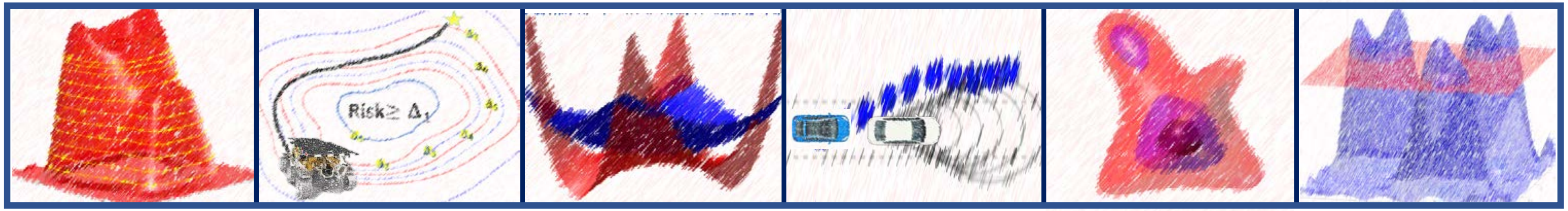
Lecture 6

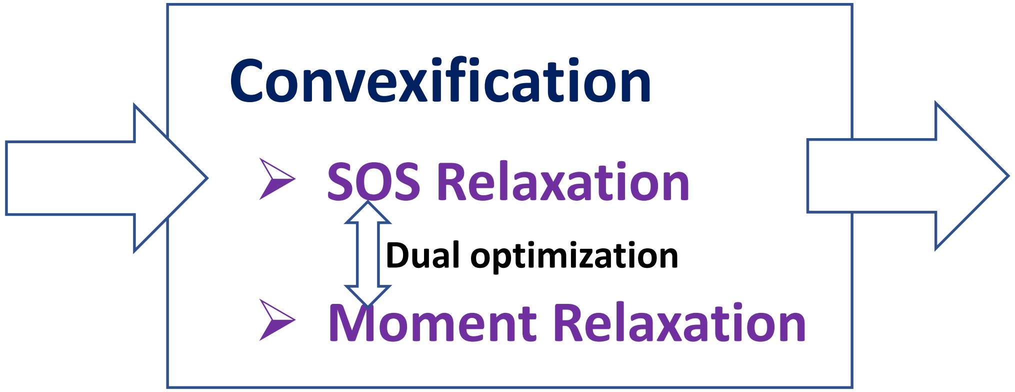
Modified Sum-of-Squares Relaxations for Large Scale Optimizations

MIT 16.S498: Risk Aware and Robust Nonlinear Planning
Fall 2019

Ashkan Jasour



Moment-SOS Relaxations



Nonlinear-Nonconvex Optimization

Convexification

- SOS Relaxation
- Moment Relaxation

Dual optimization

Semidefinite Program

Moment-SOS Relaxations: Applications in Robotics and Control

Motion Planning

- A. Majumdar, R. Tedrake, “Funnel libraries for real-time robust feedback motion planning”, international journal of robotics and research(IJRR), Volume: 36 issue: 8, page(s): 947-982, 2017
- S. Singh, A. Majumdar, J.J. Slotine, M. Pavone “Robust Online Motion Planning via Contraction Theory and Convex Optimization”, IEEE International Conference on Robotics and Automation (ICRA), 2017
- A. Majumdar, M. Tobenkin, R.Tedrake, “Algebraic verification for parameterized motion planning libraries”, American Control Conference (ACC), 2012

Planning and Controllers for UAV

- R. Deits, R. Tedrake” Efficient mixed-integer planning for UAVs in cluttered environments”, IEEE International Conference on Robotics and Automation (ICRA) 2015.
- A. J. Barry, A. Majumdar, R. Tedrake, “Safety verification of reactive controllers for UAV flight in cluttered environments using barrier certificates”, IEEE International Conference on Robotics and Automation (ICRA) 2012.

Legged Robots

- M.Posa, T. Koolen, R. Tedrake, “Balancing and Step Recovery Capturability via Sums-of-Squares Optimization”, Robotics: Science and Systems, 2017
- I. R. Manchester, M. M. Tobenkin, M. Levashov, R. Tedrake “Regions of Attraction for Hybrid Limit Cycles of Walking Robots”, 18th IFAC World Congress, Volume 44, Issue 1, Pages 5801-5806

Real-Time Planning

- A. A. Ahmadi, A. Majumdary, “Some applications of polynomial optimization in operations research and real-time decision making”, Optimization Letters, Volume 10, Issue 4, pp 709–729, 2016.

Controller Design

- A. Majumdar, A. A. Ahmadi, and R. Tedrake , “Control Design Along Trajectories via Sum of Squares Optimization” , International Conference on Robotics and Automation (ICRA), 2013
- J. Moore, R. Tedrake, “Adaptive control design for underactuated systems using sums-of-squares optimization”, American Control Conference 2014
- R. Tedrake , I. R. Manchester , M. Tobenkin , J. W. Roberts, “LQR-trees: Feedback Motion Planning via Sums-of-Squares Verification”, International Journal of Robotics Research, Volume 29 Issue 8, Pages 1038-1052, 2010

Moment-SOS Relaxations: Applications in Robotics and Control

Validation

- D. Wagner, D. Henrion, M. Hromcik. Measures and LMIs for Adaptive Control Validation. To be registered as a LAAS-CNRS Research Report, March 2019. To be presented at the IEEE Conference on Decision and Control, Nice, France, December 2019.

- A. A. Ahmadi, Pablo A Parrilo , “Sum of Squares Certificates for Stability of Planar, Homogeneous, and Switched Systems” IEEE Transactions on Automatic Control, 2017

- S. Shen, R. Tedrake, “Compositional Verification of Large-Scale Nonlinear Systems via Sums-of-Squares Optimization” , American Control Conference (ACC) 2018

Environment Representation

- A. A. Ahmadi, G. Hall, A. Makadia, and V. Sindhwani, “Sum of Squares Polynomials and Geometry of 3D Environments” Robotics: Science and Systems, 2017

Control and Analysis

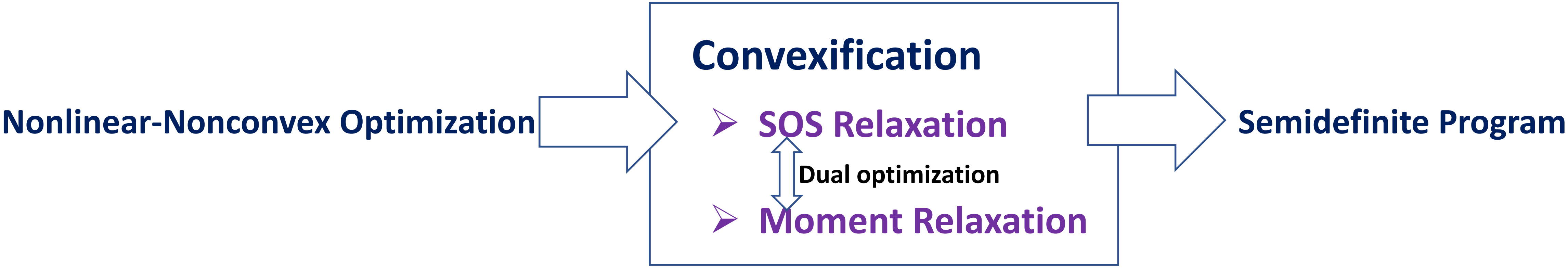
- M. Korda, D. Henrion, C. N. Jones. Controller design and region of attraction estimation for nonlinear dynamical systems. , October 2013, updated in March 2014,

- A. Oustry, M. Tacchi, D. Henrion. Inner approximations of the maximal positively invariant set for polynomial dynamical systems. HAL 02064440, March 2019. IEEE Control Systems Letters, Vol. 3, No. 3, pp. 733-738, 2019. To be presented at the IEEE Conference on Decision and Control, Nice, France, December 2019.

- M. Korda, D. Henrion, J. B. Lasserre. Moments and convex optimization for analysis and control of nonlinear partial differential equations. LAAS-CNRS Research Report 18088, April 2018. Submitted for publication. Presented at the SIAM Conference on Applications of Dynamical Systems, Snowbird, Utah, USA, May 2019.

- M. Korda, D. Henrion, C. N. Jones. Controller design and value function approximation for nonlinear dynamical systems. LAAS-CNRS Research Report 15100, March 2015. Automatica, 67(5):54-66, 2016.

Moment-SOS Relaxations



Nonlinear-Nonconvex Optimization

Convexification

- SOS Relaxation

Dual optimization

- Moment Relaxation

Semidefinite Program

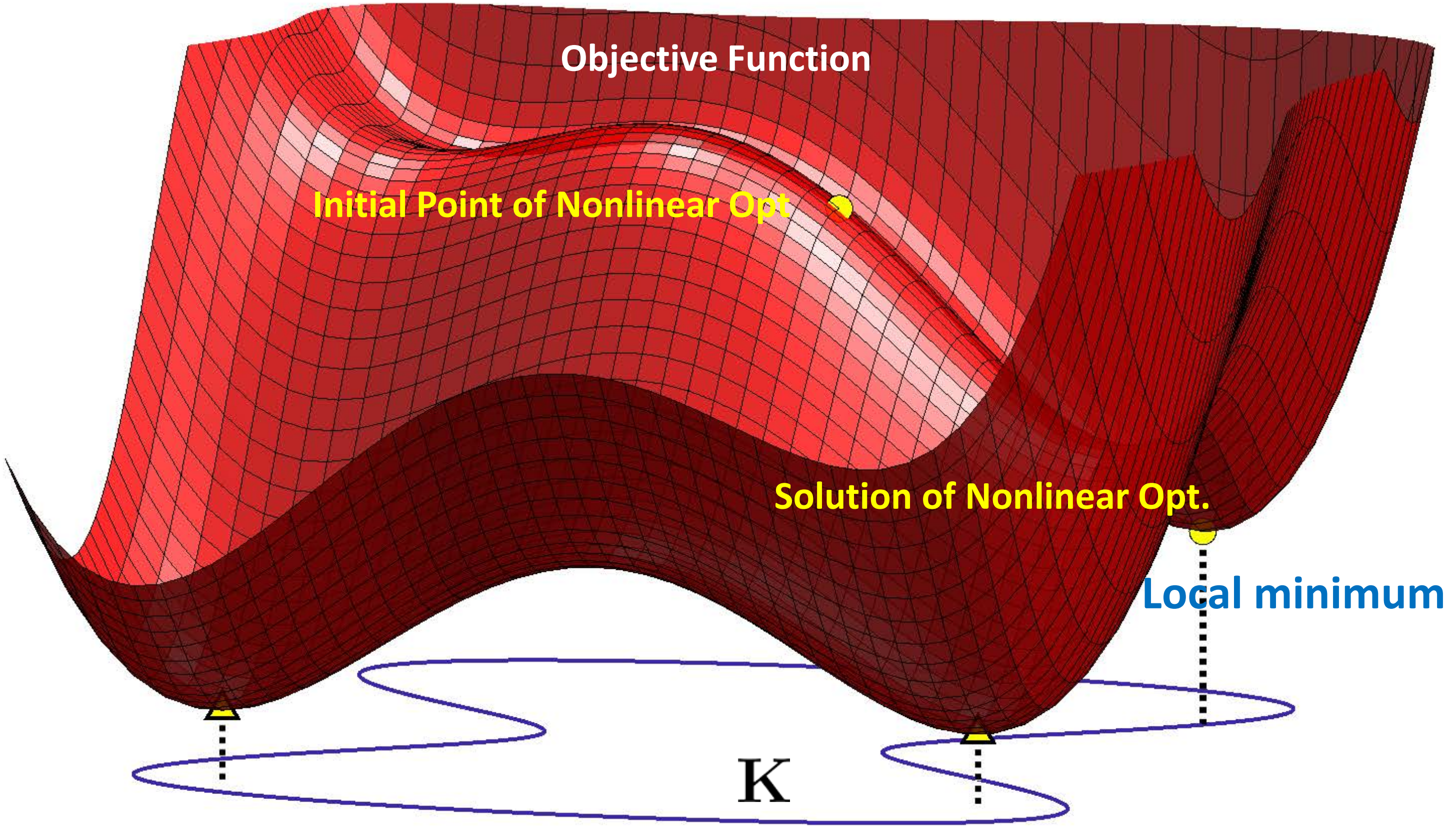
- What is the cost of convexification?

Nonlinear Optimization: variables (x_1, x_2)

$$\begin{aligned} \mathrm{P}^* = &\underset{x\in\mathbb{R}^2}{\text{minimize}} \quad \frac{1}{3}x_1^6 - \frac{21}{10}x_1^4 + 4x_1^2 + x_1x_2 + 4x_2^4 - 4x_2^2 + \frac{3}{2} \\ &\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^2 : -\frac{1}{16}x_1^4 + \frac{1}{4}x_1^3 - \frac{1}{4}x_1^2 - \frac{9}{100}x_2^2 + \frac{29}{400} \geq 0\} \end{aligned}$$

Interior-point method

⇩



Nonlinear Optimization: variables (x_1, x_2)

$$\mathrm{P}^* = \underset{x \in \mathbb{R}^2}{\text{minimize}} \quad \frac{1}{3}x_1^6 - \frac{21}{10}x_1^4 + 4x_1^2 + x_1x_2 + 4x_2^4 - 4x_2^2 + \frac{3}{2}$$

$$\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^2 : -\frac{1}{16}x_1^4 + \frac{1}{4}x_1^3 - \frac{1}{4}x_1^2 - \frac{9}{100}x_2^2 + \frac{29}{400} \geq 0\}$$

Interior-point method

⇩

Moment SDP: variables are moments $y_{\alpha_1\alpha_2} = \mathrm{E}[x_1^{\alpha_1}x_2^{\alpha_2}]$ $y = [y_\alpha, \alpha = 0, \ldots, 6]$

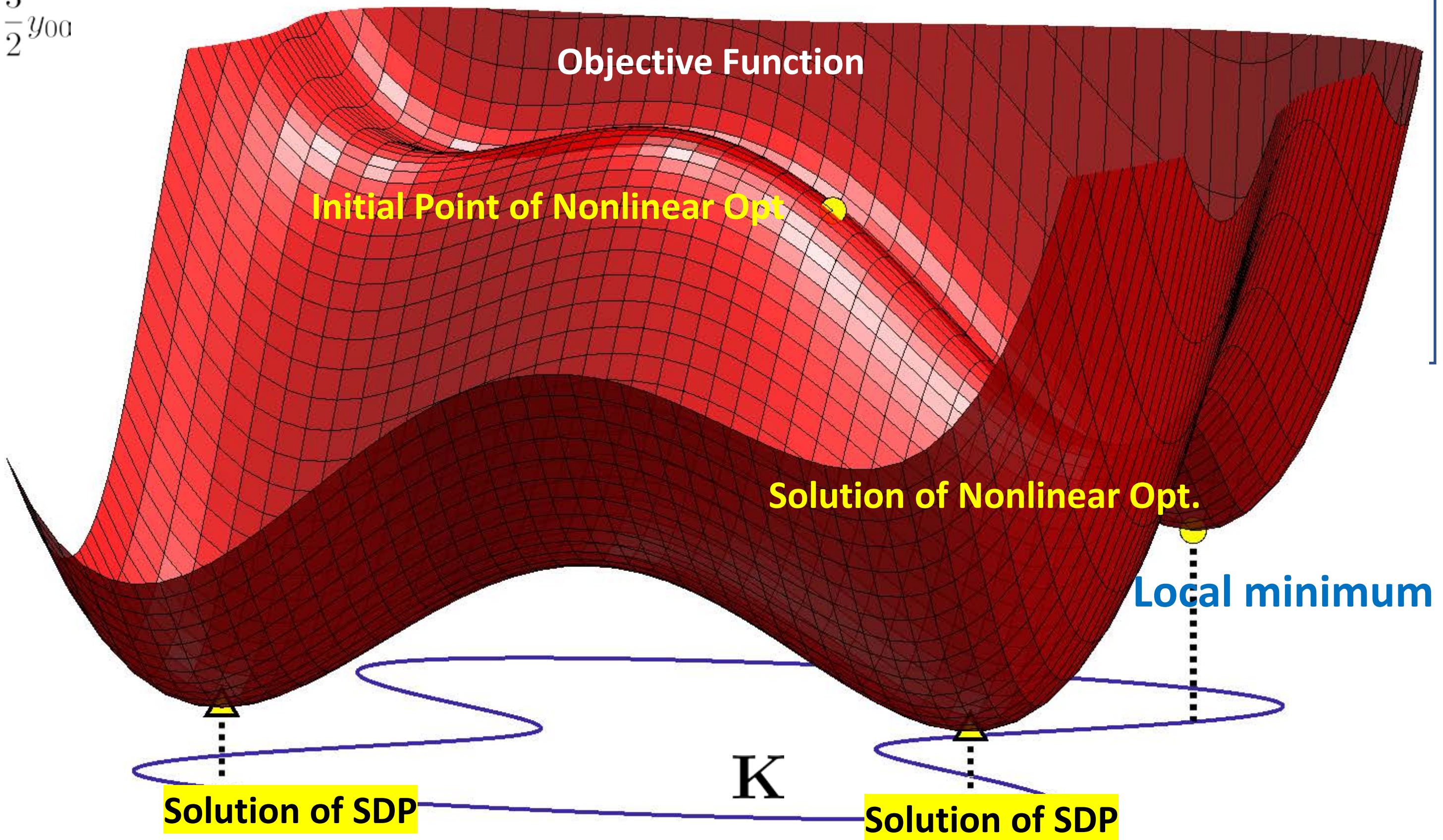
$$\mathrm{P}^{*3}_{mom} = \underset{y}{\text{minimize}} \quad \frac{1}{3}y_{60} - \frac{21}{10}y_{40} + 4y_{20} + y_{11} + 4y_{04} - 4y_{02} + \frac{3}{2}y_{00}$$

$$\text{subject to} \quad y_{00} = 1$$

$$\mathbf{M}_3(y) \succcurlyeq 0, \; \mathbf{M}_{3-2}(gy) \succcurlyeq 0$$

- Number of Moments in $\mathbb{R}^n$ up to order $2d$:

$$\binom{n+2d}{n} = \frac{(2+6)!}{2!6!} = 28$$



Nonlinear Optimization: variables (x_1, x_2)

$$\mathrm{P}^* = \underset{x\in\mathbb{R}^2}{\text{minimize}} \quad \frac{1}{3}x_1^6 - \frac{21}{10}x_1^4 + 4x_1^2 + x_1x_2 + 4x_2^4 - 4x_2^2 + \frac{3}{2}$$

$$\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^2 : -\frac{1}{16}x_1^4 + \frac{1}{4}x_1^3 - \frac{1}{4}x_1^2 - \frac{9}{100}x_2^2 + \frac{29}{400} \geq 0\}$$

Interior-point method

Moment SDP: variables are moments $y_{\alpha_1\alpha_2} = \mathrm{E}\left[x_1^{\alpha_1}x_2^{\alpha_2}\right]$ $y = [y_\alpha, \alpha = 0, \dots, 6]$

$$\mathrm{P}^{*3}_{mom} = \underset{y}{\text{minimize}} \quad \frac{1}{3}y_{60} - \frac{21}{10}y_{40} + 4y_{20} + y_{11} + 4y_{04} - 4y_{02} + \frac{3}{2}y_{00}$$

$$\text{subject to} \quad y_{00} = 1$$

$$\mathbf{M}_3(y) \succcurlyeq 0, \ \mathbf{M}_{3-2}(gy) \succcurlyeq 0$$

- Number of Moments in $\mathbb{R}^n$ up to order $2d$:

$$\binom{n+2d}{n} = \frac{(2+6)!}{2!6!} = 28$$

SOS SDP: variables are coefficients of polynomial

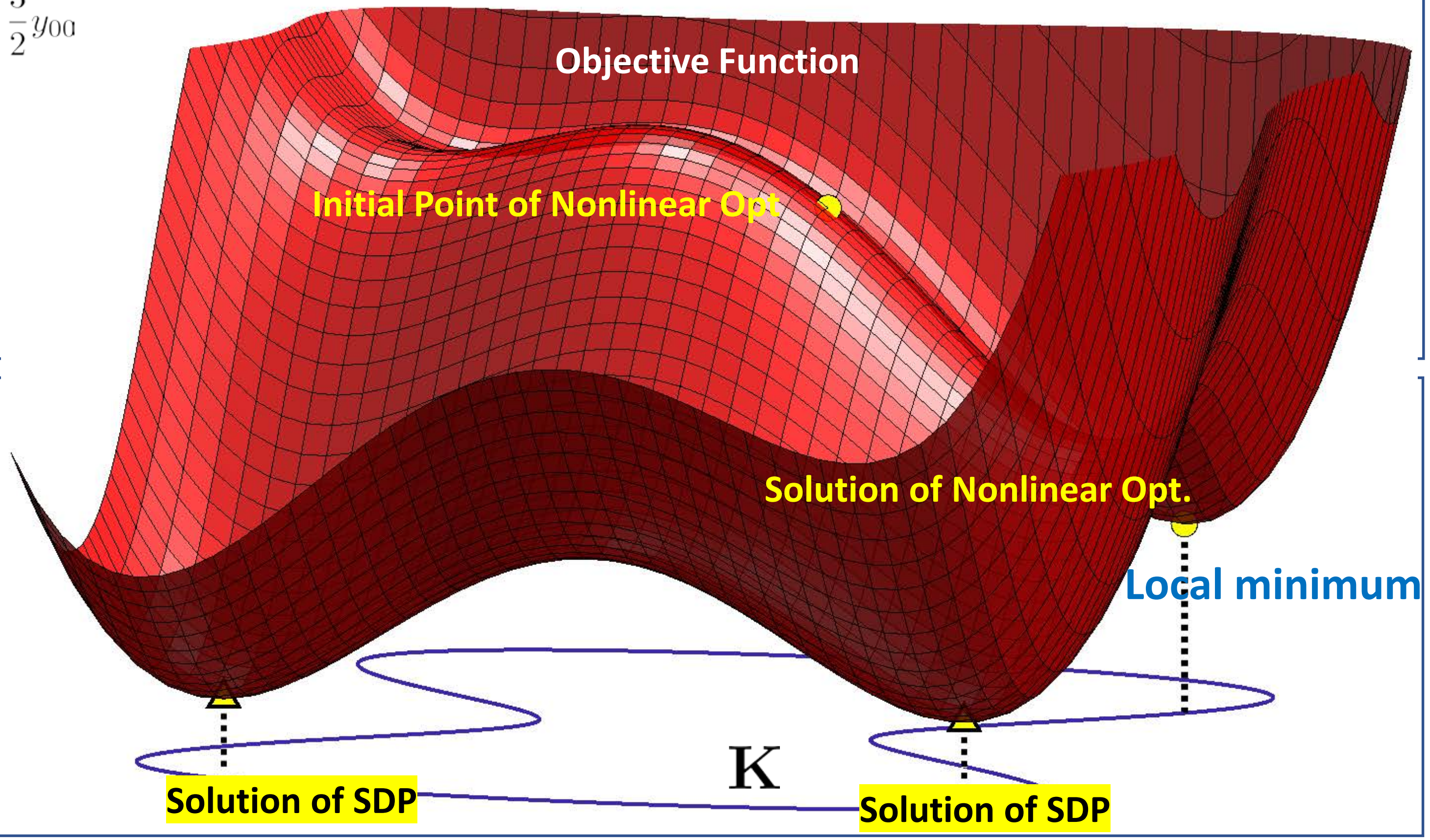
$$\mathrm{P}^{*3}_{sos} = \underset{\gamma\in\mathbb{R},\sigma_1}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad p(x) - \gamma - \sigma_1(x)g(x) \in SOS_6$$

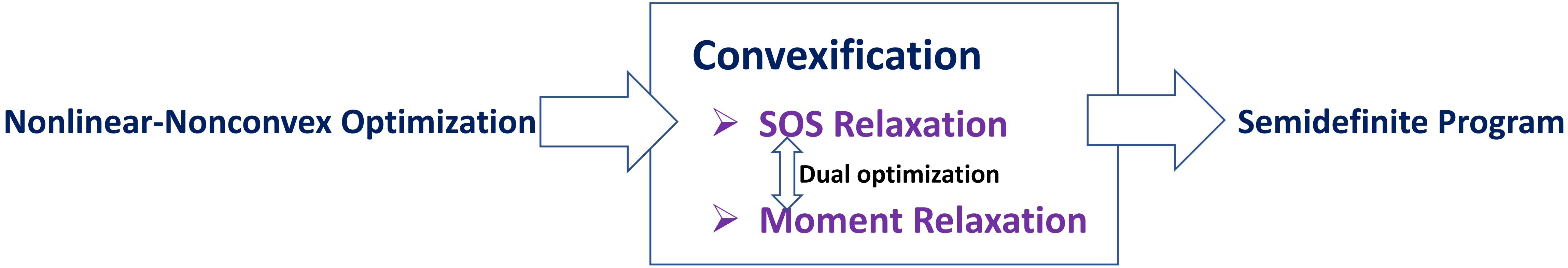
$$\sigma_1(x) \in SOS_2$$

- Number of coefficients of a $2d$-degree polynomial in $\mathbb{R}^n$:

$$\binom{n+2d}{n} = \frac{(2+6)!}{2!6!} = 28$$



Moment-SOS Relaxations



Nonlinear-Nonconvex Optimization

Convexification

- **SOS Relaxation**

Dual optimization

- **Moment Relaxation**

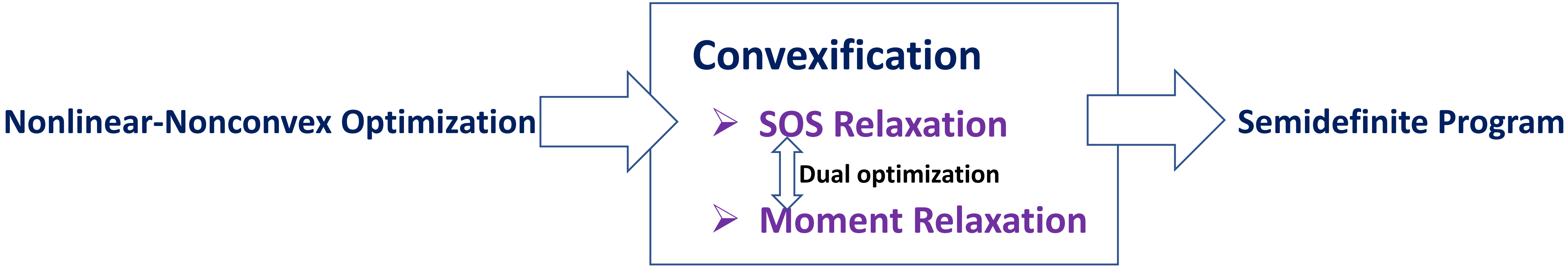
Semidefinite Program

- What is the cost of convexification ?

Convexification increases the **dimension** of the search space.

