Lecture 2: Separation and Duality

TTIC 31070 / CMSC 35470 / BUSF 36903 / STAT 31015

Convex Optimization

Prof. Zhiyuan Li

Spring 2026

The Plan for Lecture 2

Lecture 1 said convexity makes local information global.

Lecture 2 turns that slogan into one geometric picture:

> A closed convex set is exactly the intersection of all halfspaces that contain it.

The route:

1. Encode convex functions as sets via the **epigraph**
2. Develop **projection** and **separation** for convex sets
3. Translate back to functions through **subgradients**

Extended Real Line and Effective Domain

Definition 2.1 (Extended real line). We write $\mathbb{R} \cup \{+\infty\}$ for the one-sided extended real line, with $a \leq +\infty$ for every a, and the conventions

$$a + (+\infty) = +\infty, \quad \lambda(+\infty) = +\infty \ (\lambda > 0), \quad 0 \cdot (+\infty) := 0, \quad \inf \varnothing = +\infty.$$

Definition 2.2 (Effective domain). For $f : E \to \mathbb{R} \cup \{+\infty\}$, the effective domain is

$$\operatorname{dom} f := \{x \in E : f(x) < +\infty\}.$$

The value $+\infty$ encodes "outside the domain."

Convex Extended-Value Function

Definition 2.3 (Convex extended-value function). $f : E \to \mathbb{R} \cup \{+\infty\}$ is **convex** if

$$\forall x, y \in E,\ \forall \theta \in [0,1], \qquad f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y).$$

This is the same inequality as in Lecture 1, but now the domain is encoded inside f via $+\infty$: if $h : C \to \mathbb{R}$ is convex, define $\tilde{h} = h$ on C and $\tilde{h} = +\infty$ outside C. Then h is convex iff $\tilde{h}$ is.

Lecture 2 does not change the notion of convexity. It only changes the packaging.

Epigraph

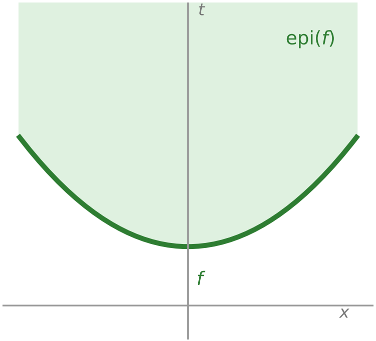
Definition 2.4 (Epigraph). For $f : E \to \mathbb{R} \cup \{+\infty\}$:

$$\text{epi}(f) := \{(x, t) \in E \times \mathbb{R} : t \geq f(x)\}.$$

Lemma 2.1 (Epigraph criterion). f is convex $\iff$ epi(f) is a convex set.

Proof idea. Convex combination of two points above the graph stays above the graph. □

After this, we leave functions and work with convex **sets** directly.



From Functions to Sets

After the epigraph criterion, the lecture switches to convex **sets**.

The first geometric input is **Euclidean projection**. Projection is what manufactures separating hyperplanes, and later those separating hyperplanes will come back as subgradients.

The chain:

$$\text{projection} \longrightarrow \text{separation} \longrightarrow \text{support} \longrightarrow \text{subgradients}$$

Halfspace, Interior, and Closure

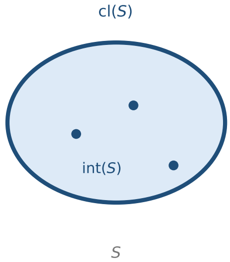
Definition 2.5 (Closed halfspace). For $\xi \in E^* \setminus \{0\}$ and $b \in \mathbb{R}$:

$$H(\xi, b) := \{x \in E : \langle \xi, x \rangle \leq b\}.$$

Definition 2.6 (Interior and closure).

$$\mathrm{int}(S) := \{x \in S : \exists\, r > 0,\ B(x, r) \subseteq S\}$$

$$\mathrm{cl}(S) := \bigcap \{F : F \text{ closed},\ S \subseteq F\}$$



$\mathrm{int}(S) \subseteq S \subseteq \mathrm{cl}(S)$. S open $\Leftrightarrow \mathrm{int}(S) = S$; S closed $\Leftrightarrow \mathrm{cl}(S) = S$.

