$i \in [m] = i \in \underline{\{1, 2, \ldots m\}}$
$[m]$

Operations that preserve convexity of sets

Proposition 3 (Intersection). *If $X_1, X_2, \ldots, X_m$ are convex sets, then $\boxed{\cap_{i\in[m]} X_i}$ is a convex set.*

Let $y_1, y_2 \in \bigcap_{i\in[m]} X_i$

$\bar{y} = \underline{\lambda y_1 + (1-\lambda) y_2}$, for some $\lambda \in [0,1]$

Example: Polyhedron $\{x \in \mathbb{R}^n | Ax \leq b\}$ for some $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$ which is an intersection of half-spaces.

we want to show $\bar{y} \in \bigcap_{i\in[m]} X_i$.

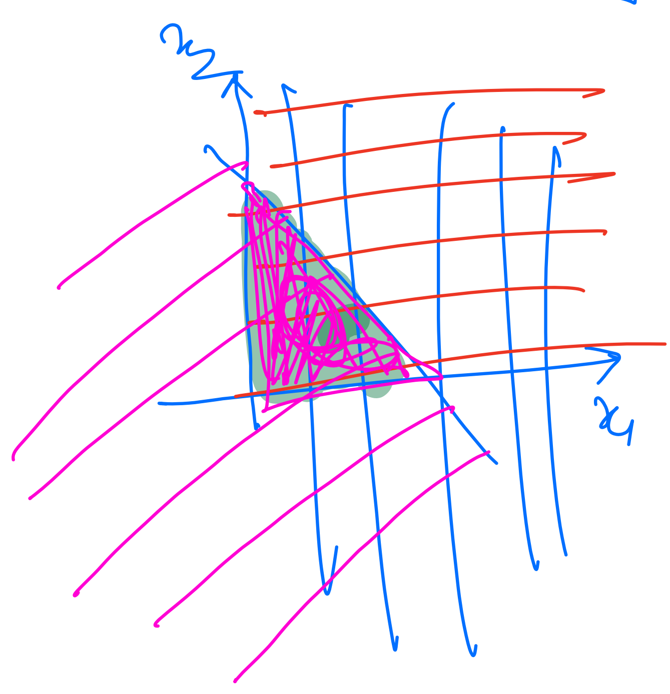
for $\bar{y} \in \bigcap_{i\in[m]} X_i$, it must belong to each X_i's.

$\iff \bar{y} \in X_1, \bar{y} \in X_2, \ldots, \bar{y} \in X_m$

since $y_1, y_2 \in X_K$ and X_K is a convex set,
$\bar{y} = \lambda y_1 + (1-\lambda) y_2 \in X_K$
true for all $K \in \{1, 2 \ldots m\}$

$\Rightarrow \boxed{\bar{y} \in \bigcap_{i\in[m]} X_i}$

$X = \left\{x \in \mathbb{R}^n \middle| \begin{array}{l} x_1 \geq 0, \\ x_2 \geq 0, \\ x_1 + x_2 \leq 1 \end{array}\right\}$



Operations that preserve convexity of sets

Proposition 4 (Affine Image). *If X is a convex set, $f(x) = Ax + b$ with $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$, then the set $f(X) := \{y | y = Ax + b \quad \text{for some} \quad x \in X\}$ is a convex set.* $\subseteq \mathbb{R}^m$

let $y_1, y_2 \in f(X)$

let $z = \lambda y_1 + (1-\lambda) y_2$, for some $\lambda \in [0,1]$

we need to show $z \in f(X) \Rightarrow \exists\, \bar{x} \in X$ s.t. $z = A\bar{x} + b$

Ellipsoid:

Proposition 5. *Let A be a symmetric positive definite matrix. Then, the set $\mathcal{E} := \{x \in \mathbb{R}^n | (x-c)^\top A^{-1}(x-c) \le 1\}$ is convex.*

Since $y_1, y_2 \in f(X)$, $\exists\, x_1 \in X : y_1 = Ax_1 + b$

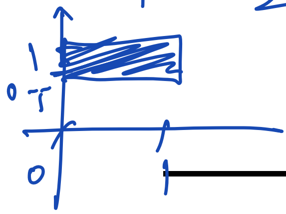
$x_2 \in X : y_2 = Ax_2 + b$

$\bar{x} = \lambda x_1 + (1-\lambda) x_2 \in X$ since X is a convex set

$$A\bar{x} + b = A(\lambda x_1 + (1-\lambda)x_2) + b$$
$$= \lambda A x_1 + (1-\lambda) A x_2 + \lambda b + (1-\lambda) b$$
$$= \lambda [Ax_1 + b] + (1-\lambda)[Ax_2 + b]$$
$$= \lambda y_1 + (1-\lambda) y_2 = z.$$

$\Rightarrow f(X)$ is a convex set.

$X_1 \times X_2 \in \mathbb{R}^2 = \{z \in \mathbb{R}^2 \mid z_1 \in X_1,\ z_2 \in X_2\}$ $\quad X_1 = [0,1] \subseteq \mathbb{R}$, $X_2 = [0.5, 1] \subseteq \mathbb{R}$



Operations that preserve convexity of sets

Proposition 6 (Product)**.** *If $X_1, X_2, \ldots, X_m$ are convex sets, then*

$$X := X_1 \times X_2 \times \ldots \times X_m := \{(x_1, x_2, \ldots, x_m) \mid x_i \in X_i, i \in [m]\}$$

is a convex set. Let $x, y \in X$, $x = (x_1, \ldots, x_m)$, $x_i \in X_i$; $y = (y_1, \ldots, y_m)$, $y_i \in X_i$

Example:

$z = \lambda x + (1-\lambda) y$

$$= \lambda \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} + (1-\lambda) \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \lambda x_1 + (1-\lambda) y_1 \\ \vdots \\ \lambda x_m + (1-\lambda) y_m \end{bmatrix} \begin{matrix} \in X_1 \\ \in X_2 \\ \in X_m \end{matrix}$$

Proposition 7 (Weighted Sum)**.** *If $X_1, X_2, \ldots, X_m$ are convex sets, then* $Y = \sum_{i \in [m]} \alpha_i X_i := \{y \mid y = \sum_{i \in [m]} \alpha_i x_i, \quad x_i \in X_i\}$ *is a convex set.* Since $X_1 \ldots X_m$ are convex sets.

