x}_1, \mathbf{x}_2 \in D$ mit $\mathbf{x}_1 \neq \mathbf{x}_2$ und alle $h \in (0, 1)$.

Die Funktion f heißt **streng konkav** in $D \subseteq \mathbb{R}^n$, falls D *konvex* ist, und

$$f((1-h)\mathbf{x}_1 + h\mathbf{x}_2) > (1-h)f(\mathbf{x}_1) + hf(\mathbf{x}_2)$$

für alle $\mathbf{x}_1, \mathbf{x}_2 \in D$ mit $\mathbf{x}_1 \neq \mathbf{x}_2$ und alle $h \in (0, 1)$.

Beispiel – Lineare Funktion

Sei $\mathbf{a} \in \mathbb{R}^n$ konstant.

Dann ist $f(\mathbf{x}) = \mathbf{a}^t \cdot \mathbf{x}$ eine lineare Funktion und es gilt:

$$\begin{aligned} f((1-h)\mathbf{x}_1 + h\mathbf{x}_2) &= \mathbf{a}^t \cdot ((1-h)\mathbf{x}_1 + h\mathbf{x}_2) \\ &= (1-h)\mathbf{a}^t \cdot \mathbf{x}_1 + h\mathbf{a}^t \cdot \mathbf{x}_2 \\ &= (1-h)f(\mathbf{x}_1) + hf(\mathbf{x}_2) \end{aligned}$$

D.h., f ist sowohl konkav als auch konvex.

Die lineare Funktion ist aber weder streng konkav noch streng konvex, da die Ungleichung niemals strikt ist.

Beispiel – Quadratische Funktion in einer Variable

Die Funktion $f(x) = x^2$ ist streng konvex:

$$\begin{aligned} & f((1-h)x + hy) - [(1-h)f(x) + hf(y)] \\ &= ((1-h)x + hy)^2 - [(1-h)x^2 + hy^2] \\ &= (1-h)^2 x^2 + 2(1-h)hxy + h^2 y^2 - (1-h)x^2 - hy^2 \\ &= -h(1-h)x^2 + 2(1-h)hxy - h(1-h)y^2 \\ &= -h(1-h)(x-y)^2 \\ &< 0 \quad \text{für } x \neq y \text{ und } 0 < h < 1 \end{aligned}$$

Also

$$f((1-h)x + hy) < (1-h)f(x) + hf(y)$$

für alle $x \neq y$ und $0 < h < 1$.

D.h. $f(x) = x^2$ ist streng konvex.

Eigenschaften

- ▶ Falls $f(\mathbf{x})$ (streng) *konvex* ist, dann ist $-f(\mathbf{x})$ (streng) *konkav* (und umgekehrt).
- ▶ Falls $f_1(\mathbf{x}), \dots, f_k(\mathbf{x})$ *konvex* (konkav) und $\alpha_1, \dots, \alpha_k > 0$ ist, dann ist auch

$$g(\mathbf{x}) = \alpha_1 f_1(\mathbf{x}) + \dots + \alpha_k f_k(\mathbf{x})$$

konvex (konkav).

- ▶ Falls (zumindest) eine der Funktionen $f_i(x)$ *streng konvex* (streng konkav) ist, dann ist auch $g(x)$ streng konvex (streng konkav).

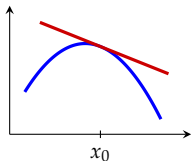
Eigenschaften

Für eine differenzierbare Funktion f gilt:

- ▶ f ist **konkav** genau dann, wenn

$$f(\mathbf{x}) - f(\mathbf{x}_0) \leq \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$

d.h., der Funktionsgraph liegt immer unterhalb der Tangente.



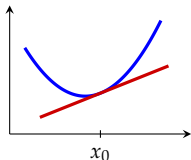
- ▶ f ist **streng konkav** genau dann, wenn

$$f(\mathbf{x}) - f(\mathbf{x}_0) < \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \quad \text{für alle } \mathbf{x} \neq \mathbf{x}_0$$

- ▶ f ist **konvex** genau dann, wenn

$$f(\mathbf{x}) - f(\mathbf{x}_0) \geq \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$

(Analog für streng konvexe Funktion.)



Quadratische Form

Sei \mathbf{A} eine symmetrische Matrix und $q_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^t \mathbf{A} \mathbf{x}$ die entsprechende quadratische Form.

Wir können \mathbf{A} diagonalisieren, d.h., es gibt eine Orthonormalbasis aus Eigenvektoren, sodass \mathbf{A} zur Diagonalmatrix \mathbf{D} aus Eigenwerten wird.

Wenn \mathbf{c} der Koordinatenvektor von \mathbf{x} bezüglich dieser Basis ist, dann ist

$$q_{\mathbf{A}}(\mathbf{x}) = c_1^2 \lambda_1 + c_2^2 \lambda_2 + \cdots + c_n^2 \lambda_n$$

- ▶ Wenn alle Eigenwerte $\lambda_i \geq 0$ sind, dann ist $q_{\mathbf{A}}$ eine Summe von konvexen Funktionen und damit selbst konvex.
- ▶ Wenn alle $\lambda_i \leq 0$ sind, dann ist $q_{\mathbf{A}}$ konkav.
- ▶ Wenn es Eigenwerte $\lambda_i > 0$ und $\lambda_i < 0$, dann ist $q_{\mathbf{A}}$ weder konvex noch konkav.

Quadratische Form

Für eine quadratische Form $q_{\mathbf{A}}$ gilt:

- ▶ *streng konvex* \Leftrightarrow *positiv definit*
- ▶ *konvex* \Leftrightarrow *positiv semidefinit*
- ▶ *streng konkav* \Leftrightarrow *negativ definit*
- ▶ *konkav* \Leftrightarrow *negativ semidefinit*
- ▶ *weder noch* \Leftrightarrow *indefinit*

Die Definitheit von \mathbf{A} kann festgestellt werden mit Hilfe

- ▶ der Eigenwerte von \mathbf{A} , oder
- ▶ der (allgemeinen) Hauptminoren von \mathbf{A} .

Beispiel – Quadratische Form

Sei $\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 2 \end{pmatrix}$. Die Hauptminoren lauten:

$$H_1 = 2 > 0$$

$$H_2 = \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = 5 > 0$$

$$H_3 = |\mathbf{A}| = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 2 \end{vmatrix} = 8 > 0$$

\mathbf{A} ist daher positiv definit.

Die quadratische Form $q_{\mathbf{A}}$ ist somit *streng konvex*.

Beispiel – Quadratische Form

Sei $\mathbf{A} = \begin{pmatrix} -1 & 0 & 1 \\ 0 & -4 & 2 \\ 1 & 2 & -2 \end{pmatrix}$. Allgemeinen Hauptminoren:

$$\tilde{H}_1 = -1$$

$$\tilde{H}_2 = -4$$

$$\tilde{H}_3 = -2$$

$$\tilde{H}_{1,2} = \begin{vmatrix} -1 & 0 \\ 0 & -4 \end{vmatrix} = 4 \quad \tilde{H}_{1,3} = \begin{vmatrix} -1 & 1 \\ 1 & -2 \end{vmatrix} = 1 \quad \tilde{H}_{2,3} = \begin{vmatrix} -4 & 2 \\ 2 & -2 \end{vmatrix} = 4$$

$$\tilde{H}_{1,2,3} = \begin{vmatrix} -1 & 0 & 1 \\ 0 & -4 & 2 \\ 1 & 2 & -2 \end{vmatrix} = 0$$

$$\tilde{H}_i \leq 0$$

$$\tilde{H}_{i,j} \geq 0$$

$$\tilde{H}_{1,2,3} \leq 0$$

\mathbf{A} ist daher negativ semidefinit.

Die quadratische Form $q_{\mathbf{A}}$ ist somit *konkav* (aber nicht streng konkav).

Krümmung differenzierbarer Funktionen

Sei $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ mit Taylorreihenentwicklung

$$f(\mathbf{x}_0 + \mathbf{h}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2} \mathbf{h}^t \cdot \mathbf{H}_f(\mathbf{x}_0) \cdot \mathbf{h}$$

Die *Hesse-Matrix* $\mathbf{H}_f(\mathbf{x}_0)$ bestimmt die *Krümmung* von f in der Nähe des Entwicklungspunkts \mathbf{x}_0 .

- ▶ $\mathbf{H}_f(\mathbf{x}_0)$ *positiv definit* $\Rightarrow f$ *streng konvex* um \mathbf{x}_0
- ▶ $\mathbf{H}_f(\mathbf{x}_0)$ *negativ definit* $\Rightarrow f$ *streng konkav* um \mathbf{x}_0

- ▶ $\mathbf{H}_f(\mathbf{x})$ *positiv semidefinit* für alle $\mathbf{x} \in D$ $\Leftrightarrow f$ *konvex* in D
- ▶ $\mathbf{H}_f(\mathbf{x})$ *negativ semidefinit* für alle $\mathbf{x} \in D$ $\Leftrightarrow f$ *konkav* in D

Vorgangsweise – streng konvex

1. Berechne Hesse-Matrix

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) & \cdots & f_{x_1x_n}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & \cdots & f_{x_2x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{x}) & f_{x_nx_2}(\mathbf{x}) & \cdots & f_{x_nx_n}(\mathbf{x}) \end{pmatrix}$$

2. Berechne die (führenden) *Hauptminoren* H_i .
3. $\blacktriangleright f$ streng konvex \Leftrightarrow alle $H_k > 0$ für (fast) **alle** $\mathbf{x} \in D$
 $\blacktriangleright f$ streng konkav \Leftrightarrow alle $(-1)^k H_k > 0$ für (fast) **alle** $\mathbf{x} \in D$
[$(-1)^k H_k > 0$ heißt: $H_1, H_3, \dots < 0$ und $H_2, H_4, \dots > 0$]
4. Andernfalls ist f in D weder streng konvex noch streng konkav.

Vorgangsweise – konvex

1. Berechne Hesse-Matrix

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) & \cdots & f_{x_1x_n}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & \cdots & f_{x_2x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{x}) & f_{x_nx_2}(\mathbf{x}) & \cdots & f_{x_nx_n}(\mathbf{x}) \end{pmatrix}$$

2. Berechne die *allgemeinen Hauptminoren* $\tilde{H}_{i_1, \dots, i_k}$.

3. ▶ f konvex \Leftrightarrow alle $\tilde{H}_{i_1, \dots, i_k} \geq 0$ für **alle** $\mathbf{x} \in D$.

▶ f konkav \Leftrightarrow alle $(-1)^k \tilde{H}_{i_1, \dots, i_k} \geq 0$ für **alle** $\mathbf{x} \in D$.

4. Andernfalls ist f in D weder konvex noch konkav.

Offensichtlich ist jede *streng* konkave (konvexe) Funktion auch konkav (bzw. konvex).

Beispiel

Ist die Funktion (streng) konkav oder konvex?

$$f(x, y) = x^4 + x^2 - 2xy + y^2$$

1. Hesse-Matrix: $\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} 12x^2 + 2 & -2 \\ -2 & 2 \end{pmatrix}$

2. Hauptminoren:

$$H_1 = 12x^2 + 2 > 0$$

$$H_2 = |\mathbf{H}_f(\mathbf{x})| = 24x^2 > 0 \quad \text{für alle } x \neq 0.$$

3. Alle Hauptminoren > 0 für (fast) alle \mathbf{x}

$\Rightarrow f$ ist *streng konvex*. (und damit auch konvex)

Beispiel – Cobb-Douglas Funktion

Sei $f(x, y) = x^\alpha y^\beta$ mit $\alpha, \beta \geq 0$ und $\alpha + \beta \leq 1$,
und $D = \{(x, y) : x, y \geq 0\}$.

Hesse-Matrix an der Stelle \mathbf{x} :

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} \alpha(\alpha - 1) x^{\alpha-2} y^\beta & \alpha\beta x^{\alpha-1} y^{\beta-1} \\ \alpha\beta x^{\alpha-1} y^{\beta-1} & \beta(\beta - 1) x^\alpha y^{\beta-2} \end{pmatrix}$$

Allgemeine Hauptminoren:

$$\tilde{H}_1 = \underbrace{\alpha}_{\geq 0} \underbrace{(\alpha - 1)}_{\leq 0} \underbrace{x^{\alpha-2} y^\beta}_{\geq 0} \leq 0$$

$$\tilde{H}_2 = \underbrace{\beta}_{\geq 0} \underbrace{(\beta - 1)}_{\leq 0} \underbrace{x^\alpha y^{\beta-2}}_{\geq 0} \leq 0$$

Beispiel – Cobb-Douglas Funktion

$$\begin{aligned}\tilde{H}_{1,2} &= |\mathbf{H}_f(\mathbf{x})| \\ &= \alpha(\alpha - 1) x^{\alpha-2} y^\beta \cdot \beta(\beta - 1) x^\alpha y^{\beta-2} - (\alpha\beta x^{\alpha-1} y^{\beta-1})^2 \\ &= \alpha(\alpha - 1) \beta(\beta - 1) x^{2\alpha-2} y^{2\beta-2} - \alpha^2 \beta^2 x^{2\alpha-2} y^{2\beta-2} \\ &= \alpha\beta [(\alpha - 1)(\beta - 1) - \alpha\beta] x^{2\alpha-2} y^{2\beta-2} \\ &= \underbrace{\alpha\beta}_{\geq 0} \underbrace{(1 - \alpha - \beta)}_{\geq 0} \underbrace{x^{2\alpha-2} y^{2\beta-2}}_{\geq 0} \quad \geq 0\end{aligned}$$

$\tilde{H}_1 \leq 0$ und $\tilde{H}_2 \leq 0$, und $\tilde{H}_{1,2} \geq 0$ für alle $(x, y) \in D$.

$f(x, y)$ ist daher *konkav* in D .

Für $0 < \alpha, \beta < 1$ und $\alpha + \beta < 1$ gilt sogar:

$H_1 = \tilde{H}_1 < 0$ und $H_2 = |\mathbf{H}_f(\mathbf{x})| > 0$ für fast alle $(x, y) \in D$.

$f(x, y)$ ist dann sogar *streng konkav*.

