Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Universität Zürich[UZH]

Mathematical Optimization, Autumn Semester 2017

Summary of the lectures in 2017

Vanessa Leite

Github (git/svn) repository page:
https://github.com/ssinhaleite/eth-mathematical-optimization-summary
Contact vrcleite@gmail.com if you have any questions.

Saturday 4th November, 2017

Professor:
Prof. Dr. Robert Weismantel
D-MATH, ETH, UZH

Summary of the lectures about Mathematical Optimization - Prof. Dr. Robert Weismantel.
You can find a referece list in the end of each chapter.

—

Contents

Introduction

Complexity Theory - short revision

Definition: an algorithm is a finite list of instructions. The running time of an algorithm is the number of operations needed to complete it.

Simplifications

- We measure the "complexity" of an algorithm with respect to the length of the input, n. $T_A(n)$ = complexity of the algorithm = worst case for same input size = calculated/defined as the maximum over all inputs of length n.
- We are only interested in how $T_A(n)$ increases w.r.t. n.
  $T_A(n) = O(n)$: "linear"
  $T_A(n) = O(n^k)$: "polynomial"
  $T_A(n) = O(2^n)$: "exponential"
- We assume that every operation of the algorithm requires the same cost ("unit cost model").

Note We need to assume that the number of bits of any stored member must be bounded by a (constant) multiple of the number of bits of the largest input value. Example: if $a \in \mathbb{Z}$ is part of the input, its input size is $O(log|a|)$. An algorithm with running time of $O(a)$ is not polynomial-time! $(O(e^{loga})$: "pseudo polynomial time" $\rightarrow O((logm \times A_{ij}) \times m \times n)$.

Decision Problems

Let $\Sigma = \{0, 1\}$ and Σ^* be a finite binary sequences (sequences on Σ). We assume that all problems are encoded as elements of Σ^*.

Definition $\mathcal{P}$ *problems*: For a set $S \subseteq \Sigma^*$, the decision problem defined by S is the problem wether a given $x \in \Sigma^*$ is in S or not.

It is a polynomial-time solvable ($S \in \mathcal{P}$) if there exists an algorithm $\mathcal{A}$ s.t. $\forall x \in \Sigma^*, \mathcal{A}(x) = 1 \iff x \in S$.

If you can solve a decision problem, usually you can solve the optimization problem.

Definition $\mathcal{NP}$ *problems*: The decision problem defined by $S \subseteq \Sigma^*$ is in $\mathcal{NP}$ if there is a polynomial time algorithm $\mathcal{A} : \Sigma^* \times \Sigma^* \mapsto \{0, 1\}$ and a polynomial p, s.t.:

- $\forall x \in S, \exists y \in \Sigma^*$ s.t. $|y| \in O(p(|x|)$ and $\mathcal{A}(x, y) = 1$
- $\forall x \in S, \forall yin\Sigma^*, \mathcal{A}(x, y) = 0$

Example of $\mathcal{P}$ problems: Let $G = (V, E)$ be an undirected graph. "does G have a cycle?"; "does G have a perfect matching?". Example of $\mathcal{NP}$ problems: "does there exist a cycle covering all vertices (hamiltonian cycle)?". This problem is easy to check, given a cycle, if it covers all vertices, however, it is not that easy find the cycle.

NP-Hard vs NP-Complete A decision problem S is reducible to $S'(S \subseteq pS') \iff$ there is a function $f : \Sigma^* \mapsto \Sigma^*$ compatible in polynomial s.t. $x \in S \iff f(x) \in S'$.

Lemma Let S, S' be decision problems. If $S' \in \mathcal{P}$ (is polynomial-time solvable). Then, $S \subseteq pS' \Rightarrow S \in \mathcal{P}$.

Definition NP-Complete: A decision problem is called NP-complete if $S \in \mathcal{NP}$ and $\forall S' \in \mathcal{NP}, S' \subseteq pS$. If you have a solution for S, you could solve all other S problems in polynomial time.

Definition NP-Hard: A decision problem S is NP-Hard if $\forall S' \in \mathcal{NP}, S' \subseteq pS$.

Optimization problems in finite dimensional space

Definition: Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. $dom(f) = \{x \in \mathbb{R}^n \mid |f(x)| < \infty\}$.

We need a function to optimize. An optimization problem is of the form:

$$\text{min/max}_{\mathbf{x} \in \mathcal{F}} \quad f(\mathbf{x})$$

Where $\mathcal{F} \subseteq \mathbb{R}^n$ is the feasible domain given implicitly or by a "membership oracle". The function $\mathcal{F} : \mathbb{R}^n \leftarrow \mathbb{R}$ is presented implicitly (i.e., is a domain you can't describe) or by an "evaluation oracle", for instance, given by a query: $x \in \mathbb{R}^n$: true ($x \in \mathcal{F}$) or false ($x \notin \mathcal{F}$).

