9. Interpretations, Lifting, SOS and Moments

- Polynomial nonnegativity
- Sum of squares (SOS) decomposition
- Example of SOS decomposition
- Computing SOS using semidefinite programming
- Convexity
- Positivity in one variable
- Background
- Global optimization
- Optimizing in parameter space
- Lyapunov functions

Interpretations

- So far, we have seen how to compute certificates of polynomial non-negativity

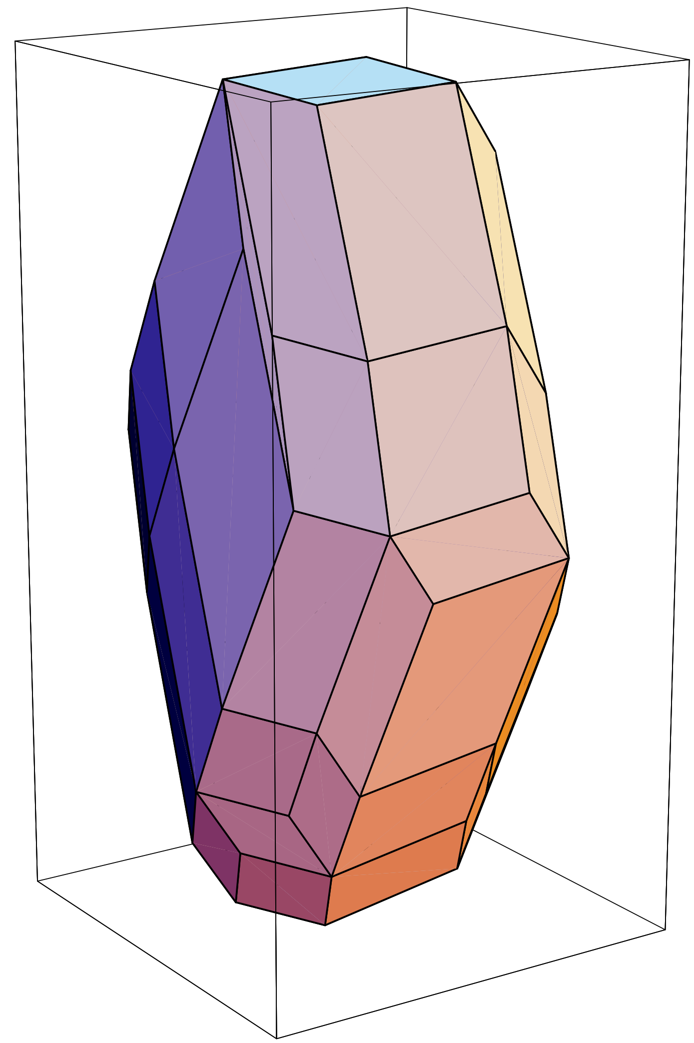
- As we will see, these are *dual SDP relaxations*

- We can also interpret the corresponding primal SDPs

- These arise through *liftings*

A General Method: Liftings

Consider this polytope in $\mathbb{R}^3$ (a zonotope). It has 56 facets, and 58 vertices.



Optimizing a linear function over this set, requires a *linear program* with 56 constraints (one per face).

However, this polyhedron is a three-dimensional *projection* of the 8-dimensional hypercube $\{x \in \mathbb{R}^8, -1 \le x_i \le 1\}$.

Therefore, by using additional variables, we can solve the same problem, by using an LP with *only 16 constraints*.

Liftings

By going to higher dimensional representations, things may become easier:

- "Complicated" sets can be the projection of much simpler ones.
- A polyhedron in $\mathbb{R}^n$ with a "small" number of faces can project to a lower dimensional space with *exponentially* many faces.
- Basic semialgebraic sets can project into non-basic semialgebraic sets.
- Feasible sets of SDPs may project to non-spectrahedral sets.

An essential technique in integer programming.

