Mathematical Foundations for Computer Science

Probability and Optimization

Chapter 3: Linear Programming

Spring 2021

Instructor: Xiaodong Gu

Optimization in Computer Science

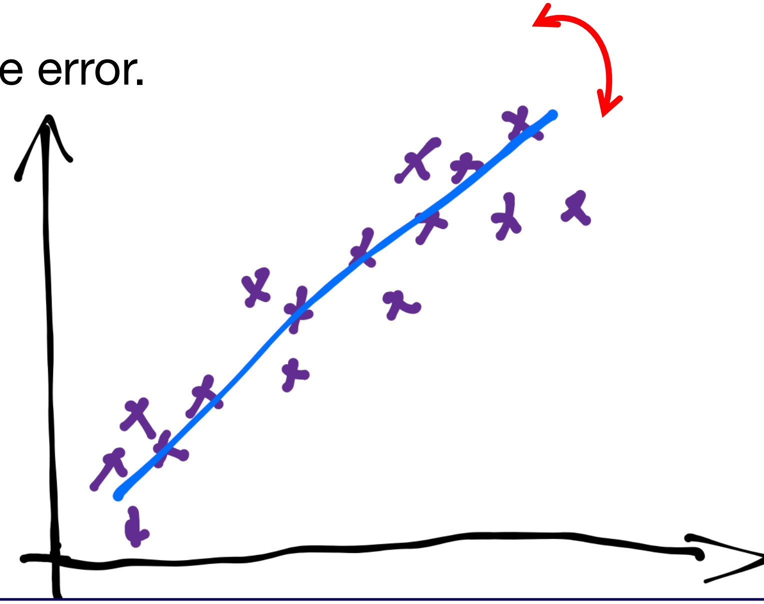
Example: Least Squares Regression

Given a set of measurements $(x_1, r_1), \ldots, (x_m, r_m)$, where $x_i \in R^n$ is the i-th input and $r_i \in R$ is the i-th output, find the linear function $f: R^n \to R$ best explaining the relationship between inputs and outputs.

Estimate $f(x) = w^T x$ for some $w \in \mathbb{R}^n$

Least squares: minimize mean-square error.

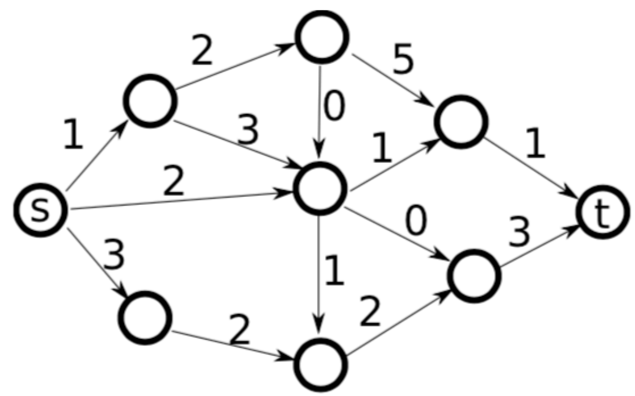
minimize $||Wx - r||_2$



Example: Shortest Path

Given a directed network G = (V, E), with a length of c_e for each edge e, find the shortest path from s to t.

x_e: edge e is selected



minimize $\sum_{e \in E} c_e x_e$

subject to

$$\sum_{e \leftarrow v} x_e = \sum_{e \rightarrow v} x_e, \quad \text{for } v \in V \setminus \{s,t\}.$$

$$\sum_{e \leftarrow s} x_e = 1$$

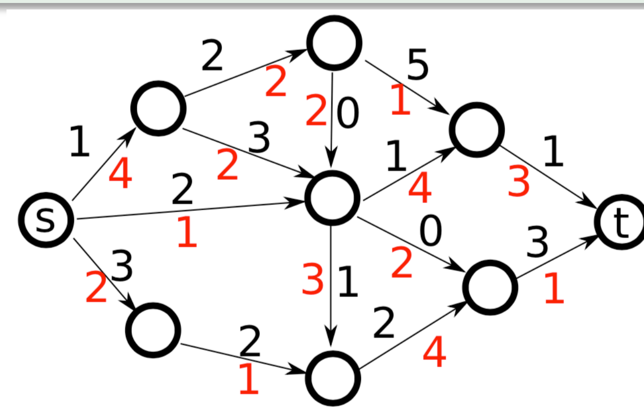
$$x_e \leq 1, \quad \text{for } e \in E.$$

$$x_e \geq 0, \quad \text{for } e \in E.$$

Optimization in Computer Science

Example: Minimum Cost Flow (网络流)

Given a directed network G = (V, E) with cost $c_e \in R^+$ per unit of traffic on edge e, and capacity d_e, find the minimum cost routing of r divisible units of traffic from s to t.



$$
\begin{array}{lll}
\text{minimize} & \sum_{e \in E} c_e x_e & \\
\text{subject to} & \sum_{e \leftarrow v} x_e = \sum_{e \rightarrow v} x_e, & \text{for } v \in V \setminus \{s, t\}\,. \\
 & \sum_{e \leftarrow s} x_e = r & \\
 & x_e \leq d_e, & \text{for } e \in E. \\
 & x_e \geq 0, & \text{for } e \in E.
\end{array}
$$

Optimization in Computer Science

Example: robot trajectory planning

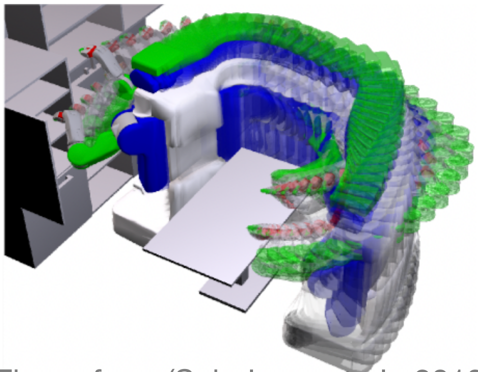


Figure from (Schulman et al., 2013)

Given robot state x_t and control inputs u_t

$$\underset{x_{1:T}, u_{1:T-1}}{\text{minimize}} \quad \sum_{i=1}^{T-1} \|x_t - x_{t+1}\|_2^2 + \|u_t\|_2^2$$

$$\text{subject to} \quad x_{t+1} = f_{\text{dynamics}}(x_t, u_t), \quad \text{(robot dynamics)}$$

