Methods and Techniques for Multifractal Spectrum Estimation in Financial Time Series

Petr Jizba and Jan Korbel

Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague
Bréhová 7, 11519, Prague, Czech Republic
(E-mail: p.jizba@fjfi.cvut.cz, korbeja2@fjfi.cvut.cz)

Abstract. In this paper, we compare two key approaches used in time series analysis, namely the Multifractal Detrended Fluctuation Analysis and Multifractal Diffusion Entropy Analysis. The comparison is done from both the theoretical and numerical point of view. To put some flesh on bare bones, we illustrate our analysis by applying both methods to three model time series. As a fourth illustration we analyze empirical time series of daily returns of S&P500 stock index recorded over the 50 years period. We argue that while the Multifractal Detrended Analysis is computationally more efficient, the Multifractal Diffusion Entropy Analysis is conceptually cleaner. In addition, the latter allows a wider applicability in cases when time series have underlying distributions that are heavy tailed.

Keywords: Multifractal spectrum, Detrended Fluctuation Analysis, Rényi Entropy.

1 Introduction

Scaling properties belong among the most important quantifiers of complexity in many real systems, including chaotic dynamical systems, biological systems and financial markets. The presence of scaling usually points to a non-trivial cooperative behavior imprinted in temporal correlations. Techniques of fractal geometry can be then applied to reveal and analyze the potential scaling exponents. In practice, however, it is rather common that systems exhibit a multiple scaling. In such a case the methods of multifractal analysis can be conveniently employed. The concept of scaling has been used routinely in theory of critical phenomena (e.g., phase transitions) and in chaotical dynamical systems since a long time. Applications in stochastic processes and ensuing time series (including financial time series) are, however, of newer vintage. At present, there are a number of works addressing this line of research. These range from seminal papers of Hurst [1], through theory of fractal geometry [2] and fractional dynamics [3] to the multifractal calculus [4–6]. Particularly, in the theory of multifractals the key rôle is played by the notion of Rényi’s entropy and ensuing generalized dimensions [7]. Recently, also the concept of Rényi’s transfer entropy [8] has been invoked in this context.

Our focus in this paper will be the study of time series from the multifractal point of view. As a rule, the presence of multifractality signalizes that the time series exhibits a complex behavior with long-range time correlations manifested on different intrinsic time scales. When considering real financial data series such as empirical series from stock market, the multifractality points, e.g., to
the onset of crises, economical cycles and other non-linear phenomena that cannot be modeled with Wiener process [9]. In this paper we compare two key techniques, namely, Multifractal Detrended Fluctuation Analysis (MF-DFA) and Multifractal Diffusion Entropy Analysis (MF-DEA). In particular MF-DEA is discussed in terms Rényi’s entropy. Our analysis reveals that while the MF-DFA is computationally efficient, the MF-DEA is conceptually cleaner. We also emphasize the fact that the MF-DEA is better suited to discuss cases when time series have underlying heavy tailed distributions.

The paper is organized as follows. In Section 2 we briefly outline some fundamentals of the multifractal calculus. In Section 3 we define a MF-DFA method. The related salient theoretical issues are presented in Section 4. Section 5 is dedicated to the MF-DEA method. Comparison of both methods for monofractal series is done in Section 6. Finally, in Section 7 are all methods demonstrated on four selected time series.

2 Some fundamentals of multifractal analysis

One way to reveal a multiscaling structure in time series is to construct a multifractal spectrum, which catalogues the involved scaling exponents. In the following we will use the standard definition, presented, e.g., in Ref. [10]. To this end we consider a time series \( \{x_i\}_{i=1}^{N} \) where \( i \) denotes a discrete times of the evolution with some specific time lag \( \tau \). We examine the scaling of the probability in different regions. The probability of a region in the space is given as 

\[
p_i = \lim_{N \to \infty} \frac{N_i}{N},
\]

where \( N_i \) is number of points lying in the region and \( N \) is total number of points. We assume that probabilities scale with some scaling exponent (Lipschitz–Hölder exponent) \( \alpha \) as 

\[
p_i \propto l^\alpha.
\]

For different regions we can obtain generally different values of \( \alpha \). Our aim is to describe a distribution of \( \alpha \), which is assumed to be in a form 

\[
d\alpha \rho(\alpha) \sim l^{-\tau(q)},
\]

where \( \tau(q) \) is value for which the expression \( qa - f(\alpha) \) is extremal. Under assumption that \( f(\alpha) \) is differentiable the latter is equivalent to the condition that \( f'(\alpha) = q \). Eq. (3) then implies that \( \tau(q) \) is nothing but the Legendre transform of \( f(\alpha) \). The generalized dimension is defined as 

\[
D_q = \lim_{l \to 0} \frac{\ln Z(q, l)}{\ln l} = \frac{\tau(q)}{q - 1},
\]
and is a generalization of a support dimension \( (q = 0) \), fractal dimension \( (q = 1) \) and correlation dimension \( (q = 2) \), see Ref. [7]. The term \( S_q = \frac{1}{q-1} \ln Z(q,l) \) is known as the Rényi entropy and it represents a one-parametric generalization of Shannon’s entropy (to which it reduces for \( q \to 1 \)). For different \( q \)'s we accentuate in the sum (1) different values of involved probabilities and so we can zoom to different regions of the distribution.

Note that if we rewrite the partition function as

\[
Z(q,l) = \sum_i p_i p_i^{q-1} = \langle p^{q-1} \rangle,
\]

the generalized dimension simply equals the scaling exponent of the generalized average of the probability distribution, namely \( s \cdot \sqrt{\langle p^{q-1} \rangle} \sim l^{D_q} \).

An alternative way to quantify a multiscaling behavior in time series is to generalize the concept of the Hurst exponent. The Hurst exponent is defined as the scaling exponent of the average increments of length of a dynamical process, i.e. \( \langle |x(t + \tau) - x(t)| \rangle \sim \tau^H \). When investigating the \( q \)-th moment of these increments, one may define the scaling:

\[
\langle |x(t + \tau) - x(t)|^q \rangle \sim \tau^{\zeta(q)}.
\]

For monofractal processes with the Hurst exponent \( H \), the \( \zeta(q) \) scaling exponent typically equals to \( \zeta(q) = Hq \). For multifractals \( \zeta(q) \) is not simply linear in \( q \) which motivates the definition of the generalized Hurst exponent \( H(q) \) as \( H(q) = \zeta(q)/q \). In this case the following scaling relation holds

\[
\sqrt{\langle |x(t + \tau) - x(t)|^q \rangle} \sim \tau^{H(q)}.
\]

For \( q = 1 \) this reduces to the standard Hurst scaling with \( H(q = 1) = H \).

### 3 Multifractal Detrended Fluctuation Analysis

The classical way how to calculate the multifractal spectrum is MF-DFA [11, 12]. The algorithm is following: if we have a noise-like series, we transform it into the walk-like series by a cumulative sum \( Y(n) = \sum_{i=1}^n (x_i - \langle x \rangle) \). The subtraction of the mean value is not important for the method itself, but it is crucial for the relation to the classical multifractal analysis, as discussed further. The method begins with dividing the series into \( N_s \) parts with length \( s \). Local linear (quadratic, cubic,...) trends \( y_\nu \) are estimated. The cornerstone of this method is fluctuation function which represents a total deviation from the trend and for each element is calculated as

\[
F(\nu, s)^2 = \sum_{i=1}^s (Y(s(\nu - 1) + i) - y_\nu(i))^2.
\]

MF-DFA method, which is a generalization of the monofractal DFA [12], deals with generalized fluctuation function dependent on \( q \) and defined as

\[
F(q, s) = \left\{ \frac{1}{N_s} \sum_{\nu=1}^{N_s} [F(s, \nu)^2]^{q/2} \right\}^{1/q},
\]

(9)
which is nothing else than $q \sqrt{\langle F(s, \cdot)^q \rangle}$, where we average over all time intervals. The fluctuation function satisfies the scaling property $F(q, s) \propto s^{h(q)}$. Because $N_s = N/s$, the following sum scales as

$$\left\{ \sum_{\nu=1}^{N_s} |F(\nu, s)|^{q/2} \right\} \sim s^{qh(q)-1}. \quad (10)$$

In passing, we may note that the exponent $h(2)$ is the exponent obtained from standard monofractal Detrended Fluctuation Analysis, while the exponent $h(1)$ is related to the R/S analysis and in the case of monofractal scaling both should be equal to the Hurst exponent $H$, cf. Ref. [12].

If we consider only a stationary, normalized and positive series $x_i$, it is possible to omit detrending procedure and we obtain that

$$\sum_{\nu=1}^{N_s} F^2_{FA}(\nu, s) = \sum_{\nu=1}^{N_s} |Y(\nu s) - Y((\nu - 1)s)|^q \sim s^{qh(q)-1}. \quad (11)$$

In order to relate the method to multifractal analysis, we define a partition sum

$$Z_q(s) = \sum_{\nu=1}^{N_s} |p_\nu(s)| \sim s^{\tau_M(q)}, \quad (12)$$

where $\tau_M(q)$ is the scaling function related to multifractal spectrum via Legendre transform and $p_\nu(s) = \sum_{k=(\nu-1)s+1}^{\nu s} x_k = (Y(\nu s) - Y((\nu - 1)s))$, see Ref. [12]. Eventually, we get the relation between $\tau_M$ and $h$

$$\tau_M(q) = q h(q) - 1. \quad (13)$$

4 Problems of $\tau_M(q)$

We have seen that the relation for $\tau_M(q)$ is problematic as it cannot be compared to standard multifractal analysis for every series. Let us now assume that we have a stationary, normalized and positive series $x_k$. Thus, the $p_\nu(s)$ should have the form $p_\nu(s) = \sum_{k=(\nu-1)s+1}^{\nu s} x_k = (Y(\nu s) - Y((\nu - 1)s))$, [13]. So, $(x - \langle x \rangle)$ represents a measure that generates the probability density $|p_\nu|$. The problem is that this measure is not a proper measure, because it does not obey axioms of measure, because we can find such set that for $A \subseteq B$ is $\mu(A) > \mu(B)$ [13]. Therefore, it can happen that relations derived in the last section might be not completely right. Apart from that, $D_q = \frac{qh(q)-1}{q-1}$ which for monofractal series, where $h(q) = H$, is equal to $D_q = \frac{qH^{-1}}{q-1}$. This leads to a singularity in $q = 1$, unless $H = 1$.

One possible way how to overcome these problems is consider the scaling function in the following form [13]:

$$\tau(q) = D(q-1) - K(q), \quad (14)$$

where the function $K(q)$ is a cumulant generating function and is related to so called fractal co-dimension and equals to $H' q - \zeta(q)$, where $H'$ is a constant to
be determined [14]. The relation between $\zeta(q)$ and $h(q)$ from MF-DFA analysis is $\zeta(q) = qH(q) = q(h(q) - 1)$. Finally, from the relation $\tau(1) = 0$, it is possible to determine the $H'$, which is equal to $H' = h(1) - 1$. For time series, we know that $D_0 = -\tau(0)$ is equal to dimension of support, which is equal to one [13], so the formula (14) takes the form

$$\tau_G(q) = qh(q) - qH' - 1 = q(h(q) - h(1) + 1) - 1.$$  

(15)

5 Diffusion Entropy Analysis and MF-DEA

The problem of MF-DFA is that in case of multifractality that originates from power-law scaling of distributions is the application of the method inappropriate, simply because the variance (and higher moments) of the time series is infinite. Of course, the real time series are finite, and hence the empirical variance is finite, but because of its non-trivial dependence of time-lag time, it does not give right predictions. In such cases it is more suitable to analyze multifractality via the self-similarity scaling ansatz. If we assume that the probability distribution of a time series has the form (as is the case, e.g., for Gaussian noise or Lévy processes)

$$p(x,t) = \frac{1}{t^B} F \left( \frac{x}{t^{\delta}} \right),$$  

(16)

then it is possible to estimate $\delta$ via Shannon’s entropy [15]

$$S(t) = -\int dx p(x,t) \ln[p(x,t)],$$  

(17)

because in this case $S(t) = A + \delta \ln t$ ($A$ is a constant).