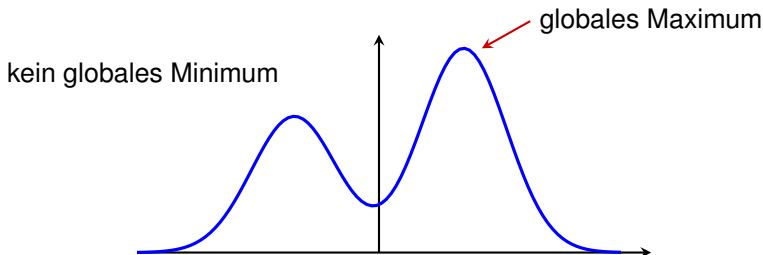
Globales Extremum (Optimum)

Ein Punkt \mathbf{x}^* heißt **globales Maximum** (*absolute Maximum*) von f , falls für alle $\mathbf{x} \in D_f$ gilt:

$$f(\mathbf{x}^*) \geq f(\mathbf{x})$$

Ein Punkt \mathbf{x}^* heißt **globales Minimum** (*absolute Minimum*) von f , falls für alle $\mathbf{x} \in D_f$ gilt:

$$f(\mathbf{x}^*) \leq f(\mathbf{x})$$



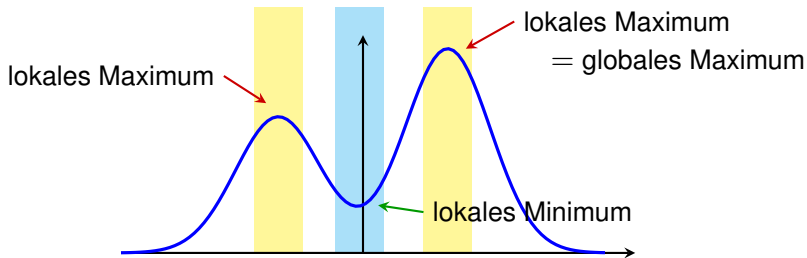
Lokales Extremum (Optimum)

Ein Punkt x_0 heißt **lokales Maximum** (*relatives Maximum*) von f , falls für alle x in einer geeigneten Umgebung von x_0 gilt:

$$f(x_0) \geq f(x)$$

Ein Punkt x_0 heißt **lokales Minimum** (*relatives Minimum*) von f , falls für alle x in einer geeigneten Umgebung von x_0 gilt:

$$f(x_0) \leq f(x)$$



Kritischer Punkt

In einem (lokalen) Maximum oder Minimum muss jede Richtungsableitung und damit auch der Gradient gleich Null sein.

Ein Punkt x_0 heißt **kritischer Punkt** (oder *stationärer Punkt*) einer Funktion f , wenn

$$\nabla f(x_0) = 0$$

Notwendige Bedingung:

Jedes Extremum von f ist ein kritischer Punkt von f .

Globales Extremum

Hinreichende Bedingung:

Sei \mathbf{x}_0 ein kritischer Punkt einer **konkaven** (oder *konvexen*) Funktion f . Dann ist \mathbf{x}_0 ein **globales Maximum** (bzw. *globales Minimum*) von f .

Falls f *streng* konkav (oder konvex) ist, dann ist das Extremum eindeutig bestimmt.

Die Aussage folgt unmittelbar aus den Eigenschaften (streng) konkaver Funktionen. Für alle $\mathbf{x} \neq \mathbf{x}_0$ gilt

$$f(\mathbf{x}) - f(\mathbf{x}_0) < \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = 0 \cdot (\mathbf{x} - \mathbf{x}_0) = 0$$

und somit

$$f(\mathbf{x}_0) > f(\mathbf{x})$$

Beispiel*

Sei $f(x) = e^x - 2x$.

Die Funktion ist streng konvex:

$$f'(x) = e^x - 2$$

$$f''(x) = e^x > 0 \quad \text{für alle } x \in \mathbb{R}$$

Kritischer Punkt:

$$f'(x) = e^x - 2 = 0 \quad \Rightarrow \quad x_0 = \ln 2$$

$x_0 = \ln 2$ ist das (eindeutig bestimmte) globale Minimum von f .

Beispiel

Sei $f: D = [0, \infty)^2 \rightarrow \mathbb{R}$, $f(x, y) = 4x^{\frac{1}{4}}y^{\frac{1}{4}} - x - y$

Hesse-Matrix an der Stelle \mathbf{x} :

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} -\frac{3}{4}x^{-\frac{7}{4}}y^{\frac{1}{4}} & \frac{1}{4}x^{-\frac{3}{4}}y^{-\frac{3}{4}} \\ \frac{1}{4}x^{-\frac{3}{4}}y^{-\frac{3}{4}} & -\frac{3}{4}x^{\frac{1}{4}}y^{-\frac{7}{4}} \end{pmatrix}$$

Hauptminoren:

$$H_1 = -\frac{3}{4}x^{-\frac{7}{4}}y^{\frac{1}{4}} < 0$$

$$H_2 = \frac{1}{2}x^{-\frac{3}{2}}y^{-\frac{3}{2}} > 0$$

f ist streng konkav in D .

Kritischer Punkt: $\nabla f = (x^{-\frac{3}{4}}y^{\frac{1}{4}} - 1, x^{\frac{1}{4}}y^{-\frac{3}{4}} - 1) = 0$

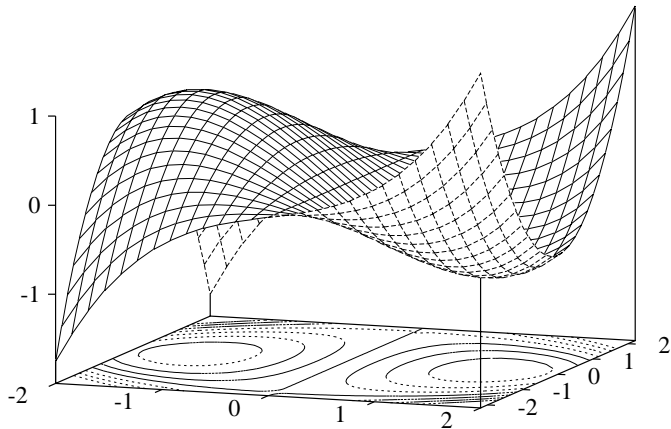
$$f_x = x^{-\frac{3}{4}}y^{\frac{1}{4}} - 1 = 0$$

$$f_y = x^{\frac{1}{4}}y^{-\frac{3}{4}} - 1 = 0$$

$$\Rightarrow \mathbf{x}_0 = (1, 1)$$

\mathbf{x}_0 ist das globale Maximum von f .

Beispiel – Lokale Extrema



$$f(x, y) = \frac{1}{6} x^3 - x + \frac{1}{4} x y^2$$

Lokal konkav

Ein Punkt \mathbf{x}_0 ist ein **lokales Maximum** (oder *Minimum*) von f , falls

- ▶ \mathbf{x}_0 ist ein **kritischer Punkt** von f ,
- ▶ f ist **konkav** (bzw. *konvex*) um \mathbf{x}_0 .

Hinreichende Bedingung:

f ist *lokal* streng konkav (konvex) um \mathbf{x}_0 ist.

Sei \mathbf{x}_0 ein kritischer Punkt von f . Dann gilt

- ▶ $\mathbf{H}_f(\mathbf{x}_0)$ *negativ definit* \Rightarrow \mathbf{x}_0 ist lokales Maximum
- ▶ $\mathbf{H}_f(\mathbf{x}_0)$ *positiv definit* \Rightarrow \mathbf{x}_0 ist lokales Minimum

Es ist ausreichend die Hesse-Matrix $\mathbf{H}_f(\mathbf{x})$ am kritischen Punkt \mathbf{x}_0 auszuwerten. (Im Gegensatz zur Bedingung für globale Extrema.)

Vorgangsweise – Univariate Funktion*

Hinreichende Bedingung

für lokale Extremwerte einer Funktion in *einer* Variablen:

1. Berechne $f'(x)$ und $f''(x)$.
2. Suche alle Punkte x_i mit $f'(x_i) = 0$ (kritischen Punkte).
3. Falls $f''(x_i) < 0 \Rightarrow x_i$ ist ein *lokales Maximum*.
Falls $f''(x_i) > 0 \Rightarrow x_i$ ist ein *lokales Minimum*.
Falls $f''(x_i) = 0 \Rightarrow$ *keine* Aussage möglich!

Falls $f''(x_i) = 0$ dann werden andere Methoden benötigt!

Z.B. kann man Terme höherer Ordnung der Taylorreihe mit Entwicklungspunkt x_0 betrachten.

Beispiel – Univariate Funktion*

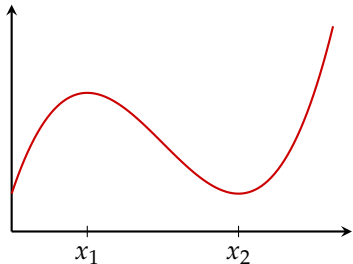
Gesucht sind die lokalen Extrema der Funktion

$$f(x) = \frac{1}{12}x^3 - x^2 + 3x + 1$$

1. $f'(x) = \frac{1}{4}x^2 - 2x + 3,$
 $f''(x) = \frac{1}{2}x - 2.$

2. $\frac{1}{4}x^2 - 2x + 3 = 0$
besitzt die Lösungen
 $x_1 = 2$ und $x_2 = 6.$

3. $f''(2) = -1 \Rightarrow x_1$ ist lokales Maximum.
 $f''(6) = 1 \Rightarrow x_2$ ist lokales Minimum.



Beispiel – Kritische Punkte

Suche alle kritischen Punkte von

$$f(x, y) = \frac{1}{6} x^3 - x + \frac{1}{4} x y^2$$

Partiellen Ableitungen:

$$(I) \quad f_x = \frac{1}{2} x^2 - 1 + \frac{1}{4} y^2 = 0$$

$$(II) \quad f_y = \frac{1}{2} x y = 0$$

$$(II) \Rightarrow \quad x = 0 \quad \text{oder} \quad y = 0$$

$$(I) \Rightarrow \quad -1 + \frac{1}{4} y^2 = 0 \quad \left| \quad \frac{1}{2} x^2 - 1 = 0 \right. \\ \left. \quad \quad \quad y = \pm 2 \quad \quad \quad \quad \quad \quad \quad \quad \quad x = \pm \sqrt{2} \right.$$

Kritische Punkte:

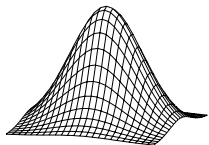
$$\mathbf{x}_1 = (0, 2)$$

$$\mathbf{x}_3 = (\sqrt{2}, 0)$$

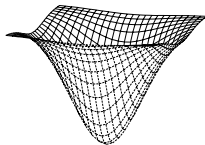
$$\mathbf{x}_2 = (0, -2)$$

$$\mathbf{x}_4 = (-\sqrt{2}, 0)$$

Kritische Punkte – Lokale Extrema

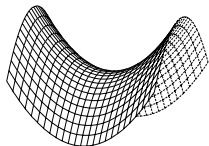


Lokales Maximum

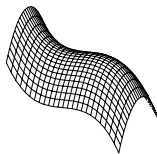


Lokales Minimum

Kritische Punkte – Sattelpunkte



Sattelpunkt



Beispiel für höhere Ordnung

Vorgangsweise – Lokale Extrema

1. Berechne Gradient ∇f und Hesse-Matrix \mathbf{H}_f .
2. Bestimme alle \mathbf{x}_i mit $\nabla f(\mathbf{x}_i) = 0$ (kritische Punkte).
3. Berechne alle Hauptminoren H_k für kritischen Punkt \mathbf{x}_0 :
 - (a) Alle Hauptminoren $H_k > 0$
 $\Rightarrow \mathbf{x}_0$ ist ein **lokales Minimum** von f .
 - (b) Für alle Hauptminoren gilt $(-1)^k H_k > 0$
[d.h., $H_1, H_3, \dots < 0$ und $H_2, H_4, \dots > 0$]
 $\Rightarrow \mathbf{x}_0$ ist ein **lokales Maximum** von f .
 - (c) $\det(\mathbf{H}_f(\mathbf{x}_0)) \neq 0$, aber weder (a) noch (b) sind erfüllt
 $\Rightarrow \mathbf{x}_0$ ist ein **Sattelpunkt** von f .
 - (d) Andernfalls ist *keine Aussage* möglich,
d.h. \mathbf{x}_0 kann ein lokales Extremum sein, muss aber nicht.

Vorgangsweise – Bivariate Funktion

1. Berechne Gradient ∇f und Hesse-Matrix \mathbf{H}_f .
2. Bestimme alle \mathbf{x}_i mit $\nabla f(\mathbf{x}_i) = 0$ (kritische Punkte).
3. Berechne die Hauptminoren H_1 und H_2 für kritischen Punkt \mathbf{x}_0 :
 - (a) $H_2 > 0$ und $H_1 > 0$
 $\Rightarrow \mathbf{x}_0$ ist ein **lokales Minimum** von f .
 - (b) $H_2 > 0$ und $H_1 < 0$
 $\Rightarrow \mathbf{x}_0$ ist ein **lokales Maximum** von f .
 - (c) $H_2 < 0$
 $\Rightarrow \mathbf{x}_0$ ist ein **Sattelpunkt** von f .
 - (d) $H_2 = \det(\mathbf{H}_f(\mathbf{x}_0)) = 0$
 \Rightarrow *keine Aussage* möglich.

Beispiel – Bivariate Funktion

Suche die lokalen Extrema von

$$f(x, y) = \frac{1}{6} x^3 - x + \frac{1}{4} x y^2$$

1. $\nabla f = \left(\frac{1}{2} x^2 - 1 + \frac{1}{4} y^2, \frac{1}{2} x y \right)$

$$\mathbf{H}_f(x, y) = \begin{pmatrix} x & \frac{1}{2} y \\ \frac{1}{2} y & \frac{1}{2} x \end{pmatrix}$$

2. Kritische Punkte:

$$\mathbf{x}_1 = (0, 2), \mathbf{x}_2 = (0, -2), \mathbf{x}_3 = (\sqrt{2}, 0), \mathbf{x}_4 = (-\sqrt{2}, 0)$$

Beispiel – Bivariate Funktion (Forts.)

3. Hauptminoren:

$$\mathbf{H}_f(\mathbf{x}_1) = \mathbf{H}_f(0, 2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$H_2 = -1 < 0 \Rightarrow \mathbf{x}_1 \text{ ist ein Sattelpunkt}$$

$$\mathbf{H}_f(\mathbf{x}_2) = \mathbf{H}_f(0, -2) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

$$H_2 = -1 < 0 \Rightarrow \mathbf{x}_2 \text{ ist ein Sattelpunkt}$$

Beispiel – Bivariate Funktion (Forts.)

