From cutting planes to cutting hypersurfaces

for convexifying quadratic programs

Amélie Lambert - Daniel Porumbel

Conservatoire National des arts et Métiers - CEDRIC-CNAM, Paris

le cnam cedric EA4629

We want the global solution of this non-convex program

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

We want the global solution of this non-convex program

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

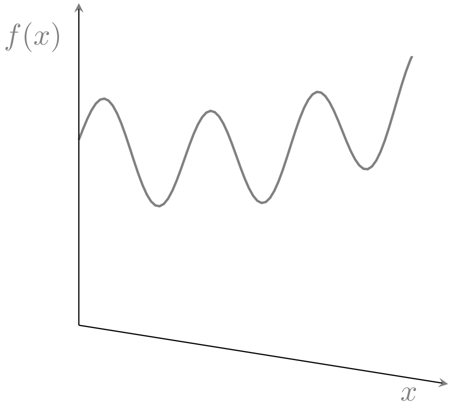
A cutting convex quadrics approach :

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

Starting point : Convexification of a quadratic function

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

Add convex function $\langle S, xx^\top \rangle$ to be canceled by linear term $\langle -S, Y \rangle$



Starting point : Convexification of a quadratic function

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

Add convex function $\langle S, xx^\top \rangle$ to be canceled by linear term $\langle -S, Y \rangle$

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

$\rightarrow$ For any matrix $S \succeq 0$, function below is convex in x and linear in Y :

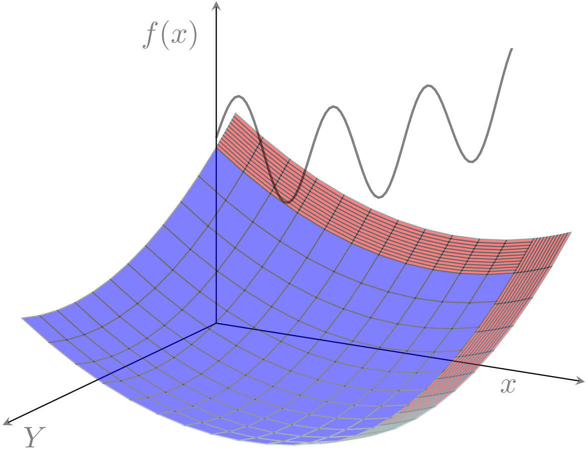
$$\begin{aligned} f_S(x, Y) =& \langle Q, Y \rangle + c^\top x + \langle S, xx^\top \rangle - \langle S, Y \rangle \\ f_S(x, Y) =& \langle Q, xx^\top \rangle + c^\top x = f(x) \quad \text{if } Y = xx^\top \end{aligned}$$

[MIQCR - Elloumi-Lambert (2019)]

Starting point : Convexification of a quadratic function

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

Add convex function $\langle S, xx^\top \rangle$ to be canceled by linear term $\langle -S, Y \rangle$

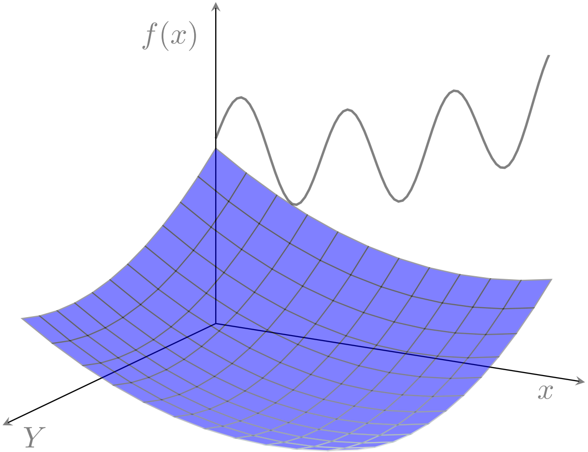


We'll add McCormick inequalities to cut some Y corresponding to no x

Starting point : Convexification of a quadratic function

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

Add convex function $\langle S, xx^\top \rangle$ to be canceled by linear term $\langle -S, Y \rangle$



