The extensions of Yuan's lemma and applications in S-lemma#

Qingzhi Yang*, Yang Zhou*, and Zhongwen Wang**

Abstract

In this paper we extend a lemma due to Yuan from several aspects. A new proof of Yuan's lemma is given. A rank-one decomposition of positive semidefinite matrix is further developed. With the extended rank-one decomposition results, we generalize the Yuan's lemma to general quadratic function systems, interval quadratic function systems and quadratic matrix function systems. Based on them, we offer several new proofs of S-lemma on quadratic functions systems, and establish a new S-lemma on quadratic matrix inequality functions systems and provide a simple proof of the strong duality of a class of quadratic matrix programming.

keywords: Yuan's lemma, S-lemma, matrix rank-one decomposition, Matrix quadratic programming, strong duality.

1.Introduction

The S-lemma is an alternative proposition on two quadratic functions systems and it plays an important role in Control theory, nonconvex quadratic optimization, numerical range, error estimations, and so on. Yakubovich[9] first presented an S-lemma on quadratic inequality system. There has been several different proofs on the fundamental S-lemma. For example, in [2], three distinct proofs are given for the S-lemma on quadratic inequality systems,

*School of Mathematical Sciences and LPMC, Nankai University.

**School of Mathematics and Statistics, Kashgar University.

#This work was supported by the National Natural Science Fund of China(Grant No. 11671217) and Natural Science Fund of Xinjiang(Grant No. 2017D01A14)

one among them is based on a lemma due to Yuan and is called "elementary proof ". The Yuan's lemma is on homogeneous quadratic function systems [12][2]. In this paper, we develop this lemma from several aspects, then offer several new proofs of several distinct S-lemmas based on the extended Yuan's lemma. In detail, we make the following contributions in the current paper: 1. provide a simple proof of Yuan's lemma and meanwhile relax the conditions in Yuan's lemma; 2. extend a matrix rank-one decomposition due to Zhang et al [11][3][10], and put forward an extended Yuan's lemma for general quadratic inequality systems, from which we deduce the S-lemma on quadratic inequality systems[2]; 3. prove the S-lemma of interval quadratic inequality system, the proof is simpler than that in [7]; 4. use the S-lemma in Item 3, deduce an S-lemma to the quadratic functions systems with equality constraint; 5. develop a fundamental result about symmetric matrix to board area, then present an S-lemma on quadratic matrix inequality functions systems. As an application of the new S-lemma, we offer a new proof of strong duality of a class of quadratic matrix programming.
This paper is organized as follows: In Section 2, we extend Yuan's lemma to general quadratic inequality system; In Section 3, we give a new proof of the S-lemma on interval quadratic system; In Section 4, we propose a weak S-lemma with equality constraint; In Section 5, we offer an S-lemma on quadratic matrix functions systems, and as an application of it, we give a simpler proof of strong duality of a class of quadratic matrix programming. We conclude this paper with some remarks in Section 6.

2.Yuan's lemma and its generalization to general quadratic system

To begin with, we recall two basic results.
Lemma 2.1 [Appendix B, [4]]: Let $A, B \in S^n$,$X \succeq 0$. Then there exists an $x \in R^n$ such that

$$A \circ X = x^T Ax \qquad B \circ X = x^T Bx$$

Lemma 2.2([2] generalized Farkas Lemma): Let $f(x), g_i(x), i = 1, \cdots$,m be the convex functions in R^n. And there exists an $\bar{x} \in R^n$ such that $g_i(\bar{x}) < 0, i = 1, \cdots, m$. $C \in R^n$ is a convex set. Then the system

$$f(x) < 0 \qquad g_i(x) \leq 0 \qquad i = 1, \cdots, m$$

is not solvable if and only if there exist $\lambda_i \geq 0, i = 1, \cdots, m$ such that

$$f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) \geq 0 \qquad \forall x \in C$$

Remark: It is easy to know that if we replace R^n by symmetric matrix space S^n, the statement is also valid.

In [12], author stated the following lemma.
Lemma 2.3(Yuan's lemma): Let $A, B \in S^n$,E, F are two closed sets in R^n such that $E \cup F = R^n$ and

$$x^T Ax \geq 0 \quad x \in E; \qquad x^T Bx \geq 0 \quad x \in F$$

Then there exists a $\lambda \in [0, 1]$ such that $\lambda A + (1 - \lambda)B \succeq 0$.

The original proof of this lemma is complex a little bit ([12][2]). In the following, we provide an easy-understanding proof of the above lemma and meanwhile relax the condition on two sets in the lemma.
Lemma 2.4: Let $A, B \in S^n$,E, F are two sets in R^n such that $E \cup F = R^n$ and

$$x^T Ax \geq 0 \quad x \in E; \qquad x^T Bx \geq 0 \quad x \in F$$

Then there exists a $\lambda \in [0, 1]$ such that $\lambda A + (1 - \lambda)B \succeq 0$.

Proof: Without loss of generality, we assume that B is not positive semidefinite. Then the system

$$A \circ X < 0, \quad B \circ X < 0, \quad X \succeq 0 \tag{1}$$

is not solvable. If the system (1) is solvable, then from Lemma 2.1, there exists an $x \in R^n$ such that

$$x^T Ax = A \circ X < 0, \qquad x^T Bx = B \circ X < 0 \tag{2}$$

Because of $E \cup F = R^n$, we must have $x \in E$ or $x \in F$. From the assumption conditions, we can get $x^T Ax \geq 0$ or $x^T Bx \geq 0$, which contradicts (2). So system (1) is not solvable. Hence for any $\epsilon > 0$, system

$$A \circ X < 0, \quad B \circ X + \epsilon \leq 0, \quad X \succeq 0$$

is not solvable. From Lemma 2.2, there is a $\lambda_\epsilon \geq 0$ such that

$$A \circ X + \lambda_\epsilon (B \circ X + \epsilon) \geq 0 \qquad \forall X \succeq 0$$

which is

$$\frac{1}{1+\lambda_\epsilon} A \circ X + \frac{\lambda_\epsilon}{1+\lambda_\epsilon}(B \circ X + \epsilon) \geq 0 \qquad \forall X \succeq 0 \tag{3}$$

One may assume that $\frac{1}{1+\lambda_\epsilon} \to \lambda(\epsilon \to 0^+)$, then $\lambda \in [0,1]$. In (3), let $\epsilon \to 0^+$, then

$$\lambda A \circ X + (1-\lambda) B \circ X \geq 0, \qquad \forall X \succeq 0$$

which yields $\lambda A + (1-\lambda)B \succeq 0$.

We may get a revised version of above lemma as following. The proof is omitted since it may be done similar to previous proof.

Lemma 2.5: Let $A, B \in S^n$,E, F are two sets in R^n such that $E \cup F = R^n$ and

$$x^T A x \geq 0 \quad x \in E; \qquad x^T B x > 0 \quad x \in F$$

Then there exists a $\lambda \geq 0$ such that $A + \lambda B \succeq 0$.

In the next, we generalize the Yuan's lemma to nonhomogeneous quadratic function system.

Theorem 2.1: Let $A_1, A_2 \in S^n$,$b_1, b_2 \in R^n$,$E_1, E_2 \subseteq R^n$ and $E_1 \cup E_2 = R^n$. If

$$\begin{array}{ll} x^T A_1 x + 2b_1^T x + c_1 \geq 0, & x \in E_1 \\ x^T A_2 x + 2b_2^T x + c_2 \geq 0, & x \in E_2 \end{array} \tag{4}$$

then there exists a $\lambda \in [0,1]$ such that

$$\lambda(A_1 \circ X + 2b_1^T x + c_1) + (1-\lambda)(A_2 \circ X + 2b_2^T x + c_2) \geq 0, \forall x \in R^n, X - xx^T \succeq 0.$$

For proving this theorem, some preparations are needed. First we recall a result of matrix rank-one decomposition as following.

Lemma 2.6[11]: Let $A \in S^n, X \succeq 0, r(X) = r$, then there exists a decomposition of X: $X = \sum_{i=1}^{r} p_i p_i^T$, where$\{p_i\}$ are linear independent such that

$$p_i^T A p_i = \frac{A \circ X}{r} \qquad i = 1, \cdots, r$$

specially, if $A \circ X \leq (<) 0$, then

$$p_i^T A p_i \leq (<) 0 \qquad i = 1, \cdots, r$$

Based on this lemma, one may get the following results.

Theorem 2.2: Let

$$\begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix} \circ Z < 0, \quad Z \in S_+^{n+1}, \quad Z_{n+1,n+1} = 1$$

then there exists a decomposition of Z:
$Z = \sum_{i=1}^{r} p^i {p^i}^T, r = r(Z) \geq 2, p_{n+1}^i \neq 0$ such that

$${p^i}^T \begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix} p^i < 0, \quad i = 1, \cdots, r$$

Proof: Using induction to show. If $r = 2$, from Lemma 2.6, there exist η^1, η^2 such that $Z = \eta^1 {\eta^1}^T + \eta^2 {\eta^2}^T$ and

$${\eta^i}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \eta^i < 0, \quad i = 1, 2$$

Obviously $(\eta_{n+1}^1)^2 + (\eta_{n+1}^2)^2 = 1$. If $\eta_{n+1}^1 \eta_{n+1}^2 \neq 0$. The statement is valid. If not, assume that $\eta_{n+1}^1 = 0$, then $(\eta_{n+1}^2)^2 = 1$, denote $y^1 = \eta^2 + u\eta^1$, because

$${\eta^2}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \eta^2 < 0$$

Denote $y^2 = \eta^1 - u\eta^2$, due to the continuity of the functions, then when $|u| > 0$ is sufficiently small, there will be

$${y^1}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} y^1 < 0, \quad {y^2}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} y^2 < 0$$

Furthermore,

$$\begin{aligned}
&y^1{y^1}^T + y^2{y^2}^T \\
=&\eta^2{\eta^2}^T + u(\eta^2{\eta^1}^T + \eta^1{\eta^2}^T) + u^2\eta^1{\eta^1}^T \\
&+ \eta^1{\eta^1}^T - u(\eta^1{\eta^2}^T + \eta^2{\eta^1}^T) + u^2\eta^2{\eta^2}^T \\
=&(1+u^2)(\eta^1{\eta^1}^T + \eta^2{\eta^2}^T) \\
=&(1+u^2)Z
\end{aligned}$$

so $Z = (y^1{y^1}^T + y^2{y^2}^T)/(1+u^2)$. Denote $z^i = y^i/\sqrt{1+u^2}, i = 1, 2$, then $Z = z^1{z^1}^T + z^2{z^2}^T$ and

$${z^i}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} z^i < 0, \quad i = 1, 2$$

From the definition of z^i, we know $z^i_{n+1} \neq 0, i = 1, 2$.
Assume that when $r(Z) = m < k$, the statement is valid. Then we have $Z = \sum_{i=1}^m \eta^i{\eta^i}^T$.

$${\eta^i}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \eta^i < 0, \quad \eta^i_{n+1} \neq 0, \quad i = 1, \cdots, m$$

For the case of $r(Z) = m + 1 \leq k$. From Lemma 2.5, there exist a set of linear independent η^i such that $Z = \sum_{i=1}^{m+1} \eta^i{\eta^i}^T$, and

$${\eta^i}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \eta^i < 0, \quad i = 1, \cdots, m+1$$

