CHARACTERISATION OF THE DISTRIBUTION OF RANDOM VARIABLES IN LINEAR STRUCTURAL RELATIONS

By C. RADHAKRISHNA RAO

Indian Statistical Institute

SUMMARY. The paper examines the nature of the latent random variables which occur in linear structural models, under the assumption that they are independent. It is shown that alternative representations with different sets of coefficients for the latent variables or different numbers of latent variables are possible when and only when some of the latent variables or some of their linear combinations have a univariate normal distribution. The investigation has applications in the problem of identifiability of parameters in linear structural relations. An important characterization of the multivariate normal distribution is deduced.

1. Introduction

In factor analysis as developed by the psychologists a linear model of the type

$$X = AF+\epsilon \qquad \ldots \quad (1.1)$$

is imposed on a vector random variable X, where F is a vector of random variables called common factors (or variables), A is a fixed matrix of coefficients and ϵ is a vector of random variables called specific factors (or variables). The common and specific variables are all uncorrelated and each column of A is assumed to contain at least two non-zero elements to maintain the distinction between common and specific factors. It can also be assumed, without loss of generality, that no two columns of A are equivalent in the sense that one is a multiple of the other. (If one column of A is a multiple of another we can obtain a reduced representation by omitting one of the columns but associating with the other, a variable which is a suitable linear combination of the variables connected with the original two columns.) We shall, therefore, assume throughout that in a representation such as (1.1) the matrix multiplying random variables does not have any two columns which are equivalent. We shall refer to the model (1.1), with the stated restrictions on the elements of A and on the random variables in F and ϵ, (which may be jointly called latent variables), as a factor analysis model. The principal problem of factor analysis is to determine a matrix A with the minimum possible rank for which a representation such as (1.1) is feasible.

We shall consider an analogous problem by demanding that the common and specific variables are independent and non-degenerate and examine the feasibility of a representation such as (1.1) with minimum rank for A.

It appears, under the hypothesis of independence of the latent variables, that the question of minimum rank is meaningful only in the context of some of the common variables having a normal distribution. For instance, let us suppose that in the representation (1.1) with some matrix A, no element of the vector variable F is normal. Then it is not possible to have an alternative representation with another matrix of a smaller rank, or even with a matrix of the same order but not equivalent to A (two matrices A and B are said to be equivalent if the columns of A are multiples of the columns of B and vice versa). Alternative representations with possible reduction in the rank of A are available only when some of the common variables have a normal distribution. Further, the existence of an alternative representation with increased rank implies that in the original representation some linear functions of the specific variables have normal components (a variable is said to have a normal component if its distribution can be expressed as the convolution of two distributions of which one is normal). In addition some of the common variables may have normal components individually or in certain linear combinations.

We infer these results by examining the conditions under which two non-equivalent representations of the type (1.1) exist. (Two representations of the type (1.1)

$$X = AF + \epsilon_1, \quad X = BG + \epsilon_2$$

are said to be equivalent if the columns of B are all multiples of those of A and vice-versa).

To begin with, we study a more general representation of the form $X = AF$ with no restrictions on A. An important result providing a characterisation of the multivariate normal distribution is as follows. Let $X = AF$ and $X = BG$ be two representations of a p-dimensional random variable X(where F and G are random vectors with independent elements) subject only to the condition that no column of A is a multiple of any column of B. Then X has a p-variate normal distribution. A different proof of this result is given earlier by the author in Rao (1965). Another interesting result is on the class C of non-equivalent representations, $X = AF$. Let C contain a member with rank A equal to the number of columns of A. Then C consists of only a single member or an infinity of members.

2. Preliminary lemmas

We prove some preliminary lemmas which are of general interest.

