Some Results on Permutations of Matrix Products

Hans Joachim Werner[1] & Ingram Olkin[2]

[1]Department of Statistics, University of Bonn, Germany hjw.de@uni-bonn.de
[2]Department of Statistics, Stanford University, USA iolkin@stat.stanford.edu

universitätbonn

Abstract

It is well-known that

$$\mathrm{trace}(AB) \geq 0$$

for real-symmetric nonnegative definite matrices A and B. However,

$$\mathrm{trace}(ABC)$$

can be positive, zero or negative, even when C is real-symmetric nonnegative definite. The genesis of the present investigation is consideration of a product of square matrices

$$A = A_1 A_2 \cdots A_n.$$

Permuting the factors of A leads to a different matrix product. We are interested in conditions under which the spectrum remains invariant. The main results are for square matrices over an arbitrary algebraically closed commutative field. The special case of real-symmetric possibly nonnegative definite matrices is also considered.

1. Introduction & Outline

Consider a product

$$p = p_1 p_2 \cdots p_n.$$

Some notations:

- DIRECT NEIGHBORS or NEXT-DOOR NEIGHBORS.
- LENGTH OF A PRODUCT.
- Let p be a product of length n, and let q be obtained from p by a permutation of its n (not necessarily distinct) factors. These two products are said to be DN-RELATED to each other if each of the n factors has in both products exactly the same direct neighbors. In which case, we write $p \sim_{\mathrm{DN}} q$.

In this note, we consider matrix products

$$A = A_1 A_2 \cdots A_n,$$

where the factors A_i $(i = 1, 2, \cdot, n)$ are square matrices over an algebraically closed commutative field $\mathbb{F}$.

Some further notations:

- Let $\mathcal{S}_n$ denote the symmetric group on the natural numbers $1, 2, \cdots, n$.
- For each $\pi := (\pi_1, \pi_2, \cdots, \pi_n) \in \mathcal{S}_n$, let $A_\pi := \prod_{i=1}^{n} A_{\pi_i}$.
- Moreover, let $\mathcal{S}_A := \{A_\pi \mid \pi \in \mathcal{S}_n\}$.

Some observations:

- $\sim_{\mathrm{DN}}$ is an equivalence relation.
- Hence the collection of equivalence classes $\mathrm{DN}(B) := \{A_\pi \,|A_\pi \sim_{\mathrm{DN}} B\}$, $B \in \mathcal{S}_A$, forms a partition of $\mathcal{S}_A$.
- There are exactly $(n-1)!/2$ disjoint equivalence classes. For observe that the equivalence class $\mathrm{DN}(A_\pi)$ of

$$A_\pi = A_{\pi_1} A_{\pi_2} \cdots A_{\pi_n} \tag{1}$$

  contains exactly all those products of the matrix symbols A_i $(i = 1, 2, \cdots, n)$ that are obtained from (1) by cyclical and/or reverse re-orderings.

Outline:

- In Section 3, we study some functions which are defined on $\mathcal{S}_A$, mainly the $\mathrm{trace}(\cdot)$ and the $\mathrm{spectrum}(\cdot)$. There, we particularly show that if A_i $(i = 1, 2, \cdots, n)$ are all symmetric $m \times m$ matrices over $\mathbb{F}$, then these functions are constant on each equivalence class.
- Section 4 deals with the set of all symmetric nonnegative definite $m \times m$ matrices over the field $\mathbb{R}$ of real numbers.
- Section 2 contains some known results, from which our findings in the subsequent two chapters easily follow.

2. Prerequisite Facts

FACT 2.1. [See, e.g., pp. 53–55 in Lancaster (1969).] *Let A be a square matrix of order m over the field $\mathbb{F}$. The characteristic polynomial of A, being defined as $c_A(\lambda) := \det(\lambda I_m - A)$, is a monic polynomial of degree m with exactly m (not necessarily distinct) roots $\lambda_1, \lambda_2, \cdots, \lambda_m \in \mathbb{F}$, called the eigenvalues of A. When writing the characteristic polynomial of A as*

$$c_A(\lambda) = \lambda^m - c_1\lambda^{m-1} + c_2\lambda^{m-2} + \cdots + (-1)^m c_m, \tag{2}$$

the following relationships hold between the coefficients c_r $(r = 1, 2, \cdots, m)$, the eigenvalues of A, the r-th compound $A^{(r)}$ and the principal minors of A:

$$c_r = \operatorname{trace}(A^{(r)}) = \sum (\text{all } r \times r \text{ principal minors}) \quad = \sum_{1 \le i_1 < i_2 < \cdots i_r \le m} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r}. \tag{3}$$

