Duality

Lecture 4

October 1, 2025

Motivation

Consider an optimization problem

$$\begin{aligned} &\text{minimize } c^\mathsf{T} x \\ &\text{such that } Ax \leq b. \end{aligned}$$

Motivation

Consider an optimization problem

$$\begin{aligned} \text{minimize } & c^\mathsf{T} x \\ \text{such that } & Ax \leq b. \end{aligned}$$

1. *Given a feasible x, how can we know* **"how good"** *it is?*
   Formally, how to quantify the gap $c^\mathsf{T} x - p^\star$ where $p^\star$ is the optimal value?

Motivation

Consider an optimization problem

$$\begin{aligned} &\text{minimize } c^\top x \\ &\text{such that } Ax \le b. \end{aligned}$$

1. *Given a feasible x, how can we know* **"how good"** *it is?*
   Formally, how to quantify the gap $c^\top x - p^\star$ where $p^\star$ is the optimal value?
2. *Without a feasible x, how to* **certify** *that $\{x : Ax \le b\}$ is empty?*

Motivation

Consider an optimization problem

$$\begin{aligned} &\text{minimize } c^\mathsf{T} x \\ &\text{such that } Ax \leq b. \end{aligned}$$

1. *Given a feasible x, how can we know* **"how good"** *it is?*
   Formally, how to quantify the gap $c^\mathsf{T} x - p^\star$ where $p^\star$ is the optimal value?
2. *Without a feasible x, how to* **certify** *that $\{x : Ax \leq b\}$ is empty?*
3. Suppose one constraint is: $a_i^\mathsf{T} x \leq 0$ where $a_i \in \mathcal{A}$ are unknown parameters. *How to find an x that is feasible* **for any** *$a_i \in \mathcal{A}$?*

Motivation

Consider an optimization problem

$$\begin{aligned} &\text{minimize } c^\mathsf{T} x \\ &\text{such that } Ax \leq b. \end{aligned}$$

1. *Given a feasible x, how can we know* **"how good"** *it is?*
   Formally, how to quantify the gap $c^\mathsf{T} x - p^\star$ where $p^\star$ is the optimal value?
2. *Without a feasible x, how to* **certify** *that $\{x : Ax \leq b\}$ is empty?*
3. Suppose one constraint is: $a_i^\mathsf{T} x \leq 0$ where $a_i \in \mathcal{A}$ are unknown parameters. *How to find an x that is feasible* **for any** $a_i \in \mathcal{A}$*?*
4. You are offered a bit more of b_i, for a "suitable price". *Is the deal worthwhile?*

Duality theory will provide answers to these questions (and more)

Outline

- Consider a **primal** optimization problem:

$$(\mathcal{P}) \quad \text{minimize } c^{\mathsf{T}} x \quad \text{such that } Ax \le b.$$

Outline

- Consider a **primal** optimization problem:

$$(\mathcal{P}) \quad \text{minimize } c^{\mathsf{T}} x \quad \text{such that } Ax \leq b.$$

- We will form a **dual** problem; also a linear program (LP):

$$(\mathcal{D}) \quad \text{maximize } \tilde{r}^{\mathsf{T}} y \quad \text{such that } \tilde{A} y \leq \tilde{b}.$$

Outline

- Consider a **primal** optimization problem:

$$(\mathcal{P}) \quad \text{minimize } c^\mathsf{T} x \quad \text{such that } Ax \leq b.$$

- We will form a **dual** problem; also a linear program (LP):

$$(\mathcal{D}) \quad \text{maximize } \tilde{r}^\mathsf{T} y \quad \text{such that } \tilde{A} y \leq \tilde{b}.$$

- We will show that the dual provides lower bounds for the primal:

$$\tilde{r}^\mathsf{T} y \leq c^\mathsf{T} x \quad \text{for any } x \text{ feasible for } (\mathcal{P}) \text{ and } y \text{ feasible for } (\mathcal{D})$$

Outline

- Consider a **primal** optimization problem:

$$(\mathcal{P}) \quad \text{minimize } c^\mathsf{T} x \quad \text{such that } Ax \le b.$$

- We will form a **dual** problem; also a linear program (LP):

$$(\mathcal{D}) \quad \text{maximize } \tilde{r}^\mathsf{T} y \quad \text{such that } \tilde{A} y \le \tilde{b}.$$

- We will show that the dual provides lower bounds for the primal:

$$\tilde{r}^\mathsf{T} y \le c^\mathsf{T} x \quad \text{for any } x \text{ feasible for } (\mathcal{P}) \text{ and } y \text{ feasible for } (\mathcal{D})$$

- If $(\mathcal{P})$ has optimal solution $x^\star$, then $(\mathcal{D})$ has optimal solution $y^\star$ and

$$c^\mathsf{T} x^\star = \tilde{r}^\mathsf{T} y^\star \quad \textbf{(strong duality)}$$

Outline

- Consider a **primal** optimization problem:

$$(\mathcal{P}) \quad \text{minimize } c^\mathsf{T} x \quad \text{such that } Ax \leq b.$$

- We will form a **dual** problem; also a linear program (LP):

$$(\mathcal{D}) \quad \text{maximize } \tilde{r}^\mathsf{T} y \quad \text{such that } \tilde{A} y \leq \tilde{b}.$$