  - Number of variables of the original nonlinear optimization: n
  - Number of variables Moment SDP: $\binom{n+2d}{n}$

Moment-SOS Relaxations



Nonlinear-Nonconvex Optimization

Convexification

- **SOS Relaxation**

Dual optimization

- **Moment Relaxation**

Semidefinite Program

- What is the cost of convexification ?

Convexification increases the **dimension** of the search space.

Cost of solving challenging problems

- Number of variables of the original nonlinear optimization: n
- Number of variables Moment SDP: $\binom{n+2d}{n}$

Pros:

- Moment-SOS relaxations solve difficult and challenging mathematical problems.
- They provide insights into challenging problems where no other solid and comprehensive approach exist.

(e.g., existing approaches for **nonlinear robust and chance constrained optimizations** work for particular class of problems,…).

Moment-SOS Relaxations

Large Scale Problems $\xrightarrow{\text{Moment-SOS Relaxations}}$ Large Scale Semidefinite Programs

- Current SDP solvers are interior-point based solvers.
- In the absence of problem structure, sum of squares problems are currently limited, roughly speaking, to a several thousands variables (variables in SDP).

- How to address large scale problems?

Moment-SOS Relaxations

➢ How to address large scale problems?

1) Modified SOS optimization to generate **i)** smaller SDP's or **ii)** other types of convex constraints like LP.

Approaches:

- i) (Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS),
- ii) Bounded degree SOS (BSOS)

Moment-SOS Relaxations

➢ How to address large scale problems?

1) Modified SOS optimization to generate **i)** smaller SDP’s or **ii)** other types of convex constraints like LP.

Approaches:

- i) (Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS),
- ii) Bounded degree SOS (BSOS)

2) Take advantage of structure of the problem (sparsity) to generate **smaller SDP’s.**

Approaches:

- i) Spars Sum-of-Squares Optimization (SSOS)
- ii) Sparse Bounded Degree Sum-of-Squares Optimization (SBSOS)

Moment-SOS Relaxations

➢ How to address large scale problems?

1) Modified SOS optimization to generate **i)** smaller SDP’s or **ii)** other types of convex constraints like LP.

Approaches:

- i) (Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS),
- ii) Bounded degree SOS (BSOS)

2) Take advantage of structure of the problem (sparsity) to generate **smaller SDP’s.**

Approaches:

- i) Spars Sum-of-Squares Optimization (SSOS)
- ii) Sparse Bounded Degree Sum-of-Squares Optimization (SBSOS)

3) Efficient Algorithms for Large Scale SDP’s (Lecture 9)

Moment-SOS Relaxations

- How to address large scale problems?

1) Modified SOS optimization to generate **i)** smaller SDP’s or **ii)** other types of convex constraints like LP.

Approaches:
  i) (Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS),
  ii) Bounded degree SOS (BSOS)

2) Take advantage of structure of the problem (sparsity) to generate **smaller SDP’s.**

Approaches:
  i) Spars Sum-of-Squares Optimization (SSOS)
  ii) Sparse Bounded Degree Sum-of-Squares Optimization (SBSOS)

3) Efficient Algorithms for Large Scale SDP’s (Lecture 9)

4) Reformulate original optimization problem to reduce the size of the optimization (Lectures 10 and 11)

Topics

1) (Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS)

Modified SOS optimization that results in LP and Second order cone program

6-link pendulum

Atlas

Applications:

Control and analyze of high dimensional systems

- A. A. Ahmadi and A. Majumdar," DSOS and SDSOS Optimization: More Tractable Alternatives to Sum of Squares and Semidefinite Optimization", *SIAM Journal on Applied Algebraic Geometry, 2019.*

Topics

1) (Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS)
   Modified SOS optimization that results in LP and Second order cone program

2) Bounded degree SOS (BSOS)
   Modified SOS optimization that results in smaller SDP’s.

- Jean B. Lasserre, Kim-Chuan Toh, Shouguang Yang, “A bounded degree SOS hierarchy for polynomial optimization”, EURO Journal on Computational Optimization March 2017, Volume 5, Issue 1–2, pp 87–117

Topics

1) (Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS)
    Modified SOS optimization that results in LP and Second order cone program

2) Bounded degree SOS (BSOS)
    Modified SOS optimization that results in smaller SDP's.

3) Sparse Sum-of-Squares Optimization (SSOS)
    Takes advantage of sparsity of the original problem to generate smaller SDP.

- H. Waki, S. Kim, M. Kojima, and M. Muramatsu, "Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity," SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

- Zheng, Y., Fantuzzi, G., & Papachristodoulou, A. "Sparse sum-of-squares (SOS) optimization: A bridge between DSOS/SDSOS and SOS optimization for sparse polynomials", arXiv preprint arXiv:1807.05463. 2018

Topics

1) (Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS)
    Modified SOS optimization that results in LP and Second order cone program

2) Bounded degree SOS (BSOS)
    Modified SOS optimization that results in smaller SDP's.

3) Sparse Sum-of-Squares Optimization (SSOS)
    Takes advantage of sparsity of the original problem to generate smaller SDP.

4) Sparse Bounded Degree Sum-of-Squares Optimization (SBSOS)
    Combination of 2 and 3

- Tillmann Weisser, Jean B. Lasserre, Kim-Chuan Toh., "Sparse-BSOS: a bounded degree SOS hierarchy for large scale polynomial optimization with sparsity", Math. Prog. Comp. (2018) 10:1–32

Topics

1) (Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS)
   Modified SOS optimization that results in LP and Second order cone program

2) Bounded degree SOS (BSOS)
   Modified SOS optimization that results in smaller SDP's.

3) Sparse Sum-of-Squares Optimization (SSOS)
   Takes advantage of sparsity of the original problem to generate smaller SDP.

4) Sparse Bounded Degree Sum-of-Squares Optimization (SBSOS)
   Combination of 2 and 3

(Scaled) Diagonally-Dominant SOS Optimization (DSOS, SDSOS)

A. A. Ahmadi and A. Majumdar,” DSOS and SDSOS Optimization: More Tractable Alternatives to Sum of Squares and Semidefinite Optimization”, *SIAM Journal on Applied Algebraic Geometry, 2019.*

Nonlinear Optimization and Nonnegative polynomials

Unconstrained Optimization:

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad p(x)$$

$$\Rightarrow$$

$$\underset{\gamma \in \mathbb{R}}{\text{maximize}} \quad \gamma \longrightarrow \text{linear function}$$

$$\text{subject to} \quad \boxed{p(x) - \gamma \geq 0, \quad \forall x \in \mathbb{R}^n}$$

Polynomial Nonnegativity Constraint

$\Downarrow$

Replace with convex constraints

Convex optimization

$p(x) \in \mathbb{R}[x]$

Constrained Optimization:

$$\begin{aligned} &\underset{x \in \mathbb{R}^n}{\text{minimize}} && p(x) \\ &\text{subject to} && g_i(x) \geq 0, \ i = 1, ..., m \end{aligned}$$

$$\Rightarrow$$

$$\underset{\gamma \in \mathbb{R}}{\text{maximize}} \quad \gamma \longrightarrow \text{linear function}$$

$$\text{subject to} \quad \boxed{p(x) - \gamma \geq 0, \quad \forall x \in \{x \in \mathbb{R}^n : g_i(x) \geq 0, \ i = 1, ..., m\}}$$

Polynomial Nonnegativity Constraint

$\Downarrow$

Replace with convex constraints

Convex optimization

$p(x), g_i(x) \in \mathbb{R}[x], \ i = 1, ..., m$

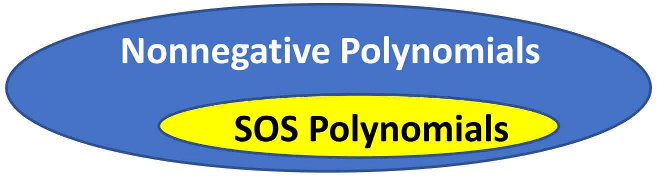
Sum of squares Polynomials

Polynomial $p(x)$ is **sum of squares** (SOS) polynomial if :

it can be written as a finite sum of squares of other polynomials.

$$p(x) \in \mathbb{R}[x] \quad \xrightarrow{\text{SOS}} \quad p(x) = \textstyle\sum_{i=1}^{\ell} h_i^2(x) \qquad h_i(x) \in \mathbb{R}[x], \quad i = 1, \dots, \ell$$

- If polynomial $p(x)$ is **SOS**, then it is $\boldsymbol{p(x) \geq 0}$ for all



PSD Matrix representation of **SOS polynomials**

$$p(x) = B(x)^T Q B(x) \qquad \underbrace{Q \in \mathcal{S}^n, \quad Q \succcurlyeq 0}_{\text{PSD Matrix}}$$

where $B(x)$:vector of monomials in x

Sum of squares Polynomials

$$p(x) \in SOS \Longleftrightarrow \underbrace{Q \in \mathcal{S}^n, \quad Q \succeq 0}_{\text{PSD Matrix}} \Longrightarrow \text{Nonnegative Eigenvalues} \Longrightarrow \text{SDP}$$

Sum of squares Polynomials

$$p(x) \in SOS \Longleftrightarrow \underbrace{Q \in \mathcal{S}^n, \quad Q \succcurlyeq 0}_{\text{PSD Matrix}} \Longrightarrow \text{Nonnegative Eigenvalues} \Longrightarrow \text{SDP}$$

- To **avoid SDP** and obtain *computationally cheap* convex optimizations, we obtain **relaxed condition** for PSD matrices.

  - For this, we use the following Results:

    1) Gershgorin Circle Theorem
    2) Diagonally Dominant Matrix (dd)

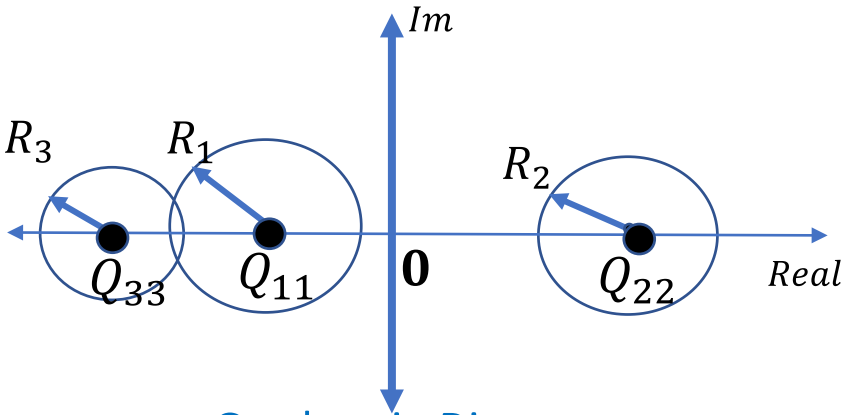
Gershgorin Circle Theorem

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

$$Disk_1(Q_{11}, R_1 = |Q_{12}| + |Q_{13}|)$$

$$Disk_2(Q_{22}, R_2 = |Q_{21}| + |Q_{23}|)$$

$$Disk_3(Q_{33}, R_3 = |Q_{31}| + |Q_{32}|)$$



Gershgorin Discs

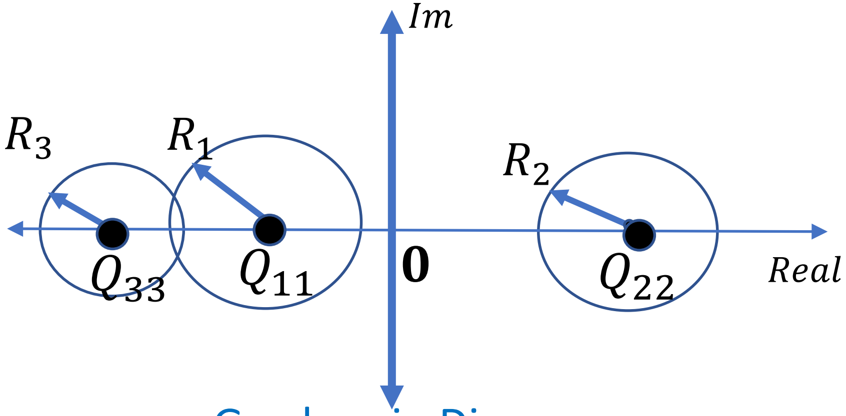
Gershgorin Circle Theorem

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

$$Disk_1(Q_{11}, R_1 = |Q_{12}| + |Q_{13}|)$$

$$Disk_2(Q_{22}, R_2 = |Q_{21}| + |Q_{23}|)$$

$$Disk_3(Q_{33}, R_3 = |Q_{31}| + |Q_{32}|)$$



Gershgorin Discs

- Eigenvalue of Q lies within the Gershgorin discs.

$$Q \in \mathbb{R}^{n\times n} \Rightarrow \text{Eigenvalues} \in \cup_{i=1}^{n} Disk_i(Q_{ii}, R_i = \sum_{i\neq j} |Q_{ij}|), \quad i = 1, \ldots, n$$

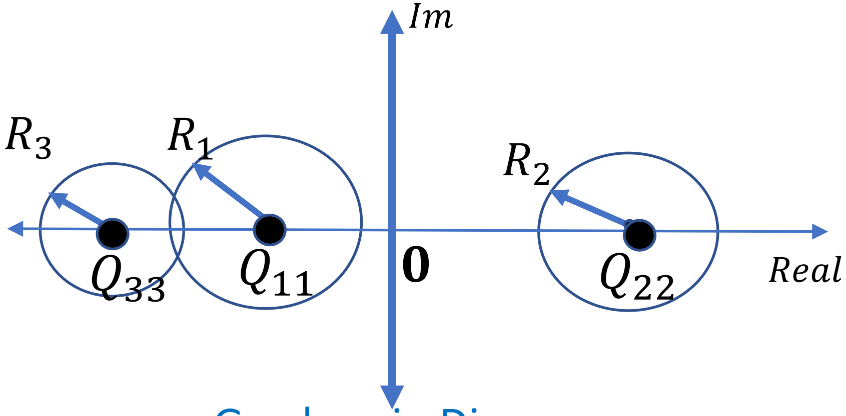
Gershgorin Circle Theorem

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

$$Disk_1(Q_{11}, R_1 = |Q_{12}| + |Q_{13}|)$$

$$Disk_2(Q_{22}, R_2 = |Q_{21}| + |Q_{23}|)$$

$$Disk_3(Q_{33}, R_3 = |Q_{31}| + |Q_{32}|)$$



Gershgorin Discs

- Eigenvalue of Q lies within the Gershgorin discs.

$$Q \in \mathbb{R}^{n\times n} \Rightarrow \text{Eigenvalues} \in \cup_{i=1}^{n} Disk_i(Q_{ii}, R_i = \textstyle\sum_{i\neq j} |Q_{ij}|), \quad i = 1, \ldots, n$$

- We use Gershgorin Circle Theorem to obtain relaxed PSD condition in terms of entries of matrices.

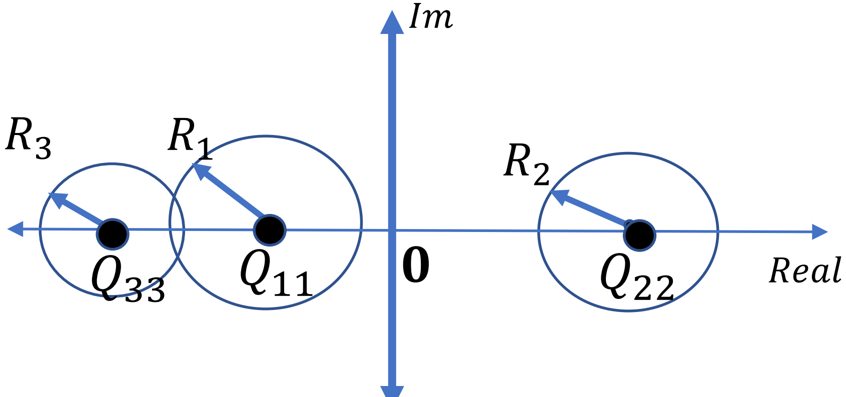
Gershgorin Circle Theorem

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

$$Disk_1(Q_{11}, R_1 = |Q_{12}| + |Q_{13}|)$$

$$Disk_2(Q_{22}, R_2 = |Q_{21}| + |Q_{23}|)$$

$$Disk_3(Q_{33}, R_3 = |Q_{31}| + |Q_{32}|)$$



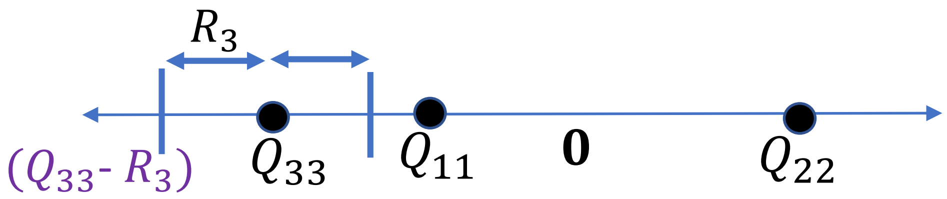
Gershgorin Discs

- ➢ Eigenvalue of Q lies within the Gershgorin discs.

$$Q \in \mathbb{R}^{n\times n} \quad \Rightarrow \quad \text{Eigenvalues} \in \ \cup_{i=1}^{n} Disk_i(Q_{ii}, R_i = \textstyle\sum_{i\neq j} |Q_{ij}|), \quad i = 1, \ldots, n$$

- We use Gershgorin Circle Theorem to obtain relaxed PSD condition in terms of entries of matrices.

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{12} & Q_{22} & Q_{23} \\ Q_{13} & Q_{23} & Q_{33} \end{bmatrix} \in \mathcal{S}^3$$



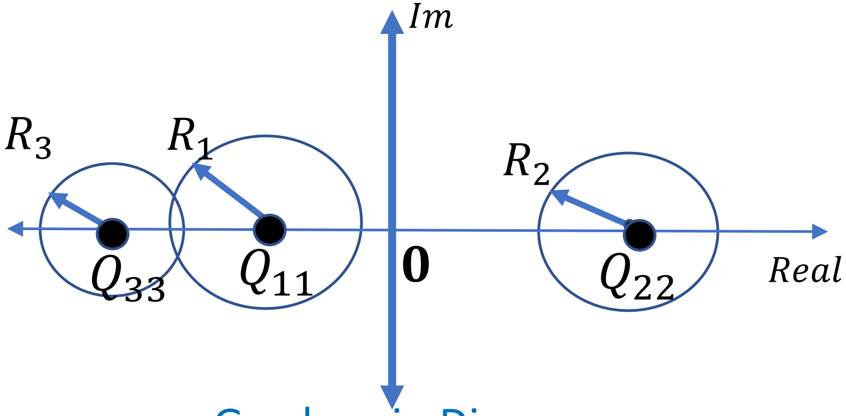
Gershgorin Circle Theorem

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

$$Disk_1(Q_{11}, R_1 = |Q_{12}| + |Q_{13}|)$$

$$Disk_2(Q_{22}, R_2 = |Q_{21}| + |Q_{23}|)$$

$$Disk_3(Q_{33}, R_3 = |Q_{31}| + |Q_{32}|)$$



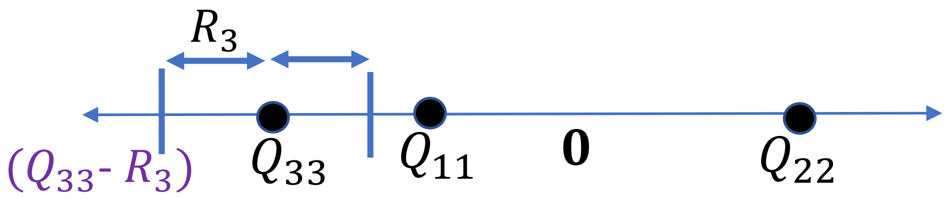
Gershgorin Discs

- ➢ Eigenvalue of Q lies within the Gershgorin discs.

$$Q \in \mathbb{R}^{n\times n} \quad \Longrightarrow \quad \text{Eigenvalues} \in \ \cup_{i=1}^{n} Disk_i(Q_{ii}, R_i = \textstyle\sum_{i\neq j} |Q_{ij}|), \quad i = 1, \ldots, n$$

- We use Gershgorin Circle Theorem to obtain relaxed PSD condition in terms of entries of matrices.

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{12} & Q_{22} & Q_{23} \\ Q_{13} & Q_{23} & Q_{33} \end{bmatrix} \in \mathcal{S}^3$$



$$\text{Smallest Eigenvalue} \geq \min_{i=1,2,3} \ (Q_{ii} - R_i)$$

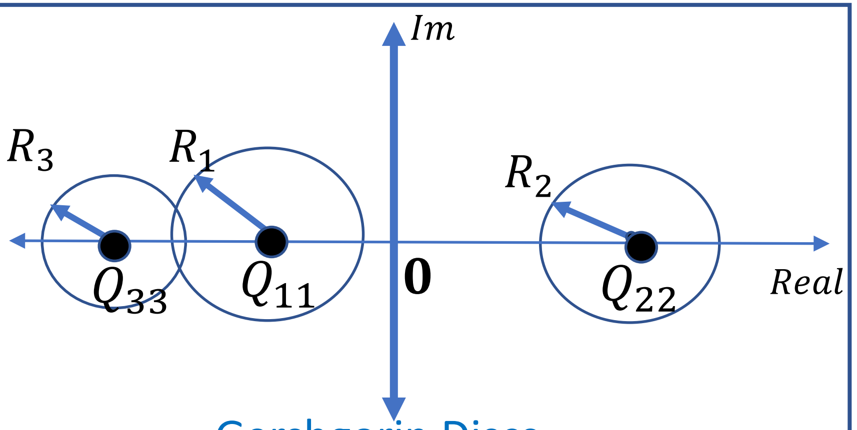
Gershgorin Circle Theorem

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \in \mathbb{R}^{3\times3}$$

$$Disk_1(Q_{11}, R_1 = |Q_{12}| + |Q_{13}|)$$

$$Disk_2(Q_{22}, R_2 = |Q_{21}| + |Q_{23}|)$$

$$Disk_3(Q_{33}, R_3 = |Q_{31}| + |Q_{32}|)$$



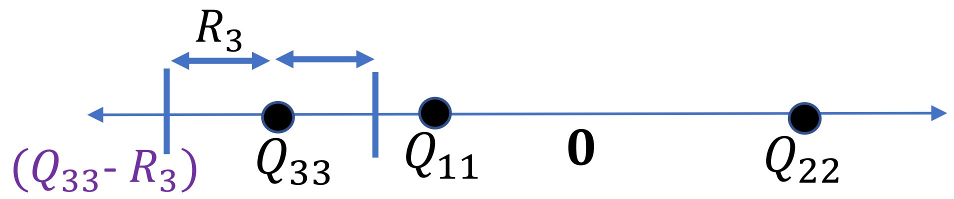
Gershgorin Discs

- ➢ Eigenvalue of Q lies within the Gershgorin discs.

$$Q \in \mathbb{R}^{n\times n} \quad \Rightarrow \quad \text{Eigenvalues} \in \ \cup_{i=1}^{n} Disk_i(Q_{ii}, R_i = \textstyle\sum_{i\neq j} |Q_{ij}|), \quad i = 1, \ldots, n$$

- We use Gershgorin Circle Theorem to obtain relaxed PSD condition in terms of entries of matrices.

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{12} & Q_{22} & Q_{23} \\ Q_{13} & Q_{23} & Q_{33} \end{bmatrix} \in \mathcal{S}^3$$



$$\text{PSD} \rightarrow \text{Smallest Eigenvalue} \geq \min_{i=1,2,3}\ (Q_{ii} - R_i) \quad \geq 0$$

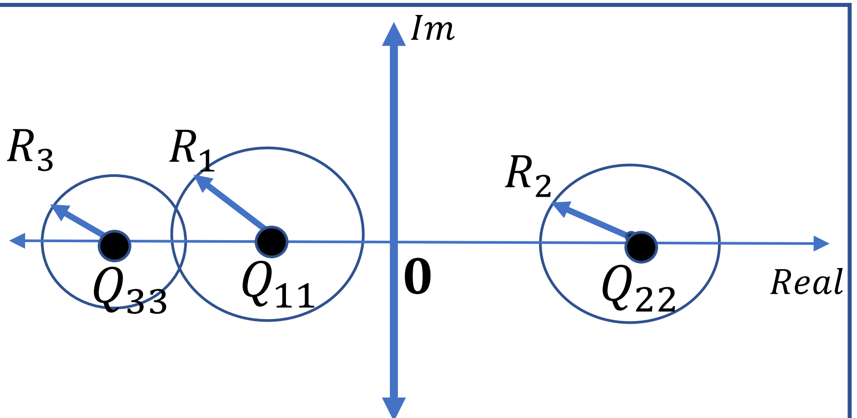
Gershgorin Circle Theorem

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

$$Disk_1(Q_{11}, R_1 = |Q_{12}| + |Q_{13}|)$$

$$Disk_2(Q_{22}, R_2 = |Q_{21}| + |Q_{23}|)$$

$$Disk_3(Q_{33}, R_3 = |Q_{31}| + |Q_{32}|)$$



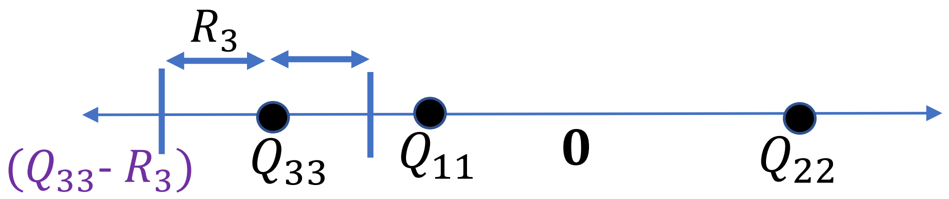
Gershgorin Discs

➢ Eigenvalue of Q lies within the Gershgorin discs.

$$Q \in \mathbb{R}^{n\times n} \Rightarrow \text{Eigenvalues} \in \cup_{i=1}^{n} Disk_i(Q_{ii}, R_i = \textstyle\sum_{i\neq j} |Q_{ij}|), \quad i = 1, \dots, n$$

- We use Gershgorin Circle Theorem to obtain relaxed PSD condition in terms of entries of matrices.

