Lecture 1
Introduction

- course overview
  - linear optimization
  - examples
  - history
  - approximate syllabus

- basic definitions
  - linear optimization in vector and matrix notation
  - halfspaces and polyhedra
  - geometrical interpretation

Linear optimization

$$
\begin{array}{ll}
\text{minimize} & \sum_{j=1}^{n} c_j x_j \\
\text{subject to} & \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, \ldots, m \\
& \sum_{j=1}^{n} d_{ij} x_j = f_i, \quad i = 1, \ldots, p
\end{array}
$$

- n optimization variables: x_1, . . . , x_n (real scalars)
- problem data (parameters): the coefficients c_j, a_{ij}, b_i, d_{ij}, f_i
- $\sum_j c_j x_j$ is the *cost function* or *objective function*
- $\sum_j a_{ij} x_j \leq b_i$ and $\sum_j d_{ij} x_j = f_i$ are inequality and equality *constraints*

called a **linear optimization problem** or **linear program** (LP)

Importance

low complexity

- problems with several thousand variables, constraints routinely solved
- much larger problems (millions of variables) if problem data are sparse
- widely available software
- theoretical worst-case complexity is polynomial

wide applicability

- originally developed for applications in economics and management
- today, used in all areas of engineering, data analysis, finance, . . .
- a key tool in combinatorial optimization

extensive theory

no simple formula for solution but extensive, useful (duality) theory

Example: open-loop control problem

single-input/single-output system (input $u(t)$, output $y(t)$ at time t)

$$y(t) = h_0 u(t) + h_1 u(t-1) + h_2 u(t-2) + h_3 u(t-3) + \cdots$$

output tracking problem: minimize deviation from desired output $y_{\rm des}(t)$

$$\max_{t=0,\ldots,N} |y(t) - y_{\rm des}(t)|$$

subject to input amplitude and slew rate constraints:

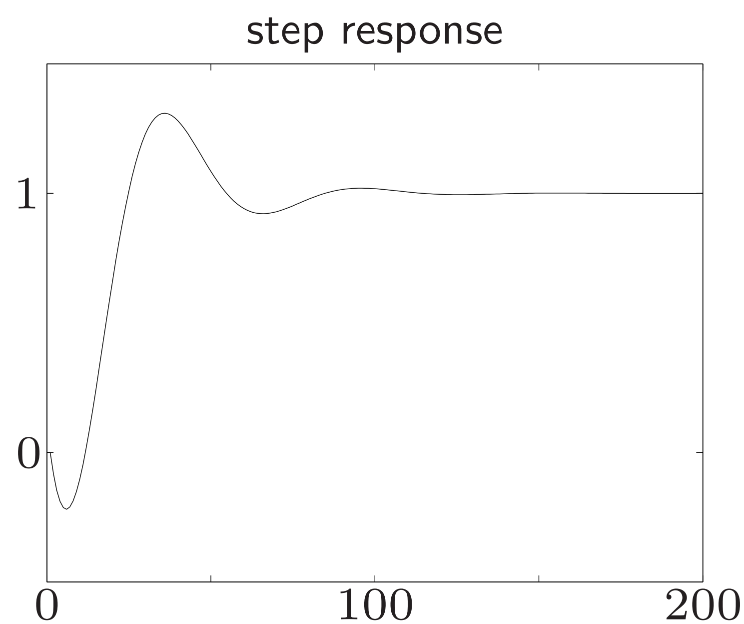
$$|u(t)| \leq U, \qquad |u(t+1) - u(t)| \leq S$$

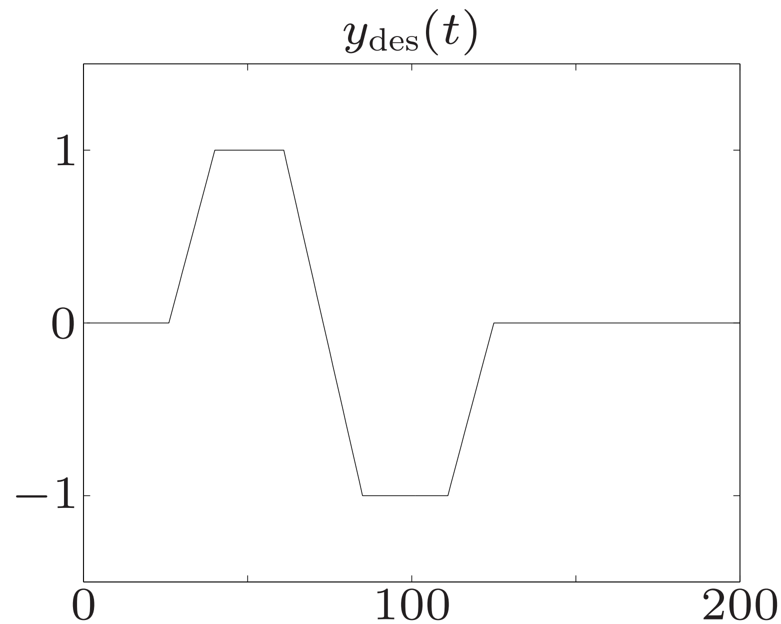
variables: $u(0)$, . . . , $u(M)$ (with $u(t) = 0$ for $t < 0$, $t > M$)

solution: can be formulated as an LP, hence easily solved (more later)

example

step response ($s(t) = h_t + \cdots + h_0$) and desired output:

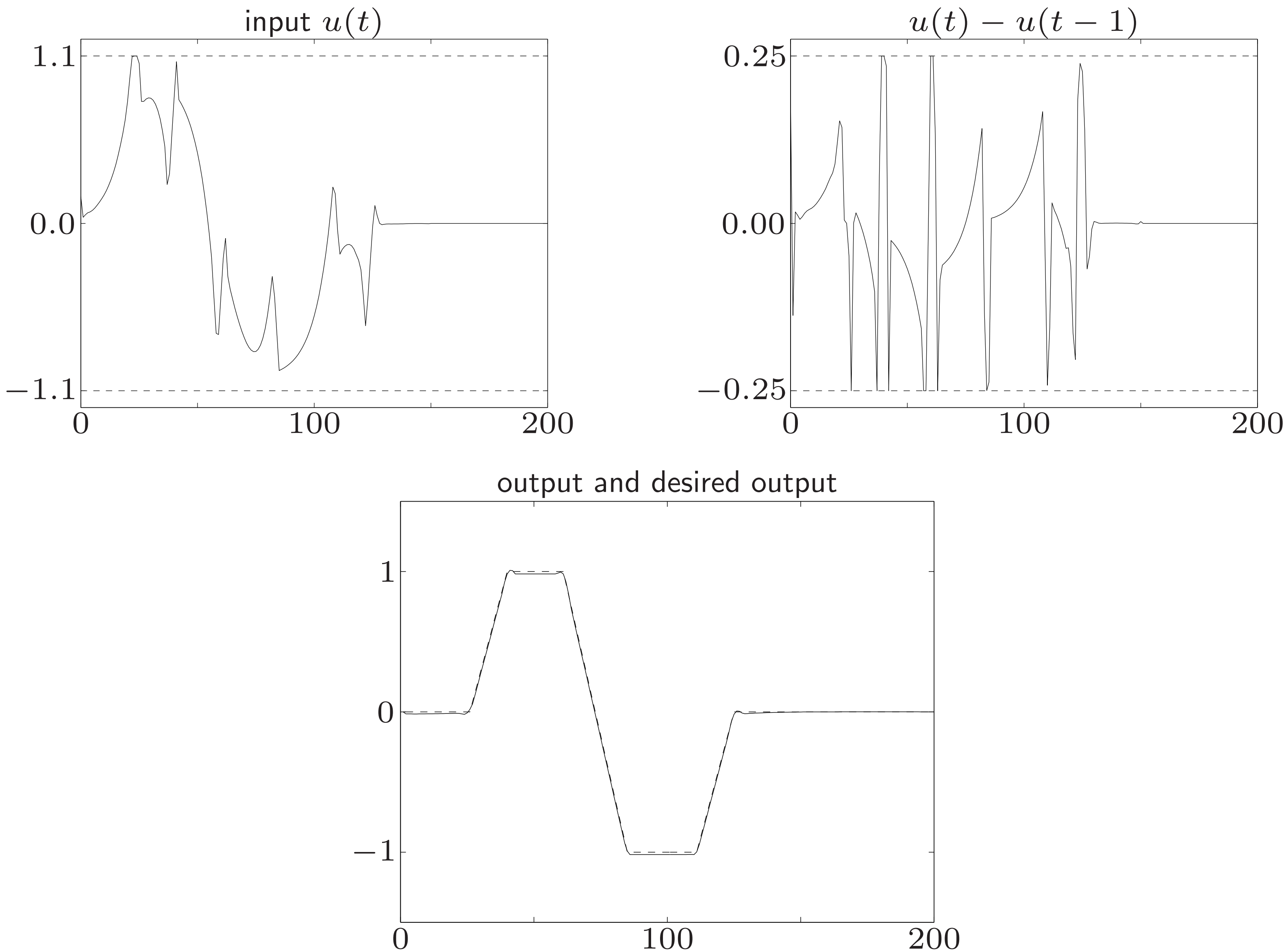




amplitude and slew rate constraint on u:

$$|u(t)| \leq 1.1, \qquad |u(t) - u(t-1)| \leq 0.25$$

optimal solution (computed via linear optimization)



Example: assignment problem

- match N people to N tasks
- each person is assigned to one task; each task assigned to one person
- cost of assigning person i to task j is a_{ij}

combinatorial formulation

$$
\begin{array}{ll}
\text{minimize} & \sum_{i,j=1}^{N} a_{ij} x_{ij} \\
\text{subject to} & \sum_{i=1}^{N} x_{ij} = 1, \quad j = 1, \ldots, N \\
& \sum_{j=1}^{N} x_{ij} = 1, \quad i = 1, \ldots, N \\
& x_{ij} \in \{0, 1\}, \quad i, j = 1, \ldots, N
\end{array}
$$

