CS 331, Fall 2024
Lecture 15 (10/21)

Today:
- Continuous optimization
- LP defs
- LP apps
- LP duality

Continuous optimization (Part VI, Section 1)

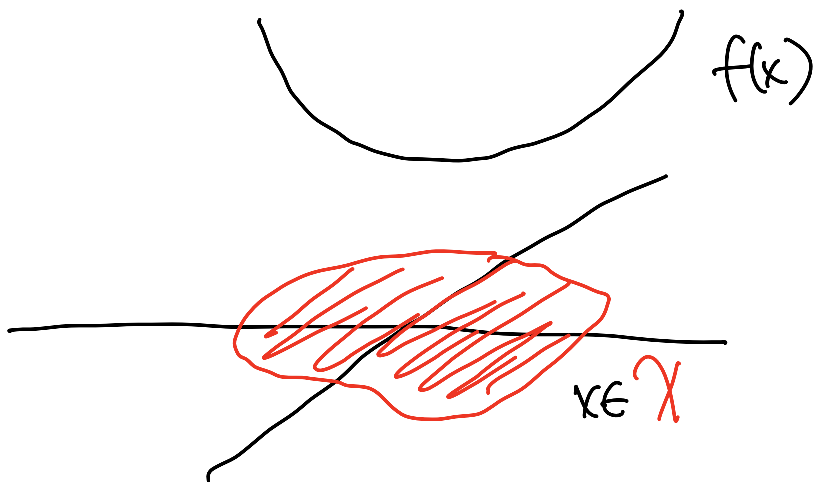
This unit:

"decision variable"

$$\min_{x \in \mathcal{X}} f(x) \quad \text{or} \quad \max_{x \in \mathcal{X}} f(x)$$

e.g.	f	$\mathcal{X}$
Scheduling	size	non-overlapping intervals
MST	total weight	spanning trees
s-t shortest path	total weight	s-t paths

Key difference: now $\underbrace{\mathcal{X}}_{\text{continuous}} \subseteq \mathbb{R}^d$



- infinite sets
- typically cannot hope for <u>exact</u>
  only high-accuracy
  (some exceptions)

e.g. s-t maxflow $\qquad \max_{x \in \mathcal{X}} f(x)$

$$f(x) = \sum_{e=(s,v)\in E} x_e$$

$$\mathcal{X} = \left\{ x \in \mathbb{R}^E \;\middle|\; \begin{array}{l} 0 \le x_e \le c_e \\ \sum_{e=(v,u)} x_e - \sum_{e=(u,v)} x_e = 0 \\ \forall v \in V \setminus \{s,t\} \end{array} \right\}$$

What are the rules of the game?

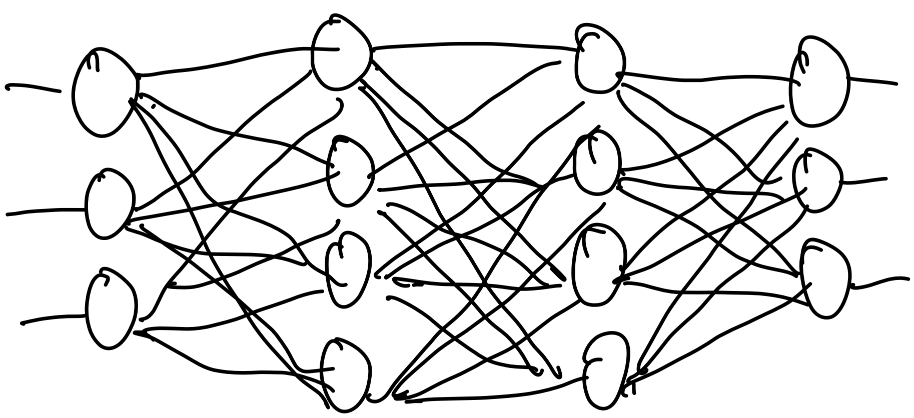
$$\min_{x \in \mathcal{X}} f(x)$$

- ability to evaluate f @ x
- ability to evaluate f' @ x reduces to

$$f'(x) \approx \frac{f(x+\delta) - f(x)}{\delta}$$

Sometimes, this is all that's reasonable.