- We will show that the dual provides lower bounds for the primal:

$$\tilde{r}^\mathsf{T} y \leq c^\mathsf{T} x \quad \text{for any } x \text{ feasible for } (\mathcal{P}) \text{ and } y \text{ feasible for } (\mathcal{D})$$

- If $(\mathcal{P})$ has optimal solution $x^\star$, then $(\mathcal{D})$ has optimal solution $y^\star$ and

$$c^\mathsf{T} x^\star = \tilde{r}^\mathsf{T} y^\star \quad \textbf{(strong duality)}$$

- In the process, will uncover some **fundamental ideas in optimization**:

separation of convex sets $\Longrightarrow$ **Farkas Lemma** $\Longrightarrow$ **strong duality**

Deriving Lower Bounds

Consider a linear optimization problem in the most general form possible:

Primal Problem

$$
\begin{aligned}
(\mathcal{P})\ \text{minimize}_x \quad & c^\mathsf{T} x \\
\text{such that} \quad & a_i^\mathsf{T} x \geq b_i, && \forall i \in I_{\text{ge}}, \\
& a_i^\mathsf{T} x \leq b_i, && \forall i \in I_{\text{le}}, \\
& a_i^\mathsf{T} x = b_i, && \forall i \in I_{\text{eq}}, \\
& x_j \geq 0, && \forall j \in J_p, \\
& x_j \leq 0, && \forall j \in J_n, \\
& x_j \text{ free}, && \forall j \in J_f \\
\text{variable} \quad & x \in \mathbb{R}^n.
\end{aligned} \tag{1}
$$

Note the mnemonic encoding...

Deriving Lower Bounds

Consider a linear optimization problem in the most general form possible:

Primal Problem

$$
\begin{array}{rlll}
(\mathcal{P})\ \text{minimize}_x & c^\mathsf{T} x & & \\
\text{such that} & a_i^\mathsf{T} x \geq b_i, & \forall i \in I_{\text{ge}}, & \\
& a_i^\mathsf{T} x \leq b_i, & \forall i \in I_{\text{le}}, & A = \begin{bmatrix} A^{\text{le}} \\ A^{\text{ge}} \\ A^{\text{eq}} \end{bmatrix} \in \mathbb{R}^{m \times n} \\
& a_i^\mathsf{T} x = b_i, & \forall i \in I_{\text{eq}}, & \\
& x_j \geq 0, & \forall j \in J_p, & \\
& x_j \leq 0, & \forall j \in J_n, & b = \begin{bmatrix} b^{\text{le}} \\ b^{\text{ge}} \\ b^{\text{eq}} \end{bmatrix} \in \mathbb{R}^{m} \\
& x_j \text{ free}, & \forall j \in J_f & \\
\text{variable} & x \in \mathbb{R}^n. & &
\end{array} \tag{1}
$$

Note the mnemonic encoding...

Deriving Lower Bounds

Consider a linear optimization problem in the most general form possible:

Primal Problem

$$
\begin{array}{rlll}
(\mathcal{P}) \text{ minimize}_x & c^\mathsf{T} x & & \\
\text{such that} & a_i^\mathsf{T} x \geq b_i, & \forall i \in I_{\text{ge}}, & \\
& a_i^\mathsf{T} x \leq b_i, & \forall i \in I_{\text{le}}, & A = \begin{bmatrix} A^{\text{le}} \\ A^{\text{ge}} \\ A^{\text{eq}} \end{bmatrix} \in \mathbb{R}^{m \times n} \\
& a_i^\mathsf{T} x = b_i, & \forall i \in I_{\text{eq}}, & \\
& x_j \geq 0, & \forall j \in J_p, & \\
& x_j \leq 0, & \forall j \in J_n, & b = \begin{bmatrix} b^{\text{le}} \\ b^{\text{ge}} \\ b^{\text{eq}} \end{bmatrix} \in \mathbb{R}^{m} \\
& x_j \text{ free}, & \forall j \in J_f & \\
\text{variable} & x \in \mathbb{R}^n. & &
\end{array}
\tag{1}
$$

Note the mnemonic encoding...

Definition

We will refer to this as the **primal problem** or problem $(\mathcal{P})$.

Let P denote its feasible set (a polyhedron), and $p^\star$ denote its optimal value.

Deriving Lower Bounds

Consider the primal problem:

$$
\begin{array}{rlll}
(\mathcal{P}) \ \text{minimize}_x & c^\mathsf{T} x & & \\
\text{such that} & a_i^\mathsf{T} x \geq b_i, & \forall i \in I_{\text{ge}}, \\
& a_i^\mathsf{T} x \leq b_i, & \forall i \in I_{\text{le}}, \\
& a_i^\mathsf{T} x = b_i, & \forall i \in I_{\text{eq}}, \\
& x_j \geq 0, & \forall j \in J_p, \\
& x_j \leq 0, & \forall j \in J_n, \\
& x_j \ \text{free}, & \forall j \in J_f \\
\text{variable} & x \in \mathbb{R}^n. & &
\end{array}
$$

$(\mathcal{P})$ is a minimization; we seek **valid lower bounds** on $(\mathcal{P})$. *Any ideas?*

