CONVEX SETS & CONVEX FUNCTIONS

Part - 3

→ Relative Interior

Recall that the affine combination of points : $x_1, x_2 \cdots x_m$ is :

$$y = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_m x_m$$

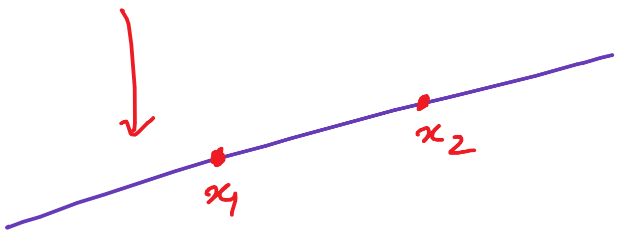
with $\theta_i \in \mathbb{R}$ & $\sum_{i=1}^{n} \theta_i = 1$

• Affine hull of a set S is :

$$\text{aff}(S) := \left\{ y = \sum_{i=1}^{m} \theta_i x_i \;\middle|\; \theta_i \in S ,\ \sum_{i=1}^{n} \theta_i = 1 \right\}$$

Let $S = \{x_1, x_2\}$ then

aff(S) is the line joining x_1 & x_2



- Relative interior Point :

Let $S \subseteq \mathbb{R}^n$ be a set / convex set.

Then, the point x is a relative interior point of S if

(i) $x \in S$ and

(ii) there exists a neighborhood $B(x, \epsilon)$ of x such that

$$\boxed{B(x,\epsilon) \cap aff(S) \subset S}$$

i.e. x is an interior point of S relative to $aff(S)$.

- The set of all relative interior points of S is "relative interior" of S & denoted as $ri(S)$,

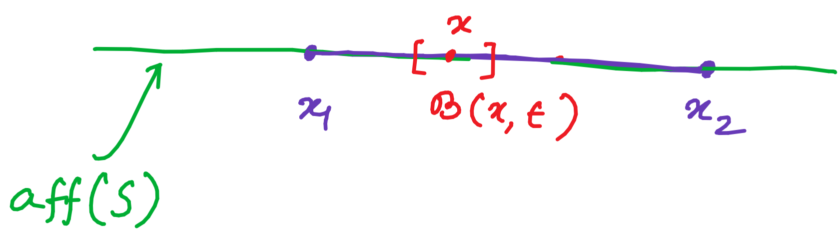
|||

$$\boxed{ri(S) := \left\{ x \in S \mid B(x,\epsilon) \cap aff(S) \subseteq S \text{ for some } \epsilon > 0 \right\}}$$

Example :

def $S := \{ x \mid x = \theta x_1 + (1-\theta) x_2 ,\ \theta \in [0,1] \}$

line segment connecting two distinct points.

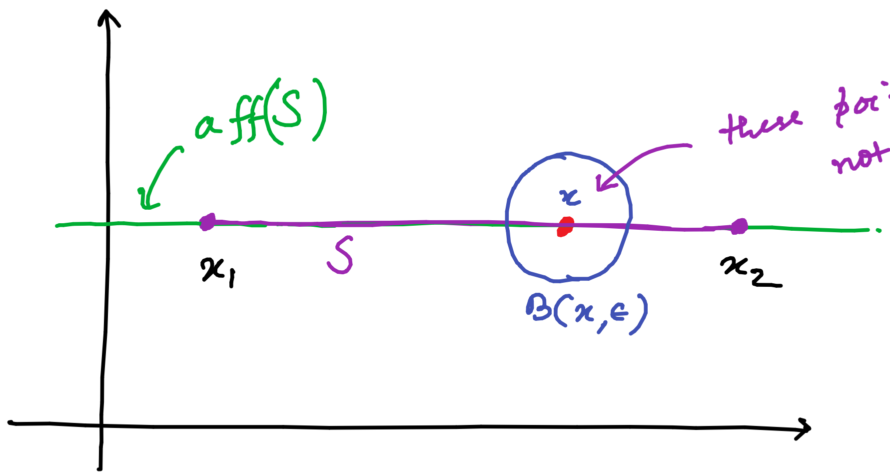


$$B(x,\epsilon) \cap \text{aff}(S) \subset S$$

Hence x is a relative interior point of S.

→ Now consider a line segment between two points in $\mathbb{R}^2$.

According to the definition



these points do not belong to the set S.

⇓

the <u>interior</u> of S is empty.