Lemma 1 : *Let $\alpha_1, \ldots, \alpha_m$ be a given set of non-null vectors in a vector space furnished with an inner product. Then there exists a vector β which is not orthogonal to any one of the set $\alpha_1, \ldots, \alpha_m$.*

Suppose β_0 is not orthogonal to the first k vectors $\alpha_1, \ldots, \alpha_k$ but is orthogonal to α_{k+1}. Then consider $\beta_0 + c\alpha_{k+1}$ where c a constant. Its inner product with $\alpha_i (i \leqslant k)$ is

$$d_i = (\beta_0, \alpha_i) + c(\alpha_{k+1}, \alpha_i) \qquad \ldots \quad (2.1)$$

which may not be zero for any c or which may be zero for $c = c_i$. Then if we choose a value of c to be different from c_i, $i = 1, \ldots, k$ and zero, none of the d_i, $i = 1, \ldots, k$, vanish. Let $\beta_1 = \beta_0 + c\alpha_{k+1}$ for such a choice of c. Then β_1 is not orthogonal to $\alpha_1, \ldots, \alpha_{k+1}$. Since we have an initial choice in $\beta_0 = \alpha_1$ with $k = 1$, the result follows by induction.

Lemma 2 : *Let A be a $p \times k$ matrix and B a $p \times m$ matrix such that the first column (without loss of generality) of A is not a multiple of any other column of A or of any column of B. Then there exists a $2 \times p$ matrix H such that the matrices*

$$C_1 = HA, \quad C_2 = HB \qquad \ldots \quad (2.2)$$

of orders $2 \times k$ and $2 \times m$ respectively satisfy the following property : the first column of C_1 is not a multiple of any other column of C_1 or of any column of C_2.

Without loss of generality, let the elements in the first rows of A and B consist of zeroes and unities only and the element in the first row and column of A be non-zero. Let the columns of A with unity as first element be denoted by $\alpha_1, \ldots, \alpha_r$ and those of B, by $\beta_1, \ldots, \beta_s$.

Let us choose the first row vector of the desired matrix H as $(1, 0, \ldots, 0)$ and the second row as $\gamma' = (0, \gamma_2, \ldots, \gamma_p)$ such that

$$\gamma'(\alpha_1 - \alpha_i) \neq 0, \quad i = 2, \ldots, r$$
$$\gamma'(\alpha_1 - \beta_i) \neq 0, \quad i = 1, \ldots, s. \qquad \ldots \quad (2.3)$$

Such a γ exists by Lemma 1. Then it is easy to verify that H so chosen satisfies the requirements of Lemma 2.

Lemma 3 : *Given two vectors α_1, α_2, which are not multiples of one another, we can find an infinity of pairs of linear combinations $\beta_1 = c_1\alpha_1 + c_2\alpha_2$ and $\beta_2 = d_1\alpha_1 + d_2\alpha_2$ such that $c_1d_1 + c_2d_2 = 0$ and neither β_1 nor β_2 is a multiple of any vector in a given finite set of vectors which may include α_1 and α_2.*

The lemma is easy to establish.

Lemma 4 : *Consider the equation*

$$\psi_1(u+b_1 v) + \psi_2(u+b_2 v) + \ldots + \psi_r(u+b_r v) = A(u) + B(v) + Q(u, v) \qquad \ldots \quad (2.4)$$

in two real variables u and v valid in $|u| < \delta_0$, $|v| < \delta_0$, where Q is a quadratic function. Let

(i) $b_1, \ldots, b_r$ be all different, and

(ii) $\psi_1, \ldots, \psi_r$, A, B be continuous functions.

Then $\psi_1, \ldots, \psi_r$, A and B are all polynomial functions of degree max $(2, r)$ *at most in a neighbourhood of the origin.*

