Analysis of Algorithms, I

CSOR W4231

Eleni Drinea
Computer Science Department

Columbia University

Linear programming

Outline

Today

1. The structure of a linear program
2. Duality
3. Examples

Why linear programming?

1. Vast range of applications
   - Resource allocation
   - Production planning
   - Military strategy forming
   - Graph theoretic problems
   - Error correction
   - ...
2. Establish the existence of polynomial-time (efficient) algorithms
3. Guide the design of approximation algorithms for computationally hard problems *(coming up in the next two weeks)*
4. Duality provides a unifying view of seemingly unrelated results and is useful in algorithm design

Today

An introductory example: profit maximization

A boutique chocolatier has two products:

- an assortment of chocolates
- an assortment of truffles

Their profit is

1. \$1 per box of chocolates
2. \$6 per box of truffles

They can produce a total of at most 400 boxes per day.
The daily demand for these products is limited

1. at most 200 boxes of chocolates per day
2. at most 300 boxes of truffles per day

What are the optimal levels of production?

The linear program (LP) for profit maximization

1. Let x_1 be the number of boxes of chocolates and x_2 the number of boxes of truffles produced daily.
2. **Objective:** maximize $x_1 + 6x_2$
3. **Constraints** on x_1, x_2: $0 \le x_1 \le 200$, $0 \le x_2 \le 300$, $x_1 + x_2 \le 400$

Linear program for chocolatier's profit

$$\begin{aligned}
\max_{x_1 \ge 0, x_2 \ge 0} \quad & x_1 + 6x_2 \\
\text{subject to} \quad & x_1 \le 200 \\
& x_2 \le 300 \\
& x_1 + x_2 \le 400
\end{aligned}$$

The general problem

Input

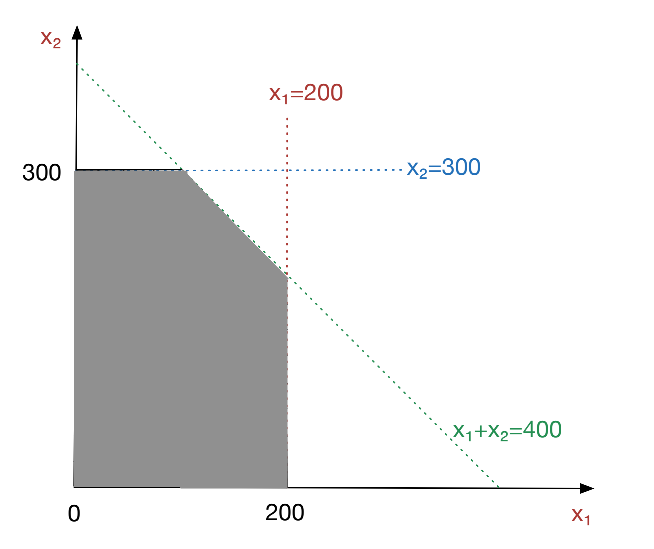
1. a set of **variables**
2. a set of *linear* **constraints** on the variables (equalities or inequalities)
3. a *linear* **objective function** to maximize (or minimize)

Output

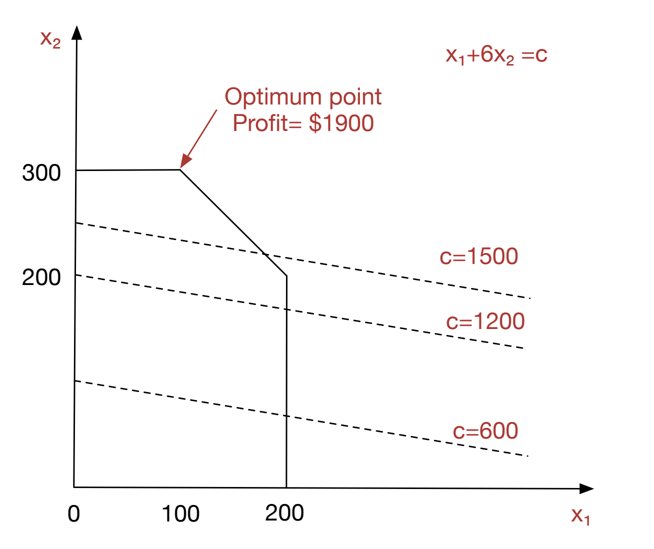
- an assignment of real values to the variables such that
  1. the constraints are satisfied;
  2. the objective function is maximized (or minimized)

A more succinct linear algebraic formulation of a linear program will appear in a later slide.