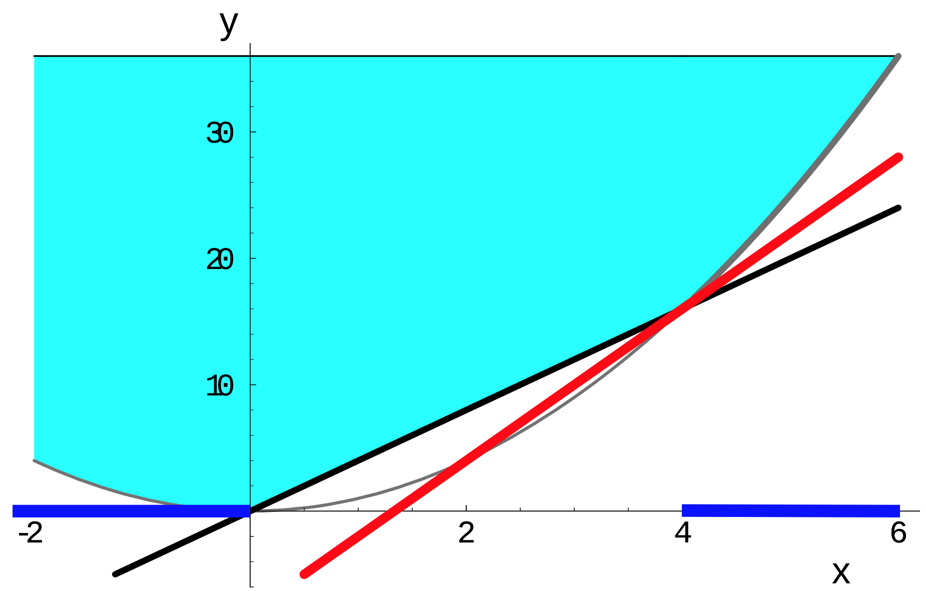
Advantages: compact representations, avoiding "case distinctions," etc.

Example

$$\begin{array}{ll} \text{minimize} & (x-3)^2 \\ \text{subject to} & x\,(x-4) \geq 0 \end{array}$$

The feasible set is $[-\infty, 0] \cup [4, \infty]$. *Not* convex, or even connected.
Consider the lifting $L : \mathbb{R} \to \mathbb{R}^2$, with $L(x) = (x, x^2) =: (x, y)$.
Rewrite the problem in terms of the lifted variables.

- For every lifted point, $\begin{bmatrix} 1 & x \\ x & y \end{bmatrix} \succeq 0$.
- Constraint becomes: $y - 4x \geq 0$
- Objective is now: $y - 6x + 9$



We "get around" nonconvexity: interior points are now on the *boundary*.

Quadratically Constrained Quadratic Programming

A general QCQP is

$$\begin{array}{ll} \text{minimize} & \begin{bmatrix} 1 \\ x \end{bmatrix}^T Q \begin{bmatrix} 1 \\ x \end{bmatrix} \\ \text{subject to} & \begin{bmatrix} 1 \\ x \end{bmatrix}^T A_i \begin{bmatrix} 1 \\ x \end{bmatrix} = 0 \quad \text{for all } i = 1, \dots, m \end{array}$$

The Lagrangian is

$$L(x, \lambda) = \begin{bmatrix} 1 \\ x \end{bmatrix}^T \left(Q - \sum_{i=1}^{m} \lambda_i A_i \right) \begin{bmatrix} 1 \\ x \end{bmatrix}^T$$

so the dual feasible set is defined by semidefinite constraints

QCQP Dual

The dual is the SDP

$$\begin{array}{ll} \text{maximize} & t \\ \text{subject to} & Q - \sum_{i=1}^{m} \lambda_i A_i \succeq t \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \end{array}$$

and the dual of the dual is

$$\begin{array}{ll} \text{minimize} & \mathbf{trace}\, QY \\ \text{subject to} & \mathbf{trace}\, A_i Y = 0 \quad \text{for all } i = 1, \ldots, m \\ & Y \succeq 0 \\ & Y_{11} = 1 \end{array}$$

Lifting

Lifting is a general approach for constructing *primal relaxations*; the idea is

- Introduce new variables Y which are polynomial in x
  This embeds the problem in a *higher dimensional* space
- Write *valid inequalities* in the new variables
- The feasible set of the original problem is the *projection* of the lifted feasible set

