Using quadratic cuts to iteratively strengthen convexifications of box quadratic programs

Amélie Lambert - Daniel Porumbel

Conservatoire National des arts et Métiers - Cédric

Presentation of the problem and goal

We consider (P) a box-constrained quadratic program :

$$(P)\left\{ \min_{x\in[\ell,u]} \quad f(x) \equiv \langle Q, xx^\top \rangle + c^\top x \right.$$

with a non-convex quadratic objective function $f(x)$.

(where $\langle Q, X\rangle = \sum_i \sum_j Q_{ij} X_{ij}$)

Our goal : solve (P) to global optimality

Presentation of the problem and goal

We consider (P) a box-constrained quadratic program :

$$(P)\left\{ \min_{x\in[\ell,u]} \quad f(x) \equiv \langle Q, xx^\top \rangle + c^\top x \right.$$

with a non-convex quadratic objective function $f(x)$.

(where $\langle Q, X \rangle = \sum_i \sum_j Q_{ij} X_{ij}$)

Our goal : solve (P) to global optimality

A cutting convex quadrics approach :

1. A family of convex piecewise quadratic relaxations
2. A cutting-quadrics algorithm to compute the "best" set of quadric cuts.
3. A spatial B&B based on the computed relaxation.

A family of convex piecewise quadratic relaxations

Literature : Convexification of a quadratic function

$$f(x) = \langle Q, xx^\top \rangle + c^\top x$$

Idea Add quadratic terms to f then compensate linearly

Literature : Convexification of a quadratic function

$$f(x) = \langle Q, xx^\top \rangle + c^\top x$$

Idea Add quadratic terms to f then compensate linearly

→ Add new variables Y_{ij} that satisfy $Y_{ij} = x_i x_j$

→ Introduce a matrix $S \succeq 0$, and the new convex function :

$$f_S(x, Y) = \langle S, xx^\top \rangle + c^\top x + \langle Q - S, Y \rangle$$
$$f_S(x, Y) = \langle Q, xx^\top \rangle + c^\top x = f(x) \quad \text{if } Y = xx^\top$$

[MIQCR - Elloumi-Lambert (2019)]

Literature : Convexification of a quadratic function

$$f(x) = \langle Q, xx^\top \rangle + c^\top x$$

Idea Add quadratic terms to f then compensate linearly

$\rightarrow$ Add new variables Y_{ij} that satisfy $Y_{ij} = x_i x_j$

$\rightarrow$ Introduce a matrix $S \succeq 0$, and the new convex function :

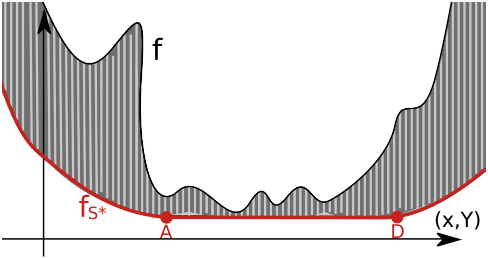
$$f_S(x, Y) = \langle S, xx^\top \rangle + c^\top x + \langle Q - S, Y \rangle$$
$$f_S(x, Y) = \langle Q, xx^\top \rangle + c^\top x = f(x) \quad \text{if } Y = xx^\top$$

[MIQCR - Elloumi-Lambert (2019)]

Observations

- $f_S(x, Y)$ is a convex quadratic function for any matrix $S \succeq 0$
- If $S = \mathbf{0}_n$, we have $f_S(x, Y) = \langle Q, Y \rangle + c^\top x$
  $\Rightarrow f_{0_n}(x, Y)$ is a linear function.

Literature : Best quadratic understimators



MIQCR [Elloumi-Lambert (2019)]

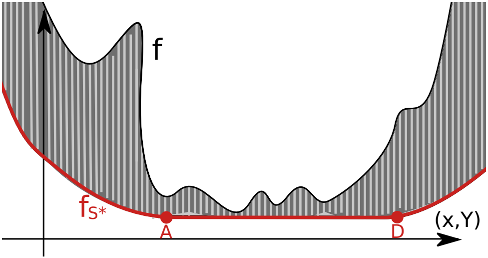
1. At the root node : compute the "best" S^* by solving (SDP) :

$$(SDP) : \left\{ \min_{x \in \mathbb{R}^n,\ Y \in \mathcal{S}_n} \langle Q, Y \rangle + c^T x \ : \ (x, Y) \in \mathcal{MC},\ Y - xx^\top \succeq 0 \right\}$$

2. Solve (P) with a spatial B&B based on this underestimator

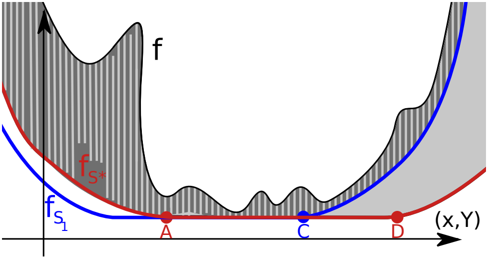
- **Good news :** the root node bound reaches the value of (SDP)
- **Bad news :** Refine S^* at each node of the B&B is too costly.

