CS 435 : Linear Optimization	Fall 2008
Lecture 8: Extreme points	
Lecturer: *Sundar Vishwanathan*	
Computer Science & Engineering	Indian Institute of Technology, Bombay

Definition 1 *Given points* $p_1, p_2, p_3, \ldots, p_n$, *the* convex hull *is the smallest convex set containing these points.*

An equivalent definition (Why?) is *convex hull*$(p_1, \ldots, p_n)$ is the set

$$\{\Sigma_i \lambda_i p_i ; \Sigma_i \lambda_i = 1; 0 \leq \lambda_i \leq 1\}.$$

The convex hull of two points gives the line segment joining the two. The convex hull of three points gives all the points in the triangle formed by these points and so on. A view we would like to recommend is that we are taking all weighted averages of the points and including these in the set. The λ_is are the weights.

To simplify the discussion of the subject, we make the following assumptions on the nature of $Ax \leq \mathbf{b}$ which hold for the rest of the course unless otherwise stated.

1 Assumptions on the nature of the convex set $Ax \leq \mathbf{b}$

- *Assumption 1:* $Ax \leq \mathbf{b}$ is bounded. There is a real number B such that for every x satisfying $Ax \leq \mathbf{b}, x_i \leq B$.
- *Assumption 2:* $Ax \leq \mathbf{b}$ has no degeneracies. This means that not more than n hyperplanes pass through a point in an n-dimensional space.
- *Assumption 3:* $Ax \leq \mathbf{b}$ should be *full dimensional.* This means that one should be able to place a n-dimensional sphere, however small, in the region defined by $Ax \leq \mathbf{b}$. For two dimensions, the convex set should have area, and for three dimensions, the convex set should have volume. In n dimensions, the convex set should have an n-dimensional volume.

Can you think of a body in three dimensions where four planes pass through a point on the boundary? Usually, perturbing the input slightly does not change the solution by much. And such perturbation leaves the input full-dimensional and non-degenerate. So, the assumptions are not unjustified.

2 Extreme Points and its Algebraic Interpretation

If we take a polygon in the plane, or a cube in three dimensions, the vertices or corners of these objects are special. It is these *corners* that we wish to identify explicitly. Indeed, if you go over the part where we saw how $c^T x$ varied over a convex body we remarked that the maximum occured at the boundary. In fact a moment's reflection should reveal that it must occur at a corner. It is this that we will prove rigorously.

Definition 2 *An extreme point is a point in a convex set that cannot be represented as a convex combination of any two distinct points in the convex set.*

Thus, an extreme point does not lie on the segment between two other distinct points of the convex set.

THEOREM 1 *Given a convex set $Ax \leq \mathbf{b}$ in n-dimensional space satisfying the assumptions of Sec. 1, the extreme points of a convex set are precisely those points in $Ax \leq \mathbf{b}$ which can be expressed as an intersection of n linearly independent hyperplanes out of the set of hyperplanes that define the convex set $Ax \leq \mathbf{b}$.*

PROOF:
Part 1: Any point which can be expressed as an intersection of n linearly independent hyperplanes out of the set of hyperplanes that define the convex set $Ax \leq \mathbf{b}$ is an extreme point.

Let x_0 be any such point. We split $Ax_0 \leq \mathbf{b}$ into two parts

$$A^{'}x_0 = \mathbf{b}^{'} \tag{1}$$

$$A^{''}x_0 < \mathbf{b}^{''} \tag{2}$$

We group the inequalities in $Ax_0 \leq \mathbf{b}$ that are equalities into the first set $A^{'}x_0 = \mathbf{b}^{'}$ and the strict inequalities into the second set $A^{''}x_0 < \mathbf{b}^{''}$.

Since x_0 is the point of intersection of n linearly independent hyperplanes, $A^{'}$ has n linearly independent rows. Indeed, x_0 must be the only solution to $A^{'}x_0 = \mathbf{b}^{'}$.

Let there be x_1 and x_2 in the given convex set such that $x_0 = \lambda x_1 + (1-\lambda)x_2$, $0 < \lambda < 1$. For a contradiction, we wish to show that such x_1 and x_2 are the same as x_0.

Exercise: Prove that if $x_1 = x_0$ then so is x_2.

How do we complete the proof? How do we use the fact that $A^{'}$ has n linearly independent rows? It is natural to aim to show that $A^{'}x_1 = \mathbf{b}^{'}$ since this would show that $x_1 = x_0$.

To this end we see that

$$\mathbf{b}' = A'x_0 \tag{3}$$

$$= \lambda A^{'}x_1 + (1-\lambda)A^{'}x_2 \tag{4}$$

$$\leq \lambda \mathbf{b}' + (1-\lambda)\mathbf{b}' \tag{5}$$

$$= \mathbf{b}' \tag{6}$$

The third inequality holds because x_1 and x_2 are part of the convex set. This implies that the third inequality is actually an equality. This can only happen if $A^{'}x_1 = \mathbf{b}^{'}$ and $A^{'}x_2 = \mathbf{b}^{'}$. Since x_0 is the only solution to this set of equalities, we have $x_0 = x_1 = x_2$.

Thus x_0 cannot be expressed as a convex combination of two other distinct points in the convex set and is therefore an extreme point.

The sequence of inequalities is the key. Why did we do that? Here is the basic fact that leads us. Suppose it is known that the average age of the students in this class is 21. Now if you are told that no student has age greater than 21 what can you infer? Yes: that the age of every student must be 21. Let us further restrict ourselves to two students. Suppose their average height is 5 and average age is 21. And suppose it is known that no student has age greater than 21 or height greater than 5. We can still infer that both have the same age and height. Suppose further that we take a weighted average, where both weights are non-zero, it is easy to see that we will reach the same conclusion.

Consider the equality: $x_0 = \lambda x_1 + (1-\lambda)x_2$, $0 < \lambda < 1$. This says that the vector x_0 is a weighted average of x_1 and x_2. Suppose we take a linear function on these points, it is easy to see that the value of the function at x_0 will be a weighted average of the values at x_1 and x_2.

Focus on one co-ordinate; say the first. This says that the first co-ordinate of x_0 is the weighted average of the first co-ordinate of x_1 and x_2. Infact the first co-ordinate of $A'x_0$ is the weighted average of $A'x_1$ and $A'x_2$. Why? Now, can you justify the sequence of inequalities?

Part 2: We need to prove the converse. To begin, say x_0 is an extreme point. We split $Ax_0 \leq \mathbf{b}$ into two parts

$$A^{'}x_0 = \mathbf{b}^{'} \tag{7}$$

$$A^{''}x_0 < \mathbf{b}^{''} \tag{8}$$

We need to show that A' has n linearly independent rows. We immediately realise that we must aim for a contradiction. In which case, we should have started from the other direction and focussed on the contrapositive. What do we mean by this? See the next lecture for a proof.

□