On Local Convexity of Quadratic Transformations

Yong Xia

Received: 23 May 2014 / Revised: 1 August 2014 / Accepted: 4 August 2014 /

Published online: 31 August 2014

© Operations Research Society of China, Periodicals Agency of Shanghai University, and Springer-Verlag Berlin Heidelberg 2014

Abstract In this paper, we improve Polyak's local convexity result for quadratic transformations. Extension and open problems are also presented.

Keywords Convexity · Quadratic transformation · Joint numerical range

1 Introduction

Let $x \in \mathbb{R}^n$ and $f(x) = (f_1(x), \dots, f_m(x))$, where

$$f_i(x) = \frac{1}{2}x^T A_i x + a_i^T x, \quad i = 1, \dots, m$$

are quadratic functions and $A_i \in \mathbb{R}^{n \times n}$ ($i = 1, \dots, m$) are symmetric. One interesting question is when the following joint numerical range

$$F_m = \{f(x) : x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$$

is convex.

This research was supported by Beijing Higher Education Young Elite Teacher Project (No. 29201442), and by the fund of State Key Laboratory of Software Development Environment (No. SKLSDE2013ZX-13).

Y. Xia (✉)

State Key Laboratory of Software Development Environment, LMIB of the Ministry of Education,
School of Mathematics and System Sciences, Beihang University, Beijing 100191, China
e-mail: dearyxia@gmail.com

Though results on the convexity of complex quadratic functions were already there since 1918, see Toeplitz [19] and Hausdorff [8], the first such result for real case is due to Dines [4] in 1941. It states that if f_1, f_2 are homogeneous quadratic functions then the set F_2 is convex. In 1971, Yakubovich [23, 24] used this basic result to prove the famous S-lemma, see [17] for a survey. Brickman [3] proved in 1961 that if f_1, f_2 are homogeneous quadratic functions and $n \geq 3$ then the set $\{(f_1(x), f_2(x)) : x \in \mathbb{R}^n, \|x\| = 1\} \subseteq \mathbb{R}^2$ is convex. Fradkov [5] proved in 1973 that if matrices $A_1, \dots, A_m$ commute and $f_1, \dots, f_m$ are homogeneous, then F_m is convex. In 1995, it was showed by Ramana and Goldman [18] that the identification of the convexity of F_m is NP-hard. In the same paper, the quadratic maps, under which the image of every linear subspace is convex, was also investigated. Based on Brickman's result, Polyak [14] proved in 1998 that if $n \geq 3$ and f_1, f_2, f_3 are homogeneous quadratic functions such that $\mu_1 A_1 + \mu_2 A_2 + \mu_3 A_3 \succ 0$ (where notation $A \succ 0$ means that A is positive definite) for some $\mu \in \mathbb{R}^3$, then the set F_3 is convex. Moreover, as shown in the same paper, when $n \geq 2$ and there exists $\mu \in \mathbb{R}^2$ such that $\mu_1 A_1 + \mu_2 A_2 \succ 0$, the set F_2 is convex. In 2007, Beck [1] showed that if $m \leq n, A_1 \succ 0$ and $A_2 = \dots = A_m = 0$, then F_m is convex. However, if $A_1 \succ 0, A_2 = \dots = A_{n+1} = 0$ and $a_2, \dots, a_{n+1}$ are linearly independent, then F_{n+1} is not convex. When $m = 2$, Beck's result reduces to be a corollary of Polyak's result. Very recently, Xia et al. [22] used the new developed S-lemma with equality to establish the necessary and sufficient condition for the convexity of F_2 for $A_2 = 0$ and arbitrary A_1 .

More generally, Polyak [15, 16] succeeded in proving a nonlinear image of a small ball in a Hilbert space is convex, provided that not only the derivative of the map is Lipschitz continuous on the ball, but also the derivative at the center of the ball is surjective and the norm of its adjoint mapping is bounded from zero. Later, Uderzo [21] extended the result to a certain subclass of uniformly convex Banach spaces. When focusing on quadratic transformations, Polyak's result reads as follows:

Theorem 1.1 [17] *Let $A = [a_1 \ \dots \ a_m] \in \mathbb{R}^{n \times m}$ and define*

$$L := \sqrt{\sum_{i=1}^m \|A_i\|^2}, \quad (1.1)$$

$$v := \sigma_{\min}(A) = \sqrt{\lambda_{\min}(A^T A)},$$

where $\|A_i\| = \sigma_{\max}(A_i) = \sqrt{\lambda_{\max}(A_i^T A_i)}$ is the spectral norm of A_i , $\sigma_{\min}(\cdot), \lambda_{\min}(\cdot), \sigma_{\max}(\cdot), \lambda_{\max}(\cdot)$, denote the smallest and largest singular value and eigenvalue, respectively. Assume $v > 0$.