Convex piecewise quadratic understimators



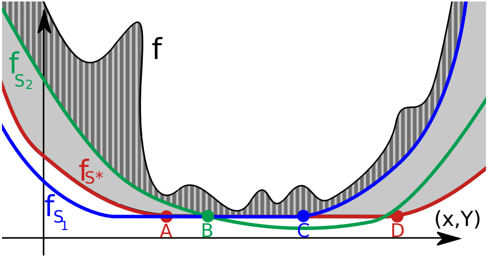
Idea Instead of a unique function f_{S^*}, use multiple functions f_{S_k}

Convex piecewise quadratic understimators



Idea Instead of a unique function f_{S^*}, use multiple functions f_{S_k}

Convex piecewise quadratic understimators

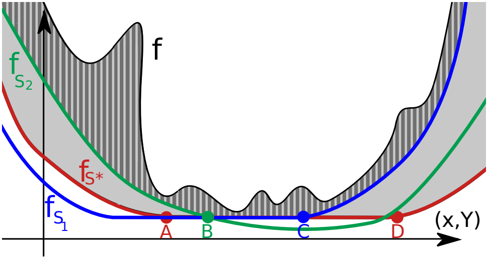


Idea Instead of a unique function f_{S^*}, use multiple functions f_{S_k}
$\Rightarrow$ we then minimize over all (x, Y)

$$f^*(x, Y) = \max_k f_{S_k}(x, Y)$$

f^* is a piecewise-quadratic convex understimator

Convex piecewise quadratic understimators



Idea Instead of a unique function f_{S^*}, use multiple functions f_{S_k}
$\Rightarrow$ we then minimize over all (x, Y)

$$f^*(x, Y) = \max_k f_{S_k}(x, Y)$$

f^* is a piecewise-quadratic convex understimator

At each sub-node :
$\rightarrow$ generate each f_{S_k} one by one as in a cutting-planes approach

A family of convex piecewise quadratic relaxations

$$(P)\left\{\ \min_{x\in[\ell,u]} \quad f(x) \equiv \langle Q, xx^\top\rangle + c^\top x\right.$$

A family of convex piecewise quadratic relaxations

$$(P)\begin{cases} \min\limits_{x\in[\ell,u]} & t \\ & t \geq \langle Q, xx^\top \rangle + c^\top x \end{cases}$$

A family of convex piecewise quadratic relaxations

$$(P)\begin{cases} \min_{x\in[\ell,u]} & t \\ & t \geq \langle Q, xx^\top\rangle + c^\top x \end{cases}$$

Let $\mathcal{K} = \{S_k \succeq 0\,, k = 1, \dots, p\}$, and a set of convex quadratic functions :

$$\langle S_k, xx^\top\rangle + c^\top x + \langle Q - S_k, Y\rangle \quad S_k \in \mathcal{K}$$

A family of convex piecewise quadratic relaxations

$$(P) \begin{cases} \min\limits_{x \in [\ell, u]} & t \\ & t \geq \langle Q, xx^\top \rangle + c^\top x \end{cases} \Leftrightarrow \begin{cases} \min\limits_{x \in [\ell, u]} & t \\ & t \geq \langle S_k, xx^\top \rangle + c^\top x + \langle Q - S_k, Y \rangle \; S_k \in \mathcal{K} \\ & \boxed{Y = xx^\top} \quad \longleftarrow \text{non-convex} \end{cases}$$

Let $\mathcal{K} = \{S_k \succeq 0 \, , k = 1, \ldots, p\}$, and a set of convex quadratic functions :

$$\langle S_k, xx^\top \rangle + c^\top x + \langle Q - S_k, Y \rangle \quad S_k \in \mathcal{K}$$

A family of convex piecewise quadratic relaxations

$$(P)\begin{cases}\min\limits_{x\in[\ell,u]} & t\\ t \geq \langle Q, xx^\top\rangle + c^\top x\end{cases} \xrightarrow{\text{relax}} (P_{\mathcal{K}})\begin{cases}\min\limits_{x\in[\ell,u]} & t\\ t \geq \langle S_k, xx^\top\rangle + c^\top x + \langle Q - S_k, Y\rangle \; S_k \in \mathcal{K}\\ (x,Y)\in \mathcal{MC}\end{cases}$$

