Semidefinite Optimization with Eigenvector Branching

Kurt M. Anstreicher

Department of Business Analytics
University of Iowa

ICERM Workshop on MINLP
February 27 - March 3, 2023

Outline

1. Semidefinite optimization
2. Eigenvector branching
3. Computational results
4. Disjunctive cuts
5. Conclusion

Interested in semidefinite programming (SDP) problems of the form

$$\begin{aligned} \text{SDP}: \quad \min \quad & C \bullet Y \\ \text{s.t.} \quad & A_i \bullet Y = b_i, \quad i \in \mathcal{E} \\ & A_i \bullet Y \le b_i, \quad i \in \mathcal{I} \\ & Y \succeq 0, \end{aligned}$$

where matrix Y has the form

$$Y = \begin{pmatrix} Y_{00} & x^T \\ x & X \end{pmatrix}.$$

- Constraint $Y \succeq 0$, with $Y_{00} = 1$, is equivalent to $X \succeq xx^T$, a convex relaxation of the rank-one condition $X = xx^T$.
- Index rows/columns of Y as $0, 1, \dots, n$, and assume that constraint for $i = 0 \in \mathcal{E}$ is $Y_{00} = 1$.
- Assume w.l.o.g. that $b_i = 0$ for all other $i > 0$.

SDP gives convex relaxation for variety of NP-hard problems, including those with binary variables and nonconvex quadratic objective and/or constraints. If $X \neq xx^T$ in solution of SDP, can either add valid inequalities to strengthen relaxation or *branch* on a variable x_i.

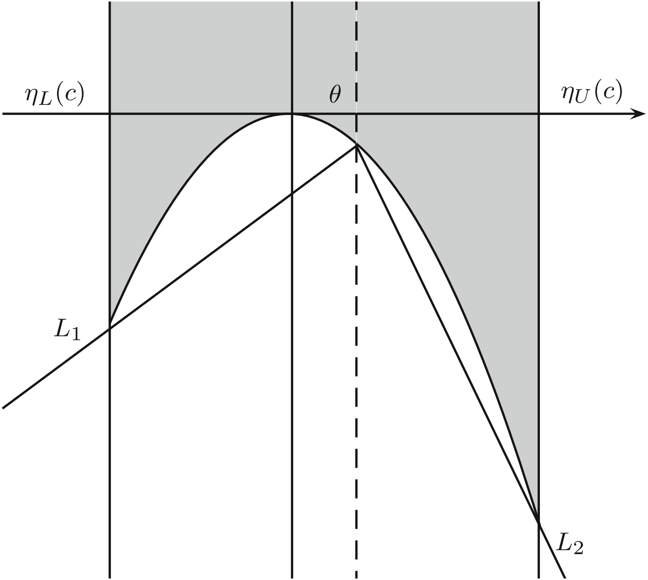
- If x_i is binary, branching creates child problems with $x_i = 0$ and $x_i = 1$.
- If x_i is continuous, typically use *spatial branching* to create child problems with $x_i \leq \theta$ and $x_i \geq \theta$, where θ is often taken to be current value of x_i. Essential to use new bound on x_i in each child problem to tighten relaxation, for example through RLT constraints.
- Spatial branching is inefficient, but nevertheless has been successfully implemented in global optimization codes such as BARON and Couenne.
- Spatial branching based on conditions $x_i \leq \theta$ and $x_i \geq \theta$ is *unrelated* to condition $X = xx^T$. Goal here is to devise spatial branching scheme that is directly connected to this condition.

Assume feasible solution to SDP has $\bar{Y} \succeq 0$, but $\bar{X} - \bar{x}\bar{x}^T \neq 0$. Then there is an $a \in \mathbb{R}^n$ with $a^T(\bar{X} - \bar{x}\bar{x}^T)a > 0$, or $a^T\bar{X}a > (a^T\bar{x})^2$.

- Constraint $a^TXa \leq (a^Tx)^2$ certainly holds for all rank-one $X = xx^T$, but is not convex in the variables (x, X).
- Can obtain a valid disjunction with convex constraints using a secant approximation of $(a^Tx)^2$, one part of which must hold for any feasible Y with $X = xx^T$.