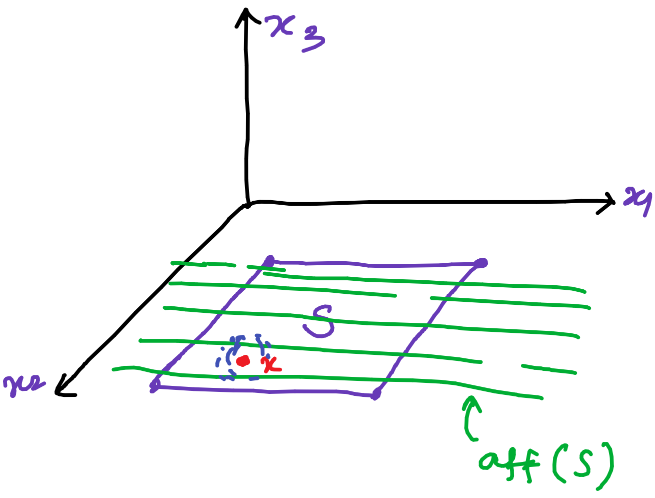
(One can not find a ball around x s.t. the entire ball is contained in S.)

However S <u>has</u> <u>relative interiors</u>.

since $B(x,\epsilon) \cap \text{aff}(S) \subset S$.

Similarly, consider a square in $\mathbb{R}^3$.

$$S = \left\{ x \in \mathbb{R}^3 \;\middle|\; \begin{array}{l} 1 \le x_1 \le 2 \quad x_3 = 0 \\ 1 \le x_2 \le 2 \end{array} \right\}$$



$$\text{aff}(S) = \{ x \in \mathbb{R}^3 \mid x_3 = 0 \}$$

$\uparrow$ entire (x_1, x_2) plane.

To satisfy $x \in S$ to be an interior of S, we need to find a solid sphere around x, i.e. $B(x, \epsilon)$ s.t. $B(x, \epsilon) \in S$.

$B(x,\epsilon)$: Not possible to find. $B(x,\epsilon) \in S$ $= \phi$

However,

$$B(x, \epsilon) \cap \text{aff}(S) \subset S$$

Hence x is a relative interior of S.

→ **Facts**

Let $S \subseteq \mathbb{R}^n$ be convex. Then,

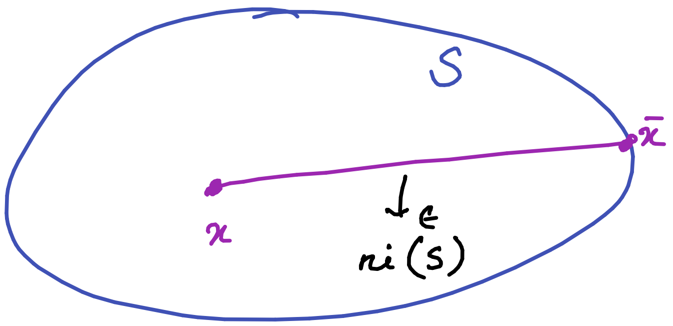
(i) the interior of S $(\text{int}(S))$

relative interior of S $(\text{ri}(S))$

2 closure of S $(\text{cl}(S))$

$\Bigg\}$ are convex.

(ii) Let $x \in ri(S)$ and $\bar{x} \in S$. Then all points on the line segment joining x & $\bar{x}$, except possibly $\bar{x}$, belong to $ri(S)$.



$$\boxed{\begin{array}{r} \theta x + (1-\theta)\bar{x} \in ri(S) \\ \theta \in (0, 1]. \end{array}}$$

→ Some Other results

1. The convex hull of a compact set is compact.

2. Let $f: \mathbb{R}^n \to \mathbb{R}$ be a convex function. Then, it is continuous.

   If $S \subset \mathbb{R}^n$ is convex and $f: S \to \mathbb{R}$ is convex, then f is continuous in the relative interior of S.

3. Let X^* be the set of minimizing points of a convex function $f: \mathbb{R}^n \to \mathbb{R}$ over a closed convex set S. Let X^* be non-empty and bounded. Then, the level sets

$$L_\gamma = \{ x \in X \mid f(x) \leq \gamma \}$$

are compact for each γ.

|||

$$\boxed{\text{If one level set of a convex function } f: \mathbb{R}^n \to \mathbb{R} \text{ is compact, then all level sets are compact.}}$$

→ Local Minimum & Global Minimum

Let $S \subseteq \mathbb{R}^n$ &

$f: S \to \mathbb{R}$ be a real valued function.

