Algebraic Preliminaries and Results about Moment Matrices

Monique Laurent

CWI, Amsterdam

ADONET Doctoral school
Optimization over Polynomials and Semidefinite Programming

Klagenfurt
September 12-16, 2005

Polynomial Ideals and Varieties

$I \subseteq \mathbb{R}[x_1, \ldots, x_n]$ ideal if $f + g \in I \quad \forall f, g \in I$ and

$$fg \in I \quad \forall f \in I, \ \forall g \in \mathbb{R}[x_1, \ldots, x_n]$$

Finite Basis Theorem: Any ideal I has a finite set of generators $h_1, \ldots, h_m$; that is, $I = \{\sum_{j=1}^m u_j h_j \mid u_j \in \mathbb{R}[x_1, \ldots, x_n]\}$

$V(I) := \{x \in \mathbb{C}^n \mid f(x) = 0 \, \forall f \in I\}$, the complex variety of I

Two ideals containing I:

$I(V(I)) := \{f \in \mathbb{R}[x_1, \ldots, x_n] \mid f(v) = 0 \, \forall v \in V(I)\}$

$\sqrt{I} := \{f \in \mathbb{R}[x_1, \ldots, x_n] \mid f^k \in I$ for some $k \geq 1\}$

Hilbert Nullstellensatz: $\boxed{I(V(I)) = \sqrt{I}}$

Radical Ideal - Zero-Dimensional Ideal

Definition: I is radical if $I = \sqrt{I}$
I is zero-dimensional if $|V(I)| < \infty$

E.g., $I = \langle x^2 \rangle$ is *not* radical, since $x \in \sqrt{I} \setminus I$

by the Nullstellensatz, I radical $\iff$ *the only polynomials vanishing at all points of $V(I)$ are the polynomials in I*

$\mathbb{R}[x_1, \ldots, x_n]/I$:= quotient vector space, whose elements are the equivalence classes $p \mod I$ (for $p \in \mathbb{R}[x_1, \ldots, x_n]$)

Theorem: I zero-dimensional $\iff \dim \mathbb{R}[x_1, \ldots, x_n]/I < \infty$

Moreover, $|V(I)| \leq \dim \mathbb{R}[x_1, \ldots, x_n]/I$, with equality if and only if I is radical

Prove: $\dim \mathbb{R}[x_1, \ldots, x_n]/I < \infty \Longrightarrow |V(I)| < \infty$

Set $k := \dim \mathbb{R}[x_1, \ldots, x_n]/I$. Consider variable x_1, say.

The set $\{1, x_1, \ldots, x_1^k\}$ is linearly dependent in $\mathbb{R}[x_1, \ldots, x_n]/I$

Hence there exist reals $\lambda_0, \ldots, \lambda_k$ (not all zero) for which the polynomial $f(x_1) := \sum_{j=0}^{k} \lambda_j x_1^j$ belongs to I

Therefore, $f(v) = 0$ for all $v \in V(I)$

This implies that $|\{v_1 \mid v \in V(I)\}| < \infty$

Because a univariate nonzero polynomial has *finitely* many roots

The same holds for the other coordinates

Prove: $|V(I)| < \infty \Longrightarrow \dim \mathbb{R}[x_1, \ldots, x_n]/I < \infty$

Define $\mathcal{B}$:= set of standard monomials w.r.t. some monomial ordering, i.e., $x^\alpha \in \mathcal{B}$ if x^α *is not divisible by the* leading term *of a polynomial in I*

Then, $\mathcal{B}$ is a basis of $\mathbb{R}[x_1, \ldots, x_n]/I$
[Use the division algorithm w.r.t. a Groebner basis of I]

Show: $|\mathcal{B}| < \infty$

Let $\{v_1 \mid v \in V(I)\} = \{a_1, \ldots, a_k\}$, $f(x) := \prod_{j=1}^{k} (x_1 - a_j)$

Then, $f \in I(V(I)) = \sqrt{I} \Longrightarrow f^m \in I$ for some $m \in \mathbb{N}$
$\Longrightarrow x_1^{km}$ is the *leading term* of a polynomial in I

Thus, $\alpha_1 < km$ for all $x^\alpha \in \mathcal{B} \quad \Longrightarrow |\mathcal{B}| < \infty$

A Useful Tool: Interpolation Polynomials

Lemma: Let $V \subseteq \mathbb{C}^n$ with $|V| < \infty$.
There exist $p_v \in \mathbb{C}[x_1, \ldots, x_n] \quad (v \in V)$
(*interpolation polynomials at the points of* V)
such that $p_v(u) = \delta_{u,v} \;\; \forall u, v \in V$

