AI2100/EE5609 Convex Optimization Theory

Lecture 4: Polyhedra, Norm Balls, Norm Cones

Scribe(s): Athira Krishnan R, Tadipatri Uday Kiran Reddy

March 27, 2022

CRISP SUMMARY

- Convex sets are set which have clear line of sight between any two points on the set. Line segment joining any two points in the set completely lies in the set.

- A set $\mathbb{C} \epsilon \mathbb{R}^n$ is affine if the line through 2 distinct points in C lies in $\mathbb{C}$.

- For convex problems local minima and global minima coincides.

- Intersection of convex sets are convex.

- If a set is convex, so are its affine combination, perspective and inverse perspective transformations.

Hyperplane

[1]A hyperplane is the set of the form, $\{x|a^\top x = b\}$, where a ϵ $\mathbb{R}^n$, $a \neq 0$ and b ϵ $\mathbb{R}$. Geometrically it can be interpreted as the set of points with a constant inner product to a given vector $\underline{a}$ or as a hyperplane with normal vector a and constant b determines the offset of the hyperplane from the origin. This can be expressed in the form, $\{x|a^\top(x - x0) = 0\}$ where xo is any point in the hyperplane.

Halfspaces

A hyperplane divides $\mathbb{R}^n$ space into 2 half spaces. A half space is a set of the form $\{x|a^\top x \geq b\}$, where $a \neq 0$. Halfspaces are convex but not affine.

Half spaces can also be represented as $\{x|a^\top(x - x0) \leq 0\}$, where xo is any point on the associated hyperplane.

Polyhedron

A polyhedron is defined as the solution set of a finite number of linear equalities and inequalities. Or a polyhedra is an intersection of halfspaces and hyperplanes. Affine sets, rays, line segments half

For example, {x:x1+3x2+x3=5} is a plane in $\mathbb{R}^3$. This can be written in the form $\{x : a\underline{X} = b\}$. This being a linear combination of x is therefore affine combination and convex.
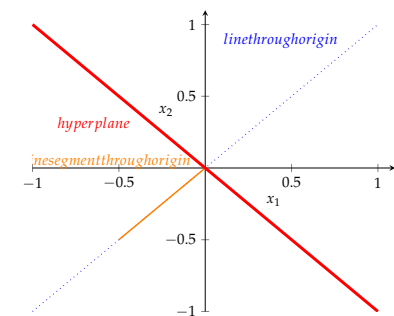
[1]

Figure 1: Hyperplane shown in red line separates the $\mathbb{R}^2$ plane into two. For a point in hyperplane set, say origin, line segment between 2 points lies inside the halfplane but the line through them may not. So by definition, hyperplane is convex but not affine.

spaces are few examples of polyhedra. Since halfspaces and hyperplanes are convex, so is polyhedra. A polyhedra need not be bounded. A bounded polyhedra is also reffered as "polytope". A polyhedra can be represented in the form, $\mathbb{P} = \{x | \mathcal{A}x \leq b, \mathcal{C}x = d\}$.

Polytope can be seen also as a convex hull formed by its vertices. Therefore a polytope can be represented in terms of edges as $\{x | \mathcal{A}\underline{x} \leq b\}$ or of vertices as $\{x | \sum_{i=1}^{N} \theta_i \mathbf{v}_i\}$. In computing the polytope using edge representation, in nD space number of planes required will be 2n, whereas in case of vertex based method, 2^n vertices would be required. Therefore as number of vertices increases, more would be the computation.
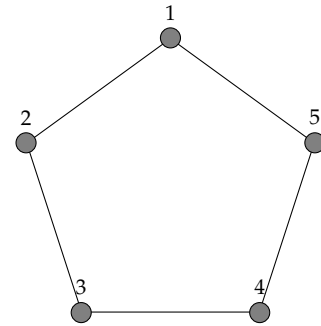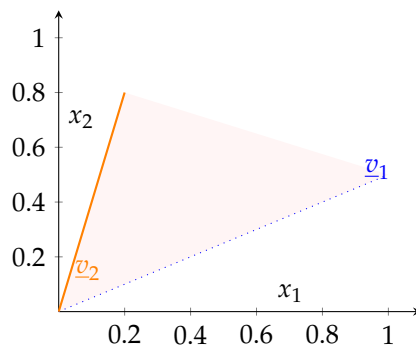

Figure 2: Polytope formed by intersection of 5 halfspaces or 5 vertices.

