Lecture 15: Linear Programming

Linear programming (LP) is a method to achieve the optimum outcome under some requirements represented by linear relationships. More precisely, LP can solve the problem of maximizing or minimizing a linear objective function subject to some linear constraints.

In general, the standard form of LP consists of

- Variables: $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)^\top$

- Objective function: $\boldsymbol{c} \cdot \boldsymbol{x}$

- Inequalities (constraints): $A\boldsymbol{x} \le \boldsymbol{b}$, where A is a $n \times d$ matrix

and we maximize the objective function subject to the constraints and $\boldsymbol{x} \ge 0$.

LP has many different applications, such as flow, shortest paths, and even politics. In this lecture, we will be covering different examples of LP, and present an algorithm for solving them. We will also learn how to convert any LP to the standard form in this lecture.

1   Examples of Linear Programming: Politics

In this example, we will be studying how to campaign to win an election. In general, there are n demographics, each with p_i people, and m issues that the voters are interested in. Given the information on how many votes can be obtained per dollar spent advertising in support of an issue, how can we guarantee victory by ensuring a majority vote in all demographics?[1]

In particular, consider the example with 3 demographics and 4 issues shown in Table 1. What is the minimum amount of money we can spend to guarantee majority in all demographics?

Let x_1, x_2, x_3, x_4 denote the dollars spent per issue. We can now formulate this problem as an LP problem:

$$
\begin{aligned}
& \text{Minimize } x_1 + x_2 + x_3 + x_4 \\
& \text{Subject to } -2x_1 + 8x_2 + 0x_3 + 10x_4 \ge 50,000 \text{ (Urban Majority)} && (1) \\
& \qquad\qquad\quad 5x_1 + 2x_2 + 0x_3 + 0x_4 \ge 100,000 \text{ (Suburban Majority)} && (2) \\
& \qquad\qquad\quad 3x_1 - 5x_2 + 10x_3 - 2x_4 \ge 25,000 \text{ (Rural Majority)} && (3) \\
& \qquad\qquad\quad x_1, x_2, x_3, x_4 \ge 0 \text{ (Can't unadvertise)} && (4)
\end{aligned}
$$

---

[1]We will assume that the votes obtained by advertising different issues are disjoint.

Policy	Demographic		
	Urban	Suburban	Rural
Building roads	-2	5	3
Gun Control	8	2	-5
Farm Subsidies	0	0	10
Gasoline Tax	10	0	2
Population	100,000	200,000	50,000

Table 1: Votes per dollar spent on advertising, and population.

1.1 Certificate of Optimality

Though we have not presented an algorithm for solving this problem, given a solution, we can verify that the solution is optimal with a proper certificate. For example,

$$x_1 = \frac{2050000}{111}$$
$$x_2 = \frac{425000}{111}$$
$$x_3 = 0$$
$$x_4 = \frac{625000}{111}$$
$$x_1 + x_2 + x_3 + x_4 = \frac{3100000}{111}$$

is a solution to this problem. Now consider the following equation (certificate):

$$\frac{25}{222} \cdot (1) + \frac{46}{222} \cdot (2) + \frac{14}{222} \cdot (3) = x_1 + x_2 + \frac{140}{222}x_3 + x_4$$
$$\geq \frac{25}{222} \cdot 50000 + \frac{46}{222} \cdot 100000 + \frac{14}{222} \cdot 25000$$
$$= \frac{3100000}{111}$$
$$\Rightarrow x_1 + x_2 + \frac{140}{222}x_3 + x_4 \geq \frac{3100000}{111}$$

We also know that $x_1 + x_2 + x_3 + x_4 \geq x_1 + x_2 + \frac{140}{222}x_3 + x_4$. Therefore, the given solution is an optimal solution to the problem.

