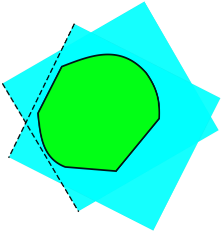
CS675: Convex and Combinatorial Optimization
Spring 2018
Duality of Convex Sets and Functions

Instructor: Shaddin Dughmi

Outline

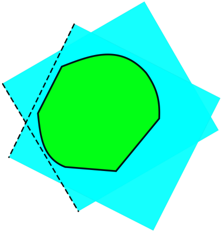
Duality Correspondances



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)

Duality Correspondances

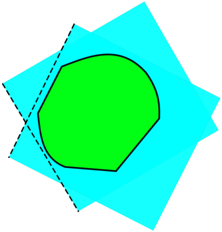


There are two equivalent ways to represent a convex set

- 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.

Duality Correspondances



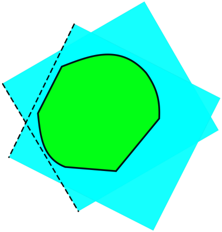
There are two equivalent ways to represent a convex set

- 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.

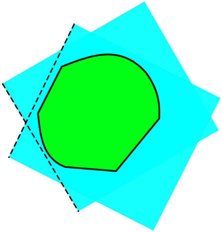
Definition

"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.

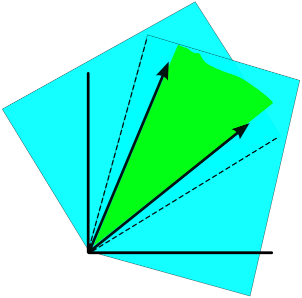


Theorem

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

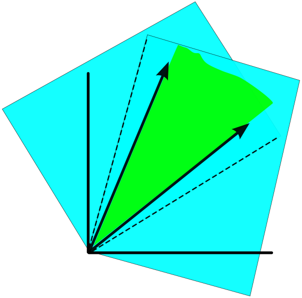
Proof

- 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 $x \notin \bigcap_{H \in \mathcal{H}} H$



Theorem

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

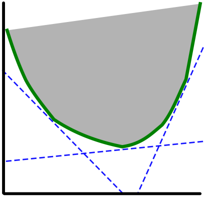


Theorem

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

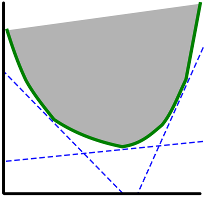
Proof

- For every non-homogeneous halfspace $a^\intercal x \leq b$ containing K, the smaller homogeneous halfspace $a^\intercal x \leq 0$ contains K as well.
- Therefore, can discard non-homogeneous halfspaces when taking the intersection



Theorem

A convex function is the point-wise supremum of all affine functions under-estimating it everywhere.



Theorem

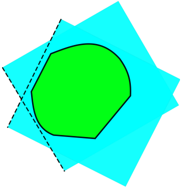
A convex function is the point-wise supremum of all affine functions under-estimating it everywhere.

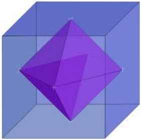
Proof

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

Outline

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:

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

Note

- Every halfspace $a^\intercal x \leq b$ with $b \neq 0$ can be written as a "normalized" inequality $y^\intercal x \leq 1$, by dividing by b.
- S° can be thought of as the normalized representations of halfspaces containing S.

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

Properties of the Polar

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

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

Properties of the Polar

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

2. Follows from representation as intersection of halfspaces
3. S contains an ϵ-ball centered at the origin, so $||y|| \leq 1/\epsilon$ for all $y \in S^\circ$.

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

Properties of the Polar

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

1.
   - Easy to see that $S \subseteq S^{\circ\circ}$
   - Take $x_\circ \notin S$, by SSHT and $0 \in S$, there is a halfspace $z^\intercal x \le 1$ containing S but not $x_\circ$ (i.e. $z^\intercal x_\circ > 1$)
   - $z \in S^\circ$, therefore $x_\circ \notin S^{\circ\circ}$

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

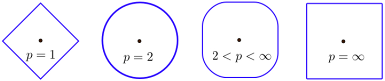
Properties of the Polar

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

Note

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

Examples

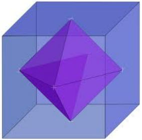


The unit sphere for different metrics: $\|x\|_{l_p} = 1$ in $\mathbb{R}^2$.

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 ∞-norm ball
- More generally, the p-norm ball is dual to the q-norm ball, where $\frac{1}{p} + \frac{1}{q} = 1$

Examples

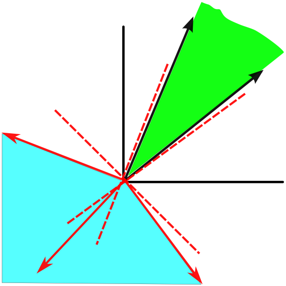


Polytopes

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

- Facets of P correspond to vertices of P°.
- Dually, vertices of P correspond to facets of P°.

