A note on the non-commutative arithmetic-geometric mean inequality

Teng Zhang

November 26, 2014

Abstract

This note proves the following inequality: if \( n = 3k \) for some positive integer \( k \), then for any \( n \) positive definite matrices \( A_1, A_2, \ldots, A_n \),

\[
\frac{1}{n^3} \left\| \sum_{j_1, j_2, j_3 = 1}^{n} A_{j_1} A_{j_2} A_{j_3} \right\| \geq \frac{(n-3)!}{n!} \left\| \sum_{j_1, j_2, j_3 = 1, \text{j_1, j_2, j_3 all distinct}}^{n} A_{j_1} A_{j_2} A_{j_3} \right\|,
\]

where \( \| \cdot \| \) represents the operator norm. This inequality is a special case of a recent conjecture by Recht and Ré.

1 Introduction

In [3], Recht and Ré conjectured that the standard arithmetic mean-geometric mean (AM-GM) inequality can be generalized to the non-commutative setting for positive definite matrices \( \{A_i \}_{i=1}^{n} \) as follows:

\[
\frac{1}{n^m} \left\| \sum_{j_1, \ldots, j_m = 1}^{n} A_{j_1} A_{j_2} \cdots A_{j_m} \right\| \geq \frac{(n-m)!}{n!} \left\| \sum_{j_1, \ldots, j_m = 1, \text{j_1, \ldots, j_m all distinct}}^{n} A_{j_1} A_{j_2} \cdots A_{j_m} \right\|.
\]

This inequality gives theoretical guarantee to the fact that, without-replacement sampling leads to faster convergence rates than with-replacement sampling for both the least mean squares and randomized Kaczmarz algorithms.

While the case \( n = m = 2 \) has been proved in [3] Proposition 3.2, to the best of our knowledge, the conjecture for the cases \( n, m > 2 \) remains open, and the main contribution of this note is a proof of the conjecture when \( m = 3 \) and \( n = 3k \) for some positive integer \( k \).

We remark that the following variant of the conjecture

\[
\frac{1}{n^m} \sum_{j_1, \ldots, j_m = 1}^{n} \left\| A_{j_1} A_{j_2} A_{j_3} \right\| \geq \frac{(n-m)!}{n!} \sum_{j_1, \ldots, j_m = 1, \text{j_1, \ldots, j_m all distinct}}^{n} \left\| A_{j_1} A_{j_2} A_{j_3} \right\|
\]

was proposed in [1] and the case \( m = 3 \) has been proved recently in [2].
1.1 Reduction of the conjecture

To prove (1.1), WLOG we assume that \( \| \sum_{j=1}^{n} A_j \| = 1 \), which is equivalent to

\[
\sum_{j=1}^{n} A_j \leq I.
\]

In the note, we write \( A \geq B \) or \( B \leq A \) if and only if \( A - B \) is positive semidefinite. Then the LHS of (1.1) is \( 1/n^m \) and it is sufficient to prove

\[
-\frac{1}{n^m} I \leq E[A_{i_1} A_{i_2} \cdots A_{i_m}] \leq \frac{1}{n^m} I,
\]

where \( \{i_1, i_2, \cdots, i_n\} \) is a random permutation of \( \{1, 2, \cdots, n\} \).

2 The proof of the conjecture for \( n = m = 3 \)

The proof is based on the following lemmas, and their proofs are deferred to Sections 2.1 and 2.2.

**Lemma 2.1.** For symmetric matrices \( A, B \) and positive semidefinite matrix \( C \),

\[
ACA + BCB \geq ACB + BCA.
\]

**Lemma 2.2.** If \( A, B \) and \( C \) are symmetric matrices and \( A \leq B \), then

\[
CAC \leq CBC.
\]

We will prove the lower bound and the upper bound in (1.2) separately. To prove the upper bound of \( E[A_{i_1} A_{i_2} A_{i_3}] \), we apply Lemma 2.1 and obtain

\[
A_{i_1} A_{i_2} A_{i_3} + A_{i_3} A_{i_2} A_{i_1} \leq A_{i_1} A_{i_2} A_{i_1} + A_{i_3} A_{i_2} A_{i_3}.
\]

