Acta Univ. Sapientiae, Mathematica, **3**, 1 (2011) 77-89

On a structured semidefinite program

Lajos László
Department of Numerical Analysis
Eötvös Loránd University, Hungary
email: laszlo@numanal.inf.elte.hu

Abstract. The nonnegative biquadratics discussed comes from the Böttcher-Wenzel inequality. It is for some matrices a sum of squares of polynomials (SOS), in other cases not, depending on the nonzero pattern of the matrices at issue. Our aim was to draw a line between them. To prove the 'not a SOS' case we solve a semidefinite programming (SDP) problem. Subsequently a two-parameter version will be investigated.

1 Introduction

The Böttcher-Wenzel inequality (see [2], [7], [3], [1], [9], [5], [4], [10]) states (in its stronger form) that for real square matrices X, Y of the same order n

$$f(X,Y) \equiv 2\,||X||^2||Y||^2 - 2\,\mathrm{trace}^2(X^TY) - ||XY-YX||^2 \geq 0, \qquad (1)$$

where the norm used is the Frobenius norm. Since all our attempts to obtain a representation for f as a sum of polynomial squares (in short: SOS) failed for $n = 3$, distinguishing between the 'good' and 'bad' cases became to a natural problem.

In case of $n = 2$ we have for $X = \begin{pmatrix} x_1 & x_3 \\ x_2 & x_4 \end{pmatrix}$, $Y = \begin{pmatrix} y_1 & y_3 \\ y_2 & y_4 \end{pmatrix}$ and with variables $z_{i,j} = x_iy_j - y_ix_j,\ 1 \leq i < j \leq 4,$ that

$$f(X,Y) = 2\,z_{1,4}^2 + (z_{1,2} - z_{2,4})^2 + (z_{1,3} - z_{3,4})^2$$

is a sum of squares of quadratics.

2010 Mathematics Subject Classification: 90C22, 15A45, 65F15
Key words and phrases: sum of squares, semidefinite programming, biquadratics

Our main result is that the nonnegative form (1) is SOS for *good* matrices X, Y, whereas it isn't SOS for general *bad* matrices, where a matrix of order n will be called *good*, if nonzero elements occur only in row 1 and column n, while it is called *bad*, if, moreover, nonzero elements occur also in the main diagonal, as shown e.g. for $n = 4$:

$$good : \begin{pmatrix} * & * & * & * \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & * \end{pmatrix} \qquad bad : \begin{pmatrix} * & * & * & * \\ 0 & * & 0 & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{pmatrix}.$$

Remark 1 For convenience we re-cite the SOS representation [8] for the *good* cases in section 2. Section 3 contains the main result: the non-possibility of an SOS-representation for the *bad* cases via SDP, while in section 4 we provide the function f with two parameters and decompose the unit square into regions with different properties.

2 SOS decomposition for good matrices

Let X, Y be *good* real n-th order 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 vectors $x = (x_i)_1^m$ and $y = (y_i)_1^m$ as

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

The SOS representation for these *good* matrices is the following.

Theorem 1 *(Theorem 1, [8])*

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

Remark 2 Indeed, (1) and (2) are identical. In particular,

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

and

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

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

3 SOS decomposition impossible for bad matrices

It suffices to prove this negative result for third order matrices. Let

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

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 the form

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

with $z_{i,j} = x_iy_j - x_jy_i,\ 1 \le i < j \le 6$.

Theorem 2 *The biquadratic form (3), nonnegative for any real* $(x_i)_1^6$, $(y_i)_1^6$, *is not a sum of squares of any quadratics!*

We will need a lemma before proving the theorem.

Lemma 1 *An sos-representation of* f *is necessarily a sum of squares of the* $z_{i,j}$*'s. Furthermore, for the variables* $z_{i,j}$ *the basic identities*

$$z_{i,j}\, z_{k,l} + z_{i,l}\, z_{j,k} - z_{i,k}\, z_{j,l} = 0, \quad 1 \le i < j < k < l \le 6 \tag{4}$$

hold, and there are no more (quadratic) relations between them.

