Optimierung

Prof. Dr. H. Maurer – WWU Münster, Wintersemester 2009/2010
Vorlesungsmitschrift von Christian Schulte zu Berge

22. Januar 2010

Inhaltsverzeichnis

Vorwort

Dieses Skript entstand als Mitschrift in der der Vorlesung „Optimierung“, gelesen im Wintersemester 2009/2010 von Prof. Dr. H. Maurer an der Universität Münster.

Es besteht keine Garantie auf Richtigkeit oder Vollständigkeit des Skriptes. Dies ist lediglich eine persönliche Mitschrift und nicht zur Veröffentlichung bestimmt. Jegliche Weitergabe an Dritte bitte nur im Einverständnis mit dem Autor. Diese Mitschrift ist noch in der Erstellungsphase und so gibt es noch laufend Ergänzungen und Korrekturen. Dies stets aktuelle Version ist auf meiner Homepage `www.cszb.net` zu finden.

Falls Fehler gefunden werden oder Fragen auftauchen, am Besten eine kurze Mail an `skript@cszb.net` schreiben.

Christian Schulte zu Berge

Kapitel 0

Einführung

Definition: (Allgemeines Optimierungsproblem)
Sei $(X, ||\cdot||)$ ein normierter Vektorraum, $D \subset X$ offen, $S \subset D \subset X$ und $f : D \to \mathbb{R}$ eine Abbildung. Dann ist ein *allgemeines Optimierungsproblem*:

$$(P) \qquad \text{Minimiere } f(x) \text{ unter der Nebenbedingung } x \in S$$

Die Funktion $f : D \to \mathbb{R}$ heißt Zielfunktion, die Menge S heißt zulässige Menge, jedes $x \in S$ heißt zulässiger Punkt.

Ein zulässiger Punkt $\bar{x} \in S$ heißt *lokale Minimalstelle* von (P), falls eine Umgebung $U_{\bar{x}} \subset X, \bar{x} \in U_{\bar{x}}$ existiert mit $\forall x \in U_{\bar{x}} \cap S : f(\bar{x}) \leq f(x)$.

$\bar{x}$ heißt *globale Minimalstelle* bzw. *optimale Lösung* von (P), falls eine Umgebung, falls $\forall x \in S : f(\bar{x}) \leq f(x)$.

$\bar{x} \in S$ heißt *strikte Minimalstelle* (lokal/global), falls $f(\bar{x}) < f(x)$ für alle entsprechenden Punkte gilt.

Notation:
Als Kurzschreibweise für ein allgemeines Optimierungsproblem verwendet man oft:

$$(P) \qquad \min\{f(x) \mid x \in S\}$$

Definition und Satz:
Der optimale Wert von (P) ist gegeben durch:

$$\inf(P) := \begin{cases} \inf\{f(x) \mid x \in S\} & \text{falls } S \neq \emptyset \\ \infty & \text{falls } S = \emptyset \end{cases}$$

Wenn (P) eine optimale Lösung besitzt, dann schreibt man auch $\min(P)$ statt $\inf(P)$.

Die Maximierung einer Funktion $g : D \to \mathbb{R}$ mit S ist äquivalent zur Minimierung von $f := -g$ auf S.

0.1 Typen von Optimierungsproblemen

0.1.1 Konvexe Optimierungsprobleme

Definition: (konvexe Mengen und Funktionen)
$D \subset X$ bzw. $D \subset \mathbb{R}^n$ heißt *konvex*, falls für alle $x, y \in D$ und $\alpha \in [0, 1]$ gilt:

$$\alpha x + (1 - \alpha) y \in D$$

Eine Funktion $f : D \to \mathbb{R}$ heißt *konvex* auf D, falls für alle $x, y \in D$ und $\alpha \in (0, 1)$ gilt:

$$f(\alpha x + (1 - \alpha) y) \leq \alpha f(x) + (1 - \alpha) f(y)$$

0.1.2 nichtlineare differenzierbare Optimierung

In dieser Vorlesung sei immer $X := \mathbb{R}^n, ||\cdot|| := ||\cdot||_2$ oder $||\cdot|| := ||\cdot||_\infty, x = (x_1, \ldots, x_n)^T$ und $f(x) = \big(f(x_1), \ldots, f(x_n)\big)^T$.

Beispiele:

1. Bestimme die größte Fläche eines Rechteckes bei gegebenem Umfang U. Maximiere also $f(x,y) = x \cdot y$ unter der Nebenbedingung $2x + 2y = U$.

   Idee: Elimination von $y = \frac{1}{2}(U - 2x)$. Lösung: Quadrat mit $x = y = \frac{U}{4}$

   $S \subset \mathbb{R}^2$ wird dabei durch die Gleichung $2x + 2y = U$ beschrieben

2. Bestimme die Extrema von $f(x,y) = \cos x + \sin y - \cos(x+y)$ unter der Nebenbedingung $x \geq 0, y \geq 0, x + y \leq 2\pi$. Dabei wird $S \subseteq \mathbb{R}^2$ beschrieben durch die linearen Ungleichungen.

3. Zeige:

   $$\max\left\{\prod_{i=1}^{n} x_i \;\middle|\; \sum_{i=1}^{n} x_i = 1, x_i \geq 0\right\} = \left(\frac{1}{n}\right)^n$$

   Folgerung:

   $$\forall x_i \geq 0 : \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} \leq \frac{x_1 + \ldots + x_n}{n}$$

4. hübsches Bildchen hier

0.1.1 Satz: *Optimalitätsbedingungen für Minimalwerte*
Sei $D \subset \mathbb{R}^n$ offen, $f : D \to \mathbb{R}$ eine C^2-Funktion, $S \subset D$. Für die lokale Minimalstelle $\bar{x} \in S$ von f gelte $\bar{x} \in \overset{\circ}{S}$.

Notwendige Optimalitätsbedingung:

$$\nabla f\big(\bar{x}\big) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} = 0$$

$$f''\big(\bar{x}\big) = \nabla^2 f\big(\bar{x}\big) = \text{Hess } f\big(\bar{x}\big) \text{ ist positiv semidefinit}$$

Hinreichende Optimalitätsbedingung:

$$\nabla f\big(\bar{x}\big) = 0$$

$$f''\big(\bar{x}\big) = \nabla^2 f\big(\bar{x}\big) = \text{Hess } f\big(\bar{x}\big) \text{ ist positiv definit}$$

0.1.3 Lineare Optimierungsprobleme

Sei $f : \mathbb{R}^n \to \mathbb{R}$ definiert durch

$$f(x) := c \cdot x = \sum_{i=1}^{n} c_i x_i$$

mit $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$ und $S \subset \mathbb{R}^n$ beschrieben durch lineare Gleichungen und/oder Ungleichungen.

Beispiel:
Produziere x Tische, y Stühle und maximiere dabei den Gewinn $f(x, y) = 6x + 8y$ unter den Nebenbedingungen

$$\begin{aligned} 3x + 2y &\leq 30 \qquad \text{(Holz)} \\ 5x + 10y &\leq 110 \qquad \text{(Arbeitsstunden)} \\ x, y &\geq 0 \end{aligned}$$

Kapitel 1

Lineare Optimierung

1.1 Beispiele

1.1.1 Produktionsmodelle

Allgemein:
Wir haben n Produkte $P_1, \ldots, P_n$ und m Aktivitäten $A_1, \ldots, A_m$ (z.B. Rohstoffe, Arbeitszeit/-stunden, ...). Das Produkt P_j enthalten a_{ij} Anteile an der Aktivität A_i $(j = 1, \ldots, n, i = 1, \ldots, m)$.
Der Reingewinn einer Einheit P_j sei $c_j > 0, x_j$ bezeichne die Produktionsmenge von P_j, b_i sei die Menge der verfügbaren Aktivität A_i.

Lineare Optimierungsproblem (LP):
Maximiere den Gewinn

$$z = \sum_{j=1}^{n} c_j x_j$$

unter den Nebenbedingungen

$$\begin{aligned} \sum_{j=1}^{n} a_{ij} x_j &\leq b_i, \qquad i = 1, \ldots, m \\ x_1, \ldots, x_n &\geq 0 \end{aligned}$$

Beispiel:
In einer Schuhfabrik werden Damen- und Herrenschuhe hergestellt. Produktionsplan:

	Damenschuh	Herrenschuh	verfügbar
Herstellungszeit:	20	10	8000
Maschinenbedarf	4	5	2000
Lederbedarf	6	15	4500
Gewinn	16	32	

Es sei x_1 die Anzahl der produzierten Damenschuhe, x_2 die Anzahl der produzierten Herrenschuhe.

LP: Maximiere $z = 16x_1 + 32x_2$ unter den Nebenbedingungen

$$\begin{aligned} 20x_1 + 10x_2 &\leq 8000 \\ 4x_1 + 5x_2 &\leq 2000 \\ 6x_1 + 15x_2 &\leq 4500 \\ x_1, x_2 &\geq 0 \end{aligned}$$

Die zulässige Menge K ist ein konvexes Polyeder. Definiere Schlupfvariablen y_1, y_2, y_3 durch:

$$\begin{aligned} 20x_1 + 10x_2 + y_1 &= 8000 \\ 4x_1 + 5x_2 + y_2 &= 2000 \\ 6x_1 + 15x_2 + y_3 &= 4500 \\ x_1, x_2, y_1, y_2, y_3 &\geq 0 \end{aligned}$$

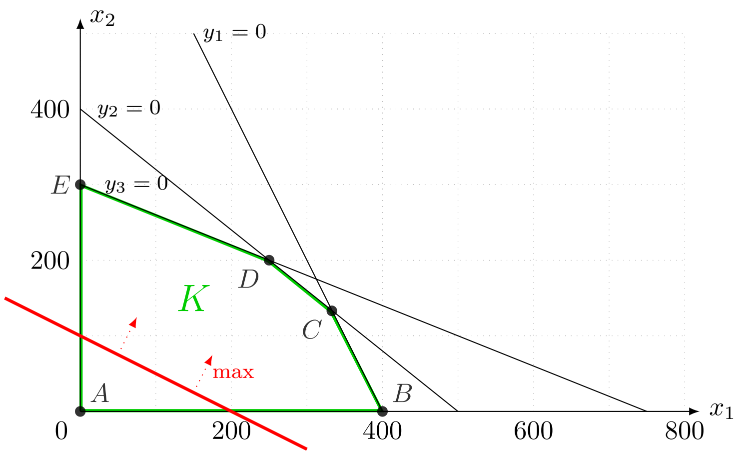


Abbildung 1.1: Grafische Lösung

Grafische Lösung: Zeichne zunächst die Geraden mit $y_1, y_2, y_3 = 0$. Betrachte die Niveaulinien der Funktion $z = 16x_1 + 32x_2 = \text{const}$. Der maximale Wert von z wird in der Ecke D angenommen.

Die Ecken von K sind:

Ecke	z	x_1	x_2	y_1	y_2	y_3
A	0	0	0	8000	2000	4500
B	6400	400	0	0	400	2100
C	9600	$\frac{1000}{3}$	$\frac{400}{3}$	0	0	500
D	10400	250	200	100	0	0
E	9600	0	300	5000	500	0

Bemerkung:
In jeder Ecke haben genau zwei Variablen von x_1, x_2, y_1, y_2, y_3 den Wert 0.

Basislösung: Ecke D, maximaler Wert $z = 10400$.

Optimale Lösung: $x_1 = 250, x_2 = 200$.

1.1.2 Mischungsproblem

Beispiel:
Drei Gase sollen so gemischt werden, dass ein möglichst biliges Mischgas mit einem Heizwert von 3000 kcal/m^3 und einem Schwefelgehalt von höchstens 3 g/m^3 entsteht.

Gasnr.	1	2	3	Einheit
Preis:	10	30	20	Euro/1000m^3
Heizwert:	1000	3000	6000	kcal/m^3
Schwefelgehalt:	8	1	2	g/m^3

Es sei x_i der Mengenanteil des Gases Nr. i am Mischgas pro m^3 Mischgas.

LP: Minimiere $z = 10x_1 + 30x_2 + 20x_3$ unter den Nebenbedingungen

$$\begin{aligned} x_1 + 3x_2 + 6x_3 &= 3 \\ 8x_1 + x_2 + 2x_3 &\leq 3 \\ x_1 + x_2 + x_3 &= 1 \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

Lösungsidee: Substituiere $x_3 = 1 - x_1 - x_2 \geq 0$ und löse das verbleibende LP in Variablen x_1, x_2.
Lösung: $x_1 = \frac{6}{23}, x_2 = \frac{13}{23}, x_3 = \frac{4}{23}$.

1.1.3 Transportprobleme

Eine Fabrik F_i produziert die Warenmenge $a_i, i = 1\ldots, m$ und beliefert damit die Abnehmer A_j, die die Warenmenge $b_j, j = 1, \ldots, n$.
Die Transportkosten pro Einheit von Fabrik F_i zu Abnehmer A_j bestimmen wir durch c_{ij}. x_{ij} seien die tatsächlich von F_i nach A_j versandten Mengen.
Das lineare Optimierungsproblem ist gegeben durch:

$$\min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

unter

$$\begin{aligned} \sum_{j=1}^{n} x_{ij} = a_i. \qquad & i = 1, \ldots, m \\ \sum_{j=1}^{n} x_{ij} = b_i. \qquad & i = 1, \ldots, m \\ x_{ij} \geq 0, \qquad & i = 1, \ldots, m, j = 1, \ldots, n \end{aligned}$$

1.2 Mathematische Formulierung linearer Optimierungsprobleme

Allgemein betrachtet können wir lineare Optimierungsprobleme formulieren durch:

$$\min \ z(x) = \sum_{j=1}^{n} c_j x_j$$

unter den Nebenbedingungen:

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m \\ x_1, \ldots, x_n & \geq & 0 \end{array}$$

1.2.1 Definition: (Standardform eines linearen Optimierungsproblemes)
Definieren wir uns die folgende vektorielle Schreibweise:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, c = (c_1, \ldots, c_n) \in \mathbb{R}^n \text{ Zeilenvektor}$$

$$b = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{R}^m, A = (a_{ij}) \ \ m \times n \text{ Matrix}$$

dann ist die Standardform eines linearen Optimierungsproblemes gegeben durch:

$$\min z(x) = cx \text{ unter } Ax = b, x \geq 0$$

Dabei nehmen wir an, dass $m < n, \ \text{rang } (A) = m$

1.2.1 Rückführung anderer Typen von LP's auf die Standardform

Behandlung von Problemen mit Ungleichungen:
Sei A eine $m \times n$ Matrix, aber es gelte nicht notwendigerweise $m < n$.

$$\min\{cx \mid AX \leq b, x \geq 0\} \tag{1.2.2}$$

Definiere folgende *Schlupfvariable* $y := b - Ax, y \in \mathbb{R}^m$, sodass $Ax \leq b \Leftrightarrow Ax + y = b, y \geq 0$.
Setze

$$\begin{aligned} \tilde{c} &= (c, 0) \in \mathbb{R}^{n+m} \\ \tilde{x} &= \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+m} \\ \tilde{A} &= (A, E_m) \in \mathbb{R}^{m \times (n+m)} \end{aligned}$$

dann ist das LP äquivalent zur Standardform

$$\min\left\{\tilde{c}\tilde{x} \;\middle|\; \tilde{A}\tilde{x} = Ax + y = b, x \geq 0, y \geq 0\right\} \tag{1.2.3}$$

Bemerkung:
Die Ungleichung $Ax \geq b$ ist äquivalent zum Gleichungssystem $Ax - y = b, y \geq 0$.

Behandlung von Problemen mit freien Variablen:
In *(1.2.1)* seien einige x_j ohne Vorzeichenbedingung zu bestimmen, z.B. gelte nicht $x_1 \geq 0$. Wir erhalten also das LP

$$\min\{cx \mid Ax = b, x_2, \ldots, x_n \geq 0\} \tag{1.2.4}$$

1. Methode: Elimination von x_1
Es gelte $a_{i1} \neq 0$. Betrachte die i-te Gleichung $a_{i1}x_1 + \ldots + a_{in}x_n = b_i$ und transformiere sie in

$$x_1 = \frac{b_i - a_{i2}x_2 - \ldots - a_{in}x_n}{a_{i1}}$$

Setze dies in Gleichungen mit Index $j \neq i$ ein:

$$Ax = b \Leftrightarrow \text{ reduziertes LGS } \tilde{A}\tilde{x} = \tilde{b}, \tilde{x} = \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} \in R^{n-1}$$

Beispiel:
$\min x_1 + 3x_2 + 4x_3$ unter

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 5 \\ 2_x1 + 3_x2 + x_3 &= 6 \\ x_2, x_3 &\geq 0 \end{aligned}$$

Eliminiere $x_1 = 5 - 2x_2 - x_3$ und erhalte das reduzierte LP: $\min x_2 + 3x_3$ [+5] unter

$$\begin{aligned} x_2 + x_3 &= 4 \\ x_2, x_3 &\geq 0 \end{aligned}$$

Als Lösung erhalten wir $x_2 = 4, x_3 = 0, x_1 = -3$.

2. Methode:

Schreibe x_1 als $x_1 = u_1 - v_1$ mit $u_1 = x_1^+ := \max\{0, x_1\} \geq 0, v_1 = x_1^- := \max\{0, -x_1\} \geq 0$. Dies ergibt ein lineares Optimierungsproblem in den Variablen $u_1, v_1, x_2, \dots, x_n \geq 0$.

Bemerkung:
Es kann in der Zielfunktion der Absolutbetrag $|x_1|$ vorkommen (dann haben wir kein Standardproblem mehr). In diesem Fall setze

$$|x_1| = u_1 + v_1, \quad u_1 = x_1^+, v_1 = x_1^-$$

und erhalte wieder ein Standardproblem.

Elementare Eigenschaften von LPs:
Die zulässige Menge $K = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ von *(1.2.1)* bzw. diejenige von *1.2.2* ist ein konvexes und abgeschlossenes Polytop.

Ein zulässiger Punkt $\bar{x} \in K$ heißt eine *optimale Lösung*, wenn

$$z(\bar{x}) = x\bar{x} = \min\{cx \mid x \in K\}$$

$\bar{x}$ also auch ein globales Minimum ist.

1.2.5 Satz:

(i) Die Menge der optimalen Lösungen ist konvex

(ii) Ist die zulässige Menge K beschränkt (also kompakt), so existiert eine optimale Lösung.

Situation für $\max\{cx \mid Ax \leq b, x \geq 0$:

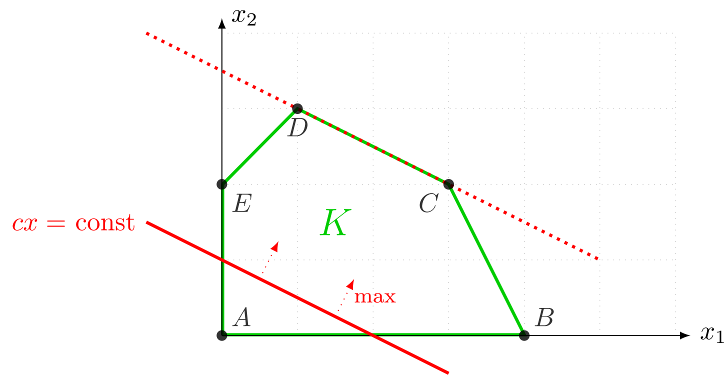


Abbildung 1.2: Grafische Anschauung

Für die folgenden Überlegungen benötigen wir:

1. Ecke einer konvexen Menge
2. Konvexe bzw. konische Hülle einer Menge
3. Trennungssatz für konvexe Menge und Kegel

1.3 Hauptsatz der Linearen Optimierung

Seien A eine $m \times n$ Matrix, $b \in \mathbb{R}^m$ und $c \in \mathbb{R}^n$ ein Zeilenvektor. Standardform eines linearen Programmes:

$$\min \left\{ cx \mid Ax = b, x \geq 0 \right\} \tag{1.3.1}$$

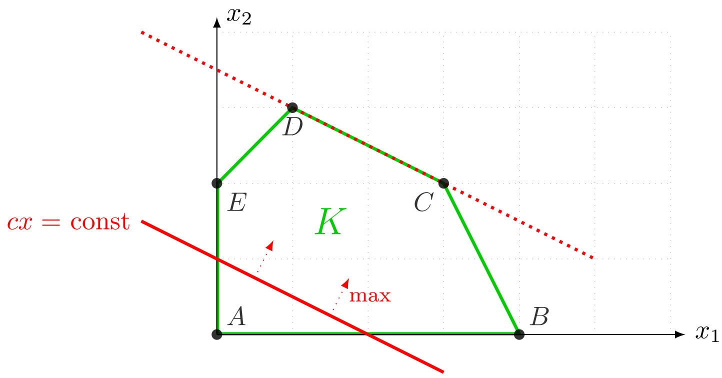


Abbildung 1.3: Situation für Ungleichungen $Ax \leq b$

Problem:
Beschreibe die Ecken der zulässigen Menge $K := \left\{ x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \right\}$.

Notation:
Wir bezeichnen die Spaltenvektoren der Matrix A mit hochgestelltem Index: $a^1, \ldots, a^n \in \mathbb{R}^m$.

$$Ax = \left(a^1 \mid \cdots \mid a^n \right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a^1 x_1 + \ldots + a^n x_n = b$$

Voraussetzung: $m < n$.

1.3.2 Satz:
$x \in K$ ist genau dann eine Ecke von K, wenn die zu positiven Komponenten $x_i > 0$ gehörenden Spaltenvektoren $a^i \in \mathbb{R}^m$ linear unabhängig sind.

Beweis:

„⇒“ Sei $x \in K$ eine Ecke von K. Ohne Einschränkung sei $x = (x_1, \ldots, x_r, 0, \ldots, 0)^T \in \mathbb{R}^n$ mit $x_i > 0$ für $i = 1, \ldots, r$. Aus trivialen Gründen gilt die Behauptung für $r = 0$. Sei also $r > 0$ und gelte $\sum_{i=1}^r a_i d_i = 0$ mit $d_i \in \mathbb{R}$. Zu zeigen ist nun $\forall i : d_i = 0$.

Wegen $x_i > 0$ für $i = 1, \ldots, r$ gibt es ein $\varepsilon > 0$ mit $x_i \pm \varepsilon d_i > 0$. Setze $d := (d_1, \ldots, d_r, 0, \ldots, 0)^T \in \mathbb{R}^n$ und $y := x + \varepsilon d$, $z := x - \varepsilon d$.

$$\Rightarrow x = \frac{1}{2}(y + z) = \frac{1}{2} y + \left(1 - \frac{1}{2} \right) z$$

Es gilt $y \geq 0$ und $z \geq 0$ sowie

$$\sum_{i=1}^n a^i \left(x_i \pm \varepsilon d_i \right) \; = \; \sum_{i=1}^r a^i x_i \pm \varepsilon \underbrace{\sum_{i=1}^r a^i d_i}_{=0} \; = \; b$$

Also sind $y, z \in K$ und $x \in K$ eine zulässige Ecke ist, so folgt mit *(3.3.1)* bereits $x = y = z$ und damit $d = 0$. Also sind $a^1, \ldots, a^r$ linear unabhängig.

„$\Leftarrow$" Sei $x = (x_1, \ldots, x_r, 0, \ldots, 0)^T \in \mathbb{R}^n$ mit $x_i > 0$ für $i = 1, \ldots, r$ sowie $a^1, \ldots, a^r$ linear unabhängig. Zu zeigen ist nun, dass $x \in K$ eine Ecke ist.

1. Fall: $r = 0$

In diesem Fall ist $x = 0 \in \mathbb{R}^n$. Es gelte $0 = x = \alpha y + (1-\alpha)z,\ y, z \in K,\ 0 < \alpha < 1$. Wegen $y, z \in K$ gilt insbesondere $y, z \geq 0$. Es folgt sofort dass $x = y = z = 0$, also x eine Ecke ist.

2. Fall: $r > 0$

Es gelte $\sum_{i=1}^r a^i x_i = b$ sowie $x = \alpha y + (1-\alpha)z$ mit $y, z \in K, 0 < \alpha < 1$. Wie in Fall $r = 0$ folgt $x_i = y_i = z_i = 0$ für $i > r$. Wegen $y, z \in K$ gilt: $Ay = b, Az = b$, also:

$$A(y-z) \ = \ \sum_{i=1}^r a^i (y_i - z_i) \ = \ 0$$

Da nun die $a^1, \ldots, a^r$ linear unabhängig sind, muss schon $y_i - z_i = 0$ also $y_i = z_i = x_i$ für $i = 1, \ldots, r$ sein.

