Lecture 10 Linear Programs

Note by Samion Suwito on 2/20/25

Module 1 Recap: Linear Algebra & LS

Theme: Linear equality constraints: $Ax = b$
Find x satisfying the constraints, exactly or approximately (in ℓ^2 sense).
Problem: How to handle something like : $Ax \leq b$ (element wise)? Like solving least squares but with a non-negative x. Or

$$\min_x \|Ax - b\|_1$$

Regression problem but with ℓ^1 norm.

Linear Programs (LPs)

New Class of Problem

Any LP can be written in “standard form”:

$$\min_x c^\top x \quad \text{s.t. } Ax \leq b$$

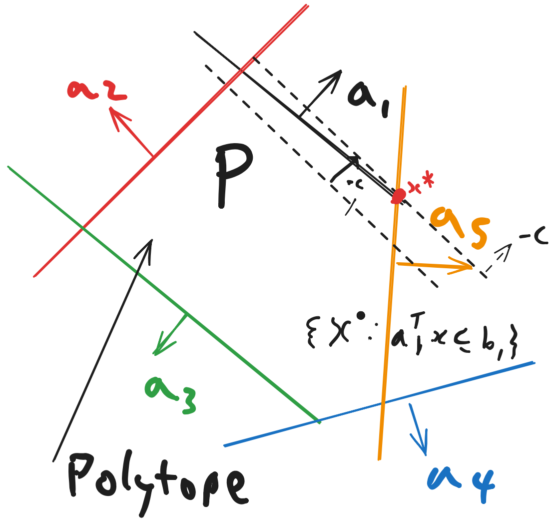
Where $c^\top x$ is a linear cost and $Ax \leq b$ is the linear equality constraints.
Other (equivalent) standard forms:

$$\min c^\top x \quad \text{s.t. } Ax = b, \;\; x \geq 0$$

”Standard Form” provides us with definition of LP = any problem that can be cast in that form..

Question: What does the “feasible set” of x s.t. $Ax \leq b$ look like?
$Ax \leq b \iff a_i^\top x \leq b_i$ where $$A=\begin{bmatrix}

- & a_{1}^{\top} & - \
  & \vdots & \
- & a^{\top}*{n} &—*
  \end{bmatrix}${x:Ax\leq b}={ x:a{i}^{\top}x\leq b*{i} ,i=1\dots n}=\cap*{i=1}^{n} {
  x:a*{i}^{\top}x\leq b*{i} }$ We call $\{x : a_i^\top x \leq b_i\}$ a half space



Intersection of halfspaces is called a “polytope”

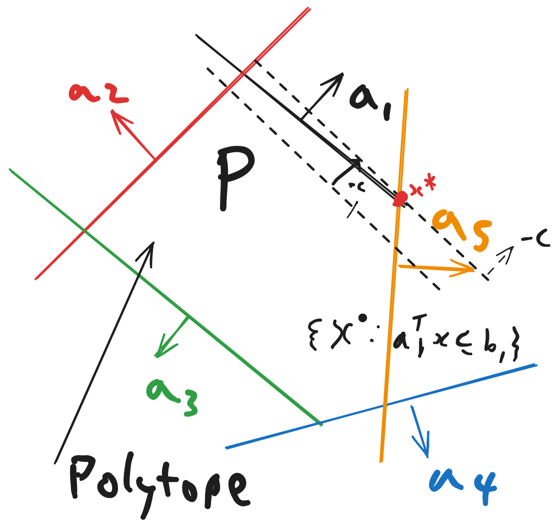
Example: Probability simplex is $\{x \in \mathbb{R}^n : x_i \geq 0, \sum^n x_i = 1\}$ Optimise probability vectors.
$x_i \geq 0 \; \forall i \implies -x \leq 0$
$\sum x_i = 1 \iff \mathbb{1}^\top x \leq 1, -\mathbb{1}^\top x \leq -1$
Which we can then write
Note: $\mathbb{1}$ represents all 1 vector.

$$\begin{bmatrix} -I \\ \mathbb{1}^\top \\ -\mathbb{1}^\top \end{bmatrix} x \leq \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{bmatrix}$$

Which represents the form $Ax \leq b$
Example 2: ℓ^1-ball $= \{x \in \mathbb{R}^n : \sum |x_i| \leq 1\}$
$= \{x \in \mathbb{R}^n : \sum s_i x_i \leq 1, \text{where } s_i \in \{-1, 1\}\}$
$= \{x \in \mathbb{R}^n : Ax \leq b\}$
A is $2^n \times n$ matrix who’s rows are all possible 2^n combination of ± 1 and $b = \mathbb{1}$

Generally Speaking: LP = $\min_x f(x)$ s..t $x \in P$
Where $f(x)$ is the affine function $c^\top x + d$ (d doesn’t matter in the optimisation since it’s constant). P is the polytope.

$\nabla f(x) = c$ which points in direction of max increase of a function so $-c$ points in direction of max decrease.



We can slowly slide $-c$ where we can find the vertex.
Really efficient in practice but in theory could be exponential and even more efficient in practice than provably optimal algorithms.

Linear programs are super common and flexible.

Many problems don't look like LPs , but can be cast as such.

