5.) Ein ROUTING Problem (Anwendung in Rechnernetzen)

Eingabe: ein gerichteter Graph $\vec{G}(V, \vec{E})$

mit ausgezeichneten Knoten

$s_1\ s_2 \ldots \ldots s_r$ (Quellen)

$t_1\ t_2 \ldots \ldots t_r$ (Senken)

Ausgabe: Bestimme gerichtete Pfade $P_1\ P_2 \ldots P_r$, so dass

- P_i in s_i beginnt, und in t_i endet
- die maximale Belastung überalle Kanten minimal ist.

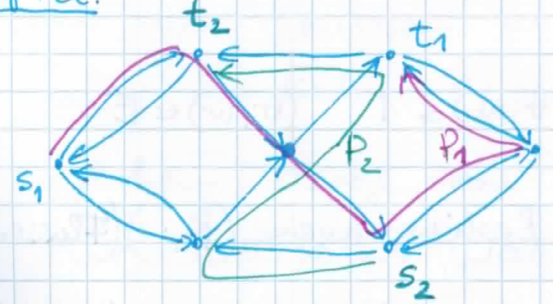
Belastung (e) = Anzahl der P_i, die über e laufen

eine IP-Formulierung des Problems

die Variablen: $x_i(e)$ für $1 \le i \le r$ und $e \in \vec{E}$

Bedeutung: $x_i(e) = \begin{cases} 1 & \text{wenn } P_i \text{ über e läuft} \\ 0 & \text{sonst} \end{cases}$

Beispiel:

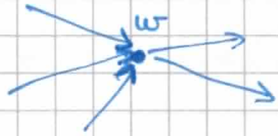


Sei i fixiert

Um auszudrücken, dass die $x_i(e)$ $(e \in E)$ einem $s_i - t_i$ Pfad P_i entsprechen sollen, haben wir für jeden Knoten $w \in V$ eine Bedingung (der Pfad P_i darf Knoten wiederholt besuchen):

Wenn $w \neq s_i, t_i$, dann soll P_i den Knoten w jedes mal auch verlassen, wenn er ihn besucht:

$$\sum_{\{e \mid e \text{ führt in } w\}} x_i(e) - \sum_{\{e \mid e \text{ startet in } w\}} x_i(e) = 0 \qquad \forall w \neq s_i, t_i \quad w \in V$$



$|V|$ viele Flusserhaltung Bedingungen

(B_i)

$$\sum_{\{e \mid e \text{ startet in } s_i\}} x_i(e) - \sum_{\{e \mid e \text{ führt in } s_i\}} x_i(e) = 1 \qquad \text{für } s_i$$

$$\sum_{\{e \mid e \text{ führt in } t_i\}} x_i(e) - \sum_{\{e \mid e \text{ startet in } t_i\}} x_i(e) = 1 \qquad \text{für } t_i$$

Die (B_i) Bedingungen brauchen wir für jedes $s_i - t_i$ Knotenpaar, diese ergeben insgesamt $|V| \cdot r$ Bedingungen (für $1 \leq i \leq r$ $w \in V$).

(Unsere Zielfunktion ist eigentlich:

Minimiere $\max\limits_{e} \underbrace{\sum_{i=1}^{r} x_i(e)}_{\text{(Belastung über Kante } e)}$ Diese ist aber keine lineare Zielfunktion! Wir führen eine neue Variable W ein die der Maximalen Belastung entspricht.)

LP 49/a

Die Zielfunktion ist dann

Minimiere W

mit den weiteren Bedingungen ($|E|$ Bedingungen):

$$\sum_{i=1}^{r} x_i(e) \leq W \qquad \forall e \in E$$

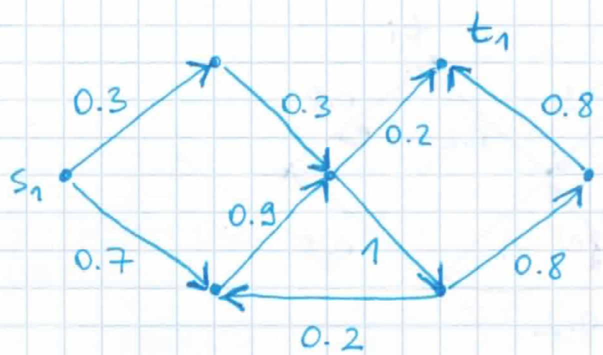
$$x_i(e) \in \{0, 1\}$$

LP-Relaxierung: $0 \leq x_i(e) \leq 1 \qquad \forall e \in E$

In der LP-Relaxierung die Bedingungen (B_i) für Flusserhaltung fordern, dass von s_i nach t_i ein Fluss von Wert 1 entsteht (eine fraktionale Lösung heißt also ‚Fluss'):

(Definition: ein Fluss von s_i nach t_i ist eine Funktion $f: E \to \mathbb{R}_{0,+}$ (hier $f = x_i$) so dass die Flusserhaltung (B_i) gilt (es gibt meistens auch Kapazitätsschranken (im allgemeinen Flussproblem) $f(e) \leq c(e)$ für jede Kante e)).

Beispiel: der folgende Fluss erfüllt die Bedingungen (B_1) (eine fraktionale Lösung der $x_1(e)$ Werte)



← (nicht optimal)

Eine (optimale) Lösung der LP-Relaxierung bestimmt also jeweils einen Fluss x_i von s_i nach t_i.

Die Idee: bestimmen wir eine ganzzahlige Lösung, dh. jeweils einen einzigen Pfad P_i aus dieser optimalen fraktionalen Lösung durch randomisiertes Runden.

⟶ wir müssen zufällig einen Pfad P_i aus mehreren möglichen Pfaden auswählen. ABER: aus welcher Menge von Pfaden?

