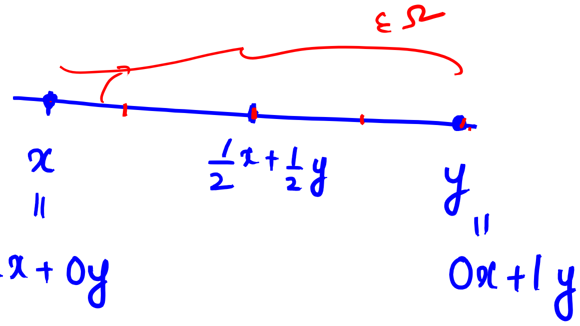
l.c: $\alpha x + \beta y$
$\alpha, \beta \in \mathbb{R}$.

CONVEX SETS

A set Ω is said to be convex if $\forall x, y \in \Omega$, $\lambda x + (1-\lambda) y \in \Omega$ $\forall \lambda \in [0,1]$

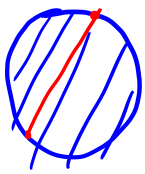
$\rightarrow$ convex combination of x & y.

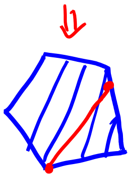


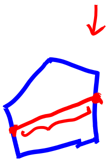
A combination of the form $\lambda x + (1-\lambda)y$ is called a convex combination of points x & y

Multiple points $x_1, \ldots x_n$: What is a convex combination here? $\sum_{i=1}^{n} \lambda_i x_i$, $\sum_{i=1}^{n} \lambda_i = 1$, $\lambda_i \geq 0$.

Examples:

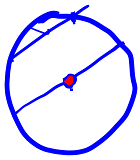




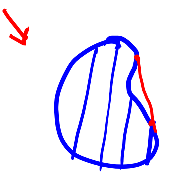


Not convex

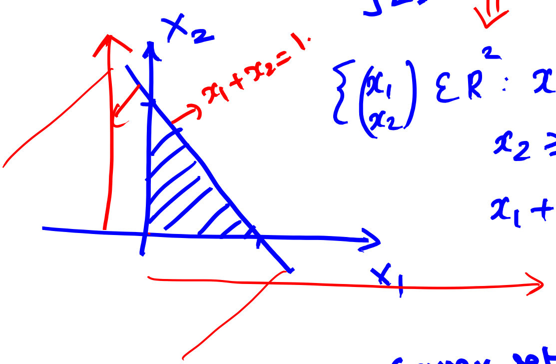
$\Omega = \{x : \text{sides of the polygon}\}$



Not convex



Not convex.

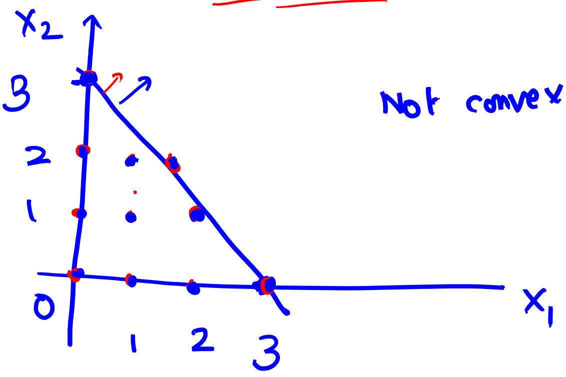


$\Omega = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_1 \geq 0, \; x_2 \geq 0, \; x_1 + x_2 \leq 1 \right\}$

Convex set

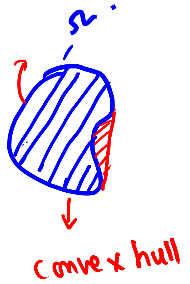
$$\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_1 \geq 0,\ x_2 \geq 0,\ x_1 + x_2 \leq 3 \right\}.$$

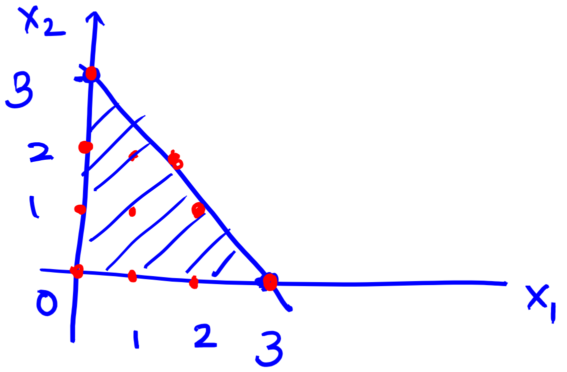
x_1, x_2 integers



Not convex

Convex hull of a set Ω: Smallest convex set containing Ω.





If Ω is already convex, its convex hull $= \Omega$.

CONVEX FUNCTIONS.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if

$\mathrm{dom}(f)$ is a convex set & $\forall x, y \in \mathrm{dom}(f)$

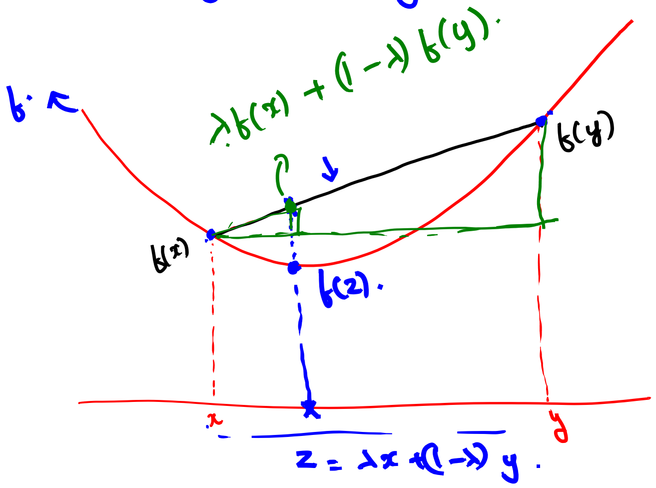
$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$$

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

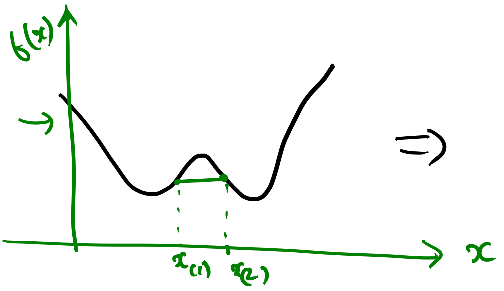
$\lambda x + (1-\lambda) y$: lies on the line connecting x & y.

z.

