Inhaltsübersicht für heute:

Inhaltsübersicht für heute:

Branch-and-Bound

Konvexe Mengen, konvexe Hülle, konvexe Funktionen

Relaxation

Branch-and-Bound

Beim systematischen Enumerieren aller Lösungen sollen möglichst viele frühzeitig durch obere und untere Schranken ausgeschlossen werden.

Bsp: $\{0,1\}$-Rucksack: Gewichte $a \in \mathbb{N}^n$, Kapazität $b \in \mathbb{N}$, Nutzen $c \in \mathbb{N}^n$,

$$\max \ c^T x \quad \text{s.t.} \quad a^T x \leq b, \ x \in \{0,1\}^n$$

Obere Schranke: $\max \ c^T x$ s.t. $a^T x \leq b, \ x \in [0,1]^n$ [LP-Relaxation]
Untere Schranke: Sortiere nach Nutzen/Gewicht und fülle danach auf

---

Ablaufskizze (für Maximierungsprobleme):
M ... Menge offener Probleme, anfangs $M = \{\text{Ursprungsproblem}\}$
$\underline{f}$ ... Wert der besten bekannten Lösung, anfangs $\underline{f} = -\infty$

1. Falls $M = \emptyset$ STOP, sonst wähle $P \in M$, $M \leftarrow M \setminus \{P\}$
2. Berechne obere Schranke $\overline{f}(P)$.
3. Falls $\overline{f}(P) < \underline{f}$ (P enthält keine OL), gehe zu 1.
4. Berechne zulässige Lösung $\hat{f}(P)$ für P (untere Schranke).
5. Ist $\hat{f}(P) > \underline{f}$ (neue beste Lösung), setze $\underline{f} \leftarrow \hat{f}(P)$
6. Ist $\overline{f}(P) = \hat{f}(P)$ (keine bessere Lösung in P), gehe zu 1.
7. Teile P in „kleinere" Teilprobleme P_i, $M \leftarrow M \cup \{P_1, \ldots, P_k\}$
8. Gehe zu 1.

Beispiel: $\{0,1\}$-Rucksackproblem

Gegenstand	A	B	C	D	E	F	Kapazität
Gewicht(a)	9	7	6	4	4	3	14
Nutzen (c)	18	6	18	7	6	5	

Sortierung Nutzen/Gewicht: $C > A > D > F > E > B$.

Obere Schranke: $\max\ c^T x$ s.t. $a^T x \leq 14,\ x \in [0,1]^6$ [LP-Relaxation]

Untere Schranke: Nach Sortierung möglichst lange auffüllen

Beispiel: $\{0,1\}$-Rucksackproblem

Gegenstand	A	B	C	D	E	F	Kapazität
Gewicht(a)	9	7	6	4	4	3	14
Nutzen (c)	18	6	18	7	6	5	

Sortierung Nutzen/Gewicht: $C > A > D > F > E > B$.

Obere Schranke: max $c^T x$ s.t. $a^T x \leq 14,\ x \in [0,1]^6$ [LP-Relaxation]

Untere Schranke: Nach Sortierung möglichst lange auffüllen

P_1: Originalproblem
OS: $C + \frac{8}{9}A = 34$
US: $C + D + F = 30$

Beispiel: $\{0,1\}$-Rucksackproblem

Gegenstand	A	B	C	D	E	F	Kapazität
Gewicht(a)	9	7	6	4	4	3	14
Nutzen (c)	18	6	18	7	6	5	

Sortierung Nutzen/Gewicht: $C > A > D > F > E > B$.

Obere Schranke: $\max\ c^T x$ s.t. $a^T x \leq 14,\ x \in [0,1]^6$ [LP-Relaxation]

Untere Schranke: Nach Sortierung möglichst lange auffüllen

P_1: Originalproblem
OS: $C+\frac{8}{9}A=34$
US: $C+D+F=30$

$1 \quad \leftarrow \quad x_A \quad \rightarrow \quad 0$

Beispiel: $\{0,1\}$-Rucksackproblem

Gegenstand	A	B	C	D	E	F	Kapazität
Gewicht(a)	9	7	6	4	4	3	14
Nutzen (c)	18	6	18	7	6	5	

Sortierung Nutzen/Gewicht: $C > A > D > F > E > B$.

Obere Schranke: $\max\ c^T x$ s.t. $a^T x \leq 14,\ x \in [0,1]^6$ [LP-Relaxation]

Untere Schranke: Nach Sortierung möglichst lange auffüllen

P_1: Originalproblem
OS: $C + \frac{8}{9}A = 34$
US: $C + D + F = 30$

$1 \quad \leftarrow \quad x_A \quad \rightarrow \quad 0$

P_2: $x_A = 1 \Rightarrow x_B = x_C = 0$
OS: $A + D + \frac{1}{3}F = 26\frac{2}{3}$
OS $< 30 \Rightarrow$ keine OL $\square$

Beispiel: $\{0,1\}$-Rucksackproblem

Gegenstand	A	B	C	D	E	F	Kapazität
Gewicht(a)	9	7	6	4	4	3	14
Nutzen (c)	18	6	18	7	6	5	

Sortierung Nutzen/Gewicht: $C > A > D > F > E > B$.

Obere Schranke: max $c^T x$ s.t. $a^T x \leq 14,\ x \in [0,1]^6$ [LP-Relaxation]

Untere Schranke: Nach Sortierung möglichst lange auffüllen

P_1: Originalproblem
OS: $C + \frac{8}{9}A = 34$
US: $C + D + F = 30$

$1 \quad \leftarrow \quad x_A \quad \rightarrow \quad 0$

P_2: $x_A = 1 \Rightarrow x_B = x_C = 0$
OS: $A + D + \frac{1}{3}F = 26\frac{2}{3}$
OS $< 30 \Rightarrow$ keine OL $\square$

P_3: $x_A = 0$
OS: $C + D + F + \frac{1}{4}E = 31\frac{1}{2}$
US: $C + D + F = 30$

Beispiel: $\{0,1\}$-Rucksackproblem

Gegenstand	A	B	C	D	E	F	Kapazität
Gewicht(a)	9	7	6	4	4	3	14
Nutzen (c)	18	6	18	7	6	5	

Sortierung Nutzen/Gewicht: $C > A > D > F > E > B$.

