

# A Representation Theorem for $(\text{tr} A^p)^{1/p}$

Jan R. Magnus

London School of Economics

Houghton Street

London, WC2A 2AE England

Submitted by Ingram Olkin

---

## ABSTRACT

An inequality for positive semidefinite matrices is proved, and from it a quasilinear representation of  $(\text{tr} A^p)^{1/p}$  is obtained. From this representation follow matrix versions of the inequalities of Hölder and Minkowski.

---

## 1. INTRODUCTION

The fundamental inequalities of Hölder and Minkowski can be proved in a variety of ways; see Hardy, Littlewood, and Pólya (1952). A particularly interesting proof of both inequalities is based on the following lemma, of interest in itself, which is easily established by the method of Lagrange or otherwise.

LEMMA 1.<sup>1</sup> Let  $p > 1$ ,  $q = p/(p - 1)$ , and  $a_i \geq 0$  ( $i = 1, \dots, n$ ). Then

$$\sum_{i=1}^n a_i x_i \leq \left( \sum_{i=1}^n a_i^p \right)^{1/p} \quad (1.1)$$

for every set of nonnegative quantities  $x_1, x_2, \dots, x_n$  satisfying  $\sum_{i=1}^n x_i^q = 1$ . Equality in (1.1) occurs if and only if  $a_1 = a_2 = \dots = a_n = 0$  or  $x_i^q = a_i^p / \sum_{j=1}^n a_j^p$  ( $i = 1, \dots, n$ ).

---

<sup>1</sup>See Beckenbach and Bellman (1961, Theorem 5, p. 23).

An immediate and easy consequence of Lemma 1 is

**HÖLDER'S INEQUALITY.** *If  $x = (x_1, \dots, x_n)$  and  $y = (y_1, \dots, y_n)$  are vectors in the nonnegative orthant of  $\mathbb{R}^n$  and  $0 < \alpha < 1$ , then*

$$\sum_{i=1}^n x_i^\alpha y_i^{1-\alpha} \leq \left( \sum_{i=1}^n x_i \right)^\alpha \left( \sum_{i=1}^n y_i \right)^{1-\alpha} \quad (1.2)$$

with equality if and only if  $x$  and  $y$  are linearly dependent.

If we formulate the inequality (1.1) as a maximization problem, we get

$$\max_R \sum_{i=1}^n a_i x_i = \left( \sum_{i=1}^n a_i^p \right)^{1/p}, \quad (1.3)$$

where  $R$  is the region defined by

$$\sum_{i=1}^n x_i^q = 1, \quad x_i \geq 0.$$

The idea of expressing a *nonlinear* function, such as  $(\sum a_i^p)^{1/p}$ , as an envelope of *linear* functions goes back to Minkowski and was used extensively by Bellman and others. This technique is called *quasilinearization*. A direct consequence of the quasilinear representation (1.3) is

**MINKOWSKI'S INEQUALITY.** *If  $x = (x_1, \dots, x_n)$  and  $y = (y_1, \dots, y_n)$  are vectors in the nonnegative orthant of  $\mathbb{R}^n$  and  $p > 1$ , then*

$$\left( \sum_{i=1}^n (x_i + y_i)^p \right)^{1/p} \leq \left( \sum_{i=1}^n x_i^p \right)^{1/p} + \left( \sum_{i=1}^n y_i^p \right)^{1/p} \quad (1.4)$$

with equality if and only if  $x$  and  $y$  are linearly dependent.

The purpose of this paper is to extend the inequality (1.1) to positive semidefinite matrices, thus obtaining a quasilinear representation of  $(\text{tr } A^p)^{1/p}$ . This is achieved in Theorem 5. Interesting matrix versions of the inequalities of Hölder and Minkowski are then easily derived from this representation (Theorems 6 and 7). Several preliminary results, of interest in themselves, are also reported.

## 2. A NECESSARY AND SUFFICIENT CONDITION FOR DIAGONALITY

Let us begin by stating the following simple but useful result concerning the conditions under which a real symmetric (or positive definite) matrix is diagonal.

