Duality of Convex Optimization Problems

Instructor: Shaddin Dughmi
Outline

1. The Lagrange Dual Problem

2. Duality
Recall: Optimization Problem in Standard Form

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

- 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

minimize \( f_0(x) \)
subject to \( f_i(x) \leq 0, \) for \( i = 1, \ldots, m. \)
\( h_i(x) = 0, \) for \( i = 1, \ldots, k. \)

- 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.
We have already seen the standard form LP below

\[
\begin{align*}
\text{maximize} & \quad c^T x \\
\text{subject to} & \quad Ax \leq b \\
& \quad x \geq 0
\end{align*}
\]

\[
\begin{align*}
\text{minimize} & \quad -c^T x \\
\text{subject to} & \quad Ax - b \leq 0 \\
& \quad -x \leq 0
\end{align*}
\]
We have already seen the standard form LP below

\[
\begin{align*}
\text{maximize} & \quad c^\top x \\
\text{subject to} & \quad Ax \leq b \\
& \quad x \geq 0
\end{align*}
\]

\[
\begin{align*}
\text{minimize} & \quad -c^\top x \\
\text{subject to} & \quad Ax - b \leq 0 \\
& \quad -x \leq 0
\end{align*}
\]

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

\[
\begin{align*}
\text{minimize} & \quad y^\top b \\
\text{subject to} & \quad A^\top y \geq c \\
& \quad y \geq 0
\end{align*}
\]
The Lagrangian

\[ \text{minimize} \quad f_0(x) \]
\[ \text{subject to} \quad f_i(x) \leq 0, \quad \text{for } i = 1, \ldots, m. \]
\[ h_i(x) = 0, \quad \text{for } i = 1, \ldots, k. \]

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

minimize \( f_0(x) \)
subject to \( f_i(x) \leq 0, \) for \( i = 1, \ldots, m. \)
\( h_i(x) = 0, \) for \( i = 1, \ldots, k. \)

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

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

- \( \lambda_i \) is Lagrange Multiplier for \( i \)'th inequality constraint
  - Required to be nonnegative
- \( \nu_i \) is Lagrange Multiplier for \( i \)'th equality constraint
  - Allowed to be of arbitrary sign

The Lagrange Dual Problem
The Lagrange Dual Function

minimize \( f_0(x) \)
subject to \( f_i(x) \leq 0, \) for \( i = 1, \ldots, m. \)
\( h_i(x) = 0, \) for \( i = 1, \ldots, k. \)

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

minimize \( f_0(x) \)

subject to \( f_i(x) \leq 0, \) for \( i = 1, \ldots, m. \)

\( h_i(x) = 0, \) for \( i = 1, \ldots, k. \)

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

The Lagrange Dual Function

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

- 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 \( \lambda \) and \( \nu \).
- By convention, domain of \( g \) is \((\lambda, \nu)\) s.t. \( g(\lambda, \nu) > -\infty \).
minimize $-c^T x$
subject to $Ax - b \leq 0$
$-x \leq 0$

First, the Lagrangian function

$$L(x, \lambda) = -c^T x + \lambda_1^T (Ax - b) - \lambda_2^T x$$
$$= (A^T \lambda_1 - c - \lambda_2)^T x - \lambda_1^T b$$
First, the Lagrangian function

\[ L(x, \lambda) = -c^\top x + \lambda_1^\top (Ax - b) - \lambda_2^\top x \]
\[ = (A^\top \lambda_1 - c - \lambda_2)^\top x - \lambda_1^\top b \]

And the Lagrange Dual

\[ g(\lambda) = \inf_x L(x, \lambda) \]
\[ = \begin{cases} 
-\infty & \text{if } A^\top \lambda_1 - c - \lambda_2 \neq 0 \\
-\lambda_1^\top b & \text{if } A^\top \lambda_1 - c - \lambda_2 = 0 
\end{cases} \]
First, the Lagrangian function

\[ L(x, \lambda) = -c^\top x + \lambda_1^\top (Ax - b) - \lambda_2^\top x \]
\[ = (A^\top \lambda_1 - c - \lambda_2)^\top x - \lambda_1^\top b \]