Let $\mathcal{K} = \{S_k \succeq 0\,, k = 1, \ldots, p\}$, and a set of convex quadratic functions :

$$\langle S_k, xx^\top\rangle + c^\top x + \langle Q - S_k, Y\rangle \quad S_k \in \mathcal{K}$$

$$Y_{ij} = x_i x_j \xrightarrow{\text{relax}} \mathcal{MC}\begin{cases} Y_{ij} - u_j x_i - \ell_i x_j + u_j \ell_i \leq 0\\ Y_{ij} - u_i x_j - \ell_j x_i + u_i \ell_j \leq 0\\ -Y_{ij} + u_j x_i + u_i x_j - u_i u_j \leq 0\\ -Y_{ij} + \ell_j x_i + \ell_i x_j - \ell_i \ell_j \leq 0\end{cases}$$

McCormick envelopes [McCormick 76]

$\Rightarrow$ $(P_{\mathcal{K}})$: family of relaxations of (P) parameterized by the set $\mathcal{K}$

Which set $\mathcal{K}$ to choose ?

$$(P_{\mathcal{K}}) \begin{cases} \min\limits_{x \in [\ell, u]} \quad t \\ t \geq \langle S_k, xx^\top \rangle + c^\top x + \langle Q - S_k, Y \rangle \quad S_k \in \mathcal{K} \\ (x, Y) \in \mathcal{MC} \end{cases}$$

Goal Calculate the set of matrices S_k that maximizes the value of $(P_{\mathcal{K}})$.

Given an integer $p = |\mathcal{K}|$, we aim to solve :

$$(LB_p) = \max_{S_k \succeq 0} \; v(P_{\mathcal{K}})$$

Which set $\mathcal{K}$ to choose ?

$$(P_{\mathcal{K}}) \begin{cases} \min\limits_{x\in[\ell,u]} \quad t \\ t \geq \langle S_k, xx^\top \rangle + c^\top x + \langle Q - S_k, Y \rangle \quad S_k \in \mathcal{K} \\ (x, Y) \in \mathcal{MC} \end{cases}$$

Goal Calculate the set of matrices S_k that maximizes the value of $(P_{\mathcal{K}})$.

Given an integer $p = |\mathcal{K}|$, we aim to solve :

$$(LB_p) = \max_{S_k \succeq 0} \; v(P_{\mathcal{K}})$$

Remark If $p = 1$, we have $v(LB_1) = v(SDP)$ [MIQCR 19] .

$$(SDP) : \left\{ \min_{x\in\mathbb{R}^n,\, Y\in\mathcal{S}_n} \langle Q, Y \rangle + c^T x \; : \; (x, Y) \in \mathcal{MC},\; Y - xx^\top \succeq 0 \right\}$$

Which set $\mathcal{K}$ to choose ?

$$(P_{\mathcal{K}}) \begin{cases} \min\limits_{x \in [\ell, u]} \quad t \\ t \geq \langle S_k, xx^\top \rangle + c^\top x + \langle Q - S_k, Y \rangle \quad S_k \in \mathcal{K} \\ (x, Y) \in \mathcal{MC} \end{cases}$$

Goal Calculate the set of matrices S_k that maximizes the value of $(P_{\mathcal{K}})$.

Given an integer $p = |\mathcal{K}|$, we aim to solve :

$$(LB_p) = \max_{S_k \succeq 0} \; v(P_{\mathcal{K}})$$

Proposition $v(LB_\infty) = v(LB_p) = \ldots = v(LB_1) = v(SDP)$.

$$(SDP) : \left\{ \min_{x \in \mathbb{R}^n,\, Y \in \mathcal{S}_n} \langle Q, Y \rangle + c^T x \; : \; (x, Y) \in \mathcal{MC},\; Y - xx^\top \succeq 0 \right\}$$

Which set $\mathcal{K}$ to choose ?

$$(P_{\mathcal{K}}) \begin{cases} \min\limits_{x \in [\ell, u]} \quad t \\ t \geq \langle S_k, xx^\top \rangle + c^\top x + \langle Q - S_k, Y \rangle \quad S_k \in \mathcal{K} \\ (x, Y) \in \mathcal{MC} \end{cases}$$

Goal Calculate the set of matrices S_k that maximizes the value of $(P_{\mathcal{K}})$.

Given an integer $p = |\mathcal{K}|$, we aim to solve :

$$(LB_p) = \max_{S_k \succeq 0} v(P_{\mathcal{K}})$$

Proposition $v(LB_\infty) = v(LB_p) = \ldots = v(LB_1) = v(SDP)$.

$$(SDP) : \left\{ \min_{x \in \mathbb{R}^n,\, Y \in \mathcal{S}_n} \langle Q, Y \rangle + c^T x \,:\, (x, Y) \in \mathcal{MC},\ Y - xx^\top \succeq 0 \right\}$$

Goal : Add quadratic cuts in course of the B&B to refine the bound

The cutting-quadrics algorithm

Basic ideas of the iterative cutting-quadrics algorithm

Initialization : A set $\mathcal{K}_0 = \mathcal{K}_0 \cup \mathbf{0}_n$ of SDP matrices