Example 1:

$$\min_x c^\top x + \lambda \|x\|_1 \text{ s.t. } Ax \le b$$

We couldn't solve the ℓ^1 regularisation in LS but it's simple in LP.
#1 Trick Introduce New variables

$$\min_{x,y} c^\top x + \lambda \mathbb{1}^\top y \text{ s.t. } Ax \le b, x \le y, -x \le y$$

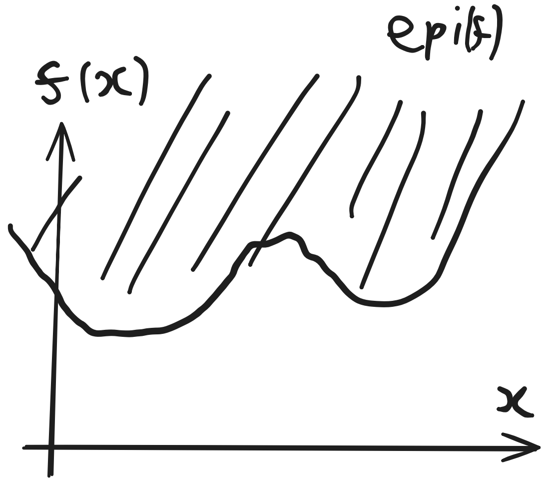
The constraints says that $x_i \le y_i$ and $-x_i \le y_i$ and if we only focus on minimising y we notice if you substitute any number one of the constraints is redundant and we get $|x_i| \le y_i$ so we can choose the smallest y_i which ends up being $|x_i|$

We can rewrite the constraints now as

$$\begin{bmatrix} A & 0 \\ I & -I \\ -I & -I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \le \begin{bmatrix} b \\ 0 \\ 0 \end{bmatrix}$$

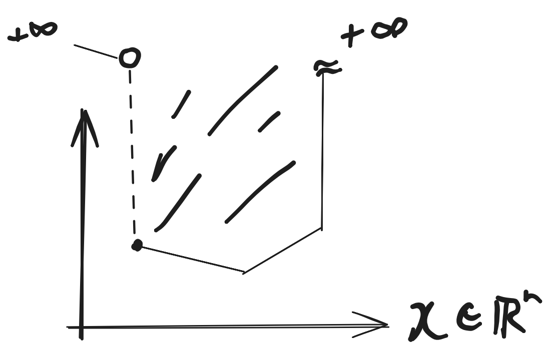
Polyhedral Functions

For $f : \mathbb{R}^n \to \mathbb{R}$ its “epigraph” is the set in $\mathbb{R}^{n+1}$ defined by the following
$\text{epi}(f) = \{(x, t) \in \mathbb{R}^{n+1} : f(x) \leq t\}$



A function is “polyhedral” if its epigraph is a polytope (polyhedron).
i.e.

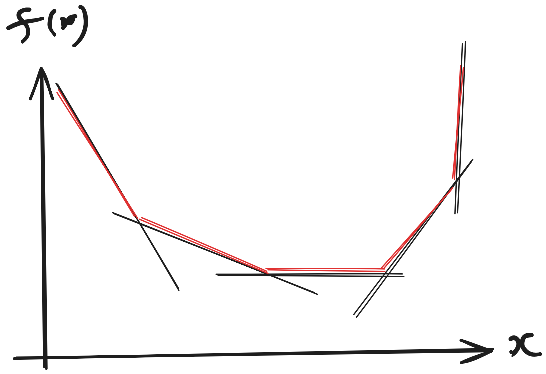
$$\text{epi}(f) = \{(x, t) \in \mathbb{R}^{n+1} : C \begin{bmatrix} x \\ t \end{bmatrix} \leq d\}$$



The epigraph deosn’t have to be bounded it just has to be the intersection of halfspaces. In fact if it is a function it will always be unbounded above. The above is a convex function

Ex: “max affine” functions:

$$f(x) = \max_{i=1...m} \{a_i^\top x + b_i\}$$



According to duality, any convex functions can be approximated by max affine functions.

Examples

Example:

$$
\begin{aligned}
f(x) &= \|x\|_\infty \\
= \max_{i=1...n} \max\{-x_i x_i\} &\leftarrow \text{max of 2n affine fn} \\
&= \text{max affine fn}
\end{aligned}
$$

Example 2:

$$f(x) = \|x\|_1 = \sum_{i=1}^{n} |x_i| = \max_{s_1 ... s_n \in \{\pm 1\}} \sum s_i x_i =$$

Max of 2^n affine functions

If f is polyhedral fn then the problem $\min_x f(x)$ is a Linear Program

$$\min_x f(x) = \min_{x,t} t \text{ s.t. } (x,t) \in \text{epi}(f) = C \begin{bmatrix} x \\ t \end{bmatrix} \leq d$$

Example of minimising polyhedral function

Example 1: ℓ_∞ regression:

$$
\begin{aligned}
\min_x \|Ax - b\|_\infty \implies \min_{x,t} t \text{ s.t. } \|Ax - b\|_\infty &\leq t \\
&= Cx \leq d >
\end{aligned}
$$

We neaed to convert $\|Ax - b\|_\infty \leq t$ to linear constraints

$$\begin{aligned} & \|Ax-b\|_\infty \le t \\ \iff & |a_i^\top x - b_i| \le t \;\; \forall i = 1 \dots m \\ \iff & \begin{cases} (a_i^\top x - b_i) \le t \\ -(a_i^\top x - b_i) \le t \end{cases} \forall i = 1 \dots m \end{aligned}$$

No need to find C, d.

Example 2: ℓ^1 regression (preferable if outliers exist)

$$\min_x \|Ax-b\|_1 = \min_{x,t} t \;\text{ s.t. } \|Ax-b\|_1 \le t$$

Again we need to convert to linear constraints

$$\begin{aligned} & \min_{x,u} \sum u_i \;\text{ s.t. } |a_i^\top x - b_i| \le u_i \; \forall i \\ \iff & \begin{cases} (a_i^\top x - b_i) \le u_i \\ -(a_i^\top x - b_i) \le u_i \end{cases} \forall i \end{aligned}$$