3. Hauptminoren:

$$\mathbf{H}_f(\mathbf{x}_3) = \mathbf{H}_f(\sqrt{2}, 0) = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{\sqrt{2}}{2} \end{pmatrix}$$

$$H_2 = 1 > 0 \quad \text{und} \quad H_1 = \sqrt{2} > 0$$

$\Rightarrow \mathbf{x}_3$ ist ein *lokales Minimum*

$$\mathbf{H}_f(\mathbf{x}_4) = \mathbf{H}_f(-\sqrt{2}, 0) = \begin{pmatrix} -\sqrt{2} & 0 \\ 0 & -\frac{\sqrt{2}}{2} \end{pmatrix}$$

$$H_2 = 1 > 0 \quad \text{und} \quad H_1 = -\sqrt{2} < 0$$

$\Rightarrow \mathbf{x}_4$ ist ein *lokales Maximum*

Konvexe Funktion und konvexe Menge

Sei f konvex.

Dann ist die **untere Niveaumenge** von f

$$\{\mathbf{x} \in D_f : f(\mathbf{x}) \leq c\}$$

konvex.

Seien $\mathbf{x}_1, \mathbf{x}_2 \in \{\mathbf{x} \in D_f : f(\mathbf{x}) \leq c\}$, d.h.

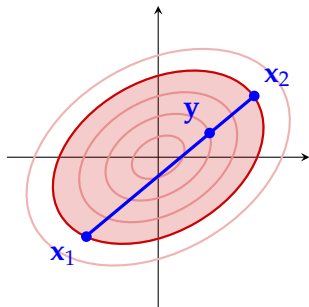
$f(\mathbf{x}_1), f(\mathbf{x}_2) \leq c$. Sei

$\mathbf{y} = (1 - h)\mathbf{x}_1 + h\mathbf{x}_2$ für ein $h \in [0, 1]$.

Dann ist

$$\begin{aligned} f(\mathbf{y}) &= f((1 - h)\mathbf{x}_1 + h\mathbf{x}_2) \\ &\leq (1 - h)f(\mathbf{x}_1) + hf(\mathbf{x}_2) \\ &\leq (1 - h)c + hc = c \end{aligned}$$

Also $\mathbf{y} \in \{\mathbf{x} \in D_f : f(\mathbf{x}) \leq c\}$.



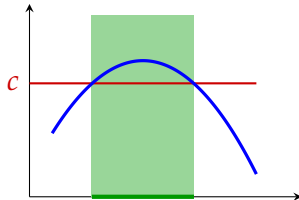
Konkave Funktion und konvexe Menge

Sei f konkav.

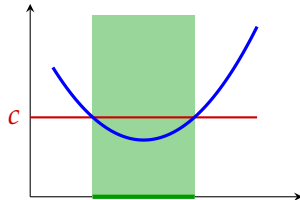
Dann ist die **obere Niveaumenge** von f

$$\{\mathbf{x} \in D_f : f(\mathbf{x}) \geq c\}$$

konvex.



obere Niveaumenge



untere Niveaumenge

Extremum und monotone Transformation

Sei $T: \mathbb{R} \rightarrow \mathbb{R}$ eine *streng monoton steigende* Funktion.

Falls \mathbf{x}^* ein *Maximum* (Minimum) von f ist,
dann ist \mathbf{x}^* auch ein Maximum (bzw. Minimum) von $T \circ f$:

Da \mathbf{x}^* ein *Maximum* von f ist, gilt

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) \text{ f\"ur alle } \mathbf{x}.$$

Da T streng monoton steigend ist, gilt

$$T(x_1) > T(x_2) \text{ falls } x_1 > x_2.$$

Wir erhalten daher

$$(T \circ f)(\mathbf{x}^*) = T(f(\mathbf{x}^*)) > T(f(\mathbf{x})) = (T \circ f)(\mathbf{x}) \text{ f\"ur alle } \mathbf{x},$$

i.e., \mathbf{x}^* ist ein Maximum von $T \circ f$.

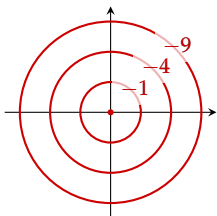
Da T injektiv ist, gilt sogar die Umkehrung:

Falls \mathbf{x}^* ein *Maximum* (Minimum) von $T \circ f$ ist, dann auch von f .

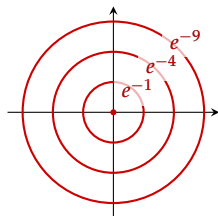
Extremum und monotone Transformation

Die streng monotone Transformation T erhält die Extrema von f .

T erhält aber auch die Niveaumengen von f .



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



$$T(f(x, y)) = \exp(-x^2 - y^2)$$

Quasi-konvex und quasi-konkav

Eine Funktion f heißt **quasi-konvex** in $D \subseteq \mathbb{R}^n$, falls D *konvex* ist und jede *untere Niveaumenge* $\{\mathbf{x} \in D_f : f(\mathbf{x}) \leq c\}$ *konvex* ist.

Eine Funktion f heißt **quasi-konkav** in $D \subseteq \mathbb{R}^n$, falls D *konvex* ist und jede *obere Niveaumenge* $\{\mathbf{x} \in D_f : f(\mathbf{x}) \geq c\}$ *konvex* ist.

Konvex und quasi-konvex

Jede *konkave* (konvexe) Funktion ist auch *quasi-konkav* (bzw. quasi-konvex).

Eine quasi-konkave Funktion muss aber nicht konkav sein.

Sei T eine streng monoton steigende Funktion. Falls eine Funktion $f(x)$ *konkav* (konvex) ist, dann ist $T \circ f$ *quasi-konkav* (bzw. quasi-konvex).

Die Funktion $g(x, y) = e^{-x^2-y^2}$ ist quasi-konkav, da $f(x, y) = -x^2 - y^2$ konkav ist, und $T(x) = e^x$ streng monoton steigend ist.

$g = T \circ f$ ist aber nicht konkav.

Konvex, quasi-konvex und Extrema

Der Begriff *quasi-konvex* ist ein **schwächerer** Begriff als *konvex*, in dem Sinne, dass jede konvexe Funktion auch quasi-konvex ist, es aber viel mehr quasi-konvexe Funktionen gibt als konvexe.

Die Bedeutung dieses Begriffs liegt darin, dass sich manche Sätze über konvexe Funktionen auf quasi-konvexe Funktionen verallgemeinern lassen.

Quasi-konvex und quasi-konkav II

- ▶ f ist *quasi-konvex* genau dann, wenn

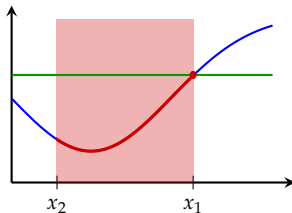
$$f((1-h)\mathbf{x}_1 + h\mathbf{x}_2) \leq \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$$

für alle $\mathbf{x}_1, \mathbf{x}_2$ und $h \in [0, 1]$.

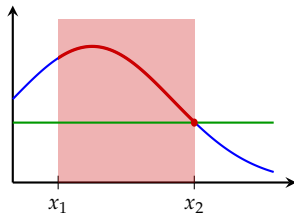
- ▶ f ist *quasi-konkav* genau dann, wenn

$$f((1-h)\mathbf{x}_1 + h\mathbf{x}_2) \geq \min\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$$

für alle $\mathbf{x}_1, \mathbf{x}_2$ und $h \in [0, 1]$.



quasi-konvex



quasi-konkav

Streng quasi-konvex und streng quasi-konkav

- ▶ Eine Funktion f heißt **streng quasi-konvex**, wenn

$$f((1-h)\mathbf{x}_1 + h\mathbf{x}_2) < \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$$

für alle $\mathbf{x}_1, \mathbf{x}_2$, mit $\mathbf{x}_1 \neq \mathbf{x}_2$, und $h \in (0, 1)$.

- ▶ Eine Funktion f heißt **streng quasi-konkav**, wenn

$$f((1-h)\mathbf{x}_1 + h\mathbf{x}_2) > \min\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$$

für alle $\mathbf{x}_1, \mathbf{x}_2$, mit $\mathbf{x}_1 \neq \mathbf{x}_2$, und $h \in (0, 1)$.

Quasi-konvex und quasi-konkav III

Für eine differenzierbare Funktion f gilt:

- ▶ f ist *quasi-konvex* genau dann, wenn

$$f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad \Rightarrow \quad \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \leq 0$$

- ▶ f ist *quasi-konkav* genau dann, wenn

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) \quad \Rightarrow \quad \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \geq 0$$

Beispiel

Seien $p, r > 0$ und

$$f: D = [0, \infty)^2 \rightarrow \mathbb{R}, \quad f(x, y) = 4x^{\frac{1}{4}}y^{\frac{1}{4}} - px - ry$$

$$\text{Hesse-Matrix: } \mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} -\frac{3}{4}x^{-\frac{7}{4}}y^{\frac{1}{4}} & \frac{1}{4}x^{-\frac{3}{4}}y^{-\frac{3}{4}} \\ \frac{1}{4}x^{-\frac{3}{4}}y^{-\frac{3}{4}} & -\frac{3}{4}x^{\frac{1}{4}}y^{-\frac{7}{4}} \end{pmatrix}$$

Hauptminoren:

$$H_1 = -\frac{3}{4}x^{-\frac{7}{4}}y^{\frac{1}{4}} < 0$$

$$H_2 = \frac{1}{2}x^{-\frac{3}{2}}y^{-\frac{3}{2}} > 0$$

f ist streng konkav in D .

$$\text{Kritischer Punkt: } \nabla f = (x^{-\frac{3}{4}}y^{\frac{1}{4}} - p, x^{\frac{1}{4}}y^{-\frac{3}{4}} - r) = 0$$

$$f_x = x^{-\frac{3}{4}}y^{\frac{1}{4}} - p = 0$$

$$f_y = x^{\frac{1}{4}}y^{-\frac{3}{4}} - r = 0$$

$$\Rightarrow \mathbf{x}_0 = \left(\sqrt{\frac{1}{rp^3}}, \sqrt{\frac{1}{r^3p}} \right)$$

\mathbf{x}_0 ist das globale Maximum von f .

Frage:

Wie ändert sich das Optimum $f^* = f(\mathbf{x}_0)$ mit den Parametern r und p ?

Umhüllungssatz (Envelope-Theorem)

Gegeben sei eine Funktion

$$\mathbf{x} \mapsto f(\mathbf{x}, \mathbf{r}) \quad \mathbf{x} = (x_1, \dots, x_n) \dots \text{Variable} \quad (\text{endogen})$$

$$\mathbf{r} = (r_1, \dots, r_k) \dots \text{Parameter} \quad (\text{exogen})$$

mit Extremum \mathbf{x}^* .

Das Extremum hängt vom Parameter \mathbf{r} ab:

$$\mathbf{x}^* = \mathbf{x}^*(\mathbf{r})$$

und damit auch der Extremalwert f^* :

$$f^*(\mathbf{r}) = f(\mathbf{x}^*(\mathbf{r}), \mathbf{r})$$

Es gilt:

$$\frac{\partial f^*(\mathbf{r})}{\partial r_j} = \frac{\partial f(\mathbf{x}, \mathbf{r})}{\partial r_j} \Big|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})}$$

Umhüllungssatz (Envelope-Theorem)

$$\begin{aligned}\frac{\partial f^*(\mathbf{r})}{\partial r_j} &= \left. \frac{\partial f(\mathbf{x}^*(\mathbf{r}), \mathbf{r})}{\partial r_j} \right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})} && \text{[Kettenregel]} \\ &= \sum_{i=1}^n \underbrace{f_{x_i}(\mathbf{x}^*(\mathbf{r}), \mathbf{r})}_{=0 \text{ da kritischer Punkt}} \cdot \frac{\partial x_i^*(\mathbf{r})}{\partial r_j} + \left. \frac{\partial f(\mathbf{x}, \mathbf{r})}{\partial r_j} \right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})} \\ &= \left. \frac{\partial f(\mathbf{x}, \mathbf{r})}{\partial r_j} \right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})}\end{aligned}$$

Beispiel

Das (eindeutig bestimmte) Maximum von

$$f: D = [0, \infty)^2 \rightarrow \mathbb{R}, \quad f(x, y) = 4x^{\frac{1}{4}}y^{\frac{1}{4}} - px - ry$$

$$\text{ist } \mathbf{x}^*(p, r) = (x^*(p, r), y^*(p, r)) = \left(\sqrt{\frac{1}{rp^3}}, \sqrt{\frac{1}{r^3p}} \right).$$

Frage:

Wie ändert sich das Optimum $f^* = f(\mathbf{x}^*)$ mit den Parametern r und p ?

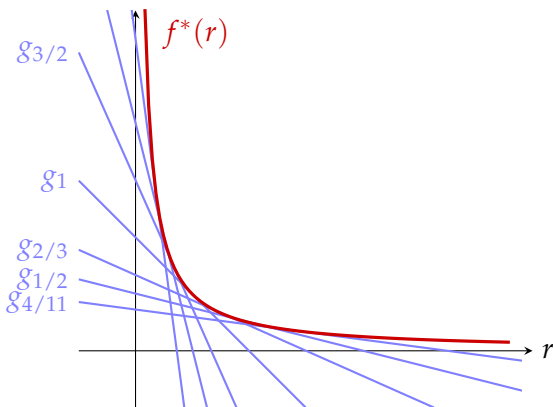
$$\frac{\partial f^*(p, r)}{\partial p} = \frac{\partial f(\mathbf{x}; p, r)}{\partial p} \Big|_{\mathbf{x}=\mathbf{x}^*(p, r)} = -x \Big|_{\mathbf{x}=\mathbf{x}^*(p, r)} = -\sqrt{\frac{1}{rp^3}}$$

$$\frac{\partial f^*(p, r)}{\partial r} = \frac{\partial f(\mathbf{x}; p, r)}{\partial r} \Big|_{\mathbf{x}=\mathbf{x}^*(p, r)} = -y \Big|_{\mathbf{x}=\mathbf{x}^*(p, r)} = -\sqrt{\frac{1}{r^3p}}$$

Eine geometrische Interpretation

Sei $f(x, r) = \sqrt{x} - rx$. Wir suchen $f^*(r) = \max_x f(x, r)$.

Zeichnen die Graphen von $g_x(r) = f(x, r)$ für verschiedene x .