One may choose

$p_{\overline{v}} = \overline{p_v}$ if V is closed under complex conjugation

Proof: Fix $v \in V$. For $u \in V \setminus \{v\}$, pick an index i_u for which $v_{i_u} \neq u_{i_u}$ and define $p_v(x) := \prod_{u \in V,\, u \neq v} \frac{x_{i_u} - u_{i_u}}{v_{i_u} - u_{i_u}}$

I **radical 0-dimensional** $\Longrightarrow \dim \mathbb{R}[x_1, \ldots, x_n]/I = |V(I)|$

$V(I)$ is closed under complex conjugation
$\Longrightarrow$ Can partition $V(I) = S \cup T \cup \overline{T}$, where $S := V(I) \cap \mathbb{R}^n$

Let $p_v \in \mathbb{C}[x_1, \ldots, x_n]$ be interpolation polynomials, where
$p_v \in \mathbb{R}[x_1, \ldots, x_n]$ ($v \in S$) and $p_{\overline{v}} = \overline{p_v}$ ($v \in T$)

Claim: The set $\mathcal{B} := \{p_v \ (v \in S)\} \cup \{\mathrm{Re}(p_v), \mathrm{Im}(p_v) \mid v \in T\}$ is a linear basis of $\mathbb{R}[x_1, \ldots, x_n]/I$

$\mathcal{B}$ generates $\mathbb{R}[x_1, \ldots, x_n]/I$: For $f \in \mathbb{R}[x_1, \ldots, x_n]$,
$f - \sum_{v \in V(I)} f(v) p_v$ vanishes on $V(I)$
$\Longrightarrow f = \sum_{v \in V(I)} f(v) p_v \mod I$ [as I radical]
$\Longrightarrow$ f is a linear combination of $\mathcal{B} \mod I$

$\mathcal{B}$ is linearly independent in $\mathbb{R}[x_1, \ldots, x_n]/I$

Application 1: Parrilo's Representation Result

Theorem [Parrilo 2002]
Assume $I = \langle h_1, \ldots, h_m \rangle$ is zero-dimensional and radical.
If $f(x) \geq 0$ on $S = \{x \in \mathbb{R}^n \mid h_1(x) = 0, \ldots, h_m(x) = 0\}$,
then f is a sum of squares modulo I

Proof: Partition $V(I) = S \cup T \cup \overline{T}$, where $S = V(I) \cap \mathbb{R}^n$
For $v \in S$, $q_v := \sqrt{f(v)} p_v \in \mathbb{R}[x_1, \ldots, x_n]$
For $v \in T$, $q_v := \sqrt{f(v)} p_v + \overline{\sqrt{f(v)} p_v} \in \mathbb{R}[x_1, \ldots, x_n]$

Then, $f - \sum_{v \in S \cup T} q_v^2$ vanishes on $V(I)$
Therefore, $f - \sum_{v \in S \cup T} q_v^2 \in I \quad$ [as I radical]

That is, $f = \sum_{v \in S \cup T} q_v^2 \mod I$

Application 2: Solving Systems of Polynomial Equations

Problem: Find the solutions to the polynomial equations:

$$\boxed{h_1(x) = 0, \ldots, h_m(x) = 0}$$

i.e., find $V(I)$, assuming $I = \langle h_1, \ldots, h_m \rangle$ is 0-dimensional radical

The eigenvalue method: For $h \in \mathbb{R}[x_1, \ldots, x_n]$, consider the multiplication operator:

$$\boxed{\begin{array}{rll} m_h : & \mathbb{R}[x_1, \ldots, x_n]/I & \rightarrow \mathbb{R}[x_1, \ldots, x_n]/I \\ & f \mod I & \mapsto fh \mod I \end{array}}$$

Let M_h be the matrix of m_h w.r.t. a basis $\mathcal{B}$ of $\mathbb{R}[x_1, \ldots, x_n]/I$

Thus, M_h has size $|\mathcal{B}| = \dim \mathbb{R}[x_1, \ldots, x_n]/I = |V(I)|$

The Eigenvalue Method - Stickelberger Theorem

Say the basis $\mathcal{B}$ of $\mathbb{R}[x_1, \ldots, x_n]/I$ is a monomial basis, i.e., has the form $\boxed{\mathcal{B} = \{x^\alpha \mid \alpha \in \mathcal{A}\}}$

For $v \in V(I)$, set $\boxed{\zeta_v = (v^\alpha)_{\alpha \in \mathcal{A}}}$

Let p_v be interpolation polynomials at the points of $V(I)$ s.t. *each p_v is a linear combination of monomials in the basis $\mathcal{B}$*

