Operations research meets data science: principles, models and algorithms

Immanuel Bomze, University of Vienna

Dip. Matematica@Univ. Padova 05 November 2025

Single-objective optimization problem

$$\min\{f(\mathbf{x}) : g_i(\mathbf{x}) \le 0\,,\ \text{all } i = 1,\ldots,m\}$$

Extension: view $g_i(\mathbf{x}) = G(\mathbf{x}, i) \le 0$, all $i = 1,\ldots,m$

Instead of number/index i now any symbol $\mathbf{u}$, look at

$$G(\mathbf{x}, \mathbf{u}) \le 0\,, \quad \text{all } \mathbf{u} \in \mathcal{U}$$

leads to *semi-infinite optimization*.

Quite general, allows for many approaches, see later.

Quick outlook on soft constraints

They come in many flavors, e.g. in last context:

$$G(\mathbf{x},\mathbf{u}) \leq 0 \quad \text{for sufficiently many } \mathbf{u} \in \mathcal{U}$$

leads to probability/chance constraints (service level guarantee).

Or *pricing* constraints $g_i(\mathbf{x})$ by p_i:

consider new objective

$$L(\mathbf{x};\mathbf{p}) := f(\mathbf{x}) + \sum_{i=1}^{m} p_i g_i(\mathbf{x})$$

leads to duality, p_i are *multipliers* or *dual variables*.

Again, a cliffhanger ...

The workhorse of optimization ...

... is the *Linear optimization Problem (LP)*:
a single linear objective, cost coefficients c_i:

$$\mathbf{c}^\top \mathbf{x} = c_1 x_1 + \cdots + c_n x_n \to \text{min !}$$

subject to m linear constraints

$$\mathsf{A}\mathbf{x} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \mathbf{b}$$

plus sign constraints $x_1 \geq 0, \ldots, x_n \geq 0$ ($\mathbf{x} \in \mathbb{R}^n_+$).

... an LP in *standard form*.

Alternative representation: LP in *canonical form*:

$$\text{min}\left\{\mathbf{c}^\top \mathbf{x} : \mathsf{A}\mathbf{x} \geq \mathbf{b}\,,\ \mathbf{x} \in \mathbb{R}^n_+\right\}\,.$$

Why workhorse ?

Why is LP figuring so prominently ?

- A simple formal structure
- Efficient scalable algorithms available
- Powerful interplay of different contexts:

  algebra/computing vs. geometry/interpretation

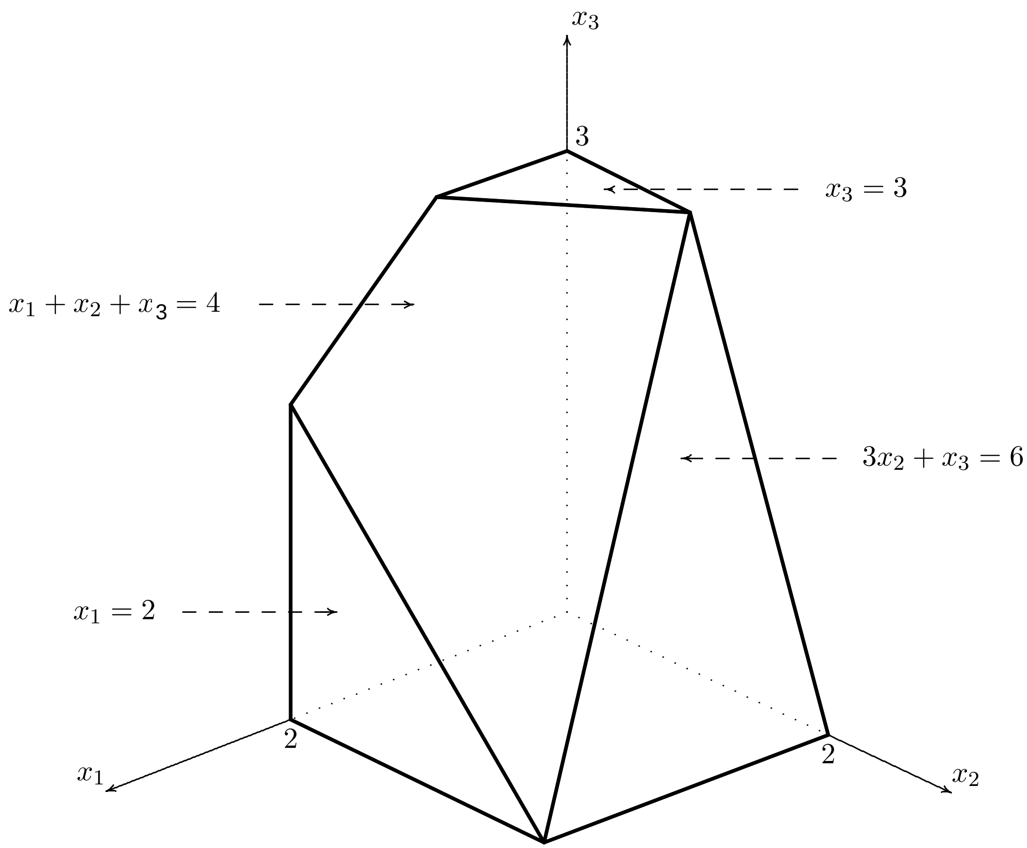
Crucial: role of (single) linear equation, e.g.,

