Filomat 33:13 (2019), 4137–4150
https://doi.org/10.2298/FIL1913137G

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

The multiparameter t'distribution

Emna Ghorbel[a], Mahdi Louati[b]

[a]*Faculty of Sciences of Sfax, B.P. 1171, CP 3038, University of Sfax, Tunisia.*
[b]*Sfax National School of Electronics and Telecommunications, B.P. 1163, CP 3018, Tunisia.*

Abstract. This research paper stands for an extension to the multivariate t'distribution introduced in 1954 by Cornish, Dunnett and Sobel, namely the multiparameter t'distribution. This distribution is expressed in two different ways. The first way invests the mixture of a normal vector with a natural extension to the Wishart distribution, that is the Riesz distribution on symmetric matrices. The second one rests upon the Cholesky decomposition of the Riesz matrix. An algorithm for generating this distribution is investigated using the Riesz distribution arising obtained through not only the distribution of the empirical normal covariance matrix for samples with monotone missing data but also through Cholesky decomposition. In addition, Some fundamentals properties of the multiparameter t'distribution such as the infinite divisibility are identified. Besides, the Expectation Maximization algorithm is used to estimate its parameters. Finally, the performance of these estimators is assessed by means of the Mean Squared Error between the true and the estimated parameters.

1. Introduction

Towords 1950's, Cornish [5] and Dunnett and Sobel [10] set forward the notion of multivariate t'distribution which is a natural generalization of the classical Student distribution. It stands for central focus of statistical inference. This distribution is elliptically symmetric and performs a basic role in terms of statistical analysis of multivariate data. It is worth noting that the tails of the multivariate t'distribution are more realistic. From this perspective, this distribution can provide a more viable alternative for the classical multivariate analysis with regard to the real data rather than the multivariate normal one. Recently, some interesting applications have been set forward in novel areas such as cluster and discriminant analysis (Andrews and McNicholas [2]), multiple regression (Arashi and Tabatabaey [3]), Bayesian prediction approach (Chien [4]) and robust projection indices (Nason [23]). From a theoretical point of view, the multivariate t' distribution has attracted the attention of researchers and whetted their interest. Therefore, several research works have been oriented towards this direction. For instance, Cornish [6] focused on the link between the multivariate t'distribution and the set of normal sample deviates. Dickey [8] explored the characterizations of the multivariate t'distribution and its inverse. Kotz and Nadarajah [19] reported that there exist few forms of multivariate t'distributions. Moreover, Lin [20] demonstrated that on the one hand this distribution arises as a mixture of a normal vector with a Chi-squared variable. On the other hand, it represents a mixture of a normal vector with a Wishart matrix. In this paper, our central focus is upon the natural extensions

2010 *Mathematics Subject Classification*. Primary 44A10; Secondary 62H12
Keywords. Cholesky decomposition; Expectation Maximization algorithm; Multivariate t'distribution; Riesz distribution.
Received: 25 January 2019; Revised: 14 August 2019; Accepted: 29 September 2019
Communicated by Aleksandar Nastić
Email addresses: emnaghorbel.fss@gmail.com (Emna Ghorbel), mahdi.louati@ipeis.rnu.tn (Mahdi Louati)

of these two forms. This allows us to characterize a natural extension to the multivariate t'distribution, namely the multiparameter t'distribution. For this purpose, we basically use the Riesz distribution in the space of symmetric matrices which represents the natural extension of the Wishart one.

Over the last decades, there has been a spate of interest in the Wishart distribution owing to its use in the graphical Gaussian models. Nobody can deny its crucial role in the context of multivariate statistics as it arises in the estimation of the covariance matrices and stands for the conjugate prior of the precisian matrix. It is significant to mention that the Wishart probability distribution is initially identified as a multivariate extension of the Chi-squared distribution. Indeed, it displays the sums of the squares of multivariate normal distribution. Nevertheless, only a few results have been obtained recently (see for instance Graczyk *et al.* [14] and Von Rosen [25]). Furthermore, this distribution proves to lack flexibility in certain situations as it has a real shape parameter. In order to overcome this deficiency based on the Gindikin [13] theorem, Hassairi and Lajmi [15] introduced an important extension which is the Riesz distribution. Recently, several interesting results concerning the latter emerged. In fact, Andersson and Klein [1] extended the definition of the Riesz distribution to homogeneous cones in association with graphical models. Moreover, Díaz-Garcia [7] gave the first two moments of the Riesz distribution using their characteristic functions. Hassairi and Louati [16, 17] and Louati [21] showed that the approach based on the mixture of distributions may be extended to the Riesz model. This allowed to build new families of distributions that are useful in modeling. Besides, Kammoun *et al.* [18] investigated the estimation of the scale parameter for the Riesz distribution. Furthermore, Louati and Masmoudi [22] used the Weyl's integration formula and the Cholesky decomposition to provide stands for the first moment of the inverse Riesz distribution which represents the natural extension of the inverse Wishart one. Over the past few decades, the applications of the Wishart distributions have been growing in different fields. They were substantial particularly in image processing, wireless communications, recognition and clustering, etc. The assumption that its natural extension may lead to several applications that are more general than the ones related to the Wishart, seemed to be the basic impetus motivating researchers for a further and thorough investigation of theoretical and applied aspects of the Riesz distribution. Within this framework, we attempt to prove that the Riesz distribution can be drawn from the square of some matrix connected to the multivariate normal distribution with missing data.

Departing from the definition and certain results concerning the Riesz distribution, we identify in Section 3, the notion of the degree of freedom of the Riesz distribution. Using the Laplace transform, we demonstrate that this distribution, in the discrete case, is related to the normal matrix with monotone missing data. We characterize the Riesz distribution, in the continuous case, by means of its Cholesky decomposition. This enables us to provide algorithms for generating the Riesz distribution. In Section 4, the multiparameter t'distribution is characterized as a mixture of normal vectors with respect to some derived Riesz matrices and certain fundamental related properties are investigated. Finally, in Section 5, we display the Expectation Maximization (EM) algorithm for maximum likelihood (ML) estimation with known degree of freedom. A simulation study is incorporated to illustrate the proposed algorithm.

2. Preliminaries

In order to present our results in their most general form, we first need to recall some notations and review some characteristic properties concerning the Riesz distributions. The notations used in this paper mostly follow those of Louati [21]. Let E be the set of (r, r) real symmetric matrices equipped with the scalar product $\langle x, y\rangle = \mathrm{tr}(xy)$. We denote by Ω the cone of positive definite elements of E. The definition of the Riesz distribution is based on the notion of generalized power of x in Ω which defined, for $s = (s_1, s_2, \dots, s_r) \in \mathbb{R}^r$, by

$$\Delta_s(x) = \Delta_1(x)^{s_1-s_2}\Delta_2(x)^{s_2-s_3}\dots\Delta_{r-1}(x)^{s_{r-1}-s_r}\Delta_r(x)^{s_r},$$

where for all $k \in \{1, \dots, r\}$, $\Delta_k(x)$ is the determinant of the (k, k) sub-matrix $P_k(x) = (x_{ij})_{1\le i,j\le k}$ of x.
We denote by $\Delta^*_k(x)$ the determinant of the (k, k) sub-matrix $P^*_k(x) = (x_{ij})_{r-k+1\le i,j\le r}$ of x. For all $s \in \mathbb{R}^r$,

$$\Delta^*_s(x) = (\Delta^*_1(x))^{s_1-s_2}(\Delta^*_2(x))^{s_2-s_3}\dots(\Delta^*_{r-1}(x))^{s_{r-1}-s_r}(\Delta^*_r(x))^{s_r}. \tag{1}$$

