Pacific Journal of Mathematics

EQUALITY IN CERTAIN INEQUALITIES

Marvin David Marcus and Afton Herbert Cayford

Vol. 13, No. 4 June 1963

EQUALITY IN CERTAIN INEQUALITIES

MARVIN MARCUS AND AFTON CAYFORD

1. Introduction. Let $\sigma = (\sigma_1, \cdots, \sigma_n)$ be a point on the unit $(n-1)$-simplex S^{n-1}: $\sum_{i=1}^n \sigma_i = 1, \sigma_i \geqq 0$. Let $0 < \lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_n$ and $\mu_1 \geqq \mu_2 \geqq \cdots \geqq \mu_n > 0$ be positive numbers and form the function on S^{n-1}

$$F(\sigma) = \sum_{i=1}^n \sigma_i \lambda_i \sum_{i=1}^n \sigma_i \mu_i \, . \tag{1.1}$$

The main purpose of this paper is to examine the structure of the set of points $\sigma \in S^{n-1}$ for which $F(\sigma)$ takes on its maximum value. In case a convex monotone decreasing function f is fitted to the points (λ_i, μ_i)(i.e. $f(\lambda_i) = \mu_i$), $i = 1, \cdots, n$, then it is not difficult to show that the maximum for $F(\sigma)$ on S^{n-1} is the upper bound given by M. Newman [4] in a recent interesting paper. In the case of the Kantorovich inequality [1] the function f is $f(t) = t^{-1}$, $\mu_i = \lambda_i^{-1}$, $i = 1, \cdots, n$. In this case a maximizing σ is $\sigma_1 = 1/2, \sigma_n = 1/2, \sigma_i = 0$, $i = 2, \cdots, n-1$, and if $\lambda_1 < \lambda_k < \lambda_n$, $k = 2, \cdots, n-1$, it is a corollary of our main result (Theorem 2) that this is the only choice possible for $\sigma \in S^{n-1}$ in order to achieve the maximum value.

We shall assume henceforth in this paper that $\mu_i = f(\lambda_i)$, $i = 1, \cdots, n$, where f is a monotone decreasing convex function defined on the closed interval $[\lambda_1, \lambda_n]$. In 2 we determine the structure of the set of $\sigma \in S^{n-1}$ for which $F(\sigma)$ is a maximum in the case in which f is assumed to be strictly convex. In 3 we investigate the structure of the set of unit vectors x for which the function

$$\varphi(x) = (Ax, x)(f(A)x, x) \tag{1.2}$$

assumes its maximum value on the unit sphere $\| x \| = 1$. Throughout, A is a positive definite hermitian transformation on an n-dimensional unitary space U with inner product (x, y). The eigenvalues of A are λ_i, $0 < \lambda_1 \leqq \cdots \leqq \lambda_n$, with corresponding orthonormal eigenvectors u_i, $Au_i = \lambda_i u_i$, $i = 1, \cdots, n$. Of particular interest in (1.2) is the choice $f(t) = t^{-p}$, $p > 0$.

Finally, in 4, we discuss the applications of the previous results to Grassmann compounds and induced power transformations associated with A. In two recent papers [2, 5] the Kantorovich inequality was applied to the compound to obtain inequalities involving principal subdeterminants of a positive definite hermitian matrix. We shall prove (Theorem 5) that these inequalities are in fact strict except in

Received October 30, 1962.

trivial cases. Similar inequalities are obtained for the permanent function together with a discussion of the cases of equality. These inequalities are believed to be new.

2. Maximum values for F. In the rest of the paper M will systematically denote the maximum value of $F(\sigma)$, $\sigma \in S^{n-1}$, and m will denote the largest of $\lambda_1\mu_1$ and $\lambda_n\mu_n$. Also, γ will denote the number $(\lambda_1\mu_n + \lambda_n\mu_1)/2$. The main result of this section is Theorem 2 which describes the structure of those σ for which $F(\sigma) = M$ when f is strictly convex. We first prove

THEOREM 1. *For any $\sigma \in S^{n-1}$ there exists a $\beta \in [0, 1]$ such that*

$$F(\sigma) \leqq (\beta\lambda_1 + (1-\beta)\lambda_n)(\beta\mu_1 + (1-\beta)\mu_n) . \tag{2.1}$$

If f is strictly convex and for some k, $1 \leqq k \leqq n$, $\lambda_1 < \lambda_k < \lambda_n$ and $\sigma_k > 0$ then there exists a $\beta \in [0, 1]$ for which (2.1) *is a strict inequality.*

To prove Theorem 1 we use the following elementary fact.

