553.665 Introduction to Convexity, Fall 2022
Section 1

Hongyu Cheng

hongyucheng@jhu.edu

September 2, 2022

Note: In the session I start from Section 4: Convex sets.

1 Recall the definitions

Let $X \subseteq \mathbb{R}^d$.

1. X is a convex set if $\forall \mathbf{x}, \mathbf{y} \in X, \ \forall \lambda \in [0,1], \lambda \mathbf{x} + (1-\lambda)\mathbf{y} \in X$.
2. $\operatorname{conv}(X)$ is the smallest convex set containing X.
3. An affine transformation is a function $T : \mathbb{R}^d \to \mathbb{R}^m$ of the form $T(x) = A\mathbf{x} + \mathbf{b}$, where A is an $m \times d$ matrix and $\mathbf{b} \in \mathbb{R}^m$.
4. X is a cone if $\forall \mathbf{x}, \mathbf{y} \in X, \forall \lambda, \gamma \geq 0, \ \lambda \mathbf{x} + \gamma \mathbf{y} \in X$.
5. X is an affine set if $\forall \lambda, \gamma \in \mathbb{R}$ s.t. $\lambda + \gamma = 1, \ \lambda \mathbf{x} + \gamma \mathbf{y} \in X$.
6. X is a linear subspace if $\forall \lambda, \gamma \in \mathbb{R}, \ \lambda \mathbf{x} + \gamma \mathbf{y} \in X$.

2 A brief review of real analysis and basic topology

Exercise 1. Let A_i, $i \in \mathcal{I}$ be an arbitrary family of sets with $\mathcal{I} \neq \varnothing$, and X is a set. Then,

1. $\forall \alpha \in \mathcal{I}$,
$$\bigcap_{i \in \mathcal{I}} A_i \subseteq A_\alpha \subseteq \bigcup_{i \in \mathcal{I}} A_i$$
2. (Distributive property)
$$X \cap \left(\bigcup_{i \in \mathcal{I}} A_i \right) = \bigcup_{i \in \mathcal{I}} (X \cap A_i)$$
$$X \cup \left(\bigcap_{i \in \mathcal{I}} A_i \right) = \bigcap_{i \in \mathcal{I}} (X \cup A_i)$$

3. (De Morgan's law)

$$X \backslash \left(\bigcap_{i \in \mathcal{I}} A_i \right) = \bigcup_{i \in \mathcal{I}} (X \backslash A_i)$$
$$X \backslash \left(\bigcup_{i \in \mathcal{I}} A_i \right) = \bigcap_{i \in \mathcal{I}} (X \backslash A_i)$$

Proof. The first equality in 2 holds since

$$\begin{aligned} x \in X \cap \left(\bigcup_{i \in \mathcal{I}} A_i \right) &\iff x \in X \text{ and } x \in \bigcup_{i \in \mathcal{I}} A_i \\ &\iff x \in X \text{ and } \exists \alpha \in \mathcal{I} \text{ such that } x \in A_\alpha \\ &\iff \exists \alpha \in \mathcal{I} \text{ such that } x \in X \cap A_\alpha \\ &\iff x \in \bigcup_{i \in \mathcal{I}} (X \cap A_i). \end{aligned}$$

The first equality in 3 holds since

$$\begin{aligned} x \in X \backslash \left(\bigcap_{i \in \mathcal{I}} A_i \right) &\iff x \in X \text{ and } x \notin \bigcap_{i \in \mathcal{I}} A_i \\ &\iff x \in X \text{ and } \exists \alpha \in \mathcal{I} \text{ such that } x \notin A_\alpha \\ &\iff \exists \alpha \in \mathcal{I} \text{ such that } x \in X \backslash A_\alpha \\ &\iff x \in \bigcup_{i \in \mathcal{I}} (X \backslash A_i). \end{aligned}$$

The proofs of the remaining arguments are very similar. □

Definition 1. For any set $X \subseteq \mathbb{R}^d$,

1. $\text{cl}(X)$ is the smallest closed set containing X (the intersection of all closed sets of $\mathbb{R}^d$ containing X).
2. $\text{int}(X)$ is the largest open set contained inside X (the union of all open sets of $\mathbb{R}^d$ contained in X).
3. $\text{bd}(X) := \text{cl}(X) \backslash \text{int}(X)$.