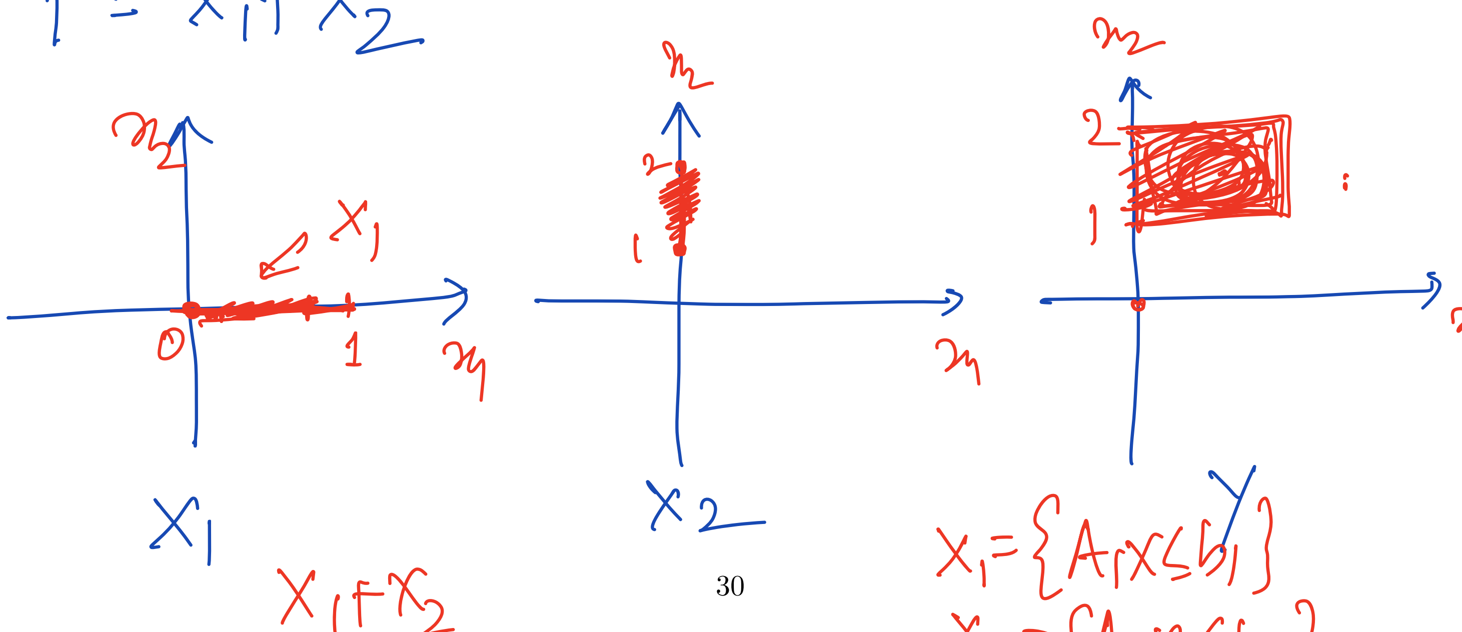
$\alpha_i \in \mathbb{R}$

Example: Minkowski sum of sets

$X_1 = \{x \in \mathbb{R}^2 \mid x_2 = 0, x_1 \in [0,1]\}$

$X_2 = \{x \in \mathbb{R}^2 \mid x_1 = 0, x_2 \in [1,2]\}$

$Y = X_1 + X_2$



$X_1 = \{A_1 x \le b_1\}$

$X_2 = \{A_2 x \le b_2\}$

Operations that preserve convexity of sets

Proposition 8 (Inverse Affine Image)**.** *Let $X \in \mathbb{R}^n$ be a convex set and $\mathcal{A} : \mathbb{R}^m \to \mathbb{R}^n$ be an affine map with $\mathcal{A}(y) = Ay + b$ for matrix A and vector b of suitable dimension. Then, the set $\mathcal{A}^{-1}(X) := \{y \in \mathbb{R}^m \mid Ay + b \in X\}$ is a convex set.*

Let $y_1, y_2 \in A^{-1}(X)$. $\Rightarrow$ $Ay_1 + b \in X$
$Ay_2 + b \in X$

$\lambda \in [0,1]$.

$z = \lambda y_1 + (1-\lambda) y_2$

We need to show $z \in A^{-1}(X)$

$\Leftrightarrow Az + b \in X$

$+ \lambda b_1 + (1-\lambda) b_2$

$\Leftrightarrow A(\lambda y_1 + (1-\lambda) y_2) \in X$

$\Leftrightarrow \underline{\lambda(Ay_1 + b)} + (1-\lambda)(Ay_2 + b) \in X$

$\underline{\lambda x_1 + (1-\lambda) x_2} \in X$ where $x_1 \in X$, $x_2 \in X$

since X is a convex set

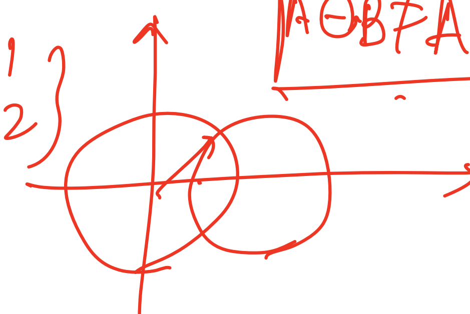
Note: i) Union of convex sets is not convex in general

ii) $\underline{X_1 \ominus X_2} = \{z \mid z = x_1 - x_2, x_1 \in X_1, x_2 \in X_2\}$ is convex

Pontryagin set difference

$\boxed{A \ominus B \neq A \setminus B}$

iii) $X_1 \setminus X_2$ is not convex.



Convex Combination

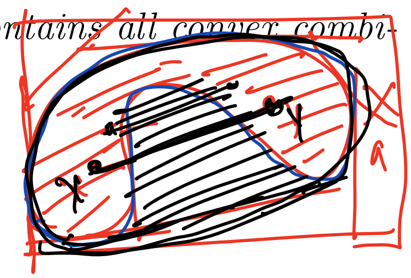
Given a collection of points $x_1, x_2, \ldots, x_k$, the combination $\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_k x_k$ is called **Convex** if $\lambda_i \geq 0$ and $\sum_{i=1}^{n} \lambda_i = 1$.

$\sum_{i=1}^{k} \lambda_i x_i$, $\lambda_i \in \mathbb{R}$

Equivalent Definition:

Definition 4 (Convex Set). *A set is convex if it contains all convex combinations of its points.*

$\mathbb{R}^n$



X is not a convex set.

Definition 5 (Convex Hull). *The convex hull of a set $X \in \mathbb{R}^n$ is the set of all convex combinations of its elements, i.e.,*

$$\mathtt{conv}(X) := \left\{ y \in \mathbb{R}^n \mid y = \sum_{i \in [k]} \lambda_i x_i, \ where \lambda_i \geq 0, \sum_{i \in [k]} \lambda_i = 1, x_i \in X \forall i \in [k], k \in \mathbb{N} \right\}.$$

Proposition 9 (Convex Hull). *The following are true.*

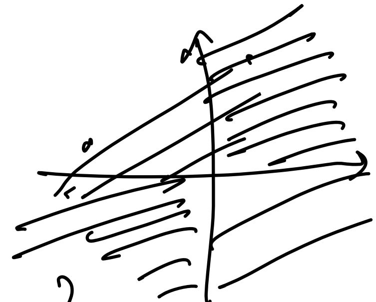
- $\mathtt{conv}(X)$ *is a convex set (even when X is not).*
- *If X is convex, then* $\mathtt{conv}(X) = X$.
- *For any set X,* $\mathtt{conv}(X)$ *is the smallest convex set containing X.*

Example:

$X = [0,1] \cup [2,3]$, is not a convex set

$\mathrm{conv}(X) = [0,3]$: convex Set

$X = \mathbb{R}^2_{++} \cup \mathbb{R}^2_{--}$



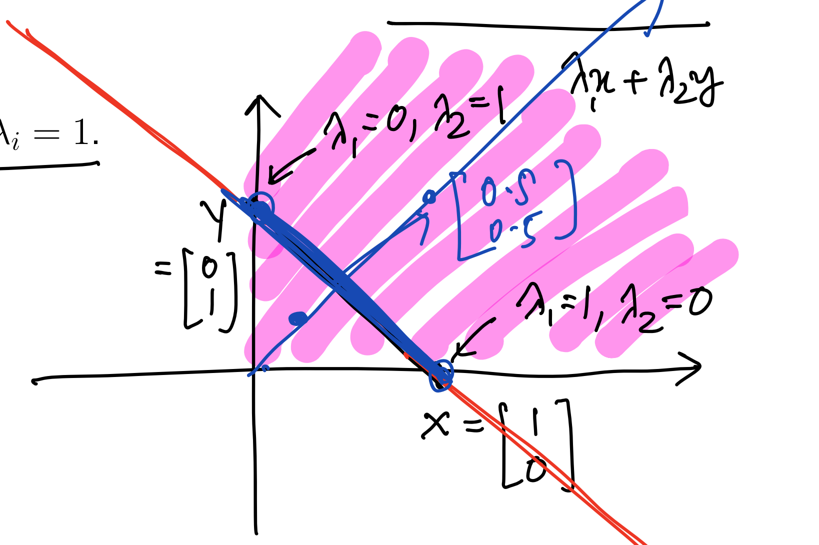
$\mathrm{conv}(X) = \mathbb{R}^2$

Combination of points

Given a collection of points $x_1, x_2, \ldots, x_k$, the combination $\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_k x_k$ is called

- **Convex** if $\lambda_i \geq 0$ and $\sum_{i=1}^{n} \lambda_i = 1$.
- **Conic** if $\lambda_i \geq 0$,
- **Affine** if $\sum_{i=1}^{n} \lambda_i = 1$,
- **Linear** if $\lambda_i \in \mathbb{R}$.

$\mathbb{R}^2$



A set is convex/ convex cone/ affine subspace/linear subspace if it contains all convex/conic/affine/linear combinations of its elements.