Simplex

It is a polyhedra formed by a certain number of non degenerates.[2] Consider k+1 points $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k \epsilon \mathbb{R}^n$ are affinely independent. Or in simple terms, $\underline{v}_1 - \underline{v}_0, \ldots, \underline{v}_k - \underline{v}_0$ are linearly independent. Now taking convex hull of all these points results in a simplex represented in the form, $\mathcal{C} = conv\{\underline{v}_0, \ldots, \underline{v}_k\} = \{\theta_0 \underline{v}_0 + \ldots + \theta_k \underline{v}_k \theta \geq 0, 1^\top \theta = 1\}$

[2] Minimum number of points in 3D to construct a polyhedra is 4, whereas its 3 in 2D.

Figure 3: Polytope using 2 linearly independent vectors.

Consider the set, $\{xx \geq 0, 1^\top x = 1\}$ (simplex polyhedra). Ensure they sum to 1 and is non negative. Let the probability vector be $\{x : x1, x2 \geq 0 x1 + x2 = 1\}$. $\underline{v}_0 = 0, \ldots \underline{v}_i = ei$. Simplex would be as shown below. Consider bounded polyhedron or a polytope P. Consider the optimization problem,

$min(a^\top x)$

subject to $\underline{x} \epsilon P$ This is a convex problem as its constraints set is convex and objective function $a^\top x$ being a linear combination. Being convex, the local minima and global minima of the problem coincides. So optima falls to one of the vertices of the polyhedra. If $a^\top v1 = -1$, $a^\top v2 = -1$ the any point on the line segment joining v1 and v2, $a^\top(\theta v1 + (1 - \theta)v2 = -1$.

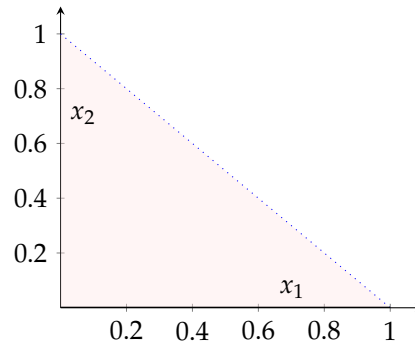*Proof.* Taking the convex combination of all n vertices $v_1, v_2 \ldots v_n$, is given by,

$a^\top(\sum_{i=1}^{N}\theta_i v_i), 0 \leq \theta \leq 1$ and $\sum_{i=1}^{N}\theta_i = 1$ As weighted average >smallest element in the set,

$a^\top(\sum_{i=1}^{N}\theta_i v_i) \geq min_{i1}^{N}\theta_i v_i$ Then value in the function would be atleast $a^\top v_i$ or value at the vertices. Therefore one of the vertices is an optima. □

Norm Ball

In mathematics, a norm is the length (or size) of a vector from origin. A function to be called as norm, it need to have following properties,

1. $||X| \geq 0$, norm of any vector is non negative. $||X|| = 0$, iff $X = 0$.

2. $||\alpha X|| = \alpha||X||$, norm follows property of homogeniety.

3. $||X + Y|| \leq ||X|| + ||Y||$, nonrm follows triangle inequality.

A Lp-norm is a norm on a finite-dimensional space of dimension N defined as, $||u||_p = (\sum_{i=1}^{N}|u_i|^p)^{1/p}$, where $p \geq 1$. This set can be represented as $\{x|||X|| \leq r\}$, where r is a constant. For p<1, convexity is not valid. The norm for p=1 is the sum of the distances a 3D computer numerical control (CNC) mill would have to move in the x,y,z directions. In 3D L1 unit norm ball takes shape of an octahedron, whereas in 2D it forms a square shape.It can be defined as 1 norm: $\{x||x1| + x2| \leq r\}$. When p=2, norm deduces to eucledian distance (Line of sight distance), now unit norm ball takes a circle shape. It can be represented as 2 norm : $\{x|x1^2 + x2^2 \leq r^2\}$. When p=∞, norm is computed as $||u_{inf=max\{u_i\}}$.

