Duality of Convex Sets and Functions

Instructor: Shaddin Dughmi
Outline

1 Convexity and Duality

2 Duality of Convex Sets

3 Duality of Convex Functions
There are two equivalent ways to represent a convex set:

- The family of points in the set (standard representation)
- The set of halfspaces containing the set (“dual” representation)
There are two equivalent ways to represent a convex set
- The family of points in the set (standard representation)
- The set of halfspaces containing the set ("dual" representation)

This equivalence between the two representations gives rise to a variety of "duality" relationships among convex sets, cones, and functions.
There are two equivalent ways to represent a convex set:
- The family of points in the set (standard representation)
- The set of halfspaces containing the set (“dual” representation)

This equivalence between the two representations gives rise to a variety of “duality” relationships among convex sets, cones, and functions.

**Definition**

“Duality” is a woefully overloaded mathematical term for a relation that groups elements of a set into “dual” pairs.
Theorem

A closed convex set $S$ is the intersection of all closed halfspaces $\mathcal{H}$ containing it.
Theorem

A closed convex set $S$ is the intersection of all closed halfspaces $\mathcal{H}$ containing it.

Proof

- Clearly, $S \subseteq \bigcap_{H \in \mathcal{H}} H$
- To prove equality, consider $x \notin S$
- By the separating hyperplane theorem, there is a hyperplane separating $S$ from $x$
- Therefore there is $H \in \mathcal{H}$ with $x \notin H$, hence $x \notin \bigcap_{H \in \mathcal{H}} H$
A closed convex cone $K$ is the intersection of all closed homogeneous halfspaces $\mathcal{H}$ containing it.
Theorem

A closed convex cone $K$ is the intersection of all closed homogeneous halfspaces $\mathcal{H}$ containing it.

Proof

- For every non-homogeneous halfspace $a^T x \leq b$ containing $K$, the smaller homogeneous halfspace $a^T x \leq 0$ contains $K$ as well.
- Therefore, can discard non-homogeneous halfspaces when taking the intersection.
A convex function is the point-wise supremum of all affine functions under-estimating it everywhere.
Theorem
A convex function is the point-wise supremum of all affine functions under-estimating it everywhere.

Proof
- \( \text{epi } f \) is convex
- Therefore \( \text{epi } f \) is the intersection of a family of halfspaces of the form \( a^\top x - t \leq b \), for some \( a \in \mathbb{R}^n \) and \( b \in \mathbb{R} \). (Why?)
- Each such halfspace constrains \( (x, t) \in \text{epi } f \) to \( a^\top x - b \leq t \)
- \( f(x) \) is the lowest \( t \) s.t. \( (x, t) \in \text{epi } f \)
- Therefore, \( f(x) \) is the point-wise maximum of \( a^\top x - b \) over all halfspaces
Outline

1 Convexity and Duality

2 Duality of Convex Sets

3 Duality of Convex Functions
Polar Duality of Convex Sets

One way of representing the all halfspaces containing a convex set.

**Polar**

Let $S \subseteq \mathbb{R}^n$ be a closed convex set containing the origin. The polar of $S$ is defined as follows:

$$S^\circ = \{ y : y^T x \leq 1 \text{ for all } x \in S \}$$

**Note**

- Every halfspace $a^T x \leq b$ with $b \neq 0$ can be written as a “normalized” inequality $y^T x \leq 1$, by dividing by $b$.
- $S^\circ$ can be thought of as the normalized representations of halfspaces containing $S$. 
\[ S^\circ = \{ y : y^\top x \leq 1 \text{ for all } x \in S \} \]

**Properties of the Polar**

1. \( S^{\circ \circ} = S \)
2. \( S^\circ \) is a closed convex set containing the origin
3. When 0 is in the interior of \( S \), then \( S^\circ \) is bounded.
\[ S^\circ = \{ y : y^\top x \leq 1 \text{ for all } x \in S \} \]

Properties of the Polar

1. \( S^{\circ \circ} = S \)
2. \( S^\circ \) is a closed convex set containing the origin
3. When 0 is in the interior of \( S \), then \( S^\circ \) is bounded.