$$\mathbf{p}^\top \mathbf{x} = p_1 x_1 + \cdots + p_n x_n = b\,,$$

geometrically describing a *(hyper-)plane* $H = \left\{\mathbf{x} : \mathbf{p}^\top \mathbf{x} = b\right\}$, separating the (decision) space into two parts (half-spaces),

$$H_+ = \left\{\mathbf{x} : \mathbf{p}^\top \mathbf{x} \geq b\right\} \quad \text{and} \quad H_- = \left\{\mathbf{x} : \mathbf{p}^\top \mathbf{x} \leq b\right\}\,.$$

Example from Bomze/Grossmann 1983



$$
\begin{aligned}
x_1 + \quad x_2 + x_3 &\le 4 \\
x_1 \qquad\qquad\quad &\le 2 \\
x_3 &\le 3 \\
3x_2 + x_3 &\le 6 \\
\\
x_1, \quad x_2, \quad x_3 &\ge 0
\end{aligned}
$$

Optimal solutions to LPs at vertices

Consider LP min $\left\{\mathbf{c}^\top \mathbf{x} : \mathsf{A}\mathbf{x} = \mathbf{b}\,,\ \mathbf{x} \in \mathbb{R}^n_+\right\}$.

Special role of vertices of the feasible set M:

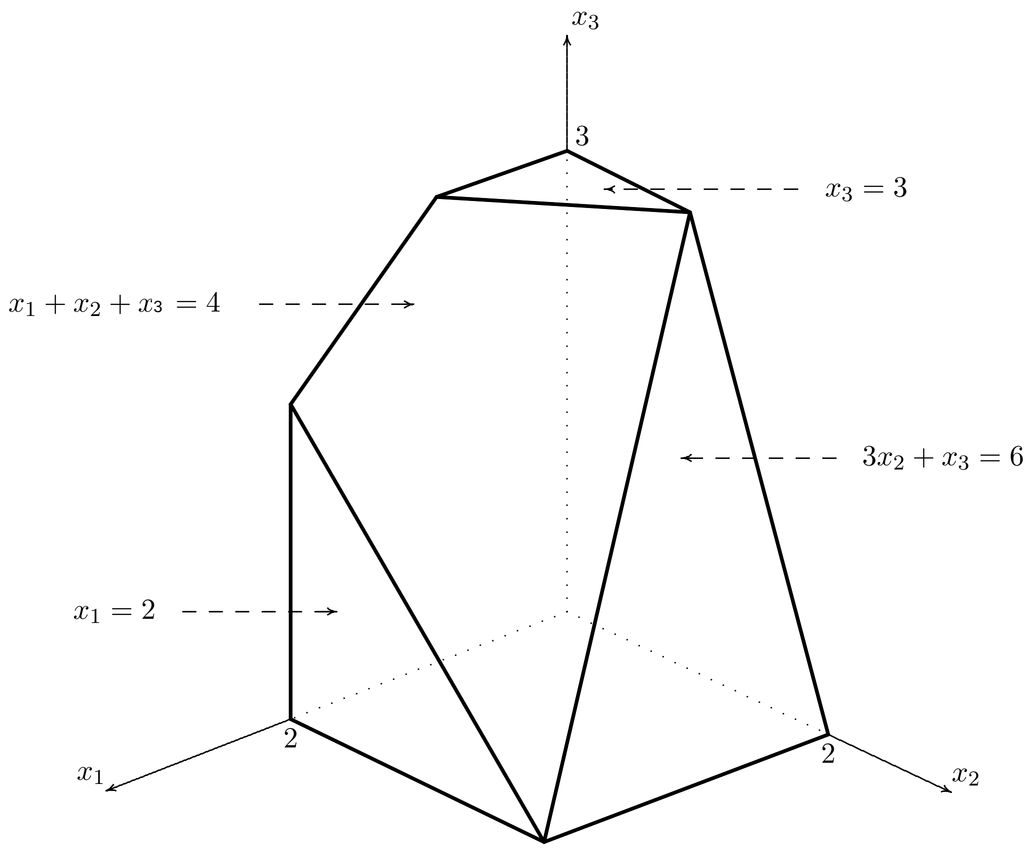
if there is an optimal solution, it must be attained at a vertex *(extreme point)* $\mathbf{x}^* \in M$, which means that $\mathbf{x}^*$ is not midpoint of any line segment in M.

Geometry: see next slide.

Algebra: for no $\mathbf{x}, \mathbf{y}$ in M different from $\mathbf{x}^*$, have $\mathbf{x}^* = \frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y}$.

Why ? Suppose $\mathbf{c}^\top \mathbf{x}^* \le \mathbf{c}^\top \mathbf{x} < \mathbf{c}^\top \mathbf{y}$ ...

