Extreme Points, Corners, and Basic Feasible Solutions

CS 149 Staff

September 25, 2008

1 Extreme Points and Convexity Example

Claim 1. $\epsilon(\{5\}) = \{5\}$

Proof sketch: since there are no other points in the set, one certainly cannot express 5 as a convex combination of two other points, so 5 is extreme. We have shown that every point in the set is extreme, so there is no need to prove that no other points are extreme.

Claim 2. *A half space has no extreme points as long as the dimension is at least two. Formally, for any vector $a \neq 0$ and scalar α, $\epsilon(\{\, x |\ a^T x \geq \alpha \,\}) = \{\}$.*

Proof. we construct a nonzero y such that $a^T y = 0$, which will show that for any x in the set $a^T(x + -y) = a^T x \geq \alpha$ and therefore x is not extreme. First consider the case where a has some zero component a_i. In that case the unit vector with that component one and all others zero satisfies $a^T y = 0$. If a has all non-zero components, let $y = (a_2, -a_1, 0 \cdots)$, which clearly satisfies $a^T y = 0$. □

Claim 3. *A half space of dimension one has a single extreme point. Formally, for any $a, \alpha \in \mathbb{R}$, we have $\epsilon(\{\, x |\ ax \geq \alpha \,\}) = \{\frac{\alpha}{a}\}$.*

Proof. In one dimension a vector has simply one component. I will prove this theorem for the case $a > 0$; the case $a < 0$ is symmetric. The inequality defining the set can be rewritten in this case as $x \geq \alpha/a$. For any $x_0 > \alpha/a$, one can express x_0 as the convex combination of α/a and $2x_0 - \alpha/a$ with weight 1/2. The former is clearly in the set and the latter follows from $2x_0 \geq 2\alpha/a$. Therefore every point satisfying $x_0 > \alpha/a$ is not extreme. If $x_0 = \alpha/a$, it suffices to show that for any nonzero y either $x_0 + y$ or $x_0 - y$ is not in the set. Suppose $x_0 + y$ is in the set. That implies $y > 0$, but that in turn implies that $x_0 - y = \alpha/a - y < \alpha/a$, and therefore $x_0 - y$ is not in the set. Therefore α/a is extreme. □

In topology of real vector spaces, a point x is called an *interior point* of a set S if and only if $\exists \delta > 0 : \forall y \in S : |y - x| < \delta \rightarrow y \in S$. That is, there exists a small ball around S that is contained within the set. A set is *open* if and only if all its points are interior points.

Claim 4. *every interior point x of a set $S \subset \mathbb{R}^n$ is not an extreme point.*

Proof. let a set S and an interior point x be given. We want to show that x is not extreme. Since x is an interior point, we can choose a $\delta > 0 : \forall y \in S : |y - x| < \delta \rightarrow y \in S$. Let u be an arbitrary vector of length 1. The points $x + u\delta/2$ and $x - u\delta/2$ show that x is a convex combination of points in the set. □

Corollary: every open subset of $\mathbb{R}^n$ has no extreme points.

2 Polyeders and Corners

A polyeder/polyhedron is a finite intersection of half spaces. A bounded polyeder is called a polytope.

It is useful to have many different equivalent characterizations of the same concept, since different versions are easier to use in different contexts. We therefore will define what will turn out to be a synonym for extreme point (at least on polyeders): a corner. Intuitively the corners of an object are the parts that one could put ink on to make a writing implement. More precisely, a point x_0 in a polyeder P is a corner iff for some nonzero vector a and real constant α, $a^T x_0 = \alpha$ and $\forall y \in P - \{x_0\} : a^T x_0 > \alpha$.

