A GENERALIZED MULTIVARIATE ANALOGUE OF THE ONE SIDED TEST

By

Akio KUDÔ and Jae-Rong CHOI

(Received April 18, 1975)

Contents

§ 1. Introduction	303
§ 2. A theorem on the closest point	305
§ 3. An algorithm for calculating $\bar{\chi}^2$ and the closest point	308
§ 4. Distribution of the likelihood ratio statistic $\bar{\chi}^2$	314
§ 5. An algorithm for computing the $\bar{\chi}^2$ -statistic and its significance level in case when rank $(A)=3$	318
§ 6. A numerical example in case when rank $(A)=3$	324
§ 7. Acknowledgements	327

§ 1. Introduction

Given a k variate normal distribution with the known variance matrix A , the problem of testing the hypothesis that all the components of the mean vector are zero against the alternative that all the components are positive was first treated by one of the authors [4] and independently by Nuesch [6]. One of the applications of this type of test is to reduce the test of homogeneity of means under ordered alternative (cf. Bartholomew [3]).

In this paper we generalize the previous result to the case when variance matrix is singular. Consider the test of homogeneity of means under the alternative hypothesis where the number of inequalities about the means exceeds the number of means involved, the problem can be again reduced to the problem of multivariate analogue of the one-sided test, but the variance matrix turns out to be a singular matrix. This is the first motivation of this paper.

Like in the previous paper, we employ the likelihood ratio criterion, and in order to compute the likelihood ratio we naturally ought to compute the maximum

likelihood estimate. In the previous paper, we described a certain method to arrive at the maximum likelihood estimate. As this method involves intuitive judgements, it is not an algorithm in the strict sense. Bremner is said to have developed an algorithm based on the Kudô's method. There is, however, said to be no proof that this algorithm always leads to the solution (cf. Barlow [3], p. 177).

The existence of the solution and its uniqueness is guaranteed as is seen in the previous paper, because of the closedness and convexity of the polyhedral cone involved. Recently one of the authors worked out with Tarumi an algorithm related to all possible regressions and discriminations [9]. After it was published, we realized that this algorithm can be modified to compute the maximum likelihood estimate under the alternative hypothesis of the present paper. This is the second motivation of this paper.

In §2, we prove a theorem about the closest point y_0 from a point y in E^p to the convex polyhedral cone $\mathcal{S}$ generated by the linear inequalities concerned. This is a generalization of Theorem 3.1 in Kudô [4], and plays an important role in the subsequent arguments.

§3 is devoted to the description of an algorithm for computing the maximum likelihood estimate of the mean vector under the alternative hypothesis and also to compute the likelihood ratio statistic $\bar{\chi}^2$ for testing the homogeneity of means under the restricted alternative.

The purpose of §4 is to derive the distribution of the likelihood ratio test statistic $\bar{\chi}^2$ under the null hypothesis. We will also describe briefly the general aspect of determining the shape of the convex polyhedral cone $\mathcal{S}$ in our present concern.

In §3 and §4, we sometimes restrict ourselves to the case when the set of p -vectors $\{\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k\}$, which determines the linear restrictions, satisfies the condition that any subset of size p out of k is independent. We call this "the rank assumption", which is sometimes called as the Haar condition.

We are able to present an algorithm, in full details, for checking the rank assumption and for removing, if any, the redundant vectors in case when $p=3$ only, which will be described in §5 along with calculating of the upper tail probability. We nevertheless feel that the rank assumption is neither too restrictive nor too difficult to check in practical cases.

A numerical example is given in §6, and acknowledgements of priority and appreciations are in §7.

§2. A theorem on the closest point

Let us consider a convex polyhedral cone defined by

$$(2.1) \quad \mathcal{S} = \{ \underline{y} : (\underline{a}'_1, \underline{y}) \geq 0, (\underline{a}'_2, \underline{y}) \geq 0, \dots, (\underline{a}'_k, \underline{y}) \geq 0 \}.$$

Now for a given point $\underline{y}_1$, we define in $\mathcal{S}$ the closest point, which attains the distance from $\underline{y}_1$ to the convex polyhedral cone $\mathcal{S}$, i.e., $\underline{y}_0$ such that

$$(2.2) \quad \min_{\underline{y} \in \mathcal{S}} \|\underline{y}_1 - \underline{y}\| = \|\underline{y}_1 - \underline{y}_0\|$$

where $\|\cdot\| = \sqrt{(\cdot, \cdot)}$ is the Euclidean distance of two vectors. The closedness of the cone confirms the existence of $\underline{y}_0$ and the convexity its uniqueness.

Let $K = \{1, 2, \dots, k\}$ and let denote M a subset of K whose number of elements is less than p , and let $\mathcal{S}_M$ be a subset of $\mathcal{S}$ with

$$(2.3) \quad \mathcal{S}_M = \{ \underline{y} : (\underline{a}'_i, \underline{y}) = 0, i \in M, (\underline{a}'_i, \underline{y}) > 0, i \in K - M \},$$

which may be an empty set in general.

Let us define the linear subspace $\mathfrak{B}_M : \mathfrak{B}_M = \{ \underline{y} : (\underline{a}'_i, \underline{y}) = 0, i \in M \}$, which is the smallest sub-space containing $\mathcal{S}_M$ if $\mathcal{S}_M \neq \phi$.

Further let us define, following closely to the notations of Kudô [4], $\underline{a}^{(j)}$ as a vector contained in the space spanned by $\{ \underline{a}_i, i \in M \}$, orthogonal to $\{ \underline{a}_i : i \in M, i \neq j \}$, and $(\underline{a}^{(j)}, \underline{a}_j) = 1$.

In order to avoid cumbersome notations, we state the theorem, without loss of generality, for $M = \{1, 2, \dots, m\}$.

THEOREM 2.1. Suppose $\{ \underline{a}_1, \underline{a}_2, \dots, \underline{a}_m \}$, be a subset of independent vectors out of $\{ \underline{a}_1, \underline{a}_2, \dots, \underline{a}_k \}$, then a necessary and sufficient condition that for a given $\underline{y}_1$, the closest point $\underline{y}_0$ of $\underline{y}_1$, is contained in $\mathcal{S}_M$ is given by

$$(2.4) \quad \begin{bmatrix} t_1 \\ \vdots \\ t_m \end{bmatrix} = \begin{bmatrix} \underline{a}^{(1)'} \\ \vdots \\ \underline{a}^{(m)'} \end{bmatrix} \underline{y}_1 = \begin{bmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{bmatrix} (\underline{a}_1, \dots, \underline{a}_m)^{-1} \begin{bmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{bmatrix} \underline{y}_1 \leq 0$$

$$(2.5) \quad \begin{bmatrix} s_{m+1} \\ \vdots \\ s_k \end{bmatrix} = \begin{bmatrix} a'_{m+1} \\ \vdots \\ a'_k \end{bmatrix} \quad \hat{y}_1 > \underline{0}.$$

where $\hat{y}_1$ is the projection of y_1 on $\mathfrak{B}_M$.

REMARK 1. In case $m=0$, the condition (2.4) is vacuous and (2.5) becomes

$$(\underline{a}'_i, \underline{y}_1) > 0, \quad i = 1, 2, \dots, k.$$

REMARK 2. In case (2.4), (2.5) are not satisfied for any set of vectors $\{a_{i_1}, \dots, a_{i_m}\}$, $0 \leq m < p$, then $\underline{y}_0 = 0 \in \mathcal{S}_K$.

REMARK 3. Given a set of independent vectors $\{a_{i_1}, \dots, a_{i_m}\}$ a necessary and sufficient condition that $\mathcal{S}_M = \phi$ is that (2.4) and (2.5) are not satisfied for any $\underline{y}_1 \in E^p$.

PROOF. If $\underline{y}_0$ the closest point to $\mathcal{S}$, belongs to $\mathcal{S}_M$, $\underline{y}_0$ is the closest point to $\mathcal{S}_M$ and also is the closest point to $\mathfrak{B}_M$, i.e., $\underline{y}_0 = \hat{y}_1$, therefore it implies condition (2.5).

For the proof of condition (2.4), we suppose that (2.4) is not true, and without loss of generality, we shall show that $t_m > 0$ leads to a contradiction.

