Traces of Matrix Products

Zhanwen Huang

Department of Mathematics and Statistics

University of Minnesota, Duluth

Duluth, MN 55812

Advisor: John Greene

Department of Mathematics and Statistics

University of Minnesota, Duluth

Duluth, MN 55812

Traces of Matrix Products

Abstract

A formula for the number of trace equivalent classes for a matrix string of 2×2 matrices which is comprised of two different matrices A and B with k $A's$ and $n-k$ $B's$ is derived. Simulations for traces of matrix products with 2 $A's$ and n $B's$ for n varying from 2 to 10 are carried out. A comparison between traces of $ABAB$ and $AABB$ and their connection to the eigenvalues of individual matrix is discussed. A formula for a special case is given and a potential application in Statistical Physics is provided.

Key Words: Trace, Matrix Products, Trace Equivalent Class

1. Introduction

The trace of a product of matrices has been given extensive study and it is well known that the trace of a product of matrices is invariant under cyclic permutations of the string of matrices [1, P.76]. For example, $Tr(ABC)=Tr(BCA)=Tr(CAB)$ where BCA and CAB are cyclic permutations of ABC. Therefore, the three permutations are equivalent when we are interested in their traces, and we define permutations have the same trace as a trace equivalent class. Thus, for a string of matrices of length n, the actual number of trace equivalent class is much less than $n!$.

In this paper, we investigate the relative size of traces of matrix products. That is, if M and N are both products of n $A's$ and $B's$, and the only difference is the order of the factors, i.e., $M=ABABB$, $N=AABBB$, what can be said about $Tr(M)$ vs. $Tr(N)$? Furthermore, in this paper, we prove for a string of 2×2 matrices which is comprised of two different matrices A and B, the trace of the product of those matrices is invariant under reversal operations. If $M=M_1M_2...M_n$ is a product of matrices, then the reversal of this product is defined as $M^R=M_nM_{n-1}...M_1$. For example, the reversal of

$ABCD$ is $DCBA$; the reversal of $AABBABA$ is $ABABBAA$. Note the reversal of a matrix string generally cannot be obtained by a cyclic permutation of the original string. Therefore, for a 2×2 matrix string of length n which contains k $A's$ and $n-k$ $B's$, the number of trace equivalent class would be cut down further.

For example, for matrix strings contain $4A's$ and $4B's$, there are in total 8 trace equivalent classes, rather than $8!$ which is about 40 thousand, or even $\frac{8!}{4!4!}$ which is 70.

Consider the following table:

Product	Trace 1	Trace 2	Trace 3
A^4B^4	203	463	13721
$A^2B^2A^2B^2$	207	479	7889
$A^2BA^2B^3$	211	495	6593
$A^3B^2AB^2$	219	559	6593
A^3BAB^3	235	655	5009
A^2BAB^2AB	243	687	2057
A^2B^2ABAB	255	767	1769
$ABABABAB$	343	1471	257

Where in Trace 1, $A=\begin{pmatrix}1 & 1\\ 3 & 0\end{pmatrix}, B=\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}$,

in Trace 2, $A=\begin{pmatrix}1 & 1\\ 5 & 0\end{pmatrix}, B=\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}$, in Trace 3, $A=\begin{pmatrix}0 & 1\\ 2 & 3\end{pmatrix}, B=\begin{pmatrix}3 & 2\\ 1 & 0\end{pmatrix}$.

We might think there is an intrinsic rank for those trace equivalent classes, where A^4B^4 and $ABABABAB$ would either be in the trace equivalent class with greatest trace or smallest trace. However, this is not the case. Simulations show there are more than 2 types of orderings, namely other trace equivalent classes might have the largest trace for other values of A and B. However, although every trace equivalent class has the chance to have the greatest trace, the probabilities for different trace equivalent classes are quite different. If the entries of A and B obey normal or uniform distributions, the trace equivalent classes containing $ABABAB...$ appears to have a much higher probability to have the greatest trace than the trace equivalent classes containing $AA...BB...$, which seem most likely to have the smallest trace.

In Section 2, the theoretical background for this paper is presented. We prove in Theorem 1 that the trace of a product of 2×2 matrices comprised of two different matrices A and B equals to the trace of its reversal. Then we obtain a formula in Theorem 2 for the number of possible trace equivalent classes given k $A's$ and $n-k$ $B's$. Then we prove when the entries of $A's$ and $B's$ are random, the probability that $Tr(ABAB) > Tr(AABB)$ is the same as the probability that $Det(AB-BA) < 0$. And if we restrict the distribution of the entries of $A's$ and $B's$ to be uniform on $[-1,1]$, $Tr(ABAB)$ is more likely to be greater than $Tr(AABB)$.

In Section 3, data from simulations for traces of matrix products is presented and the analysis is given accordingly. We give the results for traces of matrix products with 2 $A's$ and n $B's$ for n varying from 2 to 10. Also we present the simulation results of the probability that $Tr(ABAB) > Tr(AABB)$ and show it is in a good agreement to the theoretical prediction in Section 2. A special case for the traces of matrix products is solved analytically and verified with simulations.

The applications of the traces of matrix products are mainly to Physics and Statistics. Jackson and Lautrup [5] studied the infinite product of 2×2 matrices with all entries drawn at random from a distribution of zero mean and unit variance. Due to the infinity product property, the law of large number is applicable and they obtained the fact that the determinant of the product matrix is log-normally distributed. They also pointed out a potential application in statistical imagine analysis.

Also, in Statistical Physics, products of random transfer matrices [3] describe both the physics of disordered magnetic systems and localization of electronic wave functions in random potentials. The Lyapunov exponent is an important parameter for the predictablility of a dynamical system [7], and if the system has a positive maximum Lyapunov exponent, then the system is chaotic. The trace of the product of matrices has an application in the calculation of Lyapunov exponents[8].

We expect our results could be used in the study of disordered one dimensional systems. P.S. Davis[4] showed that the random binary alloy can be expressed as a product of 2×2 random matrices. The asymmetry of the probability of different trace equivalent classes obtained in our paper might serve as a prediction of the configuration of the alloy, since thermodynamic systems always move from a state of low probability to a state of high probability.

2. Theoretical Background

In this chapter, the theoretical part of the thesis is presented. This includes some theorems on the trace of a product of matrices and a formula for the number of possible trace equivalent classes given $k\ A's$ and $n-k\ B's$.

Powers of a 2×2 matrix A can always be written as a linear combination of A and the identity matrix. This follows inductively from the following lemma.

2.1.1 Lemma 1

For any 2×2 matrix A, $A^2 = Tr(A)A - Det(A)I$, where I is the 2×2 identity matrix.

Proof: Assume $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $Tr(A)A=\begin{pmatrix} a(a+d) & b(a+d) \\ c(a+d) & d(a+d) \end{pmatrix}$ and $Det(A)I=\begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix}$. Thus $Tr(A)A-Det(A)I=\begin{pmatrix} a^2+bc & b(a+d) \\ c(a+d) & d^2+bc \end{pmatrix}=A^2$

□

2.1.2 Lemma 2

If A and B are square matrices of the same size then $Tr(AB)=Tr(BA)$.

Proof: This follows directly from the definition of matrix multiplication.

$$Tr(AB)=\sum_{i=1}^{m}\sum_{j=1}^{n}A_{ij}B_{ji}=\sum_{i=1}^{m}\sum_{j=1}^{n}B_{ji}A_{ij}=\sum_{j=1}^{n}\sum_{i=1}^{m}B_{ji}A_{ij}=Tr(BA).\Box$$

The invariance of trace under cyclic permutations is a consequence of this lemma. For example, by replacing B above by BC we will obtain $Tr(ABC)=Tr(BCA)$ immediately and by replacing A above by CA we will get $Tr(CAB)=Tr(BCA)$. By induction, we can prove this is true for any finite matrix string.

Next, we show that the trace of a matrix product is invariant under reversal.

