Operations research meets data science: principles, models and algorithms

Immanuel Bomze, University of Vienna

Dip. Matematica@Univ. Padova 04 November 2025

What is Operations/Operational Research (OR) ?

Quantitative Decision Support...

Process view — the OR cycle:

Problem analysis

Model building

Develop new solution methods/algorithms

Testing, benchmarking and interpreting

The plan

Planning an OR project, typology of models

Recall Linear optimization Problems (LP)

Continuous optimization (NLP), (N)LP duality

Shortest paths, network flows, spanning trees

Discrete optimization − branch-and-bound

Integer Linear optimization Problems (ILP)

Gaps and cuts, unimodularity

Quantifying problems — variables

Numbers: binary $b \in \{0, 1\}$, integer $z \in \mathbb{Z}$,
"natural" $n \in \mathbb{N} = \{n \in \mathbb{Z} : n \geq 1\}$, rational $r \in \mathbb{Q} = \left\{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N}\right\}$
real $x \in \mathbb{R}$, complex $y \in \mathbb{C}$ ...

Collect them in vectors $\mathbf{x} = [x_i]_{i=1}^m$, $x_i \in \mathbb{R}$, usually stacked vertically — *columns* $\mathbf{x} \in \mathbb{R}^m$.

Row vector $\mathbf{x}^\top = [x_1, \ldots, x_m]$ by transposition $^\top$.

Collect n vectors $\mathbf{a}_j \in \mathbb{R}^m$ in arrays —
$m \times n$-matrices $\mathsf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_n]$.

Operations: addition, scaling, matrix-vector multiplication $\mathsf{A}\mathbf{x}$.

OR is Quantitative Decision Support

Quantitative – yes, numbers; but variables ? Range ?

Decisions ? Are coded in **values** of variables, i.e.
decide which value is taken in the range of a variable.

Instead of (unknown) variables solve for equations/relations/criteria..

Reason for looking at $A\mathbf{x} = \mathbf{b}$ where $(A, \mathbf{b})$ problem/model data,
$\mathbf{x}$ the **decision variables**.

More complex quantitative-coded decisions than vectors ?

Caution: not all variables are decision variables !

Ex.: random variable $\tilde{x}$ with probability distribution $\mathcal{N}(0, 1)$.

What is the difference ? Randomness ?

Even if unknown parameters are involved, e.g. $\mathcal{N}(\mu, \sigma^2)$.

Again: careful with terminology and model ...

Quantifying problems — decisions

Simple decision: "no/yes" or "false/true" encoded by $\{0, 1\}$.

Combining ten of them results in ...

Enables modeling complex statements:

Ex.: "if A then B", $A \Rightarrow B$, encoded by binary $a, b \in \{0, 1\}$:

$$z = (1 - a) + b - (1 - a) * b = (1 - a) * (1 - b) + b$$

Important for conditional relations/constraints.

Ex.: if A, then $x_1 + x_2 = 1$; how to model in one equation ?

Formulating (hard) constraints

Equality constraints

Ex.: $x_1 + x_2 = 1$ **Ex.:** $x = x^2$ **Ex.:** $\sin(2\pi t) = 0$ **Ex.:** $A\mathbf{x} = \mathbf{b}$.

... all can be written as $\mathbf{h}(\mathbf{x}) = \mathbf{o}$.

Inequality constraints **Ex.:** $x_1 + x_2 \leq 1$

Ex.: $A\mathbf{x} \geq \mathbf{b}$ meaning ?

$$x_1 \geq 0 \quad \text{and} \quad x_2 \geq 0\,, \quad \text{write } \mathbf{x} \geq \mathbf{o} \text{ or } \mathbf{x} \in \mathbb{R}^2_+\,.$$

... all can be written as $\mathbf{g}(\mathbf{x}) \leq \mathbf{o}$.