- A point $x \in S$ is called "local minimum" of f if there exists an $\epsilon > 0$ s.t.

$$\boxed{f(x) \leq f(y) \text{ for every } y \text{ satisfying } \|x - y\| \leq \epsilon}$$

- A vector x is called "global minimum" of f if

$$\boxed{f(x) \leq f(y) \text{ for every } y \in S}$$

→ Result

Let $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R}$ be a convex function. Then

(i) a local minimum of f is also global minimum of f,

(ii) if f is strictly convex, then there exists at most one global minimum of f.

Proof : Assignment.

- Non-negative combination of points of $S \subseteq \mathbb{R}^n$:

$$\boxed{y = \sum_{i=1}^{k} \theta_i x_i \quad \text{with} \quad \theta_i \geq 0 \;\; \& \;\; x_i \in S}$$

- Positive combination of points of $S \subseteq \mathbb{R}^n$:

$$\boxed{y = \sum_{i=1}^{k} \theta_i x_i \quad \text{with} \quad \theta_i > 0 \;\; \& \;\; x_i \in S}$$

→ Caratheodory's Theorem :

Let S be a non-empty subset of $\mathbb{R}^n$. Then,

(i) every non-zero vector y from cone(S) can be represented as a positive combination of at most 'n' linearly independent vectors from S, i.e.

For every $y \in$ cone(S), we have :

dim(S)

$$\boxed{y = \sum_{i=1}^{m} \theta_i x_i \quad \text{with} \quad \theta_i > 0 \, , \text{ for } i = 1, 2, \cdots m \leq n \;\; \& \;\; x_i \in S}$$

(ii) every vector from conv(S) can be represented as a convex combination of not more than $n+1$ vectors from S.

i.e.

convex hull of S

for $y \in conv(S)$, y can be represented as

$$\boxed{y = \sum_{i=1}^{m} \theta_i x_i \text{ with } \theta_i \geq 0 \text{ for } i = 1, 2, \cdots m \leq n+1, \quad \sum_{i=1}^{m} \theta_i = 1 \ \& \ x_i \in S}$$

Proof of Part - i

Consider a vector $y \in cone(S)$. Then y can be represented by

$$y = \sum_{i=1}^{m} \theta_i x_i \quad \text{with } \theta_i \geq 0 \ \& \ x_i \in S \cdots (*)$$

Let us assume that 'm' is the minimum number (i.e. the minimal set of m vectors $\{x_1, x_2 \cdots x_m\}$) s.t. the representation $(*)$ holds.

linearly independent

$\downarrow$

Since $\{x_1, x_2 \cdots x_m\}$ is the minimal set of m-vectors to represent $y = \sum_{i=1}^{m} \theta_i x_i$, none of the θ_i's are zero i.e. in the representation $(*)$: $\theta_i > 0$ for all $i = 1, 2 \cdots m$.

So $y = \sum_{i=1}^{m} \theta_i x_i$ with $\theta_i > 0$ & $x_i \in S$.

To prove the result by contradiction, let us assume that $\underline{m > n}$. Since $S \subseteq \mathbb{R}^n$ & $x_i \in S$, the set of vectors $\{x_1, x_2 \cdots x_m\}$ are linearly dependent.

$\Downarrow$

there exists scalars μ_i s.t.

$$\sum_{i=1}^{m} \mu_i x_i = 0 \quad \text{with not all } \mu_i = 0$$

Further, since at least one of the vectors in $\{x_1, x_2 \cdots x_m\}$ can be expressed as linear combination of others i.e

$$x_1 = -\sum_{i=2}^{m} \mu_i x_i \qquad \Big[\underset{\substack{\uparrow \\ \mu_1 = 1}}{x_1} + \sum_{i=2}^{m} \mu_i x_i = 0 \Big]$$

We can assume that at least one of the μ_i's is positive (as μ_1)

Then, arrange θ_i & μ_i as follows.

$$\underbrace{\theta_1 \geq \theta_2 \geq \cdots \geq \theta_m = \theta_{min}}_{\theta_i > 0} \qquad \underbrace{\mu_1 \leq \mu_2 \leq \cdots \leq \mu_m = \mu_{max}}_{\mu_i \text{ are combination of } +ve \text{ \& } -ve}$$

$\downarrow$

μ_m is +ve according to this order.

Since $y = \sum_{i=1}^{m} \theta_i x_i$ & $\sum_{i=1}^{m} \mu_i x_i = 0$

y can be represented as

$$y = \sum_{i=1}^{m} \underbrace{(\theta_i - \gamma \mu_i)}_{\bar{\theta}_i} x_i \quad \text{where} \quad \gamma > 0$$

$\downarrow$

In the above representation y becomes conic combination of points $x_i \in S$ if the coefficients $\bar{\theta}_i \geq 0$ (Note that $\theta_i > 0$)

$$\Downarrow$$

$$y \in \text{cone}(S)$$

To ensure that $\bar{\theta}_i \geq 0$, we need to choose the positive scalar $\gamma > 0$ appropriately.

