Duality

Lecture 4

October 2, 2024

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

$$\text{minimize } c^\intercal x$$
$$\text{such that } Ax \leq b.$$

1. *Given a feasible x, how can we know* **"how good"** *it is?*
   Formally, how to quantify the gap $c^\intercal x - z^*$ where z^* is the optimal value?

Motivation

Consider an optimization problem

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

1. *Given a feasible x, how can we know* **"how good"** *it is?*
   Formally, how to quantify the gap $c^\intercal x - z^*$ where z^* is the optimal value?

2. *Without a feasible x, how to* **certify** *that $\{x : Ax \leq b\}$ is empty?*

Motivation

Consider an optimization problem

$$\text{minimize } c^\intercal x$$
$$\text{such that } Ax \leq b.$$

1. *Given a feasible x, how can we know* **"how good"** *it is?*
   Formally, how to quantify the gap $c^\intercal x - z^*$ where z^* is the optimal value?

2. *Without a feasible x, how to* **certify** *that $\{x : Ax \leq b\}$ is empty?*

3. Suppose one constraint is: $a_i^\intercal x \leq 0$ where $a_i \in \mathcal{A}$ are unknown parameters.
   How can we ensure this constraint is feasible **for any** $a_i \in \mathcal{A}$*?*

Motivation

Consider an optimization problem

$$\text{minimize } c^\intercal x$$
$$\text{such that } Ax \le b.$$

1. *Given a feasible x, how can we know* **"how good"** *it is?*
   Formally, how to quantify the gap $c^\intercal x - z^*$ where z^* is the optimal value?

2. *Without a feasible x, how to* **certify** *that $\{x : Ax \le b\}$ is empty?*

3. Suppose one constraint is: $a_i^\intercal x \le 0$ where $a_i \in \mathcal{A}$ are unknown parameters.
   How can we ensure this constraint 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:

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

Outline

- Consider a **primal** optimization problem:

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

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

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

Outline

- Consider a **primal** optimization problem:

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

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

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

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

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

Outline

- Consider a **primal** optimization problem:

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

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

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

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

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

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

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

Outline

- Consider a **primal** optimization problem:

$$(\mathscr{P}) \quad \text{minimize } c^\intercal x \quad \text{such that } Ax \leq b.$$

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

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

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

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

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

$$c^\intercal x^* = \tilde{r}^\intercal y^* \quad \textbf{(strong duality)}$$

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

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

Notation

A_j will denote the j-th column of matrix $A \in \mathbb{R}^{m \times n}$

$$A = \begin{bmatrix} A_1 & A_2 & \dots & A_j & \dots & A_n \end{bmatrix}$$

For $S \subseteq \{1, \dots, n\}$, A_S is the submatrix obtained from columns $\{A_j\}_{j \in S}$

$$\text{e.g., for } S = \{1, 3\},\ A_S = \begin{bmatrix} A_1 & A_3 \end{bmatrix}$$

a_i^T will denote the i-th row of matrix $A \in \mathbb{R}^{m \times n}$

$$A^\mathsf{T} = \begin{bmatrix} a_1 & a_2 & \dots & a_i & \dots & a_m \end{bmatrix}$$

Notation

A_j will denote the j-th column of matrix $A \in \mathbb{R}^{m \times n}$

$$A = \left[\begin{array}{cccccc} A_1 & A_2 & \dots & A_j & \dots & A_n \end{array}\right]$$

For $S \subseteq \{1, \dots, n\}$, A_S is the submatrix obtained from columns $\{A_j\}_{j \in S}$

$$\text{e.g., for } S = \{1, 3\},\ A_S = \left[\begin{array}{cc} A_1 & A_3 \end{array}\right]$$

$a_i^\intercal$ will denote the i-th row of matrix $A \in \mathbb{R}^{m \times n}$

$$A^\intercal = \left[\begin{array}{cccccc} a_1 & a_2 & \dots & a_i & \dots & a_m \end{array}\right]$$

For $x \in \mathbb{R}^n$, two ways to view the expression Ax:

$$Ax = \sum_{j=1}^{n} A_j x_j = \left[\begin{array}{c} a_1^\intercal x \\ a_2^\intercal x \\ \vdots \\ a_m^\intercal x \end{array}\right].$$

We use $\|\cdot\|$ to denote the Euclidean norm: $\|x\| = (x^\intercal x)^{1/2}$.

Deriving the Dual Problem

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

$$
\begin{aligned}
(\mathscr{P})\ \text{minimize}_x \quad & c^\intercal x \\
& a_i^\intercal x \geq b_i, && i \in M_1, \\
& a_i^\intercal x \leq b_i, && i \in M_2, \\
& a_i^\intercal x = b_i, && i \in M_3, \\
& x_j \geq 0, && j \in N_1, \\
& x_j \leq 0, && j \in N_2, \\
& x_j \text{ free}, && j \in N_3.
\end{aligned}
\tag{1}
$$

We will refer to this as the **primal** problem, and also as problem $(\mathscr{P})$

We will also denote its feasible set with P (a polyhedron)

Let's also assume for now that $(\mathscr{P})$ has an optimal solution x^*

Deriving the Dual Problem

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

$$
\begin{array}{rlll}
(\mathscr{P}) \ \text{minimize}_x & c^\intercal x & & \\
(p_i \rightarrow) & a_i^\intercal x \geq b_i, & i \in M_1, \\
(p_i \rightarrow) & a_i^\intercal x \leq b_i, & i \in M_2, \\
(p_i \rightarrow) & a_i^\intercal x = b_i, & i \in M_3, \\
& x_j \geq 0, & j \in N_1, \\
& x_j \leq 0, & j \in N_2, \\
& x_j \ \text{free}, & j \in N_3.
\end{array}
$$

