Sums of Squares Methods Explained: Part I

Grant Olney Passmore
grant.passmore@cl.cam.ac.uk
15 JJ Thomson Ave., University of Cambridge, CB3 0FD, UK

1 Sums of Squares Methods

From a high level, these methods rely upon the following two observations:

1. The question as to whether or not a real polynomial $p(\vec{x}) \in \mathbb{R}[\vec{x}]$ is a sum of squares (SOS) of real polynomials can be reduced to a semidefinite programming problem, and
2. The search for a Positivstellensatz refutation certifying the emptiness of a semialgebraic set defined by an RCF constraint system can be reduced to a finite sequence of searches for SOS decompositions.

Below we present an expository account of the difficult part of the first observation, due to Powers and Wörmann [PW99] and building upon the key insights of Choi, Lam, and Reznick [MDCR95]. An expository account of the second observation, due to Parrilo [Par03], will be the subject of a second note.

Given a PSD real polynomial $p(\vec{x}) \in \mathbb{R}[\vec{x}]$ that is a sum of squares of real polynomials, we seek an algorithm that will compute $p_1(\vec{x}), \ldots, p_n(\vec{x}) \in \mathbb{R}[\vec{x}]$ s.t. $p(\vec{x}) = \sum_{i=1}^{n} p_i^2(\vec{x})$.

Remark. From now on, unless specified otherwise, when we write “sum of squares” or “SOS” we mean “sum of squares of real polynomials in $\mathbb{R}[x_1, \ldots, x_n]$”.

1.1 The Powers-Wörmann SOS Decomposition

Let $p(\vec{x}) \in \mathbb{R}[\vec{x}]$ be SOS in t real polynomial squares. Then, $p(\vec{x})$ must have even degree. Let $deg(p(\vec{x})) = 2k$. Then, $\exists q_1, \ldots, q_t \in \mathbb{R}[\vec{x}]$ s.t. $deg(q_i(\vec{x})) \leq k$ and $p(\vec{x}) = \sum_{i=1}^{t} q_i^2(\vec{x})$.

A key observation is that we can now exactly characterise the finitely many possible power-products that could occur in each $q_i(\vec{x})$.

Definition 1.1. Let $\Lambda_n(d) = \{\alpha = \langle \alpha_1, \ldots, \alpha_n \rangle \in \mathbb{N}^n \mid \alpha_1 + \ldots + \alpha_n \leq d\}$.

Then, as $deg(q_i(\vec{x})) \leq k$ ($\forall\ 1 \leq i \leq t$), we see that the exponent vector α for each monomial occurring in each $q_i(\vec{x})$ must be a member of $\Lambda_n(k)$.

Note that $|\Lambda_n(k)| = \binom{n+k}{n}$. Suppose $u = \binom{n+k}{n}$ and fix some order on the members of $\Lambda_n(k)$ s.t. $\Lambda_n(k) = \{\beta_1, \ldots, \beta_u\}$. Let

$$\vec{\zeta} \ = \begin{pmatrix} \vec{x}^{\beta_1} \\ \vec{x}^{\beta_2} \\ \vdots \\ \vec{x}^{\beta_u} \end{pmatrix}$$

and let $A \in \mathbb{R}^{tu}$ be the $u \times t$ matrix whose ith column is the vector of coefficients of $q_i(\vec{x})$

$$A = \begin{pmatrix} c_{\beta_1,1} & c_{\beta_1,2} & \cdots & c_{\beta_1,t} \\ c_{\beta_2,1} & c_{\beta_2,2} & \cdots & c_{\beta_2,t} \\ \vdots & \vdots & \vdots & \vdots \\ c_{\beta_u,1} & c_{\beta_u,2} & \cdots & c_{\beta_u,t} \end{pmatrix}$$

s.t. $q_i(\vec{x}) = \sum_{j=1}^{u} c_{\beta_j,i}\vec{x}^{\beta_j}$.

Then,

$$p(\vec{x}) \ = \ \sum_{i=1}^{n} q_i^2(\vec{x}) \quad \iff \quad p(\vec{x}) \ = \ \vec{\zeta}^{\,T}(AA^T)\vec{\zeta}.$$

Definition 1.2 (Gram matrix)**.** The $u \times u$ matrix $B = (AA^T) \in \mathbb{R}^{u^2}$ given above is called the *Gram matrix* for $p(\vec{x})$ w.r.t. $q_1(\vec{x}), \ldots, q_t(\vec{x})$.

