Mathematik für Wirtschaftswissenschaftler II

Prof. Dr. Johannes Schropp,
Universität Konstanz

Typesetting S. Ringmann

4. Juli 2024

Inhaltsverzeichnis

Kapitel 5

Lineare Algebra

5.1 Vektoren und Matrizen

5.1.1 Beispiel Erdbebenhilfe:

Nach einem Beben in Zentralamerika soll ein Flugzeug mit Hilfsmitteln entsandt werden, wobei die Kapazitäten in Bezug auf den Laderaum, das Abfluggewicht und die zur Verfügung stehenden Geldmittel voll ausgeschöpft werden sollen.

	Volumen (l)	Gewicht (kg)	Kosten (€)
Blutkonserven	200	150	1000
Medikamente	300	100	300
Nahrungsmittel	80	60	400
Frischwasser	60	70	400
Kapazitäten	60000	40000	150000

Frage: Wieviele Container jeder Sorte sind zu verschicken, wenn über dies Frischwasser am dringendsten benötigt wird und deshalb doppelt so viele Wasserbehälter wie Container mit Blut und Medikamenten verschickt werden sollen?

Lösungsweg:

$$\begin{aligned} x_1 &= \text{Anzahl Behälter Blutkonserven} \\ x_2 &= \text{Anzahl Behälter Medikamente} \\ x_3 &= \text{Anzahl Behälter Nahrungsmittel} \\ x_4 &= \text{Anzahl Behälter Frischwasser} \end{aligned}$$

Nun gilt:

$$\begin{array}{lrl} \text{Restriktion Volumen:} & 200x_1 + 300x_2 + 80x_3 + 60x_4 = & 60000 \\ \text{Restriktion Gewicht:} & 150x_1 + 100x_2 + 60x_3 + 70x_4 = & 40000 \\ \text{Nebenbedingung (Restriktion Wasser):} & 2x_1 + 2x_2 - x_4 = & 0 \\ \text{Restriktion Kosten:} & 1000x_1 + 300x_2 + 400x_3 + 400x_4 = & 150000 \end{array}$$

"lineares Gleichungssystem von 4 Gleichungen in 4 Variablen"

5.1.2 Matrixdarstellung linearer Gleichungssysteme

Seien $a_{ij},\ b_i,\ i=1,\dots,M,\ j=1,\dots,N$ gegeben. Dann heißt

$$\begin{array}{lcl} a_{11}\cdot x_1 + a_{12}\cdot x_2 + \ldots + a_{1N}\cdot x_N & = & b_1 \\ a_{21}\cdot x_1 + a_{22}\cdot x_2 + \ldots + a_{2N}\cdot x_N & = & b_2 \\ \vdots & & \vdots \\ a_{M1}\cdot x_1 + a_{M2}\cdot x_2 + \ldots + a_{MN}\cdot x_M & = & b_M \end{array} \qquad (*)$$

"lineares Gleichungssystem mit M-Gleichungen in N Unbekannten $x_1, x_2, \dots, x_N$"

Gesucht sind Zahlen $x_1, x_2, \dots, x_N$ derart, dass $(*)$ erfüllt ist.

Fragestellungen zu $(*)$:

- Existenz von Lösungen
- Eindeutigkeit von Lösungen
- Verfahren zur Bestimmung von Lösungen

Kompakte Schreibweise von $(*)$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \vdots & \\ a_{M1} & a_{M2} & \dots & a_{MN} \end{pmatrix} \in \mathbb{R}^{M,N}, \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{pmatrix} \in \mathbb{R}^M, \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \in \mathbb{R}^N$$

"A ist eine Matrix mit M Zeilen und N Spalten und Einträgen aus $\mathbb{R}$"

Das Problem $(*)$ ist gegeben durch $A \in \mathbb{R}^{M,N}$ und $b \in \mathbb{R}^M$

Gesucht ist $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \in \mathbb{R}^N$ mit $(*)$.

Eine Matrix mit einer Zeile / Spalte nennt man auch einen Zeilen- / Spaltenvektor.

Bezeichnungen: $\mathbb{R}^{M,N}$ Menge aller Matrizen mit M Zeilen und N Spalten
$\mathbb{R}^M$ Menge aller Spaltenvektoren im $\mathbb{R}^M$, offenbar gilt: $\mathbb{R}^M = \mathbb{R}^{M,1}$

Elementare Rechenoperation für Matrizen:

Matrizen gleicher Größe können addiert und subtrahiert werden.
Seien $A = (a_{ij}),\ B = (b_{ij}) \in \mathbb{R}^{M,N} \quad \Rightarrow \quad A + B = (a_{ij} + b_{ij}) \in \mathbb{R}^{M,N}$

Multiplikation mit Skalar $\lambda \in \mathbb{R}: \quad \lambda A = (\lambda a_{ij}) \in \mathbb{R}^{M,N}$

Beispiel:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathbb{R}^{2,3}, \quad B = \begin{pmatrix} 0 & -1 & 1 \\ 2 & 1 & -1 \end{pmatrix} \in \mathbb{R}^{2,3}, \quad A + B = \begin{pmatrix} 1 & 1 & 4 \\ 6 & 6 & 5 \end{pmatrix} \in \mathbb{R}^{2,3}$$

$$\lambda = 3, \quad 3 \cdot A = \begin{pmatrix} 3 & 6 & 9 \\ 12 & 15 & 18 \end{pmatrix}$$

Für die Matrizenaddition und die Multiplikation von Matrizen mit Skalaren gelten komponentenweise die Rechenregeln der reellen Zahlen.

Rechenregeln für Matrizen: Seien $A, B, C \in \mathbb{R}^{M,N}$, $\lambda \in \mathbb{R}$

$$\begin{aligned} \lambda(A+B) &= \lambda A + \lambda B && \text{(Distributivgesetz)} \\ A + B &= B + A && \text{(Kommutativgesetz)} \\ A + (B + C) &= (A+B) + C && \text{(Assoziativgesetz)} \end{aligned}$$

Diese Regeln folgen unmittelbar aus den entsprechenden Regeln für reelle Zahlen.

Für $\mathbb{R}^{M,1} = \mathbb{R}^M$ erhalten wir die Addition und Multiplikation mit Skalaren von Vektoren.

Man kann zueinander passende Matrizen auch multiplizieren.

Definition (Matrizenmultiplikation):

Seien $A = (a_{ij}) \in \mathbb{R}^{M,N}$ und $B = (b_{ij}) \in \mathbb{R}^{N,P}$, d.h. die Spaltenzahl von A stimmt mit der Zeilenzahl von B überein. Dann ist das Produkt von A und B die Matrix

$$C = A{\cdot}B = (c_{ij}) \in \mathbb{R}^{M,P} \quad \text{mit} \quad c_{ij} = a_{i1}{\cdot}b_{1j} + a_{i2}{\cdot}b_{2j} + a_{i3}{\cdot}b_{3j} + \ldots + a_{iN}{\cdot}b_{Nj} = \sum_{l=1}^{N} a_{il}{\cdot}b_{lj}, \quad i = 1, \ldots, M, \quad j = 1, \ldots, P$$

Bemerkung:

i) $A \cdot B$ ist nur definiert, falls die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmt!

ii) Ist $A \cdot B$ definiert, so erhält man $(A \cdot B)_{ij}$ als Skalarprodukt der i-ten Zeile von A mit der j-ten Spalte von B.

Beispiel: $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{3,2}$, $\quad B = \begin{pmatrix} 0 & 1 & 2 & 1 \\ -1 & 1 & 0 & -1 \end{pmatrix} \in \mathbb{R}^{2,4}$

Es gilt

$$A \cdot B \in \mathbb{R}^{3,4}, \qquad A \cdot B = \begin{pmatrix} -2 & 3 & 2 & -1 \\ -1 & 1 & 0 & -1 \\ 0 & -1 & -2 & -1 \end{pmatrix} \in \mathbb{R}^{3,4} \qquad \text{"Skalarprodukt Zeile mal Spalte"}$$

Satz:

a) Das Matrizenprodukt ist assoziativ und distributiv, d.h. $A(B \cdot C) = (A \cdot B) \cdot C$, $A(B + C) = A \cdot B + A \cdot C$, falls die Multiplikation möglich ist.

b) Das Matrizenprodukt ist im allgemeinen nicht kommutativ, d.h. $AB \neq BA$

Beispiel zu b):

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2,2}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2,2} \qquad \text{dann gilt:} \quad AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2,2} \quad \text{"Nullmatrix"}$$

Bemerkung:

$$I_M = \begin{pmatrix} 1 & 0 & \ldots & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \ldots & 0 & 1 & 0 \\ 0 & \ldots & \ldots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{M,M} \qquad \text{"Einheitsmatrix"}$$

Es gilt $I_M \cdot A = A \cdot I_M = A$ für jede Matrix $A \in \mathbb{R}^{M,M}$.

Gleichungssystem: $A = (a_{ij}) \in \mathbb{R}^{M,N}, \ x \in \mathbb{R}^N, \ b \in \mathbb{R}^M$

$$\begin{array}{rcl} a_{11} \cdot x_1 + a_{12} \cdot x_2 + \ldots + a_{1N} \cdot x_N & = & b_1 \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 + \ldots + a_{2N} \cdot x_N & = & b_2 \\ \vdots & & \vdots \\ a_{M1} \cdot x_1 + a_{M2} \cdot x_2 + \ldots + a_{MN} \cdot x_N & = & b_M \end{array}$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1N} \\ a_{21} & a_{22} & \ldots & a_{2N} \\ \vdots & \vdots & \vdots & \\ a_{M1} & a_{M2} & \ldots & a_{MN} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{pmatrix} \quad \Rightarrow \quad (*) \text{ ist äquivalent zu } Ax = b$$

Problem: Gegeben: $A \in \mathbb{R}^{M,N}, \ b \in \mathbb{R}^M$
Gesucht ist ein $x \in \mathbb{R}^N$ so, dass $A \cdot x = b$

5.1.3 Lineare Unabhängigkeit von Vektoren

Beispiel: $u = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \ z = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \ w = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \ u, z, w \in \mathbb{R}^3$

$$u + z = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} = w \quad \Leftrightarrow \quad u + z - w = 0$$

Man sagt: u, z, w sind linear abhängig, da sich w als Linearkombination von u und z darstellen lässt.

Definition (linear unabhängig):

Seien $v^1, v^2, \ldots, v^m \in \mathbb{R}^N$. Dann heißen $v^1, v^2, \ldots, v^m$ linear unabhängig, falls gilt:

$$\underbrace{\lambda_1 v^1 + \lambda_2 v^2 + \ldots + \lambda_m v^m}_{\text{Linearkombination von } v^1, \ldots, v^m} = 0, \quad \lambda_1, \ldots, \lambda_m \in \mathbb{R} \quad \Rightarrow \lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$$

Dies heißt, dass der Vektor $0 \in \mathbb{R}^N$ aus den Vektoren $v^1, v^2, \ldots, v^m$ nur durch $0 \cdot v^1 + 0 \cdot v^2 + \ldots + 0 \cdot v^m = 0$ erzeugt werden kann.

Beachte: u, z, w sind nicht linear unabhängig, d.h. linear abhängig, denn $1 \cdot u + 1 \cdot z - 1 \cdot w = 0$

Satz:

Vektoren $v^1, v^2, \ldots, v^m \in \mathbb{R}^N$ sind genau dann linear abhängig, falls es ein $i_0 \in \{1, \ldots, m\}$ und reelle Zahlen $\lambda_1, \ldots, \lambda_{i_0-1}, \lambda_{i_0+1}, \ldots, \lambda_m$ gibt mit

$$v^{i_0} = \lambda_1 v^1 + \ldots + \lambda_{i_0-1} v^{i_0-1} + \lambda_{i_0+1} v^{i_0+1} + \ldots + \lambda_m v^m$$

d.h. v^{i_0} lässt sich als Linearkombination der anderen Vektoren darstellen.

Beispiel: Vektor $e^i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow$ i-te Komponente

heißt i-ter Einheitsvektor bzw. i-ter kanonischer (normaler) Basisvektor.

Es gilt: Die Vektoren $e^1, e^2, \ldots, e^N \in \mathbb{R}^N$ sind linear unabhängig.

Setze an

$$0 = \lambda_1 e^1 + \lambda_2 e^2 + \ldots + \lambda_N e^N = \underbrace{\lambda_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{= \begin{pmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}} + \underbrace{\lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{= \begin{pmatrix} 0 \\ \lambda_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}} + \ldots + \underbrace{\lambda_N \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}}_{= \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_N \end{pmatrix}} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \end{pmatrix} \qquad \Rightarrow \lambda_i = 0, \; i = 1, \ldots, N.$$

Weiterhin gilt für jeden Vektor $w = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix}$ die Darstellung: $w = w_1 e^1 + w_2 e^2 + \ldots + w_N e^N = \sum\limits_{j=1}^{N} w_j e^j$

$\Rightarrow$ Jeder Vektor $w = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix} \in \mathbb{R}^N$ kann eindeutig durch die e^i mit Gewicht w_i, $i = 1, \ldots, N$ linear kombiniert werden.

Ein Satz von N linear unabhängigen Vektoren im $\mathbb{R}^N$ mit dieser Eigenschaft heißt Basis.
$e^1, \ldots, e^N$ heißt kanonische Basis von $\mathbb{R}^N$.

5.1.4 Rang einer Matrix

Sei $A = (a_{ij}) \in \mathbb{R}^{M,N}$ eine reelle Matrix mit M Zeilen und N Spalten: $A = \begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \dots & a_{MN} \end{pmatrix}$

Dann können wir die Zeilen (bzw. die Spalten) von A als Zeilenvektoren im $\mathbb{R}^N$ (Spaltenvektoren im $\mathbb{R}^M$) auffassen.

Die Maximalzahl linear unabhängiger Zeilen (Spalten) von A heißt Zeilenrang (Spaltenrang) von A.

Bemerkung: Es gilt: Zeilenrang = Spaltenrang für jede Matrix A und man spricht Rang von A und schreibt: $rg(A)$

Beachte: $rg(A) \in \mathbb{N} \qquad 0 \leq rg(A) \leq \min(N, M)$

Beispiel: $A = \begin{pmatrix} 1 & 4 & -2 \\ 2 & 5 & -1 \\ \underbrace{3}_{=a^1} & \underbrace{6}_{=a^2} & \underbrace{0}_{=a^3} \end{pmatrix} \in \mathbb{R}^{3,3}, \qquad rg(A) = ?$

Berechne: $2a^1 - a^2 = 2 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix} = a^3$

d.h. a^1, a^2, a^3 sind linear abhängig $\quad \Rightarrow \quad rg(A) \leq 2$

Rechne nach: a^1, a^2 sind linear unabhängig $\Rightarrow \quad rg(A) = 2$

$\Rightarrow$ Wir brauchen systematische Vorgehensweisen. Es fehlt aber bisher noch ein systematisches Vorgehen zur Bestimmung des Rangs von Matrizen.

Elementare Zeilenumformungen

Gegeben sei $A \in \mathbb{R}^{M,N}$: Folgende Operationen an A heißen elementare Zeilenumformungen.

i) Vertauschen der Zeile i mit der Zeile j.

ii) Ersetzen der Zeile i durch Zeile $i + \lambda \cdot$ Zeile j, $\lambda \in \mathbb{R}$, d.h. zur Zeile i wird das λ-fache der Zeile j hinzuaddiert.

iii) Multiplizieren der Zeile i mit einer Zahl $\lambda \in \mathbb{R}, \ \ \lambda \neq 0$

Satz:

Elementare Zeilenumformungen verändern den Rang einer Matrix nicht.

Mittels elementarer Zeilenumformungen kann man eine Matrix in eine neue Matrix mit besonders "schöner" Form, der sogenannten Zeilenstufenform überführen.

Einer solchen Matrix sieht man den Rang sofort an:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad a_{11}, a_{23}, a_{34} \neq 0, \quad \text{A ist in Zeilenstufenform}$$

Dann gilt: Der Rang von A ist gleich der Anzahl der von Null verschiedenen Zeilenvektoren.

$\Rightarrow \quad rg(A) = 3$

Berechnung der Zeilenstufenform mittels elementaren Zeilenumformungen

$$A = \begin{pmatrix} 0 & 2 & -1 \\ 3 & -4 & 5 \\ 6 & -6 & 12 \end{pmatrix} \underbrace{\sim}_{\text{Vertausche die 1. und die 3. Zeile}} \begin{pmatrix} 6 & -6 & 12 \\ 3 & -4 & 5 \\ 0 & 2 & -1 \end{pmatrix} \underbrace{\sim}_{\lambda=-1/2,\text{ addiere } \lambda\cdot \text{ 1. und 2. Zeile}}$$

$$\begin{pmatrix} 6 & -6 & 12 \\ 0 & -1 & -1 \\ 0 & 2 & -1 \end{pmatrix} \underbrace{\sim}_{\lambda=2,\text{ addiere } \lambda\cdot \text{ 2. und 3. Zeile}} \begin{pmatrix} 6 & -6 & 12 \\ 0 & -1 & -1 \\ 0 & 0 & -3 \end{pmatrix} \underbrace{\sim}_{\text{Zeilenstufenform erreicht}} \quad \Rightarrow \quad rg(A) = 3$$

Definition:

Eine quadratische Matrix $A \in \mathbb{R}^{N,N}$ heißt regulär, wenn rg(A)=N ist.

Es gibt dann die inverse Matrix $A^{-1} \in \mathbb{R}^{N,N}$.
Diese ist eindeutig charakterisiert durch: $A \cdot A^{-1} = A^{-1} \cdot A = I_N$

Satz: Sei $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}^{2,2}$.

Ist $D = a_{11} \cdot a_{22} - a_{21} \cdot a_{12} \neq 0$ so ist A regulär, d.h. invertierbar und es gilt:

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

Beweis: Man rechne nach, dass $A \cdot A^{-1} = I_2$ gilt.

5.2 Lösungsverfahren für lineare Gleichungssysteme

5.2.1 Erweiterte Koeffizientenmatrix

Vorgelegt sei

$$\begin{array}{lcl} a_{11}\cdot x_1 + a_{12}\cdot x_2 + \ldots + a_{1N}\cdot x_N & = & b_1 \\ a_{21}\cdot x_1 + a_{22}\cdot x_2 + \ldots + a_{2N}\cdot x_N & = & b_2 \\ \vdots & & \vdots \\ a_{M1}\cdot x_1 + a_{M2}\cdot x_2 + \ldots + a_{MN}\cdot x_N & = & b_M \end{array} \qquad (*)$$

„ M Gleichungen in N Unbekannten"

Kompakte Schreibweise:
$$\begin{array}{llll} A = (a_{ij}) & i = 1,\ldots,M & j = 1,\ldots,N & A \in \mathbb{R}^{M,N} \\ b = (b_i) & i = 1,\ldots,M & & b \in \mathbb{R}^M \\ x = (x_j) & j = 1,\ldots,N & & x \in \mathbb{R}^N \end{array}$$

Dann schreibt sich $(*)$ zu $A \cdot x = b$

Das Gleichungssystem $Ax = b$ heißt homogen, falls b $= 0$ (Nullvektor), sonst heißt es inhomogen.

Zu einem linearen Gleichungssystem $Ax = b$ heißt $(A|b) \in \mathbb{R}^{M,N+1}$

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1N} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2N} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{M1} & a_{M2} & \ldots & a_{MN} & b_M \end{array}\right) \in \mathbb{R}^{M,N+1}$$

die erweiterte Koeffizientenmatrix.

An der erweiterten Koeffizientenmatrix lässt sich das Lösungsverhalten des Systems $Ax = b$ entscheiden.

Satz:

Vorgelegt sei das lineare Gleichungssystem $Ax = b$ mit $A \in \mathbb{R}^{M,N}$, $x \in \mathbb{R}^N$, $b \in \mathbb{R}^M$

a) $Ax = b$ besitzt genau dann eine Lösung, wenn gilt: $rg(A) = rg(A|b)$

b) Die Lösung, falls eine existiert, von $Ax = b$ ist genau dann eindeutig, wenn $rg(A) = N$ =Anzahl der Unbekannten.
(Im Fall $M = N$ ist dies äquivalent zu A invertierbar)

Satz:

Elementare Zeilenumformungen an der erweiterten Koeffizientenmatrix $(A|b)$ verändern den Lösungsraum von $Ax = b$ nicht.

Bemerkung: Elementare Zeilenumformungen an $(A|b)$ ändern die Lösungsmenge eines linearen Gleichungssystems nicht.

5.2.2 Das Verfahren von Gauß

Ziel: Bringe zuerst die erweiterte Koeffizientenmatrix auf Zeilenstufenform und lese dann die Lösung ab.

Addiere (im ersten Schritt) das λ-fache der ersten Zeile zu der i-ten Zeile $i = 2, 3, \ldots, M$ und bestimme λ so, dass $\lambda a_{11} + a_{i1} = 0$.

Verfahren in den nächsten Schritten analog.

Beispiel: Erdbebenhilfe

$$(A|b) = (A^{(0)}|b^{(0)}) = \left(\begin{array}{cccc|c} 200 & 300 & 80 & 60 & 60000 \\ 150 & 100 & 60 & 70 & 40000 \\ 2 & 2 & 0 & -1 & 0 \\ 1000 & 300 & 400 & 200 & 150000 \end{array}\right) \begin{array}{l} \\ \lambda = -\frac{3}{4} \\ \lambda = -\frac{1}{100} \\ \lambda = -5 \end{array}$$

$(A^{(0)})_{11} = 200$ heißt Pivot-Element im ersten Schritt

$$\left(\begin{array}{cccc|c} 200 & 300 & 80 & 60 & 60000 \\ 0 & -125 & 0 & 25 & -5000 \\ 0 & -1 & -\frac{4}{5} & -\frac{8}{5} & -600 \\ 0 & -1200 & 0 & -100 & -150000 \end{array}\right) \begin{array}{l} \\ \\ \lambda = -\frac{1}{125} \\ \lambda = -\frac{1200}{125} \end{array} = (A^{(1)}|b^{(1)})$$

$(A^{(1)})_{22} = -125$ ist Pivot-Element im 2-ten Schritt.

$$\left(\begin{array}{cccc|c} 200 & 300 & 80 & 60 & 60000 \\ 0 & -125 & 0 & 25 & -5000 \\ 0 & 0 & -\frac{4}{5} & -\frac{9}{5} & -560 \\ 0 & 0 & 0 & -340 & -102000 \end{array}\right) = (A^{(2)}|b^{(2)}) = (\hat{A}|\hat{b})$$

"GaußVerfahren (Teil 1)" $(\hat{A}|\hat{b})$ hat Zeilenstufenform.

Nun gilt: $rg(A) = rg(\hat{A}) = 4$

$rg(A|b) = rg(\hat{A}|\hat{b}) = 4$

$\Rightarrow \quad rg(A) = rg(A|b) \quad \Rightarrow$ das lineare Gleichungssystem ist lösbar.

$4 =$ Anzahl der Unbekannten $\Rightarrow$ Lösung ist eindeutig.

Für ein allgemeines System $Ax = b$ liefert das Gauß - Verfahren eine Sequenz $(A|b) = (A^{(0)}|b^{(0)}) \sim (A^{(1)}|b^{(1)}) \sim (A^{(2)}|b^{(2)}) \sim \ldots \sim (A^{(M-1)}|b^{(M-1)}) = (\hat{A}|\hat{b})$ und $(\hat{A}|\hat{b})$ hat Zeilenstufenform. An $(\hat{A}|\hat{b})$ lässt sich die Lösbarkeit von $Ax = b$ ablesen.

Im Falle der Lösbarkeit lässt sich die Lösung von $Ax = b$ durch Rückwärtsauflösen $\hat{A}x = \hat{b}$ berechnen.

Beispiel 1 (Erdbebenhilfe):

$$\begin{array}{lll} -340x_4 = -102000, & \Rightarrow & x_4 = \frac{-102000}{-340} = 300 \\ -\frac{4}{5}x_3 - \frac{9}{5}x_4 = -\frac{4}{5}x_3 - \frac{9}{5}\cdot 300 = -560 & \Rightarrow & x_3 = 25 \\ -125x_2 + 0x_3 + 25x_4 = -125x_2 + 25\cdot 300 = -5000 & \Rightarrow & x_2 = 100 \\ 200x_1 + 300x_2 + 80x_3 + 60x_4 = 200x_1 + 300\cdot 100 + 80\cdot 25 + 60\cdot 300 = 60000 & \Rightarrow & x_1 = 50 \end{array}$$

Lösung : $x = (x_1, x_2, x_3, x_4) = (50, 100, 25, 300)$

Beispiel 2: Sei $A \in \mathbb{R}^{4,5},\ b \in \mathbb{R}^4$ wie folgt gegeben:

$$(A|b) = (A^{(0)}|b^{(0)}) = \left(\begin{array}{ccccc|c} 2 & 4 & 4 & -2 & 0 & 6 \\ 4 & 8 & 8 & -4 & 4 & 4 \\ 2 & 4 & 5 & 2 & 1 & 3 \\ 6 & 12 & 12 & -6 & 4 & 10 \end{array}\right) \begin{array}{l} \\ \lambda = -2 \\ \lambda = -1 \\ \lambda = -3 \end{array}$$

$$(A^{(1)}|b^{(1)}) = \left(\begin{array}{ccccc|c} 2 & 4 & 4 & -2 & 0 & 6 \\ 0 & 0 & 0 & 0 & 4 & -8 \\ 0 & 0 & 1 & 4 & 1 & -3 \\ 0 & 0 & 0 & 0 & 4 & -8 \end{array}\right) \text{ tausche 2. und 3. Zeile}$$

$$(A^{(2)}|b^{(2)}) = \left(\begin{array}{ccccc|c} 2 & 4 & 4 & -2 & 0 & 6 \\ 0 & 0 & 1 & 4 & 1 & -3 \\ 0 & 0 & 0 & 0 & 4 & -8 \\ 0 & 0 & 0 & 0 & 4 & -8 \end{array}\right) \begin{array}{l} \\ \\ \\ \lambda = -1 \end{array}$$

$$\left(\hat{A}|\hat{b}\right) = (A^{(3)}|b^{(3)}) = \left(\begin{array}{ccccc|c} 2 & 4 & 4 & -2 & 0 & 6 \\ 0 & 0 & 1 & 4 & 1 & -3 \\ 0 & 0 & 0 & 0 & 4 & -8 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$\Rightarrow \quad rg\left(\hat{A}\right) = rg(A) = 3,\ rg\left(\hat{A}|\hat{b}\right) = rg(A|b) = 3,$ da $\left(\hat{b}\right)_4 = 0$

$\Rightarrow \quad Ax = b$ ist lösbar. Die Lösung ist wegen $rg(A|b) = 3 < 5 =$ Anzahl der Unbekannten nicht eindeutig!

