STATISTICA, anno LXXIII, n. 3, 2013

MULTIVARIATE ELLIPTICALLY CONTOURED AUTOREGRESSIVE PROCESS

Taras Bodnar
Department of Mathematics, Humboldt-University of Berlin, Berlin, Germany
Arjun K. Gupta
Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, USA

1. INTRODUCTION

The historically first approach applied for modeling financial data is based on the normal distribution. Fama (1976) found that the normal assumption provides a good fit in case of data taken with a monthly or smaller frequency. From the other side, heavy-tailed distributions should be maintained for more frequent observations. Fama (1965) proposed to use the mixture of normal distribution, while Blattberg and Gonedes (1974) compared the ability of the multivariate t-distribution and the symmetric stable distribution to fit real data. All these distributions belong to the class of elliptically contoured distributions which has been already applied in modeling financial data. For instance, Owen and Rabinovitch (1983) extended several well-known in finance theorems, like, Tobin's separation theorem, Bawa's rules of ordering certain prospects to elliptically contoured distributions. Chamberlain (1983) showed that this family of distributions implies mean-variance utility functions. More recently, Berk (1997) proved that the one of the necessary conditions of the validity of the capital asset pricing model (CAPM) is an elliptical distribution of the asset returns. Zhou (1993) and Hodgson *et al.* (2002) suggested tests for the CAPM under the assumption of the elliptical symmetry, while Bodnar and Gupta (2009a) derived an exact confidence set for the efficient frontier assuming that the matrix of the asset returns follows a matrix variate elliptically contoured distribution.

Following Fang and Zhang (1990) a random vector $\mathbf{X}$ of dimension k is elliptically contoured distributed if its density function exists and has a form

$$f(\mathbf{x}) = det(\Sigma)^{-\frac{k}{2}}\, h((\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)), \tag{1}$$

where $h : [0,\infty) \to [0,\infty)$. This distribution is denoted by $E_k(\mu,\Sigma,h)$. In case $h(x) = (2\pi)^{-\frac{k}{2}} exp(-x/2)$ the random vector $\mathbf{X}$ has a k-variate normal distribution. Note that the density function of the elliptically contoured distributed random vectors does not obviously exist. The more general definition of the family is based on the characteristic function. However, in the present paper we restrict ourselves to the subclass of elliptical distributions that possess the density function and they are defined by Eq. (1). One of the most important property of the elliptically family, which makes it attractive for financial applications, is that the linear combinations of the vector $\mathbf{X}$ have the same type of distribution as the vector $\mathbf{X}$ itself, e.g. for each $p \times k$ dimensional matrix of constants $\mathbf{L}$ the distribution of $\mathbf{LX}$ is $E_p(\mathbf{L}\mu, \mathbf{L}\Sigma\mathbf{L}', h)$.

The stochastic representation of the random vector $\mathbf{X}$ is essential in the theory of elliptical distributions. It holds that $\mathbf{X} \sim E_k(\mu, \Sigma, h)$ if and only if $\mathbf{X}$ has the same distribution as $\mu + \tilde{R}\,\Sigma^{1/2}\,\mathbf{U}$, where $\mathbf{U}$ is a k-variate random vector uniformly distributed on the unit sphere in $I\!R^k$, $\tilde{R}$ is a nonnegative random variable, and $\tilde{R}$ and $\mathbf{U}$ are independent (see Fang and Zhang (1990)). The expression $\mu + \tilde{R}\,\Sigma^{1/2}\,\mathbf{U}$ is a stochastic representation of $\mathbf{X}$, i.e. it holds that

$$\mathbf{X} \stackrel{d}{=} \mu + \tilde{R}\,\Sigma^{1/2}\,\mathbf{U}, \tag{2}$$

where the symbol $A \stackrel{d}{=} B$ says that the two random variables A and B have the same distribution. The variable $\tilde{R}$ is called the generating variable of $\mathbf{X}$ and it fully determines the distribution of $\mathbf{X}$. The distribution of $\tilde{R}^2$ is equal to the distribution of $||\mathbf{X} - \mu||^2_\Sigma$, where $||\mathbf{y}||_\Sigma$ is the norm of the vector $\mathbf{y}$ with respect to the positive definite matrix Σ equal to $\sqrt{\mathbf{y}'\Sigma^{-1}\mathbf{y}}$. If $\mathbf{X}$ is absolutely continuous, then $\tilde{R}$ is also absolutely continuous and its density is

$$f_{\tilde{R}}(r) = \frac{2\pi^{nk/2}}{\Gamma(nk/2)}\, r^{nk-1}\, h(r^2) \tag{3}$$

for $r \geq 0$ (cf. Fang *et al.* (1990, Theorem 2.9)).

When the random vector $\mathbf{X}$ follows a mixture of normal distributions then the stochastic representation (2) transforms to

$$\mathbf{X} \stackrel{d}{=} \mu + R\,\Sigma^{1/2}\,\varepsilon, \tag{4}$$

where $R = \tilde{R}/||\varepsilon||_{\mathbf{I}_k}$, $\varepsilon \sim \mathcal{N}_k(0, \mathbf{I}_k)$, and R, $||\varepsilon||_{\mathbf{I}_k}$, ε are mutually independently distributed. $\mathbf{I}_k$ is an identity matrix of size k. In this case both densities of $\mathbf{X}$ and R do exist for all $k \geq 1$.