LEMMA. *If $0 \leqq a_1 \leqq a_2 \leqq a_3$, and $b_1 \geqq b_2 \geqq b_3 \geqq 0$ and*

$$(a_1 - a_3)(b_2 - b_3) \geqq (a_2 - a_3)(b_1 - b_3) \tag{2.2}$$

then for any $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in S^2$ there exists a $\beta \in [0, 1]$ such that

$$\sum_{i=1}^{3} \alpha_i a_i \sum_{i=1}^{3} \alpha_i b_i \leqq (\beta a_1 + (1-\beta)a_2)(\beta b_1 + (1-\beta)b_3) . \tag{2.3}$$

If the inequality (2.2) *is strict and $\alpha_2 > 0$ then there exists a $\beta \in [0, 1]$ such that* (2.3) *is strict.*

Proof. Let θ and ω in $[0, 1]$ be so chosen that $a_2 = \theta a_1 + (1-\theta)a_3$, $b_2 = \omega b_1 + (1-\omega)b_3$ and set $b_2' = \theta b_1 + (1-\theta)b_3$. Then

$$b_2' - b_2 = (\theta - \omega)(b_1 - b_3) . \tag{2.4}$$

Assume first that $a_3 > a_2$ and $b_2 > b_3$. Then $\theta = (a_2 - a_3)/(a_1 - a_3) > 0$ and $\omega = (b_2 - b_3)/(b_1 - b_3)$. Moreover $\theta \geqq \omega$ by (2.2) and if (2.2) is strict then $\theta > \omega$. From (2.4) $b_2' - b_2 \geqq 0$ and we compute that

$$\begin{aligned} L \leqq {} & ((\alpha_1 + \theta\alpha_2)a_1 + (\alpha_2(1-\theta) + \alpha_3)a_3) \\ & ((\alpha_1 + \theta\alpha_2)b_1 + (\alpha_2(1-\theta) + \alpha_3)b_3) , \end{aligned} \tag{2.5}$$

where L is the left side of (2.3). This is (2.3) with $\beta = \alpha_1 + \theta\alpha_2 \in [0, 1]$. If (2.2) is strict then $\theta > \omega$, $b_2' = b_2$, and $\alpha_2 > 0$ together imply that (2.5) is strict.

Suppose next that $a_2 = a_3$. From (2.2) and $(a_1 - a_3) \leqq 0$ we have

$(a_1 - a_3)(b_2 - b_3) = 0$ and hence $a_1 = a_3$ or $b_2 = b_3$. The first alternative yields $a_1 = a_2 = a_3$ and thus $L = a_1 \sum_{i=1}^{3} \alpha_i b_i \leqq a_1 b_1$ which is (2.3) with $\beta = 1$. If $b_2 = b_3$ then (2.3) holds with $\beta = \alpha_1$. This completes the proof of the lemma.

The proof of Theorem 1 is by induction on n. The first nontrivial case is $n = 3$. In general the convexity of f implies that

$$(\lambda_1 - \lambda_3)(\mu_2 - \mu_3) > (\lambda_2 - \lambda_3)(\mu_1 - \mu_3) \tag{2.6}$$

and (2.6) is strict if $\lambda_1 < \lambda_2 < \lambda_3$ and f is strictly convex. The inequality (2.1) follows from the lemma. If $n > 3$ we distinguish the two possibilities $\sigma_1 + \sigma_2 = 1$ and $\sigma_1 + \sigma_2 < 1$. In the first case

$$F(\sigma) = (\sigma_1\lambda_1 + \sigma_2\lambda_2)(\sigma_1\mu_1 + \sigma_2\mu_2) \,. \tag{2.7}$$

If $\mu_1 = \mu_n$ and hence $\mu_i = \mu_1 = \mu_n$, $i = 1, \cdots, n$, then $F(\sigma) \leqq \lambda_n\mu_n$ which is (2.1) with $\beta = 0$. If $\mu_1 > \mu_n$, and hence $\lambda_1 < \lambda_n$, obtain θ and ω in $[0, 1]$ so that $\lambda_2 = \theta\lambda_1 + (1 - \theta)\lambda_n$, $\mu_2 = \omega\mu_1 + (1 - \omega)\mu_n$ and set $\mu_2' = \theta\mu_1 + (1 - \theta)\mu_n$ to obtain

$$\mu_2' - \mu_2 = (\theta - \omega)(\mu_1 - \mu_n) \geqq 0 \,. \tag{2.8}$$

The convexity of f again implies that $\theta \geqq \omega$ with strictness in case f is strictly convex and $\lambda_2 > \lambda_n$. Hence

$$\begin{aligned} F(\sigma) &\leqq (\sigma_1\lambda_1 + (\theta\lambda_1 + (1 - \theta)\lambda_n)\sigma_2)(\sigma_1\mu_1 + \sigma_2\mu_2') \\ &= ((\sigma_1 + \theta\sigma_2)\lambda_1 + (1 - \theta)\sigma_2\lambda_n)((\sigma_1 + \theta\sigma_2)\mu_1 + (1 - \theta)\sigma_2\mu_n) \end{aligned}$$