Rückwärts auflösen:

$$\begin{array}{lcl} 4x_5 = -8 & \Rightarrow & x_5 = -2 \\ x_3 + 4x_4 + x_5 = x_3 + 4x_4 - 2 = -3 & \Rightarrow & x_3 = -1 - 4x_4 \\ 2x_1 + 4x_2 + 4x_3 - 2x_4 = 2x_1 + 4x_2 + 4(-1 - 4x_4) = 6 & \Rightarrow & x_1 = 5 - 2x_2 + 9x_4 \end{array}$$

Lösungsgesamtheit: $\mathbb{L} = \{x \in \mathbb{R}^5 \mid x_5 = -2,\ x_3 = -1 - 4x_4,\ x_1 = 5 - 2x_2 + 9x_4\}, \quad x_2, x_4$ frei wählbar.

5.2.3 Matrixgleichung und Inversenberechnung

Das Gauß'sche Verfahren lässt sich auch zur Lösung einer Matrixgleichung verwenden.
$AX = B,\ A \in \mathbb{R}^{M,N},\ X \in \mathbb{R}^{N,K},\ B \in \mathbb{R}^{M,K}$

Stellt man die Matrizen in der Form $X = (x^1, x^2, x^3, \dots, x^K), x^i \in \mathbb{R}^N$ i-te Spalte von X. $B = (b^1, b^2, \dots, b^K)$, $b^i \in \mathbb{R}^M$ i-te Spalte von B, $i = 1, \dots, K$ dann ist die Lösung von $AX = B$ äquivalent zur Lösung von $Ax^i = b^i,\ i = 1, \dots, K$ d.h. $AX = B$ entspricht K-gewöhnlichen linearen Gleichungssystemen mit einer Matrix A und K-rechten Seiten $b^i,\ i = 1, \dots, K$

Bringe hier für die erweiterte Matrix $(A|B)$ auf Zeilenstufenform.
Wichtiger Spezialfall: Sei $M = N = K$ und sei $A \in \mathbb{R}^{N,N}$ invertierbar. Dann ergibt sich die Inverse A^{-1} als Lösung von $A \cdot X = I_N$, d.h. ich wende den Gauß-Algorithmus auf $(A|I_N)$ an.

5.2.4 Das Gauß-Jordan Verfahren

Das Zielsystem des Gauß-Jordan Verfahrens ist nicht eine Matrix $\hat{A}$ in Zeilenstufenform, sondern die Identität. In jedem Eleminationsschritt werden nicht nur die Elemente unterhalb des Pivot-Elements transformiert, sondern auch alle Elemente oberhalb. Ferner wird das Pivot-Element auf 1 transformiert.

Das Gauß-Jordan Verfahren liefert für $A \in \mathbb{R}^{N,N}$, A invertierbar, eine Sequenz $(A|b) = \left(A^{(0)}|b^{(0)}\right) \sim \left(A^{(1)}|b^{(1)}\right) \sim \ldots \sim \left(A^{(N)}|b^{(N)}\right) = \left(I|A^{-1}b\right)$

Beispiel zur Inversenberechnung: $A = \begin{pmatrix} 3 & 1 & 6 \\ 1 & 1 & 1 \\ 2 & 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$$\left(A^{(0)}|B^{(0)}\right) = \left(\begin{array}{ccc|ccc} 3 & 1 & 6 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \lambda = -\frac{1}{3} \\ \lambda = -\frac{2}{3} \end{array} \sim \left(\begin{array}{ccc|ccc} 3 & 1 & 6 & 1 & 0 & 0 \\ 0 & \frac{2}{3} & -1 & -\frac{1}{3} & 1 & 0 \\ 0 & \frac{1}{3} & -1 & -\frac{2}{3} & 0 & 1 \end{array}\right) \sim$$

$$\left(\begin{array}{ccc|ccc} 1 & \frac{1}{3} & 2 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{2}{3} & -1 & -\frac{1}{3} & 1 & 0 \\ 0 & \frac{1}{3} & -1 & -\frac{2}{3} & 0 & 1 \end{array}\right) \begin{array}{l} \lambda = -\frac{1}{2} \\ \\ \lambda = -\frac{1}{2} \end{array} \sim \left(\begin{array}{ccc|ccc} 1 & 0 & \frac{5}{2} & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & \frac{2}{3} & -1 & -\frac{1}{3} & 1 & 0 \\ 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 1 \end{array}\right) \sim \ldots \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & -3 & 5 \\ 0 & 1 & 0 & 1 & 3 & -3 \\ 0 & 0 & 1 & 1 & 1 & -2 \end{array}\right)$$

$$\Rightarrow \quad X = A^{-1} = \begin{pmatrix} -2 & -3 & 5 \\ 1 & 3 & -3 \\ 1 & 1 & -2 \end{pmatrix}$$

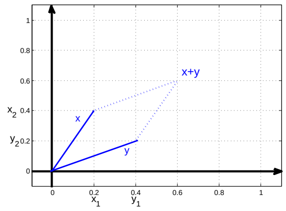
(ober- und unterhalb des Pivot-Elementes Nullen erzeugen und das Pivot-Element jedes mal normieren)

5.3 Determinanten

5.3.1 Motivation

Seien $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$

Gesucht: Fläche bzw. Volumen des Parallelogramms



Sei $V = V(x, y)$ das Volumen des Parallelogramms zu x und y. Man erhält $V(x, y) = x_1 y_2 - y_1 x_2$.

Im allgemeinen hat das Volumen V folgende Eigenschaften:

- V ist linear in jeder Variablen, d.h. $V(u + \alpha v, y) = V(u, y) + \alpha V(v, y)$ (analog in der 2. Variablen)
- Sind x und y linear abhängig, so ist $V(x, y) = 0$
- Für die Einheitsvektoren $e^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ e^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ gilt: $V(e^1, e^2) = 1$

Eine Funktion $V = V(x, y)$ mit diesen drei Eigenschaften heißt Determinante. Man schreibt dann $\det(x, y) = x_1 y_2 - x_2 y_1$.

Genauer beschreibt die Determinante das orientierte Volumen des von x und y aufgespannten Parallelogramm. Das Volumen selbst ist dann durch den Betrag der Determinante gegeben.

Für eine Matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}^{2,2}$ setze $a^i \widehat{=}$ die i-te Spalte von A, $a^1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \ a^2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$

Setze $\det(A) = \det(a^1, a^2) = (a^1)_1 (a^2)_2 - (a^1)_2 (a^2)_1 = a_{11} a_{22} - a_{12} a_{21}$

Verallgemeinerung dieses Konzepts (allgemeines N):
Seien $v^1, v^2, \ldots, v^N \in \mathbb{R}^N$

Die Determinante ist die Volumenfunktion des durch $v^1, v^2, \ldots, v^N$ in $\mathbb{R}^N$ aufgespannten "N-dimensionalen Parallelogramms"

Es gilt wieder:

- $\det(v^1, v^2, \ldots, v^N)$ ist linear in jeder Variablen
- $\det(v^1, v^2, \ldots, v^N) = 0$, falls $v^1, \ldots, v^N$ linear abhängig
- $\det(e^1, \ldots, e^N) = 1$

Für eine Matrix $A = (a^1, a^2, \ldots, a^N) \in \mathbb{R}^{N,N}, \ a^i \widehat{=}$ i-te Spalte von A, $i = 1, \ldots, N$ setzt man
$\det(A) = \det(a^1, a^2, \ldots, a^N)$

5.3.2 Eigenschaften von Determinanten

- Addiert man zu einer Zeile oder Spalte ein Vielfaches einer anderen Zeile / Spalte, so ändert sich der Wert der Determinante nicht.
- Vertauscht man zwei Zeilen oder Spalten, so multipliziert sich die Determinante mit -1.
- Multiplizieren einer Zeile / Spalte mit einem Faktor $\lambda \neq 0$ liefert das λ-fache der Determinante.

- Die Determinante einer Dreiecksmatrix ist das Produkt der Diagonalelemente.

$$\underbrace{\begin{pmatrix} * & & 0 \\ \vdots & \ddots & \\ * & \dots & * \end{pmatrix}}_{\text{linke untere Dreiecksmatrix}} \qquad \underbrace{\begin{pmatrix} * & \dots & * \\ & \ddots & \vdots \\ 0 & & * \end{pmatrix}}_{\text{rechte obere Dreiecksmatrix}}$$

<u>Beispiel:</u> $A = \begin{pmatrix} 0 & 2 & -1 \\ 3 & -4 & 5 \\ 6 & -6 & 12 \end{pmatrix}$ (vgl. Bsp. Rang aus 5.1.4)

Mit Hilfe elementarer Zeilenumformungen und einmaligem Vertauschen von zwei Zeilen haben wir A auf Zeilenstufenform
$\hat{A} = \begin{pmatrix} 6 & -6 & 12 \\ 0 & -1 & -1 \\ 0 & 0 & -3 \end{pmatrix}$
gebracht. Somit gilt
$\det(A) = -\det(\hat{A}) = \underbrace{-1}_{\text{Zeile getauscht}} \cdot (6 \cdot -1 \cdot -3) = -18$

5.3.3 La Place Entwicklungen für Determinanten

<u>Ansatz:</u>

$N = 1$ sei $A \in \mathbb{R}^{1,1}$, $A = (a_{11})$ Setze: $\det(A) = a_{11}$

$N = 2$ sei $A \in \mathbb{R}^{2,2}$, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ Setze: $\det(A) = a_{11}a_{22} - a_{21}a_{12}$

$N \geq 3$ sei $A \in \mathbb{R}^{N,N}$, $A = \begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \dots & a_{N,N} \end{pmatrix} \in \mathbb{R}^{N,N}$

Zu $A \in \mathbb{R}^{N,N}$ sei A^{ij} diejenige $(N-1) \times (N-1)$-Matrix, welche aus A durch Streichen der i-ten Zeile und j-ten Spalte hervorgeht.

$$A = \begin{pmatrix} 2 & 4 & \underline{6} \\ \underline{1} & \underline{2} & \underline{0} \\ 7 & -1 & \underline{1} \end{pmatrix} \in \mathbb{R}^{3,3} \quad \Rightarrow \quad A^{23} = \begin{pmatrix} 2 & 4 \\ 7 & -1 \end{pmatrix} \in \mathbb{R}^{2,2}$$

Dann gilt für jedes $i = 1, \dots, N$ durch "Entwicklung nach Zeile i":

$$\det(A) = \sum_{j=1}^{N} (-1)^{i+j} a_{ij} \det(A^{ij})$$

und für jedes $j = 1, \dots, N$ durch "Entwicklung nach Spalte j":

$$\det(A) = \sum_{i=1}^{N} (-1)^{i+j} a_{ij} \det(A^{ij})$$

Das Vorzeichen von $(-1)^{i+j}$ lässt sich mit Hilfe des Schachbrettmusters bestimmen:

$$\begin{pmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \end{pmatrix} \quad \text{"Schachbrettmuster"}$$

<u>Beispiel:</u> $A = \begin{pmatrix} 0 & 2 & -1 \\ 3 & -4 & 5 \\ 6 & -6 & 12 \end{pmatrix}$ Entwicklung nach Zeile 1 (wegen $a_{11} = 0$)

$$\det(A) = (-1)^{1+1} \cdot 0 \cdot \begin{vmatrix} -4 & 5 \\ -6 & 12 \end{vmatrix} + (-1)^{1+2} \cdot 2 \cdot \begin{vmatrix} 3 & 5 \\ 6 & 12 \end{vmatrix} + (-1)^{1+3} \cdot (-1) \cdot \begin{vmatrix} 3 & -4 \\ 6 & -6 \end{vmatrix}$$
$$= -2(36-30) - (-18+24) = -12 - 6 = -18$$

Das Volumen des von den Spalten der Matrix A aufgespannten Parallelogramms hat $|\det(A)| = 18$ Einheiten.

5.3.4 Spezielle Klassen von Matrizen

Sei $A \in \mathbb{R}^{M,N},\ A = (a_{ij}),\ i = 1, \ldots, M,\ j = 1, \ldots, N$

Dann ist die transponierte Matrix definiert durch: $A^T = (a_{ji}) \in \mathbb{R}^{N,M}$

Beispiel: $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \in \mathbb{R}^{3,2}, \quad A^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix} \in \mathbb{R}^{2,3}$ "Transponierte von A"

Jede Zeile wird zur Spalte / jede Spalte wird zur Zeile!

Rechenregeln für die Transposition:

- $(A^T)^T = A$,
- $(\alpha A)^T = \alpha A^T,\ \alpha \in \mathbb{R}$,
- $(A+B)^T = A^T + B^T$,
- $(A \cdot B)^T = B^T \cdot A^T$

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix} = A$$

- Eine Matrix $A \in \mathbb{R}^{N,N}$ heißt symmetrisch, falls $A = A^T$
- Eine Matrix $A \in \mathbb{R}^{N,N}$ heißt idempotent, wenn $A^2 = A \cdot A = A$
- Eine reguläre Matrix $A \in \mathbb{R}^{N,N}$ heißt orthogonal, falls $A^{-1} = A^T$

Bemerkung: Hesse-Matrizen an einer festen Stelle sind symmetrisch.

$f : D \to \mathbb{R},\ D \subset \mathbb{R}^N$ zweimal differenzierbar
$\nabla f(x) \in \mathbb{R}^N,\ x \in D$ "Gradient von f"
$\nabla^2 f(x) \in \mathbb{R}^{N,N},\ x \in D$ "Hesse-Matrix von f"

An jeder Stelle $x \in D$ ist $\nabla^2 f(x)$ symmetrisch!

Rechenregeln für Determinanten: Seien $A, B \in \mathbb{R}^{N,N}$

i) $\det(A \cdot B) = \det(A) \cdot \det(B)$

ii) Ist A invertierbar, so gilt: $1 = \det(I_N) = \det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1})$, d.h $\det(A^{-1}) = 1/\det(A)$.

iii) $\det(A) = 0$, falls $rg(A) < N$ (d.h. A nicht invertierbar)

iv) $\det(A) = \det(A^T)$

v) Sei $A \in \mathbb{R}^{N,N}$ orthogonal.

$$1 = \det(I_N) = \det(A \cdot A^{-1}) = \det(A) \cdot \det(\underbrace{A^{-1}}_{=A^T}) = \det(A) \cdot \underbrace{\det(A^T)}_{=\det(A)} = \det(A)^2$$

$$\Rightarrow \quad \det(A) = \pm 1$$

5.3.5 Lösen von linearen Gleichungssystemen mit Determinanten

Sei $A = (a_{ij}) \in \mathbb{R}^{N,N}$ invertierbar und $b = (b_i) \in \mathbb{R}^N$. Betrachte $Ax = b$

Für $j = 1, \ldots, N$ sei $A_{j,b} \in \mathbb{R}^{N,N}$ diejenige Matrix, welche entsteht, wenn in A die j-te Spalte durch den Vektor b ersetzt wird. Dann gilt für die eindeutige Lösung $x = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = A^{-1}b$ die Darstellung:

$$x_j = \frac{\det(A_{j,b})}{\det(A)}, \quad j = 1, \ldots, N \quad \text{"Cramer'sche Regel"}$$

Beachte: $\det(A) \neq 0$, da A invertierbar.

Bemerkung: Diese Methode wird z.B. angewandt, wenn man nur eine Komponente der Lösung berechnen, und nicht die gesamte Lösung ausrechnen möchte.

Beispiel: $A = \begin{pmatrix} 2 & 4 & -1 \\ 1 & -3 & 2 \\ 6 & 5 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 15 \\ -5 \\ 28 \end{pmatrix}$, Gesucht: $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

$$\det(A) = 2 \cdot \begin{vmatrix} -3 & 2 \\ 5 & 1 \end{vmatrix} - 4 \cdot \begin{vmatrix} 1 & 2 \\ 6 & 1 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & -3 \\ 6 & 5 \end{vmatrix} = 2(-3-10) - 4(1-12) - 1(5+18) = -26 + 44 - 23 = \underline{-5}$$

$$\det(A_{1,b}) = \det \begin{pmatrix} 15 & 4 & -1 \\ -5 & -3 & 2 \\ 28 & 5 & 1 \end{pmatrix} = 15 \cdot \begin{vmatrix} -3 & 2 \\ 5 & 1 \end{vmatrix} - 4 \cdot \begin{vmatrix} -5 & 2 \\ 28 & 1 \end{vmatrix} - 1 \cdot \begin{vmatrix} -5 & -3 \\ 28 & 5 \end{vmatrix}$$

$$= 15(-3-10) - 4(-5-56) - (-25+84) = -195 + 244 - 59 = -10 \quad \Rightarrow \quad x_1 = \tfrac{-10}{-5} = 2$$

$$\det(A_{2,b}) = -15 \quad \Rightarrow \quad x_2 = \tfrac{-15}{-5} = 3$$

$$\det(A_{3,b}) = -5 \quad \Rightarrow \quad x_3 = \tfrac{-5}{-5} = 1$$

$$\Rightarrow \quad x = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$$

5.3.6 Quadratische Formen

Sei $A \in \mathbb{R}^{N,N}$ eine symmetrische Matrix, d.h. $A = A^T$.
Eine Abbildung $Q : \mathbb{R}^N \to \mathbb{R}$, $Q(x) = \underbrace{x^T \cdot \underbrace{A \cdot x}_{\in \mathbb{R}^N}}_{\in \mathbb{R}}$ heißt quadratische Form in N Variablen.

Definition:

Eine quadratische Form $Q(x) = x^T \cdot A \cdot x = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} x_i x_j$ bzw. die symmetrische Matrix heißt:

i) positiv definit (negativ definit), wenn $Q(x) = x^T \cdot A \cdot x > 0 \ (< 0)$ für alle $x \in \mathbb{R}^N$, $x \neq 0$

ii) positiv semidefinit (negativ semidefinit), wenn $Q(x) = x^T \cdot A \cdot x \geq 0 \ (\leq 0)$ für alle $x \in \mathbb{R}^N$

Bemerkung: $Q(x)$ heißt indefinit, falls A regulär und $Q(x)$ weder positiv noch negativ definit ist. Ist A nicht regulär, so heißt $Q(x)$ entartet.

Beispiel: $N = 2, \quad A = \begin{pmatrix} 4 & -2 \\ -2 & 8 \end{pmatrix} \in \mathbb{R}^{2,2}$ symmetrisch , $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$$Q(x) = x^T A x = (x_1, x_2)^T \begin{pmatrix} 4 & -2 \\ -2 & 8 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1, x_2)^T \cdot \begin{pmatrix} 4x_1 - 2x_2 \\ -2x_1 + 8x_2 \end{pmatrix} = 4x_1^2 - 2x_1x_2 - 2x_1x_2 + 8x_2^2$$

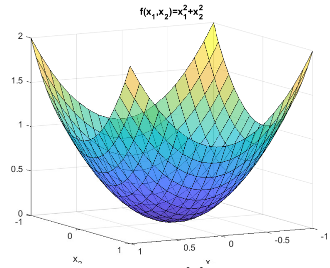
$$= 2x_1^2 + (2x_1^2 - 4x_1x_2 + 2x_2^2) + 6x_2^2 = 2x_1^2 + 2(x_1 - x_2)^2 + 6x_2^2 \geq 0 \quad \forall\, x \in \mathbb{R}^2 \quad \text{(positiv semidefinit)}$$

Sei $Q(x_1, x_2) = 2x_1^2 + 2(x_1 - x_2)^2 + 6x_2^2 = 0 \quad \Rightarrow \quad x_2 = 0,\ x_1 - x_2 = 0,\ x_1 = 0 \quad \Rightarrow \quad x = 0$

$$Q(x) = \begin{cases} > 0 & \text{für } x \in \mathbb{R}^2 \backslash \{0\} \\ 0 & \text{für } x = 0 \end{cases} \quad \Rightarrow \quad Q(x) \text{ positiv definit}$$

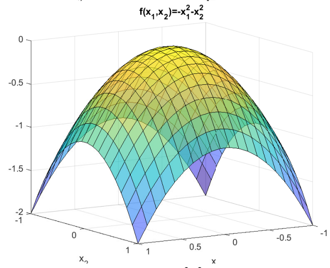
Visualisierung von Q für $N = 2$ nahe $x = 0$

$(x, Q(x)) = (x_1, x_2, Q(x_1, x_2)) \in \mathbb{R}^3$



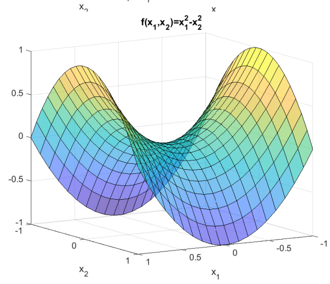
$Q(x)$ positiv definit

$(0, 0)$ lokales Minimum von Q.



$Q(x)$ negativ definit

$(0, 0)$ lokales Maximum von Q.



$Q(x)$ indefinit

$(0, 0)$ ist weder lokales Minimum noch lokales Maximum von Q

Genauer: $(0, 0)$ ist Sattelpunkt von Q.

Kriterien für Definitheit von quadratischen Formen bzw. symmetrischen Matizen

Satz:

Eine quadratische Form $Q(x) = x^T A x$, bzw. die symmetrische Matrix A ist positiv (negativ) definit genau dann, wenn

$$\det(A_K) > 0, \quad K = 1, \dots, N \quad ((-1)^K \det(A_K) > 0, \quad K = 1, \dots, N)$$

Dabei bezeichnet $A_K = \begin{pmatrix} a_{11} & \dots & a_{1K} \\ \vdots & \ddots & \vdots \\ a_{K1} & \dots & a_{KK} \end{pmatrix}, \quad K = 1, \dots, N$

die Hauptabschnittsmatrix von $A \in \mathbb{R}^{N,N}$

$$A_1 = (a_{11}), \quad A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad A_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \dots, \quad A_N = A$$

Beispiel: $A = \begin{pmatrix} 4 & -2 \\ -2 & 8 \end{pmatrix}$

$\Rightarrow \quad A_1 = 4, \det(A_1) = 4, \quad A_2 = A, \det(A_2) = 4 \cdot 8 - 4 > 0$

$\Rightarrow$ A positiv definit

Beispiel: $A = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \in \mathbb{R}^{N,N}$

Nach Aufgabe 5, Blatt 2 gilt: $\det(A_i) = i + 1, \quad i = 1, \dots, N$

Nun gilt: $\det(A_i) = i + 1 > 0$, d.h. A ist positiv definit.

5.3.7 Quadratische Formen auf Unterräumen bzw. mit Nebenbedingungen

Oft sind in der Optimierung quadratische Formen mit Nebenbedingungen zu analysieren.

Betrachte: $Q(x) = x^T Ax$ mit $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathbb{R}^{2,2} \quad \Rightarrow \quad Q(x) = ax_1^2 + 2bx_1x_2 + cx_2^2$

Angenommen die Variablen x_1, x_2 unterliegen der Nebenbedingung

$Bx = px_1 + qx_2 = 0, \quad B = (p, q)$

Löse die NB auf: $x_2 = \frac{-px_1}{q}, \quad q \neq 0$

Einsetzen liefert:

$$Q\left(x_1, \frac{-px_1}{q}\right) = ax_1^2 + 2bx_1\left(\frac{-px_1}{q}\right) + c\left(\frac{-px_1}{q}\right)^2 = ax_1^2 - \frac{2bp}{q}x_1^2 + \frac{cp^2}{q^2}x_1^2 = \frac{1}{q^2}\underbrace{(aq^2 - 2bpq + cp^2)}_{>0 \ \Leftrightarrow \ Q>0}x_1^2$$

Definiere: $C = \begin{pmatrix} 0 & B \\ B^T & A \end{pmatrix} = \begin{pmatrix} 0 & p & q \\ p & a & b \\ q & b & c \end{pmatrix}$

Nun gilt:

$$\begin{aligned} \det(C) &= \begin{vmatrix} 0 & p & q \\ p & a & b \\ q & b & c \end{vmatrix} = 0 \cdot \begin{vmatrix} a & b \\ b & c \end{vmatrix} - p \cdot \begin{vmatrix} p & b \\ q & c \end{vmatrix} + q \cdot \begin{vmatrix} p & a \\ q & b \end{vmatrix} \\ &= -p(pc - qb) + q(pb - aq) = -p^2c + qbp + qbp - aq^2 \\ &= -(p^2c - 2qbp + aq^2) \end{aligned}$$

Allgemeiner Fall: Sei $A = (a_{ij}), \quad 1 \leq i, j \leq N,$ A symmetrisch

$Q(x) = x^T Ax = \sum\limits_{i=1}^{N} \sum\limits_{j=1}^{N} a_{ij}x_ix_j$ zugehörige quadratische Form.

