Failure Rates Of The Matrix Triangle Inequality

Osman Jime and Jordan Kaseram

Abstract
The triangle inequality is a fundamental principle in mathematics. However, its validity is not always guaranteed in matrix theory, specifically, when dealing with the matrix absolute value. We investigate how often the matrix absolute value satisfies the triangle inequality for different spaces of matrices. By using numerical methods, we quantify the frequency with which the triangle inequality holds and examine possible key factors contributing to its success or failure.

1 Introduction

The triangle inequality is a foundational concept in mathematics, offering key insights into the relationships between distances within various mathematical structures. At its core, the triangle inequality asserts that for any triangle, the sum of the lengths of any two sides must be greater than or equal to the length of the remaining side. This geometric intuition applies in particular to algebraic structures, such as the real number line or the complex plane, where it invariably holds that for any $x, y \in \mathbb{C}$:

$$|x + y| \leq |x| + |y|.$$

As we expand our focus to more complicated algebraic systems such as $\mathbb{M}_n(\mathbb{C})$, the $n \times n$ matrices with complex entries, a natural question arises: Does the triangle inequality still hold in these higher-dimensional structures? Bhatia [1] and many others have noted that this fundamental inequality is not always true for the so-called *matrix absolute value*: $|A| = \sqrt{A^*A}$, a matrix with non-negative eigenvalues.

The conditions under which the triangle inequality fails for matrices are of particular interest, given their implications for the analysis of boundedness and convergence in matrix theory. This exploration provides a deeper understanding of how the structural properties of matrices influence the behavior of the absolute value in relation to the triangle inequality.

To investigate these phenomena, a thorough numerical study was conducted across specific sets, including:

- $\mathbb{M}_n(\mathbb{Z} \cap [-k, k])$,
- $\mathbb{M}_n(\{x + iy : x, y \in \mathbb{Z} \cap [-k, k]\})$,
- $\mathbb{M}_n(\mathbb{R} \cap [-k, k])$, and
- $\mathbb{M}_n(\{x + iy : x, y \in \mathbb{R} \cap [-k, k]\})$

where k is the specific entry range of interest. Using an error tolerance of 10^{-7}, for approximately zero values, the frequency of triangle inequality failures was determined within these sets. This

Vol. 8(1) | DOI: https://doi.org/10.31542/bkyzts57
  | ISSN 2369-5617

tolerance was necessary since all entry types can lead to floating-point calculation errors. The restriction of entry ranges provides a controlled environment, offering valuable insights into how the triangle inequality behaves in relation to matrix absolute values.

The structure of the paper is as follows: Section 2 discusses norms, absolute values extended on matrices, and why failures occur for the matrix absolute value. In Section 3 the chances of the triangle inequality holding were almost nonexistent yet in $\mathbb{M}_2$ with integer or real entries, the frequency of matrix pairs failing is approximately 60%; for Gaussian integer and complex numbers, this percent of failures is approximately 41%. Then, finally, Section 4 explores potential reasons for when the triangle inequality holds and the infrequency of these pairs occurring.

2 Background and Preliminaries

Human beings intuitively understand the difference between something that is "near" and "far", or "long" and "short". These concepts play a role in how we understand the world around us. The most formal description of how we understand length is as a mathematical function, $\|\cdot\| : V \to \mathbb{R}$, called a **norm** where V is a vector space over the field $\mathbb{C}$. Norms satisfy the following properties [4] for all $x, y \in V$ and all $a \in \mathbb{C}$:

(1a) $\|x\| \geq 0$ *(Negative)*

(1b) $\|x\| = 0$ if and only if $x = 0$ *(Positive)*

(2) $\|ax\| = |a|\,\|x\|$ *(Homogeneous)*

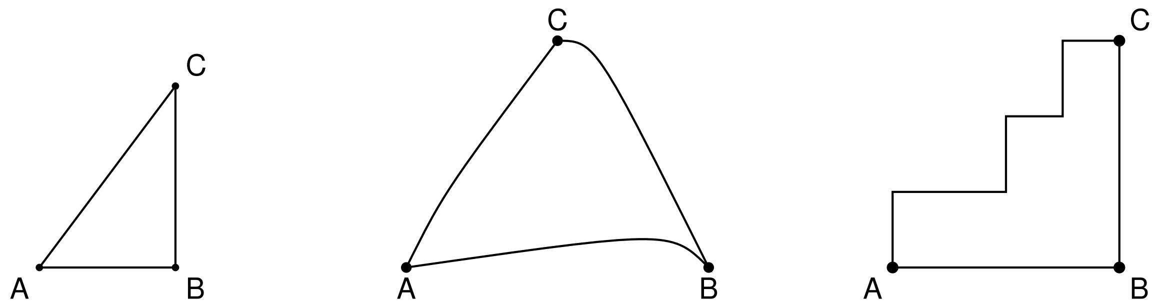
(3) $\|x + y\| \leq \|x\| + \|y\|$ *(Triangle Inequality)*

The norm can be applied to a variety of abstract objects, acting as a powerful tool that allows one to explore and navigate an abstract space as we would in a physical one. In quantum mechanics, the Uncertainty Principle, which states that an object's position and momentum can not be simultaneously known, is ultimately derived from an inequality of norms in $\mathbb{C}^n$ [10, p.140-148] . In numerical analysis, matrix norms are used when determining condition numbers, the ratio of the norm A to A^{-1}, which measures the sensitivity of a solution to perturbations [6], a significant result since ill-conditioned systems are prone to large numerical errors. In pure mathematics, having a concept of length, makes it possible to determine the convergence of sequences [8].