After these McCormick cuts, the set of feasible Y is smaller

Starting point : Convexification of a quadratic function

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

Add convex function $\langle S, xx^\top \rangle$ to be canceled by linear term $\langle -S, Y \rangle$

- If $S = \mathbf{0}_n$ our convex function

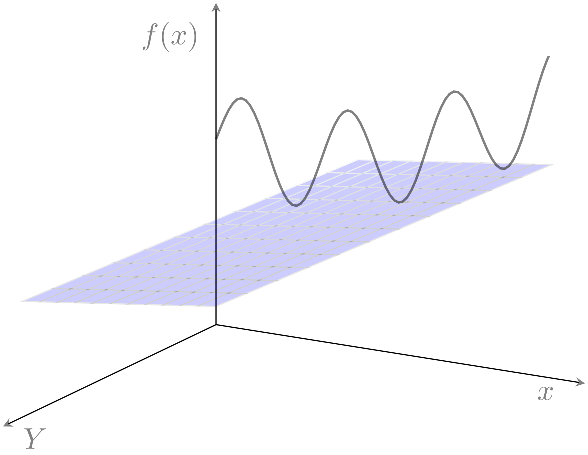
$$\langle Q, Y \rangle + c^\top x + \langle S, xx^\top \rangle - \langle S, Y \rangle$$

becomes $f_0(x, Y) = \langle Q, Y \rangle + c^\top x$, i.e., the surface a is a linear hyperplane.

Starting point : Convexification of a quadratic function

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

Add convex function $\langle S, xx^\top \rangle$ to be canceled by linear term $\langle -S, Y \rangle$



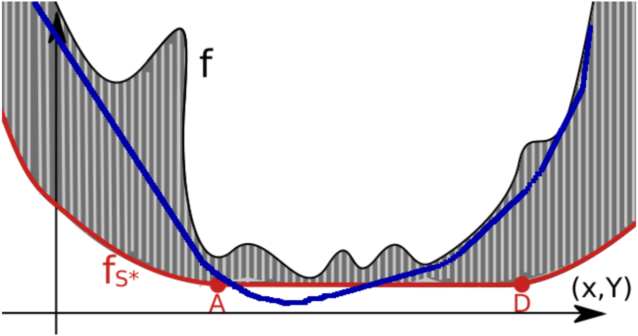
Starting point : Convexification of a quadratic function

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

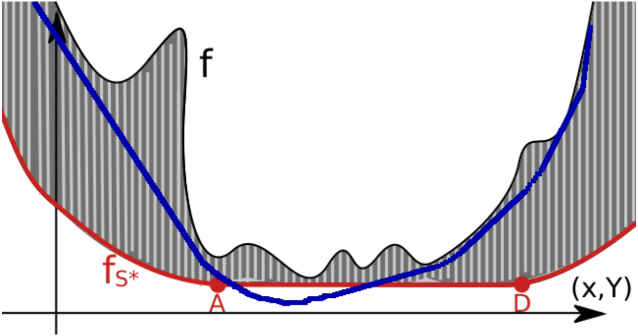
Add convex function $\langle S, xx^\top \rangle$ to be canceled by linear term $\langle -S, Y \rangle$

$$\langle Q, Y \rangle + c^\top x + \langle S, xx^\top \rangle - \langle S, Y \rangle$$

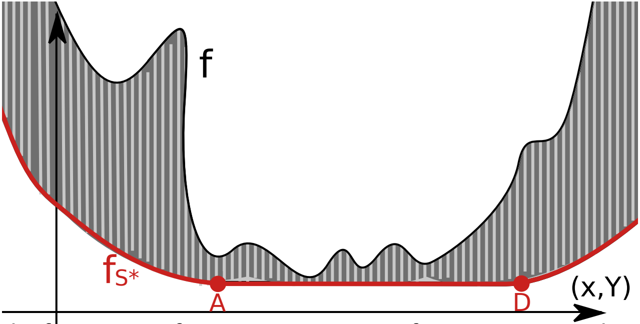
- Difficult to find the best S^*, we construct a special SDP program only for that.