Lifting QCQP

We have the QCQP

$$\begin{array}{ll} \text{minimize} & \begin{bmatrix} 1 \\ x \end{bmatrix}^T Q \begin{bmatrix} 1 \\ x \end{bmatrix} \\ \text{subject to} & \begin{bmatrix} 1 \\ x \end{bmatrix}^T A_i \begin{bmatrix} 1 \\ x \end{bmatrix} = 0 \quad \text{for all } i = 1, \ldots, m \end{array}$$

Use *lifted variables* $Y \in \mathbb{S}^n$, defined by $Y = \begin{bmatrix} 1 \\ x \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix}^T$

We have valid constraints

$$Y \succeq 0, \qquad Y_{11} = 1, \qquad \mathbf{rank}\, Y = 1$$

Every such Y corresponds to a unique x

Lifted QCQP

The lifted problem is

$$\begin{array}{rll} \text{minimize} & \mathbf{trace}\, QY & \\ \text{subject to} & \mathbf{trace}\, A_i Y = 0 & \text{for all } i = 1, \ldots, m \\ & Y \succeq 0 & \\ & Y_{11} = 1 & \\ & \mathbf{rank}\, Y = 1 & \end{array}$$

Again, we can drop the non-convex constraint to obtain a relaxation

This (happens to) give the same as the dual of the dual

QCQP Interpretation of Polynomial Programs

We can also lift *polynomial* programs; consider the example

$$\text{minimize} \qquad \sum_{k=0}^{6} a_k x^k$$

We'll choose lifted variables

$$y = \begin{bmatrix} x \\ x^2 \\ x^3 \end{bmatrix}$$

then the cost function is

$$f = a_0 + a_1 y_1 + a_2 y_2 + a_3 y_3 + a_4 y_1 y_3 + a_5 y_2 y_3 + a_6 y_3^2$$

a *quadratic* function of y (many other choices possible)

QCQP Interpretation of Polynomial Programs

We have the *equivalent QCQP*

$$\begin{array}{ll} \text{minimize} & \begin{bmatrix} 1 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix}^T \begin{bmatrix} a_0 & \frac{a_1}{2} & \frac{a_2}{2} & \frac{a_3}{2} \\ & 0 & 0 & \frac{a_4}{2} \\ & & 0 & \frac{a_5}{2} \\ & & & a_6 \end{bmatrix} \begin{bmatrix} 1 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} \\ \text{subject to} & y_2 - y_1^2 = 0 \\ & y_3 - y_1 y_2 = 0 \end{array}$$

to make the Lagrange dual tighter, we can add the *valid constraint*

$$y_2^2 - y_1 y_3 = 0$$

Every polynomial program can be expressed as an equivalent QCQP

Quadratic Constraints

The above quadratic constraints are

$$\begin{bmatrix} 1 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 2 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} = 0$$

$$\begin{bmatrix} 1 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 2 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} = 0$$

$$\begin{bmatrix} 1 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} = 0$$

Relaxations

We can now construct the SDP primal and dual relaxations of this QCQP

Example

Suppose $f = x^6 + 4x^2 + 1$, then the SDP dual relaxation is

$$\begin{array}{ll} \text{maximize} & t \\ \text{subject to} & \begin{bmatrix} 1-t & 0 & 2+\lambda_2 & -\lambda_3 \\ 0 & -2\lambda_2 & \lambda_3 & \lambda_1 \\ 2+\lambda_2 & \lambda_3 & -2\lambda_1 & 0 \\ -\lambda_3 & \lambda_1 & 0 & 1 \end{bmatrix} \succeq 0 \end{array}$$

this is exactly the condition that $f - t$ be sum of squares

The Primal Relaxation of a Polynomial Program

Since we have a QCQP, there is also an SDP *primal relaxation*, constructed via the lifting

$$Y = \begin{bmatrix} 1 \\ y \end{bmatrix} \begin{bmatrix} 1 \\ y \end{bmatrix}^T$$

