CS6501: Topics in Learning and Game Theory (Fall 2019)

Linear Programming

Instructor: Haifeng Xu

Slides of this lecture is adapted from Shaddin Dughmi at
https://www-bcf.usc.edu/~shaddin/cs675sp18/index.html

Outline

- Linear Programing Basics

- Dual Program of LP and Its Properties

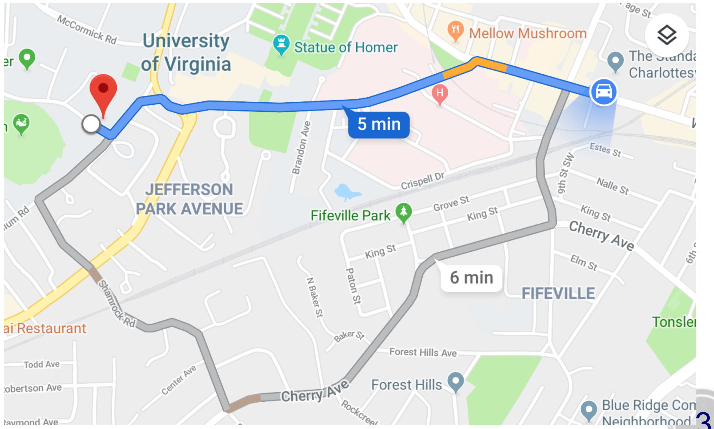
Mathematical Optimization

- The task of selecting the best configuration from a “feasible” set to optimize some objective

$$\text{minimize (or maximize)} \quad f(x)$$
$$\text{subject to} \quad x \in X$$

- x: decision variable
- $f(x)$: objective function
- X: feasible set/region
- Optimal solution, optimal value

- Example 1: minimize x^2, s.t. $x \in [-1,1]$
- Example 2: pick a road to school



Polynomial-Time Solvability

- A problem can be solved in polynomial time if there exists an algorithm that solves the problem in time polynomial in its input size
- Why care about polynomial time? Why not quadratic or linear?
  - There are studies on fined-grained complexity
  - But poly-time vs exponential time seems a fundamental separation between easy and difficult problems
  - In many cases, after a poly-time algorithm is developed, researchers can quickly reduce the polynomial degree to be small (e.g., solving LPs)
- In algorithm analysis, a significant chunk of research is devoted to studying the complexity of a problem by proving it is poly- time solvable or not (e.g., NP-hard problems)

$$\begin{array}{ll} \text{minimize (or maximize)} & f(x) \\ \text{subject to} & x \in X \end{array}$$

- Difficult to solve without any assumptions on $f(x)$ and X
- A ubiquitous and well-understood case is *linear program*

Linear Program (LP) – General Form

$$\begin{array}{lll} \text{minimize (or maximize)} & c^T \cdot x & \\ \text{subject to} & a_i \cdot x \leq b_i & \forall i \in C_1 \\ & a_i \cdot x \geq b_i & \forall i \in C_2 \\ & a_i \cdot x = b_i & \forall i \in C_3 \end{array}$$

- Decision variable: $x \in \mathbb{R}^n$
- Parameters:
  - $c \in \mathbb{R}^n$ define the linear objective
  - $a_i \in \mathbb{R}^n$ and $b_i \in \mathbb{R}$ defines the i'th linear constraint