The geometry of the solution: feasible region



The geometry of the solution: objective function



The geometry of the solution in words

- The set of all feasible solutions is the set of points in the (x_1, x_2) plane that satisfy **all** five constraints.
- A linear equation in x_1 and x_2 defines a line on that plane.
- A linear inequality defines a half-space on that plane (one side of the line).

$\Rightarrow$ The set of all feasible solutions is the intersection of the five half-spaces.

- **Goal:** Find the point in this polygon that maximizes the objective function (the profit).

Feasible region and optimal solution

Fact 1.

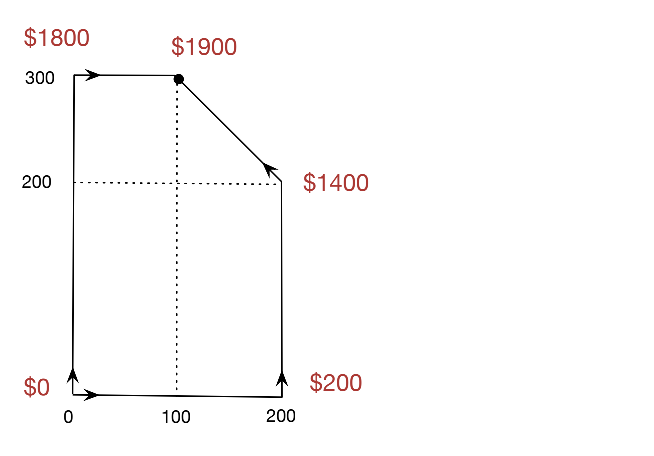
The optimum is achieved at a vertex of the feasible region.

Exceptions

1. The linear program is infeasible
   - e.g., $x \leq 1$, $x \geq 2$
2. The optimum value is unbounded

$$\max_{x_1 \geq 0, x_2 \geq 0} \quad x_1 + x_2$$

Solving LPs: the simplex method [Dantzig1947]



For more information on simplex, take, e.g., Optimization I (IEOR 6613).

Some history on LP solvers

- Simplex method [Dantzig1947]
  - fast in practice
  - exponential worst case performance
- Ellipsoid method [Khachiyan1979]
  - provably polynomial-time algorithm
  - slow in practice
- Interior-point method [Karmarkar]
  - polynomial-time algorithm
  - fast in practice
- Interior-point methods is a field of active research

Today

1. The structure of a linear program
2. Duality
3. Examples

An alternative proof that \$1900 is optimal

Recall the LP from slide 5.

$$
\begin{aligned}
\max_{x_1 \geq 0, x_2 \geq 0} \quad & x_1 + 6x_2 \\
\text{subject to} \quad & x_1 \leq 200 \\
& x_2 \leq 300 \\
& x_1 + x_2 \leq 400
\end{aligned}
$$

Using the constraints to upper bound the objective

- Multiply the first inequality by 0
- Multiply the second inequality by 5
- Multiply the third inequality by 1
- Add the new inequalities; then

$$x_1 + 6x_2 \leq 1900$$

$\Rightarrow$ the objective function cannot exceed 1900!

$\Rightarrow$ thus we indeed found the optimal solution

Where did we get the multipliers $0, 5$ and 1?

Upper bounding the objective function

The constraints themselves can help us derive an upper bound.

