Cornell University, Fall 2016 **CS 6820: Algorithms**
Supplementary lecture notes on linear programming **26 Sep–28 Sep**

1 The Simplex Method

We will present an algorithm to solve linear programs of the form

$$\begin{aligned} \text{maximize} \quad & c^\intercal x \\ \text{subject to} \quad & Ax \preceq b \\ & x \succeq 0 \end{aligned} \tag{1}$$

assuming that $b \succeq 0$, so that $x = 0$ is guaranteed to be a feasible solution. Let n denote the number of variables and let m denote the number of constraints.

A simple transformation modifies any such linear program into a form such that each variable is constrained to be non-negative, and all other linear constraints are expressed as *equations* rather than *inequalities*. The key is to introduce additioinal variables, called *slack variables* which account for the difference between and left and right sides of each inequality in the original linear program. In other words, linear program (1) is equivalent to

$$\begin{aligned} \text{maximize} \quad & c^\intercal x \\ \text{subject to} \quad & Ax + y = b \\ & x, y \succeq 0 \end{aligned} \tag{2}$$

where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$.

The solution set of $\{Ax + y = b,\ x \succeq 0,\ y \succeq 0\}$ is a polytope in the $(n+m)$-dimensional vector space of ordered pairs $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$. The simplex algorithm is an iterative algorithm to solve linear programs of the form (2) by walking from vertex to vertex, along the edges of this polytope, until arriving at a vertex which maximizes the objective function $c^\intercal x$.

To illustrate the simplex method, for concreteness we will consider the following linear program.

$$\begin{aligned} \text{maximize} \quad & 2x_1 + 3x_2 \\ \text{subject to} \quad & x_1 + x_2 \le 8 \\ & 2x_1 + x_2 \le 12 \\ & x_1 + 2x_2 \le 14 \\ & x_1, x_2 \ge 0 \end{aligned}$$

This LP has so few variables, and so few constraints, it is easy to solve it by brute-force enumeration of the vertices of the polytope, which in this case is a 2-dimensional polygon.

The vertices of the polygon are $\left[\begin{smallmatrix}0\\7\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}2\\6\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}4\\4\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}6\\0\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right]$. The objective function $2x_1 + 3x_2$ is maximized at the vertex $\left[\begin{smallmatrix}2\\6\end{smallmatrix}\right]$, where it attains the value 22. It is also easy to certify that this is the optimal value, given that the value is attained at $\left[\begin{smallmatrix}2\\6\end{smallmatrix}\right]$: simply add together the inequalities

$$\begin{aligned} x_1 + x_2 &\le 8 \\ x_1 + 2x^2 &\le 14 \end{aligned}$$

to obtain

$$2x_1 + 3x_2 \le 22,$$

which ensures that no point in the feasible set attains an objective value greater than 22.

To solve the linear program using the simplex method, we first apply the generic transformation described earlier, to rewrite it in *equational form* as

$$\begin{array}{ll} \text{maximize} & 2x_1 + 3x_2 \\ \text{subject to} & x_1 + x_2 + y_1 = 8 \\ & 2x_1 + x_2 + y_2 = 12 \\ & x_1 + 2x_2 + y_3 = 14 \\ & x_1, x_2, y_1, y_2, y_3 \ge 0 \end{array}$$

From now on, we will be choosing a subset of two of the five variables (called the *basis*), setting them equal to zero, and using the linear equations to express the remaining three variables, as well as the objective function, as a function of the two variables in the basis. Initially the basis is $\{x_1, x_2\}$ and the linear program can be written in the form

$$\begin{array}{ll} \text{maximize} & 2x_1 + 3x_2 \\ \text{subject to} y_1 = 8 - x_1 - x_2 & \\ & y_2 = 12 - 2x_1 - x_2 \\ & y_3 = 14 - x_1 - 2x_2 \\ & x_1, x_2, y_1, y_2, y_3 \ge 0 \end{array}$$

which emphasizes that each of y_1, y_2, y_3 is determined as a function of x_1, x_2. Now, as long as the basis contains a variable which has a positive coefficient in the objective function, we select one such variable and greedily increasing its value until one of the non-negativity constraints becomes tight. At that point, one of the other variables attains the value zero: it enters the basis, and the variable whose value we increased leaves the basis. For example, we could choose to increase x_1 from 0 to 6, at which point $y_2 = 0$. Then the new basis becomes $\{y_2, x_2\}$. Rewriting the equation $y_2 = 12 - 2x_1 - x_2$ as

$$x_1 = 6 - \frac{1}{2}y_2 - \frac{1}{2}x_2, \tag{3}$$

we may substitute the right side of (3) in place of x_1 everywhere in the above linear program, arriving at the equivalent form

$$\begin{aligned} \text{maximize} \quad & 12 - y_2 + 2x_2 \\ \text{subject to} \quad & y_1 = 2 + \tfrac{1}{2}y_2 - \tfrac{1}{2}x_2 \\ & x_1 = 6 - \tfrac{1}{2}y_2 - x_2 \\ & y_3 = 8 + \tfrac{1}{2}y_2 - \tfrac{3}{2}x_2 \\ & x_1, x_2, y_1, y_2, y_3 \geq 0 \end{aligned}$$

