Lecture 12: Applications of LP Duality

Rafael Oliveira

University of Waterloo
Cheriton School of Computer Science

rafael.oliveira.teaching@gmail.com

June 20, 2023

Overview

- Game Theory - Minimax Theorems
- Learning Theory - Boosting
- Combinatorics - Bipartite Matching
- Conclusion
- Acknowledgements

Ads

DIVERSE CAREERS IN COMPUTING
PANEL AND NETWORKING SESSION
The endless opportunities in tech can be quite overwhelming! UW WiCS will be hosting a Careers in Tech workshop with panelists who work in a variety of applications of computer science. We have speakers who can talk about computing careers in security, data science, software development, trading, and more!
Apply your software skill to
• Cybersecurity (OpenText)
• Medical devices (OrientaMed startup)
• Bank technology analytics (Scotiabank)
• VR/XR (Unity Technologies)
• Trading (HRT)
The Panel will be followed by a reception where attendees can talk with panelists over appetizers
Please register through Eventbrite to attend!
Open to ALL students!
22 JUNE 2023
4:00PM-6:00PM
DC 1302 - PANEL
Register using the QR code!
WiCS
WOMEN IN COMPUTER SCIENCE

More Ads


MATHINNOVATION
ENTREPRENEURSHIP AND IMPACT SERIES
ASK ME ANYTHING WITH MAYA ACKERMAN
JUNE 20, 2023 | 4:00 PM | DC 1302
BMATH '06, MMATH '07, PHD '12
Renowned researcher in generative AI and CEO & Co-founder of musical AI startup, WaveAI.
REGISTER HERE!
4:00 ASK ME ANYTHING
NETWORKING AND REFRESHMENTS TO FOLLOW
UNIVERSITY OF WATERLOO
FACULTY OF MATHEMATICS

Two-player games

Setup:

- Two players (Alice and Bob)
- Each player has a (finite) set of strategies $S_A = \{1, \ldots, m\}$ and $S_B = \{1, \ldots, n\}$

Two-player games

Setup:

- Two players (Alice and Bob)
- Each player has a (finite) set of strategies $S_A = \{1, \ldots, m\}$ and $S_B = \{1, \ldots, n\}$
- Payoff matrices $A, B \in \mathbb{R}^{m \times n}$ for Alice and Bob, respectively

Two-player games

Setup:

- Two players (Alice and Bob)
- Each player has a (finite) set of strategies $S_A = \{1, \ldots, m\}$ and $S_B = \{1, \ldots, n\}$
- Payoff matrices $A, B \in \mathbb{R}^{m \times n}$ for Alice and Bob, respectively
    - If Alice plays i and Bob plays j, then
    - Alice gets A_{ij}
    - Bob gets B_{ij}

Two-player games

Setup:

- Two players (Alice and Bob)
- Each player has a (finite) set of strategies $S_A = \{1, \ldots, m\}$ and $S_B = \{1, \ldots, n\}$
- Payoff matrices $A, B \in \mathbb{R}^{m \times n}$ for Alice and Bob, respectively
  - If Alice plays i and Bob plays j, then
  - Alice gets A_{ij}
  - Bob gets B_{ij}
- Example: *battle of the sexes* game

Two-player games

Setup:

- Two players (Alice and Bob)
- Each player has a (finite) set of strategies $S_A = \{1, \ldots, m\}$ and $S_B = \{1, \ldots, n\}$
- Payoff matrices $A, B \in \mathbb{R}^{m \times n}$ for Alice and Bob, respectively
    - If Alice plays i and Bob plays j, then
    - Alice gets A_{ij}
    - Bob gets B_{ij}
- Example: *battle of the sexes* game

	Football	Opera
Football	(2,1)	(0,0)
Opera	(0,0)	(1,2)

Table: Battle of the sexes payoff matrices

Nash Equilibrium

Assuming players are rational, i.e. want to maximize their payoffs, we have:

Definition (Nash Equilibrium)

A strategy profile (i, j) is called a Nash equilibrium if the strategy played by each player is optimal, *given the strategy of the other player*. That is:

1. $A_{ij} \geq A_{kj}$ for all $k \in S_A$
2. $B_{ij} \geq B_{i\ell}$ for all $\ell \in S_B$

Nash Equilibrium

Assuming players are rational, i.e. want to maximize their payoffs, we have:

Definition (Nash Equilibrium)

