Mathematik für Ingenieure C3:INF

30. November 2014

Dieses Skript zur Vorlesung Mathematik für Ingenieure C3:INF im Wintersemester 2013/14 bei Dr. Lars Schewe wurde von untenstehenden Studenten erarbeitet. Es ist somit offensichtlich inoffiziell und erhebt weder einen Anspruch auf Korrektheit noch auf Vollständigkeit.

Christian Bay christian.bay@studium.fau.de

Julian Brost julian.brost@studium.fau.de

Karl Werner karl.werner@studium.fau.de

GIT: `https://bitbucket.org/bayliner/mathe-c3-ws1314`

Inhaltsverzeichnis

Teil I

1 Lösen von Gleichungen und der Satz von den impliziten Funktionen

Vorkenntnise:

- Differenzierbarkeit von Funktionen $f : \mathbb{R}^n \to \mathbb{R}^m$ [1]
- Jacobi Matrizen

Beispiele:

$$f(x) = a_0 + a_1 x + a_2 x^2 \tag{1}$$

Explizite Lösung möglich (abc, pq, usw....).

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + a_5 x^5 \tag{2}$$

Keine explizite Lösung möglich. Annährung von Nullstellen mit Newton Verfahren.

Fragen:

- Wie kann man Gleichung $f(x) = 0$ näherungsweise lösen?
- Wie hängen Lösungen von Parametern ab?
- inwieweit ist eine Nullstelle differenzierbar von Parametern abhängig?

[1] Literatur: Mayberg, Vackenauer Band 1 §73 insb. §73.4; Haf, Burg, Wille Band 1 Abschnitt 6.4

1.1 Newton Verfahren in Kurzform

Leitidee: Wir erwarten, dass sich eine stetig differenzierbare Funktion lokal um einen Punkt x' so verhält wie ihre Ableitung an x'.

Aufgabe: Gegeben $F : \mathbb{R}'' \to \mathbb{R}''$
Gesucht: ein Punkt x^* mit $f(x^*) = 0$

Herleitung:

Mit Taylor Entwicklung folgt:

$$f(x^* + h) = f(x^*) + f'(x^*) \cdot h + v(h) \tag{3}$$

Herleitung des Newton Verfahrens:

Somit gilt ungefähr:

$$f(x^* + h) \approx f(x^*) + f'(x^*) \cdot h \tag{4}$$

Hierbei muss h klein sein.

Damit also $f(x^* + h) \approx 0$ gilt, sollten wir h als Lösung von

$$f'(x^*) \cdot h = -f(x) \tag{5}$$

Algorithmus für Newton Verfahren(Pseudocode):

```
1 Gegeben: Ein Orakel fuer f : R^n -> R^n
2 Ein Orakel fuer f'(x) : R^n -> R^(n×h)
3 N ⊂ N
4 x_0 ∈ R^h
5 for (i=0,\_,N-1)
6 h ← Loesung von f'(x_i) · h = -f(x_i)
7 x_{i+1} ← x_i + h
8 Gib x_N aus
```

1.2 Fixpunktoperation

Aufgabe: gegeben eine Funktion $f : \mathbb{R}^n \to \mathbb{R}^n$, gesucht $x^* \in \mathbb{R}$ mit $f(x^*) = x^*$

x^* heisst Fixpunkt von f.

Sei $\mathbb{K}$ entweder $\mathbb{R}$ oder $\mathbb{C}$:

Definition: (Normaler $\mathbb{K}$-Vektorraum). *Ein Paar $(E, ||\cdot||)$ mit $E\mathbb{K}$-Vektorraum und $||\cdot|| \cdot E \to \mathbb{R}$ heisst normierter $\mathbb{K}$-Vektorraum wenn gilt:*

- *(N1)* $\forall x \in E \quad ||x|| >= 0$
- *(N2)* $\forall x \in E \quad ||x|| = 0 \leftrightarrow x = 0$
- *(N3)* $\forall x \in E \forall \lambda \in \mathbb{K}. ||\lambda x|| = |\lambda| ||x||$
- *(N4)* $\forall x, y \in E ||x + y|| <= ||x|| + ||y||$

Beispiele:

- $(\mathbb{R}^n, ||\cdot||_2)$
- $(\mathbb{R}^n, ||\cdot||_1)$
- $(\mathbb{R}^n, ||\cdot||_\infty)$,
- $(C([a,b], \mathbb{R}), ||\cdot||_\infty)$[2], $||f|| := \sup_{x \in [a,b]} |f(x)|$

Die Definition von Konvergenz Cauchy-Folgen, Stetigkeit übertragen sich direkt aus den Definition für den $\mathbb{R}''$.

Definition: (Vollständiger, normierter $\mathbb{K}$-Vektorraum, Banachraum). *Sei $(E, ||\cdot||)$ ein normierter $\mathbb{K}$-Vektorraum, Dann heißt $(E, ||\cdot||)$ ein Banachraum, wenn $(E, ||\cdot||)$ vollständig[3] ist.*

Definition: (Kontrahierende Abbildung). *Sei $(E, ||\cdot||)$ ein normierter Vektorraum, $M \subseteq E$ und $f : M \to M$, dann heißt f kontrahierend, wenn es eine Konstante C $(0 < C < 1)$ gibt, so dass $\forall x, y \in M : ||f(x) - f(y)|| \leq C \cdot ||x - y||$*

Banach'scher Fixpunktsatz. *Sei $(E, \|\cdot\|)$ ein vollständiger normierter Vektorraum, $M \subseteq E$ eine abgeschlossene Teilmenge und $f : M \to M$ eine kontrahierende Abbildung. Dann hat f genau einen Fixpunkt $x^* \in M$.*

[2] Raum der stetigen Funktionen $f : [a, b] \to \mathbb{R}$

[3] Jede Cauchy-Folge konvergiert

Beweis. Sei $x_0 \in M$. Betrachte die Folge (x_n) mit $x_{n+1} = f(x_n)$. Wir zeigen, dass (x_n) konvergiert, indem wir zeigen, dass (x_n) Cauchyfolge ist.

Hilfsaussage: $\forall n \in \mathbb{N} : ||x_{n+1} - x_n|| \leq C^n \cdot ||x_1 - x_0||$.[4]

Für $m \geq n$ gilt:

$$\|x_m - x_n\| \leq \|\sum_{k=n}^{m-1}(x_{k+1} - x_k)\| \tag{6}$$

$$\leq \sum_{k=n}^{m-1}\|(x_{k+1} - x_k)\| \tag{7}$$

$$\leq (\sum_{k=n}^{m-1} C^k)\|(x_1 - x_0)\| \tag{8}$$

$$= C^n(\sum_{k=0}^{m-1-n} C^k)\|x_1 - x_0\| \tag{9}$$

$$\leq \frac{C^n}{1-C}\|x_1 - x_0\| \tag{10}$$

Somit ist die Folge (x_n) eine Cauchy-Folge, und da $(E, \|\cdot\|)$ vollständig ist, auch konvergent. □

Sei nun $x^* \in M$ unser Grenzwert[5]:

$$x^* = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n)^6 = f(\lim_{n\to\infty} x_n) = f(x^*)$$

Satz: (Satz von den impliziten Funktionen). *Sei $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ partiell stetig differenzierbar und $p^* \in \mathbb{R}^m$ und $x^* \in \mathbb{R}^n$ mit $f(p^*, x^*) = 0$. Wenn man die Matrix*

$$J_x(f, p^*, x^*) = (\frac{\partial f_i}{\partial x_i}(p^*, x^*))_{1\leq i\leq n, 1\leq j\leq n} \tag{11}$$

invertierbar ist, dann gibt es offene Umgebungen U um p^* und V um x^*, so dass es zu jedem $p \in U$ ein eindeutiges $x \in V$ gibt mit $f(p, x) = 0$. D.h. es gibt ein $\tilde{x} : U \to V$ mit $f(p, \tilde{x}) = 0$. Zusätzlich ist $\tilde{x}$ stetig differenzierbar.

[4] Beweis ausgelassen, Induktion

[5] Beweis der Eindeutigkeit wird ausgelassen.

[6] Jede kontrahierende Abbildung ist stetig.

Satz: (Umkehrsatz, Satz von lokalen Inversen). *Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ stetig differenzierbar und $x^* \in \mathbb{R}^n$. Wenn nun $f'(x^*)$ invertierbar ist, dan gibt es Umgebungen $U \subseteq \mathbb{R}^n$ um x^* und $V \subseteq \mathbb{R}^n$ um $y^* = f(x^*)$ und eine stetig differenzierbare Funktion $g : V \to U$, so dass für $\tilde{f} = f|_U$ gilt, dass $\tilde{f} \circ g = g \circ \tilde{f} = id$. Die Ableitung von g an der Stelle $y = f(x)$ mit $y \in V$ ist gegeben durch U: $g'(y) = (f'(x))^{-1}$*

Lemma: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^n$ stetig differenzierbar. Wenn es nun ein C gibt mit $\forall x \in U \forall h \in \mathbb{R}^n : ||f'(x)h|| \leq C \cdot ||h||$, dann gilt für alle $x, x + h \in U$: $||f(x + h) - f(x)|| \leq C \cdot ||h||$

Beweis. Ohne Beschränkung der Allgemeinheit gilt: x^*0, $f(x^*) = 0$, $f'(x^*) = id$. Setze $d(x) = f(x) - x$. Da f stetig differenzierbar, existiert eine Umgebung $D \subseteq \mathbb{R}^n$, D abgeschlossen um 0, so dass $\forall x \in D \forall h \in \mathbb{R}^n :\leq \frac{1}{2}||h||$. Sei δ der Radius der größten Kugel um 0 in D. Wähle $V = x : ||x|| \leq \frac{1}{2}\delta$. Sei $g \in V$. Definiere $\Phi_y : D \to D$ mit $\Phi_y(x) = x + y - f(x)$. (Man kann zeigen, dass $\Phi_y(D) \subseteq D$ ist). Es gilt genau dann $f(x) = y$, wenn x Fixpunkt von Φ_y ist. Wir zeigen nun, dass Φ_y eine Kontraktion ist.

$$||\Phi_y(x + h) - \Phi_y(x)|| = ||f(x + h - f(x) - h|| \tag{12}$$
$$\text{(Umstellen, Def von } \Phi\_y) \tag{13}$$
$$= ||(f(x + h) - (x + h)) - (f(x) - x)|| \tag{14}$$
$$(+x - x \text{ addieren, umstellen}) \tag{15}$$
$$= ||d(x + h) = d(x)|| \tag{16}$$
$$\text{(Definition von d)} \tag{17}$$
$$\leq \frac{1}{2}||h|| \tag{18}$$
$$\text{(Folgt aus} \tag{19}$$
$$||d'(x)h|| \leq \frac{1}{2}||h|| \tag{20}$$
$$\text{und dem Lemma)} \tag{21}$$

Somit ist Φ_y eine Kontraktion, also existiert ein eindeutiger Fixpunkt $x \in D$ von Φ_y. Daher gibt es eine eindeutige Lösung der Gleichung $f(x) = y$ in D. Somit gibt es eine Funktion $g : V \to D$ für $g(y) = x$, wenn $f(x) = y$ gilt. Nun wählen wir $U = f^{-1}(V) \cap \mathring{D}$. Wir überspringen den Beweis, dass g stets differenzierbar ist. Die Formel für die Ableitung folgt aus der Kettenregel. □

Satz: (Satz von den impliziten Funktionen). *Sei* $f : \mathbb{R}^m \times \mathbb{R}^n : \mathbb{R}^n$ *partiell stetig differenzierbar. Sei* $p^* \in \mathbb{R}^m$ *und* $x^* \in \mathbb{R}^n$ *mit* $f(p^*, x^*) = 0$

$$J_x(f, p, x) = (\frac{\partial f_i}{\partial x_j}(p, x))_{1 \leq i \leq n, 1 \leq j \leq n} \tag{22}$$