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

Deriving the Dual Problem

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

$$
\begin{array}{rlll}
(\mathscr{P}) \text{ minimize}_x & c^\intercal x & & \\
(p_i \rightarrow) & a_i^\intercal x \geq b_i, & i \in M_1, \\
(p_i \rightarrow) & a_i^\intercal x \leq b_i, & i \in M_2, \\
(p_i \rightarrow) & a_i^\intercal x = b_i, & i \in M_3, \\
& x_j \geq 0, & j \in N_1, \\
& x_j \leq 0, & j \in N_2, \\
& x_j \text{ free}, & j \in N_3.
\end{array}
$$

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

Can **remove** constraints! Drastic, and could end up with a bound of $-\infty$!

Deriving the Dual Problem

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

$$
\begin{array}{rlll}
(\mathscr{P}) \text{ minimize}_x & c^\intercal x & & \\
(p_i \rightarrow) & a_i^\intercal x \geq b_i, & i \in M_1, \\
(p_i \rightarrow) & a_i^\intercal x \leq b_i, & i \in M_2, \\
(p_i \rightarrow) & a_i^\intercal x = b_i, & i \in M_3, \\
 & x_j \geq 0, & j \in N_1, \\
 & x_j \leq 0, & j \in N_2, \\
 & x_j \text{ free}, & j \in N_3.
\end{array}
$$

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

Can **remove** constraints! Drastic, and could end up with a bound of $-\infty$!

Let's **relax** some constraints!

Deriving the Dual Problem

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

$$
\begin{array}{rlll}
(\mathscr{P}) \text{ minimize}_x & c^\mathsf{T} x & & \\
(p_i \to) & a_i^\mathsf{T} x \geq b_i, & i \in M_1, \\
(p_i \to) & a_i^\mathsf{T} x \leq b_i, & i \in M_2, \\
(p_i \to) & a_i^\mathsf{T} x = b_i, & i \in M_3, \\
& x_j \geq 0, & j \in N_1, \\
& x_j \leq 0, & j \in N_2, \\
& x_j \text{ free}, & j \in N_3.
\end{array}
$$

For every constraint i, have a **price** or **penalty** p_i that penalizes violations

Deriving the Dual Problem

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

$$
\begin{array}{rlll}
(\mathscr{P}) \text{ minimize}_x & c^\intercal x & & \\
(p_i \rightarrow) & a_i^\intercal x \geq b_i, & i \in M_1, \\
(p_i \rightarrow) & a_i^\intercal x \leq b_i, & i \in M_2, \\
(p_i \rightarrow) & a_i^\intercal x = b_i, & i \in M_3, \\
 & x_j \geq 0, & j \in N_1, \\
 & x_j \leq 0, & j \in N_2, \\
 & x_j \text{ free}, & j \in N_3.
\end{array}
$$

For every constraint i, have a **price** or **penalty** p_i that penalizes violations

Add penalized terms in the objective to formulate the **Lagrangean**:

$$
\mathcal{L}(x, p) = c^\intercal x - \sum_{i \in M_1 \cup M_2 \cup M_3} p_i{}^\intercal (a_i^\intercal x - b_i) = p^\intercal b + (c^\intercal - p^\intercal A) x.
$$

Deriving the Dual Problem

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

$$
\begin{array}{llll}
(\mathscr{P}) & \text{minimize}_x & c^\intercal x & \\
(p_i \rightarrow) & & a_i^\intercal x \geq b_i, & i \in M_1, \\
(p_i \rightarrow) & & a_i^\intercal x \leq b_i, & i \in M_2, \\
(p_i \rightarrow) & & a_i^\intercal x = b_i, & i \in M_3, \\
& & x_j \geq 0, & j \in N_1, \\
& & x_j \leq 0, & j \in N_2, \\
& & x_j \text{ free}, & j \in N_3.
\end{array}
$$

For every constraint i, have a **price** or **penalty** p_i that penalizes violations

Add penalized terms in the objective to formulate the **Lagrangean**:

$$\mathcal{L}(x, p) = c^\intercal x - \sum_{i \in M_1 \cup M_2 \cup M_3} p_i{}^\intercal (a_i^\intercal x - b_i) = p^\intercal b + (c^\intercal - p^\intercal A)x.$$

We want this to be **a valid lower bound**: $\mathcal{L}(x, p) \leq c^\intercal x, \forall x \in P$. *Is it?*

Deriving the Dual Problem

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

$$
\begin{array}{rlll}
(\mathscr{P})\ \text{minimize}_x & c^\intercal x & \\
(p_i \rightarrow) & a_i^\intercal x \geq b_i, & i \in M_1, \\
(p_i \rightarrow) & a_i^\intercal x \leq b_i, & i \in M_2, \\
(p_i \rightarrow) & a_i^\intercal x = b_i, & i \in M_3, \\
 & x_j \geq 0, & j \in N_1, \\
 & x_j \leq 0, & j \in N_2, \\
 & x_j \ \text{free}, & j \in N_3.
\end{array}
$$

We want **a valid lower bound**:

$$\mathcal{L}(x, p) := c^\intercal x - \sum_i p_i{}^\intercal (a_i^\intercal x - b_i) = p^\intercal b + (c^\intercal - p^\intercal A)x \leq c^\intercal x,\ \forall x \in P.$$