Obviously, if $\eta^i_{n+1} \neq 0, i = 1, \cdots, m+1$, the statement is valid. Otherwise, at least there exists an i such that $\eta^i_{n+1} = 0$. Assume that $\eta^1_{n+1} = 0, \eta^2_{n+1} \neq 0$. Analogy with the proof of $r = 2$ case, there exist z^1, z^2 such that $z^1{z^1}^T + z^2{z^2}^T = \eta^1{\eta^1}^T + \eta^2{\eta^2}^T$ and

$${z^i}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} z^i < 0, \quad z^i_{n+1} \neq 0, \quad i = 1, 2$$

So $Z = z^1{z^1}^T + z^2{z^2}^T + \sum_{i=3}^{m+1} \eta^1{\eta^1}^T$. Repeat this procedure at most m times, we arrive at that proposition is true.

Theorem 2.3: Assume the Slater condition holds for a single inequality system

$$x^T Ax + 2b^T x + c \le 0$$

then

$$\begin{aligned} & conv\{\begin{bmatrix} x \\ 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}^T \mid \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} < 0, x \in R^n\} \\ =& \{Z \in S_+^{n+1} \mid \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \circ Z < 0, Z_{n+1,n+1} = 1\} \triangleq C^o \end{aligned}$$

$$\begin{aligned} & conv\{\begin{bmatrix} x \\ 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}^T \mid \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \le 0, x \in R^n\} \\ =& \{Z \in S_+^{n+1} \mid \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \circ Z \le 0, Z_{n+1,n+1} = 1\} \triangleq C \end{aligned}$$

Proof: To begin with, we establish the first relationship. Obviously, both sets are nonempty and the left set is contained in the right one. So we only need to prove the converse. For a given $Z \in C^o$, assume that $r(Z) = k$, from Theorem 2.2, there exist η^i such that $Z = \sum_{i=1}^k \eta^i {\eta^i}^T$

$${\eta^i}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \eta^i < 0, \eta^i_{n+1} \neq 0, \quad i = 1, \cdots, k$$

Denote $z^i = \eta^i / \eta^i_{n+1}$, then $z^i_{n+1} = 1, i = 1, \cdots, k$

$${z^i}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} z^i < 0$$

and $Z = \sum_{i=1}^k (\eta^i_{n+1})^2 z^i {z^i}^T$. From $z_{n+1,n+1} = 1$, we know $\sum_{i=1}^k (\eta^i_{n+1})^2 = 1$, so

$$Z \in conv\{\begin{bmatrix} x \\ 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}^T \mid \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} < 0, x \in R^n\}$$

Next we prove the latter inclusion relation. Also we just need to prove the right set is contained in the left one. Suppose that $\bar{Z} \in C$, then there exists a sequence $\{Z^j\}$, $Z^j \in C^o$ such that $Z^j \to \bar{Z}(j \to \infty)$. Because

$$Z^j \in conv\{\begin{bmatrix} x \\ 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}^T \mid \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} < 0, x \in R^n\}$$

Let $j \to \infty$, then

$$\bar{Z} \in conv\{\begin{bmatrix} x \\ 1 \end{bmatrix}\begin{bmatrix} x \\ 1 \end{bmatrix}^T \mid \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix}\begin{bmatrix} x \\ 1 \end{bmatrix} \leq 0, x \in R^n\}$$

Now we are in the place of proving the Theorem 2.1.
Proof of Theorem 2.1: Without loss of generality, assume that there exists a $Z \succeq 0, Z_{n+1,n+1} = 1$ such that

$$\begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \circ Z < 0$$

Clearly (4) may be equally written as

$$\begin{aligned} \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix}\begin{bmatrix} x \\ 1 \end{bmatrix} &\geq 0, \qquad x \in E_1 \\ \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix}\begin{bmatrix} x \\ 1 \end{bmatrix} &\geq 0, \qquad x \in E_2 \end{aligned} \tag{5}$$

From (5), we can know the following system

$$\begin{aligned} &\begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix} \circ Z < 0 \\ &\begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \circ Z < 0 \\ &Z \succeq 0, Z_{n+1,n+1} = 1 \end{aligned} \tag{6}$$

is not solvable. If not, assume that (6) have a solution Z, then from Theorem 2.2 there exist a decomposition of Z, such that $Z = \sum_{i=1}^r p_i p_i^T$, $p_{n+1}^i \neq 0, i = 1, \cdots, r$ and

$$p^{iT} \begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix} p^i < 0, \qquad i = 1, \cdots, r$$

Meanwhile,

$$\begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \circ Z = \sum_{i=1}^r p^{iT} \begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} p^i < 0$$

which means that there is a p^i, says p^1, such that

$$p^{1T}\begin{bmatrix}A_2 & b_2\\ b_2^T & c_2\end{bmatrix}p^1 < 0$$

Denote

$$\frac{p^1}{p^1_{n+1}} = \begin{bmatrix}\bar{p}\\ 1\end{bmatrix}$$

then

$$\begin{bmatrix}\bar{p}\\ 1\end{bmatrix}^T\begin{bmatrix}A_1 & b_1\\ b_1^T & c_1\end{bmatrix}\begin{bmatrix}\bar{p}\\ 1\end{bmatrix} < 0, \qquad \begin{bmatrix}\bar{p}\\ 1\end{bmatrix}^T\begin{bmatrix}A_2 & b_2\\ b_2^T & c_2\end{bmatrix}\begin{bmatrix}\bar{p}\\ 1\end{bmatrix} < 0$$

