CS599: Convex and Combinatorial Optimization
Fall 2013
Lecture 11: Duality of Convex Optimization Problems

Instructor: Shaddin Dughmi

Announcements

- Yu's office hours changed to Friday 4pm-5pm
- Solutions to HW1 should be posted soon.
- HW2 coming soon
- This week: Convex Optimization Duality
  - Read all of B&V Chapter 5.

Outline

1 The Lagrange Dual Problem

2 Duality

Recall: Optimization Problem in Standard Form

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \leq 0, \quad \text{for } i = 1, \ldots, m. \\
& h_i(x) = 0, \quad \text{for } i = 1, \ldots, k.
\end{array}
$$

- For convex optimization problems in standard form, f_i is convex and h_i is affine.
- Let $\mathcal{D}$ denote the domain of all these functions (i.e. when their value is finite)

Recall: Optimization Problem in Standard Form

$$
\begin{array}{lll}
\text{minimize} & f_0(x) & \\
\text{subject to} & f_i(x) \leq 0, & \text{for } i = 1, \ldots, m. \\
& h_i(x) = 0, & \text{for } i = 1, \ldots, k.
\end{array}
$$

- For convex optimization problems in standard form, f_i is convex and h_i is affine.
- Let $\mathcal{D}$ denote the domain of all these functions (i.e. when their value is finite)

This Lecture + Next

We will develop duality theory for convex optimization problems, generalizing linear programming duality.

Running Example: Linear Programming

We have already seen the standard form LP below

$$\begin{aligned} \text{maximize} \quad & c^\intercal x \\ \text{subject to} \quad & Ax \preceq b \\ & x \succeq 0 \end{aligned}$$

$$\begin{aligned} \text{minimize} \quad & -c^\intercal x \\ \text{subject to} \quad & Ax - b \preceq 0 \\ & -x \preceq 0 \end{aligned}$$

Running Example: Linear Programming

We have already seen the standard form LP below

$$\begin{aligned} \text{maximize} \quad & c^\intercal x \\ \text{subject to} \quad & Ax \preceq b \\ & x \succeq 0 \end{aligned} \qquad \qquad \begin{aligned} \text{minimize} \quad & -c^\intercal x \\ \text{subject to} \quad & Ax - b \preceq 0 \\ & -x \preceq 0 \end{aligned}$$

Along the way, we will recover the following standard form dual

$$\begin{aligned} \text{minimize} \quad & y^\intercal b \\ \text{subject to} \quad & A^\intercal y \succeq c \\ & y \succeq 0 \end{aligned}$$

The Lagrangian

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad \text{for } i = 1, \ldots, m. \\ & h_i(x) = 0, \quad \text{for } i = 1, \ldots, k. \end{aligned}$$

Basic idea of Lagrangian duality is to relax/soften the constraints by replacing each with a linear “penalty term” or “cost” in the objective.

The Lagrangian

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & \text{for } i = 1, \ldots, m. \\ & h_i(x) = 0, & \text{for } i = 1, \ldots, k. \end{array}$$

Basic idea of Lagrangian duality is to relax/soften the constraints by replacing each with a linear “penalty term” or “cost” in the objective.

The Lagrangian Function

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{k} \nu_i h_i(x)$$

- λ_i is Lagrange Multiplier for i'th inequality constraint
  - Required to be nonnegative
- ν_i is Lagrange Multiplier for i'th equality constraint
  - Allowed to be of arbitrary sign

The Lagrange Dual Function

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & \text{for } i = 1, \ldots, m. \\ & h_i(x) = 0, & \text{for } i = 1, \ldots, k. \end{array}$$

The Lagrange dual function gives the optimal value of the primal problem subject to the softened constraints

The Lagrange Dual Function

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & \text{for } i = 1, \ldots, m. \\ & h_i(x) = 0, & \text{for } i = 1, \ldots, k. \end{array}$$

The Lagrange dual function gives the optimal value of the primal problem subject to the softened constraints

The Lagrange Dual Function

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{k} \nu_i h_i(x) \right)$$

- Observe: g is a concave function of the Lagrange multipliers
- We will see: Its quite common for the Lagrange dual to be unbounded $(-\infty)$ for some λ and ν
- By convention, domain of g is (λ, ν) s.t. $g(\lambda, \nu) > -\infty$