2.1.3 Theorem 1

If $M_1M_2...M_n$ is a product of 2×2 matrices and each M_i is either A or B, then $Tr(M_1M_2...M_n)=Tr(M_nM_{n-1}...M_1)$.

Proof: We prove this theorem by induction on n.

The cases $n=1,2,3,4$ are oblivious due to the cyclic property of trace. For example, $Tr(AAB)=Tr(BAA)$ and $Tr(ABAB)=Tr(BABA)$ because BAA is a cyclic permutation of AAB and $BABA$ is a cyclic permutation of $ABAB$.

Assume the relation hold for all products of k $A's$ and $B's$. If M is a product of $k+1$ $A's$ and $B's$, then either it is a product of matrices without consecutive $A's$ ($B's$) or a product that contains an A^2 or a B^2.

For the first case, the $A's$ and $B's$ must alternate. If $k+1$ is even, then this product is of exactly $\frac{(k+1)}{2}$ of $A's$ and $\frac{(k+1)}{2}$ of $B's$, and the reversal of the product is a cyclic permutation of original product. If $k+1$ is odd, then there are either $\frac{k}{2}$ $A's$ or $\frac{k}{2}$ $B's$, and the reversal is the same as the original product. For example, $(ABABA)^R = ABABA$.

For the case where there are consecutive $A's$ or $B's$, without losing generality, we assume there are consecutive $A's$ at i^{th} and $(i+1)'st$ positions. Then the original product can be written as $M_1M_2...M_{i-1}AAM_{i+2}...M_{n-1}M_n$, and the reversal is $M_nM_{n-1}...M_{i+2}AAM_{i-1}...M_2M_1$.

Then by the Lemma 1, the trace of the original product is

$$Tr(M_1M_2...M_{i-1}AAM_{i+2}...M_kM_{k+1}) = Tr\{M_1M_2...M_{i-1}[Tr(A)A - Det(A)I]M_{i+2}...M_kM_{k+1}\}$$
$$= Tr(A)Tr(M_1M_2...M_{i-1}AM_{i+2}...M_kM_{k+1}) - Det(A)Tr(M_1M_2...M_{i-1}M_{i+2}...M_kM_{k+1}).$$

The trace of the reversal is

$$Tr(M_{k+1}M_k...M_{i+2}AAM_{i-1}...M_2M_1) = Tr\{M_{k+1}M_k...M_{i+2}[Tr(A)A - Det(A)I]M_{i-1}...M_2M_1\}$$
$$= Tr(A)Tr(M_{k+1}M_k...M_{i+2}AM_{i-1}...M_2M_1) - Det(A)Tr(M_{k+1}M_k...M_{i+2}M_{i-1}...M_2M_1).$$

Since $M_1M_2...M_{i-1}AM_{i+2}...M_kM_{k+1}$ is a product of length k, by assumption, $Tr(M_1M_2...M_{i-1}AM_{i+2}...M_kM_{k+1}) = Tr(M_{k+1}M_k...M_{i+2}AM_{i-1}...M_2M_1)$.

Similarly, $Tr(M_1M_2...M_{i-1}M_{i+2}...M_kM_{k+1}) = Tr(M_{k+1}M_k...M_{i+2}M_{i-1}...M_2M_1)$.

Therefore, the relation is true for $n = k+1$.

By the principle of induction, the trace of a product of 2×2 matrices of $A's$ and $B's$ always equals the trace of its reversal. □

Next we figure out the number of trace equivalence classes given k $A's$ and $n-k$ $B's$. Barry Dayton [5] gives a formula for the number of equivalent classes of n beads with c colors, which is given by $\frac{1}{n}\sum_{d|n}\phi\left(\frac{n}{d}\right)c^d$, where $\phi\left(\frac{n}{d}\right)$ is the Euler phi function[6, P.80]. However, this is not exactly what we want, since it gives the total for all possible values of k.

To calculate the number of trace equivalent classes, we use Burnside's theorem[6, P.490], which states, if G is a finite group of permutations on a set S, then the number of orbits of G on S is $\frac{1}{|G|}\sum_{\phi\in G}|fix(\phi)|$, where $fix(\phi)=\{i\in S \mid \phi(i)=i, \phi\in G\}$, namely the elements of S that are invariant under certain group actions.

The group G in the theorem is the collection of symmetries on matrix products that leave the trace invariant. These are the cyclic permutations and string reversal, which generate a dihedral group.

The number of $fix(\phi)$ comes from two parts, the contribution of cyclic permutations and that of reflections. Note that any reversal operation could be obtained by cyclic permutations and reflections, for example, $(ABBAAAB)^R = BAAABBA$, it can also be obtained by reflecting with respect to the dashed line of $ABB \vdots AAAB$, but $BAAABBA$ cannot be obtained by only taking cyclic permutations to $ABBAAAB$. Considering the contribution of cyclic permutations and that of reflections, if the length of the matrix string is n and there are k $A's$, the formula for the number of trace equivalent classes is given below:

If n is even and k is even, then the number of trace equivalent classes is $\frac{1}{2n}\left[n\binom{n/2}{k/2}+\sum_{d|\gcd(n,k)}\binom{n/d}{k/d}\varphi(d)\right]$, where $\varphi(d)$ is the Euler phi function.

If n is even and k is odd, then the number is given by $\frac{1}{2n}\left[n\binom{(n-2)/2}{(k-1)/2}+\sum_{d|\gcd(n,k)}\binom{n/d}{k/d}\varphi(d)\right]$.

If n is odd and k is even, then the number is given by $\frac{1}{2n}\left[n\binom{(n-1)/2}{k/2}+\sum_{d|\gcd(n,k)}\binom{n/d}{k/d}\varphi(d)\right]$.

If n is odd and k is odd, then the number is

$$\frac{1}{2n}\left[n\binom{(n-1)/2}{(k-1)/2}+\sum_{d|\gcd(n,k)}\binom{n/d}{k/d}\varphi(d)\right].$$

2.1.4 Theorem 2

If A and B are 2×2 matrices and M is a product of $A's$ and $B's$, then $M-M^R=c(AB-BA)$ where c is a scalar.

Proof: We will prove this theorem by induction.

If the length of M is 2, namely M is AA, AB, BA or BB, the relation is obviously true with c being 0, 1, -1, or 0, respectively.

Assuming the result is true for products of k $A's$ and $B's$, then when M is a product of $k+1$ $A's$ and $B's$, there are two possibilities.

First, if there is no repetition of A and B in the string, when $k+1$ is even, either $M=(AB)^{\frac{k+1}{2}}$ or $M=(BA)^{\frac{k+1}{2}}$, and

$$M-M^R=\pm\left[(AB)^{\frac{k+1}{2}}-(BA)^{\frac{k+1}{2}}\right]$$

$$=\pm\left[Tr(AB)(AB)^{\frac{k-1}{2}}-Det(AB)(AB)^{\frac{k-3}{2}}-Tr(BA)(BA)^{\frac{k-1}{2}}+Det(BA)(BA)^{\frac{k-3}{2}}\right]$$

$$=\pm\left\{Tr(AB)\left[(AB)^{\frac{k-1}{2}}-(BA)^{\frac{k-1}{2}}\right]-Det(AB)\left[(AB)^{\frac{k-3}{2}}-(BA)^{\frac{k-3}{2}}\right]\right\}.$$

Since $2\times\frac{k-1}{2}$ and $2\times\frac{k-3}{2}$ are both less than k, then by assumption, both $(AB)^{\frac{k-1}{2}}-(BA)^{\frac{k-1}{2}}$ and $(AB)^{\frac{k-3}{2}}-(BA)^{\frac{k-3}{2}}$ can be written in the form $c(AB-BA)$, and therefore, $M-M^R=c(AB-BA)$.

When $k+1$ is odd, then either $M=ABAB...ABA$ or $M=BABA...BAB$, and in this case, $M=M^R$ and therefore, $M-M^R=0$.