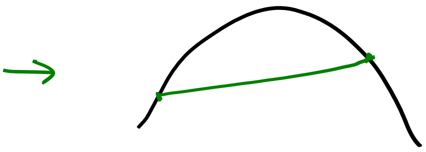
$f(\lambda x + (1-\lambda) y) \le \lambda f(x) + (1-\lambda) f(y)$.



The line joining any 2 points on the curve lies above the curve



$\Rightarrow$ Not convex



$\Rightarrow$ Not convex (Concave)

A function is concave if $-f$ is convex.

Consider the following optimization problem

$$\begin{cases} \min f(x) \\ \text{s.t. } x \in \Omega \end{cases}$$

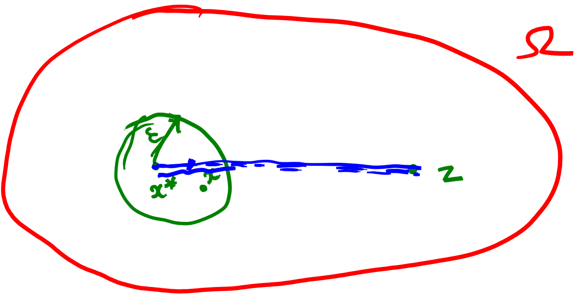
f : convex

Ω : convex set

Any local minimum is also a global minimum for this class of problems.

Proof. Let x^* be a local minimum. So $\exists$ some ε neighborhood around x^* such that

$\forall x : \|x - x^*\| \leq \varepsilon$ we have $f(x) \geq f(x^*)$



Assume x^* is not a global minimum. $f(z) < f(x^*)$

Let z be a global min

Then $f(z) \leq f(x)$ $\forall x \in \Omega$

f is convex:

$$f(\lambda x^* + (1-\lambda) z) \leq \lambda f(x^*) + (1-\lambda) \underbrace{f(z)}_{< f(x^*)}$$

$$< \lambda f(x^*) + (1-\lambda) f(x^*)$$

$$< f(x^*)$$

Not possible for λ small.

Examples of convex sets:

1) Hyperplanes:

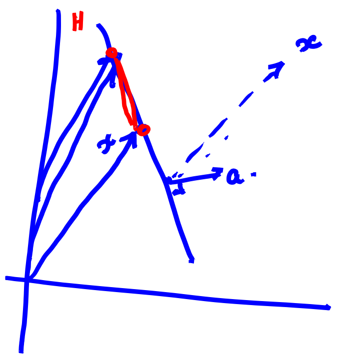
$H = \{ x \mid a^T x = b \}$, $a \in R^n$, $b \in R$, $a \neq 0$

$x_1, x_2 \in H \Rightarrow a^T x_1 = b$, $a^T x_2 = b$.

$a^T(\lambda x_1 + (1-\lambda)x_2)$

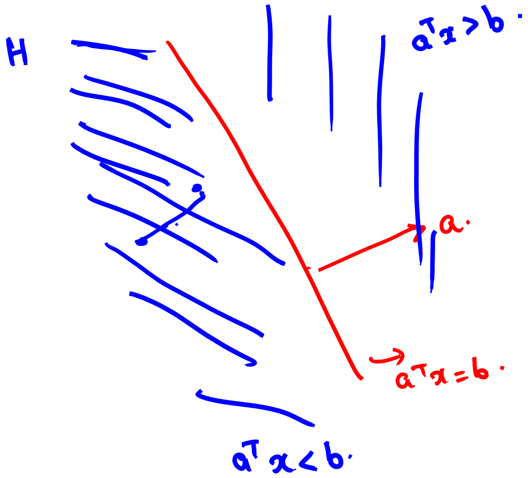
$= \lambda \underbrace{a^T x_1}_{b} + (1-\lambda)\underbrace{a^T x_2}_{=b}$

$= b$.



2) Halfspaces

$H = \{ x \mid a^T x \leq b \}$ $a \in R^n$, $b \in R$, $a \neq 0$.



$x_1, x_2 \in H$.

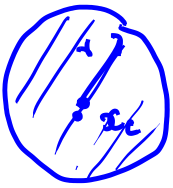
$a^T x_1 \leq b$, $a^T x_2 \leq b$.

$a^T(\lambda x_1 + (1-\lambda)x_2)$

$= \lambda \underbrace{a^T x_1}_{\leq b} + (1-\lambda)\underbrace{a^T x_2}_{\leq b}$

$\leq b$

3) Euclidean balls



$$B = \{x \mid \|x - x_c\| \le r\}$$

x_1, x_2

$\|x_1 - x_c\| \le r$

$\|x_2 - x_c\| \le r$

$$\|\lambda x_1 + (1-\lambda)x_2 - \underbrace{x_c}_{\lambda x_c + (1-\lambda) x_c}\| \overset{?}{\le} r$$

$$\le \|\lambda(x_1 - x_c)\| + \|(1-\lambda)(x_2 - x_c)\|$$

$$\le \lambda r + (1-\lambda) r \le r.$$

4) Set of psd matrices

$S_+ = \{ P \in \underline{\underline{\mathbb{R}^{n\times n}}} \mid$ P symmetric, $P \succeq 0$

P is psd

$\Leftrightarrow P = AA^T.$

$A \in \mathbb{R}^{n\times k}.$

<u>HW.</u>

$P \in S_+$

$\Rightarrow x^T P x \ge 0. \ \forall x \in \mathbb{R}^n$

$P, Q \in S_+$

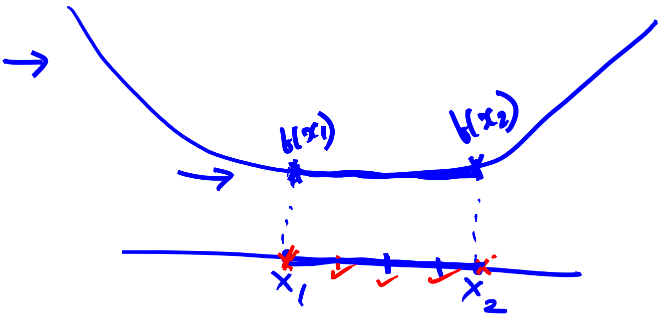
$\lambda P + (1-\lambda) Q \in S_+.$

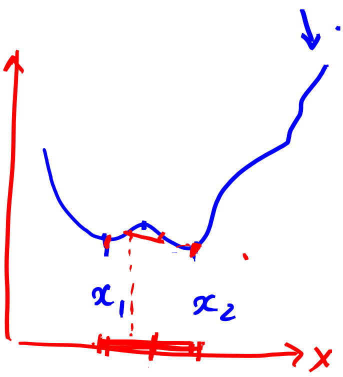
$\Rightarrow$ Set of optimal solutions to $\begin{array}{l} \min f(x) \\ \text{s.t. } x \in \Omega \end{array}$

where f: convex fn, Ω: convex set

is a convex set!