It contradicts (5). So (6) is not solvable. Then for any $\epsilon > 0$

$$\begin{aligned}\begin{bmatrix}A_1 & b_1\\ b_1^T & c_1\end{bmatrix}\circ Z < 0\\ \begin{bmatrix}A_2 & b_2\\ b_2^T & c_2\end{bmatrix}\circ Z + \epsilon \le 0\\ Z \succeq 0, Z_{n+1,n+1} = 1\end{aligned}$$

is not solvable. Denote $Q = \{Z \in S_+^{n+1}, Z_{n+1,n+1} = 1\}$, obviously Q is a convex set. From Lemma 2.2 and the proof of Lemma 2.4, there exists a $\lambda \in [0,1]$ such that

$$\lambda\begin{bmatrix}A_1 & b_1\\ b_1^T & c_1\end{bmatrix}\circ Z + (1-\lambda)\begin{bmatrix}A_2 & b_2\\ b_2^T & c_2\end{bmatrix}\circ Z \ge 0. \qquad \forall Z \in Q$$

Denote

$$Z = \begin{bmatrix}X & x\\ x^T & 1\end{bmatrix}$$

from Schur complement, we have $Z \in Q \Leftrightarrow X - xx^T \succeq 0$, then

$$\lambda(A_1\circ X + 2b_1^Tx + c_1) + (1-\lambda)(A_2\circ X + 2b_2^Tx + c_2) \ge 0, \forall x \in R^n, X - xx^T \succeq 0.$$

Remark: when $b_i = 0, c_i = 0, i = 1,2$, Theorem 2.1 reduces to Yuan's lemma.

Similar to Lemma 2.5, we can get a revised version of Theorem 2.1 as following.

Theorem 2.4: Let $A_1, A_2 \in S^n$,$b_1, b_2 \in R^n$,$E_1, E_2 \subseteq R^n$ and $E_1 \cup E_2 = R^n$. If

$$\begin{aligned} x^T A_1 x + 2b_1^T x + c_1 \geq 0, \qquad & x \in E_1 \\ x^T A_2 x + 2b_2^T x + c_2 > 0, \qquad & x \in E_2 \end{aligned} \tag{7}$$

then there exists a $\lambda \geq 0$ such that

$$(A_1 \circ X + 2b_1^T x + c_1) + \lambda(A_2 \circ X + 2b_2^T x + c_2) \geq 0, \forall x \in R^n, X - xx^T \succeq 0.$$

In [2], authors proved the fundamental S-lemma on homogeneous quadratic inequality system using Yuan's lemma and called the proof "an elementary proof ". They did not provide the proof of S-lemma of general quadratic inequality system, though they pointed out that it can be showed with analogous way of proof for homogeneous situation. In the following, we provide a simple proof of the fundamental S-lemma based on Theorem 2.1.

Theorem 2.5 (fundamental S-lemma): If $f(x) < 0, g(x) \leq 0$ is not solvable, then there exists a $\lambda \geq 0$ such that

$$f(x) + \lambda g(x) \geq 0, \qquad \forall x \in R^n$$

where $f(x), g(x)$ are quadratic functions and there exists an $\bar{x} \in R^n$ such that $g(\bar{x}) < 0$.

Proof: Denote $E = \{x : g(x) < 0\}, F = R^n \setminus E$, then from the assumption conditions one has that the system

$$f(x) < 0, \quad g(x) < 0$$

is unsolvable. This means

$$\begin{cases} f(x) \geq 0, & x \in E \\ g(x) \geq 0, & x \in F \end{cases}$$

From Theorem 2.1, there exists a $\delta \in [0, 1]$ such that

$$\delta(A_1 \circ X + 2b_1^T x + c_1) + (1 - \delta)(A_2 \circ X + 2b_2^T x + c_2) \geq 0, \forall x \in R^n, X - xx^T \succeq 0.$$

Let $X = xx^T$, then we have

$$\delta f(x) + (1-\delta)g(x) \geq 0, \quad \forall x$$

If $\delta = 0$, then $g(x) \geq 0, \forall x$, a contradiction! So $\delta \neq 0$, then

$$f(x) + \lambda g(x) \geq 0, \quad \forall x \in R^n$$

where $\lambda \geq 0$.

3. New proof of S-lemma on interval quadratic system

In this section, we discuss the S-lemma on interval quadratic system. In [7], with the S-lemma on equality system presented in [13], Wang and Xia gave a new proof of the strong duality of quadratic programming with interval constraints. Though their proof is easier than that in [8], the S-lemma involved itself is rather complex. Next we first extend the Yuan's lemma to the two sides quadratic system, then offer a new proof of S-lemma on interval quadratic system. We discuss the homogeneous and nonhomogeneous situations separately, since we use different approaches for them.

Theorem 3.1: Let $A, B \in S^n, l < u, E, F \subseteq R^n, E \cup F = R^n$

$$\begin{aligned} x^T Ax \geq 0, &\qquad x \in E \\ x^T Bx > u \quad or \quad x^T Bx < l, &\qquad x \in F \end{aligned}$$

Assume that there exists an $\bar{X} \succeq 0$ such that $l < B \circ X < u$, then there exists a $\lambda \geq 0$ such that

$$A \circ X + \lambda max\{B \circ X - u, l - B \circ X\} \geq 0, \qquad \forall X \succeq 0$$

Proof: The assumed condition means the system

$$A \circ X < 0, \quad max\{B \circ X - u, l - B \circ X\} \leq 0, \quad X \succeq 0 \tag{8}$$

is not solvable. If not, from Lemma 2.1, there exists an $x \in R^n$ such that

$$x^T Ax = A \circ X < 0, \qquad x^T Bx = B \circ X$$

so $max\{B \circ X - u, l - B \circ X\} \leq 0$. Because $E \cup F = R^n$, if $x \in E$, from the condition of theorem, $x^T Ax \geq 0$, conflict! If $x \in F$, we can get $x^T Bx > u$ or

$x^T Bx < l$. So $max\{B \circ X - u, l - B \circ X\} > 0$, conflict!
Since system (8) has no solution and $max\{B \circ X - u, l - B \circ X\}$ is the convex function in S^n_+, from Lemma 2.2 we know that there exists a $\lambda \geq 0$ such that

$$A \circ X + \lambda max\{B \circ X - u, l - B \circ X\} \geq 0, \qquad \forall X \succeq 0.$$

This completes the proof.