The monofractal Diffusion Entropy Algorithm is based on the fluctuation collection. Let us have a noise-like series $\xi_i$ and let us define $x_\tau(t) = \sum_{i=1}^{\tau} \xi_{i+t}$. Then we divide all values of $x_\tau$ into boxes of length $\epsilon$ and calculate the probability of each box as $p_i(\tau) = \frac{N_i(\tau)}{N}$, where $N_i(t)$ is number of $x_\tau(t)$ that have values in the $i$-th box. The ensuing Shannon’s entropy is of the form

$$S(\tau) = -\sum_i p_i(\tau) \ln[p_i(\tau)].$$  

(18)

The MF-DEA is a direct generalization of the previous monofractal version. If we instead of Shannon entropy use the whole class of Rényi entropies, we get a class of scaling exponents [16]

$$S_q(\tau) = B_q + H(q) \ln \tau.$$  

(19)

Here $B_q$ is $\tau$-independent constant and $H(q)$ is the generalized Hurst exponent.

6 Comparison of $\tau(q)$’s for monofractal series

So far, we have been dealing with two different scaling functions $\tau(q)$. The first one is provided via Mandelbrot’s scaling definition [2] as $\langle |X(t)|^q \rangle \propto t^{\tau_{MF}(q)+1}$. 


The second one results from the scaling of the empirical distribution and the best way to phrase it is terms of the generalized dimension as \( \tau(q) = D_q(q - 1) \) and ensuing Rényi entropy as \( \tau(q) \sim -\ln \langle P^{q-1} \rangle / \ln t \) and therefore \( \langle P^{q-1} \rangle \propto t^{-\tau_R(q)} \). Let us consider a self-similar monofractal process with probability density in the form \( p(x,t)dx = [F(x/t^H)/t^H]dx \) and calculate scaling functions in both cases

\[
\langle |X(t)|^q \rangle = \int dx |x|^q p(x,t) = \int dx |x|^q \frac{1}{t^H} F\left(\frac{x}{t^H}\right) = c(q) t^{qH}.
\]  

(20)

Since \( c(q) \) is \( t \)-independent, we have \( \tau_M(q) = qH - 1 \). The observant reader will recall that we have noticed this behavior already in Section 3. In the case of Rényi entropy we have

\[
\langle |P(t)|^{q-1} \rangle = \int dx p(x,t)^q = \int dx \frac{1}{t^{Hq}} \left[ F\left(\frac{x}{t^H}\right)\right]^q
\]

\[
= t^{H(1-q)} \int dy \left[ F(y)\right]^q = c'(q) t^{H(1-q)}.
\]

(21)

The \( t \)-independence \( c'(q) \) then implies that \( \tau_R(q) = H(q - 1) \). By comparing the scalings (20) and (21) we obtain the following relation

\[
\tau_M(q) = \frac{q}{q-1} \tau_R(q) - 1.
\]

(22)

7 Numerical comparison of multifractal techniques

To put some flesh on the bare bones we shall now compare the above methods on explicit examples. To this end we chose four sample series, three of an artificial kind and one corresponding to an empirical financial series. The first series corresponds to a Brownian motion, the second to a monofractal series of a Fractional Brownian motion (FBM) with the Hurst exponent \( H = 0.75 \), the third is an artificial multifractal process (MFP), generated from a binomial cascade (see, e.g., Ref. [17]). Finally, the last series represents the daily records of the Standard and Poor’s 500 (S&P500) financial index gathered over the period of 50 years. Log returns of these time series are plotted on Fig. 1.

On Fig. 2 we depict results of all techniques used. In case of Detrended Analysis Method (top left figure) we can observe that the spectrum of S&P500 index is similar to the Gaussian spectrum and the spectrum of MFP, which points to the fact that most of the time is the index in a “Gaussian” regime, and only in the times of crises, etc., the large fluctuations can be observed and therefore is the spectrum wider (especially in the left part). The fractional Brownian motion was purposely chosen so, that all scaling exponents are shifted to the larger values. We shall note that in the theoretical case we should observe in the case of Brownian motion and FBM a singular spectrum with a single point that corresponds to its Hurst exponent. Nevertheless, in practice we do not obtain a singular spectrum. On the top right figure is displayed a spectrum obtained from a generalized scaling function. In this case,
the spectrum serves to disclose deviations from monofractal behavior and we observe that the lines of Brownian motion and FBM practically coincide, while spectrum of S&P500 and foremost MFP reveal rich underlying multifractal behavior. The bottom left figure shows the scaling exponents of Rényi’s entropy, where we can again observe the shift of FBM, the multifractality of the MFP and the fact that S&P500 reveal a multifractal behavior, but not as strong as the artificial series. Finally, the bottom right picture shows the estimated generalized Hurst exponent. Unfortunately, for negative values of $q$ the method is unreliable due to well known instabilities in Rényi’s entropy [7]. For positive values we get a steady behavior for all discussed series, but in the case of S&P500, we see a decreasing function even for positive values. The latter signalizes that though that the series does not have a wide spectrum as, e.g., MFP, the spectrum is rounder than for other ones and scaling exponents play an important rôle.

All in all, by comparing all above methods, the most reliable method (at least from the theoretical point of view) is the Multifractal Entropy Analysis. This is because of its clear interpretation of the scaling exponent, and its stability and validity also for processes with infinite higher moments. The disadvantage of the method is the computational robustness of the fluctuation collection that is of order $O(n^2)$. On the other hand, Multifractal Detrended Fluctuation Analysis is computationally very efficient and apart from theoretical issues related to the interpretation of the scaling exponents, it provides a clear numerical tool allowing to classify the multifractality.
8 Conclusions

In this paper we have compared two key methods used in time series analysis; the Multifractal Detrended Fluctuation Analysis and Multifractal Diffusion Entropy Analysis. We have demonstrated that MF-DEA based on Rényi entropy represents a pertinent tool that allows to conveniently quantify and qualify complex structures present in numerous realistic time series, and more specifically in financial time series. We have shown under what conditions both outlined methods can successfully estimate multifractal exponents, multifractal spectrum as well as scaling functions, and when the corresponding results can be naturally translated to each other. Added advantage of the MF-DEA steams from the fact that one may rigorously discuss also systems with infinite higher moments.

By using the example of S&P500 index, we have seen that corresponding empirical time series posses interactions that are highly nonlinear, and long-ranged. This is a clear manifestation of a number of interlocked driving dynamics operating at different time scales each with its own scaling function. In an econophysic such a behavior typically points to the presence of recurrent
economic cycles, crises, large fluctuations (i.e., marginal events such as spikes or sudden jumps), and other non-linear phenomena that are out of reach of more conventional multivariate methods [9].

Acknowledgements

Authors want to thank for the financial support provided by the Grant Agency of the Czech Republic, grant No. GCP402/12/J077, and the Grant Agency of the CTU in Prague, grant No. SGS13/217/OHK4/3T/14.

References

Spatial Variation in the Relationship between Mortality Rates and Neighborhood Characteristics in South Korea

Yoohyung Joo¹ and Hee Yeon Lee²

¹ Department of Environmental Planning, Graduate School of Environmental Studies, Seoul National University, 1 Gwanak-ro, Gwanak-gu, 151-742, Seoul, Korea (E-mail: yojo78@gmail.com)
² Department of Environmental Planning, Graduate School of Environmental Studies, Seoul National University, 1 Gwanak-ro, Gwanak-gu, 151-742, Seoul, Korea (E-mail: leehyn@snu.ac.kr)

Abstract. This paper study examines spatial patterns of standardized mortality rates in South Korea and examines whether there are “spatially heterogeneous effects” of neighborhood context on mortality rates. To examine the spatial variation of neighborhood impact, this research uses geographically weighted regressions (GWR). The dependent variable is standardized mortality rates and the explanatory variables are neighborhood characteristics. They are categorized into four sectors: economic condition, social capital, housing environment, and built environment. This study uses representative indicator of each sector. The study finds that the relationship between neighborhood environments and mortality rates is spatially heterogeneous across geographical areas. This indicates that the global regression model fails to account for the spatial variation in the association between mortality rates and neighborhood environments.

Introduction

There has been growing interest in the relationship between health and the neighborhood context. Neighborhood has been regarded as a predictor of health since it mediates the way society impinges on people, finally carrying weight on health disparity (Diez Roux, 2001). Economic conditions (Boardman et al., 2005) and social capital measured as the community participation in neighborhoods (Lochner et al., 2003) have been found to be determinants of health. While the social and economic conditions of neighborhoods are important, a growing body of research has also recognized the role of the physical urban environment in shaping health and thus mortality (Ewing et al., 2003; Frank et al., 2004). Residents of highly dense or mixed-use neighborhoods walk more frequently than of less pedestrian-friendly locations (Frank & Pivo, 1995). Built environments which include intersection density, land-use mix, and proximity to the nearest public playground may be potential predictors of health, determining the degree of walking. (Frank et al., 2004; Giles-Corti et al., 2003). ‘Obesogenic’ environment is viewed as discouraging physical activities and inducing the intake of high-calorie food (Swinburn et al., 1999). In addition, the housing environment, such as the ventilation and temperature in home and air pollution in residential areas, can explain the health status of a neighborhood’s residents (Howden-Chapman, 2004).

Regression techniques have commonly been used to analyze the relationship...
between the mortality rate and neighborhood characteristics. Ordinary least square (OLS) regression assumes that the relationships between independent and dependent variables remain constant over space. However, when dealing with geographically aggregated data, the assumption of stationarity often becomes untenable. Regarding spatial data, it is inappropriate to assume a constant relationship over space. The relationship between health and neighborhood contexts can be spatially non-stationary (Holt & Lo, 2008). Since neighborhood implies diverse social and political process, the way neighborhood affects people’s health can vary across geographic areas. The magnitude or significance of the association of the relationship might geographically differ. Geographically weighted regression (GWR) then is a useful statistical model to explore spatial nonstationarity when the nature or significance of the relationships between variables differs depending on location (Fotheringham et al., 2002; Gilbert & Chakraborty, 2011). GWR measures the local relationship at each location using a spatial kernel, or a specific distance around a location. This method is technically more effective at analyzing spatial data because it weights the attributes of nearby observations within the kernel (Gilbert & Chakraborty, 2011). GWR produces a series of parameters for each observation.

There have been some studies in health using geographically weighted regression. (Congdon, 2011; Holt & Lo, 2008; Nakaya et al., 2005). They analyze the relationship between mortality and socio-economic characteristics of neighborhoods, considering the spatial variation of the relationship. However, few studies have examined simultaneously the multidimensional aspects of neighborhood characteristics, including economic conditions, the physical environment, and social capital, when analyzing the spatially varied effect of neighborhoods on health. This study aims to examine how the relationships between various neighborhood characteristics and mortality rates in South Korea vary over space. The study breaks down neighborhood context into the categories of economic condition, social capital, built environment, and housing environment. The specific objectives of this study are to 1) examine the spatial patterns of the mortality rate at the neighborhood level in South Korea; 2) analyze the important factors determining spatial variations in mortality rates; and 3) assess the spatial heterogeneity of the relationships between dependent and independent variables. Consideration of spatial heterogeneity could achieve different results and policy implications from those produced when using a global regression model.

Methods

Data and study area

This study examines the relations between neighborhood characteristics and the mortality level using municipal administrative units as indirect proxies for neighborhood feature. South Korea is composed of 230 municipal
administrative units. Out of these units, 150 units are classified as urban area and 80 units as rural area. Among them, 60 units belong to Seoul Mega City Region (SMCR), which has achieved higher economic development than other parts of the nation. The population of South Korea in 2010 was about 48 million and that of SMCR was about 24 million, indicating that half of South Korea’s total population lives in SMCR. Based on the great differences depending on geographical locations, it can be assumed that the effect of neighborhood context on health varies across regions. This analysis is based on three data sources; They are the 2010 census data from Korean Statistical Information Service (KOSIS), the 2010 municipal statistical yearbooks of 16 Special cities and provinces, and Korean Transportation Database of Korea Transport Institute.