Table 1 summarizes the above mentioned multivariate models with their univariate counterparts. They all have a similar stochastic representation, where μ denotes the location vector and Σ the scale matrix. The only difference is in the behavior of the so-called pseudo-generating variable R^2, which is independent of the standard normally distributed random vector ε_t. In case of the normal distribution it is just a constant equal to 1. For the multivariate t-distribution with n degrees of freedom it is the square root of n divided on a χ^2-distributed random variable with n-degrees of freedom. For the symmetric stable law R^2 follows an univariate non-symmetric stable distribution with the index of stability equals to $\alpha/2$, the mean 0, the variance $(cos(\pi\alpha/4))^{\frac{2}{\alpha}}$, and the skewness parameter 1 (see Samorodnitsky and Taqqu (1994)). The similar structure we also observe for the univariate GARCH process, i.e. the conditional volatility σ_t and ε_t are independently distributed (see, e.g. Bollerslev (1986), Pawlak and Schmid (2001)). We preserve this property for the suggested in the paper multivariate generalization given in Section 2. One of the main advantages of the approach is that it satisfies the necessary and sufficient conditions of the capital asset pricing model derived by Berk (1997). From

TABLE 1

Univariate and multivariate models of financial data. In the table ε_t is a random variable normally distributed with the mean 0 and the variance σ^2. $\boldsymbol{\varepsilon}_t$ has a k-dimensional normal distribution, $\mathbf{u}_t$ is a k-dimensional random vector uniformly distributed on the unit sphere in $I\!R^k$. $\mathbf{U} = (\mathbf{v}_1, ..., \mathbf{v}_n)$ with $vech(\mathbf{U})$ to be uniformly distributed on the unit sphere in $I\!R^{nk}$. $S_\alpha(\gamma, \mu, \sigma^2)$ denotes an univariate α-stable distribution with the skewness parameter γ, the mean μ and the variance σ^2.

Name	Univariate	Multivariate
Normal	$X_t = \mu + 1 \cdot \varepsilon_t$	$\mathbf{X}_t = \mu + 1 \cdot \boldsymbol{\varepsilon}_t$
t_n-distribution	$X_t = \mu + \frac{\varepsilon_t}{\sqrt{\chi_t/n}}, \chi_t \sim ii\chi_n^2$	$\mathbf{X}_t = \mu + \frac{\boldsymbol{\varepsilon}_t}{\sqrt{\chi_t/n}}, \chi \sim ii\chi_n^2$
α-symmetric stable	$X_t = \mu + \sqrt{A_t}\varepsilon_t$ $A_t \sim iiS_{\alpha/2}(1, 0, (cos(\pi\alpha/4))^{\frac{2}{\alpha}})$	$\mathbf{X}_t = \mu + \sqrt{A_t}\boldsymbol{\varepsilon}_t$, $A_t \sim iiS_{\alpha/2}(1, 0, (cos(\pi\alpha/4))^{\frac{2}{\alpha}})$
Mixture of normal	$X_t = \mu + R_t \cdot \varepsilon_t$	$\mathbf{X}_t = \mu + R_t \cdot \boldsymbol{\varepsilon}_t$
Multivariate Elliptical Contoured		$\mathbf{X}_t = \mu + \tilde{R}_t \cdot \Sigma^{1/2}\mathbf{u}_t$
Matrix Elliptical Contoured	$(X_1, ..., X_n)' = \mu + R \cdot (\varepsilon_1, ..., \varepsilon_n)'$	$(\mathbf{X}_1, ..., \mathbf{X}_n) = \mu + \tilde{R} \cdot \Sigma^{1/2}\mathbf{U}$
GARCH Process	$X_t = \sigma_t \varepsilon_t$, $\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2$	

the other side it possesses the generality of the multivariate GARCH processes keeping the time varying structure of the conditional covariance matrix.

The autoregressive method has already been applied for modeling a density function of the returns. For example Hansen (1994) used it to design the process that is based on the t-distribution with time varying degrees of freedom. Rockinger and Jondeau (2002) considered the process with time varying higher moments, i.e. the conditional time varying skewness and kurtosis were modeled. In the present paper the autoregressive technic is used to forecast the future values of the generating variable of the multivariate process, that fully specifies the unconditional distribution of the process.

The rest of the paper is organized as follows. In the next section the multivariate elliptically contoured autoregressive (MElAR) process is suggested. Its distributional properties are studied in Section 2.1. The two-stage maximum likelihood estimator of the process parameters is given in Section 3. In Section 3.1, we derived the Stein-Haff identity for the introduced stochastic model. An empirical example is presented in Section 4. Here, we fit the MElAR process to real data of the EUR/USD and EUR/JPY exchange rate returns and show that the non-diagonal elements of the dispersion matrix are slowly varying in time. Final remarks are given in Section 5.

2. Multivariate Elliptically Contoured Autoregressive Process

In this section we introduce a new class of elliptically contoured processes, that allows us to model the dependence between the process realizations in a non-trivial way. Moreover, the presented model puts together the conditional time varying properties of multivariate GARCH processes and the elliptical symmetry of mixtures of normal distributions.

Let $\mathbf{X}_t = (X_{t,1}, ..., X_{t,k})'$, $t = 1, 2, ...$ be a sequence of k-dimensional random vectors.

The model is given by

$$\mathbf{X}_t = \mu + R_t \Sigma^{1/2} \varepsilon_t, \quad (5)$$

$$R_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i R_{t-i}^2, \quad (6)$$

where $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i \in \{1, ..., p\}$.
The assertion we denote by $\{\mathbf{X}_t\} \sim MElAR_k(\mu, \Sigma, p)$. ε_t's are assumed to be independently identically normally distributed with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{I}_k$, i.e. $\varepsilon_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_k)$. Let $\mathscr{F}_t$ denote the information available up to time point t. We assume that ε_t is independent of $\mathscr{F}_{t-1}$. As a special case of this assumption we get that R_t and ε_t are independent for all $t = 1, 2, ...$ as well since R_t is fully determined by the previous information.

The idea behind the MElAR process is to replace in Eq. (4) the unobservable pseudo-generating variable R by its forecast R_t given the information available at time point t. From Eq. (5) and the independency of R_t and ε_t it follows that $\mathbf{X}_t$ is elliptically contoured distributed, i.e $\mathbf{X}_t \sim E_k(\mu, \Sigma, h)$ for some function $h(.)$. More precisely, $\mathbf{X}_t$ follows a mixture of normal distributions with an unknown shape of elliptical symmetry. Its density function does always exist and it is completely specified by the density function of R_t (see Eq. (3) and Fang *et al.* (1990, Section 2.6)).

