1. Einführung

1.1 Idee (Begriffsherkunft, Geschichte):

a) Optimierung, von lateinisch optimum, „das Beste, das Höchstmaß"

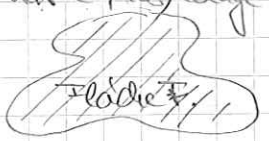
b) Geschichte

i) Euler [1744]: Variationsrechnung „Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici latissimo sensu accepti" (Eine Methode, um Kurven zu finden, denen eine Eigenschaft im höchsten oder geringsten Grade zukommt oder die Lsg. des isoperimetrischen Problems, wenn es im weitesten Sinne des Wortes aufgefasst wird)

„Zweipunkt" „Nebenbed."

$$\min \int_{t_0}^{t_1} L(y(t), y'(t), t)\,dt$$

$$\text{s.t. } y(t_0) = a,\ y(t_1) = b$$

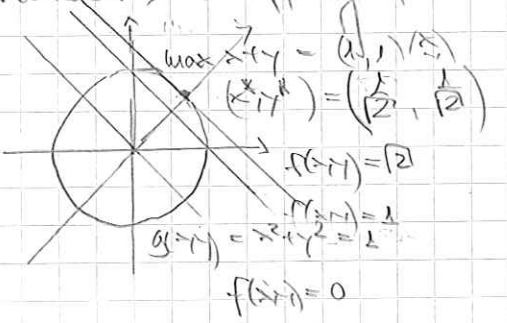


$$\max F \quad \text{s.t. } \ell(\gamma) \leq K$$

Lsg: Kreis vom Umfang K.

ii) Lagrange [1762]: notwendige Optimalitätsbedingungen

$$\max f(x)$$
$$\text{s.t. } g(x) = 0$$



x^* lokal optimal

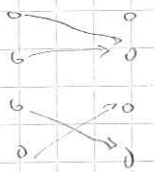
$\Rightarrow \exists \lambda^*: \nabla f(x^*) = \lambda^* g(x^*)$

$g(x^*) = 0$

$\{g=0\}$ berührt $\{f=\sqrt{2}\}$ in (x^*,y^*) tangential

$\lambda^* = \frac{1}{\sqrt{2}} \cdot (1,1) = \frac{1}{\sqrt{2}}\left(2\cdot\frac{1}{\sqrt{2}}, 2\cdot\frac{1}{\sqrt{2}}\right)$

iii) Monge [1781]: (Massen-) Transportproblem



X meßbare Räume Y

μ Wahrscheinlichkeiten ν

$\nu: X \to Y$ injektiv „Transportfunktion"

$\mu(\nu^{-1}(B)) = \nu(B)\ \forall B \subseteq Y$ meßbar „Massenerhaltung"

$c: X \times Y \to \mathbb{R}$ Transportkostenfunktion (z.B. $c(x,y) = \|x-y\|$)

$$\min \int_X c(x, \nu(x))\,d\mu(x)$$

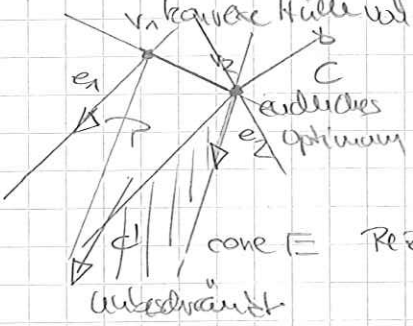
s.t. ν inj., $\mu(\nu^{-1}(B)) = \nu(B)$

iv) Tschebyscheff [1854]: „Mindestabweichung“

$$\min_{\alpha} \max_{t \in T} \left| f(t) - \sum_i \alpha_i f_i(t) \right|, \quad f_i \text{ Polynom vom Grad } i$$

log für $f \equiv 1$, $f_i(t) = t^i$, $T = [0,1]$: Tschebyscheff-Polynome

v) Minkowski [1896]: Darstellungssatz für Polyeder



$$\min cx \quad \text{s.t. } x \in P$$

vi) Farkas [1902]: Lösbarkeit von Ungleichungssystemen

$A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$

$$\{Ax \le b\} \neq \emptyset \Leftrightarrow \left\{ \begin{array}{l} b^T u < 0 \\ A u \ge 0 \\ u \ge 0 \end{array} \right\} = \emptyset \Leftrightarrow \begin{array}{l} \min b u \ge 0 \\ A^T u \ge 0 \\ u \ge 0 \end{array}$$

es ist wichtig zu prüfen, ob überhaupt eine Lsg. existiert!

Vgl. Perron-Paradoxon [1913]: $\max n \in \mathbb{N}$.

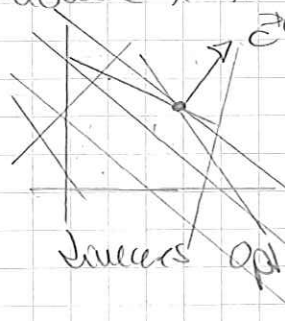
Sei $n = \operatorname{argmax}_{n \in \mathbb{N}} n$. Ann.: $n > 1 \Rightarrow n^2 > n$ ↯ $\Rightarrow n = 1$.

↑ ex. eins Max. wird „fälschlich“ angenommen

vii) Kantorovich [1939], Koopmans [1942], Dantzig [1947]:

Lineare Optimierung, Schalenpreis, Simplexalgorithmus

Nobelpreis für Ökonomie 1975 für Kantorovich und Koopmans



(LP) „primal“ $\min c^T x$, $Ax = b$, $x \ge 0$

$\Leftrightarrow \min c^T x$, $Ax \le b$, $-Ax \le -b$, $-x \le 0$

Von Neumann & Morgenstern [1944]: Spieltheorie (äquivalent)

2 Spieler mit Strategien $1, \dots, m$, Auszahlungsmatrix $A \in \mathbb{R}^{m \times n}$

$p, q \in [0,1]^m$ Wahrscheinlichkeiten für gemischte Strategien von Spieler 1, 2

↑ ↑

Minimax-Satz

$$\max_{p^T \mathbb{1} = 1, p \ge 0} \min_{q^T \mathbb{1} = 1, q \ge 0} p^T A q = \min_q \max_p p^T A q$$

zweispieler Spielbaum

viii) Karush [1939], Kuhn & Tucker [1951]: nichtlineare Opt.
„constraint qualification" für notwend. Bedingungen
min $f(x)$
$g(x) \leq 0$

ix) Khachiyan [1979]: Ellipsoidmethode (vollg. binäre Suche)
(polynomiales)
erstes überhaupt effizientes Verfahren zur Lsg. von lin. Opt.problemen

x) Karmarkar [1984]: „projektive Scaling method"
(polynomial)
1. innere-Punkte-Methode zur Lsg. von LPs

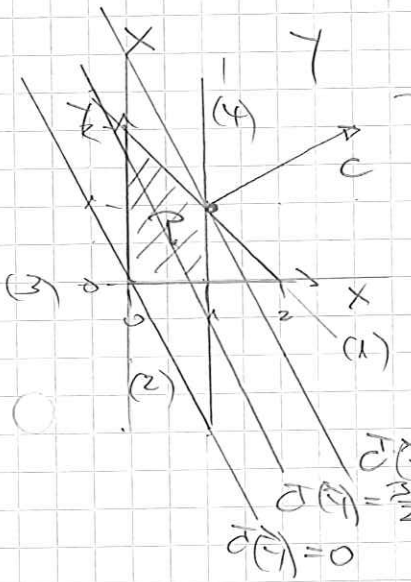
1.2 Bsp (lineares Programm, Produktion): Allerleirauh vs. AG

a)

Allerl.	AG		Aktivität
2	1	max	Nutzen/h
1	1	≤ 2	Zeitaufwand
1	0	≤ 1	
x	y	≥ 0	h/Aktivität

Lineares Programm:

$$\max\ 2x + 1\cdot y \quad (z)$$
$$x + y \leq 2 \quad (1)$$
$$x \leq 1 \quad (2)$$
$$x \geq 0 \quad (3)$$
$$y \geq 0 \quad (4)$$



$P = \{Ax \leq b, x \geq 0\}$ Polyeder

optimallsg $\begin{pmatrix}x^*\\y^*\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}$ $\Leftrightarrow \max \begin{pmatrix}2\\1\end{pmatrix}^T \begin{pmatrix}x\\y\end{pmatrix}$

Zielfktwert $c^T\begin{pmatrix}x^*\\y^*\end{pmatrix} = \begin{pmatrix}2\\1\end{pmatrix}^T\begin{pmatrix}1\\1\end{pmatrix} = 3$ $\begin{pmatrix}2 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} \leq \begin{pmatrix}2\\1\end{pmatrix}$

$x, y \geq 0$

$\Leftrightarrow \max c^T x$

$Ax \leq b$

$x \geq 0$

$\hookrightarrow \max_{x \in P} c^T x$

b) Wie können wir einfach beweisen, daß $\begin{pmatrix}x^*\\y^*\end{pmatrix}$ optimal ist?

Durch Manipulation der Ungleichungen:

$x + y \leq 2$ (1) $\quad \lambda \geq 0$

$x \leq 1$ (2) $\quad \mu \geq 0$

$2x + y \leq 3$ $\quad 1\cdot(1) + 1\cdot(2)$

$\Rightarrow \underbrace{c^T\begin{pmatrix}x\\y\end{pmatrix}}_{\text{Zielfunktion!}} \leq 3$ aber $c^T\begin{pmatrix}x^*\\y^*\end{pmatrix} = 3 \Rightarrow \begin{pmatrix}x^*\\y^*\end{pmatrix}$ optimal

↑ duale Schranke

Allgemein kann man Ungleichungen addieren und mit positiven Zahlen multiplizieren und damit eine Abschätzung der "besten" bestimmen:

$$\begin{array}{rcl} x + y & \le & 2 \quad (1) \quad \lambda \ge 0 \\ x & \le & 1 \quad (2) \quad \mu \ge 0 \end{array}$$

$$(\lambda+\mu)x + \lambda y \le 2\lambda + \mu \to \min$$

$$\lambda + \mu \ge 1, \quad \lambda \ge 1$$

$$2x + 1 \cdot y \to \max$$

"Primales LP"

$$\max c^T x \quad Ax \le b, \quad x \ge 0$$

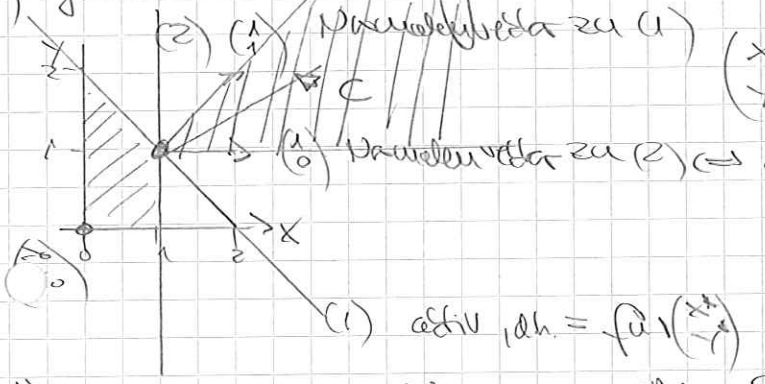
Dualitätssatz $= \min b^T y$ "Duales LP"

$$A^T y \ge c, \quad y \ge 0$$

$$\Rightarrow \min 2\lambda + \mu \quad \lambda + \mu \ge 2, \quad \lambda \ge 1, \quad \lambda \ge 0, \quad \mu \ge 0$$

$$\Leftrightarrow \min \begin{pmatrix} 2 \\ 1 \end{pmatrix}^T \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \ge \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \lambda, \mu \ge 0$$

c) geometrische Bedeutung:



$c \in$ Randkegel der für $\begin{pmatrix} x^* \\ y^* \end{pmatrix}$ aktiven Ungleichungen

d) Lösung des Optimums mittels „Simplextableau":

Führe „Schlupfvariablen" s_1, s_2 ein:

$$LP \Leftrightarrow \max 2x + y$$

$$\begin{array}{rcl} x + y + s_1 & = & 2 \\ x + s_2 & = & 1 \\ x, y, s_1, s_2 & \ge & 0 \end{array}$$

$$\Rightarrow \begin{array}{rcl} z & = & 2x + y \\ s_1 & = & 2 - x - y \\ s_2 & = & 1 - x \end{array} \quad \forall x, y \in P$$

Setze $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Lasse $x = x_0 = 0$, aber erhöhe $y \Rightarrow s_1 = 2 - y \ge 0 \Rightarrow y \le 2$

$s_2 = 1$

$z = y \to \max!$

$$\Leftrightarrow \max x - s_1 = z - 2$$

$$\begin{array}{rcl} x + y + s_1 & = & 2 \\ x + s_2 & = & 1 \\ x, y, s_1, s_2 & \ge & 0 \end{array}$$

$$\begin{array}{rcl} z & = & 0 + x - s_1 \\ y & = & 2 - x - s_1 \\ s_2 & = & 1 - x \end{array} \quad \forall x, y \in P$$

Setze $\begin{pmatrix} x_0 \\ s_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

lasse $s_1 = 0$, aber erhöhe x $\Rightarrow$ $y = 2 - x \to x \le 2$, $s_2 = 1 - x \to x \le 1$ $\Rightarrow x \le 1$

$$\begin{array}{lll} \max & -s_1 - s_2 = z - 3 & z = 3 - s_1 - s_2 \\ \Leftrightarrow & y + s_1 - s_2 = 1 & \Rightarrow \quad y = 1 - s_1 + s_2 \\ & x \quad\quad + s_2 = 1 & x = 1 \quad - s_2 \\ & x, y, s_1, s_2 \ge 0 & \end{array} \qquad \forall x, y \in \mathcal{D}$$

setze $\begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Es bringt nichts, s_1 oder s_2 zu erhöhen $\to \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ optimal // 1.3 Def. A

1.3 Def. (Optimierung, Optimierungsproblem): Sei $X \subseteq \mathbb{R}^n$ eine Menge, $f: X \to \mathbb{R}$ eine Funktion.

a) Optimierung ist die Bestimmung eines $\left\{\begin{matrix} \text{Min.} \\ \text{Max.} \end{matrix}\right\}$ von Funktion f über eine Menge X.

b) (P) $\min f(x)$ Minimierungsproblem
$\text{s.t. } x \in X$

$=: \min_{x \in X} f(x) =: \min \{f(x) : x \in X\} =: \min_{x \in X} f(x)$

X	zulässige Menge, Zulässigkeitsbereich
$x \in X$	zulässiger Punkt, zulässige Lösung
f	Zielfunktion
$X = \mathbb{R}^n$	P = unrestringiertes Opt. prob.
$X \neq \mathbb{R}^n$	P = Opt. prob. mit Nebenbedingungen
$X = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \le 0\}$ für $h: \mathbb{R}^n \to \mathbb{R}^m$, $g: \mathbb{R}^n \to \mathbb{R}^k$	P = Opt. prob. mit Gleichungen $h(x)=0$ und Ungleichungen $g(x) \le 0$
$\alpha = \min \{f(x) : x \in X\}$	Optimalwert von P
$\operatorname{Argmin}(P) = \{x \in X : f(x) = \alpha\}$	Menge der Opt. Lsgen von P

$\operatorname{argmin}(P) \subset \operatorname{Argmin}(P)$: (irgendeine) Opt. Lsg. von P.

1.4 Bsp (Opt. prob.):

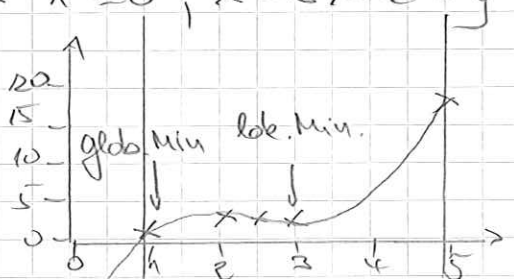
$\Leftrightarrow x(x-5) \le 0$

(P) $\min \{ x^3 - \frac{15}{2}x^2 + 18x - \frac{21}{2} : x - 1 \ge 0, x^2 - 5x \le 0 \}$

$= \min \; x^3 - \frac{15}{2}x^2 + 18x - \frac{21}{2}$

$\text{s.t.} \quad x - 1 \ge 0$

$x^2 - 5x \le 0$



$f(x) = x^3 - \frac{15}{2}x^2 + 18x - \frac{21}{2}$

$f'(x) = 3x^2 - 15x + 18 = 0 \Leftrightarrow x^2 - 5x + 6 = (x-2)(x-3) = 0$

$\Rightarrow x = 2 \vee x = 3$

$f''(x) = 6x - 15 = 0 \Leftrightarrow 2x - 5 = 0 \Leftrightarrow x = \frac{5}{2}$

$f(1) = 1$, $f(2) = \frac{7}{2} = 3{,}5$, $f(\frac{5}{2}) = \frac{11}{4} = 3{,}25$, $f(3) = 3$, $f(5) = 17$

$\Rightarrow$ Argmin $(\overline{P}) = \{1\}$.

1.5 Def. (globales und lokales Min./Max.): Sei $\bar{x} \in X \subseteq \mathbb{R}^n$, $f: X \to \mathbb{R}$.

a) $\exists \varepsilon > 0: f(\bar{x}) \leq f(x) \; \forall x \in U_\varepsilon(\bar{x}) \cap X \Leftrightarrow$ $\bar{x}$ lok. Min. punkt von f über X

b) $f(\bar{x}) \leq f(x) \; \forall x \in X \Leftrightarrow$ $\bar{x}$ glob. — " —

Bez.: glob. Max. punkte analog.

1.6 Bem. (Äquivalenz von Minimierung und Maximierung):

$\max_{x \in X} f(x) \Leftrightarrow \min_{x \in X} -f(x)$, d.h. Max. prob. lassen sich als Min. prob. lösen und umgekehrt.

1.7 Def. (wichtige Typen von Opt. prob.):

a) Sei $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $D \in \mathbb{R}^{k \times n}$, $d \in \mathbb{R}^k$, $c \in \mathbb{R}^n$.

(LP) $\min c^T x$ s.t. $Ax = b$, $Dx \leq d$ — lineares Opt. Prob., lineares Programm

$\min c^T x$ s.t. $Ax = b$, $x \geq 0$ — LP in Standardform

$X = \{Ax = b, Dx \leq d\}$, $f: X \to \mathbb{R} = c^T: X \to \mathbb{R}, x \mapsto c^T x$

b) (P) $\min f(x)$ s.t. $x \in X$, $x \in \mathbb{Z}^n$ — diskretes Opt. Prob.

(IP) $\min c^T x$, $Ax = b$, $Dx \leq d$, $x \in \mathbb{Z}^n$ — ganzzahliges lineares Programm, ganzes Programm

(BP) $\min c^T x$, $Ax = b$, $Dx \leq d$, $x \in \{0,1\}^n$ — binäres IP, kombinatorisches Opt. prob.

c) Opt. prob. ohne (explizite) Ganzzahligkeitsrestriktionen heißt kontinuierliches Opt. prob.

Beob.: $x(x-1)=0 \Leftrightarrow x \in \{0,1\} \Leftrightarrow x^2 = x$

d) (P) $\min_{\text{s.t.}} f(x)$ $x \in X$ | f, X konvex konvexes Opt. prob. (s. später)

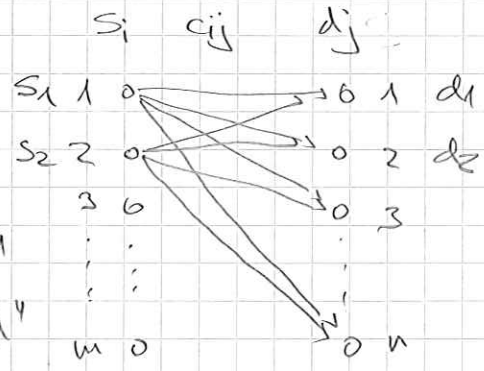
Wir werden sehen, daß bei konv. Opt. prob. lokale Opt. immer globale Opt. sind.

e) Kont. Opt. prob. ohne Konvexität heißen globale Opt. prob.

f) Sei $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}^k$, $Q \in \mathbb{R}^{n \times n}$

(QP) $\min x^T Q x$ ← quadratisches Opt. prob.
$\text{s.t. } Ax = b$
$Dx \le d$ ← linear!

1.8 Bsp. (Transportproblem):



f_i, $i = 1, \dots, m$ Fabriken

c_j, $j = 1, \dots, n$ Kunden

s_i, $i = 1, \dots, m$ Kapazität von f_i „supply"

d_j, $j = 1, \dots, n$ Bedarf von c_j „demand"

c_{ij}, $i = 1, \dots, m$, $j = 1, \dots, n$ Transportkosten/Einheit

Frage: Wie befriedigt man den Bedarf kostenminimal?

Modellierung als LP:

a) Variablen:

x_{ij}, $i = 1, \dots, m$, $j = 1, \dots, n$ Transportmenge $i \to j$

b) Zielfunktion

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

c) Nebenbed.

$$\sum_{j=1}^{n} x_{ij} \le s_i, \quad i = 1, \dots, m$$

$$\sum_{i=1}^{m} x_{ij} \ge d_j, \quad j = 1, \dots, n$$

$$x_{ij} \ge 0, \quad i = 1, \dots, m, \ j = 1, \dots, n$$

d) (LP) $\min \sum_i \sum_j c_{ij} x_{ij}$ (Transkost.)

$\text{s.t. } \sum_j x_{ij} \le s_i \ \forall i$ (Kap.)

$\sum_i x_{ij} \ge d_j \ \forall j$ (Bedarf)

$x_{ij} \ge 0 \ \forall ij$ (Vorz. neg.)

Transportproblem

1.9 Bsp (Portfoliooptimierung):

A_i, $i=1,\dots,n$ Aktien

r_i, $i=1,\dots,n$ Rendite von A_i nach einem Jahr (Zufallsvariable)

μ_i, $i=1,\dots,n$ Erwartungswert von r_i

σ_{ij}, $i,j=1,\dots,n$ Kovarianz von r_i und r_j

Frage: Welche Anteile seines Vermögens sollte man in welche Aktie investieren, um

a) bei Mindestrendite ρ das Risiko zu minimieren?

b) bei Maximalrisiko ν die Rendite zu maximieren?

Modellierung

i) Variablen: x_i, $i=1,\dots,n$ Anteil des Verm. in Aktie i

ii) Erwartete Rendite: $E\left(\sum_{i=1}^{n} r_i x_i\right) = E(r^T x) = \mu^T x$

iii) Risiko: $\mathrm{Var}(r^T x) = x^T \Sigma x$

a) Modellierung als konvexes QP

$$\begin{aligned} \min\ & x^T \Sigma x && \text{(Risiko)} \\ \text{s.t.}\ & \mu^T x \geq \rho && \text{(Min. Rendite)} \\ & \mathbb{1}^T x = 1 && \text{(Aufteilung)} \\ & x \geq 0 && \text{(Anlagerichtlinien)} \end{aligned}$$

b) Mod. als konvexes QP mit lin. und quadrat. Nebenbed.

$$\begin{aligned} \max\ & \mu^T x \\ & x^T \Sigma x \leq \nu \\ & \mathbb{1}^T x = 1 \\ & x \geq 0. \end{aligned}$$

24.04.

2. Convex Sets and Functions





2.1 Def (Convex Set): Let $C \subseteq \mathbb{R}^n$.

C convex $:\Leftrightarrow \forall x_1, x_2 \in C,\ \lambda \in [0,1]: \lambda x_1 + (1-\lambda) x_2 \in C$

i.e., the line segment connecting any two points in C is contained in C.

2.2 Ex (Convex Sets): The following sets are convex (→ ex.):

a) $\emptyset$, $\mathbb{R}^n$

b) $H := \{x \in \mathbb{R}^n : a^T x = \nu\}$, $a \in \mathbb{R}^n, \nu \in \mathbb{R}$ — hyperplane with normal vector a

c) $H^{\geq} := \{x \in \mathbb{R}^n : a^T x \geq \alpha\}$ — " — closed / open half-space generated by H

$H^{>} := \{x \in \mathbb{R}^n : a^T x > \alpha\}$ — " —

d) $B_\alpha(a) := \{x \in \mathbb{R}^n : \|x - a\| \leq \alpha\}$ — " — ball of radius α around a

2.3 Def (Convex Combination): Let $x_1, \dots, x_p \in \mathbb{R}^n$, $\lambda_1, \dots, \lambda_p \in \mathbb{R}_{\geq 0}$, $\sum_{i=1}^p \lambda_i = 1$, $p \in \mathbb{N}$. Then



$z := \sum_{i=1}^p \lambda_i x_i$ (finite) convex combination of $x_1, \dots, x_p$.

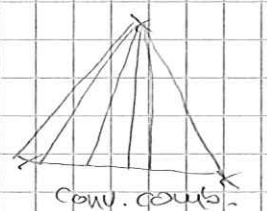
If $\lambda_1, \dots, \lambda_p > 0$, z is a proper or genuine convex comb. of $x_1, \dots, x_p$.

2.4 Thm (Char. of Convexity): A set $C \subseteq \mathbb{R}^n$ is convex if and only if all convex combinations of points in C belong to C.

Proof:

"⇒" Let $C \subseteq \mathbb{R}^n$, $z = \sum_{i=1}^p \lambda_i x_i$ a convex comb. of points $x_1, \dots, x_p \in C$.

We show $z \in C$ by induction over p.



Ind. Sat: $p = 2$ by Def. 2.1

Ind. Ass.: $z \in C$ for $p \geq 2$

Ind. Step $p \to p+1$: w.l.o.g. $\lambda_{p+1} \neq 0$.

$$z = \sum_{i=1}^{p+1} \lambda_i x_i = \sum_{i=1}^{p} \lambda_i x_i + \lambda_{p+1} x_{p+1}$$

$$= (1-\lambda_{p+1}) \underbrace{\sum_{i=1}^{p} \frac{\lambda_i}{1-\lambda_{p+1}} x_i}_{\text{convex comb.}} + \lambda_{p+1} x_{p+1} \in C,$$

Ind. Ass. $\in C$

convex comb.

"⇐" by Def. 2.1. □

2.5 Def. & Thm (Simple Prop. of Convex Sets): Let $C, D \subset \mathbb{R}^n$ be convex, $\alpha \in \mathbb{R}_{\geq 0}$. Then the following sets are convex:

a) The intersection of any number of convex sets is convex.

b) $C + D := \{x+y \in \mathbb{R}^n : x \in C, y \in D\}$ sum of C, D

c) $\alpha C := \{\alpha x \in \mathbb{R}^n : x \in C\}$ scalar multiple of C

d) $\overline{C} := C \cup \partial C$

Proof: Ex. □

2.6 Rem (Union of Convex Sets): The union of convex sets is in general not convex.

2.7 Def (Convex Hull): Let $M \subset \mathbb{R}^n$ be any set.

$$\operatorname{conv} M := \bigcap_{C \supseteq M : C \text{ convex}} C$$ convex hull of M

2.8 Thm (Convex Hull): The convex hull of a set $M \subseteq \mathbb{R}^n$ coincides with the set of convex combinations of points in M.

Proof: Ex. □

2.9 Thm (Caratheodory [1907] for compact sets, Steinitz [1914]): The convex hull of a set $M \subseteq \mathbb{R}^n$ coincides with the set of convex comb. of $n+1$ points in M.