Conversion inequalities $\rightarrow$ equalities: *slack variables* $\mathbf{s} \in \mathbb{R}^m_+$,

$$\mathbf{h}(\mathbf{x}; \mathbf{s}) := \mathbf{g}(\mathbf{x}) + \mathbf{s} = \mathbf{o}\,.$$

Objectives to select decisions – criteria

Typically, constraints still leave a choice;

many $\mathbf{x}$ satisfy them, are *feasible*.

Need criteria to select values for variables – decisions.

Single criterion/objective function f taking numerical values:

$$\min \{ f(\mathbf{x}) : \mathbf{x} \text{ feasible solution } \} .$$

Optimal solution $\mathbf{x}^*$ is feasible and satisfies

$$f(\mathbf{x}^*) \leq f(\mathbf{x}) \quad \text{for all feasible } \mathbf{x} .$$

As many inequalities as there are feasible $\mathbf{x}$.

Planning OR projects; typology of models

Planning (iterative):

- Formulation of problem (raw stage);
- Setting the goals, conceptual analysis;
- Resource planning and formulation of model(s);
- Validation of obtained solution; testing; calibration;
- Implementation, interpretation, transfer.

Typology of models

(rough, blurry boundaries and combinations):

- intuitive models – using analogies heuristically;
- empirical models – develop alternatives;
- **analytic models – prediction and optimization;**
- holistic models – emphasize system thinking, simulation

Role of time and uncertainty

- static models – one stage, one shot;
- dynamic models – several stages (not necessarily time);
- deterministic models – pretty sure about problem data;
- not so sure – stochastic models, robustness;
- strategic interaction models – conflicts, hierarchies, reactions

Back to binary constraints . . .

Recap: want to model "A implies B ($A \Rightarrow B$)" for OR handling.

Had binary variables $a, b \in \{0, 1\}$ ($a = 1 \,|\, b = 1$ iff $A \,|\, B$ holds).

Had equality constraint

$$(1 - a) * (1 - b) + b = 1\,.$$

Simpler way via inequality, involving just a, b ?

Bottom line: many formulations/models for the same fact.

Objectives to select decisions – criteria

Typically, constraints still leave a choice;

many $\mathbf{x}$ satisfy them, are *feasible (solutions)*.

Need criteria to select values for variables – decisions.

Single criterion/objective function f taking numerical values:

$$\min \{f(\mathbf{x}) : \mathbf{x} \text{ feasible solution}\} .$$

Optimal solution $\mathbf{x}^*$ satisfies

$$f(\mathbf{x}^*) \leq f(\mathbf{x}) \quad \text{for all feasible solutions} \mathbf{x} .$$

As many inequalities as there are feasible $\mathbf{x}$.

Feasibility and infeasibility

Decision $\mathbf{x}$ feasible: satisfies all constraints;

Decision $\mathbf{x}$ infeasible: at least one constraint is violated by $\mathbf{x}$.

Problem feasible: at least one feasible $\mathbf{x}$ exists;

Problem infeasible: no feasible $\mathbf{x}$ exists –
constraints are *inconsistent/contradictory.*

Ex.: $x \leq 1$ and $x \geq 2$ can never be met by same x.

Multiple criteria

Multicriteria/vector valued objective $\mathbf{f} = [f_1, \ldots, f_k]^\top$:

Utopia point $\mathbf{x}^*$ satisfies

$$f_i(\mathbf{x}^*) = \min \{ f_i(\mathbf{x}) : \mathbf{x} \text{ feasible } \} \qquad \text{for all } i\,.$$

Typically does not exist (conflicting goals).

Pareto dominance for feasible $\mathbf{x}, \mathbf{y}$:

$\mathbf{y}$ Pareto-dominates $\mathbf{x}$, $\mathbf{y} \succ_P \mathbf{x}$, if

$$f_i(\mathbf{y}) \leq f_i(\mathbf{x}) \qquad \text{for all } i\,,$$

with strict inequality for at least one i.

Pareto solution $\mathbf{x}^*$: undominated feasible $\mathbf{x}^*$,

$$\text{for all feasible } \mathbf{y} \text{ have } \quad \mathbf{y} \not\succ_P \mathbf{x}^*\,.$$