If $\varepsilon < \varepsilon^ := v/(2L)$, then the image*

$$F_m(\varepsilon) = \{f(x) : x \in \mathbb{R}^n, \|x\| \leq \varepsilon\} \quad (1.2)$$

is a convex set in $\mathbb{R}^m$.

Polyak [15, 16] used the following example to show his estimation ε^* is tight, where $n = m = 2$ and

$$f_1(x) = x_1x_2 - x_1, f_2(x) = x_1x_2 + x_2.$$

Actually, in this case, $\varepsilon^* = 1/(2\sqrt{2}) \approx 0.3536$. It is trivially verified that $F_m(\varepsilon)$ is convex for $\varepsilon \leq \varepsilon^*$ and loses convexity for $\varepsilon > \varepsilon^*$.

In this paper, we improve the above Polyak's result for quadratic transformations (i.e., Theorem 1.1) by strengthening the constant L . Then, Theorem 1.1 is extended to the image of the ball of the same radius ε centered at any point a satisfying $\|a\| < 2(\varepsilon^* - \varepsilon)$. Furthermore, we propose two new approaches for possible improvement of L .

The paper is organized as follows. In Sect. 1, we improve and extend Theorem 1.1. In Sect. 2, we discuss further possible improvements. In the final conclusion section, we propose two open questions.

Throughout the paper, all vectors are column vectors. Let $v(\cdot)$ denotes the optimal value of problem $(\cdot)$. Notation $A \succeq 0$ implies that the matrix A is positive semidefinite. $\text{vec}(A)$ denotes the vector obtained by stacking the columns of A one underneath the other. The trace of A is denoted by $\text{trace}(A) = \sum_{i=1}^n A_{ii}$. $|A|$ is the matrix having entries $|A_{ij}|$. The Kronecker product and the inner product of the matrices A and B are denoted by $A \otimes B$ and $A \bullet B = \text{trace}(AB^T) = \sum_{i,j=1}^n a_{ij}b_{ij}$, respectively. The identity matrix is denoted by I . $\|x\| = \sqrt{x^T x}$ is the ℓ_2 -norm of the vector x .

2 Main Results

In this section, we first improve Theorem 1.1 and then extend it to the ball of the same radius centered at any point close enough to the zero point.

Theorem 2.1 *Define*

$$L_{\text{new}} := \sqrt{\lambda_{\max} \left(\sum_{i=1}^m A_i^T A_i \right)}. \quad (2.1)$$

Then we have

$$L_{\text{new}} \leq L. \quad (2.2)$$

For any $\varepsilon < \varepsilon_{\text{new}}^ := v/(2L_{\text{new}})$, the image $F_m(\varepsilon)$ defined in (1.2) is convex.*

Proof Let L_b be any upper bound of the Lipschitz constant of f , i.e.,

$$\|\nabla f(x) - \nabla f(z)\| \leq L_b \|x - z\|, \quad \forall x, z \in \mathbb{R}^n. \quad (2.3)$$

According to the proof in [15], Theorem 1.1 remains true if L defined in (1.1) is replaced by L_b . It is sufficient to show that $L_b := L_{\text{new}}$ satisfies (2.3). To this end, we have

$$\begin{aligned}
 & \max_{\|x-z\|=1} \|\nabla f(x) - \nabla f(z)\| \\
 &= \max_{\|x-z\|=1} \| [A_1(x-z) \cdots A_m(x-z)] \| \\
 &= \max_{\|y\|=1} \| [A_1 y \cdots A_m y] \| \\
 &= \sqrt{\max_{\|y\|=1} \lambda_{\max}([A_1 y \cdots A_m y]^T [A_1 y \cdots A_m y])}
 \end{aligned} \tag{2.4}$$

$$\begin{aligned}
 &\leq \sqrt{\max_{\|y\|=1} \text{trace}([A_1 y \cdots A_m y]^T [A_1 y \cdots A_m y])} \\
 &= \sqrt{\max_{\|y\|=1} y^T \left(\sum_{i=1}^m A_i^T A_i \right) y} \\
 &= \sqrt{\lambda_{\max} \left(\sum_{i=1}^m A_i^T A_i \right)}.
 \end{aligned} \tag{2.5}$$

