

Mitschrift

Algorithmische Mathematik

Wintersemester 2003/2004

PD Dr. Johannes Schropp

Universität zu Köln

Mitschreiber: Thomas Wittek

Dank für Verbesserungen geht an:

- Sebastian Brocks
- Sebastian Heuel
- André Lang
- Sebastian Schiefer
- Markus Schöllgen

2. Februar 2004

Inhaltsverzeichnis

0.1	Organisatorisches	5
0.1.1	Termine	5
1	Lineare Gleichungssysteme	6
1.1	Lineare Gleichungssysteme (LGS)	6
1.2	Gaußsches Eliminationsverfahren	9
1.2.1	Grundlagen	9
1.2.2	Das Gaußsche Eliminationsverfahren	11
1.2.3	Gaußsche Eliminationsschritte	11
1.3	Die LR-Zerlegung	16
1.3.1	Zeilenvertauschung und Pivotstrategie	20
1.4	Allgemeine LGS, Matrixgleichungen, Bestimmung der Inversen	23
1.4.1	Lösbarkeit des Matrixsystems	24
1.4.2	Matrixgleichung und Inversenbildung	25
1.4.3	Das Gauß-Jordan Verfahren zur Inversenbildung	25
1.5	Kondition linearer Gleichungssysteme	28
1.5.1	Exkurs: Normen	29
1.5.2	Fehler	30
1.5.3	Kurze Wiederholung	33
2	Nichtlineare Gleichungen	34
2.1	Einführung	34
2.2	Das Bisektionsverfahren	36
2.3	Das Newtonverfahren	38
2.4	Iterationsverfahren	40
2.5	Das Sekantenverfahren	47
2.6	Nichtlineare Gleichungssysteme	48
2.6.1	Newtonverfahren zur Lösung von $F(x) = 0$	48
2.6.2	Konvergenz des Newtonverfahrens	50
2.7	Modellierung und Numerik	52
2.7.1	Einführung	52

3	Lineare Optimierung	54
3.1	Problemstellung und Normalformen	54
3.1.1	Durchführbarkeit allgemeiner linearer Optimierungsaufgaben auf einer der Normalformen	58
3.1.2	Zusammenfassung	61
3.2	Das Simplexverfahren	64
3.2.1	Einführung	64
3.2.2	Idee des Verfahrens	64
3.2.3	Beispiel für das Simplexverfahren	65
3.2.4	Formale Darstellung des Simplex-Verfahrens	68
3.2.5	Theoretischer Hintergrund des Simplex-Verfahrens	70
3.2.6	Zusammenfassung	74
3.3	Die Zweiphasenmethode	76
3.3.1	Konstruktion zulässiger Basislösungen	76
3.3.2	Die Zweiphasenmethode	78
3.3.3	Fall der Entartung	81
4	Gewöhnliche Differentialgleichungen	83
4.1	Separation der Variablen	83
4.2	Lineare Systeme von Differentialgleichungen	87
4.2.1	Eigenschaften linearer Differentialgleichungen	89
4.2.2	Kurze Zusammenfassung des Wesentlichen	89
4.3	Existenz und Eindeutigkeit von Lösungen	95
4.4	Numerische Verfahren für Anfangswertaufgaben	97

Organisatorisches

0.1 Organisatorisches

Bei erfolgreicher Teilnahme an den Übungen ($\geq 50\%$ der Punkte) wird ein Bonus von 10% der Gesamtpunktzahl der Scheinklausur gewährt.

Die Übungen werden von Frau Silvia Daun (sdaun@mi.uni-koeln.de) gehalten.

0.1.1 Termine

Sonderübung zur Klausur	Donnerstag, 29.01.2004, ab 18:30Uhr, Großer HS des Mathematischen Instituts
Klausur (Schein)	Dienstag, 03.02.2004, 14:00-16:00Uhr, Großer HS Biologie.
Klausur (Schein, Wiederholung)	Mittwoch, 18.02.2004, 08:45-10:45Uhr, Raumangabe folgt später.
Klausur (FP)	Mittwoch, 18.02.2004, 08:45-12:45Uhr, Raumangabe folgt später.
Klausur (FP, Wiederholung)	Mittwoch, 31.03.2004, 08:45-12:45Uhr, Raumangabe folgt später.

Mitzubringen ist der Studentenausweis

Zugelassene Hilfsmittel:

- Taschenrechner ohne Grafik- und Programmierfunktionen
- Mitschrift von Vorlesung und Übung, eine Formelsammlung

1 Lineare Gleichungssysteme

1.1 Lineare Gleichungssysteme (LGS)

Sei \mathbb{R} der Körper der reellen Zahlen,

\mathbb{R}^N der Vektorraum der als Spalten geschriebenen Vektoren,

$\mathbb{R}^{N,M}$ der Vektorraum der $N \times M$ -Matrizen.

Elemente: $x \in \mathbb{R}^N$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}, x_i \in \mathbb{R}, i = 1, \dots, N$$

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NM} \end{pmatrix}, a_{ij} \in \mathbb{R}; i = 1, \dots, N; j = 1, \dots, M; A \in \mathbb{R}^{N,M}$$

Definition 1.1.1

Für Matrizen $A, B \in \mathbb{R}^{N,M}$ und $\gamma \in \mathbb{R}$ ist

$$\begin{aligned} (A+B)_{i,j} &= a_{i,j} + b_{i,j}; i = 1, \dots, N; j = 1, \dots, M; \\ (\gamma A)_{i,j} &= \gamma a_{i,j}; i = 1, \dots, N; j = 1, \dots, M; \end{aligned}$$

$\mathbb{R}^{N,M}$ mit den Operationen $+$ und \cdot ist ein reeller Vektorraum.

Beispiel 1.1.1

Die Abteilung Logistik einer großen Einzelhandelskette hat die folgende Tabelle der Warenlieferungen (in 1000 €) für das 1. Quartal des Jahres aufgestellt:

Matrixdarstellung:

$$A = \begin{pmatrix} 150 & 267 & 323 & 110 \\ 356 & 56 & 167 & 455 \\ 45 & 143 & 247 & 0 \end{pmatrix} \in \mathbb{R}^{3,4}$$

A ist eine $(3,4)$ -Matrix.

Tabelle 1.1: Warenlieferungen

	Filiale			
	F_1	F_2	F_3	F_4
L_1	150	267	323	110
L_2	356	56	167	455
L_3	45	143	247	0

Der 2. Zeilenvektor beschreibt alle Auslieferungen von L_2 .

Der 2. Spaltenvektor beschreibt alle Auslieferungen an F_2 .

Die Gesamtlieferungen für 1 ganzes Jahr werden beschrieben durch $B = 4 \cdot A$

Definition 1.1.2

Sei $A \in \mathbb{R}^{N,M}$, $x \in \mathbb{R}^M$, so ist

$$Ax = \begin{pmatrix} a_{11} & \dots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \dots & a_{NM} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_M \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^M a_{1j} x_j \\ \vdots \\ \sum_{j=1}^M a_{Nj} x_j \end{pmatrix} \in \mathbb{R}^N$$

Definition 1.1.3

Jede Matrix $A \in \mathbb{R}^{N,M}$ vermittelt eine Abbildung

$$\begin{aligned} f_A : \mathbb{R}^M &\rightarrow \mathbb{R}^N \\ x &\rightarrow f_A(x) = Ax \end{aligned}$$

Diese Abbildung ist linear, d.h. es gilt:

$$f_A(\gamma x + \gamma' x') = \gamma f_A(x) + \gamma' f_A(x')$$

Definition 1.1.4

Ein lineares Gleichungssystem $Ax = b$ heißt homogen, falls $b = 0$. Ansonsten ($b \neq 0$) heißt es inhomogen.

Definition 1.1.5

Zu $A \in \mathbb{R}^{N,M}$ und $B \in \mathbb{R}^{M,l}$ ist $A \cdot B \in \mathbb{R}^{N,l}$ erklärt durch:

$$(AB)_{i,k} = \sum_{j=1}^M a_{ij} b_{jk}, i = 1, \dots, N; k = 1, \dots, l;$$

$$A = \begin{pmatrix} a_{11} & \dots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \dots & a_{NM} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & \dots & b_{1l} \\ \vdots & \ddots & \vdots \\ b_{M1} & \dots & b_{Ml} \end{pmatrix}$$

- Matrixmultiplikation ist nicht kommutativ, d.h. $AB \neq BA$.
- Das Produkt zweier von Null verschiedener Matrizen kann $\vec{0}$ ergeben.

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}$$

Definition 1.1.6 (Bedeutung eines Produktes von Matrizen)

Seien $A \in \mathbb{R}^{N,M}$, $B \in \mathbb{R}^{M,l}$ und $f_A : \mathbb{R}^M \rightarrow \mathbb{R}^N$, $f_B : \mathbb{R}^l \rightarrow \mathbb{R}^M$ die zugehörige Abbildung.

Dann ist $AB \in \mathbb{R}^{N,l}$ und für $f_{AB} : \mathbb{R}^l \rightarrow \mathbb{R}^N$ gilt $f_{AB} = f_A \circ f_B$ d.h. das Matrixprodukt ist das Hintereinanderausführen der zu den Matrixfaktoren gehörenden Abbildung.

Beispiel 1.1.2

In einem Betrieb werden zur Herstellung der Endprodukte E_1, \dots, E_4 fünf Rohstoffe R_1, \dots, R_5 benötigt. Der Verbrauch an Rohstoffen pro Tonne der Endprodukte ist wie folgt:

Es sollen 50t, 30t, 10t, 100t der Endprodukte hergestellt werden.

Tabelle 1.2: Rohstoffverbrauch

	E_1	E_2	E_3	E_4
R_1	2	1	4	1
R_2	0	5	5	2
R_3	1	3	3	2
R_4	4	1	0	0
R_5	3	2	1	1

Matrixdarstellung:

$$\underbrace{\begin{pmatrix} 2 & 1 & 4 & 1 \\ 0 & 5 & 5 & 2 \\ 1 & 3 & 3 & 2 \\ 4 & 1 & 0 & 0 \\ 3 & 2 & 1 & 1 \end{pmatrix}}_{\in \mathbb{R}^{5,4}} \cdot \underbrace{\begin{pmatrix} 50 \\ 30 \\ 10 \\ 100 \end{pmatrix}}_{\in \mathbb{R}^{4,1}} = \begin{pmatrix} 2 \cdot 50 + 1 \cdot 30 + 4 \cdot 10 + 1 \cdot 100 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} = \underbrace{\begin{pmatrix} 270 \\ 400 \\ 370 \\ 250 \\ 320 \end{pmatrix}}_{\in \mathbb{R}^{5,1}}$$

1.2 Gaußsches Eliminationsverfahren

1.2.1 Grundlagen

Untersuche, wie weit man die Matrixmultiplikation „umkehren“ kann.

Bemerkung 1.2.1

Sei $A \in \mathbb{R}^{N,M}$, $B \in \mathbb{R}^{N,k}$.

Gibt es dazu eine Matrix $X \in \mathbb{R}^{M,k}$, so dass $\underbrace{AX}_{\in \mathbb{R}^{N,k}} = \underbrace{B}_{\in \mathbb{R}^{N,k}}$ und wenn ja, wie findet man diese?

Betrachte $k = 1$, d.h. B hat genau 1 Spalte b .

Gesucht ist x , so dass $\underbrace{A}_{\in \mathbb{R}^{N,M}} \cdot \underbrace{x}_{\in \mathbb{R}^M} = \underbrace{b}_{\in \mathbb{R}^N}$.

Ausgeschrieben:

$$\begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1M}x_M & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2M}x_M & = & b_2 \\ \vdots & + & \vdots & + & \vdots & + & \vdots & = & \vdots \\ a_{N1}x_1 & + & a_{N2}x_2 & + & \cdots & + & a_{NM}x_M & = & b_N \end{array}$$

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix}, x = \begin{pmatrix} x_1 \\ \vdots \\ x_M \end{pmatrix}$$

„Lineares Gleichungssystem mit N Gleichungen in M Unbekannten.“

Beispiel 1.2.1

Nach einem Erdbeben in Mittelamerika soll dorthin ein Flugzeug mit Hilfsgütern entsandt werden, wobei die Kapazitäten bzgl. Laderaum, Abfluggewicht und zur Verfügung stehende Geldmittel voll ausgeschöpft werden sollen. Die Maßangaben (l, kg, €) sind pro Behälter zu verstehen.

Wieviele Container sind auf Reise zu schicken, wenn außerdem bekannt ist, dass Trinkwasser am dringendsten benötigt wird und deshalb doppelt so viele Wasserbehälter wie Container mit Blut und Medikamenten insgesamt verwendet werden sollen?

Setze an:

- x_1 = Anzahl von Behältern mit Blut
- x_2 = Anzahl von Behältern mit Medikamenten
- x_3 = Anzahl von Behältern mit Nahrung
- x_4 = Anzahl von Behältern mit Frischwasser

Tabelle 1.3: Beanspruchung

	Vol. (l)	Gew. (kg)	Kosten (€)
Blutkonserven	200	150	1000
Medikamente	300	100	300
Nahrung	80	60	400
Frischwasser	60	70	200
Kapazitäten	60 000	40 000	150 000

$$\begin{array}{rclclcl}
 200x_1 & + & 300x_2 & + & 80x_3 & + & 60x_4 & = & 60\,000 \\
 150x_1 & + & 100x_2 & + & 60x_3 & + & 70x_4 & = & 40\,000 \\
 1\,000x_1 & + & 300x_2 & + & 400x_3 & + & 200x_4 & = & 150\,000 \\
 2x_1 & + & 2x_2 & & & - & x_4 & = & 0
 \end{array}$$

„Lineares Gleichungssystem von 4 Gleichungen in 4 Unbekannten“

Gegeben sei ein LGS

$$Ax = b, A \in \mathbb{R}^{M,N}, x \in \mathbb{R}^N, b \in \mathbb{R}^M$$

Wir setzen $M = N$ voraus. (d.h. genauso viele Gleichungen wie Unbekannte)

$A \in \mathbb{R}^{N,N}$ sei invertierbar, falls es eine Matrix $C \in \mathbb{R}^{N,N}$ gibt mit

$$AC = CA = I \in \mathbb{R}^{N,N}, I = \begin{pmatrix} 1 & \cdots & & 0 \\ & \ddots & & \\ \vdots & & 1 & \vdots \\ & & & \ddots \\ 0 & \cdots & & 1 \end{pmatrix}$$

Man schreibt $C = A^{-1}$.

In diesem Fall gilt

$$\begin{aligned}
 Ax &= b \\
 \Leftrightarrow \underbrace{A^{-1}A}_=I x &= A^{-1}b \\
 \Leftrightarrow x &= A^{-1}b \\
 \Leftrightarrow Ib &= b \\
 \Leftrightarrow b &= b
 \end{aligned}$$

1.2.2 Das Gaußsche Eliminationsverfahren

Grundidee

Formuliere das Ausgangssystem derart um, dass die Lösungsmenge unverändert bleibt und dass das Zielsystem folgende Gestalt hat.

$$\begin{array}{cccccc} r_{11}x_1 & + & r_{12}x_2 & + & \cdots & + & r_{1N}x_N & = & y_1 \\ & & r_{22}x_2 & + & \cdots & + & r_{2N}x_N & = & y_2 \\ & & & & \ddots & + & \vdots & = & \vdots \\ & & & & & & r_{NN}x_N & = & y_N \end{array}$$

In Matrixform lautet dies $Rx = y$ mit

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1N} \\ 0 & r_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{NN} \end{pmatrix} \in \mathbb{R}^{NN} \text{ („rechte obere Dreiecksmatrix“)}$$

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}; x, y \in \mathbb{R}^N$$

Wir finden

$$\begin{aligned} x_N &= \frac{1}{r_{NN}} \cdot y_N \\ x_{N-1} &= \frac{1}{r_{N-1,N-1}} \cdot (y_{N-1} - r_{N-1,N}x_N) \\ &\vdots \\ \text{kurz: } x_i &= \frac{1}{r_{ii}} \left(y_i - \sum_{j=i+1}^N r_{ij}x_j \right); \quad i = N, N-1, \dots, 1 \end{aligned}$$

Dieser Prozess heißt *Rückwärts Auflösen*.

1.2.3 Gaußsche Eliminationsschritte

Um das System $Ax = b$ auf die gewünschte Endform $Rx = y$ zu bringen, ohne die Lösungsmenge zu ändern, stehen uns die folgenden Operationen zur Verfügung:

- *Vertauschung* von Zeilen

- *Multiplikation* einer Zeile mit einem konstanten Faktor $\neq 0$
- *Addition* einer Zeile zu einer anderen Zeile

Transformiere a_{21} zu Null. Setze dabei $a_{11} \neq 0$ voraus. Addiere ein Vielfaches λ der ersten Zeile zur zweiten Zeile und bestimme λ so, dass

$$\lambda a_{11} + a_{21} = 0$$

Dies liefert

$$\lambda = \frac{-a_{21}}{a_{11}}$$

Wir erhalten

$$\begin{aligned} a_{21}^{(1)} &= a_{21} + \lambda a_{11} = a_{21} - \frac{a_{21}}{a_{11}} \cdot a_{11} = 0 \\ a_{2j}^{(1)} &= a_{2j} - \frac{a_{21}}{a_{11}} \cdot a_{1j}; j = 2, \dots, N \\ b_2^{(1)} &= b_2 - \frac{-a_{21}}{a_{11}} \cdot b_1 \end{aligned}$$

Transformiere nun a_{31} zu Null. Wähle $\lambda = \frac{-a_{31}}{a_{11}}$ und finde

$$\begin{aligned} a_{31}^{(1)} &= a_{31} - \frac{a_{31}}{a_{11}} \cdot a_{11} = 0 \\ a_{3j}^{(1)} &= a_{3j} - \frac{a_{31}}{a_{11}} \cdot a_{1j}; j = 2, \dots, N \\ b_3^{(1)} &= b_3 - \frac{a_{31}}{a_{11}} \cdot b_1 \end{aligned}$$

Transformiere allgemein $a_{i1}; i = 2, \dots, N$ zu Null. Wähle $\lambda = \frac{-a_{i1}}{a_{11}}$ und finde:

$$\begin{aligned} a_{i1}^{(1)} &= 0 \\ a_{ij}^{(1)} &= a_{ij} - \frac{a_{i1}}{a_{11}} \cdot a_{1j}; j = 2, \dots, N \\ b_i^{(1)} &= b_i - \frac{a_{i1}}{a_{11}} \cdot b_1; i = 2, \dots, N \end{aligned}$$

$$(A^{(1)}|b^{(1)}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1N} & b_1 \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2N}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{N2}^{(1)} & \cdots & \cdots & a_{NN}^{(1)} & b_N^{(1)} \end{array} \right)$$

$$(A^{(0)}|b^{(0)}) = (A|b) = \left(\begin{array}{cccc|c} 200 & 300 & 80 & 60 & 60\,000 \\ 150 & 100 & 60 & 70 & 40\,000 \\ 2 & 2 & 0 & -1 & 0 \\ 1\,000 & 300 & 400 & 200 & 150\,000 \end{array} \right)$$

$$(A^{(1)}|b^{(1)}) = \left(\begin{array}{cccc|c} 200 & 300 & 80 & 60 & 60\,000 \\ 0 & -125 & 0 & 25 & -5\,000 \\ 0 & -1 & -\frac{4}{5} & -\frac{8}{5} & -600 \\ 0 & -1\,200 & 0 & -100 & -150\,000 \end{array} \right) \quad \begin{array}{l} a_{11} = 200 \neq 0 \\ \lambda = \frac{-150}{200} = -\frac{3}{4} \\ \lambda = \frac{-2}{200} = -\frac{1}{100} \\ \lambda = \frac{1000}{200} = 5 \end{array}$$

Die weiteren Eliminationsschritte laufen entsprechend ab. Die für $(A|b)$ beschriebene Idee wird jetzt auf $(A^{(1)}|b^{(1)})$ angewandt, präziser auf das Restsystem

$$\left(\begin{array}{cccc|c} a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2N}^{(1)} & b_2^{(1)} \\ a_{32}^{(1)} & a_{33}^{(1)} & \cdots & a_{3N}^{(1)} & b_3^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{N2}^{(1)} & a_{N3}^{(1)} & \cdots & a_{NN}^{(1)} & b_N^{(1)} \end{array} \right)$$

Unter der Voraussetzung $a_{22}^{(1)} \neq 0$ werden die Elemente $a_{i2^{(1)}}; i = 3, 4, \dots, N$ zu Null transformiert. Nach $(N-1)$ -Eliminationsschritten erhalten wir die gewünschte Form.

Schematisch:

$$(A|b) = (A^{(0)}|b^{(0)}) \rightarrow (A^{(1)}|b^{(1)}) \rightarrow (A^{(2)}|b^{(2)}) \rightarrow \dots \rightarrow (A^{(N-1)}|b^{(N-1)}) = (R|y)$$

$$\left(\begin{array}{ccc|ccc|c} a_{11}^{(k)} & \cdots & a_{1k}^{(k)} & a_{1,k+1}^{(k)} & \cdots & a_{1N}^{(k)} & b_1^{(k)} \\ & & \ddots & \vdots & & \vdots & \vdots \\ & & & a_{k,k+1}^{(k)} & \cdots & a_{kN}^{(k)} & b_k^{(k)} \\ \hline & & 0 & a_{k+1,k+1}^{(k)} & \cdots & a_{k+1,N}^{(k)} & b_{k+1}^{(k)} \\ & & & \vdots & & \vdots & \vdots \\ & & 0 & a_{N,k+1}^{(k)} & \cdots & a_{NN}^{(k)} & b_N^{(k)} \end{array} \right)$$

Satz 1.2.1 (Formelsatz)

$$\begin{aligned} a_{ij}^{(k)} &= a_{ij}^{(k-1)} - l_{ik} a_{kj}^{(k-1)}; k < i, j \leq N \\ b_i^{(k)} &= b_i^{(k-1)} - l_{ik} b_k^{(k-1)}; k < i \leq N \end{aligned}$$

mit $l_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}; k < i \leq N$.

Dieses Verfahren geht solange gut, wie $a_{kk}^{(k-1)} \neq 0; k = 1, \dots, N-1$.

Die Elemente $a_{kk}^{(k-1)}$, durch welche dividiert wird, heißen *Pivotelemente*.

Beispiel 1.2.2 (Fortführung unseres Beispiels)

$$\begin{aligned}
 (A^{(0)}|b^{(0)}) &= (A|b) = \left(\begin{array}{cccc|c} 200 & 300 & 80 & 60 & 60\,000 \\ 150 & 100 & 60 & 70 & 40\,000 \\ 2 & 2 & 0 & -1 & 0 \\ 1\,000 & 300 & 400 & 200 & 150\,000 \end{array} \right) \\
 (A^{(1)}|b^{(1)}) &= \left(\begin{array}{cccc|c} 200 & 300 & 80 & 60 & 60\,000 \\ 0 & -125 & 0 & 25 & -5\,000 \\ 0 & -1 & -\frac{4}{5} & -\frac{8}{5} & -600 \\ 0 & -1\,200 & 0 & -100 & -150\,000 \end{array} \right) \quad \begin{array}{l} \lambda = \frac{-1}{-125} \\ \lambda = -\frac{-1\,200}{-125} \end{array} \\
 (A^{(2)}|b^{(2)}) &= \left(\begin{array}{cccc|c} 200 & 300 & 80 & 60 & 60\,000 \\ 0 & -125 & 0 & 25 & -5\,000 \\ 0 & 0 & -\frac{4}{5} & -\frac{9}{5} & -560 \\ 0 & 0 & 0 & -340 & -102\,000 \end{array} \right) \\
 &= (A^{(3)}|b^{(3)}) = (R|y)
 \end{aligned}$$

In diesem Fall ist der 3-te Eliminationsschritt nicht mehr notwendig, da das System $(A^{(2)}|b^{(2)})$ bereits die gewünschte Form hat.

Rückwärts auflösen:

$$\begin{aligned}
 -340x_4 &= -102\,000 \Rightarrow x_4 = \frac{-102\,000}{-340} = 300 \\
 -\frac{4}{5}x_3 - \frac{9}{5} \cdot 300 &= -560 \\
 \dots x_2 &= 100 \\
 x_1 &= 50 \\
 &\vdots \\
 x &= (50, 100, 25, 300) \text{ erfüllt } Ax = b
 \end{aligned}$$

Beispiel 1.2.3 (Der Fall $a_{kk}^{(k-1)} = 0; k \in \{1, \dots, N-1\}$)

$$\begin{aligned}
 (\hat{A}^{(0)}|\hat{b}^{(0)}) &= \left(\begin{array}{cccc|c} \mathbf{200} & 300 & 80 & 60 & 60\,000 \\ 150 & 100 & 60 & 70 & 40\,000 \\ 1000 & 300 & 400 & 200 & 150\,000 \\ 2 & 2 & 0 & -1 & 0 \end{array} \right) \\
 (\hat{A}^{(1)}|\hat{b}^{(1)}) &= \left(\begin{array}{cccc|c} 200 & 300 & 80 & 60 & 60\,000 \\ 0 & -125 & 0 & 25 & -5\,000 \\ 0 & 0 & \mathbf{0} & -340 & -102\,000 \\ 0 & 0 & -\frac{4}{5} & -\frac{9}{5} & -560 \end{array} \right)
 \end{aligned}$$

Wir können jetzt die 3-te und die 4-te Zeile vertauschen und finden $\hat{a}_{34}^{(2)} = -\frac{4}{5} \neq 0$. Wir nehmen dieses als neues Pivotelement. Um die Rundungsfehler zu minimieren, sollte das Pivotelement betragslich möglichst groß sein.

Besonders verbreitet ist die Spaltenpivotwahl: Wähle $a_{i_0,k}^{(k-1)}$ als Pivotelement mit

$$\left| a_{i_0,k}^{(k-1)} \right| = \max \left\{ \left| a_{ik}^{(k-1)} \right| ; i = k, k+1, \dots, N \right\} \quad \text{„Spaltenpivotisierung“}$$

Tausche in diesem Fall die i -te mit der k -ten Zeile.