In the set $\{\mu_i\}_{i=1}^{m}$, some are +ve & some are −ve. For −ve μ_i the coefficients $\bar{\theta}_i > 0$ since $\theta_i > 0$. Hence, we just have to consider : +ve μ_i's.

Let K be the set of indices:

$$K := \{ k \mid k \in \{1, 2, \cdots, m\} \;\&\; \mu_k > 0 \}$$

Then, the coefficients $\bar{\theta}_i = \theta_i - \gamma \mu_i \geq 0$

if $\theta_{min} - \gamma \mu_{max} \geq 0$.

Now, select γ as follows, & denote it as $\bar{\gamma}$

$$\boxed{\bar{\gamma} = \frac{\theta_{min}}{\mu_{max}}} \quad \cdots\cdots (**)$$

$$\Downarrow$$

$$\theta_{max} - \bar{\gamma}\mu_{max} = 0$$

$$\Downarrow$$

i.e. $\theta_k - \bar{\gamma}\mu_k = 0$ for $k = m$

$\theta_k - \bar{\gamma}\mu_k \geq 0$ for $k \in K$ with $k \in \{1, 2 \cdots m-1\}$

$$\Downarrow$$

y can be represented as

$$y = \sum_{i=1}^{m-1} \bar{\theta}_i x_i \quad \text{with} \quad \bar{\theta}_i \geq 0$$

& $x_i \in S$.

$$\Downarrow$$

The assumption we had made that $m > n$ is the minimum number of linearly independent vector $\{x_1, x_2 \cdots x_m\}$ required to represent y is not correct. Hence, we arrived at a contradiction for our assumption.

$$\Downarrow$$

The minimal number $m \leq n$.

∎

→ Proof for part (ii)

$S \subseteq \mathbb{R}^n$

Let $y \in conv(S)$ (Convex hull of S)

Then y can be represented as:

$$y = \sum_{i=1}^{m} \theta_i x_i \quad \text{with} \quad \theta_i \geq 0, \quad \sum_{i=1}^{m} \theta_i = 1$$

$$\& \quad x_i \in S.$$

Assume that 'm' be the minimum number s.t. the above representation of y holds

i.e. minimal 'm' number of vectors $\{x_1, x, \cdots x_m\}$ are required to repreent y.

$\Downarrow$

Since m is minimum, in the representation

$$y = \sum_{i=1}^{m} \theta_i x_i,$$ none of the θ_i's are zero

i.e. all $\theta_i > 0$.

Now, assume the the minimum number 'm' is greater then $n+1$ i.e. $m > n+1$.

Consider the following $m-1$ vectors

$$\{x_2 - x_1, x_3 - x_1, \cdots x_m - x_1\}$$

All these vectors belong to $S \subseteq \mathbb{R}^n$.

Since $m-1 > n$ $(m > n+1)$

the set of vectors $\{x_2-x_1, x_3-x_1, \cdots, x_m-x_1\}$

in S are linearly dependent.

$$\Downarrow$$

$\exists$ some scalars μ_i (not all zero) s.t.

$$\sum_{i=2}^{m} \mu_i (x_i - x_1) = 0$$ with at least one of the μ_i's is positive.

Let $\lambda_i = \mu_i$ for $i = 2, 3, \cdots m$

& $\lambda_1 = -\sum_{i=2}^{m} \mu_i$

$$\Downarrow$$

$$\sum_{i=1}^{m} \lambda_i = 0 \quad \cdots\cdots \quad (1)$$

Further, $\sum_{i=2}^{m} \lambda_i (x_i - x_1) = 0$

$$\Rightarrow \sum_{i=2}^{m} \lambda_i x_i - \underbrace{\left(\sum_{i=2}^{m} \lambda_i\right)}_{=\lambda_1} x_1 = 0$$

$$\Rightarrow \sum_{i=1}^{m} \lambda_i x_i = 0 \quad \cdots \quad (2)$$

Since at least one of the μ_i's $i = 2, 3 \cdots m$ is +ve, we have at least one of λ_i's for $i = 2, \cdots m$ is +ve.

