SOLVING POLYNOMIAL OPTIMIZATION PROBLEMS USING REAL ALGEBRAIC GEOMETRY, MOMENTS, AND SEMIDEFINITE PROGRAMMING

Kartik Sivaramakrishnan
Department of Mathematics
North Carolina State University
kksivara@ncsu.edu
http://www4.ncsu.edu/~kksivara

Algebra Seminar
North Carolina State University
March 30, 2007

Contents

- Polynomial Optimization (POP)
- Sum of squares approach for solving POPs – Dual scheme
- Semidefinite Programming (SDP)
- Moment approach for solving POPs – Primal scheme
- One numerical example
- Exploiting sparsity in POP
- Exploiting symmetry in POP
- Decomposition approach for structured SDPs

Underlying paradigm in today's talk

1. A polynomial optimization problem (POP) is a *hard* nonconvex optimization problem.
2. We will develop a hierarchy of *convex relaxations* for the POP whose solutions converge to a real optimal solution to the POP.
3. The convex relaxations are *semidefinite programs* (SDPs) that can be solved efficiently using *interior point methods* (IPMs). However, the SDPs in the hierarchy get very large!
4. We want to exploit the *sparsity* and *symmetry* in the underlying POP to get *structured* SDP relaxations (Phase 1)
5. We will then solve these structured SDPs using *decomposition techniques* in a parallel computing environment (Phase 2).
6. The solution of POPs lies at the confluence of *pure* and *applied* mathematics; and has several *applications* in science and engineering.

Polynomial preliminaries: 1

1. $p(x) = \sum_{\alpha \in Z_+^n} p_\alpha x_1^{\alpha_1} \ldots x_n^{\alpha_n} = \sum_{\alpha \in Z_+^n} p_\alpha x^\alpha$ is a multivariate polynomial.
2. Degree of the monomial $x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ is $|\alpha| = \sum_{i=1}^{n} \alpha_i$.
3. Degree of $p(x)$ is the max degree of a monomial x^α with $p_\alpha \neq 0$.
4. $x_1^2 x_2^3 x_3 - x_2^4 + 2x_1 x_3 - 1$ is a polynomial in 3 variables with degree 6.
5. Let $S_n^d \; := \; \{\alpha \in Z_+^n : \; |\alpha| \leq d\}, \;\; |S_n^d| = \binom{n+d}{d}$.
6. Represent $p(x)$ of degree d with its sequence of coefficients $p = (p_\alpha)_{\alpha \in S_n^d}$.
7. The set $\mathbb{R}_d[x_1, \ldots, x_n]$ of real polynomials of degree $\leq d$ with real coefficients is isomorphic to $\mathbb{R}^{\binom{n+d}{d}}$.

Polynomial preliminaries: 2

1. A polynomial $p(x)$ is nonnegative if $p(x) \geq 0, \forall x \in \mathbb{R}^n$.
   Let $\mathcal{P}_{n,d} = \{p_\alpha : \ p(x) = \sum_\alpha p_\alpha x^\alpha \geq 0, \forall x \in \mathbb{R}^n$, deg$(p(x)) \leq d\}$.
2. A polynomial $p(x)$ has a sum of squares (SOS) if

$$p(x) \; = \; \sum_i p_i(x)^2.$$

   for some polynomials p_i. It is clear that $p(x)$ has even degree $2d$, and each $p_i(x)$ has degree $\leq d$. Let
   $\Sigma_{n,d} = \{p_\alpha : \ p(x) = \sum_\alpha p_\alpha$ has an SOS, deg$(p(x)) \leq d\}$.
3. **Example:** $x_1^2 + x_2^2 + 2x_1x_2 + x_3^6 = (x_1 + x_2)^2 + (x_3^3)^2$ is an SOS.
4. The sets $\mathcal{P}_{n,d}$ and $\Sigma_{n,d}$ are both closed convex cones.

When does a nonnegative polynomial have an SOS decomposition?

1. An SOS polynomial is nonnegative but not every nonnegative polynomial is an SOS.
2. Hilbert in 1888 showed that $\mathcal{P}_{n,d} = \Sigma_{n,d}$, i.e., every nonnegative polynomial of degree d in n variables is an SOS when
   - $n = 1$ ($d \geq 2$ and even)
   - $d = 2$ (all n)
   - $n = 2, d = 4$

   In all other cases, $\Sigma_{n,d} \subset \mathcal{P}_{n,d}$ (see Chapter 6 in *Squares* by Rajwade).
3. The Motzkin polynomial $x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 + 1$ is nonnegative but *not* a SOS.