Let $\mathcal{T}_l^+$ be the set of lower triangular matrices with positive diagonal elements. For $x \in \Omega$, $u \in \mathcal{T}_l^+$ and $s \in \mathbb{R}^r$ we have

$$\Delta_s(uxu^t) = \Delta_s(uu^t)\Delta_s(x) \text{ and } \Delta_s(u^{-1}x(u^{-1})^t) = \Delta_{-s}(uu^t)\Delta_s(x) \tag{2}$$

for more details, the reader can see Faraut and Korányi [11], p. 114 and Hassairi and Lajmi [15].
The Riesz probability distribution on Ω with scale parameter σ in Ω and shape parameter $s = (s_1, s_2, \ldots, s_r) \in \prod_{i=1}^r](i-1)/2, +\infty[$, has the following density function

$$R(s,\sigma)(dx) = \frac{e^{-<\sigma,x>}\Delta_{s-\frac{r+1}{2}}(x)}{\Gamma_\Omega(s)\Delta_s(\sigma^{-1})}\mathbf{1}_\Omega(x)(dx), \tag{3}$$

where

$$\Gamma_\Omega(s) = (2\pi)^{\frac{r(r-1)}{4}}\prod_{i=1}^r \Gamma(s_i - (i-1)/2)$$

is the multivariate gamma function. Its Laplace transform is equal to

$$L_{R(s,\sigma)}(\theta) = \frac{\Delta_s((\sigma-\theta)^{-1})}{\Delta_s(\sigma^{-1})}, \text{ for all } \theta \in \sigma - \Omega. \tag{4}$$

In particular, if $s_1 = s_2 = \ldots = s_r = p > (r-1)/2$, $R(s,\sigma)$ is reduced to the Wishart distribution with parameters $p \in](r-1)/2, +\infty[$ and $\sigma \in \Omega$ defined by

$$W(p,\sigma)(dx) = \frac{e^{-<\sigma,x>}\det(x)^{p-\frac{n}{r}}}{\Gamma_\Omega(p)\det(\sigma^{-p})}\mathbf{1}_\Omega(x)(dx).$$

In this case, the Laplace transform evaluated in $\sigma - \Omega$, is given by

$$L_{W(p,\sigma)}(\theta) = \left(\frac{\det(\sigma-\theta)}{det(\sigma)}\right)^{-p}. \tag{5}$$

3. Simulation of the Riesz distribution

This section is entirely devoted to the simulation of the Riesz distribution with scale parameter $\frac{\sigma^{-1}}{2}$ and shape parameter $\frac{s}{2}$. For this purpose, we demonstrate that this distribution, when the parameter s is an integer vector, is closely related to a normal matrix for samples with monotone missing data.

Let $(Y_1, Y_2, \ldots, Y_r)^t$ be a random vector with multivariate normal distribution $N(0,\sigma)$. Suppose that we have k_1 observations of $(Y_1, Y_2, \ldots, Y_r)^t$, k_2 observations of $(Y_2, \ldots, Y_r)^t$ and so k_r of Y_r. Assume that for all $j \in \{1, 2, \ldots, r\}$, we have $k_j > 0$ and consider $s_j = k_1 + k_2 + \ldots + k_j$. We denote by Y the matrix of observations where the missing data are replaced by zero. More precisely, we have

$$\begin{pmatrix} y_{1,1} & \ldots & y_{1,s_1} & 0 & & 0 & 0 & \ldots & \ldots & \ldots & 0 \\ y_{2,1} & \ldots & y_{2,s_1} & y_{2,s_1+1} & \ldots & y_{2,s_2} & 0 & \ldots & \ldots & \ldots & 0 \\ \vdots & \ldots & \vdots & \vdots & \ldots & \vdots & \ldots & \vdots & 0 & \ldots & \vdots \\ y_{r-1,1} & \ldots & y_{r-1,s_1} & y_{r-1,s_1+1} & \ldots & y_{r-1,s_2} & \ldots & y_{r-1,s_{r-1}} & 0 & \ldots & 0 \\ y_{r,1} & \ldots & y_{r,s_1} & y_{r,s_1+1} & \ldots & y_{r,s_2} & \ldots & y_{r,s_{r-1}} & y_{r,s_{r-1}+1} & \ldots & y_{r,s_r} \end{pmatrix}. \tag{6}$$

Next, we characterize the distribution of the random matrix $R = YY^t$, where Y^t denotes the transpose of the matrix Y.

Theorem 3.1. *The random matrix $R = YY^t$ is Riesz distributed with shape parameter $\frac{s}{2} = (\frac{s_1}{2}, \ldots, \frac{s_r}{2})$, scale parameter $\frac{\sigma^{-1}}{2} \in \Omega$ and degree of freedom $\nu = (s_1, s_2 - s_1, \ldots, s_r - s_{r-1}) \in \mathbb{N}^* \times \mathbb{N}^{r-1}$.*

Proof. The Laplace transform of the matrix $R = YY^t$ is indicated by

$$L_R(\theta) = \mathbb{E}\left(e^{\langle\theta,R\rangle}\right) = \mathbb{E}\left(e^{\langle\theta,YY^t\rangle}\right) = \mathbb{E}\left(\prod_{i=1}^{r} e^{\langle\theta,Z_iZ_i^t\rangle}\right),$$

with for all $i \in \{1, 2, \ldots, r\}$,

$$Z_i = \begin{pmatrix} 0_{(i-1)\times(s_i-s_{i-1})} \\ T_i \end{pmatrix}, \text{ where } T_i = \begin{pmatrix} x_{i,s_{i-1}+1} & \ldots & \ldots & x_{i,s_i} \\ \vdots & \vdots & \vdots\vdots & \\ x_{r,s_{i-1}+1} & \ldots & \ldots & x_{r,s_i} \end{pmatrix}.$$

Using the independence of the columns of T_i, we deduce that

$$L_R(\theta) = \mathbb{E}\left(\prod_{i=1}^{r} e^{\langle P^*_{r-i+1}(\theta),T_iT_i^t\rangle}\right) = \prod_{i=1}^{r} \mathbb{E}\left(e^{\langle P^*_{r-i+1}(\theta),T_iT_i^t\rangle}\right) = \prod_{i=1}^{r} L_{T_iT_i^t}(P^*_{r-i+1}(\theta)).$$

Since for all $i \in \{1, 2, \ldots, r\}$, $T_iT_i^t$ is Wishart distributed $W\left(\frac{s_i-s_{i-1}}{2}, \frac{P^*_{r-i+1}(\sigma)}{2}\right)$ (see Wishart [26]), then according to (1) and (5), we have

$$L_R(\theta) = \prod_{i=1}^{r} \left(\frac{\det\left(\frac{P^*_{r-i+1}(\sigma)}{2} - P^*_{r-i+1}(\theta)\right)}{\det\left(\frac{P^*_{r-i+1}(\sigma)}{2}\right)}\right)^{-\frac{s_i-s_{i-1}}{2}} = \frac{\Delta^*_{-\frac{s^*}{2}}(\frac{\sigma}{2} - \theta)}{\Delta^*_{-\frac{s^*}{2}}(\frac{\sigma}{2})},$$

where $\frac{s^*}{2} = (\frac{s_r}{2}, \frac{s_{r-1}}{2}, \ldots, \frac{s_1}{2})$. This, with (4) and the fact that

$$\Delta_s(x^{-1}) = \Delta^*_{-s^*}(x),$$

(see Hassairi and Lajmi [15]), gives the desired result. □

Based on (6) and Theorem 3.1, we display an algorithm for generating the Riesz distribution with an integer parameter s, namely the discrete Riesz distribution.