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{12} & Q_{22} & Q_{23} \\ Q_{13} & Q_{23} & Q_{33} \end{bmatrix} \in \mathcal{S}^3$$

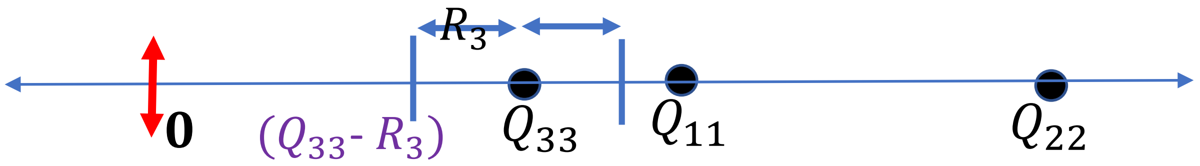


$$Q_{11} \geq R_1 = |Q_{12}| + |Q_{13}|$$

$$Q_{22} \geq R_2 = |Q_{12}| + |Q_{23}|$$

$$Q_{33} \geq R_3 = |Q_{13}| + |Q_{23}|$$

$$\text{PSD} \rightarrow \text{Smallest Eigenvalue} \geq \min_{i=1,2,3}\ (Q_{ii} - R_i) \quad \geq 0$$



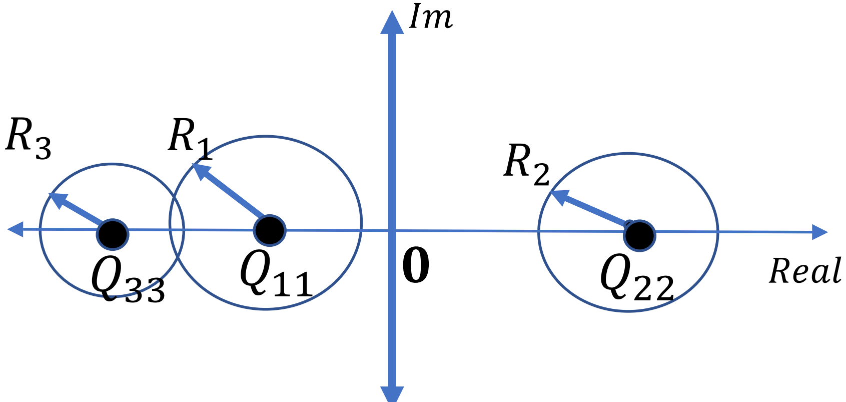
Gershgorin Circle Theorem

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

$$Disk_1(Q_{11}, R_1 = |Q_{12}| + |Q_{13}|)$$

$$Disk_2(Q_{22}, R_2 = |Q_{21}| + |Q_{23}|)$$

$$Disk_3(Q_{33}, R_3 = |Q_{31}| + |Q_{32}|)$$



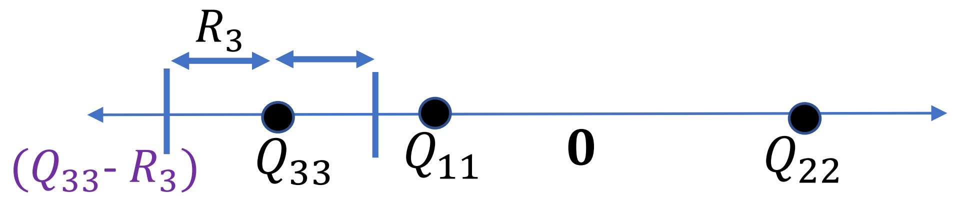
Gershgorin Discs

- Eigenvalue of Q lies within the Gershgorin discs.

$$Q \in \mathbb{R}^{n\times n} \quad \Rightarrow \quad \text{Eigenvalues} \in \ \cup_{i=1}^{n} Disk_i(Q_{ii}, R_i = \textstyle\sum_{i\neq j} |Q_{ij}|), \quad i = 1, \ldots, n$$

- We use Gershgorin Circle Theorem to obtain relaxed PSD condition in terms of entries of matrices.

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{12} & Q_{22} & Q_{23} \\ Q_{13} & Q_{23} & Q_{33} \end{bmatrix} \in \mathcal{S}^3$$



$$\text{PSD} \rightarrow \text{Smallest Eigenvalue} \geq \min_{i=1,2,3} (Q_{ii} - R_i) \quad \geq 0$$

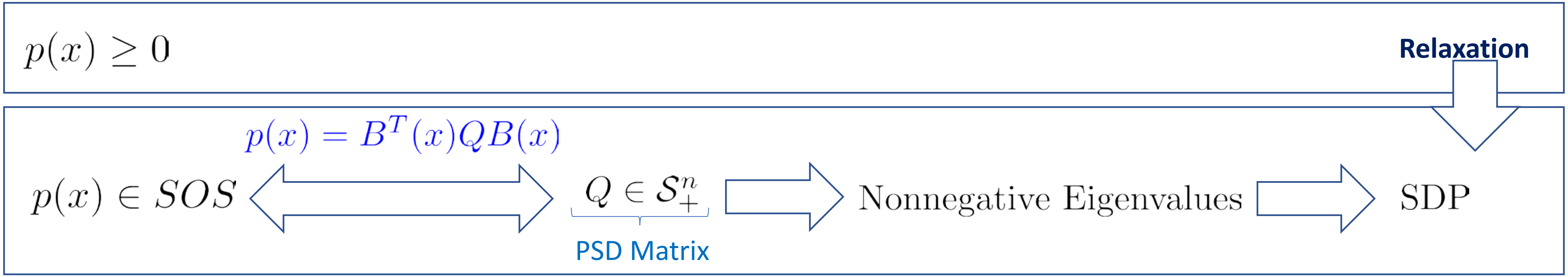
$$Q_{11} \geq R_1 = |Q_{12}| + |Q_{13}|$$

$$Q_{22} \geq R_2 = |Q_{12}| + |Q_{23}|$$

$$Q_{33} \geq R_3 = |Q_{13}| + |Q_{23}|$$

Diagonally Dominant Matrix (dd): $Q \in \mathcal{S}^n \qquad Q_{ii} \geq \sum_{i\neq j} |Q_{ij}|, \quad i = 1, \ldots, n \qquad Q \in \mathcal{S}^n_{dd} \subset \mathcal{S}^n_+$

Nonnegative Polynomials



$p(x) \geq 0$

Relaxation

$p(x) = B^T(x)QB(x)$

$$p(x) \in SOS \Longleftrightarrow \underbrace{Q \in \mathcal{S}_+^n}_{\text{PSD Matrix}} \Longrightarrow \text{Nonnegative Eigenvalues} \Longrightarrow \text{SDP}$$

A. A. Ahmadi and A. Majumdar," DSOS and SDSOS Optimization: More Tractable Alternatives to Sum of Squares and Semidefinite Optimization", *SIAM Journal on Applied Algebraic Geometry, 2019.*

Nonnegative Polynomials

$$p(x) \geq 0$$

Relaxation

$$p(x) = B^T(x)QB(x)$$

$$p(x) \in SOS \Longleftrightarrow \underbrace{Q \in \mathcal{S}^n_+}_{\text{PSD Matrix}} \Longrightarrow \text{Nonnegative Eigenvalues} \Longrightarrow \text{SDP}$$

Relaxation

$$\underbrace{Q \in \mathcal{S}^n_{dd}}_{\text{Diagonally Dominant Matrix}} \Longrightarrow Q_{ii} \geq \textstyle\sum_{i \neq j} |Q_{ij}|, \quad i = 1, \ldots, n$$

A. A. Ahmadi and A. Majumdar," DSOS and SDSOS Optimization: More Tractable Alternatives to Sum of Squares and Semidefinite Optimization", *SIAM Journal on Applied Algebraic Geometry, 2019.*

Nonnegative Polynomials

$p(x) \geq 0$

Relaxation

$p(x) = B^T(x)QB(x)$

$$p(x) \in SOS \Longleftrightarrow \underbrace{Q \in \mathcal{S}^n_+}_{\text{PSD Matrix}} \Rightarrow \text{Nonnegative Eigenvalues} \Rightarrow \text{SDP}$$

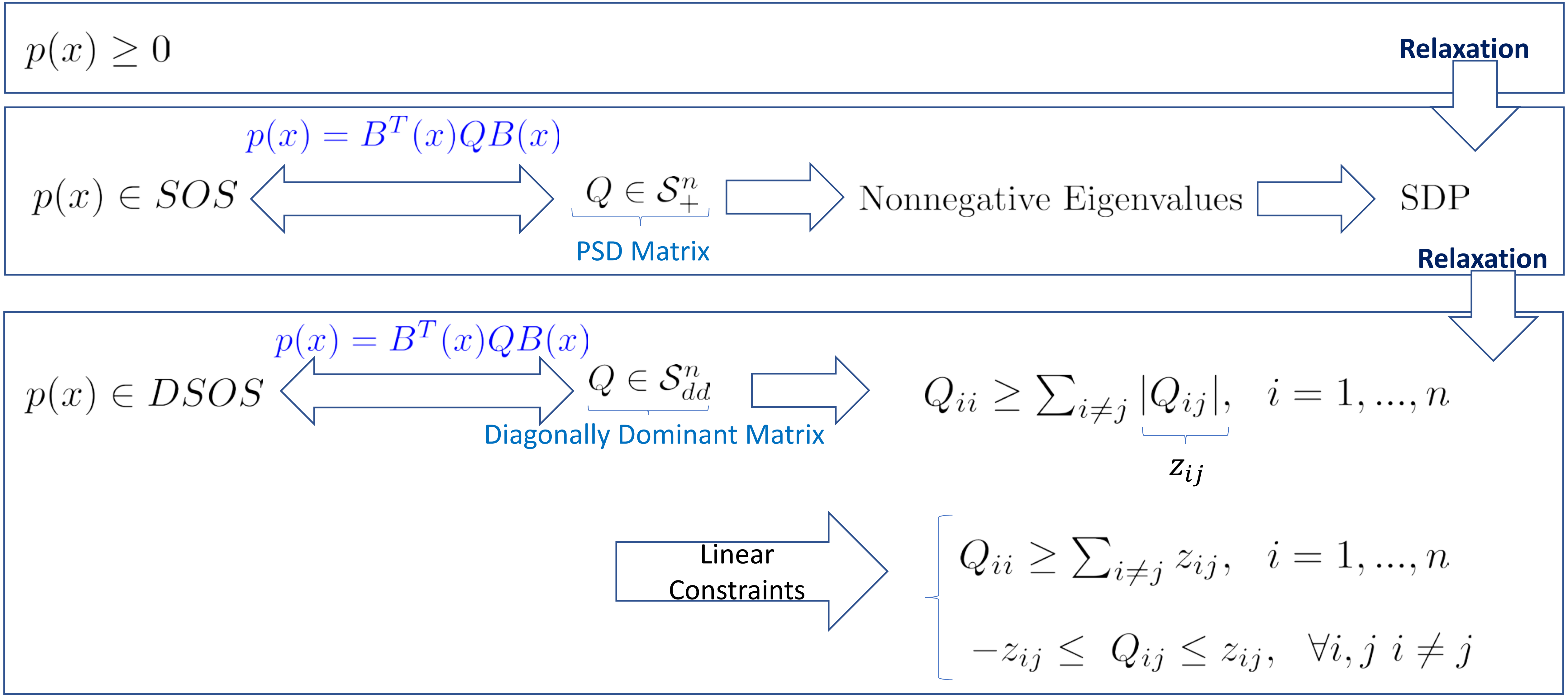
Relaxation

$$\underbrace{Q \in \mathcal{S}^n_{dd}}_{\text{Diagonally Dominant Matrix}} \Rightarrow Q_{ii} \geq \sum_{i \neq j} \underbrace{|Q_{ij}|}_{z_{ij}}, \quad i = 1, \ldots, n$$

$$\xrightarrow{\text{Linear Constraints}} \begin{cases} Q_{ii} \geq \sum_{i \neq j} z_{ij}, \quad i = 1, \ldots, n \\ -z_{ij} \leq Q_{ij} \leq z_{ij}, \quad \forall i, j \; i \neq j \end{cases}$$

A. A. Ahmadi and A. Majumdar,” DSOS and SDSOS Optimization: More Tractable Alternatives to Sum of Squares and Semidefinite Optimization”, *SIAM Journal on Applied Algebraic Geometry, 2019.*

Nonnegative Polynomials



$p(x) \geq 0$

Relaxation

$$p(x) \in SOS \Longleftrightarrow \underbrace{Q \in \mathcal{S}_+^n}_{\text{PSD Matrix}} \Longrightarrow \text{Nonnegative Eigenvalues} \Longrightarrow \text{SDP}$$

$p(x) = B^T(x) Q B(x)$

Relaxation

$$p(x) \in DSOS \Longleftrightarrow \underbrace{Q \in \mathcal{S}_{dd}^n}_{\text{Diagonally Dominant Matrix}} \Longrightarrow Q_{ii} \geq \sum_{i \neq j} \underbrace{|Q_{ij}|}_{z_{ij}}, \quad i = 1, \ldots, n$$

$p(x) = B^T(x) Q B(x)$

$$\xrightarrow{\text{Linear Constraints}} \begin{cases} Q_{ii} \geq \sum_{i \neq j} z_{ij}, \quad i = 1, \ldots, n \\ -z_{ij} \leq Q_{ij} \leq z_{ij}, \quad \forall i, j \; i \neq j \end{cases}$$

A. A. Ahmadi and A. Majumdar," DSOS and SDSOS Optimization: More Tractable Alternatives to Sum of Squares and Semidefinite Optimization", *SIAM Journal on Applied Algebraic Geometry, 2019.*

Nonnegative Polynomials

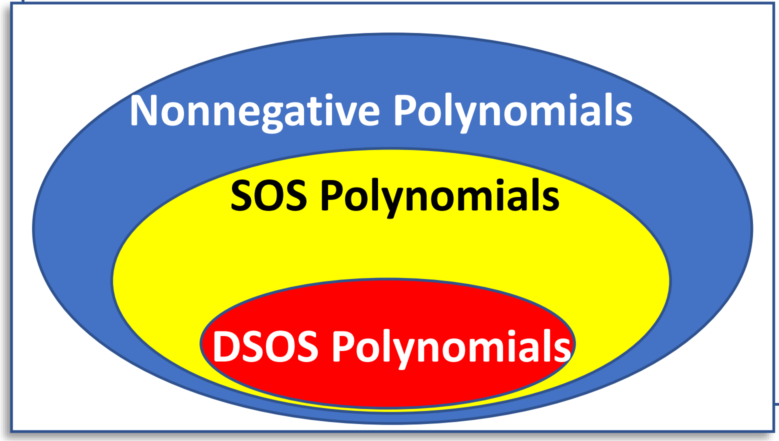
$$p(x) \geq 0$$

Relaxation

$$p(x) \in SOS \overset{p(x) = B^T(x)QB(x)}{\Longleftrightarrow} \underbrace{Q \in \mathcal{S}^n_+}_{\text{PSD Matrix}} \Longrightarrow \text{Nonnegative Eigenvalues} \Longrightarrow \text{SDP}$$

Relaxation

$$p(x) \in DSOS \overset{p(x) = B^T(x)QB(x)}{\Longleftrightarrow} \underbrace{Q \in \mathcal{S}^n_{dd}}_{\text{Diagonally Dominant Matrix}} \Longrightarrow Q_{ii} \geq \sum_{i \neq j} \underbrace{|Q_{ij}|}_{z_{ij}}, \quad i = 1, \dots, n$$



$$\xRightarrow{\text{Linear Constraints}} \begin{cases} Q_{ii} \geq \sum_{i \neq j} z_{ij}, & i = 1, \dots, n \\ -z_{ij} \leq Q_{ij} \leq z_{ij}, & \forall i, j \; i \neq j \end{cases}$$

A. A. Ahmadi and A. Majumdar," DSOS and SDSOS Optimization: More Tractable Alternatives to Sum of Squares and Semidefinite Optimization", *SIAM Journal on Applied Algebraic Geometry, 2019.*

Unconstrained optimization

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad p(x)$$

$$\begin{aligned} &\underset{\gamma\in\mathbb{R}}{\text{maximize}} && \gamma \\ &\text{subject to} && p(x)-\gamma \geq 0, \quad \forall x \in \mathbb{R}^n \end{aligned}$$

SOS Programing: SOS SDP

$$\begin{aligned} &\underset{Q\in\mathcal{S}^n,\gamma}{\text{maximize}} && \gamma \\ &\text{subject to} && p(x)-\gamma = \ B^T(x)QB(x) \\ &&& Q \in \mathcal{S}^n_+ \end{aligned}$$

DSOS Programing: Linear Program

$$\begin{aligned} &\underset{Q\in\mathcal{S}^n,\gamma}{\text{maximize}} && \gamma \\ &\text{subject to} && p(x)-\gamma = \ B^T(x)QB(x) \\ &&& Q \in \mathcal{S}^n_{dd} \end{aligned}$$

Unconstrained optimization

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad p(x)$$

$$\begin{aligned} &\underset{\gamma\in\mathbb{R}}{\text{maximize}} && \gamma \\ &\text{subject to} && p(x)-\gamma \geq 0, \quad \forall x \in \mathbb{R}^n \end{aligned}$$

SOS Programing: SOS SDP

$$\begin{aligned} &\underset{Q\in\mathcal{S}^n,\gamma}{\text{maximize}} && \gamma \\ &\text{subject to} && p(x)-\gamma = \ B^T(x)QB(x) \\ & && Q \in \mathcal{S}^n_+ \end{aligned}$$

DSOS Programing: Linear Program

$$\begin{aligned} &\underset{Q\in\mathcal{S}^n,\gamma}{\text{maximize}} && \gamma \\ &\text{subject to} && p(x)-\gamma = \ B^T(x)QB(x) \\ & && Q \in \mathcal{S}^n_{dd} \end{aligned}$$

Constrained optimization

$$\begin{aligned} &\underset{x\in\mathbb{R}^n}{\text{minimize}} && p(x) \\ &\text{subject to} && x \in \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \ i = 1, ..., m\} \end{aligned}$$

$$\begin{aligned} &\underset{\gamma\in\mathbb{R}}{\text{maximize}} && \gamma \\ &\text{subject to} && p(x)-\gamma \geq 0, \quad \forall x \in \mathbf{K} \end{aligned}$$

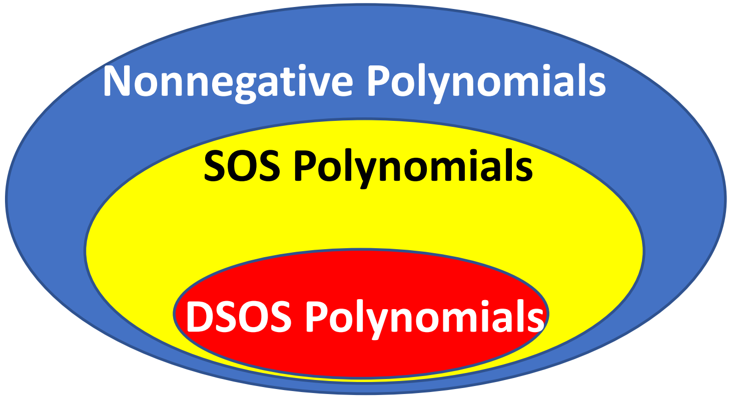
SOS Programing: SOS SDP

$$\begin{aligned} &\underset{\gamma,Q_i|_{i=0}^m}{\text{maximize}} && \gamma \\ &\text{subject to} && p(x)-\gamma-\textstyle\sum_{i=1}^m \sigma_i(x)g_i(x) = B_d(x)Q_0B_d(x) \\ & && \sigma_i(x) = B^T_{d_i}(x)Q_iB_{d_i}(x), \ i = 1, ..., m \\ & && Q_i \in \mathcal{S}^n_+, i = 0, ..., m \end{aligned}$$

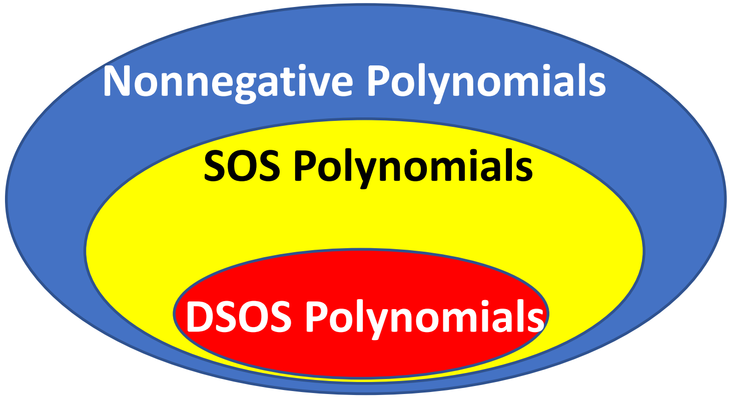
DSOS Programing: Linear Program

$$\begin{aligned} &\underset{\gamma,Q_i|_{i=0}^m}{\text{maximize}} && \gamma \\ &\text{subject to} && p(x)-\gamma-\textstyle\sum_{i=1}^m \sigma_i(x)g_i(x) = B_d(x)Q_0B_d(x) \\ & && \sigma_i(x) = B^T_{d_i}(x)Q_iB_{d_i}(x), \ i = 1, ..., m \\ & && Q_i \in \mathcal{S}^n_{dd}, i = 0, ..., m \end{aligned}$$

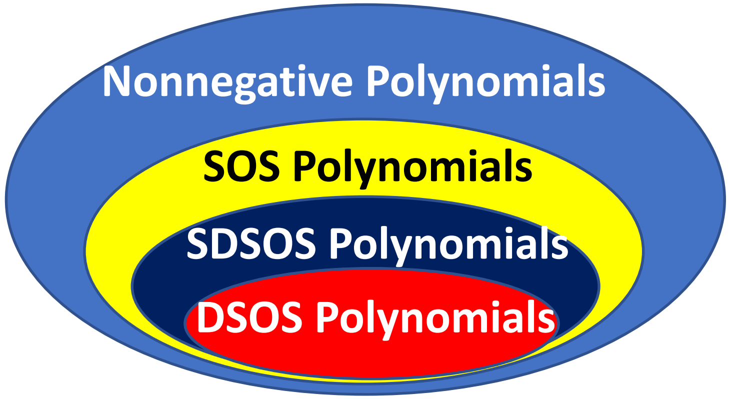
- DSOS programming searches a small subset of nonnegative polynomials set (conservative).



- DSOS programming searches a small subset of nonnegative polynomials set (conservative).



- To improve the results, we need to increase the search space.
- For this, we define “scaled-diagonally-dominant SOS” Polynomials (SDSOS).



Scaled Diagonally Dominant Matrix (sdd)

$Q \in \mathcal{S}^n$ is **sdd**, If there exist a diagonal matrix D with positive diagonal entries, such that DQD is **dd**.

Scaled Diagonally Dominant Matrix (sdd)

$Q \in \mathcal{S}^n$ is **sdd**, If there exist a diagonal matrix D with positive diagonal entries, such that DQD is **dd**.

$$Q = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 4 \end{bmatrix} \notin dd$$

$1 \geq |0| + |2|$ ☒

$3 \geq |0| + |0|$ ☑

$4 \geq |2| + |0|$ ☑

Scaled Diagonally Dominant Matrix (sdd)

$Q \in \mathcal{S}^n$ is **sdd**, If there exist a diagonal matrix D with positive diagonal entries, such that DQD is **dd**.

$$Q = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 4 \end{bmatrix} \notin dd$$

$1 \geq |0| + |2|$ ☒

$3 \geq |0| + |0|$ ☑

$4 \geq |2| + |0|$ ☑

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}}_{D \succcurlyeq 0} \underbrace{\begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 4 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}}_{D \succcurlyeq 0} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 1 \end{bmatrix} \in dd$$

$1 \geq |0| + |1|$ ☑

$3 \geq |0| + |0|$ ☑

$1 \geq |0| + |1|$ ☑

- ➢ Q is **sdd.**

Scaled Diagonally Dominant Matrix (sdd)

$Q \in \mathcal{S}^n$ is **sdd**, If there exist a diagonal matrix D with positive diagonal entries, such that DQD is **dd**.

$$Q = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 4 \end{bmatrix} \notin dd$$

$1 \geq |0| + |2|$ ☒

$3 \geq |0| + |0|$ ☑

$4 \geq |2| + |0|$ ☑

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}}_{D \succcurlyeq 0} \underbrace{\begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 4 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}}_{D \succcurlyeq 0} = \underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 1 \end{bmatrix}}_{\succcurlyeq 0} \in dd$$

$1 \geq |0| + |1|$ ☑

$3 \geq |0| + |0|$ ☑

$1 \geq |0| + |1|$ ☑

$$\boxed{\mathcal{S}^n_{dd} \subset \mathcal{S}^n_{sdd} \subset \mathcal{S}^n_{+}}$$

Every dd matrix is sdd matrix with $D = I$

- Q is **sdd.**

Scaled Diagonally Dominant Matrix (sdd)

To characterize the "sdd" matrices in terms of its element, we use the following result:

$Q \in \mathcal{S}^n$ is **sdd** if and only if it can be written as $Q = \sum_{i,j=1,\ldots,n,i<j} M^{ij}$

where, $M^{ij} \in \mathcal{S}^n$

Scaled Diagonally Dominant Matrix (sdd)

To characterize the "sdd" matrices in terms of its element, we use the following result:

$Q \in \mathcal{S}^n$ is **sdd** if and only if it can be written as $Q = \sum_{i,j=1,\ldots,n,i<j} M^{ij}$

where, $M^{ij} \in \mathcal{S}^n$ with zero every where except at most for 4 entries

$$(M^{ij})_{ii}, (M^{ij})_{ij}, (M^{ij})_{ji}, (M^{ij})_{jj}$$

which makes the 2 × 2 matrix $\begin{bmatrix} (M^{ij})_{ii} & (M^{ij})_{ij} \\ (M^{ij})_{ji} & (M^{ij})_{jj} \end{bmatrix}$ symmetric and positive semidefinite.

Scaled Diagonally Dominant Matrix (sdd)

To characterize the “sdd” matrices in terms of its element, we use the following result:

$Q \in \mathcal{S}^n$ is **sdd** if and only if it can be written as $Q = \sum_{i,j=1,\ldots,n,i<j} M^{ij}$

where, $M^{ij} \in \mathcal{S}^n$ with zero every where except at most for 4 entries

$$(M^{ij})_{ii}, (M^{ij})_{ij}, (M^{ij})_{ji}, (M^{ij})_{jj}$$

which makes the 2 × 2 matrix $\begin{bmatrix} (M^{ij})_{ii} & (M^{ij})_{ij} \\ (M^{ij})_{ji} & (M^{ij})_{jj} \end{bmatrix}$ symmetric and positive semidefinite.

Example: $Q = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 4 \end{bmatrix} \notin dd \ , \ \in sdd$ $\Rightarrow$ $Q = \sum_{i,j=1,2,3,i<j} M^{ij} = M^{12} + M^{13} + M^{23}$

$(M^{12})_{11}, (M^{12})_{12}, (M^{12})_{21}, (M^{12})_{22}$

$(M^{13})_{11}, (M^{13})_{13}, (M^{13})_{31}, (M^{13})_{33}$

$(M^{23})_{22}, (M^{23})_{23}, (M^{23})_{32}, (M^{23})_{33}$

$\Rightarrow$

$$Q = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 4 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1.5 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{M^{12}} + \underbrace{\begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 4 \end{bmatrix}}_{M^{13}} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1.5 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{M^{23}}$$

Scaled Diagonally Dominant Matrix (sdd)

To characterize the “sdd” matrices in terms of its element, we use the following result:

$Q \in \mathcal{S}^n$ is **sdd** if and only if it can be written as $Q = \sum_{i,j=1,\dots,n,i<j} M^{ij}$

where, $M^{ij} \in \mathcal{S}^n$ with zero every where except at most for 4 entries

$$(M^{ij})_{ii}, (M^{ij})_{ij}, (M^{ij})_{ji}, (M^{ij})_{jj}$$

which makes the 2 × 2 matrix $\begin{bmatrix} (M^{ij})_{ii} & (M^{ij})_{ij} \\ (M^{ij})_{ji} & (M^{ij})_{jj} \end{bmatrix}$ symmetric and positive semidefinite.