Ferner sei $B = (b_{ij}), \ 1 \leq i \leq l, \ 1 \leq j \leq N, \ B \in \mathbb{R}^{l,N}, \ l < N$ eine Matrix, welche die Nebenbedingungen

$$Bx = 0, \quad \text{d.h.} \quad \begin{aligned} b_{11}x_1 + b_{12}x_2 + \ldots + b_{1N}x_N &= 0 \\ &\vdots \\ b_{l1}x_1 + b_{l2}x_1 + \ldots + b_{lN}x_N &= 0 \end{aligned} \quad \text{festlegt.}$$

Setze die Matrix $C = \begin{pmatrix} 0 & B \\ B^T & A \end{pmatrix}$, $0 \in \mathbb{R}^{l,l}, \ B \in \mathbb{R}^{l,N}, \ B^T \in \mathbb{R}^{N,l}, \ A \in \mathbb{R}^{N,N}$, d.h.

$C \in \mathbb{R}^{N+l,N+l}$

Definition:

Eine quadratische Form $Q(x) = x^T Ax$ heißt positiv [negativ] definit unter einer Nebenbedingung $Bx = 0$, falls $x^T Ax > [<] \ 0$ für alle $x \in \mathbb{R}^N$ mit $Bx = 0, \ x \neq 0$.

Satz:

Die quadratische Form $Q(x) = x^T Ax$ ist positiv definit unter der Nebenbedingung $Bx = 0$, falls für die (k, k)-Hauptuntermatrizen h_k von C gilt:

$$(-1)^l \det(h_{2l+i}) > 0, \quad i = 1, \ldots, N - l$$

N$\hat{=}$ Anzahl Variablen, l$\hat{=}$ Anzahl Nebenbedingungen

$Q(x)$ ist negativ definit unter der Nebenbedingung $Bx = 0$, falls

$$(-1)^{l+i}\det(h_{2l+i}) > 0, \quad i = 1, \dots, N-l$$

Beispiel: $Q(x) = x^T A x = 3x_1^2 - x_2^2 + 4x_3^2, \quad A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \quad N = 3,\ l = 1$

NB: $Bx = x_1 + x_2 + x_3 = 0, \quad B = (1, 1, 1)$

Finde also $C = \begin{pmatrix} 0 & B \\ B^T & A \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 3 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 4 \end{pmatrix} \in \mathbb{R}^{4,4}$

Zu berechnen sind $\det(h_3) = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 3 & 0 \\ 1 & 0 & -1 \end{vmatrix}$ und $\det(h_4) = \det(C)$

Man findet $\det(h_3) = -2$ und $\det(h_4) = -\begin{vmatrix} 1 & 1 & 1 \\ 3 & 0 & 0 \\ 0 & -1 & 0 \end{vmatrix} + 0 + 0 + 4 \cdot \underbrace{\begin{vmatrix} 0 & 1 & 1 \\ 1 & 3 & 0 \\ 1 & 0 & -1 \end{vmatrix}}_{=\det(h_3)=-2}$

$$= (-1)\begin{vmatrix} 3 & 0 \\ 0 & -1 \end{vmatrix} - 8 = 3 - 8 = -5$$

Wegen $(-1)^1\det(h_3) = 2 > 0, \quad (-1)^1\det(h_4) = 5 > 0$ folgt, dass $Q(x) = x^T A x$ positiv definit unter der Nebenbedingung $Bx = 0$ ist.

5.4 Eigenwerte und Eigenvektoren

5.4.1 Definition und Eigenschaften

Sei $A \in \mathbb{R}^{N,N}$. Dann heißt

$$\begin{array}{lll} Ker(A) & = \{x \in \mathbb{R}^N \mid Ax = 0\} & \text{"Kern von A"} \\ Bi(A) & = \{y \in \mathbb{R}^N \mid \exists\, x \in \mathbb{R}^N : Ax = y\} & \text{"Bild von A"} \end{array}$$

Es gilt immer: $0 \in Ker(A), \quad 0 \in Bi(A)$

Satz: Für $A \in \mathbb{R}^{N,N}$ sind gleichwertig:

i) A ist invertierbar

ii) $\det(A) \neq 0$

iii) $Ker(A) = \{0\}$

iv) $Bi(A) = \mathbb{R}^N$

Definition (Eigenwert, Eigenvektor):

Sei $A \in \mathbb{R}^{N,N}$. Ein Vektor x und eine Zahl λ heißen Eigenvektor und Eigenwert von A, falls $x \neq 0$ und $Ax = \lambda x$.

Eigenvektoren x werden unter A auf ein Vielfaches von sich selbst abgebildet.

Bemerkung: Der Eigenvektor x ist nicht eindeutig, denn

$$Ax = \lambda x \quad \Leftrightarrow \quad A(\alpha x) = \lambda(\alpha x),\ \alpha \neq 0 \quad \Leftrightarrow \quad y = \alpha x,\ Ay = \lambda y$$

Bedeutung: $Ax = \lambda x \quad \Leftrightarrow \quad Ax - \lambda x = 0 \quad \Leftrightarrow \quad (A - \lambda\, I_N)x = 0,\ x \neq 0$

d.h. $x \in Ker(A - \lambda\, I_N),\ x \neq 0$

Dies bedeutet: $A - \lambda\, I_N$ ist nicht invertierbar bzw. $\det(A - \lambda\, I_N) = 0$

Die Matrix $A - \lambda I_N$ hat folgende Form:

$$A - \lambda\, I_N = \begin{pmatrix} a_{11} - \lambda & a_{12} & a_{13} & \dots & a_{1N} \\ a_{21} & a_{22} - \lambda & a_{23} & \dots & a_{2N} \\ a_{31} & a_{32} & a_{33} - \lambda & \dots & a_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & a_{N3} & \dots & a_{NN} - \lambda \end{pmatrix}$$

$p(\lambda) = \det(A - \lambda\, I_N)$ ist ein Polynom vom Grade N, das "charakteristische Polynom" der Matrix A. $p(\lambda)$ kann z.B. nach La Place berechnet werden.

Satz:

λ ist genau dann Eigenwert zu einer Matrix $A \in \mathbb{R}^{N,N}$, wenn λ Nullstelle des charakteristischen Polynoms ist. Jedes $x \in Ker(A - \lambda\, I_N),\ x \neq 0$ ist Eigenvektor von A zum Eigenwert λ.

Bezeichnung: $E(\lambda, A) = Ker(A - \lambda\, I_N),\ \lambda$ Eigenwert

heißt Eigenraum zum Eigenwert λ von A.

Beispiel: $A = \begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{pmatrix}$ Gesucht: Eigenwerte und Eigenvektoren von A

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I_3) \\ &= \begin{vmatrix} -1-\lambda & 2 & 2 \\ 2 & 2-\lambda & 2 \\ -3 & -6 & -6-\lambda \end{vmatrix} \\ &= (-1-\lambda)\begin{vmatrix} 2-\lambda & 2 \\ -6 & -6-\lambda \end{vmatrix} - 2\begin{vmatrix} 2 & 2 \\ -3 & -6-\lambda \end{vmatrix} + 2\begin{vmatrix} 2 & 2-\lambda \\ -3 & -6 \end{vmatrix} \\ &= (-1-\lambda)\,[(2-\lambda)(-6-\lambda)+12] - 2\,[2(-6-\lambda)+6] + 2\,[-12+3(2-\lambda)] \\ &= \ldots = -\lambda(\lambda+2)(\lambda+3) \end{aligned}$$

$\Rightarrow$ A hat die Eigenwerte: $\lambda_1 = 0, \quad \lambda_2 = -2, \quad \lambda_3 = -3$

$\lambda_1 = 0$: $(A - \lambda_1 I_3)\cdot v^1 = Av^1 = 0$

Gauss-Elimination: $\left(\begin{array}{ccc|c} -1 & 2 & 2 & 0 \\ 2 & 2 & 2 & 0 \\ -3 & -6 & -6 & 0 \end{array}\right) \begin{array}{l} \\ \lambda = 2 \\ \lambda = -3 \end{array} \sim \left(\begin{array}{ccc|c} -1 & 2 & 2 & 0 \\ 0 & 6 & 6 & 0 \\ 0 & -12 & -12 & 0 \end{array}\right) \begin{array}{l} \\ \\ \lambda = 2 \end{array} \sim \left(\begin{array}{ccc|c} -1 & 2 & 2 & 0 \\ 0 & 6 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$

Bei der Berechnung von Eigenvektoren muss $(A - \lambda\, I_N)$ nach Gauss-Elimination immer mindestens eine Nullzeile haben!

$$\begin{array}{lll} \text{Normiere:} & & (v^1)_3 = 1 \\ 6(v^1)_2 + 6\cdot 1 = 0 & \Rightarrow & (v^1)_2 = -1 \\ -(v^1)_1 - 2 + 2 = 0 & \Rightarrow & (v^1)_1 = 0 \end{array} \quad \Rightarrow \quad (v^1) = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \text{ Eigenvektor zu } \lambda_1 = 0$$

$\lambda_2 = -2$: $(A - \lambda_2 I_3)v^2 = (A + 2I_3)v^2 = 0$

$\left(\begin{array}{ccc|c} 1 & 2 & 2 & 0 \\ 2 & 4 & 2 & 0 \\ -3 & -6 & -4 & 0 \end{array}\right) \begin{array}{l} \\ \lambda = -2 \\ \lambda = 3 \end{array} \sim \left(\begin{array}{ccc|c} 1 & 2 & 2 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right) \begin{array}{l} \\ \\ \lambda = 1 \end{array} \sim \left(\begin{array}{ccc|c} 1 & 2 & 2 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$

$$\begin{array}{lll} -2(v^2)_3 = 0 & \Rightarrow & (v^2)_3 = 0 \\ \text{Normiere:} & & (v^2)_2 = 1 \\ (v^2)_1 + 0 + 2 = 0 & \Rightarrow & (v^2)_1 = -2 \end{array} \quad \Rightarrow \quad v^2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$$

$\lambda_3 = -3$: Analog folgt: $v^3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$

Beispiel: $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2,2}$ Gesucht: Eigenwerte und Eigenvektoren von A.

charakteristisches Polynom: $p(\lambda) = \det(A - \lambda\, I_N) = \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = (-\lambda)^2 + 1 = \lambda^2 + 1 \overset{!}{=} 0 \quad \Leftrightarrow \quad \lambda^2 = -1$

Das Polynom $p(\lambda)$ hat in $\mathbb{R}$ keine Nullstellen!

5.4.2 Komplexe Zahlen

Führe formal die imaginäre Einheit "i" ein als Lösung der Gleichung $i^2 = -1$.

In diesem Fall hat die Gleichung $p(\lambda) = \lambda^2 + 1 = 0$ die Lösung $\lambda_{1,2} = \pm i$

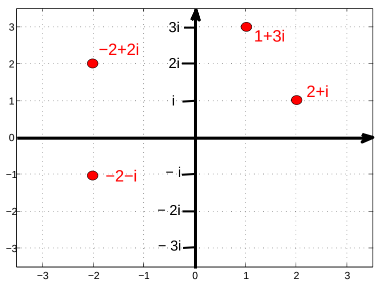
Allgemein nennt man eine Zahl $z = a + bi \quad a, b \in \mathbb{R}$, i imaginäre Einheit
eine komplexe Zahl und setzt $\mathbb{C} = \{z = a + bi,\ \ a, b \in \mathbb{R}\}$ "Menge der komplexen Zahlen"
Offensichtlich gilt $\mathbb{R} \subset \mathbb{C}$.

Bezeichnung: $a = Re(z) \hat{=}$ Realteil von z

$b = Im(z) \hat{=}$ Imaginärteil von z

Vorstellung: Stelle die beiden reellen Zahlen a,b einer komplexen Zahl $z = a + bi$ als Vektor (a, b) dar und interpretiere sie als Punkt im $\mathbb{R}^2$

"Gauß'sche Zahlenebene"



Addition komplexer Zahlen: $(a_1 + b_1 i) + (a_2 + b_2 i) := (a_1 + a_2) + i(b_1 + b_2)$

Multiplikation komplexer Zahlen:

$(a_1 + b_1 i) \cdot (a_2 + b_2 i) = a_1 a_2 + a_1 b_2 i + a_2 b_1 i + i^2 b_1 b_2 = a_1 a_2 - b_1 b_2 + i(a_1 b_2 + a_2 b_1)$

Für die Addition und Multiplikation komplexer Zahlen gelten alle Eigenschaften, welche für die reellen Zahlen gültig sind. $\mathbb{C}$ bzw. genauer $(\mathbb{C}, +, \cdot)$ heißt deshalb Körper der komplexen Zahlen.

Weitere Setzungen:

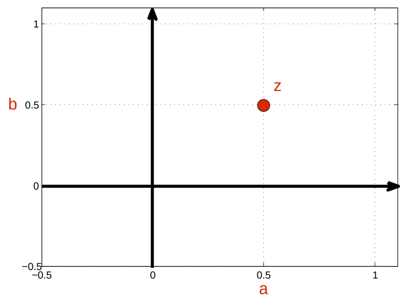
Zu $z = a + bi$ heißt $\bar{z} = a - bi$ die konjugiert komplexe Zahl

Rechenregeln:

$$\overline{z + w} = \bar{z} + \bar{w} \quad w, z \in \mathbb{C}$$

$$\overline{z \cdot w} = \bar{z} \cdot \bar{w}$$

$$\bar{\bar{z}} = z$$

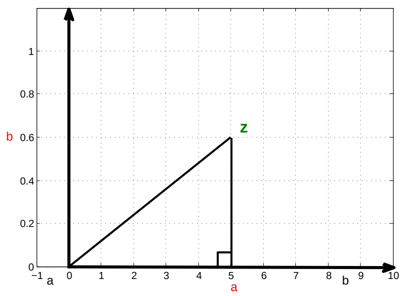


Für jedes $z \in \mathbb{C}$ gilt: $z \cdot \overline{z} = (a+bi)(a-bi) = a^2 - i^2b^2 = a^2 + b^2 \quad \in \mathbb{R}$

Man definiert

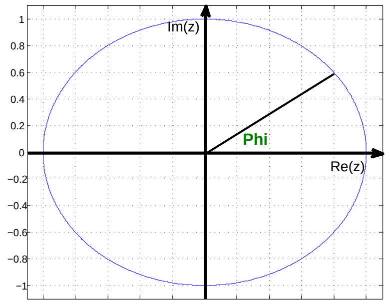
$|z| = \sqrt{z \cdot \overline{z}} = \sqrt{a^2 + b^2}$

"Betrag der komplexen Zahl z"



Behalte die Anschauung der Gauß'schen Zahlenebene bei.
Jede komplexe Zahl z mit $|z| = 1$ lässt sich schreiben als $z = E(\Phi) = \cos(\Phi) + i\sin(\Phi)$

Φ ist der Winkel zwischen $z = E(\Phi)$ und der reellen Achse im Bogenmaß



Man rechnet leicht nach, dass gilt: $E(\Phi) \cdot E(\Psi) = E(\Phi + \Psi), \quad E(0) = 1$

<u>Komplexe exp-Funktion:</u>

$$E(\Phi) = \cos(\Phi) + i\sin(\Phi) = \exp(i\Phi) = e^{i\Phi} \quad \text{"Eulersche Formel"}$$

Man kann nun komplexe Zahlen stets schreiben als:

$$\begin{aligned} z &= |z| \cdot E(\Phi) = |z| \cdot E(\Phi) = |z| \exp(i\Phi) \\ z \cdot w &= |z| \cdot |w| \cdot E(\Phi) \cdot E(\Psi) = |z||w| \cdot \exp(i\Phi) \cdot \exp(i\Psi) \\ &= |z||w| \cdot E(\Phi + \Psi) = |z||w| \cdot \exp(i(\Phi + \Psi)) \end{aligned}$$

Die Multiplikation von komplexen Zahlen ist also geometrisch eine Drehstreckung!

Sei nun $z = a + bi$.
Dann findet man $|z| = (a^2 + b^2)^{\frac{1}{2}}$ und $\tan(\Phi) = \frac{b}{a}, \quad a > 0$ d.h. $\Phi = \arctan(\frac{b}{a}), \quad a > 0$

Im allgemeinen findet man: $\begin{cases} \arctan(\frac{b}{a}), & \text{falls } a > 0 \\ \arctan(\frac{b}{a}) + \pi, & \text{falls } a < 0,\ b \geq 0 \\ \arctan(\frac{b}{a}) - \pi & \text{falls } a < 0,\ b < 0 \\ \pi/2, & \text{falls } a = 0,\ b > 0 \\ -\pi/2, & \text{falls } a = 0,\ b < 0 \end{cases}$

Somit gilt (zumindest für $a \geq 0$) :

$$z = a + bi = \underbrace{(a^2+b^2)^{\frac{1}{2}}}_{r} \cdot \left(\cos\left(\underbrace{\arctan\left(\frac{b}{a}\right)}_{\Phi} \right) + i \sin\left(\underbrace{\arctan\left(\frac{b}{a}\right)}_{\Phi} \right) \right)$$

$z = r(\cos(\Phi) + i\sin(\Phi))$

$z = r \cdot \exp(i \cdot \Phi)$ "Polarkoordinatendarstellung"

Es lässt sich zeigen, dass diese Formel auch für $a \leq 0$, d.h. für alle komplexen Zahlen $z = a + bi \neq 0$ gilt. Damit erhält man für die Potenzen von z folgende Darstellung:

$$z^n = r^n \cdot (\cos(\Phi) + i\sin(\Phi))^n = r^n \cdot \exp(i\Phi)^n = r^n \cdot \exp(i\,n\Phi) = r^n \cdot (\cos(n\Phi) + i\sin(n\Phi))$$

5.4.3 Komplexe Eigenwerte

Nach unserem Einschub über komplexe Zahlen kehren wir zum Eigenwertproblem zurück.

Wir haben gesehen, dass eine reelle Matrix also durchaus komplexe Eigenwerte haben kann.

Satz (Fundamentalsatz der Algebra) :

Jedes Polynom $p(t) = \sum_{i=0}^{N} a_i t^i$, $a_0, a_1, a_2, \ldots, a_N \in \mathbb{R}$, $a_N \neq 0$ zerfällt in komplexe Linearfaktoren über $\mathbb{C}$, d.h. es gibt $t_1, \ldots, t_N \in \mathbb{C}$, $\alpha \in \mathbb{R}$ mit

$$p(t) = \alpha(t - t_1)(t - t_2) \cdot \ldots \cdot (t - t_N)$$

Beipiel: $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2,2}$

charakteristisches Polynom: $p(\lambda) = \det(A - \lambda\, I_2)$

$\det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1 = (\lambda - i)(\lambda + i)$

Nullstellen: $\lambda_1 = i,\ \lambda_2 = -i \quad (\lambda_2 = \overline{\lambda_1})$ in $\mathbb{C}$

Bemerkung: i) Die Nullstellen eines reellen Polynoms in $\mathbb{C}$ sind immer konjugiert komplex, d.h., ist μ eine Nullstelle von diesem Polynom $p(\lambda)$, so auch $\bar{\mu}$.

ii) Sind die Eigenwerte einer Matrix komplex, so sind auch die Eigenvektoren komplex. Weiter gilt:

$$Ax = \lambda x \quad \Leftrightarrow \quad \overline{Ax} = \overline{\lambda x} \quad \Leftrightarrow \quad A\bar{x} = \bar{\lambda}\bar{x}, \quad x \in \mathbb{C}^n$$

Berechnung der Eigenvektoren (Gauß - Algorithmus in $\mathbb{C}$):

$$(A - \lambda_1\, I_2)x^1 = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \left(\begin{array}{cc|c} -i & 1 & 0 \\ -1 & -i & 0 \end{array} \right) \; \lambda = i \quad \sim \quad \left(\begin{array}{cc|c} -i & 1 & 0 \\ 0 & 0 & 0 \end{array} \right)$$

Wähle $x_2 = 1$ und finde $\begin{array}{ll} -i \cdot x_1 + x_2 & = -i \cdot x_1 + 1 = 0 \\ -i \cdot x_1 = -1 & \Leftrightarrow \quad x_1 = -i \end{array} \quad \Rightarrow \quad x^1 = \begin{pmatrix} -i \\ 1 \end{pmatrix} \in \mathbb{C}^2$

Weiter gilt

$$Ax^1 = \lambda_1 x^1 \quad \Leftrightarrow \quad A\overline{x^1} = \overline{Ax^1} = \overline{\lambda_1 x^1} = \overline{\lambda_1} \cdot \overline{x^1} = \lambda_2 \overline{x^1} \quad \Rightarrow \quad x^2 = \overline{x^1} = \begin{pmatrix} i \\ 1 \end{pmatrix}$$

,
d.h. ist x^1 komplexer Eigenvektor zum komplexen Eigenwert λ_1, so ist $x^2 = \overline{x^1}$ Eigenvektor zum Eigenwert $\lambda_2 = \overline{\lambda_1}$.

Satz:

i) Sei $A \in \mathbb{R}^{N,N}$ symmetrisch, d.h. $A = A^T$. Dann sind alle Eigenwerte λ von A reell.

ii) Sei $A \in \mathbb{R}^{N,N}$ orthogonal, d.h. $A^{-1} = A^T$. Dann gilt für alle Eigenwerte $\lambda \in \mathbb{C}$ von A: $|\lambda| = 1$

iii) Sei $A \in \mathbb{R}^{N,N}$ idempotent, d.h. $A^2 = A$. Dann hat A die Eigenwerte $\lambda = 0$ oder $\lambda = 1$.

Satz: Sei $A \in \mathbb{R}^{N,N}$ symmetrisch.

i) Dann gibt es eine orthogonale Matrix O mit $= O \cdot A \cdot O^T = \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_N \end{pmatrix}$ und die $\lambda_1, \ldots, \lambda_N \in \mathbb{R}$ sind die Eigenwerte von A.

ii) A ist positiv (negativ) definit [semidefinit] genau dann, wenn $\lambda_i > 0 \quad (\lambda_i < 0) \quad [\lambda_i \geq 0] \quad [(\lambda_i \leq 0)]$ für $i = 1, \ldots, N$. Es gilt:

$$\begin{aligned}
x^T A x &= x^T \cdot \underbrace{O^T \cdot O}_{I_N} \cdot A \cdot \underbrace{O^T \cdot O}_{I_N} \cdot x \\
&= \underbrace{(O \cdot x)^T}_{=y^T \text{ (Vektor)}} \cdot O \cdot A \cdot O^T \cdot \underbrace{O \cdot x}_{=y} \\
&= y^T \cdot \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_N \end{pmatrix} \cdot y \\
&= y^T \cdot \begin{pmatrix} \lambda_1 y_1 \\ \lambda_2 y_2 \\ \lambda_3 y_3 \\ \vdots \\ \lambda_N y_N \end{pmatrix} = \sum_{i=1}^{N} \lambda_i \cdot y_i^2 = \sum_{i=1}^{N} \lambda_i \cdot [(O \cdot x)_i]^2 \\
&= \begin{cases} > 0 & \text{falls } \lambda_i > 0, \; i = 1, \ldots, N \\ < 0 & \text{falls } \lambda_i < 0, \; i = 1, \ldots, N \end{cases}
\end{aligned}$$

Beispiel: Es sei

$$A = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \in \mathbb{R}^{N,N}.$$

Wegen $A = A^T$ sind alle Eigenwerte reell. Ferner ist A positiv definit, da alle Hauptunterdeterminanten (vgl. Aufgabe 5, Blatt 2) positiv sind. Also gilt $\lambda > 0$ für alle Eigenwerte λ von A.

Kapitel 6

Optimierung

6.1 Freie Extremwerte

6.1.1 Hinreichende Kriterien

Definition:

$f: D \to \mathbb{R},\ D \subset \mathbb{R}^N$ hat an der Stelle $y \in D$ ein relatives Maximum (relatives Minimum) falls

$$f(y) \geq f(x) \text{ für alle } x \in U_\varepsilon(y),\ x \neq y$$

$$(f(y) \leq f(x) \text{ für alle } x \in U_\varepsilon(y),\ x \neq y)$$

Gelten die Ungleichungen für alle $x \in D$, spricht man von absoluten oder globalen Extremwerten. Das relative Maximum (Minimum) heißt strikt, falls $f(y) >\ (<)\ f(x),\ x \in U_\varepsilon(y)$

Bemerkung: $U_\varepsilon(y) = \{x \in D \mid ||x-y||_2 < \varepsilon\} = \left\{ x \in D \;\middle|\; \left(\sum_{i=1}^{N}(x_i - y_i)^2\right)^{\frac{1}{2}} < \varepsilon \right\}$

Charakterisierung lokaler Extrema:

Sei $f: D \to \mathbb{R}$ zweimal differenzierbar, $D \subset \mathbb{R}^N$. An einem lokalen Extremum muss eine horizontale Tangentialebene vorliegen. Dies verlangt

$$\nabla f(y) = \left(\frac{\partial f(y)}{\partial x_1}, \frac{\partial f(y)}{\partial x_2}, \ldots, \frac{\partial f(y)}{\partial x_N} \right) = 0.$$

Weitere Bedingungen an die 2-te Ableitung, d.h. die Hesse-Matrix

$$\nabla^2 f(y) = \begin{pmatrix} \frac{\partial^2 f(y)}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f(y)}{\partial x_1 \partial x_N} \\ \vdots & & \vdots \\ \frac{\partial^2 f(y)}{\partial x_N \partial x_1} & \cdots & \frac{\partial^2 f(y)}{\partial x_N \partial x_N} \end{pmatrix} \in \mathbb{R}^{N,N} \text{ symmetrisch}$$

sind notwendig.