**Dependent variable**

This study employs Standardized mortality rates (SMRs) as health outcome. It is the age-adjusted death rate per 1,000 people. This is the calculated mortality ratio with estimated death rate per age group for the total population of Korea in 2005. While crude mortality rates are highly affected by the age structure of population, SMR controls the influence of age structure. SMR is an appropriate variable to compare mortality rates among neighborhoods in South Korea.

**Independent variables**

The independent variables which imply the neighborhood contexts are selected on the basis of previous studies. Variables are divided into four categories; economic condition, social capital, housing environment, and built environment. The study selects representative independent variables in each category which are found significantly associated with health outcomes in previous research. Regarding economic conditions, income is mainly used as predictor of health status. However, income information is not available in each municipal level in Korea. This research employs local tax base as a representative variable of the economic condition of a neighborhood. In terms of social capital, community participation is used as a representative measure. Several studies have showed that social cohesion, social trust, and community participation improve health status (Kawachi et al., 1997; Kennedy et al., 1998). In this study, community participation is calculated as the percent of the population that is affiliated with political, religious, and social organizations in each neighborhood. Housing environment is epitomized as housing quality. Old and overcrowded housing, housing with poor ventilation or poor air quality are associated with higher mortality rates and lower levels of health status (Howden-Chapman, 2004). This study has used housing with poor quality as a representative indicator which increases mortality risk. It is calculated as the percent of houses built before 1979 and sized below 66 square meters. Regarding built environment, this study uses
intersection density as a representative variable to affect health. Intersection density is often employed as an indicator of walkable environment which was revealed to improve health status. This study selects the number of intersections with more than three legs per 1,000 square meters as a criterion of walkable environment. It represents street connectivity which captures the degree to which destinations can be reached in a direct pathway and predicts the relative ease of walking (Frank et al., 2007) The higher the intersection density per area, the smaller the block size, which indicates better chances to walk. This study hypothesizes that neighborhoods with a higher local tax base, more community participation, better housing quality, and higher intersection density are associated with lower SMR; however, the impact of each neighborhood characteristic on SMR varies over space.

**Geographically weighted regression**

OLS regression, which generates ‘global’ regression coefficient, assumes that the relationships between neighborhood characteristics and health outcomes are constant. It hides any spatial variation in the relationship between independent variables and dependent variables. In order to display spatial heterogeneity, GWR called a ‘local’ regression is useful method. Unlike OLS, GWR assumes that such relationships may vary over space, generating a set of local regression coefficients for each observation point in the study area. OLS regression model can be expressed as:

\[ y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon \]  

(1)

Where \( y \) is the dependent variable, \( x_1, x_2, \ldots, x_n \) are the independent variables, \( \beta_0, \beta_1, \ldots, \beta_n \) are the parameters, and \( \varepsilon \) is the error term. GWR model can be written as:

\[ y_i = \beta_0(u_i, v_i) + \sum_{k \neq i} \beta_k(u_k, v_k) x_{ik} + \varepsilon_i \]  

(2)

Where \((u_i, v_i)\) is the location of observation i point. Each observation is weighted according to its proximity to i point (Graif and Sampson, 2009). The weighting function is represented as below.

\[ w_{ij} = \left\{ \begin{array}{ll}
1 - \left( \frac{d_{ij}}{h} \right)^2 & \text{if } d_{ij} < h \\
0 & \text{otherwise}
\end{array} \right. \]

Where \( h \) is a bandwidth, the distance of the nearest neighbor from i, where \( d_{ij} \) is the Euclidean distance between sampled points. Weight assigned to each observation point i decreases with increased distance from point i. In the calibration process, locations closer to observation i are more strongly
weighted. This study uses adaptive bandwidth as weighting function. The adaptive function allows smaller bandwidth where sampled points are dense and larger bandwidth where sampled points are sparse. Within a bandwidth, a weighting process is calibrated. If the bandwidth is too large, the spatial variation will be low and the model will tend toward the global model (Cahill & Mulligan, 2007). If bandwidth is too small, the sampled data points for estimation will be less and cause model instability. The bandwidth is calibrated until the Akaike Information Criterion (AIC) becomes minimized, which provides better fitness. Since regression is calibrated locally, each administrative unit has each parameter and t-statistic.

**Result**

As shown in Figure 1, the SMRs (Standardized Mortality Rates) of Korea exhibited the distinct spatial pattern, with regions of low mortality concentrated in Seoul Mega City Region and regions of high mortality concentrated in northeastern and southwestern parts of the nation. Each of the neighborhoods’ characteristics show the various geographic distributions (Figure 2).

A conventional OLS regression model is employed to determine the relationship between SMRs and neighborhoods’ characteristics. The regression analysis results are summarized in Table 1. The model explains the 51% variance in SMRs. Parameter estimates show that the local tax base, the social participation, and the density of the intersections are significantly and positively associated with SMRs, while the proportion of poor-quality housing is significantly and negatively related to SMRs. This indicates lower SMR rates correspond to a higher local tax base, a higher level of social participation, a higher intersection density, and less poor-quality housing. The conventional model suggests the relative effect of each independent variable. The local tax base is the most effective association with SMR, followed by intersection density, social participation, and poor-quality housing.

This OLS regression model assumes that relationships between the explanatory and dependent variables are identical in all the geographical units in the study area. However, the Koenker statistical value is 9.29, which is statistically significant (p<0.1) and indicates that there is nonstationarity in the relationship among the dependent and independent variables (Table 1). The significant Koenker test implies that the OLS model is biased and not suitable for exploring spatial variations.
Although the OLS regression model provides an overview of the relationship between dependant variable and independent variables, GWR allows us to see how and where these relationships vary spatially. Comparison of OLS model with GWR model revealed the GWR model is a statistically significant improvement over OLS model. The AIC score for the GWR model is 2,274, which indicates a better fit than the OLS model for which the AIC score is 2305. Also, the R-square estimate for OLS estimation is 0.51, while that of the GWR analysis is 0.61, which shows the GWR process explain variance of SMRs better. However, local R-square estimates ranged from 0.18 in the southern part of Korea to 0.68 in the Capital region, indicating the high degree of spatial variation (Figure 2).

![Figure 2. Local R-squared estimates.](image)

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>GWR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coefficients</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>520.52</td>
<td>520.02</td>
</tr>
<tr>
<td>Local Tax base</td>
<td>-0.8</td>
<td>-0.36***</td>
</tr>
<tr>
<td>Intersection Density</td>
<td>-3.69</td>
<td>-3.11***</td>
</tr>
<tr>
<td>Community Participation</td>
<td>-4.88</td>
<td>-2.77***</td>
</tr>
<tr>
<td>Housing with Poor Quality</td>
<td>9.97</td>
<td>.22***</td>
</tr>
<tr>
<td>Local R Squared</td>
<td>0.524</td>
<td>0.65</td>
</tr>
<tr>
<td>Adjusted R Squared</td>
<td>0.516</td>
<td>0.61</td>
</tr>
<tr>
<td>AIC</td>
<td>2304.92</td>
<td>2273.87</td>
</tr>
<tr>
<td>Moran I Index of Standard Residuals</td>
<td>0.08***</td>
<td>0.03**</td>
</tr>
<tr>
<td>Koenker statistics</td>
<td>9.29*</td>
<td>Neighbors</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>0.01</td>
<td>Bandwidth methods</td>
</tr>
<tr>
<td></td>
<td>Kernel type</td>
<td>Adaptive</td>
</tr>
</tbody>
</table>

Table 1. Coefficient of OLS and GWR Model.

Results from Table 1 justify the hypothesis that significant non-stationary relationships between mortality and each independent variables. The parameter for the local tax base is -0.8 in the OLS estimation, while the GWR model shows that the coefficients vary from -1.63 to -0.36. The intersection density has a varied range of coefficients from -5.8 to -0.18. The coefficient values for the community participation range from -11.59 to 0.78 and poor-quality housing indicates both a positive and a negative association with SMRs. Its coefficients range from -5.1 to 24.33.

Maps of GWR coefficients and t-values for each coefficient allow the visualization of the spatial heterogeneity (Figure 3, Figure 4). From figure 4, the established relationship between mortality and other neighborhood
attributes are not necessarily significant everywhere in Korea. Regarding local tax bases, the association with SMRs is significantly stronger in mideastern and southwestern parts of the nation. On the other hand, the association was found to be weaker around Seoul mega region. This implies that, in economically affluent areas, the degree of the relationship between economic conditions and SMRs might be weaker than in economically deprived areas. This suggests that the SMRs tends to be more influenced by the wealth of the neighborhood in deprived areas than other areas. The relationship between intersection density and SMRs is stronger in northeastern and mid-northern parts with statistical significance. A greater association is observed where a low density of intersections is pervasive. However, except for Seoul region and a few geographical units in the southeastern part, the intersection density is quite low. It is difficult to define what makes the SMRs of this area more closely associated with intersection density.

Figure 3. Spatially varying coefficients
Community participation is more highly related to SMRs around Seoul mega region. In the eastern part of the nation, community participation is positively associated with SMRs, which contradicts the OLS result. However, this relationship around these areas is not statistically significant (Figure 4). Poor-quality housing has been found to be more highly associated with SMR around the periphery of Seoul mega region and the northeastern part of South Korea, with a statistical significance ($p<=0.05$). Some areas show a negative relation between poor-quality housing and SMRs, but the relationship is not statistically significant.

**Conclusion**

This research analyzes the relationship between SMR and neighborhood contexts. OLS analysis indicates that higher local tax bases, intersection density, and community participation and less poor-quality housing are associated with a lower SMR. This analysis assumes that the neighborhood impact on health is constant across regions. However, the GWR analysis explores the spatially heterogeneous effect of neighborhood characteristics using the bandwidth of each sampled point. Nearby observations have a greater influence on the estimation of coefficients than observations further away.

In the GWR model, each neighborhood characteristic has a geographically varied range of parameter estimation. The influence of the local tax base on SMR is stronger in areas with smaller tax bases, suggesting that health status
is more vulnerable to economic conditions, particularly in deprived areas. Intersection density has a greater influence on SMR outside the SMCR and the northeastern region. The effect of community participation on SMR is stronger in the SMCR. The proportion of poor-quality housing has greater influence on SMR on the periphery of the SMCR and the northeastern region. The GWR model provides better fit and explanation for the variances of SMR than the OLS model. This study reveals where the mortality risk burdens are significantly related to neighborhood characteristics. In some areas, health status is more vulnerable to certain neighborhood characteristics, which underscores the need for place-specific policies.

However, this study has some limitations. It uses administrative units as indirect proxies for neighborhood features. The municipal unit differs from people’s perception of neighborhood, which makes it difficult to draw causal inferences about the effects of neighborhood on SMRs. Although the GWR analysis has improved the model, it still leaves spatially auto-correlated standard residuals. In addition, the study does not explain why certain areas show a greater association between SMR and neighborhood characteristics. The cause of the spatial heterogeneity of this relationship merits further study.

References


Queuing Systems with Two Service Operations as Mathematical Models of Reliability and Survivability

Revaz Kakubava¹, Tinatin Kaishauri², and Otar Shonia³

¹ Department of Applied Mathematics, Georgian Technical University, 77 Kostava str., 0175, Tbilisi, Georgia (E-mail: r.kakubava@gtu.ge)
² Management Information Systems, Georgian Technical University, 77 Kostava str., 0175, Tbilisi, Georgia (E-mail: kaishauri@gtu.ge)
³ Management Information Systems, Georgian Technical University, 77 Kostava str., 0175, Tbilisi, Georgia (E-mail: o.shonia@yahoo.com)

Abstract: The given paper deals with the redundancy and maintenance problem for a wide class of any territorially distributed standby systems consisting of unreliable repairable elements. Mathematical models for interaction of degradation and its compensation processes in the above mentioned systems are proposed and their possible applications are partially analysed. These models represent mixed type queuing systems for two parallel maintenance operations – replacements and repairs. The problem for optimization of said system by economic criterion is stated. The possible ways of its solution are discussed.

