EXTRACTA MATHEMATICAE Vol. **12**, Núm. 1, 15 – 18 (1997)

Characterizing the General Multivariate Normal Distribution through the Conditional Distributions

Ramón Ardanuy Albajar and Jesús F. López Fidalgo

Departamento de Matemática Pura y Aplicada, Univ. de Salamanca, Pl. Merced, 1–4, E37008-*Salamanca, Spain*

(Research paper presented by M. Molina)

AMS *Subject Class.* (1991): 60E05, 62E10, 60E10 *Received April* 8, 1997

In this paper we give a characterization of the multivariate normal distribution through the conditional distributions in the most general case, which include the singular distribution. In fact, the necessary and sufficient condition for the joint distribution of two random vectors X and Y to be normal is that X must be normal and also the conditional distribution of Y when $X = x$ must be normal with its mean vector as an affine transformation of x and the dispersion matrix constant. The necessary condition in the nonsingular case is well known ([1], [2], [3]). Muirhead ([4]) proves the necessary condition for the most general case using generalized inverses.

The originality of this paper lays on the proof of the sufficient condition in the most general case, including the singular case. For making this proof the use of the characteristic function will be needed. In the nonsingular case the proof could be made with the density function.

The nonsingular normal distribution may be defined through the density function. For the general case, including the singular one, the characteristic function may be used for its definition.

Let Σ be a real, symmetric and nonnegative definite matrix $n \times n$ and let $\mu = (\mu_1, \dots, \mu_n)^{\mathrm{T}}$ be a vector of real components. It will be said that the random vector $X = (X_1, \dots, X_n)^{\mathrm{T}}$ has a multivariate normal distribution of mean vector μ and covariance matrix Σ if its characteristic function is,

$$\varphi(t) = \mathrm{E}\Big[\exp\Big(it^{\mathrm{T}}X\Big)\Big] = \exp\Big(it^{\mathrm{T}}\mu - \frac{1}{2}t^{\mathrm{T}}\Sigma t\Big).$$

We will write $X \equiv \mathcal{N}_n(\mu, \Sigma)$. It may be proved that the mean and the covariance matrix are μ and Σ. The multivariate normal distribution is then concentrated on an affine subspace of $\Re^n$, whose dimension is the rank of Σ. In the nonsingular case the density function is,

$$f(x|\mu,\Sigma) = \frac{1}{\sqrt{(2\pi)^n \det\Sigma}} \exp\Big[-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\Big], \quad x \in \Re^n.$$

Some interesting properties are,

THEOREM 1. *Let* $X \equiv \mathcal{N}_n(\mu_X, \Sigma_X)$ *and* $Y = AX$, *where* A *is a matrix* $k \times n$. *Then the random vector* $(Y_1, \ldots, Y_k)^T$ *has a multivariate normal distribution with mean vector* $\mu_Y = A\mu_X$ *and covariance matrix* $\Sigma_Y = A\Sigma_X A^T$.

That is, every linear combination of the components of a multivariate normal distribution is normal. This property characterized the multivariate normal distribution and therefore it may be done as an alternative definition, that includes the singular case.

THEOREM 2. *A random vector is normal if, and only if, every linear combination of the components of a multivariate normal distribution is normal.*

THEOREM 3. *Let* $X_1, \ldots, X_k$ *be independent random vectors such that* $X_j \equiv \mathcal{N}_n(\mu_j, \Sigma_j)$, $j = 1, \ldots, k$, *then* $X_1 + \cdots + X_k \equiv \mathcal{N}_n(\mu, \Sigma)$, *where* $\mu = \mu_1 + \cdots + \mu_k$ *and* $\Sigma = \Sigma_1 + \cdots + \Sigma_k$.

The following theorem is the main result in this paper,

THEOREM 4. *The two following statements hold,*

1. *If the random vector* $Z = \begin{pmatrix} X \\ Y \end{pmatrix}$ *has a multivariate normal distribution* $\mathcal{N}_n(\mu, \Sigma)$, *where,*

$$\mu = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \quad \textit{and} \quad \Sigma = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix},$$

then the conditional distribution of Y *given* X *has a normal distribution of mean* $E(Y|X) = \mu_Y + B_0(X - \mu_X)$ *and covariance matrix* $\Sigma_{YY} - B_0\Sigma_{XY}$, *where* B_0 *is a solution of the equation* $\Sigma_{YX} = B\Sigma_{XX}$. *That is, the expectation of the conditional* Y *given* X *is an affine transformation of* X *and the covariance matrix of the conditional of* Y *given* X *is constant.*

2. *If* $X \equiv \mathcal{N}_r(\mu_X, \Sigma_{XX})$ *and the conditional distribution of* Y *given* X *is a* s*-dimensional normal distribution of mean* $a + BX$ *and constant covariance matrix* Σ_0*, then the joint distribution of* (X, Y) *is a* n*-dimensional normal distribution with* $n = r + s$*, whose mean vector and covariance matrix are,*

$$\begin{aligned} \mu &= \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} = \begin{pmatrix} \mu_X \\ a + B\mu_X \end{pmatrix} \\ \Sigma &= \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{XY}^T & \Sigma_{YY} \end{pmatrix} = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XX}B^T \\ B\Sigma_{XX} & \Sigma_0 + B\Sigma_{XX}B^{\mathrm{T}} \end{pmatrix}. \end{aligned}$$

Hence, Y *shall be normal as well.*

Proof. A proof of the first part of the theorem from a regression point of view can be seen in [4].