$f(x)$ = average of neural network @ 1000000 pictures???



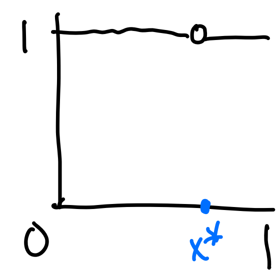
However, this is not enough. Need structure

e.g. Can we compute $\hat{x} \in \mathcal{X}$

s.t. $f(\hat{x}) \le \min_{x \in \mathcal{X}} f(x) + 0.99$,

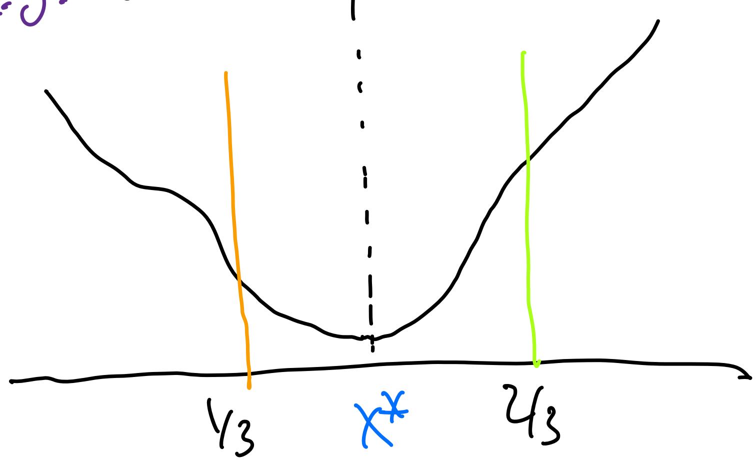
$\mathcal{X} = [0,1]$, $f : [0,1] \to [0,1]$?

No! Let $f(x) = \begin{cases} 0 & x = x^* \text{ (secret)} \\ 1 & \text{else} \end{cases}$



What structure helps?

e.g. Unimodality



"ternary search"
repeatedly throw
out 1/3 using
only f queries

Key difference: "hints" about x^*

Linear program definitions (Part VI, Section 2.1)

Linear function: $f(x) = c^T x = \sum_{i \in [d]} c_i x_i$

Examples

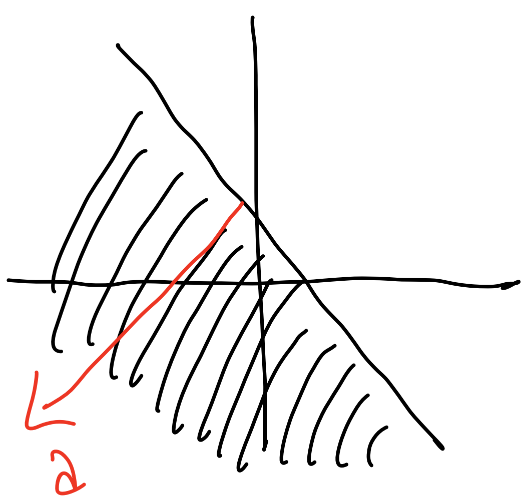
$f(x) = x_1$ $\qquad c = e_1$ ✓

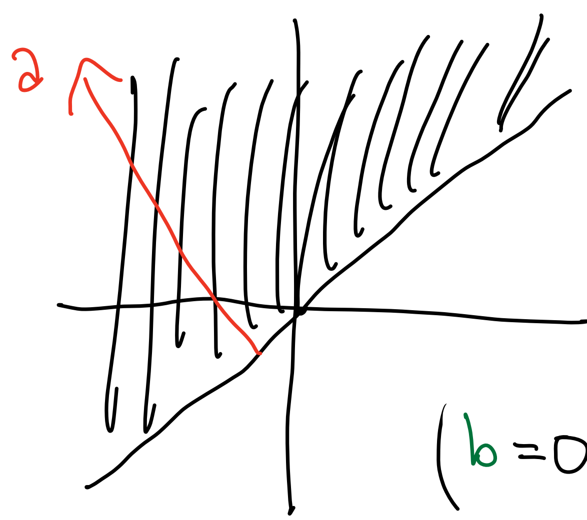
$f(x) = 5x_2 + 7x_5$ $\qquad c = \begin{pmatrix} 0 \\ 5 \\ 0 \\ 0 \\ 7 \\ \vdots \end{pmatrix}$ ✓

