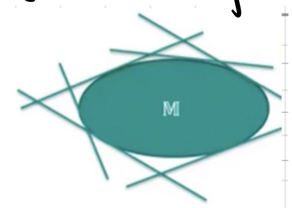
Restringierte nichtlineare Optimierung

– Schnittebenenverfahren von Kelley (numerisches Verfahren)
→ äußeres Approximations Verfahren



· Voraussetzungen:

- · f ist linear
- · g_i $(i \in I)$ ist konvex und stetig differenzierbar
- · $M = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0 \,, i \in I\}$ ist nichtleer und beschränkt

Weierstraß anwendbar

1. Initialisierung
   - Wähle ein Startpolytop $M^0 \supseteq M$.
   - Setze $k = 0$

2. Löse das Hilfsproblem LP^k durch Maximierung von $-f(x)$
   ⇝ x^h ist optimale Punkt von LP^k

3. für welches i liegt die größte Verletzung vor?
   ⇝ welches i ist $\max_i \; g_i(x^h)$

$$\Rightarrow M^{h+1} = M^k \cap \underbrace{\{x \in \mathbb{R}^2 \mid g_i(x^h) + \nabla g_i(x^h)^T (x - x^h) \leq 0}_{\text{neue NB für } LP^{h+1}}$$

4. Setze $h = h+1$ und gehe zu Schritt 2.

(Anmerkung: In der Praxis gibt es noch ein Abbruchkriterium zw. Schritt 2 und 3)

Verfahren von Franke-Wolfe (numerisches Verfahren)

⇝ Verfahren der zulässigen Richtung

- Voraussetzungen:
  - f stetig differenzierbar und konvex
  - $\mathbb{M}$ konvex, nichtleer und kompakt

$\Big\}$ Weierstraß anwendbar

1. Initialisierung
   - Wähle einen Startpunkt $x^0 \in \mathbb{M}$
   - Setze $k=0$

2. Bestimme Gradient von $\nabla f(x^k)$
3. Bestimme Suchrichtung
   - Löse Hilfsproblem Q^k: $\min \langle \nabla f(x^k), x - x^k \rangle$

     $\text{s.t. } Ax \leq 0$

   - Bestimme optimalen Punkt y^k (mit ZFW v^k) von Q^k
   - Setze $d^k = y^k - x^k$

4. Abbruchkriterium: $v^k \geq 0$
5. Bestimme Schrittweite

   $\min z = f(x^k + t d^k)$

   $\text{s.t. } t \in [0,1]$

   - Setze $x^{k+1} = x^k + t \cdot d^k$
6. Setze $k = k+1$ und gehe zu Schritt 2.

Aufgabe 1

Lösen Sie die folgenden linearen Optimierungsprobleme, die als Subprobleme in den nachfolgenden Aufgaben auftreten werden.

a)
$$\begin{array}{llcrcl} \max & x_1 & + & 2x_2 & & \\ \text{s.t.} & x_1 & & & \leq & 1 \\ & & & x_2 & \leq & 1 \\ & & & x_1, x_2 & \geq & 0 \end{array}$$

b)
$$\begin{array}{llcrcl} \max & x_1 & + & 4x_2 & & \\ \text{s.t.} & x_1 & & & \leq & 1 \\ & & & x_2 & \leq & 1 \\ & x_1 & + & x_2 & \leq & \frac{3}{2} \\ & & & x_1, x_2 & \geq & 0 \end{array}$$

c)
$$\begin{array}{llcrcrl} \max & 2x_1 & - & 2x_2 & + & 2 & \\ \text{s.t.} & x_1 & + & 2x_2 & \leq & 2 & \\ & & & x_1, x_2 & \geq & 0 & \end{array}$$

d)
$$\begin{array}{llcrcrl} \max & -\frac{2}{5}x_1 & - & \frac{4}{5}x_2 & + & 4 & \\ \text{s.t.} & x_1 & + & 2x_2 & \leq & 2 & \\ & & & x_1, x_2 & \geq & 0 & \end{array}$$

a)

