Annales Univ. Sci. Budapest., Sect. Comp. 33 (2010) 273-284

ON SUM OF SQUARES DECOMPOSITION FOR A BIQUADRATIC MATRIX FUNCTION

L. László (Budapest, Hungary)

Dedicated to Professor Ferenc Schipp on his 70th birthday and to Professor Péter Simon on his 60th birthday

Abstract. We investigate a nonnegative biquadratic form generated by some structured matrices and prove that it is a sum of squares of polynomials (SOS). On the other hand, if the underlying matrices have additional nonzero entries, the form at issue is - although still nonnegative - no more SOS.

1. Introduction

The Böttcher-Wenzel inequality (see [2], [5], [9])

$$||XY - YX||^2 \leq 2\,||X||^2||Y||^2$$

gives in its original form an upper bound for the commutator of two real square matrices X, Y of the same size in the Frobenius norm. This raises the question: is the difference $2||X||^2||Y||^2 - ||XY - YX||^2$ - as a quartic polynomial - a sum of squares of some quadratics?

We will be concerned, however, with the nonnegativeness of

$$2\;||X||^2||Y||^2 - 2\;\text{trace}^2(X^TY) - ||XY - YX||^2, \tag{1}$$

a strengthened version of the above theorem (see [2], Theorem 3.1). The reason is that subtracting the trace-term guarantees that (1) only depends on the differences $x_{i,j}\,y_{k,l} - y_{i,j}\,x_{k,l}$ – a useful property. It turns out that the answer for the question is 'yes' for some matrices of simple structure, and 'no' for matrices with a little more difficult structure.

Introduce now the relevant matrix classes. A real square matrix of order n will be called RC, if nonzero elements occur only in row 1 and column n, while it is called RCD, if nonzero elements occur in row 1, column n, and the main diagonal.

For instance, these patterns are in case of fourth order matrices:

$$\mathrm{RC}: \begin{pmatrix} * & * & * & * \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & * \end{pmatrix} \qquad \mathrm{RCD}: \begin{pmatrix} * & * & * & * \\ 0 & * & 0 & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{pmatrix},$$

where $*$ stands for an arbitrary (real) number.

Our main result is that the nonnegative form (1) is a sum of squares of polynomials (in short: SOS) for RC matrices X, Y, whereas it is not SOS for general RCD matrices. First, however, we illustrate these notions by the celebrated Motzkin polynomial, see Reznick's survey [8] for further details and references.

Example. Denote by

$$M(x,y,z) = x^4y^2 + x^2y^4 + z^6 - 3x^2y^2z^2$$

Motzkin's (first) polynomial. M is nonnegative for all real x, y, z by the A-G inequality applied for $\{x^4y^2,\ x^2y^4,\ z^6\}$. The most important property of this ternary sextic is that it is not a SOS. Nevertheless, M is a sum of squares of some rational functions, for this holds for any nonnegative polynomial by Artin's answer to Hilbert's 17-th problem. Here are two known representations:

$$(x^2+y^2)^2M(x,y,z) = (x^2-y^2)^2z^6 + x^2y^2(x^2+y^2+z^2)(x^2+y^2-2z^2)^2,$$

and

$$(x^2+y^2+z^2)M(x,y,z) = (x^2yz - yz^3)^2 + (xy^2z - xz^3)^2 + \\ + (x^2y^2 - z^4)^2 + \frac{1}{4}(xy^3 - x^3y)^2 + \frac{3}{4}(xy^3 + x^3y - 2z^2)^2.$$

Substitute $y = 1$ into Motzkin's form to get the dehomogenized polynomial

$$M(x,1,z) = x^4 + x^2 + z^6 - 3x^2z^2.$$

Parrilo [7] computed the minimum constant c, for which $M(x,1,z) + c$ is SOS, and found that

$$M(x,1,z) + \frac{729}{4096} = \left(z^3 - \frac{9}{8}z\right)^2 + \left(x^2 - \frac{3}{2}z^2 + \frac{27}{64}\right)^2 + \frac{5}{32}x^2$$

– a strange identity, knowing that $M(x,1,z) \geq 0$ for all real x,z and $M(1,1,1)=0$.