At this point, x_2 still has a positive coefficient in the objective function, so we increase x_2 from 0 to 4, at which point $y_1 = 0$. Now x_2 leaves the basis, and the new basis is $\{y_1, y_2\}$. We use the equation $x_2 = 4 + y_2 - 2y_1$ to substitute a function of the basis variables in place of x_2 everywhere it appears, arriving at the new linear program

$$\begin{aligned} \text{maximize} \quad & 20 - 4y_1 + y_2 \\ \text{subject to} \quad & x_2 = 4 + y_2 - 2y_1 \\ & x_1 = 4 - y_2 + y_1 \\ & y_3 = 2 - y_2 + 3y_1 \\ & x_1, x_2, y_1, y_2, y_3 \geq 0 \end{aligned}$$

Now we increase y_2 from 0 to 2, at which point $y_3 = 0$ and the new basis is $\{y_1, y_3\}$. Substituting $y_2 = 2 - y_3 + 3y_1$ allows us to rewrite the linear program as

$$\begin{aligned} \text{maximize} \quad & 22 - y_1 - y_3 \\ \text{subject to} \quad & x_2 = 6 + y_1 - y_3 \\ & x_1 = 2 - 2y_1 + y_3 \\ & y_2 = 2 + 3y_1 - y_3 \\ & x_1, x_2, y_1, y_2, y_3 \geq 0 \end{aligned} \tag{4}$$

At this point, there is no variable with a positive coefficient in the objective function, and we stop.

It is trivial to verify that the solution defined by the current iteration—namely, $x_1 = 2$, $x_2 = 6$, $y_1 = 0$, $y_2 = 2$, $y_3 = 0$—is optimal. The reason is that we have managed to write the objective function in the form $22 - y_1 - y_3$. Since the coefficient on each of the variables y_1, y_3 is negative, and y_1 and y_3 are constrained to take non-negative values, the largest possible value of the objective function is achieved by setting both y_1 and y_3 to zero, as our solution does.

More generally, if the simplex method terminates, it means that we have found an equivalent representation of the original linear program (2) in a form where the objective function attaches a non-positive coefficient to each of the basis variables. Since the basis variables are required to be non-negative, the objective function is maximized by setting all the basis variables to zero, which certifies that the solution at the end of the final iteration is optimal.

Note that, in our running example, the final objective function assigned coefficient -1 to both y_1 and y_3. This is closely related to the fact that the simple "certificate of optimality" described above (before we started running the simplex algorithm) we obtained by summing the first and third inequalities of the original linear program, each with a coefficient of 1. We will see in the following section that this is not a coincidence.

Before leaving this discussion of the simplex method, we must touch upon a subtle issue regarding the question of whether the algorithm always terminates. A basis is an n-element subset of $n+m$ variables, so there are at most $\binom{n+m}{n}$ bases; if we can ensure that the algorithm never returns to the same basis as in a previous iteration, then it must terminate. Note that each basis determines a unique point $(x, y) \in \mathbb{R}^{n+m}$—defined by setting the basis variables to zero and assigning to the remaining variables the unique values that satisfy the equation $Ax + b = y$—and as the algorithm proceeds from basis to basis, the objective function value at the corresponding points never decreases. If the objective function strictly increases when moving from basis B to basis B', then the algorithm is guaranteed never to return to basis B, since the objective function value is now strictly greater than its value at B, and it will never decrease. On the other hand, it is possible for the simplex algorithm to shift from one basis to a different basis with the same objective function value; this is called a *degenerate pivot*, and it happens when the set of variables whose value is 0 at the current solution is a strict superset of the basis.

There exist *pivot rules* (i.e., rules for selecting the next basis in the simplex algorithm) that are designed to avoid infinite loops of degenerate pivots. Perhaps the simplest such rule is *Bland's rule*, which always chooses to remove from the basis the lowest-numbered variable that has a positive coefficient in the objective function. (And, in case there is more than one variable that may move into the objective function to replace it, the rule also chooses the lowest-numbered such variable.) Although the rule is simple to define, proving that it avoids infinite loops is not easy, and we will omit the proof from these notes.

2 The Simplex Method and Strong Duality

An important consequence of the correctness and termination of the simplex algorithm is *linear programming duality*, which asserts that for every linear program with a maximization objective, there is a related linear program with a minimization objective whose optimum matches the optimum of the first LP.