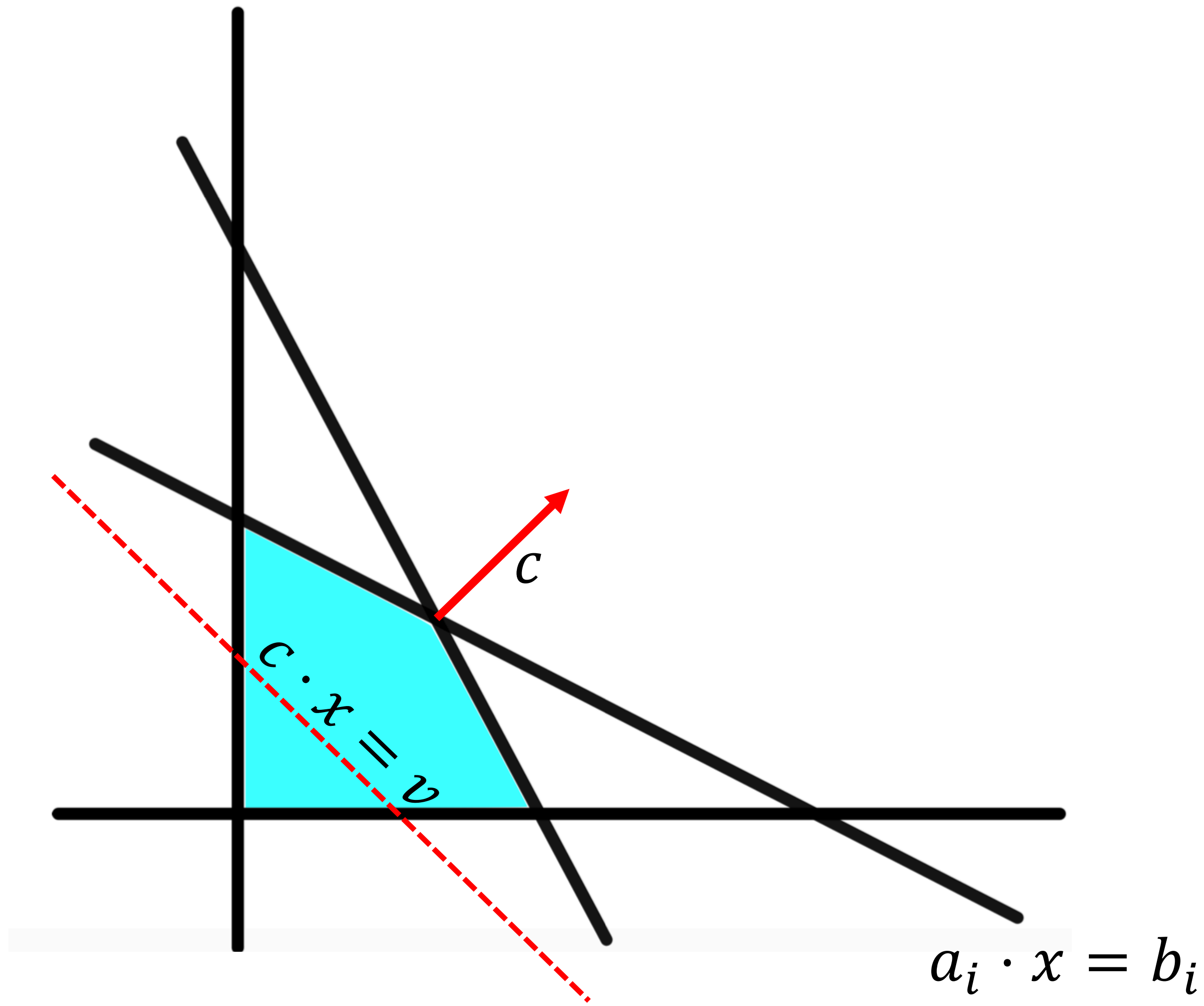
Linear Program (LP) – Standard Form

maximize	$c^T \cdot x$	
subject to	$a_i \cdot x \leq b_i$	$\forall i = 1, \cdots, m$
	$x_j \geq 0$	$\forall j = 1, \cdots, n$

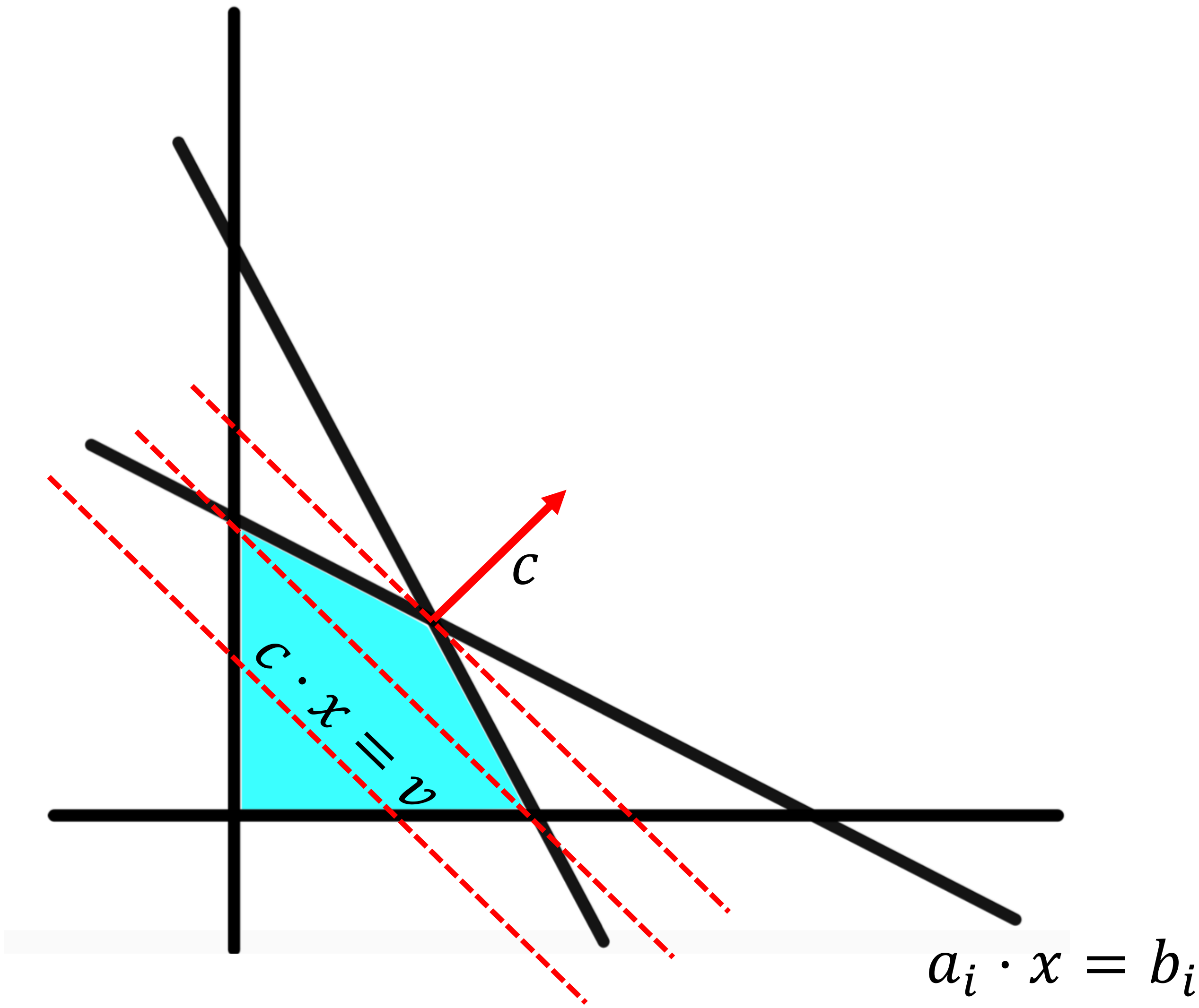
Claim. Every LP can be transformed to an *equivalent* standard form

