Mathematical Programming	**Lecture 3**
OR&IE 630, Fall 2005	**September 01, 2005**
Instructor: Mike Todd	Scribe: Yogeshwer Sharma

It is the theory that decides what can be observed [or proved].

—*Albert Einstein*

A test of a good theory is that the number of theorems in it should be at least twice the number of definitions.

—*Source unknown (quoted by Mike Todd)*

1 Some forms of linear programs

In its full generality, linear programming concerns optimizing a linear function over linear constraints (equalities and inequalities). But restricting the program to equalities or inequalities only does not reduce the power of linear programs. In this section we will see some of the more-used forms of linear programs and why they solve the linear programming problem in general.

1.1 LP in standard form

In its "standard" form, linear programming concerns itself with *miminizing* a linear function with *equality* and *non-negativity* constraints. More precisely, in standard form, a linear program looks like the following.

$$\begin{array}{llcll} \text{minimize} & c^\top x & & & \text{(minimization)} \\ \text{subject to} & Ax & = & b & \text{(equality constraints)} \\ & x & \geq & 0 & \text{(non-negativity constraints).} \end{array}$$

The dual of such an LP can be written in the following form,

$$\begin{array}{llcll} \text{maximize} & b^\top y & & & \\ \text{subject to} & A^\top y & \leq & c & \text{(inequality constraints)} \\ & y & & \text{unrestricted} & \text{(no sign constraints on variables).} \end{array}$$

1.2 LP in symmetric dual form

In the symmetric dual form, an LP problem is a minimization problem with inequality and non-negativity constraints. This is called the symmetric the dual form because dual of such an LP also has inequality and non-negativity constraints.

$$\begin{array}{llcl} \text{minimize} & c^\top x & & \\ \text{subject to} & Ax & \geq & b \\ & x & \geq & 0. \end{array}$$

And the dual of an LP problem in this form can be expressed as

$$\begin{array}{lrcl} \text{maximize} & b^\top y & & \\ \text{subject to} & A^\top y & \leq & c \\ & y & \geq & 0. \end{array}$$

1.3 Conversion from one form to another

In this section, we will consider the problem of converting one form of LP to another. In this process, it will also be clear that LP in standard or symmetric dual form can solve the linear programming problem in full generality. In Section 1.3.1, we will see the reduction from Linear Programming to Linear Programming in standard form.

Unrestricted variables to non-negative variables. If a variable (say x_j) is unrestricted, then we can replace each occurance of x_j by $s-t$ and add non-negativity constraints on s and t: $s \geq 0, t \geq 0$.

Equality to inequality constraints. Equality constrints can be changed to inequality constraints by putting both greater than and less than constraints. For example, $a^\top x = b$ can be changed to $a^\top x \leq b$ and $a^\top x \geq b$.

Inequality constraints to equality constraints. Inequality constraints can be changed to equality constraints by introducing slack variables. For example, consider the following two constraints

$$\begin{array}{rcl} a^\top x & \leq & b \\ a'^\top x & \geq & b'. \end{array}$$

We introduce two "slack" variables which correspond to how much is the slack (difference) between $a^\top x$ and b.

$$\begin{array}{rcl} a^\top x + s & = & b \\ s & \geq & 0 \\ a'^\top x - s' & = & b' \\ s' & \geq & 0. \end{array}$$

Here, s is called a slack variable and s' is called a surplus variable.

Maximization problem to minimization problem. Maximize $c^\top x$ can be transformed to minimize $-c^\top x$.

1.3.1 Transformation to LP in standard form

We consider the problem of converting LP in general form to LP in standard form. Let us call the original LP problem (LP).

1. First of all, if (LP) is a maximization problem, we multiply the objective function by -1 and make it a minimization problem.

2. As the next step, we change all the inequality constraints to equality constraints by introducing slack or surplus variables. This might introduce some non-negativity constraints too.

3. Finally, we change any unrestricted variables to non-negative variables by the trick mentioned in the last subsection.

2 Geometry of linear programs

We switch gears to a very interesting way of looking at LP now and consider the geometry of linear programs in more detail. We will look at subsets of $\mathbb{R}^n$ from two different angles and derive connections between them. The two ways of looking at the sets are the following. Let $S \subseteq \mathbb{R}^n$.

1. S can be looked as being generated by some finite number of points by taking various "combinations" of them. For example, $\mathbb{R}^2$ can be generated by $(1;0)$ and $(0;1)$ by taking all linear combinations. (We will explain the meaning of linear combination in the coming sections.)