Deriving the Dual Problem

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

$$
\begin{array}{rlll}
(\mathscr{P}) \text{ minimize}_x & c^\intercal x & & \\
(p_i \rightarrow) & a_i^\intercal x \geq b_i, & i \in M_1, \\
(p_i \rightarrow) & a_i^\intercal x \leq b_i, & i \in M_2, \\
(p_i \rightarrow) & a_i^\intercal x = b_i, & i \in M_3, \\
& x_j \geq 0, & j \in N_1, \\
& x_j \leq 0, & j \in N_2, \\
& x_j \text{ free}, & j \in N_3.
\end{array}
$$

We want **a valid lower bound**:

$$\mathcal{L}(x,p) := c^\intercal x - \sum_i p_i{}^\intercal (a_i^\intercal x - b_i) = p^\intercal b + (c^\intercal - p^\intercal A)x \leq c^\intercal x, \ \forall x \in P.$$

We must impose constraints on p: need $\text{sign}(p_i) = \text{sign}(a_i^\intercal x - b_i)$

Deriving the Dual Problem

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

$$
\begin{array}{rlll}
(\mathscr{P}) \ \text{minimize}_x & c^\intercal x & \\
(p_i \rightarrow) & a_i^\intercal x \geq b_i, & i \in M_1, \\
(p_i \rightarrow) & a_i^\intercal x \leq b_i, & i \in M_2, \\
(p_i \rightarrow) & a_i^\intercal x = b_i, & i \in M_3, \\
& x_j \geq 0, & j \in N_1, \\
& x_j \leq 0, & j \in N_2, \\
& x_j \ \text{free}, & j \in N_3.
\end{array}
$$

We want **a valid lower bound**:

$$\mathcal{L}(x, p) := c^\intercal x - \sum_i p_i{}^\intercal (a_i^\intercal x - b_i) = p^\intercal b + (c^\intercal - p^\intercal A)x \leq c^\intercal x, \ \forall x \in P.$$

We must impose constraints on p: need $\text{sign}(p_i) = \text{sign}(a_i^\intercal x - b_i)$

$$
\begin{array}{ll}
p_i \geq 0, & \forall\, i \in M_1 \\
p_i \leq 0, & \forall\, i \in M_2 \\
p_i \ \text{free}, & \forall\, i \in M_3.
\end{array}
\tag{2}
$$

Deriving the Dual Problem

Summarizing... with p satisfying (2), we have **a valid lower bound**:

$$\mathcal{L}(x, p) := c^\intercal x - \sum_i p_i{}^\intercal (a_i^\intercal x - b_i) = p^\intercal b + (c^\intercal - p^\intercal A)x \leq c^\intercal x, \ \forall x \in P.$$

How can we get a lower bound on the **optimal value** $c^\intercal x^*$ *of* $(\mathscr{P})$*?*

Deriving the Dual Problem

Summarizing... with p satisfying (2), we have **a valid lower bound**:

$$\mathcal{L}(x,p) := c^\intercal x - \sum_i p_i{}^\intercal (a_i^\intercal x - b_i) = p^\intercal b + (c^\intercal - p^\intercal A)x \leq c^\intercal x, \ \forall x \in P.$$

How can we get a lower bound on the **optimal value** $c^\intercal x^*$ *of* $(\mathscr{P})$*?*

For any p satisfying (2), let

$$\begin{aligned} g(p) := \min_x \ & \left[p^\intercal b + (c^\intercal - p^\intercal A)x\right] \\ \text{s.t. } & x_j \geq 0, \ j \in N_1, \\ & x_j \leq 0, \ j \in N_2, \\ & x_j \text{ free}, \ j \in N_3. \end{aligned} \tag{3}$$

Then, we have $g(p) \leq c^\intercal x^*$.

Can we simplify this $g(p)$ *further?*

Deriving the Dual Problem

For p satisfying (2), the value:

$$\begin{aligned} g(p) := \text{min}_x \ & \left[p^\intercal b + (c^\intercal - p^\intercal A)x\right] \\ \text{s.t. } & x_j \geq 0, \ \ j \in N_1, \\ & x_j \leq 0, \ \ j \in N_2, \\ & x_j \text{ free}, \ \ j \in N_3 \end{aligned}$$

is **a valid lower bound** on the **optimal value** of $(\mathscr{P})$: $g(p) \leq c^\intercal x^*$.

Can we simplify this $g(p)$ further?

Deriving the Dual Problem

For p satisfying (2), the value:

$$\begin{aligned} g(p) := \min_x \ & \left[p^\intercal b + (c^\intercal - p^\intercal A)x\right] \\ \text{s.t. } & x_j \geq 0, \ \ j \in N_1, \\ & x_j \leq 0, \ \ j \in N_2, \\ & x_j \text{ free}, \ \ j \in N_3 \end{aligned}$$

is **a valid lower bound** on the **optimal value** of $(\mathscr{P})$: $g(p) \leq c^\intercal x^*$.

Can we simplify this $g(p)$ further?

$$g(p) = \begin{cases} p^\intercal b, & \text{if } c_j - p^\intercal A_j \geq 0, \forall\, j \in N_1 \textbf{ and} \\ & c_j - p^\intercal A_j \leq 0, \forall\, j \in N_2 \textbf{ and} \\ & c_j - p^\intercal A_j = 0, \forall\, j \in N_3 \\ -\infty, & \text{otherwise.} \end{cases}$$

Deriving the Dual Problem

For p satisfying (2), the value:

$$g(p) = \begin{cases} p^\intercal b, & \text{if } c_j - p^\intercal A_j \geq 0, \forall\, j \in N_1 \textbf{ and} \\ & c_j - p^\intercal A_j \leq 0, \forall\, j \in N_2 \textbf{ and} \\ & c_j - p^\intercal A_j = 0, \forall\, j \in N_3 \\ -\infty, & \text{otherwise.} \end{cases}$$

is **a valid lower bound** on the **optimal value** $(\mathscr{P})$: $g(p) \leq c^\intercal x^*$.

