Vadym Omelchenko
Parameter estimation of sub-Gaussian stable distributions

Kybernetika, Vol. 50 (2014), No. 6, 929–949

Persistent URL: http://dml.cz/dmlcz/144117

KYBERNETIKA — VOLUME *50* (2014), NUMBER 6, PAGES 929–949

PARAMETER ESTIMATION OF SUB-GAUSSIAN STABLE DISTRIBUTIONS

Vadym Omelchenko

In this paper, we present a parameter estimation method for sub-Gaussian stable distributions. Our algorithm has two phases: in the first phase, we calculate the average values of harmonic functions of observations and in the second phase, we conduct the main procedure of asymptotic maximum likelihood where those average values are used as inputs. This implies that the main procedure of our method does not depend on the sample size of observations. The main idea of our method lies in representing the partial derivative of the density function with respect to the parameter that we estimate as the sum of harmonic functions and using this representation for finding this parameter. For fifteen summands we get acceptable precision. We demonstrate this methodology on estimating the tail index and the dispersion matrix of sub-Gaussian distributions.

Keywords: stable distribution, sub-Gaussian distribution, maximum likelihood, characteristic function

Classification: 93E12, 62A10

1. INTRODUCTION

Classical models in financial risk management and portfolio optimization are based on normality, but normal models are known to have a number of shortcomings and there is overwhelming empirical evidence that the normality assumption must be rejected [3, 5, 13, 19, 24]. On the other hand, the use of normality is theoretically justifiable because random effects that influence data are caused by a range of micro-effects, which add up and hence allow us to use the central limit theorem [9, 25]. The assumption of normality enables us to make many problems tractable, while replacing normality by another distribution capable of capturing more features of the time series of prices may lead to useless or even dangerous models unless the modeler can properly handle the arising complexity entailed by introducing a non-normal innovation [17]. Hence, the question is: what is the best compromise? How can we transform the model without the need to rule out normality and without dropping common-sense assumptions which entail normality? If we are to replace the normal assumption with a stable one, we will also be able to use the central limit theorem and convolution properties of the normal distribution will be also preserved [25]. However we will need to be able to

DOI: 10.14736/kyb-2014-6-0929

estimate the parameters of the stable distributions properly, which is a challenging task because stable distributions in general do not have an explicit form of either the density or the distribution function except for a few cases [9, 25]. In this paper, we take up this challenge. In the case of density functions of exponential types such as normal or exponential densities, we can easily obtain explicit estimators of parameters such as mean or variance. In such cases, we can easily obtain estimates of parameters for large numbers of observations. The main goal of this work is to develop a method based on maximum likelihood capable of handling large amounts of data and producing decent precision as in the case of exponential type densities. This method consists of two phases: in the first phase, we calculate the means of the harmonic functions of observations and in the second phase, we conduct the main procedure using those average values as inputs. The first phase enables us to deal with large amounts of data and the second phase enables us to attain high precision and is independent of the sample size of the data. This two-phase algorithm is called a method of projections which converges to the maximum likelihood methodology. We apply this to univariate stable and multivariate sub-Gaussian distributions. We will denote our methodology with *MLP* to emphasize that our algorithm is based on maximum likehood estimation and projections.

1.1. Article structure

This article is organized as follows. In the Introduction we gave a brief description to stable distributions and challenges that arise in estimating their parameters. In the next sections, we give a complete definition of stable distributions, provide their basic properties and discuss contributions of other authors to estimating their parameters. Then, we provide the description of our method of the estimation of the tail index and dispersion matrix of sub-Gaussian distributions and then move to conclusions.

2. SUB-GAUSSIAN DISTRIBUTIONS AND THEIR PROPERTIES

Sub-Gaussian distributions are a special case of stable distributions. They represent symmetric and heavy tailed distributions whose dependence structure is given by a matrix. To describe sub-Gaussian distributions, we need first to define general stable distributions.

2.1. Definition of stable distributions

There are four equivalent definitions of univariate stable distributions that concern their different statistical properties: two concern convolution properties, one concerns limit properties and the last one concerns the form of the characteristic function [25]. The definition that concerns the form of the characteristic function is the most important for us because only this function has an explicit form unlike the density and distribution functions.

Definition 2.1. The random variable X has a univariate sub-Gaussian distribution if its characteristic function is of the form:

$$\psi(u) = \exp(i\mu u)\exp\left(-\sigma^{\alpha}|u|^{\alpha}\left(1 - i\beta\cdot\operatorname{sign}(u)\tan\frac{\pi\alpha}{2}\right)\right), \qquad \alpha \neq 1, \tag{1}$$

$$\psi(u) = \exp(i\mu u)\exp\left(-\sigma|u|\left(1 + i\beta\cdot\operatorname{sign}(u)\ln(u)\frac{2}{\pi}\right)\right), \qquad \alpha = 1,$$

where $\alpha \in (0, 2]$, $\mu \in (-\infty, \infty)$, $\beta \in [-1, 1]$ and $\sigma > 0$.

Note that if $\alpha = 2$ we have a characteristic function of the normal distribution. If $\alpha < 2$ then any moment EX^a with $a \geq \alpha$ is infinite. If $a < \alpha$ then EX^a is finite [9, 25]. Hence if $\alpha < 2$, then the variance of X is infinite. A general univariate stable distribution is denoted as $S_{\alpha}(\sigma, \beta, \mu)$, where α, σ, β and μ are the tail index, the scale parameter, the skewness parameter and the location parameter respectively.

Definition 2.2. The random vector $\mathbf{X}$ has a sub-Gaussian Distribution with the location parameter μ and the matrix $\mathbf{Q}$ if its characteristic function is of the form

$$\psi(\mathbf{u}) = \exp(i\cdot\mathbf{u}^T\mu)\exp\left(-\left|\mathbf{u}^T\mathbf{Q}\mathbf{u}\right|^{\alpha/2}\right). \tag{2}$$

Such a vector is called a sub-Gaussian vector.

Here μ is the location parameter and $\mathbf{Q}$ is the matrix that determines the dependence structure between the marginals. (Another name for these distributions is *multivariate elliptical stable distributions*) $\mathbf{Q}$ is a positively definite matrix and in the case of α equal to 2 we get the multivariate normal distribution whose covariance matrix is $\mathbf{Q}$ [22]. If $\alpha > 1$ then $\mu = E(X)$. A very important property of stable distributions is the fact that any linear combination of stable random variables with the same α has a stable distribution with the same α parameter as well. In other words, if X_1, X_2,...,X_n, X are i.i.d. sub-Gaussian distributions $X_i \sim \psi(\mathbf{u}) = \exp(-\left|\mathbf{u}^T Q\mathbf{u}\right|^{\alpha/2})$, $i = 1, 2, \ldots$, then

$$X_1 + X_2 + \cdots + X_n =_d \frac{1}{n^{1/\alpha}}X,$$

[19, 22].