Wenn $J_x(f, p^*, x^*)$ *invertierbar ist, dann gibt es offene Umgebungen* $U \subseteq$ *um* p^* *und* $V \subseteq V$ *um* x^**, so dass es zu jedem* $p \in U$ *ein eindeutiges* $x \in V$ *gibt mit* $f(p, x) = 0$*. D.h. es gibt eine Funktion* $\tilde{x} : U \to V$*, so dass für alle* $p \in U$ $f(p, \tilde{x}(p)) = 0$ *gilt. Zusätzlich ist* $\tilde{x}$ *stetig differenzierbar.*

$$\forall p \in U : \tilde{x}'(p) = -J_x(f, p, \tilde{x}(p))^{-1} \cdot J_p(f, p, \tilde{x}(p)) \tag{23}$$

$$J_p(f, p, x) = (\frac{\partial f_i}{\partial p_j})_{1 \leq i \leq n, 1 \leq j \leq m} \tag{24}$$

$$\Phi(p, x) = (p, f(p, x))\Phi\mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n \tag{25}$$

2 Lokale Optima ohne Nebenbedingungen

Vorkenntnisse:

- Gradient, Hessenmatrix, Taylorentwicklung
- Symmetrische Matrizen diagonalisieren

Ziel löse folgendes Problem:
Gegeben: Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$
Gesucht: $x^* \in \mathbb{R}^n \ f(x^*) = \min_{x \in \mathbb{R}^n} f(x)$

Bemerkung:

- Das Problem hat so nicht notwendigerweise eine Lösung (f kann unbeschränkt sein)
- Analog für Maxima

Hauptanwendung: Ausgleichsprobleme (Least-Squares-Probleme)

$$\min_{x \in \mathbb{R}^n} \sum_{j \in J} ||f_j(x)||_2^2 \tag{26}$$

Wir suchen lokale Lösungen.
Definition: Lokale Optimallösung
Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine Funktion. Denn heisst x^* lokale Optimallösung von $\min_{x \in \mathbb{R}^n} f(x)$, wenn es eine Umgebung U um x^* gibt mit: $\forall x \in U f(x^*) \leq f(x)$.

Satz: Notwendige Optimalitätsbedingung erster Ordnung. *Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar und sei x^* lokales Optimum von $\min_{x \in \mathbb{R}} f(x)$. Dann gilt $\nabla f(x^*) = 0$.*

Beweis. Für $i \in \{1, \ldots, n\}$ setze $g_i(\lambda) = f(x^* + \lambda \cdot e_i)$, wobei e_i der i-te Einheitsvektor ist. Da x^* ein lokales Minimum von f ist, muss es sich bei 0 um ein lokales Minimum von g_i handeln, weshalb $g_i'(0) = 0$ gilt. Ableiten nach λ ergibt:

$$g_i'(0) = \frac{\partial f}{\partial x_i}(x^*)$$

Somit gilt:

$$\nabla f(x^*) = 0$$

□

Definitition: Stationärer Punkt, Kritischer Punkt. *Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar. Dann heisst x^* stationärer Punkt (auch kritischer Punkt) von f, wenn $\nabla f(x^*) = 0$.*

2.1 Weitere Optimalitätskriterien

Vorüberlegung: Betrachte $f(x) = x^T Q x$, wobei $x \in \mathbb{R}^n$ und Q eines symmetrischen $n \times n$ Matrix.
Feststellung: 0 ist stationärer Punkt.

Da Q symmetrisch, existiert eine Basis $v_1, \ldots, v_n$ von Eigenvektoren (zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$).
Sei v ein solcher Eigenvektor (zum Eigenwert λ). Dann gilt:

$$f(v) = v^\top \cdot Q \cdot v = v^\top (\lambda v) \tag{27}$$
$$= \lambda \cdot v^\top \cdot v = \lambda \cdot \underbrace{||v||_2^2}_{\geq 0} \tag{28}$$

Definition. *Sei $Q \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Q heißt*

- positiv semidefinit*, wenn alle Eigenwerte λ von Q nichtnegativ sind (äquivalent $\forall x \in \mathbb{R}^n : x^\top Q x \geq 0$)*
- positiv definit*, wenn alle Eigenwerte λ von Q positiv sind (äquivalent $\forall x \in \mathbb{R}^n \backslash \{0\} : x^\top Q x > 0$)*
- negativ semidefinit*, wenn alle Eigenwerte λ von Q nichtpositiv sind (äquivalent $\forall x \in \mathbb{R}^n : x^\top Q x \leq 0$)*
- negativ definit*, wenn alle Eigenwerte λ von Q negativ sind (äquivalent $\forall x \in \mathbb{R}^n \backslash \{0\} : x^\top Q x < 0$)*

Mitschrift 31.10.2013

Folgerung aus Taylorentwicklung
Sei $f : \mathbb{R}^n \to \mathbb{R}$ zweimal stetig differenzierbar.
Sei $x^* \in \mathbb{R}^n$. Dann gibt es eine stetige Funktion $r(h)$ mit $\lim_{h \to 0} r(h) = 0$, so dass für alle $h \in \mathbb{R}^n$ gilt:

$$f(x^* + h) = f(x^*) + \nabla f(x^*)^T \cdot h + \frac{1}{2} h^T \cdot \underbrace{Hf(x^*)}_{\text{Hessematrix}} \cdot h + ||h||^2 \cdot \rho(h) \tag{29}$$

Satz: (Notwendige Optimalitätsbedingung zweiter Ordnung). *Sei $R \subseteq \mathbb{R}^n$ eine offene Menge. Sei $f : R \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Sei $x^* \in \Omega$ lokales Minimum von f.*
Dann gilt:

a) $\nabla f(x^*) = 0$

b) $Hf(x^*)$ *ist positiv semidefinit*

Beweis. Behauptung a) ist die notwendige Optimalitätsbedingung erster Ordnung. Sei λ_{min} der kleinste Eigenwert von $Hf(x^*)$. Wenn $\lambda_{min} \leq 0$ gilt, ist $Hf(x^*)$ positiv semidefinit. Wenn also $\lambda_{min} = 0$ sind wird fertig, sei also $\lambda_{min} \neq 0$. Da x^* lokales Minimum, können wir annehmen, dass x^* das eindeutige Minimum in Ω ist, sonst verkleinere Ω entsprechend.
Wähle eine Umgebung $U \subseteq \Omega$ um x^*, so dass für alle $h \in \mathbb{R}^n$, so dass $x^* + h \in U$, gilt:

$$|\rho(h) \leq \frac{|\lambda_{min}|}{4} \quad \text{möglich da } \rho \text{ stetig} \tag{30}$$

Wähle $h \neq 0$ so, dass h Eigenvektor zu λ_{min} von $Hf(x^*)$ und $x^* + h \in U$.
Dann gilt:

$$0 \leq f(x^* + h) - f(x*) \quad x^* \text{ lokales Minimum} \tag{31}$$

$$= \underbrace{\nabla f(x^*)^T}_{=0} \cdot h + \frac{1}{2} h^T Hf(x^*) \cdot h + ||h||^2 \cdot \rho(h) \quad \text{Taylor Formel} \tag{32}$$

$$= \frac{1}{2}\lambda_{min}||h||^2 + ||h||^2\rho(h) \quad h \text{ ist Eigenvektor} \tag{33}$$

$$= \frac{1}{2}\lambda_{min}||h||^2 + ||h||^2 \cdot \rho(h) \tag{34}$$

$$\leq \underbrace{\frac{1}{2}||h||^2}_{\geq 0}(\lambda_{min} + \frac{1}{2}|\lambda_{min}|) \quad \text{da } \rho(h) \leq \frac{1}{4}|\lambda_{min}| \tag{35}$$

Es gilt also $\lambda_{min} + \frac{1}{2}|\lambda_{min}| \geq 0$. Daraus folgt aber $\lambda_{min} \geq 0$. □

Einschub: Symmetrische Matrizen haben immer reele Eigenwerte und sind diagonilizierbar. Hessematrix ist immer symmetrisch!

Satz: (Hinreichende Optimalitätsbedinung zweiter Ordnung). *Sei* $\Omega \subseteq \mathbb{R}^n$ *offen und* $f : \Omega \to \mathbb{R}$ *zweimal stetig differenzierbar. Wenn nun* $x^* \in \Omega$ *folgende Bedinungen erfüllt, ist* x^* *lokales Minimum von* f*.*

„Kochrezept für Minima finden “:

a) $\nabla f(x^*) = 0$

b) $Hf(x^*)$ ist positiv definit

2.2 Grundlegende Algorithmen

Definition: Abstiegsrichtug. *Sei* $f : \mathbb{R}^n \mapsto \mathbb{R}$ *stetig differenzierbar und* $x^* \in \mathbb{R}^n$.
Die Richtung $h \in \mathbb{R}^n$ *(genauer* $\frac{h}{\|h\|}$*) heisst Abstiegsrichtung, wenn für die Funktion* $\Phi(t) = f(x^* + th)$ *gilt, dass* $\Phi'(0) < 0$.
Dies ist äquivalent zu $\nabla f(x^*)^T h < 0$.

Schema: Abstiegsverfahren
Input: Funktionsorakel f, Gradientorakel g, Startpunkt x_0

```
1  x ← x_0
2  while true do
3      if g(x) ≈ 0 then
4          return x;
5       Waehle eine Abstiegsrichtung s
6       Bestimme Schrittweite λ
7       x ← x + λ · s
```

Wie finde ich eine Abstiegsrichtung?
Klassische Wahlmöglichkeiten:

a) negativer Gradient

b) Newton-Schritt, d.h. die Lösung von $Hf(x^*)s = -g(x^*)$
Anmerkung: Nicht immer eine Abstiegsrichtung!
Hinreichend: $Hf(x^*)$ positiv definit

Wahl der Schrittweite (Line-Search)
Typische Regeln:

a) Armijo-Regel

b) Powell-Wolfe Bedingung

3 Lokale Optima unter Nebenbedingungen

Problem:

$$\begin{aligned} f : \mathbb{R}^n &\mapsto \mathbb{R} \\ h_i : \mathbb{R}^n &\mapsto \mathbb{R} \text{ für } i = 1, ..., m \\ &\min_x f(x) \\ \forall i \in \{1, ...m\} \quad &h_i(x) = 0 \end{aligned}$$

Beispiel: Finde den Punkt auf der Einheitskreislinie, der am nächsten an $(1, 1)$ liegt.

$$\begin{aligned} \min_x (x-1)^2 + (y-1)^2 \\ x^2 + y^2 - 1 = 0 \end{aligned}$$

Definition: (Zulässige Menge). *Im Problem oben nennen wir die Menge*

$$Z = \{x : \forall i \in \{1, ..., m\} \quad h_i = 0\}$$

die zulässige Menge.