which is (2.1) with $\beta = \sigma_1 + \theta\sigma_2$. We proceed to the case $\sigma_1 + \sigma_2 < 1$. Let $\lambda_3' = \sum_{i=3}^{n} \sigma_i\lambda_i/(1 - \sigma_1 - \sigma_2)$, $\mu_3'' = \sum_{i=3}^{n} \sigma_i\mu_i/(1 - \sigma_1 - \sigma_2)$ and observe that $\lambda_1 \leqq \lambda_2 \leqq \lambda_3'$, $\mu_1 \geqq \mu_2 \geqq \mu_3''$ and $F(\sigma) = (\sigma_1\lambda_1 + \sigma_2\lambda_2 + (1 - \sigma_1 - \sigma_2)\lambda_3')$ $(\sigma_1\mu_1 + \sigma_2\mu_2 + (1 - \sigma_1 - \sigma_2)\mu_3'')$. We next verify that (2.2) holds for the choices $\lambda_3' = a_3$, $\lambda_2 = a_2$, $\lambda_1 = a_1$, $\mu_1 = b_1$, $\mu_2 = b_2$, $\mu_3'' = b_3$:

$$\begin{aligned} &(\lambda_1 - \lambda_3)(\mu_2 - \mu_3'') - (\mu_1 - \mu_3'')(\lambda_2 - \lambda_3') \\ &\quad = \mu_2(\lambda_1 - \lambda_3') - \mu_1(\lambda_2 - \lambda_3') + \mu_3''(\lambda_2 - \lambda_1) \,; \end{aligned} \tag{2.9}$$

and

$$\mu_3'' = \sum_{i=3}^{n} f(\lambda_i)\sigma_i/(1 - \sigma_1 - \sigma_2) \geqq f\Big(\sum_{i=3}^{n} \lambda_i\sigma_i/(1 - \sigma_1 - \sigma_2)\Big) = f(\lambda_3') = \mu_3' \,.$$

Hence the expression in (2.9) is at least

$$\mu_2(\lambda_1 - \lambda_3') - \mu_1(\lambda_2 - \lambda_3') + \mu_3'(\lambda_2 - \lambda_1) \,. \tag{2.10}$$

If $\lambda_2 = \lambda_3'$ the expression (2.10) reduces to 0 and the expression in (2.9) is nonnegative. If $\lambda_2 < \lambda_3'$ then $\lambda_1 < \lambda_3'$ and (2.10) becomes $(\lambda_1 - \lambda_3')(\lambda_2 - \lambda_3')\{(\mu_2 - \mu_3')/(\lambda_2 - \lambda_3') - (\mu_1 - \mu_3')/(\lambda_1 - \lambda_3')\} \geqq 0$. Apply the lemma to obtain $\beta_1 \in [0, 1]$ for which

$$\begin{aligned}
&(\sigma_1\lambda_1 + \sigma_2\lambda_2 + (1 - \sigma_1 - \sigma_2)\lambda_3')(\sigma_1\mu_1 + \sigma_2\mu_2 + (1 - \sigma_1 - \sigma_2)\mu_3'') \\
&\quad \leqq (\beta_1\lambda_1 + (1 - \beta_1)\lambda_3')(\beta_1\mu_1 + (1 - \beta_1)\mu_3'') \\
&\quad = \left(\beta_1\lambda_1 + \sum_{i=3}^{n}(1 - \beta_1)\sigma_1\lambda_i/(1 - \sigma_1 - \sigma_2)\right) \\
&\qquad \left(\beta_1\mu_1 + \sum_{i=3}^{n}(1 - \beta_1)\sigma_i\mu_i/(1 - \sigma_1 - \sigma_2)\right).
\end{aligned}$$

This last expression is a product of convex combinations of λ's and μ's involving only $n - 1$ terms and satisfying the induction hypothesis. Hence there exists $\beta \in [0, 1]$ such that

$$\begin{aligned}
F(\sigma) &\leqq \left(\beta_1\lambda_1 + \sum_{i=3}^{n}(1 - \beta_1)\sigma_i\lambda_i/(1 - \sigma_1 - \sigma_2)\right) \\
&\left(\beta_1\mu_1 + \sum_{i=3}^{n}(1 - \beta_1)\sigma_i\mu_i/(1 - \sigma_1 - \sigma_2)\right) \leqq (\beta\lambda_1 + (1 - \beta)\lambda_n) \\
&(\beta\mu_1 + (1 - \beta)\mu_n).
\end{aligned}$$

This establishes (2.10).

The discussion of the strictness in (2.1) requires the use of (2.1) itself. Let k be the least integer for which both $\sigma_k > 0$ and $\lambda_1 < \lambda_k < \lambda_n$. Then

$$\begin{aligned}
F(\sigma) &= (\alpha_1\lambda_1 + \alpha_k\lambda_k + \alpha_{k+p}\lambda_{k+p} + \cdots + \alpha_n\lambda_n) \\
&\quad (\alpha_1\mu_1 + \alpha_k\mu_k + \alpha_{k+p}\mu_{k+p} + \cdots + \alpha_n\mu_n)
\end{aligned} \tag{2.11}$$

in which $\alpha_1 + \alpha_k + \alpha_{k+p} + \cdots + \alpha_n = 1$, $\alpha_j = \sigma_j$, $j = k + p, \cdots, n$, and $\lambda_k < \lambda_{k+p}$. Assume

