7 Linear Programming

In Chapters 3 (Greedy Algorithms) and 4 (Dynamic Programming), we saw two general approaches to solving optimization problems that are applicable to a variety of settings (e.g., subsequences, knapsacks, graphs). In this chapter, we study another general technique known as *linear programming*.

This chapter is a bit different from previous chapters. In Sec. 7.1, we introduce the basic idea behind linear programming and give two applications. But we won't actually study any algorithms that solve linear programs; for our purposes, we can assume that we have access to an LP solver. By relying on a solver, we can focus our attention on some applications of linear programming.

7.1 Basics of LPs

A *linear program* (LP) is an optimization problem in which the objective is to maximize (or minimize) a linear function subject to constraints in the form of linear inequalities. For example, here is a linear program that has 2 variables and 1 constraint:[1]

$$\begin{aligned} \max\ & 3x_1 + 2x_2 \\ & x_1 + x_2 \leq 5 \\ & x_1, x_2 \geq 0 \end{aligned}$$

A *solution* to the LP is an assignment of each variable to a real number. It is *feasible* if it satisfies all of the constraints, and it is *optimal* if it is feasible and has the highest (or lowest, for minimization problems) objective value among all feasible solutions. (An LP can have multiple optimal solutions.) In the LP above, the solution $(x_1, x_2) = (1, 2)$ is feasible, $(5, 3)$ and $(-1, 4)$ are not, and $(5, 0)$ is optimal. The value of the optimal solution is 15.

For concise, we can use vectors and matrices to specify linear programs. If an LP has n variables and m constraints, then we let $x \in \mathbb{R}^n$ denote the vector of variables, $c \in \mathbb{R}^n$ denote the coefficients of the objective function, and $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ specify the constraints. Putting it all together, an LP in *canonical form*[2] looks like this:

$$\begin{aligned} \max\ & c^\top x \\ & Ax \leq b \\ & x \geq 0 \end{aligned}$$

Note that "$x \geq 0$" means each of the n entries of x is at least 0. Similarly, "$Ax \leq b$" means the i-th entry of Ax is at most b_i. So the first LP can be written as $x = (x_1, x_2), c = (3, 2), A = (1, 1)$, and $b = (5)$ (where x, c, b are column vectors). We now look at two well-known applications of linear programming.

[1] The constraint "$x_1, x_2 \geq 0$" is short for "$x_1 \geq 0$ and $x_2 \geq 0$", and these are not usually counted as constraints because they appear so frequently.

[2] Different sources refer to different things as canonical (or sometimes *standard*) form, but all of the forms you might find can be transformed into each other.

The Diet Problem. Suppose a store sells n food items, where each item j has cost c_j per unit. Also, there are m nutrients, and item j provides A_{ij} units of nutrient i. Our diet mandates that we eat at least b_j units of nutrient j. Our goal is to spend the least amount of money on the n food items while satisfying the m nutritional requirements.

We can model this problem as an LP as follows. Let x be an n-dimensional vector. We need to determine x_j for all $j \in \{1, \ldots, n\}$; x_j represents the amount of food j we buy. Let $c \in \mathbb{R}^n$ denote the vector containing the costs, let A be an $m \times n$ matrix whose (i, j)-th entry is A_{ij}, and let $b \in \mathbb{R}^m$ denote the vector containing the nutritional lower bounds. Our goal is to minimize $c^\top x$ such that $Ax \geq b$ and $x \geq 0$.

The Manufacturing Problem. Now suppose we're the ones doing the selling. Our factory is capable of producing n products, which are made from m different materials. Each unit of product j requires A_{ij} units of material i and yields c_j units of profit. We have b_i units of material i at our disposal, and our goal is to maximize the total profit.

In this case, x_j represents the amount of product j we produce. The n-dimensional vector c contains the profits, and $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ encode the material constraints. Our goal is to maximize $c^\top x$ such that $Ax \leq b$ and $x \geq 0$.

Remark. The manufacturing problem LP is in canonical form, but the LP for the diet problem is not. However, if we negate the vector c then the objective becomes a maximization, and if we negative all entries of A and b then the inequality signs flip. In general, by doing small transformations, we can ensure that the objective is in canonical form; we leave the details as an exercise. For our purposes, we assume that we can efficiently find an optimal solution of any linear program, even if it's not in canonical form.[3]

7.2 Vertex Cover

Problem statement. Let $G = (V, E)$ be an undirected graph. A *vertex cover* is a subset of vertices S such that every edge is incident to at least one vertex in S. Our goal is to find a vertex cover of G containing the fewest number of vertices.

Algorithm. Associate each vertex u with a variable $x_u \in \{0, 1\}$. (Intuitively, $x_u = 1$ corresponds to selecting u to be included in the solution, but this is not known by the LP solver.) Then the vertex cover problem is equivalent to the following *integer* LP:

$$\begin{aligned} \min \sum_{u \in V} x_u \\ x_u + x_v \geq 1 \quad \forall \{u, v\} \in E \\ x_u \in \{0, 1\} \quad \forall u \in V. \end{aligned}$$

This problem is almost an LP; the only difference is the $x_u \in \{0, 1\}$ constraint. Let's replace this constraint with $x_u \geq 0$ and solve the resulting LP to obtain an optimal solution $x^* \in \mathbb{R}^n$. We return $S = \{u : x_u^* \geq 1/2\}$. For the analysis, let $\mathsf{OPT} \subseteq V$ denote an optimal solution.