2.2. Simulation of sub-Gaussian distributions

If $\mathbf{Z}$ is a random vector with a characteristic function $\psi(\mathbf{u}) = \exp(-\{\mathbf{u}^T\mathbf{Q}\mathbf{u}\}^{\alpha/2})$ then

$$\mathbf{Z} = \sqrt{s}\cdot\mathbf{G}, \tag{3}$$

[22] where $s \sim S_{\alpha/2}\left(\left(\cos\left(\frac{\pi\alpha}{4}\right)\right)^{2/\alpha}, 1, 0\right)$ and $\mathbf{G} \sim N(\mathbf{0}, \mathbf{Q})$. $\mathbf{G} = \mathbf{C}^T\mathbf{Y}$, $\mathbf{C}\cdot\mathbf{C}^T = \mathbf{Q}$, $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)^T$, $Y_i \sim N(0, 1)$, $i = 1, 2, \ldots, n$ and Y_i, $i = 1, 2, \ldots, n$ are i.i.d. and finally, s and $\mathbf{G}$ are independent. If we want to simulate a sample from $S_{\alpha}(1, \beta, 0)$, we can do it as follows [2]:

$$X = S_{\alpha,\beta}\cdot\frac{\sin(\alpha(V + B_{\alpha,\beta}))}{\cos(V)^{\alpha}}\left[\frac{\cos(V - \alpha(V + B_{\alpha,\beta}))}{W}\right]^{\frac{1-\alpha}{\alpha}}, \tag{4}$$

where

$$B_{\alpha,\beta} = \frac{\arctan\left(\beta \tan\frac{\pi\alpha}{2}\right)}{\alpha},$$

$$S_{\alpha,\beta} = \left[1 + \beta^2 \tan^2\left(\frac{\pi\alpha}{2}\right)\right]^{1/2\alpha},$$

$X \sim S_\alpha(1,\beta,0)$, $W \sim \exp(1)$, $V \sim U(-\pi/2,\pi/2)$, V and W are independent. If $X \sim S_\alpha(1,\beta,0)$ then for all $\sigma > 0$ and $\mu \in \mathbf{R}$ we have $Y = \sigma X + \mu \sim S_\alpha(\sigma,\beta,\mu)$. (3) and (4) enable us to simulate any sub-Gaussian distribution. The simulation of a general multivariate stable distribution is a more complicated task and requires numerical techniques to be conducted.

3. GENERAL MULTIVARIATE STABLE DISTRIBUTIONS

As it was noted, sub-Gaussian distributions represent a special case of multivariate stable distributions and the latter are defined as follows.

Theorem 3.1. Let $0 < \alpha \leq 2$. Then $X = (X_1, X_2, \ldots, X_n)$ is a stable random vector with the tail index α iff there exists a finite measure Γ on the unit hypersphere $S_n = \{\mathbf{s} \in R^n | \|s\| = 1\}$ and a vector $\mu \in R^n$ such that for $\alpha > 0$

$$\mathbf{\Psi}_\alpha(\mathbf{u}) = \exp\left\{-\int_{S^d}\left(1 - i\cdot\operatorname{sign}((\mathbf{u},\mathbf{s}))\tan\left(\frac{\pi\alpha}{2}\right)\right)\Gamma(\mathrm{d}\mathbf{s}) + i\cdot(\mathbf{u},\mu)\right\}. \qquad (5)$$

The pair (μ,Γ) is unique. [9, 25].

The measure Γ, called the *spectral measure* of the stable random vector X, specifies the dependence structure between its marginal distributions.

Remark. Note that α is the same for all such marginal distributions.

4. APPROACHES TO PARAMETERS ESTIMATION OF STABLE DISTRIBUTIONS

In the case of univariate stable distributions, we need to estimate four parameters: α, σ, β, and μ. The most challenging task is estimating the parameter α [9]. In the case of a general stable distribution, we estimate the parameter α, vector μ and the *spectral measure* Γ, which bears all of the information about the values β and σ of their univariate marginals. It can be easily shown that the matrix $\mathbf{Q}$ is a special case of Γ [3, 9, 25]. Univariate marginals of sub-Gaussian distributions are symmetric, i. e. $\beta = 0$ for all of the marginals [25].

4.1. Parameters of univariate stable distributions

For estimating α, there are many methods. There are estimators in a form of an explicit function of observations like the Hill estimator [6, 16]. The expression for the Hill estimator is:

$$\hat{\alpha} = \frac{1}{\frac{1}{k}\sum_{j=1}^{k}\log X_{n+1-j:n} - \log X_{n-k:n}},$$

where $X_{j:n}$ is the sample j-th order statistics and k is the *window* parameter. This estimator is easy to implement, but there are a number of shortcomings associated with it because we take into account only those observations that are in the tail of the distribution whose number is a small fraction of all observations. The definition of what is in the tail and what is not there is problematic if we do not know α. For this purpose, there exists the *window* parameter of the Hill estimator whose calibration requires us to use sophisticated numerical methods and heavy dependency on the *window* size is the main problem [6, 16]. There are methods for estimating α based on quantiles. The main idea of McCulloch's methodology [14] [14] is to use differences in quantiles, properly normalized, in order to get rid of our dependence on location and scale parameters. Then, two functions on the stability index and the skewness are numerically calculated from the sample quantiles values and inverted to get the corresponding parameter estimates. This method is consistent and can be used for the estimation of all of the parameters of a univariate stable distribution, however it does not work well for certain choices of the parameters' values [14]. There are methods of estimating α and the rest of the parameters by using the empirical characteristic function. The main idea of this methodology consists in minimizing the distance between the characteristic function (CF) and the *empirical characteristic function* (ECF) in an appropriate norm [1, 10, 28]. Let us denote by ψ the characteristic function of a stable distribution and by

$$\hat{\psi}(u) = \frac{1}{n}\sum_{j=1}^{n} \exp\left(iuX_j\right)$$

its *empirical characteristic function*. Since $|\hat{\psi}(u)|$ is bounded, all moments of $\hat{\psi}(u)$ are finite for any fixed u. By the Law of Large Numbers $\hat{\psi}(u)$ is a consistent estimator of $\psi(u)$.

The method finds

$$\theta = \operatorname{argmin}_{\theta\in\Theta}\|\hat{\psi} - \psi\|,$$

where θ is a point in the parametric space Θ and $\|\cdot\|$ is a norm usually L^∞ or an L^r weighted norm with $r > 0$. The last type of the norm is more useful for implementation and it can be written as

$$h(\theta) = \int_{-\infty}^{\infty} |\hat{\psi}(u) - \psi(u,\theta)|^r W(u)\,\mathrm{d}u.$$

Here $W(\cdot)$ is a weight function and the optimal estimate is obtained as follows:

$$\theta = \operatorname{argmin}_\theta h(\theta).$$

More generally the objective function is of the form:

$$h(\theta) = \int_{-\infty}^{\infty} |\hat{\psi}(u) - \psi(u,\theta)|^r \,\mathrm{d}G(u),$$

where $G(\cdot)$ is a distribution function. When $G(\cdot)$ is a step function and $r = 2$, the objective function becomes

$$h(\theta) = \sum_{i=1}^{n} |\hat{\psi}(u_i) - \psi(u_i,\theta)|^2 g(u_i)$$

with $g(u) = \frac{\mathrm{d}G(u)}{\mathrm{d}u}$. The optimal selection of discreet points $u_1, u_2, \ldots, u_p$ is discussed in Carrasco, Madan et. al, and Schmidt [1, 12, 26]. We will denote such estimates with "*CFB*" where this abbreviation states for *characteristic function based* estimators.

DuMouchel [4], Zolotarev [29], and Nolan [19] developed methods of asymptotic maximum likelihood for estimating the parameters of stable distributions [4, 19, 29]. The main limitation is that density functions of stable distributions do not have an explicit form. Approximative methods are based on approximating functionals of density functions. For a general problem of estimating the parameters $\alpha, \sigma, \beta, \mu$ we maximize the following expression:

$$l(\alpha, \sigma, \beta, \mu) = \sum_{i=1}^{n} \log p\left(X_i; \alpha, \sigma, \beta, \mu\right),$$

where $p(\cdot)$ is the density function that we have to approximate. The reason why asymptotic maximum likelihood methods are popular in estimating parameters is the fact that maximum likelihood estimators are consistent, efficient, and have asymptotic normality. For large samples, if $\hat{\theta}_n$ is a maximum likelihood estimator of $\theta = (\alpha, \sigma, \beta, \mu)$ then

$$\hat{\theta}_n \sim N(\theta, n^{-1}\mathbf{B}),$$

where n is the sample size and $\mathbf{B}$ is the inverse of the Fisher information matrix

$$I_{ij} = \int_{-\infty}^{\infty} \frac{\partial p}{\partial \theta_i} \frac{\partial p}{\partial \theta_j} \frac{1}{p}\, \mathrm{d}x.$$

So the large sample confidence intervals for each of the parameters are:

$$\hat{\theta}_i \pm z_{1-\gamma/2} \frac{\sigma_{\hat{\theta}_i}}{\sqrt{n}},$$

where $\sigma_{\hat{\theta}_1}$, $\sigma_{\hat{\theta}_2}$, $\sigma_{\hat{\theta}_3}$, and $\sigma_{\hat{\theta}_4}$ are the square roots of the diagonal entries of $\mathbf{B}$. Asymptotical maximum likelihood estimators developed by the aforementioned authors converge to maximum likelihood and their quality is determined by the rate of the convergence and easiness of their implementation. In this paper, we present a method for estimating α that is also based upon maximum likelihood estimation and can be efficiently implemented by enumerating in one dimension.