Deriving Lower Bounds

Consider the primal problem:

$$
\begin{array}{rlll}
(\mathcal{P}) \ \text{minimize}_x & c^\mathsf{T} x & & \\
\text{such that} & a_i^\mathsf{T} x \geq b_i, & \forall i \in I_{\text{ge}}, \\
& a_i^\mathsf{T} x \leq b_i, & \forall i \in I_{\text{le}}, \\
& a_i^\mathsf{T} x = b_i, & \forall i \in I_{\text{eq}}, \\
& x_j \geq 0, & \forall j \in J_p, \\
& x_j \leq 0, & \forall j \in J_n, \\
& x_j \ \text{free}, & \forall j \in J_f \\
\text{variable} & x \in \mathbb{R}^n. & &
\end{array}
$$

$(\mathcal{P})$ is a minimization; we seek **valid lower bounds** on $(\mathcal{P})$. *Any ideas?*

Deriving Lower Bounds

Consider the primal problem:

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

For every constraint i, have a **penalty** λ_i

Construct the **lower bound** as the **Lagrangean**:

$$
\mathcal{L}(x, \lambda) = c^\mathsf{T} x - \sum_{i=1}^{m} \lambda_i (a_i^\mathsf{T} x - b_i) = c^\mathsf{T} x - \lambda^\mathsf{T} (Ax - b)
$$

Deriving Lower Bounds

Consider the primal problem:

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

For every constraint i, have a **penalty** λ_i

Construct the **lower bound** as the **Lagrangean**:

$$\mathcal{L}(x, \lambda) = c^\mathsf{T} x - \sum_{i=1}^{m} \lambda_i (a_i^\mathsf{T} x - b_i) = c^\mathsf{T} x - \lambda^\mathsf{T}(Ax - b)$$

Note: we relaxed the complicating constraints, $a_i^\mathsf{T} x$? b_i, and used a linear penalty

Not apriori clear that this will give us very good bounds...

Deriving Lower Bounds

Consider the primal problem:

$$
\begin{array}{rll}
(\mathcal{P})\ \text{minimize}_x & c^\mathsf{T} x & \\
(\lambda_i \to) & a_i^\mathsf{T} x \geq b_i, & \forall i \in I_{\text{ge}}, \\
(\lambda_i \to) & a_i^\mathsf{T} x \leq b_i, & \forall i \in I_{\text{le}}, \\
(\lambda_i \to) & a_i^\mathsf{T} x = b_i, & \forall i \in I_{\text{eq}}, \\
 & x_j \geq 0, & \forall j \in J_p, \\
 & x_j \leq 0, & \forall j \in J_n, \\
 & x_j \ \text{free}, & \forall j \in J_f.
\end{array}
$$

We want the Lagrangean to give us **a valid lower bound**:

$$
\mathcal{L}(x, \lambda) = c^\mathsf{T} x - \lambda^\mathsf{T}(Ax - b) \leq c^\mathsf{T} x,\ \forall x \in P.
$$

Deriving Lower Bounds

Consider the primal problem:

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

We want the Lagrangean to give us **a valid lower bound**:

$$\mathcal{L}(x, \lambda) = c^{\mathsf{T}}x - \lambda^{\mathsf{T}}(Ax - b) \leq c^{\mathsf{T}}x,\ \forall x \in P.$$

We must impose constraints on λ:

Deriving Lower Bounds

Consider the primal problem:

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

We want the Lagrangean to give us **a valid lower bound**:

$$
\mathcal{L}(x, \lambda) = c^{\mathsf{T}}x - \lambda^{\mathsf{T}}(Ax - b) \leq c^{\mathsf{T}}x,\ \forall x \in P.
$$

We must impose constraints on λ:

$$
\left.
\begin{array}{ll}
\lambda_i \geq 0, & \forall i \in I_{\text{ge}} \\
\lambda_i \leq 0, & \forall i \in I_{\text{le}} \\
\lambda_i \ \text{free}, & \forall i \in I_{\text{eq}}.
\end{array}
\right\} \Leftrightarrow \lambda \in \Lambda \tag{2}
$$

Deriving Lower Bounds

Summarizing... any $\lambda \in \Lambda$ produces **a valid lower bound**:

$$\mathcal{L}(x, \lambda) = c^{\mathsf{T}}x - \lambda^{\mathsf{T}}(Ax - b) \leq c^{\mathsf{T}}x, \ \forall x \in P.$$

How can we get a lower bound on the primal's **optimal value** $p^{\star}$*?*

Deriving Lower Bounds

Summarizing... any $\lambda \in \Lambda$ produces **a valid lower bound**:

$$\mathcal{L}(x, \lambda) = c^{\mathsf{T}} x - \lambda^{\mathsf{T}}(Ax - b) \leq c^{\mathsf{T}} x, \ \forall x \in P.$$

How can we get a lower bound on the primal's **optimal value** *$p^\star$?*

Claim

The function $g : \Lambda \to \mathbb{R}$ defined as:

$$\begin{aligned} g(\lambda) := min_x \ & \mathcal{L}(x, \lambda) \\ s.t. \ & x_j \geq 0, \ \forall j \in J_p \\ & x_j \leq 0, \ \forall j \in J_n \\ & x_j \ free, \ \forall j \in J_f \end{aligned} \tag{3}$$

satisfies $g(\lambda) \leq p^\star$ for any $\lambda \in \Lambda$.