A strategy profile (i, j) is called a Nash equilibrium if the strategy played by each player is optimal, *given the strategy of the other player*. That is:

1. $A_{ij} \geq A_{kj}$ for all $k \in S_A$
2. $B_{ij} \geq B_{i\ell}$ for all $\ell \in S_B$

	Football	Opera
Football	(2,1)	(0,0)
Opera	(0,0)	(1,2)

Table: Battle of the sexes payoff matrices

Nash Equilibrium

Assuming players are rational, i.e. want to maximize their payoffs, we have:

Definition (Nash Equilibrium)

A strategy profile (i, j) is called a Nash equilibrium if the strategy played by each player is optimal, *given the strategy of the other player*. That is:

1. $A_{ij} \geq A_{kj}$ for all $k \in S_A$
2. $B_{ij} \geq B_{i\ell}$ for all $\ell \in S_B$

	Football	Opera
Football	(2,1)	(0,0)
Opera	(0,0)	(1,2)

Table: Battle of the sexes payoff matrices

	Silent	Snitch
Silent	(-1,-1)	(-10,0)
Snitch	(0,-10)	(-5,-5)

Table: Prisoner's dilemma

Mixed Strategies

Definition (Mixed Strategy)

A mixed strategy is a probability distribution over a set of pure strategies S. If Alice's strategies are $S_A = \{1, \ldots, n\}$, her mixed strategies are:

$$\Delta_A := \{x \in \mathbb{R}^n \mid x \geq 0 \text{ and } \|x\|_1 = 1\}$$

Mixed Strategies

Definition (Mixed Strategy)

A mixed strategy is a probability distribution over a set of pure strategies S. If Alice's strategies are $S_A = \{1, \ldots, n\}$, her mixed strategies are:

$$\Delta_A := \{x \in \mathbb{R}^n \mid x \geq 0 \text{ and } \|x\|_1 = 1\}$$

- Models situation where players choose their strategy "at random"

Mixed Strategies

Definition (Mixed Strategy)

A mixed strategy is a probability distribution over a set of pure strategies S. If Alice's strategies are $S_A = \{1, \ldots, n\}$, her mixed strategies are:

$$\Delta_A := \{x \in \mathbb{R}^n \mid x \geq 0 \text{ and } \|x\|_1 = 1\}$$

- Models situation where players choose their strategy "at random"
- Payoffs for each player defined as *expected gain*. That is, (x, y) is the profile of mixed strategies used by Alice and Bob, we have:

Mixed Strategies

Definition (Mixed Strategy)

A mixed strategy is a probability distribution over a set of pure strategies S. If Alice's strategies are $S_A = \{1, \ldots, n\}$, her mixed strategies are:

$$\Delta_A := \{x \in \mathbb{R}^n \mid x \geq 0 \text{ and } \|x\|_1 = 1\}$$

- Models situation where players choose their strategy "at random"
- Payoffs for each player defined as *expected gain*. That is, (x, y) is the profile of mixed strategies used by Alice and Bob, we have:

$$v_A(x, y) = \sum_{(i,j) \in S_A \times S_B} A_{ij} x_i y_j = x^T A y$$

$$v_B(x, y) = \sum_{(i,j) \in S_A \times S_B} B_{ij} x_i y_j = x^T B y$$

Nash Equilibrium Mixed Strategies

Assuming players are rational, i.e. want to maximize their payoffs, we have:

Definition ((Mixed) Nash Equilibrium)

A strategy profile $x \in \Delta_A, y \in \Delta_B$ is called a (mixed) Nash equilibrium if the strategy played by each player is optimal, *given the strategy of the other player*. That is:

1. $x^T Ay \geq z^T Ay$ for all $z \in \Delta_A$
2. $x^T By \geq x^T Bw$ for all $w \in \Delta_B$

Nash Equilibrium Mixed Strategies

Assuming players are rational, i.e. want to maximize their payoffs, we have:

Definition ((Mixed) Nash Equilibrium)

A strategy profile $x \in \Delta_A, y \in \Delta_B$ is called a (mixed) Nash equilibrium if the strategy played by each player is optimal, *given the strategy of the other player*. That is:

1. $x^T Ay \geq z^T Ay$ for all $z \in \Delta_A$
2. $x^T By \geq x^T Bw$ for all $w \in \Delta_B$