$$\begin{aligned}
\alpha_1 + \alpha_k < 1, \text{ set } \lambda_{k+p}' &= \sum_{i=k+p}^{n}\sigma_i\lambda_i/(1 - \alpha_1 - \alpha_k),\ \mu_{k+p}'' \\
&= \sum_{i=k+p}^{n}\sigma_i\mu_i/(1 - \alpha_1 - \alpha_k)
\end{aligned}$$

and (2.11) becomes

$$\begin{aligned}
F(\sigma) &= (\alpha_1\lambda_1 + \alpha_k\lambda_k + (1 - \alpha_1 - \alpha_k)\lambda_{k+p}') \\
&\quad (\alpha_1\mu_1 + \alpha_k\mu_k + (1 - \alpha_1 - \alpha_k)\mu_{k+p}'').
\end{aligned} \tag{2.12}$$

Clearly $\lambda_1 < \lambda_k < \lambda_{k+p}'$ and we compute that

$$\begin{aligned}
&(\lambda_1 - \lambda_{k+p}')(\mu_k - \mu_{k+p}') - (\mu_1 - \mu_{k+p}')(\lambda_k - \lambda_{k+p}') \\
&\quad = \mu_k(\lambda_1 - \lambda_{k+p}') - \mu_1(\lambda_k - \lambda_{k+p}') + \mu_{k+p}''(\lambda_k - \lambda_1);
\end{aligned} \tag{2.13}$$

$$\mu_{k+p}'' \geqq f(\lambda_{k+p}') = \mu_{k+p}'. \tag{2.14}$$

It follows that the expression in (2.13) is at least

$$(\lambda_1 - \lambda'_{k+p})(\lambda_k - \lambda'_{k+p})\{(\mu_k - \mu'_{k+p})/(\lambda_k - \lambda'_{k+p}) - (\mu_1 - \mu'_{k+p})/(\lambda_1 - \lambda'_{k+p})\}$$

and in case f is strictly convex this whole expression is positive. The inequality (2.2) holds strictly with $\lambda_1 = a_1$, $\lambda_k = a_2$, $\lambda'_{k+p} = a_3$, $\mu_1 = b_1$, $\mu_k = b_k$, $\mu'_{k+p} = b_3$ and the strict form of the lemma together with (2.12) implies that there exists $\beta_1 \in [0, 1]$ such that

$$F(\sigma) < \left(\beta_1\lambda_1 + \sum_{i=k+p}^{n}(1 - \beta_1)\sigma_i\lambda_i/(1 - \alpha_1 - \alpha_k)\right) \left(\beta_1\mu_1 + \sum_{i=k+p}^{n}(1 - \beta_1)\sigma_i\mu_i/(1 - \alpha_1 - \alpha_k)\right) . \tag{2.15}$$

Now apply (2.1) to the right side of (2.15) to obtain a $\beta \in [0, 1]$ for which $F(\sigma) < (\beta\lambda_1 + (1 - \beta)\lambda_n)(\beta\mu_1 + (1 - \beta)\mu_n)$.

Assume now that $\alpha_1 + \alpha_k = 1$ and then $F(\sigma)$ becomes $(\alpha_1\lambda_1 + (1 - \alpha_1)\lambda_k)(\alpha_1\mu_1 + (1 - \alpha_1)\mu_k)$. Choose θ and ω in $[0, 1]$ so that $\lambda_k = \theta\lambda_1 + (1 - \theta)\lambda_n$, $\mu_k = \omega\mu_1 + (1 - \omega)\mu_n$, set $\mu''_k = \theta\mu_1 + (1 - \theta)\mu_n$ and note that $\mu''_k - \mu_k = (\theta - \omega)(\mu_1 - \mu_n)$. Then since f is monotone decreasing and strictly convex, $\theta - \omega$ and $\mu_1 - \mu_n$ are both positive. It follows that

$$(\alpha_1\lambda_1 + (1 - \alpha_1)\lambda_k)(\alpha_1\mu_1 + (1 - \alpha_1)\mu_k) < ((\alpha_1 + \theta(1 - \alpha_1))\lambda_1 + (1 - \theta)(1 - \alpha_1)\lambda_n)((\alpha_1 + \theta(1 - \alpha_1))\mu_1 + (1 - \theta)(1 - \alpha_1)\mu_n) .$$

If the quadratic polynomial in β on the right in (2.1) is maximized in $[0, 1]$ we immediately obtain our main result.