Hence, in particular,

$$c_1 = \operatorname{trace}(A) = \sum_{i=1}^{m} \lambda_i \qquad \text{and} \qquad c_m = \det(A) = \prod_{i=1}^{n} \lambda_i. \tag{4}$$

FACT 2.2. [See, e.g., Exercise 6 on p. 56 in Lancaster (1969).] *Let A be a square matrix over $\mathbb{F}$. Then*

$$c_A(\lambda) = c_{A^t}(\lambda), \tag{5}$$

where A^t denotes the transpose of A. In other words, A and its transpose A^t possess the same characteristic polynomial, and so these matrices have the same set of eigenvalues with corresponding algebraic multiplicities.

FACT 2.3. [See, e.g., Exercise 7.1.19 on p. 503 in Meyer (2000).] *Let* A *and* B *be square matrices of order* m *over the field* $\mathbb{F}$*. Then the matrices* AB *and* BA *have the same set of eigenvalues with corresponding algebraic multiplicities. Hence, in particular,*

$$\text{spectrum}(AB) = \text{spectrum}(BA). \tag{6}$$

For the sake of completeness as well as for easier reference, we also cite some well-known results for Hermitian matrices (over the field $\mathbb{C}$ of complex numbers).

FACT 2.4. [See, e.g., pp. 75–78 in Lancaster (1969).] *Let A be an Hermitian $m \times m$ matrix. Then all eigenvalues of A are real. Moreover, A is unitarily similar to the diagonal matrix $D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_m)$ of its eigenvalues, i.e., there exists an $m \times m$ (unitary) matrix $U = (u_1, u_2, \cdots, u_m)$ such that*

$$UU^* = I_m \quad \text{and} \quad A = UDU^*$$

or, equivalently,

$$\sum_{i=1}^{m} u_i u_i^* = I_m \quad \text{and} \quad A = \sum_{i=1}^{m} \lambda_i u_i u_i^*,$$

with $(\cdot)^*$ indicating as usual the conjugate transpose of $(\cdot)$. The pairs (λ_i, u_i), $i = 1, 2, \cdots, m$, are eigenpairs for A, i.e., λ_i and u_i, satisfying $Au_i = \lambda_i u_i$, are eigenvalues and eigenvectors of A, respectively.

For real-symmetric matrices the previous result allows the following version.

FACT 2.5. [See, e.g., pp. 75–78 in Lancaster (1969).] *Let A be a real-symmetric $m \times m$ matrix. Then all eigenvalues of A are real. Moreover, A is orthogonally similar to the diagonal matrix $D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_m)$ of its eigenvalues, i.e., there exists an $m \times m$ (orthogonal) real matrix $P = (p_1, p_2, \cdots, p_m)$ such that*

$$PP^t = I \quad \text{and} \quad A = PDP^t$$

or, equivalently,

$$\sum_{i=1}^{m} p_i p_i^t = I_m \quad \text{and} \quad A = \sum_{i=1}^{m} \lambda_i p_i p_i^t.$$

The pairs (λ_i, p_i), $i = 1, 2, \cdots, m$, are eigenpairs for A, i.e., λ_i and p_i, satisfying $Ap_i = \lambda_i p_i$, are eigenvalues and eigenvectors of A, respectively.

Below we will also make use of the following two results.

<u>FACT 2.6.</u> [See, e.g., p. 559 in Meyer (2000).] *Let A be a real-symmetric nonnegative definite matrix. Then all its eigenvalues are nonnegative. If all its eigenvalues are positive, then A is even a positive definite matrix.*

<u>FACT 2.7.</u> [See, e.g., Exercise 7.2.16 in Meyer (2000).] *Let A and B be diagonalizable matrices of the same order, say $m \times m$. Then A and B commute, i.e. $AB = BA$, if and only if A and B can be simultaneously diagonalized, i.e. if and only*

$$A = XD_AX^{-1} \quad \text{and} \quad B = XD_BX^{-1}$$

for some regular matrix $X = (x_1, x_2, \cdots, x_m)$ and some diagonal matrices $D_A = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_m)$ and $D_B = (\mu_1, \mu_2, \cdots, \mu_m)$. For $i = 1, 2, \cdots, m$, the pairs (λ_i, x_i) and (μ_i, x_i) are eigenpairs of A and B, respectively.