Example: $Q = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 4 \end{bmatrix} \notin dd \ , \ \in sdd \quad \Longrightarrow \quad Q = \sum_{i,j=1,2,3,i<j} M^{ij} = M^{12} + M^{13} + M^{23}$

$$Q = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1.5 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1.5 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\downarrow \qquad\qquad \downarrow \qquad\qquad \downarrow$$

$$\begin{bmatrix} 0 & 0 \\ 0 & 1.5 \end{bmatrix} \succcurlyeq 0 \qquad \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \succcurlyeq 0 \qquad \begin{bmatrix} 1.5 & 0 \\ 0 & 0 \end{bmatrix} \succcurlyeq 0$$

Scaled Diagonally Dominant Matrix (sdd)

$$\mathcal{S}^n_{dd} \subset \mathcal{S}^n_{sdd} \subset \mathcal{S}^n_{+}$$

Every dd matrix is sdd matrix with $D = I$

Every sdd matrix is sum of psd matrices M^{ij}

$$p(x) \in SOS \xLeftrightarrow{p(x) = B^T(x)QB(x)} \underbrace{Q \in \mathcal{S}^n_{+}}_{\text{PSD Matrix}} \Longrightarrow \text{SDP}$$

$$p(x) \in DSOS \xLeftrightarrow{p(x) = B^T(x)QB(x)} Q \in \mathcal{S}^n_{dd} \Longrightarrow \text{LP}$$

$$p(x) \in SDSOS \xLeftrightarrow{p(x) = B^T(x)QB(x)} Q \in \mathcal{S}^n_{sdd} \Longrightarrow \text{?}$$

Scaled Diagonally Dominant Matrix (sdd)

$Q \in \mathcal{S}^n$ is **sdd** if and only if it can be written as $Q = \sum_{i,j=1,\dots,n,i<j} M^{ij}$

where, $M^{ij} \in \mathcal{S}^n$ with zero every where except at most for 4 entries

$$(M^{ij})_{ii}, (M^{ij})_{ij}, (M^{ij})_{ji}, (M^{ij})_{jj}$$

which makes the 2 × 2 matrix $\begin{bmatrix} (M^{ij})_{ii} & (M^{ij})_{ij} \\ (M^{ij})_{ji} & (M^{ij})_{jj} \end{bmatrix}$ symmetric and positive semidefinite.

Scaled Diagonally Dominant Matrix (sdd)

$Q \in \mathcal{S}^n$ is **sdd** if and only if it can be written as $Q = \sum_{i,j=1,\ldots,n,i<j} M^{ij}$

where, $M^{ij} \in \mathcal{S}^n$ with zero every where except at most for 4 entries

$$(M^{ij})_{ii}, (M^{ij})_{ij}, (M^{ij})_{ji}, (M^{ij})_{jj}$$

which makes the 2 × 2 matrix $\begin{bmatrix} (M^{ij})_{ii} & (M^{ij})_{ij} \\ (M^{ij})_{ji} & (M^{ij})_{jj} \end{bmatrix}$ symmetric and positive semidefinite.

$$\begin{bmatrix} (M^{ij})_{ii} & (M^{ij})_{ij} \\ (M^{ij})_{ji} & (M^{ij})_{jj} \end{bmatrix} \succeq 0 \Rightarrow \boxed{\begin{array}{c} trace(.) = \lambda_1 + \lambda_2 \geq 0 \qquad det(.) = \lambda_1 \lambda_2 \geq 0 \\ \Downarrow \\ \lambda_1 \geq 0, \ \lambda_2 \geq 0 \end{array}} \Rightarrow$$

1) $(M^{ij})_{ii} + (M^{ij})_{jj} \geq 0$

2) $(M^{ij})_{ii}(M^{ij})_{jj} - (M^{ij})_{ji}(M^{ij})_{ji} \geq 0$

Scaled Diagonally Dominant Matrix (sdd)

$Q \in \mathcal{S}^n$ is **sdd** if and only if it can be written as $Q = \sum_{i,j=1,\ldots,n,i<j} M^{ij}$

where, $M^{ij} \in \mathcal{S}^n$ with zero every where except at most for 4 entries

$$(M^{ij})_{ii}, (M^{ij})_{ij}, (M^{ij})_{ji}, (M^{ij})_{jj}$$

which makes the 2 × 2 matrix $\begin{bmatrix} (M^{ij})_{ii} & (M^{ij})_{ij} \\ (M^{ij})_{ji} & (M^{ij})_{jj} \end{bmatrix}$ symmetric and positive semidefinite.

$$\begin{bmatrix} (M^{ij})_{ii} & (M^{ij})_{ij} \\ (M^{ij})_{ji} & (M^{ij})_{jj} \end{bmatrix} \succcurlyeq 0$$

$trace(.) = \lambda_1 + \lambda_2 \geq 0 \quad det(.) = \lambda_1 \lambda_2 \geq 0$

$\lambda_1 \geq 0, \ \lambda_2 \geq 0$

1) $(M^{ij})_{ii} + (M^{ij})_{jj} \geq 0$

2) $(M^{ij})_{ii}(M^{ij})_{jj} - (M^{ij})_{ji}(M^{ij})_{ji} \geq 0$

$$||C_i x + d_i||_2 \leq e_i^T x + f_i, \ i = 1, \ldots, m$$

Second Order Cone

$$\left\| \begin{bmatrix} 2(M^{ij})_{ij} \\ (M^{ij})_{ii} - (M^{ij})_{jj} \end{bmatrix} \right\|_2 \leq (M^{ij})_{ii} + (M^{ij})_{jj}$$

- F. Alizadeh and D. Goldfarb, “Second-order cone programming,” Mathematical programming, vol. 95, no. 1, pp. 3–51, 2003.

Scaled Diagonally Dominant Matrix (sdd)

$Q \in \mathcal{S}^n$ is **sdd**, If there exist a diagonal matrix D with positive diagonal entries, such that DQD is **dd**.

$$\mathcal{S}^n_{dd} \subset \mathcal{S}^n_{sdd} \subset \mathcal{S}^n_{+}$$

Every dd matrix is sdd matrix with $D = I$

Every sdd matrix is sum of psd matrices M^{ij}

$$p(x) \in SOS \xLeftrightarrow{p(x) = B^T(x)QB(x)} \underbrace{Q \in \mathcal{S}^n_+}_{\text{PSD Matrix}} \Longrightarrow \text{SDP}$$

$$p(x) \in DSOS \xLeftrightarrow{p(x) = B^T(x)QB(x)} Q \in \mathcal{S}^n_{dd} \Longrightarrow \text{LP}$$

$$p(x) \in SDSOS \xLeftrightarrow{p(x) = B^T(x)QB(x)} Q \in \mathcal{S}^n_{sdd} \Longrightarrow \text{Second Order Cone Program (SOCP)}$$

Scaled Diagonally Dominant Matrix (sdd)

$Q \in \mathcal{S}^n$ is **sdd**, If there exist a diagonal matrix D with positive diagonal entries, such that DQD is **dd**.

$$\mathcal{S}^n_{dd} \subset \mathcal{S}^n_{sdd} \subset \mathcal{S}^n_+$$

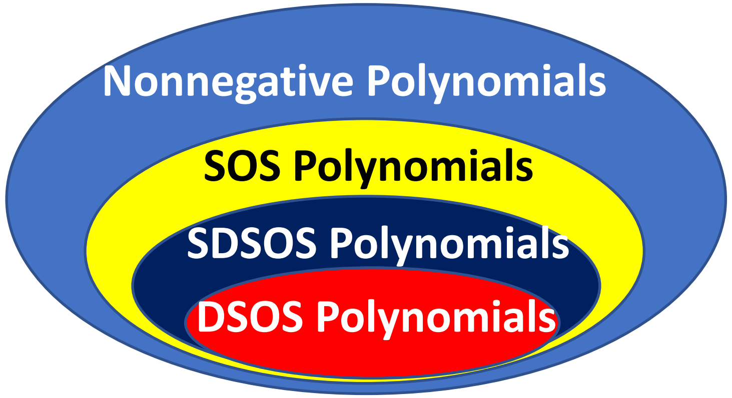
Every dd matrix is sdd matrix with $D = I$

Every sdd matrix is sum of psd matrices M^{ij}

$$p(x) \in SOS \xLeftrightarrow{p(x) = B^T(x)QB(x)} \underbrace{Q \in \mathcal{S}^n_+}_{\text{PSD Matrix}} \Longrightarrow \text{SDP}$$

$$p(x) \in DSOS \xLeftrightarrow{p(x) = B^T(x)QB(x)} Q \in \mathcal{S}^n_{dd} \Longrightarrow \text{LP}$$

$$p(x) \in SDSOS \xLeftrightarrow{p(x) = B^T(x)QB(x)} Q \in \mathcal{S}^n_{sdd} \Longrightarrow$$ **Second Order Cone Program (SOCP)**



(Appendix I)

Unconstrained optimization

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad p(x)$$

$$\begin{aligned}&\underset{\gamma\in\mathbb{R}}{\text{maximize}} && \gamma\\ &\text{subject to} && p(x)-\gamma\geq 0, \quad \forall x\in\mathbb{R}^n\end{aligned}$$

SOS Programing: SOS SDP

$$\begin{aligned}&\underset{Q\in\mathcal{S}^n,\gamma}{\text{maximize}} && \gamma\\ &\text{subject to} && p(x)-\gamma = \ B^T(x)QB(x)\\ & && Q\in\mathcal{S}^n_+\end{aligned}$$

SDSOS Programing: SOCP

$$\begin{aligned}&\underset{Q\in\mathcal{S}^n,\gamma}{\text{maximize}} && \gamma\\ &\text{subject to} && p(x)-\gamma = \ B^T(x)QB(x)\\ & && Q\in\mathcal{S}^n_{sdd}\end{aligned}$$

Constrained optimization

$$\begin{aligned}&\underset{x\in\mathbb{R}^n}{\text{minimize}} && p(x)\\ &\text{subject to} && x\in\mathbf{K}=\{x\in\mathbb{R}^n: g_i(x)\geq 0, \ i=1,...,m\}\end{aligned}$$

$$\begin{aligned}&\underset{\gamma\in\mathbb{R}}{\text{maximize}} && \gamma\\ &\text{subject to} && p(x)-\gamma\geq 0, \quad \forall x\in\mathbf{K}\end{aligned}$$

SOS Programing: SOS SDP

$$\begin{aligned}&\underset{\gamma,Q_i|_{i=0}^m}{\text{maximize}} && \gamma\\ &\text{subject to} && p(x)-\gamma-\textstyle\sum_{i=1}^m\sigma_i(x)g_i(x)=B_d(x)Q_0B_d(x)\\ & && \sigma_i(x)=B^T_{d_i}(x)Q_iB_{d_i}(x), \ i=1,...,m\\ & && Q_i\in\mathcal{S}^n_+, i=0,...,m\end{aligned}$$

SDSOS Programing: SOCP

$$\begin{aligned}&\underset{\gamma,Q_i|_{i=0}^m}{\text{maximize}} && \gamma\\ &\text{subject to} && p(x)-\gamma-\textstyle\sum_{i=1}^m\sigma_i(x)g_i(x)=B_d(x)Q_0B_d(x)\\ & && \sigma_i(x)=B^T_{d_i}(x)Q_iB_{d_i}(x), \ i=1,...,m\\ & && Q_i\in\mathcal{S}^n_{sdd}, i=0,...,m\end{aligned}$$

SDSOS/DSOS Programming

SPOTT: MATLAB package for DSOS and SDSOS optimization written using the SPOT toolbox.

- A. Ahmadi and A. Majumdar," DSOS and SDSOS Optimization: More Tractable Alternatives to Sum of Squares and Semidefinite Optimization", *SIAM Journal on Applied Algebraic Geometry, 2019.*
- A. Ahmadi, A. Majumdary, "Some applications of polynomial optimization in operations research and real-time decision making", Optimization Letters, Volume 10, Issue 4, pp 709–729, 2016.
- A. Majumdar, A. A. Ahmadi, R. Tedrake,, "Control and verification of high-dimensional systems with DSOS and SDSOS programming", 53rd IEEE Conference on Decision and Control 2014

6-link pendulum

Atlas

Applications:

Control and analyze of high dimensional systems

$$\mathrm{P}^* = \underset{x \in \mathbb{R}^2}{\text{minimize}} \quad 3 + 2x_1 + 2x_2 + 3x_1^2 + 2x_1x_2 + 3x_2^2 + x_1^4 + x_2^4$$

SDSOS Programming in SPOT

`x = msspoly('x',2);`	variables x_1, x_2
`prog = spotsosprog;`	DSOS/SDSOS Programing
`prog = prog.withIndeterminate(x);`	
`p = 3+2*x(1)+2*x(2)+3*x(1)^2+2*x(1)*x(2)+3*x(2)^2+x(1)^4+x(2)^4;`	$p(x)$
`[prog,gamma] = prog.newFree(1);`	variable γ
`prog = prog . withSDSOS (p-gamma) ;`	$p(x) - \gamma \in DSOS/SDSOS/SOS$
`sol = prog . minimize ( -gamma,@spot_mosek) ;`	SDP solver, solve SDSOS programming
`double(sol.eval(gamma))`	obtained lower bound

$\mathrm{P}^{*2}_{sos} = 2.5074 = \mathrm{P}^*$ $\quad \mathrm{P}^{*2}_{sdsos} = 2.0877 \leq \mathrm{P}^{*2}_{sos}$ $\quad \mathrm{P}^{*2}_{dsos} = 1 \leq \mathrm{P}^{*2}_{sdsos}$

https://github.com/jasour/rarnop19/blob/master/Lecture6_modified-SOS/SDSOS-DSOS/Example_SDSOS_1.m

$$P^* = \underset{x\in\mathbb{R}^2}{\text{minimize}} \quad (1+x_1x_2)^2 - x_1x_2 + (1-x_2)^2$$

$$\text{subject to} \quad x \in \mathbf{K} = \{x\in\mathbb{R}^2 : 3-2x_2-x_1^2-x_2^2 \geq 0,\ -x_1-x_2-x_1x_2 \geq 0,\ 1+x_1x_2 \geq 0\}$$

SDSOS Programming in SPOT

`d=1;`	relaxation order
`x = msspoly('x',2);`	variables x_1, x_2
`prog = spotsosprog;`	DSOS/SDSOS Programing
`prog = prog.withIndeterminate(x);`	
`p = (1+x(1)*x(2))^2-x(1)*x(2)+(1-x(2))^2 ;`	$p(x)$
`g=[3-2*x(2)-x(1)^2-x(2)^2;-x(1)-x(2)-x(1)*x(2);1+x(1)*x(2)];`	$\mathbf{K}$
`[prog,gamma] = prog.newFree(1);`	variable γ
`mos=monomials(x,0:2*d);`	vector of monomials up to order $2d$
`[prog,coeffs1] = prog.newFree(length(mos)); s1 = coeffs1'*mos;`	$\sigma_1$with coefficients c_1
`[prog,coeffs2] = prog.newFree(length(mos)); s2 = coeffs2'*mos;`	$\sigma_2$with coefficients c_2
`[prog,coeffs3] = prog.newFree(length(mos)); s3 = coeffs3'*mos;`	$\sigma_3$with coefficients c_3
`prog = prog . withSDSOS (p-gamma-[s1 s2 s3]*g);`	$p(x)-\gamma-(\sigma_1 g_1(x)+\sigma_2 g_2(x)+\sigma_3 g_3(x))$
`prog = prog . withSDSOS (s1);`	$\sigma_1 \in SDSOS$
`prog = prog . withSDSOS (s2);`	$\sigma_2 \in SDSOS$
`prog = prog . withSDSOS (s3);`	$\sigma_3 \in SDSOS$
`sol = prog . minimize ( -gamma,@spot_mosek);`	SDP solver, solve SDSOS programming
`double(sol.eval(gamma))`	obtained lower bound

$P^{*1}_{sos} = 0.7549 = P^*$

$P^{*1}_{sdsos} = 0.7549 = P^{*1}_{sos}$

$P^{*1}_{dsos} = 0.5 \leq P^{*1}_{sdsos}$ $\quad P^{*2}_{dsos} = 0.6585$ $\quad P^{*3}_{dsos} = 0.6891$ $\quad P^{*4}_{dsos} = 0.6935$ $\quad P^{*5}_{dsos} = 0.6937$ $\quad P^{*6}_{dsos} = 0.6937$

https://github.com/jasour/rarnop19/blob/master/Lecture6_modified-SOS/SDSOS-DSOS/Example_SDSOS_2.m

$$\begin{aligned} \mathrm{P}^* =& \underset{x\in\mathbb{R}^2}{\text{minimize}} \quad \frac{1}{3}x_1^6 - \frac{21}{10}x_1^4 + 4x_1^2 + x_1x_2 + 4x_2^4 - 4x_2^2 + \frac{3}{2} \\ &\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^2 : -\frac{1}{16}x_1^4 + \frac{1}{4}x_1^3 - \frac{1}{4}x_1^2 - \frac{9}{100}x_2^2 + \frac{29}{400} \geq 0\} \end{aligned}$$

$\mathrm{P}^{*3}_{sos} = 0.4684 = \mathrm{P}^*$

$\mathrm{P}^{*3}_{sdsos} = 0.3114 \leq \mathrm{P}^{*1}_{sos}$ $\quad \mathrm{P}^{*5}_{sdsos} = 0.3132$ $\quad \mathrm{P}^{*7}_{sdsos} = 0.3538$

$\mathrm{P}^{*1}_{dsos} = -0.0341 \leq \mathrm{P}^{*1}_{sdsos}$ $\quad \mathrm{P}^{*5}_{dsos} = -0.0061$ $\quad \mathrm{P}^{*7}_{dsos} = -0.0353$

https://github.com/jasour/rarnop19/blob/master/Lecture6_modified-SOS/SDSOS-DSOS/Example_SDSOS_3.m

Main Benefit:
SDSOS/DSOS can scale to problems where SOS programming ceases to run due to memory/computation constraints.

- A. Ahmadi and A. Majumdar," DSOS and SDSOS Optimization: More Tractable Alternatives to Sum of Squares and Semidefinite Optimization", *SIAM Journal on Applied Algebraic Geometry, 2019.*

Illustrative Example:

$$\mathrm{P}^* = \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad = 5 + \sum_{i=1}^{n} (x_i - 1)^2 \qquad p^* = 5, \quad x^* = [1, 1, ..., 1]^T \in \mathbb{R}^n$$

Number of variables Polynomial of order 2

- **SOS:** **Variables:200** **Relaxation Order=1** **time= 286.5458 (s)** p^***=5** **sdp solver: mosek**
- **SDSOS:** **Variables:200** **Relaxation Order=1** **time= 3.6338 (s)** p^***=5** **sdp solver: mosek**
- **DSOS:** **Variables:200** **Relaxation Order=1** **time=2.6824 (s)** p^***=5** **sdp solver: mosek**

https://github.com/jasour/rarnop19/blob/master/Lecture6_modified-SOS/SDSOS-DSOS/Example_SDSOS_compare_Uncons.m

Bounded Degree SOS

- Jean B. Lasserre, Kim-Chuan Toh, Shouguang Yang, “A bounded degree SOS hierarchy for polynomial optimization”, EURO Journal on Computational Optimization March 2017, Volume 5, Issue 1–2, pp 87–117

Nonnegative polynomial

$$p(x) \geq 0, \quad \forall x \in \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, \ldots, m\}$$

SDP Relaxation

Putinar's Positivity Certificate

$$p(x) = \sigma_0(x) + \sum_{i=1}^{m} \sigma_i(x) g_i(x)$$

$$\sigma_i(x) \in SOS_{2d_i}, \ i = 0, \ldots, m$$

SDP

$$p(x) - \sum_{i=1}^{m} \sigma_i(x) g_i(x) = B_d(x) Q_0 B_d(x)$$

$$\sigma_i(x) = B_{d_i}^T(x) Q_i B_{d_i}(x), \ i = 1, \ldots, m$$

$$Q_i \in \mathcal{S}_+^n, i = 0, \ldots, m$$

Nonnegative polynomial

$$p(x) \geq 0, \quad \forall x \in \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \;\; i = 1, \ldots, m\}$$

SDP
Relaxation

Putinar's Positivity Certificate

$$p(x) = \sigma_0(x) + \sum_{i=1}^{m} \sigma_i(x) g_i(x)$$

$$\sigma_i(x) \in SOS_{2d_i}, \;\; i = 0, \ldots, m$$

SDP

$$p(x) - \sum_{i=1}^{m} \sigma_i(x) g_i(x) = B_d(x) Q_0 B_d(x)$$

$$\sigma_i(x) = B_{d_i}^T(x) Q_i B_{d_i}(x), \;\; i = 1, \ldots, m$$

$$Q_i \in \mathcal{S}_+^n, i = 0, \ldots, m$$

LP
Relaxation

Krivine-Stengle's Positivity Certificate

$\text{Let} \;\; \mathbf{K} = \{x \in \mathbb{R}^n : 0 \leq g_i(x) \leq 1, \;\; i = 1, \ldots, m\}$ (normalized polynomials)

Nonnegative polynomial

$$p(x) \geq 0, \quad \forall x \in \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \ \ i = 1, ..., m\}$$

SDP Relaxation

Putinar’s Positivity Certificate

$$p(x) = \sigma_0(x) + \textstyle\sum_{i=1}^{m} \sigma_i(x) g_i(x)$$

$$\sigma_i(x) \in SOS_{2d_i}, \ \ i = 0, ..., m$$

SDP

$$p(x) - \textstyle\sum_{i=1}^{m} \sigma_i(x) g_i(x) = B_d(x) Q_0 B_d(x)$$

$$\sigma_i(x) = B_{d_i}^T(x) Q_i B_{d_i}(x), \ \ i = 1, ..., m$$

$$Q_i \in \mathcal{S}_+^n, i = 0, ..., m$$

LP Relaxation

Krivine-Stengle’s Positivity Certificate

Let $\mathbf{K} = \{x \in \mathbb{R}^n : 0 \leq g_i(x) \leq 1, \ \ i = 1, ..., m\}$ (normalized polynomials)

$$p(x) = \sum_{\alpha, \beta \in \mathbb{N}^m} \lambda_{\alpha\beta} g_1^{\alpha_1}(x) ... g_m^{\alpha_m}(x)(1 - g_1(x))^{\beta_1} ... (1 - g_m(x))^{\beta_m}$$

Unknowns: $\lambda_{\alpha\beta}$ Finitely many Nonnegative scalars

- *Theorem 2.23.* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Nonnegative polynomial

$$p(x) \geq 0, \quad \forall x \in \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \; i = 1, \ldots, m\}$$

SDP Relaxation

Putinar's Positivity Certificate

$$p(x) = \sigma_0(x) + \sum_{i=1}^{m} \sigma_i(x) g_i(x)$$

$$\sigma_i(x) \in SOS_{2d_i}, \quad i = 0, \ldots, m$$

SDP

$$p(x) - \sum_{i=1}^{m} \sigma_i(x) g_i(x) = B_d(x) Q_0 B_d(x)$$

$$\sigma_i(x) = B_{d_i}^T(x) Q_i B_{d_i}(x), \; i = 1, \ldots, m$$

$$Q_i \in \mathcal{S}_+^n, i = 0, \ldots, m$$

LP Relaxation

Krivine-Stengle's Positivity Certificate

$$\left.\begin{cases} x \in \mathbf{K} & p(x) = \sum \lambda_{\alpha\beta} \underbrace{g_1^{\alpha_1}(x) \ldots g_m^{\alpha_m}(x)}_{+} \underbrace{(1 - g_1(x))^{\beta_1} \ldots (1 - g_m(x))^{\beta_m}}_{+} \longrightarrow p(x) \geq 0 \\ x \notin \mathbf{K} & p(x) = \sum \lambda_{\alpha\beta} \underbrace{g_1^{\alpha_1}(x) \ldots g_m^{\alpha_m}(x)}_{+/-} \underbrace{(1 - g_1(x))^{\beta_1} \ldots (1 - g_m(x))^{\beta_m}}_{+/-} \longrightarrow p(x) \geq 0 \text{ or } p(x) \leq 0 \end{cases}\right\} p(x) \geq 0 \;\; \forall \; x \in \mathbf{K}$$

$g_1(x) \leq 0$ *or* $g_1(x) \geq 1$

- *Theorem 2.23.* Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.

Nonnegative polynomial

$$p(x) \geq 0, \quad \forall x \in \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \ i = 1, \ldots, m\}$$

SDP
Relaxation

Putinar's Positivity Certificate

$$p(x) = \sigma_0(x) + \sum_{i=1}^{m} \sigma_i(x) g_i(x)$$

$$\sigma_i(x) \in SOS_{2d_i}, \quad i = 0, \ldots, m$$

SDP

$$p(x) - \sum_{i=1}^{m} \sigma_i(x) g_i(x) = B_d(x) Q_0 B_d(x)$$

$$\sigma_i(x) = B_{d_i}^T(x) Q_i B_{d_i}(x), \ i = 1, \ldots, m$$

$$Q_i \in \mathcal{S}_+^n, i = 0, \ldots, m$$

LP
Relaxation

Krivine-Stengle's Positivity Certificate

Let $\mathbf{K} = \{x \in \mathbb{R}^n : 0 \leq g_i(x) \leq 1, \ i = 1, \ldots, m\}$ (normalized polynomials)

$$p(x) = \sum_{\alpha, \beta \in \mathbb{N}^m} \lambda_{\alpha\beta} g_1^{\alpha_1}(x) \ldots g_m^{\alpha_m}(x)(1 - g_1(x))^{\beta_1} \ldots (1 - g_m(x))^{\beta_m}$$

Linear constraints on λ

Unknowns: $\lambda_{\alpha\beta}$ Finitely many Nonnegative scalars

- Determining if $p(x) \geq 0, \quad \forall x \in \mathbf{K}$ leads to a linear optimization feasibility problem.

- *Theorem 2.23.* Jean Bernard Lasserre, "Moments, Positive Polynomials and Their Applications" Imperial College Press Optimization Series, V. 1, 2009.
- Sherali H.D., Adams W.P. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems, SIAM J. Discr. Math. 3, pp. 411–430, 1990.

$$\mathbf{P}^* = \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad p(x)$$

$$\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ \ i = 1, ..., m\}$$

$$\underset{\gamma\in\mathbb{R}}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad p(x) - \gamma \geq 0, \quad \forall x \in \mathbf{K}$$

SDP Relaxation

$$\underset{\gamma, Q_i|_{i=0}^{m}}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad p(x) - \gamma - \textstyle\sum_{i=1}^{m} \sigma_i(x) g_i(x) = B_d(x) Q_0 B_d(x)$$

$$\sigma_i(x) = B_{d_i}^T(x) Q_i B_{d_i}(x), \ \ i = 1, ..., m$$

$$Q_i \in \mathcal{S}_+^n, i = 0, ..., m$$

LP Relaxation

Let $d \in \mathbb{N}$

$$\mathbf{P}_L^{*d} = \underset{\gamma, \lambda_{\alpha\beta}\geq 0}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \ \ p(x) - \gamma = \sum_{\substack{\forall\ \alpha, \beta \in \mathbb{N}^m \\ \sum_{j=1}^{m} \alpha_j + \beta_j \leq d}} \lambda_{\alpha\beta} \ g_1^{\alpha_1}(x) ... g_m^{\alpha_m}(x)(1 - g_1(x))^{\beta_1} ... (1 - g_m(x))^{\beta_m}$$

- **Theorem:** Let $\mathbf{K}$ be compact (Archimedean). $\mathbf{P}_L^{*d} \leq \mathbf{P}_L^{*d+1}$ $\quad lim_{d\to\infty}\ \mathbf{P}_L^{*d} = \mathbf{P}^*$
- *Theorem 5.10.* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

- LP-relaxations suffer from several serious theoretical and practical drawbacks:

- The LPs of the hierarchy are numerically **ill-conditioned**.
  - It involves products of arbitrary powers of the $g_i(x)$ ’s and $(1 - g_i(x))$’s.
  - In particular, the presence of large coefficients is source of ill-conditioning and numerical instability.