Polynomial optimization

1. Let $p, g_1, \ldots, g_m$ be polynomials with real coefficients defined on $\mathbb{R}^n$. The polynomial optimization problem is

$$\min_{x \in S} \; p(x)$$

   where S is a *compact semi-algebraic* set given by

$$S \;:=\; \{x \in \mathbb{R}^n : \; g_i(x) \;\geq\; 0, \;\; i = 1, \ldots, m\}.$$

   The set S can be nonconvex and even disconnected.

2. Applications in nonconvex and combinatorial optimization, statistics, symbolic computation, nonlinear optimal control, nonlinear PDEs.

3. The general problem is NP-hard.

4. An equivalent problem is finding the *greatest* lower bound t, i.e.,

$$\max_t \{t : \; p(x) - t \;\geq\; 0 \;\; \forall x \in S\}$$

Semidefinite programming

(SDP)	max	$C \bullet X$
	s.t.	$\mathcal{A}(X) = b$
		$X \succeq 0$

(SDD)	min	$b^T y$
	s.t.	$\mathcal{A}^T y - S = C$
		$S \succeq 0$

- **Notation**
  - $X, S, C \in \mathcal{S}^n$, $b \in \mathbb{R}^m$
  - $A \bullet B = \text{trace}(AB) = \sum_{i,j=1}^n A_{ij} B_{ij}$ (Frobenius inner product)
  - The operator $\mathcal{A} : \mathcal{S}^n \to \mathbb{R}^m$ and its adjoint $\mathcal{A}^T : \mathbb{R}^m \to \mathcal{S}^n$ are

$$\mathcal{A}(X) = \begin{pmatrix} A_1 \bullet X \\ \vdots \\ A_m \bullet X \end{pmatrix}, \quad \mathcal{A}^\top y = \sum_{i=1}^m y_i A_i$$

where $A_i \in \mathcal{S}^n$, $i = 1, \ldots, m$

Connection between SOS and SDP

Given $p(x) = \sum_{\alpha \in Z_+^n, |\alpha| \leq 2d} p_\alpha x^\alpha$ of degree $2d$. The following are equivalent

1. $p(x)$ has an SOS representation
2. $p(x) = z^T M z$ for some matrix $M = (M_{\beta,\gamma})_{|\beta|,|\gamma| \leq d} \succeq 0$ where $z = (x^\beta)_{|\beta| \leq d}$.
3. The following semidefinite feasibility problem

$$\begin{aligned} M &\succeq 0, \\ \sum_{i,j \in S_n^d : i+j=\alpha} M_{i,j} &= p_\alpha, \;\; |\alpha| \leq 2d \end{aligned}$$

   is feasible.

This problem is in the primal (SDP) form with matrix size $\binom{n+d}{d}$ and $\binom{n+2d}{2d}$ equality constraints.

Example

Does $p(x) = 2x_1^4 + 2x_1^3x_2 - x_1^2x_2^2 + 5x_2^4$ have an SOS?

1. Write $p(x) = z^TMz$ where $z = [1 \;\; x_1 \;\; x_2 \;\; x_1^2 \;\; x_1x_2 \;\; x_2^2]$ with rows and columns of M indexed by $\{\{0,0\},\{1,0\},\{0,1\},\{2,0\},\{1,1\},\{0,2\}\}$.
2. Equating coefficients gives $M_{11} = 0$, $M_{12} = 0$, $M_{13} = 0$, $M_{14} = 0$, $M_{15}+M_{23} = 0$, $M_{33}+2M_{16} = 0$, $M_{24} = 0$, $M_{25}+M_{34} = 0$, $M_{26}+M_{35} = 0$, $M_{36} = 0$, $M_{44} = 2$, $M_{45} = 2$, $M_{55} + 2M_{46} = -1$, $M_{56} = 0$, and $M_{66} = 5$.
3. The SDP feasibility problem with these 15 constraints and $M \succeq 0$ has a feasible solution. So $p(x)$ has an SOS. Indeed,

$$\begin{aligned} p(x) \;=\; & 1.0036(0.6070x_1^2 - 0.7736x_1x_2 + 0.1820x_2^2)^2 + \\ & 1.9658(0.7206x_1^2 + 0.6323x_1x_2 + 0.2845x_2^2)^2 + \\ & 5.4266(-0.3351x_1^2 - 0.0416x_1x_2 + 0.9412x_2^2)^2 \end{aligned}$$

   This SOS decomposition can be constructed from the eigenvalue and eigenvector information of M.

