6. Linear Programming

- Linear Programming
- LP reduction
- Duality
- Max-flow min-cut, Zero-sum game
- Integer Programming and LP relaxation
- Maximum Bipartite Matching, Minimum weight vertex cover, Shortest paths

Linear Programming

- Overview
  - We consider the following Linear Programming (LP) problem

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

    x is the variable we optimize over, and c, b, A are given constant vectors and matrices
  - Objective function: $F(x) = c^T x$
  - Constraints: $Ax = b$ and $x \geq 0$
  - A *feasible* solution is a vector x that satisfy all the constraints
  - An *optimal* solution is a feasible solution that minimizes the objective function
  - A feasible solution might not exist
  - A solution to LP can be unbounded, i.e. for any a there exists a feasible x such that $cx \leq a$

Linear Programming

- Transformations
  - LP is a family of optimization problems with linear objective function and *affine* constraints (equality and inequality)
  - There are elementary transformations that allow us to rewrite any LP to an equivalent one
  - Maximization vs. minimization
    - ★ maximize $c^T x$ is equiv. to minimize $-c^T x$
  - Equality to inequality constraints
    - ★ $Ax = b$ is equiv. to $Ax \leq b$ and $-Ax \leq -b$.
    - ★ number of constraints doubled
  - Inequality to equality constraints
    - ★ $Ax \leq b$ is equiv. to $Ax + s = b$ and $s \geq 0$ (introducing slack variables)
    - ★ number of variables increase by the number of constraints
  - Unrestricted variables to non-negative variables
    - ★ $x_i = x_i^+ - x_i^-$
    - ★ $x_i^+ \geq 0$ and $x_i^- \geq 0$
    - ★ number of variables doubled
    - ★ one solution in the original problem corresponds to infinite number of solutions

Linear Programming

- Example

$$\begin{array}{rl} \text{maximize} & c^T x \\ \text{subject to} & Ax \le b \end{array}$$

is equivalent to

$$\begin{array}{rl} \text{minimize} & c'^T x' \\ \text{subject to} & A'x' = b' \\ & x' \ge 0 \end{array}$$

for $x' = \begin{bmatrix} x^+ \\ x^- \\ s \end{bmatrix}$, $A' = \begin{bmatrix} A \\ -A \\ I \end{bmatrix}$, $b' = b$ and $c' = \begin{bmatrix} -c \\ c \\ 0 \end{bmatrix}$

Linear Programming

- Geometric interpretation of LP
- Consider $Ax \leq b$
  - the feasible set can be empty, bounded, or unbounded
  - feasible set is always *convex*
    - ⋆ A set is *convex* if and only if for any two points in the set x and y, $\alpha x + (1-\alpha)y$ is also in the set for any $\alpha \in [0, 1]$
    - ⋆ if $Ax \leq b$ and $Ay \leq b$, then $A(\alpha x + (1-\alpha)y) = \alpha Ax + (1-\alpha)Ay \leq \alpha b + (1-\alpha)b \leq b$
- Consider $Ax = b$
  - the feasible set is a convex set on a lower dimensional space

Linear Programming

- Primal LP

$$\begin{array}{rl} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b \\ & x \geq 0 \end{array}$$

- Dual LP

$$\begin{array}{rl} \text{minimize} & b^T y \\ \text{subject to} & A^T y \geq c \\ & y \geq 0 \end{array}$$

- Each variable y_i corresponds to a constraint $A_{i\cdot}x \leq b_i$
- Each constraint $A_{\cdot i}^T y \geq c_i$ corresponds to variable x_i

Linear Programming

- Weak duality theorem
  - Given feasible x and y we can derive

    $$c^T x \leq (A^T y)^T x \leq y^T A x \leq y^T b = b^T y$$

    Hence, $c^T x \leq b^T y$
  - If we can compute a feasible solution x for the primal LP, this gives a lower bound on the Dual LP
- Strong dual theorem
  - For the optimal solutions x^* and y^*,

    $$c^T x^* = b^T y^*$$

- For any feasible x and y,

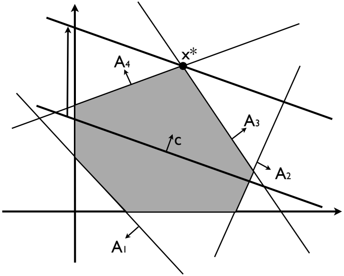
  $$c^T x \leq c^T x^* = b^T y^* \leq b^T y$$

  and this is useful when approximating the optimal solution
- This assumes both primal and dual LP are feasible with finite solution

Linear Programming

- Geometric view of LP duality
  - two dimensional $x = (x_1, x_2)$

$$
\begin{array}{ll}
\text{maximize} & c_1 x_1 + c_2 x_2 \\
\text{subject to} & A_{11}x_1 + A_{12}x_2 \leq b_1 \\
& \vdots \\
& A_{m1}x_1 + A_{m2}x_2 \leq b_m \\
& x_1 \geq 0 \\
& x_2 \geq 0
\end{array}
$$



  - the constraints A_3 and A_4 are *active* at the optimal solution x^*
  - Intuitively, at the optimal point the forces from c, A_3 and A_4 should achieve an equilibrium such that they sum to zero.

$$c - y_3^* A_3 - y_4^* A_4 = 0$$

  - Define y^* to be this vector with all zeros except for y_3^* and y_4^* that satisfy the above equation with $y_3^* \geq 0$ and $y_4^* \geq 0$

Linear Programming

- Geometric view of LP duality
  - Then, we claim that this y^* defined this way is (i) a feasible solution of the dual problem; and (ii) the optimal solution of the dual problem
  - Feasibility: From the above equation we have $A^T y^* = c$. Also, $y^* \geq 0$ by definition.
  - Optimality:

$$\begin{aligned} c^T x^* &= (A^T y^*) x^* \\ &= (y^*)^T A x^* \\ &= y_3^* A_3 x^* + y_4 A_4 x^* \\ &= y_3^* b_3 + y_4 b_4 \\ &= b^T y^* \end{aligned}$$

From weak duality we know that for any feasible y, $b^T y \geq c^T x^*$. We just showed that $c^T x^* = b^T y^*$. This proves that y^* minimizes $b^T y$ among all feasible solutions.

