Lecture 13: Introduction to Duality

Lecturer: *Sundar Vishwanathan*
Computer Science & Engineering Indian Institute of Technology, Bombay

Let x_0 be an extreme point. Suppose it is given by $A^{'}x_0 = b^{'}, A^{''}x_0 < b^{''}$. We call $A'x = b'$ as the defining hyperplanes of x_0.

We know that the neighbours of x_0 are along the columns of $-A^{'-1}$.

1 Proof of correctness of Simplex algorithm

Theorem 1 *If the cost does not increase along any of the columns of $-A^{'-1}$ then x_0 is optimal.*

Proof: The columns of $-A^{'-1}$ span R^n. Let x_{opt} be an optimal point. We need to show that $c^T x_{opt} \leq c^T x_0$. Since the columns of $-A^{'-1}$ form a basis of R^n (why?) the vector $x_{opt} - x_0$ can be represented as a linear combination of them say as given below.

$$x_{opt} - x_0 = \sum \beta_j (-A^{'-1})^{(j)} \tag{1}$$

Premultiplying both sides with A' yields

$$A'x_{opt} - A'x_0 = \sum \beta_j A'(-A^{'-1})^{(j)}. \tag{2}$$

We know that $A'x_{opt} \leq b'$ and $A'x_0 = b'$ hence $A'(x_{opt} - x_0) \leq 0$. That is, the lhs is a vector, all of whose components are non-positive. Also note that $A'(-A^{'-1})^{(j)}$ is an $n \times 1$ vector whose j^{th} element is -1 and remaining elements are 0. Hence

$$A'x_{opt} - A'x_0 = \begin{pmatrix} -\beta_1 \\ -\beta 2 \\ . \\ . \\ . \\ -\beta n \end{pmatrix} \tag{3}$$

This implies that $\beta_j \geq 0$ for all j.

Now,

$$c^T x_{opt} - c^T x_0 = \sum \beta_j c^T (-A^{'-1})^{(j)} \tag{4}$$

Since the cost decreases along the columns of $-A^{'-1}$ we have $c^T(-A^{'-1})^{(j)} \leq 0$ and since $\beta_j \geq 0$ we conclude that $\sum \beta_j c^T (-A^{'-1})^j \leq 0$. Hence $c^T x_{opt} \leq c^T x_0$, as desired. □

Note: From the above theorem we infer that when the Simplex algorithm terminates it gives us an optimal solution.

2 A Geometric Introduction to Duality

We will do some algebra before giving an intuitive geometric description.

Let x_0 be an optimal point. Using the termination condition of the simplex algorithm we know that the cost does not increase along the columns of $-A^{'-1}$. That is, $c^T(-A^{'-1}) \leq \bar{0}^T$. Multiplying both sides with -1 we can rewrite this as $c^T(A^{'-1}) \geq \bar{0}$. Hence

$$c^T(A^{'-1}) = (\alpha_1, \alpha_2, \ldots, \alpha_n) \geq \bar{0}^T. \tag{5}$$

Postmultiplying both sides by A', we get

$$c^T = (\alpha_1, \alpha_2, \ldots, \alpha_n)A'. \tag{6}$$

Where $\alpha_i \geq 0$ for each i. Taking transpose of both sides we rewrite this as

$$c = (A^{'})^T y; y \geq \bar{0} \tag{7}$$

We observe that at the optimal point the cost vector can be written as a *non-negative* linear combination of the rows of A'. Given a half-space $A_i x \leq b_i$, we note that the vector A_i is a normal to the hyperplane $A_i x = b_i$. In fact it is an *outward normal.* It points away from the feasible region. When we talk about a normal, we will always mean an outward normal.

To recap, given a vertex x, the cost is non-increasing along the directions towards the neighbours *iff* the cost vector can be written as a non-negative linear combination of the normals to the hyperplanes at x.

Hence an extreme point x_0 is optimal if and only if x_0 is feasible and the cost vector can be written as a non-negative linear combination of the rows of A', where A' are the defining inequalities of the point x_0.

This geometric fact is depicted in the figure below.

In the algebraic manipulations we did, notice that only A' plays a role. A'' does not play a role. The only place we use it is to say that x_0 is feasible.

Denote by F the set of all the points (not necessarily feasible) given by the intersection of n linearly independent hyperplanes from A where the cost vector can be written as a non-negative linear combination of the normals of the defining hyperplanes.

We define F in slightly different language below. Take any n linearly independent rows of A. This gives rise to a point say z. Note that z may not be feasible. Among all such points, we consider those z where we can write the cost vector as a positive linear combinations of the normals of the defining hyperplanes of z. This set we denote by F.

We will show first that in F, feasible points will have the lowest cost. From the previous discussion these identify the optimal points of the original LP.