3. Main Results

<u>THEOREM 3.1.</u> *For symmetric $m \times m$ matrices $A_1,\ A_2, \cdots, A_n$ over the field $\mathbb{F}$, let $A := \prod_{i=1}^n A_i$. Then we have*

$$c_A(\lambda) = c_{A_\pi}(\lambda), \tag{7}$$

irrespective of $A_\pi \in \mathrm{DN}(A)$. Consequently,

$$\begin{aligned} \mathrm{trace}(A^{(r)}) &= \sum (\text{all } r \times r \text{ principal minors of } A) \\ &= \sum (\text{all } r \times r \text{ principal minors of } A_\pi) \\ &= \mathrm{trace}(A_\pi^{(r)}), \end{aligned} \tag{8}$$

irrespective of $A_\pi \in \mathrm{DN}(A)$, and so, in particular,

$$\mathrm{trace}(A) = \mathrm{trace}(A_\pi) \quad \text{and} \quad \det(A) = \det(A_\pi) \tag{9}$$

for all A_π with $A_\pi \sim_{\mathrm{DN}} A$.

PROOF: Recall from Section 1 that $\mathrm{DN}(A)$ consists exactly of all those $2n$ matrix products that are obtainable from $A = A_1 A_2 \cdots A_n$ by cyclical and/or reverse re-orderings of the n matrix factors in A. In virtue of Fact 2.1 and Fact 2.2, the claimed results now follow easily by means of Fact 2.3 and the fact that, for instance, $(A_1 A_2 \cdots A_n)^t = A_n A_{n-1} \cdots A_1$. Details are left to the reader. $\square$

Because for any matrix product A of length 3, $\mathrm{DN}(A) = \mathcal{S}_A$, the following is an immediate corollary of the previous theorem.

<u>Corollary 3.2</u> *Let $A := A_1A_2A_3$, with A_1, A_2 and A_3 being symmetric matrices of the same order $m \times m$ over the field $\mathbb{F}$. Then*

$$c_{A_1A_2A_3}(\lambda) = c_{A_{\pi_1}A_{\pi_2}A_{\pi_3}}(\lambda) \tag{10}$$

for each permutation $\pi = (\pi_1, \pi_2, \pi_3) \in \mathcal{S}_3$. Hence, in particular,

$$\mathrm{trace}(A_1^{(r)}A_2^{(r)}A_3^{(r)}) = \mathrm{trace}(A_{\pi_1}^{(r)}A_{\pi_2}^{(r)}A_{\pi_3}^{(r)}), \tag{11}$$

irrespective of $\pi = (\pi_1, \pi_2, \pi_3) \in \mathcal{S}_3$ and $r \in \mathbb{N}_m$, where

$$\mathbb{N}_m := \{r \in \mathbb{N} : 1 \le r \le m\}.$$

If two of the three (not necessarily symmetric) square matrices A_1, A_2 and A_3 in the matrix product $A := A_1 A_2 A_3$ commute, then each matrix in $\mathcal{S}_A$ can obviously be obtained by a cyclical reordering of the factors of A and/or by the commutation of the commuting factors, and so we obtain the following.

<u>COROLLARY 3.3.</u> *Let $A := A_1 A_2 A_3$, with A_1, A_2 and A_3 being such that at least two of the three $m \times m$ matrices commute. Then*

$$c_{A_1 A_2 A_3}(\lambda) = c_{A_{\pi_1} A_{\pi_2} A_{\pi_3}}(\lambda) \tag{12}$$

and so

$$\mathrm{trace}(A_1^{(r)} A_2^{(r)} A_3^{(r)}) = \mathrm{trace}(A_{\pi_1}^{(r)} A_{\pi_2}^{(r)} A_{\pi_3}^{(r)}), \tag{13}$$

irrespective of $\pi = (\pi_1, \pi_2, \pi_3) \in \mathcal{S}_3$ and $r \in \mathbb{N}_m$.

4. Special Case: Products of Length Three of Real-Symmetric Nonnegative Definite Matrices

Consider the product

$$A_1 A_2 A_3,$$

where all three factors are real-symmetric nonnegative definite. Then, according to Corollary 3.2,

$$c_{A_1 A_2 A_3}(\lambda) = c_{A_2 A_1 A_3}(\lambda)$$

and so

$$\operatorname{trace}((A_1 A_2 + A_2 A_1) A_3) = 2 \operatorname{trace}(A_1 A_2 A_3)$$

holds true. An interesting question is, whether for given such factors anything can be said about the signum of $\operatorname{trace}((A_1 A_2 + A_2 A_1) A_3)$.