If there are consecutive $A's$ or $B's$ in the string, without losing generality, let $M = M_1 M_2 ... M_{i-1} AAM_{i+2} ... M_k M_{k+1} = M_a AAM_b$, then

$$
\begin{aligned}
M - M^R &= M_a AAM_b - M_b^r AAM_a^r \\
&= M_a \left[Tr(A) A - Det(A) I \right] M_b - M_b^r \left[Tr(A) A - Det(A) I \right] M_a^r \\
&= Tr(A) M_a AM_b - Det(A) M_a M_b - Tr(A) M_b^r AM_a^r + Det(A) M_b^r M_a^r \\
&= Tr(A) \left(M_a AM_b - M_b^r AM_a^r \right) - Det(A) \left(M_a M_b - M_b^r M_a^r \right) \\
&= Tr(A) \cdot c_1 (AB - BA) - Det(A) \cdot c_2 (AB - BA) \\
&= c(AB - BA)
\end{aligned},
$$

where c_1, c_2 and c are scalars.

Thus, by the principle of induction, the relation is true for any matrix product of $A's$ and $B's$. Note the scalar c here depends on A, B and the order of the matrices. For example, if $M = AABB$, then $M^R = BBAA$ and $M - M^R = Tr(A) Tr(B) (AB - BA)$, so here $c = Tr(A) Tr(B)$. However, if $M = ABAB$, then $M^R = BABA$ and $M - M^R = Tr(AB)(AB - BA)$, which indicates $c = Tr(AB)$. □

The following theorem was suggested by the data in the table 2 in the Section 3.

2.1.5 Theorem 3

If A and B are 2×2 matrices and A or B has complex eigenvalues, then $Det(AB - BA) \le 0$.

To prove this theorem, we first need the following lemma.

2.1.6 Lemma 3

If A is a 2×2 matrix and has two complex eigenvalues $a + bi$ and $a - bi$, $b \ne 0$, then there exists a real matrix P such that $P^{-1} AP = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

Proof: Assume $\lambda = a+bi$ and $\overline{\lambda} = a-bi$ are the two eigenvalues of A and u and $\overline{u}$ are the corresponding eigenvectors. Then $Au = \lambda u$ and $A\overline{u} = \overline{\lambda}\overline{u}$. Let $v_1 = \frac{1}{2}\left(u+\overline{u}\right)$ and $v_2 = \frac{1}{2i}\left(u-\overline{u}\right)$, then $u = v_1 + iv_2$ and $\overline{u} = v_1 - iv_2$.

We claim $P = \begin{pmatrix} v_1 & v_2 \end{pmatrix}$ works.

First, P^{-1} exists, since $\lambda \neq \overline{\lambda}$, the corresponding eigenvectors u and $\overline{u}$ are independent, which implies v_1 and v_2 are independent.

Since $P^{-1}v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $P^{-1}v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, then we have

$$
\begin{aligned}
&P^{-1}AP = P^{-1}A\begin{pmatrix} v_1 & v_2 \end{pmatrix} = P^{-1}\begin{pmatrix} Av_1 & Av_2 \end{pmatrix} \\
&= P^{-1}\begin{pmatrix} \frac{1}{2}\left(\lambda u + \overline{\lambda}\overline{u}\right) & \frac{1}{2i}\left(\lambda u - \overline{\lambda}\overline{u}\right) \end{pmatrix} \\
&= P^{-1}\begin{pmatrix} \frac{1}{2}\left[\left(\lambda+\overline{\lambda}\right)v_1 + i\left(\lambda-\overline{\lambda}\right)v_2\right] & \frac{1}{2i}\left[\left(\lambda-\overline{\lambda}\right)v_1 + i\left(\lambda+\overline{\lambda}\right)v_2\right] \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2}\left[\left(\lambda+\overline{\lambda}\right)\begin{pmatrix} 1 \\ 0 \end{pmatrix} + i\left(\lambda-\overline{\lambda}\right)\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right] & \frac{1}{2i}\left[\left(\lambda-\overline{\lambda}\right)\begin{pmatrix} 1 \\ 0 \end{pmatrix} + i\left(\lambda+\overline{\lambda}\right)\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right] \end{pmatrix} \\
&= \begin{pmatrix} a\begin{pmatrix} 1 \\ 0 \end{pmatrix} - b\begin{pmatrix} 0 \\ 1 \end{pmatrix} & b\begin{pmatrix} 1 \\ 0 \end{pmatrix} + a\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}
\end{aligned}
$$

For example, if $A = \begin{pmatrix} 2 & 3 \\ -1 & 2 \end{pmatrix}$, then the eigenvectors of A are $u = \begin{pmatrix} -i\sqrt{3} \\ 1 \end{pmatrix}$ and $\overline{u} = \begin{pmatrix} i\sqrt{3} \\ 1 \end{pmatrix}$. In this case, $P = \begin{pmatrix} 0 & -\sqrt{3} \\ 1 & 0 \end{pmatrix}$ and $P^{-1}AP = \begin{pmatrix} 2 & \sqrt{3} \\ -\sqrt{3} & 2 \end{pmatrix}$.

□

Proof of Theorem 3: Assume A has complex eigenvalues $a+bi$ and $a-bi$, then by Lemma 3, there exists a real matrix P such that $A^{'} = P^{-1}AP = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

Let $B^{'} = P^{-1}BP = \begin{pmatrix} w & x \\ y & z \end{pmatrix}$, then

$$
\begin{aligned}
&Det(AB-BA) = Det\left(P^{-1}APP^{-1}BP - P^{-1}BPP^{-1}AP\right) \\
&= Det\left(A'B' - B'A'\right) = Det\begin{pmatrix} b(x+y) & b(z-w) \\ b(z-w) & -b(x+y) \end{pmatrix} \\
&= -b^2\left[(x+y)^2 + (z-w)^2\right] \le 0
\end{aligned}
$$

□

Next, we relate the relative sizes of matrix traces to a determinant.

2.1.7 Theorem 4

The probability that $Tr(ABAB) > Tr(AABB)$ is the same as the probability that $Det(AB-BA) < 0$.

Proof: Since $Tr(AB) = Tr(BA)$, $Tr(AB-BA) = 0$. Using this and Lemma 1, we have

$$
\begin{aligned}
&(AB-BA)^2 = Tr(AB-BA)\cdot(AB-BA) - Det(AB-BA)\cdot I \\
&= -Det(AB-BA)I
\end{aligned}
$$

.

Taking the trace of each side, $Tr\left[(AB-BA)^2\right] = -2Det(AB-BA)$.

Also, since $(AB-BA)^2 = ABAB - ABBA - BAAB + BABA$,

we have $Tr\left[(AB-BA)^2\right] = 2Tr(ABAB) - 2Tr(AABB)$. Here we have used the fact that $Tr(ABAB) = Tr(BABA)$ and $Tr(ABBA) = Tr(BAAB)$.

Therefore, $2Tr(ABAB) - 2Tr(AABB) = -2Det(AB-BA)$ or

$Tr(ABAB) - Tr(AABB) = -Det(AB-BA)$. □

This theorem provides us an alternative way to calculate the probability that $Tr(ABAB) > Tr(AABB)$, which plays an important role in the following chapter when we derive the formula for the probability that $Tr(ABAB) > Tr(AABB)$ with B is restricted to have the form $\begin{pmatrix} 1 & x \\ 0 & 0 \end{pmatrix}$.

2.1.8 Theorem 5

If M is a product of 2×2 matrices A and B and $M \neq M^R$, then the probability that $Tr(M^2) > Tr(MM^R)$ is the same as the probability that $Tr(ABAB) > Tr(A^2B^2)$.

Proof: First, we show that $Tr(M^2) = Tr\left((M^R)^2\right)$.