Proof: Let $z \in \mu^{\text{Thm 2.8}} \Rightarrow z = \sum_{i=1}^{p} \lambda_i x_i$ convex comb. of points $x_1, \dots, x_p \in M$

If $p \leq n+1$ ✓. Let $p > n+1 \Leftrightarrow p-1 > n$ and set

$$y_i := x_i - x_p, \quad i = 1, \dots, p-1$$

$\Rightarrow$ y_i, $i = 1, \dots, p-1$ lin. dep.

$\Rightarrow \exists (\alpha_1, \dots, \alpha_{p-1}) \in \mathbb{R}^{p-1} \setminus \{0\} : \sum_{i=1}^{p-1} \alpha_i y_i = 0$

$$\Rightarrow 0 = \sum_{i=1}^{p-1} \alpha_i y_i = \sum_{i=1}^{p-1} \alpha_i x_i - \underbrace{\sum_{i=1}^{p-1} \alpha_i}_{=: \alpha_p} x_p = \sum_{i=1}^{p} \alpha_i x_p$$

with $(\alpha_1, \dots, \alpha_p) \neq 0$, $\sum_{i=1}^{p} \alpha_i = 0$.

$\Rightarrow \exists i \in \{1, \dots, p\} : \alpha_i > 0$ exists.

$$i_0 := \operatorname{argmin} \left\{ \frac{x_i}{\alpha_i} : \alpha_i > 0 \right\}.$$

$$\Rightarrow \lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}} \alpha_i \geq 0, \quad i = 1, \dots, p \quad (i) \left(\frac{\lambda_i}{\alpha_i} > \frac{\lambda_{i_0}}{\alpha_{i_0}} \text{ für } \alpha_i \geq 0, \text{ otherwise does} \right)$$

$$\sum_{i=1}^{p}\left(\lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\alpha_i\right) = \sum_{i=1}^{p}\lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\sum_{i=1}^{p}\alpha_i = 1, \quad \text{(ii)}$$

($\lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\alpha_i \geq 0$; $\sum_{i=1}^{p}\lambda_i = 1$; $\sum_{i=1}^{p}\alpha_i = 0$)

and

$$z = \sum_{\substack{i=1 \\ i\neq i_0}}^{p}\lambda_i x_i$$

$$0 = \sum_{\substack{i=1 \\ i\neq i_0}}^{p}\alpha_i x_i \quad \Big| \cdot\left(-\frac{\lambda_{i_0}}{\alpha_{i_0}}\right)$$

$$z = \sum_{\substack{i=1 \\ i\neq i_0}}^{p}\left(\lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\alpha_i\right)x_i = \sum_{\substack{i=1 \\ i\neq i_0}}^{p}\left(\lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\alpha_i\right)x_i \geq 1 \quad \square$$

($\lambda_i - \frac{\lambda_{i_0}}{\alpha_{i_0}}\alpha_i = 0$ for $i = i_0$)

(i), (ii): convex comb.

2.10 Def (Extreme Point, Vertex): Let $\emptyset \neq C \subseteq \mathbb{R}^n$ be convex. Then $x \in C$ is called an extreme point of C, if x cannot be constructed as a convex combination of two distinct points of C, i.e., if

$$x = \lambda x_1 + (1-\lambda)x_2,\ x_1, x_2 \in C,\ \lambda \in [0,1] \Rightarrow x_1 = x_2;$$

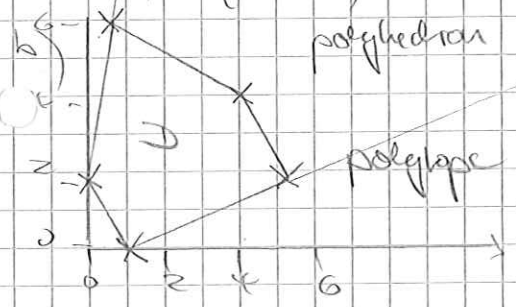
Let $\mathrm{ext}(C)$ be the set of extreme points. If $|\mathrm{ext}(C)| < \infty$, we call $\mathrm{ext}(C) =: \mathrm{vert}(C)$ vertices.

2.11 Ex (Extreme Point):

a) $C = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 \leq 1\} = B_1(0,0)$ is convex.

$\mathrm{ext}(C) = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 = 1\}$.

b)



$\mathrm{ext}(\mathbb{D}) = \{(0,2), (\frac{1}{2},6), (1,0), (4,4), (\frac{5}{2},2)\}$.

2.12 Def (Polyhedron, Polytope): Let $C \subseteq \mathbb{R}^n$. C is called

a) C convex, closed, $|\mathrm{ext}(C)| < \infty$ — polyhedron

b) C convex, compact, $|\mathrm{ext}(C)| < \infty$ — polytope

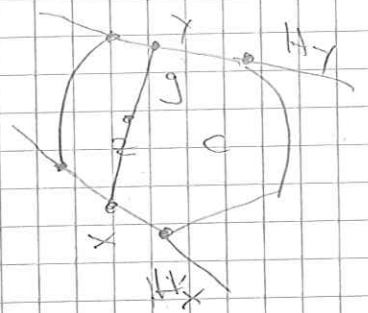
2.13 Thm (Rockafellar [1970]): Any non-empty compact convex set is the convex hull of its extreme points.

Proof: Rockafellar [1970] $\square$

2. Thm (generation of convex sets):

$C = \operatorname{conv} \operatorname{ext} C$ for any convex, compact $C \subseteq \mathbb{R}^n$.

Proof: By induction over n.



Ind. Start $n=0$ ✓

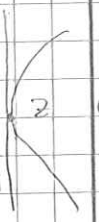
Ind Hyp. $n \geq 0$ ✓

Ind Step $n \to n+1$:

"$\supseteq$": ✓

"$\subseteq$": Let $z \in C$, g a line through z

C bounded $\Rightarrow g \cap C = \operatorname{conv}\{x_1, x_2\}$, $x_1, x_2 \in \partial C \subseteq C$ (closed - H)



Claim 1: $\forall p \in \partial C$: $\exists$ supporting hyperplane H: $C \subseteq H_{\geq}$, $p \in H$

Arg: Separating hyperplane claim ($\to$ Lotte) ◇

Claim 2: For any convex set $K \subseteq \mathbb{R}^n$ and any supp. hyperpl. H

$$\operatorname{ext}(K \cap H) = \operatorname{ext} K \cap H$$



Arg: "$\supseteq$" ✓, "$\subseteq$": Let $z \in \operatorname{ext}(K \cap H)$.

Assume $z \in \operatorname{conv}\{x_1, x_2\}$, $x_1, x_2 \in K \to x_1, x_2 \in H \to z = x_1$ ↯ ◇

Let H_{x_1}, H_{x_2} be supp. hyperplanes at $x_1, x_2 \in \partial K$.

Ind. Hyp. $\Rightarrow x \in \operatorname{conv} \operatorname{ext}(K \cap H_{x_1}) = \operatorname{conv} \operatorname{ext} K \cap H_{x_1} \subseteq \operatorname{conv} \operatorname{ext} K$

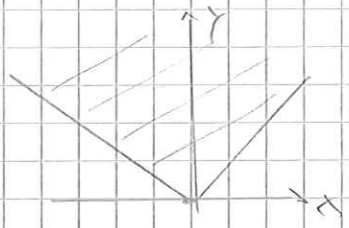
dim $H_{x_1}, H_{x_2} \leq n-1$ $y \in \operatorname{conv} \operatorname{ext}(K \cap H_{x_2}) = \operatorname{conv} \operatorname{ext} K \cap H_{x_2} \subseteq \operatorname{conv} \operatorname{ext} K$ □

26.4.17

Notation: $[x_1, x_2] = \{x_1 + \lambda(x_2 - x_1) : \lambda \in [0,1]\} = \{z \in \mathbb{R}^n : z = \lambda x_1 + (1-\lambda) x_2, \lambda \in [0,1]\}$

2.14 Ex (Extreme Point):

$C = \{(x,y) \in \mathbb{R}^2 : y \geq |x|\}$



$\Rightarrow ext(C) = \{(0,0)\} \Rightarrow C \neq conv(ext(C))$.

2.15 Def (Extreme Ray): Let $\emptyset \neq C \subseteq \mathbb{R}^n$ be closed, convex.

a) A vector $d \in \mathbb{R}^n \setminus \{0\}$ is called ray of C, if
$x + \alpha d \in C \quad \forall x \in C, \alpha \in \mathbb{R}$.

b) Two rays d_1, d_2 of C are called different, if
$d_1 \neq \beta d_2 \quad \forall \beta \in \mathbb{R}_{>0}$.

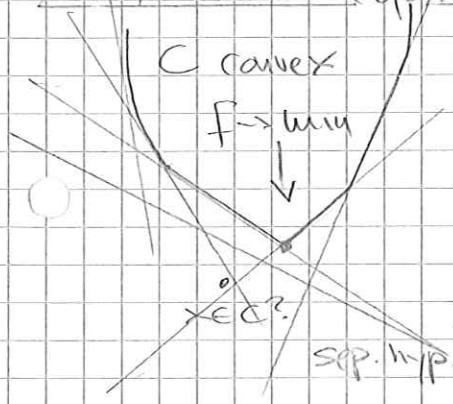
c) A ray d of C is called an extreme ray of C, if
d cannot be constructed as a conic (positive linear)
combination of two distinct rays d_1, d_2 of C, i.e., if
$d = \beta_1 d_1 + \beta_2 d_2, \ \beta_1, \beta_2 > 0 \Rightarrow \exists \mu > 0 : d_1 = \mu d_2$

2.14 Ex (cont'd):

$\{(x,y) \in \mathbb{R}^2 : \sphericalangle((x,y),(0,1)) \leq \frac{\pi}{4}\}$ rays of C

$\{(-1,1), (1,1)\}$ extreme rays of C.

2.16 Rem (optimization over convex sets):



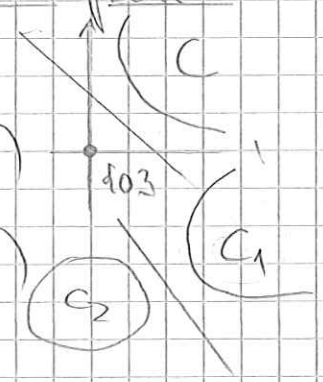
a) "encoding of C" $\Leftrightarrow x \in C$?

b) "separation" by hyperplanes $\Leftrightarrow x \notin C$

c) "outer description" by hyperplanes,
→ only partial description necessary
→ lin. ieqs useful for computation

d) "polynomial time equivalence of separation and optimization"

Program:



i) strict separation of $\{0\}$ and $C \not\ni \{0\}$

ii) strict separation of C_1 closed and C_2 compact

iii)  C separation of $\{0\}$ and $C \not\ni 0$

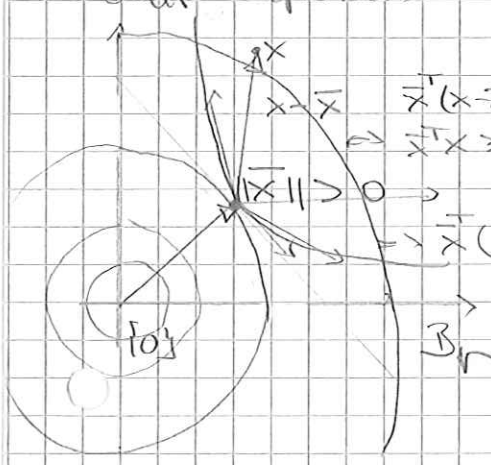
ii) C_1, C_2 separation of C_1 and C_2

2.17 Def. (separating Hyperplane): Let $M_1, M_2 \subseteq \mathbb{R}^n$ be arbitrary sets, $H := H_=(a,\alpha) \subseteq \mathbb{R}^n$ a hyperplane with opposing half spaces $H_\leq := H_\leq(a,\alpha)$, $H_\geq := H_\geq(a,\alpha)$, $H^\circ_\leq$, $H^\circ_\geq$. We say

H separates M_1, M_2 (weakly) $:\Leftrightarrow M_1 \subseteq H_\leq$, $M_2 \subseteq H_\geq$.

H separates M_1, M_2 strictly $:\Leftrightarrow M_1 \subseteq H^\circ_\leq$, $M_2 \subseteq H^\circ_\geq$.

2.18 Lemma (Sep. Hyperplanes): Let $C \subseteq \mathbb{R}^n$ non-empty, closed, convex, $0 \notin C$. Then there exists a hyperplane H that separates $\{0\}$ and C strictly.



Proof: Choose $r > 0$ s.t. $C \cap B_r(0) = \{x \in C : \|x\| \leq r\} \neq \emptyset$.

$C \cap B_r(0)$ is compact, continuous function

$\|\cdot\|^2 : C \cap B_r(0) \to \mathbb{R}_{\geq 0}$, $x \mapsto \|x\|^2$

attains min. over $C \cap B_r(0)$.

Let $a := \bar{x} \in \operatorname{Argmin}_{x \in C \cap B_r(0)} \|x\|^2$, $\alpha := \|\bar{x}\|^2$. Then

$\alpha = \|\bar{x}\|^2 > 0$ and $\|x\|^2 \geq \|\bar{x}\|^2$ $\forall x \in C$, in particular

for $\lambda x + (1-\lambda)\bar{x} \in C$, $\lambda \in (0,1]$

$\|\lambda x + (1-\lambda)\bar{x}\|^2 = \|\lambda(x-\bar{x}) + \bar{x}\|^2 \geq \|\bar{x}\|^2$

$\Leftrightarrow \lambda^2 (x-\bar{x})^T(x-\bar{x}) + 2\lambda(x-\bar{x})^T\bar{x} + \|\bar{x}\|^2 \geq \|\bar{x}\|^2$ $\cup$

$\lambda \to 0$ $\Rightarrow (x-\bar{x})^T\bar{x} \geq 0$ ($\bar{x} \in \mathbb{R}$)

$\hookrightarrow x^T\bar{x} - \bar{x}^T\bar{x} \geq 0 \Leftrightarrow \bar{x}^T x \geq \|\bar{x}\|^2 \Leftrightarrow a^T x \geq \alpha$

$\Rightarrow a^T x \geq \alpha > 0$ $\forall x \in C$ $\overset{x=0}{\Rightarrow} a^T 0 = 0 < \alpha$. $\square$

2.19 Thm (Strict Separation Thm): Let $C_1, C_2 \subset \mathbb{R}^n$ non-empty, disjoint, closed, convex, C_2 also bounded (i.e., compact). Then there exists a hyperplane H that separates C_1, C_2 strictly.

Proof: C_1, C_2 convex $\overset{\text{Thm 2.5}}{\Rightarrow}$ $C_1 + C_2$ convex

C_1, C_2 disjoint $\Rightarrow$ $C_1 - C_2 \neq \{0\}$.

C_1 closed, C_2 compact $\overset{(*)}{\Rightarrow}$ $C_1 - C_2$ closed.

$\overset{\text{Thm 2.18}}{\Rightarrow}$ $\exists\, H_=(a,\alpha)$ separating $C_1 - C_2$, hence strictly

$\Leftrightarrow$ $a^T x = a^T(u - v) \geq \alpha > 0 \quad \forall u \in C_1, v \in C_2$
($u \in C_1$, $v \in C_2$)

$\Rightarrow$ $a^T u \geq a^T v + \alpha > a^T v \quad \forall u \in C_1, v \in C_2$

$\Rightarrow$ $\inf_{u \in C_1} a^T u \geq \sup_{v \in C_2} a^T v + \alpha > \sup_{v \in C_2} a^T v \quad \forall u \in C_1, v \in C_2$ ($\alpha > 0$)

$\Rightarrow$ $\inf_{u \in C_1} a^T u > \underbrace{\inf_{u \in C_1} a^T u + \sup_{v \in C_2} a^T v}_{=:\beta} \geq \sup_{v \in C_2} a^T v$ (divide by 2)

$\Rightarrow$ $H_=(a, \beta)$ separates C_1, C_2 strictly.

(*) Let $C_1 - C_2 \ni x^i \to \bar{x}$, $x^i = x_1^i - x_2^i$ ($x_1^i \in C_1$, $x_2^i \in C_2$).

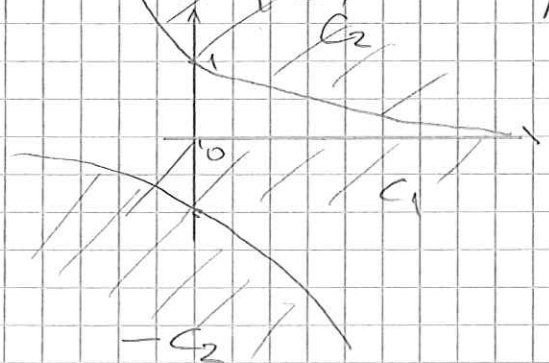
Sequential Weierstraß $\Rightarrow$ $\exists\, x_2^{ij} \to \bar{x}_2 \in C_2$

$\Rightarrow$ $x_1^{ij} - x_2^{ij} \to \bar{x} - \bar{x}_2 =: \bar{x}_1 \in C_1$

$\Rightarrow$ $\bar{x} \in C_1 - C_2$.

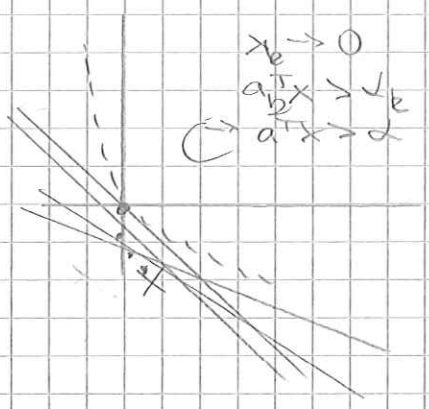
2.20 Ex (Strict separation Thm): Thm 2.19's assumption C_2 bounded cannot be omitted, cf., e.g.,

$C_1 = \{(x,y) \in \mathbb{R}^2 : y \leq 0\}$, $C_2 = \{(x,y) \in \mathbb{R}^2 : y \geq e^x\}$



C_1, C_2 non-empty, disjoint, closed, but $\nexists$ hyperplane sep. C_1, C_2 strictly.

2.21 Lem. (Weakly Sep. Hyperplane): Let $C \subseteq \mathbb{R}^n$ non-empty, convex, $0 \notin C$. Then there exists a hyperplane H that seperates $C, \{0\}$.



C convex $\Rightarrow \overline{C} = C \cup \partial C$ convex ($\to$ Ex)

a) $0 \notin \overline{C}$: Lem. 2.18

b) $0 \in \overline{C} \overset{0 \notin C}{\Rightarrow} 0 \in \partial C$

$\Rightarrow \exists (x_k)_{k=1}^{\infty} \to 0, x_k \in \overline{C} \quad \forall k \in \mathbb{N}$

$\overset{\text{Thm 2.19}}{\Rightarrow} \exists\ a_k^T x \geq \alpha_k \geq a_k^T x_k \quad \forall k \in \mathbb{N}, x \in \overline{C}$

$0 \in \overline{C} \Rightarrow a_k^T 0 \geq \alpha_k \geq a_k^T x_k \to 0$

$\Rightarrow \alpha_k \to 0$ o.B.d.A. $\|a_k\| = 1$

$a_{k_\ell} \to a, \quad \|a\| = 1$

$\Rightarrow a^T x \geq 0 \quad \forall x \in \overline{C}$. $\square$

2.22 Thm (Weak Separation Thm): Let $C_1, C_2 \subseteq \mathbb{R}^n$ be non-empty, disjoint, convex. Then there exists a hyperplane H that seperates C_1, C_2 weakly.

Proof: C_1, C_2 convex $\Rightarrow C_1 - C_2$ convex

C_1, C_2 disjoint $\Rightarrow C_1 - C_2 \not\ni 0$

$\overset{\text{Lem. 2.18}}{\Rightarrow} \exists\ H = (a, \beta)$ separating $C_1 - C_2, \{0\}$

$\Rightarrow a^T(u - v) \geq 0 \quad \forall u \in C_1, v \in C_2$

$\Rightarrow a^T u \geq a^T v \quad \forall u \in C_1, v \in C_2$

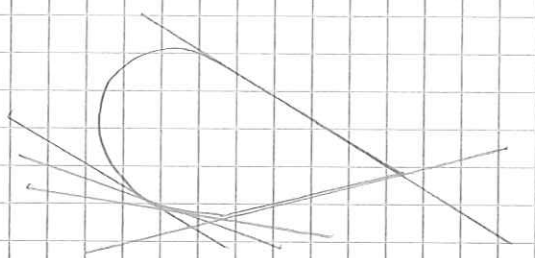
$\Rightarrow \inf_{u \in C_1} a^T u \geq \inf_{u \in C_1} a^T u + \sup_{v \in C_2} a^T v \geq \sup_{v \in C_2} a^T v$

$= : \beta$

$\Rightarrow H = (a, \beta)$ separates C_1, C_2. $\square$

02.05.17

2.23 Rem (Outer description of convex sets)



description by „supporting hyperplanes"

2.24 Def (Supporting Hyperplane): Let $C \subseteq \mathbb{R}^n$ non-empty, closed, convex, $H = H_=(a,\alpha)$ a hyperplane, $H_\leq = H_\leq(a,\alpha)$, $H_\geq = H_\geq(a,\alpha)$

H supporting hyperplane f.a. $C :\Leftrightarrow H \cap C \neq \emptyset$, $C \subseteq H_\leq \vee C \subseteq H_\geq$

$H_\leq \geq C$ resp. $H_\geq \geq C$ supporting half-space f.a. C

if $\left\{\begin{matrix} H_\leq \geq C \\ H_\geq \geq C \end{matrix}\right\}$, then $\left\{\begin{matrix} a \\ -a \end{matrix}\right\}$ is an outer normal f.a. C

2.25 Thm (Existence of Supp. Hyperplanes): Let $C \in \mathbb{R}^n$ non-empty, compact, convex, $a \in \mathbb{R}^n \setminus \{0\}$. Then there exists a supporting hyperplane with outer normal a.

Proof: $C \neq \emptyset$ compact, $f: \mathbb{R}^n \to \mathbb{R}$, $x \mapsto a^T x$ continuous.

$\Rightarrow \exists \alpha = \max_{x \in C} a^T x$

$\Rightarrow H_\leq(a,\alpha)$ is supp. hyperplane f.a. C with outer normal a. ☐

2.26 Thm (Outer description of convex sets): Every non-empty, closed, convex set $C \subseteq \mathbb{R}^n$ coincides with the intersection of its supp. halfspaces, i.e.,

$$C = \bigcap_{H_\leq \text{ supp. halfspace of } C} H_\leq$$

Proof: „$\subseteq$": „$\supseteq$": Ass. $\exists \bar{x} \in \bigcap H_\leq \setminus C$, w.l.o.g. $\bar{x} = 0$.

Thm 2.19 $\Rightarrow \exists$ hyperplane seperating $\{0\}$ and C strictly

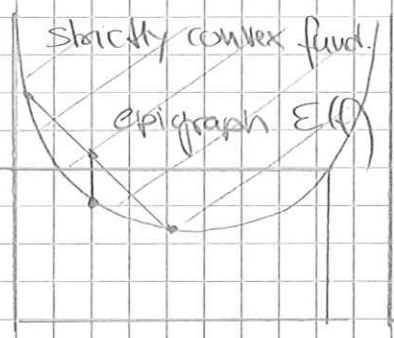
$\Rightarrow \exists r > 0 : B_r(0) \cap C = \emptyset \Rightarrow \|x\| \geq r \ \forall x \in C$.

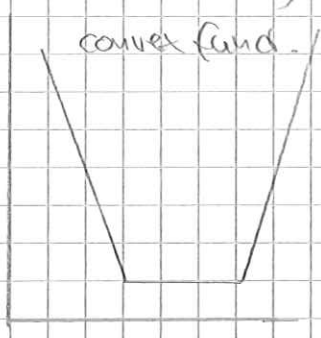
Choose $R > 0 : B_R(0) \cap C \neq \emptyset$ and set

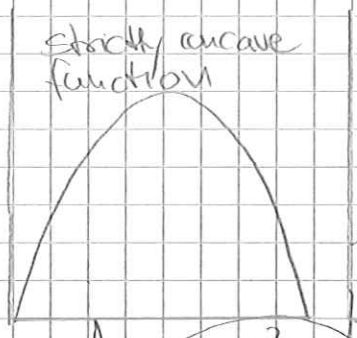
$\alpha := \min_{x \in B_R(0)} \|x\|^2$, $a \in \operatorname{Argmin}_{x \in B_R(0)} \|x\|^2$, i.e.,

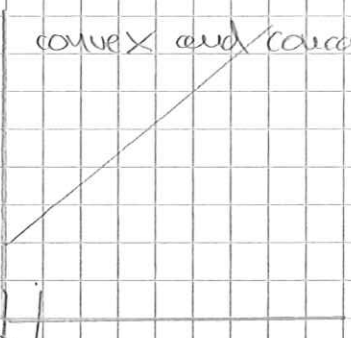
$\|x\|^2 \geq \|a\|^2 \ \forall x \in C$, in particular for

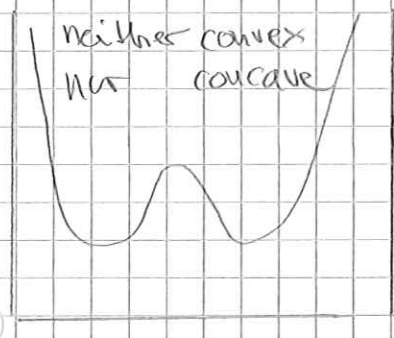
2.27 Rem (Convex Functions):



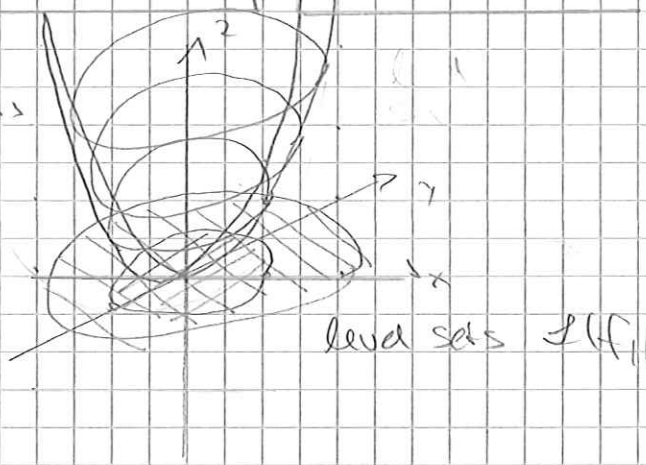












- f convex ⇔ E(f) convex
- f convex ⇒ L(f,β) convex
- f convex ⇒ f continuous in interior of its domain

2.28 Def. (Convex Function): Let $C \subseteq \mathbb{R}^n$ convex, $f: C \to \mathbb{R}$

a) f convex: ⇔ $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$ $\forall x_1, x_2 \in C$, $\lambda \in [0,1]$.

b) f strictly convex: ⇔ — " — $<$ — " — $\lambda \in (0,1)$.

c) f (strictly) concave: ⇔ $-f$ (strictly) convex.

2.29 Prop. (Convex combo and Convex Functions):

$f: C \to \mathbb{R}$ convex ⇔ $f\left(\sum_{i=1}^{p} \lambda_i x_i\right) \leq \sum_{i=1}^{p} \lambda_i f(x_i)$ $\forall x_i \in C$, $\lambda_i \geq 0$, $\sum \lambda_i = 1$.