	Jump left	Jump right
kick left	(-1,1)	(1,-1)
kick right	(1,-1)	(-1,1)

Table: Penalty Kick

Nash Equilibrium Mixed Strategies

Assuming players are rational, i.e. want to maximize their payoffs, we have:

Definition ((Mixed) Nash Equilibrium)

A strategy profile $x \in \Delta_A, y \in \Delta_B$ is called a (mixed) Nash equilibrium if the strategy played by each player is optimal, *given the strategy of the other player*. That is:

1. $x^T Ay \geq z^T Ay$ for all $z \in \Delta_A$
2. $x^T By \geq x^T Bw$ for all $w \in \Delta_B$

	Jump left	Jump right
kick left	(-1,1)	(1,-1)
kick right	(1,-1)	(-1,1)

Table: Penalty Kick

- *Zero-Sum Game*: payoff matrices satisfy $A = -B$

Nash Equilibrium Mixed Strategies

Assuming players are rational, i.e. want to maximize their payoffs, we have:

Definition ((Mixed) Nash Equilibrium)

A strategy profile $x \in \Delta_A, y \in \Delta_B$ is called a (mixed) Nash equilibrium if the strategy played by each player is optimal, *given the strategy of the other player*. That is:

1. $x^T Ay \geq z^T Ay$ for all $z \in \Delta_A$
2. $x^T By \geq x^T Bw$ for all $w \in \Delta_B$

	Jump left	Jump right
kick left	(-1,1)	(1,-1)
kick right	(1,-1)	(-1,1)

Table: Penalty Kick

- *Zero-Sum Game*: payoff matrices satisfy $A = -B$
- No pure Nash Equilibrium!

Nash Equilibrium Mixed Strategies

Assuming players are rational, i.e. want to maximize their payoffs, we have:

Definition ((Mixed) Nash Equilibrium)

A strategy profile $x \in \Delta_A, y \in \Delta_B$ is called a (mixed) Nash equilibrium if the strategy played by each player is optimal, *given the strategy of the other player*. That is:

1. $x^T Ay \geq z^T Ay$ for all $z \in \Delta_A$
2. $x^T By \geq x^T Bw$ for all $w \in \Delta_B$

	Jump left	Jump right
kick left	(-1,1)	(1,-1)
kick right	(1,-1)	(-1,1)

Table: Penalty Kick

- *Zero-Sum Game*: payoff matrices satisfy $A = -B$
- No pure Nash Equilibrium!
- One mixed Nash equilibrium: $x = y = (1/2, 1/2)$

Von Neumann's Minimax Theorem

Theorem

In a zero-sum game, for any payoff matrix $A \in \mathbb{R}^{m \times n}$*:*

$$\max_{x \in \Delta_A} \min_{y \in \Delta_B} x^T A y = \min_{y \in \Delta_B} \max_{x \in \Delta_A} x^T A y$$

Von Neumann's Minimax Theorem

Theorem

In a zero-sum game, for any payoff matrix $A \in \mathbb{R}^{m \times n}$:

$$\max_{x \in \Delta_A} \min_{y \in \Delta_B} x^T A y = \min_{y \in \Delta_B} \max_{x \in \Delta_A} x^T A y$$

For given $x \in \Delta_A$:

$$\min_{y \in \Delta_B} x^T A y = \min_{j \in S_B} (x^T A)_j$$

Von Neumann's Minimax Theorem

Theorem

In a zero-sum game, for any payoff matrix $A \in \mathbb{R}^{m \times n}$:

$$\max_{x \in \Delta_A} \min_{y \in \Delta_B} x^T A y = \min_{y \in \Delta_B} \max_{x \in \Delta_A} x^T A y$$

For given $x \in \Delta_A$:

$$\min_{y \in \Delta_B} x^T A y = \min_{j \in S_B} (x^T A)_j$$

Left hand side can be written as

$$\begin{aligned} \max \quad & s \\ \text{s.t.} \quad & s \leq (x^T A)_j \quad \text{for } j \in S_B \\ & \sum_{i \in S_A} x_i = 1 \\ & x \geq 0 \end{aligned}$$

Von Neumann's Minimax Theorem

Theorem

In a zero-sum game, for any payoff matrix $A \in \mathbb{R}^{m \times n}$:

$$\max_{x \in \Delta_A} \min_{y \in \Delta_B} x^T A y = \min_{y \in \Delta_B} \max_{x \in \Delta_A} x^T A y$$