The inequality (2.2) holds since

$$\begin{aligned}
 L_{\text{new}} &= \sqrt{\lambda_{\max} \left(\sum_{i=1}^m A_i^T A_i \right)} = \sqrt{\max_{\|y\|=1} y^T \left(\sum_{i=1}^m A_i^T A_i \right) y} \\
 &\leq \sqrt{\sum_{i=1}^m \left(\max_{\|y\|=1} y^T A_i^T A_i y \right)} = \sqrt{\sum_{i=1}^m \lambda_{\max}(A_i^T A_i)} = \sqrt{\sum_{i=1}^m \|A_i\|^2} = L.
 \end{aligned}$$

□

Theorem 2.2 For any $0 < \varepsilon < \varepsilon_{\text{new}}^* = v/(2L_{\text{new}})$ and any $a \in \mathbb{R}^n$ such that $\|a\| < 2(\varepsilon_{\text{new}}^* - \varepsilon)$, the image

$$F_m(\varepsilon, a) = \{f(x) : x \in \mathbb{R}^n, \|x - a\| \leq \varepsilon\}$$

is a convex set in $\mathbb{R}^m$.

Proof For any $a \in \mathbb{R}^n$ such that $\|a\| < 2(\varepsilon_{\text{new}}^* - \varepsilon)$, we have

$$\begin{aligned}
 & \sigma_{\min}(A + [A_1 a \cdots A_m a]) \\
 & \geq \sigma_{\min}(A) - \sigma_{\max}(-[A_1 a \cdots A_m a]) \\
 & \geq \sigma_{\min}(A) - \sup_{\|a\| < 2(\varepsilon_{\text{new}}^* - \varepsilon)} \sigma_{\max}([A_1 a \cdots A_m a]) \\
 & = \sigma_{\min}(A) - \sqrt{\sup_{\|a\| < 2(\varepsilon_{\text{new}}^* - \varepsilon)} \lambda_{\max}([A_1 a \cdots A_m a]^T [A_1 a \cdots A_m a])} \\
 & \geq \sigma_{\min}(A) - \sqrt{\sup_{\|a\| < 2(\varepsilon_{\text{new}}^* - \varepsilon)} \text{trace}([A_1 a \cdots A_m a]^T [A_1 a \cdots A_m a])} \\
 & = \sigma_{\min}(A) - \sqrt{\sup_{\|a\| < 2(\varepsilon_{\text{new}}^* - \varepsilon)} a^T \left(\sum_{i=1}^m A_i^T A_i \right) a} \\
 & = \sigma_{\min}(A) - 2(\varepsilon_{\text{new}}^* - \varepsilon) \sqrt{\lambda_{\max} \left(\sum_{i=1}^m A_i^T A_i \right)} \\
 & = \sigma_{\min}(A) - 2(\varepsilon_{\text{new}}^* - \varepsilon) L_{\text{new}} \\
 & = 2\varepsilon L_{\text{new}},
 \end{aligned} \tag{2.6}$$

where the first inequality is Weyl's inequality [10] for the singular values, see also Problem III.6.5 in [2] or Theorem 3.3.16 in [11].

Since the optimal value of the maximizing problem (2.6) is unattainable, the above inequality implies that

$$\sigma_{\min}(A + [A_1 a \cdots A_m a]) > 2\varepsilon L_{\text{new}}, \quad \forall a \in \mathbb{R}^n : \|a\| < 2(\varepsilon_{\text{new}}^* - \varepsilon). \tag{2.7}$$

Notice that

$$f_i(x) = f_i(a) + (A_i a + a_i)^T (x - a) + \frac{1}{2} (x - a)^T A_i (x - a), \quad i = 1, \dots, m.$$

