Lecture 4: Convex sets

Rajat Mittal *
IIT Kanpur

Last time we looked at the definition of convex sets. Today lets look at some of the examples of convex sets which will be useful later.

1 Hyperplane and Halfspaces

We extend the idea of a line to hyperplanes and halfspaces. A hyperplane $H$ is described by a vector $a \in \mathbb{R}^n$ and a number $b \in \mathbb{R}$

$$H = \{ x : a^T x - b = 0 \}$$

Exercise 1. Prove that a hyperplane is affine and so convex too.

Suppose we know a point $x_0$ on the hyperplane $H$. Then this equation can be changed to $a^T x = b \Rightarrow a^T x = a^T x_0$, $x_0 \in H$. We can view the hyperplane as the set of all points which have same inner product $b$ with the vector $a$. This gives a very nice geometrical picture of the hyperplane, i.e., all points in $H$ can be expressed as the sum of $x_0$ and a vector orthogonal to $a$ (we can call it $a^\perp$). So, another definition of hyperplane is

$$H = \{ x : x = x_0 + a^\perp, a^T a^\perp = 0 \}$$

Note that this definition assumes that we know a point on the hyperplane. But this point is not special in any way, for any point on the hyperplane we can define the hyperplane in the same way. The vector $a$ is called the normal vector of the hyperplane and $b$ is called the offset. Another way to think of this hyperplane is to take the set of all vectors orthogonal to $a$ (hyperplane, passing through origin) and offset them by the distance $\frac{b}{\|a\|}$.

We studied the generalization of a line in higher dimensional space, it was called a hyperplane. A hyperplane divides the space into two parts, $a^T x \geq b$ and $a^T x \leq b$. Geometrically, they are the two sides of the plane. Is a halfspace affine/convex?

2 Polytopes and Polygons

We will mostly be interested in linear equations and linear inequalities as our constraints. We already discussed the set of points which satisfy the set of all constraints is called the feasible region. When this constraints are linear inequalities and linear equalities the feasible region is called a polytope. Mathematically, $S$ is a polytope iff

$$S = \{ x : a_i^T x - b_i \leq 0 \ i = 1, 2, \cdots, m \text{ and } c_i^T x - d_i = 0 \ i = 1, 2, \cdots, p \}$$

A bounded polytope in two dimensions is called a polygon.

Exercise 2. Prove that the polytope is a convex set.

Exercise 3. What are the polytopes in one dimension?

* Thanks to books from Boyd and Vandenberghe, Dantzig and Thapa, Papadimitriou and Steiglitz
Geometrically, polytopes are intersections of hyperplanes and halfspaces. Remember that a polytope could be bounded or unbounded. Intuitively, polytopes have vertices, edges, planes and hyperplanes as their bounding surface. A bounded polytope can be thought of as the convex hull of its vertices. Actually a bounded polytope can have an alternate definition as the convex hull of a finite set of points. The statement that these two definitions (the previous one and the original definition in terms of equalities and inequalities) are the same is known as Minkowski-Weyl theorem. The proof of that is out of the scope of this course. Now we look at a special case of polytope.

We haven’t defined vertices/extremal points formally till now. It is intuitively clear that a vertex is a corner of the polytope. Formally, A vertex of a polytope is the point which cannot be expressed as the convex combination of two different points in the polytope. This implies that vertex is not inside of any line segment joining two points in the convex set.

We saw that the convex hull of a triangle and a point inside the triangle is triangle itself. Suppose we are given a convex set and we want to find out the minimal set whose convex combinations will generate the entire set. By the definition of vertices, all vertices should be in this minimal set. It turns out that the for a bounded polytope this set (set of all vertices) is enough.

Exercise 4. Suppose $S = Conv(x_1, \ldots, x_k)$. Prove that $x_i$ is not extremal/vertex if and only if it can be written as the convex combination of other $x_j$’s.

3 Cones

We have seen sets made by vertices, lines, line segments. Now we look at sets generated by rays. A set is called a cone iff every ray from origin to any element of the set is contained in the set. Hence the set $C$ is a cone iff for every $x \in C$ we have $\theta x \in C, \theta \geq 0$.

NOTE: A cone is not a set which has all possible conic combinations of all its points (Show an example). Remember the notion of conic combination. A conic combination of vectors $x_1, \ldots, x_k \in \mathbb{R}^n$ is any vector of the form $\theta_1 x_1 + \cdots + \theta_k x_k$ for $\theta_1, \ldots, \theta_k \geq 0$.