Es gilt:

$$\begin{aligned}
 R &= \tilde{L}A \\
 \Leftrightarrow (\tilde{L})^{-1}R &= \underbrace{(\tilde{L})^{-1}\tilde{L}A}_{=1} \\
 &= \underbrace{A}_{=A} \\
 \Leftrightarrow (\tilde{L})^{-1}R &= A \\
 \Leftrightarrow A &= (\tilde{L})^{-1}R \\
 \stackrel{L=(\tilde{L})^{-1}}{\Leftrightarrow} A &= L \cdot R
 \end{aligned}$$

Wir haben A zerlegt als Produkt zweier Matrizen.

- **R** ist eine rechte obere Dreiecksmatrix
- **L** ist eine linke untere Dreiecksmatrix mit $l_{ii} = 1; i = 1, \dots, N$

Bemerkung 1.3.1

Die Matrizen $\tilde{L}^{(k)}, k = 1, \dots, N$ sind linke untere Dreiecksmatrizen mit $l_{ii}^{(k)} = 1, \dots, N$. Gemäß Übung folgt $(\tilde{L}^{(k)})^{-1}; k = 1, \dots, N$ ist eine linke untere Dreiecksmatrix mit Einsen in der Hauptdiagonalen.

Situation:

$$\begin{aligned}
 L = (\tilde{L})^{-1} &= (\tilde{L}^{(N-1)} \cdot \tilde{L}^{(N-2)} \cdot \dots \cdot \tilde{L}^{(1)})^{-1} \\
 &= (\tilde{L}^{(1)})^{-1} \cdot \dots \cdot (\tilde{L}^{(N-2)})^{-1} \cdot (\tilde{L}^{(N-1)})^{-1}
 \end{aligned}$$

mit linken unteren Dreiecksmatrizen $(\tilde{L}^{(k)})^{-1}$. Also ist auch L eine linke untere Dreiecksmatrix mit $l_{ii} = 1; i = 1, \dots, N$.

Satz 1.3.1 (Dreieckszerlegung)

Sei $A \in \mathbb{R}^{NN}$ invertierbar und

sei die diagonale Pivotwahl beim Gaußschen Eliminationsverfahren möglich,

Dann gilt:

$$A = L \cdot R \text{ mit}$$

$$L = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ l_{21} & 1 & & & \vdots \\ l_{31} & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & 1 & 0 \\ l_{N1} & \cdots & \cdots & l_{N,N-1} & 1 \end{pmatrix} \in \mathbb{R}^{NN}, R = \begin{pmatrix} r_{11} & r_{12} & \cdots & \cdots & r_{1N} \\ 0 & r_{22} & & & \vdots \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & r_{NN} \end{pmatrix} \in \mathbb{R}^{NN}$$

Ferner gilt die Darstellung:

$$r_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} r_{kj}; 1 \leq i \leq j \leq N$$

$$l_{ij} = \frac{1}{r_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} r_{kj} \right); 1 \leq j < i \leq N$$

Bemerkung 1.3.2

r_{ij} und l_{ij} werden abwechselnd beginnend mit r_{ij} gebildet, da sie von einander abhängen!

Beweis 1.3.1 (Zu Satz 1.3.1: Konstruktiv mit dem Verfahren von Crout)

$$A = LR \Leftrightarrow \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ l_{N1} & \cdots & l_{N,N-1} & 1 \end{pmatrix} \cdot \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1N} \\ 0 & r_{22} & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{NN} \end{pmatrix}$$

1. Für die erste Zeile von A gilt:

$$a_{1j} = 1 \cdot r_{1j} = (LR)_{1j}; j = 1, \dots, N \quad \checkmark$$

2. Für den Rest der ersten Spalte von A gilt

$$a_{i1} = l_{i1} \cdot r_{11} = \frac{1}{r_{11}} a_{i1} \cdot r_{11}; i = 2, \dots, N$$

$$= l_{i1} \cdot r_{11} = a_{i1}; i = 2, \dots, N \quad \checkmark$$

3. Für die zweite Restzeile von A gilt

$$a_{2j} = l_{21} \cdot r_{1j} + \underbrace{l_{22}}_{=1} \cdot r_{2j}; i = 2, \dots, N$$

$$\Rightarrow r_{2j} = a_{2j} - l_{21} \cdot r_{1j} \quad \checkmark$$

4. Für die zweite Restspalte gilt

$$a_{i2} = l_{i1} \cdot r_{12} + l_{i2} \cdot r_{22}; i = 3, \dots, N$$

Dies liefert wegen $r_{22} \neq 0$

$$l_{i2} = \frac{1}{r_{22}} (a_{i2} - l_{i1} r_{12}) \quad \checkmark$$

5. Sukzessive erhält man für die i -te Restzeile und j -te Restspalte die obige Darstellung

Beispiel 1.3.1

$$A = \begin{pmatrix} 3 & 1 & 6 \\ 1 & 1 & 1 \\ 2 & 1 & 3 \end{pmatrix} \in \mathbb{R}^{3,3}$$

Die LR-Zerlegung lautet

$$L = \begin{pmatrix} 1 & & \\ \frac{1}{3} & 1 & \\ \frac{2}{3} & \frac{1}{2} & 1 \end{pmatrix} R = \begin{pmatrix} 3 & 1 & 6 \\ & \frac{2}{3} & -1 \\ & & -\frac{1}{2} \end{pmatrix}$$

Nebenrechnungen:

$$\begin{aligned} r_{1i} &= a_{1i}; i = 1, 2, 3; \text{ d.h. } r_{11} = 3, r_{12} = 1, r_{13} = 6 \\ l_{i1} &= \frac{a_{i1}}{r_{11}} = \frac{a_{i1}}{a_{11}}; i = 2, 3; l_{21} = \frac{1}{3}; l_{31} = \frac{2}{3} \\ r_{22} &= a_{22} - l_{21} \cdot r_{12} = 1 - \frac{1}{3} \cdot 1 = \frac{2}{3} \\ r_{23} &= a_{23} - l_{21} \cdot r_{13} = 1 - \frac{1}{3} \cdot 6 = -1 \\ l_{32} &= \frac{1}{r_{22}}(a_{32} - l_{31}r_{12}) = \frac{1}{\frac{2}{3}}(1 - \frac{2}{3} \cdot 1) \\ r_{33} &= a_{33} - l_{31}r_{13} - l_{32}r_{23} = 3 - \frac{2}{3} \cdot 6 - \frac{1}{2}(-1) = -\frac{1}{2} \end{aligned}$$

Bemerkung 1.3.3

Man sagt, die LR-Zerlegung ist eine „verkettete“ Gauß-Elimination.

Satz 1.3.2 (Korollar zu Satz 1.3.1)

Seien die Voraussetzungen von Satz 1.3.1 erfüllt. Dann gilt:

Die r_{ij} (bzw. l_{ij}) von Satz 1.3.1 stimmen mit den Elementen r_{ij} (bzw. l_{ij}) der Matrix $R = A^{(N-1)}$ (bzw. L) der Gauß-Elimination überein.

Auflösung des Gleichungssystems $Ax = b$ bei gegebener Zerlegung $A = LR$.

$$L \underbrace{Rx}_{=y} = Ax = b$$

Setze $Rx = y$ und löse $Ly = b$ sowie $Rx = y$.

Da L eine linke untere Dreiecksmatrix mit $l_{ii} = 1; i = 1, \dots, N$ ist, gilt:

$$(Ly)_i = \sum_{j=1}^{i-1} l_{ij}y_j + y_i = b_i; i = 1, \dots, N$$

$$\Leftrightarrow y_i = b_i - \sum_{j=1}^{i-1} l_{ij} y_j; i = 1, \dots, N \text{ „vorwärtsauflösen“}$$

$$\begin{aligned} Rx &= y \\ (Rx)_i &= y_i; i = 1, \dots, N \\ \sum_{j=1}^N r_{ij} x_j &= r_{ii} x_i + \sum_{j=i+1}^N r_{ij} x_j \\ \Leftrightarrow x_i &= \frac{1}{r_{ii}} \left(y_i - \sum_{j=i+1}^N r_{ij} x_j \right); i = N, N-1, \dots, 1 \text{ „rückwärtsauflösen“} \end{aligned}$$

Sind Gleichungssysteme $Ax^{(i)} = b^{(i)}$ für Vektoren $b^{(1)}, b^{(2)}, \dots, b^{(k)}$ zu lösen, so ist eine LR-Zerlegung von A notwendig und es wird mindestens k -mal vorwärts und rückwärts aufgelöst.

Kurzbeschreibung der Vorgehensweise

1. Ermittlung von R und L anhand von A . b wird hier noch nicht benötigt.
2. Ermittlung von y mittels $Ly = b$ (ab hier brauchen wir b).
3. Ermittlung von x mittels $Rx = y$ (nun brauchen wir das eben ermittelte y und können so die Lösung x berechnen).

Bemerkung 1.3.4 (Laufzeit)

Fast man eine Multiplikation plus eine Addition als eine wesentliche Operation zusammen, so erhält man

- $\frac{1}{3}N^3$ wesentliche Operationen für die LR-Zerlegung
- $\frac{1}{2}N^2$ wesentliche Operationen für einmal vorwärts bzw. rückwärts Auflösen.

1.3.1 Zeilenvertauschung und Pivotstrategie

Wir hatten bisher der Einfachheit halber angenommen, dass die Matrix A eine diagonale Pivotwahl erlaubt. Das trifft nicht immer zu!

Spaltenpivotwahl

Wähle

$$|a_{i_0,k}^{(k-1)}| = \max \left\{ |a_{ik}^{(k-1)}| \mid i = k, k+1, \dots, N \right\}$$

Situation im $(k-1)$ -ten Großschritt der Gaußelimination:

$$(A^{(k-1)} | b^{(k-1)}) = \left(\begin{array}{ccc|ccc|c} a_{11}^{(k-1)} & \cdots & a_{1,k-1}^{(k-1)} & a_{1k}^{(k-1)} & \cdots & a_{1N}^{(k-1)} & b_1^{(k-1)} \\ \vdots & \ddots & \vdots & \vdots & & \vdots & \\ 0 & \cdots & a_{k-1,k-1}^{(k-1)} & a_{k-1,k}^{(k-1)} & \cdots & a_{k-1,N}^{(k-1)} & b_{k-1} \\ \hline & & 0 & a_{kk}^{(k-1)} & \cdots & a_{kN}^{(k-1)} & b_k \\ & & & \vdots & & \vdots & \\ & & & a_{Nk}^{(k-1)} & \cdots & a_{NN}^{(k-1)} & b_N \end{array} \right)$$

$$\boxed{\begin{array}{c} a_{kk}^{(k-1)} \\ \vdots \\ a_{Nk}^{(k-1)} \end{array}} \hat{=} a_{ik}^{(k-1)}; i = k, \dots, N$$

Diagonale Pivotwahl

Wähle $a_{kk}^{(k-1)}$, falls dies $\neq 0$ ist.

Alternativ: *Spaltenpivotwahl*

Ist A invertierbar, so ist $\max \left\{ |a_{ik}^{(k-1)}| \mid i = k, k+1, \dots, N \right\} > 0$. In diesem Fall ist dann die i_0 -te mit der k -ten Zeile zu vertauschen.

Matrizentechnisch bedeutet dies eine Multiplikation mit einer sogenannten *Permutationsmatrix* von links.

$$P = \left(\begin{array}{ccc|cc|c} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ \hline & & & 0 & 1 & \\ & & & & \ddots & \\ & & & 1 & & 0 \\ \hline & & & & & 1 \\ 0 & & & p_{Nk} & p_{Nj} & \\ \hline & & & & & \ddots \\ & & & & & 1 \end{array} \right)$$

Mit dem Einheitsvektor

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{-te Komponente} \in \mathbb{R}^N$$

finden wir $p_{jk} = (l_1, \dots, l_{k-1}, l_j, l_{k+1}, \dots, l_{j-1}, l_k, l_{j+1}, \dots, l_N)$.

Bemerkung 1.3.5 (zur Permutationsmatrix)

Wenn man die Permutationsmatrix P von links heran mit A multipliziert, so entspricht eine Vertauschung der j -ten mit der k -ten Spalte in P einer Vertauschung der j -ten mit der k -ten Zeile in A .

Wenn man die Permutationsmatrix P von rechts heran mit A multipliziert, so entspricht eine Vertauschung der j -ten mit der k -ten Zeile in P einer Vertauschung der j -ten mit der k -ten Spalte in A .

Permutationsmatrizen sind invertierbar. (Beweis Übung)

Spaltenpivotwahl führt also insgesamt zu einer Zerlegung $PA = LR$ mit

P: Permutationsmatrix

A: Zu verändernde Matrix

L: Linke untere Dreiecksmatrix ($l_{ii} = 1; i = 1, \dots, N$)

R: Rechte obere Dreiecksmatrix

1.4 Allgemeine LGS, Matrixgleichungen, Bestimmung der Inversen

Betrachte jetzt:

$$Ax = b; A \in \mathbb{R}^{MN}, x \in \mathbb{R}^N, b \in \mathbb{R}^M, \text{„M Gleichungen in N Unbekannten“}$$

Gaußsche Elimination:

$$(A|b) = (A^{(0)}|b^{(0)}) = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1N} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{M1} & \dots & a_{MN} & b_M \end{array} \right)$$

Auch im Fall M,N beliebig läßt sich die Gaußsche Elimination anpassen.

Wir benutzen die Operationen

- Vertauschen von Zeilen
- Multiplizieren einer Zeile mit einem konstanten Faktor $\neq 0$
- Addition einer Zeile zu einer anderen

Man erhält folgendes Endschema („Zeilenstufenform“)

$$\left(\begin{array}{cccc|ccc} + & * & * & \dots & * & \dots & * & \tilde{b}_1 \\ \vdots & + & * & \dots & * & & * & \vdots \\ \vdots & & \ddots & \ddots & \vdots & & \vdots & \vdots \\ 0 & \dots & \dots & + & * & \dots & * & \tilde{b}_p \\ \hline & & & & & & & \tilde{b}_{p+1} \\ & & & 0 & & & 0 & \vdots \\ & & & & & & & \tilde{b}_M \end{array} \right); + \neq 0, * = \text{beliebig}$$

Man kann diese Matrix wie folgt aufteilen:

Zeilen	$1 - p$:	Quadrant 1 und 2
Zeilen	$p + 1 - M$:	Quadrant 3 und 4
Spalten	$1 - r$:	Quadrant 2 und 3
Spalten	$r + 1 - N$:	Quadrant 1 und 4

1.4.1 Lösbarkeit des Matrixsystems

Das gegebene System ist genau dann lösbar, wenn $\tilde{b}_i = 0; i = p + 1, \dots, M$. Dann haben wir p essentiell verschiedene Gleichungen in N Unbekannten. Es können dann $(N - p)$ Parameter frei gewählt werden.

Da die Umformungen den Rang einer Matrix nicht ändern, finden wir $rg(A) = rg(\text{Zeilenstufenmatrix}) = p$.

Ferner gilt:

Das Gleichungssystem $Ax = b$ ist lösbar genau dann, wenn $rg(A) = rg(A|b)(= p)$.

Dabei bezeichnet der Rang einer Matrix die Anzahl *linear unabhängiger* Spalten oder Zeilen.

Nebeneffekt:

Die Gaußsche Elimination bestimmt den Rang einer Matrix.

Beispiel 1.4.1

$$\begin{aligned}
 (A|b) &= \left(\begin{array}{ccccc|c} 2 & 4 & 4 & -2 & 0 & 6 \\ 4 & 8 & 8 & -4 & 4 & 4 \\ 2 & 4 & 5 & 2 & 1 & 3 \\ 6 & 12 & 12 & -6 & 4 & 10 \end{array} \right); A \in \mathbb{R}^{4,5}, x \in \mathbb{R}^5, b \in \mathbb{R}^4, M = 4, N = 5 \\
 (A^{(1)}|b^{(1)}) &= \left(\begin{array}{ccccc|c} 2 & 4 & 4 & -2 & 0 & 6 \\ 0 & \boxed{0} & \boxed{0} & 0 & 4 & -8 \\ 0 & 0 & 1 & 4 & 1 & -3 \\ 0 & 0 & 0 & 0 & 4 & -8 \end{array} \right) \\
 (A^{(2)}|b^{(2)}) &= \left(\begin{array}{ccccc|c} 2 & 4 & 4 & -2 & 0 & 6 \\ 0 & 0 & 1 & 4 & 1 & -3 \\ 0 & 0 & 0 & 0 & 4 & -8 \\ 0 & 0 & 0 & 0 & 4 & -8 \end{array} \right) \\
 (A^{(3)}|b^{(3)}) &= \left(\begin{array}{ccccc|c} \boxed{2} & 4 & 4 & -2 & 0 & 6 \\ 0 & 0 & \boxed{1} & 4 & 1 & -3 \\ 0 & 0 & 0 & 0 & \boxed{4} & -8 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \text{„Zeilenstufenform“}
 \end{aligned}$$

$\boxed{\mathbb{N}}$ entspricht den mit + gekennzeichneten Elementen

Wir lesen $rg(A) = 3$. „3 Gleichungen in 5 Unbekannten“. Das System ist lösbar, denn $(b^{(3)})_4 = 0$. Es können $(N - P) = (5 - 3) = 2$ Parameter gewählt werden.

Rückwärtsauflösen:

$$\begin{aligned}4x_5 &= -8 \\ &\Rightarrow x_5 = -2 \\ x_3 + 4x_4 + x_5 &= x_3 + 4x_4 - 2 = -3 \\ &\Rightarrow x_3 = -1 - 4x_4 \\ 2x_1 + 4x_2 + 4x_3 - 2x_4 &= 2x_1 + 4x_2 - 4 - 16x_4 - 2x_4 = 6 \\ &\Rightarrow 2x_1 = 10 - 4x_2 + 18x_4 \Rightarrow x_1 = 5 - 2x_2 + 9x_4\end{aligned}$$

2 Parameter (x_2, x_4) sind frei wählbar.

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

1.4.2 Matrixgleichung und Inversenbildung

Matrixgleichung:

$$A \cdot X = B, A \in \mathbb{R}^{M,N}, X \in \mathbb{R}^{N,k}, B \in \mathbb{R}^{M,k}$$

Stellt man die Matrizen

$$X = (x^1, \dots, x^k), B = (b^1, \dots, b^k), x^i \text{ i-te Spalte von } X, b^i \text{ i-te Spalte von } B$$

so ist $AX = B$ äquivalent zu

$$Ax^i = b^i; i = 1, \dots, k$$

Sei nun $M = N$ und sei A invertierbar, dann erhält man A^{-1} als Lösung der Matrixgleichung $A \cdot X = I$.

1.4.3 Das Gauß-Jordan Verfahren zur Inversenbildung

Das *Gauß-Jordan Verfahren* ist eine Variante des Gauß-Verfahrens. Das Zielsystem des Gauß-Jordan Verfahrens ist die Identität (I) .

Betrachte

$$Ax = b; A \in \mathbb{R}^{NN} \text{ invertierbar}; x, b \in \mathbb{R}^N$$

Stelle das Verfahren mit diagonalen Pivotwahl dar. Es werden alle Elemente in einem Großschritt oberhalb und unterhalb des Pivotelementes zu Null transformiert. Ferner wird das Pivotelement auf 1 normiert.

Sequenz:

$$(A|b) = (A^{(0)}|b^{(0)}) \Rightarrow (A^{(1)}|b^{(1)}) \Rightarrow \dots \Rightarrow (A^{(N)}|b^{(N)}) = (I|x)$$

Dabei hat $(A^{(k)}|b^{(k)}); k \in \{1, \dots, N\}$ die Form

$$(A^{(k)}|b^{(k)}) = \left(\begin{array}{ccc|ccc} 1 & & & a_{1,k+1}^{(k)} & \cdots & a_{1,N}^{(k)} & b_1^{(k)} \\ & \ddots & & \vdots & & \vdots & \vdots \\ & & 1 & a_{k,k+1}^{(k)} & \cdots & a_{k,N}^{(k)} & b_k^{(k)} \\ \hline & & & a_{k+1,k+1}^{(k)} & \cdots & a_{k+1,N}^{(k)} & b_{k+1}^{(k)} \\ & & & \vdots & & \vdots & \vdots \\ & & & a_{N,k+1}^{(k)} & \cdots & a_{N,N}^{(k)} & b_N^{(k)} \end{array} \right)$$

Analog zum Gauß-Verfahren lässt sich der Schritt

$$(A^{(k-1)}|b^{(k-1)}) \Rightarrow (A^{(k)}|b^{(k)})$$

durch Multiplikation mit einer Matrix beschreiben:

$$(A^{(k)}|b^{(k)}) = \tilde{L}^{(k)}(A^{(k-1)}|b^{(k-1)})$$

$$\tilde{L}^{(k)} = \left(\begin{array}{ccccccc} 1 & & & -l_{1,k} & & & \\ & \ddots & & \vdots & & & \\ & & 1 & -l_{k-1,k} & & & \\ & & & \frac{1}{a_{kk}^{(k-1)}} & & & \\ & & & -l_{k+1,k} & 1 & & \\ & & & \vdots & & \ddots & \\ & & & -l_{N,k} & & & 1 \end{array} \right); l_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}, i = 1, \dots, k-1, k+1, \dots, N$$

Bemerkung 1.4.1

- *Rechenaufwand: $\approx \frac{1}{2}N^3$ wesentliche Operationen. (vgl. $\approx \frac{1}{3}N^3$ beim Gauß-Verfahren).*
- *Das Verfahren ist besonders interessant zur Berechnung von A^{-1}*

Matrixgleichung $Ax = I; A, x \in \mathbb{R}^{NN}$.

Nach N Schritten folgt $(A^{(N)}, B^{(N)}) = (I, A^{-1})$

Beispiel 1.4.2

$$\begin{aligned}
 A &= \begin{pmatrix} 3 & 1 & 6 \\ 1 & 1 & 1 \\ 2 & 1 & 3 \end{pmatrix} \in \mathbb{R}^{3,3} \\
 (A^{(0)}|B^{(0)}) &= \left(\begin{array}{ccc|ccc} \boxed{3} & 1 & 6 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} \lambda = -\frac{1}{3} \\ \lambda = -\frac{1}{3} \end{array} \\
 (\tilde{A}^{(1)}|\tilde{B}^{(1)}) &= \left(\begin{array}{ccc|ccc} 3 & 1 & 6 & 1 & 0 & 0 \\ 0 & \frac{2}{3} & -1 & -\frac{1}{3} & 1 & 0 \\ 0 & \frac{1}{3} & -1 & -\frac{2}{3} & 0 & 1 \end{array} \right) \cdot \frac{1}{3} \\
 (A^{(1)}|B^{(1)}) &= \left(\begin{array}{ccc|ccc} 1 & \frac{1}{3} & 2 & \frac{1}{3} & 0 & 0 \\ 0 & \boxed{\frac{2}{3}} & -1 & -\frac{1}{3} & 1 & 0 \\ 0 & \frac{1}{3} & -1 & -\frac{2}{3} & 0 & 1 \end{array} \right) \begin{array}{l} \lambda = -\frac{1}{2} \\ \lambda = -\frac{1}{2} \end{array} \\
 (\tilde{A}^{(2)}|\tilde{B}^{(2)}) &= \left(\begin{array}{ccc|ccc} 1 & 0 & \frac{5}{2} & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & \frac{2}{3} & -1 & -\frac{1}{3} & 1 & 0 \\ 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 1 \end{array} \right) \cdot \frac{3}{2} \\
 (A^{(2)}|B^{(2)}) &= \left(\begin{array}{ccc|ccc} 1 & 0 & \frac{5}{2} & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{3}{2} & -\frac{1}{2} & \frac{3}{2} & 0 \\ 0 & 0 & \boxed{-\frac{1}{2}} & -\frac{1}{2} & -\frac{1}{2} & 1 \end{array} \right) \begin{array}{l} \lambda = 5 \\ \lambda = -3 \end{array} \\
 (\tilde{A}^{(3)}|\tilde{B}^{(3)}) &= \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & -3 & 5 \\ 0 & 1 & 0 & 1 & 3 & -3 \\ 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 1 \end{array} \right) \cdot -2 \\
 (A^{(3)}|B^{(3)}) &= \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & -3 & 5 \\ 0 & 1 & 0 & 1 & 3 & -3 \\ 0 & 0 & 1 & 1 & 1 & -2 \end{array} \right) \\
 \Rightarrow A^{-1} &= \begin{pmatrix} -2 & -3 & 5 \\ 1 & 3 & -3 \\ 1 & 1 & -2 \end{pmatrix}
 \end{aligned}$$

Probe:

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

1.5 Kondition linearer Gleichungssysteme

Fehleranalyse der Gauß-Elimination bei Darstellung auf dem Rechner. Betrachte

$$Ax = b; A \in \mathbb{R}^{N \times N}, x, b \in \mathbb{R}^N, A \text{ invertierbar}$$

Eindeutige Lösung: $x = A^{-1} \cdot b$ mit Gauß-Elimination berechenbar.

Bei der Algorithmischen Umsetzung treten jedoch Fehlerquellen auf:

- Jede reelle Zahl muss auf die nächstmögliche darstellbare Zahl gerundet werden.

$$rd : \mathbb{R} \rightarrow g, a \rightarrow rd(a) \text{ („gerundete Gitterzahl“)}$$

g bezeichnet die Menge darstellbarer Zahlen.

- Die Grundoperationen $\{+, -, \cdot, : \}$ werden auf dem Rechner nur auf der Menge g , also *approximiert*, ausgeführt.

Weitere Fehlerquellen: Die Matrix A und/oder die rechte Seite sind mit Messfehlern behaftet.