Theorem 1. *Consider any pair of linear programs of the form*

$$\begin{array}{lllll} \text{maximize} & c^{\intercal} x & & \text{minimize} & b^{\intercal} \eta \\ \text{subject to} & Ax \preceq b & \textit{and} & \text{subject to} & A^{\intercal} \eta \succeq c \\ & x \succeq 0 & & & \eta \succeq 0 \end{array} \tag{5}$$

If the optimum of the first linear program is finite, then both linear programs have the same optimum value.

Proof. Before delving into the formal proof, the following intuition is useful. If a_i denotes the i^{th} row of the matrix A, then the relation $Ax \preceq b$ can equivalently be expressed by stating that $a_i^\intercal x \leq b_i$ for $j = 1, \ldots, m$. For any m-tuple of non-negative coefficients $\eta_1, \ldots, \eta_m$, we can form a weighted sum of these inequalities,

$$\sum_{j=1}^{m} \eta_i a_i^\intercal x \leq \sum_{j=1}^{m} \eta_i b_i, \tag{6}$$

obtaining an inequality implied by $Ax \preceq b$. Depending on the choice of weights $\eta_1, \ldots, \eta_m$, the inequality (6) may or may not imply an upper bound on the quantity $c^\intercal x$, for all $x \succeq 0$. The case in which (6) implies an upper bound on $c^\intercal x$ is when, for each variable x_j ($j = 1, \ldots, n$), the coefficient of x_j on the left side of (6) is greater than or equal to the coefficient of x_j in the expression $c^\intercal x$. In other words, the case in which (6) implies an upper bound on $c^\intercal x$ for all $x \succeq 0$ is when

$$\forall j \in \{1, \ldots, n\} \quad \sum_{i=1}^{m} \eta_i a_{ij} \geq c_j. \tag{7}$$

We can express (6) and (7) more succinctly by packaging the coefficients of the weighted sum into a vector, η. Then, inequality (6) can be rewritten as

$$\eta^\intercal Ax \leq \eta^\intercal b, \tag{8}$$

and the criterion expressed by (7) can be rewritten as

$$\eta^\intercal A \succeq c^\intercal. \tag{9}$$

The reasoning surrounding inequalities (6) and (7) can now be summarized by saying that for any vector $\eta \in \mathbb{R}^m$ satisfying $\eta \succeq 0$ and $\eta^\intercal A \succeq c^\intercal$, we have

$$c^\intercal x \leq \eta^\intercal Ax \leq \eta^\intercal b \tag{10}$$

for all $x \succeq 0$ satisfying $Ax \preceq b$. (In hindsight, proving inequality (10) is trivial using the properties of the vector ordering $\preceq$ and our assumptions about x and η.)

Applying (10), we may immediately conclude that the minimum of $\eta^\intercal b$ over all $\eta \succeq 0$ satisfying $\eta^\intercal A \succeq c^\intercal$, is greater than or equal to the maximum of $c^\intercal x$ over all $x \succeq 0$ satisfying $Ax \preceq b$. That is, the optimum of the first LP in (5) is less than or equal to the optimum of the second LP in (5), a relation known as *weak duality.*

To prove that the optima of the two linear programs are equal, as asserted by the theorem, we need to furnish vectors x, η satisfying the constraints of the first and second linear programs in (5), respectively, such that $c^\intercal x = b^\intercal \eta$. To do so, we will make use of the simplex algorithm and its termination condition. At the moment of termination, the objective function has been rewritten in a form that has no positive coefficient on any variable. In other words, the objective function is written in the form $v - \xi^\intercal x - \eta^\intercal y$ for some coefficient vectors $\xi \in \mathbb{R}^n$ and $\eta \in \mathbb{R}^m$ such that $\xi, \eta \succeq 0$.

An invariant of the simplex algorithm is that whenever it rewrites the objective function, it preserves the property that the objective function value matches $c^\intercal x$ for all pairs $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ such that $Ax + y = b$. In other words, we have

$$\forall x \in \mathbb{R}^n \quad v - \xi^\intercal x - \eta^\intercal (b - Ax) = c^\intercal x. \tag{11}$$

Equating the constant terms on the left and right sides, we find that $v = \eta^\intercal b$. Equating the cofficient of x_j on the left and right sides for all j, we find that $\eta^\intercal A = \xi^\intercal + c^\intercal \succeq c^\intercal$. Thus, the vector η satisfies the constraints of the second LP in (5).

Now consider the vector (x, y) which the simplex algorithm outputs at termination. All the variables having a non-zero coefficient in the expression $-\xi^\intercal x - \eta^\intercal y$ belong to the algorithm's basis, and hence are set to zero in the solution (x, y). This means that

$$v = v - \xi^\intercal x - \eta^\intercal y = c^\intercal x$$

and hence, using the relation $v = \eta^\intercal b$ derived earlier, we have $c^\intercal x = b^\intercal \eta$ as desired. $\square$