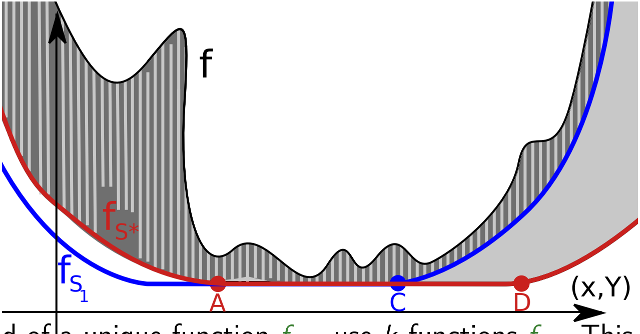
- Traditional convexification : the "best" S^* is the one that leads to a convex relaxation of the highest optimum value
- According to this criterion, the red convexification is better than the blue one. Yet the blue one is tighter if we think over the whole area, which may be useful when we start branching.



- Traditional convexification : the "best" S^* is the one that leads to a convex relaxation of the highest optimum value
- According to this criterion, the red convexification is better than the blue one. Yet the blue one is tighter if we think over the whole area, which may be useful when we start branching. In reality, we can't have such a tight convexification as the blue one with a unique S^*. That's an ideal dream to forget $\implies$ we need multiple $S_1, S_2, S_3, \ldots$.

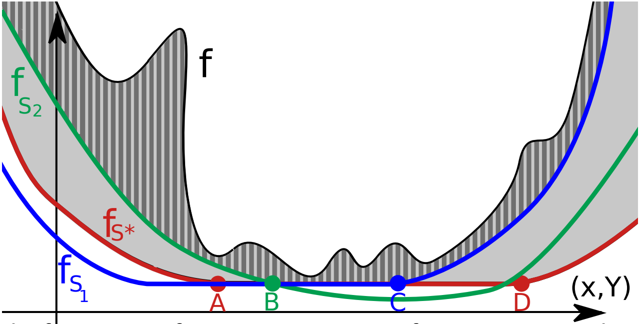


Idea Instead of a unique function f_{S^*}, use k functions f_{S_k}. This is what happens with $k = 1$



Idea Instead of a unique function f_{S^*}, use k functions f_{S_k}. This is what happens with $k = 1$

Idea This is what happens for $k = 2$ when taking the best of two functions, it's a bit better



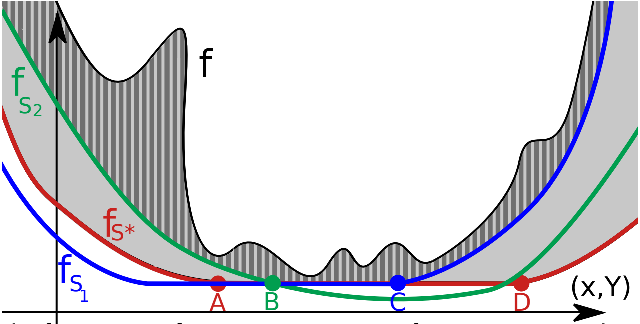
Idea Instead of a unique function f_{S^*}, use k functions f_{S_k}. This is what happens with $k = 1$

Idea This is what happens for $k = 2$ when taking the best of two functions, it's a bit better

$\Rightarrow$ This is what happens for $k = 3$. Generally,

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

f^* is a piecewise-quadratic convex understimator



Idea Instead of a unique function f_{S^*}, use k functions f_{S_k}. This is what happens with $k = 1$

Idea This is what happens for $k = 2$ when taking the best of two functions, it's a bit better

$\Rightarrow$ This is what happens for $k = 3$. Generally,

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

f^* is a piecewise-quadratic convex understimator

At each branching node : add more surfaces f_{S_k} as in a `Cutting-planes`

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

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

A family of convex piecewise quadratic relaxations

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

Let $\mathcal{K} = \{S_k \succeq 0\,, k = 1,\ldots,\overline{k}\}$ and 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 $\mathcal{K} = \{S_k \succeq 0\,, k = 1, \ldots, \overline{k}\}$ and 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 $\mathcal{K}=\{S_k\succeq 0\,,k=1,\dots,\overline{k}\}$ and convex quadratic functions :