3 Proof that corners are extreme

If $\mathbf{P}$ is a convex polyeder, the following two facts are true:

- Every corner on $\mathbf{P}$ is an extreme point
- If $\mathbf{P}$ is bounded than $\mathbf{P}$ is a convex hull of its corners

We will now prove the first remark:

1. Let x^0 be a corner of $\mathbf{P}$.
2. We know there exists an $H = \{x \,|\, a^T x = \alpha\}$ such that $\mathbf{P} \subseteq H^{\leq}$ and $\mathbf{P} \cap H = \{x^0\}$ by the definition of a corner.
3. We need to show $\forall y \in \mathbb{R}^n : x^0 + y, x^0 - y \in \mathbf{P}$ implies that $y = 0$.
4. Let $y \in \mathbb{R}^n$ such that $x^0 + y, x^0 - y \in \mathbf{P}$ be given.
5. Because $\mathbf{P} \subseteq H^{\geq}$, we have $x^0 + y, x^0 - y \in H^{\geq}$, which is equivalent to $a^T(x^0 \pm y) \geq \alpha = a^T x^0$. So, $a^T y \geq 0$ and $a^T y \leq 0$, so $a^T y = 0$.
6. So, $a^T(x^0 + y) = a^T x^0 + a^T y = a^T x^0 = \alpha$, therefore $x^0 + y \in \mathbf{P} \cap H = \{x^0\}$ and $y = 0$.
7. By definition, it follows that x^0 is an extreme point.

4 Basic Feasible Solutions

Our definitions of extreme points nor corners are so far of limited utility because we do not have a systematic way to look for them. Computers cannot reason efficiently using pictures, so it would be helpful to have a definition of a corner that is based on linear algebra rather than geometry. The concept of a basic feasible solution is the answer.

If you do Gauss Jordan elimination on a matrix and then set all variables that aren't part of the identity sub matrix equal to zero, you can read out values for the remaining variables. This is called a basic solution. If all the variables are nonnegative, it is a basic feasible solution. Basic feasible solutions correspond to corners. To get a different basic feasible solution, simply choose a different set of columns for the identity matrix.

Formally, fix a matrix $A \in \mathbb{R}^{mxn}$ with rank $m \leq n$ and a vector $b \in \mathbb{R}^m$. A function B from 1..m to 1..n indicates which columns to find the identity matrix in. In particular $B(i) = j$ means that the jth column is intended to be all zeros except for a one in row i. The function N from 1..m-n to 1..n indicates which columns are not used in the identity matrix. Each column should be used exactly once, so assume that the union of the ranges of N and M is 1..n. Let A_B denote the sub matrix of A corresponding to the columns indicated by B, that is $A_B = (a^{B(1)} \cdots a^{B(m)}$ where a^t denotes the tth column of A. In order for row operations to produce the identity matrix it is necessary and sufficient that A_B be invertible.

Given a pair of functions satisfying these conditions one can construct a basic solution as follows. Set $x_N = 0$ and $x_B = A_B^{-1}b$, where x_B and x_N are defined analogously to A_B. Let x be the vector with component x_i equal to the appropriate component of x_B or x_N depending on which the corresponding column belongs to. Such a vector is a basic solution. If it is nonnegative it is feasible and is therefore called a basic feasible solution. Note that pre-multiplying by the matrix A_B^{-1} is equivalent to doing the row operations that transform A_B into the identity matrix.

As is proved below, basic feasible solutions are equivalent to corners and extreme point of polyeders.

For example, consider the optimization problem with constraints:[1]

$$\begin{pmatrix} 1 & 7 & 22 & 0 & 5 & 0 & 1 \\ 0 & 3 & -5 & 0 & 2 & 1 & 2 \\ 0 & 1 & 0 & 1 & -4 & 0 & 3 \end{pmatrix} x = \begin{pmatrix} 7 \\ 1 \\ 2 \end{pmatrix}$$

The easiest basis to work with is the one in which A_B is the identity matrix. This corresponds to $B(1) = 1, B(2) = 6, B(3) = 4$ and $N(1) = 2, N(2) = 3, N(3) = 5, N(4) = 7$. Note that the order of the columns in B matters. To ensure $A_B = I$ the column $a^{B(i)}$ should have a 1 in component i and zeros elsewhere. The corresponding basic feasible solution is $x^T = (7, 0, 0, 2, 0, 1, 0)$.