Beispiel 1.5.1

Betrachte das parameterabhängige lineare Gleichungssystem $A(\epsilon) \cdot x(\epsilon) = b(\epsilon)$.

$$A(\epsilon) = \begin{pmatrix} 1 & 1 \\ 1 & 1 - \epsilon \end{pmatrix}, b(\epsilon) = \begin{pmatrix} 4 \\ 4 - \epsilon \end{pmatrix}; \epsilon > 0 \text{ (jedoch sehr klein)}$$

Lösung:

$$x = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

Betrachte nun eine kleine Störung auf der rechten Seite:

$$\tilde{b}(\epsilon) = \begin{pmatrix} 4 + \epsilon \\ 4 - 2\epsilon \end{pmatrix}$$

$$A(\epsilon) \cdot x(\epsilon) = \tilde{b}(\epsilon) \text{ hat die Lösung } \tilde{x} = \begin{pmatrix} 1 + \epsilon \\ 3 \end{pmatrix}.$$

Offenbar bewirkt der Fehler $b(\epsilon) - \tilde{b}(\epsilon) = \begin{pmatrix} -\epsilon \\ \epsilon \end{pmatrix}$, d.h. $\|b(\epsilon) - \tilde{b}(\epsilon)\|_\infty = \epsilon$ in der rechten Seite die riesige Abweichung

$$x(\epsilon) - \tilde{x}(\epsilon) = \begin{pmatrix} 2 - \epsilon \\ 2 \end{pmatrix}; \|x(\epsilon) - \tilde{x}(\epsilon)\|_\infty = 2$$

Man nennt solche Gleichungssysteme schlecht konditioniert.

1.5.1 Exkurs: Normen

Definition 1.5.1 (Norm)

Sei V ein Vektorraum. Eine Abbildung: $\|\cdot\| : V \rightarrow \mathbb{R}_{\geq 0}$ mit

1. $\|v\| \geq 0 \forall v \in V; \|v\| = 0$ genau dann, wenn $v = 0$
2. $\|\lambda v\| = |\lambda| \cdot \|v\|; \lambda \in \mathbb{R}, v \in V$
3. $\|u + v\| \leq \|u\| + \|v\|; u, v \in V$ („Dreiecksungleichung“)

Normen dienen der Abstandsmessung.

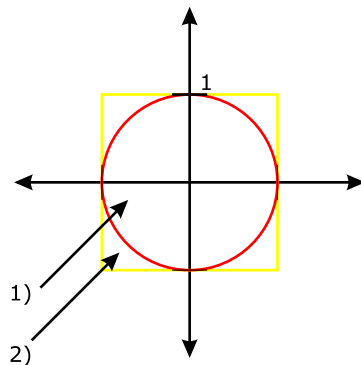
Gängige Normen auf $V = \mathbb{R}^N$:

- Euklidische Norm (2-Norm): $\|u\|_2 = (\sum_{i=1}^N u_i^2)^{\frac{1}{2}}$
Beispiel: $u = (1, 3, -7), \|u\|_2 = \sqrt{1^2 + 3^2 + (-7)^2} = \sqrt{59}$
Diese Norm entspricht der Abstandsmessung im kartesischen Koordinatensystem.
- Maximumnorm (∞ -Norm): $\|u\|_\infty = \max\{|u_i| \mid i = 1, \dots, N\}$
Beispiel: $u = (1, 3, -7), \|u\|_\infty = 7$

Betrachte $V = \mathbb{R}^2$ und

1. $\{u \in \mathbb{R}^2 \mid \|u\|_2 = 1\} \Rightarrow \|u\|_2 = 1 \Leftrightarrow \sqrt{u_1^2 + u_2^2} = 1 \Leftrightarrow u_1^2 + u_2^2 = 1$
2. $\{u \in \mathbb{R}^2 \mid \|u\|_\infty = 1\} \Rightarrow \|u\|_\infty = 1 \Leftrightarrow \max\{|u_1|, |u_2|\} = 1$

Abbildung 1.1: Grafische Darstellung der Norm



Wobei 1) die Euklidische und 2) die Maximumnorm darstellt.

Jede Norm $\|\cdot\|$ auf $V = \mathbb{R}^{NN}$ heißt *Matrixnorm*. Von besonderem Interesse sind Matrixnormen, die zu einer Vektornorm passen, d.h. es gilt

$$\|Au\| \leq \|A\| \cdot \|u\|; u \in \mathbb{R}^N, A \in \mathbb{R}^{NN}$$

Definition 1.5.2 (Natürliches Konstruktionsprinzip für derartige Matrixnormen)

Die zu einer gegebenen Vektornorm $\|\cdot\|$ zugehörige Matrixnorm $\|A\|$ ist gegeben durch:

$$\|A\| = \sup \left\{ \frac{\|Au\|}{\|u\|} \mid u \in \mathbb{R}^N \setminus \{0\} \right\}$$

Schreibe nun $\|A\|$ anstatt $\|A\|$. Man findet

$$\begin{aligned} \|A\|_\infty &= \sup \left\{ \frac{\|Au\|_\infty}{\|u\|_\infty} \mid u \in \mathbb{R}^N \setminus \{0\} \right\} \\ &= \max \left\{ \sum_{j=1}^N |a_{ij}| \mid i = 1, \dots, N \right\} \text{ „Zeilensummennorm“} \end{aligned}$$

Beispiel:

$$A = \begin{pmatrix} 1 & 2 \\ 5 & 7 \end{pmatrix} \quad \|A\|_\infty = \max \{ |1| + |2|, |5| + |7| \} = 12$$

Ferner gilt $\|A\|_2 = \sqrt{\max \{ |\lambda| \mid \lambda = \text{Eigenwert von } A^T A \}}$ „Spaltennorm“.

Bemerkung 1.5.1

A^T ist die transponierte Matrix A . Dies entspricht einer Vertauschung der Zeilen mit den Spalten.

1.5.2 Fehler

Betrachte

$$Ax = b; A \in \mathbb{R}^{N \times N}, b \in \mathbb{R}^N \text{ mit Lösung } x \in \mathbb{R}^N$$

Seien $\tilde{A} = A + \Delta A, \tilde{b} = b + \Delta b$ kleine Störungen und sei \tilde{x} die Lösung von $\tilde{A}\tilde{x} = \tilde{b}$.

Wir interessieren uns für den Fehler $\Delta x_i = \tilde{x} - x$. Man nennt $\|\Delta x\| = \|\tilde{x} - x\|$ den *absoluten Fehler* und $\frac{\|\tilde{x} - x\|}{\|x\|}$ den *relativen Fehler*.

Gesucht sind Abschätzungen für den absoluten und relativen Fehler.

Fall 1 - Störung nur in b

Störung nur in b , d.h. $\Delta A = 0$.

$$\begin{aligned} A\Delta x &= A\tilde{x} - Ax = \tilde{b} - b = \Delta b \Rightarrow \Delta x = A^{-1}\Delta b \\ \|\Delta x\| &= \underbrace{\|A^{-1}\Delta b\|}_{\text{Vektornorm}} \leq \underbrace{\|A^{-1}\|}_{\text{Matrixnorm}} \cdot \underbrace{\|\Delta b\|}_{\text{Vektornorm}} \text{ „absoluter Fehler“} \end{aligned}$$

Mit $b = Ax$ finden wir

$$\begin{aligned} \|\Delta x\| &\leq \|A^{-1}\| \cdot \|\Delta b\| \cdot \frac{\|Ax\|}{\|b\|} \\ &\leq \|A^{-1}\| \cdot \frac{\|\Delta b\|}{\|b\|} \cdot \|A\| \cdot \|x\| \\ \Rightarrow \frac{\|\Delta x\|}{\|x\|} &\leq \|A^{-1}\| \cdot \|A\| \cdot \frac{\|\Delta b\|}{\|b\|} \text{ „relativer Fehler“} \end{aligned}$$

Die Größen $\|A^{-1}\|$ bzw. $\|A\| \cdot \|A^{-1}\|$ spielen eine entscheidende Rolle.

Definition 1.5.3 (Konditionszahl)

Sei $A \in \mathbb{R}^{NN}$ invertierbar. Dann heißt

$$K(A) = \|A\| \cdot \|A^{-1}\|$$

die Konditionszahl von A .

Es gilt immer:

$$1 = \|I\| = \|A \cdot A^{-1}\| \leq \|A\| \cdot \|A^{-1}\| = K(A)$$

Ist die Konditionszahl $K(A)$ in der Größenordnung von 1 ($\gg 1$), so heißt A gut (schlecht) konditioniert.

Fall 2 - Störung von A und b

Definition 1.5.4 (Kern einer Matrix)

Sei $C \in \mathbb{R}^{NN}$.

$$\begin{aligned} \text{Ker}(C) &= \{u \in \mathbb{R}^N \mid Cu = 0\} \text{ „Kern von } C\text{“} \\ \text{Bi}(C) &= \{v \in \mathbb{R}^N \mid \exists u \in \mathbb{R}^N, Cu = v\} \text{ „Bild von } C\text{“} \\ &0 \in \text{Bi}(C) \end{aligned}$$

Es gilt:

$\text{Bi}(C)$ und $\text{Ker}(C)$ sind Unterräume von \mathbb{R}^N .

Faktum:

Sei $C \in \mathbb{R}^{NN}$, so gilt gleichzeitig:

- C ist invertierbar
- $\text{Ker}(C) = \{0\}$
- $\text{Bi}(C) = \mathbb{R}^N$

Satz 1.5.1 (Konditionszahlenschätzung)

Betrachte $Ax = b$ und das gestörte $\tilde{A}\tilde{x} = \tilde{b}$.

Es gelte:

$$\|\Delta A\| \cdot \|A^{-1}\| < 1$$

Dann folgt:

$$\|\Delta x\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|} (\|\Delta b\| + \|A^{-1}\| \cdot \|\Delta A\| \cdot \|b\|) \text{ „absoluter Fehler“}$$

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{K(A)}{1 - K(A) \frac{\|\Delta A\|}{\|A\|}} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|} \right); b \neq 0 \text{ „relativer Fehler“}$$

Beweis 1.5.1 (Zu Satz 1.5.1)

Zeige zuerst: $\tilde{A} = A + \Delta A$ ist invertierbar.

Sei nun $u \in \text{Ker}(A + \Delta A)$.

$$\begin{aligned} (A + \Delta A)u &= 0 \\ \Rightarrow \Delta Au &= -Au \\ \Rightarrow -u &= A^{-1} \cdot \Delta Au \end{aligned}$$

$$\|u\| = \|-u\| = \|A^{-1}\Delta Au\| \leq \underbrace{\|A^{-1}\| \cdot \|\Delta A\|}_{<1} \cdot \|u\|$$

$$\underbrace{(1 - \|A^{-1}\| \cdot \|A\|)}_{>0} \cdot \|u\| = 0 \Rightarrow \|u\| = 0 \Rightarrow u = 0 \Rightarrow$$

$$\Rightarrow \text{Ker}(A + \Delta A) = \{0\} \Rightarrow A + \Delta A = \tilde{A} \text{ ist invertierbar.}$$

Aus $\|\Delta A\| \cdot \|A^{-1}\| < 1$ folgt, dass \tilde{A} ebenfalls invertierbar ist.

$$\begin{aligned} (A + \Delta A)(x + \Delta x) &= b + \Delta b \\ \cancel{Ax} + A\Delta x + \Delta Ax + \Delta A \cdot \Delta x &= \cancel{b} + \Delta b \\ A\Delta x + \Delta Ax + \Delta A \cdot \Delta x &= \Delta b \\ A\Delta x &= \Delta b - \Delta Ax - \Delta A \cdot \Delta x \quad | \cdot A^{-1} \\ \Delta x &= A^{-1}\Delta b - A^{-1}\Delta Ax - A^{-1}\Delta A \cdot \Delta x \end{aligned}$$

$$\Rightarrow \|\Delta x\| \leq \|A^{-1}\| \cdot \|\Delta b\| + \|A^{-1}\| \cdot \|\Delta A\| \cdot \|x\| + \|A^{-1}\| \cdot \|\Delta A\| \cdot \|\Delta x\|$$

$$\underbrace{(1 - \underbrace{\|A^{-1}\| \cdot \|\Delta A\|}_{<0})}_{>0} \cdot \|\Delta x\| \leq \|A^{-1}\| (\|\Delta b\| + \|\Delta A\| \cdot \|x\|)$$

$$\|\Delta x\| \leq \frac{1}{1 - \|A^{-1}\| \cdot \|\Delta A\|} \cdot \|A^{-1}\| (\|\Delta b\| + \|\Delta A\| \cdot \|x\|)$$

Es gilt:

$$\|x\| = \|A^{-1}b\| \leq \|A^{-1}\| \cdot \|b\|$$

Einsetzen liefert:

$$\|\Delta x\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|} (\|\Delta b\| + \|\Delta A\| \cdot \|A^{-1}\| \cdot \|b\|)$$

Analog zeigt man auch die Abschätzung für den relativen Fehler.

1.5.3 Kurze Wiederholung

Ungestörtes System : $Ax = b$, *Ainvertierbar*

Gestörtes System : $\tilde{A}\tilde{x} = \tilde{b}$, $\tilde{A} = A + \Delta A$, $\tilde{b} = b + \Delta b$, $\tilde{x} = x + \Delta x$

Absoluter Fehler : $\|x - \tilde{x}\| = \|\Delta x\|$

Relativer Fehler : $\frac{\|x - \tilde{x}\|}{\|x\|}$

Wichtige Größe $K(A) = \|A\| \cdot \|A^{-1}\|$ „Konditionszahl von A“.

2 Nichtlineare Gleichungen

2.1 Einführung

Ziel ist die näherungsweise Berechnung einer Nullstelle von $f \in C([a, b])$, d.h. gesucht ist ein Wert $\xi \in [a, b]$; $f(\xi) = 0$.

Beispiel 2.1.1 (Effektivzins bei prämienbegünstigten Sparverträgen)

Der Sparer zahlt zu Beginn des Jahres jeweils einen gleichbleibenden festen Betrag auf ein Sparbuch ein. Dieser wird dann am Ende des Jahres mit dem Zinssatz $p\%$ verzinst und gutgeschrieben. Nach Ablauf von n Jahren (z.B. $n = 7$) zahlt das Kreditinstitut dem Sparer auf die Beiträge eine Prämie von $r\%$ und der Kunde kann über das gesamte Kapital verfügen.

Gesucht ist die Effektiv-Verzinsung, d.h. welcher Zinssatz q ist anzusetzen, damit der Kunde das gleiche Kapital ausbezahlt bekommt, wenn keine Prämie gewählt wird.

Lösung:

Unter Berücksichtigung von Zins und Zinseszins:

$$\begin{array}{ll} R(1+p) & \text{(nach 1 Jahr)} \\ R(1+p)^2 & \text{(nach 2 Jahren)} \\ & \vdots \\ R(1+p)^i & \text{(nach i Jahren)} \end{array}$$

Für unser Beispiel gilt:

$$\begin{array}{ll} \text{Einzahlung } R \text{ zu Beginn des} & \text{1. Jahres: } R(1+p)^n \\ & \text{2. Jahres: } R(1+p)^{n-1} \\ & \vdots \\ & \vdots \\ & \text{n. Jahres: } R(1+p)^1 \end{array}$$

Sparbeiträge und Zinsen:

$$S = R \left[\sum_{i=1}^n (1+p)^i \right] = R \left[\frac{(1+p)^{n+1} - (1+p)}{1+p-1} \right] = R \left[\frac{(1+p)^{n+1} - (1+p)}{p} \right]$$

Prämie auf das eingezahlte Kapital:

$$P = Rr \cdot n$$

Gesamtes Kapital:

$$K = S + P = R \left(\frac{(1+p)^{n+1} - (1+p)}{p} + r \cdot n \right)$$

Sei q der Effektivzins. Dann muss gelten:

$$K = R \cdot \sum_{i=1}^n (1+q)^i = R \cdot \frac{1}{q} ((1+q)^{n+1} - (1+q))$$

Zu lösen ist

$$\begin{aligned} f(q) &= \frac{R}{q} ((1+q)^{n+1} - (1+q)) - K \\ &= \frac{R}{q} ((1+q)^{n+1} - (1+q)) - R \left(\frac{1}{p} (1+p)^{n+1} - (1+p) + rn \right) \\ &= 0 \end{aligned}$$

q ist die Unbekannte, während R, n und p bzw. K gegeben sind.

Bemerkung 2.1.1

q ist von R unabhängig.

2.2 Das Bisektionsverfahren

Definition 2.2.1 (Zwischenwertsatz (ZWS))

Sei $f \in C([a, b], \mathbb{R})$. Dann gibt es für jedes $c \in [\min(f(a), f(b)), \max(f(a), f(b))]$ ein $x \in [a, b]$ mit $f(x) = c$.

Im Falle $f(a) \cdot f(b) < 0$ sichert der ZWS die Existenz einer Nullstelle, denn es gilt

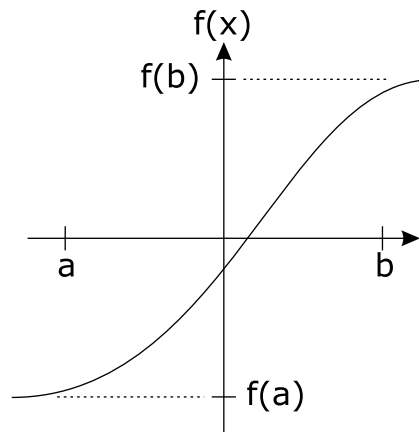
$$f(a) < 0, f(b) > 0 \text{ oder } f(a) > 0, f(b) < 0$$

Das liefert $\min(f(a), f(b)) < 0 < \max(f(a), f(b))$, d.h. 0 ist als c zugelassen.

\Rightarrow Es gibt ein $x \in [a, b]$ mit $f(x) = 0$.

Sei $f \in C([a, b])$ gegeben. Ferner gelte $f(a) \cdot f(b) < 0$

Abbildung 2.1: Funktion f



Nach dem Zwischenwertsatz existiert dann ein $\xi \in [a, b]$ mit $f(\xi) = 0$.

Definition 2.2.2 (Bisektionsverfahren)

Versuche durch Intervallhalbierung eine Nullstelle zu lokalisieren.

Erster Schritt:

$$s := \frac{1}{2}(a + b)$$

3 Fälle können auftreten:

1. $f(s) = 0$: Dann ist s eine gesuchte Nullstelle von f .

2. $f(a) \cdot f(s) < 0$: Dann liegt nach dem Zwischenwertsatz eine Nullstelle ξ in $]a, s[=]a, \frac{a+b}{2}[$.
3. $f(s) \cdot f(b) < 0$: Dann liegt nach dem Zwischenwertsatz eine Nullstelle ξ in $]s, b[=]\frac{a+b}{2}, b[$.

Wir haben dadurch das Intervall halbiert, indem eine Nullstelle liegt. Wählt man ein u_1 aus diesem Intervall als Approximation für ξ , so gilt für den Fehler

$$|\xi - u_1| \leq \frac{b - a}{2}$$

wegen 1)–3) lässt sich die Methode iterativ anwenden. Wählt man nach dem n -ten Schritt ein u_n aus dem verbleibenden Intervall als Näherung für ξ , so gilt

$$|\xi - u_n| \leq \frac{b - a}{2^n}$$

\Rightarrow Durch fortgeführte Intervallhalbierung kann eine Nullstelle von f beliebig genau berechnet werden.

Definition 2.2.3 (Algorithmus für das Bisektionsverfahren)

1. Eingabe von f, a, b sowie zwei beliebige Zahlen $\epsilon > 0, \eta > 0$. (z.B. $\epsilon = \eta = 10^{-10}$)
2. Setze $x_0 = a, y_0 = b, n = 0$.
3. Berechne $s = \frac{1}{2}(x_n + y_n)$ und $f(s)$.
Falls $|f(s)| < \eta$, so setze $z = s$ und beende das Verfahren.

Andernfalls setze

$$(x_{n+1}, y_{n+1}) = \begin{cases} (x_n, s) & \text{falls } f(x_n) \cdot f(s) < 0 \\ (s, y_n) & \text{falls } f(y_n) \cdot f(s) < 0 \end{cases}$$

Ist $|x_{n+1} - y_{n+1}| < \epsilon$, so setze $z = x_{n+1}$ und beende das Verfahren.

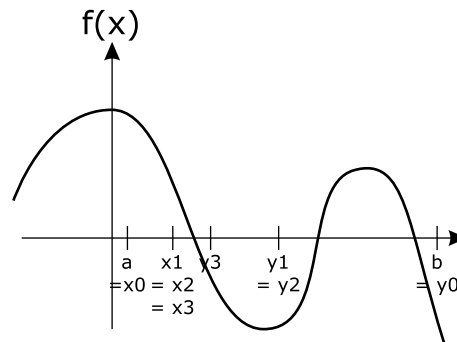
Ist dies nicht der Fall, so erhöhe n und wiederhole Schritt 3.

4. z wird als Näherungswert für eine Nullstelle ξ akzeptiert und ausgegeben.

Für z gilt:

$$|f(z)| < \eta \text{ oder } |z - \xi| < \epsilon$$

Abbildung 2.2: Grafische Darstellung des Algorithmus



2.3 Das Newtonverfahren

Es sei $f \in C^2[a, b]$ und es existiert ein $\xi \in [a, b]$ mit $f(\xi) = 0$. (z.B. da $f(a) \cdot f(b) < 0$)

Ferner gelte $f'(\xi) \neq 0$.

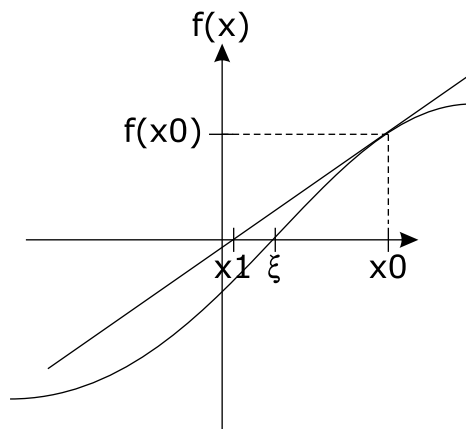
Eine solche Nullstelle (NST) heißt „reguläre Nullstelle“.

Es sei $x_0 \in [a, b]$ eine Startnäherung für ξ .

Idee:

Konstruiere die Tangente an die Kurve $y = f(x)$ im Punkt $(x_0, y_0) = (x_0, f(x_0))$. Der Schnittpunkt der Tangente mit der x-Achse ist der neue Näherungswert x_1 .

Abbildung 2.3: Geometrische Vorstellung des Newtonverfahrens



Explizite Berechnung der Geradengleichung der Tangente:

$$y = mx + b; \quad m = f'(x_0), b = f(x_0) - f'(x_0)x_0$$
$$\Rightarrow y = f'(x_0) \cdot x + f(x_0) - f'(x_0)x_0$$

Als Nullstelle der Tangente ($y = 0$) finden wir

$$0 = f'(x_0) \cdot x + f(x_0) - f'(x_0)x_0$$
$$f'(x_0)x_0 - f(x_0) = f'(x_0) \cdot x$$
$$x_0 - \frac{f(x_0)}{f'(x_0)} = x = x_1, \text{ falls } f'(x_0) \neq 0$$

Definition 2.3.1 (Newtonverfahren)

Das Newtonverfahren dient zur Lösung von $f(x) = 0$.

Wähle $x_0 \in [a, b]$ und setze

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \dots \text{ „Newtonverfahren“}$$

Das Newtonverfahren ist ein Iterationsverfahren (siehe 2.4).

Fragen zur Praxis:

- Ist die Iteration überhaupt durchführbar?
- Konvergiert die Iteration gegen eine Nullstelle von f ?
- Wie wählt man den Startwert x_0 geeignet?

Bemerkung 2.3.1

Ist die Berechnung von $f'(x_n)$ zu aufwändig, so iteriert man gemäß

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}, \quad n = 0, 1, \dots \text{ „Vereinfachtes Newtonverfahren“}$$

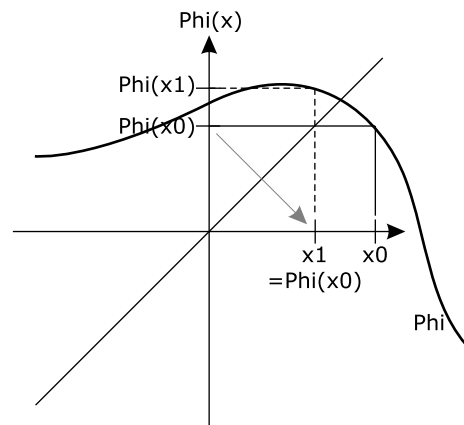
2.4 Iterationsverfahren

Allgemein nennt man ein Verfahren, bei dem ein Startwert x vorgegeben wird, und bei dem man x_{n+1} aus x_n berechnet, d.h. $x_{n+1} = \Phi(x_n)$ mit einer gewissen Funktion Φ , ein *Iterationsverfahren* (wie z.B. das *Newtonverfahren*).

Betrachte

$$x_{n+1} = \Phi(x_n), \quad \Phi : D \rightarrow D, \quad D \subset \mathbb{R}, \quad x_0 \in D$$

Abbildung 2.4: Geometrische Veranschaulichung von $x_{n+1} = \Phi(x_n)$, $n = 0, 1, 2, \dots, x_0$



Für das Newtonverfahren gilt:

$$\Phi(x) = x - \frac{f(x)}{f'(x)} \quad \text{für ein } f \in C^2(D) \text{ „Newton-Operation“}$$

Definition 2.4.1

Ein $y \in D$ heißt *Fixpunkt* von Φ , falls gilt:

$$y = \Phi(y)$$

Im Falle des Newtonverfahrens entsprechen die Fixpunkte von Φ gerade den Nullstellen von f , denn

$$\begin{aligned} \Phi(\xi) &= \xi - \frac{f(\xi)}{f'(\xi)} = \xi \\ &\Leftrightarrow \\ 0 &= \frac{f(\xi)}{f'(\xi)} \\ &\Leftrightarrow \\ f(\xi) &= 0 \quad \wedge \quad f'(\xi) \neq 0 \end{aligned}$$

D.h. die Fixpunkte entsprechen gerade den regulären Nullstellen von f .

Satz 2.4.1

Ist die Folge $(x_n)_{n \in \mathbb{N}}$ gemäß $x_{n+1} = \Phi(x_n)$ bildbar, und gilt $\lim_{n \rightarrow \infty} x_n = \xi$, so folgt $\xi = \Phi(\xi)$.

Beweis 2.4.1 (zu Satz 2.4.1)

$$\xi = \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} \Phi(x_n) \underbrace{=}_{\Phi \text{ stetig}} \Phi(\lim_{n \rightarrow \infty} x_n) = \Phi(\xi)$$

Satz 2.4.2 (Banachscher Fixpunktsatz)

Es sei Φ eine stetige Funktion auf $[a, b]$, welche $[a, b]$ in sich abbildet, d.h. $\Phi(x) \in [a, b]$ für $x \in [a, b]$. Ferner genüge Φ einer Lipschitz-Bedingung, d.h.

$$|\Phi(x) - \Phi(y)| \leq L \cdot |x - y|; \quad x, y \in [a, b]; \quad L < 1 \quad (L \text{ heißt Lipschitz-Konstante})$$

Die Lipschitz-Bedingung besagt, dass sich die zwei Punkte x und y durch die Transformation mit Φ um mindestens den Faktor L näher kommen.

