Duality - Continued

Lecture 5

October 7, 2024

Recap From Last Time

We obtained the following primal-dual pair of problems:

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

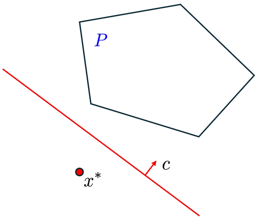
Simple rules to help you derive duals quickly:

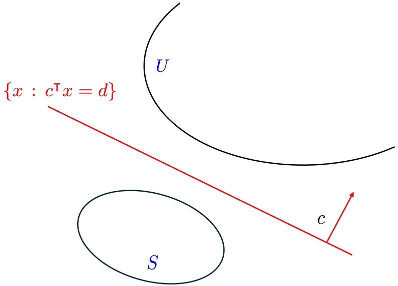
- a dual decision variable for every primal constraint (except variables signs)
  - if "$=$" constraint, dual variable is free
  - if ("$\geq$", minimize) or ("$\leq$", maximize), dual variable ≥ 0
  - if ("$\geq$", maximize) or ("$\leq$", minimize), dual variable ≤ 0
- for every decision variable in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

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 $c \in \mathbb{R}^n$ and $d \in \mathbb{R}$ such that $S \subset \{x \in \mathbb{R}^n : c^\intercal x < d\}$ and $U \subset \{x \in \mathbb{R}^n : c^\intercal x > d\}$.





Separating Hyperplane Theorem - Caveats!

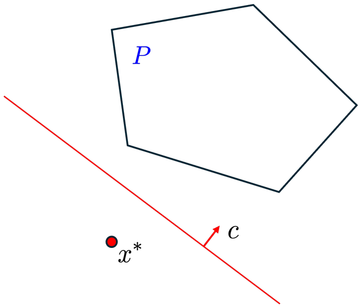
Both conditions in the theorem needed: closed and at least one bounded

Needed For Our Purposes

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

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 $c^\intercal x^* < c^\intercal x \, \forall x \in P$.



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

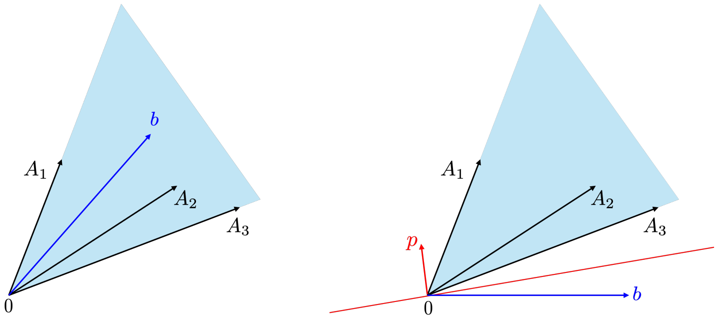
Theorem (Farkas' Lemma)

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 p such that $p^T A \geq 0$ and $p^T b < 0$.

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 p such that $p^T A \geq 0$ and $p^T b < 0$.



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 p such that $p^T A \geq 0$ and $p^T b < 0$.

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 p such that $p^T A \geq 0$ and $p^T b < 0$.

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

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

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

But then $p^T b = p^T Ax \geq 0$, 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 p such that $p^T A \geq 0$ and $p^T b < 0$.

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

Farkas Lemma

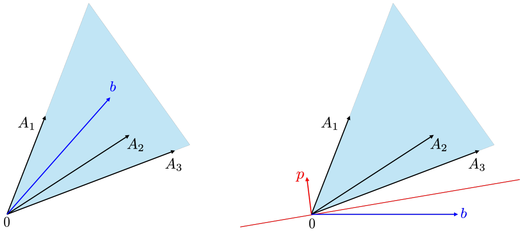
Theorem (Farkas' Lemma)

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 p such that $p^T A \geq 0$ and $p^T b < 0$.

"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 p such that $p^T A \geq 0$ and $p^T b < 0$.

"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 p such that $p^T A \geq 0$ and $p^T b < 0$.

"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 p such that $p^T A \geq 0$ and $p^T b < 0$.

"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 p such that $p^T A \geq 0$ and $p^T b < 0$.

"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 p such that $p^T A \geq 0$ and $p^T b < 0$.

"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 p such that $p^T A \geq 0$ and $p^T b < 0$.

"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 $\exists p : p^\mathsf{T} b < p^\mathsf{T} y, \forall y \in S$.

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 p such that $p^T A \geq 0$ and $p^T b < 0$.

"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 $\exists p \,:\, p^\mathsf{T} b < p^\mathsf{T} y, \forall y \in S$.
- $0 \in S \Rightarrow p^\mathsf{T} b < 0$.

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 p such that $p^T A \geq 0$ and $p^T b < 0$.

"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 $\exists p : p^\intercal b < p^\intercal y, \forall y \in S$.
- $0 \in S \Rightarrow p^\intercal b < 0$.
- Every column A_i of A satisfies $\lambda A_i \in S$ for every $\lambda > 0$, so

$$\frac{p^\intercal b}{\lambda} < p^\intercal A_i,\ \forall \lambda > 0$$

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 p such that $p^T A \geq 0$ and $p^T b < 0$.

"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 $\exists p \,:\, p^\intercal b < p^\intercal y, \forall y \in S$.
- $0 \in S \Rightarrow p^\intercal b < 0$.
- Every column A_i of A satisfies $\lambda A_i \in S$ for every $\lambda > 0$, so

$$\frac{p^\intercal b}{\lambda} < p^\intercal A_i,\ \forall \lambda > 0$$

- Limit $\lambda \to \infty$ implies $p^\intercal A_i \geq 0$. ■

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 p such that $p^T A \geq 0$ and $p^T b < 0$.

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 p such that $p^T A \geq 0$ and $p^T b < 0$.

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

This has some important implications:

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

$$
\begin{array}{rrl}
(\mathscr{P}) & \text{minimize} & c^\intercal x \\
& \text{subject to} & Ax = b \\
& & x \geq 0
\end{array}
$$

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 p such that $p^T A \geq 0$ and $p^T b < 0$.

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

This has some important implications:

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

$$
\begin{array}{lrl}
(\mathscr{P}) & \text{minimize} & c^\intercal x \\
 & \text{subject to} & Ax = b \\
 & & x \geq 0
\end{array}
$$

- Farkas Lemma states that either $(\mathscr{P})$ is feasible or ...

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 p such that $p^T A \geq 0$ and $p^T b < 0$.

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

This has some important implications:

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

$$\begin{array}{lrl} (\mathscr{P}) & \text{minimize} & c^\intercal x \\ & \text{subject to} & Ax = b \\ & & x \geq 0 \end{array}$$

- Farkas Lemma states that either $(\mathscr{P})$ is feasible or ...
  ... there exists p (satisfying $p^\intercal A \leq c^\intercal$) that is a **certificate of infeasibility**!

Strong Duality

(W.L.O.G.) Consider the following primal-dual pair:

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

Strong Duality

(W.L.O.G.) Consider the following primal-dual pair:

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

Theorem (Strong Duality)

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

Strong Duality

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

Proof.

- Assume $(\mathscr{P})$ has optimal solution x^*
- Will prove that $(\mathscr{D})$ admits feasible solution p such that $p^\intercal b = c^\intercal x^*$

Strong Duality

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

Proof.

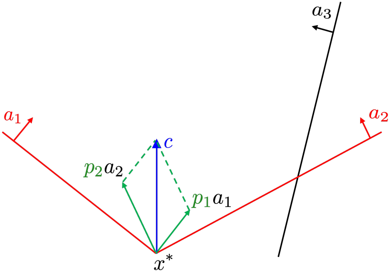
- Assume $(\mathscr{P})$ has optimal solution x^*
- Will prove that $(\mathscr{D})$ admits feasible solution p such that $p^\mathsf{T} b = c^\mathsf{T} x^*$
- Let $\mathcal{F} = \{i \mid a_i^\mathsf{T} x^* = b_i\}$ be indices of active constraints at x^*
- Show that c can be written as conic combination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

Strong Duality

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

Proof.

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



Strong Duality

$$(\mathscr{P}) \quad \text{minimize } c^\mathsf{T} x \qquad\qquad (\mathscr{D}) \quad \text{maximize } p^\mathsf{T} b$$

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

Proof.

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

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

Strong Duality

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

Proof.

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

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

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

Strong Duality

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

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 $x^* + \epsilon d \in P$ for small ϵ
  - $a_i^\mathsf{T}(x^* + \epsilon d) \geq b_i, \forall i \in \mathcal{F}$
  - $a_i^\mathsf{T}(x^* + \epsilon d) \geq b_i, \forall i \notin \mathcal{F}$ holds because $a_i^\mathsf{T} x^* > b_i \, \forall i \notin \mathcal{F}$

Strong Duality

$(\mathscr{P})$ minimize $c^\intercal x$ subject to $Ax \geq b$

$(\mathscr{D})$ maximize $p^\intercal b$ subject to $p^\intercal A = c^T, \quad p \geq 0.$

Proof.

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

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

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

Strong Duality

$(\mathscr{P})$ minimize $c^\intercal x$ $\qquad$ $(\mathscr{D})$ maximize $p^\intercal b$

subject to $Ax \geq b$ $\qquad$ subject to $p^\intercal A = c^T, \quad p \geq 0.$

Proof.

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

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

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

Strong Duality

$$(\mathscr{P}) \quad \text{minimize } c^\intercal x \qquad\qquad (\mathscr{D}) \quad \text{maximize } p^\intercal b$$

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

Proof.