Algorithm 3.2. *(The discrete Riesz simulation)*

Step 1: *For all $i \in \{1, 2, \ldots, r\}$, we generate k_i independent centered multivariate normal random variable with parameter σ*

Step 2: *For each simulated vector $Y^i = (Y^i_1, Y^i_2, \ldots, Y^i_r)^t$, we replace the set $(Y^i_1, .., Y^i_{i-1})$ by zeros.*

Step 3: *Calculate $R = YY^t$.*

Next, we exhibit a simulation of the continuous Riesz matrix which corresponds to the Riesz distribution with a shape parameter s in $\prod_{i=1}^r](i-1)/2, +\infty[$. It is well known that, if Z is an element of the cone Ω of definite positive symmetric matrices, then using the Cholesky decomposition, there exists a unique random matrix $u = (u_{ij})_{1\le i,j\le r} \in \mathcal{T}_l^+$ such that $Z = uu^t$. Besides, Z follows a Riesz distribution $R(s, I_r)$, if, and only if,

$$\begin{cases} u_{ii}^2 \text{ follows } \gamma(s_i - (i-1)/2, 1) \text{ for all } 1 \le i \le r \\ u_{ij} \text{ follows } N(0, 1/2) \text{ for all } 1 \le j < i \le r. \end{cases}$$

Consider $\sigma \in \Omega$. Then, using the Cholesky decomposition, there exists a unique v in $\mathcal{T}_l^+$ such that $\sigma^{-1} = vv^t$. It follows that, $X = vZv^t$ is a Riesz matrix with shape parameter s and scale parameter σ. In fact, The Laplace transform of X is expressed by

$$L_X(\theta) = \mathbb{E}\left(e^{\langle\theta,X\rangle}\right) = \mathbb{E}\left(e^{\langle\theta,vZv^t\rangle}\right) = \mathbb{E}\left(e^{\langle v^t\theta v,Z\rangle}\right) = L_Z(v^t\theta v).$$

Since Z is a Riesz matrix, then according to (2), (4) and the fact that $\sigma^{-1} = vv^t$, we can write

$$\begin{aligned} L_X(\theta) &= \Delta_s\left(\left(I_r - v^t\theta v\right)^{-1}\right) = \Delta_s\left(\left(v^t\left(\left(v^{-1}\right)^t v^{-1} - \theta\right)v\right)^{-1}\right) = \Delta_s\left(v^{-1}\left(\left(v^{-1}\right)^t v^{-1} - \theta\right)^{-1}\left(v^{-1}\right)^t\right) \\ &= \Delta_s\left(v^{-1}\left(v^{-1}\right)^t\right)\Delta_s\left(\left(\left(v^{-1}\right)^t v^{-1} - \theta\right)^{-1}\right) = \Delta_{-s}\left(vv^t\right)\Delta_s\left((\sigma-\theta)^{-1}\right) = \frac{\Delta_s\left((\sigma-\theta)^{-1}\right)}{\Delta_s\left(\sigma^{-1}\right)}. \end{aligned}$$

Grounded on (4), this implies that $X = vZv^t$ is $R(s, \sigma)$ (for more details, the reader can see Veleva [24]).

Algorithm 3.3. *(The continuous Riesz simulation)*

Step 1: *For all $i \in \{1, 2, \dots, r\}$, we generate u_{ii}^2 Gamma random variable with shape parameter $s_i - (i-1)/2$ and scale parameter 1.*

Step 2: *For all $1 \le j < i \le r$, we generate $\sqrt{2}u_{ij}$ standard normal random variable.*

Step 3: *Calculate $X = vuu^tv^t$.*

4. The multiparameter t'distribution

In this section, the Riesz distribution is invested to introduce the multiparameter t'distribution. It is a generalized version of the multivariate t'one and represents the marginal mixture of a normal vector with respect to a Riesz matrix. An algorithm for generating the obtained distributed vector is discussed and illustrated by a simulation study. More precisely, we have

Theorem 4.1. *Let X and Y be two random variables such that the conditional distribution of X given Y, $X|Y$ has a multivariate normal distribution $N(\mu, Y^{-1})$ with $\mu \in \mathbb{R}^r$ and Y has a Riesz distribution $R(s, \frac{\sigma}{2})$ with parameter $s \in \prod_{i=1}^r](i-1)/2, +\infty[$ and $\sigma \in \Omega$. Then, the random vector X has the following probability density function*

$$\frac{\Delta_{s+\frac{1}{2}}(((x-\mu)(x-\mu)^t + \sigma)^{-1})\Gamma_\Omega(s+\frac{1}{2})}{\pi^{\frac{r}{2}}\Delta_s(\sigma^{-1})\Gamma_\Omega(s)}(dx). \tag{7}$$

Proof. Since $X|Y$ is $N(\mu, Y^{-1})$ and Y is $R(s, \frac{\sigma}{2})$, then according to (3), the joint density of (X, Y) is expressed as

$$\begin{aligned} f_{(X,Y)}(x, y) &= f_{(X|Y=y)}(x) f_Y(y) \\ &= \frac{e^{-\frac{1}{2}<(x-\mu), y(x-\mu)>} e^{-<\frac{\sigma}{2}, y>}\Delta_{s-\frac{n}{r}}(y)}{(2\pi)^{\frac{r}{2}}\Delta_s(2\sigma^{-1})\Gamma_\Omega(s)\det(y^{-1})^{\frac{1}{2}}}. \end{aligned} \tag{8}$$

Integrating (8) with respect to y and using (3), the marginal density function of X is indicated by

$$\begin{aligned} f_X(x) &= \frac{1}{(2\pi)^{\frac{r}{2}}\Delta_s(2\sigma^{-1})\Gamma_\Omega(s)}\int_\Omega e^{-\frac{1}{2}<(x-\mu)(x-\mu)^t+\sigma, y>}\Delta_{s+\frac{1}{2}-\frac{n}{r}}(y)(dy) \\ &= \frac{\Delta_{s+\frac{1}{2}}(2((x-\mu)(x-\mu)^t + \sigma)^{-1})\Gamma_\Omega(s+\frac{1}{2})}{(2\pi)^{\frac{r}{2}}\Delta_s(2\sigma^{-1})\Gamma_\Omega(s)} \\ &= \frac{2^{\sum_{i=1}^r s_i + \frac{r}{2}}\Delta_{s+\frac{1}{2}}(((x-\mu)(x-\mu)^t + \sigma)^{-1})\Gamma_\Omega(s+\frac{1}{2})}{2^{\sum_{i=1}^r s_i}(2\pi)^{\frac{r}{2}}\Delta_s(\sigma^{-1})\Gamma_\Omega(s)}. \end{aligned}$$

Hence, we obtain the desired result. □

Subsequently, we give the definition of the multiparameter t'distribution.