- variable $x_{ij} = 1$ if person i is assigned to task j; $x_{ij} = 0$ otherwise
- $N!$ possible assignments, *i.e.*, too many to enumerate

linear optimization formulation

$$
\begin{array}{ll}
\text{minimize} & \sum_{i,j=1}^{N} a_{ij} x_{ij} \\
\text{subject to} & \sum_{i=1}^{N} x_{ij} = 1, \quad j = 1, \dots, N \\
& \sum_{j=1}^{N} x_{ij} = 1, \quad i = 1, \dots, N \\
& 0 \leq x_{ij} \leq 1, \quad i, j = 1, \dots, N
\end{array}
$$

- we have *relaxed* the constraints $x_{ij} \in \{0, 1\}$
- it can be shown that at the optimum $x_{ij} \in \{0, 1\}$ (see later)
- hence, can solve (this particular) combinatorial problem efficiently (via linear optimization or specialized methods)

Brief history

- **1940s** (Dantzig, Kantorovich, Koopmans, von Neumann, . . .)

  foundations, motivated by economics and logistics problems

- **1947** (Dantzig): simplex algorithm

- **1950s–60s:** applications in other disciplines

- **1979** (Khachiyan): ellipsoid algorithm: more efficient (polynomial-time) than simplex in worst case, much slower in practice

- **1984** (Karmarkar): projective (interior-point) algorithm: polynomial-time worst-case complexity, and efficient in practice

- **since 1984**: variations of interior-point methods (improved complexity or efficiency in practice), software for large-scale problems

Tentative syllabus

- **linear and piecewise-linear optimization**
- **polyhedral geometry**
- **duality**
- **applications**
- **algorithms:** simplex algorithm, interior-point algorithms, decomposition
- **applications in network and combinatorial optimization**
- **extensions:** linear-fractional programming
- **introduction to integer linear programming**

Vectors

vector of length n (or n-vector)

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- we also use the notation $x = (x_1, x_2, \ldots, x_n)$
- x_i is ith *component* or *element* (real unless specified otherwise)
- set of real n-vectors is denoted $\mathbf{R}^n$

special vectors (with n determined from context)

- $x = 0$ (zero vector): $x_i = 0$, $i = 1, \ldots, n$
- $x = \mathbf{1}$ (vector of all ones): $x_i = 1$, $i = 1, \ldots, n$
- $x = e_i$ (ith *basis* or *unit vector*): $x_i = 1$, $x_k = 0$ for $k \neq i$

Matrices

matrix of size $m \times n$

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}$$

- A_{ij} (or a_{ij}) is the i, j *element* (or *entry*, *coefficient*)
- set of real $m \times n$-matrices is denoted $\mathbf{R}^{m \times n}$
- vectors can be viewed as matrices with one column

special matrices (with size determined from context)

- $X = 0$ (zero matrix): $X_{ij} = 0$ for $i = 1, \ldots, m$, $j = 1, \ldots, n$
- $X = I$ (identity matrix): $m = n$ with $X_{ii} = 1$, $X_{ij} = 0$ for $i \neq j$

Operations

- matrix transpose A^T
- scalar multiplication αA
- addition $A + B$ and subtraction $A - B$ of matrices of the same size
- product $y = Ax$ of a matrix with a vector of compatible length
- product $C = AB$ of matrices of compatible size
- inner product of n-vectors:

$$x^T y = x_1 y_1 + \cdots + x_n y_n$$

LP in inner-product notation

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^{n} c_j x_j \\ \text{subject to} & \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, \ldots, m \\ & \sum_{j=1}^{n} d_{ij} x_j = f_i, \quad i = 1, \ldots, p \end{array}$$

inner-product notation

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \\ & d_i^T x = f_i, \quad i = 1, \ldots, p \end{array}$$

c, a_i, d_i are n-vectors:

$$c = (c_1, \ldots, c_n), \qquad a_i = (a_{i1}, \ldots, a_{in}), \qquad d_i = (d_{i1}, \ldots, d_{in})$$

LP in matrix notation

$$
\begin{array}{ll}
\text{minimize} & \sum_{j=1}^{n} c_j x_j \\
\text{subject to} & \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, \ldots, m \\
& \sum_{j=1}^{n} d_{ij} x_j = f_i, \quad i = 1, \ldots, p
\end{array}
$$

matrix notation

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax \leq b \\
& Dx = f
\end{array}
$$

- A is $m \times n$-matrix with elements a_{ij}, rows a_i^T
- D is $p \times n$-matrix with elements d_{ij}, rows d_i^T
- inequality is component-wise vector inequality

Terminology

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \\ & Dx = f \end{array}$$

- x is **feasible** if it satisfies the constraints $Ax \leq b$ and $Dx = f$
- **feasible set** is set of all feasible points
- $x^\star$ is **optimal** if it is feasible and $c^T x^\star \leq c^T x$ for all feasible x
- the **optimal value** of the LP is $p^\star = c^T x^\star$
- **unbounded problem**: $c^T x$ unbounded below on feasible set ($p^\star = -\infty$)
- **infeasible probem**: feasible set is empty ($p^\star = +\infty$)