Proof. Note that in addition to nonnegativity: $f(X,Y) \ge 0$, we have symmetry: $f(X,Y) = f(Y,X)$, and also zero property: $f(X,X) = 0$.

By virtue of the last property, the coefficients of x_1y_2 and y_1x_2 are opposite in all terms of the representation

$$f(X,Y) = \sum_i (\alpha_i x_1 y_2 + \beta_i y_1 x_2 + \dots)^2,$$

i.e. $\beta_i = -\alpha_i$ for all i. Hence

$$f(X,Y) = \sum_i (\alpha_i z_{1,2} + \gamma_i x_1 y_3 + \delta_i y_1 x_3 + \dots)^2$$

and the procedure can be continued.

As for the relations between the $z_{i,j}$-s, assume that there holds a nontrivial quadratic identity $g(Z) = 0$ containing the term $z_{1,2}^2$. Then also $x_1^2 y_2^2$ is present, however, this latter can only occur in $z_{1,2}^2$, therefore a term $-z_{1,2}^2$ is needed to cancel it, which contradicts the non-triviality. In a similar way we see that there is no term of type $z_{1,2}\, z_{1,3}$ occurring in a non-trivial identity.

Finally, assume we have a non-trivial identity containing the term $z_{1,2}\, z_{3,4}$. (Its coefficient can be supposed to be unity.) Then $x_1 y_2 x_3 y_4$ is a part of the (expanded) identity. In contrast to the above cases, this occurs in two additional terms: in $z_{1,4}\, z_{2,3}$ and in $z_{1,3}\, z_{2,4}$ to produce the non-trivial identity $z_{1,2}\, z_{3,4} + z_{1,4}\, z_{2,3} - z_{1,3}\, z_{2,4} = 0$.

Since $x_1 y_2 x_3 y_4$ occurs only in the expansion of the three above terms, there are no more non-trivial identities containing it. □

Before proving the theorem, we formulate the standard primal and dual semidefinite programs:

$$\min \{C \bullet X : \ X \ge 0, \ A_i \bullet X = b_i, \ 1 \le i \le m\} \qquad (Primal)$$

$$\max\ \{b^T y:\ S \equiv C - \sum_{i=1}^{m} y_i A_i \geq 0\} \qquad (Dual)$$

where all matrices are n-th order real symmetric, m is the number of constraints, C and $(A_i)_1^m$ are given, vector b of length m is also given, while the primal matrix X and the dual matrix S (the so-called 'slack' matrix - sometimes denoted by Z) together with the m-vector y are the output of the program, $\bullet$ denotes the standard scalar product $A \bullet B = \mathrm{trace}(AB)$ for symmetric matrices and $\geq$ stands for the Loewner ordering: $A \geq B$ iff $A - B$ is positive semidefinite, in short: psd.

Turning to our case, denote by $(A_i)_1^{15}$ the constraint matrices corresponding to the basic identities (4) mentioned in the Lemma. Since these are homogeneous equations, the b_i-s are zero. In an interesting way, both the order $n = \binom{6}{2}$ and the number of constraints $m = \binom{6}{4}$ equals 15.

Nevertheless we will need also the identity I as a constraint matrix to get a sum of squares decomposition, and – to emphasize its speciality – we associate it with index zero, i.e. we write $A_0 = I$ and get the *concrete* primal-dual pair of SDP programs:

$$\min\ \{C \bullet X:\ X \geq 0,\ \mathrm{tr}X = 1,\ A_i \bullet X = 0,\ 1 \leq i \leq 15\} \qquad (Primal)$$

$$\max\ \{y_0:\ S \equiv C - y_0 I - \sum_{i=1}^{15} y_i A_i \geq 0\} \qquad (Dual)$$

After this preparation we can prove our theorem.