Obere Schranke: max $c^T x$ s.t. $a^T x \leq 14,\ x \in [0,1]^6$ [LP-Relaxation]

Untere Schranke: Nach Sortierung möglichst lange auffüllen

P_1: Originalproblem
OS: $C + \frac{8}{9}A = 34$
US: $C + D + F = 30$

$1 \leftarrow x_A \rightarrow 0$

P_2: $x_A = 1 \Rightarrow x_B = x_C = 0$
OS: $A + D + \frac{1}{3}F = 26\frac{2}{3}$
OS $< 30 \Rightarrow$ keine OL $\square$

P_3: $x_A = 0$
OS: $C + D + F + \frac{1}{4}E = 31\frac{1}{2}$
US: $C + D + F = 30$

$1 \leftarrow x_E \rightarrow 0$

Beispiel: $\{0,1\}$-Rucksackproblem

Gegenstand	A	B	C	D	E	F	Kapazität
Gewicht(a)	9	7	6	4	4	3	14
Nutzen (c)	18	6	18	7	6	5	

Sortierung Nutzen/Gewicht: $C > A > D > F > E > B$.

Obere Schranke: $\max\ c^T x$ s.t. $a^T x \leq 14,\ x \in [0,1]^6$ [LP-Relaxation]

Untere Schranke: Nach Sortierung möglichst lange auffüllen

P_1: Originalproblem
OS: $C + \frac{8}{9}A = 34$
US: $C + D + F = 30$

$1 \quad \leftarrow \quad x_A \quad \rightarrow \quad 0$

P_2: $x_A = 1 \Rightarrow x_B = x_C = 0$
OS: $A + D + \frac{1}{3}F = 26\frac{2}{3}$
OS $< 30 \Rightarrow$ keine OL $\square$

P_3: $x_A = 0$
OS: $C + D + F + \frac{1}{4}E = 31\frac{1}{2}$
US: $C + D + F = 30$

$1 \quad \leftarrow \quad x_E \quad \rightarrow \quad 0$

P_4: $x_A = 0,\ x_E = 1$
OS: $E + C + D = 31$
US: $E + C + D = 31$ $\square$

Beispiel: $\{0,1\}$-Rucksackproblem

Gegenstand	A	B	C	D	E	F	Kapazität
Gewicht(a)	9	7	6	4	4	3	14
Nutzen (c)	18	6	18	7	6	5	

Sortierung Nutzen/Gewicht: $C > A > D > F > E > B$.

Obere Schranke: max $c^T x$ s.t. $a^T x \leq 14,\ x \in [0,1]^6$ [LP-Relaxation]

Untere Schranke: Nach Sortierung möglichst lange auffüllen

P_1: Originalproblem
OS: $C + \frac{8}{9}A = 34$
US: $C + D + F = 30$

$1 \quad \leftarrow \quad x_A \quad \rightarrow \quad 0$

P_2: $x_A = 1 \Rightarrow x_B = x_C = 0$
OS: $A + D + \frac{1}{3}F = 26\frac{2}{3}$
OS $< 30 \Rightarrow$ keine OL $\square$

P_3: $x_A = 0$
OS: $C + D + F + \frac{1}{4}E = 31\frac{1}{2}$
US: $C + D + F = 30$

$1 \quad \leftarrow \quad x_E \quad \rightarrow \quad 0$

P_4: $x_A = 0,\ x_E = 1$
OS: $E + C + D = 31$
US: $E + C + D = 31$ $\square$

P_5: $x_A = x_E = 0$
OS: $C + D + F + \frac{1}{7}B = 30\frac{6}{7}$
OS $< 31 \Rightarrow$ keine OL $\square$

Immer dann wird der Branch&Bound Baum groß werden, wenn viele Lösungen sehr nahe an der Optimallösung sind.

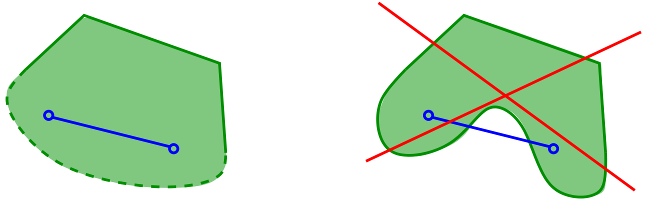
Entscheidend für den Erfolg von Branch&Bound:

Wie kommt man zu guten oberen und unteren Schranken?

Inhaltsübersicht für heute:

Konvexe Mengen und konvexe Hülle

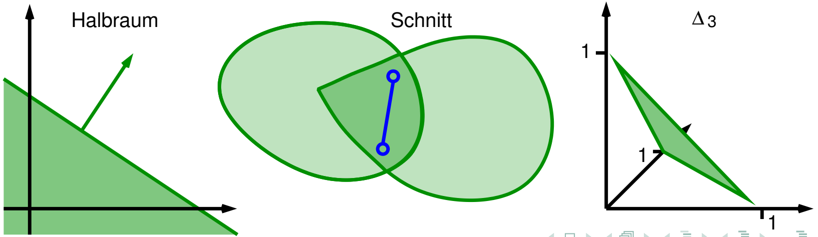
Eine Menge $C \subseteq \mathbb{R}^n$ heißt konvex, wenn für alle $x, y \in C$ auch die Verbindungsstrecke $[x, y] := \{\alpha x + (1 - \alpha) y : \alpha \in [0, 1]\}$ in C liegt.



Konvexe Mengen und konvexe Hülle

Eine Menge $C \subseteq \mathbb{R}^n$ heißt konvex, wenn für alle $x, y \in C$ auch die Verbindungsstrecke $[x, y] := \{\alpha x + (1-\alpha)y : \alpha \in [0, 1]\}$ in C liegt.

Bspe: $\emptyset$, $\mathbb{R}^n$, Halbräume, der Schnitt konvexer Mengen ist konvex, Polyeder, der k-dim. Einheitssimplex $\Delta_k := \{\alpha \in \mathbb{R}^k_+ : \sum_{i=1}^k \alpha_i = 1\}$

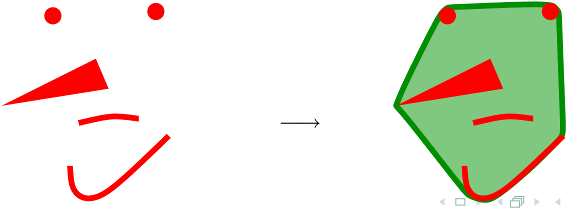


Konvexe Mengen und konvexe Hülle

Eine Menge $C \subseteq \mathbb{R}^n$ heißt konvex, wenn für alle $x, y \in C$ auch die Verbindungsstrecke $[x, y] := \{\alpha x + (1-\alpha)y : \alpha \in [0,1]\}$ in C liegt.