4.2. Parameters of multivariate stable distributions

If $\mathbf{X}$ is a k-dimensional random vector that has a multivariate stable distribution with the location parameter μ, tail index α, and the *spectral measure* Γ then for $\alpha > 1$ $\mu = E\mathbf{X}$ and

$$\hat{\alpha} = \frac{1}{k} \sum_{j=1}^{k} \hat{\alpha}_j,$$

here $\hat{\alpha}_j$ is the estimate of α obtained from the observations of the marginal X_j of $\mathbf{X}$, $j = 1, \ldots, k$ [11, 25]. However, we can not get the *spectral measure* by separate analyses

of univariate marginals because it determines the dependence structure between them [25]. There is a plethora of approaches to estimating the *spectral measure*. Nolan and Panorska [20], develop a method based upon a discrete approximation of the spectral measure. Pivato and Seco [23] estimate the *spectral measure* by its representation as the sum of spherical harmonic functions [23]. McCulloch [15] developed the method of estimating the *spectral measure* of a generalized bivariate stable distribution, based on a series of maximum likelihood (ML) estimates of the stable parameters of univariate projections of the data [15]. For a more detailed description of the *spectral measure* see Samorodnitsky [25].

4.3. Estimators of the dispersion matrix Q of sub-Gaussian

As it was mentioned above, the matrix $\mathbf{Q}$ of sub-Gaussian distributions is a special case of the spectral measure. It will be shown later that when we deal with the dispersion matrix with the rank d, we can estimate its diagonal elements q_{ii}, as the scale parameter of one-dimensional marginals X_i, $i = 1, \ldots, d$ and the elements q_{ij} with $i \neq j$, $i, j = 1, \ldots, d$ can be estimated from the bivariate random vector (X_i, X_j). This means that large dimensions are not a significant impediment in the case of sub-Gaussian distributions. However, in the case of general stable distributions, it is almost impractical to estimate the *spectral measure* if the dimension is higher than 3 [15]. In Kring et. al [11], there is a method of finding a dispersion matrix with moment type estimators. Our method, whose description is provided in the following sections, is based on the maximum likelihood.

5. PARAMETER ESTIMATION FOR THE MULTIVARIATE SUB-GAUSSIAN DISTRIBUTIONS

If we want to estimate the parameters of a multivariate sub-Gaussian distribution with parameters α and $\mathbf{Q}$, we can do it in two stages:

- Estimation of the α parameter,
- Estimation of matrix $\mathbf{Q}$ using the estimate of α as an input,

where

$$\mathbf{Q} = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 r_{12} & \cdots & \cdots & \sigma_1\sigma_d r_{1d} \\ \sigma_1\sigma_2 r_{12} & \sigma_2^2 & \cdots & \cdots & \sigma_2\sigma_d r_{2d} \\ \sigma_1\sigma_3 r_{13} & \sigma_2\sigma_3 r_{23} & \cdots & \cdots & \sigma_d\sigma_3 r_{23} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_1\sigma_d r_{1d} & \sigma_2\sigma_d r_{2d} & \cdots & \cdots & \sigma_d^2 \end{pmatrix}.$$

The former task can be conducted by analyzing the marginal distributions because all of the marginals have the same α parameter. After estimating α, we can put its estimate $\hat{\alpha}$ into the formula of the characteristic function and use it to estimate the matrix $\mathbf{Q}$. Both estimation procedures will be conducted using the methodology of projections.

5.1. Estimation of the tail index

If $\mathbf{X}$ has a d-variate sub-Gaussian distribution with the parameters α, $\mathbf{Q}$, and $\mu = \mathbf{0}$ then every marginal X_i, $i = 1, 2, \ldots, d$ of $\mathbf{X}$ has a stable distribution $S_\alpha(\sigma_i, 0, 0)$ with

the characteristic function of the form:

$$\psi(u) = \exp(-\sigma_i^{\alpha}|u|^{\alpha}).$$

If $p(x,\alpha)$ is a density function of the stable distribution then:

$$I(\alpha) = \int_{-\infty}^{\infty} J^2(x,\alpha)p(x,\alpha)\,\mathrm{d}x, \quad J(x,\alpha) = \frac{\partial L(x,\alpha)}{\partial \alpha} = \frac{\left(\frac{\partial p(x,\alpha)}{\partial \alpha}\right)}{p(x,\alpha)}, \tag{6}$$

$$\hat{\alpha}_{ML} = \left\{ \alpha : \sum_{j=1}^{n} J(X_j,\alpha) = 0 \right\}, \tag{7}$$

where $I(\alpha)$ is the Fisher information and $X_1, X_2, \ldots, X_n$ is the vector of observations. This shift to the sum follows from the expression of the maximum likelihood function

$$L(\mathbf{X},\alpha) = \frac{1}{n}\sum_{i=1}^{n} \ln p(X_j|\alpha),$$

which is an additive function.

5.2. The core of the methodology

Using the methodology of projections, we can approximate the function $J(X,\alpha)$ which enables us to obtain the estimates whose precision converges to that of the ML-estimators, and to calculate the Fisher information. We will express the approximation of the function $J(X,\alpha)$ in terms of $\{1, \exp(it_1X), \exp(it_2X), \ldots, \exp(it_kX)\}$, i. e., its approximation will be in the form $J_k(X,\alpha)$ where

$$J_k(X,\alpha) = \sum_{j=0}^{k} a_j \exp(it_jX) = \sum_{j=0}^{k} a_j \cos(it_jX) + i\sum_{j=0}^{k} a_j \sin(it_jX), \tag{8}$$

where $t_1, \ldots, t_k$ are different constants that can be chosen arbitrarily. We used the constants $t_1, \ldots, t_k$ whose absolute values are smaller than 1 and $a_1, a_2, \ldots, a_k$ are the unknown values that we need to estimate.

Remark. Note that we refer to the almost-sure convergence of $J_k(X,\alpha)$ to $J(X,\alpha)$ provided that $X \sim S_\alpha(\sigma,0,0)$ [8, 27].

