Lecture 2

Uri Feige

April 24, 2002

1 The geometry of linear programs

In Lecture 1 we presented linear programs from a linear algebra point of view. The use of a geometric representation helps build intuition about linear programming. This is easiest to visualize in two dimensions (when there are only two variables, or only two constraints), and still manageable in three dimensions. Much of the low dimension intuition is also true in higher dimensions, and can be formalized and applied for linear programs with an arbitrary number of dimensions.

1.1 Two variables or two constraints

The constraints of a linear program with two variables in canonical form can be drawn as half-planes, and the feasible region as an intersection of half-planes. The objective function can be represented either as a vector pointing in its direction, or as equivalue lines.

The vertices will correspond to bfs for the linear program, when transformed to standard form (using slack variables). A degeneracy is the result of three constraints intersecting at a point on the boundary of the feasible region. (When transforming the LP from canonical form to standard form, this will give three slack variables that are 0, which will be larger than $n - m$.)

For linear programs in standard form, the two-dimensional representation above becomes uninteresting. There can be at most two equality constraints, and the feasible region degenerates to a point.

When there are two constraints in standard form, each column of the matrix A (which corresponds to a variable) can be viewed as a point in the plane. The vector b is another point in the plane. A feasible solution is two vectors that can be added (with nonnegative coefficients) to give the point b. (In canonical form, a feasible solution is a set of vectors that can be added to give a point in the quarter plane to the right and above b.) There is a degenerate solution if a single vector points in the direction of b. (More than $n - m$ variables can be set to 0.)

1.2 Some standard terminology

A point $x \in R^n$ is a *convex combination* of points $x_1, \ldots, x_t$ if there are some $\lambda_i \geq 0$ with $\sum_{i=1}^t \lambda_i = 1$ such that $x = \sum_{i=1}^t \lambda_i x_i$. A set S is *convex* if for every two points $x, y \in S$, every convex combination of them (namely, their *convex hull*) is in S. The intersection of convex sets is a convex set.

A linear (vector) subspace in R^n is a set of vectors closed under addition and under multiplication by scalars. Equivalently, it is the set of solutions to a system of homogeneous linear equations $Ax = 0$. It is a convex set. The dimension of a linear space is the maximum number of linearly independent vectors that it contains. It can be shown that this is equal to n minus the rank (number of linearly independent rows) of the matrix A above. For a vector $b \in R^n$ and vector space $S \in R^n$, the set $\{b + s | s \in S\}$ is called an affine subspace. It has the same dimension as S. Equivalently, an affine subspace is the set of solutions to a set of linear equations $Ax = b$. The dimension of a set $T \in R^n$ is the dimension of the minimal affine subspace containing it. In particular, the dimension of the set of solutions to a linear program in standard form (with linearly independent constraints) is at most $n - m$.

For a vector $a \in R^n$, $a \neq 0$, and $b \in R$, the set of points x satisfying $a^T x = b$ is called a *hyperplane*. It is an affine subspace of dimension $n - 1$. The set of points satisfying $a^T x \geq b$ is a halfspace. Hyperplanes and halfspaces are convex. Observe that a hyperplane is the intersection of two halfspaces.

The intersection of halfspaces is a *polyhedron*. It follows that the set of contraints of a linear program define a polyhedron which is the region of feasible solutions. It is a convex set. A polyhedron P that is *bounded* (namely, for some scalar c, $x^T x \leq c$ for all $x \in P$) is a *polytope*.

We remark that polyhedra are closed sets (because our inequalities are not strict), implying that the maximum of continuous functions defined over P is attained by some point in P.

Let P be a polytope of dimension d in R^n, HS a halfspace supported by hyperplane H. If $f = P \bigcap HS$ is contained in H and not equal to P, then f is a face of P and H is a supporting hyperplane or P. A *facet* is a face of dimension $d - 1$. An *edge* is a face of dimension 1. A *vertex* is a face of dimension 0. Equivalently, we can define a vertex of a polytope (or polyhedron) P as point $x \in P$ such that for some vector $c \in R^n$, $c^T x < c^t y$ for every $y \in P - \{x\}$. The equivalence between the two definitions can be seen by taking c to be the normal of H at x, pointing inwards towards the polytope.

An *extreme point* of a polyhedron P is a point $x \in P$ such that there are no other two points $y, z \in P$ such that x is a convex combination of them.

Recall that in lecture 1 we defined basic feasible solutions for linear programs in standard form. This can be generalized to any linear program, where we define a BFS as a feasible solution for which n linearly independent constraints are tight (including nonnegativity constraints).

Lemma 1 *For a polyhedron P and a point $x \in P$, the following three statements are equivalent.*

1. *x is an extreme point of P.*
2. *x is a vertex of P.*
3. *x is a basic feasible solution.*

Proof: We shall prove the lemma for an LP in standard form. The proof for an LP in general form is left as homework.

BFS implies vertex. This is essentially Lemma 1 from Lecture 1.

Vertex implies extreme point. If w is a vertex, then for some vector c, $c^T x$ is uniquely minimized over P at w. If w were not extreme, then we could write $w = \lambda y + (1 - \lambda)z$ for some $y, z \in P - \{w\}$ and $0 < \lambda < 1$, and then for either y or z, their inner product with c would be smaller. A contradiction.

Extreme point implies BFS. We show that not BFS implies not extreme. Let $w = (w_1, \dots, w_n) \in P$ not be a BFS. Let x^+ be the set of nonzero variables in w, and let A^+ be the columns of A that correspond to the nonzero variables. The columns in A^+ are linearly dependent (as w is not a BFS). Hence there is a nonzero solution to $A^+x^+ = 0$. Call it d^+. Let d be a vector that agrees with d^+ on the variables of x^+, and is zero elsewhere. Consider the points $y = w + \epsilon d$ and $z = w - \epsilon d$, where ϵ is chosen small enough to ensure that $y \geq 0$ and $z \geq 0$. Both $y, z \in P$, and $w = y/2 + z/2$, showing that w is not an extreme point. $\square$

Another useful way of characterizing a polytope is as the convex hull of all its vertices.

Lemma 2 *For a polytope P of dimension d, every point in it is the convex hull of at most $d + 1$ of its vertices.*

Proof: The proof is by induction on d. The base case (dimension 0, a single vertex) is trivial. For the inductive step, we need the following facts:

1. Every polytope has a vertex. (We have seen something similar in lecture 1 – the existence of basic feasible solutions for LPs in standard form.)

2. Every face of a polytope is a polytope of lower dimension. (The face satisfies a linear equality that is not satisfied by some points in P.)

Take an arbitrary point $x \in P$. Take an arbitrary vertex v of the polytope. Follow the line from v through x (this portion must lie entirely within the polytope) until is hits a face of P (which is must do, as the polytope in bounded) at a point z. Then x is a convex combination of v and z. By induction, z is a convex combination of at most d vertices, showing that w is a convex combination of at most $d + 1$ vertices. $\square$