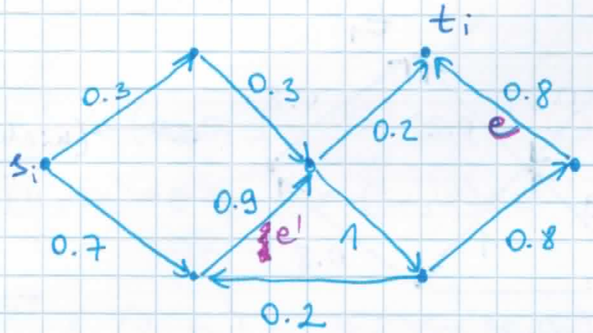
Mit welchen Wahrscheinlichkeiten?

Wie entstehen mögliche Pfade und Wahrscheinlichkeiten aus einer fraktionalen Lösung?

Pfad-Zerlegung (~~...~~ path-decomposition)

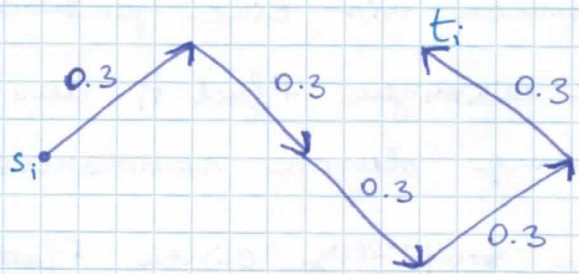
__Beispiel:__ Wir bestimmen eine Pfad-Zerlegung der obigen fraktionalen Lösung. (~~Wir lassen~~ Fixiere den Index $i (= 1)$ ~~von der Notation weg.~~)

- Betrachten wir den Teilgraph aller Kanten mit $x_i(e) > 0$.



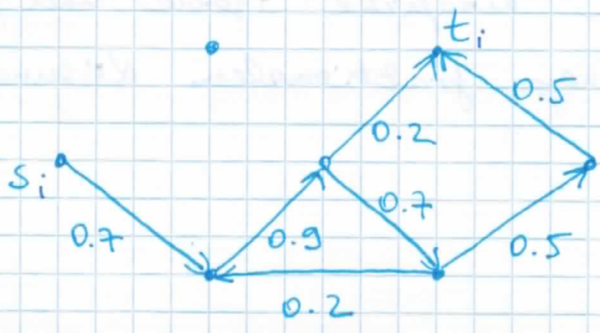
- Wir nehmen einen Pfad P_1 von s_i nach t_i und α_1 den minimalen Wert des Flusses entlang des Pfades.
  (es gibt einen s_i–t_i Pfad wegen der Flusserhaltung)

$P_{(1)}$



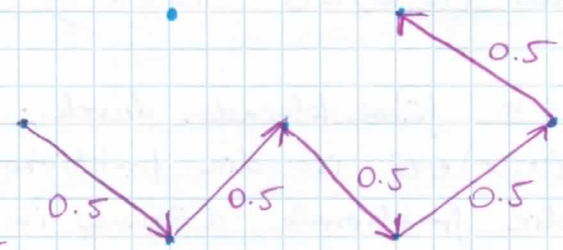
$\alpha_1 = 0.3$

schließlich reduzieren wir die entsprechenden Flusswerte um $0.3 = \alpha_1$ und erhalten einen Fluss mit Wert $1 - \alpha_1$ von s_i nach t_i

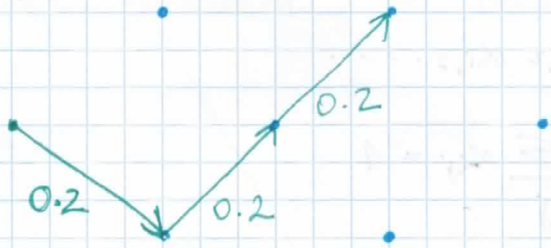


- wir iterieren diesen Prozess solange es einen $s_i - t_i$ Pfad mit positivem Flusswert gibt
  $\Longleftrightarrow$ Solange $\alpha_1 + \alpha_2 + \ldots < 1$

P_{i2}
$\alpha_2 = 0.5$

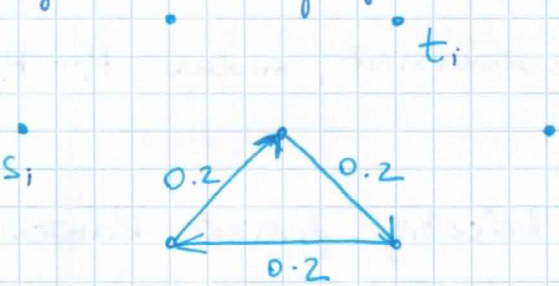


P_{i3}
$\alpha_3 = 0.2$



Es gilt: $\alpha_1 + \alpha_2 + \alpha_3 = 1$

- der gebliebene Graph enthält keinen s-t Pfad



- Wieviele Iterationen gibt es?
  → in jeder Runde wird mindestens eine Kante aus dem Graphen entfernt, und keine Kante kommt hinzu.
  ⇒ es gibt höchstens $|E|$ Iterationen.
- Wie soll der Algorithmus randomisiert einen Pfad von s_i nach t_i auswählen? → mit Prob. α_j wird P_{ij} gewählt.
- die Prob., dass Kante e benutzt wird ist dann $0.5 + 0.3 = 0.8 = x_i(e)$

- die erwartete Belastung von e <u>wegen P_i</u> ist somit

$1 \cdot 0.8 + 0 \cdot 0.2 = 1 \cdot x_i(e)$ (allgemein $\leq x_i(e)$) z.B für e' Prob $= 0.7 < x_i(e)$

Algorithmus für ROUTING (randomisiertes Runden)

→ löse die LP-Relaxierung; seien die Werte einer optimalen fraktionalen Lösung $x_i^*(e)$ (für $e \in E$, $1 \le i \le r$)