Keywords: queuing models; structural control; maintenance; replacement; renewal (repair).

1 Introduction

During last decades, in reliability theory and practice (as well as in survivability theory and practice), the problems of redundancy, maintainability and supply of large scale systems, including terrestrial ones, are becoming the main directions. This is strongly attested by leading experts, among them, distinguished scientist Igor Ushakov [1]. Other works of Igor Ushakov himself on the subject, as well as the works of other authors are referenced in [2]. Should also refer to [3, 4].

At the same time, traditional maintenance models in many cases proved to be unsuitable, and there was an urgent need for the construction and investigation of entirely new types of models for the mathematical description of the mentioned technical systems.

These models, as a rule, are distinctive symbiosis of reliability and queuing theories, along with inventory control and other parts of operations research (management science) [1-2].

One of the main reasons determining the above described statement is that in practical cases of redundancy, the main and standby elements, as a rule, were territorially concentrated and failed main element’s replacement with the standby one meant the latter’s switch, which was often automatically performed and the duration of operation was insignificantly small.

But in modern networks of above mentioned type, standby elements are not directly attached (linked) to main elements. They are located at specific storage
locations and may be tens, hundreds and sometimes thousands of kilometers away from the main elements. Therefore, the delivery duration of standby elements to the place of the failed main ones is quite substantial. At the same time, in practical cases, due to various reasons, before the beginning of standby element’s delivery operation, passes quite some time, which is often many times greater than the delivery time. Moreover, replacement operation, as a rule, is undertaken not by repair unit, but by special replacement channel. Therefore, replacement of failed main element by standby one quite naturally becomes an independent maintenance operation.

In addition, the replacement process, apart from the standby element’s delivery to the main element’s place, includes other sub-operations, whose execution is necessary in order for the standby element to continue main element’s functions. In such circumstances, the replacement operation’s average duration is not insignificant, and it often reaches 20–40% of repair operations’ average duration.

2 Object of study and its initial mathematical description

The investigation object of this article is a multi-element redundant system with repairable elements.

The system consists of identical m main and n standby elements. Standby elements are designated for permanent replacement of main elements in case of their failure. It is supposed that for the normal operation of the system, the serviceability of all m main elements is desired. However, if their number is less than m, then the system continues to function but with lower economic effectiveness.

The main elements fail with intensity $\alpha$ and the standby ones – with intensity $\beta$. A failed main element is replaced by a serviceable standby one if there is available standby element in the system. In the opposite case the replacement will be carried out after standby element’s availability. The failed elements, both the main and the standby ones, are repaired and become identical with the new ones.

There are $k$ replacement and $l$ repair units in the system. The durations of replacement and repair operations are random values with distribution functions $F(x)$ and $G(x)$, respectively. When maintenance units are busy, requests for replacement or repairs are queued. Service discipline is FCFS (first come, first served).

As we see, in a natural way we have a queuing systems with two types of maintenance operations – replacement and repair (renewal). We examine here the case, when $m$ is large number (in practice it might be tens, hundreds, thousands and more), and we will suppose that we have both infinite ($m=\infty$) and finite sources of requests and will get mixed type queuing systems.

In this system, the infinite source of requests for services is the set of main elements and finite source is the set of standby elements and service channels are replacement and repair units. At the same time, one event in a flow of
homogenous events – failure of the main element – generates requests for two parallel maintenance operations. First – the replacement of the failed main element with standby one, and the second – the repair of the very failed element.

Request for replacement occurs due to failure of the main element. The same event, coupled with standby element’s failure generates a request for repair (renewal).

To this day, both in reliability theory and queuing theory, the above problems have not been investigated in general case. At the same time, modern research methods of Markov and semi Markov processes allow us to construct and analyse such models in the framework of the mathematical theory of reliability and queuing theory [5, 6].

Only a few special cases of the described system have so far been investigated. 1) m=1, n=1; 2) m=1, n=2; 3) M/M/N – i.e. the repair time length has an exponential distribution, while the replacement time length equals zero (instant replacement); 4) some similar cases have also been investigated. One of such cases (m=1; n=1) is examined in [7].

In the last 7-8 years the specialists of Georgian Technical University (GTU) have succeeded in making considerable progress in the investigation along these lines. In particular, the models have been constructed and partly investigated for the following cases [8 10]:

1) m, n, k, l are arbitrary; the functions F(x) and G(x) are exponential;
2) m, n and the function F(x) are arbitrary; k = 1 = 1 and the function G(x) is exponential;
3) m, n and the function G(x) are arbitrary; k = 1 = 1 and function F(x) is exponential;
4) some similar statements have also been considered.

Now the investigation for other cases is under way. We call interested in it colleagues to join this work.

Conclusions

In modern large scale territorially dispersed networks standby elements (spare components) are not directly attached to main elements. They are located at specific storage locations and may be tens, hundreds and sometimes thousands of kilometers away from the main elements. Therefore, the duration of standby elements delivery to the place of the failed mains’ is quite substantial.

Also taking into account other circumstances, the replacement operation’s mean duration often reaches 20-40 % of repair operations’ mean duration and therefore, the replacement of the failed element with standby one is quite naturally becoming an independent maintenance operation.

Exactly the novel type queuing systems proposed in this work, in many cases, are the most adequate maintenance models for such networks.

References
Linear On/Off Inventory Control

Anna V. Kitaeva, and Natalya V. Stepanova

National Research Tomsk Polytechnic University, Tomsk, Russia
(E-mail: kit1157@yandex.ru)

Abstract. Single-product inventory management model with both random and controllable demand and continuous input product flow with fixed uncontrolled rate under finite storage capacity is considered. We consider the stock level process as asymptotically diffusion process and obtain its stationary distribution. The result permits us to solve the problem of minimizing the variance of the stock level process under linear on/off control and control the probabilities of the stock-out and overflow.

Keywords: On/Off Control, Stochastic Demand, Diffusion Approximation, Inventory Management.

1 The problem statement

A systematic study of inventory models incorporated uncertainly and dynamics began in the early 50s from the works by Arrow et al. [1] and Dvoretzky et al. [6]. Nowadays a set of stochastic models are available to solve the inventory control problem under various conditions encountered in practice, for example, see Ross [10], Chopra and Meindl [5], and Beyer et al. [3].

The aim of the paper is to stabilize the performance of the system under consideration. The feature of the system is exogenous (i.e., outside our control) input product flow. Let \( Q(t) \) be the stock level in the moment \( t \), the input product flow be continuous with fixed rate \( \nu_0 \), the demands be a Poisson process with constant intensity \( \lambda \), the values of purchases be i.i.d. random variables having a distribution \( F(\cdot) \) with finite the first and second moments equals respectively \( a_1 \) and \( a_2 \). The storage capacity let be bounded by \( Q_{\text{max}} \). Under certain conditions (for example, the threat of overflow) the product is delivered to outlets and the output flow is assumed to be continuous with a rate \( \nu^*(Q) \).

The aim is to stabilize the stationary process \( Q(\cdot) \) in the sense of minimum of its variance and to try to avoid the overflow and stock-out. We use the diffusion approximation of Markovian process \( Q(\cdot) \).

Diffusion methods have been applied in a variety of domains, see Janssen et al. [8]; as to application to inventory models, see, for example, Bather [3], Harrison [7], and Puterman [9].

The paper consists of two parts: the first part devotes to the approximation and in the second part we solve the optimization problem.
2 Diffusion approximation

Denote the density
\[
P(Q, t) = \frac{\Pr[Q \leq Q(t) \leq Q + dQ]}{dQ},
\]
and the rate of the product’s movement due to non-random factors
\[v_0 - v(Q) = \nu(Q)\).

Let \(P(Q, t)\) be a differentiable function of \(t\), \(\nu(Q)P(Q, t)\) be a differentiable function of \(Q\), and \(\int_0^\infty P(Q + u, t)du < \infty\).

Derive the Kolmogorov backward equation for a functional of Markov process \(Q(t)\)
\[
\varphi(Q, t) = E\{H(Q(\tau))\|Q(t) = Q\},
\]
and write the adjoint equation, which is the Kolmogorov forward equation for density function \(P(Q, t)\) as shown by Barucha-Reid [2].

Consider
\[
\varphi(Q, t - \Delta t) = E\{H(Q(\tau))\|Q(t - \Delta t) = Q\} =
\]
\[
= (1 - \lambda\Delta t) E\{H(Q(\tau))\|Q(t) = Q + \nu(Q)\Delta t\} +
\]
\[
+ \lambda\Delta t \int_0^\infty E\{H(Q(\tau))\|Q(t) = Q - u\}du + o(\Delta t) =
\]
\[
= (1 - \lambda\Delta t)\varphi(Q + \nu(Q)\Delta t, t) + \lambda\Delta t \int_0^\infty \varphi(Q - u, t)du + o(\Delta t) =
\]
\[
= \varphi(Q, t) + \nu(Q)\Delta t \frac{\partial \varphi(Q, t)}{\partial Q} - \lambda\Delta t \varphi(Q, t) + \lambda\Delta t \int_0^\infty \varphi(Q - u, t)du + o(\Delta t),
\]
which implies
\[
- \frac{\partial \varphi(Q, t)}{\partial t} = \nu(Q)\frac{\partial \varphi(Q, t)}{\partial Q} - \lambda\varphi(Q, t) + \lambda \int_0^Q \varphi(Q - u, t)du
\]
\[
= \nu(Q)P(Q, t) - \lambda\varphi(Q, t) + \lambda \int_0^Q P(Q + u, t)du.
\]

So the adjoint equation is
\[
\frac{\partial P(Q, t)}{\partial t} = - \frac{\partial \{\nu(Q)P(Q, t)\}}{\partial Q} - \lambda P(Q, t) + \lambda \int_0^Q P(Q + u, t)du. (1)
\]
To solve (1) suppose that the values of $Q(t)$ are large enough. The idea is to consider some infinitesimal parameter $\varepsilon$ so that the process $\varepsilon^2 Q(t)$ is not degenerate.

Denote

$$v_1(Q) = \sqrt{\varepsilon^2 Q}, \quad \varepsilon^2 = \tau, \quad Q\varepsilon^2 = x(\tau) + \varepsilon y, \quad P(Q, t) = \Pi(y, \tau, \varepsilon),$$

(2)

where $x(\cdot)$ is a differentiable function.

Let the limit $\lim_{\varepsilon \to 0} \Pi(y, \tau, \varepsilon) = \Pi(y, \tau)$ exists. Let $v_1(\cdot)$ be a differentiable function, $\Pi(y, \tau, \varepsilon)$ be a differentiable function with respect to $\tau$ and twice differentiable with respect to $y$.