$\Rightarrow$ every cone must include the origin

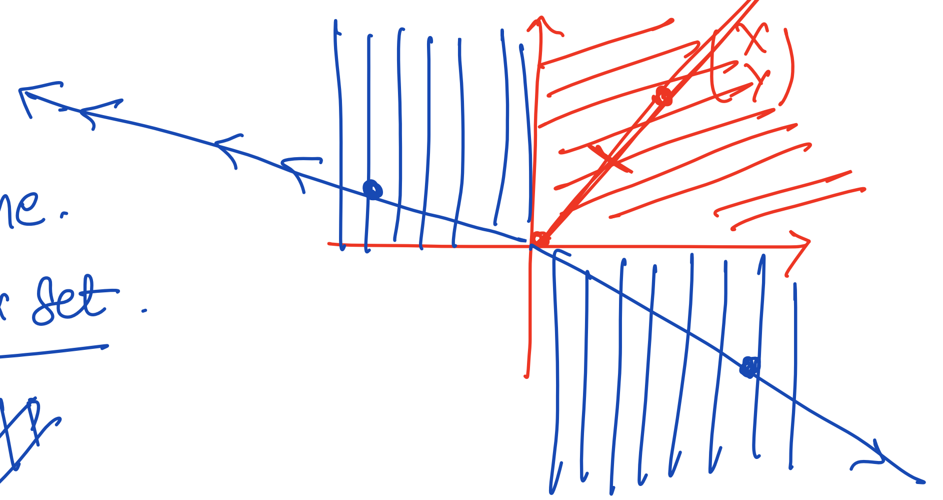
Definition 6. *A set X is a cone if for any $x \in X$, $\boxed{\alpha \geq 0}$, we have $\alpha x \in X$.*

Examples:

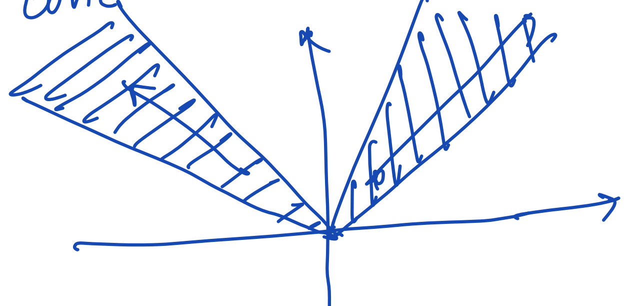
$X_1 = \mathbb{R}^2_{++}$: cone.

$X_2 = \mathbb{R}^2_{+-} \cup \mathbb{R}^2_{-+}$: cone.

$\rightarrow$ not a convex set.



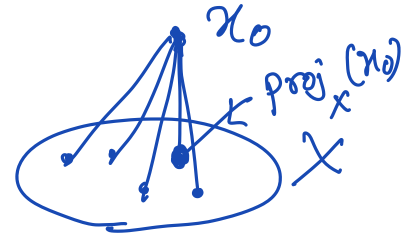
X_3: cone



Projection

Definition 7 (Projection)**.** *The projection of a point x_0 on a set X, denoted* $\texttt{proj}_X(x_0)$ *is defined as*

$$\texttt{proj}_X(x_0) := \text{argmin}_{x \in X} ||x - x_0||_2^2.$$



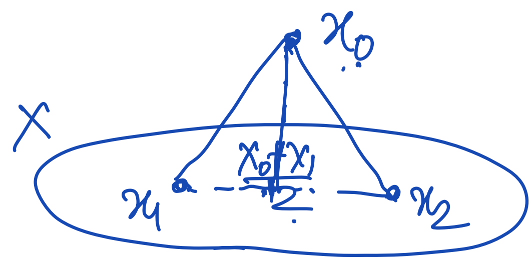
if $x_0 \in X$, $\text{proj}_X(x_0) = x_0$

Theorem 4: Projection Theorem

If X is closed and convex, then $\texttt{proj}_X(x_0)$ exists and is unique.

Main idea:

- Existence due to Weierstrass Theorem
- Uniqueness via contradiction exploiting convexity



Suppose projection is not unique, and there are two points x_1 and x_2 which minimize distance between x_0 & X.

$d_{min} = \|x_0 - x_1\|_2^2 = \|x_0 - x_2\|_2^2 \leq \|x_0 - x\|_2^2 \quad \forall x \in X$

$\|(x_0 - x_1 + x_0 - x_2)\|_2^2$ $\quad ((a+b)^T(a+b))$

$= \|x_0 - x_1\|_2^2 + \|x_0 - x_2\|_2^2 + 2(x_0 - x_1)^T(x_0 - x_2)$

$\|(x_0 - x_1) - (x_0 - x_2)\|_2^2$

$= \|x_0 - x_1\|_2^2 + \|x_0 - x_2\|_2^2 - 2(x_0 - x_1)^T(x_0 - x_2)$

(+)

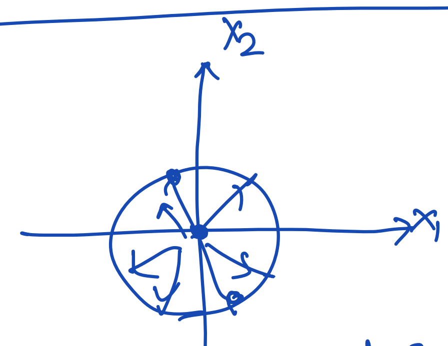
$\|x_0 - x_1 + x_0 - x_2\|_2^2 + \|x_2 - x_1\|_2^2$

$= 2\left[\|x_0 - x_1\|_2^2 + \|x_0 - x_2\|_2^2\right] = 4d_{min}$

$\Rightarrow \|2(x_0 - \frac{x_1+x_2}{2})\|_2^2 + \|x_2 - x_1\|_2^2 = 4d_{min}$

$\Rightarrow \|x_0 - (\frac{x_1+x_2}{2})\|_2^2 + \frac{1}{4}\|x_2 - x_1\|_2^2 = d_{min}$ → +ve since $x_1 \neq x_2$

$\Rightarrow \|x_0 - (\frac{x_1+x_2}{2})\|_2^2 < d_{min} \Rightarrow$ x_1 & x_2 are not projections.



$X = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$

$\text{proj}_X(0) = X$

Projection need not be unique if X is non-convex.

Class Test 1

Let K be a cone.

$$K^* = \{ y \in \mathbb{R}^n \mid \underline{x^T y \geq 0} \quad \forall x \in K \}.$$

i) Let $z \in K^* \Rightarrow x^T z \geq 0 \quad \forall x \in K$.

Let $\alpha \geq 0$, , $x^T(\alpha z) = \alpha(x^T z) \geq 0 \quad \forall x \in K$

$\Rightarrow \alpha z \in K^* \Rightarrow K^*$ is a $\underline{\text{cone}}$.

ii) Let $z_1, z_2 \in K^*$. we need to show $\lambda z_1 + (1-\lambda) z_2 \in K^*$ $\forall \lambda \in [0,1]$.

$$\Rightarrow \left.\begin{array}{l} x^T z_1 \geq 0 \quad \forall x \in K \\ x^T z_2 \geq 0 \quad \forall x \in K. \end{array}\right\}$$

$$x^T \underline{(\lambda z_1 + (1-\lambda) z_2)} = \lambda \underbrace{x^T z_1}_{\geq 0} + (1-\lambda) \underbrace{x^T z_2}_{\geq 0}$$

$$\geq 0 \quad \forall x \in K.$$

iii) If $K_1 \subseteq K_2$, then $\underline{K_2^* \subseteq K_1^*}$.

we need to show if $z \in K_2^*$, then $z \in K_1^*$

Let $z \in K_2^* \Rightarrow x^T z \geq 0 \quad \forall x \in K_2$

$\Rightarrow x^T z \geq 0 \quad \forall x \in K_1$

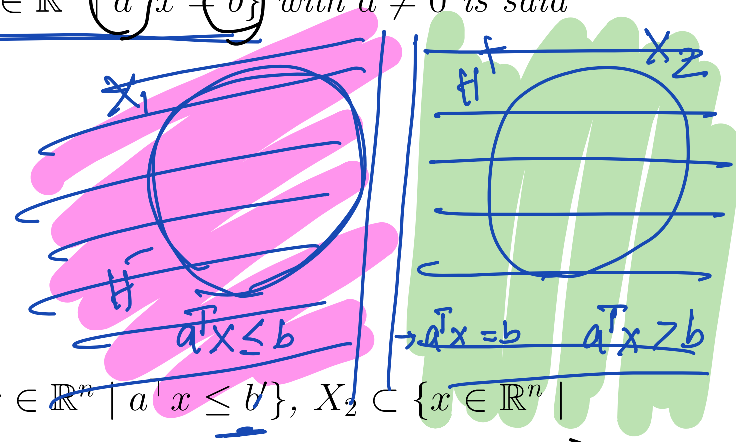
$\Rightarrow \underline{z \in K_1^*}$.

