5b. Lineare Nebenbedingungen
Ungleichungen

Optimierung SoSe 2020

Dr. Alexey Agaltsov

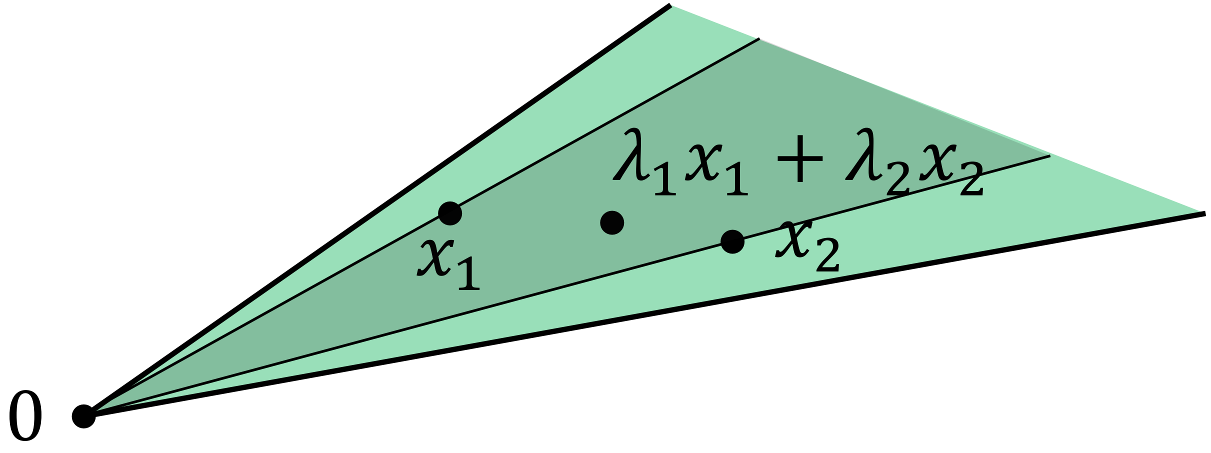
Plan

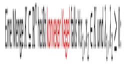
- Konvexe Kegel und Lemma von Farkas
- KKT-Bedingungen

Konvexer Kegel

Eine Menge $K \subseteq \mathbb{R}^n$ heißt konvexer Kegel falls $\forall x_1, x_2 \in K$ und $\lambda_1, \lambda_2 \geq 0$:

$$\lambda_1 x_1 + \lambda_2 x_2 \in K$$





Konische Einhüllende

- Als konische Kombination von $x_1, \dots, x_k$ bezeichnen wir:

$$v = \lambda_1 x_1 + \cdots + \lambda_k x_k \quad \text{mit } \lambda_1, \dots, \lambda_k \geq 0$$

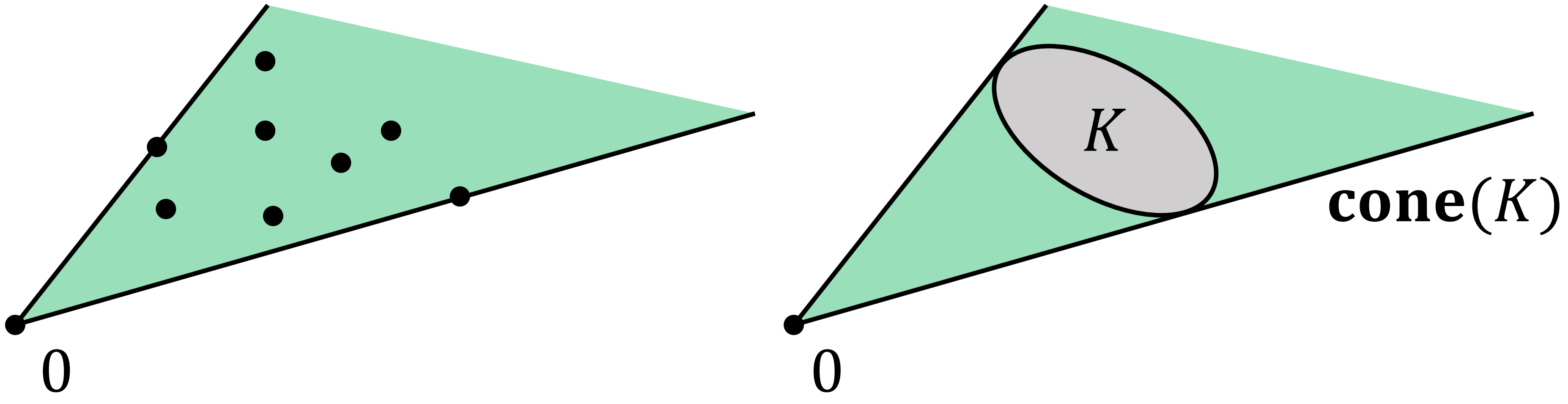
- K ist einer konvexer Kegel $\Leftrightarrow$ K enthält alle konischen Kombinationen seiner Elemente

