Lecture 1

Logistics: TAs, Office hours
grading, homework
1 midterm (in class), Final exam
Homeworks will be posted on Canvas
Deadlines for homeworks will be Friday 5 pm.

Note: This course is $\boxed{\text{MATH intense!!}}$ A solid knowledge of linear algebra, analysis, Probability and statistics is required.
⇒ undergraduates need permision

Suggested References:

Linear Programming → (first part of the class)
- Bertsimas, Tsitsiklis "Introduction to Linear optimiz..."
- R.J. Vanderbei "Linear Programming: Foundations and extension"

Convex Optimization → (second part of the class)
- Ben-Tal, Nemirovski "Lectures on Modern Convex Opt."
- Bertsekas, Nedic, Ozdaglar "Convex Analysis & optimization"
- Boyd, Vandenberghe "Convex optimization"

General formulation that we have in this course:

$$\text{minimize } f_0(x)$$
$$\text{Subject to } f_i(x) \leq b_i \quad i=1,\dots,m$$

→ $f_0: \mathbb{R}^n \to \mathbb{R}$: Our objective function.
→ $f_i: \mathbb{R}^n \to \mathbb{R}$ for $i=1,\dots,m$: constraint functions

Optimal solution x^* : $\begin{cases} x^* \text{ is feasible, i.e., } f_i(x^*) \leq b_i \ \forall i=1,\dots \\ f_0(x^*) \text{ is smaller than } f_0(\underline{x}) \text{ for any feasible } \dots \end{cases}$

In this class we focus on the case that f_0 and $f_1,\dots,f_m$ are convex
Later I'll formally define a convex function.
why should we care about this problem?
- It covers many important problems that appear in
  - ML → SVM, regression (Linear, logistic), ...
  - wireless Comm. → resourse allocation
  - Control Theory & Robotics → Multi-agent systems.
- There are severall "fast" methods for solving Convex opt.

Goal of this course is to ensure that you learn how to write the problem that you see in practice as an optimization problem. In particular, write it in one of the following forms (if possible)
- linear programming
- Semi-definite Programming
- Second-Order Cone pro...

⋮

Also, to learn how to solve this class of problems
→ Simplex, GD, Newton, ...
→ more sophisticated methods → next Semester.

Linear Programming:

It's special case of the general Convex Optimization where the objective function "f_0" and the constraint functions "f_i" $i=1,\dots,m$ are all linear functions. To be more precise,

$$\underset{\{x_1,\dots,x_n\}}{\text{minimize}} \quad \sum_{j=1}^{n} c_j x_j$$
$$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad \text{for } i=1,\dots,m$$
$$\sum_{j=1}^{n} d_{ij} x_j = e_i \quad \text{for } i=1,\dots,p$$

* $x_1,\dots,x_n$ are our decision (optimization) variables and they are real scalars, i.e., $x_1,\dots,x_n \in \mathbb{R}$.

* $c_j, a_{ij}, b_i, d_{ij}, e_i$ are all problem parameters.

→ $\sum_{j=1}^{n} c_j x_j$ is the cost function or objective func.

→ $\sum_{j=1}^{n} a_{ij} x_j \leq b_i$ inequality constraints.

→ $\sum_{j=1}^{n} d_{ij} x_j = e_i$ equality constraints.

Example: (Assignment problem) → match N people to N task
→ cost of assigning person i to task j is a_{ij}

First step: Define your variable

$x_{ij} \in \mathbb{R}$ is my variable where

$$\begin{cases} x_{ij}=1 & \text{if task } j \text{ is given to person } i \\ x_{ij}=0 & \text{o.w.} \end{cases}$$

How many x_{ij} do we have? N^2

How many of them are non zero? N

$$\boxed{\begin{aligned} &\text{minimize} \quad \sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij} x_{ij} \\ &\text{s.t.} \quad \sum_{i=1}^{N} x_{ij} = 1 \qquad j=1,\dots,N \\ &\qquad \sum_{j=1}^{N} x_{ij} = 1 \qquad i=1,\dots,N \\ &\qquad x_{ij} \in \{0,1\} \end{aligned}}$$

$N!$ feasible answers!!

relax $x_{ij} \in \{0,1\}$ to get a linear program

LP: minimize $\sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} x_{ij}$

s.t. $\sum_{i=1}^{N} x_{ij} = 1$ for $j = 1, \dots, N$

$\sum_{j=1}^{N} x_{ij} = 1$ for $i = 1, \dots, N$

$0 \leq x_{ij} \leq 1$

Solve the LP $\longrightarrow$ $x_{11}^*, \dots, x_{NN}^*$ optimal will satisfy the condition that $x_{ij} \in \{0,1\}$

$\longrightarrow$ we will prove it later. Message: we transformed a costly combinatorial problem to an LP that can be solved efficiently