Definition 1.1. A Gram matrix of a polynomial is necessarily *symmetric*.

Definition 1.3 (PSD matrix)**.** Let B be a $u \times u$ symmetric matrix. B is said to be *positive semidefinite* (PSD) iff

$$\forall \vec{r} \in \mathbb{R}^u \quad (\vec{r}\ B\ \vec{r}^{\,T} \ \geq_{\mathbb{R}} \ 0)$$

where $\vec{r} \in \mathbb{R}^u$ is taken to be a row vector. Equivalently, B is PSD iff all of its eigenvalues are non-negative.

Definition 1.2 (Gram matrices are PSD)**.** A Gram matrix for a polynomial is necessarily PSD.

The following important result, due originally to Choi, Lam, and Reznick [MDCR95], and reformulated in this convenient form by Powers and Wörmann [PW99], directly paves the way for the construction of an SOS decomposition algorithm.

Theorem 1.1. *Let $p(\vec{x}) \in \mathbb{R}[\vec{x}]$ be of degree $2k$ with $\vec{\zeta}$ as given above. Then, $p(\vec{x})$ is SOS iff there exists a symmetric PSD matrix $B \in \mathbb{R}^{u^2}$ s.t.*

$$p(\vec{x}) \ = \ \vec{\zeta}^{\,T} B \vec{\zeta}.$$

*Let B be such a matrix of rank t. Then, we can **construct** polynomials $q_1(\vec{x}), \ldots, q_t(\vec{x}) \in \mathbb{R}[\vec{x}]$ s.t.*

$$p(\vec{x}) \ = \ \sum_{i=1}^{t} q_i^2(\vec{x})$$

and B is a Gram matrix of $p(\vec{x})$ w.r.t. $q_1(\vec{x}), \ldots, q_t(\vec{x})$.

Proof. ($\Rightarrow$) If $p(\vec{x})$ is SOS, then we simply form B as (AA^T) as above.
($\Leftarrow$) Let $B \in \mathbb{R}^{u^2}$ be a symmetric PSD matrix of rank t s.t. $p(\vec{x}) = \vec{\zeta}^{\,T} B \vec{\zeta}$. Since B is real symmetric of rank t, there exists a real matrix V and a real diagonal matrix D,

$$D = \begin{pmatrix} d_1 & & & & & & \\ & d_2 & & & & & \\ & & \ddots & & & & \\ & & & d_t & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

s.t. $B = VDV^T$ and $d_i \neq 0$ ($\forall 1 \leq i \leq t$). Since B is PSD by hypothesis, it then follows that $d_i > 0$ ($\forall 1 \leq i \leq t$). So, we have

$$p(\vec{x}) \;=\; \vec{\zeta}^{\,T}(VDV^T)\vec{\zeta}. \qquad (*)$$

Suppose $V = (v_{i,j})$. Then, for $1 \leq i \leq t$ set

$$q_i(\vec{x}) \;:=\; \sqrt{d_i} \sum_{j=1}^{u} v_{j,i} \vec{x}^{\beta_i} \quad (\in \mathbb{R}[\vec{x}]).$$

By $(*)$, it then follows that $p(\vec{x}) = q_1^2(\vec{x}) + \ldots + q_t^2(\vec{x})$ as desired. □

It then follows that the task of finding an SOS decomposition of $p(\vec{x}) \in \mathbb{R}[\vec{x}]$ is equivalent to the task of finding a real symmetric, PSD matrix B s.t. $p(\vec{x}) = \vec{\zeta}^{\,T} B \vec{\zeta}$. We also see that given such a Gram matrix B, we can compute a sequence of $q_i(\vec{x})$ SOS cofactors of $p(\vec{x})$. Moreover, if we can show that no such Gram matrix B can exist, then we have shown that $p(\vec{x})$ is not SOS.

