More Applications of LP and Duality
Outline for Today

• Three applications of LP and Duality
  – Markov Decision Processes
  – Optimal Strategies in Zero Sum Games
  – Shortest Paths
• Consider a directed graph G with nodes \{1,2,...,n\} of two types: decision (nondeterministic) and random, plus two sink nodes n-1 (lose) and n (win).

• A policy \(p\) is a choice of edge from each decision node.

• Fix a policy \(p\). A run of the process starting at a given node follows edges of the graph by choosing the policy edge from a decision node and choosing an edge uniformly at random from a random node. The run ends when a sink node is reached.

• For simplicity, assume that regardless of the policy, a sink node is always reached with probability 1
LP Application: Markov Decision Processes

- Let $w_p(i)$ the probability of winning if the run starts at node $i$, and policy $p$ is used.
- We can calculate the quantities $w_p(i)$ in polynomial time by solving a set of linear equations. (There is a unique solution.)
- An *optimal* policy maximizes the probability of winning.
LP Application: Markov Decision Processes

Facts:

• An optimal policy is independent of the starting node.

• More general types of policies that can make different decisions on different visits to a node, or "mixed" policies that assign probabilities to the alternatives, are no better than the type of policy we've defined.
LP Application: Markov Decision Processes

• MDP Maximization: Given G, find an optimal policy.

• MDP Decision: Does an optimal policy have probability of winning > ½ when the start node is node 1?
LP Application: Markov Decision Processes

Applications:

• Robot navigation with randomly moving obstacles
• Server management, other planning problems in the face of uncertainty
• Tetris, many other games
LP Application: Markov Decision Processes

MDP Generalizations

- Decision and random nodes are merged (easy to separate)
- Rewards (or costs) are associated with decisions
- Rewards further into the future are discounted
LP Application: Markov Decision Processes

How to find an optimal policy?
How to find an optimal policy?

- *Exhaustive search*: For each policy, calculate the winning probabilities and choose the best one.
LP Application: Markov Decision Processes

How to find an optimal policy?

• *Exhaustive search*: For each policy, calculate the winning probabilities and choose the best one.

• *Policy iteration*: Starting with an arbitrary policy, switch the decision at a node if it improves the winning probabilities. Repeat until no further improvement is possible. Just like Simplex, this "local search" approach yields a globally optimal policy.
Can we use linear programming?

• Observation: With respect to an optimal policy, the value of a decision node is at least the values of its children, and the value of a random node is the average of the values of its children.
LP Application: Markov Decision Processes

Use one variable $x_i$ per node, to represent the optimal probability of winning from node $i$

Constraints:

• At a random node $i$ we must have
  
  $x_i = \frac{\sum_{j \in \text{C}(i)} x_j}{|\text{C}(i)|}$, where $\text{C}(i)$ are the children of node $i$
  
  (i.e., $i$’s probability of winning is the average of its children’s probability of winning)

• At a decision node $i$ we must have
  
  $x_i \geq x_j$, for each child $j$ of $i$
  
  (i.e., $i$’s probability of winning is at least as great as that of its children; see previous observation)
Counterintuitively, we minimize $\sum x_i$, subject to these constraints, plus constraints on the sink nodes:

$$\begin{align*}
\text{minimize } & \sum x_i \\
\text{subject to } & \\
& x_i = (\sum_{j \in \text{C}(i)} x_j) / |\text{C}(i)|, \text{ for each random node } i \\
& x_i \geq x_j, \text{ for each child } j \text{ of a decision node } i \\
& x_{n-1} = 0, x_n = 1
\end{align*}$$

By minimizing, we force decision nodes’ winning probabilities to be equal to the max of the winning probabilities of their children.
LP Application: Shortest Paths
LP Application: Shortest Paths

- Consider a directed, graph $G$ with weight $w(e)$ for each edge $e$
- The *length* of a path is the sum of its weights
- A path of minimum length between two nodes $s$ and $t$ is a *shortest path*
LP Application: Shortest Paths

• Again, counterintuitively, the following linear program computes the length of the shortest path from s to t:

maximize $l_t$

subject to $l_s = 0$

$l_v - l_u \leq w(u,v)$ for every edge $(u,v)$ of $G$

($l_v$ represents the length of the shortest path from s to v)

Exercise: prove that this LP is correct
LP Application: Zero Sum Games
LP Application: Zero Sum Games

• Example:

\[
G = \begin{array}{c|cc}
& m & t \\ \hline
e & 3 & -1 \\ s & -2 & 1 \\
\end{array}
\]

• The game models a presidential election scenario; moves correspond to campaign issues that candidates can focus on (economy, society, morality, and tax cut)

• More generally, a game is represented by a matrix G

Dasgupta et al.
LP Application: Zero Sum Games

• A mixed strategy (x1,x2) for Max determines the probabilities of choosing e and s

• If Max's mixed strategy is (1/2,1/2), what is Min's best strategy?