- Als konische Einhüllende von K bezeichnen wir:

$$\mathbf{cone}(K) = \{\lambda_1 x_1 + \cdots + \lambda_k x_k : x_i \in K, \lambda_i \geq 0, i = 1, \dots, k, k \geq 0\}$$

Der minimale konvexe Kegel, der K enthält

Konische Einhüllende



Aufgabe 5.7. Konvexe Kegel

Zeigen Sie, dass die folgenden Mengen konvexe Kegel sind:

1. Positiver Orthant $\mathbb{R}^n_{\geq 0} = \{x \in \mathbb{R}^n : x \geq 0\}$
2. Positiv definiter (bzw. semidefiniter) Kegel $\mathbb{S}^n_{\succ}$ (bzw. $\mathbb{S}^n_{\succcurlyeq}$)
3. Norm-Kegel $\{(x, t) \in \mathbb{R}^n \times \mathbb{R}:\ \|x\| \leq t\}$ für eine beliebige Norm $\|\cdot\|$

Satz 5.8. Satz von Carathéodory

Angenommen:

- $K \subseteq \mathbb{R}^n$ eine nichtleere Menge und $x \in \mathbf{cone}(K)$
- $x = \sum_{i=1}^{k} \lambda_i x_i$ mit $\lambda_i \geq 0, x_i \in K$ ist eine Darstellung von x mit minimalem k

Dann sind $x_1, \dots, x_k$ linear unabhängig

Beweis

$$x = \sum_{i=1}^{k} \lambda_i x_i \text{ mit } 0 < \lambda_1 \leq \cdots \leq \lambda_k; x_i \in K$$

Angenommen, $x_1, . . , x_k$ sind linear abhängig:

$$\exists \mu_1, \ldots, \mu_1 \in \mathbb{R}: \ \sum_{i=1}^{k} \mu_i x_k = 0$$

und o.B.d.A. $\mu_1 \leq \cdots \leq \mu_k$ mit $\mu_k = \lambda_1$

$$x = \sum_{i=2}^{k} \underbrace{(\lambda_i - \mu_i)}_{\geq 0} x_k$$

Widerspruch zur Annahme, dass k Minimal ist

$\square$

Endlich erzeugter Kegel

- Ein konvexer Kegel $K \subseteq \mathbb{R}^n$ heißt endlich erzeugt falls $\exists x_1, \dots, x_k \in K$:

$$K = \mathbf{cone}\{x_1, \dots, x_k\}$$

- Beispiel:

$$K = \{Ax : x \in \mathbb{R}^n_{\geq 0}\},\ A \in \mathbb{R}^{m \times n}$$

$$= \mathbf{cone}\{a_1, \dots, a_n\}$$

Spalten von A

Lemma 5.9. Endlich erzeugter Kegel ist abgeschlossen

Sei $K = \mathbf{cone}\{x_1, \ldots, x_k\}$ mit $x_1, \ldots, x_k \in \mathbb{R}^n$

Dann ist K abgeschlossen

Beweis

$$I = \{S \subseteq \{1, \dots, k\}: \ \{x_i\}_{i \in S} \text{ sind linear unabhängig}\}$$

Satz 5.8 von Carathéodory

$$K = \bigcup_{S \subseteq I} K_S$$

$$K_S = \{\textstyle\sum_{i \in S} \lambda_i x_i : \lambda_i \geq 0 \ \forall i \in S\} \cong \mathbb{R}_{\geq 0}^{|S|}$$

$$\Longrightarrow K \cong \bigcup_{S \subseteq I} \mathbb{R}_{\geq 0}^{|S|}, \quad |I| < \infty$$