Theorem: The eigenvalues of M_h are $h(v)$ $(v \in V(I))$ with respective left eigenvector p_v and right eigenvector ζ_v.
That is,

$$\boxed{M_h p_v = h(v) p_v \qquad M_h^T \zeta_v = h(v) \zeta_v \quad (v \in V(I))}$$

Prove: $M_h^T \zeta_v = h(v)\zeta_v$

$$\begin{array}{cc} & h(x)x^\alpha \\ x^\beta & \begin{pmatrix} & \vdots & \\ \cdots & \lambda_\beta & \cdots \\ & \vdots & \end{pmatrix} \end{array} = M_h$$

For $x^\alpha \in \mathcal{B}$, $\boxed{h(x)x^\alpha \equiv \sum_{\beta\in\mathcal{A}} \lambda_\beta x^\beta \mod I}$

The α-th coordinate of $M_h^T \zeta_v$ is equal to

$$\sum_{\beta\in\mathcal{A}} \lambda_\beta v^\beta = h(v)v^\alpha$$

thus equal to the α-th coordinate of $h(v)\zeta_v$

Find the points $v \in V(I)$ from the right eigenvectors ζ_v

[Corless-Gianni-Trager 1997]

1. Pick $h \in \mathbb{R}[x_1, \ldots, x_n]$ for which all values $h(v)$ ($v \in V(I)$) are distinct. True with high probability for

$$\boxed{h(x) = c_1 x_1 + \ldots + c_n x_n}$$

with random scalars $c_1, \ldots, c_n$

2. Then the right eigenspaces of M_h are 1-dimensional, thus spanned by ζ_v ($v \in V(I)$)
$\rightsquigarrow$ can compute ζ_v ($v \in V(I)$) [w.l.o.g. assume $1 \in \mathcal{B}$]

3. Easy to recover $v = (v_1, \ldots, v_n)$ from ζ_v

Either $x_i \in \mathcal{B}$; else express x_i in terms of the basis $\mathcal{B} \mod I$

Back to Moment Matrices

Objective: Discuss some results of Curto-Fialkow about moment matrices:

- Finite Rank Theorem
- Flat Extension Theorem
- Application to optimization over a semialgebraic set K

which will enable us to derive:

- the optimality certificate
- the extraction procedure for global minimizers

of Henrion-Lasserre

Recap Some Notation

μ measure $\rightsquigarrow y = (\int x^\alpha \mu(dx))_\alpha$ the *sequence of moments* of μ

μ is said to be a *representing measure* for y

$\mu = \delta_v$ *Dirac measure* at $v \rightsquigarrow y = \zeta_v := (v^\alpha)_{\alpha \in \mathbb{N}^n}$ *Zeta vector*

μ is *r-atomic nonnegative* if μ is a conic combination of r Dirac measures

$y \in \mathbb{R}^{\mathbb{N}^n} \rightsquigarrow M(y) := (y_{\alpha+\beta})_{\alpha,\beta \in \mathbb{N}^n}$ the *moment matrix* of y

Recall: y has a r-atomic nonnegative representing measure

$$\Longrightarrow M(y) \succeq 0 \text{ and } \mathrm{rank}\, M(y) \leq r$$

Finite Rank Positive Semidefinite Moment Matrices

The Finite Rank Theorem: [Curto-Fialkow 1996]
Assume $M(y) \succeq 0$ and $r := \text{rank } M(y) < \infty$.
Then y is the sequence of moments of a (unique) r-atomic nonnegative measure μ.

That is,

$$\mu = \sum_{i=1}^{r} \lambda_i \delta_{v_i}, \quad \text{where } \lambda_i > 0, \ v_i \in \mathbb{R}^n$$

$$\Longrightarrow y = \sum_{i=1}^{r} \lambda_i \zeta_{v_i} = \sum_{i=1}^{r} \lambda_i (v_i^{\alpha})_{\alpha \in \mathbb{N}^n}$$

Elementary (algebraic) proof following [Laurent 2005]

Some Calculus About Moment Matrices

Recall that we identify a polynomial

$$p(x) = \sum_\alpha p_\alpha x^\alpha \in \mathbb{R}[x_1, \ldots, x_n]$$

with the vector of its coefficients

$$p = (p_\alpha)_{\alpha \in \mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$$