We can easily show that

$$(2.6) \quad \hat{y}_1 = \underline{y}_1 - (\underline{a}_1, \dots, \underline{a}_m) \left[\begin{bmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{bmatrix} (\underline{a}_1, \dots, \underline{a}_m) \right]^{-1} \begin{bmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{bmatrix} \underline{y}_1$$

and hence

$$(2.7) \quad \begin{bmatrix} t_1 \\ \vdots \\ t_m \end{bmatrix} = \begin{bmatrix} \underline{a}^{(1)'} \\ \vdots \\ \underline{a}^{(m)'} \end{bmatrix} \underline{y}_1 = \left[\begin{bmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{bmatrix} (\underline{a}_1, \dots, \underline{a}_m) \right]^{-1} \begin{bmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{bmatrix} \underline{y}_1.$$

Now consider

$$(2.8) \quad \underline{y}_1 = \lambda_1 \underline{a}_1 + \lambda_2 \underline{a}_2 + \dots + \lambda_{m-1} \underline{a}_{m-1} + \lambda_m \underline{a}^{(m)} + \underline{b}$$

where $\underline{b} \in \mathfrak{B}_M$.

Let $\hat{y}_1^{(m)}$ be the projection of $\underline{y}_1$ on $\mathfrak{B}_N$ such that

$$\mathfrak{B}_N = \{\underline{y}: (\underline{a}'_i, \underline{y}) = 0, \quad i = 1, 2, \dots, m-1\}.$$

It can be easily seen that $\hat{y}_1 = \underline{b}$ and

$$(2.9) \quad \hat{y}_1^{(m)} = \lambda_m \underline{a}^{(m)} + \underline{b}.$$

Therefore

$$(2.10) \quad (\underline{a}^{(m)'}, \underline{y}_1) = (\underline{a}^{(m)'}, \hat{y}_1^{(m)}) = \lambda_m (\underline{a}^{(m)'}, \underline{a}^{(m)}) = t_m$$

and hence, from the assumption $(\underline{a}^{(m)'}, \underline{y}_1) = t_m > 0$ and it implies $\lambda_m > 0$. From (2.9) we get

$$(2.11) \quad (\underline{a}'_m, \hat{y}_1^{(m)}) = \lambda_m > 0.$$

Now we consider the following two cases

Case 1. Suppose $\hat{y}_1^{(m)} \in \mathcal{S}_N$, where

$$\mathcal{S}_N = \{\underline{y}: (\underline{a}'_i, \underline{y}) = 0, \quad i \in N, (\underline{a}'_i, \underline{y}) > 0, \quad i \in K - N\}.$$

Since $\mathfrak{B}_N$ is a sub-space containing $\mathfrak{B}_M$ as its proper subspace and $\hat{y}_1^{(m)} = \hat{y}_1 = \underline{y}_0$. We have

$$(2.12) \quad \|\hat{y}_1^{(m)} - \underline{y}_1\| \leq \|\underline{y}_0 - \underline{y}_1\|,$$

which implies that $\underline{y}_0$ is not the closest point in $\mathcal{S}$ to $\underline{y}_1$. This is a contradiction.

Case 2. Suppose $\hat{y}_1^{(m)} \notin \mathcal{S}_N$, i.e., $(\underline{a}'_i, \hat{y}_1^{(m)}) \leq 0$ for some i ($i = m+1, \dots, k$).

Let us consider a line between $\hat{y}_1^{(m)}$ and $\hat{y}_1$, which is the totality of points

$$\underline{y}(p) = p \hat{y}_1^{(m)} + (1-p) \hat{y}_1 \quad (0 \leq p \leq 1).$$

Then

$$(2.13) \quad (\underline{a}'_i, \underline{y}(p)) = 0, \quad i = 1, 2, \dots, m-1,$$

$$(2.14) \quad (\underline{a}'_m, \underline{y}(p)) = p(\underline{a}'_m, \hat{y}_1^{(m)}),$$

and

$$(2.15) \quad (\underline{a}'_{m+i}, \underline{y}(p)) = p(\underline{a}'_{m+i}, \hat{\underline{y}}_1^{(m)}) + (1-p)(\underline{a}'_{m+i}, \hat{\underline{y}}_1) \\ i = 1, 2, \dots, k-m.$$

As the left side of (2.14) is positive by (2.11), and those of (2.15) are all positive for $p=0$, by a suitable choice of p , say p_0 , $\underline{y}(p_0)$ can be made to belong to $\mathcal{S}_N$, and hence $\|\underline{y}(p_0) - \underline{y}_1\| \not\leq \|\underline{y}_0 - \underline{y}_1\|$. Thus we are led to a contradiction.

The sufficient condition is proved as follows.

Applying the equation (2.6), we have that the equation of the plane passing through $\hat{\underline{y}}_1$ and orthogonal to the vector $\underline{y}_1 - \hat{\underline{y}}_1$ is given by

$$(2.16) \quad \underline{y}'_1(\underline{a}_1, \dots, \underline{a}_m) \left[\begin{pmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{pmatrix} (\underline{a}_1, \dots, \underline{a}_m) \right]^{-1} \begin{pmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{pmatrix} \underline{y} = 0$$

From (2.4), we can easily get

$$(2.17) \quad (t_1, t_2, \dots, t_m) \begin{pmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{pmatrix} \underline{y} = 0$$

The fact that $\underline{y}_1$ lies in the positive side of this plane, and that if the conditions to this theorem are satisfied, all the points in $\mathcal{S}$ are either on the plane or in the negative side leads us to the sufficiency.

§3. An algorithm for calculating $\bar{\chi}^2$ and the closest point

Let $\underline{Y} = (Y_1, Y_2, \dots, Y_p)$ be distributed independently normally with the same variance 1. About the mean, we consider the null hypothesis

$$(3.1) \quad H_0: E(\underline{Y}) = \underline{0}$$

against an alternative

$$(3.2) \quad H_1: E(\underline{a}'_i, \underline{Y}) \geq 0 \quad i = 1, 2, \dots, k \quad (\geq p)$$

where inequality is strict for at least one value of i .

Let the random variable $\underline{X}$ be defined by

$$(3.3) \quad \underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix} = \begin{bmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_k \end{bmatrix} \underline{Y} = A \underline{Y},$$

then it is distributed as a singular normal distribution and our hypotheses are transformed to

$$(3.4) \quad \hat{H}_0: E(\underline{X}) = \underline{0}$$

$$(3.5) \quad \hat{H}_1: E(\underline{X}) \geq \underline{0}$$

and the variance-covariance matrix A is

$$(3.6) \quad A = AA'$$

We note here that the problems involving (H_0, H_1) and $(\hat{H}_0, \hat{H}_1)$ are essentially the same. Therefore from A we can compute Λ , and from Λ , it is possible to compute A by the pivotal condensation method, for instance, as demonstrated in Rao [7], p. 293.

The purpose of this section is to describe an algorithm for calculating the test statistic $\bar{\chi}^2$ for $(\hat{H}_0, \hat{H}_1)$ given $\underline{Y} = \underline{y}$ by making use of $A = AA'$.

We shall use either of the notations $\underline{y}$ or $\underline{x} = A\underline{y}$ whenever one is more convenient than the other.

Now consider the partition of the vector $\underline{x}' = (\underline{x}'_{(1)}, \underline{x}'_{(2)})$ and the variance matrix A and the coefficient matrix A are partitioned correspondingly.

$$(3.7) \quad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \begin{bmatrix} A'_1 & A'_2 \end{bmatrix}$$

$$= \begin{bmatrix} \begin{bmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{bmatrix} & (\underline{a}_1, \dots, \underline{a}_m) & \begin{bmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{bmatrix} & (\underline{a}_{m+1}, \dots, \underline{a}_k) \\ \begin{bmatrix} \underline{a}'_{m+1} \\ \vdots \\ \underline{a}'_k \end{bmatrix} & (\underline{a}_1, \dots, \underline{a}_m) & \begin{bmatrix} \underline{a}'_{m+1} \\ \vdots \\ \underline{a}'_k \end{bmatrix} & (\underline{a}_{m+1}, \dots, \underline{a}_k) \end{bmatrix}$$

with the condition $|A_{11}| \neq 0$ or rank of A_{11} is m .