Implicitly means that you need evaluate by a given point; continuum spectrum of points can be scaled and produce a minimum nore negative than any natural.

What does $\min\limits_{x \in \mathcal{F}} f(x)$ mean for us? This problem can be assigned to one of three meanings:

- nothing: the problem is *infeasible*; it means that the domain is empty, i.e., $\mathcal{F} = \emptyset$, there is nothing to optimize.
- the problem is *unbounded*, i.e., there exists a sequence of points $x^1, x^2, \cdots \in \mathcal{F}$ and $\forall n \in \mathbb{N}, f(x^1) < -n \Rightarrow lim_{i \to \infty} f(x^i) = -\infty$. There exists a sequence of points that for each natural number, $f(x^i < -n$, so we cannot bound the sequence.
- the problem has an *optimal solution*, i.e., $\exists x^* \in \mathcal{F}$ such that $f(x^*) \leq f(z) \forall z \in \mathcal{F}$.

Target of the course Derive conclusions(meanings from above 1.2) for optimization problem $\min\limits_{x \in \mathcal{F}} f(x)$ and give ideally a proof that our conclusion is correct. It is necessary to proof (find a proof) to hold the conclusions. What to do when the se $\mathcal{F}$ is empty? It is necessary a mathematical theory.

Basic definitions

organize the enumeration of figures in the text

- A *half space* is a set of the form $\{x \in \mathbb{R}^n \mid a^T x \leq \alpha\}$ for $a \in \mathbb{R}^n, \alpha \in \mathbb{R}$.

  A *half space rational* is a set of the form $\{x \in \mathbb{R}^n \mid a^T x \leq \alpha\}$ for $a \in \mathbb{Q}^n, \alpha \in \mathbb{Q}$. In this case the boundary is a rational line.

- A *polyhedron* is the intersection of a **finite** number of halfspaces (Figure 1.3). The intersection of halfspaces gives a region of finite points. In this case, it is called *polytope*: a bounded polyhedron. We need to use geometry and algebric forms. In algebric form: $\mathcal{P} = \{x \in \mathbb{R}^n \mid Ax \leq b\} = \{x \in \mathbb{R}^n \mid a_1^T x \leq b_1, \ldots, a_n^T x \leq b_n\}$ where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. A is a matrix and $Ax \leq b$ represents each row of A.

  Rational polyhedron: $A \in \mathbb{Q}^{m \times n}, b \in \mathbb{Q}^m$.

  one half space is also a polyhedron.

  all points inside a polyhedron are called feasible.

  $Ax \leq b$ represents the multplication of the rows of A with x. This is less equal than a value in b.

  add a graphical representation of this subitem

- A set $\mathcal{Q}$ is *convex* (Figure 1.3) if $\forall x, y \in \mathcal{Q}$ and $\lambda \in (0, 1)$ then $\lambda x + (1 - \lambda)y \in \mathcal{Q}$.

- A function f is *linear* if $f(x) = c^T x$ for $c \in \mathbb{R}^n$.

- A function f is *convex* if $f : \mathbb{R}^n \to \mathbb{R} \Rightarrow \emptyset \neq dom(\mathcal{F})$ is a convex set and $f(\lambda x + (1 - \lambda)y) \leq \lambda(f(x)) + (1 - \lambda)(f(y)) \forall x, y \in dom(\mathcal{F})$ and $\forall \lambda \in (0, 1)$.

- A function f is *strictly convex* if $f(\lambda x + (1 - \lambda)y) < \lambda(f(x)) + (1 - \lambda)(f(y)) \forall x, y \in dom(\mathcal{F})$ and $\forall \lambda \in (0, 1)$.

- lef f be continuously differentiable, then f is convex if, and only if, $f(y) \geq f(x) + \Delta f(x)^T (yx) \forall y, x \in dom(f)$. And strictly convex if, and only if, $f(y) > f(x) + \Delta f(x)^T (yx) \forall y, x \in dom(f)$ and $y \neq x$.

- If f is twice differentiable, then f is convex if, and only if, $\Delta^2 f(x) \geq 0, \forall x \in dom(f)$ (positive semidefinite).

Convex set Given vectors $(x^1, \ldots, x^t \in \mathbb{R}^n)$ the convex hull of this vectors are defined as $conv(x^1, \ldots, x^t) := \{x \in \mathbb{R}^n \mid \exists \lambda_1, \ldots, \lambda_t \geq 0\}$ such that $\mathbb{1}^T \lambda = 1$ (or $\sum_{i=1}^t \lambda_i = 1$): $x = \sum_{i=1}^t \lambda_i x^i$.