Follows from representation as intersection of halfspaces

\( S \) contains an \( \epsilon \)-ball centered at the origin, so \( \|y\| \leq 1/\epsilon \) for all \( y \in S^\circ \).
\[ S^\circ = \{ y : y^\top x \leq 1 \text{ for all } x \in S \} \]

**Properties of the Polar**

1. \( S^{\circ\circ} = S \)
2. \( S^\circ \) is a closed convex set containing the origin
3. When 0 is in the interior of \( S \), then \( S^\circ \) is bounded.

**Easy to see that \( S \subseteq S^{\circ\circ} \)**

- Take \( x_\circ \not\in S \), by SSHT and \( 0 \in S \), there is a halfspace \( z^\top x \leq 1 \) containing \( S \) but not \( x_\circ \) (i.e. \( z^\top x_\circ > 1 \))
- \( z \in S^\circ \), therefore \( x_\circ \not\in S^{\circ\circ} \)
\[
S^\circ = \{ y : y^\top x \leq 1 \text{ for all } x \in S \}
\]

**Properties of the Polar**

1. \( S^{\circ\circ} = S \)
2. \( S^\circ \) is a closed convex set containing the origin
3. When 0 is in the interior of \( S \), then \( S^\circ \) is bounded.

**Note**

When \( S \) is non-convex, \( S^\circ = (\text{convexhull}(S))^\circ \), and \( S^{\circ\circ} = \text{convexhull}(S) \).
Examples

Norm Balls

- The polar of the Euclidean unit ball is itself (we say it is self-dual)
- The polar of the $1$-norm ball is the $\infty$-norm ball
- More generally, the $p$-norm ball is dual to the $q$-norm ball, where
  \[ \frac{1}{p} + \frac{1}{q} = 1 \]
Examples

Polytopes

Given a polytope $P$ represented as $Ax \leq \vec{1}$, the polar $P^\circ$ is the convex hull of the rows of $A$.

- Facets of $P$ correspond to vertices of $P^\circ$.
- Dually, vertices of $P$ correspond to facets of $P^\circ$. 
Polar duality takes a simplified form when applied to cones

The polar of a closed convex cone $K$ is given by

$$K^\circ = \{ y : y^\top x \leq 0 \text{ for all } x \in K \}$$

Note

- If halfspace $y^\top x \leq b$ contains $K$, then so does smaller $y^\top x \leq 0$.
- $K^\circ$ represents all homogeneous halfspaces containing $K$. 
Polar duality takes a simplified form when applied to cones

**Polar**

The polar of a closed convex cone $K$ is given by

$$K^\circ = \{ y : y^\top x \leq 0 \text{ for all } x \in K \}$$

**Dual Cone**

By convention, $K^* = -K^\circ$ is referred to as the dual cone of $K$.

$$K^* = \{ y : y^\top x \geq 0 \text{ for all } x \in K \}$$
\( K^\circ = \{ y : y^\top x \leq 0 \text{ for all } x \in K \} \)

Properties of the Polar Cone

1. \( K^{\circ\circ} = K \)
2. \( K^\circ \) is a closed convex cone
$K^\circ = \{ y : y^\top x \leq 0 \text{ for all } x \in K \}$

### Properties of the Polar Cone

1. $K^{\circ\circ} = K$
2. $K^\circ$ is a closed convex cone

1. Same as before
2. Intersection of homogeneous halfspaces
Examples

- The polar of a subspace is its orthogonal complement
- The polar cone of the nonnegative orthant is the nonpositive orthant
  - Self-dual
- The polar of a polyhedral cone $Ax \preceq 0$ is the conic hull of the rows of $A$
- The polar of the cone of positive semi-definite matrices is the cone of negative semi-definite matrices
  - Self-dual
Recall: Farkas’ Lemma

Let $K$ be a closed convex cone and let $w \notin K$. There is $z \in \mathbb{R}^n$ such that $z^T x \leq 0$ for all $x \in K$, and $z^T w > 0$.

Equivalently: there is $z \in K^\circ$ with $z^T w > 0$. 
Outline

1  Convexity and Duality
2  Duality of Convex Sets
3  Duality of Convex Functions
Conjugation of Convex Functions

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be a convex function. The conjugate of $f$ is

$$f^*(y) = \sup_x (y^Tx - f(x))$$

Note

- $f^*(y)$ is the minimal value of $\beta$ such that the affine function $y^Tx - \beta$ underestimates $f(x)$ everywhere.
- Equivalently, the distance we need to lower the hyperplane $y^Tx - t = 0$ in order to get a supporting hyperplane to $\text{epi } f$.
- $y^Tx - t = f^*(y)$ are the supporting hyperplanes of $\text{epi } f$. 