Proof. To prove Theorem 2, we specify in detail the data for the SDP above and explain the results obtained. Considering the band-width of matrices C, X and S, a good order of the $z_{i,j}$'s is

$$(z_{2,5}, z_{3,4}, z_{1,2}, z_{2,6},\ z_{1,4}, z_{2,3}, z_{4,5}, z_{4,6},\ z_{1,3}, z_{3,5}, z_{2,4},\ z_{1,5}, z_{1,6}, z_{3,6}, z_{5,6}).$$

Then, denoting by z the corresponding column vector, it holds that $f(X,Y) = z^T C z$ for C appropriately defined. To this, we describe the common block-structure of the matrices C, S, X. All these matrices are block-diagonal with two 4×4 blocks and a 3×3 block, while the remaining 4×4 block is diagonal. In case of C e.g. these blocks will be denoted by C_4, C_4', C_3 and C_d. Here, C_4' is diagonally similar to C_4 through $\mathrm{diag}(1,1,1,-1)$, hence the eigenvalues of C_4' and C_4 coincide. The whole matrix is

$$C = C_4 \oplus C_4' \oplus C_3 \oplus C_d,$$

and the same direct sum representation holds for the optimal primal and dual matrices X and S. As regards C, we have

$$C_4 = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \quad C_3 = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}, \quad C_d = 2I_4.$$

Now we can explain the output of our program. The optimal value of the objective is negative: $y_0 = -\frac{1}{7}$, indicating that (1) is not a SOS, but the modified quartics

$$2\tfrac{1}{7}\,||X||^2||Y||^2 - 2\,\mathrm{trace}^2(X^TY) - ||XY - YX||^2$$

is a sum of squares, $2\frac{1}{7}$ being the smallest number with this property. The only nonzero y-s are $y_1 = y_2 = 5/7$, they correspond to the basic relations (4) with indices $(1,2,3,4)$ and $(2,3,4,5)$.

For the optimal matrix S we have $S = S_4 \oplus S_4' \oplus S_3 \oplus S_d$ with

$$S_4 = \frac{1}{7}\begin{pmatrix} 15 & -5 & 0 & 0 \\ -5 & 15 & -5 & 0 \\ 0 & -5 & 8 & -7 \\ 0 & 0 & -7 & 8 \end{pmatrix}, \quad S_3 = \frac{1}{7}\begin{pmatrix} 8 & -7 & -2 \\ -7 & 8 & -2 \\ -2 & -2 & 8 \end{pmatrix}, \quad S_d = \frac{15}{7}I_4,$$

which yields the wanted sum of squares decomposition. The optimal primal matrix is $X = X_4 \oplus X_4' \oplus X_3 \oplus X_d$, where

$$X_4 = \frac{2}{735}\begin{pmatrix} 1 & 3 & 8 & 7 \\ 3 & 9 & 24 & 21 \\ 8 & 24 & 64 & 56 \\ 7 & 21 & 56 & 49 \end{pmatrix}, \quad X_3 = \frac{9}{245}\begin{pmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{pmatrix},$$

and X_d is the fourth order zero matrix. Using the block-structure, the positive semidefiniteness of X and S and the complementarity condition $XS = 0$ can easily be checked (cf. the Karush-Kuhn-Tucker necessary conditions). Also, strict complementarity holds, in particular $\mathrm{rank}(X) = \mathrm{def}(S) = 3$. □

Notice that in general – unlike linear programming – rational data for a SDP problem does not necessarily result in rational solution!

4 On a parametric version

To get more insight into the problem, we insert two parameters α and β to investigate the SOS representability of the biquadratics