Dann ist durch $x_{n+1} = \Phi(x_n)$, $x_0 \in [a, b]$ eine Folge $(x_n)_{n \in \mathbb{N}}$ in $[a, b]$ erklärt, und es gilt $\lim_{n \rightarrow \infty} x_n = \xi$ mit $\xi = \Phi(\xi)$.

Außerdem besitzt Φ keinen weiteren Fixpunkt in $[a, b]$. Es gilt die Fehlerabschätzung

$$|\xi - x_n| \leq \frac{L^n}{1 - L} |x_1 - x_0|, \quad n = 0, 1, 2, \dots$$

Beweis 2.4.2 (zu Satz 2.4.2)

1. $(x_n)_{n \in \mathbb{N}}$ ist erklärt, da $\Phi(x) \in [a, b]$ für $x \in [a, b]$.
2. $(x_n)_{n \in \mathbb{N}}$ konvergiert gegen ein $\xi \in [a, b]$ mit $\Phi(\xi) = \xi$.

$$\begin{aligned} |x_{n-1} - x_n| &= |\Phi(x_n) - \Phi(x_{n-1})| \leq L \cdot |x_n - x_{n-1}| \\ &= L \cdot |\Phi(x_{n-1}) - \Phi(x_{n-2})| \leq L^2 |x_{n-1} - x_{n-2}| \leq \dots \leq L^n |x_1 - x_0| \end{aligned}$$

Sei nun $m, n \in \mathbb{N}, m > n$:

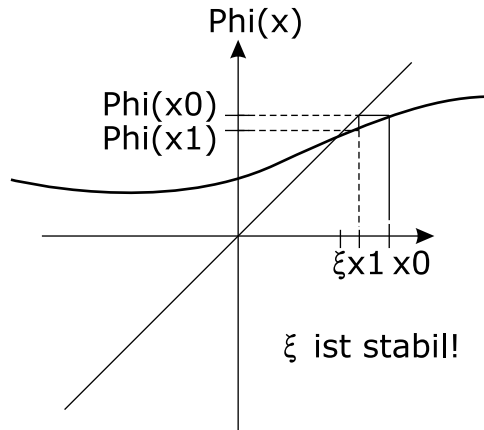
$$\begin{aligned} |x_m - x_n| &= |x_m - x_{m-1} + x_{m-1} - x_{m-2} + x_{m-2} + \dots + x_{n+1} - x_n| \\ &\leq \underbrace{|x_m - x_{m-1}|}_{L^{m-1} \cdot |x_1 - x_0|} + \underbrace{|x_{m-1} - x_{m-2}|}_{L^{m-2} \cdot |x_1 - x_0|} + \dots + \underbrace{|x_{n+1} - x_n|}_{L^n \cdot |x_1 - x_0|} \\ &\leq (L^{m-1} + L^{m-2} + \dots + L^n) |x_1 - x_0| \\ &= L^n \underbrace{(L^{m-1-n} + L^{m-2-n} + \dots + L + 1)}_{\leq \frac{1}{1-L}} \cdot |x_1 - x_0| \\ &\leq \frac{L^n}{1 - L} \cdot |x_1 - x_0| \rightarrow 0 \text{ für } n \rightarrow \infty \end{aligned}$$

Offensichtlich gilt $|x_m - x_n| \rightarrow 0$ für $m, n \rightarrow \infty$.

Also ist $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge (Siehe Definition 2.4.2), also konvergent. Es gibt also ein $\xi \in [a, b]$ mit $\lim_{n \rightarrow \infty} x_n = \xi$. Nach Satz 2.4.1 und Satz 2.4.2 folgt $\xi = \Phi(\xi)$

Grafische Veranschaulichung

Abbildung 2.5: Konvergenter Fall



Achtung:

Obwohl ein Fixpunkt ξ von Φ vorliegt, muss die Iteration nicht dagegen konvergieren! (vgl. unteres Bild „divergenter Fall“).

Definition 2.4.2 (Cauchy-Folge)

Eine Folge $(a_n)_{n \in \mathbb{N}}$ reeller Zahlen heißt Cauchy-Folge, falls es zu jedem $\epsilon > 0$ einen Index $N = N(\epsilon) \in \mathbb{N}$ gibt:

$$|a_n - a_m| < \epsilon; \forall n, m \geq N$$

In \mathbb{R} ist jede Cauchy-Folge konvergent.

Satz 2.4.3

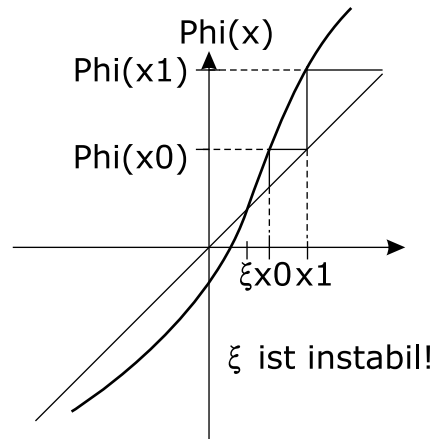
Sei $\Phi \in C^1([a, b])$ und es gelte $|\Phi'(x)| \leq L; x \in [a, b]$. Dann ist Φ Lipschitz-stetig auf $[a, b]$ mit der Konstanten L .

Beweis 2.4.3 (zu Satz 2.4.3)

Seien $x, y \in [a, b]$. Nach dem Mittelwertsatz gilt:

$$\Phi(x) - \Phi(y) = \Phi'(\eta)(x - y); \eta \in [\min(x, y), \max(x, y)]$$

Abbildung 2.6: Divergender Fall



$$\Rightarrow |\Phi(x) - \Phi(y)| = |\Phi'(\eta)| \cdot |x - y| \leq \underbrace{\max \{ |\Phi'(w)| \mid w \in [a, b] \}}_{\leq L} \cdot |x - y| \leq L |x - y|$$

Definition 2.4.3

Ein Fixpunkt ξ von Φ heißt stabil für die Iteration $x_{n+1} = \Phi(x_n)$, wenn es eine Umgebung $U =]\xi - \epsilon, \xi + \epsilon[$; $\epsilon > 0$ gibt, so dass für $x_0 \in U$ die Folge $x_{n+1} = \Phi(x_n)$ gegen ξ konvergiert.

Satz 2.4.4 (Stabilitätskriterium, ohne Beweis)

Sei $\Phi \in C^1([a, b])$ und sei $\xi \in [a, b]$ ein Fixpunkt. Dann gilt:

1. Ist $|\Phi'(\xi)| < 1$, so ist ξ stabil.
2. Ist $|\Phi'(\xi)| > 1$, so ist ξ instabil (nicht stabil).
3. Ist $|\Phi'(\xi)| = 1$, so kann ξ stabil oder instabil sein.

Satz 2.4.5 (Lokaler Konvergenzsatz für das Newtonverfahren)

Sei $f \in C^2([a, b])$ und es existiere ein $\xi \in]a, b[$ mit $f(\xi) = 0$, $f'(\xi) \neq 0$. Dann gibt es eine Umgebung U von ξ in $[a, b]$, so dass für alle Startwerte $x_0 \in U$ das Newtonverfahren $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ konvergiert.

Beweis:

Für das Newtonverfahren gilt:

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

ξ ist ein Fixpunkt für Φ , denn $\Phi(\xi) = \xi - \frac{f(\xi)}{f'(\xi)} = \xi$.

Zum Beweis der Behauptung genügt es zu zeigen, dass ξ ein stabiler Fixpunkt ist.

Nach Satz 2.4.4 1) ist ξ stabil, falls $|\Phi'(x)| < 1$ gilt.

Wir finden Φ einmal stetig differenzierbar und

$$\begin{aligned}\Phi'(x) &= 1 - \frac{f'(x) \cdot f'(x) - f(x) \cdot f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2} \\ \Rightarrow \Phi'(\xi) &= \frac{f(\xi) \cdot f''(\xi)}{f'(\xi)^2} = 0 \text{ Also ist } \xi \text{ stabil}\end{aligned}$$

Bemerkung 2.4.1

Vereinfachtes Newtonverfahren:

$$\Phi(x) = x - \frac{f(x)}{f'(x_0)}; \quad x_0 \text{ gegeben}$$

ξ mit $f(\xi) = 0$ ist Fixpunkt von Φ

$$\Phi'(x) = 1 - \frac{f'(x)}{f'(x_0)}$$

$$\Phi(\xi) = 1 - \frac{f'(\xi)}{f'(x_0)}$$

Also ist ξ stabil, falls $0 < \frac{f'(\xi)}{f'(x_0)} < 2$.

Definition 2.4.4 (Konvergenzordnung)

Sei Φ eine Iterationsfunktion mit Fixpunkt ξ , und U sei eine Umgebung von ξ . Für jeden Startwert $x_0 \in U$ erfülle die zugehörige Iterationsfolge $x_{n+1} = \Phi(x_n)$ eine Ungleichung der Form

$$|x_{n+1} - \xi| \leq C \cdot |x_n - \xi|^p; \quad n \geq 0, \quad C \geq 0, \quad p \geq 1$$

Falls $p = 1$, so gelte $0 < C < 1$. Dann heißt das Iterationsverfahren konvergent der Ordnung p .

Im Fall $p = 1$ spricht man von linearer Konvergenz und im Fall $p = 2$ spricht man von quadratischer Konvergenz (deutlich schneller).

Typisches Verhalten der iterierten x_1, x_2, x_3, \dots

1. bei linearer Konvergenz: $f(x_1) \approx 10^{-1}$, $f(x_2) \approx 10^{-2}$, $f(x_3) \approx 10^{-3}, \dots, f(x_n) \approx 10^{-n}$
2. bei quadratischer Konvergenz: $f(x_1) \approx 10^{-1}$, $f(x_2) \approx 10^{-2}$, $f(x_3) \approx 10^{-4}$, $f(x_4) \approx 10^{-8}, \dots, f(x_n) \approx 10^{-2^{n-1}}$

Satz 2.4.6

Es sei $f \in C^2([a, b])$; $f'(x) \neq 0$ für $x \in [a, b]$. Ferner existiere ein $\xi \in]a, b[$ mit $f(\xi) = 0$. Dann hat das Newtonverfahren die Konvergenzordnung 2, d.h. es gibt eine Umgebung U von ξ in $[a, b]$ so, dass die Newtoniteration $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$; $x_0 \in U$ durchführbar ist, und es gilt:

$$|x_{n+1} - \xi| \leq C \cdot |x_n - \xi|^2; \quad C > 0$$

Zum Beweis benötigen wir die Taylor-Formel.

Definition 2.4.5 (Taylor-Formel)

Man verwendet die Taylor-Formel, um Funktionen in der Umgebung eines Punktes durch Polynome anzunähern.

Sei $I \subset \mathbb{R}$ und $f : I \rightarrow \mathbb{R}$ $(n+1)$ -stetig differenzierbar. Ferner sei $a \in I$ und $x \in I$. Dann gilt

$$\begin{aligned} f(x) &= \underbrace{f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n}_{=: p_n(x) \text{ „n-tes Taylorpolynom zu f bei a“}} \\ &+ \underbrace{\frac{f^{(n+1)}(\eta)}{(n+1)!}(x-a)^{n+1}}_{=: R_{n+1}(x) \text{ „Restglied in differentieller Form“}}; \quad \eta \in [\min(a, x), \max(a, x)] \\ &= \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} + \frac{f^{(n+1)}(\eta)}{(n+1)!}(x-a)^{n+1} \end{aligned}$$

Entwicklung an Stelle a .

Beweis 2.4.4 (von Satz 2.4.6)

p_n ist ein Polynom vom Grad n . p_n approximiert f lokal bei a . Es gilt: $p^{(k)}(a) = f^{(k)}$; $k = 0, 1, \dots, n$.

Aus $f \in C^2([a, b])$ folgt:

$$\begin{aligned} C_1 &:= \max \{ |f''(x)| \mid x \in [a, b] \} < \infty \\ C_2 &:= \min \{ |f'(x)| \mid x \in [a, b] \} > 0 \end{aligned}$$

denn eine stetige Funktion auf einem kompakten Intervall nimmt ihr Maximum bzw. Minimum an.

Taylorentwicklung liefert:

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + \frac{f''(\eta)}{2!}(x - x_n)^2$$

Wähle $x = \xi$:

$$\begin{aligned}
 0 &= f(\xi) \\
 &= f(x_n) + f'(x_n)(\xi - x_n) + \frac{f''(\eta)}{2}(\xi - x_n)^2 \\
 &\Leftrightarrow \\
 f(x_n) + f'(x_n)(\xi - x_n) &= -\frac{f''(\eta)}{2}(\xi - x_n)^2 \\
 \underbrace{(\xi - x_n) + \frac{f(x_n)}{f'(x_n)}}_{\xi - (x_n - \frac{f(x_n)}{f'(x_n)})} &= \frac{-f''(\eta)}{2f'(x_n)}(\xi - x_n)^2 \\
 &= \xi - x_{n+1} \\
 |\xi - x_{n+1}| &= \left| \frac{f''(\eta)}{2f'(x_n)} \right| |\xi - x_n|^2 \\
 &\leq \frac{C_1}{2C_2} |\xi - x_n|^2 =: C|x_n - \xi|^2
 \end{aligned}$$

Bemerkung 2.4.2

$p = 2$ bedeutet sehr rasche (quadratische) Konvergenz.

Beispiel 2.4.1

Für ein $a > 0$ wollen wir \sqrt{a} berechnen. Betrachte daher die Gleichung

$$f(x) = x^2 - a = 0$$

Mit $f'(x) = 2x$ finden wir das Newtonverfahren

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

Für $a = 2$ lautet die exakte Lösung $\sqrt{2} = 1,414214\dots$

Starte das Newtonverfahren z.B. mit $x_0 = 1$.

$$\begin{aligned}
 x_1 &= \frac{1^2 + 2}{2} = \frac{3}{2} = 1,5 \\
 x_2 &= \frac{(\frac{3}{2})^2 + 2}{3} = \frac{17}{12} = 1,41666\dots \\
 x_3 &= \frac{(\frac{17}{12})^2 + 2}{\frac{34}{12}} = \frac{577}{408} \approx 1,414216
 \end{aligned}$$

Beachte die rasche Konvergenz!

Bemerkung 2.4.3

Starte mit $\tilde{x}_0 = 2$. Dann folgt $\tilde{x}_1 = \frac{2^2+2}{2} = 1,5 = x_1$. Der Vorgänger einer Iterierten ist also nicht eindeutig.

2.5 Das Sekantenverfahren

Ein Nachteil des Newton-Verfahrens besteht darin, dass in jedem Schritt die Berechnung einer Ableitung (nämlich $f'(x_n)$) erforderlich ist. Vermeide dies, indem man im Newtonverfahren die Tangente durch eine Sekante ersetzt.

Verfahren:

Sei $f \in C([a, b])$ und seien $x_0, x_1 \in [a, b]$; $x_0 \neq x_1$ gegeben. Für $n = 1, 2, 3$ ersetzen wir $f'(x_n)$ durch $\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$

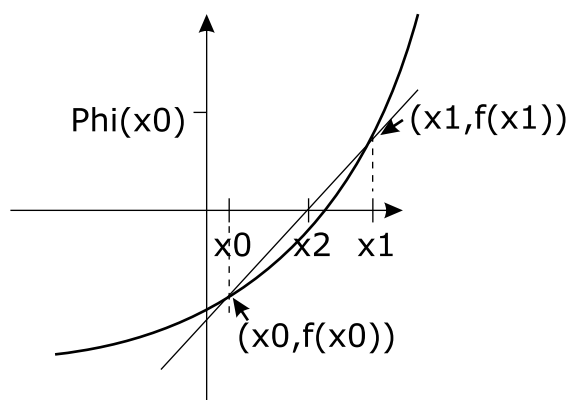
Vorschrift:

$$x_{n+1} = x_n - \frac{f(x_n)}{\left(\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}\right)} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})} = \Phi(x_n, x_{n-1})$$

d.h. die Iterationsfunktion hängt von x_n und x_{n-1} ab.

Graphisch ist x_{n+1} der Schnittpunkt der Sekante durch die Punkte $(x_{n-1}, f(x_{n-1}))$ und $(x_n, f(x_n))$ mit der x-Achse.

Abbildung 2.7: Graphische Darstellung



Bemerkung 2.5.1

1. Man erhält die Konvergenzordnung $p = \frac{1+\sqrt{5}}{2} \approx 1,618$
2. Man braucht, im Gegensatz zum Newtonverfahren, nur 1 Funktionsauswertung pro Iterationsschritt

2.6 Nichtlineare Gleichungssysteme

Sei $D \subset \mathbb{R}^N$ und $F : D \rightarrow \mathbb{R}^N$ zweimal stetig differenzierbar, d.h. es existieren alle partiellen Ableitungen $\frac{\delta^2 F(x)}{\delta x_i \delta x_j}$; $x \in D$ und sind stetig. Gesucht ist ein $\xi \in D$ mit $F(\xi) = 0$ (Nullstelle).

Komponentendarstellung

$$F = (F_1, \dots, F_N); \quad x = (x_1, \dots, x_N)$$

$$\begin{aligned} F_1(x_1, \dots, x_N) &= 0 \\ F_2(x_1, \dots, x_N) &= 0 \\ &\vdots \\ F_N(x_1, \dots, x_N) &= 0 \end{aligned}$$

2.6.1 Newtonverfahren zur Lösung von $F(x) = 0$

Beim Newtonverfahren konstruiert man wieder die „Tangente“ an die „Kurve“ $y = F(x)$ im Punkt (x^0, y^0) , $y^0 = F(x^0)$.

Als Nullstelle der Tangentengleichung finden wir

$$0 = F'(x^0)x + F(x^0) - F'(x^0)x^0 \text{ „Vektorgleichung“}$$

Mit $F'(x)$ berechnen wir die Funktionalmatrix von F

$$F'(x) = \begin{pmatrix} \frac{\delta F_1(x)}{\delta x_1} & \dots & \frac{\delta F_1(x)}{\delta x_N} \\ \vdots & & \vdots \\ \frac{\delta F_N(x)}{\delta x_1} & \dots & \frac{\delta F_N(x)}{\delta x_N} \end{pmatrix} \in \mathbb{R}^{N,N}$$

Die Funktion $F_i(x)$ ist in $y \in D$ nach x_j partiell differenzierbar falls die Funktion $g(x_j) = F_i : (y_1, \dots, y_{j-1}, x_j, y_{j+1}, \dots, y_N)$ differenzierbar ist, d.h.

$$\frac{g(x_j + \tau) - g(x_j)}{\tau} \rightarrow a; \quad \tau \rightarrow 0$$

Man schreibt

$$a = \frac{\delta F_i(y)}{\delta x_j}$$

Die neue iterierte x^1 finden wir als Lösung der Vektorgleichung.

$$\underbrace{\underbrace{F'(x^0)}_{\in \mathbb{R}^{N,N}} \underbrace{x^0}_{\in \mathbb{R}^N}}_{\in \mathbb{R}^N} - \underbrace{F(x^0)}_{\in \mathbb{R}^N} = \underbrace{\underbrace{F'(x^0)}_{\in \mathbb{R}^{N,N}} \underbrace{x}_{\in \mathbb{R}^N}}_{\in \mathbb{R}^N}$$

$$\underbrace{F'(x^0)^{-1}(F'(x^0)x^0 - F(x^0))}_{=x^0 - F'(x^0)^{-1}F(x^0)} = x (= x^1)$$

Newtonverfahren zur Lösung von $F(x) = 0$

Wähle Startpunkt $x^0 \in D$ und setze

$$x^{n+1} = x^n - F'(x^n)^{-1}F(x^n); \quad n = 0, 1, 2, \dots$$

Bemerkung 2.6.1

- Im Fall $N = 1$ ergibt sich das aus Kapitel 2.3 bekannte (eindimensionale) Newtonverfahren.
- Das Verfahren ist durchführbar, falls $F'(x_n) \in \mathbb{R}^{N,N}$ invertierbar für $n = 0, 1, 2, \dots$

Praktische Durchführung:

1. Wähle $x^0 \in D$ sowie $\epsilon_1, \epsilon_2 > 0$ beliebig.
2. Für $n = 0, 1, 2, \dots$ löse man das lineare Gleichungssystem $F'(x^n)v = -F(x^n)$ und setzt $x^{n+1} = x^n + v$.
3. Man breche die Iteration ab, falls $\|F(x^{n+1})\|_\infty < \epsilon_1$ oder $\|x^{n+1} - x^n\|_\infty < \epsilon_2$ oder falls nach einer vorgegebenen Iterationszahl keines dieser beiden Kriterien erfüllt wird.

Bemerkung 2.6.2 (Warnung)

Man invertiere niemals Matrizen, sondern man löse lineare Gleichungssysteme!

Bei den Iterationen des Newtonverfahrens ist die Lösung eines Gleichungssystems der Invertierung der Matrix $F'(x^n)$ vorzuziehen, da man hier auf die LR-Zerlegung zurückgreifen kann.

Varianten des Newtonverfahrens

1. Vereinfachtes Newtonverfahren:

Beim Newtonverfahren ist in jedem Schritt ein lineares Gleichungssystem mit der Matrix $A = F'(x^n)$ zu lösen. Zur Bereitstellung dieser Matrix sind die N^2 Funktionsauswertungen $\frac{\delta F_i}{\delta x_j}(x)$; $1 \leq i, j \leq N$ nötig.

Beim *vereinfachten Newtonverfahren* wählen wir immer $A = F'(x^0)$.

Algorithmus:

- Wähle $x^0 \in D$
- Berechne iterativ

$$F'(x^0)v^n = -F(x^n)$$

- Setze

$$x^{n+1} = x^n + v^n = x^n - DF(x^0)^{-1}F(x^n)$$

Dies erfordert nur eine LR-Zerlegung, nämlich die von $F'(x^0)$.

2. Gedämpftes Newtonverfahren:

Das *gedämpfte Newtonverfahren* unterscheidet sich vom *Newtonverfahren* durch eine neue Setzung der Iterierten.

Man setzt

$$x^{n+1} = x^n + \alpha_n v^n; \alpha_n \in]0, 1] \quad (F'(x^n)v^n = -F(x^n))$$

Hoffnung: Die Wahl einer Dämpfung erhöht die Konvergenzchancen.

2.6.2 Konvergenz des Newtonverfahrens

Bemerkung 2.6.3

Konvergiert die Newton-Iteration x_n gegen \bar{x} und ist $F'(\bar{x})$ invertierbar, so folgt $F(\bar{x}) = 0$, denn:

$$\bar{x} = \lim_{n \rightarrow \infty} x^n = \lim_{n \rightarrow \infty} x^{n+1} = \lim_{n \rightarrow \infty} (x^n - F'(x^n)^{-1}F(x^n)) = \bar{x} - F'(\bar{x})^{-1}F(\bar{x})$$

$$\Leftrightarrow 0 = F'(\bar{x})^{-1}F(\bar{x}) \Leftrightarrow F(\bar{x}) = 0$$

Satz 2.6.1 (Lokaler Konvergenzsatz, ohne Beweis)

Sei $D \subset \mathbb{R}^N$ und $F \in C^2(D, \mathbb{R}^N)$. Es existiere ein $\xi \in D$ mit $F(\xi) = 0$ und $F'(\xi)$ sei invertierbar. Dann gibt es eine Kugel:

$$K_\rho := \{x \in \mathbb{R}^N \mid \|x - \xi\|_\infty \leq \rho\} \subset D; \rho > 0$$

so dass ξ die einzige Nullstelle von F in $K_\rho(\xi)$ ist. Für jeden Anfangswert $x^0 \in K_\rho(\xi)$ existiert die Newton-Folge $x^n, x^n \in K_\rho(\xi)$ und es gilt $\lim_{n \rightarrow \infty} x^n = \xi$.

Weiter gibt es ein $C > 0$ mit

$$\|\xi - x^{n+1}\|_\infty \leq C \cdot \|\xi - x^n\|_\infty^2; n \in \mathbb{N}$$

Bemerkung 2.6.4

- Es liegt also lokal quadratische Konvergenz vor
- Problem des Newtonverfahrens:
Startet man fernab von einer Nullstelle, so divergiert das Newtonverfahren oftmals.

Beispiel 2.6.1

Gegeben sei

$$F(x) = \begin{pmatrix} F_1(x_1, x_2) \\ F_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} 2x_1^3 - x_2^2 - 1 \\ x_1x_2^3 - x_2 - 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Wir finden

$$F'(x) = \begin{pmatrix} 6x_1^2 & -2x_2 \\ x_2^3 & 3x_1x_2^2 - 1 \end{pmatrix} \in \mathbb{R}^{2,2}$$

Wähle $x^0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

$$F'(x^0) = \begin{pmatrix} 6 & -4 \\ 8 & 11 \end{pmatrix}, F(x^0) = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$$

Zu lösen ist $F'(x^0)v^0 = -F(x^0)$, d.h.

$$\begin{pmatrix} 6 & -4 \\ 8 & 11 \end{pmatrix} \cdot \begin{pmatrix} v_1^0 \\ v_2^0 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

z.B. mit dem Gaußverfahren. Die Lösung ist

$$v^0 = \begin{pmatrix} \frac{25}{98} \\ -\frac{18}{49} \end{pmatrix}$$

und somit

$$x^1 = x^0 + v^0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} \frac{25}{98} \\ -\frac{18}{49} \end{pmatrix} = \begin{pmatrix} \frac{123}{98} \\ \frac{80}{49} \end{pmatrix} \approx \begin{pmatrix} 1,255 \\ 1,633 \end{pmatrix}$$

Auf 10 Stellen genau lautet die exakte Lösung

$$\xi = \begin{pmatrix} 1,234274484 \\ 1,661526467 \end{pmatrix}$$

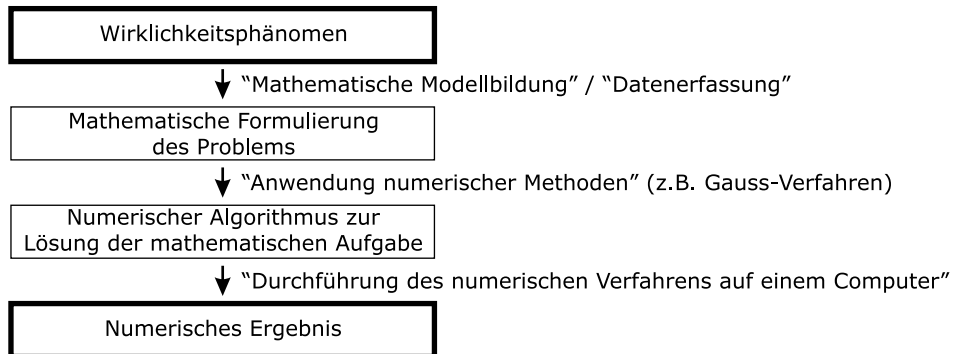
Diese Genauigkeit erreicht das Newtonverfahren mit 4 Iterationsschritten.

