Lecture 2

Geometry of LPs*

Last time we saw that, given a (minimizing) linear program in equational form, one of the following three possibilities is true:

1. The LP is infeasible.
2. The optimal value of the LP is $-\infty$ (i.e., the LP does not have a bounded optimum).
3. A basic feasible solution exists that achieves the optimal value.

2.1 Finding a basic feasible solution

Suppose we have an LP in equational form:

$$\min\{\, c^T x \mid Ax = b,\ x \geq 0 \,\}, \quad \text{where } A \text{ is an } m \times n \text{ matrix of rank } m.$$

Recall how we might attempt to find a BFS to this LP: Let $B \subseteq [n]$, with $|B| = m$, such that A_B (the set of columns of A corresponding to the elements of B) is linearly independent. (Such a set of columns exists because A has full rank.) Let $N = [n] \setminus B$ be the indices of the columns of A that are not in B. Since A_B is invertible, we can define a vector $x \in \mathbb{R}^n$ by

$$x_B = A_B^{-1} b,$$
$$x_N = 0.$$

By construction, x is a basic solution. If x is also feasible (i.e., if $x \geq 0$), then it is a BFS.

Fact 2.1. *Every LP in equational form that is feasible has a BFS. (Note that this BFS may or may not be optimal.)*

Proof. Pick some feasible point $\tilde{x} \in \mathbb{R}^n$. (In particular, since $\tilde{x}$ is feasible, $\tilde{x} \geq 0$.) Let

$$P = \{\, j \mid \tilde{x}_j > 0 \,\}$$

*Lecturer: Anupam Gupta. Scribe: Brian Kell.

be the set of coordinates of $\tilde{x}$ that are nonzero. We consider two cases, depending on whether the columns of A_P are linearly independent.

Case 1. The columns of A_P are linearly independent. Then we may extend P to a basis B, i.e., a subset $P \subseteq B \subseteq [n]$ with $|B| = m$ such that the columns of A_B are also linearly independent. Let $N = [n] \setminus B$; then $\tilde{x}_N = 0$ (because $P \subseteq B$). In addition, since $A\tilde{x} = b$, we have $A_B\tilde{x}_B = b$, so $\tilde{x} = A_B^{-1}b$. So $\tilde{x}$ is a basic solution; since it is feasible by assumption, it is a BFS. (Note, by the way, that the equation $\tilde{x} = A_B^{-1}b$ means that $\tilde{x}$ is the unique solution to $Ax = b$ having $x_N = 0$.)

Case 2. The columns of A_P are linearly dependent. Let $N = [n] \setminus P$. Then, by the definition of linear dependence, there exists a nonzero vector $w \in \mathbb{R}^n$ with $w_N = 0$ such that $A_P w_P = 0$. For any $\lambda \in \mathbb{R}$, the vector $\tilde{x} + \lambda w$ satisfies $A(\tilde{x} + \lambda w) = b$, because

$$A(\tilde{x} + \lambda w) = A\tilde{x} + \lambda Aw = b + 0 = b.$$

Because $\tilde{x}_N = 0$ and $w_N = 0$, we have $(\tilde{x} + \lambda w)_N = 0$, so $\tilde{x} + \lambda w$ has no more nonzero entries than $\tilde{x}$ does. Since $\tilde{x}_P > 0$, for sufficiently small $\epsilon > 0$ both $\tilde{x} + \epsilon w$ and $\tilde{x} - \epsilon w$ are feasible (i.e, $\tilde{x} \pm \epsilon w \geq 0$). Let $\eta = \sup\{\, \epsilon > 0 \mid \tilde{x} \pm \epsilon w \geq 0 \,\}$ be the largest such ϵ; then one of $\tilde{x} \pm \eta w$ has one more zero coordinate than $\tilde{x}$ does. We can repeat this until we find a feasible solution with no more than m nonzero coordinates, at which point Case 1 applies and we have found a BFS.

