The Annals of Mathematical Statistics
1972, Vol. 43, No. 1, 205-229

CHARACTERIZATIONS OF THE MULTIVARIATE NORMAL DISTRIBUTION USING REGRESSION PROPERTIES[1]

BY F. S. GORDON[2] AND A. M. MATHAI

C. W. Post College and McGill University

1. Introduction. In a previous paper [2], the authors obtained a series of characterizations for many univariate distributions using the property of cubic regression for a cubic statistic on a linear one. These characterizations were obtained by solving a differential equation in the characteristic function of an arbitrary population.

In order to obtain characterizations of multivariate distributions analogous to those determined in the univariate case, it is necessary to utilize a concept for the derivative of any quantity, be it scalar, vector or matrix, with respect to the vector variable $t = [t_1, \cdots, t_p]$. This derivative is similar to that presented, for example, by Wedderburn [7]. On the basis of several properties which this derivative possesses, a vector differential equation is obtained in the characteristic function using the assumption of cubic regression for a cubic statistic on a linear one. For appropriate conditions on the coefficients of this equation, a series of characterizations for the multivariate normal distribution is obtained within the class of those populations whose characteristic functions can be expressed in a certain type of infinite series expansion. This expansion is one in the vector variable t and its transpose t', with coefficients which are determined in terms of the vector derivatives of the characteristic function. In particular, in the univariate case ($p = 1$), it is shown that this infinite series reduces to the usual Taylor series expansion about the origin and therefore, the corresponding univariate class of populations includes all those whose characteristic functions are analytic.

Throughout the present work, the symbol 0 is used to represent the scalar zero element, the zero vector and the null matrix. The particular usage is always clear from the context.

2. Generalized differentiation with respect to a vector variable. We begin by defining the derivative of any scalar, any $p \times 1$ vector, or any $p \times q$ matrix with respect to the vector variable

$$t = [t_1, t_2, \cdots, t_p]\,.$$

Received June 24, 1968.

[1] This paper is adapted from a portion of the first author's doctoral dissertation written under the supervision of Professor A. M. Mathai and submitted to the Faculty of Graduate Studies and Research, McGill University, 1968.

[2] The author wishes to acknowledge support of the National Research Council of Canada.

For this purpose, we introduce the vector differential operator

$$\nabla = \left[\frac{\partial}{\partial t_1}, \frac{\partial}{\partial t_2}, \cdots, \frac{\partial}{\partial t_p}\right]$$

and define the derivative of any quantity X as ∇X, where the product is taken in the sense of the usual matrix product. Thus, for a scalar $f(t)$, we have

$$\nabla f(t) = \left[\frac{\partial f(t)}{\partial t_1}, \frac{\partial f(t)}{\partial t_2}, \cdots, \frac{\partial f(t)}{\partial t_p}\right],$$

a $1 \times p$ vector; for a $1 \times p$ vector $Y = [y_1, \cdots, y_p]$, we have

$$\nabla Y' = \left[\frac{\partial}{\partial t_1}, \frac{\partial}{\partial t_2}, \cdots, \frac{\partial}{\partial t_p}\right] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} = \sum_{j=1}^{p} \frac{\partial y_j}{\partial t_j},$$

a scalar; for a $p \times q$ matrix $X = [x_{jk}]$,

$$\nabla X = \left[\sum_{j=1}^{p} \frac{\partial x_{j1}}{\partial t_j}, \sum_{j=1}^{p} \frac{\partial x_{j2}}{\partial t_j}, \cdots, \sum_{j=1}^{p} \frac{\partial x_{jq}}{\partial t_j}\right],$$

a $1 \times q$ vector.

In addition we also introduce the transpose of this operator, $(\nabla)' = \nabla'$, a column differential operator. Using this transpose operator, we are able to consider

$$\nabla' Y = \begin{bmatrix} \frac{\partial}{\partial t_1} \\ \vdots \\ \frac{\partial}{\partial t_p} \end{bmatrix} [y_1, y_2, \cdots, y_p] = \begin{bmatrix} \frac{\partial y_1}{\partial t_1} & \cdots & \frac{\partial y_p}{\partial t_1} \\ \vdots & & \\ \frac{\partial y_1}{\partial t_p} & \cdots & \frac{\partial y_p}{\partial t_p} \end{bmatrix}.$$

We can form also such higher order combinations as $\nabla\nabla'(\cdot)$, $\nabla'\nabla(\cdot)$, $\nabla\nabla'\nabla(\cdot)$, and so forth, by alternating ∇ and ∇' to maintain consistency in size so that the respective multiplications exist.

3. Vector derivatives of multivariate characteristic functions. We now apply the preceding general theory to the characteristic function of a stochastic row vector

$$X = [x_1, x_2, \cdots, x_p],$$

namely,

$$h(t) = E(e^{itX'}),$$

where t is a real $1 \times p$ vector. For the various derivatives of the characteristic function, we readily obtain

$$\text{(1A)} \qquad \begin{aligned} \nabla h &= i[E(x_1 e^{itX'}), E(x_2 e^{itX'}), \cdots, E(x_p e^{itX'})] \\ &= iE(Xe^{itX'}), \end{aligned}$$

and, similarly,

$$\nabla' h = iE(X' e^{itX'}) . \tag{1B}$$

Moreover,

$$\nabla'\nabla h = i\nabla'[E(x_1 e^{itX'}), E(x_2 e^{itX'}), \cdots, E(x_p e^{itX'})]$$

$$= i^2 \begin{bmatrix} E(x_1^2 e^{itX'}) & E(x_1 x_2 e^{itX'}) \cdots E(x_1 x_p e^{itX'}) \\ E(x_2 x_1 e^{itX'}) & \\ \vdots & \\ E(x_p x_1 e^{itX'}) & E(x_p x_2 e^{itX'}) \cdots E(x_p^2 e^{itX'}) \end{bmatrix} \tag{1C}$$

$$= - E(X' X e^{itX'})$$

and similarly,

$$\nabla\nabla' h = i^2 E(\textstyle\sum_{j=1}^{p} x_j^2 e^{itX'}) \tag{1D}$$

$$= - E(XX' e^{itX'}) .$$

In the same manner, we can continue by applying the appropriate operator to obtain

$$\nabla\nabla'\nabla h = - iE(XX'Xe^{itX'}) \tag{1E}$$

and also

$$\nabla'\nabla\nabla' h = - iE(X'XX'e^{itX'}) . \tag{1F}$$

For future reference, rather than singling out particular groupings of the above set of six equations, we shall refer to them as equations (1).

4. Some properties of vector polynomial regression. In the ensuing theory for cubic regression, we shall require the following definition and theorem.

DEFINITION. We say that the stochastic $1 \times p$ row vector X has odd polynomial regression of order $2m + 1$ on the stochastic $1 \times p$ vector Y if

$$E(X/Y) = \textstyle\sum_{j=0}^{m} \beta_{2j+1}(YY')^j Y \qquad \text{a.e.} \tag{2}$$

where m is a nonnegative integer and the β's are real constants, provided that the conditional moment $E(X/Y) = E(X/Y = y)$ exists.

THEOREM 1. *Let X and Y be two $1 \times p$ stochastic vectors and assume that the expectations $E(X)$ and $E[(YY')^r Y]$, $r \leqslant m$, exist, for m a nonnegative integer. Then X has odd polynomial regression of order $2m + 1$ on Y if, and only if, the relation*

$$E(Xe^{itY'}) = \textstyle\sum_{j=0}^{m} \beta_{2j+1} E[(YY')^j Y e^{itY'}]$$

holds for every real $1 \times p$ vector t.

PROOF. Suppose that X has odd polynomial regression of order $2m + 1$ on Y.

We make use of the well-known property

$$E(U) = E[E(U/V)] \tag{3}$$

for all U and V to say that

$$E(Xe^{itY'}) = E[E(Xe^{itY'}/Y)] .$$

Now since $e^{itY'}$ is a scalar, once the value of Y is given, the above expression can be written as

$$E(Xe^{itY'}) = E[e^{itY'}E(X/Y)] .$$

However, using (2), the latter is equivalent to

$$\begin{aligned} E(Xe^{itY'}) &= E[e^{itY'} \textstyle\sum_{j=0}^{m} \beta_{2j+1}(YY')^j Y] \\ &= \textstyle\sum_{j=0}^{m} \beta_{2j+1} E(YY')^j Ye^{itY'}] . \end{aligned}$$

Conversely, if we are given the above equality, we equivalently have

$$E[e^{itY'}(X - \textstyle\sum_{j=0}^{m} \beta_{2j+1}(YY')^j Y)] = 0 .$$

Hence, we again use the result of (3), to say

$$E[E\{e^{itY'}(X - \textstyle\sum_{j=0}^{m} \beta_{2j+1} (YY')^j Y)\ Y\}] = 0 .$$

But, taking the integral in the Lebesgue-Stieltjes sense, this is the same as

$$\int e^{itY'} E[X - \textstyle\sum_{j=0}^{m} \beta_{2j+1}(YY')^j Y/Y] dF = 0 ,$$

where F is the distribution function of the stochastic vector Y. However, since by assumption, all the elements in the row vector $E[X - \sum_{j=0}^{m} \beta_{2j+1}(YY')^j Y/Y]$ are bounded, we take the Fourier transform applied to functions of bounded variation to obtain

$$E[X - \textstyle\sum_{j=0}^{m} \beta_{2j+1}(YY')^j Y/Y] = 0 \qquad \text{a.e.}$$

which, when evaluated, gives

$$E(X/Y) = \textstyle\sum_{j=0}^{m} \beta_{2j+1}(YY')^j Y \qquad \text{a.e.}$$

as required.