(0)	x_1	x_2	
z	-1	-2	0
x_3	1	0	1
x_4	0	$\boxed{1}$	1

(1)	x_1	x_4	
z	-1	2	2
x_3	$\boxed{1}$	0	1
x_2	0	1	1

(2)	x_3	x_4	
z	1	2	3
x_1	1	0	1
x_2	0	1	1

Optimaler Punkt $x_1 = 1, x_2 = 1$ mit $f(x) = 3$

b)

(0)	x_1	x_2	
z	-1	-2	0
x_3	1	0	1
x_4	0	$\boxed{1}$	1
x_5	1	1	$\frac{3}{2}$

(1)	x_1	x_4	
z	-1	2	2
x_3	1	0	1
x_2	0	1	1
x_5	$\boxed{1}$	-1	$\frac{1}{2}$

(2)	x_5	x_4	
z	1	1	$\frac{5}{2}$
x_3	-1	1	$\frac{1}{2}$
x_2	0	1	1
x_1	1	-1	$\frac{1}{2}$

Optimaler Punkt $x_1 = \frac{1}{2}, x_2 = 1$ mit $f(x) = \frac{5}{2}$

c)

(0)	x_1	x_2	
z	-2	2	2
x_3	$\boxed{1}$	2	2

(1)	x_3	x_2	
z	2	6	6
x_1	1	2	2

Optimaler Punkt $x_1 = 2, x_2 = 0$ mit $f(x) = 6$

d)

(0)	x_1	x_2	
z	$\frac{2}{5}$	$\frac{4}{5}$	4
x_3	1	2	2

Optimaler Punkt $x_1 = 0, x_2 = 0$ mit $f(x) = 4$

Aufgabe 2

Gegeben sei das folgende nichtlineare Optimierungsproblem:

$$
\begin{array}{rll}
P: & \min & 3x_1^2 + 2x_2^2 - x_1x_2 - 2x_1 - 6x_2 + 1 \\
 & \text{s.t.} & x_1 + 2x_2 \leq 2 \\
 & & x_1 - x_2 \leq 0 \\
 & & x_1, x_2 \geq 0
\end{array}
$$

Formulieren Sie ein zu P äquivalentes Optimierungsproblem P' mit linearer Zielfunktion. Lässt sich darauf das Schnittebenenverfahren von Kelley anwenden? Begründen Sie Ihre Antwort.

A2) Verschieben von konvexer Zielfunktion in eine Nebenbedingung.

$$
\begin{array}{rlrl}
P': & \min\limits_{(x,\alpha)\in\mathbb{R}^2\times\mathbb{R}} & \alpha & \\
 & \text{s.t.} & 3x_1^2 + 2x_2^2 - x_1x_2 - 2x_1 - 6x_2 + 1 & \leq \alpha \\
 & & x_1 + 2x_2 - 2 & \leq 0 \\
 & & x_1 - x_2 & \leq 0 \\
 & & -x_1 & \leq 0 \\
 & & -x_2 & \leq 0
\end{array}
$$

äquivalent zu "="

$\rightsquigarrow$ Die Zielfunktion hat also die Form $c^T\tilde{x}$ mit $c = (0,0,1)$ und $\tilde{x} = (x_1, x_2, \alpha)$

Prüfe Voraussetzungen:

- f linear $\longrightarrow$ erfüllt
- g_i $(i \in I)$ konvex und stetig differenzierbar

$\longrightarrow$ $g_0(x) = 3x_1^2 + 2x_2^2 - x_1x_2 - 2x_1 - 6x_2 + 1 - \alpha$

$g_1(x) = x_1 + 2x_2 - 2$

$g_2(x) = x_1 - x_2$

$g_3(x) = -x_1$

$g_4(x) = -x_2$

Als Summe stetig differenzierbarer Funktionen sind alle $g_i(x)$ $(i \in I)$ auch stetig differenzierbar.
Da $g_i(x)$ $(i \in 1,2,3,4)$ alle linear sind, sind sie auch konvex. Prüfe $g_0(x)$ über die Definitheit.