For given $x \in \Delta_A$:

$$\min_{y \in \Delta_B} x^T A y = \min_{j \in S_B} (x^T A)_j$$

For given $y \in \Delta_B$:

$$\max_{x \in \Delta_A} x^T A y = \max_{i \in S_A} (Ay)_i$$

Left hand side can be written as

$$\begin{aligned} \max \quad & s \\ \text{s.t.} \quad & s \leq (x^T A)_j \quad \text{for } j \in S_B \\ & \sum_{i \in S_A} x_i = 1 \\ & x \geq 0 \end{aligned}$$

Von Neumann's Minimax Theorem

Theorem

In a zero-sum game, for any payoff matrix $A \in \mathbb{R}^{m \times n}$:

$$\max_{x \in \Delta_A} \min_{y \in \Delta_B} x^T A y = \min_{y \in \Delta_B} \max_{x \in \Delta_A} x^T A y$$

For given $x \in \Delta_A$:

$$\min_{y \in \Delta_B} x^T A y = \min_{j \in S_B} (x^T A)_j$$

Left hand side can be written as

$$\begin{aligned} \max \quad & s \\ \text{s.t.} \quad & s \leq (x^T A)_j \quad \text{for } j \in S_B \\ & \sum_{i \in S_A} x_i = 1 \\ & x \geq 0 \end{aligned}$$

For given $y \in \Delta_B$:

$$\max_{x \in \Delta_A} x^T A y = \max_{i \in S_A} (Ay)_i$$

Right hand side can be written as

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & t \geq (Ay)_i \quad \text{for } i \in S_A \\ & \sum_{j \in S_B} y_j = 1 \\ & y \geq 0 \end{aligned}$$

Proof of Duality

- Game Theory - Minimax Theorems
- Learning Theory - Boosting
- Combinatorics - Bipartite Matching
- Conclusion
- Acknowledgements

Learning Theory

Consider classification problem over $\mathcal{X}$:

- Set of hypothesis $\mathcal{H} := \{h : \mathcal{X} \to \{0, 1\}\}$

Learning Theory

Consider classification problem over $\mathcal{X}$:

- Set of hypothesis $\mathcal{H} := \{h : \mathcal{X} \to \{0, 1\}\}$
- Each $x \in \mathcal{X}$ has a correct value $c(x) \in \{0, 1\}$

Learning Theory

Consider classification problem over $\mathcal{X}$:

- Set of hypothesis $\mathcal{H} := \{h : \mathcal{X} \to \{0, 1\}\}$
- Each $x \in \mathcal{X}$ has a correct value $c(x) \in \{0, 1\}$
- Data is sampled from unknown distribution $q \in \Delta_{\mathcal{X}}$

Learning Theory

Consider classification problem over $\mathcal{X}$:

- Set of hypothesis $\mathcal{H} := \{h : \mathcal{X} \to \{0, 1\}\}$
- Each $x \in \mathcal{X}$ has a correct value $c(x) \in \{0, 1\}$
- Data is sampled from unknown distribution $q \in \Delta_{\mathcal{X}}$
- *Weak learning assumption*:

For any distribution $q \in \Delta_{\mathcal{X}}$, there is a hypothesis $h \in \mathcal{H}$ which is wrong less than half the time.

$$\exists \gamma > 0,\ \forall q \in \Delta_{\mathcal{X}},\ \exists h \in \mathcal{H}, \quad \Pr_{x \sim q}[h(x) \neq c(x)] \leq \frac{1-\gamma}{2}$$

Learning Theory

Consider classification problem over $\mathcal{X}$:

- Set of hypothesis $\mathcal{H} := \{h : \mathcal{X} \to \{0, 1\}\}$
- Each $x \in \mathcal{X}$ has a correct value $c(x) \in \{0, 1\}$
- Data is sampled from unknown distribution $q \in \Delta_{\mathcal{X}}$
- *Weak learning assumption*:

  For any distribution $q \in \Delta_{\mathcal{X}}$, there is a hypothesis $h \in \mathcal{H}$ which is wrong less than half the time.

  $$\exists \gamma > 0,\ \forall q \in \Delta_{\mathcal{X}},\ \exists h \in \mathcal{H}, \quad \Pr_{x \sim q}[h(x) \neq c(x)] \leq \frac{1 - \gamma}{2}$$

- Surprisingly, weak learning assumption implies something much stronger: it is possible to *combine* classifiers in $\mathcal{H}$ to construct a *classifier* that is *always right* (known as *strong learning*).