Lemma 1: For $f, g, h \in \mathbb{R}[x_1, \ldots, x_n]$,

$$\boxed{f^T M(y)(gh) = y^T(fgh) = (fg)^T M(y) h}$$

Proof: Direct verification

The Kernel of $M(y)$ is a Radical Ideal

Lemma 2: If $M(y) \succeq 0$, then

$$I := \text{Ker } M(y) = \{p \in \mathbb{R}[x_1, \ldots, x_n] \mid M(y)p = 0\}$$

is a *radical ideal* in $\mathbb{R}[x_1, \ldots, x_n]$

Proof: If $f \in I$, $g \in \mathbb{R}[x_1, \ldots, x_n]$, then

$$(fg)^T M(y)(fg) = (fg^2)^T M(y) f = 0,$$

implying $fg \in I$. Thus I **is an ideal**

If $f^2 \in I$, then

$$f^T M(y) f = 1^T M(y)(f^2) = 0,$$

implying $f \in I$. Thus I **is radical**

The Kernel of $M(y)$ is a 0-dimensional Ideal

Let $\mathcal{A} \subseteq \mathbb{N}^n$ indexing a maximum nonsingular principal submatrix of $M(y)$. Thus, $|\mathcal{A}| = r = \text{rank } M(y)$.

Lemma 3: $\mathcal{B} := \{x^\alpha \mid \alpha \in \mathcal{A}\}$ is a basis of $\mathbb{R}[x_1, \ldots, x_n]/I$

Proof:

- $\mathcal{B}$ is linearly independent in $\mathbb{R}[x_1, \ldots, x_n]/I$ since the columns of $M(y)$ indexed by $\mathcal{A}$ are linearly independent
- $\mathcal{B}$ spans $\mathbb{R}[x_1, \ldots, x_n]/I$ since any column of $M(y)$ is a linear combination of its columns indexed by $\mathcal{A}$

Corollary: $\dim \mathbb{R}[x_1, \ldots, x_n]/I = |\mathcal{B}| = r = |V(I)|$

An Identity for $M(y)$

Partition $V(I) = S \cup T \cup \overline{T}$, where $S := V(I) \cap \mathbb{R}^n$

p_v:= interpolation polynomials s.t. $p_{\overline{v}} = \overline{p_v}$ for $v \in V(I)$

Z:= $|\mathbb{N}^n| \times |V(I)|$ matrix with columns $\zeta_v := (v^\alpha)_{\alpha \in \mathbb{N}^n}$

$\tilde{Z}$:= $|V(I)| \times |\mathbb{N}^n|$ matrix with rows p_v

Lemma 4: $\boxed{M(y) = Z\mathrm{diag}(\tilde{Z}y)Z^T}$

Proof: The polynomial $x^{\alpha+\beta} - \sum_{v \in V(I)} v^\alpha v^\beta p_v(x)$ vanishes at all points of $V(I)$, thus belongs to I [as I radical], which implies

$$y_{\alpha+\beta} = \sum_{v \in V(I)} v^\alpha v^\beta \sum_\gamma (p_v)_\gamma y_\gamma$$

Lemma 5: $V(I) \subseteq \mathbb{R}^n$

$$Z = \begin{array}{c} \begin{array}{ccc} S & T & \overline{T} \end{array} \\ \begin{pmatrix} A & B & \overline{B} \end{pmatrix} \end{array}, \quad \tilde{Z}y = \begin{array}{c} S \\ T \\ \overline{T} \end{array} \begin{pmatrix} a \\ b \\ \overline{b} \end{pmatrix} \quad \text{for real } A, B, a, b$$

$$\Longrightarrow \quad M(y) = A\text{diag}(a)A^T + B\text{diag}(b)B^T + \overline{B\text{diag}(b)B^T}$$

Write: $A\text{diag}(a)A^T = CC^T - DD^T$

$B\text{diag}(b)B^T + \overline{B\text{diag}(b)B^T} = EE^T - FF^T$, where

columns of C = $|S_+| := |\{v \in S \mid a_v > 0\}$
columns of D = $|S_-| := |\{v \in S \mid a_v < 0\}|$
columns of E and F = $|T|$

$$\Longrightarrow M(y) = (CC^T + EE^T) - (DD^T + FF^T)$$

$$M(y) = (CC^T + EE^T) - (DD^T + FF^T)$$

$$M(y) \succeq 0 \Longrightarrow \text{ Ker } (CC^T + EE^T) \subseteq \text{Ker } M(y)$$

$$\Longrightarrow \text{ rank } M(y) \leq \text{rank}(CC^T + EE^T) \leq |S_+| + |T|$$

But, rank $M(y) = r = |V(I)| = |S| + 2|T|$

$$\Longrightarrow T = \emptyset, \; S_- = \emptyset$$

Thus, $M(y) = \sum_{v \in V(I)} a_v \zeta_v \zeta_v^T$, where $a_v = p_v^T M(y) p_v > 0$

That is, $y = \sum_{v \in V(I)} a_v \zeta_v$

End of Proof of Finite Rank Theorem

The Flat Extension Theorem for Truncated Moment Matrices

Definition: The matrix $X = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$ is a *flat extension* of A if $\operatorname{rank}(X) = \operatorname{rank}(A)$. Then, $\boxed{X \succeq 0 \Longleftrightarrow A \succeq 0}$