SOS approach for POP: 1

1. Given a polynomial $p(x)$ in n variables and even degree $2d$.

$$\begin{aligned} p_{\min} &= \inf_{x \in \mathbb{R}^n} p(x) \\ &= \sup\{t : \; p(x) - t \geq 0 \;\; \forall x \in \mathbb{R}^n\}. \end{aligned}$$

   (an LP in one variable t and infinite number of constraints). The constraint requires $p(x) - t \in \mathcal{P}_{n,2d}$.

2. A constrained version of this problem is

$$p_{sos} = \sup\{t : \; p(x) - t \text{ has an SOS}\}.$$

   Since $p(x) - t \in \Sigma_{n,2d} \subseteq \mathcal{P}_{n,2d}$, we have $p_{\min} \geq p_{sos}$.

3. The second problem is the following SDP

$$\begin{array}{lrcl} \sup & t & & \\ \text{s.t.} & M_{0,0} + t & = & p_0, \\ & \displaystyle\sum_{i,j \in S_n^d : i+j=\alpha} M_{i,j} & = & p_\alpha \;\; \forall \alpha, \\ & M & \succeq & 0. \end{array}$$

SOS approach for POP: 2

1. Consider the sequence of SDP approximations

$$p_r = \sup\{t : \ p(x) - t \text{ has an SOS with degree } 2r\}.$$

   for $r \geq d$. Since $p(x) - t \in \Sigma_{n,2d} \subseteq \mathcal{P}_{n,2r}$, we have $p_{\min} \geq p_r \geq p_{sos}$.

2. As $r \to \infty$ we get tighter SOS approximations for the POP.

3. The second problem is the following SDP

$$\begin{array}{lrcl} \sup & t & & \\ \textbf{s.t.} & M_{0,0} + t & = & p_0, \\ & \displaystyle\sum_{i,j \in S_n^r : i+j=\alpha} M_{i,j} & = & p_\alpha \ \forall \alpha, \\ & M & \succeq & 0. \end{array}$$

   where the coefficients p_α of monomials of degree higher than $2d$ are set to 0.

Moment approach for POP: 1

1. The POP is equivalent to

$$p_{\min} = \min_{\mu} \int p(x) \mathrm{d}\mu$$

   where $\mu \in \mathbb{R}^n$ is a probability measure with support on S. (This LP is the dual to our earlier infinite LP, and has 1 equality constraint and infinite number of variables).

2. We have

$$\int p(x) \mathrm{d}\mu = \sum_{\alpha} p_{\alpha} \int x^{\alpha} \mathrm{d}\mu = p^T y$$

   where y denotes the sequence of moments of μ.

3. Let $M_r(y)$ be the truncated moment matrix indexed by S_n^r with (α, β)th entry given by $y_{\alpha+\beta}$ with $\alpha, \beta \in S_n^r$.

4. Let $M_r(g * y)$ be a localizing matrix indexed by S_n^r with (α, β)th entry given by $\sum_{\gamma} g_{\gamma} y_{\alpha+\beta+\gamma}$.

Moment approach for POP: 2

1. Let $2d_0$ or $2d_0-1$ be the degree of $p(x)$, and $2d_j-1$ or $2d_j$ be the degrees of $g_j(x)$ in POP. The *relaxation order* r is some integer $\geq \max_{i=0,\ldots,m} d_i$.

2. A necessary condition for y to be a sequence of moments of μ is

$$\begin{aligned} M_r(y) &\succeq 0 \\ M_{r-d_j}(g_j * y) &\succeq 0, \quad j = 1, \ldots, m. \end{aligned}$$

   As r increases, these necessary conditions get tighter.

3. This gives a hierarchy of SDP relaxations for POP, where the rth order relaxation is

$$\begin{array}{lrcl} \min & \sum_\alpha p_\alpha y_\alpha & & \\ \text{s.t.} & M_r(y) & \succeq & 0 \\ & M_{r-d_j}(g_j * y) & \succeq & 0, \quad j = 1, \ldots, m. \end{array}$$

   As $r \to \infty$, the solutions to these SDPs (providing lower bounds) converge to a real optimal solution of POP (Lasserre, Schweighofer).