Multiplier	Inequality		
y_1	x_1	$\leq$	200
y_2	x_2	$\leq$	300
y_3	$x_1 + x_2$	$\leq$	400

- Multipliers y_i must be non-negative (*why?*)

Add the multiplied inequalities together:

$$y_1 x_1 + y_2 x_2 + y_3(x_1 + x_2) \leq 200y_1 + 300y_2 + 400y_3$$

An upper bound for the objective

We want to upper bound the original objective

$$1x_1 + 6x_2$$

using the linear combination

$$\begin{aligned} & y_1x_1 + y_2x_2 + y_3(x_1 + x_2) \le 200y_1 + 300y_2 + 400y_3 \\ \Rightarrow \quad & (y_1 + y_3)x_1 + (y_2 + y_3)x_2 \le 200y_1 + 300y_2 + 400y_3 \end{aligned} \tag{1}$$

An upper bound for the objective

We want to upper bound the original objective

$$1x_1 + 6x_2$$

using the linear combination

$$\begin{aligned} & y_1x_1 + y_2x_2 + y_3(x_1 + x_2) \leq 200y_1 + 300y_2 + 400y_3 \\ \Rightarrow \quad & (y_1 + y_3)x_1 + (y_2 + y_3)x_2 \leq 200y_1 + 300y_2 + 400y_3 \qquad (1) \end{aligned}$$

Since $x_1, x_2 \geq 0$, if we constrain $y_1 + y_3 \geq 1$ and $y_2 + y_3 \geq 6$, then the right-hand side in (1) is an upper bound for our objective.

The dual LP

- *What is the best possible upper bound for our objective?*
  Minimize equation (1) subject to constraints on y_1, y_2, y_3.
- This is yet another LP!

$$\begin{aligned} \min_{y_1 \geq 0, y_2 \geq 0, y_3 \geq 0} \quad & 200y_1 + 300y_2 + 400y_3 \\ \text{subject to} \quad & y_1 + y_3 \geq 1 \\ & y_2 + y_3 \geq 6 \end{aligned}$$

This new LP is called the dual of the original, which is called the primal.

Weak duality

- By construction, any feasible solution for the dual LP is an upper bound on the original primal LP.
- Let V_P be the optimal objective value for the primal (a maximization)
- Let V_D be the optimal objective value for the dual (a minimization)

Theorem 2 (Weak Duality).

$$V_P \leq V_D$$

The upper bounding strategy can also be used for more general kinds of optimization problems, and weak duality again holds.

Strong duality

- Suppose we found a pair of primal and dual feasible values that are equal.
- Then they must both be optimal.
  - E.g., in our chocolatier's profit maximization problem
    - $(x_1, x_2) = (100, 300)$ is a feasible solution for the primal LP
    - $(y_1, y_2, y_3) = (0, 5, 1)$ is a feasible solution for the dual LP
    - they have the same value of 1900.

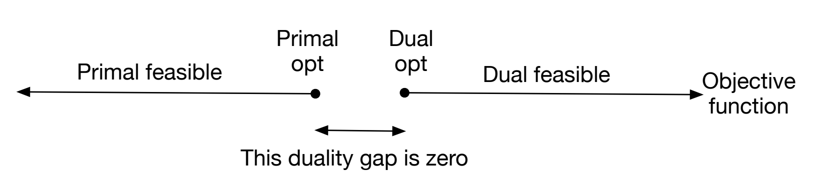
    $\Rightarrow$ Thus these solutions certify each other's optimality.

Amazingly, this is always possible for LPs with bounded optima.

Theorem 3 (Strong Duality).

$$V_P = V_D$$

Strong duality consequences



- We can alternatively solve the dual to find the optimal objective value.
- An optimal dual solution can be used to derive an optimal primal solution (complementary slackness).
- The dual may have structure making it easier to solve at scale (e.g., via parallel optimization).

Today

1. The structure of a linear program
2. Duality
3. Examples

Higher-dimensional LPs

Suppose that the chocolatier introduces a third product, *seasonal* truffles, such that

- *seasonal* truffles yield a profit of \$13 per box
- ≤ 100 boxes of *seasonal* truffles may be produced.

What are the new optimal levels of production?
What if we add a fourth line of production? A hundred-th?

- High-dimensional problem
- Simplex still works!

LP for more products

$$\begin{aligned} \max_{x_1 \geq 0, x_2 \geq 0, x_3 \geq 0} \quad & x_1 + 6x_2 + 13x_3 \\ \text{subject to} \quad & x_1 \leq 200 \\ & x_2 \leq 300 \\ & x_3 \leq 100 \\ & x_1 + x_2 + x_3 \leq 400 \end{aligned}$$

Production planning for a carpet company

- The company has 30 employees.
- Each employee makes 20 carpets per month.
- Monthly employee salary is $2000.
- Initially, no surplus of carpets.