Notation: $S_t := \{\alpha \in \mathbb{N}^n \mid |\alpha| = \sum_{i=1}^n \alpha_i \leq t\}$

The Flat Extension Theorem: [Curto-Fialkow 1996]
Let $y \in \mathbb{R}^{S_{2t}}$ for which $M_t(y) \succeq 0$ and $M_t(y)$ is a flat extension of $M_{t-1}(y)$. Then one can extend y to a (unique) vector of $\mathbb{R}^{S_{2t+2}}$, again denoted as y, in such a way that $M_{t+1}(y)$ is a flat extension of $M_t(y)$

Application 1 [Useful for Optimization]

Corollary CF: [Curto-Fialkow 1996]
Let $y \in \mathbb{R}^{S_{2t}}$ for which $M_t(y) \succeq 0$ and $M_t(y)$ is a flat extension of $M_{t-1}(y)$. Then, y is the sequence of moments of a (unique) nonnegative measure

Proof:
• Repeatedly applying the *Flat Extension Theorem*, y can be extended to a vector $y \in \mathbb{R}^{\mathbb{N}^n}$ in such a way that $M(y)$ *is a flat extension of* $M_t(y)$

• As $\operatorname{rank} M(y) = \operatorname{rank} M_t(y) < \infty$, y is the sequence of moments of a (unique) nonnegative measure by the *Finite Rank Theorem*

Ideal-like Property of the Kernel of a Truncated Moment Matrix

Lemma: Let $M_t(y) \succeq 0$ and f, g polynomials.
(i) If $\deg(fg) \leq t-1$, then $M_t(y)f = 0 \Longrightarrow M_t(y)(fg) = 0$
(ii) If $M_t(y)$ is a flat extension of $M_{t-1}(y)$ and $\deg(fg) \leq t$, then $M_t(y)f = 0 \Longrightarrow M_t(y)(fg) = 0$

The proof uses the following fact for matrix $X = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \succeq 0$:

(i) $Ax = 0 \Longrightarrow X \begin{pmatrix} x \\ 0 \end{pmatrix} = 0$

(ii) If X is a flat extension of A, then

$$Ax + By = 0 \Longrightarrow X \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

Idea of Proof for the Flat Extension Theorem

Given: $X = M_t(y) = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$ flat extension of $A = M_{t-1}(y)$

Goal: Construct matrices D, E in such a way that the matrix

$Y := \begin{pmatrix} X & D \\ D^T & E \end{pmatrix}$ *is a flat extension of* X and *a moment matrix*

Idea: Let $|\gamma| = t + 1$. Say, $\gamma_1 \geq 1$. Write $\gamma = e_1 + \beta$, where $|\beta| = t$. Then $\boxed{X_\beta = \sum_{|\alpha| \leq t-1} \lambda_\alpha X_\alpha}$

Necessarily: $\boxed{Y_\gamma = \sum_{|\alpha| \leq t-1} \lambda_\alpha Y_{\alpha+e_1}}$

Define: $\boxed{D_\gamma := \sum_{|\alpha| \leq t-1} \lambda_\alpha X_{\alpha+e_1}}$

Application 2 [Useful for Optimization with Constraints]

$$\boxed{K := \{x \in \mathbb{R}^n \mid h_1(x) \geq 0, \ldots, h_m(x) \geq 0\}}$$

$$d_j := \lceil \deg(h_j)/2 \rceil, \, d := \max_j d_j$$

Theorem CF: [Curto-Fialkow 2000]
Assume $M_t(y) \succeq 0$, $M_{t-d_j}(h_j * y) \succeq 0 \ \ \forall j \leq m$ and
$\text{rank } M_t(y) = \text{rank } M_{t-d}(y)$=:$r$.
Then, y is the sequence of moments of an r-atomic nonnegative measure on K.
That is,

$$y = \sum_{i=1}^{r} \lambda_i \zeta_{v_i} \ \text{ where } \lambda_i > 0, \ v_i \in K$$

Short Proof of Theorem CF [Laurent 2005]

By Corollary CF, we already know that y is the sequence of moments of a measure $\mu = \sum_{i=1}^{r} \lambda_i \delta_{v_i}$, where $\lambda_i > 0$ and $v_i \in \mathbb{R}^n$

Remains to show that $v_i \in K$, i.e., $h_j(v_i) \geq 0 \ \ \forall j \leq m$

Let $p_1, \ldots, p_r$ be interpolation polynomials at $v_1, \ldots, v_r$

Fact: We may assume that $\deg(p_i) \leq t - d$

Proof: let $I := \mathrm{Ker}\ M(y)$ and let $\mathcal{B} \subseteq S_{t-d}$ be a basis of $\mathbb{R}[x_1, \ldots, x_n]/I$. Then, $\{v_1, \ldots, v_r\} = V(I)$.
Simply, replace p_i by its residue modulo I w.r.t. the basis $\mathcal{B}$.