Vector-Matrix form of LP:

what was the general LP formulation?

$$\underset{\{x_1, \dots, x_n\}}{\text{minimize}} \quad \sum_{j=1}^{n} c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad \text{for } i = 1, \dots, m$$

$$\sum_{j=1}^{n} d_{ij} x_j = e_i \quad \text{for } i = 1, \dots, p$$

Using the Vector-Matrix notation we can write it as

$$\underset{x}{\text{minimize}} \quad c^T x$$

$$\text{s.t.} \quad a_i^T x \leq b_i \quad i = 1, \dots, m$$

$$d_i^T x = e_i \quad i = 1, \dots, p$$

where $c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $a_i = \begin{bmatrix} a_{i1} \\ \vdots \\ a_{in} \end{bmatrix}$, $d_i = \begin{bmatrix} d_{i1} \\ \vdots \\ d_{in} \end{bmatrix}$

Definition: For two vectors x and $y \in \mathbb{R}^n$, $x \preceq y \iff \begin{cases} x_1 \le y_1 \\ \vdots \\ x_n \le y_n \end{cases}$

Using this definition I can combine the inequality constraints as well as the equality constraints

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^T x \\ \text{s.t.} \quad & Ax \preceq b \\ & Dx = e \end{aligned}$$

where $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$ $\quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$

$D = \begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & & \vdots \\ d_{p1} & \cdots & d_{pn} \end{bmatrix}$ $\quad e = \begin{bmatrix} e_1 \\ \vdots \\ e_p \end{bmatrix}$

Terminology:

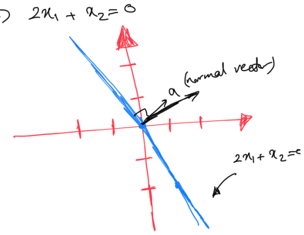
- x is feasible if it satisfies $Ax \preceq b$ & $Dx = e$
- Feasible set: set of all feasible points $\{x \mid Ax \preceq b, Dx = e\}$
- x^* is Optimal if it satisfies $c^T x^* \le c^T x$ for all feasible x
- The Optimal Value of LP is $p^* = c^T x^*$.
- Unbounded Problem: $c^T x$ unbounded below $\to p^* = -\infty$.
- Infeasible Problem: feasible set is empty $\to p^* = +\infty$

Some basic definitions:

- $\boxed{\text{Hyperplane}}$: Solution set of one linear equation with nonzero coefficient vector a $\quad a^T x = b$

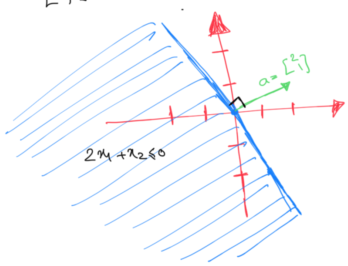
E.g.: $a = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $b = 0 \implies 2x_1 + x_2 = 0$

at least one component of vector a is nonzero



→ $\boxed{\text{Half-Space}}$: Solution set of one $\underline{\text{linear inequality}}$ with nonzero Coefficient vector $\underline{a}$ $\quad a^T x \leq b$

E.g.: $a=\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $b=0 \implies 2x_1 + x_2 \leq 0$



Here, we call $\underline{a}$ as the $\underline{\text{Normal Vector}}$ of the half-spac

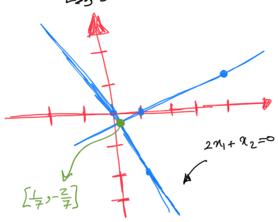
→ $\boxed{\text{affine set}}$: $\begin{cases} \text{Intersection of a set of hyper planes} \\ \text{Solution to a system of equality equations} \end{cases}$

$a_1^T x = b_1$, $a_2^T x = b_2$

$Ax = b \qquad A=\begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix} \quad b=\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$

E.g.: $a_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $b_1 = 0 \rightarrow 2x_1 + x_2 = 0$ $\quad A=\begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}$ $b=\begin{bmatrix} \\ \end{bmatrix}$

$a_2 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ $b_2 = 1 \rightarrow x_1 - 3x_2 = 1$

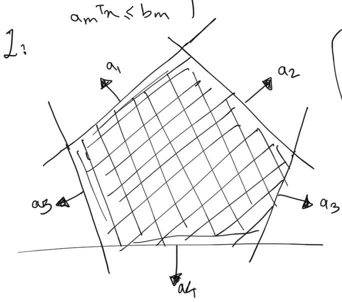


Solution set $x = [\frac{1}{7}, -\frac{2}{7}]$

$\boxed{\text{Polyhedron}}$: $\begin{cases} \text{intersection of a set of half-spaces} \\ \text{Solution of a finite Set of Linear inequalities} \end{cases}$