By substituting (2) into (1) we obtain the equation

$$\varepsilon^2 \frac{\partial \Pi(y, \tau, \varepsilon)}{\partial \tau} - \varepsilon x'(\tau) \frac{\partial \Pi(y, \tau, \varepsilon)}{\partial y} = -\varepsilon \frac{\partial}{\partial y} \left[ v_1(x(\tau) + \varepsilon y) \Pi(y, \tau, \varepsilon) \right] -$$

$$- \lambda \Pi(y, \tau, \varepsilon) + \lambda \int_0^\infty \Pi(y + \varepsilon u, \tau, \varepsilon)dF(u).$$

(3)

Rewrite (3)

$$\varepsilon^2 \frac{\partial \Pi(y, \tau, \varepsilon)}{\partial \tau} - \varepsilon x'(\tau) \frac{\partial \Pi(y, \tau, \varepsilon)}{\partial y} =$$

$$= -\varepsilon \frac{\partial}{\partial y} \left[ \left[ v_1(x(\tau)) + \varepsilon y v_1'(x(\tau)) \right] \Pi(y, \tau, \varepsilon) \right] - \lambda \Pi(y, \tau, \varepsilon) +$$

$$+ \lambda \int_0^\infty \left[ \Pi(y + \varepsilon u, \tau, \varepsilon) + \varepsilon \frac{\partial \Pi(y, \tau, \varepsilon)}{\partial y} + \varepsilon^2 u^2 \frac{\partial^2 \Pi(y, \tau, \varepsilon)}{\partial y^2} \right] dF(u) + o(\varepsilon^2).$$

It follows

$$\varepsilon^2 \frac{\partial \Pi(y, \tau, \varepsilon)}{\partial \tau} = \varepsilon \left[ x'(\tau) - v_1(x(\tau)) + \lambda \alpha_1 \frac{\partial \Pi(y, \tau, \varepsilon)}{\partial y} \right] -$$

$$- \varepsilon^2 v_1'(x(\tau)) \frac{\partial \Pi(y, \tau, \varepsilon)}{\partial y} + \varepsilon^2 \frac{\lambda}{2} \frac{\partial^2 \Pi(y, \tau, \varepsilon)}{\partial y^2} + o(\varepsilon^2).$$

(4)

Let function $x(\cdot)$ be a solution of the equation
\[
\frac{dx(t)}{dt} = v_1(x(t)) - \lambda a_t. \tag{5}
\]

Then function $\Pi(y, t)$ satisfies the Fokker-Planck equation
\[
\frac{\partial \Pi(y, t)}{\partial t} = -v_1'(x(t)) \frac{\partial \{y \Pi(y, t)\}}{\partial y} + \frac{\lambda a_2}{2} \frac{\partial^2 \Pi(y, t)}{\partial y^2}. \tag{6}
\]

Consequently the process $y(\tau, \varepsilon) = \frac{\varepsilon^2 Q(t) - x(\tau)}{\varepsilon}$ converges in distribution to the Ornstein–Uhlenbeck process $y(\cdot)$ as $\varepsilon \to 0$ satisfying the following stochastic differential equation
\[
dy(\tau) = v_1'(x(\tau))y d\tau + \sqrt{\lambda a_2} dw(\tau). \tag{6}
\]

Let $v_1(\cdot)$ be a twice differentiable function. From (5) and (6) we get that the process
\[
z(\tau) = x(\tau) + \varepsilon y(\tau) \tag{7}
\]
satisfies
\[
dz(\tau) = (v_1(z) - \lambda a_1) d\tau + \varepsilon \sqrt{\lambda a_2} dw(\tau) + \frac{\varepsilon^2}{2} R_2 d\tau, \tag{8}
\]
where $R_2 = -y^2 v_1'(x(\cdot) \theta)$. \(0 \leq \theta \leq 1\).

Indeed it is clear
\[
dz(\tau) = dx(\tau) + \varepsilon dy(\tau) = (v_1(x(\tau)) + \varepsilon y v_1'(x(\tau)) - \lambda a_1) d\tau + \varepsilon \sqrt{\lambda a_2} dw(\tau). \tag{9}
\]

By Taylor expansion with Lagrange remainder we obtain
\[
v_1(z) = v_1(x + \varepsilon y) = v_1(x) + \varepsilon y v_1'(x) + \frac{\varepsilon^2}{2} y^2 v_1''(x + \theta\varepsilon y). \tag{9}
\]

By substituting $v_1(x) + \varepsilon y v_1'(x) = v_1(z) - \frac{\varepsilon^2}{2} y^2 v_1''(x + \theta\varepsilon y)$ into (9) we obtain (8).

We use (2) and (7) to get asymptotic equation
\[
\varepsilon^2 Q(t) = x(\tau) + \varepsilon y(\tau) = z(\tau). \tag{9}
\]

From (8) we get
\[ \epsilon^2 dQ(t) = \left( v_1(\epsilon^2 Q) - \lambda a_1 \right) dt + \epsilon \sqrt{\lambda a_2} dw(t) \frac{\epsilon^2}{2} R_2 d\tau, \]

which implies using \( v(Q) = v_1(\epsilon^2 Q) \)

\[ dQ(t) = (v(Q(t)) - \lambda a_1) \frac{dt}{\epsilon^2} + \sqrt{\lambda a_2} \frac{dw(t)}{\epsilon} - \frac{1}{2} R_2 d\tau. \]

Since \( \epsilon^2 \tau = \tau \), we have the equation

\[ dQ(t) = (v(Q(t)) - \lambda a_1) dt + \sqrt{\lambda a_2} dw(t) - \frac{1}{2} R_2 d\tau. \]

So approximately the equation holds

\[ dQ(t) = (v(Q) - \alpha_0 a_0) dt + \sqrt{\lambda a_2} dw(t). \]

Because of the boundedness of \( Q(\cdot) \) the stationary distribution exists

\[ p(s) = C \cdot \exp \left( \frac{2}{a_2} \int (v(s) - \alpha_0 a_0) ds \right), \quad (10) \]

where \( C \) is the normalization constant.

### 3 Linear on/off control

If the inventory level \( Q(\cdot) \) is above the some breakdown (base-stock) level \( Q_{\text{max}} - Q_0 \) we begin to deliver the product to outlets with a rate proportional to the difference \( Q - (Q_{\text{max}} - Q_0) \) to prevent the stock’s overflow. Thus

\[ v(Q) = \begin{cases} v_0, & \text{if } Q < Q_{\text{max}} - Q_0, \\ v_0 - \beta (Q - (Q_{\text{max}} - Q_0)), & \text{if } Q > Q_{\text{max}} - Q_0, \end{cases} \quad (11) \]

and \( v_0 > a_0 \lambda, \ \beta > 0 \).

The condition \( v_0 > a_0 \lambda \) means that if the inventory level is below the base-stock level, then the stock level is replenished in the mean, i.e., the warehouse accumulates goods.

We use (10) and (11) to get
\[ p(x) = C \exp \left( \frac{2}{a^2 \lambda} (v_0 - a^2 \lambda) \left( x - (Q_{\text{max}} - Q_0) \right) \right), \text{ if } x < Q_{\text{max}} - Q_0, \]

and

\[ p(x) = C \exp \left( \frac{2}{a^2 \lambda} \left( (v_0 - a^2 \lambda) \left( x - (Q_{\text{max}} - Q_0) \right) - \beta \left( \frac{x - (Q_{\text{max}} - Q_0)^2}{2} \right) \right) \right), \text{ if } x > Q_{\text{max}} - Q_0, \]

where

\[ C^{-1} = \frac{1 - 2b \Phi(b) \exp(b^2)}{2d}, \]

\[ d = \frac{v_0 - a^2 \lambda}{a^2 \lambda} > 0, \quad b = -d \sqrt{\frac{a^2 \lambda}{\beta}} < 0, \quad \Phi(b) = \int_b^{\infty} \exp(-t^2) dt. \]

The expectation of the inventory level process is

\[ E(Q) = \frac{1}{1 - 2b \Phi(b) \exp(b^2)} (Q_{\text{max}} - Q_0 - \frac{1}{2d} \left( Q_{\text{max}} - Q_0 - \frac{b^2}{d} \right) \Phi(b) \exp(b^2) - \frac{b}{2d}). \]

It’s variance is

\[ Var(Q) = \frac{1}{2d^2} \left( b^2 + \frac{1 + b^2}{1 - 2b \Phi(b) \exp(b^2)} \right) = \frac{g(b)}{2d^2}. \quad (12) \]

The value \( b \) giving the minimal value of the variance (12) is

\[ b_0 \approx -0.563, \quad g(b_0) \approx 0.734. \quad (13) \]

The probability of the overflow is

\[ \alpha = P(Q(t) > Q_{\text{max}}) = \frac{2b \Phi \left( \frac{b - d}{b} Q_0 \right) \exp(b^2)}{2b \Phi(b) \exp(b^2) - 1}. \quad (14) \]

The probability (14) takes the maximal value under \( Q_0 = 0 \).
\[ \alpha_{\text{max}} = 1 + \frac{1}{2h_0\Phi(h_0) \exp\left(b_0^2\right) - 1}. \]

We use (13) to compute \( \alpha_{\text{max}} \approx 0.68. \)

Minimal value of \( \alpha \) under optimal control is

\[ \alpha_{\min} = -2h_0^\gamma \Phi\left(b_0 - \frac{d}{h_0}Q_{\max}\right) \exp\left(h_0^2\gamma\right) \approx 0.48\Phi\left(-0.563 + \sqrt{\frac{b_0}{\alpha_{\max}}Q_{\max}}\right). \]

Storage capacity \( Q_{\alpha_{\max}}^{0\alpha} \) given desirable \( \alpha_{\min}^0 \) is

\[ Q_{\max}^{0\alpha} \approx \left(\Psi\left(2.08\alpha_{\min}^0\right) + 0.563\right)\sqrt{\frac{d_\gamma}{\beta_0}}, \tag{15} \]

where \( \Psi() \) is the inverse function of \( \Phi(). \)

The probability of the stock-out (the warehouse is empty) is

\[ \gamma = P(Q < 0) = \frac{\exp\left(-2d\left(Q_{\max} - Q_0\right)\right)}{1 - 2h_0\Phi(h_0) \exp\left(b_0^2\right)}. \]

The optimal base-stock level given \( \gamma_0 \) is

\[ Q_{\max} - Q_0 = \frac{2\ln h_0 - \ln \gamma_0}{2d}. \]

Maximal value of \( \gamma \) is

\[ \gamma_{\max} = \frac{1}{1 - 2h_0\Phi(h_0) \exp\left(h_0^2\right)} \approx 0.316 \]

under \( Q_0 = Q_{\max}. \)

Minimal value of \( \gamma \) is

\[ \gamma_{\min} = \frac{\exp\left(-2dQ_{\max}\right)}{1 - 2h_0\Phi(h_0) \exp\left(h_0^2\right)} \approx 0.316 \exp\left(-2dQ_{\max}\right). \]

The optimal stock capacity \( Q_{\text{opt}}^{0\gamma} \) given desirable \( \gamma_{\min}^0 \) is

\[ Q_{\max}^{0\gamma} \approx -\frac{\ln \left(3.16\gamma_{\min}^0\right)}{2d}. \tag{16} \]
Combining (15) and (16) we receive the storage capacity $Q^0_{\text{max}}$ making it possible to choose $\alpha \in [\alpha_{\text{min}}; 0.68]$ and $\gamma \in [\gamma_{\text{min}}; 0.316]$

$$Q^0_{\text{max}} = \max \left( \Psi \left( 2.08 \alpha_{\text{min}} \right) + 0.563 \right) \frac{\alpha_{\text{min}}^2}{\beta_0}, \frac{\ln \left( 3.16 \gamma_{\text{min}} \right)}{2d}.$$

4 Conclusions

Trade-off between the probabilities of overflow and stock-out under linear on/off control can be overcome only by increasing of stock’s capacity assuming that the input flow is beyond our power.

Note also that the overflow’s probability $\alpha_{\text{max}} = 0.68$ more than twice as big as the stock-out’s probability $\gamma_{\text{max}} = 0.316$ under optimal control. So presumably we need to consider more complicated, nonlinear models of controlled output flow.