- The sequence of the associated optimal values converges to the global optimum only **asymptotically** and **not in finitely** many steps. (**Appendix II**)

- Finite convergence even does not hold for **convex optimizations**. (In standard SOS finite convergence takes place for SOS-convex problems)

- LP-relaxations suffer from several serious theoretical and practical drawbacks:

- The LPs of the hierarchy are numerically **ill-conditioned**.
  - It involves products of arbitrary powers of the $g_i(x)$ ’s and $(1 - g_i(x))$’s.
  - In particular, the presence of large coefficients is source of ill-conditioning and numerical instability.

- The sequence of the associated optimal values converges to the global optimum only **asymptotically** and **not in finitely** many steps. (**Appendix II**)

- Finite convergence even does not hold for **convex optimizations**. (In standard SOS finite convergence takes place for SOS-convex problems)

Bounded Degree SOS (BSOS):
Hierarchy of convex relaxations which **combines** some of the advantages of the **SOS** and **LP** hierarchies.

Bounded Degree SOS (BSOS):

Hierarchy of convex relaxations which **combines** some of the advantages of the **SOS**- and **LP**- hierarchies.

- **SOS Relaxation**

$$p(x) = \boxed{\sigma_0(x)} + \sum_{i=1}^{m} \sigma_i(x) g_i(x)$$

$$\sigma_0(x) \in SOS_{2d}$$

$$\sigma_i(x) \in SOS_{2d_i}, \quad i = 1, \ldots, m$$

- **LP Relaxation**

$$p(x) = \boxed{\sum_{\substack{\forall\, \alpha, \beta \in \mathbb{N}^m \\ \sum_{j=1}^{m} \alpha_j + \beta_j \leq i}} \lambda_{\alpha\beta}\, g_1^{\alpha_1}(x) \ldots g_m^{\alpha_m}(x) (1 - g_1(x))^{\beta_1} \ldots (1 - g_m(x))^{\beta_m}}$$

- **BSOS Relaxation**

$$p(x) = \sigma_0(x) + \sum_{\substack{\forall\, \alpha, \beta \in \mathbb{N}^m \\ \sum_{j=1}^{m} \alpha_j + \beta_j \leq d}} \lambda_{\alpha\beta} g_1^{\alpha_1}(x) \ldots g_m^{\alpha_m}(x) (1 - g_1(x))^{\beta_1} \ldots (1 - g_m(x))^{\beta_m}$$

$$\sigma_0(x) \in SOS_{2k}$$

$k \in \mathbb{N}$:Degree of SOS polynonial
Determines the size of SDP

$d \in \mathbb{N}$:degree of LP representation
Determines the number Linear Constraints

- Jean B. Lasserre, Kim-Chuan Toh, Shouguang Yang, “A bounded degree SOS hierarchy for polynomial optimization”, EURO Journal on Computational Optimization March 2017, Volume 5, Issue 1–2, pp 87–117

$$\mathbf{P}^* = \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad p(x)$$

$$\text{subject to} \quad x \in \mathbf{K} = \{x\in\mathbb{R}^n : g_i(x) \geq 0,\ \ i = 1, ..., m\}$$

$$\underset{\gamma\in\mathbb{R}}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad p(x) - \gamma \geq 0, \quad \forall x \in \mathbf{K}$$

SOS SDP Relaxation

$$\underset{\gamma, Q_i|_{i=0}^{m}}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad p(x) - \gamma - \textstyle\sum_{i=1}^{m} \sigma_i(x) g_i(x) = B_d(x) Q_0 B_d(x)$$

$$\sigma_i(x) = B_{d_i}^T(x) Q_i B_{d_i}(x), \ \ i = 1, ..., m$$

$$Q_i \in \mathcal{S}_+^n, i = 0, ..., m$$

LP Relaxation

$$\mathbf{P}_L^{*i} = \underset{\gamma, \lambda_{\alpha\beta}\geq 0}{\text{maximize}} \quad \gamma$$

subject to

$$p(x) - \gamma = \sum \lambda_{\alpha\beta} \quad g_1^{\alpha_1}(x)...g_m^{\alpha_m}(x)(1-g_1(x))^{\beta_1}...(1-g_m(x))^{\beta_m}$$

$$\forall\ \alpha, \beta \in \mathbb{N}^m$$

$$\textstyle\sum_{j=1}^{m} \alpha_j + \beta_j \leq i$$

Bounded SOS Relaxation

$$\mathrm{P}_d^{*k} = \underset{\gamma, \lambda_{\alpha\beta}\geq 0, Q_0}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad p(x) - \gamma - \sum \lambda_{\alpha\beta} g_1^{\alpha_1}(x)...g_m^{\alpha_m}(x)(1-g_1(x))^{\beta_1}...(1-g_m(x))^{\beta_m} = B_k(x) Q_0 B_k(x)$$

$$\forall\ \alpha, \beta \in \mathbb{N}^m$$

$$\textstyle\sum_{j=1}^{m} \alpha_j + \beta_j \leq d$$

$$Q_0 \in \mathcal{S}_+^n$$

- **Theorem:** Let $k \in \mathbb{N}$ be fixed. $\quad \mathbf{P}_d^{*k} \leq \mathbf{P}_{d+1}^{*k} \qquad lim_{d\to\infty}\ \mathbf{P}_d^{*k} = \mathbf{P}^*$
  - **Finite convergence (Like standard SOS)** (Finite convergence condition : Rank condition of the dual (moment) problem) (**Appendix III**)
  - Unlike standard SOS, the size of SDP is fixed $\binom{n+k}{n}$

- Section 1.1, Jean B. Lasserre, Kim-Chuan Toh, Shouguang Yang, “A bounded degree SOS hierarchy for polynomial optimization”, EURO Journal on Computational Optimization March 2017, Volume 5, Issue 1–2, pp 87–117

Example 1

$$\begin{array}{llllllllll}
(P_1) & f = & x_1^2 & -x_2^2 & & +x_3^2 & -x_4^2 & +x_1 & -x_2 & \\
\text{s.t.} & 0 \leq g_1 = & 2x_1^2 & +3x_2^2 & +2x_1x_2 & +2x_3^2 & +3x_4^2 & +2x_3x_4 & \leq 1 \\
 & 0 \leq g_2 = & 3x_1^2 & +2x_2^2 & -4x_1x_2 & +3x_3^2 & +2x_4^2 & -4x_3x_4 & \leq 1 \\
 & 0 \leq g_3 = & x_1^2 & +6x_2^2 & -4x_1x_2 & +x_3^2 & +6x_4^2 & -4x_3x_4 & \leq 1 \\
 & 0 \leq g_4 = & x_1^2 & +4x_2^2 & -3x_1x_2 & +x_3^2 & +4x_4^2 & -3x_3x_4 & \leq 1 \\
 & 0 \leq g_5 = & 2x_1^2 & +5x_2^2 & +3x_1x_2 & +2x_3^2 & +5x_4^2 & +3x_3x_4 & \leq 1 \\
 & 0 \leq x. & & & & & & &
\end{array}$$

$$\mathrm{P}_{d=1}^{*k=1} = -0.57491 = \mathrm{P}^*$$

Example 2

$$\begin{array}{lllllll}
(P_2) & f = & x_1^4x_2^2 & +x_1^2x_2^4 & -x_1^2x_2^2 & \\
\text{s.t.} & 0 \leq g_1 = & x_1^2 & +x_2^2 & & \leq 1 \\
 & 0 \leq g_2 = & 3x_1^2 & +2x_2^2 & -4x_1x_2 & \leq 1 \\
 & 0 \leq g_3 = & x_1^2 & +6x_2^4 & -8x_1x_2 + 2.5 & \leq 1 \\
 & 0 \leq g_4 = & x_1^4 & +3x_2^4 & & \leq 1 \\
 & 0 \leq g_5 = & x_1^2 & +x_2^3 & & \leq 1 \\
 & 0 \leq x_1, & 0 \leq x_2. & & &
\end{array}$$

$$\mathrm{P}^* = -0.037037$$

Fixed size of SDP

$k = 3$ $\quad \mathrm{P}_{d=1}^{*k=3} = -0.041855 \quad \mathrm{P}_{d=2}^{*k=3} = -0.037139 \quad \mathrm{P}_{d=3}^{*k=3} = -0.037087 \quad \mathrm{P}_{d=4}^{*k=3} = -0.037073 \quad \mathrm{P}_{d=5}^{*k=3} = -0.037046$

$k = 4$ $\quad \mathrm{P}_{d=1}^{*k=4} = -0.038596 \quad \mathrm{P}_{d=2}^{*k=4} = -0.037046 \quad \mathrm{P}_{d=3}^{*k=4} = -0.037040 \quad \mathrm{P}_{d=4}^{*k=4} = -0.037038 \quad \mathrm{P}_{d=5}^{*k=4} = -0.037037$

More examples: https://github.com/jasour/rarnop19/blob/master/Lecture6_modified-SOS/Bounded_Degree_SOS/BSOS_Example1.m

https://github.com/tweisser/Sparse_BSOS/tree/master/test_suite/Dense

---

Code: https://github.com/tweisser/Sparse_BSOS

- Jean B. Lasserre, Kim-Chuan Toh, Shouguang Yang, “A bounded degree SOS hierarchy for polynomial optimization”, EURO Journal on Computational Optimization March 2017, Volume 5, Issue 1–2, pp 87–117

Topics

1) (Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS)
   Modified SOS optimization that results in LP and Second order cone program

2) Bounded degree SOS (BSOS)
   Modified SOS optimization that results in smaller SDP's.

3) Spars Sum-of-Squares Optimization (SSOS)
   Takes advantage of sparsity of the original problem to generate smaller SDP.

4) Sparse Bounded Degree Sum-of-Squares Optimization (SBSOS)
   Combination of 2 and 3

Sparse SOS

- H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

- Zheng, Y., Fantuzzi, G., & Papachristodoulou, A. “Sparse sum-of-squares (SOS) optimization: A bridge between DSOS/SDSOS and SOS optimization for sparse polynomials”, arXiv preprint arXiv:1807.05463. 2018

- Take advantage of structure (sparsity) of the problem to solve smaller SDP

- ➢ Take advantage of structure (sparsity) of the problem to solve smaller SDP

1) PSD Constraint obtained form SOS/Moment Relaxation.

- (Under some conditions)We can replace Constraint of the form $Q \succcurlyeq 0$ by **PSD** constraints of set of smaller matrices.

Example:

$$Q = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0.5 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.5 & 1 \\ 0 & 1 & 2 \end{bmatrix} \Rightarrow Q \text{ is PSD becasue :}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 0.5 \end{bmatrix} \succcurlyeq 0 \qquad \begin{bmatrix} 0.5 & 1 \\ 1 & 2 \end{bmatrix} \succcurlyeq 0$$

- Take advantage of structure (sparsity) of the problem to solve smaller SDP

1) PSD Constraint obtained form SOS/Moment Relaxation.

- (Under some conditions)We can replace Constraint of the form $Q \succcurlyeq 0$ by **PSD** constraints of set of smaller matrices.

Example:

$$Q = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0.5 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.5 & 1 \\ 0 & 1 & 2 \end{bmatrix} \Rightarrow$$

Q is PSD becasue :

$$\begin{bmatrix} 2 & 1 \\ 1 & 0.5 \end{bmatrix} \succcurlyeq 0 \quad \begin{bmatrix} 0.5 & 1 \\ 1 & 2 \end{bmatrix} \succcurlyeq 0$$

2) SOS relaxation of nonnegative Polynomials

- (Under some conditions) We can replace constraint of $p(x) \in SOS$ by **SOS** constraints of low dimensional polynomials.

Example:

$$p(x_1, x_2, x_3\ ) = 2(1 + x_1 + x_3 + x_1^2 + x1\ x_2 + x_1^2 + x_2 x_3 + x_3^2) \Rightarrow$$

$$p(x_1, x_2, x_3) = p_1(x_1, x_2) + p_2(x_2, x_3)$$

$$p_1(x_1, x_2) = (1 + x_1)^2 + (x_1 + x_2)^2$$
$$p_2(x_2, x_3) = (1 + x_3^2)^2 + (x_2 + x_3)^2$$

Polynomial $p(x_1, x_2, x_3)$ is SOS because $p_1(x_1, x_2)$ abd $p_2(x_2, x_3)$ are SOS.

Sparse Polynomials

Polynomial: $p(x) : \mathbb{R}^n \to \mathbb{R}$ $p(x) = \sum_\alpha p_\alpha x^\alpha$ number of coefficients $\binom{n+d}{n} = \frac{(n+d)!}{n!d!}$

- ***Fully dense* polynomial:** Polynomial is fully dense if all the coefficients are nonzero

Sparse Polynomials

Polynomial: $p(x) : \mathbb{R}^n \to \mathbb{R}$ $\quad p(x) = \sum_\alpha p_\alpha x^\alpha$ $\quad$ number of coefficients $\binom{n+d}{n} = \frac{(n+d)!}{n!d!}$

- ***Fully dense* polynomial:** Polynomial is fully dense if all the coefficients are nonzero

- ***Sparse* polynomial:** Polynomial is sparse if the number of nonzero coefficients is much smaller than the number of the total coefficients .

Example: Sparse Polynomial $\quad p(x_1, x_2) = 0.56 + 0.5x_1 + 2x_2^2 + 0.75x_1^3x_2^2$

Number of nonzero coefficients: 4

Number of all coefficients: $\binom{2+5}{2} = \frac{(7)!}{2!5!} = 21$

Sparse Polynomials

Polynomial: $p(x) : \mathbb{R}^n \to \mathbb{R}$ $p(x) = \sum_\alpha p_\alpha x^\alpha$ number of coefficients $\binom{n+d}{n} = \frac{(n+d)!}{n!d!}$

- ***Fully dense* polynomial:** Polynomial is fully dense if all the coefficients are nonzero
- ***Sparse* polynomial:** Polynomial is sparse if the number of nonzero coefficients is much smaller than the number of the total coefficients .

Example: Sparse Polynomial $p(x_1, x_2) = 0.56 + 0.5x_1 + 2x_2^2 + 0.75x_1^3x_2^2$

Number of nonzero coefficients: 4

Number of all coefficients: $\binom{2+5}{2} = \frac{(7)!}{2!5!} = 21$

- ***Correlative Sparsity:*** It describes coupling between the variables $x_1, \dots, x_n$ of a polynomial $p(x) : \mathbb{R}^n \to \mathbb{R}$
  - Variables x_i and x_j are coupled if they appear simultaneously in a monomial of the polynomial.

Sparse Polynomials

Polynomial: $p(x) : \mathbb{R}^n \to \mathbb{R}$ $p(x) = \sum_\alpha p_\alpha x^\alpha$ number of coefficients $\binom{n+d}{n} = \frac{(n+d)!}{n!d!}$

- ***Fully dense* polynomial:** Polynomial is fully dense if all the coefficients are nonzero
- ***Sparse* polynomial:** Polynomial is sparse if the number of nonzero coefficients is much smaller than the number of the total coefficients .

Example: Sparse Polynomial $p(x_1, x_2) = 0.56 + 0.5x_1 + 2x_2^2 + 0.75x_1^3x_2^2$

Number of nonzero coefficients: 4

Number of all coefficients: $\binom{2+5}{2} = \frac{(7)!}{2!5!} = 21$

- ***Correlative Sparsity:*** It describes coupling between the variables $x_1, \dots, x_n$ of a polynomial $p(x) : \mathbb{R}^n \to \mathbb{R}$
  - Variables x_i and x_j are coupled if they appear simultaneously in a monomial of the polynomial.

Example: $p(x_1, x_2, x_3, x_4) = 0.56 + 0.5x_1 + 2x_1x_2^2 + 0.75x_3^3x_4^2$

Coupled variables: (x_1, x_2) , (x_3, x_4)

Missing Coupled variables: (x_1, x_3) , (x_1, x_4), (x_2, x_3), (x_2, x_4)

Sparse Polynomials

Polynomial: $p(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ $\quad p(x) = \sum_\alpha p_\alpha x^\alpha$ $\quad$ number of coefficients $\binom{n+d}{n} = \frac{(n+d)!}{n!d!}$

- ***Fully dense* polynomial:** Polynomial is fully dense if all the coefficients are nonzero
- ***Sparse* polynomial:** Polynomial is sparse if the number of nonzero coefficients is much smaller than the number of the total coefficients .

Example: Sparse Polynomial $\quad p(x_1, x_2) = 0.56 + 0.5x_1 + 2x_2^2 + 0.75x_1^3x_2^2$

Number of nonzero coefficients: 4

Number of all coefficients: $\binom{2+5}{2} = \frac{(7)!}{2!5!} = 21$

- ***Correlative Sparsity:*** It describes coupling between the variables $x_1, \dots, x_n$ of a polynomial $p(x) : \mathbb{R}^n \rightarrow \mathbb{R}$
  - Variables x_i and x_j are coupled if they appear simultaneously in a monomial of the polynomial.

Example: $\quad p(x_1, x_2, x_3, x_4) = 0.56 + 0.5x_1 + 2x_1x_2^2 + 0.75x_3^3x_4^2$

Coupled variables: (x_1, x_2) , (x_3, x_4)

Missing Coupled variables: (x_1, x_3) , (x_1, x_4), (x_2, x_3), (x_2, x_4)

- Number of all possible coupling between variables $x_1, \dots, x_n$: $\binom{n}{2}$
- Polynomial has **correlative sparsity** if the number of coupled variables is much smaller than the Number of all possible coupling

Sparse Polynomials

- ***Sparse polynomial***: Polynomial is **sparse** if the number of nonzero coefficients is much smaller than the number of the total coefficients.

- ***Correlative Sparsity:*** Polynomial has **correlative sparsity** if the number of coupled variables is much smaller than the Number of all possible coupling

Sparse Polynomials

- ***Sparse polynomial*****:** Polynomial is **sparse** if the number of nonzero coefficients is much smaller than the number of the total coefficients.

- ***Correlative Sparsity:*** Polynomial has **correlative sparsity** if the number of coupled variables is much smaller than the Number of all possible coupling

- Correlative sparsity is a special case of the sparsity.

- Correlative sparsity implies the sparsity, but the converse is not necessarily true.

$$p(x_1, x_2, x_3, x_4) = x_1 x_2 + x_1 x_3 + x_1^3 x_4 + x_2 x_3 + x_2 x_4 + x_3 x_4^{10}$$

Number of nonzero coefficients: 6

Number of all coefficients: $\binom{4+10}{4} = \frac{(14)!}{4!10!} = 1001$

Sparse Polynomial
With NO correlative sparsity

- (Under some conditions)Constraint of the form $p(x) \in SOS$ can be replaced by SOS constraints of **low dimensional** polynomials.

$p(x) \in SOS$ If and only if $p(x) = \sum_k p_k(X_k)$ $\quad p_k(X_k) \in SOS$

X_k: Coupled set variables of $p(x)$

H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

- (Under some conditions)Constraint of the form $p(x) \in SOS$ can be replaced by SOS constraints of **low dimensional** polynomials.

$$p(x) \in SOS$$

If and only if

$$p(x) = \sum_k p_k(X_k) \qquad p_k(X_k) \in \ SOS$$

X_k: Coupled set variables of $p(x)$

$$p(x) = B^T(x)QB(x)$$
$$Q \in \mathcal{S}_+^n$$

If and only if

$$p(x) = \sum_k z_k^T(x)Q_k z_k(x) \qquad Q_k \in \mathcal{S}_+^{C_k} \qquad C_k < n$$

$C_k \times C_k$ matirx

$C_k \times 1$ monomial vector

H. Waki, S. Kim, M. Kojima, and M. Muramatsu, "Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity," SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

- (Under some conditions)Constraint of the form $p(x) \in SOS$ can be replaced by SOS constraints of **low dimensional** polynomials.

$$p(x) \in SOS \quad \text{If and only if} \quad p(x) = \sum_k p_k(X_k) \qquad p_k(X_k) \in \ SOS$$

X_k: Coupled set variables of $p(x)$

$$\begin{gathered} p(x) = B^T(x) Q B(x) \\ Q \in \mathcal{S}_+^n \end{gathered} \quad \text{If and only if} \quad p(x) = \sum_k z_k^T(x) Q_k z_k(x) \qquad Q_k \in \mathcal{S}_+^{C_k} \qquad C_k < n$$

$C_k \times C_k$ matirx

$C_k \times 1$ monomial vector

H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

Example:

$$\begin{gathered} p(x_1, x_2, x_3) = 2x_1^2 + 2x_1x_2 + x_2^2 + 2x_2x_3 + 2x_3^2 \\ p(x) \in SOS \end{gathered} \quad \Longleftrightarrow \quad \begin{aligned} p(x_1, x_2, x_3) &= p_1(x_1, x_2) + p_2(x_2, x_3) \\ p_1(x_1, x_2) &= (\sqrt{2}x_1 + \sqrt{0.5}x_2)^2 \in SOS \\ p_2(x_2, x_3) &= (\sqrt{0.5}x_2 + \sqrt{2}x_3)^2 \in SOS \end{aligned}$$

- (Under some conditions)Constraint of the form $p(x) \in SOS$can be replaced by SOS constraints of **low dimensional** polynomials.

$$p(x) \in SOS$$

If and only if

$$p(x) = \sum_k p_k(X_k) \qquad p_k(X_k) \in \ SOS$$

X_k: Coupled set variables of $p(x)$

$$p(x) = B^T(x)QB(x)$$
$$Q \in \mathcal{S}_+^n$$

If and only if

$$p(x) = \sum_k z_k^T(x)Q_k z_k(x) \qquad Q_k \in \mathcal{S}_+^{C_k} \qquad C_k < n$$

$C_k \times C_k$ matirx

$C_k \times 1$ monomial vector

H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

Example:

$$p(x_1, x_2, x_3) = 2x_1^2 + 2x_1x_2 + x_2^2 + 2x_2x_3 + 2x_3^2$$
$$p(x) \in SOS$$

$$\Longleftrightarrow$$

$$p(x_1, x_2, x_3) = p_1(x_1, x_2) + p_2(x_2, x_3)$$
$$p_1(x_1, x_2) = (\sqrt{2}x_1 + \sqrt{0.5}x_2)^2 \in SOS$$
$$p_2(x_2, x_3) = (\sqrt{0.5}x_2 + \sqrt{2}x_3)^2 \in SOS$$

$$p(x_1, x_2, x_3) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^T \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \Longleftrightarrow p(x_1, x_2, x_3) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} 2 & 1 \\ 1 & 0.5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} x_2 \\ x_3 \end{bmatrix}^T \begin{bmatrix} 0.5 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix}$$

$$Q \in \mathcal{S}_+^3 \qquad Q_1 \in \mathcal{S}_+^2 \qquad Q_2 \in \mathcal{S}_+^2$$

- (Under some conditions)Constraint of the form $p(x) \in SOS$can be replaced by SOS constraints of **low dimensional** polynomials.

$$p(x) \in SOS \quad \Longleftrightarrow \quad p(x) = \sum_k p_k(X_k) \qquad p_k(X_k) \in \ SOS$$

If and only if

X_k: Coupled set variables of $p(x)$

$$\begin{gathered} p(x) = B^T(x)QB(x) \\ Q \in \mathcal{S}_+^n \end{gathered} \quad \Longleftrightarrow \quad p(x) = \sum_k z_k^T(x)Q_k z_k(x) \qquad Q_k \in \mathcal{S}_+^{C_k} \qquad C_k < n$$

If and only if

$C_k \times C_k$ matirx

$C_k \times 1$ monomial vector

H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

- (Under some conditions)Constraint of the form $p(x) \in SOS$ can be replaced by SOS constraints of **low dimensional** polynomials.

$$p(x) \in SOS$$

If and only if

$$p(x) = \sum_k p_k(X_k) \qquad p_k(X_k) \in SOS$$

X_k: Coupled set variables of $p(x)$

$$p(x) = B^T(x) Q B(x)$$
$$Q \in \mathcal{S}_+^n$$

If and only if

$$p(x) = \sum_k z_k^T(x) Q_k z_k(x) \qquad Q_k \in \mathcal{S}_+^{C_k} \qquad C_k < n$$

$C_k \times C_k$ matirx

$C_k \times 1$ monomial vector

H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

- (Under some conditions) Constraint of the form $X \succcurlyeq 0$ can be replaced by **PSD** constraints of smaller matrices $X_k \succcurlyeq 0$

$$X \succcurlyeq 0$$

$n \times n$ matirx

If and only if

$$X = \sum_k E_k^T X_k E_k \qquad X_k \succcurlyeq 0 \qquad C_k < n$$

$C_k \times C_k$ matirx

$C_k \times n$ constant matrix

$C_k \times C_k$ matirx

R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz, “Positive definite completions of partial Hermitian matrices,” Linear Algebra and its Applications, vol. 58, pp. 109–124, 1984.

J. Agler, W. Helton, S. McCullough, and L. Rodman, “Positive semidefinite matrices with a given sparsity pattern,” *Linear Algebra. Appl.*, vol. 107, pp. 101–149, 1988.

A. Griewank and P. L. Toint, “On the existence of convex decompositions of partially separable functions,” *Math. Prog.*, vol. 28, no. 1, pp. 25–49, 1984.

- (Under some conditions)Constraint of the form $p(x) \in SOS$can be replaced by SOS constraints of **low dimensional** polynomials.

$p(x) \in SOS$ If and only if $p(x) = \sum_k p_k(X_k)$ $p_k(X_k) \in SOS$

X_k: Coupled set variables of $p(x)$

$p(x) = B^T(x)QB(x)$
$Q \in \mathcal{S}_+^n$

If and only if

$p(x) = \sum_k z_k^T(x)Q_k z_k(x)$ $Q_k \in \mathcal{S}_+^{C_k}$ $C_k < n$

$C_k \times C_k$ matirx $C_k \times 1$ monomial vector

H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

- (Under some conditions) Constraint of the form $X \succcurlyeq 0$ can be replaced by **PSD** constraints of smaller matrices $X_k \succcurlyeq 0$

$X \succcurlyeq 0$ If and only if $X = \sum_k E_k^T X_k E_k$ $X_k \succcurlyeq 0$ $C_k < n$

$n \times n$ matirx $C_k \times C_k$ matirx $C_k \times n$ matirx $C_k \times C_k$ matirx

Example:

$$\underbrace{X = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}}_{X \succcurlyeq 0} \Longleftrightarrow \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 0.5 \end{bmatrix}}_{X_1 \succcurlyeq 0} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 0.5 & 1 \\ 1 & 2 \end{bmatrix}}_{X_2 \succcurlyeq 0} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- (Under some conditions)Constraint of the form $p(x) \in SOS$ can be replaced by SOS constraints of **low dimensional** polynomials.