Definition 1.3. Given $p(\vec{x}) = \sum c_\alpha \vec{x}^\alpha$ as above and a real symmetric, PSD matrix $B \in \mathbb{R}^{u^2}$ s.t.

$$B = \begin{pmatrix} b_{1,1} & \ldots & b_{1,u} \\ \vdots & \ldots & \vdots \\ b_{u,1} & \ldots & b_{u,u} \end{pmatrix},$$

$p(\vec{x}) = \vec{\zeta}^{\,T} B \vec{\zeta}$ iff for each exponent vector $\alpha = \langle \alpha_1, \ldots, \alpha_n \rangle \in \Lambda_n(2k)$

$$\sum_{\beta_i + \beta_j = \alpha} b_{i,j} = c_\alpha,$$

where $\Lambda_n(k) = \{\beta_1, \ldots, \beta_u\}$ as before and $\beta_i + \beta_j$ is point-wise vector addition. This simple observation, following from term inspection, yields a rather quick check to see if a candidate matrix B is indeed Gram for $p(\vec{x})$. It also leads to the reduction of a search for such a B to a linear programming problem derived from these coefficient constraints.

The preceding observation can be somewhat confusing, but its importance cannot be overstated. We will illustrate it with a simple example.

Example 1.1. Let $p(x_1, x_2) = 4x_1^2 + x_2^2$. Then,

1. $dim(p) = 2$, $deg(p) = 2 = 2k$ with $k = 1$,
2. $\Lambda_2(2k) = \Lambda_2(2) = \{\langle 0,0\rangle, \langle 1,0\rangle, \langle 0,1\rangle, \langle 1,1\rangle, \langle 2,0\rangle, \langle 0,2\rangle\}$, which are the exponent vectors for the following power-products (in this order):
$$\{1, x_1, x_2, x_1x_2, x_1^2, x_2^2\}.$$
These are all of the power-products that could occur in an arbitrary 2-dimensional polynomial of degree 2.
3. $\Lambda_2(k) = \Lambda_2(1) = \{\langle 0,0\rangle, \langle 1,0\rangle, \langle 0,1\rangle\}$ which are the exponent vectors for the following power-products (in this order):
$$\{1, x_1, x_2\}.$$
These are all of the power-products that could occur in SOS co-factors of an arbitrary 2-dimensional polynomial of degree 2.
4. We set $u = |\Lambda_2(k)| = 3$ and fix an order upon $\Lambda_2(k)$ by setting:
$$\beta_1 = \langle 0,0\rangle,\ \beta_2 = \langle 1,0\rangle,\ \beta_3 = \langle 0,1\rangle.$$
5. We then set
$$\vec{\zeta} = \begin{pmatrix} \vec{x}^{\beta_1} \\ \vec{x}^{\beta_2} \\ \vec{x}^{\beta_3} \end{pmatrix} = \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix}$$
6. Now, $p(x_1, x_2)$ is *SOS* iff we can exhibit a $u \times u$ $(= 3 \times 3)$ real, symmetric, PSD matrix B s.t.
$$p(x_1, x_2) \ = \ \vec{\zeta}^{\,T} B \vec{\zeta}.$$
That is, we are looking for some real, symmetric, PSD
$$B = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{pmatrix} \quad s.t. \quad \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} \quad = \quad 4x_1^2 + x_2^2.$$
By multiplying through, we then see that:
$$\begin{pmatrix} 1 & x_1 & x_2 \end{pmatrix} \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} \quad =$$
$$\begin{pmatrix} b_{1,1} + b_{2,1}x_1 + b_{3,1}x_2 & b_{1,2} + b_{2,2}x_1 + b_{3,2}x_2 & b_{1,3} + b_{2,3}x_1 + b_{3,3}x_2 \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} \quad =$$
$$(b_{1,2} + b_{2,1})x_1 + (b_{1,3} + b_{3,1})x_2 + (b_{2,3} + b_{3,2})x_1x_2 + b_{2,2}x_1^2 + b_{3,3}x_2^2 + b_{1,1}.$$

So, by comparing coefficients,

$$\begin{aligned} b_{1,2} + b_{2,1} &= 0, \\ b_{1,3} + b_{3,1} &= 0, \\ b_{2,3} + b_{3,2} &= 0, \\ b_{1,1} &= 0, \\ b_{2,2} &= 4, \\ b_{3,3} &= 1. \end{aligned}$$