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

Deriving the Dual Problem

For p satisfying (2), the value:

$$g(p) = \begin{cases} p^\intercal b, & \text{if } c_j - p^\intercal A_j \geq 0, \forall\, j \in N_1 \textbf{ and} \\ & c_j - p^\intercal A_j \leq 0, \forall\, j \in N_2 \textbf{ and} \\ & c_j - p^\intercal A_j = 0, \forall\, j \in N_3 \\ -\infty, & \text{otherwise.} \end{cases}$$

is **a valid lower bound** on the **optimal value** $(\mathscr{P})$: $g(p) \leq c^\intercal x^*$.

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

$$\text{maximize}_p \{g(p) \; : \; p \text{ satisfying (2)}\}. \tag{4}$$

Deriving the Dual Problem

For p satisfying (2), the value:

$$g(p) = \begin{cases} p^\intercal b, & \text{if } c_j - p^\intercal A_j \geq 0, \forall\, j \in N_1 \textbf{ and} \\ & c_j - p^\intercal A_j \leq 0, \forall\, j \in N_2 \textbf{ and} \\ & c_j - p^\intercal A_j = 0, \forall\, j \in N_3 \\ -\infty, & \text{otherwise.} \end{cases}$$

is **a valid lower bound** on the **optimal value** $(\mathscr{P})$: $g(p) \leq c^\intercal x^*$.

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

$$\text{maximize}_p \, \{g(p) \; : \; p \text{ satisfying (2)}\}. \tag{4}$$

- Because we maximize $g(p)$, we can restrict attention to p so $g(p) > -\infty$...

Deriving the Dual Problem

For p satisfying (2), the value:

$$g(p) = \begin{cases} p^\intercal b, & \text{if } c_j - p^\intercal A_j \geq 0, \forall\, j \in N_1 \textbf{ and} \\ & c_j - p^\intercal A_j \leq 0, \forall\, j \in N_2 \textbf{ and} \\ & c_j - p^\intercal A_j = 0, \forall\, j \in N_3 \\ -\infty, & \text{otherwise.} \end{cases}$$

is **a valid lower bound** on the **optimal value** $(\mathscr{P})$: $g(p) \leq c^\intercal x^*$.

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

$$\text{maximize}_p \{g(p) \;:\; p \text{ satisfying (2)}\}. \tag{4}$$

- Because we maximize $g(p)$, we can restrict attention to p so $g(p) > -\infty$...
- Recall that (2) requires:

$$\begin{aligned} p_i &\geq 0, & \forall\, i \in M_1 \\ p_i &\leq 0, & \forall\, i \in M_2 \\ p_i &\text{ free}, & \forall\, i \in M_3. \end{aligned}$$

Deriving the Dual Problem

The **best lower bound** on the optimal value of $(\mathscr{P})$ is given by:

$$
\begin{array}{rlll}
\text{maximize} & p^\intercal b & & \\
\text{subject to} & p_i \geq 0, & i \in M_1, & \\
& p_i \leq 0, & i \in M_2, & \\
& p_i \ \text{free}, & i \in M_3, & \quad (5) \\
& p^\intercal A_j \leq c_j, & j \in N_1, & \\
& p^\intercal A_j \geq c_j, & j \in N_2, & \\
& p^\intercal A_j = c_j, & j \in N_3. &
\end{array}
$$

Deriving the Dual Problem

The **best lower bound** on the optimal value of $(\mathscr{P})$ is given by:

$$
\begin{array}{rlll}
\text{maximize} & p^\intercal b & & \\
\text{subject to} & p_i \geq 0, & i \in M_1, & \\
& p_i \leq 0, & i \in M_2, & \\
& p_i \text{ free}, & i \in M_3, & \\
& p^\intercal A_j \leq c_j, & j \in N_1, & \\
& p^\intercal A_j \geq c_j, & j \in N_2, & \\
& p^\intercal A_j = c_j, & j \in N_3. &
\end{array} \tag{5}
$$

This is the **dual** of $(\mathscr{P})$, which we will also refer to as $(\mathscr{D})$.

Summarizing

We obtained the following primal-dual pair of problems:

Primal $(\mathscr{P})$

$$
\begin{array}{rlll}
\text{minimize}_x & c^\intercal x & \\
(p_i \rightarrow) & a_i^\intercal x \geq b_i, & i \in M_1, \\
(p_i \rightarrow) & a_i^\intercal x \leq b_i, & i \in M_2, \\
(p_i \rightarrow) & a_i^\intercal x = b_i, & i \in M_3, \\
& x_j \geq 0, & j \in N_1, \\
& x_j \leq 0, & j \in N_2, \\
& x_j \text{ free}, & j \in N_3.
\end{array}
$$

Dual $(\mathscr{D})$

$$
\begin{array}{rlll}
\text{maximize}_p & p^\intercal b & \\
& p_i \geq 0, & i \in M_1, \\
& p_i \leq 0, & i \in M_2, \\
& p_i \text{ free}, & i \in M_3, \\
(x_j \rightarrow) & p^\intercal A_j \leq c_j, & j \in N_1, \\
(x_j \rightarrow) & p^\intercal A_j \geq c_j, & j \in N_2, \\
(x_j \rightarrow) & p^\intercal A_j = c_j, & j \in N_3.
\end{array}
$$

