Handling Non-Linearities with Gurobi

Gurobi Live Barcelona

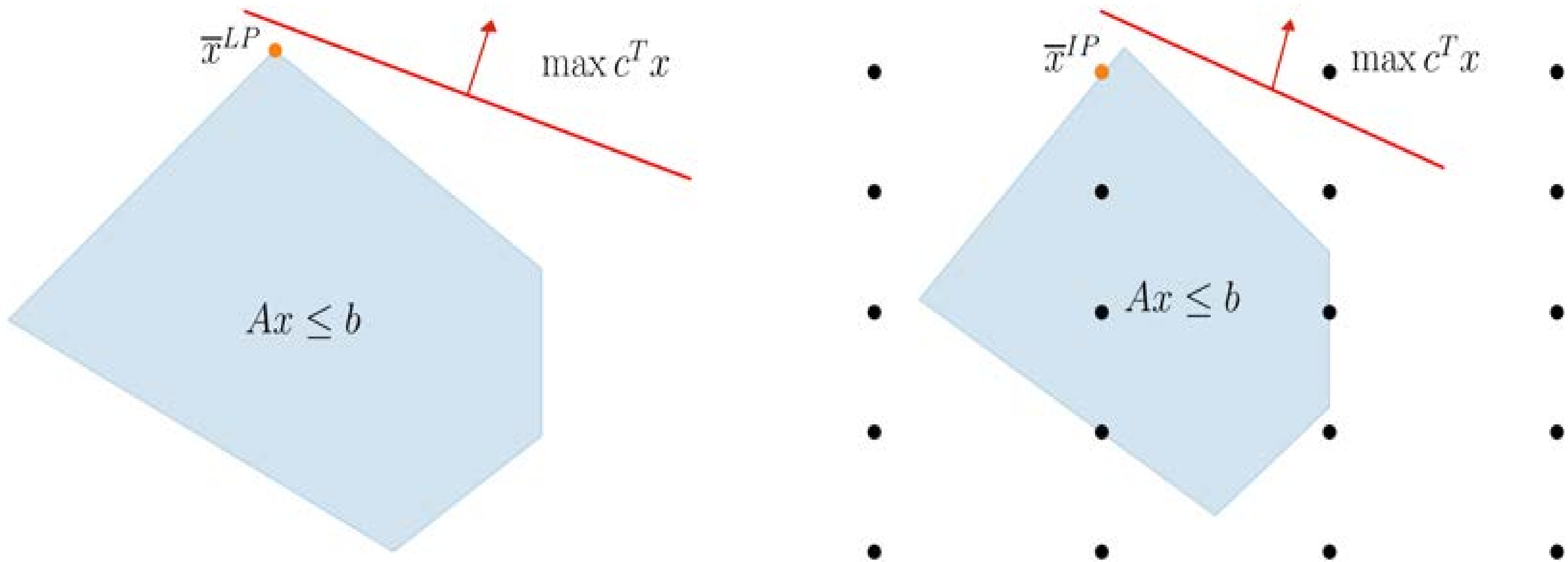
Pierre Bonami

Agenda

1. Introduction
2. Mixed-Integer Convex Quadratic Optimiziation
3. Non-Convex Quadratic Optimization
4. General Functions, piece-wise linear and beyond

GUROBI LiVE!

LP and MIP



- Base problems that Gurobi solves
- Simplex and Barrier algorithms for LP
- Branch-and-cut for MIP

Nonlinearities

Application	Phenomenon	Nonlinearity
Finance	risk	quadratic (convex)
Truss topology	physical forces	quadratic (convex)
Pooling (petrochemical, mining, agriculture)	mixing products	quadratic non-convex
electricity distribution (ACOPF)	Alternative Current	$\sin$ and $\cos$ (can be made quadratic)
machine learning	—	logistic function, $\tanh$
chemical engineering	chemical reactions	—

- many more…

The MINLP Goal

Ideally, we seek to solve

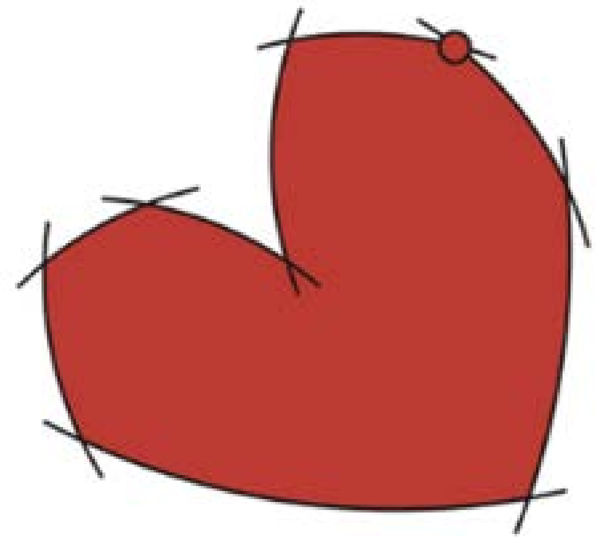
$$
\begin{aligned}
&\min f(x) \\
&\text{s.t:} \\
&g_i(x) \leq 0, i = 1, \ldots, m \\
&x_j \in \mathbb{Z}, j \in \square \\
&l \leq x \leq u
\end{aligned}
$$

- $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, reasonably smooth
- If l or u is not finite: undecidable in general
- Even assuming finiteness it’s very difficult in theory and practice
- Note that integer variables are “not convex” and can be represented as a polynomial if bounded

Convex or not convex?

Convex Region

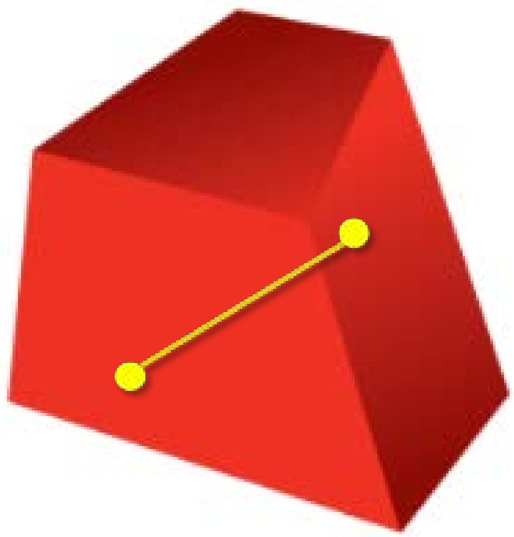
Any segment connecting two points inside the region is inside the region.



Non-Convex Region

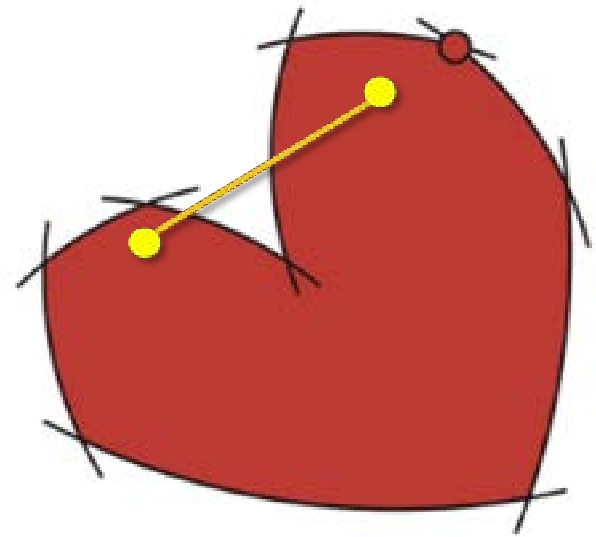
There exists two points in the region the segment connecting them is not completely in the region.