As for the dehomogenization in z, $M(x,y,1)+c$ is not SOS for any real c - another interesting property.

Note finally, that M is the special case for $n=3$ of the more general form

$$\left(\sum_{i=1}^{n-1} x_i^2 - n x_n^2\right) \prod_{i=1}^{n-1} x_i^2 + x_n^{2n},$$

which is nonnegative and non-SOS, as well.

2. SOS decomposition for RC matrices

Let X,Y be real n-th order RC matrices with $m=2n-1$ possible nonzero elements:

$$X = \begin{pmatrix} x_1 & \dots & x_{n-1} & x_n \\ 0 & \dots & 0 & x_{n+1} \\ \vdots & \ddots & \vdots & \\ 0 & \dots & 0 & x_m \end{pmatrix}, \qquad Y = \begin{pmatrix} y_1 & \dots & y_{n-1} & y_n \\ 0 & \dots & 0 & y_{n+1} \\ \vdots & \ddots & \vdots & \\ 0 & \dots & 0 & y_m \end{pmatrix},$$

and define an $m-$th order matrix Z by help of the vectors $x=(x_i)_1^m$ and $y=(y_i)_1^m$ as

$$Z = xy^T - yx^T = (z_{i,j})_{i,j=1}^m, \quad z_{i,j} = x_i y_j - y_i x_j. \tag{2}$$

Observe that using one subscript for the entries of X,Y considerably simplifies the presentation of the following theorem.

Theorem 1.

$$\begin{aligned} &||Z||^2 - \Big(\sum_{i=1}^{n} z_{i,i+n-1}\Big)^2 - \sum_{i=2}^{n-1} z_{1,i}^2 - \sum_{i=n+1}^{m-1} z_{i,m}^2 = \\ &= \sum_{i=1}^{n-1}\sum_{j=n+1}^{m} z_{i,j}^2 + \sum_{i=2}^{n-2}\sum_{j=i+1}^{n-1} z_{i,j}^2 + \sum_{i=n+1}^{2n-3}\sum_{j=i+1}^{2n-2} z_{i,j}^2 + \\ &+ \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \big(z_{i,j} - z_{i+n-1,j+n-1}\big)^2 + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \big(z_{i,j+n-1} - z_{j,i+n-1}\big)^2. \end{aligned} \tag{3}$$

Proof. There are $\binom{m}{4}$ immediately checked basic relations

$$z_{i,j}z_{k,l}+z_{i,l}z_{j,k}-z_{i,k}z_{j,l}=0,\quad 1\le i<j<k<l\le m. \tag{4}$$

Applying these with $k=i+n-1,\ l=j+n-1$ for the first subtrahend $\left(\sum_{i=1}^{n} z_{i,i+n-1}\right)^2$ yields the double products involved in the last line.

It remains to prove the equality of the (pure) squares. To this, introduce the notations

$$S_0=\sum_{i=1}^{n} z_{i,i+n-1}^2,\quad S_1=\sum_{i=2}^{n-1} z_{1,i}^2,\quad S_2=\sum_{i=n+1}^{m-1} z_{i,m}^2,$$

$$S_3=\sum_{i=1}^{n-1}\sum_{j=n+1}^{m} z_{i,j}^2=\sum_{i=n+1}^{m}\sum_{j=1}^{n-1} z_{i,j}^2,\quad S_4=\sum_{i=2}^{n-2}\sum_{j=i+1}^{n-1} z_{i,j}^2,$$

$$S_5=\sum_{i=n+1}^{2n-3}\sum_{j=i+1}^{2n-2} z_{i,j}^2,\quad S_6=\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} z_{i,j}^2=\sum_{i=2}^{n}\sum_{j=1}^{i-1} z_{i,j}^2,$$

$$S_7=\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} z_{i+n-1,j+n-1}^2=\sum_{i=n+1}^{m}\sum_{j=n}^{i-1} z_{i,j}^2,$$

$$S_8=\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} z_{i,j+n-1}^2=\sum_{i=1}^{n-1}\sum_{j=i+n}^{m} z_{i,j}^2,$$

$$S_9=\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} z_{j,i+n-1}^2=\sum_{i=2}^{n}\sum_{j=n}^{n+i-2} z_{i,j}^2.$$