References

Modelling structural changes in relations between returns of selected REIT indexes

Jozef Komorník¹, Magda Komorníková², and Cuong Nguyen³

¹ Faculty of Management, Comenius University, Odbojárov 10, P.O.BOX 95, 820 05 Bratislava, Slovakia
(E-mail: Jozef.Komornik@fm.uniba.sk)

² Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, 813 68 Bratislava, Slovakia
(E-mail: Magdalena.Komornikova@stuba.sk)

³ Faculty of Commerce, Lincoln University NZ, Canterbury, New Zealand
(E-mail: Cuong.Nguyen@lincoln.ac.nz)

Abstract. We have investigated the relations between 8 selected countries’ (USA, Canada, Japan, Australia, Hongkong, Singapore, UK and France) daily returns of the REIT (Real Estate Investment Trust) indexes in the time period January 3, 2000 – May 8, 2012, divided in 3 subperiods bounded by the recent global financial market crises (July 1, 2008 – April 30, 2009). We have observed that in the postcrisis subperiod the influence of the delayed values (by 1 day) of the returns of the US REIT index on the REIT indexes of Japan and Australia greatly increased, while the same effect did not take place for the remaining (even East Asian) REIT indexes. We used the copula approach for fitting the optimal models for the investigated relations.

Keywords: Real Estate Investment Trust (REIT), Returns of REIT indices, Copula, Archimedean copula, Reflection of copulas.

1 Introduction

The aim of this paper is to analyse the relations between the above mentioned 8 selected countries daily returns of the REIT (Real Estate Investment Trust) indexes in different time periods, determined by the recent global financial markets crises (July 1, 2008 – April 30, 2009). Our aim was to investigate the influence of the (expectedly) dominant US Real Estate (RE) market (where the crises has been initiated) on the remaining RE markets in the group, as well as its development in the different considered time subperiods. From this point of view, the most interesting change has been observed in the third (postcrisis) subperiod with considerably stronger relations of the delayed (by 1 day) returns of the US REIT index to the couple of returns of the Japan’s and Australia’s REIT indexes in comparison to the returns of the remaining 6 REIT indexes.

The paper is organized as follows. The second section contains the results of nonparametric correlation analyses (based on the Kendall coefficients) of the returns of the considered group of REIT indexes that have been filtered (in order to avoid a possible violation of the i.i.d. property) by ARMA–GARCH models (separately in the individual subperiods of time). The third section is devoted to a brief overview of the theory of copulas including the methodology of their fitting to two-dimensional time series. The fourth section contains
an overview of the best copula models for different time subperiods and selected significantly correlated pairs of returns of REIT indexes. Finally, some conclusions are presented.

2 The results of nonparametric correlation analyses of the returns of the REIT indexes filtered by ARMA-GARCH models

The results for different time subperiods are contained in the following Tables 1, 2 and 3. The considered countries are represented by the following consecutive number:

1. USA
2. Canada
3. Australia
4. Japan
5. Hongkong
6. Singapore
7. France
8. UK.

<table>
<thead>
<tr>
<th>countries</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>x</td>
<td>0.937</td>
<td>0.920</td>
<td>0.508</td>
<td>0.868</td>
<td>0.686</td>
<td>0.679</td>
<td>0.567</td>
</tr>
<tr>
<td>2</td>
<td>0.937</td>
<td>x</td>
<td>0.880</td>
<td>0.519</td>
<td>0.885</td>
<td>0.696</td>
<td>0.689</td>
<td>0.559</td>
</tr>
<tr>
<td>3</td>
<td>0.920</td>
<td>0.880</td>
<td>x</td>
<td>0.498</td>
<td>0.835</td>
<td>0.662</td>
<td>0.658</td>
<td>0.571</td>
</tr>
<tr>
<td>4</td>
<td>0.508</td>
<td>0.519</td>
<td>0.498</td>
<td>x</td>
<td>0.527</td>
<td>0.496</td>
<td>0.464</td>
<td>0.399</td>
</tr>
<tr>
<td>5</td>
<td>0.868</td>
<td>0.885</td>
<td>0.835</td>
<td>0.527</td>
<td>x</td>
<td>0.714</td>
<td>0.681</td>
<td>0.554</td>
</tr>
<tr>
<td>6</td>
<td>0.686</td>
<td>0.696</td>
<td>0.662</td>
<td>0.496</td>
<td>0.714</td>
<td>x</td>
<td>0.587</td>
<td>0.516</td>
</tr>
<tr>
<td>7</td>
<td>0.679</td>
<td>0.689</td>
<td>0.658</td>
<td>0.464</td>
<td>0.681</td>
<td>0.587</td>
<td>x</td>
<td>0.565</td>
</tr>
<tr>
<td>8</td>
<td>0.567</td>
<td>0.559</td>
<td>0.571</td>
<td>0.399</td>
<td>0.554</td>
<td>0.516</td>
<td>0.565</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 1. The values of the Kendall’s correlation coefficient $\tau$ for the first (precrisis) time period

We can observe that all values of the Kendall correlation coefficients are (quite surprisingly) high during the first two subperiods. They also remain quite high in the third period for the group comprising Canada (2), Hongkong (5), Singapore (6), France (7) and UK (8) as well as for the couple Australia (3) and Japan (4). However, their values for the remaining couple of countries are considerably diminished.

Realizing that there exists a time lag between the (presumably) most influential US RE market and another considered RE markets (except for Canada), we also calculated the values of the Kendall correlation coefficients for the delayed (by 1 day) values of the returns of the US REIT index with the other returns of REIT indexes that are presented in the Table 4. We see that their
values are very low in comparison with the corresponding values of the Kendall correlation coefficients of nondelayed values of the returns of the US REIT index with the returns of the another considered REIT indexes for the first 2 subperiods (while reaching considerably higher levels in the third subperiod only for the pairs with Australia (3) and Japan (4)).

Table 2. The values of the Kendall’s correlation coefficient \( \tau \) for the second (crisis) time period

<table>
<thead>
<tr>
<th>countries</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>x</td>
<td>0.377</td>
<td>0.301</td>
<td>0.267</td>
<td>0.313</td>
<td>0.328</td>
<td>0.350</td>
<td>0.306</td>
</tr>
<tr>
<td>2</td>
<td>0.377</td>
<td>x</td>
<td>0.692</td>
<td>0.549</td>
<td>0.868</td>
<td>0.731</td>
<td>0.515</td>
<td>0.391</td>
</tr>
<tr>
<td>3</td>
<td>0.301</td>
<td>0.692</td>
<td>x</td>
<td>0.535</td>
<td>0.667</td>
<td>0.716</td>
<td>0.486</td>
<td>0.397</td>
</tr>
<tr>
<td>4</td>
<td>0.267</td>
<td>0.549</td>
<td>0.535</td>
<td>x</td>
<td>0.586</td>
<td>0.556</td>
<td>0.488</td>
<td>0.378</td>
</tr>
<tr>
<td>5</td>
<td>0.313</td>
<td>0.868</td>
<td>0.667</td>
<td>0.586</td>
<td>x</td>
<td>0.707</td>
<td>0.503</td>
<td>0.365</td>
</tr>
<tr>
<td>6</td>
<td>0.328</td>
<td>0.731</td>
<td>0.716</td>
<td>0.556</td>
<td>0.707</td>
<td>x</td>
<td>0.496</td>
<td>0.419</td>
</tr>
<tr>
<td>7</td>
<td>0.350</td>
<td>0.515</td>
<td>0.486</td>
<td>0.488</td>
<td>0.503</td>
<td>0.496</td>
<td>x</td>
<td>0.626</td>
</tr>
<tr>
<td>8</td>
<td>0.306</td>
<td>0.391</td>
<td>0.397</td>
<td>0.378</td>
<td>0.365</td>
<td>0.419</td>
<td>0.626</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 3. The values of the Kendall’s correlation coefficient \( \tau \) for the third (postcrisis) time period

<table>
<thead>
<tr>
<th>countries</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>x</td>
<td>0.047</td>
<td>0.111</td>
<td>0.061</td>
<td>0.057</td>
<td>0.078</td>
<td>0.220</td>
<td>0.221</td>
</tr>
<tr>
<td>2</td>
<td>0.047</td>
<td>x</td>
<td>-0.005</td>
<td>-0.004</td>
<td>0.898</td>
<td>0.828</td>
<td>0.613</td>
<td>0.584</td>
</tr>
<tr>
<td>3</td>
<td>0.111</td>
<td>-0.005</td>
<td>x</td>
<td>0.222</td>
<td>0.013</td>
<td>0.052</td>
<td>0.059</td>
<td>0.087</td>
</tr>
<tr>
<td>4</td>
<td>0.061</td>
<td>-0.004</td>
<td>0.222</td>
<td>x</td>
<td>0.018</td>
<td>0.064</td>
<td>0.062</td>
<td>0.073</td>
</tr>
<tr>
<td>5</td>
<td>0.057</td>
<td>0.898</td>
<td>0.013</td>
<td>0.018</td>
<td>x</td>
<td>0.840</td>
<td>0.622</td>
<td>0.594</td>
</tr>
<tr>
<td>6</td>
<td>0.078</td>
<td>0.828</td>
<td>0.052</td>
<td>0.064</td>
<td>0.840</td>
<td>x</td>
<td>0.643</td>
<td>0.617</td>
</tr>
<tr>
<td>7</td>
<td>0.220</td>
<td>0.613</td>
<td>0.059</td>
<td>0.062</td>
<td>0.622</td>
<td>0.643</td>
<td>x</td>
<td>0.736</td>
</tr>
<tr>
<td>8</td>
<td>0.221</td>
<td>0.584</td>
<td>0.087</td>
<td>0.073</td>
<td>0.594</td>
<td>0.617</td>
<td>0.736</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 4. The values of the Kendall’s correlation coefficient \( \tau \) for the delayed values (by 1 day) of the returns of the US REIT index with those of the other considered countries in different subperiods of time

<table>
<thead>
<tr>
<th>03/01/2000–31/07/2008</th>
<th>01/08/2008–30/04/2009</th>
<th>01/05/2009–08/05/2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.011</td>
<td>0.004</td>
</tr>
<tr>
<td>3</td>
<td>0.009</td>
<td>0.091</td>
</tr>
<tr>
<td>4</td>
<td>0.009</td>
<td>0.191</td>
</tr>
<tr>
<td>5</td>
<td>0.008</td>
<td>0.023</td>
</tr>
<tr>
<td>6</td>
<td>0.009</td>
<td>0.065</td>
</tr>
<tr>
<td>7</td>
<td>0.003</td>
<td>0.154</td>
</tr>
<tr>
<td>8</td>
<td>-0.016</td>
<td>0.175</td>
</tr>
</tbody>
</table>
3 Theoretical concepts

Let \((X,Y)\) be a 2-dimensional random vector with a joint distribution \(F_{XY}\) and marginal distribution functions \(F_X, F_Y\). We will use the standard definition of a copula (see e.g., Joe[4], Nelsen[7]) \(C(u,v) : [0,1]^2 \to [0,1]\) satisfying

\[
F_{XY}(x,y) = C(F_X(x), F_Y(y))
\]  (1)

and corresponding density function (if \(C\) is absolutely continuous)

\[
c(u,v) = \frac{\partial^2}{\partial u \partial v} C(u,v).
\]  (2)

In our subsequent investigations, we mainly utilize four copula families of Archimedean class: Gumbel, strict Clayton, Frank and Joe BB1 (see e.g., Embrechts et al.[2], Joe[4], Nelsen[7]). Recall that their are given by the following expressions.

- **Gumbel family**
  \(C^G_\theta(u,v) = \exp\left(-((-\ln u)^\theta + (-\ln v)^\theta)\right)\frac{1}{\theta}\)  (3)
  where \(\theta > 1\). Let us note that \(C^G_1(u,v) = \Pi(u,v) = u \cdot v\).

- **Strict Clayton family (Kimeldorf and Sampson)**
  \(C^C_\theta(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}\)  (4)
  for \(\theta > 0\), \(C^C_0(u,v) = \Pi(u,v) = u \cdot v\).

To enrich the classes of considered models, we also consider the classes of Frank and Joe BB1 copulas.

- **Frank family**
  \(C^F_\theta(u,v) = -\frac{1}{\theta} \log \left(1 - \frac{(1 - e^{-\theta u})(1 - e^{-\theta v})}{(1 - e^{-\theta})}\right)\)  (5)
  for \(\theta > 0\), \(C^F_0(u,v) = \Pi(u,v) = u \cdot v\).