Then, we have

$$F_m(\varepsilon, a) - f(a) = \{g(y) : y \in \mathbb{R}^n, \|y\| \leq \varepsilon\} := G_m(\varepsilon, a),$$

where $g(y) = ((A_1 a + a_1)^T y + \frac{1}{2} y^T A_1 y, \dots, (A_m a + a_m)^T y + \frac{1}{2} y^T A_m y)$. According to Theorem 2.1, for any

$$\varepsilon < \sigma_{\min}(A + [A_1 a \cdots A_m a]) / (2L_{\text{new}}), \tag{2.8}$$

the image $G_m(\varepsilon, a)$ is a convex set in $\mathbb{R}^m$. The proof is complete as (2.8) is ensured by (2.7). $\square$

Remark 2.3 Theorem 2.1 is a special case of Theorem 2.2 by setting $a = 0$.

3 Further Possible Improvements

The estimation of Theorem 2.1 is still not tight. Actually, L_{new} defined in (2.1) can be further improved to be the Lipschitz constant of f , denoted by L_f . According to (2.4), we have

$$L_f^2 = \max_{\|y\|=1} \lambda_{\max} \left([A_1 y \cdots A_m y]^T [A_1 y \cdots A_m y] \right). \quad (3.1)$$

However, this is a nonlinear eigenvalue optimization problem and not easy to solve. As pointed by one referee, the problem (3.1) can be reduced to the largest singular value of a third order tensor [13], which is generally NP-hard [9].

Except for the upper bound L_{new} (2.1), we further consider the other two relaxations of (3.1). We first need two lemmas.

Lemma 3.1 [2] *Every eigenvalue of $B \in \mathbb{R}^{m \times m}$ lies within at least one of the Gershgorin disks*

$$\left\{ \lambda : |\lambda - B_{ii}| \leq \sum_{j \neq i} |B_{ij}| \right\}, i = 1, \dots, m.$$

Lemma 3.2 [7] *For any $m \times m$ matrix B , all its eigenvalues are located in the same disk*

$$\left| \lambda - \frac{\text{trace}(B)}{m} \right| \leq \sqrt{\frac{m-1}{m} \left(\text{trace}(B^T B) - \frac{(\text{trace}(B))^2}{m} \right)}. \quad (3.2)$$

Remark 3.3 Let $\lambda_i(B)$ be the i -th largest eigenvalue of B . When $B \succeq 0$, substituting the following inequality

$$\text{trace}(B^T B) = \sum_{i=1}^m \lambda_i^2(B) \leq \left(\sum_{i=1}^m \lambda_i(B) \right)^2 = (\text{trace}(B))^2$$

into (3.2), we see that Lemma 3.2 improves the inequality

$$\lambda_{\max}(B) \leq \text{trace}(B),$$

which is used in (2.5).

Now, we apply Lemmas 3.1 and 3.2 to establish two new relaxations of L_f (3.1).

Firstly, according to Lemma 3.1, we have

$$\begin{aligned}
 & \sqrt{\max_{\|y\|=1} \lambda_{\max} \left([A_1 y \cdots A_m y]^T [A_1 y \cdots A_m y] \right)} \\
 & \leq \sqrt{\max_{\|y\|=1} \max_{i=1, \dots, m} \left\{ y^T (A_i^T A_i) y + \sum_{j \neq i} y^T |A_i^T A_j| y \right\}} \\
 & = \sqrt{\max_{i=1, \dots, m} \max_{\|y\|=1} y^T \left(A_i^T A_i + \sum_{j \neq i} |A_i^T A_j| \right) y} \\
 & = \sqrt{\max_{i=1, \dots, m} \lambda_{\max} \left(A_i^T A_i + \frac{1}{2} \sum_{j \neq i} (|A_i^T A_j| + |A_j^T A_i|) \right)} \\
 & := \bar{L}_{\text{new}}.
 \end{aligned}$$

Consequently, Theorem 1.1 holds true if we replace L with $\bar{L}_{\text{new}}$.