The designed process possesses both generality of the GARCH process and symmetric properties of elliptical distributions. On the other side, the model (5) and (6) cannot be considered as a special case of a multivariate GARCH-type process. First, the idea behind the process (5) and (6) is to model the conditional density, whereas the multivariate GARCH processes are models for the conditional covariance matrix. Second, the unconditional distribution of any multivariate GARCH process is not elliptically symmetric since the generating variable R_t is replaced by a conditional covariance matrix in the definition of a multivariate GARCH process.

Similarly to the multivariate GARCH process the model (5) and (6) assumes that the conditional covariance matrix $\Sigma_{t|t-1} = R_t \Sigma$ is time varying, whereas the dispersion matrix Σ is time invariant. The last assumption can be violated by considering the time varying dispersion matrix Σ. This generalization is not treated in the paper, but presents a possible extension of the obtained results and it is left for future researches. Moreover, a further extension of the $MElAR_k(\mu, \Sigma, p)$ process can be considered. It is given by the model (5) and

$$R_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i R_{t-i}^2 + \sum_{j=1}^{q} \beta_j ||\mathbf{X}_{t-j} - \mu||_{\Sigma}^2, \quad (7)$$

where $\beta_j \geq 0$, $j = 1, ..., q$. This process we denote by $MElAR_k(\mu, \Sigma, p, q)$. The main properties of the process (5) and (7) are given in Section 2.1.

The one of the most important application of the MElAR process leads to the portfolio theory. Because of its elliptical structure the suggested process satisfies the necessary conditions of the validity of the CAPM model derived by Berk (1997). Fitting the process (5) and (6) allows us to model the expected utility portfolio with time-varying coefficient of the investor's risk aversion. Moreover, since the dispersion matrix Σ is time invariant, it holds that the expected return of the global minimum variance portfolio is time invariant. The only two parameters of the efficient frontier (the set of the all mean-variance optimal portfolios) that depend on R_t^2 are the variance of the global minimum variance portfolio and the slope parameter (see Bodnar and Gupta (2009a)).

2.1. *Distributional Properties*

The first property of the MElAR process follows directly from its elliptical behavior.

LEMMA 1. *Assume that* $\{\mathbf{X}_t\} \sim MElAR_k(\mu, \Sigma, p, q)$*. Let* $\mathbf{L}$ *be* $k \times k$ *nonsingular matrix of constants. Then* $\{\mathbf{L}\mathbf{X}_t\} \sim MElAR_k(\mathbf{L}\mu, \mathbf{L}\Sigma\mathbf{L}', p, q)$ *given by*

$$\mathbf{L}\mathbf{X}_t = \mathbf{L}\mu + R_t \mathbf{L}\Sigma^{1/2}\varepsilon_t, \tag{8}$$

$$R_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i R_{t-i}^2 + \sum_{j=1}^{q} \beta_j ||\mathbf{L}\mathbf{X}_{t-j} - \mathbf{L}\mu||_{\mathbf{L}\Sigma\mathbf{L}'}^2, \tag{9}$$

The coefficients in the recursive equations for R_t^2 are the same in both the formulas (7) and (9). Note, that the similar property does not hold for matrices $\mathbf{L}$ of order $p \times k$, $p < k$. The result follows from the fact that $||\mathbf{L}\mathbf{X}_{t-i} - \mathbf{L}\mu||_{\mathbf{L}\Sigma\mathbf{L}'}^2 = b_{p/2,(k-p)/2}||\mathbf{X}_{t-i} - \mu||_{\Sigma}^2$, where the random variable $b_{p/2,(k-p)/2}$ has a beta-distribution with parameters $p/2$ and $(k-p)/2$, and $b_{p/2,(k-p)/2}$ and $||\mathbf{X}_{t-i} - \mu||_{\Sigma}^2$ are independently distributed (see Fang *et al.* (1990, p. 39)). Correspondingly, it follows that a linear combination of the components of the $MElAR_k(\mu, \Sigma, p, q)$ process $\{\mathbf{l}'\mathbf{X}_t\}$ does not follow an univariate GARCH process. Although the process $\{\mathbf{l}'\mathbf{X}_t\}$ has a similar structure to the GARCH process, the coefficients in the conditional variance equation are positive random variables. If $\{\mathbf{X}_t\} \sim MElAR_k(\mu, \Sigma, p)$ it follows that $\{\mathbf{L}\mathbf{X}_t\} \sim MElAR_p(\mathbf{L}\mu, \mathbf{L}\Sigma\mathbf{L}', p)$ with the same coefficient α_0, α_i, $j = 1, ..., q$ for all $p \times k$ matrices $\mathbf{L}$. Moreover, $\{\mathbf{l}'\mathbf{X}_t\}$ is an univariate GARCH(0,p) process.

An important problem is a statement not only about the distribution of $\mathbf{X}_t$ but also about the joint density of the arbitrary sequence of $(\mathbf{X}_{t_1}, ..., \mathbf{X}_{t_s})$ generated by the MElAR process. In order to shed light on the problem, first, we study the independency structure of the process.

LEMMA 2. *Let* $\{\mathbf{X}_t\} \sim MElAR_k(\mu, \Sigma, p, q)$*. Then for each* s, t_1, ..., t_s *the sequences* $(R_{t_1}, ..., R_{t_s})$ *and* $(\varepsilon_{t_1}, ..., \varepsilon_{t_s})$ *are independently distributed.*

The result of the lemma follows from the fact that for each t_i the generating variable $R_{t_i}^2$ depends on ε_j, $j < t_i$ only through the quadratic forms $||\varepsilon_j||_{\Sigma}^2$.
Because $(..., ||\varepsilon_0||_{\Sigma}^2, ..., ||\varepsilon_{t_s}||_{\Sigma}^2)$ and $(..., \varepsilon_0, ..., \varepsilon_{t_s})$ are independently distributed the same holds for $(R_{t_1}, ..., R_{t_s})$ and $(\varepsilon_{t_1}, ..., \varepsilon_{t_s})$. From Lemma 2 we get that the joint unconditional distribution of $(\mathbf{X}_{t_1}, ..., \mathbf{X}_{t_s})$ is a matrix mixture of normal distributions.