$$\nabla g_0(x) = \begin{pmatrix} 6x_1 - x_2 - 2 \\ 4x_2 - x_1 - 6 \end{pmatrix} \qquad D^2 g_0(x) = \begin{pmatrix} 6 & -1 \\ -1 & 4 \end{pmatrix}$$

$$\det\left(D^2 g_0(x) - \lambda I\right) = \lambda^2 - 10\lambda + 23 \overset{!}{=} 0$$

$$\Rightarrow \lambda_{1,2} = 5 \pm \sqrt{2} \geq 0$$

Da alle Eigenwerte positiv sind, ist die Hesse-matrix positiv semidefinit und $g_0(x)$ konvex.

- $\mathbb{M}$ nichtleer und beschränkt

$\rightarrow$ nichtleer, da $x = (0,0) \in \mathbb{M}$

beschränkt, da aus $g_i(x)$ $(i = 1,3,4)$ folgt:

$x_1, x_2 \geq 0$ und $x_1 \leq 2$, $x_2 \leq 1$

$\Rightarrow \mathbb{M} \subseteq [0,2] \times [0,1]$.

Somit sind alle Voraussetzungen für das Schnitt-ebenenverfahren von Kelley erfüllt.

Aufgabe 3

Gegeben sei das folgende nichtlineare Optimierungsproblem:

$g_1(x) = x_1^2 + x_2^2 - 1$

$\nabla g_1(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$

$$\begin{array}{lrcrcl} \min & -x_1 & - & 2x_2 & & \\ \text{s.t.} & x_1^2 & + & x_2^2 & \leq & 1 \\ & & & x_1, x_2 & \geq & 0 \end{array}$$

Führen Sie das Schnittebenenverfahren von Kelleyausgehend von

$$\mathbb{M}^0 = \left\{(x_1, x_2) \in \mathbb{R}^2 \,|\, 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\right\}$$

bis zur Bestimmung der zweiten Schnittebene aus. Geben Sie insbesondere auch die Menge $\mathbb{M}^1$ und den Optimalpunkt x^1 des zugehörigen Optimierungsproblems P^1 an.

A3) Als initialer Bereich für LP^0 wird gewählt:

$\mathbb{M}^0 = \{(x_1,x_2) \in \mathbb{R}^2 \,|\, 0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1\}$

Es ergibt sich das lineare Problem LP^0

$$LP^{0\prime}: \max \quad x_1 + 2x_2 \\ x_1 \leq 1 \\ x_2 \leq 1 \\ x_1, x_2 \geq 0$$

(0)	x_1	x_2	
z	-1	-2	0
x_3	1	0	1
x_4	0	$\boxed{1}$	1

(1)	x_1	x_4	
z	-1	2	2
x_3	$\boxed{1}$	0	1
x_2	0	1	1

(2)	x_3	x_4	
z	1	2	3
x_1	1	0	1
x_2	0	1	1

Optimaler Punkt $x_1 = 1, x_2 = 1$ mit $f(x) = 3$

Der optimale Punkt von LP^0 ist also $x^* = (1,1,0,0)$ mit $z^* = -3$.

Da wegen $g_1(x^0) = 1 > 0$ der Punkt $x^0 \notin \mathbb{M}$ gilt wird eine neue Schnittebene hinzugefügt:

$$g_1(x^0) + \nabla g_1(x^0)^T (x - x^0) \leq 0$$

$$\Leftrightarrow \quad 1 + (2,2) \begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix} \leq 0$$

$$\Leftrightarrow \quad x_1 + x_2 \leq \frac{3}{2}$$

Daraus ergibt sich die neue zulässige Menge

$$\mathbb{M}^1 = \{(x_1,x_2) \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1,\ x_1 + x_2 \leq \tfrac{3}{2}\}$$