Definition 4.2. *Let s be an element of $\prod_{i=1}^r](i-1)/2, +\infty[$, $\mu \in \mathbb{R}^r$ and $\sigma \in \Omega$. A random vector X of $\mathbb{R}^r$ is said to have a multiparameter t'distribution with parameters s, μ,σ and degree of freedom $\nu = (s_1, s_2 - s_1, \dots, s_r - s_{r-1})$, denoted as $X \sim MT_\nu(\mu, \sigma, r)$ if its probability density function is defined by (7).*

Remarks 4.3. *1. If, for all $i \in \{1,2,\ldots,r\}$, $s_i = p$, then* (7) *reduces to the multivariate t'distribution which is represented in Lin [20] by associating a normal vector with an independent Wishart matrix. This refers basically to the fact that in this case, the Riesz distribution $R(s,\sigma)$ reduces to the Wishart distribution $W(p,\sigma)$. Furthermore, if $r = 1$, then the probability density function given in* (7) *defines the univariate t'distribution.*
2. The conditional distribution of Y given $X = x$ has Riesz distribution with parameters $s+\frac{1}{2}$ and $\frac{1}{2}((x-\mu)(x-\mu)^t+\sigma)$. In fact, the posterior distribution of Y given $X = x$ is equal to

$$f_{(Y|X=x)}(y) = \frac{f_{(X|Y=y)}(x) f_Y(y)}{f_X(x)}.$$

Inserting the probability density function of a normal vector, the Riesz matrix given in (3) *and* (7)*, we obtain*

$$\begin{aligned} f_{(Y|X=x)}(y) &= \frac{e^{-\frac{1}{2}<(x-\mu),y(x-\mu)>} e^{-<\frac{\sigma}{2},y>} \Delta_{s-\frac{n}{r}}(y)}{(2\pi)^{\frac{r}{2}} \Delta_s(2\sigma^{-1})\Gamma_\Omega(s)\det(y^{-1})^{\frac{1}{2}}} \times \frac{\pi^{\frac{r}{2}} \Delta_s(\sigma^{-1})\Gamma_\Omega(s)}{\Delta_{s+\frac{1}{2}}(((x-\mu)(x-\mu)^t+\sigma)^{-1})\Gamma_\Omega(s+\frac{1}{2})} \\ &= \frac{e^{-\frac{1}{2}<(x-\mu)(x-\mu)^t+\sigma,y>} \Delta_{s+\frac{1}{2}-\frac{n}{r}}(y)}{2^{\frac{r}{2}} 2^{\sum_{i=1}^r s_i} \Delta_{s+\frac{1}{2}}(((x-\mu)(x-\mu)^t+\sigma)^{-1})\Gamma_\Omega(s+\frac{1}{2})}. \end{aligned}$$

This combined with (3) *provides the desired result.*

Next, some fundamental properties of the multiparameter t'distribution are exhibited. Specifically, the property of the infinitely divisible studied and the action of the affinities is explored.

Proposition 4.4. 1. *The multiparameter t'distribution is infinitely divisible.*
2. *If X is multiparameter t'distributed $MT_\nu(\mu,\sigma,r)$, then*
   (a) *If $\sigma^{-1} = vv^t \in \Omega$ then $v'(X-\mu)$ is multiparameter t'distributed $MT_\nu(0, I_r, r)$.*
   (b) *For all $a \in \mathcal{T}_l^+$ and $b \in R^r$, $a^t X + b$ is multiparameter t'distributed $MT_\nu(a^t\mu + b, a^t\sigma a, r)$.*
   (c) *For all $\alpha \in \mathbb{R}^*$, αX is multiparameter t'distributed $MT_\nu(\alpha\mu, \alpha^2\sigma, r)$.*

Proof. 1. Let $X_1,\ldots,X_n$ be n independent copies $MT_\lambda(\frac{\mu}{n},\frac{\sigma}{n},r)$, then according to Theorem 4.1, there exists a Riesz matrix $Y \sim R(s,\frac{\sigma}{2n})$ such that the conditional distribution

$$X_i|Y \sim N(\mu/n, Y^{-1}) \text{ for all } i \in \{1,2,\ldots,r\}.$$

It follows that $\sum_{i=1}^n X_i|Y$ is $N(\mu, nY^{-1})$. Furthermore, using (4), we deduce that $\frac{Y}{n}$ is $R(s,\frac{\sigma}{2})$. This implies that

$$X = \sum_{i=1}^n X_i \sim MT_\nu(\mu,\sigma,r).$$

This proves the infinitely divisible property of the multiparameter t'distribution (see Feller [12], p. 176).
2. (a) Setting $Z = v'(X-\mu)$, then we have

$$dz = \det(v)dx = \det(\sigma)^{-\frac{1}{2}}dx.$$

Using (2) and (7), we can write

$$\begin{aligned} f_Z(z) &= \frac{\Delta_{s+\frac{1}{2}}\left(\left(\left(v^{-1}\right)^t zz^t v^{-1}+\sigma\right)^{-1}\right)\Gamma_\Omega(s+\frac{1}{2})\det(\sigma)^{\frac{1}{2}}}{\pi^{\frac{r}{2}}\Delta_s(vv^t)\Gamma_\Omega(s)} = \frac{\Delta_{s+\frac{1}{2}}\left(v(zz^t+I)^{-1}v^t\right)\Gamma_\Omega(s+\frac{1}{2})\det(\sigma)^{\frac{1}{2}}}{\pi^{\frac{r}{2}}\Delta_s(vv^t)\Gamma_\Omega(s)} \\ &= \frac{\Delta_{s+\frac{1}{2}}(vv^t)\Delta_{s+\frac{1}{2}}\left((zz^t+I_r)^{-1}\right)\Gamma_\Omega(s+\frac{1}{2})\det(\sigma)^{\frac{1}{2}}}{\pi^{\frac{r}{2}}\Delta_s(vv^t)\Gamma_\Omega(s)} = \frac{\Delta_{s+\frac{1}{2}}(zz^t+I_r)\Gamma_\Omega(s+\frac{1}{2})}{\pi^{\frac{r}{2}}\Gamma_\Omega(s)}. \end{aligned}$$

(b) Setting $W = a^tX + b$, then the probability density function of W is given by

$$\begin{aligned} f_W(w) &= \frac{\Delta_{s+\frac{1}{2}}\left(\left(\left(a^{-1}\right)^t (w-b-a^t\mu)(w-b-a^t\mu)^t a^{-1} + \sigma\right)^{-1}\right)\Gamma_\Omega\left(s+\frac{1}{2}\right)}{\pi^{\frac{r}{2}}\det(a)\Delta_s(\sigma^{-1})\Gamma_\Omega(s)} \\ &= \frac{\Delta_{s+\frac{1}{2}}\left(a\left((w-b-a^t\mu)(w-b-a^t\mu)^t + a^t\sigma a\right)^{-1} a^t\right)\Gamma_\Omega\left(s+\frac{1}{2}\right)}{\pi^{\frac{r}{2}}\det(a)\Delta_s\left(a\left(a^t\sigma a\right)^{-1}a^t\right)\Gamma_\Omega(s)}. \end{aligned}$$

This combined with (2), leads us to draw that

$$\begin{aligned} f_W(w) &= \frac{\Delta_{s+\frac{1}{2}}(aa^t)\,\Delta_{s+\frac{1}{2}}\left(\left((w-b-a^t\mu)(w-b-a^t\mu)^t + a^t\sigma a\right)^{-1}\right)\Gamma_\Omega\left(s+\frac{1}{2}\right)}{\pi^{\frac{r}{2}}\det(a)\,\Delta_s(aa^t)\,\Delta_s\left((a^t\sigma a)^{-1}\right)\Gamma_\Omega(s)} \\ &= \frac{\Delta_{s+\frac{1}{2}}(aa^t)\,\Delta_{s+\frac{1}{2}}\left(\left((w-b-a^t\mu)(w-b-a^t\mu)^t + a^t\sigma a\right)^{-1}\right)\Gamma_\Omega\left(s+\frac{1}{2}\right)}{\pi^{\frac{r}{2}}\Delta_{\frac{1}{2}}(a)\Delta_s(aa^t)\Delta_s\left((a^t\sigma a)^{-1}\right)\Gamma_\Omega(s)}. \end{aligned}$$

After a standard simplification, we deduce that

$$f_W(w) = \frac{\Delta_{s+\frac{1}{2}}(((w-b-a^t\mu)(w-b-a^t\mu)^t + a^t\sigma a)^{-1})\Gamma_\Omega(s+\frac{1}{2})}{\pi^{\frac{r}{2}}\Delta_s((a^t\sigma a)^{-1})\Gamma_\Omega(s)}.$$

Based on this equation and (7) the result holds. □

A relatively simple way to generate a multiparameter t'distribution consists in involving a sample of a Riesz matrix Y with parameters s and $\sigma/2$ using the continuous Riesz simulation algorithm and a generating multivariate normal vector X with parameters μ and Y^{-1}.