4. There is also a sufficient condition to verify optimality and a mechanism to extract solutions (Henrion & Lasserre).

Illustrative Example

1. Consider the POP

$$\begin{array}{lrcl} \min & -x_2 & & \\ \text{s.t.} & 3+2x_2-x_1^2-x_2^2 & \geq & 0 \\ & -x_1-x_2-x_1x_2 & \geq & 0 \\ & 1+x_1x_2 & \geq & 0 \end{array}$$

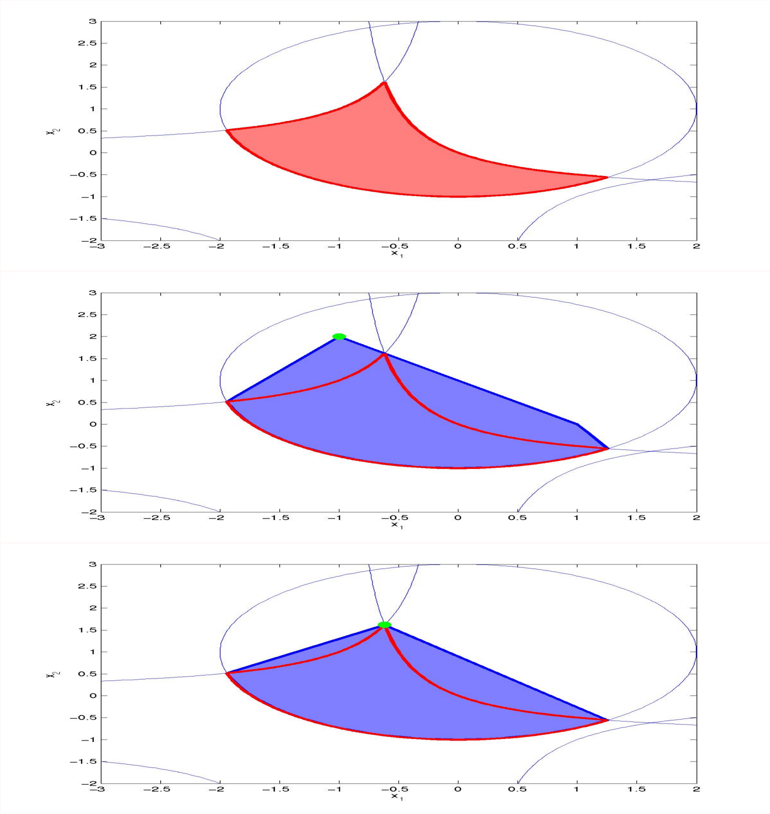
   Global optimal solution is $x^* = (-0.6180, 1.6180)$ with value -1.6180.

2. The first SDP relaxation in hierarchy (of order 1) is

$$\begin{array}{lrcl} \min & -y_{01} & & \\ \text{s.t.} & \begin{bmatrix} 1 & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{bmatrix} & \succeq & 0 \\ & 3+2y_{01}-y_{20}-y_{02} & \geq & 0 \\ & -y_{10}-y_{01}-y_{11} & \geq & 0 \\ & 1+y_{11} & \geq & 0. \end{array}$$

   Optimal objective value of the SDP is -2.

From Henrion & Lasserre (IEEE Control Magazine)



Exploiting sparsity in POP: 1

1. Consider the POP

$$
\begin{array}{lrcl}
\min & x_1x_2 + x_1x_3 + x_1x_4 & & \\
\text{s.t.} & x_1^2 + x_2^2 & \leq & 1 \\
 & x_1^2 + x_3^2 & \leq & 1 \\
 & x_1^2 + x_4^2 & \leq & 1.
\end{array}
$$

2. Sets F_k contain variables that appear in a monomial of $p(x)$.
3. Let H_j contain all variables in constraint polynomial $g_j(x)$.
4. Construct *correlative sparsity pattern* (CSP) graph from POP, where $V = \{1, 2, \ldots, n\}$ and

$$E = \{(i, j) : i, j \in F_k \text{ for some } k \text{ or } i, j \in H_l \text{ for some } l\}$$

5. Let I_i, $i = 1, \ldots, t$ be a set of maximal cliques of CSP graph. (find chordal extension of graph and then maximal cliques in extension)
6. In our example, we have $I_1 = \{1, 2\}$, $I_2 = \{1, 3\}$, $I_3 = \{1, 4\}$.

