Linear optimization

$$\begin{aligned} \min \quad & \sum_{j=1}^{n} c_j x_j \\ \text{s.t.} \quad & \sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i = 1, \dots, m \\ & \sum_{j=1}^{n} d_{ij} x_j = f_i, \quad i = 1, \dots, p \end{aligned}$$

- n optimization variables: $x_1, \dots, x_n$ (real scalars)
- problem data (parameters): the coefficients c_j, a_{ij}, b_i, d_{ij}, f_i
- $\sum_{j=1}^{n} c_j x_j$ is the cost function or objective function
- $\sum_{j=1}^{n} a_{ij} x_j \le b_i$ and $\sum_{j=1}^{n} d_{ij} x_j = f_i$ are inequality and equality constraints

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

Brief history

- 1940s (Dantzig, Kantorovich, Koopmans, von Neumann, . . .) foundations, motivated by economics and logistics problems
- 1947 (Dantzig): simplex algorithm
- 1972, (Klee and Minty): exponential example
- 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

von Neumann, 1903 – 1957

Dantzig, 1914 – 2005; Khachiyan, 1952 – 2005

Karmarkar, 1957 –

LO in inner-product notation

$$
\begin{array}{ll}
\min & \displaystyle\sum_{j=1}^{n} c_j x_j \\
\text{s.t.} & \displaystyle\sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, \ldots, m \\
& \displaystyle\sum_{j=1}^{n} d_{ij} x_j = f_i, \quad i = 1, \ldots, p
\end{array}
$$

inner-product notation

$$
\begin{array}{ll}
\min & c^T x \\
\text{s.t.} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \\
& d_i^T x = f_i, \quad i = 1, \ldots, p
\end{array}
$$

c, a_i, d_i are n-vectors:

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

LO in matrix notation

$$
\begin{aligned}
\min \quad & \sum_{j=1}^{n} c_j x_j \\
\text{s.t.} \quad & \sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i = 1, \ldots, m \\
& \sum_{j=1}^{n} d_{ij} x_j = f_i, \quad i = 1, \ldots, p
\end{aligned}
$$

matrix notation

$$
\begin{aligned}
\min \quad & c^T x \\
\text{s.t.} \quad & Ax \le b \\
& Dx = f
\end{aligned}
$$

- 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

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b \\ & Dx = f \end{aligned}$$

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

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**

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$
$$\|x\|_\infty = \max\big\{|x_1|, |x_2|, \cdots, |x_n|\big\}$$

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); $f(x) = 0 \Leftrightarrow x = 0$ (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

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

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

$$\begin{array}{l|ll} x^T y > 0 & \theta < \pi/2 & \text{(vectors make an acute angle)} \\ x^T y = 0 & \theta = \pi/2 & \text{(orthogonal vectors)} \\ x^T y < 0 & \theta > \pi/2 & \text{(vectors make an obtuse angle)} \end{array}$$

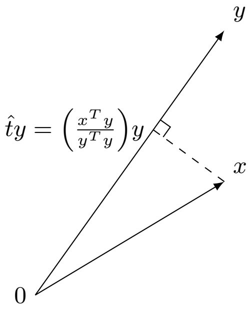
Projection

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

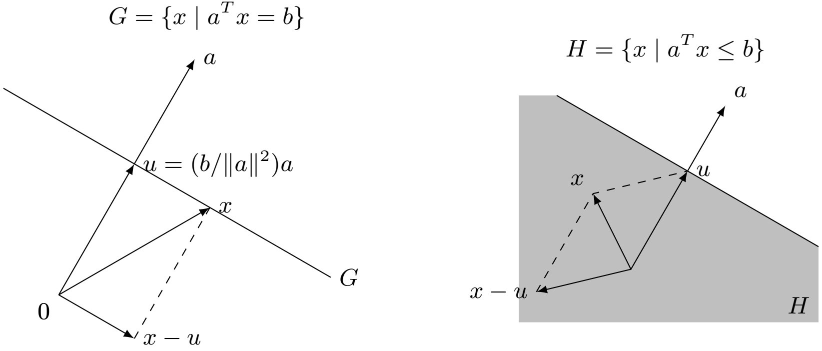
$$a^T x = b$$

halfspace
solution set of one linear inequality with nonzero coefficient vector a

$$a^T x \leq b$$

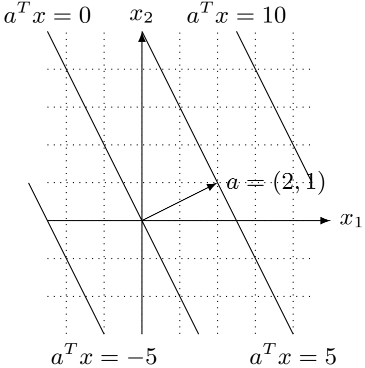
a is the **normal vector**

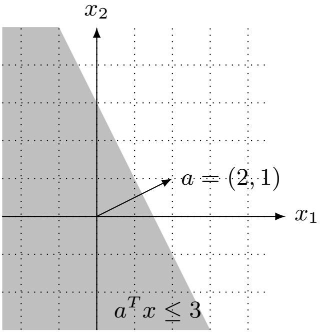
Geometrical interpretation



- 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 $a^T(x-u) \le 0$ (angle $\angle(x-u, a) \ge \pi/2$)