Proposition. (Saxeena, Bonami and Lee 2010) Suppose that $\bar{Y}$ is feasible in SDP, with $a^T\bar{x} = \theta$ and $a^T\bar{X}a > \theta^2$. Assume that $\mu^- \leq a^Tx \leq \mu^+$ for all x feasible in SDP. Then if Y is feasible in *SDP* with $X = xx^T$, one of the following must hold:

$$
\begin{array}{rcll}
a^Tx & \geq & \theta, & a^TXa \leq (\theta + \mu^+)a^Tx - \theta\mu^+, \\
a^Tx & \leq & \theta, & a^TXa \leq (\theta + \mu^-)a^Tx - \theta\mu^-.
\end{array}
$$


$\eta_L(c)$
θ
$\eta_U(c)$
L_1
L_2

In SBL (2010), disjunction is used to construct a *cut* that can be added to strengthen SDP. Here we will use the same disjunction as the basis for *spatial branching*.

- Note that $(\theta+\mu^+)a^T\bar{x}-\theta\mu^+ = (\theta+\mu^-)a^T\bar{x}-\theta\mu^- = \theta a^T\bar{x} = \theta^2$, so current values $(\bar{x}, \bar{X})$ are infeasible for both parts of disjunction.
- Added branching constraint $a^Tx \le \theta$ or $a^Tx \ge \theta$ can be combined with other constraints to further tighten child problems.
- Good candidate for a, suggested in SBL (2010), is eigenvector corresponding to maximal eigenvector of $X - xx^T$. Refer to resulting spatial branching scheme as *eigenvector branching.*
- Required values of μ^+ and μ^- could be obtained by explicitly maximizing/minimizing a^Tx for feasible (x, X). Alternatively, if constraints imply $\|x\| \le M$ then can certainly take $\mu^+ = M\|a\|$, $\mu^- = -M\|a\|$.

Consider applying eigenvector branching to instances of the Two-Trust-Region Subproblem

$$\begin{array}{lll} \mathrm{TTRS}: & \min & x^T Qx + c^T x \\ & \text{s.t.} & x^T x \le 1, \quad x^T Hx + 2h^T x \le 1, \end{array}$$

where Q is indefinite and $H \succ 0$. Constraints can also be written in SOC form as $\|x\|^2 \le 1$, $\|H^{1/2}x\|^2 + (h^T x)^2 \le (1 - h^T x)^2$.
SDP (Shor) relaxation is then

$$\begin{array}{lll} \mathrm{TTRS}_{\mathrm{SDP}}: & \min & Q \bullet X + c^T x \\ & \text{s.t.} & I \bullet X \le 1, \quad H \bullet X + 2h^T x \le 1, \\ & & X \succeq xx^T. \end{array}$$

Well known that $\mathrm{TTRS}_{\mathrm{SDP}}$ can have a nonzero optimality gap, unlike the simpler trust-region subproblem TRS with one ellipsoid constraint.

Extensive literature for TTRS, also known as the Celis-Dennis-Tapia (CDT) problem.

- SDP relaxation can be proved tight for some special cases, for example $c = h = 0$.
- Known polynomial-time algorithm based on theory of polynomial equations (Barvinok 1993, Bienstock 2016). Method based on solving KKT conditions that is polynomial-time in n and computational precision in Sakaue et al. (2016).
- No known methodology for adding convex constraints that provably tighten SDP relaxation to obtain exact solution.

Consider test problems first constructed in Burer and Anstreicher (2013) to evaluate methodology based on SOC-RLT cuts, further strengthened in Yang and Burer (2016) and Anstreicher (2017). Here use only instances *not solved* in BA (2013).

Table: Previous results on TTRS test instances

		Number of instances solved in:			
n	Instances	A (2017) only	YB (2016) only	Both	Neither
5	38	8	8	12	10
10	70	34	7	14	15
20	104	35	14	24	31
	212	77	29	50	56

All but one of these same 212 problems were solved in recent work by Consolini and Locatelli (2022) using specialized cuts and local search.

When applying eigenvector branching, problem at depth $k \geq 0$ in branching tree includes k branching inequalities that are also used to derive additional constraints. Resulting problem is solved to obtain a lower bound $z(\bar{x}, \bar{X})$ and feasible objective value $v(\bar{x}) = \bar{x}^T Q\bar{x} + c^T\bar{x}$.

If $v(\bar{x})$ is less than the best known value (BKV) then the BKV is updated. The problem is fathomed if either it is infeasible, $z(\bar{x}, \bar{X}) \geq \mathrm{BKV}$ or the relative gap condition

$$\frac{v(\bar{x}) - z(\bar{x}, \bar{X})}{|v(\bar{x})|} < 10^{-4}$$

is satisfied. If none of these conditions holds then two child problems are created using the branching inequalities $a^T x \geq \theta$, $a^T x \leq \theta$, where a with $\|a\| = 1$ is the eigenvector corresponding to the maximal eigenvalue of $\bar{X} - \bar{x}\bar{x}^T$ and $\theta = a^T\bar{x}$.

