Linear Optimization

Yinyu Ye

Department of Management Science and Engineering

Stanford University

Stanford, CA 94305, U.S.A.

http://www.stanford.edu/~yyye

Introduction to Linear Optimization

The field of optimization is concerned with the study of maximization and minimization of mathematical functions. Very often the arguments of (i.e., variables or unknowns in) these functions are subject to side conditions or constraints. By virtue of its great utility in such diverse areas as applied science, engineering, economics, finance, medicine, and statistics, optimization holds an important place in the practical world and the scientific world. Indeed, as far back as the Eighteenth Century, the famous Swiss mathematician and physicist Leonhard Euler (1707-1783) proclaimed[a] that . . . nothing at all takes place in the Universe in which some rule of maximum or minimum does not appear.

---

[a] See Leonhardo Eulero, *Methodus Inviendi Lineas Curvas Maximi Minimive Proprietate Gaudentes*, Lausanne & Geneva, 1744, p. 245.

Linear Programming History

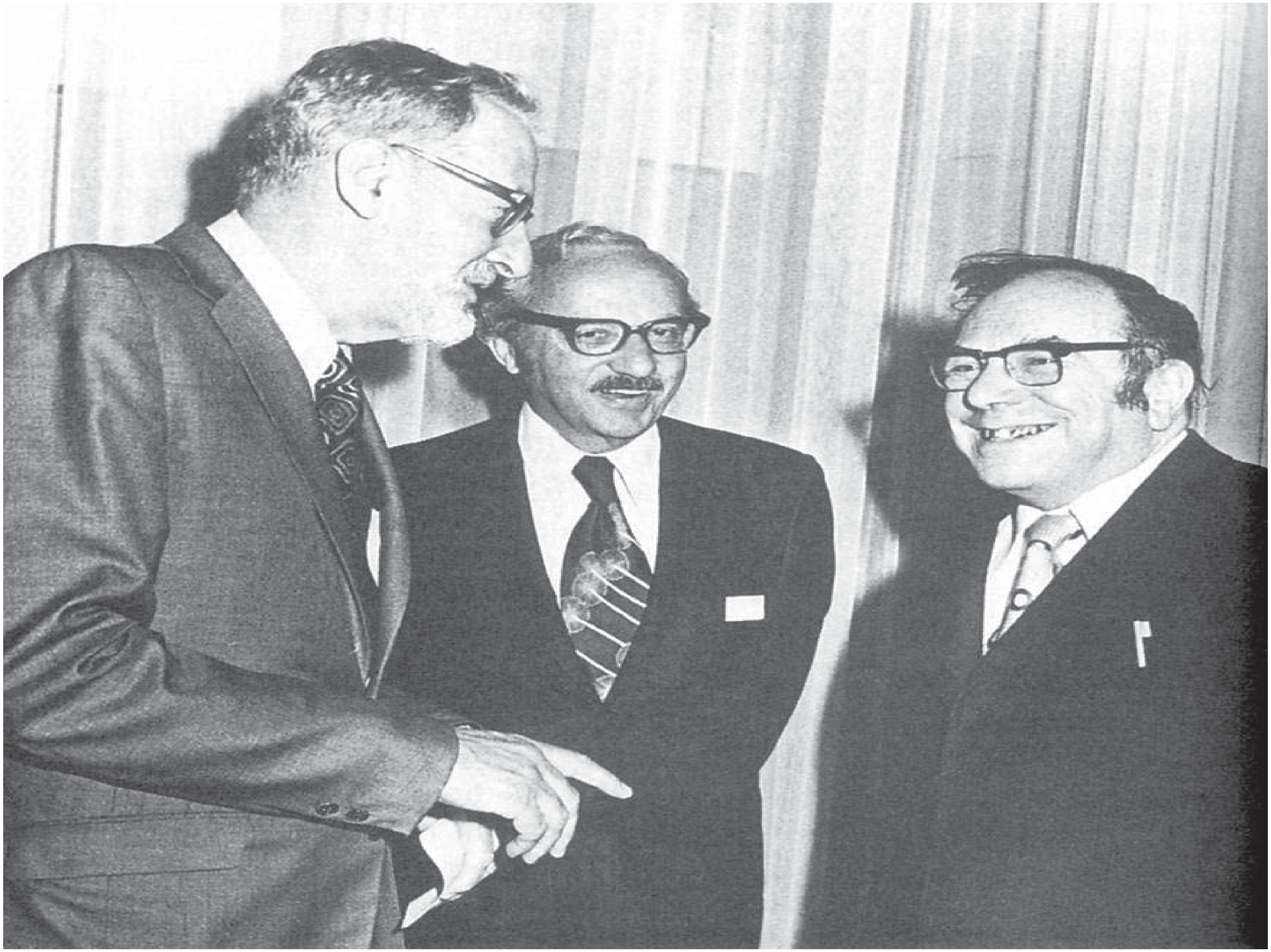

Figure 1: LP and Nobel Prize.

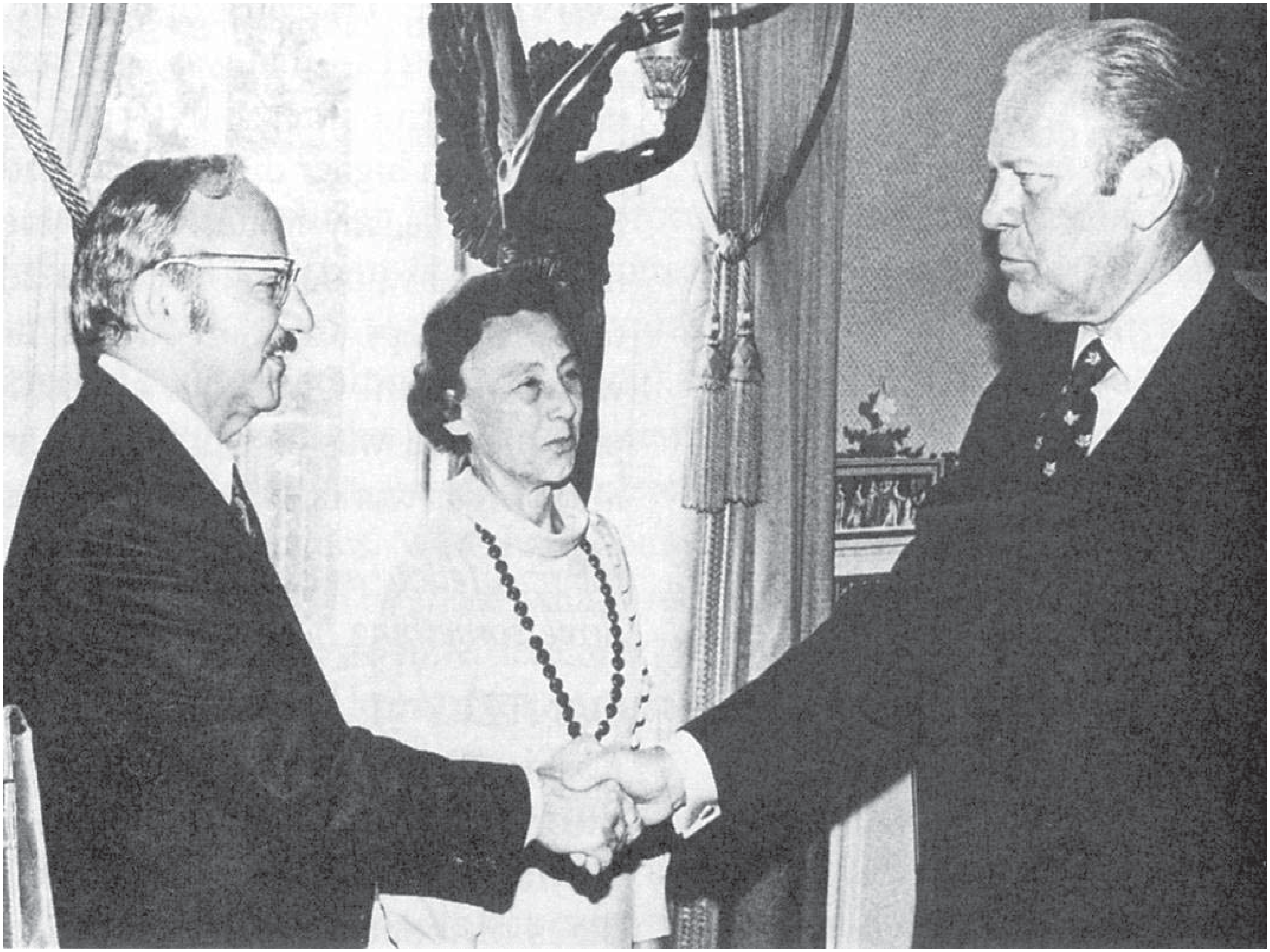

Figure 2: National Metal of Science.