Note $conv(x^1, \ldots, x^t) = \{x \in \mathbb{R}^n \mid x = \sum_{i=1}^t \lambda_i x^i, \lambda \geq 0, \sum \lambda = 1\}$ is a special case of a polyhedron (bounded polyhedron) and every bounded polyhedron is of this form.

Figure 1: A polyhedron/polytope.

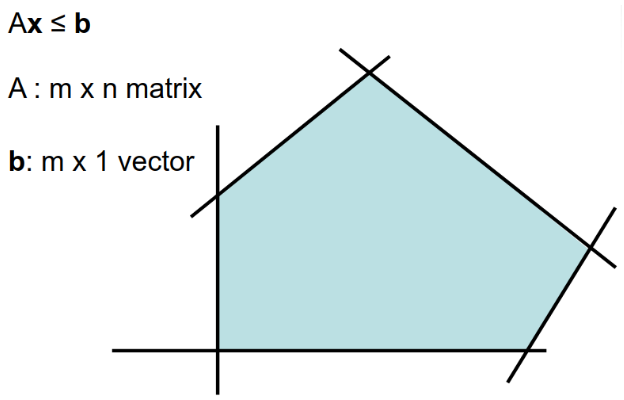
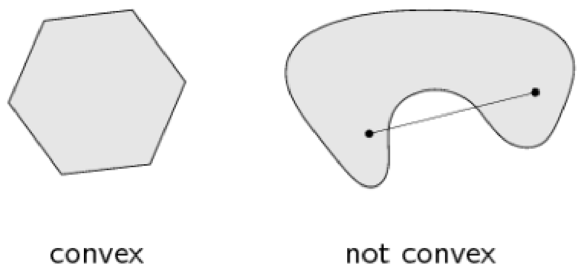


Figure 2: Convex and non-convex set.



Lemma

1. a polyhedron is a convex set
2. $conv(x^1, \ldots, x^t)$ is a convex set

Proof of 1 Let $\mathcal{P} = \{x \in \mathbb{R}^n \mid Ax \leq b\}$. Let $x, y \in \mathcal{P}, \lambda \in (0, 1)$. Show $\lambda x + (1 - \lambda)y \in \mathcal{P}$.

$$A[\lambda x + (1 - \lambda)y] = \lambda Ax + (1 - \lambda)Ay$$
$$\lambda Ax \leq \lambda b \text{ and } (1 - \lambda)Ay \leq (1 - \lambda)b$$
$$\lambda Ax + (1 - \lambda)Ay \leq \lambda b + (1 - \lambda)b = b$$
$$\Rightarrow \lambda x + (1 - \lambda)y \in \mathcal{P}.$$

Proof of 1 Let $x^1, \ldots, x^t \in \mathbb{R}^n$. Let $x, y \in conv(x^1, \ldots, x^t)$ and $0 < \lambda < 1$. Show $\lambda x + (1 - \lambda)y \in conv(x^1, \ldots, x^t)$.
x lies in $conv(x^1, \ldots, x^t)$. What does that means? That implies $x = \sum_{i=1}^{t} \mu_i x^i, \mu_i \geq 0 \forall i$ and $y = \sum_{i=1}^{t} \sigma_i x^i, \sigma_i \geq 0 \forall i$

$$
\begin{aligned}
\lambda x + (1-\lambda)y &= \sum_{i=1}^{t} \lambda\mu_i x^i + \sum_{i=1}^{t} (1-\lambda)\sigma_i x^i \\
&= \sum_{i=1}^{t} (\lambda\mu_i + (1-\lambda)\sigma_i) x^i \\
\tau_i &= (\lambda\mu_i + (1-\lambda)\sigma_i), \tau_i \geq 0, \forall i \\
&= \sum_{i=1}^{t} \tau_i = \sum_{i=1}^{t} (\lambda\mu_i + (1-\lambda)\sigma_i) \\
&= \lambda \sum_{i=1}^{t} \mu_i + (1-\lambda) \sum_{i=1}^{t} \sigma_i \\
\sum_{i=1}^{t} \mu_i &= 1 \text{ and } \sum_{i=1}^{t} \sigma_i = 1 \\
&= \lambda + (1-\lambda) = 1
\end{aligned}
$$

Classification of Optimization Problems

Definition of optimization problem: For $f : \mathbb{R}^n \mapsto \mathbb{R}$ (any function), solve $\min_{x \in \mathcal{F}} f(x)$ (objective value) or $\operatorname{argmin}_{x \in \mathcal{F}} f(x)$ (value of x that minimize $f(x)$).