Following Linnik (1960), we multiply both sides of (2.4) by $(x-u)$ and integrate with respect to u from 0 to x where $|x| < \delta_0$. This gives

$$\sum_1^r \int_0^x (x-u)\psi_j(u+b_j v)du$$

$$= \int_0^x (x-u)A(u)du + B(v) \int_0^x (x-u)du + \int_0^x (x-u)Q(u, v)du$$
$$= C(x) + x^2B_1(v) + x^3L(v) \qquad \ldots \quad (2.5)$$

where $L(v)$ is linear in v and $B_1(v)$ is a continuous function of v. Making the transformation $u + b_j v = \tau$ in the j-th term of the left member in (2.5), $|\tau| < \delta_1 < \delta_0$

$$\sum_{j=1}^r \int_{b_j v}^{x+b_j v} (x + b_j v - \tau)\, \psi_j(\tau) d\tau = x^3L(v) + x^2B_1(v) + C(x). \qquad \ldots \quad (2.6)$$

The equality (2.6) is true when $|x| < \delta_2$, $|v| < \delta_2$, $0 < \delta_2 < \delta_1$. The left hand side of (2.6) is differentiable with respect to v for each fixed value of x, $|x| < \delta_2$. This implies that $B_1(v)$ is differentiable with respect to v. Then, differentiating both sides of (2.6) with respect to v we have

$$\Sigma\, b_j \int_0^{x+b_j v} \psi_j'(\tau)d\tau = hx^3 + B_1'(v)x^2 + B_2(v)x + B_3(v) \qquad \ldots \quad (2.7)$$

where h is a constant, $B_2(v)$ and $B_3(v)$ are suitable functions of v and the equality (2.7) is true for a certain domain of v and x.

Differentiating both sides of (2.7) with respect to x and writing u for x

$$\Sigma b_j\, \psi_j'(u+b_j v) = 3hu^2 + 2uB_1'(v) + B_2(v). \qquad \ldots \quad (2.8)$$

Putting $v = 0$ in (2.8) we obtain the equation

$$\Sigma b_j\, \psi_j'(u) = P_{21}(u) \qquad \ldots \quad (2.9)$$

where $P_{21}(u)$ is a polynomial in u of degree 2.

Starting from the equation (2.8) and applying the same analysis as on the equation (2.4) we obtain

$$\Sigma\, b_j^2 \psi_j'(u) = P_{22}(u) \qquad \ldots \quad (2.10)$$

where $P_{22}(u)$ is a polynomial in u of degree 2.

Thus repeating the analysis r times we obtain the equations

$$\begin{aligned} b_1\psi_1'(u) + \cdots + b_r\, \psi_r'(u) &= P_{21}(u) \\ b_1^2\psi_1'(u) + \cdots + b_r^2\, \psi_r'(u) &= P_{22}(u) \\ &\cdots \\ b_1^r\psi_1'(u) + \cdots + b_r^r\, \psi_r'(u) &= P_{rr}(u) \end{aligned} \qquad \ldots \quad (2.11)$$

where $P_{rr}(u)$ is a polynomial in u of degree r. Since $b_1, \ldots, b_m$ are different, the equation (2.11) can be solved uniquely for $\psi_1'(u), \ldots, \psi_r'(u)$. But the solution for $\psi_i'(u)$ is a linear combination of $P_{21}(u), \ldots, P_{rr}(u)$. Hence ψ_i' is a polynomial of degree $\max(2, r)$ at most in a neighbourhood of the origin.

Corollary : *If the equation* (2.4) *of Lemma* 4 *is of the form*

$$\sum_{i=1}^{r} \psi_i(u+b_i\, v) = au+cv+d$$

with $r \leqslant 3$, *then under the conditions of Lemma* 4, ψ_i, $i = 1, \ldots, r$, *are all linear functions.*

Lemma 5 : *Let* Y *be a two-dimensional random variable* (Y_1, Y_2) *admitting two representations*

$$\left.\begin{aligned} Y_1 &= a_{11}f_1 + \cdots + a_{1k}f_k \\ Y_2 &= a_{21}f_1 + \cdots + a_{2k}f_k \end{aligned}\right\} \quad and \quad \left.\begin{aligned} Y_1 &= b_{11}g_1 + \cdots + b_{1m}g_m \\ Y_2 &= b_{21}g_1 + \cdots + b_{2m}g_m \end{aligned}\right\} \qquad \ldots \quad (2.12)$$