Separating Hyperplane

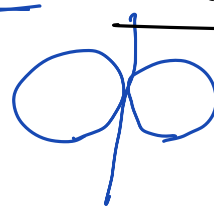
Definition 8 (Separating Hyperplane)**.** *Let X_1 and X_2 be two nonempty convex sets in $\mathbb{R}^n$. A hyperplane $H = \{x \in \mathbb{R}^n \mid a^\top x = b\}$ with $a \neq 0$ is said to separate X_1 and X_2 if*

- $X_1 \subset H^- := \{x \in \mathbb{R}^n \mid a^\top x \leq b\}$,
- $X_2 \subset H^+ := \{x \in \mathbb{R}^n \mid a^\top x \geq b\}$,
- $X_1 \cap X_2 \not\subset H$.

ruled out $X_1 = \{a^\top x = b\}$, $X_2 = \{a^\top x = b\}$



Separation is said to be **strict** *if $X_1 \subset \{x \in \mathbb{R}^n \mid a^\top x \leq b'\}$, $X_2 \subset \{x \in \mathbb{R}^n \mid a^\top x \geq b''\}$ with $b' < b''$.*



Equivalently

$$\boxed{\sup_{x \in X_1} a^\top x \leq \inf_{x \in X_2} a^\top x}$$

with the inequality being strict for strict separation.

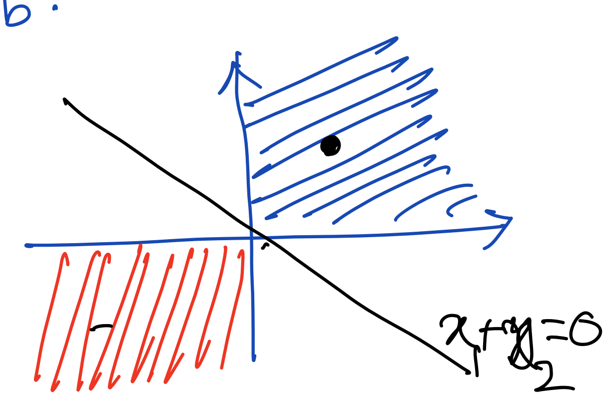
If $X_1 \subset H^-$,

$\iff \sup_{x \in X_1} a^\top x \leq b$

If $X_2 \subset H^+$, $\inf_{x \in X_2} a^\top x \geq b$.

Example: $X_1 = \mathbb{R}^2_{++}$

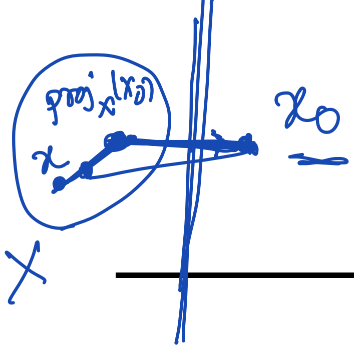
$X_2 = \mathbb{R}^2_{--}$

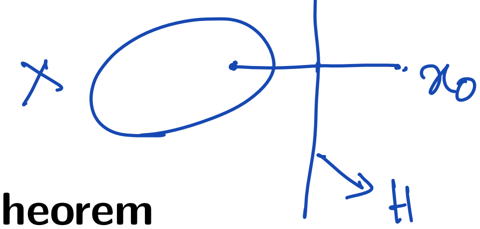


$H = \{x \in \mathbb{R}^2 \mid [1 \;\; 1]x = 0\}$

$a = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $b = 0$

$X_1 \subset H^+$, $X_2 \subset H^-$





Separating Hyperplane Theorem

Theorem 5: Separating Hyperplane Theorem

Let X be a closed convex set and $x_0 \notin X$. Then, there exists a hyperplane that strictly separates x_0 and X.

$H^+ = \{x \in \mathbb{R}^n \mid a^\top x > b\}$, $H^- = \{x \in \mathbb{R}^n \mid a^\top x < b\}$

Main Idea:

1. Let $H = \{x \in \mathbb{R}^n \mid a^\top x = b\}$ with $a = x_0 - \mathtt{proj}_X(x_0)$ and $\boxed{b = a^\top x_0 - \frac{\|a\|_2^2}{2}}$.
2. Use properties of projection and convexity of X to verify that H is indeed the separating hyperplane.

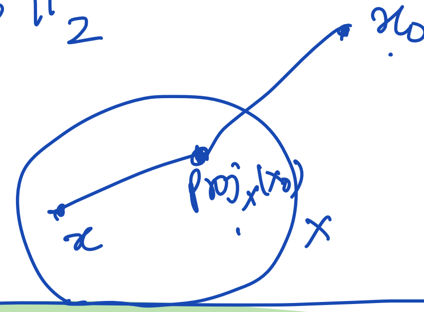
$$\frac{\|a\|_2^2}{2} > 0 \Rightarrow b < a^\top x_0 \Rightarrow x_0 \in H^+$$

To show that H separates x_0 & X, we need to show that $X \subset H^-$, i.e., if $x \in X$, then $a^\top x < b$.

Since X is a convex set, $\lambda x + (1-\lambda)\mathrm{proj}_X(x_0) \in X$ $\forall \lambda \in [0,1]$

Let $\phi(\lambda) = \|\lambda x + (1-\lambda)\mathrm{proj}_X(x_0) - x_0\|_2^2$

$\phi(0) = \|\mathrm{proj}_X(x_0) - x_0\|_2^2$



$\phi(\lambda) \geq \phi(0)$ $\forall \lambda \in [0,1]$ true from the definition of projection.

$$\Rightarrow \phi'(\lambda)\big|_{\lambda=0} \geq 0 \iff \boxed{(x - \mathrm{proj}_X(x_0))^\top (\mathrm{proj}_X(x_0) - x_0) \geq 0}$$

$$\phi(\lambda) = \|\lambda(x - \mathrm{proj}_X(x_0)) + \mathrm{proj}_X(x_0) - x_0\|_2^2$$

$$= \lambda^2 \|x - \mathrm{proj}_X(x_0)\|_2^2 + 2\lambda (x - \mathrm{proj}_X(x_0))^\top (\mathrm{proj}_X(x_0) - x_0)$$

$$\phi'(\lambda)\Big|_{\lambda=0} = \cancel{2\lambda \|x - \text{proj}_X(x_0)\|_2^2}^{\,0} + 2\,(x - \text{proj}_X(x_0))^T(\text{proj}_X(x_0) - x_0) + \|x_0 - \text{proj}_X(x_0)\|_2^2$$

Theorem of the Alternative (Farkas' Lemma)

Lemma 1 (Farkas' Lemma). *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Then, exactly one of the following sets must be empty:*

1. $\boxed{\{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\} = S_1}$
2. $\{y \in \mathbb{R}^m \mid A^\top y \leq 0, b^\top y > 0\} \in S_2$.

Let $y_1 \in \mathbb{R}^m$ s.t. $A^T y_1 \leq 0$, $b^T y_1 > 0$.

Suppose $x_1 \in S_1 \Rightarrow Ax_1 = b$, $x_1 \geq 0$

Insight: If unable to show a system of linear inequalities does not have a solution, try to show that its alternative system does.

$$0 \geq x_1^T(A^T y_1) = y_1^T A x_1 = y_1^T b > 0$$

which is a contradiction

Main Idea:

1. Easy to show that if (2) is feasible, (1) is infeasible.
2. For the converse, suppose (1) is infeasible. Then, $b \notin \boxed{\text{cone}(a_1, a_2, \ldots, a_n)}$ where a_i is the i-th column of A. Find a hyperplane separating b from $\text{cone}(a_1, a_2, \ldots, a_n)$ and show that (2) is feasible.