COROLLARY 3. *Let* $\{\mathbf{X}_t\} \sim MElAR_k(\mu, \Sigma, p, q)$*. Then the random matrix* $\mathbf{X}_{t_1,...,t_s} = (\mathbf{X}_{t_1}, ..., \mathbf{X}_{t_s})$ *is matrix elliptically distributed with the stochastic representation given by*

$$(\mathbf{X}_{t_1}, ..., \mathbf{X}_{t_s}) \stackrel{d}{=} \Sigma^{1/2}(\varepsilon_{t_1}, ..., \varepsilon_{t_s}) diag\{(R_{t_1}, ..., R_{t_s})\},$$

where $(\varepsilon_{t_1},...,\varepsilon_{t_s})$ *and* $(R_{t_1},...,R_{t_s})$ *are independently distributed. Moreover, the sequence* $(\varepsilon_{t_1},...,\varepsilon_{t_s})$ *consists of independently identically distributed random vectors* $\varepsilon_{t_i} \sim \mathcal{N}_k(\mathbf{0},\mathbf{I}_k)$.

The results of Lemma 2 are also used to derive the moments of the MElAR process.

COROLLARY 4. *Let* $\{\mathbf{X}_t\} \sim MElAR_k(\mu,\Sigma,p,q)$. *Then*
a) $E(\mathbf{X}_t) = \mu$;
b) $Var(\mathbf{X}_t) = E(R_t^2)\Sigma$, *where* $E(R_t^2) = \frac{\alpha_0}{1-\sum_{i=1}^p \alpha_i}$;
c) $Cov(\mathbf{X}_t,\mathbf{X}_{t-i}) = \mathbf{0}$.

PROOF. a) The statement follows directly from the definition of the MElAR process and the fact that $E(\varepsilon_t) = \mathbf{0}$.
b) The first assertion holds because of the elliptical properties of the MElAR process. The identity $E(R_t^2) = \alpha_0/(1-\sum_{i=1}^p \alpha_i)$ follows from the autoregressive structure of the process $\{R_t^2\}$.
c) It holds that

$$Cov(\mathbf{X}_t,\mathbf{X}_{t-i}) = E((\mathbf{X}_t-\mu)(\mathbf{X}_{t-i}-\mu)') = \Sigma^{1/2}E(R_t^2R_{t-i}^2)E(\varepsilon_t\varepsilon_{t-1})\Sigma^{1/2},$$

where the last equality follows from Lemma 2. Because $E(\varepsilon_t\varepsilon_{t-1}) = \mathbf{0}$ the statement of part c) is proved. □

Next, we consider $MElAR_k(\mu,\Sigma,1)$ process. For each t we obtain that

$$R_t^2 = \alpha_0 + \alpha_1 R_{t-1}^2 = \alpha_0(1+\alpha_1) + \alpha_1^2 R_{t-2}^2 = ... \tag{10}$$

$$= \alpha_0 \sum_{i=0}^{t-1} \alpha_1^i + \alpha_1^t R_0^2 = \frac{\alpha_0}{1-\alpha_1}(1-\alpha_1^t) + \alpha_1^t R_0^2. \tag{11}$$

where R_0^2 is the intial value of the process $\{R_t^2\}$. The last equality yields

$$(\mathbf{X}_1,...,\mathbf{X}_t) = \Sigma^{1/2}(\varepsilon_1,...,\varepsilon_t)diag\{(\sqrt{\alpha_0+\alpha_1 R_0^2},...,\sqrt{\frac{\alpha_0}{1-\alpha_1}(1-\alpha_1^t)+\alpha_1^t R_0^2})\}.$$

From the last presentation and Table 1 we conclude that the $MElAR_k(\mu,\Sigma,1)$ process is an extension of the matrix variate mixture of normal distributions that allow us to model the time varying behavior of the generating variable. Moreover, because of its similarity to the matrix variate elliptically contoured distributions, the distributional results of the latter can be extended to the suggested model. We provide a further discussion of this property in Section 3.1, where the Stein-Haff identity is generalized to the MElAR process.

The necessary and sufficient condition of the $MElAR_k(\mu,\Sigma,p,q)$ process to be weakly stationarity is given by

$$\sum_{i=1}^p \alpha_i + k\sum_{j=1}^q \beta_j < 1. \tag{12}$$

The inequality (12) follows from the proof of Theorem 1 of Bollerslev (1986) and the fact that $E(||\mathbf{X}_t||^2_\Sigma) = k$ for each t. The necessary and sufficient condition of the strictly stationarity are similar to those given in Bougerol and Picard (1992) and it is based on the top Lyapunov exponent (see Bougerol and Picard (1992) for details).

This section we finish with a further property of the MElAR process, which shows its relationship to the ARMA time series. Let $\tau_t = ||\mathbf{X}_t - \mu||^2_\Sigma - R_t^2$. Then the Eq. (7) can be rewritten in the following form

$$||\mathbf{X}_t||^2_\Sigma = \alpha_0 + \sum_{j=1}^{\max\{q,p\}} (\alpha_j + \beta_j)||\mathbf{X}_{t-j}||^2_\Sigma + \tau_t - \sum_{i=1}^{p} \alpha_i \tau_{t-i}\,. \tag{13}$$

Hence, it holds

LEMMA 5. *Let* $\{\mathbf{X}_t\} \sim MElAR_k(\mu, \Sigma, p, q)$. *Then*

$$\{||\mathbf{X}_{t-i} - \mu||^2_\Sigma\} \sim ARMA(max\{p,q\}, p).$$

3. ESTIMATION

First, we assume that the mean vector μ is known. Later, it is shown how this assumption can be violated. The rest of parameters of the $MElAR_k(\mu, \Sigma, p)$ process we denote by θ which are divided into two groups, i.e. $\theta = (\theta_1, \theta_2)$ with $\theta_1 = (\alpha_0, \alpha_1, ..., \alpha_p)'$ to be the vector of the parameters given in (6) and $\theta_2 = vech(\Sigma)$ corresponds to the elements of the dispersion matrix Σ. For estimating the parameters of the MElAR process, the two-stage quasi maximum likelihood method is applied (see, e.g. Engle (2002)). In the first stage θ_1 is estimated, whose estimator is used in the second stage for estimating θ_2.