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

(The reason for the two β's is that these terms behave similarly.) It turns out that only the first two constraints $A_1 \bullet X = 0$ and $A_2 \bullet X = 0$ will be active with $y_1 = y_2$, and $y_i = 0,\ i \ge 3$, as in the above special case of $\alpha = \beta = 1$. This means that our problem reduces to finding the optimal y_0, y_1 for a given pair $(\alpha, \beta) \in [0,1]^2$ such that

$$(2-y_0)f_0 - \alpha f_1 - \beta f_2 - y_1 F_1 - y_2 F_2 \ge 0 \quad (6)$$

and y_0 is maximum, where we use the abbreviations

$$f_0 = \sum_{i<j}^{6} z_{i,j}^2,\ f_1 = (z_{1,3}+z_{2,4}+z_{3,5})^2,\ f_2 = (z_{1,2}+z_{2,6})^2 + (z_{4,5}-z_{4,6})^2,$$

$$F_1 = 2(z_{1,2}\, z_{3,4} + z_{1,4}\, z_{2,3} - z_{1,3}\, z_{2,4}),\ F_2 = 2(z_{2,3}\, z_{4,5} + z_{2,5}\, z_{3,4} - z_{2,4}\, z_{3,5})$$

in connection with the notations

$$p = 2 - y_0, \qquad q = y_1 = y_2, \quad \text{and} \quad F = F_1 + F_2$$

to write (6) in the simpler form

$$pf_0 - \alpha f_1 - \beta f_2 - qF.$$

Observation. Assume that for some $(\alpha,\beta) \in [0,1]^2$ we know the optimal values of y_0, y_1, i.e. the optimal p and q. Then by multiplying through the coefficient vector (α, β, p, q) by $2/p$ we get $(\alpha', \beta', p', q')$ with

$$\alpha' = \frac{2\alpha}{p},\ \beta' = \frac{2\beta}{p},\ p' = 2,\ q' = \frac{2q}{p},$$

showing that for this new (α', β') we have $y_0' = 0$.

Example. Let us calculate the largest $\alpha = \beta$ for which (5) is SOS! (Theorem 2 tells us that this $\alpha < 1$.) For $\alpha = \beta = 1$ we know that $y_0 = -\frac{1}{7}$, thus $p = \frac{15}{7}$, and $q = y_1 = \frac{5}{7}$. The transformed variables are $\alpha' = \beta' = \frac{14}{15}$ and $q' = y_1' = \frac{2}{3}$.

Table 1: The unit square: optimal values

region	name	$p = 2 - y_0$	$q = y_1$	def(S)	rank(X)
$\beta = 0$		$2\,\alpha$	α	1	1
$\beta < 0.8\,\alpha$	R_1	$2\,\alpha$	α	1	1
$\beta = 0.8\,\alpha$		$2\,\alpha$	α	3	1
$\beta \in (0.8\alpha, 1.5\alpha)$	R_3	(7)	(8)	3	3
$\beta = 1.5\,\alpha$		$2\,\beta$	0	3	2
$\beta > 1.5\,\alpha$	R_2	$2\,\beta$	0	2	2
$\alpha = 0$		$2\,\beta$	0	2	2

Consequently the polynomial $2f_0 - \frac{14}{15}(f_1 + f_2)$ is not only nonnegative but also SOS (and the point $(\alpha, \beta) = (\frac{14}{15}, \frac{14}{15})$ lies on the border of the 'good' and 'bad' cases).

In the next theorem we summarize the results obtained for parameters (α, β) from the unit square.

Theorem 3 *Table 1 gives the optimal values for the parametrized problem (6).*

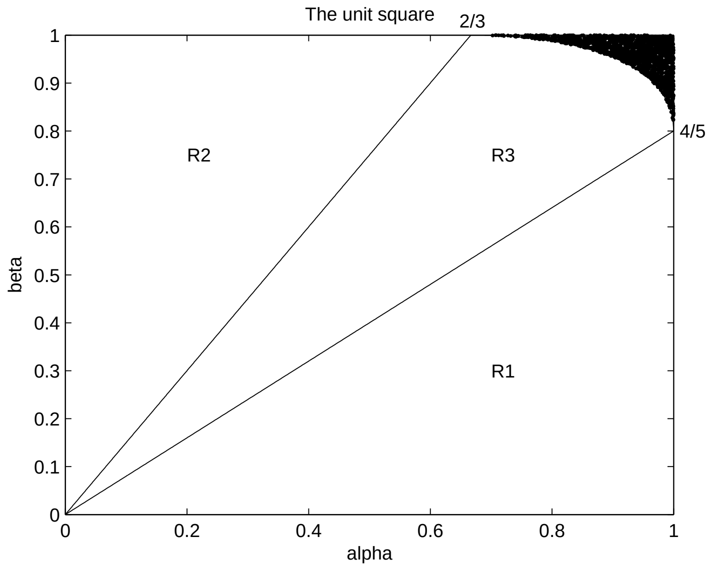
The optimal p and q for the middle sector R_3 are