Moreover, we know that B must be symmetric, so we can strengthen our linear constraint system even further:

$$\begin{aligned} 2b_{1,2} &= 0, \\ 2b_{1,3} &= 0, \\ 2b_{2,3} &= 0, \\ b_{1,1} &= 0, \\ b_{2,2} &= 4, \\ b_{3,3} &= 1, \end{aligned}$$

and by simple linear manipulations and symmetry we can derive the following additional constraints:

$$b_{1,2} = b_{2,1} = 0,$$
$$b_{1,3} = b_{3,1} = 0,$$
$$b_{2,3} = b_{3,2} = 0.$$

Thus, we see that for $p(x_1, x_2) = 4x_1^2 + x_2^2$ to be SOS,

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

must be PSD. Of course, since it is trivial to see that $p(x_1, x_2)$ is indeed SOS, B in this case must be PSD. But, for the sake example, let us presume that we do not know B is PSD. Recall one of the characterisations for a real, square symmetric matrix to be PSD: B is PSD iff all of its eigenvalues are non-negative. Thus, we can simply compute B's eigenvalues in this case and be done with it. In 3-dimensions, the characteristic eigenvalue equation is:

$$\chi(B) \;=\; x^3 - tr(B)x^2 + Ax - det(B) = 0,$$

where A is the sum of (irreflexive) pairs of diagonal elements minus the products of each opposite pair of off diagonal elements:

$$A = b_{1,1}b_{2,2} + b_{1,1}b_{3,3} + b_{2,2}b_{3,3} - b_{1,2}b_{2,1} - b_{1,3}b_{3,1} - b_{2,3}b_{3,2} \ = \ 4.$$

Recalling that the trace of a square matrix is the sum of the elements on the main diagonal, we have:

$$tr(B) = b_{1,1} + b_{2,2} + b_{3,3} \ = \ 5.$$

Using cofactor expansion for the determinant:

$$det(B) = b_{1,1} \ * \ det\begin{pmatrix} b_{2,2} & b_{2,3} \\ b_{3,2} & b_{3,3} \end{pmatrix} \ - \ b_{2,2} \ * \ det\begin{pmatrix} b_{2,1} & b_{2,3} \\ b_{3,1} & b_{3,3} \end{pmatrix} \ + \ b_{3,3} \ * \ det\begin{pmatrix} b_{2,1} & b_{2,2} \\ b_{3,1} & b_{3,3} \end{pmatrix}$$

$$= (b_{1,1}b_{2,2}b_{3,3} + b_{1,2}b_{2,3}b_{3,1} + b_{1,3}b_{2,1}b_{3,2}) - (b_{3,1}b_{2,2}b_{1,3} + b_{3,2}b_{2,3}b_{1,1} + b_{3,3}b_{2,1}b_{1,2}) \ = \ 0$$

So,

$$\chi(B) = x^3 - 5x^2 + 4x$$

and thus by the univariate cubic formula, we solve for χ's roots and see the three eigenvalues[1] of B are $\{0, 1, 4\}$. B is thus verified to be PSD, and $p(x_1, x_2)$ is SOS.

There are a few points worth reflection concerning the previous example. First, consider the expansion of $\vec{\zeta}^T B \vec{\zeta}$ that allowed us to derive the linear equational constraints in step 6. Computing this expansion is tedious and expensive, involving three separate matrix multiplications. Instead, we can use **Observation 1.3** to derive this list of coefficient constraints in a more intelligent way. To make this process precise, let us introduce the following definitions.

Definition 1.4. Given $\Lambda(2k) = \{\alpha_1, \ldots, \alpha_w\}$, $\Lambda(k) = \{\beta_1, \ldots, \beta_u\}$, and $B = (b_{i,j})$ a real symmetric $k \times k$ matrix, we define

$$\mathfrak{C}(\alpha_m) = \{(i,j) \mid \beta_i + \beta_j = \alpha_m\}.$$

Then, we see the equational constraints we need to derive are always of the form

$$\left(\sum_{(i,j) \ \in \ \mathfrak{C}(\alpha_m)} b_{i,j} \right) \ = c_{\alpha_m}$$

where $p(\vec{x}) = \sum_{\alpha \in \Lambda(2k)} c_\alpha \vec{x}^\alpha$. Moreover, by commutativity of vector addition, whenever $(i,j) \in \mathfrak{C}(\alpha_m)$ it follows that $(j,i) \in \mathfrak{C}(\alpha_m)$, and so by symmetry of B, $b_{i,j} + b_{j,i}$ can be replaced by $2b_{i,j}$, eliminating $b_{j,i}$ from the constraint system generated.

We can reexamine step 6 from our previous example using the method above.