Because $\mathbf{X}_t|\mathscr{F}_{t-1} \sim \mathscr{N}(\mu, R_t^2\Sigma)$, the quasi-likelihood function is given by

$$\begin{aligned}
QL(\theta|\mathbf{X}_t) &= -\frac{1}{2}\sum_{t=1}^{T}\Big(k\log(2\pi) + k\log(R_t^2) + \log(det(\Sigma)) + \frac{(\mathbf{X}_t-\mu)'\Sigma^{-1}(\mathbf{X}_t-\mu)}{R_t^2}\Big)\\
&= -\frac{1}{2}\sum_{t=1}^{T}\Big(k\log(2\pi) + k\log(R_t^2) + \frac{(\mathbf{X}_t-\mu)'(\mathbf{X}_t-\mu)}{R_t^2}\\
&+ \log(det(\Sigma)) + \frac{(\mathbf{X}_t-\mu)'\Sigma^{-1}(\mathbf{X}_t-\mu)}{R_t^2} - \frac{(\mathbf{X}_t-\mu)'(\mathbf{X}_t-\mu)}{R_t^2}\Big)\\
&= k l_1(\theta_1) + l_2(\theta_2|\theta_1),
\end{aligned}$$

where

$$l_1(\theta_1) = -\frac{1}{2}\sum_{t=1}^{T}\Big(\log(2\pi) + \log(R_t^2) + \frac{(\mathbf{X}_t-\mu)'(\mathbf{X}_t-\mu)/k}{R_t^2}\Big) \tag{14}$$

is the quasi log-likelihood function of the univariate GARCH(0,p) process applied to the process $\{(\mathbf{X}_t-\mu)'(\mathbf{X}_t-\mu)/k\}$.

In the second stage the dispersion matrix Σ is estimated by maximizing the normal likelihood function

$$l_2(\theta_2|\hat{\theta}_1) = -\frac{T}{2}\log(det(\Sigma)) - \frac{1}{2}\sum_{t=1}^{T}(\hat{\mathbf{X}}_t'\Sigma^{-1}\hat{\mathbf{X}}_t) \tag{15}$$

with $\hat{\mathbf{X}}_t = (\mathbf{X}_t - \mu)/\hat{R}_t$. It leads to

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}\frac{(\mathbf{X}_t - \mu)(\mathbf{X}_t - \mu)'}{\hat{R}_t^2}. \tag{16}$$

Next, the asymptotic properties of the suggested estimators are studied. Here, we use the results of Engle and Sheppard (2001) who considered the two-stage quasi maximum likelihood estimation of the DCC process and argued that this estimation procedure can be presented as a two stage GMM estimation studied by Newey and McFadden (1994) in detail. It is assumed that
Assumption 1: $\theta^0 = (\theta_1^0, \theta_2^0)'$ is an identifiably unique interior in $\mathbf{\Theta} = \mathbf{\Theta}_1 \times \mathbf{\Theta}_2$, where $\mathbf{\Theta}$ is compact. Moreover, it is assumed that Σ^0 is positive definit, $\alpha_0^0 > 0$, and $\alpha_i^0 \geq 0$ for $i = 1, ..., p$.
Assumption 2: θ_1^0 uniquely maximizes $E(ln(l_1(\theta_1)))$ and θ_2^0 uniquely maximizes $E(ln(l_2(\theta_2|\theta_1)))$.
Assumption 3: The first and second stage quasi log-likelihoods, i.e. $ln(l_1(\theta_1))$ and $ln(l_2(\theta_2|\theta_1))$ are twice continuously differentiable on θ^0.
Assumption 4: $E(sup_{\theta_1 \in \Theta_1} ||ln(l_1(\theta_1))||)$ and $E(sup_{\theta_2 \in \Theta_2} ||ln(l_2(\theta_2|\theta_1))||)$ exist and are finite.
Assumption 5a: $E(\nabla_{\theta_1} ln(l_1(\theta_1^0))) = 0$ and $E(|| \nabla_{\theta_1} ln(l_1(\theta_1^0))||^2) < \infty$.
Assumption 5b: $E(\nabla_{\theta_2} ln(l_2(\theta_2^0|\theta_1^0))) = 0$ and $E(|| \nabla_{\theta_2} ln(l_2(\theta_2^0|\theta_1^0))||^2) < \infty$.
Assumption 6a: $\mathbf{A}_{11} = E(\nabla_{\theta_1\theta_1} ln(l_1(\theta_1^0)))$ is $O(1)$ and negative definite.
Assumption 6b: $\mathbf{B}_{11} = E(\nabla_{\theta_2\theta_2} ln(l_2(\theta_2^0|\theta_1^0)))$ is $O(1)$ and negative definite.

The asymptotic distribution of $\hat{\theta}$ is given in Theorem 6.

THEOREM 6. *Let* $\{\mathbf{X}_t\} \sim MElAR_k(\mu, \Sigma, p)$. *Then*
a) under Assumptions 1-4, $\hat{\theta}_1 \xrightarrow{P} \theta_1^0$ *and* $\hat{\theta}_1 \xrightarrow{P} \theta_2^0$*;*
b) under Assumptions 1-6,

$$\sqrt{T}(\hat{\theta} - \theta_0) \overset{a}{\sim} N(0, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}'^{-1}),$$

where

$$\mathbf{A} = \begin{pmatrix} E(\nabla_{\theta_1\theta_1} ln(l_1(\theta_1^0))) & 0 \\ E(\nabla_{\theta_1\theta_2} ln(l_2(\theta_2^0|\theta_1^0))) & E(\nabla_{\theta_2\theta_2} ln(l_2(\theta_2^0|\theta_1^0))) \end{pmatrix}$$

and

$$\mathbf{B} = Var(\sum_{t=1}^{T} T^{-1/2} \nabla_{\theta_1} ln(l_1(\theta_1^0)), \sum_{t=1}^{T} T^{-1/2} \nabla_{\theta_2} ln(l_2(\theta_2^0|\theta_1^0))).$$

The proof of Theorem 6 follows from the proof of Theorems 1 and 2 of Engle and Sheppard (2001).

