

Optimization ▸ Duality

Lecture 13

- Background & Motivation
- Certificates for LPs
- Duality
- Applications of Duality





Optimization ▸ Duality

Lecture 13

- Background & Motivation
- Certificates for LPs
- Duality
- Applications of Duality

Recap: Linear Programming Framework

Linear Programs (LPs):

- Variables with linear objective function
- Linear constraints (inequalities or equalities)
- Polynomial-time solvable

How we've used LPs:

- Integer Program $\rightarrow$ LP Relaxation $\rightarrow$ $\text{OPT}_{\text{LP}} \leq \text{OPT}$
- Round LP solution to integral solution
- Compare to OPT_{LP} as certificate

Today's Lecture: A Different Focus

Today is a detour from approximation algorithms

Focus: An important theoretical property of LPs

- Not directly about algorithms or approximation
- Fundamental to optimization theory
- Beautiful mathematical structure

Why study this?

- Provides deeper understanding of LPs
- Connects to many areas of computer science and math
- **In your PSet:** Applications to approximation and algorithms





Optimization ▸ Duality

Lecture 13

- Background & Motivation
- Certificates for LPs
- Duality
- Applications of Duality

Certificates and Witnesses of Optimality

How do we prove a solution is optimal?

A recurring theme in algorithms: Finding certificates or witnesses that prove solution quality

Recall the recipe from approximation algorithms (minimization):

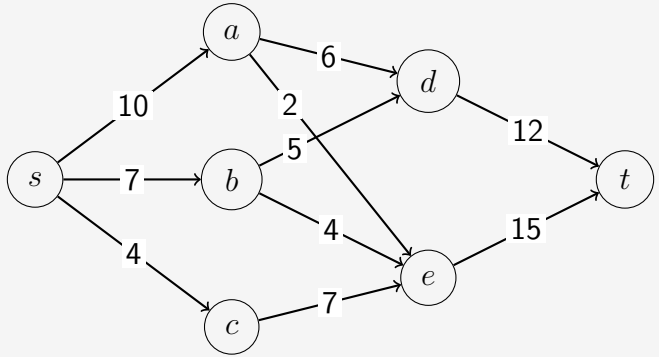
1. Find a certificate $L(I) \leq \mathrm{OPT}(I)$
2. Design an algorithm with $A(I) \leq \alpha L(I)$
3. Conclude $A(I) \leq \alpha L(I) \leq \alpha\,\mathrm{OPT}(I)$

For example, $L(I) = \mathrm{OPT}_{\mathsf{LP}}$, the optimum of an LP relaxation.

Today: We'll see how *linear programs* naturally come with ways to provide such certificates through duality

Example: Max Flow

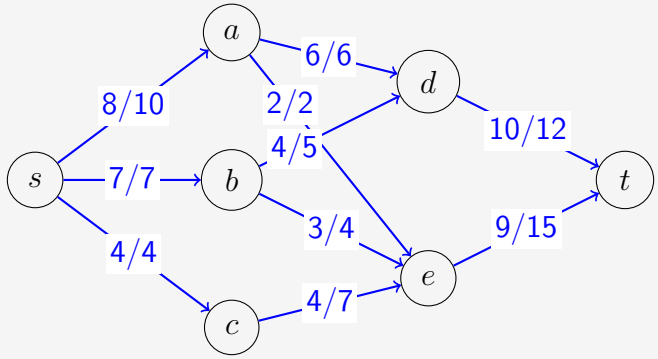
Setup: Consider the following network with capacity constraints



Suppose someone tells you the max flow is 19. How can you provide a certificate that proves this is correct?

Certifying Max Flow Value

One direction is easy: Provide a valid flow with value 19

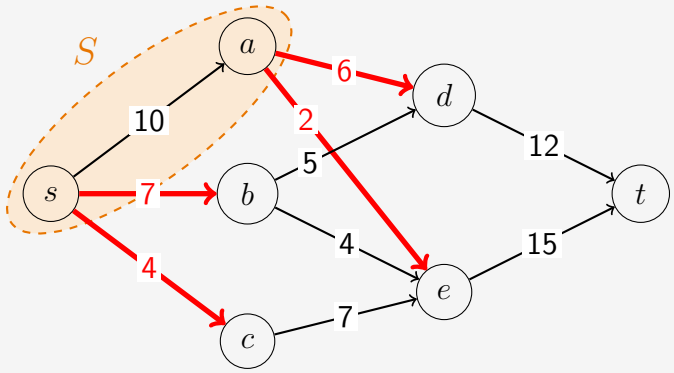


This proves max flow ≥ 19. **But how do we know it's optimal?**

Maybe there's a better flow we haven't found!