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

$$Y_{ij}=x_ix_j\xrightarrow{\text{relax}}\mathcal{MC}\begin{cases}Y_{ij}-u_jx_i-\ell_ix_j+u_j\ell_i\leq 0\\ Y_{ij}-u_ix_j-\ell_jx_i+u_i\ell_j\leq 0\\ -Y_{ij}+u_jx_i+u_ix_j-u_iu_j\leq 0\\ -Y_{ij}+\ell_jx_i+\ell_ix_j-\ell_i\ell_j\leq 0\end{cases}$$

McCormick envelopes specific to each node

A family of convex piecewise quadratic relaxations

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

We would like to cut more Y values that are too far from an outer product form $xx^\top$. Since Y is meant to satisfy $Y=xx^\top$, we may like to add $Y\succeq xx^\top$, but it's too expensive because above program is quadratic in x : applying a branch-and-bound on such program would be too slow.

Illustration : Cutting-Quadrics Algorithm

We extend with Y meant to satisfy $Y = xx^\top$

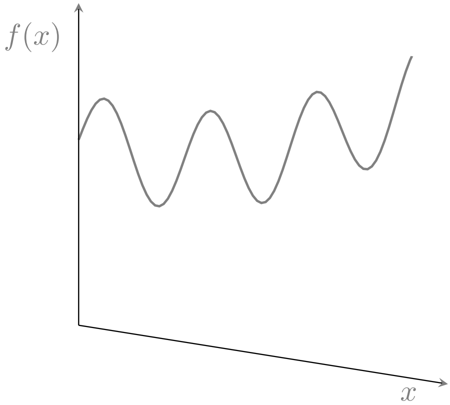


Illustration : Cutting-Quadrics Algorithm

We extend with Y meant to satisfy $Y = xx^{\top}$

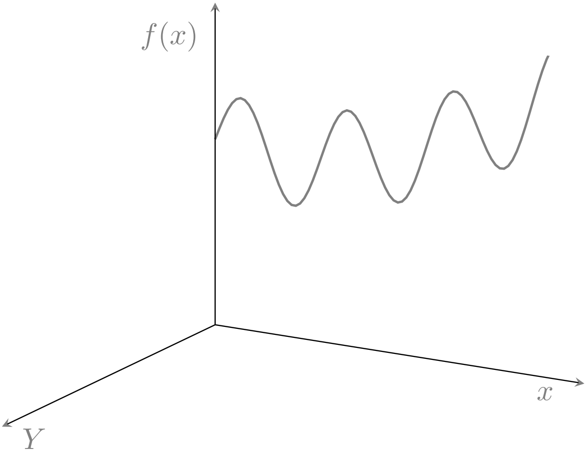


Illustration : Cutting-Quadrics Algorithm

We cut some unwanted Y using linear McCormick

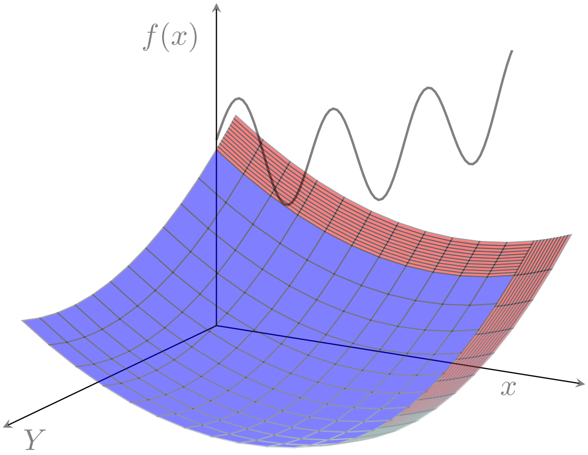


Illustration : Cutting-Quadrics Algorithm

We want $Y \succeq xx^\top$ which implies $xx^\top - Y \preceq 0$, but how to add it to the quadratic program ?