Langrange Dual of LP

$$\begin{array}{ll} \text{minimize} & -c^\intercal x \\ \text{subject to} & Ax - b \preceq 0 \\ & -x \preceq 0 \end{array}$$

First, the Lagrangian function

$$\begin{aligned} L(x, \lambda) &= -c^\intercal x + \lambda_1^\intercal (Ax - b) - \lambda_2^\intercal x \\ &= (A^\intercal \lambda_1 - c - \lambda_2)^\intercal x - \lambda_1^\intercal b \end{aligned}$$

Langrange Dual of LP

$$
\begin{array}{ll}
\text{minimize} & -c^\intercal x \\
\text{subject to} & Ax - b \preceq 0 \\
& -x \preceq 0
\end{array}
$$

First, the Lagrangian function

$$
\begin{aligned}
L(x, \lambda) &= -c^\intercal x + \lambda_1^\intercal (Ax - b) - \lambda_2^\intercal x \\
&= (A^\intercal \lambda_1 - c - \lambda_2)^\intercal x - \lambda_1^\intercal b
\end{aligned}
$$

And the Lagrange Dual

$$
\begin{aligned}
g(\lambda) &= \inf_x L(x, \lambda) \\
&= \begin{cases} -\infty & \text{if } A^\intercal \lambda_1 - c - \lambda_2 \neq 0 \\ -\lambda_1^\intercal b & A^\intercal \lambda_1 - c - \lambda_2 = 0 \end{cases}
\end{aligned}
$$

Langrange Dual of LP

$$\begin{array}{ll} \text{minimize} & -c^\intercal x \\ \text{subject to} & Ax - b \preceq 0 \\ & -x \preceq 0 \end{array}$$

First, the Lagrangian function

$$\begin{aligned} L(x, \lambda) &= -c^\intercal x + \lambda_1^\intercal (Ax - b) - \lambda_2^\intercal x \\ &= (A^\intercal \lambda_1 - c - \lambda_2)^\intercal x - \lambda_1^\intercal b \end{aligned}$$

And the Lagrange Dual

$$\begin{aligned} g(\lambda) &= \inf_x L(x, \lambda) \\ &= \begin{cases} -\infty & \text{if } A^\intercal \lambda_1 - c - \lambda_2 \neq 0 \\ -\lambda_1^\intercal b & A^\intercal \lambda_1 - c - \lambda_2 = 0 \end{cases} \end{aligned}$$

So we restrict the domain of g to λ satisfying $A^\intercal \lambda_1 - c - \lambda_2 = 0$

Interpretation: “Soft” Lower Bound

$$\begin{array}{lll} \text{min} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & \text{for } i = 1, \ldots, m. \\ & h_i(x) = 0, & \text{for } i = 1, \ldots, k. \end{array}$$

The Lagrange Dual Function

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{k} \nu_i h_i(x) \right)$$

Interpretation: "Soft" Lower Bound

$$\begin{array}{lll} \text{min} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & \text{for } i = 1, \ldots, m. \\ & h_i(x) = 0, & \text{for } i = 1, \ldots, k. \end{array}$$

The Lagrange Dual Function

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{k} \nu_i h_i(x) \right)$$

Fact

$g(\lambda, \nu)$ is a lowerbound on OPT(primal) for every $\lambda \succeq 0$ and $\nu \in \mathbb{R}^k$.

Interpretation: "Soft" Lower Bound

$$
\begin{array}{lll}
\text{min} & f_0(x) & \\
\text{subject to} & f_i(x) \leq 0, & \text{for } i = 1, \ldots, m. \\
 & h_i(x) = 0, & \text{for } i = 1, \ldots, k.
\end{array}
$$

The Lagrange Dual Function

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{k} \nu_i h_i(x) \right)$$

Fact

$g(\lambda, \nu)$ is a lowerbound on OPT(primal) for every $\lambda \succeq 0$ and $\nu \in \mathbb{R}^k$.