Secondly, according to Lemma 3.2, we have

$$\begin{aligned}
 & \lambda_{\max} \left([A_1 y \cdots A_m y]^T [A_1 y \cdots A_m y] \right) \\
 & \leq \frac{1}{m} \sqrt{\left(y^T \left(\sum_{i=1}^m A_i^T A_i \right) y \right)^2} \\
 & + \sqrt{\frac{m-1}{m} \left(\sum_{i,j=1}^m (y^T A_i^T A_j y)^2 - \frac{1}{m} \left(y^T \left(\sum_{i=1}^m A_i^T A_i \right) y \right)^2 \right)} \\
 & = \frac{1}{m} \sqrt{z^T \left(\left(\sum_{i=1}^m A_i^T A_i \right) \otimes \left(\sum_{i=1}^m A_i^T A_i \right) \right) z} + \sqrt{\frac{m-1}{m}} \\
 & \sqrt{z^T \left(\sum_{i,j=1}^m (A_i^T A_j) \otimes (A_i^T A_j) - \frac{1}{m} \left(\sum_{i=1}^m A_i^T A_i \right) \otimes \left(\sum_{i=1}^m A_i^T A_i \right) \right) z} \\
 & = \frac{1}{m} \sqrt{\left(\left(\sum_{i=1}^m A_i^T A_i \right) \otimes \left(\sum_{i=1}^m A_i^T A_i \right) \right) \bullet Z} + \sqrt{\frac{m-1}{m}} \\
 & \sqrt{\left(\sum_{i,j=1}^m (A_i^T A_j) \otimes (A_i^T A_j) - \frac{1}{m} \left(\sum_{i=1}^m A_i^T A_i \right) \otimes \left(\sum_{i=1}^m A_i^T A_i \right) \right) \bullet Z} \\
 & := B(Z)
 \end{aligned}$$

where $z = y \otimes y$ and $Z = zz^T$. Since $y^T y = 1$, we have

$$\text{trace}(Z) = z^T z = (y \otimes y)^T (y \otimes y) = (y^T y) \otimes (y^T y) = 1 \otimes 1 = 1,$$

$$\text{vec}(I)^T Z \text{vec}(I) = \left(\text{vec}(I)^T z \right)^2 = \left(\sum_{i=1}^m y_i^2 \right)^2 = 1,$$

$$\|Z \text{vec}(I)\| = \|z z^T \text{vec}(I)\| = |\text{vec}(I)^T z| \|z\| = \|z\| = \sqrt{z^T z} = 1,$$

$$Z = z z^T \succeq 0.$$

Therefore, Theorem 1.1 remains true if L is replaced by $\tilde{L}_{\text{new}}$, where

$$\begin{aligned} \tilde{L}_{\text{new}}^2 = \max \quad & B(Z) \\ \text{s.t.} \quad & \text{trace}(Z) = 1, \\ & \text{vec}(I)^T Z \text{vec}(I) = 1, \\ & \|Z \text{vec}(I)\| \leq 1, \\ & Z \succeq 0, \end{aligned}$$

which is a convex semidefinite programming (CSDP) problem, and hence can be efficiently solved. In the following examples, the CSDP problems are modeled by CVX 1.2 [6] and solved by SDPT3 [20] within CVX.

Example 3.4 Let $n = 3$, $m = 2$. Consider the two examples:

$$(E_1) :$$

$$(E_2) :$$

We can verify that

$$(E_1) :$$

$$(E_2) :$$

It is observed that neither L_{new} nor $\bar{L}_{\text{new}}$ dominates each other. Moreover, both are dominated by $\tilde{L}_{\text{new}}$.

Figure 1 shows the images of the ε -disks for (E_1) and (E_2) , respectively. It follows that $\tilde{L}_{\text{new}}$ is not tight and the convexity loses when ε is large enough.

Fig. 1 Images of ε -disks for (E_1) with $\varepsilon = 1/(2\tilde{L}_{\text{new}}) \approx 0.0394$, 0.06 , 0.14 in the left subgraph and for (E_2) with $\varepsilon = 1/(2\tilde{L}_{\text{new}}) \approx 0.0362$, 0.04 , 0.08 in the right subgraph

4 Concluding Remarks

In this paper, we improve and extend Polyak's local convexity result for quadratic transformations by providing tighter bounds for

$$\max_{\|y\|=1} \lambda_{\max} \left([A_1 y \cdots A_m y]^T [A_1 y \cdots A_m y] \right).$$

It is open whether the above nonlinear eigenvalue optimization problem can be efficiently globally solved. Moreover, we propose a convex semidefinite programming (CSDP) relaxation, which is conjectured to be the tightest among all existing upper bounds as we are unable to find a counterexample. Since the above nonlinear eigenvalue optimization problem can be reformulated as the following polynomial optimization problem

$$\max_{\|z\|=1, \|y\|=1} z^T \left([A_1 y \cdots A_m y]^T [A_1 y \cdots A_m y] \right) z,$$

it is the future work to study upper bounds on this polynomial optimization problem, for example, applying the SOS (sum of squares) method of Lasserre's SDP Hierarchy [12].