→ für jede $1 \le i \le r$ führe Folgendes durch:

- $x_i^*(e)$ $(e \in E)$ ist die fraktionale Lösung (ein Fluss) für $s_i - t_i$
- zerlege diese fraktionale Lösung in Pfade $P_{i,1}$ $P_{i,2}$ $P_{i,3}$ .... mit den entsprechenden Fluss-Werten $\alpha_{i,1}$ $\alpha_{i,2}$ $\alpha_{i,3}$ ....
- es gilt $\sum_j \alpha_{i,j} = 1$
- wähle den $s_i - t_i$ Pfad P_i randomisiert, so dass

$$P_i := P_{i,j} \text{ mit Wahrscheinlichkeit } \alpha_{i,j}$$

→ Da wir eine Pfad-Zerlegung des Flusses $x_i^*(e)$ $(e \in E)$ haben, gilt für jede einzelne Kante $e \in E$, dass

$$\sum_{\{j \mid e \in P_{i,j}\}} \alpha_{i,j} \le x_i^*(e)$$

↑ Flusswerte aller Pfade, die die Kante e benutzen

→ Die Wahrscheinlichkeit, dass der zufällig gewählte P_i eine konkrete Kante e enthält, ist genau

$$\text{Prob}(P_i \text{ benutzt } e) = \sum_{\{j \mid e \in P_{i,j}\}} \alpha_{i,j} \le x_i^*(e)$$

(die Flusswerte =) die Wahrscheinlichkeiten der Pfade, die die Kante e benutzen

→ Die erwartete Belastung über Kante e <u>wegen des Pfades P_i</u> (!) ist

(die durch P_i verursachte Belastung falls P_i benutzte e ist 1, sonst ist es 0)

$$1 \cdot \text{Prob}(P_i \text{ benutzte}) + 0 \cdot \text{Prob}(P_i \text{ benutzte nicht}) =$$

$$= \text{Prob}(P_i \text{ benutzte}) \leq x_i^*(e)$$

→ Die erwartete Belastung über e wegen aller Pfade $P_1\, P_2\, P_3 \dots P_i \dots P_r$ ist

$$\leq x_1^*(e) + x_2^*(e) + \dots + x_r^*(e) \leq W_{OPT}^{frac} \leq W_{OPT}$$

[Genauer: die Zufallsvariablen werden definiert. (e fixiert)

$$y_i = \begin{cases} 1 & \text{falls } P_i \text{ benutzte e} \\ 0 & \text{sonst} \end{cases} \quad \text{(Indikatorvariablen)}$$

$$E[y_i] = 1 \cdot \text{Prob}(P_i \text{ benutzte e}) + 0 \cdot \text{Prob}(\text{nicht})$$

$y = \sum_{i=1}^{r} y_i$ ist die Belastung über Kante e insgesamt (y ist auch Zufallsvariable).

Es gilt $E[y] = \sum_{i=1}^{r} E[y_i] \leq \sum_{i=1}^{r} x_i^*(e) \leq W_{OPT}^{frac} \leq W_{OPT}$]

Thm: Die maximale <u>erwartete</u> Belastung über alle Kanten e (max über Kanten) ist somit $\leq W_{OPT}$.

<u>Bemerkung</u>: Die erwartete maximale Belastung über Kanten (Erwartung über jeweils einer schlechtesten Kante) kann trotzdem groß sein. Aber, mit Prob $> 1-\varepsilon$ ist die maximale Belastung $\leq W_{OPT} + \sqrt{3 W_{OPT} \cdot \ln \frac{n^2}{\varepsilon}}$

Dualität in der Linearen Programmierung

Einführendes Beispiel: (Vazirani 12.1)

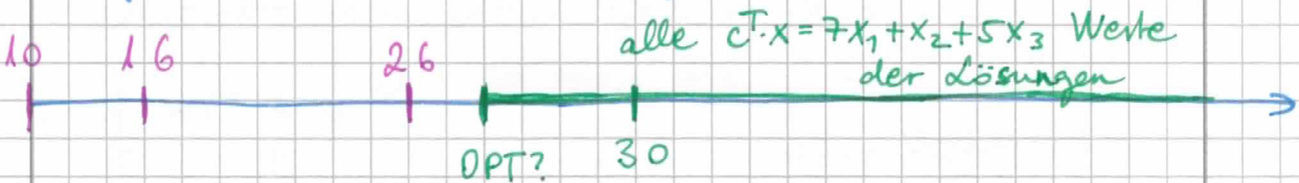
Wir betrachten das folgende LP in kanonischer Form:

minimiere $\quad 7x_1 + x_2 + 5x_3$

so dass

$$(P) \qquad \begin{aligned} x_1 - x_2 + 3x_3 &\geq 10 \\ 5x_1 + 2x_2 - x_3 &\geq 6 \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

Statt dieses LP zu lösen, werden wir schnell (obere und) untere Schranken für den optimalen Wert OPT finden



→ Jede Lösung resultiert eine obere Schranke für OPT:

z.B. $\quad x = (2, 1, 3)$

→ erfüllt die Nebenbedingungen:

$$x_1, x_2, x_3 \geq 0 \checkmark$$

$$2 - 1 + 3 \cdot 3 \geq 10 \checkmark$$

$$5 \cdot 2 + 2 \cdot 1 - 3 = 9 \geq 6 \checkmark$$

→ und hat Zielwert

$$7x_1 + x_2 + 5x_3 =$$

$$= 7 \cdot 2 + 1 + 5 \cdot 3 = 30$$

→ Wie findet man gute <u>untere Schranken</u> für OPT?