Up to now, the mean vector μ was assumed to be known. Next, we simplify this assumption. In Theorem 7, it is shown that $\hat{\mu} = \bar{\mathbf{X}} = \frac{1}{T}\sum_1^T \mathbf{X}_T$ is a consistent estimator of μ.

THEOREM 7. *Let* $\{\mathbf{X}_t\} \sim MElAR_k(\mu, \Sigma, p)$. *Then* $\hat{\mu} \xrightarrow{P} \mu$ *for* $T \to \infty$.

PROOF. First, we note that $\hat{\mu} = \bar{\mathbf{X}}$ is an unbiased estimator of μ which follows directly from the definition of $\hat{\mu}$. Next, we calculate the mean-square error of the estimator $\hat{\mu}$. It holds that

$$MSE_{\mu}(\hat{\mu}) = Var(\hat{\mu}) = \frac{1}{T^2}\sum_{i=1}^{T} Var(\mathbf{X}_i),$$

where the last equality follows from Corollary 4. Hence,

$$MSE_{\mu}(\hat{\mu}) = \frac{1}{T}\frac{\alpha_0}{1 - \sum_{i=1}^{p}\alpha_i}\Sigma \to 0 \quad \text{for} \quad T \to \infty. \quad \square$$

Because $\hat{\mu}$ is a consistent estimator of μ, the results of Theorem 6 do not change when μ is replaced by $\hat{\mu}$. Hence, the estimation of the $MElAR_k(\mu, \Sigma, p)$ process can be perform in the following three steps:

Step 1: Estimate μ by $\hat{\mu} = \bar{\mathbf{X}}$.

Step 2: Fit the univariate GARCH(0,p) process to $(\mathbf{X}_t - \hat{\mu})'(\mathbf{X}_t - \hat{\mu})/k$ for estimating θ_1.

Step 3: The dispersion matrix Σ is estimated as in (16), where $\hat{R}_t^2$ are calculated recursively with $\hat{\theta}_1$ instead of θ_1.

3.1. The Stein-Haff Identity

Since the seminal paper of Stein (1956) and James and Stein (1961), different estimators for the covariance matrix of the normal distribution have been proposed (see, e.g. Haff (1980, 1991), Dey and Srinivasan (1985)) with a detailed survey given in Kubokawa (2005). The performance of every estimator is based on a risk function. In order to simplify the comparison of the risk functions, the Stein-Haff identity was derived by Stein (1977) and Haff (1979a). This identity was extended to the inverse Wishart distribution by Haff (1979b), while Kubokawa and Srivastava (1999) and Bodnar and Gupta (2009b) obtained the Stein-Haff identity for different classes of the matrix variate elliptically contoured distribution.

For estimating the dispersion matrix of an elliptically contoured distribution the sample covariance matrix is, usually, used which is given by $\mathbf{S} = \tilde{\mathbf{X}}\tilde{\mathbf{X}}'$ where $\tilde{\mathbf{X}} = (\mathbf{X}_1 - \mu, \mathbf{X}_2 - \mu, ..., \mathbf{X}_n - \mu)$. The stochastic representation of $\mathbf{S}$ is

$$\mathbf{S} = \mathbf{X}\mathbf{X}' \stackrel{d}{=} \Sigma^{1/2}\mathbf{U}\mathbf{R}^2\mathbf{U}'\Sigma^{1/2}. \tag{17}$$

Let $\mathbf{G}(\mathbf{S})$ be a $k \times k$ matrix such that the $(i,j)^{th}$ element $g_{ij}(\mathbf{S})$ is a function of $\mathbf{S}$. Let

$$\{\mathbf{D}_S\mathbf{G}(\mathbf{S})\}_{ij} = \sum_l d_{il}g_{lj}(\mathbf{S}),$$

where

$$d_{il} = \frac{1}{2}(1 + \delta_{il})\frac{\partial}{\partial s_{il}},$$

with $\delta_{il} = 1$ for $i = l$ and $\delta_{il} = 0$ for $i \neq l$.

In Theorem 8 we generalize the results of Stein (1977) and Haff (1979a) to the case of the MElAR process.

THEOREM 8. *Let* $\{\mathbf{X}_t\} \sim MElAR_k(\mu, \Sigma, p)$. *We assume that* $g_{ij}(\mathbf{S})$, $i, j \in \{1, ..., k\}$, *is a linear function of the elements of* $\mathbf{S}$. *Let* $n > k$ *and* Σ *be positive definite. Then we have*

$$E(tr(\mathbf{G}(\mathbf{S})\Sigma^{-1})) = \frac{E(R_1^2)}{k}E_{\Sigma}^F((n-k-1)tr(\mathbf{G}(\mathbf{S})\mathbf{S}^{-1}) + 2tr(\mathbf{D}_S\mathbf{G}(\mathbf{S}))), \tag{18}$$

where

$$E_{\Sigma}^F(h(\mathbf{X}\mathbf{X}')) = \int h(\mathbf{X})|\Sigma|^{-n/2}F(tr(\Sigma^{-1}\mathbf{X}\mathbf{X}'))\,d\mathbf{X},$$

with $F(x) = \frac{1}{2}\int_x^{+\infty}\exp(-t/2)dt$.

The proof of Theorem 8 follows from the proof of Theorem 1 of Bodnar and Gupta (2009b).