(Intuitively, for sufficiently small $\epsilon > 0$, one of $\tilde{x} \pm \epsilon w$ is moving toward a nonnegativity constraint, that is, toward the boundary of the nonnegative orthant. When ϵ becomes just large enough that the point $\tilde{x} \pm \epsilon w$ reaches the boundary of the nonnegative orthant, we have made one more coordinate of the point zero.) □

2.2 Geometric definitions

Definition 2.2. Given points $x, y \in \mathbb{R}^n$, a point $z \in \mathbb{R}^n$ is a *convex combination* of x and y if

$$z = \lambda x + (1 - \lambda)y$$

for some $\lambda \in [0, 1]$.

Definition 2.3. A set $X \subseteq \mathbb{R}^n$ is *convex* if the convex combination of any two points in X is also in X; that is, for all $x, y \in X$ and all $\lambda \in [0, 1]$, the point $\lambda x + (1 - \lambda)y$ is in X.

Definition 2.4. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if for all points $x, y \in \mathbb{R}^n$ and all $\lambda \in [0, 1]$ we have

$$f\big(\lambda x + (1 - \lambda)y\big) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Fact 2.5. *If $P \subseteq \mathbb{R}^n$ is a convex set and $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function, then, for any $t \in \mathbb{R}$, the set*

$$Q = \{\, x \in P \mid f(x) \leq t \,\}$$

is also convex.

Proof. For all $x_1, x_2 \in Q$ and all $\lambda \in [0,1]$, we have

$$f\big(\lambda x_1 + (1-\lambda)x_2\big) \le \lambda f(x_1) + (1-\lambda) f(x_2) \le \lambda t + (1-\lambda)t = t,$$

so $\lambda x_1 + (1-\lambda)x_2 \in Q$. □

Fact 2.6. *The intersection of two convex sets is convex.*

Proof. Let $P, Q \subseteq \mathbb{R}^n$ be convex sets, and let $x_1, x_2 \in P \cap Q$. Let $\lambda \in [0,1]$. Because $x_1, x_2 \in P$ and P is convex, we have $\lambda x_1 + (1-\lambda)x_2 \in P$; likewise, $\lambda x_1 + (1-\lambda)x_2 \in Q$. So $\lambda x_1 + (1-\lambda)x_2 \in P \cap Q$. □

Definition 2.7. A set $S \subseteq \mathbb{R}^n$ is a *subspace* if it is closed under addition and scalar multiplication.

Equivalently, S is a subspace if $S = \{\, x \in \mathbb{R}^n \mid Ax = 0 \,\}$ for some matrix A.

Definition 2.8. The *dimension* of a subspace $S \subseteq \mathbb{R}^n$, written $\dim(S)$, is the size of the largest linearly independent set of vectors contained in S.

Equivalently, $\dim(S) = n - \operatorname{rank}(A)$.

Definition 2.9. A set $S' \subseteq \mathbb{R}^n$ is an *affine subspace* if $S' = \{\, x_0 + y \mid y \in S \,\}$ for some subspace $S \subseteq \mathbb{R}^n$ and some vector $x_0 \in \mathbb{R}^n$. In this case the *dimension* of S', written $\dim(S')$, is defined to equal the dimension of S.

Equivalently, S' is an affine subspace if $S' = \{\, x \in \mathbb{R}^n \mid Ax = b \,\}$ for some matrix A and some vector b.

Definition 2.10. The *dimension* of a set $X \subseteq \mathbb{R}^n$, written $\dim(X)$, is the dimension of the minimal affine subspace that contains X.

Note that if S'_1 and S'_2 are two affine subspaces both containing X, then their intersection $S'_1 \cap S'_2$ is an affine subspace containing X. Hence there is a unique minimal affine subspace that contains X, so $\dim(X)$ is well defined.