Your data shows that carpet demand is extremely seasonal: monthly demand d_i ranges from 440 to 920. Fluctuations in demand may be handled as follows

1. Overtime
   - they are paid 80% more than regular workers
   - workers can put in at most 30% overtime.
2. Hiring and firing
   - these cost $320 and $400 respectively per worker
3. Storing surplus production
   - costs $8 per month
   - no stored carpets at the end of the year

Goal: minimize yearly expenses for company

Variables for carpet company production planning

- w_i = number of workers during i-th month; $w_0 = 30$
- h_i, f_i = number of workers hired and fired, respectively, at beginning of month i
- x_i = number of carpets made during i-th month
- o_i = number of carpets made by overtime in i-th month
- s_i = number of carpets stored at end of month i; $s_0 = 0$

LP for carpet company production planning

Constraints (one constraint for every month $1 \leq i \leq 12$)

- $w_i, h_i, f_i, x_i, o_i, s_i \geq 0$
- $x_i = 20w_i + o_i$
- $w_i = w_{i-1} + h_i - f_i$
- $s_i = s_{i-1} + x_i - d_i$
- $o_i \leq 6w_i$

Objective:

$$\min \quad 2000 \sum_i w_i + 320 \sum_i h_i + 400 \sum_i f_i + 8 \sum_i s_i + 180 \sum_i o_i$$

What if solution is **not** *integral?*

Today

1. The structure of a linear program
2. Duality
3. Examples

LPs in matrix-vector notation

We may rewrite any LP as follows (*think about it!*).

1. It is either a maximization or a minimization
2. All constraints are inequalities in the same direction
3. All variables are non-negative

This results in an LP of the following form

$$\begin{array}{rl} \max\limits_{\mathbf{x} \geq \mathbf{0}} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & A\mathbf{x} \leq \mathbf{b} \end{array}$$

The dual in matrix-vector notation

Then the dual is given as follows (*prove this!*).

$$\begin{array}{rl} \min\limits_{\mathbf{y}\geq\mathbf{0}} & \mathbf{b}^T\mathbf{y} \\ \text{subject to} & A^T\mathbf{y} \geq \mathbf{c} \end{array}$$

By construction, we know that the any feasible solution to the primal is upper bounded by any feasible solution to the dual (**weak duality**). Hence

$$\mathbf{c}^T\mathbf{x} \leq \mathbf{b}^T\mathbf{y}$$

What if the primal is unbounded?
What if the dual is unbounded?

Feasibility vs Optimality

Finding a feasible solution of a linear program is generally computationally as difficult as finding an optimal solution.

For example, consider the primal in slide 28. Any feasible solution to the following LP (restricted to $\mathbf{x}$) is an optimal solution to the primal.

$$\begin{aligned} \max_{\mathbf{x}\geq\mathbf{0},\mathbf{y}\geq\mathbf{0}} \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} \leq \mathbf{b} \\ & A^T\mathbf{y} \geq \mathbf{c} \\ & \mathbf{c}^T\mathbf{x} \geq \mathbf{b}^T\mathbf{y} \end{aligned}$$

Textbook dualization recipe

Note that $\mathbf{b} \in \mathcal{R}^m$, $\mathbf{c} \in \mathcal{R}^n$, $A \in \mathcal{R}^{m \times n}$

	Primal LP	Dual LP
Variables	$x_1, \ldots, x_n$	$y_1, \ldots, y_m$
Matrix	A	A^T
Right-hand side	$\mathbf{b}$	$\mathbf{c}$
Objective	$\max \mathbf{c}^T \mathbf{x}$	$\min \mathbf{b}^T \mathbf{y}$
Constraints	$x_i \geq 0$	i-th constraint has $\geq$
	$x_i \leq 0$	$\leq$
	$x_i \in \mathcal{R}$	$=$
	j-th constraint has $\leq$	$y_j \geq 0$
	$\geq$	$y_j \leq 0$
	$=$	$y_j \in \mathcal{R}$