On Power Means of Positive Quadratic Forms

Erwin Lutwak
Department of Mathematics
Polytechnic Institute of New York
333 Jay Street
Brooklyn, New York 11201

Submitted by Richard A. Brualdi

ABSTRACT

Some power means of positive definite quadratic forms are closely related to the fundamental scalar functions of the matrix associated with the quadratic form. This relation can (among other things) be used to give new proofs of some of the classical matrix inequalities.

Associated with a real $n \times n$ symmetric matrix $A = (a_{ij})$ is the quadratic form

$$Q_A(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j,$$

where x denotes the vector $(x_1, \dots, x_n)$. The matrix is called positive definite if the associated quadratic form is positive for all nonzero vectors x. We shall be concerned with the restriction of Q_A to the unit sphere, S^{n-1}:

$$Q_A : S^{n-1} \to (0, \infty),$$

where $(0, \infty)$ denotes the set of positive reals.

We shall investigate some of the properties of the power means of the function Q_A. Some of the power means of Q_A are closely related to some of the fundamental scalar functions (such as the determinant and trace) of A. Among other things, this relation leads to new proofs of some of the classical matrix inequalities.

For reference regarding matrices, quadratic forms, and matrix inequalities, the reader is referred to [1]. Chapter 6 of [1] contains some material in a spirit

52 Vanderbilt Ave., New York, NY 10017 0024-3795/84/$3.00

similar to that of this note. For material concerning power means and analytic inequalities, the reader is referred to [2].

Given a continuous function

$$f: S^{n-1} \to (0, \infty),$$

and a real number $p \neq 0$, the p-mean of f, $M_p[f]$, is defined by

$$M_p[f] = \left[\frac{1}{n\omega_n} \int_{S^{n-1}} f^p(u)\, dS(u) \right]^{1/p},$$

where ω_n denotes the volume of the unit ball in Euclidean n-space, R^n, and $dS(u)$ denotes the area element of S^{n-1} at u. For $p = -\infty$, 0, or ∞, the p-mean of f is defined by

$$M_p[f] = \lim_{r \to p} M_r[f].$$

For a positive definite matrix A (or the quadratic form associated with A), we define the p-mean of A, $\Phi_p[A]$, by

$$\Phi_p[A] = M_{p/2}[Q_A].$$

Hence, for a real $p \neq 0$, we have:

$$\Phi_p[A] = \left[\frac{1}{n\omega_n} \int_{S^{n-1}} (u, Au)^{p/2}\, dS(u) \right]^{2/p}, \qquad (*)$$

where $(\cdot, \cdot)$ denotes the usual inner product in R^n.

We list some of the consequences of our definition. (All matrices A, B are assumed to be positive definite. All scalars λ are also assumed to be positive.)

1. *For all p and all positive scalars λ,*

$$\Phi_p[\lambda A] = \lambda \Phi_p[A]$$

and

$$\Phi_p[I] = 1,$$

where I denotes the unit matrix.

2. *For a fixed matrix* A, $\Phi_p[A]$ *is continuous in* p.

Proof. This follows directly from the continuity (in p) of the power means M_p (see [2, p. 143]). ∎

3. *If* $-\infty \leqslant p < q \leqslant \infty$, *then*

$$\Phi_p[A] \leqslant \Phi_q[A],$$

with equality if and only if $A = \lambda I$.

Proof. This follows from the observation that unless f is constant, $M_p[f]$ is strictly increasing in p (see [2, p. 144]) and that for symmetric A, we have Q_A constant if and only if $A = \lambda I$. ∎

4. *If* $-\infty < p < q < r < \infty$, *then*

$$\Phi_q[A]^{q(r-p)} \leqslant \Phi_p[A]^{p(r-q)} \Phi_r[A]^{r(q-p)},$$

with equality if and only if $A = \lambda I$.

Proof. If one of the indices (p, q, or r) is 0, the inequality reduces to the inequality in statement 3. If none of the indices is equal to 0, then the result is obtained by a direct application of the Hölder integral inequality [2, p. 140]. ∎

5. *If* $A \leqslant B$, *and* $-\infty < p < \infty$, *then*

$$\Phi_p[A] \leqslant \Phi_p[B],$$

with equality if and only if $A = B$.

Proof. The inequality follows easily from the representation given in (*). The conditions for equality follow from the observation that, for symmetric A and B, we have $Q_A = Q_B$ if and only if $A = B$. ∎

6. *If* $p > 2$, *then*

$$\Phi_p[A + B] \leqslant \Phi_p[A] + \Phi_p[B],$$

with equality if and only if $A = \lambda B$. *If* $p < 2$, *then*

$$\Phi_p[A+B] \geqslant \Phi_p[A] + \Phi_p[B],$$

with equality if and only if $A = \lambda B$. *For* $p = 2$,

$$\Phi_2[A+B] = \Phi_2[A] + \Phi_2[B].$$

Proof. The inequalities are simple consequences of the Minkowski integral inequality [2, p. 146] (see also [2, p. 138]). The conditions for equality follow from the observation that, for symmetric matrices A and B, we have $Q_A = \lambda Q_B$ if and only if $A = \lambda B$. ■

7. *If* A *and* B *are orthogonally equivalent, then*

$$\Phi_p[A] = \Phi_p[B]$$

for all p.