Note: including the sign constraints on x in this optimization improves the lower bound!

Deriving Lower Bounds

Let us analyze this further:

$$
\begin{aligned}
g(\lambda) = \min_x \ & \mathcal{L}(x, \lambda) & = \min_x \ & \left[\lambda^\mathsf{T} b + (c^\mathsf{T} - \lambda^\mathsf{T} A)x\right] \\
\text{s.t. } & x_j \geq 0, \ \forall j \in J_p, & \text{s.t. } & x_j \geq 0, \ \forall j \in J_p, \\
& x_j \leq 0, \ \forall j \in J_n, & & x_j \leq 0, \ \forall j \in J_n, \\
& x_j \text{ free}, \ \forall j \in J_f & & x_j \text{ free}, \ \forall j \in J_f
\end{aligned}
$$

Deriving Lower Bounds

Let us analyze this further:

$$
\begin{aligned}
g(\lambda) = \min_x \ & \mathcal{L}(x, \lambda) & = \min_x \ & \left[\lambda^\mathsf{T} b + (c^\mathsf{T} - \lambda^\mathsf{T} A)x\right] \\
\text{s.t. } & x_j \geq 0, \ \forall j \in J_p, & \text{s.t. } & x_j \geq 0, \ \forall j \in J_p, \\
& x_j \leq 0, \ \forall j \in J_n, & & x_j \leq 0, \ \forall j \in J_n, \\
& x_j \text{ free}, \ \forall j \in J_f & & x_j \text{ free}, \ \forall j \in J_f
\end{aligned}
$$

$$
g(\lambda) = \begin{cases} \lambda^\mathsf{T} b, & \text{if } \lambda^\mathsf{T} A_j \leq c_j, \forall j \in J_p \textbf{ and } \lambda^\mathsf{T} A_j \geq c_j, \forall j \in J_n \textbf{ and } \lambda^\mathsf{T} A_j = c_j, \forall j \in J_f \\ -\infty, & \text{otherwise.} \end{cases}
$$

Deriving the Dual Problem

$$g(\lambda) = \begin{cases} \lambda^{\mathsf{T}} b, & \text{if } \lambda^{\mathsf{T}} A_j \leq c_j, \forall j \in J_p \textbf{ and } \lambda^{\mathsf{T}} A_j \geq c_j, \forall j \in J_n \textbf{ and } \lambda^{\mathsf{T}} A_j = c_j, \forall j \in J_f \\ -\infty, & \text{otherwise.} \end{cases}$$

is **a valid lower bound** on the primal **optimal value**: $g(\lambda) \leq p^\star$ for any $\lambda \in \Lambda$.

How can we get the **best** *lower bound?*

Deriving the Dual Problem

$$g(\lambda) = \begin{cases} \lambda^{\mathsf{T}} b, & \text{if } \lambda^{\mathsf{T}} A_j \leq c_j, \forall j \in J_p \textbf{ and } \lambda^{\mathsf{T}} A_j \geq c_j, \forall j \in J_n \textbf{ and } \lambda^{\mathsf{T}} A_j = c_j, \forall j \in J_f \\ -\infty, & \text{otherwise.} \end{cases}$$

is **a valid lower bound** on the primal **optimal value**: $g(\lambda) \leq p^\star$ for any $\lambda \in \Lambda$.

How can we get the **best** *lower bound?*

$$\underset{\lambda \in \Lambda}{\text{maximize}}\, g(\lambda) \tag{4}$$

This is equivalent to the following optimization problem:

Deriving the Dual Problem

$$g(\lambda) = \begin{cases} \lambda^{\mathsf{T}} b, & \text{if } \lambda^{\mathsf{T}} A_j \le c_j, \forall j \in J_p \textbf{ and } \lambda^{\mathsf{T}} A_j \ge c_j, \forall j \in J_n \textbf{ and } \lambda^{\mathsf{T}} A_j = c_j, \forall j \in J_f \\ -\infty, & \text{otherwise.} \end{cases}$$

is **a valid lower bound** on the primal **optimal value**: $g(\lambda) \le p^\star$ for any $\lambda \in \Lambda$.

How can we get the **best** *lower bound?*

$$\underset{\lambda \in \Lambda}{\text{maximize}}\, g(\lambda) \tag{4}$$

This is equivalent to the following optimization problem:

Dual Problem

$$\begin{aligned} \text{maximize} \quad & \lambda^{\mathsf{T}} b \\ \text{subject to} \quad & \lambda_i \ge 0, && \forall i \in I_{\text{ge}}, \\ & \lambda_i \le 0, && \forall i \in I_{\text{le}}, \\ & \lambda_i \text{ free}, && \forall i \in I_{\text{eq}}, \\ & \lambda^{\mathsf{T}} A_j \le c_j, && \forall j \in J_p, \\ & \lambda^{\mathsf{T}} A_j \ge c_j, && \forall j \in J_n, \\ & \lambda^{\mathsf{T}} A_j = c_j, && \forall j \in J_f. \end{aligned} \tag{5}$$