1. Solve $(P_{\mathcal{K}_0})$: get $v(P_{\mathcal{K}_0})$ and (x^0, Y^0) the optimal solution.
2. **While** the matrix $(Y^i - x^i x^{i\top})$ is not feasible to (SDP) :
   i. Solve$(P_{\mathcal{K}_i})$: get $v(P_{\mathcal{K}_i})$ and (x^i, Y^i) the optimal solution.
   ii. Add a new matrix $S_{i+1} \succeq 0$ to $\mathcal{K}_i$
   iii. i=i+1

Output : (x^i, Y^i) an optimal solution to (LB_∞)

Basic ideas of the iterative cutting-quadrics algorithm

Initialization : A set $\mathcal{K}_0 = \mathcal{K}_0 \cup \mathbf{0}_n$ of SDP matrices

1. Solve $(P_{\mathcal{K}_0})$: get $v(P_{\mathcal{K}_0})$ and (x^0, Y^0) the optimal solution.
2. **While** the matrix $(Y^i - x^i x^{i\top})$ is not feasible to (SDP) :
   i. Solve$(P_{\mathcal{K}_i})$: get $v(P_{\mathcal{K}_i})$ and (x^i, Y^i) the optimal solution.
   ii. Add a new matrix $S_{i+1} \succeq 0$ to $\mathcal{K}_i$
   iii. i=i+1

Output : (x^i, Y^i) an optimal solution to (LB_∞)

We want to choose S_{i+1} such that :

1. The point (x^i, Y^i) is cut by the new function $f_{S_{i+1}}$
2. The algorithm converges to the optimal value of (SDP)

Which S_{i+1} to choose ?

1. The point (x^i, Y^i) is cut by the new function $f_{S_{i+1}}$

Which S_{i+1} to choose?

1. The point (x^i, Y^i) is cut by the new function $f_{S_{i+1}}$

let v the eigenvector of $(x^i x^{i\top} - Y^i)$ associated to its largest eigenvalue λ,

$$\lambda = \langle vv^\top, x^i x^{i\top} - Y^i \rangle = \max_{u \in \mathbb{R}^n ||u||=1} \langle uu^\top, x^i x^{i\top} - Y^i \rangle$$

Which S_{i+1} to choose?

1. The point (x^i, Y^i) is cut by the new function $f_{S_{i+1}}$

let v the eigenvector of $(x^i x^{i\top} - Y^i)$ associated to its largest eigenvalue λ,

$$\lambda = \langle vv^\top, x^i x^{i\top} - Y^i \rangle = \max_{u \in \mathbb{R}^n ||u||=1} \langle \underbrace{uu^\top}_{S = uu^\top \succeq 0}, x^i x^{i\top} - Y^i \rangle$$

$$\langle vv^\top, x^i x^{i\top} - Y^i \rangle \geq \langle S, x^i x^{i\top} - Y^i \rangle \quad \forall S \succeq 0, \ \mathrm{Tr}(S) = 1$$

Which S_{i+1} to choose ?

1. The point (x^i, Y^i) is cut by the new function $f_{S_{i+1}}$

let v the eigenvector of $(x^i x^{i\top} - Y^i)$ associated to its largest eigenvalue λ,

$$\lambda = \langle vv^\top, x^i x^{i\top} - Y^i \rangle = \max_{u \in \mathbb{R}^n ||u||=1} \langle \underbrace{uu^\top}_{S_i = uu^\top \succeq 0}, x^i x^{i\top} - Y^i \rangle$$

$$\langle vv^\top, x^i x^{i\top} - Y^i \rangle \geq \langle S_i, x^i x^{i\top} - Y^i \rangle$$

Which S_{i+1} to choose?

1. The point (x^i, Y^i) is cut by the new function $f_{S_{i+1}}$

let v the eigenvector of $(x^i x^{i\top} - Y^i)$ associated to its largest eigenvalue λ,

$$\lambda = \langle vv^\top, x^i x^{i\top} - Y^i \rangle = \max_{u \in \mathbb{R}^n ||u||=1} \langle \underbrace{uu^\top}_{S_i = uu^\top \succeq 0}, x^i x^{i\top} - Y^i \rangle$$

$$\langle \underbrace{vv^\top}_{S_{i+1}}, x^i x^{i\top} - Y^i \rangle \geq \langle S_i, x^i x^{i\top} - Y^i \rangle$$

Which S_{i+1} to choose?