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

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

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

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

Strong Duality

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

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 $x^* + \epsilon d \in P$ for small ϵ
  - $a_i^{\mathsf{T}}(x^* + \epsilon d) \geq b_i, \forall i \in \mathcal{F}$
  - $a_i^{\mathsf{T}}(x^* + \epsilon d) \geq b_i, \forall i \notin \mathcal{F}$ holds because $a_i^{\mathsf{T}} x^* > b_i \, \forall i \notin \mathcal{F}$
- $c^{\mathsf{T}} d \geq 0$ because otherwise $c^{\mathsf{T}}(x^* + \epsilon d) < c^{\mathsf{T}} x^*$ would contradict x^* 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:

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

- Let $p_i = 0$ for $i \notin \mathcal{F} \quad \Rightarrow \quad \exists p$ feasible for $(\mathscr{D})$
- $p^{\mathsf{T}} b = \sum_{i \in \mathcal{F}} p_i b_i = \sum_{i \in \mathcal{F}} p_i a_i^{\mathsf{T}} x^* = c^{\mathsf{T}} x^*$ $\blacksquare$

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	?

Example

Is this primal feasible? What is its dual?

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

Example

Is this primal feasible? What is its dual?

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

Application in Robust Optimization

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

$$a^\intercal x \leq b, \tag{1}$$

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

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

$$a^\intercal x \leq b, \; \forall \, a \in \mathcal{A} \tag{2}$$

  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:

$$a^\intercal x \leq b, \tag{1}$$

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

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

$$a^\intercal x \leq b, \; \forall \, a \in \mathcal{A} \tag{2}$$

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

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:

$$\sum_{i=1}^{k} v_{[i]} \geq b, \tag{3}$$

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:

$$\sum_{i=1}^{k} v_{[i]} \geq b, \tag{3}$$

  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?*

Optimality for Standard-Form LPs

$$(\mathscr{P}) \ \min c^\mathsf{T} x \qquad Ax = b, \quad x \geq 0$$

$$(\mathscr{D}) \ \max p^\mathsf{T} b \qquad p^\mathsf{T} A \leq c^\mathsf{T}$$

Optimality for Standard-Form LPs

$$(\mathscr{P})\ \min c^\mathsf{T} x \qquad\qquad (\mathscr{D})\ \max p^\mathsf{T} b$$

$$Ax = b, \quad x \geq 0 \qquad\qquad p^\mathsf{T} A \leq c^\mathsf{T}$$

- $(\mathscr{P})$ achieves optimality at a **basic feasible solution** x:
  - If $B \subseteq \{1, \ldots, n\}$ is a basis, the b.f.s. is: $x = [x_B, 0], \quad x_B = A_B^{-1} b$.
  - Simplex algorithm: feasibility and optimality for $(\mathscr{P})$ are given by:

Optimality for Standard-Form LPs

$$(\mathscr{P}) \ \min c^\intercal x \qquad\qquad (\mathscr{D}) \ \max p^\intercal b$$

$$Ax = b, \quad x \geq 0 \qquad\qquad p^\intercal A \leq c^\intercal$$

- $(\mathscr{P})$ achieves optimality at a **basic feasible solution** x:
  - If $B \subseteq \{1, \ldots, n\}$ is a basis, the b.f.s. is: $x = [x_B, 0], \ \ x_B = A_B^{-1} b$.
  - Simplex algorithm: feasibility and optimality for $(\mathscr{P})$ are given by:

$$\text{Feasibility-}(\mathscr{P}): \qquad x_B := A_B^{-1} b \geq 0 \tag{4a}$$

$$\text{Optimality-}(\mathscr{P}): \qquad c^\intercal - c_B^\intercal A_B^{-1} A \geq 0 \tag{4b}$$

Optimality for Standard-Form LPs

$$(\mathscr{P}) \ \min c^\intercal x \qquad\qquad (\mathscr{D}) \ \max p^\intercal b$$

$$Ax = b, \quad x \geq 0 \qquad\qquad p^\intercal A \leq c^\intercal$$

- $(\mathscr{P})$ achieves optimality at a **basic feasible solution** x:
  - If $B \subseteq \{1, \ldots, n\}$ is a basis, the b.f.s. is: $x = [x_B, 0], \quad x_B = A_B^{-1} b$.
  - Simplex algorithm: feasibility and optimality for $(\mathscr{P})$ are given by:

$$\text{Feasibility-}(\mathscr{P}): \qquad x_B := A_B^{-1} b \geq 0 \tag{4a}$$

$$\text{Optimality-}(\mathscr{P}): \qquad c^\intercal - c_B^\intercal A_B^{-1} A \geq 0 \tag{4b}$$

- $(\mathscr{D})$: same basis B can also be used to determine **a dual vector** p:

$$p^\intercal A_i = c_i, \forall\, i \in B \quad \Rightarrow \quad p^\intercal = c_B^\intercal A_B^{-1}, \forall\, i \in B.$$

  - The dual objective value of p is exactly:

Optimality for Standard-Form LPs

$$(\mathscr{P}) \;\; \min\, c^\intercal x \qquad\qquad (\mathscr{D}) \;\; \max\, p^\intercal b$$

$$Ax = b, \quad x \geq 0 \qquad\qquad p^\intercal A \leq c^\intercal$$

- $(\mathscr{P})$ achieves optimality at a **basic feasible solution** x:
  - If $B \subseteq \{1, \ldots, n\}$ is a basis, the b.f.s. is: $x = [x_B, 0], \;\; x_B = A_B^{-1} b$.
  - Simplex algorithm: feasibility and optimality for $(\mathscr{P})$ are given by:

$$\text{Feasibility-}(\mathscr{P}): \qquad x_B := A_B^{-1} b \geq 0 \tag{4a}$$

$$\text{Optimality-}(\mathscr{P}): \qquad c^\intercal - c_B^\intercal A_B^{-1} A \geq 0 \tag{4b}$$

- $(\mathscr{D})$: same basis B can also be used to determine **a dual vector** p:

$$p^\intercal A_i = c_i, \forall\, i \in B \quad \Rightarrow \quad p^\intercal = c_B^\intercal A_B^{-1}, \forall\, i \in B.$$

  - The dual objective value of p is exactly: $\;p^\intercal b = c_B^\intercal A_B^{-1} b = c^\intercal x$

Optimality for Standard-Form LPs

$$(\mathscr{P}) \ \min c^\mathsf{T} x \qquad\qquad (\mathscr{D}) \ \max p^\mathsf{T} b$$

$$Ax = b, \quad x \geq 0 \qquad\qquad p^\mathsf{T} A \leq c^\mathsf{T}$$

- $(\mathscr{P})$ achieves optimality at a **basic feasible solution** x:
  - If $B \subseteq \{1, \ldots, n\}$ is a basis, the b.f.s. is: $x = [x_B, 0], \quad x_B = A_B^{-1} b$.
  - Simplex algorithm: feasibility and optimality for $(\mathscr{P})$ are given by:

$$\text{Feasibility-}(\mathscr{P}): \qquad x_B := A_B^{-1} b \geq 0 \tag{4a}$$

$$\text{Optimality-}(\mathscr{P}): \qquad c^\mathsf{T} - c_B^\mathsf{T} A_B^{-1} A \geq 0 \tag{4b}$$

- $(\mathscr{D})$: same basis B can also be used to determine **a dual vector** p:

$$p^\mathsf{T} A_i = c_i, \forall\, i \in B \quad \Rightarrow \quad p^\mathsf{T} = c_B^\mathsf{T} A_B^{-1}, \forall\, i \in B.$$

  - The dual objective value of p is exactly: $p^\mathsf{T} b = c_B^\mathsf{T} A_B^{-1} b = c^\mathsf{T} x$
  - p is feasible in the dual if and only if:

Optimality for Standard-Form LPs

$$(\mathscr{P}) \;\; \min\, c^\intercal x \qquad\qquad (\mathscr{D}) \;\; \max\, p^\intercal b$$

$$Ax = b, \quad x \geq 0 \qquad\qquad p^\intercal A \leq c^\intercal$$

- $(\mathscr{P})$ achieves optimality at a **basic feasible solution** x:
  - If $B \subseteq \{1, \ldots, n\}$ is a basis, the b.f.s. is: $x = [x_B, 0], \;\; x_B = A_B^{-1} b$.
  - Simplex algorithm: feasibility and optimality for $(\mathscr{P})$ are given by:

$$\text{Feasibility-}(\mathscr{P}): \qquad x_B := A_B^{-1} b \geq 0 \tag{4a}$$

$$\text{Optimality-}(\mathscr{P}): \qquad c^\intercal - c_B^\intercal A_B^{-1} A \geq 0 \tag{4b}$$

- $(\mathscr{D})$: same basis B can also be used to determine **a dual vector** p:

$$p^\intercal A_i = c_i, \, \forall\, i \in B \quad \Rightarrow \quad p^\intercal = c_B^\intercal A_B^{-1}, \, \forall\, i \in B.$$

  - The dual objective value of p is exactly: $\;p^\intercal b = c_B^\intercal A_B^{-1} b = c^\intercal x$
  - p is feasible in the dual if and only if:

$$\text{Feasibility-}(\mathscr{D}): \qquad c^\intercal - p^\intercal A \geq 0 \quad \Leftrightarrow \quad c^\intercal - c_B^\intercal A_B^{-1} A \geq 0 \tag{5}$$

Primal optimality $\Leftrightarrow$ Dual feasibility

Simplex terminates when finding a dual-feasible solution!