Also ist $x = y = z$ eine Ecke.

□

1.3.3 Korollar:
Ist $x \in K$ eine Ecke, so sind höchstens m Komponenten $x_i > 0$ Daher gibt es höchstens endlich viele Ecken:

$$\#\text{Ecken} \leq \sum_{i=0}^m \binom{n}{i}$$

1.3.4 Satz:
Ist die zulässige Menge $K = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ nicht leer, so gibt es eine Ecke von K.

Beweis:
Wir bezeichnen mit $r(x)$ die Anzahl der positiven Komponenten $x_i > 0$ von $x \in K$. Sei $z \in K$ mit $r(z) = \min\{r(x) \mid x \in K\}$ (z exisitiert wegen $r(x) \in \mathbb{N}$). Wir zeigen nun, dass z eine Ecke von K ist:

Falls $r(z) = 0$, dann ist $z = 0$ eine Ecke von K (vgl. Beweis von *(1.3.2)*).

Sei also $r(z) > 0$ und ohne Einschränkung sei $z_i > 0$ für $i = 1, \ldots, r(z)$, dann ist zu zeigen, dass $a^1, \ldots, a^{r(z)}$ linear unabhängig sind.

Angenommen $a^1, \ldots, a^{r(z)}$ seien linear abhängig, das heißt $\sum_{i=1}^{r(z)} a^i d_i = 0$ mit $(d_1, \ldots, d_{r(z)})^T \neq 0$. Setze

$$\mu := \left\{ \frac{z_i}{|d_i|} \ \middle| \ d_i \neq 0 \right\} = \frac{z_k}{|d_k|}$$

Ohne Einschränkung sei $d_k > 0$. Setze $d := (d_1, \ldots, d_{r(z)}, 0 \ldots, 0)^T \in \mathbb{R}^n$ und $y := z - \mu d \in \mathbb{R}^n$. Für diesen Punkt gilt

$$Ay \ = \ Az - \mu \underbrace{Ad}_{=0} \ = \ Az \ = \ b$$

Es bleibt zu zeigen, dass $y \geq 0$. Es gilt $y_i = i_i - \frac{z_k}{d_k} d_i \geq 0$ wegen $\mu = \frac{z_k}{d_k} = \left\{ \frac{z_i}{|d_i|} \ \middle| \ d_i \neq 0 \right\}$. Also ist $y \in K$.

So erhalten wir jedoch $y_k = 0$ und damit $r(y) \geq r(z) - 1$ was einen Widerspruch zur minimalen Wahl von z. Also ist $z \in K$ eine Ecke von K. □

Bemerkung:
Sei $L \subset \mathbb{R}^n$ die Menge der optimalen Lösungen des linearen Optimierungsproblemes $\min\{cx \mid Ax = b, x \geq 0\}$. Falls $L \neq \emptyset$, setze $\beta := \min\{cx \mid Ax = b, x \geq 0\}$. Definiere

$$\tilde{A} := \begin{pmatrix} c \\ A \end{pmatrix} \ (m+1) \times n \text{ Matrix}, \quad \tilde{b} := \begin{pmatrix} \beta \\ b \end{pmatrix} \in \mathbb{R}^{m+1}$$

Damit folgt:

$$\begin{aligned} L &= \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0, cx = \beta\} \\ &= \{x \in \mathbb{R}^n \mid \tilde{A}x = \tilde{b}, x \geq 0\} \end{aligned}$$

Damit hat L die gleiche Struktur wie die zulässige Menge K.

1.3.5 Satz:
Sei $L \neq \emptyset$, dann gilt:

(i) L besitzt eine Ecke.

(ii) Jede Ecke von L ist eine Ecke von K.

Beweis:

1. Folgt direkt aus Satz *(1.3.4)*.

2. Sei $x \in L$ eine Ecke von L, etwa $x = \alpha y + (1-\alpha)z$ mit $y, z \in K, \alpha \in (0,1)$. Dann gilt

$$\beta = cx = \alpha \cdot cy + (1-\alpha) \cdot cz$$

Mit $\beta \leq cy$ und $\beta \leq cz$ folgt $\beta = cy = cz$. Also sind $y, z \in L$ optimale Läsungen. Da $x \in L$ eine Ecke von L ist, so folgt $x = y = z$. Also ist $x \in K$ eine Ecke von K.

$\square$

1.3.6 Satz: *Hauptsatz der Linearen Optimierung*
Das LP $\min\{cx \mid Ax = b, x \geq 0\}$ besitze eine optimale Lösung, d.h. $L \neq \emptyset$. Dann gibt es eine optimale Ecke in K, d.h. es gibt eine Ecke von K, welche eine optimale Lösung ist.

Beweis:
Folgt direkt aus den letzten Sätzen. $\square$

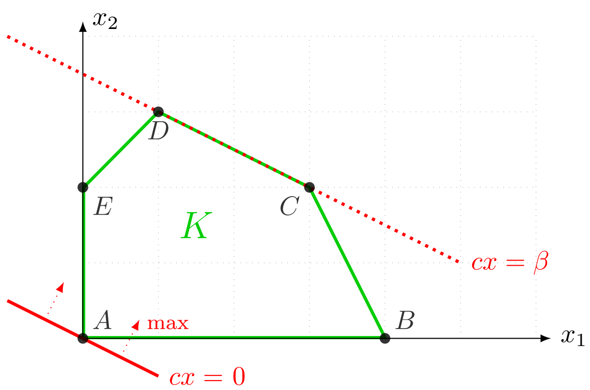


Abbildung 1.4: Situation für Ungleichungen $Ax \leq b$

Idee zur Lösung eines LP:
Berechne alle Ecken von K und bestimme cx.

Beispiel:

$$\max x_1 + 2x_2 + 3x_3 \quad \text{unter} \quad \begin{aligned} 2x_1 + x_2 + 5x_3 &= 5 \\ x_1 + 2x_2 + x_3 &= 4 \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

Ecken: Genau eine Komponente einer Ecke von K ist gleich Null.

1. Fall: $x_3 = 0$
Löse das LGS

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \quad \Rightarrow \quad x := \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \geq 0 \text{ mit } cx = 4$$

2. Fall: $x_2 = 0$
Löse das LGS

$$\begin{pmatrix} 2 & 5 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \quad \Rightarrow \quad x := \begin{pmatrix} 5 \\ 0 \\ -1 \end{pmatrix} \text{ ist nicht zulässig!}$$

3. Fall: $x_1 = 0$
Löse das LGS

$$\begin{pmatrix} 1 & 5 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \quad \Rightarrow \quad x := \begin{pmatrix} 0 \\ \frac{5}{3} \\ \frac{2}{3} \end{pmatrix} \geq 0 \text{ mit } cx = \frac{16}{3} \text{ ist optimal}$$

Einfacher Beweis des Hauptsatzen, falls die zulässige Menge K kompakt ist:
Seien $x^1, \ldots, x^k$ die Ecken von K. Mit Satz *(3.1.8.3)* folgt

$$K = \operatorname{co}\left(x^1, \ldots, x^k\right) = \left\{ \sum_{i=1}^{k} \alpha_i x^i \;\middle|\; \sum_{i=1}^{k} \alpha_i = 1, \alpha_i \geq 0 \text{ für } i = 1, \ldots, k \right\}$$

Sei $cx^j = \min_{i=1,\ldots,k} cx^i$. Zu zeigen ist nun, dass $cx^j = \min\{cx \mid x \in K\}$. $x \in K$ hat die Darstellung

$$\sum_{i=1}^{k} \alpha_i x^i, \sum_{i=1}^{k} \alpha_i = 1, \alpha_i \geq 0 \text{ geeignet}$$

Also gilt:

$$\forall x \in K : cx = \sum_{i=1}^{k} \alpha_i \cdot \underbrace{cx^i}_{\geq cx^j} \geq cx^j \cdot \sum_{i=1}^{k} \alpha_i = cx^j$$

und damit ist x^j eine optimale Ecke.

1.4 Das Simplex-Verfahren

Wir betrachten ein Standard-LP

$$\min \left\{ cx \mid Ax = b, x \geq 0 \right\} \tag{1.4.1}$$

und setzen $\operatorname{rang}(A) = m < n$ voraus. Unser Ziel ist die rechnerische Bestimmung aller Ecken von $K := \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$.
Simplex-Verfahren:

1. Schritt: Bestimmung einer *zulässigen* Basislösung.

2. Schritt: Entscheidung, ob die vorliegende Basislösung optimal ist ($r_n \leq 0$).

Notation:
$a^i \in \mathbb{R}^m, i = 1, \ldots, n$ seien die Spalten von A.

1.4.2 Definition: (Basis)
Ein Indexvektor $B := \big(i_1, \dots, i_m\big)$ mit m paarweise verschiedenen Indizes $i_k \in \{1, \dots, n\}, k = 1, \dots, m$ heißt *Basis*, wenn die Spaltenvektoren $a^{i_k}, k \in 1, \dots, m$ linear unabhängig sind.

Ein zu B komplementärer Indexvektor $N = \big(j_1, \dots, j_{n-m}\big)$ mit paarweise verschiedenen Indizes $j_k \in \{1, \dots, n\}, k = 1, \dots, n-m$ heißt *Nichtbasis*, falls $B \oplus N = \{1, \dots, n\}$.

Notation:
Für $x \in \mathbb{R}^n$ heißen $x_i, i \in B$ *Basis-Variablen* und die $x_j, j \in N$ *Nichtbasis-Variablen*.

Weiter bezeichnen wir $x_B := (x_i)_{i \in B} \in \mathbb{R}^m, x_N := (x_j)_{j \in N}$.

Die *fundamentale Basismatrix* A_B ist die $m \times m$ Matrix Spalten $a^i, i \in B$. Analog definieren wir $A_N = \big(a^{j_1} \cdots a^{j_{n-m}}\big)$. Die Reihenfolge der Indizes spielt dabei natürlich eine Rolle!

Beispiel:

$$A := \begin{pmatrix} 1 & 2 & 2 & 4 \\ 0 & 3 & 0 & 6 \end{pmatrix}$$

Dann sind mögliche Basen gegeben durch $B = (1,2), B = (1,4), B = (2,3), B = (3,4)$ aber auch $B = (2,1), B = (4,1)$, etc. Die Nichtbasen sind dann gegeben durch $N = (3,4)$ zu $B = (1,2), N = (4,1)$ zu $B = (2,3)$, etc.

Folgerung: Das LGS $Ax = b$ geht über in

$$Ax = \sum_{i \in B} a^i x_i + \sum_{j \in N} a^j x_j = A_B x_B + A_N x_N = b$$

Parametrisierung des $(n-m)$-dimensinalen Lösungsraumes von $Ax = b$:

$$x_B = A_B^{-1} b - A_B^{-1} A_N x_N \tag{1.4.3}$$

Dabei ist x_B die abhängige Variable und $x_N \in \mathbb{R}^{n-m}$ ist die unabhängige Variable.

Idee:
Reduziere das LP auf die unabhängigen Vriablen x_N. Zerlege $c \in \mathbb{R}^n$ in $c_B := (c_i)_{i \in B} \in \mathbb{R}^m$ und $c_N := (c_j)_{j \in N} \in \mathbb{R}^{n-m}$.

Darstellung der Zielfunktion:

$$cx \;=\; c_B x_B + c_N x_N \overset{1.4.3}{=} c_B A_B^{-1} b - \big(c_B A_B^{-1} A_N - c_N\big)\, x_N \;=\; z_0 - r_N X_N \tag{1.4.4}$$

mit dem Vektor der *reduzierten Kosten*:

$$z_0 = c_B A_B^{-1} b, \quad r_N := c_B A_B^{-1} A_N - c_N \in \mathbb{R}^{n-m} \text{ (Zeilenvektor)}$$

1.4.5 Definition: (Basislösung)
Sei B eine Basis. Dann heißt ein $x \in \mathbb{R}^n$ *Basislösung* von $Ax = b$, falls $x_N = 0, x_B = A_B^{-1} b$. Eine Basislösung $x \in \mathbb{R}^n$ heißt

(i) *zulässig*, falls $x_B = A_B^{-1} b \geq 0$ für $x_N = 0$.

(ii) *nicht-entartet*, falls $x_B = A_B^{-1} b > 0$ für $x_N = 0$.

(iii) *entartet*, falls $x_i = 0$ für ein $i \in B$.

1.4.6 Satz: *Hinrichendes Optimalitätskriterium*
Sei B eine Basis mit den Eigenschaften

(i) Die zugehörige Basislösung $x_B = A_B^{-1}b, x_N = 0$ ist zulässig. („primale Bedingung“)

(ii) $r_N = c_B A_B^{-1} A_N - c_N \leq 0$. („duale Bedingung“)

Dann ist $x \in \mathbb{R}^n$ optimal für das LP *1.4.1* und der optimale Wert ist gegeben durch:

$$z(x) = cx = z_0 = c_B a_B^{-1} b$$

Beweis:
Es sei $x \in K$ zulässig, also $x_B = A_B^{-1}b \geq 0, x_N = 0$. Dann gilt für die Zielfunktion nach *1.4.4*:

$$z(x) = cx = z_0 - \underbrace{\underbrace{r_N}_{\leq 0}\ \underbrace{x_N}_{\geq 0}}_{\geq 0}$$

Das Minimum wird also für $x_N = 0$ angenommen. □

1.4.7 Anwendung auf den Spezialfall: LP mit Ungleichungen
Wir hatten uns bereits überlegt:

$$\min\{cx \mid Ax \leq b, x \geq 0\} \Leftrightarrow \min\{cx + 0y \mid Ax + y = b, x \geq 0, y \geq 0\}$$

$$Ax + y = \tilde{A}\binom{x}{y}, \tilde{A} = \begin{pmatrix} A & E_m \end{pmatrix}$$

Wähle Basis $B = (n+1, \ldots, n+m)$ und Nichtbasis $N = (1, \ldots, n)$. Basislösung ist $x = 0, y = b$, zulässig falls $b \geq 0$. Die reduzierten Kosten

$$\tilde{A}_B = E_m,\ \tilde{A}_N = A,\ \tilde{c}_B = 0,\ \tilde{c}_N = c \quad \Rightarrow \quad r_N = -c$$

Grundidee des Simplex-Verfahrens:
Bestimme iterativ eine Basis B mit

(i) $x_B \geq 0, x_N = 0$

(ii) $r_N \leq 0$.

Beispiel:
$\max x_1 + 2x_2 + 3x_3$ also $\min z(x) = -x_1 - 2x_2 - 3x_3$ unter

$$Ax = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \binom{5}{4} = b, x \geq 0$$

Rechnungen für die *Tableau-Matrix*:

b	A

$$\begin{array}{|c|ccc|} \hline 5 & 2 & 1 & 5 \\ 4 & 1 & 2 & 1 \\ \hline \end{array} \rightarrow \begin{array}{|c|ccc|} \hline \frac{5}{2} & 1 & \frac{1}{2} & \frac{5}{2} \\ 4 & 1 & 2 & 1 \\ \hline \end{array} \rightarrow \begin{array}{|c|ccc|} \hline \frac{5}{2} & 1 & \frac{1}{2} & \frac{5}{2} \\ \frac{3}{2} & 0 & \frac{3}{2} & -\frac{3}{2} \\ \hline \end{array} \rightarrow \begin{array}{|c|ccc|} \hline \frac{5}{2} & 1 & \frac{1}{2} & \frac{5}{2} \\ 1 & 0 & 1 & \text{-1} \\ \hline \end{array} \rightarrow \begin{array}{|c|ccc|} \hline \text{b} & x_1 & x_2 & x_3 \\ \hline 2 & 1 & 0 & 3 \\ 1 & 0 & 1 & \text{-1} \\ \hline \end{array}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A_B^{-1} b - A_B^{-1} A_N x_N = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 3 \\ -1 \end{pmatrix} x_3$$

Die Basislösung $x = (2, 1, 0)$ mit $x_3 = 0$ ist zulässig aber nicht optimal, denn

$$\begin{aligned} r_N = r_3 &= c_B A_B^{-1} A_N - c_N \\ &= (-1, -2) \begin{pmatrix} 3 \\ -1 \end{pmatrix} - 3(-3) \\ &= 2 \geq 0 \end{aligned}$$

wegen $c = (-1, -2, -3)$ für *min*.

Wähle eine neue Basis: Für $0 \leq x_3 \leq \frac{2}{3}$ bleibt die Basislösung zulässig:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 - 3x_3 \\ 1 + x_3 \end{pmatrix} \geq 0$$

Übergang: $x_3 = 0 \rightarrow x_3 = \frac{2}{3}$ läuft entlang einer Kante der zulässigen Menge. Für den Wert $x_3 = \frac{2}{3}$ erhält man die zulässige Basislösung $x = (0, \frac{5}{3}, \frac{2}{3})^T$.

Dieser Übergang entspricht einem sogenannten Basistausch:

$$B = (1, 3) \rightarrow B' = (3, 2), \; N = (3) \rightarrow N' = (1)$$

Darstellung von $Ax = b$ bzgl. B' und N': $x_3 = \frac{2}{3} - \frac{1}{3} x_1$ Einsetzen in die 2. Gleichung:

$$\begin{aligned} x_{B'} &= \begin{pmatrix} \frac{2}{3} \\ \frac{5}{3} \end{pmatrix} - \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} x_1 \\ &\Rightarrow \; r_{N'} = r_1 = c_{B'} A_{B'}^{-1} A_{N'} - c_{N'} \\ &= \; (-3, -2) \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} - (-1) = -\frac{2}{3} < 0 \text{ optimal} \end{aligned}$$

1.4.8 Beisiel
Vorüberlegungen zum Basistausch:

Vorausgesetzt sei zur Vereinfachung $B = (1, \dots, m)$ und es gelte $A_B = E_m$ (allgemeiner sei $A_B = P_m$ Permutationsmatrix). Es gelte $x_B \geq 0$. Gilt $r_N \leq 0$, so ist $x_B = A^{-1}b = b, x_N = 0$ optimal.

Sei nun $r_S > 0$ für ein $s \in N$. Setze $x_j = 0$ für $j \in N \setminus \{s\}$ in $x_B = b - A_N x_N$.

$$x_B = b - a^s x_s \tag{1.4.9}$$

1. Fall: $a^s \leq 0$
Dann bleibt $x_B = b - a^s x_s \geq 0$ für alle $x_s \geq 0$. Aber $z(x) = cx = z_0 - r_s x_s \to -\infty$ für $x_s \to +\infty$ wegen $r_s > 0$. Also gibt es keine endliche Lösung.

2. Fall: $a_{is} > 0$ für mindestens ein $i \in \{1, \dots, m\}$
Ziel: Erhalte die Zulässigkeit $x_B \geq 0$ bei der Wahl von x_s. Dazu definieren wir uns einen Index $p \in B$ mit:

$$\frac{b_p}{a_{ps}} = \min\left\{ \frac{b_i}{a_{is}} \;\middle|\; a_{is} > 0 \right\} \tag{1.4.10}$$

Um $x_B \geq 0$ zu garantieren, muss $x_s \leq \frac{b_p}{a_{ps}} =: b'_p$. Es gilt $x_p = 0$ für $x_s = \frac{b_p}{a_{ps}}$! Dies entspricht einem Tausch der Nichtbasisvariablen x_s gegen die Basisvariable x_p.

Wir nennen $a_{ps} > 0$ das *Pivotelement.*

Fall a: $b_p = 0$
Die Basislösung ist entartet und es gilt $z(x) = z_0$. Wir erhalten also keine Verbesserung der Zielfunktion.

Fall b: $b_p > 0$ (hinreichend ist $x_B > 0$, d.h. eine nicht-entartete Basislösung) Wir folgern an dieser Stelle ein notwendiges Optimalitätskriterium: *(1.4.11).*

1.4.11 Korollar: *Notwendiges Optimalitätskriterium*
Seien die gleichen Voraussetzungen wie bei *(1.4.8)* gegeben. Die Basislösung x mit $x_B \geq 0$ und $x_N = 0$ sei optimal. Gilt $x_B > 0$, dann folgt $r_N \leq 0$.

Bemerkung:
Die Indizes $s \in N$ mit $r_s > 0$ und $p \in B$ mit *1.4.10* sind im Allgemeinen *nicht* eindeutig bestimmt.

1.4.12 Möglichkeiten zur Festlegun von s und p:

(i) Wähle $s \in N$ kleinstmöglich mit $r_s = \max\{r_j \mid j \in N\}$. Wähle den kleinsten Index $p \in B$ mit $\frac{b_p}{a_{ps}} = \min\left\{ \frac{b_i}{a_{is}} \mid a_{is} > \right\}$.

(ii) (Regel von BLAND) Wähle den kleinsten Index $s \in N$ mit $r_s > 0$. Wähle den kleinsten Index $p \in B$, sodass b_p kleinstmöglich ist mit $\frac{b_p}{a_{ps}} = \min\left\{ \frac{b_i}{a_{is}} \mid a_{is} > \right\}$.

Basistausch mit Gauß-Jordan Transformation:
Erinnerung: Sei $a_B = E_m, x_B = b - A_N x_N$.

$$x_i = b_i - \sum_{j \in N} a_{ij} x_j \tag{1.4.13}$$

Betrachte die Gleichung mit Index $p \in B$:

$$x_p = b_p - \sum_{j \in N} a_{pj} x_j = b_p - \sum_{\substack{j \in N, \\ j \neq s}} a_{pj} x_j - a_{ps} x_s$$

Auflösen nach x_s:

$$x_s = \frac{b_p}{a_{ps}} - \sum_{\substack{j \in N, \\ j \neq s}} \frac{a_{pj}}{a_{ps}} x_j - \frac{1}{a_{ps}} x_p \tag{1.4.14}$$

Setze x_s in die übrigen Gleichungen *1.4.13* mit $i \neq p$ ein:

$$x_i = b_i - a_{is}\frac{b_p}{a_{ps}} - \sum_{\substack{j\in N,\\ j\neq s}} \left(a_{ij} - a_{is}\frac{a_{pj}}{a_{ps}}\right) x_j + \frac{a_{is}}{a_{ps}} x_p \tag{1.4.15}$$

Definiere neue Basis: Übergang $B = \{1, \dots, m\} \to b' = \{1, \dots, p-1, s, p+1, \dots, m\}$. Neue Nichtbasis: $N = (m+1, \dots, n) \to N = (m+1, \dots, s-1, p, s+1, \dots, m)$.

Allgemein: $B \to B' := \big(B \setminus \{p\}\big) \cup \{s\}, N \to N' = \big(N \setminus \{s\}\big) \cup \{p\}$. Die Ausdrücke *1.4.14* und *1.4.15* haben die Form $x_{B'} = b' - a_{N'} x_{N'}$.

1.4.16 Algorithmus: Die Elemente

b'	$A_{N'}$

sind gegeben durch:

Pivotelement:

$a'_{ps} = \frac{1}{a_{ps}}$

Übrige Zeile p:

Dividiere durch das Pivotelement a_{ps}: $a'_{pj} = \frac{a_{pj}}{a_{ps}}$

Übrige Spalte:

Dividiere durch das negative Pivotelement: $a'_{is} = -\frac{a_{is}}{a_{ps}}$ für $i \neq p$

Sonstige Elemente in der i-ten Zeile:

Subtrahiere das a_{is}-fache der neuen Zeile p von der i-ten Zeile:

$a'_{ij} = a_{ij} - a_{is}\frac{a_{pj}}{a_{ps}} = a_{ij} - a_{is}a'_{pj}$ für $i \neq p, j \neq s$

$b'_i = b_i - a_{is}\frac{b_p}{a_{ps}} = b_i - a_{is}b'_p$

Merkregel:

Subtrahiere das a_{is}-fache der neuen Zeile p von der i-ten Zeile.