Linear Programs and Extensions (in Standard Form)

Linear Programming

$$\begin{array}{lll} (LP) & \text{minimize} & \mathbf{c}^T\mathbf{x} \\ & \text{subject to} & A\mathbf{x} = \mathbf{b}, \\ & & \mathbf{x} \geq \mathbf{0}. \end{array}$$

Linearly Constrained Optimization Problem

$$\begin{array}{lll} (LCOP) & \text{minimize} & f(\mathbf{x}) \\ & \text{subject to} & A\mathbf{x} = \mathbf{b}, \\ & & \mathbf{x} \geq \mathbf{0}. \end{array}$$

Linear Complementarity Problem Find nonnegative vectors $\mathbf{x} \geq \mathbf{0}, \mathbf{s} \geq \mathbf{0}$ such that

$$(LCP) \qquad \mathbf{s} = M\mathbf{x} + \mathbf{q},$$
$$\mathbf{s}^T\mathbf{x} = 0.$$

Conic Linear Programming

$$\begin{array}{lll} (CLP) & \text{minimize} & \mathbf{c} \bullet \mathbf{x} \\ & \text{subject to} & \mathbf{a}_i \bullet \mathbf{x} = b_i,\ i = 1, ..., m, \\ & & \mathbf{x} \in K, \end{array}$$

where K is a closed convex cone.

CLP: LP, SOCP, and SDP Examples

$$\begin{array}{ll} \text{minimize} & 2x_1 + x_2 + x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 1, \\ & (x_1; x_2; x_3) \geq \mathbf{0}; \end{array}$$

$$\begin{array}{ll} \text{minimize} & 2x_1 + x_2 + x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 1, \\ & \sqrt{x_2^2 + x_3^2} \leq x_1. \end{array}$$

$$\begin{array}{ll} \text{minimize} & 2x_1 + x_2 + x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 1, \\ & \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq \mathbf{0}, \end{array}$$

LP Terminology

- decision variable/activity, data/parameter
- objective/goal/target, coefficient vector
- constraint/limitation/requirement, satisfied/violated
- equality/inequality constraint, direction of inequality, non-negativity
- constraint matrix/the right-hand side
- feasible/infeasible solution, interior feasible solution
- optimizers and optimum values
- active constraint (binding constraint), inactive constraint, redundant constraint

Linear Programming Facts

- The feasible region is a convex polyhedron.
- Every linear program is either feasible/bounded, feasible/unbounded, or infeasible.
- If feasible/bounded, every local optimizer is global and all optimizers form a convex polyhedron set.
- All optimizers are on the boundary of the feasible region.
- If the feasible region has an extreme point, then there must be an extreme optimizer.
- LP possesses efficient algorithms in both practice and theory (polynomial-time).

Linear Optimization Model and Formulation

- Sort out data and parameters from the verbal description
- Define the set of decision variables
- Formulate the linear objective function of data and decision variables
- Set up linear equality and inequality constraints

LP Example: Parimutuel Call Auction Mechanism I

Given m potential states that are mutually exclusive and exactly one of them will be realized at the maturity.

An order is a bet on one or a combination of states, with a price limit (the maximum price the participant is willing to pay for one unit of the order) and a quantity limit (the maximum number of units or shares the participant is willing to accept).

A contract on an order is a paper agreement so that on maturity it is worth a notional \$$1$ dollar if the order includes the winning state and worth \$$0$ otherwise.

There are n orders submitted now.

Parimutuel Call Auction Mechanism II: order data

The ith order is given as $(\mathbf{a}_{i\cdot} \in R^m_+,\ \pi_i \in R_+,\ q_i \in R_+)$: $\mathbf{a}_{i\cdot}$ is the betting indication row vector where each component is either 1 or 0

$$\mathbf{a}_{i\cdot} = (a_{i1},\ a_{i2},\ ...,\ a_{im})$$

where 1 is winning state and 0 is non-winning state; π_i is the price limit for one unit of such a contract, and q_i is the maximum number of contract units the better like to buy.

