Lecture notes for Mar 8, 2023
Linear Programming and weighted bipartite matching

Chun-Hung Liu

March 8, 2023

1 Linear programming

Linear Programming
Input: An $m \times n$ matrix A over real numbers, a column vector $b \in \mathbb{R}^m$ and a column vector $c \in \mathbb{R}^n$.
Output: Find $\max_{x \in \mathbb{R}^n, Ax \leq b} c^T x$, or conclude that the set $\{x \in \mathbb{R}^n : Ax \leq b\}$ is empty, or conclude that the maximum does not exist (i.e. for every $\alpha \in \mathbb{R}^n$, there exists $x_0 \in \mathbb{R}^n$ with $Ax_0 \leq b$ such that $c^T x_0 > \alpha$).

Every vector x in the set $\{x \in \mathbb{R}^n : Ax \leq b\}$ is called a *feasible solution*; any feasible solution that attains the maximum is called an *optimal solution*.

Example:

1. The maximum flow of a network (D, s, t, g) is the case that A is the incidence matrix whose each row corresponds to an edge of D and whose column corresponds to a directed path from s to t, and the vector b is the vector with $|E(D)|$ entries such that for every $e \in E(D)$, the e-th entry of b equals the capacity of e, and c is the vector 1.

Unlike integer programming, linear programming can be solved in polynomial time. But the proof is too complicated to be included here.

Theorem 1 *There exists a polynomial p such that if all entries of the matrix A are rational, then Linear Programming can be solved in time $p(|A|)$, where $|A|$ is the total number of bits to represent the entries of A.*

Given the maximization problem $\max_x c^T x$ subject to $Ax \leq b$, it has a "dual problem": $\min_y y^T b$ subject to $y^T A = c^T, y \geq 0$. It is easy to see that for every feasible solution x for the primal problem and for every feasible solution for the dual problem, $y^T b \leq y^T(Ax) = (y^T A)x = c^T x$. So the optimal value for the dual is always at most the optimal value of the primal. Hence we obtain the following proposition.

Proposition 2 *Let A be an $m \times n$ matrix over $\mathbb{R}$. Let $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Then*

$$\sup_{x \in \mathbb{Z}^n, Ax \leq b} c^T x \leq \sup_{x \in \mathbb{R}^n, Ax \leq b} c^T x \leq \inf_{y \in \mathbb{R}^m, y^T A = c^T, y \geq 0} b^T y \leq \inf_{y \in \mathbb{Z}^m, y^T A = c^T, y \geq 0} b^T y.$$

In fact, the optimal values of the pair of primal and dual of linear programming always equal, as long as feasible solutions exist.

Theorem 3 (Duality Theorem) *Let A be an $m \times n$ matrix over $\mathbb{R}$. Let $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ and $D = \{y \in \mathbb{R}^m : y^T A = c^T, y \geq 0\}$. If $P \neq \emptyset \neq D$, then $\max_{x \in P} c^T x = \min_{y \in D} b^T y$.*

Proofs for the Duality Theorem can be found in most of linear optimization courses, so we do not repeat it here.

We can also obtain a "more symmetric" version.

Corollary 4 (Duality Theorem) *Let A be an $m \times n$ matrix over $\mathbb{R}$. Let $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Let $P = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$ and $D = \{y \in \mathbb{R}^m : y^T A \geq c^T, y \geq 0\}$. If $P \neq \emptyset \neq D$, then $\max_{x \in P} c^T x = \min_{y \in D} b^T y$.*

Proof. Let $B = \begin{bmatrix} A \\ -I \end{bmatrix}$, where I is the identity matrix, and let $d = \begin{bmatrix} b \\ 0 \end{bmatrix}$. Then $P = \{x \in \mathbb{R}^n : Bx \leq d\}$. Let $D' = \{y \in \mathbb{R}^{m+n} : y^T B = c^T, y \geq 0\}$. Since $P \neq \emptyset$, by Theorem 3, if $D' \neq \emptyset$, then $\max_{x \in P} c^T x = \min_{y \in D'} d^T y$.

For each $y \in \mathbb{R}^{m+n}$, let y_1 be the vector consisting of the first m entries of y, and let y_2 be the vector consisting of the last n entries of y. Then for every $y \in D'$, $c^T = y^T B = y_1^T A - y_2^T I \leq y_1^T A$, so $y_1 \in D$. Moreover, the last n entires of d are zero, so $d^T y = b^T y_1$ for every $y \in D'$. Hence $\min_{y \in D'} d^T y = \min_{y \in D'} b^T y_1 \geq \min_{z \in D} b^T z$.