Consider branching constraint $a^T x \geq \theta$. Rather than adding linear constraint based on secant approximation, combine with SOC representations of ellipsoid constraints (with $h = 0$) to get $\|(a^T x - \theta)x\| \leq a^T x - \theta$ and $\|(a^T x - \theta)H^{1/2}x\| \leq a^T x - \theta$. Substituting X for xx^T, obtain the *SOC-RLT* constraints

$$\|Xa - \theta x\| \leq a^T x - \theta, \qquad \|H^{1/2}Xa - \theta H^{1/2}x\| \leq a^T x - \theta.$$

Example: Branching tree for instance_10_607, unsolved in both A (2017) and YB (2016). Relative gap at root (SDP relaxation) is about 38%.

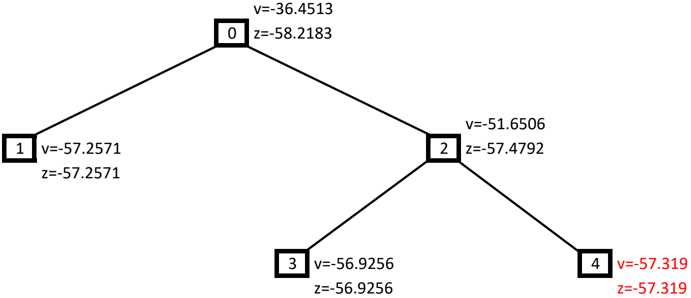


Figure: Branching tree for instance_10_607

Behavior on instance_10_607 is typical. For all 212 instances, max depth is 4 (204 problems have depth 2 or less) and max nodes is 11 (206 problems have 7 nodes or less).

Following charts give distributions for depth, nodes, and final relative gap for all 212 problems. In addition to version using original SOC-RLT constraints, consider two alternatives:

- (SOCRLT - Nullspace) Use SOC-RLT constraints, but branch using $a = Z\bar{a}$, where $\bar{a}$ is the eigenvector for the maximum eigenvalue of $Z^T \bar{X} Z$ and the columns of Z are an orthonormal basis for the nullspace of $\bar{x}^T$. Gives symmetric disjunction with $\theta = a^T \bar{x} = 0$.
- (Linear + RLT) Use linear constraints based on secant approximation, adding all RLT constraints from branching inequalities. Values of μ^+ and μ^- obtained from max/min of $a^T x$ over individual ellipsoids.