Clearly, if A_1, A_2 and A_3 are real-symmetric nonnegative definite $m \times m$ matrices, then according to Fact 2.5, these matrices are orthogonally similar to some diagonal matrices and hence, for $i = 1, 2, 3$, A_i can always be written as

$$A_i = \sum_{j=1}^{m} \lambda_{ij} x_{ij} x_{ij}^t, \tag{14}$$

where (λ_{ij}, x_{ij}) $(j = 1, 2, \cdots, m)$ are eigenpairs of A_i and $\{x_{ij} \mid j = 1, 2, \cdots, m\}$ constitutes an orthonormal basis for $\mathbb{R}^m$. Then

$$\text{trace}(A_1 A_2) = \sum_{j=1}^{m} \sum_{k=1}^{m} \lambda_{1j} \lambda_{2k} (x_{1j}^t x_{2k})^2 \geq 0, \tag{15}$$

since, in view of Fact 2.6, all eigenvalues of a real-symmetric nonnegative definite matrix are nonnegative. One might be tempted to believe that this result can be extended to three factors. That this, however, is erroneous is illustrated by our next example.

EXAMPLE 4.1. Consider the real-symmetric positive definite matrices

$$A_1 := \begin{pmatrix} 4 & -1.9 \\ -1.9 & 1 \end{pmatrix}, \quad A_2 := \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix} \quad \text{and} \quad A_3 := \begin{pmatrix} 1 & -1.4 \\ -1.4 & 2 \end{pmatrix}.$$

Then

$$A_1 A_2 A_3 = \begin{pmatrix} -0.09 & 0.194 \\ 0.006 & -0.02 \end{pmatrix} \quad \text{and} \quad A_2 A_1 A_3 = \begin{pmatrix} 3.69 & -5.206 \\ 2.694 & -3.8 \end{pmatrix},$$

and so $\mathrm{trace}(A_1 A_2 A_3) = -0.11$ and $\mathrm{trace}(A_2 A_1 A_3) = -0.11$. The traces are negative and coincide; the latter is in accordance with our findings in Section 2. The spectrum of both matrices as well as all other matrices from $\mathcal{S}_{A_1 A_2 A_3}$ is given by

$$\mathrm{spectrum}(A_1 A_2 A_3) = \{(\sqrt{61} - 55)/1000, (-\sqrt{61} - 55)/1000\} = \mathrm{spectrum}(A_2 A_1 A_3).$$

We conclude with emphasizing that therefore, without any further restrictive assumptions, nothing can be said about the signum of the trace of the product of three real-symmetric nonnegative definite matrices. The trace can be positive, negative, or even 0. If, however, the $m \times m$ matrices A_1, A_2, and A_3 are all real-symmetric nonnegative definite and, in addition, such that at least two of them commute, then, in view of Fact 2.6 and Fact 2.5, it is clear that the product of the commuting pair of matrices is itself a symmetric nonnegative definite matrix. Since in such a situation the product of $A_1A_2A_3$ can hence be considered as the product of two real-symmetric nonnegative matrices, it follows from the lines around (15) that the trace of $A_1A_2A_3$ is indeed also nonnegative. Needless to say, if A_1, A_2 and A_3 are real-symmetric positive definite and two of these matrices commute, then the trace of $A_1A_2A_3$ is necessarily positive.

A.C. AITKEN (1952). *Determinants and Matrices.* 7th edition. Oliver & Boyd, Edinburgh-London.

Wolfgang GRÖBNER (1965). *Matrizenrechnung.* Bibliographisches Institut, Mannheim.

Peter LANCASTER (1969). *Theory of Matrices.* Academic Press, New York.

Albert W. MARSHALL & Ingram OLKIN (1979). *Inequalities: Theory of Majorization and its Applications.* Academic Press, New York.

Carl D. MEYER (2000). *Matrix Analysis and Applied Linear Algebra.* SIAM, Philadelphia, USA.

Hans Joachim WERNER & Ingram OLKIN (2009). On permutations of matrix products. **In** *Statistical Inference, Econometric Analysis and Matrix Algebra* (B. Schipp & W. Krämer, eds.), pp. 359–365, Physica Verlag, Heidelberg.