Example from Bomze/Grossmann 1983



How to decide extremality ?

Theorem. $\mathbf{x}^*$ is vertex of $M = \left\{\mathbf{x} \in \mathbb{R}^n_+ : \mathsf{A}\mathbf{x} = \mathbf{b}\right\}$ if and only if there are linearly independent columns $\mathbf{a}_1, \ldots \mathbf{a}_m$ of A such that

$$\sum_{j=1}^{m} x_j^* \mathbf{a}_j = \mathbf{b}\,, \qquad\qquad (+)$$

and all other $x_k^* = 0$.

Above equation (+) has a **unique solution** $\mathbf{x}^*$, which is *sparse*, i.e., many $(n-m)$ coordinates $x_k^* = 0$, a few (m) may be nonzero.

Resulting $\mathbf{x}^*$ is feasible if and only if all $x_j^* \geq 0$.

Idea of Simplex Algorithm for solving LPs

Iteratively pass from vertex $\mathbf{x}^*$ to neighboring vertex $\mathbf{y}^*$, to improve objective, $\mathbf{c}^\top \mathbf{y}^* < \mathbf{c}^\top \mathbf{x}^*$.

$\mathbf{x}^*$ and $\mathbf{y}^*$ have almost same sparsity pattern, typically

one non-zero $x_j^* > 0$ gets zero $y_j^* = 0$, and
one zero $x_k^* = 0$ gets positive $y_k^* > 0$.

Caution: most remaining nonzeroes change too, $0 < x_j^* \neq y_j^* > 0$.

Calculation of $\mathbf{y}^*$ done by *tableau* step, similar to solving lin.eq.s.

Invention of the Simplex Algorithm for LP ...

... independently in two hemispheres $\approx$ 1940 ...

IFORS' Operational Research Hall of Fame
Leonid Vitaliyevich Kantorovich

Co-inventor of linear programming.