Solve $(\mathscr{P})$ or $(\mathscr{D})$?

$$(\mathscr{P})\ \min\ c^\intercal x \qquad Ax = b, \quad x \geq 0$$

$$(\mathscr{D})\ \max\ p^\intercal b \qquad p^\intercal A \leq c^\intercal$$

Solve $(\mathscr{P})$ or $(\mathscr{D})$?

$$(\mathscr{P}) \ \min\ c^\intercal x \qquad\qquad (\mathscr{D}) \ \max\ p^\intercal b$$
$$Ax = b, \quad x \geq 0 \qquad\qquad p^\intercal A \leq c^\intercal$$

Primal simplex

- maintain a **basic feasible solution**
- basis $B \subset \{1, \ldots, n\}$
- stopping criterion: dual feasibility

Solve $(\mathscr{P})$ or $(\mathscr{D})$?

$$(\mathscr{P}) \quad \min c^\intercal x \qquad Ax = b, \quad x \geq 0$$

$$(\mathscr{D}) \quad \max p^\intercal b \qquad p^\intercal A \leq c^\intercal$$

Primal simplex

- maintain a **basic feasible solution**
- basis $B \subset \{1, \ldots, n\}$
- stopping criterion: dual feasibility

Dual simplex

- maintain a dual feasible solution
- stopping criterion: primal feasibility
- different from primal simplex: works with an LP with inequalities

- How to choose $(\mathscr{P})$ or $(\mathscr{D})$?
- Suppose we have x^*, p^* and must solve a **larger** problem. *Any ideas?*

Solve $(\mathscr{P})$ or $(\mathscr{D})$?

$$(\mathscr{P}) \quad \min c^\mathsf{T} x \qquad Ax = b, \quad x \geq 0$$

$$(\mathscr{D}) \quad \max p^\mathsf{T} b \qquad p^\mathsf{T} A \leq c^\mathsf{T}$$

Primal simplex

- maintain a **basic feasible solution**
- basis $B \subset \{1, \ldots, n\}$
- stopping criterion: dual feasibility

Dual simplex

- maintain a dual feasible solution
- stopping criterion: primal feasibility
- different from primal simplex: works with an LP with inequalities

- How to choose $(\mathscr{P})$ or $(\mathscr{D})$?
- Suppose we have x^*, p^* and must solve a **larger** problem. *Any ideas?*
- Modern solvers include primal and dual simplex and allow concurrent runs

Dual Variables As Marginal Costs

$$(\mathscr{P})\ \min c^\intercal x \qquad\qquad (\mathscr{D})\ \max p^\intercal b$$
$$Ax = b, \quad x \geq 0 \qquad\qquad p^\intercal A \leq c^\intercal$$

- Solved the LP and obtained x^* and p^*
- Want to show that p^* is **gradient of the optimal cost with respect to** b ("almost everywhere")
- Related to **sensitivity analysis**
  How do the optimal value and solution depend on problem data A, b, c?

Global Dependency On b

$$(\mathscr{P}) \ \min c^\intercal x \qquad\qquad (\mathscr{D}) \ \max p^\intercal b$$
$$Ax = b, \quad x \geq 0 \qquad\qquad p^\intercal A \leq c^\intercal$$

- Let $P(b) := \{x : Ax = b, x \geq 0\}$ and $F(b)$ denote the optimal cost
- Assume that dual is feasible: $\{p : p^\intercal A \leq c^\intercal\} \neq \emptyset$, so $F(b) > -\infty$
- Want to show that $F(b)$ is **piecewise linear and convex**

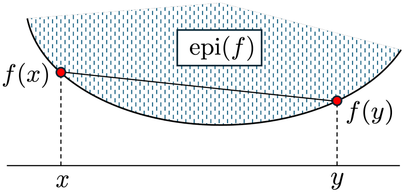
Convex and Concave Functions

Definition

$f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is **convex** if X is a convex set and

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y), \quad \forall x, y \in X \text{ and } \lambda \in [0,1]. \tag{6}$$

A function is **concave** if $-f$ is convex.



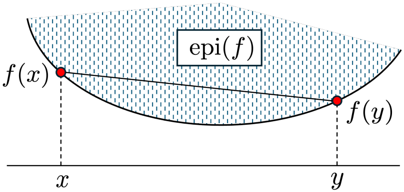
Convex and Concave Functions

Definition

$f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is **convex** if X is a convex set and

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda) f(y), \quad \forall x, y \in X \text{ and } \lambda \in [0,1]. \tag{6}$$

A function is **concave** if $-f$ is convex.



Equivalent definition in terms of **epigraph**:

$$\text{epi}(f) = \{(x,t) \in X \times \mathbb{R} \,:\, t \geq f(x)\} \tag{7}$$

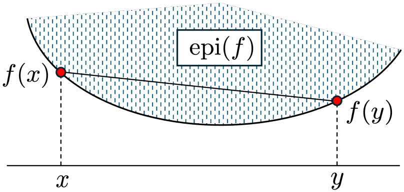
Convex and Concave Functions

Definition

$f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is **convex** if X is a convex set and

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda) f(y), \quad \forall x, y \in X \text{ and } \lambda \in [0,1]. \qquad (6)$$

A function is **concave** if $-f$ is convex.



Equivalent definition in terms of **epigraph**:

$$\text{epi}(f) = \{(x,t) \in X \times \mathbb{R} \,:\, t \geq f(x)\} \qquad (7)$$

f is convex if and only if epi(f) is a convex set.

Global Dependency On b

$$F(b) := \min\left\{c^\intercal x : Ax = b,\ x \geq 0\right\} \equiv \max\left\{p^\intercal b \,:\, p^\intercal A \leq c^\intercal\right\}$$

Theorem

$F(b)$ *is a convex and piece-wise linear function of* b *on* $S := \{b : P(b) \neq \emptyset\}$.

Global Dependency On b

$$F(b) := \min\left\{c^\mathsf{T}x : Ax = b,\ x \geq 0\right\} \equiv \max\left\{p^\mathsf{T}b\ :\ p^\mathsf{T}A \leq c^\mathsf{T}\right\}$$

Theorem

$F(b)$ is a convex and piece-wise linear function of b on $S := \{b : P(b) \neq \emptyset\}$.

Proof. Claim: S is convex.

Global Dependency On b

$$F(b) := \min\left\{c^\intercal x : Ax = b,\ x \geq 0\right\} \equiv \max\left\{p^\intercal b\ :\ p^\intercal A \leq c^\intercal\right\}$$

Theorem

$F(b)$ is a convex and piece-wise linear function of b on $S := \{b : P(b) \neq \emptyset\}$.

Proof. Claim: S is convex.

- Let $b_1, b_2 \in S$, $\lambda \in [0, 1]$, and $b := \lambda b_1 + (1 - \lambda) b_2$. Must prove that $b \in S$.
- Let $x_i \in$ argmax$\{c^\intercal x : x \geq 0, Ax = b_i\}$ and $x_\lambda := \lambda x_1 + (1 - \lambda) x_2$.
- Note that:

Global Dependency On b

$$F(b) := \min\left\{c^\intercal x : Ax = b,\ x \geq 0\right\} \equiv \max\left\{p^\intercal b\ :\ p^\intercal A \leq c^\intercal\right\}$$

Theorem

$F(b)$ is a convex and piece-wise linear function of b on $S := \{b : P(b) \neq \emptyset\}$.

Proof. Claim: S is convex.

- Let $b_1, b_2 \in S$, $\lambda \in [0,1]$, and $b := \lambda b_1 + (1-\lambda) b_2$. Must prove that $b \in S$.
- Let $x_i \in \text{argmax}\{c^\intercal x : x \geq 0, Ax = b_i\}$ and $x_\lambda := \lambda x_1 + (1-\lambda) x_2$.
- Note that:

$$x_\lambda \geq 0 \text{ and } Ax_\lambda = A(\lambda x_1 + (1-\lambda)x_2) = \lambda b_1 + (1-\lambda) b_2 := b,$$

Global Dependency On b

$$F(b) := \min\left\{c^\intercal x : Ax = b,\ x \geq 0\right\} \equiv \max\left\{p^\intercal b\ :\ p^\intercal A \leq c^\intercal\right\}$$

Theorem

$F(b)$ is a convex and piece-wise linear function of b on $S := \{b : P(b) \neq \emptyset\}$.

Proof. Claim: S is convex.

- Let $b_1, b_2 \in S$, $\lambda \in [0,1]$, and $b := \lambda b_1 + (1-\lambda) b_2$. Must prove that $b \in S$.
- Let $x_i \in$ argmax$\{c^\intercal x : x \geq 0, Ax = b_i\}$ and $x_\lambda := \lambda x_1 + (1-\lambda) x_2$.
- Note that:

$$x_\lambda \geq 0 \text{ and } Ax_\lambda = A(\lambda x_1 + (1-\lambda)x_2) = \lambda b_1 + (1-\lambda) b_2 := b,$$

$$\Rightarrow \quad x_\lambda \in P(b) \quad \Rightarrow \quad b \in S \quad \Rightarrow \quad S \text{ is convex.}$$

Global Dependency On b

$$F(b) := \min\left\{c^\intercal x : Ax = b,\ x \geq 0\right\} \equiv \max\left\{p^\intercal b \,:\, p^\intercal A \leq c^\intercal\right\}$$

Theorem

$F(b)$ is a convex and piece-wise linear function of b on $S := \{b : P(b) \neq \emptyset\}$.

Proof. Because $(\mathscr{D})$ feasible $\Rightarrow\ F(b) > -\infty$.

Global Dependency On b

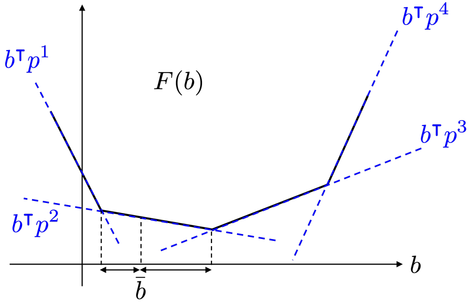
$$F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b \ : \ p^\intercal A \leq c^\intercal\}$$

Theorem

$F(b)$ is a convex and piece-wise linear function of b on $S := \{b : P(b) \neq \emptyset\}$.

Proof. Because $(\mathscr{D})$ feasible $\Rightarrow$ $F(b) > -\infty$.