Projection onto a Closed Convex Set

Theorem 2.3 (Projection). Let $C \subseteq E$ be nonempty, closed, and convex, with auxiliary Euclidean norm $\|\cdot\|_2$. For every $z \in E$ there exists a unique $p \in C$ such that

$$\|z - p\|_2 = \inf_{y \in C} \|z - y\|_2.$$

Moreover: p is the projection iff $\langle z - p, y - p \rangle_2 \leq 0$ for all $y \in C$.

Proof idea. Existence: Weierstrass on $C \cap \overline{B}(z, R)$. Uniqueness: strict convexity of $\|\cdot\|_2^2$. Characterization: first-order condition for the constrained minimum. $\square$

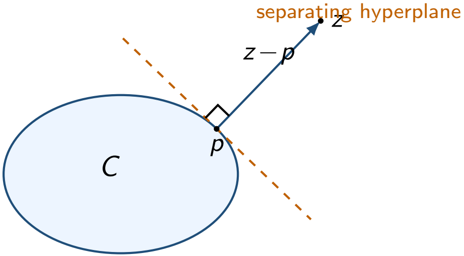
Point-to-Set Separation

How do we actually separate a point from a closed convex set? Project, and use the normal vector.

Theorem 2.2 (Point-to-set separation). Let $C \subseteq E$ be nonempty, closed, and convex, $z \notin C$. Then $\exists \xi \in E^* \setminus \{0\}$, $b \in \mathbb{R}$:

$$\langle \xi, z \rangle > b \quad \text{and} \quad \forall y \in C,\ \langle \xi, y \rangle \leq b.$$

Proof idea. Project z onto C; set $\xi := z - p$, $b := \langle \xi, p \rangle$. $\square$



Why Is Closedness Needed?

Question. Can we drop the closedness assumption in the separation theorem?

No. Let $C = (0, 1) \subseteq \mathbb{R}$ and $z = 1$.

Then $z \notin C$, but no strict separating halfspace exists: any halfspace $\{x : x \leq b\}$ that contains C must have $b \geq 1$, so it also contains z.

Closedness is needed both for the **proof** (projection may not exist) and for the **statement** (separation can genuinely fail).

Separation as a Certificate

The separation theorem gives a **certificate** that $z \notin C$:

- To *prove* $z \in C$, one must exhibit a representation of z as a feasible point.
- To *prove* $z \notin C$, it suffices to exhibit a halfspace $H(\xi, b)$ with $C \subseteq H(\xi, b)$ and $z \notin H(\xi, b)$.

This is the **certificate viewpoint** that recurs throughout the course: duality provides short, checkable proofs of infeasibility and optimality.

Lecture 3: Farkas' lemma as an infeasibility certificate for LP.

Lecture 5: KKT multipliers as an optimality certificate.

Example: The PSD Cone as a Proof System

Example. Work in symmetric matrices $\mathbb{S}^n$ with Frobenius pairing $\langle Y, Z \rangle := \operatorname{tr}(YZ)$. Then:

$$\mathbb{S}^n_+ = \bigcap_{u \in \mathbb{R}^n} \{X \in \mathbb{S}^n : \langle uu^\top, X \rangle \geq 0\}.$$

If $X \notin \mathbb{S}^n_+$, there exists u with $u^\top X u < 0$.

So $uu^\top$ is a **concrete certificate** that X is not PSD.

The separation theorem says this kind of linear proof system is **complete**: every false claim "$X \succeq 0$" can be refuted by a separating inequality.

Supporting Hyperplane Theorem

Projection gives separation for **exterior** points. What about **boundary** points? Let $z_k \to x$ approach the boundary and pass to the limit.

Theorem 2.4 (Supporting hyperplane). Let $C \subseteq E$ be nonempty and convex, $x \in \mathrm{cl}(C) \setminus \mathrm{int}(C)$. Then there exist $\xi \in E^* \setminus \{0\}$ and $b \in \mathbb{R}$ such that

$$\langle \xi, x \rangle = b \qquad \text{and} \qquad \forall y \in C, \ \langle \xi, y \rangle \leq b.$$

Proof idea. Take a sequence $z_k \to x$ with $z_k \notin C$. Separate each z_k from C, normalize the separating covectors on the unit sphere, extract a convergent subsequence (compactness), and pass to the limit. $\square$