$\{x^* : f(x^*) \le f(x) \ \forall x \in \Omega\}$





Proof

Let S = set of optimal solutions to the problem.

$$x \in S \Rightarrow f(x) \le f(\tilde{x}) \ \forall \tilde{x} \in \Omega$$

(optimal)

Let $x_1, x_2 \in S$. $\boxed{\left.\begin{array}{l} f(x_1) \le f(\tilde{x}) \\ f(x_2) \le f(\tilde{x}) \end{array}\right\} \forall \tilde{x}}$

Claim: Any $\lambda x_1 + (1-\lambda) x_2$ will also be optimal.

f is convex

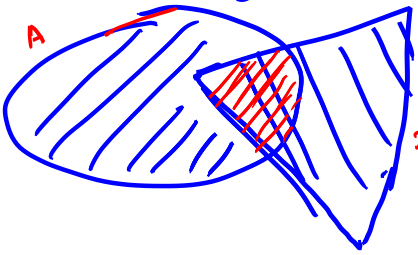
$$\Rightarrow \underbrace{f(\lambda x_1 + (1-\lambda) x_2)}_{\text{is also an optimum.}} \le \lambda f(x_1) + (1-\lambda) f(x_2)$$

$$\le \lambda f(\tilde{x}) + (1-\lambda) f(\tilde{x}) \quad \forall \tilde{x} \in \Omega$$

$$\le f(\tilde{x}). \quad \forall \tilde{x} \in \Omega$$

In general its not easy to prove that a set is convex. Thats why useful to know some operations which on sets which preserve convexity

- → Intersection of convex sets is convex



$A \cap B$ is convex.

$x \in A \cap B \Rightarrow x \in A,\ x \in B \Rightarrow$

$y \in A \cap B \Rightarrow y \in A,\ y \in B$

$\lambda x + (1-\lambda) y \overset{?}{\in} A \cap B$

eg. $\{x \mid Ax \le b\}$

$= \left\{x \;\middle|\; \begin{matrix} a_1^T x \le b_1 \\ \vdots \\ a_m^T x \le b_m \end{matrix} \right\}$

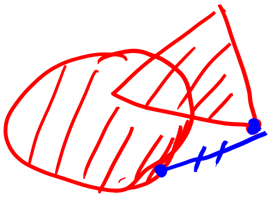
$A \cdot \begin{bmatrix} - a_1^T - \\ - a_2^T - \\ \vdots \\ - a_m^T - \end{bmatrix}$

$H_i : \{x \mid a_i^T x \le b_i\}$.

- Is union of 2 convex sets convex?

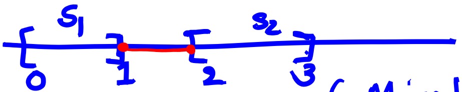
$S_1 \cup S_2 = \{x \mid x \in S_1 \text{ or } x \in S_2\}$ No.

S_1, S_2 convex.



$S_1 = [0, 1] \subseteq \mathbb{R}$.

$S_2 = [2, 3] \subseteq \mathbb{R}$.



- Sum of 2 convex sets. (Minkowski sum).

$S = \{x_1 + x_2 \mid x_1 \in S_1,\ x_2 \in S_2\} \subseteq \mathbb{R}^n$.

is convex if S_1 & S_2 are convex

CONVEX FUNCTIONS AND JENSEN'S INEQUALITY

f convex :

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

$$f\Big(\underbrace{\sum_{i=1}^{m} \lambda_i x_i}_{\int_x \lambda_x x \, dx}\Big) \leq \underbrace{\sum_{i=1}^{m} \lambda_i f(x_i)}_{\int \lambda_x f(x) \, dx}, \quad \boxed{\sum_{i=1}^{m} \lambda_i = 1 \quad \lambda_i \geq 0}$$

R.V $X = \{ x_i \text{ w.p. } \lambda_i \}$

$\hookrightarrow P(X = x_i)$

$$\sum_{i=1}^{m} \lambda_i x_i = E[X].$$

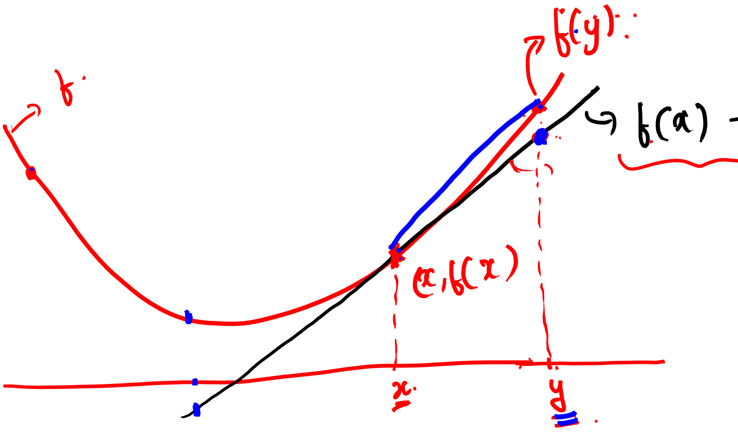
$$f(E[X]) \leq E[f(X)]$$

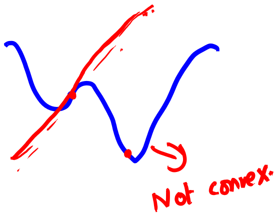
Characterization of convex functions

Consider $f: \mathbb{R}^n \to \mathbb{R}$, differentiable

f is a convex function $\iff$ dom(f) is a convex set &

(1) $$\boxed{f(y) \geq f(x) + (y-x)^T \nabla f(x)} \quad \forall x, y \in \text{dom}(f)$$

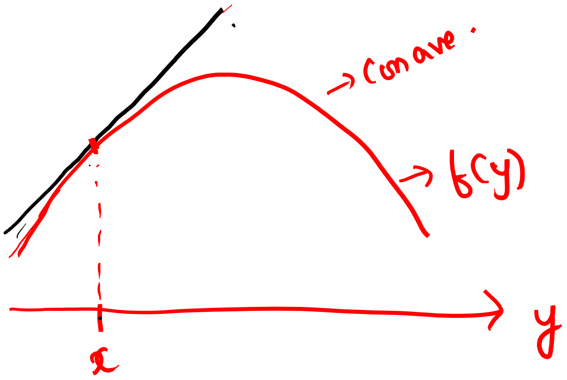




The function always lies above the first order approximation at any point.