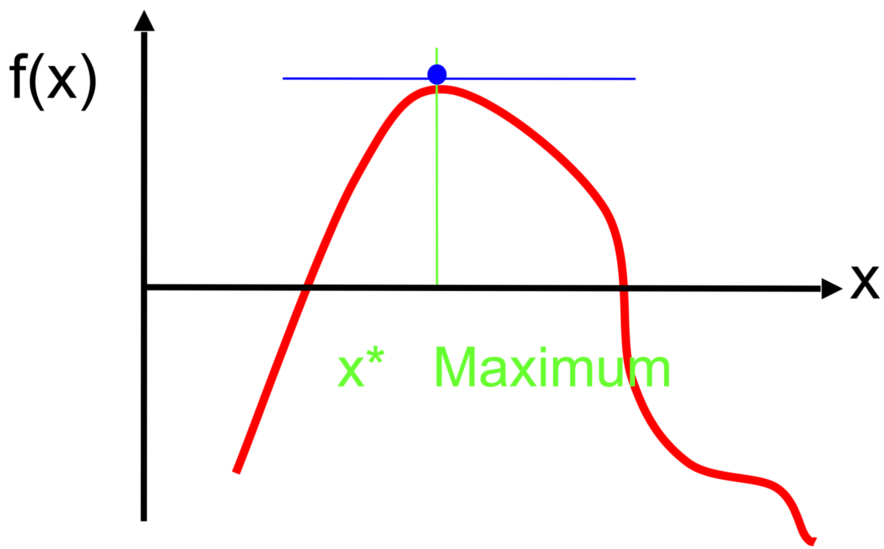
$$f_{\text{collision}}(x_t) \geq 0.1 \quad \text{(avoid collisions)}$$

$$x_1 = x_{\text{init}}, \; x_T = x_{\text{goal}}$$

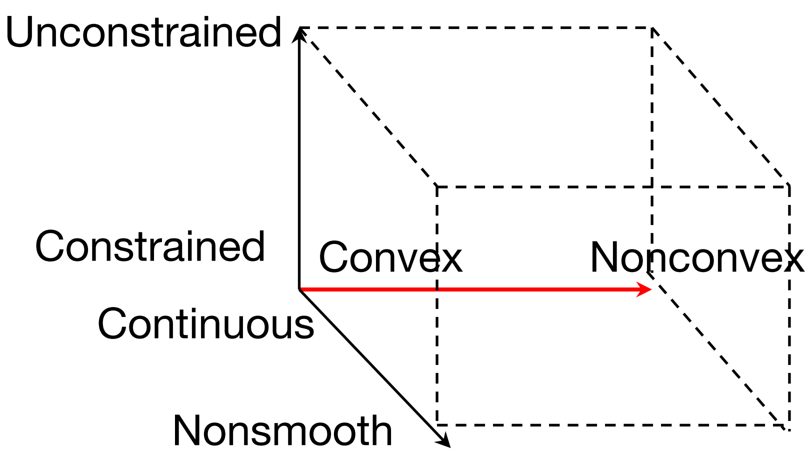
What is Optimization?

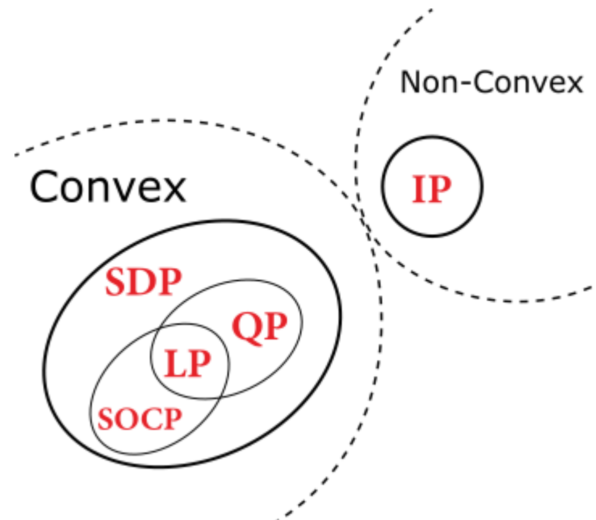
- Optimization = search for the best solution !
- In mathematical terms:

minimization or maximization of an objective function f (x) depending on variables x subject to constraints



Overview of Optimization Problems





Our focus in this course:

— **C**onstrained, **C**ontinuous, **C**onvex

Some Linear Algebra Review

Vector, Matrix, Positive (semi-) definite

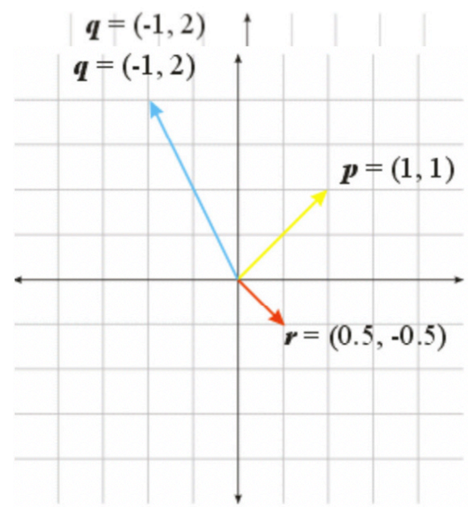
Vector

- Definition: an *n*-tuple of values
  - *n* referred to as the *dimension* of the vector
- Can be written in column form or row form

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad x^\top = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}$$

$\top$ means "transpose"

- Can think of a vector as
  - a point in space *or*
  - a directed line segment with a magnitude and direction



Vector Arithmetic

- Addition of two vectors
  - add corresponding elements

$$\mathbf{z} = \mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 & \cdots & x_n + y_n \end{pmatrix}^\top$$

- Scalar multiplication of a vector
  - multiply each element by scalar

$$\mathbf{y} = a\mathbf{x} = \begin{pmatrix} ax_1 & \cdots & ax_n \end{pmatrix}^\top$$

Vector Norms (范数)

For a vector x∈$\mathbb{R}^n$ with elements $x = (x_1, x_2, \dots, x_n)$:

- The l_2 norm, or Euclidean norm:

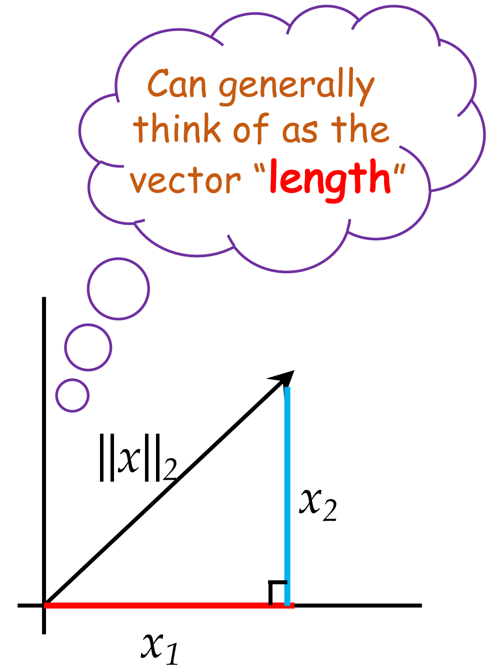
$$\|x\|_2 = \|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

- The l_1 norm:

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

- The l_p-norm:

$$\|x\|_p = \sqrt[p]{|x|_1^p + |x|_2^p + \cdots + |x|_n^p}$$



https://zhuanlan.zhihu.com/p/26884695

Matrix

A matrix $\boldsymbol{A} \in \boldsymbol{R}^{m \times n}$ is a two-dimensional array of values with m rows and n columns.