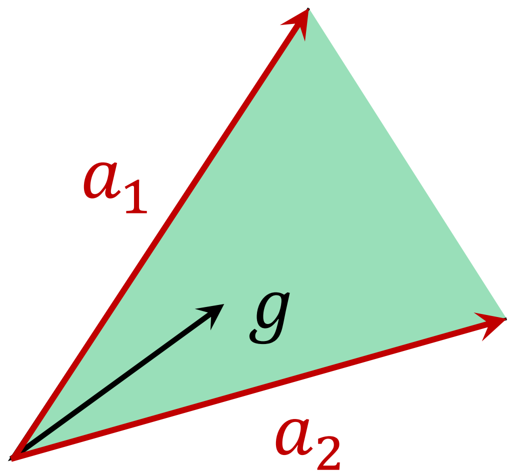
$\mathbb{R}_{\geq 0}^{|S|}$ ist abgeschlossen

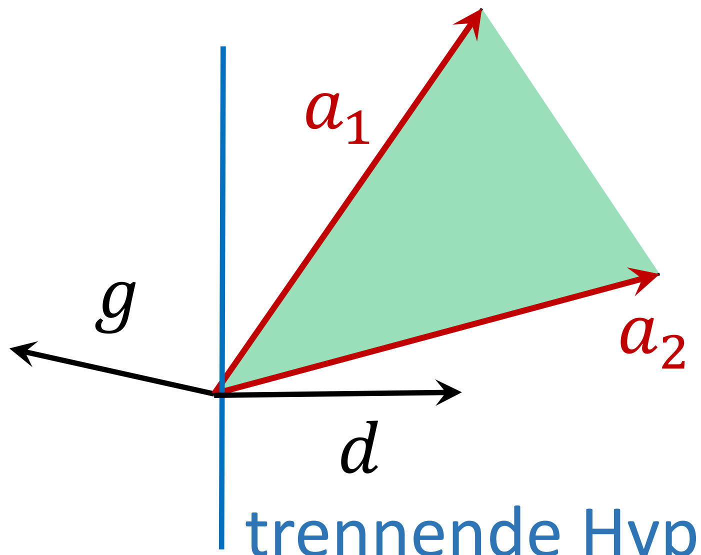
K ist abgeschlossen

$\square$

Lemma 5.10. Lemma von Farkas

Seien $a_1, \dots, a_k \in \mathbb{R}^n$ und $g \in \mathbb{R}^n$. Es gilt genau eine der folgenden Aussagen:





1. $g \in \mathbf{cone}\{a_1, \dots, a_k\}$
2. $\exists d \in \mathbb{R}^n$ so, dass $a_i^T d \geq 0 \ \forall i$, $g^T d < 0$

Beweis: $1 \Rightarrow \neg 2$

Sei $g \in \mathbf{cone}\{a_1, \dots, a_k\}$

Sei $d \in \mathbb{R}^n$: $a_1^T d \geq 0, \dots a_k^T d \geq 0$

$\exists \mu_1 \geq 0, \dots, \mu_k \geq 0$: $\quad g = \mu_1 a_1 + \cdots + \mu_k a_k$