Now consider the representation.

$$y = \sum_{i=1}^{m} \underbrace{(\theta_i - \gamma \lambda_i)}_{\bar{\theta}_i} x_i$$

In the above representation, $y \in \text{conv}(S)$

if $\underline{\bar{\theta}_i \geq 0}$ & $\underline{\sum_{i=1}^{m} \bar{\theta}_i = 1}$

Note
$$\sum_{i=1}^{m} \bar{\theta}_i = \sum_{i=1}^{m} \theta_i - \gamma \lambda_i$$

$$= \underbrace{\sum_{i=1}^{m} \theta_i}_{=1} - \gamma \underbrace{\left(\sum_{i=1}^{m} \lambda_i\right)}_{=0}$$

$$= 1$$

Hence we just need to ensure that $\bar{\theta}_i \geq 0$.

We follow the similar procedure as in proof of part - i, & choose γ as follows:

$$\bar{\gamma} = \frac{\theta_{min}}{\lambda_{max}} \quad \left(\text{arranged } \theta_i \text{ \& } \lambda_i \text{ accordingly}\right)$$

For the above selection

$$\theta_k - \bar{\gamma} \lambda_k = 0 \quad \text{for } k = m$$

& $\theta_k - \bar{\gamma} \lambda_k \geq 0$ for $k \in K$, except for m.

$\Rightarrow$ In the representation

$$y = \sum_{i=1}^{m} \bar{\theta}_i x_i \quad \text{becomes}$$

$$y = \sum_{i=1}^{m-1} \bar{\theta}_i x_i \quad \text{with } \bar{\theta}_i \geq 0$$

$$\& \sum_{i=1}^{m} \bar{\theta}_i = \sum_{i=1}^{m-1} \bar{\theta}_i = 1$$

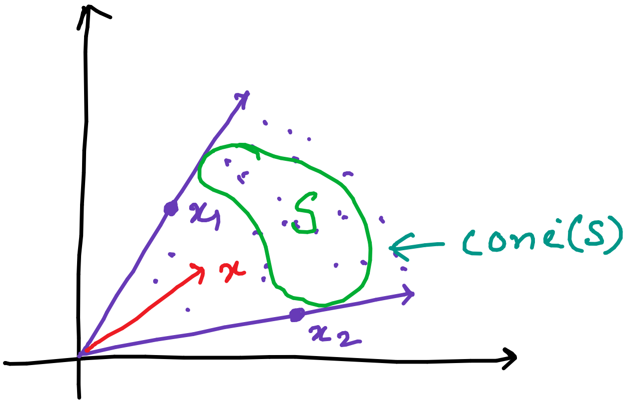
$$\Downarrow \qquad (\text{since } \bar{\theta}_m = 0)$$

The assumption that we had made on minimal number m is not true. Hence we arrived at a contradiction on our assumptions that 'm' is minimal & it is strictly greater than $n+1$.

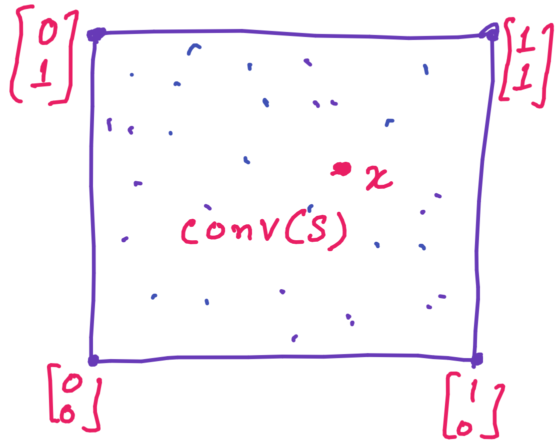
$$\Downarrow$$

To represent $y \in conv(S)$, we need convex combination of at most $n+1$ elements of S.

$\boxed{////}$



The cone generated by x_1 & x_2 in $\mathbb{R}^2$ $\mathcal{K}$. So any $x \in \mathcal{K}$ can be expressed by using at most two vectors x_1 & x_2 which are linearly independent.



Let $S = \left\{ \begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix} \right\}$

then the shaded region is conv(S)

To represent $x \in \mathrm{conv}(S)$ we need at most 3 points of S.

→ ## Radon's Theorem

Let $x_1, x_2 \cdots x_m$ be vectors in $\mathbb{R}^n$ where $m \geq n+2$. Then, there exists a partition of the index set $\{1, 2, \cdots m\}$ into two disjoint sets I & J s.t.

$$\boxed{\mathrm{conv}\left(\{x_i \mid i \in I\}\right) \cap \mathrm{conv}\left(\{x_j \mid j \in J\}\right) \neq \phi}$$

→ Helly's Theorem

Let $\mathcal{F}$ be a finite family of convex sets in $\mathbb{R}^n$. Assume that every $n+1$ sets from the family have a point in common. Then, all the sets in $\mathcal{F}$ have a point in common.