Bspe: $\emptyset$, $\mathbb{R}^n$, Halbräume, der Schnitt konvexer Mengen ist konvex, Polyeder, der k-dim. Einheitssimplex $\Delta_k := \{\alpha \in \mathbb{R}^k_+ : \sum_{i=1}^k \alpha_i = 1\}$

Für $S \subseteq \mathbb{R}^n$ ist die konvexe Hülle der Schnitt aller konvexen Mengen, die S enthalten, $\operatorname{conv} S := \bigcap\{C \text{ konvex} : S \subseteq C\}$.



Konvexe Mengen und konvexe Hülle

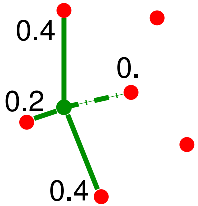
Eine Menge $C \subseteq \mathbb{R}^n$ heißt konvex, wenn für alle $x, y \in C$ auch die Verbindungsstrecke $[x, y] := \{\alpha x + (1-\alpha)y : \alpha \in [0,1]\}$ in C liegt.

Bspe: $\emptyset$, $\mathbb{R}^n$, Halbräume, der Schnitt konvexer Mengen ist konvex,

Polyeder, der k-dim. Einheitssimplex $\Delta_k := \{\alpha \in \mathbb{R}^k_+ : \sum_{i=1}^k \alpha_i = 1\}$

Für $S \subseteq \mathbb{R}^n$ ist die konvexe Hülle der Schnitt aller konvexen Mengen, die S enthalten, $\operatorname{conv} S := \bigcap\{C \text{ konvex} : S \subseteq C\}$.

Für gegebene Punkte $x^{(i)} \in \mathbb{R}^n$, $i \in \{1, \dots, k\}$ und $\alpha \in \Delta_k$ heißt $x = \sum \alpha_i x^{(i)}$ Konvexkombination der Punkte $x^{(i)}$.



Konvexe Mengen und konvexe Hülle

Eine Menge $C \subseteq \mathbb{R}^n$ heißt konvex, wenn für alle $x, y \in C$ auch die Verbindungsstrecke $[x, y] := \{\alpha x + (1-\alpha)y : \alpha \in [0,1]\}$ in C liegt.

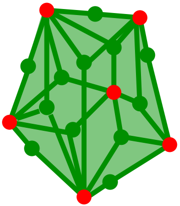
Bspe: $\emptyset$, $\mathbb{R}^n$, Halbräume, der Schnitt konvexer Mengen ist konvex,

Polyeder, der k-dim. Einheitssimplex $\Delta_k := \{\alpha \in \mathbb{R}^k_+ : \sum_{i=1}^k \alpha_i = 1\}$

Für $S \subseteq \mathbb{R}^n$ ist die konvexe Hülle der Schnitt aller konvexen Mengen, die S enthalten, $\operatorname{conv} S := \bigcap\{C \text{ konvex} : S \subseteq C\}$.

Für gegebene Punkte $x^{(i)} \in \mathbb{R}^n$, $i \in \{1, \dots, k\}$ und $\alpha \in \Delta_k$ heißt $x = \sum \alpha_i x^{(i)}$ Konvexkombination der Punkte $x^{(i)}$.

$\operatorname{conv} S$ ist die Menge aller Konvexkombinationen endlich vieler Punkte aus S, $\operatorname{conv} S = \{\sum_{i=1}^k \alpha_i x^{(i)} : x^{(i)} \in S, i = 1, \dots, k \in \mathbb{N}, \alpha \in \Delta_k\}$.



Konvexe Hülle und ganzz. Optimierung

Satz

Die konvexe Hülle endlich vieler Punkte ist ein (beschränktes) Polyeder.

Konvexe Hülle und ganzz. Optimierung

Satz

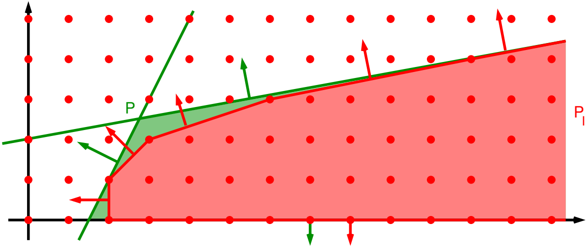
Die konvexe Hülle endlich vieler Punkte ist ein (beschränktes) Polyeder.

Die ganzzahlige Hülle eines Polyeders $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ ist die konvexe Hülle der in P enthaltenen ganzz. Punkte, $P_I := \mathrm{conv}(P \cap \mathbb{Z}^n)$.

Satz

Ist $A \in \mathbb{Q}^{m \times n}$ und $b \in \mathbb{Q}^m$, dann ist die ganzz. Hülle P_I des Polyeders $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ selbst ein Polyeder.

Problem: Beschreibung von P_I meist unbekannt oder extrem groß!



[Ausnahme z.B. für A tot. unimod., $b \in \mathbb{Z}^n$ ist $P = P_I$]

Konvexe Hülle und ganzz. Optimierung

Satz

Die konvexe Hülle endlich vieler Punkte ist ein (beschränktes) Polyeder.

Die ganzzahlige Hülle eines Polyeders $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ ist die konvexe Hülle der in P enthaltenen ganzz. Punkte, $P_I := \mathrm{conv}(P \cap \mathbb{Z}^n)$.

Satz

Ist $A \in \mathbb{Q}^{m \times n}$ und $b \in \mathbb{Q}^m$, dann ist die ganzz. Hülle P_I des Polyeders $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ selbst ein Polyeder.

Problem: Beschreibung von P_I meist unbekannt oder extrem groß!