Algorithm 4.5. ***(The multiparameter t'distribution simulation)***

Step 1: *Simulate a random matrix Y from the Riesz distribution $R(s, \frac{\sigma}{2})$.*

Step 2: *Simulate a vector X using the scale mixture of the multivariate normal distribution with the variance parameter Y^{-1}.*

Next, we characterize the multiparameter t'distribution by means of its Laplace transform. It is important to mention that the latter is expressed in terms of the generalized Bessel function defined in Faraut and Korányi [11], p. 356.

Theorem 4.6. *The Laplace transform of the multiparameter t'distribution $MT_\nu(0,\sigma,r)$ is equal to*

$$L_{MT_\nu(0,\sigma,r)}(\theta) = \frac{K_s\left(\frac{\sigma}{2}, \frac{\theta\theta^t}{2}\right)}{\Delta_s(2\sigma^{-1})\Gamma_\Omega(s)}, \text{ for all } \theta \in \mathbb{R}^r,$$

where K_s is the generalized Bessel function.

Proof. Using the notation given in Theorem 4.1, we can write

$$L_{MT_\nu(0,\sigma,r)}(\theta) = \int_{\mathbb{R}^r} e^{<\theta,x>} f_X(x)dx = \int_{\mathbb{R}^r} e^{<\theta,x>}\left(\int_\Omega f_{X|Y=y}(x)f_Y(y)dy\right)dx = \int_\Omega L_{X|Y}(\theta)f_Y(y)dy.$$

Since the conditional distribution of X given $Y = y$ is $N(0, y^{-1})$ and Y is $R(s, \frac{\sigma}{2})$, then we obtain

$$L_{MT_\nu(0,\sigma,r)}(\theta) = \int_\Omega e^{-<\frac{\theta\theta^t}{2},y^{-1}>} f_Y(y)dy = \frac{1}{\Delta_s(2\sigma^{-1})\Gamma_\Omega(s)}\int_\Omega e^{-<\frac{\theta\theta^t}{2},y^{-1}>}e^{-<\frac{\sigma}{2},y>}\Delta_{s-\frac{n}{r}}(y)dy.$$

Now, we use the fact that for $(a,b) \in \Omega^2$ and for $s \in \prod_{i=1}^{r}](i-1)/2, +\infty[$,

$$K_s(a,b) = \int_\Omega e^{-<a,y>} e^{-<b,y^{-1}>} \Delta_{s-\frac{n}{r}}(y) dy$$

(see Faraut and Korányi [11], p. 356) to get the desired result. □

Remark 4.7. *Using the conditional expectation, we deduce that if* $X \sim MT_\nu(\mu,\sigma,r)$*, then* $E(X) = \mu$*. This implies that we can estimate the parameter vector* μ *by using the method of moments. In fact, considering* j *copies* $X_1, X_2, \ldots, X_j$ *with multiparameter t'distribution* $MT_\nu(\mu,\sigma,r)$*, then*

$$\hat{\mu} = \frac{X_1 + X_2 + \ldots + x_j}{j}.$$

Furthermore,

$$\begin{aligned} Var(X) &= \mathbb{E}\left(\mathbb{E}[(X-\mathbb{E}(X))(X-\mathbb{E}(X))^t | Y=y]\right) + Var(\mathbb{E}(X|Y=y)) \\ &= \mathbb{E}\left(\mathbb{E}[(X-\mu)(X-\mu)^t | Y=y]\right) = \mathbb{E}(Y^{-1}). \end{aligned}$$

Since Y *has Riesz distribution* $R(s, \frac{\sigma}{2})$*, then* Y^{-1} *is an inverse Riesz matrix with parameter* s *and* $\frac{\sigma}{2}$*. Therefore, using the expression of the first moment of an inverse Riesz matrix given by Louati and Masmoudi [22], we deduce that*

$$Var(X) = \frac{1}{s_r - \frac{r+1}{2}} \frac{\sigma}{2} + \sum_{k=1}^{r-1} \frac{(s_{k+1} - s_k) \prod_{i=k+1}^{r} \frac{s_i - \frac{i}{2}}{s_i - \frac{i+1}{2}}}{(s_k - \frac{k+1}{2})(s_{k+1} - \frac{k+2}{2})} \left(P_k^*(2\sigma^{-1})\right)^{-1}. \quad (9)$$

A standard calculation demonstrates that, in particular, if for all $i \in \{1,2,\ldots,r\}$*,* $s_i = p$*, then (9) becomes*

$$Var(X) = \frac{\sigma}{2p - (r+1)},$$

which corresponds to the variance-covariance of the multivariate t'distribution $T_p(\mu,\sigma,r)$ *(see Cornish [5]).*

In the following theorem, we confirm that the multiparameter t'distribution can be also obtained by combining the notion of mixture and the Cholesky decomposition.

Theorem 4.8. *Let* V *be a standard normal vector and* $Y = UU^t$ *a Riesz matrix* $R(s, \frac{\sigma}{2})$ *independent of* V*, then the random vector* $X = \left(U^{-1}\right)^t V + \mu$ *has a multiparameter t'distribution with parameters* $\mu \in \mathbb{R}^r$ *and* $\sigma \in \Omega$ *and a degree of freedom* $\nu = (s_1, s_2 - s_1, \ldots, s_r - s_{r-1})$*.*

Proof. According to (3), the joint density of (Y, V) is provided by

$$f_{(Y,V)}(y,v)dydv = f_Y(y) f_V(v) dydv = \frac{e^{-\frac{1}{2}\left(\langle v,v\rangle + \langle \sigma,y\rangle\right)} \Delta_{s-\frac{n}{r}}(y)}{(2\pi)^{\frac{r}{2}} \Delta_s(2\sigma^{-1}) \Gamma_\Omega(s)} \mathbf{1}_{\Omega\times\mathbb{R}^r}(y,v) dydv.$$

Since $V = U^t(X - \mu)$, then the Jacobian is given by

$$dydv = det(u^t)dx = det(y)^{\frac{1}{2}} dydx.$$

Therefore, the joint density of (Y, X) is

$$\begin{aligned} f_{(Y,X)}(y,x)dydx &= \frac{e^{-\frac{1}{2}\left(\langle\sigma,y\rangle+\langle u^t(x-\mu),u^t(x-\mu)\rangle\right)}\Delta_{s-\frac{n}{r}}(y)det(y)^{\frac{1}{2}}}{(2\pi)^{\frac{r}{2}}\Delta_s(2\sigma^{-1})\Gamma_\Omega(s)}\mathbf{1}_{\Omega\times\mathbb{R}^r}(y,x)dydx \\ &= \frac{e^{-\frac{1}{2}\left(\langle\sigma,y\rangle+\langle(x-\mu),uu^t(x-\mu)\rangle\right)}\Delta_{s-\frac{n}{r}+\frac{1}{2}}(y)}{(2\pi)^{\frac{r}{2}}\Delta_s(2\sigma^{-1})\Gamma_\Omega(s)}\mathbf{1}_{\Omega\times\mathbb{R}^r}(y,x)dydx \\ &= \frac{e^{-\frac{1}{2}\left(\langle\sigma,y\rangle+\langle(x-\mu)(x-\mu)^t,y\rangle\right)}\Delta_{s-\frac{n}{r}+\frac{1}{2}}(y)}{(2\pi)^{\frac{r}{2}}\Delta_s(2\sigma^{-1})\Gamma_\Omega(s)}\mathbf{1}_{\Omega\times\mathbb{R}^r}(y,x)dydx \\ &= \frac{e^{-\frac{1}{2}\left(\langle\sigma+(x-\mu)(x-\mu)^t,y\rangle\right)}\Delta_{s-\frac{n}{r}+\frac{1}{2}}(y)}{(2\pi)^{\frac{r}{2}}\Delta_s(2\sigma^{-1})\Gamma_\Omega(s)}\mathbf{1}_{\Omega\times\mathbb{R}^r}(y,x)dydx. \end{aligned}$$