The Certificate: An s-t Cut

To prove optimality: Also provide an s-t cut of capacity 19



Cut $S = \{s, a\}$, $T = \{b, c, d, e, t\}$: Capacity $= (s \to b) + (s \to c) + (a \to d) + (a \to e) = 7 + 4 + 6 + 2 = 19$

Weak Max-Flow Min-Cut: Any flow value $\leq$ any cut capacity
Therefore, flow with value 19 and cut with capacity 19 $\implies$ both optimal! ✓

The Key Insight

What just happened:

- We had a maximization problem (max flow)
- We found a certificate by solving a related minimization problem (min cut)
- When their values match, both solutions are optimal!

This is the essence of duality:

- A maximization LP (max flow) has a related minimization LP (min cut)
- Feasible solutions to one provide bounds on the other
- When values match, we have certificates of optimality

Today: We'll see this works for *all* linear programs!

Example: Smoothies

Let's see this certificate idea with linear programs

Setup: You have a stock of fruits and can make three types of smoothies

	Type 1 *Orange–Banana*	**Type 2** *Strawberry–Banana*	**Type 3** *Citro–Bananaberry*
Ingredients	1 Orange 2 Bananas	1 Banana 2 Strawberries	1 Orange 1 Banana 1 Strawberry
Profit	\$5/Liter	\$4/Liter	\$6/Liter

Your stock: 3 oranges, 10 bananas, 6 strawberries

Remark

You can make fractions of liters, but no negative amounts.

Smoothie LP Formulation

Variables: x_i = liters of Smoothie Type i to make

Maximize: $5x_1 + 4x_2 + 6x_3$ (total profit)

Subject to:

$$
\begin{aligned}
x_1 + x_3 &\leq 3 \quad \text{(oranges)} \\
2x_1 + x_2 + x_3 &\leq 10 \quad \text{(bananas)} \\
2x_2 + x_3 &\leq 6 \quad \text{(strawberries)} \\
x_1, x_2, x_3 &\geq 0
\end{aligned}
$$

A Proposed Solution

Claim: The optimal solution is $x_1 = 3$, $x_2 = 3$, $x_3 = 0$

Check feasibility:

- Oranges: $3 + 0 = 3 \leq 3$
- Bananas: $2(3) + 3 + 0 = 9 \leq 10$
- Strawberries: $2(3) + 0 = 6 \leq 6$
- Non-negativity: $x_1, x_2, x_3 \geq 0$

Profit: $5(3) + 4(3) + 6(0) = 15 + 12 = \$27$

This solution is feasible and gives profit \$27. But is it *optimal*?
How can we provide a certificate that no solution can do better?

Finding an Upper Bound Certificate

Goal: Show that *any* feasible solution has profit $\leq \$27$

Idea: Combine the constraints to get an upper bound on the objective

Recall the constraints:

$$\begin{aligned} x_1 + x_3 &\leq 3 \quad \text{(oranges)} \\ 2x_1 + x_2 + x_3 &\leq 10 \quad \text{(bananas)} \\ 2x_2 + x_3 &\leq 6 \quad \text{(strawberries)} \end{aligned}$$

Can we combine these to get something that looks like $5x_1 + 4x_2 + 6x_3 \leq$ some value?

If we can show $5x_1 + 4x_2 + 6x_3 \leq 27$ for *all* feasible solutions, then we're done!

Combining Constraints

Let's try multiplying constraints by some non-negative values:

Take the orange and strawberry constraints:

$$\begin{aligned} 5 \times [x_1 + x_3 \leq 3] &\implies 5x_1 + 5x_3 \leq 15 \\ 2 \times [2x_2 + x_3 \leq 6] &\implies 4x_2 + 2x_3 \leq 12 \end{aligned}$$

Adding these up:

$$5x_1 + 4x_2 + 7x_3 \leq 27$$

Since $x_3 \geq 0$ for any feasible solution:

$$\text{Profit} = 5x_1 + 4x_2 + 6x_3 \leq 5x_1 + 4x_2 + 7x_3 \leq 27$$

Certificate of Optimality

What did we just prove?

By multiplying the constraints by $5, 0, 2$ and adding them up:

- We got $5x_1 + 4x_2 + 7x_3 \leq 27$
- Since $x_3 \geq 0$, we have $5x_1 + 4x_2 + 6x_3 \leq 27$
- Therefore: **Every** feasible solution has profit ≤ 27

We also have: A feasible solution with profit exactly 27

Conclusion: The solution $x_1 = 3, x_2 = 3, x_3 = 0$ is optimal!
The multipliers $(5, 0, 2)$ provide a certificate of optimality

Natural question: How did we know to use $5, 0, 2$? Can we find these systematically?