2.7 Modellierung und Numerik

2.7.1 Einführung

Versuche darzustellen, welche Schritte erforderlich sind, um ein in der Wirklichkeit beobachtetes Phänomen numerisch zu analysieren.

Abbildung 2.8: Schematische Darstellung



Definition 2.7.1 (Modellfehler)

Modellfehler entstehen dadurch, dass das verwendete mathematische Modell das zu untersuchende Problem vereinfacht oder nur idealisiert beschreibt.

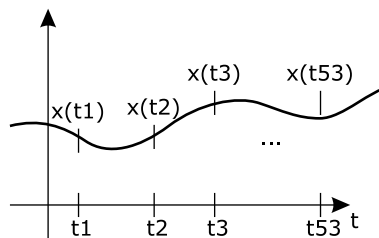
Definition 2.7.2 (Datenfehler)

Diese rühren von Messungenauigkeiten oder Messfehlern her.

Definition 2.7.3 (Prinzipielle Fehler numerischer Methoden)

Abbruchfehler bei iterativen Verfahren. Diskretisierungsfehler (falls Funktionen gesucht werden)

Abbildung 2.9: Schematische Darstellung. Funktionen müssen diskret dargestellt werden



Definition 2.7.4 (Rundungsfehler)

Rundungsfehler entstehen, weil nur endlich viele Zahlen auf dem Rechner darstellbar sind und die Elementaroperatoren $+$, $-$, \cdot , $:$ nur mit endlicher Genauigkeit ausgeführt werden.

Beispiel 2.7.1 (Erdbebenhilfe in Zentralamerika)

- Mathematisches Problem: $Ax = b$
- Numerischer Algorithmus: Gauss-Elimination
- Numerisches Ergebnis: Lösungsvektor

Gauss-Elimination ist ein direkter Algorithmus, welcher in endlich vielen Schritten zur exakten Lösung führt, d.h. keine Diskretisierungs- und keine Abbruchfehler, sehr wohl aber Rundungsfehler.

Beispiel 2.7.2 (Berechnung von Wurzeln reeller Zahlen)

- Mathematisches Problem: $f(x) = 0$ (z.B. für $\sqrt{2}$ setze $f(x) = x^2 - 2$)
- Numerischer Algorithmus: Iterationsverfahren $x_{n+1} = \Phi(x_n)$
- Numerisches Ergebnis: Skalare Lösung

Es entstehen Abbruchfehler und Rundungsfehler

Beispiel 2.7.3 (Raketenflug zum Mond)

- Mathematisches Problem: Gewöhnliche Differentialgleichung. Gesucht ist eine Funktion $u(t)$, für welche gilt

$$u'(t) = f(u(t)); u(t_0) = u_0$$

mit u_0 und f vorgegeben.

$$u \in C^1([t_0, t_{end}])$$

Hier ist eine Funktion gesucht, d.h. es liegen Diskretisierungsfehler und Rundungsfehler vor.

- Numerischer Algorithmus: (nicht angegeben)
- Numerisches Ergebnis: (nicht angegeben)

Bemerkung 2.7.1

- In jedem der 3 Beispiele benötigen wir Informationen über den Fehler. Diese gewinnt man über Fehlerabschätzung oder in Form von Konvergenzaussagen.
- Ermittlung einer „Lösung“ ohne jegliche Kenntnis des Fehlers ist sinnlos!

3 Lineare Optimierung

3.1 Problemstellung und Normalformen

In der *linearen Optimierung* werden lineare Zielfunktionen unter linearen Nebenbedingungen optimiert.

Beispiel 3.1.1

In einer Stanzerie werden aus Blechen drei verschiedene Teile T_1, T_2, T_3 gestanzt. Aufgrund der Geometrie der Teile werden vier sinnvoll erscheinende Varianten V_1, V_2, V_3, V_4 aus einer Blechtafel technologisch vorbereitet.

Beim Stanzen dieser Varianten entstehen folgende Stückzahlen:

Tabelle 3.1: Stanzvarianten

Teile \ Varianten	V_1	V_2	V_3	V_4
Anzahl T_1	1	1	0	0
Anzahl T_2	1	0	1	0
Anzahl T_3	2	4	6	8

Es ist ein Auftrag von 3 Stück T_1 , 2 Stück T_2 und 40 Stück T_3 zu stanzen. Wie oft müssen die einzelnen Varianten zur Anwendung kommen, damit möglichst wenig Bleche verbraucht werden.

Es bezeichne v_j die Anzahl der benutzten Stanzvariante V_j ; $j = 1, \dots, 4$.

Wählt man v_j oft die Variante V_j , so wird produziert:

$$\begin{pmatrix} \text{Anzahl } T_1 \\ \text{Anzahl } T_2 \\ \text{Anzahl } T_3 \end{pmatrix} = v_1 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + v_2 \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} + v_3 \begin{pmatrix} 0 \\ 1 \\ 6 \end{pmatrix} + v_4 \begin{pmatrix} 0 \\ 0 \\ 8 \end{pmatrix} = \begin{pmatrix} v_1 + v_2 \\ v_1 + v_3 \\ 2v_1 + 4v_2 + 6v_3 + 8v_4 \end{pmatrix}$$

Es muss gelten:

$$\begin{pmatrix} \text{Anzahl } T_1 \\ \text{Anzahl } T_2 \\ \text{Anzahl } T_3 \end{pmatrix} \geq \begin{pmatrix} 3 \\ 2 \\ 40 \end{pmatrix}, \text{ d.h. } \begin{array}{l} v_1 + v_2 \geq 3 \\ v_1 + v_3 \geq 2 \\ 2v_1 + 4v_2 + 6v_3 + 8v_4 \geq 40 \end{array}$$

Möglichst wenig Bleche sollen verbraucht werden, d.h. $\underbrace{v_1 + v_2 + v_3 + v_4}_{= \text{Anzahl Bleche}} \rightarrow \min..$

Ferner gilt $v_i \geq 0$; $i = 1, \dots, 4$ und (eigentlich) ganzzahlig.

Beispiel 3.1.2

Ein Landwirt beabsichtigt, ins Tiefkühlgeschäft einzusteigen. Er will 30ha seines Landes zum Anbau von Erbsen und Möhren, verwenden. Ein ha Erbsen verursachen 200€ und ein ha Möhren 100€ Saatkosten. Mehr als 5000€ sollen nicht investiert werden. Der Zeitaufwand für den Anbau eines ha Erbsen wird mit 1 Arbeitstag, für Möhren mit 2 Arbeitstagen veranschlagt. Insgesamt hat der Landwirt 50 Tage Zeit.

Gesucht sind die gewinnoptimalen Anbauflächen, falls ein ha Erbsen 400€ und ein ha Möhren 600€ Reingewinn bringen?

Lösung:

Es sei x_1 die für Erbsenanbau bzw. x_2 die für den Möhrenanbau vorgesehene Fläche.

Maximierungsproblem: $\underbrace{400x_1 + 600x_2}_{\text{Gewinn}} \rightarrow \max.$ („Zielfunktion“).

Unter den Nebenbedingungen:

$$\begin{array}{ll} x_1 + x_2 \leq 30 & \text{(Beschränkung Anbaufläche)} \\ 200x_1 + 100x_2 \leq 5000 & \text{(Beschränkung Saatkosten)} \quad ; \quad x_1, x_2 \geq 0 \\ x_1 + 2x_2 \leq 50 & \text{(Beschränkung Arbeitszeit)} \end{array}$$

Definition 3.1.1 (Strukturelle Form eines linearen Optimierungsproblems): Normalform 1:

Nebenbedingungen:

$$\sum_{j=1}^k g_{ij}x_j \leq b_i; \quad i = 1, \dots, M; \quad x_j \geq 0; \quad j = 1, \dots, k$$

mit $k = \text{Anzahl der Variablen}$, $M = \text{Anzahl der Nebenbedingungen}$.

Zielfunktion:

$$z(x) = z(x_1, \dots, x_k) = \sum_{j=1}^k c_j x_j \rightarrow \max.$$

Matrixdarstellung der Normalform 1:

$$\begin{array}{l} Gx \leq b; \quad G \in \mathbb{R}^{M,k}; \quad b \in \mathbb{R}^M; \quad x \geq 0; \quad x \in \mathbb{R}^k; \quad c \in \mathbb{R}^k \\ b = (b_1, \dots, b_M), \quad x = (x_1, \dots, x_k), \quad c = (c_1, \dots, c_k) \end{array}$$

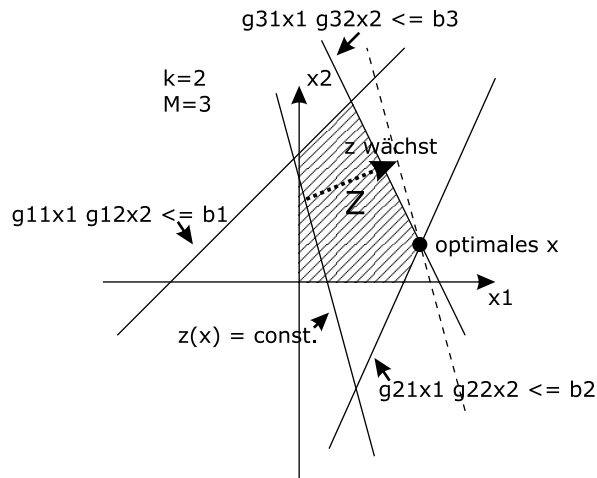
$$z(x) = c^T x \rightarrow \max.$$

Die vektorwertigen Ungleichungen sind komponentenweise zu verstehen.

$$\begin{aligned} Gx \leq b &\Leftrightarrow (Gx)_i \leq b_i; i = 1, \dots, M \\ x \geq 0 &\Leftrightarrow x_j \geq 0; j = 1, \dots, k \end{aligned}$$

Veranschaulichung einer linearen Optimierungsaufgabe im Fall $k = 2$. Jede der Ungleichungen $(Gx)_j \leq b_j; j = 1, \dots, M$ sowie $x_j \geq 0; j = 1, 2$ definiert einen linearen Halbraum.

Abbildung 3.1: Schematische Darstellung



$Z = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geq 0, x_2 \geq 0, (Gx)_j \leq b_j, j = 1, 2, 3\}$ heißt *zulässiger Bereich*. Gesucht ist das Maximum von $z(x)$ für $x \in Z$.

Zeichne nun die Gerade $z(x) = \text{const.}$ ein. Man verschiebt die Gerade $z(x) = \text{const.}$ zu immer größeren Werten, solange der Schnitt mit Z nicht leer ist, d.h. (in den meisten Fällen) bis nur noch ein Schnittpunkt existiert.

Normalform 2:

Sei $G \in \mathbb{R}^{M,k}; b \in \mathbb{R}^M; c \in \mathbb{R}^k$. Gesucht sind $x \in \mathbb{R}^k; w \in \mathbb{R}^M; x \geq 0; w \geq 0$ mit

$$Gx + w = b$$

sowie

$$z(x) = c^T x + c_0 = \max.$$

Bemerkung 3.1.1

Die Variablen w heißen auch Schlupfvariablen.

Alternativ:

$$A \in \mathbb{R}^{M, M+k}; b \in \mathbb{R}^M; p \in \mathbb{R}^{k+M}$$

Gesucht ist $y \in \mathbb{R}^{M+k}; y \geq 0$ mit

$$Ay = b \text{ und } z(y) = p^T y + c_0 \stackrel{!}{=} \max.$$

Dabei ist

$$A = (G|I_M); y = \begin{pmatrix} x \\ w \end{pmatrix}; p = \begin{pmatrix} c \\ 0 \end{pmatrix}$$

Normalform 3:

Gegeben sei $A \in \mathbb{R}^{M, M+k}, b \in \mathbb{R}^M, p \in \mathbb{R}^{M+k}, \text{rg}(A) = M$. Gesucht ist $y \in \mathbb{R}^{M+k}; y \geq 0$ mit

$$Ay = b \text{ und } z(y) = p^T y + c_0 \stackrel{!}{=} \max.$$

Bemerkung 3.1.2

- Gegenüber der Normalform 2 ist die Struktur $y = \begin{pmatrix} x \\ w \end{pmatrix}, p = \begin{pmatrix} c \\ 0 \end{pmatrix}, A = (G|I_M)$ entfallen.
- Die Voraussetzung $\text{rg}(A) = M$ sichert, dass die Menge der Lösungen von $Ay = b$ ein k -dimensionaler Unterraum des \mathbb{R}^{M+k} ist.

Die Normalformen 1, 2 und 3 sind in einem gewissen Sinne gleichwertig:

Normalform 1 \Rightarrow Normalform 3:

- Setze $w = b - Gx \in \mathbb{R}^M$ und $w \geq 0$
- Setze $A = \left(\underbrace{G}_{\in \mathbb{R}^{M,k}} \mid \underbrace{I_M}_{\in \mathbb{R}^{M,M}} \right) \in \mathbb{R}^{M, M+k}; y = (x, w) \in \mathbb{R}^{k+M}$ und erhalte

$$Ay = (G|I_M) \cdot \begin{pmatrix} x \\ w \end{pmatrix} = Gx + w = Gx + b - Gx = b; \text{rg}(A) = M$$

, da A die Teilmatrix I_M enthält.

Mit $y = (x, w)$ und $p = (c, 0)$ folgt die Äquivalenz der Zielfunktion.

Normalform 3 \Rightarrow Normalform 2:

Der Nachweis erfolgt durch das Simplexverfahren anhand von Austauschschritten.

Normalform 2 \Rightarrow Normalform 1:

klar, denn $Gx + w = b; w \geq 0 \Rightarrow Gx \leq b$ und die Zielfunktionen sind äquivalent

3.1.1 Durchführbarkeit allgemeiner linearer Optimierungsaufgaben auf einer der Normalformen

Betrachte das lineare Problem

$$\sum_{j=1}^M a_{ij}x_j = b_i; \quad i \in I_ = \{j \in \{1, \dots, M\} | (Ax)_j = b_j\}$$

$$\sum_{j=1}^M a_{ij}x_j \leq b_i; \quad i \in I_{\leq} = \{j \in \{1, \dots, M\} | (Ax)_j \leq b_j\}$$

$$\sum_{j=1}^M a_{ij}x_j \geq b_i; \quad i \in I_{\geq} = \{j \in \{1, \dots, M\} | (Ax)_j \geq b_j\}$$

$$x_j \geq 0; \quad j \in J_{\geq} = \{l \in \{1, \dots, k\} | x_l \geq 0\}$$

$$x_j \leq 0; \quad j \in J_{\leq} = \{l \in \{1, \dots, k\} | x_l \leq 0\}$$

$$x_j \in \mathbb{R}; \quad j \in J_0 = \{l \in \{1, \dots, k\} | x_l \in \mathbb{R}\}$$

$$z(x) = c_0 + \sum_{i=1}^k c_i x_i \stackrel{!}{=} \min./\max.$$

Man beachte, dass der Minimierung von z die Maximierung von $-z$ entspricht, d.h. Multiplikation der Zielfunktion mit (-1) wandelt eine Minimierungsaufgabe in eine Maximierungsaufgabe um.

Sei $i \in I_ =$:

$$\sum_{j=1}^M a_{ij}x_j = b_i$$

$$\Leftrightarrow \sum_{j=1}^M a_{ij}x_j \leq b_i \quad \text{und} \quad \sum_{j=1}^M a_{ij}x_j \geq b_i$$

$$\Leftrightarrow \sum_{j=1}^M a_{ij}x_j \leq b_i \quad \text{und} \quad -\sum_{j=1}^M a_{ij}x_j \leq -b_i$$

Sei $i \in I_{\geq}$:

$$\sum_{j=1}^M a_{ij}x_j \geq b_i \Leftrightarrow -\sum_{j=1}^M a_{ij}x_j \leq -b_i$$

Sei $i \in I_{\leq}$: ✓

Sei $j \in J_{\leq}$:

$$x_j \leq 0 \Leftrightarrow \tilde{x}_j := -x_j \geq 0$$

Sei $j \in J_{\geq}$: ✓

Sei $j \in J_0$ und sei $x_j \in \mathbb{R}$: Führe neue Variablen $x_j^{(1)}, x_j^{(2)}$ ein und ersetze x_j gemäß $x_j = x_j^{(1)} - x_j^{(2)}$; $x_j^{(1)}, x_j^{(2)} \geq 0$.

Definition 3.1.2

Die Menge der zulässigen Lösung ist definiert als

$$Z = \left\{ x \in \mathbb{R}^k \mid Gx \leq b, x \geq 0 \right\} \text{ für ein Problem in Normalform 1}$$

bzw.

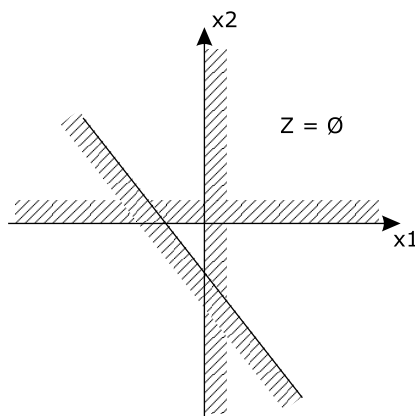
$$\bar{Z} = \left\{ y \in \mathbb{R}^{M+k} \mid Ay = b, y \geq 0 \right\} \text{ für Probleme in Normalform 2,3}$$

Definition 3.1.3

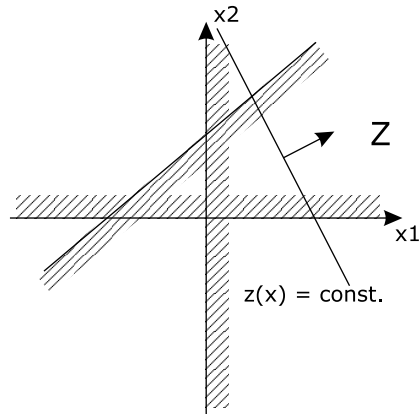
$x \in \mathbb{R}^k$ bzw. $y \in \mathbb{R}^{M+k}$ heißt optimale Lösung, wenn x bzw. y eine zulässige Lösung ist und die Zielfunktion dort ihr Minimum annimmt.

Mögliche Fälle bei linearen Optimierungsaufgaben:

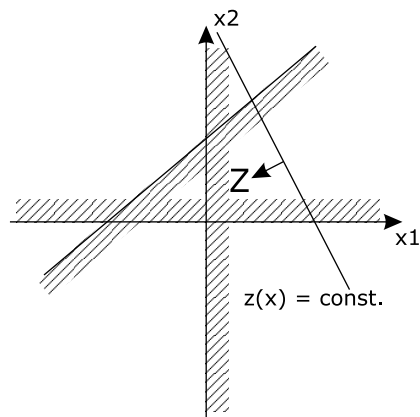
1. Es kann der Fall auftreten, dass keine zulässige Lösung existiert, d.h. $Z = \emptyset$. Dann kann es auch keine optimale Lösung geben.



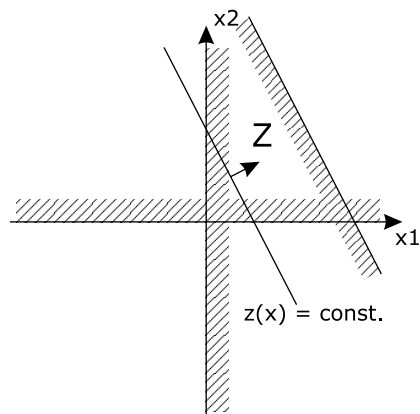
2. Die Zielfunktion z kann auf der Menge der zulässigen Lösungen Z unbeschränkt sein.



3. Auch wenn die Menge Z unbeschränkt ist, kann eine optimale Lösung x^* existieren.



4. Es können unendlich viele Lösungen existieren. Dies kann auftreten, wenn $z(x) = const$ zu einer der Geraden der Nebenbedingungen parallel ist.



Man beachte, dass das Optimum immer am Rand liegt, weil $z(x)$ linear ist und $z'(x) = c^T \neq 0$.

3.1.2 Zusammenfassung

Die Lösung einer linearen Optimierungsaufgabe $x \in \mathbb{R}^2$ lässt sich graphisch veranschaulichen. Die Ungleichung $(Gx)_i \leq b_i$; $i = 1, \dots, M$ definiert für jedes $i \in \{1, \dots, M\}$ einen Halbraum in \mathbb{R}^2 . Der zulässige Bereich $Z = \{x \in \mathbb{R}^2 \mid Gx \leq b, x \geq 0\}$ ist als Schnitt von Halbräumen ein Polyeder.

Die Isogewinnflächen $z(x) = \text{const.}$ sind Geraden. Die Optimallösung erhält man, indem man die Isogewinnfläche so weit wie möglich in Richtung wachsender Erlöse parallel verschiebt, bis sie den Polyeder nur noch berührt.

Ecken von Polyedern

Ecken sind im \mathbb{R}^2 dadurch definiert, dass zwei Ungleichungen nicht strikt, sondern mit Gleichheit angenommen werden. Diese Ungleichungen heißen dann auch *aktiv*.

Satz 3.1.1

Für ein lineares Optimierungsproblem in Normalform 1, 2 oder 3 gilt:

- Ist $Z \neq \emptyset$, so ist Z bzw. \bar{Z} ein konvexer Polyeder.
- Ist $Z \neq \emptyset$ und ist Z beschränkt, so existiert eine optimale Lösung
- Ist $Z \neq \emptyset$ und ist Z unbeschränkt, so gilt:

Es gibt eine optimale Lösung genau dann, wenn die Zielfunktion z auf Z nach oben beschränkt ist.

Bemerkung 3.1.3

Eine Menge $S \subset \mathbb{R}^k$ heißt konvex, falls für je zwei Punkte $p, q \in S$ die Strecke

$$\overline{p, q} := \left\{ v \in \mathbb{R}^k \mid v = \lambda p + (1 - \lambda)q; 0 \leq \lambda \leq 1 \right\}$$

in S liegt.

Bemerkung 3.1.4 (Austauschschritte und das Gauss-Jordan-Verfahren)

Betrachte $Ax = b$, A invertierbar, $A \in \mathbb{R}^{N, N}$.

Das Gauss-Jordan-Verfahren ist eine Sequenz

$$(A|b) = (A^{(0)}|b^{(0)}) \rightarrow (A^{(1)}|b^{(1)}) \rightarrow \dots \rightarrow (A^{(N)}|b^{(N)}) = (I_N|A^{-1}b)$$

Im zweiten Gauss-Jordan-Schritt wählen wir $(A^{(1)})_{1,2} = \frac{1}{2}$ und finden

$$(A^{(2)}|b^{(2)}) = \left(\begin{array}{ccccc|c} 0 & 1 & \frac{15}{2} & \frac{7}{2} & -\frac{1}{2} & \frac{11}{2} \\ 1 & 0 & -\frac{7}{2} & -\frac{3}{2} & \frac{1}{4} & -\frac{5}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \Rightarrow \text{rg}(A) = 2$$

Das System ist lösbar und die allgemeine Lösung $x = (x_1, \dots, x_5)$ hat $5 - 2 = 3$ Parameter. Sie kann elementar bestimmt werden.

3.2 Das Simplexverfahren

3.2.1 Einführung

Betrachte eine lineare Optimierungsaufgabe in Normalform 2:

$$Ay = b, y \geq 0; A = (G|I_M) \in \mathbb{R}^{M, M+k}, G \in \mathbb{R}^{M, k}, y = \begin{pmatrix} x \\ w \end{pmatrix} \in \mathbb{R}^{M+k}, w, b \in \mathbb{R}^M, x \in \mathbb{R}^k$$

Die Zielfunktion $z(x) = c^T x + c_0$ soll maximiert werden.

Setze zunächst $b \geq 0$ voraus. Dann ist $y = \begin{pmatrix} 0 \\ b \end{pmatrix} \geq 0$ eine zulässige Lösung, denn

$$Ay = (G|I_M) \begin{pmatrix} 0 \\ b \end{pmatrix} = G \cdot 0 + I_M \cdot b = b$$

Definition 3.2.1

1. $y_0 \in \mathbb{R}^{M+k}$ heißt Basislösung von $Ay = b$, $A \in \mathbb{R}^{M, M+k}$, $b \in \mathbb{R}^M$, wenn höchstens M Komponenten von y ungleich 0 sind und $Ay_0 = b$.
2. y_0 heißt Ecke von $\bar{Z} = \{y \in \mathbb{R}^{M+k} | Ay = b, y \geq 0\}$, wenn $y_0 \in \bar{Z}$ und y_0 Basislösung ist.

3.2.2 Idee des Verfahrens

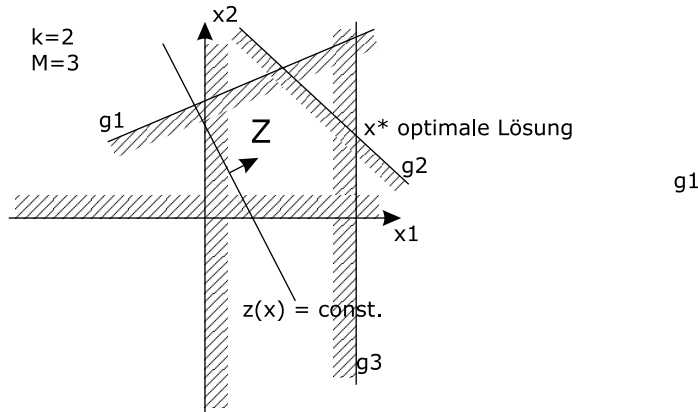
Der zulässige Bereich $Z = \{x \in \mathbb{R}^k | Gx \leq b, x \geq 0\}$ wird durch $x_i \geq 0$, $i = 1, \dots, k$ sowie durch M Halbräume bestimmt. Jeder Ecke eines Polyeders Z entspricht eine zulässige Basislösung.