And the Lagrange Dual

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

So we restrict the domain of \( g \) to \( \lambda \) satisfying \( A^\top \lambda_1 - c - \lambda_2 = 0 \)
Interpretation: “Soft” Lower Bound

\[
\begin{align*}
\min & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 0, \quad \text{for } i = 1, \ldots, m. \\
& \quad h_i(x) = 0, \quad \text{for } i = 1, \ldots, k.
\end{align*}
\]

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

\[
\begin{align*}
\min \quad & f_0(x) \\
\text{subject to} \quad & f_i(x) \leq 0, \quad \text{for } i = 1, \ldots, m. \\
& h_i(x) = 0, \quad \text{for } i = 1, \ldots, k.
\end{align*}
\]

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 \(\text{OPT(primal)}\) for every \(\lambda \succeq 0\) and \(\nu \in \mathbb{R}^k\).
Interpretation: “Soft” Lower Bound

\[
\begin{align*}
\min & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 0, \quad \text{for } i = 1, \ldots, m. \\
& \quad h_i(x) = 0, \quad \text{for } i = 1, \ldots, k.
\end{align*}
\]

The Lagrange Dual Function

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

Fact

\(g(\lambda, \nu)\) is a lowerbound on \(\text{OPT}(\text{primal})\) 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 \(g(\lambda, \nu) \leq f_0(x^*) = \text{OPT}(\text{Primal})\)
Interpretation: “Soft” Lower Bound

\[
\min \quad f_0(x) \\
\text{subject to} \quad f_i(x) \leq 0, \quad \text{for } i = 1, \ldots, m. \\
\quad h_i(x) = 0, \quad \text{for } i = 1, \ldots, k.
\]

The Lagrange Dual Function

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

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

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_1(x) \leq 0
\end{align*}
\]

Let \( G \) be attainable constraint/objective function value tuples
- i.e. \((u, t) \in G\) if there is an \( x \) such that \( f_1(x) = u \) and \( f_0(x) = t \)

\( p^* = \inf \{ t : (u, t) \in G, u \leq 0 \} \)

\( g(\lambda) = \inf \{ \lambda u + t : (u, t) \in G \} \)
Interpretation: Geometric

Most easily visualized in the presence of a single inequality constraint

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_1(x) \leq 0
\end{align*}
\]

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 \(f_0(x) = t\)
- \(p^* = \inf \{ t : (u, t) \in \mathcal{G}, u \leq 0 \}\)
- \(g(\lambda) = \inf \{ \lambda u + t : (u, t) \in \mathcal{G} \}\)

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

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

\[
\begin{align*}
\text{maximize} & \quad g(\lambda, \nu) \\
\text{subject to} & \quad \lambda \succeq 0
\end{align*}
\]

- 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. \( g(\lambda, \nu) \geq -\infty \))
- \((\lambda^*, \nu^*)\) solving the above are referred to as the dual optimal solution
Langrange Dual Problem of LP

maximize \( c^T x \)  
subject to \( Ax \preceq b \)  
\( x \succeq 0 \)

minimize \( -c^T x \)  
subject to \( Ax - b \preceq 0 \)  
\( -x \succeq 0 \)

Recall

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

\[ g(\lambda) = -\lambda_1^T b \]
**Langrange Dual Problem of LP**

maximize \( c^\top x \) 
subject to \( Ax \preceq b \) 
\( x \succeq 0 \)

minimize \( -c^\top x \) 
subject to \( Ax - b \preceq 0 \) 
\( -x \succeq 0 \)

**Recall**

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

\[ g(\lambda) = -\lambda^\top_1 b \]

The Lagrange dual problem can then be written as

maximize \( -\lambda^\top_1 b \) 
subject to \( A^\top \lambda_1 - c - \lambda_2 = 0 \) 
\( \lambda \succeq 0 \)
Langrange Dual Problem of LP

maximize \[ c^\top x \]
subject to \[ Ax \leq b \]
\[ x \geq 0 \]

minimize \[ -c^\top x \]
subject to \[ Ax - b \leq 0 \]
\[ -x \leq 0 \]

Recall

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

\[ g(\lambda) = -\lambda_1^\top b \]