Exercise 2. Let $X \subseteq \mathbb{R}^d$, then $x \in \text{bd}(X) \implies \forall r > 0, B(x, r) \cap X \neq \varnothing$ and $B(x, r) \cap (\mathbb{R}^d \backslash X) \neq \varnothing$.

Proof. Note that $\text{bd}(X)$ is defined to be $\text{cl}(X) \backslash \text{int}(X)$. So if there exists $r > 0$ such that $B(x, r) \subseteq X$, then $\text{int}(B(x, r)$ is an open set contained in X that contains x, which contradicts with the definition.

Similarly, if there exists $r > 0$ such that $B(x, r) \subseteq \mathbb{R}^d \backslash X$, then $\mathbb{R}^d \backslash \text{int}(B(x, r))$ is a closed set contains X, but not x, again contradicting the definition.

Thus, for every $r > 0$, $B(x, r)$ must contain a point in X and a point in $\mathbb{R}^d \backslash X$. □

Remark 1. The reverse direction is also true.

Exercise 3. Let $X \subseteq \mathbb{R}^d$, then $\text{int}(A) = \mathbb{R}^d \backslash \text{cl}(\mathbb{R}^d \backslash A)$.

Proof. $\forall x \in \text{int}(A) \implies x \notin \text{cl}(\mathbb{R}^d \backslash A) \implies x \in \mathbb{R}^d \backslash \text{cl}(\mathbb{R}^d \backslash A)$.

For the reverse inclusion,

$$\begin{aligned} \forall x \in \mathbb{R}^d \backslash \text{cl}(\mathbb{R}^d \backslash A) &\implies \exists r > 0 \text{ s.t. } \text{int}(B(x,r)) \subseteq \mathbb{R}^d \backslash \text{cl}(\mathbb{R}^d \backslash A) \\ &\implies \text{int}(B(x,r)) \cap \mathbb{R}^d \backslash A = \varnothing \\ &\implies \text{int}(B(x,r)) \subseteq A \\ &\implies x \in \text{int}(A). \end{aligned}$$

□

Remark 2. Similarly, $\text{cl}(A) = \mathbb{R}^d \backslash \text{int}(\mathbb{R}^d \backslash A)$.

3 A brief review of linear algebra

Definition 2. Let $A \in R^{m \times n}$, then there are several subspaces naturally associated to it:

- $\text{null}(A) = \{x \in \mathbb{R}^n : Ax = 0\}$,
- $\text{range}(A) = \{y \in \mathbb{R}^m : \exists x \in \mathbb{R}^n \text{ such that } Ax = y\}$.

Exercise 4. Let $A \in \mathbb{R}^{m \times n}$, then $\text{null}(A)$ is a subspace of $\mathbb{R}^n$.

Proof. 1. $A0 = 0 \implies 0 \in \text{null}(A)$ and $\text{null}(A)$ is nonempty.

2. $\forall x, y \in \text{null}(A)$, $A(x + y) = Ax + Ay = 0 + 0 = 0 \implies x + y \in \text{null}(A)$.

3. $\forall x \in \text{null}(A),\ \forall \lambda \in \mathbb{R},\ A(\lambda x) = \lambda Ax = \lambda 0 = 0 \implies \lambda x \in \text{null}(A)$.

□

Exercise 5. Let $A \in \mathbb{R}^{m \times n}$, then $\text{null}(A)$ is orthogonal to $\text{range}(A^\top)$.

Proof. We just need to show that $\forall x \in \text{null}(A),\ \forall y \in \text{range}(A^\top),\ \langle x, y \rangle = 0$. By definition of $\text{range}(A^\top)$, there exists $z \in \mathbb{R}^n$ such that $A^\top z = y$. Therefore,

$$\langle x, y \rangle = \langle x, A^\top z \rangle = \langle Ax, z \rangle = \langle 0, z \rangle = 0.$$

□

Theorem 1 (Rank-nullity theorem)**.** *Let $A \in \mathbb{R}^{m \times n}$, then dim range$(A)$ + dim null$(A) = n$.*

Theorem 2. *Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, then there exists a matrix $S \in \mathbb{R}^{n \times n}$ such that $S^\top S = I$ and*

$$A = S \Lambda S^\top,$$

where Λ is the diagonal matrix with the diagonal entries equal to the eigenvalues of A.

Remark 3. Both directions would hold if A is *normal*, i.e., $A^\top A = AA^\top$.

Exercise 6. Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. The rank of A is defined to be the dimension of the range of A. Prove that this is equal to the number of nonzero eigenvalues of A.