Es ergibt sich das lineare Problem LP^1

$$
LP^{1'}: \max \quad
\begin{array}{l}
x_1 + 2x_2 \\
x_1 \leq 1 \\
x_2 \leq 1 \\
x_1 + x_2 \leq \frac{3}{2} \\
x_1, x_2 \geq 0
\end{array}
$$

(0)	x_1	x_2	
z	-1	-2	0
x_3	1	0	1
x_4	0	$\boxed{1}$	1
x_5	1	1	$\frac{3}{2}$

(1)	x_1	x_4	
z	-1	2	2
x_3	1	0	1
x_2	0	1	1
x_5	$\boxed{1}$	-1	$\frac{1}{2}$

(2)	x_5	x_4	
z	1	1	$\frac{5}{2}$
x_3	-1	1	$\frac{1}{2}$
x_2	0	1	1
x_1	1	-1	$\frac{1}{2}$

Optimaler Punkt $x_1 = \frac{1}{2}, x_2 = 1$ mit $f(x) = \frac{5}{2}$

Der optimale Punkt von LP^1 ist also $x^* = (\frac{1}{2}, 1, \frac{1}{2}, 0, 0)$ mit $z^* = -\frac{5}{2}$

Da wegen $g_1(x^1) = \frac{1}{4} > 0$ der Punkt $x^1 \notin \mathbb{M}$ gilt wird eine neue Schnittebene hinzugefügt:

$$g_1(x^1) + \nabla g_1(x^1)^T (x - x^1) \leq 0$$

$$\Leftrightarrow \quad \frac{1}{4} + (1,2) \begin{pmatrix} x_1 - \frac{1}{2} \\ x_2 - 1 \end{pmatrix} \leq 0$$

$$\Leftrightarrow \quad x_1 + 2x_2 \leq \frac{9}{4}$$

Daraus ergibt sich die neue zulässige Menge

$$\mathbb{M}^2 = \{(x_1,x_2) \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1,\ x_1 + x_2 \leq \tfrac{3}{2},\ x_1 + 2x_2 \leq \tfrac{9}{4}\}$$

Als Nächstes müsste damit das lineare Optimierungsproblem LP^2 gelöst werden:

$$
\begin{array}{ll}
\min & -x_1 - 2x_2 \\
\text{s.t.} & x \in \mathbb{M}^2
\end{array}
$$

Da das Vorgehen völlig analog zur ersten Iteration verläuft, verzichten wir auf eine weitere Darstellung der numerischen Schritte. Insgesamt lässt sich festhalten, dass das Kelley-Verfahren eine Approximation des optimalen Punkts eines nichtlinearen Optimierungsproblems algorithmisch durch das Lösen einer Folge von linearen Optimierungsproblemen (die jeweils z.B. mit dem Simplex-Algorithmus gelöst werden können) ermittelt.

Aufgabe 4

Gegeben sei das nichtlineare Optimierungsproblem

$$\begin{array}{llcl} \min & f(x_1,x_2) = x_1 e^{x_1} + x_2^2 e^{x_2} - 2x_1 - 3x_2 \\ \text{s.t.} & x_1 + x_2 & \leq & 3 \\ & x_1, x_2 & \geq & 0 \end{array}$$

a) Sind die Voraussetzungen zur Anwendung des Verfahrens von Frank-Wolfe erfüllt?

A4a) • f stetig differenzierbar und konvex

→ Als Summe stetig differenzierbarer Funktionen ist f stetig differenzierbar.

Konvexität über Definitheit:

$$\nabla f(x) = \begin{pmatrix} e^{x_1} + x_1 e^{x_1} - 2 \\ 2x_2 e^{x_2} + 2x_2^2 e^{x_2} - 3 \end{pmatrix} \qquad D^2 f(x) = \begin{pmatrix} 2e^{x_1} + x_1 e^{x_1} & 0 \\ 0 & 2e^{x_2} + 4x_2 e^{x_2} + x_2^2 e^{x_2} \end{pmatrix}$$

⇒ Da $x_1, x_2 \geq 0$ ist $D^2 f(x)$ positiv semidefinit und somit konvex.