Deriving the Dual Problem

Dual Problem

$$
\begin{array}{rll}
\text{maximize} & \lambda^{\mathsf{T}} b & \\
\text{subject to} & \lambda_i \geq 0, & \forall i \in I_{\text{ge}}, \\
& \lambda_i \leq 0, & \forall i \in I_{\text{le}}, \\
& \lambda_i \text{ free}, & \forall i \in I_{\text{eq}}, \\
& \lambda^{\mathsf{T}} A_j \leq c_j, & \forall j \in J_p, \\
& \lambda^{\mathsf{T}} A_j \geq c_j, & \forall j \in J_n, \\
& \lambda^{\mathsf{T}} A_j = c_j, & \forall j \in J_f.
\end{array}
\tag{6}
$$

Deriving the Dual Problem

Dual Problem

$$
\begin{aligned}
\text{maximize} \quad & \lambda^{\mathsf{T}} b \\
\text{subject to} \quad & \lambda_i \geq 0, && \forall i \in I_{\text{ge}}, \\
& \lambda_i \leq 0, && \forall i \in I_{\text{le}}, \\
& \lambda_i \text{ free}, && \forall i \in I_{\text{eq}}, \\
& \lambda^{\mathsf{T}} A_j \leq c_j, && \forall j \in J_p, \\
& \lambda^{\mathsf{T}} A_j \geq c_j, && \forall j \in J_n, \\
& \lambda^{\mathsf{T}} A_j = c_j, && \forall j \in J_f.
\end{aligned}
\tag{6}
$$

Definition

This is the **dual** of $(\mathcal{P})$, which we will also refer to as $(\mathcal{D})$. We denote its feasible set with D and its optimal value with $d^\star$.

Note: The dual is also a linear optimization problem!

Primal-Dual Pair

Primal-Dual Pair of Problems

Primal $(\mathcal{P})$

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

Dual $(\mathcal{D})$

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

Primal-Dual Pair

Primal-Dual Pair of Problems

Primal $(\mathcal{P})$

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

Dual $(\mathcal{D})$

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

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 λ_i)

Primal-Dual Pair

Primal-Dual Pair of Problems

Primal $(\mathcal{P})$

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

Dual $(\mathcal{D})$

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

Exercise

Rewrite the dual problem as a minimization problem and construct its dual.

Primal-Dual Pair

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 $(\mathcal{D})$

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

Exercise

Rewrite the dual problem as a minimization problem and construct its dual.

Rules for Constructing the Dual of Any LP

Consider any linear optimization problem (minimization/maximization):

$$\begin{aligned}
\underset{x}{\text{minimize}} \quad / \quad \underset{x}{\text{maximize}} \quad & c^{\mathsf{T}}x \\
({\color{red}\lambda} \rightarrow) \quad & Ax \underset{\geq}{\overset{\leq}{=}} b \\
& x \underset{\geq}{\overset{\leq}{=}} 0
\end{aligned} \tag{7}$$

Rules for Constructing the Dual of Any LP

Consider any linear optimization problem (minimization/maximization):

$$\begin{array}{rl} \underset{x}{\text{minimize}} \ / \ \underset{x}{\text{maximize}} & c^{\mathsf{T}}x \\ (\lambda \rightarrow) & Ax \lesseqqgtr b \\ & x \lesseqqgtr 0 \end{array} \tag{7}$$

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.

Rules for Constructing the Dual of Any LP

Consider any linear optimization problem (minimization/maximization):

$$\begin{aligned} \underset{x}{\text{minimize}} \ / \ \underset{x}{\text{maximize}} \quad & c^{\mathsf{T}} x \\ (\lambda \rightarrow) \quad & Ax \begin{smallmatrix} \leq \\ = \\ \geq \end{smallmatrix} b \\ & x \begin{smallmatrix} \leq \\ = \\ \geq \end{smallmatrix} 0 \end{aligned} \tag{7}$$

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.

Rules for Constructing the Dual of Any LP

Consider any linear optimization problem (minimization/maximization):

$$\begin{aligned} \underset{x}{\text{minimize}} \quad / \quad \underset{x}{\text{maximize}} \quad & c^{\mathsf{T}} x \\ (\lambda \rightarrow) \quad & Ax \lesseqqgtr b \\ & x \lesseqqgtr 0 \end{aligned} \tag{7}$$

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

Rules for Constructing the Dual of Any LP

Consider any linear optimization problem (minimization/maximization):

$$\begin{array}{rr} \underset{x}{\text{minimize}} \ / \ \underset{x}{\text{maximize}} & c^\mathsf{T}x \\ (\lambda \rightarrow) & Ax \begin{smallmatrix}\leq\\=\\\geq\end{smallmatrix} b \\ & x \begin{smallmatrix}\leq\\\geq\end{smallmatrix} 0 \end{array} \tag{7}$$

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 1

$$
\begin{aligned}
(\mathcal{P}) \quad & \max\ 3x_1 + 2x_2 & \\
& \text{s.t. } x_1 + 2x_2 \leq 4 & (1)\\
& \phantom{\text{s.t. }} 3x_1 + 2x_2 \geq 6 & (2)\\
& \phantom{\text{s.t. }} x_1 - x_2 = 1 & (3)\\
& \phantom{\text{s.t. }} x_1,\ x_2 \geq 0. &
\end{aligned}
$$

Some Quick Results

Theorem ("Duals of equivalent primals")

If we transform a primal P_1 into an equivalent formulation P_2 by:

- *replacing a free variable x_i with $x_i = x_i^+ - x_i^-$,*
- *replacing an inequality with an equality by introducing a slack variable,*
- *removing linearly dependent rows $a_i^\top$ for a* **feasible** *LP in standard form,*

then the duals of (P_1) and (P_2) are **equivalent**, *i.e., they are either both infeasible or they have the same optimal objective.*

Weak duality

	Primal $(\mathcal{P})$			**Dual** $(\mathcal{D})$	
minimize_{x}	$c^{\mathsf{T}} x$		$\text{maximize}_{\lambda}$	$\lambda^{\mathsf{T}} b$	
$(\lambda_i \rightarrow)$	$a_i^{\mathsf{T}} x \geq b_i,$	$\forall i \in I_{\text{ge}},$		$\lambda_i \geq 0,$	$\forall i \in I_{\text{ge}},$
$(\lambda_i \rightarrow)$	$a_i^{\mathsf{T}} x \leq b_i,$	$\forall i \in I_{\text{le}},$		$\lambda_i \leq 0,$	$\forall i \in I_{\text{le}},$
$(\lambda_i \rightarrow)$	$a_i^{\mathsf{T}} x = b_i,$	$\forall i \in I_{\text{eq}},$		λ_i 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 free,	$\forall j \in J_f.$	$(x_j \rightarrow)$	$\lambda^{\mathsf{T}} A_j = c_j,$	$\forall j \in J_f.$

Weak duality

	Primal $(\mathcal{P})$			**Dual** $(\mathcal{D})$	
minimize_x	$c^\mathsf{T} x$		maximize_λ	$\lambda^\mathsf{T} b$	
$(\lambda_i \to)$	$a_i^\mathsf{T} x \geq b_i,$	$\forall i \in I_{\text{ge}},$		$\lambda_i \geq 0,$	$\forall i \in I_{\text{ge}},$
$(\lambda_i \to)$	$a_i^\mathsf{T} x \leq b_i,$	$\forall i \in I_{\text{le}},$		$\lambda_i \leq 0,$	$\forall i \in I_{\text{le}},$
$(\lambda_i \to)$	$a_i^\mathsf{T} x = b_i,$	$\forall i \in I_{\text{eq}},$		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,	$\forall j \in J_f.$	$(x_j \to)$	$\lambda^\mathsf{T} A_j = c_j,$	$\forall j \in J_f.$

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:

(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?

Implications of Weak Duality

Corollary

The following results hold:

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

(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?

Implications of Weak Duality

Corollary

The following results hold:

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

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

(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?

Implications of Weak Duality

Corollary

The following results hold:

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

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

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

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

(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?

Implications of Weak Duality

Corollary

The following results hold:

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

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

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

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

(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...

Implications of Weak Duality

Corollary

The following results hold:

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

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

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

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

(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, $p^\star = d^\star$.

Strong duality

Theorem (Strong duality)

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

Proof. Many proofs possible...

- See Bertsimas & Tsitsiklis for a proof involving the simplex algorithm
- We provide a more general proof (some ideas work for **convex** optimization)

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

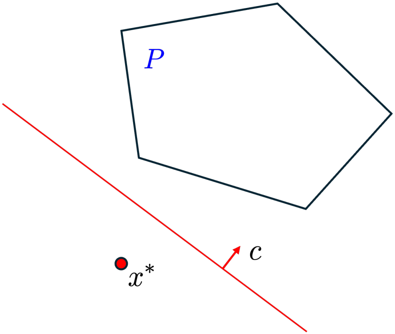
Separating Hyperplane Theorem

The first **fundamental result in optimization**

Separating Hyperplane Theorem

Theorem (Simple Separating Hyperplane Theorem)

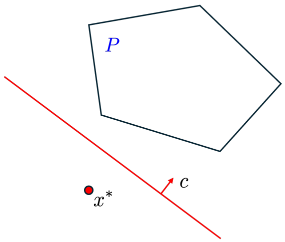
Consider a point $x^\star$ and a polyhedron P. If $x^\star \notin P$, then there exists a vector $c \in \mathbb{R}^n$ such that $c \neq 0$ and $c^\intercal x^\star < c^\intercal y$ holds for all $y \in P$.

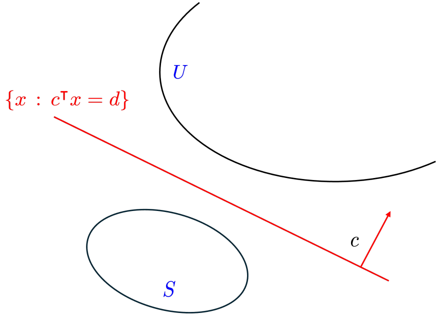


Separating Hyperplane Theorem

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 $U \subset \{x \in \mathbb{R}^n : c^\mathsf{T} x > d\}$.

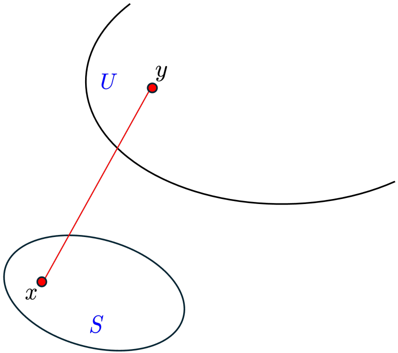
Proof.