Boosting

Theorem

Let $\mathcal{H}$ be a set of hypotheses satisfying weak learning assumption. Then there is distribution $p \in \Delta_{\mathcal{H}}$ such that the weighed majority classifier

$$c_p(x) := \begin{cases} 1, & \text{if } \sum_{h \in \mathcal{H}} p_h \cdot h(x) \geq 1/2 \\ 0, & \text{otherwise} \end{cases}$$

is always correct. That is, $c_p(x) = c(x)$ for all $x \in \mathcal{X}$

Boosting

Theorem

Let $\mathcal{H}$ be a set of hypotheses satisfying weak learning assumption. Then there is distribution $p \in \Delta_{\mathcal{H}}$ such that the weighed majority classifier

$$c_p(x) := \begin{cases} 1, & \text{if } \sum_{h \in \mathcal{H}} p_h \cdot h(x) \geq 1/2 \\ 0, & \text{otherwise} \end{cases}$$

is always correct. That is, $c_p(x) = c(x)$ for all $x \in \mathcal{X}$

- Let $M \in \{-1, 1\}^{m \times n}$, where $m = |\mathcal{X}|$ and $n = |\mathcal{H}|$.

$$M_{ij} = \begin{cases} +1, & \text{if classifier } h_j \text{ wrong on } x_i \\ -1, & \text{otherwise} \end{cases}$$

Boosting

Theorem

Let $\mathcal{H}$ be a set of hypotheses satisfying weak learning assumption. Then there is distribution $p \in \Delta_{\mathcal{H}}$ such that the weighed majority classifier

$$c_p(x) := \begin{cases} 1, & \text{if } \sum_{h \in \mathcal{H}} p_h \cdot h(x) \geq 1/2 \\ 0, & \text{otherwise} \end{cases}$$

is always correct. That is, $c_p(x) = c(x)$ for all $x \in \mathcal{X}$

- Let $M \in \{-1, 1\}^{m \times n}$, where $m = |\mathcal{X}|$ and $n = |\mathcal{H}|$.

$$M_{ij} = \begin{cases} +1, & \text{if classifier } h_j \text{ wrong on } x_i \\ -1, & \text{otherwise} \end{cases}$$

- Weak learning:

$$\sum_{1 \leq i \leq n} q_j \cdot \delta_{h_j(x_i) \neq c(x_i)} \leq \frac{1-\gamma}{2}$$

Boosting - Proof

Let $M \in \{-1, 1\}^{m \times n}$,
where $m = |\mathcal{X}|$ and $n = |\mathcal{H}|$.

$$M_{ij} = \begin{cases} +1, & \text{if } h_j \text{ wrong on } x_i \\ -1, & \text{otherwise} \end{cases}$$

Weak learning:

$$\sum_{1 \leq i \leq n} q_j \cdot \delta_{h_j(x_i) \neq c(x_i)} \leq \frac{1 - \gamma}{2}$$

Boosting - Proof

Let $M \in \{-1,1\}^{m \times n}$,
where $m = |\mathcal{X}|$ and $n = |\mathcal{H}|$.

$$M_{ij} = \begin{cases} +1, & \text{if } h_j \text{ wrong on } x_i \\ -1, & \text{otherwise} \end{cases}$$

Weak learning:

$$\sum_{1 \le i \le n} q_j \cdot \delta_{h_j(x_i) \neq c(x_i)} \le \frac{1-\gamma}{2}$$

- Note that $M_{ij} = 2 \cdot \delta_{h_j(x_i) \neq c(x_i)} - 1$

$$q^T M e_j \le -\gamma \quad \Rightarrow \quad q^T M p \le -\gamma$$

for any $p \in \Delta_{\mathcal{H}}$.

