Semidefinite and Second Order Cone Programming Seminar
Fall 2012
Lecture 10

Instructor: Farid Alizadeh
Scribe: Joonhee Lee

11/19/2012

1 Overview

In this lecture, we show that a number of sets and functions related to the notion of *sum-of-squares (SOS)* are SD-representable. We will start with positive polynomials. Then, we introduce a general algebraic framework in which the notion of sum-of-squares can be formulated in very general setting.

2 Polynomials

Recall the cone of nonnegative univariate polynomials:

$$\mathcal{P}_{2d}[t] = \{p(t) = p_0 + p_1 t + p_2 t^2 + \cdots + p_{2d} t^{2d} \geqslant 0 \quad \forall t \in \mathbb{R}$$

Earlier we have examined this cone and have shown that it is SD-representable. We now consider the case of multivariate polynomials. The set of nonnegative polynomials is

$$\mathcal{P}_{n,d}[t_1, \ldots, t_d] = \{ \ P(t_1, \ldots, t_n) \geq 0 \qquad \forall t \in \mathbb{R}\}$$

Recall that in the case of univariate polynomials, a polynomial $p(t) \geqslant 0$ for all $t \in \mathbb{R}$ if and only if there are two polynomials $p_1(t), p_2(t)$ such that $p(t) = p_1^2(t) + p_2^2(t)$. In other words, a univariate polynomial is nonnegative over the real line if and only if it is a sum of squares. For multivariate polynomials this is no longer true.

Example 1 (Motzkin Polynomial) consider the following polynomial:

$$p(x, y) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1$$

This polynomial is nonnegative for all $x, y \in \mathbb{R}$ because

$$\frac{x^4y^2 + x^2y^4 + 1}{3} \geq \sqrt[3]{x^6y^6}$$

by the arithmetic-geometric inequality. On the other hand it cannot be a sum of two square polynomials. If there were polynomials $p_1(x, y), \ldots, p_n(x, y)$ such that $\sum_i p_i^2(x, y) = p(x, y)$, then we must have that each $p_i(x, y)$ is of the form

$$x^2\big(a_iy^2 + b_iy + c_i\big) + x(d_iy^2 + e_iy + f_i) + g_iy^2 + h_iy + k_i$$

But it is immediate that $\sum a_i^2 = \sum c_i^2 = \sum f_i^2 = \sum g_i^2 = \sum h_i^2 = 0$, and thus, $a_i = c_i = f_i = g_i = h_i = 0$ for all $i = 1, \ldots, n$, otherwise the sum of squares will contains terms which do not appear in $p(x, y)$. This leaves us with

$$p(x, y) = \sum_i \big(b_ix^2y + d_ixy^2 + e_ixy + k_i\big)^2$$

But this implies that $\sum_i e_i^2 = -3$ which is impossible. This completes the proof that the Motzkin polynomial is not sum of squares of other polynomials. ∎

Define

$$\Sigma_{n,d}[t_1, \ldots, t_d] \stackrel{\text{def}}{=} \{p(t_1, \ldots, t_n) \mid p(t_1, \ldots, t_n) = \sum_i^N p_i^2(t_1, \ldots, t_d)\}$$

It is clear that $\Sigma_{n,2d} \subseteq \mathcal{P}_{n,2d}$, since any sum-of-square polynomial is nonnegative. The Motzkin example shows that the inclusion is proper.

When can we guarantee P= Σ? It turns out that the following are the only possible cases where $\mathcal{P} = \Sigma$:

1. **Case 1:**$n = 1$**.** This is the case of univariate polynomials which we have proved earlier
2. **Case 2:** d**=2.** All polynomials of degree two can be written in the form of $\mathbf{t}^\top A\mathbf{t} + \mathbf{b}^\top \mathbf{t} + c$, with $A \succ 0$ (A is nonsingular). To see this, first suppose $\mathbf{d}$ is a vector satisfying $\mathbf{d} = B^{-\top}\mathbf{b}$, and let $A = B^\top B$ Then

   $$\mathbf{t}^\top B^\top B\mathbf{t} + \mathbf{b}^\top \mathbf{t} + c = \big(B\mathbf{t} + \tfrac{\mathbf{d}^\top \mathbf{t}}{2}\big)^\top \big(B\mathbf{t} + \tfrac{\mathbf{d}^\top \mathbf{t}}{2}\big)$$

   which is a sum-of-squares.
3. **Case 3:** $d = 4, n = 3$**.** A very special case is three-variable polynomials of degree four, where nonnegative polynomials are always sum-of-squares. The proof is somewhat hairy.