1.) Eine einfache untere Schranke:

$$10 \leq x_1 - x_2 + 3x_3 \leq 7x_1 + x_2 + 5x_3$$

Nebenbedingung

gilt Term für Term, weil $x_1, x_2, x_3 \geq 0$ und für die Koeffizienten gilt die Ungleichung einzeln

10 ist eine untere Schranke für den Zielwert jeder Lösung, also auch für OPT

2.) eine bessere untere Schranke:

$$16 = 10 + 6 \leq \overbrace{(x_1 - x_2 + 3x_3) + (5x_1 + 2x_2 - x_3)}^{6x_1 + x_2 + 2x_3} \leq 7x_1 + x_2 + 5x_3$$

laut Nebenbedingungen

gilt Term für Term

3.) eine noch bessere untere Schranke für OPT

$$26 = 2 \cdot 10 + 1 \cdot 6 \leq 2 \cdot (x_1 - x_2 + 3x_3) + 1 \cdot (5x_1 + 2x_2 - x_3) \leq 7x_1 + x_2 + 5x_3$$

Nebenbedingungen

gilt Term für Term für die Koeffizienten von $x_1, x_2, x_3 \geq 0$

Allgemeine Idee: finden wir Multiplikatoren $y_1 \geq 0$ und $y_2 \geq 0$ so dass

$$\underbrace{y_1 \cdot 10 + y_2 \cdot 6}_{y^T \cdot b} \leq \underbrace{y_1 \cdot (x_1 - x_2 + 3x_3) + y_2 \cdot (5x_1 + 2x_2 - x_3)}_{y^T \cdot A \cdot x} \leq \underbrace{7x_1 + x_2 + 5x_3}_{c^T \cdot x}$$

dann ist $y_1 \cdot 10 + y_2 \cdot 6$ eine untere Schranke für OPT

Wir brauchen also:

$$y_1, y_2 \geq 0$$

Und, wegen der Koeffizienten von x_1, x_2, x_3

$$\begin{aligned} y_1 + 5y_2 &\leq 7 \\ -y_1 + 2y_2 &\leq 1 \qquad (D) \\ 3y_1 + (-y_2) &\leq 5 \end{aligned}$$

und $10 \cdot y_1 + 6y_2$ so groß wie möglich, i.e. wir maximieren $10y_1 + 6y_2$ unter den obigen Bedingungen.

Dies ist ein anderes LP, ein Maximierungsproblem.

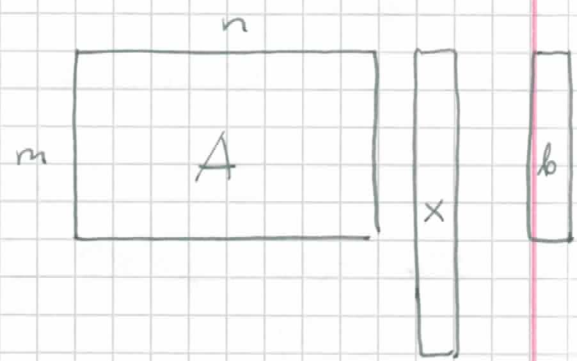
(D) ist das sog. duale LP von (P). (und umgekehrt: (P) ist auch das duale LP von (D)).

g.) Dualität in der linearen Programmierung

i.) Einführung

Beispiel: Sei das folgende LP gegeben (wir nennen ihn (P)):

(P) minimiere $c^T \cdot x$ so dass $A \cdot x \geq b$ und $x \geq 0$



I.

Für beliebigen nichtnegativen m-dimensionalen Vektor $y \geq 0$

$$y^T = (y_1 \; y_2 \; \dots \; y_m) \quad (y_i \geq 0 \; \forall i)$$

und jede Lösung x gilt dann

(denke an y_i und x_j als an Zahlen)

$$b \leq A \cdot x$$

$$y^T \cdot b \leq y^T \cdot A \cdot x$$

($b \leq A \cdot x$ ist die kurze Form der Ungleichungen

$$b_1 \leq a_1^T \cdot x = \sum_{j=1}^{n} a_{1j} \cdot x_j$$

$$b_2 \leq a_2^T \cdot x$$

$$\vdots$$

$$b_m \leq a_m^T \cdot x$$

Wenn die i-te Ungleichung durch $y_i \geq 0$ multipliziert wird, erhalten wir $y_i \cdot b_i \leq y_i \cdot (a_i^T \cdot x)$

LPD 2.

Wenn die letzteren m Ungleichungen dann summiert werden:

$$y^T \cdot b = \sum_{i=1}^{m} y_i \cdot b_i \leq \sum_{i=1}^{m} y_i \, (a_i^T \cdot x) = y^T \cdot A \cdot x \quad)$$

II. Für jede Lösung x von (P) gilt $x \geq 0$.

Wenn noch für ein $y \geq 0$ auch $y^T \cdot A \leq c^T$ gilt, dann

$$y^T \cdot A \cdot x \leq c^T \cdot x$$

Also, für alle $y \in \mathbb{R}^m$ so dass $y \geq 0$ und $y^T \cdot A \leq c^T$

$$y^T \cdot b \leq y^T \cdot A \cdot x \leq c^T \cdot x$$

dh. $y^T \cdot b$ ist eine untere Schranke für jeden (auch für den optimalen) Zielwert von (P)

Suchen wir eine beste (höchste) solche untere Schranke, dann müssen wir das folgende Problem lösen:

$$\text{(D)} \qquad \begin{aligned} &\text{maximiere} && y^T \cdot b \\ &\text{so dass} && y^T \cdot A \leq c^T \\ &\text{und} && y \geq 0 \end{aligned}$$

Das ist auch ein LP, wir nennen ihn (D)
(P) steht für 'Primal', (D) steht für 'Dual'

(Statt $y^T \cdot b$ hätten wir ruhig $b^T \cdot y$ und statt $y^T \cdot A \leq c^T$ $A^T \cdot y \leq c$ schreiben können, da y die Variablen von (D) enthält. Es wird bequemer jedoch mit der Notation y^T zu bleiben.)