Satz:

Sei $f : D \to \mathbb{R}$ zweimal total differenzierbar, und sei $y \in D$ mit $\nabla f(y) = 0$.
Ist überdies die quadratische Form $Q(w) = w^T \nabla^2 f(y) w$ bzw. die Matrix $\nabla^2 f(y)$ positiv (negativ) definit, so liegt in $y \in D$ ein relatives Minimum (Maximum) vor.

Ist die Hesse-Matrix bzw. quadratische Form $\nabla^2 f(y)$ indefinit, so liegt ein Sattelpunkt vor.

Aus der linearen Algebra ist bekannt: Äquivalent sind

i) $\nabla^2 f(y)$ bzw. $Q(w) = w^T \nabla^2 f(y) w$ ist pos. (neg.) definit, d.h. $w^T \nabla^2 f(y) w > \ (<)\ 0$ für $w \in \mathbb{R}^N$, $w \neq 0$

ii) Alle Eigenwerte λ von $\nabla^2 f(y)$ erfüllen $\lambda > \ (<) \, 0$

iii) Die Determinanten der $(i \times i)$-Hauptuntermatrizen h_i der Hessematrix $\nabla^2 f(y)$ erfüllen $\det(h_i) > 0 \ \left((-1)^i ; \ \det(h_i) > 0\right), \ i = 1, \dots, N$

Zur Berechnung eines lokalen Extremwertes von f ist also die Gleichung $\nabla f(x) = 0$ zu lösen und für eine Lösung y ist eine der drei Bedingungen i)-iii) für die Matrix $\nabla^2 f(y)$ nachzuweisen.

Beispiel: $N = 3, \quad f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 + 2x_1x_3, \quad x = (x_1, x_2, x_3) \in \mathbb{R}^3$

$$\begin{aligned} \frac{\partial f(x_1,x_2,x_3)}{\partial x_1} &= 2x_1 + 2x_2 + 2x_3 \stackrel{!}{=} 0 \\ \frac{\partial f(x_1,x_2,x_3)}{\partial x_2} &= 4x_2 + 2x_1 \stackrel{!}{=} 0 \\ \frac{\partial f(x_1,x_2,x_3)}{\partial x_3} &= 6x_3 + 2x_1 \stackrel{!}{=} 0 \end{aligned} \quad \Rightarrow \nabla f(x_1, x_2, x_3) = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 4 & 0 \\ 2 & 0 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Ausrechnen liefert

$$(A|0) = \left(\begin{array}{ccc|c} 2 & 2 & 2 & 0 \\ 2 & 4 & 0 & 0 \\ 2 & 0 & 6 & 0 \end{array}\right) \begin{array}{c} \\ \lambda = -1 \\ \lambda = -1 \end{array} \sim \left(\begin{array}{ccc|c} 2 & 2 & 2 & 0 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & 4 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 2 & 2 & 2 & 0 \\ 0 & 2 & -2 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right) = (\hat{A}|0)$$

$$\begin{aligned} 2x_3 = 0 &\Rightarrow x_3 = 0 \\ 2x_2 = 0 &\Rightarrow x_2 = 0 \\ 2x_1 = 0 &\Rightarrow x_1 = 0 \end{aligned} \Rightarrow \quad y = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \nabla f(y) = 0$$

Für die Hesse-Matrix erhält man

$$\nabla^2 f(x_1, x_2, x_3) = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 4 & 0 \\ 2 & 0 & 6 \end{pmatrix} = \nabla^2 f(y) = \nabla^2 f(0).$$

Prüfung auf Definitheit: Die Determinanten der Hauptuntermatrizen h_i, $i = 1, 2, 3$ von A und die von $\hat{A}$ stimmen überein! Da nur elementare Zeilenumformungen durchgeführt wurden folgt

$\det(h_1) = 2 > 0, \quad \det(h_2) = 2 \cdot 2 = 4 > 0, \quad \det(h_3) = 2 \cdot 2 \cdot 2 = 8 > 0$
Also ist die Hessematrix $\nabla^2 f(0)$ positiv definit und $y = 0$ ist lokales Minimum für f.

6.1.2 Konvexe und konkave Funktionen

Definition:

a) Eine Menge $D \subset \mathbb{R}^N$ heißt konvex, falls für je zwei Punkte $x, y \in D$ und die Strecke $\lambda x + (1-\lambda)y,\ 0 \leq \lambda \leq 1$ in D ist.

b) $f : D \to \mathbb{R}$ heißt konvex (konkav), falls $D \subset \mathbb{R}^N$ eine konvexe Menge ist und $\lambda f(x) + (1-\lambda)f(y) \geq\ (\leq)\ f(\lambda x + (1-\lambda)y),\ \ 0 \leq \lambda \leq 1,\ x, y \in D$

Gilt die entsprechende Ungleichung mit $<\ \ (>)$ für $0 < \lambda < 1$, so heißt f strikt konkav (strikt konvex).

Beachte: Im Fall $N = 1$ gilt, dass die Mengen $[a, b],\ \mathbb{R}, \mathbb{R}^+$ konvex sind.

Satz: Sei $f : D \to \mathbb{R},\ D \subset \mathbb{R}^N$ konvex zweimal differenzierbar. Dann gilt:

f ist konvex auf D $\Leftrightarrow$ $\nabla^2 f(x)$ ist positiv semidefinit für $x \in D$

f ist konkav auf D $\Leftrightarrow$ $\nabla^2 f(x)$ ist negativ semidefinit für $x \in D$

f ist stikt konvex auf D $\Leftarrow$ $\nabla^2 f(x)$ ist positiv definit für $x \in D$

f ist strikt konkav auf D $\Leftarrow$ $\nabla^2 f(x)$ ist negativ definit für $x \in D$

Satz:

Sei $f : D \to \mathbb{R}$ zweimal differenzierbar und konvex (konkav). Es sei $y \in D$ mit $\nabla f(y) = 0$.
Dann ist $y \in D$ globales Minimum (Maximum) von f in D.

Bemerkung: i) Bei konvexen (konkaven) Funktionen entfällt die Überprüfung auf lokale Minima / Maxima.

ii) Seien $f, g : D \to \mathbb{R},\ D \subset \mathbb{R}^N$ konvex. Sind f,g konvex (konkav) und $a, b \geq 0 \quad \Rightarrow \quad af + bg$ konvex (konkav).

Beispiel: $f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 + 2x_1x_3, \quad \nabla^2 f(x_1, x_2, x_3) = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 4 & 0 \\ 2 & 0 & 6 \end{pmatrix}$

Schon gezeigt: $\nabla^2 f(x_1, x_2, x_3)$ ist positiv definit für alle $x \in \mathbb{R}^3$ und aus vorigem Satz gilt: $y = (0, 0, 0)$ ist globales Minimum für f in $\mathbb{R}^3$.

Abschwächung des Konvexitätsbegriffs

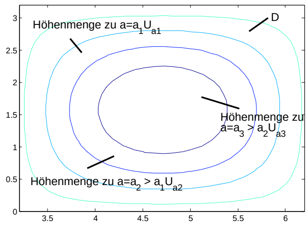
Situation: $f : D \to \mathbb{R},\ D \subset \mathbb{R}^N$ konvex. Zu $\alpha \in \mathbb{R}$ setze $U_\alpha = \{x \in D \mid f(x) \geq \alpha\},\ U_\alpha \subset D$

"obere Höhenmenge von f zum Level α"

Definition:

$f : D \to \mathbb{R},\ D \subset \mathbb{R}^N$ heißt quasikonkav, falls die obere Höhenmenge U_α von f konvex ist für jedes $\alpha \in \mathbb{R}$.

Visualisierung $(N = 2)$



Bemerkung: f heißt quasikonvex, falls

$L_\alpha = \{x \in D \mid f(x) \leq \alpha\},\ L_\alpha \subset D$

"untere Höhenmenge von f zum Level α" für jedes $\alpha \in \mathbb{R}$ eine konvexe Menge ist.

Lemma: Ist $f : D \to \mathbb{R}$ konvex (konkav), so auch quasikonvex (quasikonkav).

Genauer gilt die folgende Charakterisierung:

f ist quasikonkav genau dann, wenn $f(\lambda x + (1 - \lambda)y) \geq min\{f(x), f(y)\}$ für $x, y \in D,\ 0 \leq \lambda \leq 1$.

In der folgenden Situation kann man von quasikonkav auf konkav zurückschließen:

Definition:
i) $K \subset \mathbb{R}^N$ heißt Kegel, falls für $x \in K$ und $t \geq 0$ gilt: $tx \in K$.
Man beachte $0 \in K$, da $0 \cdot x = 0$
ii) Sei $f : K \to \mathbb{R},\ K \subset \mathbb{R}^N$ Kegel. f heißt homogen vom Grade q, falls gilt:

$$f(tx) = t^q f(x) \ \text{ für } \ x \in K,\ t > 0.$$

Satz:

Sei $K \subset \mathbb{R}^N$ Kegel, sei $f : K \to \mathbb{R}$ homogen vom Grade q mit $0 < q \leq 1$. Ist f quasikonkav, so ist f konkav.

Beispiel: Cobb-Douglas-Funktion $f(x_1, x_2) = Ax_1^{\alpha_1}x_2^{\alpha_2}$

$f : D \to \mathbb{R},\ D = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geq 0,\ x_2 \geq 0\}$, D ist ein Kegel.

Homogenität: $f(tx_1, tx_2) = A(tx_1)^{\alpha_1}(tx_2)^{\alpha_2} = At^{\alpha_1}x_1^{\alpha_1}t^{\alpha_2}x_2^{\alpha_2} = t^{\alpha_1+\alpha_2}\underbrace{Ax_1^{\alpha_1}x_2^{\alpha_2}}_{=f(x_1,x_2)} = t^{\alpha_1+\alpha_2} \cdot f(x_1, x_2)$

$\Rightarrow$ f ist homogen vom Grade $q = \alpha_1 + \alpha_2$

Somit ist die Cobb-Douglas Funktion konkav, falls $\alpha_1 + \alpha_2 \leq 1$.

6.2 Funktionalmatrix, Kettenregel und implizite Funktionen

6.2.1 Funktionalmatrix

Bisher: $f : D \to \mathbb{R},\ D \subset \mathbb{R}^N$ (Ableiten einer Variablen)

Verallgemeinere den Wertebereich. Betrachte nun $f : D \to \mathbb{R}^l,\ D \subset \mathbb{R}^N$

$$\text{Finde } f(x) = \overbrace{f(\underbrace{x_1, \dots, x_N}_{\in \mathbb{R}^N})}^{\in \mathbb{R}^l} = \begin{pmatrix} f_1(x_1, \dots, x_N) \\ f_2(x_1, \dots, x_N) \\ \vdots \\ f_l(x_1, \dots, x_N) \end{pmatrix} \in \mathbb{R}^l$$

f heißt vektorwertige Funktion in N Variablen.

Beispiel: $N = 2,\ l = 3$

$$f(x) = f(x_1, x_2) = \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \\ f_3(x_1, x_2) \end{pmatrix} = \begin{pmatrix} x_1^2 - 2x_2 \\ \exp(x_1) - x_2^2 \\ \sin(x_1) + x_1 x_2 \end{pmatrix} \qquad D = \mathbb{R}^2,\ f : \mathbb{R}^2 \to \mathbb{R}^3$$

Vektorwertige Funktionen sind eine Kollektion von l reellwertigen Funktionen $f_1, f_2, \dots, f_l$, welche als Vektor $f = \begin{pmatrix} f_1 \\ \vdots \\ f_l \end{pmatrix}$ geschrieben werden. Der Definitionsbereich von f ist der gemeinsame Definitionsbereich von $f_1, \dots, f_l$.

$f = (f_1, \dots, f_l)$ heißt im Punkt $y = (y_1, \dots, y_N) \in D$ partiell differenzierbar nach x_j, falls die l reellwertigen Funktionen einer Variablen $g_i(x_j) = f_i(y_1, \dots, y_{j-1}, x_j, y_{j+1}, \dots, y_N)$ im Punkt y_j differenzierbar ist, $i = 1, \dots, l$, d.h.

$$\frac{g_i(y_j + \tau) - g_i(y_j)}{\tau} = \frac{f_i(y_1, \dots, y_{j-1}, y_j + \tau, y_{j+1}, \dots, y_N) - f_i(y_1, \dots, y_N)}{\tau} \longrightarrow a_{ij} \quad \text{für} \quad \tau \to 0$$

Man schreibt $a_{ij} = \frac{\partial f_i}{\partial x_j}(y_1, \dots, y_N), \quad i = 1, \dots, l, \quad j = 1, \dots, N$

Ist f nach jeder Variablen $x_j,\ j = 1, \dots, N$ partiell differenzierbar, so entstehen N partielle Ableitungen

$$\frac{\partial f}{\partial x_j}(y_1, \dots, y_N) = \begin{pmatrix} \frac{\partial f_1}{\partial x_j}(y_1, \dots, y_N) \\ \frac{\partial f_2}{\partial x_j}(y_1, \dots, y_N) \\ \vdots \\ \frac{\partial f_l}{\partial x_j}(y_1, \dots, y_N) \end{pmatrix}, \quad j = 1, \dots, N$$

Die Matrix, welche $\frac{\partial f}{\partial x_j}(y_1, \dots, y_N)$ als j-te Spalte besitzt, heißt Funktionalmatrix von f.

$$\begin{aligned} Df(y_1, \dots, y_N) &= \left(\frac{\partial f}{\partial x_1}(y_1, \dots, y_N), \frac{\partial f}{\partial x_2}(y_1, \dots, y_N), \dots, \frac{\partial f}{\partial x_N}(y_1, \dots, y_N) \right) \\ &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(y_1, \dots, y_N) & \frac{\partial f_1}{\partial x_2}(y_1, \dots, y_N) & \dots \frac{\partial f_1}{\partial x_N}(y_1, \dots, y_N) \\ \frac{\partial f_2}{\partial x_1}(y_1, \dots, y_N) & \frac{\partial f_2}{\partial x_2}(y_1, \dots, y_N) & \dots \frac{\partial f_2}{\partial x_N}(y_1, \dots, y_N) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_l}{\partial x_1}(y_1, \dots, y_N) & \frac{\partial f_l}{\partial x_2}(y_1, \dots, y_N) & \dots \frac{\partial f_l}{\partial x_N}(y_1, \dots, y_N) \end{pmatrix} \end{aligned}$$

Funktionalmatrix oder Jacobi-Matrix von f an der Stelle $y \in D$

$f : D \to \mathbb{R}^l,\ D \subset \mathbb{R}^N$ heißt total differenzierbar in D, falls für jedes $y \in D,\ j \in \{1, \ldots, N\},\ i \in \{1, \ldots, l\}$ die Funktionen f_i partiell nach x_j differenzierbar sind und die partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}(x_1, \ldots, x_N),\ i = 1, \ldots, l,\ j = 1, \ldots, N$ stetig in D sind.
Ist $f : D \to \mathbb{R}^l,\ D \subset \mathbb{R}^N$ differenzierbar, so ist $Df(x) \in \mathbb{R}^{l,N},\ x \in D$.

Betrachte $f : \mathbb{R}^2 \to \mathbb{R}^3,\ f(x_1, x_2) = \begin{pmatrix} x_1^2 - 2x_2 \\ \exp(x_1) - x_2^2 \\ \sin(x_1) + x_1 x_2 \end{pmatrix} \in \mathbb{R}^3$

f ist differenzierbar und

$$f'(x_1, x_2) = Df(x_1, x_2) = \begin{pmatrix} 2x_1 & -2 \\ \exp(x_1) & -2x_2 \\ \cos(x_1) + x_2 & x_1 \end{pmatrix} \in \mathbb{R}^{3,2}.$$

<u>Beispiel:</u> $f(x_1, x_2, x_3) = \begin{pmatrix} 2x_1 - 5x_2 + 7x_3 \\ -x_1 + 2x_2 - 4x_3 \end{pmatrix} = \begin{pmatrix} 2 & -5 & 7 \\ -1 & 2 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = Ax$

$$Df(x_1, x_2, x_3) = \begin{pmatrix} 2 & -5 & 7 \\ -1 & 2 & -4 \end{pmatrix} \in \mathbb{R}^{2,3}$$

$f : \mathbb{R}^3 \to \mathbb{R}^2$ ist differenzierbar und $f'(x_1, x_2, x_3) = \begin{pmatrix} 2 & -5 & 7 \\ -1 & 2 & -4 \end{pmatrix} = A$

Dies gilt auch allgemein.
Sei $f : \mathbb{R}^N \to \mathbb{R}^M,\ f(x) = Ax$ mit $A = (a_{ij}) \in \mathbb{R}^{M,N}$. Dann gilt: f ist an jeder Stelle $x \in \mathbb{R}^N$ differenzierbar und $f'(x) = Df(x) = A$ für $x \in \mathbb{R}^N$.

Matrizen stehen also für lineare Abbildungen, d.h. durch $A \in \mathbb{R}^{M,N}$ und die lineare Abbildung $f(x) = Ax,\ f : \mathbb{R}^N \to \mathbb{R}^M$ beschrieben.
f ist differenzierbar und $f'(x) = A$ für $x \in \mathbb{R}^N$.

6.2.2 Kettenregel

Sei nun $f : D \to E$ und $g : F \to H$ mit $H, D \subset \mathbb{R}^l,\ H \subset D,\ F \subset \mathbb{R}^N,\ E \subset \mathbb{R}^M$. Wegen $g(x) \in H \subset D$ ist dann $f \circ g : F \to H \subset D \to E\ \ (F \to E)$ bildbar und wir finden

$$(f \circ g)(x) = f(g(x)) = f(g_1(x_1, \ldots, x_N), g_2(x_1, \ldots, x_N), \ldots, g_l(x_1, \ldots, x_N)) = \begin{pmatrix} f_1(g_1(x_1, \ldots, x_N), \ldots, g_l(x_1, \ldots, x_N)) \\ f_2(g_1(x_1, \ldots, x_N), \ldots, g_l(x_1, \ldots, x_N)) \\ \vdots \\ f_M(g_1(x_1, \ldots, x_N), \ldots, g_l(x_1, \ldots, x_N)) \end{pmatrix}$$

Sind f,g differenzierbar, so ist auch $f \circ g : F \to E$ differenzierbar und es gilt:

$$D(f \circ g)(x) = Df(g(x)) \cdot Dg(x) \quad \text{bzw.} \quad (f \circ g)'(x) = f'(g(x)) \cdot g'(x)$$

Beachte: $Df(g(x))$ und $Dg(x)$ sind Matrizen, d.h.$''\cdot''$ bedeuted Matrizenmultiplikation.

Beispiel: $f(y_1, y_2) = 2y_1 + \ln(y_2^2), \quad g(x_1, x_2, x_3) = \begin{pmatrix} g_1(x_1, x_2, x_3) \\ g_2(x_1, x_2, x_3) \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(x_2^2 + x_3^2) - x_1 \\ \exp(x_1 + x_2 x_3) \end{pmatrix}$

$(l = 2,\ N = 3,\ M = 1)$

$$\begin{aligned} (f \circ g)(x_1, x_2, x_3) \quad &= f(\underbrace{g_1(x_1, x_2, x_3)}_{=y_1}, \underbrace{g_2(x_1, x_2, x_3)}_{=y_2}) = 2g_1(x_1, x_2, x_3) + \ln(g_2(x_1, x_2, x_3)^2) \\ &= (x_2^2 + x_3^2) - 2x_1 + \ln(\exp(x_1 + x_2 x_3)^2) \\ &= (x_2^2 + x_3^2) - 2x_1 + 2\ln(\exp(x_1 + x_2 x_3)) \\ &= (x_2^2 + x_3^2) - 2x_1 + 2(x_1 + x_2 x_3) \\ &= x_2^2 + x_3^2 + 2x_2 x_3 \\ &= (x_2 + x_3)^2 \end{aligned}$$

Finde durch direktes Ableiten: $D(f \circ g)(x) = \left(\frac{\partial (f \circ g)(x)}{\partial x_1}, \frac{\partial (f \circ g)(x)}{\partial x_2}, \frac{\partial (f \circ g)(x)}{\partial x_3} \right) = \big(0, 2(x_2 + x_3), 2(x_2 + x_3)\big) \in \mathbb{R}^{1,3}$

Berechne nun $D(f \circ g)(x)$ mit Hilfe der Kettenregel:

$$\begin{aligned} &Dg(x_1, x_2, x_3) = \begin{pmatrix} -1 & x_2 & x_3 \\ \exp(x_1 + x_2 x_3) & \exp(x_1 + x_2 x_3) x_3 & \exp(x_1 + x_2 x_3) x_2 \end{pmatrix} \quad \in \mathbb{R}^{2,3} \\ &Df(y_1, y_2) = \left(2,\ \frac{2}{y_2}\right) \\ &\text{Mit } y = (y_1, y_2) = g(x) = (g_1(x_1, x_2, x_3), g_2(x_1, x_2, x_3)) \text{ folgt} \\ &Df(g(x_1, x_2, x_3)) = \left(2,\ \frac{2}{g_2(x_1, x_2, x_3)}\right) = \left(2,\ \frac{2}{\exp(x_1 + x_2 x_3)}\right) \qquad \in \mathbb{R}^{1,2} \end{aligned}$$

Bilde nun $Df(g(x_1, x_2, x_3)) \cdot D(g(x_1, x_2, x_3))$:

$$\begin{aligned} Df(g(x))Dg(x) \quad &= \left(-2 + \frac{2}{\exp(x_1 + x_2 x_3)} \cdot \exp(x_1 + x_2 x_3),\ 2x_2 + \frac{2x_3 \cdot \exp(x_1 + x_2 x_3)}{\exp(x_1 + x_2 x_3)},\ 2x_3 + \frac{2x_2 \cdot \exp(x_1 + x_2 x_3)}{\exp(x_1 + x_2 x_3)} \right) \\ &= (-2 + 2,\ 2x_2 + 2x_3,\ 2x_3 + 2x_2) \\ &= (0,\ 2(x_2 + x_3),\ 2(x_2 + x_3)) \ \in \mathbb{R}^{1,3} \end{aligned}$$

6.2.3 Implizite Funktionen

Beispiel: Sei $f : \mathbb{R}^3 \to \mathbb{R}^2$ definiert durch $f(x_1, x_2, x_3) = \begin{pmatrix} x_1^2 + x_2^2 + x_3^2 - 1 \\ x_1 - 2x_2 \end{pmatrix}$

Betrachte wieder die Gleichung $f(x_1, x_2, x_3) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \mathbb{R}^2$

Hieraus wird implizit eine neue Funktion definiert:

$$\begin{aligned} &x_1^2 + x_2^2 + x_3^2 - 1 \overset{!}{=} 0 && x_1 - 2x_2 \overset{!}{=} 0 \\ &x_3^2 = 1 - x_1^2 - x_2^2 && x_1 = 2x_2 \\ &x_3 = \pm\sqrt{1 - x_1^2 - x_2^2} && x_2 = \tfrac{1}{2}x_1 = \varphi_1(x_1) \\ &\Rightarrow \quad x_3 = \pm\sqrt{1 - x_1^2 - \tfrac{1}{4}x_1^2} = \pm\sqrt{1 - \tfrac{5}{4}x_1^2} = \pm\varphi_2(x_1) \end{aligned}$$

Auflösende Funktion $\varphi \pm (x_1) = (\varphi_1(x_1), \pm\varphi_2(x_1))$
Also finden wir $f(x_1, \varphi_1(x_1), \pm\varphi_2(x_1)) = 0$ nach Konstruktion von φ_1 und φ_2

Allgemein gilt für die Auflösbarkeit von nichtlinearen Gleichungen folgender Satz:

Satz (Implizite Funktionen):

Sei $F : D \to \mathbb{R}^l,\ D \subset \mathbb{R}^{N+l}$ differenzierbar. Es sei $(x^0, y^0) = (x_1^0, \dots, x_N^0, y_1^0, \dots, y_l^0)$ mit $F(x^0, y^0) = F(x_1^0, \dots, x_N^0, y_1^0, \dots, y_l^0) = 0$. Ist dann $\frac{\partial F(x^0,y^0)}{\partial(y_1,\dots,y_l)}$ oder kurz geschrieben $\frac{\partial F(x^0,y^0)}{\partial y} \in \mathbb{R}^{l,l}$ invertierbar, so lässt sich die Gleichung $F(x_1, \dots, x_N, y_1, \dots, y_l) = 0$ lokal bei $(x_1^0, \dots, x_N^0, y_1^0, \dots, y_l^0)$ auflösen, d.h. es gibt eine differenzierbare Funktion $\varphi : U_\varepsilon(x^0) \to \mathbb{R}^l$ mit $\varphi(x^0) = y^0$ und $F(x_1, \dots, x_N, \varphi(x_1, \dots, x_N)) = 0$ für $x = (x_1, \dots, x_N) \in U_\varepsilon(x^0)$.