Note $\max_{x \in \mathcal{F}} f(x) = \min_{x \in \mathcal{F}} -f(x)$

- *linear optimization problem*: given $A \in \mathcal{Q}^{m \times n}, b \in \mathcal{Q}^m, c \in \mathcal{Q}^n$. $\min_{x \in \mathcal{P}} c^T x$ where $\mathcal{P}$ is a polyhedron defined by $\{x \in \mathcal{R}^n \mid x_i \geq 0 \forall i, Ax = b\}$ or $\max_{x \in \mathcal{P}} c^T x$ where $\mathcal{P} = \{x \in \mathcal{R}^n \mid Ax \leq b\}$
  $c^T x$ is the objective (linear) function.
  $Ax \leq b$ is the constraint/feasible region.
  Can be solved quickly and in polinomial time.

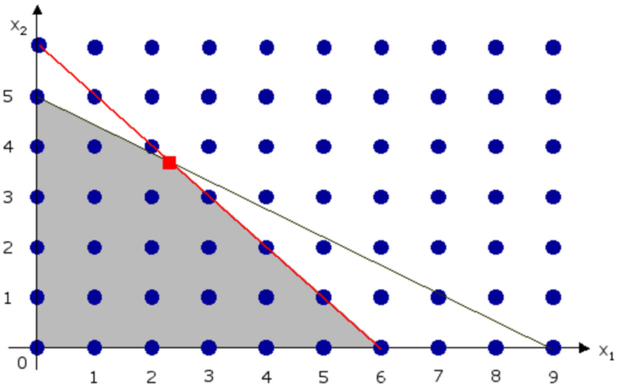
  integer linear optimization problem (Figure 1.4). $max/min\{c^T x \mid Ax \leq b, x \in \mathbb{Z}^n\}$. Solve a integer optimization problem can lead to a better result than just round linear solutions.

  mixed integer linear optimization problem $max/min\{c^T x \mid Ax \leq b, x \in \mathbb{Z}^{n-d} \times \mathbb{R}^d\}$.

- *convex optimization problem*: $\mathcal{F}$ is a convex set, $f : \mathcal{R}^n \to \mathcal{R}$ is a convex function and $\mathcal{F} \subseteq dom(f)$
  $f(x)$ convex and $\mathcal{F}$ convex set. $\forall x, y \in \mathbb{R}^n, \lambda \in [0,1], f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$.
  $\forall x, y \in \mathcal{F}, \lambda x + (1-\lambda)y \in \mathcal{F}$.

  Example of implicitly: knapsack problem: collection are all feasible solutions of knapsack problem.

Figure 3: ILP



- *combinatorial optimization problem*: given finite groundset $\mathcal{E}$ and a collection $\mathcal{I}$ of subsets of $\mathcal{E}$ (tipically, implicitly given: you cannot write it down). For $c : \mathcal{E} \mapsto \mathcal{R}$ find a member in the collection $(I \in \mathcal{I})$ such that $\sum_{i \in I} c_i$ is minimal/maximal. Example: $G = (V, E), E \subseteq V \times V$ $S \subseteq V$ stable implies that for all pairs in $S(\forall i, j \in S)$, the corresponding edge i, j is not present $(i \neq j, (i, j) \notin E)$. Groundset: V
  $\mathcal{I} \subseteq 2^V$
  $I \in \mathcal{I} \iff I$ is stable in G (implicitly given)
  $c : V \mapsto \mathcal{R}$; suppose $c(v) = 1, \forall c \in V$
  find a maximal (w.r.t. cardinality) stable set in G

- *general integer optimization problem*:
  $\max\limits_{x \in \mathcal{P} \cap \mathbb{Z}^n} c^T x, \mathcal{P} = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ or $\min\limits_{x \in \mathcal{P} \cap \mathbb{Z}^n} c^T x, \mathcal{P} = \{x \in \mathbb{R}^n \mid x_i \geq 0 \forall i, Ax = b\}$

Linear Optimization and Extreme Points I

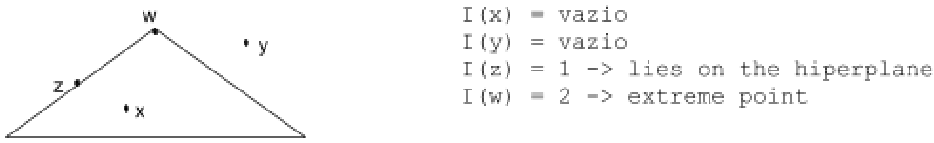
Geometry can be misleading. We try to get some intuitions from geometry but what we want to do is came up with the algebra.