Hilbert's seventeenth problem states that all nonnegative polynomials are sum-of-squares of *rational functions.* In other words for each $p(\mathbf{t}) = p(t_1, \ldots, t_n) \geqslant 0$ for all $t_i \geqslant 0$, there are polynomials $p_1(\mathbf{t}), \ldots, p_N(\mathbf{t}), q(\mathbf{t})$ such that

$$q^2(\mathbf{t})p(\mathbf{t}) = p_1^2(\mathbf{t}) + \cdots + p_N^2(\mathbf{t})$$

Hilbert's seventeenth problem was proved by E. Artin in 1923, and in the process laid out the foundation of a field of algebra known as *real algebraic geometry*.

Assuming that the greatest common denominator of $q(\mathbf{t})$ and $p_i(\mathbf{t})$ is 1, there is still no satisfactory bound on the number N, the number of squares, and D, the largest degree among $q(\mathbf{t})$ and $p_i(\mathbf{t})$. In fact, bounds that are known for N and D are not only depend on n, the number of variable, and d the degree of $p(\mathbf{t})$, but also the coefficients of p as well.

3 General Algebra

Consider (A, B, $\diamond$) where $\diamond$ is a bilinear operator that is $\diamond$: A $\times$ A $\rightarrow$ B. A and B are finite-dimensional real linear spaces with $\dim A = m$ and $\dim(B) = n$. Note that bilinearity assumption is equivalent to the distributive law:

- $a \diamond (\alpha b + \beta c) = \alpha a \diamond b + \beta a \diamond c$
- $(\alpha b + \beta c) \diamond a = \alpha b \diamond a + \beta c \diamond a$

Note also that bilinearity means that there are matrices Q_i such that $(a \diamond b)_i$ $= a^T Q_i b$. Indeed, to each element $\mathbf{a} \in A$ we can associate a linear transformation $L_{\mathbf{a}}$ mapping $A \rightarrow B$, that is $L_{\mathbf{a}}\mathbf{b} = \mathbf{a} \diamond \mathbf{b}$. The linear transformation $L_{\mathbf{a}}$ may be represented by a $n \times m$ matrix (also written as $L\mathbf{a}$, whose entries are linear forms in a_i.

Our object of interest is the following set:

$$\Sigma_\diamond = \left\{ \sum \mathbf{a}_i \diamond \mathbf{a}_i \mid \mathbf{a}_i \in A \right\} \subseteq B.$$

which we call the *sum-of-squares cone* (or the SOS cone) associated with the algebra. Note that $\Sigma_\diamond$ is a convex cone, since adding to sums of squares creates another sum of squares.

Example 2 Suppose $A = \mathcal{P}_d[t]$, the set of degree d univariate polynomials, and $B = \mathcal{P}_{2d}[t]$, the set of degree $2d$ univariate polynomials. Then the $(\mathcal{P}_d[t], \mathcal{P}_{2d}[t], *)$ forms an algebra, with $*$ indicating the multiplication of polynomials. If we represent each polynomial by the vector of its coefficients, then $*$ is the *convolution* operation:

$$(p_0, p_1, \ldots, p_d) * (q_0, q_1, \ldots, q_d) = (p_0 q_0, p_0 q1 + p_1 q_0, \ldots, p_0 q_k + p_1 q_{k-1} + \cdots, \ldots, p_d q_d)$$

For this algebra Σ_* is the set of polynomials of degree $2d$ which are sum of squares of polynomials. As we know, in this case this cone is exactly the cone of polynomials of degree $2d$ which are nonnegative for every $t \in \mathbb{R}$.