THEOREM 2. *If*

$$\gamma \geqq m \text{ and } \lambda_1 < \lambda_n \text{ and } \mu_1 > \mu_n \tag{2.16}$$

then

$$M = (\lambda_n\mu_1 - \lambda_1\mu_n)/4(\lambda_n - \lambda_1)(\mu_1 - \mu_n) . \tag{2.17}$$

If

$$\gamma \leqq m \text{ or } \lambda_1 = \lambda_n \text{ or } \mu_1 = \mu_n \tag{2.18}$$

then

$$M = m . \tag{2.19}$$

Let f be strictly convex and suppose that

$$\lambda_1 = \cdots = \lambda_p < \lambda_{p+1} \leqq \cdots \leqq \lambda_{n-q} < \lambda_{n-q+1} = \cdots = \lambda_n .$$

Then $F(\sigma) = M$, $\sigma \in S^{n-1}$, if and only if σ has the form

$$\sigma = (\sigma_1, \cdots \sigma_p, 0, \cdots, 0, \sigma_{n-q+1}, \cdots, \sigma_n) ,$$

$\sum_{j=1}^{p} \sigma_j = \beta_0$, $\sum_{j=n-q+1}^{n} \sigma_j = 1 - \beta_0$, *where*

$$\beta_0 = \begin{cases} (\gamma - \lambda_n \mu_n)/(\lambda_n - \lambda_1)(\mu_1 - \mu_n) \text{ if (2.16) holds,} \\ 0 \text{ or } 1 \text{ if (2.18) holds.} \end{cases} \tag{2.20}$$

We remark that if $\gamma = m$ then the expression on the right in (2.17) reduces to m.

3. Applications.

As customary $f(A)$ will designate the linear transformation defined for any $x \in U$ by

$$f(A)x = \sum_{i=1}^{n} \mu_i (x, u_i) u_i, \; (\mu_i = f(\lambda_i)) \; . \tag{3.1}$$

On the unit sphere $\|x\| = 1$ define the real valued function

$$\varphi(x) = (Ax, x)(f(A)x, x) \; . \tag{3.2}$$

We compute directly from (3.1) that

$$\varphi(x) = \sum_{i=1}^{n} \lambda_i \,|(x, u_i)|^2 \sum_{i=1}^{n} \mu_i \,|(x, u_i)|^2 \tag{3.3}$$

and by setting $\sigma_i = |(x, u_i)|^2$, $i = 1, \cdots, n$, we have $\sigma = (\sigma_1, \cdots, \sigma_n) \in S^{n-1}$ and

$$\varphi(x) = F(\sigma) \; . \tag{3.4}$$

Thus by direct application of Theorem 2 we have

THEOREM 3. *Then maximum value of $\varphi(x)$ for x on the unit sphere $\|x\| = 1$ is the number M in the statement of Theorem 2. Moreover $\varphi(x_0) = M$ can always be achieved with a unit vector x_0 in the subspace spanned by those eigenvectors of A corresponding to λ_1 and λ_n. If f is strictly convex and $\varphi(x_0) = M$ then x_0 must lie in the sum of the null spaces of $A - \lambda_1 I$ and $A - \lambda_n I$. In particular, if λ_1 and λ_n are simple eigenvalues of A, f is strictly convex and $\varphi(x_0) = M$ then x_0 must lie in the two dimensional subspace spanned by u_1 and u_n.*

In Theorem 3 take $f(t) = t^{-p}$, $p > 0$. Let $\theta = \lambda_1/\lambda_n$ denote the condition number of A. Assume that $\theta < 1$ (otherwise $\lambda_1 = \lambda_n$ and A is a multiple of the identity). There are two cases to consider: $p > 1$; $p \leqq 1$. In case $p > 1$, $m = \lambda_1^{1-p}$ and the condition (2.16), $\gamma \geqq m$, becomes

$$g(\theta) = \theta^{p+1} - 2\theta + 1 \geqq 0 \; . \tag{3.5}$$

We note that g is convex, $g(1) = 0$, $g'(\theta) = 0$ for $\theta = (2/(p+1))^{1/p}$, and hence g has precisely one root in $(0, 1)$, call it θ_p. It is easy to see that $\theta_p > 1/2$ for all $p > 1$. In general, if $0 < \theta \leqq \theta_p$ then Theorem 2 yields

$$M = \lambda_1^{1-p}(\theta^{p+1} - 1)^2/4\theta(\theta - 1)(\theta^p - 1) \; ; \tag{3.6}$$

and if $1 \geqq \theta > \theta_p$ then

$$M = \lambda_1^{1-p} \; . \tag{3.7}$$

In case $p \leqq 1$, $m = \lambda_n^{1-p}$ and the condition (2.16), $\gamma \geqq m$, becomes $g(\eta) \geqq 0$ where $\eta = \theta^{-1}$. But $g(\eta) \geqq 0$ for $\eta \geqq 1$ and $\eta = \theta^{-1} \geqq 1$ so the upper bound for $F(\sigma)$ is M given in (3.6).