Equivalently, given $x_0 \in X$, the dimension of X is the largest number k for which there exist points $x_1, x_2, \ldots, x_k \in X$ such that the set $\{x_1 - x_0, x_2 - x_0, \ldots, x_k - x_0\}$ is linearly independent.

Note that the definition of the dimension of a set X agrees with the definition of the dimension of an affine subspace if X happens to be an affine subspace, and the definition of the dimension of an affine subspace S' agrees with the definition of the dimension of a subspace if S' happens to be a subspace.

Definition 2.11. A set $H \subseteq \mathbb{R}^n$ is a *hyperplane* if $H = \{\, x \in \mathbb{R}^n \mid a^T x = b \,\}$ for some nonzero $a \in \mathbb{R}^n$ and some $b \in \mathbb{R}$.

A hyperplane is an affine subspace of dimension $n - 1$.

Definition 2.12. A set $H' \subseteq \mathbb{R}^n$ is a (closed) *halfspace* if $H' = \{\, x \in \mathbb{R}^n \mid a^T \ge b \,\}$ for some nonzero $a \in \mathbb{R}^n$ and some $b \in \mathbb{R}$.

A hyperplane can be written as the intersection of two halfspaces:

$$\{\, x \in \mathbb{R}^n \mid a^T x = b \,\} = \{\, x \in \mathbb{R}^n \mid a^T x \geq b \,\} \cap \{\, x \in \mathbb{R}^n \mid -a^T x \geq -b \,\}.$$

Both hyperplanes and halfspaces are convex sets. Therefore the feasible region of an LP is convex, because it is the intersection of halfspaces and hyperplanes. The dimension of the feasible region of an LP in equational form, having n variables and m linearly independent constraints (equalities), is no greater than $n - m$, because it is contained in the intersection of m distinct hyperplanes, each of which is an affine subspace of dimension $n - 1$. (The dimension of the feasible region may be less than $n - m$, because of the nonnegativity constraints, for instance.)

For example, the region in $\mathbb{R}^3$ defined by

$$\begin{aligned} x_1 + x_2 + x_3 &= 1, \\ x &\geq 0 \end{aligned}$$

is a 2-dimensional triangle; here, $n - m = 3 - 1 = 2$. (Note, however, that if the constraint were $x_1 + x_2 + x_3 = 0$, the region would have dimension 0.)

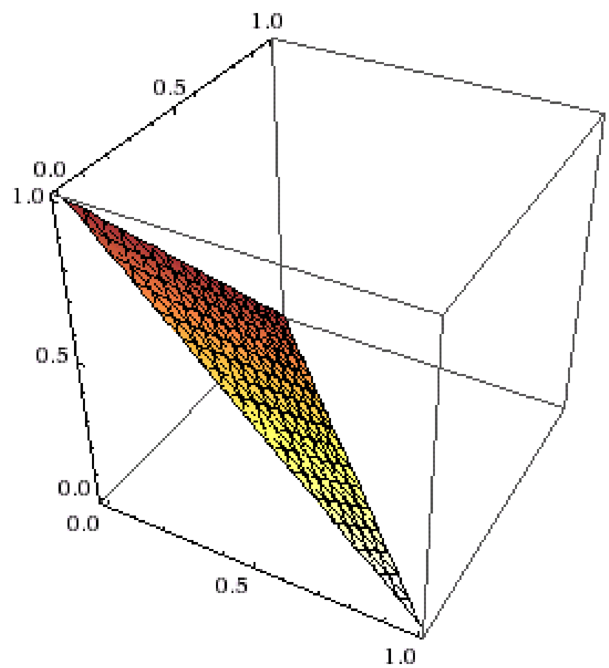

Figure 2.1: The region $\{\, x \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 1,\ x \geq 0 \,\}$.

Definition 2.13. A *polyhedron* in $\mathbb{R}^n$ is the intersection of finitely many halfspaces.

For example, feasible regions of LPs are polyhedra.