Beispiel für eine "hässliche"zulässige Menge

$$h(x, y) = x^2 - y^2$$

Definition: (Notwendige Optimalitätsbedingung erster Ordnung). *Sei $x^* \in \mathbb{R}^n$ eine lokale Optimallösung von (ECP), sodass die Vektoren $\nabla h_i(x^*)(i = 1, ..., m)$ linear unabhängig sind.*
Dann gibt es ein $\mu \in \mathbb{R}^m$, so dass

a) $\forall i \in \{1, ..., m\} \quad h_i(x^*) = 0$

b) $\nabla f(x^*) + \sum_{i=1}^{m} = h_i(x^*) = 0$

Beispiel: (banal)

$$\begin{aligned} \min_{st} (x-1)^2 + (y-1)^2 \\ y = 0 \end{aligned}$$

Beweis. Beweisskizze für die notwendige Optimalitätsbedingung erster Ordnung:
Betrachte $h(x) := \begin{pmatrix} h_i(x) \\ ... \\ h_m(x) \end{pmatrix}$
Da $\nabla h_i(x^*), ..., \nabla h_m(x^*)$ linear unabhängig sind, hat die Jacobi-Matrix von h an der Stelle x^* vollen Rang.
Somit existiert lokal um x^* eine Auflösungsfunktion, sodass für neue Variablen $y \in \mathbb{R}^n$ (ECP) die Form

$$\begin{aligned} \min_{st^2} & \, g(z) \\ z_1 &= 0 \\ ... & \\ z_m &= 0 \end{aligned}$$

Nach Einsetzen von $z_1 = 0, ..., z_m = 0$ erhalte ich eine Optimierungsproblem ohne Nebenbedingungen, sodass $\frac{\partial g(z^*)}{\partial z_{m+1}} = 0, ..., \frac{\partial g(z^*)}{\partial z_m} = 0$ für ein lokales Minimum z^*x gelten muss.
Rücktransformation dieser Bedingungen ergibt die Behauptung des Satzes.
□

Definition: (Lagrangefunktion). *Sei ein Problem der (ECP) gegeben, dann heisst*

$$L : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R} \qquad L(x, \mu) = f(x) + \sum_{i=1}^{m} \mu_i h_i(x)$$

Lagrangefunktion des Problems

Satz: (Notwendige Optimalitätsbedingung erster Ordnung). *Formulierung mit Lagrangefunktion*
Sei $x^* \in \mathbb{R}^n$ eine lokale Optimallösung von (ECP), sodass $\nabla h_1(x^*), ..., \nabla h_m(x^*)$ linear unabhängig sind. Dann gilt für die Lagrangefunktion $L(x, \mu)$ von (ECP), dass es ein $\mu^* \in \mathbb{R}^m$ gibt mit:

a) $\nabla_\mu L(x^*, \mu^*) = 0$

b) $\nabla_x L(x^*, \mu^*) = 0$

4 Konvexität

4.1 4.1 noch zu definieren

Definition: (Konvexe Menge). *Eine Menge $K \subseteq \mathbb{R}^n$ heisst konvex, wenn für alle $x, y \in K$ gilt, dass auch alle Punkte*

$$z(\lambda) = \lambda x + (1 - \lambda)y \text{ für } \lambda \in [0, 1]$$

in K liegen.

Definition: (Konvexe Funktion). *Sei $f : \mathbb{R}^n \mapsto \mathbb{R}^n$. f ist konvex, wenn für alle $x, y \in \mathbb{R}^n$ gilt, dass für alle $\lambda \in [0, 1]$*

$$\lambda f(x) + (1 - \lambda)f(y) \geq f(\lambda x + (1 - \lambda)y)$$

Definition 4.1. *Eine Menge $K \subseteq \mathbb{R}^n$ heißt konvex, wenn folgende Bed. erfüllt ist:*

$$\forall x, y \in K \forall \lambda \in [0, 1], \lambda x + (1 - \lambda)y \in K$$

Definition 4.2. *Eine konvexe Fkt. $f : \Omega \to R$ (mit $\Omega \subseteq \mathbb{R}^n$ konvex) heißt, konvex, wenn folgende Bed. gilt:*

$$\forall x, y, \forall \lambda \in [0, 1] \lambda f(x) + (1 - \lambda)f(y) \geq f(\lambda + (1 - \lambda)y)$$

Satz 4.1. *Sei $K \subseteq \mathbb{R}^n$ eine konvexe Menge und $f : K \to \mathbb{R}$ eine konvexe Funktion. Dann ist jedes lokale Minimum von f auch globales Minimum.*

Beweis. Sei $x^* \in K$ ein lokales Optimum von f. Sei $x \in K$ ein weiterer Punkt mit $x \neq x^*$. Betrachte nun Punkte $x^\lambda = x^* + \lambda(x - x^*)$. Da K konvex ist, liegen alle Punkte x^λ in K.
Da nun x^* lokales Minimum von f ist, gibt es ein $\overline{\lambda}$, so dass für alle $\lambda \in [0, \overline{\lambda}]$ gilt, dass $f(x) \leq f(x^\lambda)$ gilt. Sei nun $\lambda \in [0, \overline{\lambda}]$. Dann gilt:

$$\begin{aligned} f(x^*) &\leq f(x^\lambda) \\ &= f(x^* + \lambda(x - x^*)) \\ &= f((1 - \lambda)x^* + \lambda x)) \\ &\leq (1 - \lambda f(x^*) + \lambda f(x)) \end{aligned}$$

Nach Umstellen erhalten wir nun $\lambda f(x^*) \leq \lambda f(x)$ Für $\lambda \neq 0$ erhalten wir also $f(x^*) \leq f(x)$. Da x beliebig war, ist x^* also globales Minimum von f auf K. □

Lemma 4.2. *Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine konvexe Fkt. und $c \in \mathbb{R}$. Dann ist die Menge $K = x : f(x) \leq c$ konvex*

Lemma 4.3. *Sei $K, L \subseteq \mathbb{R}^n$ konvexe Mengen, dann ist $K \cap L$ konvex.*

Lemma 4.4.

a) *Seien $g, h : \Omega \to \mathbb{R}$ konvexe Funktionen und $\lambda, \mu >= 0$. Dann $f = \lambda g + \mu h$.*

b) *Seien $g, h : \Omega \to \mathbb{R}$ konvexe Funktionen dann ist auch $f(x) := sup\{g(x), h(x)\}$.*

c) *Sei $h : \Omega \to \mathbb{R}$ eine konvexe Fkt. und $g : \mathbb{R} \to \mathbb{R}$ konvex und monoton wachsend, dann ist $f = g \circ h = g(h(x))$ konvex.*

d) *Sei $g : \Omega \to \mathbb{R}$ eine konvexe Fkt. und $A \in \mathbb{R}^n x \mathbb{R}^n$ und $b \in \mathbb{R}^n$, dann $f(x) := g(Ax + b)$ eine konvexe Fkt.*

Wenn die Funktion f zweimal stetig diff'bar ist, gibt es einen einfachen Test für Konvexität.

Lemma 4.5. *Sei $\Omega \subseteq \mathbb{R}^n$ konvex mit nichtleeren Innern und $f : \Omega \to \mathbb{R}$ stetig auf ganz Ω und zweimal stetig diff'bar auf dem inneren von Ω. Wenn nun für alle x im Inneren von Ω gilt, dass $Hf(x)$ positiv semidefinit ist, dann ist f konvex.*

4.2 Konvexe Optimierungsprobleme

Wenn wir obigen Resultate zusammenfassen, sehen wir, dass wir für konvexe Opt.-probleme gute Aussagen über *globale* Optimalität helfen können.

$$\begin{aligned} &\min f(x) \\ &g_i(x) \leq 0 \quad \forall i \in I \\ &h_j(x) = 0 \quad \forall j \in J \end{aligned}$$

Dabei müssen f und g_i konvexe und die h_j affin lineare Fkt. Betrachte eine Beispielklasse: Linear-Quadratische Programm. Sei $Q \in \mathbb{R}^n x \mathbb{R}^n$ symmetrisch positiv definit, $d \in \mathbb{R}^n$ und sei $A \in \mathbb{R}^n x \mathbb{R}^n$ von vollem Rang und $b \in \mathbb{R}^n$.

$$\begin{aligned} \min \frac{1}{2} * x^t Q x + d^t x \\ Ax = b \end{aligned}$$

Sei also $L(x, \mu)$ sei die Lagrangefkt, unseres Problems:

$$L(x, \mu) = \frac{1}{2} x^T Q x + d^T x + \mu^T (b - Ax)$$

Wenn wir $\nabla L = 0$ lösen wollen, erhalten wir

$$0 = \nabla L(x, \mu) = \begin{pmatrix} x^T Q + d^T + \mu^T A \\ b - Ax \end{pmatrix}$$

Das führt auf folgendes (*lineares!*) Gleichungssystem
$\begin{pmatrix} Q & A^T \\ A & T \end{pmatrix} \times \begin{pmatrix} x \\ \mu \end{pmatrix} = \begin{pmatrix} -d \\ b \end{pmatrix}$
Da nun Q und A vollen Rang haben, sehen wir, dass dieses Gleichungssystem genau eine Lsg. hat.

Mitschrift vom 15.11.13

Konvexe Optimierung (Fortsetzung) .... TODO.

$$\begin{array}{ll} \min\limits_x(f(x)) & f \text{ konvex} \\ \forall i \in I : g_i(x) \leq 0 & g_i \text{ konvex} \\ \forall j \in J : h_j(x) = 0 & h_j \text{ affin linear} \end{array}$$

4.3. Dualität. *Motivation: Zwei Personen Nullsummenspiel (Gefangenendilemma)*

Spieler P	*Spieler D*
darf $x \in X$ *wählen*	*darf* $y \in Y$ *wählen*

Auszahlungsfunktion $P : X \times Y \to \mathbb{R}$ *mit der Interpretation, dass Spieler P* $P(x, y)$ *Spieler D auszahlt.*
Hier schließen sich zwei Optimierungsprobleme an:

$$\inf_{x \in X} \sup_{y \in Y} P(x, y)$$
$$\sup_{y \in Y} \inf_{x \in X} P(x, y)$$

Lemma. *Angenommen beide Werte in folgender Ungleichung existieren. Dann gilt:*

$$\max_{y \in Y} \min_{x \in X} P(x, y) \leq \min_{x \in X} \max_{y \in Y} P(x, y)$$

Beweis. Sei $\underline{x}, \underline{y} \in X \times Y$ mit $P(\underline{x}, \underline{y}) = \max\limits_{y \in Y} \min\limits_{x \in X} P(x, y)$
Sei $\bar{x}, \bar{y} \in Y$ mit $P(\bar{x}, \bar{y}) = \min\limits_{x \in X} \max\limits_{y \in Y} P(x, y)$
Dann gilt:

$$P(\underline{x}, \underline{y}) \leq P(\bar{x}, \underline{y}) \leq P(\bar{x}, \bar{y})$$

□

Definition (Sattelpunkt). *Ein Punkt* $(x^*, y^*) \in X \times Y$ *heißt* Sattelpunkt *wenn*

1. $\forall x \in X : P(x, y^*) \geq P(x^*, y^*)$
2. $\forall y \in Y : P(x^*, y) \leq P(x^*, y^*)$

Idee. *Interpretiere die Lagrangefunktion eines Optimierungsproblems als Auszahlungsfunktion eines Spiels.*

Beispiel(Linear-quadratische Optimierungsprobleme).

$\min_x \frac{1}{2} x^T Q x + d^T x \quad Q \in \mathbb{R}^{n \times n}$ *positiv definit*
$Ax = b \qquad\qquad\quad A \in \mathbb{R}^{m \times n}$ *A hat vollen Rang*

Setze nun $X = \mathbb{R}^n$, $Y \in \mathbb{R}^m$ und

$$P(x, \mu) = L(x, \mu) = \frac{1}{2} x^T Q x + d^T x + \mu^T (b - Ax)$$

Schauen wir uns das Optimierungsproblem von Spieler P an:

$$\min_x \max_\mu L(x, \mu)$$

Es ist notwendig, dass $x \in X Ax = b$ erfüllt, sonst ist der Wert des Spiels unbeschränkt.
Somit ist das Problem für Spieler P:

$$\min_x \frac{1}{2} x^T Q x + d^T x$$
$$Ax = b$$

Schauen wir uns das Optimierungsproblem von Spieler D an:

$$\max_\mu \min_x L(x, \mu) \quad \text{Setze } \Phi(\mu) = \min_{x \in X} L(x, \mu)$$

Für festes μ muss für ein Minimum x^* gelten, dass

$$Qx^* + d - \mu^T A = 0 \quad \text{Optimalitätsbedingung erster Ordnung}$$

Da Q positiv definit gilt also: $-A^T \mu$

$$x^* = Q^{-1}(A^T \mu - d)$$

Das heißt .....

$$\Phi(\mu) = \min_{x \in X} L(x, \mu) = L(x^*(\mu), \mu)$$

Somit gilt:

$$\Phi(\mu) = \frac{1}{2}(Q^{-1}(A^T\mu - d))^T Q(Q^{-1}(A^T\mu - d)) + ...$$

Dies führt auf ein Problem der Form

$$\Phi(\mu) = -\tfrac{1}{2}\mu^T\tilde{Q}\mu + \tilde{d}^T + \mu + \tilde{c} \quad \text{wobei } \tilde{Q} \text{ positiv definit}$$

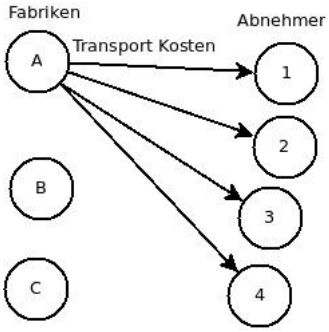
Mitschrift 21.11.2013

5 Lineare Programmierung

Ein lineares Programm in Standardform ist

$$\begin{array}{lll} \min c^T x & c \in \mathbb{R}^n i & \text{A vollen Zeilenrang} \\ Ax = b & A \in \mathbb{R}^{m \times n} & \\ x \geq 0 & b \in \mathbb{R}^m & \end{array}$$

Beispiel.