Convex or not convex?



Convex Region

Any segment connecting two points inside the region is inside the region.

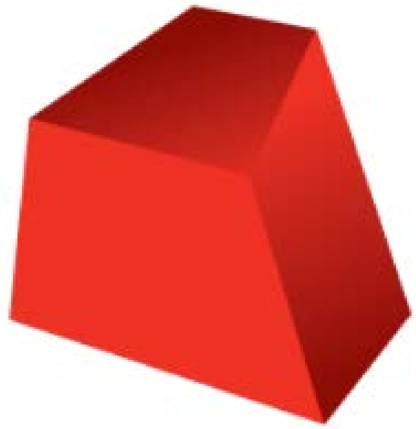


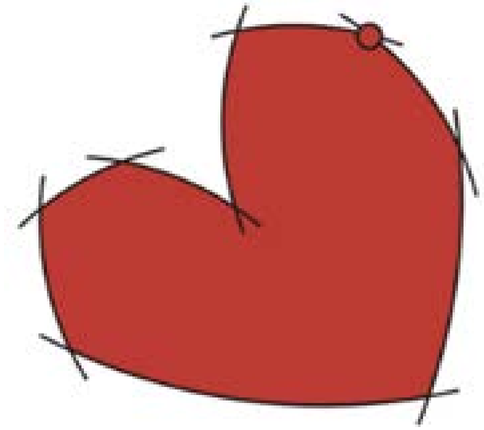
Non-Convex Region

There exists two points in the region the segment connecting them is not completely in the region.

GUROBI LiVE!

Why?

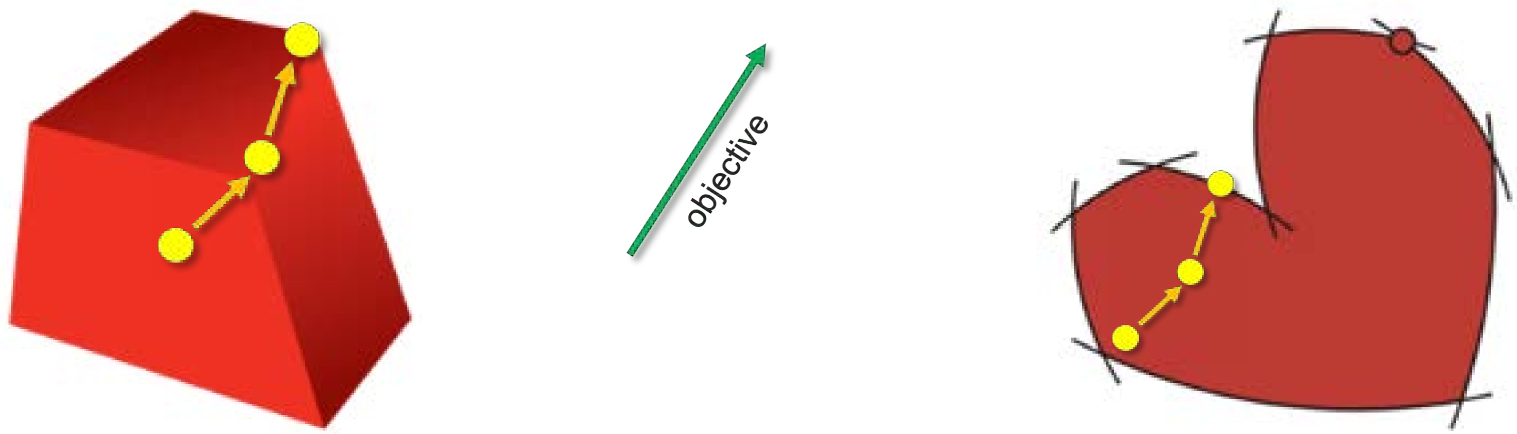




- Optimization "easy" (usually)
  - Start from any point in the region
  - Take steps inside the region improving objective
  - When there is no more step **global optimum**
- Interior point methods for any closed convex region
- Simplex algorithm for polyhedra

- Optimization "hard"
  - Start from any point in the region
  - Take steps inside the region improving objective
  - When there is no more step **local optimum** is reached
- Need a divide-and-conquer algorithm to find a **global optimum**.

Why?



- Optimization “easy” (usually)
  - Start from any point in the region
  - Take steps inside the region improving objective
  - When there is no more step **global optimum**
- Interior point methods for any closed convex region
- Simplex algorithm for polyhedra

- Optimization “hard”
  - Start from any point in the region
  - Take steps inside the region improving objective
  - When there is no more step **local optimum** is reached
- Need a divide-and-conquer algorithm to find a **global optimum**.

Nonlinearities in Gurobi

Convex

- Quadratic objective: $\min c^T x + x^T Q x$ with $Q \succeq 0$
- Quadratic constraints: $a^T x + x^T Q x \leq b$ describing a convex region
  - $Q \succeq 0$ is a simple case, can be more complex

Non-Convex

- Discrete objects: integer variables, SOS constraints
- Bilinear terms: $z_{ij} = x_i x_j$
  - Non-Convex quadratic forms: $a^T x + x^T Q x \leq b$
- General functions: $\exp$, $\log$, $\cos$,
  - Reformulated as PWL in Gurobi 9 and 10
  - Treated directly

Nonlinearities in Gurobi

Convex

- Quadratic objective: $\min c^T x + x^T Q x$ with $Q \succeq 0$
- Quadratic constraints: $a^T x + x^T Q x \leq b$ describing a convex region
  - $Q \succeq 0$ is a simple case, can be more complex

Non-Convex

- Discrete objects: integer variables, SOS constraints
- Bilinear terms: $z_{ij} = x_i x_j$
  - Non-Convex quadratic forms: $a^T x + x^T Q x \leq b$
- General functions: $\exp, \log, \cos,$
  - Reformulated as PWL in Gurobi 9 and 10
  - Treated directly

New in Gurobi 11

Agenda

1. Introduction
2. Mixed-Integer Convex Quadratic Optimization
3. Non-Convex Quadratic Optimization
4. General Functions, piece-wise linear and beyond

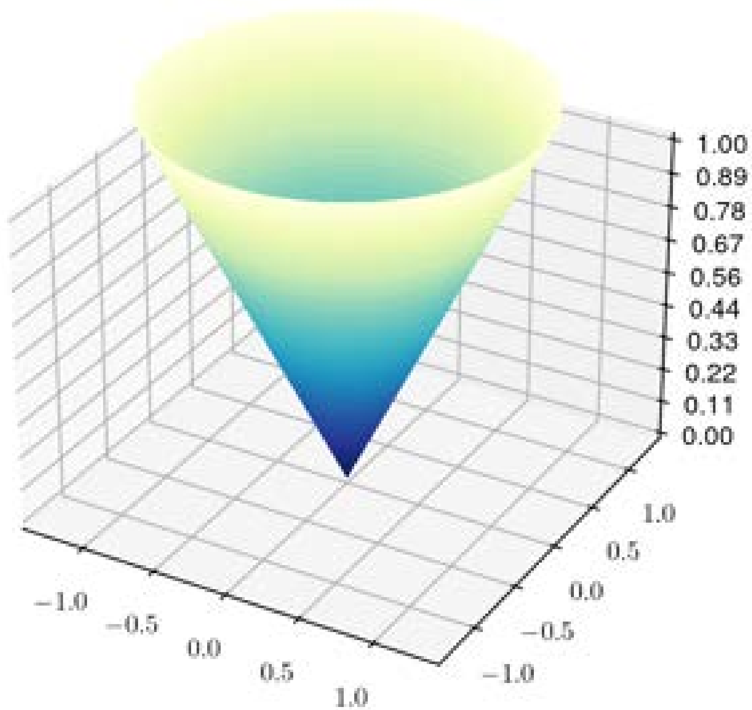
GUROBI LIVE!