Definition 2.14. A *polytope* is a bounded polyhedron, that is, a polyhedron P for which there exists $B \in \mathbb{R}^+$ such that $\|x\|_2 \leq B$ for all $x \in P$.

Both polyhedra and polytopes are convex.

Definition 2.15. Given a polyhedron $P \subseteq \mathbb{R}^n$, a point $x \in P$ is a *vertex* of P if there exists $c \in \mathbb{R}^n$ such that $c^T x < c^T y$ for all $y \in P$, $y \neq x$.

Suppose $\hat{x}$ is a vertex of a polyhedron $P \subseteq \mathbb{R}^n$. Let c be as in the definition above. Take $K = c^T \hat{x}$. Then for all $y \in P$ we have $c^T y \geq K$, so the polyhedron P is contained in the halfspace $\{ x \in \mathbb{R}^n \mid c^T x \geq K \}$, i.e., P lies entirely on one side of the hyperplane $\{ x \in \mathbb{R}^n \mid c^T x = K \}$. Furthermore, the vertex $\hat{x}$ is the *unique* minimizer of the function $c^T z$ for $z \in P$.

Definition 2.16. Given a polyhedron $P \subseteq \mathbb{R}^n$, a point $x \in P$ is an *extreme point* of P if there do not exist points $u, v \neq x$ in P such that x is a convex combination of u and v.

In other words, x is an extreme point of P if, for all $u, v \in P$,

$$\big(x = \lambda u + (1-\lambda)v \text{ for some } \lambda \in [0,1]\big) \quad \text{implies} \quad u = v = x.$$

2.3 Equivalence of vertices, extreme points, and basic feasible solutions

In fact, vertices and extreme points are the same thing, and for an LP the vertices (i.e., extreme points) of its feasible region are precisely its basic feasible solutions. This is shown in the following theorem.

Theorem 2.17. *Consider an LP in equational form, i.e.,* $\min\{ c^T x \mid Ax = b,\, x \geq 0 \}$, *and let* K *be its feasible region. Then the following are equivalent:*

1. *The point* x *is a vertex of* K.
2. *The point* x *is an extreme point of* K.
3. *The point* x *is a BFS of the LP.*

Proof. *(1)* $\Rightarrow$ *(2)*. Let x be a vertex of K. Then there exists $c \in \mathbb{R}^n$ such that $c^T x < c^T y$ for all $y \in K$, $y \neq x$. Suppose for the sake of contradiction that x is not an extreme point of K, so there exist some points $u, w \in K$ with $u, w \neq x$ and some $\lambda \in [0,1]$ such that $x = \lambda u + (1-\lambda)w$. Then

$$c^T x = \lambda c^T u + (1-\lambda) c^T w < \lambda c^T x + (1-\lambda) c^T x = c^T x,$$

which is a contradiction. Hence x is an extreme point of K.

(2) $\Rightarrow$ *(3)*. Let x be an extreme point of K. (In particular, x is a feasible solution for the LP, so $x \geq 0$.) Let $P = \{ j \mid x_j > 0 \}$ be the set of nonzero coordinates of x. We consider two cases, depending on whether A_P (the set of columns of A corresponding to P) is linearly independent.

Case 1. The columns of A_P are linearly independent. Then x is a BFS. (This is the same as in the proof of Fact 2.1: Extend P to a basis B, and let $N = [n] \setminus B$; then $x_B = A_B^{-1}b$ and $x_N = 0$.)

Case 2. The columns of A_P are linearly dependent. Then there exists a nonzero vector w_P such that $A_P w_P = 0$. Let $N = [n] \setminus P$ and take $w_N = 0$. Then $Aw = A_P w_P + A_N w_N = 0$. Now consider the points $y^+ = x + \lambda w$ and $y^- = x - \lambda w$, where $\lambda > 0$ is sufficiently small so that $y^+, y^- \geq 0$. Then

$$\begin{aligned} Ay^+ &= A(x + \lambda w) = Ax + \lambda Aw = b + 0 = b, \\ Ay^- &= A(x - \lambda w) = Ax - \lambda Aw = b - 0 = b, \end{aligned}$$

so y^+ and y^- are feasible, i.e., $y^+, y^- \in K$. But $x = (y^+ + y^-)/2$ is a convex combination of y^+ and y^-, which contradicts the assumption that x is an extreme point of K. So Case 2 is impossible.