Konsequenz: Man bekommt an einer Nullstelle (x^0, y^0) lokal eine ganze Fläche $(x, \varphi(x))$ von Nullstellen.

Betrachte nochmal die Gleichung $F(x_1, \dots, x_n, \varphi(x_1, \dots, x_N)) = 0 \quad x = (x_1, \dots, x_N) \in U_\varepsilon(x^0) \quad \Leftrightarrow$ $F(x, \varphi(x)) = 0,\ x \in U_\varepsilon(x^0)$

Nach der Kettenregel gilt: $\frac{\partial F(x,\varphi(x))}{\partial x} + \frac{\partial F(x,\varphi(x))}{\partial y} \cdot \varphi'(x) = 0 \quad \Leftrightarrow \quad \frac{\partial F(x,\varphi(x))}{\partial y} \cdot \varphi'(x) = -\frac{\partial F(x,\varphi(x))}{\partial x}$

Ist $\frac{\partial F(x,\varphi(x))}{\partial y}$ invertierbar, so folgt

$\varphi'(x) = -\left[\frac{\partial F(x,\varphi(x))}{\partial y}\right]^{-1} \cdot \frac{\partial F(x,\varphi(x))}{\partial x}, \quad x \in U_\varepsilon(x_0).$

Betrachte nochmals unser Beispiel:

Mit $x_2 = y_1,\ x_3 = y_2$ finde $f(x_1, y_1, y_2) = \begin{pmatrix} x_1^2 + y_1^2 + y_2^2 - 1 \\ x_1 - 2y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad l = 2,\ N = 1,\ N + l = 3$

Setze $(x_1^0, y_1^0, y_2^0) = \left(\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right)$ und finde

$$f\left(\tfrac{2}{3}, \tfrac{1}{3}, \tfrac{2}{3}\right) = \begin{pmatrix} \left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{2}{3}\right)^2 - 1 \\ \frac{2}{3} - 2 \cdot \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{4+1+4}{9} - 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\frac{\partial f(x_1, y_1, y_2)}{\partial(y_1, y_2)} = \begin{pmatrix} \frac{\partial f_1(x_1,y_1,y_2)}{\partial y_1} & \frac{\partial f_1(x_1,y_1,y_2)}{\partial y_2} \\ \frac{\partial f_2(x_1,y_1,y_2)}{\partial y_1} & \frac{\partial f_2(x_1,y_1,y_2)}{\partial y_2} \end{pmatrix} = \begin{pmatrix} 2y_1 & 2y_2 \\ -2 & 0 \end{pmatrix}$$

$$\frac{\partial f\left(\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right)}{\partial(y_1, y_2)} = \begin{pmatrix} \frac{2}{3} & \frac{4}{3} \\ -2 & 0 \end{pmatrix} \in \mathbb{R}^{2,2}$$

Wegen $\det\left(\frac{\partial f\left(\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right)}{\partial(y_1, y_2)}\right) = \det\begin{pmatrix} \frac{2}{3} & \frac{4}{3} \\ -2 & 0 \end{pmatrix} = \frac{8}{3} \neq 0$ ist $\frac{\partial f\left(\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right)}{\partial(y_1, y_2)}$ invertierbar, d.h. es existiert eine auflösende Funktion $\varphi :]x_1^0 - \varepsilon, x_1^0 + \varepsilon[\to \mathbb{R}^2$ mit $\varphi(\frac{2}{3}) = (\frac{1}{3}, \frac{2}{3})$.

6.3 Extremwerte mit Gleichheitsnebenbedingungen

6.3.1 Problemstellung

Seien $f : D \to \mathbb{R},\ g_i : D \to \mathbb{R},\ i = 1, \dots, l,\ D \subset \mathbb{R}^N,\ l < N$ differenzierbare reellwertige Funktionen.
Problem: Maximiere / minimiere $f(x) = f(x_1, \dots, x_N)$ in D unter der Nebenbedingung $g_1(x) = 0,\ g_2(x) = 0, \dots, g_N(x) = 0$

$Z = \{x \in D \mid g_1(x) = g_2(x) = \dots = g_l(x) = 0\}$ "Menge der zulässigen Punkte"

Gesucht ist ein lokales / globales Maximum / Minimum von der Funktion f auf Z.

Sei $g(x) = g(x_1, \dots, x_N) = \begin{pmatrix} g_1(x_1, \dots, x_N) \\ \vdots \\ g_l(x_1, \dots, x_N) \end{pmatrix}$, $g : D \to \mathbb{R}^l$ differenzierbar mit Funktionalmatrix

$$Dg(x) = \begin{pmatrix} \frac{\partial g_1(x)}{\partial x_1} & \frac{\partial g_1(x)}{\partial x_2} & \dots & \frac{\partial g_1(x)}{\partial x_N} \\ \frac{\partial g_2(x)}{\partial x_1} & \frac{\partial g_2(x)}{\partial x_2} & \dots & \frac{\partial g_2(x)}{\partial x_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial g_l(x)}{\partial x_1} & \frac{\partial g_l(x)}{\partial x_2} & \dots & \frac{\partial g_l(x)}{\partial x_N} \end{pmatrix} \in \mathbb{R}^{l,N}$$

Mögliche Ränge für $Dg(x) : 0, 1, 2, \dots, l$

Bemerkung: x heißt regulär für g, falls $rg(Dg(x)) = l$ d.h. die Vektoren $\nabla g_1(x), \nabla g_2(x), \dots, g_l(x)$ sind linear unabhängig.

Satz:

Gilt $rg(Dg(x)) = l$ [maximal] für alle $x \in Z$, so ist Z eine (N-l)-dimensionale "Fläche" im $\mathbb{R}^N$.

6.3.2 Lösung des Problems durch Variablenelimination

Betrachte die Nebenbedingung $g(x_1, \dots, x_N) = 0 \in \mathbb{R}^l$

Setze $y_1 = x_{N-l+1},\ \ y_2 = x_{N-l+2}, \dots, y_l = x_N \quad \Rightarrow \quad g(x_1, \dots, x_N) = g(x_1, x_2, \dots, x_{N-l}, y_1, \dots, y_l)$

Löse diese Gleichung mit dem Satz über implizite Funktionen auf. Sei $(x_1^0, \dots, x_{N-l}^0, y_1^0, \dots, y_l^0) \in Z$, d.h. $g(x_1^0, \dots, x_{N-l}^0, y_1^0, \dots, y_l^0) = 0$

Gilt $\frac{\partial g}{\partial y}(x_1^0, \dots, x_{N-l}^0, y_1^0, \dots, y_l^0) \in \mathbb{R}^{l,l}$ invertierbar, so ist die Gleichung
$g(x_1, \dots, x_{N-l}, y_1, \dots, y_l) = 0$ lokal an der Stelle $(x_1^0, \dots, x_{N-l}^0, y_1^0, \dots, y_l^0)$ nach $(y_1, \dots, y_l)$ auflösbar.
Dann finden wir eine Funktion $\varphi = \varphi(x_1, \dots, x_{N-l})$ mit $g(x_1, \dots, x_{N-l}, \varphi(x_1, \dots, x_{N-l})) = 0$.
Man untersucht dann $h(x_1, \dots, x_{N-l}) = f(x_1, \dots, x_{N-l}, \varphi(x_1, \dots, x_{N-l}))$ auf freie Extremwerte, d.h. man sucht $(x_1^*, \dots, x_{N-l}^*)$ mit $\nabla h(x_1^*, \dots, x_{N-l}^*) = 0$, $\nabla^2 h(x_1^*, \dots, x_{N-l}^*)$ positiv / negativ definit und $(x_1^*, \dots, x_{N-l}^*, \varphi(x_1^*, \dots, x_{N-l}^*)) \in Z$. Dann ist $(x_1^*, \dots, x_{N-l}^*, \varphi(x_1^*, \dots, x_{N-l}^*))$ eine Lösung des Ausgangsproblems.

6.3.3 Die Methode von Lagrange

Problem: Maximiere / minimiere $f(x),\ x \in D$ unter der NB: $g_1(x) = \ldots = g_l(x) = 0 \quad (*)$

Dann heißt

$$L(\lambda, x) = L(\lambda_1, \ldots, \lambda_l, x_1, \ldots, x_N) = f(x_1, \ldots, x_N) + \sum_{i=1}^{l} \lambda_i \cdot g_i(x_1, \ldots, x_N)$$

die Lagrange-Funktion des Problems und die Zahlen $\lambda_i,\ i = 1, \ldots, l$ heißen Lagrangemultiplikatoren

Satz:

Vorgelegt sei das Problem $(*)$, und es sei $L(\lambda, x) = f(x) + \sum\limits_{i=1}^{l} \lambda_i g_i(x)$ die zugehörige Lagrange-Funktion. Dann lauten die notwendigen Bedingungen erster Ordnung für ein lokales Extremum unter Nebenbedingungen

$$\frac{\partial L}{\partial \lambda_i}(\lambda, x) = g_i(x) = 0, \quad i = 1, \ldots, l$$
$$\frac{\partial L}{\partial x_j}(\lambda, x) = \frac{\partial f}{\partial x_j}(x) + \sum_{i=1}^{l} \lambda_i \cdot \frac{\partial g_i}{\partial x_j}(x) = 0, \quad j = 1, \ldots, N$$

"$(N + l)$-Gleichungen für $(N + l)$-Unbekannte"

Lösungen (x^*, λ^*) des obigen Systems nennt man auch stationäre Punkte der Lagrange-Funktion.

Beispiel: $f(x_1, x_2, x_3) = x_1 + 2x_3 \quad (N = 3,\ l = 2)$

NB: $g_1(x_1, x_2, x_3) = x_1 + x_2 +_3 -1 = 0, \quad g_2(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3 - \frac{7}{4} = 0$

Lagrangefunktion: $L(\lambda_1, \lambda_2, x_1, x_2, x_3) = x_1 + 2x_3 + \lambda_1(x_1 + x_2 + x_3 - 1) + \lambda_2(x_1^2 + x_2^2 + x_3 - \frac{7}{4})$

Notwendige Bedingungen 1-ter Ordnung für lokale Extrema:

$$\begin{array}{lll}
a) & \frac{\partial L}{\partial \lambda_1}(\lambda, x) = & x_1 + x_2 + x_3 - 1 \overset{!}{=} 0 \\
b) & \frac{\partial L}{\partial \lambda_2}(\lambda, x) = & x_1^2 + x_2^2 + x_3 - \frac{7}{4} \overset{!}{=} 0 \\
c) & \frac{\partial L}{\partial x_1}(\lambda, x) = & 1 + \lambda_1 + 2\lambda_2 x_1 \overset{!}{=} 0 \\
d) & \frac{\partial L}{\partial x_2}(\lambda, x) = & \lambda_1 + 2\lambda_2 x_2 \overset{!}{=} 0 \\
e) & \frac{\partial L}{\partial x_3}(\lambda, x) = & 2 + \lambda_1 + \lambda_2 \overset{!}{=} 0
\end{array}$$

$$\begin{array}{ll}
(e) \Rightarrow & \lambda_2 = -2 - \lambda_1 \\
(d) \Rightarrow & 0 = \lambda_1 + 2(-2 - \lambda_1)x_2 = \lambda_1 - 4x_2 - 2\lambda_1 x_2 \\
 & = \lambda_1(1 - 2x_2) - 4x_2 \qquad \Rightarrow \quad \lambda_1 = \dfrac{4x_2}{1 - 2x_2} \\
(c) \Rightarrow & 0 = 1 + \frac{4x_2}{1-2x_2} + 2x_1\left(-2 - \frac{4x_2}{1-2x_2}\right) \\
\Leftrightarrow & 0 = 1 - 2x_2 + 4x_2 - 4x_1(1 - 2x_2) - 8x_1 x_2 = 1 + 2x_2 - 4x_1 \\
 & \Rightarrow \quad x_2 = 2x_1 - \frac{1}{2}
\end{array}$$

Einsetzen in (a) liefert: $1 = x_1 + x_2 + x_3 = x_1 + 2x_1 - \frac{1}{2} + x_3 \quad \Leftrightarrow \quad x_3 = \frac{3}{2} - 3x_1$

Einsetzen in (b) liefert:
$$\begin{aligned}
& \tfrac{7}{4} = x_1^2 + x_2^2 + x_3 = x_1^2 + \left(2x_1 - \tfrac{1}{2}\right)^2 + \tfrac{3}{2} - 3x_1 \\
\Leftrightarrow\ & \tfrac{1}{4} = x_1^2 + 4x_1^2 - 2x_1 + \frac{1}{4} - 3x_1 \\
\Leftrightarrow\ & 0 = 5x_1^2 - 5x_1 = 5x_1(x_1 - 1)
\end{aligned}$$

Also folgt: $x_1 = 1$ oder $x_1 = 0$

$\underline{x_1 = 0 \text{ liefert:}}$ $x_3 = \dfrac{3}{2}$, $\quad x_2 = -\frac{1}{2}$, $\quad \lambda_1 = \dfrac{4(-\frac{1}{2})}{1-2(-\frac{1}{2})} = \frac{-2}{2} = -1$, $\quad \lambda_2 = -2-(-1) = -1$

$\underline{x_1 = 1 \text{ liefert:}}$ $x_3 = \frac{3}{2} - 3 = -\frac{3}{2}$, $\quad x_2 = 2 \cdot 1 - \frac{1}{2} = \frac{3}{2}$, $\quad \lambda_1 = -3$, $\quad \lambda_2 = 1$

Zwei stationäre Punkte der Lagrangefunktion: $x^* = \left(0, -\frac{1}{2}, \frac{3}{2}\right)$, $\quad \lambda^* = (-1,-1)$
$\hat{x} = \left(1, \frac{3}{2}, -\frac{3}{2}\right)$, $\quad \hat{\lambda} = (-3, 1)$

Weitere Untersuchungen sind notwendig um zu bestimmen, ob dies lokale Minima / Maxima unter Nebenbedingungen sind.

6.3.4 Bemerkungen zur Lösbarkeit der Gleichungen

Wir haben $\dfrac{\partial L}{\partial x_j}(\lambda, x) = \dfrac{\partial f}{\partial x_j}(x) + \sum\limits_{i=1}^{l} \lambda_i \cdot \dfrac{\partial g_i}{\partial x_j}(x) = 0, \quad j = 1, \ldots, N.$

Mit $\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_N} \end{pmatrix}$ und $\nabla g_i(x) = \begin{pmatrix} \frac{\partial g_i(x)}{\partial x_1} \\ \vdots \\ \frac{\partial g_i(x)}{\partial x_N} \end{pmatrix}$, $\quad i = 1, \ldots, l$ lautet obige Gleichung

$$\nabla f(x) + \sum_{i=1}^{l} \lambda_i \nabla g_i(x) = 0 \quad \Leftrightarrow \quad \nabla f(x) = -\sum_{i=1}^{l} \lambda_i \nabla g_i(x)$$

d.h. die $(l \times 1)$-Vektoren $\nabla f(x), \nabla g_1(x), \ldots, \nabla g_l(x)$ sind linear abhängig.

<u>Lemma:</u>

Ist nun $x^* \in Z$ eine lokale Lösung des Optimierungsproblems mit $rg(Dg(x^*)) = l$ (d.h. eine reguläre Lösung bzw. $\nabla g_1(x^*), \ldots, \nabla g_l(x^*)$ sind linear unabhängig).
Dann existieren die Lagrangemultiplikatoren $\lambda_1^*, \ldots, \lambda_l^*$ mit

$$\nabla f(x^*) + \sum_{i=1}^{l} \lambda_i^* \nabla g_i(x^*) = \frac{\partial L(x^*, \lambda^*)}{\partial x} = 0.$$

<u>Bemerkung:</u> An einer Lösung x^* mit $rg(Dg(x^*)) = l$ findet man also immer die zugehörigen Lagrangemultiplikatoren λ_i^*, $i = 1, \ldots, l$.

6.3.5 Hinreichende Bedingungen 2-ter Ordnung für lokale Extrema unter Nebenbedingungen

Seien nun x^*, λ^* Lösung der Gleichung

$$\frac{\partial L}{\partial \lambda}(\lambda, x) = g(x) = 0, \qquad \frac{\partial L}{\partial x}(\lambda, x) = \nabla f(x) + \sum_{i=1}^{l} \lambda_i \nabla g_i(x) = 0.$$

Mit $\lambda^* = (\lambda_1^*, \ldots, \lambda_l^*)$ setzen wir

$\hat{L}(x) = L(\lambda^*, x) = f(x) + \sum\limits_{i=1}^{l} \lambda_i^* g_i(x)$ und untersuchen die Funktion $\hat{L}$ in der Nähe von x^*.

$\Rightarrow \quad \nabla \hat{L}(x) = \dfrac{\partial L}{\partial x}(\lambda^*, x) = \nabla f(x) + \sum\limits_{i=1}^{l} \lambda_i^* \nabla g_i(x)$

Es gilt $\nabla \hat{L}(x^*) = \dfrac{\partial L}{\partial x}(\lambda^*, x^*) = \nabla f(x^*) + \sum\limits_{i=1}^{l} \lambda_i^* \nabla g_i(x^*) = 0,$ da x^* stationärer Punkt der Lagrangefunktion. Weiter folgt

$$\nabla^2 \hat{L}(x) = \nabla^2 f(x) + \sum_{i=1}^{l} \lambda_i^* \nabla^2 g_i(x), \qquad \nabla^2 \hat{L}(x^*) = \nabla^2 f(x^*) + \sum_{i=1}^{l} \lambda_i^* \nabla^2 g_i(x^*) \in \mathbb{R}^{N,N} \text{ symmetrisch.}$$

Hinreichendes Kriterium 2-ter Ordnung für lokale Minima / Maxima:

Die quadratische Form $Q(y) = y^T \nabla^2 \hat{L}(x^*) y$ ist positiv / negativ definit bezüglich der NB

$Dg(x^*)y = 0$.

Also wird die Definitheit entschieden an den Hauptunterdeterminanten von

$$C(x^*) = \begin{pmatrix} 0 & Dg(x^*) \\ Dg(x^*)^T & \nabla^2 \hat{L}(x^*) \end{pmatrix} = \left(\begin{array}{ccc|ccc} 0 & \dots & 0 & \frac{\partial g_1(x)}{\partial x_1} & \dots & \frac{\partial g_1(x)}{\partial x_N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & o & \frac{\partial g_l(x)}{\partial x_1} & \dots & \frac{\partial g_l(x)}{\partial x_N} \\ \hline \frac{\partial g_1(x)}{\partial x_1} & \dots & \frac{\partial g_l(x)}{\partial x_1} & \frac{\partial^2 \hat{L}(x)}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 \hat{L}(x)}{\partial x_1 \partial x_N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1(x)}{\partial x_N} & \dots & \frac{\partial g_l(x)}{\partial x_N} & \frac{\partial^2 \hat{L}(x)}{\partial x_N \partial x_1} & \dots & \frac{\partial^2 \hat{L}(x)}{\partial x_N \partial x_N} \end{array}\right) \in \mathbb{R}^{N+l,N+l}$$

Satz:

Sei (x^*, λ^*) eine Lösung von $\dfrac{\partial L}{\partial \lambda}(x, \lambda) = g(x) = 0,\ \dfrac{\partial L}{\partial x}(\lambda, x) = \nabla f(x) + \sum\limits_{i=1}^{l} \lambda_i \nabla g_i(x) = 0,$ und sei $\hat{L}(x) = L(x, \lambda^*)$. Dann ist x^* ein lokales Minimum [Maximum] von f unter der Nebenbedingung $g_1(x) = \dots, g_l(x) = 0,$ falls $(-1)^l \det(h_{2l+i}) > 0, \quad i = 1, \dots, N-l$ $\big[(-1)^{l+i} \det(h_{2l+i}) > 0, \quad i = 1, \dots, N-l\big]$.

Dabei bezeichnen h_j die $(j \times j)$-Hauptuntermatrizen der Matrix

$$C(x^*) = \begin{pmatrix} 0 & Dg(x^*) \\ Dg(x^*)^T & \nabla^2 \hat{L}(x^*) \end{pmatrix} \in \mathbb{R}^{N+l,N+l}.$$

Betrachte unser Beispiel:

$f(x_1, x_2, x_3) = x_1 + 2x_3 \overset{!}{=} min \,/\, max$

$g_1(x_1, x_2, x_3) = x_1 + x_2 + x_3 - 1 = 0, \quad g_2(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3 - \frac{7}{4} = 0$

Betrachte den Punkt $x^* = \left(0, -\frac{1}{2}, \frac{3}{2}\right), \quad \lambda^* = (-1, -1)$.

(x^*, λ^*) ist stationärer Punkt der Lagrange-Funktion, d.h. erfüllt $\frac{\partial L}{\partial x}(\lambda^*, x^*) = 0, \quad \frac{\partial L}{\partial \lambda}(\lambda^*, x^*) = 0$: Berechne nun die dazugehörige Matrix $C(x^*)$. Benötige hierfür die Matrizen $Dg(x^*)$ und $\nabla^2 \hat{L}(x^*)$. Ausrechnen liefert

$$Dg(x) = \begin{pmatrix} 1 & 1 & 1 \\ 2x_1 & 2x_2 & 1 \end{pmatrix}, \quad Dg(x^*) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}, \quad \nabla^2 \hat{L}(x^*) = \nabla^2 f(x^*) + \sum_{i=1}^{l} \lambda_i^* \nabla^2 g_i(x), \quad \nabla f(x) = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix},$$

$$\nabla^2 f(x) = 0 \in \mathbb{R}^{3,3}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \nabla^2 g_1(x) = 0 \in \mathbb{R}^{3,3}, \quad \nabla g_2(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \\ 1 \end{pmatrix}, \quad \nabla^2 g_2(x) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \nabla^2 g_2(x^*),$$

$$\nabla^2 \hat{L}(x^*) = \underbrace{\nabla^2 f(x^*)}_{=0} + \lambda_1^* \underbrace{\nabla^2 g_1(x^*)}_{=0} + \underbrace{\lambda_2^*}_{=-1} \nabla^2 g_2(x^*) = -\nabla^2 g_2(x^*) = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Aufstellen der C-Matrix: $C(x^*) = \begin{pmatrix} 0 & Dg(x^*) \\ Dg(x^*)^T & \nabla^2 \hat{L}(x^*) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & -2 & 0 & 0 \\ 1 & -1 & 0 & -2 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}$ $\quad N = 3$, $l = 2$, $N - l = 1$

Benötigen hier nur $\det(h_{2l+i}) = \det(h_5) = \det(C(x^*))$.

Also liegt nach obigem Satz ein lokales Minimum vor, falls $\det(C(x^*)) > 0$, bzw. ein lokales Maximum, falls $\det(C(x^*)) < 0$, ist.

Man berechnet nun $\det(C(x^*)) = \det(h_5) = -10 < 0 \quad \Rightarrow \quad x^*$ ist ein lokales Maximum von f unter den Nebenbedingungen $g_1(x) = g_2(x) = 0$.

6.3.6 Konvexe Lagrange-Funktionen

Seien wieder x^*, λ^* Lösungen von $\frac{\partial L}{\partial \lambda}(\lambda, x) = g(x) = 0, \quad \frac{\partial L}{\partial x}(\lambda, x) = \nabla f(x) + \sum_{i=1}^{l} \lambda_i g_i(x) = 0$

Setze wieder $\hat{L}(x) = L(\lambda^*, x) = f(x) + \sum_{i=1}^{l} \lambda_i^* g_i(x), \quad \hat{L} : D \to \mathbb{R},\ D \subset \mathbb{R}^M$

Satz:

Ist die Funktion $\hat{L}$ konkav (konvex), so ist x^* ein lokales Maximum (Minimum) von f unter der NB $g_1(x), \ldots, g_l(x) = 0$.

Kehre zu unserem Beispiel zurück:

$$\nabla^2 \hat{L}(x) = \underbrace{\nabla^2 f(x)}_{=0} + \lambda_1^* \cdot \underbrace{\nabla^2 g_1(x)}_{=0} + \underbrace{\lambda_2^*}_{=-1} \cdot \nabla^2 g_2(x) = -\nabla^2 g_2(x) = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad x \in \mathbb{R}^3, \quad D = \mathbb{R}^5 \text{ konvex}$$

$\nabla^2 \hat{L}(x)$ hat die Eigenwerte $-2, 0$. Also ist die zugehörige Form $Q(y) = y^T \nabla^2 \hat{L}(x) y$ negativ semidefinit für alle $x \in \mathbb{R}^3$, d.h. $\hat{L}$ ist konkav auf $\mathbb{R}^3$ und nach vorigem Satz ist $x^* = \left(0, -\frac{1}{2}, \frac{3}{2}\right)$ ein lokales Maximum.

6.3.7 Bemerkungen zur Lagrange-Funktion

Betrachte $f(x) \overset{!}{=} \min / \max$ unter der NB $g_1(x) = \ldots = g_l(x) = 0$.

Zugehörige Lagrange-Funktion: $L(\lambda, x) = f(x) + \sum_{i=1}^{l} \lambda_i g_i(x)$

Wegen $g_i(x) = g_i(x_1, \ldots, x_N) = 0 \quad \Leftrightarrow \quad -g_i(x) = g_i(x_1, \ldots, x_N) = 0, \quad i = 1, \ldots, l$ gilt:

$\hat{L}(x, \lambda) = f(x) + \sum_{i=1}^{l} \lambda_i(-g_1(x)) = f(x) - \sum_{i=1}^{l} \lambda_i g_i(x)$.

