Oleksandr Romanko, Ph.D.
Senior Research Analyst, Risk Analytics, Watson Financial Services, IBM Canada
Adjunct Professor, University of Toronto

MIE1624H – Introduction to Data Science and Analytics

Lecture 8 – Optimization

University of Toronto
November 13, 2018

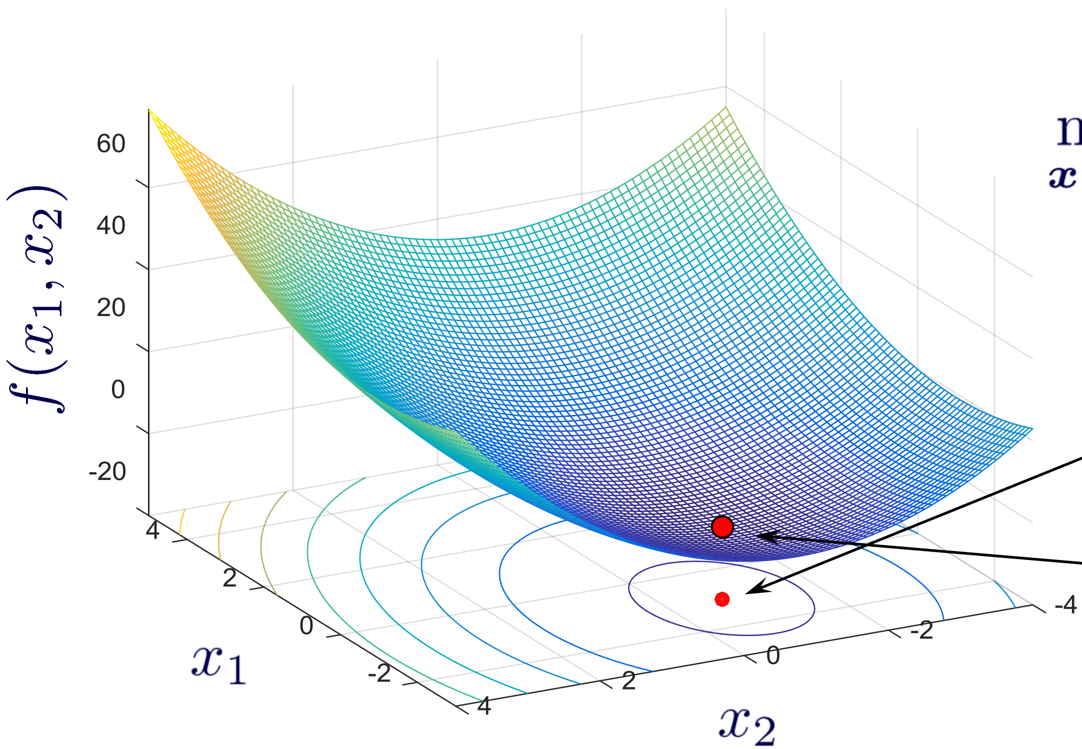
📖 Overview of Optimization

Optimization

- **Optimization problem**

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega \end{array}$$

- Examples:
  – Minimize **cost**
  – Maximize **profit**



$$\min_{\boldsymbol{x}\in\mathbb{R}^2} \quad x_1^2 + 2x_2^2 + 2x_1 + 3x_2$$

solve

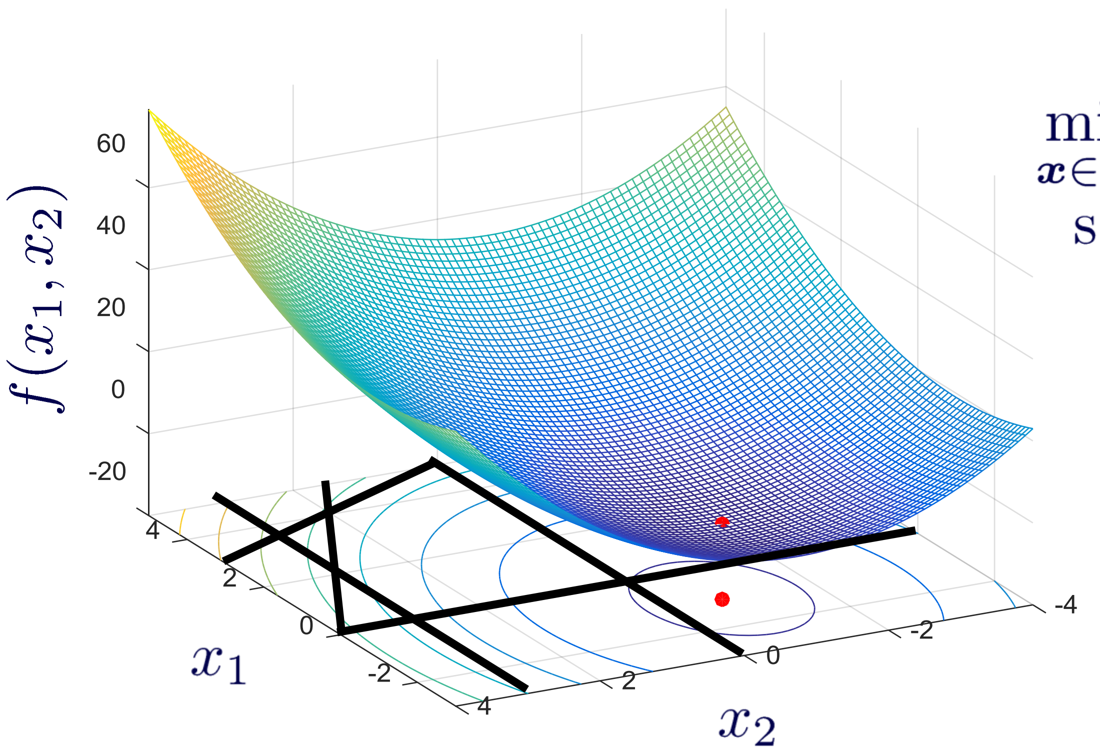
$$\boldsymbol{x}^* = \begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} -1.00 \\ -0.75 \end{pmatrix}$$

$$f(\boldsymbol{x}^*) = -2.125$$

Optimization

- **Optimization problem**

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega \end{array}$$



$$\begin{array}{rl} \min\limits_{\boldsymbol{x}\in\mathbb{R}^2} & x_1^2 + 2x_2^2 + 2x_1 + 3x_2 \\ \text{s.t.} & x_1 + 2x_2 \leq 8 \\ & 2x_1 + \ \ x_2 \leq 10 \\ & x_2 \leq 3 \\ & x_1,\ x_2 \geq 0 \end{array}$$

Optimization

- **Optimization problem**

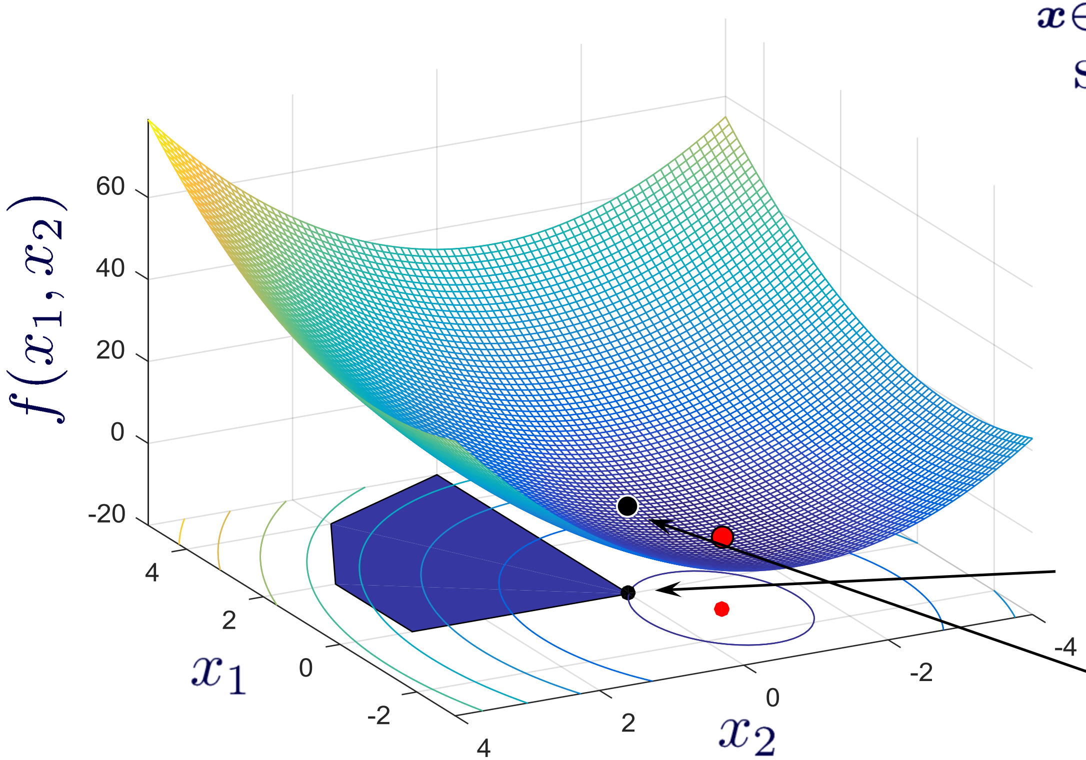
$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega \end{array}$$

$$\begin{array}{rl} \min\limits_{\boldsymbol{x}\in\mathbb{R}^2} & x_1^2 + 2x_2^2 + 2x_1 + 3x_2 \\ \text{s.t.} & x_1 + 2x_2 \le 8 \\ & 2x_1 + x_2 \le 10 \\ & x_2 \le 3 \\ & x_1,\, x_2 \ge 0 \end{array}$$

solve

$$\boldsymbol{x}^* = \begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$f(\boldsymbol{x}^*) = 0$$