$g^T d = \mu_1 a_1^T d + \cdots + \mu_k a_k^T d \geq 0$ Multipliziere mit d

Beweis: $\neg 1 \Longrightarrow 2$

Sei $N = \mathbf{cone}\{a_1, \dots, a_k\}$ und $g \notin N$

N ist konvex, abgeschlossen nach Lemma 5.9

Projektionssatz 2.8

$$g_0 \coloneqq \operatorname{argmin}_{x \in N} \|g - h\|_2$$

$$-d^T(h - g_0) \leq 0 \;\; \forall h \in N \;\; \text{für } d = g_0 - g$$

$$h = 2g_0 \Longrightarrow -d^T g_0 \leq 0$$

$$h = \frac{1}{2} g_0 \Longrightarrow d^T g_0 \leq 0$$

$$\Longrightarrow d^T g_0 = 0$$

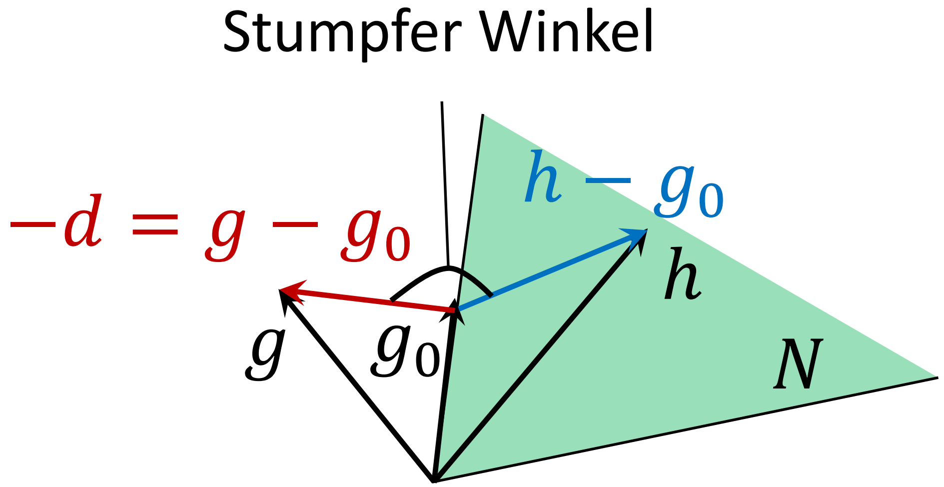
$$-d^T h \leq 0 \;\; \forall h \in N$$

$h = a_i, \dots, a_k$

$$d^T a_1 \geq 0, \dots, d^T a_k \geq 0$$

$$d^T g = d^T(g_0 - d) = -\|d\|_2^2 < 0$$

$\square$



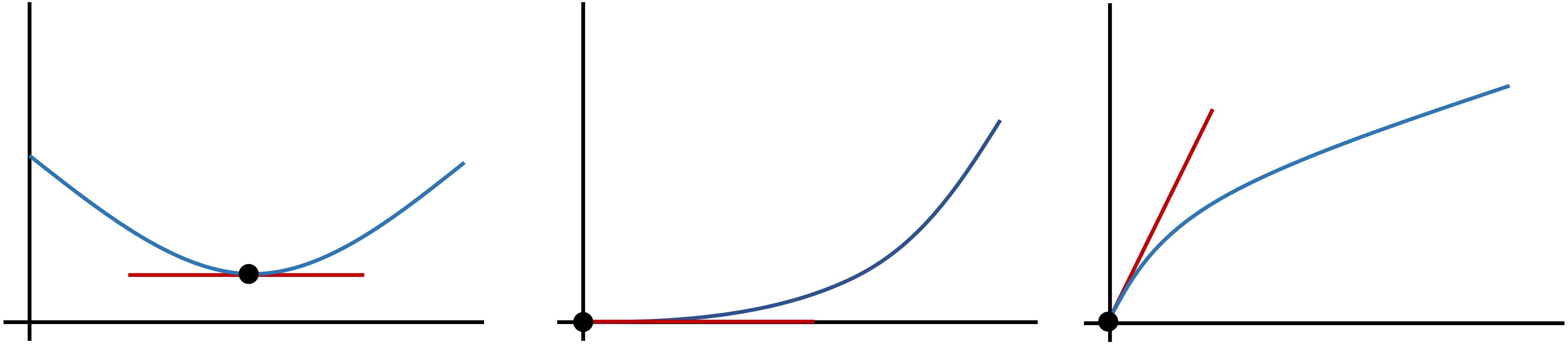
Plan

- Konvexe Kegel und Lemma von Farkas
- KKT-Bedingungen

Beispiel

$$\begin{aligned} &\text{Minimiere} \quad f(x), \\ &\text{u.d.N.} \quad x \geq 0 \\ &f \in C^1(\mathbb{R}) \end{aligned}$$

Ist x_* ein lokales Minimum, so gilt eine der folgenden Aussagen:



$x_* > 0, f'(x_*) = 0$ | $x_* = 0, f'(x_*) = 0$ | $x_* = 0, f'(x_*) > 0$

Optimalitätsbedingungen

$$\text{Minimiere} \quad f(x) \;\text{ über } x \in \mathbb{R}$$
$$\text{u.d.N.} \quad x \geq 0$$
$$f \in C^1(\mathbb{R})$$

