CUBO 8 , 91-97 (1992)
Séptima Jornada de Matemática de la Zona Sur.

A trace inequality with a subtracted term

H. Miranda *and Robert C. Thompson †

Abstract.
For fixed real or complex matrices A and B, the well known von Neumann trace inequality identifies the maximum of $| tr(UAVB) |$, as U and V range over the unitary group, the maximum being a bilinear expression in the singular values of A y B. This paper establishes the analogue of this inequality for real matrices A and B when U and V range over the proper (real) orthogonal group. The maximum is again a bilinear expression in the singular values but there is a subtracted term when A and B have determinants of opposite sign.

John von Neumann [1] proved a half century ago that if A and B are square matrices with complex elements, then

$$sup_{U,V\in U(n)} \mid tr(UAVB) \mid = \alpha_1\beta_1 + \alpha_2\beta_2 + ... + \alpha_n\beta_n ,$$

where $\alpha_1 \geq ... \geq \alpha_n$ are the singular values of A and $\beta_1 \geq ... \geq B_n$ the singular values of B, with the *sup* taken over all matrices U and V in the $n \times n$ unitary group $U(n)$. This theorem has attracted interest in applied linear algebra, including mathematical physics [7], psychology [9], the hyperelasticity of isotropic materials [18], and elsewhere, including [17, 19]. In this paper we consider matrices A and B with real elements, and we locate the value of $sup\, tr(UAVB)$ as the *sup* is taken over all elements U, V of $SO(n)$, the real proper orthogonal group.

A list [2-11] of articles simplifying the original von Neumann proof, or expanding the scope of the result, appears at the end of this paper. The earliest of these is the Fan paper [8]. There is a detailed analysis of the case of equality in [7]. New results related to the theorem are probably worthwhile, and ours are a natural

*Expositor

†Partially supported by a National Science Foundation grant

counterpart to the original theorem and seem not to be in the literature, at least not in [2-11]. Our results add to the slowly growing class of spectral inequalities having subtracted terms.

We augment the last sentence by explaninig that spectral inequalities with subtracted terms often occur in the study of singular values. See [13, 14, 15] for some examples. Many of these seemingly curious inequalities are best understood in terms of the properties of the root systems associated with the classical simple Lie groups and algebras.

Our proof technique in this paper is elementary, using no Lie theory, instead using a maximization technique often employed to establish spectral inequalities.

Theorem 1 *Taking the singular values α_i, β_i of the real matrices A and B in weakly decreasing order,*

$$sup_{P}, Q \in SO(n) tr(PAQB) = \alpha_1\beta_1 + \alpha_2\beta_2 + \ldots + \alpha_{n-1}\beta_{n-1} + sign(det(AB))\alpha_n\beta_n.$$

In particular, when A and B have determinants of opposite sign,

$$sup_{P,Q \in SO(n)} tr(PAQB) = \alpha_1\beta_1 + \alpha_2\beta_2 + \ldots + \alpha_n\beta_n.$$

Proof Since $SO(n) \times SO(n)$ is compact and trace is a continuous function, the sup in Theorem 1 is attained. We show that it at most is the value claimed in the theorem. Let P_0 and Q_0 be elements of $SO(n)$ at which the sup is attained. We are going to perturb the matrix P_0AQ_0B by a rotation and deduce certain information. Let $R_{ij}(\theta)$ be a rotation matrix, that is, an identity matrix apart from elements $cos\theta, sin\theta, -sin\theta, cos\theta$ in positions (i, i), (i, j), (j ,i) and (j, j), respectively. Then $tr(R_{ij}(\theta)P_0AQ_0B)$ achieves a maximum at $\theta = 0$, so that its derivative with respect to θ at $\theta = 0$. A simple computation shows that the (i, j) and (j, i) elements of P_0AQ_0B are the same. Application of this fact for all i and j shows that P_0AQ_0B is symmetric.

Since $tr(P_0AQ_0B) = tr(Q_0BP_0A)$, a similar computation shows that Q_0BP_0A is symmetric.