Exploiting sparsity: 2

For our example, the 1st order sparse SDP relaxation is

$$
\begin{array}{lrcl}
\min & y_{1100} + y_{1010} + y_{1001} & & \\
\text{s.t.} & \begin{bmatrix} 1 & y_{1000} & y_{0100} \\ y_{1000} & y_{2000} & y_{1100} \\ y_{0100} & y_{1100} & y_{0200} \end{bmatrix} & \succeq & 0 \\
 & \begin{bmatrix} 1 & y_{1000} & y_{0010} \\ y_{1000} & y_{2000} & y_{1010} \\ y_{0010} & y_{1010} & y_{0020} \end{bmatrix} & \succeq & 0 \\
 & \begin{bmatrix} 1 & y_{1000} & y_{0001} \\ y_{1000} & y_{2000} & y_{1001} \\ y_{0001} & y_{1001} & y_{0002} \end{bmatrix} & \succeq & 0 \\
 & 1 - y_{2000} - y_{0200} & \geq & 0 \\
 & 1 - y_{2000} - y_{0020} & \geq & 0 \\
 & 1 - y_{2000} - y_{0002} & \geq & 0.
\end{array}
$$

which has 11 variables and 3 SDP blocks of size 3.
(15 variables and 1 SDP block of size 5 for the regular SDP relaxation).

Exploiting symmetry: 1

1. Consider the unconstrained POP of minimizing

$$p(x) \;=\; x_1^6 + x_2^6 - x_1^4 x_2^2 - x_2^4 x_1^2 - x_1^4 - x_2^4 - x_1^2 - x_2^2 + 3x_1^2 x_2^2 + 1.$$

2. The objective $p(x)$ is invariant under the dihedral group D_4 of order 8.
3. The moment matrix $M_3(y)$ in the first SDP relaxation has following structure

$$M_3(y) \;=\; \begin{bmatrix} 1 & 0 & 0 & y_{20} & 0 & y_{20} & 0 & 0 & 0 & 0 \\ 0 & y_{20} & 0 & 0 & 0 & 0 & y_{40} & 0 & y_{22} & 0 \\ 0 & 0 & y_{20} & 0 & 0 & 0 & 0 & y_{22} & 0 & y_{40} \\ y_{20} & 0 & 0 & y_{40} & 0 & y_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & y_{22} & 0 & 0 & 0 & 0 & 0 \\ y_{20} & 0 & 0 & y_{22} & 0 & y_{40} & 0 & 0 & 0 & 0 \\ 0 & y_{40} & 0 & 0 & 0 & 0 & y_{60} & 0 & y_{42} & 0 \\ 0 & 0 & y_{22} & 0 & 0 & 0 & 0 & y_{42} & 0 & y_{42} \\ 0 & y_{22} & 0 & 0 & 0 & 0 & y_{42} & 0 & y_{42} & 0 \\ 0 & 0 & y_{40} & 0 & 0 & 0 & 0 & y_{42} & 0 & y_{60} \end{bmatrix}$$

Exploiting symmetry: 2

1. One can compute an appropriate *symmetry adapted basis* that exploits the symmetry (Gatermann & Parrilo, Fässler & Stiefel).

2. After the symmetry reduction, an *equivalent* SDP relaxation is

$$
\begin{array}{lrcl}
\min & 2y_{60} - 2y_{42} - 2y_{40} - 2y_{20} + 3y_{22} + 1 & & \\
\text{s.t.} & \begin{bmatrix} 1 & \frac{y_{20}}{\sqrt{2}} \\ \frac{y_{20}}{\sqrt{2}} & y_{22} + y_{40} \end{bmatrix} & \succeq & 0 \\
& y_{22} & \geq & 0 \\
& y_{40} - y_{22} & \geq & 0 \\
& \begin{bmatrix} y_{20} & y_{40} & y_{22} \\ y_{40} & y_{60} & y_{42} \\ y_{22} & y_{42} & y_{42} \end{bmatrix} & \succeq & 0
\end{array}
$$

with 5 variables and 4 blocks of sizes 2, 1, 1, and 3 respectively (compare with 55 variables and 1 SDP block of size 10).