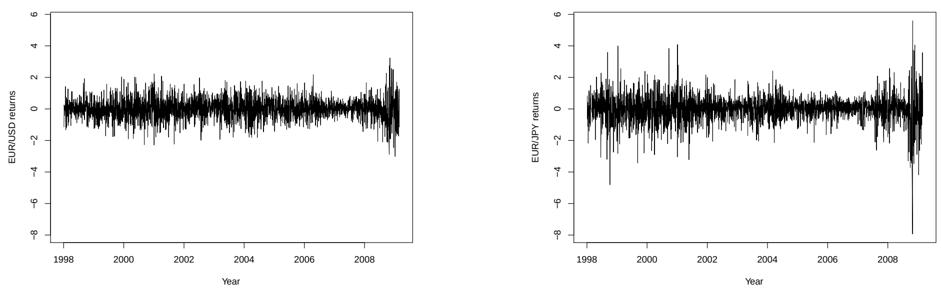

Figure 1 – Daily EUR/USA and EUR/JPY exchange rate returns for the period from January 2, 1999 to November 30, 2009.

4. Estimating the Dispersion Matrix of EUR/USA and JPY/USA Exchange Rate Returns

Two of the most important conditions of a model application are an easiness in estimating of the model's parameters and an ability to forecast future values. In Section 3, we propose the two-stage maximum likelihood method for estimating the parameters of the MElAR process. In the present section, it is shown how the suggested model can be applied to real data by estimating the dispersion matrix of the EUR/USA ($i = 1$) and EUR/JPY ($i = 2$) exchange rate returns. For these purposes, we consider daily EUR/USA and EUR/JPY exchange rate returns data for the period from January 2, 1999 to November 30, 2009 (see Figure 1).

For the comparison purposes we plot the rolling estimator of the unconditional covariance matrix in Figure 2 given by

$$\hat{\mathbf{V}} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{X}-\bar{\mathbf{X}})(\mathbf{X}-\bar{\mathbf{X}})'.$$

The estimation window is set to be 250. In the figure, we observe that the estimator of the covariance matrix is time varying. Significant increases in the elements of $\hat{\mathbf{V}}$ are present at the end of 2001 and during the financial crisis in 2008.

The $MElAR(\mu, \Sigma, 1)$ process is fitted to the considered data by using the two-stage maximum likelihood estimator of Section 3. It holds that $\alpha_0 = 0.6944$ $\alpha_1 = 0.6605$ which are both significant at 1% level of significance. Using these values, we calculate the realizations of the pseudo-generating variable $\hat{R}^2$. The dispersion matrix $\hat{\Sigma}$ is estimated as in (16) by using the rolling estimation with the estimation window of 250 days. The results are presented in Figure 3. We observe that a large amount of the time variable behavior of the covariance matrix of the exchange rate returns is included in the pseudo-generating variable of the process. It increases significantly in the time periods, when the elements of the unconditional covariance matrix are larger. From the other side we get the considerable reduction of the time variability in the dispersion matrix, especially, in the non-diagonal elements of the dispersion matrix which appears to be slowly varying with time.

5. Summary

In the present paper we suggest a new class of multivariate elliptically contoured processes. The proposed process possesses both generality of the GARCH models and symmetric properties of

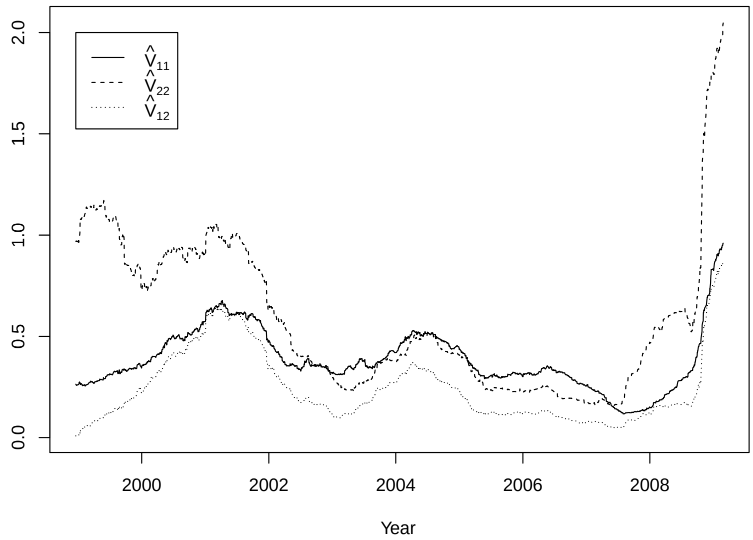

Figure 2 – The rolling estimator for the unconditional covariance matrix with the estimation window of 250 days.

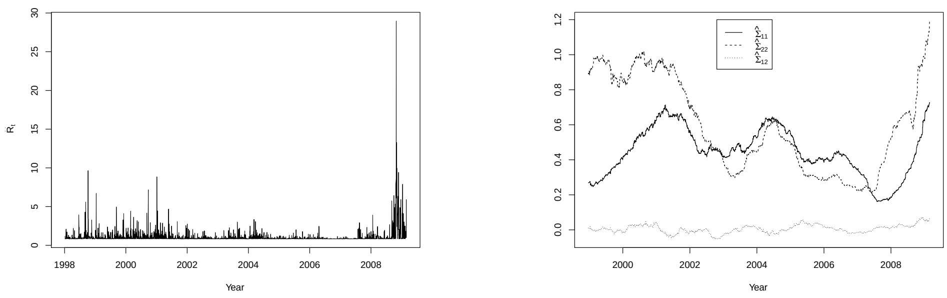

Figure 3 – Estimators for the generating variable and the dispersion matrix of the daily EUR/USA and EUR/JPY exchange rate returns for the period from January 2, 1999 to November 30, 2009. For estimating of the dispersion matrix the rolling estimator is used with the estimation window equals to 250 days.

elliptically contoured distributions. Moreover, the number of unknown parameters of the model is significantly reduced in comparison to other multivariate GARCH processes.

In the empirical study the daily EUR/USD and EUR/JAP exchange rate returns are used for estimating the dispersion matrix of the MElAR process. The parameters of the process are estimated by the two-stage quasi maximum likelihood method. The obtained results do not support the volatile behavior of the dispersion matrix. The time variability of the covariance matrix is explained by the time varying behavior of the generating variable that influences the coefficient of the investor's risk aversion (see, e.g. Bodnar and Gupta (2009a)).