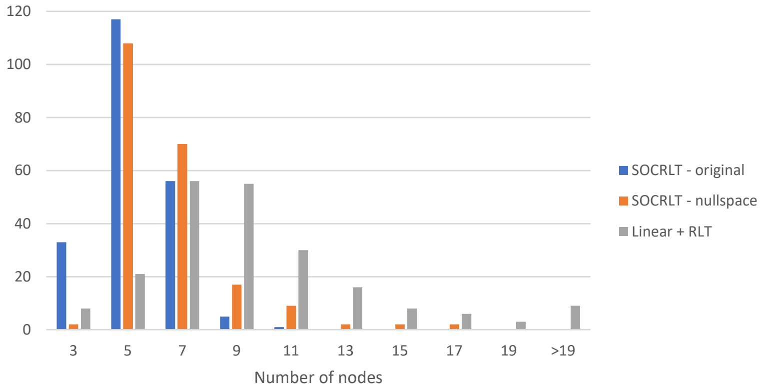


Figure: Number of nodes using eigenvector branching

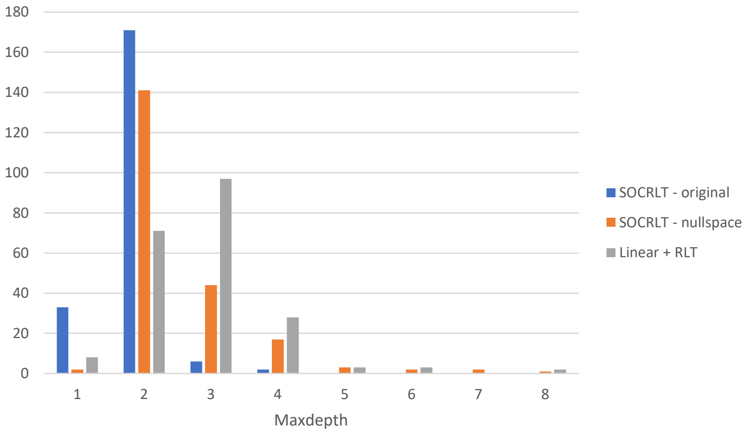


Figure: Maximum depth using eigenvector branching

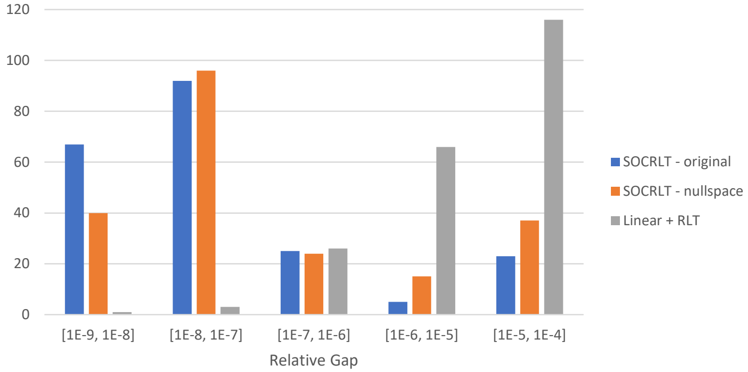


Figure: Relative gap using eigenvector branching

Bottom line: All 3 versions solve all 212 problems, but performance is substantially better using SOC-RLT constraints, and original version is better than nullspace version.

Distribution of relative gaps using SOC-RLT constraints is particularly notable since fathoming criterion is based on a relative gap of 1E-4. For original version, 184 of 212 problems have relative gaps of 1E-6 or less due to large number of terminal nodes with solutions that are *numerically rank-one*, as in the case of instance_10_607.

In SBL (2010), disjunction based on a with $a^T\bar{X}a > (a^T\bar{x})^2$ is used to construct a valid cut that can be added to SDP, without branching. Methodology in SBL (2010) is based on linear constraints only, but can be extended to enforce semidefinite and SOC constraints.

Consider $\mathrm{TTRS}_{\mathrm{SDP}}$ with $h = 0$ as problem of the form SDP with one equality constraint $Y_{00} = 1$ and two homogenous inequality constraints $A_i \bullet Y \le 0$, $i = 1, 2$. Assume $\bar{Y} \succeq 0$, but $a^T\bar{X}a > (a^T\bar{x})^2 = \theta^2$. Applying the Proposition from SBL (2010), but using SOC-RLT constraints as above, if $\bar{Y}$ is in convex hull of rank-one solutions then there must be Y^+ and Y^- so that

$$Y^+ + Y^- = \bar{Y},$$

where $Y^+ \succeq 0$ satisfies the constraints

$$\begin{aligned} A_i \bullet Y^+ &\le 0, \quad i = 1, \ldots, m, \\ \|X^+a - \theta x^+\| &\le a^T x^+ - \theta Y^+_{00}, \\ \|H^{1/2}X^+a - \theta H^{1/2}x^+\| &\le a^T x^+ - \theta Y^+_{00}, \end{aligned}$$

and Y^- satisfies a similar set of constraints with the right-hand-sides $a^T x^+ - \theta Y^+_{00}$ replaced by $\theta Y^-_{00} - a^T x^-$. Note that above constraints are homogenized using variables Y^+_{00} and Y^-_{00}. Initially have $m = 2$, but will eventually have additional linear constraints based on added cuts.

To determine if above constraints are feasible, add artificial variables $(u^+, v^+, w^+) \in \mathbb{R}^m_+ \times \mathbb{R}^2_+ \times \mathbb{R}_+$ to obtain constraints

$$
\begin{aligned}
A_i \bullet Y^+ &\leq u^+_i, \quad i = 1, \ldots, m, \\
\|X^+ a - \theta x^+\| &\leq a^T x^+ - \theta Y^+_{00} + v^+_1, \\
\|H^{1/2} X^+ a - \theta H^{1/2} x^+\| &\leq a^T x^+ - \theta Y^+_{00} + v^+_2, \\
Y^+ + w^+ I &\succeq 0,
\end{aligned}
$$

and similarly add variables (u^-, v^-, w^-) to constraints for Y^-.

Lemma. Assume that $\bar{Y} \succeq 0$, $\theta = a^T \bar{x}$, $a^T \bar{X} a > \theta^2$. Consider the problem to minimize $e^T u^+ + e^T u^- + e^T v^+ + e^T v^- + w^+ + w^-$ subject to the above constraints. Let $\bar{S}$ be the dual solution matrix for the constraint $Y^+ + Y^- = \bar{Y}$. Then $\bar{S} \bullet Y \leq 0$ for any Y that is in the convex hull of rank-one solutions of $\mathrm{TTRS}_{\mathrm{SDP}}$.