Integrating this with respect to y and using (3), we infer that the marginal density function of X is

$$\frac{2^{\frac{r}{2}+\sum_{i=1}^r s_i}\Gamma_\Omega\left(s+\frac{1}{2}\right)\Delta_{s+\frac{1}{2}}\left(((x-\mu)(x-\mu)^t+\sigma)^{-1}\right)}{(2\pi)^{\frac{r}{2}}2^{\sum_{i=1}^r s_i}\Delta_s(\sigma^{-1})\Gamma_\Omega(s)}.$$

Grounded upon this equation and (7), we obtained the desired result. □

Remark 4.9. *It is worth mentioning that, it is possible to apply this representation to obtain the expectation and the variance of the random vector with the multiparameter t'distribution. In fact, since V (a centered normal vector) and $Y = UU^t$ are independent, then we can assert*

$$\mathbb{E}(X) = \mathbb{E}(U^{-1}V + \mu) = \mu.$$

The variance of X is given by

$$\begin{aligned} Var(X) &= \mathbb{E}\left((X-\mathbb{E}(X))(X-\mathbb{E}(X))^t\right) = \mathbb{E}\left(\left(\left(U^{-1}\right)^t V\right)\left(\left(U^{-1}\right)^t V\right)^t\right) = \mathbb{E}\left(\left(U^{-1}\right)^t VV^tU^{-1}\right) \\ &= \int_{\mathcal{T}_l^+}\int_{\mathbb{R}^r}\left(u^{-1}\right)^t vv^tu^{-1}f_U(u)f_V(v)dudv = \int_{\mathcal{T}_l^+}\left(u^{-1}\right)^t\left(\int_{\mathbb{R}^r}vv^tf_V(v)dv\right)u^{-1}f_U(u)du. \end{aligned}$$

Moreover, since V is a centered normal vector with $Var(V) = I_r$, then

$$Var(X) = \int_{\mathcal{T}_l^+}\left(u^{-1}\right)^t I_ru^{-1}f_U(u)du = \mathbb{E}\left(\left(U^{-1}\right)^t U^{-1}\right) = \mathbb{E}\left(\left(UU^t\right)^{-1}\right) = \mathbb{E}(Y^{-1}).$$

The fact that Y is Riesz distributed $R(s, \frac{\sigma}{2})$ gives

$$Var(X) = \frac{1}{s_r - \frac{r+1}{2}}\frac{\sigma}{2} + \sum_{k=1}^{r-1}\frac{(s_{k+1}-s_k)\prod_{i=k+1}^r\frac{s_i-\frac{i}{2}}{s_i-\frac{i+1}{2}}}{(s_k-\frac{k+1}{2})(s_{k+1}-\frac{k+2}{2})}\left(P_k^*(2\sigma^{-1})\right)^{-1},$$

(see Louati and Masmoudi (2015)).

Next, we invest the representation given in Theorem 4.8 to generate the multiparameter t'distribution with parameters $\mu \in \mathbb{R}^r$ and $\sigma \in \Omega$ and a degree of freedom $\nu = (s_1, s_2 - s_1, \dots, s_r - s_{r-1})$.

Algorithm 4.10. ***(The multiparameter t'distribution simulation)***

Step 1: *Simulate a random matrix Y from the Riesz distribution $R(s, \frac{\sigma}{2})$.*

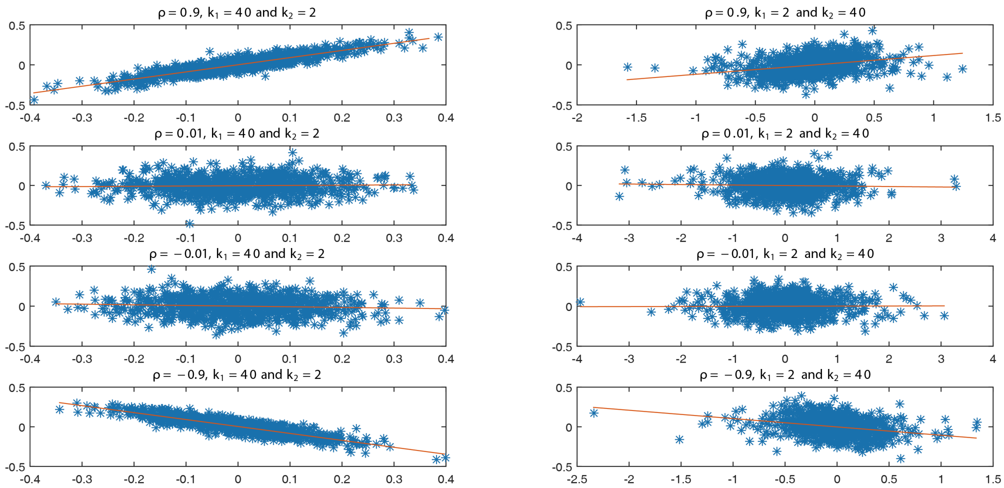
Step 2: *Compute the Cholesky decomposition U of Y and Calculate its inverse U^{-1}.*

Step 3: *Simulate a standard normal vector V.*

Step 4: *Calculate $X = \left(U^{-1}\right)^t V + \mu$.*

Remark 4.11. *Subsequently, to clarify the shape of the multiparameter t′distribution, we attempt to repeat the above algorithm for a chosen sample size $N = 1000$ with parameters $\mu = (0,0)^t$ and $\sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, where $\rho \in \{-0.9, -0.01, 0.01, 0.9\}$ for different values of degree of freedom $\nu = (k_1, k_2)$.*

Figure 1: Simulation studies with different degrees of freedom



We represent the obtained simulated data in Figure 1 as a scatter plot which seems to have an ellipsoid shape. Its longitudinal axis is slightly inclined when the correlation coefficient is around -1 or 1. This implies that a linear equation describes the relationship between X_1 and X_2 perfectly, with all data points lying on a line. Besides, if the correlation coefficient is around 0, then there is no linear correlation between the two variables.

5. Estimation of the parameters

5.1. The Expectation Maximization Algorithm

The maximum likelihood estimation of the multivariate t′distribution, especially with known degree of freedom, has been based on the Expectation Maximization (EM) algorithm. Note that Dempster *et al.* [9] showed that the EM algorithm can be used to find maximum likelihood estimates when the observation is viewed as incomplete data and with a fixed degree of freedom. It consists of an Expectation step (E-step) followed by a Maximization step (M-step). We apply the EM algorithm for the multiparameter t′distribution in order to estimate the set of parameters $\Theta = \{\mu, \sigma\}$.

Let $X_1, X_2, \ldots, X_N$ be N independent random vectors such that

$$X_i|Y_i \sim N(\mu, Y_i^{-1}) \text{ and } Y_i \sim R(s, \sigma/2), \text{ for all } i = 1, 2, \ldots, N.$$

The log-likelihood function of $(X_1, \ldots, X_N, Y_1, \ldots, Y_N)$ is equal to

$$l(X_1, \ldots, X_N, Y_1, \ldots, Y_N, \Theta) = \sum_{i=1}^{N} \ln\left(f_{X_i|Y_i=y_i}(x_i|\Theta) f_{Y_i}(y_i|\Theta)\right).$$

The first step of the EM algorithm consists of finding the expected value of the log-likelihood

$$\mathbb{Q}(\Theta,\Theta^{(l)}) \quad = \quad \mathbb{E}(l(X_1,\dots,X_N,Y_1,\dots,Y_N,\Theta)|X_1,\dots,X_N,\Theta^{(l)}). \tag{10}$$

where $\Theta^{(l)} = \{\mu^{(l)},\sigma^{(l)}\}$ is the set of the current parameters estimates. The second step of the EM algorithm is to maximize $\mathbb{Q}(\Theta,\Theta^{(l)})$.
Next, we introduce the recursive expressions of the estimators of $\mu^{(l+1)}$ and $\sigma^{(l+1)}$ at the iteration $(l+1)$.