Linear Programming

- Perturbation and sensitivity analysis
- (unperturbed) LP and its dual (with optimal solution $x^*(0)$ and $y^*(0)$)

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b \\ & x \geq 0 \end{array} \qquad \begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \geq c \\ & y \geq 0 \end{array}$$

- perturbed LP and its dual (with optimal solution $x^*(\Delta)$ and $y^*(\Delta)$)

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b + \Delta \\ & x \geq 0 \end{array} \qquad \begin{array}{ll} \text{minimize} & b^T y + \Delta^T y \\ \text{subject to} & A^T y \geq c \\ & y \geq 0 \end{array}$$

- From weak duality, $c^T x^*(\Delta) \leq b^T y^*(0) + \Delta^T y^*(0)$
- If $y_i^*(0)$ is large, changing i-th constraint changes the optimal value greatly (the LP is sensitive to the i-th constraint)

Linear Programming

- local sensitivity
  - if $x^*(\Delta)$ is differentiable at 0, then

$$y_i^*(0) = \frac{\partial\{c^T x^*(\Delta)\}}{\partial \Delta_i}\Big|_{\Delta=0}$$

  - proof. From previous slide,

$$\begin{aligned}
\frac{\partial\{c^T x^*(\Delta)\}}{\partial \Delta_i}\Big|_{\Delta=0} &= \lim_{t \searrow 0} \frac{c^T x^*(te_i) - c^T x^*(0)}{t} \geq y_i^*(0) \\
\frac{\partial\{c^T x^*(\Delta)\}}{\partial \Delta_i}\Big|_{\Delta=0} &= \lim_{t \nearrow 0} \frac{c^T x^*(te_i) - c^T x^*(0)}{t} \leq y_i^*(0)
\end{aligned}$$

Exercise: LP duality and LP relaxation

- Primal LP

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \le b \\ & x \ge 0 \end{array}$$

- Dual LP

$$\begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \ge c \\ & y \ge 0 \end{array}$$

- Exercise

$$\begin{array}{ll} \text{maximize} & 0.1x + 0.3y + 0.4z \\ \text{subject to} & x + y + z \le 1 \\ & x + y - 0.3z \le 1 \\ & x + 0.5y + z \le 0.8 \\ & x - y \le 1.3 \\ & x, y, z \ge 0 \end{array}$$

- What is the dual LP?
- Is it feasible?

Complementary slackness

LP Duality

for some problems, the strong duality theorem for LP gives fresh insights into the primal problem

- Max-flow Min-cut
- Shortest paths
- Zero-sum games

for some problems that can be formulated as integer program (IP), we can solve the LP relaxations and still find the optimal solutions of IP

- Maximum spanning trees
- Maximum bipartite matching

for some problems that can be formulated as integer program (IP), LP relaxations give approximation algorithm with precise approximation guarantees

- Minimum vertex cover

Max-flow and min-cut

- A weighted directed graph $G = (V, E, C)$ with source s and sink t
- There are two ways to formulate this problem
  - variable for each edge: f_{ij}
  - variable for each path p from s to t: $f(p)$
- first one has smaller number of variables, but second one is more intuitive
- for each path $p \in \mathcal{P}$ ($\mathcal{P}$ is the set of all paths from s to t)

$$
\begin{array}{ll}
\text{maximize} & \displaystyle\sum_{p \in \mathcal{P}} f(p) \\
\text{subject to} & \displaystyle\sum_{p:(i,j) \in p} f(p) \le C_{ij}, \ \forall (i,j) \in E \\
& f(p) \ge 0, \ \forall p
\end{array}
$$

- consider a $|\mathcal{P}| \times |E|$ matrix $A = [A_{p,(i,j)}]$ such that $A_{p,(i,j)} = 1$ if $(i,j) \in p$
- then, the constraint is $A^T f \le C$
- let ℓ_{ij}'s be the dual variable for each edge

Exercise: LP duality and LP relaxation

- Primal LP

$$\begin{aligned} \max \quad & \sum_{p \in \mathcal{P}} f_p \\ \text{s. t.} \quad & \sum_{p:(i,j)\in p} f_p \le C_{ij},\ \forall (i,j) \in E \\ & f_p \ge 0,\ \forall p \end{aligned}$$

- variable vector: $f_p \in \mathbb{R}^{|\mathcal{P}|}$
- objective: $c^T f$ with $c = [1, \ldots, 1] \in \mathbb{R}^{|\mathcal{P}|}$
- constraints: $Af \le b$ with $A \in \mathbb{R}^{|\mathcal{P}| \times |E|}$

$$A_{e,p} = \begin{cases} 1 & \text{if } e \in p \\ 0 & \text{otherwise} \end{cases}$$

- $b_e = C_{ij} \in \mathbb{R}^{|E|}$ for $e = (i,j)$
- What is the dual LP?

Max-flow and min-cut

- Dual LP:

$$
\begin{array}{ll}
\text{minimize} & \displaystyle\sum_{(i,j)\in E} \ell_{ij}\, C_{ij} \\
\text{subject to} & \displaystyle\sum_{(i,j)\in p} \ell_{ij} \geq 1,\ \forall p \in \mathcal{P} \\
& \ell_{ij} \geq 0,\ \forall (i,j) \in E
\end{array}
$$

- Interpretation
  - consider ℓ_{ij} as (virtual) length of the edge (i,j)
  - C_{ij} as the volume per unit length of the edge (i,j)
  - minimize the total volume of the pipes
  - while ensuring that for all paths, the length of that path is at least one
- One way to construct a feasible dual ℓ_{ij} is to use graph cuts
  - Consider any cut (S, S^c) with $s \in S$ and $t \in S^c$
  - Let ℓ_{ij} be one if it crosses the cut and zero otherwise
  - This is feasible: any path goes through one of the edges across the cut at least once
  - the value of the cut is $\sum_{(i,j)\in E} \ell_{ij}\, C_{ij}$

weak duality theorem implies

$$\sum_{p \in \mathcal{P}} f(p) \leq \sum_{(i,j) \in E} \boldsymbol{\ell}_{ij} C_{ij}$$