$f(x) = 3x_2^2 + 6x_2x_3$ ✗

Halfspace: $\{ x \in \mathbb{R}^d \mid a^T x \le b \}$

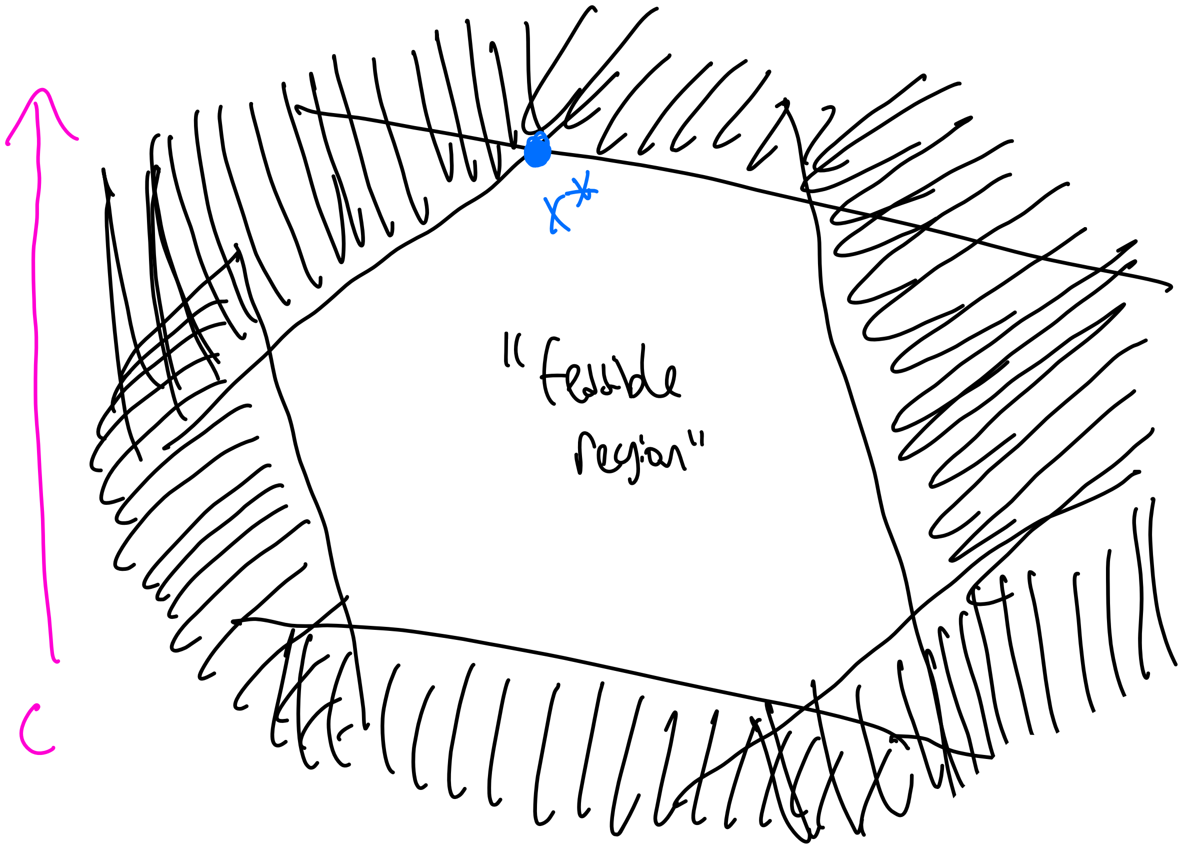
Examples $\qquad \mathbb{R}^2, \quad a_1 x_1 + a_2 x_2 \le b$