Proof

- Every primal feasible x incurs nonpositive penalty by $L(x, \lambda, \nu)$
- Therefore, $L(x^*, \lambda, \nu) \leq f_0(x^*)$
- So $g(\lambda, \nu) \leq f_0(x^*) = OPT(Primal)$

Interpretation: "Soft" Lower Bound

$$\begin{array}{lll} \text{min} & f_0(x) & \\ \text{subject to} & f_i(x) \le 0, & \text{for } i = 1, \ldots, m. \\ & h_i(x) = 0, & \text{for } i = 1, \ldots, k. \end{array}$$

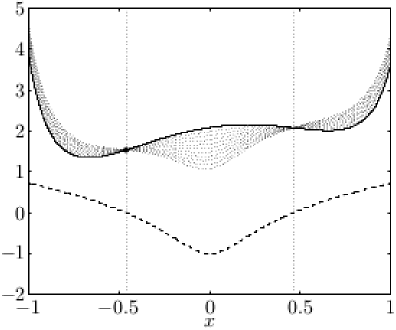
The Lagrange Dual Function

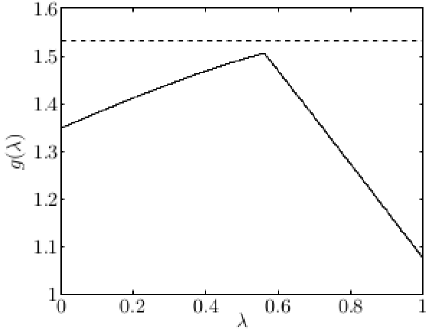
$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{k} \nu_i h_i(x) \right)$$

Interpretation

- A "hard" feasibility constraint can be thought of as imposing a penalty of $+\infty$ if violated
- Lagrangian imposes a "soft" linear penalty for violating a constraint, and a reward for slack
- Lagrange dual finds the optimal subject to these soft constraints

Interpretation: "Soft" Lower Bound

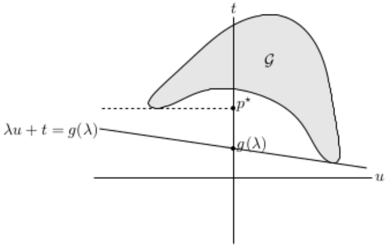




Interpretation: Geometric

Most easily visualized in the presence of a single inequality constraint

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_1(x) \le 0 \end{array}$$

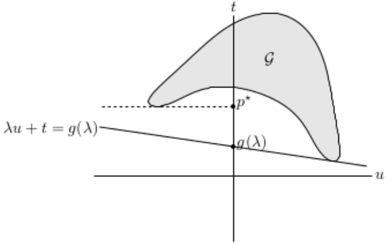


- Let $\mathcal{G}$ be attainable constraint/objective function value tuples
  - i.e. $(u, t) \in \mathcal{G}$ if there is an x such that $f_1(x) = u$ and $f_0(x) = t$
- $p^* = \inf \{t : (u, t) \in \mathcal{G}, u \le 0\}$
- $g(\lambda) = \inf \{\lambda u + t \ : \ (u, t) \in \mathcal{G}\}$

Interpretation: Geometric

Most easily visualized in the presence of a single inequality constraint

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_1(x) \le 0 \end{array}$$



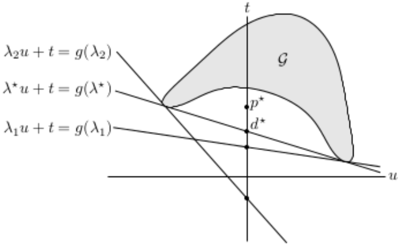
- Let $\mathcal{G}$ be attainable constraint/objective function value tuples
  - i.e. $(u, t) \in \mathcal{G}$ if there is an x such that $f_1(x) = u$ and $f_0(x) = t$
- $p^* = \inf \{t : (u, t) \in \mathcal{G}, u \le 0\}$
- $g(\lambda) = \inf \{\lambda u + t \ : \ (u, t) \in \mathcal{G}\}$

- $\lambda u + t = g(\lambda)$ is a supporting hyperplane to $\mathcal{G}$ pointing northeast
- Must intersect vertical axis below p^*
- Therefore $g(\lambda) \le p^*$

The Lagrange Dual Problem

This is the problem of finding the best lower bound on OPT(primal) implied by the Lagrange dual function

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$



- Note: this is a convex optimization problem, regardless of whether primal problem was convex
- By convention, sometimes we add "dual feasibility" constraints to impose "nontrivial" lowerbounds (i.e. $g(\lambda, \nu) \geq -\infty$)
- (λ^*, ν^*) solving the above are referred to as the dual optimal solution