Zusammenfassung

- ▶ Konvexe Menge
- ▶ Konvex und konkav
- ▶ Quasi-konvex und quasi-konkav
- ▶ Lokales und globales Extremum
- ▶ Minimum, Maximum und Sattelpunkt
- ▶ Kritischer Punkt
- ▶ Hesse-Matrix und Hauptminor
- ▶ Umhüllungssatz (Envelope-Theorem)

Kapitel 12

Lagrange-Funktion

Optimierung unter Nebenbedingungen

Aufgabe:

Berechne die Extrema der Funktion

$$f(x, y)$$

unter der Nebenbedingung

$$g(x, y) = c$$

Beispiel:

Wir suchen die lokalen Extrema der Funktion

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

unter der Nebenbedingung

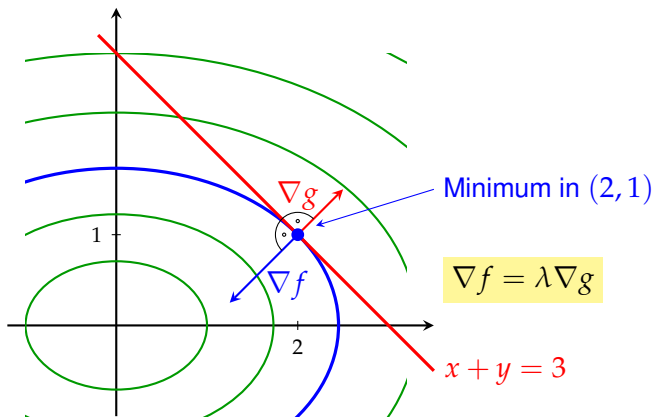
$$g(x, y) = x + y = 3$$

Graphische Lösung

Im Falle von zwei Variablen können wir das Problem graphisch „lösen“.

1. Zeichne die Nebenbedingung $g(x, y) = c$ in die xy -Ebene ein.
(Kurve in der Ebene)
2. Zeichne „geeignete“ Niveaulinien der zu optimierenden Funktion $f(x, y)$ ein.
3. Untersuche an Hand der Zeichnung welche Niveaulinien den zulässigen Bereich schneiden und bestimme die ungefähre Lage der Extrema.

Beispiel – Graphische Lösung



Extrema von $f(x, y) = x^2 + 2y^2$ gegeben $g(x, y) = x + y = 3$

Lagrange-Ansatz

Sei \mathbf{x}^* ein Extremum von $f(x, y)$ gegeben $g(x, y) = c$.
Dann müssen $\nabla f(\mathbf{x}^*)$ und $\nabla g(\mathbf{x}^*)$ proportional sein:

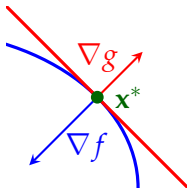
$$\nabla f(\mathbf{x}^*) = \lambda \nabla g(\mathbf{x}^*)$$

wobei λ eine geeignete Proportionalitätskonstante ist.

$$f_x(\mathbf{x}^*) = \lambda g_x(\mathbf{x}^*)$$

$$f_y(\mathbf{x}^*) = \lambda g_y(\mathbf{x}^*)$$

$$g(\mathbf{x}^*) = c$$



Umformen ergibt

$$f_x(\mathbf{x}^*) - \lambda g_x(\mathbf{x}^*) = 0$$

$$f_y(\mathbf{x}^*) - \lambda g_y(\mathbf{x}^*) = 0$$

$$c - g(\mathbf{x}^*) = 0$$

Linke Seite ist Gradient von $\mathcal{L}(x, y; \lambda) = f(x, y) + \lambda (c - g(x, y))$

Lagrange-Funktion

Wir erzeugen uns aus f , g und einer Hilfsvariablen λ eine neue Funktion, die **Lagrange-Funktion**:

$$\mathcal{L}(x, y; \lambda) = f(x, y) + \lambda (c - g(x, y))$$

Die Hilfsvariable λ heißt **Lagrange-Multiplikator**.

Lokale Extrema von f gegeben $g(x, y) = c$ sind kritische Punkte der Lagrangefunktion \mathcal{L} :

$$\mathcal{L}_x = f_x - \lambda g_x = 0$$

$$\mathcal{L}_y = f_y - \lambda g_y = 0$$

$$\mathcal{L}_\lambda = c - g(x, y) = 0$$

Beispiel – Lagrangefunktion

Wir suchen die lokalen Extrema der

$$f(x, y) = x^2 + 2y^2 \quad \text{gegeben} \quad g(x, y) = x + y = 3$$

Lagrangefunktion:

$$\mathcal{L}(x, y, \lambda) = (x^2 + 2y^2) + \lambda(3 - (x + y))$$

Kritische Punkte:

$$\mathcal{L}_x = 2x - \lambda = 0$$

$$\mathcal{L}_y = 4y - \lambda = 0$$

$$\mathcal{L}_\lambda = 3 - x - y = 0$$

⇒ einziger kritischer Punkt: $(\mathbf{x}_0; \lambda_0) = (2, 1; 4)$

Geränderte Hesse-Matrix

Die Matrix

$$\bar{\mathbf{H}}(\mathbf{x}; \lambda) = \begin{pmatrix} 0 & g_x & g_y \\ g_x & \mathcal{L}_{xx} & \mathcal{L}_{xy} \\ g_y & \mathcal{L}_{yx} & \mathcal{L}_{yy} \end{pmatrix}$$

heißt **geränderte Hesse-Matrix**.

Hinreichende Bedingung für lokales Extremum:

Sei $(\mathbf{x}_0; \lambda_0)$ ein kritischer Punkt von \mathcal{L} .

- ▶ $|\bar{\mathbf{H}}(\mathbf{x}_0; \lambda_0)| > 0 \Rightarrow \mathbf{x}_0$ ist *lokales Maximum*
- ▶ $|\bar{\mathbf{H}}(\mathbf{x}_0; \lambda_0)| < 0 \Rightarrow \mathbf{x}_0$ ist *lokales Minimum*
- ▶ $|\bar{\mathbf{H}}(\mathbf{x}_0; \lambda_0)| = 0 \Rightarrow$ keine Aussage möglich

Beispiel – Geränderte Hesse-Matrix

Wir suchen die lokalen Extrema der

$$f(x, y) = x^2 + 2y^2 \quad \text{gegeben} \quad g(x, y) = x + y = 3$$

Lagrangefunktion: $\mathcal{L}(x, y, \lambda) = (x^2 + 2y^2) + \lambda(3 - x - y)$

Kritischer Punkt: $(\mathbf{x}_0; \lambda_0) = (2, 1; 4)$

Determinante der geränderten Hesse-Matrix:

$$|\bar{\mathbf{H}}(\mathbf{x}_0; \lambda_0)| = \begin{vmatrix} 0 & g_x & g_y \\ g_x & \mathcal{L}_{xx} & \mathcal{L}_{xy} \\ g_y & \mathcal{L}_{yx} & \mathcal{L}_{yy} \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 4 \end{vmatrix} = -6 < 0$$

$\Rightarrow \mathbf{x}_0 = (2, 1)$ ist ein lokales Minimum.

Viele Variablen und Gleichungen

Berechne die Extrema der Funktion

$$f(x_1, \dots, x_n)$$

unter den Nebenbedingungen

$$g_1(x_1, \dots, x_n) = c_1$$

$$\vdots$$

$$(k < n)$$

$$g_k(x_1, \dots, x_k) = c_k$$

Optimierungsproblem: $\min / \max f(\mathbf{x})$ gegeben $\mathbf{g}(\mathbf{x}) = \mathbf{c}$.

Lagrange-Funktion

$$\mathcal{L}(\mathbf{x}; \lambda_1, \dots, \lambda_k) = f(\mathbf{x}) + \sum_{i=1}^k \lambda_i (c_i - g_i(\mathbf{x}))$$

Vorgangsweise – Kritische Punkte

1. Stelle Lagrange-Funktion \mathcal{L} auf:

$$\begin{aligned}\mathcal{L}(x_1, \dots, x_n; \lambda_1, \dots, \lambda_k) \\ = f(x_1, \dots, x_n) + \sum_{i=1}^k \lambda_i (c_i - g_i(x_1, \dots, x_n))\end{aligned}$$

2. Berechne die ersten partiellen Ableitungen von \mathcal{L} .
3. Setze alle ersten partiellen Ableitungen gleich Null und löse das so entstandene Gleichungssystem mit $n + k$ Unbekannten in $n + k$ Gleichungen.
4. Die ersten n Komponenten (x_1, \dots, x_n) sind die Koordinaten der gesuchten kritischen Punkte.

Beispiel – Kritische Punkte

Wir suchen die kritischen Punkte von

$$f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 - 2)^2 + 2x_3^2$$

unter den Nebenbedingungen

$$x_1 + 2x_2 = 2 \quad \text{und} \quad x_2 - x_3 = 3$$

Lagrange-Funktion:

$$\begin{aligned} \mathcal{L}(x_1, x_2, x_3; \lambda_1, \lambda_2) = & ((x_1 - 1)^2 + (x_2 - 2)^2 + 2x_3^2) \\ & + \lambda_1(2 - x_1 - 2x_2) + \lambda_2(3 - x_2 + x_3) \end{aligned}$$

Beispiel – Kritische Punkte

Partielle Ableitungen (Gradient):

$$\mathcal{L}_{x_1} = 2(x_1 - 1) - \lambda_1 = 0$$

$$\mathcal{L}_{x_2} = 2(x_2 - 2) - 2\lambda_1 - \lambda_2 = 0$$

$$\mathcal{L}_{x_3} = 4x_3 + \lambda_2 = 0$$

$$\mathcal{L}_{\lambda_1} = 2 - x_1 - 2x_2 = 0$$

$$\mathcal{L}_{\lambda_2} = 3 - x_2 + x_3 = 0$$

Die kritischen Punkte von \mathcal{L} erhalten wir durch Nullsetzen der ersten partiellen Ableitungen: (lineares Gleichungssystem)

$$x_1 = -\frac{6}{7}, x_2 = \frac{10}{7}, x_3 = -\frac{11}{7}; \lambda_1 = -\frac{26}{7}, \lambda_2 = \frac{44}{7}.$$

Der einzige kritische Punkt von f unter dieser Nebenbedingungen ist somit

$$\mathbf{x}_0 = \left(-\frac{6}{7}, \frac{10}{7}, -\frac{11}{7}\right).$$

Geränderte Hesse-Matrix

$$\bar{\mathbf{H}}(\mathbf{x}; \boldsymbol{\lambda}) = \begin{pmatrix} 0 & \cdots & 0 & \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{\partial g_k}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_1} & \mathcal{L}_{x_1 x_1} & \cdots & \mathcal{L}_{x_1 x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial x_n} & \cdots & \frac{\partial g_k}{\partial x_n} & \mathcal{L}_{x_n x_1} & \cdots & \mathcal{L}_{x_n x_n} \end{pmatrix}$$

Sei $B_r(\mathbf{x}; \boldsymbol{\lambda})$ der $(k + r)$ -te führende Hauptminor von $\bar{\mathbf{H}}(\mathbf{x}; \boldsymbol{\lambda})$.

Hinreichende Bedingung

Sei $(\mathbf{x}_0; \lambda_0)$ ein kritischer Punkt von \mathcal{L} .

- ▶ $(-1)^k B_r(\mathbf{x}_0; \lambda_0) > 0$ für alle $r = k + 1, \dots, n$
 $\Rightarrow \mathbf{x}_0$ ist *lokales Minimum*
- ▶ $(-1)^r B_r(\mathbf{x}_0; \lambda_0) > 0$ für alle $r = k + 1, \dots, n$
 $\Rightarrow \mathbf{x}_0$ ist *lokales Maximum*

(n ist die Anzahl der Variablen x_i und
 k ist die Anzahl der Nebenbedingungen.)

Beispiel – Hinreichende Bedingung

Suchen Extrema von $f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 - 2)^2 + 2x_3^2$
unter den Nebenbedingungen $x_1 + 2x_2 = 2$ und $x_2 - x_3 = 3$

Lagrange-Funktion:

$$\mathcal{L}(x_1, x_2, x_3; \lambda_1, \lambda_2) = ((x_1 - 1)^2 + (x_2 - 2)^2 + 2x_3^2) \\ + \lambda_1(2 - x_1 - 2x_2) + \lambda_2(3 - x_2 + x_3)$$

Kritischer Punkt von \mathcal{L} :

$$x_1 = -\frac{6}{7}, x_2 = \frac{10}{7}, x_3 = -\frac{11}{7}; \lambda_1 = -\frac{26}{7}, \lambda_2 = \frac{44}{7}.$$

Beispiel – Hinreichende Bedingung

Geränderte Hesse-Matrix:

$$\tilde{\mathbf{H}}(\mathbf{x}; \lambda) = \begin{pmatrix} 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 2 & 0 & 0 \\ 2 & 1 & 0 & 2 & 0 \\ 0 & -1 & 0 & 0 & 4 \end{pmatrix}$$

3 Variablen, 2 Nebenbedingungen: $n = 3, k = 2 \Rightarrow r = 3$

$$B_3 = |\tilde{\mathbf{H}}(\mathbf{x}; \lambda)| = 14$$

$$(-1)^k B_r = (-1)^2 B_3 = 14 > 0 \quad \text{Bedingung erfüllt}$$

$$(-1)^r B_r = (-1)^3 B_3 = -14 < 0 \quad \text{nicht erfüllt}$$

Der kritische Punkt $\mathbf{x}_0 = \left(-\frac{6}{7}, \frac{10}{7}, -\frac{11}{7}\right)$ ist ein *lokales Minimum*.

Globale Extrema

Sei $(\mathbf{x}^*, \lambda^*)$ ein kritischer Punkt der Lagrange-Funktion \mathcal{L} des Optimierungsproblems

$$\min / \max f(\mathbf{x}) \text{ gegeben } \mathbf{g}(\mathbf{x}) = \mathbf{c}$$

Falls $\mathcal{L}(\mathbf{x}, \lambda^*)$ *konkav* (konvex) in \mathbf{x} ist, dann ist \mathbf{x}^* ein **globales Maximum** (globales Minimum) von $f(\mathbf{x})$ gegeben $\mathbf{g}(\mathbf{x}) = \mathbf{c}$.

Beispiel – Globale Extrema

$(x^*, y^*; \lambda^*) = (2, 1; 4)$ ist ein kritischer Punkt der Lagrange-Funktion des Optimierungsproblems

$$\min / \max f(x, y) = x^2 + 2y^2 \quad \text{gegeben} \quad g(x, y) = x + y = 3$$

Lagrangefunktion:

$$\mathcal{L}(x, y, \lambda^*) = (x^2 + 2y^2) + 4 \cdot (3 - (x + y))$$

Hesse-Matrix:

$$\mathbf{H}_{\mathcal{L}}(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} \quad \begin{array}{l} H_1 = 2 > 0 \\ H_2 = 8 > 0 \end{array}$$

\mathcal{L} ist konvex in (x, y) .

$(x^*, y^*) = (2, 1)$ ist ein globales Minimum.