Gleichungssystem $Ay = b$

$$(G|I) \cdot \begin{pmatrix} y \\ w \end{pmatrix} = b$$

Offensichtlich ist $y = (0, b)$ bzw. $x = 0$, $w = b$ eine zulässige Basislösung.

Dieser Basislösung entspricht die Ecke $x = 0$ des Polyeders Z .



Durch Jordansche Austauschschritte in dem System $Ay = b$ bzw. $Gx + w = b$ geht das Simplex-Verfahren von einer Ecke zu einer benachbarten Ecke, beginnend mit der Startecke $x = 0, w = b$.

Dabei sind die Austauschschritte nur dann sinnvoll, wenn gilt:

1. Die neue Ecke ist wieder zulässig, d.h. für die transformierte rechte Seite $\tilde{b} \geq 0$.
2. Die Zielfunktion z soll beim Übergang zu einer neuen Ecke zumindest nicht kleiner werden.

3.2.3 Beispiel für das Simplexverfahren

$$G = \begin{pmatrix} 3 & -1 & 2 & 1 \\ -2 & 4 & 0 & 2 \\ -4 & 3 & 8 & 2 \end{pmatrix} \in \mathbb{R}^{3,4}, \quad b = \begin{pmatrix} 7 \\ 12 \\ 10 \end{pmatrix} \in \mathbb{R}^3; \quad z(x) = c^T x = \max. \quad \text{mit } c = \begin{pmatrix} -1 \\ 3 \\ -2 \\ 1 \end{pmatrix}$$

Schema:

$$H^{(0)} = \left(\begin{array}{cccc|c} A & & & & b \\ p^T & & & & 0 \end{array} \right) = \left(\begin{array}{cccc|ccc|c} G & I & & & b \\ c^T & 0 & 0 & 0 & 0 \end{array} \right)$$

$$H^{(0)} = \left(\begin{array}{cccc|ccc|c} 3 & -1 & 2 & 1 & 1 & 0 & 0 & 7 \\ -2 & \boxed{4} & 0 & 2 & 0 & 1 & 0 & 12 \\ -4 & 3 & 8 & 2 & 0 & 0 & 1 & 10 \\ \hline -1 & 3 & -2 & 1 & 0 & 0 & 0 & 0 \end{array} \right)$$

!!habe b und $H^{(0)}$ leicht geändert. ok?

Beachte:

Die ersten 3 Zeilen entsprechen $Ay = b$. Die letzte Zeile entspricht $z(x) = c^T x = 0$. Man startet mit $y = \begin{pmatrix} 0 \\ b \end{pmatrix}$.

Es gilt:

$$z(\text{Startwert}) = z(0) = 0$$

Bezeichnungen:

$$\begin{aligned} y = (x, w) &= (x_1, \dots, x_4, w_1, w_2, w_3) \\ &= (y_1, \dots, y_7) \end{aligned}$$

Man nennt $y_5 = w_1$, $y_6 = w_2$, $y_7 = w_3$ die Basisvariablen. Die restlichen Variablen heißen Nichtbasisvariablen.

Der Übergang zu einer neuen Basis erfolgt durch einen Jordanschen Austauschschritt.

Definition 3.2.2 (Fakt)

Gilt $c_i \leq 0$, $i = 1, \dots, 4$ so hat $z(x) = c^T x$ für $x \geq 0$ nur nicht-positive Werte. Dann ist die Zielfunktion nicht mehr zu verbessern und $x_1 = x_2 = x_3 = x_4 = 0$ ist die optimale Lösung.

Wähle deshalb das Pivot-Element aus einer Spalte s , in der $c_s > 0$, z.B. $a_{rs} = a_{22} = 4$, da $c_2 = 3 > 0$. !!Spalte r nach Spalte s geändert

Beachte:

Wähle das Pivot-Element a_{rs} so, dass die neue transformierte Seite die Relation $\tilde{b} > 0$ erfüllt. Dies impliziert

1. $a_{rs} > 0$, da sonst $\tilde{b}_r < 0$ würde
2. Es muss gelten: $a_{rs}b_i - a_{is}b_r = \tilde{b}_i \cdot a_{rs} \geq 0$; $i = 1, \dots, r-1, r+1, \dots, M$

In diesem Beispiel wählen wir $a_{rs} = a_{22} = 4$ als Pivot-Element. Dann gilt:

$$\begin{aligned} a_{22} = 4 &> 0 \\ 4 \cdot 7 - (-1) \cdot 12 &\geq 0 \\ 4 \cdot 10 - 3 \cdot 12 &\geq 0 \end{aligned}$$

Die Durchführung eines Austauschschrittes für $H^{(0)}$ liefert

$$\tilde{H}^{(0)} = \left(\begin{array}{cccc|ccc|c} \frac{5}{2} & 0 & 2 & \frac{3}{2} & 1 & \frac{1}{4} & 0 & 10 \\ -2 & 4 & 0 & 2 & 0 & 1 & 0 & 12 \\ -\frac{5}{2} & 0 & 8 & -\frac{1}{2} & 0 & -\frac{3}{4} & 1 & 1 \\ \hline \frac{1}{2} & 0 & -2 & -\frac{1}{2} & 0 & -\frac{3}{4} & 0 & -9 \end{array} \right)$$

$$H^{(1)} = \left(\begin{array}{ccc|ccc|c} \boxed{\frac{5}{2}} & 0 & 2 & \frac{3}{2} & 1 & \frac{1}{4} & 0 & 10 \\ -\frac{1}{2} & 1 & 0 & \frac{1}{2} & 0 & \frac{1}{4} & 0 & 3 \\ -\frac{3}{2} & 0 & 8 & -\frac{1}{2} & 0 & -\frac{3}{4} & 1 & 1 \\ \hline \frac{1}{2} & 0 & -2 & -\frac{1}{2} & 0 & -\frac{3}{4} & 0 & -9 \end{array} \right)$$

Die Basisvariablen zu $H^{(0)}$ waren durch die Indizes 5, 6 und 7 charakterisiert. Die Basisindizes für $H^{(1)}$ sind 2, 5 und 7 (die Einheitsspalten von $H^{(1)}$).

Die Basislösung erhält man, indem man die 4 Nichtbasisvariablen y_1, y_3, y_4, y_6 zu Null setzt und die Werte der Basisvariablen abliest.

$$\text{Aus } \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_2^{(1)} \\ y_5^{(1)} \\ y_7^{(1)} \end{pmatrix} = \begin{pmatrix} 10 \\ 3 \\ 1 \end{pmatrix} \text{ erhalten wir } y_5^{(1)} = 10, y_2^{(1)} = 3, y_7^{(1)} = 1$$

,d.h. $y^{(1)} = (0, 3, 0, 0, 10, 0, 1)$.

Ferner haben wir

$$(p^{(1)})^T = \left(\frac{1}{2}, 0, -2, -\frac{1}{2}, 0, -\frac{3}{4}, 0 \right)$$

sowie

$$z^{(1)} = c^T x^{(1)} = 3 \cdot 3 = 9 = -(H^{(1)})_{4,8} = \underbrace{(-1, 3, -2, 1)}_{=c \text{ „Zielfunktionskoeffizienten“}} \underbrace{\begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \end{pmatrix}}_{=x=y_{1,\dots,4}^{(1)}}$$

Das Minuszeichen erklärt sich dadurch, dass diese Zahl auf der rechten Seite des GLS steht.

Pivotwahl für den nächsten Austauschschritt:

$$c_1^{(1)} = \frac{1}{2} > 0$$

und in der ersten Spalte von $H^{(1)}$ finden wir $(H^{(1)})_{1,1} = \frac{5}{2} > 0$ als Pivot-Element.

Ausführen des Pivot-Schrittes liefert:

$$\tilde{H}^{(1)} = \left(\begin{array}{ccc|ccc|c} \frac{5}{2} & 0 & 2 & \frac{3}{2} & 1 & \frac{1}{4} & 0 & 10 \\ 0 & 1 & \frac{2}{5} & \frac{4}{5} & \frac{1}{5} & \frac{3}{10} & 0 & 5 \\ 0 & 0 & 10 & 1 & 1 & -\frac{1}{2} & 1 & 11 \\ \hline 0 & 0 & -\frac{12}{5} & -\frac{4}{5} & -\frac{1}{5} & -\frac{4}{5} & 0 & -11 \end{array} \right)$$

$$H^{(2)} = \left(\begin{array}{ccc|cc|cc|c} 1 & 0 & 4 & \frac{3}{5} & 2 & \frac{1}{10} & 0 & 4 \\ 0 & 1 & \frac{2}{5} & \frac{4}{5} & \frac{1}{5} & \frac{3}{10} & 0 & 5 \\ 0 & 0 & 10 & 1 & 1 & -\frac{1}{10} & 1 & 11 \\ \hline 0 & 0 & -\frac{12}{5} & -\frac{4}{5} & -\frac{1}{5} & -\frac{4}{5} & 0 & -11 \end{array} \right)$$

Die Basisvariablen haben nun die Indizes 1, 2 und 7. Setzen wir die Nichtbasisvariablen zu Null, so folgt:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1^{(2)} \\ y_2^{(2)} \\ y_7^{(2)} \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 11 \end{pmatrix}, \text{ d.h. } y_1^{(2)} = 4, y_2^{(2)} = 5, y_7^{(2)} = 11$$

und somit $y^{(2)} = (4, 5, 0, 0, 0, 0, 11)$.

Da $(p^{(2)})^T = (0, 0, -\frac{12}{5}, -\frac{4}{5}, -\frac{1}{5}, -\frac{4}{5}, 0) \leq 0$ ist, haben wir eine optimale Lösung gefunden. Wir erhalten $y^{(2)} = y^* = (x^*, w^*)$ mit $x^* = (4, 5, 0, 0)$ und $w^* = (0, 0, 11)$, sowie die optimale Zielfunktion

$$z^* = c^T x^* = -(H^{(2)})_{4,8} = 11$$

3.2.4 Formale Darstellung des Simplex-Verfahrens

Im allgemeinen Fall werden hintereinander Matrizen $H^{(i)}$, $i = 0, 1, 2$ der Form

$$H^{(i)} = \left(\begin{array}{c|c} A^{(i)} & b^{(i)} \\ \hline p^{(i)T} & \alpha^{(i)} \end{array} \right)$$

berechnet.

Dies entspricht dem linearen Gleichungssystem

$$A^{(i)}y = b^{(i)}$$

und der Zielfunktion

$$z^{(i)} = (p^{(i)})^T y^{(i)} - \alpha^{(i)} (= 0)$$

Es ist

$$H^{(0)} = \left(\begin{array}{c|c} A & b \\ \hline p^T & 0 \end{array} \right) = \left(\begin{array}{c|c|c} G & I_M & b \\ \hline c^T & 0 & 0 \end{array} \right)$$

und für $i = 0$ sind y_{k+1}, \dots, y_{k+M} die Basisvariablen. Damit ergibt sich für $i = 0$ die Basislösung

$$y^{(0)} = \begin{pmatrix} 0 \\ b \end{pmatrix} \text{ und } z^{(0)} = 0$$

Simplex-Schritt (Austauschschritt)

!!kann mal jemand dieses verfahren pruefen? da gibt es viele unsicherheiten

1. Prüfe, ob alle $p_l^{(i)} \leq 0$; $l = 1, \dots, M + k$ (Optimalität). Falls ja, ist $y^{(i)} = y^*$ optimale Basislösung. Falls nein, gehe zu 2).
2. Wähle ein $s \in \{1, \dots, M + k\}$ mit $p_s > 0$ (z.B. das mit dem kleinsten Index oder dem größten Betrag)
3. Prüfe, ob alle $a_{ls}^{(i)} \leq 0$; $l = 1, \dots, M$ (Nichtlösbarkeit). Falls ja, so ist kein Austausch mehr erforderlich. Das Problem hat keine Lösung. Falls nein, gehe zu 4).
4. Wähle $r \in \{1, \dots, M\}$ mit !!habe $r \in \{1, \dots, k\}$ nach $r \in \{1, \dots, M\}$ geändert. Wir wählen doch ein Pivot-Element a_{rs} mit $s = \text{Spalte}$ und $r = \text{Reihe}$. Wobei wir M Reihen haben und k bzw. $M + k$ Spalten und in diesem Schritt die *Reihe* und nicht die Spalte wählen.

$$a_{rs}^{(i)} > 0,$$

$$a_{rs}^{(i)} b_l^{(i)} - a_{ls}^{(i)} b_r^{(i)} \geq 0; l = 1, \dots, r - 1, r + 1, \dots, M \text{ (Zulässigkeit)}$$

!!änderung von $l = 1, \dots, k$ nach $l = 1, \dots, r - 1, r + 1, \dots, M$. zum einen ist es gut möglich (wie auch in diesem beispiel), dass wir weniger nebenbedingungen als variablen haben ($M < k$) und b deshalb nur M elemente hat und $b_l, l = k$ gar nicht definiert ist. zum anderen wählen wir ja eine zeile und keine spalte. wir haben M zeilen und nicht k zeilen... ausserdem brauchen wir den fall $l = r$ nicht überprüfen, da $a_{rs}^{(i)} b_r^{(i)} - a_{rs}^{(i)} b_r^{(i)} = 0 \geq 0$ immer erfüllt ist. zudem: siehe 3.2.6 zusammenfassung. dort hat er $r \in 1, \dots, M$ definiert.

5. Man wende auf $H^{(i)}$ einen Jordanschen Austauschschritt mit Pivot-Element $a_{rs}^{(i)}$ an. Die neue Matrix ist $H^{(i+1)}$.
6. Man lese die Basisindizes für $H^{(i+1)}$ ab. s ist neuer Basisindex. Alle Basisindizes von $H^{(i)}$ sind auch Basisindizes von $H^{(i+1)}$ außer t . Dabei ist t dadurch charakterisiert, dass in der t -Spalte von $A^{(i)}$ der r -te Einheitsvektor steht.
7. Man lese die neue Basislösung $y^{(i+1)}$ ab.

Alternativer Simplex-Schritt (entliehen aus Operations Research)

1. Prüfe, ob alle $p_l^{(i)} \leq 0$; $l = 1, \dots, M + k$ (Optimalität). Falls ja, ist $y^{(i)} = y^*$ optimale Basislösung. Falls nein, gehe zu 2).
2. Wähle ein s (Spalte) mit $p_s = \max \{p_s \mid s = 1, \dots, M + k; p_s > 0\}$ (Spalte)
3. Wähle ein r (Zeile/Reihe) mit $\frac{b_r}{a_{rs}} = \min \left\{ \frac{b_r}{a_{rs}} \mid r = 1, \dots, M; \frac{b_r}{a_{rs}} \geq 0 \right\}$

4. Man wende auf $H^{(i)}$ einen Jordanschen Austauschschritt mit Pivot-Element $a_{rs}^{(i)}$ an. Die neue Matrix ist $H^{(i+1)}$.
5. Man lese die Basisindizes für $H^{(i+1)}$ ab. s ist neuer Basisindex. Alle Basisindizes von $H^{(i)}$ sind auch Basisindizes von $H^{(i+1)}$ außer t . Dabei ist t dadurch charakterisiert, dass in der t -Spalte von $A^{(i)}$ der r -te Einheitsvektor steht.
6. Man lese die neue Basislösung $y^{(i+1)}$ ab.

3.2.5 Theoretischer Hintergrund des Simplex-Verfahrens

Satz 3.2.1

Vorgelegt sei eine lineare Optimierungsaufgabe $z(x) = c^T x \stackrel{!}{=} \max$ unter den Nebenbedingungen $Gx + w = b; x \geq 0$.

Dann gilt:

1. Jedes Maximum der Zielfunktion wird auf dem Rand von Z , dem Polyeder der zulässigen Lösungen, angenommen, falls $c \neq 0$.
2. Das Maximum der Zielfunktion wird in einer Ecke, d.h. in einer zulässigen Basislösung angenommen (ohne Beweis).

Beweis zu 1):

Eine differenzierbare Funktion kann ihr Maximum im Inneren einer Menge nur dann annehmen, falls der Gradient verschwindet (Satz).

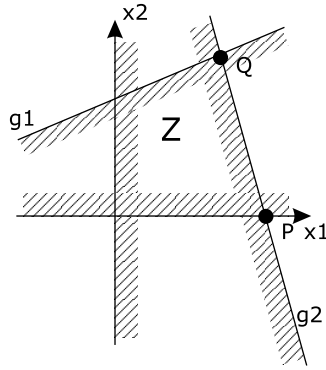
Für $z(x) = c^T x$ bedeutet dies

$$0 \stackrel{!}{=} \nabla z(x) = c^T$$

Dies ist nicht möglich!

Definition 3.2.3 (Der Gradient von f in a)

$$\text{grad } f(a) = \begin{pmatrix} \delta_1 f(a) \\ \vdots \\ \delta_n f(a) \end{pmatrix} := \nabla f(a)$$



Jeder Ecke entspricht genau eine Basislösung, z.B. in $P(x_1 = 0, g_2 = 0)$ oder $Q(g_1 = 0, g_2 = 0)$

Entartete und nicht entartete Basislösungen

Gemäß Definition nennen wir $y_0 \in \mathbb{R}^{M+k}$ eine Basislösung von $Ay = b$, $A \in \mathbb{R}^{M,k}$, $b \in \mathbb{R}^M$ wenn höchstens M Komponenten $\neq 0$ sind.

Grundsätzlich können die folgenden zwei Fälle auftreten.

Beispiel 3.2.1 (Entartete und nicht entartete Basislösungen)

Optimalitätskriterium

Wir betrachten eine lineare Optimierungsaufgabe in der Form

$$Ay = b, y \geq 0, \text{rg}(A) = M, z(y) = p_0 + p^T y \text{ (Normalform 3)}$$

Zusätzlich gelte y^* sei eine zulässige Basislösung; $z(y)$ sei als Funktion der Nichtbasisvariablen von y^* geschrieben, d.h. $z(y) = \sum_{l=1}^{M+k} p_l y_l$ mit $p_l = 0$, falls l Basisindex von y^* .

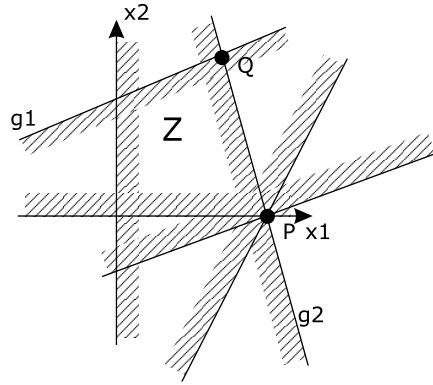
Gilt dann $p_l \leq 0$ für $l = 1, \dots, M+k$, so ist y^* die optimale Lösung.

Beweis 3.2.1

Es gilt $z(y) = p_0 + \sum_{l=1}^{M+k} \underbrace{p_l}_{\leq 0} \underbrace{y_l}_{\geq 0}$ für jede zulässige Lösung y .

Andererseits gilt: $z(y^*) = p_0$, denn für $y^* \neq 0$ gilt $p_l = 0$ und für Nichtbasisindizes ist $y_l^* = 0$.

Zusammen mit Satz 3.2.1 folgt die Behauptung.



Hier liegt in P eine Entartung vor. Die Basislösungen

$$x_1 = 0, g_2 = 0$$

$$g_3 = 0, g_4 = 0$$

$$g_4 = 0, x_1 = 0$$

usw. entsprechen alle dem Punkt P .

Nichtlösbarkeit

Satz 3.2.2

Vorgelegt sei $Ay = b$, $y \geq 0$, $rg(A) = M$, $A \in \mathbb{R}^{M, M+k}$, $z(y) = p_0 + p^T y \stackrel{!}{=} \max$.

Es gebe einen Nichtbasisindex $s \in \{1, \dots, M+k\}$ mit $p_s > 0$ und $a_{is} \leq 0$ für $i = 1, \dots, M$. Dann besitzt die lineare Optimierungsaufgabe keine Lösung.

Beweis 3.2.2

Es sei $y(\lambda_0)$ die Lösung $Ay = b$ mit $y_s(\lambda_0) = \lambda_0 \geq 0$ und $y_l(\lambda_0) = 0$ für alle Nichtbasisindizes. Für $\lambda \geq \lambda_0$ finden wir eine Lösung von $Ay = b$ mit $y_l(\lambda) = 0$ für Nichtbasisindizes aus

$$\left(\begin{array}{c|ccc} a_{1S} & l_{i1} & \dots & l_{iM} \\ \vdots & & & \\ a_{MS} & & & \end{array} \right) \begin{pmatrix} y_s \\ y_{i1} \\ \vdots \\ y_{iM} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_M \end{pmatrix}$$

y_{ij} für $j = 1, \dots, M$ sind die Basisvariablen.

$$\Leftrightarrow \begin{pmatrix} l_{i1} & \dots & l_{iM} \end{pmatrix} \begin{pmatrix} y_{i1} \\ \vdots \\ y_{iM} \end{pmatrix} = \begin{pmatrix} b_1 - \underbrace{\lambda}_{\geq 0} \underbrace{a_{1S}}_{\leq 0} \\ \vdots \\ b_M - \underbrace{\lambda}_{\geq 0} \underbrace{a_{MS}}_{\leq 0} \end{pmatrix}$$

Somit folgt $y_{i1}, \dots, y_{iM} \geq 0$, $y_s = \lambda \geq 0$, d.h. $y \geq 0$ für $\lambda \geq \lambda_0$, d.h. $y(\lambda)$ ist zulässig für $\lambda \geq \lambda_0$.

Ferner gilt $z(y(\lambda)) = p_0 + p_s \lambda$, denn y verschwindet für alle Nichtbasisindizes $\neq S$ und p verschwindet für alle Basisindizes, somit folgt

$$z(y(\lambda)) \rightarrow +\infty \text{ für } \lambda \rightarrow +\infty$$

d.h. es gibt kein Maximum.

Übergang von einer Basislösung zur nächsten

Satz 3.2.3

Vorgelegt sei $Ay = b$, $y \geq 0$, $A \in \mathbb{R}^{M, M+k}$, $y \in \mathbb{R}^{M+k}$, $b \in \mathbb{R}^k$, $rg(A) = M$ mit der Zielfunktion $z(y) = p_0 + p^T y \stackrel{!}{=} \max$.

Seien \bar{y} und \tilde{y} zulässige Basislösungen und seien

$$\bar{J} = \{\bar{j}_1, \dots, \bar{j}_M\} \subset \{1, \dots, M+k\}$$

$$\tilde{J} = \{\tilde{j}_1, \dots, \tilde{j}_M\} \subset \{1, \dots, M+k\}$$

die Indexmengen der Basisvariablen.

Es gelte $\bar{p}_s > 0$ und $\bar{a}_{rs} > 0$, $\bar{b}_i \bar{a}_{rs} - \bar{b}_r \bar{a}_{is} \geq 0$, $i = 1, \dots, M$.

Dann ist die durch einen Austauschschritt des Simplex-Verfahrens erhaltene Basislösung \tilde{y} wieder zulässig und es gilt:

$$z(\tilde{y}) \geq z(\bar{y})$$

Beweis 3.2.3

Zunächst gilt \tilde{y} ist eine zulässige Basislösung $\Leftrightarrow \tilde{b} \geq 0$.

Nach Definition des Austauschschrittes und der Voraussetzung gilt:

$$0 \leq \frac{\overbrace{\bar{b}_i \bar{a}_{rs} - \bar{b}_r \bar{a}_{is}}^{\geq 0}}{\underbrace{\bar{a}_{rs}}_{> 0}} = \bar{b}_i - \frac{\bar{a}_{is}}{\bar{a}_{rs}} \bar{b}_r = \tilde{b}_i$$

Für den Wert der Zielfunktion gilt:

$$-z(\tilde{y}) = -\bar{c}_0 - \frac{\overbrace{\bar{p}_s}^{> 0}}{\underbrace{\bar{a}_{rs}}_{> 0}} \underbrace{\bar{b}_r}_{\geq 0} \leq -\bar{c}_0 = -z(\bar{y})$$

, d.h. $z(\bar{y}) \leq z(\tilde{y})$

3.2.6 Zusammenfassung

Gegeben sei ein Lineares Optimierungsproblem der Form

$$Ay = b, y \geq 0, \operatorname{rg}(A) = M, A \in \mathbb{R}^{M, M+k}, b \in \mathbb{R}^M, y \in \mathbb{R}^{M+k}, z(y) = p_0 + p^T y \stackrel{!}{=} \max.$$

y^* sei zulässige Basislösung, d.h. $y^* \in \bar{Z} = \{y \in \mathbb{R}^{M+k} | Ay = b, y \geq 0\}$ und y^* besitzt höchstens M von Null verschiedene Komponenten.

Ferner sei $p_l = 0$, falls l der Basisindex ist.

Fakten:

- Gilt $p_l \leq 0, l = 1, \dots, M+k \Rightarrow y^*$ ist die *optimale Lösung*.
- Existiert ein Nichtbasisindex s mit $a_{is} \leq 0, i = 1, \dots, M$, so ist das lineare Optimierungsproblem *unlösbar*.
- Existieren $r \in \{1, \dots, M\}$ und $s \in \{1, \dots, M+k\}$ mit $p_s > 0, a_{rs} > 0$ und $b_i a_{rs} - b_r a_{is} \geq 0, i = 1, \dots, M$, so ist die durch den Austauschschritt berechnete neue Basislösung \tilde{y} zulässig und es gilt $z(\tilde{y}) \geq z(y)$.