Falls die ganzzahlige Hülle gut linear beschreibbar ist, kann man das ganzzahlige Optimierungsproblem mit Simplex lösen:

Satz

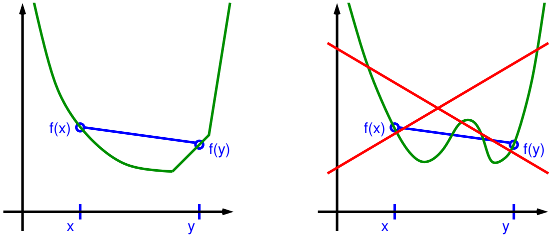
Ist für $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ die ganzzahlige Hülle durch $P_I = \{x \in \mathbb{R}^n : A_I x \leq b_I\}$ gegeben, so gilt:

$$\begin{aligned} \sup\{c^T x : A_I x \leq b_I, x \in \mathbb{R}^n\} &= \sup\{c^T x : Ax \leq b, x \in \mathbb{Z}^n\}, \\ \mathrm{Argmin}\{c^T x : A_I x \leq b_I, x \in \mathbb{R}^n\} &= \mathrm{conv}\,\mathrm{Argmin}\{c^T x : Ax \leq b, x \in \mathbb{Z}^n\}. \end{aligned}$$

Konvexe Funktionen

Eine Funktion $f : \mathbb{R}^n \to \bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ heißt konvex, wenn $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$ für $x, y \in \mathbb{R}^n$ und $\alpha \in [0,1]$.

Sie heißt streng konvex, wenn $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y)$ für $x, y \in \mathbb{R}^n$, $x \neq y$ und $\alpha \in (0,1)$.



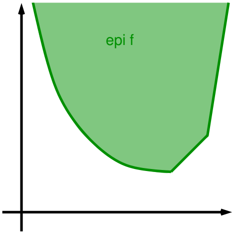
Konvexe Funktionen

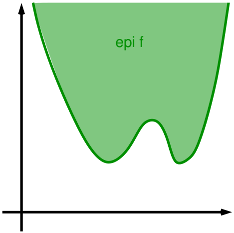
Eine Funktion $f : \mathbb{R}^n \to \bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ heißt konvex, wenn
$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$ für $x, y \in \mathbb{R}^n$ und $\alpha \in [0,1]$.

Sie heißt streng konvex, wenn $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y)$ für $x, y \in \mathbb{R}^n$, $x \neq y$ und $\alpha \in (0,1)$.

Der Epigraph einer Funktion $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ ist die Menge

$$\operatorname{epi} f := \left\{ \begin{pmatrix} x \\ r \end{pmatrix} : x \in \mathbb{R}^n, r \geq f(x) \right\}$$ [die Punkte „oberhalb“von $f(x)$]





Konvexe Funktionen

Eine Funktion $f : \mathbb{R}^n \to \bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ heißt konvex, wenn
$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$ für $x, y \in \mathbb{R}^n$ und $\alpha \in [0,1]$.

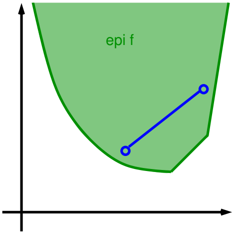
Sie heißt streng konvex, wenn $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y)$ für $x, y \in \mathbb{R}^n$, $x \neq y$ und $\alpha \in (0,1)$.

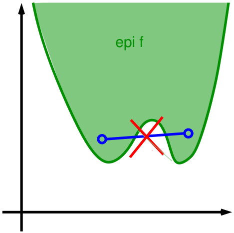
Der Epigraph einer Funktion $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ ist die Menge

$$\operatorname{epi} f := \left\{ \begin{pmatrix} x \\ r \end{pmatrix} : x \in \mathbb{R}^n, r \geq f(x) \right\}$$ [die Punkte „oberhalb“von $f(x)$]

Satz

Eine Funktion ist genau dann konvex, wenn ihr Epigraph konvex ist.





Konvexe Funktionen

Eine Funktion $f : \mathbb{R}^n \to \bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ heißt konvex, wenn $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$ für $x, y \in \mathbb{R}^n$ und $\alpha \in [0,1]$.

Sie heißt streng konvex, wenn $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y)$ für $x, y \in \mathbb{R}^n$, $x \neq y$ und $\alpha \in (0,1)$.

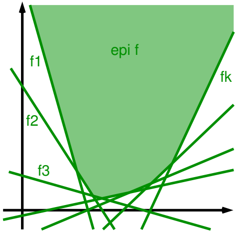
Der Epigraph einer Funktion $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ ist die Menge

$$\operatorname{epi} f := \left\{ \begin{pmatrix} x \\ r \end{pmatrix} : x \in \mathbb{R}^n, r \geq f(x) \right\}$$ [die Punkte „oberhalb“von $f(x)$]

Satz

Eine Funktion ist genau dann konvex, wenn ihr Epigraph konvex ist.

Bsp: Sind $f_i : \mathbb{R}^n \to \bar{\mathbb{R}}$ konvex, so auch f mit $f(x) := \sup_i f_i(x)$, $x \in \mathbb{R}^n$.



Konvexe Funktionen

Eine Funktion $f : \mathbb{R}^n \to \bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ heißt konvex, wenn $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$ für $x, y \in \mathbb{R}^n$ und $\alpha \in [0,1]$.

Sie heißt streng konvex, wenn $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y)$ für $x, y \in \mathbb{R}^n$, $x \neq y$ und $\alpha \in (0,1)$.

Der Epigraph einer Funktion $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ ist die Menge

$$\operatorname{epi} f := \left\{ \begin{pmatrix} x \\ r \end{pmatrix} : x \in \mathbb{R}^n, r \geq f(x) \right\}$$ [die Punkte „oberhalb“von $f(x)$]

Satz

Eine Funktion ist genau dann konvex, wenn ihr Epigraph konvex ist.

Bsp: Sind $f_i : \mathbb{R}^n \to \bar{\mathbb{R}}$ konvex, so auch f mit $f(x) := \sup_i f_i(x)$, $x \in \mathbb{R}^n$.

Satz

Jedes lokale Minimum einer konvexen Funktion ist auch globales Minimum, und für streng konvexe Funktionen ist es das einzige.

Für konvexe Funktionen gibt es gute Optimierungsverfahren.

Konvexe Funktionen

Eine Funktion $f : \mathbb{R}^n \to \bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ heißt konvex, wenn $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$ für $x, y \in \mathbb{R}^n$ und $\alpha \in [0, 1]$.

Sie heißt streng konvex, wenn $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y)$ für $x, y \in \mathbb{R}^n$, $x \neq y$ und $\alpha \in (0, 1)$.

Der Epigraph einer Funktion $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ ist die Menge

$$\operatorname{epi} f := \left\{ \begin{pmatrix} x \\ r \end{pmatrix} : x \in \mathbb{R}^n, r \geq f(x) \right\}$$ [die Punkte „oberhalb“von $f(x)$]

Satz

Eine Funktion ist genau dann konvex, wenn ihr Epigraph konvex ist.