Optimization

- **Optimization problem**

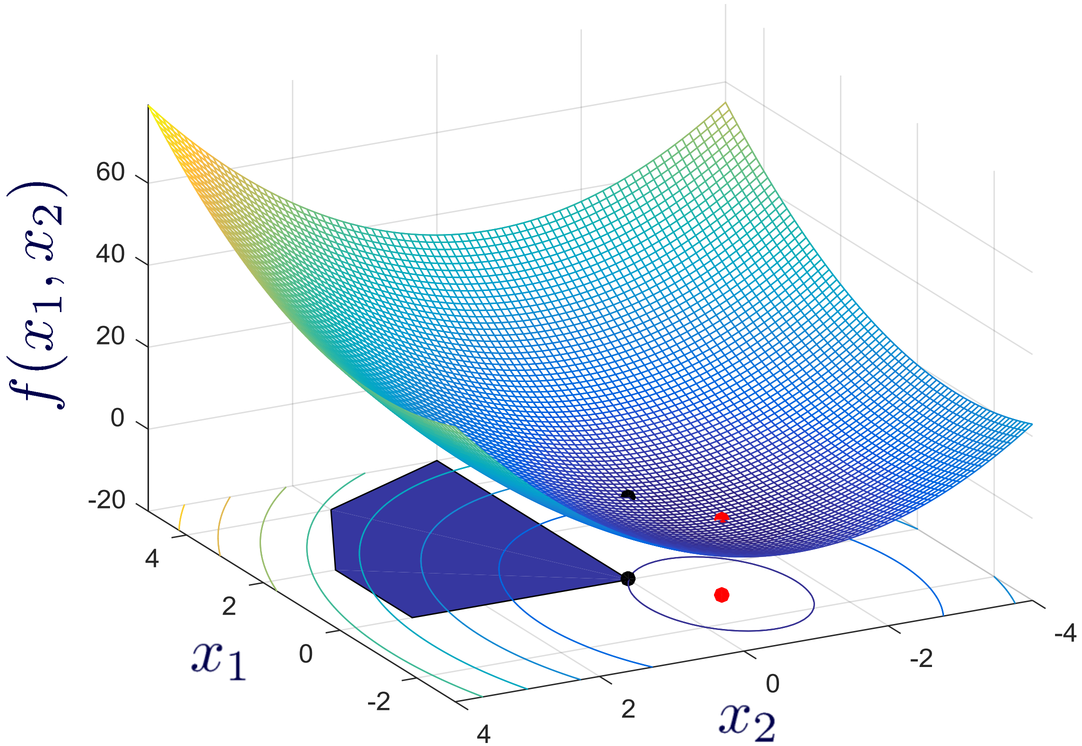
$$\begin{aligned} \text{minimize}_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{x} \in \Omega \end{aligned}$$

- **Minimizing convex quadratic (QP) objective function** over a **polyhedron** (linear constraints)

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \boldsymbol{c}^T\boldsymbol{x} + \tfrac{1}{2}\boldsymbol{x}^T Q \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{l} \le A\boldsymbol{x} \le \boldsymbol{u} \\ & \boldsymbol{l}_b \le \boldsymbol{x} \le \boldsymbol{u}_b \end{aligned}$$

general form

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^2} \quad & x_1^2 + 2x_2^2 + 2x_1 + 3x_2 \\ \text{s.t.} \quad & x_1 + 2x_2 \le 8 \\ & 2x_1 + x_2 \le 10 \\ & x_2 \le 3 \\ & x_1,\, x_2 \ge 0 \end{aligned}$$



Optimization

- **Optimization problem**

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega \end{array}$$

- Examples:
  - Minimize **cost**
  - Maximize **profit**

- **Multi-objective optimization**: simultaneously optimizing two or more conflicting objectives subject to certain constraints

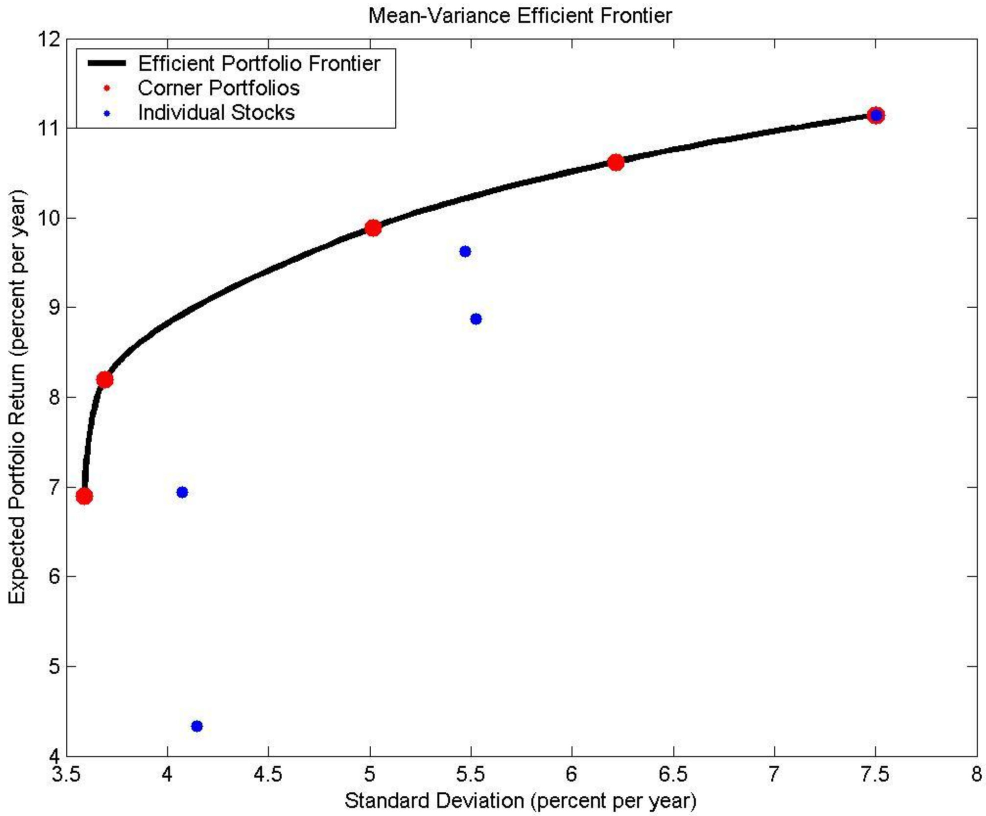
$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & F(\boldsymbol{x}) = [f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_k(\boldsymbol{x})]^T \\ \text{s.t} & \boldsymbol{x} \in \Omega \end{array}$$

- Examples:
  - Minimize **cost** & Minimize **environmental impact**
  - Minimize **risk** & Maximize **return**

Multi-objective optimization

- Solving **multi-objective optimization** problems:

$$\begin{array}{rl} \text{minimize} & \textbf{risk} \\ \text{subject to} & \textbf{return} \geq \textbf{target} \\ & \textbf{other constraints} \end{array}$$



Linear Optimization Examples

Workforce planning

- Restaurant is open 7 days a week
- Based on past experience, the number of workers needed on a particular day:

Day	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Number	14	13	15	16	19	18	11

- Every worker works five consecutive days, and then takes two days off
- Minimize the number of workers that staff the restaurant
- Decision variables:
  - (**wrong**) x_i is the number of workers that work on day i
  - (**right**) x_i is the number of workers who begin their five-day shift on day i
- Objective function:

$$\min_{\boldsymbol{x} \in \mathbb{R}^7} \quad x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7$$

- Bounds on variables:

$$x_i \geq 0 \qquad \forall \ i$$

Workforce planning

- Constraints:
  - Consider the constraint for Monday's staffing level of 14
  - Who works on Monday? Those who start their shift on Monday (x_1)
  - Those who start on Tuesday (x_2) do not work on Monday, nor do those who start on Wednesday (x_3)
  - Those who start on Thursday (x_4) do work on Monday, as do those who start on Friday (x_5), Saturday (x_6), and Sunday (x_7)
  - This gives the constraint:

$$x_1 + x_4 + x_5 + x_6 + x_7 \geq 14$$

- Similar argument give a total formulation

Workforce planning – expression formulation

- Constraints:
  - Consider the constraint for Monday's staffing level of 14
  - Who works on Monday? Those who start their shift on Monday (x_1)
  - Those who start on Tuesday (x_2) do not work on Monday, nor do those who start on Wednesday (x_3)
  - Those who start on Thursday (x_4) do work on Monday, as do those who start on Friday (x_5), Saturday (x_6), and Sunday (x_7)
- Similar argument give a total formulation:

$$\begin{aligned}
\min_{\boldsymbol{x} \in \mathbb{R}^7} \quad & x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 \\
\text{s.t.} \quad & x_1 + x_4 + x_5 + x_6 + x_7 \geq 14 \\
& x_1 + x_2 + x_5 + x_6 + x_7 \geq 13 \\
& x_1 + x_2 + x_3 + x_6 + x_7 \geq 15 \\
& x_1 + x_2 + x_3 + x_4 + x_7 \geq 16 \\
& x_1 + x_2 + x_3 + x_4 + x_5 \geq 19 \\
& x_2 + x_3 + x_4 + x_5 + x_6 \geq 18 \\
& x_3 + x_4 + x_5 + x_6 + x_7 \geq 11 \\
& x_i \geq 0 \quad \forall\, i
\end{aligned}$$

Workforce planning – matrix formulation

Mathematical formulation

$$
\begin{aligned}
\min_{\boldsymbol{x}\in\mathbb{R}^7} \quad & x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 \\
\text{s.t.} \quad & x_1 + x_4 + x_5 + x_6 + x_7 \geq 14 \\
& x_1 + x_2 + x_5 + x_6 + x_7 \geq 13 \\
& x_1 + x_2 + x_3 + x_6 + x_7 \geq 15 \\
& x_1 + x_2 + x_3 + x_4 + x_7 \geq 16 \\
& x_1 + x_2 + x_3 + x_4 + x_5 \geq 19 \\
& x_2 + x_3 + x_4 + x_5 + x_6 \geq 18 \\
& x_3 + x_4 + x_5 + x_6 + x_7 \geq 11 \\
& x_i \geq 0 \quad \forall\, i
\end{aligned}
$$

$$\Rightarrow \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} \Rightarrow$$