Boosting - Proof

Let $M \in \{-1, 1\}^{m \times n}$,
where $m = |\mathcal{X}|$ and $n = |\mathcal{H}|$.

$$M_{ij} = \begin{cases} +1, & \text{if } h_j \text{ wrong on } x_i \\ -1, & \text{otherwise} \end{cases}$$

Weak learning:

$$\sum_{1 \leq i \leq n} q_j \cdot \delta_{h_j(x_i) \neq c(x_i)} \leq \frac{1-\gamma}{2}$$

- Note that $M_{ij} = 2 \cdot \delta_{h_j(x_i) \neq c(x_i)} - 1$

$$q^T M e_j \leq -\gamma \quad \Rightarrow \quad q^T M p \leq -\gamma$$

  for any $p \in \Delta_{\mathcal{H}}$.
- By minimax, we have:

$$\max_{q \in \Delta_{\mathcal{X}}} \min_{p \in \Delta_{\mathcal{H}}} q^T M p = \min_{p \in \Delta_{\mathcal{H}}} \max_{q \in \Delta_{\mathcal{X}}} q^T M p \leq -\gamma$$

Boosting - Proof

Let $M \in \{-1,1\}^{m\times n}$,
where $m = |\mathcal{X}|$ and $n = |\mathcal{H}|$.

$$M_{ij} = \begin{cases} +1, & \text{if } h_j \text{ wrong on } x_i \\ -1, & \text{otherwise} \end{cases}$$

Weak learning:

$$\sum_{1\le i\le n} q_j \cdot \delta_{h_j(x_i)\neq c(x_i)} \le \frac{1-\gamma}{2}$$

- Note that $M_{ij} = 2\cdot\delta_{h_j(x_i)\neq c(x_i)} - 1$

$$q^T M e_j \le -\gamma \quad \Rightarrow \quad q^T M p \le -\gamma$$

  for any $p \in \Delta_{\mathcal{H}}$.
- By minimax, we have:

$$\max_{q\in\Delta_{\mathcal{X}}} \min_{p\in\Delta_{\mathcal{H}}} q^T M p = \min_{p\in\Delta_{\mathcal{H}}} \max_{q\in\Delta_{\mathcal{X}}} q^T M p \le -\gamma$$

- In particular, right hand side implies weighted classifier given by optimum solution p^* *always* correct.

Proof of Correctness of Classifier

- Game Theory - Minimax Theorems
- Learning Theory - Boosting
- Combinatorics - Bipartite Matching
- Conclusion
- Acknowledgements

Bipartite Matching

- Given a bipartite graph $G(L \sqcup R, E)$, does it have a perfect matching?

Bipartite Matching

- Given a bipartite graph $G(L \sqcup R, E)$, does it have a perfect matching?
- We saw in lecture 7 that we can randomly isolate a perfect matching, if one exists

Bipartite Matching

- Given a bipartite graph $G(L \sqcup R, E)$, does it have a perfect matching?
- We saw in lecture 7 that we can randomly isolate a perfect matching, if one exists
- Can we remove the randomness in that process? This would lead to a fast parallel algorithm for matching.

Bipartite Matching

- Given a bipartite graph $G(L \sqcup R, E)$, does it have a perfect matching?
- We saw in lecture 7 that we can randomly isolate a perfect matching, if one exists
- Can we remove the randomness in that process? This would lead to a fast parallel algorithm for matching.
- Breakthrough result of [Fenner, Gurjar and Thierauf 2019]

Bipartite Matching

- Given a bipartite graph $G(L \sqcup R, E)$, does it have a perfect matching?
- We saw in lecture 7 that we can randomly isolate a perfect matching, if one exists
- Can we remove the randomness in that process? This would lead to a fast parallel algorithm for matching.
- Breakthrough result of [Fenner, Gurjar and Thierauf 2019]
- We will see just a piece of the proof

Bipartite Matching & Circulation

- Given an even cycle $C = (e_1, e_2, \ldots, e_{2k})$, we say that the *circulation* of C is given by

$$circ(C) = |w(e_1) - w(e_2) + \ldots + w(e_{2k-1}) - w(e_{2k})|$$

Bipartite Matching & Circulation

- Given an even cycle $C = (e_1, e_2, \ldots, e_{2k})$, we say that the *circulation* of C is given by

$$circ(C) = |w(e_1) - w(e_2) + \ldots + w(e_{2k-1}) - w(e_{2k})|$$

- Lemma: if we assign weights $w(e_i)$ such that $circ(C) \neq 0$ for each cycle C of the bipartite graph G, then we get that the minimum weight PM is unique!