**THEOREM 1.** *A positive definite matrix is diagonal if and only if the product of its diagonal elements is equal to its determinant. A real symmetric matrix is diagonal if and only if its diagonal elements and its eigenvalues coincide.*

*Proof.* A proof by induction and perturbation is straightforward and therefore omitted. A variety of other proofs can also be constructed. ■

## 3. KARAMATA'S INEQUALITY

Let  $x = (x_1, x_2, \dots, x_n)$  and  $y = (y_1, y_2, \dots, y_n)$  be two vectors in  $\mathbb{R}^n$ . We shall say that  $x$  is *majorized* by  $y$ , and write

$$(x_1, \dots, x_n) < (y_1, \dots, y_n),$$

when the following three relations are satisfied:

$$x_1 + x_2 + \dots + x_n = y_1 + y_2 + \dots + y_n;$$

$$x_1 \geq x_2 \geq \dots \geq x_n, \quad y_1 \geq y_2 \geq \dots \geq y_n;$$

$$x_1 + x_2 + \dots + x_k \leq y_1 + y_2 + \dots + y_k \quad (1 \leq k \leq n-1).$$

The theory of majorization originated with the work of Schur (1923) and Hardy, Littlewood, and Pólya (1929), and is now firmly established; see Marshall and Olkin (1979). The following theorem was essentially proved by Schur (1923). A continuous analogue was proved by Hardy, Littlewood, and Pólya (1929), and an important generalization provided by Karamata (1932).

**THEOREM 2.** *Let  $\phi$  be a real-valued convex function defined on an interval  $I \subset \mathbb{R}$ . If  $(x_1, \dots, x_n) < (y_1, \dots, y_n)$  on  $I^n$ , then*

$$\sum_{i=1}^n \phi(x_i) \leq \sum_{i=1}^n \phi(y_i). \quad (3.1)$$

If, in addition,  $\phi$  is strictly convex on  $I$ , then equality in (3.1) occurs if and only if  $x_i = y_i$  ( $i = 1, \dots, n$ ).

*Proof.* See Marshall and Olkin (1979, Propositions 3.C.1 and 3.C.1.a). ■

An important application of Karamata's inequality is given in Theorem 3.

**THEOREM 3.** Let  $A = (a_{ij})$  be a real symmetric  $n \times n$  matrix with eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$ . Then for any convex function  $\phi: \mathbb{R} \rightarrow \mathbb{R}$  we have

$$\sum_{i=1}^n \phi(a_{ii}) \leq \sum_{i=1}^n \phi(\lambda_i). \quad (3.2)$$

Moreover, if  $\phi$  is strictly convex, then equality in (3.2) occurs if and only if  $A$  is diagonal.

*Proof.* Without loss of generality we may assume that

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n, \quad a_{11} \geq a_{22} \geq \dots \geq a_{nn}.$$

It is well known that  $(a_{11}, \dots, a_{nn})$  is majorized by  $(\lambda_1, \dots, \lambda_n)$ ; see Schur (1923) or Marshall and Olkin (1979, Theorem 9.B.1). Thus (3.2) follows from Theorem 2. For strictly convex  $\phi$ , Theorem 2 implies that equality in (3.2) holds if and only if  $a_{ii} = \lambda_i$  for  $i = 1, \dots, n$ , and, by Theorem 1, this is the case if and only if  $A$  is diagonal. ■

As a corollary of Theorem 3, let us establish the following useful inequalities concerning positive semidefinite matrices, which we shall use in proving Theorem 5.

**THEOREM 4.** Let  $A = (a_{ij})$  be a positive semidefinite  $n \times n$  matrix. Then

$$\operatorname{tr} A^p \geq \sum_{i=1}^n a_{ii}^p \quad (p > 1)$$

and

$$\operatorname{tr} A^p \leq \sum_{i=1}^n a_{ii}^p \quad (0 < p < 1)$$

with equality if and only if  $A$  is diagonal.