It is the SDP

$$\begin{aligned} \text{minimize} \quad & \mathbf{trace} \begin{bmatrix} a_0 & \frac{a_1}{2} & \frac{a_2}{2} & \frac{a_3}{2} \\ & 0 & 0 & \frac{a_4}{2} \\ & & 0 & \frac{a_5}{2} \\ & & & a_6 \end{bmatrix} Y \\ \text{subject to} \quad & Y \succeq 0 \\ & Y_{11} = 1 \qquad Y_{24} = Y_{33} \\ & Y_{22} = Y_{13} \qquad Y_{14} = Y_{23} \end{aligned}$$

The Primal Relaxation of a Polynomial Program

This is constructed by

$$Y = \begin{bmatrix}1\\ y\end{bmatrix}\begin{bmatrix}1\\ y\end{bmatrix}^T = \begin{bmatrix}1\\ x\\ x^2\\ x^3\end{bmatrix}\begin{bmatrix}1\\ x\\ x^2\\ x^3\end{bmatrix}^T = \begin{bmatrix}1 & x & x^2 & x^3\\ x & x^2 & x^3 & x^4\\ x^2 & x^3 & x^4 & x^5\\ x^3 & x^4 & x^5 & x^6\end{bmatrix}$$

- One may construct this directly from the polynomial program
- Direct extensions to the multivariable case
- The feasible set of Y may be projected to give a feasible set of x
- If the optimal Y has $\mathbf{rank}\, Y = 1$ then the relaxation is *exact*

Lifting

Higher dimensional representations have several possible advantages

- One may find *simpler representations*, e.g., polytopes
- Basic semialgebraic sets may project to non-basic ones
- Adding new variables via lifting allows new valid inequalities, which tightens the dual
- Using polynomial lifting allows more constraints to be represented in LP or SDP form
- Lifting *wraps* the feasible set onto a higher dimensional variety; this tends to map interior points to boundary points

Outer Approximation of Semialgebraic Sets

The primal SDP relaxation allows us to construct outer approximation of a semialgebraic set

For example, one can compute an outer approximation of the epigraph

$$S = \left\{ (x_1, x_2) \mid x_2 \geq f(x_1) \right\}$$

In one variable, the SDP relaxation gives exactly the *convex hull*, since S is contained in a halfspace