Therefore,

\[
E[A_{i_1} A_{i_2} A_{i_3}] = \frac{1}{2} E[A_{i_1} A_{i_2} A_{i_3} + A_{i_3} A_{i_2} A_{i_1}] \leq \frac{1}{4} E[A_{i_1} A_{i_2} A_{i_3} + A_{i_3} A_{i_2} A_{i_1} + A_{i_1} A_{i_2} A_{i_1} + A_{i_3} A_{i_2} A_{i_3}] = \frac{1}{4} E[(A_{i_1} + A_{i_3}) A_{i_2} (A_{i_1} + A_{i_3})].
\]

(2.1)

Since \( 0 \leq A_{i_1} + A_{i_3} \leq I \), max_{0 \leq a \leq 1} \( a^2 (1 - a) = 4/27 \) and \( A_{i_2} \leq I - A_{i_1} - A_{i_3} \), applying Lemma 2.2 we have

\[
(A_{i_1} + A_{i_3}) A_{i_2} (A_{i_1} + A_{i_3}) \leq (A_{i_1} + A_{i_3}) (I - (A_{i_1} + A_{i_3})) (A_{i_1} + A_{i_3}) \leq \frac{4}{27} I.
\]

(2.2)

Combining (2.1) and (2.2), the upper bound of \( E[A_{i_1} A_{i_2} A_{i_3}] \) in (1.2) is proved.
To prove the lower bound of $\mathbb{E}[A_i A_{i2} A_{i3}]$ in (1.2), we again apply Lemma 2.1 and obtain
\[ -A_i A_{i2} A_{i3} + A_{i3} A_{i2} A_i \leq A_i A_{i2} A_i + A_{i3} A_{i2} A_{i3}. \]

Similar to (2.2), we have
\[ \mathbb{E}[-A_i A_{i2} A_{i3}] = \frac{1}{2} \mathbb{E}[A_{i3} A_{i2} A_i - A_{i3} A_{i2} A_{i3}] \]
\[ \leq \frac{1}{4} \mathbb{E}[-A_i A_{i2} A_{i3} - A_{i3} A_{i2} A_{i3} + A_{i3} A_{i2} A_{i3}] \]
\[ = \frac{1}{4} \mathbb{E}[(A_i - A_{i3}) A_{i2} (A_i - A_{i3})]. \]

Therefore, to prove the lower bound in (1.2), it is sufficient to show
\[ \mathbb{E}[(A_i - A_{i3}) A_{i2} (A_i - A_{i3})] \leq \frac{4}{27} I. \]

To prove (2.4), we only need to consider the case
\[ \sum_{i=1}^{3} A_i = I, \]

since the triple $\{\hat{A}_1, \hat{A}_2, \hat{A}_3\} = \{A_1 + D, A_2 + D, A_3 + D\}$ with $D = (I - A_1 - A_2 - A_3)/3$ satisfies $\hat{A}_1 + \hat{A}_2 + \hat{A}_3 = I$ and by Lemma 2.2,
\[ \mathbb{E}[(A_i - A_{i3}) A_{i2} (A_i - A_{i3})] \leq \mathbb{E}[(\hat{\hat{A}}_i - \hat{\hat{A}}_{i3}) \hat{\hat{A}}_{i2} (\hat{\hat{A}}_i - \hat{\hat{A}}_{i3})]. \]

Under the assumption (2.5), we have
\[ \mathbb{E}[(\hat{\hat{A}}_i - \hat{\hat{A}}_{i3}) A_{i2} (\hat{\hat{A}}_i - \hat{\hat{A}}_{i3})] \]
\[ = \mathbb{E}[2 A_{i2} A_{i} A_{i2} A_{i} - (A_{i} + A_{i3}) A_{i2} (A_{i} + A_{i3})] \]
\[ = \mathbb{E}[2 A_{i2} (A_{i} + A_{i3}) A_{i2} - (A_{i} + A_{i3}) A_{i2} (A_{i} + A_{i3})] \]
\[ = \mathbb{E}[2 A_{i2} (I - A_{i2}) A_{i2} - (I - A_{i2}) A_{i2} (I - A_{i2})] \]
\[ = \mathbb{E}[-A_{i2} + 4 A_{i2}^2 - 3 A_{i2}^3] = \frac{4}{27} I + \mathbb{E}[-\frac{13}{9} A_{i2} + 4 A_{i2}^2 - 3 A_{i2}^3], \]

where the last step applies $\mathbb{E}[A_{i2}] = \frac{1}{3} I$, which follows from (2.5). Since $\max_{0 \leq x \leq 1} -\frac{13}{9} x + 4 x^2 - 3 x^3 = 0$ and $0 \leq A_{i2} \leq I$, we have
\[ -\frac{13}{9} A_{i2} + 4 A_{i2}^2 - 3 A_{i2}^3 \leq 0. \]

Combining it with (2.6), we proved (2.4) and therefore the lower bound in (1.2).