Bsp: Sind $f_i : \mathbb{R}^n \to \bar{\mathbb{R}}$ konvex, so auch f mit $f(x) := \sup_i f_i(x)$, $x \in \mathbb{R}^n$.

Satz

Jedes lokale Minimum einer konvexen Funktion ist auch globales Minimum, und für streng konvexe Funktionen ist es das einzige.

Für konvexe Funktionen gibt es gute Optimierungsverfahren.

Eine Funktion f heißt konkav, wenn $-f$ konvex ist.
(Jedes lokale Maximum einer konkaven Funktion ist auch globales.)

Inhaltsübersicht für heute:

Branch-and-Bound

Konvexe Mengen, konvexe Hülle, konvexe Funktionen

Relaxation

Relaxation

Konzept auf beliebige Optimierungsprobleme anwendbar (hier Maximieren):

Definition

Gegeben zwei Optimierungsprobleme mit $\mathcal{X}, \mathcal{W} \subseteq \mathbb{R}^n$ *und* $f, f' : \mathbb{R}^n \to \mathbb{R}$

$$(OP) \quad \max \; f(x) \text{ s.t. } x \in \mathcal{X} \qquad \textit{bzw.} \qquad (RP) \quad \max \; f'(x) \text{ s.t. } x \in \mathcal{W},$$

(RP) *heißt* *Relaxation* *von* (OP), *falls*

1. $\mathcal{X} \subseteq \mathcal{W}$,
2. $f(x) \leq f'(x)$ *für alle* $x \in \mathcal{X}$.

Relaxation

Konzept auf beliebige Optimierungsprobleme anwendbar (hier Maximieren):

Definition

Gegeben zwei Optimierungsprobleme mit $\mathcal{X}, \mathcal{W} \subseteq \mathbb{R}^n$ und $f, f' : \mathbb{R}^n \to \mathbb{R}$

$$(OP) \quad \max \; f(x) \text{ s.t. } x \in \mathcal{X} \qquad \textit{bzw.} \qquad (RP) \quad \max \; f'(x) \text{ s.t. } x \in \mathcal{W},$$

*(RP) heißt **Relaxation** von (OP), falls*

1. *$\mathcal{X} \subseteq \mathcal{W}$,*
2. *$f(x) \leq f'(x)$ für alle $x \in \mathcal{X}$.*

Sei (RP) eine Relaxation von (OP).

Beobachtung

1. *$v(RP) \geq v(OP)$.* *[(RP) liefert obere Schranke]*
2. *Ist (RP) unzulässig, so auch (OP),*
3. *Ist x^* OL von (RP) und gilt $x^* \in \mathcal{X}$ sowie $f'(x^*) = f(x^*)$, dann ist x^* OL von (OP).*

Relaxation

Konzept auf beliebige Optimierungsprobleme anwendbar (hier Maximieren):

Definition

Gegeben zwei Optimierungsprobleme mit $\mathcal{X}, \mathcal{W} \subseteq \mathbb{R}^n$ und $f, f' : \mathbb{R}^n \to \mathbb{R}$

$$(OP) \quad \max \; f(x) \text{ s.t. } x \in \mathcal{X} \qquad \textit{bzw.} \qquad (RP) \quad \max \; f'(x) \text{ s.t. } x \in \mathcal{W},$$

(RP) heißt Relaxation von (OP), falls

1. *$\mathcal{X} \subseteq \mathcal{W}$,*
2. *$f(x) \leq f'(x)$ für alle $x \in \mathcal{X}$.*

Sei (RP) eine Relaxation von (OP).

Beobachtung

1. *$v(RP) \geq v(OP)$.* *[(RP) liefert obere Schranke]*
2. *Ist (RP) unzulässig, so auch (OP),*
3. *Ist x^* OL von (RP) und gilt $x^* \in \mathcal{X}$ sowie $f'(x^*) = f(x^*)$, dann ist x^* OL von (OP).*

Man sucht nun möglichst „kleines“ $\mathcal{W} \supseteq \mathcal{X}$ und $f' \geq f$ so, dass (RP) noch gut lösbar ist.

Eine Relaxation (RP) von (OP) heißt exakt, falls $v(OP) = v(RP)$ gilt.

Allgemein: Konvexe Relaxation

Wird Konvexität für $\mathcal{W}$ und (bei max) Konkavität für f' gefordert, spricht man von konvexer Relaxation. Meist (aber nicht immer!) dient die Konvexität als Garant für vernünftige Lösbarkeit der Relaxation.

Allgemein: Konvexe Relaxation

Wird Konvexität für $\mathcal{W}$ und (bei max) Konkavität für f' gefordert, spricht man von konvexer Relaxation. Meist (aber nicht immer!) dient die Konvexität als Garant für vernünftige Lösbarkeit der Relaxation.

Bsp: Für ein kombinatorisches Max.-Problem mit endl. Grundmenge Ω, zulässigen Lösungen $\mathcal{F} \subseteq 2^\Omega$ und linearer Zielfunktion $c \in \mathbb{R}^\Omega$ ist

$$\max c^T x \text{ s.t. } x \in \operatorname{conv}\{\chi_\Omega(F) : F \in \mathcal{F}\}$$

eine exakte konvexe (sogar lineare) Relaxation, aber nur dann nützlich, wenn das Polyeder $\operatorname{conv}\{\chi(F) : F \in \mathcal{F}\}$ gut durch ein nicht zu großes Ungleichungssystem $Ax \leq b$ darstellbar ist.

Allgemein: Konvexe Relaxation

Wird Konvexität für $\mathcal{W}$ und (bei max) Konkavität für f' gefordert, spricht man von konvexer Relaxation. Meist (aber nicht immer!) dient die Konvexität als Garant für vernünftige Lösbarkeit der Relaxation.

Bsp: Für ein kombinatorisches Max.-Problem mit endl. Grundmenge Ω, zulässigen Lösungen $\mathcal{F} \subseteq 2^\Omega$ und linearer Zielfunktion $c \in \mathbb{R}^\Omega$ ist

$$\max c^T x \text{ s.t. } x \in \operatorname{conv}\{\chi_\Omega(F) : F \in \mathcal{F}\}$$

eine exakte konvexe (sogar lineare) Relaxation, aber nur dann nützlich, wenn das Polyeder $\operatorname{conv}\{\chi(F) : F \in \mathcal{F}\}$ gut durch ein nicht zu großes Ungleichungssystem $Ax \leq b$ darstellbar ist.