$$\{ x \in \mathbb{R}^2 \mid a^T x \leq b \}$$

if and only if the following polynomial inequality holds

$$a_1 x + a_2 f(x) \leq b \text{ for all } x$$

Example: Outer Approximation of the Epigraph

Let's look at the univariate example

$$f = \frac{1}{2}(x-1)(x-2)(x-3)(x-5)$$

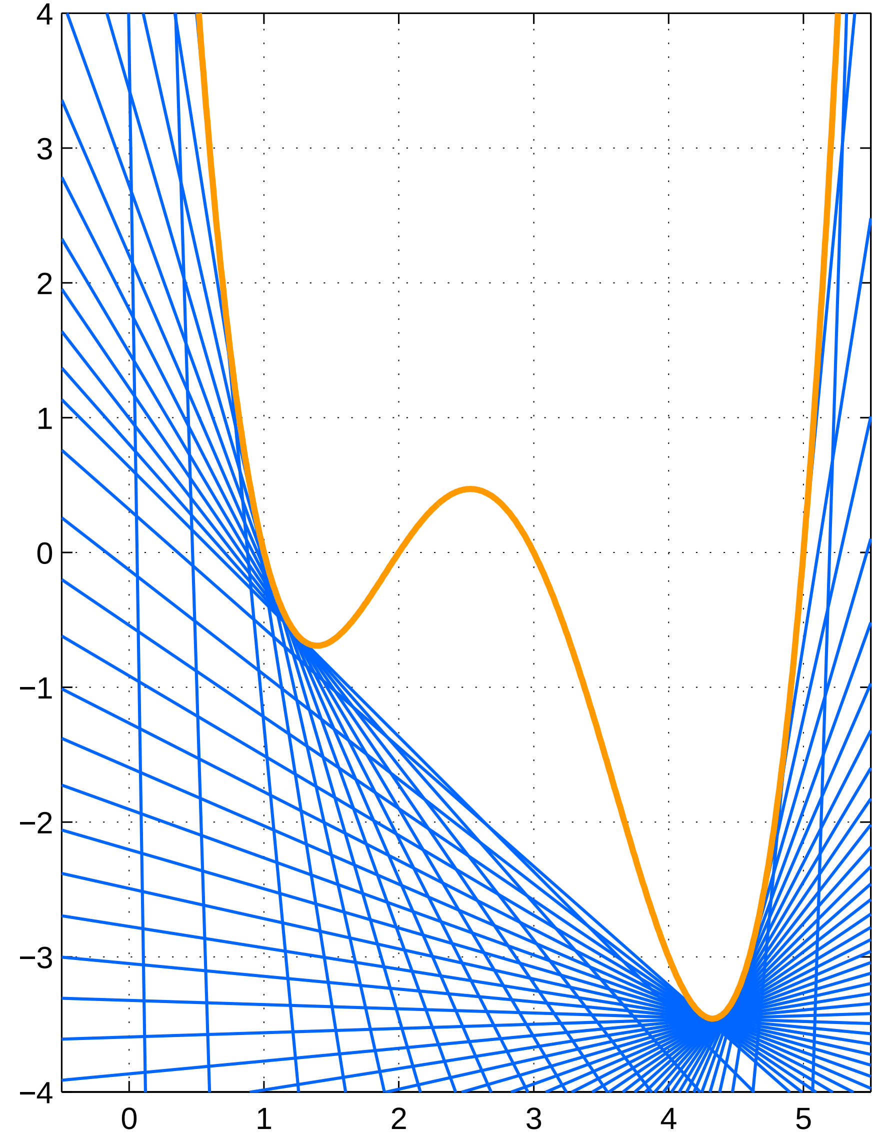
If $y \geq f(x)$ then the following SDP is feasible

$$y \geq \frac{1}{4}\,\mathbf{trace}\begin{bmatrix} 60 & -61 & 41 \\ -61 & 0 & -11 \\ 41 & -11 & 2 \end{bmatrix} X$$

$$X \succeq 0$$

$$X_{22} = 2X_{12} \quad X_{11} = 1$$

$$X_{12} = x$$



Moments Interpretation of the Primal Relaxation

Instead of trying to minimize directly f, we can solve

$$\begin{array}{ll} \text{minimize} & \mathbf{E}\, f = \displaystyle\int_{\mathbb{R}^n} f(x)p(x)\,dx \\ \text{subject to} & p \text{ is a probability distribution on } \mathbb{R}^n \end{array}$$

- This is a *dual* problem to minimizing f
- If f has a unique minimum at x_0, then the optimal will be a point measure at x_0
- Essentially due to Lasserre

Moments Interpretation of the Primal Relaxation

suppose $y = \begin{bmatrix} 1 & x & y & xy & x^2 & \dots \end{bmatrix}^T$, then $f = c^T y$ and

$$\mathbf{E}\, f = c^T\, \mathbf{E}\, y$$

$\mathbf{E}\, y$ is the *vector of moments* of the distribution

so we have the equivalent problem

$$\begin{array}{ll} \text{minimize} & c^T z \\ \text{subject to} & z \text{ is a vector of moments of } y \end{array}$$

Example

Since $\mathbf{E}\, yy^T \succeq 0$ for any distribution, we have *valid inequalities*