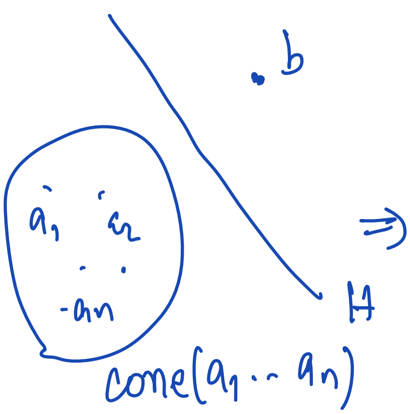
$\{z \in \mathbb{R}^m \mid z = \sum_{i=1}^n a_i x_i,\ x_i \geq 0\}$

$\Downarrow$

is a convex set

can be shown easily.

$Ax = b \iff \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} x = b$,



$\Rightarrow \sum_{i=1}^n a_i x_i = b$

conic combination of columns of the matrix A

since $b \notin \text{cone}(a_1 \cdots a_n)$, $\exists$ hyperplane $H = \{z \in \mathbb{R}^n \mid \alpha^T z = \beta\}$ which strictly separates b from $\text{cone}(a_1 \cdots a_n)$.

$\alpha^T b > \beta$, $\boxed{\alpha^T z \leq \beta}$ $\forall z \in \text{cone}(a_1 \cdots a_n)$

Note that: $\beta \geq 0$ since $0 \in$ cone.

Claim: $\boxed{\beta = 0}$ suppose $\bar{z} \in$ cone for which $\boxed{0 < \alpha^T \bar{z} < \beta}$

$\Rightarrow$ $\alpha^T z \leq 0 \ \forall z \in \text{cone}$

& $\alpha^T b > 0$

$\Rightarrow \alpha \in S_2$

not possible since if $\bar{z} \in$ cone, then a large constant M s.t. $M\bar{z} \in$ cone.

$\alpha^T(M\bar{z}) = M(\alpha^T\bar{z}) > \beta$.

Application: Linear Programming Duality

Consider the following pair of linear optimization problems.

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^\top x \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0. \end{aligned} \tag{P}$$

$$\begin{aligned} \max_{y\in\mathbb{R}^m} \quad & b^\top y \\ \text{s.t.} \quad & A^\top y \leq c, \end{aligned} \tag{D}$$

Theorem 6: LP Duality

If (P) has a finite optimal value, then (D) also has a finite optimal value and both optimal values are equal to each other.

Domain of a Function

- We consider *extended real-valued* functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\} =: \bar{\mathbb{R}}$.

- The (effective) domain of f, denoted $\texttt{dom}(f)$, is the set $\{x \in \mathbb{R}^n \mid |f(x)| < +\infty\}$.

- Example: $f(x) = \frac{1}{x}$. What is $\texttt{dom}(f)$? $\text{dom}(f) = \mathbb{R} \setminus \{0\}$

- $f(x) = \sum_{i=1}^{n} x_i \log(x_i)$. What is $\texttt{dom}(f)$? $= \{x \in \mathbb{R}^n \mid x_i \geq 0 \ \forall i = 1, 2, \dots, n\}$

  $x \in \mathbb{R}^n$

- When $\texttt{dom}(f) \neq \phi$, we say that the function f is *proper*.

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

let $g(x) = \log x$, $\text{dom}(g) = \{x \in \mathbb{R} \mid x > 0\}$

$x \in \mathbb{R}$

$$\lim_{x \to 0} (x \log x) = \lim_{x \to 0} \frac{\log x}{(1/x)} = \lim_{x \to 0} \frac{1/x}{-1/x^2}$$

$$= 0.$$

Convex Functions

Definition 9 (Convex Function)**.** *A function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is convex if*

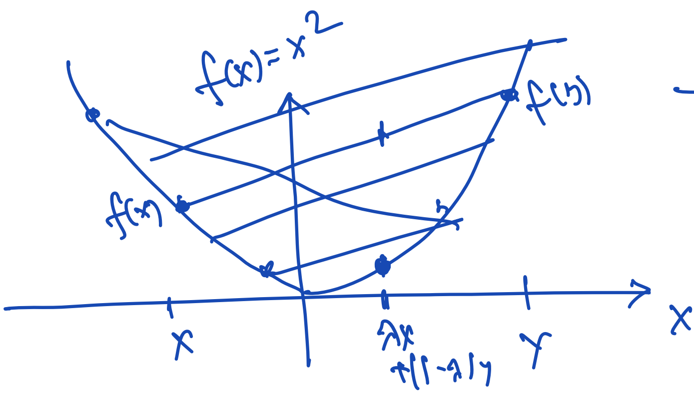
1. $\texttt{dom}(f) \subseteq \mathbb{R}^n$ *is a convex set, and*
2. **for every** $x, y \in \texttt{dom}(f), \lambda \in [0,1]$, *we have* $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$.

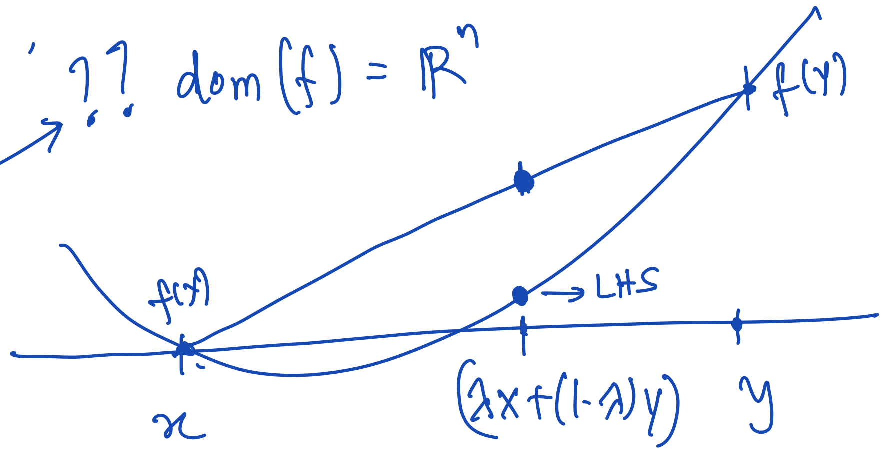
The Line segment joining $(x, f(x))$ and $(y, f(y))$ lies "above" the function.

Examples:

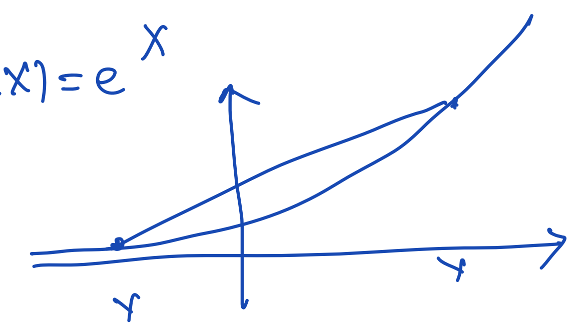
- $f(x) = x^2$, $\mathrm{dom}(f) = \mathbb{R}$,
- $\boxed{f(x) = e^x}$
- $\boxed{f(x) = a^\top x + b \text{ for } x \in \mathbb{R}^n}$

?? $\mathrm{dom}(f) = \mathbb{R}^n$





$f(x) = e^x$



$$
\begin{aligned}
f(\lambda x + (1-\lambda)y) &= a^T(\lambda x + (1-\lambda)y) + b \\
&= \lambda a^T x + (1-\lambda) a^T y + \lambda b + (1-\lambda) b \\
&= \lambda [a^T x + b] + (1-\lambda)[a^T y + b] \\
&= \lambda f(x) + (1-\lambda) f(y) \quad \forall x, y, \lambda \in [0,1]
\end{aligned}
$$

$\Rightarrow a^T x + b$ is a convex function.