1. The point (x^i, Y^i) is cut by the new function $f_{S_{i+1}}$

let v the eigenvector of $(x^i x^{i\top} - Y^i)$ associated to its largest eigenvalue λ,

$$\lambda = \langle vv^\top, x^i x^{i\top} - Y^i \rangle = \max_{u \in \mathbb{R}^n ||u||=1} \langle \underbrace{uu^\top}_{S_i = uu^\top \succeq 0}, x^i x^{i\top} - Y^i \rangle$$

$$\langle \underbrace{vv^\top}_{S_{i+1}}, x^i x^{i\top} - Y^i \rangle \geq \langle S_i, x^i x^{i\top} - Y^i \rangle$$

$$\underbrace{\langle S_{i+1}, x^i x^{i\top} - Y^i \rangle + c^\top x^i + \langle Q, Y^i \rangle}_{f_{S_{i+1}}(x^i, Y^i)} \geq \underbrace{\langle S_i, x^i x^{i\top} - Y^i \rangle + c^\top x^i + \langle Q, Y^i \rangle}_{f_{S_i}(x^i, Y^i)}$$

Which S_{i+1} to choose ?

1. The point (x^i, Y^i) is cut by the new function $f_{S_{i+1}}$

Remark :

- Let $r \geq 1$, in practice we can choose

$$S_{i+1} = r \sum_{j=1}^{s} v_j {v_j}^\top$$

with v_j the eigenvectors associated to positive eigenvalues $(x^i x^{i\top} - Y^i)$

Which S_{i+1} to choose ?

1. The point (x^i, Y^i) is cut by the new function $f_{S_{i+1}}$

- The additional quadric penalizes (very heavily) the optimal solution from the previous iteration (x^i, Y^i).
- For a given $r > 0$, we take $S_{i+1} = r \cdot v_{max} v_{max}^\top$, where v_{max} is the eigenvector of matrix $(x^i x^{i\top} - Y^i)$ of maximum eigenvalue.
- This choice ensures that solution (x^i, Y^i) will become sub-optimal in the new program $(P_{\mathcal{K}_{i+1}})$.

The choice of matrix S_{k+1} is not unique, and in practice, we can also define $S_{k+1} = r^k \sum v_i v_i^T$, where the sum is carried out over all eigenvectors v_i having a positive eigenvalue. This will enable the new quadric to penalize larger areas of the (x, Y) space, i.e.,, more solutions (x, Y) such that $\langle S_{k+1}, xx^\top - Y \rangle > 0$.

Which S_{i+1} to choose ?

2. The algorithm converges to the optimal value of (SDP)

$$(SDP) : \left\{ \min_{x \in \mathbb{R}^n,\, Y \in \mathcal{S}_n} \langle Q, Y \rangle + c^T x \; : \; (x, Y) \in \mathcal{MC},\; Y - xx^\top \succeq 0 \right\}$$

Sketch of proof : $v(P_{\mathcal{K}_i}) \geq v(SDP)$

Which S_{i+1} to choose ?

2. The algorithm converges to the optimal value of (SDP)

$$(SDP) : \left\{ \min_{x \in \mathbb{R}^n,\, Y \in \mathcal{S}_n} \langle Q, Y \rangle + c^T x \; : \; (x, Y) \in \mathcal{MC},\; Y - xx^\top \succeq 0 \right\}$$

Sketch of proof : $v(P_{\mathcal{K}_i}) \geq v(SDP)$

- At last iteration $(x^i x^{i\top} - Y^i) \preceq 0$ is feasible for (SDP) of value

$$\langle Q, Y^i \rangle + c^\top x^i \geq v(SDP)$$

Which S_{i+1} to choose ?

2. The algorithm converges to the optimal value of (SDP)

$$(SDP) : \left\{ \min_{x \in \mathbb{R}^n,\, Y \in \mathcal{S}_n} \langle Q, Y \rangle + c^T x \;:\; (x, Y) \in \mathcal{MC},\; Y - xx^\top \succeq 0 \right\}$$

Sketch of proof : $v(P_{\mathcal{K}_i}) \geq v(SDP)$

- At last iteration $(x^i x^{i\top} - Y^i) \preceq 0$ is feasible for (SDP) of value

$$\langle Q, Y^i \rangle + c^\top x^i \geq v(SDP)$$

$$v(P_{\mathcal{K}_i}) = \max_{S_{k \in \mathcal{K}_i}} \underbrace{\langle \underbrace{S_i}_{\succeq 0}, \underbrace{x^i x^{i\top} - Y^i}_{\preceq 0} \rangle}_{=0 \text{ since } \mathbf{0_n} \in \mathcal{K}_0} + \underbrace{c^\top x^i + \langle Q, Y^i \rangle}_{\text{val. } (SDP) \text{ at } (x^i, Y^i)} \geq v(SDP)$$

The Cutting-Quadradric Algorithm

Initialization : A set $\mathcal{K}_0 = \mathcal{K}_0 \cup \mathbf{0}_n$ of SDP matrices

1. Solve $(P_{\mathcal{K}_0})$: get $v(P_{\mathcal{K}_0})$ and (x^0, Y^0) the optimal solution.
2. **While** the matrix $(Y^i - x^i x^{i\top})$ is not feasible to (SDP) :
   i. Solve$(P_{\mathcal{K}_i})$: get $v(P_{\mathcal{K}_i})$ and (x^i, Y^i) the optimal solution.
   ii. Add a new matrix $S_{i+1} \succeq 0$ to $\mathcal{K}_i$
   iii. i=i+1

Output : (x^i, Y^i) an optimal solution to (LB_∞)

Theorem When $k \to \infty$, $v(P_{\mathcal{K}_i})$ converges to the $v(SDP)$.