Born: 19 January 1912 in St. Petersburg, Russia.

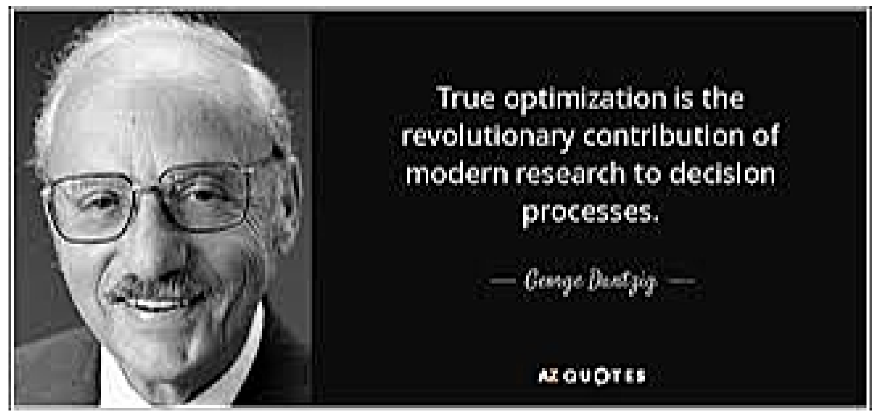
Died: 7 April 1986 in USSR.

Education: D.Sc. Mathematics, Leningrad University, 1935.

Key positions: Professor of Mathematics, Leningrad University, 1934–1960; Chair of Mathematics and Economics, Siberian branch of the USSR Academy of Sciences, 1961–1971; Director of Research, Moscow's Institute of National Economic Planning, 1971–1976.

Awards: State Prize, 1949; Lenin Prize, 1965; Nobel Prize in Economics, 1975; Silver Medal of the Operational Research Society, 1986.

Leonid Vitaliyevich Kantorovich (1912–1986) invented linear programming and published his initial results in 1939. This was eight years before its independent re-discovery by George Dantzig.



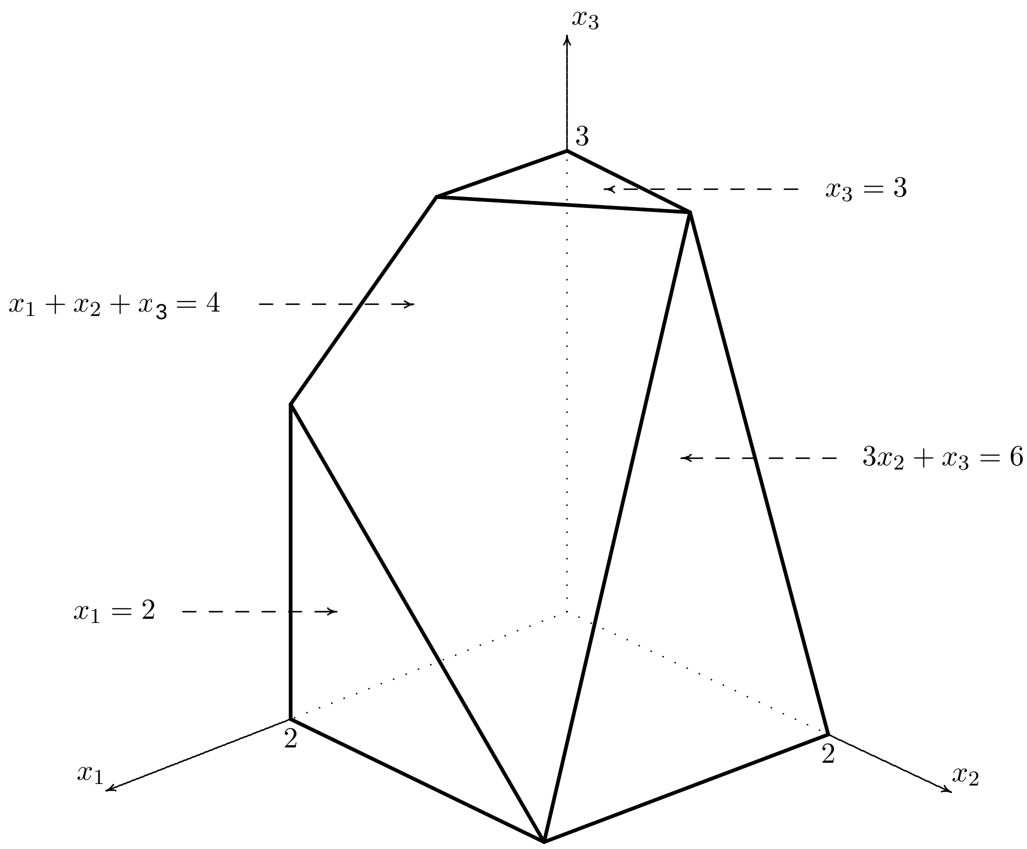
George Bernard Dantzig
(Portland 1914 – Stanford 2005)