We make three assumptions on the algebra $(A, B, \diamond)$ without loss of generality, which will make the presentation cleaner and more streamlined.

scribe: Joonhee Lee
Lecture 10 **Date:** 11/19/2012

1. $\diamond$ **is commutative:** $\mathbf{a} \diamond \mathbf{b} = \mathbf{b} \diamond \mathbf{a}$. If $\diamond$ it is not commutative, then we can replace it with its *anti-commutator*:

$$\mathbf{a} \bar{\diamond} \mathbf{b} = \frac{\mathbf{a} \diamond \mathbf{b} + \mathbf{b} \diamond \mathbf{a}}{2}$$

Note that for the algebras $(A, B, \diamond)$ and $(A, B, \bar{\diamond})$ we have $\Sigma_\diamond = \Sigma_{\bar{\diamond}}$.

2. $B = \text{span}(A \diamond A)$ This assumption ensures that B does not contain elements that are not somehow generated by elements from A, and in turn results in $\Sigma_\diamond$ to be full dimensional in B.

3. **The mapping** $L : A \to \mathbb{R}^{n \times m}$ **is injective.**

$$L_{\mathbf{x}_1} = L_{\mathbf{x}_2} \Rightarrow \mathbf{x}_1 = \mathbf{x}_2.$$

Without the this assumption, there is no way to distinguish $\mathbf{x}_1$ and $\mathbf{x}_2$. In this case we observe that the relation $L_{\mathbf{x}_1} = L_{\mathbf{x}_2}$ defines an equivalence relation on A:

$$\mathbf{x}_1 \simeq \mathbf{x}_2 \Leftrightarrow L_{\mathbf{x}_1} = L_{\mathbf{x}_2}$$

Then, by replacing A, with $A/\simeq$, the set of equivalence classes, and defining on $A/\simeq$:

$$[\mathbf{x}] \diamond \mathbf{y}] \stackrel{\text{def}}{=} [\mathbf{x} \diamond \mathbf{y}]$$

Using commutativity it is easy to see that this definition is consistent, that is, if $\mathbf{x}_1 \simeq \mathbf{x}_2$ and $\mathbf{y}_1 \simeq \mathbf{y}_2$ then $\mathbf{x}_1 \diamond \mathbf{y}_1 \simeq \mathbf{x}_2 \diamond \mathbf{y}_2$. Therefore, $(A/\simeq, B, \diamond)$ satisfies the third assumption.

Lemma 1 *With assumptions 1, 2, and 3,*

1. $\Sigma_\diamond$ *is full dimensional convex cone.*
2. *Every element in* $\Sigma_\diamond$ *is sum of at most* $n = \dim(B)$ *squares.*

Proof:
1) Since the sum of two sums of squares is a another sum of square, $\Sigma_\diamond$ is convex. To prove it is full-dimensional we claim that $B = \Sigma_\diamond - \Sigma_\diamond$. First note that for any $\mathbf{a}, \mathbf{b} \in A, \mathbf{a} \diamond \mathbf{b} = (\frac{a+b}{2})^{\diamond 2} - (\frac{a-b}{2})^{\diamond 2}$; this follows from commutativity. On the other hand by assumption 2, every element in B is of the form $\sum_i \mathbf{a}_i \diamond \mathbf{b}_i$. This shows that $B = \Sigma_\diamond - \Sigma_\diamond$, and thus, Σ is full dimensional. 2) By Caratheodory's Theorem for cones, every element of $\Sigma_\diamond$ is sum of at most n extreme rays. But the extreme rays of Σ are among perfect squares $\mathbf{a} \diamond \mathbf{a}$, so each element of $\Sigma_\diamond$ is sum of at most n squares. ■

(A Trivial Example) $A = B = \mathbb{C}$: Complex number under ordinary multiplication. Then, $\Sigma_\diamond = \mathbb{C}$

Lemma 2 *If (A, B, $\diamond$) is formally real[1],then $\Sigma_\diamond$ is pointed. Conversely, if $\Sigma_\diamond$ is pointed and there are no nilpotent[2] elements, then (A, B, $\diamond$) is formally real.*