The general interval quadratic function system means:

$$\begin{aligned} &\begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} < 0, \quad l \leq \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A_2 & b_2 \\ b_2^T & 0 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \leq u \\ &-\infty < l < u < +\infty \end{aligned} \tag{9}$$

Theorem 3.2: Let $E, F \subseteq R^n$,$E \cup F = R^n$.

$$\begin{aligned} f(x) &= x^T A_1 x + 2b_1^T x + c_1 = \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \\ g(x) &= x^T A_2 x + 2b_2^T x + c_2 = \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \\ l &< u \end{aligned}$$

if

$$f(x) \geq 0 \quad x \in E; \qquad max\{g(x) - u, l - g(x)\} > 0 \quad x \in F$$

and Slater condition is satisfied. Then there exists a $\lambda \geq 0$ such that

$$\begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix} \circ Z + \lambda max\{\begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \circ Z - u, l - \begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \circ Z\} \geq 0 \qquad \forall Z \in Q$$

Proof: From the assumption conditions, the system

$$\begin{aligned} &\begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix} \circ Z < 0 \\ &max\{\begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \circ Z - u, l - \begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \circ Z\} \leq 0 \\ &Z \succeq 0, \quad Z_{n+1,n+1} = 1 \end{aligned} \tag{10}$$

has no solution. If not, this will lead to a contradiction. Consider the following interval quadratic programming:

$$\begin{aligned} &min \;\; f(x) \\ &s.t \;\; l \le g(x) \le u \end{aligned} \tag{11}$$

Because the Slater condition holds, from [14] we know that the strong duality holds for (11). Since (10) is feasible, that implies that the optimal value of (11) is negative. So there exists an x such that $f(x) < 0$, $max\{g(x) - u, l - g(x)\} \le 0$. This contradicts the assumptions. So (10) is unsolvable. Since the Slater condition is satisfied. Denote $Q = \{Z \in S_+^{n+1}, Z_{n+1,n+1} = 1\}$, then from Lemma 2.2, there exists a $\lambda \ge 0$ such that

$$\begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix} \circ Z + \lambda max\{ \begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \circ Z - u, l - \begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \circ Z\} \ge 0 \qquad \forall Z \in Q$$

In the following, we provide a new proof of the S-lemma on the interval quadratic function system based on the previous theorem.

Theorem 3.3: Assume that there exists an x_0 such that $l < g(x_0) < u$. Then system (9) is unsolvable if and only if there is $\lambda \ge 0$ such that

$$f(x) + \lambda max\{g(x) - u, l - g(x)\} \ge 0, \qquad \forall x \in R^n$$

Proof: In Theorem 3.2, take $E = \{x \in R^n : max\{g(x) - u, l - g(x)\} \le 0\}$ and $F = R^n \setminus E$. Then one has

$$f(x) \ge 0 \quad x \in E; \qquad max\{g(x) - u, l - g(x)\} > 0 \quad x \in F$$

So from previous theorem we have that there is $\lambda \ge 0$ such that

$$\begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix} \circ Z + \lambda max\{ \begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \circ Z - u, l - \begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \circ Z\} \ge 0 \qquad \forall Z \in Q$$

For any $x \in R^n$, taking $Z = \begin{bmatrix} xx^T & x \\ x^T & 1 \end{bmatrix}$, then required conclusion follows.

4. S-lemma with equality constraint

Consider the following quadratic system with equality

$$\begin{aligned} (P_2) \qquad & f(x) = x^T Ax + 2a^T x + c_1 < 0 \\ & g(x) = x^T Bx + 2b^T x + c_2 = 0 \end{aligned}$$

In this section we present a weak S-lemma on this system as following.
Theorem 4.1(S-lemma with equality): Let $f, g : R^n \to R$ be quadratic functions and we assume that $g(x) = 0$ is solvable. Then the following two statements are equivalent:
(i) The system

$$f(x) < 0, \quad g(x) = 0$$

has no solution.
(ii) There exist a $\lambda \in [0, 1]$ and a multiplier δ(not constant) such that

$$\lambda f(x) + \delta g(x) \geq 0 \quad \forall x \in R^n$$

Proof: Obviously, (P_2) is not solvable means that

$$(P_3) \qquad \begin{array}{l} f(x) \geq 0 \\ g(x) = 0 \end{array}$$

is solvable. Then for any $\epsilon > 0$, the system

$$(P_3)_\epsilon \qquad \begin{array}{l} f(x) \geq 0 \\ -\epsilon \leq g(x) \leq \epsilon \end{array}$$

is solvable. Then from S-lemma with two sides restriction in the previous section, there exists a $\lambda_\epsilon \geq 0$ such that

$$f(x) + \lambda_\epsilon max\{g(x) - \epsilon, -\epsilon - g(x)\} \geq 0 \qquad \forall x \in R^n$$

then

$$\frac{1}{1+\lambda_\epsilon} f(x) + \frac{\lambda_\epsilon}{1+\lambda_\epsilon} max\{g(x) - \epsilon, -\epsilon - g(x)\} \geq 0 \qquad \forall x \in R^n$$