Proof: ex. □

2.30 Thm (Computation Rules for convex Functions)

Let $C \subseteq \mathbb{R}^n$ convex, $f_1, f_2: C \to \mathbb{R}$ convex, $\alpha > 0$.

a) αf_1 is convex.

b) $f_1 + f_2$ is convex.

c) $\max\{f_1, f_2\}$ is convex.

Proof: ex. □

2.31 Rem (Comp. Rules for Convex Functions): Difference, product and minimum of convex functions are in general not convex.

2.32 Def. (Epigraph, level set): Let $C \in \mathbb{R}^n$, $f: C \to \mathbb{R}$, $\beta \in \mathbb{R}$

a) $\mathcal{E}(f) := \{(x,\nu) \in C \times \mathbb{R} : f(x) \leq \nu\}$ epigraph of f

b) $\mathcal{L}(f,\beta) := \{x \in C : f(x) \leq \beta\}$ level set of f at level β

2.33 Thm (Epigraph, level set): Let $C \in \mathbb{R}^n$ convex, $f: C \to \mathbb{R}$.

Then:

a) f convex $\Leftrightarrow$ $\mathcal{E}(f)$ convex

b) f convex $\Rightarrow$ $\mathcal{L}(f,\beta)$ convex $\forall \beta \in \mathbb{R}$.

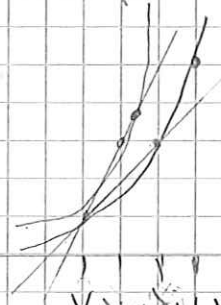
The converse of b) does not hold

Proof: $\Leftrightarrow$. □

2.34 Rem (Quasi-convex Functions): Functions with convex level sets are called quasi-convex.

2.35 Thm (Continuity of convex functions): A convex, bounded function $f: C \to \mathbb{R}$ is continuous in $\mathring{C}$.

Proof:



$\forall u, v, x \in C;\ \lambda \in (0,1)$:

$x = \lambda u + (1-\lambda) v \Rightarrow f(x) \leq \lambda f(u) + (1-\lambda) f(v)$

$\Leftrightarrow u = \frac{1}{\lambda} x + \left(1 - \frac{1}{\lambda}\right) v \Rightarrow f(u) \geq \frac{1}{\lambda} f(x) + \left(1 - \frac{1}{\lambda}\right) f(v)$

Let $M \geq f(v)$.

Let $v \in \mathring{C}$ and $B_\varepsilon(v) \subseteq \mathring{C}$, $\varepsilon > 0$.

Ass.: f not continuous in v.

$\Rightarrow \exists \varepsilon > 0 : \forall \delta > 0 : \exists x : \|x - v\| < \delta,\ \|f(x) - f(v)\| > \varepsilon$.

Choose $\lambda \in (0,1)$: $\frac{1}{\lambda} - 1 > \frac{M - f(v)}{\varepsilon} \Leftrightarrow f(v) + \frac{\varepsilon}{\lambda} > M + \varepsilon$

and $\delta := \lambda\varepsilon < \varepsilon$ $\Leftrightarrow \left(\frac{1}{\lambda} - 1\right)\varepsilon > M - f(v)$

$\Rightarrow \exists x : \|x - v\| < \delta,\ \|f(x) - f(v)\| > \varepsilon$

a) $f(x) > f(v) + \varepsilon$

$u := \frac{1}{\lambda} x + \left(1 - \frac{1}{\lambda}\right) v$ $\left(\Leftrightarrow x = \lambda u + (1-\lambda) v\right)$

$\Rightarrow \|u - v\| = \left\| \frac{1}{\lambda} x + \left(1 - \frac{1}{\lambda}\right) v - v \right\| = \frac{1}{\lambda} \underbrace{\|x - v\|}_{< \lambda\varepsilon} < \varepsilon \Rightarrow u \in B_\varepsilon(v) \subseteq \mathring{C}$

$\Rightarrow f(u) \geq \frac{1}{\lambda} \underbrace{f(x)}_{> f(v) + \varepsilon} + (1 - \frac{1}{\lambda}) f(v) = f(v) + \frac{\varepsilon}{\lambda} > M$ ↯

b) $f(x) < f(v) - \varepsilon$

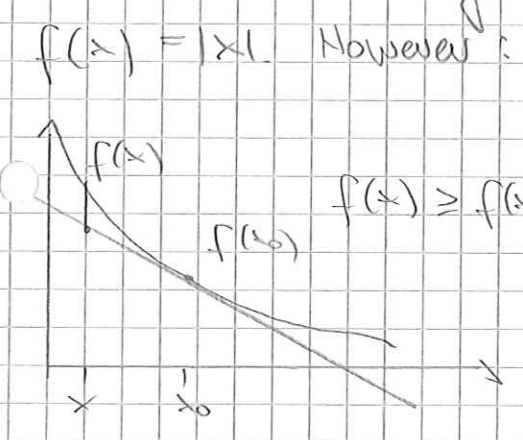
$u := \frac{\lambda}{\Lambda} v + \left(\lambda - \frac{\lambda}{\Lambda}\right) x$

$\Rightarrow \|u - v\| = \left\| \frac{\lambda}{\Lambda} v + \left(\lambda - \frac{\lambda}{\Lambda}\right) x - v \right\| = \left| \frac{\lambda - \Lambda}{\Lambda} \right| \underbrace{\|x - v\|}_{< \lambda e_1} \leq (\lambda - \lambda) e_1 \Rightarrow u \in \overset{\circ}{C}$

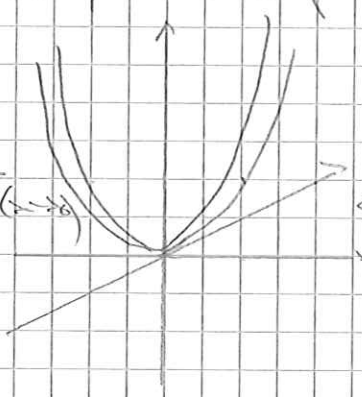
$\Rightarrow f(u) \geq \frac{\lambda}{\Lambda} f(v) + \underbrace{\left(\lambda - \frac{\lambda}{\Lambda}\right)}_{<0} \underbrace{f(x)}_{< f(v) - \varepsilon} > f(v) + \left(\frac{\lambda}{\Lambda} - \lambda\right)\varepsilon > \lambda \lightning$ □

ad a)

2.37 Rem. (Diff'bility of convex Functions): Convex functions are in general not everywhere diff'able, e.g. $f(x) = |x|$. However:



$f(x) \geq f(x_0) + \nabla f(x_0)^T (x - x_0)$



f convex $\Leftrightarrow$ $\nabla^2 f$ pos. semidef.

2.38 Thm (Diff'able Convex Functions): Let $C \subseteq \mathbb{R}^n$ convex, $f: \mathbb{R}^n \to \mathbb{R}$ cont. diff'able. Then:

a) f convex on $C \Leftrightarrow f(x) \geq f(x_0) + \nabla f(x_0)^T (x - x_0) \quad \forall x_0, x \in C$

b) f strictly — $\Leftrightarrow$ — $\Rightarrow$ — "

In particular, $H \geq (\nabla f(x_0), f(x_0))$ supports $E(f)$ at x_0.

Proof:

"$\Leftarrow$" Let $x_1, x_2 \in C$, $\lambda \in (0,1) \Rightarrow z := \lambda x_1 + (1-\lambda) x_2 \in C$

$\lambda \left[f(x_1) \geq f(z) + \nabla f(z)^T (x_1 - z) \right]$

$+ (1-\lambda) \left[f(x_2) \geq f(z) + \nabla f(z)^T (x_2 - z) \right]$

$\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(z) + \nabla f(z)^T \left(\underbrace{\lambda x_1 + (1-\lambda) x_2}_{= z} - z \right)$

$= f(\lambda x_1 + (1-\lambda) x_2)$.

"$\Rightarrow$" Let $h: \mathbb{R} \to \mathbb{R}$, $\lambda \mapsto (1-\lambda) f(x_0) + \lambda f(x) - f((1-\lambda) x_0 + \lambda x)$

f convex $\Rightarrow h(\lambda) \geq 0$, $\lambda \in [0,1]$, $h(0) = 0$

$\Rightarrow \left. \frac{dh}{d\lambda} \right|_{\lambda = 0} = -f(x_0) + f(x) - \nabla f(x_0)^T (x - x_0) \geq 0$

$\Rightarrow f(x) \geq f(x_0) + \nabla f(x_0)^T (x - x_0)$.

b) "$\Leftarrow$": Analogous.

"$\Rightarrow$": Suppose $f(x) \geq f(x_0) + \nabla f(x_0)^T (x - x_0)$ $\forall x$

$\Rightarrow f(\lambda x + (1-\lambda) x_0) \begin{cases} \geq \\ \leq \end{cases} \lambda f(x) + (1-\lambda) f(x_0) \geq f(x_0) + \nabla f(x_0)^T \lambda (x - x_0)$ ↯. □

2.39 Thm (Convexity and the Hessian): Let $C \subseteq \mathbb{R}^n$ be convex open, $f: C \to \mathbb{R} \in C^2$, i.e., twice cont. diff'able. Then:

a) f convex $\Leftrightarrow \nabla^2 f(x)$ pos. semidef. $\forall x \in \mathbb{R}^n$

b) f strictly convex $\Leftarrow \nabla^2 f(x)$ pos. def. $\forall x \in \mathbb{R}^n$

Proof:

a) "$\Leftarrow$": Taylor expansion

$$f(x) = f(x_0) + \nabla f(x_0)^T (x - x_0) + \frac{1}{2} \underbrace{(x - x_0)^T \nabla^2 f((x_0 + t(x - x_0))(x - x_0)}_{\geq 0} \quad t \in [0,1]$$

$$\overset{(*)}{\geq} f(x_0) + \nabla f(x_0)^T (x - x_0)$$

"$\Rightarrow$": Ass.: $\exists x_0 \in \mathbb{R}^n$: $\nabla^2 f(x_0)$ not pos. semidef.

$\Rightarrow \exists x \in \mathbb{R}^n$: $(x - x_0)^T \nabla^2 f(x_0)(x - x_0) < 0$

$\Rightarrow t (x - x_0)^T \nabla^2 f(x_0) t (x - x_0) < 0 \quad \forall t > 0$

$\Rightarrow \exists x \in \mathbb{R}^n \quad (x - x_0)^T \nabla^2 f(x_0 + t(x - x_0))(x - x_0) < 0$ ↯ (*). ($\nabla^2 f \in C$, $t \to 0$)

b) Analogously. □

2.40 Rem (Convexity and the Hessian): f strictly convex $\Rightarrow \nabla^2 f(x)$ pos. def. $\forall x \in C$ does in general not hold, e.g., $f(x) = x^4$.

2.41 Ex (Quadratic Functions):

a) $f(x) = \alpha + a^T x + \frac{1}{2} x^T Q x$, $\nabla^2 f = Q$

f (strictly) convex $\Leftrightarrow Q$ pos. (semi) def.

f (strictly) concave $\Leftrightarrow Q$ neg. (semi) def.

b) $f(x_1, x_2) = x_1^2 + x_2^2 - 4 x_1 x_2 = (x_1 + x_2) \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$$= \frac{1}{2} (x_1 + x_2) \begin{pmatrix} 2 & -4 \\ -4 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{2} x^T Q x$$

$$\begin{vmatrix} 2-\lambda & -4 \\ -4 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 16 = (\lambda + 2)(\lambda - 6) = 0$$

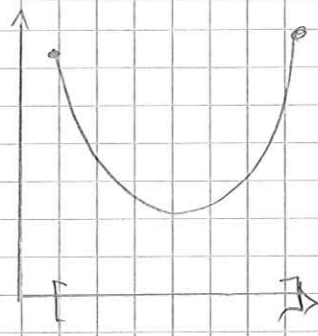
$\Rightarrow \Lambda(Q) = \{-2, 6\}$, i.e., f is neither convex nor concave

10.05.17

c) $f(x_1,x_2) = x_1^2 + x_2^2 = \frac{1}{2}(x_1,x_2)\begin{pmatrix}2&0\\0&2\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \frac{1}{2}x^TQx$

$\Rightarrow \Delta(Q) = \{2\} \Rightarrow f$ strictly convex

2.4 Theorem (Optimality results for convex functions):



- local optimum = global optimum
- strictly convex $\Rightarrow$ unique optimum
  convex $\Rightarrow$ Argmin f convex
- f diff'able, $x_0 \in \text{Argmin } f \Rightarrow \nabla f(x_0) = 0$
- Argmax $f: C \to \mathbb{R}$ $\cap$ ext$(C) \neq \emptyset$.

2.43 Thm (local and global Optima): Let $f: C \to \mathbb{R}$ convex.

x_0 local min of f on $C \Leftrightarrow x_0$ global min of f on C.

Proof: Let $x_0 \in C$ be a local min of f. Ass.: $\exists \bar{x} \in C: f(\bar{x}) < f(x_0)$

$\Rightarrow f(x_0 + t(\bar{x} - x_0)) \leq t f(\bar{x}) + (1-t) f(x_0) < f(x_0) \quad \forall t \in (0,1]$ ↯ □

2.44 Thm (Set of Minimizers): Let $f: C \to \mathbb{R}^n$ be convex. Then

Argmin $f(C)$ is convex.

Proof: ⇒ □

2.45 Cor (Unique Optimum): Let $f: C \to \mathbb{R}^n$ be strictly convex. Then

Argmin $f(C)$ consists of a single point.

Proof: Ass.: $x_1, x_2 \in \text{Argmin } f(C)$, $x_1 \neq x_2$

$\Rightarrow f(\lambda x_1 + (1-\lambda)x_2) < \lambda f(x_1) + (1-\lambda) x_2 \quad \forall \lambda \in (0,1)$ ↯ □

2.46 Thm (1st order condition): Let $f: \mathbb{R}^n \to \mathbb{R}$ convex, diff'able. Then

$x_0 \in \text{Argmin } f(C) \Leftrightarrow \nabla f(x_0) = 0$.

Proof: "$\Rightarrow$" Calculus

"$\Leftarrow$" $f(x) \geq f(x_0) + \nabla f(x_0)^T (x - x_0) = f(x_0)$. □

2.47 Thm (Max. of convex functions): Let $C \in \mathbb{R}^n$ be

convex, compact, $f: C \to \mathbb{R}$ convex. Then

Argmin $f(C) \cap$ ext$(C) \neq \emptyset$

i.e., f always its max. in an extreme point of C.

Proof: Let $x \in C$, then $\exists v_0, \dots, v_n: x = \sum_{i=0}^{n} \lambda_i v_i$, $\lambda_i \geq 0$, $\sum \lambda_i = 1$

3.6 Rem (Representation of Polyhedra): The solution set of the following system is a polyhedron:

$$Bx + Cy = c \qquad Bx + Cy \le c$$
$$Dx + Ey \le d \quad \leftrightarrow \quad -Bx - Cy \le -c \quad \leftrightarrow \quad \begin{pmatrix} B & C \\ -B & -C \\ D & E \\ -I & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \le \begin{pmatrix} c \\ -c \\ d \\ 0 \end{pmatrix}$$
$$x \ge 0 \qquad Dx + Ey \le d$$
$$x \in \mathbb{R}^p \qquad -x \le 0$$
$$y \in \mathbb{R}^q \qquad (x,y) \in \mathbb{R}^{p+q} \qquad \underbrace{\begin{pmatrix} B & C \\ -B & -C \\ D & E \\ -I & 0 \end{pmatrix}}_{=:A} \le \underbrace{\begin{pmatrix} c \\ -c \\ d \\ 0 \end{pmatrix}}_{=:b}$$

$B \in \mathbb{R}^{r \times p}$, $C \in \mathbb{R}^{r \times q}$, $c \in \mathbb{R}^r$, $D \in \mathbb{R}^{s \times p}$, $E \in \mathbb{R}^{s \times q}$, $d \in \mathbb{R}^s$

3.7 Rem. (Transformations of Polyhedra): Let $A \in \mathbb{R}^{m \times n}$, $a, b \in \mathbb{R}^m$, $\alpha \in \mathbb{R}$, $P^= (A,b) := \{x \in \mathbb{R}^n : Ax = b, x \ge 0\}$.

Rule I (Introduction of Slack Variables):

$$\{x \in \mathbb{R}^n : a^T x \le \alpha\} = \{x \in \mathbb{R}^n : a^T x + s = \alpha, s \ge 0\}.$$

Rem II (Introduction of sign-restricted variables):

$$\{x \in \mathbb{R}^n\} = \{x \in \mathbb{R}^n : x = x^+ - x^-, x^+, x^- \ge 0\}.$$

Rule III (Replacing equations by inequalities):

$$\{x \in \mathbb{R}^n : a^T x = \alpha\} = \{x \in \mathbb{R}^n : a^T x \le \alpha, -a^T x \le -\alpha\}.$$

Using rules I – III, a polyhedron of the form $P(D,d)$ can be transformed into the form $P^=(A,b)$ and vice versa:

$$P(D,d) \underset{\text{II}}{\overset{\text{I, II}}{\rightleftarrows}} P^=(A,b).$$

3.8 Def (Cone): Let $K \subseteq \mathbb{R}^n$.

K cone $:\Leftrightarrow \forall x \in K \ \forall \alpha \ge 0 : \alpha x \in K$ $(\Rightarrow 0 \in K)$

K polyhedral cone $:\Leftrightarrow K$ cone $\wedge$ K polyhedron

3.9 Rem (Representation of Polyhedral Cones): Let $K \subseteq \mathbb{R}^n$.

K polyhedral cone $\Leftrightarrow \exists A \in \mathbb{R}^{m \times n} : K = P(A,0)$.

Proof: "$\Leftarrow$" ✓ "$\Rightarrow$": Let $K = P(A,b)$. $0 \in K \Rightarrow b \ge 0$.

Claim: $b = 0$. Ass.: $\exists i : b_i \ne 0$. Let $A_i = (a_{i1}, \dots, a_{in})$ be the i-th row of A. Ass.: $\exists \bar{x} \in K : A_i \bar{x} = b_i > 0$

$\Rightarrow A_i \cdot \alpha \bar{x} \xrightarrow[\alpha \to \infty]{} \infty > b_i \Rightarrow A_i \cdot \bar{x} \le 0 \ \forall \bar{x} \in K \Rightarrow b_i = 0$. ☐

3.2 Def. (Polyhedron): Let $P \subseteq \mathbb{R}^n$

P polyhedron $\Leftrightarrow$ P intersection of finite number of halfspaces

P polytope $\Leftrightarrow$ P bounded polyhedron



Greek	Engl.
πολύς	many
hedra	seat (ing area)
topos	location

3.3 Rem. (Convexity of Polyhedra): Polyhedra are convex.

Proof: The intersection of convex sets is convex. □

3.4 Rem. (Representation of Polyhedra by Systems of Linear Inequalities): Any polyhedron can be represented as

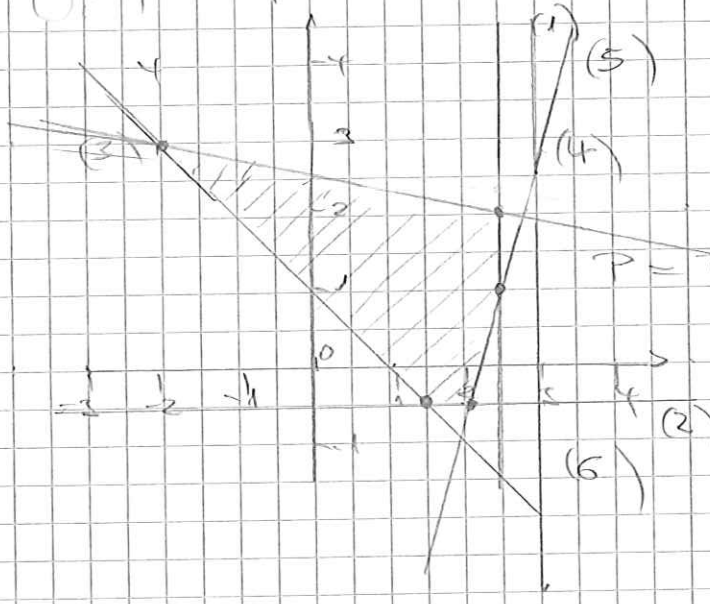
$$P = P(A,b) = \{x \in \mathbb{R}^n : Ax \leq b\},$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $m \in \mathbb{N}$.

3.5 Ex. (Representation of Polyhedra):

$$\begin{array}{lrcl}
(1) & 2x_1 & \leq 5 & \\
(2) & -2x_2 & \leq 1 & \\
(3) & -x_1 - x_2 & \leq -1 & \Leftrightarrow \\
(4) & 2x_1 + 9x_2 & \leq 23 & \\
(5) & 6x_1 - 2x_2 & \leq 13 & \\
(6) & x_1 & \leq 3 &
\end{array}
\quad
\begin{pmatrix} 2 & 0 \\ 0 & -2 \\ -1 & -1 \\ 2 & 9 \\ 6 & -2 \\ 1 & 0 \end{pmatrix}
\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}
\leq
\begin{pmatrix} 5 \\ 1 \\ -1 \\ 23 \\ 13 \\ 3 \end{pmatrix}$$

$$\begin{array}{cc|c} -1 & -1 & -1 \\ 0 & 7 & 21 \\ & 1 & 3 \end{array}$$



$P = P(A,b)$

6) can be dropped, i.e., is redundant

$$\text{ext}(P) = \left\{ (-2,3), \left(\tfrac{3}{2}, -\tfrac{1}{2}\right), \left(2, -\tfrac{1}{2}\right), \left(\tfrac{5}{2}, 1\right), \left(\tfrac{5}{2}, \tfrac{1}{2}\right) \right\}$$

$\Rightarrow f(x) = f\left(\sum \lambda_i v_i\right) \leq \sum \lambda_i f(v_i) \leq \max_{i=0}^{4} f(v_i)$. □

2.48 Cor. (Min / Max of a linear function): A linear function attains its min and max on a compact set in an extreme point.

Proof: Thm. 2.47 for $f_1 - f$. □

2.49 Thm (Max. on a compact set): Let $M \subseteq \mathbb{R}^4$ compact, $f: \mathbb{R}^4 \to \mathbb{R}$ convex. Then

$$\max f(M) = \max f(\text{ext } M).$$

Proof: Cf. □

3. Polytopes and Polyhedra

3.1 Ex (Lin. Optimization Problem):

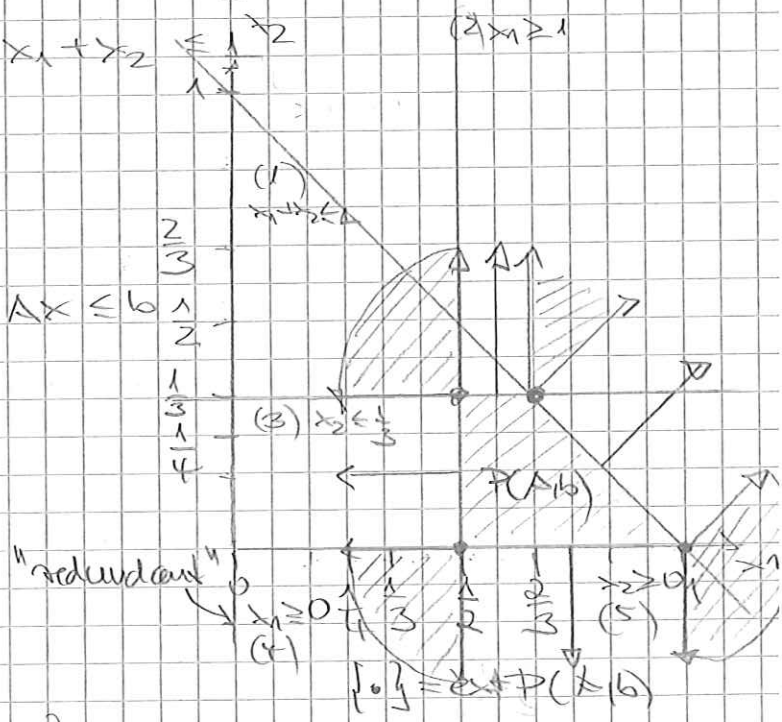
x_1 fraction of time studying: $\frac{1}{2} \leq x_1 \leq 1$

x_2 " job $0 \leq x_2 \leq \frac{1}{3}$

total time available $x_1 + x_2 \leq 1$

$$\max \; c_1 x_1 + c_2 x_2$$

$$\text{s.t.} \quad \begin{aligned} x_1 + x_2 &\leq 1 \quad (1) \\ -x_1 &\leq -\tfrac{1}{2} \quad (2) \\ x_2 &\leq \tfrac{1}{3} \quad (3) \\ -x_1 &\leq 0 \quad (4) \\ -x_2 &\leq 0 \quad (5) \end{aligned} \qquad Ax \leq b$$



$$\{Ax \leq b\} = P(A,b)$$

$$= \min -c_1 x_1 - c_2 x_2$$

$$\left.\begin{aligned} x_1 + x_2 &= 1 \\ -x_1 + s_1 &= -\tfrac{1}{2} \\ x_2 + s_2 &= \tfrac{1}{3} \end{aligned}\right\} \; A'x = b' \qquad x' \geq 0$$

$$x_1, x_2, s_1, s_2 \geq 0$$

$$\{A'x' = b', x' \geq 0\} = P^=(A', b')$$

3.10 Def. (valid inequality): Let $M \subseteq \mathbb{R}^n$, $a \in \mathbb{R}^n$ ($a \neq 0$ possible), $\alpha \in \mathbb{R}$

$a^T x \leq \alpha$ / (a, α) } valid for $M :\Leftrightarrow M \subseteq \{x \in \mathbb{R}^n : a^T x \leq \alpha\}$.

3.11 Def. (Face): Let $P \subseteq \mathbb{R}^n$ polyhedron, $F \subseteq P$.

a) F face of $P :\Leftrightarrow \exists a^T x \leq \alpha$ valid for $P : F = P \cap \{a^T x = \alpha\}$, $F \neq \emptyset$

b) F proper — " — $:\Leftrightarrow F$ face of P, $F \neq P$

c) F non-trivial — " — $:\Leftrightarrow F$ face of P, $\emptyset \neq F \neq P$.

d) $a^T x \leq \alpha$ valid for $P \Rightarrow F := P \cap \{a^T x = \alpha\}$ face induced by $a^T x \leq \alpha$

3.5 Ex (could):

a) $F_1 = \operatorname{conv}\left\{\left(\frac{5}{2}, 1\right), \left(\frac{5}{2}, 2\right)\right\} = P \cap \{(x_1, x_2) \in \mathbb{R}^2 : 2x_1 = 5\}$

is induced by (1) $2x_1 \leq 5$.

b) $F_2 = \{(-2, 3)\} = P \cap \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 3\}$ ind. by $x_2 \leq 3$

$= P \cap \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = -2\}$ ind. by $x_1 \geq -2$

$= P \cap \{(x_1, x_2) \in \mathbb{R}^2 : -x_1 - x_2 = -1,$

$2x_1 + 9x_2 = -23\}$ ind by (3), (4). !!

i.e., faces can be induced by different ineqs.

c) $F_3 = P = P \cap \{(x_1, x_2) \in \mathbb{R}^2 : 0x_1 + 0x_2 = 0\}$ is ind. by $0^T x \leq 0$

d) $F_4 = \emptyset$ is not induced by an inequality.