Figure 5: Minima of polyhedra occurs at one of its vertex

Theorem 1. *Norm balls are convex.*

Proof. Consider 2 points in set X, such that $||X1|| \leq rand||X2|| \leq r$.Thenbytriangularinequality,$||\theta x1 + (1 − \theta)x2|| \leq ||\theta x1|| + ||\theta x2||$ Since θ is a scalar,by scalar multiplication, $||\theta x1 + (1 − \theta)x2|| \leq \theta||x1|| + \theta||x2||$ By definition, $||\theta x1 + (1 − \theta)x2|| \leq \theta r + \theta r \leq r$ Therefore norm balls are also convex. □
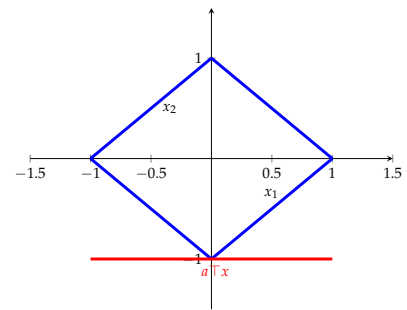
Theorem 2. *Given transformation matrix P is positive semi definite, Norm balls preserves convexity during stretching/compressing the set.*

Proof. Consider a convex set $\{x|x^t op.x \leq 1\}$ representing 2 norm. $x|vec x^\top P^{-1}\underline{x}$, assuming P^{-1} is symmetric. By spectral decomposition theorem, $P = VDV^\top$. By taking inverse on both sides, $P^{-1} = VD^{-1}V^\top$, then set x can be rewritten as, $\{x|x^\top VD^{-1}V^\top x \leq 1\}$. Taking $V^\top x$ as y, problem reduces to form, $\{y|y^\top D^{-1}y \leq 1\}$. $D =$

$$\begin{bmatrix} \lambda 1 & 0 & ... & 0 \\ 0 & \lambda 2 & ... & 0 \\ . & . & ... & . \\ . & . & ... & . \\ 0 & 0 & ... & \lambda n \end{bmatrix}$$ Assuming all $\lambda_i > 0$, solution deduces to form,

$\{y|\sum_{i=1}^{N} y_i^2/\lambda_i^2\}$. This represents an ellipsoid in 2D of the form, $y1^2/a^2 + y2^2/b^2 \leq 1$. Therefore given that P is positive semi definite, even after stretching or compressing the set remains convex.

$\square$

Norm Cones

Set $\mathbb{C}$ is a cone if for any $x\epsilon\mathbb{C}$, $\theta x\epsilon\mathbb{C}$ for any $\theta \geq 0$. θx represents the line joining x and origin or in other words, it represents ray pointing in the direction of x from origin. For any point in C the ray lies in C. Cones are convex sets. In $\mathbb{R}^2$, cones are represented as $z = \sqrt{x^2 + y^2}=$ L2 norm of first 2 coordinates. As any point on the cone is given by $z = \sqrt{x^2 + y^2}$, any point inside the cone can be represented as $z \geq \sqrt{x^2 + y^2}$. Therefore in $\mathbb{R}^n$, norm of top n-1 coordinate $\leq n^{th}$ coordinate. In $\mathbb{R}^{n+1}$ a cone is $\{(x,t)|x\epsilon\mathbb{R}^n, t\epsilon\mathbb{R}, ||x|| \leq t\}$. 2 norm cones are generally referred as Lorentz cones. All norm cones are convex.

Proof. Assume a set $\{(x,t)|x\epsilon\mathbb{R}^n, t\epsilon\mathbb{R}, ||x|| \leq t\}$, then $||x|| \leq t, this means, ||x||/t \leq 1$. This represents the perspective transform of (x,t). This is of the form $||X|| \leq 1$. As norm is convex, so is norm cones.

   Question.