$(b = 0)$

Polytope: intersection of halfspaces



LP: Optimize linear function over polytope

$$\max \; c^T x \qquad A = \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{pmatrix} \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$Ax \le b \qquad\qquad n \times d \qquad\qquad n \times 1$$

a.k.a. $a_i^T x \le b_i$ for all $i \in [n]$

Much more powerful! $\max c^T x$
$Ax \le b$

- $\min c^T x$ OK:

  $\max -c^T x$ gives same answer

- $a^T x \ge b$ OK:

  add constraint $-a^T x \le -b$

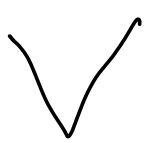
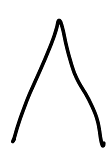
- $a^T x = b$ OK:

  add two constraints $-a^T x \le -b$
  $a^T x \le b$

- $\min c^T |x|$ OK if $c \in \mathbb{R}^d_{\ge 0}$:

  $z = \begin{pmatrix} x \\ y \end{pmatrix}$ add constraints: $y_i \le -x_i$, $y_i \le x_i$ $\forall i \in [d]$

  new objective: $\max c^T y$

Intuition:  OK to min.  not OK

LP applications (Part VI, Section 2.2)

Every LP has an equivalent __dual__.

Asymmetric form

$$\max_{\substack{x \in \mathbb{R}^d \\ Ax \le b}} c^T x \quad = \quad \min_{\substack{y \in \mathbb{R}^n_{\ge 0} \\ A^T y = c}} b^T y$$

(primal) (dual)

Symmetric form

$$\max_{\substack{x \in \mathbb{R}^d_{\ge 0} \\ Ax \le b}} c^T x \quad = \quad \min_{\substack{y \in \mathbb{R}^n_{\ge 0} \\ A^T y \ge c}} b^T y$$

(primal) (dual)

Exercise: rewrite symmetric as asymmetric, check dual is consistent.

Example: Resource allocation

$$\max_{\substack{x \in \mathbb{R}^d_{\geq 0} \\ Ax \leq b}} c^T x$$

d products

n materials

$A_{:j}$

"recipe" for product j

b = materials available

c = market price per product unit

x = amount of each product

$$Ax = \sum_{j \in [d]} A_{:j} x_j \leq b$$

Primal LP

Dual LP

$$\min_{\substack{y \in \mathbb{R}^n_{\geq 0} \\ A^T y \geq c}} b^T y$$

y = offered price per material unit

$$[A^T y]_j = \sum_{i \in [n]} A_{ij} y_i \geq c_j$$

(don't sell unless beats market price)

Interpretation of duality: "packing - covering LP"

$$\max_{\substack{x \in \mathbb{R}^d_{\geq 0} \\ Ax \leq b}} c^T x \quad = \quad \min_{\substack{y \in \mathbb{R}^n_{\geq 0} \\ A^T y \geq c}} b^T y$$

Competitor can't beat the market if we do our best.

Example: Zero-sum games

Payoff matrix

$$A = \underbrace{\begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}}_{\text{Bob action}} \Bigg\} \text{Alice action}$$

If actions $= (i,j) \in [n] \times [d]$, Alice wins A_{ij}

Bob wins $-A_{ij}$

What if random strategies?

Version 0: independent

Alice picks $y \in \mathbb{R}^n_{\geq 0}$ $\quad \sum_{i \in [n]} y_i = 1$

Bob picks $x \in \mathbb{R}^d_{\geq 0}$ $\quad \sum_{j \in [d]} x_j = 1$

Expected score:
$$\sum_{\substack{i \in [n] \\ j \in [d]}} A_{ij} x_j y_i = y^T (Ax) = x^T (A^T y)$$

$$(Ax)_i = \sum_{j \in [d]} A_{ij} x_j$$

Version 1: Bob-favored (can pick <u>after</u> y)