Summarizing

We obtained the following primal-dual pair of problems:

Primal $(\mathscr{P})$

$$
\begin{array}{rlr}
\text{minimize}_x & c^\intercal x & \\
(p_i \rightarrow) & a_i^\intercal x \geq b_i, & i \in M_1, \\
(p_i \rightarrow) & a_i^\intercal x \leq b_i, & i \in M_2, \\
(p_i \rightarrow) & a_i^\intercal x = b_i, & i \in M_3, \\
 & x_j \geq 0, & j \in N_1, \\
 & x_j \leq 0, & j \in N_2, \\
 & x_j \text{ free}, & j \in N_3.
\end{array}
$$

Dual $(\mathscr{D})$

$$
\begin{array}{rlr}
\text{maximize}_p & p^\intercal b & \\
 & p_i \geq 0, & i \in M_1, \\
 & p_i \leq 0, & i \in M_2, \\
 & p_i \text{ free}, & i \in M_3, \\
(x_j \rightarrow) & p^\intercal A_j \leq c_j, & j \in N_1, \\
(x_j \rightarrow) & p^\intercal A_j \geq c_j, & j \in N_2, \\
(x_j \rightarrow) & p^\intercal A_j = c_j, & j \in N_3.
\end{array}
$$

Simple rules to help you derive duals quickly:

- a dual decision variable for every primal constraint (except variables signs)
  - if "$=$" constraint, dual variable is free
  - if ("$\geq$", minimize) or ("$\leq$", maximize), dual variable ≥ 0
  - if ("$\geq$", maximize) or ("$\leq$", minimize), dual variable ≤ 0
- for every decision variable in the primal, there is a constraint in the dual
  - signs for the constraint derived by reversing the above

Example 1

$$
\begin{aligned}
\min\ & x_1 + 2x_2 + 3x_4 \\
& -x_1 + 3x_2 = 5 \\
& 2x_1 - x_2 + 3x_3 \geq 6 \\
& x_3 \leq 4 \\
& x_1 \geq 0 \\
& x_2 \leq 0 \\
& x_3 \text{ free}
\end{aligned}
$$

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 $a_i^\intercal$ 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.*

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 $a_i^\intercal$ 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.*

Theorem (The dual of the dual is the primal)

If we transform the dual into an equivalent minimization problem and then form its dual, we obtain a problem equivalent to the original primal problem.

Weak duality

Primal $(\mathscr{P})$

minimize_x	$c^\mathsf{T} x$	
$(p_i \to)$	$a_i^\mathsf{T} x \geq b_i,$	$i \in M_1,$
$(p_i \to)$	$a_i^\mathsf{T} x \leq b_i,$	$i \in M_2,$
$(p_i \to)$	$a_i^\mathsf{T} x = b_i,$	$i \in M_3,$
	$x_j \geq 0,$	$j \in N_1,$
	$x_j \leq 0,$	$j \in N_2,$
	x_j free,	$j \in N_3.$

Dual $(\mathscr{D})$

maximize_p	$p^\mathsf{T} b$	
	$p_i \geq 0,$	$i \in M_1,$
	$p_i \leq 0,$	$i \in M_2,$
	p_i free,	$i \in M_3,$
$(x_j \to)$	$p^\mathsf{T} A_j \leq c_j,$	$j \in N_1,$
$(x_j \to)$	$p^\mathsf{T} A_j \geq c_j,$	$j \in N_2,$
$(x_j \to)$	$p^\mathsf{T} A_j = c_j,$	$j \in N_3.$

Weak duality

Primal $(\mathscr{P})$				**Dual** $(\mathscr{D})$		
minimize$_x$	$c^\top x$			maximize$_p$	$p^\top b$	
$(p_i \to)$	$a_i^\top x \geq b_i,$	$i \in M_1,$			$p_i \geq 0,$	$i \in M_1,$
$(p_i \to)$	$a_i^\top x \leq b_i,$	$i \in M_2,$			$p_i \leq 0,$	$i \in M_2,$
$(p_i \to)$	$a_i^\top x = b_i,$	$i \in M_3,$			p_i free,	$i \in M_3,$
	$x_j \geq 0,$	$j \in N_1,$		$(x_j \to)$	$p^\top A_j \leq c_j,$	$j \in N_1,$
	$x_j \leq 0,$	$j \in N_2,$		$(x_j \to)$	$p^\top A_j \geq c_j,$	$j \in N_2,$
	x_j free,	$j \in N_3.$		$(x_j \to)$	$p^\top A_j = c_j,$	$j \in N_3.$

Theorem (Weak duality)

If x is feasible for $(\mathscr{P})$ and p is feasible for $(\mathscr{D})$, then $p^\top b \leq c^\top x$.

Proof.

Weak duality

Primal $(\mathscr{P})$			**Dual** $(\mathscr{D})$		
minimize_x	$c^\intercal x$		maximize_p	$p^\intercal b$	
$(p_i \rightarrow)$	$a_i^\intercal x \geq b_i,$	$i \in M_1,$		$p_i \geq 0,$	$i \in M_1,$
$(p_i \rightarrow)$	$a_i^\intercal x \leq b_i,$	$i \in M_2,$		$p_i \leq 0,$	$i \in M_2,$
$(p_i \rightarrow)$	$a_i^\intercal x = b_i,$	$i \in M_3,$		p_i free,	$i \in M_3,$
	$x_j \geq 0,$	$j \in N_1,$	$(x_j \rightarrow)$	$p^\intercal A_j \leq c_j,$	$j \in N_1,$
	$x_j \leq 0,$	$j \in N_2,$	$(x_j \rightarrow)$	$p^\intercal A_j \geq c_j,$	$j \in N_2,$
	x_j free,	$j \in N_3.$	$(x_j \rightarrow)$	$p^\intercal A_j = c_j,$	$j \in N_3.$

Theorem (Weak duality)

If x is feasible for $(\mathscr{P})$ and p is feasible for $(\mathscr{D})$, then $p^\intercal b \leq c^\intercal x$.