Bipartite Matching & Circulation

- Given an even cycle $C = (e_1, e_2, \ldots, e_{2k})$, we say that the *circulation* of C is given by

$$circ(C) = |w(e_1) - w(e_2) + \ldots + w(e_{2k-1}) - w(e_{2k})|$$

- Lemma: if we assign weights $w(e_i)$ such that $circ(C) \neq 0$ for each cycle C of the bipartite graph G, then we get that the minimum weight PM is unique!
- The approach of [Fenner, Gurjar and Thierauf 2019] is to construct a set of weights which make all circulations non-zero!
  - To do that, they iteratively construct a weight assignment that kills small cycles (i.e., make their circulation non-zero)
  - Once we have a bipartite graph with no cycles of length $2k$, then in next iteration we kill cycles of length up to $4k$
  - show that no cycles of length $2k \Rightarrow$ *few cycles* of length $4k$ – similar to Karger's min cut lemma!

Bipartite Matching

- Suppose we have a weight assignment w. Let G_w be the subgraph of G given by the union of *all min w-weight perfect matchings in G*.

Bipartite Matching

- Suppose we have a weight assignment w. Let G_w be the subgraph of G given by the union of *all min w-weight perfect matchings in G*.
- Claim: circulation of each (even) cycle in G_w is zero

Bipartite Matching

- Suppose we have a weight assignment w. Let G_w be the subgraph of G given by the union of *all min w-weight perfect matchings in G*.
- Claim: circulation of each (even) cycle in G_w is zero
- Proof: LP duality! (complementary slackness)
  - Linear programs:

Primal

$$\begin{aligned} \text{min} \quad & \sum_{e \in E} w_e x_e \\ \text{s.t.} \quad & x \geq 0 \\ & \sum_{e \in \delta(u)} x_e = 1 \\ & \text{for } u \in L \sqcup R \end{aligned}$$

Bipartite Matching

- Suppose we have a weight assignment w. Let G_w be the subgraph of G given by the union of *all min w-weight perfect matchings in G*.
- Claim: circulation of each (even) cycle in G_w is zero
- Proof: LP duality! (complementary slackness)
  - Linear programs:

Primal

$$\begin{aligned} \min \quad & \sum_{e \in E} w_e x_e \\ \text{s.t.} \quad & x \geq 0 \\ & \sum_{e \in \delta(u)} x_e = 1 \\ & \text{for } u \in L \sqcup R \end{aligned}$$

Dual

$$\begin{aligned} \max \quad & \sum_{u \in L \sqcup R} y_u \\ \text{s.t.} \quad & y_u + y_v \leq w_e \\ & \text{for } e = \{u, v\} \in E \end{aligned}$$

  - Complementary slackness says $x_e \neq 0$ in primal, where $e = \{u, v\}$ $\Rightarrow y_u + y_v = w_e$ in dual optimal.

Bipartite Matching - Dual

Bipartite Matching - Circulation

Conclusion

- Mathematical programming - very general, and pervasive in Algorithmic life
- General mathematical programming very hard (how hard do you think it is?)
- Special cases have very striking applications!

Today and last lecture: *Linear Programming*

Conclusion

- Mathematical programming - very general, and pervasive in Algorithmic life
- General mathematical programming very hard (how hard do you think it is?)
- Special cases have very striking applications!

Today and last lecture: *Linear Programming*

- Linear Programming and Duality - fundamental concepts, lots of applications!
  - Applications in Combinatorial Optimization (a lot of it happened here at UW!)
  - Applications in Game Theory (minimax theorem)
  - Applications in Learning Theory (boosting)
  - Applications in Parallel Computation/Derandomization (Perfect Matching)
  - many more

Acknowledgement

- Lecture based largely on:
  - Lectures 3-6 of Yarom Singer's Advanced Optimization class
  - [Schrijver 1986, Chapter 7]
  - Personal Communication with Rohit
- See Yarom's notes at `https://people.seas.harvard.edu/~yaron/AM221-S16/schedule.html`

References I

Schirjver, Alexander (1986)
Theory of Linear and Integer Programming

Fourier, J. B. 1826
Analyse des travaux de l'Académie Royale des Sciences pendant l'année 1823.
Partie mathématique (1826)

Fourier, J. B. 1827
Analyse des travaux de l'Académie Royale des Sciences pendant l'année 1824.
Partie mathématique (1827)

Fenner, Stephen, and Gurjar, Rohit, and Thierauf, Thomas (2019)
Bipartite perfect matching is in quasi-NC
SIAM Journal on Computing, 2019