Inventors of the Simplex Algorithm met ...

https://www.math.spbu.ru/user/jvr/LVK_html/70_dkk.html (cut out)

Photo taken 1970 at IIASA, Laxenburg, Austria

Example from Bomze/Grossmann 1983



$$
\begin{aligned}
x_1 + \quad x_2 + x_3 &\leq 4 \\
x_1 \qquad\qquad\quad &\leq 2 \\
x_3 &\leq 3 \\
3x_2 + x_3 &\leq 6 \\
\\
x_1, \quad x_2, \quad x_3 &\geq 0
\end{aligned}
$$

Example, continued

Write LP in standard form: enrich $\mathbf{z}^\top = [\mathbf{x}^\top \,|\, \mathbf{s}^\top]$ by slacks $\mathbf{s} \in \mathbb{R}^4_+$,

$$\begin{aligned}
x_1 + \;\; x_2 + x_3 + s_1 \qquad\qquad\qquad\quad &= 4 \\
x_1 \qquad\qquad\qquad\quad + s_2 \qquad\qquad &= 2 \\
x_3 \qquad\qquad + s_3 \qquad &= 3 \\
3x_2 + x_3 \qquad\qquad\qquad\qquad + s_4 &= 6
\end{aligned}$$

In matrix form $\mathsf{A}\mathbf{z} = \mathbf{b}$ with

$$\mathsf{A} = \left[\begin{array}{ccc|cccc} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 & 1 \end{array}\right], \; \mathbf{b} = \left[\begin{array}{c} 4 \\ 2 \\ 3 \\ 6 \end{array}\right].$$

More generally ...

$$\mathsf{A} = [\,\mathsf{N}\,|\,\mathsf{B}\,] \quad \text{and} \quad \mathbf{z} = \begin{bmatrix} \mathbf{x} \\ -- \\ \mathbf{s} \end{bmatrix} \text{ feasible implies } \mathsf{A}\mathbf{z} = \mathbf{b}$$

which means (block multiplication is like ordinary multiplication!)

$$\mathsf{N}\mathbf{x} + \mathsf{B}\mathbf{s} = \mathbf{b} \quad \text{or} \quad \mathsf{B}\mathbf{s} = \mathbf{b} - \mathsf{N}\mathbf{x} \quad \text{or} \quad \mathbf{s} = \mathsf{B}^{-1}\mathbf{b} - \mathsf{B}^{-1}\mathsf{N}\mathbf{x}$$

which gives for a cost vector (objective) $\mathbf{c}^\top = [\,\mathbf{c}_\mathsf{N}^\top\,|\,\mathbf{c}_\mathsf{B}^\top\,]$

$$\mathbf{c}^\top\mathbf{z} = \mathbf{c}_\mathsf{N}^\top\mathbf{x} + \mathbf{c}_\mathsf{B}^\top\mathbf{s} = \mathbf{c}_\mathsf{B}^\top\mathsf{B}^{-1}\mathbf{b} + (\mathbf{c}_\mathsf{N}^\top - \mathbf{c}_\mathsf{B}^\top\mathsf{B}^{-1}\mathsf{N})\mathbf{x}\,.$$

If current vertex is $\mathbf{z}_{old}^\top = [\mathbf{o}^\top|\mathbf{s}_{old}^\top]$, value $\mathbf{c}^\top\mathbf{z}_{old} = \mathbf{c}_\mathsf{B}^\top\mathsf{B}^{-1}\mathbf{b}$.