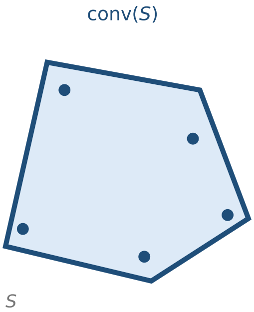
Convex Hull and Closed Convex Hull

Definition 2.7 (Convex hull).

$$\operatorname{conv}(S) := \bigcap \{C \subseteq E : C \text{ convex}, \ S \subseteq C\}.$$

Definition 2.8 (Closed convex hull).

$$\operatorname{cl}\operatorname{conv}(S) := \bigcap \{C \subseteq E : C \text{ closed, convex}, \ S \subseteq C\}.$$



Basic Properties of Closure, Convex Hull, and Closed Convex Hull

Proposition 2.5. Let $S \subseteq E$. Then:

1. $\mathrm{cl}(S)$ is closed, $\mathrm{conv}(S)$ is convex, $S \subseteq \mathrm{cl}(S) \cap \mathrm{conv}(S)$.
2. $\mathrm{cl\,conv}(S) = \mathrm{cl}(\mathrm{conv}(S))$.
3. $\mathrm{cl\,conv}(S)$ is closed and convex, and $S \subseteq \mathrm{cl\,conv}(S)$.
4. If C is closed and convex, then $\mathrm{cl\,conv}(C) = C$.
5. $\mathrm{cl\,conv}(\mathrm{cl\,conv}(S)) = \mathrm{cl\,conv}(S)$.

The proof is left as an exercise.

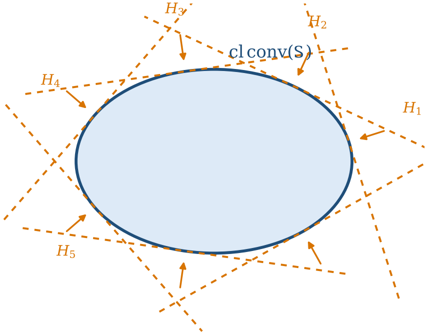
Closed Convex Hull = Intersection of Halfspaces

Once supporting hyperplanes are available, can we reconstruct a closed convex set from **all** the halfspaces that contain it?

Theorem 2.6. For any $S \subseteq E$:

$$\operatorname{cl}\operatorname{conv}(S) = \bigcap \{H(\xi, b) : S \subseteq H(\xi, b)\}.$$

Proof idea. $\subseteq$: trivial. $\supseteq$: if $z \notin \operatorname{cl}\operatorname{conv}(S)$, separate. $\square$



Intersecting all containing halfspaces recovers $\operatorname{cl}\operatorname{conv}(S)$.

Convex Cones and Duality

The halfspace-hull picture works for general sets. Cones are the special case where the bias term b vanishes, so halfspaces become **half-spaces through the origin**.

Definition 2.9 (Convex cone, dual cone, polar cone). $K \subseteq E$ is a convex cone if

$$\forall x_1, x_2 \in K,\ \forall \alpha_1, \alpha_2 \geq 0, \qquad \alpha_1 x_1 + \alpha_2 x_2 \in K.$$

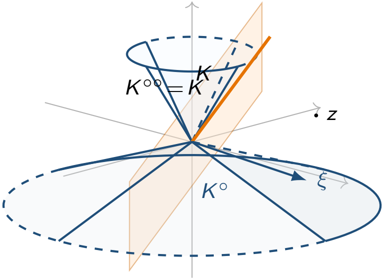
Its dual and polar cones are

$$K^* := \{\xi \in E^* : \forall x \in K,\ \langle \xi, x\rangle \geq 0\}, \qquad K^\circ := \{\xi \in E^* : \forall x \in K,\ \langle \xi, x\rangle \leq 0\}.$$

Cone Separation and Bipolarity

Theorem 2.7 (Cone separation and bipolarity). Let K be a nonempty closed convex cone, $z \notin K$. Then $\exists \xi \in K^\circ$: $\langle \xi, z \rangle > 0$.

Consequently, $K^{\circ\circ} = K$.