Assume that $\frac{1}{1+\lambda_\epsilon} \to \lambda \in [0, 1] (\epsilon \to 0^+)$. In the above inequality, let $\epsilon \to 0^+$, then we have

$$\lambda f(x) + (1 - \lambda) max\{g(x), -g(x)\} \geq 0 \qquad \forall x \in R^n$$

Denote

$$\delta = \begin{cases} 1 - \lambda, & g(x) \geq 0 \\ \lambda - 1, & g(x) < 0 \end{cases}$$

then the above equation is equal to

$$\lambda f(x) + \delta g(x) \geq 0 \qquad \forall x \in R^n$$

Remark: In [2], S-lemma with equality is stated. Our result needs weaker assumption than that in [2] and provides an interesting proof way. Of course our conclusion is weaker than that in [2]. In our weak S-lemma, the coefficient of $g(x)$ is dependent of x. So our S-lemma with equality may be regarded as special version in quadratic form of S-lemma in [5].

5. S-lemma on quadratic matrix functions system

This section will propose an S-lemma on quadratic inequality matrix functions system. Based on this S-lemma, we give a new proof of strong duality of a class of quadratic matrix programming. First we extend Lemma 2.1 to wider range. In this section, without loss of generality we assume $n > 2$.

Lemma 5.1: Let $A_i \in S^n, i = 1, \cdots, m$, $1 < m < n$. For any $X \in S^n_+$, there exists an $x \in R^{n\times(m-1)}$ such that

$$A_i \circ X = tr(x^T A_i x), \quad i = 1, \cdots, m$$

Proof: For any $X \in S^n_+$, denote

$$v = \begin{bmatrix} A_1 \circ X \\ \vdots \\ A_m \circ X \end{bmatrix}$$

if $v = 0$, denote $x = 0$, then $rank(x) = 0 \leq (m-1)$.
if $v \neq 0$, assume $v_m \neq 0$, denote $D \triangleq \frac{A_m}{v_m}$, then

$$\begin{aligned} &A_i \circ X = v_i, \qquad i = 1, \cdots, m-1 \\ &D \circ X = 1 \end{aligned}$$

then we have

$$(A_{m-1} - v_{m-1}D) \circ X = 0$$

According to Lemma 2.6, there is a rank-one decomposition of X, $X = \sum_{j=1}^{r} x_j x_j^T$ with $r = rank(X)$ such that

$$(A_{m-1} - v_{m-1}D) \circ x_j x_j^T = 0, \quad j = 1, \cdots, r$$

Define an $(m-1) \times r$ matrix $B = (b)_{ij}$ with entries given by

$$b_{ij} = (A_i - v_i D) \circ x_j x_j^T, \quad i = 1, \cdots, m-2, j = 1, \cdots, r$$
$$b_{m-1,j} = D \circ x_j x_j^T, \qquad j = 1, \cdots, r$$

and consider the following equation

$$Bt = e_{m-1}, t > 0 \tag{12}$$

where e_{m-1} is the $(m-1)$ dimensional unit vector whose last component is 1 and 0 elsewhere. By the construction of the matrix B we have $Be = e_{m-1}$ with e being the vector of all ones. Then there must exist some $\hat{t} \geq 0$ satisfying (12) such that the number of its components does not exceed $m-1$.

When $r \leq (m-1)$, we have $X = PP^T, P \in R^{r \times m}$, statement is valid.
When $r = m, B \in R^{(m-1)\times m}$, $By = 0$ have non-zero solution. Let $\bar{y}$ is a solution with positive component and assumed that $\bar{y}_1 = max\{\bar{y}_i\}$. Then we have $e - \frac{\bar{y}}{\bar{y}_1} \geq 0$. Denote $\hat{t} = e - \frac{\bar{y}}{\bar{y}_1}$, then $\hat{t} \geq 0, \hat{t}_1 = 0$ and the number of its components of $\hat{t}$ does not exceed $m - 1.Be = e_{m-1}$, so $B\hat{t} = e_{m-1}$. Denote

$$\hat{X} = \sum_{j=1}^{m} \hat{t}_j x_j x_j^T = \sum_{j=2}^{m} \hat{t}_j x_j x_j^T$$

then $\hat{X} \in S_+^n, r(\hat{X}) \leq (m-1)$, and

$$v = (A_1 \circ \hat{X}, \cdots, A_m \circ \hat{X})$$

then there exists a $P \in R^{n \times (m-1)}$ such that $\hat{X} = PP^T$. We have

$$(A_1 \circ \hat{X}, \cdots, A_m \circ \hat{X}) = (tr(P^T A_1 P), \cdots, tr(P^T A_m P))$$

When $r > m$, by induction, assume that $X = \bar{X} + yy^T$ where $\bar{X} \in S_+^n, rank(\bar{X}) = r(X) - 1$. For $\bar{X}$, there exists a $\bar{P} \in R^{n\times(m-1)}$ such that

$$A_i \circ \bar{X} = tr(\bar{P}^T A_i \bar{P}) \qquad i = 1, \cdots, m$$

so

$$A_i \circ X = A_i \circ (\bar{P}\bar{P}^T + yy^T) \qquad i = 1, \cdots, m$$

and $rank(\bar{P}\bar{P}^T + yy^T) \leq m$. From the case of $r \leq m$, there exists a $P \in R^{n\times(m-1)}$ such that

$$A_i \circ X = tr(P^T A_i P) \qquad i = 1, \cdots, m.$$

This completes the proof.

The following lemma gives the version of Yuan's lemma on homogeneous quadratic matrix function system.

