DS 320 Algorithms for Data Science	**Lecture #20**
Spring 2024	Prof. Kira Goldner

Linear Programming III: Duality Theory and Zero-Sum Games

Conditions for Optimality

Weak Duality

Theorem 1 (Weak Duality)**.** *If* $\mathbf{x}$ *is feasible in (P) and* $\mathbf{y}$ *is feasible in (D) then* $\mathbf{c}^T\mathbf{x} \leq \mathbf{b}^T\mathbf{y}$.

This theorem says that *any* feasible solution to the primal is a *lower bound* to *any* feasible solution to the dual, and likewise, any feasible solution to the dual is an *upper bound* to the primal.

That is, fractional vertex cover gives an upper bound on how large the (fractional) maximum matching can be, and likewise, fractional maximum matching gives a lower bound on how small the minimum (fractional) vertex cover can be.

Strong Duality

Strong duality states that everything in fact needs to hold with equality to be optimal.

Theorem 2 (Strong Duality)**.** *A pair of solutions* $(\mathbf{x}^*, \mathbf{y}^*)$ *are optimal for the primal and dual respectively if and only if* $\mathbf{c}^T\mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$.

Proof. ($\Leftarrow$) The *if* direction is easy to see: we know that the dual gives an upper bound on the primal, so if these objectives are equal, then the primal objective that we are trying to maximize could not possible get any larger, as it's always *at most* the dual's objective. This is *as tight as possible.*

($\Rightarrow$) The *only if* direction is harder to prove, and we'll skip it for now. □

Complementary Slackness

We rewrite the primal and dual with each constraint separated, and then formalize another condition for optimality called *complementary slackness*, which states that for each corresponding constraint and variable, at most one can be slack in an optimal solution.

Primal (P):

$$\begin{aligned} \max \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \sum_i a_{ji}x_i \leq b_j \quad \forall j \quad (y_j) \\ & x_i \geq 0 \quad \forall i \end{aligned}$$

Dual (D):

$$\begin{aligned} \min \quad & \mathbf{b}^T\mathbf{y} \\ \text{subject to} \quad & \sum_i a_{ij}y_i \geq c_i \quad \forall i \quad (x_i) \\ & y_j \geq 0 \quad \forall j \end{aligned}$$

Theorem 3 (Complementary Slackness). *A pair of solutions* $(\mathbf{x}^*, \mathbf{y}^*)$ *are optimal for the primal and dual respectively if and only if the following complementary slackness conditions (1) and (2) hold:*

$$\sum_i a_{ji}x_i = b_j \quad \text{or} \quad y_j = 0 \qquad (1) \qquad\qquad \sum_i a_{ij}y_i = c_i \quad \text{or} \quad x_i = 0. \qquad (2)$$

Proof. ($\Rightarrow$) According to complementary slackness, by rearranging our constraint, either $\sum_i a_{ji}x_i - b_j = 0$ or $y_j = 0$. This ensures that the multiplied quantity $(\sum_i a_{ji}x_i - b_j)\, y_j = 0$, as *one* of the two terms on the left-hand side must be 0. Then multiplying out and rearranging gives that $y_j \sum_i a_{ji}x_i = y_j b_j$. This process with all rows gives the equality from complementary slackness that $\mathbf{y}^T\mathbf{A}\mathbf{x} = \mathbf{y}^T\mathbf{b}$.

Similarly, using the condition that $\sum_i a_{ij}y_i = c_i$ or $x_i = 0$ gives that $\mathbf{c}^T\mathbf{x} = (\mathbf{A}^T\mathbf{y})\mathbf{x}$.

Then following our inequalities in the proof of weak duality, they now all hold with equality, so by Strong Duality, $(\mathbf{x}, \mathbf{y})$ are optimal solutions to the primal and dual.

$$\mathbf{c}^T\mathbf{x} \quad = \quad (\mathbf{A}^T\mathbf{y})\mathbf{x} \quad = \quad \mathbf{y}^T\mathbf{A}\mathbf{x} \quad = \quad \mathbf{y}^T\mathbf{b} \quad = \quad \mathbf{b}^T\mathbf{y}.$$

($\Leftarrow$) Similarly, if Strong Duality holds, the above inequalities hold with equality, in which case it must be that $y_j \sum_i a_{ji}x_i = y_j b_j$ for all j and $\sum_i a_{ij}y_i x_i = c_i x_i$ for all i, and hence that either $\sum_i a_{ji}x_i - b_j = 0$ or $y_j = 0$ for all j and that either $\sum_i a_{ij}y_i = c_i$ or $x_i = 0$ for all i. □

Approximation

Using Linear Programming for a Vertex Cover Approximation Algorithm

$$\begin{aligned} \min \sum_{i \in V} w_i x_i & & \\ \text{s.t.} \quad x_i + x_j \geq 1 & \qquad\qquad & (i,j) \in E \\ x_i \in [0,1] & & i \in V. \end{aligned}$$

Claim 1. Let S^* denote the optimal vertex cover of minimum weight, and let x^* denote the optimal solution to the Linear Program. Then $\sum_{i \in V} w_i x_i^* \leq w(S^*) = \text{OPT}$.

Proof. The vertex cover problem is equivalent to the integer program, whereas the linear program is a *relaxation*. Then there are simply more solutions allowed to the linear program, so the minimum can only be smaller. □

Claim 2. The set $S = \{i : x_i \geq 0.5\}$ is a vertex cover, and $w(S) \leq 2 \sum_{i \in V} w_i x_i^*$.