Proof:

$$\begin{aligned} & \text{If } -\sum_i a_i^{\diamond 2} \in \Sigma_\diamond, \text{then} \\ & \exists\, b_i : \sum_i b_i^{\diamond 2} = -\sum_i a_i^{\diamond 2} \Rightarrow \sum_i b_i^{\diamond 2} + \sum_i a_i^{\diamond 2} = 0 \\ \Rightarrow\ & a_i = 0 \text{ and } b_i = 0 \text{ (by formally real)} \end{aligned}$$

$$\sum_i a_i^{\diamond 2} = 0 \Rightarrow a_i^{\diamond 2} = 0 \Rightarrow a_i = 0 \text{ (by nilpotent)}$$

■

The dual of $\Sigma_\diamond$ is $\Sigma_\diamond^* = \{ z \mid \langle a, z\rangle \geq 0, \forall a \in\ \Sigma_\diamond\}$. Then,

Theorem 3 *$\Sigma_\diamond$ is a proper cone iff $\Sigma_\diamond^*$ is.*

Define 1 *Λ and Λ^* operators*

$$\begin{aligned} & (a, b, \diamond), \Sigma_\diamond, \Sigma_\diamond^* \\ & \Lambda_\diamond : B \rightarrow \mathbb{S}_A \textit{ where } \mathbb{S}_A \textit{ is a set of symmetric bilinear forms} \\ & (\Lambda_\diamond(w), a, b) \triangleq \langle w, a \diamond b\rangle_B \\ & a^T \Lambda b = b^T \Lambda a, (a, b) \in A, w \in B \end{aligned}$$

Theorem 4 *$w \in \Sigma_\diamond^*$ iff $\Lambda_\diamond(w) \succeq 0$.*

Proof: $\Rightarrow$:

$$\begin{aligned} & \Lambda_\diamond(w) \succeq 0 \\ & \Lambda_\diamond(w)(a, a) = \langle w, a \diamond a\rangle \geq 0 \\ & (\because w \in \Sigma_\diamond^* \text{ and } a \diamond a \in \Sigma_\diamond) \\ & \ \Rightarrow \Lambda(w) \geq 0 \end{aligned}$$

Next, we need to show ($\Leftarrow$)

$$\begin{aligned} \Lambda(w) \geq 0 \quad & \Rightarrow \quad \forall a \in A \text{ such that } \Lambda_\diamond(w)(a, a) \geq 0 \\ & \Rightarrow \quad \langle w, x\rangle \geq 0 \forall x \in \Sigma_\diamond \\ & \Rightarrow \quad w\ \in\ \Sigma_\diamond^* \end{aligned}$$

[1] $\sum_i a_i^{\diamond 2} = 0 \Rightarrow a_i = 0$
[2] $a_i^{\diamond 2} = 0$

■

Note that, if $\Lambda^* : \mathbb{S} \to \mathrm{B}$, then

$$\langle x, \Lambda(w)\rangle_{\mathbb{S}_A} = \langle x, \Lambda^*(x)\rangle_{\mathrm{B}} \quad \forall w \in \mathrm{B} \text{ and } x \in \mathbb{S}_A$$

Theorem 5 $u \in \Sigma_\diamond$ *iff* $\exists Y \succeq 0$ *such that* $u = \Lambda^*(Y)$.

Proof:
Show ($\Rightarrow$)

$$\begin{aligned} & \text{if } Y \succeq 0 \text{ and } \Lambda^*(mathitY) = u, \text{ then} \\ & \forall v \in \Sigma_\diamond^*,\ \langle u, v\rangle_{\mathrm{B}} = \langle \Lambda^*(Y, v\rangle_{\mathrm{B}} = \langle Y, \Lambda v\rangle_{\mathbb{S}_A} \geq 0 \\ & (\because Y \geq 0 \text{ and } \Lambda(v) \geq 0) \\ \Rightarrow\ & u \in \Sigma_\diamond^{**} = \Sigma_\diamond \end{aligned}$$