Now:
$0 \leq p_i^T M_{t-d}(h_j * y) p_i = \sum_{k=1}^{r} \lambda_k (p_i(v_k))^2 h_j(v_k) = \lambda_i h_j(v_i)$,

implying $h_j(v_i) \geq 0$

Back to the Optimality Certificate of Henrion-Lasserre

Problem:

$$p_{\min} := \inf\ p(x)\ \text{ s.t. } x \in K = \{x \in \mathbb{R}^n \mid h_1(x) \geq 0 \ldots h_m(x) \geq 0\}$$

Moment relaxation (MOMt):

$$p_t^* := \inf\ p^T y\ \text{ s.t. } \quad y_0 = 1,\ M_t(y) \succeq 0$$
$$M_{t-d_j}(h_j * y) \succeq 0\ (j = 1, \ldots, m)$$

for $t \geq \lceil \deg(p)/2 \rceil,\ d_j := \lceil \deg(h_j)/2 \rceil,\ d := \max_j d_j$

Optimality Certificate Theorem: [Henrion-Lasserre 2005]

Let y be an optimum solution to **(MOMt)**. Then,

$$\text{rank } M_t(y) = \text{rank } M_{t-d}(y) \Longrightarrow p_t^* = p_{\min}$$

Extracting Global Minimizers

Proof: Directly from *Theorem CF*:
Indeed, $y = \sum_{i=1}^{r} \lambda_i \delta_{v_i}$ where $\lambda_i > 0$, $\sum_i \lambda_i = 1$ and $v_i \in K$,
implying $p^T y = \sum_i \lambda_i p(v_i) \geq p_{\min}$ and thus $p^T y = p_{\min}$

Moreover, $\boxed{v_1, \ldots, v_r \text{ are global minimizers over } K}$

Finding the global minimizers $v_1, \ldots, v_r$:
$\{v_1, \ldots, v_r\} = V(I)$, where $I = \text{Ker } M(y)$ is a radical
0-dimensional ideal

Therefore, one can apply the **eigenvalue method**

Implementing the Extraction Procedure [HL 05] [Jibetean-L 05]

1. Construct (in a greedy way) a set $\mathcal{A}$ indexing a maximum nonsingular principal submatrix of $M_t(y)$. Thus, $\mathcal{A} \subseteq S_{t-d}$, $\mathcal{B} := \{x^\alpha \mid \alpha \in \mathcal{A}\}$ is a basis of $\mathbb{R}[x_1, \ldots, x_n]/I$

2. Construct the multiplication matrix M_{x_i} w.r.t. basis $\mathcal{B}$ in the following way:
M_0:= principal submatrix of $M_t(y)$ indexed by $\mathcal{A}$
U_i:= submatrix of $M_t(y)$ with rows in $\mathcal{A}$ and columns in $x_i\mathcal{A} := \{\alpha + e_i \mid \alpha \in \mathcal{A}\}$
Then, $M_{x_i} = {M_0}^{-1} U_i$

3. Compute the right eigenvectors $\zeta_{v_i} = (v_i^\alpha)_{\alpha \in \mathcal{A}}$ $(i = 1, \ldots, r)$ of $M_h = c_1 M_{x_1} + \ldots + c_n M_{x_n}$ for a random linear combination $h(x) = c_1 x_1 + \ldots + c_n x_n$

4. Deduce $v_1, \ldots, v_r$ from $\zeta_{v_1}, \ldots, \zeta_{v_r}$

Finite Convergence in the Finite Variety Case

Theorem: [Laurent 2002]
Assume the equations: $h_1(x) = 0 \ldots h_k(x) = 0$ are present in the description of K and

$$V := \{x \in \mathbb{C}^n \mid h_1(x) = 0 \ldots h_k(x) = 0\} \text{ is finite}$$

Then, $p_t^* = p_{\min}$ for t large enough

Sketch of proof: Use the Flat Extension Theorem

$J := \langle h_1, \ldots, h_k \rangle$ is 0-dimensional

$\mathcal{B} := \{b_1, \ldots, b_s\}$ a basis of $\mathbb{R}[x_1, \ldots, x_n]/J$

$d_{\mathcal{B}} := max(\deg(b_1), \ldots, \deg(b_s))$

Assume $M_t(y) \succeq 0$, $M_{t-d_j}(h_j * y) = 0$ **for** $j = 1, \ldots, k$

Show: For t large enough, y has a representing measure on K

For $\alpha \in S_t$, $x^\alpha \equiv \sum_{r=1}^s \lambda_r b_r \mod J$ for some scalars λ_r