- minimize $c^T \cdot x \quad \Leftrightarrow \quad$ maximize $-c^T \cdot x$
- $a_i \cdot x \geq b_i \;\Leftrightarrow\; -a_i \cdot x \leq -b_i$
- $a_i \cdot x = b_i \;\Leftrightarrow\; a_i \cdot x \leq b_i$ and $-a_i \cdot x \leq -b_i$
- Any unconstrained x_j can be replaced by $x_j^+ - x_j^-$ with $x_j^+, x_j^- \geq 0$

Geometric Interpretation

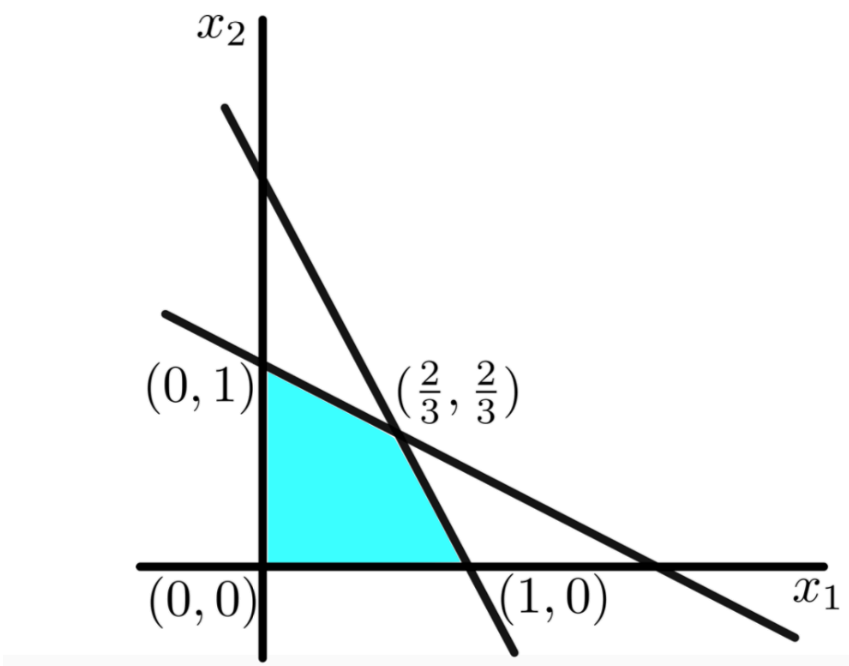


Geometric Interpretation



A 2-D Example

$$\begin{array}{ll} \text{maximize} & x_1 + x_2 \\ \text{subject to} & x_1 + 2x_2 \leq 2 \\ & 2x_1 + x_2 \leq 2 \\ & x_1, x_2 \geq 0 \end{array}$$



Application: Optimal Production

- n products, m raw materials
- Every unit of product j uses a_{ij} units of raw material i
- There are b_i units of material i available
- Product j yields profit c_j per unit
- Factory wants to maximize profit subject to available raw materials

j: product index
i: material index

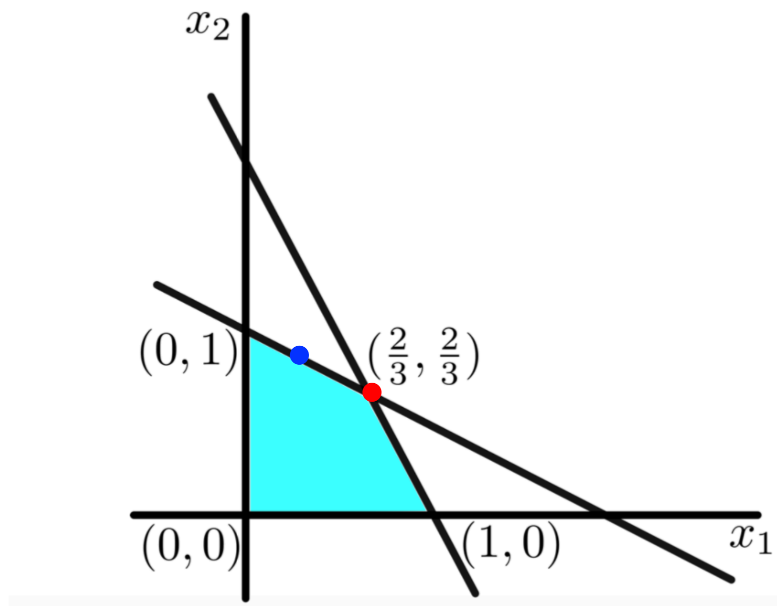
maximize	$c^T \cdot x$	
subject to	$a_i \cdot x \leq b_i$	$\forall i = 1, \cdots, m$
	$x_j \geq 0$	$\forall j = 1, \cdots, n$

where variable x_j = # units of product j

Terminology

- Hyperplane: The region defined by a linear equality $a_i \cdot x = b_i$
- Halfspace: The region defined by a linear inequality $a_i \cdot x \leq b_i$
- Polyhedron: The intersection of a set of linear inequalities
  - Feasible region of an LP is a polyhedron
- Polytope: *Bounded* polyhedron
- Vertex: A point x is a vertex of polyhedron P if $\nexists\, y \neq 0$ with $x + y \in P$ and $x - y \in P$