where $f_1, \ldots, f_k$ *are independent random variables,* $g_1, \ldots, g_m$ *are independent random variables and* a_{ij}, b_{ij} *are constants. Let the r-th column* $\begin{pmatrix} a_{1r} \\ a_{2r} \end{pmatrix}$ *in the first representation of* (2.12) *be not a multiple of any column of the type* $\begin{pmatrix} a_{1j} \\ a_{2j} \end{pmatrix}$, $j \neq r$ *or of any column of the type* $\begin{pmatrix} b_{1j} \\ b_{2j} \end{pmatrix}$, $j = 1, \ldots, m$. *Then the variable* f_r *has a univariate normal distribution.*

Let ψ_i be the second characteristic (i.e., principal branch of the logarithm of the characteristic function, which is uniquely defined in a neighbourhood of the origin) of f_i, and ϕ_j be that of g_j. Considering the joint characteristic function of Y_1, Y_2 from the two representations, taking logarithms and equating them, we have for $|u|<\delta_0$, $|v|<\delta_0$

$$\begin{aligned}\log [E\exp(i\,u\,Y_1+i\,v\,Y_2)] &= \psi_1(a_{11}u+a_{21}v)+\ldots+\psi_k(a_{1k}u+a_{2k}v)\\ &= \phi_1(b_{11}u+b_{21}v)+\ldots+\phi_m(b_{1m}u+b_{2m}v).\end{aligned} \quad \ldots \quad (2.13)$$

We may, without loss of generality, assume that $a_{1r} \neq 0$ and replace non-zero a_{1i} and b_{1j} by unity, which only means scaling the original variables f_i and g_j. Then using the condition of Lemma 5, the equation (2.13) reduces to

$$\psi_r(u+a_{2r}v)+\eta_1(u+c_{21}v)+\ldots+\eta_s(u+c_{2s}v) = A(u)+B(v) \quad \ldots \quad (2.14)$$

if $a_{2r} \neq 0$, and to

$$\eta_1(u+c_{21}v)+\ldots+\eta_s(u+c_{2s}v) = \psi_r(u)+B(v) \quad \ldots \quad (2.15)$$

if $a_{2r} = 0$. In (2.14) and (2.15) every η function is obtained by adding the ψ functions and subtracting the ϕ functions having a common coefficient for v in (2.13). Observe that $a_{2r}, c_{21}, \ldots, c_{2s}$ can all be taken to be different and the function ψ_r which is the second characteristic of f_r, cannot be combined with any others.

Now, applying Lemma 4, ψ_r is a polynomial of degree at most s in a neighbourhood of the origin. Hence ψ_r must be quadratic if f_r is a non-degenerate random variable. Thus f_r has a univariate normal distribution, and Lemma 5 is established.

Lemma 6 : *Let the i-th column in the first representation of* (2.12) *be a multiple of the j-th column in the second representation of* (2.12) *and not of any other column in either representation, then the second characteristics of f_i and g_j differ by a polynomial in a neighbourhood of the origin.*

We can take $i=j$ without loss of generality and observe that one of the η functions in (2.14) and (2.15) is indeed the difference between the second characteristics of f_i and g_j. But every η function is a polynomial by an application of Lemma 4.

An example showing that the difference of two second characteristics can be a non-quadratic polynomial, due to Dr. B. Ramachandran, is as follows.

Let $f(t) = e^{-|t|-t^4/4}$. It is easy to show that $f(t)$ is convex in $t>0$ and otherwise satisfies Polya's conditions and is thus a c.f. Choosing $g(t) = e^{-|t|}$, which is a c.f., we find

$$f(t)/g(t) = e^{-t^4/4}, \ \log f(t) - \log g(t) = -t^4/4$$

yielding the desired result.

Let us observe that Lemma 6 provides only the relationship in a neighbourhood of the origin between the second characteristics of two random variables, which are themselves, in general, defined only in the neighbourhood of the origin. This result does not, therefore, provide the relationship between the characteristic functions of the two variables over the entire real line but only in an interval round the origin.