Conversely, if $z \in D$, then $z \in \mathbb{R}^m$ with $z^T A \geq c^T$, so there exists $w \in \mathbb{R}^n$ such that $w \geq 0$ and $z^T A - w^T I = c^T$, and hence $\begin{bmatrix} z \\ w \end{bmatrix} \in D'$ and $b^T z = d^T \begin{bmatrix} z \\ w \end{bmatrix} \geq \min_{y \in D'} d^T y$. So $\min_{z \in D} b^T z \geq \min_{y \in D'} d^T y$. Therefore, $\min_{z \in D} b^T z = \min_{y \in D'} d^T y$. Moreover, $D \neq \emptyset$ implies $D' \neq \emptyset$, so $\max_{x \in P} c^T x = \min_{y \in D'} d^T y = \min_{z \in D} b^T z$. ■

Remark:

1. Theorem 3 shows that the middle inequality in Proposition 2 is an equality, assuming the feasibility. But the other two inequalities usually are unequal, as we will see below.

2. Consider the matrix A and vectors b and c for matching. That is, A is the vertex-edge incidence matrix of a graph G, and b and c equal 1. Now assume that G is the cycle of length k for some integer k. Then the integral version for maximization has optimal value $\lfloor \frac{k}{2} \rfloor$, because it is equals the maximum size of a matching. And the vector that assigns each entry $\frac{1}{2}$ is a feasible solution for $Ax \leq b$, so the optimal value for the fractional maximization problem is at least $\frac{k}{2}$, which is strictly greater than the integral maximum when k is odd.

3. Again consider the matrix A and vectors b and c for matching. That is, A is the vertex-edge incidence matrix of a graph G, and b and c equal 1. Now we consider the "symmetric" version. Note that having the condition $x \geq 0$ or not does not affect the optimal value because $c > 0$. Each feasible solution y of the integral version of the dual (i.e. $y^T A \geq c^T$ and $y \geq 0$ is integral) is equivalent to assigning each vertex 0 or 1 such that each edge has at least one end assigned by 1, and hence it is equivalent to a vertex-cover of G. Now assume that G is the cycle of length k for some integer k. Then the integral version for minimization has optimal value $\lceil \frac{k}{2} \rceil$, because it is equals the minimum size of a vertex-cover. And the vector that assigns each entry $\frac{1}{2}$ is a feasible solution for $y^T A \geq c^T$, so the optimal value for the fractional minimization problem is at most $\frac{k}{2}$, which is strictly smaller than the integral minimum when k is odd.

4. König's theorem states that if G is bipartite, then the maximum size of a matching in G equals the minimum size of a vertex-cover in G. Therefore, if G is bipartite, then the four optima in Proposition 2 are equal.

5. Consider the matrix A and vectors b and c used to formulate the maximum value of a flow of a network (D, s, t, g) as a maximization problem of a linear programming.

   - Every feasible solution y of the dual is equivalent to assigning each row of A (i.e. each edge of D) a nonnegative real number y_e such that for each column of A (i.e. a directed path from s to t), the sum of the numbers on the edges of this path is at least 1. If y is integral, then it is equivalent to a subset Y of $E(D)$ such that there exists no directed path in $D - Y$ from s to t, so Y contains an edge-cut separating s and t, and we can repeatedly remove edges from Y until Y equals an edge-cut without increasing $b^T y$ (since the capacity is nonnegative). Conversely, if Y is an edge-cut separating s and t, then the corresponding vector y is an integral feasible solution for the dual. Hence the integral minimization problem is equivalent to finding the minimum capacity of a cut in the network. By the maximum-flow-minimum-cut theorem, the minimum capacity of a cut equals the maximum value of a flow. So the second and the third inequalities in Proposition 2 are equalities.

   - When the capacity is not integral, it is possible that every maximum flow is not an integral flow, so the first inequality in Proposition 2 can be strict.

   - Assume that the capacity is integral. We have seen from the Ford-Fulkerson algorithm that there exists an integral optimal solution for the maximization problem. So the first inequality in Proposition 2 is an equality.

Here is a useful corollary of the Duality Theorem.

Corollary 5 *Let A be an $m \times n$ matrix over $\mathbb{R}$. Let $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ and $D = \{y \in \mathbb{R}^m : y^T A = c^T, y \geq 0\}$. Let $x^* \in P$ and $y^* \in D$. Then the following are equivalent:*

1. $c^T x^* = \max_{x\in P} c^T x$ *and* $b^T y^* = \min_{y\in D} b^T y$.
2. $c^T x^* = b^T y^*$.
3. $y^{*T}(b - Ax^*) = 0$.

Proof. The equivalence between the first two statements follow from Theorem 3. And $c^T x^* - b^T y^* = (y^{*T}A)x^* - y^{*T}b = y^{*T}(Ax^* - b)$. So Statements 2 and 3 are equivalent. ∎

Statement 3 in Corollary 5 is called the *complementary slackness condition*. We can obtain a similar result for the symmetric version.