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

- a_{ij} (or A_{ij}) denotes the element in the *i*-th row and *j*-th column;
- a_i refers to the *i*-th column and $a^T{}_i$ refers to the i-th row.

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix} = \begin{bmatrix} - & a_1^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix}$$

Matrix Addition/Subtraction

Addition/subtraction of matrices is defined as addition/subtraction of corresponding elements.

for $A, B \in R^{m \times n}$,

$$C \in R^{m \times n} = A + B \Longleftrightarrow c_{ij} = a_{ij} + b_{ij}$$

Multiply a Vector by a Matrix

Multiplying a matrix $\mathbf{A} = [a_{ij}]_{m\times n}$ with a vector $\mathbf{x} = (x_1,\ldots,x_n)$ can be written as a weighted sum of **A**'s column vectors

$$\mathbf{Ax} = \mathbf{y}$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

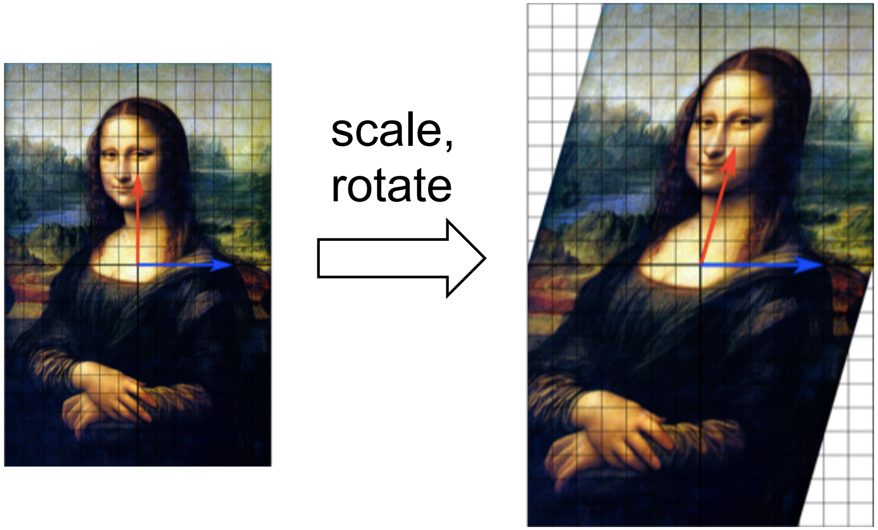
$$y_i = \sum_{j=1}^{n} a_{ij}$$

Matrix Multiplication

If $\mathbf{C}_{m\times n} = \mathbf{A}_{m\times p}\mathbf{B}_{p\times n}$, then $[c_{ij}] = \sum_{k=1}^{p} a_{ik}b_{kj}$

- in general, non-commutative: $\mathbf{AB} \neq \mathbf{BA}$
- associative: $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
- distributive: $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$

Matrix multiplication scales/rotates a geometric plane



Transpose

- Matrix transpose (denoted $^\top$)
  - swap columns and rows
  - $m \times n$ matrix becomes n x m matrix
  - example:

$$\mathbf{A} = \begin{pmatrix} 2 & 7 & -1 & 0 & 3 \\ 4 & 6 & -3 & 1 & 8 \end{pmatrix} \qquad \mathbf{A}^\top = \begin{pmatrix} 2 & 4 \\ 7 & 6 \\ -1 & -3 \\ 0 & 1 \\ 3 & 8 \end{pmatrix}$$

Inner Product

The inner product (dot product) of two vectors **x**, y ∈ R^n:

$$\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x}^T\boldsymbol{y} = \boldsymbol{y}^T\boldsymbol{x} = \sum_i x_i y_i$$

- **Example**: $\mathbf{x} = (4, 6, 1)^T$, $\mathbf{y} = (1, 3, 1)^T$, $\mathbf{x}^T\mathbf{y} = 4\times1+6\times3+1\times1 = 23$
- If $\mathbf{x}^T\mathbf{y} = 0$, then **x** and **y** are orthogonal.

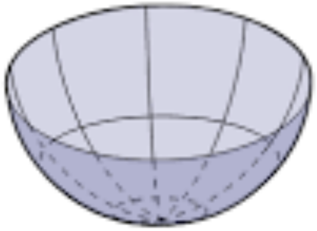
Positive (semi-)definite matrices

- A symmetric matrix A is positive semi-definite(PSD) if for all $\mathbf{x} \in \mathbb{R}^n, \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$

Example:

- The identity matrix $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is positive–definite

$$z^T I z = [x \quad y] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + y^2 \geq 0$$



- A symmetric matrix A is positive definite(PD) if for all nonzero $\mathbf{x} \in \mathbb{R}^n, \mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$

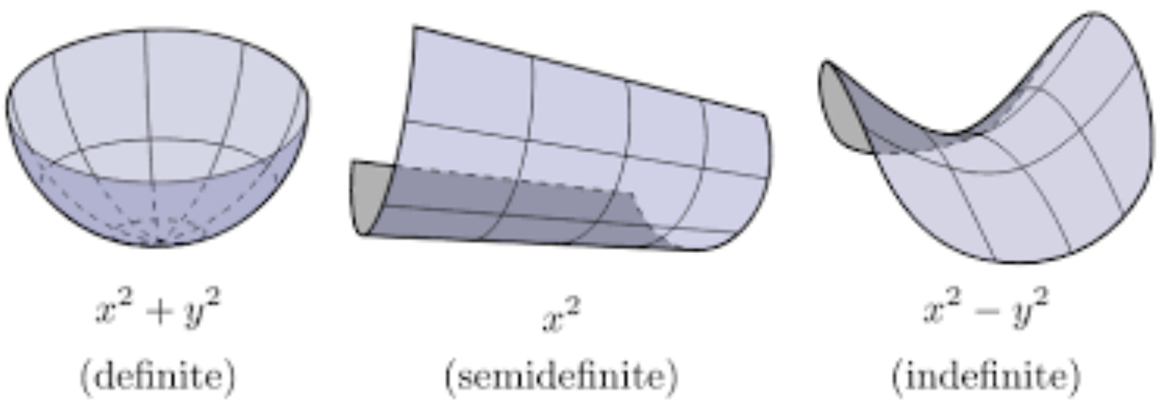
Notation: $\mathbf{A} \succeq 0$ if A is PSD, $\mathbf{A} \succ 0$ if A is PD

Examples

$$z^T A z = [x \quad y] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + y^2$$

$$z^T A z = [x \quad y] \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x^2$$

$$z^T A z = [x \quad y] \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x^2 - y^2$$