Beispiel:

$$\begin{array}{c|c|ccc|}
 & b & x_4 & x_5 & x_6 \\ \hline
x_1 & 5 & 1 & 1 & 1 \\
x_2 & 3 & \boxed{2} & -3 & 1 \\
x_3 & -1 & -1 & 2 & -1 \\ \hline
\end{array}
\to
\begin{array}{|c|ccc|} \hline
\frac{7}{2} & -\frac{1}{2} & \frac{5}{2} & -\frac{3}{2} \\
\frac{3}{2} & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \hline
\end{array}$$

Basistausch:
Zielfunktion und reduzierte Kosten:

$$\begin{aligned}
z(x) &= z_0 - \sum_{j \in N} r_j x_j \\
&\overset{1.4.14}{=} z_0 - \sum_{\substack{j \in N \\ j \neq s}} r_j x_j - r_s \left(\frac{b_p}{a_{ps}} - \sum_{\substack{j \in N \\ j \neq s}} \frac{a_{pj}}{a_{ps}} x_j - \frac{1}{a_{ps}} x_p \right) \\
&= z_0 - r_s \frac{b_p}{a_{ps}} - \sum_{\substack{j \in N \\ j \neq s}} \left(r_j - r_s \frac{a_{pj}}{a_{ps}} \right) x_j - \left(\frac{-r_s}{a_{ps}} \right) x_p \\
&= z_0' \sum_{j \in N'} r_j' x_j
\end{aligned}$$

$$z_0' = z_0 - r_s \frac{b_p}{a_{ps}}, \quad r_j' = r_j - r_s \frac{a_{pj}}{a_{ps}} \text{ für } j \neq p \tag{1.4.17}$$

Folgerung:
Die Zeile $\boxed{z_0 \mid r_N}$ wird wie die Zeile von $\boxed{b \mid A_N}$ transformiert!

1.4.18 Simplex-Tableau:

		$x_j (j \in N)$
	z_0	r_N
$x_i (i \in B)$	b	A_N

Beispiel:
(vgl. *(1.4.8)*): $\min z(x) = -x_1 - 2x_2 - 3x_3$ unter

$$Ax = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} = b, x \geq 0$$

Basis $B = (1,2), N = (3)$. Transformation des LGS $Ax = b$: [...]

$$\boxed{b \mid A_N} = \begin{array}{|c|c|} \hline 2 & 3 \\ 1 & -1 \\ \hline \end{array}, \quad x_B = \binom{2}{1} > 0$$

Aber: $r_3 = 2 > 0$ also nicht optimal

Simplex-Tableau:

	z_0	x_3
	-4	2
x_1	2	$\boxed{3}$
x_2	1	-1

Tausche x_3 ggn. x_1 $\Longrightarrow$

	z_0	x_1
	$-\frac{16}{3}$	$-\frac{2}{3}$
x_3	$\frac{2}{3}$	$\frac{1}{3}$
x_2	$\frac{5}{3}$	$\frac{1}{3}$

$B' = (3,2), N' = (1)$. Die Lösung ist optimal mit $x_1 = 0, x_2 = \frac{5}{3}, x_3 = \frac{2}{3}$, denn $r_1' = -\frac{2}{3} < 0$.

1.4.19 Tableau-Matrix:

$$T = (t_{ij}) = \begin{array}{|c|c|} \hline z_0 & r_N \\ \hline b & A_N \\ \hline \end{array}, \quad 0 \leq i \leq m, 0 \leq j \leq n-m.$$

1.4.20 Simplex-Algorithmus: (George Dantzig, 1947)

(i) Start: Sei B eine Basis mit $A_B = E_m$, die zugehörige Basislösung sei zulässig, d.h. es gelte $x_B = b \geq 0$. Berechne mit einer Nichtbasis N die Tableau-Matrix in *(1.4.19)*.

(ii) Gilt $t_{0s} > 0$ (d.h. $r_{s'} > 0$ für ein $s' \in N$), so gehe zu (iii). Andernfalls ist die Basislösung optimal, d.h. $r_N \leq 0$. Setze $x_{B(i)} := t_{i0}$ für $1 \leq i \leq m, x_{N(j)} := 0$ für $1 \leq j \leq n-m, z(x) = t_{00}$.

(iii) Bestimmung einer Austauschspalte: Wähle Index $s \in \{1, \ldots, m\}$ mit $t_{0s} > 0, (1 \leq s \leq n-m)$ z.B. gemäß Regel *(1.4.12)*.

(iv) Sind alle $t_{is} \leq 0$ für $i = 1, \ldots, m$ so gibt es keine endliche Lösung. STOP.
Gibt es eim $1 \leq i \leq m$ mit $t_{is} > 0$, so gehe zu (v).

(v) Bestimmung einer Austauschzeile: Wähle $p \in \{1, \ldots, m\}$ mit

$$\frac{t_{p0}}{t_{ps}} = \min \left\{ \frac{t_{io}}{t_{is}} \;\middle|\; t_{is} > 0, 1 \leq i \leq m \right\}$$

z.B. gemäß *(1.4.12)*.

(vi) Führe eine Pivotoperation (Austauschscritt) mit dem Pivotelement $t_{ps} > 0$ durch: vertausche den p-ten Index von B mit dem s-ten Index von N, vgl. *(1.4.16)*.

- Pivotelement: $t'_{ps} = \frac{1}{t_{ps}}$
- übrige Pivotzeile: $t'_{pj} = \frac{t_{pj}}{t_{ps}},\ j \neq s$
- übrige Pivotspalte: $t'_{is} = -\frac{t_{is}}{t_{ps}},\ i \neq s$
- sonstige Elemente: $t'_{ij} = t_{ij} - t_{is}\frac{t_{pj}}{t_{ps}},\ i \neq p, j \neq s$

Setze $t_{ij} = t'_{ij}$ für $0 \leq i \leq m, 0 \leq j \leq n-m$ und gehe zu (ii).

Anwendung auf LPs mit Ungleichungen:

$$\min\{cx \mid Ax \leq b, x \geq 0\} \;\Leftrightarrow\; \min\{cx + 0y \mid Ax + y, x \geq 0, y \geq 0\}$$

Zum Start wähle $B = (n+1, \ldots, n+m)$ und $N = (1, \ldots n)$. Als Basislösung ist dann durch $y = b, x = 0$ gegeben. Berechne $r_N = -c, z_0 = 0$. Simplex-Tableau:

$$\begin{array}{c|c|c|} & & x \\ \hline & z_0 = 0 & r_N = -c \\ \hline y & b & A \\ \hline \end{array}$$

Beispiel:
Vgl. Beispiel Damen- und Herrenschuhe: $\max 16x_1 + 32x_2$ also $\min -16x_1 - 32x_2$ unter

$$\begin{aligned} 20x_1 + 10x_2 + y_1 &= 8000 \\ 4x_1 + 5x_2 + y_2 &= 2000 \\ 6x_1 + 15x_2 + y_3 &= 4500 \\ x_1, x_2, x_3, y_1, y_2 &\geq 0 \end{aligned}$$

Das Simplexverfahren führt nun zu einem Übergang der Ecken $A \to B \to C \to D$ (vgl. damalige Grafik).
Tableau:

		x_1	x_2
$z(x)$	0	16	32
y_1	8000	$\boxed{20}$	10
y_2	2000	4	5
y_3	4500	6	15

Tausche y_1 und x_1: Ecke $A \to B$

		y_1	x_2
$z(x)$	-6400	$-\frac{4}{5}$	24
x_1	400	$\frac{1}{20}$	$\frac{1}{2}$
y_2	400	$-\frac{1}{5}$	$\boxed{3}$
y_3	2100	$-\frac{3}{10}$	12

Tausche y_2 und x_2: Ecke $B \to C$

		y_1	y_2
$z(x)$	-9600	$\frac{4}{5}$	-8
x_1	$\frac{1000}{3}$	$\frac{1}{12}$	$-\frac{1}{6}$
x_2	$\frac{400}{3}$	$-\frac{1}{15}$	$\frac{1}{3}$
y_3	500	$\boxed{\frac{1}{2}}$	-4

Tausche y_3 und y_1: Ecke $C \to D$

		y_3	y_2
$z(x)$	-10400	$-\frac{8}{5}$	$-\frac{8}{5}$
x_1	250	$-\frac{1}{6}$	$\frac{1}{2}$
x_2	250	$\frac{2}{15}$	$-\frac{1}{5}$
y_1	1000	2	-8

Das Tableau ist optimal, denn $r_N = (-\frac{8}{5}, -\frac{8}{5}) < 0$. Basislösung ist $x = (250, 200)^T, y_1 = 1000, y_2 = 0, y_3 = 0$. Der Gewinn beträgt 10400.

1.5 Die Zweiphasenmethode

Die Zweiphasenmethode dient der Bestimmung einer zulässigen Basislösung. Dazu sei das LP

$$\min\{cx \mid AX = b, x \geq 0\} \tag{1.5.1}$$

gegeben, sowie $b_i \geq 0$ für $i = 1, \ldots, m$ gegeben. (Falls $b_i < 0$ für ein i gilt, so multiplizieren wir die i-te Gleichung mit -1. Zum Start des Simplex-Verfahrens:

1.5.2 Problem:
Berechne eine zulässige Basislösung von $Ax = b, x \geq 0$, zu einer geeigneten Basis B.

Dazu formulieren wir das Hilfsproblem

$$\min \sum_{i=1}^{m} y_i \text{ unter } Ax + y = b,\ x, y \geq 0 \tag{1.5.3}$$

Dann gibt es zwei Möglichkeiten:

(i) Das Hilfsproblem *1.5.3* besitzt eine Lösung mit Zielfunktionswert 0. In diesem Fall ist die zugehörige Basislösung zulässig, d.h. *(1.5.2)*, gelöst.

(ii) Das Hilfsproblem *1.5.3* besitzt eine Lösung mit Zielfunktionswert > 0. Dann besitzt das System $Ax = b, x \geq 0$ keine zulässige Lösung.

Wir betrachten zunächst Fall (i): Zur Lösung von *1.5.3* mit dem Simplex-Verfahren beginnen wir mit der Wahl von $B = (n+1, \dots, n+m)$ und $N = (1, \dots, n)$. Dann erhalten wir $c_B = (1, \dots, 1) \in \mathbb{R}^m, c_N = 0, y = b, x = 0$ und

$$z_0 = c_B b = \sum_{i=1}^{m} b_i, \quad r_N = c_B \underbrace{\tilde{A}_b^{-1}}_{E_m} \underbrace{\tilde{A}_N}_{A} - c_N = c_B A \Rightarrow r_j = \sum_{i=1}^{m} a_{ij} \tag{1.5.4}$$

1.5.5 Starttableau:

		x
$z(x,y)$	z_0	$r_N = c_B A$
y	b	A

Zweiphasenmethode zur Lösung des LP *1.5.1*:

Phase 1:

Löse das Hilfsproblem *1.5.3* mit dem Starttableau *(1.5.5)*. Die Basisvariablen $y_1, \dots, y_m$ müssen im optimalen Tableau zu Nichtbasisvariablen werden, d.h. setze $y = 0$.

Phase 2:

Berechne das Starttableau für die Lösung des LP *1.5.1* durch

1. Streichung der unter y_i stehenden Spalten
2. Berechnung der Werte z_0 und r_N

Beispiel:

Wir rechnen unser running example: $\min 4x_1 + x_2 + x_3$ unter

$$\begin{pmatrix} 2 & 1 & 2 \\ 3 & 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \quad x_1, x_2, x_3 \geq 0$$

Phase 1: $B = (4,5), N = (1,2,3)$ und damit $z_0 = b_1 + b_2 = 7$ sowie $r_N = \left(\sum_{i=1}^{2} a_{ij}\right)_{j=1,2,3} = (5,4,3)$

$$\begin{array}{c|c|ccc} & & x_1 & x_2 & x_3 \\ \hline & 7 & 5 & 4 & 3 \\ \hline y_1 & 4 & 2 & 1 & 2 \\ y_2 & 3 & \boxed{3} & 3 & 1 \end{array} \xrightarrow{y_2 \leftrightarrow x_1} \begin{array}{c|c|ccc} & & y_2 & x_2 & x_3 \\ \hline & 2 & -\frac{5}{3} & -1 & \frac{4}{3} \\ \hline y_1 & 2 & -\frac{2}{3} & -1 & \boxed{\frac{4}{3}} \\ x_1 & 1 & \frac{1}{3} & 1 & \frac{1}{3} \end{array} \xrightarrow{y_1 \leftrightarrow x_3} \begin{array}{c|c|ccc} & & y_2 & x_2 & y_1 \\ \hline & 0 & -1 & 0 & -1 \\ \hline x_3 & \frac{3}{2} & -\frac{1}{2} & -\frac{3}{4} & \frac{4}{3} \\ x_1 & \frac{1}{2} & \frac{1}{2} & \frac{5}{4} & -\frac{1}{4} \end{array}$$

Es gilt $r_N \leq 0$. y_1, y_2 sind Nichtbasisvariablen. Die zulässige Basislösung ist $x_1 = \frac{1}{2}, x_2 = 0, \frac{3}{2}$. Aktuelle Basis ist $(3,1)$ und $N = (5,2,4)$.

Phase 2: Wir streichen zunächst die Spalten unter y_1 und y_2, sodass nur noch $N = (2)$. Nun berechnen wir die neuen Werte zu $B = (3, 1)$ und $N = (2)$.

$$z_0 = c_B b = (1, 4)\begin{pmatrix} \frac{3}{2} \\ \frac{1}{2} \end{pmatrix} = \frac{7}{2}$$

und

$$r_N = c_B A_N - c_N = (1, 4)\begin{pmatrix} -\frac{3}{4} \\ \frac{5}{4} \end{pmatrix} - 1 = \frac{13}{4} > 0$$

Unser Tableau was wir bekommen haben ist also noch nicht optimal.

		x_2
	$\frac{7}{2}$	$\frac{13}{4}$
x_3	$\frac{3}{2}$	$-\frac{3}{4}$
x_1	$\frac{1}{2}$	$\boxed{\frac{5}{4}}$

$\xrightarrow{x_1 \leftrightarrow x_2}$

		x_1
	$\frac{11}{5}$	-$\frac{13}{5}$
x_3	$\frac{9}{5}$	$\frac{3}{5}$
x_2	$\frac{2}{5}$	$\frac{4}{5}$

Dieses Tableau ist optimal mit der Basis $B = (3, 2)$, Nichtbasis $N = (1)$ und Lösungsvektor $x = \left(0, \frac{2}{5}, \frac{9}{5}\right)$.

Behandlung von Gleichungen und Ungleichungen:

Die Idee ist sehr einfach: Alles was mit Ungleichungen zu tun hat wird mit Schlupfvariablen vesehen, für die ursprünglichen Gleichungen muss unter Umständen mit der Zweiphasenmethode eine Basislösung gefunden werden.

Teil 1

Seien a_i die Zeilen von A mit $i = 1, \ldots, m$. Die Gleichungen des Problems seien gegeben durch $a_i x = b_i, i = 1, \ldots, k$ und die Ungleichungen durch $a_i x \leq b_i, i = k+1, \ldots, m$. Führe Schlupfvariablen $y_1, \ldots, y_k, y_{k+1}, \ldots, y_m$ ein. So erhalten wir:

$$a_i x + y_i = b_i, \;\; i = 1, \ldots, m$$

Löse das Hilfsproblem

$$\min \sum_{i=1}^{k} y_i \quad \text{unter} \quad Ax + y = b, \;\; x \geq 0, y \geq 0$$

Beispiel:

Optimales Gasgemisch, vgl. *1.1.2*:

$$\min z(x) = 10x_1 + 30x_2 + 20x_3 \quad \text{unter} \quad \begin{array}{rcl} x_1 + 3x_2 + 6x_3 & = & 3 \\ 8x_1 + x_2 + 2x_2 & \leq & 3 \\ x_1 + x_2 + x_3 & = & 1 \\ x_1, x_2, x_2 & \geq & 0 \end{array}$$

Phase 1:

Löse das Hilfsproblem mit Zielfunktion $z = y_1 + y_3$ (nur Gleichungen werden berücksichtigt!):

$$\mathit{1.5.4} \Rightarrow \begin{cases} z_0 = b_1 + b_3 = 4 \\ r_N = (a_{1j} + a_{3j})_{j=1,2,3} = (2, 4, 5) \end{cases}$$

Wir stellen unsere Tableaus auf:

		x_1	x_2	x_3
	4	2	4	7
y_1	3	1	3	$\boxed{6}$
y_2	3	8	1	2
y_3	1	1	1	1

$\xrightarrow{y_1 \leftrightarrow x_3}$

		x_1	x_2	y_1
	$\frac{1}{2}$	$\frac{5}{6}$	$\frac{1}{2}$	$*$
x_3	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{2}$	$*$
y_2	2	$\frac{23}{3}$	0	$*$
y_3	$\frac{1}{2}$	$\frac{5}{6}$	$\boxed{\frac{1}{2}}$	$*$

$\xrightarrow{y_3 \leftrightarrow x_2}$

		x_1	y_3	y_1
	0	0	$*$	$*$
x_3	0	$-\frac{2}{3}$	$*$	$*$
y_2	2	$\frac{23}{3}$	$*$	$*$
x_2	1	$\frac{5}{3}$	$*$	$*$

Dieses Tableau ist optimal für Phase 1 (das y_2 stört hier nicht, da es aus einer Ungleichung stammt!).

Phase 2

Wir haben $B = (3,4,2)$ und $N = (1)$ und berechnen

$$z_0 = c_B b = (20,0,30)\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = 30, \quad r_N = c_B A_N - c_N = (20,0,30)\begin{pmatrix} -\frac{2}{3} \\ \frac{23}{3} \\ \frac{5}{3} \end{pmatrix} - 10 = \frac{80}{3} > 0$$

Also ist die x_B Basislösung nicht optimal.

$$\begin{array}{c|c|c} & & x_1 \\ \hline & 30 & \frac{80}{3} \\ \hline x_3 & 0 & -\frac{2}{3} \\ y_2 & 2 & \boxed{\frac{23}{3}} \\ x_2 & 1 & \frac{5}{3} \end{array} \xrightarrow{x_1 \leftrightarrow y_2} \begin{array}{c|c|c} & & y_2 \\ \hline & \frac{530}{23} & -\frac{80}{3} \\ \hline x_3 & \frac{12}{69} & * \\ x_1 & \frac{6}{23} & \frac{3}{23} \\ x_2 & \frac{39}{69} & * \end{array}$$

Die Optimale Lösung ist also

$$x = \begin{pmatrix} \frac{6}{23} \\ \frac{39}{39} \\ \frac{12}{69} \end{pmatrix}$$

Gleichungen und Ungleichungen mit $\geq$

Etwa

$$\begin{array}{rcll} a_i x & \geq & b_i, & i = 1, \dots, k \\ y_i x & \leq & b_i & i = k+1, \dots, m \end{array}$$

ist äquivalent zu:

$$\begin{array}{lcll} a_i x - x_{n+i} & = & b_i, & i = 1, \dots, k \\ a_i x & \leq & b_i, & i = k+1, \dots, m \\ x_i & \geq & 0, & i = 1, \dots, n, n+1, \dots, n+k \end{array} \tag{1.5.6}$$

Das Hilfsproblem *1.5.3* für *1.5.6* lautet:

$$\min \sum_{i=1}^{k} y_i \quad \text{unter} \quad \begin{array}{lcll} a_i x - x_{n+i} + y_i & = & b_i, & i = 1, \dots, k \\ a_i x + y_i & = & b_i, & i = k+1, \dots, m \\ x_1, \dots, x_{n+k}, y_1, \dots, y_m & \geq & 0 & \end{array}$$

Bestimmung eines Starttableaus:

$$N = (1, \dots, n, n+1, \dots, n+k) \cong x_1, \dots, x_n, x_{n+1}, \dots, x_{n+k}$$
$$B = (n+k+1, \dots, n+k+m) \cong y_1, \dots, y_m$$

$$\begin{array}{c|c|c|c} & & x_1, \dots, x_n & x_{n+1}, \dots, x_{n+k} \\ \hline & z_0 = \sum_{i=1}^k b_i & \sum_{i=1}^k a_{ij} & -1 \cdots -1 \\ \hline y_1 & b_1 & a_1 & \\ \vdots & \vdots & \vdots & -E_k \\ y_k & b_k & a_k & \\ \cline{3-4} y_{k+1} & b_{k+1} & a_{k+1} & \\ \vdots & \vdots & \vdots & 0 \\ y_m & b_m & a_m & \end{array}$$

Begründung:

$$\begin{aligned}
c_b &= \big(\underbrace{1,\ldots,1}_{k-\text{mal}},0,\ldots,0\big)\\
z_0 &= c_b b = \sum_{i=1}^{k} b_i\\
r_N &= \big(r_j\big)_{j\in N}\\
r_j &= c_B a^j - c_j = \begin{cases} \sum_{i=1}^{k} a_{ij} & \text{für } i = 1,\ldots,n\\ -1 & \text{für } j = n+1,\ldots,n+k \end{cases}
\end{aligned}$$

Beispiel:
Betrachte das Problem

$$\min z(x) = x_1 + 4x_2 + 5x_3 \quad \text{unter} \quad \begin{array}{lcl} 3x_1 + 6x_2 + x_3 & \geq & 40\\ 2x_1 + x_2 + 3x_3 & \geq & 30\\ x_1 + 4x_3 & \geq & 20\\ x_1, x_2, x_3 & \geq & 0 \end{array}$$

Start: $N = (1,\ldots,6), B = (7,8,9)$.

$$z_0 = c_B b = 90, \quad r_N = (6,7,8,-1,-1,-1)$$

Starttableau

		x_1	x_2	x_3	x_4	x_5	x_6
	90	6	7	8	−1	−1	−1
y_1	40	3	6	1	−1	0	0
y_2	30	2	3	3	0	−1	0
y_3	20	1	0	4	0	0	−1

Nach 3 Simplex-Schritten sind y_1, y_2, y_3 zu Nichtbasisvariablen geworden: Streiche die Spalten unter y_i und erhalte:

		x_2	x_4	x_6
	0	0	0	0
x_5	$\frac{10}{11}$	$\frac{13}{11}$	$-\frac{5}{11}$	$-\frac{7}{11}$
x_1	$\frac{140}{11}$	$\frac{24}{11}$	$-\frac{4}{11}$	$\frac{1}{11}$
x_3	$\frac{29}{11}$	$-\frac{6}{11}$	$\frac{1}{11}$	$-\frac{3}{11}$

Phase 2: Berechne neue Werte für $B = (5,3,1), N = (2,4,6)$

$$z_0 = c_B b = \frac{380}{11}, \qquad r_N = c_B A_N - c_N = \left(-\frac{26}{11}, -\frac{3}{11}, -\frac{13}{11}\right) < 0$$

Optimale Basislösung:

$$x_1 = \frac{140}{11}, x_2 = 0, x_3 = \frac{29}{11}, x_4 = 0, x_5 = \frac{10}{11}, x_6 = 0$$

Bemerkung:
Ungleichungen $a_i x \geq b_i, i = 1,\ldots,k$ können einfacher behandelt werden mit dem dualen Simplex-Verfahren.

1.6 Grundidee des revidierten Simplexverfahrens:

Löse das LP $\min\{cx \mid Ax = b, x \geq 0\}$ wobei A eine $m \times n$-Matrix und A_N eine $m \times (n-m)$ Matrix. Im Fall $m << n$ ist es jedoch numerisch günstiger mit der $m \times n$ Matrix A_B zu arbeiten.

1.6.1 Grundidee des revidierten Simplexverfahrens
Die Spalten von A seien $a^1, \ldots, a^n \in \mathbb{R}^m$. Wähle eine Basis B

(i) Berechne die Baislösung $x_B = \tilde{b} = A_B^{-1} b$ als Lösung des LGS $A_B x_B = b$. Die Basislösung x_B sei zulässig, d.h. $x_B \geq 0$

(ii) Löse das LGS $\lambda A_B = c_B$ mit einem Zeilenvektor $\lambda \in \mathbb{R}^m$, d.h. $\lambda = c_B A_B^{-1}$. Berechne $r_N = \lambda A_N - c_N = c_B A_B^{-1} A_N - c_N$. Gilt $r_N \leq 0$: STOP. Dann ist $x_B = \tilde{b} = A_B^{-1}, x_N = 0$ optimal. Sonst gehe zu (iii).

(iii) Wähle $s \in N$ mit $r_s > 0$, etwa gemäß *(1.4.12)*.

(iv) Löse das LGS $A_B \tilde{a}^s = \tilde{a}^s$.

(v) Ist $\tilde{a}^s \leq 0$, dann gibt es keine endliche Lösung: STOP.

(vi) Bestimme Index $p \in B$ mit

$$\frac{\tilde{b}_p}{\tilde{a}_{ps}} = \min\left\{ \frac{\tilde{b}_p}{\tilde{a}_{ps}} \;\middle|\; \tilde{a}_{is} > 0 \right\}$$

Wähle etwa $p \in B$ gemäß *(1.4.12)*.

(vii) Bestimme neue Basis $\overline{B}$ durch Vertauschung des Index $p \in B$ mit dem Index $s \in N$. Ebenso für neue Nichtbasis $\overline{N}$.