Duality of Convex Functions 12/14
\[ f^*(y) = \sup_{x} (y^T x - f(x)) \]

**Properties of the Conjugate**

1. \( f^{**} = f \) when \( f \) is convex
2. \( f^* \) is a convex function
3. \( xy \leq f(x) + f^*(y) \) for all \( x, y \in \mathbb{R}^n \) (Fenchel’s Inequality)
\[ f^*(y) = \sup_{x} (y^T x - f(x)) \]

### Properties of the Conjugate

1. \( f^{**} = f \) when \( f \) is convex
2. \( f^* \) is a convex function
3. \( xy \leq f(x) + f^*(y) \) for all \( x, y \in \mathbb{R}^n \) (Fenchel’s Inequality)

- Supremum of affine functions of \( y \)
- By definition of \( f^*(y) \)
\[ f^*(y) = \sup_x (y^\top x - f(x)) \]

**Properties of the Conjugate**

1. \( f^{**} = f \) when \( f \) is convex
2. \( f^* \) is a convex function
3. \( xy \leq f(x) + f^*(y) \) for all \( x, y \in \mathbb{R}^n \) (Fenchel’s Inequality)

1. \( f^{**}(x) = \max_y y^\top x - f^*(y) \) when \( f \) is convex
\[
f^*(y) = \sup_x (y^T x - f(x))
\]

### Properties of the Conjugate

1. \( f^{**} = f \) when \( f \) is convex
2. \( f^* \) is a convex function
3. \( xy \leq f(x) + f^*(y) \) for all \( x, y \in \mathbb{R}^n \) (Fenchel’s Inequality)

- \( f^{**}(x) = \max_y y^T x - f^*(y) \) when \( f \) is convex
- For fixed \( y \), \( f^*(y) \) is minimal \( \beta \) such that \( y^T x - \beta \) underestimates \( f \).
\[ f^*(y) = \sup_x (y^\top x - f(x)) \]

**Properties of the Conjugate**

1. \( f^{**} = f \) when \( f \) is convex
2. \( f^* \) is a convex function
3. \( xy \leq f(x) + f^*(y) \) for all \( x, y \in \mathbb{R}^n \) (Fenchel’s Inequality)

- \( f^{**}(x) = \max_y y^\top x - f^*(y) \) when \( f \) is convex
- For fixed \( y \), \( f^*(y) \) is minimal \( \beta \) such that \( y^\top x - \beta \) underestimates \( f \).
- Therefore \( f^{**}(x) \) is the maximum, over all \( y \), of affine underestimates \( y^\top x - \beta \) of \( f \).
\( f^*(y) = \sup_x (y^T x - f(x)) \)

Properties of the Conjugate:

1. \( f^{**} = f \) when \( f \) is convex
2. \( f^* \) is a convex function
3. \( xy \leq f(x) + f^*(y) \) for all \( x, y \in \mathbb{R}^n \) (Fenchel’s Inequality)

Additional notes:

- \( f^{**}(x) = \max_y y^T x - f^*(y) \) when \( f \) is convex
- For fixed \( y \), \( f^*(y) \) is minimal \( \beta \) such that \( y^T x - \beta \) underestimates \( f \).
- Therefore \( f^{**}(x) \) is the maximum, over all \( y \), of affine underestimates \( y^T x - \beta \) of \( f \).
- By our characterization early in this lecture, this is equal to \( f \).
Examples

- **Affine function**: \( f(x) = ax + b \). Conjugate has \( f^*(a) = -b \), and \( \infty \) elsewhere.

- \( f(x) = x^2/2 \) is self-conjugate.

- **Exponential**: \( f(x) = e^x \). Conjugate has \( f^*(y) = y \log y - y \) for \( y \in \mathbb{R}_+ \), and \( \infty \) elsewhere.

- **Quadratic**: \( f(x) = \frac{1}{2} x^\top Q x \) with \( Q \succeq 0 \). Self conjugate.

- **Log-sum-exp**: \( f(x) = \log(\sum_i e^{x_i}) \). Conjugate has \( f^*(y) = \sum_i y_i \log y_i \) for \( y \succeq 0 \) and \( 1^\top y = 1 \), \( \infty \) otherwise.