Let us see the second part of the theorem, that is the original part of this paper. For this we will compute the joint characteristic function of X and Y:

$$\begin{aligned} \varphi_{X,Y}(t,u) &= \mathrm{E}[\exp(it^{\mathrm{T}}X + iu^{\mathrm{T}}Y)] = \mathrm{E}\{\mathrm{E}[\exp(it^{\mathrm{T}}X + iu^{\mathrm{T}}Y) \mid X]\} \\ &= \mathrm{E}\{\exp(it^{\mathrm{T}}X)\mathrm{E}[\exp(iu^{\mathrm{T}}Y) \mid X]\} = \mathrm{E}[\exp(it^{\mathrm{T}}X)\varphi_{Y|X}(u)]. \end{aligned}$$

Since the conditional distribution of Y given $X = x$ is $\mathcal{N}_s(a + Bx, \Sigma_0)$, its characteristic function is

$$\varphi_{Y|X=x}(u) = \exp\Big[iu^{\mathrm{T}}(a + Bx) - \frac{1}{2}u^{\mathrm{T}}\Sigma_0 u\Big],$$

and the characteristic function of X is

$$\varphi_X(t) = \exp\Big(it^{\mathrm{T}}\mu_X - \frac{1}{2}\mu_X^{\mathrm{T}}\Sigma_{XX}\mu_X\Big).$$

Therefore, the joint characteristic function of X and Y is

$$\begin{aligned} \varphi_{X,Y}(t,u) &= \mathrm{E}\Big\{\exp\Big[it^{\mathrm{T}}X + iu^{\mathrm{T}}(a + BX) - \frac{1}{2}u^{\mathrm{T}}\Sigma_0 u\Big]\Big\} \\ &= \exp\Big[iu^{\mathrm{T}}a - \frac{1}{2}u^{\mathrm{T}}\Sigma_0 u\Big]\mathrm{E}\Big\{\exp\Big[i(t + B^{\mathrm{T}}u)^{\mathrm{T}}X\Big]\Big\}. \end{aligned}$$

Taking into account the shape of the characteristic function of X then,

$$\begin{aligned}
\varphi_{X,Y}(t,u) &= \exp\Big(iu^{\mathrm{T}}a - \frac{1}{2}u^{\mathrm{T}}\Sigma_0 u\Big)\varphi_X(t+B^{\mathrm{T}}u) \\
&= \exp\Big[iu^{\mathrm{T}}a - \frac{1}{2}u^{\mathrm{T}}\Sigma_0 u + i(t+B^{\mathrm{T}}u)^{\mathrm{T}}\mu_X \\
&\quad -\frac{1}{2}(t+B^{\mathrm{T}}u)^{\mathrm{T}}\Sigma_{XX}(t+B^{\mathrm{T}}u)\Big] \\
&= \exp\Big(is^{\mathrm{T}}\mu - \frac{1}{2}s^{\mathrm{T}}\Sigma s\Big),
\end{aligned}$$

where,

$$s = \begin{pmatrix} t \\ u \end{pmatrix}; \quad \mu = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} = \begin{pmatrix} \mu_X \\ a + B\mu_X \end{pmatrix};$$

$$\Sigma = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{XY}^{\mathrm{T}} & \Sigma_{YY} \end{pmatrix} = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XX}B^{\mathrm{T}} \\ B\Sigma_{XX} & \Sigma_0 + B\Sigma_{XX}B^{\mathrm{T}} \end{pmatrix}.$$

Thus the joint distribution of X and Y is normal $(r+s)$-dimensional with the mean vector and the covariance matrix given above. ∎

References

[1] ANDERSON, T.W., "An Introduction to Multivariate Statistical Analysis", 2nd Ed, John Wiley and Sons, New York, 1984.

[2] DE GROOT, M.H., "Optimal Statistical Decisions", McGraw-Hill, New York, 1970.

[3] FERGUSON, T.S., "Mathematical Statistics: A Decision Theoretic Approach", Academic Press, New York, 1967.

[4] MUIRHEAD, R.J., "Aspects of Multivariate Statistical Theory", John Wiley and Sons, New York, 1982.