$$p = -\alpha\beta\,\frac{2\alpha + 9\beta + \sqrt{(8\alpha - 9\beta)^2 + 15\beta^2}}{4\alpha^2 - 12\alpha\beta + \beta^2} \tag{7}$$

$$q = \frac{\beta(p(60\alpha^2 - 4\alpha\beta - 9\beta^2) + 15\alpha\beta(\beta - 6\alpha)}{4(4\alpha^2 - 12\alpha\beta + \beta^2)(3\beta - 2p)} \tag{8}$$

In addition to this "radial" characterization, the level lines also can be described: these are curves, along which the optimal y_0 is constant. The case $y_0 = 0$, i.e. $p = 2$ can be seen on the figure, cf. also (10). (For another instance, the dotted line starting at $\alpha = 0$, $\beta = 0.5$ and ending at $\alpha = 0.5$, $\beta = 0$ belongs to $p = 1$.) The indices of region names correspond to the defects of the optimal dual matrix S. The right upper darkened region within R_3 refers to points (α, β) for which (5), i.e. $2f_0 - \alpha f_1 - \beta f_2$ is not a sum of squares.

Proof. The block-structure of the special case $\alpha = \beta = 1$ keeps on holding,



and the blocks at issue depend just on one of the parameters α and β :

$$S_3(\alpha) = \begin{pmatrix} p-\alpha & -\alpha & q-\alpha \\ -\alpha & p-\alpha & q-\alpha \\ q-\alpha & q-\alpha & p-\alpha \end{pmatrix}, \; S_4(\beta) = \begin{pmatrix} p & -q & 0 & 0 \\ -q & p & -q & 0 \\ 0 & -q & p-\beta & -\beta \\ 0 & 0 & -\beta & p-\beta \end{pmatrix},$$

as it easily follows by considering the polynomial (6).

Let us begin with the case $\beta = 0$. Then S_4 (and also $S_4{}'$) is psd, and our 'work matrix' is S_3. Its (1,1) entry, $p - \alpha \geq 0$, and the nonnegativity of the left upper 2-minor implies $|p - \alpha| \geq |\alpha|$. Since both sides are nonnegative, it follows that $p \geq 2\alpha$. On the other hand, the determinant of S_3 equals

$$|S_3| \equiv \det(S_3) = p\,\{p^2 - 2q^2 - \alpha(3p - 4q)\},$$

which must vanish at the optimal variables, therefore

$$\big(p - \frac{3}{2}\alpha\big)^2 = 2\big(q - \alpha\big)^2 + \frac{1}{4}\alpha^2.$$

From this we get $|p - \frac{3}{2}\alpha| \geq |\frac{1}{2}|$, and, since this holds without absolute value as well, we conclude that the optimal variables are $q = 0$ and $p = 2\alpha$ (note that maximizing y_0 is equivalent to minimizing $p = 2 - y_0$).