Was hätte sich im obigen Argument geändert, wenn wir weiterhin $y^T \cdot b \leq c^T \cdot x$ gewollt hätten,

aber im (P)... in (D)

(statt $Ax \geq b$)

→ $Ax \leq b$ steht → dann sollte $y \leq 0$ sein, so gilt

$y^T \cdot b \leq y^T \cdot A \cdot x$ in I.

→ $Ax = b$ steht → dann hätte y keine Vorzeichenbedingung, und in I wäre

$b = A \cdot x$

$y^T \cdot b = y^T \cdot A \cdot x$

(statt $x \geq 0$)

→ $x \leq 0$ steht → dann in II. würde man

$y^T \cdot A \geq c^T$ fordern, und so

$y^T \cdot A \cdot x \leq c^T \cdot x$ würde gelten

→ x ohne Vorzeichenbedingung. → dann würde man in II

$y^T \cdot A = c^T$ fordern, und

$y^T \cdot A \cdot x = c^T x$ würde gelten

L+D 4:

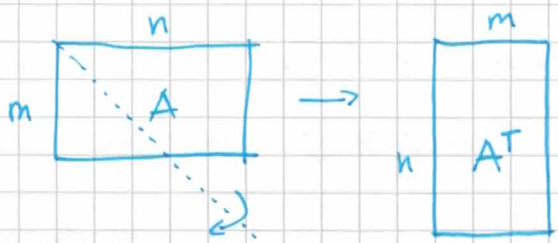
Wir sehen: Lineare Programme treten in Paaren auf!

Zu jedem LP Minimierungsproblem gehört ein Maximierungsproblem mit der transponierten (gespiegelten) Matrix, so dass jede Lösung y des Maximierungsproblems hat einen Zielwert $\leq$ als jede Lösung des Minimierungsproblems; es gilt sogar: die optimalen Zielwerte sind gleich

↓ nicht trivial, siehe später

Zielwerte $y^T \cdot b$ des Maximierungsproblems

Zielwerte $c^T \cdot x$ des Minimierungsproblems



ii.) Das duale LP zu einem LP in Standardform

(Sei $A \in \mathbb{R}^{m \times n}$ c, x n-dimensional b m-dimensional)

Zu dem primalen LP

(P) minimiere $c^T \cdot x$ s.d. $Ax = b$ $x \geq 0$

gehört das duale LP

(D) maximiere $y^T \cdot b$ s.d. $y^T \cdot A \leq c^T$

mit $y^T = (y_1\ y_2 \ldots y_m)$

(Statt A zu transponieren, wird y^T von links mit A multipliziert)

Beispiel 1: Minimiere $13x_1 + 10x_2 + 6x_3$

so dass
$$5x_1 + x_2 + 3x_3 = 8$$
$$3x_1 + x_2 = 3$$
$$x_1, x_2, x_3 \geq 0$$

$\downarrow$

(P) $$\begin{bmatrix} 5 & 1 & 3 \\ 3 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

$$c^T = \begin{bmatrix} 13 & 10 & 6 \end{bmatrix}$$

(D) maximiere $\begin{bmatrix} y_1 & y_2 \end{bmatrix} \cdot \begin{bmatrix} 8 \\ 3 \end{bmatrix}$

s.d. $$\begin{bmatrix} y_1 & y_2 \end{bmatrix} \cdot \begin{bmatrix} 5 & 1 & 3 \\ 3 & 1 & 0 \end{bmatrix} \leq \begin{bmatrix} 13 & 10 & 6 \end{bmatrix}$$

$\downarrow$

maximiere $8y_1 + 3y_2$

so dass
$$5y_1 + 3y_2 \leq 13$$
$$y_1 + y_2 \leq 10$$
$$3y_1 \leq 6$$
$$y_1\ y_2 \text{ frei}$$

Beachte: Jede Bedingung $a_i^T \cdot x = b_i$ im (P) entspricht einer Variable y_i in (D); (die m Zeilen von A)

jede Variable $x_j \geq 0$ in (P) entspricht einer Bedingung $y^T \cdot a^j \leq c_j$. (die n Spalten von A)

LPD 6.

Bemerkung: (P) und (D) (in der Standardform) können in der folgenden kompakten Form ausgedrückt werden:

(D) bestimme $\max\limits_{y} \min\limits_{x\geq 0} \left(c^T\cdot x + y^T\cdot b - y^T\cdot A\cdot x\right)$

(P) bestimme $\min\limits_{x\geq 0} \max\limits_{y} \left(c^T\cdot x + y^T\cdot b - y^T\cdot A\cdot x\right)$

(Übrigens: dasselbe gilt für beliebiges (P) und (D) Paar mit den nötigen Einschränkungen für x und y in min und max)

Warum?

Wir zeigen für die Standardform: für (D)

(*) $\min\limits_{x\geq 0} \left(c^T\cdot x + y^T\cdot b - y^T\cdot A\cdot x\right) = y^T\cdot b + \min\limits_{x\geq 0} \left(c^T - y^T\cdot A\right)\cdot x$

$$\min\limits_{x\geq 0} \left(c^T - y^T\cdot A\right)\cdot x = \begin{cases} 0 & \text{falls } c^T - y^T A \geq 0 \\ & \text{d.h. } c_j - y^T\cdot a^j \geq 0 \ \forall j \\ -\infty & \text{sonst} \\ & \text{d.h. wenn } c_j - y^T\cdot a^j < 0 \\ & \text{für mind. ein } j \end{cases}$$

Deshalb $(*) = \begin{cases} y^T\cdot b & \text{falls } c^T - y^T\cdot A \geq 0 \\ -\infty & \text{sonst} \end{cases}$

Das Maximieren von (*) über alle y bedeutet somit:

"maximiere $y^T\cdot b$ unter allen y so dass $c^T \geq y^T\cdot A$."