2 Linear Programming Duality

The short certificate provided in the last section is not a coincidence, but a consequence of **duality** of LP problems. For every primal LP problem in the form of

$$\text{Maximize } \boldsymbol{c} \cdot \boldsymbol{x}$$
$$\text{Subject to } A\boldsymbol{x} \leq \boldsymbol{b}, \boldsymbol{x} \geq 0,$$

there exists an equivalent dual LP problem

$$\text{Minimize } \boldsymbol{b} \cdot \boldsymbol{y}$$
$$\text{Subject to } A^{\top}\boldsymbol{y} \geq \boldsymbol{c}, \boldsymbol{y} \geq 0.$$

This property of LP can be used show many important theorems. For instance, the max-flow min-cut theorem can be proven by formulating the max-flow problem as the primal LP problem.

3 Converting to Standard Form

The natural LP formulation of a problem may not result in the standard LP form. In these cases, we can convert the problem to standard LP form without affecting the answers by using the following rules.

- **Minimize an objective function**: Negate the coefficients and maximize.

- **Variable x_j does not have a non-negativity constraint**: Replace x_j with $x_j' - x_j''$, and $x_j', x_j'' \geq 0$.

- **Equality constraints**: Split into two different constraints; $x = b \Rightarrow x \leq b, x \geq b$.

- **Greater than or equal to constraints**: Negate the coefficients, and translate to less than or equal to constraint.

4 Formulating LP Problems

In this section, we will give brief descriptions of how to formulate some problems seen previously in this class as LP problems. Once we have a LP formulation, we can convert the problem into the standard form as described in Section 3.

4.1 Maximum Flow

We can model the max flow problem as maximization of sum of flows, under some constraints which will model different properties of the flow. Given $G(V, E)$, the capacity $c(e)$ for each $e \in E$, the source s, and the sink t,

$$\text{Maximize } \sum_{v \in V} f(s, v)$$
$$\text{Subject to } f(u, v) = -f(v, u) \ \forall u, v \in V \ \text{ skew symmetry}$$
$$\sum_{v \in V} f(u, v) = 0 \ \forall u \in V - \{s, t\} \ \text{ conservation}$$
$$f(u, v) \leq c(u, v) \ \forall u, v \in V \ \text{ capacity.}$$

4.2 Shortest Paths

We can model the shortest paths problem as minimization of the sum of all distances from a node. Note that this sum is minimized only when all distances are minimized. Given $G(V, E)$, the weight $w(e)$ for each $e \in E$, and the source s,

$$\text{Maximize } \sum_{v \in V} d(v)$$
$$\text{Subject to } d(v) - d(u) \leq w(u, v) \forall u, v \in V \text{ triangular inequality}$$
$$\sum_{v \in V} d(s) = 0.$$

Note the maximization above, so all distances don't end up being zero. There is no solution to this LP if and only if there exists a negative weight cycle reachable from s.

5 Algorithms for LP

There are many algorithms for solving LP problems:

- **Simplex algorithm**: In the feasible region, $\boldsymbol{x}$ moves from vertex to vertex in direction of $\boldsymbol{c}$. The algorithm is simple, but runs in exponential time in the worst case.

- **Ellipsoid algorithm**: It starts with an ellipsoid that includes the optimal solution, and keeps shrinking the ellipsoid until the optimal solution is found. This was the first poly-time algorithm, and was a theoretical breakthrough. However, the algorithm is impractical in practice.

- **Interior Point Method**: x moves inside the polytope following c. This algorithm runs in poly-time, and is practical.

In this lecture, we will study only the simplex algorithm.

5.1   Simplex Algorithm

As mentioned before, the simplex algorithm works well in practice, but runs in exponential time in the worst case. At a high level, the algorithm works as Gaussian elimination on the inequalities or constraints. The simplex algorithm works as follows:

- Represent LP in "slack" form.

- Convert one slack form into an equivalent slack form, while likely increasing the value of the objective function, and ensuring that the value does not decrease.

- Repeat until the optimal solution becomes apparent.