$p(x) \in SOS$ If and only if $p(x) = \sum_k p_k(X_k) \qquad p_k(X_k) \in \ SOS$

X_k: Coupled set variables of $p(x)$

$$p(x) = B^T(x)QB(x)$$
$$Q \in \mathcal{S}_+^n$$

If and only if

$$p(x) = \sum_k z_k^T(x) Q_k z_k(x) \qquad Q_k \in \mathcal{S}_+^{C_k} \qquad C_k < n$$

$C_k \times C_k$ matirx

$C_k \times 1$ monomial vector

H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

- (Under some conditions) Constraint of the form $X \succcurlyeq 0$ can be replaced by **PSD** constraints of smaller matrices $X_k \succcurlyeq 0$

$X \succcurlyeq 0$ If and only if $X = \sum_k E_k^T X_k E_k \qquad X_k \succcurlyeq 0 \qquad C_k < n$

$n \times n$ matirx

$C_k \times C_k$ matirx

$C_k \times n$ matirx

$C_k \times C_k$ matirx

R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz, “Positive definite completions of partial Hermitian matrices,” Linear Algebra and its Applications, vol. 58, pp. 109–124, 1984.
J. Agler, W. Helton, S. McCullough, and L. Rodman, “Positive semidefinite matrices with a given sparsity pattern,” *Linear Algebra. Appl.*, vol. 107, pp. 101–149, 1988.
A. Griewank and P. L. Toint, “On the existence of convex decompositions of partially separable functions,” *Math. Prog.*, vol. 28, no. 1, pp. 25–49, 1984.

- (Under some conditions)Constraint of the form $p(x) \in SOS$ can be replaced by SOS constraints of **low dimensional** polynomials.

$p(x) \in SOS$ If and only if $p(x) = \sum_k p_k(X_k)$ $p_k(X_k) \in SOS$

X_k: Coupled set variables of $p(x)$

$p(x) = B^T(x)QB(x)$
$Q \in \mathcal{S}_+^n$

If and only if

$p(x) = \sum_k z_k^T(x)Q_k z_k(x)$ $Q_k \in \mathcal{S}_+^{C_k}$ $C_k < n$

$C_k \times C_k$ matirx

$C_k \times 1$ monomial vector

H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

- (Under some conditions) Constraint of the form $X \succcurlyeq 0$ can be replaced by **PSD** constraints of smaller matrices $X_k \succcurlyeq 0$

$X \succcurlyeq 0$ If and only if $X = \sum_k E_k^T X_k E_k$ $X_k \succcurlyeq 0$ $C_k < n$

$n \times n$ matirx

$C_k \times C_k$ matirx

$C_k \times n$ matirx

$C_k \times C_k$ matirx

R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz, “Positive definite completions of partial Hermitian matrices,” Linear Algebra and its Applications, vol. 58, pp. 109–124, 1984.
J. Agler, W. Helton, S. McCullough, and L. Rodman, “Positive semidefinite matrices with a given sparsity pattern,” *Linear Algebra. Appl.*, vol. 107, pp. 101–149, 1988.
A. Griewank and P. L. Toint, “On the existence of convex decompositions of partially separable functions,” *Math. Prog.*, vol. 28, no. 1, pp. 25–49, 1984.

- Results rely on **sparsity pattern** of polynomials and Matrices and its **graph representation,** and **Chordality of sparsity graph** (the classical theory of graph and cliques).

Undirected Graph

- Undirected graph $\mathcal{G}$

$\mathcal{V}$ Set of nodes of the graph

$\mathcal{E}$ Set of edges of the graph

Undirected Graph

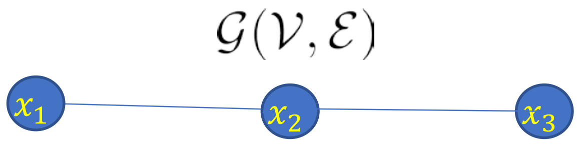
- ➢ Undirected graph $\mathcal{G}$
- $\mathcal{V}$ Set of nodes of the graph
- $\mathcal{E}$ Set of edges of the graph

- We use undirected graph to represent polynomials and symmetric matrices.

$$p(x_1, x_2, x_3\ ) = 1 + x_1 + x_3 + x_1^2 + x_1\ x_2 + x_1^2 + x_2 x_3 + x_3^2$$

Coupled variables: (x_1, x_2), (x_2, x_3)

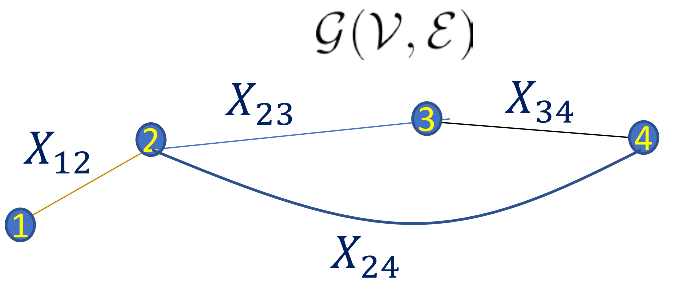
Edges between coupled variables



sparsity pattern of polynomial

$$X = \begin{array}{cccc|c} ① & ② & ③ & ④ & \\ X_{11} & X_{12} & 0 & 0 & ① \\ X_{12} & X_{22} & X_{23} & X_{24} & ② \\ 0 & X_{23} & X_{33} & X_{34} & ③ \\ 0 & X_{24} & X_{34} & X_{44} & ④ \end{array}$$

Edges: Nonzero entries of matrix

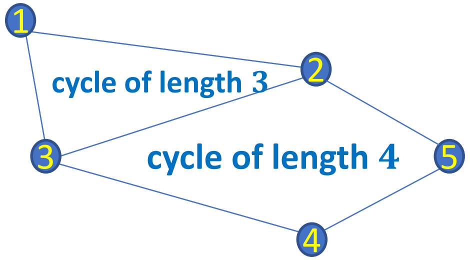


sparsity pattern of matrix

Undirected Graph

- Undirected graph $\mathcal{G}$ $\mathcal{V}$ Set of nodes of the graph $\mathcal{E}$ Set of edges of the graph

Cycle: A cycle of length k in a undirected graph is a sequence of nodes $(v_1, v_2, \dots, v_k)$ such that (v_i, v_{i+1}) $i = 1, \dots, k-1$ and (v_1, v_k) are the edges.



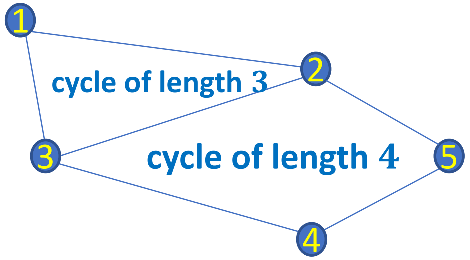
Undirected Graph

- Undirected graph $\mathcal{G}$

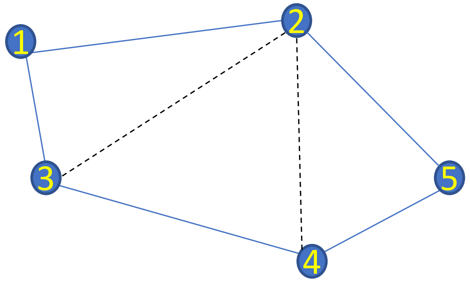
$\mathcal{V}$ Set of nodes of the graph

$\mathcal{E}$ Set of edges of the graph

Cycle: A cycle of length k in a undirected graph is a sequence of nodes $(v_1, v_2, \dots, v_k)$ such that (v_i, v_{i+1}) $i = 1, \dots, k-1$ and (v_1, v_k) are the edges.



Chord: is an edge that connects 2 nonadjacent nodes in a **cycle**.



cycle of length 5 with 2 chords

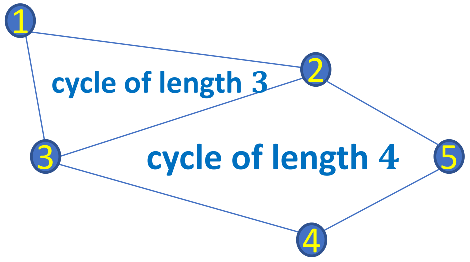
Undirected Graph

- Undirected graph $\mathcal{G}$

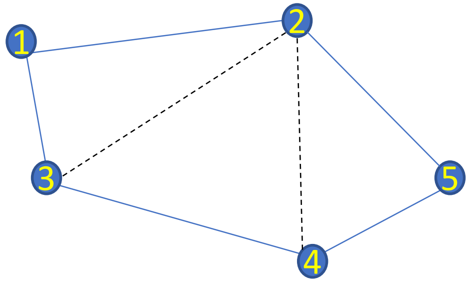
$\mathcal{V}$ Set of nodes of the graph

$\mathcal{E}$ Set of edges of the graph

Cycle: A cycle of length k in a undirected graph is a sequence of nodes $(v_1, v_2, \ldots, v_k)$ such that (v_i, v_{i+1}) $i = 1, \ldots, k-1$ and (v_1, v_k) are the edges.



Chord: is an edge that connects 2 nonadjacent nodes in a **cycle**.



cycle of length 5 with 2 chords

Chordal graph

Chordal Graph: An undirected graph is chordal if every **cycle** of the length $k \geq 4$ has a **chord**,
(if there are no cycles of length $\geq$ 4)

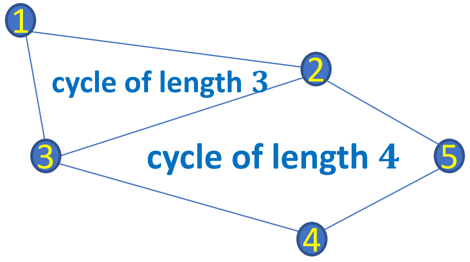
Undirected Graph

- Undirected graph $\mathcal{G}$

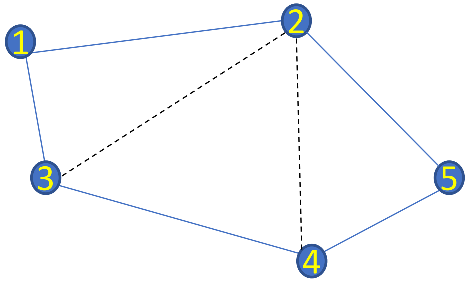
$\mathcal{V}$ Set of nodes of the graph

$\mathcal{E}$ Set of edges of the graph

Cycle: A cycle of length k in a undirected graph is a sequence of nodes $(v_1, v_2, \ldots, v_k)$ such that (v_i, v_{i+1}) $i = 1, \ldots, k-1$ and (v_1, v_k) are the edges.



Chord: is an edge that connects 2 nonadjacent nodes in a **cycle**.

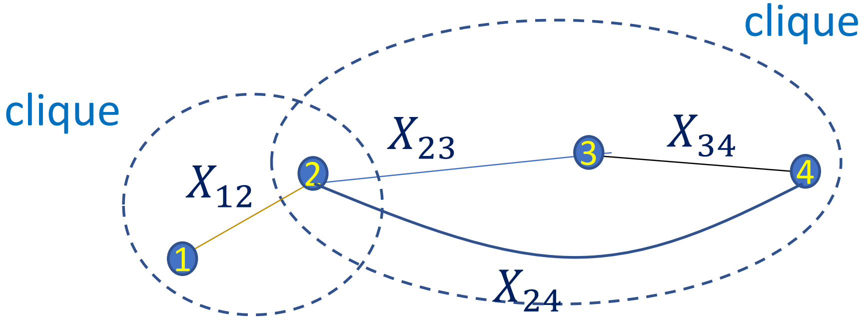


cycle of length 5 with 2 chords

Chordal graph

Chordal Graph: An undirected graph is chordal if every **cycle** of the length $k \geq 4$ has a **chord**, (if there are no cycles of length ≥ 4)

Clique: a clique of a graph is a subset of nodes that construct a complete graph (i.e. each node in the clique is connected to all the nodes in the clique.)



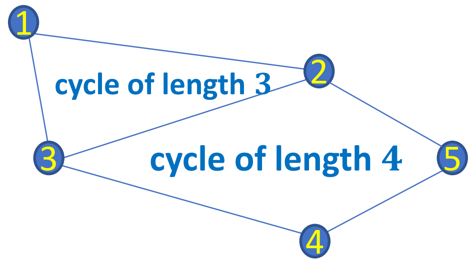
Undirected Graph

- Undirected graph $\mathcal{G}$

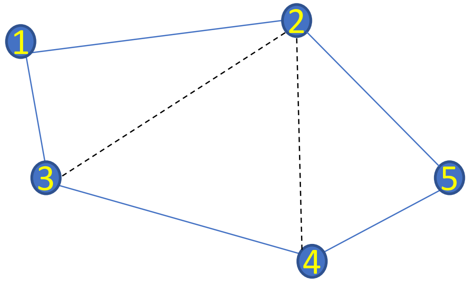
$\mathcal{V}$ Set of nodes of the graph

$\mathcal{E}$ Set of edges of the graph

Cycle: A cycle of length k in a undirected graph is a sequence of nodes $(v_1, v_2, \dots, v_k)$ such that (v_i, v_{i+1}) $i = 1, \dots, k-1$ and (v_1, v_k) are the edges.



Chord: is an edge that connects 2 nonadjacent nodes in a **cycle**.



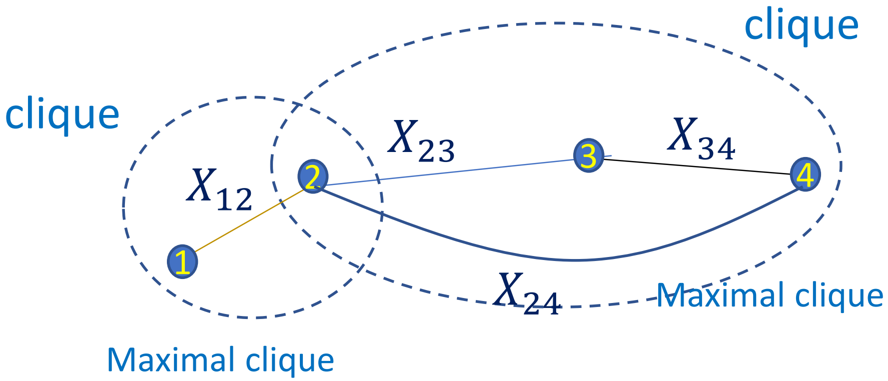
cycle of length 5 with 2 chords

Chordal graph

Chordal Graph: An undirected graph is chordal if every **cycle** of the length $k \geq 4$ has a **chord**,
(if there are no cycles of length ≥ 4)

Clique: a clique of a graph is a subset of nodes that construct a complete graph (i.e. each node in the clique is connected to all the nodes in the clique.)

Maximal Clique: a clique is maximal if it is not a subset of another clique.



Theorem

Let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be a **chordal graph**[1] with maximal cliques $\{\mathcal{C}_1, \mathcal{C}-2, ..., \mathcal{C}_t\}$.Then, Matrix $X \in \mathcal{S}^n$ with sparsity pattern $\mathcal{G}(\mathcal{V},\mathcal{E})$ is **PSD** if and only if there exist PSD matrices $X_k \in \mathcal{S}^{|\mathcal{C}_k|} \succcurlyeq 0$

$$X \succcurlyeq 0 \quad \xLeftrightarrow{\text{If and only if}} \quad X = \sum_k E^T_{\mathcal{C}_k} X_k E_{\mathcal{C}_k} \qquad X_k \in \mathcal{S}^{|\mathcal{C}_k|} \succcurlyeq 0 \qquad |\mathcal{C}_k| < n$$

Matrices constructed form the maximal Cliques

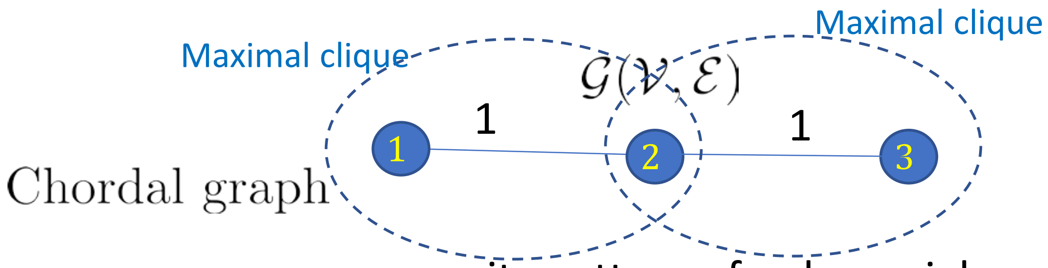
Number of the nodes in maximal Cliques

- Constraint of the form $X \succcurlyeq 0$can be replaced by **PSD** constraints of smaller matrices $X_k \succcurlyeq 0$

Example:

$$X = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

$$X \succcurlyeq 0$$



sparsity pattern of polynomial

---

[1] Perfect Elimination and Completable Graph theory, Theorem 7, R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz, “Positive definite completions of partial Hermitian matrices,” Linear Algebra and its Applications, vol. 58, pp. 109–124, 1984.

Theorem

Let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be a **chordal graph**[1] with maximal cliques $\{\mathcal{C}_1,\mathcal{C}-2,...,\mathcal{C}_t\}$.Then, Matrix $X\in\mathcal{S}^n$ with sparsity pattern $\mathcal{G}(\mathcal{V},\mathcal{E})$ is **PSD** if and only if there exist PSD matrices $X_k\in\mathcal{S}^{|\mathcal{C}_k|}\succcurlyeq 0$

$$X\succcurlyeq 0 \quad \text{If and only if} \quad X=\sum_k E_{\mathcal{C}_k}^T X_k E_{\mathcal{C}_k}$$

$X_k\in\mathcal{S}^{|\mathcal{C}_k|}\succcurlyeq 0$

Matrices constructed form the maximal Cliques

$|\mathcal{C}_k|<n$

Number of the nodes in maximal Cliques

- Constraint of the form $X\succcurlyeq 0$can be replaced by **PSD** constraints of smaller matrices $X_k\succcurlyeq 0$

Example:

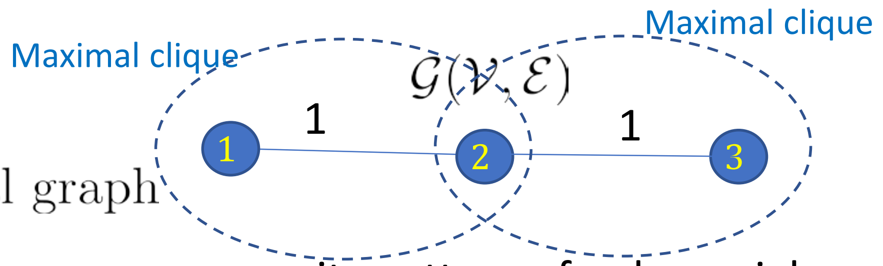
$$X=\begin{bmatrix}2&1&0\\1&1&1\\0&1&2\end{bmatrix}$$

$$X\succcurlyeq 0$$

$$\exists \quad X_1\in\mathcal{S}^{|\mathcal{C}_1|}\succcurlyeq 0 \quad X_2\in\mathcal{S}^{|\mathcal{C}_2|}\succcurlyeq 0 \quad \text{Iff} \quad X\succcurlyeq 0$$

$$X=\sum_k E_{\mathcal{C}_k}^T X_k E_{\mathcal{C}_k}$$

Chordal graph



sparsity pattern of polynomial

---

[1] Perfect Elimination and Completable Graph theory, Theorem 7, R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz, “Positive definite completions of partial Hermitian matrices,” Linear Algebra and its Applications, vol. 58, pp. 109–124, 1984.

Theorem

Let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be a **chordal graph**[1] with maximal cliques $\{\mathcal{C}_1, \mathcal{C}-2, ..., \mathcal{C}_t\}$.Then, Matrix $X \in \mathcal{S}^n$ with sparsity pattern $\mathcal{G}(\mathcal{V},\mathcal{E})$ is **PSD** if and only if there exist PSD matrices $X_k \in \mathcal{S}^{|\mathcal{C}_k|} \succcurlyeq 0$

$$X \succcurlyeq 0 \quad \Longleftrightarrow \quad X = \sum_k E^T_{\mathcal{C}_k} X_k E_{\mathcal{C}_k}$$

If and only if

$X_k \in \mathcal{S}^{|\mathcal{C}_k|} \succcurlyeq 0$

Matrices constructed form the maximal Cliques

$|\mathcal{C}_k| < n$

Number of the nodes in maximal Cliques

- Constraint of the form $X \succcurlyeq 0$can be replaced by **PSD** constraints of smaller matrices $X_k \succcurlyeq 0$

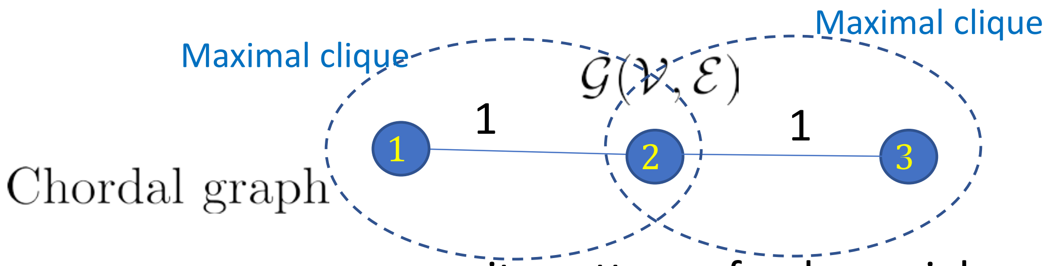
Example:

$$X = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0.5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.5 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$X \succcurlyeq 0$ $\qquad$ $X_1 \succcurlyeq 0$ $\qquad$ $X_2 \succcurlyeq 0$

$$\exists \quad X_1 \in \mathcal{S}^{|\mathcal{C}_1|} \succcurlyeq 0 \quad X_2 \in \mathcal{S}^{|\mathcal{C}_2|} \succcurlyeq 0 \quad \xLeftrightarrow{\text{Iff}} \quad X \succcurlyeq 0$$

$$X = \sum_k E^T_{\mathcal{C}_k} X_k E_{\mathcal{C}_k}$$



sparsity pattern of polynomial

---

[1] 1: Perfect Elimination and Completable Graph theory, Theorem 7, R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz, “Positive definite completions of partial Hermitian matrices,” Linear Algebra and its Applications, vol. 58, pp. 109–124, 1984.

Theorem

Let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be a **chordal graph**[1] with maximal cliques $\{\mathcal{C}_1,\mathcal{C}-2,...,\mathcal{C}_t\}$.Then, Matrix $X\in\mathcal{S}^n$ with sparsity pattern $\mathcal{G}(\mathcal{V},\mathcal{E})$

is **PSD** if and only if there exist PSD matrices $X_k\in\mathcal{S}^{|\mathcal{C}_k|}\succcurlyeq 0$

$$X\succcurlyeq 0 \xLeftrightarrow{\text{If and only if}} X=\sum_k E_{\mathcal{C}_k}^T X_k E_{\mathcal{C}_k}$$

$X_k\in\mathcal{S}^{|\mathcal{C}_k|}\succcurlyeq 0$ Matrices constructed form the maximal Cliques

$|\mathcal{C}_k|<n$ $\downarrow$ Number of the nodes in maximal Cliques

- Constraint of the form $X\succcurlyeq 0$can be replaced by **PSD** constraints of smaller matrices $X_k\succcurlyeq 0$

Example:

(**Appendix IV**)

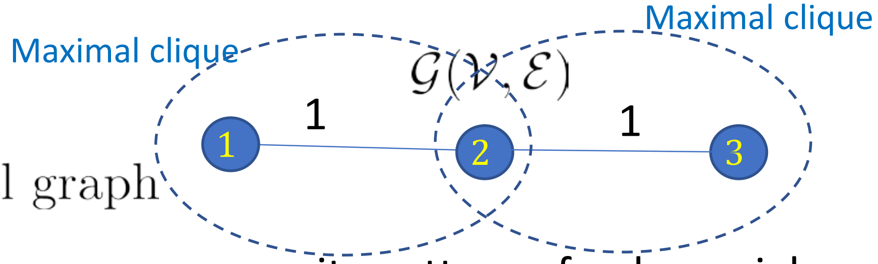
$$X=\begin{bmatrix}2&1&0\\1&1&1\\0&1&2\end{bmatrix} \Longleftrightarrow \begin{bmatrix}2&1&0\\1&1&1\\0&1&2\end{bmatrix}=\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}\underset{X_1\succcurlyeq 0}{\begin{bmatrix}2&1\\1&0.5\end{bmatrix}}\begin{bmatrix}1&0&0\\0&1&0\end{bmatrix}+\begin{bmatrix}0&0\\1&0\\0&1\end{bmatrix}\underset{X_2\succcurlyeq 0}{\begin{bmatrix}0.5&1\\1&2\end{bmatrix}}\begin{bmatrix}0&1&0\\0&0&1\end{bmatrix}$$

$$X\succcurlyeq 0$$

$$\exists \quad X_1\in\mathcal{S}^{|\mathcal{C}_1|}\succcurlyeq 0 \quad X_2\in\mathcal{S}^{|\mathcal{C}_2|}\succcurlyeq 0 \xLeftrightarrow{\text{Iff}} X\succcurlyeq 0$$

$$X=\sum_k E_{\mathcal{C}_k}^T X_k E_{\mathcal{C}_k}$$

Chordal graph



sparsity pattern of polynomial

---

[1] Perfect Elimination and Completable Graph theory, Theorem 7, R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz, "Positive definite completions of partial Hermitian matrices," Linear Algebra and its Applications, vol. 58, pp. 109–124, 1984.

SDP

$$\begin{array}{ll} \underset{X}{\text{minimize}} & C \bullet X \\ \text{subject to} & A_i \bullet X = b_i \quad i = 1, \ldots, m. \\ & X \succcurlyeq 0. \qquad X \in \mathcal{S}^n \end{array}$$

Sparsity pattern of matrix X : Chordal graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$

$$\begin{array}{ll} \underset{X}{\text{minimize}} & C \bullet X \\ \text{subject to} & A_i \bullet \left(\sum_k E_{\mathcal{C}_k}^T X_k E_{\mathcal{C}_k} \right) = b_i \quad i = 1, \ldots, m. \\ & X_k \succcurlyeq 0, k = 1, 2, \ldots \qquad X_k \in \mathcal{S}^{|\mathcal{C}_k|} \end{array}$$

Theorem

Let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be a **chordal graph** obtained from the polynomial $p(x)$ with maximal cliques $\{\mathcal{C}_1, \mathcal{C}-2, ..., \mathcal{C}_t\}$

Then, polynomial $p(x)$ is SOS if and only if:

$$p(x) \in SOS \quad \xLeftrightarrow{\text{If and only if}} \quad p(x) = \sum_k p_k(X_k) \qquad p_k(X_k) \in \ SOS$$

X_k: Nodes in clique $\mathcal{C}_k$

- Constraint of the form $p(x) \in SOS$ can be replaced by SOS constraints on **low dimensional** polynomials.