The equivalent formulae for S_3 and S_6 to S_9 (as consequences of the identity $z_{i,j}^2=z_{j,i}^2$) enable us to put each term in a *distinct* position, illustrated e.g. for $n=5,\ m=2n-1=9$ by the tableau

$$\begin{pmatrix}
- & 1 & 1 & 1 & 0 & 8 & 8 & 8 & 8\\
6 & - & 4 & 4 & 9 & 0 & 8 & 8 & 8\\
6 & 6 & - & 4 & 9 & 9 & 0 & 8 & 8\\
6 & 6 & 6 & - & 9 & 9 & 9 & 0 & 8\\
6 & 6 & 6 & 6 & - & 9 & 9 & 9 & 0\\
3 & 3 & 3 & 3 & 7 & - & 5 & 5 & 2\\
3 & 3 & 3 & 3 & 7 & 7 & - & 5 & 2\\
3 & 3 & 3 & 3 & 7 & 7 & 7 & - & 2\\
3 & 3 & 3 & 3 & 7 & 7 & 7 & 7 & -
\end{pmatrix},$$

where an integer k in the position (i,j) shows that $z_{i,j}^2$ is contained in S_k, while the '$-$' characters on the main diagonal stand for the zeros (due to $z_{i,i}=0$). This auxiliary matrix helps us to see that

$$||Z||^2 = \sum_{i=0}^{9} S_i.$$

Since (3) is quadratic in the $z_{i,j}$-s, and both the squares and the double products of the equality coincide, the theorem is true.

Corollary. This yields the SOS representation required. Indeed, (1) and (2) are identical. In particular,

$$||Z||^2 = 2\,||X||^2||Y||^2 - 2\,\mathrm{trace}^2(X^TY), \tag{5}$$

and

$$\left(\sum_{i=1}^{n} z_{i,i+n-1}\right)^2 + \sum_{i=2}^{n-1} z_{1,i}^2 + \sum_{i=n+1}^{m} z_{i,m}^2 = ||XY-YX||^2,$$

where the first is Lagrange's identity, the second is straightforward.

3. SOS decomposition impossible for RCD matrices

It suffices to prove this for third order matrices. (For matrices of order $n>3$ choose an index i, $1<i<n$, and annihilate the rows and columns with indices, different from $\{1,i,n\}$.) To be compatible with our former notation, we write

$$X = \begin{pmatrix} x_1 & x_2 & x_3 \\ 0 & x_6 & x_4 \\ 0 & 0 & x_5 \end{pmatrix}, \qquad Y = \begin{pmatrix} y_1 & y_2 & y_3 \\ 0 & y_6 & y_4 \\ 0 & 0 & y_5 \end{pmatrix}.$$

It turns out that the presence of x_6 and y_6 causes the impossibility of an SOS representation for (1). Since

$$XY - YX = \begin{pmatrix} 0 & z_{1,2}+z_{2,6} & z_{1,3}+z_{2,4}+z_{3,5} \\ 0 & 0 & z_{4,5}-z_{4,6} \\ 0 & 0 & 0 \end{pmatrix},$$

the nonnegative form (1) to be discussed assumes

$$(6)\qquad 2\sum_{1\le i<j\le 6} z_{i,j}^2-(z_{1,2}+z_{2,6})^2-(z_{4,5}-z_{4,6})^2-(z_{1,3}+z_{2,4}+z_{3,5})^2.$$

First we briefly recall the Gram matrix method [4] widely used in the literature. Let $g\in\mathbb{R}[u_1,\dots u_k]$ be a nonnegative form (form stands for homogeneous polynomial) of degree $2d$, and let $z=(z_i)_1^K$ be the appropriately ordered vector of all monomials of the half degree d. If there exists a positive semidefinite matrix G_0 of order K such that $g=z^TG_0z$, then g is SOS.