($\Leftarrow$)

$$\text{if } u \in \Sigma_\diamond, \text{ then } \exists a_i \text{ such that } u = \sum_i a_i^{\diamond 2}. \text{ Let } v \in B$$

$$\begin{aligned} \langle u, v\rangle_{\mathrm{B}} &= \langle \sum_i a_i^{\diamond 2}, v\rangle_{\mathrm{B}} = \sum_i \langle a_i^{\diamond 2}, v\rangle_{\mathrm{B}} \\ &= \sum_i \Lambda(v)(a_i, a_i) = \sum_i \langle \Lambda(v), a_i a_i^T\rangle_{\mathbb{S}_A}, \\ &\overset{(\text{Let } Y = \sum_i a_i a_i^T)}{=} \langle \Lambda(v), Y\rangle_{\mathbb{S}_A} = \langle v, \Lambda^*(Y\rangle_{\mathrm{B}} \\ &\Rightarrow u = \Lambda^*(Y) \end{aligned}$$

■

Example 3 For the algebra of univariate polynomials, the Λ operator can be computed as follows. If we represent $\Lambda(\boldsymbol{w})$ as a matrix, then its i, j entry, by definition, is given by $\langle \boldsymbol{w}, \boldsymbol{e}_i * \boldsymbol{e}_j\rangle$. But, $\boldsymbol{e}_i * \boldsymbol{e}_j = \boldsymbol{e}_{i+j}$, thus $\big(\Lambda(\boldsymbol{w})\big)_{ij} = w_{i+j}$. This means that the i, j entry of Λ in this case only depends on $i + j$, that is, all entries of $\Lambda(\boldsymbol{w})$ with the same $i + j$ must be equal. It follows that Λ has all entries on the reverse diagonals are equal.

$$\Lambda(\boldsymbol{w}) = \begin{pmatrix} w_0 & w_1 & \cdots & w_n \\ w_1 & w_2 & \cdots & w_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ w_n & w_{n+1} & \cdots & w_{2n} \end{pmatrix}$$

3.1 Squared functional systems

Of particular interest are algebras that are induced by *functional spaces.* Let $\mathcal{F} = \{f_1, f_2, \cdots, f_m\}$ where each $f_i : \Delta \to \mathbb{R}$, is a real-valued function. Let $F = \text{span}(\mathcal{F}) = \sum_i \alpha_i f_i(x), \quad \forall x \in \Delta$ be the linear space spanned by f_i, where $(\alpha_i f_i + \alpha_j f_j)(x) = \alpha_i f_i(x) + \alpha_j f_j(x)$.

Now define $\mathcal{S} = \{f_i, f_j\}$ and $(f_i, f_j)(x) = f_i(x) f_j(x)$

Let $S = \text{span}(\mathcal{I})$ $(F, S, \dot{)}$ is an algebra and $\Sigma_{\mathcal{F}} = \{\sum_i g_i^2 : g_i \in F\}$. The algebra $(F, S, \cdot)$ along with its SOS cone $\Sigma_{\mathcal{F}}$ is called *a squared function of system.* The univariate polynomial example given earlier is a special case where $\mathcal{F} = \{1, t, t^2, \ldots, t^d\}$, and $S = \{1, t, t^2, \ldots, t^{2d}\}$.

3.2 The semidefinite and second order cones as SOS cones

For two $p \times q$ matrices A and B define the *Cracovian multiplication* as follows:

$$A \diamond B = AB^T \text{ and } A \bar{\diamond} B = \frac{A \diamond B + B \diamond A}{2}$$

Then, for the algebra $(\mathbb{R}^{p \times q}, \mathbb{R}^{p \times p}, \bar{\diamond})$, the SOS cone $\Sigma_\diamond$ is exactly the cone of positive semidefinite $p \times p$ matrices.

4 Operations on algebras and their SOS cones

4.1 Bijective linear transformations

Let $(A, B, \diamond)$ be an algebra which as usual satisfies assumptions 1,2, and 3, and let C be another linear space, such that the linear transformation $F \mid B \to C$ is bijective. (Note that this means that necessarily $\dim(B) = n = \dim(C)$). Define a new binary operation $\circ \mid A \times A \to C$ by $L_\circ = FL_\diamond$. Then $(A, C, \circ)$ is an algebra, and if satisfying assumptions 1,2, and 3. Furthermore, $\Sigma_\circ = F\Sigma_\diamond$.

Definition 6 *Two cones* K_1 *and* K_2 *are linearly isomorphic if* $\exists$ *bijective linear transformation F :* $K_1 = FK_2$.