Example

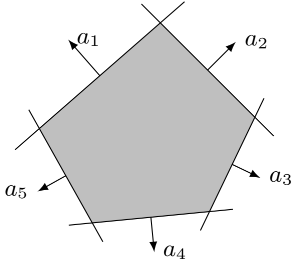




Polyhedron

solution set of a finite number of linear inequalities

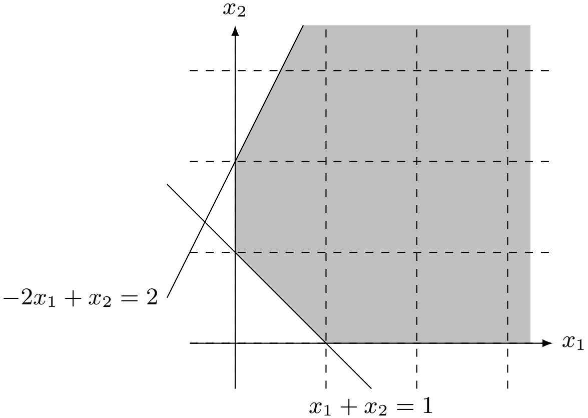
$$a_1^T \leq b_1 \qquad a_2^T x \leq b_2 \qquad \ldots \qquad a_m^T x \leq b_m$$



- 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: $Dx = f$ is equivalent to $Dx \leq f$, $Dx \geq f$

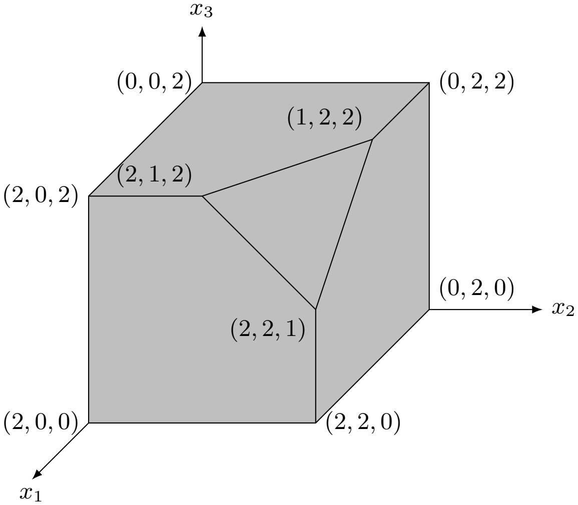
Example

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



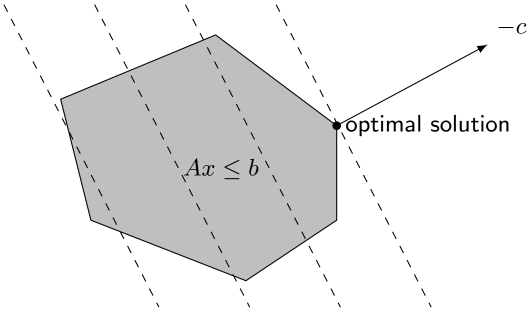
Example

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



Geometrical interpretation of LO

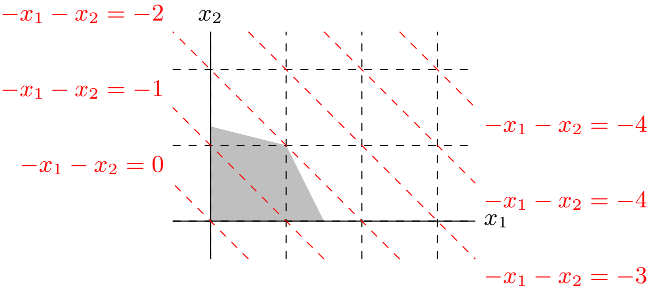
$$\begin{array}{rl} \min & c^T x \\ \text{s.t.} & Ax \le b \end{array}$$



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

Example

$$\begin{array}{rl} \min & -x_1 - x_2 \\ \text{s.t.} & 2x_1 + x_2 \le 3 \\ & x_1 + 4x_2 \le 5 \\ & x_1 \ge 0, \quad x_2 \ge 0 \end{array}$$



optimal solution is $(1, 1)$