$\hat{L}$ besitzt dieselben stationären Punkte wie L, aber die Lagrangemultiplikatoren werden mit (-1) durchmultipliziert. In der Literatur sind beide Varianten für die Lagrange-Funktion gebräuchlich.

6.3.8 Verhalten bei Parametern (Envelope-Theorem)

Seien $\alpha = (\alpha_1, \ldots, \alpha_M), \quad \beta = (\beta_1, \ldots, \beta_l)$ äußere Parameter.
Zielfunktion: $f = f(x, \alpha) = f(x_1, \ldots, x_N, \alpha_1, \ldots, \alpha_M)$
Nebenbedingung: $g_i(x) = \beta_i, \quad i = 1, \ldots, l$

Die zum parameterabhängigen Problem gehörige Lagrange-Funktion lautet dann $L = L(\lambda, x, \alpha, \beta) = f(x, \alpha) + \sum_{i=1}^{l} \lambda_i (g_i(x) - \beta_i)$.

Es seien $x^* = x^*(\alpha, \beta)$, $\lambda^* = \lambda^*(\alpha, \beta)$, $\alpha \in A$, $\beta \in B$ eine Schar lokaler Extrema von $f(x, \alpha) \overset{!}{=}$ min / max unter der NB $g_i(x) - \beta_i = 0$, $i = 1, \ldots, l$,
d.h. $\frac{\partial L}{\partial \lambda}(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta) = 0$, $\frac{\partial L}{\partial x}(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta) = 0$.

Setze dann $\Phi(\alpha, \beta) = L(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta)$, $\alpha \in A$, $\beta \in B$ "Optimalwertfunkion"

Kleine Änderungen der Optimalwertfunktion Φ bei Änderung von α, β zu $\alpha + \triangle\alpha$, $\beta + \triangle\beta$ werden in 1-ter Näherung durch die Größen $\frac{\partial \Phi}{\partial \alpha}(\alpha, \beta) \cdot \triangle\alpha$ bzw. $\frac{\partial \Phi}{\partial \beta}(\alpha, \beta) \cdot \triangle\beta$ gegeben.

$$\begin{aligned}
\frac{\partial \Phi}{\partial \alpha_j}(\alpha, \beta) &= \frac{\partial}{\partial \alpha_j}(L(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta)) \\
\text{(Kettenregel)} \quad &= \underbrace{\frac{\partial L}{\partial \lambda}(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta)}_{=0} \cdot \frac{\partial \lambda^*}{\partial \alpha_j}(\alpha, \beta) + \underbrace{\frac{\partial L}{\partial x}(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta)}_{=0} \cdot \frac{\partial x^*}{\partial \alpha_j}(\alpha, \beta) \\
&\quad + \frac{\partial L}{\partial \alpha_j}(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta) \cdot 1 \\
&= \frac{\partial L}{\partial \alpha_j}(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta) \\
&= \frac{\partial f}{\partial \alpha_j}(x^*(\alpha, \beta)), \quad j = 1, \ldots, l.
\end{aligned}$$

Die partielle Ableitung von f nach α_j reicht aus, um die Änderung der Optimalwertfunktion Φ in Abhängigkeit von α_j angeben.

$$\begin{aligned}
\frac{\partial \Phi}{\partial \beta_j}(\alpha, \beta) &= \frac{\partial}{\partial \beta_j}(L(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta)) \\
\text{(Kettenregel)} \quad &= \underbrace{\frac{\partial L}{\partial \lambda}(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta)}_{=0} \cdot \frac{\partial \lambda^*}{\partial \beta_j}(\alpha, \beta) + \underbrace{\frac{\partial L}{\partial x}(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta)}_{=0} \cdot \frac{\partial x^*}{\partial \beta_j}(\alpha, \beta) \\
&\quad + \frac{\partial L}{\partial \beta_j}(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta) \cdot 1 \\
&= \frac{\partial L}{\partial \beta_j}(\lambda^*(\alpha, \beta), x^*(\alpha, \beta), \alpha, \beta) \\
&= -\lambda_j^*(\alpha, \beta), \quad j = 1, \ldots, l.
\end{aligned}$$

Der negative Lagrangemultiplikator $-\lambda_j$ gibt die Änderung der Optimalwertfunktion Φ in Abhängigkeit von β_j an.

"Envelope-Theorem"

6.4 Extremwerte unter Gleichheits- und Ungleichheitsnebenbedingungen

6.4.1 Allgemeiner Fall einer Optimierungsaufgabe

Seien $f : D \to \mathbb{R},\ g_i : D \to \mathbb{R},\ i = 1, \dots, l\ ,\ l < N$ sowie $k_i : D \to \mathbb{R},\ i = 1, \dots, m,\ \ D \subset \mathbb{R}^N$ gegeben.

Problem: $f(x) \overset{!}{=} max,\ x \in D$ unter der NB $g_i(x) = 0,\ i = 1, \dots, l$ und $k_i(x) \geq 0,\ i = 1, \dots, m$

Bemerkung: Möchte man lokale Minima suchen, so ist f durch $-f$ zu ersetzen.

Wir setzen
$Z = \{x \in D \mid g_i(x) = 0,\ i = 1, \dots, l,\ \ k_i(x) \geq 0,\ i = 1, \dots, m\}$ "Menge der zulässigen Punkte"

Gesucht ist also ein lokales Maximum von f auf der Menge Z. Z hat jetzt keine so schöne geometrische Struktur mehr, wie im Fall von nur Gleichheitsnebenbedingungen.

Sei $\tilde{x} \in Z$ beliebig. Dann heißt $J_{\tilde{x}} = \{i \in \{1, \dots, m\} \mid k_i(\tilde{x}) = 0\}$ die Menge der bei $\tilde{x} \in Z$ aktiven Ungleichheitsrestriktionen.
Seien weiter $f_i, g_i\ \ i = 1, \dots, l,\ \ k_i,\ i = 1, \dots, m$ zweimal differenzierbar.

$\tilde{x} \in D$ heißt regulär, falls die Vektoren $\nabla g_1(\tilde{x}), \dots, \nabla g_i(\tilde{x}), \nabla k_i(\tilde{x}),\ \ \ i \in J_{\tilde{x}}$ linear unabhängig sind.

6.4.2 Zusammenhang zwischen Gleichheits- und Ungleichheitsrestriktionen

Betrachte die Ungleichheitsrestriktionen $k_i(x) \geq 0,\ \ \ i = 1, \dots, m$

Führe neue Variablen $\underbrace{y_1, \dots, y_m}_{\text{Schlupfvariablen}}$ ein und ersetze $k_i(x) \geq 0$ durch $k_i(x) - y_i^2 = 0,\ \ \ i = 1, \dots, m$

Schreibe auf diese Weise das Problem um in ein Optimierungsproblem mit nur noch Gleichheitsrestriktionen. Dabei erhöht sich die Anzahl der Variablen um m.

Setze $y = (y_1, \dots, y_m)$

$$\hat{f}(x, y) = f(x),\quad \hat{g}_1(x, y) = g_1(x), \dots, \hat{g}_l(x, y) = g_l(x),$$

,

$$\hat{g}_{l+1}(x, y) = k_1(x) - y_1^2, \dots, \hat{g}_{l+m}(x, y) = k_m(x) - y_m^2,\quad \hat{g} = (\hat{g}_1, \dots, \hat{g}_{l+m})$$

Die zum Optimierungsproblem gleichwertige Aufgabe lautet: $\hat{f}(x, y) \overset{!}{=} max$ unter der NB $\hat{g}(x, y) = 0$

Zugehörige Lagrange-Funktion:
$L(\alpha, x, y) = \hat{f}(x, y) + \sum\limits_{i=1}^{m+l} \underbrace{\alpha_i}_{\lambda_i,\ i=1,\dots,l} \hat{g}_i(x, y) = f(x) + \sum\limits_{i=1}^{l} \alpha_i g_i(x) + \sum\limits_{i=1}^{m} \underbrace{\alpha_{l+i}}_{\mu_i,\ i=1,\dots,m} (k_i(x) - y_i^2)$

$L(\lambda, \mu, x, y) = f(x) + \sum\limits_{i=1}^{l} \lambda_i g_i(x) + \sum\limits_{i=1}^{m} \mu_i(k_i(x) - y_i^2)$

Die notwendige Bedingung $1-$ter Ordnung für ein lokales Maximum lautet: $0 = \nabla L(\lambda, \mu, x, y)$

Ausgeschrieben bedeutet dies:

$$\begin{array}{llll}
a) & \frac{\partial L}{\partial \lambda_i}(\lambda,\mu,x,y) & = g_i(x) = 0 & i=1,\dots,l \\
b) & \frac{\partial L}{\partial \mu_i}(\lambda,\mu,x,y) & = k_i(x) - y_i^2 = 0 & i=1,\dots,m \\
c) & \frac{\partial L}{\partial x_j}(\lambda,\mu,x,y) & = \frac{\partial f}{\partial x_j}(x) + \sum_{i=1}^{l} \lambda_i \frac{\partial g_i(x)}{\partial x_j} + \sum_{i=1}^{m} \mu_i \frac{\partial k_i(x)}{\partial x_j} = 0 & j=1,\dots,N \\
d) & \frac{\partial L}{\partial y_i}(\lambda,\mu,x,y) & = -2\mu_i y_i = 0 & i=1,\dots,m
\end{array}$$

$(N+2m+l)$-Gleichungen für $(N+2m+l)$-Variablen $x_1,\dots,x_N,y_1,\dots,y_m,\lambda_1,\dots,\lambda_l,\mu_1,\dots,\mu_m$

Elimination der Schlupfvariablen

Multipliziere (d) mit y_i und erhalte $0 = -2\mu_i y_i^2 = -2\mu_i k_i(x) \quad \Leftrightarrow \quad \mu_i \cdot k_i(x) = 0, \quad i=1,\dots,m$

Ersetze (b) durch $k_i(x) \geq 0, \quad i=1,\dots,m$

Ferner benötigt man $\mu_i \geq 0, \quad i=1,\dots,m$

Erhalte dann

$$\left.\begin{array}{l}
\nabla f(x) + \sum_{i=1}^{l} \lambda_i \nabla g_i(x) + \sum_{i=1}^{m} \mu_i \nabla k_i(x) = 0 \\
\mu_i \cdot k_i(x) = 0, \quad i=1,\dots,m \\
g(x)=0, \quad k(x) \geq 0, \quad \mu_i \geq 0,\ i=1,\dots,m
\end{array}\right\} \text{"Karush-Kuhn-Tucker (KKT) Gleichungen"}$$

Lösbarkeit der KKT-Gleichungen

Satz:

Sei x^* eine lokale Lösung unseres Optimierungsproblems mit $f, g_i, \ i=1,\dots,l, \quad k_i, \ i=1,\dots,m$ zweimal differenzierbar. Ferner sei $x^* \in Z$ regulär. Dann existieren eindeutig Zahlen $\lambda_1^*,\dots,\lambda_l^*, \ \mu_1^*,\dots,\mu_m^*, \quad \mu_i^* \geq 0, \quad i=1,\dots,m$ mit
$\nabla f(x^*) + \sum_{i=1}^{l} \lambda_i^* \nabla g_i(x^*) + \sum_{i=1}^{m} \mu_i^* \nabla k_i(x^*) = 0, \quad \mu_i^* \cdot k_i(x^*) = 0, \quad i=1,\dots,m$
"Karush-Kuhn-Tucker-Bedingung"

Bemerkung: Wenn das Optimierungsproblem eine reguläre Lösung x^* hat, dann haben die KKT-Gleichungen eine Lösung, d.h. man findet die zugehörigen Lagrange-Multiplikatoren λ^* und μ^*.

Beispiel: $f(x) = \frac{1}{2}x_1 - x_2 \stackrel{!}{=} max, \quad D = \mathbb{R}^3, \ N=3, \ l=0, \ m=2$

unter der NB $k_1(x) = -x_1 - \exp(-x_1) - x_3^2 + x_2 \geq 0, \quad k_2(x) = x_1 \geq 0$

$$\nabla f(x) = \begin{pmatrix} \frac{1}{2} \\ -1 \\ 0 \end{pmatrix}, \quad \nabla k_1(x) = \begin{pmatrix} -1+\exp(-x_1) \\ 1 \\ -2x_3 \end{pmatrix}, \quad \nabla k_2(x) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad x = (x_1,x_2,x_3)$$

KKT-Gleichungen

$$\begin{array}{lll}
a) & \frac{1}{2} + \mu_1(-1+\exp(-x_1)) + \mu_2 & = 0 \\
b) & -1 + \mu_1 & = 0 \\
c) & -2x_3\mu_1 & = 0 \\
d) & \mu_1(-x_1 - \exp(-x_1) - x_3^2 + x_2) & = 0 \\
e) & \mu_2 x_1 & = 0
\end{array} \quad \text{(d), (e): Schlupfbedingungen}$$

$$\mu_1, \mu_2 \geq 0, \quad k_1(x), k_2(x) \geq 0$$

b) liefert $\mu_1 = 1$ und mit c) folgt: $x_3 = 0$. Mit d) erhalten wir $\Rightarrow \quad -x_1 - \exp(-x_1) + x_2 = 0 \quad \Rightarrow \quad x_2 = x_1 + \exp(-x_1)$

Zum "Lösen" der Ungleichungen d), e) benutzt man Elemente der mathematischen Logik.

Prinzip der Logik: Man unterstellt eine Annahme, zeigt dass sich durch diese Annahme ein Widerspruch ergibt. Dann ist das Gegenteil der Annahme wahr.

Annahme: Sei $x_1 = 0$.

Dann gilt: $x_2 = 0 + \exp(0) = 1$. Einsetzen in a): $\frac{1}{2} + 1(-1+1) + \mu_2 = 0 \quad \Rightarrow \quad \mu_2 = -\frac{1}{2} < 0$ Widerspruch.
Damit ist das Gegenteil von $x_1 = 0$ wahr, d.h. $x_1 \neq 0$. Weiter gilt dann mit e) sofort $\mu_2 = 0$.

Einsetzen in a) liefert:

$\frac{1}{2} + 1(-1 + \exp(-x_1)) + 0 = 0 \;\Rightarrow\; \exp(-x_1) = \frac{1}{2} \;\Leftrightarrow\; -x_1 = \ln\left(\frac{1}{2}\right) \;\Leftrightarrow\; x_1 = -\ln\left(\frac{1}{2}\right) = \ln\left(\left(\frac{1}{2}\right)^{-1}\right) = \ln(2)$

Ferner folgt: $x_2 = \ln(2) + \exp(-\ln(2)) = \ln(2) + \exp(\ln\left(\frac{1}{2}\right)) = \ln(2) + \frac{1}{2}$

Einzige Lösung: $x^* = \left(\ln(2), \ln(2) + \frac{1}{2}, 0\right), \quad \mu^* = (1, 0)$

Für diese Lösung gilt $k_1(x^*) = 0, \quad k_2(x^*) = \ln(2) > 0$, d.h. Ungleichung k_1 ist aktiv und Ungleichung k_2 ist inaktiv.
(x^*, μ^*) erfüllt die notwendige Bedingung 1-ter Ordnung für ein lokales Maximum.

6.4.3 Hinreichende Bedingungen 2-ter Ordnung für lokale Maxima

Es sei $x^* \in Z = \{x \in D \mid g(x) = 0,\ h(x) \geq 0\}$ und es seien $\lambda^* = (\lambda_1^*, \dots, \lambda_l^*)$, $\mu^* = (\mu_1^*, \dots, \mu_m^*)$, $\mu_i^* \geq 0, \quad i = 1, \dots, m$ mit $\nabla f(x^*) + \sum\limits_{i=1}^{l} \lambda_i^* \nabla g_i(x^*) + \sum\limits_{i=1}^{m} \mu_i^* \nabla k_i(x^*) = 0, \quad \mu_i^* k_i(x^*) = 0, \quad i = 1, \dots, m$, d.h. die KKT-Bedingungen seien erfüllt.

Ferner sei $\mu_i^* > 0$, falls $k_i(x^*) = 0$.

Setze dann $\hat{L}(x) = f(x) + \sum\limits_{i=1}^{l} \lambda_i^* g_i(x) + \sum\limits_{i=1}^{m} \mu_i^* k_i(x)$

Berechne $\nabla^2 \hat{L}(x) = \nabla^2 f(x) + \sum\limits_{i=1}^{l} \lambda_i^* \nabla^2 g_i(x) + \sum\limits_{i=1}^{m} \mu_i^* \nabla^2 k_i(x)$ und betrachte

die quadratische Form $Q(y) = y^T \nabla^2 \hat{L}(x^*) y = y^T \left(\nabla^2 f(x^*) + \sum\limits_{i=1}^{l} \lambda_i^* \nabla^2 g_i(x^*) + \sum\limits_{i=1}^{m} \mu_i^* \nabla^2 k_i(x^*) \right) y$.

Dann liegt ein lokales Maximum x^* von f unter den Nebenbedingungen $g(x) = 0$, $k(x) \geq 0$ vor, falls die quadratische Form $Q(y) = y^T \nabla^2 \hat{L}(x^*) y$ negativ definit unter den Nebenbedingungen

$\nabla g_i(x^*)^T y = 0, \ i = 1, \dots, l, \qquad \nabla k_i(x^*)^T y = 0$ für $i = \{1, \dots, m\}$ mit $k_i(x^*) = 0$ ist.

Seien nun $i_1, \dots, i_r \in \{1, \dots, m\}$ die Indices der bei x^* aktiven Ungleichungen.

Setze $R(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_l(x) \\ k_{i_1}(x) \\ \vdots \\ k_{i_r}(x) \end{pmatrix} \in \mathbb{R}^{l+m}$. Dann muss $Q(y) = y^T (\nabla^2 \hat{L}(x^*)) y$ negativ definit sein bzgl. der Nebenbedingungen $DR(x^*) y = 0$. Also wird die Definitheit von Q an der Determinante der Hauptuntermatrizen von $C(x^*) = \begin{pmatrix} 0 & DR(x^*) \\ DR(x^*)^T & \nabla^2 \hat{L}(x^*) \end{pmatrix} \in \mathbb{R}^{l+r+N,l+r+N}$ entschieden. Ein lokales Maximum liegt vor, falls $(-1)^{l+r+i} \det(h_{2(l+r)+i}) > 0,\ i = 1, \dots N-l-r$. Dabei bezeichnet h_i die $(i \times i)$-Hauptuntermatrix von $C(x^*)$.

6.4.4 Ein Beispiel

Betrachte wieder unser Beispiel: $f(x) = \frac{1}{2}x_1 - x_2 \overset{!}{=} max$
Nebenbedingungen: $k_1(x) = -x_1 - \exp(-x_1) - x_3^2 + x_2 \geq 0, \quad k_2(x) = x_1 \geq 0$
Lösung der KKT-Gleichung: $x^* = \left(\ln(2), \frac{1}{2} + \ln(2), 0\right),\ \mu^* = (1,0),\ k_1(x^*) = 0,\ k_2(x^*) > 0$

Ungleichung k_1 ist aktiv bei x^*, Ungleichung k_2 ist nicht aktiv.
Also folgt: $R(x) = k_1(x)$

$\hat{L}(x) = f(x) + \mu_1^* k_1(x) + \mu_2^* k_2(x), \quad \nabla^2 \hat{L}(x^*) = \nabla^2 f(x) + \underbrace{\mu_1^*}_{=1} \nabla^2 k_1(x^*) + \underbrace{\mu_2^*}_{=0} \nabla^2 k_2(x^*) = \nabla^2 f(x^*) + \nabla^2 k_1(x^*)$

Finde $\nabla^2 f(x) = 0 \in \mathbb{R}^{3,3}$ und $\quad \nabla^2 k_1(x) = \begin{pmatrix} -\exp(-x_1) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}$

Einsetzen liefert: $\nabla^2 \hat{L}(x^*) = \begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}$

$R(x) = k_1(x) = -x_1 - \exp(-x_1) - x_3^2 + x_2, \quad DR(x) = (-1 + \exp(-x_1), 1, -2x_3) \quad \Rightarrow \quad DR(x^*) = \left(-\frac{1}{2}, 1, 0\right)$

Also folgt: $C(x^*) = \begin{pmatrix} 0 & -\frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} \qquad \begin{matrix} N = 3 \\ l = 0 \\ r = 1 \end{matrix}$

Benötige für lokales Maximum $\det(h_3) > 0, \quad \det(h_4) = \det(C(x^*)) < 0$.

$$\det(h_3) = 1 \begin{vmatrix} -\frac{1}{2} & 1 \\ -\frac{1}{2} & 0 \end{vmatrix} = \frac{1}{2} > 0, \qquad \det(h_4) = -2 \underbrace{\begin{vmatrix} 0 & -\frac{1}{2} & 1 \\ -\frac{1}{2} & -\frac{1}{2} & 0 \\ 1 & 0 & 0 \end{vmatrix}}_{\det(h_3)} = -2 \cdot \frac{1}{2} = -1 < 0$$

Also liegt in x^* ein lokales Maximum von f unter den Nebenbedingungen $k_1(x) \geq 0, \quad k_2(x) \geq 0$ vor.

6.4.5 Konkave Optimierungsprobleme

Seien $f : D \to \mathbb{R},\ g_i : D \to \mathbb{R},\ i = 1, \dots, l$ und
$k_i : D \to \mathbb{R},\ i = 1, \dots, m,\ D \subset \mathbb{R}^N$ gegeben.

Definition: Ein Optimierungsproblem $f(x) \overset{!}{=} max$ unter den Nebenbedingungen $g_i(x) = 0,\ i = 1, \dots, l$ und $k_i(x) \geq 0,\ i = 1, \dots, m$ heißt konkaves Optimierungsproblem, falls $D \subset \mathbb{R}^N$ konvexe Menge, f konkav auf D, $g_i(x) = b_i^T x - \beta_i = 0,\ b_i \in \mathbb{R}^N,\ \beta_i \in \mathbb{R},\ i = 1, \dots, l$, k_i konkav auf D, $i = 1, \dots, m$.

Satz:

Vorgelegt sei ein konkaves Optimierungsproblem mit $f, g_i,\ i = 1, \dots, l,\ k_i,\ i = 1, \dots, m$

zweimal differenzierbar. Erfüllt x^* zusammen mit den zugehörigen Lagrange-Multiplikatoren λ^*, μ^* die KKT-Bedingungen, so ist x^* globales Maximum von f unter den Nebenbedingungen $g(x) = 0,\ \ k(x) \geq 0$.

Betrachte wieder unser Beispiel: $f(x) = \frac{1}{2}x_1 - x_2 \overset{!}{=} \ max$

$k_1(x) = -x_1 - \exp(-x_1) - x_3^2 + x_2 \geq 0, \quad k_2(x) = x_1 \geq 0$

KKT-Punkt: $x^* = \big(\ln(2), \frac{1}{2} + \ln(2), 0\big)\,, \ \ \mu^* = (1,0)$

Überprüfe nun, ob ein konkaves Optimierungsproblem vorliegt.

$D = \mathbb{R}^3$, d.h. D ist konvex.
Es gilt: $\nabla^2 f(x) = 0 \in \mathbb{R}^{3,3} \quad \Rightarrow \quad$ f konkav

$$\nabla^2 k_1(x) = \begin{pmatrix} \overbrace{-\exp(-x_1)}^{<0} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Diese Matrix hat die Eigenwerte $-\exp(-x_1), 0, -2$ ist also negativ semidefinit $\quad \Rightarrow \quad k_1$ konkav

$\nabla^2 k_2(x) = 0 \in \mathbb{R}^{3,3} \quad \Rightarrow \quad k_2$ konkav.

Es liegt also ein konkaves Optimierungsproblem vor. Also ist nach vorigem Satz x^* globales Maximum von f unter den Nebenbedingungen $k_i(x) \geq 0$, $i = 1, 2$.

Kapitel 7

Differenzialgleichungen

7.1 Prozesse in der Zeit und skalare Differenzialgleichungen

7.1.1 Modelle für Prozesse in der Zeit

Beispiel 1:

$K(t)$ sei das verfügbare Kapital einer Firma zur Zeit t. Angenommen, das Kapital wurde mit der Rate $\delta > 0$ abgeschrieben und der Kapitalmarktzins sei I.
Dann lässt sich zeigen: Die Änderung der Kapitalmenge ist dann: $I - \delta K(t)$ Andererseits ist die Änderung der Kapitalmenge nach Definition gegeben durch $K'(t)$
Man erhält: $K'(t) = I - \delta K(t), \quad I, \delta > 0, \ \ t \in$ Intervall
So etwas nennt man "skalare Differenzialgleichung".
Gesucht wird eine stetig differenzierbare Funktion $K(t)$, für welche an jedem Zeitpunkt t die Größen $K'(t)$ und $I - \delta K(t)$ übereinstimmen.
Manchmal ist auch die Kapitalmenge zu einem festem Zeitpunkt t_0 bekannt, d.h. es gilt:

$K(t_0) = K_0$ "Anfangsbedingung"

Man betrachtet dann die Aufgabe:

$K'(t) = I - \delta K(t), \quad I, \delta > 0, \ \ t \in$ Intervall

$K(t_0) = K_0$ "Anfangswertaufgabe" (AWA)

Bemerkung: Wenn das Problem (AWA) eine Lösung hat, dann hat die Differenzialgleichung viele Lösungen. Präziser sagt man "skalare AWA", da $K(t) \in \mathbb{R}$.