Insgesamt: Löse die folgenden drei linearen Gleichungssystemen:

$$A_B x_B = b, \quad \lambda A_B = c_B, \quad A_B \tilde{a}^s = a^s \tag{1.6.2}$$

Idee:
Berechnung von A_B^{-1} durch die Betrachtung benachbarter Basen. Übergang von B zu $\overline{B}$: Berechnung $A_{\overline{B}}^{-1}$ durch den Übergang

$$\boxed{\tilde{a}^s \;\Big|\; A_B^{-1}} \overset{\text{Elimination}}{\Longrightarrow} \boxed{e_p \;\Big|\; A_{\overline{B}}^{-1}}, \quad \text{Basis } \overline{B}, e_p = \begin{pmatrix} \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}$$

Beachte:
$\lambda = c_B A_B^{-1}$ ist der sogenannte Simplex-Multiplikator, der eine Lösung des dualen Optimierungsproblemes ist, vgl *(1.7.10)*.

1.7 Dualität

Jedem Primalproblem (P) wird ein Dualproblem (P^*) zugeordnet.

1.7.1 Duale lineare Programme (symmetrische Form):
Primalproblem (P): $\min\{cx \mid Ax \geq b, x \geq 0\}$
Dualproblem (P^*): $\max\{\lambda b \mid \lambda A \leq c, \lambda \geq 0\}$.

Hierbei ist A eine $m \times n$ Matrix, $b \in \mathbb{R}^m, \lambda \in \mathbb{R}^m$ Zeilenvektor. Die zulässigen Mengen sind dabei

$$K := \{x \in \mathbb{R}^n \mid Ax \geq b, x \geq 0\}, \quad K^* : \{\lambda \in \mathbb{R}^m \mid \lambda A \leq c, \lambda \geq 0\}$$

Beispiel:
Sei das Primalproblem gegeben durch: $\min 3x_1 + 4x_2 + 5x_3$ unter

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &\geq 5 \\ 2x_1 + 2x_2 + x_3 &\geq 6 \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

Dann ist das zugehörige Dualproblem gegeben durch $\max 5\lambda_1 + 6\lambda_2$ unter

$$\begin{aligned} \lambda_1 + 2\lambda_2 &\leq 3 \\ 2\lambda_1 + 2\lambda_2 &\leq 4 \\ 3\lambda_1 + \lambda_2 &\leq 5 \\ \lambda_1, \lambda_2 &\geq 0 \end{aligned}$$

Bemerkung:
Beachte: (P^*) ist einfacher zu lösen als (P), auf das die Zweiphasenmethode angewandt werden muss.

Das Dualproblem des Dualproblems (P^*) ist das Primalproblem, d.h. es gilt $(P^{**}) = (P)$.

Berechnung des Dualproblems zum Standard-LP $\min\{cx \mid Ax = b, x \geq 0\}$. $Ax = b$ ist äquivalent zu $Ax \geq b \wedge -Ax \geq -b$.

$$Ax = b \Leftrightarrow \begin{pmatrix} A \\ -A \end{pmatrix} x \geq \begin{pmatrix} b \\ -b \end{pmatrix}$$

Das Dualproblem (P^*) zu

$$\min \left\{ cx \;\middle|\; \begin{pmatrix} A \\ -A \end{pmatrix} x \leq \begin{pmatrix} b \\ -b \end{pmatrix}, x \geq 0 \right\}$$

lautet also mit einer Variablen $\omega \in \mathbb{R}^{2m}$

$$\max \left\{ \omega \begin{pmatrix} b \\ -b \end{pmatrix} \;\middle|\; \omega \begin{pmatrix} A \\ -A \end{pmatrix} \leq c, \omega \geq 0 \right\}$$

Setze $\omega = (u, v)$ mit $u, v \in \mathbb{R}^m$ und erhalte

$$\max \left\{ ub - vb \;\middle|\; uA - vA \leq c, u, v \geq 0 \right\}$$

Setze $\lambda := u - v \in \mathbb{R}^m$ und erhalte

$$\max \left\{ \lambda b \;\middle|\; \lambda A \leq c \right\} \tag{P^*}$$

Aber, $\lambda = u - v$ ist nun *nicht* mehr vorzeichenbeschränkt.

1.7.2 Duale lineare Programme (unsymmetrische Form):
Primalproblem:

$$\min \left\{ cx \mid Ax = b, x \geq 0 \right\} \tag{P}$$

Dualproblem:

$$\max \left\{ \lambda b \mid \lambda A \leq c, \lambda \text{ frei } \right\} \tag{P^*}$$

Merkregel:

Primalproblem	Dualproblem
$a_i x \geq b_i, x \geq 0$	$\lambda a^i \leq c_i, \lambda_i \geq 0$
$a_i x = b_i, x \geq 0$	$\lambda a^i \leq c_i, \lambda_i$ frei
x_i frei	$\lambda a^i = c_i$

Beispiel:
Betrachten wir das Primalproblem

$$\min(3,1,-2)x \text{ unter } \begin{array}{rcl} 4x_1 + 5x_2 - 2x_3 & = & 1 \\ -x_1 + 3x_2 + x_3 & \geq & 1 \\ x_2 \text{ frei}, x_1, x_3 & \geq & 0 \end{array}$$

so ist das Dualproblem gegeben durch

$$\min \lambda_1 + \lambda_2 \text{ unter } \begin{array}{rcl} 4\lambda_1 - \lambda_2 & \leq & 3 \\ 5\lambda_1 + 3\lambda_2 & = & 1 \ (x_2 \text{ frei}) \\ -2\lambda_1 + \lambda_3 & \leq & -2 \\ \lambda_1 \text{ frei}, \lambda 2 & \geq & 0 \end{array}$$

1.7.1 Der Dualitätssatz

Die zulässige Mengen für die linearen Optimierungsprobleme in *(1.7.1)* bzw. *(1.7.2)* sind gegeben durch:

$$K = \left\{ x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \right\}$$

für (P) sowie

$$K^* := \left\{ \lambda \in \mathbb{R}^m \mid \lambda A \leq c, \lambda \text{ frei } \right\}$$

für (P^*).

1.7.3 Satz: *Schwacher Dualitätssatz*
Für alle zulässigen Punkte $x \in K$ und $\lambda \in K^*$ gilt:

$$\lambda b \leq cx$$

Beweis:
Ja also seien $x \in K$ und $\lambda \in K^*$, also $Ax = b, x \geq 0$ und $\lambda A \leq c$ und damit:

$$\lambda b \ = \ \lambda(Ax) \ = \ (\lambda A)x \ \leq \ cx \text{ wegen } x \geq 0$$

□

1.7.4 Korollar:
Sind $x \in K$ und $\lambda \in K^*$ zulässige Punkte mit $\lambda b = cx$, dann sind x und λ optimale Lösungen von (P) und (P^*).

1.7.5 Satz: *Dualitätssatz der linearen Optimierung*
Besitzt eines der Probleme (P) oder (P^*) eine endliche Lösung, so besitzt auch das jeweils andere Problem eine endliche Lösung. Die Optimalwerte beider Probleme stimmen überein.

Beweis:
Ohne Einschränkung habe das primale Problem (P) eine endliche Lösung $\overline{x} \in K$. Sei $z_0 = c\overline{x}$. Definiere den *konvexen und abgeschlossenen Kegel*

$$C := \left\{ \begin{pmatrix} r \\ w \end{pmatrix} = t \cdot \begin{pmatrix} cx - c\overline{x} \\ b - Ax \end{pmatrix} \;\middle|\; x \geq 0, t > 0 \right\} \subset \mathbb{R} \times \mathbb{R}^m$$

Es gilt $\binom{0}{0} \in C$ (wähle $x = \overline{x}$).

Es gilt $\binom{-1}{0} \notin C$, denn *angenommen* es gelte $\binom{-1}{0} \in C$, dann gibt es ein $t_0 > 0$ und ein $x_0 \geq 0$ mit

$$\begin{pmatrix} -1 \\ 0 \end{pmatrix} = t_0 \cdot \begin{pmatrix} cx_0 - c\overline{x} \\ b - Ax_0 \end{pmatrix}$$

Dann muss wegen $t_0 > 0$ bereits $b - Ax_0 = 0$ gelten. Also ist $x_0 \in K$ und daher $cx_0 \geq c\overline{x}$. Dies ist jedoch ein Widerspruch zur Annahme $-1 = t_0(cx_0 - c\overline{x})$.

Wir führen eine Skalierung von C durch:

$$t(cx - c\overline{x}) = c\underbrace{(tx)}_{=:y} - tz_0 = cy - tz_0 \text{ mit } y \geq 0$$
$$t(b - Ax) = tb - A\underbrace{(tx)}_{=:y} = tb - Ay$$

und erhalten

$$C := \left\{ \begin{pmatrix} cy - tz_0 \\ tb - Ay \end{pmatrix} \;\middle|\; x \geq 0, t > 0 \right\}$$

Wegen $\binom{-1}{0} \notin C$ gibt es nach dem Trennungssatz für Kegel *(3.2.5)* einen Zeilenvektor $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m, (\lambda_0, \lambda) \neq (0, 0_m)$, mit

$$(\lambda_0, \lambda)\begin{pmatrix} -1 \\ 0_m \end{pmatrix} = -\lambda_0 < 0 \leq (\lambda_0, \lambda)\begin{pmatrix} cy - tz_0 \\ tb - Ay \end{pmatrix}$$

für alle $y \geq 0, t > 0$. Es folgt $\lambda_0 > 0$. Ohne Einschränkung sei $\lambda_0 = 1$ (zu erreichen durch Multiplikation mit $\frac{1}{\lambda_0} > 0$) und damit

$$\forall y \geq 0, t > 0 : 0 \leq cy - tz_0 + \lambda(tb - Ay) = (c - \lambda A)y - tz_0 + t\lambda b$$

Für $t \searrow 0$ gilt dann

$$\forall y \geq 0 : 0 \leq (c - \lambda A)y \Rightarrow 0 \leq c - \lambda A \Rightarrow \lambda A \leq c$$

Also ist $\lambda \in K^*$ und λ ist frei. Betrachten wir nun die Ungleichung für $t = 1, y = 0$ so erhalten wir

$$0 \leq -z_0 + \lambda b \;\Leftrightarrow\; z_0 = c\overline{x} \leq \lambda b$$

Nach dem schwachen Dualitätssatz *(1.7.3)* muss dann schon $z_0 = c\overline{x} = \lambda b$ gelten und nach *(1.7.4)* sind dann auch $\overline{x} \in K$ und $\lambda \in K^*$ optimal. □

Bemerkung:

1. Der Dualitätssatz *(1.7.5)* benutzt *nicht* die Rangbedingung rg $A = m$
2. Der Beweis zum Dualitätssatz *(1.7.5)* lässt sich auch mit dem Lemma von Farkas *(3.2.6)* durchführen.
3. Ein einfacher Beweis von *(1.7.5)* lässt sich unter stärkeren Voraussetzungen mit dem Simplex-Multiplikator durchführen $\lambda = c_B A_B^{-1}$ vgl. *(1.7.10)*.

1.7.6 Satz: *Existenzsatz*

(i) Besitzen (P) und (P*) zulässige Punkte, dann besitzen sie auch endliche optimale Lösungen.

(ii) Besitzt nur eines der Problem (P) und (P*) zulässige Punkte, so besitzen beide Probleme keine endlichen Lösungen.

(iii) Besitzt ein Problem zulässige Punkte, aber keine endliche Lösung, so hat das jeweils andere Problem keine zulässigen Punkte.

1.7.7 Satz: *Komplementarität der unsymmetrischen Form*
Seien x, λ zulässig. Dann ist äquivalent:

(i) x, λ sind optimal, d.h. $\lambda b = cx$

(ii) $(\lambda A - c)x = 0$

(iii) Es gilt $x_i > 0 \Rightarrow \lambda a^i = c_i$ sowie $\lambda a^i < c_i \Rightarrow x_i = 0$

1.7.8 Satz: *Komplementarität der symmetrischen Form*
Seien x, λ zulässig. Dann ist äquivalent:

(i) x und λ sind optimal, d.h. $cx = \lambda b$

(ii) $(\lambda A - c)x = 0$ und $\lambda(Ax - b) = 0$

(iii) Es gilt

$$\begin{aligned}
x_i > 0 &\quad\Rightarrow\quad \lambda a^i = c_i \\
\lambda a^i < c_i &\quad\Rightarrow\quad x_i = 0 \\
\lambda_j > 0 &\quad\Rightarrow\quad a_j x = b_j \\
a_j x > b_j &\quad\Rightarrow\quad \lambda_j = 0
\end{aligned}$$

für $i = 1, \ldots, n$ und $j = 1, \ldots, m$.

Beispiel:
Wir betrachten das Primalproblem $\min 3x_1 + 4x_2 + 5x_3$ unter

$$\begin{aligned}
x_1 + 2x_2 + 3x_3 &\geq 5 \\
2x_1 + 2x_2 + x_3 &\geq 6 \\
x_1, x_2, x_3 &\geq 0
\end{aligned}$$

Dann ist das Dualproblem gegeben durch $\max 5\lambda_1 + 6\lambda_2$ unter

$$\begin{aligned}
\lambda_1 + 2\lambda_2 &\leq 3 \\
2\lambda_1 + 2\lambda_2 &\leq 4 \\
3\lambda_1 + \lambda_2 &\leq 5 \\
\lambda_1, \lambda_2, &\geq 0
\end{aligned}$$

Die optimale Lösung ist $\lambda_1 = \lambda_2 = 1$. Mit dem Komplementaritätssatz *(1.7.8)* gilt nun $x_3 = 0$. $x_1 = 1, x_2 = 2$ können wir nun durch einfaches Einsetzen ausrechnen.

1.7.2 Das duale Simplexverfahren

$$\min \left\{ cx \mid Ax = b, x \geq 0 \right\} \tag{P}$$

$$\max \left\{ \lambda b \mid \lambda A \leq c, \lambda \text{ frei} \right\} \tag{P^*}$$

Vorausgesetzt sei $\mathrm{rg}(A) = m$. Wir erinnern uns, dass die Basislösung zu einer Basis B gegeben ist durch

$$x_B = A_B^{-1}b, \quad x_N = 0$$

1.7.9 Definition: ()
Der Zeilenvektor

$$\lambda := c_B A_B^{-1} \in \mathbb{R}^m$$

heißt *Simplex-Multiplikator* (vgl. rev SV §6)

Bemerkung:
Reduzierte Kosten:

$$r_N = c_B A_B^{-1} A_N - c_N = \lambda A_N - c_N$$

1.7.10 Satz: *Dualitätssatz*
Sei $\mathrm{rg}(A) = m$. Dann gilt

(i) $\lambda = c_B A_B^{-1}$ ist zulässig für (P^*) genau dann wenn $r_N \leq 0$.

(ii) Die Basislösung x sei optimal und nicht entartet, d.h. $x_B = A_B^{-1}b > 0, x_N = 0$. Dann ist $\lambda = c_B A_B^{-1}$ und es gilt

$$z_0 = c_B x_B = c_B A_B^{-1} b = \lambda b$$

Beweis:

1. Sei $\lambda = c_B A_B^{-1}$. Dann

$$\begin{aligned} \lambda A \leq c \quad &\Leftrightarrow \quad (\lambda A_B, \lambda A_N) \leq (c_B, c_N) \\ &\overset{\lambda A_B = c_B}{\Leftrightarrow} \quad \lambda A_N \leq c_N \\ &\Leftrightarrow \quad r_N = \lambda A_N - c_N \leq 0 \end{aligned}$$

2. Sei x optimal und nicht entartet, dann folgt mit *(1.4.11)*, dass $r_N \leq 0$. Mit Teil (i) folgt, dass auch $\lambda = c_B A_B^{-1}$ zulässig ist, d.h. $\lambda A \leq c$.

$$\lambda b = c_B A_B^{-1} b = c_B x_B \overset{x_N = 0}{=} cx = z_0$$

also ist λ optimal.

□

Idee des dualen Simplexverfahrens:
Es liege eine Basis B und eine Basislösung x vor, die

(a) nicht (primal) zulässig ist, d.h. es gilt nicht $x_B = A_B^{-1} b \geq 0$

(b) dual zulässig ist, d.h. es gilt $r_N \leq 0$.

Man geht vom „alten“ Simplextableau *(1.4.18)* aus und löst (P) vom dualen Standpunkt:

1. erhalte die Optimalitätsbedingung $r_N \leq 0$
2. erreiche die Zulässigkeit von x, d.h. $x_B = A_B^{-1} b \geq 0$
3. Lasse die Zielfunktion von (P^*) wachsen.

1.7.11 Starttableau:

$$\left.\begin{array}{r|c|c|}
 & & x_j, j \in N \\ \hline
 & z_0 & r_N \\ \hline
x_i, i \in B & b & A_N \\ \hline
\end{array}\right\} = T = \big(t_{ij}\big)_{\substack{0\le i\le m\\ 0\le j\le n-m}}$$

mit $r_N = c_B A_N - c_N$. Dieses Tableau ist

(i) primal zulässig, wenn $b \geq 0$

(ii) dual zulässig, wenn $r_N \leq 0$

(iii) optimal, falls $b \geq 0$ und $r_N \leq 0$

1.7.12 Das duale Simplexverfahren zur Lösung von (P):

1. Start: (P) sei *dual zulässig* zur Basis B und Nichtbasis N, d.h. $r_N \leq 0$. Berechne die Tableau-Matrix T aus *(1.7.11)*.
2. Ist $t_{i0} \geq 0$ für $i = 1, \ldots, m$, so ist die Basislösung optimal. Setze $x_{B(i)} = t_{i0}, x_{N(j)} = 0$ und $z_0 = t_{00}$. Sonst gehe zu 3.
3. Bestimme die Austauschzeile: Wähle Index p mit $t_{p0} < 0$, z.B.
$$t_{p0} = \min\big\{t_{i0} \mid t_{i0} < 0\big\}$$
4. Gilt $t_{pj} \geq 0$ für $1 \leq j \leq n - m$, so besitzt (P^*) keine endliche Lösung. STOP
5. Bestimme die Austauschspalte: Wähle Index s mit
$$\frac{t_{0s}}{t_{ps}} = \min\left\{\frac{t_{0j}}{t_{pj}} \;\middle|\; t_{pj} < 0, j = 1, \ldots, n - m\right\}$$
6. Vertausche das s-te Element von N mit dem p-ten Element von B. Führe an der Tableau-Matrix $T = \big(t_{ij}\big)$ eine Pivotoperation mit dem Pivotelement t_{ps} durch wie in *(1.4.20)* (*Amerkung:* Beachte Vorzeichen!). Gehe zu 2.

Anwendung auf LPs in symmetrischer Form:
Betrachten wir das LP $\min\big\{cx \mid Ax \geq b, x \geq 0\big\}$, dann gilt

$$Ax \geq b \Leftrightarrow -Ax \leq -b$$

Führe Schlupfvariablen ein und betrachte $-Ax + y = -b, x \geq 0, y \geq 0$.

Vorausgesetzt, dass $c \geq 0$, wähle Basis $B = (n + 1, \ldots, n + m)$ und Nichtbasis $N = (1, \ldots, n)$.

$$r_N = \tilde{c}_B A_N - \tilde{c}_N = 0 - c = -c, \quad z_0 = 0$$

Lösung von Beispiel 1:
Starttableau:

$$\begin{array}{r|r|rrr}
 & & x_1 & x_2 & x_3 \\ \hline
 & 0 & -3 & -4 & -5 \\ \hline
y_1 & -5 & -1 & -2 & -3 \\
y_2 & -6 & \boxed{-2} & -2 & -1
\end{array}
\xrightarrow{y_2 \leftrightarrow x_1}
\begin{array}{r|r|rrr}
 & & y_2 & x_2 & x_3 \\ \hline
 & 9 & -\frac{3}{2} & -1 & -\frac{7}{2} \\ \hline
y_1 & -2 & -\frac{1}{2} & \boxed{-1} & -\frac{5}{2} \\
x_1 & 3 & -\frac{1}{2} & 1 & \frac{1}{2}
\end{array}
\xrightarrow{y_1 \leftrightarrow x_2}
\begin{array}{r|r|rrr}
 & & y_2 & y_1 & x_3 \\ \hline
 & 11 & -1 & -1 & -1 \\ \hline
x_2 & 2 & \frac{1}{2} & -1 & \frac{5}{2} \\
x_1 & 1 & -1 & 1 & -2
\end{array}$$

Das letzte Tableau ist optimal. Die optimale Lösung ist $x = (1, 2, 0)^T$ mit $cx = 11$.

Nachtrag zum dualen Simplexverfahren:

$$\min cx \text{ unter } \begin{array}{c} a_i x \geq b_i,\ i = 1, \ldots, k \\ a_i x \leq b_i,\ i = k+1, \ldots, m \end{array}$$

führt auf das System von Ungleichungen:

$$\begin{array}{rcl} -a_i x & \leq & b_i,\ i = 1, \ldots, k \\ a_i x & \leq & b_i,\ i = k+1, \ldots, m \end{array}$$

Starttableau:

$z_0 = 0$	$r_n = -c$
$-b_1$	$-a_1$
$\vdots$	$\vdots$
$-b_k$	$-a_k$
b_{k+1}	a_{k+1}
$\vdots$	$\vdots$
b_m	a_m

Beispiel:

$$\min \left\{ x_1 + x_2 \mid x_1 - x_2 \geq 2, x_1 + 2x_2 \leq 4, x_1, x_2 \geq 0 \right\}$$

Lösung durch Raten: $x_1 = 2, x_2 = 0$

Ausgangstableau:

		x_1	x_2
	0	-1	-1
y_1	-2	-1	1
y_2	4	1	2

1.7.3 Sensitivität und Schattenpreise

Das LP $\min\{cx \mid Ax = b, x \geq 0\}$ habe die optimale Basis B und die Lösung sei nicht entartet, also $x_B = A_B^{-1} b > 0$. Die optimale duale Lösung ist dann $\lambda = c_B A_B^{-1}$. Doch was ist die Bedeutung von λ?

Idee:
Bei kleinen Störungen $b \to \Delta b$ bleibt die Basis B optimal für das gestörte Optimierungsproblem $\min\{cx \mid Ax = b + \Delta b, x \geq 0\}$.

Begründung:
Es gilt $\tilde{x}_B = A_B^{-1}(b+\Delta b), \tilde{b}_N = 0$ und wegen $x_B > 0$ kann $\varepsilon > 0$ gewählt werden mit $\tilde{x}_B = A_B^{-1}(b+\delta b) \geq 0$ für $||\Delta b||_\infty \leq \varepsilon$. Da $r_N = c_B A_B^{-1} A_N - c_N < 0$ nicht von b abhängt, sind B und N optimal für das gestörte Optimierungsproblem.

Änderung der Zielfunktion:

$$\begin{array}{rcl} \Delta z & = & z(\tilde{x}) - z(x) = c_B \tilde{x}_B - c_B x_B \\ & = & c_B A_B^{-1}(b + \Delta b) - c_B A_B^{-1} b \\ & = & \underbrace{c_B A_B^{-1}}_{=\lambda} \Delta b = \lambda \Delta b \end{array}$$

$\lambda = c_B A_B^{-1}$ war dabei die duale Lösung (Simplex-Multiplikator)

1.7.13 Korollar: *Schattenpreisformel*
Sei $w(b) = \min\{cx \mid Ax = b, x \geq 0\}$ der optimale Wert bezüglich $b \in \mathbb{R}^m$. Dann ist w differenzierbar und es gilt

$$\frac{\partial w}{\partial b_j} = \lambda_j, j = 1, \ldots, m$$

Ausblick: Parametrische Lineare Optimierung:
Abhängigkeit der optimalen Lösung $x \in \mathbb{R}^n$ und des optimalen Wertes w in Abhängigkeit von den Daten $b \in \mathbb{R}^m, c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}$. Wie der Zusammenhang mit b in etwa ausschaut, haben wir bereits gesehen. Schauen wir uns den Zusammenhang mit c an:

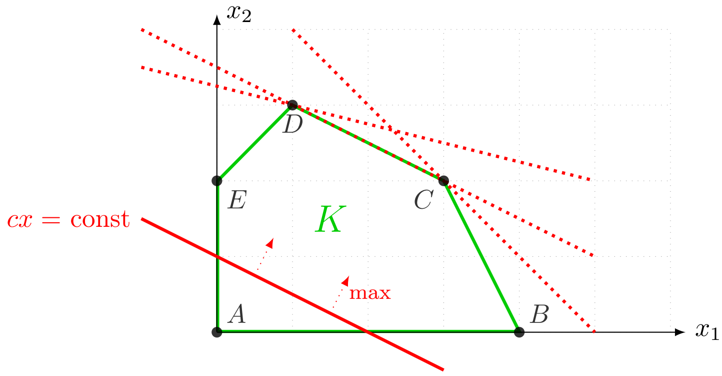


Abbildung 1.5: Beispiel: Änderung von c

Genaueres heben wir uns jedoch für die nichtlineare Optimierung auf.