Since $Tr(M - M^R) = Tr(M) - Tr(M^R) = 0$, by Lemma 1,

$Tr(M^2) = Tr(Tr(M)M - Det(M)I) = (Tr(M))^2 - 2Det(M)$,

$Tr\left((M^R)^2\right) = Tr\left(Tr(M^R)M^R - Det(M^R)I\right) = \left(Tr(M^R)\right)^2 - 2Det(M^R)$.

By Theorem 1, $(Tr(M))^2 = \left(Tr(M^R)\right)^2$, and $Det(M) = Det(M^R)$ by the definition of determinant. Thus $Tr(M^2) = Tr\left((M^R)^2\right)$.

Since $(M - M^R)^2 = M^2 + (M^R)^2 - MM^R - M^RM$, and combined with the fact that $Tr(MM^R) = Tr(M^RM)$, we have $Tr\left((M - M^R)^2\right) = 2Tr(M^2) - 2Tr(MM^R)$.

By Lemma 1,

$$Tr\left((M - M^R)^2\right) = Tr\left(Tr(M - M^R)(M - M^R) - Det(M - M^R)I\right)$$
$$= \left(Tr(M - M^R)\right)^2 - 2Det(M - M^R) = -2Det(M - M^R) \qquad ,$$

where we have used the fact that $Tr(M - M^R) = Tr(M) - Tr(M^R) = 0$.

Thus $2Tr(M^2) - 2Tr(MM^R) = -2Det(M - M^R)$ or the probability that $Tr(M^2) > Tr(MM^R)$ is the same as the probability that $Det(M - M^R) < 0$.

By Theorem 2, $Det\left(M-M^{R}\right)=c^{2}Det\left(AB-BA\right)$ for some scalar c. And by Theorem 4, the probability that $Tr\left(ABAB\right)>Tr\left(AABB\right)$ is the same as the probability that $Det\left(AB-BA\right)<0$.

Therefore, the probability that $Tr\left(M^{2}\right)>Tr\left(MM^{R}\right)$ is the same as the probability that $Tr\left(ABAB\right)>Tr\left(A^{2}B^{2}\right)$. □

2.1.9 Theorem 6

If the entries in a 2×2 matrix A are uniformly random on $\left[-1,1\right]$, then A has real eigenvalues with a probability of $\frac{49}{72}$.

Proof: Assume $A=\begin{pmatrix} w & x \\ y & z \end{pmatrix}$, then

$Det\left(\lambda I-A\right)=Det\begin{pmatrix} \lambda-w & -x \\ -y & \lambda-z \end{pmatrix}=\lambda^{2}-\left(w+z\right)\lambda+wz-xy$. The discriminant of this quadratic is $\left(w+z\right)^{2}-4\left(wz-xy\right)=\left(w-z\right)^{2}+4xy$. Since $w,x,y,z\in U\left[-1,1\right]$, by symmetry, $\left(w-z\right)^{2}+4xy>0$ is equivalent to $\left(w+z\right)^{2}>4xy$.

Since there are four unknowns and thus it is a four dimensional problem, the support set is a subset of a four dimensional unit cube. We can divide the hypercube of side 2 to $2^{4}=16$ unit cubes. In these 16 small cubes, 8 of them are entirely contained in the solution set of the inequality, since when the right hand side is negative, the inequality is automatically true. That is, when $xy<0$, either $x<0$ and $y>0$ or $x>0$ and $y<0$, and simultaneously w and z can either be positive or negative, so in total there are $2\times2\times2=8$ of them.

When $x>0$ and $y>0$, the contribution of the volumes is given by $\int_{-1}^{1}\int_{-1}^{1}\int_{0}^{1}\int_{0}^{\min\left(1,\frac{\left(w+z\right)^{2}}{4y}\right)}dxdydwdz$. The reason we use $\min\left(1,\frac{\left(w+z\right)^{2}}{4y}\right)$ for the upper limit of x is because $\frac{\left(w+z\right)^{2}}{4y}$ might be greater than 1 and that volume is outside of our hypercube. When $\min\left(1,\frac{(w+z)^2}{4y}\right)=1$, or $0<y<\frac{(w+z)^2}{4}$,

$\int_{-1}^{1}\int_{-1}^{1}\int_{0}^{1}\int_{0}^{\min\left(1,\frac{(w+z)^2}{4y}\right)}dxdydwdz$ reduces to $\int_{-1}^{1}\int_{-1}^{1}\int_{0}^{\frac{(w+z)^2}{4}}1dydwdz$. When

$\min\left(1,\frac{(w+z)^2}{4y}\right)=\frac{(w+z)^2}{4y}$, or $\frac{(w+z)^2}{4}<y<1$, $\int_{-1}^{1}\int_{-1}^{1}\int_{0}^{1}\int_{0}^{\min\left(1,\frac{(w+z)^2}{4y}\right)}dxdydwdz$ reduces

to $\int_{-1}^{1}\int_{-1}^{1}\int_{\frac{(w+z)^2}{4}}^{1}\frac{(w+z)^2}{4y}dydwdz$.

By symmetry, when $x<0$ and $y<0$, the volume would be the same.

In total, the effective volume is

$$8+2\int_{-1}^{1}\int_{-1}^{1}\int_{0}^{1}\int_{0}^{\min\left(1,\frac{(w+z)^2}{4y}\right)}dxdydwdz=8+2\left(\int_{-1}^{1}\int_{-1}^{1}\int_{\frac{(w+z)^2}{4}}^{1}\frac{(w+z)^2}{4y}dydwdz+\int_{-1}^{1}\int_{-1}^{1}\int_{0}^{\frac{(w+z)^2}{4}}1dydwdz\right)$$

$$=8+2\left(\int_{-1}^{1}\int_{-1}^{1}-\frac{(w+z)^2}{4}\left(\ln\frac{(w+z)^2}{4}\right)dwdz+\int_{-1}^{1}\int_{-1}^{1}\frac{(w+z)^2}{4}dwdz\right)=8+\frac{20}{9}+\frac{2}{3}=\frac{98}{9}$$

And thus the probability that the discriminant is positive is $\frac{98}{9}\div 16=\frac{49}{72}$. □

2.1.10 Theorem 7

If the entries in 2×2 matrices A and B are uniformly random on $[-1,1]$, then $Tr(ABAB)>Tr(AABB)$ with a probability of $P>0.5$.

Proof: By Theorem 3 and Theorem 6, the probability that $Det(AB-BA)\le 0$ is at least $1-\left(\frac{49}{72}\right)^2\approx 53.68\%$, by Theorem 4, this is the same as the probability that $Tr(ABAB)>Tr(AABB)$, therefore $Tr(ABAB)>Tr(AABB)$ with a probability of $P>0.5$. □

This theorem gives us a lower bound of the probability that $Det(AB-BA)\le 0$ or $Tr(ABAB)>Tr(AABB)$.

3. Data and analysis of data

In this chapter we investigate traces for products of matrix strings with a fixed number of $A's$.

Having obtained a formula for the number of trace equivalent classes, we now are interested in their relationship. For example, for a matrix string of length 6 with 2 $A's$, there are three possible trace equivalent classes, denoted by α, β, γ, and to all of these distinct traces we obtained, one might think if we vary the entries of A and B, the rank of those traces would achieve all orderings : $\alpha<\beta<\gamma$, $\beta<\gamma<\alpha$, $\gamma<\alpha<\beta$, $\alpha<\gamma<\beta$, $\beta<\alpha<\gamma$, $\gamma<\beta<\alpha$ with equal likelihood. However, in the data obtained from our simulations, there is one ordering that never appeared, and among the orderings that occurred, not all of them have the same frequency.

Analyzing data, we noticed several patterns for the ordering of trace equivalent classes and their frequency.

Let $2AnB$ denote strings of matrices with 2 $A's$ and n $B's$ in any order. if n is odd, then each ordering appears to have the same frequency as its reversal ordering, i.e., the frequency of $\alpha<\beta<\gamma$ is the same as $\gamma<\beta<\alpha$. If n is even, then this symmetry is broken.

