

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

¹A hyperplane is the set of the form, $\{x | a^\top x = b\}$, where $a \in \mathbb{R}^n$, $a \neq 0$ and $b \in \mathbb{R}$. Geometrically it can be interpreted as the set of points with a constant inner product to a given vector 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, $\{x | a^\top (x - x_0) = 0\}$ where x_0 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 $\{x | a^\top (x - x_0) \leq 0\}$, where x_0 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, $\{x : x_1 + 3x_2 + x_3 = 5\}$ is a plane in \mathbb{R}^3 . This can be written in the form $\{x : a^\top x = b\}$. This being a linear combination of x is therefore affine combination and convex.

¹

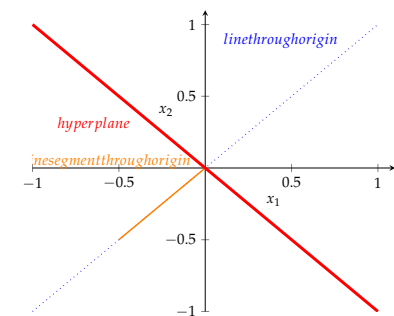


Figure 1: Hyperplane shown in red line separates the \mathbb{R}^2 plane into two. For a point in hyperplane set, say origin, line segment between 2 points lies inside the halfplane but the line through them may not. So by definition, hyperplane is convex but not affine.

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 referred as "polytope". A polyhedra can be represented in the form, $P = \{x | Ax \leq b, Cx = d\}$.

Polytope can be seen also as a convex hull formed by its vertices. Therefore a polytope can be represented in terms of edges as $\{x | Ax \leq b\}$ or of vertices as $\{x | \sum_{i=1}^N \theta_i v_i\}$. 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.

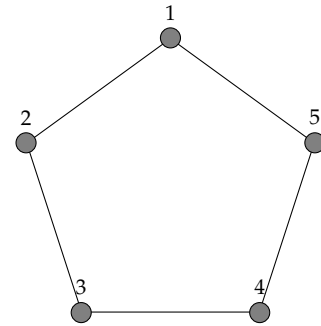
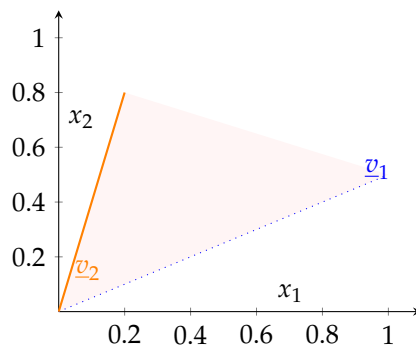


Figure 2: Polytope formed by intersection of 5 halfspaces or 5 vertices.

Simplex

It is a polyhedra formed by a certain number of non degenerates.² Consider $k+1$ points $v_1, v_2, \dots, v_k \in \mathbb{R}^n$ are affinely independent. Or in simple terms, $v_1 - v_0, \dots, v_k - v_0$ are linearly independent. Now taking convex hull of all these points results in a simplex represented in the form, $C = \text{conv}\{v_0, \dots, v_k\} = \{\theta_0 v_0 + \dots + \theta_k v_k \mid \theta_i \geq 0, \sum \theta_i = 1\}$



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, $\{x \mid x \geq 0, 1^T x = 1\}$ (simplex polyhedra). Ensure they sum to 1 and is non negative. Let the probability vector be $\{x : x_1, x_2 \geq 0, x_1 + x_2 = 1\}$. $v_0 = 0, \dots, v_i = e_i$. Simplex would be as shown below. Consider bounded polyhedron or a polytope P . Consider the optimization problem,

$$\min(a^T x)$$

subject to $x \in P$ This is a convex problem as its constraints set is convex and objective function $a^T x$ 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 $a^T v_1 = -1, a^T v_2 = -1$ the any point on the line segment joining v_1 and $v_2, a^T(\theta v_1 + (1 - \theta)v_2) = -1$.

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

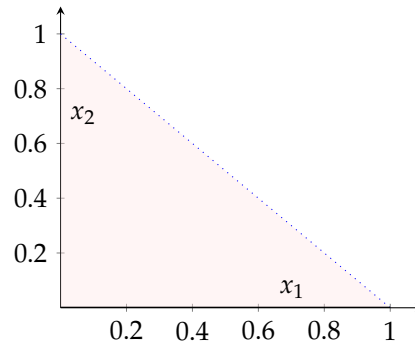


Figure 4: Polytope using 2 linearly independent vectors.