1.7.4 Die Dreiphasenmethode

Nebenbedingungen:

$$\begin{array}{rcl} a_i x & = & b_i,\ i = 1, \ldots, k \\ a_i x & \geq & b_i,\ i = k_0 + 1, \ldots, k_1 \\ a_i x & \leq & b_i,\ i = k_1 + 1, \ldots, m \\ x & \geq & 0 \end{array}$$

Wir multiplizieren $a_i x \geq b_i$ für $i = k_0 + 1, \ldots, k_1$ mit -1 und heralten die Normalform des linearen Optimierungsproblemes:

$$\min cx \text{ unter } \begin{array}{rcl} a_i x & = & b_i,\ i = 1, \ldots, k \\ a_i x & \leq & b_i,\ i = k + 1, \ldots, k_1 \\ x & \geq & 0 \end{array}$$

Dabei gilt nicht notwendigerweise $b_i \geq 0$. Wir füllen das System mit Schlupfvariablen auf:

$$\min cx \text{ unter } \begin{array}{rcl} a_i x + y_i & = & b_i,\ i = 1, \ldots, k \\ a_i x + y_i & = & b_i,\ i = k + 1, \ldots, k_1 \\ x, y & \geq & 0 \end{array}$$

Ausgangstableau:

		x
	$z_0 = 0$	$r_N = c_B A_N - c_N$
y	b	A

Die Dreiphasenmethode besteht nun aus drei Schritten (Phasen):

Phase 1:

Mache die zu den Gleichungen gehörenden Basisvariablen $y_1, \dots y_k$ zu Nichtbasisvariablen (Es wird keine Hilfsfunktion $\sum_{i=1}^{k} y_i$ wie in §5 betrachtet).

Einzelner Schritt: Wähle Pivotspalte $s \in \{1, \dots, n\}$ und Pivotzeile $p \in \{1, \dots, m\}$ mit $a_{ps} \neq 0$, wobei y_p Basisvariable ist. Führe Austauschschritt durch und streiche die Spalte s unter y_p. Dabei wird keine spezielle Regel zur Wahl von s und p gefordert.

Beachte: Das Tableau am Ende der Phase 1 ist im Allgemeinen weder primal noch zulässig.

Phase 2:

Führe solange *duale Simplexschritte* durch bis eine primal zulässige Lösung $x_B \geq 0$ erreicht ist.

Einzelner Schritt:

1. Wähle Pivotzeile p mit $b_p < 0$, etwa $b_p = \min b_i$.
2. Wähle Pivotspalte s mit $\frac{r_s}{a_{ps}} = \min \left\{ \frac{r_j}{a_{pj}} \;\middle|\; a_{pj} < 0 \right\}$. Beachte: Die Quotienten können auch *negativ* sein. Falls alle $a_{pj} \geq 0$ sind, dann gibt es keine endliche Lösung.

Phase 3:

Führe das primale Simplexverfahren *(1.4.20)* mit der in Phase 2 berechneten primal zulässigen Lösung durch.

Beispiel: Kostenminimale Gasmischung:

$$\min z(x) = 10x_1 + 30x_2 + 20x_3 \quad \text{unter} \quad \begin{array}{rcl} x_1 + 3x_2 + 6x_3 & \geq & 3 \\ 8x_1 + x_2 + 2x_2 & \leq & 3 \\ x_1 + x_2 + x_3 & = & 1 \\ x_1, x_2, x_2 & \geq & 0 \end{array}$$

Phase 1: Ausgangstableau:

$$\begin{array}{c|c|ccc} & & x_1 & x_2 & x_3 \\ \hline & 0 & -10 & -30 & -20 \\ \hline y_1 & -3 & -1 & -3 & -6 \\ y_2 & 3 & 8 & 1 & 2 \\ y_3 & 1 & \boxed{1} & 1 & 1 \end{array} \quad \xrightarrow{y_3 \Leftrightarrow x_1} \quad \begin{array}{c|c|cc} & & x_2 & x_3 \\ \hline & 10 & -20 & -10 \\ \hline y_1 & -2 & -2 & -5 \\ y_2 & -5 & -7 & -6 \\ x_1 & 1 & 1 & 1 \end{array}$$

Phase 2: Das Starttableau ist bereits dual zulässig, also entfällt die Phase 3.

$$\begin{array}{c|c|cc} & & x_2 & x_3 \\ \hline & 10 & -20 & -10 \\ \hline y_1 & -2 & -2 & -5 \\ y_2 & -5 & -7 & \boxed{-6} \\ x_1 & 1 & 1 & 1 \end{array} \quad \xrightarrow{y_2 \Leftrightarrow x_3} \quad \begin{array}{c|c|cc} & & x_2 & x_3 \\ \hline & 18\frac{1}{3} & -\frac{25}{3} & -\frac{5}{3} \\ \hline y_1 & \frac{13}{6} & * & * \\ y_2 & \frac{5}{6} & * & * \\ x_1 & \frac{1}{6} & * & * \end{array}$$

Diese Tableau ist primal und dual zulässig.

Beispiel:
Mischung von 4 Brennstoffen mit *maximalem* Heizwert:

$$\max z(x) = 7x_1 + 4x_2 + 5x_3 + 6x_4 \quad \text{unter} \quad \begin{array}{rcl} 2x_1 + x_2 + 1.2x_3 + 1.6x_4 & \leq & 4000 \\ x_1 + x_2 + x_3 + x_4 & \leq & 3000 \\ x_3 & \leq & 500 \\ x_1 & \geq & 500 \\ x_1, x_2, x_3, x_4 & \geq & 0 \end{array}$$

Phase 1: entfällt, da keine Gleichungen vorhanden.

Phase 2: Das Starttableau ist weder primal noch dual zulässig:

		x_1	x_2	x_3	x_4
	0	7	4	5	6
y_1	4000	2	1	$1{,}2$	$1,6$
y_2	3000	1	1	1	1
y_3	500	0	0	1	0
y_4	-500	$\boxed{-1}$	0	0	0

$\xrightarrow{x_1 \leftrightarrow y_4}$

		y_4	x_2	x_3	x_4
	-3500	7	4	5	6
y_1	3000	2	1	$1{,}2$	$1{,}6$
y_2	2500	1	1	1	1
y_3	500	0	0	1	0
x_1	500	-1	0	0	0

Phase 3: Das Tableau ist nun primal zulässig. Nach vier weiteren primalen Simplexschritten erhält man das optimale Tableau:

		y_1	y_2	y_3	y_4
	$-15333\frac{1}{3}$	$-\frac{10}{3}$	$-\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$
x_4	$666\frac{1}{3}$	$*$	$*$	$*$	$*$
x_2	$1333\frac{1}{3}$	$*$	$*$	$*$	$*$
x_3	500	$*$	$*$	$*$	$*$
x_1	500	$*$	$*$	$*$	$*$

1.8 Zusammenfassung: Simplexverfahren

1. Primales Simplexverfahren

   (a) Grundform: $\min\{cx \mid Ax = b, x \geq 0\}$. Vor.: $b \geq 0, B$ Basis mit $A_B = E_m$, Tableau:

$$\begin{array}{c|c|c}
 & & x_n \\
\hline
 & z_0 & r_N \\
\hline
x_b & b & A_N
\end{array}
\qquad
\begin{array}{l}
z_0 = c_B b \\
r_N = c_B A_N - c_N
\end{array}$$

   (b) Zweiphasenmethode:

$$\min cx \text{ unter } \begin{array}{rcl}
a_i x & \geq & b_i, i = 1, \ldots, l \\
a_i x & = & b_i, i = l+1, \ldots, k \\
a_i x & \leq & b_i, i = k+1, \ldots, m \\
x & \geq & 0
\end{array}$$

   Definiere Hilfsfunktion $\min \sum_{i=1}^{k} y_i$ und erweitere die Nebenbedingungen zu:

$$\begin{array}{rcl}
a_i x - x_{n+i} + y_i & = & b_i, i = 1, \ldots, l \\
a_i x + y_i & = & b_i, i = l+1, \ldots, k \\
a_i x + y_i & = & b_i, i = k+1, \ldots, m \\
x_1, \ldots, x_{n+l}, y_1, \ldots, y_m & \geq & 0
\end{array}$$

   Phase 1.:
   Wähle Nichtbasis $N = (1, \ldots, n, n+1, n+l)$ und Basis $B = (n+l+1, \ldots, n+l+m)$.

$$z_0 = c_B b = \sum_{i=1}^{k} b_i, \quad r_j = (c_B A_N - c_N)_j = \begin{cases} \sum_{i=1}^{k} a_{ij} & j = 1, \ldots, n \\ -1 & j = n+1, \ldots, n+l \end{cases}$$

   Starttableau:

$$\begin{array}{c|c|c|c}
 & & x_1, \ldots, x_n & x_{n+1}, \ldots, x_{n+l} \\
\hline
 & z_0 & r_N & -1 \cdots -1 \\
\hline
y_1 & b_1 & a_1 & \\
\vdots & \vdots & \vdots & -E_l \\
y_l & b_k & a_l & \\
\hline
y_{l+1} & b_{k+1} & a_{k+1} & \\
\vdots & \vdots & \vdots & 0 \\
y_m & b_m & a_m &
\end{array}$$

Bringe die Schlupfvariablen $y_1, \ldots, y_k$ in die Nichtbasis und streiche die Spalten unter $y_1, \ldots, y_k$.

Phase 2.:

Berechne die neuen Werte $z_0 = c_B b$ und $r_N = c_B A_N - c_N$ mit der neuen Basis B und Nichtbasis N. Wende dann das Simplexverfahren *(1.4.20)* an.

2. Duales Simplexverfahren: Voraussetzung: $r_N \leq 0$, d.h. die Basislösung ist zulässig.

   (a) Grundform: $\min\{cx \mid Ax = b, x \geq 0\}$, Annahme: Sei B eine Basis mit $A_B = E_m$. Benutze das duale Simplexverfahren *(1.7.12)* mit dem Tableau:

		x_N
	$z_0 = c_B b$	$r_N = c_B A_N - c_N$
x_B	b	A_N

   (b) Ungleichungen: Voraussetzung: $c \geq 0$.

      i. $Ax \geq b \Leftrightarrow -Ax + y = -b, y \geq 0$. Tableau:

		x
	$z_0 = 0$	$r_N = -c \leq 0$
y	$-b$	$-A$

Wende *(1.7.12)* an

$$\max\{\lambda b \mid \lambda A \leq c, \lambda \geq 0\} \tag{P*}$$

$\lambda_j = -r_{N(j)}$ unter y_i als Nichtbasisvariable.

      ii.

$$\begin{aligned} a_i x &\geq b_i, i = 1, \ldots, k \\ a_i x &\leq b_i, i = k+1, \ldots, m \\ \Rightarrow -a_i x &\leq -b_i, i = 1, \ldots, k \end{aligned}$$

Tableau:

		x
	$z_0 = 0$	$r_n = -c$
y_1	$-b_1$	$-a_1$
$\vdots$	$\vdots$	$\vdots$
v_k	$-b_k$	$-a_k$
y_{k+1}	b_{k+1}	a_{k+1}
$\vdots$	$\vdots$	$\vdots$
y_m	b_m	a_m

3. Gemischte Verfahren

   (a) Drephasenmethode: Es wird weder $b \geq 0$ noch $r_N \leq 0$ vorausgesetzt.

$$\min cx \quad \text{unter} \quad \begin{aligned} a_i x &= b_i, i = 1, \ldots, k \\ a_i x &\leq b_i, i = 1, \ldots, m \\ x &\geq 0 \end{aligned}$$

Mit Schlupfvariablen bringt man diese Gleichungen alle auf die Form $a_i x + y_i = b_i$ für $i = 1, \ldots, m$.

Phase 1:

Bringe irgendwie $y_1, \ldots, y_k$ in die Nichtbasis. Einzige Bedingung an das Pivotelement ist, dass $a_{ps} \neq 0$. Streiche dann die Spalten unter den $y_1, \ldots, y_k$.

Phase 2:
Berechne mit dem dualen Simplexverfahren eine primal zulässige Lösung $x_B \geq 0$ *(1.7.12)*.
Phase 3:
Wende das primale Simplexverfahren *(1.4.20)* an.

Weitere Themen zur linearen Optimierung:

- Transportprobleme
- Flüsse in Netzwerken
- Ganzzahlige lineare Optimierung mit $x \in \mathbb{Z}^m$
- Parametrisierte lineare Optimierung
- Innere-Punkt-Verfahren (interior point methods)

Kapitel 2

Nichtlineare Optimierung

2.1 Unbeschränkte Optimierungsprobleme

2.1.1 Probleme und Beispiele

Sei $D \subset \mathbb{R}^n$ offen und sei $f : D \to \mathbb{R}$ eine C^k-Funktion ($k \geq 0$). Dann ist das Optimierungsproblem gegeben durch

$$\min \left\{ f(x) \mid x \in D \right\} \tag{2.1.1}$$

Ein Punkt $\bar{x} \in D$ heißt *lokale* bzw. *streng lokale Minimalstelle* von *2.1.1*, wenn es eine Umgebung $U \subset \mathbb{R}^n$ von $\bar{x}$ gibt mit

$$\forall x \in U : f(x) \geq f(\bar{x})$$

bzw.

$$\forall x \in U, x \neq \bar{x} : f(x) > f(\bar{x})$$

Beispiel: Testfunktion von Rosenbrock (1960)

$$f(x_1, x_2) = \left(100x_2 - x_1^2\right)^2 + \left(1 - x_1\right)^2, \quad D = \mathbb{R}^2$$

Die Lösung ist offensichtlich gegeben durch $x_1 = 1, x_2 = x_1^2 = 1$. Interessant sind dabei die Höhenlinien:

$$\left\{ (x_1, x_2)^T \in \mathbb{R}^2 \mid f(x_1, x_2) = \text{const} \right\}$$

Beispiel: Testfunktion von Himmelblau

$$f(x_1, x_2) = \left(x_1^2 + x_2 - 11\right)^2 + \left(x_1 + x_2^2 - 7\right)^2, \quad D = \mathbb{R}^2$$

Dies ist ein Polynom vierten Grades. Es gibt vier lokale Minimalstellen, vier Sattelpunkte, eine lokale Maximalstelle in $(-0.270845, -0.923039)^T$

Beispiel: Quadratische Probleme
Sei Q eine symmetrische $n \times n$ Marix und sei $b \in \mathbb{R}^n$.

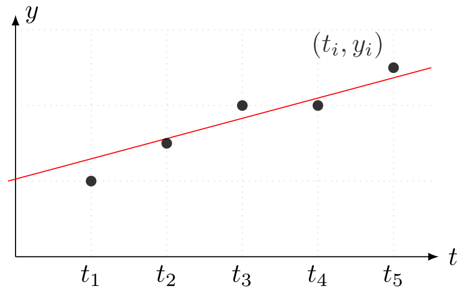
$$f(x) = \frac{1}{2} x^T Q x + b^T x$$

Beispiel: Nichtlineare Ausgleichsrechnung
Gegeben sind Messdaten $\left(t_i, y_i\right) \in \mathbb{R} \times \mathbb{R}$ für $i = 1, \ldots, m$. Gesucht ist eine Approximationsfunktion $g(x, t)$ mit $x \in \mathbb{R}^n, t \in \mathbb{R}$. Das dazugehörige Optimierungsproblem ist gegeben durch

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^{m} \left(y_i - g(x, t_i)\right)^2$$

Beispiele:

Ausgleichsgerade: $g(x_1, x_2, t) = x_1 + x_2 \cdot t$



Im allgemeinen erhalten wir das Ausgleichsproblem:

$$g(x,t) = x_1 + x_2 t + x_3 t^2 + \ldots + x_n t^{n-1}, m > n$$

Exponentielle Ausgleichsfunktion: $g(x_1, x_2, x_3, t) = x_1 e^{x_2 \cdot t} + x_3$

2.1.2 Existenz von Lösungen

2.1.2 Definition: (Niveaumenge)
Sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$. Für $\alpha \in \mathbb{R}$ heißen die Mengen

$$N(f, \alpha) := \{x \in D \mid f(x) \leq \alpha\}$$

die *Niveaumengen* von f.

2.1.3 Satz:
Die Funktion $f : D \to \mathbb{R}$ sei stetig. Für ein $y \in D$ sei die Niveaumenge $N(f, f(y)) = \{x \in D \mid f(x) \leq f(y)\}$ kompakt. Dann besitzt f auf D (mindestens) ein globales Minimum.

Beweis:
Der Satz von Weierstraß garantiert uns, dass es eine konvergente Minimalfolge und damit auch ein Minimum $\bar{x}$ auf der kompakten Menge $N(f, f(y))$ gibt. Für $x \in D \setminus N(f, f(y))$ gilt per Definition $f(x) > f(y)$. Also ist $\bar{x}$ auch ein globales Minimum. $\square$

2.1.4 Korollar:
Die Funktion $f : \mathbb{R}^n \to \mathbb{R}$ sei stetig und es gelte $\lim_{||x||_2 \to \infty} f(x) = +\infty$. Dann ist für alle $y \in \mathbb{R}^n$ die Niveaumenge $N(f, f(y))$ kompakt. Also gibt es (mindestens) ein globales Minimum von f auf $\mathbb{R}^n$

Beispiel:
Sei Q eine symmetrische und positiv definite $n \times n$ Matrix und $b \in \mathbb{R}^n$. Für $f(x) := \frac{1}{2} x^T Q x + b^T x$ gilt $\lim_{||x||_2 \to \infty} f(x) = +\infty$. Also hat f ein globales Minimum.
Notwendige Bedingung: $\nabla f(\bar{x}) = Q\bar{x} + b = 0, \bar{x} = -Q^{-1} b, \nabla^2 f(\bar{x}) = Q > 0$.

Lineare Ausgleichsrechnung:
Sei A eine $m \times n$ Matrix mit $m \geq n$ und sei $b \in \mathbb{R}^m$.

$$\min_{x \in \mathbb{R}^n} f(x) := \frac{1}{2} ||Ax - b||_2^2 = \frac{1}{2} x^T A^T A x - x^T A^T b + \frac{1}{2} ||b||_2^2$$

Vorausgesetzt $\text{rg}(A) = n \leq m$ gilt $\nabla f(\bar{x}) = A^T A \bar{x} - A^T b = 0$ und $\nabla^2 f(\bar{x}) = A^T A$ ist positiv definite $n \times n$ Matrix

2.1.3 Optimalitätsbedingungen

Notation:

- $f'(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)\right)$ Zeilenvektor
- $\nabla f(x) = f'(x)^T$ Spaltenvektor
- Falls f zweimal differenzierbar ist, so ist die symmetrische $n \times n$ Hesse-Matrix gegeben durch

$$\nabla^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{1 \le i,j \le n}$$

2.1.5 Satz: *Notwendige Optimalitätsbedingungen erster und zweiter Ordnung*
Sei $D \subset \mathbb{R}^n$ offen und sei $f : D \to \mathbb{R}$. Sei $\bar{x} \in D$ eine lokale Minimalstelle von f.

(i) Ist f differenzierbar in $\bar{x}$, so gilt $\nabla f(\bar{x}) = 0 = f'(\bar{x})$.

(ii) Ist f zweimal stetig differenzierbar in einer Umgebung von $\bar{x}$, so gelten die Bedingungen
   (a) $f'(\bar{x}) = 0$
   (b) $\nabla^2 f(x)$ ist positiv semidefinit, also $\forall v \in \mathbb{R}^n : v^T(\nabla^2 f(\bar{x}))v \ge 0$.

Beweis:

1. Für $v \in \mathbb{R}^n$ gilt $f(\bar{x} + tv) \ge f(\bar{x})$ für alle $t \ge 0$ genügend klein. Also gilt

$$\forall v \in \mathbb{R}^n : f(\bar{x})v = \lim_{t \searrow 0} \frac{f(\bar{x} + tv) - f(\bar{x})}{t} \ge 0$$

   Also muss schon $f'(\bar{x}) = 0$.

□

2.1.6 Satz: *Hinreichende Bedingung zweiter Ordnung (second-order sufficient conditions, ssc)*
Die Funktion $f : D \to \mathbb{R}$ sei zwei mal stetig differenzierbar in einer Umgebung eines Punktes $\bar{x} \in D$. In $\bar{x}$ gelten die folgenden Bedingungen:

(i) $f(\bar{x}) = 0$

(ii) $\forall v \in \mathbb{R}^n \setminus \{0\} : v^T f''(\bar{x})v > 0$, d.h. $f''(\bar{x})$ ist positiv definit.

Dann gibt es Konstanten $c > 0$ und $\alpha > 0$ mit

$$\forall\ ||x - \bar{x}||_2 \le \alpha : f(x) \ge f(\bar{x}) + c \cdot ||x - \bar{x}||_2^2$$

Insbesondere ist $\bar{x} \in D$ eine strikte lokale Minimalstelle.

Beweis:
Sei $\lambda_{\min} \in \mathbb{R}$ der kleinste Eigenwert von $f''(\bar{x})$. Da $f''(\bar{x})$ positiv definit ist, gilt $\lambda_{\min} > 0$ und

$$\forall v \in \mathbb{R}^n : v^T f''(\bar{x})v \ \le\ \lambda_{\min} \cdot v^T v$$

Mit Taylorentwicklung von $f(x)$ um $\bar{x} \in D$ folgt:

$$\begin{aligned} f(x) &= f(\bar{x}) + \underbrace{f'(\bar{x})(x - \bar{x})}_{=0} + \frac{1}{2}(x - \bar{x})^T f''(\bar{x})(x - \bar{x}) + \mathcal{O}\Big(||x - \bar{x}||_2^2\Big) \\ &\ge f(\bar{x}) + \frac{1}{2}\lambda_{\min}||x - \bar{x}||_2^2 + \mathcal{O}\Big(||x - \bar{x}||_2^2\Big) \end{aligned}$$

Zu beliebigem $\varepsilon > 0$ mit $0 < \varepsilon < \frac{1}{2}\lambda_{\min}$ gibt es ein $\alpha > 0$ mit

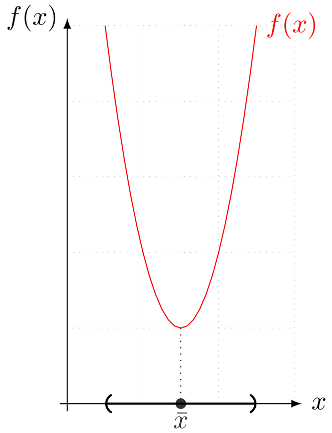
$$\forall\ ||x-\bar{x}||_2 \leq \alpha : \left|\mathcal{O}\Big(||x-\bar{x}||_2^2\Big)\right| \leq \varepsilon \cdot ||x-\bar{x}||_2^2$$

Also ist

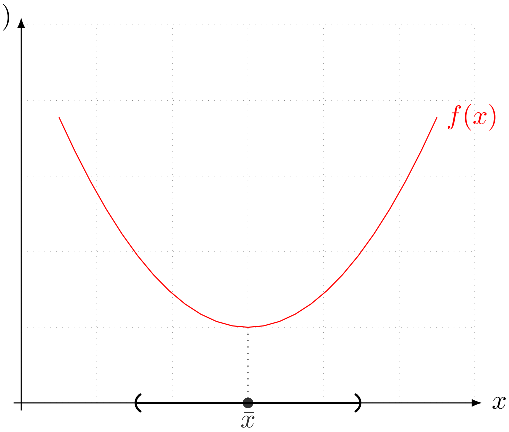
$$\begin{aligned} f(x) &\geq f(\bar{x}) + \frac{1}{2}\lambda_{\min}||x-\bar{x}||_2^2 - \varepsilon||y-\bar{x}||_2^2 \\ &= f(\bar{x}) + \underbrace{\left(\frac{1}{2}\lambda_{\min} - \varepsilon\right)}_{:=c>0} \cdot ||x-\bar{x}||_2^2 \end{aligned}$$

für alle $||x-\bar{x}||_2 \leq \alpha$. $\square$

Geometrische Bedeutung von $c \approx \frac{1}{2}\lambda_{\min}$:



c ist groß $\Rightarrow$ starke Krümmung, schnelle Konvergenz



c ist klein $\Rightarrow$ kleine Krümmung, langsame Konvergenz.