The Lagrange dual problem can then be written as

maximize \[ -\lambda_1^\top b \]
subject to \[ A^\top \lambda_1 - c - \lambda_2 = 0 \]
\[ A^\top \lambda_1 \geq c \]
\[ \lambda \geq 0 \]
Langrange Dual Problem of LP

maximize \( c^\top x \)
subject to \( Ax \leq b \)
\( x \geq 0 \)

minimize \(-c^\top x\)
subject to \( Ax - b \leq 0 \)
\(-x \leq 0\)

Recall

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

\[ g(\lambda) = -\lambda_1^\top b \]

The Lagrange dual problem can then be written as

minimize \( y^\top b \)
subject to \( A^\top y \geq c \)
\( y \geq 0 \)

maximize \(-\lambda_1^\top b\)
subject to \( A^\top \lambda_1 - c - \lambda_2 = 0 \)
\( A^\top \lambda_1 \geq c \)
\( \lambda \geq 0 \)
Another Example: Conic Optimization Problem

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax = b \\
& \quad x \in K
\end{align*}
\]

\(x \in K\) can equivalently be written as \(z^T x \leq 0, \forall z \in K^\circ\)

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

\[
= (c - A^T \nu + \sum_{z \in K^\circ} \lambda_z \cdot z)^T x + \nu^T b
\]
Another Example: Conic Optimization Problem

minimize \( c^T x \)
subject to \( Ax = b \)
\( x \in K \)

\( x \in K \) can equivalently be written as \( z^T x \leq 0, \forall z \in K^\circ \)

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

\[
= (c - A^T \nu + \sum_{z \in K^\circ} \lambda_z \cdot z)^T x + \nu^T b
\]

Can think of \( \lambda \succeq 0 \) as choosing some \( s \in K^\circ \)

\[
L(x, s, \nu) = (c - A^T \nu + s)^T x + \nu^T b
\]
Another Example: Conic Optimization Problem

minimize \( c^\top x \)
subject to \( Ax = b \)
\( x \in K \)

\( x \in K \) can equivalently be written as \( z^\top x \leq 0, \forall z \in K^\circ \)

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

\[
= (c - A^\top \nu + \sum_{z \in K^\circ} \lambda_z \cdot z)^\top x + \nu^\top b
\]

Can think of \( \lambda \succeq 0 \) as choosing some \( s \in K^\circ \)

\[
L(x, s, \nu) = (c - A^\top \nu + s)^\top x + \nu^\top b
\]

Lagrange dual function \( g(s, \nu) \) is bounded when coefficient of \( x \) is zero, in which case it has value \( \nu^\top b \)
Another Example: Conic Optimization Problem

minimize \( c^T x \)
subject to \( Ax = b \)
\( x \in K \)

maximize \( \nu^T b \)
subject to \( A^T \nu - c \in K^\circ \)

- \( x \in K \) can equivalently be written as \( z^T x \leq 0, \forall z \in K^\circ \)

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

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

\[
L(x, s, \nu) = (c - A^T \nu + s)^T x + \nu^T b
\]

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

1. The Lagrange Dual Problem
2. Duality
Weak Duality

Primal Problem

\[ \min f_0(x) \]
\[ \text{s.t.} \]
\[ f_i(x) \leq 0, \quad \forall i = 1, \ldots, m. \]
\[ h_i(x) = 0, \quad \forall i = 1, \ldots, k. \]

Dual Problem

\[ \max g(\lambda, \nu) \]
\[ \text{s.t.} \]
\[ \lambda \succeq 0 \]

Duality Gap: difference between optimal dual and primal values
Weak Duality

Primal Problem

\[
\begin{align*}
\min \ f_0(x) \\
\text{s.t.} \\
& f_i(x) \leq 0, \quad \forall i = 1, \ldots, m. \\
& h_i(x) = 0, \quad \forall i = 1, \ldots, k.
\end{align*}
\]

Dual Problem

\[
\begin{align*}
\max \ g(\lambda, \nu) \\
\text{s.t.} \\
& \lambda \succeq 0
\end{align*}
\]

We have already argued that it 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

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_1(x) \leq 0
\end{align*}
\]