Red point: vertex
Blue point: not a vertex

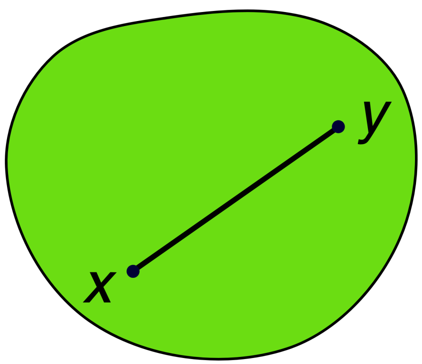


Terminology

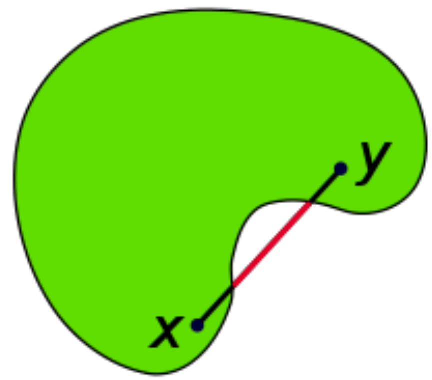
Convex set: A set S is convex if $\forall x, y \in S$ and $\forall p \in [0,1]$, we have

$$p \cdot x + (1-p) \cdot y \in S$$

- Inherently related to convex functions



convex



Non-convex

Terminology

Convex set: A set S is convex if $\forall x, y \in S$ and $\forall p \in [0,1]$, we have

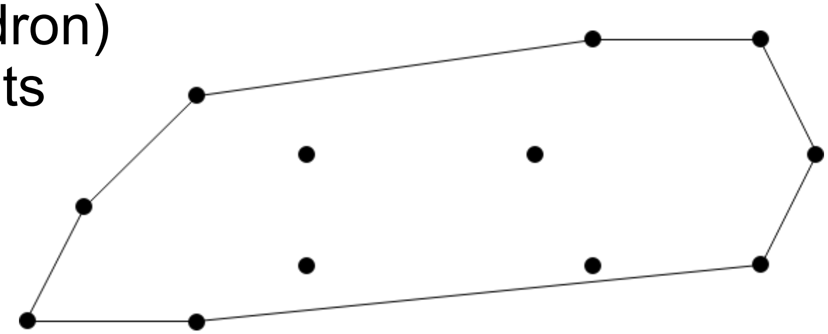
$$p \cdot x + (1-p) \cdot y \in S$$

Convex hull: the convex hull of points $\mathrm{x}_1, \cdots, x_m \in \mathbb{R}$ is

$$\mathrm{convhull}(x_1, \cdots, x_n) = \left\{ \mathrm{x} = \sum_{i=1}^{n} p_i x_i : \forall p \in \mathbb{R}_+^n \;\; s.t. \;\; \sum p_i = 1 \right\}$$

That is, $\mathrm{convhull}(x_1, \cdots, x_n)$ includes all points that can be written as expectation of $x_1, \cdots, x_n$ under some distribution p.

- Any polytope (i.e., a bounded polyhedron) is the convex hull of a finite set of points

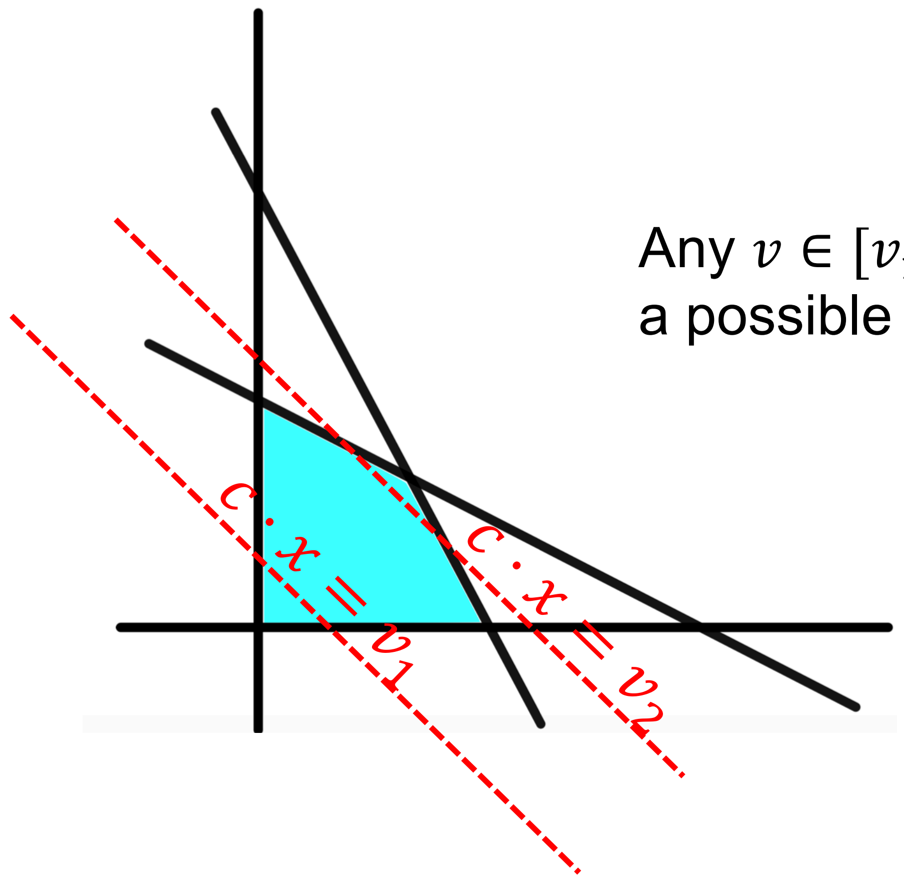


Geometric visualization of convex hull

Basic Facts about LPs and Polyhedrons

Fact: The feasible region of any LP (a polyhedron) is a convex set. All possible objective values form an interval (possibly unbounded).