2. S can also be specified by finite number of constraints, which every point in S must satisfy. For example, the non-negative quadrant of $\mathbb{R}^2$ can be written as the subset of $\mathbb{R}^2$ satisfying $x_1 \geqslant 0$ and $x_2 \geqslant 0$.

But, what type of combinations of points we can take? What kind of constraints are allowed? Let us define some concepts now which will allow us to answer these questions.

Definition 1. *Let x_1, x_2, ..., x_k be arbitrary k vectors in $\mathbb{R}^n$, and λ_1, λ_2, ..., $\lambda_k \in \mathbb{R}$. A vector x defined as*

$$x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k = \sum_{i=1}^{k} \lambda_i x_i$$

is called a linear combination *of the vectors $x_i, i = 1, ..., k$.*

1. *A linear combination is called an* affine combination *if $\sum_{i=1}^{k} \lambda_i = 1$.*

2. *A linear combination if called a* convex combination *if all λ_i sum to one and they are all non-negative. That is, $\sum_{i=1}^{k} \lambda_i = 1$ and $\lambda_i \geq 0$ for all i.*

3. *A linear combination if called a* conical combination (or non-negative combination) *if all the λ_i are non-negative. That is, $\lambda_i \geq 0$ for all i.*

4. *A linear combination if called a* positive combination *if all the λ_i's are positive. That is, $\lambda_i > 0$ for all i.*

To make the above ideas more concrete, we consider a simple example and consider various combinations.

Example 2. Let us consider the following two vectors in $\mathbb{R}^3$: $x_1 = (1; 0; 0)$ and $x_2 = (0; 1; 0)$. The set of all linear combinations of x_1 and x_2 is $\{x \in \mathbb{R}^3 | x_3 = 0\}$, a plane in $\mathbb{R}^3$. The set of all affine combinations is the line passing through the tip of the two vectors (through points (1; 0; 0) and (0; 1; 0)). The set of all convex combinations is the part of that line falling between the two points (For any point x, λ_1 and λ_2 are related to the distances of the point from x_2 and x_1). The set of all conical combinations is the non-negative quadrant of the plane $x_3 = 0$, or more formally, it is the set of all points $x = (x_1; x_2; x_3)$ satisfying $x_1 \geq 0, x_2 \geq 0, x_3 = 0$. The set of all positive combinations is the same set as that of conical combinations except for leaving out the boundary. More precisely, it is the set of all points $(x_1; x_2; x_3)$ satisfying $x_1 > 0, x_2 > 0, x_3 = 0$.

Comment about duality In the coming lecture, we will see a lot of duality in the sense that the same set (subspace) can be written as being generated by finite number of points by taking all combinations of some sort or it can also be written as specifying the constraints that must be satisfied in order for a point to be in the set (subspace). In the above example, we saw many such examples. All the sets were generated by taking different kinds of combinations for two points and we also described the set by the set of constraints.

We see that the names of the combinations corresponds to the geometric picture that is generated by the particular kind of combination. For example, the set of all conical combinations is a cone. The set of all convex combinations is always a convex set (in particular, the convex hull of the original points). For drawing the correspondence for linear and affine combination, we define some subspaces below.

We introduce the idea of subspace now.

Definition 3. *A set $S \subseteq \mathbb{R}^n$ is called a* linear subspace *if it is closed under taking the linear combinations of pairs of its elements. More precisely, S is called a linear subspace if for any vectors $x, y \in S$ and $\lambda, \mu \in \mathbb{R}$, $\lambda x + \mu y \in S$.*

Similarly, S is called an affine subspace (or affine manifold or flat) *if it is closed under taking affine combinations of any pair of its elements.*

S is called a convex set *if it is closed under taking convex combinations of pairs of its elements and it is called a* convex cone *if it is closed under taking positive combinations of pairs of its elemtns.*

2.1 Null space and range space of $A \in \mathbb{R}^{m \times n}$

Let A be an $m \times n$ real matrix ($A \in \mathbb{R}^{m \times n}$), b be a real m-dimensional column vector and c be a real n-dimensional column vector. The *null space* of matrix A is defined to be set of all vectors that are mapped to 0 when premultiplied by A. More precisely, the null space of matrix A (denoted by $N(A)$) is

$$N(A) = \{x \in \mathbb{R}^n | Ax = 0\}$$

The *range space* of a matrix A is defined to be set of all vectors which can be obtained by premultiplying some vector by A. More formally, range space of a matrix A (denoted by $R(A)$) can be written as

$$R(A) \;=\; \{Ay \,|\, y \in \mathbb{R}^n\}.$$

Definition 4. (Hyperplane and half-space) *If* $0 \neq a \in \mathbb{R}^n$, $\beta \in \mathbb{R}$, *then the set*

$$\{x|a^\top x = \beta\}$$

is called a hyperplane*, and the set*

$$\{x|a^\top x \leq \beta\}$$

is called a (closed) halfspace*.*

Remark 5. A half space $\{x|a^\top x \leq \beta\}$ is a convex set. This can be easily verified by taking two arbitraty vectors satisfying the inequality and showing that any convex combination of them also satisfies the inequality.