Solver formulation

$$
\begin{aligned}
\min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \boldsymbol{c}^T\boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{l} \leq A\boldsymbol{x} \leq \boldsymbol{u} \\
& \boldsymbol{l}_b \leq \boldsymbol{x} \leq \boldsymbol{u}_b
\end{aligned}
$$

$$\boldsymbol{c} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}^T \qquad \boldsymbol{l}_b = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^T \qquad \boldsymbol{u}_b = \boldsymbol{u}$$

$$
\underbrace{\begin{pmatrix} 14 \\ 13 \\ 15 \\ 16 \\ 19 \\ 18 \\ 11 \end{pmatrix}}_{\boldsymbol{l}} \leq
\underbrace{\begin{pmatrix}
1 & 0 & 0 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & 0 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 & 0 & 1 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}}_{A} \cdot \boldsymbol{x} \leq
\underbrace{\begin{pmatrix} +\infty \\ +\infty \\ +\infty \\ +\infty \\ +\infty \\ +\infty \\ +\infty \end{pmatrix}}_{\boldsymbol{u}}
$$

Road design problem

- Given point **A** and **B**, and a map, find a road that will be the cheapest to construct/maintain

Source: Warren L. Hare, A Multi-Haul Model for Vertical Alignment Optimization of Road Design

Road design problem

- Given point **A** and **B**, and a map, find a road that will be the cheapest to construct/maintain

Source: Warren L. Hare, A Multi-Haul Model for Vertical Alignment Optimization of Road Design

Road design problem

- Given point **A** and **B**, and a map, find a road that will be the cheapest to construct/maintain

Source: Warren L. Hare, A Multi-Haul Model for Vertical Alignment Optimization of Road Design

Road design problem - costs (easiest to hardest to compute)

- Cost depends on many factors
  - Land acquisition ≈ 0% to 25%
  - Bridges, tunnels, etc. ≈ 0% to 20%
  - Per km cost (final paving, maintenance) ≈ 5% to 15%
  - Earthwork ≈ 20% to 50%
  - Long term economic costs ≈ ?
  - Environmental issues ≈ ?

Source: Warren L. Hare, A Multi-Haul Model for Vertical Alignment Optimization of Road Design

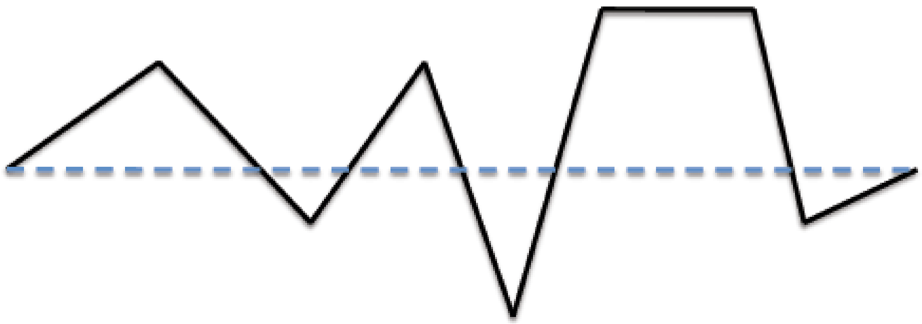
Road optimization

- Level 1: Horizontal Alignment (HA)
  - Look at a map and select a path for the road to follow
  - Evaluate HA based on safety constraints and Vertical Alignment
- Level 2: Vertical Alignment (VA)
  - Use HA to build a cross section of the terrain
  - Determine a VA for the future road
  - Evaluate VA based safety constraints and Earthwork
- Level 3: Earthwork (EW)
  - Determine how to move earth in order to make the terrain fit the VA
  - Cost based on minimization

Vertical Alignment and Earthwork can be modelled and solved simultaneously as a mixed-integer linear optimization problem

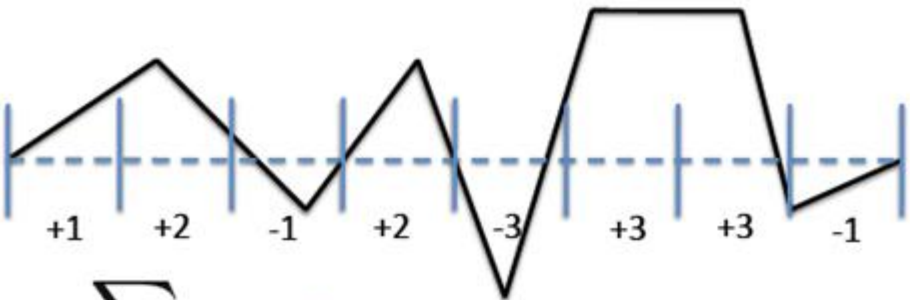
Source: Warren L. Hare, A Multi-Haul Model for Vertical Alignment Optimization of Road Design

Road optimization - earthwork



Source: Warren L. Hare, A Multi-Haul Model for Vertical Alignment Optimization of Road Design

Road optimization - earthwork



$$\min \sum_{(i,j)\in\mathcal{M}^2} c_{ij} x_{ij}$$

$$\text{s.t.} \quad \sum_{j\in\mathcal{M}^i_{\rightarrow}} x_{ij} - \sum_{j\in\mathcal{M}^i_{\leftarrow}} x_{ji} = V_i \quad \text{for all } i \in \mathcal{S}$$

$$x_{ij} \geq 0 \quad \text{for all } i,j \in \mathcal{S} i \neq j$$

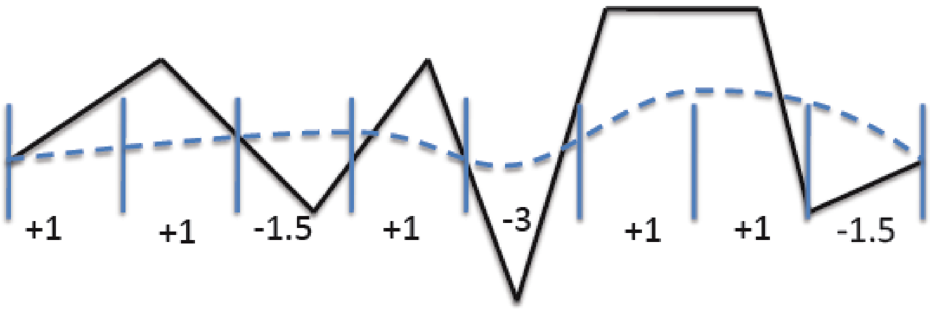
Feasibility:

- Only feasible if $\sum V_i = 0$
- Fixed by introducing "borrow pits" and "waste pits"
- Essentially extra sections that don't need to be balanced

Source: Warren L. Hare, A Multi-Haul Model for Vertical Alignment Optimization of Road Design

Road optimization - earthwork

A flat road is unlikely to minimize costs



$$
\begin{aligned}
\min \quad & \sum_{(i,j)\in\mathcal{M}^2} c_{ij} x_{ij} \\
\text{s.t.} \quad & \textstyle\sum_{j\in\mathcal{M}^i_\rightarrow} x_{ij} - \sum_{j\in\mathcal{M}^i_\leftarrow} x_{ji} = V_i && \text{for all } i \in \mathcal{S} \\
& x_{ij} \geq 0 && \text{for all } i,j \in \mathcal{S}, i \neq j \\
& V_i = L(h_i, u_i) && \text{for all } i \in \mathcal{S}
\end{aligned}
$$

Notation: (variables in blue, constants in black)

x_{ij} = earth moved from i to j

c_{ij} = cost

V_i = target volume

L = a linear function

$\mathcal{S}$ = sections

$\mathcal{M}^i_\rightarrow, \mathcal{M}^i_\leftarrow, \mathcal{M}^2$ = move lists

u_i = height of ground at section i

h_i = height of road at section i

Source: Warren L. Hare, A Multi-Haul Model for Vertical Alignment Optimization of Road Design

Overview of Optimization Techniques

Optimization

- **Optimization problem**

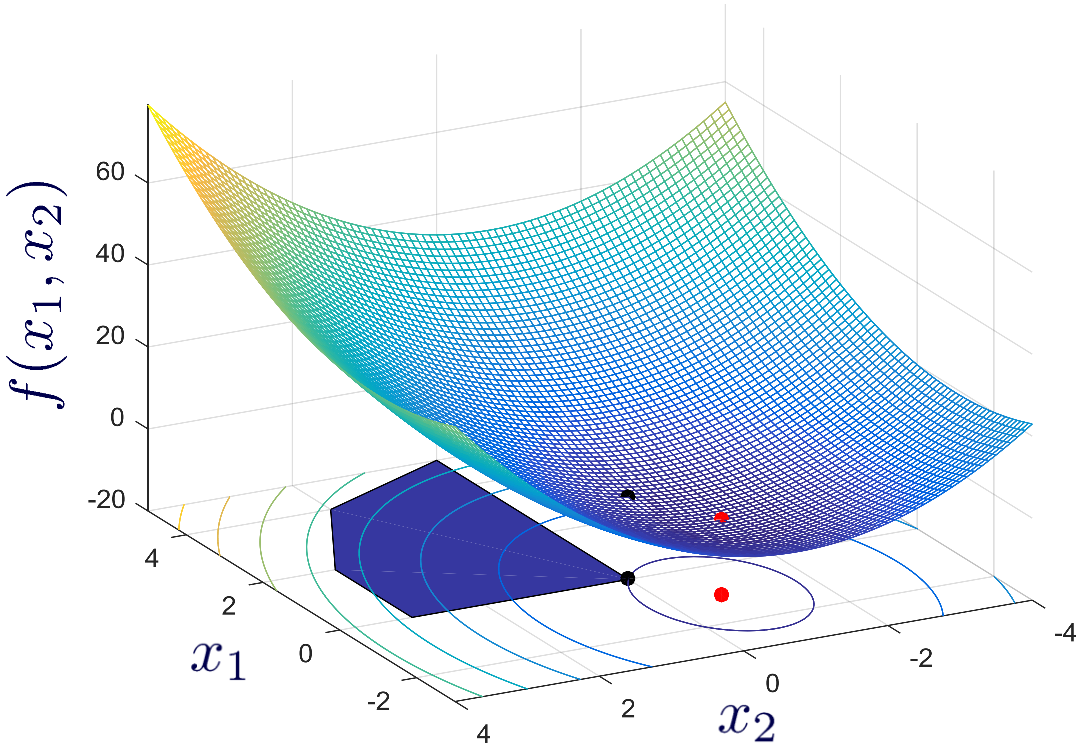
$$\begin{aligned} \text{minimize}_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{x} \in \Omega \end{aligned}$$

