Lecture 4: Convex Conjugates and Marginal Duality

TTIC 31070 / CMSC 35470 / BUSF 36903 / STAT 31015

Convex Optimization

Prof. Zhiyuan Li

Spring 2026

From LP to Convex Duality

Lecture 3: duality for **linear programs** via Farkas' lemma.

Today: duality for **general convex programs**.

Key idea. The convex conjugate f^* records all **affine lower bounds** of f.

For each slope $\xi \in E^*$, the affine function

$$x \mapsto \langle \xi, x \rangle - f^*(\xi)$$

is the **tightest** affine function with slope ξ that lies below f.

The biconjugate f^{**} takes the supremum over all such bounds — and for closed convex f, recovers f exactly.

Convex Conjugate

Definition 4.1 (Convex conjugate). Let $f : E \to \mathbb{R} \cup \{+\infty\}$ be proper. Its *convex conjugate* is

$$f^*(\xi) := \sup_{x \in E}\{\langle \xi, x\rangle - f(x)\}, \qquad \xi \in E^*.$$

Properties:

- f^* is always convex and closed (as a supremum of affines in ξ), even if f is not convex
- $f^*(\xi) =$ smallest shift so that $x \mapsto \langle \xi, x\rangle - f^*(\xi)$ lies below f everywhere
- Under our convention (f proper), f^* is valued in $\mathbb{R} \cup \{+\infty\}$

Interactive animation: Lecture 4 on course website.

Biconjugate

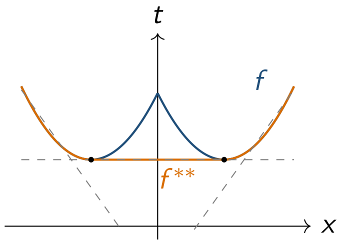
Definition 4.1 (continued). The *biconjugate* is

$$f^{**}(x) := \sup_{\xi \in E^*} \{\langle \xi, x\rangle - f^*(\xi)\}, \qquad x \in E.$$

$f^{**}(x)$ = supremum of **all** affine functions lying below f, evaluated at x.

Equivalently: f^{**} is the greatest closed convex function $\leq f$.

We identify $E \simeq E^{**}$ via $\iota_E(x)(\xi) = \langle \xi, x\rangle$ (finite-dim. isomorphism), so f^{**} lives on E.



f^{**} agrees with f on the convex parts and flattens the non-convex gap.

Closure

Definition 4.2 (Closure and closed function). The *closure* of $f : E \to \mathbb{R} \cup \{+\infty\}$ is the function $\operatorname{cl} f$ satisfying

$$\operatorname{epi}(\operatorname{cl} f) = \operatorname{cl}(\operatorname{epi}(f)).$$

f is *closed* if $\operatorname{epi}(f)$ is closed, equivalently $\operatorname{cl} f = f$.

In finite dimensions: **closed** $\Leftrightarrow$ **lower semicontinuous**.

A proper, closed, convex function has a closed convex epigraph — so the separation tools from Lecture 2 apply directly to $\operatorname{epi}(f)$.

Fenchel–Young Inequality

Lemma 4.1 (Fenchel–Young inequality). Let $f : E \to \mathbb{R} \cup \{+\infty\}$ be proper. Then

$$f(x) + f^*(\xi) \geq \langle \xi, x \rangle \qquad \forall x \in E,\ \xi \in E^*.$$

If f is also closed and convex, then

$$f(x) + f^*(\xi) = \langle \xi, x \rangle \quad \iff \quad \xi \in \partial f(x).$$

Proof idea. $f^*(\xi) \geq \langle \xi, x \rangle - f(x)$ by definition. Rearranging gives the inequality. Equality means the affine lower bound with slope ξ touches f at x — exactly the subgradient condition. □

Biconjugate Inequality and Fenchel–Moreau

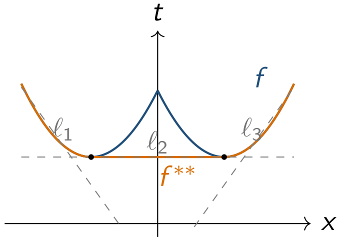
Lemma 4.2. $f^{**}(x) \le f(x)$ for all $x \in E$.