Corollary The algorithm solves problem (SDP).

Illustration : Cutting-Quadrics Algorithm

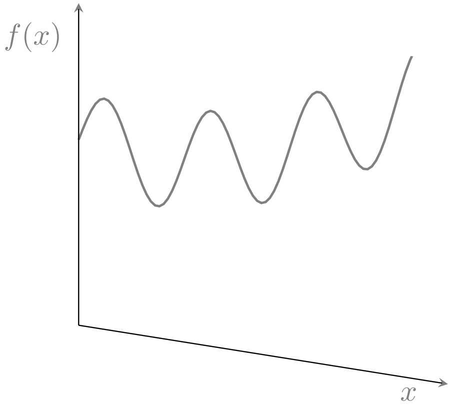


Illustration : Cutting-Quadrics Algorithm

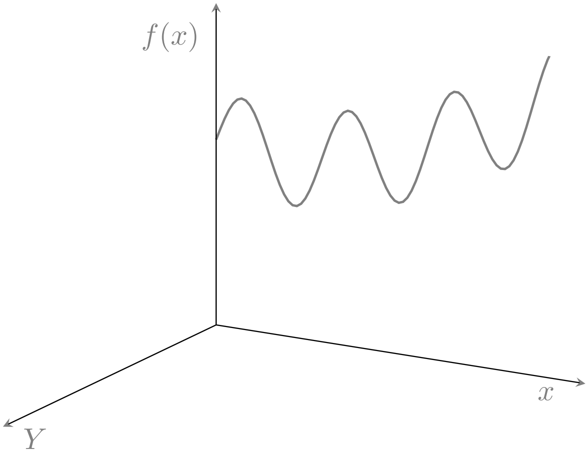


Illustration : Cutting-Quadrics Algorithm

$\mathcal{K} = \{S_1\}$

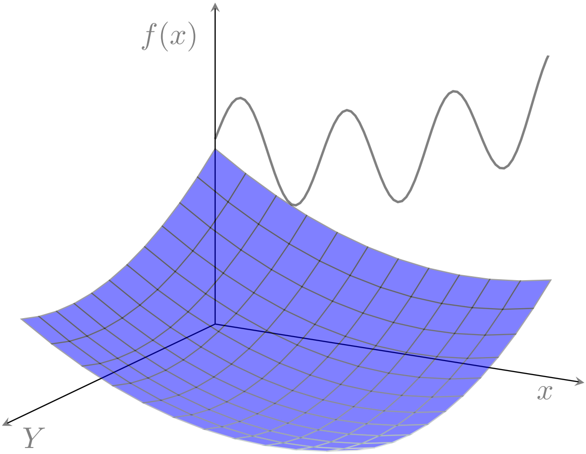


Illustration : Cutting-Quadrics Algorithm

$\mathcal{K} = \{S_1\}$

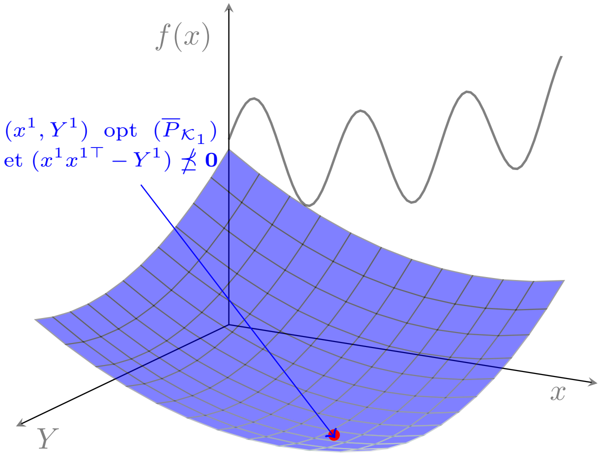


Illustration : Cutting-Quadrics Algorithm

$\mathcal{K} = \{S_1\}$

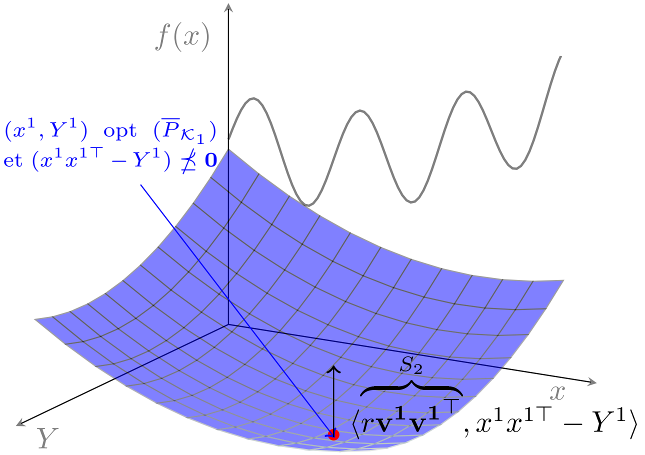