Note: intervals are the only convex sets in $\mathbb{R}$



Any $v \in [v_1, v_2]$ must also be a possible objective value

Basic Facts about LPs and Polyhedrons

Fact: The feasible region of any LP (a polyhedron) is a convex set. All possible objective values form an interval (possibly unbounded).

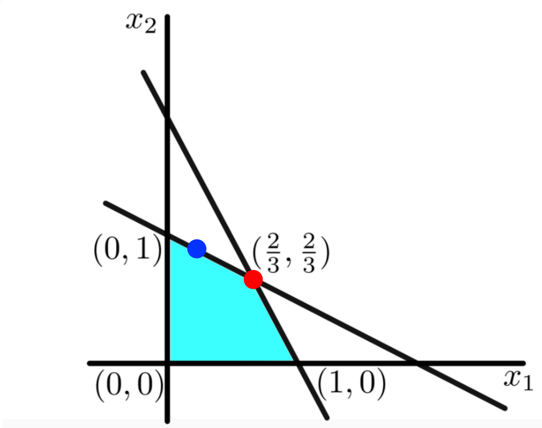
Note: intervals are the only convex sets in $\mathbb{R}$

Fact: The set of optimal solutions of any LP is a convex set.

- It is the intersection of feasible region and hyperplane $c^T \cdot x = OPT$

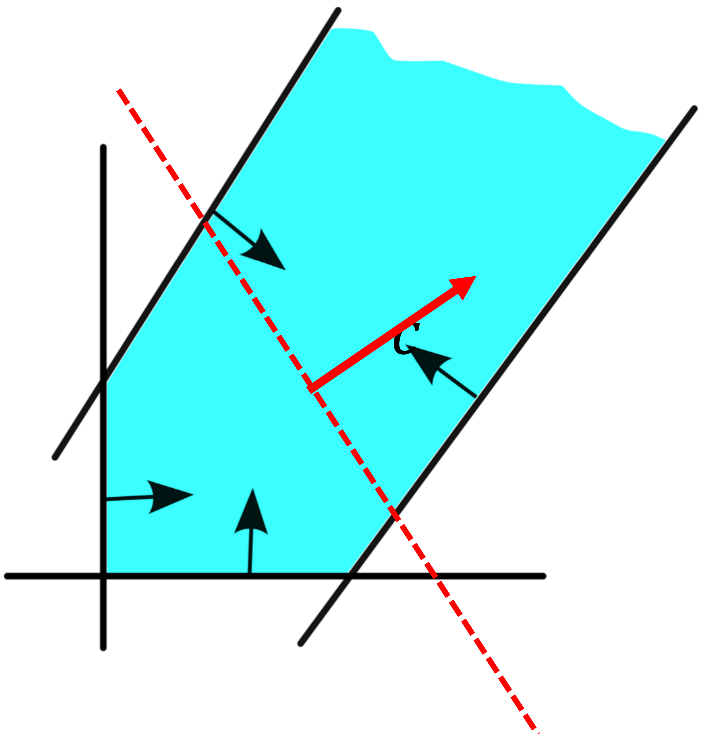
Fact: At a vertex, n linearly independent constraints are satisfied with equality (a.k.a., tight).