Langrange Dual Problem of LP

maximize	$c^\intercal x$
subject to	$Ax \preceq b$
	$x \succeq 0$

minimize	$-c^\intercal x$
subject to	$Ax - b \preceq 0$
	$-x \preceq 0$

Recall

Our Lagrange dual function for the above LP (to the right), defined over the domain $A^\intercal \lambda_1 - c - \lambda_2 = 0$.

$$g(\lambda) = -\lambda_1^\intercal b$$

Langrange Dual Problem of LP

$$\begin{array}{ll} \text{maximize} & c^\intercal x \\ \text{subject to} & Ax \preceq b \\ & x \succeq 0 \end{array}$$

$$\begin{array}{ll} \text{minimize} & -c^\intercal x \\ \text{subject to} & Ax - b \preceq 0 \\ & -x \preceq 0 \end{array}$$

Recall

Our Lagrange dual function for the above LP (to the right), defined over the domain $A^\intercal \lambda_1 - c - \lambda_2 = 0$.

$$g(\lambda) = -\lambda_1^\intercal b$$

The Lagrange dual problem can then be written as

$$\begin{array}{ll} \text{maximize} & -\lambda_1^\intercal b \\ \text{subject to} & A^\intercal \lambda_1 - c - \lambda_2 = 0 \\ & \\ & \lambda \succeq 0 \end{array}$$

Langrange Dual Problem of LP

$$\begin{array}{ll} \text{maximize} & c^\intercal x \\ \text{subject to} & Ax \preceq b \\ & x \succeq 0 \end{array}$$

$$\begin{array}{ll} \text{minimize} & -c^\intercal x \\ \text{subject to} & Ax - b \preceq 0 \\ & -x \preceq 0 \end{array}$$

Recall

Our Lagrange dual function for the above LP (to the right), defined over the domain $A^\intercal \lambda_1 - c - \lambda_2 = 0$.

$$g(\lambda) = -\lambda_1^\intercal b$$

The Lagrange dual problem can then be written as

$$\begin{array}{ll} \text{maximize} & -\lambda_1^\intercal b \\ \text{subject to} & \sout{A^\intercal \lambda_1 - c - \lambda_2 = 0} \\ & A^\intercal \lambda_1 \succeq c \\ & \lambda \succeq 0 \end{array}$$

Langrange Dual Problem of LP

$$\begin{array}{ll} \text{maximize} & c^\intercal x \\ \text{subject to} & Ax \preceq b \\ & x \succeq 0 \end{array}$$

$$\begin{array}{ll} \text{minimize} & -c^\intercal x \\ \text{subject to} & Ax - b \preceq 0 \\ & -x \preceq 0 \end{array}$$

Recall

Our Lagrange dual function for the above LP (to the right), defined over the domain $A^\intercal \lambda_1 - c - \lambda_2 = 0$.

$$g(\lambda) = -\lambda_1^\intercal b$$

The Lagrange dual problem can then be written as

$$\begin{array}{ll} \text{minimize} & y^\intercal b \\ \text{subject to} & A^\intercal y \succeq c \\ & y \succeq 0 \end{array}$$

$$\begin{array}{ll} \text{maximize} & -\lambda_1^\intercal b \\ \text{subject to} & \cancel{A^\intercal \lambda_1 - c - \lambda_2 = 0} \\ & A^\intercal \lambda_1 \succeq c \\ & \lambda \succeq 0 \end{array}$$

Another Example: Conic Optimization Problem

$$\begin{array}{ll} \text{minimize} & c^\intercal x \\ \text{subject to} & Ax = b \\ & x \in K \end{array}$$

- $x \in K$ can equivalently be written as $z^\intercal x \leq 0, \forall z \in K^\circ$

$$\begin{aligned} L(x, \lambda, \nu) &= c^\intercal x + \nu^\intercal (Ax - b) + \sum_{z \in K^\circ} \lambda_z \cdot z^\intercal x \\ &= (c - A^\intercal \nu + \sum_{z \in K^\circ} \lambda_z \cdot z)^\intercal x + \nu^\intercal b \end{aligned}$$