Beispiel 3.2.2

$$\begin{aligned} 2x_1 + x_2 &\leq 200 \\ 2x_1 + 3x_2 &\leq 800 \\ x_2 &\leq 80 \\ x_1 &\leq 90 \\ x_1, x_2 &\geq 0 \end{aligned}$$

Zielfunktion:

$$z(x) = 4x_1 + 3x_2 \stackrel{!}{=} \max.$$

Führe die Schlupfvariablen w_1, \dots, w_4 ein und finde:

$$\begin{aligned} 2x_1 + x_2 + w_1 &= 200 \\ 2x_1 + 3x_2 + w_2 &= 800 \\ x_2 + w_3 &= 80 \\ x_1 + w_4 &= 90 \quad x_1, x_2 \geq 0 \end{aligned}$$

$$A = (G|I) = \begin{pmatrix} 2 & 1 & 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} b = \begin{pmatrix} 200 \\ 800 \\ 80 \\ 90 \end{pmatrix}$$

Basisvariablen: w_1, \dots, w_4

$$H^{(0)} = \left(\begin{array}{c|c|c} G & I & b \\ \hline c^T & 0 & 0 \end{array} \right) = \left(\begin{array}{cccccc|c} 2 & 1 & \bar{1} & \bar{0} & \bar{0} & \bar{0} & 200 \\ 2 & 3 & 0 & 1 & 0 & 0 & 800 \\ 0 & 1 & 0 & 0 & 1 & 0 & 80 \\ \boxed{1} & 0 & 0 & 0 & 0 & 1 & 90 \\ \hline 4 & 3 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \begin{array}{l} \lambda = -2 \\ \lambda = -2 \\ \\ \lambda = -4 \end{array}$$

Basisvariablen: x_1, w_1, w_2, w_3

$$H^{(1)} = \left(\begin{array}{cccccc|c} \bar{0} & \boxed{1} & \bar{1} & \bar{0} & \bar{0} & -2 & 20 \\ 0 & 3 & 0 & 1 & 0 & -2 & 620 \\ 0 & 1 & 0 & 0 & 1 & 0 & 80 \\ 1 & 0 & 0 & 0 & 0 & 1 & 90 \\ \hline 0 & 3 & 0 & 0 & 0 & -4 & -360 \end{array} \right) \begin{array}{l} \lambda = -3 \\ \lambda = -1 \\ \\ \lambda = -3 \end{array}$$

Basisvariablen: x_1, x_2, w_2, w_3

$$H^{(2)} = \left(\begin{array}{cccccc|c} \bar{0} & \bar{1} & 1 & \bar{0} & \bar{0} & -2 & 20 \\ 0 & 0 & -3 & 1 & 0 & 4 & 580 \\ 0 & 0 & -1 & 0 & 1 & \boxed{2} & 60 \\ 1 & 0 & 0 & 0 & 0 & 1 & 90 \\ \hline 0 & 0 & -3 & 0 & 0 & 2 & -420 \end{array} \right) \begin{array}{l} \lambda = 1 \\ \lambda = -2 \\ \\ \lambda = -1 \end{array}$$

Man findet:

Basisvariablen: x_1, x_2, w_2, w_4

$$H^{(3)} = \left(\begin{array}{cccccc|c} \bar{0} & \bar{1} & 0 & \bar{0} & 1 & \bar{0} & 80 \\ 0 & 0 & -1 & 1 & -2 & 0 & 440 \\ 0 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 1 & 30 \\ 1 & 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 0 & 60 \\ \hline 0 & 0 & -2 & 0 & -1 & 0 & -480 \end{array} \right)$$

Wegen $p^{(3)} \leq 0$ ist das Optimalitätskriterium erfüllt.

Setze die Nichtbasisvariablen zu Null und finde

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ w_2 \\ w_4 \end{pmatrix} = \begin{pmatrix} 80 \\ 440 \\ 30 \\ 60 \end{pmatrix}$$

d.h. $x_2 = 80, w_2 = 440, w_4 = 30, x_1 = 60$.

Lösung:

$y^* = y^{(3)} = (60, 80, 0, 440, 0, 30) = (x^*, w^*), z(x^*) = 480$.

Basislösung zu $H^{(0)}$:

$$x_1 = x_2 = 0, w = b$$

3.3 Die Zweiphasenmethode

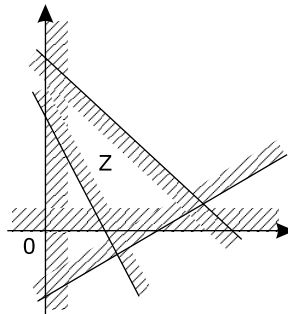
Das Simplexverfahren startet in der Ecke $x = 0, w = b$.

Betrachte $Ay = b, y \geq 0$ und $A = (G|I), y = \begin{pmatrix} x \\ w \end{pmatrix}$ und die Zielfunktion $z(y) = c_0 + c^T y \stackrel{!}{=} \max$.

1. Fall 1: Sei $b \geq 0$. Dann ist $y = (0, b)$ eine zulässige Basislösung, an welcher das Simplexverfahren gestartet werden kann.
2. Fall 2: Es gibt ein $i_0 \in \{1, \dots, M\}$ mit $bi_0 < 0$. Was ist zu tun, falls $b \geq 0$ nicht erfüllt ist?

Geometrisch bedeutet dies das Folgende:

Abbildung 3.2: Geometrische Bedeutung



mit $Z = \{x \in \mathbb{R}^k \mid Gx \leq b, x \geq 0\}$ und $0 \notin Z$.

Benötige eine Basislösung zum Start des Simplex-Verfahrens! Konstruiere mit Hilfe eines Hilfsproblems eine zulässige Basislösung.

3.3.1 Konstruktion zulässiger Basislösungen

Betrachte die Lineare Optimierung in Normalform 2, d.h.

$$Gx + w = b, x \geq 0, w \geq 0$$

$$z(x) = c_0 + c^T x \stackrel{!}{=} \max.$$

Annahme: Es gibt ein i mit $b_i < 0$

Führe die Hilfsvariable ξ ein und betrachte

$$\begin{aligned} \sum_{j=1}^M g_{ij}x_j + w_i &= b_i \forall i \text{ mit } b_i \geq 0 \\ - \sum_{j=1}^M g_{ij}x_j - w_i + \xi_i &= -b_i \forall i \text{ mit } b_i < 0 \end{aligned}$$

$$\text{und } x \geq 0, w \geq 0, \xi = 0$$

sowie

$$z(x) = c_0 + c^T x \stackrel{!}{=} \max$$

Wegen der Bedingung $\xi = 0$ ist dies gleichwertig zum *Ausgangsproblem*.

Hilfsproblem

$$\begin{aligned} \sum_{j=1}^M g_{ij}x_j + w_i &= b_i \forall i \text{ mit } b_i \geq 0 \\ - \sum_{j=1}^M g_{ij}x_j - w_i + \xi_i &= -b_i \forall i \text{ mit } b_i < 0 \end{aligned}$$

$$x \geq 0, w \geq 0, \xi \begin{matrix} \square \\ \geq \end{matrix} 0$$

mit der neuen Zielfunktion

$$\tilde{z} = - \sum_{i \text{ mit } b_i < 0} \xi_i$$

Wobei $-\xi_i$ jeweils aus den entsprechenden Zeilen des entstandenen Gleichungssystems gewonnen wird. Die Zielfunktion des Hilfsproblems ist die Summe der nach $-\xi_i$ aufgelösten Gleichungen.

Man multipliziert also jede Zeile, die die Bedingung $b_i \geq 0$ verletzt, mit -1 und ergänzt dort je Zeile eine Spalte ξ_i , die dem Einheitsvektor e_i entspricht. Wir haben nun ein Gleichungssystem folgender Gestalt:

$$\left(\begin{array}{c|c|c|c} G & I_M & \xi & b \\ \hline \tilde{z} & & 0 & -\tilde{z}(y) \end{array} \right)$$

Wobei ξ die hinzugefügten Spalten und \tilde{y} die Koeffizienten der Zielfunktion \tilde{z} sind. Der Vektor b hat jetzt nur noch nicht-negative Werte.

Eine optimale Lösung des Hilfsproblems ist eine zulässige Basislösung des Ausgangsproblems.

Eigenschaften des Hilfsproblems:

- (x, w, ξ) ist optimale Lösung des Hilfsproblems $\Leftrightarrow \xi = 0$.
- $\xi = 0 \Leftrightarrow (x, w)$ ist zulässige Lösung des Ausgangsproblems.

Genauer gilt: (x, w) ist zulässige Basislösung des Ausgangsproblems, da die optimalen Lösungen von lösbaren Linearen Optimierungen immer in den Ecken des zulässigen Polyeders liegen.

ξ ist genau dann $= 0$, wenn alle ξ_i zu Nichtbasisvariablen transformiert wurden.

3.3.2 Die Zweiphasenmethode

Lösung von allgemeinen Linearen Optimierungsaufgaben:

Falls nicht gilt $b \geq 0$, bilde das zugehörige *Hilfsproblem*. Löse dies mit dem *Simplexverfahren* und erhalte die Lösung $(x^H, w^H, \xi^H), \xi^H = 0$. Dann ist (x^H, w^H) eine zulässige Basislösung des Originalproblems (1-te Phase).

Die zweite Phase besteht dann in der Lösung des Ausgangsproblems mit (x^H, w^H) als zulässige Basislösung.

Beispiel 3.3.1

Betrachte

$$\begin{aligned} 2x_1 + x_2 &\leq 200 \\ -x_1 + x_2 &\leq 40 \\ x_1 + x_2 &\geq 50 \quad , \quad x_1, x_2 \geq 0 \end{aligned}$$

mit der Zielfunktion $z(x) = -x_1 + 2x_2 \stackrel{!}{=} \max..$

Führe Schlupfvariablen ein:

$$\begin{aligned} 2x_1 + x_2 + w_1 &= 200 \\ -x_1 + x_2 + w_2 &= 40 \\ -x_1 - x_2 + w_3 &= -50 \quad \leftarrow b > 0 \text{ hier verletzt!} \end{aligned}$$

$$x_1, x_2, w_1, w_2, w_3 \geq 0, \quad z(x) = -x_1 + 2x_2 \stackrel{!}{=} \max.$$

$y = (0, b)$ ist keine zulässige Basislösung!

Hilfsproblem:

$$\begin{aligned} 2x_1 + x_2 + w_1 &= 200 \\ -x_1 + x_2 + w_2 &= 40 \\ x_1 + x_2 - w_3 + \xi_3 &= 50 \end{aligned}$$

$$x_1, x_2, w_1, w_2, w_3, \xi_3 \geq 0$$

Zielfunktion

$$\tilde{z} = -\xi_3 = x_1 + x_2 - w_3 - 50 \stackrel{!}{=} \max.$$

Lösung des Hilfsproblems mit dem Simplex-Algorithmus.

$$H_H^{(0)} = \left(\begin{array}{cccccc|c} 2 & 1 & 1 & 0 & 0 & 0 & 200 \\ -1 & 1 & 0 & 1 & 0 & 0 & 40 \\ \boxed{1} & 1 & 0 & 0 & -1 & 1 & 50 \\ \hline 1 & 1 & 0 & 0 & -1 & 0 & 50 \end{array} \right) \begin{array}{l} \lambda = -2 \\ \lambda = 1 \\ \lambda = -1 \end{array}$$

Basis: $y = (0, b)$, $z = c^T x + 0 \cdot w$

$$\tilde{z}(0) = -50$$

Basisvariablen von $H_H^{(0)}$: w_1, w_2, ξ_3 .

Nichtbasisvariablen von $H_H^{(0)}$: x_1, x_2, w_3 .

$$H_H^{(1)} = \left(\begin{array}{cccccc|c} 0 & -1 & 1 & 0 & 2 & -2 & 100 \\ 0 & 2 & 0 & 1 & -1 & 1 & 90 \\ 1 & 1 & 0 & 0 & -1 & 1 & 50 \\ \hline 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{array} \right)$$

Wegen $p_H^{(1)} \leq 0$ haben wir die optimale Lösung des Hilfsproblems gefunden.

Basisvariablen: x_1, w_1, w_2

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 100 \\ 90 \\ 50 \end{pmatrix}, \text{ d.h. } w_1 = 100, w_2 = 90, x_1 = 50$$

Lösung des Hilfsproblems: $(x^H, w^H, \xi^H) = (\underbrace{50}_{=x^H}, \underbrace{0, 100, 90, 0}_{=w^H}, \underbrace{0}_{=\xi^H})$.

$(x^H, w^H) = (50, 0, 100, 90, 0)$ ist zulässige Basislösung des Ausgangsproblems.

Finde die aus dem Hilfsproblem resultierende Ausgangsproblem:

$$H^{(0)} = \left(\begin{array}{ccccc|c} 0 & -1 & 1 & 0 & 2 & 100 \\ 0 & 2 & 0 & 1 & -1 & 90 \\ 1 & 1 & 0 & 0 & -1 & 50 \\ \hline & & & & & \end{array} \right) (\hat{=} H_H^{(1)} \text{ ohne } \xi)$$

Basisvariablen: x_1, w_1, w_2 .

Nichtbasisvariablen: x_2, w_3 .

Zielfunktion: $z(x) = -x_1 + 2x_2$

Achtung:

Die Zielfunktion (des Ausgangsproblems) muss in den Nichtbasisvariablen geschrieben werden!

Daher versuchen wir die Basisvariablen der Zielfunktion (des Ausgangsproblems) zu eliminieren, indem wir beliebige Gleichungen aus $H^{(0)}$, die möglichst wenige weitere Basisvariablen enthalten, nach der zu eliminierenden Basisvariable auflösen und dann in die Zielfunktion einsetzen.

Neue Zielfunktion:

$$\begin{aligned} z(x) &= -x_1 + 2x_2 \\ 50 &= x_1 + x_2 - w_3 \\ \Rightarrow -x_1 &= x_2 - w_3 - 50 \\ \Rightarrow z &= x_2 - w_3 - 50 + 2x_2 = 3x_2 - w_3 - 50 \end{aligned}$$

$$z(y) = z(50, 0, 100, 90, 0) = z(0) = -50$$

$z(y) = z(0)$, da z nur noch Nichtbasisvariablen enthält und alle Nichtbasisvariablen von y per Definition = 0 sind.

Die neue Zielfunktion ergänzt die Matrix unseres Problems:

$$H^{(0)} = \left(\begin{array}{ccccc|c} 0 & -1 & 1 & 0 & 2 & 100 \\ 0 & \boxed{2} & 0 & 1 & -1 & 90 \\ 1 & 1 & 0 & 0 & -1 & 50 \\ \hline 0 & 3 & 0 & 0 & -1 & 50 \end{array} \right)$$

$$\tilde{H}^{(0)} = \left(\begin{array}{ccccc|c} 0 & 0 & 1 & \frac{1}{2} & \frac{3}{2} & 145 \\ 0 & 2 & 0 & 1 & -1 & 90 \\ 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & 5 \\ \hline 0 & 0 & 0 & -\frac{3}{2} & \frac{1}{2} & -85 \end{array} \right) \begin{array}{l} \lambda = \frac{1}{2} \\ \lambda = -\frac{1}{2} \\ \lambda = -\frac{3}{2} \end{array}$$

$$H^{(1)} = \left(\begin{array}{cccc|c} 0 & 0 & 1 & \frac{1}{2} & \frac{3}{2} & 145 \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & 45 \\ 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & 5 \\ \hline 0 & 0 & 0 & -\frac{3}{2} & \frac{1}{2} & -85 \end{array} \right)$$

Ein weiterer Simplexschritt liefert $H^{(2)}$.

$$H^{(2)} = \left(\begin{array}{cccc|c} 0 & 0 & \frac{2}{3} & \frac{1}{3} & 1 & \frac{290}{3} \\ 0 & 1 & \frac{1}{3} & \frac{2}{3} & 0 & \frac{280}{3} \\ 1 & 0 & \frac{1}{3} & -\frac{1}{3} & 0 & \frac{160}{3} \\ \hline 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & 0 & -\frac{400}{3} \end{array} \right)$$

Basisvariablen: x_1, x_2, w_3 .

Wegen $p^{(2)} \leq 0$ ist das Optimalitätskriterium erfüllt.

Wir lesen ab:

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} \frac{290}{3} \\ \frac{280}{3} \\ \frac{160}{3} \end{pmatrix}$$

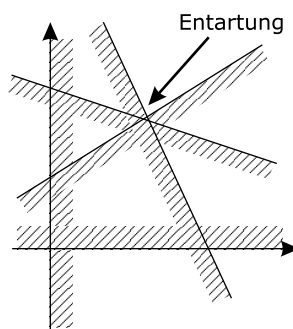
, d.h. $w_3^* = \frac{290}{3}, x_2^* = \frac{280}{3}, x_1^* = \frac{160}{3}$

Lösung:

$$(x^*, w^*) = \left(\frac{160}{3}, \frac{280}{3}, 0, 0, \frac{290}{3} \right) \text{ mit } z(x^*) = \frac{400}{3}$$

3.3.3 Fall der Entartung

Abbildung 3.3: Entartung



Die Ecke ist entartet, da ein Schnittpunkt von 3 Nebenbedingungen im \mathbb{R}^3 vorliegt.

In diesem Fall gehören zu \tilde{x} mehrere Basislösungen, nämlich $g_1 = 0, g_2 = 0$ oder $g_1 = 0, g_3 = 0$ oder $g_2 = 0, g_3 = 0$.

In solch einem Fall kann das Simplex-Verfahren in Zyklen laufen und hängen bleiben. Die Zielfunktion bleibt dadurch immer gleich.

Gegenmassnahme:

Man legt eine Reihenfolge beim Durchlaufen der zugehörigen Basislösungen bei der Wahl des Pivotelementes fest, wenn die Zielfunktion sich nicht ändert.

Konvergenzergebnis:

Läuft das Simplex-Verfahren ohne Zyklen ab und ist die Lineare Optimierungsaufgabe lösbar, so führt das Simplex-Verfahren nach endlich vielen Schritten zum optimalen Ergebnis.

Rechenaufwand:

Für ein Problem in k Variablen und M Nebenbedingungen gilt:

$$\text{Aufwand} \sim \binom{M+k}{M} = \frac{(M+k)!}{M!k!}$$

Sind M und k gleich groß, so folgt:

$$\text{Aufwand} \sim \binom{2k}{k} \sim 2^k$$

d.h. der Aufwand wächst exponentiell in k .

Die Anzahl der Operationen ist nicht polynomial (NP) in k und M .

Man sagt: Der Simplex-Algorithmus ist ein NP-Algorithmus (Nicht-Polynomial).

4 Gewöhnliche Differentialgleichungen

4.1 Separation der Variablen

Beispiel 4.1.1 (Zinsrechnung)

Bei einfacher Zinszahlung für ein Kapital K_0 am Ende eines Jahres gilt für das Kapital $K(t)$ zur Zeit t in Jahren mit Zinssatz q

$$\begin{aligned}\frac{d}{dt}K(t) &= qK_0 \\ K(0) &= K_0\end{aligned}$$

Einfache Integration liefert

$$\begin{aligned}\underbrace{\int_0^t \frac{d}{d\tau}K(\tau)d\tau}_{=K(t)-K(0)} &= \underbrace{\int_0^t qK_0d\tau}_{=[qK_0\tau]_{\tau=0}^t=qK_0t} \\ \Rightarrow K(t) - \underbrace{K(0)}_{=K_0} &= qK_0t \\ K(t) &= K_0 + qK_0t = K_0(1 + qt)\end{aligned}$$

Für den kontinuierlichen Zinseszins (Zinssatz q) gilt

$$\frac{d}{dt}K(t) = qK(t), K(0) = K_0$$

Definition 4.1.1

Gleichungen, welche die Ableitung einer (zeitabhängigen) Funktion in Beziehung zur Funktion selbst setzen, heißen gewöhnliche Differentialgleichungen.

Gibt man überdies noch den Wert der gesuchten Funktion zu einem Zeitpunkt (z.B. $t = 0$) vor, so nennt man dies eine Anfangswertaufgabe.

Lösungsansatz

$$K(t) = K_0 \cdot \exp(qt)$$

Nachweis durch Differentiation

$$\frac{d}{dt}K(t) = \underbrace{K_0 \cdot \exp(qt)}_{=K(t)} \cdot q = q \cdot K(t)$$

$$t = 0 : \quad K(0) = K_0 \cdot \underbrace{\exp(0)}_{=1} = K_0$$

$\Rightarrow (t)$ ist eine Lösung der Anfangswertaufgabe.

Beispiel 4.1.2 (Populationsmodelle)

$x(t)$ bezeichne die Anzahl von Individuen zum Zeitpunkt t .

Veränderungsrate $x'(t) \sim$ Populationsumfang $x(t)$, d.h.

$$\begin{aligned} x'(t) &= R \cdot x(t) \quad R \hat{=} \text{Replikationsrate} \\ x(0) &= x_0 \geq 0 \end{aligned}$$

Realistischeres Modell („Verhülst-Gleichung“):

$$x'(t) = Rx(t) - \frac{R}{K}x(t)^2 \quad , \quad K > 0 \text{ modelliert die Umweltbedingungen}$$

Gesucht: Konstruktive Lösungsmethode

Betrache die Anfangswertaufgabe (AWA)

$$x'(t) = f(x(t)), \quad x(0) = x_0$$

mit $f \in C^1(I, \mathbb{R})$, $I \subset \mathbb{R}$ Intervall, $x_0 \in I$.

Sei $\bar{x}(t)$ eine Lösung, und sei $F(x)$ eine Stammfunktion von $\frac{1}{f(x)}$, d.h.

$$F'(x) = \frac{1}{f(x)}, \quad x \in J, \quad J \subset I$$

Dann gilt:

$$(F(\bar{x}(t)))' = \underbrace{F'(\bar{x}(t))}_{=\frac{1}{f(\bar{x}(t))}} \cdot \underbrace{\bar{x}'(t)}_{=f(\bar{x}(t))} = \frac{1}{f(\bar{x}(t))} \cdot f(\bar{x}(t)) = 1$$

Integration liefert

$$F(\bar{x}(t)) = t + c$$

An der Stelle $t = 0$ erhalte man:

$$\underbrace{F(\bar{x}(0))}_{F(x_0)} = 0 + c = c$$

Dies liefert

$$\begin{aligned} F(\bar{x}(t)) &= t + F(x_0) \\ (\star) \quad F(\bar{x}(t)) - F(x_0) &= t \quad \text{„Separation der Variablen“} \end{aligned}$$

(\star) ist eine implizite Gleichung für die gesuchte Lösung $\bar{x}(t)$. Lässt sich diese Gleichung nach $\bar{x}(t)$ auflösen, so hätten wir $\bar{x}(t)$ berechnet.

Beispiel 4.1.3 (Fortsetzung von Beispiel 4.1.2)

$$x'(t) = \underbrace{Rx(t)}_{=f(x(t))}, x(0) = x_0 \geq 0 \quad \text{Populationsmodell}$$

$$f(x) = Rx, F(x) = \int \frac{dx}{f(x)} = \int \frac{dx}{Rx} = \frac{1}{R} \ln(x) + C$$

Separation der Variablen liefert

$$F(\bar{x}(t)) - F(x_0) = t$$

$$\frac{1}{R} \ln(\bar{x}(t)) + C - \left(\frac{1}{R} \ln(x_0) + C\right) = t$$

$$\frac{1}{R} \ln(\bar{x}(t)) - \left(\frac{1}{R} \ln(x_0)\right) = t$$

$$\frac{1}{R} (\ln(\bar{x}(t)) - \ln(x_0)) = t$$

$$\ln\left(\frac{\bar{x}(t)}{x_0}\right) = Rt, \quad \frac{\bar{x}(t)}{x_0} = \exp(Rt)$$

$$\Rightarrow \bar{x}(t) = x_0 \cdot \exp(Rt), t \geq 0, x_0 > 0$$

Hinweis:

Man ist gut beraten, das erhaltene Ergebnis durch Einsetzen in die Differentialgleichung zu prüfen.

In Beispiel 4.1.2 erkennt man, dass $\bar{x}(t) = x_0 \cdot \exp(Rt)$ die Differentiale für alle $t \in \mathbb{R}$ und $x_0 \in \mathbb{R}$ löst.

Differentialgleichung:

$$x'(t) = f(x(t)), f \in C^1(I, \mathbb{R}), I \subset \mathbb{R}$$

„viele Lösungen“

Anfangswertproblem:

$$x'(t) = f(x(t)), x(0) = x_0$$

„eindeutige Lösung“

Lösung der AWA (Anfangswertaufgabe) durch *Separation der Variablen*. Sei $F(x) = \int \frac{1}{f(x)} dx$. Setze $F(\bar{x}(t)) - F(x_0) = t$ und löse diese nach $\bar{x}(t)$ auf. „Freies Rechnen“.

Beispiel 4.1.4

$$x'(t) = Rx(t), x(0) = x_0$$

Lösung

$$\bar{x}(t) = x_0 \cdot \exp(Rt), x_0 \in \mathbb{R}, t \geq 0$$

Beispiel 4.1.5

$$x'(t) = x(t)^2, x(0) = x_0 > 0, f(x) = x^2$$

$$F(x) = \int \frac{1}{x^2} dx = -\frac{1}{x} + c$$

Separation der Variablen:

$$-\frac{1}{\bar{x}(t)} + \frac{1}{x_0} = t$$

$$\frac{1}{\bar{x}(t)} = \frac{1}{x_0} - t = \frac{1 - tx_0}{x_0}$$

$$\bar{x}(t) = \frac{x_0}{1 - tx_0}, 0 \leq t < \frac{1}{x_0}$$

Probe:

$$\bar{x}(0) = \frac{x_0}{1} = x_0, \bar{x}'(t) = -\frac{x_0}{(1 - tx_0)^2}(-x_0) = \frac{x_0^2}{(1 - tx_0)^2} = \bar{x}(t)^2$$

4.2 Lineare Systeme von Differentialgleichungen

$x'(t) = f(x(t))$ heißt linear, falls $f(x) = ax + b$.

Ist $f \in C^1(\mathbb{R}, \mathbb{R})$, so können wir Lösungen durch Separation der Variablen berechnen.

Sei nun $x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_N(t) \end{pmatrix}$ eine vektorwertige Funktion und sei $A \in \mathbb{R}^{N,N}$, $b \in \mathbb{R}^N$. Dann

heißt $x'(t) = Ax(t) + b$ ein *lineares System von Differentialgleichungen* bzw. $x'(t) = Ax(t) + b$, $x(0) = c \in \mathbb{R}^N$ eine *lineare Anfangswertaufgabe*.