- Consider any any ℓ_{ij}'s derived from a cut (S, S^c)
- Consider any flow $f(p)$'s
- the weak duality implies that the value of a cut is larger than or equal to value of a flow (we already know this, but this is an alternative way for proving it)

proving max-flow min-cut theorem using **strong duality theorem**

we can prove a much stronger result (which we already proved using Ford-Fulkerson algorithm in previous lectures) of max-flow min-cut theorem using LP duality, stating that there exists a cut whose value is equal to the maximum flow

theorem. for all feasible solution ℓ of the dual LP of max-flow problem, there exists a cut (S, S^c) such that if we let

$$\ell^*_{ij} = \begin{cases} 1 & \text{if } i \in S, j \in S^c \\ 0 & \text{otherwise} \end{cases}$$

then $c(S, S^c) = \sum_{(i,j)\in E} \ell^*_{ij} C_{ij} \leq \sum_{(i,j)\in E} \ell_{ij} C_{ij}$

corollary. [max-flow min-cut theorem] there exists a cut (S, S^c) such that the value of the cut is equal to the value of the maximum flow

proof of the corollary. from strong duality, we know that there is a dual feasible solution such that

$$\sum_{p \in \mathcal{P}} f_p = \sum_{(i,j) \in E} \ell^*_{ij} C_{ij}$$

from the adobe theorem, we know that there exists a cut with value at most $\sum_{(i,j)\in E} \ell^*_{ij} C_{ij}$ and this proves the max-flow min-cut theorem (notice that the cut value cannot be smaller due to weak duality theorem)

proof of the theorem. for each node i, let d_i be the shortest path distance from s to i according to the lengths defined by ℓ_{ij}'s, then it follows that $d_t \geq 1$
for a $\rho \in [0, 1)$, let

$$S_\rho = \{i \in V | d_i \leq \rho\}$$

then (S_ρ, S_ρ^c) is a cut in G

now suppose we choose ρ uniformly at random from $[0,1)$. If we can show that

$$\mathbb{E}[c(S_\rho, S_\rho^c)] \le \sum_{(i,j)\in E} \ell_{ij} C_{ij}$$

then it proves that there exists at least one cut (S_ρ, S_ρ^c) such that

$$c(S_\rho, S_\rho^c) \le \sum_{(i,j)\in E} \ell_{ij} C_{ij}$$

to prove the bound on the expectation, note that

$$\mathbb{E}[c(S_\rho, S_\rho^c)] = \sum_{(i,j)\in E} C_{ij} \mathbb{P}\big((i,j) \in c(S_\rho, S_\rho^c)\big)$$

since $\mathbb{P}\big((i,j) \in c(S_\rho, S_\rho^c)\big) = \mathbb{P}(\rho \in [d_i, d_j)) = d_j - d_i$ and $d_j \le d_i + \ell_{ij}$, we have

$$\mathbb{P}\big((i,j) \in c(S_\rho, S_\rho^c)\big) \le \ell_{ij}$$

putting the inequalities together we get that there exists a cut corresponding to a value ρ^* such that

$$\begin{array}{rcl} c(S_{\rho^*}, S_{\rho^*}^c) & \leq & \mathbb{E}_\rho[c(S_\rho, S_\rho^c)] \\ & = & \displaystyle\sum_{(i,j)\in E} C_{ij}\mathbb{P}\big((i,j) \in c(S_\rho, S_\rho^c)\big) \\ & \leq & \displaystyle\sum_{(i,j)\in E} C_{ij}\ell_{ij} \end{array}$$

this finishes the proof of the theorem

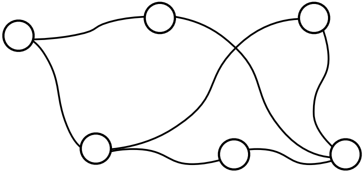
complementary slackness

if $\ell_{ij} > 0$ then $\sum_{p:(i,j)\in p} f_p = C_{ij}$ (if an edge is in the minimum cut in the dual, then the edge is saturated in the primal)

if $f_p > 0$ then $\sum_{(i,j)\in p} \ell_{ij} = 1$ (all paths carrying non-zero flow much be one of many shortest paths and pass through a min-cut once)

Shortest Paths

- Pair-wise Shortest Paths Problem:
  - Given an undirected weighted graph $G = (V, E, w)$
  - Find the shortest path from node s to node t
- Consider solving the shortest paths problem using the following *physical device*
  - The physical device has nodes V connected by strings of length w_{ij} connecting node i and node j
  - To find the shortest path from s to t, one only needs to pull s away from t until one can no longer pull them away
  - Intuitively, this finds the shortest path from s to t



- Question: why are we pulling (maximization) when we want to minimize the distance?

- LP formulation as a *min-cost flow* problem with infinite capacity

$$\begin{array}{rl} \text{minimize} & \displaystyle\sum_{(i,j)\in E} f_{ij}\,w_{ij} \\ \text{subject to} & f_{ij} \geq 0, \;\; \forall (i,j) \in E \\ & \displaystyle\sum_{k:(ki)\in E} f_{ki} - \sum_{k:(ik)\in E} f_{ik} = \begin{cases} 0 & \forall i \neq s,t \\ 1 & \text{if } i = t \\ -1 & \text{if } i = s \end{cases} \end{array}$$

  - feasible set: set of all flows of value one
  - LP: bring 1 unit flow from s to t with minimum cost
  - equivalent to finding a single path from s to t with minimum distance
  - **Claim.** there always exists a *single* path with min-cost flow

- Transformation for writing the dual

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{(i,j)\in E} f_{ij} w_{ij} \\ \text{subject to} & f_{ij} \geq 0, \;\; \forall (i,j) \in E \\ & \displaystyle\sum_{k:(ki)\in E} f_{ki} - \sum_{k:(ik)\in E} f_{ik} = 0, \;\; \forall i \notin \{s,t\} \\ & \displaystyle\sum_{k:(kt)\in E} f_{kt} - \sum_{k:(tk)\in E} f_{tk} = 1 \\ & \displaystyle\sum_{k:(ks)\in E} f_{ks} - \sum_{k:(sk)\in E} f_{sk} = -1 \end{array}$$