Formal proofs: homework exercise



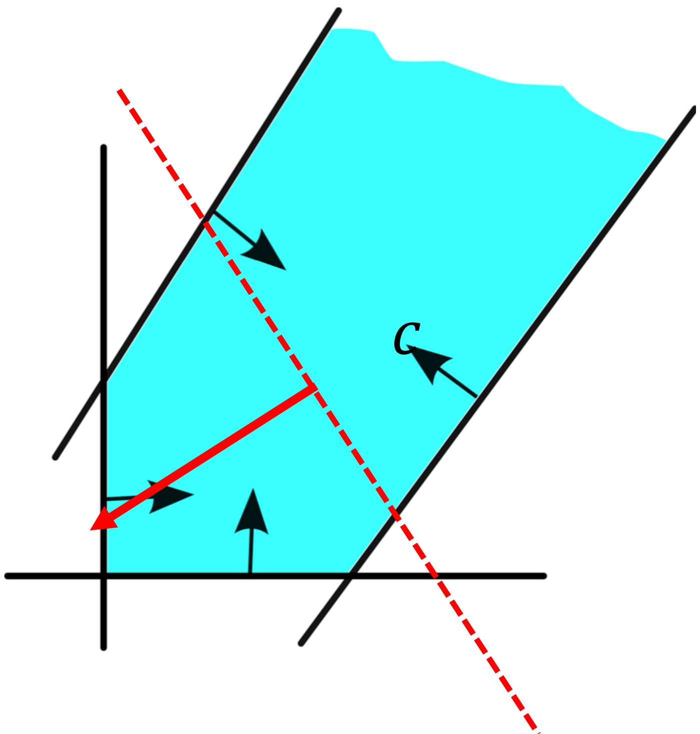
Basic Facts about LPs and Polyhedrons

Fact: An LP either has an optimal solution, or is unbounded or infeasible



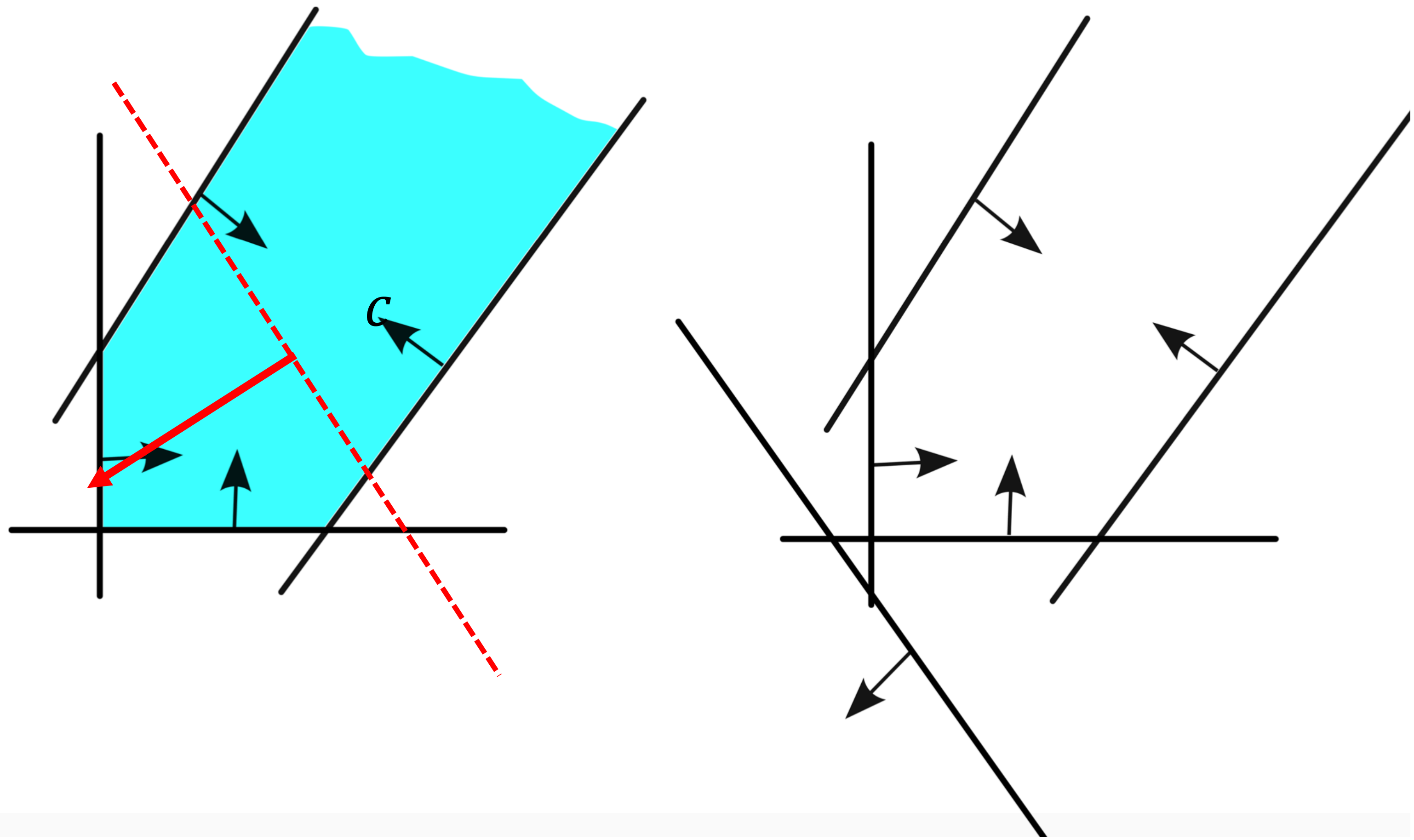
Basic Facts about LPs and Polyhedrons

Fact: An LP either has an optimal solution, or is unbounded or infeasible



Basic Facts about LPs and Polyhedrons

Fact: An LP either has an optimal solution, or is unbounded or infeasible



Fundamental Theorem of LP

Theorem: if an LP in standard form has an optimal solution, then it has a vertex optimal solution.

Proof

- Assume not, and take a non-vertex optimal solution $\bar{x}$ with the maximum number of tight constraints
- There is $y \neq 0$ s.t. $\bar{x} \pm y$ are feasible
- y is orthogonal to objective function and all tight constraints at $\bar{x}$
  - i.e. $c^T \cdot y = 0$, and $a_i^T \cdot y = 0$ whenever the i'th constraint is tight for $\bar{x}$

a) Arguments for $a_i^T \cdot y = 0$
  - $\bar{x} \pm y$ feasible $\Rightarrow a_i^T \cdot (\bar{x} \pm y) \leq b_i$
  - $\bar{x}$ is tight at constraint $i \Rightarrow a_i^T \cdot \bar{x} = b_i$
  - These together yield $a_i^T \cdot (\pm y) \leq 0 \Rightarrow a_i^T \cdot y = 0$

b) Similarly, $\bar{x}$ optimal implies $c^T(\bar{x} \pm y) \leq c^T\bar{x} \Rightarrow c^t y = 0$

Fundamental Theorem of LP

Theorem: if an LP in standard form has an optimal solution, then it has a vertex optimal solution.

Proof

- Assume not, and take a non-vertex optimal solution x with the maximum number of tight constraints
- There is $y \neq 0$ s.t. $x \pm y$ are feasible
- y is orthogonal to objective function and all tight constraints at x
  - i.e. $c^T \cdot y = 0$, and $a_i^T \cdot y = 0$ whenever the i'th constraint is tight for x
- Can choose y s.t. $y_j < 0$ for some j
- Let α be the largest constant such that $x + \alpha y$ is feasible
  - Such an α exists (since $x_j + \alpha y_j < 0$ if α very large)
- An additional constraint becomes tight at $x + \alpha y$, contradiction

Fundamental Theorem of LP

Theorem: if an LP in standard form has an optimal solution, then it has a vertex optimal solution.

Corollary [counting non-zero variables]: If an LP in standard form has an optimal solution, then there is an optimal solution with at most m non-zero variables.

$$\begin{array}{lll} \text{maximize} & c^T \cdot x & \\ \text{subject to} & a_i \cdot x \leq b_i & \forall i = 1, \cdots, m \\ & x_j \geq 0 & \forall j = 1, \cdots, n \end{array}$$

- Meaningful when $m < n$
- E.g. for optimal production with $n = 10$ products and $m = 3$ raw materials, there is an optimal plan using at most 3 products.

Poly-Time Solvability of LP

Theorem: any linear program with n variables and m constraints can be solved in $\mathrm{poly}(m, n)$ time.

- Original proof gives an algorithm with very high polynomial degree
- Now, the fastest algorithm with guarantee takes $\sqrt{\min(n, m)} \cdot T$ where $T =$ time of solving linear equation systems of the same size
- In practice, Simplex Algorithm runs extremely fast though in (extremely rare) worst case it still takes exponential time
- We will not cover these algorithms; Instead, *we use them as building blocks to solve other problems*

Brief History of Linear Optimization

- The forefather of convex optimization problems, and the most ubiquitous.
- Developed by Kantorovich during World War II (1939) for planning the Soviet army's expenditures and returns. Kept secret.
- Discovered a few years later by George Dantzig, who in 1947 developed the simplex method for solving linear programs
- John von Neumann developed LP duality in 1947, and applied it to game theory
- Polynomial-time algorithms: Ellipsoid method (Khachiyan 1979), interior point methods (Karmarkar 1984).

Outline

- Linear Programing Basics
- Dual Program of LP and Its Properties

Dual Linear Program: General Form

Primal LP

$$
\begin{aligned}
\text{max} \quad & c^T \cdot x \\
\text{s.t.} \quad & \\
y_i: \quad & a_i^T x \leq b_i, & \forall i \in C_1 \\
y_i: \quad & a_i^T x = b_i, & \forall i \in C_2 \\
& x_j \geq 0, & \forall j \in D_1 \\
& x_j \in \mathbb{R}, & \forall j \in D_2
\end{aligned}
$$

Dual LP

$$
\begin{aligned}
\text{min} \quad & b^T \cdot y \\
\text{s.t.} \quad & \\
x_j: \quad & \bar{a}_j y \geq c_j, & \forall j \in D_1 \\
x_j: \quad & \bar{a}_j\, y = c_j, & \forall j \in D_2 \\
& y_i \geq 0, & \forall i \in C_1 \\
& y_i \in \mathbb{R}, & \forall i \in C_2
\end{aligned}
$$

- y_i is the dual variable corresponding to primal constraint $a_i^T x \leq (\text{or} =) b_i$
  - Loose constraint (i.e. inequality) ⇒ tight dual variable (i.e. nonnegative)
  - Tight constraint (i.e. equality) ⇒ loose dual variable (i.e. unconstrained)
- $\bar{a}_j y \geq (\text{or} =) c_j$ is the dual constraint corresponding to primal variable x_j
  - Loose variable (i.e. unconstrained) ⇒ tight dual constraint (i.e. equality)
  - Tight variable (i.e. nonnegative) ⇒ loose dual constraint (i.e. inequality)

Dual Linear Program: General Form

Primal LP

max $\quad c^T \cdot x$

s.t.

$$
\begin{aligned}
y_i:&\quad a_i^T x \leq b_i, && \forall i \in C_1\\
y_i:&\quad a_i^T x = b_i, && \forall i \in C_2\\
&\quad x_j \geq 0, && \forall j \in D_1\\
&\quad x_j \in \mathbb{R}, && \forall j \in D_2
\end{aligned}
$$

Dual LP

min $\quad b^T \cdot y$

s.t.

$$
\begin{aligned}
x_j:&\quad \bar{a}_j y \geq c_j, && \forall j \in D_1\\
x_j:&\quad \bar{a}_j\, y = c_j, && \forall j \in D_2\\
&\quad y_i \geq 0, && \forall i \in C_1\\
&\quad y_i \in \mathbb{R}, && \forall i \in C_2
\end{aligned}
$$

$\bar{a}_j$ (column) $\uparrow$

a_i^{T} (row) $\leftarrow$

$$
\begin{vmatrix}
a_{11} & a_{12} & a_{13} & a_{14}\\
a_{21} & a_{22} & a_{23} & a_{24}\\
a_{31} & a_{32} & a_{33} & a_{34}
\end{vmatrix}
$$

Dual Linear Program: Standard Form

Primal LP

max	$c^T \cdot x$
s.t.	$Ax \leq b$
	$x \geq 0$

Dual LP

min	$b^T \cdot y$
s.t.	$A^T y \geq c$
	$y \geq 0$

- $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$
- y_i is the dual variable corresponding to primal constraint $A_i x \leq b_i$
- $A_j^T\, y \geq c_j$ is the dual constraint corresponding to primal variable x_j

Interpretation 1: Economic Interpretation

Recall the optimal production problem

- n products, m raw materials
- Every unit of product j uses a_{ij} units of raw material i
- There are b_i units of material i available
- Product j yields profit c_j per unit
- Factory wants to maximize profit subject to available raw materials

Interpretation I: Economic Interpretation

Primal LP

$$\begin{aligned} \max \quad & c^T \cdot x \\ \text{s.t.} \quad & \textstyle\sum_{j=1}^{n} a_{ij}\, x_j \le b_i, \quad \forall i \in [m] \\ & x_j \ge 0, \quad \forall j \in [n] \end{aligned}$$

Dual LP

$$\begin{aligned} \min \quad & b^T \cdot y \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^{m} a_{ij}\, y_i \ge c_j, \quad \forall j \in [n] \\ & y_i \ge 0, \quad \forall i \in [m] \end{aligned}$$

j: product index
i: material index

Dual LP corresponds to the buyer's optimization problem, as follows:

- Buyer wants to directly buy the raw material
- Dual variable y_i is buyer's proposed price per unit of raw material i
- Dual price vector is feasible if factory is incentivized to sell materials
- Buyer wants to spend as little as possible to buy raw materials

Interpretation 1: Economic Interpretation

Primal LP

$$\begin{array}{lll} \max & c^T \cdot x & \\ \text{s.t.} & \sum_{j=1}^{n} a_{ij}\, x_j \le b_i, & \forall i \in [m] \\ & x_j \ge 0, & \forall j \in [n] \end{array}$$

Dual LP

$$\begin{array}{lll} \min & b^T \cdot y & \\ \text{s.t.} & \sum_{i=1}^{m} a_{ij}\, y_i \ge c_j, & \forall j \in [n] \\ & y_i \ge 0, & \forall i \in [m] \end{array}$$

	x_1	x_2	x_3	x_4	
y_1	a_{11}	a_{12}	a_{13}	a_{14}	b_1
y_2	a_{21}	a_{22}	a_{23}	a_{24}	b_2
y_3	a_{31}	a_{32}	a_{33}	a_{34}	b_3
	c_1	c_2	c_3	c_4	

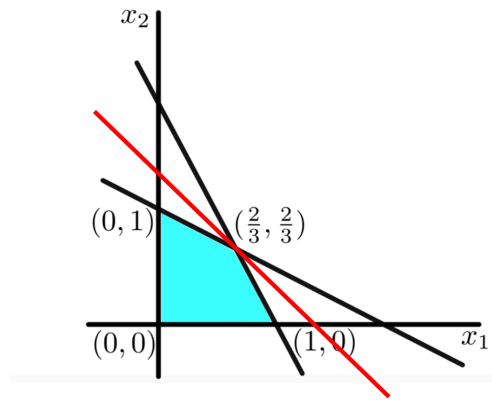
price of material (← y_1, y_2, y_3)

units of products ($x_1, \dots, x_4$ →)

Interpretation II: Finding Best Upperbound

- Consider the simple LP from previous 2-D example

$$\begin{aligned} \text{maximize} \quad & x_1 + x_2 \\ \text{subject to} \quad & x_1 + 2x_2 \le 2 \\ & 2x_1 + x_2 \le 2 \\ & x_1, x_2 \ge 0 \end{aligned}$$



- We found that the optimal solution was at $(\frac{2}{3},\frac{2}{3})$ with an optimal value of $\frac{4}{3}$.
- What if, instead of finding the optimal solution, we sought to find an upperbound on its value by combining inequalities?
  - Each inequality implies an upper bound of 2
  - Multiplying each by 1 and summing gives $x_1 + x_2 \le 4/3$.

Interpretation II: Finding Best Upperbound

Primal LP

max $c^T \cdot x$
s.t. $Ax \leq b$
$x \geq 0$

Dual LP

min $b^T \cdot y$
s.t. $A^T y \geq c$
$y \geq 0$

- Multiplying each row i by y_i and summing gives the inequality

$$y^T A x \leq y^T b$$

(now we see why $y_i \geq 0$ when $a_i x \leq b_i$ but $y_i \in \mathbb{R}$ when $a_i x = b_i$)

- When $c^T \leq y^T A$, the right hand side of the inequality is an upper bound on $c^T x$ for every feasible x, because

$$c^T x \leq y^T A x \leq y^T b$$

- The dual LP can be interpreted as finding the best upperbound on the primal that can be achieved this way.

Properties of Duals

- Duality is an inversion

Fact: Given any primal LP, the dual of its dual is itself.

Proof: homework exercise

Primal LP

$$\begin{aligned} \max \quad & c^T \cdot x \\ \text{s.t.} \quad & a_i^T x \leq b_i, \quad \forall i \in C_1 \\ & a_i^T x = b_i, \quad \forall i \in C_2 \\ & x_j \geq 0, \quad \forall j \in D_1 \\ & x_j \in \mathbb{R}, \quad \forall j \in D_2 \end{aligned}$$

Dual LP

$$\begin{aligned} \min \quad & b^T \cdot y \\ \text{s.t.} \quad & \bar{a}_j y \geq c_j, \quad \forall j \in D_1 \\ & \bar{a}_j y = c_j, \quad \forall j \in D_2 \\ & y_i \geq 0, \quad \forall i \in C_1 \\ & y_i \in \mathbb{R}, \quad \forall i \in C_2 \end{aligned}$$

Thank You

Haifeng Xu
University of Virginia
hx4ad@virginia.edu