Thus, $x^\alpha - \sum_{r=1}^s \lambda_r b_r = \sum_{j=1}^k u_j h_j$ for some polynomials u_j

For t **large enough,**

$$M_t(y) h_j = 0 \ \ \forall j \leq k \Longrightarrow M_t(y)(\textstyle\sum_j u_j h_j) = 0$$

Thus, the α-th column of $M_t(y)$ is a linear combination of its columns indexed by $S_{d_{\mathcal{B}}}$; i.e., $M_t(y)$ is a flat extension of $M_{d_{\mathcal{B}}}(y)$

$\rightsquigarrow$ exists a representing measure μ for y

Verify that the support of μ is contained in K

Instance: The 0/1 Case

Assume the equations $h_i := x_i^2 - x_i = 0$ $(i = 1, \ldots, n)$ are present in the description of K

Then one can use the conditions $M_t(h_i * y) = 0$ to decrease the number of variables:

$$y_\alpha = y_{\text{Supp}(\alpha)} \text{ for all } \alpha$$

$\leadsto$ The moment relaxation (MOMt) can be written as a program involving only variables y_α with $\alpha \in \{0, 1\}^n$

$\leadsto$ Finite convergence in n steps

$\leadsto$ The hierarchy of bounds p_t^* refines the SDP hierarchies of Sherali-Adams [1990], Lovász-Schrijver [1991]

Example: The Stable Set Problem

For a graph $G = (V, E)$,

$$\alpha(G) = \max \sum_{i \in V} x_i \text{ s.t. } x_i x_j = 0 \ (ij \in E), \ x_i^2 = x_i \ (i \in V)$$

The moment relaxation of order $t = 1$:

$$\max \sum_{i \in V} y_i \text{ s.t. } M_1(y) \succeq 0, \ y_{ij} = 0 \ (ij \in E), \ y_{\emptyset} = 1$$

gives Lovász' theta number $\vartheta(G)$

Finite convergence in $\alpha(G)$ steps

Example: The Max-Cut Problem

For a graph $G = (V, E)$ with edge weights $w \in \mathbb{R}^E$, the maximum weight of a cut is given by

$$\max \sum_{ij \in E} w_{ij}(1 - x_i x_j) \;\text{ s.t. } x_i^2 = 1 \; (i \in V)$$

The moment relaxation of order $t = 1$:

$$\max \sum_{ij \in E} w_{ij}(1 - X_{ij}) \;\text{ s.t. } X \succeq 0,\; X_{ii} = 1 \; (i \in V)$$

is the basic SDP relaxation used, e.g., by Goemans-Williamson for their 0.878-approximation algorithm

Finite convergence in n steps
Proven: $n/2$ steps are necessary
Conjecture: $n/2$ steps are sufficient

Bounded Moment Matrices

Theorem [Berg-Christensen-Ressel 1976]
Let $y \in \mathbb{R}^{\mathbb{N}^n}$ with, say, $y_0 = 1$
Assume $M(y) \succeq 0$ and y is bounded
Then, y has a representing measure on $[-1,1]^n$

Lemma 1: $|y_\alpha| \leq C \;\; \forall \alpha \in \mathbb{N}^n \Longrightarrow |y_\alpha| \leq 1 \;\; \forall \alpha \in \mathbb{N}^n$

Proof: $y_\alpha^2 \leq y_0 y_{2\alpha} = y_{2\alpha}$, which implies $|y_\alpha| \leq y_{2\alpha}^{1/2}$

Therefore, $|y_\alpha| \leq (y_{2^k\alpha})^{1/2^k} \leq C^{1/2^k}$ for any $k \geq 1$.

Letting $k \to \infty$, we find that $|y_\alpha| \leq 1$.

Lemma 2: Assume y is an extreme point of
$\mathcal{S} := \{y \in \mathbb{R}^{\mathbb{N}^n} \mid M(y) \succeq 0, y_0 = 1, y \text{ bounded}\}$
Then y has a representing measure on $[-1,1]^n$

Proof: Fix $\alpha_0 \in \mathbb{N}^n$. For $\epsilon = \pm 1$, define $z^{(\epsilon)} \in \mathbb{R}^{\mathbb{N}^n}$ with

$$\boxed{z_\alpha^{(\epsilon)} := y_\alpha + \epsilon y_{\alpha+\alpha_0} \; (\alpha \in \mathbb{N}^n)}$$

Claim: $M(z^{(\epsilon)}) \succeq 0$

Proof: Show $p^T M(z^{(\epsilon)}) p \geq 0$. Set $x := (p^2) * y$. Then, $M(x) \succeq 0$, since $q^T M(x) q = (pq)^T M(y)(pq)$. As x is bounded, $|x_{\alpha_0}| \leq x_0$ by L.1. This gives $p^T M(z^{(\epsilon)}) p \geq 0$.