[1]Indeed, it is good to know that for a *diagonal matrix*, the eigenvalues are always precisely the entries along the diagonal. B in this case is diagonal, so one could save some effort by reading off its eigenvalues directly. But, B in this example is highly atypical and much simpler than most Gram matrices we will derive in practice.

Example 1.2 (Step 6 of Example 2.9 revisited)**.** We are looking for some real, symmetric, PSD matrix B s.t.

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{pmatrix} \quad s.t. \quad \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} \quad = \quad 4x_1^2 + x_2^2.$$

Recalling that in this case:

$$\Lambda_2(2k) = \Lambda_2(2) = \{\langle 0,0\rangle, \langle 1,0\rangle, \langle 0,1\rangle, \langle 1,1\rangle, \langle 2,0\rangle, \langle 0,2\rangle\}$$

which are the exponent vectors for the following power-products:

$$\{1, x_1, x_2, x_1x_2, x_1^2, x_2^2\},$$

and

$$\Lambda_2(k) = \Lambda_2(1) = \{\langle 0,0\rangle, \langle 1,0\rangle, \langle 0,1\rangle\}$$

which are the exponent vectors for the following power-products:

$$\{1, x_1, x_2\}.$$

We now proceed to compute $\mathfrak{C}(\alpha)$ for each $\alpha \in \Lambda_2(2k)$:

$$\begin{aligned}
\mathfrak{C}(\langle 0,0\rangle) &= \{(i,j) \mid \beta_i + \beta_j = \langle 0,0\rangle\} &= \{(1,1)\}, \\
\mathfrak{C}(\langle 1,0\rangle) &= \{(i,j) \mid \beta_i + \beta_j = \langle 1,0\rangle\} &= \{(1,2),(2,1)\}, \\
\mathfrak{C}(\langle 0,1\rangle) &= \{(i,j) \mid \beta_i + \beta_j = \langle 0,1\rangle\} &= \{(1,3),(3,1)\}, \\
\mathfrak{C}(\langle 1,1\rangle) &= \{(i,j) \mid \beta_i + \beta_j = \langle 0,1\rangle\} &= \{(2,3),(3,2)\}, \\
\mathfrak{C}(\langle 2,0\rangle) &= \{(i,j) \mid \beta_i + \beta_j = \langle 2,0\rangle\} &= \{(2,2)\}, \\
\mathfrak{C}(\langle 0,2\rangle) &= \{(i,j) \mid \beta_i + \beta_j = \langle 0,2\rangle\} &= \{(3,3)\}.
\end{aligned}$$

This, in conjunction with the coefficients c_α for $p(x_1, x_2) = 4x_1^2 + x_2^2$:

$$\begin{array}{lll}
c_{\langle 0,0\rangle} = 0 & c_{\langle 1,0\rangle} = 0 & c_{\langle 0,1\rangle} = 0 \\
c_{\langle 1,1\rangle} = 0 & c_{\langle 2,0\rangle} = 4 & c_{\langle 0,2\rangle} = 1
\end{array}$$

yields, modulo symmetry of B, the following linear constraints:

$$\begin{array}{rrr}
b_{1,1} = 0 & 2b_{1,2} = 0 & 2b_{1,3} = 0 \\
2b_{3,2} = 0 & b_{2,2} = 4 & b_{3,3} = 1
\end{array}$$

which is exactly what was derived by the more costly product expansion in the previous example.

From this example, we can extrapolate a general algorithm, due to Powers and Wörmann, for computing SOS decompositions when they exist. An important caveat, though, is that this technique *presupposes* the existence of an algorithm for testing the emptiness of a semialgebraic set defined in the following form:

$$S = \{(r_1, \ldots, r_n) \mid p_1(r_1, \ldots, r_n) \geq 0 \ \wedge \ \ldots \wedge \ p_m(r_1, \ldots, r_n) \geq 0\},$$

where $p_i \in \mathbb{R}[x_1, \ldots, x_n]$. That is, the feasibility of the Powers-Wörmann technique reduces to the feasibility of testing the emptiness of an intersection of *closed* subsets of $\mathbb{R}^n$ for which a collection of real polynomials are PSD. Nevertheless, this algorithm contains many important ideas. In particular, it provides a systematic way to *construct* SOS decompositions when they exist, allows one to exactly parameterise the set of all Gram matrices for a real polynomial using only intersections of closed sets explicitly characterised semialgebraically, and makes the crucial observation that the search for SOS decompositions reduces to a search for a PSD matrix modulo linear constraints. As we will see in the next note, this crucial observation is what paves the way for the use of powerful, feasible convex optimisation techniques for the task of Gram matrix derivation.