- If $p^1, p^2, \ldots, p^r$ are the extreme points of the dual feasible set, then: $F(b) = \max_{i=1,\ldots,r} b^\intercal p^i$, $\forall\, b \in S$



Global Dependency On b

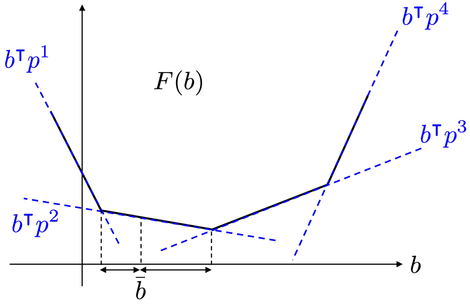
$$F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b \,:\, p^\intercal A \leq c^\intercal\}$$

Theorem

$F(b)$ is a convex and piece-wise linear function of b on $S := \{b : P(b) \neq \emptyset\}$.

Proof. Because $(\mathscr{D})$ feasible $\Rightarrow$ $F(b) > -\infty$.

- If $p^1, p^2, \ldots, p^r$ are the extreme points of the dual feasible set, then: $F(b) = \max_{i=1,\ldots,r} b^\intercal p^i$, $\forall\, b \in S$



How to complete proof that $F(b)$ is convex?

Global Dependency On b

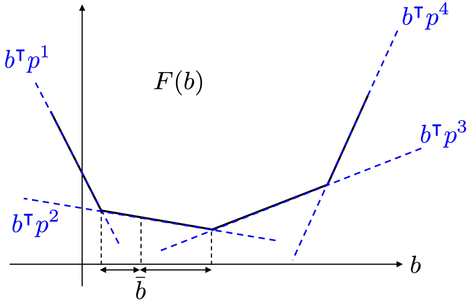
$$F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b \ : \ p^\intercal A \leq c^\intercal\}$$

Theorem

$F(b)$ is a convex and piece-wise linear function of b on $S := \{b : P(b) \neq \emptyset\}$.

Proof. Because $(\mathscr{D})$ feasible $\Rightarrow$ $F(b) > -\infty$.

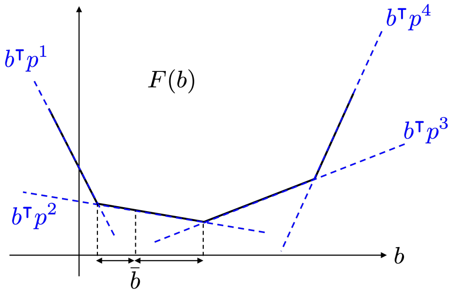
- If $p^1, p^2, \ldots, p^r$ are the extreme points of the dual feasible set, then: $F(b) = \max_{i=1,\ldots,r} b^\intercal p^i$, $\forall\, b \in S$



How to complete proof that $F(b)$ is convex?

Global Dependency On b - Implications

$$F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b \,:\, p^\intercal A \leq c^\intercal\}$$

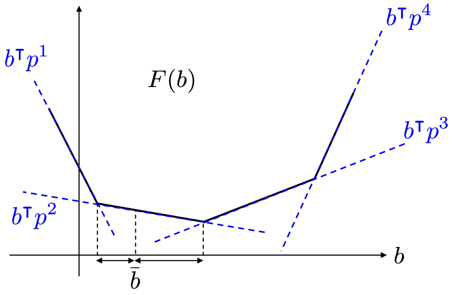


- At any $b = \bar{b}$ where $F(b)$ is differentiable, p^* **is the gradient of** $F(b)$
- p_i^* acts as a **marginal cost** or **shadow price** for the i-th constraint r.h.s. b_i
- p_i allows estimating **exact change in** $F(b)$ **in a range around** $\bar{b}$
- Modern solvers give direct access to p_i^* and the range

  Gurobipy: for constraint **c**, the attribute **c.Pi** is p_i^* and the range is from **c.SARHSLow** to **c.SARHSUp**

Global Dependency On b - Implications

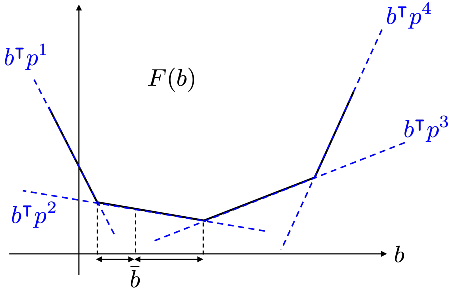
$$F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b : p^\intercal A \leq c^\intercal\}$$



- At b where $F(b)$ is not differentiable, several p^i are optimal
- All such p^i are valid **subgradients** of $F(b)$

Global Dependency On b - Implications

$$F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b \,:\, p^\intercal A \leq c^\intercal\}$$



- At b where $F(b)$ is not differentiable, several p^i are optimal
- All such p^i are valid **subgradients** of $F(b)$

Definition (Subgradient.)

F convex, defined on (convex) set S. A vector p is a **subgradient** of F at $\bar{b} \in S$ if

$$F(\bar{b}) + p^\intercal (b - \bar{b}) \leq F(b), \quad \forall b \in S.$$

Optimal Duals As Subgradients

Theorem

Suppose $F(b) := \min\{c^\mathsf{T} x : Ax = b,\ x \geq 0\} \equiv \max\{p^\mathsf{T} b \,:\, p^\mathsf{T} A \leq c^\mathsf{T}\} > -\infty$*.*
Then p *is optimal for the dual* **if and only if** *it is a subgradient of* F *at* $\bar{b}$*.*

Proof. First show that any dual optimal p is a valid subgradient.

Optimal Duals As Subgradients

Theorem

Suppose $F(b) := \min\{c^\mathsf{T} x : Ax = b,\ x \geq 0\} \equiv \max\{p^\mathsf{T} b \,:\, p^\mathsf{T} A \leq c^\mathsf{T}\} > -\infty$*.*
Then p *is optimal for the dual* **if and only if** *it is a subgradient of* F *at* $\bar{b}$*.*

Proof. First show that any dual optimal p is a valid subgradient.

- Suppose that p is optimal for the dual
- Strong duality implies $p^\mathsf{T}\bar{b} = F(\bar{b})$

Optimal Duals As Subgradients

Theorem

Suppose $F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b \, : \, p^\intercal A \leq c^\intercal\} > -\infty$*.*
Then p *is optimal for the dual* **if and only if** *it is a subgradient of* F *at* $\bar{b}$*.*

Proof. First show that any dual optimal p is a valid subgradient.

- Suppose that p is optimal for the dual
- Strong duality implies $p^\intercal \bar{b} = F(\bar{b})$
- Consider arbitrary $b \in S$
- For any feasible solution $x \in P(b)$, weak duality yields $p^\intercal b \leq c^\intercal x$
- This implies $p^\intercal b \leq F(b)$

Optimal Duals As Subgradients

Theorem

Suppose $F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b \,:\, p^\intercal A \leq c^\intercal\} > -\infty$. *Then* p *is optimal for the dual* **if and only if** *it is a subgradient of* F *at* $\bar{b}$.

Proof. First show that any dual optimal p is a valid subgradient.

- Suppose that p is optimal for the dual
- Strong duality implies $p^\intercal \bar{b} = F(\bar{b})$
- Consider arbitrary $b \in S$
- For any feasible solution $x \in P(b)$, weak duality yields $p^\intercal b \leq c^\intercal x$
- This implies $p^\intercal b \leq F(b)$
- But then, $p^\intercal b - p^\intercal \bar{b} \leq F(b) - F(\bar{b})$

We conclude that p is a subgradient of F at $\bar{b}$

Optimal Duals As Subgradients

Theorem

Suppose $F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b \,:\, p^\intercal A \leq c^\intercal\} > -\infty$*.*
Then p *is optimal for the dual* **if and only if** *it is a subgradient of* F *at* $\bar{b}$*.*

Proof. For the reverse direction, let p be a subgradient of F at $\bar{b}$, that is,

$$F(\bar{b}) + p^\intercal(b - \bar{b}) \leq F(b), \quad \forall b \in S. \tag{8}$$

Optimal Duals As Subgradients

Theorem

Suppose $F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b\ :\ p^\intercal A \leq c^\intercal\} > -\infty$*.*
Then p *is optimal for the dual* **if and only if** *it is a subgradient of* F *at* $\bar{b}$*.*

Proof. For the reverse direction, let p be a subgradient of F at $\bar{b}$, that is,

$$F(\bar{b}) + p^\intercal(b - \bar{b}) \leq F(b), \quad \forall b \in S. \tag{8}$$

- Pick some $x \geq 0$ and let $b = Ax$, which implies $x \in P(b)$ and $F(b) \leq c^\intercal x$.

Optimal Duals As Subgradients

Theorem

Suppose $F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b \ :\ p^\intercal A \leq c^\intercal\} > -\infty$*. Then* p *is optimal for the dual* **if and only if** *it is a subgradient of* F *at* $\bar{b}$*.*

Proof. For the reverse direction, let p be a subgradient of F at $\bar{b}$, that is,

$$F(\bar{b}) + p^\intercal(b - \bar{b}) \leq F(b), \quad \forall b \in S. \tag{8}$$

- Pick some $x \geq 0$ and let $b = Ax$, which implies $x \in P(b)$ and $F(b) \leq c^\intercal x$.
- By (8), we have: $p^\intercal Ax = p^\intercal b \leq F(b) - F(\bar{b}) + p^\intercal \bar{b} \leq c^\intercal x - F(\bar{b}) + p^\intercal \bar{b}$.
- Because this is true for any $x \geq 0$, we must have $p^\intercal A \leq c^\intercal$. *Why?*

Optimal Duals As Subgradients

Theorem

Suppose $F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b \, : \, p^\intercal A \leq c^\intercal\} > -\infty$*.*
Then p *is optimal for the dual* **if and only if** *it is a subgradient of* F *at* $\bar{b}$*.*

Proof. For the reverse direction, let p be a subgradient of F at $\bar{b}$, that is,

$$F(\bar{b}) + p^\intercal(b - \bar{b}) \leq F(b), \quad \forall b \in S. \tag{8}$$

- Pick some $x \geq 0$ and let $b = Ax$, which implies $x \in P(b)$ and $F(b) \leq c^\intercal x$.
- By (8), we have: $p^\intercal Ax = p^\intercal b \leq F(b) - F(\bar{b}) + p^\intercal \bar{b} \leq c^\intercal x - F(\bar{b}) + p^\intercal \bar{b}$.
- Because this is true for any $x \geq 0$, we must have $p^\intercal A \leq c^\intercal$. *Why?*
- This implies that p is dual-feasible
- With $x = 0$, we obtain $F(\bar{b}) \leq p^\intercal \bar{b}$

Optimal Duals As Subgradients

Theorem

Suppose $F(b) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b\ :\ p^\intercal A \leq c^\intercal\} > -\infty$.
Then p *is optimal for the dual* **if and only if** *it is a subgradient of* F *at* $\bar{b}$.