   Consider a set in $\mathbb{R}^n$ given by $\{x : x^\top P^{-1}x \leq (\mathbb{C}^\top x)^2, \mathbb{C}^\top x > 0\}$. Assume P is positive definite, S_{++}^n. Let a new variable $t = \mathbb{C}.^\top x$, set becomes $\{x : x^\top P^{-1}x \leq t^2, t > 0, t = \mathbb{C}.^\top x\}$

   Taking square root on both sides, $\{x : \sqrt{x^\top P^{-1}x} \leq t, t > 0, t = \mathbb{C}.^\top x\}$

   When last coordinate is positive and $t = \mathbb{C}.^\top x$, this becomes equivalent to linear combination of first n values, this represents a hyperplane.
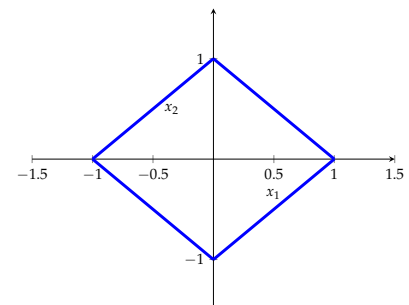
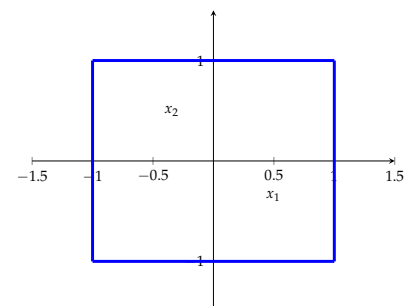Sketch 1 norm and infinite norm balls

Figure 6: 1 norm ball.

Figure 7: infinite norm ball.

If $\sqrt{x^\top P^{-1} x} \le t$ is convex (being a hyperplane), rest also becomes convex. Assume $P = VDV^\top$ is symmetric positive definite matrix, then $x^t op V D^{1/2} D^{1/2} V^t op x$.

If $y = D^{1/2} V^t op x$, this reduces to $\sqrt{y^\top y} \le t$, which means $||y|| \le t$, represents a cone.

Therefore a norm cone is convex it is intersection of hyperplane and cone. $\qquad\square$

Separating hyperplane theorem

Suppose C_1 and C_2 are nonempty disjoint convex sets, i.e, $C_1 \cap C_2 \in \Phi$. Then there exists $\bar{a} \ne 0$ and b such that $\bar{a}^T \bar{x} \le b; \forall \bar{x} \in C_1$ and $\bar{a}^T \bar{x} \ge b; \forall \bar{x} \in C_2$. The hyperplane $\{\bar{x}|\bar{a}^T \bar{x} = b\}$ is called *separating hyperplane* for the sets C_1 and C_2.

Proof. Take $\bar{a} \in C_1$ and $\bar{b} \in C_2$ such that these two vectors form a pair that belongs to the set $inf\{||\bar{u} - \bar{v}||_2; \bar{u} \in C_1, \bar{v} \in C_2\}$

Then we have to show that affine function $f(\bar{x}) = \bar{c}^T \bar{x} - d$ is the separating hyperplane such that $\bar{c} = \bar{a} - \bar{b}$ and $d = \frac{||a||_2^2 - ||b||_2^2}{2}$.

Now since f is non-negative on C_1 and non-positive on C_2, Suppose there is point $\bar{u} \in C_1$ such that $f(\bar{u}) < 0$.

$$f(\bar{u}) = (\bar{a} - \bar{b})^T \bar{u} - \frac{||a||_2^2 - ||b||_2^2}{2} = (\bar{a} - \bar{b})^T (\bar{u} - \bar{b}) + 0.5||\bar{d} - \bar{c}||_2^2$$

$$(\bar{a} - \bar{b})^T (\bar{u} - \bar{a}) < 0$$

$$(\bar{b} - \bar{a})^T (\bar{u} - \bar{a}) > 0$$

The above inequality means that some point in C_1 makes acute angle with line joining closed of two sets which possible only when $\bar{u}$ is more closer to $\bar{b}$ than $\bar{a}$ is. As $\bar{a}$ itself is closest thus, such a $\bar{u}$ cannot exists. $\qquad\square$

Property: Every (closed) convex set has a unique point with smallest norm.