*Proof.* Let  $p > 1$ , and define  $\phi(x) = x^p$  ( $x \geq 0$ ). The function  $\phi$  is strictly convex. Hence Theorem 3 implies that

$$\operatorname{tr} A^p = \sum_{i=1}^n \lambda_i^p(A) = \sum_{i=1}^n \phi(\lambda_i(A)) \geq \sum_{i=1}^n \phi(a_{ii}) = \sum_{i=1}^n a_{ii}^p$$

with equality if and only if  $A$  is diagonal. Next, let  $0 < p < 1$ . Then  $\psi(x) \equiv -x^p$  ( $x \geq 0$ ) is strictly convex, and the second result follows in the same way. ■

#### 4. QUASILINEAR REPRESENTATION OF $(\operatorname{tr} A^p)^{1/p}$

We now have all the ingredients to prove the matrix analogue of Lemma 1.

**THEOREM 5.** *Let  $p > 1$ ,  $q = p/(p - 1)$ , and let  $A \neq 0$  be a positive semidefinite  $n \times n$  matrix. Then*

$$\operatorname{tr} AX \leq (\operatorname{tr} A^p)^{1/p} \tag{4.1}$$

for every positive semidefinite  $n \times n$  matrix  $X$  satisfying  $\operatorname{tr} X^q = 1$ . Equality in (4.1) occurs if and only if

$$X^q = \frac{1}{\operatorname{tr} A^p} A^p. \tag{4.2}$$

**REMARK.** If we let  $R$  be the region

$$R = \{ X : X \in \mathbb{R}^{n \times n}, X \text{ positive semidefinite, } \operatorname{tr} X^q = 1 \},$$

then we can state Theorem 5 equivalently as

$$\max_R \operatorname{tr} AX = (\operatorname{tr} A^p)^{1/p} \tag{4.3}$$

for every positive semidefinite  $n \times n$  matrix  $A$ . Thus we can express  $(\operatorname{tr} A^p)^{1/p}$  (a *nonlinear* function of  $A$ ) as an envelope of *linear* functions of  $A$ . In other

words, we obtain a quasilinear representation of  $(\operatorname{tr} A^p)^{1/p}$ .

*Proof.* Let  $X$  be an arbitrary positive semidefinite  $n \times n$  matrix satisfying  $\operatorname{tr} X^q = 1$ . Let  $S$  be an orthogonal matrix such that  $S'XS = \Lambda$ , where  $\Lambda$  is diagonal and has the eigenvalues of  $X$  as its diagonal elements. Define  $B = (b_{ij}) = S'AS$ . Then

$$\operatorname{tr} AX = \operatorname{tr} B\Lambda = \sum b_{ii}\lambda_i$$

and

$$\operatorname{tr} X^q = \operatorname{tr} \Lambda^q = \sum \lambda_i^q.$$

Hence, by Lemma 1,

$$\operatorname{tr} AX = \sum b_{ii}\lambda_i \leq \left( \sum b_{ii}^p \right)^{1/p}. \quad (4.4)$$

Since  $A$  is positive semidefinite, so is  $B$ , and Theorem 4 thus implies that

$$\sum b_{ii}^p \leq \operatorname{tr} B^p. \quad (4.5)$$

Combining (4.4) and (4.5) we obtain

$$\operatorname{tr} AX \leq (\operatorname{tr} B^p)^{1/p} = (\operatorname{tr} A^p)^{1/p}.$$

Equality in (4.4) occurs if and only if

$$\lambda_i^q = \frac{b_{ii}^p}{\sum b_{ii}^p} \quad (i = 1, \dots, n),$$

and equality in (4.5) if and only if  $B$  is diagonal. Hence, equality in (4.1) occurs if and only if

$$\Lambda^q = \frac{B^p}{\operatorname{tr} B^p},$$

which is equivalent to (4.2). This concludes the proof. ■

5. MATRIX ANALOGUES OF THE INEQUALITIES OF HÖLDER AND MINKOWSKI

An immediate consequence of Theorem 5 is the matrix analogue of Hölder’s inequality (1.2).