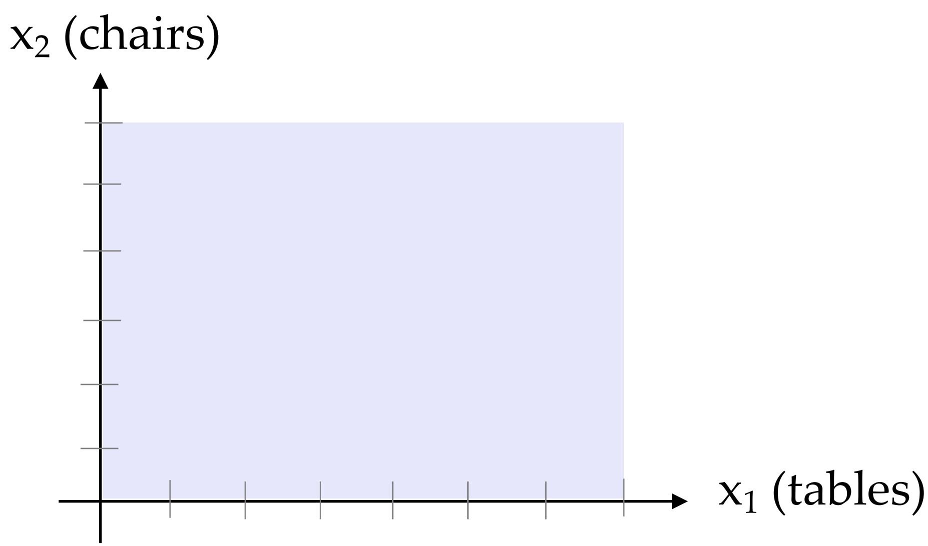
Linear Programming

Example: Optimal Manufacturing

A large factory makes tables and chairs. Each table returns a profit of \$200 and each chair a profit of \$100. Each table takes 1 unit of metal and 3 units of wood and each chair takes 2 units of metal and 1 unit of wood. The factory has 6K units of metal and 9K units of wood. How many tables and chairs should the factory make to maximize profit?

Example: Optimal Manufacturing

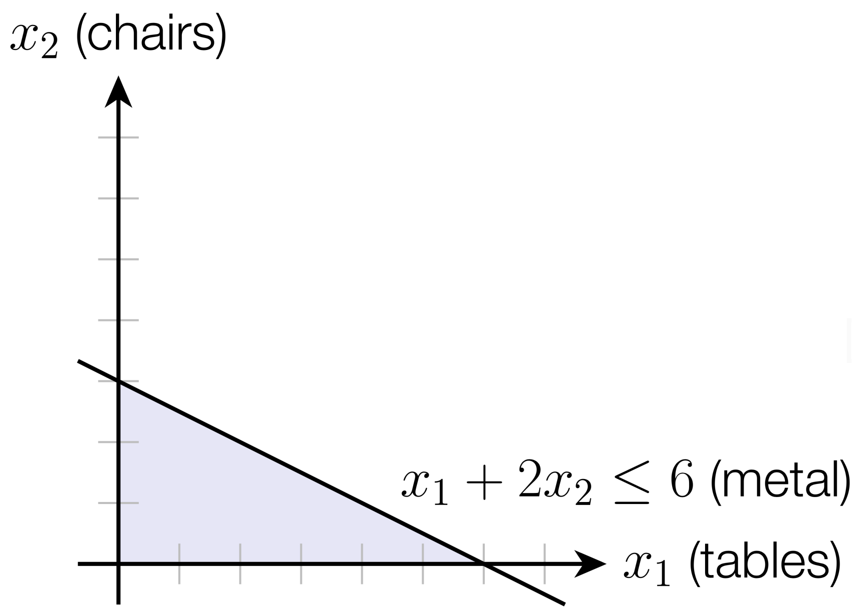
A large factory makes tables and chairs. Each table returns a profit of \$200 and each chair a profit of \$100. Each table takes 1 unit of metal and 3 units of wood and each chair takes 2 units of metal and 1 unit of wood. The factory has 6K units of metal and 9K units of wood. How many tables and chairs should the factory make to maximize profit?



Example: Optimal Manufacturing

A large factory makes tables and chairs. Each table returns a profit of $200 and each chair a profit of $100. Each table takes 1 unit of metal and 3 units of wood and each chair takes 2 units of metal and 1 unit of wood. The factory has 6K units of metal and 9K units of wood. How many tables and chairs should the factory make to maximize profit?

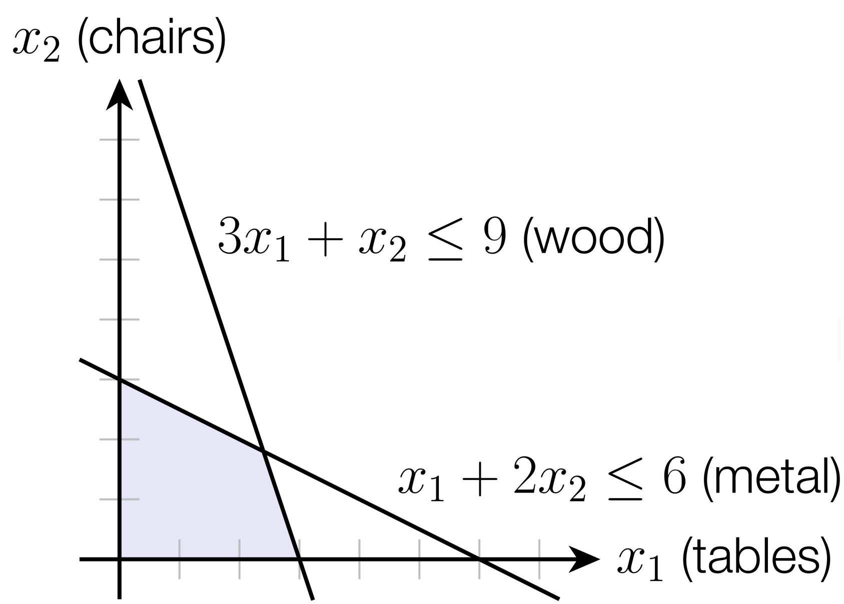
	table x_1	chair x_2	amt.
metal	1	2	6



Example: Optimal Manufacturing

A large factory makes tables and chairs. Each table returns a profit of \$200 and each chair a profit of \$100. Each table takes 1 unit of metal and 3 units of wood and each chair takes 2 units of metal and 1 unit of wood. The factory has 6K units of metal and 9K units of wood. How many tables and chairs should the factory make to maximize profit?

	table x_1	chair x_2	amt.
metal	1	2	6
wood	3	1	9



Example: Optimal Manufacturing

A large factory makes tables and chairs. Each table returns a profit of $200 and each chair a profit of $100. Each table takes 1 unit of metal and 3 units of wood and each chair takes 2 units of metal and 1 unit of wood. The factory has 6K units of metal and 9K units of wood. How many tables and chairs should the factory make to maximize profit?