Theorem 5.1. *The estimators of μ and σ at the iteration $(l+1)$ are*

1. $\sigma^{(l+1)} = \sum_{k=1}^{r}(s_{r-k+1}-s_{r-k})P_k^*\left(\frac{1}{2N}\sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\dots,X_N,\Theta^{(l)}\right)\right)^{-1}$
2. $\mu^{(l+1)} = \left(\sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\dots,X_N,\Theta^{(l)}\right)\right)^{-1}\sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\dots,X_N,\Theta^{(l)}\right)x_i.$

Proof. According to (3) and (10), we assert

$$\mathbb{Q}(\Theta,\Theta^{(l)}) = \mathbb{E}\left((f(\mu)+f(\sigma)+H)\,\middle|X_1,\dots,X_N,\Theta^{(l)}\right),$$

where

$$f(\mu) = \mu^t\sum_{i=1}^{N}\left(y_ix_i - \frac{y_i\mu}{2}\right),$$

$$f(\sigma) = -N\log(\Delta_s(2\sigma^{-1})) - \sum_{i=1}^{N}\langle\frac{\sigma}{2},y_i\rangle$$

and

$$H = -\frac{N}{2}\left(r\log(2\pi)+2\ln\left(\Gamma_\Omega(s)\right)\right)+\frac{1}{2}\left(\sum_{i=1}^{N}\log(\det(y_i)) - \sum_{i=1}^{N}x_i^ty_ix_i + \sum_{i=1}^{N}2\ln\left(\Delta_{s-\frac{n}{r}}(y_i)\right)\right).$$

It follows that

$$\mathbb{Q}(\Theta,\Theta^{(l)}) = \mu^t\left(\sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\dots,X_N,\Theta^{(l)}\right)x_i - \frac{1}{2}\mathbb{E}\left(y_i|X_1,\dots,X_N,\Theta^{(l)}\right)\mu\right)$$

$$-N\log(\Delta_s(2\sigma^{-1})) - \sum_{i=1}^{N}\left\langle\frac{\sigma}{2},\mathbb{E}\left(y_i|X_1,\dots,X_N,\Theta^{(l)}\right)\right\rangle + \mathbb{E}\left(H|X_1,\dots,X_N,\Theta^{(l)}\right).$$

1. Differentiating $\mathbb{Q}(\Theta,\Theta^{(l)})$ with respect to σ and using the fact that for all $x\in\Omega$, we have

$$\frac{\partial\log(\Delta_k^*(x))}{\partial x} = (P_k^*(x))^{-1},\ \text{for all } k\in\{1,2,\dots,r\}.$$

We obtain

$$0 = \frac{\partial\left(\mathbb{Q}(\Theta,\Theta^{(l)})\right)}{\partial\sigma} = -N\sum_{k=1}^{r}(s_{r-k}-s_{r-k+1})(P_k^*(\sigma))^{-1} - \frac{1}{2}\sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\dots,X_N,\Theta^l\right),$$

with $s_0=0$. Therefore,

$$\sum_{k=1}^{r}(s_{r-k+1}-s_{r-k})(P_k^*(\sigma))^{-1} = \frac{1}{2N}\sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\dots,X_N,\Theta^{(l)}\right). \tag{11}$$

Using the fact that for all $\theta\in-\Omega$,

$$k'_{R_s}(\theta) = \sum_{i=1}^{r}(s_{r-i+1}-s_{r-i})\left(P_i^*(-\theta)\right)^{-1},$$

(see Hassairi and Lajmi [15]), (11) becomes

$$k'_{R_s}(-\sigma) = \frac{1}{2N}\sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\ldots,X_N,\Theta^{(l)}\right).$$

Hence,

$$\sigma^{(l+1)} = -k'^{-1}_{R_s}\left(\frac{1}{2N}\sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\ldots,X_N,\Theta^{(l)}\right)\right).$$

Based on this equation and the fact that for all $m \in \Omega$,

$$k'^{-1}_{R_s}(m) = \sum_{k=1}^{r}(s_{r-k} - s_{r-k+1})\left(P^*_k(m)\right)^{-1},$$

(see Kammoun *et al.* [18]), we give the expression $\sigma^{(l+1)}$.

2. Differentiating $\mathbb{Q}(\Theta,\Theta^{(l)})$ with respect to μ, we obtain

$$0 = \frac{\partial\left(\mathbb{Q}(\Theta,\Theta^{(l)})\right)}{\partial\mu} = \sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\ldots,X_N,\Theta^{(l)}\right)x_i - \sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\ldots,X_N,\Theta^{(l)}\right)\mu.$$

This implies that

$$\sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\ldots,X_N,\Theta^{(l)}\right)x_i = \sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\ldots,X_N,\Theta^{(l)}\right)\mu.$$

Multiplying by $\left(\sum_{i=1}^{N}\mathbb{E}\left(y_i|X_1,\ldots,X_N,\Theta^{(l)}\right)\right)^{-1}$, we get the expression of the estimator $\mu^{(l+1)}$. □

5.2. Numerical estimation

This subsection is devoted to the evaluation of the performance of the proposed EM estimation approach. For this purpose, we suppose that $r = 5$ and we consider $s = (5, 7.5, 8.5, 12.5, 16)$, $\mu = (-1, 2, 0, 1, -1)^t$ and

$$\sigma = \begin{pmatrix} 27.5 & 2 & 0.5 & 0 & 2.75 \\ 2 & 27.5 & 1.25 & 2.5 & 1.25 \\ 0.5 & 1.25 & 52.5 & 0 & 3.75 \\ 0 & 2.5 & 0 & 25 & 0.25 \\ 2.75 & 1.25 & 3.75 & 0.25 & 36.25 \end{pmatrix}.$$

We repeat the above EM algorithm for different samples sizes $N \in \{500, 1000, 2000, 5000\}$. We assume that the number of simulated samples n was chosen to be 1000 and we take as initial value of the mean vector $\mu^{(0)} = \frac{X_1+X_2+\ldots+X_j}{j}$, where $X_1, X_2, \ldots, X_j$ are j copies with multiparameter t'distribution $MT_\nu(\mu,\sigma,r)$. The estimated values $\hat{\mu}$ and $\hat{\sigma}$ are given by

$$\hat{\mu} = (-0.9539, 2.0526, 0.0003, 1.0036, -0.9815)^t \text{ and } \hat{\sigma} = \begin{pmatrix} 26.2081 & 0.1331 & 0.1177 & -0.0111 & 0.0339 \\ 0.1331 & 26.8061 & 0.2186 & 0.1554 & 0.0089 \\ 0.1177 & 0.2186 & 51.1257 & -0.0284 & 0.1347 \\ -0.0111 & 0.1554 & -0.0284 & 25.4496 & 0.0317 \\ 0.0339 & 0.0089 & 0.1347 & 0.0317 & 35.6550 \end{pmatrix}.$$

The performance of these estimators is assessed by means of the Mean Squared Error between the true and the estimated parameters defined by $MSE(\hat{\mu}) = \mathbb{E}\left((\mu-\hat{\mu})^2\right)$ and $MSE(\hat{\sigma}) = \mathbb{E}\left((\sigma-\hat{\sigma})^2\right)$. The obtained numerical results are depicted in the following table.