Thus, if Σ_1 is an SOS cone, and Σ_2 is a cone linearly isomorphic to Σ_1 then Σ_2 is also an SOS cone for some algebra.

4.2 Isomorphism and linear isomorphism among algebras

Let $(A_1, B_1, \diamond)$ and $(A_2, B_2, \diamond)$ be two algebras, and let there be two linear transformation F and G such that

$$\left.\begin{array}{l} F : A_1 \to A_2 \\ G : B_1 \to B_2 \end{array}\right\}$$

If both F and G are bijective, and we have

$$G(a \diamond_1 b) = F(a) \diamond_2 F(b)$$

we say that these algebras are *isomorphic*. Note that as opposed to ordinary algebraic structures we need two maps two define isomorphism.

Lemma 7 *If* $(A_1, B_1, \diamond_1)$ *and* $(A_2, B_2, \diamond_2)$ *are two algebras isomorphic to each other, then their SOS cones* $\Sigma_{\diamond_1}$ *and* $\Sigma_{\diamond_2}$ *are linearly isomorphic.*

Let $y \in \Sigma_{\diamond_2}$. Then

$$\begin{aligned} y = \sum_i y_i \diamond_2 y_i &= \sum_i F(x_i) \diamond_2 F(x_i) \quad \text{(for some } x_i \in A\text{, since } F \text{ is surjective)} \\ &= \sum_i G(x_i \diamond_1 x_i) \quad \text{(by definition of homomorphism)} \\ &= G\Big(\sum_i x_i \diamond_1 x_i\Big) \in G\big(\Sigma_{\diamond_1}\big) \quad \text{(by linearity).} \end{aligned}$$

The sequence of implications above goes through in both directions, establishing that $\Sigma_{\diamond_2} = G\big(\Sigma_{\diamond_1}\big)$. By definition, if G is bijective, then it is also a linear isomorphism between $\Sigma_{\diamond_1}$ and $\Sigma_{\diamond_2}$.

4.3 Direct sums of algebras

For k algebras $(A_1, B_1, \diamond_1), \cdots, (A_k, B_k, \diamond_k)$ define a new algebra

$$(A_1 \times \cdots \times A_k, B_1 \times \cdots \times B_k, \diamond)$$

with

$$\begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} \diamond \begin{pmatrix} b_1 \\ \vdots \\ b_k \end{pmatrix} = \begin{pmatrix} a_1 \diamond_1 b_1 \\ \vdots \\ a_k \diamond_k b_k \end{pmatrix}$$

This new algebra is called the *direct sum algebra*. It is immediate that

$$\Sigma_\diamond = \Sigma_{\diamond_1} \times \cdots \times \Sigma_{\diamond_k}$$

And the Λ operator is given by

$$\Lambda_\diamond\big((\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k)\big) = \Lambda_1 \oplus \cdots \oplus \Lambda_k.$$

4.4 Minkowski sum of algebras

Consider the algebras $(A_1, B, \diamond_1), \cdots, (A_k, B, \diamond_2)$ with possible different A_i, but all having the same B. The *Minkowski sum algebra* is the algebra $(A_1 \times \cdots \times A_k, B, \diamond)$, with

$$\begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} \diamond \begin{pmatrix} b_1 \\ \vdots \\ b_k \end{pmatrix} = a_1 \diamond_1 b_1 + \cdots + a_k \diamond_k b_k$$

Then we have

$$\Sigma_\diamond = \Sigma_{\diamond_1} + \cdots + \Sigma_{\diamond_k} \text{ and } \Lambda_\diamond(\boldsymbol{w}) = \Lambda_{\diamond_1}(\boldsymbol{w}) + \cdots + \Lambda_{\diamond_k}(\boldsymbol{w})$$

scribe: Joonhee Lee
Lecture 10 **Date:** 11/19/2012

4.5 Weighted sum-of squares

Combining the Minkowski sum and linear transformations we can show that a kind of *Weighted Sums Of Squares (WSOS)* is also in fact a sum-of squares. This follows from the following observation: Let $(A_i, B_i, \diamond_i)$, for $\imath = 1, \dots, k$ be algebras, and let F_i be linear transformations each mapping B_i to a common set B. Then the cone

$$F_1\Sigma_{\diamond_1} + \cdots + F_k\Sigma_{\diamond_k}$$

is also an SOS cone. Here is an example:

Example 4 Let $\mathbb{P}^d_{[0,\infty)} = \{p_d(t) \mid p(t) \geq \forall t \geq 0\}$, where $\deg(\mathbb{P}) = d$. Clearly this set is a convex cone. Let us show that it is in fact ans SOS cone.

$$\mathcal{P}^d_{[0,\infty)} = \begin{cases} \mathcal{P}^d + t\mathcal{P}^{d-2}, & \text{d is even} \\ \mathcal{P}^{d-1} + t\mathcal{P}^{d-1}, & \text{d is odd} \end{cases}$$

Proof:
We have : $p(t) \geq 0 \; \forall t \geq 0 \Leftrightarrow p(t^2) \geq 0 \; \forall t \in \mathbb{R}$. Thus, $q(t) = p(t^2) = p_1^2(t) + p_2^2(t)$ for some polynomials p_1 and p_2. We have,

$$\begin{aligned} p(t^2) &= q^2(t) + r^2(t) \quad \text{separating odd and even degree terms in } q(t) \text{ and } r(t): \\ &= \left(q_1(t^2) + tq_2(t^2)\right)^2 + \left(r_1(t^2) + tr_2(t^2)\right)^2 \\ &= q_1^2(t^2) + t^2q_2^2(t^2) + \underbrace{2tq_1(t^2)q_2(t^2)}_{=0} + r_1^2(t^2) + t^2r_2(t^2) + \underbrace{2tr_1(t^2)r_2(t^2)}_{=0} \\ &\quad (\divideontimes\ 2tq_1(t^2)q_2(t^2) = 2tr_1(t^2)r_2(t^2) = 0 \text{ since all terms of } p(t^2) \text{ have even degree}) \\ &\Rightarrow p(t) = p_1^2(t) + tp_2^2(t) \text{ (By changing } t^2 \text{ to } t) \end{aligned}$$

■

Thus, we have shown that the cone $\mathcal{P}^d_{[0,\infty]}[t]$ is a weighted sum of squares. However, note that the operation of multiplying by t is a bijective linear transformation mapping the space of degree d polynomials, to the space of degree $d+1$ polynomials with a zero constant term. Thus, $t\mathcal{P}_{\mathbb{R}}[t]$ is an SOS cone, and, its Minkowski sum $\mathcal{P}[t] + t\mathcal{P}[t]$ is also an SOS cone[3].

4.6 Isomorphism by change of basis and change of variables

Our presentation of algebras has been *basis-free* in that all arguments are made independent of any particular basis for A or B in the algebra $(A, B, \diamond)$. Of course in practice, the multiplication operator $L_\diamond$ is represented by a matrix, and this representation in turn depends on the particular basis chosen for A

[3] More precisely, if d is even then $\mathcal{P}^d_{[0,\infty]}[t] = \mathcal{P}^d[t] + t\mathcal{P}^{d-2}[t]$, and if d is odd $\mathcal{P}^d_{[0,\infty]}[t] = \mathcal{P}^{d-2}[t] + t\mathcal{P}^{d-1}[t]$.

and B. When we change basis for A, it is tantamount to replacing $L_\diamond(\mathbf{x})$ with $L_\diamond(F\mathbf{x})$ where F is the change of basis matrix. Similarly changing basis in B is the same as replacing $L_\diamond(\mathbf{x})$ with $L_\diamond(\mathbf{x})G$, where G is the change of basis matrix in B. Needless to say, the resulting algebras are all isomorphic to each other, and thus, the resulting SOS cones are linearly isomorphic.

Polynomials, and in general squared functional systems, are functional linear spaces. As such, there are many different ways of choosing a basis for them. For instance, for polynomials, in the ordinary representation $p_0 + p_1 t + \cdots + p_d t^d$ we are using the basis $\{1, t, t^2, \ldots, t^d\}$. However, there are many other bases: for instance $\{1, t+1, (t+1)^2, \ldots, (t+1)^d\}$ is another basis, and use of orthogonal polynomials (such as Chebychev, Legendre, Laguerre, etc.) are other ways of representing polynomials. Clearly, changing the basis in which polynomials are represented does not affect the SOS cone, nor does it change the fact that it equals the cone of nonnegative polynomials in the univariate case.