All this holds for R_1, i.e. for β "small" – until we arrive at $|S_4(\beta)| = 0$. Before that moment we still have $p = 2\alpha,\ q = \alpha$ and

$$S_3 = \alpha \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

giving one rank decrease. As regards S_4, for $\beta \le (4/5)\,\alpha$ we have

$$S_4 \ge \begin{pmatrix} p & -\alpha & 0 & 0 \\ -\alpha & p & -\alpha & 0 \\ 0 & -\alpha & p - \frac{4}{5}\alpha & -\frac{4}{5}\alpha \\ 0 & 0 & -\frac{4}{5}\alpha & p - \frac{4}{5}\alpha \end{pmatrix} = \frac{\alpha}{5} \begin{pmatrix} 10 & -5 & 0 & 0 \\ -5 & 10 & -5 & 0 \\ 0 & -5 & 6 & -4 \\ 0 & 0 & 4 & 6 \end{pmatrix}$$

in the sense of the semidefinite (Loewner) ordering, which easily follows by (6). Matrix S_4 becomes active, if equality is attained: $\beta = (4/5)\alpha$. Then S_4 becomes a psd matrix with rank 3, therefore S_4 and S_4' yield two further rank losses. (The situation on the border of R_1 and R_3 can be understood by thinking of the continuity of the roots depending on the parameters.)

The cases $\alpha = 0$ and $\beta > 1.5\,\alpha$ are similar, hence we give only the necessary formulas. The determinant of S_4 is

$$|S_4| = p\,\{p^2(p - 2\beta) - q^2(2p - 3\beta)\}.$$

Inequality $p \ge 2\beta$ follows similarly to the case $p \ge 2\alpha$. On the other hand, $|S_4| = 0$ can be rewritten as

$$p(p-\beta)^2 + 3q^2\beta = p(2q^2 + \beta^2),$$

whence we conclude $(p-\beta)^2 \le 2q^2 + \beta^2$, i.e. $p(p - 2\beta) \le 2q^2$, and it follows that the optimal values ($p \to \min!$) are $q = 0$ and $p = 2\beta$. The matrix

$$S_4 = \beta \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

is psd with defect one, giving (together with ${S_4}'$) rank loss=2 for the whole matrix S. As regards S_3, for $(\alpha, \beta) \in R_2$ it equals

$$S_3(\alpha) = pI_3 - \alpha\, ee^T = \alpha\big(\tfrac{p}{\alpha} I_3 - ee^T\big), \quad e = (1,1,1,1)^T,$$

which is positive definite for $p/\alpha > 3$, i.e. for $\beta > (3/2)\,\alpha$, and psd (with rank=2) if equality holds:

$$S_3 = \alpha \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

In this latter case, i.e. if $\beta = (3/2)\,\alpha$, the defect of S becomes 3, since not only S_4 and S_4', but also S_3 yields a rank loss.

The middle region R_3 can be characterized by that the SDP program annihilates *both* determinants. The equations $|S_3(\alpha)| = 0$ and $|S_4(\beta)| = 0$ are equivalent to

$$p^2 - 2q^2 = \alpha(3p - 4q) \quad \text{and} \quad p(p^2 - 2q^2) = \beta(2p^2 - 3q^2), \tag{9}$$

from which we can express p and q by means of α and β to get (7) and (8), and also β by help of α and p :

$$\varphi_p(\alpha) = \tfrac{2p}{3} - \tfrac{2p^3}{3}\big(p^2 + 9\alpha p - 12\alpha^2 + 6\alpha\sqrt{2(p-\alpha)(p-2\alpha)}\,\big)^{-1}$$

In the important special case $p = 2$ we have the function $\varphi : [\frac{2}{3}, 1] \to [\frac{4}{5}, 1]$ defined by

$$\varphi(\alpha) \equiv \varphi_2(\alpha) = \tfrac{4}{3} - \tfrac{8}{3}\big(2 + 9\alpha - 6\alpha^2 + 6\alpha\sqrt{(1-\alpha)(2-\alpha)}\,\big)^{-1}. \tag{10}$$

The critical values of φ are:

$$\varphi(\tfrac{2}{3}) = 1,\ \varphi'(\tfrac{2}{3}) = 0,\ \varphi(1) = \tfrac{4}{5},\ \varphi'(1) = -\infty.$$

The set of points in the unit square, for which (5) is a SOS are delimited by the α and β axes, the horizontal line segment $(0, 1)$ to $(\frac{2}{3}, 1)$, the graph of φ and the vertical line segment $(1, 0)$ to $(1, \frac{4}{5})$. All other curves with different p (e.g. that with $p = 1$ dotted on the figure) are proportional to this one, since equations (9) are homogeneous. □