Lemma 5.2: Let $A_i \in S^n, i = 1, \cdots, m, 1 < m \leq n, E_i \subseteq R^{n\times(m-1)}, i = 1, \cdots, m$ and $\cup_{i=1}^m E_i = R^{n\times(m-1)}$. If

$$tr(x^T A_i x) \geq 0, \quad x \in E_i, \quad i = 1, \cdots, m$$

then there exist $\lambda_i \geq 0, i = 1, \cdots, m$,$\sum_{i=1}^m \lambda_i = 1$ such that

$$\sum_{i=1}^m \lambda_i A_i \succeq 0$$

Proof: From the conditions we know the system

$$A_i \circ X < 0 \qquad X \succeq 0, \quad i = 1, \cdots, m \tag{13}$$

is unsolvable. If (13) have the solution X^*, from Lemma 5.1, there exists an $x^* \in R^{n\times(m-1)}$ such that

$$Tr(x^{*T} A_i x^*) = A_i \circ X^*, \quad i = 1, \cdots, m$$

Conflict. Analogy with proof of Lemma 2.4, there exist $\lambda_i \geq 0, i = 1, \cdots, m$, $\sum_{i=1}^m \lambda_i = 1$ such that

$$\sum_{i=1}^m \lambda_i A_i \succeq 0.$$

Like Lemma 2.5 or Theorem 2.4, in this lemma, if one inequality is strict, then the corresponding coefficient in conclusion is positive.

Based on Lemma 5.1, we present an S-lemma on quadratic inequality matrix function system as following.

Theorem 5.1: Assume $1 \leq r \leq n-1$. Let $x \in R^{n\times r}, f(x) = tr(x^T A x) +$

$2tr(a^Tx) + c, g_i(x) = tr(x^TB_ix) + 2tr(b_i^Tx) + c_i, i = 1, \cdots, r$. And there exists an $\bar{x} \in R^{n\times r}$ such that $g_i(\bar{x}) < 0, i = 1, \cdots, r$, where $A, B_i \in S^n, a, b_i \in R^{n\times r}, i = 1, \cdots, r$. If

$$f(x) < 0, \quad g_i(x) \le 0 \quad i = 1, \cdots, r$$

is not solvable, then there exist $\lambda_i \ge 0, i = 1, \cdots, r$ such that

$$\begin{bmatrix} A & a \\ a^T & \frac{c}{r}I \end{bmatrix} \circ Z + \sum_{i=1}^{r} \lambda_i \begin{bmatrix} B_i & b_i \\ b_i^T & \frac{c_i}{r}I \end{bmatrix} \circ Z \ge 0 \qquad \forall Z \in Q$$

where

$$Q = \{Z \in S_{n+r}^+ : Z_{n+1}^r = I_r\}$$

where Z_{n+1}^r denotes the lower-right $r \times r$ sub-matrix of Z .
Proof: Denote

$$H_f = \begin{bmatrix} A & a \\ a^T & \frac{c}{r}I \end{bmatrix}, \quad H_i = \begin{bmatrix} B_i & b_i \\ b_i^T & \frac{c_i}{r}I \end{bmatrix}$$

From the assumptions, one has the following system

$$H_f \circ Z < 0, \;\; H_i \circ Z \le 0, \;\; i = 1, ..., r, \;\; Z \succeq 0, \;\; Z_{n+1}^r = I_r \qquad (14)$$

has no solution. If not, assume Z^* is a solution of this system with rank k. Since Z^* is positive semidefinite, there is $P \in R^{(n+r)\times k}$ such that $Z^* = PP^T$. Because $Z_{n+1}^r = I_r$, then $k \ge r$. If $k = r$, denote

$$P = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}, \quad P_1 \in R^{n\times r}, \quad P_2 \in R^{r\times r}.$$

Then $P_2P_2^T = I_r$. It is easy to check that $f(P_1P_2^T) = H_f \circ Z^* < 0, \; g_i(P_1P_2^T) \le 0, \;\; i = 1, \cdots, r$. This contradicts the assumptions.
If $k > r$, use the same steps in the proof of Lemma 5.1, we can find a solution of system (14) with rank r. From the above argument, this will lead to a contradiction. So the system has no solution. From Lemma 2.2, the required conclusion follows.

Beck[1] first studied the following quadratic matrix programming and proved the strong duality under some conditions. But his proof is quite complex, here we offer a simple easy-understanding proof based on Theorem 5.1.

Theorem 5.2: Consider the following quadratic matrix programming problem

$$(P) \qquad min \quad tr(x^T Ax) + 2tr(a^T x) + c$$
$$s.t. \quad tr(x^T B_i x) + 2tr(b_i^T x) + c_i \leq 0. \quad i = 1, \cdots, r$$

where $x \in R^{n\times r}$ and there exists an $\bar{x}$ such that $g_i(\bar{x}) < 0, i = 1, \cdots, r$. then strong duality of (P) is satisfied which is $V(P) = V(D)$.