Beispiel – Globale Extrema

$$(\mathbf{x}^*; \boldsymbol{\lambda}^*) = \left(-\frac{6}{7}, \frac{10}{7}, -\frac{11}{7}; -\frac{26}{7}, \frac{44}{7}\right)$$

ist ein kritischer Punkt der Lagrange-Funktion des Optimierungsproblems

$$\min / \max \quad f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 - 2)^2 + 2x_3^2$$

$$\text{gegeben} \quad g_1(x_1, x_2, x_3) = x_1 + 2x_2 = 2$$

$$g_2(x_1, x_2, x_3) = x_2 - x_3 = 3$$

Lagrangefunktion:

$$\begin{aligned} \mathcal{L}(\mathbf{x}; \boldsymbol{\lambda}^*) = & ((x_1 - 1)^2 + (x_2 - 2)^2 + 2x_3^2) \\ & - \frac{26}{7} (2 - x_1 - 2x_2) + \frac{44}{7} (3 - x_2 + x_3) \end{aligned}$$

Beispiel – Globale Extrema

Hesse-Matrix:

$$\mathbf{H}_{\mathcal{L}}(x, y) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad \begin{array}{l} H_1 = 2 > 0 \\ H_2 = 4 > 0 \\ H_3 = 16 > 0 \end{array}$$

\mathcal{L} ist konvex in \mathbf{x} .

$\mathbf{x}^* = \left(-\frac{6}{7}, \frac{10}{7}, -\frac{11}{7}\right)$ ist ein globales Minimum.

Interpretation des Lagrange-Multiplikators

Die Lage des Extremums \mathbf{x}^* des Optimierungsproblems

$$\min / \max f(\mathbf{x}) \text{ gegeben } \mathbf{g}(\mathbf{x}) = \mathbf{c}$$

hängt von \mathbf{c} ab, $\mathbf{x}^* = \mathbf{x}^*(\mathbf{c})$, und somit auch der Extremalwert von f :

$$f^*(\mathbf{c}) = f(\mathbf{x}^*(\mathbf{c}))$$

Wie ändert sich $f^*(\mathbf{c})$ mit \mathbf{c} ?

$$\frac{\partial f^*}{\partial c_j}(\mathbf{c}) = \lambda_j^*(\mathbf{c})$$

Der Lagrange-Multiplikator λ_j gibt also an, wie sich der Extremalwert ändert, wenn die Konstante c_j der Nebenbedingung $g_j(\mathbf{x}) = c_j$ verändert wird.

Herleitung

Im Optimum stimmen \mathcal{L} und f überein. Daher ist

$$\begin{aligned}\frac{\partial f^*(\mathbf{c})}{\partial c_j} &= \frac{\partial \mathcal{L}(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}(\mathbf{c}))}{\partial c_j} \quad [\text{Kettenregel}] \\ &= \sum_{i=1}^n \underbrace{\mathcal{L}_{x_i}(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}(\mathbf{c}))}_{=0 \text{ da kritischer Punkt}} \cdot \frac{\partial x_i^*(\mathbf{c})}{\partial c_j} + \left. \frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{c})}{\partial c_j} \right|_{(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}^*(\mathbf{c}))} \\ &= \left. \frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{c})}{\partial c_j} \right|_{(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}^*(\mathbf{c}))} \\ &= \left. \frac{\partial}{\partial c_j} \left(f(\mathbf{x}) + \sum_{i=1}^k \lambda_i (c_i - g_i(\mathbf{x})) \right) \right|_{(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}^*(\mathbf{c}))} \\ &= \lambda_j^*(\mathbf{c})\end{aligned}$$

Beispiel – Lagrange-Multiplikator

$(x^*, y^*) = (2, 1)$ ist ein Minimum des Optimierungsproblems

min / max $f(x, y) = x^2 + 2y^2$ gegeben $g(x, y) = x + y = c = 3$
mit $\lambda^* = 4$.

Wie ändert sich der Minimalwert f^* von f , wenn sich c ändert?

$$\frac{df^*}{dc} = \lambda^* = 4$$

Umhüllungssatz (Envelope-Theorem)

Wie ändert sich das Extremum f^* des Optimierungsproblems

$$\min / \max f(\mathbf{x}, \mathbf{p}) \text{ gegeben } \mathbf{g}(\mathbf{x}, \mathbf{p}) = \mathbf{c}$$

wenn sich der Parameter (die exogene Variable) \mathbf{p} ändert?

$$\frac{\partial f^*(\mathbf{p})}{\partial p_j} = \frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{p})}{\partial p_j} \Big|_{(\mathbf{x}^*(\mathbf{p}), \lambda^*(\mathbf{p}))}$$

Beispiel – Roys Identität

Maximiere Nutzenfunktion

$$\max U(\mathbf{x}) \quad \text{gegeben} \quad \mathbf{p}^t \cdot \mathbf{x} = m$$

Maximaler Nutzen hängt von Preisen \mathbf{p} und Einkommen m ab:

$$U^* = U^*(\mathbf{p}, m) \quad [\text{indirekte Nutzenfunktion}]$$

Lagrange-Funktion $\mathcal{L}(\mathbf{x}, \lambda) = U(\mathbf{x}) + \lambda (m - \mathbf{p}^t \cdot \mathbf{x})$

$$\frac{\partial U^*}{\partial p_j} = \frac{\partial \mathcal{L}}{\partial p_j} = -\lambda^* x_j^* \quad \text{und} \quad \frac{\partial U^*}{\partial m} = \frac{\partial \mathcal{L}}{\partial m} = \lambda^*$$

und somit

$$x_j^* = - \frac{\partial U^* / \partial p_j}{\partial U^* / \partial m} \quad [\text{Marshallsche Nachfragefunktion}]$$

Beispiel – Shephards Lemma

Minimiere Ausgaben

$$\min \mathbf{p}^t \cdot \mathbf{x} = m \quad \text{gegeben} \quad U(\mathbf{x}) = \bar{u}$$

(Minimale) Ausgabenfunktion hängt von \mathbf{p} und \bar{u} ab: $e = e(\mathbf{p}, \bar{u})$

Lagrange-Funktion $\mathcal{L}(\mathbf{x}, \lambda) = \mathbf{p}^t \cdot \mathbf{x} + \lambda (\bar{u} - U(\mathbf{x}))$

$$\frac{\partial e}{\partial p_j} = \frac{\partial \mathcal{L}}{\partial p_j} = x_j^* \quad [\text{Hicks'sche Nachfragefunktion}]$$

Zusammenfassung

- ▶ Optimierung unter Nebenbedingungen
- ▶ Graphische Lösung
- ▶ Lagrange-Funktion und Lagrange-Multiplikator
- ▶ Extremum und kritischer Punkt
- ▶ Geränderte Hesse-Matrix
- ▶ Globale Extrema
- ▶ Interpretation des Lagrange-Multiplikators
- ▶ Umhüllungssatz

Kapitel 13

Kuhn-Tucker Bedingung

Optimierung unter Nebenbedingungen

Aufgabe:

Berechne das Maximum der Funktion

$$f(x, y)$$

unter den Nebenbedingungen

$$g(x, y) \leq c, \quad x, y \geq 0$$

Beispiel:

Wir suchen das Maximum von

$$f(x, y) = -(x - 5)^2 - (y - 5)^2$$

unter den Nebenbedingungen

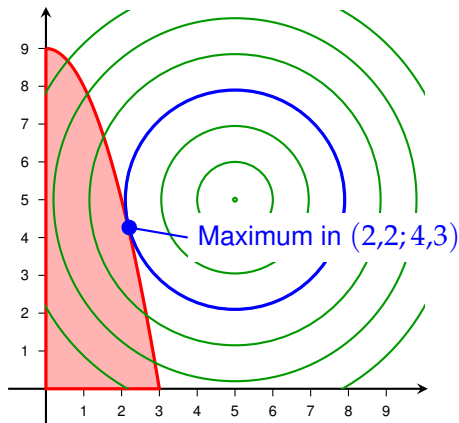
$$x^2 + y \leq 9, \quad x, y \geq 0$$

Graphische Lösung

Im Falle von zwei Variablen können wir das Problem graphisch „lösen“.

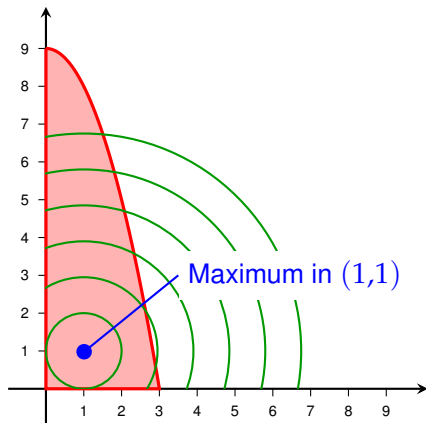
1. Zeichne die Nebenbedingung $g(x, y) \leq c$ in die xy -Ebene ein.
(Fläche in der Ebene)
2. Zeichne „geeignete“ Niveaulinien der zu optimierenden Funktion $f(x, y)$ ein.
3. Untersuche an Hand der Zeichnung welche Niveaulinien den zulässigen Bereich schneiden und bestimme die ungefähre Lage des Maximums.

Beispiel – Graphische Lösung



Maximum von $f(x,y) = -(x-5)^2 - (y-5)^2$
gegeben $g(x,y) = x^2 + y \leq 9, \quad x, y \geq 0.$

Beispiel – Graphische Lösung



Maximum von $f(x, y) = -(x - 1)^2 - (y - 1)^2$
gegeben $g(x, y) = x^2 + y \leq 9, \quad x, y \geq 0.$

Optimierung unter Nebenbedingungen

Berechne das Maximum der Funktion

$$f(x_1, \dots, x_n)$$

unter den Nebenbedingungen

$$g_1(x_1, \dots, x_n) \leq c_1$$

$$\vdots$$

$$g_k(x_1, \dots, x_n) \leq c_k$$

$$x_1, \dots, x_n \geq 0 \quad \text{(Nichtnegativitätsbedingung)}$$

Optimierungsproblem:

$$\max f(\mathbf{x}) \quad \text{gegeben} \quad \mathbf{g}(\mathbf{x}) \leq \mathbf{c} \quad \text{und} \quad \mathbf{x} \geq 0.$$

Nichtnegativitätsbedingung

Funktion f in einer Variable mit Nichtnegativitätsbedingung.

Für ein Maximum x^* gilt:

- ▶ x^* liegt im Inneren des zulässigen Bereichs:

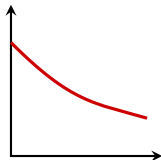
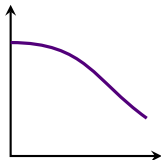
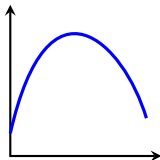
$$x^* > 0 \text{ und } f'(x^*) = 0.$$

- ▶ x^* liegt am Rand:

$$x^* = 0 \text{ und } f'(x^*) \leq 0.$$

Zusammengefasst:

$$f'(x^*) \leq 0, \quad x^* \geq 0 \quad \text{und} \quad x^* f'(x^*) = 0$$



Nichtnegativitätsbedingung

Im Falle einer Funktion $f(\mathbf{x})$ in mehreren Variablen erhalten wir für jede Variable x_j so eine Bedingung:

$$f_{x_j}(x^*) \leq 0, \quad x_j^* \geq 0 \quad \text{und} \quad x_j^* f_{x_j}(x^*) = 0$$

Schlupfvariable

Maximiere

$$f(x_1, \dots, x_n)$$

unter den Nebenbedingungen

$$g_1(x_1, \dots, x_n) + \mathbf{s}_1 = c_1$$

\vdots

$$g_k(x_1, \dots, x_n) + \mathbf{s}_k = c_k$$

$$x_1, \dots, x_n \geq 0$$

$$\mathbf{s}_1, \dots, \mathbf{s}_k \geq 0 \quad (\text{neue Nichtnegativitätsbedingung})$$

Lagrange-Funktion:

$$\tilde{\mathcal{L}}(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}) = f(x_1, \dots, x_n) + \sum_{i=1}^k \lambda_i (c_i - g_i(x_1, \dots, x_n) - \mathbf{s}_i)$$

Schlupfvariable

$$\tilde{\mathcal{L}}(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}) = f(x_1, \dots, x_n) + \sum_{i=1}^k \lambda_i (c_i - g_i(x_1, \dots, x_n) - s_i)$$

Berücksichtigen der Nichtnegativitätsbedingung:

$$\frac{\partial \tilde{\mathcal{L}}}{\partial x_j} \leq 0, \quad x_j \geq 0 \quad \text{und} \quad x_j \frac{\partial \tilde{\mathcal{L}}}{\partial x_j} = 0$$

$$\frac{\partial \tilde{\mathcal{L}}}{\partial s_i} \leq 0, \quad s_i \geq 0 \quad \text{und} \quad s_i \frac{\partial \tilde{\mathcal{L}}}{\partial s_i} = 0$$

$$\frac{\partial \tilde{\mathcal{L}}}{\partial \lambda_i} = 0 \quad \text{(keine Nichtnegativitätsbedingung)}$$

Eliminieren der Schlupfvariablen

Wegen $\frac{\partial \tilde{\mathcal{L}}}{\partial s_i} = -\lambda_i$ wird die zweite Zeile zu

$$\lambda_i \geq 0, \quad s_i \geq 0 \quad \text{und} \quad \lambda_i s_i = 0$$

Aus $\frac{\partial \tilde{\mathcal{L}}}{\partial \lambda_i} = c_i - g_i(\mathbf{x}) - s_i = 0$ folgt $s_i = c_i - g_i(\mathbf{x})$

und wir erhalten daher:

$$\lambda_i \geq 0, \quad c_i - g_i(\mathbf{x}) \geq 0 \quad \text{und} \quad \lambda_i (c_i - g_i(\mathbf{x})) = 0$$

Jetzt brauchen wir die Schlupfvariablen nicht mehr.

Eliminieren der Schlupfvariablen

Verwenden statt $\tilde{\mathcal{L}}$ die Lagrange-Funktion

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(x_1, \dots, x_n) + \sum_{i=1}^k \lambda_i (c_i - g_i(x_1, \dots, x_n))$$

Es gilt:

$$\frac{\partial \mathcal{L}}{\partial x_j} = \frac{\partial \tilde{\mathcal{L}}}{\partial x_j} \quad \text{und} \quad \frac{\partial \mathcal{L}}{\partial \lambda_i} = c_i - g_i(\mathbf{x})$$

Wir können daher die zweite Zeile der Bedingungen für ein lokales Maximum unter Nebenbedingungen schreiben als

$$\lambda_i \geq 0, \quad \frac{\partial \mathcal{L}}{\partial \lambda_i} \geq 0 \quad \text{und} \quad \lambda_i \frac{\partial \mathcal{L}}{\partial \lambda_i} = 0$$

Kuhn-Tucker Bedingung

$$\mathcal{L}(\mathbf{x}, \lambda) = f(x_1, \dots, x_n) + \sum_{i=1}^k \lambda_i (c_i - g_i(x_1, \dots, x_n))$$

Die **Kuhn-Tucker Bedingung** für ein (lokales) Maximum lauten:

$$\frac{\partial \mathcal{L}}{\partial x_j} \leq 0, \quad x_j \geq 0 \quad \text{und} \quad x_j \frac{\partial \mathcal{L}}{\partial x_j} = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} \geq 0, \quad \lambda_i \geq 0 \quad \text{und} \quad \lambda_i \frac{\partial \mathcal{L}}{\partial \lambda_i} = 0$$

Die Kuhn-Tucker Bedingung ist nicht hinreichend.
(Analog zu kritischen Punkten).

