Duality

Lecture 5

October 6, 2025

Recap From Last Time

Primal-Dual Pair of Problems

Primal $(\mathcal{P})$

$$
\begin{array}{rll}
\underset{x}{\text{minimize}} & c^\mathsf{T} x & \\
(\lambda_i \rightarrow) & a_i^\mathsf{T} x \geq b_i, & \forall i \in I_{\text{ge}} \\
(\lambda_i \rightarrow) & a_i^\mathsf{T} x \leq b_i, & \forall i \in I_{\text{le}} \\
(\lambda_i \rightarrow) & a_i^\mathsf{T} x = b_i, & \forall i \in I_{\text{eq}} \\
 & x_j \geq 0, & \forall j \in J_p \\
 & x_j \leq 0, & \forall j \in J_n \\
 & x_j \text{ free}, & \forall j \in J_f \\
\text{variables} & x \in \mathbb{R}^n &
\end{array}
$$

We seek **lower bounds** on $\lambda^\star$

Recap From Last Time

Primal-Dual Pair of Problems

Primal $(\mathcal{P})$			**Dual** $(\mathcal{D})$		
$\underset{x}{\text{minimize}}$	$c^{\mathsf{T}} x$		$\underset{\lambda}{\text{maximize}}$	$\lambda^{\mathsf{T}} b$	
$(\lambda_i \to)$	$a_i^{\mathsf{T}} x \geq b_i,$	$\forall i \in I_{\text{ge}}$		$\lambda_i \geq 0,$	$\forall i \in I_{\text{ge}}$
$(\lambda_i \to)$	$a_i^{\mathsf{T}} x \leq b_i,$	$\forall i \in I_{\text{le}}$		$\lambda_i \leq 0,$	$\forall i \in I_{\text{le}}$
$(\lambda_i \to)$	$a_i^{\mathsf{T}} x = b_i,$	$\forall i \in I_{\text{eq}}$		λ_i free,	$\forall i \in I_{\text{eq}}$
	$x_j \geq 0,$	$\forall j \in J_p$	$(x_j \to)$	$\lambda^{\mathsf{T}} A_j \leq c_j,$	$\forall j \in J_p$
	$x_j \leq 0,$	$\forall j \in J_n$	$(x_j \to)$	$\lambda^{\mathsf{T}} A_j \geq c_j,$	$\forall j \in J_n$
	x_j free,	$\forall j \in J_f$	$(x_j \to)$	$\lambda^{\mathsf{T}} A_j = c_j,$	$\forall j \in J_f$
variables	$x \in \mathbb{R}^n$		variables	$\lambda \in \mathbb{R}^m.$	

We seek **lower bounds** on $\lambda^\star$

Recap From Last Time

Primal-Dual Pair of Problems

Primal $(\mathcal{P})$

$$
\begin{array}{rlll}
\underset{x}{\text{minimize}} & c^{\mathsf{T}} x & & \\
(\lambda_i \rightarrow) & a_i^{\mathsf{T}} x \geq b_i, & \forall i \in I_{\text{ge}} \\
(\lambda_i \rightarrow) & a_i^{\mathsf{T}} x \leq b_i, & \forall i \in I_{\text{le}} \\
(\lambda_i \rightarrow) & a_i^{\mathsf{T}} x = b_i, & \forall i \in I_{\text{eq}} \\
 & x_j \geq 0, & \forall j \in J_p \\
 & x_j \leq 0, & \forall j \in J_n \\
 & x_j \text{ free}, & \forall j \in J_f \\
\text{variables} & x \in \mathbb{R}^n &
\end{array}
$$

Dual $(\mathcal{D})$

$$
\begin{array}{rlll}
\underset{\lambda}{\text{maximize}} & \lambda^{\mathsf{T}} b & & \\
 & \lambda_i \geq 0, & \forall i \in I_{\text{ge}} \\
 & \lambda_i \leq 0, & \forall i \in I_{\text{le}} \\
 & \lambda_i \text{ free}, & \forall i \in I_{\text{eq}} \\
(x_j \rightarrow) & \lambda^{\mathsf{T}} A_j \leq c_j, & \forall j \in J_p \\
(x_j \rightarrow) & \lambda^{\mathsf{T}} A_j \geq c_j, & \forall j \in J_n \\
(x_j \rightarrow) & \lambda^{\mathsf{T}} A_j = c_j, & \forall j \in J_f \\
\text{variables} & \lambda \in \mathbb{R}^m. &
\end{array}
$$

We seek **lower bounds** on $\lambda^\star$

Recall the procedure for deriving the dual:

- a dual decision variable λ_i for every primal constraint (except variable signs)
- constrain λ_i to ensure lower bound: λ_i ? 0
- for every primal **decision** x_j, add a dual **constraint** in the form $\lambda^{\mathsf{T}} A_j$? c_j (involving the **column** A_j and the **objective coefficient** c_j corresponding to x_j)

Rules for Constructing the Dual of **Any LP**

Consider **any** linear optimization problem (minimization/maximization):

$$\begin{aligned} \underset{x}{\text{minimize}} \quad / \quad \underset{x}{\text{maximize}} \quad & c^{\mathsf{T}} x \\ (\lambda \rightarrow) \quad & Ax \substack{\leq \\ = \\ \geq} b \\ & x \substack{\leq \\ = \\ \geq} 0 \end{aligned} \tag{1}$$

R1: A dual variable λ_i for every constraint, i.e., every row a_i^{T} of A.
λ_i **free** for **equality** constraints ($a_i^{\mathsf{T}} x = b_i$). Otherwise: λ_i ? 0.

R2: In the dual, add a constraint for *every primal variable* x_j
If x_j is **free**, write this as $\lambda^{\mathsf{T}} A_j = c_j$. Otherwise: $\lambda^{\mathsf{T}} A_j$? c_j.

R3: To determine the signs ?, use this rule of thumb:

the dual variable λ_i is the (sub)gradient of the optimal objective value with respect to the constraint's right-hand-side b_i

- in a minimization, for a "$\leq$" constraint, the dual variable is ≤ 0
- in a minimization, for a "$\geq$" constraint, the dual variable is ≥ 0
- in a maximization, for a "$\leq$" constraint, the dual variable is ≥ 0
- in a maximization, for a "$\geq$" constraint, the dual variable is ≤ 0.