Ist x_* ein lokales Minimum, so gelten die folgenden Bedingungen:

1. $\frac{df}{dx}(x_*) \geq 0$
2. $x_* \geq 0$
3. $x_* \dfrac{df}{dx}(x_*) = 0$ *Komplementaritätsbedingung*

Lineare Ungleichungsnebenbedingungen

Minimiere $f(x)$ über $x \in \mathbb{R}^n$

u.d.N. $a_i^T x \geq b_i, \quad i = 1, \ldots, m$

$f \in C^1(\mathbb{R}^n)$

$a_i \in \mathbb{R}^n, b_i \in \mathbb{R}$

- Sei x_* ein lokales Minimum
- Wie kann man x_* kennzeichnen?

Aktive und inaktive Mengen

- Als aktive Menge bei x_* bezeichnen wir:

aktive UNB

$$\mathcal{A}(x_*) = \{i = 1, \ldots, m\colon\ a_i^T x_* = b_i\}$$

Aktive NB lassen sich bei weitem ähnlich zu GNB betrachten

- Als inaktive Menge bei x_* bezeichnen wir:

inaktive UNB

$$\mathcal{I}(x_*) = \{i = 1, \ldots, m\colon\ a_i^T x_* > b_i\}$$

Inaktive UNB lassen sich bei der lokalen Analyse weglassen

Optimalitätsbedingung

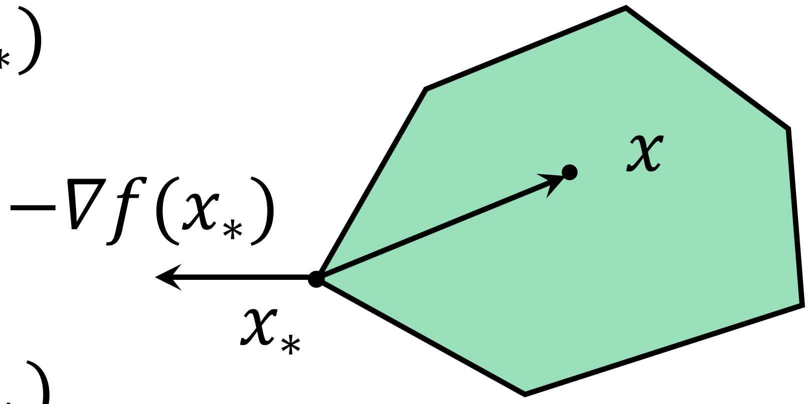
- x_* ist ebenfalls eine optimale Lösung des Programms

$$\text{Minimiere } f(x) \text{ über } x \in \mathbb{R}^n$$
$$\text{u.d.N. } a_i^T x \geq b_i, \quad i \in \mathcal{A}(x_*)$$

Satz 2.20

OB für konvexe Zulässigkeitsbereiche

$$-\nabla f(x_*)^T (x - x_*) \leq 0$$
$$\forall x \in \mathbb{R}^n \text{ mit } a_i^T x \geq b_i, i \in \mathcal{A}(x_*)$$



Satz 2.20

- Ist f konvex und x_* zulässig, so ist dies ebenfalls hinreichend für die Optimalität von x_*

Optimalitätsbedingung

$$-\nabla f(x_*)^T \underbrace{(x - x_*)}_{d} \le 0 \quad \forall x \in \mathbb{R}^n \text{ mit } a_i^T x \ge b_i, i \in \mathcal{A}(x_*)$$