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 $x \in S, y \in U$

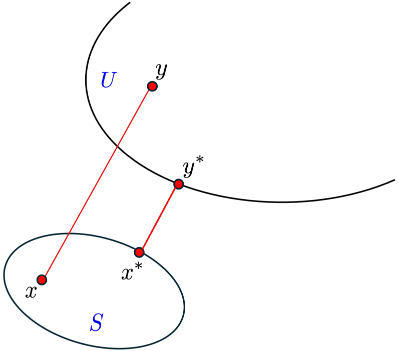


Separating Hyperplane Theorem

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 $U \subset \{x \in \mathbb{R}^n : c^\mathsf{T} x > d\}$.

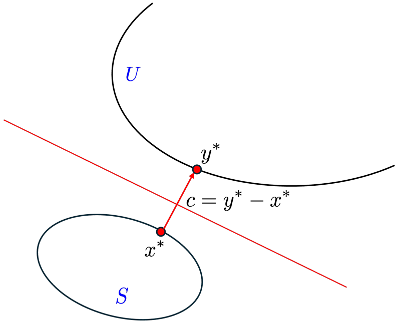
Proof. Argue that the minimum is achieved, at $x^\star, y^\star$



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 bounded

Separating Hyperplane Theorem - Caveats!

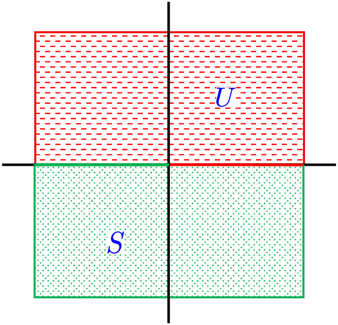
Both conditions in the theorem needed: closed and at least one bounded

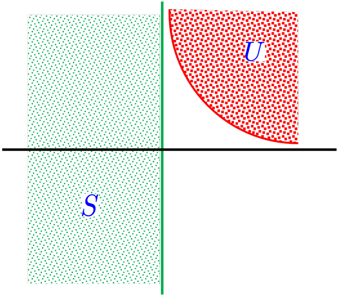
- **Left:** two convex sets that are **not closed** but are both bounded:

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

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

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



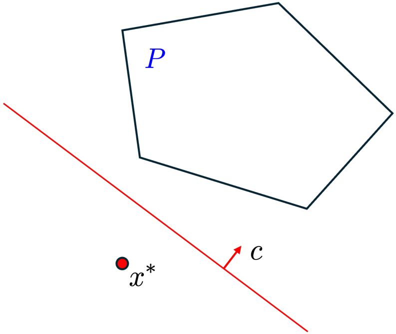


Needed For Our Purposes

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 $x \in P$.



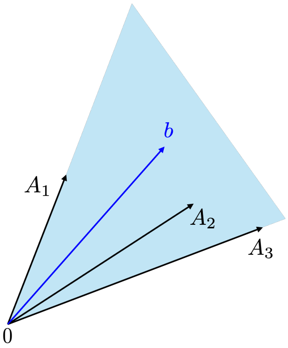
Farkas Lemma

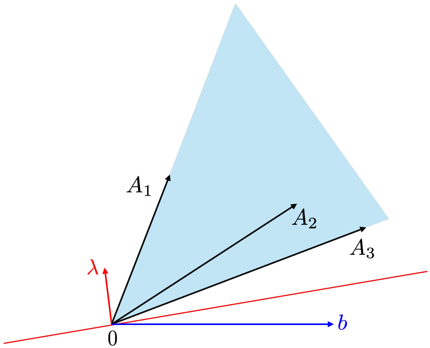
Time for the **second fundamental result in optimization**!

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 $Ax = b$.
(b) There exists some vector λ such that $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

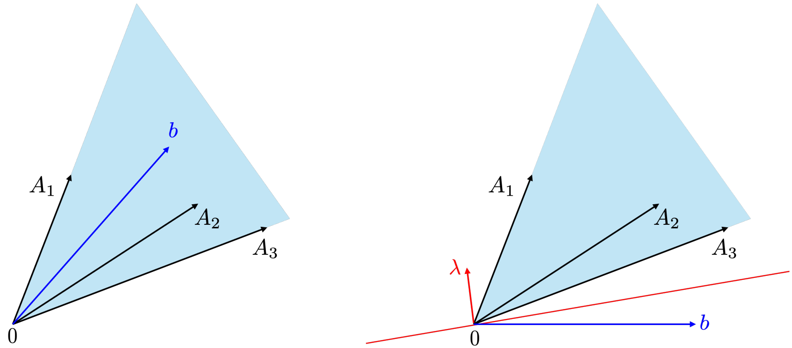
Farkas Lemma

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 $\lambda^T A \geq 0$ and $\lambda^T b < 0$.



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 $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

Proof. "(a) $\Rightarrow$ not (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 $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

Proof. "(a) $\Rightarrow$ not (b)."

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

(b) implies $\exists \lambda : \lambda^T A \geq 0$.

But then $\lambda^T b = \lambda^T Ax \geq 0$, so (b) cannot hold.

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* $\lambda^T A \geq 0$ *and* $\lambda^T b < 0$.

"not (a) $\Rightarrow$ **(b)."** Want to use the separating hyperplane theorem.