I Menge aller Fabriken
J Menge aller Abnehmer
c_{ij} *Kosten des Transports von Fabrik zu Abnehmer* j pro Einheit des Guts
p_i *Produktionsmenge von Fabrik* i
b_i *Bedarf von Abnehmer* f
x_{ij} *Menge, welche von Fabrik* i *an Abnehmer* j *geliefert wird*

$$\begin{array}{ll} \min \sum_{i \in J} \sum_{j \in J} c_{ij} x_{ij} & \textit{Minimiere die Gesamttransportkosten} \\ \forall i \in I \sum_i x_{ij} = p_i & \textit{Transportiere nur das, was produziert wurde (aber auch nicht weniger)} \\ \forall j \in J \sum_i x_{ij} = b_j & \textit{Alle Bedarfe werden erfüllt} \\ \forall i \in T \forall j \in J x_{ij} \geq 0 & \end{array}$$

Notation. *Seien Vektoren* $x, y \in \mathbb{R}^n$*, dann schreiben wir* $x \leq y$*, wenn* $\forall i \in \{1,, n\} : x_i \leq y_i$

Achtung! $\begin{pmatrix} -1 \\ 1 \end{pmatrix} \not\leq \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ und $\begin{pmatrix} 2 \\ -1 \end{pmatrix} \not\leq \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

5.1 Das duale lineare Programm

Sei

$$\begin{aligned} &\min c^T x \\ Ax &= b \text{ (P)} \\ x &\geq 0 \end{aligned}$$

das *primale* Problem.
Dann heißt

$$\begin{array}{ll} \max\limits_{y} b^T y \text{ (D')} \text{ bzw.} & \max\limits_{y,z} b^T y \\ A^T y \leq c & A^T y + z = c \text{ (D)} \\ & z \geq 0 \end{array}$$

das *duale* Problem zu (P).

Lemma (schwache Realität). *Sei $\bar{x} \in \mathbb{R}^n$ eine zulässige Lösung von (P) und $(\bar{y}, \bar{z})$ eine zulässige Lösung von (D),*
dann gilt:

$$c^T \bar{x} \geq b^T \bar{y}$$

Insbesondere gilt: Wenn sowohl (P) als auch (D) zulässige Lösungen besitzen, gilt dass der Optimalwert von (P) immer größer oder gleich dem Optimalwert von (D) ist.
Es gilt auch, dass wenn $b^T \bar{y} = c^T \bar{x}$ gilt, dass $\bar{x}$ und $(\bar{y}, \bar{z})$ Optimallösungen von (P) bzw. (D) sind.

Beweis.

$c^T \bar{x} = (A^T \bar{y} + \bar{z})^T \bar{x}$	$(\bar{y}, \bar{z})$ zulässig für (D)
$= \bar{y}^T A \bar{x} + \bar{z}^T \bar{x}$	Umstellen
$= \bar{y}^T b + \bar{z}^T \bar{x}$	$\bar{x}$ zulässig für (P)
$\geq \bar{y}^T b = b^T \bar{y}$	wg. $\bar{x} \geq 0$ und $\bar{z} \geq 0$ gilt $\bar{x}^T \bar{z} \geq 0$

□

5.2 Der primale Simplex-Algorithmus

Notation. *Sei $S \subseteq \{1, \dots, n\}$ eine geordnete Menge, dann nenne A_s die $m \times |S|$ Untermatrix von A, die aus den Spalten A_j mit $j \in S$ in der Reihenfolge der j in S.*

Mitschrift 22.11.2013

Definition (Simplex-) Basis. *Eine (Simplex-) Basis ist eine geordnete Menge B aus $\{1, \dots, n\}$, so dass A_B nicht singulär ist.*

Definition Basis Lösung. *Sei B eine Basis, dann x Basislösung zu Basis B, wenn $x = \begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} A_B^{-1} b \\ 0 \end{pmatrix}$. Wir sagen, die Basislösung ist zulässig, wenn $x_B \geq 0$*

Feststellung *Sei x eine Basislösung zur Basis B, dann gilt:*

$$Ax = \begin{pmatrix} A_B & A_N \end{pmatrix} \begin{pmatrix} x_B \\ x_N \end{pmatrix} = A_B \underbrace{x_B}_{=A_B^{-1}b} + \underbrace{A_\mu x_\mu}_{=0 \text{ } x \text{ Basislösung}} = A_B A_B^{-1} b = b$$

Primale revidierte Simplexalgorithmus. *Input: Eine zulässige Basislösung zu $\bar{x}$ zur Basis B*

```
1  While TRUE
2          Loese ȳ^T a_B = c_B^T fuer ȳ (Backward Transformation) (BTRAN)
3          z̄ ← c_N − A_N^T ȳ (Pricing)
4          If (z̄ ≥ 0) then return OPTIMAL (Pricing)
5          Waehle j mit z̄_j < 0 (Pricing)
6          Loese a_B w = A_j fuer w (FTRAN)
                         (A_j: j. Spalte von A)
7          If w ≤ 0 then return UNBESCHRAENKT (Ratio Test)
8          y ← min{ X̄_{B_x} / W_K W_K > 0} (Ratio Test)
9          Waehle ein i mit X̄_{B_i} / W_i = γ (Ratio Test)
10         x̄_B ← x̄_B − γW (Update)
11         N ← N ∪ {B_i}\{j}
12         B_i ← j
13         X̄_{B_i} − γ
```

Wenn der Algorithmus UNBESCHRÄNKT zurückgibt, ist das Problem (P) unbeschränkt

Beweis. Für $\lambda \geq 0$ setze

$$\bar{x}^\lambda = \begin{pmatrix} \bar{x}_B \\ \bar{x}_j \\ \bar{x}_{N \setminus \{j\}} \end{pmatrix} - \lambda \begin{pmatrix} w \\ -1 \\ 0 \end{pmatrix}$$

Wir zeigen zunächst $A\bar{x}^\lambda = b$
$A\bar{x}^\lambda = A_B\bar{x}_B^\lambda + A_{\cdot j}\bar{x}_j^\lambda + \underbrace{A_{N\setminus\{j\}}\bar{x}^\lambda_{\mu\setminus\{j\}}}_{=0} = A_B(\bar{x}_B - \lambda w) + \lambda A_{\cdot j}$
$= b - \lambda(A_B w - A_{\cdot j}) = b - \lambda(A_{\cdot j} - A_{\cdot j}) = b$
Wir zeigen $\bar{x}^\lambda \geq 0$:
Fallunterscheidung:

$k \in N\setminus\{j\}$	$\bar{x}_k^\lambda = 0$
$k = j$	$\bar{x}_j^\lambda = \lambda \geq 0$
$k \in B$	$\bar{x}_k^\lambda = \bar{x}_k - \lambda w_k \vert w_k \leq 0$
	$\geq \bar{x}_k \geq 0$

Betrachte:

$$\begin{aligned} c^T\bar{x}^\lambda &= c_B^T(\bar{x}_B - \lambda w) + \lambda c_j \\ &= \cdots\cdots\cdots \\ &= \bar{y}^T b + \lambda\bar{z}_j \end{aligned}$$

d.h für $\lambda > 0$ gilt: $c^T\bar{x} > c^T\bar{x}^\lambda$

□

Exkurs
Transformation auf Standardform
$a^T x \leq b \rightsquigarrow a^T x + z = b$ z Schlupfvariable b Slackvariable
$z \geq 0$
x frei, d.h. keine Vorzeichenbeschränkung $\rightsquigarrow x = x^+ - x^-$ $x^+, x^- \geq 0$

Mitschrift 28.11.2013

Offene Fragen:

- Ist das Problem optimal gelöst, wenn der Algorithmus OPTIMAL zurückgibt?
- Ist das Problem unbeschränkt, wenn der Algorithmus UNBESCHRÄNKT zurückgibt? (siehe oben)
- Ist nach dem Update-Schritt $\bar{x}$ eine Basislösung zur Basis B?
- Terminiert der Algorithmus?
- Wie erhalte ich eine Startlösung?
- Wie treffe ich die Wahlen im Algorithmus geschickt?

Hilfssatz: . *Wenn der Algorithmus OPTIMAL zurückgibt ist $\bar{x}$ eine Optimallösung von (P) und $(\bar{y}, \bar{z})$ eine Optimallösung von (D), wobei $\bar{z}_B = 0$.*

Beweis. $\bar{x}$ ist eine zulässige (Basis)lösung für (P) nach Vorraussetzung.
Es gilt:

$$A_y^T + \bar{z} = \begin{pmatrix} A_B^T \bar{y} \\ A_N^T \bar{y} \end{pmatrix} + \begin{pmatrix} \bar{z}_B \\ \bar{z}_N \end{pmatrix} = \begin{pmatrix} c_B \\ A_N^T y \end{pmatrix} + \begin{pmatrix} 0 \\ \bar{z}_N \end{pmatrix} = \begin{pmatrix} c_B \\ c_N \end{pmatrix} = c$$

Zu zeigen: $\bar{z} > 0$.
Klar für $j \in B \qquad z_j = 0$.
für $j \notin B \Rightarrow j \in N \qquad \bar{z}_j \geq 0$, da der Algorithmus OPTIMAL zurückgibt.
Somit ist $(\bar{y}, \bar{z})$ eine zulässige Lösung von (D).
Aus dem Beweis der schwachen Dualität wissen wir, dass $c^T \bar{x} = b^T \bar{y} + \bar{x}^T \bar{z}$ gilt.
Nun ist aber $\bar{x}^T \bar{z} = 0$, da für $j \in B \bar{z}_j = 0$ und für $j \in N \bar{x}_j = 0$. □

Hilfssatz: . *Nach dem Update-Schritt ist $\bar{x}$ Basislösung zur Basis B, d.h.*

- *B ist eine Basis*
- $\bar{x}_B = A_B^{-1} b$
- $\bar{x}_N = 0$

Beweis wird übersprungen.

Terminiert der Algorithmus?

Es gibt nur endlich viele Basen, also endlich viele Basislösungen.
Somit kann der Algorithmus nur dann nicht terminieren, wenn er "zykelt", d.h. Basen unendlich oft wieder besucht.
Dies kann vorkommen bei ungeschickter Wahl von j bzw. i.
Folgende Regeln verhindern das:

- Bland: Pricing: Wähle den kleinsten Index j
  RatioTest: Wähle den kleinsten Index i

- lexicographische Regel: Pricing: Wähle beliebiges j mit $\bar{z}_j < 0$
  Ratio Test: Wähle i so, dass $\frac{(A_B^{-1} A)_i}{w_i}$ lexicographisch minimal ist.