Proof. By Fenchel–Young, $\langle \xi, x\rangle - f^*(\xi) \le f(x)$ for every ξ. Taking $\sup_\xi$: $f^{**}(x) \le f(x)$. □

Theorem 4.3 (Fenchel–Moreau). Let $f : E \to \mathbb{R} \cup \{+\infty\}$ be proper, closed, and convex. Then

$$f^{**}(x) = f(x) \qquad \forall\, x \in E.$$

More generally: $(\mathrm{cl}\, f)^* = f^*$, so $f^{**} = \mathrm{cl}\, f$.



f^{**} = sup of all affine lower bounds ℓ_i. Agrees with f on convex parts; flattens the non-convex gap.

Proof of Fenchel–Moreau

Fix x_0 and $t_0 < f(x_0)$. Then $(x_0, t_0) \notin \mathrm{epi}(f)$.

Since $\mathrm{epi}(f)$ is closed convex, **separation** (Theorem 2.6) gives a halfspace $H_{\xi,a} = \{(x,t) : t \geq \langle \xi, x \rangle - a\}$ containing $\mathrm{epi}(f)$ but not (x_0, t_0).

This halfspace is an affine lower bound for f. Since the tightest affine lower bound with slope ξ has intercept $f^*(\xi)$:

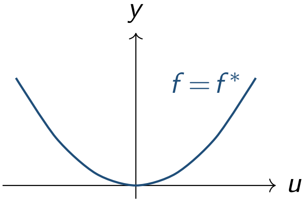
$$t_0 < \langle \xi, x_0 \rangle - a \leq \langle \xi, x_0 \rangle - f^*(\xi) \leq f^{**}(x_0).$$

Since $t_0 < f(x_0)$ was arbitrary: $f(x_0) \leq f^{**}(x_0)$. $\square$

Example: Power-Law Conjugate Pair

Example 4.1. Let $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, and $f(x) = \frac{|x|^p}{p}$ on $\mathbb{R}$. Then

$$f^*(\xi) = \frac{|\xi|^q}{q}.$$



Proof. Maximize $\varphi_\xi(x) = x\xi - \frac{x^p}{p}$ for $x, \xi \geq 0$. Setting $\varphi'_\xi(x) = \xi - x^{p-1} = 0$ gives $x = \xi^{q-1}$. Substituting: $f^*(\xi) = \xi^q - \frac{\xi^q}{p} = \frac{\xi^q}{q}$. □

Fenchel–Young: $x\xi \leq \dfrac{|x|^p}{p} + \dfrac{|\xi|^q}{q}$ (**Young inequality**).

Special case $p = q = 2$: $\frac{1}{2}u^2$ is self-conjugate.

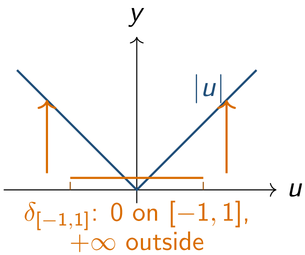
Indicator Functions and Support Functions

Definition 4.3 (Indicator function).
$$\delta_C(x) := \begin{cases} 0, & x \in C, \\ +\infty, & x \notin C. \end{cases}$$

Example 4.2. Let $C \subseteq E$ be nonempty, with support function $\sigma_C(\xi) := \sup_{x \in C} \langle \xi, x \rangle$. Then $(\delta_C)^* = \sigma_C$.

1D case: $C = [-1, 1]$.

$$(\delta_{[-1,1]})^* = |\cdot|, \quad |\cdot|^* = \delta_{[-1,1]}.$$



Constraints ↔ penalties.

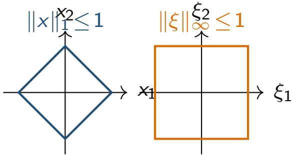
Norms and Dual Unit Balls

Example 4.3. Let $\|\cdot\|$ be a norm on E with dual norm $\|\xi\|_* := \sup_{\|x\|\le 1}\langle \xi, x\rangle$. Then

$$\|\cdot\|^* = \delta_{B_*}, \qquad B_* := \{\xi \in E^* : \|\xi\|_* \le 1\}.$$

Proof sketch.

- $\|\xi\|_* \le 1$: $\langle \xi, x\rangle \le \|\xi\|_* \|x\| \le \|x\|$, so $\langle \xi, x\rangle - \|x\| \le 0$. Hence $\|\cdot\|^*(\xi) = 0$.
- $\|\xi\|_* > 1$: pick x_0 with $\langle \xi, x_0\rangle > \|x_0\|$. Scale $tx_0 \to \infty$. $\square$



ℓ_1-ball $\leftrightarrow$ ℓ_∞-ball.