Definition 1.5 (Powers-Wörmann Real SOS Decomposition Algorithm)**.** Given $p(x_1, \ldots, x_n) = \sum c_\alpha \vec{x}^\alpha \in \mathbb{R}[x_1, \ldots, x_n]$ s.t. $deg(p) = 2k$:

1. Fix an order on the set of all possible generic degree $\leq 2k$ exponent vectors $\Lambda_n(2k)$ s.t. $\Lambda_n(2k) = \{\alpha_1, \ldots, \alpha_{\binom{n+2k}{n}}\}$.
2. Fix an order on the set of all possible SOS cofactor exponent vectors $\Lambda_n(k)$ s.t. $\Lambda_n(k) = \{\beta_1, \ldots, \beta_u\}$.
3. Let $\vec{\zeta}$ be the column vector of all possible SOS cofactor monomials, and $B = (b_{i,j})$ be a real, symmetric $u \times u$ matrix with variable entries:

$$\vec{\zeta} = \begin{pmatrix} \vec{x}^{\beta_1} \\ \vec{x}^{\beta_2} \\ \vdots \\ \vec{x}^{\beta_u} \end{pmatrix} \quad B = \begin{pmatrix} b_{1,1} & \ldots & b_{1,u} \\ \vdots & \ldots & \vdots \\ b_{u,1} & \ldots & b_{u,u} \end{pmatrix} \quad b_{i,j} = b_{j,i}.$$

4. Derive $\vec{\mathfrak{C}}(\Lambda_n(2k))$, the Gram coefficient contribution table of B w.r.t. $\Lambda_n(2k)$ induced by the assumption $\vec{\zeta}^T B \vec{\zeta} = p(\vec{x})$:

$$\begin{aligned} \mathfrak{C}(\alpha_1) &= \{(i,j) \mid \beta_i + \beta_j = \alpha_1\}, \\ \vdots &= \vdots \\ \mathfrak{C}(\alpha_{\binom{n+2k}{n}}) &= \{(i,j) \mid \beta_i + \beta_j = \alpha_{\binom{n+2k}{n}}\}. \end{aligned}$$

5. Derive $\mathcal{L}(\vec{\mathfrak{C}}(p), c)$, the linear constraint system induced by $\vec{\mathfrak{C}}(p)$, $c : \Lambda_n(2k) \to \mathbb{R}$ (the coefficients of p), and **Observation 1.3**:

$$\begin{array}{rcl} \sum_{(i,j) \,\in\, \mathfrak{C}(\alpha_1)} b_{i,j} & = & c_{\alpha_1}, \\ \vdots & = & \vdots \\ \sum_{(i,j) \,\in\, \mathfrak{C}(\alpha_{\binom{n+2k}{n}})} b_{i,j} & = & c_{\alpha_{\binom{n+2k}{n}}} \end{array}$$

(modulo the symmetry of B - e.g., replace subterms $b_{i,j} + b_{j,i}$ by $2b_{i,j}$).

6. Note that each B variable $b_{i,j}$ only exists in one linear equation in $\mathcal{L}(\vec{\mathfrak{C}}(p), c)$. Thus, to solve $\mathcal{L}(\vec{\mathfrak{C}}(p), c)$ we can work pointwise upon each constraint $\mathfrak{l}_i$ independently by replacing every variable except one in $\mathfrak{l}_i$ with a fresh parameter λ_j and solving for the remaining variable. Then, all candidate solution matrices B can be given as a sum of parameter-scaled matrices:

$$B = B_0 + \lambda_1 B_1 + \ldots + \lambda_l B_l$$

where each B_m is a real, symmetric $u \times u$ matrix s.t. the nonzero entries of B_0 are precisely the entries of B which contain no parameters, and the nonzero entries of B_m $(m > 0)$ are precisely the values $\frac{b_{i,j}}{\lambda_m}$ when $b_{i,j}$ contains the parameter λ_m.