3. The main results

Theorem 1 is concerned with representations of the form $X = AF$, with no restrictions on A except that it is in a reduced form (without loss of generality) with no two of its columns being equivalent.

Theorem 1 : *Let $X = AF$ and $X = BG$ be two representations of a p-dimensional random variable X, where A and B are fixed matrices, and, F and G are vectors of independent non-degenerate random variables. We shall take A to be a $p \times r$ matrix and B to be a $p \times s$ matrix without any restrictions on ranks of A and B, and indicate the elements of the vector variable F by $f_1, \ldots, f_r$ and those of G by $g_1, \ldots, g_s$. Then the following are true :*

(i) *The ranks of A and B are the same.*

(ii) *If the i-th column of A is not a multiple of any column of B then f_i has univariate normal distribution.*

(iii) *If the i-th column of A is a multiple of the j-th column of B, then the second characteristics of the variables f_i and g_j differ by a polynomial in a neighbourhood of the origin.*

Proof: Let α be a column vector which is orthogonal to the columns of A. Then

$$\alpha' X = \alpha' AF = 0. \qquad \ldots \quad (3.1)$$

Using the second representation

$$0 = \alpha' X = \alpha' BG. \qquad \ldots \quad (3.2)$$

But G is a vector of non-degenerate independent random variables so that no linear combination of the elements of G is non-degenerate except when all the coefficients are zero. This implies that $\alpha' B = 0$. Thus $\alpha' A = 0 \Longrightarrow \alpha' B = 0$, and vice versa; hence A and B have the same rank, establishing (i).

To prove (ii), let us assume, without loss of generality, that the first column of A is not a multiple of any column of B. Then by Lemma 2, we can find a $2 \times p$ matrix H to give two matrices $C_1 = HA$, $C_2 = HB$ such that the first column of C_1 is not a multiple of any other column of C_1 or of any column of C_2. Now

$$HX = Y = C_1 F \quad \text{and} \quad HX = Y = C_2 G \qquad \ldots \quad (3.3)$$

and applying the result of Lemma 5, we find that f_1 has a univariate normal distribution.

To prove (iii), let us assume, without loss of generality, that the first column of A is the same as the first column of B. Then by Lemma 2, we can find a $2 \times p$ matrix H which provides two matrices, $C_1 = HA$ and $C_2 = HB$ such that the first column of C_1 is not equivalent to any other column of C_1 or any column of C_2 except the first. Now

$$HX = Y = C_1 F,\ HX = Y = C_2 G \qquad \ldots \quad (3.4)$$

and applying Lemma 6, we find that the second characteristics of f_1 and g_1 differ by a polynomial near the origin. We have shown in Section 2, that there exist examples of such second characteristics where the degree of the polynomial can be higher than 2.

Corollary 2 : *If no column of A is a multiple of any column of B then X has a p-variate normal distribution.*

The result follows since each f_i has a univariate normal distribution.

Theorem 2 : *Let C be the class of all non-equivalent representations of a p-dimensional random variable X. (Two representations are equivalent if the columns in the matrix of one representation are multiples of the columns in the other and vice versa). Furthermore, let C have a member of the type $X = AF$ where the rank of A is equal to the number of columns of A. Then C has one member or infinitely many.*

Let $X = BG$ be another representation. If there are two columns of B which are not multiples of any columns of A, then at least two elements of G have a univariate normal distribution. Then making orthogonal transformations on such elements the two columns of B can be altered in infinitely many non-equivalent ways, by lemma 3.