Score:
$$\max_{\substack{y \in \mathbb{R}^n_{\geq 0} \\ \sum_{i \in [n]} y_i = 1}} \underbrace{\min_{j \in [d]} [A^T y]_j}_{\text{only pick the lowest-scoring move}}$$

Version 2: Alice-favored (can pick <u>after</u> x)

Score: $\displaystyle \min_{\substack{x \in \mathbb{R}^n_{\geq 0} \\ \sum_{j \in [n]} x_j = 1}} \max_{i \in [n]} [Ax]_i$

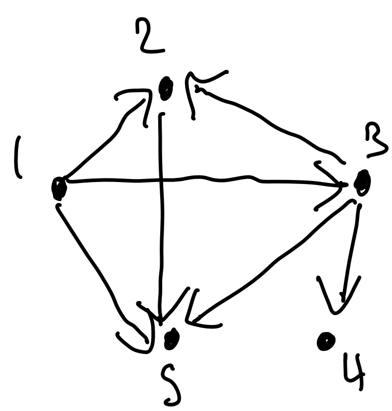
LP duality: Version 1 score = Version 2 score

(basic version of Nash equilibrium)

Example: maxflow-mincut

$A \in \mathbb{R}^{V \times E}$ edge $e=(a,b) \Rightarrow$ column $\begin{array}{cc} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} & \begin{array}{c} \\ a \\ \\ b \\ \\ \end{array} \end{array}$

$$\begin{array}{c|ccccccc} & & & & & & & \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 & 1 & -1 & 0 & 0 \\ 3 & 0 & -1 & 0 & 0 & 1 & 1 & 1 \\ 4 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 5 & 0 & 0 & -1 & -1 & 0 & 0 & -1 \\ & (1,2) & (1,3) & (1,5) & (2,5) & (3,2) & (3,4) & (3,5) \end{array}$$



"incidence matrix"

Observation:

$$[Ax]_v = \sum_{(v,u)\in E} x_{(v,u)} - \sum_{(u,v)\in E} x_{(u,v)}$$

$$= \partial x(v) \quad (\text{netflow @ } v)$$

Hence maxflow LP is

$$\max_{\substack{x\in\mathbb{R}^E_{\geq 0}\\ x\leq c\\ [Ax]_U=0}} [Ax]_s \quad \text{where } U = V\setminus\{s,t\}$$

Dual LP is

$$\min_{\substack{y\in\mathbb{R}^E_{\geq 0},\, z\in\mathbb{R}^V\\ z_s=1,\ z_t=0\\ y_e\geq z_u-z_v\ \forall e=(u,v)\in E}} c^T y$$

Interpretation

z indicates whether vertices belong to mincut

$y_e = 1$ if crosses cut

0 else

LP duality (Part VI, Section 2.3)

Weak duality easier. Let:

$x \in \mathbb{R}^d_{\geq 0}$, $Ax \leq b$ (primal feasible)

$y \in \mathbb{R}^n_{\geq 0}$, $A^T y \geq c$ (dual feasible)

$$Ax \leq b \Rightarrow \sum_{i \in [n]} y_i (Ax)_i \leq \sum_{i \in [n]} y_i b_i$$

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

$$A^T y \geq c \Rightarrow \sum_{j \in [d]} x_j [A^T y]_j \geq \sum_{j \in [d]} x_j c_j$$

