ORIE 6300 Mathematical Programming I August 23, 2016

Lecture 1

Lecturer: Damek Davis *Scribe: Andrew Loeb*

Much of the course will be devoted to linear programming (LP), the study of the optimization of a linear function of several variables subject to linear inequality constraints. Here "programming" should be understood in the sense of planning — more like TV programming than computer programming — and linear refers to the types of functions involved.

Some examples are

- Finding the shortest path between a series of hubs,
- Investment planning,
- Learning from data, e.g. for recommendation systems

Mathematical programming is the art and science of transforming these types of problems into mathematical language, and solving them. We often use computers to solve mathematical programming problems, but the practice of mathematical programming existed long before computers were widespread.

There are many forms such a problem can take. We start with a (column) vector $x \in \mathbf{R}^n$ of *decision variables*. We want to maximize a linear *objective function* $c^T x$ for $c \in \mathbf{R}^n$ subject to linear inequalities $Ax \le b$ for $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$. The inequality is componentwise; if a_i is the ith row of A, and b_i is the ith component of b, then we want to have $a_i x \le b_i$ for $i = 1, \ldots, m$. Any decision vector x for which $Ax \le b$ is called a (*basic*) *feasible point* (or a *feasible solution*). We call the set

$$Q(A, b) := \{x \in \mathbf{R}^n : Ax \le b\}$$

of points satisfying all the constraints the *feasible region* or *feasible set*. For any feasible solution x, $c^T x$ is the *value* of x. A feasible solution x^* is *optimal* if it attains the maximum value (if it exists) among all feasible solutions. For any $x \in Q$, $c^T x$ is a *value* of the linear program, and $c^T x^*$, for x^* optimal, is the *maximal/optimal value* of the linear program. Note that all our vectors are columns, and that a subscripted letter could be a component of a vector (like b_i) or itself a vector (like a_i). We write:

$$\begin{array}{rrcl} \max & c^T x & & \\ \text{subject to} & Ax & \le & b, \end{array}$$

or sometimes just $\max(c^T x : Ax \le b)$.

We will study the following things:

- What is the geometry of the feasible region $Q(A, b)$?
- What form do optimal solutions take?
- How can we know if a given solution is optimal?

- How can we efficiently find an optimal solution?

Let's consider a concrete example.

Example 1 *(Product Mix): The Marie-Antoinette bakery makes high-end bread and cakes. Each loaf requires 3 pounds of flour and 2 hours of oven time, while each cake requires just 1 pound of flour but 4 hours of oven time. There are 7 pounds of flour and 8 hours of oven time available, and all other ingredients are in ample supply. (Note that this is a very small operation, and the oven can handle only one bakery product at a time!) Each loaf and each cake makes a \$5 profit. How many loaves and how many cakes should be made to maximize the bakery's profit?*

If we let x_1 and x_2 denote the numbers of loaves and cakes made (our decision variables), then the objective function to be maximized is $5x_1 + 5x_2$. The flour constraint is $3x_1 + x_2 \leq 7$, while the oven constraint is $2x_1 + 4x_2 \leq 8$. Are these all the constraints? No. The numbers of loaves and cakes cannot be negative, so we get

$$\begin{array}{llcll} \max & 5x_1 & + & 5x_2 & \\ & 3x_1 & + & x_2 & \leq 7, \\ & 2x_1 & + & 4x_2 & \leq 8, \\ & x_1 \geq 0, & & x_2 \geq 0. & \end{array}$$

We might argue that x_1 and x_2 should be integers, but this makes our problem an integer linear programming problem, which is potentially much harder to solve. So for now we allow x_1 and x_2 to take on any real values. This might be a reasonable approximation for a problem instance of more realistic size: perhaps x_j is the number of batches (say of 100 loaves or 100 cakes) made, so fractions are possible.

Our problem above is of the form $\max\{c^T x : Ax \leq b\}$, where $A = \begin{bmatrix} 3 & 2 & -1 & 0 \\ 1 & 4 & 0 & -1 \end{bmatrix}^T$ (note that the rows of A give the coefficients of the constraints), $c = \begin{pmatrix} 5 \\ 5 \end{pmatrix} = (5,5)^T$ and $b = (7,8,0,0)^T$.

We can solve such a small problem graphically, by drawing the feasible region in $\mathbf{R}^2$:

We can also draw "isoprofit" lines of the form $z = 5x_1 + 5x_2$, each of which show points of equal profit. By moving the isoprofit line up and to the right as much as possible, we see that $(2;1)$ looks like a good point; it gives a profit of \$15.

This is pretty convincing, but can we get an *algebraic* proof that this solution is optimal, which might work even if we can't draw a picture? Yes! *Any* feasible point must satisfy the two constraints

$$\begin{array}{lcll} 3x_1 & + & x_2 & \leq 7, \\ 2x_1 & + & 4x_2 & \leq 8, \end{array}$$

and so satisfies their sum: $5x_1 + 5x_2 \leq 15$. But we also have a feasible point, $x = (2;1)$, which gives objective function value 15, therefore it must be optimal!