Beispiel

Wir suchen das Maximum von

$$f(x, y) = -(x - 5)^2 - (y - 5)^2$$

unter den Nebenbedingungen

$$x^2 + y \leq 9, \quad x, y \geq 0$$

Beispiel

Lagrange-Funktion:

$$\mathcal{L}(x, y; \lambda) = -(x - 5)^2 - (y - 5)^2 + \lambda(9 - x^2 - y)$$

Kuhn-Tucker-Bedingung:

$$(A) \quad \mathcal{L}_x = -2(x - 5) - 2\lambda x \leq 0$$

$$(B) \quad \mathcal{L}_y = -2(y - 5) - \lambda \leq 0$$

$$(C) \quad \mathcal{L}_\lambda = 9 - x^2 - y \geq 0$$

$$(N) \quad x, y, \lambda \geq 0$$

$$(I) \quad x \mathcal{L}_x = -x(2(x - 5) + 2\lambda x) = 0$$

$$(II) \quad y \mathcal{L}_y = -y(2(y - 5) + \lambda) = 0$$

$$(III) \quad \lambda \mathcal{L}_\lambda = \lambda(9 - x^2 - y) = 0$$

Beispiel

Schreiben (I)–(III) an als

$$(I) \quad x = 0 \quad \text{oder} \quad 2(x - 5) + 2\lambda x = 0$$

$$(II) \quad y = 0 \quad \text{oder} \quad 2(y - 5) + \lambda = 0$$

$$(III) \quad \lambda = 0 \quad \text{oder} \quad 9 - x^2 - y = 0$$

Müssen nun alle 8 Möglichkeiten ausrechnen, und überprüfen ob die entsprechenden Lösungen die Ungleichungen (A), (B), (C) und (N) erfüllen.

- Falls $\lambda = 0$ (III, links), dann gibt es wegen (I) und (II) vier Lösungen für $(x, y; \lambda)$:

$$(0,0;0), (5,0;0), (0,5;0) \text{ und } (5,5;0).$$

Keiner dieser Punkte erfüllt alle Ungleichungen (A), (B) und (C).

Daher: $\lambda \neq 0$.

Beispiel

Wenn $\lambda \neq 0$, dann gilt wegen (III, rechts): $y = 9 - x^2$.

- ▶ Wenn nun $\lambda \neq 0$ und $x = 0$, dann ist $y = 9$ und wegen (II, rechts), $\lambda = -8$. Ein Widerspruch zu (N).
- ▶ Wenn $\lambda \neq 0$ und $y = 0$, dann ist $x = 3$ and wegen (I, rechts), $\lambda = \frac{2}{3}$. Ein Widerspruch zu (B).
- ▶ Alle drei Variablen müssen daher ungleich 0 sein.

Daher ist $y = 9 - x^2$ und $\lambda = -2(y - 5) = -2(4 - x^2)$.

Eingesetzt in (I) erhalten wir $2(x - 5) - 4(4 - x^2)x = 0$ und

$$\begin{aligned}x &= \frac{\sqrt{11}+1}{2} \approx 2,158 & y &= \frac{12-\sqrt{11}}{2} \approx 4,342 \\ \lambda &= \sqrt{11} - 2 \approx 1,317\end{aligned}$$

Die Kuhn-Tucker-Bedingung wird daher nur vom Punkt

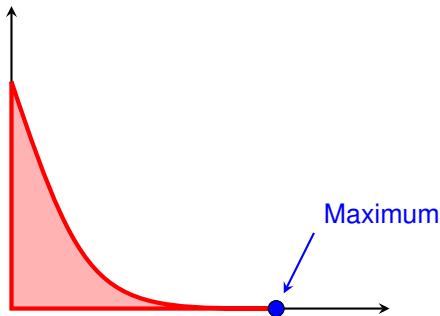
$$(x, y; \lambda) = \left(\frac{\sqrt{11}+1}{2}, \frac{12-\sqrt{11}}{2}; \sqrt{11} - 2 \right)$$

erfüllt.

Kuhn-Tucker Bedingung

Die Kuhn-Tucker Bedingung ist leider auch nicht notwendig!

D.h., es gibt Optimierungsprobleme, in denen das Maximum die Kuhn-Tucker Bedingung nicht erfüllt.



Der Satz von Kuhn-Tucker

Wir brauchen ein Werkzeug, um festzustellen, ob ein Punkt ein (globales) Maximum ist. Das ist aber nicht immer leicht.

Der **Satz von Kuhn-Tucker** gibt für einen Spezialfall eine *hinreichende* Bedingung:

- (1) Die Zielfunktion $f(\mathbf{x})$ sei differenzierbar und **konkav**.
- (2) Die Funktionen der Nebenbedingungen $g_i(\mathbf{x})$, $i = 1, \dots, k$, seien alle differenzierbar und **konvex**.
- (3) Der Punkt \mathbf{x}^* erfüllt die Kuhn-Tucker-Bedingung.

Dann ist \mathbf{x}^* ein *globales Maximum* von f unter den Nebenbedingungen $g_i \leq c_i$.

Das Maximum ist eindeutig, wenn die Funktion f *streng konkav* ist.

Beispiel

Wir suchen das Maximum von

$$f(x, y) = -(x - 5)^2 - (y - 5)^2$$

unter den Nebenbedingungen

$$x^2 + y \leq 9, \quad x, y \geq 0$$

Die Hessematrizen von $f(x, y)$ und $g(x, y) = x^2 + y$ lauten

$$\mathbf{H}_f = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \quad \text{und} \quad \mathbf{H}_g = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

Beispiel

$$\mathbf{H}_f = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \quad \text{und} \quad \mathbf{H}_g = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

- (1) f ist streng konkav.
- (2) g ist konvex.
- (3) Der Punkt $(x, y; \lambda) = \left(\frac{\sqrt{11}+1}{2}, \frac{12-\sqrt{11}}{2}; \sqrt{11} - 2 \right)$ erfüllt die Kuhn-Tucker-Bedingung.

Daher ist nach dem Satz von Kuhn-Tucker $\mathbf{x}^* = \left(\frac{\sqrt{11}+1}{2}, \frac{12-\sqrt{11}}{2} \right)$ das gesuchte globale Maximum.

Zusammenfassung

- ▶ Optimierung unter Nebenbedingungen
- ▶ Graphische Lösung
- ▶ Lagrange-Funktion
- ▶ Kuhn-Tucker Bedingung
- ▶ Satz von Kuhn-Tucker

Kapitel 14

Differentialgleichungen

Ein einfaches Modell (Domar)

Im Domar Wachstumsmodell treffen wir die folgenden Annahmen:

(1) Erhöhung der Investitionsrate $I(t)$ erhöht das Einkommen $Y(t)$:

$$\frac{dY}{dt} = \frac{1}{s} \cdot \frac{dI}{dt} \quad (s = \text{konstant})$$

(2) Verhältnis von Kapitalstock $K(t)$ zur Produktionskapazität $\kappa(t)$ sei konstant:

$$\frac{\kappa(t)}{K(t)} = \varrho \quad (= \text{konstant})$$

(E) In Gleichgewichtszustand gilt:

$$Y = \kappa$$

Frage: Welche Investitionsrate erhält das Modell für all Zeiten $t \geq 0$ im Gleichgewicht.

Ein einfaches Modell (Domar)

Wir suchen eine Funktion $I(t)$, die zu allen Zeiten die Modellvoraussetzungen und die Gleichgewichtsbedingung erfüllt.

$Y(t) = \kappa(t)$ für alle t impliziert, dass auch $Y'(t) = \kappa'(t)$.

Wir erhalten daher

$$\frac{1}{s} \cdot \frac{dI}{dt} \stackrel{(1)}{=} \frac{dY}{dt} \stackrel{(E)}{=} \frac{d\kappa}{dt} \stackrel{(2)}{=} \varrho \frac{dK}{dt} = \varrho I(t)$$

oder kurz

$$\frac{1}{s} \cdot \frac{dI}{dt} = \varrho I(t)$$

Die Gleichung enthält eine **Funktion** und deren **Ableitung** und muss für all t gelten. Die Unbekannte dieser Gleichung ist eine **Funktion**.

Differentialgleichung erster Ordnung

Eine **gewöhnliche Differentialgleichung (DG) erster Ordnung** ist eine Gleichung in der die Unbekannte eine Funktion in einer Variable ist und die die (erste) Ableitung dieser Funktion enthält.

$$y' = a y$$

$$y' + a y = b$$

$$y' + a y = b y^2$$

sind Differentialgleichungen erster Ordnung, die exponentielles, beschränktes, bzw. logistisches Wachstum beschreiben.

Allgemein

$$y' = F(t, y)$$

Lösung des Domar Modells

Durch Umformung der Differentialgleichung erhalten wir

$$\frac{1}{I(t)} I'(t) = \varrho s$$

Diese Gleichung muss für alle t gelten:

$$\ln(I) = \int \frac{1}{I} dI = \int \frac{1}{I(t)} I'(t) dt = \int \varrho s dt = \varrho s t + c$$

$$\text{Substitution: } I = I(t) \quad \Rightarrow \quad dI = I'(t) dt$$

Wir erhalten daher

$$I(t) = e^{\varrho s t} \cdot e^c = C e^{\varrho s t} \quad (C > 0)$$

Allgemeine Lösung

Alle Lösungen der DG $I' = \varrho s I$ lassen sich darstellen als

$$I(t) = C e^{\varrho s t} \quad (C > 0)$$

Diese Darstellung heißt die **allgemeine Lösung** der DG.

Wir erhalten *unendlich viele* verschiedene Lösungen!

Wir können uns von der Gültigkeit der Lösung durch Probe überzeugen:

$$\frac{dI}{dt} = \varrho s \cdot C e^{\varrho s t} = \varrho s \cdot I(t)$$

Anfangswertproblem

In unserem Modell ist die Investitionsrate zum Zeitpunkt $t = 0$ („jetzt“) bekannt. Wir erhalten somit zwei Gleichungen:

$$\begin{cases} I'(t) = qs \cdot I \\ I(0) = I_0 \end{cases}$$

Wir müssen daher eine Funktion $I(t)$ finden, die sowohl die DG als auch den Anfangswert erfüllt, i.e., wir müssen das sogenannte **Anfangswertproblem** lösen.

Wir erhalten die **spezielle Lösung** des Anfangswertproblems durch Einsetzen in die allgemeine Lösung.

Lösung des Domar Modells

Die speziell Lösung des Anfangswertproblems

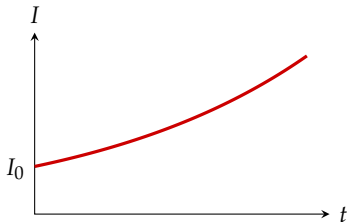
$$\begin{cases} I'(t) = \rho s \cdot I \\ I(0) = I_0 \end{cases}$$

erhalten wir durch Einsetzen in die allgemeine Lösung:

$$I_0 = I(0) = C e^{\rho s 0} = C$$

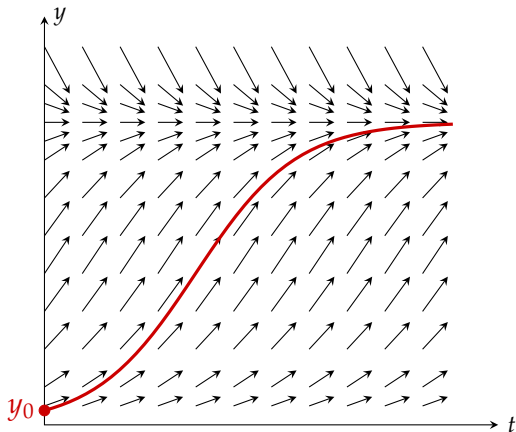
und daher

$$I(t) = I_0 e^{\rho s t}$$



Eine graphische Interpretation

Die Gleichung $y' = F(t, y)$ ordnet jedem Punkt (t, y) den Anstieg der Tangente zu. Wir erhalten ein sogenanntes **Vektorfeld**.



Trennung der Variablen

Differentialgleichungen der Form

$$y' = f(t) \cdot g(y)$$

lassen sich formal durch **Trennung der Variablen** lösen:

$$\frac{dy}{dt} = f(t) \cdot g(y) \iff \frac{1}{g(y)} dy = f(t) dt$$

Integration auf beiden Seiten ergibt

$$\int \frac{1}{g(y)} dy = \int f(t) dt + c$$

Wir erhalten dadurch eine Lösung der DG in *impliziter* Form.

Die DG des Domar Modells haben wir mit dieser Methode gelöst.

Beispiel – Trennung der Variablen

Wir suchen die Lösung der DG

$$y' + t y^2 = 0$$

Trennung der Variablen:

$$\frac{dy}{dt} = -t y^2 \quad \Rightarrow \quad -\frac{dy}{y^2} = t dt$$

Integrieren ergibt

$$-\int \frac{dy}{y^2} = \int t dt + c \quad \Rightarrow \quad \frac{1}{y} = \frac{1}{2}t^2 + c$$

und wir erhalten die allgemeine Lösung

$$y(t) = \frac{2}{t^2 + 2c}$$

Beispiel – Anfangswertproblem

Wir suchen die Lösung des Anfangswertproblems

$$y' + t y^2 = 0, \quad y(0) = 1$$

Spezielle Lösung durch Einsetzen:

$$1 = y(0) = \frac{2}{0^2 + 2c} \Rightarrow c = 1$$

und daher

$$y(t) = \frac{2}{t^2 + 2}$$

Lineare DG erster Ordnung

Ein **lineare Differential Gleichung erster Ordnung** hat die Gestalt

$$y'(t) + a(t)y(t) = s(t)$$

- ▶ **homogene** DG, falls $s = 0$.
- ▶ **inhomogene** DG, falls $s \neq 0$.

Homogene lineare DG lassen sich durch Trennung der Variablen lösen.