Also: Für eine optimale duale Lösung y^*

$$y^{*T}\cdot b = \max\limits_{y} \min\limits_{x\geq 0} \left(y^T\cdot b + c^T\cdot x - y^T\cdot A x\right)$$

Analog gilt: Für eine optimale primale Lösung x^*

$$c^T \cdot x^* = \min_{x \geq 0} \max_{y} \left(y^T \cdot b + c^T \cdot x - y^T \cdot A \cdot x \right)$$

(überlege den analogen Beweis für $c^T \cdot x^*$)

---

iii.) Zusammenfassung, die allgemeine Form von (P) und (D)

Wir haben in der Einführung gesehen, wie die (Un)gleichungen in den Bedingungen den dualen Vorzeichenbedingungen entsprechen, und umgekehrt. Diese Regel können sogar Zeile für Zeile bzw. Spalte für Spalte verwendet werden. Als Konvention, nehmen wir (hier) an, dass (P) ein Minimierungsproblem, und (D) ein Maximierungsproblem ist (nicht wesentlich), dann gilt die allgemeine Definition in folgender Form:

	(P)	(D)
	minimiere $c^T \cdot x$	maximiere $y^T \cdot b$
	so dass	so dass
Zeilen	$a_i^T \cdot x \geq b_i$	$y_i \geq 0$
	$a_i^T \cdot x \leq b_i$	$y_i \leq 0$
	$a_i^T \cdot x = b_i$	y_i frei
Spalten	$x_j \geq 0$	$y^T \cdot a^j \leq c_j$
	$x_j \leq 0$	$y^T \cdot a^j \geq c_j$
	x_j frei	$y^T \cdot a^j = c_j$

Jede Variable y_i in (D) entspricht einer Bedingung in (P), und jede Bedingung in (D) entspricht einer Variable x_j. Weiterhin werden die Bedingungen so definiert, dass

$$\boxed{\begin{gathered} (c_j - y^T \cdot a^j) \cdot x_j \geq 0 \quad \text{für} \quad j = 1, 2, \ldots, n \\ \text{und} \\ y_i \cdot (a_i^T \cdot x - b_i) \geq 0 \quad \text{für} \quad i = 1, 2, \ldots, m \end{gathered}}$$

Als die Summe dieser Terme über alle j bzw. über alle i erhalten wir jeweils

$$\sum_{j=1}^{n}(c_j - y^T a^j)\cdot x_j = (c^T - y^T\cdot A)\cdot x \geq 0$$

$$\sum_{i=1}^{m} y_i\cdot(a_i^T\cdot x - b_i) = y^T\cdot(A\cdot x - b) \geq 0$$

(✿)

für beliebige Lösungen x für (P) und y für (D)

Die beiden Ungleichungen ergeben den schwachen Dualitätssatz $c^T x \geq y^T\cdot b$ (siehe später)

Beobachtungen:

1. Das duale LP des dualen LP (D) ist das primale LP (P).

2. Wenn wir das primale LP (P) in ein äquivalentes LP transformieren (z.B. durch die Einführung von Slackvariablen, die Eliminierung von linear abhängigen Gleichungen, usw.), dann ist das duale LP des so erhaltenen LP äquivalent mit dem dualen von (P).

Beispiel 2: Bestimme das duale (D) von

$$(P)\quad \begin{aligned} \text{minimiere}\quad & 7x_1 - 11x_2 \\ \text{so dass}\quad & x_1 + 2x_2 \geq 1 \\ & 2x_1 + x_2 = 5 \\ & 3x_1 + x_2 \leq -5 \\ & x_1 \leq 0 \end{aligned}$$

Wir wissen: maximiere $y^T \cdot b$

jede Bedingung i in (P) $\longleftrightarrow$ Variable y_i in (D)

jede Variable x_j in (P) $\Leftrightarrow$ Bedingung mit c_j $y^T \cdot a^j$ in (D)

Weiterhin sollen wir alle Bedingungen und Vorzeichenbedingungen so setzen, dass $(c_j - y^T \cdot \underline{a}^j) \cdot x_j \geq 0$ und

$y_i \cdot (\underline{a}_i^T \underline{x} - b_i) \geq 0$ gelten

$$
\begin{array}{l|cc c c}
 & \begin{matrix} 0 \\ \text{\rotatebox{90}{$\leq$}} \\ x_1 \end{matrix} & \begin{matrix} \text{frei} \\ x_2 \end{matrix} & & b \\
\hline
y_1 \geq 0 & 1 & 2 & \geq & 1 \\
y_2 \text{ frei} & 2 & 1 & = & 5 \\
y_3 \leq 0 & 3 & 1 & \leq & -5 \\
\hline
 & \geq & = & & \\
c^T = & 7 & -11 & &
\end{array}
\qquad A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}
$$

(D) maximiere $y_1 + 5y_2 - 5y_3$

so dass $y_1 + 2y_2 + 3y_3 \geq 7$

$2y_1 + y_2 + y_3 = -11$

$y_1 \geq 0$

$y_3 \leq 0$

LPD 10.

IV.) Dualitätssätze

Schwache Dualität:

Theorem: Wenn x eine Lösung des primalen LP, und y eine Lösung des dualen LP ist, dann gilt

$$y^T \cdot b \le c^T \cdot x$$

Beweis: Laut Definition von (D) gelten die Ungleichungen (✿), deshalb $c^T \cdot x \geq y^T \cdot A \cdot x \geq y^T \cdot b$ für beliebige Lösungen x und y von (P) bzw. (D) □

Korollar 1: Wenn für die Lösungen x und y von (P) bzw. (D) $y^T \cdot b = c^T \cdot x$ gilt, dann sind x und y optimale Lösungen von (P) bzw. von (D).