Proof: It is easy to get the dual problem of (P) as following.

$$(D) \qquad max \quad \gamma$$
$$s.t. \quad \begin{bmatrix} A + \sum_{i=1}^r \lambda_i B_i & a + \sum_{i=1}^r \lambda_i b_i \\ (a + \sum_{i=1}^r \lambda_i b_i)^T & \frac{c + \sum_{i=1}^r \lambda_i c_i - \gamma}{r} I \end{bmatrix} \succeq 0$$
$$\lambda \geq 0$$

Lagrange function of (D) is

$$L = -\gamma - \begin{bmatrix} A + \sum_{i=1}^r \lambda_i B_i & a + \sum_{i=1}^r \lambda_i b_i \\ (a + \sum_{i=1}^r \lambda_i b_i)^T & \frac{c + \sum_{i=1}^r \lambda_i c_i - \gamma}{r} I \end{bmatrix} \circ Z - \lambda^T \delta$$

then

$$\min_{\lambda,\gamma}\{-\begin{bmatrix} A & a \\ a^T & \frac{c}{r} I \end{bmatrix} \circ Z - \sum_{i=1}^r \lambda_i(\begin{bmatrix} B_i & b_i \\ b_i^T & \frac{c_i}{r} I \end{bmatrix} \circ Z + \delta_i) + \frac{\gamma}{r}(\sum_{j=n+1}^{n+r} Z_{j,j} - r)\}$$
$$= \begin{cases} -\begin{bmatrix} A & a \\ a^T & \frac{c}{r} I \end{bmatrix} \circ Z & \begin{bmatrix} B_i & b_i \\ b_i^T & \frac{c_i}{r} I \end{bmatrix} \circ Z + \delta_i = 0, i = 1, \cdots, r, Z \succeq 0, \sum_{j=n+1}^{n+r} Z_{j,j} = r \\ -\infty & else \end{cases}$$

So the dual problem of (D) is

$$(RD) \qquad max \quad -\begin{bmatrix} A & a \\ a^T & \frac{c}{r} I \end{bmatrix} \circ Z$$
$$s.t. \quad \begin{bmatrix} B_i & b_i \\ b_i^T & \frac{c_i}{r} I \end{bmatrix} \circ Z \leq 0, \qquad i = 1, \cdots, r$$
$$Z \succeq 0, \sum_{j=n+1}^{n+r} Z_{j,j} = r$$

Supposed that $V(P)$ is the optimal value of (P), then

$$\begin{aligned} &f(x) - V(P) < 0 \\ &g_i(x) \leq 0 \quad i = 1, \cdots, r \end{aligned}$$

has no solution. Then from Theorem 5.1, there exist $\lambda_i \geq 0$ such that

$$\begin{bmatrix} A & a \\ a^T & \frac{c}{r}I \end{bmatrix} \circ Z + \sum_{i=1}^{r} \lambda_i \begin{bmatrix} B_i & b_i \\ b_i^T & \frac{c_i}{r}I \end{bmatrix} \circ Z \geq V(P) \qquad \forall Z \in Q$$

We take the minimum for the left side of above inequality. Then one has $V(RD) \geq V(P)$, where (RD) is the dual problem of (D). Since both (D) and (RD) are convex optimization problem, so $V(D) = V(RD)$, then $V(D) \geq V(P)$. From the weak duality, we know $V(D) \leq V(P)$. So $V(D) = V(P)$. This completes the proof.

6. Conclusion

In this paper, by extending a matrix rank-one decomposition, we develop Yuan's lemma from several aspects. In particular we extend Yuan's lemma to general quadratic function system and quadratic matrix function system, from which several S-lemmas are proved with the easy-understanding way or proposed. As an application of S-lemma on inequality quadratic matrix system, we provide a simple proof of the strong duality of a class of quadratic matrix programming.
In [6] authors provided a tractable extension of Yuan's theorem of the alternative to the symmetric tensor, and showed that the solution of a polynomial optimization problem with suitable structure can be found by solving a single semi-definite programming problem. We believe that the results of the current paper can be extended to at least fourth-tensor system and applied to establish the strong duality of some tensor optimization problems. We will make the further study on this problem in the separate paper.

Reference

[1] A. Beck. Quadratic matrix programming. *SIAM Journal on Optimization 17 (4)*, (2007), 1224–1238.

[2] I. P'olik and T. Terlaky. A survey of the S-lemma. *SIAM Review. 49(3)*, (2007), 371–418.

[3] J.F Sturm and S. Zhang. On cones of nonnegative quadratic functions. *Mathematics of Operations Research 28*, (2003), 246–267.

[4] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press (2004).

[5] S. He, M.H. Wong and S. Zhang. On the S-lemma for univariate polynomials. *Mathematics of Operations Research, (2016).*

[6] S. Hu, G. Li and L. Qi. A tensor analogy of Yuan's theorem of the alternative and polynomial optimization with sign structure. *Journal of Optimization Theory and Applications*, (2016), 446–474.

[7] S. Wang and Y. Xia. Strong Duality for Generalized Trust Region Subproblem: S-Lemma with Interval Bounds. *Optimization Letters 9 (6)*, (2015), 1063–1073.

[8] T. Pong and H. Wolkowicz. Generalizations of the trust region subproblem. *Comput. Optim. Appl. 58(2)*, (2014), 273–322.

[9] V.A. Yakubovich. S-procedure in nonlinear control theory. *Vestnik Leningrad. Univ. 4*, (1977), 73–93.

[10] W. Ai, Y. Huang and S. Zhang. New results on Hermitian matrix rank-one decomposition. *Mathematical programming 128 (1)*, (2011), 253–283.

[11] Y. Huang and S. Zhang. Complex Matrix Decomposition and Quadratic Programming. *Mathematics of Operations Research 32*, (2007), 758 – 768.

[12] Y. X. Yuan. On a subproblem of trust region algorithms for constrained optimization. *Mathematical Programming 47 (1)*, (1990), 53–63.

[13] Y. Xia, S. Wang and R. L Sheu. S-lemma with equality and its applications. *Mathematical Programming 156 (1-2)*, (2016), 513–547.

[14] Y. Ye and S. Zhang. New results on quadratic minimization. *SIAM Journal on Optimization 14 (1)*, (2001), 245–267.

[15] Z.-Q Luo, J. F Sturm and S. Zhang. Multivariate nonnegative quadratic mappings. *SIAM J. Optim. 14*, (2004), 1140–1162.