An improved probability bound for the Approximate S-Lemma

Kürşad Derinkuyu, Mustafa Ç. Pınar*, Ahmet Camcı

Department of Industrial Engineering, Bilkent University, 06800 Bilkent, Ankara, Turkey

Received 22 February 2006; accepted 9 February 2007
Available online 28 February 2007

Abstract

The purpose of this note is to give a probability bound on symmetric matrices to improve an error bound in the Approximate S-Lemma used in establishing levels of conservatism results for approximate robust counterparts.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Robust optimization; S-Lemma

1. Introduction

The purpose of this note is to prove the following result:

Lemma 1. Let $B$ denote a symmetric $n \times n$ matrix and $\xi = \{\xi_1, \ldots, \xi_n\} \in \mathbb{R}^n$. If the coordinates $\xi_i$ of $\xi$ are independently identically distributed random variables with

$$\Pr(\xi_i = 1) = \Pr(\xi_i = -1) = 1/2$$

then one has

$$\Pr(\xi^T B \xi \leq \text{Tr } B) \geq \frac{1}{2 \log_2(n)} > \frac{1}{2n}. $$

The above result improves Lemma A.4 by Ben-Tal et al. [1] which stated

$$\Pr(\xi^T B \xi \leq \text{Tr } B) \geq \frac{1}{8n^2},$$

and where the authors conjectured that the right-hand side could be improved to $\frac{1}{4}$. Ben-Tal et al. [1] used Lemma A.4 to give the Approximate S-Lemma used in levels of conservatism results for approximate robust counterparts of uncertain convex programs. Our Lemma 1 above improves the error bound in the Approximate S-Lemma of [1] to

$$\rho := \left(2 \log \left(4n \sum_{k=1}^{K} \text{rank } R_k\right)\right)^{1/2}$$

from

$$\rho := \left(2 \log \left(16n^2 \sum_{k=1}^{K} \text{rank } R_k\right)\right)^{1/2}. $$

* Corresponding author. Tel.: +90 312 290 1514; fax: +90 312 266 4054.
E-mail addresses: kursad@mail.utexas.edu (K. Derinkuyu), mustafap@bilkent.edu.tr (M. Ç. Pınar), camci@bilkent.edu.tr (A. Camcı).

0167-6377/ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.orl.2007.02.003
2. Proof of the main result

Our proof, which is based on contradiction, recursively eliminates the non-zero entries of a symmetric matrix while the proof of [1] uses moments. We arrive at the proof of Lemma 1 after giving three intermediate results.

First, since \( \text{Tr } B = \zeta^T \text{diag } B \zeta \) for any \( \zeta \in \{-1, 1\}^n \) it follows that
\[
\Pr(\zeta^T B \zeta \leq \text{Tr } B) = \Pr(\zeta^T B \zeta - \text{Tr } B \leq 0) = \Pr(\zeta^T (B - \text{diag } B) \zeta \leq 0).
\]

This enables us to restrict ourselves to the case that the matrix under consideration is a symmetric matrix with zero diagonal since \( B - \text{diag } B \) is a matrix with this property. Therefore, in order to prove Lemma 1 we need to show that for any symmetric matrix \( B \) with zero diagonal, and for \( \zeta \) as defined in Lemma 1 we have
\[
\Pr(\zeta^T B \zeta \leq 0) \geq \frac{1}{2\lceil \log_2(n) \rceil}.
\] (5)

Now, we will give three intermediate results which lead to the proof of Lemma 1.

**Lemma 2.** Let \( X \) be a finite set. Then for any pair of subsets \( U \) and \( V \) of \( X \), one has
\[
|U \cap V| \leq |U| + |V| - |X|.
\]

**Proof.** Using the inclusion–exclusion principle we have \( |U| + |V| - |U \cap V| = |U \cup V| \leq |X| \). After rearranging the right and left sides of the inequality we get the desired result. \( \square \)

**Lemma 3.** Let \( f : \mathbb{N} \to \mathbb{N} \) be a function such that \( f(n) = \lceil n/2 \rceil \). If \( k = \lceil \log_2(n) \rceil \), then \( f^k(n) = f(f(\ldots(f(n))\ldots)) \leq 1 \).

**Proof.** By the definition of \( k \) we have \( k - 1 < \log_2(n) \leq k \), which implies \( n \leq 2^k \). Since \( f \) is a non-decreasing function, we have \( f^k(n) \leq f^k(2^k) \). It can be seen that \( f^k(2^k) = 1 \). Therefore the result holds. \( \square \)

In the remaining part of the paper for any \( q \in \mathbb{R}^n \) such that \( q(i) \in \{-1, 1\} \) for any \( i \in \{1, \ldots, n\} \) we denote \( \text{diag}(q) \) by \( Q \). Here, \( q(i) \) is the \( i \)th entry of vector \( q \). For any such \( Q \) and any symmetric matrix \( B \) having zero diagonal entries we define
\[
B^q = \frac{1}{2}(B + QBQ).
\]
The matrix \( QBQ \) is a symmetric matrix with zero diagonal. Hence, \( B^q \) is a symmetric matrix with zero diagonal. Since \( q(i)q(j) \in \{-1, 1\} \) and the \((i, j)\) entry of \( QBQ \) is given by \( q(i)q(j)B_{ij} \) we have
\[
B^q_{ij} = \begin{cases} B_{ij} & \text{if } q(i)q(j) = 1, \\ 0 & \text{if } q(i)q(j) = -1. \end{cases}
\]

**Lemma 4.** Let \( \zeta \) and \( B \) defined as in Lemma 1. Moreover, let \( Q = \text{diag}(q) \), with \( q \in \mathbb{R}^n \) such that \( q_i \in \{-1, 1\} \) and \( B^q \) as defined above. Then one has
\[
\Pr(\zeta^T B \zeta > 0) = \Pr(\zeta^T QBQ \zeta > 0), \tag{6}
\]
and
\[
\Pr(\zeta^T B^q \zeta > 0) \geq 2 \Pr(\zeta^T B \zeta > 0) - 1. \tag{7}
\]

**Proof.** We have
\[
(\zeta^T Q \zeta) \cdot QBQ \cdot Q \zeta = \zeta^T Q^2 B Q^2 \zeta = \zeta^T B \zeta,
\]
since \( Q^2 = I_n \), where \( I_n \) is the \( n \times n \) identity matrix. Hence
\[
\Pr(\zeta^T B \zeta > 0) = \Pr((Q \zeta)^T \cdot QBQ \cdot Q \zeta > 0).
\]
Since \( \zeta \) and \( Q \zeta \) occur with the same probability this implies (6). To prove (7) we use the fact
\[
\Pr(\zeta^T B^q \zeta > 0) = \Pr(\zeta^T (B + QBQ) \zeta > 0) \geq \Pr(\zeta^T B \zeta > 0 \land \zeta^T QBQ \zeta > 0).
\]
Then using Lemma 2 we get
\[
\Pr(\zeta^T B \zeta > 0 \land \zeta^T QBQ \zeta > 0) \geq \Pr(\zeta^T B \zeta > 0) + \Pr(\zeta^T QBQ \zeta > 0) - 1
\]
\[
= 2 \Pr(\zeta^T B \zeta > 0) - 1,
\]
where the last equality follows from (6). Therefore we get inequality (7). \( \square \)

At this point, using our result in Lemma 4, we are ready to prove Lemma 1.
Proof of Lemma 1. Assume to the contrary that Lemma 1 is false. Then, one can see from the derivation of inequality (5) that there exists a symmetric $n \times n$ matrix $B$ having zero diagonal such that

$$\Pr(\xi^T B \xi \leq 0) < \frac{1}{2^{\lceil \log_2(n) \rceil}}$$

which is equivalent to

$$\Pr(\xi^T B \xi > 0) > 1 - \frac{1}{2^{\lceil \log_2(n) \rceil}}. \quad (9)$$

We construct a sequence of block diagonal matrices $B_i$ having zero diagonal such that