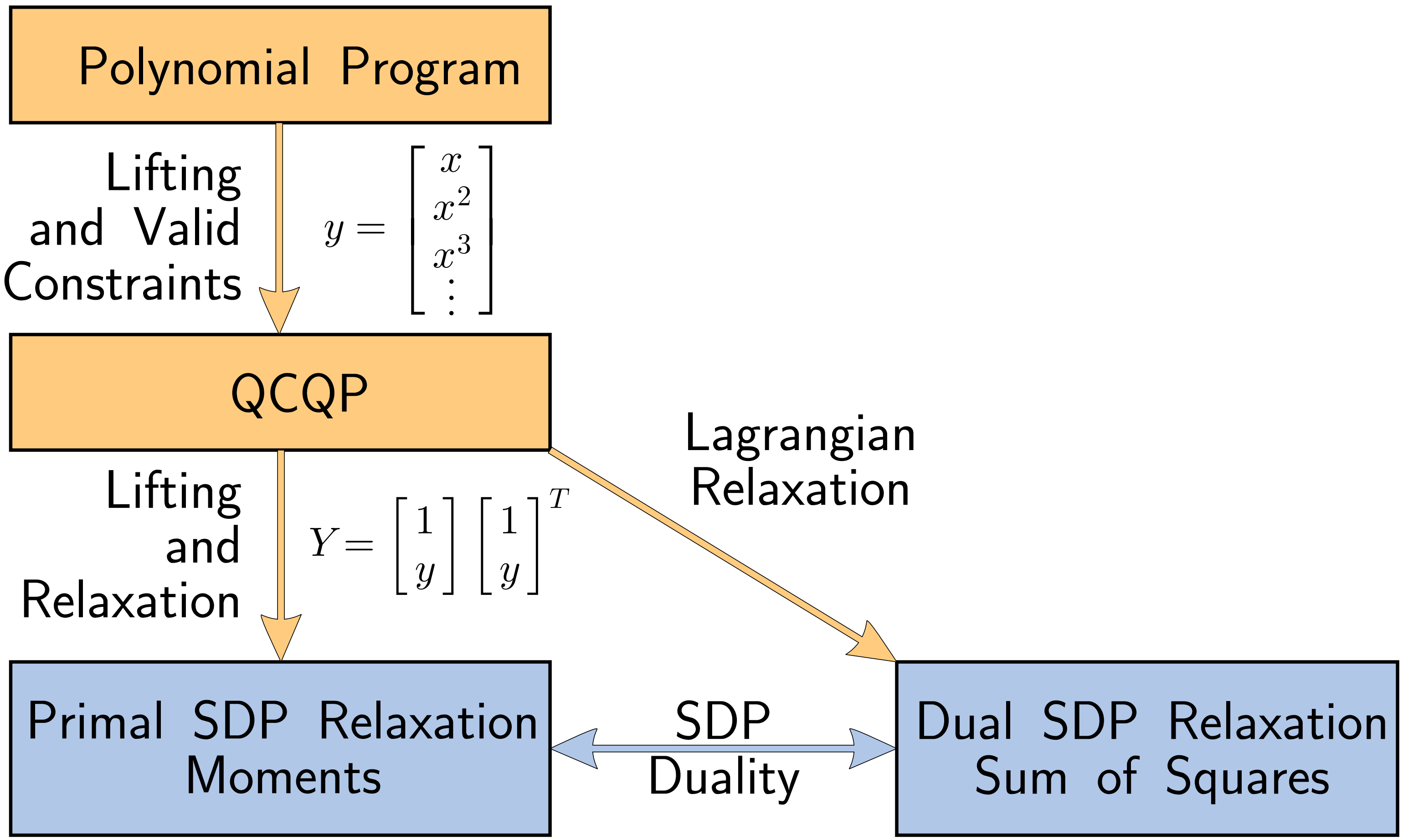
$$\mathbf{E} \begin{bmatrix} 1 \\ x \\ y \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \end{bmatrix}^T = \mathbf{E} \begin{bmatrix} 1 & x & y \\ x & x^2 & xy \\ y & xy & y^2 \end{bmatrix} \succeq 0$$

so to find a lower bound $x^2 + 2xy + 3y^2$ we solve the SDP

$$\begin{aligned} \text{minimize} \quad & \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} z \\ \text{subject to} \quad & M \succeq 0 \\ & z_1 = M_{22},\ z_2 = M_{12},\ z_3 = M_{22} \end{aligned}$$

- This is exactly the *primal SDP relaxation*; the dual of SOS
- Similar to MAXCUT, where the SDP relaxation may be viewed as a *covariance matrix*

A General Scheme



- *Primal:* the solution to the lifted problem *may* suggest candidate points where the polynomial is negative.
- *Dual:* the sum of squares *certifies* or *proves* polynomial nonnegativity.