In der globalen Optimierung werden auf Teilintervallen nichtlineare Funktionen nach unten durch konvexe Funktionen abgeschätzt.

Bsp: Betrachte (OP) $\min f(x) := \frac{1}{2}x^T Qx + q^T x$ s.t. $x \in [0,1]^n$ mit f nicht konvex, d.h., $\lambda_{\min}(Q) < 0$. [$\lambda_{\min}\ldots$ minimaler Eigenwert]
Es ist $Q - \lambda_{\min}(Q)I$ positiv semidefinit und wegen $x_i^2 \leq x_i$ auf $[0,1]^n$ gilt
$f'(x) := \frac{1}{2}x^T(Q - \lambda_{\min}(Q)I)x + (q + \lambda_{\min}(Q)\mathbf{1})^T x \;\leq\; f(x) \quad \forall x \in [0,1]^n$.
Damit ist (RP) $\min f'(x)$ s.t. $x \in [0,1]^n$ eine konvexe Relaxation von (OP).

Die LP-Relaxation für Ganzzahlige Programme

Für ein ganzzahliges Programm $\max\ c^T x$ s.t. $Ax \le b,\ x \in \mathbb{Z}^n$
entsteht die LP-Relaxation durch Weglassen der Ganzz.-Bedingung,

$$\max\ c^T x \text{ s.t. } Ax \le b,\ x \in \mathbb{R}^n.$$

Ist Relaxation, denn

$$\mathcal{X} := \{x \in \mathbb{Z}^n : Ax \le b\} \ \subseteq\ \{x \in \mathbb{R}^n : Ax \le b\} =: \mathcal{W}.$$

[wird von allen Standardlösern für gemischt-ganzz. Programme verwendet]

Die LP-Relaxation für Ganzzahlige Programme

Für ein ganzzahliges Programm $\max\ c^T x \text{ s.t. } Ax \leq b,\ x \in \mathbb{Z}^n$
entsteht die LP-Relaxation durch Weglassen der Ganzz.-Bedingung,

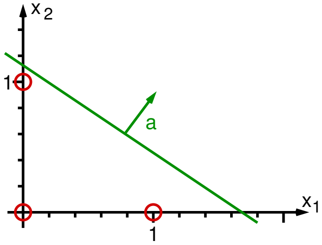
$$\max\ c^T x \text{ s.t. } Ax \leq b,\ x \in \mathbb{R}^n.$$

Ist Relaxation, denn

$$\mathcal{X} := \{x \in \mathbb{Z}^n : Ax \leq b\} \ \subseteq\ \{x \in \mathbb{R}^n : Ax \leq b\} =: \mathcal{W}.$$

[wird von allen Standardlösern für gemischt-ganzz. Programme verwendet]

Bsp: Rucksackproblem: $n = 2$, Gewichte $a = (6, 8)^T$, Kapazität $b = 10$,
$\max\ c^T x \text{ s.t. } a^T x \leq b,\ x \in \mathbb{Z}^n_+ \quad \rightarrow \quad \max\ c^T x \text{ s.t. } a^T x \leq b,\ x \geq 0,$



zulässige ganzzahlige Punkte: ◯

Die LP-Relaxation für Ganzzahlige Programme

Für ein ganzzahliges Programm $\max\ c^T x$ s.t. $Ax \le b,\ x \in \mathbb{Z}^n$
entsteht die LP-Relaxation durch Weglassen der Ganzz.-Bedingung,

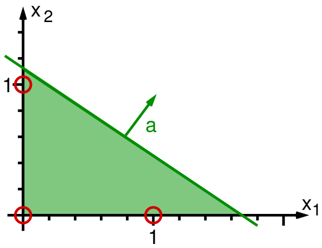
$$\max\ c^T x \text{ s.t. } Ax \le b,\ x \in \mathbb{R}^n.$$

Ist Relaxation, denn

$$\mathcal{X} := \{x \in \mathbb{Z}^n : Ax \le b\} \subseteq \{x \in \mathbb{R}^n : Ax \le b\} =: \mathcal{W}.$$

[wird von allen Standardlösern für gemischt-ganzz. Programme verwendet]

Bsp: Rucksackproblem: $n = 2$, Gewichte $a = (6,8)^T$, Kapazität $b = 10$,
$\max\ c^T x$ s.t. $a^T x \le b,\ x \in \mathbb{Z}^n_+ \quad \rightarrow \quad \max\ c^T x$ s.t. $a^T x \le b,\ x \ge 0$,



zulässige ganzzahlige Punkte: ◯

LP-Relaxation: grün

Die LP-Relaxation für Ganzzahlige Programme

Für ein ganzzahliges Programm $\max\ c^T x$ s.t. $Ax \le b,\ x \in \mathbb{Z}^n$
entsteht die LP-Relaxation durch Weglassen der Ganzz.-Bedingung,

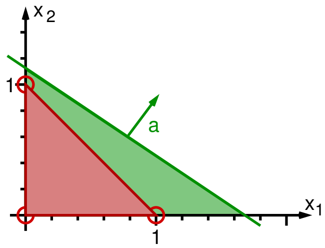
$$\max\ c^T x \text{ s.t. } Ax \le b,\ x \in \mathbb{R}^n.$$

Ist Relaxation, denn

$$\mathcal{X} := \{x \in \mathbb{Z}^n : Ax \le b\} \ \subseteq\ \{x \in \mathbb{R}^n : Ax \le b\} =: \mathcal{W}.$$

[wird von allen Standardlösern für gemischt-ganzz. Programme verwendet]

Bsp: Rucksackproblem: $n = 2$, Gewichte $a = (6, 8)^T$, Kapazität $b = 10$,
$\max\ c^T x$ s.t. $a^T x \le b,\ x \in \mathbb{Z}^n_+ \quad \rightarrow \quad \max\ c^T x$ s.t. $a^T x \le b,\ x \ge 0$,



zulässige ganzzahlige Punkte: ◯

LP-Relaxation: grün

beste Relaxation: die konvexe Hülle

Die LP-Relaxation für Ganzzahlige Programme

Für ein ganzzahliges Programm $\max\ c^T x \text{ s.t. } Ax \le b,\ x \in \mathbb{Z}^n$
entsteht die LP-Relaxation durch Weglassen der Ganzz.-Bedingung,

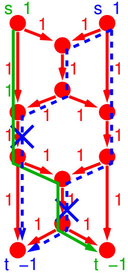
$$\max\ c^T x \text{ s.t. } Ax \le b,\ x \in \mathbb{R}^n.$$

Ist Relaxation, denn

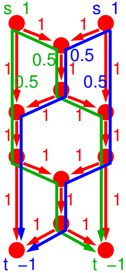
$$\mathcal{X} := \{x \in \mathbb{Z}^n : Ax \le b\} \ \subseteq\ \{x \in \mathbb{R}^n : Ax \le b\} =: \mathcal{W}.$$

[wird von allen Standardlösern für gemischt-ganzz. Programme verwendet]

Bsp: ganzzahliger Mehrgüterfluss $\rightarrow$ gebrochener Mehrgüterfluss



ganzz. geht nicht, $\mathcal{X} = \emptyset$



gebrochen geht, $\mathcal{W} \neq \emptyset$

auch zu groß, bräuchten die konvexe Hülle

Lagrange-Relaxation

[Allg. für restr. Opt. verwendbar, hier vorerst nur Unglgs.-Nebenbed.]