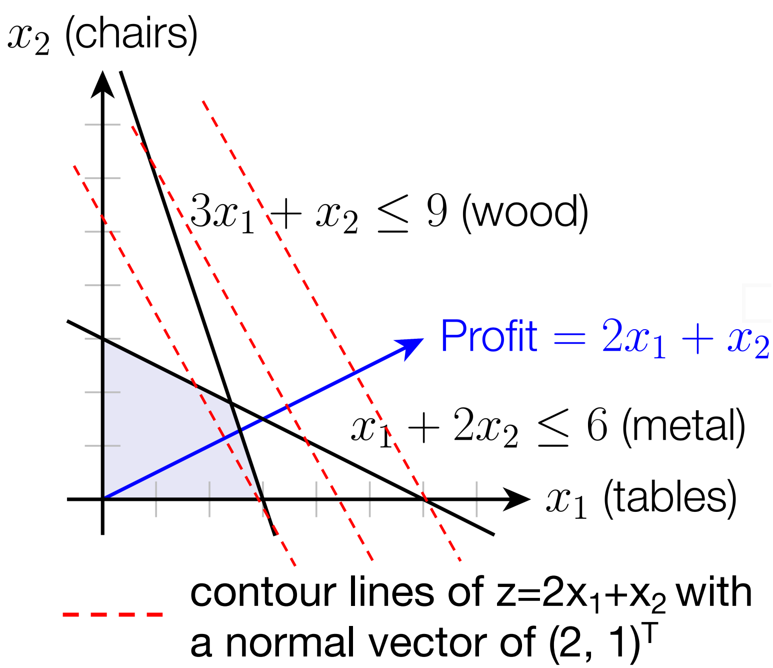
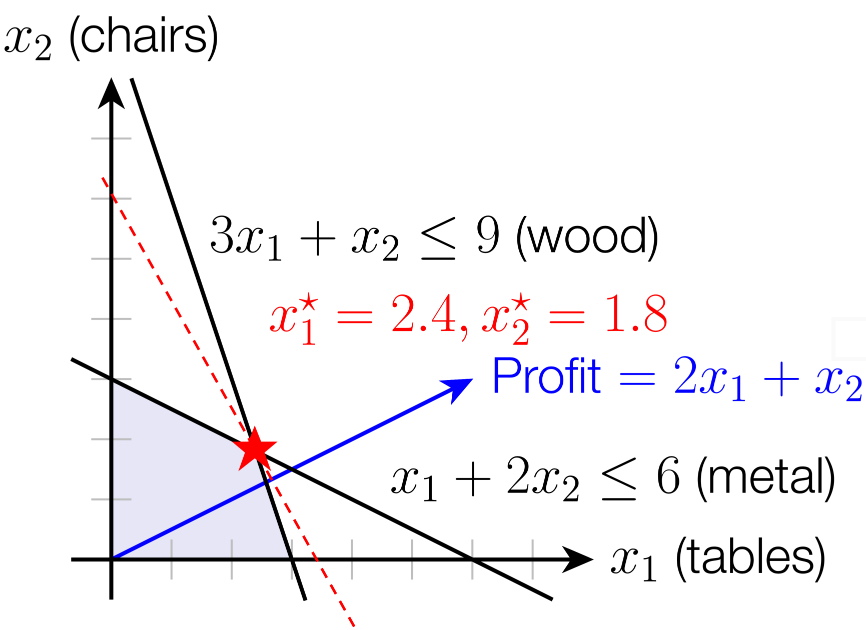


	table x_1	chair x_2	amt.
metal	1	2	6
wood	3	1	9
	$200	$100	

Which contour line stands for the optimal profit?

Example: Optimal Manufacturing

A large factory makes tables and chairs. Each table returns a profit of \$200 and each chair a profit of \$100. Each table takes 1 unit of metal and 3 units of wood and each chair takes 2 units of metal and 1 unit of wood. The factory has 6K units of metal and 9K units of wood. How many tables and chairs should the factory make to maximize profit?



$$\begin{aligned} \underset{x_1, x_2}{\text{maximize}} \quad & 2x_1 + x_2 \\ \text{subject to} \quad & x_1 + 2x_2 \leq 6 \\ & 3x_1 + x_2 \leq 9 \\ & x_1, x_2 \geq 0 \end{aligned}$$

$$c = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 6 \\ 9 \end{pmatrix}$$

$$\Downarrow$$

$$\begin{aligned} \text{maximize} \quad & c^T x \\ \text{subject to} \quad & Ax \preceq b \\ & x \succeq 0 \end{aligned}$$

Example: Manufacturing

General Form

- 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
- Facility wants to maximize profit subject to available raw materials

	x_1	x_2	x_3	x_4	
m_1	a_{11}	a_{12}	a_{13}	a_{14}	b_1
m_2	a_{21}	a_{22}	a_{23}	a_{24}	b_2
m_3	a_{31}	a_{32}	a_{33}	a_{34}	b_3
	c_1	c_2	c_3	c_4	

$$
\begin{array}{ll}
\text{maximize} & z = c_1x_1 + c_2x_2 + \cdots + c_nx_n \\
\text{subject to} & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \le b_1 \\
& \cdots \\
& a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \le b_m \\
\text{and} & x_i \ge 0, \quad \text{for } i = 1, \ldots, n
\end{array}
$$

Linear Programming: Standard Form

$$\begin{array}{lll} \text{maximize} & c^T x & \\ \text{subject to} & a_i^T x \le b_i, & \text{for } i = 1, \dots, m \\ & x_j \ge 0, & \text{for } j = 1, \dots, n \end{array}$$

- Every LP can be transformed to the standard form:
  - Minimizing $c^T x$ is equivalent to maximizing $-c^T x$
  - $\geq$ constraints can be flipped by multiplying by -1
  - Each equality constraint can be replaced by two inequalities ($a_i^T x \le b_i$ and $-a_i^T x \le b_i$)
  - Unconstrained variables x_j can be replaced by $x_j^+ - x_j^-$, where both x_j^+ and x_j^- are constrained to be nonnegative.

Example:

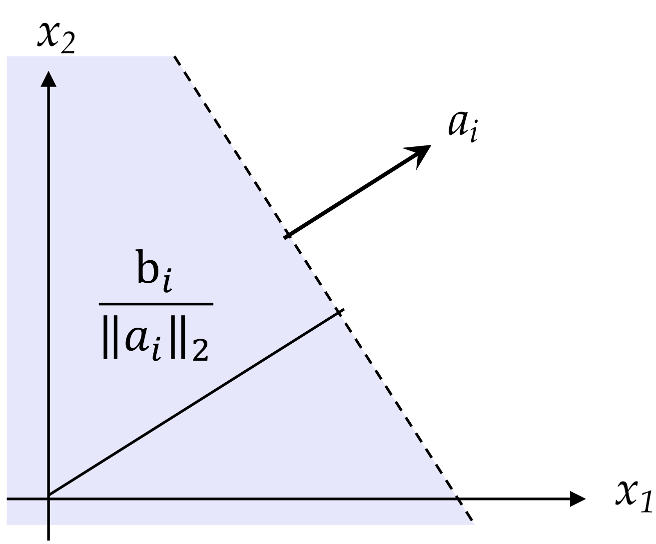
$$\begin{array}{ll} \text{minimize} & 3x_1+x_2 \\ \text{subject to} & x_1 > x_2+5 \\ & x_1+3x_2=10 \end{array} \Longrightarrow \begin{array}{ll} \text{maximize} & -3x_1-x_2 \\ \text{subject to} & -x_1+x_2 < -5 \\ & x_1-3x_2 \le 10 \\ & -x_1+3x_2 \le -10 \end{array}$$