Proof. For the reverse direction, let p be a subgradient of F at $\bar{b}$, that is,

$$F(\bar{b}) + p^\intercal(b - \bar{b}) \leq F(b), \quad \forall b \in S. \tag{8}$$

- Pick some $x \geq 0$ and let $b = Ax$, which implies $x \in P(b)$ and $F(b) \leq c^\intercal x$.
- By (8), we have: $p^\intercal Ax = p^\intercal b \leq F(b) - F(\bar{b}) + p^\intercal \bar{b} \leq c^\intercal x - F(\bar{b}) + p^\intercal \bar{b}$.
- Because this is true for any $x \geq 0$, we must have $p^\intercal A \leq c^\intercal$. *Why?*
- This implies that p is dual-feasible
- With $x = 0$, we obtain $F(\bar{b}) \leq p^\intercal \bar{b}$
- Using weak duality, every dual-feasible q satisfies $q^\intercal \bar{b} \leq F(\bar{b}) \leq p^\intercal \bar{b}$

We conclude that p is optimal.

Global Dependency On c

Let $G(c) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b \,:\, p^\intercal A \leq c^\intercal\}$

Theorem

For an LP in standard form,

1. *The set $T := \{c : G(c) > -\infty\}$ is convex.*
2. *$G(c)$ is a* **concave** *function of c on the set T.*
3. *If for some c the LP has a* **unique** *optimal solution x^*, then G is linear in the vicinity of c and its gradient is x^*.*

Global Dependency On c

Let $G(c) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b\ :\ p^\intercal A \leq c^\intercal\}$

Theorem

For an LP in standard form,

1. *The set $T := \{c : G(c) > -\infty\}$ is convex.*
2. *$G(c)$ is a* **concave** *function of c on the set T.*
3. *If for some c the LP has a* **unique** *optimal solution x^*, then G is linear in the vicinity of c and its gradient is x^*.*

Proof. Analogous ideas applied to the dual - omitted.

Global Dependency On c

Let $G(c) := \min\{c^\intercal x : Ax = b,\ x \geq 0\} \equiv \max\{p^\intercal b \,:\, p^\intercal A \leq c^\intercal\}$

Theorem

For an LP in standard form,

1. *The set $T := \{c : G(c) > -\infty\}$ is convex.*
2. *$G(c)$ is a* **concave** *function of c on the set T.*
3. *If for some c the LP has a* **unique** *optimal solution x^*, then G is linear in the vicinity of c and its gradient is x^*.*

Proof. Analogous ideas applied to the dual - omitted.

- The optimal primal solution x^* is a **shadow price for the dual constraints**
- x^* remains optimal for a range of change in each objective coefficient c_j
- Modern solvers also allow obtaining the range directly

  Gurobipy: attributes **SAObjLow** and **SAObjUp** for each decision variable

Signs of Dual Variables Revisited

These ideas carry over directly to **primals in general form**:

$$
\begin{array}{llll}
F(b,c) := \min_x & c^\intercal x & & \\
& a_i^\intercal x \geq b_i, & i \in M_1, & \\
& a_i^\intercal x \leq b_i, & i \in M_2, & \\
& a_i^\intercal x = b_i, & i \in M_3, & \\
& x_j \geq 0, & j \in N_1, & \\
& x_j \leq 0, & j \in N_2, & \\
& x_j \text{ free}, & j \in N_3. &
\end{array}
\qquad
\begin{array}{lll}
\max_p & p^\intercal b & \\
& p_i \geq 0, & i \in M_1, \\
& p_i \leq 0, & i \in M_2, \\
& p_i \text{ free}, & i \in M_3, \\
& p^\intercal A_j \leq c_j, & j \in N_1, \\
& p^\intercal A_j \geq c_j, & j \in N_2, \\
& p^\intercal A_j = c_j, & j \in N_3.
\end{array}
$$

Signs of Dual Variables Revisited

These ideas carry over directly to **primals in general form**:

$$
\begin{array}{llll}
F(b,c) := \min_x & c^\mathsf{T} x & & \max_p & p^\mathsf{T} b & \\
& a_i^\mathsf{T} x \geq b_i, & i \in M_1, & & p_i \geq 0, & i \in M_1, \\
& a_i^\mathsf{T} x \leq b_i, & i \in M_2, & & p_i \leq 0, & i \in M_2, \\
& a_i^\mathsf{T} x = b_i, & i \in M_3, & & p_i \text{ free}, & i \in M_3, \\
& x_j \geq 0, & j \in N_1, & & p^\mathsf{T} A_j \leq c_j, & j \in N_1, \\
& x_j \leq 0, & j \in N_2, & & p^\mathsf{T} A_j \geq c_j, & j \in N_2, \\
& x_j \text{ free}, & j \in N_3. & & p^\mathsf{T} A_j = c_j, & j \in N_3.
\end{array}
$$

- $F(b,c)$ is piece-wise linear, convex in b and piece-wise linear, concave in c
- p^* are subgradients for $F(b,c)$ with respect to b
- x^* are subgradients for $-F(b,c)$ with respect to c

Signs of Dual Variables Revisited

These ideas carry over directly to **primals in general form**:

$$
\begin{array}{llll}
F(b,c) := \min_x & c^\intercal x & & \\
& a_i^\intercal x \geq b_i, & i \in M_1, \\
& a_i^\intercal x \leq b_i, & i \in M_2, \\
& a_i^\intercal x = b_i, & i \in M_3, \\
& x_j \geq 0, & j \in N_1, \\
& x_j \leq 0, & j \in N_2, \\
& x_j \text{ free}, & j \in N_3.
\end{array}
\qquad
\begin{array}{lll}
\max_p & p^\intercal b & \\
& p_i \geq 0, & i \in M_1, \\
& p_i \leq 0, & i \in M_2, \\
& p_i \text{ free}, & i \in M_3, \\
& p^\intercal A_j \leq c_j, & j \in N_1, \\
& p^\intercal A_j \geq c_j, & j \in N_2, \\
& p^\intercal A_j = c_j, & j \in N_3.
\end{array}
$$

- $F(b,c)$ is piece-wise linear, convex in b and piece-wise linear, concave in c
- p^* are subgradients for $F(b,c)$ with respect to b
- x^* are subgradients for $-F(b,c)$ with respect to c
- There is a direct connection between:
  - **the optimization problem** (max/min)
  - the **constraint type** ($\leq$, $\geq$)
  - **the signs of the shadow prices**

Signs of Dual Variables Revisited

- There is a direct connection between:
  - **the optimization problem** (max/min)
  - the **constraint type** ($\leq$, $\geq$)
  - **the signs of the shadow prices**
- Given two of these, can figure out the third one!
- *What is the sign of the shadow price for a ...*

  $\leq$ constraint in a **minimization** problem ?

  $\geq$ constraint in a **minimization** problem ?

  $\leq$ constraint in a **maximization** problem ?

  $\leq$ constraint in a **maximization** problem ?
- *What is the dependency of the optimal objective on the r.h.s. of a ...*

  $\leq$ constraint in a **minimization** problem ?

  $\geq$ constraint in a **minimization** problem ?

  $\leq$ constraint in a **maximization** problem ?

  $\leq$ constraint in a **maximization** problem ?

Signs of Dual Variables Revisited

- There is a direct connection between:
  - **the optimization problem** (max/min)
  - the **constraint type** ($\leq$, $\geq$)
  - **the signs of the shadow prices**
- Given two of these, can figure out the third one!