Warum? Laut schwache Dualität ist $y^T \cdot b$ eine untere Schranke für den Zielwert für alle x in der Lösungsmenge von (P), also wenn $y^T \cdot b = c^T \cdot x$, dann ist $c^T \cdot x$ minimal. Analog: $y^T \cdot b$ ist maximal.

Korollar 2: Wenn das Minimum im primalen LP $-\infty$ ist, dann ist das duale LP unlösbar, wenn das Maximum des dualen LP ∞ ist, dann ist das primale LP unlösbar.

Starke Dualität:

Theorem: Wenn ein lineares Programm eine (endliche) optimale Lösung x^* hat, dann hat sein duales Programm auch eine optimale Lösung y^*, und

$$y^{*T} \cdot b = c^T \cdot x^*$$

(P) ist dabei ein Minimierungsproblem oder ein Maximierungsproblem, alles ist symmetrisch

~~Der Vollständigkeit halber wird der Beweis hier später aufgeführt.~~ Statt Beweis, zeigen wir zunächst zwei Illustrationen zum starken Dualitätssatz.

Illustration 1: → In der Bemerkung haben wir gesehen, dass für (P) in Standardform, das Optimum $y^{*T} \cdot b$ von (D) ist

$$y^{*T} \cdot b = \max_{y} \min_{x \geq 0} \left(c^T \cdot x + y^T \cdot b - y^T \cdot A \cdot x \right)$$

und das Optimum $c^T \cdot x^*$ von (P) ist

$$c^T \cdot x^* = \min_{x \geq 0} \max_{y} \left(c^T \cdot x + y^T \cdot b - y^T \cdot A \cdot x \right)$$

→ (Seien $S \subseteq \mathbb{R}^n$ und $T \subseteq \mathbb{R}^m$) Für beliebige Funktion $f : S \times T \to \mathbb{R}$ gilt trivial (siehe Übung)

$$\max_{y \in T} \min_{x \in S} f(x, y) \leq \min_{x \in S} \max_{y \in T} f(x, y)$$

(dies würde bei unserem Fall also $y^{*T} \cdot b \leq c^T \cdot x^*$ ergeben)

→ Für eine lineare Funktion f und gut gewählte S und T gilt aber sogar

$$\max_{y} \min_{x} f(x, y) = \min_{x} \max_{y} f(x, y)$$

Siehe die 3D Fläche (Graph) einer bivariaten linearen Funktion $g(x, y) = c \cdot x + by + a \cdot x \cdot y$ die dem Fall $n = m = 1$ entspricht, und seien $S \subseteq \mathbb{R}$ und $T \subseteq \mathbb{R}$ Intervalle

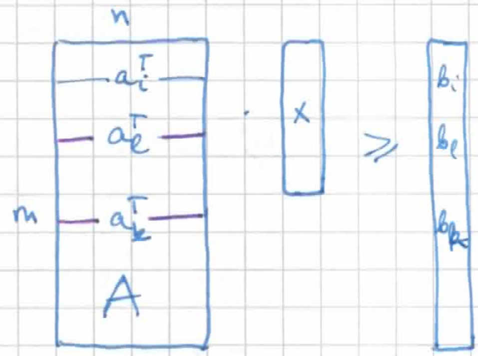
LPD 12.

Illustration 2.

Sei (P) ein LP in kanonischer Form:

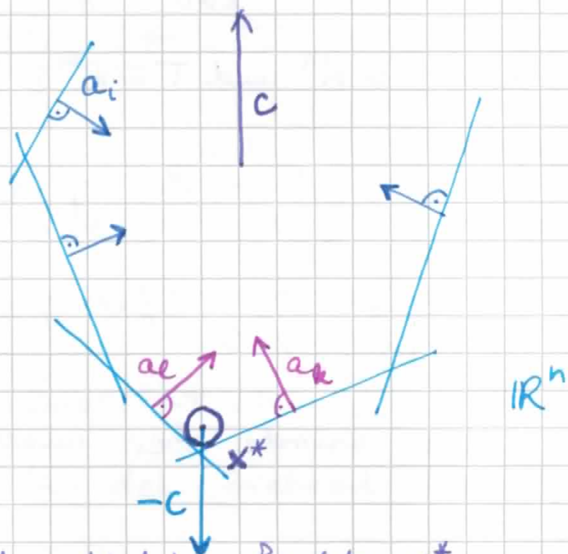
$$\text{minimiere } \quad c^T \cdot x$$

$$\text{so dass } \quad A x \geq b$$



Sei $n=2$ oder $n=3$, und nehmen wir o.B.d.A. an, dass c vertikal nach oben zeigt.

→ die Kanten des Lösungspolyeders haben jeweils eine Zeile a_i^T der Matrix als Normalvektor (weil das Polyeder durch die Bedingungen $a_i^T \cdot x \geq b_i$ definiert wird).



→ $c^T \cdot x$ zu minimieren bedeutet den tiefsten Punkt x^* im Lösungspolyeder zu finden.

→ Wir stellen uns vor, dass wir eine (unendlich kleine) Kugel in das Polyeder werfen, und $-c$ ist die Gravitationskraft die auf die Kugel wirkt.

Die Kugel wird natürlich im tiefsten Punkt, also im Punkt x^* zur Ruhe kommen.

Das duale Programm zu (P) ist

(D) maximiere $y^T \cdot b$

so dass $y^T \cdot A = c^T$

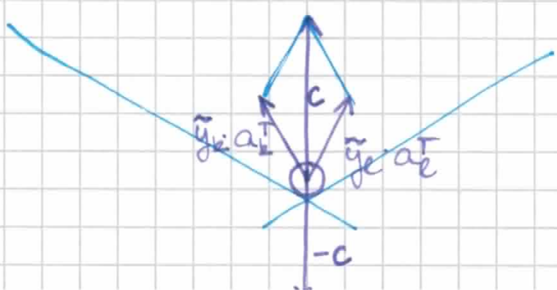
$y \geq 0$

→ Für jede Lösung y von (D) $y^T \cdot A = c^T$ bedeutet

$\sum_{i=1}^{m} y_i \cdot \underline{a}_i^T = \underline{c}^T$, d.h. c^T wird als eine Linearkombination der Zeilenvektoren a_i^T erzeugt und die y_i der dualen Lösung sind die Koeffizienten zu den a_i^T Vektoren.