Example

What is the dual of this problem? Are the two problems feasible?

$$
\begin{aligned}
\text{minimize} \quad & x_1 + 2x_2 \\
\text{subject to} \quad & x_1 + x_2 = 1 \\
& 2x_1 + 2x_2 = 3.
\end{aligned}
$$

Example

What is the dual of this problem? Are the two problems feasible?

$$\begin{aligned} \text{minimize} \quad & x_1 + 2x_2 \\ \text{subject to} \quad & x_1 + x_2 = 1 \\ & 2x_1 + 2x_2 = 3. \end{aligned}$$

Weak duality

$$
\begin{array}{llll|llll}
 & \textbf{Primal } (\mathcal{P}) & & & & \textbf{Dual } (\mathcal{D}) & & \\
\text{minimize}_{x} & c^{\mathsf{T}} x & & & \text{maximize}_{\lambda} & \lambda^{\mathsf{T}} b & & \\
(\lambda_i \rightarrow) & a_i^{\mathsf{T}} x \geq b_i, & \forall i \in I_{\text{ge}}, & & & \lambda_i \geq 0, & \forall i \in I_{\text{ge}}, & \\
(\lambda_i \rightarrow) & a_i^{\mathsf{T}} x \leq b_i, & \forall i \in I_{\text{le}}, & & & \lambda_i \leq 0, & \forall i \in I_{\text{le}}, & \\
(\lambda_i \rightarrow) & a_i^{\mathsf{T}} x = b_i, & \forall i \in I_{\text{eq}}, & & & \lambda_i \text{ free}, & \forall i \in I_{\text{eq}}, & \\
 & x_j \geq 0, & \forall j \in J_p, & & (x_j \rightarrow) & \lambda^{\mathsf{T}} A_j \leq c_j, & \forall j \in J_p, & \\
 & x_j \leq 0, & \forall j \in J_n, & & (x_j \rightarrow) & \lambda^{\mathsf{T}} A_j \geq c_j, & \forall j \in J_n, & \\
 & x_j \text{ free}, & \forall j \in J_f. & & (x_j \rightarrow) & \lambda^{\mathsf{T}} A_j = c_j, & \forall j \in J_f. &
\end{array}
$$

Weak duality

$$
\begin{array}{lll|lll}
 & \textbf{Primal } (\mathcal{P}) & & & \textbf{Dual } (\mathcal{D}) & \\
\text{minimize}_x & c^\mathsf{T} x & & \text{maximize}_\lambda & \lambda^\mathsf{T} b & \\
(\lambda_i \to) & a_i^\mathsf{T} x \geq b_i, & \forall i \in I_{\text{ge}}, & & \lambda_i \geq 0, & \forall i \in I_{\text{ge}}, \\
(\lambda_i \to) & a_i^\mathsf{T} x \leq b_i, & \forall i \in I_{\text{le}}, & & \lambda_i \leq 0, & \forall i \in I_{\text{le}}, \\
(\lambda_i \to) & a_i^\mathsf{T} x = b_i, & \forall i \in I_{\text{eq}}, & & \lambda_i \text{ free}, & \forall i \in I_{\text{eq}}, \\
 & x_j \geq 0, & \forall j \in J_p, & (x_j \to) & \lambda^\mathsf{T} A_j \leq c_j, & \forall j \in J_p, \\
 & x_j \leq 0, & \forall j \in J_n, & (x_j \to) & \lambda^\mathsf{T} A_j \geq c_j, & \forall j \in J_n, \\
 & x_j \text{ free}, & \forall j \in J_f. & (x_j \to) & \lambda^\mathsf{T} A_j = c_j, & \forall j \in J_f.
\end{array}
$$

Theorem (Weak duality)

If x is feasible for $(\mathcal{P})$ and λ is feasible for $(\mathcal{D})$, then $\lambda^\mathsf{T} b \leq c^\mathsf{T} x$.

Proof. Trivially true from our construction – omitted.

Implications of Weak Duality

Corollary

The following results hold:

(a) If the optimal objective in $(\mathcal{P})$ *is* $-\infty$*, then* $(\mathcal{D})$ *...*

(b) If the optimal objective in $(\mathcal{D})$ *is* $+\infty$*, then* $(\mathcal{P})$ *...*

Implications of Weak Duality

Corollary

The following results hold:

(a) If the optimal objective in $(\mathcal{P})$ *is* $-\infty$*, then* $(\mathcal{D})$ *...*

(b) If the optimal objective in $(\mathcal{D})$ *is* $+\infty$*, then* $(\mathcal{P})$ *...*

Implications of Weak Duality

Corollary

The following results hold:

(a) If the optimal objective in $(\mathcal{P})$ *is* $-\infty$*, then* $(\mathcal{D})$ *...*

(b) If the optimal objective in $(\mathcal{D})$ *is* $+\infty$*, then* $(\mathcal{P})$ *...*

(c) If $x \in P$ *and* $\lambda \in D$*, then:*

$$c^\mathsf{T} x - \lambda^\star \le c^\mathsf{T} x - \lambda^\mathsf{T} b \quad \textbf{and} \quad d^\star - \lambda^\mathsf{T} b \le c^\mathsf{T} x - \lambda^\mathsf{T} b.$$

Implications of Weak Duality

Corollary

The following results hold:

(a) If the optimal objective in $(\mathcal{P})$ *is* $-\infty$*, then* $(\mathcal{D})$ *...*

(b) If the optimal objective in $(\mathcal{D})$ *is* $+\infty$*, then* $(\mathcal{P})$ *...*

(c) If $x \in P$ *and* $\lambda \in D$*, then:*

$$c^\mathsf{T}x - \lambda^\star \le c^\mathsf{T}x - \lambda^\mathsf{T}b \;\; \textbf{and} \;\; d^\star - \lambda^\mathsf{T}b \le c^\mathsf{T}x - \lambda^\mathsf{T}b.$$

(d) If $x \in P$*,* $\lambda \in D$*, and* $\lambda^\mathsf{T}b = c^\mathsf{T}x$*, then* x **optimal** *for* $(\mathcal{P})$ *and* λ **optimal** *for* $(\mathcal{D})$*.*

Implications of Weak Duality

Corollary

The following results hold:

(a) If the optimal objective in $(\mathcal{P})$ *is* $-\infty$*, then* $(\mathcal{D})$ *...*

(b) If the optimal objective in $(\mathcal{D})$ *is* $+\infty$*, then* $(\mathcal{P})$ *...*

(c) If $x \in P$ *and* $\lambda \in D$*, then:*

$$c^{\mathsf{T}}x - \lambda^{\star} \leq c^{\mathsf{T}}x - \lambda^{\mathsf{T}}b \;\; \textbf{and} \;\; d^{\star} - \lambda^{\mathsf{T}}b \leq c^{\mathsf{T}}x - \lambda^{\mathsf{T}}b.$$

(d) If $x \in P$*,* $\lambda \in D$*, and* $\lambda^{\mathsf{T}}b = c^{\mathsf{T}}x$*, then* x **optimal** *for* $(\mathcal{P})$ *and* λ **optimal** *for* $(\mathcal{D})$*.*

(c) and (d) provide (sub)optimality certificates, but...

How do we know that the gaps in (c) are not very large?

How do we know that x and λ satisfying (d) even exist?

Strong duality

Theorem (Strong duality)

If $(\mathcal{P})$ has an optimal solution, so does $(\mathcal{D})$, and the optimal values are equal, $\lambda^\star = d^\star$.

Strong duality

Theorem (Strong duality)

If $(\mathcal{P})$ has an optimal solution, so does $(\mathcal{D})$, and the optimal values are equal, $\lambda^\star = d^\star$.

Proof. Many proofs possible...

- See Bertsimas & Tsitsiklis for a proof involving the simplex algorithm
- We provide a more general proof, in three steps:
  1. The **separating hyperplane theorem** (for convex sets)
  2. The Farkas Lemma
  3. Strong duality

Need a tiny bit of **real analysis** background...

A Few Real Analysis Results

Definition (Closed Set)

A set $S \subseteq \mathbb{R}^n$ is called **closed** if it contains the limit of any sequence of elements of S. That is, if $x_n \in S$, $\forall n \geq 1$ and $x_n \to x^\star$, then $x^\star \in S$.

A Few Real Analysis Results

Definition (Closed Set)

A set $S \subseteq \mathbb{R}^n$ is called **closed** if it contains the limit of any sequence of elements of S. That is, if $x_n \in S$, $\forall n \geq 1$ and $x_n \to x^\star$, then $x^\star \in S$.

Theorem

Every polyhedron is closed.

A Few Real Analysis Results

Definition (Closed Set)

A set $S \subseteq \mathbb{R}^n$ is called **closed** if it contains the limit of any sequence of elements of S. That is, if $x_n \in S$, $\forall n \geq 1$ and $x_n \to x^\star$, then $x^\star \in S$.

Theorem

Every polyhedron is closed.

Proof.

- Consider $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ (representation is w.l.o.g.)
- Suppose that $\{x_n\}_{n \geq 1}$ is a sequence with $x_n \in S$ for every n, and $x_n \to x^\star$.
- For each k, we have $x_k \in P$, and therefore, $Ax_k \geq b$.
- Then, $Ax^\star = A\left(\lim_{k\to\infty} x_k\right) = \lim_{k\to\infty} Ax_k \geq b$, so $x^\star$ belongs to P. $\square$

A Few Real Analysis Results

Definition (Closed Set)

A set $S \subseteq \mathbb{R}^n$ is called **closed** if it contains the limit of any sequence of elements of S. That is, if $x_n \in S$, $\forall n \geq 1$ and $x_n \to x^\star$, then $x^\star \in S$.

Theorem

Every polyhedron is closed.

Is every **convex set** *closed?*

A Few Real Analysis Results

Definition (Closed Set)

A set $S \subseteq \mathbb{R}^n$ is called **closed** if it contains the limit of any sequence of elements of S. That is, if $x_n \in S$, $\forall n \geq 1$ and $x_n \to x^\star$, then $x^\star \in S$.

Theorem

Every polyhedron is closed.

Theorem (Weierstrass' Theorem)

If $f : \mathbb{R}^n \to \mathbb{R}$ is a continuous function, and if S is a nonempty, closed, and bounded subset of $\mathbb{R}^n$, then there exist $\underline{x}, \bar{x} \in S$ such that $f(\underline{x}) \leq f(x) \leq f(\bar{x})$ for all $x \in S$.

i.e., a continuous function achieves its minimum and maximum

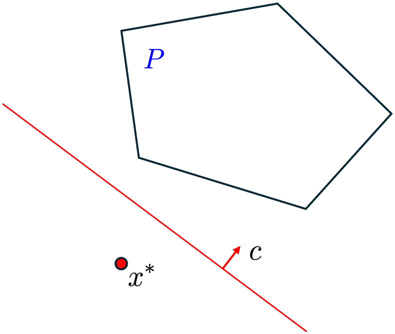
Separating Hyperplane Theorem

The first **fundamental result in optimization**

Separating Hyperplane Theorem

Theorem (**Simple** Separating Hyperplane Theorem)

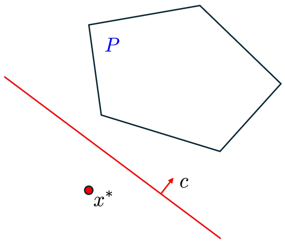
Consider a point $x^\star$ and a polyhedron P. If $x^\star \notin P$, then there exists a vector $c \in \mathbb{R}^n$ such that $c \neq 0$ and $c^\mathsf{T} x^\star < c^\mathsf{T} y$ holds for all $y \in P$.

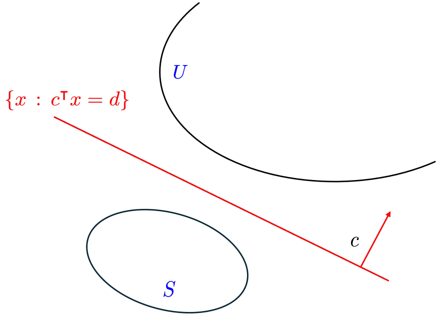


Separating Hyperplane Theorem

Theorem (Separating Hyperplane Theorem for Convex Sets)

Let S and U be two nonempty, closed, convex subsets of $\mathbb{R}^n$ such that $S \cap U = \emptyset$ and S is bounded. Then, there exists $c \in \mathbb{R}^n$ and $d \in \mathbb{R}$ such that $S \subset \{x \in \mathbb{R}^n : c^\mathsf{T} x < d\}$ and $U \subset \{x \in \mathbb{R}^n : c^\mathsf{T} x > d\}$.