  - Rewriting equality constraints as inequalities

$$\sum_{k:(ki)\in E} f_{ki} - \sum_{k:(ik)\in E} f_{ik} \leq 0$$

$$\sum_{k:(ki)\in E} f_{ki} - \sum_{k:(ik)\in E} f_{ik} \geq 0$$

    each corresponding to dual variables x_i^- and x_i^+

- Primal LP

$$\begin{array}{ll} \text{maximize} & c^T f \\ \text{subject to} & Af \leq b \\ & f \geq 0 \end{array}$$

  - $c_{ij} = -w_{ij}$
  - $A = \begin{bmatrix} \widetilde{A} \\ -\widetilde{A} \end{bmatrix}$, with $\widetilde{A}_{k,(i,j)} = \begin{cases} 1 & \text{if } k = j \\ -1 & \text{if } k = i \end{cases}$
  - $b = \begin{bmatrix} \tilde{b} \\ -\tilde{b} \end{bmatrix}$, with $\tilde{b}_i = \begin{cases} 0 & \forall i \neq s, t \\ 1 & \text{if } i = t \\ -1 & \text{if } i = s \end{cases}$

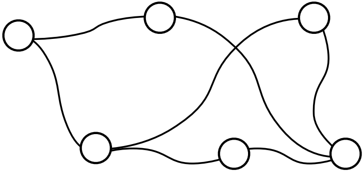
- Dual LP

$$
\begin{array}{ll}
\text{minimize} & -(x_t^+ - x_t^-) + (x_s^+ - x_s^-) \\
\text{subject to} & x_i^+, x_i^- \geq 0, \ \ \forall i \\
& -(x_i^+ - x_i^-) + (x_j^+ - x_j^-) \geq -w_{ij}, \ \ \forall (i,j) \in E
\end{array}
$$

- Can be rewritten as (transformation)

$$
\begin{array}{ll}
\text{maximize} & x_t - x_s \\
\text{subject to} & x_i - x_j \leq w_{ij}, \ \ \forall (i,j) \in E
\end{array}
$$

- Intuitively, this is stretching s and t as far apart as possible, subject to each endpoint of any edge (i,j) are separated by at most w_{ij}
- This is expected from the intuition from the physical device with strings



- More generally, we have proven (and used) the following LP duality

  - Primal LP

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

  - Dual LP

$$\begin{aligned} \text{maximize} \quad & b^T y \\ \text{subject to} \quad & A^T y \leq c \end{aligned}$$

Zero-sum game

- A zero-sum game is represented by a payoff matrix M
- Example: rock-paper-scissors

	r	p	s
r	0	-1	1
p	1	0	-1
s	-1	1	0

Table : Payoff matrix for Player A

- if player A uses a fixed strategy, then player B can countermove
- so we consider *mixed strategies*: player A chooses action i with probability x_i, player B chooses action j with probability y_j
- the expected payoff to player A (row player) is

$$\sum_{i,j} M_{ij} x_i y_j$$

- Player A wants to maximize it, player B wants to minimize it an optimal strategy for rock-paper-scissors: uniformly random

consider A's strategy $x = [1/3, 1/3, 1/3]$

fact 1. the expected payoff is 0 regardless of B's strategy

claim. This is optimal strategy for A, in the sense that there exists a strategy for B (that does not depend on A's strategy) where A cannot get an expected payoff larger than 0.

proof. Consider the best scenario for A, where B chooses a mixed strategy and then A choose her mixed strategy knowing B's strategy. Clearly, this can only be better than when A chooses her strategy without knowing B's. We claim that even in this best case scenario, A cannot do better than 0.

Let B play uniform strategy $[1/3, 1/3, 1/3]$, then B can also achieve 0 regardless of A's actions.

claim. this optimal strategy is unique in the sense that for any other strategy of A, there exists a strategy for B (that depends on the strategy of A) such that the payoff is strictly negative

proof. let $[x_1, x_2, x_3]$ be A's strategy and $[x_3, x_2, x_1]$ be B's strategy. Then the payoff is $-x_1^2 - x_2^2 - x_3^2 + x_1x_2 + x_2x_3 + x_1x_3 = -\frac{1}{2}\left\{(x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_1 - x_3)^2\right\}$. This is strictly negative unless $x_1 = x_2 = x_3 = 1/3$.

question: for general zero-sum games, if A can choose her strategy after B, can she do better than if B can choose her action after A?
the one choosing strategy after has obvious advantage
however, we will see that if both players play optimally, there is no gain in choosing strategy after the other player

- Non-symmetric Example:

	a	b
c	3	-1
d	-2	1

- Once player A's strategy is fixed $x = [x_1, x_2]$, there is always a pure strategy that is optimal for B
  - ★ B chooses action a if and only if a gives A smaller payoff
  - ★ payoff to A when action a: $(3x_1 - 2x_2)$
  - ★ payoff to A when action b: $(-x_1 + x_2)$

$$\text{payoff}(x_1, x_2) \;=\; \min\ \{3x_1 - 2x_2, -x_1 + x_2\}$$

hence, A can predict B's optimal strategy and choose x accordingly to

$$\max_{x_1, x_2}\ \min\ \{3x_1 - 2x_2, -x_1 + x_2\}$$

the maximum is achieved when $3x_1 - 2x_2 = -x_1 + x_2$ (and $x_1 + x_2 = 1$), which gives $x_1^* = 3/7$ and $x_2^* = 4/7$; the optimal expected payoff for A is $1/7$

- In general, we can formulate $\max_{x_1,x_2} \min\{3x_1 - 2x_2, -x_1 + x_2\}$ as an LP

$$\begin{aligned} \text{maximize}_{\ x_1,x_2,z} \quad & z \\ \text{subject to} \quad & z \leq 3x_1 - 2x_2 \\ & z \leq -x_1 + x_2 \\ & x_1 + x_2 = 1 \\ & x_1, x_2 >= 0 \end{aligned}$$