3.6 Rem. (Faces of Polyhedra): Let $P \subseteq \mathbb{R}^n$ be a polyhedron:

a) P is a face of P

b) $\emptyset$ — " —

c) $F = P \cap \{x : d^T x = \delta\}$ non-trivial face of $P \Rightarrow d \neq 0$

Proof:

a) $P = P \cap \{0^T x = 0\}$

b) By def.

c) Ass.: $d = 0$, $0^T x \leq \delta$ valid $\overset{a)}{\Rightarrow} \delta \geq 0 \Rightarrow F = P$ ↯. □

3.7 Thm (Optimal Face): Let $P = P(A, b) \subseteq \mathbb{R}^n$, $P \neq \emptyset$, $c \in \mathbb{R}^n$.

$X := \operatorname{Argmax}(c, P) := \{x \in P : c^T x = \max c^T x, x \in P\}$.

$\gamma = \begin{cases} c^T x & \text{if } X \neq \emptyset \\ \infty & \text{else} \end{cases}$, $\max\{c^T x : x \in P\}$

Then:

a) $\overline{X} \neq \emptyset \Rightarrow \overline{X} = P \cap \{ c^T x = z \}$, i.e. the set $\overline{X}$ of opt. sol. of an LP is a non-empty face of P induced by the supporting hyperplane $c^T x \leq z$.

b) $\overline{X}$ is a face of P.

Proof:

a) $c^T x \leq z$ is valid for P b) $\overline{X} = \emptyset$ ✓, $\overline{X} \neq \emptyset$ a) □

3.8 Def (Row/Col. Ind Set): Let $A \in \mathbb{R}^{m \times n}$, $M := [m] := \{1, \dots, m\}$, $I \subseteq M$, $N := [n] := \{1, \dots, n\}$, $J \subseteq N$

$A_{i\cdot} := (a_{i1}, \dots, a_{in})$, $i \in M$ i-th row of A

$A_{\cdot j} := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$, $j \in N$ j-th column of A

$A_{I\cdot} := (A_{i\cdot})_{i \in I}$ submatrix of A with rows in I

$A_{\cdot J} := (A_{\cdot j})_{j \in J}$ — " — cols in J

$A_{IJ} := (a_{ij})_{\substack{i \in I \\ j \in J}}$ — " — rows in I, cols in J.

3.9 Def (Binding or Active Restrictions of a Face):

Let $P = P(A,b) \subseteq \mathbb{R}^n$, $M = [m]$ row index set of A, $F \subseteq P$ a face of P, $I \subseteq M$.

$eq(F) := \{ i \in M : A_{i\cdot} x = b_i \ \forall x \in F \}$ equality set of F

$fa(I) := \{ x \in P : A_{I\cdot} x = b_I \}$ face of P induced by I

3.10 Rem (Equality Set):

a) $eq(F)$ is the set of inequalities that are binding $\forall x \in F$.

b) $fa(I)$ is a face of P and hence well-defined, because

$A_{i\cdot} x = b_i \ \forall i \in I \Leftrightarrow \exists \lambda_i \ A_{i\cdot} x = \sum_{i \in I} \lambda_i b_i$, i.e.

$fa(I) = P \cap \{ x \in \mathbb{R}^n : \sum_{i \in I} A_{i\cdot} x = \sum_{i \in I} b_i \}$.

3.5 Ex (cont'd): $fa(\{3,4\}) = \{(-2,3)\}$,

$eq(\text{conv}\{(\tfrac{5}{2},1), (\tfrac{5}{2},2)\}) = \{1\}$.

3.11 Def (Aff. Independence and Combination): Let $x_0, \dots, x_p \in \mathbb{R}^n$

a) $\sum_{i=0}^p \lambda_i x_i = 0 \ \wedge \ \sum_{i=0}^p \lambda_i = 0 \Rightarrow \lambda_1 = \dots = \lambda_p = 0 \Leftrightarrow x_0, \dots, x_p$ aff. independent.

b) $\exists (\lambda_0, \dots, \lambda_p) \neq (0, \dots, 0) : \sum_{i=0}^p \lambda_i x_i = 0, \ \sum_{i=0}^p \lambda_i = 0 \Leftrightarrow x_0, \dots, x_p$ aff. dep. (26)

c) $x_0 = \sum_{i=1}^{p} \lambda_i x_i$, $\sum_{i=1}^{p} \lambda_i = 1$ $\rightsquigarrow$ x_0 aff. combination of $x_1, \dots, x_p$

d) $\operatorname{aff}(M) := \{$ "x is aff. comb. of points in M" $\}$ affine hull of M

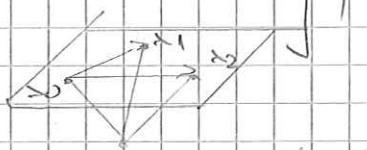
3.12 Obs. (Aff. Ind. and Comb.): Let $x_0, \dots, x_p \in \mathbb{R}^n$.

a) $x_0 = \sum_{i=1}^{p} \lambda_i x_i$, $\sum_{i=1}^{p} \lambda_i = 1$ $\overset{\lambda_0 := -1}{\Leftrightarrow}$ $\sum_{i=0}^{p} \lambda_i x_i = 0$, $\sum_{i=0}^{p} \lambda_i = 0$, $(\lambda_0, \dots, \lambda_p) \neq 0$

b) $\sum_{i=0}^{p} \lambda_i x_i = 0$, $\sum_{i=0}^{p} \lambda_i = 0$, $\lambda \neq 0 \overset{\text{W.l.o.g. } \lambda_0 = -1}{\Rightarrow}$ $x_0 = \sum_{i=1}^{p} \lambda_i x_i$, $\sum_{i=1}^{p} \lambda_i = 1$.

c) $\left[\sum_{i=0}^{p} \lambda_i x_i = 0,\ \sum_{i=0}^{p} \lambda_i = 0 \Rightarrow \lambda = 0 \right] \Leftrightarrow \left[\sum \lambda_i \begin{pmatrix} x_i \\ 1 \end{pmatrix} = 0 \Rightarrow \lambda = 0 \right]$,

i.e., x_i aff. ind. $\Leftrightarrow$ $\begin{pmatrix} x_i \\ 1 \end{pmatrix}$ lin. ind.



d) $x_0, \dots, x_p$ aff. ind. $\Leftrightarrow$ $x_1 - x_0, \dots, x_p - x_0$ lin. ind.

Proof: a), b), c) ✓ d)

$\Rightarrow$: Ass. $x_1 - x_0, \dots, x_p - x_0$ lin. dep. $\Rightarrow \exists (\lambda_1, \dots, \lambda_p) \neq (0, \dots, 0)$:

$\sum_{i=1}^{p} \lambda_i (x_i - x_0) = 0 \Leftrightarrow \sum_{i=1}^{p} \lambda_i x_i = \sum_{i=1}^{p} \lambda_i x_0$

i) $\sum_{i=1}^{p} \lambda_i \neq 0 \Rightarrow x_0 = \sum_{i=1}^{p} \frac{\lambda_i}{\sum \lambda_i} x_i$, $\sum_{i=1}^{p} \lambda_i = 1$ ↯.

ii) $\sum_{i=1}^{p} \lambda_i = 0 \Rightarrow \sum_{i=1}^{p} \lambda_i x_i = 0$, $\sum_{i=1}^{p} \lambda_i = 0$, $(\lambda_1, \dots, \lambda_p) \neq (0, \dots, 0)$,

i.e., $x_1, \dots, x_p$ aff. dep. ↯

$\Leftarrow$: Ass. $x_0, \dots, x_p$ aff. dep. $\Rightarrow \exists (\lambda_0, \dots, \lambda_p) \neq (0, \dots, 0)$:

$\sum_{i=0}^{p} \lambda_i x_i = 0$, $\sum_{i=0}^{p} \lambda_i = 0$.

i) $\lambda_0 \neq 0$: $\lambda_0 = \sum_{i=1}^{p} (-\lambda_i)$, $x_0 = \sum_{i=1}^{p} \frac{(-\lambda_i)}{\sum (-\lambda_i)} x_i$ ↯.

ii) $\lambda_0 = 0$: $\sum_{i=1}^{p} \lambda_i (x_i - x_0) = \underbrace{\sum_{i=1}^{p} \lambda_i x_i}_{=0} + \underbrace{\sum_{i=1}^{p} \lambda_i}_{=0} x_0 = 0$ ↯. □

3.13 Def. (Dimension of a convex set): Let $\emptyset \neq C \subseteq \mathbb{R}^n$ be convex.

a) $\dim C := \max \{ d \in \mathbb{N}_0 : x_0, \dots, x_p$ aff. independent $\}$

dimension of C

b) C full dimensional $:\Leftrightarrow \dim C = n$.

c) $\dim \emptyset := -1$.

3.14 Obs. (Dimension of a convex set): Let $\emptyset \neq C \subseteq \mathbb{R}^n$.

$\dim C = \dim \operatorname{aff}(C)$.

3.15 Def (redundancy): Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $D \in \mathbb{R}^{k \times n}$, $d \in \mathbb{R}^k$, $m = [m]$, $k = [k]$, F a face of $P = P(A,b)$.

a) $A_I x \leq b_I$ redundant $(w.r.t.\ Ax \leq b) \Leftrightarrow P(A_{M \setminus I}, b_{M \setminus I}) = P(A,b)$

b) $A_i x \leq b_i$ — " — $\Leftrightarrow P(A_{M \setminus i}, b_{M \setminus i}) = P(A,b)$

c) $Ax \leq b$ $\left\{\begin{matrix} \text{redundant} \\ \text{non-redundant} \end{matrix}\right\} \Leftrightarrow \left\{\begin{matrix} \exists i \in M \\ \nexists i \in M \end{matrix}\right\}$ $A_i x \leq a$ redundant

d) $A_i x \leq b_i$ implicit equation $\Leftrightarrow i \in eq(P(A,b))$.

e) $\begin{matrix} Ax \leq a \\ Dx = d \end{matrix}$ non-redundant $\Leftrightarrow$ rank $D = k$ and

$$P\left(\begin{pmatrix} A_{M \setminus i} \\ D \end{pmatrix}\begin{pmatrix} b_{M \setminus i} \\ d \end{pmatrix}\right) \overset{?}{\neq} P\left(\begin{pmatrix} A \\ D \\ -D \end{pmatrix}\begin{pmatrix} b \\ d \\ -d \end{pmatrix}\right) \quad \forall i \in M$$

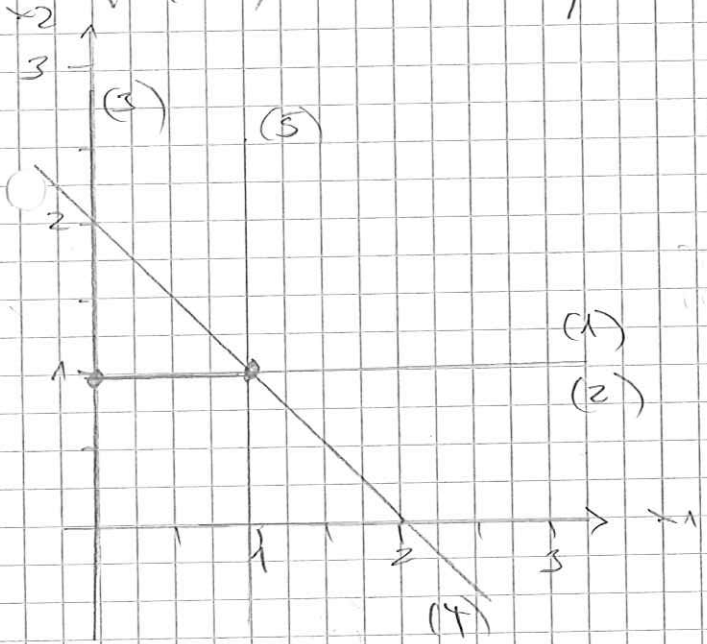
f) F facet (of P) $:\Leftrightarrow [F \subsetneq G \neq P$, G face of $P \Rightarrow F = G]$

g) F vertex (of P) $:\Leftrightarrow \dim F = 0$

3.16 Ex (redundancy, facets, vertices): Consider $P(A,b)$ for

$$A = \begin{pmatrix} 0 & 1 \\ 0 & -1 \\ -1 & 0 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 2 \\ 1 \end{pmatrix}$$

(1) $x_2 \leq 1$
(2) $-x_2 \leq -1$ $\Leftrightarrow x_2 = 1$
(3) $-x_1 \leq 0$
(4) $x_1 + x_2 \leq 2$ redundant
(5) $x_1 \leq 1$



(4) is redundant

(5) is redundant

(4), (5) is not redundant

(1), (2), (3) are "essential"

(1), (2), (3), (4) is irredundant

$x_2 = 2$, (3), (4) is irredundant

$\dim P(A,b) = 1$

$\{(0,1)\}$ is a facet and a vertex

$\{(1,1)\}$ is a facet and a vertex

3.17 Def (Vertex): Let $P \subseteq \mathbb{R}^n$ be a polyhedron, $F \subseteq P$ a face

F vertex (of P) $:\Leftrightarrow \dim F = 0$.

3.18 Thm (Vertices of Polyhedra): The following statements are equiv.:

a) x is a vertex of P

b) $\{x\}$ is a 0-dim. face of P

c) $\forall y,z \in P, y \neq z : x \in (y,z)$, i.e., x is not a proper convex combination of points in P

d) $\operatorname{rank} A_{\operatorname{eq}(\{x\})} = n$

e) $\exists c \in \mathbb{R}^n \setminus \{0\} : \{x\} = \operatorname{Argmax}(c, P)$ i.e., x is the unique opt. sol. of the LP $\max c^T x, x \in P$.

Proof:

a) $\Leftrightarrow$ b): Def. 3.17

b) $\Rightarrow$ e): $\{x\}$ face of $\Rightarrow \exists$ valid ineq. $c^T x \leq \gamma : \{x\} = \{y \in P : c^T y = \gamma\}$

$\Rightarrow \{x\} = \operatorname{Argmax}(c, P)$.

i) $c \neq 0$ ✓

ii) $c = 0 \Rightarrow \gamma = 0 \Rightarrow P = \{0\} = \mathbb{R}^0$. Choose $c = 1$. ✓

e) $\Rightarrow$ c): Ass.: $\exists y,z \in P, y \neq z, \lambda \in (0,1) : x = \lambda y + (1-\lambda) z$

$\Rightarrow \gamma = c^T x = c^T[\lambda y + (1-\lambda) z] = \lambda \underbrace{c^T y}_{\leq \gamma} + (1-\lambda) \underbrace{c^T z}_{\leq \gamma}$

$\Rightarrow c^T y = c^T z = \gamma$ ↯

c) $\Rightarrow$ d): Ass.: $\exists v \neq 0 : A_{\operatorname{eq}(\{x\})} v = 0$

$\Rightarrow A(x + tv) \leq b$ for small $|t|$

$\Rightarrow x = \frac{1}{2}\underbrace{(x+tv)}_{=:y \in P} + \frac{1}{2}\underbrace{(x-tv)}_{=:z \in P}$ ↯.

d) $\Rightarrow$ b): $\dim(\operatorname{eq}(\{x\})) = \{y \in P : A_{\operatorname{eq}(\{x\})} y = b_{\operatorname{eq}(\{x\})}\}$

$= \{x\}$, $\dim\{x\} = 0$. □

3.19 Def (Support of a vector): Let $x \in \mathbb{R}^n$

$\operatorname{supp}(x) = \{j \in \{1,\dots,n\} : x_j \neq 0\}$ support of x

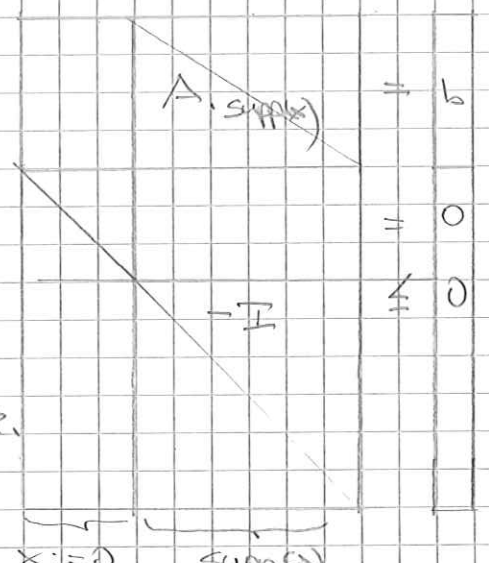
3.20 Thm (Vertices of Polyhedra $P=(A,b)$): Let $P=P(A,b)\subseteq\mathbb{R}^n$ be a polyhedron, $x\in P$. The following statements are equivalent:

a) x is a vertex of $P(A,b)$

b) $\operatorname{rank} A_{\cdot \operatorname{supp}(x)} = |\operatorname{supp}(x)|$

c) $\{A_{\cdot j}\}_{j\in\operatorname{supp}(x)}$ is lin. independent.

Proof: ex. □



3.21 Def (Extreme Lines and Rays):

Let $P\subseteq\mathbb{R}^n$ be a polyhedron, $F\subseteq P$ a face.

a) F edge of P :$\Leftrightarrow \dim F = 1$.

b) $\dim F=1$, $\exists z\in\mathbb{R}^n\setminus\{0\}\ \forall x\in P$

$F = \{x\} + \underbrace{\operatorname{lin}(\{z\})}_{=\langle z\rangle} = \{x+\alpha z : \alpha\in\mathbb{R}\}$ F extreme line of P

$F = \{x\} + \operatorname{cone}(\{z\}) = \{x+\alpha z : \alpha\in\mathbb{R}_{\geq 0}\}$ F extreme ray of P

c) x,y vertices of P, $x,y\in F$, F edge of P :$\Leftrightarrow$ x,y adjacent

3.22 Def (Inner Point of a Polyhedron): Let $P(A,b)\subseteq\mathbb{R}^n$ be a polyhedron, $x\in P(A,b)$.

x inner point of P :$\Leftrightarrow x\notin F$ $\forall$ proper face $F\subseteq P$.

Rem.: Inner points are inner points w.r.t. the relative topology of P.

3.23 Thm (Dimension of Polyhedra, Dimension Formula):

Let $A\in\mathbb{R}^{m\times n}$, $M=[m]$, $N=[n]$, $P=P(A,b)$ then:

a) $\exists\, x\in P$: $A_{\operatorname{eq}(P),\cdot}x = b_{\operatorname{eq}(P)}$, $A_{M\setminus\operatorname{eq}(P),\cdot}x < b_{M\setminus\operatorname{eq}(P)}$

b) x as in a) is an inner point of P

c) $P(A,b)\neq\emptyset \Rightarrow \exists$ inner point $x\in P$

d) $\dim P = n - \operatorname{rank} A_{\operatorname{eq}(P),\cdot} = \dim \ker A_{\operatorname{eq}(P),\cdot}$.

Proof: Let $I := \operatorname{eq}(P)$, $J := M\setminus I$

$\Rightarrow$ $Ax\leq b \Leftrightarrow A_{I,\cdot}x = b_I$, $A_{J,\cdot}x\leq b_J$

a) $\forall j\in J$: $\exists x^j\in P$: $A_{j,\cdot}x^j < b_j$.

$\Rightarrow A_I \cdot \left(\sum_{j\in J} \bar{x}_j\right) = b_I \;,\; A_J \cdot \left(\sum_{j\in J} \bar{x}_j\right) < b_J$

If $J = \emptyset$: choose any $x \in F$.

b) Let $F = \{ c^T x = \beta \} \cap P$ be a face of P.

We show: $\bar{x} \in F \Rightarrow F = P$.

$A_I \cdot \bar{x} = b_I \;,\; A_J \cdot \bar{x} < b_J$

$\Rightarrow \exists \varepsilon > 0 : A_I \cdot (\bar{x}+y) = b_I \;,\; A_J (\bar{x}+y) < b_J \quad \forall y \in \ker A_I \cap B_\varepsilon(0)$

$\Rightarrow \bar{x}+y \in P \quad \forall y \in \ker A_I \cap B_\varepsilon(0)$

$\Rightarrow c^T(\bar{x}+y) \le \beta \quad \forall y \in \ker A_I \cap B_\varepsilon(0)$

$\Rightarrow c^T y \le 0 \quad \forall y \in \ker A_I$

$\Rightarrow c^T y = 0 \quad \forall y \in \ker A_I$

$\Rightarrow P \subseteq \{ x : A_I x = b_I \} = \{ \bar{x}+y : y \in \ker A_I \} \subseteq \{ x : c^T x = \beta \}$

$\Rightarrow P \subseteq F$

c) $\Leftarrow$ a), b).

d) $\{ \bar{x}+y : y \in \ker A_I \cap B_\varepsilon(0) \} \subseteq F \subseteq \{ \bar{x}+y : y \in \ker A_I \}$

$\Rightarrow \dim \ker A_I \le \dim F \le \dim \ker A_I$ □

3.24 Thm (Dimension of Face): Let $F \subseteq P(A,b)$ be a face of a polyhedron P. Then

$$\dim F = \operatorname{rank} A_{eq(F)}.$$

Proof: Let $A \in \mathbb{R}^{m\times n}$, $M = [m]$, $I = eq(F)$, $J = M \setminus I$, $\bar{x} \in \operatorname{relint} F$ (3.25.a)

F induced by $c^T x \le \beta$. Then:

$$F = \{ c^T x = \beta,\; Ax \le b \}$$

$$= \{ c^T x = \beta,\; A_I x = b_I,\; A_J x < b_J \}$$

Thm 3.23

$\Rightarrow \exists \bar{x} : c^T \bar{x} = \beta,\; A_I \bar{x} = b_I,\; A_J \bar{x} < b_J$ and

$$\dim F = \dim \ker \begin{pmatrix} A_I \\ c^T \end{pmatrix}.$$

We show: $\dim \ker \begin{pmatrix} A_I \\ c^T \end{pmatrix} = \dim \ker A_I$.

"$\le$": $\exists y \in \ker A_I : c^T y \neq 0$, w.l.o.g. $c^T y > 0$. Then

$\exists \varepsilon > 0 : A_I(\bar{x}+\varepsilon y) = b_I \;,\; A_J(\bar{x}+\varepsilon y) < b_J \;,\; c^T(\bar{x}+\varepsilon y) > \beta$

$\rightarrow \bar{x}+\varepsilon y \in P(A,b)$, but $c^T x \le \beta$ not valid ↯ □

3.25 (a) (Dimension of Polyhedra and Faces): Let $\emptyset \subsetneq_{\neq} P = P(A,b) \subseteq \mathbb{R}^n$

a) $\text{eq}(P) = \emptyset \Rightarrow \dim P = n$

b) $F \subsetneq P$ face $\Rightarrow F = f_A(\text{eq}(F)) = \{x \in P : A_{\text{eq}(F)} x = b_{\text{eq}(F)}\}$.

c) $F \subsetneq P$ proper face $\Rightarrow \dim F \leq \dim P - 1$.

Proof: Thm 3.24. ☐

3.26 Thm (Interior Points): Let $F \subseteq P(A,b)$ be a face of $P(A,b)$, $\bar{x} \in F$. Then:

$\bar{x}$ interior point of $F \Leftrightarrow \text{eq}(\{\bar{x}\}) = \text{eq}(F)$.

Proof: Let G be the minimal face of F containing $\bar{x}$. Then:

$\bar{x}$ int. point of $F \overset{\text{3.25 a}}{\Leftrightarrow} F = G \Leftrightarrow f_A(\text{eq}(F)) = f_A(\text{eq}(\{\bar{x}\}))$

$\Leftrightarrow \text{eq}(F) = \text{eq}(\{\bar{x}\})$. ☐

3.27 Thm (Affine Hull of a Face): Let $\emptyset \neq F \subseteq P(A,b) = P$ be a face of P. Then:

a) $F = f_A(\text{eq}(F)) = \{x \in P : A_{\text{eq}(F)} x = b_{\text{eq}(F)}\}$.

b) $\text{aff } F = \{x \in \mathbb{R}^n : A_{\text{eq}(F)} x = b_{\text{eq}(F)}\}$.

Proof:

a) Kor. 3.25 b.

b) Let $I = \text{eq}(F)$, $T = \{x : A_I x = b_I\} = \text{aff } T$.

$F \subseteq T \Rightarrow \text{aff } F \subseteq \text{aff } T = T \overset{\dim =}{\Rightarrow} \text{aff } F = T$. ☐

$= \dim \text{aff } F$

$= \dim T$

$= \text{rk } A_I$.

4. Linear Optimization

4.1 Rem (Motivation): Consider a linear program

Primal LP: $\min c^T x$, $Ax \geq b$

Weak Bound: $Ax \geq b \Rightarrow \forall y \in \mathbb{R}^m_{\geq 0} : y^T A x \geq y^T b \Rightarrow c^T x \geq y^T b$ if $y^T A = c^T$

Dual LP: $\max y^T b$, $y^T A = c^T$, $y \geq 0$.

4.2 Thm (Weak Duality): Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$

(P) $\min c^T x$, $Ax \geq b$ (D) $\max y^T b$, $A^T y = c$, $y \geq 0$

Primal LP Dual LP

If x is feasible for P, and y for D, then

$$c^T x \geq b^T y$$

i.e. $\{b^T y\}$ is $\{$a lower$\}$ bound for the optimal objective value of $\{c^T x\}$ $\{$an upper$\}$.

Proof: $c^T x = (A^T y + s)^T x = y^T A x + s^T x \geq y^T b$. $\square$
$\geq 0 \quad \geq 0$

4.3 Cor (Optimality Criterion for LPs): given the prerequisites of Thm 4.2, then:

If x is feasible for (P), and y for (D), and $c^T x = b^T y$, then x is optimal for (P), and y for (D).

4.4 Cor. (Dual LPs): The following LPs are dual:

(P) $\min c^T x + a^T y$ s.t. $Ax + Dy \geq b$ (u), $Cx + Dy = d$ (v), $x \geq 0$

(D) $\max b^T u + d^T v$ s.t. $A^T u + C^T v \leq c$ (x), $D^T u + D^T v = a$ (y), $u \geq 0$

Mnemonics:

P	$\longleftrightarrow$	D
min	$\longleftrightarrow$	max
$\geq$	$\longleftrightarrow$	$\leq$
ineq.	$\longleftrightarrow$	sign restricted var.
equation	$\longleftrightarrow$	free variable
sign restr. var	$\longleftrightarrow$	$\geq$
free var.	$\longleftrightarrow$	=

Proof: Transform to standard form and verify. $\square$

4.6 Def (Dual LPs in standard form): Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.

(P) $\min c^T x$ s.t. $Ax \geq b$, $x \geq 0$

(D) $\max b^T y$ s.t. $A^T y \leq c$, $y \geq 0$ $\Leftrightarrow$ $\max y^T b$ s.t. $y^T A \leq c^T$, $y \geq 0$

$\left\{\begin{matrix}(P)\\(D)\end{matrix}\right\}$ is the $\left\{\begin{matrix}\text{primal}\\\text{dual}\end{matrix}\right\}$ LP in standard form.