Wie finden wir eine Startbasis?

"Phase I des Simplex-Algorithmus"

$$\begin{aligned} \min c^T x \\ Ax = b \\ x \geq 0 \end{aligned}$$

Wir können annehmnen, dass $b \geq 0$
Betrachte folgendes Hilfsproblem:

$$\begin{aligned} \min \sum_{i=1}^{n} t_i \\ Ax + t = b \\ x_j t \geq 0 \end{aligned}$$

$$\left(A \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \right) \begin{pmatrix} x \\ t \end{pmatrix} = b$$

Feststellung: Für das Hilfsproblem ist es einfach eine zulässige Basislösung anzugeben, nämlich

$$x = 0, t = b$$

Wenn der Optimalwert des Hilfsproblems von 0 verschieden ist, ist unser Ausgangsproblem unzulässig.
Wenn der Optimalwert 0 ist, dann haben wir einen zulässigen Punkt von (P) gefunden.
Die so gefunden Basis lässt sich zu einer Startbasis von (P) oder

$$\begin{aligned} \min c^T x \\ Ax + t = b \\ t = 0 \\ x \geq 0 \end{aligned}$$

modifizieren
Mitschrift 29.11.2013

Geometrische Interpretation. *Wir betrachten lineare Programme in natürlicher (oder kanonischer) Form.*

$$\begin{aligned} \max c^T x \\ Ax \leq b \end{aligned}$$

Eine Menge $H = \{x : a^T x \leq b\}$ *mit* $a \in \mathbb{R}^n, b \in \mathbb{R}$ *heißt* Halbraum.

Definition Eine Menge $P \subseteq \mathbb{R}^n$ heißt *Polyeder*, wenn es eine Matrix $A \in \mathbb{R}^{n \times m}$ und ein $b \in \mathbb{R}^m$ gibt, so dass $P = \{x : Ax \leq b\}$ gilt.

Interpretation. *Der Simplexalgorithmus startet von einer zulässigen Ecke und geht von dort aus zu einer mindestens gleich guten neuen Ecke, bis eine Optimallösung gefunden ist oder eine unbeschränkte Richtung.*

Hinweis Es kann mehrere Basen $B_1, B_2, ...$ zur gleichen Basislösung existieren.

Definition (Ecke). *Sei $P \subseteq \mathbb{R}^n$ ein Polyeder und $v \in P$. Dann heißt v* Ecke *von P, wenn es keinen Vektor $h \in \mathbb{R}^n$ gibt, so dass $v + h \in P$* und *$v - h \in P$ gilt.*

Definition (geradenfrei). *Eine Menge $X \subseteq \mathbb{R}^n$ heißt* geradenfrei*, wenn sie keine Gerade enthält.*
Lemma *Jeder geradenfreie Polyeder hat eine Ecke.*
Feststellung *Mengen der Form $P = \{x : Ax = b, x \geq 0\}$ sind geradenfreie Polyeder.*
Lemma *Sei $P = \{x : Ax = b, x \geq 0\}$, wobei A vollen Zeilenrang hat. Dann ist $v \in P$ genau dann Ecke von P, wenn v eine zulässige Basislösung von*

$$\begin{aligned} \min c^T x \\ Ax = b \\ x \geq 0 \end{aligned}$$

Hirsch-Vermutung Schewe nett zum anschauen, aber 2010 widerlegt

Was hat das duale Problem mit der Lagrangefunktion zu tun?

$$\begin{array}{ll} \min\limits_{x} c^T x & \min\limits_{x,s} c^T x \\ Ax = b & \rightsquigarrow \quad Ax = b \\ x \geq 0 & \forall i \in \{1, \ldots, n\}, x_i - \frac{1}{2} s_i^2 = 0 \end{array}$$

$$L(x, s, \mu, \lambda) = c^T x + \mu^T (b - Ax) + \sum_{i=1} \lambda_i (\frac{1}{2} s_i^2 - x_i)$$

$$\begin{aligned} \frac{\partial L}{\partial \mu} &= b - Ax \\ \frac{\partial L}{\partial \lambda_i} &= \frac{1}{2} s_i^2 - x_i \\ \frac{\partial L}{\partial x} &= c^T - \mu^T A - \lambda^T \\ \frac{\partial L}{\partial s_i} &= \lambda_i s_i \end{aligned}$$

Schreibe nun $y = \mu$ und $z = \lambda$. Dann sind die Optimalitätsbedingungen erster Ordnung

$$\begin{aligned} Ax &= b \\ x_i - \frac{1}{2}s_i^2 &= 0 \\ A^T y + z &= c \\ \forall i \in \{1, \dots, n\} : x_i z_i &= 0 \end{aligned}$$

Wir erhalten das duale Problem mit Ausnahme der Bedingung $z \geq 0$. Diese folgt aus den Optimalitätsbedingungen zweiter Ordnung.

Mitschrift 5.12.2013

Teil II
Algebra

6 Elementare Zahlentheorie

Definition (Teilbarkeit). *Sei $d, n \in \mathbb{Z}$. Wir sagen d teilt n und schreiben $d \mid n$, wenn es ein $k \in \mathbb{Z}$ gibt, sodass $n = k * d$. (Negation $\nmid$)*

Definition (Division mit Rest). *Sei $x, n \in \mathbb{Z}$ mit $n \neq 0$, dann existieren eindeutige $q, r \in \mathbb{Z}$ mit $0 \leq r < n$, sodass $x = qn + r$.*

Beweis. Existenz:
Setze $q = \lfloor \frac{x}{n} \rfloor$ und $r = x - \lfloor \frac{x}{n} \rfloor n$. Dann gilt: $x - \lfloor \frac{x}{n} \rfloor n = n(\frac{x}{n} = \lfloor \frac{x}{n} \rfloor)$. Damit gilt für $r \quad 0 \leq r < n$.
Eindeutigkeit:
Sei $q, r \in \mathbb{Z}$ mit $x = qn + r \qquad 0 \leq r < n$
und $q', r' \in \mathbb{Z}$ mit $x = q'n + r' \qquad 0 \leq r' < n'$.
Darüber hinaus sei $r \geq r$.
Dann gilt:
$qn + r = x = q'n + r$. Somit $(q - q')n = (r' - r)$. Es gilt $0 = r' - r < n$.
Aus der Gleichung folgt $n \mid (r' - r)$. Daraus folgt $r' - r = 0$ somit auch $q' - q = 0$. □

Wir nennen die obige Zerlegung Division mit Rest und schreiben

$$x \mod n := r$$

Definition (Primzahl). *Eine Zahl $p \in \mathbb{N}$ mit $p \geq 2$ heisst Primzahl, wenn für alle $n, m \in \mathbb{N}$ mit $p = nm$ gilt, dass $n = 1$ oder $m = 1$.*

Lemma. *Seien $n, m \in \mathbb{Z}$ und sei $p \in \mathbb{N}$ prim.*
Wenn nun $p \mid nm$, dann muss $p \mid m$ oder $p \mid n$ gelten.

Beweis. Wende Division mit Rest auf n und m an:

$$\begin{aligned} n &= q_n p + r_n & 0 \leq r_n < p \\ m &= q_m p + r_m & 0 \leq r_m < p \end{aligned}$$

Dann gilt:

$$\begin{aligned} nm &= (q_n p + r_n)(q_m p + r_m) = \\ &= (q_n q_m p + q_n r_n + q_m r_n)p + r_n r_m \end{aligned}$$

Aus $p \mid nm$ folgt aber $p \mid r_n r_m$.
Es gilt aber $0 \leq r_n r_m < p^2$, da p prim ist, folgt daraus $r_n r_m = 0$. □

Satz (Satz über die Primfaktorzerlegung. *Sei $n \in \mathbb{N}$ mit $n \geq 2$. Dann gibt es Primzahlen $p_1, ...p_r \in \mathbb{N}$ und Exponenten $e_1, ..., e_r \in \mathbb{Z}_+$, sodass:*

$$n = \prod_{i=1}^{r} p_i^{e_i}$$
$$p_1 < p_2 < ... < p_r$$

Diese Darstellung ist eindeutig.

Beweis. (Primfaktorzerlegung)
Existenz: Induktion nach n.
$n = 2$ klar, da 2 Primzahl.
$n > 2$ wenn n prim, dann ist die Behauptung wahr. Ansonsten existiert $x, y \in \mathbb{Z}$ mit $n = xy$ und $x \neq 1$ und $y \neq 1$. Dann haben aber x und y nach Induktionsvorraussetzung Primfaktorzerlegung und somit auch n.
Eindeutigkeit: Widerspruchsbeweis.

Sei $n \in \mathbb{N}$ die kleinste Zahl, deren Primfaktorzerlegung nicht eindeutig ist, d.h.

$$n = \prod_{i=1}^{r} p_i^{e_i} \qquad \text{mit } p_i \text{ prim}$$
$$\prod_{j=1}^{r} q_j^{f_j} \qquad \text{mit } q_j \text{ prim}$$

Es kann kein i, j geben mit $p_i = q_j$, denn sonst ist $\frac{n}{p_i}$ eine kleinere Zahl ohne eindeutige Zerlegung.
Da aber $p_1 \mid n$ muss es nach dem Lemma ein i geben mit $p_1 \mid q_j$. Da aber p_1 und q_j prim sind, muss $p_1 = q_1$ gelten. Widerspruch. □

Satz. *Es gibt unendlich viele Primzahlen.*

Definition (ggT, gcd; kgV, lcm). [7]
Sei $x, y \in \mathbb{Z}$.

$$gcd(x, y) = \max\{d \quad d \mid x \wedge d \mid y\}$$
$$lcm(x, y) = \min\{m \quad m \geq 0 \wedge x \mid m \wedge y \mid m\}$$

Wir nennen $x, y \in \mathbb{Z}$ teilerfremd, wenn $gcd(x, y) = 1$.

Lemma.

a) $\forall z \in \mathbb{Z}$ $\quad gcd(z, 0) = |z|$

b) $\forall a, b \in \mathbb{Z}$ $\quad gcd(a, b) = gcd(b, a)$

c) $\forall a, b \in \mathbb{Z}$ $\quad a \leq b \quad \Rightarrow gcd(a, b) = gcd(b \mod a, a)$

Euklidischer Algorithmus

Input: $0 \leq a \leq b$
Output: $gcd(a, b)$

```
1  r, s ← a, b
2  While r ≠ 0
3          r, s ← s mod r, r
4  return s
```

[7] greatest common divisor, least common multiple

Es gilt immer $gcd(r,s) = gcd(a,b)$.
Nach jedem Durchlauf der Schleife gilt $r = 0$ oder r ist echt kleiner geworden.