Remark 3 Consider the ellipse with center in $(\frac{2}{3}, \frac{4}{5})$ and vertices $(1, \frac{4}{5})$, $(\frac{2}{3}, 1)$, the right upper quarter of which is close to the graph of φ. The elementary equality $(\frac{14}{15} - \frac{2}{3})^2 + (1 - \frac{4}{5})^2 = (1 - \frac{2}{3})^2$ shows that the projection of the 'border point' $(\frac{14}{15}, \frac{14}{15})$ onto the longer axis of the ellipse is just its focus $(\frac{14}{15}, \frac{4}{5})$! As for their measures, the approximate area of the 'bad' (shadowed) region is 0.0121, while that above the ellipse amounts to $\frac{1}{15} - \frac{\pi}{60} \approx 0.0143$.

Remark 4 For the interested reader we give some 'nice' rational solutions: in addition to the known quadruples $(\alpha, \beta, p, q) = (1, 1, \frac{15}{7}, \frac{5}{7})$ and $(\frac{14}{15}, \frac{14}{15}, 2, \frac{2}{3})$ we have e.g. $(\alpha, \beta, p, q) = (\frac{34}{35}, \frac{17}{19}, 2, \frac{4}{5}), (\frac{17}{31}, \frac{68}{23}, 6, 1)$, or $(\frac{31}{16}, \frac{124}{50}, 5, \frac{5}{8})$.

Considering the derivatives also is of interest: the slope of φ_p for $p = 2\frac{1}{7}$ at $\alpha = 1$ equals $-\frac{3}{4}$, which can be used to define a new problem with the same solution! Replace to this the identity $A_0 = I$ by two matrices A_{16} and A_{17} associated with the quadratic forms f_1 and f_2, and set the corresponding coordinates of b equal to 3 and 4 (coming from the numerator and denominator of the ratio $-\frac{3}{4}$ above). The result will coincide with that of Theorem 2.

Remark 5 In [6] the authors write: "Unfortunately, the nature of a parametric SDP is far more complicated [than LP] due to regions of nonlinearity of $\phi(\gamma)$." (The function $\phi(\gamma) = C(\gamma) \bullet X(\gamma)$ is the primal objective depending on the parameter.) In light of this, present problem seems to be a refreshing exception: the nonlinearity (cf. the functions φ_p) can be handled by means of elementary functions.

References

[1] K. Audenaert, Variance bounds, with an application to norm bounds for commutators, *Linear Algebra Appl.*, **432** (2010), 1126–1143.

[2] A. Böttcher, D. Wenzel, How big can the commutator of two matrices be and how big is it typically?, *Linear Algebra Appl.*, **403** (2005), 216–228.

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

[4] Ch-M. Cheng, S-W. Vong, D. Wenzel, Commutators with maximal Frobenius norm, *Linear Algebra Appl.*, **432** (2010), 292–306.

[5] K-S. Fong, Ch-M. Cheng, I-K. Lok, Another unitarily norm attaining the minimum norm bound for commutators, *Linear Algebra Appl.*, **433**, (2010), 1793–1797.

[6] D. Goldfarb, K. Scheinberg, On parametric semidefinite programming, *Appl. Numer. Math.*, **29** (1999), 361–377.

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

[8] L. László, On Sum of Squares Decomposition for a Biquadratic Matrix Function, *Ann. Univ. Sci. Budapest. Sect. Comput.*, **33** (2010), 273–284.

[9] S-W. Vong, X-Q. Jin, Proof of Böttcher and Wenzel's conjecture, *Oper. Matrices*, **2** (2008), 435–442.

[10] D. Wenzel, K. M. R. Audenaert, Impressions of convexity - An illustration for commutator bounds, *Linear Algebra Appl.*, **433** (2010), 1726–1759.

Received: 22, January, 2011