Arguably the most important property of a norm is the third property, commonly known as the **triangle inequality**, as it spatially captures how we expect a length to behave. The triangle inequality gives rise to the comparison of different paths, the shortest distance between two points, and in $\mathbb{R}^2$, the right angle triangle inequality also inspires the canonical Euclidean Norm, $\|\cdot\| : \mathbb{R}^2 \to \mathbb{R}, \|x\| = \sqrt{x \cdot x}$ (Figure 1).

Taking the square root of the product of something with itself is a common operation defined for any vector space equipped with an inner product. For example, a norm defined in $\mathbb{C}$,
$|\cdot| : \mathbb{C} \to \mathbb{R}, |x| = \sqrt{x^*x}$, for all $x \in \mathbb{C}$, gives rise to the absolute value function. In fact, any function $|\cdot| : V \to \mathbb{V}$, with this structure, is commonly referred to as the "absolute value" or "modulus" in V, relating elements of V to their non-negative counterparts in V.

However, a function that behaves like an absolute value does not necessarily satisfy all the properties of a norm. The set of matrices, $\mathbb{M}_n$ is itself a vector space of dimension n^2 and the



(a) Euclidean Norm (b) Some Random Norm (c) Taxicab Norm

Figure 1: Comparison of different norms on $\mathbb{R}^2$

size of a matrix can be measured by using any norm on $\mathbb{C}^{n^2}$ [4]. However, the absolute value in $\mathbb{M}_n$ fails the triangle inequality property. Consequently, the matrix absolute value fails to define a norm.

To fully understand why the matrix absolute value fails to be a norm, we begin with the following definitions in matrix theory:

Definition 2.1. *A matrix A is* ***Hermitian*** *(or* ***self-adjoint****) if and only if A equals its conjugate transpose, denoted A^*.*

Definition 2.2. *A matrix $A \in \mathbb{M}_n(\mathbb{C})$ is* ***positive semidefinite*** *(PSD) if and only if $x^*Ax \geq 0$ for all $x \in \mathbb{C}^n$. Such matrices are equivalent to being Hermitian with non-negative eigenvalues.*

Definition 2.3. *Given Hermitian matrices $A, B \in \mathbb{M}_n(\mathbb{C})$, if $B - A$ is PSD then it is denoted by $A \leq B$. This defines the usual partial order on matrices known as the* ***Loewner order*** *[4].*

Remark 2.4. *The relation "$\leq$" denotes a* ***partial order*** *on matrices. This means that certain pairs of matrices may be ordered relative to each other but not all matrix pairs are necessarily comparable. A partial order on matrices $A, B, C \in \mathbb{M}_n(\mathbb{C})$ must satisfy the following properties [4]:*

(i) Reflexive: $A \leq A$

(ii) Anti-symmetric: If $A \leq B$ and $B \leq A$ then $A = B$

(iii) Transitive: If $A \leq B$ and $B \leq C$ then $A \leq C$.

It is well-known that the Loewner order ($\leq$) is not a total order. An example of non-comparable matrix pairs is given by:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \nleq \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \nleq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Definition 2.5 (Horn & Johnson, Exercise 1.2 p.7, [4])**.** *A matrix $A \in \mathbb{M}_n$ has a* ***square root*** *$B \in \mathbb{M}_n$ if $A = B^2$. Moreover, if A is PSD, then there is a unique PSD B, denoted $B = \sqrt{A}$.*

Just as the absolute value for the field $\mathbb{C}$ can be expressed as a square root function: $|x| = \sqrt{\overline{x}x}$, the absolute value in $\mathbb{M}_n(\mathbb{C})$ has a similar form.

Definition 2.6. *For $A \in \mathbb{M}_n(\mathbb{C})$ we say that $|A| = \sqrt{A^*A}$ is the* ***matrix absolute value*** *(or* ***matrix modulus****). Such matrices are PSD.*

The usual method to calculate the matrix absolute value is through the spectral decomposition algorithm. This representation involves applying the square root to the diagonal entries of matrix D in the decomposition, $A^*A = UDU^*$, giving:

$$|A| = (A^*A)^{\frac{1}{2}} = UD^{\frac{1}{2}}U^*,$$

and it is noted that A^*A is a PSD matrix.