Mitschrift vom 6.12.2013

Definition: (Größter gemeinsamer Teiler). $gcd(x,y) = \max\{d : d|x \wedge d|y\}$
Lemma

a) $\forall z \in \mathbb{Z}\ gcd(0,z) = |z|$

b) $\forall a,b \in \mathbb{Z}\ gcd(a,b) = gcd(b,a)$

c) $\forall a,b \in \mathbb{N}\ a \in b \Rightarrow gcd(a,b) = gcd(a, b\%a)$

Beweis. Beweis a),b) klar
c) Wir zeigen: Jeder gemeinsame Teiler von a und b teilt r *und* jeder gemeinsame Teiler von a und r teilt b. $r := b\%a$
Gelte also $x|a$ und $x|b$. Dann gilt $a = qx$ für $g \in \mathbb{Z}$
Nun ist also $b = na + r$ für $n \in \mathbb{Z}$. Damit gilt: $b = n(qx) + r$. Nach Annahme gilt $x|b$. Daraus folgt $x|r$. Die zweite Bek. folgt analog. □

Euklidischer Algorithmus in Matrixform Input: $0 \leq a \leq b$
Output: $gcd(a,b)$
$\begin{pmatrix} r \\ s \end{pmatrix} \leftarrow \begin{pmatrix} a \\ b \end{pmatrix}$

```
1 While r ≠ 0
2         q ← ⌊s/r⌋
```

$$\begin{pmatrix} r \\ s \end{pmatrix} \leftarrow \begin{pmatrix} -g & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r \\ s \end{pmatrix}$$

Satz. *Seien $x,y \in \mathbb{Z}$ Dann gilt:*

$$gcd(x,y) = \min\{m > 0 : \exists s,t \in \mathbb{Z}\ \ m = sx + ty\}$$

Beweis.

$$m = \min\{m > 0 : \exists s,t \in \mathbb{Z}\ m' = xs + ty\}$$

```
1 Zunaechst zeigen wir
2         gcd(x,y)|m Da es s,t∈ℤ gibt mit m=sx+ty und
              gcd(x,y)|x,gcd(x,y)|y gilt , folgt gcd(x,y)|m
3         Daraus folgt gcd(x,y) ≤ m
```

```
1 Wir zeigen nun m|x und m|y
2         Zu zeigen m|x (Beweis fuer m|y analog)
3         Sei nach Division mit Rest x = qm + r mit 0 ≤ r < m
4         Es gilt auch m = sx + ty, d.h.
5         x = qsx + qty + r ⇒ (1 − qs)x − qty = r, d.h. r ist
              Linearkombination von x und y
6         Nach Definition von m ist m die kleinste positive
              Linearkombination. Da aber r < m ist, muss r = 0 gelten
7 Somit gilt: m|x.
```

□

Erweiterter Euklidischer Algorithmus Input: $0 \leq a \leq b$
Output: d, u, v mit $gcd(a, b) = d = ua + vb$
$r, s \leftarrow a, b$
$U \leftarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

```
1 While r ≠ 0
2         q ← ⌊s/r⌋
3         Q ← (−q  1)
              ( 1  0)
4         (r)     (r)
          (s) ← Q (s)
5         U ← QU
6         return s, U21, U22
```

Zur Korrektheit:
Es gilt immer $\begin{pmatrix} r \\ s \end{pmatrix} = U \begin{pmatrix} a \\ b \end{pmatrix}$
Wie beim euklidischen Algorithmus gilt am Ende $gcd(a, b) = s$. Daraus folgt die Korrektheit.

Feststellung Seien $x, n \in \mathbb{Z}$. Dann sind x und n genau dann teilerfremd, wenn es ein $x_I \in \mathbb{Z}$ und $q \in \mathbb{Z}$, sodass $xx_I = 1 + qn$
Beweis von der Richtung $gcd(x, n) = 1 \Rightarrow \exists x_I, q \in \mathbb{Z}\ xx_I = 1 + qn$
Nach dem Lemma gilt: $1 = sx + tn$
Setze $x_I = s$ und $q = -t$.

Def. (die Relation $. \equiv .(mod\ n)$). *Seien $x, y, n \in \mathbb{Z}$, dann schreiben wir $\underbrace{x \equiv y}_{kongruent}\ \underbrace{mod\ n}_{modulo}$, wenn es ein $k \in \mathbb{Z}$ mit $kn = (x - y)$ (d.h. $n|(x - y)$)*

Satz (kleine Satz von Fermat). *Sei $a \in \mathbb{Z}$ und p eine Primzahl. Dann*

gilt:
$a^P \equiv a(mod\ p)$.

Definition (Euler'sche ϕ-Funktion) Die Funktion $\phi : \mathbb{Z} \to \mathbb{N}$ ist definiert als:

$$\phi(z) = |\{k :\ 1 \leq k \leq z\ \wedge gcd(k, z)\}|$$

Satz (Satz von Euler, Satz von Euler-Fermat) Sei $a, n \in \mathbb{Z}$ und a und n teilferfremd. Dann gilt:

$$a^{\phi(n)} \equiv 1(mod\ n)$$

Mitschrift vom 12.12.2013

7 Überblick über die gebräuchlichsten algebraischen Strukturen

Def Sei H eine Menge und $o : H \times H \to H$ eine binäre Operation
Dann heißt das Tupel (H, o) eine *Halbgruppe* (semigroup), wenn gilt:
$\forall a, b, c \in H : (a \circ b) \circ c = a \circ (b \circ c)$ (Assoziativität)

Def (Halbgruppen homomorphismus) Seien (H, o_H) und (L, o_L) Halbgruppen und $\varphi : H \to L$. Dann heißt φ Halbgruppenhomomorphismus, wenn:

$$\forall x, y \in H : \varphi(xo_H y) = \varphi(x) o_L \varphi(y)$$

Def (Monoid) Sei M eine Menge und $o : M \times M \to M$. Dann heißt (M, o) Monoid, wenn gilt:

(M1) $\forall a, b, c \in M : (a \circ b) \circ c = a \circ (b \circ c)$

(M2) $\exists e \in M, \forall x \in M : x \circ e = e \circ x = x$

Ein solches Element e heißt *Neutralelement*.
Definition von Monoid homomorphismen analog zu Halbgruppen homomorphismen.
Bsp

Sei $\underbrace{\Sigma}_{\text{Alphabet}}$ eine Menge und betrachten $\Sigma^* = \underbrace{\bigcup_{n=0}^{\infty} \Sigma^n}_{\text{Wörter über } \Sigma}$

Setze $o : \Sigma^* \times \Sigma^* \to \Sigma^*$
$v \circ w = (v_1, \ldots, v_k) \circ (w_1, \ldots, w_l) := (v_1, \ldots, v_k, w_1, \ldots, w_l)$ ("hintereinander schreiben")

Definition (Gruppe) Sei G eine Menge und $o : G \times G \to G$ Dann heißt (G, o) Gruppe, wenn gilt:

(G1) $\forall a, b, c \in G : (a \circ b) \circ c = a \circ (b \circ c)$

(G2) $\exists e \in G \; \forall g \in G : e \circ g = g \circ e = g$

(G3) $\forall g \in G \; \exists g^{-1} \in G : g \circ g^{-1} = g^{-1} \circ g = e$ (Existenz von Inversen)

Zur Notation: Typischerweise werden Gruppen *multiplikativ*, d.h. mit Operation $*$ notiert bzw. die Multiplikation durch hintereinanderschreiben, d.h. $x * y =: xy$, notiert.
Ausnahme: Wenn die Gruppe *abelsch* (d.h. kommutativ) ist, wird sie oft *additiv*, d.h. mit Operation $+$, Neutralelement e oder 0 und Inversen $-x$ notiert.

Def (Abelsche Gruppe, Kommutative Gruppe) Eine Gruppe (G, o) heißt *abelsch*, wenn $\forall g, h \in G : g \circ h = h \circ g$ (Kommutativität)
Bsp.
Setze $GL(n, \mathbb{R}) = \{A : \lambda \in \mathbb{R}^{n \times n}$ und A ist invertierbar$\}$.
Dann ist $GL(n, \mathbb{R})$ eine Gruppe, aber *keine* abelsche Gruppe.

Def Gruppenhomomorphismus erfolgt analog.

Lemma. *Sei (M, o) ein Monoid. Dann ist das Neutralelement eindeutig.*

Beweis. Seien e, e' Neutralelemente:

$$e \underbrace{=}_{e' \text{ Neutralelement}}^{(M2)} e \circ e' \underbrace{=}_{e \text{ ist Neutralelement}}^{(M2)} e'$$

□

Notation Wenn G endlich ist, schreiben wir die Abbildung $o : G \times G \to G$ oft als Tabelle (Verknüpfungstabelle oder Gruppentafel)

o	g	h
g		$g \circ h$
h	$h \circ g$	

Def (Ring) Sei R eine Menge $+ : R \times R \to R$, $* : R \times R \to R$
Dann heißt R Ring, wenn es ein Element $0 \in R$ und ein Element 1 in R gibt, sodass

(R1) $(R, +, 0)$ eine abelsche Gruppe (mit Neutralelement 0) ist.

(R2) $(R, *, 1)$ ein Monoid ist und

(R3) (Distributivität)

a) $\forall x, y, z \in \mathbb{R} : \ (x + y)z = xz + yz$

b) $\forall x, y, z \in \mathbb{R} : \ z(x + y) = zx + zy$

Def Ein Ring $(R, +, *)$ heißt *kommutativ*, wenn auch $\forall x, y \in \mathbb{R} x * y = y * x$ gilt.
Bsp. $M(n, \mathbb{R}) = \{A : A \in \mathbb{R}^{n \times n}\}$
$\mathbb{R}[x] = \{$Menge aller Polynome in einer Variable mit Koeffizienten in $\mathbb{R}\}$

Def Körper (engl. Field) Ein kommutativer Ring $(k, +, *)$ heißt *Körper*, wenn gilt:

$$\forall x \in K \backslash \{0\}. \ \exists x^{-1} \ xx^{-1} = 1$$

Mitschrift vom 13.12.2013

8 $\mathbb{Z}_n$

Def. (Restklasse) Sei $n \in \mathbb{Z} \backslash \{0\}$. Dann heißt für $x \in \mathbb{Z}$

$[x]_n := \{z : x = z(mod n)\}$ die Restklasse von x bezüglich n.

Bemerkung: Wir lassen den Index n weg, wenn er aus dem Kontext klar ist.
Andere übliche Notation $x + n\mathbb{Z}$
Bsp.

$$[0]_6 = \{\dots, -12, -6, 0, 6, 12, \dots\}$$
$$[1]_6 = \{\dots, -11, -5, 1, 7, 13, \dots\}$$

Hilfssatz Sei $n \in \mathbb{Z}\backslash\{0\}$ und $x, y \in \mathbb{Z}$. Dann gilt: Entweder $[x]_n = [y]_n$ oder $[x]_n \cap [y]_n = \emptyset$.

Beweis. Betrachte $r_x = x \mod n$ und $r_y = y \mod n$. Es gilt: $r_x \in [x]_n$ und $r_y \in [y]_n$.
Behauptung: Es gilt $[x]_n = [r_x]_n$ und $[y_n] = [r_y]_n$
Das folt aus $x = r_x + q_x n$ (nach Def von mod) Analog für y.
Nun gilt aber $[r_x]_n = [r_y]_n$ genau dann, wenn $r_x = r_y$. Wenn aber $r_x \neq r_y$ haben $[r_x]_n$ und $[r_y]_n$ keine gemeinsamen Elemente in $\{0, \ldots, n-1\}$. Somit haben sie aber *keine* gemeinsamen Elemente in ganz $\mathbb{Z}$ □

Feststellung Für $n \in \mathbb{Z}\backslash\{0\}$ gibt es genau n verschiedene Restklassen, nämlich $[0]_n, [1]_n, \ldots, [n-1]_n$

Lemma Sei $n \in \mathbb{Z}\backslash\{0\}$ und seien $a, a' \in \mathbb{Z}$ und $b, b' \in \mathbb{Z}$ mit $[a]_n = [a']_n$ und $[b]_n = [b']_n$.
Dann gilt:

a) $[a + b]_n = [a' + b']_n$

b) $[a * b]_n = [a' * b']_n$

Beweis. Nach Voraussetzung gilt:

$$a = r_a + q_a n,\ a' = r_a + q_{a'} n \text{ mit } r_a \in \{0, \ldots, n-1\}$$
$$b = r_b + q_b n,\ b' = r_b + q_{b'} n \text{ mit } r_b \in \{0, \ldots, n-1\}$$

Somit gilt:

$$a + b = r_a + r_b + (q_a + q_b)n$$
$$a' + b' = r_a + r_b + (q_{a'} + q_{b'})n$$

Nach Definition von $[\ ]_n$ gilt dann aber: $[a + b]_n = [r_a + r_b]_n = [a' + b']_n$ □

Definition Für $n \in \mathbb{Z}\backslash\{0\}$ und $a, b \in \mathbb{Z}$ setze

$$[a]_n + [b]_n := [a + b]_n$$
$$[a]_n * [b]_n := [a * b]_n$$

Definition Für $n \in \mathbb{Z}\backslash\{0\}$ heißt die Menge aller Restklassen bezüglich n $\mathbb{Z}_n$ (Alternativnotation: $\mathbb{Z}/n\mathbb{Z}$)

Feststellung $(\mathbb{Z}_n, +, [0]_n)$ ist eine abelsche Gruppe für alle $n \in \mathbb{Z}\backslash\{0\}$
$(\mathbb{Z}_n, *, [1]_n)$ ist ein kommutativer Monoid für alle $n \in \mathbb{Z}\backslash\{0\}$

Bsp

$\mathbb{Z}_4$

$*$	[0]	[1]	[2]	[3]
[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]
[2]	[0]	[2]	[0]	[2]
[3]	[0]	[3]	[2]	[1]

$\mathbb{Z}_n$ ist *kein* Körper
([2] nicht invertierbar)

$\mathbb{Z}_5$

$*$	[0]	[1]	[2]	[3]	[4]
[0]	[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]
[2]	[0]	[2]	[4]	[1]	[3]
[3]	[0]	[3]	[1]	[4]	[2]
[4]	[0]	[4]	[3]	[2]	[1]

$(\mathbb{Z}_5, +, *)$ ist ein *Körper*

Lemma Sei $n \in \mathbb{Z}\backslash\{0\}$ Dann hat $[x]_n$ für $x \in \mathbb{Z}$ genau dann ein multiplikatives Inverses, wenn x und n teilerfremd sind.