Another Example: Conic Optimization Problem

$$\begin{array}{ll} \text{minimize} & c^\intercal x \\ \text{subject to} & Ax = b \\ & x \in K \end{array}$$

- $x \in K$ can equivalently be written as $z^\intercal x \leq 0$, $\forall z \in K^\circ$

$$\begin{aligned} L(x, \lambda, \nu) &= c^\intercal x + \nu^\intercal (Ax - b) + \sum_{z \in K^\circ} \lambda_z \cdot z^\intercal x \\ &= (c - A^\intercal \nu + \sum_{z \in K^\circ} \lambda_z \cdot z)^\intercal x + \nu^\intercal b \end{aligned}$$

- Can think of $\lambda \succeq 0$ as choosing some $s \in K^\circ$

$$L(x, s, \nu) = (c - A^\intercal \nu + s)^\intercal x + \nu^\intercal b$$

Another Example: Conic Optimization Problem

$$\begin{array}{ll} \text{minimize} & c^\intercal x \\ \text{subject to} & Ax = b \\ & x \in K \end{array}$$

- $x \in K$ can equivalently be written as $z^\intercal x \leq 0$, $\forall z \in K^\circ$

$$\begin{aligned} L(x, \lambda, \nu) &= c^\intercal x + \nu^\intercal (Ax - b) + \sum_{z \in K^\circ} \lambda_z \cdot z^\intercal x \\ &= (c - A^\intercal \nu + \sum_{z \in K^\circ} \lambda_z \cdot z)^\intercal x + \nu^\intercal b \end{aligned}$$

- Can think of $\lambda \succeq 0$ as choosing some $s \in K^\circ$

$$L(x, s, \nu) = (c - A^\intercal \nu + s)^\intercal x + \nu^\intercal b$$

- Lagrange dual function $g(s, \nu)$ is bounded when coefficient of x is zero, in which case it has value $\nu^\intercal b$

Another Example: Conic Optimization Problem

minimize	$c^\intercal x$
subject to	$Ax = b$
	$x \in K$

maximize	$\nu^\intercal b$
subject to	$A^\intercal \nu - c \in K^\circ$

- $x \in K$ can equivalently be written as $z^\intercal x \leq 0$, $\forall z \in K^\circ$

$$L(x, \lambda, \nu) = c^\intercal x + \nu^\intercal (Ax - b) + \sum_{z \in K^\circ} \lambda_z \cdot z^\intercal x$$
$$= (c - A^\intercal \nu + \sum_{z \in K^\circ} \lambda_z \cdot z)^\intercal x + \nu^\intercal b$$

- Can think of $\lambda \succeq 0$ as choosing some $s \in K^\circ$

$$L(x, s, \nu) = (c - A^\intercal \nu + s)^\intercal x + \nu^\intercal b$$

- Lagrange dual function $g(s, \nu)$ is bounded when coefficient of x is zero, in which case it has value $\nu^\intercal b$

Outline

1 The Lagrange Dual Problem

2 Duality

Weak Duality

Primal Problem

min $f_0(x)$
s.t.
$f_i(x) \leq 0, \quad \forall i = 1, \dots, m.$
$h_i(x) = 0, \quad \forall i = 1, \dots, k.$

Dual Problem

max $g(\lambda, \nu)$
s.t.
$\lambda \succeq 0$

Weak Duality

Primal Problem

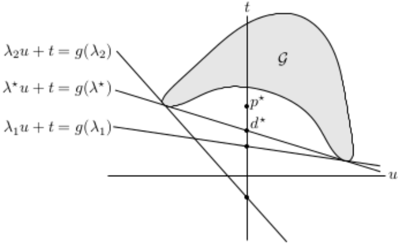
min $f_0(x)$
s.t.
$f_i(x) \leq 0, \quad \forall i = 1, \ldots, m.$
$h_i(x) = 0, \quad \forall i = 1, \ldots, k.$

Dual Problem

max $g(\lambda, \nu)$
s.t.
$\lambda \succeq 0$

Weak Duality

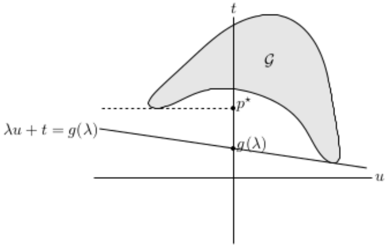
$OPT(dual) \leq OPT(primal)$.