Optimality Conditions and Complementary Slackness

$$
\begin{array}{llll}
\min_x & c^\top x & & \\
& a_i^\top x \ge b_i, & i \in M_1, \\
& a_i^\top x \le b_i, & i \in M_2, \\
& a_i^\top x = b_i, & i \in M_3, \\
& x_j \ge 0, & j \in N_1, \\
& x_j \le 0, & j \in N_2, \\
& x_j \text{ free}, & j \in N_3.
\end{array}
\qquad
\begin{array}{llll}
\max_p & p^\top b & & \\
& p_i \ge 0, & i \in M_1, \\
& p_i \le 0, & i \in M_2, \\
& p_i \text{ free}, & i \in M_3, \\
& p^\top A_j \le c_j, & j \in N_1, \\
& p^\top A_j \ge c_j, & j \in N_2, \\
& p^\top A_j = c_j, & j \in N_3.
\end{array}
$$

Sometimes, we just want to **characterize** the optimal solutions

Optimality Conditions and Complementary Slackness

$$
\begin{array}{llll}
\min_x & c^\intercal x & & \\
 & a_i^\intercal x \geq b_i, & i \in M_1, \\
 & a_i^\intercal x \leq b_i, & i \in M_2, \\
 & a_i^\intercal x = b_i, & i \in M_3, \\
 & x_j \geq 0, & j \in N_1, \\
 & x_j \leq 0, & j \in N_2, \\
 & x_j \text{ free}, & j \in N_3.
\end{array}
\qquad
\begin{array}{llll}
\max_p & p^\intercal b & & \\
 & p_i \geq 0, & i \in M_1, \\
 & p_i \leq 0, & i \in M_2, \\
 & p_i \text{ free}, & i \in M_3, \\
 & p^\intercal A_j \leq c_j, & j \in N_1, \\
 & p^\intercal A_j \geq c_j, & j \in N_2, \\
 & p^\intercal A_j = c_j, & j \in N_3.
\end{array}
$$

Sometimes, we just want to **characterize** the optimal solutions

Theorem (Complementary Slackness)

Let x and p be feasible solutions for $(\mathscr{P})$ and $(\mathscr{D})$, respectively. Then x and p **are optimal solutions** *for $(\mathscr{P})$ and $(\mathscr{D})$* **if and only if***:*

$$p_i(a_i^\intercal x - b_i) = 0, \; \forall i$$

$$(c_j - p^\intercal A_j)x_j = 0, \; \forall j.$$

Optimality Conditions and Complementary Slackness

Theorem (**General** Complementary Slackness)

Let x and p be feasible solutions for $(\mathscr{P})$ and $(\mathscr{D})$, respectively. Then x and p **are optimal solutions** *for $(\mathscr{P})$ and $(\mathscr{D})$* **if and only if***:*

$$p_i(a_i^\intercal x - b_i) = 0, \ \forall i$$
$$(c_j - p^\intercal A_j)x_j = 0, \ \forall j.$$

Optimality Conditions and Complementary Slackness

Theorem (**General** Complementary Slackness)

Let x and p be feasible solutions for $(\mathscr{P})$ and $(\mathscr{D})$, respectively. Then x and p **are optimal solutions** *for $(\mathscr{P})$ and $(\mathscr{D})$* **if and only if***:*

$$p_i(a_i^\intercal x - b_i) = 0,\ \forall i$$
$$(c_j - p^\intercal A_j)x_j = 0,\ \forall j.$$

Theorem (**Strict** C.S. **Standard-Form LPs**)

Consider the following primal-dual pair of LPs:

$$(\mathscr{P})\ \min c^\intercal x \qquad\qquad (\mathscr{D})\ \max p^\intercal b$$
$$Ax = b, x \geq 0 \qquad\qquad p^\intercal A \leq c^\intercal$$

If $(\mathscr{P})$ and $(\mathscr{D})$ are feasible, they admit optimal solutions x^* and p^* satisfying **strict complementarity***: $x_j^* > 0 \ \Leftrightarrow\ p^\intercal A_j = c_j$.*

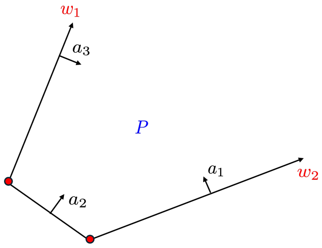
Representation of Polyhedra

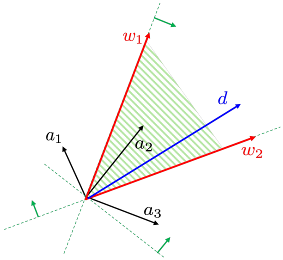
Important consequence of duality: alternative representation of all polyhedra

Definition (Extreme rays of a polyhedron)

Consider a nonempty polyhedron $P = \{x \in \mathbb{R}^n \,:\, Ax \geq b\}$. Then:

1. $\mathcal{C} := \{d \in \mathbb{R}^n \,:\, Ad \geq 0\}$ is called the **recession cone** of P.
2. Any $d \in \mathcal{C}$ with $d \neq 0$ is called a **ray** of P.
3. Any ray d that satisfies $a_i^\intercal d = 0$ for $n-1$ linearly independent a_i is called an **extreme ray** of P.





Representation of Polyhedra

Theorem (Resolution Theorem)

Let $P = \{x \in \mathbb{R}^n \,:\, Ax \geq b\}$ be a non-empty polyhedron, $x^1, x^2, \ldots, x^k$ be its extreme points, and $w^1, w^2, \ldots, w^r$ be its extreme rays. Then $P = Q$, where

$$Q := \left\{ \sum_{i=1}^{k} \lambda_i x^i + \sum_{j=1}^{r} \theta_j w^j \;:\; \lambda \geq 0,\ \theta \geq 0,\ e^\mathsf{T} \lambda = 1 \right\}.$$

Proof. Proving $Q \subseteq P$ is immediate. To prove $P \subseteq Q$, assume $\exists z \in P$ with $z \notin Q$. Consider the following primal-dual pair:

$$(\mathscr{P}) \quad \max_{\lambda \geq 0, \theta \geq 0} \; \sum_{i=1}^{k} 0\lambda_i + \sum_{j=1}^{r} 0\theta_j$$

$$\sum_{i=1}^{k} \lambda_i x^i + \sum_{j=1}^{r} \theta_j w^j = z$$

$$\sum_{i=1}^{k} \lambda_i = 1$$

$$(\mathscr{D}) \quad \min_{p,q} \; p^\mathsf{T} z + q$$

$$p^\mathsf{T} x_i + q \geq 0, \quad i = 1, \ldots, k,$$

$$p^\mathsf{T} w_j \geq 0, \qquad j = 1, \ldots, r,$$

Is $(\mathscr{P})$ feasible? Is $(\mathscr{D})$ feasible? What are the optimal values?

Representation of Polyhedra - cntd

$$P := \{x \in \mathbb{R}^n : Ax \geq b\} = Q := \left\{ \sum_{i=1}^k \lambda_i x^i + \sum_{j=1}^r \theta_j w^j : \lambda \geq 0, \theta \geq 0, e^\mathsf{T} \lambda = 1 \right\}.$$

Proof - cont'd. Assume $\exists z \in P$ with $z \notin Q$. Consider the following primal-dual pair:

$$(\mathscr{P}) \quad \max_{\lambda \geq 0, \theta \geq 0} \sum_{i=1}^k 0\lambda_i + \sum_{j=1}^r 0\theta_j \qquad\qquad (\mathscr{D}) \quad \min_{p,q} p^\mathsf{T} z + q$$

$$\sum_{i=1}^k \lambda_i x^i + \sum_{j=1}^r \theta_j w^j = z \qquad\qquad p^\mathsf{T} x_i + q \geq 0, \quad i = 1, \ldots, k,$$

$$\sum_{i=1}^k \lambda_i = 1 \qquad\qquad p^\mathsf{T} w_j \geq 0, \quad j = 1, \ldots, r,$$

- $(\mathscr{P})$ is infeasible because $z \notin Q$
- $(\mathscr{D})$ is feasible with $p = q = 0$, so its optimal value is $-\infty \Rightarrow \exists (p,q) : p^\mathsf{T} z + q < 0$

Representation of Polyhedra - cntd

$P := \{x \in \mathbb{R}^n : Ax \geq b\} = Q := \left\{ \sum_{i=1}^{k} \lambda_i x^i + \sum_{j=1}^{r} \theta_j w^j : \lambda \geq 0, \theta \geq 0, e^\intercal \lambda = 1 \right\}$.

Proof - cont'd. Assume $\exists z \in P$ with $z \notin Q$. Consider the following primal-dual pair:

$$(\mathscr{P}) \quad \max_{\lambda \geq 0, \theta \geq 0} \sum_{i=1}^{k} 0\lambda_i + \sum_{j=1}^{r} 0\theta_j \qquad \sum_{i=1}^{k} \lambda_i x^i + \sum_{j=1}^{r} \theta_j w^j = z \qquad \sum_{i=1}^{k} \lambda_i = 1$$

$$(\mathscr{D}) \quad \min_{p,q} p^\intercal z + q \qquad p^\intercal x_i + q \geq 0, \quad i = 1, \ldots, k, \qquad p^\intercal w_j \geq 0, \quad j = 1, \ldots, r,$$

- $(\mathscr{P})$ is infeasible because $z \notin Q$
- $(\mathscr{D})$ is feasible with $p = q = 0$, so its optimal value is $-\infty \Rightarrow \exists (p, q) : p^\intercal z + q < 0$
- (p, q) feasible $\Rightarrow p^\intercal z < -q \leq p^\intercal x_i$ for any $i = 1, \ldots, k$ and $p^\intercal w_i \geq 0$
- With p as above, consider the LP $\min_x \{p^\intercal x : Ax \geq b\}$

Representation of Polyhedra - cntd

$P := \{x \in \mathbb{R}^n : Ax \geq b\} = Q := \left\{ \sum_{i=1}^k \lambda_i x^i + \sum_{j=1}^r \theta_j w^j : \lambda \geq 0, \theta \geq 0, e^\intercal \lambda = 1 \right\}.$

Proof - cont'd. Assume $\exists z \in P$ with $z \notin Q$. Consider the following primal-dual pair:

$$(\mathscr{P}) \quad \max_{\lambda \geq 0, \theta \geq 0} \sum_{i=1}^k 0\lambda_i + \sum_{j=1}^r 0\theta_j \qquad\qquad (\mathscr{D}) \quad \min_{p,q} p^\intercal z + q$$

$$\sum_{i=1}^k \lambda_i x^i + \sum_{j=1}^r \theta_j w^j = z \qquad\qquad \begin{aligned} p^\intercal x_i + q &\geq 0, \quad i = 1, \ldots, k, \\ p^\intercal w_j &\geq 0, \quad j = 1, \ldots, r, \end{aligned}$$

$$\sum_{i=1}^k \lambda_i = 1$$

- $(\mathscr{P})$ is infeasible because $z \notin Q$
- $(\mathscr{D})$ is feasible with $p = q = 0$, so its optimal value is $-\infty \Rightarrow \exists (p, q) : p^\intercal z + q < 0$
- (p, q) feasible $\Rightarrow p^\intercal z < -q \leq p^\intercal x_i$ for any $i = 1, \ldots, k$ and $p^\intercal w_i \geq 0$
- With p as above, consider the LP $\min_x \{p^\intercal x : Ax \geq b\}$
- If optimal cost finite, $\exists x^i$ optimal. But $z \in P$ and $p^\intercal z < p^\intercal x_i$ lead to $\lightning$
- If cost is $-\infty$, $\exists w^j : p^\intercal w^j < 0$, which is also a $\lightning$