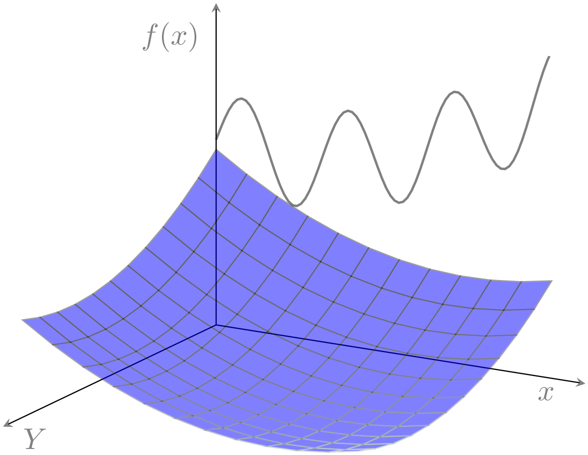


Illustration : Cutting-Quadrics Algorithm

We want $Y \succeq xx^\top$ which implies $xx^\top - Y \preceq 0$, but how to add it to the quadratic program ?

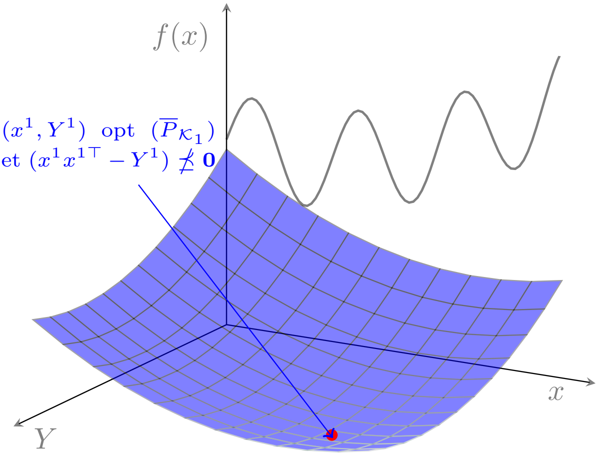


Illustration : Cutting-Quadrics Algorithm

We want $Y \succeq xx^\top$ which implies $xx^\top - Y \preceq 0$, but how to add it to the quadratic program ? Construct $S_2 = v^1 {v^1}^\top$, where v^1 corresponds to the maximum eigenvalue of $x^1 {x^1}^\top - Y^1$ of first iterate (x^1, Y^1)

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

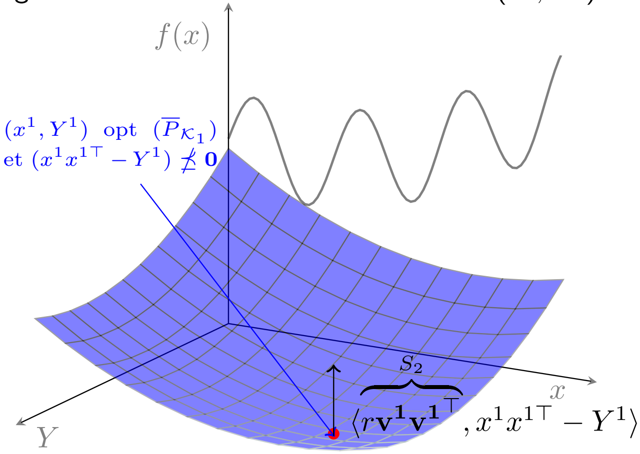


Illustration : Cutting-Quadrics Algorithm

We want $Y \succeq xx^\top$ which implies $xx^\top - Y \preceq 0$, but how to add it to the quadratic program ?

This S_2 generates a new hyper-surface !