Consider $\max\limits_{s.t. x \in \mathcal{P}} c^T x, P = \{x \in \mathcal{R}^n \mid Ax \leq b\}, A \in \mathbb{Q}^{m \times n}, b \in \mathbb{Q}^m, c \in \mathbb{Q}^n$

From linear algebra we know:
$A \in \mathbb{Q}^{m \times n}$
$n = dim(Ker(A)) + dim(Im(A))$, where $Ker(A)$ is the kernel space and $Im(A)$ is the image space.
$Ker(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$
$Im(A) = \{x \in \mathbb{R}^n \mid x^T = y^T A$, for some $y \in \mathbb{R}^n\}$
$Ker(A) \perp Im(A)$ (kernel space is orthogonal/perpendicular to image space).
$\forall z \in Im(A)$ and $\forall x \in Ker(A)$ (where z and x are vectors), $z^T x = y^T Ax = 0$ (the dot product is zero).

definition of vector space

Definition: Let $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ be a polyhedron.



1. $x \in P$ is an extreme point if there is no representation of the form $x = \lambda y + (1-\lambda)z$ for $y, z \in P$, $y \neq z$ and $\lambda(0,1)$.
2. $x \in P$ is an vertex if there exists $c \in \mathbb{R}^n$ such that $c^T x > c^T y$, $\forall y \in P \setminus \{x\}$
3. For a point $x \in \mathbb{R}^n$, the index set has tight constraints: $I(x) = \{i \in \{1, ..., m\} \mid A_{i\cdot}x = b_i\}$
4. x is a basic solution if $dim(\{A_{i\cdot} \mid i \in I(x)\} = n$, i.e., the vectors are linearly independent. For a basic feaseble solution, $x \in P$.

Theorem: let $P = \{x \in \mathbb{R}^n \mid Ax \leq b\} \neq \emptyset$. Let $x^* \in P$. The following statements are equivalent:

1. x^* is a vertex
2. x^* is an extreme point
3. x^* is a basic feasible solution

Proof (1 $\rightarrow$ 2): Take a vertex x^*. That means that exist some vector $c \in \mathbb{R}^n$ such that $c^T x^* > c^T x$, $\forall x \in P \setminus \{x^*\}$. Suppose x^* is not an extreme point. $\exists y, z \in P$, $y \neq z$ and $\lambda \in (0,1)$ such that $x^* = \lambda y + (1-\lambda)z$

$$\begin{gathered}
c^T x^* = c^T(\lambda y + (1-\lambda)z) \\
= \lambda c^T y + (1-\lambda)c^T z \\
< \lambda c^T x^* + (1-\lambda)c^T x^* = c^T x^* \\
c^T x^* < c^T x^* \rightarrow contradiction!
\end{gathered}$$

Proof (2 $\rightarrow$ 3) or ($\neg$3 $\rightarrow$ $\neg$2): Let $x^* \in P$ (so, it is feasible), assume x^* is not a basic solution, i.e, $dim(\{A_{i\cdot} \mid i \in I(x^*)\}) < n$. Linear algebra tells us $\exists z \in \mathbb{R}^{n \setminus \{0\}}$ such that $A_{i\cdot}z = 0$, $\forall i \in I(x^*)$.

$$\begin{gathered}
A_{i\cdot}x^* = b_i, \forall i \in I(x^*) \\
A_{i\cdot}x^* < b_i, \forall i \notin I(x^*)
\end{gathered}$$

Let

$$\varepsilon = \begin{cases} 1, & \text{if } A_{i\cdot}z = 0, \forall i \notin I(x^*). \\ \min\{\frac{b_i - A_{i\cdot}x^*}{|A_{i\cdot}z|} \mid i \notin I(x^*) \text{ such that } A_{i\cdot}z \neq 0 & \text{otherwise.} \end{cases}$$

Note $\varepsilon > 0$.

Claim: $y^+ = x^* + \varepsilon z \in P$ and $y^- = x^* - \varepsilon z \in P$. Result follows then, because $x^* = \frac{1}{2}y^+ + \frac{1}{2}y^-$.

Proof of the claim: wlog, $y^+ \in P$. $\forall i \in I(x^*) : A_{i\cdot}y^+ = A_{i\cdot}x^* + \varepsilon \underbrace{A_{i\cdot}z}_{=0} = b_i$

- $\forall i \notin I(x^*)$ such that $A_{i\cdot}z \leq 0$: $A_{i\cdot}y^+ = A_{i\cdot}x^* + \underbrace{\overbrace{\varepsilon}^{>0} A_{i\cdot}z}_{\leq 0} < b_i$
- $\forall i \notin I(x^*)$ such that $A_{i\cdot}z > 0$: Then $\varepsilon \leq \frac{b_i - A_{i\cdot}x^+}{A_{i\cdot}z}$ and hence, $A_{i\cdot}y^+ = A_{i\cdot}x^* + \varepsilon A_{i\cdot}z \leq A_{i\cdot}x^* + \frac{b_i - A_{i\cdot}x^*}{A_{i\cdot}z} A_{i\cdot}z = b_i$

do the same proof for y^-

Proof (3 → 1): $x^* \in P$ is a basic feasible solution. $dim(\{A_{i\cdot} \mid i \in I(x^*)\}) = n$