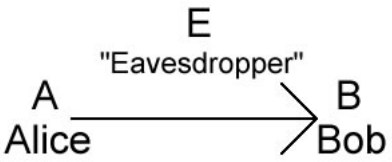
Beweis. Da x und n teilerfremd, liefert uns der erweiterte euklidische Algorithmus $a, q \in \mathbb{Z}$ mit $ax + qn = 1$
Somit gilt: $[a]_n[x]_n = [ax]_n = [1 - qn]_n = [1]_n$
Dh. $[a]_n$ ist multiplikatives Inverses von $[x]_n$.
Wenn x und n nicht teilerfremd sind, kann es keine solche linearkombination geben, und mit einem analogen Argument auch kein multiplikatives Inverses von $[x]_n$ □

Folgerung $(\mathbb{Z}, +, *)$ ist genau dann ein Körper, wenn n Primzahl ist.

Hilfssatz Die Menge $(\mathbb{Z}_n)^* = \{[x]_n : x \in \{1, \ldots, n-1\}$ und x und n teilerfremd$\}$ mit den Verknüpfungen $+$ und $*$ wie oben bildet eine Gruppe.

Mitschrift vom 19.12.2013

9 Anwendungen der Algebra in Kryptographie und Kanalkodierung



Symmetrische Verschlüsselungsverfahren
Public-Key-Verfahren (asymmetrische Verschlüsselungsverfahren): Diffie-Hellman 1976

Diffie-Hellman-Schlüsselaustausch

Alice wählt eine Primzahl p und wählt ein $g \in \mathbb{Z}_p^*$. Alice schickt (p, g) an Bob.
Alice wählt zufällig ein $a \in \mathbb{Z}$ und schickt $g^a \mod p$ an Bob.
Bob wählt zufällig $b \in \mathbb{Z}$ und schickt $g^b \mod p$ an Alice.
Nun berechnen beide $g^{ab} \mod p$ (Alice als $(g^b)^a \mod p$ und Bob als $(g^a)^b \mod p$).
Was weiss Eve? $p, g, g^a \mod p, g^b \mod p$
Diskretes Logarithmusproblem: Gegeben p, g und $x = g^a \mod p$, bestimme a.
Verallgemeinerung: Wähle zunächst eine zyklische Gruppe G und einem Erzeuger $g \in G$.

$$A \xrightarrow{G,g} B$$

$$a\ A \xrightarrow{g^a} B$$

$$a\ A \xleftarrow{g^b} B\ b$$

Definition: (zyklische, Erzeuger). *Eine Gruppe G heisst zyklisch, wenn es ein Element $g \in G$ gibt (g heisst Erzeuger), sodass für alle $x \in G$ ein $a \in \mathbb{Z}$ existiert mit $x = g^a$.*

Typische Quellen für Gruppen sind z.B. elliptische Kurven.

9.1 RSA-Verfahren (Rivext, Shamir, Adleman)

Schlüsselerzeugung

- Wähle zufällig zwei Primzahlen p, q.
- Berechne $\phi(n) = (p-1)(q-1)$.
- Wähle e zufällig mit $gcd(\phi(n), e) = 1$.
- Berechne $d \equiv e^{-1} \mod \phi(n)$.
- $de \equiv 1 \mod \phi(n)$

Öffentlicher Schlüssel ist (n, e).
Geheimer Schlüssel ist (d, p, g).

Verschlüsselung

Sei die Nachricht $m \in \mathbb{Z}$. Setze $c = m^e \mod n$. Sende c.

Entschlüsselung

Setze $m' = c^d \mod n$

Eve kennt (n, e), $n^e \mod n$, aber <u>nicht</u> $\phi(n)$.

9.2 Satz von Euler

Sei $n \in \mathbb{Z}$ und sei $a \in \mathbb{Z}_n^*$. Dann gilt:

$$a^{\Phi(n)} = 1 \quad \text{in } \mathbb{Z}_n^*$$

Beweis: Wir betrachten die Abbildung $m_a : \mathbb{Z}_n^* \to \mathbb{Z}_n^*, m_a(x) = a \cdot x$. Da $\mathbb{Z}_n^*$ eine Gruppe ist, ist m_a eine Bijektion (Umkehrabbildung m_{a-1}).
Nun gilt

$$\prod_{x \in \mathbb{Z}_n^*} x = \prod_{x \in \mathbb{Z}_n^*} m_a(x) = \prod_{x \in \mathbb{Z}_n^*} ax = a^{|\mathbb{Z}_n^*|} \prod_{x \in \mathbb{Z}_n^*} x = a^{\Phi(n)} \prod_{x \in \mathbb{Z}_n^*} x \Rightarrow a^{\Phi(n)} = 1 \quad \text{in } \mathbb{Z}_n^*$$

Hilfssatz: Für $m \in \mathbb{Z}_n^*$ und e mit $gcd(e, \Phi(n)) = 1$ und $d = e^{-1} \mod \Phi(n)$ gilt:

$$(m^e)^d = m \mod n$$

Beweis:

$$(m^e)^d = m^{ed} = m^{1+K\Phi(n)} = m \cdot m^{K\Phi(n)} = m \cdot (m^{\Phi(n)})^K = m \cdot 1^K = m$$

Berechnung von $m^e \mod n$

Gegeben: $a \in \mathbb{Z}_n^*$ und $b \in \mathbb{N}$. Sei $(b_k, \ldots, b_0)$ die Binärdarstellung von b, d.h. $b = \sum_{i=0}^{k} 2^i b_i$.

Berechne a^b in$\mathbb{Z}_n^*$

```
1 r = 1
2 for i = k ... 0:
3         r = r^2
4         if b_i:
5                 r = a * r
6 return r
```

9.3 Finden einer Primzahl / Primzahltests

Fermattest: Gegeben $n \in \mathbb{Z}$, wähle zufällig a teilerfremd zu n. Berechne $a^{n-1} \mod n =: t$. Wenn $t \neq 1$, antworte n ist nicht prim.

Problem: Es existieren die sog. Carmickaelzahlen, welche nicht prim sind, aber trotzdem $a^{n-1} \mod n = 1$ für alle a gilt.

Hilfssatz: Gegeben $n \in \mathbb{Z}$ mit n ungerade und $n > 1$. Setze $T_n = \{\alpha \in \{1, \ldots, n-1\} : a^{t2^h} = 1 \text{ und } \forall j \in \{0, \ldots, h\} : \alpha^{t2^j} = \pm 1$.

Dann gilt: Wenn n prim, $T_n = \{1, \ldots, n-1\}$, sonst $|T_n| \leq \frac{n-1}{4}$. Beweis siehe Shoup.

Aus dem Hilfssatz lässt sich ein Primzahltest konstruieren, der sog. Miller-Rabin-Test.

Bemerkung: Es existiert ein Polynomialzeitalgorithmus, um zu entscheiden, ob eine Zahl prim ist.[8]

10 Fehlt

Mitschrift vom 16.01.2014

[8] Agrawal et al.: PRIMES is in P

Teil III
Gewöhnliche Differentialgleichungen

11 Einführung

Definition. *Eine (explizite] gewöhnliche Differentialgleichung erster Ordnung ist eine Gleichung der Form*

$$y'(t) = f(t, y(t))$$

mit $y : I \subseteq \mathbb{R}$ *und* $f : \mathbb{R} \times I \mapsto \mathbb{R}$ *(*oder so ähnlich*)*

Bsp.:
$y(t) = y'(t)$
Gesucht: y
Lösungen sind: $y(t) = ce^t$

Definition: (Anfangswertprobleme). *Ein Anfangswertproblem für gewöhlniche Differentialgleichungen erster Ordnung besteht aus einer gew. Differentialgleichung erster Ordnung und einem Paar* (t_0, y_0).
Eine Lösungs des Anfangswertproblems ist ein Funktion $y : I \mapsto \mathbb{R}$*, sodass die Differentialgleichung erfüllst ist und* $y(t_0) = y_0$ *(mit* $t_0 \in I$*).*

Definition. *Eine implizite gewöhnliche Differentialgleichung erster Ordnung ist von der Form* $0 = f(y'(t), y(t), t)$ *mit* $f : \mathbb{R} \times \mathbb{R} \times I \mapsto \mathbb{R}$

Bezeichnungen:
gewöhnliche Differentialgleichung: ordinary differential equation (ODE)
Anfangswertproblem: initial value problem (NP)

Definition: (System von gew. Differentialgleichungen erster Ordnung. *Ein System von gew. Differentialgleichungen erster Ordnung ist von der Form*

$$y_1'(t) = f_1(y_1(t), y_2(t), ..., y_n(t), t)$$
$$...$$
$$y_n'(t) = f_n(y_1(t), y_2(t), ..., y_n(t), t)$$

Wir schreiben auch
$y'(t) = f(y(t), t)$ *mit* $y : I \subseteq \mathbb{R} \mapsto \mathbb{R}^n$ *und* $f : \mathbb{R}^n \times I \mapsto \mathbb{R}^n$

Definition: (Anfangswertproblem für Systeme erster Ordnung. *Ein Anfangswertproblem von Systemen von gew. Differentialgleichungen erster Ordnung besteht aus einem System gew. Differentialgleichungen erster Ordnung*
$y't(t) = f(y(t), t)$ *mit* $f : \mathbb{R}^n\$ \times I \mapsto \mathbb{R}$
und einem Anfangswert (t_0, y_0), $y_0 \in \mathbb{R}^n$.
Eine Lösung des Anfangswertproblems ist eine Funktion $y : I \mapsto \mathbb{R}^n$, *die die Gleichung erfüllt und für die* $y(t_0) = y_0$ *gilt.*

Definition. *Eine gewöhnliche Differentialgleichung n-ter Ordnung ist von der Form*

$$y^{(n)}(t) = f(y^{(n-1)}(t), y^{(n-2)}(t), ..., y'(t), y(t), t)$$

mit $y : I \mapsto \mathbb{R}$ *und* $f : \mathbb{R}^n \times I \mapsto \mathbb{R}$