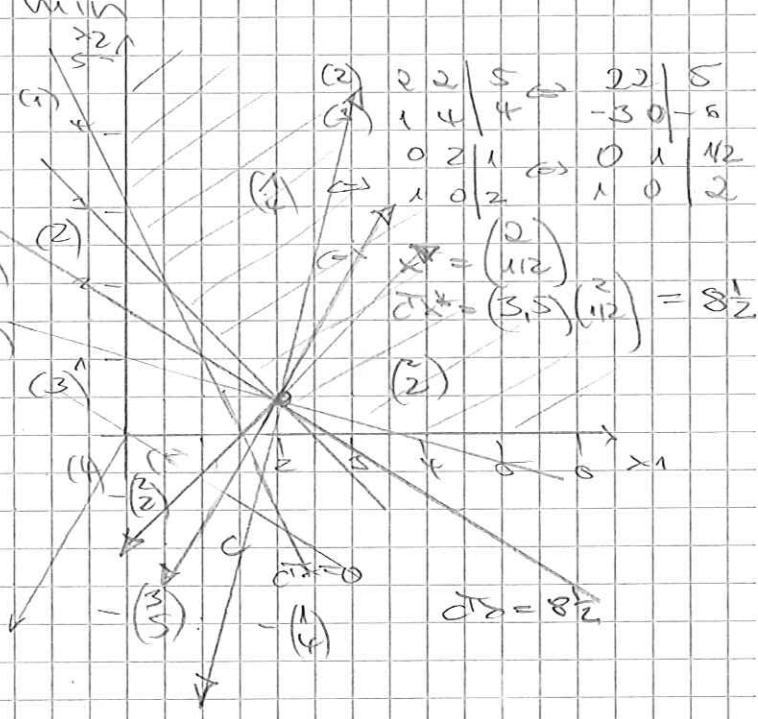
4.5 Cor (Dual LPs): The primal LP is the dual of the dual LP.

Proof: $\max b^T y$ s.t. $A^T y \leq c$, $y \geq 0$ $\Leftrightarrow \min -b^T y$ s.t. $-A^T y \geq -c$, $y \geq 0$ — the dual

$\max -c^T x$ s.t. $-Ax \leq -b$, $x \geq 0$ $\Leftrightarrow \min c^T x$ s.t. $Ax \geq b$, $x \geq 0$. $\square$

4.7 Ex (Oil Refinery): Crude processes p_1, p_2 et levels x_1, x_2

output / prices	p_1	p_2	demand b	price y
heavy oil	2	1	3	y_1
medium fuel oil	2	2	5	y_2
light fuel oil	1	4	4	y_3
cost c	3	5		max
activities	x_1	x_2		min



(P) $\min \; 3x_1 + 5x_2$

(y1) s.t. $2x_1 + x_2 \geq 3$ (1)

(y2) $2x_1 + 2x_2 \geq 5$ (2)

(y3) $x_1 + 4x_2 \geq 4$ (3)

(4) (5)

$x_1, x_2 \geq 0$

$$y_2^* \begin{pmatrix} 2 \\ 2 \end{pmatrix} + y_3^* \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix} = c$$

$$\Leftrightarrow \left(\begin{array}{cc|c} 2 & 1 & 3 \\ 2 & 4 & 5 \end{array}\right) \Leftrightarrow \left(\begin{array}{cc|c} 2 & 1 & 3 \\ 0 & 3 & 2 \end{array}\right)$$

$$\Leftrightarrow \left(\begin{array}{cc|c} 2 & 0 & 7/3 \\ 0 & 1 & 2/3 \end{array}\right) \Leftrightarrow \left(\begin{array}{cc|c} 1 & 0 & 7/6 \\ 0 & 1 & 2/3 \end{array}\right) \Rightarrow y^* = \left(0, \frac{7}{6}, \frac{2}{3}\right)$$

Note: $y_1^* = 0$, i.e., we are not willing to pay for heavy oil (surplus of 1½ units)

$0 \cdot (2x_1 + x_2 \geq 3)$

$+ 7/6 \cdot (2x_1 + 2x_2 \geq 5)$

$+ 2/3 \cdot (x_1 + 4x_2 \geq 4)$

$$\left(\frac{7}{3} + \frac{2}{3}\right)x_1 + \left(\frac{7}{3} + \frac{8}{3}\right)x_2 \geq \frac{35}{6} + \frac{8}{3}$$

$$\Leftrightarrow \quad 3x_1 + 5x_2 \geq \frac{51}{6} = 8\frac{3}{6} = 8\frac{1}{2}.$$

(D) $\max \; 3y_1 + 5y_2 + 4y_3$ profit max

$2y_1 + 2y_2 + y_3 \leq 3$

$y_1 + 2y_2 + 4y_3 \leq 5$ } competitiveness

$y_1, y_2, y_3 \geq 0$

non-neg. prices for heavy oil, med. fuel oil, light fuel oil

Economic interpretation: Oil supplier, prices y_1, y_2, y_3

process/product	hourly oil	wild fuel oil	refined fuel oil	max price
p_1	2	2	1	3
p_2	1	2	4	5
demand	3	5	4	
price	y_1	y_2	y_3	

y_1, y_2, y_3 are called shadow prices of the resources assoc. with the constraints.

4.8 Thm (1. Nec. and Suff. Optimality Condition for LPS)

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$. Then:

$\bar{x}$ opt. for $\min c^T x$, $Ax \geq b$ $\iff$ $c^T s \geq 0$ $\forall s \in T(-A_{eq(\bar{x})}, 0) =: C(\bar{x})$, $A\bar{x} \geq b$



Proof $s \in C(\bar{x}) \iff -A_{eq(\bar{x})} s \leq 0$

$\iff A_{eq(\bar{x})} \cdot s \geq 0$

$A_{eq(\bar{x})} \cdot \bar{x} = b_{eq(\bar{x})}$

$A_i \cdot \bar{x} > b_i \quad \forall i \notin eq(\bar{x})$

$\iff \exists \lambda > 0 : A(\bar{x} + \lambda s) \geq b$, i.e.

$C(\bar{x})$ is the cone of feasible directions, in which one can move starting from $\bar{x}$.

"$\Rightarrow$": $\bar{x}$ opt, $s \in C(\bar{x}) \to \exists \lambda > 0 : x = \bar{x} + \lambda s$ feas.

$\Rightarrow c^T x = c^T \bar{x} + \lambda c^T s \geq c^T \bar{x} \to c^T s \geq 0$.

"$\Leftarrow$": x feas., $s := x - \bar{x} \to A_{eq(\bar{x})} \cdot x \geq b_{eq(\bar{x})}$

$- A_{eq(\bar{x})} \cdot \bar{x} = b_{eq(\bar{x})}$

$A_{eq(\bar{x})} s \geq 0$

$c^T s \geq 0 \iff c^T (x - \bar{x}) \geq 0 \iff c^T x \geq c^T \bar{x}$. □

4.9 Rem (Feasibility of Systems of Lin. Ineq.):

$$\max\ c^T x \ (=0) \quad \Leftrightarrow \quad \min\ b^T y \ (\geq 0) \qquad (y=0)$$
$$\text{s.t. } Ax = b \qquad\qquad A^T y \geq 0$$
$$x \geq 0$$

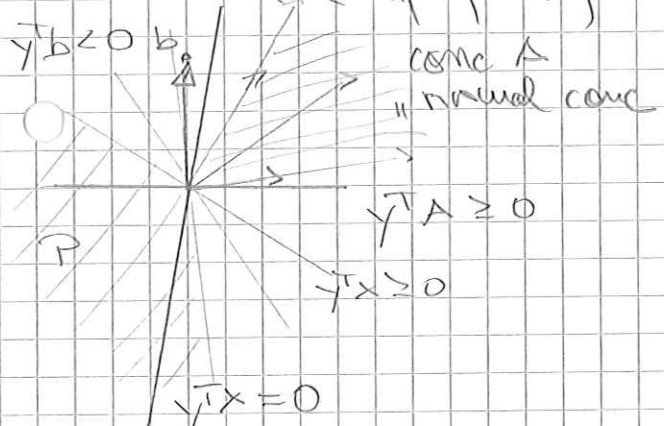
$\Leftrightarrow$ $Ax = b$, $x \geq 0$, feasible $\quad \dot{\vee}$ $b^T y < 0$, $A^T y \geq 0$, feasible

alternative Theorem: „Farkas Lemma"

geometric interpretation:

$b \in \operatorname{cone} A$ $\quad \dot{\vee}$ $\exists$ hyperplane $y^T x = 0$, through 0, that separates b from $\operatorname{cone} A$.

$= \operatorname{cone}(A_{\cdot 1}, \dots, A_{\cdot n})$



4.10 Thm (Farkas-Lemma, Farkas [1902]): Let $A \in \mathbb{K}^{m \times n}$, $b \in \mathbb{K}^m$. Then exactly one of the following systems has a solution:

(1) $Ax = b$, $x \geq 0$ $\quad \dot{\vee}$ $\quad$ (2) $y^T b < 0$, $y^T A \geq 0$.

Proof:

„$\circ$": Supp. (1) & (2): $0 = 0^T x \leq y^T A x = y^T b < 0$ ↯ ($x \geq 0$)

„$\vee$": (1) feas. $\Leftrightarrow b \in \operatorname{cone} A = \{Ax : x \geq 0\} =: K$

K polyhedral cone, i.e., in particular convex and closed (as a finite intersection of closed sets).

Case 1: $b \in \operatorname{cone} A \Rightarrow$ (1)

Case 2: $b \notin \operatorname{cone} A \overset{\text{strong separation Theorem}}{\Rightarrow} \exists$ hyperplane $H^=(y, \delta)$ sep. b, K

$\Rightarrow \exists H^=(y,\delta)^{2.9}: y^T z \geq \delta \ \forall z \in K$, $y^T z < \delta$

$\Rightarrow \exists (y, \delta): y^T A \geq \delta$, $y^T z < \delta$

$0 \in K \Rightarrow y^T 0 = 0 \geq \delta$

Claim: $y^T A \geq 0$.

Pf: Supp $z \in K$: $y^T z < 0 \Rightarrow y^T (\lambda z) \xrightarrow[\lambda \to \infty]{} -\infty \geq 0$ ↯ □

4.11 Cor (Farkas Lemma, Alternative Form I): Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Then exactly one of following systems has a solution:

$$Ax \leq b \quad \text{or} \quad y^T A = 0,\ y \geq 0,\ y^T b < 0$$

Proof: ⇔. □

4.12 Cor (Farkas Lemma, Alternative Form II): Let A, B, C, D and a, b be compatible matrices and vectors. Then exactly one of the following systems has a solution:

$$Ax + By \leq a,\quad Cx + Dy = b,\quad x \geq 0$$

or

$$u^T A + v^T C \geq 0,\quad u^T B + v^T D = b,\quad u \geq 0,\quad u^T a + v^T b < 0$$

Proof: ⇔. □

4.13 Thm (Complementarity): Consider

$$(P) \quad \min_{Ax \geq b} c^T x \qquad (D) \quad \max_{y^T A = c^T,\ y \geq 0} y^T b$$

a) $c^T x \geq y^T b$ $\forall x, y$ (resp. for (P), (D)) weak duality

b) $c^T x = y^T b$ for x, y ⟺ x, y opt strong duality

c) $c^T x \geq y^T b \Leftrightarrow y^T \underbrace{(Ax - b)}_{=: s} \geq 0$

$y^T (Ax - b) = 0 \Leftrightarrow y^T s = 0 \Leftrightarrow y_i = 0 \vee s_i = 0$ $\forall i$ complementarity

($y \geq 0$, $s \geq 0$) dual slack primal slack

"complementary slackness"

4.14 Thm (Necessary and sufficient KKT optimality condition for LPs): Consider the following pair of dual LPs, where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$:

(P) $\min c^T x$, $Ax \geq b$ (D) $\max y^T b$, $y^T A = c^T$, $y \geq 0$.

Then $\begin{bmatrix} x \\ y \end{bmatrix}$ is optimal for $\begin{bmatrix} (P) \\ (D) \end{bmatrix}$ if and only if there exists $\begin{pmatrix} x \\ y \end{pmatrix}$ s.t. the following conditions hold:

(KKT) $Ax \geq b$ — primal feas
$y^T A = c^T$, $y \geq 0$ — dual feas
$y^T(b - Ax) = 0$ — complementary slackness

Proof: We show the statement for x, then y follows by duality.

"$\Rightarrow$": Let x be opt. for (P), $I := eq(x)$ the set of act. ineqs.

x opt. for (P) $\overset{\text{Thm 4.8}}{\Leftrightarrow}$ $c^T s < 0$, $A_{I\cdot} s \geq 0$ infeasible

$\Leftrightarrow$ $s^T c < 0$, $s^T A_{I\cdot}^T \geq 0$ infeasible

$\overset{\text{Farkas}}{\Leftrightarrow}$ $A_{I\cdot}^T z = c$, $z \geq 0$ feasible

$\overset{\text{Lemma 4.10}}{\Leftrightarrow}$ $z^T A_{I\cdot} = c^T$, $z \geq 0$

Set $y^T = (y_I^T, y_{[m]\setminus I}^T) = (z^T, 0^T)$, then

$Ax \geq b$

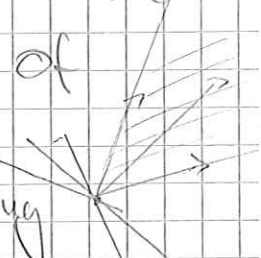
$y^T A = z^T A_{I\cdot} = c^T$, $y \geq 0$ $\Leftrightarrow$ $Ax \geq b$, $y^T A = c^T$, $y \geq 0$

$x_i x - b_i = 0$ $\forall i \in I$, $y_i = 0$ $\forall i \notin I$ $y^T(Ax - b) = 0$.

"$\Leftarrow$": KKT $\Rightarrow$ $\begin{bmatrix} x \\ y \end{bmatrix}$ primal dual feasible.

$c^T x - y^T b = y^T Ax - y^T b = y^T(Ax - b) = 0 \Rightarrow x$ opt. □

4.14 Rem (Complementarity): Thm 4.14 guarantees, that the objective can be combined as a conic combination of the active constraints of the optimal solution.



4.15 Thm (Strong Duality Theorem): The following statements are equivalent for a pair of dual LPs (P) $\min c^T x$, $Ax \geq b$ and (D) $\max y^T b$, $y^T A = c^T$, $y \geq 0$

a) (P) and (D) have feasible solutions.

b) (P) has a finite optimum.

c) (P) and (D) have optimal solutions with the same objective.

d) (P) has a finite optimum.

e) (P) or (D) has a finite optimum.

f) (P) has feasible solutions and its objective is bounded from below

or (D) " " above.

Proof:

a) $\Rightarrow$ b): $c^T x \geq y^T b \quad \forall x, y$ feas for (P), (D) $\Rightarrow$ (P) bounded

Let $p^* = \inf \{ c^T x \mid Ax \geq b \}$.

Claim: $\exists x : Ax \geq b,\ c^T x = p^*$.

Arg.: Supp not. Then

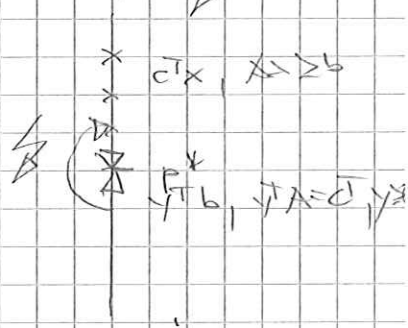
$$\begin{array}{ll} -Ax \leq -b & (y) \\ c^T x \leq p^* & (\mu) \end{array} \quad \text{inf.} \overset{\text{Farkas Lemma}}{\Leftrightarrow} \quad \begin{array}{l} -y^T A + \mu c^T = 0 \quad (1) \\ y, \mu \geq 0 \quad (2) \text{ feas.} \\ -y^T b + \mu p^* < 0 \quad (3) \end{array}$$

$(1) \cdot x$

$$\Rightarrow 0 = -y^T Ax + \mu c^T x \overset{(3)}{<} -y^T b + \mu c^T x < -\mu p^* + \mu c^T x$$

$\mu = 0 \Rightarrow -y^T A = 0$, $y \geq 0$, $-y^T b < 0$ feas. $\rightarrow$ $Ax \geq b$ inf.

$\mu > 0 \Rightarrow (y^T/\mu) A = c^T$, $(y/\mu) \geq 0$, $(y^T/\mu) b > p^*$ ↯



b) $\Rightarrow$ c): KKT conditions.

c) $\Rightarrow$ d) $\Rightarrow$ e) $\Rightarrow$ f): trivial.

f) $\Rightarrow$ a): Case i): (P) feas. bounded $\Rightarrow$ (D) feas.

Supp not: " " , (P) inf.

$\rightsquigarrow$ $y^T A = c^T$, $y \geq 0$ inf. $\overset{\text{Farkas Lemma}}{\Leftrightarrow}$ $As \geq 0$, $c^T s < 0$ feas.

$\Rightarrow$ $x + \lambda s$ feas for (P) $\forall \lambda > 0$ and

$c^T (x + \lambda s) = c^T x + \lambda \underbrace{c^T s}_{<0} \to -\infty$ ↯ $\lambda \to \infty$

Case ii) is analogous. □

4.16 Thm (Duality for unbounded and infeasible LPs): Consider the following pair of dual LPs:

(P) $\min\ c^T x$ s.t. $Ax \geq b$

(D) $\max\ y^T b$ s.t. $y^T A = c^T$, $y \geq 0$

dual feasible sets: $P = \{x : Ax \geq b\}$ and $D = \{y : y^T A = c^T, y \geq 0\}$, and let

$$\inf(P) := \begin{cases} +\infty, & P = \emptyset \\ -\infty, & \text{if } (P) \text{ is unbounded} \\ \min c^T x \text{ s.t. } Ax \geq b, & \text{else} \end{cases}, \quad \sup(D) := \begin{cases} -\infty, & \text{if } D = \emptyset \\ +\infty, & \text{if } (D) \text{ is unb.} \\ \max y^T b \text{ s.t. } y^T A = c^T, y \geq 0, & \text{else} \end{cases}$$

Then:

a) $\inf(P) = -\infty \Rightarrow D = \emptyset$

b) $\sup(D) = +\infty \Rightarrow P = \emptyset$

c) $P = \emptyset \Rightarrow \sup(D) = +\infty \vee D = \emptyset$

d) $D = \emptyset \Rightarrow \inf(P) = +\infty \vee P = \emptyset$

Proof:

a) $\sup D \neq \emptyset \Leftrightarrow \exists y \in D : -\infty < y^T b \leq c^T x \ \forall x \in P$ ↯.

b) Analogous to a)

c) $\sup(D) < +\infty \wedge D \neq \emptyset \overset{\text{Thm 4.15 f)}}{\Rightarrow} P \neq \emptyset$

d) Analogous. □

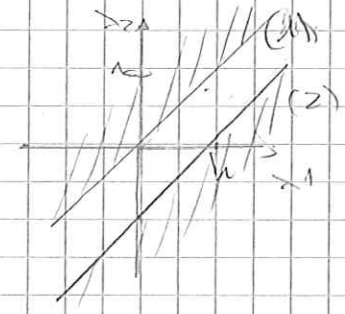
It is true: Thm 4.18 also holds for LPs in general form.

4.20 Ex (Duality for ungl. LPs):

a) (P) $\min\ -x_1 - x_2$ s.t.

$-x_1 + x_2 \geq 0$ (1)

$x_1 - x_2 \geq 1$ (2)

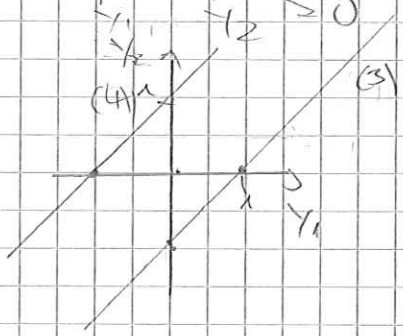


(D) $\max\ 0y_1 + y_2$ s.t.

$-y_1 + y_2 = -1$ (3)

$y_1 - y_2 = -1$ (4)

$y_1, y_2 \geq 0$



i) (P) and (D) are dual

ii) (1) + (2) $\Leftrightarrow 0 \geq 1 \Rightarrow P = \emptyset$

iii) (3) + (4) $\Leftrightarrow 0 = -2 \Rightarrow D = \emptyset$

The case $P = D = \emptyset$ is possible.

b) (P) $\min\ -x_1 - x_2$
$-x_1 + x_2 \geq 0$ (1)

(D) $\max\ 0 y_1$
$-y_1 = -1$ (3)
$y_1 = -1$ (4)
$y \geq 0$

i) (P) and (D) are dual

ii) $\lambda \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in P \ \forall \lambda \to c^T \lambda \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = -2\lambda \xrightarrow{\lambda \to \infty} -\infty$

iii) (3) + (4) $\leadsto 0 = -2 \Rightarrow D = \emptyset$.

4.21 Thm (Weak Complementary Slackness Theorem): Let $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$

(P) $\min\ c^T x$
$Ax \geq b$

(D) $\max\ y^T b$
$y^T A = c^T$
$y \geq 0$

be a pair of dual LPs with solutions $\bar{x}, \bar{y}$. The following statements are equivalent:

a) $\bar{x}$ opt. for (P), $\bar{y}$ opt. for (D)

b) $\bar{y}^T (A\bar{x} - b) = 0$

c) $\bar{y}_i = 0 \ \vee \ A_{i\cdot} x - b_i = 0, \quad i = 1, \dots, m$

Proof:

b) $\leadsto$ c) ✓

b) $\Rightarrow$ a): $\bar{y}^T (A\bar{x} - b) = 0 \leadsto c^T \bar{x} = \bar{y}^T b$ $\overset{\text{weak duality}}{\Rightarrow}$ a)

a) $\Rightarrow$ b): b) $\overset{\text{def}}{\Rightarrow}$ a). ☐

4.22 Rem (Weak Complementary Slackness Thm): In Thm 4.21 c),
it is possible that

$$\bar{y}_i = 0 \ \wedge \ A_{i\cdot} x = b_i$$

holds for some i. We will show (later): $\exists \bar{x}, \bar{y}$ feasible:

$$\bar{y}_i = 0 \ \vee \ A_{i\cdot} \bar{x} = b_i.$$

4.23 Thm (Weak Complementary Slackness (in general LPs))

For a pair of dual LPs

(P) $\min\ c^T x + d^T y$
$Ax + By \geq a$
$Cx + Dy = b$
$x \geq 0$

(D) $\max\ u^T a + v^T b$
$u^T A + v^T C \leq c^T$
$u^T B + v^T D = d^T$
$u \geq 0$

and feasible solutions $\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}$, $\begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix}$ the following statements are equiv.:

a) $\begin{pmatrix} \bar{x} \\ \bar{r} \end{pmatrix}$ is opt. for (P), $\begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix}$ is opt. for (D)

b) $(\bar{c}^T - \bar{u}^T A - \bar{v}^T C)\bar{x} + \bar{u}^T(A\bar{x} + B\bar{y} - a) = 0$

c) $\bar{u}_i = 0 \vee A_{i\cdot}\bar{x} + B_{i\cdot}\bar{y} = a_i$,
$\bar{x}_j = 0 \vee \bar{u}^T A_{\cdot j} + \bar{v}^T C_{\cdot j} = c_j$.

Proof:

$(P) \leftrightarrow (P')$ where $\begin{pmatrix} A & B \\ C & D \\ -C & -D \\ I & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \geq \begin{pmatrix} a \\ b \\ -b \\ 0 \end{pmatrix}$, $(D) \leftrightarrow (D')$ where $\max \begin{pmatrix} a \\ -b \\ 0 \end{pmatrix}^T \begin{pmatrix} u \\ v \\ w \end{pmatrix}$ s.t. $(u', u'', v'', w')\begin{pmatrix} A & B \\ C & D \\ -C & -D \\ I & 0 \end{pmatrix} = (c', d')$, $u, u', v', v'', w \geq 0$

$\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}, \begin{pmatrix} \bar{u} \\ \bar{v} \\ \bar{w} \end{pmatrix}$ opt. for (P'), (D')

Thm. 4.21 $\Rightarrow$ $(\bar{u}', \bar{u}'', \bar{v}'', \bar{w}^T)\left[\begin{pmatrix} A & B \\ C & D \\ -C & -D \\ I & 0 \end{pmatrix}\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} - \begin{pmatrix} a \\ b \\ -b \\ 0 \end{pmatrix}\right] = 0$

$\leftrightarrow \bar{u}^T(A\bar{x} + B\bar{y} - a) + \underbrace{(\bar{v}' - \bar{v}'')}_{=:\bar{v}}\underbrace{(C\bar{x} + D\bar{y} - b)}_{=0} + \bar{w}^T(I\bar{x} + 0) = 0$, $\bar{w}^T = c' - \bar{u}^T A - \bar{v}^T C$

$\Rightarrow \bar{u}^T(A\bar{x} + B\bar{y} - a) + (c^T - \bar{u}^T A - \bar{v}^T C)\bar{x} = 0$ ☐

4.24 Thm (Strong Complementary Slackness Thm): Consider the pair of dual LPs

(P) $\min\ c^T x$ s.t. $Ax \geq b$ $\qquad$ (D) $\max\ y^T b$ s.t. $y^T A = c^T$, $y \geq 0$

If (P), (D) have feasible solutions, then there exist opt. solutions $\bar{x}, \bar{y}$ s.t.

$\bar{y}_i > 0 \leftrightarrow A_{i\cdot}\bar{x} = b_i$
$\bar{y}_i = 0 \leftrightarrow A_{i\cdot}\bar{x} > b_i$
$\Big\}$ $i = 1, \dots, m$

Proof: Set

$N := \{i \in [m] : \exists$ opt. sol. $\bar{x}$ of (P): $A_{i\cdot}\bar{x} > b_i\}$

$B := \{i \in [m] : \exists$ opt. sol. $\bar{y}$ of (D): $y_i > 0\}$.

$\bar{x} := \sum_{i \in N} \frac{1}{|N|} \bar{x}^i \Rightarrow \{ A_N \cdot \bar{x} > b_N \mid \bar{y}_N = 0 \} \Rightarrow \exists N = \emptyset$

$\bar{y} := \sum_{i \in \bar{I}} \frac{1}{|\bar{I}|} \bar{y}^i \Rightarrow \{ \bar{y}_{\bar{I}} > 0 \mid A_{\bar{I}} \cdot \bar{x} = b_{\bar{I}} \}$

Claim: $\exists \cup N = [m]$.

Ass. $\exists \neq I := [m] \setminus (\exists \cup N) \neq \emptyset, \bar{y}_I = A_I \cdot \bar{x} - b_I = 0$.

For $i \in I$, consider

$$A_{\exists} \cdot s = 0 \mid A_{=\{i\}} \cdot s \geq 0 \mid A_i \cdot s > 0. \quad (*)$$

(i) feas.: Let $x' := \bar{x} + \lambda s$ then for small λ

$Ax' \geq b \Rightarrow x'$ feas.
$\bar{y}^T (Ax' - b) = 0 \Rightarrow x'$ opt.
$\Rightarrow i \in N$ ↯. (standard form)

(ii) inf.: $A_{\exists} s = 0 \mid A_{=\{i\}} \cdot s \geq 0 \mid -A_i \cdot s < 0$ inf.