Let $S = P_0A$ and $T = Q_0B$. Then ST and TS are real symmetric matrices, with S having as its singular values those of A and T those B. By the singular value decomposition for real matrices, matrices O_1 and O_2 in $SO(n)$ exist such that

$$O_1SO_2 = diag(s_1, \ldots, s_n).$$

We may assume that the diagonal elements s_i in O_1SO_2 are nonnegative, except perhaps for the last, and are arranged in order of weakly decreasing absolute values. Thus $s_1 = \alpha_1, \ldots, s_{n-1} = \alpha_{n-1}, s_n = sign(det(A))\alpha_n$. Note that

$$tr(P_0AQ_0B) = tr(ST) = tr((O_1P_0)(AO_2)(O_2^{-1}Q_0)(BO_1^{-1})).$$

Renaming O_1P_0 as P_0, AO_2 as A, $O_2^{-1}asQ_0$ as BO_1^{-1} as B, O_1SO_2 as S, and $O_2^{-1}TO_1^{-1}$ as T, we now have $S = P_0A = diag(\alpha_1, \ldots, \alpha_{n-1}, sign(det(A))\alpha_n), T = Q_0B$, with ST and TS symmetric. Let $T = [t_{ij}]$.

We assert that the *trace* of ST is the *trace* of a product

$$diag(\alpha_1, ..., \alpha_{n-1}, sing(det(A))\alpha_n) diag(\pm\beta_{p(1)}, ..., \pm\beta_{p(n)}),$$

with ρ a permutation of $1, ..., n$, an with the product fo the $\pm$ signs in the right factor giving the sign of $detB$.

The symmetry of ST and TS implies that $s_i t_{ij} = s_j t_{ji}$ and $t_{ij} s_j = t_{ji} s_i$. Hence $(s_i^2 - s_j^2) t_{ij} = 0$. If $s_i^2 \neq s_j^2$ then $t_{ij} = 0$. If S^2 has distinct diagonal elements, then T must be diagonal. Because the diagonal elements of T ara $\pm$ the singular values of B and $detST = detAB$, our assertion is immediate, even if A or B is singular.

Since an inequality is being proved, we could avoid the in which S^2 has nondistint singular values case by appealing to the distinct singular value case and continuity. We prefer to give a direct analysis. Let S^2 have nondistinct diagonal elements. Then T splits as a direct sum of blocks: $T = diag(T_1, T_2, ..., T_{k-1}, T_k)$, say, corresponding to $S = diag(\sigma_1 I_1, \sigma_2 I_2, ..., \sigma_{k-1} I_{k-1}, \sigma_k D_k)$, with $\sigma_1 > \sigma_2 > ... > \sigma_{k-1} > \sigma_k \geq 0$. Here each I_i is an identity matrix but D_k departs from an identity in that the last diagonal entry is -1 exactly when $detA$ is negative. A simultaneous block diagonal similarity of S and T, with proper orthogonal diagonal blocks, permits us to take $T_1, ..., T_{k-1}$ to be diagonal, and also T_k when D_k is an identity matrix and σ_k is nonzero. If $\sigma_k = 0$, we may replace T_k by $P_k T_k Q_k$ where P_k and Q_k are proper orthogonal matrices diagonalizing T_k, and leave the products ST and TS unchanged. We only have to show how to replace T_k by a diagonal matrix when $detA$ is negative and D_k has -1 as its last diagonal element. The matrix $\sigma_k D_k T_k$ is symmetric, and its trace, as the sum of its eigenvalues, is a sum of terms each of which is the singular value σ_k of S times $\pm$ a singular value of $D_k T_k$, that is, times $\pm$ a singular value of T_k. The product of the $\pm$ signs is the sign of the product of the eigenvalues of $D_k T_k$ and therefore is the sign of $detD_k T_k$. Because the last diagonal element of D_k is -1, the product of the $\pm$ signs is the sign of $-detT_k = detD_k T_k$. Hence the *trace* of $\sigma_k D_k T_k$ is the *trace* a product ($\sigma_k D_k$ times a diagonal matrix of signed singular values of T_k), in which the signs of the singular values of T_k are those on then singular values of $D_k T_k$, except for the sign on one singular value, which for T_k is apposite to that of $D_k T_k$. From these facts, our assertion follows without any need to effect a diagonalization of T_k.

Thus

$$tr(P_0 A Q_0 B) = tr(ST) = \sum_{i=1}^{n} (\pm\alpha_i)(\pm\beta_{p(i)}).$$

with only α_n among α_i perhaps carrying a negative sign. If $detAB$ is negative, the positions of the negative entries on the α_i and on the $\beta_{p(i)}$ cannot completely be the same, so that at least one term $\alpha_i \beta_{p(i)}$ carries a negative sign. A simple rearrangement argument shows that when $det(AB)$ is nonnegative the sum cannot exceed.

$$\sum_{i=1}^{n} \alpha_i \beta_i$$

and when $det(AB)$ is negative

$$\sum_{i=1}^{n-1} \alpha_i\beta_i - \alpha_n\beta_n.$$

Returning to the original matrices A and B, before the notational changes, we have proved that the expressions just displayed are upper bounds for $tr(PAQB)$.