- **Minimizing convex quadratic (QP) objective function** over a **polyhedron** (linear constraints)

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}^n} \quad & \boldsymbol{c}^T \boldsymbol{x} + \tfrac{1}{2} \boldsymbol{x}^T Q \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{l} \leq A\boldsymbol{x} \leq \boldsymbol{u} \\ & \boldsymbol{l}_b \leq \boldsymbol{x} \leq \boldsymbol{u}_b \end{aligned}$$

general form

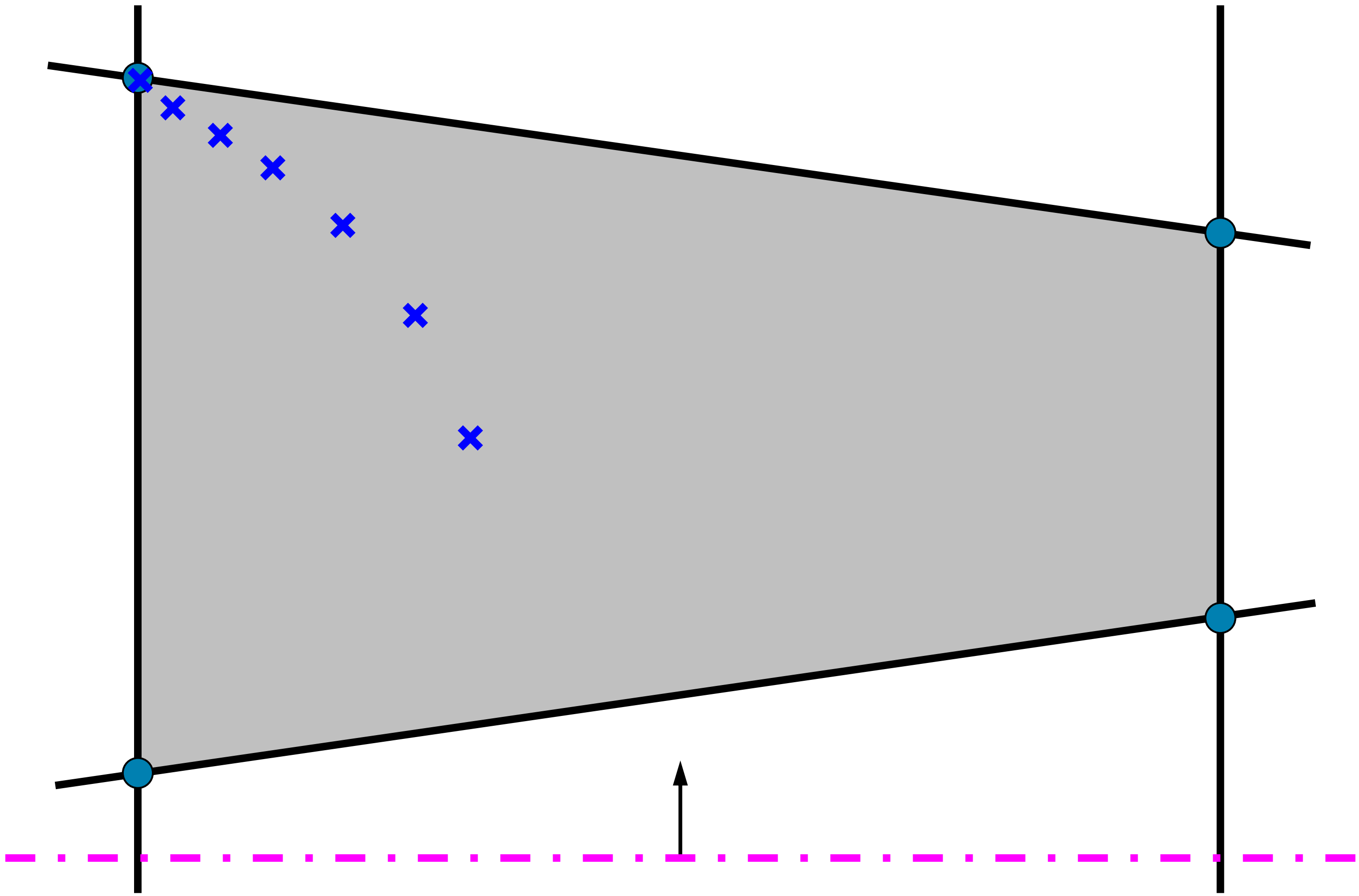
$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}^2} \quad & x_1^2 + 2x_2^2 + 2x_1 + 3x_2 \\ \text{s.t.} \quad & x_1 + 2x_2 \leq 8 \\ & 2x_1 + x_2 \leq 10 \\ & x_2 \leq 3 \\ & x_1,\, x_2 \geq 0 \end{aligned}$$



Solving linear optimization problems

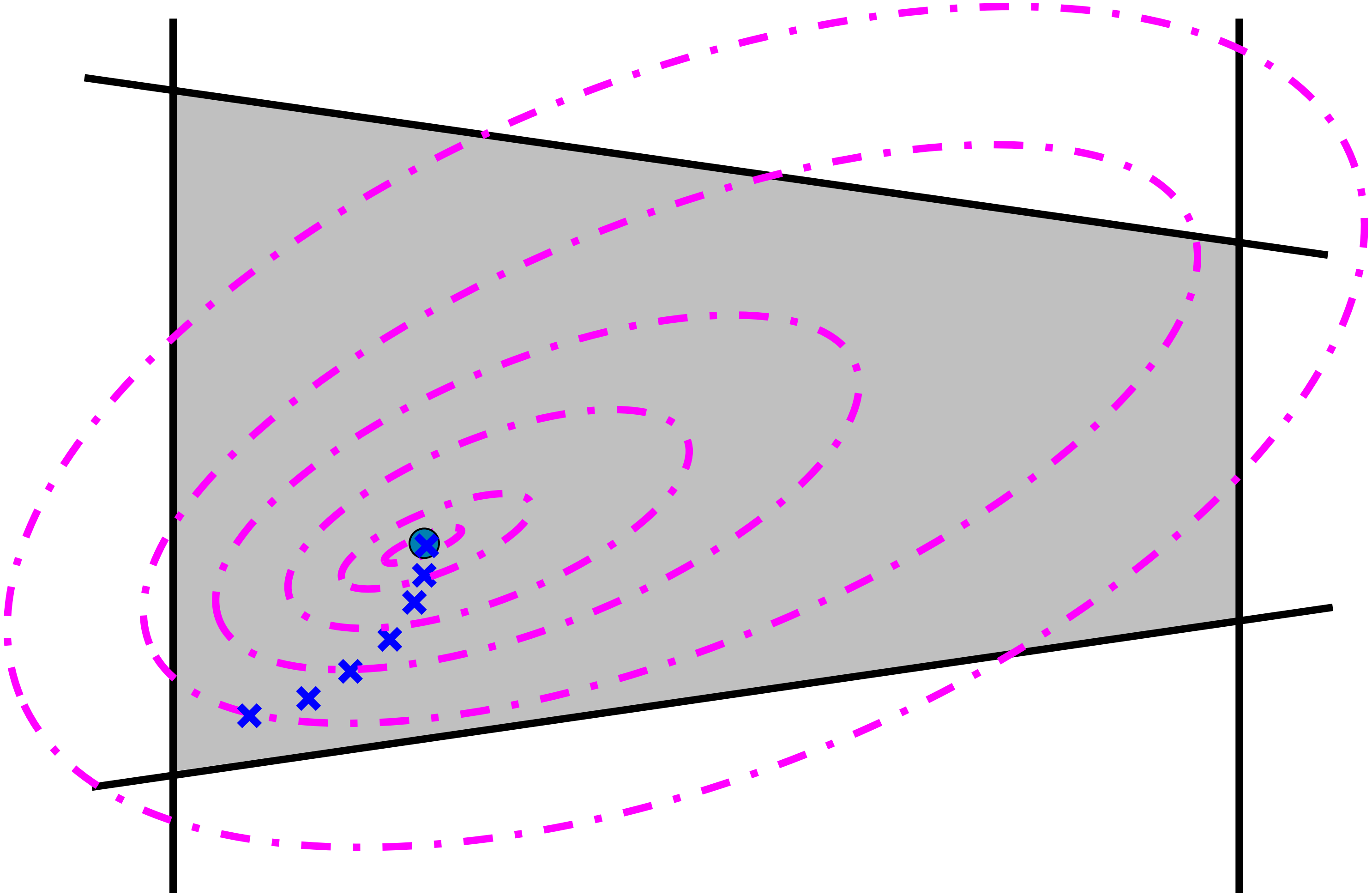
- Maximizing/minimizing linear (LP) function over a polyhedron
- Interior Point Methods vs. Simplex-type Methods

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \quad \boldsymbol{c}^T\boldsymbol{x}$$
$$\text{s.t.} \quad \boldsymbol{l} \le A\boldsymbol{x} \le \boldsymbol{u}$$



Solving non-linear optimization problems

- Convex non-linear objective function (NLP), linear or non-linear constraints
- Illustrated solution technique – Interior Point Methods
- Other solution techniques – gradient methods, Newton and Quasi-Newton