Beispiel – Homogene lineare DG

Wir suchen die allgemeine Lösung der homogenen linearen DG

$$y' + 3t^2y = 0$$

Trennung der Variablen

$$\frac{dy}{dt} = -3t^2y \quad \Rightarrow \quad \frac{1}{y}dy = -3t^2 dt \quad \Rightarrow \quad \ln y = -t^3 + c$$

Die allgemeine Lösung lautet daher

$$y(t) = C e^{-t^3}$$

Inhomogene lineare DG erster Ordnung

Wir wollen hier nur den Fall betrachten, in dem die Koeffizienten a und s der DG *konstant* und *ungleich 0* sind.

$$y'(t) + a y(t) = s$$

Wir erhalten dann die allgemeine Lösung als

$$y(t) = C e^{-at} + \frac{s}{a}$$

Für das Anfangswertproblem

$$y'(t) + a y(t) = s, \quad y(0) = y_0$$

erhalten wir die spezielle Lösung

$$y(t) = (y_0 - \bar{y}) e^{-at} + \bar{y} \quad \text{mit } \bar{y} = \frac{s}{a}$$

Beispiel – Inhomogene lineare DG

Wir suchen die Lösung des Anfangswertproblems

$$y' - 3y = 6, \quad y(0) = 1$$

Wir erhalten

$$\bar{y} = \frac{s}{a} = \frac{6}{3} = 2$$

$$y(t) = (y_0 - \bar{y}) e^{-at} + \bar{y} = (1 - 2) e^{3t} + 2 = e^{3t} + 2$$

Die spezielle Lösung lautet daher

$$y(t) = e^{3t} + 2$$

Modell – Marktdynamik

Nachfrage- und Angebotsfunktion seien linear:

$$q_d(t) = \alpha - \beta p(t) \quad (\alpha, \beta > 0)$$

$$q_s(t) = -\gamma + \delta p(t) \quad (\gamma, \delta > 0)$$

Die Preisanpassung sei direkt proportional zur Differenz ($q_d - q_s$):

$$\frac{dp}{dt} = j (q_d(t) - q_s(t)) \quad (j > 0)$$

Wie entwickelt sich der Preis $p(t)$ in Laufe der Zeit?

$$\begin{aligned} \frac{dp}{dt} &= j (q_d - q_s) = j (\alpha - \beta p - (-\gamma + \delta p)) \\ &= j (\alpha + \gamma) - j (\beta + \delta) p \end{aligned}$$

und wir erhalten die inhomogene lineare DG erster Ordnung

$$p'(t) + j (\beta + \delta) p(t) = j (\alpha + \gamma)$$

Modell – Marktdynamik

Die Lösung des Anfangswertproblems

$$p'(t) + j(\beta + \delta)p(t) = j(\alpha + \gamma), \quad p(0) = p_0$$

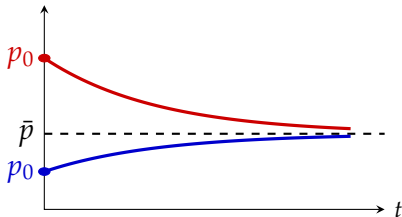
lautet

$$p(t) = (p_0 - \bar{p})e^{-j(\beta+\delta)t} + \bar{p}$$

mit

$$\bar{p} = \frac{s}{a} = \frac{j(\alpha + \gamma)}{j(\beta + \delta)} = \frac{\alpha + \gamma}{\beta + \delta}$$

\bar{p} ist gerade der Preis
im Marktgleichgewicht.



Logistische Differentialgleichung

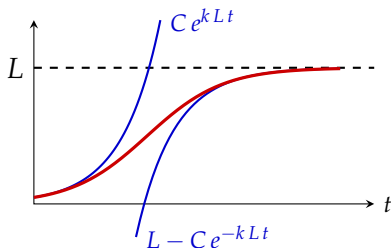
Eine **logistische Differentialgleichung** hat die Form

$$y'(t) - k y(t) (L - y(t)) = 0$$

wobei $k > 0$ und $0 \leq y(t) \leq L$.

► $y \approx 0$: $y'(t) - k L y(t) \approx 0 \Rightarrow y(t) \approx C e^{k L t}$

► $y \approx L$: $y'(t) + k L y(t) \approx k L^2 \Rightarrow y(t) \approx L - C e^{-k L t}$

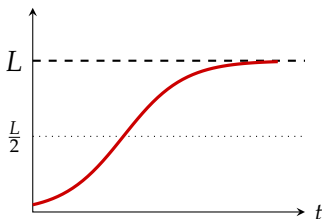


Logistische Differentialgleichung

Die exakte Lösung dieser DG kann durch Trennung der Variablen gefunden werden. Wir erhalten:

$$y(t) = \frac{L}{1 + C e^{-Lkt}}$$

Alle Lösungen haben einen Wendepunkt in $y = \frac{L}{2}$.



Beispiel

In einer Stadt mit 8100 Einwohnern ist eine Grippeepidemie ausgebrochen. Als die Grippewelle erkannt wird, sind bereits 100 Personen infiziert. 20 Tage später sind es bereits 1000. Es wird erwartet, dass alle Einwohner infiziert werden. Beschreiben Sie den Verlauf der Grippeepidemie.

Wir verwenden eine logistische DG mit $L = 8100$.

$q(t)$ sei die Anzahl der Infizierten,
wobei $q(0) = 100$ und $q(20) = 1000$.

Die allgemeine Lösung der DG lautet

$$q(t) = \frac{8100}{1 + C e^{-8100kt}}$$

Wir müssen k und C bestimmen.

Beispiel

$$q(0) = 100 \quad \Rightarrow \quad \frac{8100}{1 + C} = 100 \quad \Rightarrow \quad C = 80$$

$$q(20) = 1000 \quad \Rightarrow \quad \frac{8100}{1 + 80 e^{-8100 \cdot 20 k}} = 1000 \quad \Rightarrow \quad k = 0,00001495$$

Die Ausbreitung der Epidemie kann beschrieben werden durch die Funktion

$$q(t) = \frac{8100}{1 + 80 e^{-0,121 t}}$$

Differentialgleichungen 2. Ordnung

Eine **gewöhnliche Differentialgleichung zweiter Ordnung** ist eine Gleichung in der die Unbekannte eine Funktion in einer Variable ist und die die erste und zweite Ableitung dieser Funktion enthält:

$$y'' = F(t, y, y')$$

Wir beschränken uns hier auf **lineare Differentialgleichungen zweiter Ordnung mit konstanten Koeffizienten**:

$$y''(t) + a_1 y'(t) + a_2 y(t) = s$$

Homogene lineare DG 2. Ordnung

Wir erhalten allgemeine Lösung der homogenen linearen DG

$$y''(t) + a_1 y'(t) + a_2 y(t) = 0$$

mit dem Ansatz

$$y(t) = C e^{\lambda t}$$

wobei die Konstante λ die sogenannte **charakteristische Gleichung** erfüllen muss:

$$\lambda^2 + a_1 \lambda + a_2 = 0$$

Diese Bedingung folgt unmittelbar aus

$$\begin{aligned} y''(t) + a_1 y'(t) + a_2 y(t) &= \lambda^2 C e^{\lambda t} + a_1 \lambda C e^{\lambda t} + a_2 C e^{\lambda t} \\ &= C e^{\lambda t} (\lambda^2 + a_1 \lambda + a_2) = 0 \end{aligned}$$

Charakteristische Gleichung

Die charakteristische Gleichung

$$\lambda^2 + a_1\lambda + a_2 = 0$$

hat die Lösungen

$$\lambda_{1,2} = -\frac{a_1}{2} \pm \sqrt{\frac{a_1^2}{4} - a_2}$$

Es gibt drei Fälle:

1. $\frac{a_1^2}{4} - a_2 > 0$: zwei reelle Lösungen
2. $\frac{a_1^2}{4} - a_2 = 0$: eine reelle Lösung
3. $\frac{a_1^2}{4} - a_2 < 0$: zwei komplexe (nicht reelle) Lösungen

Fall: $\frac{a_1^2}{4} - a_2 > 0$

Die allgemeine Lösung der homogenen DG lautet

$$y(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}, \quad \text{mit } \lambda_{1,2} = -\frac{a_1}{2} \pm \sqrt{\frac{a_1^2}{4} - a_2}$$

Beispiel

Wir suchen die allgemeine Lösung von

$$y'' - y' - 2y = 0$$

Die charakteristische Gleichung

$$\lambda^2 - \lambda - 2 = 0$$

hat die beiden reellen Lösungen

$$\lambda_1 = -1 \quad \text{und} \quad \lambda_2 = 2$$

Die allgemeine Lösung der homogenen DG lautet daher

$$y(t) = C_1 e^{-t} + C_2 e^{2t}$$

Fall: $\frac{a_1^2}{4} - a_2 = 0$

Die allgemeine Lösung der homogenen DG lautet

$$y(t) = (C_1 + C_2 t) e^{\lambda t}, \quad \text{mit } \lambda = -\frac{a_1}{2}$$

Von der Gültigkeit der Lösung $t e^{\lambda t}$ kann man sich durch Nachrechnen überzeugen.

Beispiel

Wir suchen die allgemeine Lösung von

$$y'' + 4y' + 4y = 0$$

Die charakteristische Gleichung

$$\lambda^2 + 4\lambda + 4 = 0$$

hat die einzige (reelle) Lösung

$$\lambda = -2$$

Die allgemeine Lösung der homogenen DG lautet daher

$$y(t) = (C_1 + C_2 t) e^{-2t}$$

Fall: $\frac{a_1^2}{4} - a_2 < 0$

$\sqrt{\frac{a_1^2}{4} - a_2}$ ist in diesem Fall nicht reell (sondern eine imaginäre Zahl).

Aus den Rechenregeln für die komplexen Zahlen lässt sich aber eine rein reelle Lösung herleiten:

$$y(t) = e^{at} [C_1 \cos(bt) + C_2 \sin(bt)]$$

$$\text{mit } a = -\frac{a_1}{2} \quad \text{und} \quad b = \sqrt{\left| \frac{a_1^2}{4} - a_2 \right|}$$

Es sei darauf hingewiesen, dass a gerade der Realteil der Lösung der charakteristischen Gleichung ist, und b der Imaginärteil. Wir können aber hier nicht auf das Rechnen mit komplexen Zahlen eingehen.

Beispiel

Wir suchen die allgemeine Lösung von

$$y'' + y' + y = 0$$

Die charakteristische Gleichung

$$\lambda^2 + \lambda + 1 = 0$$

hat keine reellen Lösungen, da $\frac{a_1^2}{4} - a_2 = \frac{1}{4} - 1 = -\frac{3}{4} < 0$ ist.

$$a = -\frac{a_1}{2} = -\frac{1}{2} \quad \text{und} \quad b = \sqrt{\left|\frac{a_1^2}{4} - a_2\right|} = \sqrt{\frac{3}{4}} = \frac{\sqrt{3}}{2}$$

Die allgemeine Lösung der homogenen DG lautet daher

$$y(t) = e^{-\frac{1}{2}t} \left[C_1 \cos\left(\frac{\sqrt{3}}{2}t\right) + C_2 \sin\left(\frac{\sqrt{3}}{2}t\right) \right]$$

Inhomogene lineare DG 2. Ordnung

Die allgemeine Lösung der inhomogenen linearen DG

$$y''(t) + a_1 y'(t) + a_2 y(t) = s$$

hat die Form (falls $a_2 \neq 0$)

$$y(t) = y_h(t) + \frac{s}{a_2}$$

wobei $y_h(t)$ die allgemeine Lösung der entsprechenden homogenen DG ist:

$$y_h''(t) + a_1 y_h'(t) + a_2 y_h(t) = 0$$

Beispiel

Wir suchen die allgemeine Lösung der inhomogenen DG

$$y''(t) + y'(t) - 2y(t) = -10$$

Die charakteristische Gleichung der homogenen DG

$$\lambda^2 + \lambda - 2 = 0$$

hat die beiden reellen Lösungen

$$\lambda_1 = 1 \quad \text{und} \quad \lambda_2 = -2$$

Die allgemeine Lösung der inhomogenen DG lautet daher

$$y(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} + \frac{s}{a_2} = C_1 e^t + C_2 e^{-2t} + \frac{-10}{-2}$$

Anfangswertproblem

Alle allgemeinen Lösungen enthalten zwei unabhängige Koeffizienten C_1 und C_2 .

Für das Anfangswertproblem müssen wir daher zwei Werte vorgeben:

$$\begin{cases} y''(t) + a_1 y'(t) + a_2 y(t) = s \\ y(t_0) = y_0 \\ y'(t_0) = y'_0 \end{cases}$$

Beispiel

Wir suchen die spezielle Lösung des Anfangswertproblems

$$y''(t) + y'(t) - 2y(t) = -10, \quad y(0) = 12, \quad y'(0) = -2$$

Die allgemeine Lösung der DG lautet

$$y(t) = C_1 e^t + C_2 e^{-2t} + 5$$
$$y'(t) = C_1 e^t - 2C_2 e^{-2t}$$

Durch Einsetzen der Anfangswerte erhalten wir die Gleichungen

$$12 = y(0) = C_1 + C_2 e^{-2 \cdot 0}$$
$$-2 = y'(0) = C_1 - 2C_2$$

mit der Lösung $C_1 = 4$ und $C_2 = 3$. Die spezielle Lösung lautet daher

$$y(t) = 4e^t + 3e^{-2t} + 5$$

Fixpunkt einer Differentialgleichung

Die inhomogene lineare DG

$$y''(t) + a_1 y'(t) + a_2 y(t) = s$$

besitzt die spezielle konstante Lösung

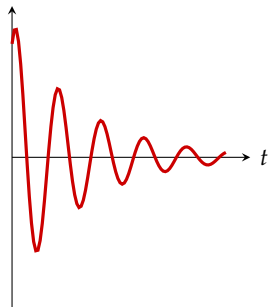
$$y(t) = \bar{y} = \frac{s}{a_2} \quad (= \text{konstant})$$

Der Punkt \bar{y} heißt **Fixpunkt** oder **stationärer Zustand** der DG.

Stabile und instabile Fixpunkte

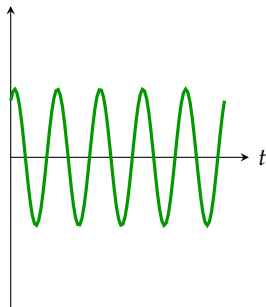
Der Wert von a bestimmt das qualitative Verhalten der Lösung

$$y(t) = e^{at} [C_1 \cos(bt) + C_2 \sin(bt)] + \bar{y}$$

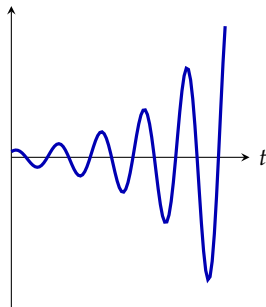


$$a < 0$$

stabiler Fixpunkt



$$a = 0$$



$$a > 0$$

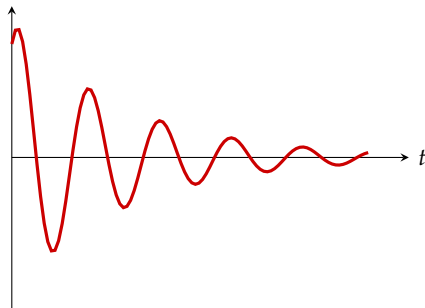
instabiler Fixpunkt

Asymptotisch stabiler Fixpunkt

Falls $a < 0$, dann konvergiert jede Lösung

$$y(t) = e^{at} [C_1 \cos(bt) + C_2 \sin(bt)] + \bar{y}$$

gegen \bar{y} . Der Fixpunkt \bar{y} heißt **asymptotisch stabil**.