Farkas $\Rightarrow \exists t: t^T A_{\exists} + t^T A_{=\{i\}} = -x_i \mid t_i = 1$

$\Rightarrow \exists t: t^T A = 0 \mid t_{=\{i\}} \geq 0 \mid t_i > 0$

Let $y' := \bar{y} + \lambda t$ then for small λ

$y'^T A = c^T \mid y' \geq 0 \Rightarrow y'$ feas
$y'^T (A\bar{x} - b) = 0 \Rightarrow y'$ opt
$\Rightarrow i \in \exists$ ↯. □

4.25 Rem: The strong complementary slackness theorem also holds for LPs in general form.

5.2 Def (Basis): Consider the LP in standard form

(P) $\min c^T x$, $Ax = b$, $x \geq 0$ with dual (D) $\max y^T b$, $y^T A + z^T = c^T$, $z \geq 0$

Let $A \in \mathbb{R}^{m \times n}$ have full row rank, B an m-tuple of pairwise diff. indices from $[n]$, N an $(n-m)$-tuple of the remaining indices. We also use B, N as sets.

a) B basis (of P) $:\Leftrightarrow$ $A_{\cdot B}$ regular

$A_{\cdot B}$ basis (matrix), x_B basic variables

$A_{\cdot N}$ non-basis (matrix), x_N non basic variables.

b) $(\bar{x}_B, \bar{x}_N) := (A_{\cdot B}^{-1} b, 0)$ (primal)

$\bar{y}^T := c_B^T A_{\cdot B}^{-1}$ (dual)

$\bar{c}^T := c^T - \underbrace{c_B^T A_{\cdot B}^{-1}}_{=\bar{y}^T} A$ reduced costs

} basic solution w.r.t basis B

c) $(\bar{x}_B, \bar{x}_N)$ (primal) feasible $:\Leftrightarrow \bar{x}_B \geq 0 \Leftrightarrow$: B primal feas.

$\bar{y}$ (dual) feasible $:\Leftrightarrow \bar{c} \geq 0 \Leftrightarrow$: B dual feas.

d) $(\bar{x}_B, \bar{x}_N)$ non-degenerate $:\Leftrightarrow \bar{x}_B > 0$

$(\bar{x}_B, \bar{x}_N)$ degenerate $:\Leftrightarrow \exists j \in B : \bar{x}_j = 0$.

5.3 Thm (Vertices and Basic Solutions): Let $A \in \mathbb{R}^{m \times n}$, rank $A = m$, $b \in \mathbb{R}^m$, $x \in P^=(A,b)$. Then the following statements are equivalent:

a) x is a vertex of $P^=(A,b)$

b) x is a basic feasible sol. of (P)

Proof:

"$\Rightarrow$": Let $I = \operatorname{supp}(x)$ $\overset{\text{Thm. 3.20}}{\Rightarrow}$ $A_{\cdot I}$ lin. indep.

rank $A = m$ $\Rightarrow$ $\exists\, J \subseteq [n] \setminus I : A_{\cdot I \cup J}$ regular

$\Rightarrow$ $B := I \cup J$ basis, $N := [n] \setminus B$ non-basis

$\Rightarrow$ $x_B = A_{\cdot B}^{-1} b \geq 0$ ($x \in P^=(A,b)$), $x_N = 0$.

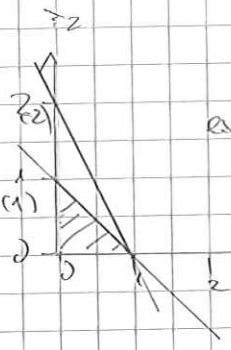
"$\Leftarrow$" Thm. 3.20 $\square$

5.4 Rem (Vertices and bases): Note:

$x \in X$ $\overset{1:1}{\longleftrightarrow}$ feas. basic sol $\overset{\text{not}}{\underset{1:1}{\longleftrightarrow}}$ feas. bases

5.5 Ex (Degeneration):

a) $x_1 + x_2 \leq 1 \quad (1)$
$2x_1 + x_2 \leq 2 \quad (2)$
$x_1, x_2 \geq 0$



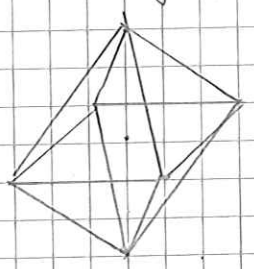
$\mathrm{ext}(P) = \left\{\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$

c) $x_1 + x_2 + x_3 \leq 1$
$2x_1 + x_2 + x_4 = 2$
$x_1, x_2, x_3, x_4 \geq 0$

B	$A_{\cdot B}$	b	$\begin{pmatrix}x_1\\x_2\end{pmatrix}$	N	$x_N = 0$
(1,2)	$\begin{pmatrix}1&1\\2&1\end{pmatrix}$	$\begin{pmatrix}1\\2\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$	(3,4)	$x_3 = 0$, $x_4 = 0$
(1,3)	$\begin{pmatrix}1&1\\2&0\end{pmatrix}$	$\begin{pmatrix}1\\2\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$	(2,4)	$x_2 = 0$, $x_4 = 0$
(1,4)	$\begin{pmatrix}1&0\\2&1\end{pmatrix}$	$\begin{pmatrix}1\\2\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$	(2,3)	$x_2 = 0$, $x_3 = 0$
(2,3)	$\begin{pmatrix}1&1\\1&0\end{pmatrix}$	$\begin{pmatrix}1\\2\end{pmatrix}$	$\begin{pmatrix}1\\-1\end{pmatrix}$	(1,4)	$x_1 = 0$, $x_4 = 0$
(2,4)	$\begin{pmatrix}1&0\\1&1\end{pmatrix}$	$\begin{pmatrix}1\\2\end{pmatrix}$	$\begin{pmatrix}1\\1\end{pmatrix}$	(1,3)	$x_1 = 0$, $x_3 = 0$
(3,4)	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\begin{pmatrix}1\\2\end{pmatrix}$	$\begin{pmatrix}1\\2\end{pmatrix}$	(1,2)	$x_1 = 0$, $x_2 = 0$

Reason: $\begin{pmatrix}1\\0\end{pmatrix}$ is defined as an intersection of three lines. Then, in $\mathbb{R}^2$, one lin. is redundant.

b) In $\mathbb{R}^3$ genuine degeneration exists:



all vertices are degenerate

5.6 Rem: $x_B = A_{\cdot B}^{-1} b \geq 0 \Rightarrow B$ is unique.

5.7 Ex: The converse of Rem. 5.6 is not true:

$$A = \begin{pmatrix}1&1&0\\0&0&1\end{pmatrix}, \quad b = \begin{pmatrix}1\\0\end{pmatrix}$$

$\mathcal{I}$	$A_{\cdot\mathcal{I}}$	$x_{\mathcal{I}}$	x
$(1,2)$	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\\0\end{pmatrix}$
$(2,3)$	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\begin{pmatrix}1\\0\end{pmatrix}$	$\begin{pmatrix}0\\1\\0\end{pmatrix}$

The basic solutions are degenerate, but the bases unique.

5.2 Alg (Simplex Algorithm, basic version):

Input: LP (P) with $c^T x$, $Ax=b$, $x \geq 0$ in standard form
∃ primal feasible basis $\mathcal{I}$, $\bar{x}_{\mathcal{I}} = A_{\cdot\mathcal{I}}^{-1} b$

Output: $\bar{x}$ = optimum (P) or "(P) unbounded"

1. Backward transformation (BTRAN):
   $\bar{y}^T \leftarrow A_{\cdot\mathcal{I}}^{-T} c_{\mathcal{I}}$ (dual sol.) — solve $\bar{y}^T A_{\cdot\mathcal{I}} = c_{\mathcal{I}}^T$
2. Pricing (PRICE):
   $\bar{c}_N^T \leftarrow c_N^T - \bar{y}^T A_{\cdot N}$ (red. costs)
   if $(\bar{c}_N^T \geq 0)$ then output $(\bar{x}_{\mathcal{I}}, 0)$, stop
   $j \leftarrow \bar{c}_j < 0$ (variable entering the basis)
3. Forward Transformation (FTRAN):
   $\bar{A}_{\cdot j} \leftarrow A_{\cdot\mathcal{I}}^{-1} A_{\cdot j}$ — solve $A_{\cdot\mathcal{I}} \bar{A}_{\cdot j} = A_{\cdot j}$
4. Ratio Test (CHUZR):
   if $(\bar{A}_{ij} \leq 0)$ then output "(P) unbounded", stop
   $i \leftarrow \underset{k \in 1,\dots,m}{\text{argmin}} \left\{ \frac{\bar{x}_{\mathcal{I}k}}{\bar{a}_{kj}} \mid \bar{a}_{kj} > 0 \right\}$ (var leaving the basis)
   $\theta \leftarrow \frac{\bar{x}_{\mathcal{I}i}}{\bar{a}_{ij}}$ (step size, $\theta \geq 0$, $\theta = 0$ possible)
   $\bar{a}_{ij}$ pivot (element) $(\bar{a}_{ij} > 0)$
5. Update (UPDATE):
   $\bar{x}_j \leftarrow \theta$
   $\bar{x}_{\mathcal{I}} \leftarrow \bar{x}_{\mathcal{I}} - \theta \bar{A}_{\cdot j}$
   $\mathcal{I} \leftarrow (\mathcal{I} \setminus \mathcal{I}_i) \cup \{j\}$ (replace i by j)
   $N \leftarrow \{1,\dots,n\} \setminus \mathcal{I}$ (in any order)
   goto 1

S.9 Thm (Correctness of Alg. S.8): When Alg. S.8 terminates, it yields a correct solution.

Proof: Let $\bar{a}_{ij}$ be a pivot, $\mathcal{B}' = (\mathcal{B} \setminus \{B_i\}) \cup \{j\}$, $\mathcal{N}' = \mathcal{N} \setminus \mathcal{B}'$.

Claim 1: $A_{\cdot \mathcal{B}'}$ is regular

Pf. Supp not $\Rightarrow$ $A_{\cdot j}$ dep lin. on $A_{\cdot \mathcal{B} \setminus \{B_i\}}$.

$$\Rightarrow \exists \lambda: \begin{cases} A_{\cdot j} = A_{\cdot \mathcal{B} \setminus \{B_i\}} \lambda \\ A_{\cdot j} = A_{\cdot \mathcal{B}} \bar{A}_{\cdot j} \end{cases} \Rightarrow \bar{A}_{\cdot j} = \begin{pmatrix} \lambda \\ 0 \end{pmatrix} \Rightarrow \bar{a}_{ij} = 0 \; \lightning$$

Claim 2: $\bar{x}'_{\mathcal{B}'} \geq 0$, $\bar{x}'_{\mathcal{N}'} = 0$

Pf.: $\bar{x}'_{\mathcal{B}'_k} = \bar{x}_{\mathcal{B}_k} - \mu \bar{a}_{kj} \geq 0$

$$\Leftrightarrow \begin{cases} \bar{a}_{kj} \leq 0 \\ \dfrac{\bar{x}_{\mathcal{B}_k}}{\bar{a}_{kj}} \geq \mu = \min\limits_{l=1}^{m} \left\{ \dfrac{\bar{x}_{\mathcal{B}_l}}{\bar{a}_{lj}}, \bar{a}_{lj} \geq 0 \right\} \end{cases}, \; \bar{a}_{kj} \geq 0$$

in particular $\dfrac{\bar{x}_{\mathcal{B}_i}}{\bar{a}_{ij}} = \mu$

$$\Rightarrow \bar{x}'_{\mathcal{B}_i} = \bar{x}_{\mathcal{B}_i} - \frac{\bar{x}_{\mathcal{B}_i}}{\bar{a}_{ij}} \bar{a}_{ij} = 0.$$

Claim 3: $\bar{x}'_{\mathcal{B}'} = A^{-1}_{\cdot \mathcal{B}'} b$

Pf.: $A_{\cdot \mathcal{B}'} \bar{x}'_{\mathcal{B}'} = A_{\cdot \mathcal{B} \setminus \{B_i\}} \bar{x}'_{\mathcal{B} \setminus \{B_i\}} + A_{\cdot j} \bar{x}'_j + A_{\cdot \mathcal{B}} \underbrace{\bar{x}'_{B_i}}_{=0}$

$$= A_{\cdot \mathcal{B}} \underbrace{\bar{x}'_{\mathcal{B}}}_{} + A_{\cdot j} \underbrace{\bar{x}'_j}_{}$$

$$= \bar{x}_{\mathcal{B}} - \mu \bar{A}_{\cdot j} \qquad = \mu$$

$$= A_{\mathcal{B}} \left(\bar{x}_{\mathcal{B}} - \mu \cancel{A^{-1}_{\mathcal{B}} A_{\cdot j}} \right) + \cancel{A_{\cdot j} \mu}$$

$$= A_{\cdot \mathcal{B}} \bar{x}_{\mathcal{B}}$$

$$= b.$$

Claim 4: Stop in PEICE $\Rightarrow$ $\bar{x}, \bar{y}$ opt.

Pf.: $\bar{c}^T = \left(\underbrace{c_{\mathcal{B}}^T - \bar{y}^T A_{\cdot \mathcal{B}}}_{=0}, \underbrace{c_{\mathcal{N}}^T - \bar{y}^T A_{\cdot \mathcal{N}}}_{=\bar{z} \geq 0} \right) \geq 0 \Rightarrow \bar{y}$ dual feas.

compl. slackness

$$\underbrace{\bar{c}_{\mathcal{B}}^T}_{=0} \bar{x}_{\mathcal{B}} + \bar{c}_{\mathcal{N}}^T \underbrace{\bar{x}_{\mathcal{N}}}_{=0} = \bar{z}_{\mathcal{B}}^T \bar{x}_{\mathcal{B}} + \bar{z}_{\mathcal{N}}^T \bar{x}_{\mathcal{N}} = 0 \Rightarrow \bar{x}, \bar{y} \text{ opt}$$

Claim 5: Step in CHUZR $\Rightarrow$ $\mathcal{P}$ unbounded

Pf: $\bar{x}'_B = \bar{x}_B - \mu \bar{A}_{\cdot j} \geq 0$

($\bar{x}_B \geq 0$, $\mu \geq 0$, $\bar{A}_{\cdot j} \leq 0$)

$\bar{x}'_j = \mu \qquad \geq 0$

$\bar{c}^T \bar{x}' = \bar{c}_B^T (\bar{x}_B - \mu \bar{A}_{\cdot j}) + c_j \mu$

$= \bar{c}_B^T \bar{A}_{\cdot B}^{-1} b + (c_j - \underbrace{\bar{c}_B^T \bar{A}_{\cdot B}^{-1}}_{= \bar{y}^T} A_{\cdot j}) \mu \xrightarrow{\mu \to \infty} -\infty$,

$= \bar{c}_j < 0$

§10 Thm (Finiteness in the Non-degenerate Case):

If $\mathcal{P}(A,b) \neq \emptyset$ and all basic solutions are non-degenerate, Alg. S. 8 terminates.

Proof: $c^T \bar{x}' = c_B^T \bar{x}_B + \bar{c}_j \mu < c^T \bar{x}$,

($\bar{c}_j < 0$: dual non-degenerate; $\mu > 0$: primal non-degenerate)

i.e., Alg § 8 produces a sequence of bases with strictly decreasing objective. As there is only a finite number of bases, the algorithm terminates. □

5.11 Thm (Anti-Degeneracy Pivot Rules): Alg. S. 8 terminates if one of the following pivot rules is used:

a) Bland Rule (Smallest Index Rule):

PRICE: Choose j: $\bar{c}_j < 0$ arbitrarily

CHUZR: Choose $i \leftarrow \underset{k \in B,\ \bar{a}_{kj} > 0}{\operatorname{argmin}} \left\{ \frac{\bar{x}_{Bk}}{\bar{a}_{kj}} \right\}$

b) ~~Lexicographic~~ Row Index Rule:

PRICE: Choose j: $\bar{c}_j < 0$ arbitrarily

CHUZR: Choose $i \leftarrow \underset{k=1,\ \bar{a}_{kj} > 0}{\operatorname{argmin}} \left\{ \frac{\bar{x}_{Bk}}{\bar{a}_{kj}} \right\}$

5.12 Rem (Lexicographic Row Index Rule): Let $\varepsilon > 0$, $\varepsilon = \begin{pmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^m \end{pmatrix}$, then $\varepsilon_m \ll \varepsilon_{m-1} \ll \dots \ll \varepsilon_1$. Then the perturbed LP

(P_ε) $\min c^T x$, $Ax = b + \varepsilon$, $x \geq 0$

has the following properties for some coeff. vector ε:

a) $A_B^{-1}(b+\varepsilon) = A_B^{-1}b + A_B^{-1}\varepsilon$ is non-degenerate $\forall$ basis B

b) $\exists$ feasible basis for $(P_\varepsilon) \Rightarrow \exists$ feas. basis for (P)

c) $\exists$ opt. — " — $\Rightarrow \exists$ opt. basis for (P)

Solve (P_ε) instead of (P):

$$\min \; 0^T x_B + (c_N^T - c_B^T A_B^{-1} A_N) x_N$$

$$x_B + A_B^{-1}A_N x_N = A_B^{-1}b + A_B^{-1}\varepsilon = \bar{x}_B + \bar{\varepsilon}$$

CHUZR: $B_i = \underset{k=1}{\overset{m}{\operatorname{argmin}}} \left\{ \dfrac{\bar{x}_{B_k} + \bar{\varepsilon}}{\bar{a}_{kj}} \;,\; \bar{a}_{kj} > 0 \right\}$,

i.e., in case of a tie in $\dfrac{\bar{x}_{B_k}}{\bar{a}_{kj}}$, the pivot with the smallest (= lex. smallest) value $\dfrac{\bar{x}_{B_k} + \sum_{e=1}^{n} a^{-1}_{ke}\varepsilon^e}{\bar{a}_{kj}}$ is chosen.

S.13 Ex (Worst case Runtime Example, Klee-Minty-Cube):

$$(KM_n) \quad \min -\sum_{j=1}^{n} 10^{n-j} x_j$$

$$2\sum_{j=1}^{i-1} 10^{i-j} x_j + x_i \le 100^{i-1}, \quad i = 1, \dots, n$$

$$x \ge 0$$

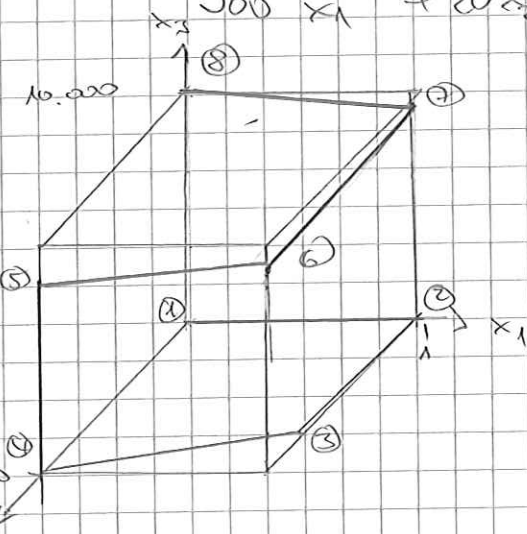
$$(KM_3) \quad \min -100x_1 - 10x_2 - 1x_3$$

$$x_1 \le 1$$

$$20x_1 + x_2 \le 100$$

$$200x_1 + 20x_2 + x_3 \le 10\,000$$

$$x \ge 0$$



Klee-Minty-Cube in $\mathbb{R}^3$

The Simplex-Algorithm takes 2^n pivots to solve (KM_n) starting from the slack basis, if the most negative reduced cost rule is used.

| | | |

8.14

5.14 Rem (Phase I of the Simplex Algorithm):

(P) $\min c^T x$
$Ax = b$ w.l.o.g $b \geq 0$
$x \geq 0$

(PI) $\min \mathbb{1}^T y$
$Ax + y = b$ Note: $\operatorname{rank}(A, I) = m$
$x, y \geq 0$
artificial variables

5.15 Thm (Artificial Variables): Let $A \in \mathbb{R}^{m \times n}$, $\operatorname{rank} A = m$, $b \in \mathbb{R}^m$

(P) $\min c^T x$, $Ax = b$, $x \geq 0$

(PI) $\min \mathbb{1}^T y$, $Ax + Iy = b$, $x, y \geq 0$.

Then:

a) $\mathcal{B} = \{n+1, \dots, n+m\}$ is a primal feas. basis for P with basic solution $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}$.

b) Let $\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \operatorname{argmin}(PI)$. Then:

i) $y \neq 0 \Rightarrow$ (P) infeasible

ii) $y = 0 \Rightarrow x^*$ primal feasible basic sol. of (P).

Proof: a) ✓ b) i) ✓ ii) $y^* = 0 \Rightarrow x^*$ primal feasible for (P).

Let $\mathcal{B}$ be the basis assoc. with $\begin{pmatrix} x^* \\ 0 \end{pmatrix}$, $\mathcal{B}_x = \mathcal{B} \cap [n]$,

$\mathcal{B}_y = [n+m] \setminus \mathcal{B}_x$, $\mathcal{N}_x = [n] \setminus \mathcal{B}_x$, $\mathcal{N}_y = \{n+1, \dots, n+m\} \setminus \mathcal{B}$.

$\Rightarrow A_{\cdot \mathcal{B}_x}$ has lin. indep. columns, $\operatorname{rank} A = \operatorname{rank} A_{\mathcal{B}_x} A_{\mathcal{N}_x} = m$

$\Rightarrow A_{\cdot \mathcal{B}_x}$ can be extended to a basis using columns from $A_{\cdot \mathcal{N}_x}$. □

5.16 Alg (Phase I of the Simplex Algorithm):

Input: LP (P) $\min c^T x$, $Ax = b \geq 0$, $x \geq 0$

Output: i) Primal feasible basis $\mathcal{B}$ with basic sol. $\bar{x}$

ii) Message "(P) infeasible"

iii) Message "$\operatorname{rank}(A) < m$".

1. $\begin{pmatrix} x^* \\ y^* \end{pmatrix} \leftarrow \operatorname{argmin}(PI)$ started with basis $\{n+1, \dots, n+m\}$, $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}$.
2. if ($y^* \neq 0$) then output "(P) infeasible", stop
3. $\mathcal{B}_x \leftarrow \mathcal{B} \cap [n]$, $\mathcal{B}_y \leftarrow \mathcal{B} \setminus \mathcal{B}_x$
$\mathcal{N}_x \leftarrow [n] \setminus \mathcal{B}_x$

while ($|B_x| < m$) do

if ($\exists j \in N_x : A_{\cdot j} = \sum_{k \in B_x} w_k A_{\cdot k} + \sum_{i \in B_y} w_i e_{i-m}$

$\wedge \exists k \in B_y : w_k \neq 0$) then

$B_x \leftarrow B_x \cup \{j\}$, $B_y \leftarrow B_y \setminus \{j\}$, $N_x \leftarrow N_x \setminus \{j\}$

else output "rank(A) < m", stop.

endif

endwhile

$B \leftarrow B_x$, stop

5.17 Prop. (Phase I) Alg 5.16 is correct.

Proof: Basis extension

$(A_{\cdot B_x}, I_{B_y - m\mathbb{1}})$ regular $\Leftrightarrow$ $(A_{\cdot B_x}, I_{B_y - m\mathbb{1}})\, w = A_{\cdot j}$ feas.

$w_{B_y - m\mathbb{1}} \neq 0 \Rightarrow \exists k \in B_y : w_k \neq 0$

$\Rightarrow$ basis exchange $(A_{\cdot B_x \cup \{j\}}, I_{\cdot (B_y \setminus \{i\}) - m\mathbb{1}})$ regular. □

5.18 Ex (Phase I)

(P) $\min -x_1 - x_2 - x_3$

$$\begin{aligned} -2x_1 - 2x_2 + x_3 &= -4 \\ x_1 + x_2 + x_3 &= 2 \\ x_1 + x_3 &= 1 \\ x &\geq 0 \end{aligned}$$

(PI) $\min x_4 + x_5 + x_6$

$$\begin{aligned} +2x_1 + 2x_2 - x_3 + x_4 &= 4 \\ x_1 + x_2 + x_3 + x_5 &= 2 \\ x_1 + x_3 + x_6 &= 1 \end{aligned}$$

			4	5	6	
0	0	0	1	1	1	0
2	2	-1	1	0	0	4
1	1	1	0	1	0	2
1	0	1	0	0	1	1
			5	6		
0	-3	3	0	0	4	-3
0	2	-3	1	0	-2	2
0	(1)	0	0	1	0	1
1	0	1	0	0	1	1

			4	5	6	
-4	-3	-1	0	0	0	-7
2	2	-1	1	0	0	4
1	1	1	0	1	0	2
(1)	0	1	0	0	1	1
			5			
0	0	3	0	3	4	0
0	0	-3	1	2	-2	0
0	1	0	0	1	0	1
1	0	1	0	0	1	1

$A_{\cdot B_x}$ $A_{\cdot B_y}$ $A_{\cdot 3}$

$$\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = 1 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} - 3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$\neq 0$

5.19 Satz (BTRAN, FTRAN Implementation):

BTRAN: $y^T A_{\cdot B} = c_B^T$

FTRAN: $A_{\cdot B} \bar{A}_{\cdot j} = A_{\cdot j}$

} solution of a system of lin. equations

Basis inverse A_B^{-1}: Computation takes $O(n^{2+\varepsilon})$ steps, $\varepsilon \in (0,1]$, numerically unstable

LU factorization: $LU = PAQ$, $PP^T = QQ^T = I$

L: lower Δ-matrix; U: upper Δ-matrix; P: row permutation matrices; Q: col permutation matrices

a) $b = A_{\cdot B} x \Leftrightarrow Pb = PA_{\cdot B} \underbrace{Q}_{=x} \underbrace{z}_{=:w} = LUz$

$\Rightarrow$
1. $Lw = Pb$ forward substitution
2. $Uz = w$ backward substitution
3. $Qz = x$ permutation

} $O(n^2)$

and less, if L, U are sparse

b) $c_B^T = y^T A_{\cdot B}$ analogous

c) LU update

$A_{\cdot B} = LU$, $B' = (B \setminus \{B_i\}) \cup \{j\}$, $A_{\cdot B} w = A_{\cdot j}$ $(w_i \neq 0)$

$$F = \begin{pmatrix} 1 & & w_1 & & \\ & \ddots & \vdots & & \\ & & w_i & & \\ & & \vdots & \ddots & \\ & & w_m & & 1 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & & -w_1/w_i & & \\ & \ddots & \vdots & & \\ & & 1/w_i & & \\ & & \vdots & \ddots & \\ & & -w_m/w_i & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & \eta_1 & & \\ & \ddots & \vdots & & \\ & & \eta_i & & \\ & & \vdots & \ddots & \\ & & \eta_m & & 1 \end{pmatrix}$$

eta-matrix

where $\eta_k = \begin{cases} -w_k/w_i, & k \neq i \\ 1/w_i, & k = i \end{cases}$. Then:

i) $F \cdot E = I \Leftrightarrow E = F^{-1}$

ii) $A_{\cdot B} \cdot F = A_{\cdot B'}$

iii) $A_{\cdot B'}^{-1} = (A_{\cdot B} \cdot F)^{-1} = F^{-1} A_{\cdot B}^{-1} = E A_{\cdot B}^{-1}$

d) FTRAN: $A_{\cdot j} = A_{\cdot B'} w' = A_B \underbrace{F w'}_{=:w}$, $w = F w'$

$B' = (B_1 | \cdots | \underbrace{j}_{\text{leaving}} | \cdots | B_m) =: B^{(1)}$, analogously $E^{(k)}, F^{(k)}$