Vector norms

Euclidean norm

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

ℓ_1-norm and ℓ_∞-norm

$$\begin{aligned}\|x\|_1 &= |x_1| + |x_2| + \cdots + |x_n| \\ \|x\|_\infty &= \max\{|x_1|, |x_2|, \ldots, |x_n|\}\end{aligned}$$

properties (satisfied by any norm $f(x)$)

- $f(\alpha x) = |\alpha| f(x)$ (homogeneity)
- $f(x+y) \leq f(x) + f(y)$ (triangle inequality)
- $f(x) \geq 0$ (nonnegativity); $f(x) = 0$ if only if $x = 0$ (definiteness)

Cauchy-Schwarz inequality

$$-\|x\|\|y\| \leq x^T y \leq \|x\|\|y\|$$

- holds for all vectors x, y of the same size
- $x^T y = \|x\|\|y\|$ iff x and y are aligned (nonnegative multiples)
- $x^T y = -\|x\|\|y\|$ iff x and y are opposed (nonpositive multiples)
- implies many useful inequalities as special cases, for example,

$$-\sqrt{n}\,\|x\| \leq \sum_{i=1}^{n} x_i \leq \sqrt{n}\,\|x\|$$

Angle between vectors

the angle $\theta = \angle(x, y)$ between nonzero vectors x and y is defined as

$$\theta = \arccos \frac{x^T y}{\|x\| \|y\|} \qquad (i.e., \ x^T y = \|x\| \|y\| \cos\theta)$$

- we normalize θ so that $0 \leq \theta \leq \pi$
- relation between sign of inner product and angle

$x^T y > 0$	$\theta < \frac{\pi}{2}$	(vectors make an acute angle)
$x^T y = 0$	$\theta = \frac{\pi}{2}$	(orthogonal vectors)
$x^T y < 0$	$\theta > \frac{\pi}{2}$	(vectors make an obtuse angle)

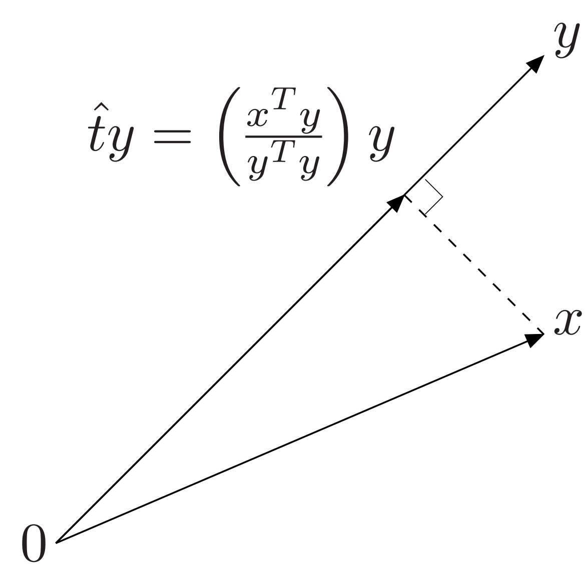
Projection

projection of x on the line defined by nonzero y: the vector $\hat{t}y$ with

$$\hat{t} = \underset{t}{\operatorname{argmin}} \|x - ty\|$$

expression for $\hat{t}$:

$$\hat{t} = \frac{x^T y}{\|y\|^2} = \frac{\|x\| \cos\theta}{\|y\|}$$



Hyperplanes and halfspaces

hyperplane

solution set of one linear equation with nonzero coefficient vector a

$$a^T x = b$$

halfspace

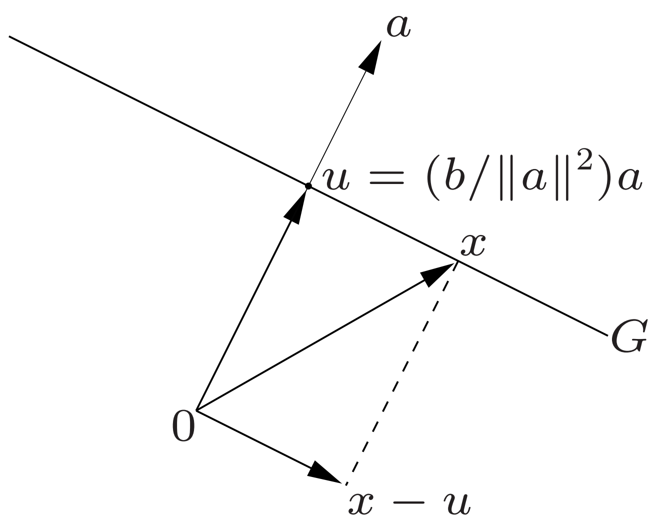
solution set of one linear inequality with nonzero coefficient vector a