Acknowledgments The author is grateful to the anonymous referee whose comments improved this paper.

References

- [1] Beck, A.: On the convexity of a class of quadratic mappings and its application to the problem of finding the smallest ball enclosing a given intersection of balls. *J. Glob. Optim.* **39**(1), 113–126 (2007)
- [2] Bhatia, R.: *Matrix Analysis*. Springer-Verlag, New York (1997)
- [3] Brickman, L.: On the field of values of a matrix. *Proc. AMS* **12**, 61–66 (1961)
- [4] Dines, L.L.: On the mapping of quadratic forms. *Bull. AMS* **47**, 494–498 (1941)
- [5] Fradkov, A.L.: Duality theorems for certain nonconvex extremum problems. *Sib. Math. J.* **14**, 247–264 (1973)
- [6] M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 1.21 (2010). <http://cvxr.com/cvx>
- [7] Gu, Y.: The distribution of eigenvalues of a matrix. *Acta Math. Appl. Sin.* **17**(4), 501–511 (1994)
- [8] Hausdorff, F.: Der Wervorrat einer Bilinearform. *Mathematische Zeitschrift* **3**, 314–316 (1919)
- [9] He, S., Li, Z., Zhang, S.: Approximation algorithms for homogeneous polynomial optimization with quadratic constraints. *Math. Program. Ser. B* **125**, 353–383 (2010)
- [10] Horn, R.A., Johnson, C.R.: *Matrix Analysis*. Cambridge University Press, Cambridge (1985)
- [11] Horn, R.A., Johnson, C.R.: *Topics in Matrix Analysis*. Cambridge University Press, Cambridge (1991)
- [12] Lasserre, J.B.: Global optimization with polynomials and the problem of moments. *SIAM J. Optim.* **11**, 796–817 (2001)
- [13] Lim, L.-H.: Singular values and eigenvalues of tensors: a variational approach. In: *Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, vol. 1, pp. 129–132 (2005)
- [14] Polyak, B.T.: Convexity of quadratic transformations and its use in control and optimization. *J. Optim. Theory Appl.* **99**, 553–583 (1998)
- [15] Polyak, B.T.: Convexity of nonlinear image of a small ball with applications to optimization. *Set-Valued Anal.* **9**, 159–168 (2001)

- [16] Polyak, B.T.: The convexity principle and its applications. *Bull. Br. az. Math.Soc. (N.S.)* **34**(1), 59–75 (2003)
- [17] Pólik, I., Terlaky, T.: A survey of S-lemma. *SIAM Rev.* **49**(3), 371–418 (2007)
- [18] Ramana, M., Goldman, A.J.: Quadratic maps with convex images, Report 36-94, Rutgers Center for Operations Research, Rutgers, The State University of New Jersey (1994)
- [19] Toeplitz, O.: Das algebraische Analogen zu einem Satz von Fejer. *Mathematische Zeitschrift* **2**, 187–197 (1918)
- [20] Toh, K.C., Todd, M.J., Tutuncu, R.H.: SDPT3—a Matlab software package for semidefinite programming. *Optim. Methods Softw.* **11**, 545–581 (1999)
- [21] Uderzo, A.: On the Polyak convexity principle and its application to variational analysis. *Nonlinear Anal.* **91**, 60–71 (2013)
- [22] Y. Xia, S. Wang, R.L. Sheu, S-Lemma with equality and its applications. [arXiv:1403.2816v2](https://arxiv.org/abs/1403.2816v2) (2014) <http://arxiv.org/abs/1403.2816>
- [23] Yakubovich, V.A.: S-procedure in nonlinear control theory. *Vestnik Leningrad. Univ.* **1**, 62–77 (1971). (in Russian)
- [24] Yakubovich, V.A.: S-procedure in nonlinear control theory, *Vestnik Leningrad. Univ.* **4**, 73–93 (1977) (English translation)