Asset Pricing and No-Arbitrage

- Investment world with $n+1$ securities indexed by $i = 0, \ldots, n$
- $i = 0$ denotes cash; the other securities can be anything (stocks, derivatives, ...)
- We have two periods: current period **c**, future period **f**

Asset Pricing and No-Arbitrage

- Investment world with $n+1$ securities indexed by $i = 0, \ldots, n$
- $i = 0$ denotes cash; the other securities can be anything (stocks, derivatives, ...)
- We have two periods: current period **c**, future period **f**
- Current period: prices of securities are S_i^c for $i = 1, \ldots, n$; cash: $S_0^c = 1$

Asset Pricing and No-Arbitrage

- Investment world with $n+1$ securities indexed by $i = 0, \ldots, n$
- $i = 0$ denotes cash; the other securities can be anything (stocks, derivatives, ...)
- We have two periods: current period **c**, future period **f**
- Current period: prices of securities are S_i^c for $i = 1, \ldots, n$; cash: $S_0^c = 1$
- Future period: prices are uncertain; there are m possible states of the world $\Omega = \{\omega_1, \omega_2, \ldots, \omega_m\}$, each occurring with positive probability, and prices are:
  - cash is riskless: $S_0^f = R = 1 + r$, where r is the risk-free rate of return
  - security $i > 1$ will have price $S_i^f(\omega_j)$ in state of world ω_j

Asset Pricing and No-Arbitrage

- Investment world with $n + 1$ securities indexed by $i = 0, \ldots, n$
- $i = 0$ denotes cash; the other securities can be anything (stocks, derivatives, ...)
- We have two periods: current period **c**, future period **f**
- Current period: prices of securities are S_i^c for $i = 1, \ldots, n$; cash: $S_0^c = 1$
- Future period: prices are uncertain; there are m possible states of the world $\Omega = \{\omega_1, \omega_2, \ldots, \omega_m\}$, each occurring with positive probability, and prices are:
  - cash is riskless: $S_0^f = R = 1 + r$, where r is the risk-free rate of return
  - security $i > 1$ will have price $S_i^f(\omega_j)$ in state of world ω_j
- If we purchase x_i of each security i:
  - we incur immediate cost $\sum_{i=0}^n S_i^c x_i$
  - we have future cashflow $\sum_{i=0}^n S_i^f(\omega) \cdot x_i$ if state of world is $\omega \in \Omega$

Asset Pricing and No-Arbitrage

Definition (Arbitrage)

An **arbitrage** is a trading strategy that either has a positive initial cashflow and has no risk of a loss later (type A) or that requires no initial cash input, has no risk of loss, and has a positive probability of making profits in the future (type B).

Asset Pricing and No-Arbitrage

Definition (Arbitrage)

An **arbitrage** is a trading strategy that either has a positive initial cashflow and has no risk of a loss later (type A) or that requires no initial cash input, has no risk of loss, and has a positive probability of making profits in the future (type B).

- a type-A arbitrage means $\exists x$ such that:

$$\begin{aligned} \sum_{i=0}^{n} S_i^c \cdot x_i &< 0 && \text{(positive initial cashflow)} \\ \sum_{i=0}^{n} S_i^f(\omega) \cdot x_i &\geq 0, \, \forall \, \omega \in \Omega && \text{(no risk of loss)} \end{aligned} \tag{9}$$

Asset Pricing and No-Arbitrage

Definition (Arbitrage)

An **arbitrage** is a trading strategy that either has a positive initial cashflow and has no risk of a loss later (type A) or that requires no initial cash input, has no risk of loss, and has a positive probability of making profits in the future (type B).

- a type-A arbitrage means $\exists x$ such that:

$$\begin{aligned} \sum_{i=0}^{n} S_i^c \cdot x_i < 0 & \qquad \text{(positive initial cashflow)} \\ \sum_{i=0}^{n} S_i^f(\omega) \cdot x_i \geq 0, \, \forall \, \omega \in \Omega & \qquad \text{(no risk of loss)} \end{aligned} \tag{9}$$

- a type-B arbitrage means $\exists x$ such that:

$$\begin{aligned} \sum_{i=0}^{n} S_i^c \cdot x_i = 0 & \qquad \text{(no initial cash input)} \\ \sum_{i=0}^{n} S_i^f(\omega) \cdot x_i \geq 0, \, \forall \, \omega \in \Omega & \qquad \text{(no risk of loss)} \\ \exists \omega \in \Omega \, : \, \sum_{i=0}^{n} S_i^f(\omega) \cdot x_i > 0, & \qquad \text{(positive probability of profit).} \end{aligned} \tag{10}$$

Asset Pricing and No-Arbitrage

Definition (R.N.P.M.)

A **risk-neutral probability measure** on the set $\Omega = \{\omega_1, \omega_2, \ldots, \omega_m\}$ is a vector $p \in \mathbb{R}^m$ so that $p > 0$ and $\sum_{j=1}^m p_j = 1$ and for every security $S_i, i = 0, \ldots, n$,

$$S_i^c = \frac{1}{R}\left(\sum_{j=1}^{m} p_j S_i^f(\omega_j)\right) = \frac{1}{R}\mathbb{E}_p[S_i^f].$$

- Above, $\mathbb{E}_p[S]$ is the expected value of the random variable S under the probability distribution $p := (p_1, p_2, \ldots, p_m)$
- The definition states that the current price/value of every asset, S_i^c, exactly equals **the discounted expected price/value in the future**
- The expectation is taken with respect to the R.N.P.M.
- Discounting is done at the risk-free interest rate R

Asset Pricing and No-Arbitrage

Theorem (Asset Pricing Theorem)

A risk-neutral probability measure exists **if and only if** *there is no arbitrage.*

Proof. Consider the following linear program with variables x_i, for $i = 0, \ldots, n$:

$$\begin{aligned} \min_x \; & \sum_{i=0}^{n} S_i^c \cdot x_i \\ \text{s.t.} \; & \sum_{i=0}^{n} S_i^f(\omega_j) \cdot x_i \geq 0, j = 1, \ldots, m. \end{aligned} \tag{11}$$

Asset Pricing and No-Arbitrage

Theorem (Asset Pricing Theorem)

A risk-neutral probability measure exists **if and only if** *there is no arbitrage.*

Proof. Consider the following linear program with variables x_i, for $i = 0, \ldots, n$:

$$\begin{aligned} \min_x \ & \sum_{i=0}^{n} S_i^c \cdot x_i \\ \text{s.t.} \ & \sum_{i=0}^{n} S_i^f(\omega_j) \cdot x_i \geq 0, j = 1, \ldots, m. \end{aligned} \tag{11}$$

- Type-A arbitrage: $\exists x : \sum S_i^0 x_i < 0$
- Constraints are homogeneous, so if $\exists x : \sum S_i^0 x_i < 0$, the objective is $-\infty$
- $x = 0$ is feasible, so the optimal objective value is ≤ 0
- No type-A arbitrage if and only if the optimal objective value of this LP is 0

Asset Pricing and No-Arbitrage

Theorem (Asset Pricing Theorem)

A risk-neutral probability measure exists **if and only if** *there is no arbitrage.*

Proof. Consider the following linear program with variables x_i, for $i = 0, \ldots, n$:

$$\begin{aligned} \min_x \ & \sum_{i=0}^{n} S_i^c \cdot x_i \\ \text{s.t. } & \sum_{i=0}^{n} S_i^f(\omega_j) \cdot x_i \geq 0, j = 1, \ldots, m. \end{aligned} \tag{11}$$

- Type-A arbitrage: $\exists x : \sum S_i^0 x_i < 0$
- Constraints are homogeneous, so if $\exists x : \sum S_i^0 x_i < 0$, the objective is $-\infty$
- $x = 0$ is feasible, so the optimal objective value is ≤ 0
- No type-A arbitrage if and only if the optimal objective value of this LP is 0
- Suppose no type-A arbitrage. Then, no type-B arbitrage if and only if all constraints are tight for all optimal solutions of (11): $\sum_{i=0}^{n} S_i^f(\omega_j) \cdot x_i^* = 0$, for $j = 1, \ldots, m$

Asset Pricing and No-Arbitrage

Theorem (Asset Pricing Theorem)

A risk-neutral probability measure exists **if and only if** *there is no arbitrage.*

Proof. Consider the dual of this LP.

$$
\begin{aligned}
&\max_p 0 \\
&\text{s.t. } \sum_{j=1}^{m} p_j \cdot S_i^f(\omega_j) = S_i^c,\ i = 0, \ldots, n, \\
&\qquad p_j \geq 0.
\end{aligned}
$$

- If no type-A arbitrage, optimal value in primal and dual must be 0, so dual has a feasible solution p^* (that is also optimal)

Asset Pricing and No-Arbitrage

Theorem (Asset Pricing Theorem)

A risk-neutral probability measure exists **if and only if** *there is no arbitrage.*

Proof. Consider the dual of this LP.

$$
\begin{aligned}
\max_{p}\ & 0 \\
\text{s.t. } & \sum_{j=1}^{m} p_j \cdot S_i^f(\omega_j) = S_i^c,\ i = 0, \ldots, n, \\
& p_j \geq 0.
\end{aligned}
$$

- If no type-A arbitrage, optimal value in primal and dual must be 0, so dual has a feasible solution p^* (that is also optimal)
- No type-B arbitrage means $\sum_{i=0}^{n} S_i^f(\omega_j) \cdot x_i^* = 0$, for $j = 1, \ldots, m$. Because dual is standard-form LP, Theorem 19 (**strict** complem. slack.) implies $\exists p^* : p^* > 0$.