**THEOREM 6.** *For any two positive semidefinite matrices  $A$  and  $B$  of the same order,  $A \neq 0$ ,  $B \neq 0$ , and  $0 < \alpha < 1$ , we have*

$$\text{tr } A^\alpha B^{1-\alpha} \leq (\text{tr } A)^\alpha (\text{tr } B)^{1-\alpha}, \tag{5.1}$$

*with equality if and only if  $B = \mu A$  for some scalar  $\mu > 0$ .*

*Proof.* Let  $p = 1/\alpha$ ,  $q = 1/(1 - \alpha)$ , and assume  $B \neq 0$ . Now define

$$X = \frac{B^{1/q}}{(\text{tr } B)^{1/q}}.$$

Then  $\text{tr } X^q = 1$ , and hence Theorem 5 applied to  $A^{1/p}$  yields

$$\text{tr } A^{1/p} B^{1/q} \leq (\text{tr } A)^{1/p} (\text{tr } B)^{1/q},$$

which is (5.1). According to Theorem 5, equality in (5.1) can only occur if  $X^q = (1/\text{tr } A)A$ , that is, if  $B = \mu A$  for some  $\mu > 0$ . ■

Another consequence of Theorem 5, more specifically of its quasilinear representation (4.3), is the matrix version of Minkowski’s inequality (1.4).

**THEOREM 7.** *For any two positive semidefinite matrices  $A$  and  $B$  of the same order ( $A \neq 0$ ,  $B \neq 0$ ), and  $p > 1$ , we have*

$$[\text{tr } (A + B)^p]^{1/p} \leq (\text{tr } A^p)^{1/p} + (\text{tr } B^p)^{1/p}. \tag{5.2}$$

*with equality if and only if  $A = \mu B$  for some  $\mu > 0$ .*

*Proof.* Let  $p > 1$ ,  $q = p/(p - 1)$ , and let  $R$  be the region

$$R = \{ X : X \in \mathbb{R}^{n \times n}, X \text{ positive semidefinite, } \text{tr } X^q = 1 \}.$$

Using the quasilinear representation (4.3) of  $(\operatorname{tr} A^p)^{1/p}$ ,

$$\max_R \operatorname{tr} AX = (\operatorname{tr} A^p)^{1/p},$$

we obtain

$$\begin{aligned} [\operatorname{tr} (A + B)^p]^{1/p} &= \max_R \operatorname{tr} (A + B)X \\ &\leq \max_R \operatorname{tr} AX + \max_R \operatorname{tr} BX \\ &= (\operatorname{tr} A^p)^{1/p} + (\operatorname{tr} B^p)^{1/p}. \end{aligned} \quad (5.3)$$

Equality in (5.3) can only occur if the same  $X$  maximizes  $\operatorname{tr} AX$ ,  $\operatorname{tr} BX$ , and  $\operatorname{tr} (A + B)X$ , which implies, by Theorem 5, that  $A^p$ ,  $B^p$ , and  $(A + B)^p$  are proportional, and hence that  $A$  and  $B$  are proportional. ■

*I am very grateful to Ken Binmore for his constructive remarks.*

#### REFERENCES

- Beckenbach, E. F. and Bellman, R. 1961. *Inequalities*, Springer, Berlin.
- Hardy, G. H., Littlewood, J. E., and Pólya, G. 1929. Some simple inequalities satisfied by convex functions, *Messenger Math.* 58:145–152.
- Hardy, G. H., Littlewood, J. E., and Pólya, G. 1952. *Inequalities* (2nd ed.), Cambridge U.P., Cambridge.
- Karamata, J. 1932. Sur une inégalité relative aux fonctions convexes, *Publ. Math. Univ. Belgrade* 1:145–148.
- Marshall, A. W. and Olkin, I. 1979. *Inequalities: Theory of Majorization and Its Applications*, Academic, New York.
- Schur, I. 1923. Ueber eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie, *Sitzungsber. Berlin. Math. Ges.* 22:9–20.

*Received 23 December 1985; revised 17 November 1986*