Consider any point $z \in$ F. Let B' denote the matrix of the defining hyperplanes of z. The crucial observation that proves the above statement is this. The feasible region of the LP lies in the region (called the cone) generated by the the columns of $-B'^{-1}$. That is, each feasible point can be written as a non-negative linear combination of the columns of $-B'^{-1}$. (Hint to prove this: See an earlier part of this lecture.) Notice also that the cost is non-increasing along these directions. (Hint to prove this: See an earlier part of this lecture.)

Hence no feasible point will have cost greater than that of z. If z were also feasible, then we see that it will be an optimum vertex in the original LP.

Indeed we see that among all points in F a point of minimum cost will be feasible, and this feasible point is an optimal point for the LP.

Let us investigate the cost at points in F. Consider such a point z which is characterised by $A'z = \mathbf{b}'$. Note that z may not satisfy some of the other constraints. We also know that there exists some

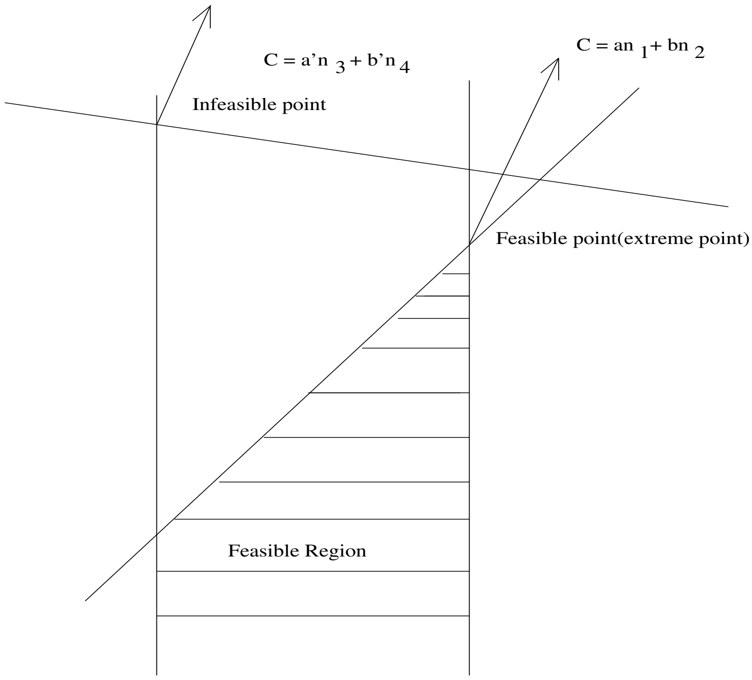

Figure 1: Cost vector (the arrow) can be written as a positive linear combination of normals to the hyperplanes. Normals are not shown though.

$\alpha_1, \alpha_2, \ldots, \alpha_n$ such that $\alpha_i \geq 0$ for each i and $c^T = \Sigma_{i=1}^{n} \alpha_i A'_i$. Consider a vector y with m co-ordinates. The jth co-ordinate of y is α_s if the jth row of A is the sth row of A' and zero otherwise. For such a y, note that $c^T = \Sigma_{i=1}^{m} y_i A_i$ and $y_i \geq 0$. The cost at z is $c^T z = \Sigma_{i=1}^{m} y_i A_i z$. Now note that whenever $y_i > 0$, $A_i z = b_i$. Hence the cost $c^T z = \Sigma_{i=1}^{m} y_i b_i = y^T b$.

This motivates the definition of the following LP called the dual:

$$\text{minimize} : y^T \mathbf{b} \tag{8}$$

$$A^T y = c \tag{9}$$

$$y \geq 0 \tag{10}$$

We have shown above that for each point in F there is a feasible point in the new LP with the same cost. We will also show a kind of converse. Consider the extreme points of this LP. We will show that to each extreme point we can find a point in F of the original LP with the same cost. Note that the costs are in different domains.

What can we say about the extreme points of this LP? We note that the dimension of the LP is m. There are n equalities which every feasible y must satisfy, viz, $A^T y = c$. Hence at an extreme point at least $m - n$ of the other inequalities must be equalities, that is at least $m - n$ of the y_is must be zero.

To each such extreme point, we can associate a point z in F as follows. At most n of the y_is are greater than zero. Pick a subset of hyperplanes from *lpset*, as follows. Include $A_i x = b_i$ if $y_i > 0$. There are at most n of them. Add more if needed to get n linearly independent hyperplanes. These will define a point $z \in$ F corresponding to y. Why? As before one can prove $c^T z = b^T y$. Do this. Note that there may be many points corresponding to one point y.

We note that even if we assume that the primal is non-degenerate, the dual may be degenerate. Suppose we say that the dual is also non-degenerate. What modifications will you make to the discussion above?

We will prove formally in the next lecture that the optimum points of both LPs will coincide and the costs will be equal. It will be worthwhile trying to prove this before reading ahead. This, but for technicalities, is called the duality theorem in Linear Programming.