Solving linear optimization problems

- Simplex Method – graphical view

$$
\begin{aligned}
\max_{x\in\mathbb{R}^3} \quad & 3x_1 + 2x_2 + 2x_3 \\
\text{s.t.} \quad & x_1 + \quad x_3 \le 8 \\
& x_1 + x_2 \qquad \le 7 \\
& x_1 + 2x_2 \qquad \le 12 \\
& x_1,\, x_2,\, x_3 \ge 0
\end{aligned}
$$

standard form

$$
\begin{aligned}
\max_{x\in\mathbb{R}^3} \quad & 3x_1 + 2x_2 + 2x_3 \\
\text{s.t.} \quad & x_1 + \quad x_3 + x_4 = 8 \\
& x_1 + x_2 \qquad + x_5 = 7 \\
& x_1 + 2x_2 \qquad + x_6 = 12 \\
& x_1,\, x_2,\, x_3 \ge 0 \\
& x_4,\, x_5,\, x_6 \ge 0
\end{aligned}
$$

Source: Ed Klotz, Interior Point Algorithm Fundamentals

Solving linear optimization problems

- Simplex Method – graphical view

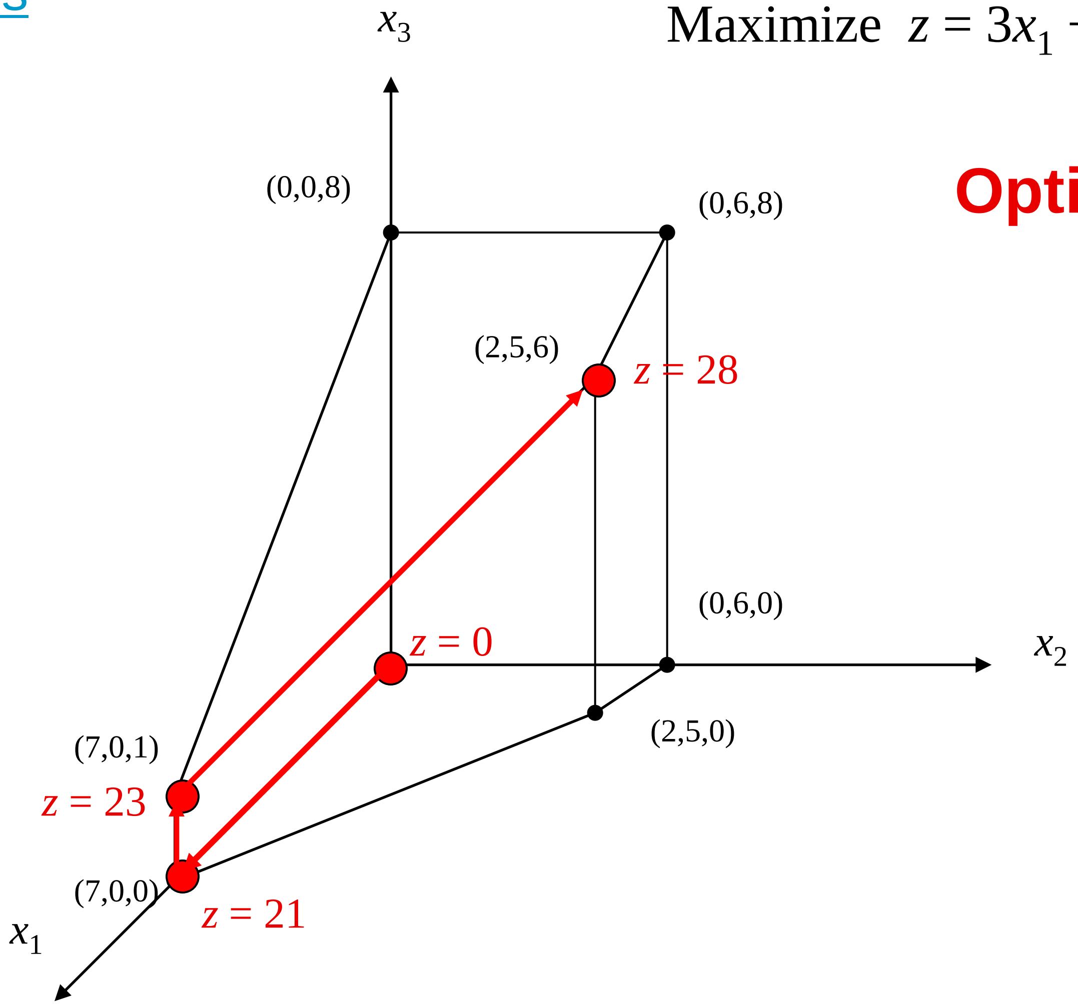
Current Basis

~~(x_4, x_5, x_6)~~

~~(x_4, x_1, x_6)~~

~~(x_3, x_1, x_6)~~

(x_3, x_1, x_2)



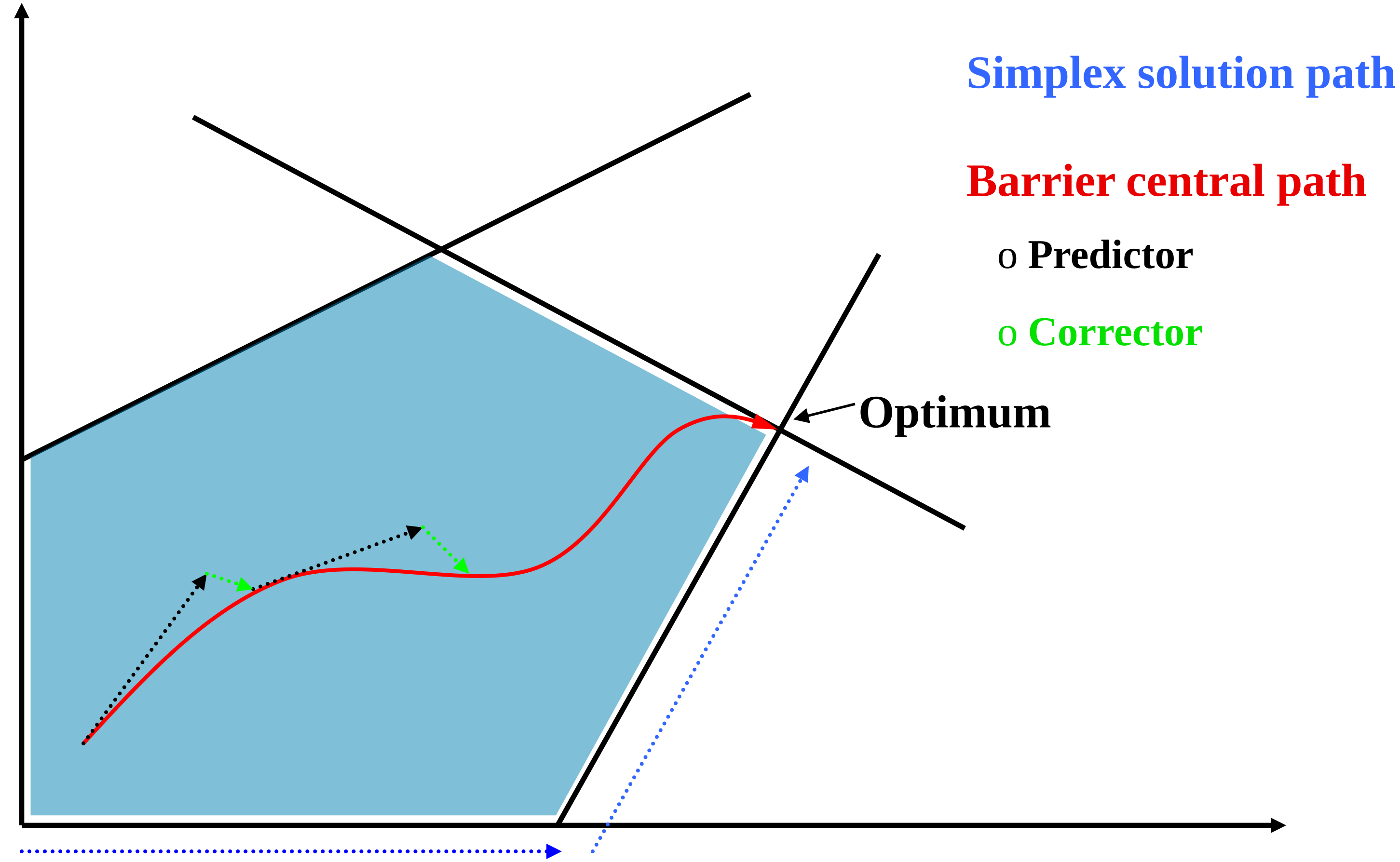
Maximize $z = 3x_1 + 2x_2 + 2x_3$

Optimal!

Source: Ed Klotz, Interior Point Algorithm Fundamentals

Solving linear optimization problems

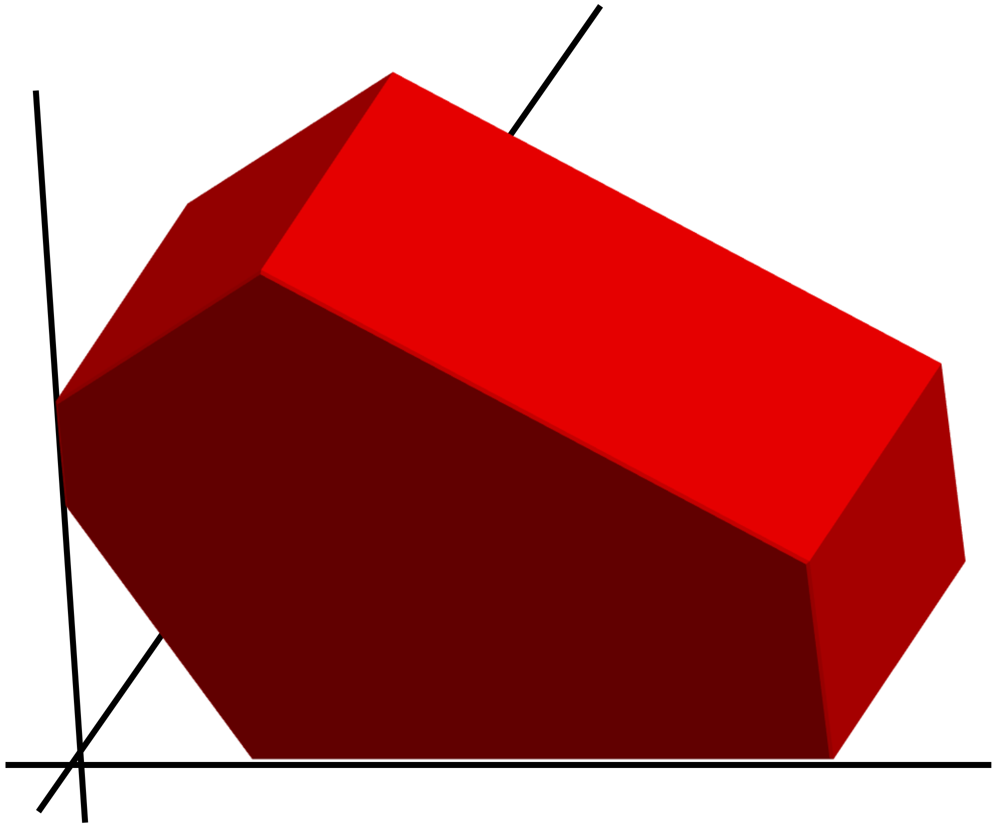
- Interior Point Method (barrier algorithm in CPLEX)