Problem definition

$$\min c^T x + x^T Q^0 x$$
$$\text{s.t:}$$
$$a_k{}^T x + x^T Q^k x \leq b_k, k = 1, \ldots, m$$
$$x_j \in \mathbb{Z}, j \in \square$$
$$l \leq x \leq u$$

- $Q^0, Q^1, \ldots, Q^m$ are assumed to be symmetric
  - Q^0 is positive semi definite
  - The quadratic forms $a_k^T + x^T Q^k x - b_k$ are *second order cone representable*.

The second order cone



$\square^n = \{x \in \mathbb{R}^{n+1} : \sum_{j=1}^{n} x_j^2 \leq x_0^2, x_0 \geq 0\}$

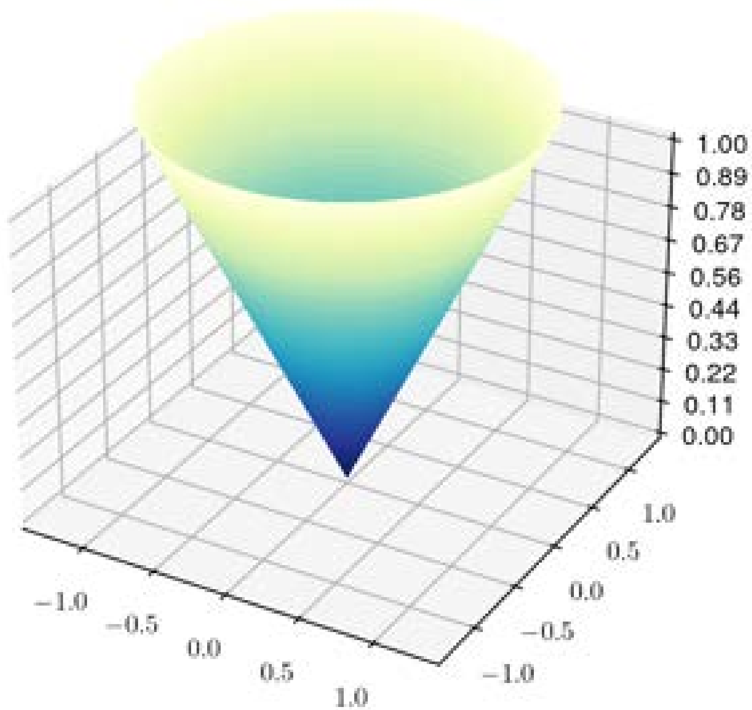
Through simple algebra, can be represented as SOC:

- $\sum_{i=1}^{n} x_i^2 \leq x_0^2$, with $x_0 \geq 0$
- $\sum_{i=2}^{n} x_i^2 \leq x_0 x_1$, with $x_0, x_1 \geq 0$ (rotated SOC)
- $a^T x + x^T Q x \leq b$, with $Q \geq 0$
- $x^T Q x \leq y^2$, with $Q \geq 0, y \geq 0$

Very powerful but modeling sometimes far from obvious.

Not all forms recognized by solvers

The second order cone



$\square^n = \{x \in \mathbb{R}^{n+1} : \sum_{j=1}^{n} x_j^2 \leq x_0^2, x_0 \geq 0\}$

Through simple algebra, can be represented as SOC:

- $\sum_{i=1}^{n} x_i^2 \leq x_0^2$, with $x_0 \geq 0$
- $\sum_{i=2}^{n} x_i^2 \leq x_0 x_1$, with $x_0, x_1 \geq 0$ (rotated SOC)
- $a^T x + x^T Q x \leq b$, with $Q \succeq 0$
- $x^T Q x \leq y^2$, with $Q \succeq 0, y \geq 0$

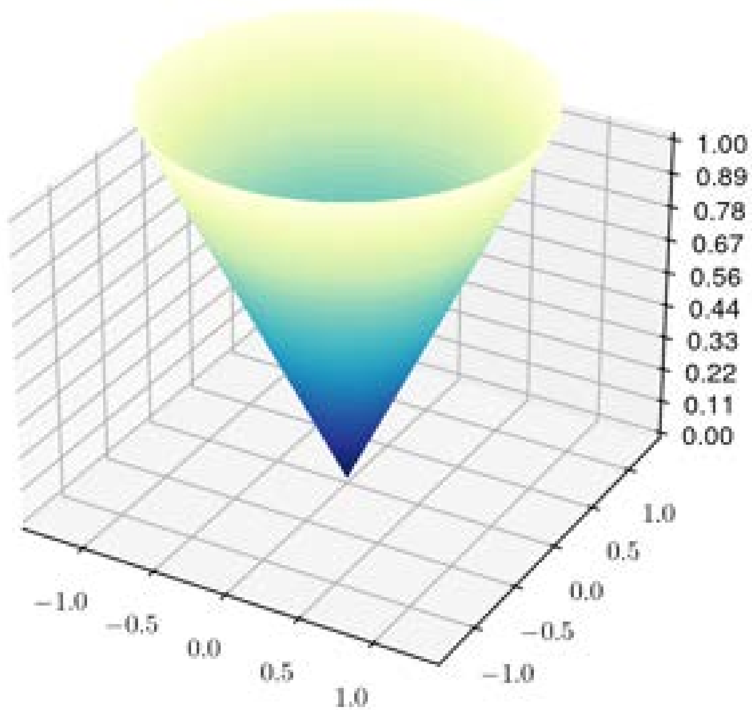
Very powerful but modeling sometimes far from obvious.

New in Gurobi 11

Not all forms recognized by solvers

GUROBI LiVE!

The second order cone



$\square^n = \{x \in \mathbb{R}^{n+1} : \sum_{j=1}^{n} x_j^2 \leq x_0^2, x_0 \geq 0\}$

Through simple algebra, can be represented as SOC:

- $\sum_{i=1}^{n} x_i^2 \leq x_0^2$, with $x_0 \geq 0$
- $\sum_{i=2}^{n} x_i^2 \leq x_0 x_1$, with $x_0, x_1 \geq 0$ (rotated SOC)
- $a^T x + x^T Q x \leq b$, with $Q \succeq 0$
- $x^T Q x \leq y^2$, with $Q \succeq 0, y \geq 0$

Very powerful but modeling sometimes far from obvious.