However, to get a condition, not only sufficient but also necessary for g to be SOS, in case of $d>1$ we have to collect *all* possible quadratic relations holding for the z_i-s (like e.g. $z_1z_2-z_3^2=0$ for $z_1=u_1^2, z_2=u_2^2, z_3=u_1u_2, d=$ $=2$). These relations can be written as $z^TG_iz=0$ $(i=1,\dots,L)$, where G_i are symmetric matrices of the same size K. Then, g is SOS if and only if there exist real numbers $(\alpha_i)_1^L$ such that $G_0+\sum\limits_{i=1}^{L}\alpha_iG_i$ is positive semidefinite.

For further reading see the Lecture Notes by Parrilo [6].

We make use of this method via formulating the following condition, sufficient for *not* being SOS. The above notations will be used.

Lemma 1. *Let $g=z^TG_0z$ be a real nonnegative form with spectral decomposition $G_0=V\Lambda V^T$, V orthogonal, Λ diagonal. Assume that $r>0$ eigenvalues of G_0 (the first r diagonal entries of Λ) are negative and denote by W the $K\times r$ matrix of the associated eigenvectors (the first r columns of V). Let $(G_i)_1^L$ be a full set of the matrices of quadratic relations.*

If there exists a diagonal positive definite matrix $D=D_0^2$ of order r such that for the reduced matrices $R_i=W^TG_iW$ it holds that

$$\operatorname{trace}(R_iD)=0,\ 1\le i\le L,$$

then g is not SOS.

Proof. Since $g=z^TG_0z$, and $z^TG_iz=0$, $1\le i\le L$, all possible representations for g are given by $g=z^TGz$, where

$$G=G_0+\sum_{i=1}^{L}\alpha_iG_i.$$

Pre-multiplying with W^T and post-multiplying with W gives an equation for the reduced matrices:

$$R=R_0+\sum_{i=1}^{L}\alpha_iR_i.$$

Here R_0 is an r-th order diagonal matrix with negative eigenvalues. Assume by contradiction that, there exists a sequence of α_i's such that G is positive semidefinite. Then, for these weights, the reduced matrix $R = W^T G W$ is positive semidefinite, and the matrix $\sum_{i=1}^{L} \alpha_i R_i = R + (-R_0)$ – as the sum of a positive semidefinite and a positive definite matrix – is positive definite as well, and so is the matrix

$$D_0^T \left(\sum_{i=1}^{L} \alpha_i R_i \right) D_0 = \sum_{i=1}^{L} \alpha_i D_0^T R_i D_0.$$

On the other hand, its trace equals to

$$\sum_{i=1}^{L} \alpha_i \operatorname{trace}(D_0^T R_i D_0) = \sum_{i=1}^{L} \alpha_i \operatorname{trace}(R_i D) = 0,$$

which contradicts the positive definiteness.

Remark 1. A quite similar sufficient condition for not being SOS is described in [7], however with the dual problem involved.

The idea of Lemma 1 is closely related to that of [3], however, we need here only a suitable *sufficient* condition for a nonnegative form to be not a SOS, while G. Chesi sets a full characterization of the problem. Further he writes: "The results proposed in this paper represent a first step in the characterization of the existing gap between positive polynomials and SOS of polynomials, about which only few isolated examples were known until now."

Our second theorem gives another example: the RCD matrices.