Theorem 6. *A* $\begin{pmatrix} \textit{linear subspace} \\ \textit{affine subspace} \\ \textit{convex set} \\ \textit{convex cone} \end{pmatrix}$ *contains all* $\begin{pmatrix} \textit{linear combinations} \\ \textit{affine combinations} \\ \textit{convex combinations} \\ \textit{positive combinations} \end{pmatrix}$ *of any (finite) number of its vectors.*

Note that the theorem does not follow from the definitions. Linear subspace (respectively affine subspace, convex set, convex cone) is defined to contain all linear combinations (respectively affine combinations, convex combinations, positive combinations) of any two of its vectors. The property we want to prove holds for all combinations of any number of vectors, not just two of them.

Proof. The proofs for all the cases are analogous. We will prove the claim only for convex sets.

Consider an arbitrary convex combination of finitely many vectors of a convex set S. Let this be $\lambda_1 x_1 + \cdots + \lambda_k x_k$. Here $\sum \lambda_i = 1$ and $\lambda_i \geq 0$. We want to prove the $\sum_{i=1}^k \lambda_i x_i \in S$.

We will prove the claim by induction on k. If $k = 1$, the combination is the point itself and the point trivially belongs to the set. If $k = 2$, then the combination is contained in the convex set by definition.

If $k > 2$, we can write the above combination as (without loss of generality, assume that $\lambda_k < 1$, otherwise we can rename the vectors)

$$x = \sum_{i=1}^k \lambda_i x_i = (1 - \lambda_k) \underbrace{\sum_{i=1}^{k-1} \frac{\lambda_i}{1 - \lambda_k} x_i}_{\text{Call this vector } y} + \lambda_k x_k$$

By the induction hypothesis, the underbraced vector belongs to the convex set S (it is a convex combination of $k-1$ vectors). Call this vector $y\in S$. Now, $x=(1-\lambda_k)y+\lambda_k x_k$ which is written as convex combination of two vectors in S again. By definition, a convex set contains convex combinations of all of its pairs, so $x\in S$, proving the claim. □

Theorem 7. *An arbitraty interesection of* $\begin{pmatrix}\textit{linear subspaces}\\ \textit{affine subspaces}\\ \textit{convex sets}\\ \textit{convex cones}\end{pmatrix}$ *again is a* $\begin{pmatrix}\textit{linear subspace}\\ \textit{affine subspace}\\ \textit{convex set}\\ \textit{convex cone}\end{pmatrix}$.

Proof. The proofs are analogous for all four cases. We will prove it for convex cones.

Let $(C_i)_{i\in I}$ be an arbitrary collection of convex cones and $C=\bigcap_{i\in I}C_i$. If $C=\emptyset$, the claim follows because the empty set if a convex cone by definition (There are no points to take combinations). If C is a singleton set, the claim again follows. So, assume that $|C|\geq 2$.

We will prove that C contains all positive combinations of any two of its points. Let x, $y\in C$ and $\lambda x+\mu y, \lambda>0, \mu>0$ be an arbitrary positive combination. For every i, x and y are contained in C_i (because C is the intersection of all C_i's). Therefore, $\lambda x+\mu y\in C_i$ for all $i\in I$ (because, C_i is a convex cone for all $i\in I$). Hence, $\lambda x+\mu y\in\bigcap_{i\in I}C_i=C$, proving the claim. □

2.2 Two views of the same set

In the coming lectures, we will exploit the connection between two views of looking at the same subset of $\mathbb{R}^n$. These two views are the following.

Let $S\subseteq\mathbb{R}^n$. For many such sets of interests, we can find a set $T\subseteq\mathbb{R}^n$ such that every point in S can be written as $*$-combination of points in T where $*$ stands form linear, affine, convex, conical or positive. If we can find a finite T, then the set S is said to be finitely generated by T. This is one view of looking at a set. In this case, we only have to look at T to get all the information about S. (This is good because T is a finite set while S could have been a large or even infinite set.) An example: The non-negative quadrant in $\mathbb{R}^2$ can be written as all conic combinations of two points $(1,0)$ and $(0,1)$.