$\Rightarrow$
$A_{\cdot B^{(k)}} = A_{\cdot B} F^1 \cdots F^k$
$A_{\cdot B^{(k)}}^{-1} = E^k \cdots E^1 A_{\cdot B}^{-1}$

$$A_{\cdot, j^{(k)}} = A_{\cdot, J^{(k)}} w^k = A_{\cdot B} \underbrace{F^1 \cdots F^k w^k}_{=: w_0} \Rightarrow w_0 = F^1 \cdots F^k w^k$$

$$\Rightarrow w_k = E^k \cdots E^1 w_0$$

w_1

w_k

FTRAN

$$y^{kT} = c_{B^k}^T A_{B^k}^{-1} = \underbrace{c_{B^k}^T E^k}_{=: v^{kT}} \cdots E^1 A_{\cdot B}^{-1}$$

BTRAN

y^{kT}

recompute LU factorization periodically, e.g., every 200 iterations, to avoid numerical problems

5.20 Rem (Pivot rules):

a) Smallest index rule
   $j \leftarrow \min \{ j : \bar{c}_j < 0 \}$
   part of Bland anti-cycling rule

b) Dantzig rule (most negative coeff., full pricing)
   $j \leftarrow \operatorname{argmin} \{ \bar{c}_j < 0 \}$
   popular, can be expensive

c) partial pricing
   $j \leftarrow \operatorname{argmin}_{j \in J} \{ \bar{c}_j < 0 \}$, $J \subseteq \{1, \dots, n\}$
   J is updated on exhaustion

d) multiple pricing
   first consider variables that had negative reduced costs in earlier iterations

e) steepest edge pricing

$$\bar{x}' \leftarrow \begin{pmatrix} \bar{x}_B \\ \bar{x}_N \end{pmatrix} + \mu \underbrace{\begin{pmatrix} -\bar{A}_{\cdot j} \\ e_j \end{pmatrix}}_{=\bar{d}_j \text{ direction}}, \quad \bar{A}_{\cdot j} = A_{\cdot B}^{-1} A_{\cdot j}$$

$$c^T \bar{x}' = c^T \bar{x} + \bar{c}_j \mu = c^T \bar{x} + \mu \cdot c^T \begin{pmatrix} -\bar{A}_{\cdot j} \\ e_j \end{pmatrix} = c^T \bar{x} + \mu c^T \bar{d}_j$$

progress in direction $\bar{d}_j$

$$j \leftarrow \operatorname{argmin} \left\{ \frac{\bar{c}_j}{\|(-\bar{A}_{\cdot j}, e_j)\|} \mid \bar{c}_j < 0 \right\}$$ "biggest possible progress in direction $\bar{d}_j$"

$$= \operatorname{argmin} \left\{ \frac{\bar{c}_j}{\|\bar{d}_j\|} \mid \bar{c}_j < 0 \right\}$$

efficient implementation using update formula for $\frac{\bar{d}_j}{\|\bar{d}_j\|}$ or approximation of it (Devex pricing)

5.21 Rem (Ratio Test):

$$\mu = \min \left\{ \frac{\bar{b}_i}{\bar{a}_{ij}} \mid \bar{a}_{ij} > 0 \right\}$$ suppose, if $\bar{b}_i, \bar{a}_{ij}$ small

Ex.:
$$\left(\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 10^6 & 0 \end{array}\right) \rightarrow \left(\begin{array}{cc|c} -10^6 & 0 & 1 \\ 10^6 & 1 & 0 \end{array}\right)$$

$$t \leftarrow \min \left\{ \frac{\bar{b}_i + \varepsilon}{\bar{a}_{ij}} \mid \bar{a}_{ij} > 0 \right\}$$

$$j \in \operatorname{argmax} \left\{ \bar{a}_{ij} : \frac{\bar{b}_i}{\bar{a}_{ij}} \leq t, \bar{a}_{ij} > 0 \right\}$$

Ex. (cont'd): $\varepsilon = 10^{-5}$

$$\left(\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 10^{-6} & 0 \end{array}\right) \left| \begin{array}{c} +10^{-5} \\ +10^{-1} \end{array} \right. \rightarrow \left(\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 0 & -10^{-6} \end{array}\right) < 0$$

If $\begin{pmatrix} \bar{x}_B \\ \bar{x}_N \end{pmatrix} + \mu \begin{pmatrix} -\bar{A}_{\cdot j} \\ e_j \end{pmatrix} \neq 0$

then $\bar{x}_i \geq 0 \rightsquigarrow \bar{x}_i \geq -\varepsilon$ (shifting)

Remove shift later (with bounding technique, see later).

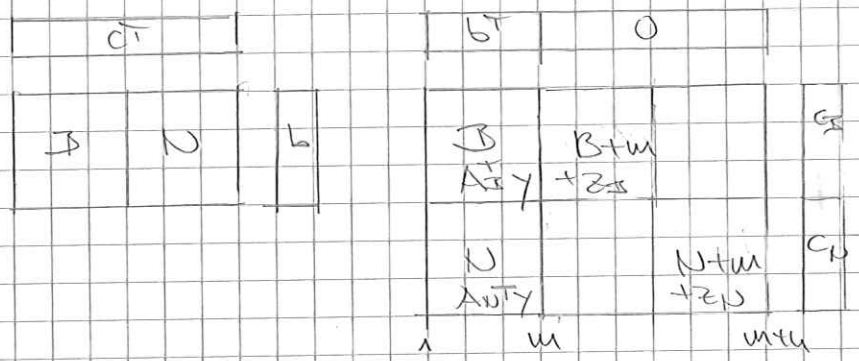
5.22 Rem (Idea of the Dual Simplex Alg.):

(P) $\min c^T x$ s.t. $Ax = b$, $x \geq 0$ (P) $\max y^T b$ s.t. $y^T A \leq c^T$

$= \max b^T y$ s.t. $A^T y + z = c$, $z \geq 0$, y free (cannot leave basis)

$D = (A^T, I)$

$= \max b^T y$ s.t. $D \begin{pmatrix} y \\ z \end{pmatrix} = c$, $z \geq 0$



dual basis $B \cup (N+m) =: H$ s.t. D_H is regular

5.23 Thm:

Let (P) and (D) be a pair of dual LPs as in 5.22 and rank $A = m$ be full.

Then (D) is unbounded or has an optimal basis $H \supseteq \{1, \dots, m\}$.

Proof: s. D

5.24 Prop (Primal and Dual Feasibility):

Let

(P) $\min c^T x$ s.t. $Ax = b$, $x \geq 0$ (D) $\max b^T y$ s.t. $A^T y + z = D \begin{pmatrix} y \\ z \end{pmatrix} = c$, $z \geq 0$

rank $A = m$, $H \supseteq \{1, \dots, m\}$ be a primal feas basis for (P),

$H = \{1, \dots, m\} \cup (N+m)$, $B = \{1, \dots, n\} \setminus N$.

Then: H primal feasible basis for (D)

$\hookrightarrow$ B dual feasible basis for (P)

Proof: H primal feas. basis for (P)

$\Leftrightarrow A_B^T y + z_B = c_B$, $\Leftrightarrow c_B^T - y^T A_B - z_B^T = 0$

$A_N^T y + z_N = c_N$ $\quad c_N^T - y^T A_N = z_N^T \geq 0$

$z_B = 0$

$z_N \geq 0$

$\Leftrightarrow$ B dual feas. basis for (P). $\square$

5.25 Rem (Idea of Dual Simplex Algorithm):

Apply Simplex Algorithm to (D) starting from basis $H = [m] \cup (n+m)$

START: $\bar{x}^T \begin{pmatrix} A_B^T & 0 \\ A_N^T & I_N \end{pmatrix} = (b^T, 0)$

$\Leftrightarrow \bar{x}_N = 0, \; A_B^T \bar{x}_B = b$

PRICE: $i_i = \arg\max \{ 0_B^T - (\bar{x}_B^T, \bar{x}_N^T) \begin{pmatrix} e_i \\ 0 \end{pmatrix} = -\bar{x}_B \}$

if $\bar{x}_B \geq 0$ stop, optimal

PIVOT: $\begin{pmatrix} A_B^T & 0 \\ A_N^T & I_N \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} e_i \\ 0 \end{pmatrix}$

$\Leftrightarrow v^T A_B = e_i^T \quad v^T A_N + w^T = 0$

$\Leftrightarrow v^T = e_i^T A_B^{-1}, \quad w^T = -e_i^T A_B^{-1} A_N = -\bar{A}_{i\cdot}$

$= A_{B^{-1}i\cdot}$

CHUZE: $\begin{pmatrix} y \\ z_N \end{pmatrix} \leftarrow \begin{pmatrix} y \\ z_N \end{pmatrix} - \mu \begin{pmatrix} A_{B^{-1}i\cdot} \\ -\bar{A}_{i\cdot} \end{pmatrix}$ free ≥ 0

$\Rightarrow \mu = \frac{z_j}{\bar{a}_{ij}} = \min \left\{ \frac{z_j}{\bar{a}_{ij}} : \bar{a}_{ij} < 0 \right\} = \min \left[\frac{\bar{c}_j}{\bar{a}_{ij}} \mid a_{ij} < \right.$

if $\mu = +\infty$ stop, unbounded

UPDATE: $\begin{pmatrix} y \\ z_N \end{pmatrix} = \begin{pmatrix} y \\ z_N \end{pmatrix} + \mu \begin{pmatrix} -A_B^{-T} \\ \bar{A}_{i\cdot}^T \end{pmatrix}$

$\bar{z}_i \leftarrow \mu$

$H \leftarrow (H \setminus \{j\}) \cup \{i\}$.

5.26 Alg (Dual Simplex Algorithm, Lemke [1954]):

Input: (LP) $\min c^T x, \; Ax = b, \; x \geq 0$,

dual feas basis B, i.e.,

$\bar{c}_N^T = z_N^T - c_B^T - c_B^T A_B^{-1} A_N \geq 0$

Output: i) opt. sol. $\bar{x}$ for (P)

resp. opt sol $\bar{y}^T = \bar{c}_B^T A_B^{-1}$ for (D)

resp. opt sol $\bar{y}^T, \bar{z}_N \geq 0, \bar{z}_B = 0$ for (D)

ii) Message "$P=(A,b) = \emptyset$ or (D) unbounded"

1. BTRAN: $\bar{x}_B = A_B^{-1} b$

2. PRICE: if $\bar{x}_B \geq 0$ stop optimal.
$i \leftarrow \bar{x}_{B_i} < 0$.

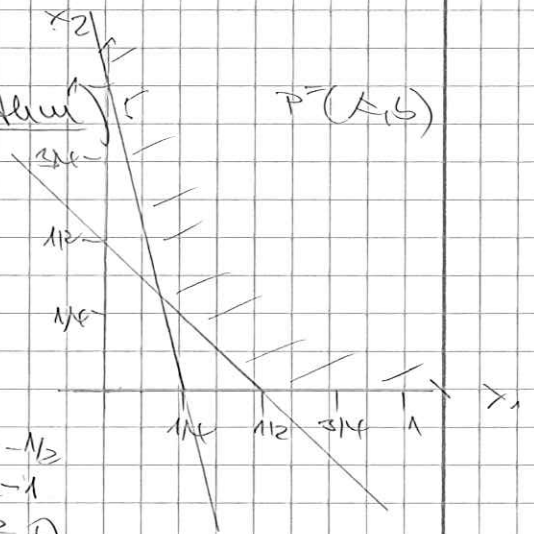
3. FTRAN: $\bar{v}^T \leftarrow e_i^T A_B^{-1}$, $\bar{w}^T \leftarrow -\bar{v}^T A_N = -e_i^T A_B^{-1} A_N = -\bar{a}_i$.

4. CHUZR: if $\bar{w}^T \geq 0$ stop "$P=(A,b) = \emptyset$ or (D) unbounded"
$j \leftarrow \frac{\bar{z}_j}{|\bar{a}_{ij}|} = \min\left[\frac{\bar{z}_j}{|\bar{a}_{ij}|}, \bar{a}_{ij} < 0\right]$

5. UPDATE: $N \leftarrow (N \setminus \{j\}) \cup B_i$
$B_i \leftarrow j$
$\bar{y} \leftarrow \bar{y} - \theta \bar{v}$
$\bar{z}_N \leftarrow \bar{z}_N - \theta \bar{w}$
$\bar{z}_{B_i} \leftarrow \theta$

SDA Ex (Dual Simplex Algorithm):

$$\begin{aligned} \min\ & 2x_1 + x_2 \\ & -x_1 - x_2 \leq -1/2 \quad (1) \\ & -4x_1 - x_2 \leq -1 \quad (2) \\ & x_1 \geq 0 \quad (3) \\ & x_2 \geq 0 \quad (4) \end{aligned}$$



$\Rightarrow$ $\min\ 2x_1 + x_2$

$$\begin{aligned} -x_1 - x_2 + x_3 &= -1/2 \\ -4x_1 - x_2 + x_4 &= -1 \\ x &\geq 0 \end{aligned}$$

		3	4			1		3		
2	1	0	0	0		0	1/2	0	1/2	-1/2
-1	-1	1	0	-1/2	$\rightarrow$	0	(-3/4)	1	-1/4	-1/4
(-4)	-1	0	1	-1		1	1/4	0	-1/4	1/4

≥ 0, i.e., dual feas.

0	0	2/3	1/3	-2/3
0	1	4/3	1/3	1/3
1	0	1/3	-1/3	1/6

opt $\bar{x}^* = \left(\frac{1}{6}, \frac{1}{3}, 0, 0\right)^T$

$\bar{y} = \left(-\frac{2}{3}, -\frac{1}{3}\right)$

$\bar{z} = \left(0, 0, \frac{2}{3}, \frac{1}{3}\right)$

	4	2			
	2	1	0	0	0
3	-1	-1	1	0	-1
4	(-4)	-1	0	1	-1/2 ←

	1	2 (2/3)	3	4 (2)	
	0	1/2	0	1/2	-1/2
3	0	(-3/4)	1	-1/4	-1/4 ←
1	1	1/4	0	-1/4	1/4

	0	0	2/3	1/3	-2/3
2	0	1	4/3	1/3	1/3
1	1	0	1/3	-1/3	1/6

opt. $\Rightarrow \bar{x}^* = (\frac{1}{6}, \frac{1}{3}, 0, 0)$

5.21 Rem (Upper and lower bounds):

$$(P_{eu}) \quad \min c^T x \quad Ax = b \quad l \le x \le u \qquad l \in (\mathbb{R} \cup \{-\infty\})^n, \; u \in (\mathbb{R} \cup \{+\infty\})^n$$

Transformations: $x_j = x_j^+ - x_j^-$

i) $-\infty = l_j \le x_j \le u_j = +\infty \Rightarrow 0 \le x_j^+, x_j^- \le +\infty$

ii) $-\infty < l_j \le x_j \le u_j$ $\quad x_j' := x_j - l_j \Rightarrow 0 \le x_j' \le u_j - l_j$

iii) $-\infty = l_j \le x_j \le u_j < +\infty$ $\quad x_j' = u_j - x_j \Rightarrow 0 \le x_j' \le +\infty$

5.22 Def (Standard form with upper bounds):

$$(P_u) \quad \min c^T x \quad Ax = b \quad 0 \le x \le u$$

standard form with upper bounds

5.23 Rem (Solution of LPS with upper bounds):

Goal: Treatment of upper bounds without increasing the basis

Idea: Subdivision of the non-basis $N = N_e \cup N_u$

$$N_e = \{j \in N : x_j = 0\}$$

$$N_u = \{j \in N : x_j = u_j\}$$

$$\bar{x}_B = A_B^{-1} b - A_B^{-1} A_{N_u} \underbrace{\bar{x}_{N_u}}_{=u_{N_u}} - A_B^{-1} A_{N_e} \underbrace{\bar{x}_{N_e}}_{=0}$$

$$= \bar{b} - \bar{A}_{\cdot N_u} \underbrace{\bar{x}_{N_u}}_{=u_{N_u}} - \bar{A}_{\cdot N_e} \bar{x}_{N_e}$$ — primal basic solution

$$0 \le \bar{x}_B = \bar{b} - \bar{A}_{\cdot N_u} u_{N_u} \le u_B$$ — primal feas.

$$c^T \bar{x} = c_B^T \bar{x}_B + c_{N_u}^T x_{N_u} + c_{N_e}^T \bar{x}_{N_e}$$

$$= c_B^T \bar{b} - c_B^T \bar{A}_{\cdot N_u} \bar{x}_{N_u} - c_B^T A_{\cdot N_e} \bar{x}_{N_e} + c_{N_u}^T \bar{x}_{N_u} + c_{N_e}^T \bar{x}_{N_e}$$

$$= c_B^T \bar{b} + \underbrace{\bar{c}_{N_u}^T}_{\le 0} \bar{x}_{N_u} + \underbrace{\bar{c}_{N_e}^T}_{\ge 0} \bar{x}_{N_e}$$

opt. $\Rightarrow$ dual feasibility

$$(D_u)\quad \max\ y^T b - v^T u$$
$$y^T A - v^T \le c^T$$
$$v \ge 0$$

$$= \max\ y^T b - v^T u$$
$$(c^T - y^T A) + v^T \ge 0$$
$$v \ge 0$$

$$y_B^T = c_B^T A_B^{-1}$$
$$v_B^T = 0$$
$$v_{N_e}^T = 0$$
$$v_{N_u}^T = -(c_{N_u}^T - y^T A_{\cdot N_u}) = -\bar{c}_{N_u}^T$$

dual basic solution

$$\Rightarrow\quad v \ge 0 \iff v_{N_u} = -\bar{c}_{N_u} \ge 0 \iff \bar{c}_{N_u} \le 0$$
$$(c^T - y^T A) + v^T = \bar{c}^T + v^T \ge 0 \iff \bar{c}_{N_e} \ge 0$$

i) $j \in N_e,\ \bar{x}_j = 0 \to x_j = \theta \le u_j$

$$0 \le \bar{x}_{B_i} = b_i - \theta \bar{a}_{ij} \le u_i \quad \forall i \in B$$

$$\Rightarrow \theta = \min\left\{ u_j,\ \underbrace{\min\left\{ \frac{b_i}{\bar{a}_{ij}} : a_{ij} > 0 \right\}}_{=:\theta_e},\ \underbrace{\min\left\{ \frac{u_i - b_i}{|\bar{a}_{ij}|} : \bar{a}_{ij} < 0 \right\}}_{=:\theta_u} \right\}$$

$$= \min\{u_j, \theta_e, \theta_u\}$$

ii) $j \in N_u,\ \bar{x}_j = u_j \to x_j = u_j - \theta$

$$0 \le \bar{x}_{B_i} = b_i + \theta \bar{a}_{ij} \le u_i \quad \forall i \in B$$

$$\Rightarrow \theta = \min\left\{ u_j,\ \underbrace{\min\left\{ \frac{u_i - b_i}{\bar{a}_{ij}} : \bar{a}_{ij} > 0 \right\}}_{=\theta_u},\ \underbrace{\min\left\{ \frac{b_i}{|\bar{a}_{ij}|} : a_{ij} < 0 \right\}}_{=\theta_e} \right\}$$

$$= \min\{u_j, \theta_u, \theta_e\}$$

S.x. Alg. (Simplex Alg. with Upper Bounding Technique):

Input: (D_u), primal feas. ext. basis (B, N_u) with sol $\bar{x}$.

Output: Opt sol x or message "(P_u) unbounded"

1. DUAL: $y^T = c_B^T A_B^{-1}$
2. PRICE: $\bar{c}^T = c^T - y^T A$

if ($\bar{c}_{N_e}^T \ge 0 \wedge \bar{c}_{N_u}^T \le 0$) then stop, output $\bar{x}$, opt

choose $j \in N_e$: $\bar{c}_j < 0$ or $j \in N_u$: $\bar{c}_j > 0$

3. TRAN: $\overline{A}_{\cdot j} = w = A_B^{-1} A_{\cdot j}$

4. CHUZR: if $(j \in N_e)$ then

$$\theta_e = \min\left\{\frac{\bar{b}_i}{\bar{a}_{ij}} : a_{ij} > 0\right\}, \quad \theta_u = \min\left\{\frac{u_i - \bar{b}_i}{|\bar{a}_{ij}|} : a_{ij} < 0\right\}$$

else if $(j \in N_u)$ then

$$\theta = \min\left\{\frac{\bar{b}_i}{|\bar{a}_{ij}|} : \bar{a}_{ij} < 0\right\}, \quad \theta_u = \min\left\{\frac{u_i - \bar{b}_i}{\bar{a}_{ij}} : a_{ij} > 0\right\}$$

endif

$\theta = \min\{u_j, \theta_e, \theta_u\}$

if $(\theta = +\infty)$ then stop "problem unbounded"

B_i = anyone with θ

5. UPDATE: if $(j \in N_e)$ then

$\bar{b} \leftarrow \bar{b} - \theta \bar{A}_{\cdot j}$

$\bar{x}_j \leftarrow \theta$

if $(\theta = u_j)$ then

$N_e \leftarrow N_e \setminus \{j\}, \quad N_u \leftarrow N_u \cup \{j\}$

else if $(\theta = \theta_e)$

$B \leftarrow (B \setminus \{B_i\}) \cup \{j\}$

$N_e \leftarrow (N \setminus \{j\}) \cup \{B_i\}$

else

$B \leftarrow (B \setminus \{B_i\}) \cup \{j\}$

$N_u \leftarrow N_u \cup \{B_i\}$

$N_e \leftarrow N_e \setminus \{j\}$

endif

else

...

endif

5.25 Ex (Upper Bounding Technique):

$$\min\ x_1 - 2x_2$$
$$x_1 + x_2 \le 3$$
$$2x_1 + x_2 \le 5$$
$$0 \le x_1 \le 2$$
$$0 \le x_2 \le 2$$

$$= \min\ x_1 - 2x_2$$
$$x_1 + x_2 + x_3 = 3$$
$$2x_1 + x_2 + x_4 = 5$$
$$0 \le x_1, x_2 \le 2$$
$$x_3, x_4 \ge 0$$

1	2												
-1	-2	0	0	0			-1	-2	0	0	4		
1	(1)	1	0	3	3		(1)	1	1	0	1	3	
2	1	0	1	5	4		2	1	0	1	3	4	
	-2	-2											
0	-1	1	0	5									
1	1	1	0	1	1								
0	-1	-2	1	1	4								

$\Rightarrow x^* = (1, 2, 0, 1)^T$

Notation: N_u is denoted with negative indices.

5. Rem (Sensitivity Analysis):

(P) $\min c^T x$ opt basis basis B
$Ax = b$
$x \geq 0$ opt. basis solution $\bar{x} = (\bar{x}_B, \bar{x}_N) = (\bar{b}, 0)$

Change in (P):

(D) $\max y^T b$
$y^T A + z^T = c^T$
$z \geq 0$.

a) Change in objective c
b) — " — rhs b } are easy
c) — " — $A_{\cdot N}$

d) Adding a variable
e) Adding a constraint

How can we "reoptimize"?

a) Change in objective $c \to \tilde{c}$

$\tilde{x}_B = A_{\cdot B}^{-1} b = \bar{b}$ unchanged, primal feasible
$\tilde{y}^T = \tilde{c}_B^T A_{\cdot B}^{-1}$
$\tilde{d}^T = \tilde{c}_N^T - \tilde{y}^T A_{\cdot N}$
$= \tilde{c}_N^T - \tilde{c}_B^T A_{\cdot N} \begin{cases} \geq 0 & \text{optimal} \\ \not\geq 0 & \text{primal Simplex alg.} \end{cases}$

b) Change of rhs $b \to \tilde{b}$

$\bar{y}^T = c_B^T A_B^{-1}$ unchanged
$\bar{c}_N^T = c_N^T - \bar{y} A_{\cdot N} \geq 0$ unchanged
$\tilde{x}_B = A_B^{-1} \tilde{b} \begin{cases} \geq 0 & \text{optimal} \\ \not\geq 0 & \text{dual Simplex alg.} \end{cases}$

c) Change in A_N $a_{ij} \to \tilde{a}_{ij}$

$x_B = A_B^{-1} b \geq$ unchanged
$\bar{y}^T = c_B^T A_B^{-1}$ "
$\tilde{c}_j = c_j - \bar{y}^T \tilde{A}_{\cdot j} \begin{cases} \geq 0 & \text{optimal} \\ \not\geq 0 & \text{primal Simplex alg.} \end{cases}$

d) Adding a variable $x_{n+1} \geq 0$, c_{n+1}, $A_{\cdot n+1}$
$\tilde{N} \leftarrow N \cup \{n+1\}$
$x_B = A_B^{-1} b \geq$ unchanged
$\bar{y}^T = c_B^T A_B^{-1}$ "

$\bar{c}_{n+1} = c_{n+1} - y^T A_{\cdot n+1} \begin{cases} \geq 0 & \text{optimal} \\ < 0 & \text{primal simplex alg.} \end{cases}$

e) Adding a constraint $A_{m+1,\cdot} x + x_{n+1} = b_{m+1}$, $x_{n+1} \geq 0$

$\tilde{B} \leftarrow B \cup \{n+1\}$

$\tilde{x}_B = A_{\tilde{B}}^{-1} b = \begin{pmatrix} A_B^{-1} b \\ b_{m+1} - A_{m+1,\cdot} A_B^{-1} b \end{pmatrix} \geq 0$

$\Leftrightarrow \tilde{x}_{n+1} = b_{m+1} - A_{m+1,\cdot} A_B^{-1} b \geq 0$

$\tilde{y}^T = (c_B^T, 0) \begin{pmatrix} A_B^{-1} & 0 \\ -A_{m+1,\cdot} A_B^{-1} & 1 \end{pmatrix} = (c_B^T A_B^{-1}, 0) = (y^T, 0)$

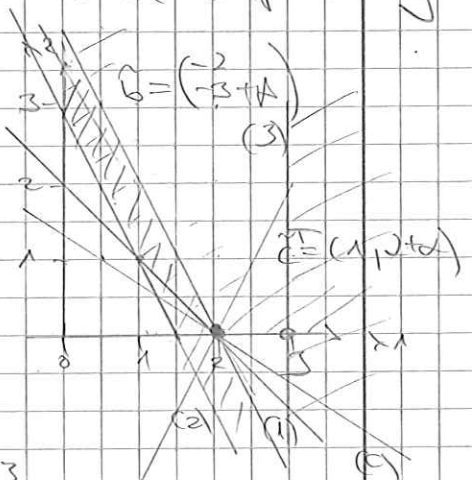
$\tilde{c}_N^T = \tilde{c}_N^T - (y^T, 0) A_{\tilde{B}} = (c_B^T, 0) \geq 0$ dual feas.

$\tilde{x}_{n+1} = b_{m+1} - A_{m+1,\cdot} A_B^{-1} b \begin{cases} \geq 0 & \text{optimal} \\ < 0 & \text{dual simplex alg.} \end{cases}$

5. Ex (Sensitivity Analysis):

(1) $\min\ x_1 + 2x_2$ (0)
$-x_1 - x_2 \leq -2$ (1)
$-2x_1 - x_2 \leq -3$ (2)
$x \geq 0$

$= \min\ x_1 + 2x_2$
$-x_1 - x_2 + x_3 = -2$
$-2x_1 - x_2 + x_4 = -3$
$x \geq 0$



dual feas.