(3) $\Rightarrow$ *(1)*. Suppose x is a BFS for the LP. We aim to show that x is a vertex of K, that is, there exists $c \in \mathbb{R}^n$ such that $c^T x < c^T y$ for all $y \in P$, $y \neq x$. Since x is a BFS, there exists a set $B \subseteq [n]$ with $|B| = m$ such that the columns of A_B are linearly independent, $A_B x_B = b$, and $x_N = 0$ (where $N = [n] \setminus B$). For $j = 1, \ldots, n$, define

$$c_j = \begin{cases} 0, & \text{if } j \in B; \\ 1, & \text{if } j \in N. \end{cases}$$

Note that $c^T x = c_B^T x_B + c_N^T x_N = 0^T x_B + c_N^T 0 = 0$. For $y \in K$, we have $y \geq 0$ (since y is feasible), and clearly $c \geq 0$, so $c^T y \geq 0 = c^T x$. Furthermore, if $c^T y = 0$, then y_j must be 0 for all $j \in N$, so $A_B y_B = b = A_B x_B$. Multiplying on the left by A_B^{-1} gives $y_B = A_B^{-1} b = x_B$. So x is the unique point in K for which $c^T x = 0$. Hence x is a vertex of K. □

Definition 2.18. A polyhedron is *pointed* if it contains at least one vertex.

Note that a polyhedron contains a (bi-infinite) line if there exist vectors $u, d \in \mathbb{R}^n$ such that $u + \lambda d \in K$ for all $\lambda \in \mathbb{R}$.

Theorem 2.19. *Let $K \subseteq \mathbb{R}^n$ be a polyhedron. Then K is pointed if and only if K does not contain a (bi-infinite) line.*

Note that the feasible region of an LP with nonnegativity constraints, such as an LP in equational form, cannot contain a line. So this theorem shows (again) that every LP in equational form that is feasible has a BFS (Fact 2.1).

2.4 Basic feasible solutions for general LPs

Note that we've defined basic feasible solutions for LPs in equational form, but not for general LPs. Before we do that, let us make an observation about equational LPs, and the number of *tight* constraints (i.e., those constraints that are satisifed at equality).

Consider an LP in equational form with n variables and m constraints, and let x be a BFS. Then x satisfies all m equality constraints of the form $a_i x = b_i$. Since $x_N = 0$, we see that x additionally satisfies at least $n - m$ nonnegativity constraints at equality. A constraint is said to be *tight* if it is satisfied at equality, so we have the following fact.

Fact 2.20. *If x is a BFS of an LP in equational form with n variables, then x has at least n tight constraints.*

We can use this idea to extend the definition of BFS to LPs that are not in equational form.

Definition 2.21. For a general LP with n variables, i.e., an LP of the form

$$\min\{\, c^T x \mid Ax \geq b,\, x \in \mathbb{R}^n \,\},$$

a point $x \in \mathbb{R}^n$ is a *basic feasible solution* if x is feasible and there exist some n linearly independent constraints that are tight (hold at equality) for x.

Proposition 2.22. *For an LP in equational form, this definition of BFS and the previous definition of BFS are equivalent.*

(You may want to prove this for yourself.) Using this definition, one can now reprove Theorem 2.17 for general LPs: i.e., show the equivalence between BFSs, vertices, and extreme points holds not just for LPs in equational form, but for general LPs. We can use this fact to find optimal solutions for LPs whose feasible regions are pointed polyhedra (and LPs in equational form are one special case of this).