Theorem 2. *The biquadratic form (6), although nonnegative for any real* $(x_i)_1^6$, $(y_i)_1^6$, *is not a sum of squares of any quadratics!*

Proof. It can easily be shown that if (6) is SOS, then it is a sum of squares depending only on the $z_{i,j}$-s. The number of the $z_{i,j}$-s with $i < j$ is now $\binom{6}{2} = 15$. The vector

$$z = [z_{1,2}, z_{2,3}, z_{3,4}, z_{4,5}, z_{5,6}, z_{1,3}, z_{2,4}, z_{3,5}, z_{4,6}, z_{1,4}, z_{2,5}, z_{3,6}, z_{1,5}, z_{2,6}, z_{1,6}]^T$$

of these differences uniquely determines the initial matrix G_{in} by the equality $z^T G_{in} z = f(x, y)$. (For the matrix G_{in} see Remark 2.)

In view of $\binom{6}{2} = \binom{6}{4}$, the size K of the matrices coincides with the number L of the basic relations $z^T G_i z = 0$. (There are no more relations for the $z_{i,j}$'s than (4), as can easily be shown.) These matrices have 6 nonzero elements.

We display the following important ones by giving only their nonzero upper diagonal elements:

$$G_{1234}(1,3)=1,\quad G_{1234}(2,10)=1,\quad G_{1234}(6,7)=-1,$$

$$G_{2345}(2,4)=1,\quad G_{2345}(3,11)=1,\quad G_{2345}(7,8)=-1,$$

$$G_{2346}(2,9)=1,\quad G_{2346}(3,14)=1,\quad G_{2346}(7,12)=-1.$$

Here G_{ijkl} stands for the 15-th order symmetric matrix, for which

$$z^T G_{ijkl}\, z = z_{i,j}z_{k,l} + z_{i,l}z_{j,k} - z_{i,k}z_{j,l} \;(1 \le i < j < k < l \le 6).$$

Note that the matrix G_{in} cannot play the role of the basic matrix G_0 in Lemma 1 (see also Remark 2 below), hence we try choosing

$$G_0 = G_{in} + p(G_{1234} + G_{2345}),$$

where $p \neq 0$ is a real parameter to be determined later. With $q = p - 1$ the matrix G_{in} has the form

$$\begin{pmatrix}
1 & 0 & -p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0\\
0 & 2 & 0 & -p & 0 & 0 & 0 & 0 & 0 & -p & 0 & 0 & 0 & 0 & 0\\
-p & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -p & 0 & 0 & 0 & 0\\
0 & -p & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 1 & q & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & q & 1 & q & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & -1 & q & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & -p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & -p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0\\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2
\end{pmatrix}$$

while its characteristic polynomial in the variable x equals

$$(x-2)^7\Big(x^2 - x - 2(p-1)^2\Big)\Big(x(x-2)^2 - p^2(2x-1)\Big)^2.$$

First we solve the quadratics $\varphi_2(x) = x^2 - x - 2(p-1)^2 = 0$ for the root

$$\lambda = \lambda(p) = \frac{1 - \sqrt{1 + 8(p-1)^2}}{2} = \frac{-\,4(p-1)^2}{1 + \sqrt{1 + 8(p-1)^2}},$$

negative for $p \neq 1$.

As for the roots of the cubic $\varphi_3(x) = x(x-2)^2 - p^2(2x-1)$, the equality

$$x(x-2)^2 = p^2(2x-1)$$

implies that $\mathrm{sign}(x) = \mathrm{sign}(2x-1)$, i.e. $x \notin (0, 1/2)$. On the other hand, from the values $\varphi_3(0) = p^2 > 0$, $\varphi_3(2) = -3p^2 < 0$, and the fact that the main coefficient is positive we conclude that φ_3 has exactly one negative root $\mu = \mu(p)$ for all $p \neq 0$. To sum up, the negative eigenvalues of $G_0 = G_0(p)$ are λ and the double eigenvalue μ.