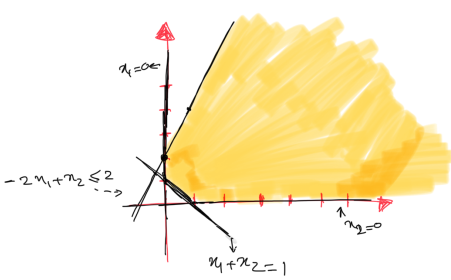
$$\left.\begin{array}{l} a_1^T x \leq b_1 \\ a_2^T x \leq b_2 \\ \vdots \\ a_m^T x \leq b_m \end{array}\right\} \implies Ax \leq b \qquad A=\begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix} \quad b=\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

Example 1:



$\boxed{\text{(*) it can also include linear equalities as } Cx = d \iff \begin{array}{l} Cx \leq d \\ -Cx \leq -d \end{array}}$

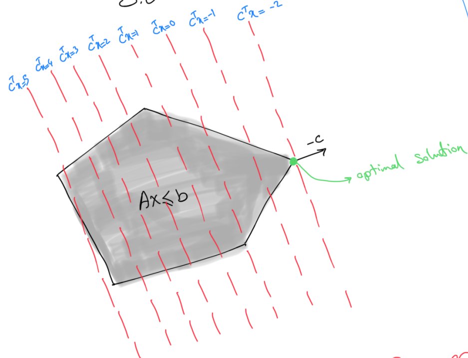
Example 2: $x_1 + x_2 \geq 1$, $-2x_1 + x_2 \leq 2$, $x_1 \geq 0$, $x_2 \geq 0$



Side note

Geometric Interpretation of LP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{s.t.} & Ax \leq b \end{array}$$



dashed red lines (hyperplanes) are level sets $c^Tx = \alpha$ for different

a line in $\mathbb{R}^n$ is defined as the set of points $x = [x_1, \dots, x_n]$ such that

$$\frac{x_1 - \alpha_1}{\beta_1} = \frac{x_2 - \alpha_2}{\beta_2} = \dots = \frac{x_n - \alpha_n}{\beta_n}$$

where $(\alpha_1, \dots, \alpha_n)$ and $(\beta_1, \dots, \beta_n)$ are scalar

In the ~~rest~~ rest of the lecture we consider a polyhedron

$$P = \{x \mid Ax \leq b,\ Cx = d\}$$

* We further assume that P is non-empty

→ $\boxed{\text{Lineality Space:}}$

The lineality space of a polyhedron $\underline{P}$ is defined as

$$L = \text{nullspace}\left(\begin{bmatrix} A \\ C \end{bmatrix}\right) \quad \begin{array}{l} \to m\times n \\ \to p\times n \end{array} \quad (m+p)\times n$$

Note: if $x \in P$, then $x + v \in P$ for all $v \in L$

why?
$$\left.\begin{array}{l} x \in P \Rightarrow Ax \leq b,\ Cx = d \\ v \in L \Rightarrow Ax = 0 \quad Cx = 0 \end{array}\right\} \Rightarrow \begin{array}{l} A(x+v) \leq b \\ C(x+v) = d \end{array}$$

→ $\boxed{\text{Pointed Polyhedron}}$:

A polyhedron P with lineality space $\{0\}$ is called point

In other words, for $P = \{x \mid Ax \leq b,\ Cx = d\}$,

$$P \text{ is a pointed polyhedron} \iff \text{null space}\left(\begin{bmatrix} A \\ C \end{bmatrix}\right) = \{0\}$$

- A polyhedron is <u>pointed</u> if it <u>does NOT</u> contain an entire li
- " " " <u>Not Pointed</u> if it contains an entire line.

Two ways to check if a polyhedron is pointed or not
1_ check it lineality space   2_ check if it contains a line

Examples:

① a halfspace $\{x \mid a^T x \leq b\}$ $(n \geq 2)$ is not pointed
as its lineality space is $\{x \mid a^T x = 0\}$

② a slab $\{x \mid -1 \leq a^T x \leq 1\}$ $(n \geq 2)$ is not pointed
as its lineality space is $\{x \mid a^T x = 0\}$

③ $\{(x, y, z) \mid |x| \leq 1, |y| \leq 1\}$ is not pointed as its
has lineality space $S = \{(0, 0, z) \mid z \in R\}$

④ probability simplex $\{x \in R^n \mid 1^T x = 1, x \geq 0\}$
we have to check nullspace of $\begin{bmatrix} -1 & 0 & \cdots & 0 \\ 0 & -1 & 0 \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots 0 & -1 \\ 1 & 1 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} -I_n \\ 1 \cdots 1 \end{bmatrix}$

$\Rightarrow$ lineality space $= \{0\}$ $\Rightarrow$ pointed

⑤ $\{(x, y, z) \mid |x| \leq z, |y| \leq z\}$

$\boxed{\text{Homework}}$.