The previous paragraph implies that a cone is not necessarily convex (Give example of a cone which is not convex). A set which is a cone and is convex is called a convex cone. In this course we will mostly be concerned with convex cones. Mathematically, a convex cone $C$ is a set where for all $\forall x_1, x_2 \in C$ and $\theta_1, \theta_2 \geq 0$, $\theta_1 x_1 + \theta_2 x_2 \in C$. So we get the definition, A cone is convex iff it contains all the conic combinations of its elements.

Convex hulls and cones are closely related.
Exercise 5. Take $x_i$'s as row vectors. Prove,

$$x \in \text{Conv}(x_1,x_2,\cdots,x_k) \Leftrightarrow (x,1) \in \text{Cone}((x_1,1),\cdots,(x_k,1))$$

Note: Some authors define cones as sets closed under positive scalar multiplication. We have defined cones as sets closed under non-negative scalar multiplication.

Clearly the set $\text{Cone}(x_1,x_2,\cdots,x_k) := \{\theta_1 x_1 + \cdots + \theta_k x_k : \forall i \theta_i \geq 0, \ x_i \in \mathbb{R}^n\}$ is a cone. It is called a finitely generated cone because it is generated by finite number of vectors. A convex finitely generated cone is also a polytope. Next theorem gives a characterization of a finitely generated cone.

**Theorem 1. Weyl's Theorem**: A non-empty finitely generated convex cone is a polytope.

**Proof.** Suppose the set of generators for cone $C$ are $x_1,\cdots,x_k$. We can define a matrix $X$ which has $x_i$'s as the columns. Then the cone $C$ can be written as

$$C = \{x : x = X\theta, \ \theta \in \mathbb{R}_+^k\}$$

Converting equalities into inequalities

$$C = \{x : x - X\theta \leq 0, \ X\theta - x \leq 0, \ -\theta \leq 0\}$$

Now $\theta$ can be eliminated from these inequalities using something known as Fourier-Motzkin elimination.

**Lemma 1.** Let $Ax \leq b$ be a system of $m$ inequalities in $n$ variables. This system can be converted into another equivalent system $A'x \leq b'$ with $n - 1$ variables and polynomial in $m$ many inequalities. Here equivalent means any solution $x$ of old system will be a solution of the new system ignoring the removed variable. Also given any solution $x$ of new system ($A'x \leq b'$), we can find a solution $(x_0,x)$ of old system.

**Proof.** Suppose the variable to be removed is $x_0$. We divide all the inequalities into three sets depending upon whether the coefficient of $x_0$ is positive ($P$), negative ($N$) or zero ($Z$). Divide the inequalities in $P$ and $N$ by the modulus of the coefficient of $x_0$. The inequalities in the new system are the inequalities from $Z$ and every inequality of the form $p_i + n_j \leq 0, p_i \in P, n_j \in N$.

**Exercise 6.** Prove that this construction works.

With the $\theta$ eliminated from the system of equations which define the cone, we get

$$C = \{x : Ax \leq 0\}$$

Hence it is a polytope.

**NOTE:** A general polytope is $Ax \leq b$ and we will see that a finitely generated cone is $Ax \leq 0$.

4 Characterization of polytope

We saw last week that cones and bounded polytope has two representations. One in terms of linear inequalities and one in terms of convex combinations (conic combination). There is a similar theorem for polytope.

**Theorem 2.** Let there be a polytope defined by a set of inequalities, $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. Then there exist vectors $x_1,\cdots,x_k \in \mathbb{R}^n$ and $y_1,\cdots,y_l \in \mathbb{R}^n$, s.t.,

$$P = \text{Cone}(x_1,\cdots,x_k) + \text{Conv}(y_1,\cdots,y_l)$$
NOTE: The sum defined here is Minkowski’s sum. It will be defined in the next class. For intuition $S_1 + S_2 = x_1 + x_2 : x_1 \in S_1, x_2 \in S_2$.

This is known as the Affine Minkowski-Weyl theorem. We will not do the proof of this theorem in this course. Notice that now there are equivalent characterizations of cone/polytope/bounded polytope in terms of convex/conic hulls or linear inequalities. It is instructive to remember the special forms of linear inequalities and hulls required to make these shapes.

<table>
<thead>
<tr>
<th>Linear inequalities</th>
<th>Convexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bounded polytope</td>
<td>Bounded and $Ax \leq b$</td>
</tr>
<tr>
<td>Finitely generated Cone</td>
<td>$Ax \leq 0$</td>
</tr>
<tr>
<td>Polytope</td>
<td>$Ax \leq b$</td>
</tr>
</tbody>
</table>

We defined these sets polytopes, cones etc.. They are interesting because the feasible regions of our optimization problems will be intersections of these various kind of convex sets.