Farkas Lemma

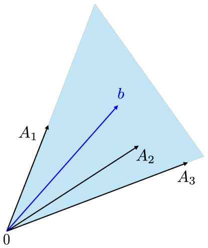
Theorem (Farkas' Lemma)

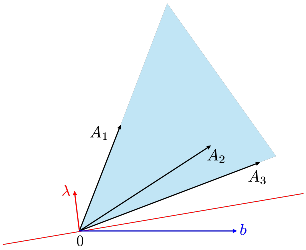
For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, exactly one of the following two alternatives holds:
(a) There exists some $x \geq 0$ such that $Ax = b$.
(b) There exists some vector λ such that $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

"not (a) $\Rightarrow$ (b)." Want to use the separating hyperplane theorem.

- Assume $\nexists x \geq 0 : Ax = b$. This implies that $b \notin S$ where:

$$S := \{Ax \, : \, x \geq 0\} = \{y \, : \, \exists \, x \geq 0 \, \text{such that} \, y = Ax\}.$$





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 $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

"not (a) $\Rightarrow$ (b)." Want to use the separating hyperplane theorem.

- Assume $\nexists x \geq 0 : Ax = b$. This implies that $b \notin S$ where:

$$S := \{Ax \, : \, x \geq 0\} = \{y \, : \, \exists\, x \geq 0 \text{ such that } y = Ax\}.$$

- S is convex.
- To apply separating hyperplane theorem, need S **closed**!

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 $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

"not (a) $\Rightarrow$ (b)." Want to use the separating hyperplane theorem.

- Assume $\nexists x \geq 0 : Ax = b$. This implies that $b \notin S$ where:

$$S := \{Ax \,:\, x \geq 0\} = \{y \,:\, \exists\, x \geq 0 \text{ such that } y = Ax\}.$$

- S is convex.
- To apply separating hyperplane theorem, need S **closed**!
  - S is the projection of $\bar{S} := \{(x, y) \,:\, x \geq 0,\, y = Ax\}$ on the y variables.

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 $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

"not (a) $\Rightarrow$ (b)." Want to use the separating hyperplane theorem.

- Assume $\nexists x \geq 0 : Ax = b$. This implies that $b \notin S$ where:

$$S := \{Ax \,:\, x \geq 0\} = \{y \,:\, \exists\, x \geq 0 \,\text{such that}\, y = Ax\}.$$

- S is convex.
- To apply separating hyperplane theorem, need S **closed**!
  - S is the projection of $\bar{S} := \{(x, y) \,:\, x \geq 0,\, y = Ax\}$ on the y variables.
  - The projection of a polyhedron is another polyhedron.

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 $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

"not (a) $\Rightarrow$ (b)." Want to use the separating hyperplane theorem.

- Assume $\nexists x \geq 0 : Ax = b$. This implies that $b \notin S$ where:

$$S := \{Ax \, : \, x \geq 0\} = \{y \, : \, \exists \, x \geq 0 \text{ such that } y = Ax\}.$$

- S is convex.
- To apply separating hyperplane theorem, need S **closed**!
  - S is the projection of $\bar{S} := \{(x, y) \, : \, x \geq 0, \, y = Ax\}$ on the y variables.
  - The projection of a polyhedron is another polyhedron.
  - Every polyhedron is closed.

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 $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

"not (a) $\Rightarrow$ (b)." Want to use the separating hyperplane theorem.

- Assume $\nexists x \geq 0 : Ax = b$. This implies that $b \notin S$ where:

$$S := \{Ax \,:\, x \geq 0\} = \{y \,:\, \exists\, x \geq 0 \,\text{such that}\, y = Ax\}.$$

- S is convex.
- To apply separating hyperplane theorem, need S **closed**!
  - S is the projection of $\bar{S} := \{(x, y) \,:\, x \geq 0,\, y = Ax\}$ on the y variables.
  - The projection of a polyhedron is another polyhedron.
  - Every polyhedron is closed.

$\Rightarrow$ S is closed.

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 $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

"not (a) $\Rightarrow$ (b)." (cont'd)

- $S := \{Ax : x \geq 0\} = \{y : \exists\, x \geq 0 \text{ such that } y = Ax\}$ is convex and closed.
- Sep. Hyp. Thm. 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 $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

"not (a) $\Rightarrow$ (b)." (cont'd)

- $S := \{Ax : x \geq 0\} = \{y : \exists\, x \geq 0 \text{ such that } y = Ax\}$ is convex and closed.
- Sep. Hyp. Thm. 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 $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

"not (a) $\Rightarrow$ (b)." (cont'd)

- $S := \{Ax : x \geq 0\} = \{y : \exists\, x \geq 0 \,\text{such that}\, y = Ax\}$ is convex and closed.
- Sep. Hyp. Thm. 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 $\lambda A_i \in S$ for every $\lambda > 0$, so

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

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 $\lambda^T A \geq 0$ and $\lambda^T b < 0$.

"not (a) $\Rightarrow$ (b)." (cont'd)

- $S := \{Ax : x \geq 0\} = \{y : \exists\, x \geq 0 \,\text{such that}\, y = Ax\}$ is convex and closed.
- Sep. Hyp. Thm. 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 $\lambda A_i \in S$ for every $\lambda > 0$, so

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

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