Alternative feasible $\mathbf{z}_{new}$ has $\mathbf{c}^\top\mathbf{z}_{new} = \mathbf{c}^\top\mathbf{z}_{old} + (\mathbf{c}_\mathsf{N}^\top - \mathbf{c}_\mathsf{B}^\top\mathsf{B}^{-1}\mathsf{N})\mathbf{x}$.

Improvement if $0 > \mathbf{c}^\top\mathbf{z}_{new} - \mathbf{c}^\top\mathbf{z}_{old} = (\mathbf{c}_\mathsf{N}^\top - \mathbf{c}_\mathsf{B}^\top\mathsf{B}^{-1}\mathsf{N})\mathbf{x}$.

No improvement if $(\mathbf{c}_\mathsf{N}^\top - \mathbf{c}_\mathsf{B}^\top\mathsf{B}^{-1}\mathsf{N}) \geq \mathbf{o}^\top$ because $\mathbf{x} \geq \mathbf{o}$ ...

Stopping criterion/updates for simplex algorithm

At current vertex $\mathbf{z}_{old}$, if the *reduced cost vector*

$$\mathbf{c}^{red} := (\mathbf{c}_\mathsf{N}^\top - \mathbf{c}_\mathsf{B}^\top \mathsf{B}^{-1}\mathsf{N})^\top \geq \mathbf{o}$$

then **stop:** $\mathbf{z}^* = \mathbf{z}_{old}$ is optimal solution to LP.

Otherwise, select i with $c_i^{red} < 0$ and update to neighboring vertex $\mathbf{z}_{new}$ with $z_i^{new} > 0 = z_i^{old}$.

Back to example

$$\min\left\{\mathbf{c}^\top \mathbf{z} : \mathsf{A}\mathbf{z} = \mathbf{b}\,,\ \mathbf{z} \in \mathbb{R}_+^7\right\} \ldots$$

Example, continued

... with, say, $\mathbf{c}^\top = [-1, 2, 3 \,|\, 0, 0, 0, 0]$ (recall: last four are slacks)

$$\mathsf{A} = \begin{bmatrix} 1 & 1 & 1 & | & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & | & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & | & 0 & 0 & 1 & 0 \\ 0 & 3 & 1 & | & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4 \\ 2 \\ 3 \\ 6 \end{bmatrix}.$$

Start with vertex $\mathbf{z}_{old} = [\mathbf{o}^\top | \mathbf{s}_{old}^\top]$, so here $\mathbf{s}_{old} = \mathbf{b}$ and $\mathbf{c}_\mathsf{B} = \mathbf{o}$.

Calculate $\mathbf{c}^{red} = (\mathbf{c}_\mathsf{N}^\top - \mathbf{c}_\mathsf{B}^\top \mathsf{B}^{-1}\mathsf{N})^\top = \mathbf{c}_\mathsf{N} = [-1, 2, 3]^\top \not\geq \mathbf{o}$.

Hence $\mathbf{z}_{old}$ not optimal, pass to next vertex with $z_1^{new} > 0$

Only such vertex is $\mathbf{z}_{new}^\top = [2, 0, 0 \,|\, 2, 0, 3, 6]$. Compare with previous vertex: $\mathbf{z}_{old}^\top = [0, 0, 0 \,|\, 4, 2, 3, 6]$.

New reduced cost vector is $\mathbf{c}_{new}^{red} = [1, 2, 3]^\top \geq \mathbf{o}$, so **stop:**

$\mathbf{x}^* = [2, 0, 0]^\top$ is optimal (first part of $\mathbf{z}_{new}$).