Now, $y = \frac{z_0^{(1)}}{2} \underbrace{\frac{z^{(1)}}{z_0^{(1)}}}_{\in \mathcal{S}} + \frac{z_0^{(-1)}}{2} \underbrace{\frac{z^{(-1)}}{z_0^{(-1)}}}_{\in \mathcal{S}}$.

Thus, $y = \frac{z^{(1)}}{z_0^{(1)}}$

This implies: $y_{\alpha+\alpha_0} = y_\alpha y_{\alpha_0}$ for any α, α_0

Therefore, $y = \zeta_v$, where $v := (y_{e_i})_{i=1}^n \in [-1,1]^n$

End of Proof

Note: Find again finite convergence in 0/1 case: If $M_n(y) \succeq 0$ with $y_\alpha = y_{\text{Supp}(\alpha)}$ $\forall \alpha$, then y has a representing measure on $\{0,1\}^n$

Proof: Extend y to $y \in \mathbb{R}^{\mathbb{N}^n}$ by $y_\alpha = y_{\text{Supp}(\alpha)}$ $\forall \alpha$. Then, $M(y) \succeq 0$

and y is bounded.

Characterizing Sequences of Moments of Measures with Compact Support

Theorem: [Berg-Maserick 1984]
$y \in \mathbb{R}^{\mathbb{N}^n}$ has a representing measure with compact support

$$\Longleftrightarrow$$

$M(y) \succeq 0$ and y is bounded w.r.t. some absolute value $\sigma(.)$,
i.e., $\exists C$ constant s.t. $|y_\alpha| \leq C\sigma(\alpha) \quad \forall \alpha \in \mathbb{N}^n$

Absolute value: $\sigma : \mathbb{N}^n \to \mathbb{R}_+$ with $\sigma(0) = 1$ and
$\sigma(\alpha + \beta) \leq \sigma(\alpha)\sigma(\beta) \;\; \forall \alpha, \beta \in \mathbb{N}^n$

Proof of $\Longrightarrow$: If $y_\alpha = \int_K x^\alpha \mu(dx)$, then $|y_\alpha| \leq \underbrace{\mu(K)}_{:=C} \underbrace{\sup_{x \in K} |x^\alpha|}_{:=\sigma(\alpha)}$

Proof of $\Longleftarrow$: Analogous to case $\sigma(.) = 1$

Sometimes Local (Component-Wise) Implies Global: Two Fundamental Tools for Lasserre's SOS Representation Result

$y \in \mathbb{R}^{\mathbb{N}^n} \rightsquigarrow n$ monovariate sequences $y^{(i)} \in \mathbb{R}^{\mathbb{N}}$ $\quad (i = 1, \ldots, n)$

$$\boxed{y_k^{(i)} := y_{ke_i} \quad \text{for } k \in \mathbb{N}}$$

Carleman's Condition: [Nussbaum 1966]

Assume $\sum_{k=0}^{\infty} (y_{2k}^{(i)})^{-1/2k} = \infty \quad$ for $i = 1, \ldots, n$

Then, y has a representing measure

Lemma: [Lasserre 2004] Assume $M_t(y) \succeq 0$ and $y_{2k}^{(i)} \leq C$ for $k = 1, \ldots, t$, $i = 1, \ldots, n$.

Then, $|y_\alpha| \leq C \quad$ for all $|\alpha| \leq 2t$

Lasserre's SOS Representation Result

Theorem: [Lasserre 2004]

Assume $p \geq 0$ on $\mathbb{R}^n$.

Then, $\forall \epsilon > 0 \; \exists t \in \mathbb{N}$ for which $\underbrace{p + \epsilon \left(\sum_{k=0}^{t} \sum_{i=1}^{n} \frac{x_i^{2k}}{k!} \right)}_{p_{\epsilon,t}}$ is SOS

$\rightsquigarrow$ New relaxation scheme for $\boxed{p_{\min} = \inf_{x \in \mathbb{R}^n} p(x)}$

$$\boxed{p^*_{\epsilon,t} := \inf \; y^T p_{t,\epsilon} \;\; \text{s.t. } M_t(y) \succeq 0, \;\; y_0 = 1}$$

For $\epsilon > 0$, $p_{\min} \leq p^*_{\epsilon,t}$ [for t large enough] $\leq p_{\min} + \epsilon \sum_{i=1}^n e^{x_i^*}$ [if $p_{\min} = p(x^*)$]