In the table below, we list our simulation results for $2A2B$ to $2A10B$. We used numbers to represent those different orderings. For example, in the $2A4B$ case, we denote the trace equivalent class of $AABBBB$ with 1, $ABABBB$ with 2, and $ABBABB$ with 3. The ordering 132 means the ordering of $Tr(ABABBB)<Tr(ABBABB)<Tr(AABBBB)$, and the frequency of ordering 132 is 2193 means the ordering $Tr(ABABBB)<Tr(ABBABB)<Tr(AABBBB)$ happens 2193 times out of $10,000$ trials.

Table 1a

2A2B		2A3B		2A4B	
Ordering	Frequency	Ordering	Frequency	Ordering	Frequency
21	0.2761	21	0.4987	132	0.2135
12	0.7239	12	0.5013	231	0.2845
				312	0.1089
				321	0.3329
				123	0.0601
1=AABB		1=AABBB		1=AABBBB	
2=ABAB		2=ABABB		2=ABABBB	
				3=ABBABB	
2A5B		2A6B		2A7B	
Ordering	Frequency	Ordering	Frequency	Ordering	Frequency
123	0.1588	1234	0.0618	1234	0.1206
321	0.1569	1342	0.2153	4312	0.0133
213	0.0394	4123	0.0493	1342	0.2762
312	0.0382	2431	0.2806	2431	0.2871
132	0.2992	4321	0.2088	4132	0.0237
231	0.3073	4213	0.0646	3142	0.0233
		4312	0.0432	2413	0.0243
		4132	0.074	3241	0.0368
				4321	0.1201
				2314	0.0234
				2134	0.0121
				1423	0.0377
1=AABBBBB		1=AABBBBBB		1=AABBBBBBB	
2=ABABBBB		2=ABABBBBB		2=ABABBBBBB	
3=ABBABBB		3=ABBABBBB		3=ABBABBBBB	
		4=ABBBABBB		4=ABBBABBBB	

The ordering 213cannot occur in the $2A4B$ case. This will be proven in the appendix.

Table 1b

2A8B		2A9B		2A10B	
Ordering	Frequency	Ordering	Frequency	Ordering	Frequency
52143	0.0399	24531	0.2682	123456	0.058
24531	0.2799	32415	0.0126	651423	0.0171
51423	0.0426	54321	0.1039	654321	0.1409
13542	0.2119	41352	0.021	641235	0.0099
54321	0.1611	13542	0.2679	135642	0.1916
54132	0.0259	54132	0.0073	246531	0.2586
12345	0.0663	35214	0.0153	631452	0.0298
53124	0.049	12345	0.1104	612534	0.0196
53214	0.0178	51423	0.0121	632541	0.0369
52341	0.0436	25134	0.0176	614325	0.0207
51243	0.031	41253	0.0147	642135	0.0249
54312	0.0167	34251	0.0237	654132	0.0103
		15243	0.0233	621534	0.0237
		23145	0.0079	613452	0.0177
		25314	0.0019	615243	0.0229
		21345	0.0048	623514	0.0305
		42153	0.0089	654312	0.0051
		35124	0.0076		
		31542	0.008		
		24513	0.0094		
		43152	0.0152		
		54312	0.0044		
1=AABBBBBBBB		1=AABBBBBBBBB		1=AABBBBBBBBBB	
2=ABABBBBBBB		2=ABABBBBBBBB		2=ABABBBBBBBBB	
3=ABBABBBBBB		3=ABBABBBBBBB		3=ABBABBBBBBBB	
4=ABBBABBBBB		4=ABBBABBBBBB		4=ABBBABBBBBBB	
5=ABBBBABBBB		5=ABBBBABBBBB		5=ABBBBABBBBBB	
				6=ABBBBBABBBBB	

Note that, although $2A4B$ and $2A5B$ both have 3 trace equivalent classes, the number of possible orderings given by $2A4B$ is less than the number of $2A5B$. Where $2A4B$ has 5 possible orderings, $2A5B$ has 6 possible orderings. This is also true for $2A6B$ vs $2A7B$, and for $2A8B$ vs $2A9B$.

Based on our simulations, we conjecture that for a matrix string whose length is even, if it contains an even number of $B's$, then each ordering does not necessary have the same frequency as the reversal ordering, and more likely, the frequencies are different. By the result of simulations, this conjecture is true for the matrix strings of length up to 15. The data are attached in the appendix of this paper.

Another pattern in the data in the table is that the trace equivalent classes of form $AA...BB...$ appear to be most likely the trace equivalent classes with the smallest trace. For example, with 2 $A's$ and 6 $B's$, A^2B^6 has the smallest trace in 4829 of $10,000$ trials. This also appears to be true when we change to the distribution of the entries of the $A's$ and $B's$ to a normal distribution with mean equal to 0.

If we have a matrix string of length 4, and with 2 $A's$ and 2 $B's$, then there are only two trace equivalent classes, $AABB$ and $ABAB$, and a direct simulation shows if the entries of A and B are uniformly distributed between -1 and 1, $ABAB$ has a chance of 72.1% to be the larger trace equivalent class. This probability is strongly related to the eigenvalues of A and B.

The table below shows the simulation result that $Tr(ABAB) > Tr(AABB)$ in 720660 of $1,000,000$ trials.

Table 2

Eigenvalues of A	Eigenvalues of B	$Tr(ABAB) > Tr(AABB)$	Frequency
Real	Real	No	279340
Real	Real	Yes	183701
Real	Complex	No	0
Real	Complex	Yes	217715
Complex	Real	No	0
Complex	Real	Yes	217542
Complex	Complex	No	0
Complex	Complex	Yes	101702

By Theorem 4 and Theorem 5 in Chapter 2, we can see the simulation result agrees with the theoretical prediction. Namely, If A and B are 2×2 matrices and A or B have complex eigenvalues, then $Det(AB-BA)<0$, and the probability that $Tr(ABAB) > Tr(AABB)$ is the same as the probability that $Det(AB-BA)<0$.

By Theorem 7, we know that the probability that $Tr(ABAB) > Tr(AABB)$ should be no less than $1-\left(\frac{49}{72}\right)^2 = 53.68\%$. But this is still different from the result we obtained from the simulation. We do not have a theorem corresponding to Theorem 3 when the eigenvalues are real and A and B are general random matrices but we have worked out general special cases. Let $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $B=\begin{pmatrix} x & y \\ z & w \end{pmatrix}$. By Theorem 4, the probability that $Tr(ABAB) > Tr(AABB)$ is equivalent to find the probability that $Det(AB-BA)<0$ with the given A and B. However, this gives rise to an inequality with 8 random variables and an exact solution would be hard to obtain. In this paper, we solve the problem completely when z and w equal to 0, namely $B=\begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix}$. For the efficiency of calculation, we let $B=\begin{pmatrix} 1 & x \\ 0 & 0 \end{pmatrix}$

where x is an arbitrary real number. This is equivalent to $B=\begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix}$ for x and y are uniformly random variables on $[-1,1]$.

If $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and B has the form of $\begin{pmatrix} 1 & x \\ 0 & 0 \end{pmatrix}$, then $Det(AB-BA)<0$ is equivalent to $c^2x^2 \geq c\left[b+x(d-a)\right]$ (*).Below we give the formula of the probability that $c^2x^2 < c\left[b+x(d-a)\right]$ and (*) is the complementary event, so the probability is just 1 minus the probability we obtained.

Here we will only list the results with the details of the proof given in the appendix. For various range of x there are probability function describing when $Det(AB-BA)>0$.