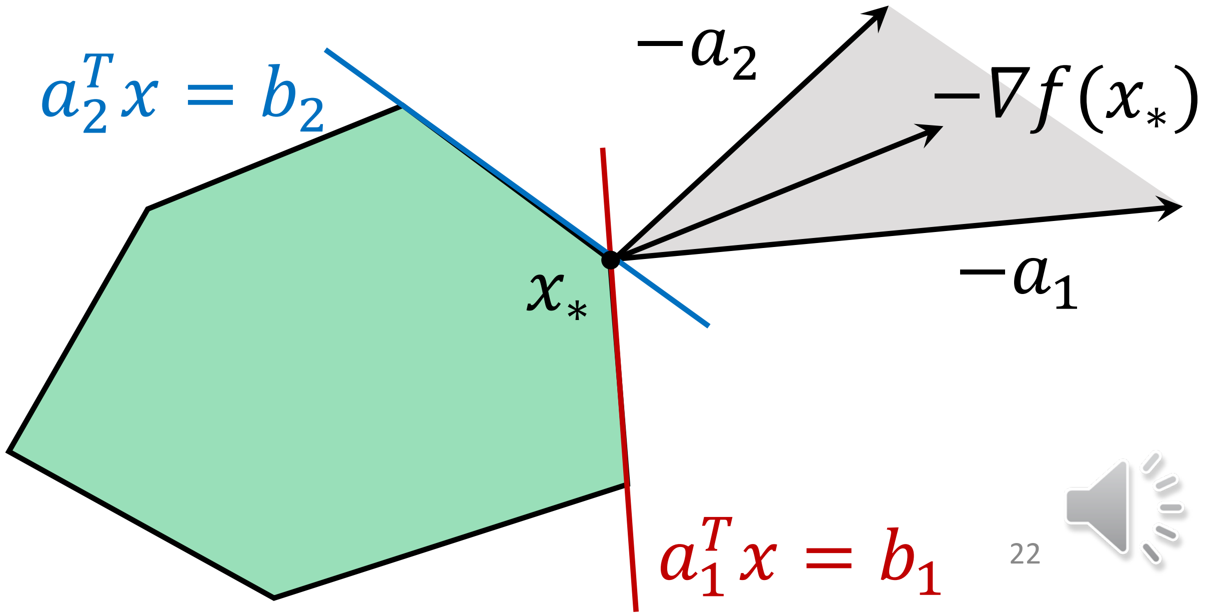
$$\Updownarrow \quad a_i^T x_* = b_i$$

$$\nabla f(x_*)^T d \ge 0 \quad \forall\, d \in \mathbb{R}^n \text{ mit } a_i^T d \ge 0, i \in \mathcal{A}(x_*)$$

$\Updownarrow$ Lemma 5.10 von Farkas

$$\nabla f(x_*) = \sum\nolimits_{i \in \mathcal{A}(x_*)} \mu_{*,i} a_i$$

$$\mu_{*,i} \ge 0, i \in \mathcal{A}(x_*)$$



Lemma 5.11: OB für lineare UNB

Minimiere $f(x)$ über $x \in \mathbb{R}^n$ $\quad f \in C^1(\mathbb{R}^n)$

u.d.N. $a_i^T x \geq b_i, \quad i = 1, \dots, m$ $\quad a_i \in \mathbb{R}^n, b_i \in \mathbb{R}$

Lagrange-Multiplikator

- Ist x_* eine optimale Lösung, so $\exists \mu_{*,i} \geq 0, i = 1, \dots, m$ so, dass:

$$\nabla f(x_*) = \sum_{i=1}^m \mu_{*,i} a_i$$

$$\mu_{*,i}(a_i^T x_* - b_i) = 0, \; i = 1, \dots, m$$ *Komplementaritätsbedingung*

- Sei f konvex und sei x_* zulässig. Seien $\mu_{*,i} \geq 0, i = 1, \dots, m$ so, dass die obigen Bedingungen gelten. Dann ist x_* eine optimale Lösung

Allgemeine lineare Nebenbedingungen

Minimiere $\quad f(x)$ über $x \in \mathbb{R}^n$

u.d.N. $a_{U,i}^T x \geq b_{U,i},\ i = 1,..,m_U$

$a_{G,j}^T x = b_{G,i},\ j = 1,..,m_G$

Minimiere $\quad f(x)$ über $x \in \mathbb{R}^n$

u.d.N. $a_{U,i}^T x \geq b_{U,i},\ i = 1,..,m_U$

$a_{G,j}^T x \geq b_{G,i},\ j = 1,..,m_G$

$-a_{G,j}^T x \geq -b_{G,i},\ j = 1,..,m_G$

Optimalitätsbedingungen

- Sei x_* eine optimale Lösung. Setze:

$$\mathcal{A}(x_*) = \{i = 1, \dots, m_U : a_{U,i}^T x_* = b_{U,i}\} \quad \textit{aktive ursprüngliche UNB}$$