By sweeping out the following matrix taking A_{11} as the pivot

$$(3.8) \quad \begin{bmatrix} A_{11} & A_{12} & \underline{x}_{(1)} \\ A_{21} & A_{22} & \underline{x}_{(2)} \\ \underline{x}'_{(1)} & \underline{x}'_{(2)} & 0 \end{bmatrix},$$

we have

$$(3.9) \quad \begin{bmatrix} A_{11}^{-1} & 0 & 0 \\ -A_{21}A_{11}^{-1} & I & 0 \\ -\underline{x}'_{(1)}A_{11}^{-1} & 0 & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} & \underline{x}_{(1)} \\ A_{21} & A_{22} & \underline{x}_{(2)} \\ \underline{x}'_{(1)} & \underline{x}'_{(2)} & 0 \end{bmatrix} \\ = \begin{bmatrix} I & A_{11}^{-1}A_{12} & A_{11}^{-1}\underline{x}_{(1)} \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} & \underline{x}_{(2)} - A_{21}A_{11}^{-1}\underline{x}_{(1)} \\ 0 & \underline{x}'_{(2)} - \underline{x}'_{(1)}A_{11}^{-1}A_{12} & -\underline{x}'_{(1)}A_{11}^{-1}\underline{x}_{(1)} \end{bmatrix}$$

According to the theorem 2.1, we can easily show that, if the closest point $\underline{y}_0$ is located on $\mathcal{S}_M$,

$$(3.10) \quad A_{11}^{-1}\underline{x}_{(1)} = \left[\begin{bmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{bmatrix} (\underline{a}_1, \dots, \underline{a}_m) \right]^{-1} \begin{bmatrix} \underline{a}'_1 \\ \vdots \\ \underline{a}'_m \end{bmatrix} \underline{y} \leq \underline{0}$$

and

$$(3.11) \quad \underline{x}_{(2)} - A_{21}A_{11}^{-1}\underline{x}_{(1)} = \begin{bmatrix} \underline{a}'_{m+1} \\ \vdots \\ \underline{a}'_k \end{bmatrix} \underline{\hat{y}} > \underline{0}.$$

Next we find that $\underline{x}'_{(1)}A_{11}^{-1}\underline{x}_{(1)}$ is equal to $(\underline{y} - \underline{\hat{y}})'(\underline{y} - \underline{\hat{y}})$ by observing

$$(3.12) \quad (\underline{y} - \underline{\hat{y}})'(\underline{y} - \underline{\hat{y}}) = (A_1\underline{y})'(A_1A_1')^{-1}(A_1\underline{y}) \\ = \underline{x}'_{(1)}A_{11}^{-1}\underline{x}_{(1)}$$

In case the rank assumption is valid, any principal minor of size $p \times p$ is nonsingular. In particular

$$(3.13) \quad \begin{bmatrix} A & \underline{x} \\ \underline{x}' & 0 \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \underline{\tilde{x}}_{(1)} \\ \Sigma_{21} & \Sigma_{22} & \underline{\tilde{x}}_{(2)} \\ \underline{\tilde{x}}'_{(1)} & \underline{\tilde{x}}'_{(2)} & 0 \end{bmatrix} = \begin{bmatrix} B_1 B'_1 & B_1 B'_2 & \underline{\tilde{x}}_{(1)} \\ B_2 B'_1 & B_2 B'_2 & \underline{\tilde{x}}_{(2)} \\ \underline{\tilde{x}}'_{(1)} & \underline{\tilde{x}}'_{(2)} & 0 \end{bmatrix}$$

where

$$A = \begin{bmatrix} a'_1 \\ \vdots \\ a'_p \\ a'_{p+1} \\ \vdots \\ a'_k \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

and $\text{rank}(\Sigma_{11}) = p$.

By sweeping out (3.13) taking Σ_{11} as the pivot, we have

$$(3.14) \quad \begin{bmatrix} \Sigma_{11}^{-1} & 0 & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I & 0 \\ -\underline{\tilde{x}}'_{(1)}\Sigma_{11}^{-1} & 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \underline{\tilde{x}}_{(1)} \\ \Sigma_{21} & \Sigma_{22} & \underline{\tilde{x}}_{(2)} \\ \underline{\tilde{x}}'_{(1)} & \underline{\tilde{x}}'_{(2)} & 0 \end{bmatrix} \\ = \begin{bmatrix} I_p & \Sigma_{11}^{-1}\Sigma_{12} & \Sigma_{11}^{-1}\underline{\tilde{x}}_{(1)} \\ 0 & \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} & \underline{\tilde{x}}_{(2)} - \Sigma_{21}\Sigma_{11}^{-1}\underline{\tilde{x}}_{(1)} \\ 0 & \underline{\tilde{x}}'_{(2)} - \underline{\tilde{x}}'_{(1)}\Sigma_{11}^{-1}\Sigma_{12} & -\underline{\tilde{x}}'_{(1)}\Sigma_{11}^{-1}\underline{\tilde{x}}_{(1)} \end{bmatrix}$$

In (3.14), A is of rank p and matrix $\begin{bmatrix} \Sigma_{11}^{-1} & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{bmatrix}$ is of rank $k (\geq p)$, and

hence we must have

$$(3.15) \quad \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} = 0.$$

The above observation is quite important for computation. If we sweep out a matrix of the form (3.13), where A is singular, successively, we eventually come to the stage of (3.14) with (3.15), where there is no non-zero diagonal elements and thus no possibility of further sweep out remains. At this stage the extreme right bottom element is nothing but $-\underline{y}'\underline{y}$. Incidentally this rule can be

used without the rank assumption or in the absence of knowledge on the exact rank p of A .

If $x_i > 0$ ($i = 1, \dots, k$), $M = \phi$ and $\bar{\chi}^2 = \underline{y}'\underline{y} = \tilde{\underline{x}}_{(1)}'\Sigma_{11}^{-1}\tilde{\underline{x}}_{(1)}$, otherwise some calculations are needed. For a given $\underline{x}$, we ought to compute $\bar{\chi}^2$, $\hat{\underline{x}} = A\underline{\hat{y}}$, the closest point and M .

If it fails to satisfy the conditions (3.10) and (3.11) for all possible partitions, whose numerical verification is not a very heavy task when k is not too large, then we conclude $M = K$, $\underline{\hat{y}} = \underline{0}$, and $\bar{\chi}^2 = 0$.

In case when there turns out an intermediate situation where after making permutations of the first k rows and columns of (3.8) simultaneously, we have, for some A_{11} with $m \times m$,

$$A_{11}^{-1}\underline{x}_{(1)} \leq \underline{0}$$

and

$$\underline{x}_{(2)} - A_{21}A_{11}^{-1}\underline{x}_{(1)} > \underline{0},$$

then the closest point is

$$\underline{\hat{y}} = \underline{y} - A_1' A_{11}^{-1} \underline{x}_{(1)} = [I_p - A_1' A_{11}^{-1} A_1] \underline{y}$$

or in terms of $\underline{x}$,

$$(3.16) \quad \hat{\underline{x}} = \begin{bmatrix} 0 \\ \underline{x}_{(2)} - A_{21}A_{11}^{-1}\underline{x}_{(1)} \end{bmatrix},$$

and

$$(3.17) \quad \begin{aligned} \bar{\chi}^2 &= \underline{y}'\underline{y} - (\underline{y} - \underline{\hat{y}})'(\underline{y} - \underline{\hat{y}}) \\ &= \tilde{\underline{x}}_{(1)}'\Sigma_{11}^{-1}\tilde{\underline{x}}_{(1)} - \underline{x}_{(1)}'A_{11}^{-1}\underline{x}_{(1)} \end{aligned}$$

The subset M of K is the set of suffices of the components of $\hat{\underline{x}}$ in (3.16) equal to zeros.

It is interesting to note the following theorem.