- We have already argued holds for every optimization problem
- Duality Gap: difference between optimal dual and primal values

Recall: Geometric Interpretation of Weak Duality

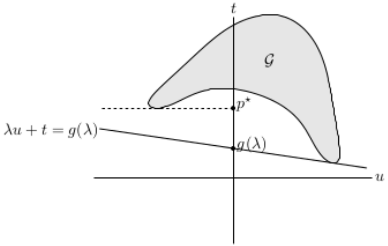
minimize	$f_0(x)$
subject to	$f_1(x) \leq 0$



- Let $\mathcal{G}$ be attainable constraint/objective function value tuples
  - i.e. $(u, t) \in \mathcal{G}$ if there is an x such that $f_1(x) = u$ and $f_0(x) = t$
- $p^* = \inf\{t : (u, t) \in \mathcal{G}, u \leq 0\}$
- $g(\lambda) = \inf\{\lambda u + t \;:\; (u, t) \in \mathcal{G}\}$

Recall: Geometric Interpretation of Weak Duality

minimize	$f_0(x)$
subject to	$f_1(x) \leq 0$



- Let $\mathcal{G}$ be attainable constraint/objective function value tuples
  - i.e. $(u, t) \in \mathcal{G}$ if there is an x such that $f_1(x) = u$ and $f_0(x) = t$
- $p^* = \inf \{t : (u, t) \in \mathcal{G}, u \leq 0\}$
- $g(\lambda) = \inf \{\lambda u + t \ : \ (u, t) \in \mathcal{G}\}$

Fact

The equation $\lambda u + t = g(\lambda)$ defines a supporting hyperplane to $\mathcal{G}$, intersecting t axis at $g(\lambda) \leq p^*$.

Strong Duality

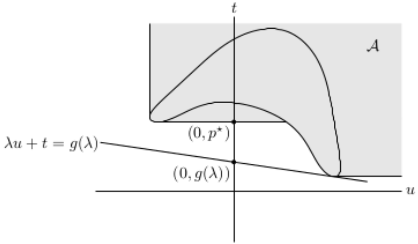
Strong Duality

We say strong duality holds if $OPT(dual) = OPT(primal)$.

- Equivalently: there exists a setting of Lagrange multipliers so that $g(\lambda, \nu)$ gives a tight lowerbound on primal optimal value.
- In general, does not hold for non-convex optimization problems
- Usually, but not always, holds for convex optimization problems.
  - Mild assumptions, such as Slater's condition, needed.

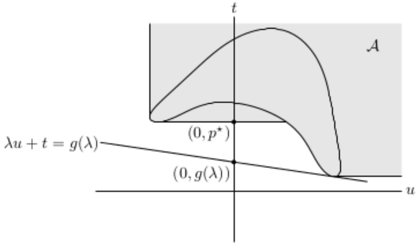
Geometric Proof of Strong Duality

minimize $\quad f_0(x)$
subject to $\quad f_1(x) \leq 0$



- Let $\mathcal{A}$ be everything northeast (i.e. "worse") than $\mathcal{G}$
  - i.e. $(u, t) \in \mathcal{A}$ if there is an x such that $f_1(x) \leq u$ and $f_0(x) \leq t$
- $p^* = \inf \{t : (0, t) \in \mathcal{A}\}$
- $g(\lambda) = \inf \{\lambda u + t \ : \ (u, t) \in \mathcal{A}\}$

Geometric Proof of Strong Duality



minimize	$f_0(x)$
subject to	$f_1(x) \leq 0$

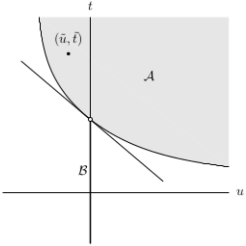
- Let $\mathcal{A}$ be everything northeast (i.e. "worse") than $\mathcal{G}$
  - i.e. $(u, t) \in \mathcal{A}$ if there is an x such that $f_1(x) \leq u$ and $f_0(x) \leq t$
- $p^* = \inf \{t : (0, t) \in \mathcal{A}\}$
- $g(\lambda) = \inf \{\lambda u + t \ : \ (u, t) \in \mathcal{A}\}$

Fact

The equation $\lambda u + t = g(\lambda)$ defines a supporting hyperplane to $\mathcal{G}$, intersecting t axis at $g(\lambda) \leq p^*$.