Corollary 6 *Let A be an $m \times n$ matrix over $\mathbb{R}$. Let $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Let $P = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$ and $D = \{y \in \mathbb{R}^m : y^T A \geq c^T, y \geq 0\}$. Let $x^* \in P$ and $y^* \in D$. Then the following are equivalent:*

1. $c^T x^* = \max_{x\in P} c^T x$ *and* $b^T y^* = \min_{y\in D} b^T y$.
2. $c^T x^* = b^T y^*$.
3. $y^{*T}(b - Ax^*) = 0$ *and* $(c^T - y^{*T}A)x^* = 0$.

Proof. The equivalence between the first two statements follow from Theorem 3.

Note that $y^{*T}(b-Ax^*) \geq 0 \geq (c^T - y^{*T}A)x^*$, and Statement 2 is equivalent to $y^{*T}(b - Ax^*) = (c^T - y^{*T}A)x^*$. So Statement 2 is equivalent to $y^{*T}(b - Ax^*) = 0 = (c^T - y^{*T}A)x^*$ which is Statement 3. ∎

Statement 3 in Corollary 6 is also called the *complementary slackness condition*.

2 Weighted matching

In this section we consider the maximum weighted matching in a weighted graph. That is, given an (edge-)weighted graph (G, w), we want to find a matching M in G maximizing $\sum_{e\in M} w(e)$. Note that every edge with negative weight cannot be in any matching with maximum weight. So we can delete all edges with negative weight from G without changing the answer. Hence we may assume that w is nonnegative without changing the maximum weight of a matching.

2.1 Bipartite graphs

We first consider the case when G is bipartite. Even though we will consider the case for general G and provide a more efficient algorithm than the bipartite case stated in this section, this bipartite case is simpler and the general case uses some ideas from it. So we still state the bipartite case here.

Recall the integer/linear programming formulation for the matching and vertex-cover problem. They are easily generalized to weighted graphs, as described below.

Let (G, w) be a nonnegative weighted graph, and let A be the vertex-edge incidence matrix of G. The goal is to find a matching M in G such that $\sum_{e \in M} w(e)$ is as large as possible. Note that we can treat w as a vector in $\mathbb{R}^{|E(G)|}$ indexed by $E(G)$.

- Every matching in (G, w) corresponds to a vector $x \in \{0,1\}^{|E(G)|}$ satisfying $Ax \leq 1$. And weight of a maximum weighted matching is $\max_{x \in \{0,1\}^{|E(G)|}, Ax \leq 1} w^T x$. Note that it is also equal to

$$\max_{x \in \{0,1\}^{|E(G)|}, Ax \leq 1, x \geq 0} w^T x$$

  since w is a nonnegative function.

- The LP-relaxation of the above problem is

$$\max_{x \in \mathbb{R}^{|E(G)|}, Ax \leq 1, x \geq 0} w^T x.$$

- The dual of the LP-relaxation of the above problem is

$$\min_{y \in \mathbb{R}^{|V(G)|}, y^T A \geq w^T, y \geq 0} 1^T y.$$

  Note that it is the minimum of "weighted vertex-cover" of G. In particular, if $w = 1$, then it is exactly the minimum vertex-cover. We call every vector $y \in \mathbb{R}^{|V(G)|}$ satisfying $y^T A \geq w^T$ and $y \geq 0$ a *fractional vertex-cover* for (G, w).

- By the weak duality,

$$\max_{x \in \{0,1\}^{|E(G)|}, Ax \leq 1, x \geq 0} w^T x \leq \max_{x \in \mathbb{R}^{|E(G)|}, Ax \leq 1, x \geq 0} w^T x \leq \min_{y \in \mathbb{R}^{|V(G)|}, y^T A \geq w^T, y \geq 0} 1^T y.$$

- The key idea of the algorithm is to find a 0-1 solution x^* for the maximization problem (for matching) and a (possibly non-integral) solution y^* of the minimization problem (for fractional vertex-cover) such that $w^T x^* = 1^T y^*$. This implies that all inequalities are equalities, and hence x^* is a maximum weighted matching.

- When x^* corresponds to a perfect matching, we can relax the condition $y \geq 0$ as stated in fractional vertex-cover, as we will see in Proposition 7.

A *weak fractional vertex-cover* of a graph G is a function $f : V(G) \to \mathbb{R}$ such that for every $e = uv \in E(G)$, $f(u) + f(v) \geq w(e)$.

Proposition 7 *Let (G, w) be a weighted graph. If M is a perfect matching of G and f is a weak fractional vertex-cover of G, then $\sum_{e \in M} w(e) \leq \sum_{v \in V(G)} f(v)$.*

Proof. For each edge $e \in M$, let u_e, v_e be the ends of e. Since M is a perfect matching, $\{u_e, v_e : e \in M\} = V(G)$. Then $\sum_{e \in M} w(e) \leq \sum_{e \in M}(f(u_e) + f(v_e)) = \sum_{v \in V(G)} f(v)$. ■

We first consider the case that G is a complete bipartite graph whose two parts have the same size. Note that such G has a perfect matching, so Proposition 7 can apply. The problem for finding a maximum weighted matching in such G is called the *Assignment Problem*.