N	500	1000	2000	5000
$MSE(\hat{\mu})$	0.0044	0.0016	0.0013	0.0004
$MSE(\hat{\sigma})$	0.0132	0.0115	0.0102	0.0099

Table 1: MSE of the estimated parameters $\hat{\mu}$ and $\hat{\sigma}$

Table 1 portrays that the MSE of the estimated parameters decreases when N increases. It is very close to zero. In fact, the $MSE(\hat{\mu})$ is around 4×10^{-4} and the $MSE(\hat{\sigma})$ is about 9.9×10^{-3} for $N = 5000$. This allows us to confirm the performance of the proposed EM estimation approach. Once MSE is calculated, the confidence interval of the MSE is assessed and confidence level is chosen to be 95%. The obtained results are presented in the following figure.

Figure 2: MSE and confidence interval MSE error of the estimated parameters $\hat{\mu}$ and $\hat{\sigma}$

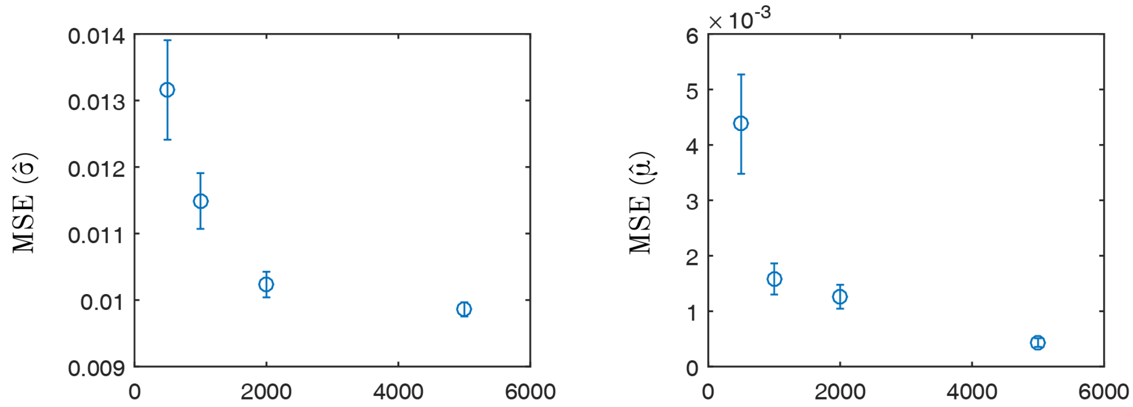


Figure 2 highlights that the confidence intervals of the MSE cover the true parameters values. Besides, the size of data influences the quality of the estimators. In fact, As MSE decreases, the width of the interval decreases. As we increase the sample size n, the width of the interval decreases. This proves that the estimators $\hat{\mu}$ and $\hat{\sigma}$ are close to the true parameters.

6. Conclusion

To this extent, there is no doubt about the significant role played by the multivariate t'distribution in the context of multivariate statistics. Our central focus is upon the natural extension of the the multivariate t'distribution, namely the multiparameter t'distribution. From this perspective, we first invested the natural extension of the Wishart distribution, that is the Riesz distribution, as well as the Cholesky decomposition to represent this distribution. It is noteworthy that unlike the multivariate t'distribution, the multiparameter t'distribution acquired a shape parameter comprising r different element. We offered an algorithm to generate the new distribution involving a sampling of a Riesz matrix. We introduced the EM algorithm to estimate its parameters. Some simulations were carried out in different sizes so as to illustrate the performance of their estimators. At this stage of analysis, we would assert that our work is a step that may be taken further as it lays the ground and paves the way for future works to enact promising applications in such areas as multiple regression and Bayesian approach.

References

[1] S.A. Andersson, T. Klein, On Riesz and Wishart distributions associated with decomposable undirected graphs. J. Multivariate Anal. 101(4) (2010) 789–810.
[2] J.L. Andrews, P.D. McNicholas, Model-based clustering, classification, and discriminant analysis via mixtures of multivariate t-distributions. Stat. Comput. 22(5) (2012) 1021–1029.
[3] M. Arashi, S.M.M. Tabatabaey, Stein-type improvement under stochastic constraints: Use of multivariate Student-t model in regression. Stat. Probabil. Lett. 78(14) (2008) 2142–2153.
[4] J.T. Chien, A Bayesian prediction approach to robust speech recognition and online environmental testing. Speech Communication 37 (2002) 321-334.
[5] E.A. Cornish, The multivariate t-distribution associated with a set of normal standard deviates. Aust. J. Phys., 7 (1954) 531–542.
[6] E.A. Cornish, The sampling distributions of statistics derived from the multivariate t-distribution. Aust. J. Phys., 8 (1954) 193–199.
[7] J.A. Díaz-Garcia, A note on the moments of the Riesz distribution. J. Stat. Plan. Inference 143(11) (2013) 1880–1886.
[8] J.M. Dickey, Matricvariate generalizations of the multivariate t distribution and the inverted multivariate *t* distribution. Ann. Math. Stat., 38 (1967) 511–518.
[9] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, J. R. Stat. Soc. Series B, 39(1) (1977), 1–38.
[10] C.W. Dunnett, M. Sobel, A bivariate generalization of Student's t-distribution with tables for certain special cases. Biometrika. 41 (1954) 153–169.
[11] J. Faraut, A. Korányi, Analysis on symmetric cones. Oxford University Press, Oxford, 1994.
[12] W. Feller, An introduction to probability theory and its applications. Vol II, Second Edition. John Wiley & sons, New York, 1970.
[13] S.G. Gindikin, Analysis in homogeneous domains. Russian Math. Surveys, 19(4) (1964) 1–89.
[14] P. Graczyk, G. Letac, H. Massam, The moments of the complex Wishart ditribution and the symmetric group. Ann. Stat. 41 (2003) 287–309.
[15] A. Hassairi, S. Lajmi, Riesz exponential families on symmetric cones. J. Theoret. Probab. 14(4) (2001) 927–948.
[16] A. Hassairi, M. Louati, Multivariate stable exponential families and Tweedie scale. J. Stat. Plann. Inf. 139(2) (2009) 143–158.
[17] A. Hassairi, M. Louati, Mixture of the Riesz distribution with respect to a multivariate Poisson. Commun. Stat. Theory & Methods. 42(6) (2013) 1124–1140.
[18] K. Kammoun, M. Louati, A. Masmoudi, Maximum likelihood estimator of the scale parameter for the Riesz distribution. Stat. Probabil. Lett. 126 (2017) 127–131.
[19] S. Kotz, S. Nadarajah, Multivariate T-distributions and their applications. Cambridge University Press, Cambridge, 2004.
[20] P. Lin, Some characterizations of the multivariate *t* distribution. J. Multivariate Anal. 2 (1972) 339–344.
[21] M. Louati, Mixture of the Riesz distribution with respect to the generalized multivariate gamma distribution, J. Korean Statist. Soc. 42 (2013) 83–93.
[22] M. Louati, A. Masmoudi, Moment for the inverse Riesz distributions, Stat. Probabil. lett. 102 (2015) 30–37.
[23] G.P. Nason, Robust projection indices. J. R. Stat. Soc. Series B, 63 (2001) 551–567.
[24] E. Veleva, Stochastic representations of the Bellman gamma distribution. In Proceeding of international conference on theory and applications in mathematics and informatics, ICTAMI 2009. Alba Iulia, Romania. (2009) 463–474.
[25] D. Von Rosen, Moments for the inverted Wishart distribution. Scand. J. Stat. 15 (1988) 97–109.
[26] J. Wishart, The generalized product moment distribution in samples from a normal multivariate population. Biometrics 20 (1928) 32–52.