[3]There are entire courses and books devoted to linear programming and its algorithms; if you are interested in learning about them, I recommend looking at other resources online.

Theorem 7.1. *The set S is a vertex cover and satisfies $|S| \leq 2 \cdot |\mathsf{OPT}|$.*

Before proving the theorem, we first prove that the optimal value of the linear program is at most the number of vertices in an optimal vertex cover.

Lemma 7.2. *The LP solution x^* and OPT satisfy $\sum_{u \in V} x_u^* \leq |\mathsf{OPT}|$.*

Proof. The integer linear program above (i.e., with the $x_u \in \{0, 1\}$ constraints) is equivalent to the vertex cover problem, so its optimal value is equal to $|\mathsf{OPT}|$. By relaxing the constraints $x_u \in \{0, 1\}$ to $x_u \geq 0$, the optimal value of the program can only improve, i.e., decrease. Thus, the optimal value of the LP is at most $|\mathsf{OPT}|$, as desired. □

Proof of Theorem 7.1. To show that S is a vertex cover, we must show that every edge $e = \{u, v\}$ is covered. Since x^* is a feasible solution to the linear program, we must have $x_u^* + x_v^* \geq 1$. Thus, at least one the values in $\{x_u^*, x_v^*\}$ must be at least $1/2$, so that corresponding vertex gets included in S, which means it covers the edge e.

Now we bound the size of S. Since every $u \in S$ satisfies $x_u^* \geq 1/2$, we have

$$|S| = \sum_{u \in S} 1 \leq \sum_{u \in S} 2x_u^* \leq 2 \cdot \sum_{u \in V} x_u^* \leq 2 \cdot |\mathsf{OPT}|,$$

where the final inequality follows from Lemma 7.2. □

7.3 Zero-Sum Games

Problem statement. Alice and Bob are playing a game; Alice has m options denoted by $[m] = \{1, \ldots, m\}$ while Bob has n options denoted by $[n] = \{1, \ldots, n\}$. They must make their choices simultaneously. If Alice chooses i and Bob chooses j, then Alice receives A_{ij} dollars from Bob (or she gives him this amount, if A_{ij} is negative). Naturally, Alice wants to maximize her payoff, and Bob wants to minimize it. Our goal is find a strategy for Alice that maximizes the amount she can *guarantee* for herself.[4]

Algorithm. A *pure* strategy is one that always selects exactly one option, while a *mixed* strategy is one that specifies a probability distribution over the options. If Alice chooses a pure strategy of option i, then the payoff she can guarantee is $\min_j A_{ij}$, since Bob can pick the strategy j that achieves this minimum value. So Alice's best pure strategy i is the one that maximizes $\min_j A_{ij}$.

Now let's consider Alice's guaranteed payoff if she chooses a mixed strategy. The key is that Alice can solve for her guaranteed payoff value by modeling this problem as a linear

[4]This game is known as zero-sum because the sum of Alice's payoff and Bob's payoff is always zero. In other words, Alice's gain is directly equal to Bob's loss, and vice versa.

program, where x_i denotes the probability that she should choose option i:

$$\max u$$
$$\sum_{i=1}^{n} x_i \cdot A_{ij} \geq u \quad \forall j \in [n]$$
$$\sum_i x_i = 1$$
$$x_i \geq 0 \quad \forall i \in [m]$$

Constraint j says that, if Bob chooses option j, then in expectation (over Alice's options), Alice's payoff must be at least u. When combined, these constraints ensure that no matter what Bob does (even if uses a mixed strategy), Alice can guarantees herself a payoff of u. The optimal solution of the above LP is known as the *value* of the game.

Remark. Let u^* denote the value of the game, and suppose Alice is forced to announce her strategy before playing. By solving the LP, she can guarantee herself a payoff of at least u^*. Can she do better if she hides her strategy? Surprisingly, the answer is no. Due to a concept known as *duality*, there is a mixed strategy for Bob that guarantees Alice's payoff is at most u^*. Thus, if Bob plays this strategy, then even if Alice hides her strategy, her payoff will still be at most u^*.

7.4 Duality

Recall that the linear program below, on the left, is in canonical form:

$\max c^\top x$	$\min b^\top y$
$Ax \leq b$	$A^\top y \geq c$
$x \geq 0$	$y \geq 0$

The linear program on the right is the *dual* of the LP in canonical form. When we take the dual of an LP, we refer to the original LP as the *primal* LP. Notice that if the primal LP has n variables and m constraints (ignoring $x \geq 0$), then the dual LP has m variables and n constraints; also, the dual of a dual is the primal.

Motivation. The dual is aesthetically pleasing as the "opposite" of the primal: the objective flips from max to min, the constraints become variables, and the variables become constraints. But there's another reason that motivates the consideration of the dual; here's a motivating example. Consider the following (primal) LP:

$$\max 7x_1 + 4x_2$$
$$2x_1 + x_2 \leq 5$$
$$x_1 \leq 2$$
$$x_2 \leq 3$$
$$x \geq 0$$

The optimal value is 19; it is achieved when $x = (1, 3)$. One proof that this is optimal is to multiply the first inequality by $7/2$, the second by 0, the third by $1/2$, and adding. The result is the constraint $7x_1 + 4x_2 \leq 19$, which matches the optimal value. But where did these magical multipliers $(7/2, 0, 1/2)$ come from, and is it always possible to find them? This is the motivation behind duality.

Let $y = (y_1, y_2, y_3)$ represent the "magical multipliers" we want to find. The proof above works only if $y = (y_1, y_2, y_3) \geq 0$ and, after multiplying and adding the inequalities, the coefficient of each x_i is at least its coefficient in the objective function. In other words, if $y \geq 0$, $2y_1 + y_2 \geq 7$, and $y_2 + y_3 \geq 4$, then $7x_1 + 4x_2 \leq 5y_1 + 2y_2 + 3y_3$. Since we want the right-hand side to be as small as possible, our goal of finding the "magical multipliers" reduces to solving the dual LP:

$$\begin{aligned} \min\, & 5y_1 + 2y_2 + 3y_3 \\ & 2y_1 + y_2 \geq 7 \\ & y_2 + y_3 \geq 4 \\ & y \geq 0 \end{aligned}$$

So by construction, the value of any feasible dual solution is an upper bound on the value of any feasible primal solution. In fact, the optimal dual solution has the same value as the optimal primal solution! More specifically, we have the following known as the *Strong Duality Theorem*:

Theorem 7.3. *If x^* is an optimal solution for an LP in canonical form, then there exists an optimal solution y^* for its dual such that $c^\top x^* = b^\top y^*$.*

This theorem is a generalization of the Max-Flow Min-Cut Theorem (Theorem 6.4); we omit the proof. Furthermore, as mentioned in the previous section, duality explains why it doesn't help a player to hide their strategy in a zero-sum game.

7.5 Exercises

7.1. Describe an LP that has infinitely many feasible solutions, but no optimal solution. Then describe an LP that has infinitely many optimal solutions.

7.2. Explain why it might be impossible to optimally solve a linear program if one of its constraints is a strict inequality (e.g., $x_1 + x_2 < 5$).

7.3. Consider the Vertex Cover LP from Sec. 7.2. Give an example of a graph such that the optimal LP value is less than the size of the optimal vertex cover.

7.4. In the weighted vertex cover problem, every vertex u has weight $w_u \geq 0$ and our goal is to find a vertex cover of minimum weight. Show that a small modification of the algorithm from Sec. 7.2 produces a solution whose total weight is at most twice the optimal weight.

7.5. In Sec. 7.1, we claimed that any linear program can be converted into canonical form. In this exercise, we consider how this can be done.

(a) Suppose that we want to minimize $c^\top x$ (rather than maximize).

(b) Suppose we have a constraint of the form $\alpha^\top x \geq \beta$ or $\alpha^\top x = \beta$, where $\alpha, x \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$.

(c) Suppose we have a variable x_i that is unrestricted in sign, that is, x_i is allowed to be positive or negative (or 0).

For each situation above, explain how we can modify the LP such that an optimal solution of the new LP yields an optimal solution to the original LP.

7.6. Let $x^*, y^* \in \mathbb{R}^n$ be two distinct optimal solutions of an LP in canonical form. Prove that $(x^* + y^*)/2 \in \mathbb{R}^n$ is also an optimal solution of the LP. Then show that, in fact, the LP has infinitely many optimal solutions.

7.7. Recall from Sec. 7.2 that Vertex Cover can be modeled by an LP. However, the LP didn't specify a matrix A, or vectors b, c, x, as shown in Sec. 7.1. For this problem, determine what A, b, c, and x should be.

7.8. Consider the Maximum Flow problem from Ch. 6. Model this problem as an LP, including the exact definitions of the matrix A and vectors b, c, x.

7.9. Consider the minimum spanning tree problem from Sec. 2.3; we shall model this problem using an *integer* LP (ILP, see Sec. 7.2) in two different ways. Both formulations contain a variable $x_e \in \{0, 1\}$ for every edge $e \in E$. The objective of both is to minimize $\sum_{e \in E} w_e x_e$, but they have different constraints.

(a) For every $S \subset V$, any spanning tree must contain at least one edge with exactly one endpoint in S. Express this idea as a set of constraints in the ILP.

(b) Any spanning tree contains exactly $n-1$ edges, and for every $S \subseteq V$, any spanning tree contains at most $|S| - 1$ edges within S. Express these ideas as constraints in the ILP.

Suppose we have an optimal solution to the first ILP. Show how we can convert this to into a spanning tree, and prove that the result is in fact an MST. Repeat this for the second ILP. Finally, for exactly one of the formulations, if we replace $x_e \in \{0, 1\}$ with $0 \leq x \leq 1$, then the optimal value can decrease — which formulation is it?