Separating Hyperplane Theorem

Theorem (Separating Hyperplane Theorem **for Convex Sets**)

Let S and U be two nonempty, closed, convex subsets of $\mathbb{R}^n$ such that $S \cap U = \emptyset$ and S is bounded. Then, there exists $c \in \mathbb{R}^n$ and $d \in \mathbb{R}$ such that $S \subset \{x \in \mathbb{R}^n : c^\mathsf{T} x < d\}$ and $U \subset \{x \in \mathbb{R}^n : c^\mathsf{T} x > d\}$.

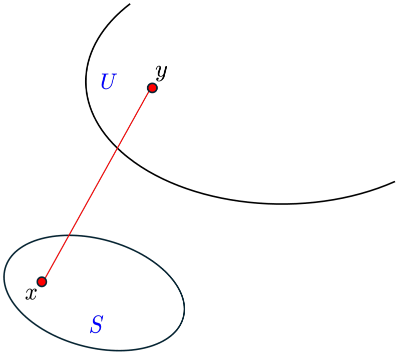
Proof.

Separating Hyperplane Theorem

Theorem (Separating Hyperplane Theorem **for Convex Sets**)

Let S and U be two nonempty, closed, convex subsets of $\mathbb{R}^n$ such that $S \cap U = \emptyset$ and S is bounded. Then, there exists $c \in \mathbb{R}^n$ and $d \in \mathbb{R}$ such that $S \subset \{x \in \mathbb{R}^n : c^\mathsf{T} x < d\}$ and $U \subset \{x \in \mathbb{R}^n : c^\mathsf{T} x > d\}$.

Proof. Consider $\|x - y\|$ with $x \in S, y \in U$

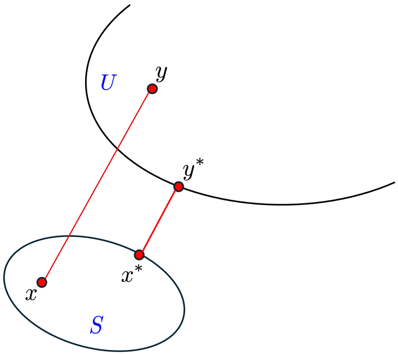


Separating Hyperplane Theorem

Theorem (Separating Hyperplane Theorem **for Convex Sets**)

Let S and U be two nonempty, closed, convex subsets of $\mathbb{R}^n$ such that $S \cap U = \emptyset$ and S is bounded. Then, there exists $c \in \mathbb{R}^n$ and $d \in \mathbb{R}$ such that $S \subset \{x \in \mathbb{R}^n : c^\mathsf{T} x < d\}$ and $U \subset \{x \in \mathbb{R}^n : c^\mathsf{T} x > d\}$.

Proof. Argue that the minimum is achieved, at $x^\star, y^\star$

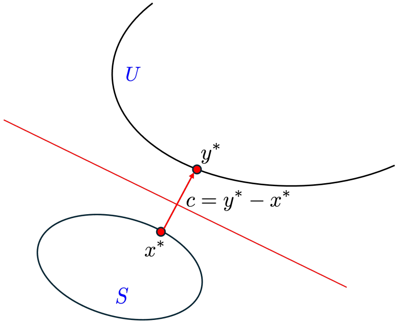


Separating Hyperplane Theorem

Theorem (Separating Hyperplane Theorem **for Convex Sets**)

Let S and U be two nonempty, closed, convex subsets of $\mathbb{R}^n$ such that $S \cap U = \emptyset$ and S is bounded. Then, there exists $c \in \mathbb{R}^n$ and $d \in \mathbb{R}$ such that $S \subset \{x \in \mathbb{R}^n : c^\mathsf{T} x < d\}$ and $U \subset \{x \in \mathbb{R}^n : c^\mathsf{T} x > d\}$.

Proof. Argue that $c = y^\star - x^\star$ and $d = \frac{c^\mathsf{T}(x^\star + y^\star)}{2}$ give strict separating hyperplane



Separating Hyperplane Theorem - Caveats!

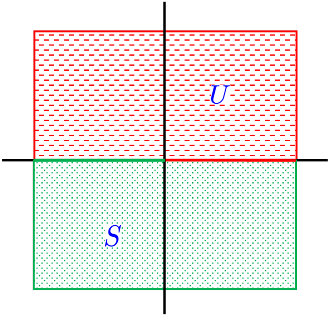
Both conditions in the theorem needed: closed and at least one set bounded

Separating Hyperplane Theorem - Caveats!

Both conditions in the theorem needed: closed and at least one set bounded

- **Left:** two convex sets that are **not closed** but are both bounded:

$$S = [-1,1] \times [-1,0) \cup \{(x,y) : x \in [-1,0], y = 0\}, \quad U = [-1,1]^2 \setminus S$$



Separating Hyperplane Theorem - Caveats!

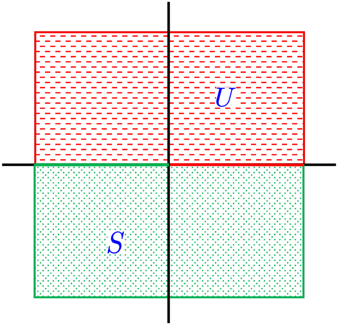
Both conditions in the theorem needed: closed and at least one set bounded

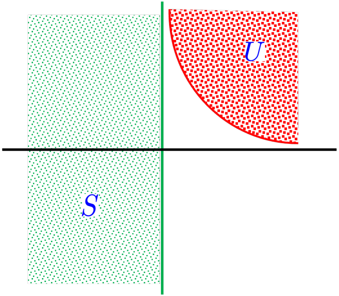
- **Left:** two convex sets that are **not closed** but are both bounded:

$$S = [-1,1] \times [-1,0) \cup \{(x,y) : x \in [-1,0], y = 0\}, \quad U = [-1,1]^2 \setminus S$$

- **Right:** two convex sets that are both closed but are **unbounded**

$$S = \{(x,y) : x \leq 0\}, \quad U = \{(x,y) : x \geq 0, y \geq 1/x\}$$



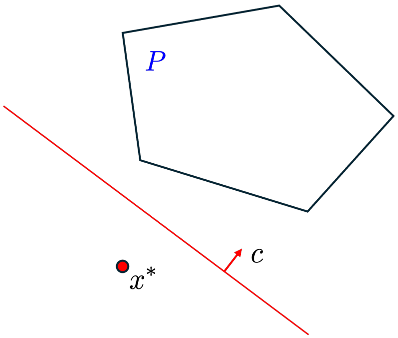


Needed For Our Purposes

We proved the first **fundamental result in optimization**!

