Minimal Sums of Squares in a free $*$-algebra

Martin Harrison

March 2, 2010

Introduction

$\mathbb{R}\langle X, X^* \rangle$ denotes the space of polynomials in the non-commuting variables $X_1, \ldots, X_n$, $X_1^*, \ldots, X_n^*$ over the reals.

$\mathbb{R}_d\langle X, X^* \rangle$: those of degree at most d

$\beta = \{m_1, \ldots, m_N\}$ is a basis of monomials for $\mathbb{R}_d\langle X, X^* \rangle$

$V = (m_1, \ldots, m_N)^t$ is the tautological vector, $V^* = (m_1^*, \ldots, m_N^*)$

Representing a SOS

- A single square:

$$f^*f = \Big(\sum_{i=1}^{N} c_i X_i\Big)^* \Big(\sum_{i=1}^{N} c_i X_i\Big) = V^* C C^t V, \qquad C = (c_1, \ldots, c_N)^t$$

Representing a SOS

- A single square:

$$f^* f = \Big(\sum_{i=1}^{N} c_i X_i\Big)^* \Big(\sum_{i=1}^{N} c_i X_i\Big) = V^* C C^t V, \qquad C = (c_1, \dots, c_N)^t$$

- A SOS:

$$\sum_{j=1}^{M} f_j^* f_j = \sum_{j=1}^{M} V^* C_j C_j^t V = V^* \Big(\sum_{j=1}^{M} C_j C_j^t\Big) V, \quad A = \sum_{j=1}^{M} C_j C_j^t$$

Representing a SOS

- A single square:

$$f^*f = \Big(\sum_{i=1}^{N} c_i X_i\Big)^* \Big(\sum_{i=1}^{N} c_i X_i\Big) = V^* C C^t V, \qquad C = (c_1, \ldots, c_N)^t$$

- A SOS:

$$\sum_{j=1}^{M} f_j^* f_j = \sum_{j=1}^{M} V^* C_j C_j^t V = V^* \Big(\sum_{j=1}^{M} C_j C_j^t\Big) V, \qquad A = \sum_{j=1}^{M} C_j C_j^t$$

- The matrix A is PSD. The correspondence

$$\textit{PSD matrix} \leftrightarrow \textit{SOS}$$

is *not* one-to-one.

An Example

- Let $V = (1, X, X^*)^T$, and $P = V^*V = X^*X + XX^* + 1$.

An Example

- Let $V = (1, X, X^*)^T$, and $P = V^*V = X^*X + XX^* + 1$.
- Define $M = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$

An Example

- Let $V = (1, X, X^*)^T$, and $P = V^*V = X^*X + XX^* + 1$.
- Define $M = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$
- $P = V^*(I - tM)V$ for any $t \in \mathbb{R}$

An Example

- Let $V = (1, X, X^*)^T$, and $P = V^*V = X^*X + XX^* + 1$.
- Define $M = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$
- $P = V^*(I - tM)V$ for any $t \in \mathbb{R}$
- $P = \left(X + \frac{\sqrt{2}}{2}\right)^* \left(X + \frac{\sqrt{2}}{2}\right) + \left(X^* - \frac{\sqrt{2}}{2}\right)^* \left(X^* - \frac{\sqrt{2}}{2}\right)$ is given by $I + \frac{1}{\sqrt{2}}M$

The Question

Question: In general, what number of squares will suffice for an arbitrary SOS? Can we neatly characterize this **minimal number** in terms of degree and dimension? How to compute?

Upper Bounds

- Carathéodory's convex hull theorem: $N(2d) + 1$ squares.

Upper Bounds

- Carathéodory's convex hull theorem: $N(2d) + 1$ squares.
- Gram matrix diagonalization gives N(d) squares:

Upper Bounds

- Carathéodory's convex hull theorem: $N(2d) + 1$ squares.
- Gram matrix diagonalization gives N(d) squares:
- Write $f_i = (FV)_i$, Compute Cholesky decomposition $F^T F$. May have full rank, but further reduction is possible...

Upper Bounds

- Recall M from example satisfying $V^*MV = 0$. Such M are symmetric, are not PSD (or NSD), and exist for all $n, d \geq 1$.

