The Annals of Mathematical Statistics
1968, Vol. 39, No. 3, 859–863

MINIMIZATION OF EIGENVALUES OF A MATRIX AND OPTIMALITY OF PRINCIPAL COMPONENTS

BY MASASHI OKAMOTO[1] AND MITSUYO KANAZAWA

Osaka University

1. Introduction. Let $x' = (x_1, x_2, \ldots, x_p)$ be a random vector with mean vector $E(x) = 0$ and variance matrix $E(xx') = \Sigma$. Let $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_p \geqq 0$ be the eigenvalues of Σ in order of decreasing magnitude, and $v_1, v_2, \cdots, v_p$ be the corresponding orthonormal eigenvectors.

The principal components of x, namely $v_1'x, v_2'x, \cdots, v_p'x$ were introduced by Hotelling [3], and since then characterized by various optimal properties. Almost all of these optimal properties, however, are stated in terms of linear functions of $x_1, x_2, \cdots, x_p$. For example, Rao [4] characterizes the first $k (\leqq p)$ principal components as a linear form $y = T'x$ with a $p \times k$ matrix T which minimizes the trace or the Euclidean norm of the residual variance matrix of x after subtracting its best linear predictor based on y. The unique exception is Darroch [2] who deals with the optimality within the class of all random variables with at most k dimensions.

The purpose of this paper is to characterize the first k principal components by a more general optimal property containing those due to Rao or Darroch as special cases. Lemma 3 in Section 2 dealing with simultaneous minimization of the eigenvalues of a non-negative definite matrix is of an algebraic character, and may be interesting by itself.

2. Notation and lemmas. Let $\mathfrak{A} = \mathfrak{A}_p$ be the set of all real non-negative definite matrices of order p. A partial order in the set $\mathfrak{A}$ is defined as usual; $A \geqq B$ if and only if $A - B \varepsilon \mathfrak{A}$. For any $A \varepsilon \mathfrak{A}$ let $\lambda_1(A) \geqq \lambda_2(A) \geqq \cdots \geqq \lambda_p(A)$ be the eigenvalues of A in order of decreasing magnitude. The following two Lemmas will be stated without proof.

LEMMA 1. *A necessary and sufficient condition for a real-valued function $f(A)$ defined on $\mathfrak{A}$ to be*

(i) *strictly increasing, that is, $f(A) \geqq f(B)$ if $A \geqq B$, and $f(A) > f(B)$ if moreover $A \neq B$, and*

(ii) *invariant under orthogonal transformation, that is, $f(P'AP) = f(A)$ for any orthogonal matrix P,*

is that $f(A)$ is identical to some function $g(\lambda_1(A), \cdots, \lambda_p(A))$ of the eigenvalues of A which is strictly increasing in each argument.

It is noted that the trace as well as the Euclidean norm of a matrix enjoys this property of a function f.

Now we denote by $M_k(A)$ $(k = 1, 2, \cdots, p)$ the linear subspace spanned by

Received April 28, 1967.

[1] Now visiting professor at the Statistical Laboratory and Department of Statistics, Iowa State University.