Polar Duality of Convex Cones



Polar duality takes a simplified form when applied to cones

Polar

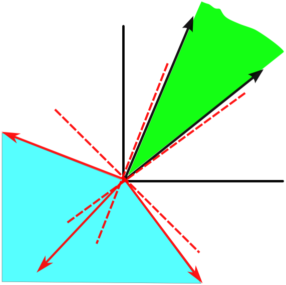
The polar of a closed convex cone K is given by

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

Note

- If halfspace $y^\intercal x \leq b$ contains K, then so does smaller $y^\intercal x \leq 0$.
- K° represents all homogeneous halfspaces containing K.

Polar Duality of Convex Cones



Polar duality takes a simplified form when applied to cones

Polar

The polar of a closed convex cone K is given by

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

Dual Cone

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

$$K^* = \{y : y^\intercal x \ge 0 \text{ for all } x \in K\}$$

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

Properties of the Polar Cone

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

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

Properties of the Polar Cone

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

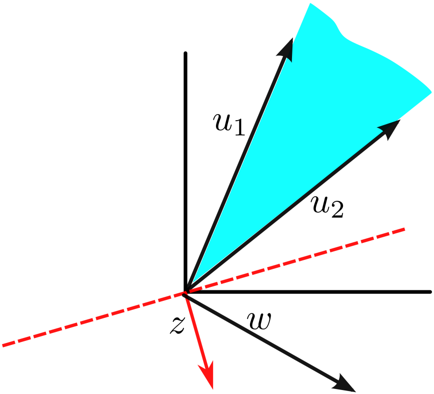
1. Same as before
2. Intersection of homogeneous halfspaces

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

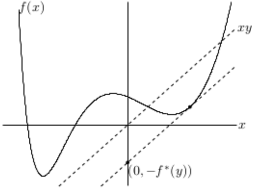
Let K be a closed convex cone and let $w \notin K$. There is $z \in \mathbb{R}^n$ such that $z^\intercal x \leq 0$ for all $x \in K$, and $z^\intercal w > 0$.



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

Outline

Conjugation of Convex Functions



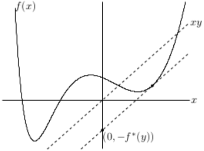
Conjugate

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

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

Note

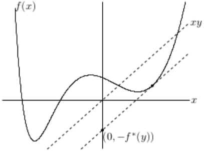
- $f^*(y)$ is the minimal value of β such that the affine function $y^T x - \beta$ underestimates $f(x)$ everywhere.
- Equivalently, the distance we need to lower the hyperplane $y^\intercal x - t = 0$ in order to get a supporting hyperplane to $\mathbf{epi}\, f$.
- $y^\intercal x - t = f^*(y)$ are the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



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

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)

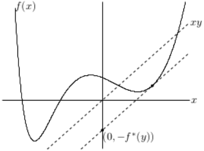


$$f^*(y) = \sup_x (y^\intercal x - f(x))$$

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)

2. Supremum of affine functions of y
3. By definition of $f^*(y)$

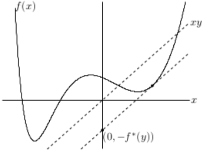


$$f^*(y) = \sup_x (y^\intercal x - f(x))$$

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.
   - $f^{**}(x) = \max_y y^\intercal x - f^*(y)$ when f is convex



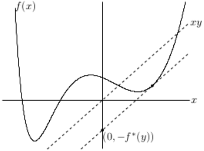
$$f^*(y) = \sup_x (y^\intercal x - f(x))$$

Properties of the Conjugate

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

1

- $f^{**}(x) = \max_y y^\intercal x - f^*(y)$ when f is convex
- For fixed y, $f^*(y)$ is minimal β such that $y^\intercal x - \beta$ underestimates f.

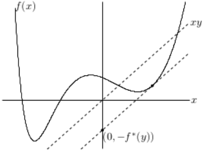


$$f^*(y) = \sup_x (y^\intercal x - f(x))$$

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



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

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.
- $f^{**}(x) = \max_y y^\intercal x - f^*(y)$ when f is convex
- For fixed y, $f^*(y)$ is minimal β such that $y^\intercal x - \beta$ underestimates f.
- Therefore $f^{**}(x)$ is the maximum, over all y, of affine underestimates $y^\intercal x - \beta$ 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 ∞ 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 ∞ elsewhere.
- Quadratic: $f(x) = \frac{1}{2} x^\intercal Q x$ 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 $1^\intercal y = 1$, ∞ otherwise.