Corollary (Needed for our purposes...)

If P is a polyhedron and $x^\star \notin P$, there exists a hyperplane that strictly separates $x^\star$ from P, i.e., $\exists c \neq 0$ such that $c^\top x^\star < c^\top x$ for any $x \in P$.



Farkas Lemma

Time for the **second fundamental result in optimization**!

Farkas Lemma

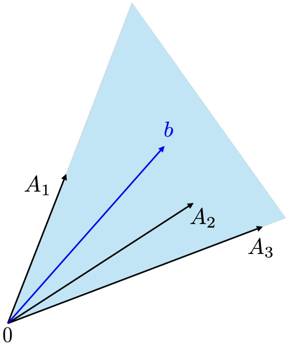
Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$*, exactly one of the following two alternatives holds:*

Farkas Lemma

Theorem (Farkas' Lemma)

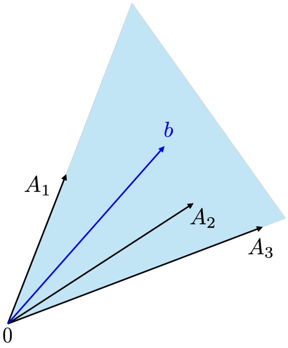
For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.

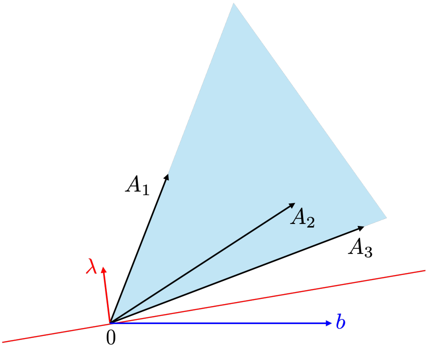


Farkas Lemma

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^\mathsf{T} A \geq 0$ and $\lambda^\mathsf{T} b < 0$.





Farkas Lemma

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^{\mathsf{T}} A \geq 0$ and $\lambda^{\mathsf{T}} b < 0$.

Proof. "(a) true implies (b) false."

Farkas Lemma

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^\mathsf{T} A \geq 0$ and $\lambda^\mathsf{T} b < 0$.

Proof. "(a) true implies (b) false."

(a) true means $\exists x \geq 0 : Ax = b$.

(b) true means $\exists \lambda : \lambda^\mathsf{T} A \geq 0$ and $\lambda^\mathsf{T} b < 0$.

If (a) and (b) both true, then $\lambda^\mathsf{T} b = \lambda^\mathsf{T} A x \geq 0$, which is a contradiction.

Farkas Lemma

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^\mathsf{T} A \geq 0$ and $\lambda^\mathsf{T} b < 0$.

"(a) false implies (b) true." Want to use the separating hyperplane theorem.

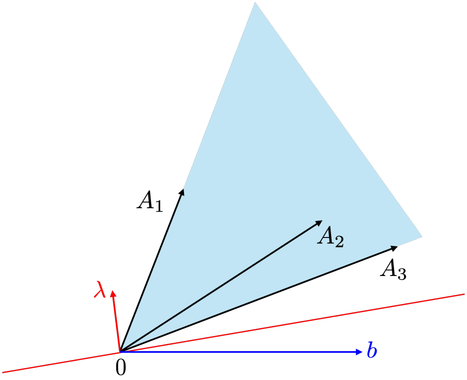
Farkas Lemma

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^{\mathsf{T}}A \geq 0$ and $\lambda^{\mathsf{T}}b < 0$.

"(a) false implies (b) true." Want to use the separating hyperplane theorem.

- (a) false implies that $b \notin \{y : \exists x \geq 0 \text{ such that } y = Ax\} := S$.



Farkas Lemma

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^\mathsf{T} A \geq 0$ and $\lambda^\mathsf{T} b < 0$.

"(a) false implies (b) true." Want to use the separating hyperplane theorem.

- (a) false implies that $b \notin \{y : \exists x \geq 0 \text{ such that } y = Ax\} := S$.
- S is a convex and **closed** set

Farkas Lemma

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^\mathsf{T} A \geq 0$ and $\lambda^\mathsf{T} b < 0$.

"(a) false implies (b) true." Want to use the separating hyperplane theorem.

- (a) false implies that $b \notin \{y : \exists x \geq 0 \text{ such that } y = Ax\} := S$.
- S is a convex and **closed** set (S is polyhedral)

Farkas Lemma

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^\mathsf{T} A \geq 0$ and $\lambda^\mathsf{T} b < 0$.

"(a) false implies (b) true." Want to use the separating hyperplane theorem.

- (a) false implies that $b \notin \{y : \exists x \geq 0 \text{ such that } y = Ax\} := S$.
- S is a convex and **closed** set (S is polyhedral)
- Separating Hyperplane Theorem implies $\exists \lambda : \lambda^\mathsf{T} b < \lambda^\mathsf{T} y, \forall y \in S$

Farkas Lemma

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^\mathsf{T} A \geq 0$ and $\lambda^\mathsf{T} b < 0$.

"(a) false implies (b) true." Want to use the separating hyperplane theorem.

- (a) false implies that $b \notin \{y : \exists\, x \geq 0 \text{ such that } y = Ax\} := S$.
- S is a convex and **closed** set (S is polyhedral)
- Separating Hyperplane Theorem implies $\exists \lambda : \lambda^\mathsf{T} b < \lambda^\mathsf{T} y, \forall y \in S$
- $0 \in S \Rightarrow \lambda^\mathsf{T} b < 0$

Farkas Lemma

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^\mathsf{T} A \geq 0$ and $\lambda^\mathsf{T} b < 0$.

"(a) false implies (b) true." Want to use the separating hyperplane theorem.