Now we come to the discussion of the eigenvectors. The following matrix W contains as columns the eigenvectors associated with these eigenvalues. (The first column corresponds to λ, while the second and third vectors constitute a basis for the 2-dimensional subspace associated with μ) :

$$W = \begin{pmatrix} 0 & 0 & p\ (1-\mu) \\ 0 & \mu\ (2-\mu) & 0 \\ 0 & 0 & \mu\ (\mu-2) \\ 0 & p\ (\mu-1) & 0 \\ 0 & 0 & 0 \\ p-1 & 0 & 0 \\ \lambda & 0 & 0 \\ p-1 & 0 & 0 \\ 0 & p & 0 \\ 0 & p\mu & 0 \\ 0 & 0 & -p\mu \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & p \\ 0 & 0 & 0 \end{pmatrix}$$

To check the first eigenvector, it suffices to calculate

$$\begin{pmatrix} 1 & p-1 & -1 \\ p-1 & 1 & p-1 \\ -1 & p-1 & 1 \end{pmatrix} \begin{pmatrix} p-1 \\ \lambda \\ p-1 \end{pmatrix} = \begin{pmatrix} \lambda(p-1) \\ \lambda + 2(p-1)^2 \\ \lambda(p-1) \end{pmatrix}$$

and observe that $\lambda + 2(p-1)^2 = \lambda^2$ by the simple fact that $\varphi_2(\lambda) = 0$. The case of the double eigenvalue μ is analogous, but then the equality $\varphi_3(\mu) = 0$ is used two times.

We are now ready to determine the reduced Gramians. There are three of them with nonzero diagonals, namely those corresponding to the above defined $G_{1234}, G_{2345}, G_{2346}$. We display only their diagonals:

$$\mathrm{diag}(R_{1234}) = 2\,\Big(\lambda(1-p),\ p\,\mu^2(2-\mu),\ p\,\mu(\mu-1)(2-\mu)\Big),$$

$$\mathrm{diag}(R_{2345}) = 2\,\Big(\lambda(1-p),\ p\,\mu(\mu-1)(2-\mu),\ p\,\mu^2(2-\mu)\Big),$$

$$\mathrm{diag}(R_{2346}) = 2\,\Big(0,\ p\,\mu(2-\mu),\ -p\,\mu(2-\mu)\Big).$$

Now we need a positive vector, which is orthogonal to all these diagonals as vectors. The form of the second and third coordinates suggests to choose $(\delta,1,1)$. Then for the corresponding diagonal matrix

$$D = \begin{pmatrix} \delta & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

it holds that $\mathrm{trace}(R_{2346}D) = 0$, and $\mathrm{trace}(R_{1234}D) = \mathrm{trace}(R_{2345}D)$ with the common value

$$t\lambda(1-p) - p\mu(2-\mu)(2\mu-1).$$

This value equals zero, if and only if

$$t = \frac{p\,\mu(\mu-2)(2\mu-1)}{\lambda(1-p)}.$$

Since both λ and μ are negative, the requirement $t>0$ is equivalent to $p(1-$ $-p)(\mu-2)(2\mu-1)>0$. However, $\mu<0$ involves $(\mu-2)(2\mu-1)>0$, giving $p(1-p)>0$, i.e. $p\in(0,1)$. Choosing p from this interval, t will be positive, D positive definite, and

$$\mathrm{trace}(R_{ijkl}D) = 0, \quad 1 \le i < j < k < l \le 6.$$

This in connection with Lemma 1 proves the Theorem.

Remark 2. The matrix G_{in} can be obtained from $G_0 = G_0(p)$ by substituting $p=0$. It has only one negative eigenvalue, $\lambda=-1$. However, the associated eigenvector (say, v) is not enough to satisfy the requirements of Lemma 1, since then all the reduced matrices – now scalars! – should vanish, while $R_{1234} = v^T G_{1234} v$ (and also $R_{2345} = v^T G_{2345} v$) are nonzero. This is why we had to search for another matrix.