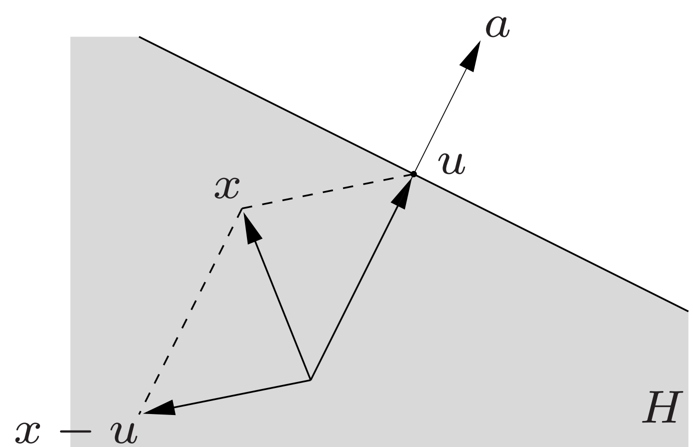
$$a^T x \leq b$$

a is the **normal vector**

Geometrical interpretation

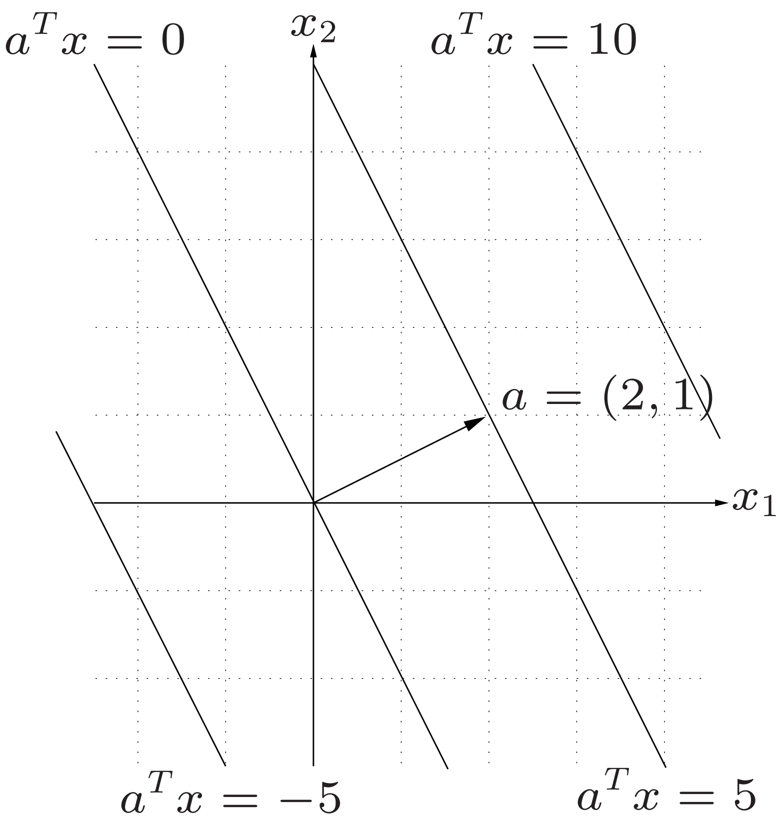
$$G = \{x \mid a^T x = b\} \qquad\qquad H = \{x \mid a^T x \leq b\}$$

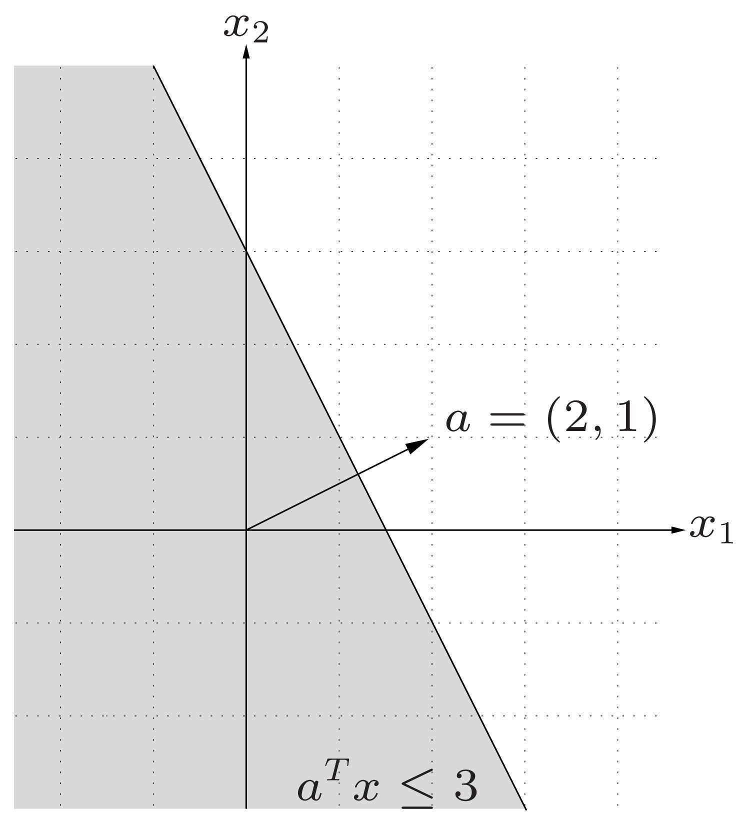




- the vector $u = (b/\|a\|^2)a$ satisfies $a^T u = b$
- x is in hyperplane G if $a^T(x - u) = 0$ ($x - u$ is orthogonal to a)
- x is in halfspace H if $a^T(x - u) \leq 0$ (angle $\angle(x - u, a) \geq \pi/2$)