Assume now that λ_1 and λ_n are both simple eigenvalues of A and we examine the structure of the vector x_0 that maximizes $\varphi(x) = (Ax, x)(A^{-p}x, x)$ on the unit sphere $||\, x \,|| = 1$. By Theorem 3 the maximum value of $\varphi(x) = F(\sigma)$ can only occur for $\sigma_2 = \cdots = \sigma_{n-1} = 0$. Moreover by (2.20) $F(\sigma) = M$ for the unique values

$$\left.\begin{aligned} &(3.8) \quad \sigma_n = \sigma_n(\theta) = g(\theta)/2(1 - \theta)(1 - \theta^p) \\ &(3.9) \quad \sigma_1 = \sigma_1(\theta) = \sigma_n(\theta^{-1}) \end{aligned}\right\} \text{if } g(\theta) \geqq 0 \text{ or } p = 1 \; ;$$

and

$$\sigma_1 = 1, \sigma_n = 0 \text{ if } g(\theta) < 0 \text{ and } p > 1 \; . \tag{3.10}$$

Summing up these results we have

THEOREM 4. *Let θ designate the condition number of A, $\theta = \lambda_1/\lambda_n$. If either $0 < p \leqq 1$, or $p > 1$ and $0 \leqq \theta \leqq \theta_p$, then for $||\, x \,|| = 1$*

$$(Ax, x)(A^{-p}x, x) \leqq \lambda_1^{1-p}(\theta^{p+1} - 1)^2/4\theta(\theta - 1)(\theta^p - 1) \; . \tag{3.11}$$

If $p > 1$ and $\theta_p < \theta$ then for $||\, x \,|| = 1$

$$(Ax, x)(A^{-p}x, x) \leqq \lambda_1^{1-p} \; . \tag{3.12}$$

If λ_1 and λ_n are simple eigenvalues of A then the upper bound in (3.11) *is only achieved for unit vectors of the form*

$$x_0 = \sqrt{\sigma_n(\theta^{-1})}\, e^{i\omega_1}u_1 + \sqrt{\sigma_n(\theta)}\, e^{i\omega_2}u_n \; , \tag{3.13}$$

ω_1, ω_2 real. The upper bound in (3.12) *is achieved only for unit vectors of the form*

$$x_0 = e^{i\omega}u_1 \; .$$

In case $p = 1$ we have the Kantorovich inequality. In this case (3.11) becomes (for $||\, x \,|| = 1$)

(3.14) $$(Ax, x)(A^{-1}x, x) \leqq (\sqrt{\theta} + \sqrt{\theta^{-1}})^2/4 .$$

If λ_1 *and* λ_n *are simple eigenvalues then the inequality* (3.14) *is strict unless*

(3.15) $$x = x_0 = (e^{i\omega_1}u_1 + e^{i\omega_2}u_n)/\sqrt{2}, \omega_1, \omega_2 \text{ real} .$$

4. Determinants and permanents. In this section we specialize by taking U to be the unitary space of n-tuples with inner product $(x, y) = \sum_{i=1}^{n} x_i\bar{y}_i$ and A to be an n-square hermitian positive semi-definite matrix. If $1 \leqq k \leqq n$ then $C_k(A)$ will denote the kth compound of A and if $x_1, \cdots, x_k$ are vectors in U then $x_1 \wedge \cdots \wedge x_k$ is the Grassmann product of these vectors, sometimes called a pure vector of grade k [6, p. 16]. The eigenvalues of $C_k(A)$ are all $\binom{n}{k}$ numbers $\lambda_{i_1} \cdots \lambda_{i_k}$, with corresponding eigenvectors $u_{i_1} \wedge \cdots \wedge u_{i_k}$, $1 \leqq i_1 < \cdots < i_k \leqq n$. The smallest and largest of these eigenvalues are $\prod_{j=1}^{k} \lambda_j$ and $\prod_{j=1}^{k} \lambda_{n-j+1}$ respectively. It has been noted in [2] and [5] that the Kantorovich inequality applied to $C_k(A)$ yields

(4.1) $$\det A[i_1, \cdots, i_k] \det A^{-1}[i_1, \cdots, i_k] \leqq (\sqrt{\varDelta} + \sqrt{\varDelta^{-1}})^2/4$$

where $\varDelta = \prod_{j=1}^{k} \lambda_j \lambda_{n-j+1}^{-1}$ *and* $A[i_1, \cdots, i_k]$ *is the principal submatrix of* A *lying in rows and columns numbered* $i_1, \cdots, i_k$.

We prove

THEOREM 5. *If* $1 \leqq k < n - 1$ *and* $\lambda_1, \cdots, \lambda_k$ *together with* $\lambda_n, \cdots, \lambda_{n-k+1}$ *are simple eigenvalues of* A *then the inequality* (4.1) *is always strict.*

Proof. The number $\det A[i_1, \cdots, i_k] \det A^{-1}[i_1, \cdots, i_k]$ is a value of the product of quadratic forms associated with $C_k(A)$ and $C_k(A^{-1})$,

(4.2) $$\begin{gathered}(C_k(A)x_1 \wedge \cdots \wedge x_k, x_1 \wedge \cdots \wedge x_k) \\ (C_k(A^{-1})x_1 \wedge \cdots \wedge x_k, x_1 \wedge \cdots \wedge x_k) ,\end{gathered}$$

and according to (3.15), (4.1) will be strict unless

(4.3) $$x_1 \wedge \cdots \wedge x_k = \frac{1}{\sqrt{2}}(e^{i\omega_1}u_1 \wedge \cdots \wedge u_k + e^{i\omega_2}u_n \wedge \cdots \wedge u_{n-k+1}) .$$