Theorem

Let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be a **chordal graph** obtained from the polynomial $p(x)$ with maximal cliques $\{\mathcal{C}_1, \mathcal{C}-2, \ldots, \mathcal{C}_t\}$

Then, polynomial $p(x)$ is SOS if and only if:

$p(x) \in SOS$ ⟺ If and only if ⟺ $p(x) = \sum_k p_k(X_k)$ $\quad p_k(X_k) \in \ SOS$

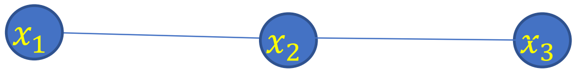
X_k: Nodes in clique $\mathcal{C}_k$

- Constraint of the form $p(x) \in SOS$ can be replaced by SOS constraints on **low dimensional** polynomials.

$$p(x_1, x_2, x_3\ ) = 2(1 + x_1 + x_3 + x_1^2 + x1\ x_2 + x_1^2 + x_2x_3 + x_3^2)$$

Coupled variables: (x_1, x_2), (x_2, x_3)

Edges between coupled variables



Polynomial with sparsity pattern $\mathcal{G}(\mathcal{V},\mathcal{E})$

Theorem

Let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be a **chordal graph** obtained from the polynomial $p(x)$ with maximal cliques $\{\mathcal{C}_1, \mathcal{C}-2, ..., \mathcal{C}_t\}$

Then, polynomial $p(x)$ is SOS if and only if:

$$p(x) \in SOS \quad \Longleftrightarrow \quad p(x) = \sum_k p_k(X_k) \qquad p_k(X_k) \in \ SOS$$

If and only if

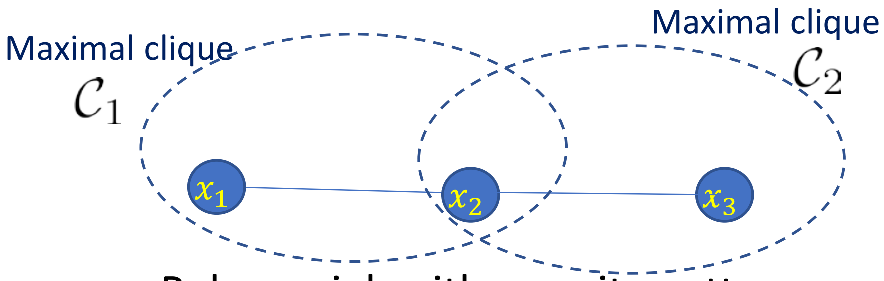
X_k: Nodes in clique $\mathcal{C}_k$

- Constraint of the form $p(x) \in SOS$ can be replaced by SOS constraints on **low dimensional** polynomials.

$$p(x_1, x_2, x_3\ ) = 2(1 + x_1\ + x_3 + x_1^2\ + x1\ x_2 + x_1^2\ + x_2x_3\ + x_3^2)$$

Coupled variables: (x_1, x_2), (x_2, x_3)

Edges between coupled variables



Polynomial with sparsity pattern

$\mathcal{G}(\mathcal{V},\mathcal{E})$

Theorem

Let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be a **chordal graph** obtained from the polynomial $p(x)$ with maximal cliques $\{\mathcal{C}_1, \mathcal{C}-2, ..., \mathcal{C}_t\}$

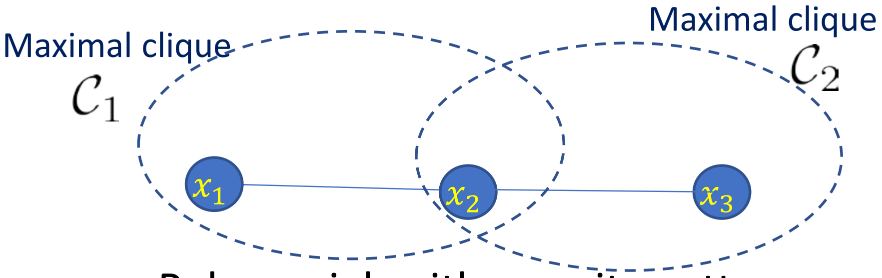
Then, polynomial $p(x)$ is SOS if and only if:

$$p(x) \in SOS \xLeftrightarrow{\text{If and only if}} p(x) = \sum_k p_k(X_k) \quad p_k(X_k) \in \; SOS$$

X_k: Nodes in clique $\mathcal{C}_k$

- ➢ Constraint of the form $p(x) \in SOS$ can be replaced by SOS constraints on **low dimensional** polynomials.

$$p(x_1, x_2, x_3\ ) = 2(1 + x_1 + x_3 + x_1^2 + x1\ x_2 + x_1^2 + x_2x_3 + x_3^2)$$



Coupled variables: (x_1, x_2), (x_2, x_3)
Edges between coupled variables

Polynomial with sparsity pattern
$\mathcal{G}(\mathcal{V},\mathcal{E})$

$$p(x_1, x_2, x_3\ ) \in SOS \quad iff \quad p(x_1, x_2) = p_1(x_1, x_2) \in SOS + p_2(x_2, x_3) \in SOS$$

$$p(x_1, x_2, x_3) = \overbrace{(1 + x_1)^2 + (x_1 + x_2)^2} + \overbrace{(1 + x_3^2)^2 + (x_2 + x_3)^2}$$

$\mathcal{G}(\mathcal{V},\mathcal{E})$ is a Chordal graph

Theorem

Let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be a **chordal graph** obtained from the polynomial $p(x)$ with maximal cliques $\{\mathcal{C}_1, \mathcal{C} - 2, \ldots, \mathcal{C}_t\}$

Then, polynomial $p(x)$ is SOS if and only if:

$$p(x) \in SOS \quad \xLeftrightarrow{\text{If and only if}} \quad p(x) = \sum_k p_k(X_k) \qquad p_k(X_k) \in \ SOS$$

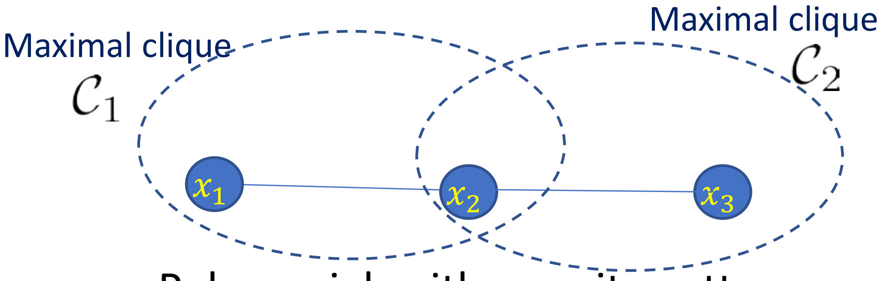
X_k: Nodes in clique $\mathcal{C}_k$

- Constraint of the form $p(x) \in SOS$ can be replaced by SOS constraints on **low dimensional** polynomials.

$$p(x_1, x_2, x_3\ ) = 2(1 + x_1 + x_3 + x_1^2 + x1\ x_2 + x_1^2 + x_2 x_3 + x_3^2)$$

Coupled variables: (x_1, x_2), (x_2, x_3)

Edges between coupled variables



Polynomial with sparsity pattern $\mathcal{G}(\mathcal{V},\mathcal{E})$

$$p(x_1, x_2, x_3\ ) \in SOS \quad iff \quad p(x_1, x_2) = p_1(x_1, x_2) \in SOS + p_2(x_2, x_3) \in SOS$$

$$p(x_1, x_2, x_3) = \overbrace{(1 + x_1)^2 + (x_1 + x_2)^2} + \overbrace{(1 + x_3^2)^2 + (x_2 + x_3)^2}$$

$\mathcal{G}(\mathcal{V},\mathcal{E})$ is a Chordal graph

$$p(x_1, x_2, x_3) \in SOS \longrightarrow p(x_1, x_2, x_3) \in SSOS$$

Unconstrained optimization

$$\underset{x}{\text{minimize}} \quad p(x)$$

SOS Program:

$$\begin{aligned} \underset{Q \in \mathcal{S}^n, \gamma}{\text{maximize}} \quad & \gamma \\ \text{subject to} \quad & p(x) - \gamma \in SOS \end{aligned}$$

SSOS Program:

$$\begin{aligned} \underset{Q \in \mathcal{S}^n, \gamma}{\text{maximize}} \quad & \gamma \\ \text{subject to} \quad & p(x) - \gamma \in SSOS \end{aligned}$$

Unconstrained optimization

$$\underset{x}{\text{minimize}} \quad p(x)$$

SOS Program:

$$\begin{aligned} \underset{Q\in\mathcal{S}^n,\gamma}{\text{maximize}} \quad & \gamma \\ \text{subject to} \quad & p(x)-\gamma \in SOS \end{aligned}$$

SSOS Program:

$$\begin{aligned} \underset{Q\in\mathcal{S}^n,\gamma}{\text{maximize}} \quad & \gamma \\ \text{subject to} \quad & p(x)-\gamma \in SSOS \end{aligned}$$

Constrained optimization

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & p(x) \\ \text{subject to} \quad & g_i(x) \geq 0,\ \ i = 1, \ldots, n \end{aligned}$$

SOS Program:

$$\begin{aligned} \underset{\gamma,\sigma_i}{\text{maximize}} \quad & \gamma \\ \text{subject to} \quad & p(x)-\gamma-\sum_{i=1}^{m}\sigma_i(x)g_i(x) \in SOS \\ & \sigma_i(x) \in SOS_{2d_i},\ \ i = 0, \ldots, m \end{aligned}$$

SSOS Program:

$$\begin{aligned} \underset{\gamma,\sigma_i}{\text{maximize}} \quad & \gamma \\ \text{subject to} \quad & p(x)-\gamma-\sum_{i=1}^{m}\sigma_i(x)g_i(x) \in SSOS \\ & \sigma_i(x) \in SOS_{2d_i},\ \ i = 0, \ldots, m \end{aligned}$$

should preserve the correlative sparsity of g_i

$$p(x) - \gamma - \sum_{i=1}^{m} \sigma_i(x) g_i(x) \in SSOS$$

$$\sigma_i(x) \in SOS_{2d_i}, \; i = 0, \ldots, m$$

- ➢ $\sigma_i(x)$ should preserve the correlative sparsity of $g_i(x)$

- ➢ Example:

  $g_i(\tilde{x})$: is a polynomial in terms of subset of variables $\tilde{x}$
  $\sigma_i(\tilde{x})$: SOS polynomial in terms of variables $\tilde{x}$

More information:

- **Section 4.2:** H. Waki, S. Kim, M. Kojima, and M. Muramatsu, "Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity," SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

- **Lemma 3**: , Tillmann Weisser, Jean B. Lasserre, Kim-Chuan Toh., "Sparse-BSOS: a bounded degree SOS hierarchy for large scale polynomial optimization with sparsity", Math. Prog. Comp. (2018) 10:1–32

Example: https://github.com/jasour/rarnop19/blob/master/Lecture6_modified-SOS/Sparse_SOS/Example_SSOS_compare_Cons.m

Sparse SOS using Yalmip

1) Copy “corrsparsity.m” to the folder of /modules/sos, and replace the original corrsparsity.m.

https://github.com/zhengy09/sos_csp

2) Add the “`ops.sos.csp = 1`” to the Yalmip SOS optimization code.

- Zheng, Y., Fantuzzi, G., & Papachristodoulou, A. “Sparse sum-of-squares (SOS) optimization: A bridge between DSOS/SDSOS and SOS optimization for sparse polynomials”, arXiv preprint arXiv:1807.05463. 2018
- Zheng, Y., Fantuzzi, G., & Papachristodoulou, A. (2018, December). Decomposition and completion of sum-of-squares matrices. In 2018 IEEE Conference on Decision and Control (CDC) (pp. 4026-4031). IEEE.

---

sparsePOP 3.03 (MATLAB Package)

This package also provides the optimal solution x^* of SSOS optimization.

https://sourceforge.net/projects/sparsepop/

- H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

Example 1: Unconstrained Optimization

$$f_{cs}(x) = \sum_{i \in J} \left((x_i + 10x_{i+1})^2 + 5(x_{i+2} - x_{i+3})^2 + (x_{i+1} - 2x_{i+2})^4 + 10(x_i - 10x_{i+3})^4 \right),$$

$$J = \{1, 3, 5, \ldots, n-3\}$$

(Number of Clique)*(Size 0f the Clique)

cpu time (sparseSOS)

cpu time (SOS)

Number of variables

n	cl.str	ϵ_{obj}	sparse	dense
16	3*14	3.5e-7	0.6	3059.5
40	3*38	8.4e-7	1.4	—
100	3*98	5.5e-7	3.8	—
200	3*198	3.0e-7	8.4	—
400	3*398	3.6e-7	19.3	—

Objective function

Example 2: Unconstrained Optimization

$$f_{\mathrm{Bb}}(x) = \sum_{i=1}^{n} \left(x_i(2 + 5x_i^2) + 1 - \sum_{j \in J_i} (1 + x_j)x_j \right)^2,$$

$$J_i = \{ j \mid j \neq i, \max(1, i-5) \leq j \leq \min(n, i+1) \}.$$

Broyden banded function				
n	cl.str	ϵ_{obj}	sparse	dense
6	6*1	8.0e-9	11.3	11.6
7	7*1	1.9e-8	69.5	69.5
8	7*2	2.8e-8	164.1	373.7
9	7*3	9.1e-8	240.3	1835.6
10	7*4	6.2e-8	348.7	8399.4

- H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

Illustrative Example:

$$\mathrm{P}^* = \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad = 5 + \sum_{i=1}^{n}(x_i - 1)^2 \qquad\qquad p^* = 5, \quad x^* = [1, 1, ..., 1]^T \in \mathbb{R}^n$$

Number of variables   Polynomial of order 2

- **SOS:** **Variables:200** **Relaxation Order=1** **time= 286.5458 (s)** $\boldsymbol{p^*}$**=5** **sdp solver: mosek**
- **SDSOS:** **Variables:200** **Relaxation Order=1** **time= 3.6338 (s)** $\boldsymbol{p^*}$**=5** **sdp solver: mosek**
- **DSOS:** **Variables:200** **Relaxation Order=1** **time=2.6824 (s)** $\boldsymbol{p^*}$**=5** **sdp solver: mosek**

- **Spars SOS:** **Variables:200** **Relaxation Order=1** **time=0.2374 (s)** $\boldsymbol{p^*}$**=5** **sdp solver: mosek**
- **SparsPOP:** **Variables:200** **Relaxation Order=1** **time=0.95 (s)** $\boldsymbol{p^*}$**=5** $\boldsymbol{x^*=[1,\dots,1]}$ **sdpt3**

https://github.com/jasour/rarnop19/blob/master/Lecture6_modified-SOS/SDSOS-DSOS/Example_SDSOS_compare_Uncons.m

https://github.com/jasour/rarnop19/blob/master/Lecture6_modified-SOS/Sparse_SOS/Example_SSOS_compare_Uncons.m

Topics

1) (Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS)
   Modified SOS optimization that results in LP and Second order cone program

2) Bounded degree SOS (BSOS)
   Modified SOS optimization that results in smaller SDP's.

3) Spars Sum-of-Squares Optimization (SSOS)
   Takes advantage of sparsity of the original problem to generate smaller SDP.

4) Sparse Bounded Degree Sum-of-Squares Optimization (SBSOS)
   Combination of 2 and 3

Sparse Bounded Degree Sum-of-Squares Optimization (SBSOS)

- ➢ **Combines** Bounded degree SOS (BSOS) and Chordal-Sparse SOS.
- Tillmann Weisser, Jean B. Lasserre, Kim-Chuan Toh., “Sparse-BSOS: a bounded degree SOS hierarchy for large scale polynomial optimization with sparsity”, Math. Prog. Comp. (2018) 10:1–32

- ➢ Takes advantages of sparsity of the original problem to reduce the size of the bounded degree SOS.

- ➢ It relies on “Running Intersection Property” (Chordal sparsity of the graph)
- M. Tacchi, T. Weisser, J. B. Lasserre, D. Henrion,”Exploiting Sparsity for Semi-Algebraic Set Volume Computation”, https://arxiv.org/abs/1902.02976
- J. R. S. Blair, B. Peyton. An introduction to chordal graphs and clique trees. Pages 1–29 in Graph Theory and Sparse Matrix Computation, Springer, New York, 1993

- Example: https://github.com/jasour/rarnop19/blob/master/Lecture6_modified-SOS/Sparse_Bounded_Degree_SOS/SBSOS_Example1.m
- ➢ MATLAB Code

https://github.com/tweisser/Sparse_BSOS

This package also provides the optimal solution x^* of SBSOS optimization.

Example 1: Constrained Optimization (Chained Singular Function)

$$f := \sum_{j\in H}\left((x_j+10x_{j+1})^2+5(x_{j+2}-x_{j+3})^2+(x_{j+1}-2x_{j+2})^4+10(x_j-x_{j+3})^4\right)$$

$$H := \{2i-1 \ : \ i = 1,\ldots,n/2-1\}$$

$$\mathbf{K} = \left\{x\in\mathbb{R}^n \ : \ 1-\sum_{i\in I_\ell}x_i \geq 0, \quad \ell = 1,\ldots,p; \quad x_i\geq 0, \quad i = 1,\ldots,n\right\},$$

- Tillmann Weisser, Jean B. Lasserre, Kim-Chuan Toh., “Sparse-BSOS: a bounded degree SOS hierarchy for large scale polynomial optimization with sparsity”, Math. Prog. Comp. (2018) 10:1–32

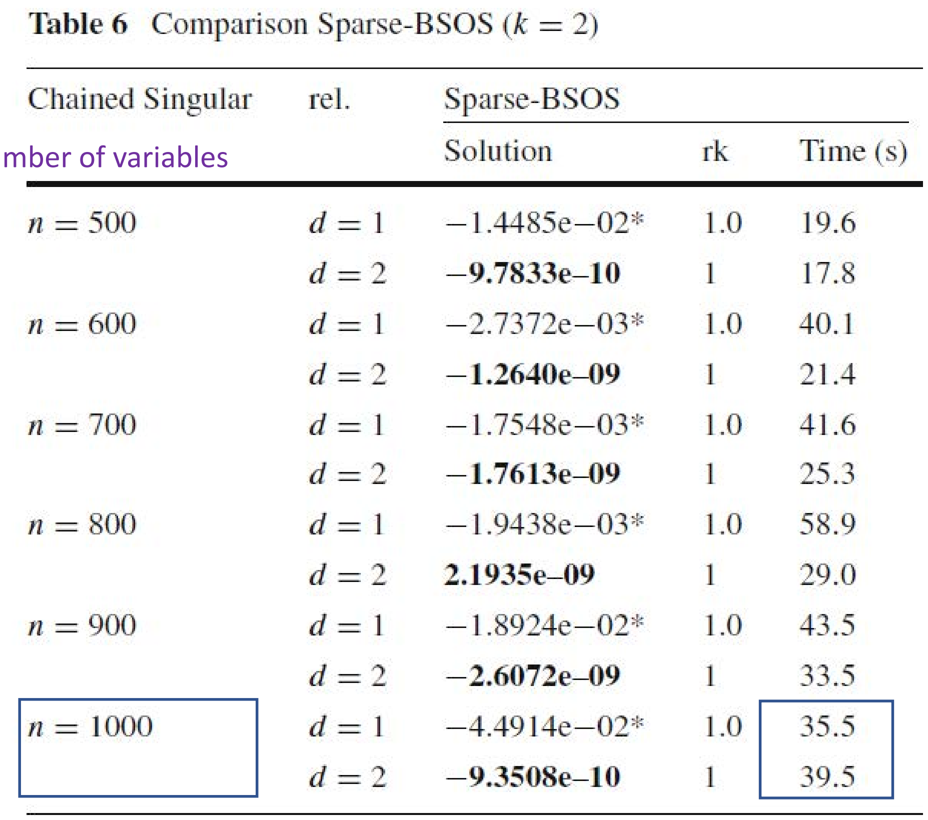


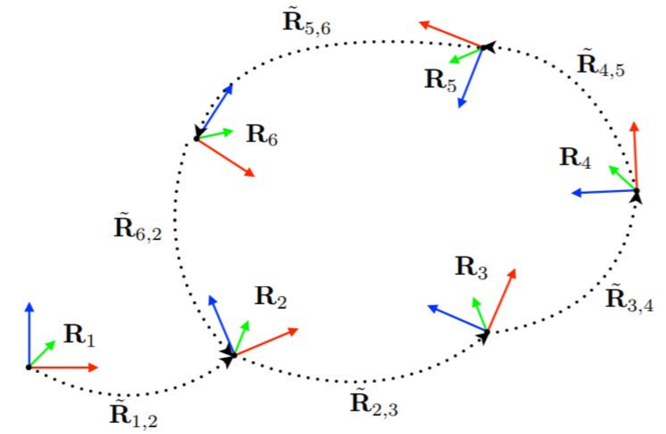
Table 6 Comparison Sparse-BSOS ($k = 2$)

Chained Singular	rel.	Sparse-BSOS		
		Solution	rk	Time (s)
$n = 500$	$d = 1$	−1.4485e−02*	1.0	19.6
	$d = 2$	**−9.7833e–10**	1	17.8
$n = 600$	$d = 1$	−2.7372e−03*	1.0	40.1
	$d = 2$	**−1.2640e–09**	1	21.4
$n = 700$	$d = 1$	−1.7548e−03*	1.0	41.6
	$d = 2$	**−1.7613e–09**	1	25.3
$n = 800$	$d = 1$	−1.9438e−03*	1.0	58.9
	$d = 2$	**2.1935e–09**	1	29.0
$n = 900$	$d = 1$	−1.8924e−02*	1.0	43.5
	$d = 2$	**−2.6072e–09**	1	33.5
$n = 1000$	$d = 1$	−4.4914e−02*	1.0	35.5
	$d = 2$	**−9.3508e–10**	1	39.5

Number of variables

Application

M. Giamou, F. Maric, V. Peretroukhin, J. Kelly “Sparse Bounded Degree Sum of Squares Optimization for Certifiably Globally Optimal Rotation Averaging”, https://arxiv.org/pdf/1904.01645.pdf, 2019



1) (Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS)

Modified SOS optimization that results in LP and Second order cone program

2) Bounded degree SOS (BSOS)

Modified SOS optimization that results in smaller SDP’s.

3) Spars Sum-of-Squares Optimization (SSOS)

Takes advantage of sparsity of the original problem to generate smaller SDP.

4) Sparse Bounded Degree Sum-of-Squares Optimization (SBSOS)

Combination of 2 and 3

(Scaled) Diagonally Dominant Sum-of-Squares Optimization (DSOS, SDSOS)

- A. Ahmadi and A. Majumdar," DSOS and SDSOS Optimization: More Tractable Alternatives to Sum of Squares and Semidefinite Optimization", *SIAM Journal on Applied Algebraic Geometry, 2019.*

  Code: https://github.com/anirudhamajumdar/spotless/tree/spotless_isos

Bounded Degree Sum-of-Squares Optimization (BSOS)

- Jean B. Lasserre, Kim-Chuan Toh, Shouguang Yang, "A bounded degree SOS hierarchy for polynomial optimization", EURO Journal on Computational Optimization March 2017, Volume 5, Issue 1–2, pp 87–117

  Code: https://github.com/tweisser/Sparse_BSOS

Sparse Sum-of-Squares Optimization (SSOS)

- H. Waki, S. Kim, M. Kojima, and M. Muramatsu, "Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity," SIAM Journal on Optimization, vol. 17, no. 1, pp. 218–242, 2006.