Geometric Proof of Strong Duality

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_1(x) \leq 0 \end{array}$$

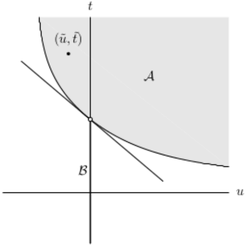


Fact

When f_0 and f_1 are convex, $\mathcal{A}$ is convex.

Geometric Proof of Strong Duality

minimize $\quad f_0(x)$
subject to $\quad f_1(x) \le 0$



Fact

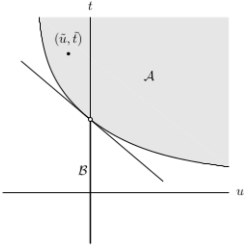
When f_0 and f_1 are convex, $\mathcal{A}$ is convex.

Proof

- Assume (u, t) and (u', t') are in $\mathcal{A}$

Geometric Proof of Strong Duality

minimize $f_0(x)$
subject to $f_1(x) \leq 0$



Fact

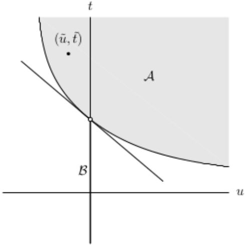
When f_0 and f_1 are convex, $\mathcal{A}$ is convex.

Proof

- Assume (u, t) and (u', t') are in $\mathcal{A}$
- $\exists x, x'$ with $(f_1(x), f_0(x)) \leq (u, t)$ and $(f_1(x'), f_0(x')) \leq (u', t')$.

Geometric Proof of Strong Duality

minimize $\quad f_0(x)$
subject to $\quad f_1(x) \le 0$



Fact

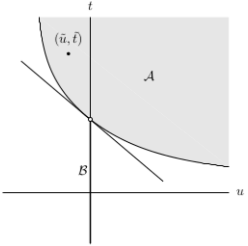
When f_0 and f_1 are convex, $\mathcal{A}$ is convex.

Proof

- Assume (u, t) and (u', t') are in $\mathcal{A}$
- $\exists x, x'$ with $(f_1(x), f_0(x)) \le (u, t)$ and $(f_1(x'), f_0(x')) \le (u', t')$.
- By Jensen's inequality $(f_1(\frac{x+x'}{2}), f_0(\frac{x+x'}{2})) \le (\frac{u+u'}{2}, \frac{t+t'}{2})$

Geometric Proof of Strong Duality

minimize $\quad f_0(x)$
subject to $\quad f_1(x) \leq 0$



Fact

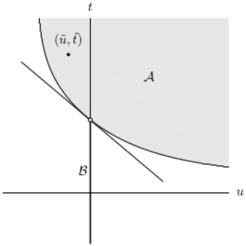
When f_0 and f_1 are convex, $\mathcal{A}$ is convex.

Proof

- Assume (u, t) and (u', t') are in $\mathcal{A}$
- $\exists x, x'$ with $(f_1(x), f_0(x)) \leq (u, t)$ and $(f_1(x'), f_0(x')) \leq (u', t')$.
- By Jensen's inequality $(f_1(\frac{x+x'}{2}), f_0(\frac{x+x'}{2})) \leq (\frac{u+u'}{2}, \frac{t+t'}{2})$
- Therefore, midpoint of (u, t) and (u', t') also in $\mathcal{A}$.

Geometric Proof of Strong Duality

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_1(x) \leq 0 \end{array}$$

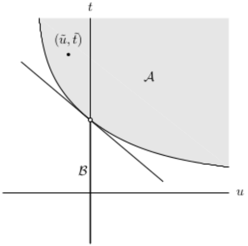


Theorem (Informal)

There is a choice of λ so that $g(\lambda) = p^*$. Therefore, strong duality holds.

Geometric Proof of Strong Duality

minimize $f_0(x)$
subject to $f_1(x) \leq 0$



Theorem (Informal)

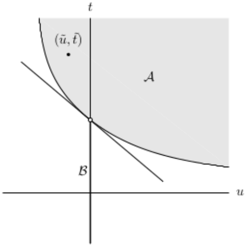
There is a choice of λ so that $g(\lambda) = p^*$. Therefore, strong duality holds.