If f is concave instead, we have dom(f) is a convex set &

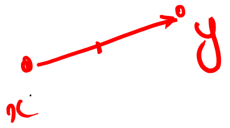
$$f(y) \leq f(x) + (y-x)^T \nabla f(x) \quad \forall x, y \in \text{dom}(f)$$



Proof of 1st order condition for convexity

Part I

Given f is convex & take any $x, y \in \text{dom}(f)$.



$$f(x + t(y-x)) = f(x(1-t) + ty)$$

$$\leq (1-t) f(x) + t f(y) \quad \forall t \in (0,1)$$

Divide by t on both sides

$$\frac{f(x + t(y-x))}{t} \leq f(y) + \left(\frac{1-t}{t}\right) f(x)$$

$$f(y) \geq f(x) + \frac{1}{t}\left[f(x + t(y-x)) - f(x) \right]$$

$$f(y) \geq f(x) + \lim_{t \to 0} \frac{1}{t}\left(f(x + t(y-x)) - f(x) \right)$$

Recall $\nabla f(x)^T d = \lim_{t \to 0} \frac{f(x+td) - f(x)}{t}$

(Directional derivative)

Part 2.

Suppose dom(f) is convex & $\forall y, x \in$ dom(f)

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

Take any $z = \lambda x + (1-\lambda) y$ $\quad \lambda \in (0,1)$

$$f(y) \geq f(z) + (y-z)' \nabla f(z) \quad \times 1-\lambda$$
$$f(x) \geq f(z) + (x-z)' \nabla f(z) \quad \times \lambda$$

Second-Order conditions.

Let $f: R^n \to R$ be twice differentiable.

Then f is convex $\iff$

Dom (f) is convex & $\nabla^2 f(x) \succeq 0$ at every $x \in dom(f)$

psd.

If $f: R \to R$ it just boils down to $f''(x) \geq 0$ at all $x \in dom(f)$.

- Increasing slope.

Examples

$\Rightarrow$ $f(x) = e^{ax}$ for any a.

$\Rightarrow$ $f(x) = \log x$, $x > 0$ (Concave)

$\Rightarrow$ Norms are all convex

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

$\Rightarrow$ $f(x) = \max(x_1, .., x_n)$

$f(\lambda x + (1-\lambda) y) = \max_i (\lambda x_i + (1-\lambda) y_i)$

$\leq \max_i \lambda x_i + \max_i (1-\lambda) y_i$

4) $f(x) = \sum_i c_i x_i$ $\Rightarrow$ Both convex & concave

Operations preserving convexity of functions

(Read Boyd's book for more operations/details)

1). $f(x) = f_1(x) + f_2(x)$. f_1, f_2 convex & defined on same domain

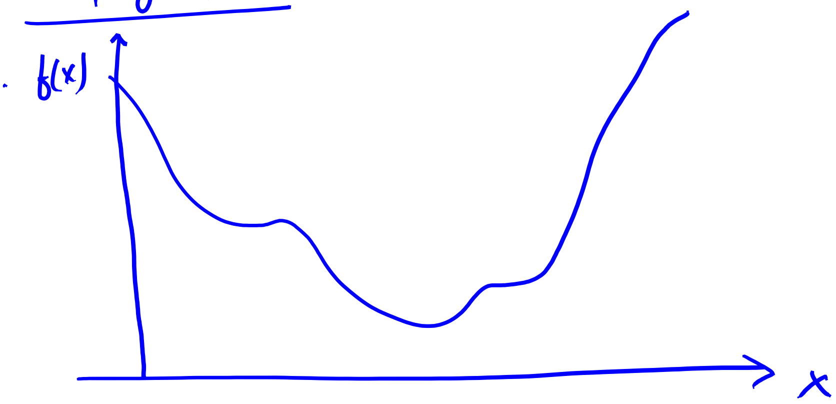
2) $f(x) = \max(f_1(x), f_2(x))$

3). $f(x) = \exp(g(x))$ with $g(x)$ convex

Connecting convex functions & convex sets.

Epigraph.

$f : \mathbb{R}^n \to \mathbb{R}$.



$Epi(f) = \{(x,t) \mid x \in dom(f),\ f(x) \leq t\}$

Region above f.

$Epi(f) \subseteq \mathbb{R}^{n+1}$

f is convex function iff $epi(f)$ is a convex set.

α-sublevel sets.

$S_\alpha = \{x \mid f(x) \leq \alpha\}$

If a function is convex, all its sublevel sets are convex

Proof: Take any $x, y \in S_\alpha$.

Quasi convex : A function with all α-sublevel sets convex.

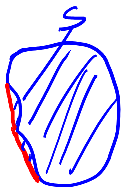
eg. $f(x) = x^3$

MORE ON CONVEX HULLS.

$$\mathrm{conv}(S) = \left\{ \sum_{i=1}^{m} \lambda_i x_i \;\middle|\; x_i \in S \ \forall i,\ \lambda_i \geq 0,\ \sum_{i=1}^{m} \lambda_i = 1 \right\}$$

- should contain every convex combination of its points.

Caratheodory's Theorem: Let $S \subseteq R^n$ & $\mathrm{conv}(S)$ is the convex hull of S. Every point in $\mathrm{conv}(S)$ can be written as a convex combination of at max $n+1$ points in S.



Proof.

Take any $x \in \mathrm{conv}(S)$ & let

$$x = \alpha_1 y_1 + \alpha_2 y_2 + \dots + \alpha_m y_m$$

$$\sum_{i=1}^{m} \alpha_i = 1 \qquad \alpha_i \geq 0.$$

If $m \leq n+1$ then we are done.

If $m > n+1$, try to show that at least one of $y_1, \dots y_m$ is redundant

Consider the system

$$\sum_{i=1}^{m} \beta_i y_i = 0$$
$$\sum^{m} \beta_i = 0$$

Find βs.

Check if this system has a solution other than 0.