$$\Rightarrow y^T A x \geq c^T y$$

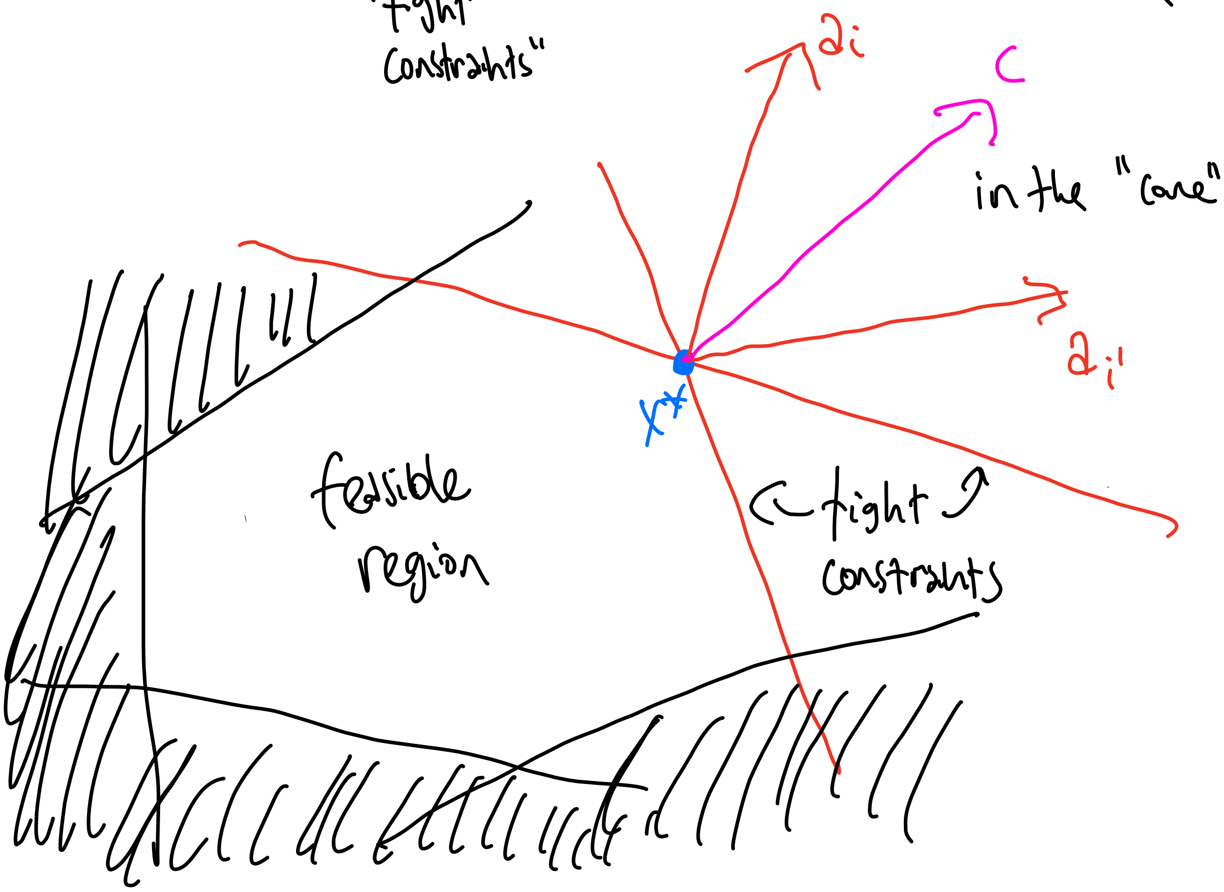
Hence we have $\displaystyle \max_{\substack{x \in \mathbb{R}^d_{\geq 0} \\ Ax \leq b}} c^T x \leq \min_{\substack{y \in \mathbb{R}^n_{\geq 0} \\ A^T y \geq c}} b^T y$

For strong duality, asymmetric form easier.

Let x^* be optimal for $\max c^T x$
$Ax \le b$

Claim: $c = \sum_{i \in I} y_i a_i \qquad y \in \mathbb{R}^n_{\ge 0}$
(Farkas' lemma)

where $I = \{ i \in [n] \mid [Ax]_i = b_i \}$
"tight constraints"



If we believe Farkas' lemma...

Let $c = \sum_{i \in [n]} y_i^* a_i$ where $y_i^* = \begin{cases} y_i & i \in I \\ 0 & i \notin I \end{cases}$

$a_i^T x^* = b_i$ for $i \in I$ (tight)

Then, $b^T y^* = \sum_{i \in I} b_i y_i^*$

$= \sum_{i \in I} (Ax)_i y_i^*$

$= \sum_{i \in [n]} (Ax)_i y_i^*$

$= y^{*T} A x^* = c^T x^*$ ■