THEOREM 2.2. *Let $\underline{x}$ be distributed as a singular normal distribution with the known variance-covariance matrix A of (3.7). The maximum likelihood*

estimate $\hat{x}$ of the mean vector $\underline{\theta} = \begin{bmatrix} \underline{\theta}_{(1)} \\ \underline{\theta}_{(2)} \end{bmatrix}$ under the restriction

$$\underline{\theta}_{(1)} = \underline{0}, \text{ i.e., } \theta_1 = \theta_2 = \dots = \theta_m = 0 \quad (m \leq p)$$

is given by

$$(3.18) \quad \hat{x} = \begin{bmatrix} 0 \\ \underline{x}_{(2)} - A_{21} A_{11}^{-1} \underline{x}_{(1)} \end{bmatrix}$$

where $\underline{x}_{(1)}$, $\underline{x}_{(2)}$, A_{11} , A_{12} are those of (3.9) and its variance-covariance matrix is

$$(3.19) \quad A_{22} - A_{21} A_{11}^{-1} A_{12}$$

PROOF. As $\underline{X}$ is distributed as a singular normal distribution $N(\underline{\theta}, A)$, there exists $k \times p$ matrix A such that $\underline{X} = AY$, $\underline{Y} \sim N(\underline{\mu}, I_p)$, and $\hat{\underline{\theta}} = \widehat{(A\underline{\mu})} = A\hat{\underline{\mu}}$, where $\hat{\underline{\theta}}$, $\hat{\underline{\mu}}$ are M. L. E. of $\underline{\theta}$, $\underline{\mu}$ respectively. The maximum of the likelihood function given $\underline{y}$, $L_{\underline{y}}$ is

$$\begin{aligned} \text{Max}_{\substack{(\underline{a}'_i, \underline{\mu}) = \underline{0} \\ i \in M}} L_{\underline{y}} &= \max_{\substack{(\underline{a}'_i, \underline{\mu}) = \underline{0} \\ i \in M}} C \cdot \exp \left\{ -\frac{1}{2} (\underline{y} - \underline{\mu})' (\underline{y} - \underline{\mu}) \right\} \\ &= C \cdot \exp \left[-\frac{1}{2} \min_{\substack{(\underline{a}'_i, \underline{\mu}) = \underline{0} \\ i \in M}} (\underline{y} - \underline{\mu})' (\underline{y} - \underline{\mu}) \right], \end{aligned}$$

where C is a constant.

Therefore we have

$$\hat{\underline{\mu}} = \{ \underline{\mu} : \min_{\substack{(\underline{a}'_i, \underline{\mu}) = \underline{0} \\ i \in M}} (\underline{y} - \underline{\mu})' (\underline{y} - \underline{\mu}) \},$$

i.e., $\hat{\underline{\mu}}$ is the projection of $\underline{y}$ to $\mathfrak{B}_M$ and then $\hat{\underline{\mu}}$ is equal to $\hat{\underline{y}}$ in (2.6), proving (3.18). (3.19) can be easily computed from (3.16).

Our algorithm can be said, in short, that if we compute the M. L. E. of $\underline{\theta}$, assuming $\underline{\theta}_i = \underline{0}$, $i \in M$, for all subset M of K with $n(M) < p$, then there will be found uniquely, if any, a set M , $0 < n(M) < p$, such that (3.10) and (3.11) are satis-

fied, and this M is the set we are looking for.

A special mention deserves to be made here on the paper of Tarumi and Kudô [9] where they described an algorithm to perform the Gaussian eliminations taking all possible principal minors as pivots. They also described the method to limit the principal minors of less than a prescribed size. In our case we can apply their method to the matrix of the form (3.8).

In case the rank assumption is valid, and p is known, what we need to modify their algorithm is merely to examine the signs of components of the last column vector at each step to see whether the condition (3.10) and (3.11) are satisfied or not.

Their algorithm can be applied in case when there is enough storage space to memorize a number of matrices in the computer at a same time. In addition, if a computer is equipped with enough accuracy, and we can be sure whether the pivot at each step is equal to zero or not, we need not know the exact value of p . In this case, moreover, we can remove the rank assumption of this section.

§4. Distribution of the likelihood ratio statistics $\bar{\chi}^2$

In order to evaluate the upper tail probability of $\bar{\chi}^2$, we ought to calculate the probability measure of a set determined by a convex polyhedral cone (2.1).

In the case of Kudô [4], the cone is determined by p linear inequalities, whereas in our case the number of inequalities is more than p , and thus a special care is needed to determine the shape of the convex polyhedral cone. For each M , we ought to determine whether the subset $\mathcal{S}_M$ or the convex polyhedral cone $\mathcal{S}$ is empty or not.

Now we write $\mathcal{S}^{(+)} = \mathcal{S}$, i.e.,

$$(4.1) \quad \mathcal{S}^{(+)} = \{ \underline{y} : (\underline{a}'_i, \underline{y}) \geq 0, i \in K \}$$

and define

$$(4.2) \quad \mathcal{S}^{(-)} = \{ \underline{y} : (-\underline{a}'_i, \underline{y}) \geq 0, i \in K \}.$$

$\mathcal{S}^{(+)}$ and $\mathcal{S}^{(-)}$ are symmetric at the origin to each other. A line L determined by $p-1$ independent set

$$\{\underline{a}_{i_1}, \dots, \underline{a}_{i_{p-1}}\}: L = \{\underline{y}: (\underline{a}'_{i_j}, \underline{y}) = 0, j = 1, 2, \dots, p-1\}$$

determines a ray of $\mathcal{S} = \mathcal{S}^{(+)}$ if $L \subset \mathcal{S}^{(+)} \cup \mathcal{S}^{(-)}$. If we can search out all rays of the polyhedral convex cone $\mathcal{S}^{(+)}$, we can determine the shape of $\mathcal{S}^{(+)} = \mathcal{S}$ and hence we can select M with $\mathcal{S}_M \neq \emptyset$ from the totality of the subsets $K = \{1, 2, \dots, k\}$.

Now consider a line L , determined by the planes, say, $(\underline{a}'_{i_j}, \underline{y}) = 0, j = 1, 2, \dots, p-1$, and take a point $\underline{y} \neq \underline{0}$. And let $\underline{y}^{(L)}$ be the projection of $\underline{y}$ on L . If the line L determines a ray, then $\underline{y}^{(L)} \in \mathcal{S}^{(+)} \cup \mathcal{S}^{(-)}$, and if $\underline{y}^{(L)} \neq \underline{0}$ and $\underline{y}^{(L)} \in \mathcal{S}^{(+)}$,

$$(4.3) \quad (\underline{a}'_{i_j}, \underline{y}^{(L)}) > 0, \quad \text{for all } i_j \neq i_1, \dots, i_{p-1}.$$

If $\underline{y}^{(L)} \neq \underline{0}$ and $\underline{y}^{(L)} \in \mathcal{S}^{(-)}$,

$$(4.4) \quad (\underline{a}_{i_j}, \underline{y}^{(L)}) < 0, \quad \text{for all } i_j \neq i_1, \dots, i_{p-1}.$$

We can check whether the inequalities (4.3) or (4.4) hold or not by the algorithm of § 3.

Taking a point $\underline{y} \neq \underline{0}$ arbitrary, and let $x = A\underline{y}$, we first apply (2.6) to check $\hat{y} = \underline{0}$ or not, and if not, we apply the form of (3.8), (3.9) and (3.11) in § 3 with $m = p-1$. For convenience, the point $\underline{y}$ is initially chosen as $\underline{y}' = (n, n, \dots, n)$ where n is a positive integer of adequate magnitude such as $n = 2$ or 3 . In case $\hat{y} = \underline{0}$, we change the sign of the first component of $\underline{y}$ and compute the corresponding procedure. If this revision of $\underline{y}$ is not successful, we change the sign of the second component of $\underline{y}$. In this way we continue to revise the point $\underline{y}$, in general at the i -th step ($i < p$). From Remark 3 in § 2 it is quite clear that we should get the case of $\hat{y} \neq \underline{0}$ before the p -th revision is needed.