Preprocessing to obtain structured SDPs - Phase 1

1. One can exploit both sparsity and symmetry in POP to obtain structured SDP relaxations.
2. Exploiting sparsity gives structured SDP relaxations that are weaker, and one has increase the order r of SDP relaxations to solve the underlying POP.
3. The weaker SDP relaxations also converge to an optimal solution of POP, albeit more slowly (Lasserre).
4. Exploiting symmetry gives an equivalent structured SDP relaxation; the number and size of blocks in new SDP are determined entirely by the *irreducible representations* of the group (Schur's lemma!)
5. Kartik is currently developing *decomposition techniques* that solve the structured SDPs in a parallel & distributed high performance computing environment.

Decomposition approach for structured SDPs - Phase 2

1. The structured SDPs have the form

$$\begin{array}{ll} \min\limits_{y,z_i} & b^T y + \sum\limits_{i=1}^{r} d_i^T z_i \\ \text{s.t.} & A_i^T y + B_i^T z_i \succeq C_i, \quad i = 1, \ldots, r \end{array}$$

2. If y variables were absent, the problem decomposes into r independent subproblems in z_i variables.
3. We are developing a *conic Benders decomposition* approach to solve the structured SDP in an iterative fashion between a master problem (over y variables); and r subproblems (over z_i variables) which are solved in parallel.
4. We have implemented the decomposition algorithm on the distributed *Henry2* cluster at NC State University.