Table 3

Range for x	$P(x)$
$x \leq -\left(1+\sqrt{2}\right)$ or $x \geq 1+\sqrt{2}$	$P(x)=\frac{32x^3+8x-1}{96x^4}$
$-\left(1+\sqrt{2}\right) \leq x \leq -\sqrt{2}$ or $\sqrt{2} \leq x \leq 1+\sqrt{2}$	$P(x)=-\frac{x^8-8x^7+20x^6-8x^5-26x^4-56x^3+20x^2-8x+3}{192x^4}$
$-\sqrt{2} < x \leq -\frac{1}{2}$ or $\frac{1}{2} < x \leq \sqrt{2}$	$P(x)=-\frac{x^8-8x^7+20x^6-8x^5-26x^4-56x^3+20x^2-24x+8}{192x^4}$
$-\frac{1}{2} < x \leq 1-\sqrt{2}$ or $\sqrt{2}-1 < x \leq \frac{1}{2}$	$P(x)=\frac{x^8+8x^7-28x^6+8x^5+70x^4-8x^3+20x^2-8x+1}{192x^4}$
$1-\sqrt{2} \leq x \leq 0$ or $0 < x \leq \sqrt{2}-1$	$P(x)=\frac{1}{2}-\frac{x^2}{4}$

Table 4 is a comparison of the probability that $Det(AB-BA)>0$ predicted by the formula and simulation results. For each value of x, we performed $100,000$ trials.

Table 4

Different x	Probability given by formula	Probability given by simulation
$x=3$	$P(3)\approx 0.114$	$\frac{11,444}{100,000}$
$x=10$	$P(10)\approx 0.033$	$\frac{3304}{100,000}$
$x=1.5$	$P(1.5)\approx 0.235$	$\frac{23,763}{100,000}$
$x=2$	$P(10)\approx 0.176$	$\frac{17,621}{100,000}$
$x=1$	$P(1)\approx 0.323$	$\frac{32,346}{100,000}$
$x=0.8$	$P(0.8)\approx 0.366$	$\frac{36,678}{100,000}$
$x=0.43$	$P(0.46)\approx 0.447$	$\frac{45,224}{100,000}$
$x=0.46$	$P(0.46)\approx 0.447$	$\frac{44,757}{100,000}$
$x=0.2$	$P(0.2)\approx 0.49$	$\frac{46,953}{100,000}$
$x=0.3$	$P(0.3)\approx 0.478$	$\frac{47,796}{100,000}$

Thus, from the table above, we can see the there is a good agreement between the prediction from formula and the numerical results.

4. Conclusion:

By using Burnside's theorem we derived a formula for the number of trace equivalent classes for a matrix string of 2×2 matrices which is comprised of two different matrices A and B with k $A's$ and $n-k$ $B's$. We carried out the simulations for traces of matrix product with 2 $A's$ and n $B's$ for n varying from 2 to 10. Data show when n is odd, then each ordering of trace equivalent classes by trace appears to have the same frequency as its reversal ordering. If n is even, then this symmetry is broken.

The relationship between the probability that $Tr(ABAB) > Tr(AABB)$ and the probability that $Det(AB - BA) < 0$ was analyzed theoretically and was verified numerically. Data from simulations show there is a probability about 72% that $Tr(ABAB) > Tr(AABB)$ if the entries of A and B are uniformly distributed on $[-1,1]$. We could not derive the 72% but we could prove the probability was greater than 50% by a combination of Theorem 3, 4 and 5. Namely, we first examined the influence of eigenvalues of A and B on this inequality. We found out if A or B has complex eigenvalues, then $Tr(ABAB) > Tr(AABB)$. Furthermore, we proved if the entries in 2×2 matrix A are uniformly random on $[-1,1]$, then A has real eigenvalues with a probability of $\frac{49}{72}$, which immediately implies $Tr(ABAB) > Tr(AABB)$ at least with a probability of 53.68%.

We have also given the formula for the probability that $Tr(ABAB) > Tr(AABB)$ when B is restricted to the form $\begin{pmatrix} 1 & x \\ 0 & 0 \end{pmatrix}$ for different x and verified the formula by simulation and showed there is a good agreement between the prediction from the formula and the numerical result.

Missing from our analysis is the gap between 72% and 53.68%, which should be contributed by cases when A and B have real eigenvalues.

It would be interesting to see these ideas applied to Statistical Physics.

5. Appendix

Here we present the details of the deduction of the formula in Section 3 and the proof for the contradiction of missing ordering 213.

5.1 We first deduce the general formula for the PDF of the linear combination of two random variables obeying uniform distribution.

Assume $X, Y \in U[-1,1]$, where X and Y are independent. Then let $U = X$, $V = m_1 X + m_2 Y$, where the m_1 and m_2 are constant. Then the marginal distribution [2] of V can be obtained by integrating the joint PDF of U and V over the support set of V.

Since $X = U$, $Y = \dfrac{V - m_1 U}{m_2}$, the Jacobian $J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -\dfrac{m_1}{m_2} & \dfrac{1}{m_2} \end{vmatrix} = \dfrac{1}{m_2}$, and the joint PDF of U and V is $f_{uv}(u,v) = \dfrac{1}{2} \cdot \dfrac{1}{2} \cdot \left| \dfrac{1}{m_2} \right| = \left| \dfrac{1}{4m_2} \right|$.

When $0 < m_2 < m_1$, then

$$f_v(V) = \int_\Omega f_{uv}(u,v)\,du = \begin{cases} \displaystyle\int_{-1}^{\frac{m_2+v}{m_1}} \frac{1}{4m_2} du, & -m_1 - m_2 \le v < m_2 - m_1 \\ \displaystyle\int_{\frac{-m_2+v}{m_1}}^{\frac{m_2+v}{m_1}} \frac{1}{4m_2} du, & m_2 - m_1 \le v < m_1 - m_2 \\ \displaystyle\int_{\frac{-m_2+v}{m_1}}^{1} \frac{1}{4m_2} du, & m_1 - m_2 \le v \le m_2 + m_1 \end{cases}$$

$$= \begin{cases} \dfrac{1}{4m_2}\left(\dfrac{m_2+v}{m_1} + 1\right), & -m_1 - m_2 \le v < m_2 - m_1 \\ \dfrac{1}{2m_1}, & m_2 - m_1 \le v < m_1 - m_2 \\ \dfrac{1}{4m_2}\left(1 - \dfrac{v - m_2}{m_1}\right), & m_1 - m_2 \le v \le m_2 + m_1 \end{cases}.$$

When $0 < m_1 < m_2$, then $f_v(V) = \begin{cases} \dfrac{1}{4m_2}\left(\dfrac{m_2+v}{m_1}+1\right), & -m_1-m_2 \le v < m_1-m_2 \\ \dfrac{1}{2m_2}, & m_1-m_2 \le v < m_2-m_1 \\ \dfrac{1}{4m_2}\left(1-\dfrac{v-m_2}{m_1}\right), & m_2-m_1 \le v \le m_2+m_1 \end{cases}$.