Umwandlung von expliziten gew. Differentialgleichungen n-ter Ordnung in ein System erster Ordnung:
Gegeben:

$$y^{(n)}(t) = f(y^{(n-1)}(t), ..., y'(t), y(t), t)$$

Setze:
$\bar{y_1}(t) = y(t)$
$\bar{y_2}(t) = y'(t)$
$\bar{y_3}(t) = y''(t)$
...
$\bar{y_n}(t) = y^{(n-1)})(t)$
Es gilt:
$\bar{y}_1'(t) = \bar{y}_2(t)$
$\bar{y}_2'(t) = \bar{y}_3(t)$
...
$\bar{y}_{n-1}'(t) = \bar{y}_n(t)$
$\bar{y}_n'(t) = y^{(n)}(t) = f_{n+1}(\bar{y}_n(t), \bar{y}_{n-1}(t), \bar{y}_{n-2}(t), ..., \bar{y}_1(t), t)$
Beispiel:
Bezeichne mit $x(t)$ die Positionen eines Endes der Feder und sei x_p ein fester Punkt.
Es gilt (Newton): $F(t) = m * x''(t)$

Nach Hook'schem Gesetz gilt: $F(t) = -K(x - x_0)$ (Rückstellkraft)

Mitschrift 17.01.2014

12 Elementare Lösungsansätze für gew. Differentialgleichungen

Gegeben: $y'(t) = f(y(t), t)$
Richtungsfeld: (Graph)
$y'(t) = y(t)(1 - y(t))$

12.1 Direkte Berechnung der Stammfunktion

Betrachte: $y'(t) = f(t) \qquad y(t_0) = y_0 \qquad f\ stetig$
Lösung: $y(t) = y_0 + \int_{t_0}^{\tau} f(\tau) d\tau$

12.2 Trennung der Variablen

Gegeben: $y'(t) = g(y(t)) * h(t) \qquad g, h\ stetig \qquad y(t_0) = y = 0$
Herleitung des Lösungsverfahrens:
Wenn $g(y_0) = 0$ ist ($y(t) = y_0$ für alle t Lösung des AWP
Sei also $\gamma : I \mapsto \mathbb{R}$, sodass $g(y(t)) \neq 0$ für $t \in I$.
Schreibe nun die Gleichung um

$$\frac{\gamma'(t)}{g(\gamma(t))} = h(t)$$

Setze $\tilde{g}(x) = \frac{1}{g(x)}, \tilde{G}$ sei eine Stammfunktion von $\tilde{g}$ und H eine Stammfunktion von h.

Nun gilt für $s, t \in I$:

$$H(t) - H(s) = \int_s^t h(\tau)d\tau =$$

$$= \int_s^t \frac{\gamma'(\tau)}{g(\gamma(\tau))}d\tau = \int_s^t \tilde{g}(\gamma(\tau)) * y'(tau)d\tau =$$

$$= \int_{\gamma(s)}^{\gamma(t)} \tilde{g}(x)dx = \tilde{G}(y(t)) - \tilde{G}(y(s))$$

Da nun $g(x) \neq 0$ ist auch $\tilde{g}(x) \neq 0$, somit $\tilde{G}$ monoton, genauer streng monoton.
Damit hat $\tilde{G}$ eine Umkehrfunktion.
Somit ist (mit $c = \tilde{G}(y(s)) - H(s)$)
$\forall t \in I : \qquad \gamma(t) = \tilde{G}^{-1}(H(t) + c)$
Beispiel:
$y't(t) = t * y^2(t) \qquad y(t_0) = y_0$
Für $y_0 = 0$ ist $y(t) = 0$ Lösung des AWP.
Sei g eine Lösung, dann gilt (mit $y_0 \neq 0$)
$\frac{y'(t)}{y^2(t)} = t$ und somit

$${t_0}^{\tau} \frac{y'(\tau)}{y^2(\tau)}d\tau = \int_{t_0}^{\tau} \tau d\tau$$

Nach Transfomrationsformel gilt:

$$\int_{y_0}^{y(t)} = \frac{1}{x^2}dx = \int_{t_0}^{t} \tau d\tau$$

Dies ist

$$-\frac{1}{y(t)} + \frac{1}{y_0} = \frac{1}{2}(t^2 - t_0^2)$$

Nun löst man diese Gleichung nach $y(t)$ auf.

Schema für die Trennung der Variablen

1. Überprüfen der Nullstellen von g
2. Schreibe das Problem als $\frac{y'(t)}{g(y(t))} = h(t)$

3. Integriere diese Gleichung auf beiden Seiten (Transformationsformel)
4. Löse nach $y(t)$ auf
5. PROBE!

Beispiel: Homogene lineare Differentialgleichung
$y'(t) = a(t)y(t)$
Lösung:
$y(t) = y_0 \qquad \exp\left(\int_{t_0}^{t} a(t)dt\right)$
Mitschrift 23.01.14

12.3 Lineare skalare Differentialgleichungen erster Ordnung

$$y'(t) = a(t)y(t) + b(t) \qquad y(t_0) = y_0$$

Homogenes Problem:

$$y'(t) = a(t)y(t) \qquad y(t_0) = y_0$$

Lösung für das homogene Problem:

$$y(t) = y_0 \exp\left(\int_{t_0}^{t} a(\tau)d\tau\right)$$

Lösung der inhomegenen Gleichung: Angenommen $y_H(t)$ ist eine Lösung der homogenen Gleichung. Idee: Variation der Konstanten

$$y_P(t) = c(t)y_H(t)$$

Dann gilt

$$\begin{aligned} y_P'(t) = c(t)y_H'(t) + c'(t)y_H(t) = c)t = (a(t)y_H(t)) + c'(t)y_H(t) = \\ = a(t)(c(t)y_H(t)) + c'(t)y_H(t) \\ = a(t)y_P(t) + c'(t)y_H(t) \end{aligned}$$

Wenn $b(t) = c'(t)y_H(t)$ gilt, dann ist $y_P(t)$ Lösung des inhomogenen Problems.
Wenn $y_?(t) \neq 0$ gilt, muss $c(t)$

$$c'(t) = \frac{b(t)}{y_H(t)}$$

erfüllen.
D.h.
$c(t)$ ist Stammfunktion von $\frac{b(t)}{y_H(t)}$
Zusammen ergibt sich für das inhomogene Anfangswertproblem

$$y(t) = \int_{t_0}^{t} b(\tau) \exp\left(\int_{\tau}^{t} a(s)ds\right) d\tau + y_0 \exp(\int_{t_0}^{t} a(\tau)d\tau$$

Lemma. *Sei $y_1 : \mathbb{R} \mapsto \mathbb{R}$ eine Lösung von $y'(t) = a(t)y(t)$.*

a) Für alle $\lambda \in \mathbb{R}$ ist auch $\lambda y_1(t)$ eine Lösung der Differentialgleichung

b) Wenn $y_2 : \mathbb{R} \mapsto \mathbb{R}$ ebenfalls Lösung der Differentialgleichung ist, dann ist auch $y_1(t) + y_2(t)$ eine Lösung der Differentialgleichung

Lemma. *Seien $y_1 : \mathbb{R} \mapsto \mathbb{R}$ und $y_2 : \mathbb{R} \mapsto \mathbb{R}$ Lösungen von*
$y'(t) = a(t)y(t) + b(t)$.
Dann ist $y_1(t)...y_2(t)$ Lösung von $y'(t) = a(t)y(t)$.

Beweis. $(y_1(t) - y_2(t))' = y_1'(t) - y_2'(t) = a(t)y_1(t)b(t) - (a(t)y_2(t) =$
$= a(t)(y_1(t) - y_2(t))$

□

Satz. *Sei $y_P : \mathbb{R} \mapsto \mathbb{R}$. eine Lösung von $y'(t) = a(t)y(t) + b(t)$.*
Dann ist jede Lösung der Differentialgleichung von der Form
$y(t) = y_P(t) + y_H(t)$, wobei
$y_H(t)$ eine Lösung von $y'(t) = a(t)y(t)$ ist.

Schema:

1. Löse die homogene Differentialgleichung durch Trennung der Variablen
2. Finde eine spezielle Lösung durch Variation der Konstanten
3. Die gesuchte Lösung hat nun die Form $y_P(t) + cy_H(t)$. Zur Bestimmung von c nutze die Anfangsbedingung.
4. Probe!

Mitschrift vom 07.02.14
$y'(t) = Ay(t) y(0) = y_0 A \in \mathbb{R}^{n*n}, y_0 \in \mathbb{R}^n$
Lösung: $y(t) = e^{-LA} y_0$
Es gilt: $exp \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} e^A & 0 \\ 0 & e^B \end{pmatrix}$
$exp(S^{-1}AS) = S^{-1} exp(A) S$

Definition Hauptvektor

Sei $A \in \mathbb{C}^{n*n}$. Dann heißt $v \in \mathbb{C}^n$ Hauptvektor zu $\lambda \in \mathbb{C}$ der Stufe K, wenn $(A - \lambda I)^K v = 0$ und $(A - \lambda I)^{v-1} v \neq 0$
Beispiel Eigenvektoren sind Hauptvektoren der Stufe 1.
$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$
Gleichungssystem für EV zum EW 1. $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} v = v \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} v = 0$
Lösungsmenge $\{\forall v \lambda e_1\}$ Um einen HV der Stufe 2 zu finden, löse $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
Lösung ist $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Somit ist $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ eine Basis aus Hauptvektoren

Feststellung

Sei v HV zu λ der Stufe k.
Sei $w = (A - \lambda I)v$. Dann ist w HV der Stufe $k - 1$ zu λ

Satz

Sei $A \in \mathbb{C}^{nxn}$ und seien $\lambda_1, \cdots, \lambda_k$ die Eigenwerte von A. Dann existiert eine Basis aus Hauptvektoren, sodass A bzgl. dieser Basis wie folgt aussieht:

$$\begin{bmatrix} \boxed{\begin{matrix} \lambda & \lambda \\ \lambda & \lambda \end{matrix}} & \cdots & 0 \\ \vdots & \boxed{\begin{matrix} \lambda & \lambda \\ \lambda & \lambda \end{matrix}} & \vdots \\ 0 & 0 & 0 \end{bmatrix}$$

D.h. um $exp(A)$ zu berechnen, reicht Matrizen der Form $A = \begin{pmatrix} \lambda & & * \\ & \ddots & \\ 0 & & \lambda \end{pmatrix} =$

$$\begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix} + \underbrace{\begin{pmatrix} 0 & \dots & . \\ & \ddots & * \\ 0 & & 0 \end{pmatrix}}_{N}$$

$$e^{tA} = exp(t\lambda I + tN) = exp(t\lambda I)exp(tN) = \begin{pmatrix} e^{\lambda t} & & \\ & \ddots & \\ & & e^{\lambda t} \end{pmatrix} \underbrace{P_N(tN)}_{\text{Polynom von Grad n}}$$

Teil IV
Zusammenfassung Semester

Lösen von Gleichungen
Satz von den impliziten Funktionen

I Optimierung
Lokale Optima ohne Nebenbedingungen $\min_{x \in \mathbb{R}} f(x)$
Lokale Optima unter Nebenbedingungen $\min f(x)\ g(x) = 0$
Konvexe Optimierungsprobleme
Dualität
Lineare Programmierung (Spezialfall von konvex Opt) (GLOBALE LÖSUNG)
Simplexalgorithmus
II Algebra, Zahlentheorie
Elementare Zahlentheorie
Überblick über algebraische Strukturen
Endliche Körper
Abelsche Gruppen
Beispiel: Ein wenig Kryptographie
III Differentialgleichungen
Elementare Differentialgleichungen
Existenz und Eindeutigkeit
Systeme linearer Differentialgleichungen

Klausur 90 min
keine 6x6 matrizen-aufgabe

mehrere kleine Aufgaben
Schriftliche Unterlagen erlaubt
Empfehlung ”Keine Halbe Bibliothek mitnehmen” ”Kurze Klausur. Machen sie sich ein 1DIN A4 Blatt oder 2”