- (a) false implies that $b \notin \{y : \exists x \geq 0 \text{ such that } y = Ax\} := S$.
- S is a convex and **closed** set (S is polyhedral)
- Separating Hyperplane Theorem implies $\exists \lambda : \lambda^\mathsf{T} b < \lambda^\mathsf{T} y, \forall y \in S$
- $0 \in S \Rightarrow \lambda^\mathsf{T} b < 0$
- Every column A_i of A satisfies $\theta A_i \in S$ for every $\theta > 0$, so

$$\frac{\lambda^\mathsf{T} b}{\theta} < \lambda^\mathsf{T} A_i, \ \forall \theta > 0$$

Farkas Lemma

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^\mathsf{T} A \geq 0$ and $\lambda^\mathsf{T} b < 0$.

"(a) false implies (b) true." Want to use the separating hyperplane theorem.

- (a) false implies that $b \notin \{y : \exists x \geq 0 \text{ such that } y = Ax\} := S$.
- S is a convex and **closed** set (S is polyhedral)
- Separating Hyperplane Theorem implies $\exists \lambda : \lambda^\mathsf{T} b < \lambda^\mathsf{T} y, \forall y \in S$
- $0 \in S \Rightarrow \lambda^\mathsf{T} b < 0$
- Every column A_i of A satisfies $\theta A_i \in S$ for every $\theta > 0$, so

$$\frac{\lambda^\mathsf{T} b}{\theta} < \lambda^\mathsf{T} A_i,\ \forall \theta > 0$$

- Limit $\theta \to \infty$ implies $\lambda^\mathsf{T} A_i \geq 0$. ■

Farkas Lemma Implications

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^\mathsf{T} A \geq 0$ and $\lambda^\mathsf{T} b < 0$.

We proved the **second fundamental result in optimization**!

Farkas Lemma Implications

Theorem (Farkas' Lemma)

For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^\mathsf{T} A \geq 0$ and $\lambda^\mathsf{T} b < 0$.

We proved the **second fundamental result in optimization**!

- Suppose your primal problem $(\mathcal{P})$ was the standard-form LP:

$$
\begin{array}{lll}
(\mathcal{P}) & \text{minimize} & c^\mathsf{T} x \\
 & \text{subject to} & Ax = b \\
 & & x \geq 0
\end{array}
$$

- What does the Farkas Lemma state about this?

Strong Duality

Consider the following primal-dual pair:

$$
\begin{aligned}
(\mathcal{P}) \quad & \text{minimize } c^{\mathsf{T}} x \\
& \text{subject to } Ax \geq b
\end{aligned}
$$

Strong Duality

Consider the following primal-dual pair:

$$(\mathcal{P})\ \ \text{minimize}\ c^{\mathsf{T}}x \quad \text{subject to}\ Ax \geq b$$

Theorem (Strong Duality)

If $(\mathcal{P})$ has an optimal solution, so does $(\mathcal{D})$, and their optimal values are equal.

Strong Duality

$$(\mathcal{P}) \quad \text{minimize } c^{\mathsf{T}}x \qquad\qquad (\mathcal{D}) \quad \text{maximize } \lambda^{\mathsf{T}}b$$
$$\text{subject to } Ax \geq b \qquad\qquad \text{subject to } \lambda^{\mathsf{T}}A = c^{\mathsf{T}}, \quad \lambda \geq 0.$$

Proof.

- Assume $(\mathcal{P})$ has optimal solution $x^\star$
- Will prove that $(\mathcal{D})$ admits feasible solution λ such that $\lambda^{\mathsf{T}}b = c^{\mathsf{T}}x^\star$

Strong Duality

$$(\mathcal{P}) \quad \text{minimize } c^{\mathsf{T}}x \qquad\qquad (\mathcal{D}) \quad \text{maximize } \lambda^{\mathsf{T}}b$$
$$\text{subject to } Ax \geq b \qquad\qquad \text{subject to } \lambda^{\mathsf{T}}A = c^{\mathsf{T}}, \quad \lambda \geq 0.$$

Proof.

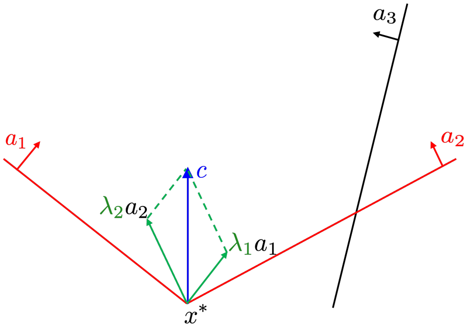
- Assume $(\mathcal{P})$ has optimal solution $x^\star$
- Will prove that $(\mathcal{D})$ admits feasible solution λ such that $\lambda^{\mathsf{T}}b = c^{\mathsf{T}}x^\star$
- Let $\mathcal{F} = \{i \mid a_i^{\mathsf{T}}x^\star = b_i\}$ denote the indices of active constraints at $x^\star$
- Show that c can be written as conic combination of constraints $\{a_i : i \in \mathcal{F}\}$

Strong Duality

$$(\mathcal{P}) \quad \text{minimize } c^\mathsf{T} x \qquad\qquad (\mathcal{D}) \quad \text{maximize } \lambda^\mathsf{T} b$$
$$\text{subject to } Ax \geq b \qquad\qquad \text{subject to } \lambda^\mathsf{T} A = c^\mathsf{T}, \quad \lambda \geq 0.$$

Proof.

- Assume $(\mathcal{P})$ has optimal solution $x^\star$
- Will prove that $(\mathcal{D})$ admits feasible solution λ such that $\lambda^\mathsf{T} b = c^\mathsf{T} x^\star$
- Let $\mathcal{F} = \{i \mid a_i^\mathsf{T} x^\star = b_i\}$ denote the indices of active constraints at $x^\star$
- Show that c can be written as conic combination of constraints $\{a_i : i \in \mathcal{F}\}$



Strong Duality

$$(\mathcal{P}) \quad \begin{array}{ll} \text{minimize} & c^\mathsf{T} x \\ \text{subject to} & Ax \geq b \end{array} \qquad (\mathcal{D}) \quad \begin{array}{ll} \text{maximize} & \lambda^\mathsf{T} b \\ \text{subject to} & \lambda^\mathsf{T} A = c^\mathsf{T}, \quad \lambda \geq 0. \end{array}$$

Proof.