Proof.

By construction, the (optimal) dual objective provides a lower bound on the (optimal) primal objective...

Implications of Weak Duality

Corollary

The following results hold:

Implications of Weak Duality

Corollary

The following results hold:

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

Implications of Weak Duality

Corollary

The following results hold:

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

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

Implications of Weak Duality

Corollary

The following results hold:

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

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

(c) If x *feasible for* $(\mathscr{P})$ *and* p *feasible for* $(\mathscr{D})$*, then:*

$$c^{\intercal}x - c^{\intercal}x^{*} \leq c^{\intercal}x - p^{\intercal}b \quad \textbf{and} \quad (p^{*})^{\intercal}b - p^{\intercal}b \leq c^{\intercal}x - p^{\intercal}b.$$

Implications of Weak Duality

Corollary

The following results hold:

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

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

(c) If x *feasible for* $(\mathscr{P})$ *and* p *feasible for* $(\mathscr{D})$*, then:*

$$c^\intercal x - c^\intercal x^* \leq c^\intercal x - p^\intercal b \quad \textbf{and} \quad (p^*)^\intercal b - p^\intercal b \leq c^\intercal x - p^\intercal b.$$

(d) Under the premises in (c), if $p^\intercal b = c^\intercal x$ *holds, then* x *and* p *are* **optimal** *solutions to* $(\mathscr{P})$ *and* $(\mathscr{D})$*, respectively.*

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

Implications of Weak Duality

Corollary

The following results hold:

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

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

(c) If x *feasible for* $(\mathscr{P})$ *and* p *feasible for* $(\mathscr{D})$*, then:*

$$c^\intercal x - c^\intercal x^* \leq c^\intercal x - p^\intercal b \quad \textbf{and} \quad (p^*)^\intercal b - p^\intercal b \leq c^\intercal x - p^\intercal b.$$

(d) Under the premises in (c), if $p^\intercal b = c^\intercal x$ *holds, then* x *and* p *are* **optimal** *solutions to* $(\mathscr{P})$ *and* $(\mathscr{D})$*, respectively.*

(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 p satisfying (d) even exist?

Strong duality

Theorem (Strong duality)

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

Strong duality

Theorem (Strong duality)

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

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 $x_n \to x^*$, then $x^* \in S$.

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 $x_n \to x^*$, then $x^* \in S$.

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 $x_n \to x^*$, then $x^* \in S$.

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 $x_n \to x^*$.
- For each k, we have $x_k \in P$, and therefore, $Ax_k \geq b$.
- Then, $Ax^* = A\left(\lim_{k\to\infty} x_k\right) = \lim_{k\to\infty} Ax_k \geq b$, so x^* 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 $x_n \to x^*$, then $x^* \in S$.

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 $x_n \to x^*$, then $x^* \in S$.

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 exists some $\underline{x} \in S$ such that $f(\underline{x}) \leq f(x)$ for all $x \in S$ and there exists some $\bar{x} \in S$ such that $f(\bar{x}) \geq f(x)$ for all $x \in S$.

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

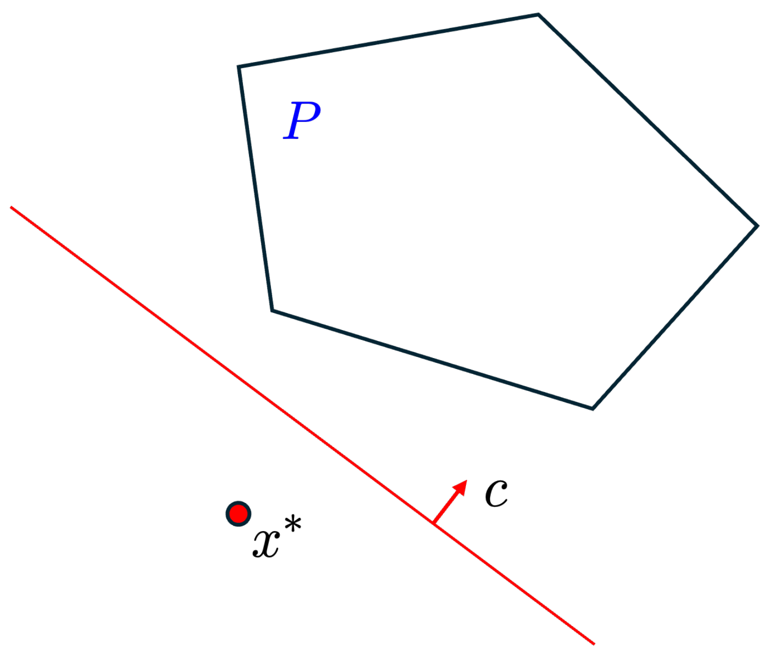
Separating Hyperplane Theorem

The first **fundamental result in optimization**

Separating Hyperplane Theorem

Theorem (**Simple** Separating Hyperplane Theorem)