- Die UNB, die den GNB entsprechen, sind automatisch aktiv bei x_*

- Lemma 5.11 $\Rightarrow$ $\exists \mu_{*,i} \geq 0, i = 1, \dots, m_U$ und $\mu_{\pm,i} \geq 0$, i $= 1, \dots m_G$ sodass:

$$\nabla f(x_*) = \sum_{i=1}^{m} \mu_{*,i} a_{U,i} + \sum_{i=1}^{m_U} \underbrace{\left(\mu_{+,i} - \mu_{-,i}\right)}_{\lambda_{*,i} \in \mathbb{R}} a_{G,i}$$

$$\mu_{*,i}\left(a_{U,i}^T x_* - b_i\right) = 0, \; i = 1, \dots, m_U$$

KKT-Bedingungen

- (x_*, μ_*, λ_*) erfüllt die KKT-Bedingungen (nach Karush, Kuhn und Tucker):

$$
\begin{gathered}
\nabla f(x_*) = \sum_{i=1}^{m_U} \mu_{*,i} a_{U,i} + \sum_{i=1}^{m_G} \lambda_{*,i} a_{G,i} \\
a_{U,i}^T x_* \geq b_{U,i}, \; i = 1, . . , m_U \\
a_{G,i}^T x_* = b_{G,i}, \; i = 1, . . , m_G \\
\mu_{*,i} \geq 0, \quad i = 1, \dots, m_U \\
\mu_{*,i}\left(a_{U,i}^T x_* - b_{U,i}\right) = 0, \quad i = 1, \dots, m_U
\end{gathered}
$$

- Wir bezeichnen (x, μ, λ) als KKT-Punkt falls die folgenden Bedingungen für (x, μ, λ) gelten

Satz 5.12. Optimalitätsbedingungen für lineare NB

Minimiere $f(x)$ über $x \in \mathbb{R}^n$ $\qquad f \in C^1(\mathbb{R}^n)$

u.d.N. $a_{U,i}^T x \geq b_{U,i},\ i = 1,..,m_U \qquad a_{U,i} \in \mathbb{R}^n, b_{U,i} \in \mathbb{R}$

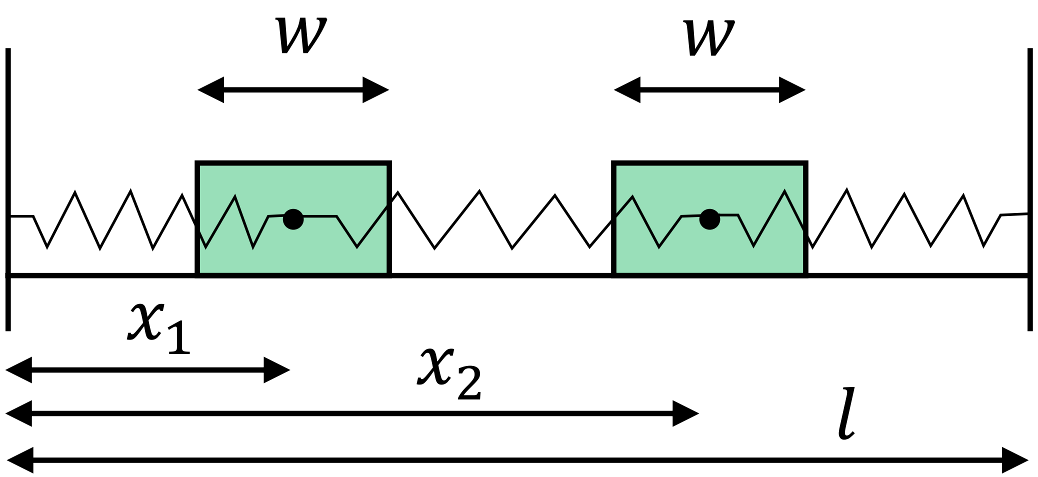
$a_{G,j}^T x = b_{G,i},\ j = 1,..,m_G \qquad a_{G,i} \in \mathbb{R}^n, b_{G,i} \in \mathbb{R}$

Lagrange-Multiplikatoren

- Sei x_* eine optimale Lösung. Dann $\exists \mu_* \in \mathbb{R}^{m_U}, \lambda_* \in \mathbb{R}^{m_G}$ so, dass (x_*, μ_*, λ_*) ein KKT-Punkt ist