Another view of looking at the set is by considering the constraints which must be satisfied by all points in the set. For example, the non-negative quadrant in $\mathbb{R}^2$ can be written as the set of all points satisfying the following two constraints:

$$x_1\geq 0, x_2\geq 0.$$

Remark 8. Any closed convex set can be written as the intersection of half-spaces and also as convex combinations of some number of points. Are these two numbers (number of half-spaces and number of points) equal?

For example, consider a circle C. Call the interior and boundary of C combined as D (for disc). Any point of D can be expressed as a convex combination of points in C. D can also be written as intersection of infinitely many half-spaces, one each for a point in C (see Figure 1).

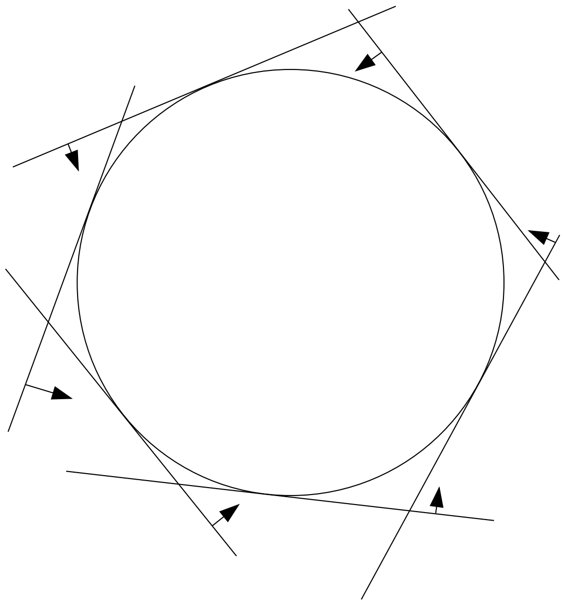

Figure 1. A circle can be written as the intersection of infinitely many half-spaces. The lines with arrows on them show the half-space.

Is the cardinality of the set of half-spaces and the set of points equal? In this case, it is equal, as both are equal to the cardinality of the real number set $\mathbb{R}$.

But in genaral, this is not the case. The easiest example is the cube in three dimensions. It requires eight points for it to be written as a convex combination of them, but six half-spaces suffice, one for each face of the cube.

In general, the number of points can be exponentially larger than the number of half-spaces or vice versa. Consider the hypercube in n dimensions. It requires 2^n points but only $2n$ half-spaces suffice. Conversely, the L_1-cube

$$\{x \in \mathbb{R}^n | \|x\|_1 \leq 1\}$$

requires only $2n$ points but 2^n half-spaces.

3 Polyhedra and polytopes

Definition 9. *A set $S \subseteq \mathbb{R}^n$ is called a* polyhedron *if it can be written as*

$$\{y \in \mathbb{R}^n | Ay \leq c\}$$

for some $A \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^m$.

Corollary 10. *A polyhedron is a convex set.*

Proof. The proof is straightforward. A polyhedron is the intersection of finitely many half-spaces, each of which is a convex set. From Theorem 7, it follows that a polyhedron is a convex set. □

Definition 11. *A set $S \subseteq \mathbb{R}^n$ is called a* polytope *if there exists a finite number of points $x_1, \ldots, x_k \in \mathbb{R}^n$ such that S is the set of all convex combinations of $x_1 \ldots, x_k$.*

Both polytopes and polyhedra are convex sets. In lower dimensions, just by visualization their geometry looks similar. Can we represent a polytope as a polyhedron? How about the other way? Is every polytope a polyhedron? Is every polyhedron a polytope? Let us consider the following two questions regarding polyhedra and polytopes.

1. When is a polyhedron a polytope?
2. When is a polytope a polyhedron?

The answer to the first question is *almost always*. If a polyhedron is bounded, then it can be written as a polytope too. But, if it is unbounded, as in Figure 2, it cannot be represented as a polytope.

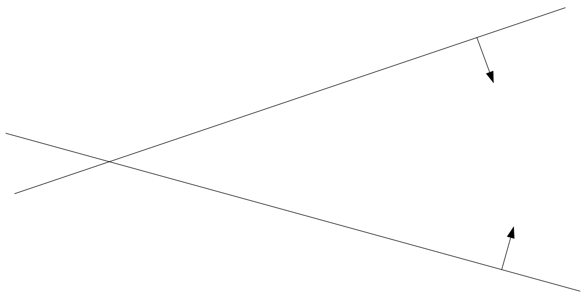

Figure 2. A polytope is also a polyhedron, but a polyhedron might not be a polytope. This is an example of a polyhedron, which cannot be expressed as a polytope.

The answer to the second question is *always*. But, in general, the number of half-spaces required to represent a polytope may be exponential in the number of points required to represent the polytope.