  - since we are maximizing z, z is always saturating at least one of the constraints such that $z = \min\{3x_1 - 2x_2, -x_1 + x_2\}$
  - hence, we are effectively maximizing $\min\{3x_1 - 2x_2, -x_1 + x_2\}$
  - we can do this for any zero-sum games with any number of actions (as long as it is finite)

- Optimal strategy for A

$$\begin{aligned} \text{maximize} \quad & z \\ \text{subject to} \quad & z \le 3x_1 - 2x_2 \\ & z \le -x_1 + x_2 \\ & x_1 + x_2 = 1 \\ & x_1, x_2 \ge 0 \end{aligned}$$

- Optimal strategy for B

$$\begin{aligned} \text{minimize} \quad & w \\ \text{subject to} \quad & w \ge 3y_1 - y_2 \\ & w \ge -2y_1 + y_2 \\ & y_1 + y_2 = 1 \\ & y_1, y_2 \ge 0 \end{aligned}$$

- Notice two LP's are dual
  - ⋆ $x = \begin{bmatrix} x_1 \\ x_2 \\ z^+ \\ z^- \end{bmatrix}$, $c = [0, 0, 1, -1]$, $A = \begin{bmatrix} -3 & 2 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$
- By strong duality, they have the same optimal value
- This proves that for general zero-sum games with finite action space, there exists optimal mixed strategies for both players and they achieve the same value (a fundamental result in game theory known as *min-max theorem*)

$$\max_x \min_{y(x)} \sum_{i,j} M_{ij} x_i y_j = \min_y \max_{x(y)} \sum_{i,j} M_{ij} x_i y_j$$

Maximum weight perfect matching in bipartite graphs

consider an undirected weighted bipartite graph $G = (U, V, E)$ with $|U| = |V|$ and non-negative weights w_{ij}'s on the edges

a *perfect matching* is a matching that includes all the nodes in the graph

max-weight bipartite perfect matching problem can be formulated as the following integer programming:

$$
\begin{array}{ll}
\text{maximize} & \displaystyle\sum_{(i,j)\in E} w_{ij} x_{ij} \\
\text{subject to} & \displaystyle\sum_{j:(i,j)\in E} x_{ij} = 1\,, \quad \text{for all } i \in U \\
& \displaystyle\sum_{i:(i,j)\in E} x_{ij} = 1\,, \quad \text{for all } j \in V \\
& x_{ij} \in \{0, 1\}\,, \quad \text{for all } (i,j) \in E
\end{array}
$$

the LP relaxation is

$$
\begin{array}{ll}
\text{maximize} & \displaystyle\sum_{(i,j)\in E} w_{ij} x_{ij} \\
\text{subject to} & \displaystyle\sum_{j:(i,j)\in E} x_{ij} = 1\,, \quad \text{for all } i \in U \\
& \displaystyle\sum_{i:(i,j)\in E} x_{ij} = 1\,, \quad \text{for all } j \in V \\
& x_{ij} \geq 0\,, \quad \text{for all } (i,j) \in E
\end{array}
$$

fact 1. we do not need upper bounds on x_{ij}'s

theorem. this LP always has an integer optimal solution

corollary. there is no loss in LP relaxation: we can (efficiently) solve the LP relaxation to get the optimal solution of the original IP

proof of the theorem. we will prove this by showing that, if the LP relaxation has a non-integral optimal solution, then we can always find an integral solution with the same value of $w^T x = \sum w_{ij} x_{ij}$.

Suppose there is an optimal, feasible and non-integral solution $\tilde{x}$. Then, there exists an edge (i, j) with non-integral solution $x_{ij} \neq 0$ and $x_{ij} \neq 1$. Then by the equality constraint on node i, node i is connected to another edge with non-integral solution. Similarly, j is connected to another edge with non-integral solution. Similarly, we can continue extending this line of non-integral solution edges until they form a cycle C with even number of edges, since the graph is bipartite.

Let $\epsilon = \min\{\min_{(i,j)\in C}(x_{ij}), \min_{(i,j)\in C}(1 - x_{ij})\}$, such that it is the minimum distance of the solution of the edges in the cycle to its closest integral solution. Let x^+ be the same as $\tilde{x}$ but with ϵ added to the even edges and $-\epsilon$ added to the odd edges. Let x^- be the same but now ϵ added to the odd edges and $-\epsilon$ added to the even edges. Note that $x = \frac{1}{2}(x^+ + x^-)$.

Further, since $\tilde{x}$ is optimal solution and $w^T\tilde{x} = \frac{1}{2}(w^Tx^+ + w^Tx^-)$, it follows that both x^+ and x^- achieve the optimal value:

$$w^Tx^+ = w^Tx^- = w^T\tilde{x}$$

Either x^+ or x^- have at least one less non-integral edge. One can continue this process until all non-integral edges are eliminated, resulting in a integral solution with the same objective value.

Minimum/maximum spanning tree

Integer Program (IP) formulation of the MST problem

$$
\begin{aligned}
\text{maximize} \quad & \sum_{(i,j)\in E} w_{ij} x_{ij} \\
\text{subject to} \quad & \sum_{(i,j)\in E_S} x_{ij} \;\le\; |S| - 1\,,\ \text{ for all } S \subseteq V \\
& x_{ij} \in \{0,1\}\,,\ \text{ for all } (i,j) \in E
\end{aligned}
$$