- Sei f konvex und sei (x_*, μ_*, λ_*) ein KKT-Punkt. Dann ist x_* optimal

Mechanische Interpretation



$$x_1 - \frac{w}{2} \geq 0$$

$$x_2 - x_1 - w \geq 0$$

$$l - x_2 - \frac{w}{2} \geq 0$$

- Zwei Blöcke, zwei Wände und drei (Null-Länge) Federn mit Feder-Konstanten k_1, k_2, k_3
- Potentialenergie der drei Federn:

$$f(x_1, x_2) = \frac{1}{2} k_1 x_1^2 + \frac{1}{2} k_2 (x_2 - x_1)^2 + \frac{1}{2} k_3 (l - x_2)^2$$

- Im Gleichgewicht (x_1^*, x_2^*) ist die Potentialenergie minimal

$\Longrightarrow$ [Bo, 5.5.3]

Mathematisches Programm

Minimiere $\quad f(x_1, x_2) = \frac{1}{2}k_1 x_1^2 + \frac{1}{2}k_2(x_2 - x_1)^2 + \frac{1}{2}k_3(l - x_2)^2$

u.d.N. $\quad x_1 - \frac{w}{2} \geq 0$

$$x_2 - x_1 - w \geq 0$$

$$l - x_2 - \frac{w}{2} \geq 0$$

KKT-Bedingungen

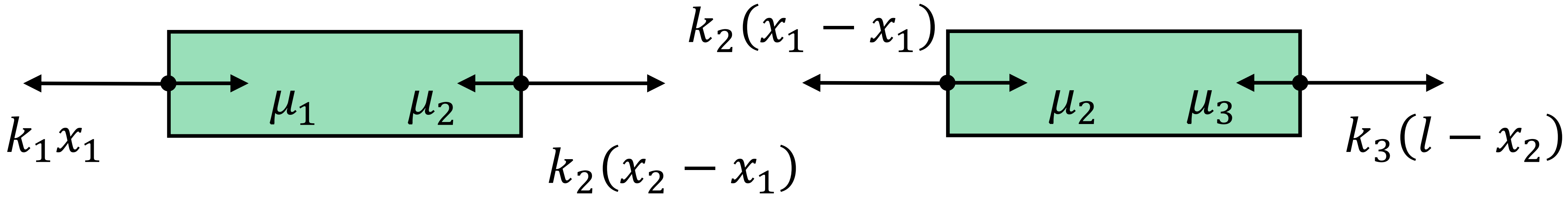
$$\begin{bmatrix} k_1 x_1 - k_2(x_2 - x_1) \\ k_2(x_2 - x_1) - k_3(l - x_2) \end{bmatrix} = \overbrace{\mu_1}^{\geq 0} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \overbrace{\mu_2}^{\geq 0} \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \overbrace{\mu_3}^{\geq 0} \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

$$\mu_1 \left(x_1 - \frac{w}{2} \right) = 0, \quad \mu_2(x_2 - x_1 + w) = 0, \quad \mu_3 \left(l - x_2 - \frac{w}{2} \right) = 0$$

Reaktionskräfte

$$\begin{bmatrix} k_1 x_1 - k_2(x_2 - x_1) \\ k_2(x_2 - x_1) - k_3(l - x_2) \end{bmatrix} = \mu_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \mu_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \mu_3 \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

- Ausgleichsgleichungen für die Kräfte, die auf die Blöcke wirken



- μ_1 ist eine von der linken Wand ausgeübte Reaktionskraft

  $\mu_1 > 0$ *nur wenn es einen Kontakt mit der Wand gibt*

- μ_2 entspricht der Reaktionskraft für einen Kontakt der Blöcke

Komplementaritätsbedingungen

$$\mu_1\left(x_1-\frac{w}{2}\right)=0,\quad \mu_2(x_2-x_1+w)=0,\quad \mu_3\left(l-x_2-\frac{w}{2}\right)=0$$

Reaktionskraft ist Null, wenn es keinen Kontakt gibt

Zusammenfassung

- Konvexe Kegel und Lemma von Farkas
- KKT-Bedingungen

Nächstes Video

- 5c. Lineare Nebenbedingungen: Lagrange-Dualität