Geometric View

Consider the inequality constraints of the linear program written out explicitly:

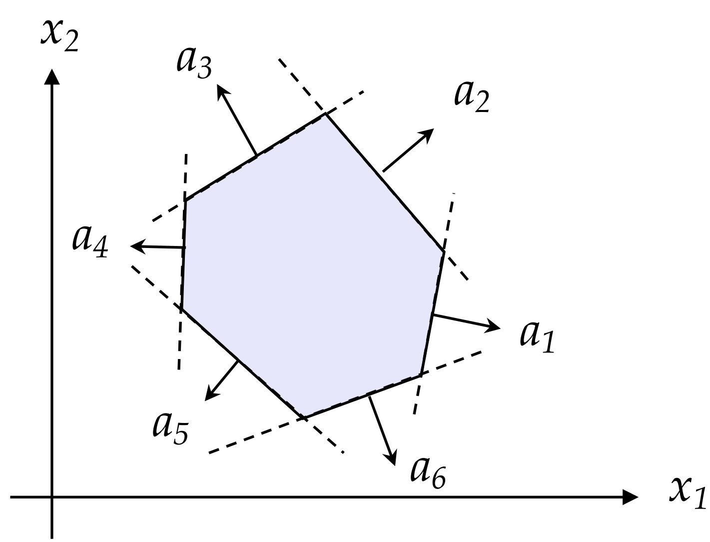
$$\begin{aligned} \underset{x}{\text{maximize}} \quad & c^T x \\ \text{subject to} \quad & a_i^T x \le b_i, \; i = 1,\ldots,m \end{aligned}$$

Each linear inequality constraint $a_i^T x \le b_i$ defines a half-space



Geometric View

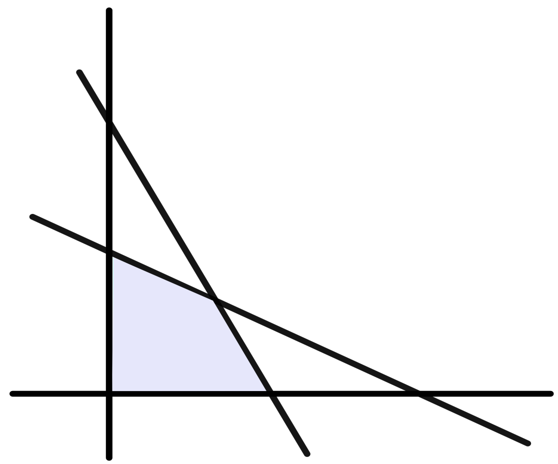
Multiple halfspace constraints, $a_i^T x \le b_i$, i = 1, . . . , m (or equivalently $Ax \le b$), define what is called a polyhedron.



Feasible region of an LP is a polyhedron.

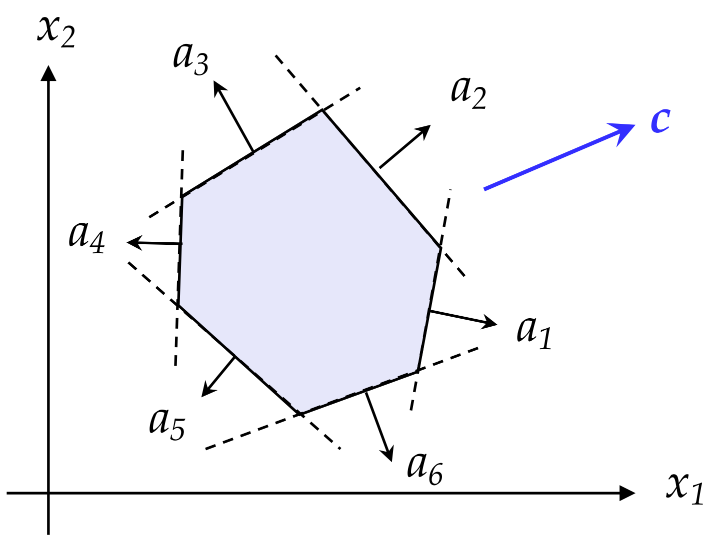
Geometric View

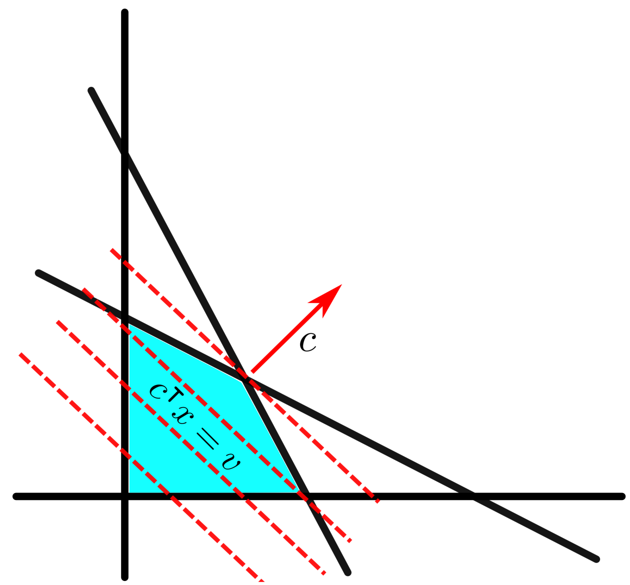
Vertex (Corner): a point x is a vertex of polyhedron P if $!\exists y \neq 0$ with $x+y \in P$ and $x-y \in P$



Geometric View

So linear programming is equivalent to maximizing some direction ($\mathbf{c}^T\mathbf{x}$) over a polyhedron.