Primal LP relaxation of the MST problem

$$
\begin{aligned}
\text{maximize} \quad & \sum_{(i,j)\in E} w_{ij} x_{ij} \\
\text{subject to} \quad & \sum_{(i,j)\in E_S} x_{ij} \;\le\; |S| - 1\,,\ \text{ for all } S \subseteq V \\
& x_{ij} \ge 0\,,\ \text{ for all } (i,j) \in E
\end{aligned}
$$

fact 1. $x_{ij} > 1$ is not feasible (we do not need the upper bound on x_{ij}'s)
fact 2. a spanning tree satisfies the constraints with equality

theorem. there exists an optimal *integer* solution of the primal LP, and it achieves the same weight as the Kruskal's algorithm

proof. we construct the dual LP

$$\begin{array}{ll} \text{maximize} & \displaystyle\sum_{S \subseteq V} r(S) y_S \\ \text{subject to} & \displaystyle\sum_{S:(i,j)\in E_S} y_S \ \geq \ w_{ij}\ , \ \text{for all } (i,j) \in E \\ & y_S \geq 0\ , \ \text{for all } S \subseteq V \end{array}$$

where $r(S) = |S| - 1$

we will find a solution $y^* = \{y_S^*\}_{S \subseteq V}$ of this dual problem, and show

lemma 1. y^* is a feasible solution of the dual

lemma 2. $\sum_{S \subseteq V} r(S) y_S^* = \sum_{(i,j)\in MST} w_{ij}$

where MST is a Maximum Spanning Tree solution of the Kruskal's algorithm by weak duality theorem, a feasible solution of the dual provides an upper bound on the primal problem

since we have found a pair of feasible primal and dual solutions (the MST of Kruskal's algorithm for the primal and y^* for the dual) that achieve the same objective value, this proves that the pair of solutions are optimal for the primal and dual LP, respectively

we can conclude that the Kruskal's Maximum Spanning Tree (MST) solution is an optimal solution of the Primal LP relaxation (as well as the optimal solution of the original IP formulation)

now we construct a dual solution y^*

Maximum bipartite matching

- Bipartite graph $G = (L, R, E)$
- Variable: $|L| \times |R|$ matrix X such that $X_{ij} = 1$ indicate that we include edge (i, j) in the matching
- Integer Programming (IP) formulation

$$
\begin{array}{ll}
\text{maximize} & \sum_{ij} X_{ij} \\
\text{subject to} & X_{ij} \in \{0, 1\}, \ \forall (i, j) \in E \\
& \sum_{j} X_{ij} \leq 1, \ \forall i \\
& \sum_{i} X_{ij} \leq 1, \ \forall j
\end{array}
$$

- IP: non-convex feasible set
- LP relaxation: take convex hull of the feasible set
- Since feasible set of LP relaxation includes the feasible set of IP, the optimal value of LP relaxation can only be larger

$$P^*_{IP} \leq P^*_{LP}$$

Exercise: LP duality and LP relaxation

- Integer Program (IP)

$$
\begin{aligned}
\text{maximize} \quad & \sum_{ij} X_{ij} \\
\text{subject to} \quad & X_{ij} \in \{0,1\},\ \forall (i,j) \in E \\
& \sum_{j} X_{ij} \le 1,\ \forall i \\
& \sum_{i} X_{ij} \le 1,\ \forall j
\end{aligned}
$$

- LP relaxation

$$
\begin{aligned}
\text{maximize} \quad & \sum_{ij} X_{ij} \\
\text{subject to} \quad & 0 \le X_{ij},\ \forall (i,j) \in E \\
& \sum_{j} X_{ij} \le 1,\ \forall i \\
& \sum_{i} X_{ij} \le 1,\ \forall j
\end{aligned}
$$

- variable vector: $X_{ij} \in \mathbb{R}^{|E|}$
- objective: $c^T X$ with $c = [1, \dots, 1] \in \mathbb{R}^{|E|}$
- constraints: $AX \le b$ with $A \in \mathbb{R}^{(|L|+|R|)\times|E|}$

$$
A_{i,e} = \begin{cases} 1 & \text{if } i \in e \\ 0 & \text{otherwise} \end{cases}
$$

- $b = [1, \dots, 1] \in \mathbb{R}^{|L|+|R|}$
- What is the dual LP?

Maximum bipartite matching

- Linear Programming relaxation

$$\begin{array}{ll} \text{maximize} & \sum_{ij} X_{ij} \\ \text{subject to} & 0 \leq X_{ij},\ \forall (i,j) \in E \\ & \sum_{j} X_{ij} \leq 1,\ \forall i \\ & \sum_{i} X_{ij} \leq 1,\ \forall j \end{array}$$

- Dual problem

$$\begin{array}{ll} \min & \sum_{i} Y_i + \sum_{j} Z_j \\ \text{s.t.} & 0 \leq Y_i,\ \forall i \\ & 0 \leq Z_j,\ \forall j \\ & Y_i + Z_j \geq 1,\ \forall (i,j) \in E \end{array}$$

- The LP relaxation does not require $X_{ij} \leq 1$.
- Dual problem is solving minimum vertex cover: find smallest set of nodes S such that at least one end of each edge is in S
- From strong duality theorem, we know $P^*_{LP} = D^*_{LP}$
- Consider IP formulation of the dual where $Y_i, Z_j \in \{0, 1\}$, then

$$P^*_{IP} \leq P^*_{LP} = D^*_{LP} \leq D^*_{IP}$$

- This implies that minimum vertex cover is at least as large as maximum bipartite matching

Minimum weight vertex cover

- Problem
  - An undirected graph $G = (V, E)$ with node weights w_i's
  - A *vertex cover* is a set of nodes S such that each edge has at least one end in S
  - The weight of a vertex cover is sum of all weights of nodes in the cover
  - Find the vertex cover with minimum weight

- Integer Program (IP): $x_i = 1$ if node i is in vertex cover

$$
\begin{array}{ll}
\text{minimize} & \sum_{i \in V} x_i w_i \\
\text{subject to} & x_i + x_j \geq 1, \ \forall (i,j) \in E \\
& x_i \in \{0,1\}, \forall i \in V
\end{array}
$$

- LP relaxation

$$
\begin{array}{ll}
\text{minimize} & \sum_{i \in V} x_i w_i \\
\text{subject to} & x_i + x_j \geq 1, \ \forall (i,j) \in E \\
& x_i \geq 0, \forall i \in V
\end{array}
$$

- LP relaxation

$$
\begin{array}{ll}
\text{minimize} & \sum_{i \in V} x_i w_i \\
\text{subject to} & x_i + x_j \geq 1, \ \forall (i,j) \in E \\
& x_i \geq 0 \ , \forall i \in V
\end{array}
$$