Let \( G \) be attainable constraint/objective function value tuples

- i.e. \((u, t) \in G\) if there is an \( x \) such that \( f_1(x) = u \) and \( f_0(x) = t \)

- \( p^* = \inf \{ t : (u, t) \in G, u \leq 0 \} \)

- \( g(\lambda) = \inf \{ \lambda u + t : (u, t) \in G \} \)

**Fact**

The equation \( \lambda u + t = g(\lambda) \) defines a supporting hyperplane to \( G \), intersecting \( t \) axis at \( g(\lambda) \leq p^* \).
We say strong duality holds if $\text{OPT}(\text{dual}) = \text{OPT}(\text{primal})$.

- 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

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_1(x) \leq 0
\end{align*}
\]

- Let \( \mathcal{A} \) be everything northeast (i.e. “worse”) than \( G \)
  - 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 \( G \), intersecting \( t \) axis at \( g(\lambda) \leq p^* \).
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

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

**Fact**

When \( f_0 \) and \( f_1 \) are convex, \( \mathcal{A} \) is convex.
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 \( \mathcal{A} \)
Geometric Proof of Strong Duality

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_1(x) \leq 0
\end{align*}
\]

**Fact**

When \( f_0 \) and \( f_1 \) are convex, \( A \) is convex.

**Proof**

- Assume \((u, t)\) and \((u', t')\) are in \( A \)
- \( \exists x, x' \) with \((f_1(x), f_0(x)) \leq (u, t)\) and \((f_1(x'), f_0(x')) \leq (u', t')\).
minimize $f_0(x)$
subject to $f_1(x) \leq 0$

Fact
When $f_0$ and $f_1$ are convex, $A$ is convex.

Proof
- Assume $(u, t)$ and $(u', t')$ are in $A$
- $\exists x, x'$ with $(f_1(x), f_0(x)) \leq (u, t)$ and $(f_1(x'), f_0(x')) \leq (u', t')$.
- By Jensen's inequality
  $$(f_1(\alpha x + (1-\alpha)x'), f_0(\alpha x + (1-\alpha)x')) \leq (\alpha u + (1-\alpha)u', \alpha t + (1-\alpha)t')$$
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, \(A\) is convex.

Proof

- Assume \((u, t)\) and \((u', t')\) are in \(A\)
- \(\exists x, x'\) with \((f_1(x), f_0(x)) \leq (u, t)\) and \((f_1(x'), f_0(x')) \leq (u', t')\).
- By Jensen’s inequality
  
  \[
  (f_1(\alpha x + (1-\alpha)x'), f_0(\alpha x + (1-\alpha)x')) \leq (\alpha u + (1-\alpha)u', \alpha t + (1-\alpha)t')
  \]
- Therefore, midpoint of \((u, t)\) and \((u', t')\) also in \(A\).
Geometric Proof of Strong Duality

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

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

Proof
- Recall \((0, p^*)\) is on the boundary of \( A \)
- By the supporting hyperplane theorem, there is a supporting hyperplane to \( A \) at \((0, p^*)\)
- Direction of the supporting hyperplane gives us an appropriate \( \lambda \)
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.
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 $\lambda$ exists such that $(\lambda, 1)$ is normal to $H$. 

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

minimize \( e^{-x} \)  
subject to \( \frac{x^2}{y} \leq 0 \)

- Let domain of constraint be region \( y \geq 1 \)
- Problem is convex, with feasible region given by \( x = 0 \)
- Optimal value is 1, at \( x = 0 \) and \( y = 1 \)
Violation of Strong Duality

minimize \( e^{-x} \)
subject to \( \frac{x^2}{y} \leq 0 \)

- Let domain of constraint be region \( y \geq 1 \)
- Problem is convex, with feasible region given by \( x = 0 \)
- Optimal value is 1, at \( x = 0 \) and \( y = 1 \)
- \( \mathcal{A} = \mathbb{R}_+^2 \cup (\{0\} \times [1, \infty]) \)
- Therefore, any supporting hyperplane to \( \mathcal{A} \) at \((0, 1)\) must be vertical.
Slater’s Condition

There exists a point \( x \in D \) 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 $x \in D$ 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