Asset Pricing and No-Arbitrage

Theorem (Asset Pricing Theorem)

A risk-neutral probability measure exists **if and only if** *there is no arbitrage.*

Proof. Consider the dual of this LP.

$$\max_{p}\ 0$$

$$\text{s.t. } \sum_{j=1}^{m} p_j \cdot S_i^f(\omega_j) = S_i^c,\ i = 0, \ldots, n,$$

$$p_j \geq 0.$$

- If no type-A arbitrage, optimal value in primal and dual must be 0, so dual has a feasible solution p^* (that is also optimal)
- No type-B arbitrage means $\sum_{i=0}^{n} S_i^f(\omega_j) \cdot x_i^* = 0$, for $j = 1, \ldots, m$. Because dual is standard-form LP, Theorem 19 (**strict** complem. slack.) implies $\exists p^* : p^* > 0$.
- Dual constraint for $i = 0$ implies $\sum_{j=1}^{m} p_j^* = \frac{1}{R}$, so taking $p^* \cdot R$ yields a RNPM.

The converse direction is proved in an identical manner. ■

Network Revenue Management

- Airline revenue management ("yield management"): setting **booking limits** to control how many tickets of each type are sold

Network Revenue Management

- Airline revenue management ("yield management"): setting **booking limits** to control how many tickets of each type are sold
- Airline is planning operations for a specific day in the future

Network Revenue Management

- Airline revenue management ("yield management"): setting **booking limits** to control how many tickets of each type are sold
- Airline is planning operations for a specific day in the future
- Airline operates a set F of direct flights in its (hub-and-spoke) network

Network Revenue Management

- Airline revenue management ("yield management"): setting **booking limits** to control how many tickets of each type are sold
- Airline is planning operations for a specific day in the future
- Airline operates a set F of direct flights in its (hub-and-spoke) network
- For each flight leg $f \in F$, we know the capacity of the aircraft c_f

Network Revenue Management

- Airline revenue management ("yield management"): setting **booking limits** to control how many tickets of each type are sold
- Airline is planning operations for a specific day in the future
- Airline operates a set F of direct flights in its (hub-and-spoke) network
- For each flight leg $f \in F$, we know the capacity of the aircraft c_f
- The airline can offer a large number of "products" (i.e., itineraries) I:
  - each itinerary refers to an origin-destination-fare class combination
  - each itinerary i has a price r_i that is fixed
  - for each itinerary, the airline estimates the demand d_i
  - each itinerary requires a seat on several flight legs operated by the airline

Network Revenue Management

- Airline revenue management ("yield management"): setting **booking limits** to control how many tickets of each type are sold
- Airline is planning operations for a specific day in the future
- Airline operates a set F of direct flights in its (hub-and-spoke) network
- For each flight leg $f \in F$, we know the capacity of the aircraft c_f
- The airline can offer a large number of "products" (i.e., itineraries) I:
  - each itinerary refers to an origin-destination-fare class combination
  - each itinerary i has a price r_i that is fixed
  - for each itinerary, the airline estimates the demand d_i
  - each itinerary requires a seat on several flight legs operated by the airline
- Requirements: $A \in \{0,1\}^{F \cdot I}$ with $A_{f,i} = 1 \Leftrightarrow$ itinerary i needs seat on flight leg f

Resource matrix A :

	Itinerary 1	Itinerary 2	...	Itinerary $\|I\|$
Flight leg 1	1	0	...	1
Flight leg 2	0	1	...	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Flight leg $\|F\|$	1	1	...	0

Network Revenue Management

- Airline revenue management ("yield management"): setting **booking limits** to control how many tickets of each type are sold
- Airline is planning operations for a specific day in the future
- Airline operates a set F of direct flights in its (hub-and-spoke) network
- For each flight leg $f \in F$, we know the capacity of the aircraft c_f
- The airline can offer a large number of "products" (i.e., itineraries) I:
  - each itinerary refers to an origin-destination-fare class combination
  - each itinerary i has a price r_i that is fixed
  - for each itinerary, the airline estimates the demand d_i
  - each itinerary requires a seat on several flight legs operated by the airline
- Requirements: $A \in \{0, 1\}^{F \cdot I}$ with $A_{f,i} = 1 \Leftrightarrow$ itinerary i needs seat on flight leg f

Resource matrix A :

	Itinerary 1	Itinerary 2	...	Itinerary $\|I\|$
Flight leg 1	1	0	...	1
Flight leg 2	0	1	...	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Flight leg $\|F\|$	1	1	...	0

- Goal: decide how many itineraries of each type to sell to maximize revenue

Network Revenue Management

- Let x_i denote the number of itineraries of type i that the airline plans to sell, and let x be the vector with components x_i

Network Revenue Management

- Let x_i denote the number of itineraries of type i that the airline plans to sell, and let x be the vector with components x_i
- The problem can be formulated as follows:

$$\max_{x \in \mathbb{R}^I} \left\{ r^\intercal x : Ax \leq c,\ x \leq d \right\}$$

Network Revenue Management

- Let x_i denote the number of itineraries of type i that the airline plans to sell, and let x be the vector with components x_i
- The problem can be formulated as follows:

$$\max_{x \in \mathbb{R}^I} \left\{ r^\intercal x : Ax \leq c,\ x \leq d \right\}$$

- $Ax \leq c$ capture the constraints on plane capacity
- $x \leq d$ states that the planned sales cannot exceed the demand
- In practice, an approach that includes **all possible itineraries** encounters challenges

Network Revenue Management

- Let x_i denote the number of itineraries of type i that the airline plans to sell, and let x be the vector with components x_i
- The problem can be formulated as follows:

$$\max_{x \in \mathbb{R}^I} \left\{ r^\intercal x : Ax \leq c,\ x \leq d \right\}$$

- $Ax \leq c$ capture the constraints on plane capacity
- $x \leq d$ states that the planned sales cannot exceed the demand
- In practice, an approach that includes **all possible itineraries** encounters challenges
  - gargantuan LP
  - poor demand estimates for some itineraries

Network Revenue Management

- Let x_i denote the number of itineraries of type i that the airline plans to sell, and let x be the vector with components x_i
- The problem can be formulated as follows:

$$\max_{x \in \mathbb{R}^I} \left\{ r^\intercal x : Ax \le c,\ x \le d \right\}$$

- $Ax \le c$ capture the constraints on plane capacity
- $x \le d$ states that the planned sales cannot exceed the demand
- In practice, an approach that includes **all possible itineraries** encounters challenges
  - gargantuan LP
  - poor demand estimates for some itineraries
- To sell "exotic itineraries", use the **shadow prices for the capacity constraints**

Network Revenue Management

- Let x_i denote the number of itineraries of type i that the airline plans to sell, and let x be the vector with components x_i
- The problem can be formulated as follows:

$$\max_{x \in \mathbb{R}^I} \left\{ r^\intercal x : Ax \leq c,\ x \leq d \right\}$$

- $Ax \leq c$ capture the constraints on plane capacity
- $x \leq d$ states that the planned sales cannot exceed the demand
- In practice, an approach that includes **all possible itineraries** encounters challenges
  - gargantuan LP
  - poor demand estimates for some itineraries
- To sell "exotic itineraries", use the **shadow prices for the capacity constraints**
  - $p \in \mathbb{R}^F$: dual variables for capacity constraints $Ax \leq c$
  - At optimality, p_f is marginal revenue lost if airline loses one seat on flight f

Network Revenue Management

- Let x_i denote the number of itineraries of type i that the airline plans to sell, and let x be the vector with components x_i
- The problem can be formulated as follows:

$$\max_{x \in \mathbb{R}^I} \left\{ r^\intercal x : Ax \leq c,\ x \leq d \right\}$$

- $Ax \leq c$ capture the constraints on plane capacity
- $x \leq d$ states that the planned sales cannot exceed the demand
- In practice, an approach that includes **all possible itineraries** encounters challenges
  - gargantuan LP
  - poor demand estimates for some itineraries
- To sell "exotic itineraries", use the **shadow prices for the capacity constraints**
  - $p \in \mathbb{R}^F$: dual variables for capacity constraints $Ax \leq c$
  - At optimality, p_f is marginal revenue lost if airline loses one seat on flight f
  - For an "exotic" itinerary that requires seats on several flights $f \in E$, the **minimum price** to charge is given by the sum of the shadow prices, $\sum_{f \in E} p_f$

Network Revenue Management

- Let x_i denote the number of itineraries of type i that the airline plans to sell, and let x be the vector with components x_i
- The problem can be formulated as follows:

$$\max_{x \in \mathbb{R}^I} \left\{ r^\intercal x : Ax \leq c,\ x \leq d \right\}$$

- $Ax \leq c$ capture the constraints on plane capacity
- $x \leq d$ states that the planned sales cannot exceed the demand
- In practice, an approach that includes **all possible itineraries** encounters challenges
  - gargantuan LP
  - poor demand estimates for some itineraries
- To sell "exotic itineraries", use the **shadow prices for the capacity constraints**
  - $p \in \mathbb{R}^F$: dual variables for capacity constraints $Ax \leq c$
  - At optimality, p_f is marginal revenue lost if airline loses one seat on flight f
  - For an "exotic" itinerary that requires seats on several flights $f \in E$, the **minimum price** to charge is given by the sum of the shadow prices, $\sum_{f \in E} p_f$
- **Bid-price heuristic** in network revenue management
- Broader principle of how to price "products" through resource usage/cost