$$B_1 = B, \quad B_{i+1} = B_i^{Q_i}, \quad i = 1, 2, \ldots, k.$$ 

We have $k = \lceil \log_2(n) \rceil$, and $q_i$'s are chosen according to the following process. For $q_1$ we take the first $\lceil n/2 \rceil$ entries as 1's and the remaining entries as $-1$'s. Let us call these two parts of $q_1$ as segments of $q_1$. We illustrate this for $n = 13$ with two segments separated by the symbol “|” again:

$$q_1 = [1\ 1\ 1\ 1\ 1\ 1\ 1\ -1\ -1\ -1\ -1\ -1\ -1].$$

For $q_{i+1}$, consider each segment of $q_i$. If the length of a segment is $l$ we take the first $\lceil l/2 \rceil$ entries as 1's and the remaining entries in the segment as $-1$'s. Let us call these two parts segments again. Note that if $l = 1$ for a segment the process will produce only one part of length 1 out of the segment. The resulting vector is $q_{i+1}$ with its segments defined as above. To illustrate it for $n = 13$, we show $q_2$ obtained from $q_1$. Here, $q_2$ has four segments separated by the symbol “|” again:

$$q_2 = [1\ 1\ 1\ 1\ -1\ -1\ -1\ 1\ 1\ 1\ -1\ -1\ -1].$$

Now, let $S$ denote the first principal submatrix of $B$ with size $\lceil n/2 \rceil \times \lceil n/2 \rceil$, and let $T$ denote the last principal submatrix of $B$ with size $\lfloor n/2 \rfloor \times \lfloor n/2 \rfloor$. Denote the remaining matrix at the upper right corner of $B$ by $R$, and the remaining matrix at the lower left corner of $B$ becomes $R^T$ since $B$ is symmetric. Then $B^{Q_1}$ is obtained from $B$ by replacing all entries of $R$ and $R^T$ by zeros. In other words

$$B_1 = B = \begin{bmatrix} S & R^T \\ \end{bmatrix} \Rightarrow Q_1 B_1 Q_1 = \begin{bmatrix} S & -R \\ -R^T & T \\ \end{bmatrix}$$

$$\Rightarrow B_2 = B^{Q_1} = \begin{bmatrix} S & 0 \\ 0 & T \\ \end{bmatrix},$$

where $Q_1$ is the diagonal matrix with the vector $q_1$ as the diagonal. Now using Lemma 4 and (9) we obtain

$$\Pr(\xi^T B_2 \xi > 0) > 2 \left( 1 - \frac{1}{2^{\lceil \log_2(n) \rceil}} \right) - 1$$

$$= 1 - \frac{2}{2^{\lceil \log_2(n) \rceil}}. \quad (10)$$

Note that the block matrices along the diagonal of $B_2$ have sizes $\lceil n/2 \rceil$ and $\lfloor n/2 \rfloor$. Hence, the sizes do not exceed $f(n)$ of Lemma 3 which was defined as $f(n) = \lceil n/2 \rceil$. We repeat the above procedure using $q_2$ which was shown before. Thus we obtain $B_3 = B_2^{Q_2}$ which has the form

$$B_3 = \begin{bmatrix} D_1 & & & \\ & D_2 & & \\ & & D_3 & \\ & & & D_4 \end{bmatrix},$$

where $D_1, D_2, D_3$ and $D_4$ constitute the symmetric, zero-diagonal blocks of the block diagonal matrix $B_3$. These block matrices have dimensions $\lceil \frac{1}{2} \lceil n/2 \rceil \rceil \times \lceil \frac{1}{2} \lceil n/2 \rceil \rceil$, $\lfloor \frac{1}{2} \lfloor n/2 \rfloor \rfloor \times \lceil \frac{1}{2} \lceil n/2 \rceil \rceil$, $\lfloor \frac{1}{2} \lfloor n/2 \rfloor \rfloor \times \lfloor \frac{1}{2} \lfloor n/2 \rfloor \rfloor$, respectively.