Finding the Best Multipliers

Multiply the orange, banana, and strawberry constraints by $y_1, y_2, y_3 \geq 0$ and add them up:

$$(y_1 + 2y_2)\, x_1 + (y_2 + 2y_3)\, x_2 + (y_1 + y_2 + y_3)\, x_3 \leq 3y_1 + 10y_2 + 6y_3$$

Since $x_1, x_2, x_3 \geq 0$, this bounds the profit $5x_1 + 4x_2 + 6x_3$ whenever each coefficient is large enough:

$$y_1 + 2y_2 \geq 5, \qquad y_2 + 2y_3 \geq 4, \qquad y_1 + y_2 + y_3 \geq 6$$

Key observation: Finding the best (smallest) bound is itself an LP!
Minimize: $3y_1 + 10y_2 + 6y_3$ subject to the constraints above and $y_i \geq 0$

This is the dual LP. Our multipliers $(5, 0, 2)$ are feasible, with value 27.



Optimization ▸ Duality

Lecture 13

- Background & Motivation
- Certificates for LPs
- Duality
- Applications of Duality

The Primal LP

Consider a general linear program:

Maximize: $\sum_j c_j x_j$

Subject to:

$$\sum_j a_{i,j} x_j \leq b_i \quad \forall i$$

$$x_j \geq 0 \quad \forall j$$

Notation: We call this the primal LP

- Variables: x_j (indexed by j)
- Constraints: indexed by i
- Coefficients: $a_{i,j}$ (constraint i, variable j), b_i, c_j

Goal: Find $\vec{x}$ that maximizes the objective subject to feasibility

Constructing the Dual: Motivation

Recall: We want an upper bound on the objective value

Key idea: Take a nonnegative weight $y_i \geq 0$ for each constraint i and form a linear combination:

$$\sum_i y_i \Big(\sum_j a_{i,j} x_j \Big) \leq \sum_i y_i b_i$$

This gives us an upper bound if we choose y_i such that:

$$\sum_i a_{i,j} y_i \geq c_j \quad \forall j$$

Why? Since $x_j \geq 0$ for all j, we can bound the objective:

$$\begin{aligned} \sum_j c_j x_j &\leq \sum_j \Big(\sum_i a_{i,j} y_i \Big) x_j \\ &= \sum_i y_i \Big(\sum_j a_{i,j} x_j \Big) \leq \sum_i y_i b_i \end{aligned}$$

The Dual LP

To get the tightest upper bound, we minimize over all valid weights:

Minimize: $\sum_i b_i y_i$

Subject to:

$$\sum_i a_{i,j} y_i \geq c_j \quad \forall j$$

$$y_i \geq 0 \quad \forall i$$

This is the dual LP!

Key observation:

- j indexes primal variables and dual constraints
- i indexes primal constraints and dual variables
- Taking the dual "transposes" the LP structure

Weak Duality Theorem

Theorem: Weak Duality

If $\vec{x}$ is feasible for the primal and $\vec{y}$ is feasible for the dual, then $\sum_j c_j x_j \leq \sum_i b_i y_i$.

Proof

$$\begin{aligned}
\sum_j c_j x_j &\leq \sum_j \Big(\sum_i a_{i,j} y_i \Big) x_j \quad \text{(dual feasibility, } x_j \geq 0\text{)} \\
&= \sum_i y_i \Big(\sum_j a_{i,j} x_j \Big) \quad \text{(rearrange)} \\
&\leq \sum_i y_i b_i \quad \text{(primal feasibility, } y_i \geq 0\text{)} \quad \square
\end{aligned}$$

Key takeaway: Every dual solution provides an upper bound on the primal objective!

Exercise: Dual of Dual is Primal

Challenge: Take the dual LP and construct *its* dual

What you should get: The original primal LP!

Sketch of why:

- Dual is a minimization problem with $\geq$ constraints
- Apply the same construction directly, with the inequalities flipped
- Combine its constraints with weights ≥ 0 to get a *lower* bound on its objective
- Maximize that lower bound
- Get back the original primal!

Remark

This shows primal/dual is a *symmetric* relationship. Neither is more "fundamental" than the other.

Strong Duality Theorem

Weak duality says: primal value $\leq$ dual value (when both feasible)

Question: Are they *equal* at optimality?

Theorem: Strong Duality

If the primal LP has an optimal solution $\vec{x}^*$, then the dual LP also has an optimal solution $\vec{y}^*$, and their optimal values are equal:

$$\sum_j c_j x_j^* = \sum_i b_i y_i^*$$

Remark

We won't prove this theorem. The proof is a bit long and uses some theory of LPs. See an optimization course (like ORF 307) for details.