Proof idea.
Separation: Apply point-to-set separation to K. Since $0 \in K$, we get $b \geq 0$. Since K is a cone, $\langle \xi, x \rangle \leq 0$ for all $x \in K$ (otherwise scale x to violate b). So $\xi \in K^\circ$.
Bipolarity: $K \subseteq K^{\circ\circ}$ is immediate. If $z \notin K$, separation gives $\xi \in K^\circ$ with $\langle \xi, z \rangle > 0$, so $z \notin K^{\circ\circ}$. □

Affine Sets, Affine Hull, and Relative Interior

The cone theorem is the last set-level result. Before returning to functions, we need **relative interior**: the subgradient existence theorem lives on $\mathrm{ri}(\mathrm{dom}\, f)$, not $\mathrm{dom}\, f$.

Definition 2.10 (Affine set). $A \subseteq E$ is **affine** if $\forall x, y \in A,\ \forall \theta \in \mathbb{R},\ \theta x + (1-\theta)y \in A$.

Definition 2.11 (Affine hull and relative interior).

$$\mathrm{aff}(S) := \bigcap \{A \subseteq E : A \text{ is affine and } S \subseteq A\}.$$

If $C \subseteq E$, its relative interior is

$$\mathrm{ri}(C) := \{x \in C : \exists\, r > 0,\ \{y \in \mathrm{aff}(C) : \|y - x\| < r\} \subseteq C\}.$$

Why affine? Convexity and separation are fundamentally affine notions — no preferred origin is needed. We work in E but pass locally to affine subspaces when needed.

Subgradient

We now translate the geometry back into function language. A supporting hyperplane to $\operatorname{epi}(f)$ through $(x, f(x))$ is exactly a global affine lower bound on f — and that is what a subgradient is.

Definition 2.12 (Subgradient and subdifferential). Let $f : E \to \mathbb{R} \cup \{+\infty\}$, $x \in \operatorname{dom} f$. A covector $g \in E^*$ is a **subgradient** of f at x if

$$\forall y \in E, \qquad f(y) \geq f(x) + \langle g, y - x \rangle.$$

The set of all such covectors is $\partial f(x)$.

Lemma 2.8 (Subgradients = supporting hyperplanes of epi).
$g \in \partial f(x) \iff (g, -1)$ supports $\operatorname{epi}(f)$ at $(x, f(x))$.

Proof idea. Unpack both definitions; they say the same thing. □

Proper Function

Definition 2.13 (Proper extended-value function). $f : E \to \mathbb{R} \cup \{+\infty\}$ is **proper** if

$$\operatorname{dom} f \neq \varnothing.$$

Equivalently, there exists $x \in E$ with $f(x) < +\infty$.

Properness rules out the trivial case $f \equiv +\infty$. The subgradient existence theorem requires properness.

Differentiable Convex ⇒ Unique Subgradient

Proposition 2.9. Let $f : E \to \mathbb{R}$ be convex and differentiable. Then for every $x \in E$,

$$\partial f(x) = \{\nabla f(x)\}.$$

So for smooth convex functions, the subgradient is just the gradient. Subgradients become genuinely new only for **nonsmooth** convex functions.

Existence of Subgradients

Theorem 2.10 (Subgradient existence on the relative interior). Let $f : E \to \mathbb{R} \cup \{+\infty\}$ be proper and convex. Then

$$\forall x \in \mathrm{ri}(\mathrm{dom}\, f), \qquad \partial f(x) \neq \varnothing.$$

Proof idea. The point $(x, f(x))$ is a boundary point of $\mathrm{epi}(f)$ inside $\mathrm{aff}(\mathrm{dom}\, f) \times \mathbb{R}$. Apply supporting hyperplane there and read the result as a subgradient. □

Can we weaken $x \in \mathrm{ri}(\mathrm{dom}\, f)$ to $x \in \mathrm{dom}\, f$?

No: $f(x) = x \log x$ on $[0, \infty)$ has $\partial f(0) = \varnothing$, since $\log y \geq g$ for all $y > 0$ is impossible.

Summary & What's Next

Today:

- Epigraph converts function convexity to set convexity
- Projection $\rightarrow$ point-to-set separation $\rightarrow$ supporting hyperplanes
- $\operatorname{cl}\operatorname{conv}(S) = \bigcap$ containing halfspaces
- Cone separation and bipolarity: $K^{\circ\circ} = K$
- Subgradient $=$ supporting hyperplane of the epigraph
- Subgradients exist on $\operatorname{ri}(\operatorname{dom} f)$

Next lecture:

- Linear programming as the first application of duality and separation
- Farkas' lemma as a certificate alternative
- Strong duality and complementary slackness