Suppose, there is only one column in B which is not a multiple of any column of A. Then we have the explicit representations (without loss of generality)

$$\left.\begin{array}{l} X_1 = a_{11}f_1+\dots+a_{1k}f_k \\ \dots \quad \dots \quad \dots \\ X_p = a_{p1}f_1+\dots+a_{pk}f_k \end{array}\right\} \quad \left.\begin{array}{l} X_1 = a_{11}g_1+\dots+a_{1k}g_k+a_{1k+1}\,g_{k+1} \\ \dots \quad \dots \quad \dots \\ X_p = a_{p1}g_1+\dots+a_{pk}g_k+a_{pk+1}\,g_{k+1} \end{array}\right\} \quad \dots \quad (3.5)$$

where the rank of the matrix (a_{ij}), $i = 1, \dots, p$, $j = 1, \dots, k$ is k. Since, by assumption the column associated with the variable g_{k+1} is not a multiple of any other column, it follows that g_{k+1} has a univariate normal distribution. Then writing down the joint (first) characteristic function of $(X_1, \dots, X_p)$ in two ways (corresponding to the two representations)

$$\phi_1(a_{11}t_1+\dots+a_{p1}t_p) \dots \phi_k(a_{1k}t_1+\dots+a_{pk}t_p) = \psi_1(a_{11}t_1+\dots+a_{p1}t_p) \dots \psi_k(a_{1k}t_1+\dots+a_{pk}t_p)$$
$$\times \exp[-\lambda(a_{1k+1}t_1+\dots+a_{pk+1}t_p)^2] \quad \dots \quad (3.6)$$

where ϕ_i is the characteristic function of f_i, ψ_i that of g_i, and λ is the variance of g_{k+1} whose expectation is taken as zero without loss of generality. Making the transformation

$$\tau_i = a_{1i}t_1+\dots+a_{pi}t_p, \quad i = 1, \dots, k \qquad \dots \quad (3.7)$$

the equation (3.6) reduces to

$$\phi(\tau_1) \dots \phi_k(\tau_k) = \psi_1(\tau_1) \dots \psi_k(\tau_k)[\exp(-d_1\tau_1-\dots-d_k\tau_k)^2] \qquad \dots \quad (3.8)$$

since the column of coefficients $(a_{1k+1}, \dots, a_{pk+1})$ must depend on the other columns, by result (i) of Theorem 1. Putting $\tau_j = 0$, $j = 1, \dots, i-1, i+1, \dots, k$ in (3.8),

$$\phi_i(\tau_i) = \psi_i(\tau_i)[\exp(-d_i^2\tau_i^2)] \quad i = 1, \dots, k. \qquad \dots \quad (3.9)$$

Then it follows that the variable f_i has a normal component which is non-degenerate if $d_i \neq 0$. At least two of the d_i are not zero, for otherwise it would mean that the column of coefficients of g_{k+1} is a multiple of some other column. *Thus at least two of the f_i, say f_1 and f_2, have non-degenerate normal components.*

Writing $f_1 = g_1+z_1$, $f_2 = g_2+z_2$, $f_3 = g_3, \ldots, f_k = g_k$ we have the representation

$$X_1 = a_{11}g_1+\ldots+a_{1k}g_k+a_{11}z_1+a_{12}z_2, \qquad \ldots \qquad \ldots \qquad X_p = a_{p1}g_1+\ldots+a_{pk}g_k+a_{p1}z_1+a_{p2}z_2 \quad \ldots (3.10)$$

derived from the first representation of (3.5). By making orthogonal transformations on the independent normal variables z_1, z_2 we can obtain representations where the last two columns are not multiples of the other columns, by Lemma 3. Thus an infinity of non-equivalent representations can be obtained.

Note that the case where the number of columns in the two representations of (3.5) are the same but only $(k-1)$ of the columns are equivalent can be ruled out.

We have shown that if there exists a second member in C other than the type assumed, then there must be an infinity of them. On the other hand, there may not be any second member.

The third theorem concerns factor analysis models. Let

$$X = AF+\epsilon_1 \quad \text{and} \quad X = BG+\epsilon_2 \quad \ldots (3.11)$$

be two representations where A is a $p\times r$ matrix of rank m, B is a $p\times s$ matrix of rank n, and ϵ_1 and ϵ_2 are vectors of specific variables (factors) and F and G represent common factors. We represent the elements of F by $f_1, \ldots, f_r$, of G by $g_1, \ldots, g_s$, of ϵ_1 by $e_1, \ldots, e_p$ and of ϵ_2 by $e'_1, \ldots, e'_p$. With these notations we have Theorem 3.