4 Convex sets

Theorem 3 (Operations that preserve convexity)**.** *The following are all true.*

1. *Let $X_i, i \in \mathcal{I}$ be an arbitrary family of convex sets. Then $\bigcap_{i\in\mathcal{I}} X_i$ is a convex set.*
2. *Let X be a convex set and $\alpha \in \mathbb{R}$, then αX is a convex set.*
3. *Let X, Y be convex sets, then $X + Y$ is convex.*
4. *Let $T : \mathbb{R}^d \to \mathbb{R}^m$ be any affine transformation.*
   (a) *If $X \subseteq \mathbb{R}^d$ is convex, then $T(X)$ is a convex set.*
   (b) *If $Y \subseteq \mathbb{R}^m$ is convex, then $T^{-1}(Y)$ is convex.*

Proof.
1. See lecture notes.
2. $\forall \mathbf{x}^1, \mathbf{y}^1 \in \alpha X,\ \forall \lambda \in [0,1]$, there exist $\mathbf{x}^2, \mathbf{y}^2 \in X$ such that $\mathbf{x}^1 = \alpha\mathbf{x}^2, \mathbf{y}^1 = \alpha\mathbf{y}^2$. Therefore, $\lambda\mathbf{x}^1 + (1-\lambda)\mathbf{y}^1 = \lambda(\alpha\mathbf{x}^2) + (1-\lambda)(\alpha\mathbf{y}^2) = \alpha(\underbrace{\lambda\mathbf{x}^2 + (1-\lambda)\mathbf{y}^2}_{\in X}) \in \alpha X$.
3. $\forall \mathbf{a}, \mathbf{b} \in X + Y,\ \forall \lambda \in [0,1]$, there exist $\mathbf{x}^1, \mathbf{x}^2 \in X$ and $\mathbf{y}^1, \mathbf{y}^2 \in Y$ such that $\mathbf{a} = \mathbf{x}^1 + \mathbf{y}^1$ and $\mathbf{b} = \mathbf{x}^2 + \mathbf{y}^2$. Therefore,
$$\begin{aligned}\lambda\mathbf{a} + (1-\lambda)\mathbf{b} &= \lambda(\mathbf{x}^1 + \mathbf{y}^1) + (1-\lambda)(\mathbf{x}^2 + \mathbf{y}^2)\\ &= \underbrace{\lambda\mathbf{x}^1 + (1-\lambda)\mathbf{x}^2}_{\in X} + \underbrace{\lambda\mathbf{y}^1 + (1-\lambda)\mathbf{y}^2}_{\in Y} \in X + Y.\end{aligned}$$
4. (a) $\forall \mathbf{x}^1, \mathbf{y}^1 \in T(X),\ \forall \lambda \in [0,1]$, there exists $\mathbf{x}^2, \mathbf{y}^2 \in X$ such that $\mathbf{x}^1 = A\mathbf{x}^2 + b$ and $\mathbf{y}^1 = A\mathbf{y}^2 + b$, where $A \in \mathbb{R}^{m\times d},\ b \in \mathbb{R}^m$. Therefore,
$$\begin{aligned}\lambda\mathbf{x}^1 + (1-\lambda)\mathbf{y}^1 &= \lambda(A\mathbf{x}^2 + b) + (1-\lambda)(A\mathbf{y}^2 + b)\\ &= \lambda A\mathbf{x}^2 + (1-\lambda)A\mathbf{y}^2 + \lambda b + (1-\lambda)b\\ &= A(\underbrace{\lambda\mathbf{x}^2 + (1-\lambda)\mathbf{y}^2}_{\in X}) + b \in T(X).\end{aligned}$$
   (b) (We may need to consider the case that $Y = \varnothing$)
   Note that $T^{-1}(Y) = \{\mathbf{x} \in \mathbb{R}^d : T(x) \in Y\}$, then $\forall \mathbf{x}^1, \mathbf{x}^2 \in T^{-1}(Y),\ T(\mathbf{x}^1) \in Y$ and $T(\mathbf{x}^2) \in Y$. Therefore, for any $\lambda \in [0,1]$,
$$T(\lambda\mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) = \lambda\underbrace{T(\mathbf{x}^1)}_{\in Y} + (1-\lambda)\underbrace{T(\mathbf{x}^2)}_{\in Y} \in Y,$$
   this proves $\lambda\mathbf{x}^1 + (1-\lambda)\mathbf{x}^2 \in T^{-1}(Y)$.