Lemma 2.7 (Horn & Johnson, 2.4.1 p.108 [4])**.** *Let $A \in \mathbb{M}_n$. Then det$A = \prod_i^n \lambda_i$ and tr$A = \sum_i^n \lambda_i$.*

Remark 2.8. *For any Hermitian matrix, $A \in \mathbb{M}_2$, if the trace and determinant A are non-negative, then by the property of Lemma 2.7 the eigenvalues of A are non-negative and A is PSD.*

In $M_2(\mathbb{C})$ matrices can be quickly calculated using the following short cut:

Theorem 2.9 (Horn & Johnson, Exercise 7.3 p.26 [4])**.** *Let $A \in \mathbb{M}_2(\mathbb{C})$ be PSD, and nonzero. Let*

$$\tau = \left(\mathrm{tr}A + 2\sqrt{\det A}\right)^{\frac{1}{2}}$$

Then the square root of a matrix is the following:

$$A^{\frac{1}{2}} = \tau^{-1}(A + \sqrt{\det A}I_2).$$

Proof. Since A is PSD, it has the general form $A = \begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix}$, where $a, c \in \mathbb{R}$ and $b \in \mathbb{C}$. Then:

$$A^{\frac{1}{2}} = \frac{1}{\sqrt{a + c + 2\sqrt{ac - b\bar{b}}}} \begin{bmatrix} a + \sqrt{ac - b\bar{b}} & b \\ \bar{b} & c + \sqrt{ac - b\bar{b}}, \end{bmatrix}$$

which is clearly Hermitian. By squaring this result, we obtain the following:

$$\begin{aligned}
\left(A^{\frac{1}{2}}\right)^2 &= \left(\frac{1}{\sqrt{a + c + 2\sqrt{ac - b\bar{b}}}} \begin{bmatrix} a + \sqrt{ac - b\bar{b}} & b \\ \bar{b} & c + \sqrt{ac - b\bar{b}} \end{bmatrix}\right)^2 \\
&= \frac{1}{a + c + 2\sqrt{ac - b\bar{b}}} \begin{bmatrix} a + \sqrt{ac - b\bar{b}} & b \\ \bar{b} & c + \sqrt{ac - b\bar{b}} \end{bmatrix}^2 \\
&= \frac{1}{a + c + 2\sqrt{ac - b\bar{b}}} \begin{bmatrix} a^2 + ac + 2a\sqrt{ac - b\bar{b}} & ab + ba + 2b\sqrt{ac - b\bar{b}} \\ a\bar{b} + \bar{b}c + \bar{b}\sqrt{ac - b\bar{b}} & c^2 + ac + 2c\sqrt{ac - b\bar{b}} \end{bmatrix} \\
&= \frac{1}{a + c + 2\sqrt{ac - b\bar{b}}} \begin{bmatrix} a(a + c + 2\sqrt{ac - b\bar{b}}) & b(a + c + 2\sqrt{ac - b\bar{b}}) \\ \bar{b}(a + c + 2\sqrt{ac - b\bar{b}}) & c(a + c + 2\sqrt{ac - b\bar{b}}) \end{bmatrix} \\
&= \begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix} \\
&= A.
\end{aligned}$$

By Remark 2.8, we need only check the trace and determinant to determine the matrix positive definiteness. Note that $\mathrm{tr}A \geq 0$ and $\det A \geq 0$ since A is PSD. It follows then that,

$$\begin{aligned}
\mathrm{tr}A^{\frac{1}{2}} &= \frac{1}{\sqrt{\mathrm{tr}A + 2\sqrt{\det A}}} \cdot \left((a+c) + 2\sqrt{ac - b\bar{b}}\right) \\
&= \frac{1}{\sqrt{\mathrm{tr}A + 2\sqrt{\det A}}} \cdot \left(\mathrm{tr}A + 2\sqrt{\det A}\right) \\
&= \sqrt{\mathrm{tr}A + 2\sqrt{\det A}} \\
&\geq 0.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\det A^{\frac{1}{2}} &= \frac{1}{\mathrm{tr}A + 2\sqrt{\det A}} \cdot \left(2\left(ac - b\bar{b}\right) + (a+c)\sqrt{ac - b\bar{b}}\right) \\
&= \frac{1}{\mathrm{tr}A + 2\sqrt{\det A}} \cdot \left(2\det A + \mathrm{tr}A\sqrt{\det A}\right) \\
&= \sqrt{\det A} \\
&\geq 0.
\end{aligned}$$

□

This matrix square root formula for a 2×2 matrix can be derived using Schur's or spectral decomposition, as the characteristic polynomials, $C_A(\lambda)$, of matrices with degree $n \leq 4$ have radical roots, enabling the existence of such formulas. However, as Eberhard [3] conjectured for matrices $\mathbb{M}_n$ with entries from a finitely supported distribution in $\mathbb{Z}$, the polynomial $C_A(\lambda)$ becomes irreducible with high probability as n becomes arbitrarily large. Since polynomials with degree $n \geq 5$ are generally irreducible by radicals as discussed in [Hungerford [5], Section 12.3]; therefore, a matrix square root formula is unlikely for $n \geq 5$.

The concept of matrix absolute value is closely tied to the partial order defined on the space of PSD matrices. Since the Hermitian matrices are a partially ordered set, there are instances where matrices cannot be directly compared, therefore, $|AB| \not\leq |A||B|$ and $|A+B| \not\leq |A| + |B|$ for $A, B \in \mathbb{M}_n$ generally. However, Mortad [9, Theorem 2.2] has shown that the equality can hold under certain conditions, specifically: $|AB| = |A||B|$, when both A and B are Hermitian such that the matrix AB is normal.