Source: Ed Klotz, Interior Point Algorithm Fundamentals

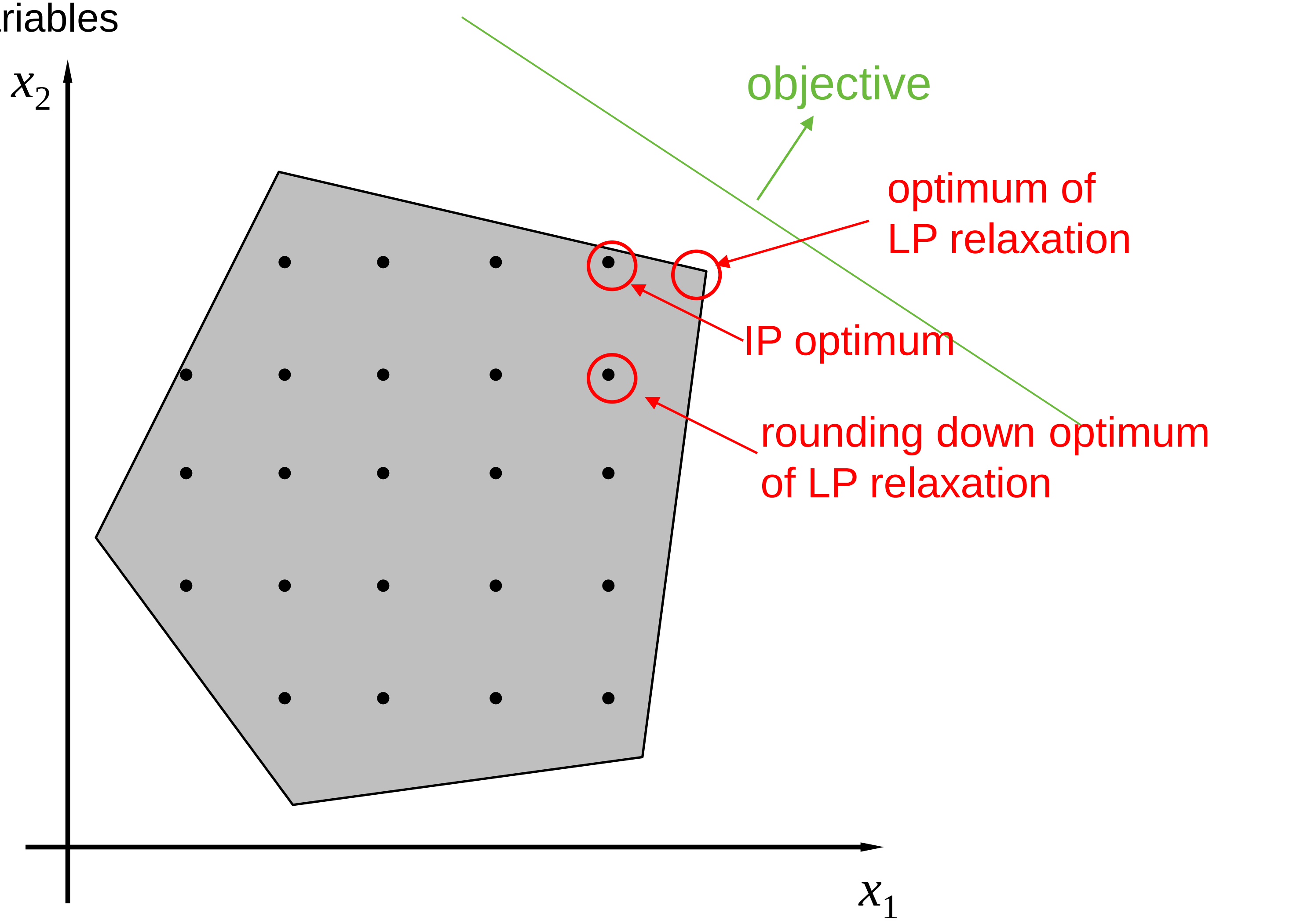
Solving linear optimization problems

- Higher dimensions



Solving mixed-integer optimization problems

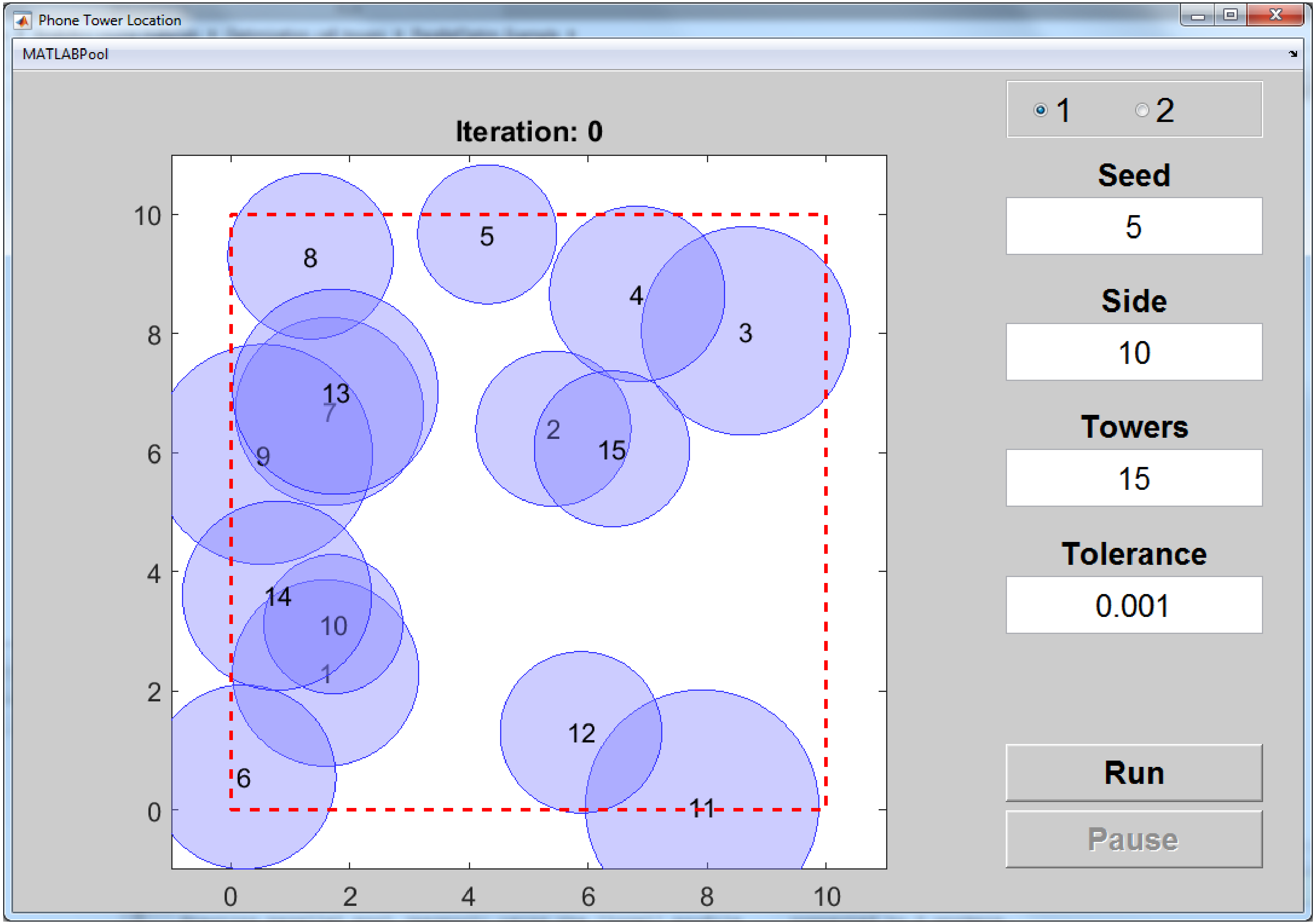
- Mixed-integer optimization problems (MIP)
  - continuous variables
  - integer variables