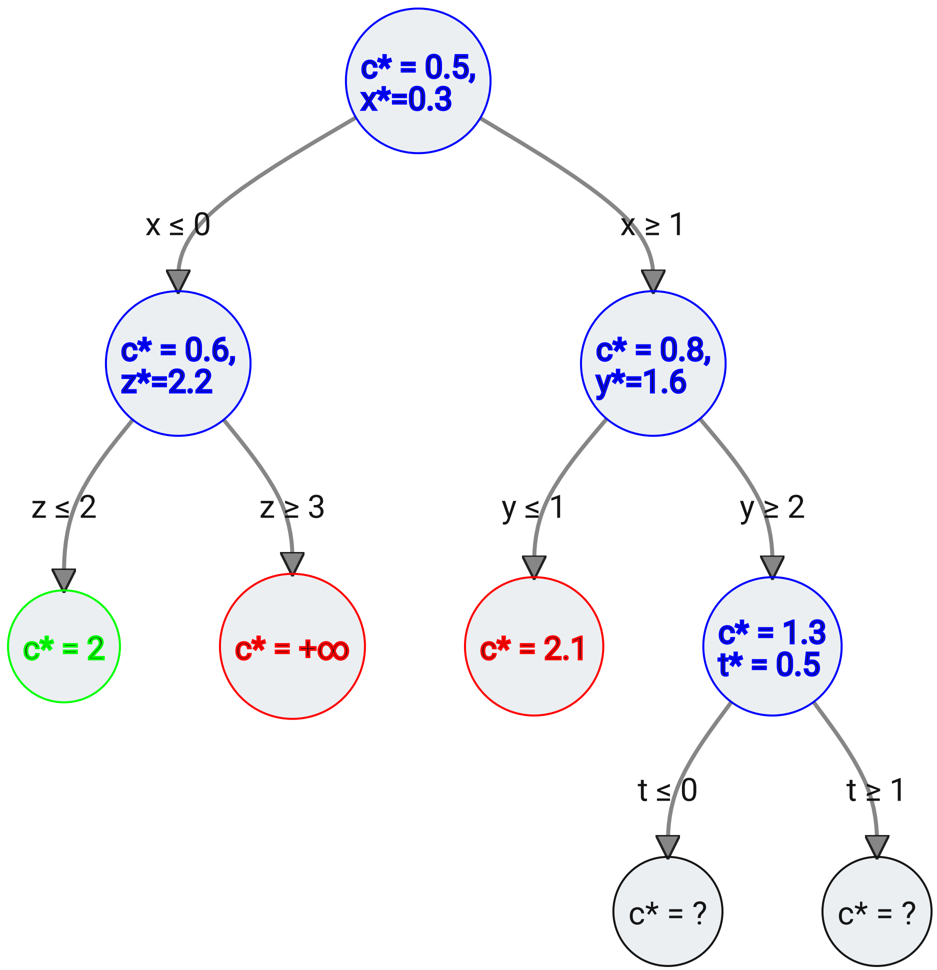
Not all forms recognized by solvers

New in Gurobi 11

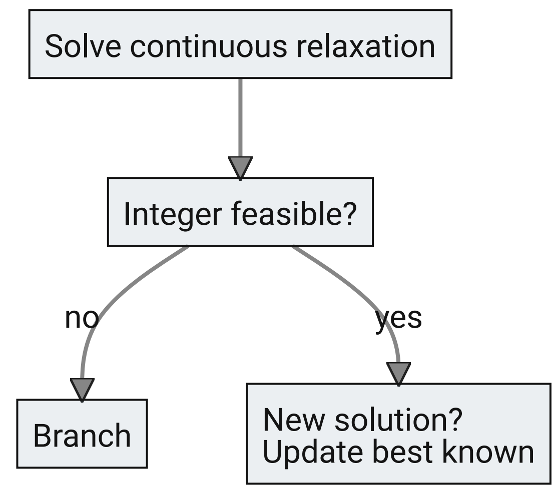
Note

GUROBI LiVE!

The basic branch-and-bound algorithm



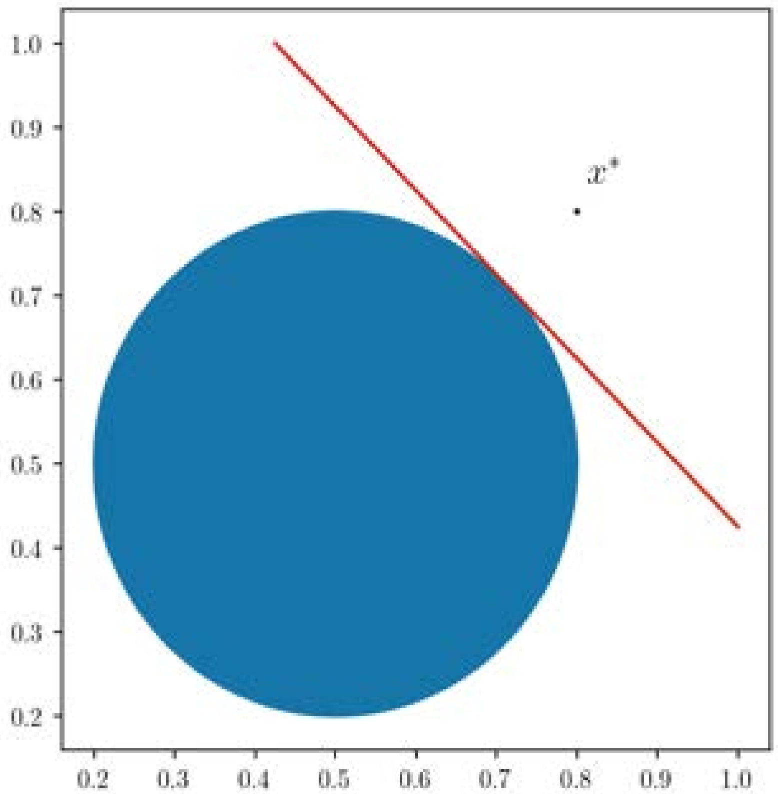
At each node of the tree:



In MILP and MIQP continuous relaxation usually solved by simplex.

In MISOCP/MIQCP, continuous relaxation solved by barrier.

The outer approximation cut



- Let $C = \{g(x) \leq b : x \in R^n\}$, with g a convex function
- For any $x^* \in \mathbb{R}^n$, the constraint:

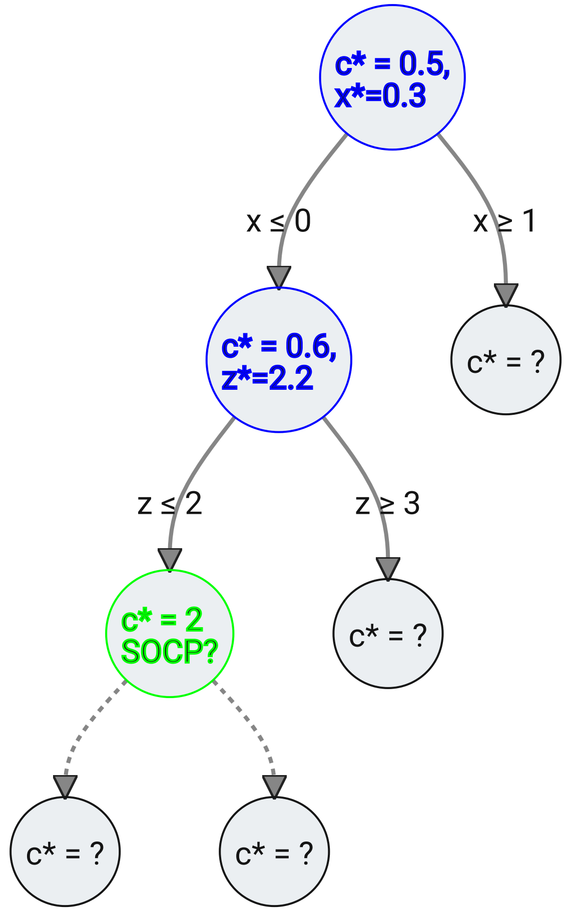
$$\nabla g(x^*)(x - x^*) + g(x^*) \leq 0$$

  is valid
- If $x^* \notin C$, it cuts x^*:

$$\nabla g(x^*)(x^* - x^*) + g(x^*) > 0$$

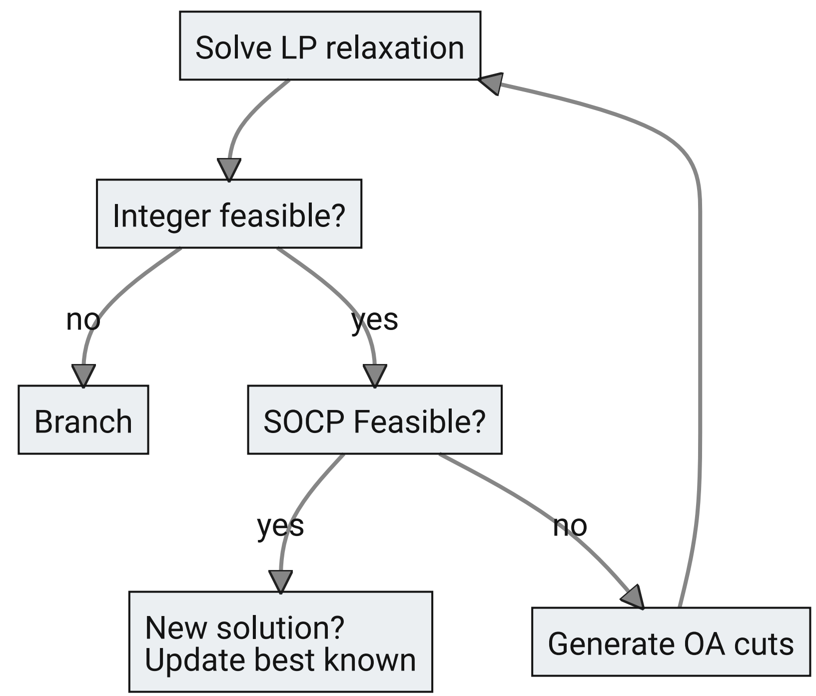
GUROBI LiVE!

Outer approximation branch-and-cut



Drop quadratic constraints and solve an LP relaxation at each node.

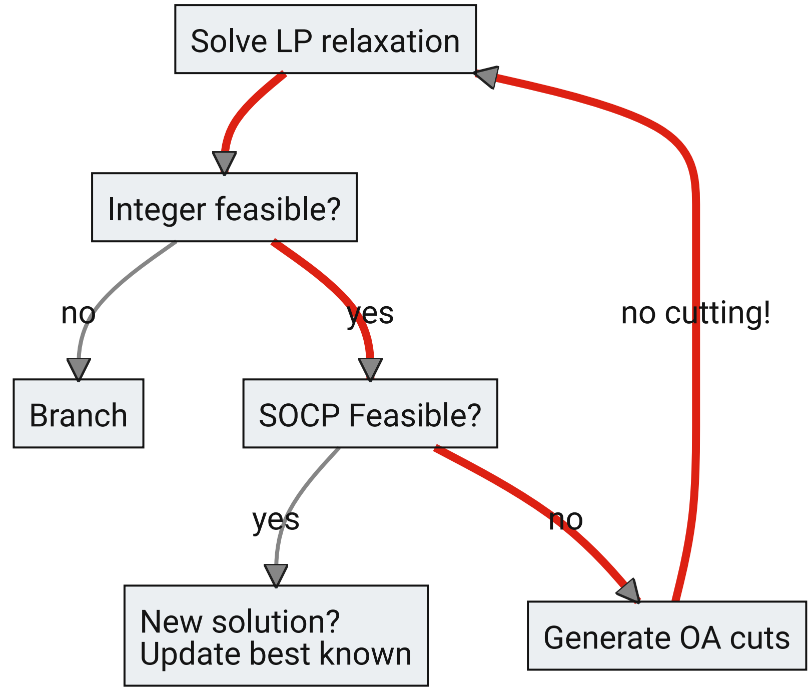
Integer feasible nodes are not necessarily solutions.



GUROBI LiVE!

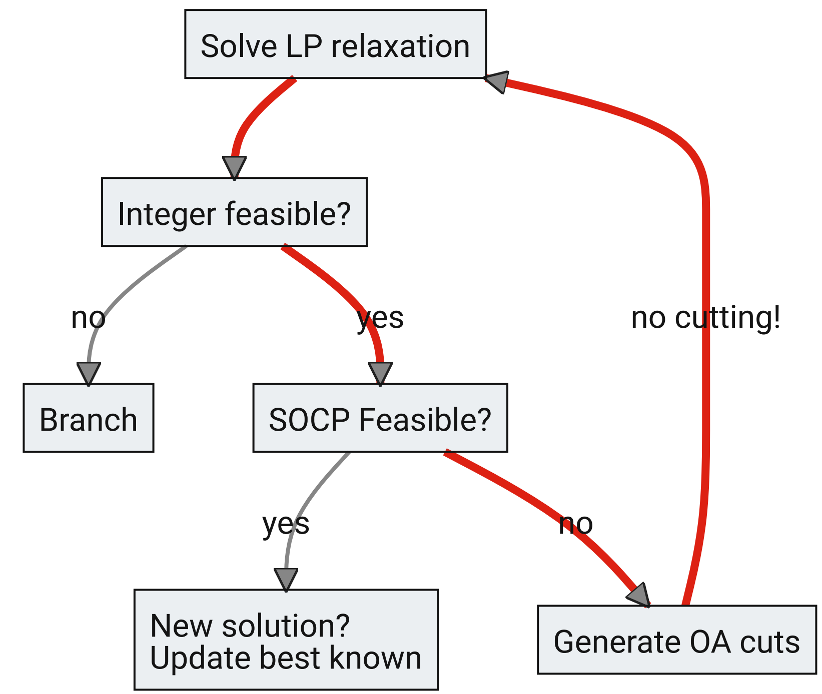
Numerical difficulties

- OA branch-and-cut builds a cutting plane approximation of smooth functions
- It can happen that node solution:
  - is integer feasible
  - is not SOC feasible
  - OA cuts are not cutting enough



Numerical difficulties

- OA branch-and-cut builds a cutting plane approximation of smooth functions
- It can happen that node solution:
  - is integer feasible
  - is not SOC feasible
  - OA cuts are not cutting enough