1	2	0	0	0		0	1	1	0	-2
(-1)	-1	1	0	-2	3	1	1	-1	0	2	1
-2	-1	0	1	-3	4	0	1	-2	1	1	4
1	2									

a) $\tilde{c}_2 \leftarrow c_2 + \alpha = 2 + \alpha$

0	1+2	1	0	-2
1	1	-1	0	2
0	1	-2	1	1

$1 + \alpha \geq 0 \Leftrightarrow \alpha \geq -1 \Leftrightarrow \tilde{c}_2 \geq 1$

b) $\tilde{b}_2 = b_2 + \beta = -3 + \beta$

$A_B^{-1} \tilde{b} = \begin{pmatrix} -1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ -3+\beta \end{pmatrix} = \begin{pmatrix} 2 \\ 1+\beta \end{pmatrix}$

$1 + \beta \geq 0 \Leftrightarrow \beta \geq -1 \Leftrightarrow \tilde{b}_2 \geq -4$

c) $\tilde{a}_{12} = a_{12} + \gamma = -1 + \gamma$

$A_B^{-1} A_{\cdot 2} = \begin{pmatrix} -1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} -1+\gamma \\ -1 \end{pmatrix} = \begin{pmatrix} 1-\gamma \\ 1-2\gamma \end{pmatrix} = \tilde{A}_{\cdot 2}$

$\tilde{c}_2 = c_2 - y^T A_2 = 2 - (-1,0)\begin{pmatrix} -1+\mu \\ -1 \end{pmatrix} = 1+\mu \geq 0 \Leftrightarrow \mu \geq -1$

$\Leftrightarrow \tilde{a}_{12} \geq -2$

d) c_5, a_{15}, a_{25}

$\tilde{A}_5 = \begin{pmatrix} -1 & 0 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} a_{15} \\ a_{25} \end{pmatrix} = \begin{pmatrix} -a_{15} \\ -2a_{15}+a_{25} \end{pmatrix}$

$\tilde{c}_5 = c_5 - y^T A_5 = c_5 - (-1,0)\begin{pmatrix} a_{15} \\ a_{25} \end{pmatrix}$

$= c_5 + a_{15} \geq 0 \Leftrightarrow c_5 \geq -a_{15}$

0	1	1	0	c_5+a_{15}	2
1	1	-1	0	$-a_{15}$	2
0	1	-2	1	$-2a_{15}+a_{25}$	1

e) $x_1 \geq 0 \Leftrightarrow -x_1 \leq -3$ (3)

								2			5		
0	1	1	0	0	-2	0	2	0	0	1	-3		
1	1	-1	0	0	2	1	0	0	0	-1	3	1	
0	1	-2	1	0	1	0	-1	0	1	-2	3	4	
-1	0	0	0	1	-3	0	-1	1	0	-1	1	3	
0	1	(-1)	0	1	-1								

5. Def (Perturbation Function): Let $P^=(A,b) \in \mathbb{R}^n$.

$Z = \{b \in \mathbb{R}^m : P^=(A,b) \neq \emptyset\}$, $0 \in Z \Leftrightarrow P^=(A,0) \neq \emptyset$

$V : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$, $b \mapsto \inf_{x \in P^=(A,b)} c^T x$ Perturbation function assoc. with $P^=(A,b)$

5. Thm (Perturbation Function): Let

$V : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$, $b \mapsto \inf_{x \in P^=(A,b)} c^T x$ be the perturbation function assoc. with the (LP) (P) ...

a) $\mathcal{E}(V)$ is a polyhedral cone

b) V is continuous on Z°.

Proof: a) $\mathcal{E}(V) = \text{proj}_{(b,\mu)}\{(x,b,\mu) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} : \underbrace{Ax=b, x \geq 0, c^T x \leq \mu}_{\text{Polyhedron}}\}$

Polyhedron

$(b,\mu) \in \mathcal{E}(V) \Rightarrow (\lambda b, \lambda \mu) \in \mathcal{E}(V)$, $\lambda \in \mathbb{R}_{\geq 0}$ ✓

b) $\mathcal{E}(V)$ cone $\Rightarrow \mathcal{E}(V)$ convex $\Rightarrow V$ convex $\Rightarrow V$ cont. in Z° ☐

6 Optimality Conditions for Nonlinear Problems

6.1 Def (Nonlinear Programming Problem): Let $C^1 \ni f, g_i, h_j : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$, $j = 1, \dots, k$ be functions

$$(NLP) \quad \min f(x) \quad \text{s.t.} \quad g_i(x) \le 0, \; i = 1, \dots, m \quad h_j(x) = 0, \; j = 1, \dots, k$$

f = objective; $g_i(x) \le 0$ = inequality constraints; $h_j(x) = 0$ = equality constraints; (General) Nonlinear Optimization Problem

f, g_i, h_j are not necessarily convex or linear.

$$X := X(g,h) = \{ x \in \mathbb{R}^n : g_i(x) \le 0, \; i = 1, \dots, m; \; h_j(x) = 0, \; j = 1, \dots, k \}$$

— feasible set (of NLP)

(NLP) $\min f(x)$, $g_i(x) \le 0$, $i = 1, \dots, m$

Special case: $X = P(A|b)$, i.e.,

(NP) $\min f(x)$ s.t. $Ax \le b$

Non-linear Optimization Problem with linear constraints

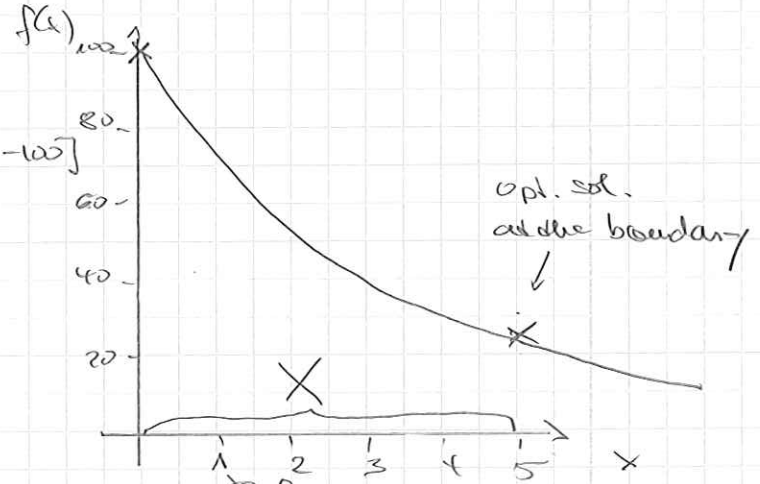
6.2 Ex (NLP):

a) $\min x^2 - 20x \quad [+100 - 100]$

$0 \le x \le 5$

$\Leftrightarrow \min (x-10)^2$

$0 \le x \le 5$

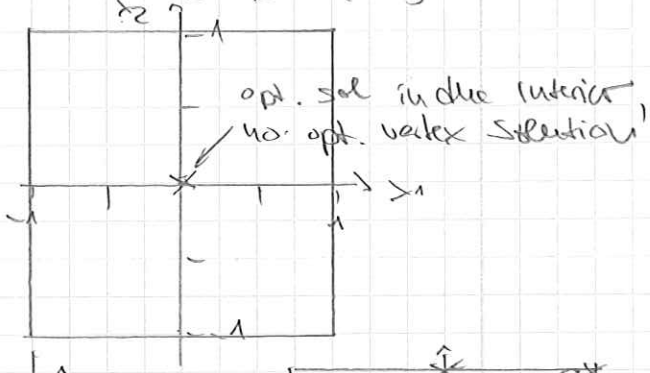


b) $\min \|x\|_2^2$

$-1 \le x_1, x_2 \le 1$

$\Leftrightarrow \min x_1^2 + x_2^2$

$-1 \le x_1, x_2 \le 1$

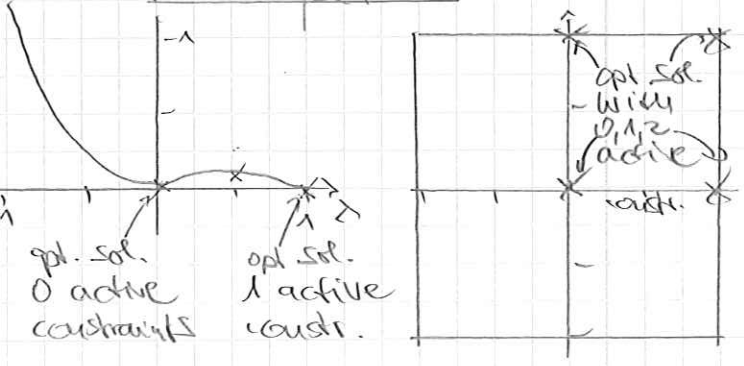


c) $\min x^2(1-x)^2$

$-1 \le x \le 1$

$\min x^2(1-x)^2 + y^2(1-y)^2$

$-1 \le x, y \le 1$



d) $\min x(x-1)(x-2)$

$-1 \leq x \leq 3$

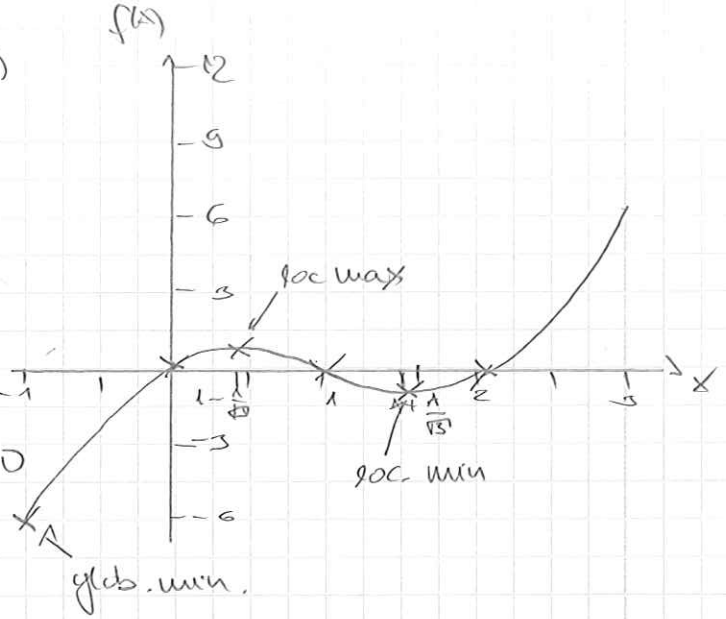
$\Leftrightarrow \min x^3 - 3x^2 + 2x$

$-1 \leq x \leq 3$

$f'(x) = 3x^2 - 6x + 2$

$= 3(x-1)^2 - 1 = 0$

$\Leftrightarrow x = 1 \pm \frac{1}{\sqrt{3}}$



6.3 Rem (general Properties of NLPs): For NLPs hold:

a) A local optimum is in general not a global optimum.

b) The global opt. is not necessarily at the boundary of the feasible region, and if so, not necessarily at an extreme point.

c) There is no algorithm that can compute the optimal solution of a general NLP in a finite number of steps.

6.4 Rem (Idea for optimality conditions for NLPs):

x^* loc opt. for NLP — feasible directions S at x^* in $U_\varepsilon(x^*)$

$\hookrightarrow \exists \varepsilon > 0: f(x^* + s) \geq f(x^*) \quad \forall x + s \in U_\varepsilon(x^*) \cap X$

$\Leftrightarrow \exists \varepsilon > 0: f(x) \geq f(x^*) \quad \forall x \in U_\varepsilon(x^*) \cap X$

6.5 Def (Feasible Direction): Let $x^* \in X$ be feasible for NLP.

a) $S \in \mathbb{R}^n: \exists \bar{\lambda} > 0: x^* + \lambda s \in X \; \forall \lambda \in [0, \bar{\lambda}]$ — feasible direction in x^*

b) $Z(x^*) = \{ s \in \mathbb{R}^n : s \text{ feas. dir. in } x^* \}$ — tangent cone or cone of feas. dir. in x^*

6.6 Rem (Feasible directions for NLPs):

For NLPs, i.e., $X = P(A, b)$, it holds:

$Z(x^*) = P(A_{eq(x^*)}, 0) = \{ A_{eq(x^*)} \cdot s \leq 0 \}$.

6.7 Thm (Necessary Optimality Condition): For NLP with $f \in C^1$:

$x^* \in \operatorname{Argmin}_{x \in X} f(x) \Rightarrow x^* \in X \wedge \nabla f(x^*)^T s \geq 0 \quad \forall s \in Z(x^*)$.

Proof: Let s be a feasible direction in x^*

$\Rightarrow \exists \bar{\lambda} > 0 : x^* + \lambda s \in X, \; f(x^* + \lambda s) \geq f(x^*) \; \forall \lambda \in [0, \bar{\lambda}]$

$\Rightarrow 0 \leq \frac{f(x^* + \lambda s) - f(x^*)}{\lambda} \xrightarrow[\lambda \searrow 0]{} D_s^+ f(x^*) \overset{\text{analy.}}{=} \nabla f(x^*)^T s$

$\Rightarrow \nabla f(x^*)^T s \geq 0 \quad \forall$ feas. dir. s in x^*

$\Rightarrow \nabla f(x^*)^T s \geq 0 \quad \forall s \in Z(x^*)$. □

6.8 Def (Active constraints): Consider (NLPI), $X(g) := \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \dots, m\}$, $x^* \in X(g)$.

a) $g_i(x^*) = 0 :\Leftrightarrow g_i(x) \leq 0$ {active / binding} at x^*

b) $eq(\{x^*\}) =: eq(x^*) = \{i \in [m] : g_i(x) \leq 0 \text{ active at } x^*\}$

equivalent / index set of active constr. at x^*

6.9 Ex (Active constraints):

(NLPI) min $f(x)$

$g_1(x) = x_1^2 + x_2^2 - 1 \leq 0 \quad (1)$

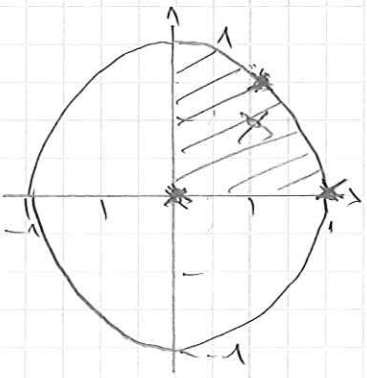
$g_2(x) = -x_1 \leq 0 \quad (2)$

$g_3(x) = -x_2 \leq 0 \quad (3)$

$\Leftrightarrow$ min $f(x)$

$x_1^2 + x_2^2 \leq 1$

$x_1, x_2 \geq 0$



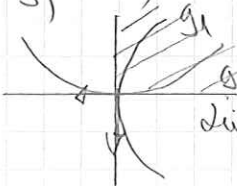
$eq(1,0) = \{1, 3\}$ $\quad eq(\frac{3}{5}, \frac{4}{5}) = \{1\}$

$eq(0,0) = \{2, 3\}$ $\quad eq(\frac{1}{2}, \frac{1}{2}) = \emptyset$

6.10 Def (Linearization cone): For NLPI, let $x^* \in X(g) = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \dots, m\}$. Then

$\mathcal{L}_{X(g)}(x^*) = \{s \in \mathbb{R}^n : \nabla g_i(x^*)^T s \leq 0 \; \forall i \in eq(x^*)\}$

linearization cone of $X(g)$ at x^*



6.11 Lem (tangent and linearization cone for linear constraints): Consider NPI, i.e., $X(g) = P(A,b)$, $x^* \in X(g) = P(A,b)$. Then

$$Z(x^*) = \mathcal{L}_{P(A,b)}(x^*) = P(A_{eq(x^*)}, 0).$$

Proof: $\nabla g(x^*) = A \Rightarrow \mathcal{L}_{P(A,b)}(x^*) = P(A_{eq(x^*)}, 0)$.

Claim: $Z(x^*) = \mathcal{L}_{P(A,b)}(x^*)$.

Pf.:

"$\subseteq$" let s be a feas. dir. in x^*

$\Rightarrow \exists \bar{\lambda} > 0 : x^* + \lambda s \in X \quad \forall \lambda \in [0, \bar{\lambda}]$

$\Rightarrow$ —"— $A(x^* + \lambda s) \leq b \quad \forall \lambda \in [0, \bar{\lambda}]$

$\Rightarrow$ —"— $A_{eq(x^*)} x^* + \lambda A_{eq(x^*)} s \leq b \quad \forall \lambda \in [0, \bar{\lambda}]$

$\Rightarrow A_{eq(x^*)} \cdot s \leq 0 \quad \forall s$ feas dir. in x^*

$\overset{\text{closure}}{\Rightarrow} A_{eq(x^*)} \cdot s \leq 0 \quad \forall s \in Z(x^*)$.

"$\supseteq$" let $s \in \mathcal{L}_{P(A,b)}(x^*) = \{ A_{eq(x^*)} s \leq 0 \}$

$\Rightarrow A_{eq(x^*)} \cdot (x^* + \lambda s) = \underbrace{A_{eq(x^*)} x^*}_{=b} + \underbrace{\lambda}_{\geq 0} \underbrace{A_{eq(x^*)} s}_{\leq 0} \leq b$

$A_{\overline{eq(x^*)}} (x^* + \lambda s) = \underbrace{A_{\overline{eq(x^*)}} x^*}_{< b_{\overline{eq(x^*)}}} + \lambda A_{eq(x^*)} s \leq b$ $\quad \forall \lambda$ small enough.

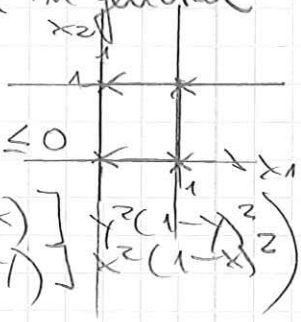
6.12 Rem (tangent and linearization cone for NLPs)

a) $Z(x^*) \subseteq \mathcal{L}_{X(g)}(x^*)$ always holds

b) $Z(x^*) \supseteq \mathcal{L}_{X(g)}(x^*)$ does not hold in general

$\min f(x,y)$

$g(x,y) = x^2(1-x)^2 \cdot y^2(1-y)^2 \leq 0$



$$\nabla g(x,y) = \begin{pmatrix} [2x(1-x)^2 - 2x^2(1-x)]\, y^2(1-y)^2 \\ [2y(1-y)^2 - 2y^2(1-y)]\, x^2(1-x)^2 \end{pmatrix}$$

$\Rightarrow \nabla g(0,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\Rightarrow \mathcal{L}_{X(g)}(0,0) = \mathbb{R}^2 \neq Z(0,0) = \text{aff}(1,0) \cup \text{aff}(0,1)$.

$\leadsto Z(x^*) \supseteq \mathcal{L}_{X(g)}(x^*)$ "constraint qualification"

$\rightarrow$ Opt III.

6.13 Prop. (Tangent and linearization cone):

For NLPI and $x^* \in X(g)$

$$Z(x^*) \subseteq \mathcal{L}_{X(g)}(x^*).$$

Proof: Let s be a feasible direction in x^*.

$\Rightarrow \exists \bar{\lambda} > 0 : \; x^* + \lambda s \in X(g) \quad \forall \lambda \in [0, \bar{\lambda}]$

$\Rightarrow \dfrac{g_i(x^* + \lambda s) - \overbrace{g_i(x^*)}^{=0}}{\lambda} \leq 0 \quad \forall \lambda \in [0, \bar{\lambda}], \; i \in eq(x^*)$

$\Rightarrow D_s^+ g_i(x^*) = \nabla g_i(x^*)^T s \leq 0 \quad \forall i \in eq(x^*)$

closure $\Rightarrow \nabla g_i(x^*)^T s \leq 0 \quad \forall s \in Z(x^*), \; i \in eq(x^*)$ ☐

6.14 Cor (Necessary optimality condition for linear constraints): For NLPI holds:

$$x^* \in \operatorname{Argmin} f(x) \;\Rightarrow\; Ax^* \leq b, \quad \nabla f(x^*)^T s \geq 0 \;\; \forall s : A_{eq(x^*)} s \leq 0$$
$$x \in P(A,b)$$

Proof: Rem 6.6, Thm 6.7. ☐

6.15 Rem (Idea of Karush-Kuhn-Tucker conditions):

$$x^* \in \operatorname*{Argmin}_{Ax \leq b} f(x) \;\overset{(*)}{\Rightarrow}\; Ax^* \leq b, \quad \nabla f(x^*)^T s \geq 0 \;\; \forall s : A_{eq(x^*)} s \leq 0$$

$\Leftrightarrow \min \nabla f(x^*)^T s \geq 0$ (direction gradient ≥ 0 f.a. feas. directions)
$\quad -A_{eq(x^*)} s \geq 0$

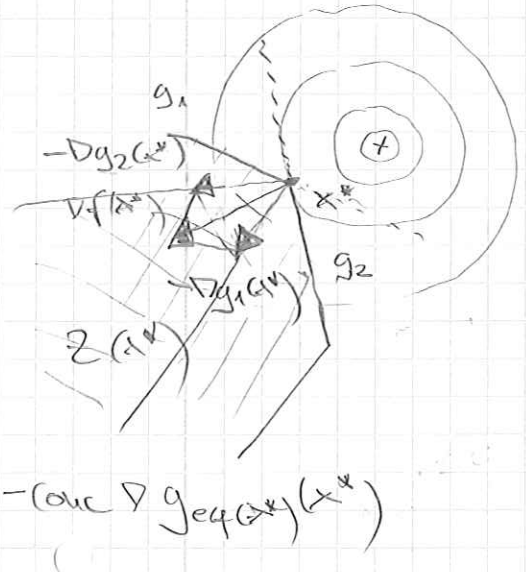
Duality $\Leftrightarrow \max\; 0^T \lambda$

Thm

$$-\lambda^T A_{eq(x^*)} = \nabla f(x^*)^T$$
$$\lambda \geq 0$$
$$A_{eq(x^*)} x^* = b_{eq(x^*)}$$
$$A_{\overline{eq(x^*)}} x^* < b_{\overline{eq(x^*)}}$$

$\Leftrightarrow \exists \lambda \in \mathbb{R}^{eq(x^*)}_{\geq 0}$

$$-\lambda^T A_{eq(x^*)} = \nabla f(x^*)$$
$$A_{eq(x^*)} x^* = b_{eq(x^*)}$$
$$A_{\overline{eq(x^*)}} x^* < b_{eq(x^*)}$$



6.16 Thm (Karush-Kuhn-Tucker conditions): Let
$f \in C^1 \Rightarrow f: \mathbb{R}^n \to \mathbb{R}$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\bar{x}$ a local minimizer of

$$\min f(x) \quad | \quad Ax \leq b.$$

Then there exists a vector $\bar{u} \in \mathbb{R}^m_{\geq 0}$ s.t.

(KKT) a) $\bar{u}^T A + \nabla f(\bar{x})^T = 0^T$ multiplier rule

KKT b) $A\bar{x} \leq b$ primal } feas.

system c) $\bar{u} \geq 0$ dual

d) $\bar{u}^T(A\bar{x} - b) = 0$ complementarity

Speech: $\bar{u}$ (vector of) Lagrange multipliers

$\bar{x}$ Karush-Kuhn-Tucker point (KKT point)

Proof: Let $I := eq(\bar{x})$, $J := \overline{eq(\bar{x})}$. Then by Cor. 6.14

$-A_{I\cdot} s \geq 0$, $\nabla f(\bar{x})^T s < 0$ infeasible

Farkas $\Rightarrow \exists z \geq 0 : -z^T A_{I\cdot} = \nabla f(\bar{x})^T$

Lemma

Let $\bar{u}_I := z$, $\bar{u}_J := 0$

$\Rightarrow \bar{u} \geq 0$

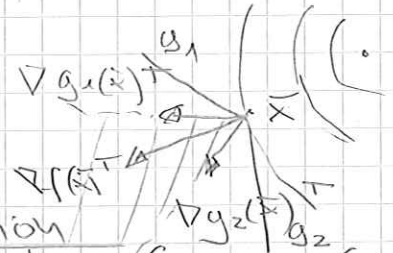
$\bar{u}^T A = z^T A_{I\cdot} = -\nabla f(\bar{x})^T \Rightarrow \bar{u}^T A_{I\cdot} + \nabla f(\bar{x})^T = 0$

$\bar{u}^T(A\bar{x} - b) = z^T \underbrace{(A_{I\cdot}\bar{x} - b_I)}_{=0} + 0^T \underbrace{(A_{J\cdot}\bar{x} - b_J)}_{<0} = 0$ □

6.17 Rem (geometric interpretation of KKT conditions):

$\bar{u}^T A = -\nabla f(\bar{x})^T$, $u \geq 0$

$\Leftrightarrow -\nabla f(\bar{x})^T \in \operatorname{cone} A_{eq(\bar{x})}^T$



condition

6.18 Thm (Sufficient KKT in the convex case): Let $f \in C^1 \Rightarrow f: \mathbb{R}^n \to \mathbb{R}$ be convex in NPI. Then:

$\bar{x}, \bar{u}$ satisfy (KKT) $\Leftrightarrow$ $\bar{x}$ (local = global) minimizer of NPI

Proof: $f(x) - f(\bar{x}) \overset{f \text{ convex}}{\geq} \nabla f(\bar{x})^T (x - \bar{x})$

$= -\bar{u}^T A (x - \bar{x})$

$= -\bar{u}^T [(Ax - b) - (A\bar{x} - b)]$

≤ 0 compl.

$= -\bar{u}^T (A\bar{x} - b) = 0$. □

6.19 Rem: Thm. 6.18 can be generalized to DLTI for $f, g \in \mathcal{C}^1$, see optimization III.

6.20 Thm (KKT conditions for inequality and equality constraints): Let $\mathcal{C}^1 \ni f: \mathbb{R}^n \to \mathbb{R}$, $A \in \mathbb{R}^{m \times n}$, $D \in \mathbb{R}^{k \times n}$, $b \in \mathbb{R}^m$, $d \in \mathbb{R}^k$, $\bar{x}$ local minimizer of

(NPFE) $\min f(x)$, $Ax \leq b$, $Dx = d$.

Then there exist vectors $\bar{u} \in \mathbb{R}^m_{\geq 0}$, $\bar{v} \in \mathbb{R}^k$ s.t.

$$\begin{aligned} \bar{u}^T A + \bar{v}^T D + \nabla f(\bar{x})^T &= 0 \\ A\bar{x} &\leq b \\ D\bar{x} &= d \\ \bar{u} &\geq 0 \\ \bar{u}^T(A\bar{x} - b) + \bar{v}^T(D\bar{x} - d) &= 0 \end{aligned}$$

Proof: $Dx = d \Leftrightarrow Dx \leq d$, $-Dx \leq -d$.

$\overset{\text{Thm. 6.16}}{\Rightarrow}$ $\exists\, \bar{u}, \bar{v}^+, \bar{v}^-$ s.t.

$$\begin{aligned} \bar{u}^T A + (\bar{v}^+ - \bar{v}^-)^T D + \nabla f(\bar{x}) &= 0 \\ A\bar{x} &\leq b \\ D\bar{x} &\leq d \\ -D\bar{x} &\leq -d \\ \bar{u}, \bar{v}^+, \bar{v}^- &\geq 0 \\ \bar{u}^T(A\bar{x} - b) + (\bar{v}^+ - \bar{v}^-)^T(D\bar{x} - d) &= 0 \end{aligned}$$

Let $\bar{v} := \bar{v}^+ - \bar{v}^-$. □

6.20 ex (KKT conditions): see page 79, 80