When $m_2 > 0, m_1 < 0, |m_2| < |m_1|$, then

$$f_v(V) = \begin{cases} \dfrac{1}{4m_2}\left(1-\dfrac{v+m_2}{m_1}\right), & m_1-m_2 \le v < m_1+m_2 \\ -\dfrac{1}{2m_1}, & m_1+m_2 \le v < -m_2-m_1 \\ \dfrac{1}{4m_2}\left(1+\dfrac{v-m_2}{m_1}\right), & -m_2-m_1 \le v \le m_2-m_1 \end{cases}.$$

When $m_2 > 0, m_1 < 0, |m_1| < |m_2|$, then

$$f_v(V) = \begin{cases} \dfrac{1}{4m_2}\left(1-\dfrac{v+m_2}{m_1}\right), & m_1-m_2 \le v < -m_1-m_2 \\ \dfrac{1}{2m_2}, & -m_1-m_2 \le v < m_2+m_1 \\ \dfrac{1}{4m_2}\left(1+\dfrac{v-m_2}{m_1}\right), & m_2+m_1 \le v \le m_2-m_1 \end{cases}.$$

When $m_2 < 0, m_1 > 0, |m_1| > |m_2|$, then

$$f_v(V)=\begin{cases}-\dfrac{1}{4m_2}\left(1+\dfrac{v-m_2}{m_1}\right), & -m_1+m_2\le v<-m_1-m_2\\ \dfrac{1}{2m_1}, & -m_1-m_2\le v<m_2+m_1\\ -\dfrac{1}{4m_2}\left(1-\dfrac{v+m_2}{m_1}\right), & m_2+m_1\le v\le m_1-m_2\end{cases}.$$

When $m_2<0, m_1>0, |m_2|>|m_1|$, then

$$f_v(V)=\begin{cases}-\dfrac{1}{4m_2}\left(1+\dfrac{v-m_2}{m_1}\right), & -m_1+m_2\le v<m_1+m_2\\ -\dfrac{1}{2m_2}, & m_1+m_2\le v<-m_2-m_1\\ -\dfrac{1}{4m_2}\left(1-\dfrac{v+m_2}{m_1}\right), & -m_2-m_1\le v\le m_1-m_2\end{cases}.$$

When $m_1<m_2<0$, then $f_v(V)=\begin{cases}-\dfrac{1}{4m_2}\left(1-\dfrac{v-m_2}{m_1}\right), & m_1+m_2\le v<m_1-m_2\\ -\dfrac{1}{2m_1}, & m_1-m_2\le v<m_2-m_1\\ -\dfrac{1}{4m_2}\left(1+\dfrac{v+m_2}{m_1}\right), & m_2-m_1\le v\le m_1-m_2\end{cases}.$

When $m_2<m_1<0$, then $f_v(V)=\begin{cases}-\dfrac{1}{4m_2}\left(1-\dfrac{v-m_2}{m_1}\right), & m_1+m_2\le v<-m_1+m_2\\ -\dfrac{1}{2m_2}, & -m_1+m_2\le v<m_1-m_2\\ -\dfrac{1}{4m_2}\left(1+\dfrac{v+m_2}{m_1}\right), & m_1-m_2\le v\le -m_1-m_2\end{cases}.$

Now we carry out the calculation for the probability that $Tr(ABAB)>Tr(AABB)$ with $B=\begin{pmatrix}1 & x\\ 0 & 0\end{pmatrix}$.

If $c>0$, (*) in Section 3 reduces to $cx^2 \le b + x(d-a)$ or $cx^2 - b \le x(d-a)$.

If we look back to (*), we will see the probability that $f(x) = c^2x^2 - c\left[b + x(d-a)\right] \le 0$ for a fixed x is the same as when we replace x with $-x$, since simultaneously we can change $(d-a)$ to $(a-d)$. That is $f(-x) = c^2x^2 - c\left[b + x(a-d)\right]$, and when a and d are uniformly distributed between -1 and 1, $(d-a)$ has the same distribution as $(a-d)$. Thus $P\left(f(x) \le 0\right) = P\left(f(-x) \le 0\right)$. Therefore, without lost of generality, we assume $x>0$. With this assumption, we can further reduce (*) to $xc - \frac{1}{x}b \le d-a$ or $m_1c + m_2b \le d-a$ where $m_1 = x, m_2 = -\frac{1}{x}$.

Since $c \in U[0,1]$, $e = 2\left(c - \frac{1}{2}\right) \in U[-1,1]$ with $c = \frac{e}{2} + \frac{1}{2}$. Let $m_1' = \frac{m_1}{2}, m_2' = \frac{m_2}{2}$.

Then (*) becomes $m_1'e + m_2'b + m_1' \le d-a$ (#) where $m_1' = \frac{m_1}{2} = \frac{x}{2}, m_2' = \frac{m_2}{2} = -\frac{1}{x}$.

By our assumption, $m_1' > 0, m_2' < 0$, and there are two cases.

First suppose $|m_1'| > |m_2'|$. This happens when $x > \sqrt{2}$. Let $v = m_1'e + m_2'b + m_1'$, $u = d-a$. The marginal Probability Density Functions(PDF) are

$$f(v) = \begin{cases} \frac{1}{2}\left(v + \frac{1}{x}\right), & -\frac{1}{x} \le v < \frac{1}{x} \\ \frac{1}{x}, & \frac{1}{x} \le v < x - \frac{1}{x} \\ \frac{1}{2}\left(v + \frac{1}{x}\right), & x - \frac{1}{x} \le v \le x + \frac{1}{x} \end{cases},$$

$$f(u) = \begin{cases} \frac{1}{4}(u+2), & -2 \le u < 0 \\ \frac{1}{4}(-u+2), & 0 \le u \le 2 \end{cases},$$

The probability of $v \le u$ is $\int_\Omega f(u,v)\,dudv$ where Ω is the region with $v \le u$.

When $x \ge 1+\sqrt{2}$, $\Omega = A_1 + A_2 + A_3 + A_4$, and

$$P_{A_1}(x) = \int_{-\frac{1}{x}}^{0}\int_{v}^{0} \frac{1}{4}(u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dudv,$$

$$P_{A_2}(x) = \int_{-\frac{1}{x}}^{0}\int_{0}^{2} \frac{1}{4}(-u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dudv,$$

$$P_{A_3}(x) = \int_{0}^{\frac{1}{x}}\int_{v}^{2} \frac{1}{4}(-u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dudv,$$

$$P_{A_4}(x) = \int_{\frac{1}{x}}^{2}\int_{v}^{2} \frac{1}{4}(-u+2)\times\frac{1}{x}dudv.$$

Therefore, $P(x) = P_{A_1}(x) + P_{A_2}(x) + P_{A_3}(x) + P_{A_4}(x) = \dfrac{32x^3+8x-1}{96x^4}$.

When $\sqrt{2} \le x \le 1+\sqrt{2}$, $\Omega = A_1 + A_2 + A_3 + A_4 + A_5$, and

$$P_{A_1}(x) = \int_{-\frac{1}{x}}^{0}\int_{v}^{0} \frac{1}{4}(u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dudv,$$

$$P_{A_2}(x) = \int_{-\frac{1}{x}}^{0}\int_{0}^{2} \frac{1}{4}(-u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dudv,$$

$$P_{A_3}(x) = \int_{0}^{\frac{1}{x}}\int_{v}^{2} \frac{1}{4}(-u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dudv,$$

$$P_{A_4}(x) = \int_{x}^{x-\frac{1}{x}}\int_{v}^{2} \frac{1}{4}(-u+2)\times\frac{1}{x}dudv,$$

$$P_{A_5}(x) = \int_{x-\frac{1}{x}}^{2}\int_{v}^{2} \frac{1}{4}(-u+2)\times\frac{1}{2}\left(x+\frac{1}{x}-v\right)dudv.$$

Therefore,

$$P(x)=P_{A_1}(x)+P_{A_2}(x)+P_{A_3}(x)+P_{A_4}(x)+P_{A_5}(x)$$
$$=-\frac{x^8-8x^7+20x^6-8x^5-26x^4-56x^3+20x^2-8x+3}{192x^4}.$$

In the second case $\left|m_1^{'}\right|\le\left|m_2^{'}\right|$. This happens when $0<x\le\sqrt{2}$. Then

$$f(v)=\begin{cases}\frac{1}{2}\left(v+\frac{1}{x}\right), & -\frac{1}{x}\le v<x-\frac{1}{x}\\ \frac{x}{2}, & x-\frac{1}{x}\le v<\frac{1}{x}\\ \frac{1}{2}\left(x+\frac{1}{x}-v\right), & \frac{1}{x}\le v\le x+\frac{1}{x}\end{cases},$$

$$f(u)=\begin{cases}\frac{1}{4}(u+2), & -2\le u<0\\ \frac{1}{4}(-u+2), & 0\le u\le 2\end{cases},$$

When $0<x\le\sqrt{2}-1$, $\Omega=A_1+A_2+A_3+A_4$, and

$$P_{A_1}(x)=\int_{-\frac{1}{x}}^{x-\frac{1}{x}}\int_{-2}^{0}\frac{1}{4}(u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dudv,$$

$$P_{A_2}(x)=\int_{-\frac{1}{x}}^{x-\frac{1}{x}}\int_{0}^{2}\frac{1}{4}(-u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dudv,$$

$$P_{A_3}(x)=\int_{-2}^{0}\int_{x-\frac{1}{x}}^{u}\frac{1}{4}(u+2)\times\frac{x}{2}dvdu,$$

$$P_{A_4}(x)=\int_{0}^{2}\int_{x-\frac{1}{x}}^{u}\frac{1}{4}(-u+2)\times\frac{x}{2}dvdu.$$

Therefore, $P(x)=P_{A_1}(x)+P_{A_2}(x)+P_{A_3}(x)+P_{A_4}(x)=\frac{1}{2}-\frac{x^2}{4}$.