The idea of the method described below was proposed by Kagan [8] however, it was applied to other types of distributions where power projections were used instead of trigonometric ones. We project onto the space with a scalar product defined as follows:

If $X \sim S_\alpha(1,0,0)$ and t_m, t_n are constants, then a product between $\exp(it_mX)$ and $\exp(it_nX)$ is defined as

$$\begin{aligned}\langle \exp(it_mX), \exp(it_nX)\rangle &= E\exp(it_mX)\cdot\exp(it_nX) = E\exp(iX(t_m+t_n)) \\ &= \int_{-\infty}^{\infty} p(x,\alpha)\exp(ix(t_m+t_n))\,\mathrm{d}x = \exp\left(-|t_m+t_n|^{\alpha}\right).\end{aligned}$$

For any projection $J_k(x,\alpha)$ it holds:

$$(J_k(x,\alpha) - J(x,\alpha)) \perp \exp(it_j x), j = 1,2,\ldots,k, \tag{9}$$

in other words:

$$\langle (J_k(X,\alpha) - J(X,\alpha)), \exp(it_j X)\rangle = 0, j = 1,2,\ldots,k \tag{10}$$

or

$$E((J_k(X,\alpha) - J(X,\alpha)) \cdot \exp(it_j X)) = 0, j = 1,2,\ldots,k. \tag{11}$$

Hence, we have:

$$\int_{-\infty}^{\infty} (J_k(x,\alpha) - J(x,\alpha)) p(x,\alpha) \exp(it_j x)\,\mathrm{d}x = 0$$

or

$$\int_{-\infty}^{\infty} J_k(x,\alpha) p(x,\alpha) \exp(it_j x)\,\mathrm{d}x = \int_{-\infty}^{\infty} J(x,\alpha) p(x,\alpha) \exp(it_j x)\,\mathrm{d}x.$$

Let us calculate each integral in the above equality separately:

$$\int_{-\infty}^{\infty} J_k(x,\alpha) p(x,\alpha) \exp(it_j x)\,\mathrm{d}x = \int_{-\infty}^{\infty} \sum_{v=0}^{k} a_v \exp(it_v x) p(x,\alpha) \exp(it_j x)\,\mathrm{d}x$$

$$= \sum_{v=0}^{k} a_v \int_{-\infty}^{\infty} \exp(it_v x) p(x,\alpha) \exp(it_j x)\,\mathrm{d}x = \sum_{v=0}^{k} a_v \int_{-\infty}^{\infty} p(x,\alpha) \exp(ix(t_j + t_v))\,\mathrm{d}x$$

$$= \sum_{v=0}^{k} a_v \exp\left(-|t_v + t_j|^{\alpha}\right),\ j = 1,2,\ldots,k.$$

We can reverse the order of the sum and the integral because the number of the items in the sum is finite. The second integral will be calculated as follows:

$$\int_{-\infty}^{\infty} J(x,\alpha) p(x,\alpha) \exp(it_j x)\,\mathrm{d}x = \int_{-\infty}^{\infty} \frac{\left(\frac{\partial p(x,\alpha)}{\partial \alpha}\right)}{p(x,\alpha)} p(x,\alpha) \exp(it_j x)\,\mathrm{d}x$$

$$= \int_{-\infty}^{\infty} \frac{\partial p(x,\alpha)}{\partial \alpha} \exp(it_j x)\,\mathrm{d}x = \frac{\partial}{\partial \alpha} \int_{-\infty}^{\infty} p(x,\alpha) \exp(it_j x) = \frac{\partial}{\partial \alpha} \exp\left(-|t_j|\right)^{\alpha}.$$

The integral and derivative can be interchanged because of the Leibnitz rule and the fact that $p(x,\alpha) \geq 0$. Hence we get the following system of linear equations:

$$\sum_{v=0}^{k} a_v \exp\left(-|t_v + t_j|\right)^{\alpha} = \frac{\partial}{\partial \alpha} \exp\left(-|t_j|\right)^{\alpha}, j = 1,2,\ldots,k,$$

which can be written in the form:

$$\frac{\partial}{\partial \alpha} \exp\left(-|t_j|\right)^{\alpha} = -\exp\left(-|t_j|\right)^{\alpha} \cdot |t_j|^{\alpha} \cdot \ln|t_j|,$$

$$\sum_{v=0}^{k} a_v \exp\left(-|t_v + t_j|^{\alpha}\right) = -\exp\left(-|t_j|\right)^{\alpha} \cdot |t_j|^{\alpha} \cdot \ln|t_j|, j = 1, 2, \ldots, k.$$

In the matrix form, the system looks as follows:

$$\begin{pmatrix} 1 & e^{-|t_1|^{\alpha}} & \ldots & e^{-|t_k|^{\alpha}} \\ e^{-|t_1|^{\alpha}} & e^{-|t_1+t_1|^{\alpha}} & \ldots & e^{-|t_1+t_k|^{\alpha}} \\ \ldots & \ldots & & \ldots \\ \ldots & \ldots & & \ldots \\ \ldots & \ldots & & \ldots \\ e^{-|t_k|^{\alpha}} & e^{-|t_k+t_1|^{\alpha}} & \ldots & e^{-|t_k+t_k|^{\alpha}} \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \ldots \\ \ldots \\ a_k \end{pmatrix} = \begin{pmatrix} 0 \\ -|t_1|^{\alpha} \ln|t_1| e^{-|t_1|^{\alpha}} \\ -|t_2|^{\alpha} \ln|t_2| e^{-|t_2|^{\alpha}} \\ \ldots \\ \ldots \\ -|t_k|^{\alpha} \ln|t_k| e^{-|t_k|^{\alpha}} \end{pmatrix}.$$

We take into account the fact that $\alpha \in (1, 2]$ in financial applications, and substitute its values into the system of equations where $a_1, a_2, \ldots, a_k$ are unknown. We solve this system by manipulating the values of α and choosing such a value of the tail index for which $\left|\sum_{j=1}^{n} J_k(X_j, \alpha)\right|$ is minimal. Note that in this system of equations only $a_1, a_2, \ldots, a_k$ are unknown and nothing else.

If we denote the components of the previous equation as follows

$$\mathbf{A}(\alpha) = \begin{pmatrix} 1 & e^{-|t_1|^{\alpha}} & \ldots & e^{-|t_k|^{\alpha}} \\ e^{-|t_1|^{\alpha}} & e^{-|t_1+t_1|^{\alpha}} & \ldots & e^{-|t_1+t_k|^{\alpha}} \\ \ldots & \ldots & & \ldots \\ \ldots & \ldots & & \ldots \\ \ldots & \ldots & & \ldots \\ e^{-|t_k|^{\alpha}} & e^{-|t_k+t_1|^{\alpha}} & \ldots & e^{-|t_k+t_k|^{\alpha}} \end{pmatrix}, \quad \mathbf{b}(\alpha) = \begin{pmatrix} 0 \\ -|t_1|^{\alpha} \ln|t_1| e^{-|t_1|^{\alpha}} \\ -|t_2|^{\alpha} \ln|t_2| e^{-|t_2|^{\alpha}} \\ \ldots \\ \ldots \\ -|t_k|^{\alpha} \ln|t_k| e^{-|t_k|^{\alpha}} \end{pmatrix},$$

$\mathbf{a}(\alpha) = (\mathbf{A}(\alpha))^{-1} \cdot \mathbf{b}(\alpha)$, and $\overline{F(\mathbf{X})} = \{1, \frac{1}{n}\sum_{m=1}^{n} \exp(it_1 X_m), \ldots, \frac{1}{n}\sum_{m=1}^{n} \exp(it_k X_m)\}$, then

$$\sum_{j=1}^{n} J_k(X_j, \alpha) = \sum_{j=1}^{k} a_j \sum_{m=1}^{n} \exp(it_j X_m) = n \cdot \overline{F(\mathbf{X})} \cdot \mathbf{a}(\alpha). \tag{12}$$

Summary of the methodology of estimating α

Note that $\overline{F(\mathbf{X})}$ can be calculated externally and used as an input into the methodology of projections. This means that the MLP procedure does not depend on the sample size. Let us denote with Ω the finite set of all considered values of α. Then we get a two-phase algorithm:

- 1) Calculation of the $\overline{F(\mathbf{X})}$,
- 2) Choosing such a value of α from Ω which minimizes $|\overline{F(\mathbf{X})} \cdot \mathbf{a}(\alpha)|$.

