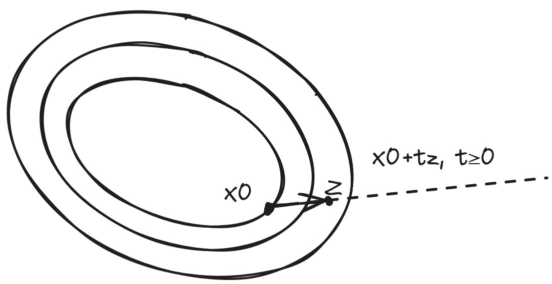
Lecture 18 Duality

Note by Samion Suwito on 20/3/25

Last Time: Convex Sets & Functions

Some characteristics of convexity of functions:

1. First order: $f(x) \geq f(y) + \nabla f(y)^\top (x - y)$
2. 2nd order: $\nabla^2 f(x) \geq 0 \quad \forall x \in \mathrm{dom} f$
3. Convexity along a line. (below is an image of the level sets)



$t \geq 0 \mapsto f(x_0 + tz)$ is convex $\forall t \geq 0$ is sufficiently small s.t. $x_0 + tz \in \mathrm{dom} f$
Example: $X \mapsto \log\det X^{-1}$ is convex on $X \succ 0$
Show: $X_0 + tZ$, $X_0 \succ 0$ then choose $Z \in \mathbb{S}^{n\times n}$

$$\begin{aligned}
\log\det(X_0 + tZ)^{-1} &= -\log\det(X_0 + tZ) \\
&= -\log\det(X_0^{1/2}(I + tX^{-1/2} Z X_0^{-1/2}) X_0^{1/2}) \\
&= -\log\det(X_0) - \log\det(I + tX_0^{-1/2} Z X_0^{1/2}) \\
&= -\log\det(X_0) - \sum_{i=1}^{n} \log(1 + t\lambda_i(X_0^{-1/2} Z X_0^{-1/2}))
\end{aligned}$$

Determinant takes eigenvalues values and inverse makes eigenvalues inverse. Since determinant is the product of eigenvalues we can then separate it into the sum. $\lambda_i(X_0^{-1/2} Z X_0^{-1/2})$ is the ith eigenvalue not multiplication. We know $-\log(\cdot)$ is convex on $\mathbb{R}_{>0}$ meaning that we can say the last sum is a sum of convex functions which are therefore convex in t.

$\log\det$ useful in entropy.
We check only for t small enough because convexity is a local property.

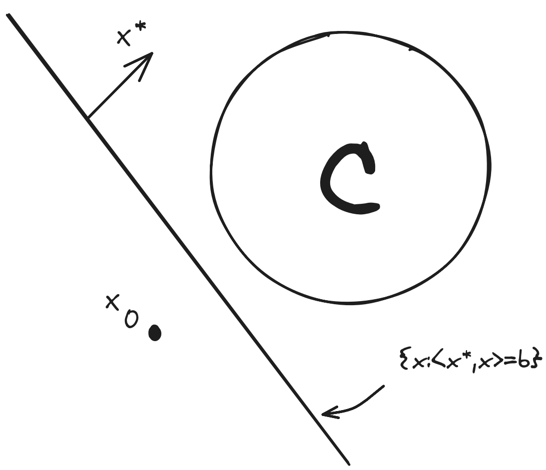
Convex Duality

Term **duality** comes from **"dual space"** $X^* =$ all linear functions on a vector space $X = \mathbb{R}^n$ for our case always finite dimensional. $X^* \equiv \mathbb{R}^n$ as we saw that any linear function can be represented as a vector.

Separation Theorem

Most fundamental duality result
Separation Theorem: Let $C \subset X$ be closed, convex set and let $x_0 \notin C$. There exists a function $x^* \in X^*$ (star is not optimal its duality) and a $\delta > 0$ s.t.

$$\underbrace{\langle x^*, x_0 \rangle}_{x^*(x_0)} + \delta \leq \underbrace{\langle x^*, x \rangle}_{x^*(x)} \forall x \in C$$



δ is the margin. Essentially we can find a hyperplane that separates C and any x_0. Also the direction of the inequality is not relevant and is just the fact that there's a separation.

$f^* : x \to \mathbb{R}$ is linear then $x^* = (f(e_1) \ \cdots \ f(e_n))^\top$

Example 1: Farkas' Lemma

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Exactly one of the following is true:

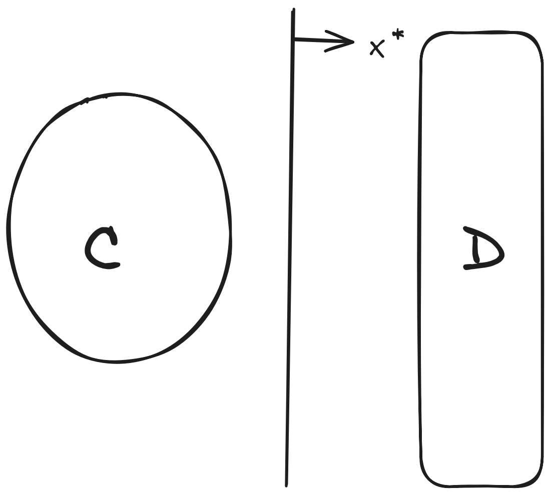
1. $\exists x \geq 0 \text{ s.t. } Ax = b$
2. $\exists y \text{ s.t. } A^\top y \geq 0 \text{ and } b^\top y < 0$
   We can certify that there is no nonnegative solution for the first statement using the second statement as a certificate.