- **Joe’s family**
  \(C^J_\theta(u,v) = 1 - \left((1 - u)^\theta + (1 - v)^\theta - (1 - u)^\theta(1 - v)^\theta\right)^{1/\theta}\)  (6)
  for \(\theta \geq 1\).

A rich overview of Archimedean copulas is presented in Embrechts et al.[2], Genest and Favre[3], Joe[4] and Nelsen[7].

Let us recall that for a given copula \(C(u,v)\) the lower (left) and upper (right) tail dependence coefficients are defined by

\[
\lambda_L(C) = \lim_{\delta \to 0} Pr(F_Y(y) \leq \delta \mid F_X(x) \leq \delta) = \lim_{\delta \to 0} \frac{C(\delta, \delta)}{\delta} = \lim_{\delta \to 0} Pr(F_X(x) \leq \delta \mid F_Y(y) \leq \delta) \]  (7)
and
\[
\lambda_R(C) = \lim_{\delta \to 0} P_r(F_Y(y) \geq 1 - \delta \mid F_X(x) \geq 1 - \delta) = \lim_{\delta \to 0} \frac{2\delta - 1 + C(1 - \delta, 1 - \delta)}{\delta} = \\
= \lim_{\delta \to 0} P_r(F_X(x) \geq 1 - \delta \mid F_Y(y) \geq 1 - \delta). \tag{8}
\]

It is well known (see e.g. Joe[4] or Nelsen[7]) that the Gumbel copula \(C_G\), Clayton copula \(C_C\) and Joe copula \(C_J\) satisfy the relation
\[
\lambda_L(C_G) = 0, \quad \lambda_R(C_G) = 2 - 2^{\frac{1}{\theta}},
\]
\[
\lambda_L(C_C) = 2 - \frac{1}{\theta}, \quad \lambda_R(C_C) = 0,
\]
and
\[
\lambda_L(C_J) = 0, \quad \lambda_R(C_J) = 2 - 2^{\frac{1}{\theta}}.
\]

It is also well known (see Embrechts et al[2]) that the values of \(\lambda_R\) and \(\lambda_L\) for Frank copulas are equal to 0.

We follow the approach of Patton[10] and consider a so–called survival copula derived from a given copula \(C(u,v)\) corresponding to the couple \((X,Y)\) by
\[
SC(u,v) = u + v - 1 + C(1 - u, 1 - v) \tag{9}
\]
which is the copula related to the couple \((-X,-Y)\) with the marginal distribution functions
\[
F_{-X}(x) = 1 - F_X(-x^+) \quad \text{and} \quad F_{-Y}(y) = 1 - F_Y(-y^+). \tag{10}
\]

Obviously, if a copula \(C\) represents the right or left tail dependence, its survival copula \(SC\) represents the opposite one.

Convex combinations of copulas and corresponding survival copulas has been successfully applied for modelling of exchange rates dependences (e.g. in Patton[10] and Ning[8],[9]).

Applying reflections of copulas (left, right and composed), we can construct new copulas that exhibit interesting properties concerning additional coefficients of tail dependencies (see Komorník and Komorníková[5,6]).

It is well known that for the convex sums of copulas, the corresponding density function is the convex sum (with the same weights) of incoming density functions. The same kind of mixing property holds for the above mentioned coefficients of tail dependencies.

### 3.1 Fitting of copulas

In practical fitting of the data we have utilized the maximum pseudolikelihood method (MPL) of parameter estimation (with initial parameters estimates received by the minimalization of the mean square distance to the empirical copula \(C_n\) presented e.g. in Genest and Favre[3]). It requires that the copula \(C_\theta(u,v)\) is absolutely continuous with density \(c_\theta(u,v) = \frac{\partial^2}{\partial u \partial v} C_\theta(u,v)\). This
method (described e.g. in Genest and Favre[3]) involves maximizing a rank-based log-likelihood of the form

\[ L(\theta) = \sum_{i=1}^{n} \ln \left( c_\theta \left( \frac{R_i}{n+1}, \frac{S_i}{n+1} \right) \right) \]

where \( n \) is the sample size, \( R_i \) stands for the rank of \( X_i \) among \( X_1, \ldots, X_n \), \( S_i \) stands for the rank of \( Y_i \) among \( Y_1, \ldots, Y_n \) and \( \theta \) is vector of parameters in the model. Note that arguments \( \frac{R_i}{n+1}, \frac{S_i}{n+1} \) equal to the corresponding values of the empirical marginal distributional functions of random variables \( X \) and \( Y \).

4 Application to real data modelling by copula models

We considered models from Frank \((C_F^\theta)\), Joe \((C_J^\theta)\), strict Clayton \((C_C^\theta)\) and Gumbel \((C_G^\theta)\) families and their pairwise convex combinations (we will call them mixed 2 families copulas as well as the convex combinations with their survival copulas (we will call them mixed 1 family copulas).

For selecting the optimal models we applied the Kolmogorov – Smirnov Anderson–Darling (KSAD, for which we use the abbreviation AD) test statistic defined e.g. in Berg and Bakken[1] (for which we also constructed a GoF simulation based test), when comparing models with their submodels and different families of models.

Surprisingly, quite successful in this selection process (especially for the crisis period) have been symmetric Gumbel model of the type

\[ 0.5 \left( C_G^\theta + SC_G^\theta \right) \]

(that have equal values of the left and right tail dependence coefficients \( \lambda_L \) and \( \lambda_R \)).

The best fitted copulas for selected couples for the first subperiods are presented in Table 5 (with the returns of the nondelayed US REIT index) and Table 6 contains the results for the third subperiod (with the returns of the US REIT index delayed by 1 day).

4.1 Best models for the first period.

In this period, we identified only one symmetric Gumbel type model (for the couple Japan & Hongkong). Several optimal models have the form of mixed 1 family copulas. However, quite many optimal models for this period have the form of convex combinations of strict Clayton copulas with Gumbel copulas.

4.2 Best models for the second period.

For the crisis period we can observe mainly optimal models in the form of mixed Gumbel copulas, many of them symmetric. Exceptionally, for the couple of Canada & Hongkong, the optimal copula is of the Gumbel type.
dependence structure at the same time. Different structures among variables, including both right and/or left tail copulas. The flexibility of using mixed copulas is that it allows us to capture different structures among variables, including both right and/or left tail dependence structure at the same time.

4.3 Best models for the third period.

Here we can observe just one symmetric Gumbel optimal model (for the couple France & UK), several mixed 2 families models (for Joe BB1, Clayton and Gumbel families) as well as convex combinations of Clayton and Gumbel copulas.

5 Concluding remarks

This paper uses copula models to examine the dependence structure between returns of international RE markets in 3 different time periods (separated by the recent financial crisis). The paper employs Archimedean and mixed copula models of Clayton, Gumbel, Frank and Joe families as well as their survival copulas. The flexibility of using mixed copulas is that it allows us to capture different structures among variables, including both right and/or left tail dependence structure at the same time.

Table 5. The overview of optimal types of copulas for selected pairs of the (filtered) returns of REIT indexes for precrisis and crisis subperiods (with the returns of the nondelayed US REIT index)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 &amp; 2</td>
<td>α · C_{Ω1}^{α} + (1 − α) · SC_{Ω1}^{α}</td>
<td>α · C_{Ω1}^{α} + (1 − α) · SC_{Ω2}^{α}</td>
</tr>
<tr>
<td>1 &amp; 3</td>
<td>α · C_{Ω2}^{α} + (1 − α) · C_{Ω3}^{α}</td>
<td>0.5 (C_{Ω2}^{α} + SC_{Ω3}^{α})</td>
</tr>
<tr>
<td>1 &amp; 4</td>
<td>α · C_{Ω2}^{α} + (1 − α) · C_{Ω4}^{α}</td>
<td>0.5 (C_{Ω2}^{α} + SC_{Ω4}^{α})</td>
</tr>
<tr>
<td>1 &amp; 5</td>
<td>α · C_{Ω1}^{α} + (1 − α) · SC_{Ω4}^{α}</td>
<td>0.5 (C_{Ω1}^{α} + SC_{Ω4}^{α})</td>
</tr>
<tr>
<td>2 &amp; 3</td>
<td>α · C_{Ω2}^{α} + (1 − α) · SC_{Ω3}^{α}</td>
<td>0.5 (C_{Ω2}^{α} + SC_{Ω3}^{α})</td>
</tr>
<tr>
<td>2 &amp; 5</td>
<td>α · C_{Ω2}^{α} + (1 − α) · C_{Ω5}^{α}</td>
<td>C_{Ω2}^{α}</td>
</tr>
<tr>
<td>3 &amp; 5</td>
<td>α · C_{Ω1}^{α} + (1 − α) · SC_{Ω5}^{α}</td>
<td>α · C_{Ω1}^{α} + (1 − α) · SC_{Ω5}^{α}</td>
</tr>
<tr>
<td>4 &amp; 5</td>
<td>0.5 (C_{Ω2}^{α} + SC_{Ω5}^{α})</td>
<td>0.5 (C_{Ω2}^{α} + SC_{Ω5}^{α})</td>
</tr>
<tr>
<td>5 &amp; 6</td>
<td>α · C_{Ω1}^{α} + (1 − α) · C_{Ω6}^{α}</td>
<td>α · C_{Ω1}^{α} + (1 − α) · SC_{Ω6}^{α}</td>
</tr>
<tr>
<td>5 &amp; 7</td>
<td>α · C_{Ω1}^{α} + (1 − α) · C_{Ω7}^{α}</td>
<td>C_{Ω1}^{α}</td>
</tr>
<tr>
<td>5 &amp; 8</td>
<td>α · C_{Ω1}^{α} + (1 − α) · C_{Ω8}^{α}</td>
<td>C_{Ω1}^{α}</td>
</tr>
<tr>
<td>6 &amp; 7</td>
<td>α · C_{Ω6}^{α} + (1 − α) · C_{Ω7}^{α}</td>
<td>C_{Ω6}^{α}</td>
</tr>
<tr>
<td>6 &amp; 8</td>
<td>α · C_{Ω6}^{α} + (1 − α) · C_{Ω8}^{α}</td>
<td>C_{Ω6}^{α}</td>
</tr>
<tr>
<td>7 &amp; 8</td>
<td>0.5 (C_{Ω7}^{α} + SC_{Ω8}^{α})</td>
<td>0.5 (C_{Ω7}^{α} + SC_{Ω8}^{α})</td>
</tr>
</tbody>
</table>

Table 6. The overview of optimal types of copulas for selected pairs of the (filtered) returns of REIT indexes for postcrisis subperiod (with the returns of the US REIT index delayed by 1 day)

<table>
<thead>
<tr>
<th>Couple</th>
<th>01/05/2009–08/05/2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 &amp; 3</td>
<td>α · C_{Ω2}^{α} + (1 − α) · SC_{Ω3}^{α}</td>
</tr>
<tr>
<td>1 &amp; 4</td>
<td>α · C_{Ω2}^{α} + (1 − α) · SC_{Ω4}^{α}</td>
</tr>
<tr>
<td>5 &amp; 6</td>
<td>α · C_{Ω1}^{α} + (1 − α) · C_{Ω5}^{α}</td>
</tr>
<tr>
<td>5 &amp; 7</td>
<td>α · C_{Ω1}^{α} + (1 − α) · SC_{Ω6}^{α}</td>
</tr>
<tr>
<td>5 &amp; 8</td>
<td>α · C_{Ω1}^{α} + (1 − α) · SC_{Ω7}^{α}</td>
</tr>
<tr>
<td>6 &amp; 7</td>
<td>α · C_{Ω6}^{α} + (1 − α) · C_{Ω7}^{α}</td>
</tr>
<tr>
<td>6 &amp; 8</td>
<td>α · C_{Ω6}^{α} + (1 − α) · C_{Ω8}^{α}</td>
</tr>
<tr>
<td>7 &amp; 8</td>
<td>0.5 (C_{Ω7}^{α} + SC_{Ω8}^{α})</td>
</tr>
</tbody>
</table>
Our results prove that using copula, especially mixed copula, to capture dependency is a useful and flexible approach. There are several opportunities for future research, such as extending the number of REIT indices and/or examining the relationship between different investment in commodity markets, bond and stock markets, using the proposed method in this paper (possibly also with an extended range of copula families, for example asymmetric logistic model copula (ALM), and mixed asymmetric logistic model copula (MALM)).