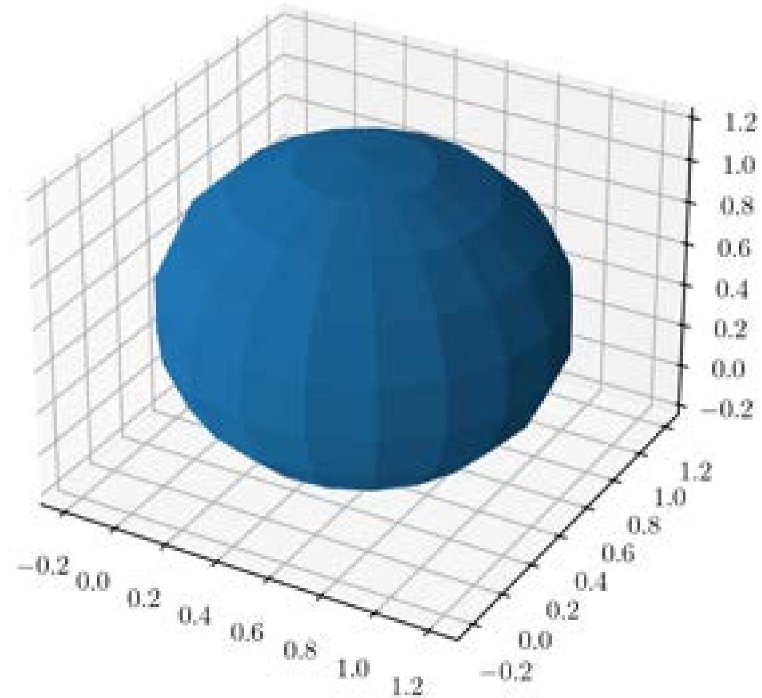
New in Gurobi 11

Rely on barrier algorithm for those nodes (usually very few).

GUROBI LIVE!

Cone disaggregation and outer approximation

An exponential number of cutting planes is needed to approximate a convex quadratic form.



Cone disaggregation

From

$$\sum_{i=1}^{n} x_i^2 \leq x_0^2, x_0 \geq 0$$

- Create variables $y_i \geq 0$, such that $x_i^2 \leq y_i x_0$ (rotated SOC)
- Replace initial constraint with $\sum_{i=1}^{n} y_i \leq x_0$

Pitfalls of disaggregation

$$(A)\begin{cases}\sum_{i=1}^{n} x_i^2 \le x_0^2, \\ x_0 \ge 0\end{cases} \qquad (B)\begin{cases}\sum_{i=1}^{n} y_i \le x_0 \\ x_i^2 \le y_i x_0 & i = 1, \dots, n \\ y \ge 0, x \ge 0\end{cases}$$

- The reformulation is correct: every solution of (B) translates to (A)
- But a solution of (B) with a small infeasibility can have a large one in (A):
- Suppose $x_0 = 1, y_i = \frac{1}{n}$ and $x_i = \sqrt{\frac{1}{n} + \epsilon}$:
  - Infeasibility in (B) is ϵ
  - Infeasibility in (A) is $n \cdot \epsilon$

Gurobi tries to deal with it but can be an issue.

GUROBI LiVE!

Options for MISOCP/MIQCQP

MIQCPMethod

- -1 Automatic choice (default)
- 0 Use QCP branch-and-bound
- 1 Use Outer Approximation

PreMIQCPForm

- -1 Automatic choice (default)
- 0 Leave the model as is (for B&B)
- 1 Reformulate to SOC
- 2 Reformulate to SOC and disaggregate

Example

Portfolio Optimization

GUROBI LiVE!

Agenda

1. Introduction
2. Mixed-Integer Convex Quadratic Optimization
3. Non-Convex Quadratic Optimization
4. General Functions, piece-wise linear and beyond

GUROBI LIVE!

Stepping into a non-convex world

Non-Convex MIQCQP

$$\min c^T x + x^T Q^0 x$$

$$\text{s.t:}$$

$$a_k{}^T x + x^T Q^k x \leq b_k, \qquad k = 1, \ldots, m$$

$$x_j \in \mathbb{Z}, j \in \square$$

$$l \leq x \leq u$$

- $Q^0, Q^1, \ldots, Q^m$ are assumed to be symmetric
- Continuous relaxation is NP-hard!
- Solution strategy:
  - Build a convex relaxation
  - Refine it through branching.

Non-Convex MIQCQP

$$\min c^T x + x^T Q^0 x$$

s.t:

$$a_k^T x + x^T Q^k x \le b_k, \qquad k = 1, \dots, m$$

$$x_j \in \mathbb{Z}, j \in \square$$

$$l \le x \le u$$

- $Q^0, Q^1, \dots, Q^m$ are assumed to be symmetric
- Continuous relaxation is NP-hard!
- Solution strategy:
  - Build a convex relaxation
  - Refine it through branching.

Note

NonConvex parameter in Gurobi

NonConvex

- -1 automatic (default)
- 0 Return error if original model has non-convex Q objective or constraints
- 1 Return error if presolved model has non-convex Q that cannot be linearized
- 2 Accept non-convex Q by building a bilinear formulation

New in Gurobi 11

Default behavior change: New default 2 (was 1)).

Bilinear formulation

- For each product $x_i x_j$ in the model
  - Introduce a new variable z_{ij}
  - Add the **bilinear** constraint $z_{ij} = x_i x_j$
  - Replace product with z_{ij}

$$\min c^T x + \langle Q^0, Z \rangle$$

s.t:

$$a_k{}^T x + \langle Q^k, Z \rangle \leq b_k, \qquad k = 1, \ldots, m$$

$$Z = xx^T$$

$$l \leq x \leq u$$

$$(\langle Q, Z \rangle = \sum_i \sum_j q_{ij} z_{ij})$$

GUROBI LiVE!

More details on bilinear formulation

- Try as much as possible to avoid creating bilinear terms:
  - if one variable is fixed
  - if one variable is binary (can be reformulated)
  - square of binary $x^2 = x$
  - square term $q_{ii}x^2$ with $q_i i > 0$ is convex
- If $x_i x_j$ always appears in inequalities with q_{ij} of same sign relax to:
  - $z_{ij} \geq x_i x_j$, if $q_{ij} > 0$
  - $z_{ij} \leq x_i x_j$, if $q_{ij} < 0$

More details on bilinear formulation

- Try as much as possible to avoid creating bilinear terms:
  - if one variable is fixed
  - if one variable is binary (can be reformulated)
  - square of binary $x^2 = x$
  - square term $q_{ii}x^2$ with $q_i i > 0$ is convex
- If $x_i x_j$ always appears in inequalities with q_{ij} of same sign relax to:
  - $z_{ij} \geq x_i x_j$, if $q_{ij} > 0$
  - $z_{ij} \leq x_i x_j$, if $q_{ij} < 0$

⚠ **Warning**

Bilinear relaxation

- Relax non-convex constraint $Z = xx^T$ using convex enveloppes.

$$\min c^T x + \langle Q^0, Z \rangle$$

s.t:

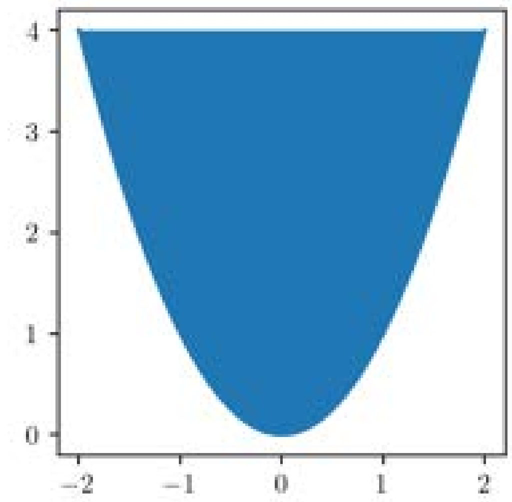
$$a_k^T x + \langle Q^k, Z \rangle \leq b_k, \qquad k = 1, \ldots, m$$

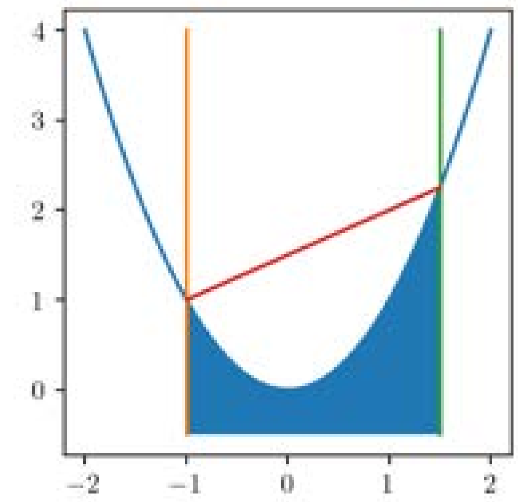
$$z^-(x_i, x_j) \leq z_{ij} \leq z^+(x_i, x_j)$$

$$l \leq x \leq u$$

Convex envelopes: parabola

Consider the square case: $z = x^2$





$z \geq x^2$

It is convex:

$$z^-(x_i, x_i) = x_i^2$$

Can be dealt with by OA.

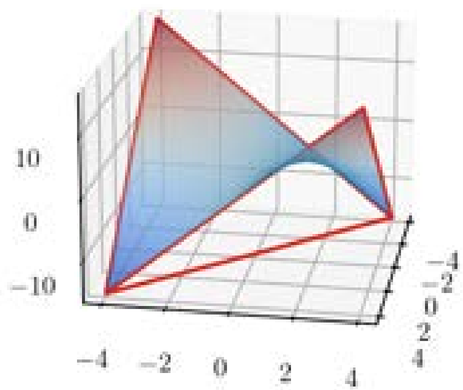
$z \leq x^2, -1 \leq x \leq 1.5$

$z^+(x_i, x_i)$ is given by the secant:

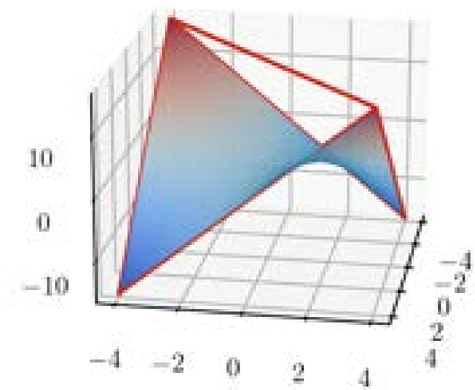
$$z_{ii}^+ = (u+1)x_i - 1 \cdot u$$

GUROBI LIVE!

Convex envelopes: products (McCormick)



Lower enveloppe z_{ij}^-

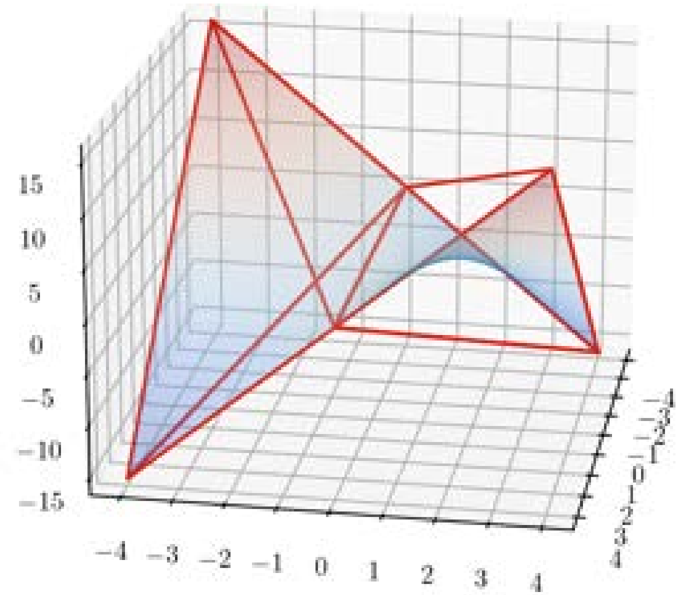


Upper enveloppe z_{ij}^+

$$z_{ij}^- = \max\left\{\begin{array}{l} l_j x_i + l_i x_j - l_i l_j \\ u_j x_i + u_i x_j - u_i u_j \end{array}\right\} \leq z_{ij} \leq z_{ij}^+ = \min\left\{\begin{array}{l} l_j x_i + u_i x_j - u_i l_j \\ u_j x_i + l_i x_j - l_i u_j \end{array}\right\}$$

Spatial branching

- Let (x^*, z^*) be the solution of the bilinear relaxation
- If not integer feasible, can branch on an integer variable
- Otherwise:
  - If $z_{ij}^* = x_i^* x_j^*$, for all bilinear term, we have a solution
  - Otherwise refine our bilinear relaxation:
    - Pick x_i or x_j s.t. $z_{ij}^* \neq x_i^* x_j^*$
    - Create two child nodes with $x_i \leq x_i^*$ and $x_i \geq x_i^*$
    - Refine bilinear relaxation in the two nodes



GUROBI LiVE!

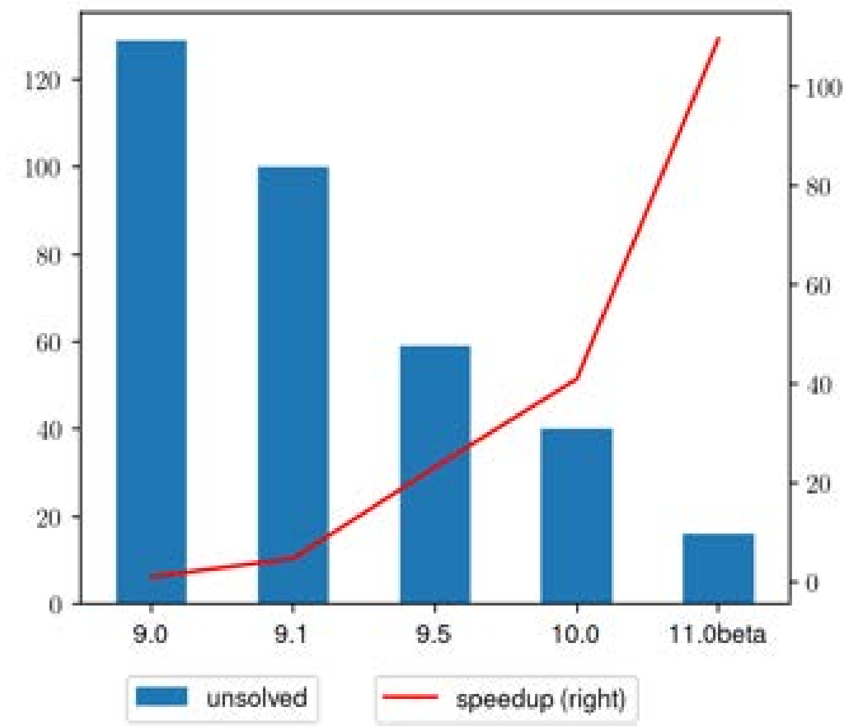
Other techniques targeted at non-convex MIQCQP

- `RLTCuts`
  - Reformulation Linearization Technique (Sherali and Adams, 1990)
  - Multiply linear constraint by a variable, linearize resulting products
- `BQPCuts`
  - Facets of the *Binary Quadratic Polytope* (Padberg 1989)
  - Clique cuts from the paper
- `SDPCuts`
  - Relax $Z = xx^T$ to $Z \succeq xx^T$, and outer approximate resulting cone
- `OBBT`:
  - Optimization based bound tightening
  - Infer tighter bound on variables involved in products by LP.