Unbequeme Nebenbedingungen werden mit einem Lagrangemultiplikator, der die Verletzung bestraft, in die Zielfunktion gehoben ($g : \mathbb{R}^n \to \mathbb{R}^k$):

$$(OP) \quad \begin{array}{ll} \max & f(x) \\ \text{s.t.} & g(x) \leq 0 \quad | \cdot \lambda \geq 0 \\ & x \in \Omega \end{array} \quad \to \quad (RP_\lambda) \quad \begin{array}{ll} \max & c^T x - \lambda^T g(x) \\ \text{s.t.} & x \in \Omega \end{array}$$

Ist Relaxation: $\mathcal{X} := \{x \in \Omega : g(x) \leq 0\} \subseteq \{x \in \Omega\} =: \mathcal{W}$ und für $x \in \mathcal{X}$, $\lambda \geq 0$ gilt $f(x) \leq f(x) - \lambda^T g(x) =: f'(x)$ wegen $g(x) \leq 0$.

Lagrange-Relaxation

[Allg. für restr. Opt. verwendbar, hier vorerst nur Unglgs.-Nebenbed.]

Unbequeme Nebenbedingungen werden mit einem Lagrangemultiplikator, der die Verletzung bestraft, in die Zielfunktion gehoben ($g : \mathbb{R}^n \to \mathbb{R}^k$):

$$(OP) \quad \begin{array}{ll} \max & f(x) \\ \text{s.t.} & g(x) \leq 0 \quad | \cdot \lambda \geq 0 \\ & x \in \Omega \end{array} \quad \rightarrow \quad (RP_\lambda) \quad \begin{array}{ll} \max & c^T x - \lambda^T g(x) \\ \text{s.t.} & x \in \Omega \end{array}$$

Ist Relaxation: $\mathcal{X} := \{x \in \Omega : g(x) \leq 0\} \subseteq \{x \in \Omega\} =: \mathcal{W}$ und für $x \in \mathcal{X}$, $\lambda \geq 0$ gilt $f(x) \leq f(x) - \lambda^T g(x) =: f'(x)$ wegen $g(x) \leq 0$.

Definiere die duale Funktion $\varphi(\lambda) := \sup_{x \in \Omega} \left[f(x) - g(x)^T \lambda \right] = v(RP_\lambda)$

[für jedes feste x linear in λ]

- Für jedes $\lambda \geq 0$ gilt $\varphi(\lambda) \geq v(OP)$ [obere Schranke]
- $\varphi(\lambda)$ gut berechenbar, falls (RP_λ) „leicht“ lösbar
- φ ist als sup von in λ linearen Funktionen konvex
- Beste Schranke ist $\inf\{\varphi(\lambda) : \lambda \geq 0\}$ [konvexes Problem!]

gut mit Verfahren der konvexen Optimierung bestimmbar!

Beispiel: Ganzzahliger Mehrgüterfluss

A sei die Knoten-Kanten-Inz.-Matrix zu $D=(V,E)$, 2 Güter,
Lagrange-Relaxation der koppelnden Kapazitätsbedingungen mit $\lambda \geq 0$:

$$
\begin{array}{llcll}
 & \min\ {c^{(1)}}^T x^{(1)} & + & {c^{(2)}}^T x^{(2)} & \\
 & \text{s.t.}\quad A x^{(1)} & & & = b^{(1)} \\
 & & & A x^{(2)} & = b^{(2)} \\
\lambda \cdot | & \quad x^{(1)} & + & x^{(2)} & \leq w \\
 & \quad x^{(1)} \leq w, & & x^{(2)} \leq w & \\
 & \quad x^{(1)} \in \mathbb{Z}^E_+, & & x^{(2)} \in \mathbb{Z}^E_+. &
\end{array}
\quad \rightarrow \quad
\begin{array}{lllr}
\min & (c^{(1)}+\lambda)^T x^{(1)} + & (c^{(2)}+\lambda)^T x^{(2)} - \lambda^T w & \\
\text{s.t.} & \quad A x^{(1)} & & = b^{(1)} \\
 & & A x^{(2)} & = b^{(2)} \\
 & x^{(1)} \leq w, & x^{(2)} \leq w, & \\
 & x^{(1)} \in \mathbb{Z}^E_+, & x^{(2)} \in \mathbb{Z}^E_+. &
\end{array}
$$

Relaxation zerfällt in zwei unabhängige Minimale-Kosten-Fluss-Probleme

$$(RP^{(i)}_\lambda) \quad \min\,(c^{(i)}+\lambda)^T x^{(i)} \quad \text{s.t.} \quad A x^{(i)} = b^{(i)},\ w \geq x^{(i)} \in \mathbb{Z}^E_+ \quad i \in \{1,2\}$$

Diese sind effizient ganzzahlig lösbar!