Parimutuel Call Auction Mechanism III: order fills

Let x_i be the number of units or shares awarded to the ith order. Then, the ith bidder will pay the amount $\pi_i \cdot x_i$ and the total amount collected would be $\pi^T \mathbf{x} = \sum_i \pi_i \cdot x_i$.

If the jth state is the winning state, then the auction organizer need to pay the winning bidders

$$\left(\sum_{i=1}^{n} a_{ij} x_i \right) = \mathbf{a}_{.j}^T \mathbf{x}$$

where column vector

$$\mathbf{a}_{.j} = (a_{1j};\ a_{2j};\ \ldots;\ a_{nj})$$

The question is, how to decide $\mathbf{x} \in R^n$, that is, how to fill the orders.

Parimutuel Call Auction Mechanism IV: worst-case profit maximization

$$\begin{array}{lrl} \max & \pi^T\mathbf{x} - \max_j\{\mathbf{a}_{\cdot j}^T\mathbf{x}\} & \\ \text{s.t.} & \mathbf{x} & \leq \mathbf{q}, \\ & \mathbf{x} & \geq \mathbf{0}. \end{array}$$

$$\begin{array}{lrl} \max & \pi^T\mathbf{x} - \max(A^T\mathbf{x}) & \\ \text{s.t.} & \mathbf{x} & \leq \mathbf{q}, \\ & \mathbf{x} & \geq \mathbf{0}. \end{array}$$

This is NOT a linear program.

Parimutuel Call Auction Mechanism V: linear programming

However, the problem can be rewritten as

$$\begin{array}{lrl} \max & \pi^T \mathbf{x} - y & \\ \text{s.t.} & A^T \mathbf{x} - \mathbf{e} \cdot y & \leq \mathbf{0}, \\ & \mathbf{x} & \leq \mathbf{q}, \\ & \mathbf{x} & \geq \mathbf{0}, \end{array}$$

where $\mathbf{e}$ is the vector of all ones. This is a linear program.

$$\begin{array}{lrl} \max & \pi^T \mathbf{x} - y & \\ \text{s.t.} & A^T \mathbf{x} - \mathbf{e} \cdot y + s_0 & = \mathbf{0}, \\ & \mathbf{x} + \mathbf{s} & = \mathbf{q}, \\ & (\mathbf{x}, s_0, \mathbf{s}) & \geq \mathbf{0}, \quad y \text{ free}, \end{array}$$

Real n-Space; Euclidean Space

- $\mathcal{R}, \quad \mathcal{R}_+, \quad \text{int}\,\mathcal{R}_+$
- $\mathcal{R}^n, \quad \mathcal{R}^n_+, \quad \text{int}\,\mathcal{R}^n_+$
- $\mathbf{x} \geq \mathbf{y}$ means $x_j \geq y_j$ for $j = 1, 2, ..., n$
- $\mathbf{0}$: all zero vector; and $\mathbf{e}$: all one vector
- Inner-Product:

$$\mathbf{x} \bullet \mathbf{y} := \mathbf{x}^T\mathbf{y} = \sum_{j=1}^{n} x_j y_j$$

- Norm: $\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^T\mathbf{x}}$, $\|\mathbf{x}\|_\infty = \max\{|x_1|, |x_2|, ..., |x_n|\}$, $\|\mathbf{x}\|_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p}$
- The dual of the p norm, denoted by $\|.\|^*$, is the q norm, where $\frac{1}{p} + \frac{1}{q} = 1$

- Column vector:

$$\mathbf{x} = (x_1; x_2; \ldots; x_n)$$

and row vector:

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)$$

- A set of vectors $\mathbf{a}_1, ..., \mathbf{a}_m$ is said to be linearly dependent if there are scalars $\lambda_1, ..., \lambda_m$, not all zero, such that the linear combination

$$\sum_{i=1}^{m} \lambda_i \mathbf{a}_i = \mathbf{0}$$

- A linearly independent set of vectors that span R^n is a basis.