In (3.9), if $\underline{x}_{(2)} - A_{21}A_1^{-1}\underline{x}_{(1)} \geq \underline{0}$, $\underline{y}^{(L)} \in \mathcal{S}^{(+)}$, and if $\underline{x}_{(2)} - A_{21}A_1^{-1}\underline{x}_{(1)} \leq \underline{0}$, $\underline{y}^{(L)} \in \mathcal{S}^{(-)}$, and we conclude the point $\underline{y}^{(L)}$ determines a ray of $\mathcal{S}$, otherwise $\underline{y}^{(L)}$ does not. In this manner we repeatedly make use of the algorithm of § 3 in order to examine all of the lines determined by the $(p-1)$ independent vectors of $\{\underline{a}_i, i \in K\}$.

As any point in $\mathcal{S}$ can be written as a positive linear form of rays, i.e.,

$$(4.6) \quad \underline{y} = \sum \lambda_i \underline{y}_{(i)}^{(L)}, \quad \text{where } \underline{y}_{(i)}^{(L)} \text{ is a ray, and } \lambda_i \geq 0.$$

$\mathcal{S}_N \neq \phi$ implies $\mathcal{S}_M \neq \phi$ for all $M \subset N$ when the rank assumption is valid. If this assumption is not valid, the following rule might be useful. Let $\mathcal{S}_N \neq \phi$, and let M be subset of N with the size $n(M)$, then $\mathcal{S}_M \neq \phi$ when there exist N_i ($i=1, \dots, n(M)$) such that $\mathcal{S}_{N_i} \neq \phi$, $M = \bigcap_{i=1}^{n(M)} N_i$. This rule enables us to remove redundant vectors unnecessary for the determination of the shape of $\mathcal{S}$.

In this manner, for any M we can determine $\mathcal{S}_M \neq \phi$ or not, and theorem 2.1 leads us to the expression of the distribution of $\bar{\chi}^2$ statistic in terms of $\underline{Y}$, $\underline{\hat{Y}}$, $\underline{a}_i$ and $\underline{a}^{(i)}$,

$$\begin{aligned}
 (4.7) \quad Pr(\bar{\chi}^2 \geq c^2) &= \sum_{\phi \subseteq M \subseteq K} Pr(\underline{\hat{Y}}' \underline{\hat{Y}} \geq c^2, \quad \underline{\hat{Y}} \in \mathcal{S}_M) \\
 &= \sum_{\phi \subseteq M \subseteq K} Pr\{\underline{\hat{Y}}' \underline{\hat{Y}} \geq c^2, \quad (\underline{a}'_i, \underline{\hat{Y}}) > 0, \quad i \in K-M, \\
 &\quad (\underline{a}^{(i)'}, \underline{Y}) \leq 0, \quad i \in M\}
 \end{aligned}$$

where the summation runs over all the subset M of K with $\mathcal{S}_M \neq \phi$. Without loss of generality, let $M = \{1, 2, \dots, m\}$ and consider the term $Pr(\underline{\hat{Y}}' \underline{\hat{Y}} \geq c^2, (\underline{a}'_j, \underline{\hat{Y}}) > 0, j \in K-M, (\underline{a}^{(i)'}, \underline{Y}) \leq 0, i \in M)$. Setting the variables $S_{m+i} = (\underline{a}'_{m+i}, \underline{\hat{Y}})$, we get

$$(4.8) \quad \begin{bmatrix} S_{m+1} \\ \vdots \\ S_k \end{bmatrix} = \begin{bmatrix} \underline{a}'_{m+1} \\ \vdots \\ \underline{a}'_k \end{bmatrix} \underline{\hat{Y}} = \underline{X}_{(2)} - \underline{A}_{21} \underline{A}_1^{-1} \underline{X}_{(1)}$$

from (3.11) and the variance-covariance matrix is

$$(4.9) \quad \underline{A}_{22} - \underline{A}_{21} \underline{A}_1^{-1} \underline{A}_{12}.$$

(4.8) and (4.9) show that S_{m+i} 's are distributed in the multivariate normal (singular) form with the variance-covariance matrix (4.9) which is the conditional one of $(X_{m+1}, \dots, X_k)$ when $X_1 = \dots = X_m = 0$.

Next we recall that the pivotal condensation method enables us to make the coefficient matrix A partitioned in such a way as $A = \begin{bmatrix} B \\ D \end{bmatrix}$, where $B(p \times p)$ is a lower triangle matrix. Using the coefficient matrix above, we have another expression in terms of the new coordinates

$$\underline{X}'_{(1)} \underline{A}_1^{-1} \underline{X}_{(1)} = Y_1^2 + \dots + Y_m^2 \quad (m \leq p)$$

Because of (3.17) we have

$$(4.10) \quad \bar{\chi}^2 = Y_{m+1}^2 + \cdots + Y_p^2.$$

We can easily find that $(\underline{a}^{(i)'}, \underline{Y})$, $i \in M$, are distributed to a normal distribution $N(0, A_1^{-1})$, and are functions of $Y_1, Y_2, \dots, Y_m$ only, and hence statistically independent of $S_{m+1}, \dots, S_k$. Therefore we can write

$$\begin{aligned} & \Pr\{\underline{\hat{Y}}' \underline{\hat{Y}} \geq c^2, S_j > 0, j \in K-M, (\underline{a}^{(i)'}, \underline{Y}) \leq 0, i \in M\} \\ &= \Pr\{\underline{\hat{Y}}' \underline{\hat{Y}} \geq c^2, S_j > 0, j \in K-M\} \cdot \Pr\{(\underline{a}^{(i)'}, \underline{Y}) \leq 0, i \in M\}. \end{aligned}$$

Finally by making use of lemma 3.2 in Kudô [4], we can get the following theorem.

THEOREM 4.1. *Let $(X_1, \dots, X_k)$ be distributed in a multivariate normal (singular) distribution with mean zero, and the variance-covariance matrix A , the probability that the value of $\bar{\chi}^2$ statistic, computed taking the null hypothesis as $\hat{H}_0: E(X_i) = 0$ ($i = 1, \dots, k$) and the alternative as $\hat{H}_1: E(X_i) \geq 0$ ($i = 1, \dots, k$) where the inequality is strict for at least one value of i , exceeds c^2 is given by*

$$\begin{aligned} (4.12) \quad \Pr(\bar{\chi}^2 \geq c^2) &= \sum_{\substack{\phi \subseteq M \subseteq K \\ \mathcal{S}_M \neq \phi}} \Pr(\chi_{p-n(M)}^2 \geq c^2) \cdot \Pr\{(\underline{a}_i', \underline{\hat{Y}}) > 0, i \in K-M\} \\ &\quad \cdot \Pr\{(\underline{a}^{(i)'}, \underline{Y}) \leq 0, i \in M\} \\ &= \sum_{\substack{\phi \subseteq M \subseteq K \\ \mathcal{S}_M \neq \phi}} \Pr(\chi_{p-n(M)}^2 \geq c^2) \cdot \Pr\{(A_M)^{-1}\} \cdot \Pr\{A_{K-M:M}\} \end{aligned}$$

where the summation runs over all the subsets M of $K = \{1, \dots, k\}$ with $\mathcal{S}_M \neq \phi$, $n(M)$ is the number of elements in M , A_M is the variance matrix of X_i , $i \in M$, $A_{K-M:M}$ is the variance matrix of X_i , $i \in K-M$ under the condition $X_j = 0$, $j \in M$, and $\Pr(\Sigma)$ is the probability that the variables distributed in a multivariate normal distribution with mean zero and variance-covariance matrix Σ are all positive. $\chi_{p-n(M)}^2$ has the χ^2 distribution with $p-n(M)$ degrees of freedom, where $\chi_{p-n(M)}^2$ is to be understood as a constant zero if $n(M) = p$ and $c^2 > 0$, if $n(M) = p$, $\Pr(A_{K-M:M}) = 1$ and $p\{(A_\phi)^{-1}\} = 1$.

REMARK. The formula (4.12) can be written as

$$(4.13) \quad \Pr(\bar{\chi}^2 \geq c^2) = \sum_{i=0}^p P(i; A) \Pr(\chi_i^2 \geq c^2).$$

The number $P(i; A)$ is called as the i -th coefficient to the upper tail probability

of χ^2 with degree of freedom i .