[1] This example is based on the example given in class but is not identical.

5 Proof that corners, basic feasible solutions and extreme points are the same

Given $x^0 \in \mathbf{P} = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$, it is equivalent to say:

1. x^0 is an extreme point of $\mathbf{P}$.
2. $\{a^j \,|\, x_j^0 > 0\}$ are linearly independent, $[A = (a^1, ..., a^m)]$.
3. x^0 is a basic feasible soluition (bfs).
4. x^0 is a corner of $\mathbf{P}$.

5.1 Statement 4 → Statement 1

We wish to use the previously proven result that every corner of a convex polyeder is an extreme point. So, we have to show that $\mathbf{P}$ is a convex polyeder:

$$\mathbf{P} = \{x \in \mathbb{R}^n \,|\, x \geq 0, Ax = b\} = \{x \in \mathbb{R}^n \,|\, \begin{bmatrix} A \\ -A \\ I \end{bmatrix} x \geq \begin{bmatrix} b \\ -b \\ 0 \end{bmatrix}\}$$

5.2 Statement 1 → Statement 2

Let us say that the first r components of x^0 are non-negative, and $x_{r+1}^0, \ldots, x_n^0 = 0$. For $r = 0$, $x^0 = 0$, which implies $\{a^j \,|\, x_j^0 > 0\} = \emptyset$. For $r > 0$, assume $\{a^j \,|\, x_j^0 > 0\}$ is linearly dependent. This implies that we can find a linear combination of these vectors with coefficients $\alpha_1, \ldots, \alpha_r$.

Without loss of generality, assume $\forall i, 1 \leq i \leq r, |\alpha_i| < |x_i^0|$, and consider $y^t = (\alpha_1, \ldots, \alpha_r, 0, \ldots, 0)$. This implies $Ay = 0$, and $A(x^0 \pm y) = b$ and $x^0 \pm y \geq 0$ follows from the statement $|\alpha_i| < |x_i^0|$ for all i.

So, $x^0 \pm y \in \mathbf{P}$, which implies $y = 0$ and thus $\alpha_i = 0$, $\forall_i$.

5.3 Statement 2 → Statement 3

Assume we have our linearly independent set $\{a^j \,|\, x_j^0 > 0\}$. Now, we know there exists a basis B (with linearly independent columns) such that $B(N_m) \supseteq \{j \,|\, x_j^0 > 0\}$.

Then we have $A_B^{-1} b = A_B^{-1} A x^0 = A_B^{-1} A_B x_B^0 + A_B^{-1} A_N x_N^0$. We also know $A_N x_N^0 = 0$, so $A_B^{-1} b = A_B^{-1} A_B x_B^0 = x_B^0 \to x^0$ is bfs.

5.4 Statement 3 → Statement 4

Let x^0 be a bfs. $x^0 = \begin{bmatrix} x_B^0 \\ x_N^0 \end{bmatrix}$. Let $a = \begin{bmatrix} a_B \\ a_N \end{bmatrix}$ with $a_B = 0$, and $a_N^t = (1, \ldots, 1)$. Let us also introduce $H = \{x \in \mathbb{R}^n \,|\, a^t x = 0\}$.

a) Now we choose any $x \in \mathbf{P}$. We have $a^t x = a_B^t x_B + a_N^t x_N = 1_{n-m} x_n \geq 0$, so $\mathbf{P} \subseteq H^{\geq}$.

b) By definition, we have $a^t x^0 = a_B^t x_B^0 + a_N^T x_N^0 = 0$, so $x^0 \in H \cap \mathbf{P}$.

c) Finally, we need to show $\mathbf{P} \cap H \subseteq \{x^0\}$. Choose $y \in \mathbf{P} \cap H$. $a^t y = 0$, or $0 = a_B^t y_B + a_N^t y_N = 1_{n-m} y_n$, so $y_n = 0$. Furthermore, $b = Ay = A_B y_B$, so $y_B = A_B^{-1} b = x^0$, so $y = x^0$.