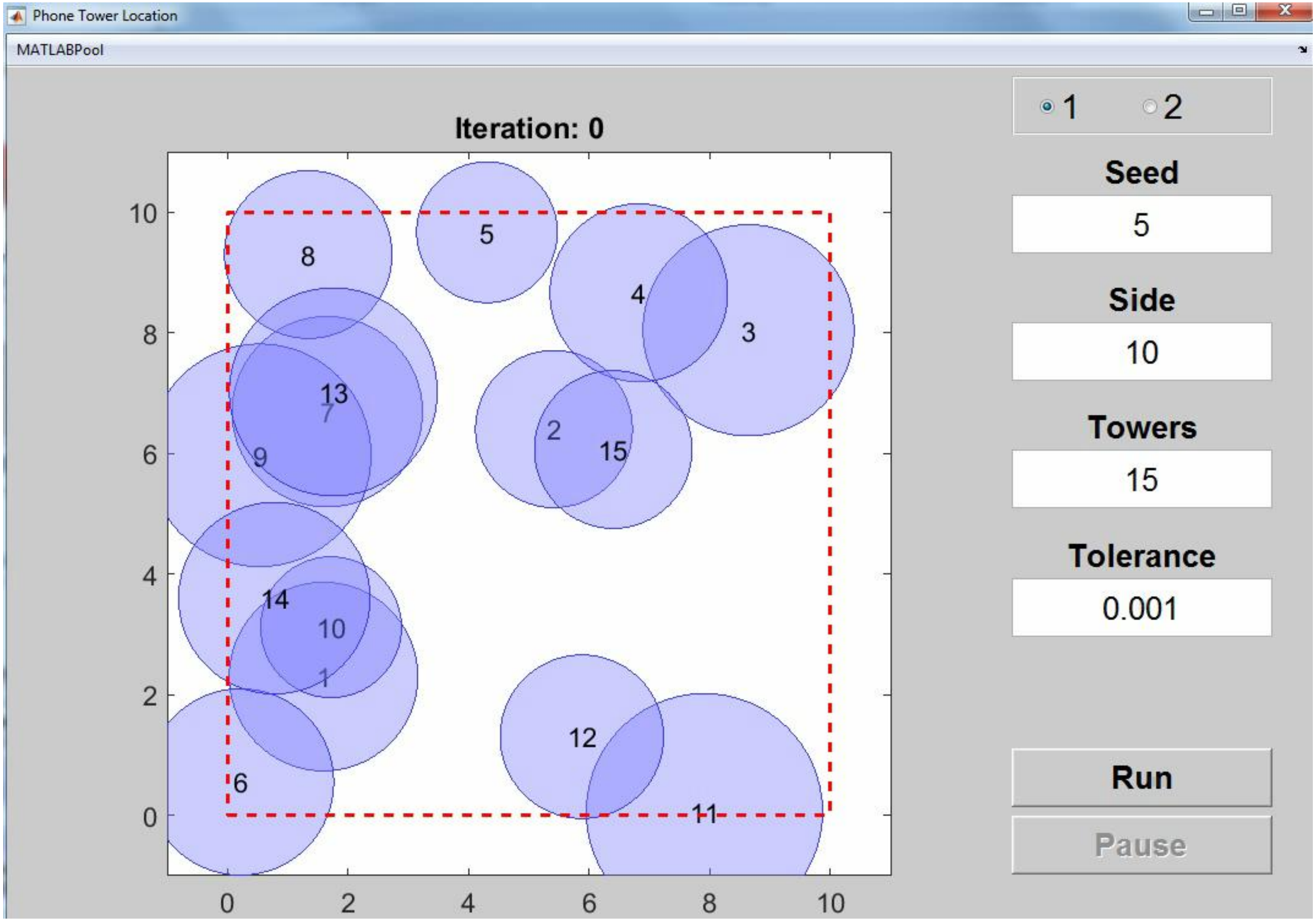
feasible
solutions = •

📖 Optimization Examples

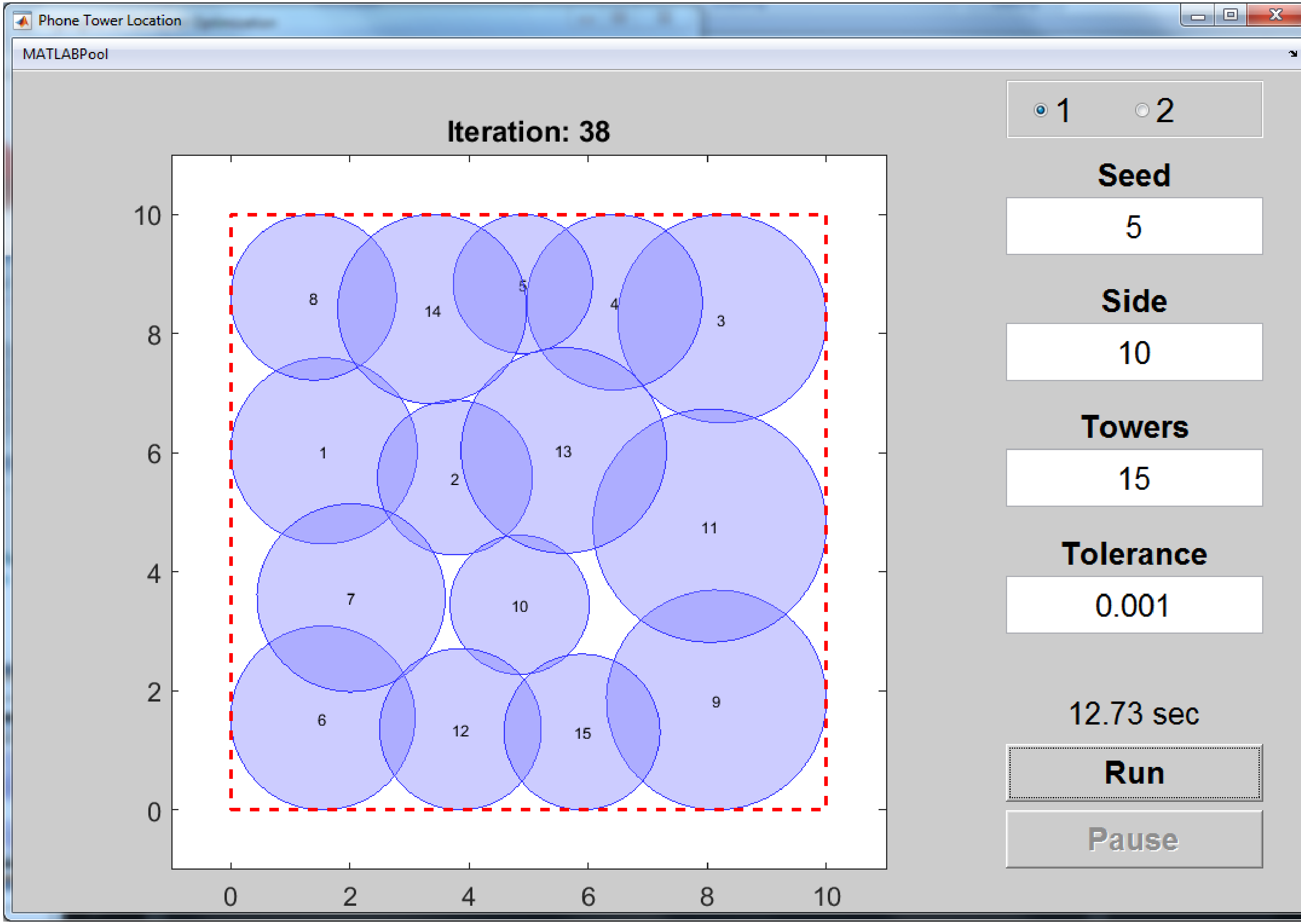
Cell tower location optimization



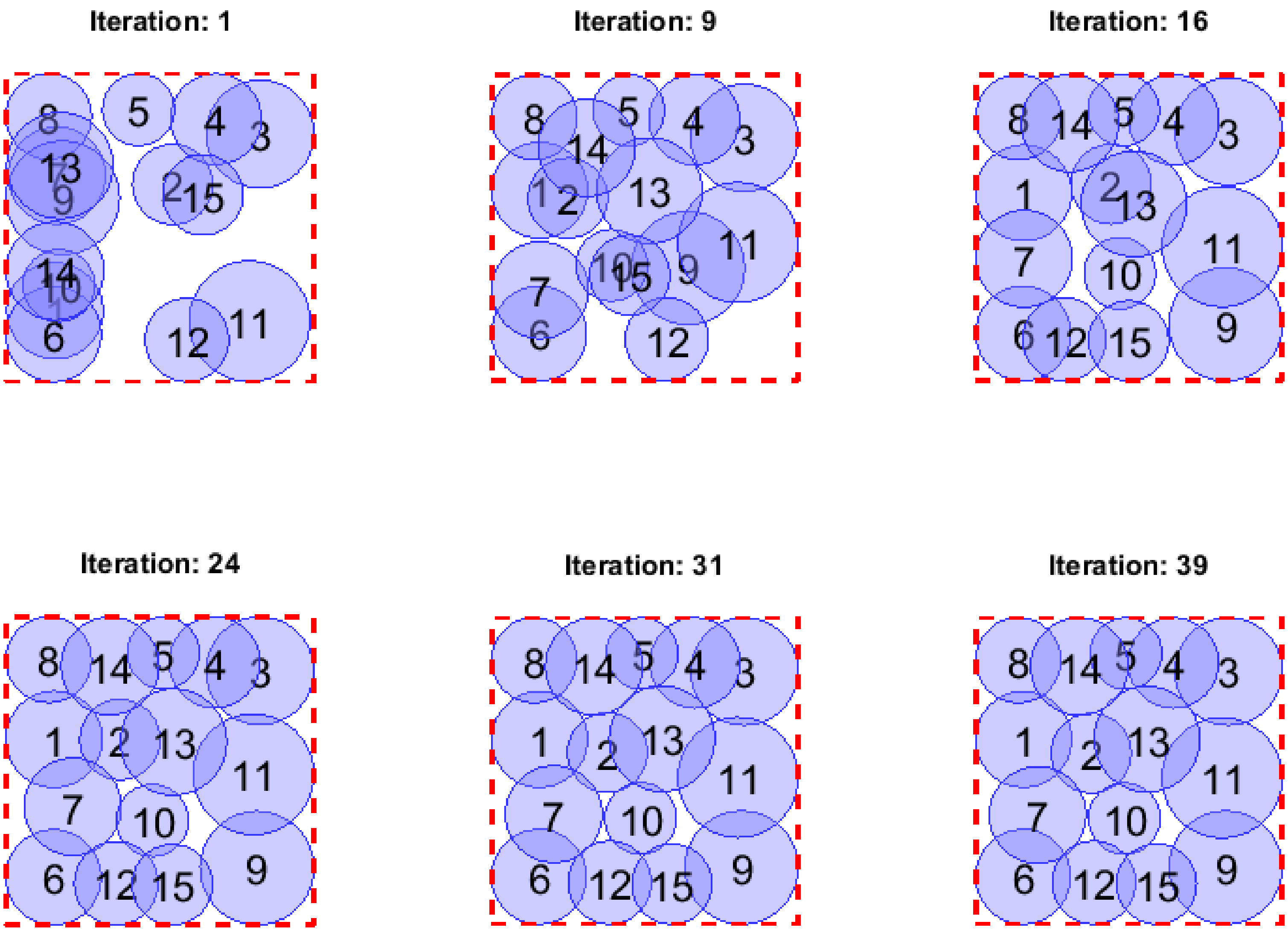
Cell tower location optimization



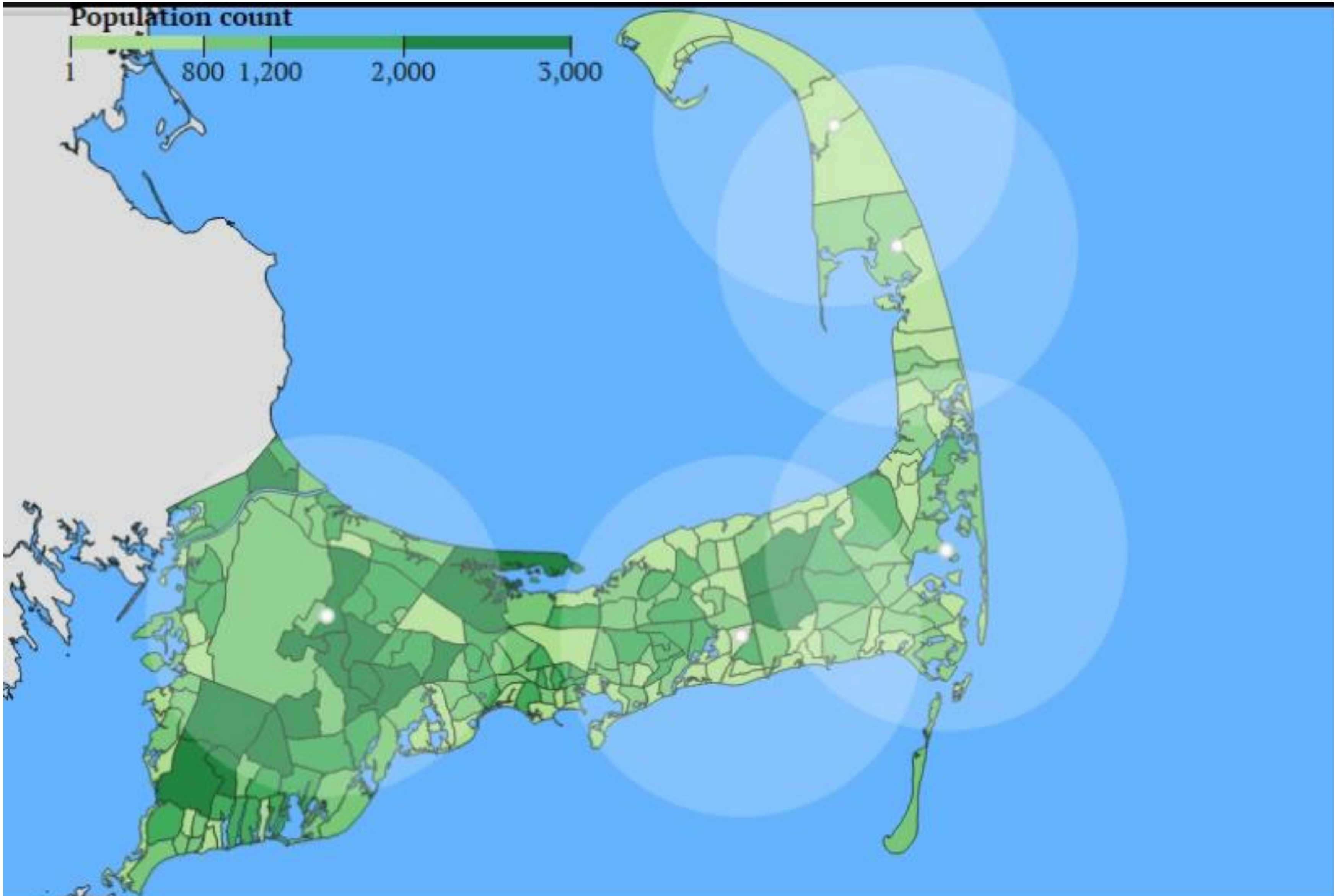
Cell tower location optimization



Cell tower location optimization



Cell tower location optimization



Aircraft conflict avoidance

Aircraft i and j are **in conflict** if

- their horizontal distance is less than d:

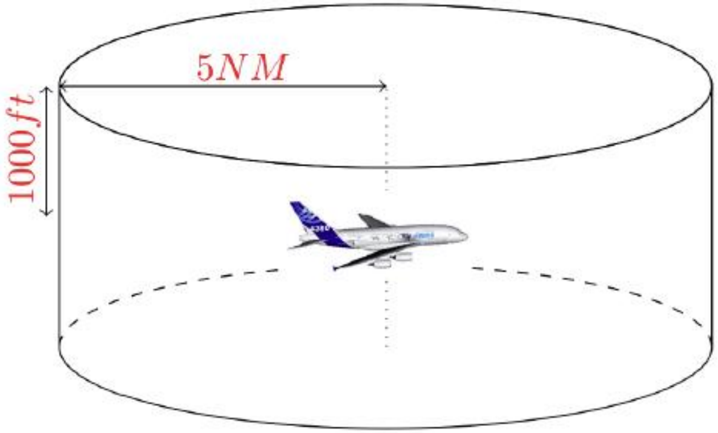