§5. An algorithm for computing the $\bar{\chi}^2$ -statistic and its significance level in case when $\text{rank}(A)=3$

When the variance-covariance A is of rank 3, we can describe an algorithm to determine the rays of the convex polyhedral cone $\mathcal{S}$ and also remove redundant variables unnecessarily to determine the shape of $\mathcal{S}$. This algorithm is based on the geometrical interpretation of the polyhedral convex cone $\mathcal{S}$ in the three dimensional Euclidean space E^3 .

Let $\underline{X}$ be normally distributed random vector with mean zeros and the variance-covariance matrix $A(k \times k)$, where $\text{rank}(A)=3$. By the pivotal condensation method, we can get the coefficient matrix $A(k \times 3)$ and the variable $\underline{Y}$: $\underline{Y} \sim N(0, I_3)$ with

$$(5.1) \quad A = AA', \quad \underline{X} = A\underline{Y}.$$

There are as many choice of A 's as the number of feasible permutations of components of $\underline{x}$, and we may consider our problem in terms of $\underline{y}$ in E^3 instead of $\underline{x}$ in E^k .

On the surface of the unit sphere, the cone $\mathcal{S}$ is represented by the polygon $\bar{\mathcal{S}}$, which is the intersection of $\mathcal{S}$ and the unit sphere. A ray L of $\mathcal{S}$ corresponds to a vertex of the polygon $\bar{\mathcal{S}}$. When we find the vertices and their circular order, the shape of $\mathcal{S}$ is determined and at the same time the redundant vectors of A are removed.

After determining the shape of the polygon, we replace A by the new coefficient matrix $\tilde{A}$ by removing the redundant vectors and also the variance-covariance matrix A is revised as $\tilde{A}$ so that our rank assumption is satisfied.

By removing the redundants, we have a new variable $\tilde{\underline{X}}$ distributed as $N(0, \tilde{A})$ and the problem for $\tilde{\underline{X}}$ is equivalent to our problem for $\underline{X} \sim N(0, A)$ and we can easily compute the χ^2 -statistic and its significance level by making use of $\tilde{A}$ and the corresponding polygon.

Our algorithm can be summerized in the following five steps although we can start from the second step in most of the practical situations and the last was already described in §3.

- (1) *Determination of the coefficient matrix A .*
- (2) *Determination of the shape of the polyhedral cone.*
 - (a) *The polygon $\bar{\mathcal{P}}$ on the unit sphere.*
 - (b) *The vertices cycle on the unit sphere.*
- (3) *Calculation of the coefficient $P(1; \Lambda)$ and $P(2; \Lambda)$*
 - (a) *Determination of the inside planes of the cone, or the inside line in the polygon.*
 - (b) *Determination of the coefficients.*
- (4) *Calculation of the coefficient $P(3; \Lambda)$.*
- (5) *Computation of $\bar{\chi}^2$ -statistic (§ 3).*

(1) *Determination of the coefficient matrix A .* We can easily compute the coefficient matrix A , satisfying $\underline{X} = A\underline{Y}$, $\underline{Y} \sim N(0, I_3)$ by the pivotal condensation method (cf. Rao [7], p. 293) keeping in mind that A may be singular.

(2) *Determination of the shape of the polyhedral convex cone $\mathcal{S}$.*

Step 1. We start with finding an independent pair of vectors, say $(\underline{a}_1, \underline{a}_2)$, and calculate a point $\underline{p}_1$ of intersection on the unit sphere, i.e.,

$$(5.2) \quad \underline{p}_1 = \{ \underline{y} : (\underline{a}'_1, \underline{y}) = 0, (\underline{a}'_2, \underline{y}) = 0, \underline{y}'\underline{y} = 1 \}$$

For $\underline{p}_1$, there are the following three cases.

Case 1.

$$(5.3) \quad (\underline{a}'_i, \underline{p}_1) \geq 0 \quad \text{for all } i \in K.$$

In this case $\underline{p}_1$ is located in the polygon $\bar{\mathcal{P}}$, where $\bar{\mathcal{P}} = \mathcal{S}^{(+)} \cap \{ \underline{y} : \underline{y}'\underline{y} = 1 \}$ and $\underline{p}_1$ is a vertex of $\bar{\mathcal{P}}$ and also the vector $\underline{p}_1$ is a ray of $\mathcal{S}$. Then we go to step 2.

Case 2.

$$(5.4) \quad (\underline{a}'_i, \underline{p}_1) \leq 0 \quad \text{for all } i \in K.$$

$\underline{p}_1$ is located in $\mathcal{S}^{(-)} \cap \{ \underline{y} : \underline{y}'\underline{y} = 1 \}$, and in this case, by merely changing the signs of the elements of $\underline{p}_1$, we can make $\underline{p}_1$ to satisfy the condition of case 1, and proceed to step 2.

Case 3. The point $\underline{p}_1$ fails to satisfy either of the inequalities (5.3) or (5.4), $\underline{p}_1$ is not a vertex of $\bar{\mathcal{P}}$, and in this case, we examine the other pairs $(\underline{a}_1, \underline{a}_3)$,

$(\underline{a}_1, \underline{a}_4), \dots, (\underline{a}_1, \underline{a}_k)$. If (5.3) or (5.4) fails to hold with all the above pairs, the condition $(\underline{a}'_1, \underline{y}) \geq 0$ or the vector $\underline{a}_1$ is judged as redundant.

When $\underline{a}_1$ is a redundant vector, we choose another vector, say $\underline{a}_2$ and repeat the step 1 to examine the condition (5.3) or (5.4) for the pairs $(\underline{a}_2, \underline{a}_3), (\underline{a}_2, \underline{a}_4), \dots$.

In this manner, we repeat the procedure of step 1 until we find a pair, say $(\underline{a}_{i_1}, \underline{a}_{i_2})$, which satisfies (5.3) or (5.4) and then we proceed to the step 2.

If we fail to find such a pair $(\underline{a}_{i_1}, \underline{a}_{i_2})$ that (5.3) or (5.4) are satisfied, the convex set $\mathcal{S}$ is empty, and the process is terminated here.

Step 2. When we enter step 2, we are given a vertex $\underline{p}_1$ and a set of vectors satisfying $(\underline{a}'_i, \underline{p}_1) = 0$, which we denote as

$$(5.5) \quad \{\underline{a}_{i_1}, \underline{a}_{i_2}, \dots, \underline{a}_{i_m}\},$$

and let

$$(5.6) \quad M_1 = \{i_1, i_2, \dots, i_m\} = \{i: (\underline{a}'_i, \underline{p}_1) = 0\}$$

In step 2, we examine, in a systematic manner, the pairs $(\underline{a}_\mu, \underline{a}_\nu)$, one from $M_1; \mu \in M_1$ and the other from $K - M_1; \nu \in K - M_1$, about the inequalities (5.3) and (5.4).

Assume that there exists a pair $(\underline{a}_\mu, \underline{a}_\nu)$, $\mu \in M_1, \nu \in K - M_1$, such that the intersecting point $\underline{p}_2$ for the pair $(\underline{a}_\mu, \underline{a}_\nu)$, satisfies (5.2) and (5.3) then $\underline{p}_2$ is the second ray, and $\underline{a}_\mu$ is a vector which determines the facet between the rays $\underline{p}_1$ and $\underline{p}_2$. We denote $\underline{a}_\mu$ as $\underline{a}_{j_1}$.

If we fail to find the other ray, $\mathcal{S}$ consists of the single ray $\underline{p}_1$ only, and the process terminates here.

For the vertex $\underline{p}_2$, we determine M_2 similarly.

$$(5.7) \quad M_2 = \{i: (\underline{a}'_i, \underline{p}_2) = 0, i \in K\}$$

and proceed to step 3.

Step 3. Like in the step 2 we examine, in a systematic manner, the pairs $(\underline{a}_\mu, \underline{a}_\nu)$, $\mu \in M_2, \nu \in K - M_2$ and look for the third ray $\underline{p}_3$, determined by a pair $(\underline{a}_\mu, \underline{a}_\nu)$. If we succeed to find the third ray $\underline{p}_3$, we denote such $\underline{a}_\mu$ as

$\underline{a}_{j_2}$ and determine M_3 again similarly

$$M_3 = \{i: (\underline{a}_i, \underline{p}_3) = 0, i \in K\}.$$

Here we note that the vectors in $\{\underline{a}_i: i \in M_2, \underline{a}_i \neq \underline{a}_{j_1}, \underline{a}_{j_2}\}$ are redundant ones.