$a^T(\sum_{i=1}^N \theta_i v_i)$, $0 \leq \theta_i \leq 1$ and $\sum_{i=1}^N \theta_i = 1$ As weighted average
 >smallest element in the set,
 $a^T(\sum_{i=1}^N \theta_i v_i) \geq \min_i a^T v_i$ Then value in the function would be
 atleast $a^T v_i$ or value at the vertices. Therefore one of the vertices is an
 optima. \square

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

A L_p -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 $\{x \mid \|x\| \leq r\}$, 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 L_1 unit norm ball takes shape of an octahedron, whereas in 2D it forms a square shape. It can be defined as 1 norm: $\{x \mid |x_1| + |x_2| \leq r\}$. When $p=2$, norm deduces to euclidean distance (Line of sight distance), now unit norm ball takes a circle shape. It can be represented as 2 norm: $\{x \mid x_1^2 + x_2^2 \leq r^2\}$. When $p=\infty$, norm is computed as $\|u\|_{\infty} = \max\{u_i\}$.

Theorem 1. Norm balls are convex.

Proof. Consider 2 points in set X , such that $\|X_1\| \leq r$ and $\|X_2\| \leq r$. Then by triangular inequality, $\|\theta x_1 + (1 - \theta)x_2\| \leq \|\theta x_1\| + \|(1 - \theta)x_2\|$. Since θ is a scalar, by scalar multiplication, $\|\theta x_1 + (1 - \theta)x_2\| \leq \theta \|x_1\| + (1 - \theta) \|x_2\|$. By definition, $\|\theta x_1 + (1 - \theta)x_2\| \leq \theta r + (1 - \theta)r \leq r$. Therefore norm balls are also convex. \square

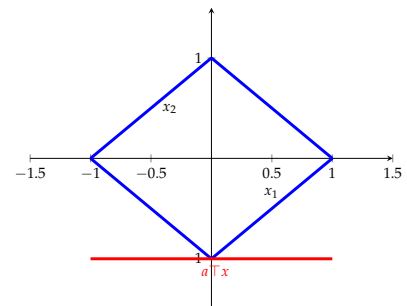


Figure 5: Minima of polyhedra occurs at one of its vertex

Theorem 2. Given transformation matrix P is positive semi definite, Norm balls preserves convexity during stretching/compressing the set.

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

$$\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \quad \text{Assuming all } \lambda_i > 0, \text{ solution deduces to form,}$$

$\{y | \sum_{i=1}^N y_i^2 / \lambda_i^2\}$. This represents an ellipsoid in 2D of the form, $y_1^2/a^2 + y_2^2/b^2 \leq 1$. Therefore given that P is positive semi definite, even after stretching or compressing the set remains convex.

□

Norm Cones

Set C is a cone if for any $x \in C$, $\theta x \in C$ 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 $\{(x, t) | x \in \mathbb{R}^n, t \in \mathbb{R}, \|x\| \leq t\}$. 2 norm cones are generally referred as Lorentz cones. All norm cones are convex.

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

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

When last coordinate is positive and $t = C^T x$, this becomes equivalent to linear combination of first n values, this represents a hyperplane.

Sketch 1 norm and infinite norm balls

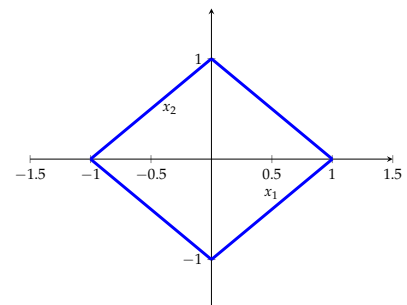


Figure 6: 1 norm ball.

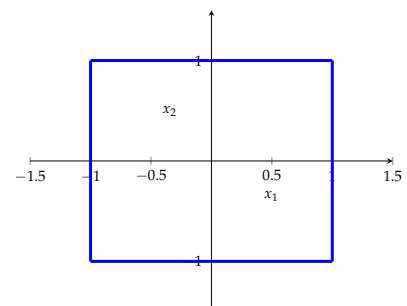


Figure 7: infinite norm ball.

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

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

Therefore a norm cone is convex it is intersection of hyperplane and cone. \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 $\bar{a} \neq 0$ and b such that $\bar{a}^\top \bar{x} \leq b; \forall \bar{x} \in C_1$ and $\bar{a}^\top \bar{x} \geq b; \forall \bar{x} \in C_2$. The hyperplane $\{\bar{x} | \bar{a}^\top \bar{x} = b\}$ is called *separating hyperplane* for the sets C_1 and C_2 .

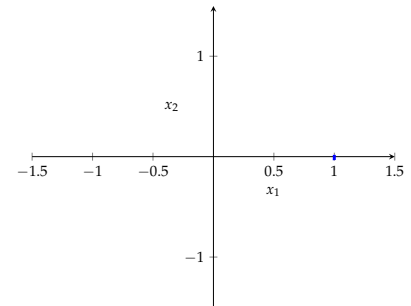


Figure 8: infinite norm ball.