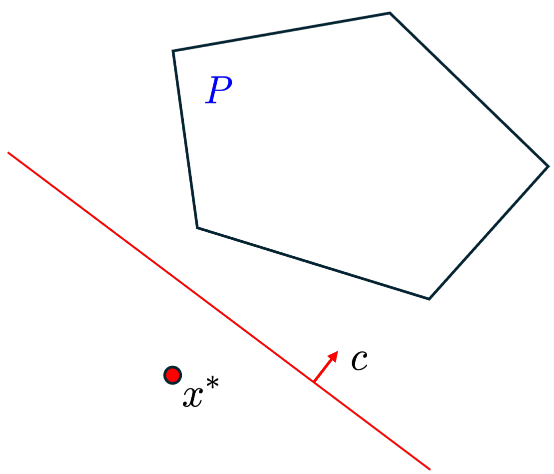
Consider a point x^* and a polyhedron P. If $x^* \notin P$, the there exists a vector $c \in \mathbb{R}^n$ such that $c \neq 0$ and $c^\intercal x^* < c^\intercal y$ holds for all $y \in P$.

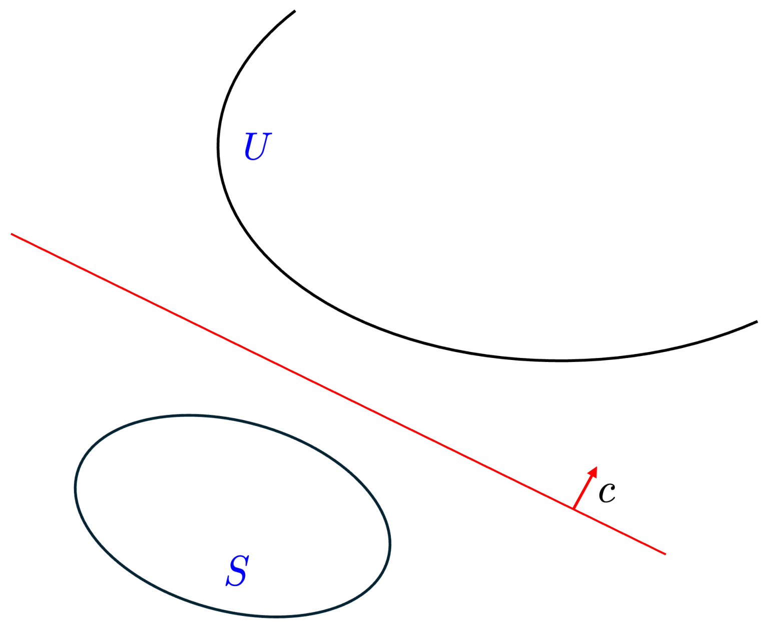


Separating Hyperplane Theorem

Theorem (Separating Hyperplane Theorem for Convex Sets)

Let S and U be two nonempty, closed, convex subsets of $\mathbb{R}^n$ such that S is bounded. Then, there exists a vector $c \in \mathbb{R}^n$ such that $c \neq 0$ and $c^\intercal x < c^\intercal y$ holds for all $x \in S$ and $y \in U$.





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 a vector $c \in \mathbb{R}^n$ such that $c^\intercal x < c^\intercal y$ holds for all $x \in S$ and $y \in U$.

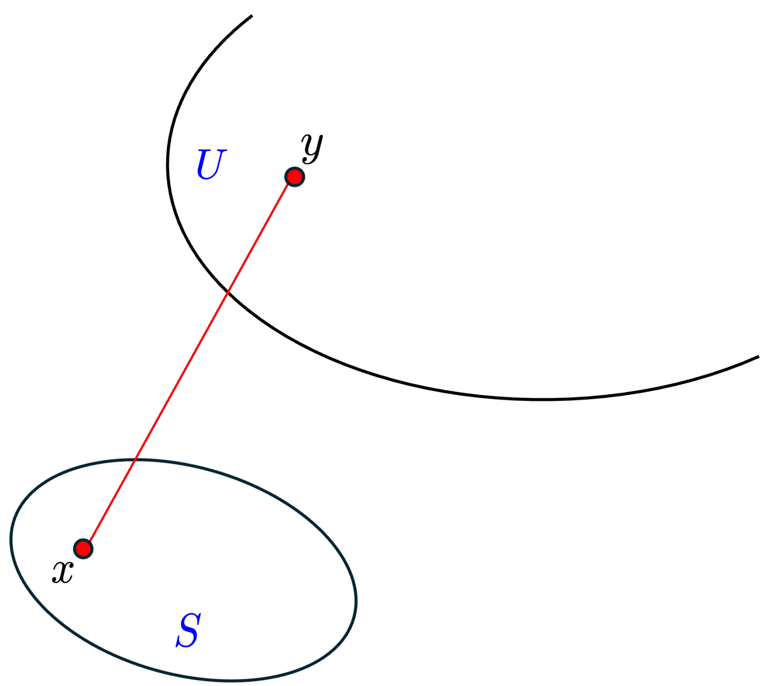
Proof.

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 a vector $c \in \mathbb{R}^n$ such that $c^\intercal x < c^\intercal y$ holds for all $x \in S$ and $y \in U$.

Proof.