• $\mathbb{M}$ kompakt, konvex und nichtleer

→ Die Nebenbedingungen beschreiben den Schnitt von 3 abgeschlossenen Halbräumen. Somit ist $\mathbb{M}$ eine abgeschlossene konvexe Menge.

$\mathbb{M}$ ist nichtleer, da $x = (0,0) \in \mathbb{M}$.

$\mathbb{M}$ ist durch die Nebenbedingung beschränkt:

$\mathbb{M} \subseteq [0,3] \times [0,3]$

Somit sind alle Voraussetzungen für das Franke-Wolfe-Verfahren erfüllt.

b) Bestimmen Sie ausgehend von $x^0 = (1, 1)$ die erste Suchrichtung und stellen Sie das Optimierungsproblem zur Bestimmung der ersten Schrittweite auf.

A4b) Bestimme Suchrichtung mit Hilfsproblem Q^0:

$$Q^0: \min \langle \nabla f(x^0), x - x^0 \rangle$$

$$= \langle (2e-2, 3e-3), (x_1 - 1, x_2 - 1) \rangle$$

$$= (2e-2)x_1 + (3e-3)x_2 - 5e + 5$$

$$\text{s.t. } x_1 + x_2 \leq 3$$

$$x_1, x_2 \geq 0$$

Da sich die Ecken von M sehr einfach angeben lassen, führen wir in diesem Fall eine Enumeration der Ecken anstatt dem Simplex aus.

Ecke (x_1, x_2)	ZFW von Q^0
$(0,0)$	$-5e+5 \approx -8.59$
$(3,0)$	$e-1 \approx 1.72$
$(0,3)$	$4e-4 \approx 6.87$

Q^0 besitzt den Optimalpunkt $y^0 = (0,0)$ mit Optimalwert $v^0 \approx -8.59$.

Prüfe auf Optimalität: $v^0 \approx -8.59 < 0$

$\Rightarrow$ Optimalpunkt nicht gefunden

Suchrichtung: $d^0 = y^0 - x^0 = (0,0) - (1,1) = (-1,-1)$

Bestimmung der Schrittweite t

$$\min \quad z = f(x^0 + td^0) = f(1-t, 1-t)$$
$$= (1-t)e^{1-t} + (1-t)^2 e^{1-t} - 5(1-t)$$

$$\text{s.t.} \quad t \in [0,1]$$

Aufgabe 5

Gegeben sei das nichtlineare Optimierungsproblem

$$\begin{array}{llcl} \min & f(x_1, x_2) = x_1^2 - 2x_1 + x_2^2 + 1 & & \\ \text{s.t.} & x_1 \;+\; 2x_2 & \leq & 2 \\ & \qquad\quad x_1, x_2 & \geq & 0 \end{array}$$

Führen Sie ausgehend von $x^0 = (0, 1)$ zwei Iterationsschritte des Verfahrens von Frank-Wolfe aus, sofern die Voraussetzungen hierfür gegeben sind.

A5) Prüfe Voraussetzungen

- f stetig differenzierbar und konvex

$\rightarrow$ Als Summe stetig differenzierbarer Funktionen ist f stetig differenzierbar.

Konvexität über Definitheit:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - 2 \\ 2x_2 \end{pmatrix} \qquad D^2 f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$\Rightarrow$ $D^2 f(x)$ ist positiv semidefinit und somit konvex

- $\mathbb{M}$ kompakt, konvex und nichtleer

$\rightarrow$ Die Nebenbedingungen beschreiben den Schnitt von 3 abgeschlossenen Halbräumen. Somit ist $\mathbb{M}$ eine abgeschlossene konvexe Menge.

$\mathbb{M}$ ist nichtleer, da $x = (0,0) \in \mathbb{M}$.

$\mathbb{M}$ ist durch die Nebenbedingung beschränkt:

$\mathbb{M} \subseteq [0,2] \times [0,1]$

Somit sind alle Voraussetzungen für das Franke-Wolfe-Verfahren erfüllt.