$\therefore$ There exists β_i not all zero such that

$$\boxed{\sum_i \beta_i y_i = 0, \quad \sum_i \beta_i = 0}$$

We also have $x = \sum_{i=1}^{m} \alpha_i y_i$

Idea is to construct a point

$$\hat{x} = \sum_{i=1}^{m} (\alpha_i - \varepsilon \beta_i) y_i$$

for some $\varepsilon > 0$ so that one of $\alpha_i - \varepsilon \beta_i = 0$.

To do this, set $\varepsilon = \min_{i: \beta_i > 0} \left\{ \frac{\alpha_i}{\beta_i} \right\}$

Does it still give a convex combination?

- usefulness of Convex hulls.

$$\min \sum_{i=1}^{m} c_i x_i \qquad \Longleftrightarrow \qquad \min \sum c_i x_i$$
$$\text{s.t} \quad x \in S \qquad\qquad\qquad \text{s.t} \quad x \in \text{conv}(S).$$

New optimization problem over conv(S) is not necessarily easier. But in some cases if conv(S) can be easily represented, then the problem can be solved easily.

Proof.

Out of the 2 sets S & $\text{Conv}(S)$ which is larger?

$$S \subseteq \text{conv}(S).$$

so
$$\begin{array}{ll} \min & f(x) \\ \text{st} & x \in S \end{array} \geq \begin{array}{ll} \min & f(x) \\ \text{st} & x \in \text{conv}(S) \end{array}$$

We need to show

$$\min_{x \in \text{Conv}(S)} \sum_i c_i x_i \geq \min_{x \in S} \sum_i c_i x_i$$

Then we are done.

Let $\tilde{x} = \underset{x \in \text{conv}(S)}{\text{argmin}} \sum_i c_i x_i$

$$\tilde{x} = \sum_{j=1}^{n+1} \lambda_j y_j$$

$$f(\tilde{x}) = \sum_i c_i \tilde{x}_i = \sum_i c_i \left(\sum_{j=1}^{n+1} \lambda_j y_{ji} \right)$$

$$y_j = \begin{bmatrix} y_{j1} \\ \vdots \\ y_{jn} \end{bmatrix}$$

$$= \sum_{j=1}^{n+1} \lambda_j \left(\sum_i c_i y_{ji} \right)$$

$$\geq \sum_{j=1}^{n+1} \lambda_j \left(\min_{z \in S} c^T z \right)$$

$$= \min_{z \in S} c^T z$$

Are convex problems easy?

Not Always!

To see this, take any optimization problem

$$\boxed{\begin{aligned} \min\quad & f(x) \\ \text{s.t}\quad & x \in \Omega \end{aligned}} \iff \boxed{\begin{aligned} \min\quad & t \\ \text{s.t}\quad & t \geq f(x) \\ & x \in \Omega \end{aligned}}$$

linear optimization prob: over (x, t)

$$\min \quad t$$

$$\text{s.t} \ (x, t) \in \operatorname{conv}\left\{ \begin{array}{l} x \in \Omega, \\ f(x) \leq t \end{array} \right\}$$

This is a problem where the constraint set is convex. Objective is linear & hence convex. If convex optimization problems could be solved easily in general, every optimization problem can be solved easily!!!

A convex optimization problem usually refers to:

$$\min f(x)$$

$$\text{s.t.} \quad g_i(x) \leq 0 \quad \forall i = 1 \ldots m$$

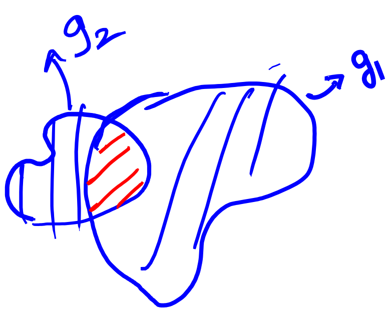
$$h_j(x) = 0 \quad \forall j = 1 \ldots k$$

where f, g_i, h_j are all convex functions. (so h_j has to be affine)

Ω is a convex set if representation is given in this way.

But if Ω is convex set, doesnt mean the functions g_i, h_j have to be convex.

eg. Intersection of non convex regions



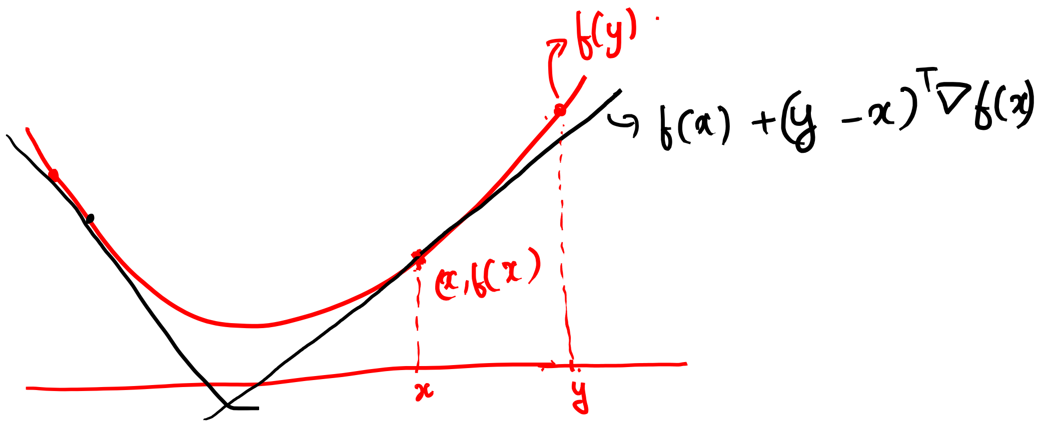
2) $\Omega = \{x \in R \mid x^3 \leq 0\}$

Characterization of convex functions

Consider $f: R^n \to R$, differentiable

f is a convex function $\iff$ dom(f) is a convex set &

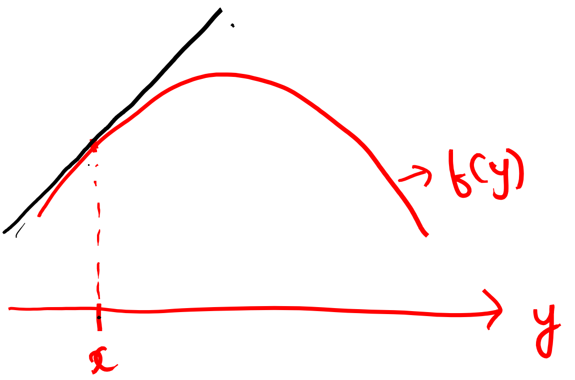
$$(1) \quad \boxed{f(y) \geq f(x) + (y-x)^T \nabla f(x) \quad \forall x, y \in \text{dom}(f)}$$



The function always lies above the first order approximation at any point.