□

Exercise 7. Show that if $A \subseteq B$ then $\operatorname{conv}(A) \subseteq \operatorname{conv}(B)$. Does the converse hold?

Proof. Observe that $A \subseteq B \subseteq \operatorname{conv}(B)$, therefore $\operatorname{conv}(B)$ is a convex set that contains A. By definition of convex hull, $\operatorname{conv}(A) \subseteq \operatorname{conv}(B)$. However, the converse is not true. Consider $A = \mathbb{Q},\ B = \mathbb{R}\backslash\mathbb{Q}$, then $\operatorname{conv}(A) = \operatorname{conv}(B) = \mathbb{R}$, but $A \cap B = \varnothing$. □

Exercise 8. Show that $\operatorname{conv}(A\cap B)\subseteq \operatorname{conv}(A)\cap\operatorname{conv}(B)$. Is the containment strict?

Proof.

$$
\begin{aligned}
A\cap B\subseteq \operatorname{conv}(A) \text{ and } A\cap B\subseteq \operatorname{conv}(B) &\implies \operatorname{conv}(A) \text{ and } \operatorname{conv}(B) \text{ are convex sets containing } A\cap B\\
&\implies \operatorname{conv}(A\cap B)\subseteq \operatorname{conv}(A) \text{ and } \operatorname{conv}(A\cap B)\subseteq \operatorname{conv}(B)\\
&\implies \operatorname{conv}(A\cap B)\subseteq \operatorname{conv}(A)\cap\operatorname{conv}(B).
\end{aligned}
$$

The containment is strict. Consider $A=\mathbb{Q}$, $B=\mathbb{R}\backslash\mathbb{Q}$, then $\operatorname{conv}(A)\cap\operatorname{conv}(B)=\mathbb{R}\cap\mathbb{R}=\mathbb{R}$, but $\operatorname{conv}(A\cap B)=\operatorname{conv}(\varnothing)=\varnothing$. □

Exercise 9. Let $A\in\mathbb{R}^{d\times d}$ be a positive definite matrix and $\mathbf{c}\in\mathbb{R}^d$. Show that the ellipsoid

$$E(A,c):=\{\mathbf{x}\in\mathbb{R}^d:(\mathbf{x}-\mathbf{c})^\top A^{-1}(\mathbf{x}-\mathbf{c})\leq 1\}$$

is a convex set.

Proof. By Theorem 2, we can let $\mathbf{y}=A^{-\frac{1}{2}}(\mathbf{x}-\mathbf{c})$, then the argument follows from 4(a) in Theorem 3. □

Exercise 10. Let $A\in\mathbb{R}^{n\times n}$ be a positive definite matrix, and $C:=\{\mathbf{x}\in\mathbb{R}^n:\mathbf{x}^\top A\mathbf{x}+\mathbf{b}^\top\mathbf{x}+\mathbf{c}\leq 0\}$. Show that C is a convex set.

Proof. $\forall \mathbf{x},\mathbf{y}\in C$, $\forall\lambda\in[0,1]$, one can easily obtain the following:

$$\mathbf{x}^\top A\mathbf{x}+\mathbf{b}^\top\mathbf{x}+\mathbf{c}\leq 0, \tag{1}$$

$$\mathbf{y}^\top A\mathbf{y}+\mathbf{b}^\top\mathbf{y}+\mathbf{c}\leq 0, \tag{2}$$

$$2\mathbf{x}^\top A\mathbf{y}<\mathbf{x}^\top A\mathbf{x}+\mathbf{y}^\top A\mathbf{y}, \tag{3}$$

where the last inequality comes from $(\mathbf{x}-\mathbf{y})^\top A(\mathbf{x}-\mathbf{y})>0$ since $A\succ 0$. Therefore,