Let $p = \min\{k, n - k\}$, $q = \max\{k + 1, n - k + 1\}$ and compute successively the Grassmann products of both sides of (4.3) with $u_1, \cdots, u_p$ and $u_n, \cdots, u_q$. We obtain

(4.4) $$x_1 \wedge \cdots \wedge x_k \wedge u_j = \frac{e^{i\omega_2}}{\sqrt{2}}(u_n \wedge \cdots \wedge u_{n-k+1} \wedge u_j), j = 1, \cdots, p ,$$

and

$$(4.5)\quad x_1 \wedge \cdots \wedge x_k \wedge u_j = \frac{e^{i\omega_2}}{\sqrt{2}}(u_1 \wedge \cdots \wedge u_k \wedge u_j),\ j = q, \cdots, n\ .$$

Since $u_1, \cdots, u_n$ are linearly independent it follows that the right sides of (4.4) and (4.5) are not 0. Thus

$$(4.6)\qquad < x_1, \cdots, x_k, u_j > = < u_1, \cdots, u_k, u_j >, j = 1, \cdots, p\ ,$$

and

$$(4.7)\qquad < x_1, \cdots, x_k, u_j > = < u_1, \cdots, u_k, u_j >, j = q, \cdots, n\ ,$$

where $< x_1, \cdots, x_k, u_j >$ denotes the subspace spanned by the vectors inside the brackets. Intersect the p subspaces on the left in (4.6) and observe that $< x_1, \cdots, x_k >$ is a subspace of the intersection. Similarly $< x_1, \cdots, x_k >$ is a subspace of the intersection of the $n - q + 1$ spaces on the left in (4.7). On the other hand

$$\bigcap_{j=1}^{p} < u_n, \cdots, u_{n-k+1}, u_j > = < u_n, \cdots, u_{n-k+1} >$$

and

$$\bigcap_{j=q}^{n} < u_1, \cdots, u_k, u_j > = < u_1, \cdots, u_k >\ .$$

Hence

$$(4.8)\quad \begin{aligned} &\dim \{< u_1, \cdots, u_k > \cap < u_n, \cdots, u_{n-k+1} >\} \\ &\quad = \dim \left\{\bigcap_{j=1}^{p} < x_1, \cdots, x_k, u_j > \cap \bigcap_{j=q}^{n} < x_1, \cdots, x_k, u_j >\right\} > k\ . \end{aligned}$$

The subspace $< u_1, \cdots, u_k > \cap < u_n, \cdots, u_{n-k+1} >$ is nonempty if and only if $n - k + 1 \leqq k$ in which case its dimension is $2k - n$. But the inequality $2k - n \geqq k$ implies that $k \geqq n$, a contradiction. Thus (4.3) cannot hold and (4.1) is strict.

We remark that in case $k = n - 1$ then $p = 1, q = n, x_1 \wedge \cdots \wedge x_k \wedge u_1 = u_n \wedge \cdots \wedge u_2 \wedge u_1, x_1 \wedge \cdots \wedge x_k \wedge u_n = u_1 \wedge \cdots \wedge u_{n-1} \wedge u_n$ and the above argument fails. In fact, it is not difficult to construct examples for which (4.1) is equality.

Once again, if $1 \leqq k \leqq n$ then $P_k(A)$ will denote the kth induced power matrix of A and if $x_1, \cdots, x_k$ are vectors in U then $x_1 \cdots x_k$ will denote the symmetric or dot product of these vectors [3, p. 49]. The eigenvalues of $P_k(A)$ are all $\binom{n+k-1}{k}$ homogeneous products $\lambda_{i_1} \cdots \lambda_{i_k}$ with corresponding eigenvectors $u_{i_1} \cdots u_{i_k}, 1 \leqq i_1 \leqq \cdots \leqq i_k \leqq n$. Suppose $x_1, \cdots, x_n$ are orthonormal vectors and the multiplicities of the distinct integers in the sequence $i_1 \leqq \cdots \leqq i_k$ are respectively $m_1, \cdots, m_p$. Let $\mu = \mu(i_1, \cdots, i_k) = m_1! \cdots m_p!$. Then the square of the length of the symmetric product $x_{i_1} \cdots x_{i_k}$ is $\mu(i_1, \cdots, i_k)$ [3, p. 50]. Applying the Kantorovich inequality to $P_k(A)$ yields

$$(P_k(A)x_i \cdots x_{i_k}, x_{i_1} \cdots x_{i_k})(P_k(A^{-1})x_{i_1} \cdots x_{i_k}, x_{i_1} \cdots x_{i_k}) \\ \leqq \mu^2(\sqrt{\delta} + \sqrt{\delta^{-1}})^2/4, 1 \leqq i_1 \leqq \cdots \leqq i_k \leqq n \,, \tag{4.9}$$

where $\delta = (\lambda_1 \lambda_n^{-1})^k$, *and* $x_1, \cdots, x_n$ *is an orthonormal basis of* U. In particular if we let $x_i = e_i$, the unit vector with 1 in the ith position, 0 elsewhere, then (4.9) becomes

$$\operatorname{per} A[i_1, \cdots, i_k] \operatorname{per} A^{-1}[i_1, \cdots, i_k] \leqq \mu^2(\sqrt{\delta} + \sqrt{\delta^{-1}})^2/4 \,, \tag{4.10}$$

where $A[i_1, \cdots, i_k]$ *is the k-square matrix whose* (s, t) *entry is* $a_{i_s i_t}$, $s, t = 1, \cdots, k$.