Example: Exponential and Negative Entropy

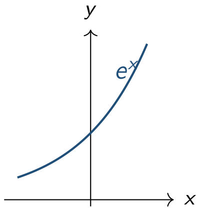
Example 4.4. Let $f(x) = e^x$ on $\mathbb{R}$. Then

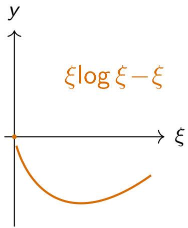
$$f^*(\xi) = \begin{cases} \xi \log \xi - \xi, & \xi \geq 0, \\ +\infty, & \xi < 0, \end{cases}$$

with the convention $0 \log 0 := 0$.

Proof. For $\xi > 0$: maximize $\xi x - e^x$. Setting $\xi - e^x = 0$ gives $x = \log \xi$, so $f^*(\xi) = \xi \log \xi - \xi$. □

This pair recurs in mirror descent and information geometry.





The exponential/negative-entropy conjugate pair.

The Perturbation Viewpoint

Lecture 3: the LP value function $V(u) = \inf\{c^\top x : Ax \geq u,\ x \geq 0\}$ is convex, and dual variables are supporting covectors of V.

General template. Given a convex perturbation function $\Phi : X \times U \to \mathbb{R} \cup \{+\infty\}$, define the **marginal value function**

$$p(u) := \inf_{x \in X} \Phi(x, u).$$

- p is convex (projection of epi(Φ) preserves convexity)
- $p(0) = \inf_x \Phi(x, 0)$ is the original problem value
- Dual variables = supporting covectors of p at $u = 0$

This is the LP value function from Lecture 3, now for arbitrary convex programs.

Marginal Duality: Weak Duality

Theorem 4.4 (Marginal duality). Let $\Phi : X \times U \to \mathbb{R} \cup \{+\infty\}$ be convex. Define $p(u) := \inf_{x \in X} \Phi(x, u)$.

(1) For every $y \in U^*$: $\quad p^*(y) = \Phi^*(0, y)$.

Consequently,

$$p(u) \geq -\Phi^*(0, y) + \langle y, u \rangle \qquad \forall\, u \in U,\ y \in U^*.$$

Proof. $p^*(y) = \sup_u \{\langle y, u \rangle - \inf_x \Phi(x, u)\} = \sup_{x,u} \{\langle y, u \rangle - \Phi(x, u)\} = \Phi^*(0, y)$. □

Setting $u = 0$: $\quad \sup_{y \in U^*} \{-\Phi^*(0, y)\} \leq p(0)$. $\quad$ **(weak duality)**

Marginal Duality: Strong Duality

Theorem 4.4 (continued).

(2) If $p(0) = -\infty$, then $\sup_{y \in U^*}\{-\Phi^*(0, y)\} = -\infty$.

(3) If $p(0) > -\infty$ and $\partial p(0) \neq \varnothing$, then

$$p(0) = \max_{y \in U^*}\{-\Phi^*(0, y)\}.$$

(4) If $p(0) \in \mathbb{R}$ and $0 \in \operatorname{ri}(\operatorname{dom} p)$, then the same holds.

Proof of (2). For every $M > 0$, choose x_M with $\Phi(x_M, 0) \leq -M$. Then $\Phi^*(0, y) \geq M$ for every y, so the sup is $-\infty$. □

Proof of (3). Pick $y \in \partial p(0)$. Then $p^*(y) = \langle y, 0\rangle - p(0) = -p(0)$, so $-\Phi^*(0, y) = p(0)$. □

Proof of (4). $0 \in \operatorname{ri}(\operatorname{dom} p)$ + convexity ensures $\partial p(0) \neq \varnothing$ (Theorem 2.10), then apply (3). Part (4) gives a checkable **constraint qualification**.

Application: Recovering the LP Dual

Example 4.5. $\inf\{c^\top x : Ax \geq b,\ x \geq 0\}$.
Perturbation: $\Phi(x,u) = c^\top x + \delta_{\mathbb{R}^n_+}(x) + \delta_{\mathbb{R}^m_+}(Ax - b - u)$.

$p(u) = \inf\{c^\top x : Ax \geq b + u,\ x \geq 0\}$.

Conjugate computation:

$$\Phi^*(0,y) = \begin{cases} -b^\top y, & A^\top y \leq c,\ y \geq 0, \\ +\infty, & \text{otherwise.} \end{cases}$$

Theorem 4.4 gives: $\sup\{b^\top y : A^\top y \leq c,\ y \geq 0\} \leq \inf\{c^\top x : Ax \geq b,\ x \geq 0\}$.