Moreover, these expressions are achievable values for $tr(PAQB)$ as P and Q range over $SO(n)$. Indeed, we may take $A = diag(\alpha_1, ..., \alpha_{n-1}, sign(det(A))\alpha_n)$, $B = diag(\beta_1, ..., \beta_{n-1}, sign(det(B))\beta_n)$, and the take $P = Q = I$. ∎

It is easy to see that in $f_{U,V\in U(n)} \mid tr(UAVB) \mid = 0$. Because of the absence of absolute values, the inf parallel to the sup in Theorem 1 is generally nonzero, and its value is left to the reader.

The generalization of Theorem 1 to more than two matrices is the content of Theorem 2. Its proof will give an alternative demonstration of Theorem 1.

Theorem 2 *Let $A_1, ..., A_m$ be matrices with real entries. Take the singular values of A_j to be $s_1(A_j) \geq ... \geq s_n(A_j)$ for $j = 1, ..., m$. Then, as matrices $P_1, ..., Pm$ range over $SO(n)$.*

$$sup_{P_1\in SO(n),...,P_m\in SO(n)} tr(P_1A_1...P_mA_m)$$

$$= \sum_{i=1}^{n-1}\prod_{j=1}^{m} s_i(A_j) + sign(det(A_1...A_m))\prod_{j=1}^{m} s_n(A_j).$$

Proof We shall use induction on m. No use is made of Theroem 1. The following argument includes then $m = 1$ case starting the induction.

Without loss of generality, we may suppose that $A_1, ..., A_m$ are diagonal, with the diagonal elements of each A_j in order of decreasing absolute values, and only the last possibly negative. As before, the sup is attained, so suppose that matrices $P_1, P_2, ..., P_m$ in $SO(n)$ achieve it. The matrices $A_1, ..., A_m$ may have multiple or zero singular values. Suppose, as an initial case, that each diagonal matrix A_i has simple nonzero singular values. Set $M = P_1A_1...P_{m-1}A_{m-1}P_m$. Let $R_{ij}(\theta)$ be a rotation matrix as before. Then $tr(R_{ij}(\theta)MA_m)$ has a maximum at $\theta = 0$, and so does $tr(MR_{ij}(\theta)A_m) = tr(R_{ij}(\theta)A_mM)$. Therefore MA_m is symmetric, and so is A_mM. Let $A_m = diag(\sigma_1, ..., \sigma_n)$, where the σ_i are distinct in absolute value. Then $M_{ij}\sigma_i$ and $\sigma_iM_{ij} = \sigma_jM_ji$. Therefore $(\sigma^2 - \sigma^2)M_{ij} = 0$, whence M is diagonal. Moreover, the diagonal elements of M are in order of weakly decreasing absolute values and only the last is possibly negative. For if not, by simple rearrangement inequalities, $tr(R^{-1}MRA_m)$ would be increased by a suitable choice of the generalized permutation matrix R in $SO(n)$.

When $m = 1$, by proper orthogonality the matrix $M = P_1$ must now be the identity, and the value of the sup in clear. Let $m > 1$.

Let $N = P_mA_mP_1A_1...P_{m-1}$. Then $trMA_m = trNA_{m-1}$, and by the same argument N is a diagonal matrix, with diagonal elements in order of weakly decreasing absolute values and only the last possibly negative.

Now

$$P_mA_mMP_m^{-1} = NA_{m-1}(= P_mA_mP_1A_1...P_{m-1}A_{m-1}).$$

Both A_mM and NA_{m-1} are diagonal matrices with nonzero diagonal elements which are in order of strictly decreasing absolute values since this is true for A_m and weakly so for M, and for A_{m-1} and weakly for N. Thus the similar diagonal matrices A_mM and NA_{m-1} have their necessarily simple eigenvalues appearing on the diagonal in the same order. Consequently the matrix P_m effecting the similarity must be a diagonal matrix, and therefore commutes with the diagonal matrix A_{m-1}. Hence

$$P_1A_1...P_{m-1}A_{m-1}P_mA_m = P_1A_1...(P_{m-1}P_m)(A_{m-1}A_m).$$

We are now in a position to apply induction. AS $P_1, ..., P_{m-2}, (P_{m-1}P_m)$ range over $SO(n)$,