3 Triangle Inequality Computations

To determine whether the triangle inequality has failed, one must examine the difference matrix: $\Delta := |A| + |B| - |A+B|$. The failure is confirmed if this matrix is not PSD, which can be verified by analyzing its eigenvalues or checking the determinants of its principal minors. We will use Remark 2.8 to determine if a 2-dimensional matrix is PSD. Let us first consider an example where the matrix triangle inequality failure occurs.

Example 3.1. *Consider:*

$$A = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We wish to verify that: $|A + I_2| \leq |A| + |I_2|$ is not true. It is noted that $\sqrt{I_2} = I_2$ and using Theorem 2.9 allows simple computation of the matrix absolute value for $A + I_2$ and A.

$$\begin{aligned} |A + I_2| &= \left| \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = \left| \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} \right| \\ &= \left(\begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} \right)^{\frac{1}{2}} = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}^{\frac{1}{2}} \\ &= \frac{1}{\sqrt{8}} \begin{bmatrix} 2 & -2 \\ -2 & 6 \end{bmatrix} \\ &= \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} & \frac{3\sqrt{2}}{2} \end{bmatrix} \end{aligned}$$

Similarly,

$$\begin{aligned} |A| &= \left| \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \right| \\ &= \left(\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \right)^{\frac{1}{2}} = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}^{\frac{1}{2}} \\ &= \frac{1}{2} \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \end{aligned}$$

Therefore, the difference matrix is given by:

$$\Delta := |A| + I_2 - |A + I_2| = \begin{bmatrix} 2 - \frac{\sqrt{2}}{2} & -1 + \frac{\sqrt{2}}{2} \\ -1 + \frac{\sqrt{2}}{2} & 2 - \frac{3\sqrt{2}}{2} \end{bmatrix}.$$

Finally, we check $tr\Delta, \det \Delta \geq 0$:

$$tr\Delta = 4 - 2\sqrt{2} \approx 1.17$$

and

$$\det \Delta = 4 - 3\sqrt{2} \approx -0.234.$$

Since the determinant is negative, then the difference matrix is not PSD and consequently, the triangle inequality fails for the chosen matrices A and I_2.

3.1 Computational Methods

The Triangle Inequality of the matrix absolute value was investigated using Python. Three key tasks were implemented into their own separate functions for this paper.

1. Verifying a Triangle Inequality
2. Computing the Matrix Absolute Value of any Matrix, $A \in \mathbb{M}_n$
3. Verifying the Positivity of Any Matrix, $A \in \mathbb{M}_n$

The complete Python project can be found in the Github repository referenced in this paper [7]. In this section we provide the pseudocode for each key function, without the usage of helper functions that would be found in the complete project.

(i) Implementing a test for the Triangle Inequality The procedure of testing the triangle inequality could have simply been included in our statistics scripts, but we wanted the flexibility of being able to test different functions as lengths, as well as changing the notion of positivity if deemed necessary.

Algorithm 1 Check Triangle Inequality, `triangle_inequality()`

Require: Two elements, $A, B \in V$, for V a vector space over $\mathbb{C}$
Require: A function to act as length, $d : V \to W$
Require: A function to act as positivity, $positive : W \to$ `bool`
Compute lengths $d(A)$, $d(B)$, $d(A+B)$
difference = $d(A) + d(B) - d(A+B)$
if $positive$(difference) $==$ `true` **then**
  return True
else
  return False
end if

(ii) Implementing the Matrix Absolute Value We defined a python function for the matrix absolute value, which would serve as our length function in our triangle inequality function. The absolute value of a matrix was implemented in two ways, where the first implementation determines the square root of A^*A using the Spectral Decomposition. This way is considerably faster and it exploits the fact that A^*A is always diagonalizable. The second implementation uses a SciPy function, `scipy.linalg.absm`, to determine the square root using the Schur Decomposition. This method is the most numerically stable way of computing a square root of any matrix regardless of whether it is diagonalizable although it is not unique.[2].

Algorithm 2 Compute Matrix Absolute Value (Spectral), `matrix_absolute_value()`

Require: Single Hermitian matrix, $A \in \mathbb{M}_n$,
Compute A^*A
Compute U, D of the spectral decomposition $A^*A = UDU^*$
Compute $D^{\frac{1}{2}}$ by taking the square root each element of D
Compute $|A| = UD^{\frac{1}{2}}U^*$
return $|A|$

(iii) Implementing a test for Matrix Positivity The matrix positivity was implemented in two different ways. The first implementation of determining if a matrix is PSD, solves for all the eigenvalues, and then asserts the minimum eigenvalue is greater than or equal to zero.

Algorithm 3 Check Matrix Positivity (Eigenvalues), `matrix_positivity()`

Require: Single Hermitian matrix, $A \in \mathbb{M}_n$,
Require: Tolerance, `tol`,
Compute $\sigma(A)$, the set of all eigenvalues of A
Determine $\lambda_{\min} = \min(\sigma(A))$
if $\lambda_{\min} >$ `-tol` **then**
  return True
else
  return False
end if

The second implementation for matrices in $\mathbb{M}_2$ utilizes the property of Lemma 2.7 to check for positive definiteness.

Algorithm 4 Check Matrix Positivity of $\mathbb{M}_2$ (Trace and Determinant), `matrix_positivity()`

Require: Single matrix, $A \in \mathbb{M}_2$,
Require: Tolerance, `tol`,
Compute tr$A = A[0,0] + A[1,1]$, the Trace of A
Compute det$A = (A[0,0] \times A[1,1]) - (A[0,1] \times A[1,0])$, the determinant of A
if $|$tr$A| <$ `tol` **and** $|$det$A| <$ `tol` **then**
  return True
else if tr$A >$ `tol` **and** det$A >$ `tol` **then**
  return True
else
  return False
end if

Each implementation requires a tolerance to determine if the minimum eigenvalue (or trace/determinant) was approximately zero. If the minimum eigenvalue is minuscule, then checking the determinant might be less likely to be near our tolerance as it is the product of all eigenvalues. If all the eigenvalues are fractional, then directly checking the minimum eigenvalue will be less likely to be near our tolerance. In the complete project, alternating between different implementations of determining PSD and of the matrix absolute value were useful in demonstrating that our results were not sensitive to differences in numerical stability and reproducible through different approaches [7].

3.2 Results

The percentage of Triangle Inequality failures was investigated for matrix pairs, $A, B \in \mathbb{M}_n$ with varying entry types and dimension. For real ($\mathbb{M}_n(\mathbb{R})$) or integer ($\mathbb{M}_n(\mathbb{Z})$) type entries, the values range from $[-k, k]$, $k \in \mathbb{Z}$. For Gaussian integer ($\mathbb{M}_n(\mathbb{Z}[i])$) or complex ($\mathbb{M}_n(\mathbb{C})$) type entries, the values are of the form $x + iy$, where x, y range from $[-k, k]$, $k \in \mathbb{Z}$. Sample failure percentages for each entry types were calculated from 100,000 randomly chosen pairs within a fixed entry range $[-k, k]$. True failure percentages were also calculated for pairs $A, B \in \mathbb{M}_2(\mathbb{Z})$ by exhaustively determining all possible pairs with entry ranges $[-k, k]$, $k \geq 7 \in \mathbb{Z}$. (Figure 2)

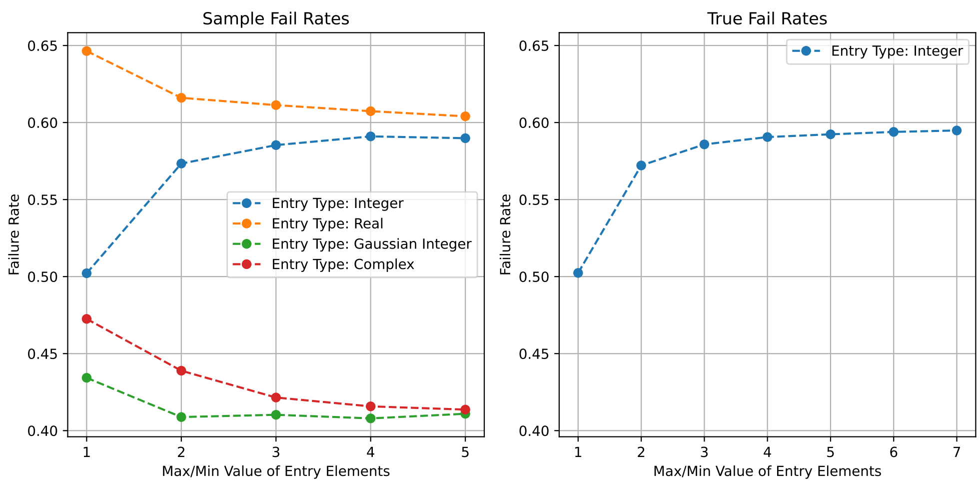


Figure 2: Comparison of sample and true triangle inequality failure rates of $\mathbb{M}_2$. Each sample point contains 100,000 randomly chosen matrix pairs, with standard tolerance

The effect of increasing the dimensions of the matrix pairs was also investigated. Sample Failure Rates for all entry types at a standard tolerance of 10^{-7} were produced for $2 \leq n \leq 4$. (Figure 3)

