The Annals of Statistics
1973, Vol. 1, No. 4, 763-765

DISTINCTNESS OF THE EIGENVALUES OF A QUADRATIC FORM IN A MULTIVARIATE SAMPLE

BY MASASHI OKAMOTO

Osaka University

This paper shows that a quadratic form in a multivariate sample has a certain rank and its nonzero eigenvalues are distinct with probability one under the assumption that the matrix defining the quadratic form satisfies a certain rank condition and that the underlying distribution of the sample is absolutely continuous with respect to Lebesgue measure.

1. Introduction. Let $\mathbf{X}_1, \cdots, \mathbf{X}_n$ be a random sample from a p-variate non-singular normal distribution and let $\bar{\mathbf{X}}$ be the sample mean. It was proved by Dykstra (1970) that the random matrix

$$\mathbf{S} = \sum_{\alpha=1}^{n} (\mathbf{X}_\alpha - \bar{\mathbf{X}})(\mathbf{X}_\alpha - \bar{\mathbf{X}})'$$

is positive definite with probability one if and only if $n > p$. This result was extended by Das Gupta (1971) and Eaton and Perlman (1973) so as to remove the assumption of normality and to generalize the matrix $\mathbf{S}$ into a quadratic form in a multivariate sample.

The purpose of the present paper is to proceed further by considering the rank of the quadratic form and the distinctness of its nonzero eigenvalues. It is noted, however, that in the important special case when the $\mathbf{X}_i$'s are independently and identically distributed our assumption of the absolute continuity of the joint distribution of the $\mathbf{X}_i$'s is stronger than the assumption imposed by Das Gupta as well as by Eaton and Perlman that any hyperplane has probability zero. Thus the question of whether our theorem holds still under their assumption is left open.

2. The result. We begin with a simple lemma which might be well-known but could not be found in literature.

LEMMA. *If $f(x_1, \cdots, x_m)$ is a polynomial in real variables $x_1, \cdots, x_m$, which is not identically zero, then the subset*

$$N_m = \{(x_1, \cdots, x_m) \mid f(x_1, \cdots, x_m) = 0\}$$

of the Euclidean m-space R^m has Lebesgue measure zero.

PROOF. We use mathematical induction. Denote by μ_m the Lebesgue measure over R^m and by $C = C(x_1, \cdots, x_{m-1})$ the cross-section of the set N_m at the point $(x_1, \cdots, x_{m-1})$, or

$$C = \{x_m \mid f(x_1, \cdots, x_m) = 0\}.$$

Received October 1972; revised December 1972.

AMS 1970 *subject classifications.* Primary 62H10; Secondary 62E15.

Key words and phrases. Distinctness of eigenvalues, quadratic form in a multivariate sample.

Let d be the degree of the polynomial f with respect to the variable x_m and let $f_i(x_1, \cdots, x_{m-1})$ for $i = 0, 1, \cdots, d$ be the coefficients of $x_m{}^i$. Put

$$N_{m-1} = \{(x_1, \cdots, x_{m-1}) \mid f_i(x_1, \cdots, x_{m-1}) = 0 \quad \text{for} \quad i = 0, 1, \cdots, d\} .$$

Then it holds that

$$\begin{aligned}\mu_m(N_m) &= \int_{R^{m-1}} \mu_1(C(x_1, \cdots, x_{m-1}))\, d\mu_{m-1}(x_1, \cdots, x_{m-1}) \\ &= \int_{N_{m-1}} \mu_1(C)\, d\mu_{m-1} + \int_{N^c_{m-1}} \mu_1(C)\, d\mu_{m-1} .\end{aligned}$$

The first term of the right-hand side vanishes, since $\mu_{m-1}(N_{m-1}) = 0$ by the assumption of induction. The second term vanishes, since for $(x_1, \cdots, x_{m-1}) \notin N_{m-1}$ the set C is finite and hence $\mu_1(C) = 0$. Thus $\mu_m(N_m) = 0$, completing the proof.

THEOREM. *Let* $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)$, *a* $p \times n$ *matrix, where the joint distribution of the p-variate random vectors* $\mathbf{X}_1, \cdots, \mathbf{X}_n$ *is absolutely continuous (with respect to np-dimensional Lebesgue measure) and let* $\mathbf{A}$ *be a real symmetric* $n \times n$ *matrix of rank r. Then, for the random matrix*

$$\mathbf{S} = \mathbf{XAX}' ,$$

the following statement holds with probability one: rank $\mathbf{S} = \min(p, r)$ *and the nonzero eigenvalues of* $\mathbf{S}$ *are distinct.*

PROOF. The eigenvalues of the matrix $\mathbf{S}$ are the roots of the equation $g(\lambda) = 0$, where

$$g(\lambda) = |\mathbf{S} - \lambda \mathbf{I}_p| .$$

Since $g(\lambda)$ is a polynomial of degree p in λ, it can be written as

$$g(\lambda) = \sum_{i=0}^{p} a_i(\mathbf{S}) \lambda^{p-i} ,$$

each coefficient $a_i(\mathbf{S})$ being a polynomial in the elements of $\mathbf{S}$. Since

$$\operatorname{rank} \mathbf{S} \leqq \min(\operatorname{rank} \mathbf{X}, \operatorname{rank} \mathbf{A}) \leqq \min(p, r) \qquad (= m, \text{ say}) ,$$

the equation $g(\lambda) = 0$ has at least $p - m$ roots of value zero, which implies that

$$a_i(\mathbf{S}) = 0 \qquad \text{for any} \quad i > m .$$

Thus it holds that

$$\operatorname{rank} \mathbf{S} = m \iff a_m(\mathbf{S}) \neq 0 . \tag{1}$$

Denote by $D(\mathbf{S})$ the discriminant of the polynomial

$$\sum_{i=0}^{m} a_i(\mathbf{S}) \lambda^{m-i} .$$

By a well-known theorem in algebra (see, e.g., van der Waerden (1949), page 82) $D(\mathbf{S})$ is a polynomial in the elements of $\mathbf{S}$ and

$$\text{the nonzero eigenvalues of } \mathbf{S} \text{ are distinct} \iff D(\mathbf{S}) \neq 0 . \tag{2}$$

Define

$$h(\mathbf{S}) = a_m(\mathbf{S}) D(\mathbf{S}) .$$

Then, in view of (1) and (2) it is sufficient to prove that

$$h(\mathbf{S}) \neq 0 \qquad \text{with probability one.}$$

By definition every element of $\mathbf{S}$ is a polynomial in the elements of the matrix $\mathbf{X}$; and hence $h(\mathbf{S}) = f(\mathbf{X})$, a polynomial in the elements of $\mathbf{X}$. Using the Lemma, we have only to show that the function $f(\mathbf{X})$ is not identically zero.

Write $\mathbf{A} = \mathbf{GDG}'$, where $\mathbf{G}$ is a column-orthogonal $n \times r$ matrix ($\mathbf{G}'\mathbf{G} = \mathbf{I}_r$) and $\mathbf{D}$ is a nonsingular diagonal $r \times r$ matrix. Now we distinguish the following two cases: (i) $r \geqq p$ and (ii) $r < p$.

Case (i). Partition

$$\mathbf{G} = \begin{array}{c}(\mathbf{G}_1 : \mathbf{G}_2) \\ \;\; p \quad r-p\end{array} \quad \text{and} \quad \mathbf{D} = \begin{array}{cl} \begin{pmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{pmatrix} & \begin{array}{l} p \\ r-p \end{array} \\ \;\; p \quad r-p & \end{array}$$

and let $\mathbf{X} = \mathbf{D}_0\mathbf{G}_1'$, where $\mathbf{D}_0$ is a nonsingular diagonal $p \times p$ matrix to be determined later. Then

$$\mathbf{S} = \mathbf{D}_0\mathbf{G}_1'\mathbf{GDG}'\mathbf{G}_1\mathbf{D}_0 = \mathbf{D}_0{}^2\mathbf{D}_1 ,$$

which is a nonsingular matrix. If we choose any $\mathbf{D}_0$ for which the diagonal elements of $\mathbf{D}_0{}^2\mathbf{D}_1$ are distinct, then $f(\mathbf{X}) = h(\mathbf{S}) \neq 0$.

Case (ii). Let

$$\mathbf{X} = \begin{array}{r} r \\ p-r \end{array} \begin{pmatrix} \mathbf{D}_0\mathbf{G}' \\ \mathbf{0} \end{pmatrix},$$

where $\mathbf{D}_0$ is a nonsingular diagonal $r \times r$ matrix. Then

$$\mathbf{S} = \begin{pmatrix} \mathbf{D}_0\mathbf{G}' \\ \mathbf{0} \end{pmatrix} \mathbf{GDG}'(\mathbf{GD}_0 : \mathbf{0}) = \begin{pmatrix} \mathbf{D}_0{}^2\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix};$$

and hence rank $\mathbf{S} = r$ and the nonzero eigenvalues of $\mathbf{S}$ are distinct, provided $\mathbf{D}_0$ is chosen so that the diagonal elements of $\mathbf{D}_0{}^2\mathbf{D}$ are distinct. This means again $f(\mathbf{X}) \neq 0$. □

REFERENCES

[1] DAS GUPTA, S. (1971). Nonsingularity of the sample covariance matrix. *Sankhyā Ser. A* **33** 475–478.

[2] DYKSTRA, R. L. (1970). Establishing the positive definiteness of the sample covariance matrix. *Ann. Math. Statist.* **41** 2153–2154.

[3] EATON, M. L. and PERLMAN, M. D. (1973). The non-singularity of generalized sample covariance matrices. *Ann. Statist.* 710–717.

[4] VAN DER WAERDEN, B. L. (1949). *Modern Algebra*, **1**. Ungar, New York.

FACULTY OF ENGINEERING SCIENCE
OSAKA UNIVERSITY
TOYONAKA, OSAKA, JAPAN