Result: $\bar{S} \bullet Y \leq 0$ is a valid cut that can be added to SDP.

Consider same 212 TTRS problems, but now applying disjunctive cuts instead of eigenvector branching. Add up to 25 cuts, using same termination criterion based on relative gap. Update constraints included in cut-generation problem every 5 iterations.

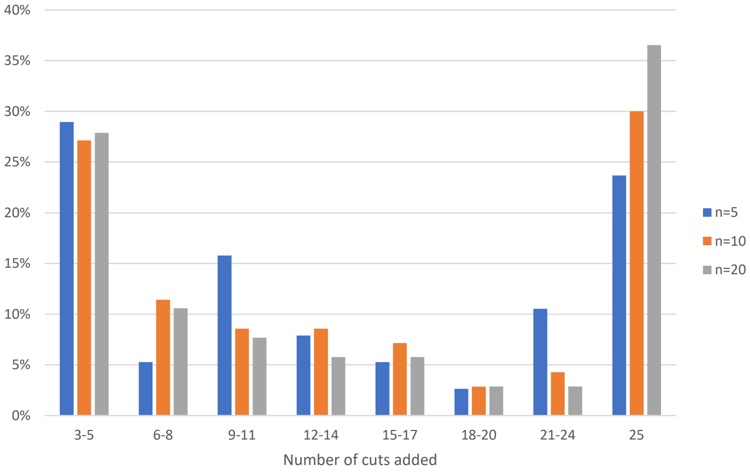


Figure: Number of disjunctive cuts

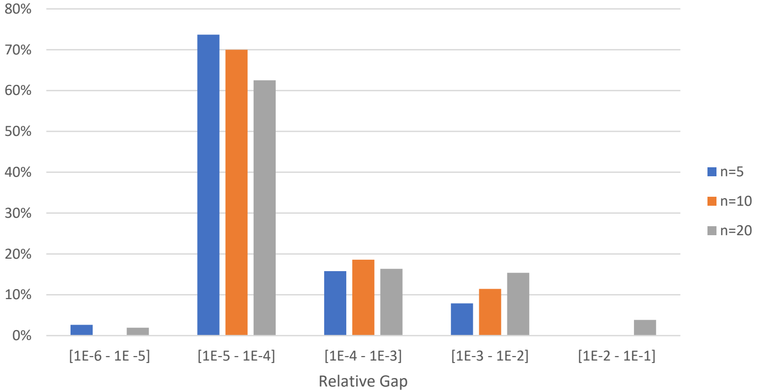
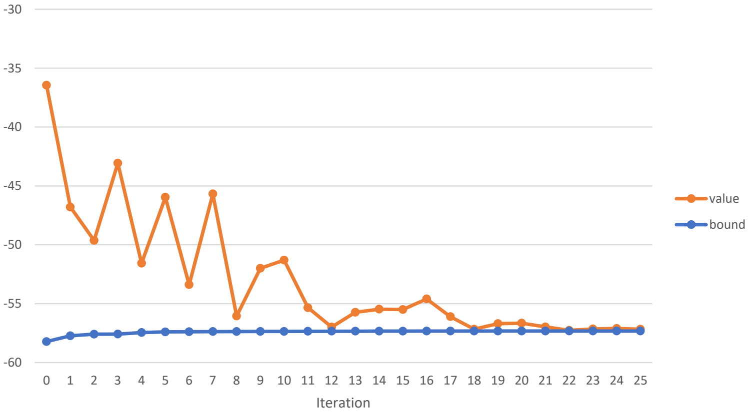


Figure: Relative gap using disjunctive cuts

Overall solve 145 of 212 problems using up to 25 cuts; increases to 164 using up to 50 cuts. Comparing results with eigenvector branching, find that for unsolved instances, difficulty is always due to *failure to generate near-optimal solution* rather than tight lower bound. For example, consider iteration sequence for instance_10_607:



- Eigenvector branching is a spatial branching scheme based on failure of the rank-one condition $X = xx^T$.
- Linear branching inequalities can be combined with SOC constraints to form SOC-RLT constraints in child problems.
- Performance of eigenvector branching on difficult TTRS instances is excellent, especially using SOC-RLT constraints.
- Using disjunctive cuts instead on eigenvector branching on same TTRS problems generates tight lower bounds but may fail to generate near-optimal solutions. Possible remedy is to add local search, as in Consolini and Locatelli (2022).

Thank You