The possible generalization of the obtained results can be done in two ways. First, heavy tailed distributions, like the multivariate t-distribution, can be considered as a model for the error process. Second, the time varying dispersion matrix can be incorporated into the model. It could be done by using the exponential smoothing or to model the time varying dispersion matrix keeping the idea of the DCC process (see Engle (2002)).

ACKNOWLEDGEMENTS

The authors are thankful to the Referee and the Editor for their suggestions which have improved the presentation in the paper.

REFERENCES

J. BERK (1997). *Necessary conditions for the capm*. Journal of Economic Theory, 73, pp. 245–257.

R. BLATTBERG, N. GONEDES (1974). *A comparison of the stable and student distributions as statistical models for stock prices*. The Journal of Business, 47, pp. 244–280.

T. BODNAR, A. GUPTA (2009a). *Construction and inferences of the efficient frontier in elliptical models*. Journal of the Japan Statistical Society, 39, pp. 193–207.

T. BODNAR, A. GUPTA (2009b). *An identity for multivariate elliptically contoured matrix distribution*. Statistics and Probability Letters, 79, pp. 1327–1330.

T. BOLLERSLEV (1986). *Generalized autoregressive conditional heteroscedasticity*. Journal of Econometrics, 31, pp. 307–327.

P. BOUGEROL, N. PICARD (1992). *Stationarity of garch processes and of some nonnegative time series*. Journal of Econometrics, 52, pp. 115–127.

G. CHAMBERLAIN (1983). *A characterization of the distributions that imply mean-variance utility functions*. Journal of Economic Theory, 29, pp. 185–201.

D. DEY, C. SRINIVASAN (1985). *Estimation of covariance matrix under stein's loss*. Annals of Statistics, 13, pp. 1581–1591.

R. ENGLE (2002). *Dynamic conditional correlation – a simple class of multivariate garch models*. Journal of Business and Economic Statistics, 20, pp. 339–350.

R. ENGLE, K. SHEPPARD (2001). *Theoretical and empirical properties of dynamic conditional correlation multivariate garch*. NBER Working Paper 8554.

E. FAMA (1965). *The behavior of stock market price*. Journal of Business, 38, pp. 34–105.

E. FAMA (1976). *Foundations of Finance*. Basic Books: New York.

K. FANG, S. KOTZ, K. NG (1990). *Symmetric Multivariate and Related Distributions*. London: Chapman and Hall.

K. FANG, Y. ZHANG (1990). *Generalized Multivariate Analysis*. Berlin: Springer-Verlag and Beijing: Science Press.

L. HAFF (1979a). *An identity for the wishart distribution with applications*. Journal of Multivariate Analysis, 9, pp. 531–542.

L. HAFF (1979b). *Estimation of the inverse convariance matrix: Random mixtures of the inverse wishart matrix and the identity*. Annals of Statistics, 7, pp. 1264–1276.

L. HAFF (1980). *Empirical bayes estimation of the multivariate normal covariance matrix*. Annals of Statistics, 8, pp. 586–597.

L. HAFF (1991). *The variational form of certain bayes estimators*. Annals of Statistics, 19, pp. 1163–1190.

B. HANSEN (1994). *Autoregressive conditional density estimation*. International Economic Review, 35, pp. 705–730.

D. HODGSON, O. LINTON, K. VORKINK (2002). *Testing the capital asset pricing model efficiency under elliptical symmetry: a semiparametric approach*. Journal of Applied Econometrics, 17, pp. 617–639.

W. JAMES, C. STEIN (1961). *Estimation with quadratic loss*. In *Proc. Fourth Berkeley Symp. Math. Statist. Prob. 1*. Univ. California Press, Berkeley, pp. 361–380.

T. KUBOKAWA (2005). *A revisit to estimation of the precision matrix of the wishart distribution*. Journal of Statistical Research, 39, pp. 91–114.

T. KUBOKAWA, M. SRIVASTAVA (1999). *Robust improvement in estimation of a covariance matrix in an elliptically contoured distribution*. Annals of Statistics, 27, pp. 600–609.

W. NEWEY, D. MCFADDEN (1994). *Large sample estimation and hypothesis testing*. In R. ENGLE, D. MCFADDEN (eds.), *Handbook of Econometrics, vol 4*, Elsevier Science, pp. 2111–2245.

J. OWEN, R. RABINOVITCH (1983). *On the class of elliptical distributions and their applications to the theory of portfolio choice*. The Journal of Finance, 38, pp. 745–752.

M. PAWLAK, W. SCHMID (2001). *On the distributional properties of garch processes*. Journal of Time Series Analysis, 22, pp. 339–352.

M. ROCKINGER, E. JONDEAU (2002). *Entropy densities with an application to autoregressive conditional skewness and kurtosis*. Journal of Econometrics, 106, pp. 119–142.

G. SAMORODNITSKY, M. S. TAQQU (1994). *Stable Non-Gaussian Random Processes, Stochastic Models with Infinite Variance*. Chapman and Hall: New York, London.

C. STEIN (1956). *Inadmissibility of the usual estimator for the mean of a multivariate normal distribution*. In *Proc. Third Berkeley Symp. Math. Statist. Prob. 1*. Univ. California Press, Berkeley, pp. 197–206.

C. STEIN (1977). *Unpublished notes on estimating the covariance matrix*.

G. ZHOU (1993). *Asset-pricing tests under alternative distributions*. The Journal of Finance, 48, pp. 1927–1942.

SUMMARY

Multivariate elliptically contoured autoregressive process

In this paper, we introduce a new class of elliptically contoured processes. The suggested process possesses both the generality of the conditional heteroscedastic autoregressive process and the elliptical symmetry of the elliptically contoured distributions. In the empirical study we find the link between the conditional time varying behavior of the covariance matrix of the returns and the time variability of the investor's coefficient of risk aversion. Moreover, it is shown that the non-diagonal elements of the dispersion matrix are slowly varying in time.

Keywords: multivariate autoregressive process; elliptically contoured distribution; Stein-Haff identity