Matrices

- $\mathcal{R}^{m\times n}$, $\mathbf{a}_{i.}$, $\mathbf{a}_{.j}$, a_{ij}
- A_I denotes the submatrix of A whose rows belong to I, A_J denotes the submatrix of whose columns belong to J, A_{IJ}.
- $\mathbf{0}$: all zero matrix, and I: the identity matrix
- $\mathcal{N}(A)$, $\mathcal{R}(A)$:

  Theorem 1 *Each linear subspace of $\mathcal{R}^n$ is generated by finitely many vectors, and is also the intersection of finitely many hyperplanes; that is, for each linear subspace of L of $\mathcal{R}^n$ there are matrices A and C such that $L = \mathcal{N}(A) = \mathcal{R}(C)$.*
- $\det(A)$, $\text{tr}(A)$

- Inner Product:

$$A \bullet B = \text{tr} A^T B = \sum_{i,j} a_{ij} b_{ij}$$

- The operator norm of $\|A\|$:

$$\|A\|^2 := \max_{\mathbf{0} \neq \mathbf{x} \in \mathcal{R}^n} \frac{\|A\mathbf{x}\|^2}{\|\mathbf{x}\|^2}$$

- Sometimes we use $X = \text{diag}(\mathbf{x})$
- Eigenvalues and eigenvectors

$$A\mathbf{v} = \lambda \cdot \mathbf{v}$$

Symmetric Matrices

- $\mathcal{S}^n$
- The Frobenius norm:

$$\|X\|_f = \sqrt{\text{tr} X^T X} = \sqrt{X \bullet X}$$

- Positive Definite (PD): $Q \succ \mathbf{0}$ iff $\mathbf{x}^T Q \mathbf{x} > 0, \quad \text{for all} \quad \mathbf{x} \neq \mathbf{0}$
- Positive SemiDefinite (PSD): $Q \succeq \mathbf{0}$ iff $\mathbf{x}^T Q \mathbf{x} \geq 0, \quad \text{for all} \quad \mathbf{x}$
- PSD set: $\mathcal{S}^n_+, \quad \text{int}\, \mathcal{S}^n_+$

Known Inequalities

- Cauchy-Schwarz: given $\mathbf{x}, \mathbf{y} \in \mathcal{R}^n$, $\mathbf{x}^T\mathbf{y} \le \|\mathbf{x}\|\|\mathbf{y}\|.$
- Triangle: given $\mathbf{x}, \mathbf{y} \in \mathcal{R}^n$, $\|\mathbf{x}+\mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|.$
- Arithmetic-geometric mean: given $\mathbf{x} \in \mathcal{R}^n_+$,