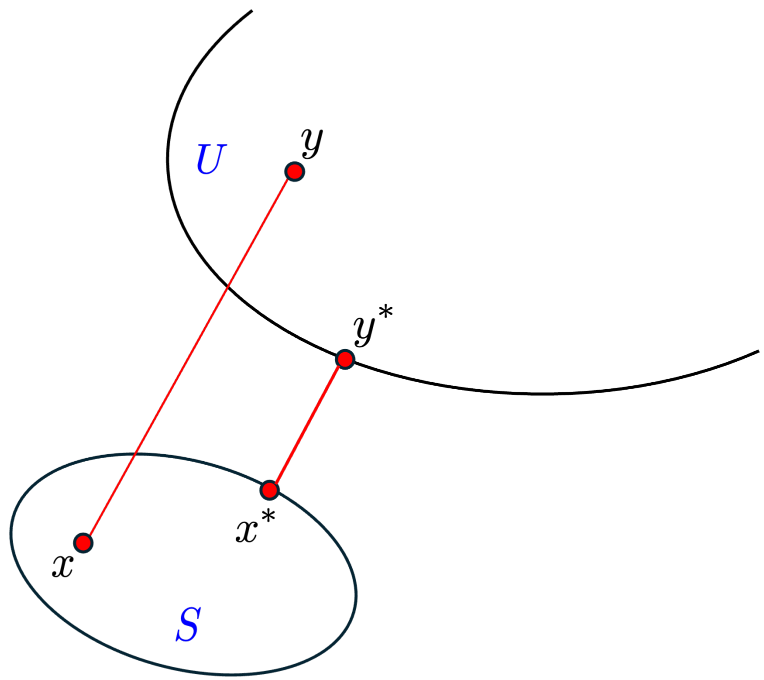


Separating Hyperplane Theorem

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 a vector $c \in \mathbb{R}^n$ such that $c^\intercal x < c^\intercal y$ holds for all $x \in S$ and $y \in U$.

Proof.

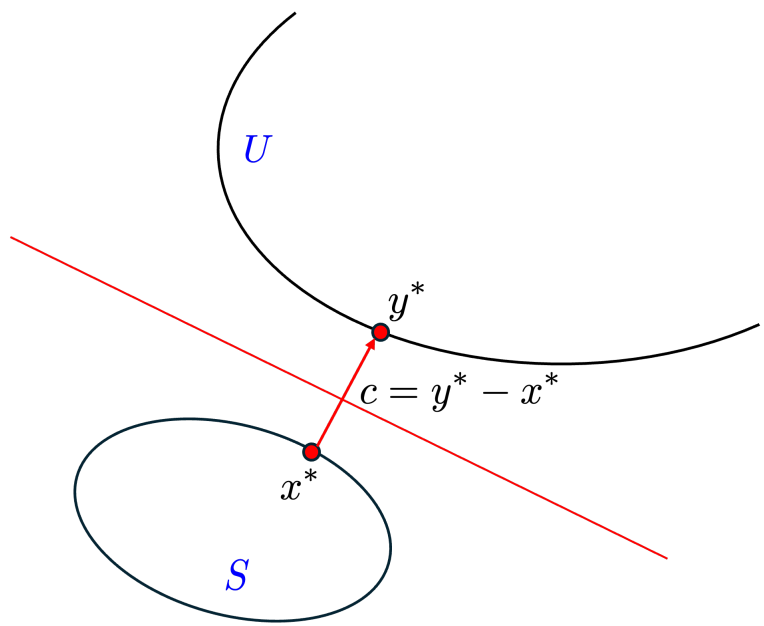


Separating Hyperplane Theorem

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 a vector $c \in \mathbb{R}^n$ such that $c^\intercal x < c^\intercal y$ holds for all $x \in S$ and $y \in U$.

Proof.

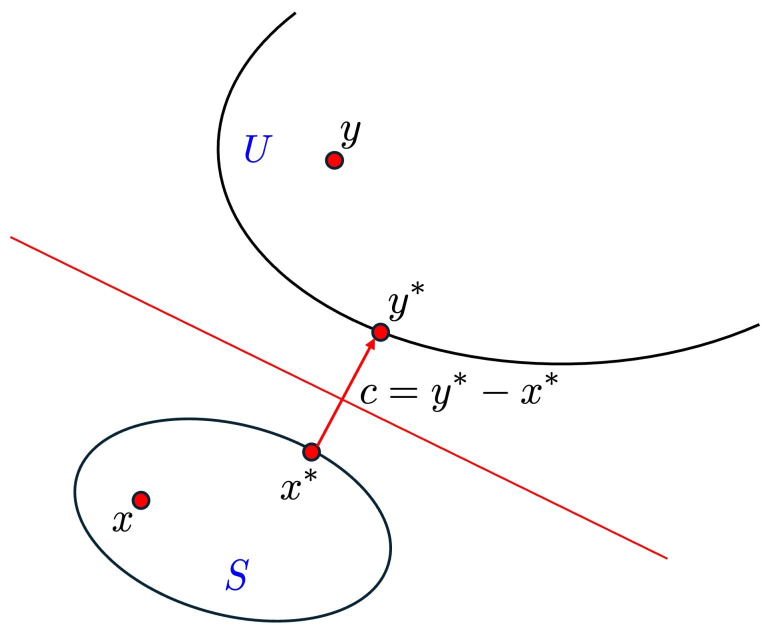


Separating Hyperplane Theorem

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 a vector $c \in \mathbb{R}^n$ such that $c^\intercal x < c^\intercal y$ holds for all $x \in S$ and $y \in U$.

Proof.



Separating Hyperplane Theorem - Done!

We proved the first **fundamental result in optimization**! The Separating Hyperplane Theorem for **convex sets** will be very useful later!

Separating Hyperplane Theorem - Done!

We proved the first **fundamental result in optimization**! The Separating Hyperplane Theorem for **convex sets** will be very useful later!

Corollary (Needed for our purposes...)

If P is a polyhedron and x^* satisfies $x \notin P$, there exists a hyperplane that strictly separates x from P, i.e., $\exists c \neq 0$ such that $c^\intercal x^* < c^\intercal x \, \forall x \in P$.