If f is concave instead, we have dom(f) is a convex set &

$$f(y) \leq f(x) + (y-x)^T \nabla f(x) \quad \forall x, y \in \text{dom}(f)$$



$\min_{x \in R^n} f(x)$: f is convex

$$\underbrace{f(y) \geq f(x) + \nabla f(x)^T (y-x)}$$

$$\nabla f(x^*) = 0 \Rightarrow \underbrace{f(y) \geq f(x^*) \quad \forall y \in R^n}$$

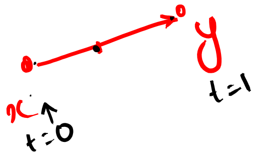
x^* is a global minimum.

Proof of 1st order condition for convexity

Part 1

Given f is convex & take any $x, y \in \operatorname{dom}(f)$.

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



$$f(\underbrace{x + t(y-x)}) = f(x(1-t) + ty)$$

$$\leq (1-t) f(x) + t f(y) \quad \forall t \in (0,1)$$

Divide by t on both sides

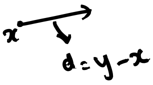
$$\frac{f(x + t(y-x))}{t} \leq f(y) + \left(\frac{1-t}{t}\right) f(x)$$

$$\left(\frac{1}{t} - 1\right) f(x)$$

$$f(y) \geq f(x) + \frac{1}{t}\left[f(x + t(y-x)) - f(x) \right] \quad \forall t \in (0,1)$$

$$f(y) \geq f(x) + \lim_{t \to 0} \frac{1}{t}\left(f(x + t(\overset{d}{\overbrace{y-x}})) - f(x) \right)$$

$$\boxed{f(y) \geq f(x) + \nabla f(x)^T d}$$



Recall $\nabla f(x)^T d = \lim_{t \to 0} \dfrac{f(x + td) - f(x)}{t}$

(Directional derivative)

Part 2.

Suppose dom(f) is convex & $\forall y, x \in$ dom(f)

$$\boxed{\Rightarrow f(y) \geq f(x) + (y-x)^T \nabla f(x)}$$

Take any $z = \lambda x + (1-\lambda) y \quad \lambda \in (0,1)$

Want: $\boxed{f(z) \leq \lambda f(x) + (1-\lambda) f(y).}$

(y,z) $\quad f(y) \geq f(z) + (y-z)' \nabla f(z) \rightarrow \times\, 1-\lambda$

(x,z) $\quad f(x) \geq f(z) + (x-z)' \nabla f(z) \rightarrow \times\, \lambda$

$$\lambda f(x) + (1-\lambda) f(y) \geq \underbrace{\lambda f(z) + (1-\lambda) f(z)}_{f(z)}$$

$$+ (y-z)^T \nabla f(z) (1-\lambda)$$

$$+ \lambda (x-z)^T \nabla f(z)$$

$$\lambda x^T \nabla f(z) - \lambda z^T \nabla f(z)$$

$$+ y^T \nabla f(z) - \lambda y^T \nabla f(z)$$

$$- z^T \nabla f(z) + \lambda z^T \nabla f(z)$$

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

$$\underbrace{(\lambda x + (1-\lambda) y - z)^T \nabla f(z)}_{= 0.}$$

Second-Order conditions.

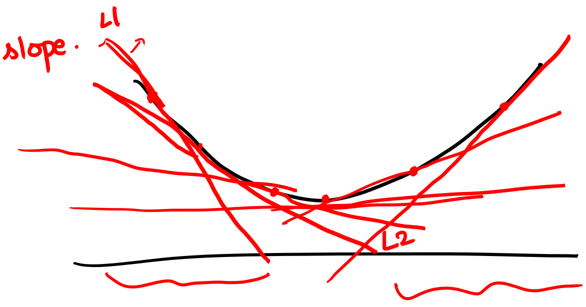
Let $f: R^n \to R$ be twice differentiable.

Then f is convex $\iff$ Dom(f) is convex & $\nabla^2 f(x) \succeq 0$ at every $x \in \text{dom}(f)$

$\nabla^2 f(x)$ ↓ matrix; $\succeq 0$ ↳ psd.

If $f: R \to R$ it just boils down to $f''(x) \geq 0$ at all $x \in \text{dom}(f)$.

- Increasing slope.



Examples

⇒ $f(x) = e^{ax}$ for any a.

$f'(x) = ae^{ax}$

$f''(x) = \underbrace{a^2}_{\geq 0} \underbrace{e^{ax}}_{\geq 0} \geq 0 \quad \forall x.$

⇒ $f(x) = \log x$, $x > 0$ (Concave)

$f'(x) = \frac{1}{x}$ $\quad f''(x) = \frac{-1}{x^2} < 0$

⇒ Norms are all convex

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

⇒ Triangle inequality.

$x \in R^n$.

⇒ $f(x) = \max(x_1, .., x_n)$

$f(\lambda x + (1-\lambda)y) = \max_i(\lambda x_i + (1-\lambda)y_i)$

$\leq \max_i \lambda x_i + \max_i (1-\lambda)y_i$

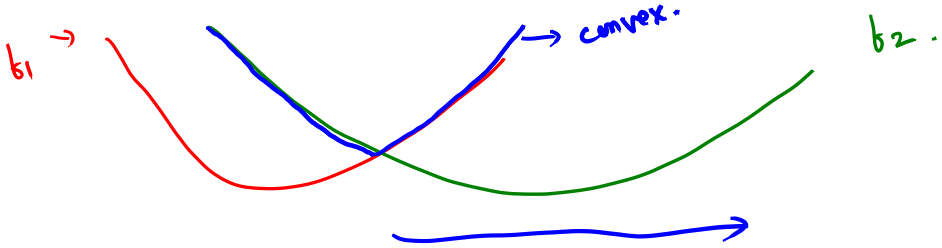
4) $f(x) = \sum_i c_i x_i$ ⇒ Both convex & concave

Operations preserving convexity of functions

(Read Boyd's book for more operations/details)

1). $f(x) = f_1(x) + f_2(x)$. $\quad f_1, f_2$ convex & defined on same domain

2) $f(x) = \max(f_1(x), f_2(x))$.