$$\frac{\sum x_j}{n} \ge \left(\prod x_j\right)^{1/n}.$$

Hyper plane and Half-spaces

$$H = \{\mathbf{x} : \mathbf{a}\mathbf{x} = \sum_{j=1}^{n} a_j x_j = b\}$$

$$H^+ = \{\mathbf{x} : \mathbf{a}\mathbf{x} = \sum_{j=1}^{n} a_j x_j \leq b\}$$

$$H^- = \{\mathbf{x} : \mathbf{a}\mathbf{x} = \sum_{j=1}^{n} a_j x_j \geq b\}$$

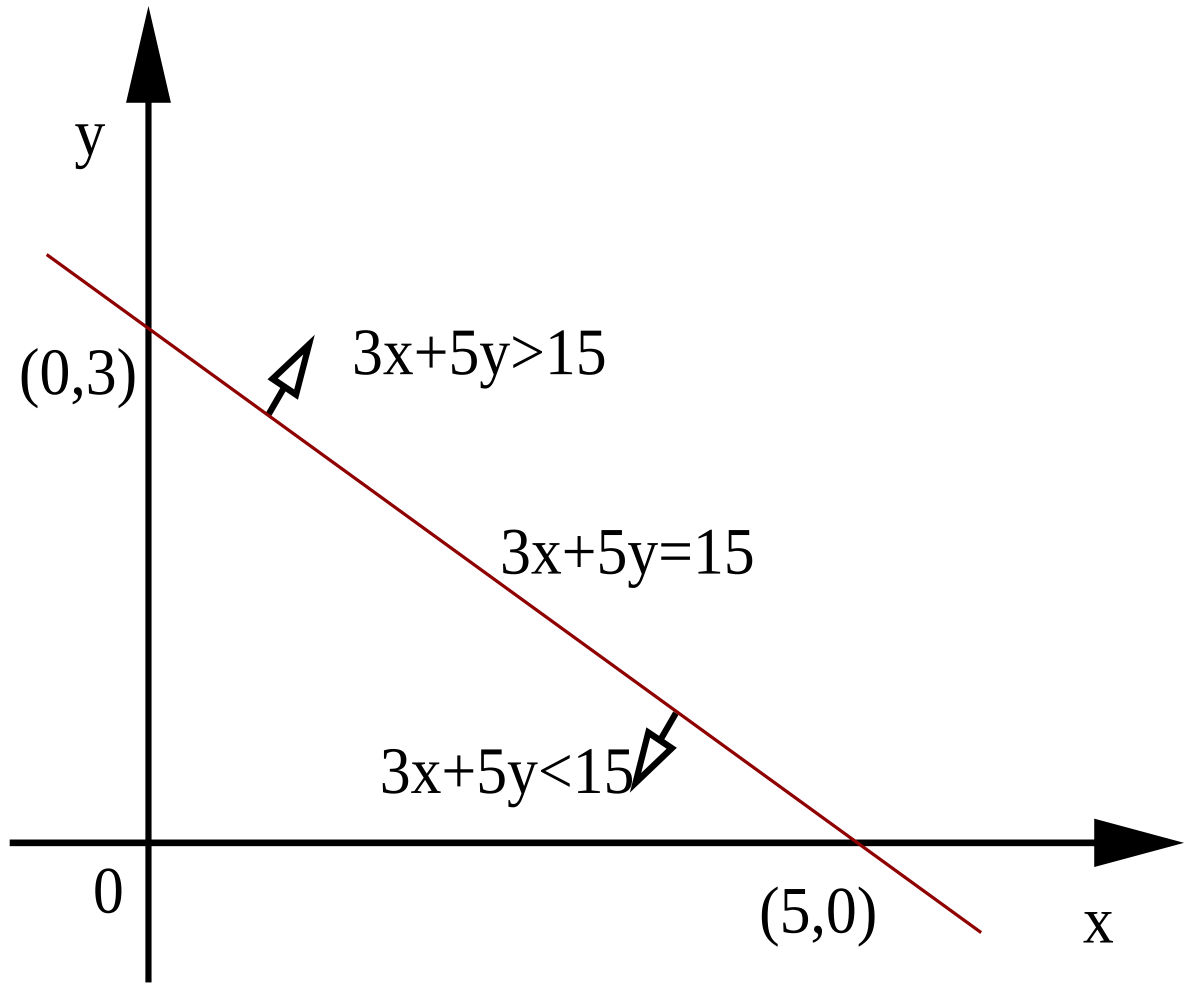

Figure 3: Plane and Half-Spaces

System of Linear Equations

Solve for $\mathbf{x} \in \mathcal{R}^n$ from:

$$\left.\begin{array}{rcl} \mathbf{a}_1\mathbf{x} & = & b_1 \\ \mathbf{a}_2\mathbf{x} & = & b_2 \\ \ldots & . & . \\ \mathbf{a}_m\mathbf{x} & = & b_m \end{array}\right. \Rightarrow A\mathbf{x} = \mathbf{b}$$