Iteration 1: $x^0 = (0,1)$ $\quad f(x^0) = 2$

- Bestimmung Gradient $\nabla f(x^0) = \nabla f(0,1) = (-2,2)$
- Bestimmung Suchrichtung mit Hilfsproblem Q^0

Q^0: $\min \langle \nabla f(x^0), x - x^0 \rangle = -2x_1 + 2x_2 - 2$

s.t. $\quad x_1 + 2x_2 \leq 2$

$\quad x_1, x_2 \geq 0$

(0)	x_1	x_2	
z	-2	2	2
x_3	$\boxed{1}$	2	2

(1)	x_3	x_2	
z	2	6	6
x_1	1	2	2

Optimaler Punkt $x_1 = 2, x_2 = 0$ mit $f(x) = 6$

$\Rightarrow y^0 = (2,0)$ mit $v^0 = -6$

- Suchrichtung bestimmen
  - $d^0 = y^0 - x^0 = (2,0) - (0,1) = (2,-1)$
- Test auf Optimalität: $v^0 < 0$ $\lightning$
- Berechnung der Schrittweite t

$\min \quad z = g(x^0 + td^0) = 5t^2 - 6t + 2$

s.t. $\quad t \in [0,1]$

$\leadsto$ Minimum einer Funktion = Ableiten und Nullsetzen

$\Rightarrow g'(t) = 10t - 6 \stackrel{!}{=} 0 \quad \Rightarrow t^0 = \frac{3}{5}$

neuer Standpunkt: $x^1 = x^0 + t^0 d^0$

$$= (0,1) + \frac{3}{5}\cdot(2,-1) = \left(\frac{6}{5}, \frac{2}{5}\right)$$

mit $f(x^1) = \frac{1}{5}$

Iteration 2:

- Bestimme Gradient $\nabla f(x^1) = \left(\frac{2}{5}, \frac{4}{5}\right)$
- Bestimme Suchrichtung mit Hilfsproblem Q^1

Q^1: $\min \langle \nabla f(x^1), x - x^1 \rangle = \frac{2}{5}x_1 + \frac{4}{5}x_2 - 4$

s.t. $x_1 + 2x_2 \leq 2$

$x_1, x_2 \geq 0$

(0)	x_1	x_2	
z	$\frac{2}{5}$	$\frac{4}{5}$	4
x_3	1	2	2

Optimaler Punkt $x_1 = 0, x_2 = 0$ mit $f(x) = 4$

$\Rightarrow y^1 = (0,0)$ mit $v^1 = -4$

- Suchrichtung bestimmen

$d^1 = y^1 - x^1 = (0,0) - \left(\frac{6}{5}, \frac{2}{5}\right) = \left(-\frac{6}{5}, -\frac{2}{5}\right)$

- Test auf Optimalität: $v^1 < 0$ ↯
- Berechnung der Schrittweite t

$\min\ z = g(x^1 + td^1) = \frac{8}{5}t^2 - \frac{4}{5}t + \frac{1}{5}$

s.t. $t \in [0,1]$

⇝ Minimum einer Funktion = Ableiten und Nullsetzen

$\Rightarrow g'(t) = \frac{16}{5}t - \frac{4}{5} \stackrel{!}{=} 0 \quad \Rightarrow t^1 = \frac{1}{4}$

neuer Standpunkt: $x^2 = x^1 + t^1 d^1$

$$= \left(\frac{6}{5}, \frac{2}{5}\right) + \frac{1}{4}\left(-\frac{6}{5}, -\frac{2}{5}\right) = \left(\frac{9}{10}, \frac{3}{10}\right)$$

mit $f(x^2) = \frac{1}{10}$.

Hinweis: Führt man das Verfahren fort, so approximiert man den globalen Minimalpunkt $x^* = (1,0)$ mit $z^* = 0$.

nächstes Tutorium:

- Straftermverfahren
- Barriereverfahren
- stochastische lineare Optimierung