The resulting estimator is

$$\alpha_{\text{MLP}} = \arg\min_{\alpha \in \Omega} \left|\overline{F(\mathbf{X})} \cdot \mathbf{a}(\alpha)\right|. \tag{13}$$

In Table 1, we compare the maximum likelihood projections method for $k = 15$ with the CFB estimates. The empirical mean and variance of the estimates of α are obtained from simulating 100 samples with stable distributions, and every sample contains 5000 elements. From those 100 estimates we calculate the mean and variance. In Table 1, μ_{est} and σ_{est} denote the mean and the standard deviation of the estimates respectively.

α	Type of estimator	μ_{est}	σ^2_{est}	$\mu_{est} \pm 2\sigma_{est}$
1.1	α_{MLP}	1.1052	0.017	[1.070,1.139]
	α_{CFB}	1.0997	0.02	[1.059,1.140]
1.2	α_{MLP}	1.2054	0.018	[1.167,1.243]
	α_{CFB}	1.2005	0.021	[1.059.1.140]
1.3	α_{MLP}	1.3000	0.018	[1.262,1.337]
	α_{CFB}	1.3000	0.022	[1.254,1.345]
1.4	α_{MLP}	1.4074	0.022	[1.363,1.451]
	α_{CFB}	1.3997	0.023	[1.350,1.440]
1.5	α_{MLP}	1.5010	0.021	[1.458,1.543]
	α_{CFB}	1.5002	0.024	[1.450,1.540]
1.7	α_{MLP}	1.6988	0.020	[1.638,1.759]
	α_{CFB}	1.7002	0.021	[1.637,1.762]
1.8	α_{MLP}	1.8004	0.018	[1.764,1.836]
	α_{CFB}	1.8006	0.020	[1.755,1.845]
1.9	α_{MLP}	1.8988	0.015	[1.868,1.929]
	α_{CFB}	1.9000	0.019	[1.860,1.931]

Tab. 1. Comparison of the Mean and Standard Deviation of the MLP for $k = 15$ and the CFB Estimates. Simulated values from samples with 5000 elements.

5.3. Auxiliary statements

The following lemma facilitates the estimation of the dependence structure of sub-Gaussian distributions because if we use it, we will be able to find $\mathbf{Q}$ by using pairs of marginal distributions.

Lemma 5.1. The estimates of parameters r_{ij}, $i, j = 1, 2, \ldots, d$ can be obtained only from the marginals X_i and X_j, $i, j = 1, 2, \ldots, d$. If $\sigma_i = \sigma_j = 1$ then

$$\Psi_{ij}(u_i, u_j) = \exp\left\{-\left|(u_i, \quad u_j) \cdot \begin{pmatrix} 1 & r_{ij} \\ r_{ij} & 1 \end{pmatrix} \cdot \begin{pmatrix} u_i \\ u_j \end{pmatrix}\right|^{\alpha/2}\right\}.$$

In particular, if $r_{ij} = 1$ then $X_j = X_i$ a.s. If $r_{ij} = -1$ then $X_j = -X_i$ a.s. However $r_{ij} = 0$ does not imply independence. r_{ij} is a measure of the linear dependence between the marginals X_i and X_j, $i, j = 1, 2, \ldots, d$ [11].

Proof.

$$\mathbf{\Psi}(0,0,\dots,u_i,0,\dots,0,u_j,0,\dots 0) = \Psi_{ij}(u_i,u_j)$$

$$= \exp\left\{-\left|(u_i, \quad u_j)\cdot\begin{pmatrix} 1 & r_{ij} \\ r_{ij} & 1 \end{pmatrix}\cdot\begin{pmatrix} u_i \\ u_j \end{pmatrix}\right|^{\alpha/2}\right\},$$

which follows from the elementary operations of products of matrices and proves the first statement. If X is a random variable with the characteristic function

$$\Psi_X(\mathbf{u}) = \exp\left(-|\mathbf{u}^T\mathbf{Q}\mathbf{u}|^{\alpha/2}\right)$$

with

$$\mathbf{Q} = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 r_{12} & \cdots & \cdots & \sigma_1\sigma_d r_{1d} \\ \sigma_1\sigma_2 r_{12} & \sigma_2^2 & \cdots & \cdots & \sigma_2\sigma_d r_{2d} \\ \sigma_1\sigma_3 r_{13} & \sigma_2\sigma_3 r_{23} & \cdots & \cdots & \sigma_d\sigma_3 r_{23} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_1\sigma_d r_{1d} & \sigma_2\sigma_d r_{2d} & \cdots & \cdots & \sigma_d^2 \end{pmatrix},$$

then the random vector $Y = (X_1/\sigma_1,\dots,X_d/\sigma_d)$ has the dependence structure of the form

$$\mathbf{q} = \begin{pmatrix} 1 & r_{12} & \cdots & \cdots & r_{1d} \\ r_{12} & 1 & \cdots & \cdots & r_{2d} \\ r_{13} & r_{23} & \cdots & \cdots & r_{23} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{1d} & r_{2d} & \cdots & \cdots & 1 \end{pmatrix},$$

where $r_{i,j} \in [-1,1]$, $i,j = 1,2,\dots,d$.

In other words, if $X = (X_1, X_2, \dots, X_d)$ has a symmetric stable distribution with matrix $\mathbf{Q}$ and the tail index α, then the random vector $Y = (1/\sigma_1 X_1, \dots, 1/\sigma_d X_d)$ has a symmetric stable distribution with the matrix $\mathbf{q}$ and the tail index α.

If $G \sim N(\mathbf{0},\mathbf{Q})$ and $s \sim S_{\alpha/2}\left(\left(\cos\left(\frac{\pi\alpha}{4}\right)\right)^{2/\alpha}, 1, 0\right)$, and s and G independent, then

$$X = \sqrt{s}G \sim S_\alpha(\mathbf{0},\mathbf{Q}),$$

i. e., its characteristic function is $\Psi_X(\mathbf{u}) = \exp\left(-|\mathbf{u}^T\mathbf{Q}\mathbf{u}|^{\alpha/2}\right)$ If $\mathbf{G}$ is 2-dimensional, with $\mathbf{Q}$ of the form

$$\begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix},$$

then according to the Cholesky factorization [7],

$$\mathbf{G} = (G_1, G_2), \quad G_1 = Z_1, \quad G_2 = rZ_1 + \sqrt{1-r^2}Z_2, \quad Z_1, Z_2 \quad i.i.d. \quad Z_1, Z_2 \sim N(0,1)$$

and

$$\mathbf{X} = \sqrt{s}\mathbf{G}, \quad X_1 = \sqrt{s}Z_1, \quad X_2 = \sqrt{s}(rZ_1 + \sqrt{1-r^2}Z_2).$$

Hence $r = 1$ implies $X_1 = X_2$ and $r = -1$ implies $X_1 = -X_2$.

Let $X_1 \sim S_\alpha(1,0,0)$ and $X_2 \sim S_\alpha(1,0,0)$ be independent. Then the characteristic function of the random vector $X = (X_1, X_2)$ is of the form

$$\Psi_X(\mathbf{u}) = \exp\left\{-|u_1|^\alpha - |u_2|^\alpha\right\}.$$

However, the characteristic function of the sub-Gaussian random vector Y with the tail index α and

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is of the form

$$\Psi_Y(\mathbf{u}) = \exp\left\{-|u_1^2 + u_2^2|^{\alpha/2}\right\}, \quad \Psi_X(\mathbf{u}) \neq \Psi_Y(\mathbf{u}).$$

□

Remark. The exactness of the parameter estimates of matrix $\mathbf{Q}$ does not diminish as the dimension increases. The problem of estimating $\mathbf{Q}$ is parallelizable because Lemma 5.1 enables us to concentrate on pairs of marginal distributions.