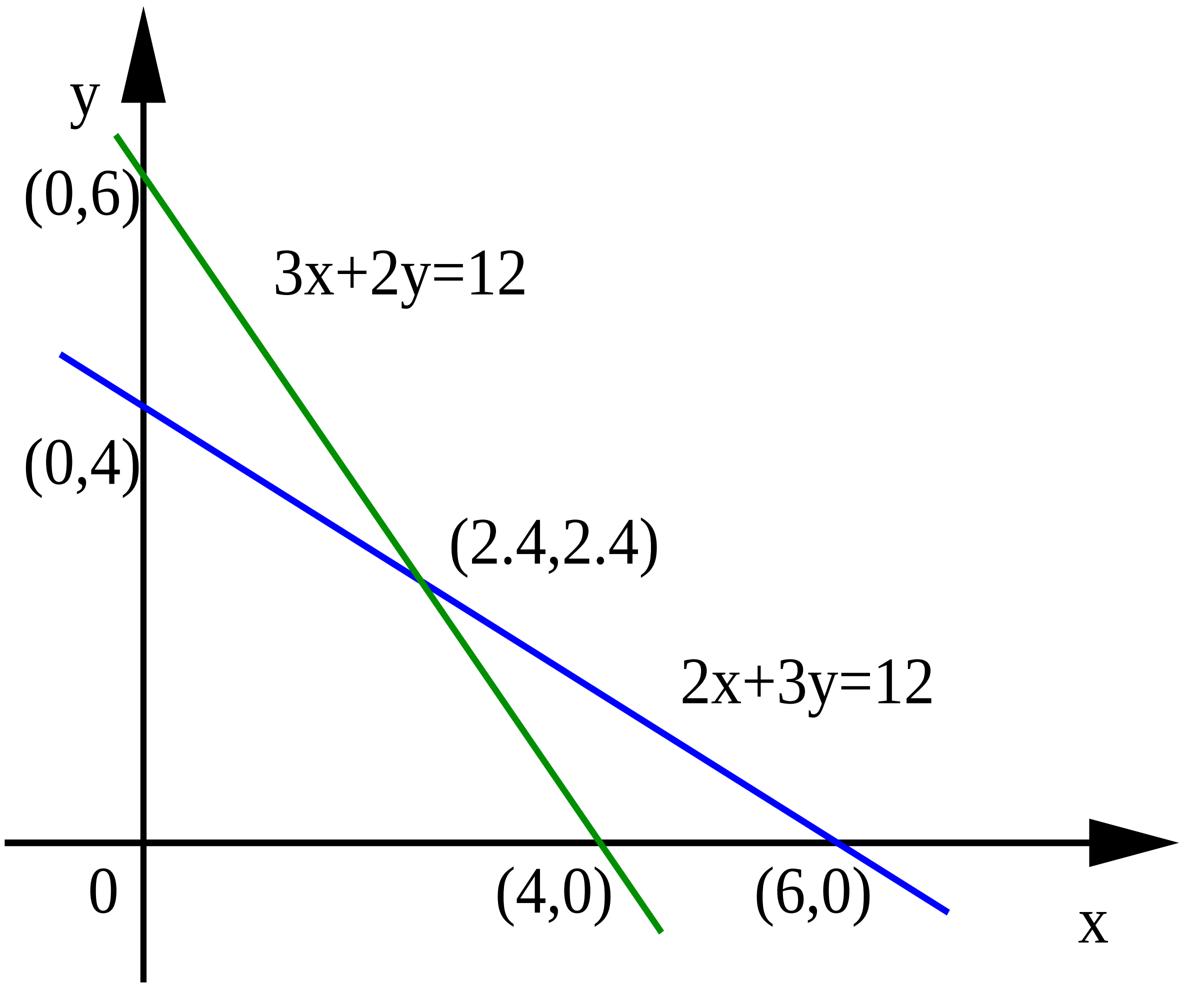

Figure 4: System of Linear Equations

Fundamental Theorem of Linear Equations

Theorem 2 *Given* $A \in \mathcal{R}^{m\times n}$ *and* $\mathbf{b} \in \mathcal{R}^m$*, the system* $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}\}$ *has a solution if and only if that* $A^T\mathbf{y} = \mathbf{0}$ *and* $\mathbf{b}^T\mathbf{y} \neq 0$ *has no solution.*

A vector $\mathbf{y}$, with $A^T\mathbf{y} = 0$ and $\mathbf{b}^T\mathbf{y} \neq 0$, is called an infeasibility certificate for the system.

Example Let $A = (1; -1)$ and $\mathbf{b} = (1; 1)$. Then, $\mathbf{y} = (1/2; 1/2)$ is an infeasibility certificate.

Alternative systems: $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}\}$ and $\{\mathbf{y} : \ A^T\mathbf{y} = \mathbf{0}, \ \mathbf{b}^T\mathbf{y} \neq 0\}$.