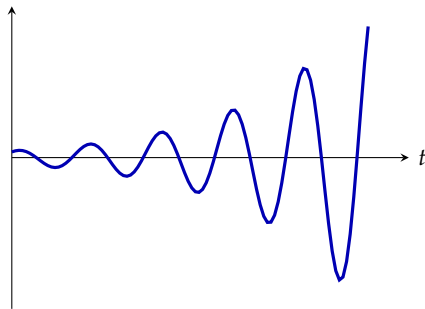
Asymptotisch stabiler Fixpunkt

Falls $a > 0$, dann wird jede Lösung

$$y(t) = e^{at} [C_1 \cos(bt) + C_2 \sin(bt)] + \bar{y}$$

mit Anfangswert $y(0) = y_0 \neq \bar{y}$ divergieren.

Der Fixpunkt \bar{y} heißt **instabil**.



Beispiel

Die allgemeine Lösung von

$$y'' + y' + y = 2$$

lautet (vgl. oben)

$$y(t) = 2 + e^{-\frac{1}{2}t} \left[C_1 \cos\left(\frac{\sqrt{3}}{2}t\right) + C_2 \sin\left(\frac{\sqrt{3}}{2}t\right) \right]$$

Der Fixpunkt $\bar{y} = 2$ ist asymptotisch stabil, da $a = -\frac{1}{2} < 0$.

Zusammenfassung

- ▶ Differentialgleichung 1. Ordnung
- ▶ Vektorfeld
- ▶ Trennung der Variablen
- ▶ Homogene und inhomogene lineare DG 1. Ordnung
- ▶ Logistische DG
- ▶ Homogene und inhomogene lineare DG 2. Ordnung
- ▶ Stabile und instabile Fixpunkte

Kapitel 15

Kontrolltheorie

Wirtschaftswachstum

Aufgabe: Maximiere Konsum im Zeitraum $[0, T]$:

$$\max_{0 \leq s(t) \leq 1} \int_0^T (1 - s(t)) f(k(t)) dt$$

$f(k)$... Produktionsfunktion

$k(t)$... Kapitalstock zum Zeitpunkt t

$s(t)$... Investitionsrate zum Zeitpunkt t , $s \in [0,1]$

Wir können nur $s(t)$ zu jedem Zeitpunkt frei wählen.

s heißt **Kontrollvariable**.

$k(t)$ folgt der Differentialgleichung

$$k'(t) = s(t) f(k(t)), \quad k(0) = k_0, \quad k(T) \geq k_T$$

Ölförderung

$y(t)$... Ölmenge in Ölfeld zum Zeitpunkt t

$u(t)$... Fördermenge zum Zeitpunkt t : $y'(t) = -u(t)$

$p(t)$... Ölpreis zum Zeitpunkt t

$C(t, y, u)$... Förderkosten

r ... Zinssatz (konstant)

Aufgabe I: Maximiere Gewinn im fixierten Zeitraum $[0, T]$:

$$\max_{u(t) \geq 0} \int_0^T [p(t)u(t) - C(t, y(t), u(t))] e^{-rt} dt$$

Wir können nur $u(t)$ zu jedem Zeitpunkt frei wählen, wobei $u(t) \geq 0$.

$y(t)$ folgt der Differentialgleichung:

$$y'(t) = -u(t), \quad y(0) = K, \quad y(T) \geq 0$$

Ölförderung

Aufgabe I:

Finde *Ölförderprogramm* $u(t)$, dass den Gewinn in einem fixierten Zeitraum $[0, T]$ maximiert.

Aufgabe II:

Finde *Ölförderprogramm* $u(t)$ und *Förderzeit* T , dass den Gewinn im Zeitraum $[0, T]$ maximiert.

Das Standardproblem (T fest)

1. Finde Maximum von

$$\max_u \int_0^T f(t, y, u) dt, \quad u \in \mathcal{U} \subseteq \mathbb{R}$$

u heißt **Kontrollvariable**, \mathcal{U} ist der **Kontrollbereich**.

2. *Kontrollierte Differentialgleichung* (Anfangswertproblem)

$$y' = g(t, y, u), \quad y(0) = y_0$$

3. Endwert

(a) $y(T) = y_1$

(b) $y(T) \geq y_1$ [oder: $y(T) \leq y_1$]

(c) $y(T)$ frei

(y, u) heißt **zulässiges Paar** falls (2) und (3) erfüllt sind.

Hamiltonfunktion

Analog zur Lagrangefunktion definieren wir die Funktion

$$\mathcal{H}(t, y, u, \lambda) = \lambda_0 f(t, y, u) + \lambda(t)g(t, y, u)$$

Diese Funktion wird als **Hamilton-Funktion** bezeichnet.

Die Funktion $\lambda(t)$ heißt die **adjungierte Variable**.

Die Zahl $\lambda_0 \in \{0,1\}$ kann bis auf wenige Ausnahmen gleich 1 gesetzt werden.

Wir werden daher im folgenden stets $\lambda_0 = 1$ voraussetzen:

$$\mathcal{H}(t, y, u, \lambda) = f(t, y, u) + \lambda(t)g(t, y, u)$$

Maximumsprinzip

Sei (y^*, u^*) ein optimales Paar für das Standardproblem.

Dann existiert eine stetige Funktion $\lambda(t)$, sodass für alle $t \in [0, T]$ gilt:

(i) u^* maximiert \mathcal{H} bezüglich u , i.e.,

$$\mathcal{H}(t, y^*, u^*, \lambda) \geq \mathcal{H}(t, y^*, u, \lambda) \quad \text{für alle } u \in \mathcal{U}$$

(ii) λ erfüllt die Differentialgleichung

$$\lambda' = -\frac{\partial}{\partial y} \mathcal{H}(t, y^*, u^*, \lambda)$$

(iii) **Transversalitätsbedingung**

(a) $y(T) = y_1$: $\lambda(T)$ frei

(b) $y(T) \geq y_1$: $\lambda(T) \geq 0$ [mit $\lambda(T) = 0$ falls $y^*(T) > y_1$]

(c) $y(T)$ frei: $\lambda(T) = 0$

Eine notwendige Bedingung

Das Maximumsprinzip beschreibt eine *notwendige* Bedingung für ein **optimales Paar** des Standardproblems, i.e., einem zulässigen Paar, dass dieses dynamische Optimierungsproblem löst.

D.h., für jedes optimale Paar lässt sich so eine Funktion $\lambda(t)$ finden.

Andererseits, falls wir so eine Funktion für ein zulässiges Paar (y^*, u^*) finden können, dann muss (y^*, u^*) nicht automatisch optimal sein. Es ist aber ein möglicher Kandidat für ein optimales Paar.

(Vgl. die Rolle der stationären Punkte in statischen Optimierungsproblemen.)

Eine hinreichende Bedingung

Sei (y^*, u^*) ein zulässiges Paar des Standardproblem und $\lambda(t)$ eine Funktion, die das Maximumsprinzip erfüllt.

Falls \mathcal{U} konvex und $\mathcal{H}(t, y, u, \lambda)$ konkav in (y, u) für alle $t \in [0, T]$ ist, dann ist (y^*, u^*) ein optimales Paar.

Vorgangsweise

1. Für jedes Tripel (t, y, λ) suche ein (globales) Maximum $\hat{u}(t, y, \lambda)$ von $\mathcal{H}(t, y, u, \lambda)$ bzgl. u .
2. Löse die Differentialgleichungen

$$y' = g(t, y, \hat{u}(t, y, \lambda), \lambda)$$

$$\lambda' = -H_y(t, y, \hat{u}(t, y, \lambda), \lambda)$$

3. Finde spezielle Lösungen $y^*(t)$ und $\lambda^*(t)$, die die Anfangsbedingung $y(0) = y_0$ bzw. die Transversalitätsbedingung erfüllen.
4. Wir erhalten einen Kandidaten für ein optimales Paar durch $y^*(t)$ und $u^*(t) = \hat{u}(t, y^*, \lambda^*)$.
5. Falls \mathcal{U} konvex und $\mathcal{H}(t, y, u, \lambda^*)$ konkav in (y, u) ist, dann ist (y^*, u^*) ein optimales Paar.

Beispiel 1

Wir suchen die optimale Kontrollfunktion u^* für

$$\max \int_0^1 y(t) dt, \quad u \in [0,1]$$

$$y' = y + u, \quad y(0) = 0, \quad y(1) \text{ frei}$$

Heuristisch:

Die Zielfunktion und damit u sollten möglichst groß sein.

Daher ist $u^*(t) = 1$ für alle t .

Hamiltonfunktion:

$$\mathcal{H}(t, y, u, \lambda) = f(t, y, u) + \lambda g(t, y, u) = y + \lambda(y + u)$$

Beispiel 1

$$\mathcal{H}(t, y, u, \lambda) = y + \lambda(y + u)$$

Maximum \hat{u} von \mathcal{H} bzgl. u :

$$\hat{u} = \begin{cases} 1 & \text{falls } \lambda \geq 0, \\ 0 & \text{falls } \lambda < 0 \end{cases}$$

Lösung der (inhomogen lineare) DG

$$\lambda' = -H_y = -(1 + \lambda), \quad \lambda(1) = 0$$

$$\Rightarrow \lambda^*(t) = e^{1-t} - 1$$

Da $\lambda^*(t) = e^{1-t} - 1 \geq 0$ für alle $t \geq 0$ gilt: $\hat{u}(t) = 1$.

Beispiel 1

Löse (inhomogene lineare) DG

$$y' = y + \hat{u} = y + 1, \quad y(0) = 0$$

$$\Rightarrow y^*(t) = e^t - 1$$

Wir erhalten daher

$$u^*(t) = \hat{u}(t) = 1$$

Die Hamiltonfunktion $\mathcal{H}(t, y, u, \lambda) = y + \lambda(y + u)$ ist linear und damit konkav in (y, u) .

$u^*(t) = 1$ ist die gesuchte optimale Kontrollfunktion.

Beispiel 2

Wir suchen die optimale Kontrollfunktion u^* für

$$\min \int_0^T [y^2(t) + cu^2(t)] dt, \quad u \in \mathbb{R}, \quad c > 0$$

$$y' = u, \quad y(0) = y_0, \quad y(T) \text{ frei}$$

Wir lösen das Maximierungsproblem

$$\max \int_0^T - [y^2(t) + cu^2(t)] dt$$

Hamiltonfunktion:

$$\mathcal{H}(t, y, u, \lambda) = f(t, y, u) + \lambda g(t, y, u) = -y^2 - cu^2 + \lambda u$$

Beispiel 2

Maximum \hat{u} von \mathcal{H} bzgl. u :

$$0 = H_u = -2c\hat{u} + \lambda \quad \Rightarrow \quad \hat{u} = \frac{\lambda}{2c}$$

Lösungen der Differentialgleichungen

$$y' = \hat{u} = \frac{\lambda}{2c}$$

$$\lambda' = -H_y = 2y$$

Durch Differenzieren der zweiten DG erhalten wir

$$\lambda'' = 2y' = \frac{\lambda}{c} \quad \Rightarrow \quad \lambda'' - \frac{1}{c}\lambda = 0$$

Die Lösung dieser homogenen linearen DG 2. Ordnung lautet

$$\lambda^*(t) = C_1 e^{rt} + C_2 e^{-rt}, \quad \text{mit } r = \frac{1}{\sqrt{c}}$$

($\pm \frac{1}{\sqrt{c}}$ sind die beiden Nullstellen des charakteristischen Polynoms.)

Beispiel 2

Anfangswert und Transversalitätsbedingung liefern

$$\lambda^{*'}(0) = 2y(0) = 2y_0$$

$$\lambda^*(T) = 0$$

und somit

$$r(C_1 - C_2) = 2y_0$$

$$C_1 e^{rT} + C_2 e^{-rT} = 0$$

mit der Lösung

$$C_1 = \frac{2y_0 e^{-rT}}{r(e^{rT} + e^{-rT})}, \quad C_2 = -\frac{2y_0 e^{rT}}{r(e^{rT} + e^{-rT})}$$

Beispiel 2

Wir erhalten somit

$$\lambda^*(t) = \frac{2y_0}{r(e^{rT} + e^{-rT})} \left(e^{-r(T-t)} - e^{r(T-t)} \right)$$

$$y^*(t) = \frac{1}{2} \lambda^*(t) = y_0 \frac{e^{-r(T-t)} - e^{r(T-t)}}{r(e^{rT} + e^{-rT})}$$

$$u^*(t) = \hat{u}(t, y^*, \lambda^*) = \frac{1}{2c} \lambda^*(t) = \frac{y_0}{c} \frac{e^{-r(T-t)} - e^{r(T-t)}}{r(e^{rT} + e^{-rT})}$$

Mittel Hesse-Matrix lässt sich leicht prüfen, dass die Hamiltonfunktion $\mathcal{H}(t, y, u, \lambda) = -y^2 - cu^2 + \lambda u$ konkav in y und u ist.

$u^*(t) = \frac{y_0}{c} \frac{e^{-r(T-t)} - e^{r(T-t)}}{r(e^{rT} + e^{-rT})}$ ist die gesuchte optimale Kontrollfunktion.

Das Standardproblem (T variabel)

Wenn der Zeitraum $[0, T]$ nicht a priori festgelegt wird, so muss außer der optimalen Kontrollvariable u^* auch das optimale Zeitintervall $[0, T^*]$ bestimmt werden.

Die Vorgangsweise ist vollkommen analog zum bereits behandelten Fall. Allerdings müssen wir noch folgende Bedingung *zusätzlich* zu (i)–(iii) zum Maximumsprinzip dazufügen:

(iv)

$$\mathcal{H}(T^*, y^*(T^*), u^*(T^*), \lambda(T^*)) = 0$$

Zusammenfassung

- ▶ Standardproblem
- ▶ Hamiltonfunktion
- ▶ Maximumsprinzip
- ▶ Hinreichende Bedingung