- First, we show that for any vector d, the following implication holds:

$$a_i^\mathsf{T} d \geq 0, \forall\, i \in \mathcal{F} \quad \Rightarrow \quad c^\mathsf{T} d \geq 0.$$

Strong Duality

$$(\mathcal{P}) \quad \text{minimize } c^{\mathsf{T}}x \qquad\qquad (\mathcal{D}) \quad \text{maximize } \lambda^{\mathsf{T}}b$$
$$\text{subject to } Ax \geq b \qquad\qquad \text{subject to } \lambda^{\mathsf{T}}A = c^{\mathsf{T}}, \quad \lambda \geq 0.$$

Proof.

- First, we show that for any vector d, the following implication holds:

$$a_i^{\mathsf{T}}d \geq 0, \forall\, i \in \mathcal{F} \quad \Rightarrow \quad c^{\mathsf{T}}d \geq 0.$$

- For any such d, we claim that $x^\star + \epsilon d \in P$ for small ϵ

Strong Duality

$$(\mathcal{P}) \quad \text{minimize } c^{\mathsf{T}}x \qquad\qquad (\mathcal{D}) \quad \text{maximize } \lambda^{\mathsf{T}}b$$
$$\text{subject to } Ax \geq b \qquad\qquad \text{subject to } \lambda^{\mathsf{T}}A = c^{\mathsf{T}}, \quad \lambda \geq 0.$$

Proof.

- First, we show that for any vector d, the following implication holds:

$$a_i^{\mathsf{T}}d \geq 0, \forall\, i \in \mathcal{F} \quad \Rightarrow \quad c^{\mathsf{T}}d \geq 0.$$

- For any such d, we claim that $x^\star + \epsilon d \in P$ for small ϵ
  - $a_i^{\mathsf{T}}(x^\star + \epsilon d) \geq b_i, \forall i \in \mathcal{F}$ **for any** $\epsilon > 0$
  - $a_i^{\mathsf{T}}x^\star > b_i\ \forall i \notin \mathcal{F}$ implies that $\exists \epsilon > 0$ such that $a_i^{\mathsf{T}}(x^\star + \epsilon d) \geq b_i, \forall i \notin \mathcal{F}$

Strong Duality

$$(\mathcal{P}) \ \ \text{minimize } c^{\mathsf{T}}x \qquad\qquad (\mathcal{D}) \ \ \text{maximize } \lambda^{\mathsf{T}}b$$
$$\text{subject to } Ax \geq b \qquad\qquad \text{subject to } \lambda^{\mathsf{T}}A = c^{\mathsf{T}}, \quad \lambda \geq 0.$$

Proof.

- First, we show that for any vector d, the following implication holds:

$$a_i^{\mathsf{T}}d \geq 0, \forall\, i \in \mathcal{F} \quad \Rightarrow \quad c^{\mathsf{T}}d \geq 0.$$

- For any such d, we claim that $x^\star + \epsilon d \in P$ for small ϵ
  - $a_i^{\mathsf{T}}(x^\star + \epsilon d) \geq b_i, \forall i \in \mathcal{F}$ **for any** $\epsilon > 0$
  - $a_i^{\mathsf{T}}x^\star > b_i \ \forall i \notin \mathcal{F}$ implies that $\exists \epsilon > 0$ such that $a_i^{\mathsf{T}}(x^\star + \epsilon d) \geq b_i, \forall i \notin \mathcal{F}$

- $c^{\mathsf{T}}d \geq 0$ because otherwise $c^{\mathsf{T}}(x^\star + \epsilon d) < c^{\mathsf{T}}x^\star$ would contradict $x^\star$ optimal

Strong Duality

$$(\mathcal{P}) \quad \text{minimize } c^{\mathsf{T}}x \qquad\qquad (\mathcal{D}) \quad \text{maximize } \lambda^{\mathsf{T}}b$$
$$\text{subject to } Ax \geq b \qquad\qquad \text{subject to } \lambda^{\mathsf{T}}A = c^{\mathsf{T}}, \quad \lambda \geq 0.$$

Proof.

- First, we show that for any vector d, the following implication holds:

$$a_i^{\mathsf{T}}d \geq 0, \forall\, i \in \mathcal{F} \quad \Rightarrow \quad c^{\mathsf{T}}d \geq 0.$$

- For any such d, we claim that $x^\star + \epsilon d \in P$ for small ϵ
  - $a_i^{\mathsf{T}}(x^\star + \epsilon d) \geq b_i, \forall i \in \mathcal{F}$ **for any** $\epsilon > 0$
  - $a_i^{\mathsf{T}}x^\star > b_i\ \forall i \notin \mathcal{F}$ implies that $\exists \epsilon > 0$ such that $a_i^{\mathsf{T}}(x^\star + \epsilon d) \geq b_i, \forall i \notin \mathcal{F}$
- $c^{\mathsf{T}}d \geq 0$ because otherwise $c^{\mathsf{T}}(x^\star + \epsilon d) < c^{\mathsf{T}}x^\star$ would contradict $x^\star$ optimal
- So $\nexists d : a_i^{\mathsf{T}}d \geq 0, \forall\, i \in \mathcal{F},\ c^{\mathsf{T}}d < 0$

Strong Duality

$$(\mathcal{P}) \ \text{minimize } c^{\mathsf{T}}x \qquad\qquad (\mathcal{D}) \ \text{maximize } \lambda^{\mathsf{T}}b$$
$$\text{subject to } Ax \geq b \qquad\qquad \text{subject to } \lambda^{\mathsf{T}}A = c^{\mathsf{T}}, \quad \lambda \geq 0.$$

Proof.