Upper Bounds

- Recall M from example satisfying $V^*MV = 0$. Such M are symmetric, are not PSD (or NSD), and exist for all $n, d \geq 1$.
- Take $A \succeq 0$, $P = V^*AV$. Some combination $A + cM$ is outside the PSD cone, so for some t we have

  $$tA + (1-t)(A + cM) \in \partial PSD$$

  so $P = V^*AV = V^*(A + tcM)V$ is the sum of $N(d) - 1$ squares.

Upper Bounds

- Recall M from example satisfying $V^*MV = 0$. Such M are symmetric, are not PSD (or NSD), and exist for all $n, d \geq 1$.
- Take $A \succeq 0$, $P = V^*AV$. Some combination $A + cM$ is outside the PSD cone, so for some t we have

$$tA + (1-t)(A + cM) \in \partial PSD$$

so $P = V^*AV = V^*(A + tcM)V$ is the sum of $N(d) - 1$ squares.
- For any dimension, and degree $d \leq 2$, this is the best we can do:

$$\sum_{i \geq 2} m_i^* m_i$$

always requires $N(d) - 1$, and for $d = 1$, so does the full sum of squares of monomials

Commutative Case

The sum $1 + X_1^2 + X_2^2 + \ldots + X_N^2$ *cannot* be expressed as the sum of N squares.

More on the Bounds

The sum of lowest (positive) degree and highest degree monomial squares cannot be reduced.
(since $m_i m_j = m_l m_k$ requires $m_i = m_l$ and $m_j = m_k$)

The bound is tight for *hereditary* SOS for the same reason.

This lower bound agrees with the the upper bound for $d \leq 2$, but is much smaller in general.

Semidefinite programming

- We have exactly the problem of minimizing the rank of the Gram matrix subject to positivity constraints:

$$\begin{aligned} \min \quad & \operatorname{rank} X \\ \text{s.t.} \quad & V^* X V = P, \\ & X \succeq 0 \end{aligned}$$

Semidefinite programming

- We have exactly the problem of minimizing the rank of the Gram matrix subject to positivity constraints:

$$\begin{aligned} \min \quad & \operatorname{rank} X \\ \text{s.t.} \quad & V^* X V = P, \\ & X \succeq 0 \end{aligned}$$

- Rank is not convex.

Semidefinite programming

- We have exactly the problem of minimizing the rank of the Gram matrix subject to positivity constraints:

$$\begin{aligned} \min \quad & \operatorname{rank} X \\ \text{s.t.} \quad & V^* X V = P, \\ & X \succeq 0 \end{aligned}$$

- Rank is not convex.
- The trace heuristic: trace minimization will recover a minimal rank solution under certain conditions

Trace Heuristic

Restricted isometry condition for the trace heuristic.[Fazel, Parrilo, Recht]

For a map $\mathcal{A} : \mathbb{R}^{n\times m} \to \mathbb{R}^M$ define the *r-restricted isometry constant* δ_r to be the smallest value d for which

$$(1-d)\|X\| \leq \|\mathcal{A}(X)\| \leq (1+d)\|X\|$$

whenever X is of rank at most r.

Newton Chip

- Reduce the monomial basis first. Based on Newton polytope.
  Input: $f = \sum a_w w$, a SOS
  1. Set $W = \emptyset$
  2. For each word $w^* w$ in the support of f:
  2.1. For each $0 \leq i \leq deg(w)$, if $rc(w, i)$ is admissible (satisfies certain degree bounds), append it to W.

Newton Chip

- Reduce the monomial basis first. Based on Newton polytope.
  Input: $f = \sum a_w w$, a SOS
  1. Set $W = \emptyset$
  2. For each word $w^* w$ in the support of f:
  2.1. For each $0 \leq i \leq deg(w)$, if $rc(w, i)$ is admissible (satisfies certain degree bounds), append it to W.
- *Theorem* [Klep, Povh] A symmetric f is SOS exactly when there is a PSD G satisfying

$$W^* G W = f$$

Newton Chip

- Reduce the monomial basis first. Based on Newton polytope.
  Input: $f = \sum a_w w$, a SOS
  1. Set $W = \emptyset$
  2. For each word $w^* w$ in the support of f:
  2.1. For each $0 \leq i \leq deg(w)$, if $rc(w, i)$ is admissible (satisfies certain degree bounds), append it to W.
- *Theorem* [Klep, Povh] A symmetric f is SOS exactly when there is a PSD G satisfying

$$W^* G W = f$$

- Augmented Newton Chip: If $a_{w^* w} = 0$ and $w^* w \neq v^* z$ for $v \neq z$ in W (obtained from Newton chip), then throw out u.