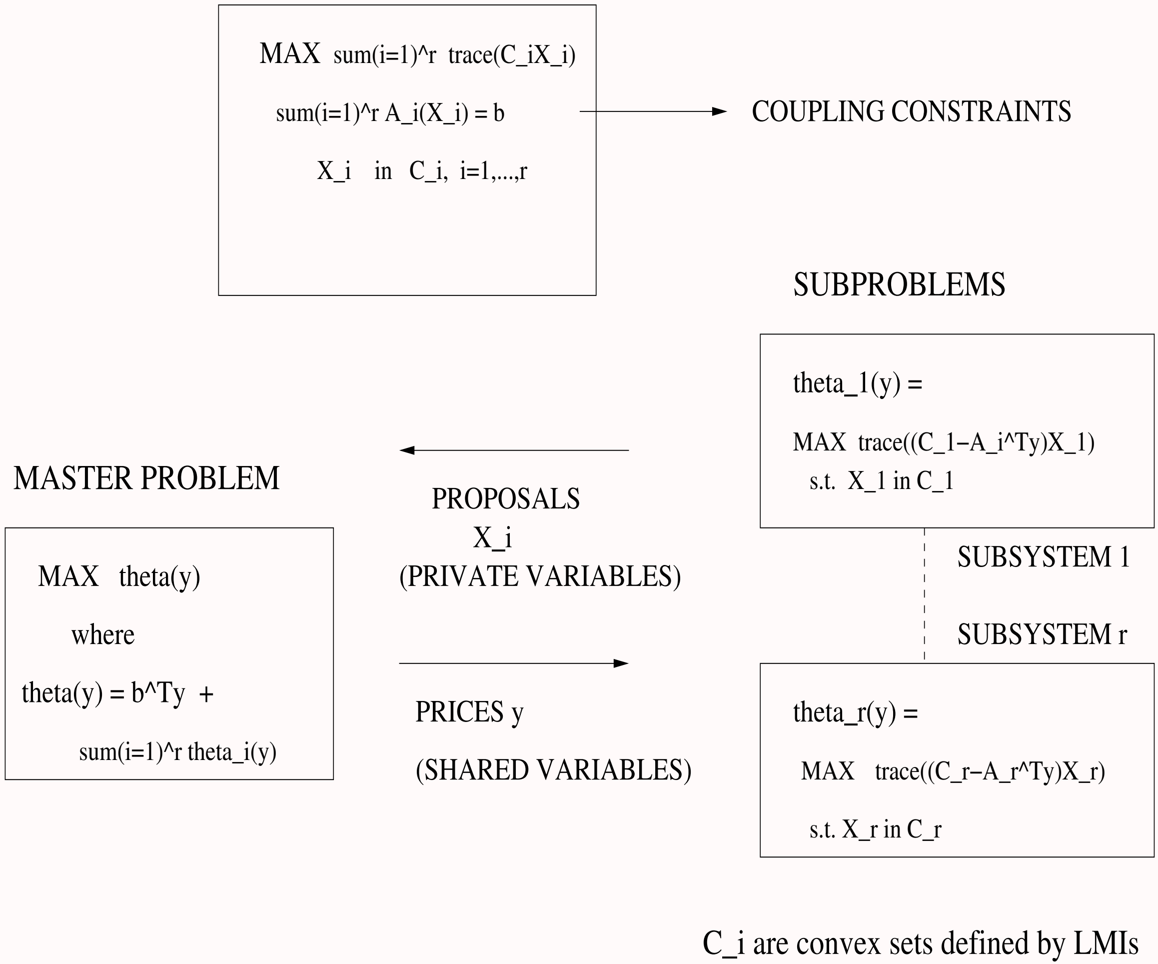


Figure 1: Decomposition by prices

Thank you for your attention!.

Questions, Comments, Suggestions ?

The slides from this talk are available online at

http://www4.ncsu.edu/∼kksivara/publications/kartik-algebra.pdf

A technical report appears at

http://www4.ncsu.edu/∼kksivara/publications/parallel-conic-blockangular.pdf

Bibliography: 1

1. K. Sivaramakrishnan, *A parallel conic interior point decomposition approach for block angular semidefinite programs*, submitted. Available at http://www4.ncsu.edu/∼kksivara/publications/parallel-conic-blockangular.pdf

2. M.J. Todd, *Semidefinite optimization*, Acta Numerica, 10(2001), pp. 515-560.

3. P.A. Parrilo, *Semidefinite programming relaxations for semi-algebraic problems*, Mathematical Programming, 96(2003), pp. 293-320.

4. J.B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM Journal on Optimization, 11(2001), pp. 796-817.

5. D. Henrion and J.B. Lasserre, *GLOPTIPOLY: Global optimization over polynomials with MATLAB and SeDuMi*, http://www.laas.fr/∼henrion/papers/gloptipoly.pdf

6. A. Fässler and E. Stiefel, *Group Theoretical Methods and Their Applications*, Birkhäuser, Boston, 1992.

7. K. Gatermann and P.A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, Journal of Pure and Applied Algebra 192(2004), pp. 95-128.

8. H. Waki, S. Kim, M. Kojima, and M. Muramatsu, *Sums of squares and semidefinite programming relaxations for polynomial problems with structured sparsity*, SIAM Journal on Optimization, 17(2006), pp. 218-242.

Bibliography: 2

9. D. Henrion and J.B. Lasserre, *Detecting global optimality and extracting solutions in GLOPTIPOLY*, In Positive Polynomials in Control, edited by D. Henrion and A. Garulli, Lecture Notes in Control and Information Science, Volume 312, Springer Verlag, Berlin, 2005, pp. 293-310.

10. D. Henrion and J.B. Lasserre, *Solving nonconvex optimization problems*, IEEE Control Systems Magazine, 24(2004), pp. 72-83.

11. E. de Klerk, *Exploiting algebraic symmetry in semidefinite programming: a review*, Talk at Delft, the Netherlands, October, 2005. Slides available at `http://www4.ncsu.edu/~kksivara/etienne-symmetry.pdf`

12. K.C. Kiwiel, *Proximity control in bundle methods for convex nondifferentiable optimization*, Mathematical Programming, 46(1990), pp. 105-122.

13. A. Ruszczyński, *Nonlinear Optimization*, Princeton University Press, 2006.

14. J.B. Lasserre, *Convergent SDP relaxations in polynomial optimization with sparsity*, Technical Report #05-612, LAAS, Toulouse, France, 2005. Available at `http://www.optimization-online.org/DB_HTML/2006/04/1367.html`.

15. M. Schweighofer, *Optimization of polynomials on compact semialgebraic sets*, SIAM Journal on Optimization, 15(2005), pp. 805-825.

16. H. Waki, S. Kim, M. Kojima, and M. Muramatsu, *SparsePOP: A sparse semidefinite programming relaxation of polynomial optimization problems*, Research Report B-414, Dept. of Mathematical and Comuting Sciences, Tokyo Institute of Technology, 2005. Available at `http://www.optimization-online.org/DB_HTML/2005/03/1091.html`.