To estimate $\mathbf{Q}$, we can determine r_{ij} from the pairs X_i and X_j, $i, j = 1, 2, \ldots, d$. First, we estimate parameters $\sigma_i = r_{ii}$ $i = 1, 2, \ldots, d$ from the univariate marginal distributions and then we estimate r_{ij}, $i, j = 1, 2, \ldots, d$. The number of such pairs is

$$\begin{pmatrix} d \\ 2 \end{pmatrix} = \frac{d(d-1)}{2}.$$

Values r_{ij} can be calculated separately and if the number of parallel processors is at least $\frac{d(d-1)}{2}$ then the calculation will be accelerated $\frac{d(d-1)}{2}$ times, which equals the number of the parameters $r_{i,j} = r_{j,i}$ $i, j = 1, 2, \ldots, d$ of the sub-Gaussian distribution.

$$r = 0 \quad \Rightarrow X_1 = \sqrt{s}Z_1, \quad X_2 = \sqrt{s}Z_2, \quad Z_1, Z_2 \sim N(0,1), \quad i.i.d.$$

$$\begin{pmatrix} d \\ 2 \end{pmatrix} = \frac{d(d-1)}{2} = O(d^2).$$

5.4. Parameter estimation of a two-dimensional sub-Gaussian distribution with known α and unknown Q

Without loss of generality, we can assume that matrix $\mathbf{Q}$ is of the form

$$\mathbf{Q} = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}$$

with unknown r because we can always get this form by scaling and translating of the initial random vector. Lemma 5.1 enables us to concentrate only on pairs of marginal distributions of the sub-Gaussian distribution to be able to estimate the whole matrix $\mathbf{Q}$. Let us denote the density function of the two-dimensional random vector by $p(x, y; r)$. Suppose that α is known because its estimate can be obtained from the univariate marginal distributions.

This random vector has the characteristic function of the form:

$$\psi(u_1, u_1; r) = \exp\left\{-(2(1+r))^{\alpha/2} u_1^{\alpha}\right\}. \tag{14}$$

We will use the method of projections to obtain the MLP estimate of r as follows:

$$I(r) = E\left[\left(\frac{\partial p(X, Y; r)}{\partial r}\right)^2\right], \tag{15}$$

$$J(x, y, r) = \frac{\partial p(x, y, r)}{\partial r}. \tag{16}$$

The ML estimate of r will be obtained by solving the equation

$$\hat{r}_{ML} = \left\{ r : \sum_{j=1}^{n} J(X_j, Y_j, r) = 0 \right\}, \tag{17}$$

where $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ are the observations. A technique analogous to the one that we used for estimating function J for α can be used for estimating r [8]. In this case, we will obtain to another system of equations but the core of the approach is much the same. Let us project the random function $J(X, Y, r)$ to the space $\{1, \mathrm{e}^{-it_1(X+Y)}, \ldots, \mathrm{e}^{-it_k(X+Y)}\}$, $k \in N$. We assume that $t_0 = 0$. Let us define the scalar product in this space in a way similar to the case of estimating α:

$$\left\langle \mathrm{e}^{-it_i(X+Y)}, \mathrm{e}^{-it_j(X+Y)} \right\rangle = E\left[\mathrm{e}^{-it_i(X+Y)} \cdot \mathrm{e}^{-it_j(X+Y)}\right]$$

$$= E\left[\mathrm{e}^{-i(t_i+t_j)(X+Y)}\right] = \psi(t_i + t_j, t_i + t_j)$$

$$\psi(t_i + t_j, t_i + t_j; r) = \exp\left\{(2(1+r))^{\alpha/2} |t_i + t_j|^{\alpha}\right\}. \tag{18}$$

Let us approximate $J(X, Y, r)$ by

$$J_k(X, Y, r) = \sum_{j=0}^{k} a_j \cdot \mathrm{e}^{it_j(X+Y)}. \tag{19}$$

Because we project $J(X, Y, r)$ to the space $\{1, \mathrm{e}^{-it_1(X+Y)}, \ldots, \mathrm{e}^{-it_k(X+Y)}\}$, $k \in N$, we have

$$(J(X, Y, r) - J_k(X, Y, r)) \perp \mathrm{e}^{it_j(X+Y)}, \quad j = 1, 2, \ldots, k. \tag{20}$$

Hence we have

$$\int_{-\infty}^{\infty} (J(x, y, r) - J_k(x, y, r))\, \mathrm{e}^{it_j(x+y)} p(x, y, r)\, \mathrm{d}x\mathrm{d}y = 0$$

or

$$\int_{-\infty}^{\infty} J(x,y,r)\mathrm{e}^{it_j(x+y)}p(x,y,r)\,\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{\infty} J_k(x,y,r)\mathrm{e}^{it_j(x+y)}p(x,y,r)\,\mathrm{d}x\mathrm{d}y.$$

Calculating the first integral yields

$$\int_{-\infty}^{\infty} J(x,y,r)\mathrm{e}^{it_j(x+y)}p(x,y,r)\,\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{\infty} \frac{\left(\frac{\partial p(x,y,r)}{\partial r}\right)}{p(x,y,r)}\mathrm{e}^{it_j(x+y)}p(x,y,r)\,\mathrm{d}x\mathrm{d}y$$

$$= \int_{-\infty}^{\infty}\left(\frac{\partial p(x,y,r)}{\partial r}\right)\mathrm{e}^{it_j(x+y)}\,\mathrm{d}x\mathrm{d}y = \frac{\partial}{\partial r}\int_{-\infty}^{\infty} p(x,y,r)\mathrm{e}^{it_j(x+y)}\,\mathrm{d}x\mathrm{d}y = \frac{\partial}{\partial r}\psi(t_j,t_j;r)$$

where

$$\frac{\partial}{\partial r}\psi(t_j,t_j;r) = \frac{\partial}{\partial r}\exp\left\{-(2(1+r))^{\alpha/2}t_j^{\alpha}\right\}$$

$$= -\exp\left\{-(2(1+r))^{\alpha/2}t_j^{\alpha}\right\}2^{\alpha/2-1}\alpha t_j^{\alpha}(1+r)^{\alpha/2-1}.$$

Calculating the second integral yields

$$\int_{-\infty}^{\infty} J_k(x,y,r)p(x,y,r)\mathrm{e}^{it_j(x+y)}\,\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{\infty}\sum_{m=0}^{k} a_m\mathrm{e}^{it_m(x+y)}\mathrm{e}^{it_j(x+y)}p(x,y,r)\,\mathrm{d}x\mathrm{d}y$$

$$= \sum_{m=0}^{k} a_m\int_{-\infty}^{\infty}\mathrm{e}^{i(t_m+t_j)(x+y)}p(x,y,r)\,\mathrm{d}x\mathrm{d}y = \sum_{m=0}^{k} a_m\psi(t_j+t_m,t_j+t_m;r)$$

and

$$\sum_{m=0}^{k} a_m\psi(t_j+t_m,t_j+t_m;r) = \sum_{m=0}^{k} a_m\exp\left(-|t_j+t_m|^{\alpha}(2(1+r))^{\alpha/2}\right) \tag{21}$$

and finally we get the following system of equations

$$\sum_{m=0}^{k} a_m\exp\left(-|t_j+t_m|^{\alpha}(2(1+r))^{\alpha/2}\right) = -\exp\left\{-(2(1+r))^{\alpha/2}t_j^{\alpha}\right\}\alpha t_j^{\alpha}(2(1+r))^{\alpha/2-1}, \tag{22}$$

where $j = 0,1,\ldots,k$.