Proof. The proof is straightforward. For orthogonally equivalent matrices A and B, there is an orthogonal matrix P such that $A = P'BP$ (where P' denotes the transpose of P). For real $p \neq 0$

$$\Phi_p[A] = \left[\frac{1}{n\omega_n}\int_{S^{n-1}} (u, P'BPu)^{p/2}\, dS(u)\right]^{2/p}$$
$$= \left[\frac{1}{n\omega_n}\int_{S^{n-1}} (Pu, BPu)^{p/2}\, dS(u)\right]^{2/p}.$$

We make the change of variables $v = Pu$ and note that since P is orthogonal, the transformation maps S^{n-1} bijectively onto S^{n-1} and has a Jacobian of absolute value 1. Hence, we have

$$\Phi_p[A] = \left[\frac{1}{n\omega_n}\int_{S^{n-1}} (v, Bv)^{p/2}\, dS(v)\right]^{2/p}$$
$$= \Phi_p[B].$$

For $p = -\infty$, 0, or ∞, a limit argument now yields the desired result. ■

8.

$$\phi_2[A] = \frac{1}{n}\operatorname{tr}(A),$$

where tr(A) *denotes the trace of* A.

Proof. Since every positive definite matrix is orthogonally equivalent to a positive definite diagonal matrix, by the previously established result we need prove this only for diagonal matrices. For a positive definite diagonal matrix D with diagonal entries $\lambda_1,\dots,\lambda_n$ we have

$$\Phi_2[D] = \frac{1}{n\omega_n}\int_{S^{n-1}}(u, Du)\,dS(u)$$

$$= \frac{1}{n\omega_n}\int_{S^{n-1}}\sum_{i=1}^{n}\lambda_i u_i^2\,dS(u),$$

where $u = (u_1,\dots,u_n)$. Since

$$\frac{1}{\omega_n}\int_{S^{n-1}}u_i^2\,dS(u) = 1,$$

the result follows. ∎

9.

$$\Phi_\infty[A] = \lambda_1(A) \quad \textit{and} \quad \Phi_{-\infty}[A] = \lambda_n(A),$$

where $\lambda_1(A)$ *and* $\lambda_n(A)$ *denote the largest and smallest of the eigenvalues of* A, *respectively*.

Proof. This follows from the observations (see [2, p. 144]) that

$$M_\infty[f] = \max\{\, f(u) \mid u \in S^{n-1}\},$$

while

$$M_{-\infty}[f] = \min\{\, f(u) \mid u \in S^{n-1}\},$$

and (see [1, p. 113]) that

$$\lambda_1(A) = \max\{ (u, Au) \mid u \in S^{n-1} \},$$

while

$$\lambda_n(A) = \min\{ (u, Au) \mid u \in S^{n-1} \}. \qquad \blacksquare$$

10.

$$\Phi_{-n}[A] = |A|^{1/n},$$

where $|A|$ *denotes the determinant of* A.

Proof. To prove this we first observe that a positive definite matrix A determines an ellipsoid, $\varepsilon(A)$, defined by

$$\varepsilon(A) = \{ x \mid (x, Ax) \leqslant 1 \}.$$

It is trivial to verify that the polar coordinate equation of the surface of this ellipsoid is given by

$$r(u) = (u, Au)^{-1/2} \qquad [u \in S^{n-1}].$$

Thus, if we use the formula for the (n-dimensional) volume in polar coordinates, we have

$$\mathrm{Vol}[\varepsilon(A)] = \frac{1}{n} \int_{S^{n-1}} (u, Au)^{-n/2} \, dS(u).$$

On the other hand, the linear transformation $y = \sqrt{A}\,x$, where $\sqrt{A}$ denotes the positive square root of A, magnifies (n-dimensional) volume by a factor of $|\sqrt{A}| = |A|^{1/2}$. Since this transformation maps $\varepsilon(A)$ into the unit ball (which has volume ω_n), it follows that

$$|A|^{1/2} \mathrm{Vol}[\varepsilon(A)] = \omega_n.$$

Hence, we have

$$|A|^{-1/2} = \frac{1}{n\omega_n}\int_{S^{n-1}} (u, Au)^{-n/2}\, dS(u). \qquad (**)$$

from which the result follows. ■

Several of the classical matrix inequalities are immediate consequences of the results developed above. For example, by combining statements 3, 8, and 10, we obtain

$$|A|^{1/n} \leqslant \frac{1}{n}\operatorname{tr}(A),$$

with equality if and only if $A = \lambda I$.

If we combine 6 and 10, we obtain the Minkowski determinant inequality:

$$|A+B|^{1/n} \geqslant |A|^{1/n} + |B|^{1/n},$$

with equality if and only if $A = \lambda B$.

Thus, the Minkowski determinant inequality can be viewed as a special case of the Minkowski integral inequality.

Relations between the power means of a matrix and the fundamental scalar functions of A (other than the ones obtained above) can be obtained by using 10 [or its equivalent $(**)$], substituting $A + \lambda I$ for A, expanding, and equating coefficients.

REFERENCES

1 R. Bellman, *Introduction to Matrix Analysis*, 2nd ed., McGraw-Hill, New York, 1970.
2 G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge U.P., 1934.

Received 9 August 1982