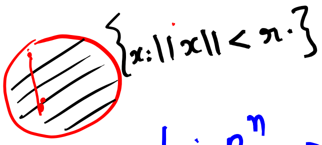
3). $f(x) = \exp(g(x))$ $\quad$ with $g(x)$ convex

$f'(x) = e^{g(x)} \cdot g'(x)$

$$f''(x) = \underbrace{g''(x)\, e^{g(x)}}_{\geq 0} + \underbrace{(g'(x))^2 e^{g(x)}}_{\geq 0}$$

≥ 0.

Connecting convex functions & convex sets.



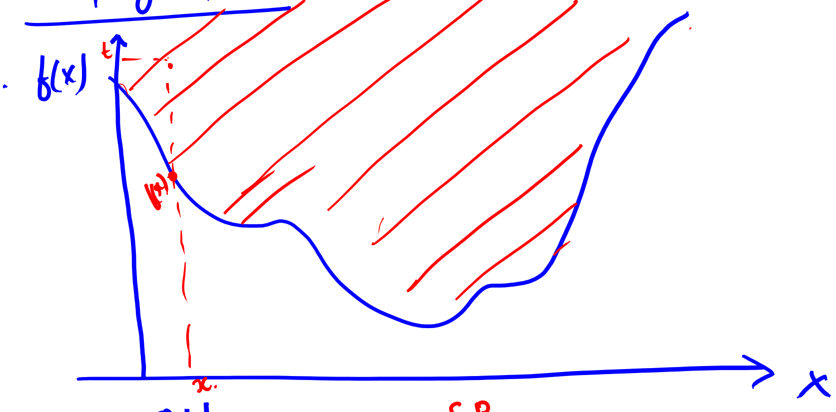
Epigraph.

$f : \mathbb{R}^n \to \mathbb{R}$.

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

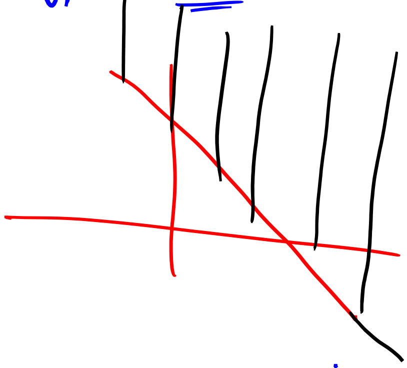
$(x, f(x))$.

(x, t)
$t \geq f(x)$



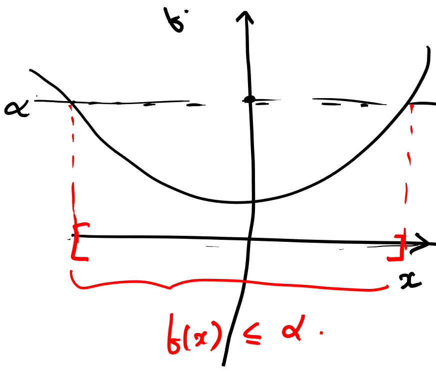
$$\text{Epi}(f) = \{ (\underbrace{x}_{\in \mathbb{R}^{n+1}}, \underbrace{t}_{\in \mathbb{R}}) \mid x \in \text{dom}(f),\ f(x) \leq t \}$$

Region above f.



$\text{Epi}(f) \subseteq \mathbb{R}^{n+1}$

f is convex function iff epi(f) is a convex set.



α-sublevel sets.

$$S_\alpha = \{ x \mid f(x) \leq \alpha \}$$

If a function is convex, $\Rightarrow$ all its sublevel sets are convex.

Proof: Take any $x, y \in S_\alpha$.

f is convex. Does $\lambda x + (1-\lambda) y \in S_\alpha$?

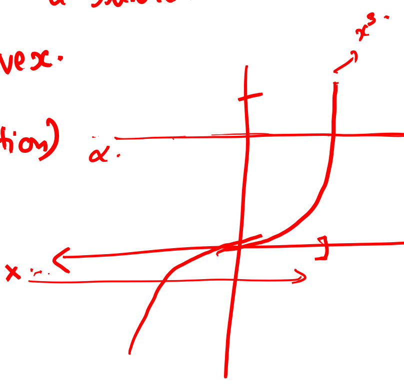
$$f(\lambda x + (1-\lambda) y) \leq \lambda \underbrace{f(x)}_{\leq \alpha} + (1-\lambda) \underbrace{f(y)}_{\leq \alpha}$$

$$\leq \alpha.$$

Quasi convex: A function with all α-sublevel sets convex.

eg. $f(x) = x^3$ (not convex function)

$S_\alpha = \{ x : x^3 \leq \alpha \}$ $\to$ convex set.



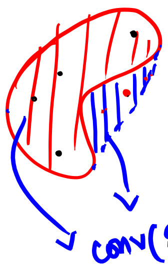
MORE ON CONVEX HULLS.

$$\text{conv}(S) = \left\{ \sum_{i=1}^{\boxed{m}} \lambda_i x_i \;\middle|\; x_i \in S \ \forall i,\ \lambda_i \geq 0,\ \sum_{i=1}^{m} \lambda_i = 1 \right\}$$

- should contain every convex combination of its points.

Caratheodory's Theorem: Let $S \subseteq R^n$ & conv(S) is the convex hull of S. Every point in conv(S) can be written as a convex combination of at most n+1 points in S.

S



conv(S).

$x \in R^2$

$S \subseteq R^2$.

n+2 points in S.

$x_1 \ \ldots \ x_{n+2}$

$\sum_i \lambda_i x_i \in \text{conv}(S)$.

$x = \underbrace{\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + \lambda_4 x_4}_{\sum \lambda_i = 1}$

$\downarrow$

conv(S)

Proof.

Take any $x \in \text{conv}(S)$ & let $S \subseteq R^n$.

$$x = \alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_m y_m$$

$y_1, \ldots y_m \in S$.

$$\sum_{i=1}^{m} \alpha_i = 1 \qquad \alpha_i \geq 0.$$

If $m \leq n+1$ then we are done.

If $\boxed{m > n+1}$, try to show that at least one of $y_1, \dots y_m$ is redundant

Consider the system

$\beta_1, \dots \beta_m \in \mathbb{R}$.

$$\sum_{i=1}^{m} \beta_i y_i = 0$$

$$\sum_{i=1}^{m} \beta_i = 0$$

Find βs.

Check if this system has a solution other than 0.