Non-convex MIQCQP performance history

New in Gurobi 11

- Improved recognition of convexity
- Branching improvements
  - Strong branching for bilinear terms
  - Better choice for deciding to branch on an integer or bilinear



874 models – discarded: 40 due to inconsistent answers, 278 that none of the versions can solve

Example solve

Pooling problem from MINLPLIB

```
m.optimize()
```

```
Gurobi Optimizer version 11.0.0 build v11.0.0beta2 (mac64[x86] - macOS
13.6 22G120)

CPU model: Intel(R) Core(TM) i5-1038NG7 CPU @ 2.00GHz
Thread count: 4 physical cores, 8 logical processors, using up to 8
threads

Optimize a model with 662 rows, 403 columns and 2229 nonzeros
Model fingerprint: 0x883de6ff
Model has 70 quadratic constraints
Variable types: 295 continuous, 108 integer (108 binary)
Coefficient statistics:
  Matrix range     [2e-03, 1e+03]
  QMatrix range    [1e+00, 1e+00]
  QLMatrix range   [1e+00, 1e+00]
```

Agenda

1. Introduction
2. Mixed-Integer Convex Quadratic Optimization
3. Non-Convex Quadratic Optimization
4. General Functions, piece-wise linear and beyond

Function constraints in Gurobi

Since Gurobi 9.0. Allow to state $y = f(x)$

- f is a predefined function
- y and x are one-dimensional variables

Library of predefined functions include:

- e^x, a^x, $\ln(x)$, $\log_a(x)$, logistic
- $\sin(x)$, $\cos(x)$, $\tan(x)$,
- monomials x^a, polynomials of one variable $a_0 + a_1 x^1 + a_2 x^2 + \ldots$

Function constraints in Gurobi

Since Gurobi 9.0. Allow to state $y = f(x)$

- f is a predefined function
- y and x are one-dimensional variables

Library of predefined functions include:

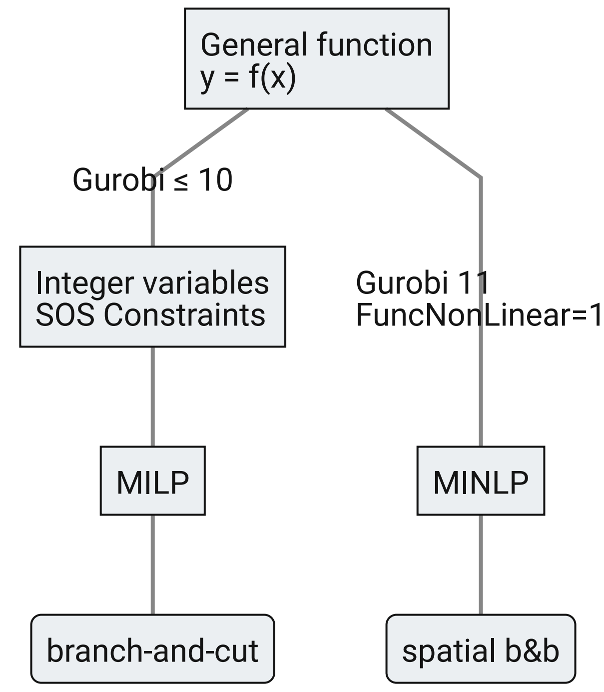
- e^x, a^x, $\ln(x)$, $\log_a(x)$, logistic
- $\sin(x)$, $\cos(x)$, $\tan(x)$,
- monomials x^a, polynomials of one variable $a_0 + a_1 x^1 + a_2 x^2 + \ldots$

Example:

```
m = gp.Model()
x = m.addVar()
y = m.addVar()
m.addGenConstrLog(x, y)
```

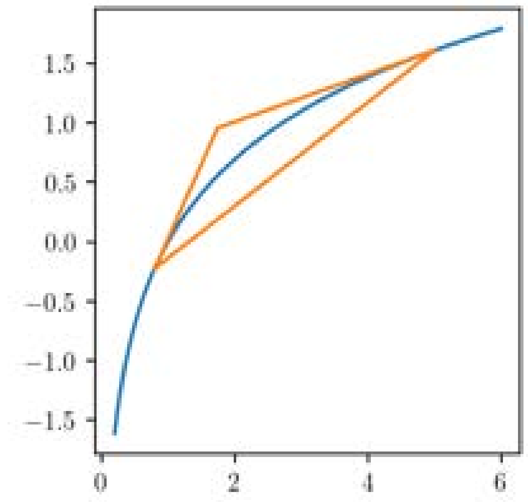
New treatment in Gurobi 11

- Gurobi 9.0-10.0: nonlinear functions replaced during presolve by a piece-wise linear approximation.
- Gurobi 11, can treat nonlinear functions directly:
  - Set `FuncNonLinear=1`
  - No other changes to users' code



Algorithmic approach

- Similar to bilinear formulation/relaxation
- For each function, compute lower/upper envelope
- Spatial branching to refine them
- Additional difficulties:
  - detect when functions are convex/concave
  - functions can be locally convex



$y = \log(x), 0.8 \le x \le 5$

Great powers and great responsabilities

- Gurobi 11.0 handles select univariate nonlinear functions
- But, those can be composed

E.g.: Consider for $x \geq 0$:

$$f(x) = \sqrt{1 + x^2} + \ln(x + \sqrt{1 + x^2}) \leq 2$$

We can formulate it:

- introduce auxiliary variables $u, v, w, z \geq 0$
- add constraints $u = 1 + x^2, v = \sqrt{u}, w = x + v, z = \ln(w)$
- $f(x) = v + z$

Caution

Feasibility tolerances!

Decomposition leading to large infeasibility

$$y = f(x) = \frac{x}{\sin(x)}$$

a solution is $x = 0.0001, y = 1.0000000016666666.$

Now decompose $f(x)$: $u = \sin(x), v = \frac{1}{u}, y = x \cdot v.$

And consider:

- $\overline{x} = 0.0001$
- $\overline{u} = 0.00009899999983333343$ $(\sin(\overline{x}) - 10^{-6})$
- $\overline{v} = 10101.010118015167$
- $\overline{y} = 1.0101010118015167$

⚠ 😱

We have $\overline{y} - f(\overline{x}) \approxeq 10^{-2}$

Example

Logistic Regression

Thank You!

For more information: www.gurobi.com

GUROBI LiVE!