The second observation is the effect of change of variable even in a nonlinear way. In general, consider the set of polynomials of degree d which are nonnegative over a set $\Delta \subseteq \mathbb{R}$. Let $H : \Omega \to \Delta$ be an onto mapping from a set Ω to Δ. Note that Ω need not be a subset of $\mathbb{R}$; it is entirely arbitrary. Then the cone

$$\mathcal{P}_\Omega[H] = \{f(x) \mid f(x) = p_0 + p_1 H(x) + \ldots + p_d H^d(x) \geq 0 \forall x \in \Omega\}$$

is a convex cone linearly isomorphic to $\mathcal{P}_\Delta[t]$.

$$\begin{aligned} f(x) \geq 0 \quad &\Leftrightarrow \quad p_0 + p_1 H(x) + \cdots + p_d H^d(x) \geq 0 \quad \forall x \in \Omega \\ &\Leftrightarrow \quad p(t) \geq 0 \quad \forall x \in \Delta \end{aligned}$$

This means that from polynomials, we can construct other sets of SOS cones by functional composition and possibly by change of basis. Two examples follow:

Example 5 Consider the set of polynomials which are nonnegative over a finite interval say $[0, 1]$. $\mathcal{P}^d_{[0,1]} = \{\, p_d(t) \mid p(t) \geq \forall t \in [0, 1]\,\}$
Setting $H(t) = \frac{t}{1+t}$, we see that $H : [0, \infty] \to [0, 1]$. Then a polynomial $p(t) \geq 0$ for all $t \in [0, \infty]$ iff $p(s) \geq 0$ for all $s \in [0, \infty]$, and this is equivalent to $p\big(\frac{t}{1+t}\big) \geq 0$ for all $t \in [0, 1]$. But expanding this function, and observing that multiplying it by $(1 + t)^d$ does not change its sign over $[0, \infty]$, we see that

$$\forall t \in [\,0, \infty)\; p\left(\frac{t}{1+t}\right) = \frac{q(t)}{(1+t)^d} \geq 0, \Leftrightarrow \quad \forall t \in [0, 1] \Leftrightarrow q(t) \geq 0 \quad \forall t \in [0, 1]$$

This implies that $\mathcal{P}_{[0,1]}[t] \simeq \mathcal{P}_{[0,\infty]}[t]$, and both are weighted SOS, and thus SOS cones.

Example 6 *Cosine polynomials* are functions of the form $p_0 + p_1 \cos(t) + \cdots + p_d \cos(dt)$. We are interested in the cone

$$\mathcal{P}_{\cos}[t] = \{\mathbf{p} \mid p(t) = p_0 + p_1 \cos(t) + \cdots p_d \cos(td) \geq 0 \forall t \in \mathbb{R}\}$$

Since the domain of the function cos is the set [-1,1], and since each cos(kt) is a polynomial of degree k in cos(t) (known as the Chebychev polynomials of the first kind), we conclude that

$$\mathcal{P}_{\cos}[\mathsf{t}] \simeq \mathcal{P}_{[-1,1]}[\mathsf{t}] \simeq \mathcal{P}_{[0,1]}[\mathsf{t}] \simeq \mathcal{P}_{[0,\infty]}[\mathsf{t}]$$

References

[1] Nesterov, Y. *Squared functional systems and optimization problems*, In *High Performance Optimization*, pp. 405-440, Kluwer Acad. Publ., Dordrecht, 2000.

[2] Faybusovich, L. *On te Nesterov's Approach to Semi-Infinte Programming*, Acta. Appli. Math., 74, (2002), pp. 195-215.

[3] Papp, D, and Alizadeh, F., *Semidefinite Characterization of Sum-Of-Squares Cones in Algebras*, RUTCOR-rrr Report no 11-14, RUTCOR, Rutgers-State University of New Jersey, accepted for publication in *SIAM J. on Optimization.*