Let us modify our example a bit. What if the profit per loaf becomes \$7 and per cake \$4? The objective function is now $7x_1 + 4x_2$. Adding the constraints no longer works, but we could take positive multiples of them first:

$$\begin{array}{rcc|lcll} & 2 & \times & 3x_1 & + & x_2 & \leq 7, \\ + & 1/2 & \times & 2x_1 & + & 4x_2 & \leq 8, \\ \hline & & & 7x_1 & + & 4x_2 & \leq 18, \end{array}$$

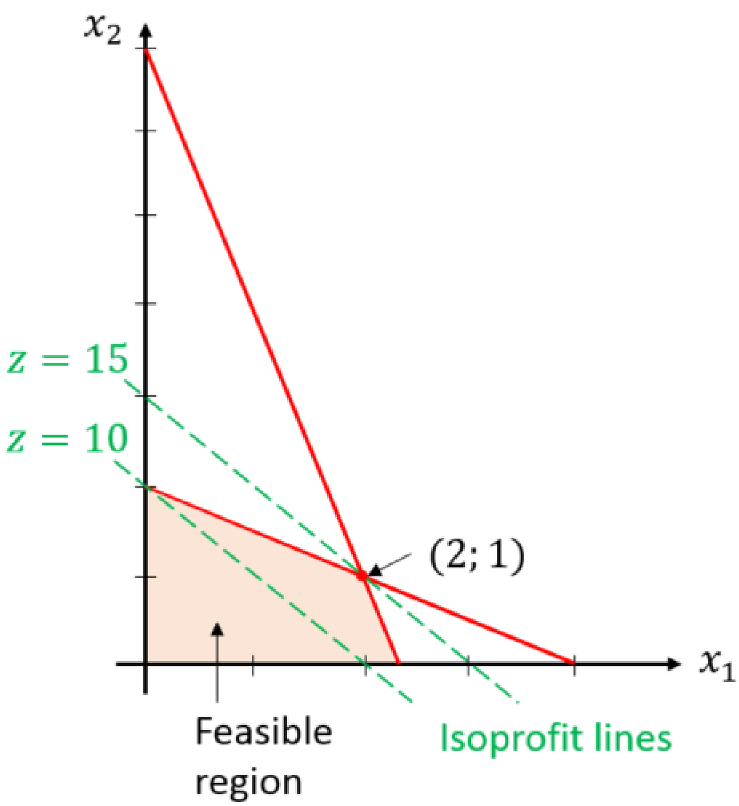

Figure 1: The feasible region and an isoprofit line for Example 1.

and the feasible point $(2;1)$ gives exactly \$18 revenue, so is still optimal.

What if the objective function becomes $x_1 + 4x_2$? Simple algebra suggests

$$\begin{array}{rrcrcrcl} & -2/5 & \times & 3x_1 & + & x_2 & \leq & 7, \\ + & 11/10 & \times & 2x_1 & + & 4x_2 & \leq & 8, \\ \hline & & & x_1 & + & 4x_2 & \leq & 6?? \end{array}$$

Is this valid? No!! Multiplying an inequality by a negative number changes its sense, and we can't then add the resulting inequalities.

Instead, we can proceed as follows:

$$\begin{array}{rrcrcrcl} & 1 & \times & 2x_1 & + & 4x_2 & \leq & 8, \\ + & 1 & \times & -x_1 & & & \leq & 0, \\ \hline & & & x_1 & + & 4x_2 & \leq & 8, \end{array}$$

and $x = (0;2)$ is feasible and gives an objective function value \$8.

Let's generalize this discussion. For each constraint $a_i x \leq b_i$, we want to multiply it by $y_i \geq 0$, so that we have

$$\begin{array}{rrcrcl} & y_1 & \times & (a_1 x & \leq & b_1) \\ & y_2 & \times & (a_2 x & \leq & b_2) \\ & & & \vdots & & \\ + & y_m & \times & (a_m x & \leq & b_m) \\ \hline & & & c^T x & \leq & y^T b \end{array}$$

So we generate an upperbound of $\sum_i y_i b_i = y^T b$. Also, for each x_i, we want to have exactly c_i copies of x_i. Thus we want

$$\begin{aligned} y_1 a_{11} + y_2 a_{21} + \cdots + y_m a_{m1} &= c_1 \\ y_1 a_{12} + y_2 a_{22} + \cdots + y_m a_{m2} &= c_2 \\ &\vdots \\ y_1 a_{1n} + y_2 a_{2n} + \cdots + y_m a_{mn} &= c_n, \end{aligned}$$

or $A^T y = c$. Then by the same arguments as above we have that for any feasible x, $c^T x \leq y^T b$. We summarize this argument in the following lemma.

Lemma 1 *Let y be a column vector that satisfies $y \geq 0$ and $A^T y = c$, then for any x satisfying $Ax \leq b$ we have that $c^T x \leq y^T b$.*

Proof: We know that $b_i \geq a_i x$ for all i. Multiplying this inequality by the non-negative y_i, we get that $y_i b_i \geq y_i(a_i x)$. Adding these constraints up for all i, we get

$$\begin{aligned} \sum_{i=1}^m y_i b_i &\geq \sum_{i=1}^m y_i (a_i x) \\ &= \sum_{i=1}^m y_i \sum_{j=1}^n a_{ij} x_j \\ &= \sum_{j=1}^n \left(\sum_{i=1}^m y_i a_{ij} \right) x_j \\ &= \sum_{j=1}^n c_j x_j. \end{aligned}$$

More compactly,

$$y^T b \geq y^T (Ax) = (y^T A) x = (A^T y)^T x = c^T x.$$

□

Our goal is to derive the best upper bound for the linear program that can be derived this way, i.e., the upper bound $w = y^T b$ that has the smallest value of the problem. We can write this as another linear program: $\min(y^T b : A^T y = c, y \geq 0)$ which is called the *dual* linear program. The original linear program is called the *primal*. The lemma above then implies what is known as the weak duality theorem.