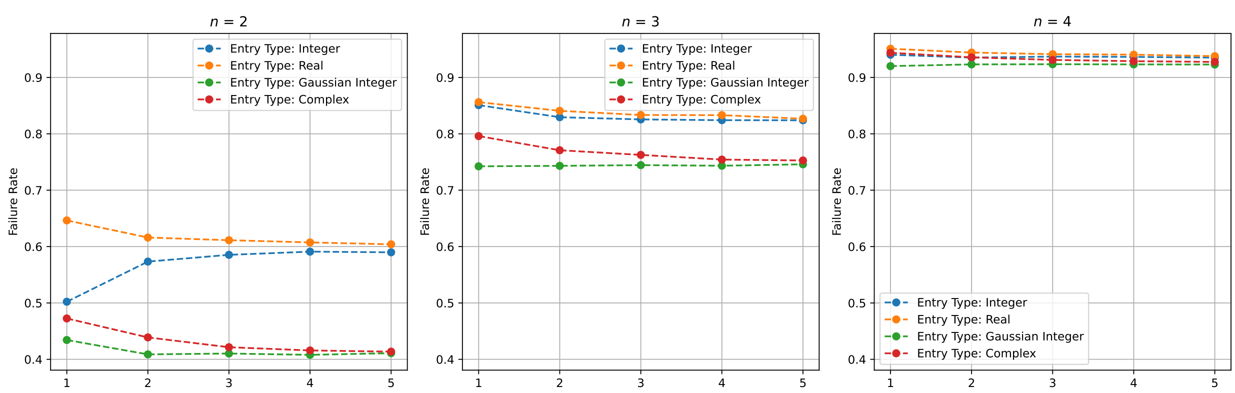


Figure 3: Changes in fail rate displaying a clear trend to an extremely high failure rate as n increases. Each sample point contains 100,000 randomly chosen matrix pairs, with standard tolerance.

Conjecture 3.2. *The rate of failure of the matrix triangle inequality for*

$\mathbb{M}_2(\mathbb{Z}\cap[-k,k])$ *and* $\mathbb{M}_2(\mathbb{R}\cap[-k,k])$ *is approximately* 60%.

Conjecture 3.3. *The rate of failure of the matrix triangle inequality for* $\mathbb{M}_2(\{x+iy : x,y\in\mathbb{Z}\cap[-k,k]\})$ *and* $\mathbb{M}_2(\{x+iy : x,y\in\mathbb{R}\cap[-k,k]\})$ *is approximately* 41%.

Conjecture 3.4. *For* $\mathbb{M}_n(\mathbb{C})$*, as* $n\to\infty$*, the matrix triangle inequality fails with high probability.*

4 Possible Reasons for Success

In this section, we discuss possible reasons for valid matrix triangle inequality pairs.

Proposition 4.1. *If* $A=B$ *for* $A,B\in\mathbb{M}_n(\mathbb{C})$ *then the triangle inequality will be valid.*

Proof. $|A+A|=|2A|=2|A|=|A|+|A|$ □

Proposition 4.2. *If* $B=-A$ *for* $A,B\in\mathbb{M}_n(\mathbb{C})$*, then the triangle inequality will be valid.*

Proof. It is noted that $\sqrt{B^*B}=\sqrt{-A^*(-A)}=\sqrt{A^*A}$. This shows that $|A|=|B|$. Therefore,

$$|A+B|=0\leq 2|A|=|A|+|B|.$$

□

Remark 4.3. *From the above proposition, if we include that* A *is invertible then it is possible to get an even stronger conclusion. The matrix square root function is continuous and by continuity, if a matrix* B *is "close" to* $-A$ *then the triangle inequality will still be true. The probability of* A *being invertible is* 1 *when working over* $\mathbb{C}$*, which then implies that there is always a positive probability for success, even if it is small.*

Proposition 4.4. *If* A *and* B *are diagonal for* $A,B\in\mathbb{M}_n(\mathbb{C})$*, then the triangle inequality will be valid.*

Proof. Let $A=\operatorname{diag}(a_{ii})$ and $B=\operatorname{diag}(b_{ii})$ for $a_{ii},b_{ii}\in\mathbb{C}$. Then, $A^*A=\operatorname{diag}\left(|a_{ii}|^2\right)$ and $B^*B=\operatorname{diag}\left(|b_{ii}|^2\right)$, so it follows that: $|A|=\operatorname{diag}(|a_{ii}|)$ and $|B|=\operatorname{diag}(|b_{ii}|)$. Similarly, $|A+B|=|\operatorname{diag}(a_{ii}+b_{ii})|=\sqrt{\operatorname{diag}(|a_{ii}+b_{ii}|)^2}=\operatorname{diag}(|a_{ii}+b_{ii}|)$. Then the difference matrix is given by:

$$|A|+|B|-|A+B|=\operatorname{diag}(|a_{ii}|+|b_{ii}|-|a_{ii}+b_{ii}|).$$

Since the triangle inequality holds for the modulus in $\mathbb{C}$, we have $|a_{ii}+b_{ii}|\leq|a_{ii}|+|b_{ii}|$ for each diagonal entry. Therefore, every eigenvalue is non-negative which implies that the matrix is PSD. □

While these cases might be obvious, Mortad also discusses less obvious cases where the triangle inequality is satisfied.

Theorem 4.5 (Mortad [9])**.** *Let* $A,B\in\mathbb{M}_n(\mathbb{C})$ *such that:*

1. $AA^*=A^*A$ *(Normal)*

2. $BB^* \leq B^*B$ *(Hyponormal)*

3. $AB = BA$ *(Commutative),*

then $|A + B| \leq |A| + |B|$.