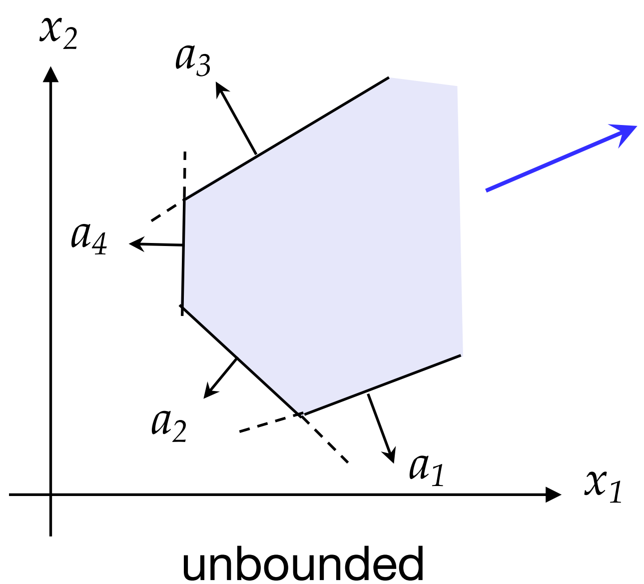


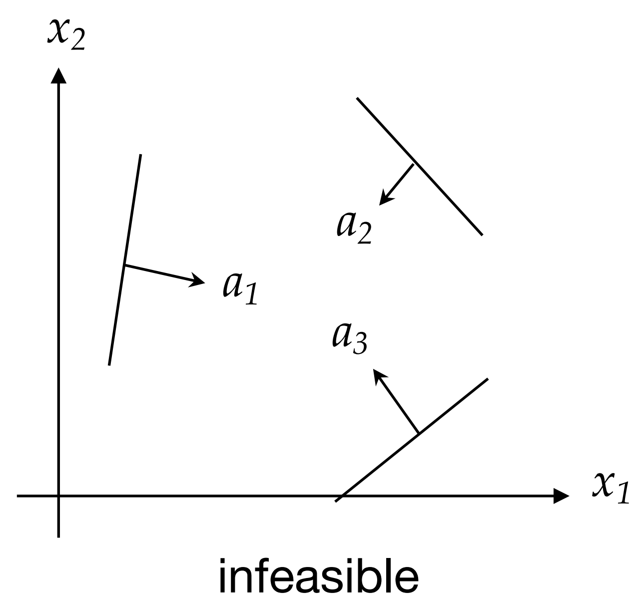


Important point: note that a maximum will always occur at a "corner" of the polyhedron (this is exactly the property that the simplex algorithm will exploit)

Unbounded and Infeasible Problems

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



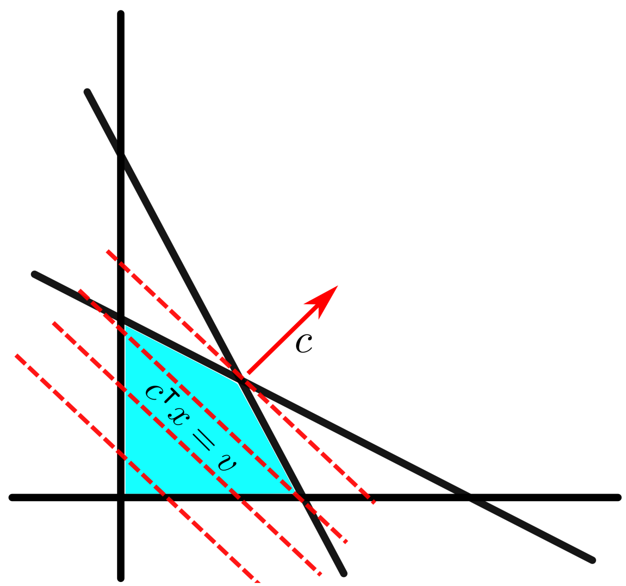


Example: suppose we have both the constraints $x_1 \geq 5$ and $x_1 \leq 4$

Fundamental Theorem of LP

If an LP in standard form has an optimal solution, then it has a vertex optimal solution.

Informal Proof: the optimal contour line cuts the polyhedron at a vertex.



Solving LP

Algorithms to Solve LP

- Simplex (单纯形法)
- Ellipsoid method (椭球法)

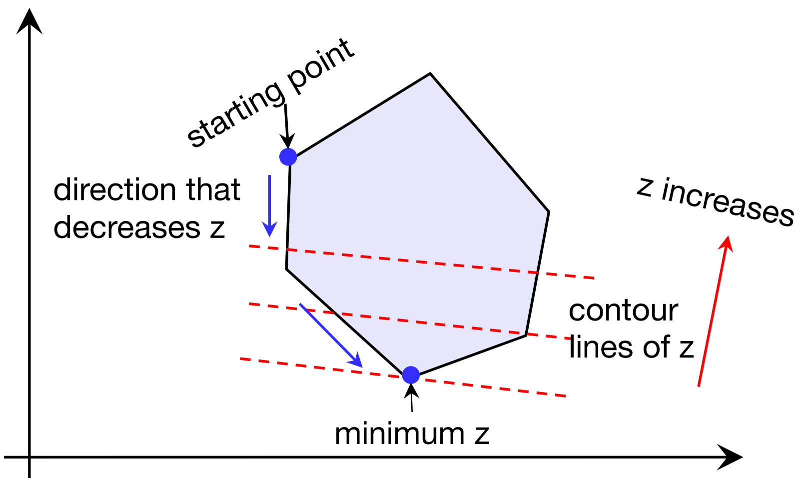
specifically designed for LP
(not important for this course)

- Unconstrained Optimization Methods
  - Gradient Descent (梯度下降)
  - …
- …

for convex optimization
(to introduce later)

Simplex – The Idea

- First methodical procedure for solving linear programs.
- Efficient in practice, leading to conjectures that it runs in polynomial time.



① starting from a vertex (with two tight constrains, i.e., only two x_i's > 0),

② finding a neighbor vertex that decreases z (replacing a x_i ∈ basic solutions with x_j∈ non basic solutions) ,

③ until reaching the minimum z ($z = \Sigma_{x_j \in \text{non basic}} w_j x_j$, with no $x_j<0$)

Duality in LP

An Alternative View of LP

Finding the upper bound of the original objective function

$$
\begin{aligned}
\text{maximize} \quad & 2x_1+3x_2 = z^* \\
\text{subject to} \quad & 4x_1+8x_2 \le 12, \\
& 2x_1+x_2 \le 3, \\
& 3x_1+2x_2 \le 4, \\
& x_1 \ge 0, x_2 \ge 0
\end{aligned}
$$

let z* denote the original objective

What is the upper bound of z*?

Let's try to bound z* from the original constrains:

$$2x_1+3x_2 \le 4x_1+8x_2 \le 12$$

$$2x_1+3x_2 \le \tfrac{1}{2}(4x_1+8x_2) \le \tfrac{1}{2}(12) = 6$$

So we have improved our upper bound to 6!

$$2x_1+3x_2 = \tfrac{1}{3}[(4x_1+8x_2)+(2x_1+x_2)] \le \tfrac{1}{3}(12+3) = 5$$

We further improved our upper bound to 5!

By continuously improving the upper bound, can we reach the original optimal?

An Alternative View of LP

We are finding suitable multipliers y_1, y_2, y_3 that are all non-negative to the original constrains

$$y_1 \times 4x_1+8x_2 \le 12$$
$$y_2 \times 2x_1+x_2 \le 3$$
$$y_3 \times 3x_1+2x_2 \le 4$$

which yields

$$(4y_1+2y_2+3y_3)x_1+(8y_1+y_2+3y_3)x_2 \le 12y_1+3y_2+4y_3$$

Now, we want

$$z^*=2x_1+3x_2 \le (4y_1+2y_2+3y_3)x_1+(8y_1+y_2+2y_3)x_2$$

which is equivalent to the dual optimization problem:

$$\begin{array}{ll} \text{minimize} & 12y_1+3y_2+4y_3 \\ \text{subject to} & 4y_1+2y_2+3y_3 \ge 2 \\ & 8y_1+y_2+2y_3 \ge 3 \\ & y_1,y_2,y_3 \ge 0 \end{array}$$

Linear Programming Duality

Primal LP

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \preccurlyeq b \end{array}$$

Dual LP

$$\begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \succcurlyeq c \\ & y \succcurlyeq 0 \end{array}$$

	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	

- $A \in \mathbb{R}^{m\times n},\ c \in \mathbb{R}^n,\ b \in \mathbb{R}^m$
- 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
- Every feasible solution y of dual LP provides an upper bound on the maximum of the objective function of primal LP

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
- Facility wants to maximize profit subject to available raw materials.