When $\sqrt{2}-1<x\le\frac{1}{2}$, $\Omega=A_1+A_2+A_3+A_4+A_5$ and

$$P_{A_1}(x)=\int_{-2}^{0}\int_{-\frac{1}{x}}^{-2}\frac{1}{4}(u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dvdu,$$

$$P_{A_2}(x)=\int_{-2}^{x-\frac{1}{x}}\int_{v}^{0}\frac{1}{4}(u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dudv,$$

$$P_{A_3}(x)=\int_{0}^{2}\int_{-\frac{1}{x}}^{x-\frac{1}{x}}\frac{1}{4}(-u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dvdu,$$

$$P_{A_4}(x)=\int_{x-\frac{1}{x}}^{0}\int_{v}^{0}\frac{1}{4}(u+2)\times\frac{x}{2}dudv,$$

$$P_{A_5}(x)=\int_{0}^{2}\int_{x-\frac{1}{x}}^{u}\frac{1}{4}(-u+2)\times\frac{x}{2}dvdu.$$

Therefore,

$$P(x)=P_{A_1}(x)+P_{A_2}(x)+P_{A_3}(x)+P_{A_4}(x)+P_{A_5}(x)$$
$$=\frac{x^8+8x^7-28x^6+8x^5+70x^4-8x^3+20x^2-8x+1}{192x^4}.$$

When $\frac{1}{2}<x\le\sqrt{2}$, $\Omega=A_1+A_2+A_3+A_4+A_5$, and

$$P_{A_1}(x)=\int_{-\frac{1}{x}}^{0}\int_{v}^{0}\frac{1}{4}(u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dudv,$$

$$P_{A_2}(x)=\int_0^2\int_{-\frac{1}{x}}^{0}\frac{1}{4}(-u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dvdu,$$

$$P_{A_3}(x)=\int_0^{x-\frac{1}{x}}\int_v^2\frac{1}{4}(-u+2)\times\frac{1}{2}\left(v+\frac{1}{x}\right)dudv,$$

$$P_{A_4}(x)=\int_{x-\frac{1}{x}}^{\frac{1}{x}}\int_v^2\frac{1}{4}(-u+2)\times\frac{x}{2}dudv,$$

$$P_{A_5}(x)=\int_{\frac{1}{x}}^{2}\int_v^2\frac{1}{4}(-u+2)\times\frac{1}{2}\left(x+\frac{1}{x}-v\right)dudv.$$

Therefore,

$$\begin{aligned}P(x)&=P_{A_1}(x)+P_{A_2}(x)+P_{A_3}(x)+P_{A_4}(x)+P_{A_5}(x)\\&=-\frac{x^8-8x^7+20x^6-8x^5-26x^4-56x^3+20x^2-24x+8}{192x^4}\end{aligned}$$

.

5.2 We show here that in the $2A4B$ case, the ordering 213, corresponding to $Tr(ABABBB)>Tr(AABBBB)>Tr(ABBABB)$ cannot occur. We show this by contradiction.

Proof: By Lemma 1, $BB=Tr(B)B-Det(B)I$. Thus, the three traces can be reduced as follows.

$$Tr(ABABBB)=\left((Tr(B))^2-Det(B)\right)Tr(ABAB)-Tr(B)Det(B)Tr(ABA),$$

$$Tr(AABBBB)=\left((Tr(B))^2-Det(B)\right)Tr(AABB)-Tr(B)Det(B)Tr(AAB), \text{ and}$$

$$Tr(ABBABB)=(Tr(B))^2Tr(ABAB)-Det(B)Tr(ABBA)-Tr(B)Det(B)Tr(AAB)$$

respectively.

Therefore, $Tr(ABABBB) > Tr(AABBBB)$ when
$\left((Tr(B))^2 - Det(B)\right)\left(Tr(ABAB) - Tr(AABB)\right) > 0$ and $Tr(AABBBB) > Tr(ABBABB)$ when $Tr(ABAB) - Tr(AABB) < 0$. By Theorem 4, $Tr(ABAB) - Tr(AABB) < 0$ implies $Det(AB - BA) > 0$.

Thus, $Tr(ABABBB) > Tr(AABBBB) > Tr(ABBABB)$ implies $Det(AB - BA) > 0$ and $(Tr(B))^2 - Det(B) < 0$. By Theorem 3, A must have real eigenvalues, u, v. Let P be a diagonalizing matrix such that $A' = P^{-1}AP = \begin{pmatrix} u & 0 \\ 0 & v \end{pmatrix}$ and $B' = P^{-1}BP = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Therefore,

$$\begin{aligned}
&Det(AB - BA) = Det\left(PP^{-1}APP^{-1}BPP^{-1} - PP^{-1}BPP^{-1}APP^{-1}\right) \\
&= Det\left(PA'B'P^{-1} - PB'A'P^{-1}\right) = Det(P)\,Det\left(A'B' - B'A'\right)Det\left(P^{-1}\right) \\
&= Det\left(A'B' - B'A'\right) = bc(u - v)^2 > 0 \text{ or } bc > 0
\end{aligned}.$$

The trace is invariant under similarity transformation, so $Tr(B) = Tr(B')$. Also by $Det(B) = Det\left(PB'P^{-1}\right) = Det(B')$, $(Tr(B))^2 - Det(B) < 0$ implies $(Tr(B'))^2 - Det(B') < 0$ or $(a + d)^2 < ad - bc$. Since $bc > 0$, $(a + d)^2 - ad = a^2 + ad + d^2 < -bc < 0$. However, $a^2 + d^2 = |a|^2 + |d|^2 \geq 2|a||d|$ for all a, d, $a^2 + ad + d^2 < 0$ cannot be true. This contradiction is due to the assumption that $Tr(ABABBB) > Tr(AABBBB) > Tr(ABBABB)$ and thus $Tr(ABABBB) > Tr(AABBBB)$ and $Tr(AABBBB) > Tr(ABBABB)$ cannot be true simultaneously.

Acknowledgement

I wish to thank Professor John Greene, Professor Joseph Gallian, Professor Ron Regal and Professor Barry James for their great help with this thesis.

References

1. Howard Anton, Elementary Linear Algebra, John Wiley &Sons, 7th ed, 1994
2. Lee Bain, Introduction to Probability and Mathematical Statistics, Duxbury Resource Center, 2nd ed, 2007
3. A. Crisanti, G. Paladin, A. Vulpiani, Products of Random Matrices, 1993
4. P.S. Davis, Lyapunov Exponent and Transfer-Matrix Spectrum of the Random Binary Alloy, Physics Review B, Vol 52, Number 6, 1995
5. Barry Dayton, Ghosts and Necklaces, Northeastern Illinois University, 1994
6. Joseph Gallian, Contemporary Abstract Algebra, CENGAGE Learning, 6th ed, 2006
7. A. Jackson, B. Lautrup, Products of Random Matrices, Physics Review E 66, (066124), 2002
8. http://en.wikipedia.org/wiki/Lyapunov_exponent