Example 4.6. *Let*

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}$$

It is clear that A *and* B *are commutative, and* A *is normal. We check that* B *is hyponormal as follows:*

$$\begin{aligned} B^*B - BB^* &= \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} - \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \\ &= 0 \end{aligned}$$

Thus, B *is normal; therefore, by Theorem 4.5 the triangle inequality results will be valid. This claim is verified with*

$$|A| = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, |B| = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}, \text{ and } |A + B| = \begin{bmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix}$$

which produces the difference matrix:

$$|A| + |B| - |A + B| = \begin{bmatrix} 1 + \sqrt{2} - \sqrt{5} & 0 \\ 0 & 1 + \sqrt{2} - \sqrt{5} \end{bmatrix} \approx \begin{bmatrix} 0.178 & 0 \\ 0 & 0.178 \end{bmatrix}.$$

This matrix is clearly positive semidefinite, which shows the triangle inequality holds.

While these matrix pairs will always satisfy the triangle inequality, they constitute only a small fraction of the total pairs. An exhaustive search of the set $\mathbb{M}_n(\mathbb{Z} \cap \{-1, 0, 1\})$ revealed that 337 pairs satisfied Theorem 4.5 out of a total of 3265 pairs tested, accounting for approximately 10.3% of successful cases. When the range is expanded to include integers between -2 and 2, 3345 such pairs were found; however, their frequency dropped to 2%. Further increasing the integer range caused this frequency to fall dramatically to less than 1%.

A random search method produced similar results for integers when the same entry range was applied to a sample of $100,000$ random matrix pairs. For matrices with complex integer entries between -1 and 1, the success rate was approximately 0.24% and this dropped to 0% as the entry range increased. However, this data does not imply that such pairs cease to exist, as $\{-1, 0, 1\} \subseteq \{-2, -1, 0, 1, 2\}$ and successful pairs were found in the -1 to 1 range. Instead, it suggests that these pairs become exceedingly sparse.

Similarly, when considering matrices with real or complex number entries, the percentage of successful pairs was found to be 0%.

These results may seem to be a numerical anomaly at first, however, the probability that a random matrix pair commutes is absolutely 0. This idea is because if we fix the matrix A, the set of B matrices that commutes with it is a linear subspace, namely the kernel for the linear mapping

$B \mapsto AB - BA$. Thus, unless A commutes with everything, then the set of B matrices which commute with A has zero volume. This behavior explains the frequency we saw when considering matrices with entries in $\mathbb{R}$ and $\mathbb{C}$, since Theorem 4.5 requires commuting matrix pairs.

Another scenario where the triangle inequality is guaranteed to hold is described by the following theorem:

Theorem 4.7 (Mortad [9]). *Let $A, B \in \mathbb{M}_n(\mathbb{C})$ such that:*

1. $AA^* = A^*A$ *(Normal)*
2. $AB = BA$ *(Commutative)*
3. $A^*B + B^*A \leq 0$,

then $|A + B| \leq |A| + |B|$.

Example 4.8 (Using Theorem 4.7). *Consider:*

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$$

Clearly A and B commute, and checking the inequality results in:

$$\begin{aligned} A^*B + B^*A &= \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix} \\ &\leq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Then the triangle inequality is guaranteed to succeed. We verify with the results with the following: $|B| = \begin{bmatrix} \frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} & \frac{3\sqrt{5}}{5} \end{bmatrix}$ *and* $|A + B| = \begin{bmatrix} \frac{3\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \end{bmatrix}$ *and with*

$$\begin{aligned} |A| + |B| - |A + B| &= \begin{bmatrix} 1 - \frac{\sqrt{5}}{5} & 0 \\ 0 & 1 + \frac{\sqrt{5}}{5} \end{bmatrix} \approx \begin{bmatrix} 0.553 & 0 \\ 0 & 1.447 \end{bmatrix} \\ &\geq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

By using an exhaustive search method for matrices with integer entries between -1 and 1, we found 377 pairs that satisfy this theorem, representing approximately 11.5% of the total pairs that passed the triangle inequality. When the range of integer entries was increased to between -2 and 2, a total of 4145 pairs satisfied Theorem 4.7, accounting for about 2.5% of the total successful pairs. As the entry range increased further, the frequency of these matrix pairs dropped below 1%. The random search method yielded similar results when applied to a sample size of 100,000 random matrix pairs. For matrices with complex integer entries between -1 and 1, approximately

3.3% of the pairs satisfied the theorem. However, as the entry range increased, the frequency dropped to less than 1%. For matrices over the real and complex fields, the success rate was 0% across all tested entry ranges, likely due to this theorem requiring commuting matrix pairs, which have probability 0%, as previously discussed.