Beispiel: Ganzzahliger Mehrgüterfluss

A sei die Knoten-Kanten-Inz.-Matrix zu $D = (V, E)$, 2 Güter, Lagrange-Relaxation der koppelnden Kapazitätsbedingungen mit $\lambda \geq 0$:

$$\begin{array}{lllll}
& \min {c^{(1)}}^T x^{(1)} & + \; {c^{(2)}}^T x^{(2)} & & \\
& \text{s.t.} \quad Ax^{(1)} & & = & b^{(1)} \\
& & \quad Ax^{(2)} & = & b^{(2)} \\
\lambda \cdot | & \quad x^{(1)} & + \quad x^{(2)} & \leq & w \\
& x^{(1)} \leq w, & x^{(2)} \leq w & & \\
& x^{(1)} \in \mathbb{Z}^E_+, & x^{(2)} \in \mathbb{Z}^E_+. & &
\end{array}
\quad \rightarrow \quad
\begin{array}{llll}
\min & (c^{(1)} + \lambda)^T x^{(1)} + (c^{(2)} + \lambda)^T x^{(2)} - \lambda^T w & & \\
\text{s.t.} & Ax^{(1)} & & = b^{(1)} \\
& & Ax^{(2)} & = b^{(2)} \\
& x^{(1)} \leq w, & x^{(2)} \leq w, & \\
& x^{(1)} \in \mathbb{Z}^E_+, & x^{(2)} \in \mathbb{Z}^E_+. &
\end{array}$$

Relaxation zerfällt in zwei unabhängige Minimale-Kosten-Fluss-Probleme

$$(RP^{(i)}_\lambda) \quad \min\,(c^{(i)} + \lambda)^T x^{(i)} \quad \text{s.t.} \quad Ax^{(i)} = b^{(i)},\; w \geq x^{(i)} \in \mathbb{Z}^E_+ \quad i \in \{1, 2\}$$

Diese sind effizient ganzzahlig lösbar!

Zerfällt das Problem bei Lagrange-Relaxation in unabhängige Teilprobleme, nennt man das Lagrange-Dekomposition. Damit können oft deutlich größere Probleme sehr effizient gelöst werden.

Bekommt man dadurch auch eine bessere Schranke?

Vergleich Lagrange- und LP-Relaxation

Sei $\Omega \subset \mathbb{Z}^n$ endlich und $D \in \mathbb{Q}^{k \times n}$, $d \in \mathbb{Q}^k$,

$$(OP) \quad \begin{array}{ll} \max & c^T x \\ \text{s.t.} & Dx \leq d \quad | \cdot \lambda \geq 0 \\ & x \in \Omega \end{array} \quad \rightarrow \quad (RP_\lambda) \quad \begin{array}{ll} \max & c^T x + \lambda^T (d - Dx) \\ \text{s.t.} & x \in \Omega \end{array}$$

Im Bsp: $\Omega = \Omega^{(1)} \times \Omega^{(2)}$ mit $\Omega^{(i)} = \{x \in \mathbb{Z}_+^E : Ax = b^{(i)}, x \leq w\}$, $i \in \{1, 2\}$.

Satz

$$\inf_{\lambda \geq 0} v(RP_\lambda) = \sup\{c^T x : Dx \leq d, x \in \operatorname{conv} \Omega\}.$$

Ist $\operatorname{conv} \Omega$ gleich der zulässigen Menge der LP-Relaxation von Ω, sind die Werte der besten Lagrange- und der LP-Relaxation gleich!

Vergleich Lagrange- und LP-Relaxation

Sei $\Omega \subset \mathbb{Z}^n$ endlich und $D \in \mathbb{Q}^{k\times n}$, $d \in \mathbb{Q}^k$,

$$(OP)\quad \begin{array}{ll} \max & c^T x \\ \text{s.t.} & Dx \le d \quad \color{blue}{| \cdot \lambda \ge 0} \\ & x \in \Omega \end{array} \qquad \rightarrow \qquad (RP_\lambda)\quad \begin{array}{ll} \max & c^T x + \lambda^T (d - Dx) \\ \text{s.t.} & x \in \Omega \end{array}$$

Im Bsp: $\Omega = \Omega^{(1)} \times \Omega^{(2)}$ mit $\Omega^{(i)} = \{x \in \mathbb{Z}_+^E : Ax = b^{(i)}, x \le w\}$, $i \in \{1, 2\}$.

Satz

$$\inf_{\lambda \ge 0} v(RP_\lambda) = \sup\{c^T x : Dx \le d, x \in \operatorname{conv} \Omega\}.$$

Ist $\operatorname{conv} \Omega$ gleich der zulässigen Menge der LP-Relaxation von Ω, sind die Werte der besten Lagrange- und der LP-Relaxation gleich!

Im Bsp ist A total unimodular, daher gilt für $i \in \{1, 2\}$ und $w \in \mathbb{Z}^E$

$$\operatorname{conv}\{x \in \mathbb{Z}_+^E : Ax = b^{(i)}, x \le w\} = \{x \ge 0 : Ax = b^{(i)}, x \le w\}.$$

Das beste λ ergibt den Wert des gebrochenen Mehrgüterflussproblems!

Vergleich Lagrange- und LP-Relaxation

Sei $\Omega \subset \mathbb{Z}^n$ endlich und $D \in \mathbb{Q}^{k \times n}$, $d \in \mathbb{Q}^k$,

$$(OP) \quad \begin{array}{ll} \max & c^T x \\ \text{s.t.} & Dx \le d \quad | \cdot \lambda \ge 0 \\ & x \in \Omega \end{array} \quad \rightarrow \quad (RP_\lambda) \quad \begin{array}{ll} \max & c^T x + \lambda^T (d - Dx) \\ \text{s.t.} & x \in \Omega \end{array}$$

Im Bsp: $\Omega = \Omega^{(1)} \times \Omega^{(2)}$ mit $\Omega^{(i)} = \{x \in \mathbb{Z}_+^E : Ax = b^{(i)}, x \le w\}$, $i \in \{1, 2\}$.

Satz

$$\inf_{\lambda \ge 0} v(RP_\lambda) = \sup\{c^T x : Dx \le d, x \in \operatorname{conv} \Omega\}.$$

Ist $\operatorname{conv} \Omega$ gleich der zulässigen Menge der LP-Relaxation von Ω, sind die Werte der besten Lagrange- und der LP-Relaxation gleich!

Im Bsp ist A total unimodular, daher gilt für $i \in \{1, 2\}$ und $w \in \mathbb{Z}^E$

$$\operatorname{conv}\{x \in \mathbb{Z}_+^E : Ax = b^{(i)}, x \le w\} = \{x \ge 0 : Ax = b^{(i)}, x \le w\}.$$

Das beste λ ergibt den Wert des gebrochenen Mehrgüterflussproblems!

Allgemein: Sei $\{x \in \mathbb{Z}^n : Ax \le b\} = \Omega$ eine Formulierung von Ω. Nur falls $\{x \in \mathbb{R}^n : Ax \le b\} \ne \operatorname{conv} \Omega$, kann die Lagrange-Relaxation einen besseren Wert liefern als die LP-Relaxation.