→ Sei jetzt $\tilde{y}_i \cdot \underline{a}_i^T$ die Stützkraft, die die "Wand" i auf unsere Kugel ausübt (Diese Kraft ist orthogonal zur Facette (Wand) i genau wie der Vektor $\underline{a}_i^T$)

Dann ist $\tilde{y}_i = 0$ für Wände die die Kugel nicht berührt, und $\tilde{y}_i > 0$ für Wände die die Kugel stützen, weiterhin gilt $c^T = \sum_{i=1}^{m} \tilde{y}_i a_i^T$, da sich alle Kräfte (auf die Kugel) ausgleichen. Deshalb ist $\tilde{y} = (\tilde{y}_1 \ldots \tilde{y}_m)$ eine Lösung von (D) In unserem Beispiel $c^T = \tilde{y}_\ell \cdot a_\ell^T + \tilde{y}_k \cdot a_k^T$



LPD 14.

Wir zeigen, dass für $\tilde{y}$ sogar

$$\tilde{y}^T \cdot b = c^T \cdot x^* \quad \text{gilt}$$

Dieses Beispiel aus der Physik illustriert somit den starken Dualitätssatz, wir haben eine Lösung von (D) gefunden mit dem gleichen Zielwert wie der von x^*.

Wie für jede (P) und (D) gelten die Ungleichungen

$$\left(c^T - \tilde{y}^T \cdot A\right) \cdot x^* \geq 0$$

$$\tilde{y}^T \cdot (A \cdot x^* - b) \geq 0 \qquad \text{für die Lösungen } x^* \text{ und } \tilde{y}.$$

Wir zeigen, dass beide genau $= 0$ sind, also

$$c^T \cdot x^* = \tilde{y}^T \cdot A \cdot x^* = \tilde{y}^T \cdot b.$$

Warum? $\left(c^T - \tilde{y}^T \cdot A\right) \cdot x^* = 0$ ist in diesem Fall (für jede Lösung y) trivial, weil $c^T = \tilde{y}^T \cdot A$ (sogar $c^T - \tilde{y}^T \cdot A = \underline{0}$) $\Rightarrow \underline{0} \cdot x^* = 0$

$$\tilde{y}^T \cdot (A x^* - b) = \sum_{i=1}^{m} \tilde{y}_i \cdot \left(a_i^T \cdot x^* - b_i\right) = 0$$

$\hookrightarrow$ für Stützwände

$a_i^T \cdot x^* = b_i$ wird exakt erfüllt in der Ecke x^* hier also $a_i^T \cdot x^* - b_i = 0$

Für nicht-Stützwände

$\tilde{y}_i = 0$

Zusammenfassung:

Der schwache u. starke Dualitätssatz lässt die folgenden Möglichkeiten zu:

(D) \ (P)	endliches OPT	OPT = $-\infty$	unlösbar
endliches OPT	+		
OPT = ∞			+
unlösbar		+	+

(Ein Beispiel in dem (P) und (D) beide unlösbar sind:

(P)

$$\begin{aligned} &\text{minimiere} && x_1 + 2x_2 \\ &\text{s.d.} && x_1 + x_2 = 1 \\ & && 2x_1 + 2x_2 = 3 \end{aligned}$$

(D)

$$\begin{aligned} &\text{maximiere} && y_1 + 3y_2 \\ &\text{s.d.} && y_1 + 2y_2 = 1 \\ & && y_1 + 2y_2 = 2 \end{aligned}$$

)

V.) Komplementäre Slackness

Wir betrachten wieder die Ungleichungen die für jedes (P)–(D) Paar gelten sollen für jede Lösung x von (P) und jede Lösung y von (D):

$$(c_j - y^T \cdot a^j) \cdot x_j \geq 0 \qquad \text{für } j = 1, 2, \ldots, n$$

und

$$y_i \cdot (a_i^T \cdot x - b_i) \geq 0 \qquad \text{für } i = 1, 2, \ldots, m$$

Der Beweis der schwachen Dualität war im Grunde genommen

$$c^T \cdot x - y^T \cdot b = \sum_{j=1}^{n} (c_j - y^T \cdot a^j) \cdot x_j + \sum_{i=1}^{m} y_i \cdot (a_i^T \cdot x - b_i) \geq 0$$

Hier gilt also ≥ 0 für jeden der $n+m$ Summanden einzeln! Dies impliziert, dass x und y genau dann **beide** optimal sind, d.h. $c^T \cdot x = y^T \cdot b$ genau dann gilt, wenn

$$(c_j - y^T \cdot a^j) \cdot x_j = 0 \qquad \text{für alle } j$$

$$\text{und } \; y_i \cdot (a_i^T \cdot x - b_i) = 0 \qquad \text{für alle } i$$

$$\Updownarrow$$

PKS: Für alle j: entweder $x_j = 0$ oder die entsprechende duale Ungleichung $c_j \geq y^T \cdot a^j$ exakt erfüllt wird $c_j = y^T \cdot a^j$

DKS: und analog für alle i: entweder $y_i = 0$ oder $a_i^T \cdot x = b_i$

Diese Bedingungen nennt man Komplementäre Slackness Bedingungen (complementary slackness conditions)

Für beliebige Lösungen x und y, die Dualitätslücke ist $c^T \cdot x - y^T \cdot b$ (duality gap)