Example

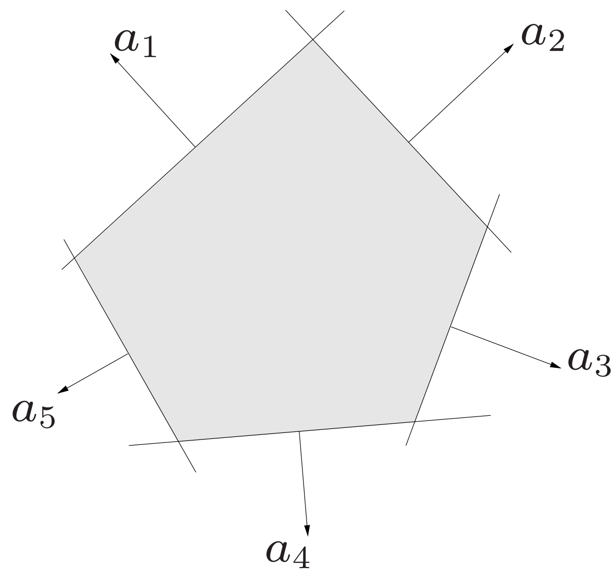




Polyhedron

solution set of a finite number of linear inequalities

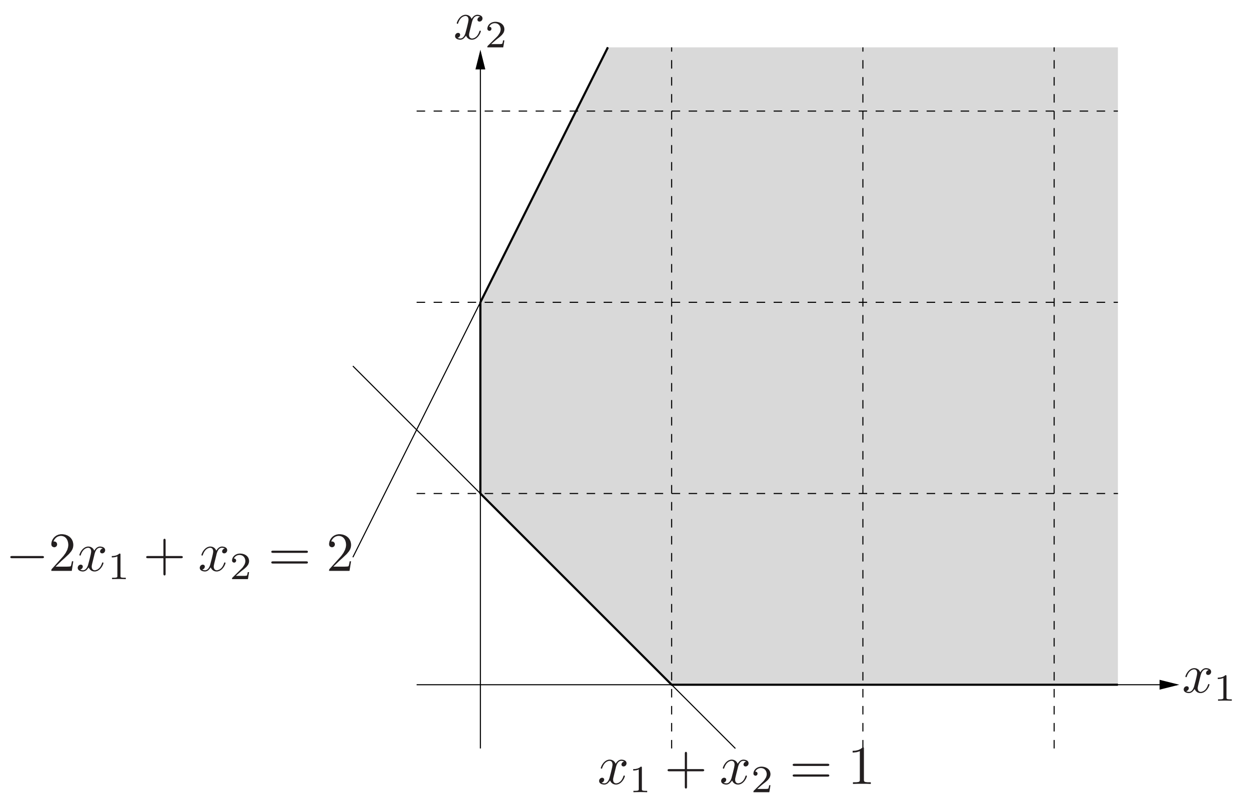
$$a_1^T x \leq b_1, \qquad a_2^T x \leq b_2, \qquad \ldots, \qquad a_m^T x \leq b_m$$