Consequence: Optimal dual solutions are *certificates* of primal optimality!





Optimization ▶ Duality

Lecture 13

- Background & Motivation
- Certificates for LPs
- Duality
- Applications of Duality

Application: Matching and Vertex Cover

Setup: An undirected graph $G = (V, E)$

Definitions:

- Matching $=$ set of edges with no shared endpoints
- Vertex cover $=$ set of vertices touching all edges

Recall (Lecture 11): For every matching M and every vertex cover S, $|M| \leq |S|$. In particular,

$$\text{Max matching size} \leq \text{Min vertex cover size}$$

Today: We'll see this inequality again, this time through LP duality!

Matching IP and its LP Relaxation

Integer Program for Maximum Matching ($x_e = 1$ means e is in the matching)

$$\textbf{Maximize:} \quad \sum_{e \in E} x_e$$

$$\textbf{Subject to:} \quad \sum_{e \text{ incident to } v} x_e \leq 1 \quad \forall v \in V$$

$$x_e \in \{0, 1\} \quad \forall e \in E$$

LP Relaxation: Replace $x_e \in \{0, 1\}$ with $0 \leq x_e \leq 1$

As usual: $\text{OPT}_{\text{IP}} \leq \text{OPT}_{\text{LP}}$, so the LP relaxation gives an **upper bound** on the maximum matching size.

Vertex Cover IP and its LP Relaxation

Integer Program for Minimum Vertex Cover ($y_v = 1$ means v is in the cover)

$$\textbf{Minimize:} \quad \sum_{v \in V} y_v$$

$$\textbf{Subject to:} \quad y_u + y_v \geq 1 \quad \forall (u, v) \in E$$

$$y_v \in \{0, 1\} \quad \forall v \in V$$

LP Relaxation: Replace $y_v \in \{0, 1\}$ with $0 \leq y_v \leq 1$

As usual: $\text{OPT}_{\text{LP}} \leq \text{OPT}_{\text{IP}}$, so the LP relaxation gives a **lower bound** on the minimum vertex cover size.

The Two LPs are Dual!

Matching LP (Primal):

- **Maximize:** $\sum_{e \in E} x_e$
- **Subject to:** $\sum_{e \text{ incident to } v} x_e \leq 1$ for all v, and $x_e \geq 0$

Vertex Cover LP (Dual):

- **Minimize:** $\sum_{v \in V} y_v$
- **Subject to:** $y_u + y_v \geq 1$ for all edges (u, v), and $y_v \geq 0$

Check: These are dual LPs!

- Matching variables $\leftrightarrow$ Vertex cover constraints (one per edge)
- Matching constraints $\leftrightarrow$ Vertex cover variables (one per vertex)

Deriving the Dual from Matching LP

Matching LP:

$$\text{Maximize:} \quad \sum_{e \in E} x_e$$

$$\text{Subject to:} \quad \sum_{e \text{ incident to } v} x_e \leq 1 \quad \forall v \in V, \qquad x_e \geq 0 \quad \forall e \in E$$

Combine: Multiply the constraint of each vertex v by $y_v \geq 0$ and add up:

$$\sum_{v \in V} y_v \Big(\sum_{e \text{ incident to } v} x_e \Big) \leq \sum_{v \in V} y_v$$

Regroup by edges: Each edge $e = (u, v)$ appears twice, once for u and once for v:

$$\sum_{v \in V} y_v \Big(\sum_{e \text{ incident to } v} x_e \Big) = \sum_{e=(u,v) \in E} (y_u + y_v)\, x_e$$

Since $x_e \geq 0$, this bounds the objective $\sum_e x_e$ when $y_u + y_v \geq 1$ for every edge (u, v). Minimizing the bound $\sum_v y_v$ gives the **Vertex Cover LP**!

Duality Connects the Two Problems

By Strong Duality:

$$\text{Matching LP optimum} = \text{Vertex Cover LP optimum}$$

Combining with IP-LP relationships:

$$\begin{aligned} \text{Max Matching (IP)} &\leq \text{Matching LP} \\ &= \text{Vertex Cover LP} \quad \text{(Strong Duality)} \\ &\leq \text{Min Vertex Cover (IP)} \end{aligned}$$

Conclusion: Max matching $\leq$ min vertex cover in every graph: the same inequality as Lecture 11, now as a consequence of duality!

The inequalities can be strict: in a triangle K_3, max matching $= 1$, both LP optima $= 3/2$ (all variables $1/2$), and min vertex cover $= 2$.