THEOREM 6. *If* λ_1 *and* λ_n *are simple eigenvalues of* A *and there are at least three distinct integers in the sequence* $i_1 \leqq \cdots \leqq i_k$ *then the inequality* (4.10) *is strict.*

Proof. According to (3.15), (4.10) will be strict unless

$$e_{i_1} \cdots e_{i_k} = \frac{e^{i\omega_1}}{\sqrt{2k!}} u_1 \cdots u_1 + \frac{e^{i\omega_2}}{\sqrt{2k!}} u_n \cdots u_n \,. \tag{4.11}$$

Let y be an arbitrary vector and compute the inner product of both sides of (4.11) with $y \cdots y$ to obtain

$$\prod_{j=1}^{k} (e_{i_j}, y) = \frac{e^{i\omega_1}}{\sqrt{2k!}} (u_1, y)^k + \frac{e^{i\omega_2}}{\sqrt{2k!}} (u_n, y)^k \,. \tag{4.12}$$

Set

$$v_1 = \left(\frac{e^{i\omega_1}}{\sqrt{2k!}}\right)^{1/k} u_1, \; v_2 = \left(\frac{e^{i\omega_2}}{\sqrt{2k!}}\right)^{1/k} u_n \,,$$

and write $e_{i_j} = \alpha_j v_1 + w_j$, $w_j \in \langle v_1 \rangle^{\perp}$, $j = 1, \cdots, k$. Then for y any vector in $\langle v_1 \rangle^{\perp}$, (4.12) becomes

$$\prod_{j=1}^{k} (e_{i_j}, y) = \prod_{j=1}^{k} (w_j, y) = (v_2, y)^k \,, \tag{4.13}$$

in which w_j, v_2, y are in $\langle v_1 \rangle^{\perp}$, $j = 1, \cdots, k$. But then from [3, Theorem 3] we conclude that $w_j = \beta_j v_2$, $j = 1, \cdots, k$, for appropriate scalars $\beta_1, \cdots, \beta_k$ and hence $e_{i_j} \in \langle v_1, v_2 \rangle$, $j = 1, \cdots, k$. Since there are at least three linearly independent e_{i_j}, (4.11) must fail and hence (4.10) is strict.

REFERENCES

1. L. V. Kantorovich and V. I. Krylov, *Approximate methods of higher analysis*, New York, Interscience (1958).

2. Marvin Marcus, and N. A. Khan *Some generalizations of Kantorovich's inequality*, Portugaliae Math. **20**, 1, (1961), 33-38.

3. Marvin Marcus and Morris Newman. *Inequalities for the permanent function*, Ann of Math., **75**, 1, (1962), 47-62.

4. Morris Newman. *Kantorovich's inequality*, J. Research Nat. Bur. Standards, **64** (B), (1960), 33-34.

5. Andreas H. Schopf,. *On the Kantorovich inequality*, Numerische Math., **2** (1960), 344-346.

6. J. H. M. Wedderburn,. *Lectures on matrices*, Amer, Math. Soc. Coll. Publ., **17** (1934).

UNIVERSITY OF CALIFORNIA, SANTA BARBARA
AND
UNIVERSITY OF BRITISH COLUMBIA, CANADA

PACIFIC JOURNAL OF MATHEMATICS

Mathematical papers intended for publication in the *Pacific Journal of Mathematics* should be typewritten (double spaced), and the author should keep a complete copy. Manuscripts may be sent to any one of the four editors. All other communications to the editors should be addressed to the managing editor, L. J. Paige at the University of California, Los Angeles 24, California.

50 reprints per author of each article are furnished free of charge; additional copies may be obtained at cost in multiples of 50.

The *Pacific Journal of Mathematics* is published quarterly, in March, June, September, and December. Effective with Volume 13 the price per volume (4 numbers) is \$18.00; single issues, \$5.00. Special price for current issues to individual faculty members of supporting institutions and to individual members of the American Mathematical Society: \$8.00 per volume; single issues \$2.50. Back numbers are available.

Subscriptions, orders for back numbers, and changes of address should be sent to Pacific Journal of Mathematics, 103 Highland Boulevard, Berkeley 8, California.

Printed at Kokusai Bunken Insatsusha (International Academic Printing Co., Ltd.), No. 6, 2-chome, Fujimi-cho, Chiyoda-ku, Tokyo, Japan.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS, A NON-PROFIT CORPORATION

The Supporting Institutions listed above contribute to the cost of publication of this Journal, but they are not owners or publishers and have no responsibility for its content or policies.

Pacific Journal of Mathematics

Vol. 13, No. 4 June, 1963