Beispiel:

Lineare Ausgleichung:
Sei A eine $m \times n$ Matrix, $n \leq m, b \in \mathbb{R}^m$, $\operatorname{rang}(A) = n$.

$$\min_{x\in\mathbb{R}^n} f(x) := \frac{1}{2}||Ax-b||_2^2, f'(\bar{x}) = 0 \Rightarrow A^TA\bar{x} = A^Tb$$

$$f''(\bar{x}) = A^TA > 0 \text{ positiv definit}$$

Rosenbrock-Funktion: $f(x_1,x_2) = 100(x_2 - x_1^2)^2 + (1-x_1)^2, \bar{x} = \binom{1}{1}$.

$$f''(x) = \begin{pmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}, \quad f''(\bar{x}) = \begin{pmatrix} 802 & -400 \\ -400 & 200 \end{pmatrix}$$

Zusammenhang der Bedingung $f''(\bar{x}) > 0$ mit dem Newtonverfahren:
Zu lösen ist die nichtlineare Gleichung $F(x) := \nabla f(x) = 0$ für $F : D \to \mathbb{R}^n$:

Start: $x^0 \in D$

Iteration: $x^{k+1} = x^k - F'\big(x^k\big)^{-1}F\big(x^k\big) = x^k - f''\big(x^k\big)^{-1}\nabla f\big(x^k\big),\ k = 0,1,\ldots$

Voraussetzung dabei ist, dass $F'\big(x^k\big) = f''\big(x^k\big)$ regulär ist. Die ist zu erfüllen mit der Voraussetzung $f''(\bar{x}) > 0$ positiv definit, da dann $f''\big(x^k\big) > 0$ für $||x^k - \bar{x}||_2$ klein.

Numerische Berechnung:

Setze $d^k := x^{k+1} - x^k$ und löse das lineare Gleichungssystem $f''\big(x^k\big)d^k = -\nabla f\big(x^k\big)$ und setze $x^{k+1} = x^k + d^k$. Führe außerdem eine Modifikation mit Schrittweite λ_k ein:

$$x^{k+1} = x^k + \lambda_k d^k,\ 0 < \lambda_k \leq 1$$

Berechnung von λ_k etwa durch $\min_{\lambda \geq 0} f\big(x^k + +\lambda d^k\big)$

2.1.4 Sensitivitätsanalyse

2.1.7 Satz: *Satz über implizite Funktionen*
Sei $F : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n$ eine Abbildung, welche in einer Umgebung eines Punktes $(x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^k$ mindestens j-mal stetig differenzierbar ist ($j \geq 1$). Weiter gelte

(i) $F(x_0, y_0) = 0$

(ii) $\frac{\partial F}{\partial x}(x_0, y_0)$ ist eine reguläre $n \times n$-Matrix

Dann gibt es Umgebungen $U(y_0)$ von y_0 und $V(x_0)$ von x_0 und eindeutig bestimmte j mal stetig differenzierbare Abbildung $x : U(y_0) \to V(x_0)$ mit der Eigenschaft

$$\forall y \in U(y_0) : F(x(y), y) = 0$$

Außerdem gilt

$$\frac{dx}{dy} = x'(y) = -\left(\frac{\partial F}{\partial x}\Big(x(y), y\Big)\right)^{-1} \frac{\partial F}{\partial y}\Big(x(y), y\Big)$$

für alle $y \in U(y_0)$.

Beweis:
siehe Analysis II. □

Anwendung auf die Optimierung: Parametrische Optimierung:
Sei $p \in \mathbb{R}^k$ ein Störungsparameter bzw. ein Sensitivitätsparameter. Sei weiter $f : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}, f = f(x,p), x \in \mathbb{R}^n, p \in \mathbb{R}^k$. Dann ist für festes p das Parametrische Optimierungsproblem gegeben durch:

$$\min_{x \in \mathbb{R}^n} f(x,p) \qquad \text{OP}(p)$$

Sei $\bar{x} = x_0$ eine lokale Minimalstelle des Problems $\text{OP}(p_0)$ für einen nominellen Parameter $p_0 \in \mathbb{R}^k$.
Vorausgesetzt $f \in C^2$ bezüglich (x,p) in einer Umgebung (x_0, p_0)

2.1.8 Satz: *Sensitivitätssatz*
$x_0 \in \mathbb{R}^n$ genüge den hinreichenden Bedingungen zweiter Ordnung *(2.1.6)* für das Problem $\text{OP}(p_0)$ für ein $p_0 \in \mathbb{R}^k$, d.h. $f_x(x_0, p_0) = 0$ und $f_{xx}(x_0, p_0) > 0$.

Dann gibt es eine Umgebung $P_0 \subset \mathbb{R}^k$ von p_0 und eine stetig differenierbare Abbildung $x : P_0 \to \mathbb{R}^n$ mit

(i) $x(p_0) = x_0$

(ii) $x(p)$ ist eine strenge lokale Minimalstelle des Problems $\text{OP}(p)$ $\min_{x \in \mathbb{R}^n} f(x,p)$ für alle $p \in P_0$.

(iii) $\forall p \in P_0 : \frac{d}{dp} f(x(p), p) = \frac{\partial f}{\partial p}(x(p), p) = f_p(x(p), p)$

(iv) $\frac{dx}{dp}(p) = -f_{xx}(x(p), p)^{-1} f_{xp}(x(p), p)$

2.2 Optimierungsprobleme mit Nebenbedingungen: Problemformulierung und Optimalitätsbedingung

2.2.1 Definition: (Allgemeines Optimierungsproblem)
Sei $D \subset \mathbb{R}^n$ offen und sei $S \subset D$, $f : D \to \mathbb{R}$. Dann ist das allgemeine Optimierungsproblem gegeben durch

$$\min \{ f(x) \mid x \in S \}$$

In den Anwendungen: S ist definiert durch Gleichungen und Ungleichungen. Sei $g : D \to \mathbb{R}^m, g = (g_1, \ldots, g_m)^T$, $k \in \mathbb{N}$ mit $0 \leq k \leq m$. Definiere

$$S := \{ x \in D \mid g_i(x) \leq 0 \text{ für } i = 1, \ldots, k,\ g_i(x) = 0 \text{ für } k+1, \ldots, m \}$$

So erhalten wir:

2.2.2 Definition: (Standardproblem der nichtlinearen Optimierung)
Das Standardproblem der nichtlinearen Optimierung lautet

$$\min \{ f(x) \mid g_i(x) \leq 0 (i = 1, \ldots, k),\ g_i(x) = 0 (i = k+1, \ldots, m) \}$$

Andere Schreibweise::
Sei $K \subset \mathbb{R}^m$ der konvexe und abgeschlossene Kegel

$$K := \mathbb{R}^k_- \times \{0_{m+k}\}, \quad \mathbb{R}^k_- = \{ y \in \mathbb{R}^k \mid y_i \leq 0, i = 1, \ldots, k \}$$

Dann ist *(2.2.2)* äquivalent zu

$$\min \{ f(x) \mid g(x) \in K \} \tag{2.2.3}$$

Diese Darstellung hat den Vorteil, dass sie koordinatenfrei ist und sich so auch auf Banachräume verallgemeinern lässt.

Optimierungsprobleme in Banachräumen:
Seien X und Y Banachräume, $D \subset X$ offen, $f : D \to \mathbb{R}$ sowie $g : D \to Y$. Weiter sei $K \subset Y$ eine konvexe und abgeschlossene Menge.

$$\min \{ f(x) \mid g(x) \in K \} \tag{2.2.4}$$

Der Vorteil in dieser Darstellung liebt auch wieder in der koordinatenfreien Formulierung. Anwenden lässt sich dies zum Beispiel bei der Optimierung von Steuerprozessen für Gewöhnliche und Partielle Differentialgleichungen.

2.2.1 Notwendige Optimalitätsbedingung

Ausgangspunkt ist unser Standardproblem *(2.2.2)*. Sei $\bar{x} \in S$ zulässig, d.h. $g_i(\bar{x}) \leq 0$ für $i = 1, \ldots, k$ und $g_i(\bar{x}) = 0$ für $i = k+1, \ldots, m$. Wir definieren uns die folgenden aktiven Indexmengen:

$$I(\bar{x}) := \{ i \in \{1, \ldots, k\} \mid g_i(\bar{x}) = 0 \}, \quad J(\bar{x}) := I(\bar{x}) \cup \{k+1, \ldots, m\}$$

Dann gilt

$$\forall i \in J(\bar{x}) : g_i(\bar{x}) = 0$$

2.2.5 Definition: (Lagrange-Funktion)
Für $\lambda_0 \in \mathbb{R}$ und $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ gilt

$$L(x, \lambda_0, \lambda) := \lambda_0 f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) = \lambda_0 f(x) + \lambda g(x)$$

Man benötigt Regularitätsbedingungen, um $\lambda_0 = 1$ setzen zu können.

2.2.6 Definition: (Regularitätsbedingung)

(i) Ein Punkt $\bar{x} \in S$ heißt *normal*, wenn die Gradienten $g_i'(\bar{x})$ für alle $i \in J(\bar{x})$ linear unabhängig sind. (Dies ist nichts anderes als die Standardbedingung der linearen Optimierung (rg $A = m$) im nichtlinearen Kontext)

(ii) Ein Punkt $\bar{x} \in S$ heißt *regulär*, wenn die Bedingung von Mangasarian/Fromovitz gilt:

  (a) die Gradienten $g_i'(\bar{x})$ sind für $i = k+1, \ldots, m$ linear unabhängig

  (b) Es gibt $v_0 \in \mathbb{R}^n$ mit $g_i'(\bar{x})v_0 < 0$ für $i \in I(\bar{x})$ sowie $g_i'(\bar{x})v_0 = 0$ für $i = k+1, \ldots, m$

2.2.7 Satz: *Notwendige Optimalitätsbedingungen von John und Karush/Kuhn/Tucker*
Sei $\bar{x} \in S$ eine lokale Minimalstelle des Standardproblems *(2.2.2)*. Weiter sei $f : D \to \mathbb{R}$ differenzierbar in $\bar{x}$ und $g : D \to \mathbb{R}^m$ stetig differenzierbar in einer Umgebung von $\bar{x}$.

Dann gibt es $\lambda_0 \geq 0$ und einen Zeilenvektor $\bar{\lambda} = \bar{\lambda}_1, \ldots, \bar{\lambda}_m \in \mathbb{R}^m$ mit $(\lambda_0, \bar{\lambda}) \neq (0, 0_m)$ und den Eigenschaften

(i) $L_x(\bar{x}, \lambda_0, \bar{\lambda}) = \lambda_0 f'(\bar{x}) + \bar{\lambda} g'(\bar{x}) = \lambda_0 f'(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i g_i'(\bar{x}) = 0$

(ii) $\bar{\lambda}_i = 0$ für $i \notin J(\bar{x})$, d.h. $i \in \{1, \ldots, k\}$ mit $g_i(\bar{x}) < 0$

(iii) $\bar{\lambda}_i \geq 0$ für $i \in I(\bar{x})$

Es gilt $\lambda_0 > 0$, wenn $\bar{x}$ regulär ist. Dann kann $\lambda_0 = 1$ gesetzt werden (teile die Lagrange-Funktion durch λ_0). Falls $\bar{x}$ normal ist, dann ist $\bar{\lambda} \in \mathbb{R}^m$ mit $\lambda_0 = 1$ eindeutig bestimmt.

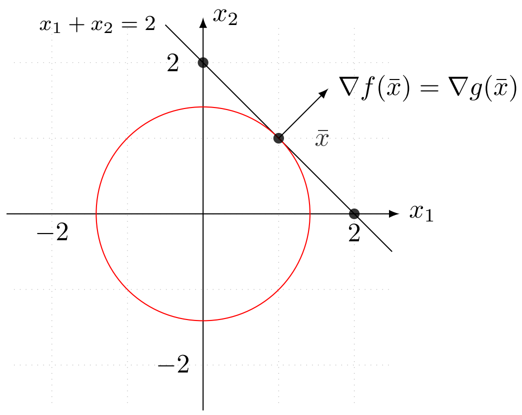
2.2.2 Beispiele

Vorüberlegung:
Sei $f : D \to \mathbb{R}$ differenzierbar, z.B. $f(x_1, x_2) = x_1^2 + x_2^2$. Wir betrachten die Niveauflächen $\{x \in \mathbb{R}^n \mid f(x) = \text{const}\}$.

[Hier ein hübsches Bild mit nem Kreis, und dem Gradienten der Senkrecht steht]

Merke: Der Gradient $\nabla f(\bar{x})$ steht senkrecht auf der Niveaufläche im Punkt $\bar{x}$

Beispiel:
$\min \left\{ f(x_1, x_2) = x_1^2 + x_2^2 \mid g(x_1, x_2) = x_1 + x_2 - 2 = 0 \right\}$

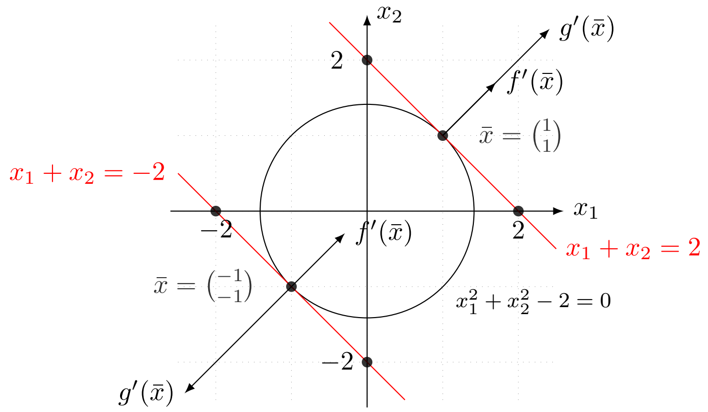


Idee: Niveaulinie berührt die Nebenbedingung in $\bar{x}$, der Punkt $\bar{x} = \binom{1}{1}$ ist optimal. Die Gradienten $f'(\bar{x})$ und $g'(\bar{x})$ sind parallel, d.h. es gibt $\bar{\lambda} \in \mathbb{R}$ mit $f'(\bar{x}) + \lambda g'(\bar{x}) = 0$. Wir berechnen $\bar{\lambda}$:

$$\begin{aligned} f'(x) &= 2(x_1, x_2), f'(\bar{x}) = (2,2), \\ g'(x) &= (1,1), f'(\bar{x}) + \bar{\lambda} g'(\bar{x}) = (2,2) + \bar{\lambda}(1,1) = 0 \\ &\Rightarrow \bar{\lambda} = -2 \end{aligned}$$

$\bar{x} = \binom{1}{1}$ ist normal, da $g'(\bar{x}) = (1,1) \neq (0,0)$.

Beispiel:
$\min \left\{ f(x_1, x_2) = x_1 + x_2 \mid g(x_1, x_2) = x_1^2 + x_2^2 - 2 = 0 \right\}$



Kandidaten für ein Optimum: $\bar{x} = \binom{1}{1}$ und $\bar{x} = \binom{-1}{-1}$. Es ist $f'(x) = (1,1), g'(x) = 2(x_1, x_2)$, also:

$$\begin{aligned} \bar{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}: & \quad f'(\bar{x}) = (1,1), g'(\bar{x}) = (2,2) \neq (0,0) \\ &\Rightarrow \bar{x} \text{ ist normal} \\ f'(\bar{x}) + \bar{\lambda} g'(\bar{x}) &= (1 + \bar{\lambda} \cdot 2, 1 + \bar{\lambda} \cdot 2) \stackrel{!}{=} (0,0) \\ &\Rightarrow \bar{\lambda} = -\frac{1}{2} \\ \bar{x} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}: & \quad f'(\bar{x}) = (1,1), g'(\bar{x}) = (-2,-2) \neq (0,0) \\ &\Rightarrow \bar{x} \text{ ist normal} \\ f'(\bar{x}) + \bar{\lambda} g'(\bar{x}) &= (1 - \bar{\lambda} \cdot 2, 1 - \bar{\lambda} \cdot 2) \stackrel{!}{=} (0,0) \\ &\Rightarrow \bar{\lambda} = \frac{1}{2} \end{aligned}$$

Also ist $\bar{x} = \binom{1}{1}$ ein Maximum und $\bar{x} = \binom{-1}{-1}$ ein Minimum, *aber* dies kann mit den KKT-Bedingungen erster Ordnung nicht entschieden werden. Benötigt werden die hinreichenden Bedingungen zweiter Ordnung: SSC.

Betrachte nun die Nebenbedinung $g(x_1, x_2) = x_1^2 + x_2^2 - 2 \leq 0$, also den vollen Kreis. Hier muss nun bei den KKT-Bedingungen $\bar{\lambda} \geq 0$ gelten. Daher ist $\bar{x} = \binom{-1}{-1}$ der einzige Punkt, der diese Bedingungen erfüllt. Dies liegt an der Konvexität der zulässigen Menge und wir haben daher einen Spezialfall einen konvexen Optimierungsproblemes. Dort hatten wir bereits gesehen, dass jedes lokale Minimum auch ein globales Minimum ist.

Beispiel: Geometrisches Mittel $\leq$ Arithmetisches Mittel

$$\left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}} \leq \frac{1}{n} \sum_{i=1}^{n} x_i$$

1. Behauptung: $\max\left\{\prod_{i=1}^n x_i \mid \sum_{i=1}^n x_i = 1, x_i \geq 0\right\} = \left(\frac{1}{n}\right)^n$, mit $x_i > 0, i = 1, \ldots, n$

Lagrange-Funktion:

$$L(x, \lambda) = \prod_{i=1}^n x_i + \lambda\left(\sum_{i=1}^n x_i - 1\right), \; x \in \mathbb{R}^n, \lambda \in \mathbb{R}$$

Aus der KKT-Bedingung (i)

$$\frac{\partial L}{\partial x_k}(\bar{x}, \lambda x) = \prod_{i \neq k} \bar{x}_i + \bar{\lambda} = 0 \text{ für } k = 1, \ldots, n$$

folgt $\bar{x}_1 = \ldots = \bar{x}_n$. Wegen $\sum_{i=1}^n \bar{x}_i = 1$ folgt $\bar{x}_i = \frac{1}{n}$ für $i = 1, \ldots, n$, also die Behauptung.

2. Behauptung: $\forall x_i \geq 0 : \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^n x_i$

Seien $x_i \in \mathbb{R}_+$ beliebig. Setze $y_i := \frac{x_i}{x_1 + \ldots + x_n} \geq 0$, dann gilt $\sum_{i=1}^n y_i = 1$. Mit der ersten Behauptung gilt:

$$\prod_{i=1}^n y_i \leq \left(\frac{1}{n}\right)^n \iff \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^n x_i$$

Direkter Beweis:

$f(x) = \ln x$ ist eine konkave Funktion für $x > 0$, d.h. nach der Jensenschen Ungleichung gilt

$$\forall x_i > 0 \;\; \forall \alpha_i \geq 0, \sum_{i=1}^n \alpha_i = 1 \; : \; \ln\left(\sum_{i=1}^n \alpha_i x_i\right) \geq \sum_{i=1}^n \alpha_i \ln x_i$$

Mit $\alpha_i = \frac{1}{n}$ folgt

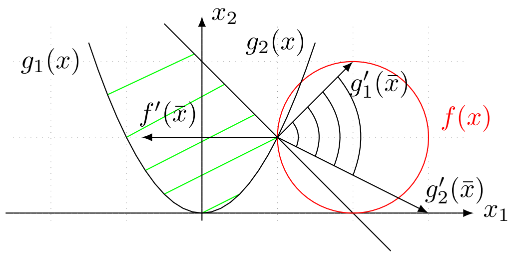
$$\ln\left(\frac{1}{n}\sum_{i=1}^n x_i\right) \geq \frac{1}{n}\sum_{i=1}^n \ln x_i = \ln\left(\left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}\right)$$

und damit bereits die Behauptung.

Beispiel:

Wir betrachten

$$\min\left\{f(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 1)^2 \mid g_1(x_1, x_2) = x_1^2 - x_2 \leq 0, g_2(x_1, x_2) = x_1 + x_2 - 2 \leq 0\right\}$$



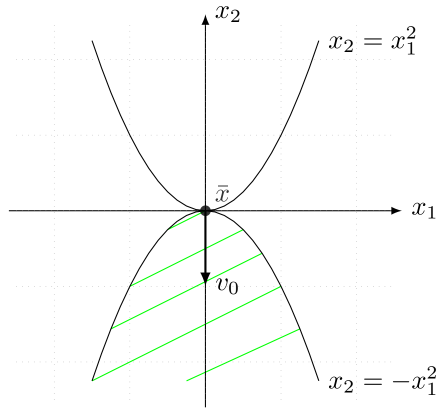
Beh.: $\bar{x} = \binom{1}{1}$ erfüllt die KKT-Bedingungen:

$$\begin{aligned} f'(\bar{x}) &= \big(2(\bar{x}_1 - 2), 2(\bar{x}_2 - 1)\big) = (-2, 0) \\ g_1'(\bar{x}) &= (2\bar{x}_1, -1) = (-2, 1) \\ g_2'(\bar{x}) &= (1, 1) \end{aligned}$$

$\bar{x}$ ist normal, da $g_1'(\bar{x}) = (2, -1), g_2'(\bar{x}) = (1, 1)$ linear unabhängig sind. Es gilt $I(\bar{x}) = \{1, 2\}$. Also sagen die KKT-Bedingungen:

$$f'(\bar{x}) + \bar{\lambda}_1 g_1'(\bar{x}) + \bar{\lambda}_2 g_2'(\bar{x}) = 0 \;\Rightarrow\; \bar{\lambda}_1 = \bar{\lambda}_2 = \frac{2}{3} > 0$$

Beispiel:
$\min\left\{f(x_1,x_2) = -x_2 \mid g_1(x_1,x_2) = x_2 + x_1^2 \le 0,\; g_2(x_1,x_2) = x_2 - x_1^2 \le 0\right\}$



Offensichtlich ist $\bar{x} = \binom{0}{0}$ eine globale Minimalstelle und $g_1'(\bar{x}) = (0,1), g_2'(\bar{x}) = (0,1)$ sind linear abhängig. Also ist $\bar{x}$ nicht normal, aber $\bar{x}$ ist regulär: $I(\bar{x}) = \{1,2\}$ und für $v_0 = \binom{0}{-1} \in \mathbb{R}^2$ gilt:

$$g_i'(\bar{x})v_0 = -1 < 0 \text{ für } i \in I(\bar{x})$$

Betrachten wir die KKT-Bedingungen:

$$f'(\bar{x}) + \bar{\lambda}_1 g_1'(\bar{x}) + \bar{\lambda}_2 g_2'(\bar{x}) = (0, -1 + \bar{\lambda}_1 + \bar{\lambda}_2) = (0,0), \quad \bar{\lambda}_1, \bar{\lambda}_2 \ge 0$$

Die Menge der KKT-Multiplikatoren:

$$\Lambda(\bar{x}) = \left\{\bar{\lambda} \in \mathbb{R}^2 \mid \bar{\lambda}_1 + \bar{\lambda} = 1,\; \bar{\lambda}_1, \bar{\lambda}_2 \ge 0\right\}$$

ist nicht einpunktig aber kompakt

2.2.3 Hinreichende Bedingungen 2. Ordnung (SSC) in einem Spezialfall

2.2.8 Satz: *SSC im Spezialfall*
Seien die Abbildungen $f: D \to \mathbb{R}$ und $g: D \to \mathbb{R}^m$ in einer Umgebung eines zulässigen Punktes $\bar{x} \in S \subset D$ zwei mal stetig differenzierbar. Der Punkt $\bar{x} \in S$ sei normal, d.h. $g_i'(\bar{x}), i \in J(\bar{x})$ sind linear unabhängig. Weiter existiere ein Multiplikator $\bar{\lambda} \in \mathbb{R}^m$, für den die drei folgenden Bedingungen gelten:

(i) $L_x(\bar{x}, \bar{\lambda}) = f'(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i g_i'(\bar{x}) = 0$

(ii) $\bar{\lambda}_j = 0$ für $j \notin J(\bar{x})$ und $\bar{\lambda}_i > 0$ für $i \in I(\bar{x})$ (strikte Komplementarität)

(iii) $v^T L_{xx}(\bar{x}, \bar{\lambda})v = v^T\left(f''(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i g_i''(\bar{x})\right)v > 0$ für alle $v \in \mathbb{R}^n \setminus \{0\}$ mit $g_i'(\bar{x})v = 0$ für $i \in J(\bar{x})$.