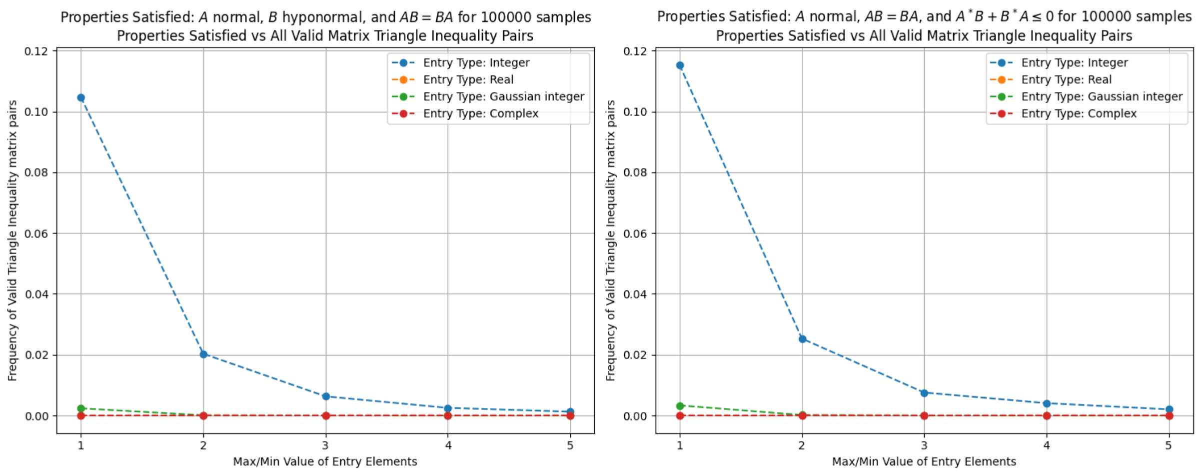


Figure 4: Random search method for matrix pairs that satisfy Theorems 4.5 and 4.7 for $A, B \in \mathbb{M}_2$.

For higher-dimensional matrices satisfying Mortad's theorems, we relied exclusively on the random search method due to hardware limitations. The number of matrix pairs satisfying these theorems was negligible compared to the total number of pairs that inherently satisfy the triangle inequality. Specifically, Mortad's pairs constituted less than 1% in $\mathbb{M}_3$ and 0% in $\mathbb{M}_n$ for $n \geq 4$. In $\mathbb{M}_3(\mathbb{Z} \cap \{-1, 0, 1\})$, the frequency for Theorem 4.5 was approximately 0.03%, while for Theorem 4.7, it was 0.13%. This comparison shows that Theorem 4.7 has more than four times as many pairs satisfying the triangle inequality as Theorem 4.5. A similar pattern was observed in $\mathbb{M}_2$, suggesting that Theorem 4.7 generally yields a higher frequency of valid matrix pairs across $\mathbb{M}_n$.

While Mortad's [9] theorems provide conditions under which the triangle inequality is valid, the matrix pairs that satisfy these theorems are sparse and contribute only minimally to the overall set of valid matrix pairs that uphold the triangle inequality.

Acknowledgements

Both authors would like to express their sincere gratitude to Dr. Micheal Francis and Dr. Christopher Ramsey for their invaluable insights and contributions to our discussions for this topic. Their expertise greatly enriched the development of this work. Both authors would like to express our sincere gratitude to the reviewer for their insightful comments and careful review, which greatly improved the clarity and quality of this paper. We also wish to acknowledge the land on which we gathered and learned is Treaty 6 territory and a traditional meeting ground and home for many Indigenous Peoples, including Cree, Saulteaux, Niitsitapi (Blackfoot), Métis, and Nakota Sioux Peoples. This project was supported by the NSERC Discovery Grant 2019-05430.

References

[1] Bhatia, R., Kittaneh, F. *The Matrix Arithmetic-Geometric Mean Inequality Revisited*, Linear Algebra and its Applications **428** (2008), 2177–2191, DOI 10.1016/j.laa.2007.11.030.

[2] Deadman, E., Higham, N., Ralha, R., *Blocked Schur Algorithms for Computing the Matrix Square Root*, Lecture Notes in Computer Science. **7782** (2013), 171-182, DOI 10.1007/978-3-642-36803-5_12.

[3] Eberhard, S. *The Characteristic Polynomial of a Random Matrix*, Combinatorica, **42** (2022), no. 4, 491-527, DOI 10.1007/s00493-020-4657-0.

[4] Horn, R. A., Johnson, C. R. *Matrix Analysis: Second Edition*, Cambridge university press. (2013).

[5] Hungerford, T.W. *Abstract Algebra: An Introduction*. Cengage Learning, Boston. (2012).

[6] Isaacson, E., Keller, H. B. *Analysis of Numerical Methods*, Courier Corporation. (1994).

[7] Jime, O., Kaseram, J. *Triangle Inequality Tests of Matrix Absolute Value, Version 1.0,* MacEwan University, Edmonton AB Canada (2024), https://github.com/ohjime/trieq-matrix-absolute-value.git.

[8] Li, C. *Norms, Isometries, and Isometry Groups*, The American Mathematical Monthly, **107** (2000), no. 4, pp. 334–40, DOI https://doi.org/10.1080/00029890.2000.12005201.

[9] Mortad, M.H. *On the Absolute Value of the Product and the Sum of Linear Operators*, Rend. Circ. Mat. Palermo, II. **68** (2019), no. 2, 247–257, DOI 10.1007/s12215-018-0356-8.

[10] Susskind, L. and Friedman, A. *Quantum Mechanics: The Theoretical Minimum*, The Theoretical Minimum Series, Basic Books, (2014).