If we denote the components of the previous equation in the following way:

$$A(r) = \begin{pmatrix} \exp\left(-|t_0+t_0|^{\alpha}(2(1+r))^{\alpha/2}\right) & \cdots & \cdots & \cdots & \exp\left(-|t_0+t_k|^{\alpha}(2(1+r))^{\alpha/2}\right) \\ \exp\left(-|t_1+t_0|^{\alpha}(2(1+r))^{\alpha/2}\right) & \cdots & \cdots & \cdots & \exp\left(-|t_1+t_k|^{\alpha}(2(1+r))^{\alpha/2}\right) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \exp\left(-|t_k+t_0|^{\alpha}(2(1+r))^{\alpha/2}\right) & \cdots & \cdots & \cdots & \exp\left(-|t_k+t_k|^{\alpha}(2(1+r))^{\alpha/2}\right) \end{pmatrix},$$

$$\mathbf{a}(r) = \begin{pmatrix} a_0 \\ a_1 \\ \cdots \\ \cdots \\ a_k \end{pmatrix}, \mathbf{b}(r) = \begin{pmatrix} -\exp\left\{-(2(1+r))^{\alpha/2} t_0^\alpha\right\} 2^{\alpha/2-1} \alpha t_0^\alpha (1+r)^{\alpha/2-1} \\ -\exp\left\{-(2(1+r))^{\alpha/2} t_1^\alpha\right\} 2^{\alpha/2-1} \alpha t_1^\alpha (1+r)^{\alpha/2-1} \\ \cdots \\ \cdots \\ -\exp\left\{-(2(1+r))^{\alpha/2} t_k^\alpha\right\} 2^{\alpha/2-1} \alpha t_k^\alpha (1+r)^{\alpha/2-1} \end{pmatrix},$$

then the equations of the MLP methodology will be as follows

$$\mathbf{A}(r)\mathbf{a} = \mathbf{b}(r), \quad \mathbf{a}(r) = (\mathbf{A}(r))^{-1}\mathbf{b}(r).$$

We can get $J_k(X, Y; r)$ without the imaginary part after making the following assumptions:
1)

$$J_k(X, Y; r) = \sum_{j=-k}^{k} a_j \exp(it_j(X+Y));$$

2) $a_j = a_{-j}$, $t_j = -t_{-j}$; and
3) $t_0 = 0$.

Under these assumptions, we have

$$J_k(X, Y; r) = \sum_{j=-k}^{k} a_j \exp(it_j(X+Y))$$

$$= a_0 \exp(it_0(X+Y)) + \sum_{j=1}^{k} (a_j \exp(it_j(X+Y)) + a_j \exp(-it_j(X+Y)))$$

$$= a_0 \exp(it_0(X+Y)) + \sum_{j=1}^{k} a_j 2\cos((X+Y)t_j).$$

If we assume $t_0 = 0$, the imaginary part will equal zero:

$$J_k(X, Y; r) = a_0 + \sum_{j=1}^{k} a_j 2\cos((X+Y)t_j)$$

or, if we denote the vector $(1, 2\cos((X+Y)t_1), \ldots, 2\cos((X+Y)t_k))$ by $F(X, Y)$ the function J_k will take on the following matrix form

$$J_k(X, Y, r) = (1, 2\cos((X+Y)t_1), \ldots, 2\cos((X+Y)t_k)) \begin{pmatrix} a_0 \\ a_1 \\ \cdots \\ a_k \end{pmatrix} = F(X, Y)(A(r))^{-1} b(r).$$

If we solve the equation

$$\left\{ r : \sum_{j=1}^{n} J_k(X_j, Y_j, r) = 0 \right\}$$

or

$$\left\{ r : \left(\sum_{j=1}^{n} F(X_j, Y_j) \right) (\mathbf{A}(r))^{-1} \mathbf{b}(r) = 0 \right\}. \tag{23}$$

It is equivalent to replacing the sum with the mean:

$$\left\{ r : \overline{F(\mathbf{X}, \mathbf{Y})} (\mathbf{A}(r))^{-1} \mathbf{b}(r) = 0 \right\}, \tag{24}$$

where $\overline{F(\mathbf{X}, \mathbf{Y})} = \frac{1}{n} \sum_{j=1}^{n} F(X_j, Y_j)$.

The latter means that the level of complexity of the problem does not rise with the number of observations. We only need to calculate the mean of k samples $\cos(t_j(X_1 + Y_1)), \ldots, \cos(t_j(X_n + Y_n))$, $j = 1, 2, \ldots, k$ and substitute those values in the equation (23). Then, taking into account that $r \in (-1, 1)$, we will partition the interval and choose the value of r for which the absolute value of the expression in (23) is the smallest. The cases in which $r = 1$ or $r = -1$ are trivial because if $r = 1$ then $X = Y$ a.s. and if $r = -1$ then $X = -Y$ a.s.

Summary of the methodology of estimating r

Let us denote with Ω the finite set of all considered values of r. Then we get a two-phase algorithm:

- 1) Calculation of the $\overline{F(\mathbf{X}, \mathbf{Y})}$,
- 2) Choosing such a value of r from Ω which minimizes $|\overline{F(\mathbf{X}, \mathbf{Y})} \cdot \mathbf{a}(r)|$.

The resulting estimator is

$$r_{\text{MLP}} = \arg\min_{r \in \Omega} \left| \overline{F(\mathbf{X}, \mathbf{Y})} \cdot \mathbf{a}(r) \right|. \tag{25}$$

In Table 2 we compare the estimates of r by MLP methodology for $k = 15$ with those obtained by CFB. We simulated one hundred samples with 5000 elements of the observations of sub-Gaussian distributions where all of their univariate marginals have the distribution $S_{1.5}(1, 0, 0)$ to estimate the dispersion matrix by means of the two aforementioned methodologies. Column "r" shows the real value of the parameter, column μ_{est} shows the means of the estimates of the 100 parameters whose real value is in column "r". Column σ_{est} shows the standard deviations of the 100 estimates by MLP methodology. Column $\mu_{est} \pm 2\sigma_{est}$ shows 2-σ confidence intervals of the estimates which have a normal distribution [8, 14, 19]. In the majority of cases, the 2-σ interval for MLP are thinner than those of CFB.

Convergence rate

The convergence rate of MLP estimates of α presented in Table 3. We simulate 100 samples from $S_{1.5}(1, 0, 0)$ with 5000 elements and estimate α by means of the MLP

r	Type of estimator	μ_{est}	σ_{est}	$\mu_{est} \pm 2\sigma_{est}$
0.0	r_{MLP}	-0.0027	0.034	[-0.070,0.041]
	r_{CFB}	0.00012	0.045	[-0.089,0.090]
0.1	r_{MLP}	0.10123	0.036	[0.029,0.173]
	r_{CFB}	0.10212	0.043	[0.016,0.188]
0.2	r_{MLP}	0.19981	0.032	[0.135,0.263]
	r_{CFB}	0.20123	0.039	[0.123,0.279]
0.3	r_{MLP}	0.30121	0.029	[0.243,0.359]
	r_{CFB}	0.29986	0.037	[0.225,0.373]
0.4	r_{MLP}	0.40012	0.028	[0.344,0.456]
	r_{CFB}	0.39876	0.033	[0.332,0.464]
0.5	r_{MLP}	0.50087	0.028	[0.444,0.556]
	r_{CFB}	0.50176	0.031	[0.439,0.563]
0.6	r_{MLP}	0.59987	0.026	[0.547,0.651]
	r_{CFB}	0.60002	0.025	[0.550,0.650]
0.7	r_{MLP}	0.70012	0.025	[0.650,0.750]
	r_{CFB}	0.69900	0.027	[0.645,0.753]
0.8	r_{MLP}	0.78891	0.026	[0.736,0.840]
	r_{CFB}	0.80078	0.027	[0.746,0.854]
0.9	r_{MLP}	0.89989	0.018	[0.861,0.937]
	r_{CFB}	0.90003	0.019	[0.862,0.938]

Tab. 2. Comparison of the mean and standard deviation of the MLP estimates of r for $k = 15$ to those of the CFB estimates for $\alpha = 1.5$.

methodology in the software *Mathematica 9*. When $k \geq 15$, we can not see any improvement in terms of the precision of the mean of the estimates, the smallness of the standard deviation of the estimates and the thinness of the $\mu \pm 3\sigma$ intervals. For the estimates of r, we also can not observe any improvement in precision for $k \geq 15$ and have analogous results.