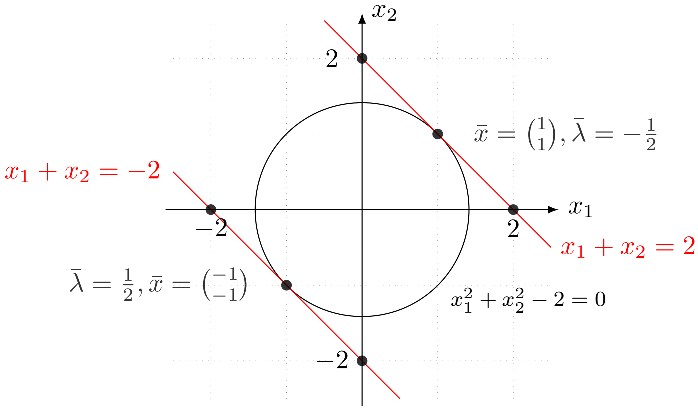
Dann gibt es $c > 0$ und $\alpha > 0$ mit:

$$\forall ||x - \bar{x}||_2 \le \alpha : f(x) \ge f(\bar{x}) + c \cdot ||x - \bar{x}||_2^2$$

Daher ist $\bar{x}$ eine strikte lokale Minimalstelle.

Betrachten wir einige Beispiele zu Anwendung von *(2.2.7.iii)*:

Beispiel: vgl. Beispiel 2 zur KKT-Bedingung
$\min\left\{f(x_1,x_2) = x_1 + x_2 \mid g(x_1,x_2) = x_1^2 + x_2^2 - 2 = 0\right\}$

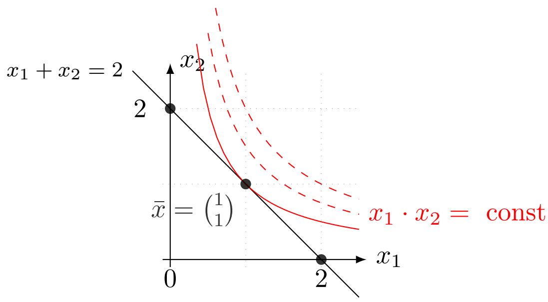


Lagrange-Funktion mit $\lambda_0 = 1$:

$$\begin{aligned} L(x, \lambda) &= x_1 + x_2 + \lambda\big(x_1^2 + x_2^2 - 2\big) \\ L_{xx}(\bar{x}, \bar{\lambda}) &= \bar{\lambda} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \end{aligned}$$

Nur für $\bar{\lambda} > 0$ kann gelten $v^T L_{xx}(\bar{x}, \bar{\lambda}) v = 2\bar{\lambda}(v_1^2 + v_2^2) > 0$, für $v \neq 0$ mit $g'(\bar{x})v = 2\big(\bar{x}_1 v_1 + \bar{x}_2 v_2\big) = 0$. Daher ist $\bar{x} = \binom{-1}{-1}$ mit $\bar{\lambda} = \frac{1}{2}$ die (globale) Minimalstelle.

Beispiel:
$\max\left\{x_1 \cdot x_2 \mid x_1 + x_2 - 2 = 0,\ x_1, x_2 > 0\right\}$



Hier: $\min f(x_1, x_2) = -x_1 \cdot x_2$ unter ...

$$L(x, \lambda) = -x_1 \cdot x_2 + \lambda(x_1 + x_2 - 2)$$

$$(0, 0) = L_x(\bar{x}, \bar{\lambda}) = (-\bar{x}_2 + \bar{\lambda}, -\bar{x}_1 + \bar{\lambda}) \ \Rightarrow\ \bar{x}_1 = \bar{x}_2 = 1, \bar{\lambda} = 1$$

Auswertung der SSC:

$$L_{xx}(\bar{x}, \bar{\lambda}) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \text{ weder positiv noch negativ definit}$$

Berechne $v^T L_{xx}(\bar{x}, \bar{\lambda}) v = (-v_2, -v_1)\binom{v_1}{v_2} = -v_2 v_1 - v_1 v_2$. Jedoch gilt $g'(\bar{x})v = v_1 + v_2 = 0$ also $v_2 = -v_1$

$$\Rightarrow\ v^T L_{xx}(\bar{x}, \bar{\lambda}) v = v_1^2 + v_1^2 = 2v_1^2 > 0 \text{ für } v \in \mathbb{R}^2 \setminus \{0\} \text{ mit } v_1 + v_2 = 0$$

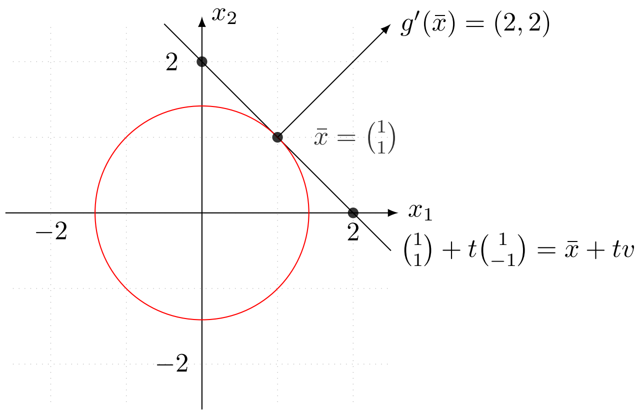
2.3 Tangentialkegel, Linearisierender Kegel und Regularität

2.3.1 Einführung

An einigen Beispielen werden wir die Begriffe Rangentialraum, Tangentialkegel und Konische Hülle kennenlernen:

Beispiel: Gleichungen: $g : \mathbb{R}^n \to \mathbb{R}^m$ sei eine C^1-Abbildung, $S = \{x \in \mathbb{R}^n \mid g(x) = 0\}$

1. $n = 2, m = 1$, z.B. $g(x_1, x_2) = x_1^2 + x_2^2 - 2 = 0$



   Es ist $\bar{x} = \binom{1}{1}, g'(\bar{x}) = (2, 2)$
   Der *Tangentialraum* ist gegeben durch alle $v \in \mathbb{R}^2$ mit $g'(\bar{x})v = 0$, d.h. $v_1 + v_2 = 0$:

$$\mathbb{R} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \left\{ t \begin{pmatrix} 1 \\ -1 \end{pmatrix} \;\middle|\; t \in \mathbb{R} \right\}$$

   Parametrisierung der Tangente: $\bar{x} + tv = \bar{x} + t\binom{1}{-1} = \binom{1+t}{1-t}, t \in \mathbb{R}$.

2. $n = 3, m = 1$, z.B. $g(x) = x_1^2 + x_2^2 + x_3^2 - r^2 = 0$

   [hier ein hübsches dreidimensionales Bildchen mit Kugel, Tangential ebene und senkrechtem $g'(\bar{x})$]

   Hier ist der Tangentialraum in $\bar{x}$ gegeben durch

$$\left\{ v \in \mathbb{R}^3 \;\middle|\; g'(\bar{x})v = 2(\bar{x}_1 v_1 + b x_2 v_2 + \bar{x}_3 v_3) = 0 \right\}$$

   und bildet einen zweidimensionalen Unterraum.

Bemerkung:
Wir zeigen später: Ist $\mathrm{rg}(g'(\bar{x})) = m \Rightarrow$ es gibt eine Kurve $x : [-\alpha, \alpha] \to \mathbb{R}^n, \alpha > 0$ mit der Eigenschaft

$$\begin{aligned} \forall t \in [-\alpha, \alpha] \quad &: \quad g(x(t)) = 0,\ x(0) = \bar{x} \\ v \quad &= \quad \lim_{t \to 0} \frac{x(t) - \bar{x}}{t} \end{aligned}$$

Beispiel: Ungleichungen: $S = \{x \in \mathbb{R}^n \mid g(x) \leq 0\}$

1. $m = 1$: $\{v \in \mathbb{R}^n \mid g'(\bar{x})v \leq 0\}$ falls $g(\bar{x}) = 0$.
   [hier Bildchen Ungleichungen 1]

2. $m = 2$:
   Falls $g_1(\bar{x}) = g_2(\bar{x}) = 0$, so ist $\{v \in \mathbb{R}^n \mid g_1'(\bar{x})v \leq \wedge g_2'(\bar{x})v \leq 0\}$. Ist $\bar{x}$ regulär, also exisitiert ein $v_0 \in \mathbb{R}^n \setminus \{0\}$ mit $g_i'(\bar{x})v_i < 0$ für $i = 1, 2$, so gilt sogar Tangentialkegel = Linearisierender Kegel.
   [hier Bildchen Ungleichungen 2]

Konische Hülle:
Sei $M \subset \mathbb{R}^n$ eine nichtleere Menge und sei $\bar{x} \in M$. Die konische Hülle von M in $\bar{x}$ ist definiet durch

$$\begin{aligned} M(\bar{x}) &:= \Big\{\alpha(x - \bar{x}) \mid x \in M, \alpha \geq 0\Big\} \\ &= \bigcup_{\alpha \geq 0} \alpha(M - \bar{x}) \end{aligned}$$

Falls $M = K$ konvex ist, dann ist die konische Hülle $K(\bar{x})$ konvex.

Hauptanwendungen:

1. Sei A eine $m \times n$ Matrix, $b \in \mathbb{R}^n$, $K := \{x \in \mathbb{R}^n \mid Ax = b\}$. Zeige: $K(\bar{x}) = \ker(A)$ für alle $\bar{x} \in K$.
   Beweis: Für $x, \bar{x} \in K$ gilt $Ax = A\bar{x} = b$, also $A(x - \bar{x}) = 0$.

2. Sei $K = \mathbb{R}^k_-$:
   Für $\bar{x} \in \mathbb{R}^k_-$ ist $K(\bar{x}) = \big\{x \in \mathbb{R}^k \mid y_i \leq 0 \text{ falls } \bar{x}_i = 0\big\}$

Beispiel: $k = 2, K = \mathbb{R}^2_-$: [Bild Bsp127]

1. Fall : Ist $\bar{x}_i < 0$ für $i = 1, 2$, so ist $K(\bar{x}) = \mathbb{R}^2$

2. Fall : Ist $\bar{x}_1 = 0, \bar{x}_2 < 0$, so ist $K(\bar{x}) = \{y \in \mathbb{R}^2 \mid y_1 \leq 0\}$.

3. Fall : Ist $\bar{x}_1 < 0, \bar{x}_2 = 0$, so ist $K(\bar{x}) = \{y \in \mathbb{R}^2 \mid y_2 \leq 0\}$.

4. Fall : Ist $\bar{x} = 0$, so ist $K(\bar{x}) = \mathbb{R}^2_-$.

Bemerkung:
Wir merken uns: Wenn wir im $\mathbb{R}^k_-$ eine Komponente von $\bar{x} = 0$ (aktiv) haben, dann haben wir die Einschränkung, dass die zugehörige Komponente in der konischen Hülle kleiner gleich Null ist.

Anwendung auf Situation $g : \mathbb{R}^n \to \mathbb{R}^m, K = \mathbb{R}^k_- \times \{0_{m-k}$:
Dann haben wir $S = \{x \in \mathbb{R}^n \mid g(x) \in K\}$ als ggf. koordinatenfreie Darstellung. Dann ist die konische Hülle gegeben durch:

$$K\big(g(\bar{x})\big) = \Big\{y \in \mathbb{R}^m \;\Big|\; y_i \leq 0 \text{ falls } g_i(\bar{x}) = 0 \text{ für } i = 1, \ldots, k,\; y_i = 0 \text{ für } i = (k+1, \ldots, m)\Big\}$$

2.3.2 Allgemein

2.3.1 Definition: (Tangentialkegel)
Sei $S \subset \mathbb{R}^n$ nichtleer und sei $\bar{x} \in S$. Der Tangentialkegel $T(S, \bar{x})$ ist definiert durch

$$T(S,\bar{x}) := \left\{ v \in \mathbb{R}^n \;\middle|\; \exists x_i \in S\ \exists t_i > 0 : \lim_{i\to\infty} t_i = 0 \wedge \lim_{i\to\infty} \frac{x_i - \bar{x}}{t_i} = v \right\}$$

Umformulierung von $v \in T(S, \bar{x})$:

$$x_i = \bar{x} + t_i v + r_i, \quad r_i \in \mathbb{R}^n, \lim_{i\to\infty} \frac{r_i}{t_i} = 0 \tag{2.3.2}$$

Insbesondere gilt

$$\lim_{i\to\infty} x_i = \bar{x}, \quad \lim_{i\to\infty} \frac{||x_i - \bar{x}||}{t_i} = ||v||$$

2.3.3 Lemma:
Für jede nichtleere Menge $S \subset \mathbb{R}^n$ und $\bar{x} \in S$ ist der Tangentialkegel $T(S, \bar{x})$ ein abgeschlossener Kegel mit Spitze im Ursprung.

Beweisidee:
Offensichtlich ist $0 \in T(S, \bar{x})$ und $v \in T(S, \bar{x})$ ist auch $\alpha v \in T(S, \bar{x})$ für $\alpha > 0$. Für die Abgeschlossenheit betrachten wir eine Folge $(v_k)_{k\in\mathbb{N}}$ mit $v = \lim_{k\to\infty} v_k$. Zu jedem v_k gibt es nun eine Folge $\{x_{i,k})\} \subset S, t_{i,k} > 0$ mit

$$v_k = \lim_{i\to\infty} \frac{x_{i,k} - \bar{x}}{t_{i,k}}, \quad \lim_{i\to\infty} t_{i,k} = 0$$

Wir betrachten nun die Diagonalfolge für die mit geeigneten ε, δ gilt:

$$v = \lim_{i\to\infty} \frac{x_{k,k} - \bar{x}}{t_{k,k}} \in T(S, \bar{x})$$

2.3.4 Lemma:
Sei $K \subset \mathbb{R}^n$ konvex und $\bar{x} \in K$. Dann ist der Tangentialkegel $T(K, \bar{x})$ der Abschluss der konischen Hülle von K, in $\bar{x}$, d.h.

$$T(K,\bar{x}) = \overline{K(\bar{x})} = \operatorname{cl}\left(\bigcup_{\alpha \geq 0} \alpha(K - \bar{x}) \right)$$

Beweis:
„$\overline{K(\bar{x})} \subset T(K, \bar{x})$":

Sei $x \in K$ beliebig und $\alpha > 0$. Zu zeigen ist nun, dass $\alpha(x - \bar{x}) \in T(K, \bar{x})$: Da K konvex ist und $x \in K$ gilt, folgt

$$\begin{aligned} x_i &= \bar{x} + \frac{1}{i}\alpha(x - \bar{x}) \in K \text{ für } i > \alpha \\ &\Rightarrow \frac{x_i - \bar{x}}{\frac{1}{i}} = \alpha(x - \bar{x}) \\ &\Rightarrow \alpha(x - \bar{x}) = v = \lim_{i\to\infty} \frac{x_i - \bar{x}}{\frac{1}{i}} \in T(K, \bar{x}) \end{aligned}$$

Mit der Abgeschlossenheit von $T(K, \bar{x})$ folgt nun $\overline{K(\bar{x})} \subset T(K, \bar{x})$.

„$T(K,\bar{x}) \subset \overline{K(\bar{x})}$":

Zu $v \in T(K,\bar{x})$ gibt es $x_i \in K$, $t_i > 0$ mit $t_i \to 0$ und $x_i = \bar{x} + t_i v + r_i$, wobei $\lim_{i\to\infty} \frac{r_i}{t_i} = 0$. Daraus folgt nun

$$\underbrace{\frac{x_i - \bar{x}}{t_i}}_{\in K(\bar{x})} = v + \frac{r_i}{t_i} \quad \Rightarrow \quad v = \lim_{i\to\infty} \frac{x_i - \bar{x}}{t_i} \in \overline{K(\bar{x})}$$

□

Lemma:

Sei $D \subset \mathbb{R}^n$ offen, $S \subset D, \bar{x} \in \S$. Sei weiter $h : D \to \mathbb{R}$ eine beliebige in $\bar{x}$ differenzierbare Abbildung und $v \in T(S,\bar{x})$. Dann gilt:

$$v = \lim_{i\to\infty} \frac{x_i - \bar{x}}{t_i} \Rightarrow h'(\bar{x})v = \lim_{i\to\infty} \frac{h(x_i) - h(\bar{x})}{t_i}$$

Beweis:

Mit *2.3.2* folgt

$$\begin{aligned} h(x_i) &= h(\bar{x} + t_i v + r_i) = H(\bar{x}) + t_i h'(\bar{x})v + h'(\bar{x})r_i + o(t_i) \\ &= h(\bar{x}) + t_i h'(\bar{x})v + o(t_i) \end{aligned}$$

Also folgt direkt die Behauptung

$$\lim_{i\to\infty} \frac{h(x_i) - h(\bar{x})}{t_i} = h'(\bar{x})v + \lim_{i\to\infty} \frac{o(t_i)}{t_i} = h'(\bar{x})v$$

□

2.3.5 Satz: *Variationsungleichung*

Sei $D \subset \mathbb{R}^n$ offen, $S \subset D$, $f : D \to \mathbb{R}$, $\bar{x} \in S$. Weiter sei die Funktion f differenzierbar in $\bar{x}$.

Ist $\bar{x} \in S$ eine lokale Minimalstelle des Problems

$$\min \left\{ f(x) \mid x \in S \right\}$$

dann gilt die Variationsungleichung

$$\forall v \in T(S,\bar{x}) : f'(\bar{x})v \geq 0$$

Bemerkung:

Anschaulich sagt dieser Satz soviel wie: „Wenn man in $\bar{x}$ ein Minimum hat, so kann die Funktion f in Richtung v nur wachsen." Dabei ist die grundlegende Theorie sogar so allgemein, dass sie ohne weitere Anstrengungen auf Banach- und Funktionenräume erweitert werden kann.

Beweis:

Sei $v \in T(S,\bar{x})$ mit $v = \lim_{i\to\infty} \frac{x_i - \bar{x}}{t_i}$ mit $x_i \in S, t_i > 0, t_i \to 0$. Nach dem vorigen Lemma gilt

$$f'(\bar{x})v = \lim_{i\to\infty} \underbrace{\frac{f(x_i) - f(\bar{x})}{t_i}}_{\geq 0 \text{ da } \bar{x} \text{ Minimum}} \geq 0$$

□

Ziel:

Einfache Berechnung von $T(S,\bar{x})$ für die zulässige Menge

$$S = \left\{ x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \ldots, k, \quad g_i(x) = 0, i = k+1, \ldots, m \right\}$$

Dazu benötigen wir der Linearisierende Kegel

$$L(S, \bar{x}) := \left\{ v \in \mathbb{R}^n \;\middle|\; \begin{array}{l} g_i'(\bar{x})v \leq 0 \text{ für } i \in I(\bar{x}) \\ g_i'(\bar{x})v = 0 \text{ für } i = k+1, \dots m \end{array} \right\} \tag{2.3.6}$$

[hier noch ein schickes Bildchen zu linearisierenden Kegel]

Umformulierung:
Sei $K = R_-^k \times \{0_{m-k}\} \not\subset \mathbb{R}^m$ ein Kegel. Dann ist

$$K\big(g(\bar{x})\big) = \big\{ y \in \mathbb{R}^m \mid y_i \leq 0, i \in I(\bar{x}), \quad y_i = 0, i = k+1, \dots m \big\}$$

$$L(S, \bar{x}) = \big\{ v \in \mathbb{R}^n \mid g'(\bar{x})v \in K(g(\bar{x})) \big\}$$

2.3.7 Lemma:
Es gilt stets $T(S, \bar{x}) \subseteq L(S, \bar{x})$.

Beweis:
Sei $v \in T(S, \bar{x})$ mit $x_i = \bar{x} + t_i v + r_i, r_i = o(t_i)$. Für $j \in I(\bar{x})$ gilt nun:

$$g_j'(\bar{x})v = \lim_{i \to \infty} \frac{g_j(x_i) - g_j(\bar{x})}{t_i} = \lim_{i \to \infty} \underbrace{\frac{g_j(x_i)}{t_i}}_{\leq 0 \text{ da } x_i \in S} \leq 0$$

Für $j = k+1, \dots, m$ gilt:

$$g_j'(\bar{x})v = \lim_{i \to \infty} \underbrace{\frac{g_j(x_i) - g_j(\bar{x})}{t_i}}_{= 0 \text{ da } x_i \in S} = 0$$

□

Bemerkung:
Achtung: Es kann durchaus $T(S, \bar{x}) \subsetneqq L(S; \bar{x})$ gelten.

Beispiel:

1. Sei $g(x) = x^2, x \in \mathbb{R}, S = \big\{ x \in \mathbb{R} \mid g(x) = x^2 = 0 \big\} = \{0\}$. Für $\bar{x} = 0$ ist dann $T(S, \bar{x}) = \{0\}$ und $L(S, \bar{x}) = \big\{ v \in \mathbb{R} \mid g'(\bar{x})v = 2 \cdot 0 \cdot v = 0 \big\} = \mathbb{R}$ wegen $g'(\bar{x}) = 0$.

   In diesem Fall ist $\bar{x}$ nicht regulär = normal.

2. Diesmal nehmen wir die bessere Funktion $g(x) = x, S = \big\{ x \in \mathbb{R} \mid g(x) = x = 0 \big\} = \{0\}$. Nun ist $L(S, \bar{x}) = \big\{ v \in \mathbb{R} \mid g'(\bar{x})v = 1 \cdot v = 0 \big\} = \{0\} = T(S, \bar{x})$.

   Hier ist nun $\bar{x} = 0$ regulär = normal, da $g'(\bar{x}) = 1 \neq 0$.

Ziel:
Finde Regularitätsbedingungen in $\bar{x} \in S$, sodass gilt $T(S, \bar{x}) = L(S, \bar{x})$.

Wir betrachten zunächst den Spezialfall von Gleichungen, also $S = \{ x \in \mathbb{R}^n \mid g(x) = 0_m \}, g : D \to \mathbb{R}^m$:

2.3.8 Definition: (Normalität, Regularität)
Ein Punkt $\bar{x} \in S$ heißt regulär (normal), wenn die Gradienten $g_i'(\bar{x}), i = 1, \dots, m$ linear unabhängig sind.

2.3.9 Alternative Bedingung:
Die $m \times m$-Matrix $g'(\bar{x})g'(\bar{x})^T$ ist regulär, sogar positiv definit (äquivalent, da $\text{im}(g'(\bar{x})) = \mathbb{R}^m$).

2.3.10 Lemma:
Es gilt genau eine der beiden folgenden Aussagen:

(i) $\bar{x} \in S$ ist regulär

(ii) Es gibt einen Zeilenvektor $\lambda \in \mathbb{R}^m \setminus \{0\}$ mit

$$\lambda g'(\bar{x}) = \sum_{i=1}^{m} \lambda_i g_i'(\bar{x}) = 0$$

2.3.11 Satz:
Sei $S = \{x \in \mathbb{R}^n \mid g(x) = 0\}$ und sei $\bar{x} \in S$ regulär.

(i) Zu $v \in \mathbb{R}^n$ mit $g'(\bar{x})v = 0$, d.h. $v \in L(S, \bar{x})$, gibt es $\alpha > 0$ und eine Kurve $x : [-\alpha, \alpha] \to S$ mit

$$\forall t \in [-\alpha, \alpha] : g(x(t)) = 0, \quad x(0) = \bar{x}, \quad v = \lim_{i \to 0} \frac{x(t) - \bar{x}}{t}$$

(ii) $T(S, \bar{x}) = L(S, \bar{x}) = \big\{v \in \mathbb{R}^n \mid g'(\bar{x})v = 0\big\}$.

Beweis:

zu (i):

Betrachte die Abbildung $F : \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^m$ mit $F(y,t) = g\big(\bar{x} + tv + g'(\bar{x})^T y\big)$. Ziel ist nun die Bestimmung von $y(t)$ mit $F(y(t), t) = 0$.

Es gilt $F(0_m, 0) = g(\bar{x}) = 0$, und die $m \times m$-Matrix

$$\frac{\partial F}{\partial y}(0_m, 0) = g'(\bar{x})g'(\bar{x})^T$$

ist regulär, denn $\bar{x} \in S$ ist regulär.