By linear algebra: $\{z \in \mathbb{R}^n \mid A_{i\cdot}z = 0, \forall i \in I(x^*)\} = \{0\}$ (*)

Define $c^T = \sum_{i \in I(x^*)} A_{i\cdot} \in \mathbb{R}^n$

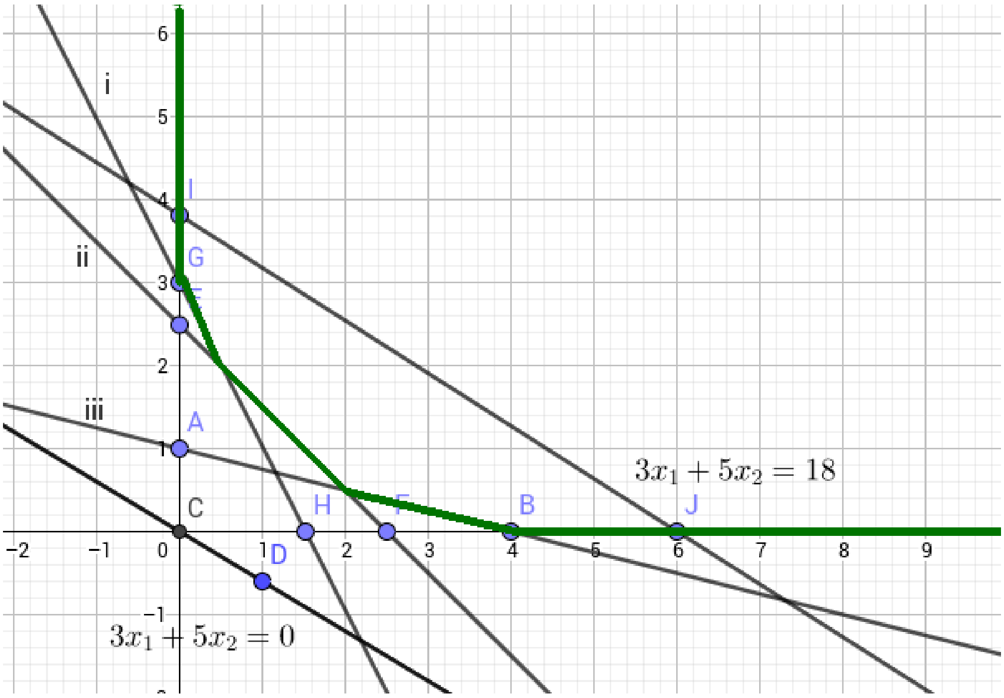
$c^T x^* = \sum_{i \in I(x^*)} A_{i\cdot}x^* = \sum_{i \in I(x^*)} b_i \underbrace{\geq}_{\forall x \in P} \sum_{i \in I(x^*)} A_{i\cdot}x = c^T x$

When is $c^T x^* = c^T x$? $A_{i\cdot}x = A_{i\cdot}x^*$, $\forall i \in I(x^*)$ $A_{i\cdot}(x - x^*) = 0$, $\forall i \in I(x^*) \rightarrow$ (*) $x = x^*$

Corollary: The number of vertices (extreme points or basic feasible solutions) in a Polyhedron P is finite. **Proof:** $A \in Q^{mxn}$. The number of basic feasible solution is smaller than m^n, and it is a finite number.

Why are extreme points interesting for linear optimization? Consider $\min 3x_1 + 5x_2$

Figure 4: unbounded polytope



$$P = \left\{ x \in \mathbb{R}^2 \mid \begin{array}{c} 2x_1 + x_2 \geq 3 \text{ (i)} \\ 2x_1 + 2x_2 \geq 5 \text{ (ii)} \\ x_1 + 4x_2 \geq 4 \text{ (iii)} \\ x_1, x_2 \geq 0 \end{array} \right\}$$

In the 2D case we can "guess" the right solution. If someone claims that the solution of $\min 3x_1+5x_2 = 0$, we can prove it is not by summing up the constraints. From (i) + (iii), we know $\min 3x_1 + 5x_2 \geq 7$. We can find the solution pushing the "lines" until the equation leaves the feasible solution area (marked in green in Figure 2). Geometrically, $(2, \frac{1}{2})$ is the optimal solution and $(3,5)^T \begin{pmatrix} 2 \\ \frac{1}{2} \end{pmatrix} = 8.5$

Constraint (ii) times 1.5 $\rightarrow 3x_1 + 3x_2 \geq 7.5$ and $2x_2 \geq 0 \rightarrow 3x_1 + 5x_2 \geq 7.5$. Take (ii) times $\frac{7}{6}$ and add (iii) multiplied with $\frac{2}{3}$: $3x_1 + 5x_2 \geq 8.5$