$$sup\ tr(P_1A_1...P_{m-2}A_{m-2}(P_{m-1}P_m)(A_{m-1}A_m))$$

$$= \sum_{i=1}^{n-1} s_i(A_1)...s_i(A_{m-2})s_i(A_{m-1}A_m)$$

$$+sign(det(A_1...A_{m-2}(A_{m-1}A_m)))s_n(A_1)...s_n(A_{m-2})s_n(A_{m-1}A_m)$$

$$= \sum_{i=1}^{n-1} s_i(A_1)...s_i(A_{m-1})s_i(A_m)$$

$$+sign(det(A_1...A_{m-1}A_m))s_n(A_1)...s_n(A_{m-1}s_n(A_m)$$

Therefore, for $m \geq 1$, the sup is its claimed value when the A_i have simple nonzero singular values. Now suppose the A_i do not all have simple nonzero singular values. Choose the P_i so that the sup is attained, and then perturb the A_i to have simple nonzero singular values. The upper bound on the *trace* is then valid for the chosen P_i and the perturbed A_i. By continuity it continues to be an upper bound as the perturbations approach zero, whence it is an upper bound for the original matrices A_i.

It is clear that the upper bound on the *trace* is achieved for suitable matrices P_i in $SO(n)$. ■

The case of equality in the von Neumannn theorem seems to be analysed only in [7]. The full result is somewhat intricate, but becomes a bit simpler if the von Neumann theorem is stated in another way. A later paper examining cases of equality in the von Neumann result and our proper orthogonal version of it will be prepared if sufficiently significat results are found.

References

[1] von Neumann J., *Some matrix inequalities and metrization of matriz s-pace*, Tomsk. Univ. Rev. **1** **286-300** **(1937)**; Collected Works, Pergamon Press, New York, **4** **205-219** **(1962)**.

[2] Alberti P.M., Uhlmann A., *Stochasticity and partial order*, Reidel, Dordrec, 1982.

[3] Eaton M.L., *Lecture on topics in probability inequalities*, Centrum voor Wiskunde en Informatica, Amsterdam, 1987.

[4] Horn R., Johnson C.R., *Matriz analysis*, Cambridge Univ. Press, Cambridge, 1985.

[5] Mirsky L., *A trace inequality of John von Neumann*, Monat. fur Math., 79 303-306 (1975).

[6] Schatten R., *Norm ideals of completely continuons operators*, Springer-Verlag, Berlin, 1960.

[7] Capel H.W., Tindemans P.A.J., *An inequality for the trace of matriz products*, Report on Mathematical Physics bf 6 225-235 (1974).

[8] Ky Fan, *Maximun properties and inequalities for the eigenvalues of completely continous operators*, Proc. Nat. Aca. Sci. USA 37 760-766 (1951).

[9] Kristoff W., *A theorem on the trace of certain matriz products and some applications*, J. Math. Psych. 7 515-530 (1970).

[10] Marcus M., Moyls B.N., *On the maximum principle of Ky Fan*, Canada. J. Math. 9 313-320 (1957).

[11] Olkin I., Marshall A., *Inequalities: Theory of majorization and its applications*, Academic Press, New York, (1979).

[12] ,*Doubly stochastic matrices and the diagonal of a rotation matrix*, Amer. J. Math. 76 620-630 (1954).

[13] Thompson R.C., *Singular values, diagonal elements, and convexity*, SIAMJ. Appl. Math. 32 39-63 (1977).

[14] Thompson R.C., *The intersection of the convex hulls of proper and improper real matrices with prescribed singular values*, Linear Multilinear Alg., 3 155-160 (1975).

[15] Thompson R.C., *Singular values and diagonal elements of complex symmetric matrices*, Linear Alg. Appl. 26 65-106 (1979).

[16] Tromberg B. and Waldenstrom S., *Bounds on the diagonal elements of unitary matrices*, Linear Alg. Appl. 20 189-195 (1978).

[17] Brockett R.W., *Dynamical systems that sort lists, diagonalize matrices, and solve linear propramming problems*, Linear. Appl. 146 79-91 (1991).

[18] Le Dret H., *Sur les functions de matrices convexes et isotropes*, C. R: Acad. Paris 310 Series bf I 617-620 (1990).

[19] Stewart G.W., Ji-guang Sun, *Matrix perturbation theory*, Academic Press, Boston, (1990).

Item 1 is the original paper, items 2-6 include proofs of the von Neumann inequality for two matrices, and 7-11 include the generalization to an arbitrary number of matrices. Item 7 discusses the case of equatlity. None of 1-11 investigate the case in which the sup is taken over the proper orthogonal group. items12-16 relate to spectral inequalities having subtracted terms, and 17-19 contain applications of the theorem.

Dirección de los autores:
Department of Mathematics
University of California
Santa Barbara, CA 9316, USA