7. Given such a parameterised B, the task at hand reduces to that of finding values for $\lambda_1, \ldots, \lambda_j$ that make $B = B_0 + \lambda_1 B_1 + \ldots \lambda_l B_l$ PSD. We again utilise the eigenvalue characterisation of PSD for symmetric matrices. Let $\chi(B)$ be the characteristic eigenvalue equation for B:

$$\chi(B) = det(xI - B)$$

where I is the $u \times u$ identity matrix and the determinant is taken over $\mathbb{R}(x)$, the field of rational functions in x. $\chi(B)$ is thus some polynomial $F(x) = x^s + a_{s-1}x^{s-1} + \ldots + a_0 \in \mathbb{R}[x, \lambda_1, \ldots, \lambda_l]$. It suffices to derive conditions upon λ_i s.t. the roots of $F(x)$ are non-negative.

8. By Descartes' Rule of Signs, $F(x)$ has only non-negative roots iff

$$(-1)^{i+u} a_i \geq 0 \ \ (\forall 0 \leq i \leq k-1).$$

As each coefficient a_i lies in the algebra $\mathbb{R}[\lambda_1, \ldots, \lambda_l]$, each a_i can itself be evaluated as a polynomial, e.g., $a_i(\lambda_1, \ldots, \lambda_l) : \mathbb{R}^l \to \mathbb{R}$. Thus, B is PSD iff the semialgebraic set

$$S = \{(\lambda_1, \ldots, \lambda_l) \in \mathbb{R}^l \mid (-1)^{i+u} a_i(\lambda_1, \ldots, \lambda_l) \geq 0 \mid 0 \leq i \leq u\}$$

is nonempty. Moreover, each point $(\lambda_1, \ldots, \lambda_l) \in S$ corresponds to a Gram matrix for $p(\vec{x})$ w.r.t. $\vec{\zeta}$.

9. At this point, one must utilise a specialised algorithm for testing the emptiness of S, e.g., cylindrical algebraic decomposition [ACM84] or virtual substitution [Wei97]. Of course, such algorithms are in general feasible only for semialgebraic constraints in very small dimensions.

10. If S is found to be nonempty and contains a point $(\lambda_1, \ldots, \lambda_l) \in \mathbb{R}^l$, then

$$B = B_0 + \lambda_1 B_1 + \ldots + \lambda_l B_l$$

can be used to construct an explicit SOS cofactor expansion of $p(\vec{x})$ using a combination of square-root-free Cholesky decomposition and the method given in the final step of the proof of **Theorem 1.1**.

   (a) Set $t = rank(B)$.

   (b) Since B is real symmetric of rank t, by square-root-free Cholesky decomposition there exists a real matrix $V = (v_{i,j})$ and a real diagonal matrix $D = diag(d_1, \ldots, d_t, 0 \ldots, 0)$ s.t. $B = VDV^T$ and $d_i \neq 0$ $(\forall 1 \leq i \leq t)$. Use square-root-free Cholesky decomposition to compute V and D.

   (c) Since B is PSD, it then follows that $d_i > 0$ $(\forall 1 \leq i \leq t)$. So, we have

$$p(\vec{x}) \;=\; \vec{\zeta}^{\,T}(VDV^T)\vec{\zeta}. \qquad (*)$$

   (d) For $1 \leq i \leq t$ set

$$q_i(\vec{x}) \;:=\; \sqrt{d_i}\sum_{j=1}^{u} v_{j,i}\vec{x}^{\beta_i} \quad (\in \mathbb{R}[\vec{x}]).$$

   (e) By $(*)$, it then follows that $p(\vec{x}) = q_1^2(\vec{x}) + \ldots + q_t^2(\vec{x})$ as desired.

References

[ACM84] Dennis S. Arnon, George E. Collins, and Scott McCallum. Cylindrical algebraic decomposition. I. the basic algorithm. *SIAM Journal on Computing*, 13(4):865–877, November 1984.

[MDCR95] T. Y. Lam M. D. Choi and B. Reznick. Sums of squares of real polynomials. In *Proc. Sympos. Pure Math.*, pages 103–126, 1995.

[Par03] Pablo A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Math. Program.*, 96(2):293–320, 2003.

[PW99] Victoria Powers and Thorsten Wörmann. An algorithm for sums of squares of real polynomials. *Journal of Pure and Applied Algebra*, 127:99–104, 1999.

[Wei97] Volker Weispfenning. Quantifier elimination for real algebra - the quadratic case and beyond. *Appl. Algebra Eng. Commun. Comput.*, 8(2):85–101, 1997.