Example: Norms

Definition 10 (Norms). *A function $\pi : \mathbb{R}^n \to \bar{\mathbb{R}}$ is a norm if*

- $\pi(x) \geq 0, \quad \forall x$ *and* $\pi(x) = 0$ *if and only if* $x = 0$,
- $\pi(\alpha x) = |\alpha|\pi(x)$ *for all* $\alpha \in \mathbb{R}$,
- $\pi(x + y) \leq \pi(x) + \pi(y)$.

Examples:

- $\boxed{||x||_p := (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}} \text{ for } p \geq 1.}$
- $||x||_Q := \sqrt{x^\top Q x}$ where Q is a positive definite matrix.
- $||A||_F := (\sum_{i=1}^m \sum_{j=1}^n |A_{i,j}|^2)^{1/2}$ Frobenius norm on $\mathbb{R}^{m\times n}$.

$p=2$: $||x||_2 = \left(\sum_{i=1}^n x_i^2\right)^{1/2}$

$p=1$: $||x||_1 = \sum_{i=1}^n |x_i|$

$p=\infty$: $||x||_\infty = \max_i |x_i|$

when $Q = I$, we get $||\cdot||_2$

Proposition 10. *A Norm is a convex function.*

we need to show that for any $x, y \in \mathbb{R}^n$, $\lambda \in [0,1]$,

$$\pi(\lambda x + (1-\lambda) y) \leq \pi(\lambda x) + \pi((1-\lambda)y)$$

$$= \lambda \pi(x) + (1-\lambda)\pi(y)$$

$\Rightarrow$ π is a convex function.

Example: Indicator Function

Definition 11. *Indicator function $I_C(x)$ of a set C is defined as*

$$I_C(x) := \begin{cases} 0, & x \in C, \\ \infty, & x \notin C. \end{cases}$$

Proposition 11. *Indicator function $I_C(x)$ is* a *convex* function *if the set C is a convex set.*

$\mathrm{dom}(I_C) = C$ is a convex set.

Let $x, y \in C$, $\lambda \in [0,1]$.

$$I_C(\underbrace{\lambda x + (1-\lambda)y}_{\in C}) = 0$$

$$\le \lambda \underbrace{I_C(x)}_{=0} + (1-\lambda)\underbrace{I_C(y)}_{=0} = 0$$

$\Rightarrow I_C$ is a convex function.

Example: Support Function

Proposition 12. *Support function of a set C is defined as $I_C^*(x) := \sup_{y \in C} x^\top y$. Support function of a set is always a convex function.*

Let $x_1, x_2 \in \text{dom}(I_C^*)$, let $\lambda \in [0, 1]$

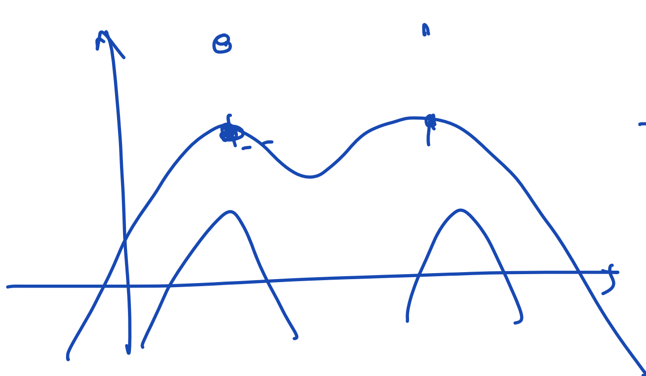
we need to show

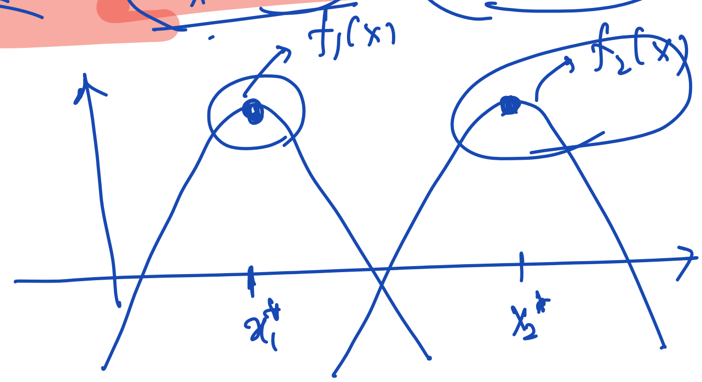
$$I_C^*(\lambda x_1 + (1-\lambda) x_2) \leq \lambda I_C^*(x_1) + (1-\lambda) I_C^*(x_2)$$

$$= \sup_{y \in C} \left[(\lambda x_1 + (1-\lambda) x_2)^\top y \right]$$

$$= \sup_{y \in C} \left[\lambda x_1^\top y + (1-\lambda) x_2^\top y \right] \leq \lambda \sup_{y \in C} x_1^\top y + (1-\lambda) \sup_{y \in C} x_2^\top y$$

$$\max_x \left[f_1(x) + f_2(x) \right] \leq \max_x f_1(x) + \max_x f_2(x)$$





Special Types of Convex Functions

$\forall x, y \in \mathrm{dom}(f),\ \lambda \in (0,1)$
$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y)$

Definition 12. *A function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is*

- **strictly convex** *if property (2) above holds with strict inequality for $\lambda \in (0,1)$,*
- *μ-**strongly convex** if $f(x) - \mu\frac{||x||_2^2}{2}$ is convex, and* for $\mu > 0$.
- $\boxed{\textbf{concave } \textit{if } -f(x) \textit{ is convex.}}$

ex: $f(x) = a^T x + b$ is not strictly convex.

$f(x) = x^2$ with $\mu = 1/2$ because

$x^2 - \frac{1}{4}||x||_2^2 = \frac{3}{4}x^2$ is also a convex function

ex: $f(x) = x^2$ is strictly convex.

If $f(x)$ is a convex function, then $-f(x)$ is a concave function.

ex: $-x^2$ is concave
$-e^x$ is concave

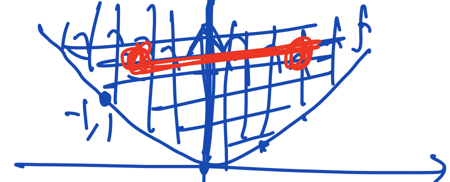
Jensen's Inequality

Proposition 13. *For a convex function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$, for any collection of points $\{x_1, x_2, \ldots, x_k\}$, we have $f(\sum_{i=1}^k \lambda_i x_i) \leq \sum_{i=1}^k \lambda_i f(x_i)$ when $\lambda_i \geq 0$ and $\sum_{i=1}^k \lambda_i = 1$.*

Proof is straightforward via induction.

$f(x)=x^2$, $f:\mathbb{R}\to\mathbb{R}$

sketch epi(f).



Epigraph Characterization

$(-1,-1)$

$\Rightarrow f(-1)=1 \nleq -1$

Definition 13. *A epigraph of a function* $f:\mathbb{R}^n\to\bar{\mathbb{R}}$ *is defined as the set*

$$\texttt{epi}(f) := \{(x,t)\in\mathbb{R}^{n+1} | f(x)\le t\}. \subseteq \mathbb{R}^{n+1}$$

$x\in\mathbb{R}^n$

$t\in\mathbb{R}$

Example: Norm cone: $\boxed{\{(x,t)|||x||\le t\} \text{ is a convex set.}}$

because it is epigraph of $||x||$ and norm is a function function.

Proposition 14. *Function* $f:\mathbb{R}^n\to\bar{\mathbb{R}}$ *is convex in* $\mathbb{R}^n$ *if and only if its epigraph is a convex set in* $\mathbb{R}^{n+1}$.

$\Rightarrow$ Let f be a convex function. Show that epi(f) is a convex set.

Let $(x_1,t_1)\in$ epi(f) $\Rightarrow f(x_1)\le t_1$

$(x_2,t_2)\in$ epi(f) $\Rightarrow f(x_2)\le t_2$

we need to show $\lambda\begin{pmatrix}x_1\\t_1\end{pmatrix}+(1-\lambda)\begin{pmatrix}x_2\\t_2\end{pmatrix}\in$ epi(f)

or $f(\lambda x_1+(1-\lambda)x_2)\le\lambda t_1+(1-\lambda)t_2$

$\forall\lambda\in[0,1]$.