$$
\begin{aligned}
&[\lambda\mathbf{x}+(1-\lambda)\mathbf{y}]^\top A\,[\lambda\mathbf{x}+(1-\lambda)\mathbf{y}]+\mathbf{b}^\top[\lambda\mathbf{x}+(1-\lambda)\mathbf{y}]+\mathbf{c}\\
=&\lambda^2\mathbf{x}^\top A\mathbf{x}+(1-\lambda)^2\mathbf{y}^\top A\mathbf{y}+2\lambda(1-\lambda)\mathbf{x}^\top A\mathbf{y}+\lambda\mathbf{b}^\top\mathbf{x}+(1-\lambda)\mathbf{b}^\top\mathbf{y}+\mathbf{c}\\
\overset{(3)}{<}&\lambda^2\mathbf{x}^\top A\mathbf{x}+(1-\lambda)^2\mathbf{y}^\top A\mathbf{y}+\lambda(1-\lambda)(\mathbf{x}^\top A\mathbf{x}+\mathbf{y}^\top A\mathbf{y})+\lambda\mathbf{b}^\top\mathbf{x}+(1-\lambda)\mathbf{b}^\top\mathbf{y}+\mathbf{c}\\
=&\left[\lambda^2+\lambda(1-\lambda)\right]\mathbf{x}^\top A\mathbf{x}+\left[(1-\lambda)^2+\lambda(1-\lambda)\right]\mathbf{y}^\top A\mathbf{y}+\lambda\mathbf{b}^\top\mathbf{x}+(1-\lambda)\mathbf{b}^\top\mathbf{y}+\mathbf{c}\\
=&\lambda\left(\mathbf{x}^\top A\mathbf{x}+\mathbf{b}^\top\mathbf{x}+\mathbf{c}\right)+(1-\lambda)\left(\mathbf{y}^\top A\mathbf{y}+\mathbf{b}^\top\mathbf{y}+\mathbf{c}\right)\overset{(1),(2)}{\leq}0.
\end{aligned}
$$

□

Exercise 11. Let $X\subseteq\mathbb{R}^n$ be a finite set and let $\mathbf{c}\in\mathbb{R}^n$. Show that $\max\{\mathbf{c}^\top\mathbf{x}:\mathbf{x}\in X\}=\max\{\mathbf{c}^\top\mathbf{x}:\mathbf{x}\in\operatorname{conv}(X)\}$.

Proof. $X\subseteq\operatorname{conv}(X)\implies\max\{\mathbf{c}^\top\mathbf{x}:\mathbf{x}\in X\}\leq\max\{\mathbf{c}^\top\mathbf{x}:\mathbf{x}\in\operatorname{conv}(X)\}$, then it suffices to show that $\max\{\mathbf{c}^\top\mathbf{x}:\mathbf{x}\in X\}\geq\max\{\mathbf{c}^\top\mathbf{x}:\mathbf{x}\in\operatorname{conv}(X)\}$.

X is a finite set $\implies\operatorname{conv}(X)$ is a polytope (Minkowski-Weyl) $\implies\operatorname{conv}(X)$ is compact. Then Weierstrass theorem yields that there exists $\mathbf{x}^*\in\operatorname{conv}(X)$ such that

$$\mathbf{c}^\top\mathbf{x}^*=\max\{\mathbf{c}^\top\mathbf{x}:\mathbf{x}\in\operatorname{conv}(X)\}.$$

Therefore, there exist $\mathbf{x}^1, \ldots, \mathbf{x}^t \in X, \lambda_1, \ldots, \lambda_t \in [0,1]$ with $\sum_{i=1}^t \lambda_i = 1$ such that

$$\begin{aligned}
\max\{\mathbf{c}^\top \mathbf{x} : \mathbf{x} \in \operatorname{conv}(X)\} &= \mathbf{c}^\top \mathbf{x}^* \\
&= \mathbf{c}^\top \left(\sum_{j=1}^t \lambda_j \mathbf{x}^j \right) \\
&= \sum_{j=1}^t \lambda_j \mathbf{c}^\top \mathbf{x}^j \\
&\leq \sum_{j=1}^t \lambda_j \max\{\mathbf{c}^\top \mathbf{x} : \mathbf{x} \in X\} \\
&= \max\{\mathbf{c}^\top \mathbf{x} : \mathbf{x} \in X\}.
\end{aligned}$$

□

Exercise 12. Let $X \subseteq \mathbb{R}^d$. Show that X is a hyperplane $\iff$ X is an affine set of dimension $d-1$.

Proof. By definition there exist d affinely independent vectors $\mathbf{v}^1, \ldots, \mathbf{v}^d \in X$ such that

$$X = \operatorname{aff}\left(\{\mathbf{v}^1, \ldots, \mathbf{v}^d\}\right).$$

Then Theorem 2.2.6 from the lecture notes yields that there exists a matrix $A \in \mathbb{R}^{(d-(d-1))\times d}$ and a scalar b such that $X = \{\mathbf{x} \in \mathbb{R}^d : A\mathbf{x} = b\}$, so X is a hyperplane since A is just a row vector. □