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

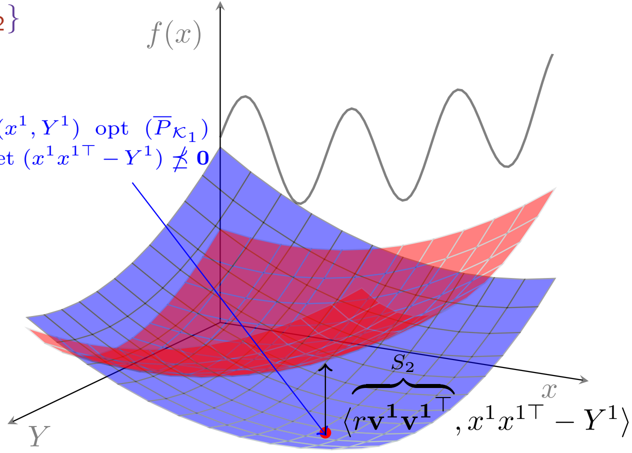


Illustration : Cutting-Quadrics Algorithm

We want $Y \succeq xx^\top$ which implies $xx^\top - Y \preceq 0$, but how to add it to the quadratic program ?

The optimum values moves to (x^2, Y^2) !

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

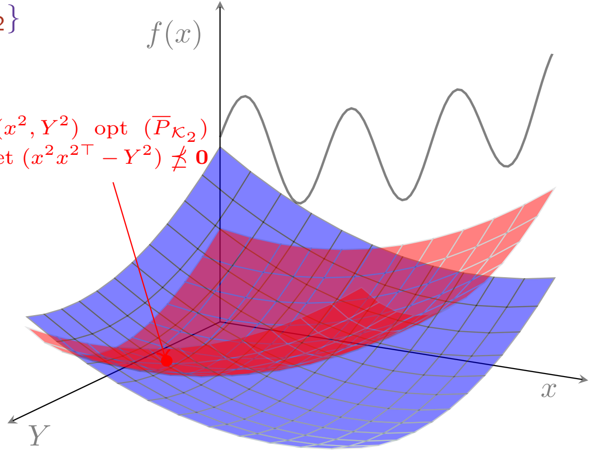


Illustration : Cutting-Quadrics Algorithm

We want $Y \succeq xx^\top$ which implies $xx^\top - Y \preceq 0$, but how to add it to the quadratic program ? Construct $S_3 = v^2 {v^2}^\top$, where v^2 corresponds to the maximum eigenvalue of $x^2 {x^2}^\top - Y^2$

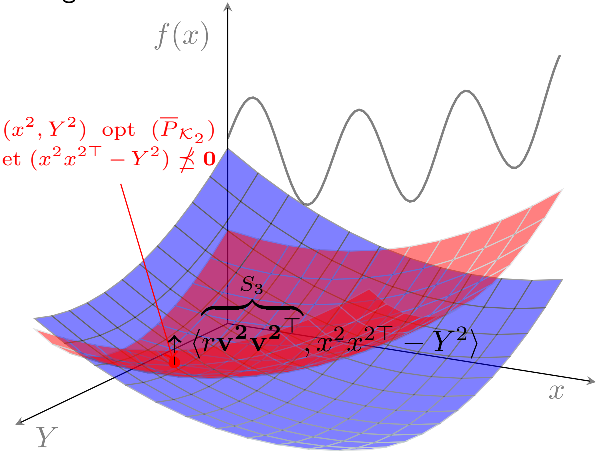


Illustration : Cutting-Quadrics Algorithm

We want $Y \succeq xx^\top$ which implies $xx^\top - Y \preceq 0$, but how to add it to the quadratic program ?