Since f is a convex function,

$f(\lambda x_1+(1-\lambda)x_2)\le\lambda f(x_1)+(1-\lambda)f(x_2)$

$\le\lambda t_1+(1-\lambda)t_2$

converse: HW

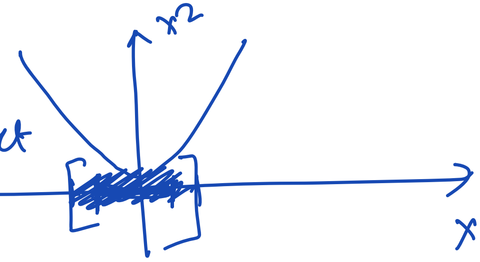
$\Rightarrow \lambda\begin{pmatrix}x_1\\t_1\end{pmatrix}+(1-\lambda)\begin{pmatrix}x_2\\t_2\end{pmatrix}\in$ epi(f)

(try to prove by contradiction).

$\mathrm{lev}_2(f) = (-\infty, \ln 2]$ → convex set

$f(x) = x^2$

$\mathrm{lev}_1(f) = [-1, 1]$ is a convex set



Level-set Characterization

Definition 14. *For any $\alpha \in \mathbb{R}$, the level set of function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ at level α is defined as*

$$\mathtt{lev}_\alpha(f) := \{x \in \mathtt{dom}(f) | f(x) \leq \alpha\}.$$

Proposition 15. *If a function f is a convex function, then* **every** *level set of f is a convex set.*

|| If $\mathrm{lev}_\alpha(f)$ is not a convex set for some $\alpha \in \mathbb{R}$, then f is not a convex function.

Converse is not true. A function is called quasi-convex if its domain and all level sets are convex sets.

Ex: $\sqrt{|x|}$: not a convex function.

HW: Give an example of a function which is quasi-convex but not convex.

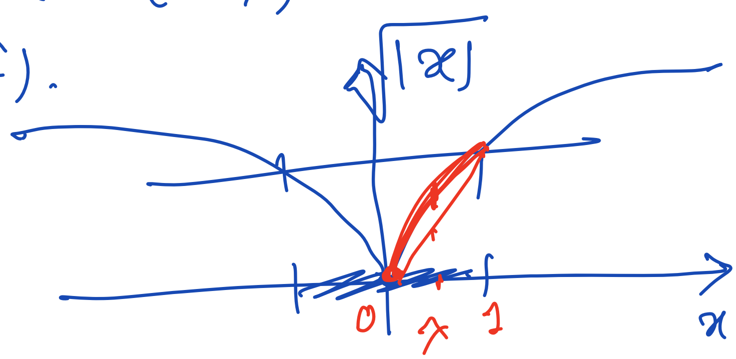
Let f be a convex function. Let $\alpha \in \mathbb{R}$.
We wish to show $\mathrm{lev}_\alpha(f)$ is a convex set.

Let $x_1, x_2 \in \mathrm{lev}_\alpha(f) \Rightarrow f(x_1) \leq \alpha$, $f(x_2) \leq \alpha$

need to show $\lambda x_1 + (1-\lambda) x_2 \in \mathrm{lev}_\alpha(f)$ $\forall \lambda \in [0,1]$.

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$
$$\leq \lambda \alpha + (1-\lambda)\alpha = \alpha.$$

$\Rightarrow \lambda x_1 + (1-\lambda) x_2 \in \mathrm{lev}_\alpha(f)$.

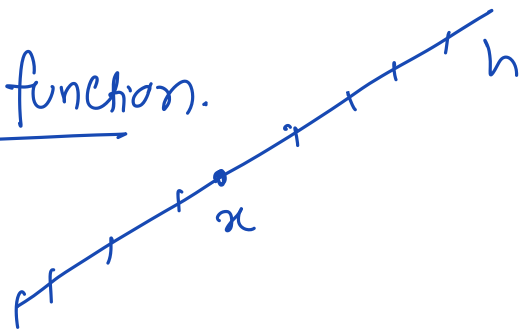


Restriction of a Convex Function on a Line

Proposition 16. *If a function f is convex if and only if for any $x, h \in \mathbb{R}^n$, the function $\phi(t) = f(x + th)$ is a convex function on $\mathbb{R}$.*

$t \in \mathbb{R}$

If we know how to check convexity of functions defined on $\mathbb{R}$, then we can check convexity of functions defined on $\mathbb{R}^n$.



($\Rightarrow$) Suppose f is a convex function.
we want to show that
ϕ is a convex function
for some $x, h \in \mathbb{R}^n$.

We need to show

$$\phi(\lambda t_1 + (1-\lambda) t_2) \leq \lambda \phi(t_1) + (1-\lambda)\phi(t_2) \quad \forall\, t_1, t_2 \in \mathrm{dom}(\phi) \;\&\; \lambda \in [0,1].$$

$$\begin{aligned}
\phi(\lambda t_1 + (1-\lambda) t_2) &= f\big(x + (\lambda t_1 + (1-\lambda) t_2) h\big) \\
&= f\big(\lambda x + (1-\lambda) x + \lambda t_1 h + (1-\lambda) t_2 h\big) \\
&= f\big(\lambda [x + t_1 h] + (1-\lambda)[x + t_2 h]\big) \\
&\leq \lambda f[x + t_1 h] + (1-\lambda) f[x + t_2 h] \\
&= \lambda \phi(t_1) + (1-\lambda)\phi(t_2). \quad \blacksquare
\end{aligned}$$

($\Leftarrow$) HW

First Order Condition

Proposition 17. *If a function f is differentiable, then it is convex if and only if* $\mathtt{dom}(f)$ *is a convex set and for any* $x, y \in \mathtt{dom}(f)$, *we have*

$$f(y) \geq \boxed{f(x) + \nabla f(x)^\top (y - x)}.$$

$$g(y) = f(x) + \nabla f(x)^T (y - x)$$

$$\boxed{f(y) \geq g(y)}$$

$\forall y \in \mathrm{dom}(f)$.

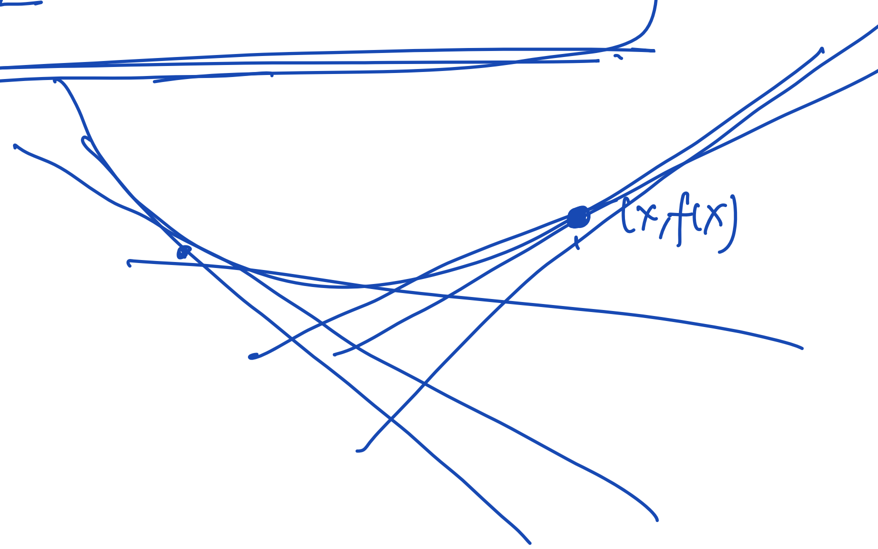
A global lower bound on the function can be obtained at any point based on local information $(f(x), \nabla f(x))$.

$h(x)$ is a convex function of x

$\Vert$ convex in x for every y.

$$f(x) = \boxed{\sup_{y \in \mathrm{dom}(f)} \Big[\overbrace{\nabla f(y)^T (x - y) + f(y)}^{g(x,y)} \Big]}$$

this is probably true for convex functions, but need to verify.