	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	

Interpretation 1: Economic Interpretation

Primal LP

$\max \ \sum_{j=1}^{n} c_j x_j$

s.t. $\sum_{j=1}^{n} a_{ij} x_j \le b_i,$ for $i \in [m]$

$x_j \ge 0,$ for j $\in [n]$

Dual LP

$\min \ \sum_{i=1}^{m} b_i y_i$

s.t. $\sum_{i=1}^{m} a_{ij} y_i \ge c_j,$ for j $\in [n]$

$y_i \ge 0,$ for i $\in [m]$

	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	

- Dual variable y_i is a proposed price per unit of raw material i
- Dual price vector is feasible if facility has incentive to sell materials
- Buyer wants to spend as little as possible to buy materials

Interpretation 2: Finding the Best Upperbound

Consider the simple LP from examples

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

- We found that the optimal solution was at (⅔ , ⅔), with an optimal value of 4/3.
- What if, instead of finding the optimal solution, we saught to find an upperbound on its value by combining inequalities?
  - Each inequality implies an upper bound of 2
  - Multiplying each by ⅓ and summing gives $x_1 + x_2 \le 4/3$.

Interpretation 2: Finding the Best Upperbound

	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	

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

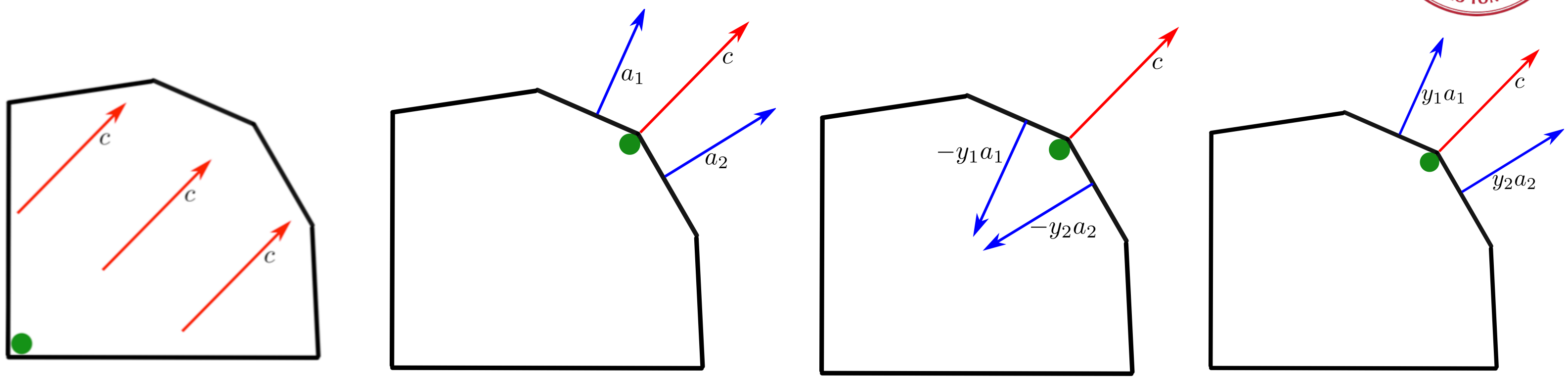
$$y^T A x \leq y^T b$$

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

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

The dual LP can be thought of as trying to find the best upper bound on the primal that can be achieved this way.

Interpretation 3: Physical Forces



- Apply force field c to a ball inside bounded polytope $Ax \preceq b$.
- Eventually, ball will come to rest against the walls of the polytope.
- Wall $a_i x \le b_i$ applies some force $-y_i a_i$ to the ball
- Since the ball is still, $c^T = \Sigma_i y_i a_i = y^T A$.
- Dual can be thought of as trying to minimize "work" $\Sigma_i y_i b_i$ ($\Sigma_i y_i b_i \geq \Sigma_i y_i a_i x = c^T x$) to bring ball back to origin by moving polytope.
- We will see that, at optimality, only the walls adjacent to the ball push (Complementary Slackness: $y_i(a_i x - b_i) = 0$)

Weak Duality

Primal LP

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \preccurlyeq b \end{array}$$

Dual LP

$$\begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \succcurlyeq c \\ & y \succcurlyeq 0 \end{array}$$

Theorem (Weak Duality)

For every primal feasible x and dual feasible y, we have $c^T x \le b^T y$.

Proof: $c^T x \le y^T A x \le y^T b$

Interpretation of Weak Duality

Economic Interpretation

If selling the raw materials is more profitable than making any individual product, then total money collected from sale of raw materials ($b^T y$) would exceed profit from production ($c^T x$).

Upperbound Interpretation

Self explanatory

Physical Interpretation

Work required to bring ball back to origin by pulling polytope is at least potential energy difference between origin and primal optimum.

Strong Duality

Primal LP

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \preccurlyeq b \end{array}$$

Dual LP

$$\begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \succcurlyeq c \\ & y \succcurlyeq 0 \end{array}$$

Theorem (Strong Duality)

If either the primal or dual is feasible and bounded, then so is the other and OPT (Primal) = OPT (Dual). (i.e., $c^T x^* = b^T y^*$)

Interpretation of Strong Duality

Economic Interpretation

Buyer can offer prices for raw materials that would make facility indifferent between production and sale (没有中间商赚差价).

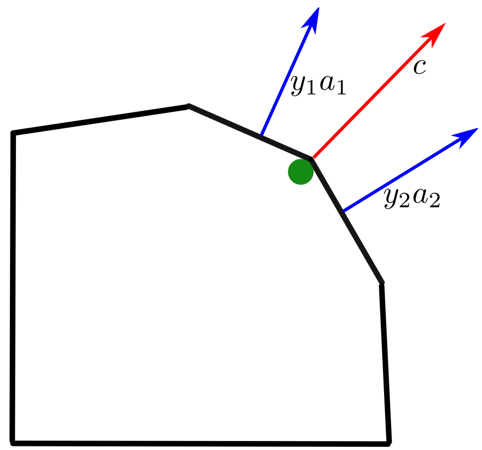
Upperbound Interpretation

The method of scaling and summing inequalities yields a tight upperbound on the primal optimal value.

Physical Interpretation

There is an assignment of forces to the walls of the polytope that brings ball back to the origin without wasting energy.

Informal Proof of Strong Duality



Recall the physical interpretation of duality
When ball is stationary at **x**, we expect force **c** to be neutralized only by constraints that are tight. i.e. force multipliers **y** $\geqslant 0$ s.t.

- $y^T A = c$
- $y_i(b_i - a_i x) = 0$ (complementary slackness)
- $y^T b - c^T x = y^T b - y^T A x = \Sigma_i y_i (b_i - a_i x) = 0$

We found a primal and dual solution that are equal in value!