Proof. First, S is a vertex cover: for any edge $e = (i,j)$, at least one of i or j must be in S, because of our constraint $x_i + x_j \geq 1$, which forces at least one of these variables to be $\geq \frac{1}{2}$ and thus in S.

With respect to weight:

$$\sum_{i \in V} w_i x_i^* \geq \sum_{i \in S} w_i x_i^* \geq \frac{1}{2} \sum_{i \in S} w_i = \frac{1}{2} w(S).$$

□

Then our algorithm of running an LP and rounding it to give the vertex cover S is a 2-approximation to the optimal vertex cover S^*, as $w(S) \leq 2\, w(S^*)$ by Claims 1 and 2.

Zero-Sum Games and the Minimax Theorem

Consider the game *Rock-Paper-Scissors*, where as usual, paper covers rock, scissors cuts paper, and rock breaks scissors (that is: the former beats the latter in the comparison). In a face-off, the winner earns +1 and the loser earns -1. If two of the same type face each other, then there is a tie, and both earn 0.

The matrix below shows the game of Rock-Paper Scissors depicted as a *zero-sum-game*. Suppose that brothers Ron and Charlie Weasley are facing off. Each brother must choose a strategy. In the language of the *payoff matrix* below, Ron is the *row player*, and he must choose a row to play as his strategy. Similarly, Charlie is the *column player* and he just choose which column to play. If Ron chooses row i and Charlie chooses column j, then the payoff to Ron will be a_{ij}, and the payoff to Charlie will be $-a_{ij}$, hence the term "zero-sum." Thus, the row and column players prefer bigger and smaller numbers, respectively.

	Rock	Paper	Scissors
Rock	0	-1	1
Paper	1	0	-1
Scissors	-1	1	0

Order of Turns

- Typically, RPS is played by both players simultaneously choosing their strategies.
- But what if I made you go first? That's obviously unfair—whatever you do, I can respond with the winning move.
- Now what if I only forced you to commit to a *probability distribution* over rock, paper, and scissors? (Then I respond choosing a strategy, and *then* nature flips coins on your behalf.)

  You can protect yourself by randomizing uniformly among the three options—then, no matter what I do, I'm equally likely to win, lose, or tie.

The *minimax theorem* states that, in general games of "pure competition," a player moving first can always protect herself by randomizing appropriately.

The Minimax Theorem

Notation:

- $m \times n$ payoff matrix $\mathbf{A}$—a_{ij} is the row player's payoff for outcome (i, j) when row player plays strategy i and column player plays strategy j
- mixed row strategy $\mathbf{x}$ (a distribution over rows)
- mixed column strategy $\mathbf{y}$ (a distribution over columns)

Expected payoff of the row player:

$$
\begin{aligned}
\sum_{i=1}^{m}\sum_{j=1}^{n}\Pr[\text{outcome } (i,j)]\, a_{ij} &= \sum_{i=1}^{m}\sum_{j=1}^{n}\underbrace{\Pr[\text{row } i \text{ chosen}]}_{=x_i}\,\underbrace{\Pr[\text{column } j \text{ chosen}]}_{=y_j}\, a_{ij} \\
&= \mathbf{x}^T\mathbf{A}\mathbf{y}
\end{aligned}
$$

The minimax theorem is the amazing statement that turn order *doesn't matter*.

Theorem 4 (Minimax Theorem)**.** *For every two-player zero-sum game* $\mathbf{A}$*,*

$$
\max_{\mathbf{x}}\left(\min_{\mathbf{y}} \mathbf{x}^T\mathbf{A}\mathbf{y}\right) = \min_{\mathbf{y}}\left(\max_{\mathbf{x}} \mathbf{x}^T\mathbf{A}\mathbf{y}\right). \tag{1}
$$

On the left, the row player goes first, choosing a strategy to maximize their payoff and protect against the fact that the column player goes second and adapts to their strategy. The right is the opposite situation. The *value of the game* (value that both sides will equal) is 0 in this case: the first player will play randomly and the second will respond arbitrarily.

From LP Duality to Minimax

This is not the original or only argument, but we will now derive Theorem 4 from LP duality arguments. The first step is to formalize the problem of computing the best strategy for the player forced to go first.

Two issues: (1) the nested min/max, and (2) the quadratic (nonlinear) character of $\mathbf{x}^T\mathbf{A}\mathbf{y}$ in the decision variables $\mathbf{x}$, $\mathbf{y}$.

Observation 5. *The second player never needs to randomize. If the row player goes first and chooses any distribution* $\mathbf{x}$*, the column player can then simply compute the expected payoff (with respect to* $\mathbf{x}$*) of each column and choose the best.*

In math, we have argued that

$$
\begin{aligned}
\max_{\mathbf{x}}\left(\min_{\mathbf{y}} \mathbf{x}^T\mathbf{A}\mathbf{y}\right) &= \max_{\mathbf{x}}\left(\min_{j=1}^{n} \mathbf{x}^T\mathbf{A}\mathbf{e}_j\right) && (2)\\
&= \max_{\mathbf{x}}\left(\min_{j=1}^{n} \sum_{i=1}^{m} a_{ij}x_i\right) && (3)
\end{aligned}
$$

where $\mathbf{e}_j$ is the jth standard basis vector, corresponding to the column player deterministically choosing column j.

We've solved one of our problems by getting rid of $\mathbf{y}$. But there is still the nested max/min.