- intersection of a finite number of halfspaces
- in matrix notation: $Ax \leq b$ if A is a matrix with rows a_i^T
- can include equalities: $Fx = g$ is equivalent to $Fx \leq g$, $-Fx \leq -g$

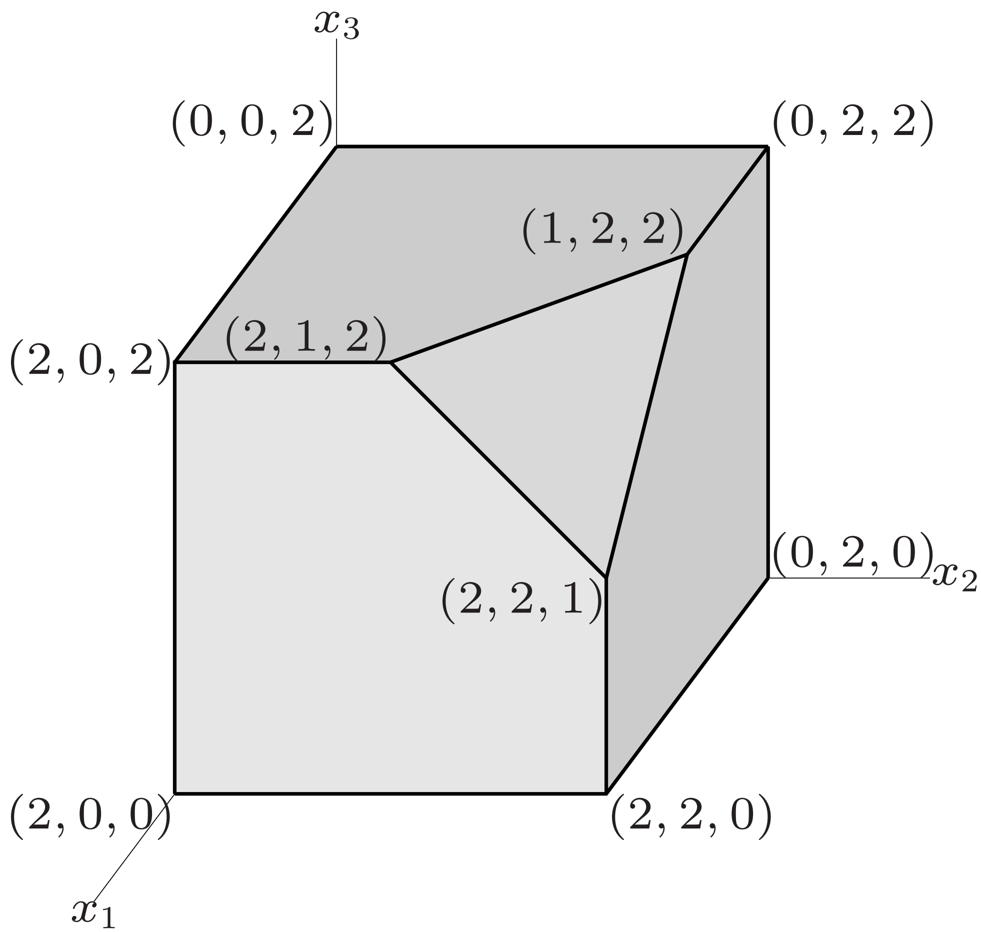
Example

$$x_1 + x_2 \geq 1, \qquad -2x_1 + x_2 \leq 2, \qquad x_1 \geq 0, \qquad x_2 \geq 0$$



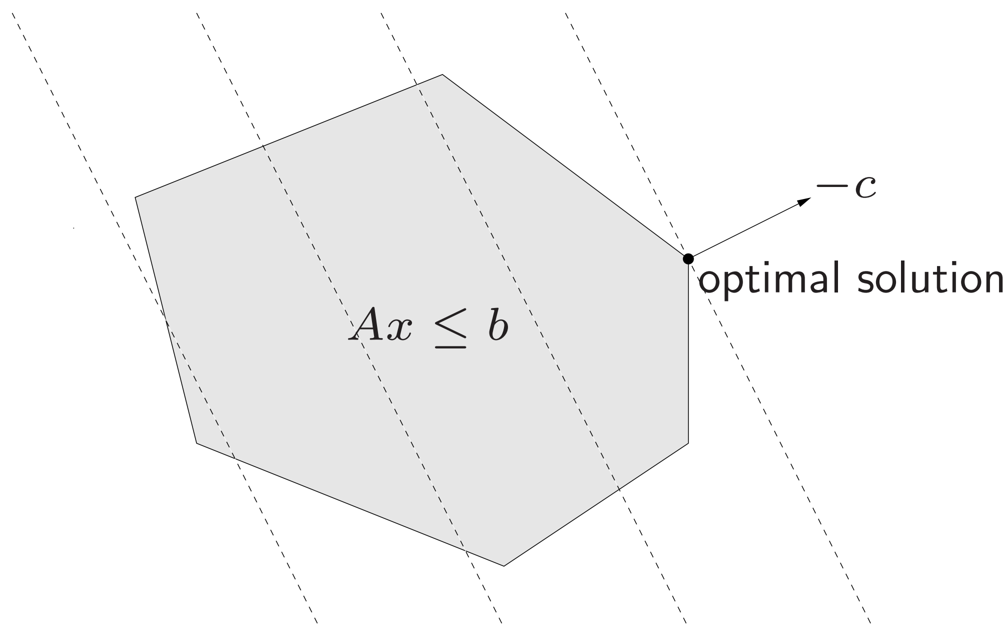
Example

$$0 \le x_1 \le 2, \qquad 0 \le x_2 \le 2, \qquad 0 \le x_3 \le 2, \qquad x_1 + x_2 + x_3 \le 5$$