4. Generalization, problem setting

Denote by $\mathcal{B}_{m,n}$ the set of nonnegative biquadratic forms, i.e. nonnegative homogeneous polynomials $f \in \mathbb{R}[x_1, \dots x_m, y_1, \dots y_n]$ with the property that $f(x,y)$ is a quadratic form in $y = (y_i)_1^n$ for fixed $x = (x_i)_1^m$ and vice versa. The polynomials, discussed in this article, obviously belong to $\mathcal{B}_{m,m}$, in fact, to the smaller set

$$\mathcal{Z}_m = \{f \in \mathcal{B}_{m,m} : f \text{ is a quadratic form of the } z'_{i,j}s\}.$$

Alternatively, if $f \in \mathcal{B}_{m,m}$, then $f(x,y) = z^T G z$ for a symmetric G. To be precise, we define the order of coordinates of the vector z as

$$z = \left((z_{i,i+k})_{i=1}^{m-k} \right)_{i=1}^{m-1}.$$

Notice that the above "inner" and the following "outer" definition of $\mathcal{Z}_m$ are equivalent:

$$\mathcal{Z}_m^* = \{f \in \mathcal{B}_{m,m} : f(x,y) = f(y,x),\ f(x,x) = 0,\ x, y \in \mathbb{R}^m\}.$$

($\mathcal{Z}_m \subset \mathcal{Z}_m^*$ is trivial, $\mathcal{Z}_m^* \subset \mathcal{Z}_m$ follows by elementary considerations.)

The polynomial (6) for instance, which occurs in Theorem 2, belongs to $\mathcal{Z}_6$. Looking at the quite long proof of that theorem, raises the issue: is not there a shorter way of proving? As an analogue, consider Calderon's result [1]: If $f \in \mathcal{B}_{m,n}$, and $m = 2$ or $n = 2$, then f is SOS.

We guess that the more special case $\mathcal{Z}_m$ also is manageable.

Problem. Characterize the SOS polynomials $f \in \mathcal{Z}_m$!

Remark 3. If $m = 3$, there are no basic relations like (4), hence nonnegativity coincides with SOS property. Therefore the smallest nontrivial problem is provided by the case $m = 4$. Then there is exactly one basic relation, namely $z_{1,2}z_{3,4} + z_{1,4}z_{2,3} = z_{1,3}z_{2,4}$, Hence not all vectors $z \in \mathbb{R}^6$ can be represented as $z = (z_{1,2}, z_{2,3}, z_{3,4}, z_{1,3}, z_{2,4}, z_{1,4})^T$ with $z_{i,j} = x_i y_j - y_i x_j$, giving the probable reason for the difficulties.

References

[1] **Calderon, A.P.,** A note on biquadratic forms, *Linear Algebra Appl.,* **7** (1973), 175-177.

[2] **Böttcher, A. and Wenzel, D.,** The Frobenius norm and the commutator, *Linear Algebra Appl.,* **429** (2008), 1864-1885.

[3] **Chesi, G.,** On the gap of positive polynomials which are not SOS, *IEEE Trans. on Automatic Control,* **52** (6) (2007), 1066-1072. `http://www.nd.edu/~ieeetac/index_2007.pdf`

[4] **Choi, M.D., Lam, T.Y. and Reznick, B.,** Sums of squares of real polynomials, *Proc. Sympos. Pure Math.,* **58.2** (1995), 103-126.

[5] **László, L.,** Proof of Böttcher and Wenzel's conjecture on commutator norms for 3-by-3 matrices, *Linear Algebra Appl.,* **422** (2007), 659-663.

[6] **Parrilo, P.,** `ocw.mit.edu/OcwWeb/Electrical-Engineering-and-Computer-Science/6-972Spring-2006/CourseHome/index.htm`

[7] **Parrilo, P.,** Semidefinite programming relaxations for semialgebraic problems, *Math Program. Ser. B,* **96** (2003), 293-320.

[8] **Reznick, B.,** Some concrete aspects of Hilbert's 17th problem, *In Contemporary Mathematics,* AMS, (2000), 251-272., see also: `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.49.8114.`

[9] **Seak-Weng Vong and Xiao-Qing Jin,** Proof of Böttcher and Wenzel's conjecture, *OAM,* **2** (3) (2008), 435-442.

L. László
Department of Numerical Analysis
Eötvös Loránd University
Pázmány Péter sét. 1/C
H-1117 Budapest, Hungary
`Laszlo@numanal.inf.elte.hu`