In case we fail to find the third ray, $\mathcal{S}$ consists of only two rays $\underline{p}_1$ and $\underline{p}_2$, and we terminate the procedure, otherwise we proceed.

After obtaining the third ray, we repeat in the similar manner, the procedures in step 3 until we return to the first ray $\underline{p}_1$, say at the $(N+1)$ -th step: $\underline{p}_{N+1} = \underline{p}_1$.

We thus have the cycle of vertices, (or rays)

$$(5.8) \quad \{\underline{p}_1, \underline{p}_2, \dots, \underline{p}_{N+1} = \underline{p}_1\}$$

and that of vectors

$$(5.9) \quad \{\underline{a}_{j_1}, \underline{a}_{j_2}, \dots, \underline{a}_{j_N}\}$$

At the same time we have excluded the redundant vectors, which are not included in the above cycle (5.9).

We have described the algorithm as if a computer were equipped with enough accuracy. Our experience suggests that it is enough to execute only one or two additional steps, for extra assurances, after returning to the original ray $\underline{p}_1$, and any more computations such as going through the cycle in the reverse order is not needed.

We now have the simplified representation of the convex cone $\mathcal{S}$ by the vectors of (5.9), i.e.,

$$(5.10) \quad \mathcal{S} = \{\underline{y}: (\underline{a}'_{j_1}, \underline{y}) \geq 0, \dots, (\underline{a}'_{j_N}, \underline{y}) \geq 0\}$$

Let

$$(5.11) \quad \tilde{A}' = (\underline{a}_{j_1}, \dots, \underline{a}_{j_N}).$$

and

$$(5.12) \quad \tilde{A} = \tilde{A}\tilde{A}'.$$

We can also reduce the given sample value $\underline{x}$ to $\tilde{\underline{x}}$ accordingly. Then our hypotheses are reduced to

$$(5.13) \quad \tilde{H}_0: E(\tilde{X}) = \underline{0}$$

$$(5.14) \quad \tilde{H}_1: E(\tilde{X}) \geq \underline{0}.$$

There is no essential loss in considering the reduced problem $(\tilde{H}_0, \tilde{H}_1)$ instead of the original test (H_0, H_1) .

For convenience in the subsequent descriptions we denote the vectors (5.9) and the rays in (5.8) as

$$(5.15) \quad \{\underline{a}_1, \underline{a}_2, \dots, \underline{a}_N\}$$

and

$$(5.16) \quad \{\underline{p}_{12}, \underline{p}_{23}, \dots, \underline{p}_{N1}\}$$

$$\text{where} \quad \underline{p}_{i,i+1} = \{\underline{y}: (\underline{a}'_i, \underline{y}) = 0, (\underline{a}'_{i+1}, \underline{y}) = 0, (\underline{a}'_j, \underline{y}) > 0, \\ j \neq i, i+1, (\underline{y}', \underline{y}) = 1\}$$

In accordance with (5.15) and (5.16), we denote (5.11) and (5.12) by

$$A' = (\underline{a}_1, \underline{a}_2, \dots, \underline{a}_N)$$

$$A = AA'.$$

(3) *Calculation of the $P(1; A)$ and $P(2; A)$.* Making use of the cycle (5.15) and (5.16), we can compute the coefficient to the upper tail probabilities of χ^2_1 and χ^2_2 of (4.13) in theorem 4.1.

Consider a plane, passing through the vertices $\underline{p}_{i-1,i}$, $\underline{p}_{i+1,i+2}$, and the origin, whose coefficient vector $\underline{c}_i$ is so oriented that the vertex $\underline{p}_{i,i+1}$ is located in the positive side of this plane. The spherical triangle $[\underline{p}_{i-1,i}, \underline{p}_{i,i+1}, \underline{p}_{i+1,i+2}]$ corresponds to the trihedral convex cone

$$(5.17) \quad \mathcal{S}_{(i,i+1)} = \{\underline{y}: (\underline{a}'_i, \underline{y}) \geq 0, (\underline{a}'_{i+1}, \underline{y}) \geq 0, (\underline{c}'_i, \underline{y}) \geq 0\}$$

The following lemma enables us to compute all the coefficients to the upper tail probabilities of χ^2 with degree of freedom 1 and 2 in (4.13). Each of the values turns out to be equal to the half of a bivariate orthant probability. The values of $P(1; A)$ and $P(2; A)$ in (4.13) can be obtained by summing up these terms respectively.

We now state the lemma.

LEMMA.

(i)

$$(5.18) \quad \begin{aligned} &Pr\{(\underline{a}^{(i)'}, \underline{Y}) \leq 0\} \cdot Pr\{(\underline{a}'_j, \underline{\hat{Y}}) > 0, j \in N - \{i\}\} \\ &= \frac{1}{2} Pr\{(\underline{a}'_{i+1}, \underline{\hat{Y}}) > 0, (\underline{c}'_i, \underline{\hat{Y}}) > 0\}, \end{aligned}$$

where $(\underline{a}^{(i)'}, \underline{a}_i) = 1$, $(\underline{a}^{(i)'}, \underline{a}_j) = 0$, $i \neq j$, and $\underline{\hat{Y}}$ is the projection of $\underline{Y}$ on the plane $\mathfrak{B} = \{\underline{y} : (\underline{a}'_i, \underline{y}) = 0\}$.

(ii)

$$(5.19) \quad \begin{aligned} &Pr\{(\underline{a}^{(i)'}, \underline{Y}) \leq 0, (\underline{a}^{(i+1)'}, \underline{Y}) \leq 0\} \cdot Pr\{(\underline{a}'_j, \underline{\hat{Y}}) > 0, j \in N - \{i, i+1\}\} \\ &= Pr\{(\underline{a}^{(i)'}, \underline{Y}) \leq 0, (\underline{a}^{(i+1)'}, \underline{Y}) \leq 0\} \cdot Pr\{(\underline{c}'_i, \underline{\hat{Y}}) > 0\} \\ &= \frac{1}{2} Pr\{(\underline{a}^{(i)'}, \underline{Y}) \leq 0, (\underline{a}^{(i+1)'}, \underline{Y}) \leq 0\} \end{aligned}$$

where $\underline{a}^{(i)}$, $\underline{a}^{(i+1)}$, and $\underline{\hat{Y}}$ are the same analogue to (i).

PROOF. We give the proof of (i) only, as that of (ii) is quite similar.

Consider the rays $\underline{p}_{i-1,i}$ and $\underline{p}_{i,i+1}$, then

$$\underline{\hat{Y}} = \alpha \underline{p}_{i-1,i} + \beta \underline{p}_{i,i+1}, \quad \alpha, \beta > 0.$$

Therefore $\{(\underline{a}'_j, \underline{\hat{Y}}) > 0, j \in N - \{i\}\}$ is equivalent to $\{(\underline{a}'_{i+1}, \underline{\hat{Y}}) > 0, (\underline{c}'_i, \underline{\hat{Y}}) > 0\}$.

(4) Calculation of the coefficient $P(3; A)$. The spherical polygon whose vertices are (5.16) is equal to the disjoint sum of the following spherical triangles

$$(5.20) \quad \begin{aligned} &[\underline{p}_{12}, \underline{p}_{23}, \underline{p}_{34}], [\underline{p}_{12}, \underline{p}_{34}, \underline{p}_{45}], \dots \\ &\dots, [\underline{p}_{12}, \underline{p}_{i-1,i}, \underline{p}_{i,i+1}], \dots [\underline{p}_{12}, \underline{p}_{N-1,N}, \underline{p}_{N1}] \end{aligned}$$

The area of each one of them is proportional to the orthant probability of the nonsingular trivariate normal distribution $N(0, \Phi)$, where Φ is determined by the three planes which determine the respective spherical triangle.

The algorithm for dividing the polygon into these triangles is obvious, and the trivariate orthant probability equals to the area of the spherical triangle

divided by 4π and is given by the well known formula;

$$(5.21) \quad Pr\{A_{(1,2,3)}\} = \frac{1}{4\pi} [2\pi - \cos^{-1}\rho_{12} - \cos^{-1}\rho_{13} - \cos^{-1}\rho_{23}]$$

where $A_{(1,2,3)} = (\rho_{ij})$, $\rho_{ii} = 1$, $i, j = 1, 2, 3$.