Theorem 3 : *Let $m < n$. Then*

(i) *at least $(n-m)$ of the common variables in the second representation of (3.11) are normally distributed,*

(ii) *there are linear combinations $\gamma'\epsilon_1$, $\gamma'\epsilon_2$ of the specific variables which differ by a non-degenerate normal component, and*

(iii) *the second characteristics of at least one of the pairs (e_i, e'_i), $i = 1, \ldots, p$ differ by a second degree polynomial in a neighbourhood of the origin.*

Proof: The condition $m < n$ implies that there are at least $(n-m)$ columns of B which are not multiples of columns of A. Hence by (ii) of Theorem 1, the associated random variables are normal, which proves the assertion (i) of Theorem 3.

To prove (ii), let $g_i, g_j \ldots$ be the random variables such that the associated columns of B are linearly independent of the columns of A. Let us choose a vector $\gamma' = (\gamma_1, \ldots, \gamma_p)$ such that it is orthogonal to all the columns of A but not to every column of B. Let

$$Y = \gamma'X = \gamma'AF+\gamma'\epsilon_1 = \gamma'\epsilon_1 \quad \ldots (3.12)$$

$$Y = \gamma'X = \gamma'BG+\gamma'\epsilon_2 = c_ig_i+c_jg_j+\ldots+\gamma'\epsilon_2 \quad \ldots (3.13)$$

where $c_i, c_j, \ldots$ are constants, *not all zero*, and $g_i, g_j, \ldots$ are all necessarily normal variables by virtue of (ii) of Theorem 1. The equations (3.12) and (3.13) show that $\gamma'\epsilon_1$ and $\gamma'\epsilon_2$ differ by a normal component which proves (ii) of Theorem 3.

To prove (iii) consider an element, say X_q of X for which $e_q \neq 0$. Writing X_q in two ways,

$$X_q = a_{q1}f_1 + \ldots + a_{qr}f_r + e_q \qquad \ldots \quad (3.14)$$

$$X_q = b_{q1}g_1 + \ldots + b_{qs}g_s + e'_q \qquad \ldots \quad (3.15)$$

the joint second characteristic of (X_q, Y) can again be written in two ways by considering the representations in the pairs of equations (3.12), (3.14) and (3.13), (3.15)

$$\log E[\exp(iuX_q + ivY)]$$

$$= \sum_{j=1}^{r} \phi_j(a_{j1}\, u) + \xi_1(u + \gamma_q v) + \eta_1(v) \qquad \ldots \quad (3.16)$$

$$= \psi_1(b_{q1}\, u) + \ldots + \psi_i(b_{qi}u + c_i v) + \ldots + \xi_2(u + \gamma_q v) + \eta_2(v) \qquad \ldots \quad (3.17)$$

where $\phi_j, \psi_k, \xi_1, \xi_2, \eta_1$ and η_2 are the second characteristics of random variables $f_j, g_k, e_q, e'_q, \gamma'\epsilon_1 - \gamma_q e_q$, and $\gamma'\epsilon_2 - \gamma_q e'_q$ respectively. Observing that $\psi_i, \psi_j, \ldots$ are all quadratic, we obtain on equating (3.16) and (3.17) that

$$\xi_1(u + \gamma_q v) - \xi_2(u + \gamma_q v) = Q(u, v) + A(u) + B(v) \qquad \ldots \quad (3.18)$$

where Q is quadratic and $A(u)$ and $B(v)$ are suitable functions. Then, applying Lemma 4, $\xi_1(u) \cdot \xi_2(u)$ is a polynomial of degree two at most in a neighbourhood of the origin. We proceed to show that for at least one of the pairs (e_q, e'_q), $1 < q < p$ the second characteristics differ exactly by a second degree polynomial in a neighbourhood of the origin.