Second Order Condition

Proposition 18. *If a function f is twice differentiable, then it is convex if and only if* $\texttt{dom}(f)$ *is a convex set and* $\nabla^2 f(y) \succeq 0$ *for every* $y \in \texttt{dom}(f)$.

---

① $f_1(x) = x_1 x_2$, $\text{dom}(f) = \{x_1 > 0, x_2 > 0\}$.

$\nabla f_1(x) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$, $\nabla^2 f_1(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, eigenvalues $= \{1, -1\}$

$\Rightarrow \nabla^2 f(x)$ not positive semidefinite

$\Rightarrow f_1$ is not a convex function.

② $f_2(x) = \|Ax - b\|_2^2$, $\nabla^2 f_2(x) = 2A^T A \rightarrow$ positive semidefinite

$\Rightarrow f_2$ is a convex function.

③ $f_3(x) = x^T A x$, $\nabla^2 f_3(x) = A + A^T$, f_3 is convex only when $A + A^T$ is positive semidefinite

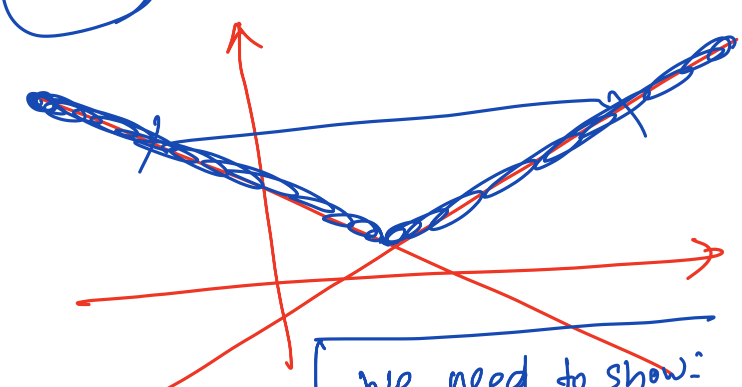
④ $f_4(x) = \max\left[a_1^T x + b_1, a_2^T x + b_2\right]$

Yes.

Can we show that $\text{epi}(f_4)$ is a convex set?

Let $\begin{pmatrix} z_1 \\ t_1 \end{pmatrix}, \begin{pmatrix} z_2 \\ t_2 \end{pmatrix} \in \text{epi}(f_4)$

$\Rightarrow f_4(z_1) \le t_1$, $f_4(z_2) \le t_2$



We need to show:

$\Rightarrow \max\left[a_1^T z_1 + b_1,\ a_2^T z_2 + b_2\right] \leq t_1$

$\Rightarrow a_1^T z_1 + b_1 \leq t_1$ & similarly $a_2^T z_2 + b_2 \leq t_2$

$a_2^T z_1 + b_2 \leq t_1$ $\qquad a_1^T z_2 + b_1 \leq t_2$

$f_4(\lambda z_1 + (1-\lambda) z_2) \leq \lambda t_1 + (1-\lambda) t_2$

→ shown on Board

Convexity Preserving Operations

Proposition 19 (Conic Combination). *Let $\{f_i(x)\}_{i\in I}$ be a collection of convex functions and let $\alpha_i \geq 0$ for all $i \in I$. Then, $g(x) := \sum_{i\in I} \alpha_i f_i(x)$ is a convex function.*

Ex: $g(x) = x^2 + e^x$ $\qquad g(x) = -x^2$

$\alpha_i < 0$ is not allowed since negative of a convex function is not convex, rather concave.

Proposition 20 (Affine Composition). *If $f : \mathbb{R}^m \to \mathbb{R}$ is a convex function, then $g(x) := f(Ax+b)$ is also a convex function where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$.*

$$g(\lambda x_1 + (1-\lambda) x_2) = f(A(\lambda x_1 + (1-\lambda) x_2) + b)$$
$$= f(\lambda(Ax_1 + b) + (1-\lambda)(Ax_2 + b))$$
$$\leq \lambda g(x_1) + (1-\lambda) g(x_2).$$

Ex: $g(x) = \|Ax+b\|_2^2 = f(Ax+b)$, $\quad f(y) = \|y\|_2^2$

convex

$$\boxed{h(x) = -\sum_{i=1}^n \log(a_i^T x + b)} = \sum_{i=1}^n \underbrace{\left(-\log(a_i^T x + b_i)\right)}$$

It is sufficient to show that $-\log(a_i^T x + b_i)$ is convex.

further, it is sufficient to show that $-\log x$ is convex.

$h(x) = -\log x$, $\quad h'(x) = -\frac{1}{x}$, $\quad h''(x) = \frac{1}{x^2} > 0$

$\Rightarrow$ h is convex.

Convexity Preserving Operations

Proposition 21 (Pointwise Maximum). *Let $\{f_i(x)\}_{i \in I}$ is a collection of convex functions, then $g(x) := \max_{i \in I} f_i(x)$ is a convex function.*

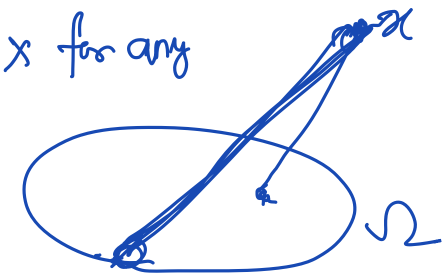
$$g(x) = \max\left(x, 1, x^2, -\log x\right)$$
$$= \max\left(a_1^T x + b_1 \; \ldots \; a_n^T x + b\right)$$

The set I need not be a finite set.

Proposition 22 (Pointwise Supremum). *Let $f(x, \omega)$ is convex in x for any $\omega \in \Omega$, then $\boxed{g(x) := \sup_{\omega \in \Omega} f(x, \omega)}$ is convex in x.*

$f(x, \omega) = \|x - \omega\|_2^2$, convex in x for any ω

$$\boxed{g(x) = \sup_{\omega \in \Omega} \|x - \omega\|_2^2}$$



The function f need not be convex in ω and the set Ω need not be a convex set.

$g(x) = F\begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}$

Convexity Preserving Operations

Proposition 23 (Scalar Composition). *If a function f is convex in $\mathbb{R}^n$, and F is a convex non-decreasing function on $\mathbb{R}$, then $g(x) := F(f(x))$ is convex.*

If $f, g : \mathbb{R} \to \mathbb{R}$, and are twice differentiable, then

$$g'(x) = F'(f(x))\, f'(x)$$
$$g''(x) = F''(f(x))\, \boxed{(f'(x))^2} + F'(f(x))\, f''(x)$$

Proposition 24 (Vector Composition). *Let $\{f_i\}_{i \in \{1,2,\ldots m\}}$ are convex functions on $\mathbb{R}^n$, and $F : \mathbb{R}^m \to \mathbb{R}$ is a convex function and non-decreasing in each argument, then the function $g(x) = F(f(x))$ is convex.*

If $\underbrace{F''(\cdot) \geq 0}_{\Downarrow \text{ convex } f^n}$, $\underbrace{F'(\cdot) \geq 0}_{\Downarrow \text{ monotonically increasing}}$, $\underbrace{f''(\cdot) \geq 0}_{\Downarrow \text{ convex function}}$, then $\underbrace{g''(\cdot) \geq 0}_{\text{is convex}}$

Ex: $g(x) = e^{(a^Tx+b)}$, $F : e^{(\cdot)}$, $f : a^Tx + b$

$e^{\|x\|_2^2}$

$+\begin{cases} F''(\cdot) \geq 0 \\ F'(\cdot) \geq 0, \Rightarrow g''(\cdot) \geq 0 \;\&\; g \text{ is convex} \\ f''(\cdot) \leq 0 \end{cases}$

$g(x) = \dfrac{1}{\log x}$, $F : \dfrac{1}{x}$, $f : \log x$

$f'' = -\dfrac{1}{x^2} < 0$.

F is convex, decreasing, f is concave $\Rightarrow$ g is convex.

If $\underbrace{F'' \leq 0,\ F'(\cdot) \leq 0}_{\text{concave}, \downarrow}$, $\underbrace{f''(\cdot) \geq 0}_{\text{convex}} \Rightarrow \underbrace{g''(\cdot) \leq 0}_{\text{concave}}$