Illustration : Cutting-Quadrics Algorithm

$\mathcal{K} = \{S_1, S_2\}$

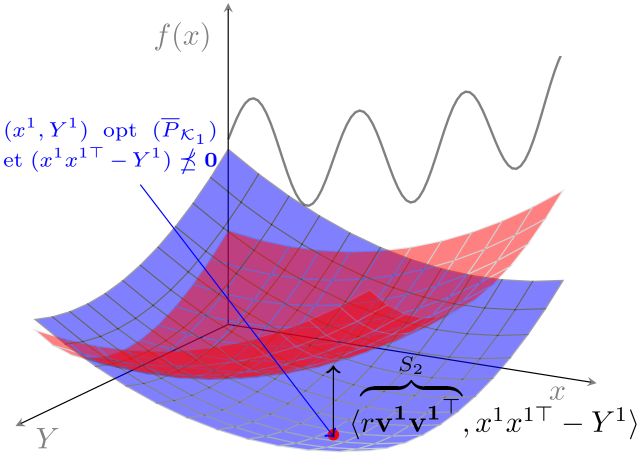


Illustration : Cutting-Quadrics Algorithm

$\mathcal{K} = \{S_1, S_2\}$

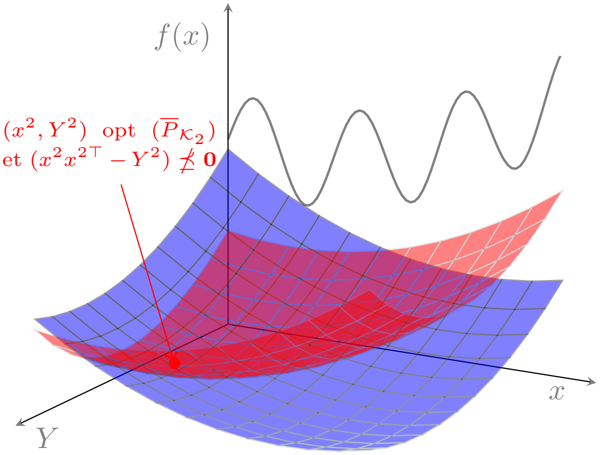


Illustration : Cutting-Quadrics Algorithm

$\mathcal{K} = \{S_1, S_2\}$

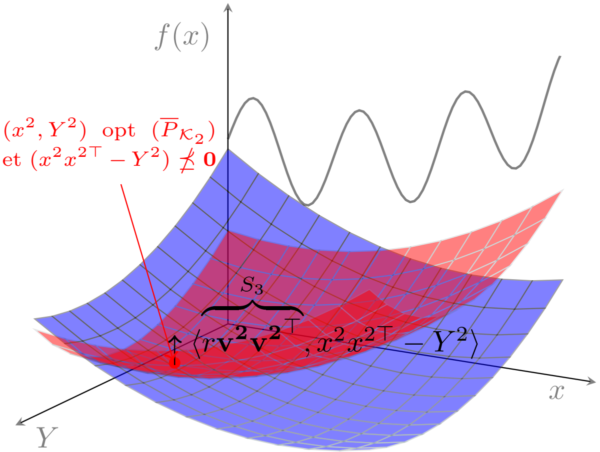


Illustration : Cutting-Quadrics Algorithm

$\mathcal{K} = \{S_1, S_2, S_3\}$

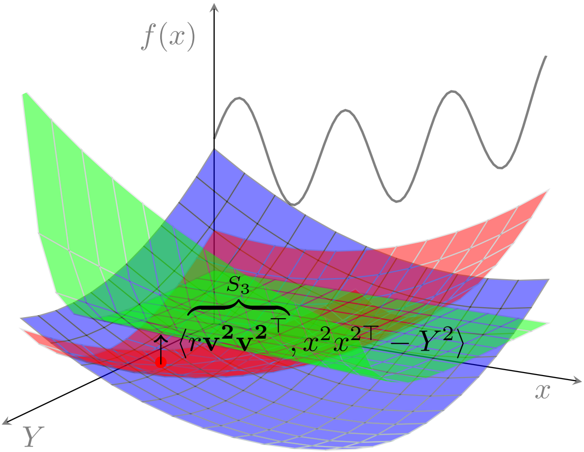
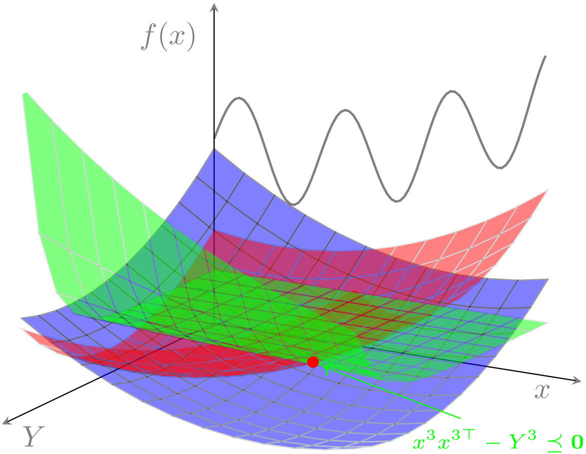