- In the LP relaxation, we do not need $x_i \leq 1$, since the optimal solution x^* of the LP does not change whether we have the constraint $x_i \leq 1$ to not.
- Here is a simple proof by contradiction
  - Suppose there exists an index i such that the optimal solution of the above LP x_i^* is strictly larger than one.
  - Then, let x^{**} be a vector which is same as x^* except for $x_i^{**} = 1 < x_i^*$.
  - This x^{**} satisfies all the constraints, and the objective function is smaller: $\sum_k x_k^{**} w_k < \sum_k x_k^* w_k$.
  - This contradicts the assumption that x^* is the optimal solution of the minimization

- When we solve the relaxed LP, we typically get fractional solutions with $0 < x_i^* < 1$
- One way to get integer solution is *rounding*:

$$x_i' = \begin{cases} 0 & \text{if } x_i^* < 0.5 \\ 1 & \text{if } x_i^* \geq 0.5 \end{cases}$$

- **Claim.** The rounded solution x' is feasible in the original problem
  - We need to show that $x_i' + x_j' \geq 1$ for all $(i, j) \in E$.
  - This is true if one of x_i^* or x_j^* is larger than or equal to 0.5 .
  - Again, this is also true since x^* is a feasible solution of the LP relaxation, which implies that $x_i^* + x_j^* \geq 1$. It follows that one of x_i^* or x_j^* is larger than or equal to 0.5 .
  - Then, either one of x_i' or x_j' is one.

- **Claim.** The weight of the vertex cover we get from rounding is at most twice as large as the minimum weight vertex cover.
  - Notice that
  $$x'_i = \min(\lfloor 2x^*_i \rfloor, 1)$$
  - Let P^*_{IP} be the optimal solution for IP, and P^*_{LP} be the optimal solution for the LP relaxation
  - Since any feasible solution for IP is also feasible in LP,
  $$P^*_{LP} \leq P^*_{IP}$$
  - The rounded solution x' satisfy

$$\begin{aligned} \sum_i x'_i w_i &= \sum_i \min(\lfloor 2x^*_i \rfloor, 1) w_i && \text{(by construction)} \\ &\leq \sum_i 2x^*_i w_i && \text{(by definition)} \\ &= 2P^*_{LP} && \text{(by definition)} \\ &\leq 2P^*_{IP} && \text{(since relaxtion increases the feasible set)} \end{aligned}$$

- Hence, $OPT_{LP} \leq 2OPT_{IP}$. LP relaxation of the minimum weight vertex cover is a *2-approximation* with the *approximation ratio* of two.

- Solving LP can be quite time consuming
- A faster approach using dual LP

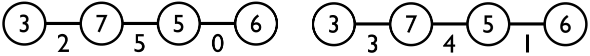
- LP relaxation

$$\begin{aligned} \text{minimize} \quad & \sum_{i \in V} x_i w_i \\ \text{subject to} \quad & x_i + x_j \geq 1, \ \forall (i,j) \in E \\ & x_i \geq 0 \,, \forall i \in V \end{aligned}$$

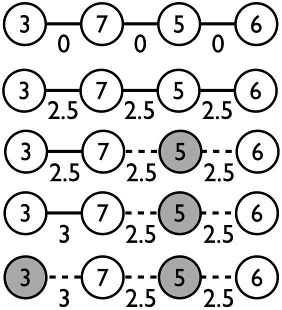
- objective: maximize $c^T x$ with $c_i = -w_i$
- constraint: $Ax \leq b$ with $A_{e,i} = -1$ if $i \in e$.
- $b_e = -1$

- Dual LP

$$\begin{aligned} \text{maximize} \quad & \sum_{(i,j)\in E} \ell_{ij} \\ \text{subject to} \quad & \sum_{j:(i,j)\in E} \ell_{ij} \le w_i,\ \forall i \in V \\ & \ell_{ij} \ge 0\ ,\forall (i,j) \in E \end{aligned}$$



- Algorithm: Primal/Dual approach



1. Initially set all ℓ_{ij} to 0 and all x_i to 0. And unfreeze all edges.
2. Uniformly increment all unfrozen ℓ_{ij}'s until for some node i we hit dual constraint $\sum_{j:(i,j)\in E} \ell_{ij} \le w_i$.
3. Freeze edges adjacent to the newly saturated node i.
4. Set $x_i = 1$
5. While there are still unfrozen edges, go back to step 2

Minimum weight vertex cover

- **Claim.** The resulting x_i's are primal feasible: $x_i + x_j \geq 1$, $\forall (i,j) \in E$.
  - The algorithm continues until all edges are frozen
  - edges are frozen only if one of the end node is saturated, and hence $x_i = 1$
- **Claim.** The weight of the vertex cover found using primal/dual approach is at most twice the value of the minimum weight vertex cover.
  - Let S be the vertex cover found using primal/dual approach.
  - Let ℓ^*_{ij} be the solution of primal/dual approach.
  - Then, for all $i \in S$ (i.e. for all the frozen nodes that are saturated),

$$w_i = \sum_{j:(i,j)\in E} \ell^*_{ij}$$

$$\sum_{i\in S} w_i = \sum_{i\in S} \sum_{j:(i,j)\in E} \ell^*_{ij}$$

  - Each ℓ_{ij} term can appear at most twice in the summation

$$\sum_{i\in S} \sum_{j:(i,j)\in E} \ell^*_{ij} \leq 2 \sum_{(i,j)\in E} \ell^*_{ij}$$

- By weak duality theorem, a dual feasible ℓ^* of the maximization problem (dual LP) is a lower bound on any primal feasible x

$$\sum_{(i,j)\in E} \ell^*_{ij} \leq \sum_i x_i w_i$$

- Also, since any feasible solution of the primal IP (x^*) is a feasible solution for the LP relaxation (x)

$$\sum_i x_i w_i \leq \sum_i x^*_i w_i$$

- putting it all together

$$\sum_{i\in S} w_i = \sum_{i\in S} \sum_{j:(i,j)\in E} \ell_{ij} \leq 2 \sum_{(i,j)\in E} \ell_{ij} \leq 2 \sum_i x_i w_i \leq 2 \sum_i x^*_i w_i$$

- Hence, again primal/dual algorithm is 2-approximation algorithm of minimum weight vertex cover problem.

there are implementations of the primal/dual algorithm with rum-time linear in the number of edges

Clarkson's greedy algorithm

Input: graph $G = (V, E)$ and weights w on V

Output: vertex cover C

- **for all** $i \in V$ set $W_i \leftarrow w_i$
- **for all** $i \in V$ set $D_i \leftarrow d_i$, which is the degree of node i
- initialize $C \leftarrow \{\}$
- **while** $E \neq \emptyset$ **do**

  $i = \arg\min_{v \in V \setminus C} \frac{W_v}{D_v}$

  for all neighbors $j \in N(i)$, do

  $E \leftarrow E \setminus (i, j)$

  $W_j \leftarrow W_j - \frac{W_i}{D_i}$

  $D_j \leftarrow D_j - 1$

  end for.