If $\exists$ solution

$$x = \sum_{i=1}^{m} \alpha_i y_i$$

$$\boxed{\sum_i (\alpha_i - \varepsilon \beta_i)} y_i = x$$

$$= \sum_i \alpha_i y_i - \varepsilon \underbrace{\sum_i \beta_i y_i}_{=0}$$

$$\begin{array}{c} \mathbb{R}^n \\ 1 \end{array} \left\{ \begin{bmatrix} y_1 & \dots & y_m \\ 1 & \dots & 1 \end{bmatrix} \right. \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = 0.$$

$A \in \mathbb{R}^{(n+1) \times m}$.

$$\boxed{A\beta = 0}$$

$m > n+1$. $\mathcal{N}(A)$ has non-zero vectors.

$\beta_1, \dots \beta_m$

$$\sum \alpha_i = 1,$$

want

$$\sum_i (\alpha_i - \varepsilon \beta_i) = 1$$

$$= \underbrace{\sum \alpha_i}_{1} - \varepsilon \underbrace{\sum \beta_i}_{0}$$

$\therefore$ There exists β, not all zero such that

$$\boxed{\sum_i \beta_i y_i = 0, \quad \sum_i \beta_i = 0}$$

we also have $x = \sum_{i=1}^{m} \alpha_i y_i$

Idea is to construct a representation

$$x = \sum_{i=1}^{m} (\alpha_i - \varepsilon \beta_i) y_i$$

for some $\varepsilon > 0$ so that one of $\alpha_i - \varepsilon \beta_i = 0$.

To do this, set $\varepsilon = \min_{i: \beta_i > 0} \left\{ \frac{\alpha_i}{\beta_i} \right\}$

i^*: $\beta_{i^*} > 0$, $\varepsilon = \frac{\alpha_{i^*}}{\beta_{i^*}}$

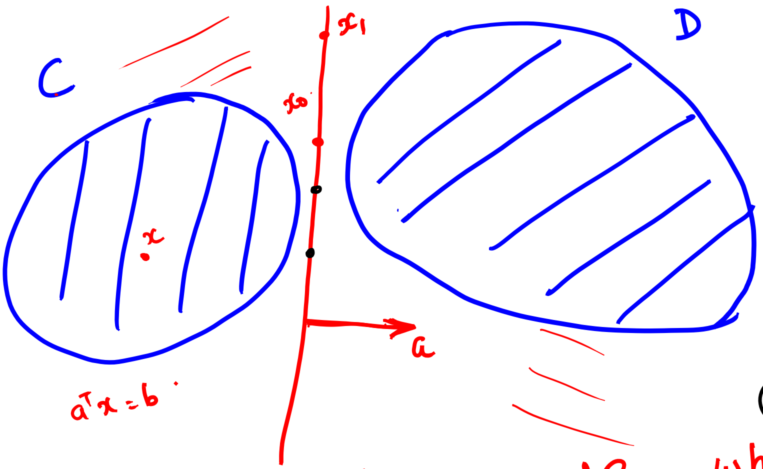
$$\alpha_{i^*} - \varepsilon \beta_{i^*} = 0$$

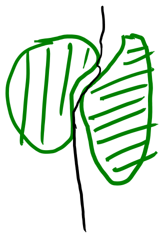
Separating Hyperplane Theorem.

Let C & $\mathcal{D}$ be convex sets in R^n that don't intersect. Then $\exists\, a \in R^n$ & $b \in R$ $(a \neq 0)$ such that $a^T x \leq b \quad \forall x \in C$ and

$a^T x \geq b \quad \forall x \in \mathcal{D}$.

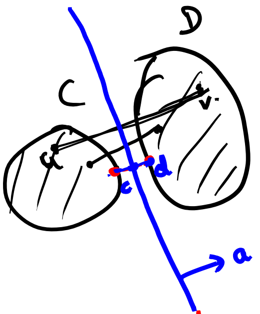
$a^T x_0 = a^T x_1 = b$.





We will prove it for the case where C & $\mathcal{D}$ are closed convex & bounded sets.

$\Rightarrow \text{dist}(C, \mathcal{D}) = \inf\limits_{\substack{u \in C \\ v \in \mathcal{D}}} \|u - v\|_2$



- Continuous objective.
- Feasible set is closed, bounded.

Weirstrass.

Infimum is achieved.

& optimal value is positive.

$\exists c \in C, d \in D$.

Let $c \in C$ & $d \in \mathcal{D}$ achieve the optimum.

Claim: The hyperplane $a^T x = b$ with $\boxed{a = d - c}$ & $b = \dfrac{\|d\|^2 - \|c\|^2}{2}$ separates C & $\mathcal{D}$.

i.e $a^T x - b \geq 0 \quad \forall x \in \mathcal{D}$

$a^T x - b \leq 0 \quad \forall x \in C$

For midpoint of c & d, $x = \frac{c+d}{2}$

$$a^T x = \frac{(d-c)^T (d+c)}{2} = \frac{d^T d + d^T c - c^T d - c^T c}{2}$$

$$= \underbrace{\frac{\|d\|^2 - \|c\|^2}{2}}_{} \leftarrow b$$

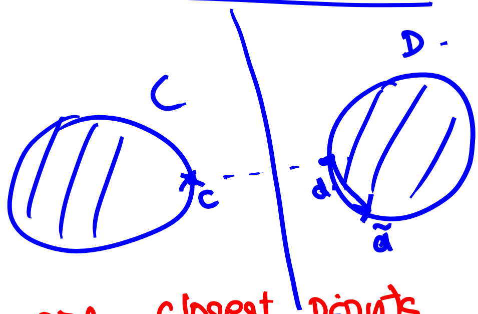
We will show $\boxed{a^T x \geq b \quad \forall x \in D}$.

($a^T x \leq b \quad \forall x \in C$ will follow along similar lines)

$a^T \tilde{d} < b$.

Suppose not. $\exists \tilde{d} \in D$ s.t $a^T \tilde{d} - b < 0$

$$\boxed{\Rightarrow \underbrace{(d-c)^T}_{a} \tilde{d} < \underbrace{\frac{\|d\|^2 - \|c\|^2}{2}}_{b}}$$



Fact: since c & d are closest points, if you minimize $\boxed{g(x) = \|x - c\|^2}$ over $x \in D$, optimum should be achieved at point d.

Idea is to show that $\tilde{d} - d$ is a descent direction for $\boxed{\min_{x \in D} \|x - c\|^2}$ at $\underline{d}$.

d.

$\boxed{\text{Prove. } (\tilde{d} - d) \text{ is a descent direction for } g(x) = \|x - c\|^2 \text{ at } x = d.}$