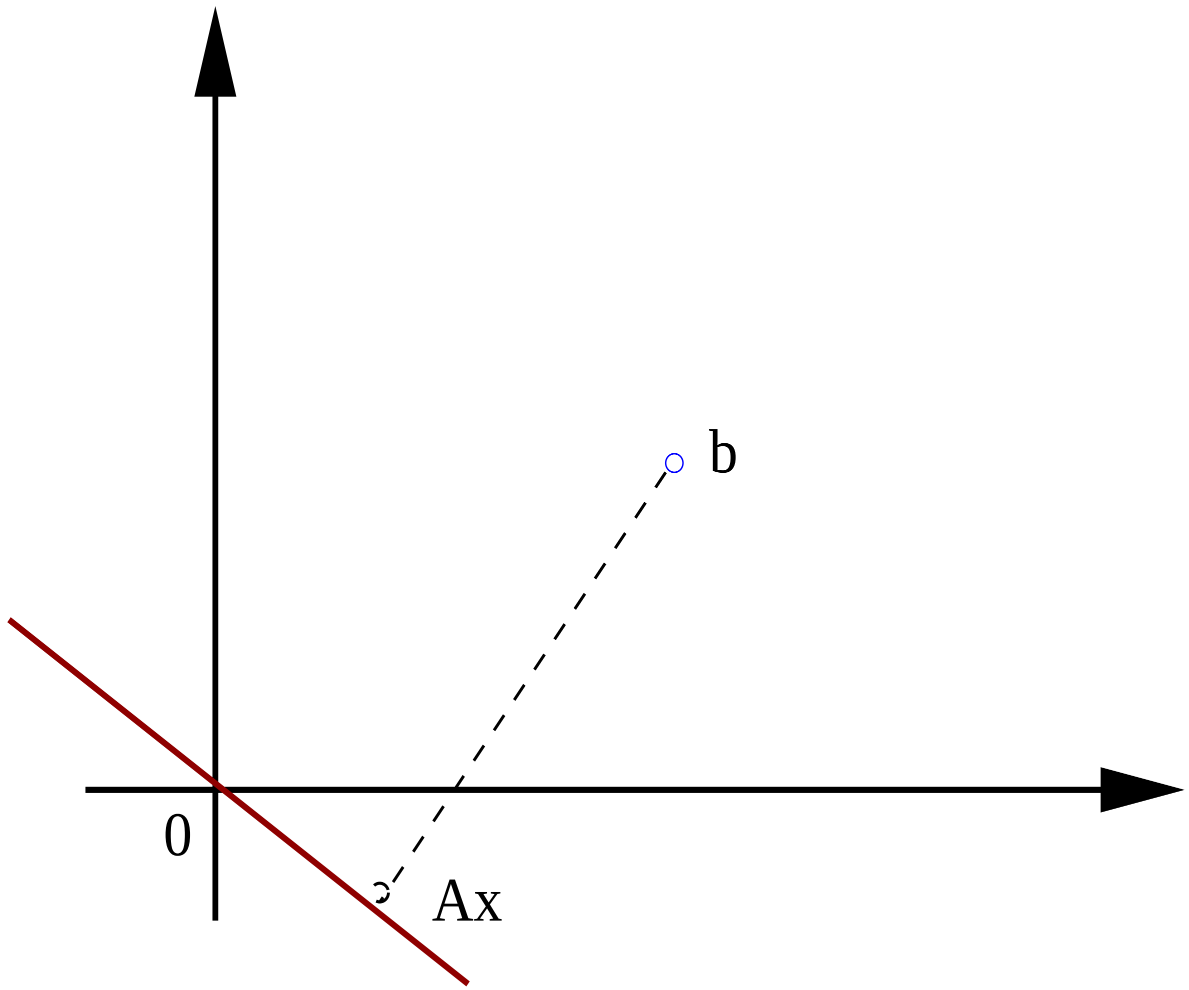

Figure 5: $\mathbf{b}$ is not in the set $\{A\mathbf{x}:\ \mathbf{x}\}$, and $\mathbf{y}$ is the distance vector from $\mathbf{b}$ to the set.

Gaussian elimination method

$$\begin{pmatrix} a_{11} & A_{1.} \\ 0 & A' \end{pmatrix} \begin{pmatrix} x_1 \\ x' \end{pmatrix} = \begin{pmatrix} b_1 \\ b' \end{pmatrix}.$$

$$A = L \begin{pmatrix} U & C \\ 0 & 0 \end{pmatrix}$$