CONCLUSION

In this paper, we presented sub-Gaussian distributions and the technique of estimating their parameters that is based on the maximum likelihood and yields estimators that converge to ML ones according to Kagan [8]. According to our results, this convergence is fast and for $k = 15$, it outperforms CFB in precision. The methodology is based on operations that are elementary for modern computers: finding average values, finding inverses of matrices and product of matrices. This means that neither a large number of observations nor a large k are significant impediments. The methodology takes into account special properties of sub-Gaussian distributions and enables us to find estimates of dispersion matrix of any dimension.

k	Computation Time	μ_{est}	σ_{est}	$\mu_{est} \pm 3\sigma_{est}$
1	0.265	1.49322	0.03872	[1.3704,1.6094]
2	0.390	1.50028	0.02610	[1.4219,1.5786]
3	0.570	1.49900	0.02632	[1.4001,1.5779]
4	0.780	1.49709	0.02370	[1.4259,1.5682]
5	1.061	1.50309	0.02259	[1.4353,1.5708]
6	1.357	1.50359	0.02105	[1.4402,1.5667]
7	1.702	1.50109	0.02094	[1.4382,1.5682]
8	2.106	1.50241	0.02116	[1.4389,1.5658]
9	2.558	1.49919	0.02269	[1.4312,1.5671]
10	2.995	1.49844	0.02216	[1.4319,1.5649]
15	6.225	1.50003	0.02333	[1.4300,1.5700]
20	10.514	1.50119	0.01979	[1.4418,1.5605]
30	21.528	1.50197	0.02146	[1.4375,1.5663]
40	37.191	1.50025	0.02181	[1.4348,1.5656]
50	57.408	1.49341	0.02085	[1.4308,1.5559]
100	249.43	1.49775	0.02236	[1.4306,1.5648]
150	515.33	1.49906	0.02088	[1.4364,1.5617]
200	896.18	1.49500	0.02112	[1.4312,1.5587]
400	3513.91	1.49916	0.02192	[1.4333,1.5649]
1000	21970.6	1.49800	0.02043	[1.4367,1.5593]

Tab. 3. The rate of convergence of the estimates of α when its true value is 1.5. Simulated values from samples with 5000 elements.

ACKNOWLEDGEMENT

This work was supported in part by Institute of Information Theory and Automation, Academy of sciences of the Czech Republic. The author is grateful to Prof. Lev Klebanov for his supervision during my research for diploma thesis, which inspired the present paper, and to Vlasta Kaňková, CSc. for valuable advice and discussion. This research was partially supported by the Czech Science Foundation under Grant No. 13-144458.

(Received December 13, 2011)

REFERENCES

[1] M. Carrasco and J. Florens: Generalization of GMM to a continuum moment condition. Econom. Theory *16* (2000), 767–834.

[2] J. M. Chambers, C L. Mallows, and B. W. Stuck: A method for simulating stable random variables. J. Amer. Statist. Assoc. *71* 1976), 340–344.

[3] B. N. Cheng and S. Rachev: Multivariate stable securities in financial markets. Math. Finance *5* (1995), 133–153.

[4] W. H. DuMouchel: Stable Distributions in Statistical Inference. PhD. Thesis, University of Ann Arbor, Ann Arbor 1971.

[5] E. Fama: Portfolio analysis in a stable Paretian market. Management Sci. *11* (1965), 404–419.

[6] B. M. Hill: A simple general approach to inference about the tail of a stable distribution. Ann. Stat. *3* (1975), 5, 1163–1174.

[7] R. A. Horn and C. R. Johnson: Matrix Analysis. Cambridge University Press 1985.

[8] A. Kagan: Fisher information contained in a finite-dimensional linear space, and a properly formulated version of the method of moments (in Russian). Problemy Peredachi Informatsii *12* (2009), 15–29.

[9] L. Klebanov: Heavy Tailed Distributions. Matfyzpress, Prague 2003.

[10] I. A. Koutrovelis: Regression-type estimation of the parameters of stable laws. J. Amer. Statist. Assoc. *75* (1980), 918–928.

[11] S. Kring, S. Rachev, M. Höchstötter, and F. J. Fabozzi: Estimation of Alpha-Stable Sub-Gaussian Distributions for Asset Returns. In: Risk Assessment: Decisions in Banking and Finance. Physica-Verlag, Heidelberg 2008, pp. 111–152.

[12] D. B. Madan and E. Seneta: The variance gamma model from shared market returns. J. Bus. *63* (1990), 511–524.

[13] B. Mandelbrot: The variation of certain speculative prices. J. Bus. *26* (1963), 394–419.

[14] J. H. McCulloch: Simple consistent estimators of stable distribution parameters. Commun. Statist. — Simula *15* (1986), 1109–1136.

[15] J. H. McCulloch: Estimation of the bivariate stable spectral representation by the projection method. Comput. Econom. *16* (2000), 47–62.

[16] S. Mittnik and S. Rachev: Tail estmation of the stable index alpha. Applied mathematics. Letters *9* (1996), 3, 53–56.

[17] S. Mittnik and M. S. Paolella: Prediction of Financial Downside-Risk with Heavy-Tailed Conditional Distributions. In: Handbook of Heavy Tailed Distributions in Finance: Book *1* (2003), pp. 385–403.

[18] J. P. Nolan: Modeling Financial Data with Stable Distributions. In: Handbook of Heavy Tailed Distributions in Finance, Handbooks in Finance: Book *1* (2003), pp. 105–130.

[19] J. P. Nolan: Maximum likelihood estimation and diagnostics for stable distributions. In: Lévy Processes (O. E. Barndorff-Nielsen, T. Mikosch, and S. Resnick, eds.), Brikhauser, Boston 2001.

[20] J. P. Nolan and A. K. Panorska: Data analysis for heavy tailed multivariate samples. Commun. Statist.: Stochastic Models (1997), 687—702.

[21] V. Omelchenko: Elliptical stable distributions. In: Mathematical Methods in Economics 2010 (M. Houda and J. Friebelova, eds.), pp. 483-488.

[22] S. Ortobelli, I. Huber, S. Rachev, and E. S. Schwarz: Portfolio Choice Theory with Non-Gaussian Distributed Return. In: Handbook of Heavy Tailed Distributions in Finance, Handbooks in Finance: Book *1* (2003), pp. 547–594.

[23] M. Pivato and L. Seco: Estimating the spectral measure of a multivariate stable distribution via spherical harmonic analysis. J. Multivariate Anal. *87* (2003), 2, 219–240.

[24] S. T. Rachev, E. S. Schwarz, and I. Khindanova: Stable Modeling of Market and Credit Value at Risk. In: Handbook of Heavy Tailed Distributions in Finance, Handbooks in Finance: Book *1* (2003), pp. 255–264.

[25] G. Samorodnitsky and M. S. Taqqu: Stable Non-Gaussian Random Processes. Chapman and Hall 1994.

[26] P. Schmidt: An improved version of Quandt–Ramsey MGF estimator for mixtures of normal distributions and switching regressions. Econometrica *50* (1982), 501–524.

[27] L. Slámová and L. Klebanov: Modeling financial returns by discrete stable distributions. In: Proc. 30th International Conference Mathematical Methods in Economics 2012.

[28] K. C. Tran: Estimating mixtures of normal distribution via empirical characteristic function. Econom. Rev. *17* (1998), 167–83.

[29] V. Zolotarev: On representation of stable laws by integrals selected translation. Math. Statist. Probab. *6* (1986), 84–88.

Vadym Omelchenko, Institute of Information Theory and Automation — Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
e-mail: omelchenko@utia.cas.cz