• More generally, if Max's mixed strategy is (x1,x2), what is Min's payoff on pure strategy m? On pure strategy t?
LP Application: Zero Sum Games

• A mixed strategy \((x_1, x_2)\) for Max determines the probabilities of choosing e and s

• If Max's mixed strategy is \((1/2, 1/2)\), what is Min's best strategy?

  \[\text{Answer: The pure strategy } t\]

• More generally, if Max's mixed strategy is \((x_1, x_2)\), what is Min's payoff on pure strategy \(m\)? On pure strategy \(t\)?

  \[\text{Answers: } 3x_1 - 2x_2 \text{ and } -x_1 + x_2, \text{ respectively}\]
LP Application: Zero Sum Games

• Suppose Max announces its strategy \((x_1,x_2)\). The best payoff Max can expect is \(\min\{3x_1-2x_2, -x_1+x_2\}\)

• So, Max's best "defensive" strategy (the strategy with the best expected guarantee) is the strategy maximizes \(\min\{3x_1-2x_2, -x_1+x_2\}\).

• We want an LP that determines Max's best strategy
LP Application: Zero Sum Games

• LP to determine Max’s best strategy \((x_1, x_2)\)

\[
\begin{align*}
\text{max} & \quad z \\
-3x_1 + 2x_2 + z & \leq 0 \\
x_1 - x_2 + z & \leq 0 \\
x_1 + x_2 & = 1 \\
x_1, x_2 & \geq 0
\end{align*}
\]
LP Application: Zero Sum Games

- LP to determine Max’s best strategy \((x_1, x_2)\)
- LP to determine Min’s best strategy?

\[
\begin{align*}
\max & \quad z \\
-3x_1 + 2x_2 + z & \leq 0 \\
x_1 - x_2 + z & \leq 0 \\
x_1 + x_2 & = 1 \\
x_1, x_2 & \geq 0
\end{align*}
\]
LP Application: Zero Sum Games

- LP to determine Max’s best strategy \((x_1, x_2)\)
- LP to determine Min’s best strategy?

\[
\begin{align*}
\text{max} & \quad z \\
-3x_1 & + 2x_2 & + z & \leq 0 \\
x_1 & - x_2 & + z & \leq 0 \\
x_1 & + x_2 & &= 1 \\
x_1, x_2 & \geq 0
\end{align*}
\]

\[
\begin{align*}
\text{min} & \quad w \\
-3y_1 & + y_2 & + w & \geq 0 \\
2y_1 & - y_2 & + w & \geq 0 \\
y_1 & + y_2 & &= 1 \\
y_1, y_2 & \geq 0
\end{align*}
\]
LP Application: Zero Sum Games

- LP to determine Max’s best strategy \((x_1, x_2)\)
- LP to determine Min’s best strategy?
- These two LPs are duals of each other!

\[
\begin{align*}
\text{max} & \quad z \\
-3x_1 & + 2x_2 + z \leq 0 \\
x_1 & - x_2 + z \leq 0 \\
x_1 + x_2 & = 1 \\
x_1, x_2 & \geq 0 \\
\end{align*}
\]

\[
\begin{align*}
\text{min} & \quad w \\
-3y_1 & + y_2 + w \geq 0 \\
2y_1 & - y_2 + w \geq 0 \\
y_1 + y_2 & = 1 \\
y_1, y_2 & \geq 0 \\
\end{align*}
\]

Primal LP:

\[
\begin{align*}
\text{max} & \quad c_1x_1 + \cdots + c_nx_n \\
a_{i1}x_1 + \cdots + a_{in}x_n & \leq b_i \quad \text{for } i \in I \\
a_{i1}x_1 + \cdots + a_{in}x_n & = b_i \quad \text{for } i \in E \\
x_j & \geq 0 \quad \text{for } j \in N \\
\end{align*}
\]

Dual LP:

\[
\begin{align*}
\text{min} & \quad b_1y_1 + \cdots + b_my_m \\
a_{1j}y_1 + \cdots + a_{mj}y_m & \geq c_j \quad \text{for } j \in N \\
a_{1j}y_1 + \cdots + a_{mj}y_m & = c_j \quad \text{for } j \notin N \\
y_i & \geq 0 \quad \text{for } i \in I \\
\end{align*}
\]
LP Application: Zero Sum Games

Dual LPs have the same optimum value $V$. What does this tell us?

- Max has a strategy that guarantees an expected gain of at least $V$ (no matter what Min does)
- Min has a strategy can guarantee an expected loss at most $V$, (no matter what Max does)
- Both strategies can be determined by solving LPs that are dual to each other
Dual LPs have the same optimum value $V$. What does this tell us?

- Max has a strategy that guarantees an expected gain of at least $V$ (no matter what Min does)
- Min has a strategy can guarantee an expected loss at most $V$, (no matter what Max does)
- Both strategies can be determined by solving LPs that are dual to each other

In our example, $V = 1/7$; Max's optimal strategy is $(3/7, 4/7)$ and Min's is $(2/7, 5/7)$
Generalizing to an arbitrary zero-sum game $G$, we have the min-max theorem:

$$\max_x \min_y \sum_{i,j} G_{ij} x_i y_j = \min_y \max_x \sum_{i,j} G_{ij} x_i y_j$$