Supporting hyperplane theorem

Take a point on the boundary of a convex set, drawing tangent through it will make the whole convex set lie on only one side of the line.

Farkas Lemma

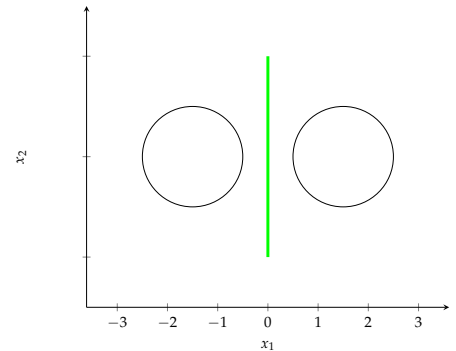Consider the below two statements,

Figure 8: Seperating hyperplane

1. We can find $\bar{x} \geq \bar{0}$ such that $\mathbf{A}\bar{x} = \bar{b}$.

2. We cam find $\bar{a}$ such that $\bar{a}^T\bar{b} \geq 0$ ans $\mathbf{A}^T\bar{a} \leq \bar{0}$

Farkas lemma states that exactly only one of the above statements is true.

Proof. Let $\bar{x} \geq \bar{0}$, s.t $\mathbf{A}\bar{x} = \bar{b}$ then,

$$\bar{a}^T\mathbf{A}\bar{x} = \bar{a}^T\bar{b}$$
$$\implies (\mathbf{A}^T\bar{a})^T\bar{x} = \bar{a}^T\bar{b}$$

Assume Statement 1) is true.
   Let $\bar{a}^T\bar{b} > 0$, then not all entries of $(\mathbf{A}^T\bar{a})^T$ are negative which means that statement 2) is false.

   Assume Statement 2) is true.
   Then we cannot have all entries of $\bar{x}$ to be positive some of them must be negative to satisfy the above equation.
   Thus Both Statement 1) and 2) cannot be true together.
   We can also show this result using separating hyperplane theorem take convex set $C_1 = \{\mathbf{A}\bar{x} : \bar{x} \geq 0\}$, this set is cone. And another set $C_2 = \{\bar{b}\}$. Now we can always find a hyper plane which separates these two sets, note that set C_1 is infinite set thus plane can never form a acute angle. Thus plane would be oriented towards $\bar{b}$. □

   Problem
   Consider the following optimisation problem,
   min $\bar{c}^T\bar{x}$
   such that $\mathbf{A}\bar{x} = \bar{b}$

   Solution
   Let $\bar{r}$ satisfy $\mathbf{A}\bar{r} = \bar{b}$, then general solution of the above equation is $\bar{r} + \overline{x_0}$ where $\overline{x_0}$ is in null space of $\mathbf{A}$. This means that solutions to the equation is just null space shifted by $\bar{r}$. If $\bar{c}$ is perpendicular to null space of then $\bar{c}^T\bar{x} = \bar{c}^T\bar{r} + \overline{x_0} = \bar{c}^T\bar{r}$. If $\bar{c}$ is not perpendicular to null space then value is $-\infty$.

Linear programming problem

min $\bar{c}^T\bar{x}$
   such that $\mathbf{A}\bar{x} = \bar{b}$ and $\bar{x} \geq 0$

   Any linear programming problem can be boiled down to above optimisation problem. As a consequence of farkas lemma we can show that there can exists a solution to above problem otherwise we

can give a certificate that above problem has no such $\bar{x}$ which satisfies the above problem.

Ordering of Vectors

If $\bar{x} < \bar{y}$ then $\bar{y} - \bar{x} \in \mathcal{K}$, where $\mathcal{K}$ is a cone. Ordering changes with the choice of the cone. We usually consider first quadrant of $\mathbb{R}^2$ as our choice of cone.

We can always find a pair of vectors where both $\bar{x} - \bar{y}$ and $\bar{y} - \bar{x} \notin \mathcal{K}$, such vectors are not comparable and every choice of cone can such pairs.