This is LP weak duality from Lecture 3. For strong duality via part (4), one additionally needs $0 \in \operatorname{ri}(\operatorname{dom} p)$; Lecture 3 avoids this assumption by using polyhedral structure directly.

Computing $\Phi^*(0, y)$ for the LP

By definition of the conjugate,

$$\Phi^*(0,y) = \sup_{x\in\mathbb{R}^n,\, u\in\mathbb{R}^m} \{y^\top u - c^\top x - \delta_{\mathbb{R}^n_+}(x) - \delta_{\mathbb{R}^m_+}(Ax - b - u)\}.$$

The indicators force $x \geq 0$ and $Ax - b - u \geq 0$, i.e. $u \leq Ax - b$:

$$\Phi^*(0,y) = \sup_{x\geq 0}\left\{ \sup_{u\leq Ax-b} y^\top u \; - \; c^\top x\right\}.$$

Inner sup over u: $y \geq 0 \Rightarrow$ maximized at $u = Ax - b$. Some $y_i < 0 \Rightarrow u_i \to -\infty$: $+\infty$.

For $y \geq 0$: $\Phi^*(0,y) = -b^\top y + \sup_{x\geq 0}\{x^\top(A^\top y - c)\}$.

Sup over x: $A^\top y \leq c \Rightarrow 0$; some $(A^\top y)_j > c_j \Rightarrow +\infty$.

$$\boxed{\Phi^*(0,y) = \begin{cases} -b^\top y, & A^\top y \leq c,\ y \geq 0, \\ +\infty, & \text{otherwise.} \end{cases}}$$

Application: Norm Regularization Duality

Example 4.6. Let f be proper closed convex, $\lambda > 0$, $A : X \to Y$ linear.

$$\inf_{x \in X} \{f(x) + \lambda \|Ax\|\} = \max_{\substack{y \in Y^* \\ \|y\|_* \leq \lambda}} \{-f^*(-A^* y)\}.$$

Key steps:

- Perturbation: $\Phi(x, u) = f(x) + \lambda \|Ax + u\|$
- Since $\lambda \|\cdot\|$ is finite everywhere, $\operatorname{dom} p = Y$, so $0 \in \operatorname{ri}(\operatorname{dom} p)$
- $\Phi^*(0, y) = f^*(-A^* y) + \delta_{\{\|y\|_* \leq \lambda\}}(y)$ (using Example 4.3)

Computing $\Phi^*(0, y)$ for Norm Regularization

By definition, with $g(v) := \lambda\|v\|$,

$$\Phi^*(0, y) = \sup_{x \in X,\, u \in Y} \{\langle y, u\rangle - f(x) - \lambda\|Ax + u\|\}.$$

Substitute $z := Ax + u$ (u free, so bijective):

$$\begin{aligned}\Phi^*(0, y) &= \sup_{x,\, z}\{\langle y, z - Ax\rangle - f(x) - \lambda\|z\|\} \\ &= \underbrace{\sup_{x}\{\langle -A^* y, x\rangle - f(x)\}}_{f^*(-A^* y)} + \underbrace{\sup_{z}\{\langle y, z\rangle - \lambda\|z\|\}}_{g^*(y)}.\end{aligned}$$

Computing g^*: $\|y\|_* \le \lambda$: $\langle y, z\rangle \le \lambda\|z\|$, so $g^*(y) = 0$. $\|y\|_* > \lambda$: scale $tz_0 \to \infty$, $g^*(y) = +\infty$.

$$\boxed{\Phi^*(0, y) = f^*(-A^* y) + \delta_{\{\|y\|_* \le \lambda\}}(y).}$$

Summary & What's Next

Today:

- Convex conjugate f^*: encodes all affine lower bounds of f
- Fenchel–Young: $f(x) + f^*(\xi) \geq \langle \xi, x \rangle$, with equality iff $\xi \in \partial f(x)$
- Fenchel–Moreau: $f^{**} = f$ for proper closed convex f
- Conjugate pairs: powers, indicator/support, norm/dual ball, exp/entropy
- Marginal duality: $\sup_y \{-\Phi^*(0, y)\} \leq p(0)$, with equality under constraint qualification
- Applications: LP dual and norm regularization as special cases

Next lecture:

- KKT conditions from the perturbation viewpoint
- From LP complementary slackness to nonlinear optimality conditions