### 2.1 Proof of Lemma 2.1

The difference of its LHS and RHS can be written as the product of a matrix with its transpose as follows:
\[ A C A + B C B - A C B - B C A = (A - B) C (A - B) = (A - B) C^{0.5} (A - B) C^{0.5}^T, \]

which is clearly positive definite.
2.2 Proof of Lemma 2.2

Since \( B - A \) is positive definite, we can assume \( B - A = H H^T \) for some matrix \( H \). Therefore, \( CBC - CAC = C(B - A)C = (CH)(CH)^T \) is positive definite.

3 Generalization to \( n = 3k \)

It is possible to extend the proof from the case \((n, m) = (3, 3)\) to the cases where \( m = 3 \) and \( n = 3k \) for some positive integer \( k \). The proof follows directly from the following observation.

Lemma 3.1. If (1.2) holds for \((n, m) = (n_0, m_0)\), then it also holds for \((n, m) = (kn_0, m_0)\) with any positive integer \( k \).

Proof. If \( n = kn_0 \) and \( m = m_0 \), then

\[
E[A_{i_1} A_{i_2} \cdots A_{i_m}] = \frac{1}{k^m} E \left[ \sum_{j=1}^{k} A_{i_j} \cdots \sum_{j=(m-1)k+1}^{2k} A_{i_j} \cdots \sum_{j=(m-1)k+1}^{3k} A_{i_j} \cdots \sum_{j=(m-1)k+1}^{mk} A_{i_j} \right]
\]

\[
= \frac{1}{k^m} E \left[ \sum_{l_1}^{l_1k} A_{i_j} \cdots \sum_{j=(l_1-1)k+1}^{l_2k} A_{i_j} \cdots \sum_{j=(l_2-1)k+1}^{l_3k} A_{i_j} \cdots \sum_{j=(l_m-1)k+1}^{l_mk} A_{i_j} \right]
\]

where \( \{l_1, l_2, \cdots, l_n\} \) is a random permutation of \( \{1, 2, \cdots, n\} \).

Apply (1.2) with \((n, m) = (n_0, m_0)\) to \( n_0 \) positive definite matrices \( \{ \sum_{j=(l-1)k+1}^{l_k} A_{i_j} \}_{l=1}^{n_0} \), we have

\[
-\frac{1}{n_0} I \leq \mathbb{E}_{l_1, l_2, \cdots, l_{n_0}} \left[ \sum_{j=(l_1-1)k+1}^{l_1k} A_{i_j} \cdots \sum_{j=(l_2-1)k+1}^{l_2k} A_{i_j} \cdots \sum_{j=(l_m-1)k+1}^{l_mk} A_{i_j} \right] \leq \frac{1}{n_0} I
\]

(3.2)

Combining (3.1) and (3.2), we proved (1.2) for \((n, m) = (kn_0, m_0)\).

We remark that since the conjecture for \((n, m) = (2, 2)\) has been proved in [3, Proposition 3.2], Lemma 3.1 implies that the conjecture also holds for \((n, m) = (2k, 2)\) when \( n \) is even.

References

[1] J. C. Duchi. Commentary on "toward a noncommutative arithmetic-geometric mean inequality: Conjectures, case-studies, and consequences". In S. Mannor, N. Srebro, and R. C. Williamson, editors, COLT, volume 23 of JMLR Proceedings, pages 11.25–11.27. JMLR.org, 2012.

[2] A. Israel, F. Krahmer, and R. Ward. An arithmetic-geometric mean inequality for products of three matrices. arXiv preprint arXiv:1411.0333, 2014.
[3] B. Recht and C. Ré. Beneath the valley of the noncommutative arithmetic-geometric mean inequality: conjectures, case-studies, and consequences. \textit{arXiv preprint arXiv:1202.4184}, 2012.