Now, again by Lemma 4 and (10) $B_3$ satisfies

$$\Pr(\xi^T B_3 \xi > 0) > 2 \left( 1 - \frac{2}{2^{\lceil \log_2(n) \rceil}} \right) - 1$$

$$= 1 - \frac{2^2}{2^{\lceil \log_2(n) \rceil}}. \quad (11)$$

Note that the sizes of the block diagonal matrices along the diagonal of $B_3$ can be at most $\lceil \frac{1}{2} \lceil n/2 \rceil \rceil$ which does not exceed $f^2(n)$. We construct $q_3$ in the same way as before. For $n = 13$ this gives

$$q_3 = [1\ 1\ | -1\ -1\ | 1\ 1\ 1\ | -1\ | 1\ 1\ | -1\ | 1\ 1\ | -1\ | 1\ 1\ | -1].$$

Again by using Lemma 4 and (11) we obtain for $B_4$ that

$$\Pr(\xi^T B_4 \xi > 0) > 2 \left( 1 - \frac{2^2}{2^{\lceil \log_2(n) \rceil}} \right) - 1$$

$$= 1 - \frac{2^3}{2^{\lceil \log_2(n) \rceil}}. \quad (12)$$
This time the sizes of the block diagonal matrices along the diagonal of \( B_4 \) do not exceed \( f^3(n) \). Then, \( q_4 \) is constructed in the same manner, and for \( n = 13 \) we have

\[
q_4 = [1 \mid -1 \mid 1 \mid -1 \mid 1 \mid -1 \mid 1 \mid -1 \mid 1 \mid 1 \mid -1 \mid 1 \mid -1 \mid 1 \mid 1 \mid -1].
\]

Hence, at the next step we get

\[
\Pr(\xi^T B_5 \xi > 0) \geq 2 \left( 1 - \frac{2^3}{2^{\lceil \log_2(n) \rceil}} \right) - 1
\]

\[
= 1 - \frac{2^4}{2^{\lceil \log_2(n) \rceil}},
\] (13)

and the sizes of the block diagonal matrices along the diagonal of \( B_5 \) do not exceed \( f^4(n) \). Note that for \( n = 13 \) these block matrices all have size 1. In the general case we proceed in the same way and after \( k \) steps we obtain

\[
\Pr(\xi^T B_{k+1} \xi > 0) \geq 1 - \frac{2^k}{2^{\lceil \log_2(n) \rceil}},
\] (14)

and the block diagonal matrices along the diagonal of \( B_{k+1} \) have sizes that do not exceed \( f^k(n) \). Now Lemma 3 implies that if \( k = \lceil \log_2(n) \rceil \), then \( f^k(n) \leq 1 \). In that case the right hand side of (14) is equal to 0. Also, the block diagonal matrices along the diagonal of \( B_{k+1} \) have sizes at most 1. We know from the construction procedure of \( B_{k+1} \) that it has zero diagonal. Hence, \( B_{k+1} \) becomes a matrix of zeros. But then the left-hand side of (14) is also equal to 0. Therefore, we arrive at the contradiction \( 0 > 0 \). This completes the proof of Lemma 1. \( \square \)

Now, it suffices to observe that equipped with the result of the previous lemma, one has to solve Eq. (A.38) pp. 559 of [1] using the probability bound \( 1/2n \) to obtain the improved bound (3).

Although we were not able to prove the conjecture of Ben-Tal et al. in [1] that would help us remove the factor \( n \) under the logarithm altogether, we offered an improvement from \( n^2 \) to \( n \) under the logarithm. While this paper was under review, we learned of a recent result [2] where it is shown that

\[
\Pr(\xi^T B \xi \leq \text{Tr } B) \geq \frac{1}{17}.
\]

Our result in Lemma 1 remains better in the range \( 3 \leq n \leq 64 \).

Acknowledgments

We are grateful to an anonymous referee whose very detailed comments helped reorganize our arguments and led to a better presentation.

References