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

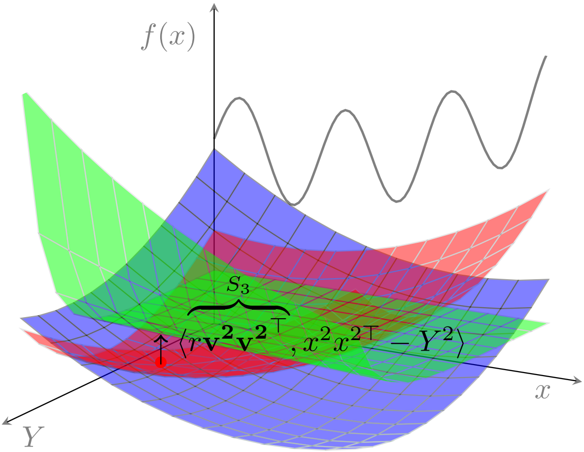
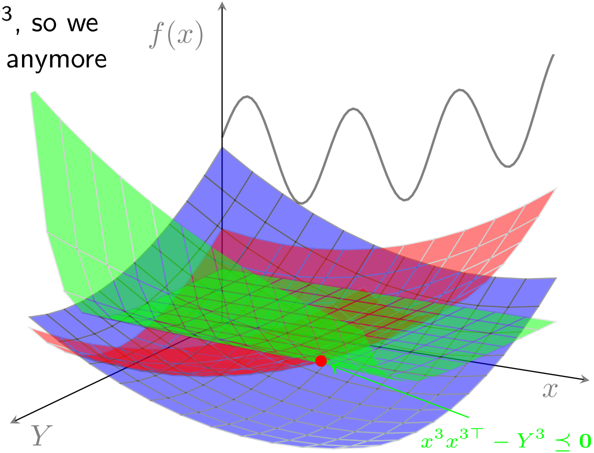


Illustration : Cutting-Quadrics Algorithm

We want $Y \succeq xx^\top$ which implies $xx^\top - Y \preceq 0$, but how to add it to the quadratic program ?

$\mathcal{K} = \{S_1, S_2, S_3\}$ The optimum moves to (x^3, Y^3) that does satisfy $x^3 {x^3}^\top \succeq Y^3$, so we can't cut it anymore



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` : fewer iterations per node
- `CQBB-2` : more iterations per node

instance	Initial method (0 iterations)			CQBB-1 (few iterations)			CQBB-2 (more iterations)		
	Nodes	*CPU*	*#calls*	*Nodes*	*CPU*	*#calls*	*Nodes*	*CPU*	*#calls*
050-030-1	13	0.8	13	9	0.5	11	7	0.8	12
050-030-2	19	1.1	19	17	0.9	19	15	1.4	24
050-030-3	31	1.6	31	27	1.4	35	13	1.2	20
050-040-1	9	0.7	9	9	0.6	12	5	0.6	8
050-040-2	43	2.5	43	31	1.3	31	25	2.1	40
050-040-3	7	0.5	7	7	0.4	7	5	0.6	8
050-050-1	5207	381	5207	4549	218	5088	4447	455	8286
050-050-2	65	3.7	65	67	3.2	77	53	4.5	88
050-050-3	155	8	155	121	7	180	63	5.8	114
060-020-1	13	1.2	13	5	0.5	5	5	0.8	7
060-020-2	11	1.0	11	5	0.5	5	5	0.8	7
060-020-3	61	4.7	61	33	2.1	33	33	4.2	54
070-025-1	77	8	77	41	3.7	41	33	6.5	57
070-025-2	199	22	199	129	11	133	109	21	209
070-025-3	225	25	225	145	15	168	105	17	174
070-050-1	183	21	183	159	17	184	145	26	255
070-050-2	29	3.6	29	25	3.0	31	23	4.7	40
070-050-3	9	1.4	9	11	1.6	15	7	1.7	11
070-075-1	65	7	65	59	6.1	63	55	11	108
070-075-2	1735	203	1735	1435	161	1750	1401	263	2592
070-075-3	819	96	819	729	86	990	585	113	1148
080-025-1	391	61	391	227	41	412	23	8	53

#calls is the number of calls of the convex quadratic solver

Conclusions et prospects

- We designed quadratic-convex variant of `Cutting-Planes`
- We uses it to update the convexification at branching node

  $\rightarrow$ Experimentally : more efficient than performing no iteration
- It is very expensive to re-optimize after adding a cutting hypersurface (unlike a dual simplex Pivot when adding a hyperplane in LP). Any improvement in such re-optimization tasks can speed-up it all
- just accepted by `Journal of Global Optimization` ; pdf here : `cedric.cnam.fr/~porumbed/pconvex.pdf`