More generally, $\min ax_1 + bx_2$:

- Suppose $a < 0$. Take the sequence of points $(k, 0)$ for $k \in \mathbb{N}$, $k \geq 4$

  problem is unbounded

- Suppose $b < 0$. Take the sequence of points $(0, k)$ for $k \in \mathbb{N}$, $k \geq 3$

  problem is unbounded

Linear Optimization and Extreme Points II

Definition: A polyhedron P contains a line if there exists $d \in \mathbb{R}^n \setminus \{0\}$ and $y \in P$ such that $y + \lambda d \in P \forall \lambda \in \mathbb{R}$

Lemma: Let $P = \{x \in \mathbb{R}^n \mid Ax \leq b\} \neq 0$, $d \in \mathbb{R}^n \setminus \{0\}$ is a line $\leftrightarrow Ad = 0$

Proof: If $d \in \mathbb{R}^n \setminus \{0\}$ and $Ad = 0$, then take any $y \in P$, i.e, $Ay \leq b$. Then $\forall \lambda \in \mathbb{R}$, $A(y + \lambda d) = Ay + \lambda \underbrace{\underbrace{Ad}_{=0}}_{=0} \leq b$. Conversely, suppose $d \in \mathbb{R}^n \setminus \{0\}$ is a line in P. Then $\exists y \in P$ such that $y + \lambda d \in P$, $\forall \lambda \in \mathbb{R}$. Suppose, $\exists i \in \{1, \ldots, m\}$ such that $A_i d \neq 0$ wlog: $A_i d > 0$. Let $\lambda^* = \frac{b_i - A_i y}{A_i d}$, for $\lambda > \lambda^*$, then $A_i(y + \lambda d) = A_i y + \lambda A_i d > A_i y + \lambda^* A_i d = A_i y + \frac{b_i - A_i y}{A_i d} A_i d = b_i$

Observation: Suppose I give you a polyhedron non empty ($P = \{x \in \mathbb{R}^n \mid Ax \leq b\} \neq 0$ and $m < n$ then P contains a line!

Theorem: Let $P = \{x \in \mathbb{R}^n \mid Ax \leq b\} \neq 0$. P has an extreme point iff P does not contain a line.

Proof: "$\leftarrow$" P does not contain a line then P has an extreme point Let $x \in P$ such that $l = dim(\{A_{i\cdot} \mid i \in I(x)\})$. If $l = n$ (maximum), then x is a basic feasible solution, thus, an extreme point. Suppose $l < n$. Then from linear algebra, there is a vector $d \in \mathbb{R}^n \setminus \{0\}$ such that $A_{i\cdot}.d = 0$, $\forall i \in I(x)$. d is not a line $\rightarrow \exists j \notin I(x)$ such that $A_{j\cdot}.d \neq 0$ wlog: $A_{j\cdot}.d \geq 0$. Let $J = \{j \notin I(x)\} \mid A_{j\cdot}.d > 0\} \neq 0$. Notice that for $j \in J$, $A_{j\cdot}$ is linearly independent from $\{A_{i\cdot} \mid i \in I(x)\}$

Linearly independent: Suppose $A_j = \sum_{i \in I(x)} A_i \lambda_i$, then $\underbrace{A_{j\cdot}.d}_{=0} = \sum_{i \in I(x)} \lambda_i A_{i\cdot}.d = 0$. So, it is linearly independent.

Let, $\lambda^* = \min\{\frac{b_j - A_{j\cdot}}{A_{j\cdot}.d} \mid j \in J\}$. Then $x + \lambda^* d \in P$. Moreover, we observe $A_{i\cdot}.(x + \lambda^* d) = b_i$, $\forall i \in I(x)$. Let $j \in J$ such that $\underbrace{\lambda^* = \frac{b_j - A_{j\cdot}.x}{A_{j\cdot}.d}}_{\text{the minimum}}$, then $A_{j\cdot}.(x + \lambda^* d) = A_{j\cdot}.x + \frac{b_j - A_{j\cdot}.x}{A_{j\cdot}.d} A_{j\cdot}.d = b_j$.

$A_{j\cdot}$ is linearly independent from $\{A_i \mid i \in I(x)\} \rightarrow dim(\{A_{k\cdot} \mid k \in I(x + \lambda^* d)\}) = l + 1$, so, this is a contradiction of l be maximal.

Prood: "$\rightarrow$" P has an extreme point then P does not contain a line Let x^* be an extreme point, then $dim(\{A_{i\cdot} \mid i \in I(x^*)\}) = n$. Then, $A_{i\cdot}.d = 0$, $\forall i \in I(x^*)$ implies $d = 0$. Therefore, $Ad = 0$ implies $d = 0$, by using the previous lemma, there is no line in P.

