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

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.

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 $\xi = \{\xi_1, .., \xi_n\} \in \mathbb{R}^n$. If the coordinates ξ_i of ξ are independently identically distributed random variables with*

$$Pr(\xi_i = 1) = Pr(\xi_i = -1) = 1/2 \tag{1}$$

then one has

$$Pr(\xi^T B\xi \leq TrB) \geq \frac{1}{2^{\lceil \log_2(n) \rceil}} > \frac{1}{2n}. \tag{2}$$

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

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

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

$$\rho := \left(\ 2log(\ 4n \textstyle\sum_{k=1}^{K} \mathrm{rank} R_k \ ) \ \right)^{\frac{1}{2}} \tag{3}$$

from

$$\rho := \left(\ 2log(\ 16n^2 \textstyle\sum_{k=1}^{K} \mathrm{rank} \ R_k \ ) \ \right)^{\frac{1}{2}}. \tag{4}$$

* kursad@mail.utexas.edu
† Corresponding author, mustafap@bilkent.edu.tr
‡ camci@bilkent.edu.tr

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 $\mathrm{Tr}B = \xi^T \mathrm{diag}B \xi$ for any $\xi \in \{-1, 1\}^n$ it follows that

$$\Pr(\xi^T B\xi \leq \mathrm{Tr}B) = \Pr(\xi^T B\xi - \mathrm{Tr}B \leq 0) = \Pr(\xi^T (B - \mathrm{diag}B)\xi \leq 0).$$

This enables us to restrict ourselves to the case that the matrix under consideration is a symmetric matrix with zero diagonal since $B - \mathrm{diag}B$ 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 ξ as defined in Lemma 1 we have

$$\Pr(\xi^T B\xi \leq 0) \geq \frac{1}{2^{\lceil \log_2(n) \rceil}}. \tag{5}$$

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*

$$|U \cap V| \geq |U| + |V| - |X|.$$

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. ∎

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

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. ∎

In the remaining part of the paper for any $q \in \mathbb{R}^n$ such that $q(i) \in \{-1, 1\}$ for any $i \in \{1, .., n\}$ we denote $\mathrm{diag}(q)$ 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

$$B^q_{ij} = \begin{cases} B_{ij} & \text{if} \quad q(i)q(j) = 1 \\ 0 & \text{if} \quad q(i)q(j) = -1. \end{cases}$$

Lemma 4 *Let* ξ *and* B *defined as in Lemma 1. Moreover, let* $Q = diag(q)$, *with* $q \in \Re^n$ *such that* $q_i \in \{-1, 1\}$ *and* B^q *as defined above. Then one has*

$$Pr(\xi^T B\xi > 0) = Pr(\xi^T QBQ\xi > 0), \tag{6}$$

and

$$Pr(\xi^T B^q \xi > 0) \geq 2Pr(\xi^T B\xi > 0) - 1. \tag{7}$$

Proof: We have

$$(Q\xi)^T \cdot QBQ \cdot Q\xi = \xi^T Q^2 BQ^2 \xi = \xi^T B\xi,$$

since $Q^2 = I_n$, where I_n is the $n \times n$ identity matrix. Hence

$$\Pr(\xi^T B\xi > 0) = \Pr((Q\xi)^T \cdot QBQ \cdot Q\xi > 0).$$

Since ξ and $Q\xi$ occur with the same probability this implies (6). To prove (7) we use the fact

$$\Pr(\xi^T B^q \xi > 0) = \Pr(\xi^T (B + QBQ)\xi > 0) \geq \Pr(\xi^T B\xi > 0 \ \& \ \xi^T QBQ\xi > 0).$$

Then using Lemma 2 we get

$$\Pr(\xi^T B\xi > 0 \ \& \ \xi^T QBQ\xi > 0) \geq \Pr(\xi^T B\xi > 0) + \Pr(\xi^T QBQ\xi > 0) - 1 = 2\Pr(\xi^T B\xi > 0) - 1,$$

where the last equality follows from (6). Therefore we get inequality (7). ∎

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

$$\Pr(\xi^T B\xi \leq 0) < \frac{1}{2^{\lceil \log_2(n) \rceil}} \tag{8}$$

which is equivalent to

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

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 $\lceil \frac{n}{2} \rceil$ 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 " | ".