5.1.1   Simplex Example

Consider the following example:

$$\text{Minimize } 3x_1 + x_2 + x_3$$
$$\text{Subject to } x_1 + x_2 + 3x_2 \leq 30$$
$$2x_1 + 2x_2 + 5x_3 \leq 24$$
$$4x_1 + x_2 + 2x_3 \leq 36$$
$$x_1, x_2, x_3 \geq 0$$

Change the given LP problem to slack form, consisting of the original variables called *nonbasic* variables, and new variables representing slack called *basic* variables.

$$z = 3x_1 + x_2 + 2x_3$$
$$x_4 = 30 - x_1 - x_2 - 3x_3$$
$$x_5 = 24 - 2x_1 - 2x_2 - 5x_3$$
$$x_6 = 36 - 4x_1 - x_2 - 2x_3$$

We start with a *basic solution*: we set all nonbasic variables on the right hand side to some feasible value, and compute the values of the basic variables. For instance, we can set $x_1 = x_2 = x_3 = 0$. Note that the all 0 solution satisfies all constraints in this problem, but may not do so in the general case.

We now perform the pivoting step:

- Select a nonbasic variable x_e whose coefficient in the objective function is positive.

- Increase the value of x_e as much as possible without violating any constraints.

- Set x_e to be basic, while some other basic variable becomes nonbasic.

In this example, we can increase the value of x_1. The third constraint will limit the value of x_1 to 9. We then get

$$x_1 = 9 - \frac{x_2}{4} - \frac{x_3}{2} - \frac{x_6}{4}.$$

Now rewrite the other constraints with x_6 on the right hand side.

$$z = 27 + \frac{x_2}{4} + \frac{x_3}{2} - \frac{3x_6}{4}$$
$$x_1 = 9 - \frac{x_2}{4} - \frac{x_3}{2} - \frac{x_6}{4}$$
$$x_4 = 21 - \frac{3x_2}{4} - \frac{5x_3}{2} + \frac{x_6}{4}$$
$$x_5 = 6 - \frac{3x_2}{2} - 4x_3 + \frac{x_6}{2}$$

We note the equivalence of the solutions. That is, the original basic solution $(0, 0, 0, 30, 24, 36)$ satisfies the rewritten constraints, and has the objective value of 0. The second basic solution $(9, 0, 0, 21, 6, 0)$ has the objective value of 27.

At this point, pivoting on x_6 will actually cause the objective value to decrease (though the computation is not shown here). Thus let us pick x_3 as the next pivot to get

$$z = \frac{111}{4} + \frac{x_2}{16} - \frac{x_5}{8} - \frac{11x_6}{16}$$
$$x_1 = \frac{33}{4} - \frac{x_2}{16} + \frac{x_5}{8} - \frac{5x_6}{16}$$
$$x_2 = \frac{3}{2} - \frac{3x_2}{8} - \frac{x_5}{4} + -\frac{x_6}{8}$$
$$x_4 = \frac{69}{4} + \frac{3x_2}{16} + \frac{5x_5}{8} - \frac{x_6}{16}$$

which results in basic solution $\left(\frac{33}{4}, 0, \frac{3}{2}, \frac{69}{4}, 0, 0\right)$ with objective value of $\frac{111}{4}$.

Finally, pivoting on x_2 yields

$$z = 28 - \frac{x_3}{6} - \frac{x_5}{6} - \frac{2x_6}{3}$$
$$x_1 = 8 + \frac{x_3}{6} + \frac{x_5}{6} - \frac{x_6}{3}$$
$$x_2 = 4 - \frac{8x_3}{3} - \frac{2x_5}{3} + \frac{x_6}{3}$$
$$x_4 = 18 - \frac{x_3}{2} + \frac{x_5}{2}$$

Though we will not prove the correctness of this algorithm in this lecture, when all coefficients of all nonbasic variables are negative, the simplex algorithm has found the optimal solution. In general, simplex algorithm is guaranteed to converge in $\binom{n+m}{n}$ iterations where n is the number of variables, and $n+m$ is the number of constraints. This for general n and m can be exponential.

6   More Topics of LP

There are several important questions regarding LP that were not discussed in this lecture:

- How do we determine if LP is feasible?

- What if LP is feasible, but the initial basic solution is infeasible?

- How do we determine if the LP is unbounded?

- How do we choose the pivot?

These questions are answered in the textbook and other LP literature.

6.046J / 18.410J Design and Analysis of Algorithms
Spring 2015