Lemma: optimization problem $\rightarrow$ maximize a linear function let $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ and $c \in \mathbb{R}^n$. Suppose that i) P has an extreme point and ii) there exists an optimal solution to $\max_{x \in P} c^T x$. Then, there exists an extreme point solution attaining the optimal solution.

Proof: Let $v^* = \max_{x \in P} c^T x =$ optimal value. Let $Q = \{x \in P \mid c^T x = v^* = \{x \in \mathbb{R}^n \mid Ax \leq b, c^T x \leq v^*, -c^T x \leq -v^*\}$. Q is a polyhedron and $Q \neq 0$ (there is an optimal solution). $Q \subset P$, since P has an extreme point. P has no line $\rightarrow$ Q has no line $\rightarrow$ Q has an extreme point x^*. Claim: x^* is an extreme point in P

Proof of the claim: Suppose $\exists y, z \in P$, $y \neq z$ and $\lambda \in (0,1)$ such that $x^* = \lambda y + (1-\lambda)$. $v^* = c^T x^* = \lambda c^T y + (1-\lambda) c^T z \leq \lambda c^T x^* + (1-\lambda) c^T x^* = c^T x^* \rightarrow \underbrace{c^T y = v^* \text{and } c^T z = v^*}_{\text{to have an equality in previous equation}} \rightarrow$ $y, z \in Q$, x^* is an extreme point in Q.

Theorem: Let $P \neq 0$ be a polyhedron not containing a line. Then, $\max_{x \in P} c^T x$ is either equal to $+\infty$ or there exists an extreme point in P attaining the optimal value.

Proof: We must show that if the optimal value is not infinite, then there exists an optimal solution. In fact, we proof a stronger claim.
Claim: if the max value is not infinite then, for every $x \in P$, there exists an extreme point, $w \in P$, such that $c^T x \leq c^T w$. From the claim, the statement follows: Let $\{w^1, \ldots, w^r\}$ be all extreme points in P, not empty, also r is finite. Let $\underbrace{w}_{\text{is the maximum}} \in \{w^1, \ldots, w^r\}$ allows $max\{c^T w^1, \ldots, c^T w^r\}$. Then, $\forall x \in P$, there exists w^i such that $c^T x \leq c^T w^i$, but by definition, $\leq c^T w$, then, w is an optimal solution.

Proof of the claim: Let $x^* \in P$ and $dim(\{A_{i\cdot} \mid i \in I(x^*)\}) = k < n$. If it were equals to n, we could use the point itself. $\exists d \in \mathbb{R}^n \setminus \{0\}$ such that $A_{i\cdot} d = 0$, $\forall i \in I(x^*)$. wlog: $c^T d > 0$.

- if $A_{j\cdot} \leq 0$, $\forall j \in I(x^*)$, then

  $x^* + \lambda d \in P$, $\forall \lambda \geq 0$. If max value is not infinite then, $c^T d = 0$. $\exists j \in I(x^*)$ such that $A_{j\cdot} d < 0$ (otherwise, d is a line). We observe $A_{j\cdot}$ is linearly independent from the constraints $\{A_{i\cdot} \mid i \in I(x^*)\}$. Let $\lambda^* = min\{\frac{b_j - A_{j\cdot} x^*}{|A_{j\cdot} d|} \mid A_{j\cdot} d < 0\}$. Then, following in the negative direction, $x^* - \lambda^* d \in P$ and $c^T(x^* - \lambda^* d) = c^T x^* - \lambda^* c^T d = c^T x^*$. Then, $dim(\{A_{i\cdot} \mid i \in I(x^* - \lambda^* d)\} = k+1$. And now we can iterate with $x^* - \lambda^* d$ in place of x^*.

- if there exists $j \notin I(x^*)$ such that $A_{j\cdot} d > 0$

  $A_{j\cdot}$ is linearly independent from $\{A_{i\cdot} \mid i \in I(x^*)\}$ and then $\lambda^* = min\{\frac{b_j - A_{i\cdot} x^*}{A_{i\cdot} d} \mid A_{j\cdot} d > 0\}$ and $x^* + \lambda^* d \in P$ and $dim(\{A_{i\cdot} \mid i \in I(x^*)\}) = k+1$ and we can iterate with $x^* + \lambda^* d$ in place of x^*.

Glossary

Term	Brief explanation of the term
$A_{i\cdot}$	row access on a matrix A.
$A_{\cdot j}$	column access on a matrix A.

TODO

This is the chapter on what still has to be improved in the summary. Please update this list by writing the todos directly into the summary where necessary by writing:

```
\todo{Something that still has to be done}
```

For a missing figure, please write:

```
\missingfigure{A missing figure}
```

Everyone that picks up the work on this summary will thank you for keeping this up-to-date properly.

Todo list