Note that we have considered only the odd powers in the polynomial expression in order to maintain consistency in size. The even power terms will be considered separately in a later section as a problem in quadratic regression.

A necessary and sufficient condition for the existence of odd polynomial regression of order $2m + 1$ for the stochastic vector X on the stochastic vector Y can also be obtained in terms of the function

$$f(t_1, t_2) = E\{\exp[i(t_1 X' + t_2 Y')]\} ,$$

where

$$t_1 = [t_{11}, t_{12}, \cdots, t_{1p}] ,$$
$$t_2 = [t_{21}, t_{22}, \cdots, t_{2p}]$$

are $1 \times p$ row vectors of arbitrary real constants. We shall find it convenient to let ∇_1 and ∇_2 represent the differential operators corresponding to t_1 and t_2, respectively.

THEOREM 2. *Under the conditions of Theorem* 1, *a necessary and sufficient condition for X to have odd polynomial regression of order $2m+1$ on Y is that*

$$\nabla_1 f(t_1, t_2)|_{t_1=0} = \textstyle\sum_{j=0}^{m} (-1)^j \beta_{2j+1} (\nabla_2 \nabla_2^{1})^j \nabla_2 f(0, t_2) . \tag{4}$$

PROOF. We have

$$\nabla_1 f(t_1, t_2) = i\{E[x_1 \exp(i(t_1 X' + t_2 Y'))], \cdots, E[x_p \exp(i(t_1 X' + t_2 Y'))]\} ,$$

so that, evaluating this expression at $t_1 = 0$, we find

$$\nabla_1 f(t_1, t_2)|_{t_1=0} = iE(Xe^{it_2 Y'}) . \tag{5}$$

In order to evaluate the right-hand side of (4), we first consider

$$\begin{aligned}\nabla_2 f(0, t_2) &= i[E(y_1 e^{it_2 Y'}), E(y_2 e^{it_2 Y'}), \cdots, E(y_p e^{it_2 Y'})] \\ &= iE(Ye^{it_2 Y'}) .\end{aligned}$$

If we further apply ∇_2' to this equation, we have

$$\nabla_2' \nabla_2 f(0, t_2) = -E(Y'Ye^{it_2 Y'}) .$$

Proceeding in this manner, successively applying first ∇_2 and then ∇_2' to $f(0, t_2)$, we finally obtain

$$(\nabla_2 \nabla_2')^j \nabla_2 f(0, t_2) = (-1)^j iE[(YY')^j Ye^{it_2 Y'}] .$$

Hence, since this statement is true for all integral j, we can evaluate the right-hand side of (4) as

$$\textstyle\sum_{j=0}^{m} (-1)^j \beta_{2j+1} (\nabla_2 \nabla_2')^j \nabla_2 f(0, t_2) = i \sum_{j=0}^{m} \beta_{2j+1} E[(YY')^j Ye^{it_2 Y'}] . \tag{6}$$

However, from Theorem 1, we know that a necessary and sufficient condition for X to have odd polynomial regression of order $2m+1$ on Y is

$$E(Xe^{itY'}) = \textstyle\sum_{j=0}^{m} \beta_{2j+1} E[(YY')^j Ye^{itY'}] .$$

Comparing with (5) and (6), we see that this last statement is equivalent to

$$\nabla_1 f(t_1, t_2)|_{t_1=0} = \textstyle\sum_{j=0}^{m} (-1)^j \beta_{2j+1} (\nabla_2 \nabla_2')^j \nabla_2 f(0, t_2) ,$$

which proves the theorem.

5. Derivation of the fundamental vector differential equation. Suppose that we are given a set of n stochastic $1 \times p$ row vectors $X_1, \cdots, X_n$ which are independently and identically distributed. We begin our study of the cubic statistic

$$S = \textstyle\sum_{j,k,m} a_{jkm} X_j X_k' X_m + \sum_j c_j X_j ,$$

a stochastic $1 \times p$ row vector, where a_{jkm} and c_j (for all $j, k, m = 1, \cdots, n$)

are real constants. We assume the statistic S has regression on the linear statistic

$$L = \sum_j X_j ,$$

a stochastic $1 \times p$ vector, of the cubic form

$$E(S/L) = \beta_1 L + \beta_3 LL'L \qquad \text{a.e.} \tag{7}$$

where β_1 and β_3 are real constants.

Using Theorem 1, (7) holds if, and only if,

$$E(Se^{itL'}) = \beta_1 E(Le^{itL'}) + \beta_3 E(LL'Le^{itL'}) . \tag{8}$$

We first derive the differential equation which results from the cubic regression assumption in terms of the characteristic function of the population.

To begin, we see that the left-hand side of (8) becomes

$$E(Se^{itL'}) = \sum_{j,k,m} a_{jkm} E(X_j X_k' X_m e^{itL'}) + \sum_j c_j E(X_j e^{itL'}) .$$

Now, since the $X_j (j = 1, \cdots, n)$ are independent stochastic vectors, this reduces to

$$\begin{aligned} E(Se^{itL'}) = &\sum_j a_{jjj} E(X_j X_j' X_j e^{itX_j'}) E[\exp i \textstyle\sum_{k \neq j} tX_k'] \\ &+ \textstyle\sum_{j \neq k} a_{jjk} E(X_j X_j' e^{itX_j'}) E(X_k e^{itX_k'}) E[\exp i \textstyle\sum_{m \neq j,k} tX_m'] \\ &+ \textstyle\sum_{j \neq k} a_{kjj} E(X_k e^{itX_k'}) E(X_j' X_j e^{itX_j'}) E[\exp i \textstyle\sum_{m \neq j,k} tX_m'] \\ &+ \textstyle\sum_{j \neq k} a_{jkj} E(X_j X_k' X_j e^{itL'}) \\ &+ \textstyle\sum_{j \neq k \neq m} a_{jkm} E(X_j e^{itX_j'}) E(X_k' e^{itX_k'}) E(X_m e^{itX_m'}) \\ &\qquad \times E[\exp i \textstyle\sum_{r \neq j,k,m} tX_r'] \\ &+ \textstyle\sum_j c_j E(X_j e^{itX_j'}) E(\exp i \textstyle\sum_{k \neq j} tX_k') . \end{aligned}$$

However, because the $X_j (j = 1, \cdots, n)$ are also identically distributed, we can make use of the results of (1) to obtain the simplification

$$\begin{aligned} E(Se^{itL'}) = &\textstyle\sum_j a_{jjj} i(\nabla\nabla'\nabla h) h^{n-1} + \textstyle\sum_{j \neq k} a_{jjk} (-\nabla\nabla' h)(-i\nabla h) h^{n-2} \\ &+ \textstyle\sum_{j \neq k} a_{kjj} (-i\nabla h)(-\nabla'\nabla h) h^{n-2} + Z \\ &+ \textstyle\sum_{j \neq k \neq m} a_{jkm} (-i\nabla h)(-i\nabla' h)(-i\nabla h) h^{n-3} \\ &+ \textstyle\sum_j c_j (-i\nabla h) h^{n-1} \\ = &\; iA_1 (\nabla\nabla'\nabla h) h^{n-1} + iA_2 (\nabla\nabla' h)(\nabla h) h^{n-2} \\ &+ iA_3 (\nabla h)(\nabla'\nabla h) h^{n-2} + Z \\ &+ iA_4 (\nabla h)(\nabla' h)(\nabla h) h^{n-3} - iC(\nabla h) h^{n-1} , \end{aligned}$$

where

$$Z = \textstyle\sum_{j \neq k} a_{jkj} E(X_j X_k' X_j e^{itL'}) ,$$

$$A_1 = \textstyle\sum_j a_{jjj} , \qquad A_2 = \textstyle\sum_{j \neq k} a_{jjk} ,$$

$$A_3 = \textstyle\sum_{j \neq k} a_{kjj} , \qquad A_4 = \textstyle\sum_{j \neq k \neq m} a_{jkm} ,$$

$$C = \textstyle\sum_j c_j .$$

We now consider the right-hand side of (8) and, again introducing the characteristic function and its derivatives, we obtain

$$\begin{aligned}
&\beta_1 E(Le^{itL'}) + \beta_3 E(LL'Le^{itL'}) \\
&\quad = \beta_1 \textstyle\sum_j E(X_j e^{itX_j'})h^{n-1} + \beta_3 \sum_j E(X_j X_j' X_j e^{itX_j'})h^{n-1} \\
&\qquad + \beta_3 \textstyle\sum_{j\neq k} E(X_j X_j' e^{itX_j'})E(X_k e^{itX_k'})h^{n-2} \\
&\qquad + \beta_3 \textstyle\sum_{j\neq k} E(X_k e^{itX_k'})E(X_j' X_j e^{itX_j'})h^{n-2} \\
&\qquad + \beta_3 \textstyle\sum_{j\neq k} E(X_j X_k' X_j e^{itL'}) \\
&\qquad + \beta_3 \textstyle\sum_{j\neq k\neq m} E(X_j e^{itX_j'})E(X_k' e^{itX_k'})E(X_m e^{itX_m'})h^{n-3} \\
&\quad = -in\beta_1(\nabla h)h^{n-1} + in\beta_3(\nabla\nabla'\nabla h)h^{n-1} \\
&\qquad + in(n-1)\beta_3(\nabla\nabla' h)(\nabla h)h^{n-2} + in(n-1)\beta_3(\nabla h)(\nabla'\nabla h)h^{n-2} \\
&\qquad + W + in(n-1)(n-2)\beta_3(\nabla h)(\nabla' h)(\nabla h)h^{n-3} ,
\end{aligned}$$

where

$$W = \textstyle\sum_{j\neq k} \beta_3 E(X_j X_k' X_j e^{itL'}) .$$

Substituting these results into (8) while simultaneously dividing each expression by h^n for convenience, we easily obtain

$$\begin{aligned}
&i(n\beta_3 - A_1)(\nabla\nabla'\nabla h)/h + i[n(n-1)\beta_3 - A_2][(\nabla\nabla' h)/h][(\nabla h)/h] \\
&\qquad + i[n(n-1)\beta_3 - A_3][(\nabla h)/h][(\nabla'\nabla h)/h] \\
&\qquad + i[n(n-1)(n-2)\beta_3 - A_4][(\nabla h)/h][(\nabla' h)/h][(\nabla h)/h] \\
&\qquad + i(C - n\beta_1)(\nabla h)/h + (W - Z)/h^n = 0 .
\end{aligned}$$

We now impose the condition $a_{jkj} = \beta_3$ (for all $j, k = 1, \cdots, n$) in order that the term $(W - Z)/h^n$ disappears. The above equation therefore becomes

$$\begin{aligned}
&(n\beta_3 - A_1)(\nabla\nabla'\nabla h)/h + [n(n-1)\beta_3 - A_2][(\nabla\nabla' h)/h][(\nabla h)/h] \\
&\qquad + [n(n-1)\beta_3 - A_3][(\nabla h)/h][(\nabla'\nabla h)/h] \\
&\qquad + [n(n-1)(n-2)\beta_3 - A_4][(\nabla h)/h][(\nabla' h)/h][(\nabla h)/h] \\
&\qquad + (C - n\beta_1)(\nabla h)/h = 0 .
\end{aligned} \tag{9}$$

Before attempting to solve this differential equation, it is convenient to introduce the transformation

$$h = e^g \quad \text{or} \quad g = \log(h) .$$

It is easy to see that the first derivatives of h can therefore be written in terms of g as

$$\nabla h = e^g \nabla g \qquad \nabla' h = e^g \nabla' g .$$

Applying the appropriate operator, it is also possible to evaluate the higher order derivatives of h directly from the definition of ∇, in terms of g. However, the calculations are very tedious and can be eliminated by recourse to a series of five lemmas which we now discuss. The first four of these lemmas can be verified by direct calculation. The proof of the fifth is not so apparent and, accordingly, we outline the details.

LEMMA 1. *If $f(t)$ is any scalar function of the $1 \times p$ vector $t = [t_1, \cdots, t_p]$ and if $Y(t) = [y_1, \cdots, y_p]$ is any $1 \times p$ vector function of t, then*

$$\nabla(fY') = (\nabla f)Y' + f(\nabla Y') .$$

LEMMA 2. *If $f(t)$ is any scalar function of the $1 \times p$ vector t and if $Y(t) = [y, \cdots, y_p]$ is any $1 \times p$ vector function of t, then*

$$\nabla'(fY) = (\nabla' f)Y + f(\nabla' Y) .$$

COROLLARY. *If $f(t)$ and $g(t)$ are any two scalar functions of the $1 \times p$ vector t, then*

$$\nabla'(fg) = (\nabla' f)g + f(\nabla' g) .$$

LEMMA 3. *If $f(t)$ is any scalar function of the $1 \times p$ vector t and if $A(t) = [a_{ij}(t)]$ is any $p \times p$ matrix function of t, then*

$$\nabla(fA) = (\nabla f)A + f(\nabla A) .$$

LEMMA 4. *If $Y(t)$ is any $1 \times p$ vector function of the $1 \times p$ vector t, then*

$$\nabla'(YY') = 2(\nabla' Y)Y' .$$

LEMMA 5. *If $Y(t)$ is any $1 \times p$ vector function of the $1 \times p$ vector t, then*

$$\nabla(Y'Y) = (\nabla Y')Y + Y(\nabla' Y) .$$

PROOF. Consider

$$\begin{aligned}
\nabla(Y'Y) &= \left[\frac{\partial}{\partial t_1}, \frac{\partial}{\partial t_2}, \cdots, \frac{\partial}{\partial t_p}\right] \begin{bmatrix} y_1^2 & y_1 y_2 \cdots y_1 y_p \\ y_2 y_1 & y_2^2 \cdots y_2 y_p \\ \vdots & \\ y_p y_1 & y_p y_2 \cdots y_p^2 \end{bmatrix} \\
&= \Big[2y_1 \frac{\partial y_1}{\partial t_1} + \textstyle\sum_{j\neq 1} \frac{\partial y_j}{\partial t_j} y_1 + \sum_{j \neq 1} y_j \frac{\partial y_1}{\partial t_j}, \cdots, \sum_{j \neq p} \frac{\partial y_j}{\partial t_j} y_p \\
&\qquad + \textstyle\sum_{j \neq p} y_j \frac{\partial y_p}{\partial t_j} + 2y_p \frac{\partial y_p}{\partial t_p} \Big] \\
&= \Big[\textstyle\sum_j \frac{\partial y_j}{\partial t_j} y_1, \cdots, \sum_j \frac{\partial y_j}{\partial t_j} y_p \Big] \\
&\qquad + \Big[\textstyle\sum_j y_j \frac{\partial y_1}{\partial t_j}, \cdots, \sum_j y_j \frac{\partial y_p}{\partial t_j} \Big] \\
&= (\nabla Y')Y + Y(\nabla' Y) .
\end{aligned}$$

With the help of these five lemmas, it is now a simple matter to evaluate the higher order derivatives of h in terms of $g(=\log h)$. We have already seen that

$$\nabla h = e^g \nabla g \,, \tag{10}$$

$$\nabla' h = e^g \nabla' g \,. \tag{11}$$

Now, applying Lemma 2, to (10), we see that

$$\nabla'\nabla h = e^g(\nabla' g)(\nabla g) + e^g(\nabla'\nabla g) \,. \tag{12}$$

Similarly, applying Lemma 1 to (11), we find

$$\nabla\nabla' h = e^g(\nabla g)(\nabla' g) + e^g(\nabla\nabla' g) \,. \tag{13}$$

Further, using Lemma 3 on (12), we obtain

$$\nabla\nabla'\nabla h = e^g(\nabla g)(\nabla' g)(\nabla g) + e^g\nabla[(\nabla' g)(\nabla g)] + e^g(\nabla g)(\nabla'\nabla g) + e^g(\nabla\nabla'\nabla g) \,,$$

since, obviously, $\nabla(A + B) = \nabla A + \nabla B$ for any matrices A and B of the correct size. In order to evaluate the term $\nabla[(\nabla' g)(\nabla g)]$ we apply Lemma 5. This leads to

$$\nabla[(\nabla' g)(\nabla g)] = (\nabla\nabla' g)(\nabla g) + (\nabla g)(\nabla'\nabla g) \,.$$

Therefore, we find that

$$\nabla\nabla'\nabla h = e^g(\nabla g)(\nabla' g)(\nabla g) + e^g(\nabla\nabla' g)(\nabla g) + 2e^g(\nabla g)(\nabla'\nabla g) + e^g(\nabla\nabla'\nabla g) \,. \tag{14}$$

Moreover, applying the Corollary following Lemma 2 to (13), we have

$$\nabla'\nabla\nabla' h = e^g(\nabla' g)(\nabla g)(\nabla' g) + e^g\nabla'[(\nabla g)(\nabla' g)] + e^g(\nabla' g)(\nabla\nabla' g) + e^g(\nabla'\nabla\nabla' g) \,.$$

We use Lemma 4 in order to evaluate the term $\nabla'[(\nabla g)(\nabla' g)]$ and so obtain

$$\nabla'[(\nabla g)(\nabla' g)] = 2(\nabla'\nabla g)(\nabla' g) \,,$$

and, as a consequence,

$$\begin{aligned} \nabla'\nabla\nabla' h = {} & e^g(\nabla' g)(\nabla g)(\nabla' g) + 2e^g(\nabla'\nabla g)(\nabla' g) \\ & + e^g(\nabla' g)(\nabla\nabla' g) + e^g(\nabla'\nabla\nabla' g) \,. \end{aligned} \tag{15}$$

Introducing the substitution $h = e^g$ together with its derivatives given by equations (10) through (15), the differential equation (9) finally assumes the simpler form

$$\begin{aligned} & (n\beta_3 - A_1)(\nabla\nabla'\nabla g) + [(n^2\beta_3 - A_1 - A_2)](\nabla\nabla' g)(\nabla g) \\ & \quad + [n(n+1)\beta_3 - 2A_1 - A_3](\nabla g)(\nabla'\nabla g) \\ & \quad + [n(n^2 - n + 1)\beta_3 - A_1 - A_2 - A_3 - A_4](\nabla g)(\nabla' g)(\nabla g) \\ & \quad + (C - n\beta_1)(\nabla g) = 0 \,. \end{aligned} \tag{16}$$

Before proceeding, let us determine the initial conditions on the functions h and g. We know that

$$h(t) = E(e^{itX'}) \,,$$

so that

$$h(0) = 1 \ . \tag{17}$$

Further, from (1),

$$\nabla h(0) = iE(X) = i\mu \ ,$$

say, while

$$\nabla' h(0) = iE(X') = i\mu' \ .$$

Also, for $\Sigma = [\sigma_{ij}] = E(X - \mu)'(X - \mu)$, it is easy to show that

$$\nabla'\nabla h(0) = -E(X'X) = -\Sigma - \mu'\mu \ ,$$

$$\nabla\nabla' h(0) = -E(XX') = -\sum_{j=1}^{p} \sigma_{jj} - \mu\mu'$$

using the fact that $\sigma_{jj} = E(x_j - \mu_j)^2$. Finally, the initial conditions on $g = \log h$ can be obtained from the above initial conditions on h and from (10) through (13). Thus

$$g(0) = 0 \ , \tag{18}$$

$$\nabla g(0) = \nabla h(0) = i\mu \ , \tag{19}$$

$$\nabla' g(0) = \nabla' h(0) = i\mu' \ , \tag{20}$$

$$\nabla'\nabla g(0) = \nabla'\nabla h(0) - [\nabla' g(0)][\nabla g(0)] = -\Sigma \ , \tag{21}$$

$$\nabla\nabla' g(0) = \nabla\nabla' h(0) - [\nabla g(0)][\nabla' g(0)] = -\sum_{j=1}^{p} \sigma_{jj} \ . \tag{22}$$

6. Fundamental results on "integration" of vector differential equations. In this section, we develop a number of mathematical tools which will prove essential in obtaining solutions to the fundamental vector differential equation (16) to give certain characterizations of multivariate distributions.

We indicate, in the following two lemmas, the conditions under which it is possible to "integrate" the differential operator ∇ to obtain a unique answer.

LEMMA 6. *If $s(t)$ is a scalar function of the vector variable $t = [t_1, \cdots, t_p]$, then a necessary and sufficient condition for the relation $\nabla s = 0$ to hold is that s is a constant.*

PROOF. Suppose that $\nabla s = 0$, the zero vector. Therefore,

$$\frac{\partial s}{\partial t_1} = \frac{\partial s}{\partial t_2} = \cdots = \frac{\partial s}{\partial t_p} = 0 \ .$$

which implies that s is independent of each of the t_i and hence is a constant. Conversely, if s is a constant, it is obvious that

$$\nabla s = 0 \ .$$

We immediately have the following corollary.

COROLLARY. *If $s(t)$ is a scalar function of the vector variable $t = [t_1, \cdots, t_p]$,* *then a necessary and sufficient condition for the relation $\nabla' s = 0$ to hold is that s is a constant.*

LEMMA 7. *If $Y = [y_1, \cdots, y_p]$ is a vector function of the vector variable $t = [t_1, \cdots, t_p]$, then a necessary and sufficient condition for the relation $\nabla' Y = 0$ to hold is that Y is a constant vector.*

PROOF. We consider $\nabla' Y = 0$, the zero matrix. Since the operator ∇' is applied in turn to each of the p scalar components of the row vector Y, the corollary following Lemma 6 guarantees that each of the components of Y must be constant.

As before, the converse is obvious.

Unfortunately, the above "integration" process breaks down as an inverse to the derivative of any vector or matrix in which addition of components occurs when the differential operator ∇ is applied. To circumvent this difficulty, we restrict ourselves to a particular class of scalar functions, those of the vector variable t for which an infinite series expansion in the variables t and t' exists and is convergent. In order to meaningfully discuss such infinite series, the following set of results on interchanging infinite summation with ∇-derivatives is essential.

LEMMA 8. *Let $Y_j = [y_{j1}, \cdots, y_{jp}]$ be a $1 \times p$ row vector for every $j = 1, 2, \cdots$. For any given k, $(k = 1, \cdots, p)$ let $\sum_{j=1}^{\infty} y_{jk}$ be a convergent series in the differentiable scalar functions y_{jk} $(j = 1, 2, \cdots)$ of the vector variable $t = [t_1, \cdots, t_p]$ such that $\partial/\partial t_k \sum_{j=1}^{n} y_{jk} = z_{n,k}$ converge uniformly. Then*

$$\nabla(\textstyle\sum_{j=1}^{\infty} Y_j') = \sum_{j=1}^{\infty} \nabla(Y_j') \quad \textit{and} \quad \nabla'(\sum_{j=1}^{\infty} Y_j) = \sum_{j=1}^{\infty} \nabla' Y_j \,. \tag{23}$$

PROOF. For finite n, it is easy to show that

$$\textstyle\sum_{j=1}^{n} \nabla(Y_j') = \nabla(\sum_{j=1}^{n} Y_j') \,.$$

Thus, we can reduce the infinite summation to

$$\textstyle\sum_{j=1}^{\infty} \nabla(Y_j') = \lim_{n\to\infty} \sum_{j=1}^{n} \nabla(Y_j') = \lim_{n\to\infty} \nabla(\sum_{j=1}^{n} Y_j') \,, \tag{24}$$

so that, comparing equations (23) and (24), we need only show that

$$\nabla(\textstyle\sum_{j=1}^{\infty} Y_j') = \lim_{n\to\infty} \nabla(\sum_{j=1}^{n} Y_j') \,. \tag{25}$$

For the right-hand side of this equation, we note that

$$\nabla(\textstyle\sum_{j=1}^{n} Y_j') = \sum_{j=1}^{n} \dfrac{\partial y_{j1}}{\partial t_1} + \sum_{j=1}^{n} \dfrac{\partial y_{j2}}{\partial t_2} + \cdots + \sum_{j=1}^{n} \dfrac{\partial y_{jp}}{\partial t_p} \,.$$

Therefore, in the limit, the right-hand side of (25) becomes

$$\lim_{n\to\infty} \nabla(\textstyle\sum_{j=1}^{n} Y_j') = \sum_{j=1}^{\infty} \dfrac{\partial y_{j1}}{\partial t_1} + \cdots + \sum_{j=1}^{\infty} \dfrac{\partial y_{jp}}{\partial t_p} \,.$$

For the left-hand side we have

$$\nabla(\sum_{j=1}^{\infty} Y_j') = \frac{\partial}{\partial t_1} \sum_{j=1}^{\infty} y_{j1} + \frac{\partial}{\partial t_2} \sum_{j=1}^{\infty} y_{j2} + \cdots + \frac{\partial}{\partial t_p} \sum_{j=1}^{\infty} y_{jp} .$$

Since the y_{jk} are scalars for any choices of j and k, the problem therefore reduces to showing

$$\frac{\partial}{\partial t_k} \sum_{j=1}^{\infty} y_{jk} = \sum_{j=1}^{\infty} \frac{\partial y_{jk}}{\partial t_k}$$

for every $k = 1, \cdots, p$. We introduce the substitutions

$$\sum_{j=1}^{n} y_{jk} = f_{nk}(t) \qquad \sum_{j=1}^{\infty} y_{jk} = f_k(t)$$

for any fixed $k = 1, \cdots, p$. For fixed k, the sequence of functions $\{f_{nk}\}$ satisfies all of the conditions in Theorem 7.17, Rudin [6], and so

$$\frac{\partial}{\partial t_k} f_k(t) = \lim_{n\to\infty} \frac{\partial}{\partial t_k} f_{nk}(t) ,$$

which completes the proof of the first equation (23). The second is proved similarly.

COROLLARY 1. *Let $\sum_{j=1}^{\infty} s_j$ be a convergent series in the differentiable scalar functions s_j $(j = 1, 2, \cdots)$ of the vector variable $t = [t_1, \cdots, t_p]$ and such that*

$$\frac{\partial}{\partial t_k} \sum_{j=1}^{n} s_j = z_{n,k}$$

converges uniformly for each $k = 1, \cdots, p$. Then

$$\nabla(\sum_{j=1}^{\infty} s_j) = \sum_{j=1}^{\infty} \nabla(s_j)$$

and

$$\nabla'(\sum_{j=1}^{\infty} s_j) = \sum_{j=1}^{\infty} \nabla'(s_j) .$$

COROLLARY 2. *Let $A_r = [a_{jk}^r]$ be a $p \times p$ matrix for every $r = 1, 2, \cdots$. For any fixed j, k $(j, k = 1, \cdots, p)$, let $\sum_{r=1}^{\infty} a_{jk}^r$ be a convergent series in the differentiable scalar functions a_{jk}^r $(r = 1, 2, \cdots)$ of the vector variable $t = [t_1, \cdots, t_p]$ such that, for each fixed j and k,*

$$\frac{\partial}{\partial t_j} \sum_{r=1}^{n} a_{jk}^r$$

converges uniformly. Then

$$\nabla(\sum_{r=1}^{\infty} A_r) = \sum_{r=1}^{\infty} \nabla(A_r) .$$

PROOF. The proof of this corollary follows directly from Lemma 8 by considering each of the matrices A_r $(r = 1, 2, \cdots)$ as a system of p column vectors.

We now consider the possibility of expressing a scalar function $f(t)$ of the vector variable $t = [t_1, \cdots, t_p]$ as an infinite series of the form

$$f(t) = a_0 + b_1 t' + a_2 tt' + b_3 t'tt' + a_4(tt')^2 + \cdots \\ = \textstyle\sum_{k=0}^{\infty} a_{2k}(tt')^k + \sum_{k=0}^{\infty} b_{2k+1} t'(tt')^k \tag{26}$$

where the a_j are scalar constants and the b_j are $1 \times p$ vector constants for each j. In order for this series expansion to be meaningful, we must be able to uniquely determine each of the coefficients and to be assured that the resulting series is convergent for all values of t under consideration.

In particular, if we evaluate the given series at $t = 0$, we immediately have

$$a_0 = f(0) .$$

The remaining coefficients will be obtained in terms of the ∇-derivatives of $f(t)$ evaluated at $t = 0$.

In order to evaluate the first derivative of $f(t)$, we must be able to differentiate such terms as $a(tt')^k$ and $bt'(tt')^k$, where a is a scalar constant and $b = [b_1, \cdots, b_p]$ is a vector constant. Corollary 1 following Lemma 8 guarantees that such termwise differentiation of the infinite series is valid. Thus, we see that

$$\nabla'[a(tt')^k] = a\nabla'(\textstyle\sum_{j=1}^{p} t_j^2)^k \\ = 2kat'(tt')^{k-1} , \tag{27}$$

while

$$\nabla'[(bt')(tt')^k] = [\nabla'(bt')](tt')^k + (bt')\nabla'[(tt')^k] \\ = b'(tt')^k + 2k(tb')t'(tt')^{k-1} \tag{28}$$

using (27) as well as the fact that $bt' = tb' = \sum_{j=1}^{p} b_j t_j$, a scalar.

Hence, the termwise differentiation of the infinite series for $f(t)$ leads to

$$\nabla' f(t) = \textstyle\sum_{k=1}^{\infty} 2ka_{2k} t'(tt')^{k-1} + \sum_{k=0}^{\infty} b'_{2k+1}(tt')^k + \sum_{k=1}^{\infty} 2k(tb'_{2k+1})t'(tt')^{k-1} \tag{29}$$

which, for the case $t = 0$, reduces to

$$\nabla' f(0) = b_1' .$$

That is,

$$b_1 = \nabla f(0) .$$

We now evaluate the second derivative of $f(t)$, $\nabla\nabla' f(t)$, and, for this purpose, we must consider the derivatives of terms of the form $t'(tt')^{k-1}$, $b'(tt')^k$ and $(tb')t'(tt')^{k-1}$. Termwise differentiation of he infinite series for $\nabla' f(t)$ given in (29) is justified in this case by referring to Lemma 8. As a result of Lemma 1, we have

$$\nabla[t'(tt')^{k-1}] = (p + 2k - 2)(tt')^{k-1} \tag{30}$$

$$\nabla[b'(tt')^k] = 2k(bt')(tt')^{k-1} \tag{31}$$

$$\nabla[(tb')(tt')^{k-1}t'] = (p + 2k - 1)(bt')(tt')^{k-1} \tag{32}$$

after some simple calculations. Therefore, by differentiating the infinite series for $\nabla' f(t)$ given by (29) termwise, we obtain

$$\begin{aligned} \nabla\nabla' f(t) &= \textstyle\sum_{k=1}^{\infty} 2ka_{2k}(p+2k-2)(tt')^{k-1} + \sum_{k=1}^{\infty} 2k(b_{2k+1}t')(tt')^{k-1} \\ &\quad + \textstyle\sum_{k=1}^{\infty} 2k(p+2k-1)(b_{2k+1}t')(tt')^{k-1} \qquad (33) \\ &= \textstyle\sum_{k=1}^{\infty} 2ka_{2k}(p+2k-2)(tt')^{k-1} \\ &\quad + \textstyle\sum_{k=1}^{\infty} 2k(p+2k)(b_{2k+1}t')(tt')^{k-1} . \end{aligned}$$

Evaluating this expression at $t = 0$, we find

$$a_2 = (1/2p)\nabla\nabla' f(0) .$$

We now notice that, except for constant scalar coefficients, the forms of $f(t)$ and $\nabla\nabla' f(t)$ are identical.

Thus, we find that the general expressions for the higher order derivatives are given by

$$\begin{aligned} (34)\quad (\nabla\nabla')^r f(t) &= \sum_{k=r}^{\infty} 2^r \frac{k!}{(k-r)!} \prod_{j=1}^{r} [p+2(k-j)]a_{2k}(tt')^{k-r} \\ &\quad + \sum_{k=r}^{\infty} 2^r \frac{k!}{(k-r)!} \prod_{j=1}^{r} [p+2(k-j+1)](b_{2k+1}t')(tt')^{k-r} \end{aligned}$$

and

$$\begin{aligned} \nabla'(\nabla\nabla')^r f(t) &= \sum_{k=r+1}^{\infty} 2^{r+1} \frac{k!}{[k-(r+1)]!} \prod_{j=1}^{r} [p+2(k-j)]a_{2k}t'(tt')^{k-(r+1)} \\ &\quad + \sum_{k=r}^{\infty} 2^r \frac{k!}{(k-r)!} \prod_{j=1}^{r} [p+2(k-j+1)]b'_{2k+1}(tt')^{k-r} \qquad (35) \\ &\quad + \sum_{k=r+1}^{\infty} 2^{r+1} \frac{k!}{[k-(r+1)]!} \prod_{j=1}^{r} [p+2(k-j+1)](b_{2k+1}t') \\ &\quad \times (t')(tt')^{k-(r+1)} . \end{aligned}$$

These two formulae are proved by a lengthy, though relatively straightforward induction argument.

We are now able to calculate the coefficients in the expansion for $f(t)$ given in (26) by using (34) and (35). Thus, if we evaluate the expression (34) at $t = 0$, we find that the only nonzero term in this series expansion occurs for $k = r$ and gives

$$(36)\qquad a_{2r} = \frac{1}{2^r r! \prod_{j=1}^{r} [p+2(r-j)]} (\nabla\nabla')^r f(0) .$$

Similarly, we evaluate (35) at $t = 0$ and again find that the only nonzero term occurs for $k = r$ and is

$$(37)\qquad b_{2r+1} = \frac{1}{2^r r! \prod_{j=1}^{r} [p+2(r-j+1)]} \nabla(\nabla\nabla')^r f(0) ,$$

a $1 \times p$ vector constant.

Hence, the expansion for $f(t)$ becomes

$$
\begin{aligned}
f(t) &= f(0) + \nabla f(0)t' \\
&\quad + \textstyle\sum_{k=1}^{\infty} \dfrac{1}{2^k k! \prod_{j=1}^{k}[p+2(k-j)]}[(\nabla\nabla')^k f(0)](tt')^k \\
&\quad + \textstyle\sum_{k=1}^{\infty} \dfrac{1}{2^k k! \prod_{j=1}^{k}[p+2(k-j+1)]}[\nabla(\nabla\nabla')^k f(0)]t'(tt')^k .
\end{aligned}
\tag{38}
$$

We restrict our attention in future discussions to those functions $f(t)$ for which the above series expansion exists and converges. We refer to the class of all such scalar functions of the variable t as the class of pseudo-analytic functions of type I of the vector variable t.

We now digress for a moment to show that the series expansion for the function $f(t)$ given by (38) is, in reality, a generalization of the standard Taylor series expansion for a function of a single variable about the origin. That is, if we choose $p = 1$, the function $f(s)$ depends on the single variable s and, furthermore, the vector constants b_{2k+1} $(k = 1, 2, \cdots)$ reduce to scalars. Obviously, for $p = 1$, the ∇-derivative reduces to the usual derivative d/ds.

If we therefore set $p = 1$ in the defining equation (36) for the a_{2k} $(k = 0, 1, 2, \cdots)$, we find that

$$a_{2k} = \frac{1}{2^k k! \prod_{j=1}^{k}[1+2(k-j)]} \frac{d^{2k}}{ds^{2k}} f(0) .$$

However, we note that

$$(2k)! = 2k(2k-1)(2K-2)(2k-3)\cdots 4\cdot 3\cdot 2\cdot 1 = 2^k k! \textstyle\prod_{j=1}^{k}(2k-2j+1) .$$

As a consequence, we have that

$$a_{2k} = \frac{1}{(2k)!} \frac{d^{2k}}{ds^{2k}} f(0) .$$

Similarly, we set $p = 1$ in (37) to obtain

$$b_{2k+1} = \frac{1}{2^k k! \prod_{j=1}^{k}[1+2(k-j+1)]} \frac{d^{2k+1}}{ds^{2k+1}} f(0) .$$

However, as above, we find that

$$(2k+1)! = 2^k k! \textstyle\prod_{j=1}^{k}(2k-2j+3) ,$$

so that

$$b_{2k+1} = \frac{1}{(2k+1)!} \frac{d^{2k+1}}{ds^{2k+1}} f(0) .$$

Substituting these coefficients into the expansion for $f(s)$, we find that

$$
\begin{aligned}
f(s) &= \textstyle\sum_{k=0}^{\infty} \dfrac{1}{(2k)!}\left[\dfrac{d^{2k}}{ds^{2k}} f(0)\right] s^{2k} + \textstyle\sum_{k=0}^{\infty} \dfrac{1}{(2k+1)!}\left[\dfrac{d^{2k+1}}{ds^{2k+1}} f(0)\right] s^{2k+1} \\
&= \textstyle\sum_{k=0}^{\infty} \dfrac{1}{k!} \dfrac{d^k}{ds^k} f(0)s^k ,
\end{aligned}
$$

which is precisely the Taylor expansion for the function $f(s)$ about the point $s=0$.

It is important to realize that, in calculating the coefficients of the given series expansion (26), the order of differentiation was arbitrarily chosen. That is, we could just as easily have begun by first applying the operator ∇ to the expansion for $f(t)$, instead of the operator ∇' as was done in the previous development. We now study the corresponding class of scalar functions of the vector variable $t=[t_1,\cdots,t_p], f(t)$, whose infinite series expansion is given by

$$f(t)=\sum_{k=0}^{\infty}a_{2k}(tt')^k+\sum_{k=0}^{\infty}b_{2k+1}t'(tt')^k\,, \tag{39}$$

where the a_j are scalar constants and the b_j are $1\times p$ vector constants for each j. In this case, we determine these coefficients by first applying the operator ∇ to $f(t)$, then ∇', and so forth. We shall call the class of functions $f(t)$ for which the resulting series expansion exists and is convergent as the class of pseudo-analytic functions of type II of the vector variable t.

Following a derivation which closely parallels the one leading to (38), we obtain as the series expansion for the second class of functions

$$\begin{aligned} f(t)&=f(0)+[\nabla f(0)]t'+(1/2)t[\nabla'\nabla f(0)]t' \\ &+\sum_{k=2}^{\infty}\frac{1}{2^k k!\prod_{j=1}^{k-1}[p+2(k-j)]}t[(\nabla'\nabla)^k f(0)]t'(tt')^{k-1} \\ &+\sum_{k=1}^{\infty}\frac{1}{2^k k!\prod_{j=1}^{k}[p+2(k-j+1)]}[\nabla(\nabla'\nabla)^k f(0)]t'(tt')^k\,. \end{aligned} \tag{40}$$

7. Some characterization theorems based on the fundamental vector differential equation. We now characterize the multivariate normal distribution by using the property of cubic regression for the vector statistics S on L discussed in Section 5. For some suitable choices for the coefficients in the fundamental vector differential equation (16), (that is, for some particular relations between the coefficients of the statistic S and the regression coefficients), we obtain such characterizations by using the results of the previous section which ensure that a unique inverse exists, in certain instances, for the differential operator ∇.

THEOREM 3. *Let $X_1,\cdots,X_n$ be a set of n $(1\times p)$ independently and identically distributed stochastic row vectors (a sample from a population), each with finite moments $E(X)$, $E(X'X)$ and $E(XX'X)$. Suppose that*

$$\begin{aligned} n\beta_3-A_1&=0\,,\\ n(n+1)\beta_3-2A_1-A_3&=0\,,\\ n(n^2-n+1)\beta_3-A_1-A_2-A_3-A_4&=0\,,\\ C-n\beta_1&=\sigma(n^2\beta_3-A_1-A_2)\neq 0\,, \end{aligned}$$

where $\sigma=\sum_{j=1}^{p}\sigma_{jj}$. Then, in the class of populations whose characteristic *functions are pseudo-analytic of type* I *of the vector variable* t, *a necessary and sufficient condition for* S *to have cubic regression on* L *is that the population be multivariate normal with characteristic function*

$$h(t) = \exp[i\mu t' - t((1/p)\sigma I)t'] .$$

PROOF. If we impose the given conditions on the coefficients of the fundamental vector differential equation, it reduces to

$$(\nabla\nabla' g)(\nabla g) = -\sigma(\nabla g) .$$

We post-multiply this equation by $(\nabla' g)$, thereby obtaining

$$(\nabla\nabla' g) \sum_{j=1}^{p} \left(\frac{\partial g}{\partial t_j}\right)^2 = -\sigma \sum_{j=1}^{p} \left(\frac{\partial g}{\partial t_j}\right)^2 . \tag{4.1}$$

However, $\sum_{j=1}^{p} (\partial g/\partial t_j)^2 > 0$, for if not, we would have necessarily that

$$\frac{\partial g}{\partial t_j} = 0$$

for every $j = 1, \cdots, p$. That is, $\nabla g = 0$, which implies that g, and consequently h, is a constant. Therefore, the differential equation (41) assumes the simpler form

$$\nabla\nabla' g = -\sigma . \tag{42}$$

Applying the differential operator ∇' to this equation, we obtain $\nabla'\nabla\nabla' g = 0$, and consequently,

$$(\nabla\nabla')^r g = 0 \quad \text{for} \quad r \geqslant 2$$

and

$$\nabla'(\nabla\nabla')^r g = 0 \quad \text{for} \quad r \geqslant 1 .$$

Therefore, $g(t)$ can be written as an infinite series expansion of the type given by (38), which reduces to the convergent expression

$$g(t) = g(0) + [\nabla g(0)]t' + (1/2p)[\nabla\nabla' g(0)]tt' . \tag{43}$$

The initial conditions on $g(t)$ given by (18), (19) and (22) enable us to simplify (43) and finally obtain

$$h(t) = \exp[g(t)] = \exp[i\mu t' - \tfrac{1}{2}t((1/p)\sigma I)t'] ,$$

which is the characteristic function of a multivariate normal distribution with variance-covariance matrix

$$\Sigma = (1/p)\sigma I .$$

Conversely, suppose that the population is multivariate normal and has characteristic function

$$h(t) = \exp[i\mu t' - \tfrac{1}{2}t((1/p)\sigma I)t'] .$$

Thus,

$$g(t) = \log h(t) = i\mu t' - \tfrac{1}{2}t((1/p)\sigma I)t' ,$$

so that

$$\nabla g = i\mu - (1/p)\sigma t .$$

Taking the transpose, we obtain

$$\nabla' g = i\mu' - (1/p)\sigma t' ,$$

from which we can also calculate

$$\nabla\nabla' g = -\sigma .$$

Substituting these results into (16) and using the conditions on the coefficients given in the hypothesis, we see that the function $g(t)$ satisfies the fundamental vector differential equation and therefore S has cubic regression on L, which proves the theorem.

THEOREM 4. *Let $X_1, \cdots, X_n$ be a set of $n(1 \times p)$ independently and identically distributed stochastic row vectors, each with finite moments $E(X)$, $E(X')$, $E(XX'X)$. Suppose that*

$$n\beta_3 - A_1 = 0 ,$$
$$n^2\beta_3 - A_1 - A_2 = 0 ,$$
$$n(n^2 - n + 1)\beta_3 - A_1 - A_2 - A_3 - A_4 = 0 ,$$
$$(C - n\beta_1)I = [n(n+1)\beta_3 - 2A_1 - A_3]\Sigma \neq 0 .$$

Then, in the class of populations whose characteristic functions are pseudo-analytic of type II *of the vector variable t, a necessary and sufficient condition for S to have cubic regression on L is that the population be multivariate normal with characteristic function*

$$h(t) = \exp[i\mu t' - \tfrac{1}{2}t\Sigma t'] ,$$

where Σ is a constant multiple of the identity.

PROOF. If we impose the given conditions on the coefficients of the fundamental vector differential equation, it reduces to

$$(\nabla g)(\nabla'\nabla g) = d(\nabla g) , \tag{44}$$

where

$$d = \frac{n\beta_1 - C}{n(n+1)\beta_3 - 2A_1 - A_3} .$$

We now attempt to "integrate" the above equation. Using Lemma 4, we find that,

$$\nabla'(\nabla g\nabla' g) = 2(\nabla'\nabla g)(\nabla' g) ,$$

and taking the transpose of both sides, we obtain

$$\nabla(\nabla g\nabla' g) = 2(\nabla g)(\nabla'\nabla g) ,$$

since the matrix $\nabla'\nabla g$ is symmetric. Hence, the differential equation (44) can be rewritten as

$$\nabla(\nabla g \nabla' g - 2dg) = 0 .$$

However, since $\nabla g \nabla' g - 2dg$ is a scalar, Lemma 6 guarantees that

$$\nabla g \nabla' g - 2dg = c ,$$

a constant. From the initial conditions (18), (19) and (20), we see that $c = -\mu\mu'$ and therefore,

$$\nabla g \nabla' g = 2dg - \mu\mu' . \tag{45}$$

Within the class of pseudo-analytic functions of type II, the function $g(t)$ can be written in the form

$$g(t) = \textstyle\sum_{k=0}^{\infty} a_{2k}(tt')^k + \sum_{k=0}^{\infty} b_{2k+1}t'(tt')^k , \tag{46}$$

where the first three coefficients can be evaluated, using the conditions (18), (19) and (21), as

$$\begin{aligned} a_0 &= g(0) = 0 \\ b_1 &= \nabla g(0) = i\mu \\ a_2 I &= \tfrac{1}{2}\nabla'\nabla g(0) = -\tfrac{1}{2}\Sigma . \end{aligned} \tag{47}$$

We now determine the remaining coefficients to see that the series indeed converges. If we apply the differential operator ∇ to (46), we see that

$$\begin{aligned} \nabla g &= \textstyle\sum_{k=0}^{\infty} b_{2k+1}(tt')^k + \sum_{k=1}^{\infty} 2kb_{2k+1}t't(tt')^{k-1} + \sum_{k=1}^{\infty} 2ka_{2k}t(tt')^{k-1} \\ &= b_1 + 2a_2 t + (b_3 tt' + 2b_3 t't) + 4a_4 t(tt') \\ &\quad + [b_5(tt')^2 + 4b_5 t't(tt')] + \cdots . \end{aligned} \tag{48}$$

Similarly, it is easy to obtain

$$\begin{aligned} \nabla' g = b_1' + 2a_2 t' + (b_3' tt' + 2b_3 t't') + 4a_4 t'tt' \\ + [b_5'(tt')^2 + 4b_5 t't'tt'] + \cdots . \end{aligned} \tag{49}$$

We now substitute (46), (48) and (49) into (45) and consider "powers" of t. The constant term is

$$b_1 b_1' = -\mu\mu' + 2da_0 ,$$

which is consistent using (47). The first order term is

$$2a_2 b_1 t' + 2a_2 t b_1' = 2db_1 t' ,$$

or equivalently,

$$2a_2 b_1 t' = db_1 t' .$$

However, comparing the final condition of the hypothesis with (47) we see that $d = 2a_2$, so that the above equation for the first order term is also consistent. In order to determine the remaining coefficients in the series expansion for $g(t)$, it is necessary to distinguish between two possible situations:

$$\text{(i)} \quad \mu \not\equiv 0 \qquad \text{(ii)} \quad \mu \equiv 0 .$$

For the case $\mu \not\equiv 0$, we calculate the second and third order terms in (45). The second order term reduces to

$$2b_1 b_3' tt' + 4b_1 t' b_3 t' + 4a_2^2 tt' = 2da_2 tt' ,$$

which, by virtue of the relation $d = 2a_2$, is equivalent to

(50) $$b_1 b_3' tt' = -2b_1 t' b_3 t' .$$

However, since $\mu \not\equiv 0$, it follows that $b_1 \not\equiv 0$ also, from (47). By direct evaluation of (50), it is then easy to verify that we must have

$$b_3 = 0 .$$

Further, the third order term reduces to

$$4a_4(b_1 t')tt' + 6a_2(b_3 t')(tt') = d(b_3 t')tt'$$

and again, using $d = 2a_2$, the above equation becomes

(51) $$a_4(b_1 t')tt' = -a_2(b_3 t')tt' .$$

Since $b_3 = 0$, $b_1 \neq 0$ and t is arbitrary, we must have

$$a_4 = 0 .$$

For the case $\mu \equiv 0$, we again consider the second and third order terms, as well as the fourth order term. The second order term is given by (50). Since $b_1 = i\mu = 0$, this equation is satisfied for arbitrary b_3. The third order term is given by (51). Since $b_1 = 0$ and $a_2 \neq 0$, it follows that

$$b_3 = 0$$

and a_4 is arbitrary. Since $b_1 = b_3 = 0$, the fourth order term reduces to

$$4a_2 a_4 (tt')^2 = a_2 a_4 (tt')^2 ,$$

which implies that

$$a_4 = 0 .$$

Therefore, in both cases, we obtain

$$b_3 = 0 \qquad a_4 = 0 .$$

We now introduce the function

$$\begin{aligned} Z_n &= a_n \quad \text{for } n \text{ even} \\ &= b_n \quad \text{for } n \text{ odd} . \end{aligned}$$

To calculate Z_n for $n \geqq 5$, we note that it will involve all products of the form

$$Z_j Z_{n-j} \qquad (j = 0, 1, \cdots, n) .$$

However, for $n \geqslant 5$, either j or $n-j$ must be greater than or equal to 3, so that, by induction,

$$Z_n = 0 \quad \text{for all} \quad n \geqslant 5 .$$

As a result, using (47), the expression (46) for $g(t)$ reduces to

$$g(t) = i\mu t' - \tfrac{1}{2} t\Sigma t' ,$$

where

$$\Sigma = -2a_2 I .$$

Hence,

$$h(t) = \exp[i\mu t' - \tfrac{1}{2} t\Sigma t'] ,$$

which is the characteristic function of the multivariate normal distribution. As in Theorem 3, the converse is obtained by direct calculation.

In order to investigate more fully the condition that Σ be a constant multiple of the identity matrix, we substitute the normal characteristic function

$$h(t) = \exp[i\mu t' - \tfrac{1}{2} t\Sigma t']$$

into the fundamental vector differential equation (16). For $g = \log h$, we can easily calculate

$$\begin{aligned} \nabla g &= i\mu - t\Sigma & \nabla' g &= i\mu' - \Sigma t' \\ \nabla'\nabla g &= -\Sigma & \nabla\nabla' g &= -\sigma \\ \nabla\nabla'\nabla g &= 0 & \nabla'\nabla\nabla' g &= 0 . \end{aligned} \tag{52}$$

Denote the coefficients of the terms in the fundamental vector differential equation (16) by d_1, d_2, d_3, d_4 and d_5, respectively, so that for the case of the multivariate normal distribution, this equation reduces to

$$\begin{aligned} &-d_2\sigma(i\mu - t\Sigma) - d_3(i\mu - t\Sigma)\Sigma \\ &\qquad + d_4(i\mu - t\Sigma)(i\mu' - \Sigma t')(i\mu - t\Sigma) + d_5(i\mu - t\Sigma) = 0 . \end{aligned} \tag{53}$$

We look at "powers" of $t\Sigma$ in order to determine what types of relationships are necessary between the coefficients, where, by "powers", in this case, we mean the alternating terms $(t\Sigma)$, $(t\Sigma)(t\Sigma)'$, $(t\Sigma)(t\Sigma)'(t\Sigma)$, and so forth. Equation (53) must hold true for every value of t. Therefore, since the coefficient of $(t\Sigma)(t\Sigma)'(t\Sigma)$ is d_4, it is necessary that $d_4 = 0$. Substituting this value for d_4 into (53), we find

$$(t\Sigma)(d_2\sigma I + d_3\Sigma - d_5 I) - i\mu(d_2\sigma I + d_3\Sigma - d_5 I) = 0 .$$

Again, because this latter equation must hold true for every possible value of t, the condition which we must necessarily have is

$$d_2\sigma I + d_3\Sigma - d_5 I = 0 ,$$

or equivalently,

$$\Sigma = (1/d_3)(d_5 - d_2\sigma)I .$$

That is, Σ must be a constant multiple of the identity matrix.

THEOREM 5. *Let $X_1, \cdots, X_n$ be a set of n $(1 \times p)$ independently and identically distributed stochastic row vectors, each with finite moments $E(X)$, $E(X'X)$, $E(XX'X)$. Suppose that*

$$\begin{aligned} n^2\beta_3 - A_1 - A_2 &= 0\,, \\ n(n+1)\beta_3 - 2A_1 - A_3 &= 0\,, \\ n(n^2 - n + 1)\beta_3 - A_1 - A_2 - A_3 - A_4 &= 0\,, \\ \mathrm{C} - n\beta_1 &= 0\,, \\ n\beta_3 - A_1 &\neq 0\,. \end{aligned}$$

Then, in the class of populations whose characteristic functions are pseudo-analytic of type II *of the vector variable t, a necessary and sufficient condition for S to have cubic regression on L is that population be multivariate normal.*

PROOF. If we impose the given conditions on the coefficients of the fundamental vector differential equation (16), it reduces to $\nabla\nabla'\nabla g = 0$. We therefore have

$$\begin{aligned} (\nabla'\nabla)^r g &= 0\,, & r &\geqslant 2 \\ \nabla(\nabla'\nabla)^r g &= 0 & r &\geqslant 1\,. \end{aligned}$$

Thus, we can expand $g(t)$ as an infinite series of the type given by (40), which reduces to

$$g(t) = g(0) + [\nabla g(0)]t' + \tfrac{1}{2}t[\nabla'\nabla g(0)]t'\,.$$

Since this expansion exists and converges, $g(t)$ is a member of the class of pseudo-analytic functions of type II. The initial conditions on $g(t)$ given by (18), (19), and (21) enable us to simplify the above expansion still further to obtain

$$g(t) = i\mu t' - \tfrac{1}{2}t\Sigma t'\,.$$

As a consequence,

$$h(t) = \exp[i\mu t' - \tfrac{1}{2}t\Sigma t']\,,$$

which is the characteristic function of a multivariate normal distribution. The converse follows easily from the given conditions on the coefficients of the fundamental vector differential equation and from (52).

8. Characterizations of multivariate normal distributions using quadratic and constant regression. For our preceding study of cubic regression, it was necessary, in order to maintain consistency in size, to consider only the odd powers in the polynomial expansion for $E(X/Y)$, where X and Y are row vectors. It has already been pointed out that it is possible to consider separately the even terms. For this purpose, we introduce the following definition.

DEFINITION. The stochastic $p \times p$ matrix Z has even polynomial regression of order $2m$ on the stochastic $1 \times p$ vector Y if

$$E(Z/Y) = \sum_{j=0}^{m} \beta_{2j} Y'(YY')^{j-1}Y, \qquad \text{a.e.}$$

where m is a nonnegative integer and the β's are real constants, provided that the conditional moment $E(Z/Y) = E(Z/Y = y)$ exists. This definition is entirely analogous to the definition of Section 4 for odd polynomial regression. Corresponding to Theorem 1, we also have the following result.

THEOREM 6. *Let Z be a $p \times p$ stochastic matrix and Y be a $1 \times p$ stochastic vector and assume that the expectations $E(Z)$ and $E[(YY')^{r-1}Y'Y]$ exist, for all nonnegative integers $r \leqslant m$. Then Z has even polynomial regression of order $2m$ on Y if, and only if, the relation*

$$E(Ze^{itY'}) = \sum_{j=0}^{m} \beta_{2j}[E\{(YY')^{j-1}Y'Ye^{itY'}\}]$$

holds for every real $1 \times p$ vector t.

The proof of this theorem is similar to that for Theorem 1 and hence we will not go through the details.

We now suppose that we have a set of n stochastic $1 \times p$ vectors $X_1, \cdots, X_n$ which are independently and identically distributed. We shall study the quadratic statistic

$$Q = \sum_{j,k=1}^{n} a_{jk} X_j' X_k ,$$

a stochastic $p \times p$ matrix, whose regression on the linear statistic

$$L = \sum_{j=1}^{n} X_j ,$$

a stochastic $1 \times p$ row vector, is of the quadratic form

$$E(Q/L) = \beta_2 L'L + \beta_0 I \qquad \text{a.e.} \tag{54}$$

where a_{jk} $(j, k = 1, \cdots, n)$, β_0 and β_2 are real constants. Using Theorem 6, a necessary and sufficient condition for (54) to hold true is

$$E(Qe^{itL'}) = \beta_2 E(L'Le^{itL'}) + \beta_0 E(Ie^{itL'}) . \tag{55}$$

It is possible to derive the differential equation which results from the quadratic regression assumption in terms of the characteristic function $h(t)$ for the population and its derivatives, ∇h, $\nabla' h$ and $\nabla'\nabla h$, studied in Section 3. As in the cubic regression considerations of Section 5, we substitute for Q and L in (55) and, upon evaluation of both sides, we easily obtain the vector differential equation

$$(n\beta_2 - A_1)(\nabla'\nabla h)/h + [n(n-1)\beta_2 - A_2][(\nabla' h)/h][(\nabla h)/h] = \beta_0 I , \tag{56}$$

where

$$A_1 = \Sigma_j a_{jj} , \qquad A_2 = \Sigma_{j \neq k} a_{jk} .$$

Rather than attempting to solve the vector differential equation (56), it is again convenient to introduce the transformation $h = e^g$. Using (10), (11) and (12), we therefore find

$$(n\beta_2 - A_1)(\nabla'\nabla g) + (n^2\beta_2 - A_1 - A_2)(\nabla' g)(\nabla g) = \beta_0 I , \tag{57}$$

which is the fundamental vector differential equation for the quadratic case. For a suitable choice of the coefficients in this fundamental vector differential equation (that is, for some particular relations between the coefficients of the statistic Q and the regression coefficients), it is possible to obtain a characterization for the multivariate normal distribution by using the property of quadratic regression for the statistics Q on L. In this case, the characterization is again obtained by using the results of Section 6 to uniquely "integrate" the operator ∇.

THEOREM 7. *Let $X_1, \cdots, X_n$ be a set of $n(1 \times p)$ independently and identically distributed stochastic row vectors, each with finite variance-covariance matrix $\Sigma = E(X-\mu)'(X-\mu)$ and mean $\mu = E(X)$. Suppose that*

$$n^2\beta_2 - A_1 - A_2 = 0 , \qquad \beta_0 I = (A_1 - n\beta_2)\Sigma \neq 0 .$$

Then in the class of populations whose characteristic functions are pseudo-analytic of type II *of the vector variable t, a necessary and sufficient condition for Q to have quadratic regression on L is that the population be multivariate normal.*

PROOF. If we impose the given conditions on the coefficients, the fundamental vector differential equation (57) reduces to

$$\nabla'\nabla g = c_0 I ,$$

where

$$c_0 = \beta_0/(n\beta_2 - A_1) .$$

Applying the differential operator ∇ to this equation, we have

$$\nabla\nabla'\nabla g = 0 ,$$

which is precisely the same differential equation as that obtained in Theorem 5 and hence has the solution

$$g(t) = i\mu t' - \tfrac{1}{2}t\Sigma t' ,$$

which in turn leads to

$$h(t) = \exp[i\mu t' - \tfrac{1}{2}t\Sigma t'] ,$$

the characteristic function of the multivariate normal distribution with $\Sigma = -c_0 I$.

The converse follows easily by direct calculation.

The above theorem is a multivariate generalization of a theorem first given by Laha and Lukacs [3] in their study of univariate quadratic regression. Furthermore, if we choose $\beta_2 = 0$ in the vector differential equation (56), we are led to the following theorem, a multivariate generalization of a theorem which was first studied by Lukacs and Laha [5] as a problem in univariate constant regression.

THEOREM 8. *Let* $X_1, \cdots, X_n$ *be a set of* $n(1 \times p)$ *independently and identically distributed stochastic row vectors, each with finite variance-covariance matrix* Σ, *and mean* μ. *Suppose that*

$$A_1 + A_2 = 0\,, \qquad \beta_0 I = A_1 \Sigma \neq 0\,.$$

Then, in the class of populations whose characteristic functions are pseudo-analytic of type II *of the vector variable* t, *a necessary and sufficient condition for* Q *to have constant regression on* L *is that the population be multivariate normal.*

REFERENCES

[1] GORDON, F. S. (1968). Characterizations of univariate and multivariate distributions using regression properties. Ph. D. thesis, McGill Univ.

[2] GORDON, F. S. and MATHAI, A. M. Characterizations of distributions using regression. Submitted for publication.

[3] LAHA, R. G. and LUKACS, E. (1960). On some characterization problems connected with quadratic regression. *Biometrika* **47** 335-343.

[4] LUKACS, E. (1960). *Characteristic Functions*. Charles Griffin and Co., London.

[5] LUKACS, E. and LAHA, R. G. (1964). *Applications of Characteristic Functions*. Charles Griffin and Co., London.

[6] RUDIN, W. (1953). *Principles of Mathematical Analysis*. McGraw Hill, New York.

[7] WEDDERBURN, J. H. M. (1934). *Lectures on Matrices*. Amer. Math. Soc. Coll. Publ. **17** New York.