Ausgeschrieben:

$$\begin{pmatrix} x'_1(t) \\ x'_2(t) \\ \vdots \\ x'_N(t) \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \dots & a_{NN} \end{pmatrix} \cdot \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_N(t) \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix}$$

$x'(t) = Ax(t) + b$ heißt *homogen*, falls $b = 0$, ansonsten *inhomogen*.

Definition 4.2.1

andere bedeutung hat.

Zu einer Matrix $A \in \mathbb{R}^{N,N}$ heißt ein Vektor $v \in \mathbb{C}^N$ *Eigenvektor* zum *Eigenwert* $\lambda \in \mathbb{C}$, falls $v \neq 0$ und $Av = \lambda v$.

Bemerkung 4.2.1

$\lambda \in \mathbb{C}$ ist ein *Eigenwert* von A genau dann, wenn $p(\lambda) := \det(A - \lambda I) = 0$, denn $Av = \lambda v$, $v \neq 0 \Leftrightarrow (A - \lambda I)v = 0$, $v \neq 0 \Leftrightarrow v \in \text{Ker}(A - \lambda I)$, $v \neq 0 \Leftrightarrow \underbrace{\det(A - \lambda I)}_{=p(\lambda)} = 0$

(„charakteristisches Polynom“).

Berechne dann den *Eigenvektor* v , indem man die Gleichung $(A - \lambda I)v = 0$, z.B. durch *Gauss-Elimination* löst.

Betrachte zunächst homogene lineare Systeme

$$x'(t) = Ax(t), A \in \mathbb{R}^{N,N}, t \in \mathbb{R}$$

Lösung durch einen Exponential-Ansatz:

$$\bar{x}(t) = \begin{pmatrix} \bar{x}_1(t) \\ \vdots \\ \bar{x}_N(t) \end{pmatrix} = \exp(\lambda t) \cdot v = \begin{pmatrix} \exp(\lambda t)v_1 \\ \vdots \\ \exp(\lambda t)v_N \end{pmatrix} \text{ mit } v \neq 0, Av = \lambda v$$

Dann gilt:

$$\begin{aligned}
 \bar{x}'(t) &= (\exp(\lambda t)v)' \\
 &= (\exp(\lambda t))'v \\
 &= \exp(\lambda t) \cdot \underbrace{\lambda v}_{=Av} \\
 &= \exp(\lambda t) \cdot Av \\
 &= A \cdot \exp(\lambda t) \cdot v \\
 &= A \cdot \bar{x}(t)
 \end{aligned}$$

Bilanz:

- Zu jedem reellen Eigenwert λ mit reellem Eigenvektor v erhalten wir eine reelle Lösung.
- Im komplexen Fall erhalten wir eine komplexe Lösung.

Konstruktion von reellen Lösungen im komplexen Fall:

Sei $\lambda = \mu + i \cdot \nu \in \mathbb{C}$ Eigenwert und $v = a + ib \in \mathbb{C}^N$ Eigenvektor von $A \in \mathbb{R}^{N,N}$.

$\bar{x}(t) = \exp(\lambda t)v$ ist eine komplexe Lösung.

Sei $\bar{x}(t) = u(t) + iv(t)$ eine Zerlegung von $\bar{x}(t)$ in Real- und Imaginärteil (\Re und \Im).
Dann gilt:

$$\begin{aligned}
 \bar{x}'(t) &= u'(t) + iv'(t) \\
 \underbrace{A\bar{x}(t)}_{=\bar{x}'(t)} &= A(u(t) + iv(t)) \\
 &= Au(t) + iAv(t), t \in \mathbb{R}
 \end{aligned}$$

Vergleich der Real- und Imaginärteile liefert $u'(t) = A \cdot u(t)$, $v'(t) = A \cdot v(t)$, d.h. findet man eine komplexe Lösung $\bar{x}(t)$ von $x'(t) = Ax(t)$, so sind $\Re(\bar{x}(t))$ und $\Im(\bar{x}(t))$ reelle Lösungen.

Seien $\lambda \in \mathbb{C}$, $v \in \mathbb{C}^N$.

$$\begin{aligned}
 \bar{x}(t) &= \exp(\lambda t)v \\
 &= \exp((\mu + i\nu)t)(a + ib) \\
 &= \exp(\mu t) \cdot \underbrace{\exp(i\nu t)}_{=(\cos(\nu t) + i \sin(\nu t)) \text{ „Eulersche Formel“}} \cdot (a + ib) \\
 &= \exp(\mu t) \cdot (\cos(\nu t) + i \sin(\nu t)) \cdot (a + ib) \\
 &= \underbrace{[\exp(\mu t)(\cos(\nu t)a - \sin(\nu t)b)]}_{=\Re(\bar{x}(t))} + i \cdot \underbrace{[\exp(\mu t)(\sin(\nu t)a + \cos(\nu t)b)]}_{=\Im(\bar{x}(t))}
 \end{aligned}$$

4.2.1 Eigenschaften linearer Differentialgleichungen

Seien $x^i(t)$, $i = 1, \dots, l$ Lösungen von $x'(t) = Ax(t)$. Dann ist auch $\bar{x}(t) = \sum_{i=1}^l \alpha_i x^i(t)$ eine Lösung, denn

$$\begin{aligned} \bar{x}'(t) &= \left(\sum_{i=1}^l \alpha_i x^i(t) \right)' \\ &= \sum_{i=1}^l \alpha_i \cdot \underbrace{(x^i(t))'}_{=Ax^i(t)} \\ &= \sum_{i=1}^l \alpha_i \cdot Ax^i(t) \\ &= A \cdot \left(\sum_{i=1}^l \alpha_i \cdot x^i(t) \right) \\ &= A\bar{x}(t) \end{aligned}$$

für $\alpha_1, \dots, \alpha_l \in \mathbb{R}$

Gilt nun $l = N$ und sind $x^i(0) \in \mathbb{R}^N$, $i = 1, \dots, N$ linear unabhängig, d.h. eine Basis von \mathbb{R}^N , so heißt $\bar{x}(t) = \sum_{i=1}^N \alpha_i x^i(t)$ die allgemeine Lösung von $x'(t) = Ax(t)$. Es lässt sich zeigen, dass dies alle Lösungen sind.

Ist eine AWA $x'(t) = Ax(t)$, $x(0) = c \in \mathbb{R}^N$ vorgegeben und sind $x^i(t)$, $i = 1, \dots, N$ Lösungen von $x'(t) = Ax(t)$ mit $x^i(0)$, $i = 1, \dots, N$ linear unabhängig, so folgt:

Löse das Gleichungssystem:

$$\begin{pmatrix} (x^1(0))_1 & (x^2(0))_1 & \dots & (x^N(0))_1 \\ \vdots & & & \vdots \\ (x^1(0))_N & (x^2(0))_N & \dots & (x^N(0))_N \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_N \end{pmatrix}$$

Dann löst $\bar{x}(t) = \sum_{i=1}^N \lambda_i \cdot x^i(t)$ die AWA, d.h.

$$\bar{x}(0) = \sum_{i=1}^N \lambda_i x^i(0) = c, \quad \bar{x}'(t) = A\bar{x}(t)$$

4.2.2 Kurze Zusammenfassung des Wesentlichen

$$x'(t) = Ax(t), \quad A \in \mathbb{R}^{N,N}$$

Berechnung der allgemeinen Lösung

1. Berechne alle Eigenwerte und Eigenvektoren von A
2. Für jeden reellen Eigenwert λ und Eigenvektor v ist $\bar{x}(t) = \exp(\lambda t)v$ eine Basislösung.
3. Für jeden komplexen Eigenwert $\lambda = \mu + i\nu$ und Eigenvektor $v = a + ib$ sind

$$\begin{aligned}\bar{x}^1(t) &= \exp(\mu t) [\cos(\nu t)a - \sin(\nu t)b] \\ \bar{x}^2(t) &= \exp(\mu t) [\cos(\nu t)b + \sin(\nu t)a]\end{aligned}$$

Basislösungen.

4. Es seien auf diese Weise die Basislösungen $\bar{x}^i(t)$, $i = 1, \dots, l$ bestimmt.

$$\text{Lösungsgesamtheit: } \mathbb{L} = \left\{ \bar{x}(t) \mid \bar{x}(t) = \sum_{i=1}^l \alpha_i \bar{x}^i(t), \alpha_1, \dots, \alpha_l \in \mathbb{R} \right\}$$

Gilt $l = N$, so sind dies alle Lösungen.

5. Lösung der Anfangswertaufgabe: $x'(t) = Ax(t)$, $x(0) = c$.

Bestimme aus \mathbb{L} die $\alpha_1, \dots, \alpha_l$ so, dass gilt $\sum_{i=1}^l \alpha_i \bar{x}^i(0) = c$.

Beispiel 4.2.1

$$x'(t) = Ax(t), A = \begin{pmatrix} 1 & -2 & 0 \\ 2 & 0 & 1 \\ -4 & -2 & -1 \end{pmatrix}$$

1. a) Berechnung der Eigenwerte:

$$\begin{aligned}p(\lambda) &= \det(A - \lambda I) = \det \begin{pmatrix} 1 - \lambda & -2 & 0 \\ 2 & -\lambda & 1 \\ -4 & -2 & -1 - \lambda \end{pmatrix} \\ &= (1 - \lambda) \cdot \det \begin{pmatrix} -\lambda & 1 \\ -2 & -1 - \lambda \end{pmatrix} - (-2) \cdot \det \begin{pmatrix} 2 & 1 \\ -4 & -1 - \lambda \end{pmatrix} \\ &= (1 - \lambda) \cdot \underbrace{[(-\lambda)(-1 - \lambda) + 2]}_{\lambda + \lambda^2 + 2} + 2 \cdot \underbrace{[2(-1 - \lambda) + 4]}_{-2 - 2\lambda + 4 = 2 - 2\lambda = 2 \cdot (1 - \lambda)} \\ &= (1 - \lambda) [\lambda + \lambda^2 + 2 + 4] \\ &= (1 - \lambda)(\lambda^2 + \lambda + 6)\end{aligned}$$

Nullstellen:

$$\lambda_1 = 1 \quad \lambda_{2,3} = -\frac{1}{2} \pm i\sqrt{\alpha}, \quad \alpha = \frac{23}{4}$$

b) Berechnung der Eigenvektoren:

$$\lambda_1 = 1 : \begin{pmatrix} 0 & -2 & 0 \\ 2 & -1 & 1 \\ -4 & -2 & -2 \end{pmatrix} \cdot \begin{pmatrix} (v^1)_1 \\ (v^1)_2 \\ (v^1)_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Gauss-Elimination liefert $v^1 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$, falls $(v^1)_1 = 1$ fixiert wird.

$$\lambda_2 = -\frac{1}{2} \pm i\sqrt{\alpha} \begin{pmatrix} \frac{3}{2} - i\sqrt{\alpha} & -2 & 0 \\ 2 & +\frac{1}{2} - i\sqrt{\alpha} & 1 \\ -4 & -2 & -\frac{1}{2} - i\sqrt{\alpha} \end{pmatrix} \cdot \begin{pmatrix} (v^2)_1 \\ (v^2)_2 \\ (v^2)_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Gauss-Elimination liefert $v^2 = \frac{1}{12} \begin{pmatrix} 1 - 2i\sqrt{\alpha} \\ -5 - 2i\sqrt{\alpha} \\ 12 \end{pmatrix}$, falls $(v^2)_3 = 1$ normiert wird.

2. und...

3. ...Basislösungen:

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

$$\begin{aligned} \bar{x}^2(t) &= \exp(\mu t) \cdot [\cos(\nu t)a - \sin(\nu t)b] \\ &= \exp(-\frac{1}{2}t) \cdot \left[\cos(\sqrt{\alpha}t) \cdot \frac{1}{12} \begin{pmatrix} 1 \\ -5 \\ 12 \end{pmatrix} - \sin(\sqrt{\alpha}t) \cdot \frac{1}{12} \begin{pmatrix} -2\sqrt{\alpha} \\ -2\sqrt{\alpha} \\ 0 \end{pmatrix} \right] \end{aligned}$$

$$\begin{aligned} \bar{x}^3(t) &= \exp(\mu t) [\cos(\nu t)b + \sin(\nu t)a] \\ &= \exp(-\frac{1}{2}t) \left[\cos(\sqrt{\alpha}t) \cdot \frac{1}{12} \begin{pmatrix} -2\sqrt{\alpha} \\ -2\sqrt{\alpha} \\ 0 \end{pmatrix} + \sin(\sqrt{\alpha}t) \cdot \frac{1}{12} \begin{pmatrix} 1 \\ -5 \\ 12 \end{pmatrix} \right] \end{aligned}$$

4. $\mathbb{L} = \{ \bar{x}(t) \mid \bar{x}(t) = \alpha_1 \bar{x}^1(t) + \alpha_2 \bar{x}^2(t) + \alpha_3 \bar{x}^3(t), \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R} \}$.

Dies sind alle Lösungen, da $N = 3$.

5.

$$x'(t) = \begin{pmatrix} 1 & -2 & 0 \\ 2 & 0 & 1 \\ -4 & -2 & -1 \end{pmatrix} x(t), \quad x(0) = c = \begin{pmatrix} 3 \\ -10 \\ 22 \end{pmatrix}$$

$$\bar{x}^1(0) = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \quad \bar{x}^2(0) = \frac{1}{12} \begin{pmatrix} 1 \\ -5 \\ 12 \end{pmatrix}, \quad \bar{x}^3(0) = \frac{1}{12} \begin{pmatrix} -2\sqrt{\alpha} \\ -2\sqrt{\alpha} \\ 0 \end{pmatrix}$$

$$(\bar{x}^1(0), \bar{x}^2(0), \bar{x}^3(0)) \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{12} & -\frac{1}{6}\sqrt{\alpha} \\ 0 & -\frac{5}{12} & -\frac{1}{6}\sqrt{\alpha} \\ -2 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = c = \begin{pmatrix} 3 \\ -10 \\ 22 \end{pmatrix}$$

Löse dies durch Gauss-Elimination und finde $\alpha_1 = 1$, $\alpha_2 = 24$, $\alpha_3 = 0$.

Lösung:

$$\begin{aligned} \bar{x}(t) &= \bar{x}^1(t) + 24\bar{x}^2(t) + 0 \cdot \bar{x}^3(t) \\ &= \exp(t) \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + 2\exp(-\frac{1}{2}t) \left[\cos(\sqrt{\alpha}t) \begin{pmatrix} 1 \\ -5 \\ 12 \end{pmatrix} - \sin(\sqrt{\alpha}t) \begin{pmatrix} -2\sqrt{\alpha} \\ -2\sqrt{\alpha} \\ 0 \end{pmatrix} \right] \end{aligned}$$

Lösung inhomogener linearer Systeme:

Betrachte $x'(t) = Ax(t) + b$, $A \in \mathbb{R}^{N,N}$, $b \in \mathbb{R}^N$ (\star).

Sei $x^P(t)$ eine Lösung und $x(t)$ irgendeine Lösung von $x'(t) = Ax(t)$. Dann gilt:

$$\begin{aligned} \bar{x}(t) &= x^P(t) + x(t) \text{ löst } (\star), \text{ denn} \\ \bar{x}'(t) &= (x^P(t))' + x'(t) = Ax^P(t) + b + Ax(t) = A(x^P(t) + x(t)) + b \end{aligned}$$

$x^P(t)$ heißt partikuläre Lösung von (\star).

Bemerkung 4.2.2

Man erhält die allgemeine Lösung von (\star), indem man zu einer Lösung $x^P(t)$ von (\star) die allgemeine Lösung des homogenen Systems $x'(t) = Ax(t)$ hinzuaddiert, d.h.

$$\mathbb{L} = \left\{ \bar{x}(t) \mid \bar{x}(t) = x^P(t) + \sum_{i=1}^N \alpha_i \bar{x}^i(t), \alpha_1, \dots, \alpha_N \in \mathbb{R} \right\}$$

mit $x^P(t)$ Lösung von (\star) und $\bar{x}^i(t)$ Lösung von $x'(t) = Ax(t)$, $i = 1, \dots, N$.

Bestimmung der partikulären Lösung $x^P(t)$:

Ist $A \in \mathbb{R}^{N,N}$ invertierbar, so gilt $x^P(t) = -A^{-1}b$, denn $(x^P(t))' = 0 = -b + b = A \underbrace{(-A^{-1}b)}_{=x^P(t)} + b = Ax^P(t) + b$, $t \in \mathbb{R}$.

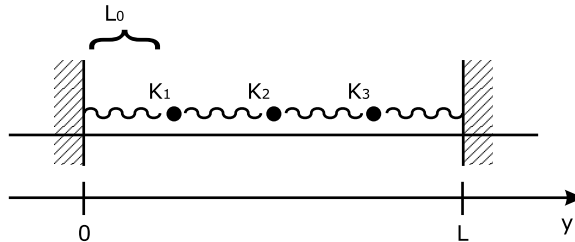
Berechnung der allgemeinen Lösung von (\star)

1. Berechne die allgemeine Lösung von $x'(t) = Ax(t)$.
2. Setze $x^P(t) = -A^{-1}b$, falls A invertierbar ist.
3. $\bar{x}(t) = x^P(t) + x(t)$

Beispiel 4.2.2 (Differentialgleichung)

Drei gleich schwere Kugeln K_1, K_2, K_3 der Masse 1 seien durch Federn gekoppelt (wie im Bild).

Abbildung 4.1: Durch Federn gekoppelte Kugeln



Die Federn haben die Ruhelänge L_0 und die Federkonstante c_i , $i = 0, 1, 2, 3$; $c_i > 0$, $i = 1, 2, 3$.

$y_i(t)$ bezeichne die Position der i -ten Kugel zum Zeitpunkt t .

Modelliert man dieses System nach den Newtonschen Prinzip, so erhält man ein „Differentialgleichungssystem 2-ter Ordnung“:

$$\begin{aligned} y_1''(t) &= (c_0 - c_1)L_0 + (c_0 + c_1)y_1(t) - c_2y_2(t) \\ y_2''(t) &= (c_1 - c_2)L_0 - c_1y_1(t) + (c_1 + c_2)y_2(t) - c_3y_3(t) \\ y_3''(t) &= (c_2 - c_3)L_0 + c_3L - c_2y_2(t) + (c_2 + c_3)y_3(t) \end{aligned}$$

Mit $y(t) = (y_1(t), y_2(t), y_3(t))$ erhalten wir

$$y''(t) = \underbrace{\begin{pmatrix} (c_0 - c_1)L_0 \\ (c_1 - c_2)L_0 \\ (c_2 - c_3)L_0 + c_3L \end{pmatrix}}_{=d \in \mathbb{R}^3} = \underbrace{\begin{pmatrix} (c_0 + c_1) & -c_2 & \\ -c_1 & (c_1 + c_2) & -c_3 \\ & -c_2 & c_2 + c_3 \end{pmatrix}}_{=C \in \mathbb{R}^{3,3}} y(t) = d + Cy(t)$$

Setze $g(t) = y'(t) = (y_1'(t), y_2'(t), y_3'(t)) \in \mathbb{R}^6$ und $x(t) = (y(t), g(t)) \in \mathbb{R}^6$

$$\begin{aligned} x'(t) &= \begin{pmatrix} y'(t) \\ g'(t) \end{pmatrix} = \begin{pmatrix} g(t) \\ y''(t) \end{pmatrix} = \begin{pmatrix} g(t) \\ d + Cy(t) \end{pmatrix} = \\ &= \underbrace{\begin{pmatrix} 0 & I_3 \\ C & 0 \end{pmatrix}}_{=A} \cdot \underbrace{\begin{pmatrix} y'(t) \\ g(t) \end{pmatrix}}_{=x(t)} + \underbrace{\begin{pmatrix} 0 \\ d \end{pmatrix}}_{=b} = Ax(t) + b, \quad A \in \mathbb{R}^{6,6}, b \in \mathbb{R}^6 \end{aligned}$$

Die ist ein 6-dimensionales lineares inhomogenes Differentialgleichungssystem. (Könnte man sich auch gut als Klausuraufgabe vorstellen...)

Die allgemeine Lösung finden wir als eine partikuläre Lösung x^P des inhomogenen Systems plus die Gesamtheit der Lösungen des homogenen Systems $x'(t) = Ax(t)$.

Da A invertierbar ist, finden wir $x^P(t) = -A^{-1}b$. x^P heißt Ruhelage des Systems.

Lösungen des homogenen Systems

Es lässt sich zeigen, dass A rein imaginäre Eigenwerte besitzt.

Diese haben die Form $\pm iw_1, \pm iw_2, \pm iw_3$, $w_1, w_2, w_3 > 0$.

Die zugehörigen Eigenvektoren seien $a_i \pm ib_i$, $i = 1, 2, 3$.

Dies liefert die Lösungen:

$$\begin{aligned}x_i^1(t) &= \cos(w_it)a_i - \sin(w_it)b_i, \\x_i^2(t) &= \cos(w_it)b_i + \sin(w_it)a_i, \quad i = 1, 2, 3\end{aligned}$$

Die allgemeine Lösung des homogenen Systems:

$$x(t) = \sum_{i=1}^3 \alpha_i x_i^1(t) + \beta_i x_i^2(t), \quad \alpha_i, \beta_i \in \mathbb{R}, \quad i = 1, 2, 3$$

Die allgemeine Lösung des inhomogenen Systems $x'(t) = Ax(t) + b$ lautet:

$$\bar{x}(t) = x^P(t) + x(t) = -A^{-1}b + \sum_{i=1}^3 \alpha_i x_i^1(t) + \beta_i x_i^2(t), \quad \alpha_i, \beta_i \in \mathbb{R}, \quad i = 1, 2, 3$$

Dies entspricht Schwingungen um die Ruhelage.

Die Koeffizienten α_i, β_i , $i = 1, 2, 3$ wurden durch eine Anfangsvorgabe $x(0) = e$ bestimmt.

4.3 Existenz und Eindeutigkeit von Lösungen

Vorgelegt sei eine skalare Anfangswertaufgabe

$$x'(t) = f(x(t)), x(0) = x_0 \in \mathbb{R}$$

Definition 4.3.1

$f : I \rightarrow \mathbb{R}$, $I \subset \mathbb{R}$ Intervall genügt in I einer lokalen Lipschitzbedingung, wenn es zu jedem $y_0 \in I$ eine Umgebung $U = U(y_0)$ und ein $L = L(y_0) \geq 0$ gibt, so dass gilt:

$$|f(y) - f(\tilde{y})| \leq L|y - \tilde{y}| \text{ für } y, \tilde{y} \in U \quad \text{„lokale Lipschitzbedingung für } f\text{“}$$

Satz 4.3.1

Ist $I \subset \mathbb{R}$ ein offenes Intervall und besitzt f in I eine stetige Ableitung, so genügt f in I einer lokalen Lipschitzbedingung.

Satz 4.3.2 (Lokaler Existenz- und Eindeutigkeitssatz; Satz von Picard-Lindelöf)

Vorgelegt sei für $f : I \rightarrow \mathbb{R}$ die Anfangswertaufgabe $x'(t) = f(x(t))$, $x(0) = x_0 \in I$.

Genügt f in I einer lokalen Lipschitzbedingung, so ist das Anfangswertproblem lokal eindeutig lösbar, d.h. es gibt ein $\epsilon > 0$, so dass $\bar{x}(t)$, $t \in]-\epsilon, \epsilon[$ eine Lösung des Anfangswertproblems ist.

Beispiel 4.3.1

$$x'(t) = \cos(x(t)), x(0) = 5$$

Wir setzen $f(x) = \cos(x)$, $f \in C^1(\mathbb{R}, \mathbb{R})$ nach Satz refsatz431 genügt f einer lokalen Lipschitzbedingung in \mathbb{R} . Also besitzt dieses Problem lokal genau eine Lösung auf einem t -Intervall der Form $]-\epsilon, \epsilon[$, $\epsilon > 0$

Satz 4.3.3

Vorgelegt sei $x'(t) = f(x(t))$, $x(0) = x_0 \in \mathbb{R}$, $f : \mathbb{R} \rightarrow \mathbb{R}$.

f sei lokal lipschitz-stetig auf ganz \mathbb{R} und $\alpha, \beta \geq 0$ mögen existieren mit

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

Dann existiert die Lösung $\bar{x}(t)$ von (t) global, d.h. für $t \in \mathbb{R}$.

Beispiel 4.3.2

$$x'(t) = \cos(x(t)), x(0) = 5$$

$$f(x) = \cos(x)$$

$$|f(x)| \leq |\cos(x)| \leq 1 = 0 \cdot |x| + 1$$

Also ist Satz 4.3.3 mit $\alpha = 0, \beta = 1$ anwendbar und liefert die Existenz der Lösung $\bar{x}(t)$ für $t \in \mathbb{R}$.

$$\begin{aligned}x'(t) &= ax(t) + b, a, b \in \mathbb{R}, x(0) = 7 \\f(x) &= ax + b \\|f(y) - f(\tilde{y})| &= |ay + b - (a\tilde{y} + b)| = |a(y - \tilde{y})| = |a| |y - \tilde{y}|, y, \tilde{y} \in \mathbb{R}\end{aligned}$$

f ist lokal Lipschitz-stetig.

$|f(x)| = |a \cdot x + b| \leq |a| |x| + |b|$, d.h. Satz 4.3.3 ist erfüllt mit $\alpha = |a|, \beta = |b|$ und liefert die Existenz einer Lösung $\bar{x}(t)$ für $t \in \mathbb{R}$.

Noch zu analysieren bleibt die Frage der Lösbarkeit für lineare Systeme der Form

$$x'(t) = Ax(t) + b, x(0) = x_0 \in \mathbb{R}^N, A \in \mathbb{R}^{N \times N}, b \in \mathbb{R}^N, x(t) = (x_1(t), \dots, x_N(t))$$

Es lässt sich zeigen, dass diese Aufgabe für jedes $x_0 \in \mathbb{R}^N$ genau eine Lösung $\bar{x}(t)$ besitzt.

Diese Lösung existiert für $t \in \mathbb{R}$.

4.4 Numerische Verfahren für Anfangswertaufgaben

(Vorlesungsende)