  $C \leftarrow C \sup\{i\}$
- **end while.**
- **output** C

Exercise

In a facility location problem, there is a set of facilities and a set of cities, and our job is to choose a subset of facilities to open, and to connect every city to one of the open facilities. There is a nonnegative cost f_j for opening a facility j, and a nonnegative connection cost $c_{i,j}$ for connecting city i to facility j. Given these as input, we look for a solution that minimizes the total cost. Formulate this facility location problem as an integer programming problem, and find its linear programming relaxation.

Homework 6

- Problem 1. [Problem 7.29 from Algorithms by Dasupta, Papadimitriou, and Vazirani]

  A film producer is seeking actors and investors for his new movie. There are n available actors; actor i charges s_i dollars. for funding, there are m available investors. Investor j will provide p_j dollars, but only on the condition that certain actors $L_j \subseteq \{1, \ldots, n\}$ are included in the cast (all of these actors L_j must be chosen in order to receive funding from investor j).

  The producer's profit is the sum of the payments from investors minus the payments to actors. The goal is to maximize this profit.

  (a) Express this problem as an integer program in which the variables take on values in $\{0, 1\}$. [hint: an inequality constraint $x \le \min\{y_1, y_2, \ldots, y_k\}$ is equivalent as k linear inequality constraints $x \le y_1$, $x \le y_2$, $\ldots$, $x \le y_k$]

  (b) Now write the linear program relaxation of this problem.

  (c) Explain that there must in fact be an *integral* optimal solution.

  (d) Write the dual linear program of the LP relaxation found in part (b).

Homework 6

- Problem 2.

  In the most general case of linear programming, we have a set I of inequalities and a set E of equality constraints over n variables (Notice E is a set of integers between 1 and n, and not a set of edges as we are used to using it.). Among those n variables, a subset of N are constrained to be non-negative. The dual has $m = |I| + |E|$ variables, of which only a subset are constrained to be non-negative.

$$\begin{array}{ll} \text{maximize} & a_1 z_1 + \cdots + a_n z_n \\ \text{subject to} & D_{i1} z_1 + \cdots + D_{in} z_n \leq d_i, \ \ \forall i \in I \\ & D_{j1} z_1 + \cdots + D_{jn} z_n = d_j, \ \ \forall j \in E \\ & z_k \geq 0, \ \ \forall k \in N \end{array}$$

  Recall the canonical form of LP and its dual we studied in class:

★ (canonical) Primal LP

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b \\ & x \geq 0 \end{array}$$

★ (canonical) Dual LP

$$\begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \geq c \\ & y \geq 0 \end{array}$$

Homework 6

- Problem 2. (continued)

  (a) Rewrite the general LP in the form of (canonical) Primal LP. What is c, A, b, and x in terms of a, D, d, and z?

  (b) Write the (canonical) Dual LP of the problem found in part (a).

  (c) Simplify the problem found in part (b) as

$$\begin{array}{rl} \text{minimize} & g_1 w_1 + \cdots + g_m w_m \\ \text{subject to} & F_{1i} w_1 + \cdots + F_{mi} w_m \geq f_i, \ \ \forall i \in N \\ & F_{1j} w_1 + \cdots + F_{mj} w_m = f_j, \ \ \forall j \notin N \\ & w_k \geq 0, \ \ \forall k \in I \end{array}$$

  In other words, write g, F, f in terms of a, D, d.

Homework 6

Problem 3. (The dual of maximum flow)

In the maximum flow problem, we are given a directed graph $G = (V, E)$ with a source node s and a sink node t. Each edge $(i, j) \in E$ is associated with a capacity c_{ij}. A flow consists of a vector valued variable $f = \{f_{ij}\}_{(i,j)\in E}$, satisfying the capacity condition ($0 \le f_{ij} \le c_{ij}$ for all edges) and conservation condition (total incoming flow is equal to the total outgoing flow at each node that is not s or t). The value of the flow is the total quantity leaving the source (or equivalently arriving the sink):

$$\mathit{size}(f) = \sum_{i:(s,i)\in E} f_{si}$$

This can be formulated as a linear program:

$$\begin{array}{ll} \text{maximize} & \mathit{size}(f) \\ \text{subject to} & 0 \le f_{ij} \le c_{ij}, \ \ \forall (i,j) \in E \\ & \displaystyle\sum_{k:(k,i)\in E} f_{ki} = \sum_{k:(i,k)\in E} f_{ik}, \ \ \forall i \notin \{s,t\} \end{array}$$

Homework 6

Problem 3. (continued)

Consider a general directed network $G = (V, E)$, with edge capacities c_{ij}'s.

(a) Write down the dual of the general flow LP above. Use a variable y_{ij} for each directed edge (i, j), and x_i for each node $i \notin \{s, t\}$.

(b) Show that any solution to the general dual LP must satisfy the following property: for any directed path from s to t in the network, the sum of the y_{ij} values along the path must be at least 1.

(c) What are the intuitive meaning of the dual variables? Show that any $s - t$ cut in the network can be translated into a dual feasible solution whose cost is exactly the capacity of that cut. More precisely, given a $s - t$ cut, construct y_{ij}'s and x_i's from the cut, such that the dual variables satisfy all the constraints in the dual LP.