- First, we show that for any vector d, the following implication holds:

$$a_i^{\mathsf{T}}d \geq 0, \forall\, i \in \mathcal{F} \quad \Rightarrow \quad c^{\mathsf{T}}d \geq 0.$$

- For any such d, we claim that $x^\star + \epsilon d \in P$ for small ϵ
  - $a_i^{\mathsf{T}}(x^\star + \epsilon d) \geq b_i, \forall i \in \mathcal{F}$ **for any** $\epsilon > 0$
  - $a_i^{\mathsf{T}}x^\star > b_i \ \forall i \notin \mathcal{F}$ implies that $\exists \epsilon > 0$ such that $a_i^{\mathsf{T}}(x^\star + \epsilon d) \geq b_i, \forall i \notin \mathcal{F}$
- $c^{\mathsf{T}}d \geq 0$ because otherwise $c^{\mathsf{T}}(x^\star + \epsilon d) < c^{\mathsf{T}}x^\star$ would contradict $x^\star$ optimal
- So $\nexists d : a_i^{\mathsf{T}}d \geq 0, \forall\, i \in \mathcal{F}, \ c^{\mathsf{T}}d < 0$
- Farkas Lemma : alternative (b) is not true, so alternative (a) must be true:

$$\exists \{\lambda_i\}_{i \in \mathcal{F}} : \lambda_i \geq 0, \ c = \sum_{i \in \mathcal{F}} \lambda_i a_i$$

Strong Duality

$$(\mathcal{P}) \ \text{minimize} \ c^{\mathsf{T}}x \qquad\qquad (\mathcal{D}) \ \text{maximize} \ \lambda^{\mathsf{T}}b$$

$$\text{subject to } Ax \geq b \qquad\qquad \text{subject to } \lambda^{\mathsf{T}}A = c^{\mathsf{T}}, \quad \lambda \geq 0.$$

Proof.

- First, we show that for any vector d, the following implication holds:

$$a_i^{\mathsf{T}}d \geq 0, \, \forall \, i \in \mathcal{F} \quad \Rightarrow \quad c^{\mathsf{T}}d \geq 0.$$

- For any such d, we claim that $x^\star + \epsilon d \in P$ for small ϵ
  - $a_i^{\mathsf{T}}(x^\star + \epsilon d) \geq b_i, \forall i \in \mathcal{F}$ **for any** $\epsilon > 0$
  - $a_i^{\mathsf{T}}x^\star > b_i \ \forall i \notin \mathcal{F}$ implies that $\exists \epsilon > 0$ such that $a_i^{\mathsf{T}}(x^\star + \epsilon d) \geq b_i, \forall i \notin \mathcal{F}$
- $c^{\mathsf{T}}d \geq 0$ because otherwise $c^{\mathsf{T}}(x^\star + \epsilon d) < c^{\mathsf{T}}x^\star$ would contradict $x^\star$ optimal
- So $\nexists d : a_i^{\mathsf{T}}d \geq 0, \forall \, i \in \mathcal{F}, \, c^{\mathsf{T}}d < 0$
- Farkas Lemma : alternative (b) is not true, so alternative (a) must be true:

$$\exists \{\lambda_i\}_{i \in \mathcal{F}} : \lambda_i \geq 0, \; c = \sum_{i \in \mathcal{F}} \lambda_i a_i$$

- Let $\lambda_i = 0$ for $i \notin \mathcal{F} \quad \Rightarrow \quad \exists \lambda$ feasible for $(\mathcal{D})$
- $\lambda^{\mathsf{T}}b = \sum_{i \in \mathcal{F}} \lambda_i b_i = \sum_{i \in \mathcal{F}} \lambda_i a_i^{\mathsf{T}} x^\star = c^{\mathsf{T}}x^\star \qquad \blacksquare$

Implications

Strong duality leaves only a few possibilities for a primal-dual pair:

		Dual		
		Finite Optimum	Unbounded	Infeasible
Primal	Finite Optimum	?	?	?
	Unbounded	?	?	?
	Infeasible	?	?	?

Implications

Strong duality leaves only a few possibilities for a primal-dual pair:

		Dual		
		Finite Optimum	Unbounded	Infeasible
Primal	Finite Optimum	?	?	?
	Unbounded	?	?	?
	Infeasible	?	?	?

		Dual		
		Finite Optimum	Unbounded	Infeasible
Primal	Finite Optimum	Possible	Impossible	Impossible
	Unbounded	Impossible	Impossible	Possible
	Infeasible	Impossible	Possible	?

Polynomially-Sized CVaR Representation

- Recall homework: ensure CVaR of portfolio payoff exceeds a lower limit
- CVaR was defined as the average over the k-smallest values (for suitable integer k)
- If payoffs in the scenarios are $v_1, v_2, \ldots, v_n$, the key constraint is:

$$\sum_{i=1}^{k} v_{[i]} \geq b, \tag{2}$$

where $v_{[1]} \leq v_{[2]} \leq \cdots \leq v_{[n]}$ is the sorted vector of payoffs.

Polynomially-Sized CVaR Representation

- Recall homework: ensure CVaR of portfolio payoff exceeds a lower limit
- CVaR was defined as the average over the k-smallest values (for suitable integer k)
- If payoffs in the scenarios are $v_1, v_2, \ldots, v_n$, the key constraint is:

$$\sum_{i=1}^{k} v_{[i]} \geq b, \tag{2}$$

  where $v_{[1]} \leq v_{[2]} \leq \cdots \leq v_{[n]}$ is the sorted vector of payoffs.
- Can write one constraint for each vector in $\{0, 1\}^n$ with exactly k values of 1.
- *How to formulate with a polynomial number of variables and constraints?*

Application in Robust Optimization

- We have LP with constraints $Ax \leq b$. One of the constraints is:

$$a^{\top}x \leq b, \tag{3}$$

where a satisfies $a \in \mathcal{A}$ and $\mathcal{A}$ is polyhedral

- We seek decisions x that are **robustly feasible**, i.e.,

$$a^{\top}x \leq b, \; \forall\, a \in \mathcal{A} := \{a \in \mathbb{R}^n \;:\; Ca \leq d\} \tag{4}$$

Infinitely many constraints : "semi-infinite" LP. *Any ideas?*

Application in Robust Optimization

- We have LP with constraints $Ax \leq b$. One of the constraints is:

$$a^{\mathsf{T}}x \leq b, \tag{3}$$

where a satisfies $a \in \mathcal{A}$ and $\mathcal{A}$ is polyhedral

- We seek decisions x that are **robustly feasible**, i.e.,

$$a^{\mathsf{T}}x \leq b, \; \forall\, a \in \mathcal{A} := \{a \in \mathbb{R}^n \;:\; Ca \leq d\} \tag{4}$$

Infinitely many constraints : "semi-infinite" LP. *Any ideas?*