  Code: https://sourceforge.net/projects/sparsepop/

- Zheng, Y., Fantuzzi, G., & Papachristodoulou, A. "Sparse sum-of-squares (SOS) optimization: A bridge between DSOS/SDSOS and SOS optimization for sparse polynomials", arXiv preprint arXiv:1807.05463. 2018

  Code: https://github.com/zhengy09/sos_csp

Sparse Bounded Degree Sum-of-Squares Optimization (SBSOS)

- Tillmann Weisser, Jean B. Lasserre, Kim-Chuan Toh., "Sparse-BSOS: a bounded degree SOS hierarchy for large scale polynomial optimization with sparsity", Math. Prog. Comp. (2018) 10:1–32

  Code: https://github.com/tweisser/Sparse_BSOS

Appendix I: SDSOS/DSOS Polynomials

Polynomial $p(x) \in \mathbb{R}[x]$ **Nonnegative Polynomial** $p(x) \geq 0$

Sum-Of-Squares Polynomials

$$p(x) \in SOS \longrightarrow p(x) = \sum_{i=1}^{\ell} h_i^2(x) \Longleftrightarrow p(x) = B(x)^T Q B(x)$$

where $h_i(x) \in \mathbb{R}[x], \quad i = 1, \ldots, \ell$

where $Q \in \mathcal{S}_+^n$

Diagonally-Dominant-Sum-Of-Squares Polynomials

$$p(x) \in DSOS$$

$$\downarrow$$

$$p(x) = \sum_i \alpha_i m_i^2(x) + \sum_{i,j} \beta_{ij}^+ (m_i(x) + m_j(x))^2 + \sum_{i,j} \beta_{ij}^- (m_i(x) - m_j(x))^2 \Longleftrightarrow p(x) = B(x)^T Q B(x)$$

for some nonnegative scalars $\alpha_i, \beta_{ij}^+, \beta_{ij}^-$ for some polynomials $m_i(x), m_i(x)$

where $Q \in \mathcal{S}_{dd}^n$

Scaled-Diagonally-Dominant-Sum-Of-Squares Polynomials

$$p(x) \in SDSOS \Longleftrightarrow p(x) = B(x)^T Q B(x)$$

where $Q \in \mathcal{S}_{sdd}^n$

$$\downarrow$$

$$p(x) = \sum_i \alpha_i m_i^2(x) + \sum_{i,j} (\hat{\beta}_{ij}^+ m_i(x) + \tilde{\beta}_{ij}^+ m_j(x))^2 + \sum_{i,j} (\hat{\beta}_{ij}^- m_i(x) - \tilde{\beta}_{ij}^- m_j(x))^2$$

for some scalars $\alpha_i \geq 0, \hat{\beta}_{ij}^+, \tilde{\beta}_{ij}^+, \hat{\beta}_{ij}^-, \tilde{\beta}_{ij}^-$ for some polynomials $m_i(x), m_i(x)$

Appendix II: Convergence of LP Relaxation

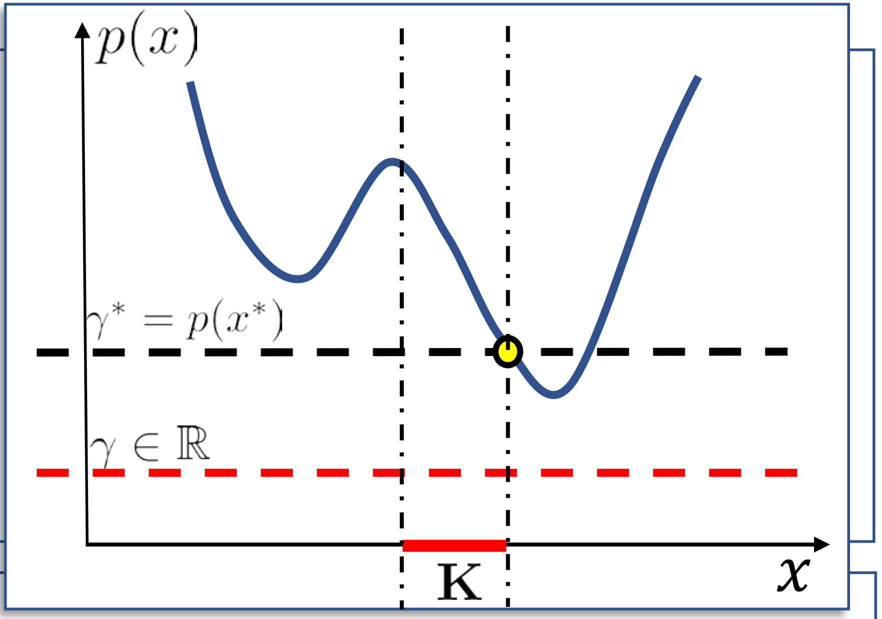
$$\mathrm{P}^* = \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad p(x)$$

$$\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, \dots, m\}$$

$$\Updownarrow$$

$$\mathrm{P}^* = \underset{\gamma\in\mathbb{R}}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad p(x) - \gamma \geq 0, \quad \forall x \in \mathbf{K} \xrightarrow[\textit{optimal solution}]{} \gamma^* = p(x^*)$$



SDP Relaxation

$$p(x) - \gamma^* = \sigma_0(x) + \textstyle\sum_{i=1}^m \sigma_i(x) g_i(x) \qquad \exists\ \gamma^* \in \mathbb{R}, \sigma_0(x) \in SOS_{2d},\ \sigma_i(x) \in SOS_{2d_i}, i = 1, \dots, m$$

$$\text{if } \gamma^* = p(x^*) = \mathrm{P}^* \xrightarrow{p(x^*) - \gamma^* = 0} \sigma_0(x^*) + \textstyle\sum_{i=1}^m \sigma_i(x^*) g_i(x^*) = 0$$

$$\mathrm{P}^* = \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad p(x)$$

$$\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, \ldots, m\}$$

$$\Updownarrow$$

$$\mathrm{P}^* = \underset{\gamma \in \mathbb{R}}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad p(x) - \gamma \geq 0, \quad \forall x \in \mathbf{K} \xrightarrow[\textit{optimal solution}]{} \gamma^* = p(x^*)$$

$$p(x) - \gamma^* = \sigma_0(x) + \textstyle\sum_{i=1}^m \sigma_i(x) g_i(x) \qquad \exists\ \gamma^* \in \mathbb{R}, \sigma_0(x) \in SOS_{2d},\ \sigma_i(x) \in SOS_{2d_i}, i = 1, \ldots, m$$

$$\text{if } \gamma^* = p(x^*) = \mathrm{P}^* \xrightarrow{p(x^*) - \gamma^* = 0} \sigma_0(x^*) + \textstyle\sum_{i=1}^m \sigma_i(x^*) g_i(x^*) = 0$$

$$\text{if} \quad x^* \in int\mathbf{K} \longrightarrow \sigma_0(x^*) + \textstyle\sum_{i=1}^m \sigma_i(x^*) \underbrace{g_i(x^*)}_{g_i(x^*) > 0} = 0 \quad \Rightarrow$$

Hence, This constraint is imposed by

$$\sigma_i(x^*) \;\; i = 0, \ldots, m$$

(The same situation for $x^* \in \partial\mathbf{K}$)

$$\begin{aligned} \mathrm{P}^* = & \underset{x}{\text{minimize}} \quad x^2 - 2x + 2 \\ & \text{subject to} \quad x \in \mathbf{K} = \{x : x(2-x) \geq 0\} \end{aligned}$$

$\Big\Downarrow$ SOS Relaxation

$$\begin{aligned} \mathrm{P}^*_{sos} = & \underset{\gamma \in \mathbb{R}, \sigma_0(x) \in SOS, \sigma_1(x) \in SOS}{\text{maximize}} \quad \gamma \\ & \text{subject to} \quad x^2 - 2x + 2 - \gamma = \sigma_0(x) + \sigma_1(x)x(2-x) \end{aligned}$$

$\Big\Downarrow$ Obtained solution (Yalmip)

$$\gamma^* = 1 \longrightarrow x^* = 1$$

$$\sigma_0(x) = (-0.291570596593 - 0.0571934472478x1 + 0.348740011438x1^2)^2 + (-0.956549252584 + 1.50888962843x1 - 0.552282590362x1^2)^2$$

$$\sigma_1(x) = (-0.653185546681 + 0.653173513801x1)^2$$

- At $x^* = 1 \in int\ \mathbf{K}$

$$p(x^*) - \gamma^* = 0 \qquad \underbrace{\sigma_0(x^*)}_{=0} + \underbrace{\sigma_1(x^*)}_{=0}\underbrace{x^*(2-x^*)}_{=1} = 0$$

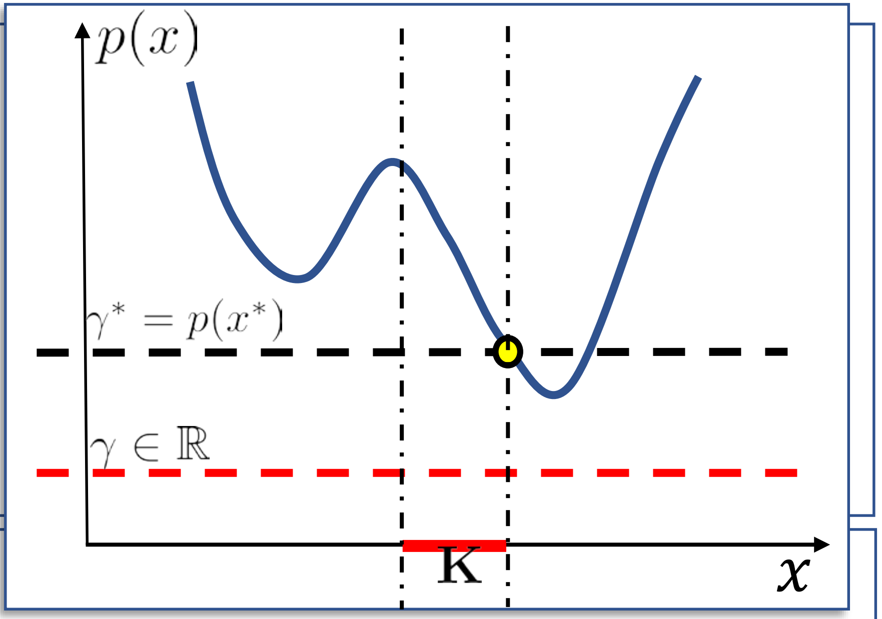
$$\mathrm{P}^* = \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad p(x)$$

subject to $\quad x \in \mathbf{K} = \{x\in\mathbb{R}^n : 0\le g_i(x)\le 1,\ \ i=1,\dots,m\}$

$$\Updownarrow$$

$$\mathrm{P}^* = \underset{\gamma\in\mathbb{R}}{\text{maximize}} \quad \gamma$$

subject to $\quad p(x)-\gamma\ge 0,\ \ \forall x\in\mathbf{K} \xrightarrow[\textit{optimal solution}]{} \gamma^* = p(x^*)$



LP Relaxation

$$p(x)-\gamma^* = \sum \lambda^*_{\alpha\beta}\, g_1^{\alpha_1}(x)\dots g_m^{\alpha_m}(x)(1-g_1(x))^{\beta_1}\dots(1-g_m(x))^{\beta_m} \qquad \exists\ \gamma^*\in\mathbb{R},\ \lambda^*_{\alpha\beta}\ge 0$$

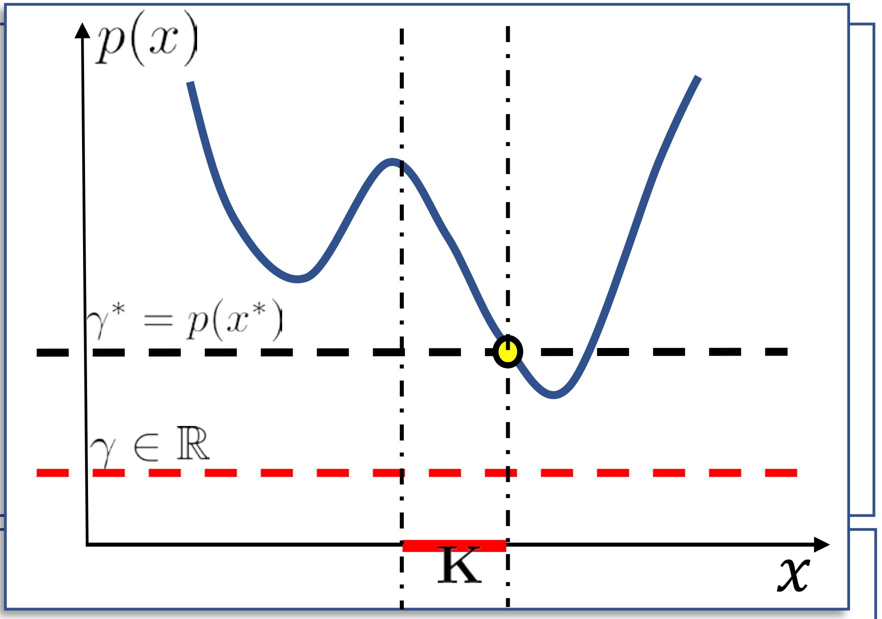
$$\text{if } \gamma^* = p(x^*) = \mathrm{P}^* \xrightarrow{p(x^*)-\gamma^*=0} \sum \lambda^*_{\alpha\beta}\, g_1^{\alpha_1}(x^*)\dots g_m^{\alpha_m}(x^*)(1-g_1(x^*))^{\beta_1}\dots(1-g_m(x^*))^{\beta_m} = 0$$

- *Section 5.4.2,* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

$$\begin{aligned} \mathrm{P}^* = &\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad p(x) \\ &\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^n : 0 \le g_i(x) \le 1,\ i = 1, \ldots, m\} \end{aligned}$$

$$\Updownarrow$$

$$\begin{aligned} \mathrm{P}^* = &\underset{\gamma\in\mathbb{R}}{\text{maximize}} \quad \gamma \\ &\text{subject to} \quad p(x) - \gamma \ge 0, \quad \forall x \in \mathbf{K} \xrightarrow[\textit{optimal solution}]{} \gamma^* = p(x^*) \end{aligned}$$



LP Relaxation

$$p(x) - \gamma^* = \sum \lambda^*_{\alpha\beta}\, g_1^{\alpha_1}(x) \ldots g_m^{\alpha_m}(x)(1 - g_1(x))^{\beta_1} \ldots (1 - g_m(x))^{\beta_m} \qquad \exists\ \gamma^* \in \mathbb{R},\ \lambda^*_{\alpha\beta} \ge 0$$

$$\text{if } \gamma^* = p(x^*) = \mathrm{P}^* \xrightarrow{p(x^*) - \gamma^* = 0} \sum \lambda^*_{\alpha\beta}\, g_1^{\alpha_1}(x^*) \ldots g_m^{\alpha_m}(x^*)(1 - g_1(x^*))^{\beta_1} \ldots (1 - g_m(x^*))^{\beta_m} = 0$$

$$\text{if } \ x^* \in int\mathbf{K} \longrightarrow \sum \lambda_{\alpha\beta}\, \underbrace{g_1^{\alpha_1}(x^*)}_{g_i(x^*) > 0} \ldots g_m^{\alpha_m}(x^*)\underbrace{(1 - g_1(x^*))^{\beta_1}}_{1 - g_i(x^*) > 0} \ldots (1 - g_m(x^*))^{\beta_m} > 0$$

- *Section 5.4.2,* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

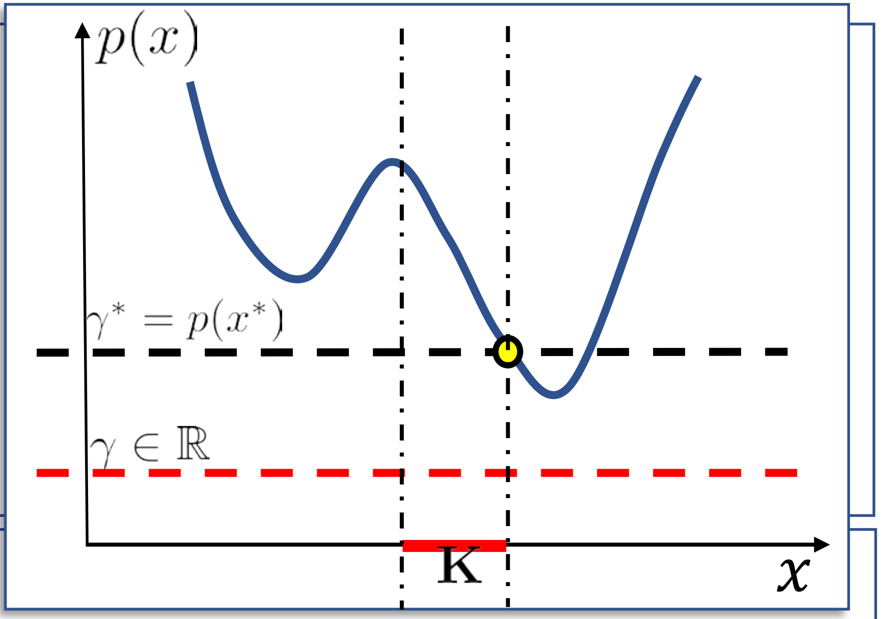
$$\mathrm{P}^* = \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad p(x)$$

$$\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^n : 0 \le g_i(x) \le 1,\ i = 1, ..., m\}$$

⇕

$$\mathrm{P}^* = \underset{\gamma\in\mathbb{R}}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad p(x) - \gamma \ge 0, \quad \forall x \in \mathbf{K} \xrightarrow[\textit{optimal solution}]{} \gamma^* = p(x^*)$$



LP Relaxation

$$p(x) - \gamma^* = \sum \lambda^*_{\alpha\beta} g_1^{\alpha_1}(x)...g_m^{\alpha_m}(x)(1-g_1(x))^{\beta_1}...(1-g_m(x))^{\beta_m} \qquad \exists\ \gamma^* \in \mathbb{R},\ \lambda^*_{\alpha\beta} \ge 0$$

$$\text{if } \gamma^* = p(x^*) = \mathrm{P}^* \xrightarrow{p(x^*)-\gamma^*=0} \sum \lambda^*_{\alpha\beta} g_1^{\alpha_1}(x^*)...g_m^{\alpha_m}(x^*)(1-g_1(x^*))^{\beta_1}...(1-g_m(x^*))^{\beta_m} = 0$$

$$\text{if} \quad x^* \in int\mathbf{K} \longrightarrow \sum \lambda_{\alpha\beta} \underbrace{g_1^{\alpha_1}(x^*)}_{g_i(x^*)>0}...g_m^{\alpha_m}(x^*)\underbrace{(1-g_1(x^*))}_{1-g_i(x^*)>0}{}^{\beta_1}...(1-g_m(x^*))^{\beta_m} > 0 \xrightarrow[\sum_{j=1}^m \alpha_j+\beta_j \le d]{d\to\infty}$$

Convergences to zero

- Hence, γ^* (optimal solution of the original problem) can not be attained.
- convergence cannot be finite $lim_{d\to\infty}\ \mathbf{P}_L^{*d} = \mathbf{P}^*$

- *Section 5.4.2,* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Example:

$$\begin{aligned} \mathrm{P}^* = &\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad p(x) = x^2 - x \\ &\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^n : g_1(x) = x \geq 0,\ g_2(x) = 1 - x \geq 0\} \end{aligned} \quad \Longrightarrow \quad \begin{aligned} &x^* = \tfrac{1}{2} \in int\mathbf{K} \\ &p(x^*) = -0.25 \end{aligned}$$

LP Relaxation

$$\begin{aligned} \mathbf{P}_L^{*i} = &\underset{\gamma,\lambda_{\alpha\beta}\geq 0}{\text{maximize}} \quad \gamma \\ &\text{subject to} \;\; p(x) - \gamma = \sum_{\substack{\forall\, \alpha,\beta \in \mathbb{N}^m \\ \sum_{j=1}^m \alpha_j + \beta_j \leq i}} \lambda_{\alpha\beta}\, g_1^{\alpha_1}(x) \ldots g_m^{\alpha_m}(x)(1 - g_1(x))^{\beta_1} \ldots (1 - g_m(x))^{\beta_m} \end{aligned}$$

Slow monotone convergence to -0.25:

$$\mathbf{P}_L^{*2} = -\frac{1}{3} \qquad \mathbf{P}_L^{*4} = -\frac{1}{3} \qquad \mathbf{P}_L^{*6} = -0.3 \qquad \mathbf{P}_L^{*10} = -0.27 \qquad \mathbf{P}_L^{*15} = -0.2695$$

Example:

$$\begin{aligned} \mathrm{P}^* = &\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad p(x) = x - x^2 \\ &\text{subject to} \quad x \in \mathbf{K} = \{x \in \mathbb{R}^n : g_1(x) = x \geq 0,\ g_2(x) = 1 - x \geq 0\} \end{aligned} \quad \Longrightarrow \quad \begin{aligned} &x^* = 0, 1 \in \partial\mathbf{K} \\ &p(x^*) = 0 \end{aligned}$$

LP Representation

$$p(x) - \gamma^* = g_1(x) g_2(x) \longrightarrow x - x^2 = x(1 - x)$$

Some of $g_i(x)'s$, $(1 - g_i(x))'s$ are zero. Hence, finite convergence **can** take place.

- *Example 5.5.* Jean Bernard Lasserre, “Moments, Positive Polynomials and Their Applications” Imperial College Press Optimization Series, V. 1, 2009.

Appendix III: Bounded Degree SOS Lagrangian Perspective

To gain more insight into how the BSOS optimization works, consider the following Nonlinear optimization and its dual:

$$\mathbf{P}^* = \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad p(x)$$

$$\text{subject to} \quad g_1^{\alpha_1}(x) ... g_m^{\alpha_m}(x)(1-g_1(x))^{\beta_1} ... (1-g_m(x))^{\beta_m} \geq 0, \qquad \forall \textstyle\sum_{j=1}^m \alpha_j + \beta_j \leq d$$

Lagrange multipliers

Lagrange function $$L(\lambda, x) = p(x) - \sum_{\sum_{j=1}^m \alpha_j + \beta_j \leq d} \lambda_{\alpha\beta}\, g_1^{\alpha_1}(x) ... g_m^{\alpha_m}(x)(1-g_1(x))^{\beta_1} ... (1-g_m(x))^{\beta_m}$$

Dual Optimization: $$\mathbf{P}^*_{dual} = \underset{\lambda}{\text{maximize}} \ \boxed{\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad L(x,\lambda)}$$ nonlinear optimization

$$\text{subject to } \lambda \geq 0$$

To solve $\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad L(x,\lambda)$, we can use SOS relaxation.

$$\underset{\gamma}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad L(x,\lambda) - \gamma \geq 0$$

⇒

$$\boxed{\begin{aligned} &\underset{\gamma, Q_0 \succcurlyeq 0}{\text{maximize}} \quad \gamma \\ &\qquad \text{subject to} \quad L(x,\lambda) - \gamma \in SOS_k \end{aligned}}$$

⇒ This results in BSOS formulation

$$\underset{\gamma, Q_0 \succcurlyeq 0}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad L(x,\lambda) - \gamma \in SOS_k$$

- For $k = 0$, this is results in “Krivine-Stengle’s Positivity Certificate” based LP.

(brutal simplification of $\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad L(x,\lambda)$)

- For $k > 0$, this is results in “BSOS” relaxation.

(tractable simplification of $\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad L(x,\lambda)$)

$$\underset{\gamma, Q_0 \succcurlyeq 0}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad L(x,\lambda) - \gamma \in SOS_k$$

$$\Downarrow$$

- For $k = 0$, this is results in “Krivine-Stengle’s Positivity Certificate” based LP.

(brutal simplification of $\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad L(x,\lambda)$)

- For $k > 0$, this is results in “BSOS” relaxation.

(tractable simplification of $\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad L(x,\lambda)$)

- Hence, $\lambda_{\alpha\beta}$ in LP and BSOS are approximation of the Lagrange multipliers.
- Based on KKT optimality condition: $\lambda_{\alpha\beta}\, g_1^{\alpha_1}(x^*)...g_m^{\alpha_m}(x^*)(1-g_1(x^*))^{\beta_1}...(1-g_m(x^*))^{\beta_m} = 0$
- Hence, when finite convergence in BSOS occurs : $\lambda_{\alpha\beta}\, g_1^{\alpha_1}(x^*)...g_m^{\alpha_m}(x^*)(1-g_1(x^*))^{\beta_1}...(1-g_m(x^*))^{\beta_m} = 0$

$$\Rightarrow \quad p(x^*) - \gamma^* = 0$$

$$\underset{\gamma, Q_0 \succcurlyeq 0}{\text{maximize}} \quad \gamma$$

$$\text{subject to} \quad L(x,\lambda) - \gamma \in SOS_k$$

$$\Downarrow$$

- For $k = 0$, this is results in "Krivine-Stengle's Positivity Certificate" based LP.

(brutal simplification of $\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad L(x,\lambda)$)

- For $k > 0$, this is results in "BSOS" relaxation.

(tractable simplification of $\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad L(x,\lambda)$)

➢ Hence, $\lambda_{\alpha\beta}$ in LP and BSOS are approximation of the Lagrange multipliers.

➢ Based on KKT optimality condition: $\lambda_{\alpha\beta}\, g_1^{\alpha_1}(x^*)...g_m^{\alpha_m}(x^*)(1-g_1(x^*))^{\beta_1}...(1-g_m(x^*))^{\beta_m} = 0$

➢ Hence, when finite convergence in BSOS occurs : $\lambda_{\alpha\beta}\, g_1^{\alpha_1}(x^*)...g_m^{\alpha_m}(x^*)(1-g_1(x^*))^{\beta_1}...(1-g_m(x^*))^{\beta_m} = 0$

$$\Rightarrow \quad p(x^*) - \gamma^* = 0$$

- Section 9.2: Jean B. Lasserre,"An Introduction to Polynomial and Semi-Algebraic Optimization", Cambridge University Press, 2015
- Jean B. Lasserre, Kim-Chuan Toh, Shouguang Yang, "A bounded degree SOS hierarchy for polynomial optimization", EURO Journal on Computational Optimization March 2017, Volume 5, Issue 1–2, pp 87–117

Appendix IV:
Maximal Clique and Principal Submatrix

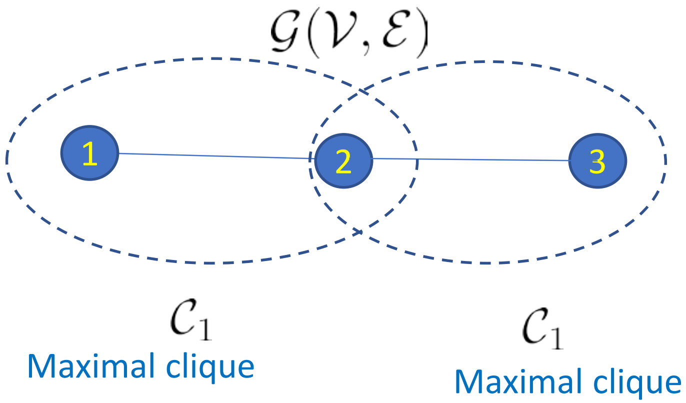
Maximal Clique and Principal Submatrix

$$X = \begin{bmatrix}2 & 1 & 0\\1 & 1 & 1\\0 & 1 & 2\end{bmatrix}\begin{array}{c}\text{①}\\\text{②}\\\text{③}\end{array}$$

- Matrix $X \in \mathcal{S}^n$ with sparsity pattern defined by Graph $\mathcal{G}(\mathcal{V},\mathcal{E})$
- $\mathcal{C}_k$ is maximal clique of graph $\mathcal{G}(\mathcal{V},\mathcal{E})$ with $|\mathcal{C}_k|$ nodes.
- Define matrix $E_{\mathcal{C}_k} \in \mathbb{R}^{|\mathcal{C}_k|\times n}$ as follows:

$$E_{\mathcal{C}_k} \in \mathbb{R}^{|\mathcal{C}_k|\times n} \qquad [E_{\mathcal{C}_k}]_{ij} = \begin{cases}1, & \text{if } \mathcal{C}_k(i) = j\\ 0, & \text{otherwise}\end{cases}$$

Where $\mathcal{C}_k(i)$ is $i-th$ node in $\mathcal{C}_k$



① ② ③ nodes in the graph

$$E_{\mathcal{C}_1} \in \mathbb{R}^{2\times 3} \qquad E_{\mathcal{C}_1} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\end{bmatrix}\begin{array}{c}\text{①}\\\text{②}\end{array} \text{ nodes in } \mathcal{C}_1 \qquad X_{\mathcal{C}_1} = E_{\mathcal{C}_1} X E_{\mathcal{C}_1}^T = \begin{bmatrix}2 & 1\\1 & 0.5\end{bmatrix}$$

$$E_{\mathcal{C}_2} \in \mathbb{R}^{2\times 3} \qquad E_{\mathcal{C}_2} = \begin{bmatrix}0 & 1 & 0\\0 & 0 & 1\end{bmatrix}\begin{array}{c}\text{②}\\\text{③}\end{array} \text{ nodes in } \mathcal{C}_2 \qquad X_{\mathcal{C}_2} = E_{\mathcal{C}_2} X E_{\mathcal{C}_2}^T = \begin{bmatrix}0.5 & 1\\1 & 2\end{bmatrix}$$

Extracts the Principal submatrix of X defined by the indices in cliques $\mathcal{C}_1,\ \mathcal{C}_2$

$$X = E_{\mathcal{C}_1}^T X_1 E_{\mathcal{C}_1} + E_{\mathcal{C}_2}^T X_2 E_{\mathcal{C}_2} \Rightarrow \begin{bmatrix}2 & 1 & 0\\1 & 1 & 1\\0 & 1 & 2\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\end{bmatrix}^T \begin{bmatrix}2 & 1\\1 & 0.5\end{bmatrix} \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\end{bmatrix} + \begin{bmatrix}0 & 1 & 0\\0 & 0 & 1\end{bmatrix}^T \begin{bmatrix}0.5 & 1\\1 & 2\end{bmatrix} \begin{bmatrix}0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$$

16.S498 Risk Aware and Robust Nonlinear Planning
Fall 2019