Geometrical interpretation of LP

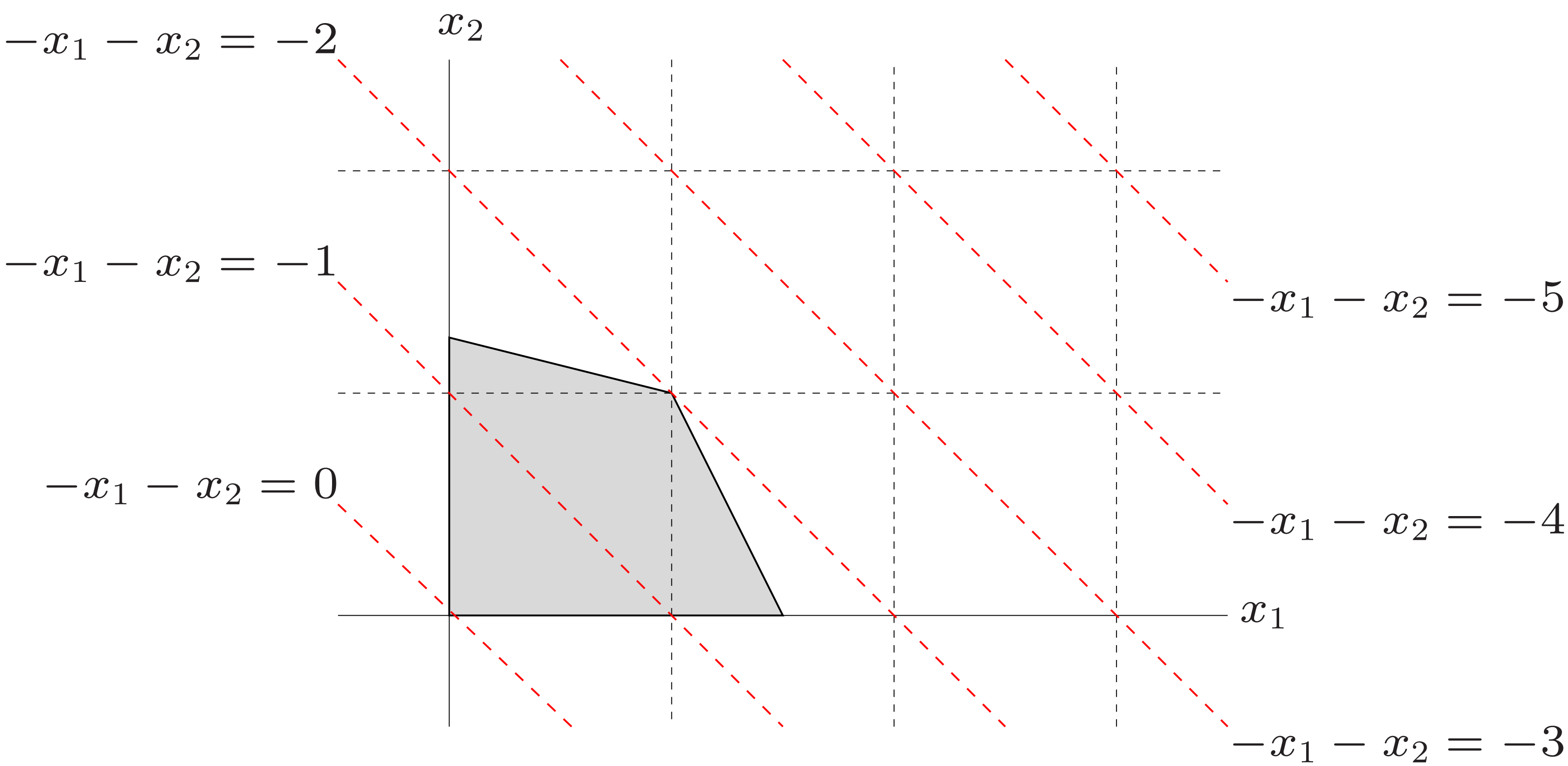
$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \le b \end{array}$$



dashed lines (hyperplanes) are level sets $c^T x = \alpha$ for different α

Example

$$\begin{array}{ll} \text{minimize} & -x_1 - x_2 \\ \text{subject to} & 2x_1 + x_2 \le 3 \\ & x_1 + 4x_2 \le 5 \\ & x_1 \ge 0, \quad x_2 \ge 0 \end{array}$$



optimal solution is $(1,1)$