Beispiel 2:

Sei $x(t)$ das Bruttosozialprodukt eines Volkes, $K(t)$ das Aktienkapital und $L(t)$ die Anzahl der Arbeitenden zur Zeit t. Es gelte:

- Die Änderung des Kapitals ist proportional zum BSP, d.h. $K'(t) = sx(t)$
- Das BSP werde durch eine Cobb-Douglas-Funktion modelliert, d.h. $x(t) = a \cdot K(t)^{1-\alpha} L(t)^{\alpha}, \quad 0 < \alpha < 1$
- Die Arbeitskraft steige exponentiell mit der Anzahl der Arbeitenden, d.h. $L(t) = L_0 \exp(\lambda t)$

Man findet dann

$$K'(t) = sx(t) = s \cdot a \cdot K(t)^{1-\alpha} L(t)^{\alpha} = s \cdot a \cdot L_0^{\alpha} \exp(\alpha \lambda t) \cdot K(t)^{1-\alpha}, \quad t \in I, \ \ s, a, L_0 > 0, \ \ 0 < \alpha < 1$$

"skalare, nichtautonome Differentialgleichung"

Nichtautonom heißt, die unabhängige Variable t tritt explizit auf.
Zu Beginn gelte: $K(t_0) = K_0$

AWA: $K'(t) = s \cdot a \cdot L_0^{\alpha} \exp(\alpha \lambda t) K(t)^{1-\alpha}, \quad t \in I, \ \ K(t_0) = K_0$

Eine Differentialgleichung (Dgl.) heißt nichtautonom, falls die rechte Seite explizit von der Zeit abhängt, sonst heißt sie autonom.

7.1.2 Allgemeine Struktur

Die auftretenden Gleichungen sind Spezialfälle der folgenden Aufgabe:

$$x'(t) = f(t, x(t)), \quad t \in I, \ \ x(t) \in D, \ \ f : I \times D \to \mathbb{R}, \ \ I, D \subset \mathbb{R} \text{ Intervalle.} \qquad (*)$$

Ferner sei f zweimal differenzierbar.

Manchmal schreibt man kurz: $x' = f(t, x), \ \ t \in I, \ \ x \in D$

Eine Lösung von $(*)$ ist eine differenzierbare Funktion $x : I \to D$, die $x'(t) = f(t, x(t)), \ \ t \in I$ erfüllt.
Die Funktion $f : I \times D \to \mathbb{R}$ ist dabei gegeben.

f heißt die rechte Seite des Problems. f verkörpert inhaltlich gesehen den unterliegenden Mechanismus. Manchmal wird die Aufgabe $(*)$ durch eine Anfangsbedingung $x(t_0) = \alpha, \quad t_0 \in I, \quad \alpha \in D \qquad (**)$ ergänzt.

Das Problem $(*), (**)$ heißt AWA.

Die Differenzialgleichung $(*)$ bzw. die Anfangswertaufgabe $(*)$, $(**)$ heißt autonom, falls f nicht explizit von t abhängt ansonsten nichtautonom.

Bemerkung: Einfache Lösungen von Differenzialgleichungen
Jede Nullstelle von f, d.h. jedes $u \in D$ mit $f(t, u) = 0$, $t \in I$ definiert die konstante Lösung $x(t) = u$, $t \in \mathbb{R}$, denn es gilt:

$$x'(t) = \frac{\partial}{\partial t}(u) = 0 = f(t, u) = f(t, x(t)), \ \ t \in \mathbb{R}$$

$x(t) = u$ heißt stationäre Lösung der Differenzialgleichung.

7.2 Lösungsmethoden für skalare Differenzialgleichungen

7.2.1 Separation der Variablen

Vorgelegt sei die allgemeine Aufgabe: $x'(t) = f(t, x(t)),\ \ f : I \times D \ \to \mathbb{R},\ \ I \times D = \{(t,x) \mid t \in I,\ x \in D\}$, f differenzierbar, mit der Anfangsbedingung $x(t_0) = x_0,\ \ t_0 \in I,\ \ x_0 \in D$

Nun gelte zusätzlich: $f(t,x) = h(t) \cdot g(x)$.

<u>Ziel:</u> Löse die Anfangswertaufgabe $x'(t) = h(t) \cdot g(x(t)),\ \ x(t_0) = x_0$ explizit.

Sei $G(x)$ eine Stammfunktion für $\frac{1}{g(x)}$, und sei $H(t)$ eine Stammfunktion von $h(t)$.

Sei nun $\bar{x}(t)$ eine Lösung, so gilt das folgende:

$$\frac{d}{dt} G(\bar{x}(t)) \overset{\text{Kettenregel}}{=} \underbrace{G'(\bar{x}(t))}_{\frac{1}{g(\bar{x}(t))}} \cdot \underbrace{\bar{x}'(t)}_{=f(t,\bar{x}(t))=h(t)g(\bar{x}(t))} = \frac{1}{g(\bar{x}(t))} \cdot g(\bar{x}(t)) \cdot h(t) = h(t) = \frac{d}{dt} H(t)$$

Also folgt durch Integration: $G(\bar{x}(t)) = H(t) + C$

Setze $t = t_0$ ein und finde dann: $G(\bar{x}(t_0)) = H(t_0) + C \quad \Leftrightarrow \quad G(x_0) = H(t_0) + C \quad \Leftrightarrow \quad C = G(x_0) - H(t_0)$

Wir erhalten also insgesamt:

$$G(\bar{x}(t)) - G(x_0) = H(t) - H(t_0) \qquad (*)$$

$$\int_{x_0}^{\bar{x}(t)} \frac{1}{g(x)} dx = \int_{t_0}^{t} h(\tau) d\tau$$

"Separation der Variablen"

$(*)$ ist eine implizite Gleichung für die gesuchte Lösung $\bar{x}(t)$. Lässt sich $(*)$ nach $\bar{x}(t)$ auflösen, so haben wir die Lösung bestimmt.

7.2.2 Behandle jetzt unsere Beispiele

<u>Beispiel 1:</u> Kapitalmodell: $K'(t) = I - \delta \cdot K(t) = f(K(t)),\ \ K(0) = K_0,\ \ t \geq 0$

Die Variable t taucht explizit gar nicht auf. Also hat $f(K)$ die Produktform mit $f = g$ und $h = 1$, d.h. $f(K) = g(K) = I - \delta K$ und somit

$$H(t) = \int 1\ dt = t$$

$$G(K) = \int \frac{1}{g(K)}\ dK = \int \frac{dK}{I - \delta K} = -\frac{1}{\delta} \ln(I - \delta K)$$

Wir finden:

$$\begin{aligned}
G(K(t)) - G(K_0) &= H(t) - H(t_0) \qquad\qquad t_0 = 0\\
-\frac{1}{\delta}\ln(I - \delta K(t)) + \frac{1}{\delta}\ln(I - \delta K_0) &= t - 0 \qquad \mid \cdot (-\delta)\\
\ln(I - \delta K(t)) - \ln(I - \delta K_0) &= -\delta t\\
\ln\left(\frac{I - \delta K(t)}{I - \delta K_0}\right) &= -\delta t\\
\frac{I - \delta K(t)}{I - \delta K_0} &= \exp(-\delta t)\\
I - \delta K(t) &= (I - \delta K_0)\exp(-\delta t)\\
K(t) &= \left(K_0 - \frac{I}{\delta}\right)\exp(-\delta t) + \frac{I}{\delta}, \quad t \geq 0
\end{aligned}$$

Das Ergebnis ist zu prüfen durch Einsetzen in die Gleichung (Probe). Dabei kommen mögliche Einschränkungen bzgl. t automatisch an den Tag.
Probe:

$$\begin{aligned}
K(0) &= \left(K_0 - \frac{I}{\delta}\right) \cdot 1 + \frac{I}{\delta} = K_0\\
K'(t) &= \left(K_0 - \frac{I}{\delta}\right) \cdot \exp(-\delta t)(-\delta) = (I - \delta K_0)\exp(-\delta t)\\
I - \delta K(t) &= I - (\delta K_0 - I)\exp(-\delta t) - I = (I - \delta K_0)\exp(-\delta t) = K'(t), \qquad t \geq 0
\end{aligned}$$

Die Anfangswertaufgabe (AWA) $K'(t) = I \cdot \delta K(t), \quad I, \delta > 0, \ \ t \geq 0, \ \ K(0) = K_0$ hat die Lösung
$K(t) = \left(K_0 - \frac{I}{\delta}\right)\exp(-\delta t) + \frac{I}{\delta}, \quad t \geq 0$

Beispiel 2:

AWA: $K'(t) = \underbrace{s \cdot a \cdot L_0^\alpha \cdot \exp(\alpha\lambda t)}_{=h(t)} \cdot \underbrace{K(t)^{1-\alpha}}_{g(K(t))}, \qquad t \geq 0$

Produktstruktur vorhanden. Separation der Variablen ist prinzipiell anwendbar.

Benötige: $\displaystyle\int_{K_0}^{K(t)} \frac{d\kappa}{\kappa^{1-\alpha}}, \quad \int_0^t saL_0^\alpha \exp(\alpha\lambda\tau)\ d\tau$

$$\int_{K_0}^{K(t)} \frac{d\kappa}{\kappa^{1-\alpha}} = \int_{K_0}^{K(t)} \kappa^{\alpha-1}\ d\kappa = \left[\frac{1}{\alpha}\kappa^\alpha\right]_{\kappa=K_0}^{\kappa=K(t)} = \frac{1}{\alpha}\left(K(t)^\alpha - K_0^\alpha\right)$$

$$\int_0^t saL_0^\alpha \exp(\alpha\lambda\tau)\ d\tau = saL_0^\alpha \cdot \int_0^t \exp(\alpha\lambda\tau)\ d\tau = saL_0^\alpha\left[\frac{1}{\alpha\lambda}\exp(\alpha\lambda\tau)\right]_{\tau=0}^{\tau=t} = \frac{saL_0^\alpha}{\alpha\lambda}\left[\exp(\alpha\lambda t) - 1\right]$$

Implizite Gleichung ($*$) aufstellen:

$$\begin{aligned}
\frac{1}{\alpha}\left(K(t)^\alpha - K_0^\alpha\right) &= \frac{saL_0^\alpha}{\alpha\lambda}\left[\exp(\alpha\lambda t) - 1\right]\\
K(t)^\alpha - K_0^\alpha &= \frac{saL_0^\alpha}{\lambda}\left(\exp(\alpha\lambda t) - 1\right)\\
K(t)^\alpha &= \frac{saL_0^\alpha}{\lambda}\left(\exp(\alpha\lambda t) - 1\right) + K_0^\alpha\\
K(t) &= \left[\frac{saL_0^\alpha}{\lambda}\left(\exp(\alpha\lambda t) - 1\right) + K_0^\alpha\right]^{\frac{1}{\alpha}}, \qquad t \geq 0
\end{aligned}$$

Probe:

$$
\begin{aligned}
K(0) &= (K_0^\alpha)^{\frac{1}{\alpha}} = K_0 \\
K'(t) &= \frac{1}{\alpha}\left[\frac{saL_0^\alpha}{\lambda}(\exp(\alpha\lambda t)-1)+K_0^\alpha\right]^{\frac{1}{\alpha}-1}\cdot\frac{saL_0^\alpha}{\lambda}\exp(\alpha\lambda t)\cdot\alpha\lambda \\
&= \left[\frac{saL_0^\alpha}{\lambda}(\exp(\alpha\lambda t)-1)+K_0^\alpha\right]^{\frac{1-\alpha}{\alpha}}\cdot\exp(\alpha\lambda t)\cdot saL_0^\alpha \\
&= \left[\underbrace{\left[\frac{saL_0^\alpha}{\lambda}(\exp(\alpha\lambda t)-1)+K_0^\alpha\right]^{\frac{1}{\alpha}}}_{=K(t)}\right]^{1-\alpha}\cdot saL_0^\alpha\exp(\alpha\lambda t) \\
&= K(t)^{1-\alpha}\cdot saL_0^\alpha\exp(\alpha\lambda t), \qquad t\geq 0 \\
\Rightarrow \quad K(t) &= \left[\frac{saL_0^\alpha}{\lambda}(\exp(\alpha\lambda t)-1)+K_0^\alpha\right]^{\frac{1}{\alpha}}, \qquad t\geq 0 \quad \text{ist Lösung der AWA}
\end{aligned}
$$

Bis jetzt wurde ausschließlich mit Separation der Variablen gearbeitet. Nun wird ein weiteres Verfahren zum Lösen von Differenzialgleichungen vorgestellt.

7.2.3 Lineare nichtautonome Differentialgleichungen

Betrachte die AWA: $x'(t) = f(t,x(t)) = a(t)x(t)+b(t)$ mit stetigen Funktionen $a,b: I\to\mathbb{R}$ mit der Anfangsbedingung $x(t_0)=x_0$ für ein $t_0\in I$, $I\subset\mathbb{R}$ Intervall.

Bemerkung: Separation der Variablen ist i.a. nicht anwendbar.

Die obige AWA heißt homogen, falls $b(t)=0$, $t\in I$, d.h. $x'(t)=a(t)x(t)$, $x(t_0)=x_0$ $\qquad (**)$
Die Lösung von $(**)$ kann man direkt angeben. Sie lautet

$$
\begin{aligned}
x(t) &= x_0\exp\left(\int_{t_0}^{t} a(\tau)\,d\tau\right), \qquad t\in I,\ \text{denn} \\
x(t_0) &= x_0\exp\left(\int_{t_0}^{t_0} a(\tau)\,d\tau\right) = x_0\exp(0)=x_0 \\
x'(t) &= \left(x_0\exp(\int_{t_0}^{t} a(\tau)\,d\tau)\right)' \\
&\overset{\text{Kettenregel}}{=} \underbrace{x_0\exp(\int_{t_0}^{t} a(\tau)\,d\tau)}_{=x(t)}\cdot\underbrace{\left(\int_{t_0}^{t} a(\tau)\,d\tau\right)'}_{=a(t)} \\
&= a(t)x(t), \qquad t\in I.
\end{aligned}
$$

Zum Verständnis des letzten Gleichheitszeichen betrachte man $F(t)=\int_a^t f(\tau)\,d\tau$ und f stetig.

Nach dem Hauptsatz der Differenzial- und Integralrechnung ist dann $F(t)$ eine differenzierbare Funktion und es gilt: $F'(t)=f(t)$

Beachte nun etwas allgemeiner: $F(t)=\int_a^t f(t,\tau)\,d\tau$, d.h. der Integrand f hängt ebenfalls von t ab.

Dann gilt: Ist f stetig in τ und t und existiert $\frac{\partial f}{\partial t}(t,\tau)$, so ist F differenzierbar mit

$$F'(t)=f(t,t)+\int_a^t \frac{\partial f}{\partial t}f(t,\tau)\,d\tau, \qquad t\in I$$

Betrachte jetzt die AWA $x'(t) = a(t)x(t) + b(t), \quad x(t_0) = x_0 \qquad (***)$

Es sei $\quad y(t) = \exp\left(\int\limits_{t_0}^{t} a(\tau)\, d\tau\right), \qquad t \in I$

$y(t),\ t \in I$ löst die AWA $x'(t) = a(t)x(t), \quad x(t_0) = 1$

Man nennt $y(t)$ auch die Fundamentallösung von $(**)$

Die Lösung von $(***)$ hat dann die Form $x(t) = y(t) \cdot x_0 + \int\limits_{t_0}^{t} y(t)y(s)^{-1}b(s)\, ds, \qquad t \in I$

"Variation der Konstanten"

Dies ist eine explizite Lösungsformel für AWAs der Form $x'(t) = a(t)x(t) + b(t), \quad x(t_0) = x_0$
Eine Probe ist folglich nicht notwendig.

Es gilt:

$$x(t_0) = \underbrace{y(t_0)}_{=1} \cdot x_0 + \underbrace{\int_{t_0}^{t_0} y(t)y(s)^{-1}b(s)\, ds}_{=0} = x_0$$

$$\begin{aligned}
x'(t) &= y'(t)x_0 + \underbrace{\left(\int_{t_0}^{t} y(t)y(s)^{-1}b(s)\, ds\right)'}_{=F(t)} \\
&= \underbrace{y'(t)x_0}_{=a(t)y(t)} + \underbrace{y(t)y(t)^{-1}b(t)}_{=b(t)} + \int_{t_0}^{t} \underbrace{\frac{\partial}{\partial t}\left(y(t)y(s)^{-1}b(s)\right)}_{=y'(t)y(s)^{-1}b(s)=a(t)y(t)y(s)^{-1}b(s)} ds \\
&= a(t)y(t)x_0 + b(t) + \int_{t_0}^{t} a(t)y(t)y(s)^{-1}b(s)\, ds \\
&= a(t) \underbrace{\left(y(t)x_0 + \int_{t_0}^{t} y(t)y(s)^{-1}b(s) ds\right)}_{=x(t)} + b(t) \\
&= a(t)x(t) + b(t), \qquad t \in I
\end{aligned}$$

Beispiel:

Vorglegt sei $x'(t) = -2x(t) + \exp(-t), \quad t \in \mathbb{R}, \qquad x(0) = 3$

$x'(t) = -2x(t) + \exp(-t) = a(t)x(t) + b(t)$ mit $a(t) = -2,\ b(t) = \exp(-t) \qquad t_0 = 0, \quad x_0 = 3$

Lösung durch Variation der Konstanten:

Berechne die Fundamentallösung $y(t) = \exp\left(\int\limits_{0}^{t} -2\, d\tau\right) = \exp([-2\tau]_{\tau=0}^{\tau=t}) = \exp(-2t)$

Wende die Lösungsformel an:

$$\begin{aligned}
x(t) &= \underbrace{\exp(-2t)}_{y(t)} \cdot \underbrace{3}_{x_0} + \int_0^t \underbrace{\exp(-2t)}_{y(t)} \underbrace{\frac{\overbrace{1}^{y(s)}}{\underbrace{\exp(-2s)}_{=\exp(2s)}} \cdot \overbrace{\exp(-s)}^{b(s)}}_{=\exp(s)} \, ds \\
&= 3\exp(-2t) + \exp(-2t) \cdot \int_0^t \exp(s) \, ds \\
&= 3\exp(-2t) + \exp(-2t) \cdot \underbrace{[\exp(s)]_{s=0}^{s=t}}_{=\exp(t)-\exp(0)=\exp(t)-1} \\
&= 3\exp(-2t) + \exp(-2t) \cdot (\exp(t) - 1) \\
&= 2\exp(-2t) + \exp(-t), \qquad t \in I = \mathbb{R}.
\end{aligned}$$

<u>Bemerkung:</u> Da zur Lösung von AWA der Form $x'(t) = a(t)x(t) + b(t),\ \ t \in I,\ \ x(t_0) = x_0$ eine explizite Lösungsformel zur Verfügung steht, ist eine Probe überflüssig.

7.3 Lineare Systeme von Differenzialgleichungen

7.3.1 Beispiel

Beispiel eines Marktmodells: Betrachte ein Marktmodell, in dem der Preis eines Produktes angepasst wird an den Nachfrage- und Lieferüberschuss und die Anzahl der Firmen, welche das Produkt herstellen, je nachdem ob Gewinne oder Verluste erzielt werden.

Sei $p(t)$ der Produktpreis und $N(t)$ Anzahl der Firmen, die das Produkt herstellen zur Zeit t.
Modellannahmen:

- Die Veränderung des Preises $p'(t)$ ist (proportional zu) $\sim \quad q^N - q^L$ mit
  $q^N = a + bp$ Nachfrage
  $q^L = mN$ Lieferfunktion
- Es sei $\bar{c}$ ein fester durchschnittlicher Produktionspreis
- $N'(t) \sim p - \bar{c}$
  $\rightarrow \quad N'(t)\left\{ \begin{matrix} > \\ < \end{matrix} \right\} 0$ genau dann, wenn Produktpreis $\left\{ \begin{matrix} > \\ < \end{matrix} \right\}$ Produktionspreis
  $N'(t) \sim p - \bar{c}$: Änderung Anzahl der Firmen ist proportional zum Produktpreis.

$N'(t) = \gamma(p(t) - \bar{c})$
$p'(t) = \alpha\left(q^N - q^L\right) = \alpha(a + bp(t) - mN(t))$ mit Konstanten: $\alpha, \gamma, \bar{c}, a > 0$

Dabei sind α, γ Proportionalitätsfaktoren. Diese regeln die Geschwindigkeit der Anpassung an die Marktsituation. Kompakt geschrieben lautet dies (als Matrix):

$$\begin{pmatrix} N'(t) \\ p'(t) \end{pmatrix} = \begin{pmatrix} 0 & \gamma \\ -m\alpha & \alpha b \end{pmatrix} \cdot \begin{pmatrix} N(t) \\ p(t) \end{pmatrix} + \underbrace{\begin{pmatrix} -\gamma\bar{c} \\ \alpha a \end{pmatrix}}_{=b\in\mathbb{R}^2}, \qquad t \geq 0$$

"zweidimensionales lineares Differentialgleichungssystem"

7.3.2 Allgemeine Problemstruktur

Sei $A \in \mathbb{R}^{N,N}$ und sei $b \in \mathbb{R}^N$ gegeben. Betrachte das lineare Differentialgleichungssystem
$x'(t) = Ax(t) + b$, d.h.

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_N'(t) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{pmatrix} \cdot \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_N(t) \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix}$$

Vergleiche Beispiel: $N = 2$

$x_1(t) = N(t), \quad x_2(t) = p(t), \quad a_{11} = 0, \quad a_{12} = \gamma, \quad a_{21} = -m\alpha, \quad a_{22} = \alpha b, \quad b_1 = -\gamma\bar{c}, \quad b_2 = \alpha a$

Dieses System nennt man homogen, wenn der Vektor b der Nullvektor ist ($b = 0 \in \mathbb{R}^N$), und ansonsten inhomogen.

Betrachte zunächst das homogene lineare System:
$x'(t) = Ax(t), \quad A \in \mathbb{R}^{N,N}, \qquad x(t) = (x_1(t), \dots, x_N(t)), \quad t \in \mathbb{R}$

Wiederholung: Zu einer Matrix $A \in \mathbb{R}^{N,N}$ heißt ein Vektor $x \in \mathbb{C}^N$ Eigenvektor zum Eigenwert $\lambda \in \mathbb{C}$, falls $x \neq 0$ und $Ax = \lambda x$. $\lambda \in \mathbb{C}$ ist Eigenwert von A genau dann, wenn $p(\lambda) = \det(A - \lambda I_N) = 0$.

Lösung von $x'(t) = Ax(t)$ mit dem Exponentialansatz:
Sei $x(t) = \exp(\lambda t)v$ mit $Av = \lambda v, \quad v \neq 0$, (d.h. λ ist Eigenwert und v Eigenvektor von A). Dann gilt:

$$\begin{aligned} x'(t) &= \lambda \exp(\lambda t)v = \exp(\lambda t)\lambda v = \exp(\lambda t)Av \\ &= A\underbrace{\exp(\lambda t)v}_{=x(t)} \\ &= Ax(t), \qquad t \in \mathbb{R} \end{aligned}$$

Zu jedem reellen Eigenwert λ mit reellem Eigenvektor v erhalten wir damit eine reelle Lösung.

Jetzt mit komplexen Zahlen:

Sei nun $\lambda = \mu + i\nu$ ein komplexer Eigenwert und $v = a + ib, \quad a, b \in \mathbb{R}^N$ zugehöriger komplexer Eigenvektor von A.
Dann liefert $x(t) = \exp(\lambda t)v$ eine komplexe Lösung.
Man ist aber normalerweise an reellen Lösungen interessiert. Betrachte dazu die Zerlegung von $x(t) = u(t) + iw(t)$ mit $u(t) = Re(x(t))$ und $w(t) = Im(x(t))$ in Real- und Imaginärteil.
Dann gilt:

$$x'(t) = u'(t) + iw'(t) \quad = \quad Ax(t) = A(u(t) + iw(t)) = Au(t) + iAw(t)$$

Vergleich der Real- und Imaginärteile liefert: $u'(t) = Au(t), \quad w'(t) = Aw(t), \quad t \in \mathbb{R}$

Konsequenz: Findet man eine komplexe Lösung $x(t)$ von $x'(t) = Ax(t)$, so sind $Re(x(t)) = u(t)$ und $Im(x(t)) = w(t)$ jeweils reelle Lösungen.

Konkret erhält man:

Sei $\lambda = \mu + i\nu, \quad v = a + ib$

$$\begin{aligned} x(t) &= \exp(\lambda t)v = \exp((\mu + i\nu)t)(a + ib) = \exp(\mu t)\exp(i\nu t)(a + ib) = \exp(\mu t)(\cos(\nu t) + i\sin(\nu t))(a + ib) \\ &= \exp(\mu t)(\cos(\nu t)a + i\sin(\nu t)a + i\cos(\nu t)b - \sin(\nu t)b) \\ &= \underbrace{\exp(\mu t)(\cos(\nu t)a - \sin(\nu t)b)}_{=Re(x(t))=u(t)\in\mathbb{R}} + i\underbrace{\exp(\mu t)(\sin(\nu t)a + \cos(\nu t)b)}_{=Im(x(t))=w(t)\in\mathbb{R}} \end{aligned}$$

Erhalte aus $x(t)$ die beiden reellen Lösungen

$$\begin{aligned} u(t) &= \exp(\mu t)(\cos(\nu t)a - \sin(\nu t)b), \qquad t \in \mathbb{R}, \qquad \mu = Re(\lambda) \\ w(t) &= \exp(\mu t)(\cos(\nu t)b + \sin(\nu t)a), \qquad t \in \mathbb{R}, \qquad \nu = Im(\lambda) \end{aligned}$$

7.3.3 Weitere Eigenschaften von linearen Differenzialgleichungen

Seien jetzt $x^i(t) \in \mathbb{R}^N, \quad i = 1, \ldots, l$ Lösungen von $x'(t) = Ax(t)$.