  $\|x_i(t) - x_j(t)\| \le d \quad \forall t \qquad (d = 5\text{NM})$

- their altitude difference is less than h:

  $\|h_i(t) - h_j(t)\| \le h \quad \forall t \qquad (h = 1000\text{ft})$

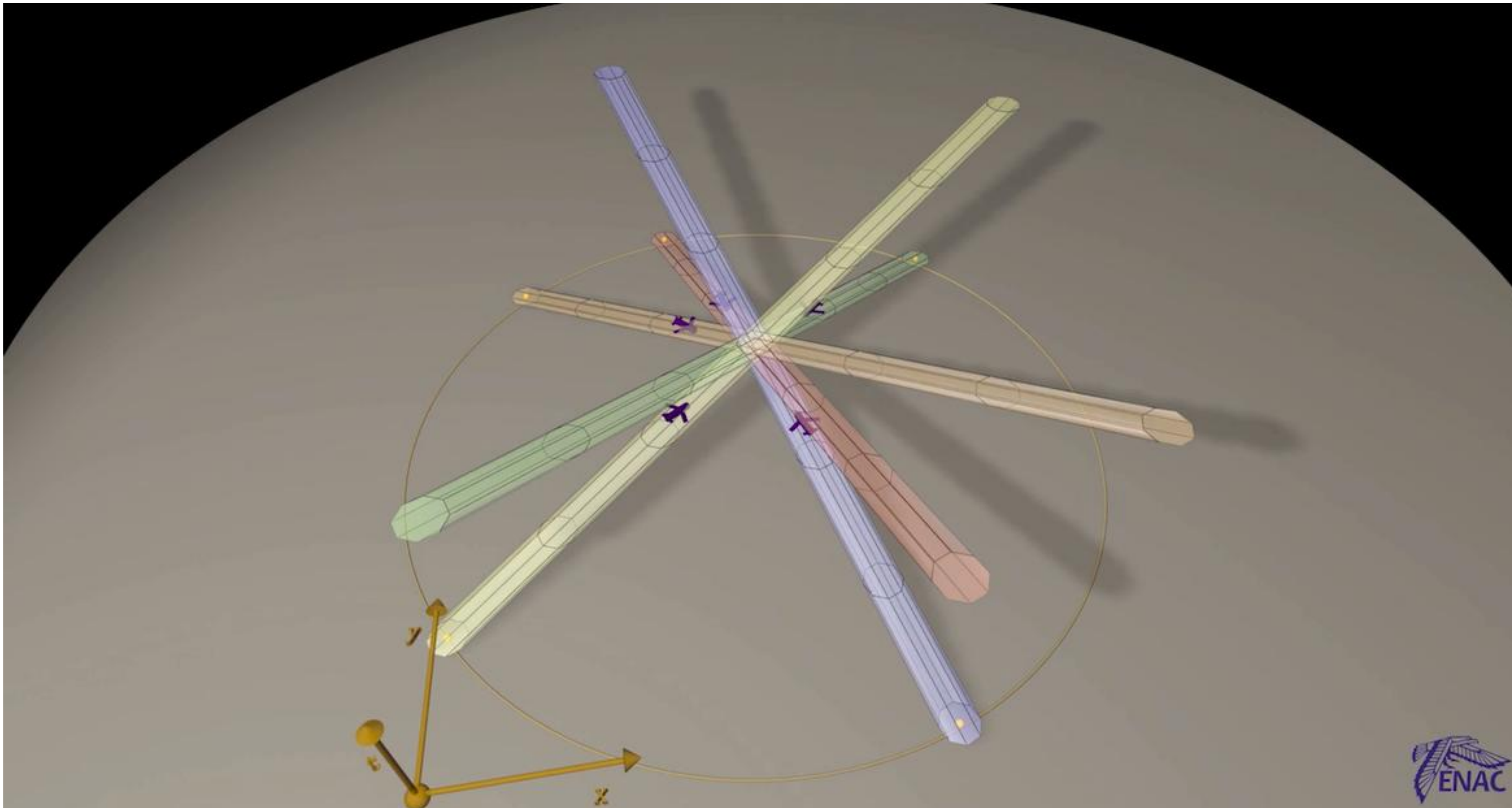
1 NM (nautical mile) = 1852 m

1 ft (feet) = 0.3048 m



Source: Marcel Mongeau, Solving aircraft conflicts by continuous optimization and mixed-integer nonlinear programming

Aircraft conflict avoidance



Source: Marcel Mongeau, Solving aircraft conflicts by continuous optimization and mixed-integer nonlinear programming