Proof

- Recall $(0, p^*) \in \mathcal{A}$
- By the supporting hyperplane theorem, there is a supporting hyperplane to $\mathcal{A}$ at $(0, p^*)$
- Direction of the supporting hyperplane gives us an appropriate λ

I Lied (A little)

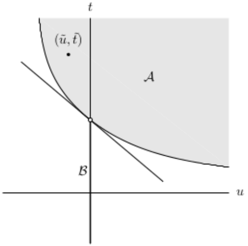
minimize $\quad f_0(x)$
subject to $\quad f_1(x) \leq 0$



- In our proof, we ignored a technicality that can prevent strong duality from holding.

I Lied (A little)

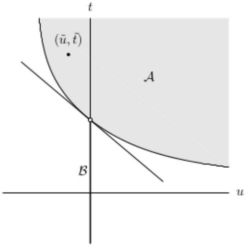
minimize $\quad f_0(x)$
subject to $\quad f_1(x) \leq 0$



- In our proof, we ignored a technicality that can prevent strong duality from holding.
- If our supporting hyperplane H at $(0, p^*)$ is vertical, then no finite λ exists such that $(\lambda, 1)$ is normal to H.

I Lied (A little)

minimize $f_0(x)$
subject to $f_1(x) \leq 0$



- In our proof, we ignored a technicality that can prevent strong duality from holding.
- If our supporting hyperplane H at $(0, p^*)$ is vertical, then no finite λ exists such that $(\lambda, 1)$ is normal to H.
- Somewhat counterintuitively, this can happen even in simple convex optimization problems (though its somewhat rare in practice)

Violation of Strong Duality

$$\begin{array}{ll} \text{minimize} & e^{-x} \\ \text{subject to} & \frac{x^2}{y} \leq 0 \\ & y \geq 1 \end{array}$$

- Problem is convex, with feasible region given by $x = 0$ and $y \geq 1$
- Optimal value is 1, at $x = 0$ and $y = 1$

Violation of Strong Duality

$$\begin{array}{ll} \text{minimize} & e^{-x} \\ \text{subject to} & \frac{x^2}{y} \leq 0 \\ & y \geq 1 \end{array}$$

- Problem is convex, with feasible region given by $x = 0$ and $y \geq 1$
- Optimal value is 1, at $x = 0$ and $y = 1$
- Consider $\mathcal{A}$ restricted to the objective and the first constraint

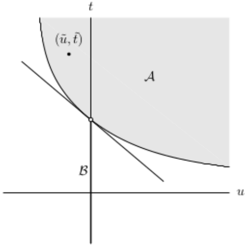
Violation of Strong Duality

$$\begin{array}{ll} \text{minimize} & e^{-x} \\ \text{subject to} & \frac{x^2}{y} \leq 0 \\ & y \geq 1 \end{array}$$

- Problem is convex, with feasible region given by $x = 0$ and $y \geq 1$
- Optimal value is 1, at $x = 0$ and $y = 1$
- Consider $\mathcal{A}$ restricted to the objective and the first constraint
- $\mathcal{A} = \mathbb{R}^2_{++} \bigcup \{0\} \times [1, \infty]$
- Therefore, any supporting hyperplane to $\mathcal{A}$ at $(0, 1)$ must be vertical.

Slater's Condition

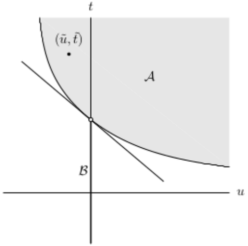
There exists a point $x \in \mathcal{D}$ where all inequality constraints are strictly satisfied (i.e. $f_i(x) < 0$). I.e. the optimization problem is strictly feasible.



- A sufficient condition for strong duality.
- Forces supporting hyperplane to be non-vertical

Slater's Condition

There exists a point $x \in \mathcal{D}$ where all inequality constraints are strictly satisfied (i.e. $f_i(x) < 0$). I.e. the optimization problem is strictly feasible.



- A sufficient condition for strong duality.
- Forces supporting hyperplane to be non-vertical
- Can be weakened to requiring strict feasibility only of non-affine constraints