Suppose the second characteristics of all the pairs (e_q, e'_q) for which $\gamma_q \neq 0$ differ locally by a linear function. Taking the second characteristics of Y in two ways corresponding to the representations (3.12), (3.13) and equating them, we find that the second characteristic of $c_i g_i + c_j g_j + \ldots$ is a linear function, which cannot be, since $c_i, c_j, \ldots$ are not all zero and $g_i, g_j \ldots$ are all non-degenerate random variables by assumption. Hence the assertion (iii) of Theorem 3 follows.

Theorem 4 : *Let k be the maximal integer such that k columns of A are multiples of k columns of B. Without loss of generality, we shall take these to be the first k columns of A and the corresponding columns of B. If the first k columns of A (and hence those of B) are independent, then the second characteristics of each pair (e_i, e'_i), $i = 1, \ldots, p$ differ at most by a polynomial of degree two in the neighbourhood of the origin and further the second characteristics of each pair (f_i, g_i), $i = 1, \ldots, k$ differ at most by a polynomial of degree two in a neighbourhood of the origin.*

Theorem 4 is easy to establish using the arguments of the previous theorems and lemmas. It is not possible, in general, to assert that (e_i, e'_i) or (f_i, g_i) differ by a normal component since the results of Theorem 4 provide only a local relationship between the characteristic functions, (i.e., the ratio of the characteristic functions may have a normal factor locally but not over the entire real line). An example due to Dr. B. Ramachandran clarifies the point.

Consider the functions

$$C_1(t) = \begin{cases} e^{-2|t|-t^2/2} & \text{for } |t| < \delta \\ h_1 e^{-|t|} & \text{for } |t| > \delta \end{cases} \quad \ldots \quad (3.19)$$

$$C_2(t) = \begin{cases} e^{-2|t|} & \text{for } |t| < \delta \\ h_2 e^{-|t|} & \text{for } |t| > \delta \end{cases} \quad \ldots \quad (3.20)$$

where $$h_1 = e^{-\delta-\delta^2/2}, \quad h_2 = e^{-\delta}.$$

Thus it is easy to see that $C_1(t)$ and $C_2(t)$ are characteristic functions and

$$\log C_1(t) - \log C_2(t_2) = -t^2/2 \text{ for } |t| < \delta.$$

But

$$\frac{C_1(t)}{C_2(t)} = \begin{cases} e^{-t^2/2} & \text{for } |t| < \delta \\ h_1/h_2 & \text{for } |t| > \delta \end{cases}$$

so that $C_1(t)$ does not differ from $C_2(t)$ by a normal factor.

The results of this paper can be used in making statements on identifiability of parameters in a linear structural relation, a problem considered by Reiersol (1950). For instance let (X_1, X_2) have the two representations

$$\left.\begin{aligned} X_1 &= f+e_1, & X_1 &= g+e_1' \\ X_2 &= f+e_2, & X_2 &= ag+e_2' \end{aligned}\right\} \quad \ldots \quad (3.21)$$

where $a \neq 0, \neq 1$, f, e_1, e_2 are independent variables and g, e_1', e_2' are independent variables. Theorem 1 shows that $a \neq 0, \neq 1$ implies that f and g are normally distributed. Then it follows that (e_1, e_1') differ by a normal component and so also (e_2, e_2'). If f is not normal and not degenerate, then a must have value 1.

In conclusion I have great pleasure in thanking Dr. B. Ramachandran for the useful discussions I had with him during the preparation of this paper.

References

LINNIK, YU. V. (1960): *Decomposition of Probability Distributions*, English translation, 1964, Oliver and Boyd.

RADHAKRISHNA RAO, C. (1965): *Linear Statistical Inference and Its Applications*, John Wiley and Sons.

REIERSOL, O. (1950): Identifiability of a linear relation between variables which are subject to error. *Econometrica*, 18, 375-380.

Paper received : May, 1966.