Let $\mathcal{A} =$ conic hull of $\{a_1, \ldots, a_n\}$ (columns of A) = $\{Ax, x \geq 0\}$

1. $\exists x \geq 0$ s.t. $Ax = b \iff b \in \mathcal{A}$
2. $b \notin \mathcal{A} \implies \exists x^*$ s.t. $\langle y, b\rangle < \langle y, a\rangle \quad \forall a \in \mathcal{A}$ Since $0 \in \mathcal{A} \implies b^\top y < 0$ and $b^\top y < \sum x_i \langle a_i, y\rangle \quad \forall x \geq 0 \iff a_i^\top y \geq 0 \quad \forall i$ We can claim the last statement As we know that $b^\top y$ is strictly less than 0 meaning that the sum on the left must be greater than or equal to 0

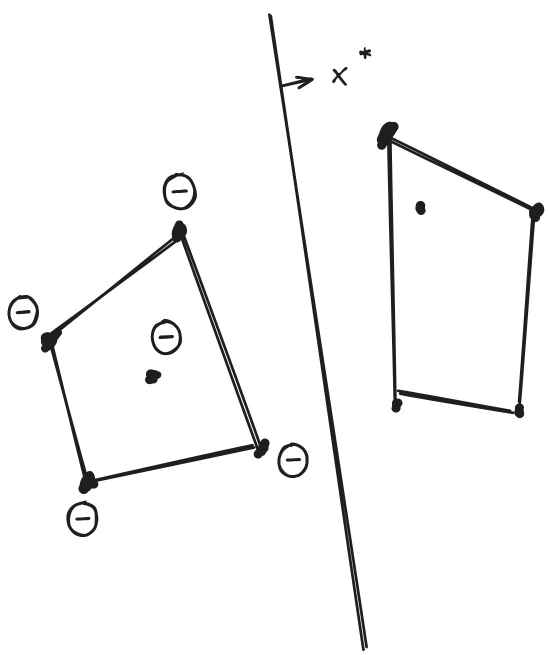
Example 2

C, D are disjoint closed convex sets and bounded.
Claim: There exists an $x^* \in X^*, \delta > 0$ s.t. $\langle x^*, c\rangle + \delta \leq \langle x^*, d\rangle$



Let's say you have two large high dimensional data and you need to separate the data between C and D which we could do with a lookup table, instead we could just use x^*

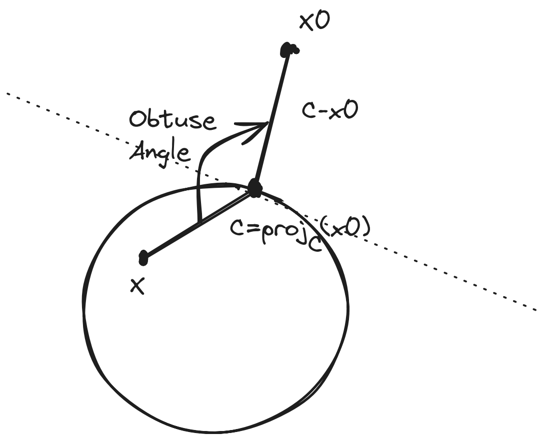
This is the idea behind SVM, that separates the convex hull of two sets of training data.



Define $C - D = \{c - d : c \in C, d \in D\}$ Which is closed and convex. $0 \notin C - D$ since the sets are disjoint. $\implies \exists x^*, \delta > 0$ s.t. $\langle x^*, y\rangle + \delta \leq \langle x^*, 0\rangle \forall y \in C - D$ and using the linearity of inner product we can get the original statement.

Projections

Theorem: Projections onto convex sets. For C closed, convex set. $x_0 \in X$. $\exists$ unique point $\text{proj}_C(x_0) \in C \text{ s.t.} \|\text{proj}_C(x_0) - x_0\|_2 \leq \|x - x_0\|_2 \quad \forall x \in C$



Moreover $c = \text{proj}_C(x_0) \iff c \in C$ and $\langle c - x_0, c - x\rangle \leq 0 \quad \forall x \in C$
Example: If C is a subspace (which is closed and convex)

$$\langle \text{proj}_C(x_0) - x_0, \text{proj}_C(x_0) - x\rangle \leq 0 \quad \forall x \in C$$

$$\begin{aligned} &\iff \langle \text{proj}_C(x_0) - x_0, v\rangle \leq 0 \quad v \in C \\ &\iff \langle \text{proj}_C(x_0) - x_0, v\rangle = 0 \quad \forall v \in C \end{aligned}$$

For set C define support function:
$h_C(x^*) = \sup_{x \in C}\langle x^*, x\rangle$(sup = max)
$\implies$ "Dual" characterisation of convex set.
Theorem: For closed convex C,

$$C = \bigcap_{x^* \in X^*} \{x : \langle x^*, x\rangle \leq h_C(x^*)\}$$

Any convex set can be written a s the intersection of half spaces.
Proof:

$$\begin{aligned} C &\subset \{x : \langle x^*, x\rangle \leq h_c(x^*)\} \\ \implies C &\subset \bigcap_{x^* \in X^*} \{x : \langle x^*, x\rangle \leq h_c(x^*)\} \end{aligned}$$

Now suppose $x_0 \notin C \implies \exists x^*$ which seperates x_0 from C i.e.
$\langle x^*, x\rangle < \langle x^*, x_0\rangle \quad \forall x \in C$
$\implies h_C(x^*) < \langle x^*, x_0\rangle \implies x_0 \notin \{x : \langle x^*, x\rangle \leq h_c(x^*)\}$

$$\text{\xymatrix}A$$