Nach dem Satz über Implizite Funktionen gibt es ein $\alpha > 0$ und eine Kurve $y : [-\alpha, \alpha) \to \mathbb{R}^m$ mit $y(0) = 0$ und $F(y(t), t) = 0$ für alle $t \in [-\alpha, \alpha]$. y ist differenzierbar in $t = 0$ mit $\dot{y}(0) = \lim_{t \to 0} \frac{y(t)}{t}$.
Schließlich ist $\dot{y}(0) = 0$, denn

$$F(y(t), t) = 0 \;\Rightarrow\; 0 = \frac{d}{dt} F(y(t), t)\Big|_{t=0} = \underbrace{g'(\bar{x})v}_{=0} + g'(\bar{x})g'(\bar{x})^T \dot{y}(0)$$

Da $g'(\bar{x})g'(\bar{x})^T$ regulär ist, folgt $\dot{y}(0) = 0$. Setze

$$x(t) := \bar{x} + tv + g'(\bar{x})^T y(t),\; x(0) = \bar{x},\; v = \dot{x}(0) = \lim_{i \to 0} \frac{x(t) - \bar{x}}{t}$$

zu (ii):

folgt direkt aus (i) mit *(2.3.1)* und $T(S, \bar{x})$ direkt aus der Limesbeziehung aus dem Satz.

□

Allgemeiner Fall:

$$S = \big\{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, k, \quad g_i(x) = 0, i = k+1, \dots, m\big\}$$

Linearisierender Kegel:

$$L(S, \bar{x}) := \left\{ v \in \mathbb{R}^n \;\middle|\; \begin{array}{l} g_i'(\bar{x})v \leq 0 \text{ für } i \in I(\bar{x}) \\ g_i'(\bar{x})v = 0 \text{ für } i = k+1, \dots m \end{array} \right\}$$

Ziel:
$T(S,\bar{x}) = L(S,\bar{x})$

2.3.12 Definition: (Regularität)
Ein Punkt $\bar{x} \in S$ heißt *regulär*, wenn die Gradienten $g_i'(\bar{x})$ für $i = k+1,\dots,m$ linear unabhängig sind und ein $v_0 \in \mathbb{R}^n$ existiert mit

$$g_i'(\bar{x})v_0 \begin{cases} < 0 & \text{für } i \in I(\bar{x}) \\ = 0 & \text{für } i = k+1,\dots,m \end{cases}$$

Umformulierung:
Sei $I(\bar{x}) = \{1,\dots,k_0\}$ mit $k_0 \le k$. Betrachte den konvexen, abgeschlossenen Kegel

$$C := \left\{ \begin{pmatrix} g_1'(\bar{x})v + r_1 \\ \vdots \\ g_{k_0}'(\bar{x})v + r_{k_0} \\ g_{k_0+1}'(\bar{x})v \\ \vdots \\ g_m'(\bar{x})v \end{pmatrix} \;\middle|\; v \in \mathbb{R}^n, r_1,\dots,r_{k_0} \ge 0 \right\}$$

Man überlegt sich, dass die Regularität von $\bar{x} \in S$ äquivalent dazu ist, dass

$$\begin{pmatrix} 0_{k_0} \\ 0_{m-k} \end{pmatrix} \in \mathring{C} \tag{2.3.13}$$

Dies ist äquivalent zur Regularitätsbedingungen von Robinson, Zowe und Kurcyusz:

$$\operatorname{im}\big(g'(\bar{x})\big) - K\big(g(\bar{x})\big) = \mathbb{R}^m, \quad K = \mathbb{R}^k_- \times \{0_{m-k}\} \tag{2.3.14}$$

wobei $K(g(\bar{x})) = \{y \in \mathbb{R}^m \mid y_i \le 0, i \in I(\bar{x}), \quad y_i = 0, i = k+1,\dots,m\}$

2.3.15 Lemma:
Es gilt genau eine der beiden folgenden Aussagen:

(i) $\bar{x} \in S$ ist regulär

(ii) Es gibt $\bar{\lambda}_i \in \mathbb{R}$ für $i \in J(\bar{x})$ mit $\big(\bar{\lambda}_i\big)_{i\in J(\bar{x})} \ne 0$ und

$$\sum_{i\in J(\bar{x})} \bar{\lambda}_i g_i'(\bar{x}) = 0 \quad \wedge \quad \bar{\lambda}_i \ge 0 \text{ für } i \in I(\bar{x})$$

Beweis:
Wenn $\bar{x} \in S$ nicht regulär ist, dann gilt $0 \notin \mathring{C}$ mit C wie in *2.3.13*. Nach dem Trennungssatz *(3.2.3)* gibt es nun ein $\bar{\lambda} = \big(\bar{\lambda}_i\big)_{i\in J(\bar{x})} \ne 0$ mit

$$\forall y \in C : 0 = \bar{\lambda} \cdot 0 \le \bar{\lambda} \cdot y$$

Setze $r_i = 0$ für $i \in I(\bar{x})$:

$$\Rightarrow \ \forall v \in \mathbb{R}^n : 0 \le \sum_{i\in J(\bar{x})} \bar{\lambda}_i g_i'(\bar{x})v$$

Da das für *alle* $v \in \mathbb{R}^n$ gilt (also insbesondere auch sowohl für positive als auch für negative), muss schon

$$\sum_{i\in J(\bar{x})} \bar{\lambda}_i g_i'(\bar{x}) = 0$$

Setze $v = 0$ und $r_i \geq 0$ beliebig:

$$0 \leq \sum_{i \in I(\bar{x})} \bar{\lambda}_i r_i \text{ für alle } r_i \geq 0, i \in I(\bar{x}) \quad \Rightarrow \quad \bar{\lambda}_i \geq 0 \text{ für } i \in I(\bar{x})$$

□

Bemerkung:
Lemma *(2.3.15.2)* sind die KKT-Bedingungen mit $\lambda_0 = 0$.

2.3.16 Satz:
Sei $\bar{x} \in S$ regulär.

(i) Zu $v_0 \in L(S, \bar{x})$ mit $g_i'(\bar{x}) v_0 < 0$ für $i \in I(\bar{x})$ gibt es $\alpha > 0$ eine Kurve $x : [0, \alpha] \to S$ mit

$$x(0) = \bar{x}, \quad v_0 = \lim_{t \searrow 0} \frac{x(t) - \bar{x}}{t}$$

(ii) Es gilt $T(S, \bar{x}) = L(S, \bar{x})$.

Beweis:

zu (i): Falls $k = m$, so setze $x(t) := \bar{x} + t v_0$.

Ist $k < m$, so gibt es nach *(2.3.11)* $\alpha > 0$ und eine Kurve $x : [-\alpha, \alpha] \to \mathbb{R}^n$ mit $g_i(x(t)) = 0$ für $i = k+1, \ldots, m$, $x(0) = \bar{x}$ und $v_0 = \lim_{t \to 0} \frac{x(t) - \bar{x}}{t}$. Außerdem gilt $g_i(x(t)) \leq 0$ für $i = 1, \ldots, k$ und $0 \leq t \leq \alpha_0$ für α_0 genügend klein. Taylorentwicklung liefert

$$g_i(x(t)) \;=\; g_i(\bar{x}) + t g_i'(\bar{x}) v_0 + o(t) \;\leq\; 0$$

für t genügend klein, denn für alle $i \in I(\bar{x})$ ist $g_i(\bar{x}) = 0, g_i'(\bar{x}) v_0 < 0$ und für $i \in \{1, \ldots, k\} \setminus I(\bar{x})$ ist $g_i(\bar{x}) < 0$.

zu (ii): Sei $v \in L(S, \bar{x})$ und v_0 wie in der Regularitätsbedingung von $\bar{x}$. Betrachte $v_\lambda = \lambda v_0 + (1 - \lambda) v$ für $\lambda \in [0, 1]$. Für $\lambda > 0$ gilt

$$\begin{aligned} g_i'(\bar{x}) v_\lambda &= g_i'(\bar{x}) \big(\lambda v_0 + (1 - \lambda) v \big) \\ &= \underbrace{\lambda g_i'(\bar{x}) v_0}_{<0} + (1 - \lambda) \underbrace{g_i'(\bar{x}) v}_{\leq 0} \quad < \quad 0 \end{aligned}$$

für $i \in I(\bar{x})$ und $g_i'(\bar{x}) v_\lambda = 0$ für $i = k+1, \ldots, m$.

Nach (i) gilt, dass $v_\lambda \in T(S, \bar{x})$ ein abgeschlossener Kegel ist.

$$\Rightarrow \quad v = \lim_{\lambda \searrow 0} v_\lambda \in T(S, \bar{x})$$

□

Kapitel 3

Anhang

3.1 Konvexe Mengen und Kegel

3.1.1 Bezeichnungen für Mengen

Sei $M \subset \mathbb{R}^n$ eine Menge.

3.1.1 Definition: (Spann)
span(M) sie der kleinsten, M enthaltende *lineare* Teilraum im $\mathbb{R}^n$.

3.1.2 Definition: (affiner Teilraum)
aff(M) sei der kleinste M enthaltende *affine* Teilraum im $\mathbb{R}^n$.

3.1.3 Definition: (topologisch Inneres)
$\text{int}(M) = \overset{\circ}{M}$ sei das *topologische Innere* von M bezüglich R^n.

3.1.4 Definition: (relatives topologisches Inneres)
M^i sei das relative topologische Innere von M bezüglich aff(M).

3.1.5 Definition: (topologischer Abschluss)
$\overline{M}$ sei der topologische Abschluss von M in $\mathbb{R}^n$.

3.1.6 Definition: (Rand)
$\partial M = \overline{M} \setminus \overset{\circ}{M}$ sei der Rand von M.

Beispiel:
Betrachte die Menge $M = \{\binom{x_1}{x_2} \in \mathbb{R}^2 \mid x_1 + x_2 = 1, x_1, x_2 \geq 0\}$.

Dann ist:

$$\begin{aligned}
\text{span}(M) &= \mathbb{R}^2 \\
\text{aff}(M) &= \{x \in \mathbb{R}^2 \mid x_1 + x_2 = 1\} \\
\text{int} &= \overset{\circ}{M} = \emptyset \\
M^i &= \{x \in \mathbb{R}^2 \mid x_1 + x_2 = 1, x_1 > 0, x_2 > 0\} \\
\overline{M} &= M \\
\partial M &= M
\end{aligned}$$

3.1.7 Definition: (konvexe Menge)

(i) Eine Teilmenge $K \subset \mathbb{R}^n$ heißt *konvex*, wenn für alle $x, y \in K$ und $0 \leq \alpha \leq 1$ gilt:

$$\alpha x + (1-\alpha)y \in K \quad \text{bzw.} \quad x + \alpha(y-x) \in K$$

(ii) Für eine Menge $M \subset \mathbb{R}^n$ heißt

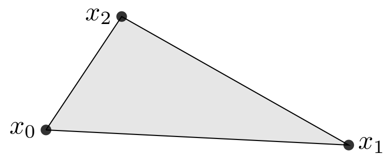
$$\operatorname{co}(M) := \bigcap_{K \subset \mathbb{R}^n} \big\{K \mid M \subset K, K \text{ konvex}\big\}$$

die *konvexe Hülle* von M. Dies ist die kleinste M enthaltende konvexe Menge.

(iii) Für gegebene Vektoren $x^0, x^0, \ldots, x^n \in \mathbb{R}^m, n < m$ heißt

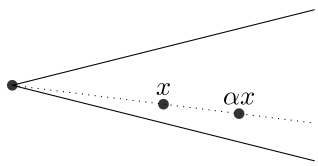
$$\operatorname{co}\left(x^0, \ldots, x^n\right) = \operatorname{co}\left(\left\{x^0, \ldots, x^n\right\}\right)$$

das von $x^0, \ldots, x^n$ *aufgespannte Simplex.*



Sind die Vetoren $x^1 - x^0, \ldots, x^n - x^0$ linear unabhängig, so heißt das Simplex *nichtentartet.*

(iv) Eine Teilmenge $K \subset \mathbb{R}^n$ heißt Kegel (mit Spitze im Ursprung), wenn für alle $x \in K$ der positive Halbstrahl $\alpha x, \alpha \geq 0$ in K liegt.



Bemerkung:
Auch ein Halbraum ist ein Kegel.

3.1.8 Lemma:

(i) Der Durchschnitt einer Familie konvexer Mengen ist konvex (vgl. *(3.1.7.2)*).

(ii) Sind $K_1, K_2 \subset \mathbb{R}^n$ konvex, dann ist für $\alpha, \beta \in \mathbb{R}$ die Menge

$$\alpha K_1 + \beta K_2 = \{\alpha x + \beta y \mid x \in K_1, y \in K_2\}$$

konvex.

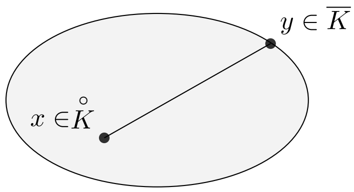
(iii) Die konvexe Hülle $\mathrm{co}(M)$ einer Menge $M \subset \mathbb{R}^n$ ist gegeben durch

$$\mathrm{co}(M) = \left\{ \sum_{i=1}^{k} \alpha_i x_i \;\middle|\; x_i \in M, \alpha_i \geq 0, \sum_{i=1}^{k} \alpha_i = 1, k \in \mathbb{N} \right\}$$

(iv) Ist $K \subset \mathbb{R}^n$ konvex, dann sind die Mengen $\overline{K}, \mathring{K}$ und K^i konvex.

(v) Sei $K \subset \mathbb{R}^n$ konvex und seien $x \in \mathring{K}$ und $y \in \overline{K}$. Dann gilt für $0 \leq \alpha < 1$ dass

$$(1-\alpha)x + \alpha y \in \mathring{K}$$



(vi) Für eine konvexe Menge K mit $\mathring{K} \neq \emptyset$ gilt: $\mathrm{int}\left(\overline{K}\right) = \mathring{\overline{K}} = \mathring{K}$.

(vii) Eine Teilmenge $K \subset \mathbb{R}^n$ ist genau dann ein konvexer Kegel, wenn

$$\forall \alpha, \beta > 0 : \alpha K + \beta K = \{\alpha x + \beta y \mid x, y \in K\} \subset K$$

3.1.9 Lemma:
Ist $K \subset \mathbb{R}^n$ eine nichtleere konvexe Menge, dann ist $K^i \neq \emptyset$.

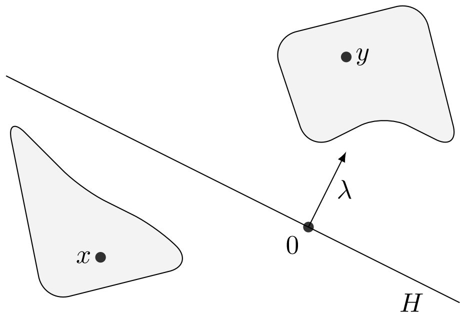
3.2 Trennbarkeit konvexer Mengen

3.2.1 Definition: (Trennbarkeit konvexer Mengen)
Zwei Teilmengen $K_1, K_2 \subset \mathbb{R}^n$ heißen *trennbar*, wenn es einen Zeilenvektor $\lambda \in \mathbb{R}^n \setminus \{0\}$ und $\alpha \in \mathbb{R}$ gibt mit:

$$\forall x \in K_1 \forall y \in K_2 : \lambda x \leq \alpha \leq \lambda y$$

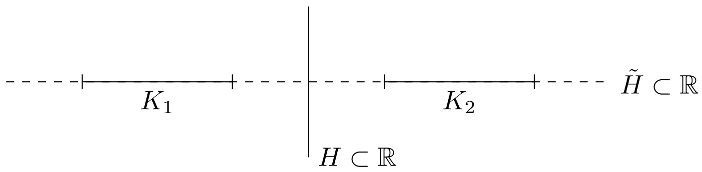
Die Hyperebene $H := \{z \in \mathbb{R}^n \mid \lambda z = \alpha\}$ heißt trennende Hyperebene.

Beispiel:



3.2.2 Definition: (echte Trennbarkeit)
Die Mengen K_1, K_2 heißen *echt trennbar* durch die Hyperebene H, falls $K_1 \cup K_2$ nicht nicht in H enthalten ist.

Beispiel:
Die Mengen K_1 und K_2 sind echt trennbar durch H jedoch nicht echt trennbar durch $\tilde{H}$:



Bemerkung:
In der Definition *(3.2.1)* hat man die folgende Situation: Die Hyperebene $H = \{z \in \mathbb{R}^n \mid \lambda z = \alpha\}$ definiert die Halbräume $H^+ := \{z \in \mathbb{R}^n \mid \lambda z \geq \alpha\}$ und $H^- := \{z \in \mathbb{R}^n \mid \lambda z \leq \alpha\}$. Dann gilt $K_1 \subset H^-$ und $K - 2 \subset H^+$.

3.2.3 Satz: *Trennungssatz*
Sei $K \subset \mathbb{R}^n$ eine nichtleere, konvexe Menge und sei $y \notin \mathring{K}$. Dann sind die Mengen $\{y\}$ und K trennbar, d.h. es gibt einen Zeilenvektor $\lambda \in \mathbb{R}^n \emptyset \{0\}$ mit $\lambda y \leq \lambda x$ für alle $x \in K$.

Ist $\mathring{K} \neq \emptyset$, so sind $\{y\}$ und $\mathring{K}$ echt trennbar, d.h. $\lambda y < \lambda x$ für alle $x \in \mathring{K}$.

Beweis:

1. Fall: $y \notin \overline{K} = \mathrm{cl}(K)$
Sei $d := \inf_{\bar{x} \in \overline{K}} ||y - x||_2$. Die Funktion $f(x) := ||y - x||_2, x \in \mathbb{R}^n$ ist stetig und nimmt daher auf der kompakten Menge $\overline{K} \cap \{x \in \mathbb{R}^n \mid ||x - y||_2 \leq 2d\}$ ein globales Minimum $x_0 \in \overline{K}$ an. Es gilt

$$d = ||y - x_0||_2 = \min_{x \in \overline{K}} ||y - x||_2 > 0$$

da $y \notin \overline{K}$.
Beh.: $\lambda := (x_0 - y)^T \neq 0$ erfüllt die Behauptung des Satzes.
Sei $x \in K$ beliebig. Wegen der Konvexität von K gilt $x_0 s(x - x_0) \in \overline{K}$ für $0 \leq s \leq 1$. Es folgt

$$\begin{array}{rrcl}
\forall s \in [0,1]: & ||x_0 + s(x - x_0) - y||_2^2 & \geq & ||x_0 - y||_2^2 \\
\Rightarrow & ||x_0 - y||_2^2 + 2s(x_0 - y)^T(x - x_0) + s^2||x - x_0||_2^2 & \geq & ||x_0 - y||_2^2 \\
\Rightarrow & 2s(x_0 - y)^T(x - x_0) + s^2||x - x_0||_2^2 & \geq & 0
\end{array}$$

Für $s > 0$ folgt direkt:

$$\forall s \in (0,1] : 2(x_0 - y)^T(x - x_0) + s||x - x_0||_2^2 \ \geq \ 0$$

Andernfalls betrachten wir den Grenzwert $s \searrow 0$ und erhalten die Variationsungleichung $2(x_0 - y)^T(x - x_0) \geq 0$.

Also gilt

$$\forall x \in K : (x_0 - y)^T(x - x_0) \geq 0$$

Wir setzen nun $\lambda := (x_0 - y)^T \neq 0$ und $d := ||x_0 - y||_2^2 = ||\lambda||^2 > 0$ und erhalten

$$\forall x \in K : \ \lambda x \geq \lambda x_0 = \lambda(x_0 - y + x) = \underbrace{\lambda(x_0 - y)}_{=d^2>0} + \lambda y = d^2 + \lambda y > \lambda y$$

Also ist $H := \{z \in \mathbb{R}^n \mid \lambda z = \lambda x_0 =: \alpha\}$ eine Hyperebene, die $\{y\}$ und K trennt.

2. Fall: $y \in \partial K$

Sei $(y_k)_{k\in\mathbb{N}}$ eine Folge mit $\forall k : y_k \notin \overline{K}$ und $\lim_{k\to\infty} y_k = y$. Nach Fall 1 gibt es eine Folge $(\lambda_k)_{k\in\mathbb{N}}, \lambda k \in \mathbb{R}^n \setminus \{0\}$ mit

$$\Big(\forall x \in K : \lambda_k y_k < \lambda_k x\Big) \implies \left(\forall x \in K : \frac{\lambda_k}{||\lambda_k||_2} y_k < \frac{\lambda_k}{||\lambda_k||_2 x}\right)$$

Ohne Einschränkung sei $||\lambda_k||_2 = 1$. Da die Einheitskugel kompakt ist, existiert ein $\lambda := \lim_{k\to\infty} \lambda_k$ mit $||\lambda||_2 = 1$. Insbesondere ist $\lambda \neq 0$. Im Grenzübergang gilt nun

$$\forall x \in K : \lambda y = \lim_{k\to\infty} \lambda_k y_k \leq \lim_{k\to\infty} \lambda_k x = \lambda x$$

Für $x \in \overset{\circ}{K}$ gilt hier sogar echte Ungleichheit.

Zusammen folgt $\forall x \in \overset{\circ}{K}$: $\lambda y < \lambda x$ □

3.2.4 Korollar: *Trennungssatz für zwei konvexe Mengen*
Seien $K_1, K_2 \subset \mathbb{R}^n$ nichtleere, konvexe Mengen mit $K_1 \cap K_2 = \emptyset$. Dann sind K_1 und K_2 trennbar, d.h., es gibt $\lambda \in \mathbb{R}^n \setminus \{0\}$ und $\alpha \in \mathbb{R}$ mit

$$\forall x \in K_1, y \in K_2 : \lambda x \leq \alpha \leq \lambda y$$

Beweis:
siehe Übung. □

3.2.5 Satz: *Trennungssatz für Kegel*
Sei $K \subset \mathbb{R}^n$ ein nichtleerer, abgeschlossener und konvexer Kegel. Zu $y \notin K$ gibt es dann ein $\lambda \in \mathbb{R}^n \setminus \{0\}$ mit $\forall x \in K : \lambda y < 0 \leq \lambda x$.

Beweis:
Wegen $K = \overline{K}$ gibt es nach dem ersten Fall von *(3.2.3)* einen Zeilenvektor $\lambda \in \mathbb{R}^n \setminus \{0\}$ mit $\forall x \in K : \lambda y < \lambda x$. Aus $0 \in K$ folgt $\lambda y < \lambda 0 = 0$.

Angenommen es existiere ein $x \in K$ mit $\lambda x < 0$ dann wäre $\forall \alpha \geq 0$ auch

$$\lambda(\alpha x_0) = \alpha(\lambda x_0) \overset{\alpha\to\infty}{\longrightarrow} -\infty$$

Dies ist ein Widerspruch zu Beschränktheit von x durch λy nach unten. Also gilt.

$$\forall x \in K : \lambda y < 0 \leq \lambda x$$

□

3.2.6 Satz: *Lemma von Farkas*
Sei B eine $k \times n$ Matrix und $d \in \mathbb{R}^k$. Dann gilt *genau eine* der folgenden Aussagen:

(i) Das System $Bx = d, x \geq 0$ hat eine Lösung $x \in \mathbb{R}^n$

(ii) Das System $\lambda B \geq 0, \lambda d < 0$ hat eine Lösung $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k$.

Beweis:
Der Kegel $K = \{Bx \mid x \in \mathbb{R}^n, x \geq 0\} \subset \mathbb{R}^k$ ist abgeschlossen, konvex und nicht leer. Also gilt entweder $d \in K$ oder $d \notin K$.

$d \in K$ ist äquivalent zu $Bx = d$ mit $x \geq 0$ also Aussage (i). Für $d \notin K$ gibt es nach *(3.2.5)* ein $\lambda \in \mathbb{R}^k \setminus \{0\}$ mit $\forall z \in K : \lambda d < 0 \leq \lambda z$. Also auch

$$\forall x \geq 0 \in K : 0 \leq \lambda z = \lambda Bx = (\lambda B)x$$

$\square$

3.3 Extremalpunkte, Ecken

3.3.1 Definition: (Extremalpunkt)
Sei $K \subset \mathbb{R}^n$ konvex. Ein Punkt $x \in K$ heißt *Extremalpunkt* oder *Ecke*, wenn es nicht zwei von x verschiedene Punkte $y, z \in K,\ y, z \neq x$ und $0 < \alpha < 1$ gibt mit

$$x = \alpha y + (1 - \alpha)z$$

Alternative Definition:
Aus der Darstellung $x = \alpha y + (1 - \alpha)z$ mit $y, z \in K$ und $0 < \alpha < 1$ folgt bereits $x = y = z$.

3.3.2 Satz: *(ohne Beweis)*
Eine konvexe und kompakte Menge $K \subset \mathbb{R}^n$ ist der Abschluss der konvexen Hülle seiner Extremalpunkte.