After describing the algorithm, we state a theorem without proof, which we believe, gives us an insight on the nature of our algorithm.

THEOREM 5.1. *Given a singular $k \times k$ matrix A with rank 3, the orthant probability $Pr\{A\} = Pr\{X_i \geq 0, i = 1, 2, \dots, k\}$ can be expressed in the following form. There is a chain of variables $X_{i_1}, X_{i_2}, \dots, X_{i_N}$ ($N \leq k$) and $N-3$ dummy variables $u_1, \dots, u_{N-3}$ such that*

$$(5.22) \quad \begin{aligned} Pr\{A\} = & Pr\{X_{i_1} > 0, X_{i_2} > 0, u_1 > 0\} \\ & + Pr\{-u_1 > 0, X_{i_3} > 0, u_2 > 0\} \\ & + \dots \dots \dots \\ & + Pr\{-u_j > 0, X_{i_{j+2}} > 0, u_{j+1} > 0\} \\ & + \dots \dots \dots \\ & + Pr\{-u_{N-3} > 0, X_{i_{N-1}} > 0, X_{i_N} > 0\} \end{aligned}$$

where each term is the trivariate orthant probability.

§6. A numerical example in case when rank $(A) = 3$

Let the coefficient matrix A be

$$A = \begin{bmatrix} \underline{a}'_1 \\ \underline{a}'_2 \\ \underline{a}'_3 \\ \underline{a}'_4 \\ \underline{a}'_5 \\ \underline{a}'_6 \\ \underline{a}'_7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & -2 \\ 1 & 1 & 1 \\ 1 & 2 & -2 \end{bmatrix}$$

and the sample value be $\underline{y}' = (2.5, -1.5, 0.5)$, then the corresponding variance-covariance matrix is

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 & 2 & 1 & 1 \\ & 1 & 0 & 1 & 1 & 1 & 2 \\ & & 1 & 0 & -2 & 1 & -2 \\ & & & 2 & 3 & 2 & 3 \\ & & & & 9 & 1 & 8 \\ & & & & & 3 & 1 \\ & & & & & & 9 \end{bmatrix}$$

and the corresponding vector $\underline{x}$ is given by

$$\underline{x}' = (2.5, -1.5, 0.5, 1.0, 2.5, 1.5, -1.5).$$

$\underline{a}'_4, \underline{a}'_6$ are found to be redundant, and A and Λ are reduced to

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & -2 \\ 1 & 2 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} 1 & 2 & 1 & 0 & 0 \\ & 9 & 8 & 1 & -2 \\ & & 9 & 2 & -2 \\ & & & 1 & 0 \\ & & & & 1 \end{bmatrix}$$

and $\underline{x}$ is also revised;

$$\underline{x} = (2.5, 2.5, -1.5, -1.5, 0.5).$$

The coefficients are found to be

$$\begin{aligned} P(1; A) &= 0.06693 + 0.03787 + 0.06693 \\ &\quad + 0.12500 + 0.12500 \\ &= 0.42173 \end{aligned}$$

$$\begin{aligned} P(2; A) &= 0.03690 + 0.05639 + 0.05639 \\ &\quad + 0.03690 + 0.12500 \\ &= 0.31158 \end{aligned}$$

$$\begin{aligned} P(3; A) &= 0.01098 + 0.03039 + 0.03690 \\ &= 0.07827. \end{aligned}$$

After a permutation, we start with the matrix

$$\begin{array}{c} \begin{array}{ccccc} [4] & [1] & [2] & [3] & [5] \end{array} \\ \begin{array}{c} [4] \\ [1] \\ [2] \\ [3] \\ [5] \end{array} \left[\begin{array}{ccccc} 1 & 0 & 1 & 2 & 0 & -1.5 \\ 0 & 1 & 2 & 1 & 0 & 2.5 \\ 1 & 2 & 9 & 8 & -2 & 2.5 \\ 2 & 1 & 8 & 9 & -2 & -1.5 \\ 0 & 0 & -2 & -2 & 1 & 0.5 \\ -1.5 & 2.5 & 2.5 & -1.5 & 0.5 & 0 \end{array} \right]. \end{array}$$

The first step of sweepout yields

$$\left[\begin{array}{ccccc} 1 & 0 & 1 & 2 & 0 & -1.5 \\ 0 & 1 & 2 & 1 & 0 & 2.5 \\ 0 & 2 & 8 & 6 & -2 & 4 \\ 0 & 1 & 6 & 5 & -2 & 1.5 \\ 0 & 0 & -2 & -2 & 1 & 0.5 \\ 0 & 2.5 & 4 & 1.5 & 0.5 & -2.25 \end{array} \right].$$

This shows that the maximum likelihood estimate is

$$\underline{\hat{x}}' = (2.5, 4, 1.5, 0, 0.5) \quad \text{and} \quad \underline{\hat{y}}' = (2.5, 0, 0.5).$$

After the third step of sweepout, it turns out

$$\begin{bmatrix} 1 & * & * & * & * & * \\ 0 & 1 & * & * & * & * \\ 0 & 0 & 1 & 1 & -0.5 & -0.25 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -8.75 \end{bmatrix}.$$

The value of $\bar{\chi}^2$ statistic can be found by subtraction,

$$\begin{aligned} \bar{\chi}^2 &= 8.7500 - 2.2500 \\ &= 6.5000. \end{aligned}$$

The significance level is

$$\begin{aligned} Pr(\bar{\chi}_2 \geq 6.5) &= P(1; A) \cdot Pr(\chi_1^2 \geq 6.5) + P(2; A) \cdot Pr(\chi_2^2 \geq 6.5) \\ &\quad + P(3; A) \cdot Pr(\chi_3^2 \geq 6.5) \\ &= 0.42173 \times 0.01079 \\ &\quad + 0.31157 \times 0.03877 \\ &\quad + 0.07827 \times 0.08966 \\ &= 0.02364. \end{aligned}$$

§7. Acknowledgements

The generalization of the paper of one of the authors [5] was initiated by Neng-che Yeh, which has been left unpublished because of unnecessary conditions imposed on the theorem. The authors acknowledge the priority and also express their indebtedness to the referee of *Biometrika* for his critical comments when it was returned.

The authors' thanks are to Miss Y. Tanaka for her help in computer programming and Mrs. T. Hukamachi for typewriting.

One of the authors, Jae-Rong Choi, is supported by a scholarship of the Japanese Government on leave of absence from Dong-A University, Pusan, Korea. He is very much grateful to the generosities of both of them.

References

- [1] ANDERSON, T. W.: *An Introduction to Multivariate Statistical Analysis*, New York. John Wiley and Sons Inc. (1958).
- [2] BARLOW, BARTHOLOMEW, BREMNER, and BRUNK: *Statistical Inference under order Restrictions*, New York, John Wiley and Sons Inc. (1972).
- [3] BARTHOLOMEW, D. J.: *A Test of Homogeneity of Means under Restricted Alternatives*, J. R. Statist. Soc. B, **2**, **3** (1961), pp. 239–281.
- [4] KUDÔ, A.: *A Multivariate Analogue of the One Sided Test*, Biometrika, **50**, **3** and **4** (1961), 403–418.
- [5] KUDÔ, A. and NENG-CHE YEH: *A Multivariate Analogue of the One Side Test-(II)*, Unpublished (1968).
- [6] NÜESCH, P. E.: *On the Problem of Testing Location in Multivariate Populations for Restricted Alternatives*, Ann. Math. Statist., **37** (1966), pp. 113–119.
- [7] RAO, C. R.: *Advanced Statistical Methods in Biometric Research*, New York, John Wiley and Sons Inc. (1952).
- [8] ROCKAFELLAR, R. T.: *Convex Analysis*, Princeton Univ. Press (1970).
- [9] TARUMI, T. and KUDÔ, A.: *An Algorithm Related to All Possible Regression and Discriminant Analysis*, Jour. Japan Statist. Soc., **4**, **2** (1974), pp. 47–56.