Illustration : Cutting-Quadrics Algorithm

$\mathcal{K} = \{S_1, S_2, S_3\}$



Computational results

Instances *boxqp* [Burer et al. 09]

Description of *boxqp* :

- 99 purely continuous quadratic instances with $x \in [0, 1]$
- Sizes vary from $n = 20$ to 125, densities of Q from 20%, to 100%.

Two versions of `Cutting Quadrics - B&B (CQBB)` with stopping criteria :

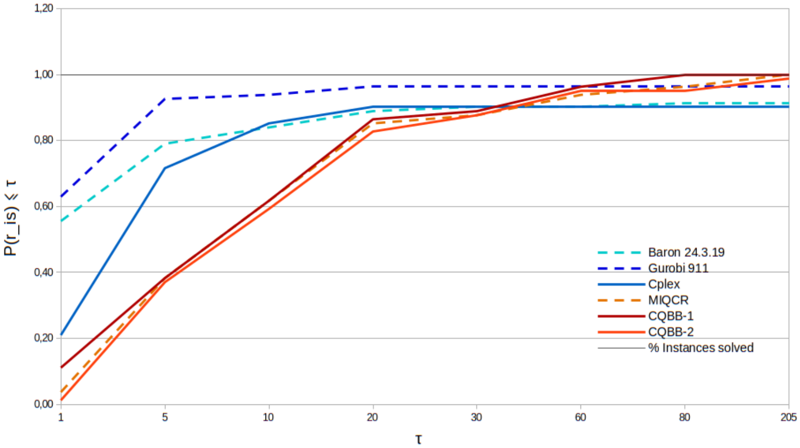
- `CQBB-1` : we stop if $\lambda_{\max}(x^k {x^k}^\top - Y^k) > 0.8$.
- `CQBB-2` : we force at maximum one iteration per node.

For both we take $\mathcal{K}_0 = 0_n \cup S^*$ (i.e. best matrix for (LB_1))

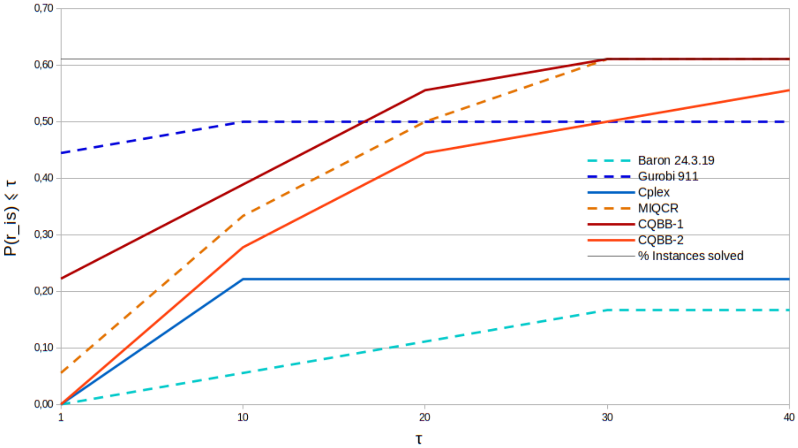
Comparison methods :

- `MIQCR`, `Cplex 12.9`, `Baron 21.1.13`, `Gurobi 9.1.1`.

Performance profile : CPU time - time limit 1 hour, $n \leq 100$



Performance profile : CPU time - time limit 1 hour, $n \geq 100$



Total number of nodes visited during the B&B

Instance	MIQCR	CQBB-1	CQBB-2
100-025-1	665	505	429
100-025-2	315	261	191
100-025-3	269	191	199
100-050-3	811	755	663
100-075-1	12001	11929	11367
100-075-2	9479	9517	9077
100-075-3	8867	8829	8521

Conclusion and perspectives

Conclusions et perspectives

Conclusions

- A cutting (convex) quadric algorithm.
- Allows to recalculate the SDP bound at each node of the B&B.

  → Experimentally : more efficient than `MIQCR`

Future work

- How to determine a good initial set of matrices ?
- Aggregating matrices to limit the number of quadrics ?
- Handling problems with constraints (and integers)