Dann ist auch $x(t) = \alpha_1 x^1(t) + \alpha_2 x^2(t) + \ldots + \alpha_l x^l(t) = \sum_{i=1}^{l} \alpha_i x^i(t)$ eine Lösung, denn

$$\begin{aligned} x'(t) &= \alpha_1 \underbrace{(x^1(t))'}_{=Ax^1(t)} + \alpha_2 \underbrace{(x^2(t))'}_{=Ax^2(t)} + \ldots + \alpha_l \underbrace{(x^l(t))'}_{=Ax^l(t)} \\ &= \alpha_1 \cdot Ax^1(t) + \alpha_2 \cdot Ax^2(t) + \ldots + \alpha_l \cdot Ax^l(t) \\ &= A(\alpha_1 x^1(t) + \alpha_2 x^2(t) + \ldots + a_l x^l(t)) \\ &= Ax(t) \end{aligned}$$

Gilt nun $l = N \ \left(A \in \mathbb{R}^{N,N}\right)$ und sind die Vektoren $x^1(0), x^2(0), \ldots, x^N(0) \in \mathbb{R}^N$ linear unabhängig, d.h. eine Basis des $\mathbb{R}^N$, so heißt

$$(+) \quad x(t) = \alpha_1 x^1(t) + \alpha_2 x^2(t) + \ldots + \alpha_N x^N(t), \quad t \in \mathbb{R}, \quad \alpha_1, \ldots, \alpha_N \in \mathbb{R}$$

die allgemeine Lösung von $x'(t) = Ax(t)$, d.h. jede Lösung hat diese Form.

Betrachte jetzt die dazugehörige Anfangswertaufgabe: $x'(t) = Ax(t), \quad x(0) = c \in \mathbb{R}^N, \quad A \in \mathbb{R}^{N,N}$.

Seien $x^1(t), \ldots, x^N(t)$ die Basislösungen und $x^1(0), \ldots, x^N(0)$ eine Basis des $\mathbb{R}^N$. Dann gilt:

$$c = \sum_{i=1}^{N} \beta_i x^i(0) \qquad \text{(Basiseigenschaft)}$$

Dann ist $\bar{x}(t) = \sum\limits_{i=1}^{N} \beta_i x^i(t)$ die Lösung der Anfangswertaufgabe ($x'(t) = Ax(t),\ x(0) = c$), denn es hat die Form (+) und $\bar{x}(0) = \sum\limits_{i=1}^{N} \beta_i x^i(0) = c$.

Beispiel: $N = 3, \quad x'(t) = Ax(t), \quad x(t) = (x_1(t), x_2(t), x_3(t)), \quad A = \begin{pmatrix} 1 & -2 & 0 \\ 2 & 0 & 1 \\ -4 & -2 & -1 \end{pmatrix} \in \mathbb{R}^{3,3}$

Berechnung der Eigenwerte: $p(\lambda) = \det(A - \lambda I_3) = \det \begin{pmatrix} 1-\lambda & -2 & 0 \\ 2 & -\lambda & 1 \\ -4 & -2 & -1-\lambda \end{pmatrix} \stackrel{!}{=} 0$

$$\begin{aligned}
\Rightarrow \quad 0 &= (1-\lambda)\left((-\lambda)(-1-\lambda) + 2\right) + 2(2(-1-\lambda) + 4) \\
&= (1-\lambda)(\lambda + \lambda^2 + 2) + 2 \cdot 2(1-\lambda) \\
&= (1-\lambda)(\lambda^2 + \lambda + 6) \qquad \Rightarrow \quad \lambda_1 = 1 \\
\lambda_{2,3} &= \frac{-1 \pm \sqrt{1 - 4\cdot 6}}{2} = \frac{-1 \pm i\sqrt{23}}{2} = -\frac{1}{2} \pm i\sqrt{\frac{23}{4}}, \qquad \alpha = \frac{23}{4} \\
\lambda_{2,3} &= -\frac{1}{2} \pm i\sqrt{\alpha}, \qquad \bar{\lambda}_3 = \lambda_2, \qquad \lambda_1 \in \mathbb{R},\ \lambda_2, \lambda_3 \in \mathbb{C}
\end{aligned}$$

Berechnung der Eigenvektoren v^i, $i = 1,2,3$: $(A - \lambda_i I)v^i = 0, \qquad i = 1,2,3$

$$\Rightarrow \quad v^1 = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix}, \quad v^2 = \frac{1}{12}\begin{pmatrix} 1 - 2i\sqrt{\alpha} \\ -5 - 2i\sqrt{\alpha} \\ 12 \end{pmatrix}, \quad v^3 = \bar{v}^2 = \frac{1}{12}\begin{pmatrix} 1 + 2i\sqrt{\alpha} \\ -5 + 2i\sqrt{\alpha} \\ 12 \end{pmatrix}$$

Dabei ist jeweils die 3-te Komponente der Eigenvektoren auf 1 normiert.

Zugehörige Basislösungen: $x^1(t) = \exp(\lambda_1 t)v^1 = \exp(t) \cdot \begin{pmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix}$

Hat man ein paar konjugierte komplexe Eigenwerte $\lambda_{2,3}$, $\quad \bar{\lambda}_2 = \lambda_3$ mit Eigenvektoren v^2, v^3, $\quad \bar{v}^3 = v^2$ so erhält man zwei Basislösungen aus λ_2 und v^2. Das Paar λ_3 und v^3 braucht nicht analysiert werden, es liefert keine neuen Basislösungen.

Konstruktion der 3 Basislösungen:

Reeller Eigenwert: $\quad \left\{ x^1(t) = \exp(\lambda_1 t)v^1 = \exp(t)\begin{pmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \right.$

Komplexer Eigenwert: $\quad \begin{cases} x^2(t) = Re(\exp(\lambda_2 t)v^2) = \exp(\mu t)(\cos(\nu t)a - \sin(\nu t)b) \\ x^3(t) = Im(\exp(\lambda_2 t)v^2) = \exp(\mu t)(\cos(\nu t)b + \sin(\nu t)a) \end{cases}$

mit

$$\lambda_2 = \underbrace{-\frac{1}{2}}_{=\mu} + i\underbrace{\sqrt{\alpha}}_{=\nu} = \mu + i\nu, \qquad v^2 = \frac{1}{12}\begin{pmatrix} 1 - 2i\sqrt{\alpha} \\ 5 - 2i\sqrt{\alpha} \\ 12 \end{pmatrix} = \underbrace{\frac{1}{12}\begin{pmatrix} 1 \\ -5 \\ 12 \end{pmatrix}}_{=a} + i \cdot \underbrace{\frac{1}{12}\begin{pmatrix} -2\sqrt{\alpha} \\ -2\sqrt{\alpha} \\ 0 \end{pmatrix}}_{=b} = a + ib.$$

Somit folgt

$$x^2(t) = \exp(-\tfrac{1}{2}t)\left(\cos(\sqrt{\alpha}t)\cdot\tfrac{1}{12}\begin{pmatrix}1\\-5\\12\end{pmatrix} - \sin(\sqrt{\alpha}t)\cdot\tfrac{1}{12}\begin{pmatrix}-2\sqrt{\alpha}\\-2\sqrt{\alpha}\\0\end{pmatrix}\right), \qquad t\in\mathbb{R},$$

$$x^3(t) = \exp(-\tfrac{1}{2}t)\left(\cos(\sqrt{\alpha}t)\cdot\tfrac{1}{12}\begin{pmatrix}-2\sqrt{\alpha}\\-2\sqrt{\alpha}\\0\end{pmatrix} + \sin(\sqrt{\alpha}t)\cdot\tfrac{1}{12}\begin{pmatrix}1\\-5\\12\end{pmatrix}\right), \qquad t\in\mathbb{R}.$$

Lösungsgesamtheit: $\mathbb{L} = \left\{x(t)\;\middle|\; x(t) = \alpha_1 x^1(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t), \quad t\in\mathbb{R}, \quad \alpha_1,\alpha_2,\alpha_3\in\mathbb{R}\right\}$

Dies sind alle Lösungen, da $x^1(0) = \begin{pmatrix}-\frac{1}{2}\\0\\1\end{pmatrix}$, $x^2(0) = \frac{1}{12}\begin{pmatrix}1\\-5\\12\end{pmatrix}$, $x^3(0) = \frac{1}{12}\begin{pmatrix}-2\sqrt{\alpha}\\-2\sqrt{\alpha}\\0\end{pmatrix}$ linear unabhängig sind.

Lösung der Anfangswertaufgabe:

Wähle aus der Lösungsgesamtheit $\mathbb{L}$ diejenige Lösung aus, welche $x(0) = c$ erfüllt, d.h. $x(0) = \alpha_1 x^1(0) + \alpha_2 x^2(0) + \alpha_3 x^3(0) = c \in \mathbb{R}^3$
Als lineares Gleichungssystem geschrieben lautet dies

$$\begin{pmatrix}x^1(0) & x^2(0) & x^3(0)\end{pmatrix}\cdot\begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix} = \begin{pmatrix}(x^1(0))_1 & (x^2(0))_1 & (x^3(0))_1\\(x^1(0))_2 & (x^2(0))_2 & (x^3(0))_2\\(x^1(0))_3 & (x^2(0))_3 & (x^3(0))_3\end{pmatrix}\cdot\begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix} = \begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}$$

In der i-ten Spalte der Matrix des zu lösenden GLS steht $x^i(0), \quad i = 1,\dots,N$
Für unser Beispiel mit $c = (3,-10,22)$ folgt:

$$\begin{pmatrix}x^1(0) & x^2(0) & x^3(0)\end{pmatrix}\cdot\begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix} = \begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix} \qquad\Leftrightarrow\qquad \begin{pmatrix}-\frac{1}{2} & \frac{1}{12} & -\frac{1}{6}\sqrt{\alpha}\\ 0 & -\frac{5}{12} & -\frac{1}{6}\sqrt{\alpha}\\ 1 & 1 & 0\end{pmatrix}\cdot\begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix} = \begin{pmatrix}3\\-10\\22\end{pmatrix}$$

Beachte, da die $x^i(0), \quad i = 1,\dots,N$ linear unabhängig sind, ist obige Matrix invertierbar und die AWA ist eindeutig lösbar. Diese ist mit dem Gauß-Verfahren berechenbar.

In unserem Beispiel erhalten wir die Lösung $\alpha_1 = -2, \quad \alpha_2 = 24, \quad \alpha_3 = 0$
Die Lösung des Anfangswertproblems lautet:

$$\begin{aligned} x(t) &= -2x^1(t) + 24x^2(t) + 0\cdot x^3(t)\\ &= -2\exp(t)\cdot\begin{pmatrix}-\frac{1}{2}\\0\\1\end{pmatrix} + 2\exp(-\tfrac{1}{2}t)\left(\cos(\sqrt{\alpha}t)\begin{pmatrix}1\\-5\\12\end{pmatrix} - \sin(\sqrt{\alpha}t)\begin{pmatrix}-2\sqrt{\alpha}\\-2\sqrt{\alpha}\\0\end{pmatrix}\right), \qquad t\in\mathbb{R}.\end{aligned}$$

7.3.4 Lösung inhomogener linearer Systeme

Gegeben: $A\in\mathbb{R}^{N,N}, \quad b\in\mathbb{R}^N$
Gesucht: $x(t) = (x_1(t),\dots,x_N(t))$ mit $x'(t) = Ax(t) + b$

Sei $x^p(t)$ eine Lösung von $x'(t) = Ax(t) + b$ und sei $x(t)$ irgendeine Lösung des homogenen Systems $x'(t) = Ax(t)$. Setze $\bar{x}(t) = x^p(t) + x(t)$.

Dann gilt:

$$\bar{x}'(t) = (x^p(t))' + x'(t) = Ax^p(t) + b + Ax(t) = A(x^p(t) + x(t)) + b = A\bar{x}(t) + b, \qquad t\in\mathbb{R}$$

$x^p(t)$ nennt man partikuläre Lösung des inhomogenen Systems.

Bemerkung: Addiert man zu $x^p(t)$ irgendeine Lösung des homogenen Systems, so erhält man wieder eine

Lösung des inhomogenen Systems.

Man erhält die allgemeine Lösung des inhomogenen Systems, indem man zu einer Lösung $x^p(t)$ des inhomogenen Systems alle Lösungen des homogenen Systems hinzuaddiert. Somit folgt

$\mathbb{L} = \left\{ x(t) \;\middle|\; x(t) = x^p(t) + \sum_{i=1}^{N} \alpha_i x^i(t), \quad t \in \mathbb{R}, \ \alpha_i \in \mathbb{R}, \ i = 1, \ldots, N \right\}$ mit den Basislösungen $x^i(t), \ i = 1, \ldots, N$ des homogenen Systems.

<u>Ansatz für eine partikuläre Lösung $x^p(t)$:</u>

Suche eine Lösung in der Form $x^p(t) = u = \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix}, \qquad t \in \mathbb{R}$, d.h. als stationäre Lösung.

Einsetzen in die Differenzialgleichung liefert: $0 = Au + b \quad \Leftrightarrow \quad Au = -b$

Ist $A \in \mathbb{R}^{N,N}$ invertierbar, so erhalten wir $x^p(t) = u = -A^{-1}b$ als partikuläre Lösung.

7.3.5 Lineare Differenzialgleichungen höherer Ordnung

Betrachte die Aufgabe:

$$(*) \qquad x^{(m)}(t) + a_{m-1}x^{(m-1)}(t) + \ldots + a_1 x'(t) + a_0 x(t) = b, \qquad t \in \mathbb{R}$$

$a_0, a_1, \ldots, a_{m-1}, b \in \mathbb{R}$ seien gegeben.

$(*)$ heißt homogen, falls $b = 0$, sonst inhomogen.

<u>a) Homogene Gleichungen</u> $(b = 0)$

Lösungsansatz: $x(t) = \exp(\lambda t), \qquad t \in \mathbb{R}$

$x'(t) = \lambda \exp(\lambda t), \ldots, \quad$ allgemein: $x^{(i)}(t) = \lambda^i \exp(\lambda t), \qquad i = 1, \ldots, m$

$$x^{(m)}(t) + a_{m-1}x^{(m-1)}(t) + \ldots + a_1 x'(t) + a_0 x(t) = \lambda^m \exp(\lambda t) + a_{m-1}\lambda^{m-1}\exp(\lambda t) + \ldots + a_1 \lambda \exp(\lambda t) + a_0 \exp(\lambda t)$$
$$= \big(\underbrace{\lambda^m + a_{m-1}\lambda^{m-1} + \ldots + a_1\lambda + a_0}_{=:p(\lambda)} \big) \exp(\lambda t) \overset{!}{=} 0$$

$\Rightarrow \quad x(t) = \exp(\lambda t), \quad t \in \mathbb{R}$ liefert eine Lösung von $(*)$ mit $b = 0$,
falls $p(\lambda) = a_0 + a_1\lambda + \ldots + a_{m-1}\lambda^{m-1} + \lambda^m = 0$

$p(\lambda)$ heißt charakteristisches Polynom zu $(*)$ mit $b = 0$

Konsequenz:

- Zu jeder reellen Nullstelle λ von p erhalten wir die reelle Basislösung $x(t) = \exp(\lambda t), \quad t \in \mathbb{R}$
- Zu jeder komplexen Nullstelle $\lambda = \mu + i\nu$ erhalten wir:
  $x(t) = \exp((\mu + i\nu)t) = \exp(\mu t)\exp(i\nu t) = \exp(\mu t)(\cos(\nu t) + i\sin(\nu t)) = \exp(\mu t)\cos(\nu t) + i\exp(\mu t)\sin(\nu t)$,
  $t \in \mathbb{R}$, d.h. zwei reelle Basislösungen
  $x^1(t) = \exp(\mu t)\cos(\nu t)$
  $x^2(t) = \exp(\mu t)\sin(\nu t), \qquad t \in \mathbb{R}$

<u>Bemerkung:</u> Mit $\lambda = \mu + i\nu$ ist auch $\bar{\lambda} = \mu - i\nu$ Nullstelle von p. $\bar{\lambda}$ liefert keine neuen Basislösungen.
Auf diese Weise erhält man zu einem charakteristischen Polynom p mit m-verschiedenen Nullstellen in $\mathbb{C}$ m Basislösungen $x^1(t), \ldots, x^m(t)$.

Die allgemeine Lösung hat dann die Form $x(t) = \sum_{i=1}^{m} \alpha_i x^i(t), \quad t \in \mathbb{R}, \ \alpha_1, \ldots, \alpha_m \in \mathbb{R}$.

Zur Festlegung der Koeffizienten $\alpha_i, \ i = 1, \ldots, m$ sind Anfangsbedingungen vorzugeben. Genauer gibt man für eine Dgl. m-ter Ordnung $x(0) = c_0, \quad x'(0) = c_1, \quad x''(0) = c_2, \ldots, x^{(m-1)}(0) = c_{m-1}$ vor.

Aus dieser Vorgabe lassen sich die freien Koeffizienten $\alpha_1, \dots, \alpha_m$ bestimmen.

b) Inhomogene Gleichungen $(b \neq 0)$

Gegeben: $x^{(m)}(t) + a_{m-1}x^{(m-1)}(t) + \ldots + a_1 x'(t) + a_0 x(t) = b, \quad t \in \mathbb{R}, \quad a_0, \dots, a_{m-1}, b \in \mathbb{R}$

Es gilt: Ist $\hat{x}(t)$ eine beliebige Lösung der homogenen Gleichung und ist $x^p(t)$ eine partikuläre Lösung der inhomogenen Gleichung, so löst auch $x(t) = \hat{x}(t) + x^p(t), \quad t \in \mathbb{R}$ die inhomogene Gleichung.

Ansatz für eine partikuläre Lösung x^p:
Setze eine stationäre Lösung an, d.h.

$x^p(t) = \gamma,\ t \in \mathbb{R}, \quad (x^p)'(t) = 0, \quad (x^p)''(t) = 0, \dots, (x^p)^{(m)}(t) = 0$

Einsetzen liefert: $\underbrace{(x^p)^{(m)}(t)}_{=0} + a_{m-1}\underbrace{(x^p)^{(m-1)}(t)}_{=0} + \ldots + a_1 \underbrace{(x^p)'(t)}_{=0} + a_0 \underbrace{x^p(t)}_{=\gamma} = a_0\gamma \overset{!}{=} b$

$\Rightarrow \quad x^p(t) = \gamma = \dfrac{b}{a_0}$ für $a_0 \neq 0$ ist eine partikuläre Lösung.

Allgemeine Lösung $\mathbb{L} = \left\{ x(t) \;\middle|\; x(t) = x^p(t) + \sum\limits_{i=1}^{m} \alpha_i x^i(t), \quad \alpha_i \in \mathbb{R}, \quad i = 1, \dots, m, \quad t \in \mathbb{R} \right\}$

Dabei sind $x^i(t), \quad i = 1, \dots, m$ die Basislösungen der homogenen Gleichung.

Beispiel: Betrachte die inhomogene AWA zweiter Ordnung

$x''(t) - 3x'(t) - 4x(t) = 10,\ t \in \mathbb{R},\ x(0) = 2,\ x'(0) = 3.$

Dazugehörige homogene Gleichung: $x''(t) - 3x'(t) - 4x(t) = 0$

Ansatz: $x(t) = \exp(\lambda t), \quad t \in \mathbb{R}$

$\Rightarrow \quad \lambda^2 \exp(\lambda t) - 3\lambda \exp(\lambda t) - 4\exp(\lambda t) = \underbrace{(\lambda^2 - 3\lambda - 4)}_{=p(\lambda)} \exp(\lambda t) \overset{!}{=} 0$

$\Rightarrow \quad \lambda_{1,2} = \dfrac{3 \pm \sqrt{9+16}}{2} = \dfrac{3 \pm 5}{2} \quad \Leftrightarrow \quad \lambda_1 = 4,\ \lambda_2 = -1$

Dies liefert die Basislösungen $x^1(t) = \exp(4t), \quad x^2(t) = \exp(-t)$.

Lösung der inhomogenen Gleichung: Ansatz: $x(t) = \gamma$

Einsetzen liefert: $-4\gamma = 10, \quad \gamma = x^p(t) = -\frac{5}{2}$

Allgemeine Lösung: $\mathbb{L} = \left\{ x(t) \;\middle|\; x(t) = \alpha_1 \exp(4t) + \alpha_2 \exp(-t) - \frac{5}{2}, \quad t \in \mathbb{R},\ \alpha_1, \alpha_2 \in \mathbb{R} \right\}$

Bestimmung von α_1, α_2 über die Anfangsbedingung:

$$
\begin{aligned}
t = 0: \quad x(0) &= \alpha_1 x^1(0) + \alpha_2 x^2(0) + x^p(0) = \alpha_1 \cdot 1 + \alpha_2 \cdot 1 - \tfrac{5}{2} = 2 \\
x'(0) &= \alpha_1 (x^1)'(0) + \alpha_2 (x^2)'(0) + x_p'(0) \\
&= \alpha_1 \cdot 4 \cdot \exp(4 \cdot 0) + \alpha_2 (-1) \exp(-1 \cdot 0) + 0 \\
&= 3
\end{aligned}
$$

$$
\Rightarrow \quad \begin{aligned} \alpha_1 + \alpha_2 &= 2 + \tfrac{5}{2} \\ 4\alpha_1 - \alpha_2 &= 3 \end{aligned} \quad \Leftrightarrow \quad \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \frac{9}{2} \\ 3 \end{pmatrix}
$$

Lösung: $\alpha_1 = \frac{3}{2}, \quad \alpha_2 = 3$

Die Lösung der AWA lautet somit $x(t) = \frac{3}{2}\exp(4t) + 3\exp(-t) - \frac{5}{2}, \quad t \in \mathbb{R}$.

Bemerkung: Die Aufgabe $x''(t) - 3x'(t) - 4x(t) = 10, \quad t \in \mathbb{R}, \quad x(0) = 2, \quad x'(0) = 3$ ist gleichwertig zu folgendem zweidimensionalen System:

Setze $y_1(t) = x(t)$, $y_2(t) = x'(t)$ und finde $y_1'(t) = x'(t) = y_2(t)$, $y_2'(t) = x''(t) = 10 + \underbrace{3x'(t)}_{=y_2(t)} + \underbrace{4x(t)}_{y_1(t)} =$ $10 + 3y_2(t) + 4y_1(t)$

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix}(t) = \begin{pmatrix} 0 & 1 \\ 4 & 3 \end{pmatrix} \cdot \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 10 \end{pmatrix}$$

$y_1(0) = x(0) = 2, \quad y_2(0) = x'(0) = 3, \qquad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}(0) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$

Bemerkung: Jede Differenzialgleichung höherer Ordnung lässt sich als ein System 1-ter Ordnung schreiben.

7.3.6 Existenz und Eindeutigkeit von Lösungen

Betrachte die skalare AWA $x'(t) = f(t, x(t)), \ \ x(t_0) = x_0$

Definition:

$f : I \times D \to \mathbb{R}, \quad I, D \subset \mathbb{R}$ Intervalle genügt einer lokalen Lipschitzbedingung, wenn es zu jedem $s_0 \in I, \ y_0 \in D$ eine Umgebung $U = U(s_0, y_0)$ und ein $L = L(s_0, y_0)$ gibt mit

$$|f(s, y) - f(s, \hat{y})| \leq L|y - \bar{y}| \quad \text{für} \quad (s, y), (\bar{s}, \bar{y}) \in U$$

"lokale Lipschitzbedingung für f bzgl. der 2-ten Variablen"

Lemma:

Ist $f : I \times D \to \mathbb{R}$ differenzierbar und ist die Ableitung stetig, so genügt f eine lokalen Lipschitzbedingung.

Satz (lokaler Existenz- und Eindeutigkeitssatz):

Vorgegeben sei $x'(t) = f(t, x(t)), \quad x(t_0) = x\_0, \quad f : I \times D \to \mathbb{R}, \quad t_0 \in I, \ \ x_0 \in D \qquad (***)$
Genügt f einer lokalen Lipschitzbedingung, so ist die AWA $(***)$ lokal eindeutig lösbar, d.h. es gibt $\varepsilon > 0$ und eine eindeutige Lösung $x :]t_0 - \varepsilon, t_0 + \varepsilon[\to D$. Gilt überdies sogar:

$$|f(t, x)| \leq \alpha|x| + \beta, \quad \alpha, \beta \in \mathbb{R},$$

so existiert die Lösung $x(t)$ von $(***)$ für alle $t \in I$.

Bemerkung: Also existiert die Lösung von $x'(t) = a(t)x(t) + b(t)$ für alle $t \in \mathbb{R}$, falls $a(t)$ und $b(t)$ beschränkte Funktionen sind. Ebenso existieren die Lösungen von $x'(t) = Ax(t) + b$ für alle $t \in \mathbb{R}$.