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

For q_{i+1}, consider each segment of q_i. If the length of a segment is l we take the first $\lceil \frac{l}{2} \rceil$ 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 $\lceil \frac{n}{2} \rceil \times \lceil \frac{n}{2} \rceil$, and let T denote the last principal submatrix of B with size $\lfloor \frac{n}{2} \rfloor \times \lfloor \frac{n}{2} \rfloor$. 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,

$$B_1 = B = \begin{bmatrix} S & R \\ R^T & T \end{bmatrix} \Rightarrow Q_1 B_1 Q_1 = \begin{bmatrix} S & -R \\ -R^T & T \end{bmatrix} \Rightarrow B_2 = B^{q_1} = \begin{bmatrix} S & 0 \\ 0 & T \end{bmatrix}$$

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

$$\Pr(\xi^T B_2 \xi > 0) > 2(1 - \frac{1}{2^{\lceil \log_2(n) \rceil}}) - 1 = 1 - \frac{2}{2^{\lceil \log_2(n) \rceil}}. \tag{10}$$

Note that the block matrices along the diagonal of B_2 have sizes $\lceil \frac{n}{2} \rceil$ and $\lfloor \frac{n}{2} \rfloor$. Hence, the sizes do not exceed $f(n)$ of Lemma 3 which was defined as $f(n) = \lceil \frac{n}{2} \rceil$. 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 $\lceil \frac{1}{2} \lceil \frac{n}{2} \rceil \rceil \times \lceil \frac{1}{2} \lceil \frac{n}{2} \rceil \rceil$, $\lfloor \frac{1}{2} \lceil \frac{n}{2} \rceil \rfloor \times \lfloor \frac{1}{2} \lceil \frac{n}{2} \rceil \rfloor$, $\lceil \frac{1}{2} \lfloor \frac{n}{2} \rfloor \rceil \times \lceil \frac{1}{2} \lfloor \frac{n}{2} \rfloor \rceil$, $\lfloor \frac{1}{2} \lfloor \frac{n}{2} \rfloor \rfloor \times \lfloor \frac{1}{2} \lfloor \frac{n}{2} \rfloor \rfloor$, respectively.

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

$$\Pr(\xi^T B_3 \xi > 0) > 2(1 - \frac{2}{2^{\lceil \log_2(n) \rceil}}) - 1 = 1 - \frac{2^2}{2^{\lceil \log_2(n) \rceil}}. \tag{11}$$

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

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

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

$$\Pr(\xi^T B_4 \xi > 0) > 2(1 - \frac{2^2}{2^{\lceil \log_2(n) \rceil}}) - 1 = 1 - \frac{2^3}{2^{\lceil \log_2(n) \rceil}}. \tag{12}$$

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

$$q_4 = [\ 1\ \ |\ \ -1\ \ |\ \ 1\ \ |\ \ -1\ \ |\ \ 1\ \ |\ \ -1\ \ |\ \ 1\ \ |\ \ 1\ \ |\ \ -1\ \ |\ \ 1\ \ |\ \ -1\ \ |\ \ 1\ \ |\ \ -1\ ]\ .$$

Hence, at the next step we get

$$\Pr(\xi^T B_5 \xi > 0) > 2(1 - \frac{2^3}{2^{\lceil \log_2(n) \rceil}}) - 1 = 1 - \frac{2^4}{2^{\lceil \log_2(n) \rceil}}, \tag{13}$$

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

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

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. ∎

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 $\frac{1}{2n}$ 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

$$\Pr(\xi^T B \xi \leq \mathrm{Tr} B) \geq \frac{1}{87}.$$

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

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

References

[1] A. Ben-Tal, A. Nemirovski, and C. Roos, 2002, Robust Solutions of Uncertain Quadratic and Conic-Quadratic Problems, *SIAM J. on Optimization*, Vol. 13, pp. 535–560.

[2] S. He, Z.-Q. Luo, J. Nie and S. Zhang, 2007, Semidefinite Relaxation Bounds for Indefinite Homogeneous Quadratic Optimization, Technical Report SEEM2007-01, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong.