SOME PROBLEMS IN THE CHARACTERIZATION OF THE WISHART DISTRIBUTION

Nadia Abud Habeeb AL-Mousaway
Dept.of Mathematics\ College of Science of Computer and Mathematics\ University of Thi-Qar

ABSTRACT:

Under the multivariate linear model $\{Y, X\beta, \sum \otimes V\}$, A number of characterization of the distribution of X_i have been made based on the properties of the statistics Y_1 and Y_2 when Y_1 and Y_2 be two linear functions defined on R^1 as follows $Y_1 = a_1X_1 + ... + a_nX_n$ and $Y_2 = b_1X_1 + ... + b_nX_n$. Generalizations of these problems to the multivariate case have been made by several authors by extending the techniques used in the univariate case. In my paper I shall consider some other generalization, which possibly require development of new techniques ,if X_1, X_2 be independent and identically distributed p-vector r.v.'s such that $E(X_1 - AX_2 \setminus X_1 + B'X_2) = 0$,where A and B are nonsingular matrices. In special case when $A = B^{-1}$, A is symmetric and the egen values do not take values ± 1. Under these conditions that X_1 has an m.n.d.
in the present paper we shall consider a few other cases.

1. Introduction: Let $X_1,.....,X_n$ be n independent random variables defined on R^1 and $Y_1 = a_1X_1,.....,a_nX_n$ and $Y_2 = b_1X_1,.....,b_nX_n$ be two linear functions. A number of characterization of the distribution of X_i have been made based on the properties of the statistics Y_1 and Y_2 of which the following are a few examples:

(i) Y_1 and Y_2 are independently distributed (Darmois-skitovic);
(ii) Y_1 and Y_2 are independently distributed(Linnik , [8]);
(iii) E(Y_1/Y_2)=constant (Ramachandran and Rao ,[9,10]);
(iv) conditional distribution of Y_1 and Y_2 is symmetric (Heyde , [3]).

Generalizations of these problems to the multivariate case have been made by several authors by extending the techniques used in the univariate case. In my paper I shall consider some other generalization, which possibly require development of new techniques.
The following definitions , notations and abbreviations are used through out the paper.

Nonsingular distribution: A random variable X is said to have a nonsingular distribution if no linear combination of X has a degenerate distribution.

Homoscedasticity: Let X_1 and X_2 be the random variables. The conditional distribution of X_1 given X_2 is said to be homoscedastic if the conditional distribution of X_2 - E(X_2/X_1) given X_1 does not depend on X_1.

Weak Homoscedasticity: The conditional distribution of X_2 given X_1 is said to be weakly homoscedastic if D(X_2/X_1) =Σ is independent of X_1 where D stands for the dispersion operator.

Abbreviations:

r.v. = random variable
m.n.d. = multivariate normal distribution,

w.d = wishart distribution
i.d = independently distribution
c.f = characteristic function
s.c.f = logarithm of c.f. defined in the neighborhood of origin, also called the second characteristic function.

2. The ever green Cauchy equation.

The famous Cauchy equation

$$f(X+Y)=f(X)+f(Y) \qquad (2.1)$$

Where V is some space and f is a function defined on V . Under some mild conditions on f , the solution is known to be linear.

Suppose we restrict the validity of (2.1) not to all $X,Y \in V$, but only to pairs X,Y satisfying some condition. Does the solution remain linear ? As a specific problem, let V be a vector space furnished with an inner product and the restriction be such as (x,y) = 0 , i.e., the inner product vanishes . In such a case f in (2.1) can be a quadratic function as shown in lemma 1.

Lemma1: let V a vector furnished with an inner product and f be a continuous complex valued function defined on V such that

$$f(X+Y)=f(X)+f(Y) \qquad (2.2)$$

$\forall X,Y \in V$ such that (X,Y)=0

Then f is a polynomial of degree not greater than two.

Proof: Let us consider any two dimensional subspace V_2 .Suppose that there exist two pairs of unit vectors e_1,e_2 and h_1,h_2 each constituting a basis of V_2 , such that $e_1 \neq \pm h_1$ or $\pm h_2$ and

$$f(t_1h_1+t_2h_2)=f(t_1h_1)+f(t_2h_2) \quad \forall \quad t_1,t_2 \in F \qquad (2.3)$$

$$f(u_1e_1+u_2e_2)=f(u_1e_1)+f(u_2e_2) \quad \forall \quad u_1,u_2 \in F \qquad (2.4)$$

Where F is the scalar filed associated with V.

There exist $a_1,a_2,b_1,b_2 \in F$ such that $h_1=a_1e_1+a_2e_2$ and $h_2=b_1e_1+b_2e_2$.Substituting for h_1 and h_2 in (2.3) and using (2.4) we obtain the equation

$$\begin{aligned} &f[(t_1a_1+t_2b_1)e_1]+f[(t_1a_2+t_2b_2)e_2] \\ &=f(t_1a_1e_1)+f(t_2b_1e_1)+f(t_1a_2e_2)+f(t_2b_2e_2) \end{aligned} \qquad (2.5)$$

Denoting $f(ce_i)=g_i(c\ )$, we have from (2.5)

$$\begin{aligned} &g_1(t_1a_1+t_2b_1)+g_2(t_1a_2+t_2b_2) \\ &=A(t_1) \quad + \quad B(t_2) \end{aligned} \qquad (2.6)$$

Where A and B are suitably defined functions. Applying lemma 1.5.1 of KLR (page 29), g_1 and g_2 are polynomials of degree not greater than two, unless $b_1=1$, $b_2=0$, in which case g_1 is linear. Under the conditions of lemma 1 we can choose e_1,e_2 and h_1, h_2 to be two different orthogonal pairs. In which case $g_i(c)$ is possibly quadratic in c, say .

$$u_ic^2+V_ic+W_i \qquad (2.7)$$

Where u_i,v_i and w_i may depend on e_i.

Let x,y be may two vectors belongs to V and e_1,e_2 be a pair of orthonormal vectors in the plane determined by x,y . Then

$$X=c_1e_1+c_2e_2 \quad , \quad Y=d_1e_1+d_2e_2 .$$

Using (2.7) for i=1 and 2 , it is easy to show that

$$2f(x+y)+f(x-y)+f(y-x)=2[f(x)+f(-x)+f(y)+f(-y)] \qquad (2.8)$$

Valid for all $X,Y\in V$. The equation (2.8) can be written in the form

$$f(x+y)+g(x-y)=A(x)+B(y) \qquad \forall x,y\in V \qquad (2.9)$$

Where A and B are suitably defined functions . from (2.9) we conclude that f is a polynomial of degree not greater two .lemma 1 is established.

Note 1:

In order to establish the possibly quadratic nature of the solution of (2.2) , we used only the condition that on any given two dimensional subspace , for any given pair of vectors e_1,e_2 there exist another pair h_1, h_2 such that (2.3) and (2.4) hold . Such a condition may replace the condition that (2.2) holds for all orthogonal vectors.

Note 2:

It is seen from the proof of lemma1 that the solution of (2.2) is linear if it is satisfied for all pairs x,y such that $(X,Y)l\|X\| \quad \|Y\|=\text{constant}\neq 0$, since in such a case there is the possibility of choosing $b_1=1$, $b_2=0$ in (2.6) . It appears that the vanishing of the inner product is a crucial condition which produce a nonlinear solution.

Note 3:

If the equation (2.2) is valid for all orthogonal pairs x,y in a neighborhood V^0 of the origin in V , then f is possibly a quadratic function in V^0.

3. Characterization through independence of linear form:

As an application of lemma 1 we have theorem 3.1 characterizing a wishart distribution on a real Hilbert space.

Theorem 3.1:

Let $X=\sum_{j=1}^{k}<x_j,x_j>$ be a random variable defined on a real Hilbert space H such that $Y_1(x,a)$ and $Y_2(b,x)$ are i.d. (independently distribution) for all $a,b\in H$ such that (a,b)=0. Then X has a wishart distribution on H.

Proof :

By hypothesis

$$E[\exp(it_1y_1+it_2y_2)]=E(\exp it_1y_1).E(\exp it_2y_2). \qquad (3.1)$$

Substituting fory_1 andy_2 in terms of x and denoting by C(t)=E[exp i(t,x)], the c.f. (characteristic function) of x we obtain from (3.1)

$$C(t_1a+t_2b)=C(t_1a) - C(t_2b) \qquad \forall(a,b)=0 \qquad (3.2)$$

or

$$C(x+y)=C(x) - C(y) \qquad \forall(x,y)=0 \qquad (3.3)$$

In terms of f(x)= log c(x) defined in a neighborhood V^0 of the origin

$$f(x+y)= f(x)+f(y) \qquad (3.4)$$

$\forall x, y \in V^0$ where $Y = \sum_{i=1}^{k} < y_i, y_i >$ such that (x,y)=0

Applying lemma1, f(x) is a polynomial of the second degree utmost in V^0.
Hence C is the c.f of a wishart distribution and the Theorem is established
Theorem 3.1requires that (a,x) and (b,x) should be i.d. whenever (a,b)=0 ,

Theorem 3.2:
Let X be abivariate r.v.(with components $\in R^1$) , A and B be given 2×2 nonsingular matrices such that $A^{-1}B$ or $B^{-1}A$ has no zero element . If the components of BX are independtly distributed and so also the components of AX , then X has a b.n.d.(bivariate normal distribution).

Proof :
Let Y=BX , where ($Y = \sum_{i=1}^{k} < y_i, y_i >$, $X = \sum_{j=1}^{k} < x_j, x_j >$)

Then $AX = AB^{-1}Y = CY(say)$.By hypothesis the components of Y are independent and so also are the component of CY .Hence by applying Darmois-Skitovic theorem , the components of Y are a wishart distributed in which case.

$X = B^{-1}Y$ has a b.n.d.

Note
Theorem 3.2 can be considered in a more general context where the two

components of X are r.v.'s belonging to a more general space than R^1.
Theorem 3.2 shows that to assert bivariate a normal distribution of X, it was only necessary to find just two pairs of linear functions such that the elements in each pair are independently distributed. For a very wide class of pairs of linear functions.

4.Characterization Through Regression:

Let X_1 ,X_2 be independent and identically distributed p-vector r.v.'s such that

$$E(X_1 - AX_2 \setminus X_1 + B'X_2) = 0 \qquad (4.1)$$

For given nonsingular matrices A and B . What can be said about the distribution of X_1 ?we may suppose that X_1 has first moment.
The problem was solved in the special case when $A = B^{-1}$, A is symmetric and the egen values do not take values ± 1. Under these conditions it was shown by Rao [11] that X_1 has an m.n.d.
in the present paper we shall consider a few other cases . A study of the problem (4.1) when p=2 was made by Klebanov [7] and solution have been obtained in a number of particular cases.

Let g(t) be the S.c.f. of X_1 and define by $G(t) = \frac{\partial g}{\partial t}$, the vector of partial derivates of g with respect to the components of t . Then it is easy to show that (4.1)

$$G(t) = AG(Bt) \text{ or } A^{-1}G(t) = G(Bt) \qquad (4.2)$$

The problem is to solve the equation (4.1) for g(t) . It is interesting to note that equation of the type (4.2) occur in a study of " optimization problems and structural stability by Andronov and Pontrjagin (see Robbins [14]) .In their problem A^{-1} =(D say) and B stand for c' diffeomorphisms from a smooth manifold μ onto itself and G is a homeomorphisms such that $D \cdot G = G \cdot B$ in which case B and D are said to be topologically conjugate. Theorem 4.1 considers that special case when A^{-1} =B

Theorem 4.1:

$$\text{Let } B = \delta_1 Q_1 P_1' + \dots + \delta_r Q_r P_r' = QDP' \tag{4.3}$$

Be the singular value decomposition of B , where Q_i and P_i are matrices of order $P \times m_i$ with orthonormal vectors corresponding to the multiplicity m_i of the singular value δ_i . If A= B^{-1} then the solution g(t) of (4.2) is of the form

$$g(pt) = h_1(t_1) + \dots\dots + h_r(t_r) \tag{4.4}$$

$$g(Qt) = \delta_1{}^2 h_1(\delta_1{}^{-1} t) + \dots\dots\dots + \delta_r{}^2 h_r(\delta_r{}^{-1} t) \tag{4.5}$$

Where $t' = (t_1' : \dots\dots : t_r')$ and t_i is a vector of order m_i , and h_i are suitable functions. Then h_i in (4.4) and (4.5) satisfy the equation.

$$\sum_i^r h_i(P_1 Q_1 t_1 + \dots\dots\dots + P_r Q_r t_r) = \sum_i^r \delta_i{}^2 h_i(\delta_i{}^{-1} t_i) \tag{4.6}$$

Proof :

Substituting $B = QDP'$ in (4.2) with $A^{-1} = B$, we have

$$QDP'G(t) = G(QDP't)$$

$$\Rightarrow QDP'G(Pt) = G(QDt)$$

$$\Rightarrow D^2 \frac{\partial}{\partial t} g(Pt) = G(QDt)$$

$$\Rightarrow \delta_i{}^2 \frac{\partial}{\partial t_i{}^2} g(Pt) = \frac{\partial}{\partial t_i} g(QDt) \qquad , i = 1, \dots, r \tag{4.7}$$

$$\Rightarrow \delta_i{}^2 g(Pt) = g(QDt) + f_i(t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_r) \tag{4.8}$$

From (4.8) it is easy to show that f_i is of the form $f_i = \sum_i^r (\delta_i{}^2 - \delta_j{}^2) h_j(t)$ where h_j are suitable function .Then

$$g(Pt) = \sum h_i(t_i), \tag{4.9}$$

And consequently

$$g(Qt) = \sum \delta_i{}^2 h_i(\delta_i{}^{-1} t_i). \tag{4.10}$$

The equation (4.6) follows from (4.9) and (4.10) , and Theorem (4.1) is established.
Note that the equation (4.6) is of the form discussed by Khatri and Rao[5] but not solved in generality .It appears that the nature of the solution of (4.6) depends on the relationships among the singular values $\delta_1, \dots, \delta_r$ and on the values of the matrices $P_i' Q_j$ .We shall consider some special cases.

(i) suppose that all singular values of B are of multiplicity one. If each column of $P'Q$ or $Q'P$ has at least two non-zero elements then $\sum_i < X_i, X_i >$ has a w.d.

Proof :

From (4.4) ,(4.5) , we conclude that the components of $\sum_{i=1}^{k} P' < X_i, X_i >$ and $\sum_{i=1}^{k} Q' < X_i, X_i >$ are independently .If $Y = \sum_{j=1}^{k} P' < X_j, X_j >$, then the components of $Q'PY$ are indepently distributed . If each column of $Q'P$ has at least two non-zero elements, then by an application of Darmoise-Skitovio theorem each component of Y is Wishart-distribution. Hence $\sum_i < X_i, X_i >$ has a w.d. Similarly if every column of $P'Q$ has at least two non-zero elements, then $\sum_i < X_i, X_i >$ has an w.d., since the components of $P'QZ$ are i.d., where $Z = \sum_{i=1}^{k} Q < X_i, X_i >$.

The result (i) established.

In order to understand the complications that may arise in solving the equation (4.6) for the general case

Let us consider p=2 .Then (4.6) becomes

$$h_1(a_{11}t_1 + a_{12}t_{12}) + h_2(a_{21}t_1 + a_{22}t_2) = \delta_1{}^2 h_1(s_1{}^{-1}t_1) + \delta_2{}^2 h_2(\delta_2{}^{-1}t_2) \qquad (4.11)$$

When there are two distinct singular values and when there are two distinct values and

$$h(t) = \delta^2 h(\delta^{-1}t) \qquad (4.12)$$

When there is only one , when in (4.12) t is a 2-vector .

(ii) consider (4.11) . If none of the a_{ij} is zero , then h_1 and h_2 are quadratic polynomials and $\sum_i < X_i, X_i >$ has a b.w.d.

If $a_{12}=0$, then h_1 and h_2 are quadratic polynomials provided δ_1 and δ_2 are different from ± 1 and hence $\sum_i < X_i, X_i >$ has a w.d.

If $a_{11}=0$, then

$$h_1(t) = \delta_2{}^2 h_2(\delta_2{}^{-1}t) \quad , \quad h_2(t) = \delta_1{}^2 h_1(\delta_1{}^{-1}t) \qquad (4.13)$$

A gain h_1 and h_2 are quadratic polynomials if $\delta_1\delta_2 \neq \pm 1$. if $\delta_1\delta_2 = \pm 1$, then h_1 can be arbitrary and h_2 depends on h_1 as in (4.13).

The results of (ii) are easy to prove.

(iii)consider (4.12). In this case $\sum_i < X_i, X_i >$ follows abivariate stable law of the type described by Eaton and Pathk[1].

Thus we have a complete solution for p=2 and for general P in a very special case.

The result in (i) for general P can be extended to cases where in some columns of $P'Q$ there is only one non-zero element. A number of cases may have to be considered some leading to quadratic solutions for all h_i and some to arbitrary solutions to a subset of h_i .However. The general problem may be stated as follows

References

[1] Eaton Morris, L.and Pathak, P.K.(1969). A characterization of the normal law in Hilbert space. Sannkya 31,259-268.

[2] Dan Geiger, David Heckerman(1980). A characterization of the bivariate Wishart distribution probality and Mathematical Statistics Vol. 18, Fasc. 1 (1980), pp. 119–131

[3] Hede, C. C. (1970). Characterization of normal law by the symmetry of a certain conditional distribution. Sankyha 32,115-118.

[4] Kagan, A .M., Linnik, Yu.V. and Rao , C.R.(1972,1973).Characterization problems of Mathematical Statistics. Russian edition , 1972, Academy Nauk, Mosscow ; English edition , 1973, John Wiley , New York.

[5] Kharri , C.G.and Rao , C.R.(1972). Functional equations and characterization of probability laws through linear functions of random variables . J. Multivariate Analysis 2,162-173.

[6] Kharri , C.G.and Rao , C.R. (1974).Characterization of multivariate normal distribution by constancy of regression of one linear function on another (under preparation).

[7] Klebanov, L.B.(1970).on a functional equation . Sankhya 32,387-392.

[8] Linnik, Yu,V. .(1953). Linear forms ans statistical criteria. I and II (in Russian). Ukrain .Mat .Zhurnal 5,207-243 and 247-290.

[9] Ramachandran , B. and Rao,C.R.(1968). Some results on characterizations of the normal and generalized stable laws.Sankyha A 30, 125-140.

[10] Ramachandran , B. and Rao,C.R. (1970).Solutions of a functional equation arising in some regression problems and a characterization of the Cauchy law, Sankhya A 32,1-30.

[11] Rao ,C.R.(1969).Some characterizations of the multivariate normal distribution. Multivariate Analysis-II.P321-328 Ed .P.R.Krishnaih , Academic Press ,New York.

[12] Rao ,C.R.((1973). Linear Statistical Inference and its Applications. Jhon Wiley , New York.

[13] Rao ,C.R.(1974c) .Functional equations and characterization of probability distributions.(Invited address,world Congress of Mathematics, 1974, Vancouver).

[14] Robbins ,J.W.(1971).A structural stability theorem.T.A.M.S.,447-493.

بعض المسائل في تميزات توزيع وشارت

نادية عبد حبيب الموسوي

قسم الرياضيات/ كلية علوم الحاسبات والرياضيات/ جامعة ذي قار

الخلاصة

في ظل النموذج الخطي متعدد المتغيرات{ $y,Xb,\sum xv$} عدد من التميزات لتوزيع X_2 قد تمت دراستها مبينة على خصائص المقلمات Y_1 و Y_2 حيث Y_1 و Y_2 هما دالتان خطيتان في هذا البحث ندرس تعميمات التي تتطلب تقنيات حديثة