Acknowledgement

This work was partially supported by Slovak Research and Development Agency under contract No. APVV-0073-10 and by VEGA 1/0143/11.

References

3. C. Genest and A.C. Favre. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering* 12, 347–368, 2007.
Reliability Evaluation of Multi-Camera Motion Detector by using Monte-Carlo Simulator

Samuel Kosolapov

Signal and Image Processing Laboratory,
ORT Braude Academic College of Engineering, Karmiel, Israel
E-mail: ksamuel@braude.ac.il

Abstract. Motion Detectors (MD) utilizing one digital camera, are well-known and widely used for detection of physical objects intrusion into protected zone. However, one-camera MD operation is limited to the protection of 2D region of fixed size, which significantly limit practical usage of MD of this kind. Stereo Motion Detectors (SMD utilizing two video cameras, can detect physical violation of the user specified 3D volume, but, as it was shown by earlier research, two-camera setup has low reliability for some motion paths, which lower total SMD reliability. In this research reliability of Multi-Camera Motion Detector (MCMD) was evaluated by using Monte-Carlo software simulator (implemented by using MAPLE script). Reliability of a number of practically interested setups was analyzed.

Keywords: Image Processing, 3D Imaging, Stereo Camera, Motion Detector, Monte-Carlo simulation, MAPLE

1 Introduction

Motion Detectors (MD) utilizing one digital camera, are well-known and widely used for detection of physical objects intrusion into protected zone [1]. For most MD any significant change in the content of the frame grabbed by digital camera is treated as “security violation event”. This means that MD operation is limited to the protection of 2D region of fixed size, which significantly limits practical usage of MD. Stereo Motion Detectors (SMD) utilizing two video cameras, can detect physical violation of the user specified 3D volume [2,3], but, as it was shown by earlier research, two-camera setup has low reliability for some motion paths, which lower total SMD reliability [4]. Constantly dropping prices on high-resolution digital cameras makes implementation of Multi-Camera Motion Detector (MCMD) practical, at least for the case of 3 or 4 digital cameras [5]. It seems obvious that increase in the number of cameras increases MCMD reliability, however, not every multi-camera setup is practically feasible because of camera(s) calibration need. Operation of SMD and MCMD in most cases requires some kind of calibration, which, in some cases, is problematic in the real-life conditions. Hence, it would be preferable to utilize setups that can be aligned during assembly, thus, effectively eliminating need for ”after-assembly” calibration. In order to evaluate accuracy and reliability of the selected “aligned” MCMD setup, software simulator was designed and implemented by using MAPLE.
## 2 SMD and MCMD Exemplary Setups

Simple exemplary scene selected for this research is shown on Fig. 1. “Gold Ring” is the object to be protected. No “alarm” must be raised when visitors are moving in the vicinity of the “Glass barrier”. “Alarm” must be raised only if the physical object (say, hand of the visitor) is moving through the plane of the “Glass Barrier”.

![Exemplary scene.

Fig. 1. Exemplary scene.](image)

![Exemplary two-camera (SMD) Configuration

Fig. 2. Exemplary two-camera (SMD) Configuration](image)
On the Fig. 2 exemplary two-camera (stereo) SMD setup is presented. This setup was analyzed in the previous work [3, 4]. Important that main optical axes of both cameras pass point “O” (origin) selected by operator during alignment step. This alignment can be easily achieved with adequate accuracy by using motorized cameras (as shown in Fig. 2).

Fig. 3 presents geometry of SMD setup in the XY plane (created by points (“A”, “B”, “O”)). From this figure relations between physical coordinates \([X,Y]\) of the exemplary point “T” and columns of the image of the point “T” on the sensors of the Left and Right cameras \([\text{ColR}, \text{ColL}]\) can be derived (Fig.4).

Camera parameters \([W, FL, FR, ps]\) are known from camera specifications. In the frames of “Alignment Instead of Calibration” approach, geometric distances of the setup (like AC, BC) are not measured, but assumed as known with some tolerance.

\[
\text{Eq1} = \frac{OC + Y}{AC + X} - \frac{AC}{AC} = \frac{\text{ColL} - \frac{1}{2} W}{FL} ps
\]

\[
\text{Eq2} := \frac{OC + Y}{BC - X} - \frac{BC}{BC} = \frac{\frac{1}{2} W - \text{ColR}}{FR} ps
\]

Fig.3. Two-camera Setup Geometry (XY Plane)

Fig.4. Equations \([X,Y] \leftrightarrow \{\text{ColL}, \text{ColR}\}\)
Geometry of the setup in the Z plane and relevant equations are trivial (standard lens equations) and thus not shown here.

By using equations presented on the Fig. 4 and equations for the Z plane, we can evaluate physical coordinates of the exemplary point “T” \( \{X,Y,Z\} \) by row and columns of the image of this point on the sensors of both cameras \( \{\text{Row}L, \text{Col}L\} \) and \( \{\text{Row}R, \text{Col}R\} \).

When point “T” is moving, \( \{\text{Row}L, \text{Col}L\} \) and \( \{\text{Row}R, \text{Col}R\} \) are changing, and thus, we can detect if motion in the user specified region was happen. However, from the analysis of Fig.2 and Fig. 3 can be seen that if point T is moving in the direction close to the direction optical axis of the (say) Left camera, changes of \( \{\text{Row}L, \text{Col}L\} \) and very small. In this case, 3D motion detection become non-reliable: actually, 3D two-camera setup operates as 2D one-camera setup.

In attempt to eliminate the problem of this directional sensitivity, three-camera MCMD setup (see Fig. 5) was tested. It is clear, that three cameras operates like three SMD: SMD#1 (Left and Right Cameras), SMD#2 (Left and Central Cameras), SMD#3 (Central and Right Cameras). In this situation, motion in the direction of the optical axis of (say) Left camera makes operation of the SMD#1 and SMD#2 non-reliable, but SMD#3 operates in the reliable way. So, in case one of tree SMD raises “alarm”, violation is considered as detected.

In order to validate this statement for the selected scene and for the selected setup in the quantitive manner, Monte-Carlo software simulation was executed.

---

![Fig. 5. Exemplary three-camera MCMD Configuration](image)
3 Monte-Carlo Simulation Procedure

In order to evaluate the feasibility of the selected “Alignment Instead of Calibration” approach, for the MCMD configuration in test, MAPLE-based software simulations were performed. On the first stage, developed simulator calculates series of digital images of the objects of the scene including “violating object” for all three cameras. By using calculated coordinates of the “violating object” in accordance with selected “violation path”, pseudo-video of the violation is generated for all three cameras as a sequence of digital images (frames) in accordance with equations presented on Fig. 4. The simplest one-camera motion detector detects motion of the “violating object” by processing images created as pixel-by-pixel differences of the consequent frames. In case no changes in the scene happened, difference image is “black”. Practically, camera noise is present. This noise is one of the factors leading to lower reliability of the camera-based motion detector [1].

On the next step, positions of “non-black” regions of two cameras images are used to evaluate [row-column] pairs and their correspondent [X, Y, Z] coordinates. In case any of [X, Y, Z] is inside the user defined protection zone, “alarm” must be raised. Simulator takes into account camera parameters and noise, effect of digitization and assembly errors of MCMD setup. Well-known mathematical models of different 3D setups cannot be used directly to evaluate accuracy and reliability of MCMD, because a number of parameters cannot be measured exactly for non-calibrated setup. Thus, classical Monte-Carlo approach was used to evaluate accuracy and reliability of the MCMD configuration in test. Simulator is operated a number of times, while, for every simulation run, values of the selected set of setup parameters are modified in a pseudo-random way. User can select parameters and assembly tolerances of the selected MCMD setup, geometry of 3D volume to be protected, intrusion object size and its motion path. As stated before, first unit of this simulator generates plurality of digital images in the situation when 3D object of specified size and shape is moving over specified 3D path. First unit utilizes a pseudo-random set of parameters – emulating assembly errors. “Restoration” is executed by operating the second simulator unit using “exact” setup parameters – as if assembly was ideal. This organization enables to estimate as accuracy as reliability of the “Alignment Instead of Calibration” approach. Simulation results in evaluation of the “true positive”, “true negative”, “false positive” and “false negative” of the selected setup. Additionally, results of simulations can be organized as a set of 2D and 3D plots enabling to recommend customer-tailored MCMD configuration for the specified volume to be protected.
4. Simulation Results

It is clear that absence of calibration leads to “Fast Positive Error”: protected volume was not violated, but “alarm” was raised. One of the goals of the Monte-Carlo simulator was to evaluate “False Positive Rate” (FPR). In the following example FPR was evaluated for the distances AC, BC, OC 5m with 2% tolerance (10 cm assembly error). Camera with 10 mm (10% tolerance) lens having VGA resolution was used. Error in point “O” alignment was set as 10 pixels; error of motion detector was set to 10 pixels. Simulator was operated 5000 runs. In case “dangerous zone margin” was set to 1cm, FPR value was 0.42 (that is in 2088 of 5000 cases motion detector generated false “alarm” – which is unacceptable). However, for the “dangerous zone margin” 10 cm, FPR was 0 (for all 5000 runs result was correct). For the specific example (see exemplary scene of Fig. 1) 10 cm accuracy can be considered as more than adequate.

Conclusions

Monte-Carlo simulator enables to evaluate accuracy and reliability of the three-camera MCMD, thus, eliminating the need for the tedious field tests. Accuracy and reliability of the three-camera MCMD in the selected exemplary configuration was estimated as adequate for the exemplary practical scene presented at Fig. 1. Usage of modern high-resolution cameras will additionally increase accuracy and reliability of MCMD.

Acknowledgment

This work was supported by a grant from research committee of ORT Braude Academic College of Engineering, Karmiel, Israel

References

BPSO Versions with Chi-squared Distribution for MKP Resolution

Raida KTARI¹², Habib CHABCHOUB¹

¹ Sfax University, L.O.G.I.Q, Tunisia
² Aix-Marseille Université, CNRS LIF UMR 7279, France
(E-mail: raidaktari@yahoo.com, habib.chabchoub@fsegs.rnu.tn)

Abstract. This paper deals with the application of BPSO (Binary Particle Swarm Optimization), EPSO (Essential Particle Swarm Optimization) and EPSOq (Essential Particle Swarm Optimization queen) to the Multidimensional Knapsack Problem (MKP) which is well-known to be NP-hard Combinatorial Optimization problem. The particularity of this paper consists in proposing a novel random number generator based on the Chi-squared distribution for the particles’ positions and velocities in the initialization step. A repair operator is also utilized to change an unfeasible solution to a feasible one. The performance assessment of the BPSO, EPSO and EPSOq is a critical point in this paper. That is why, we experiment these approaches on a variety of MKP instances from OR-Library. Then, we compare our results with those of other previous works and with the best known results in the literature.

Keywords: MKP, PSO, Chi-squared distribution, randomness, repair operator, performance.

1 Introduction

Particle Swarm Optimization (PSO) is one of the evolutionary optimization methods inspired by nature which include evolutionary strategy (ES), evolutionary programming (EP), genetic algorithm (GA), and genetic programming (GP). PSO was originally designed and introduced by Eberhart and Kennedy [7], [9] in 1995. It is a population based search algorithm based on the simulation of the social behaviour of birds, bees or a school of fishes.

The PSO algorithm is based on the exchange of information between individuals, so called particles, of the population, so called swarm. Indeed, each particle adjusts its own position towards its previous experience and towards the best previous position obtained in the swarm. Memorizing its best own position establishes the particle’s experience implying a local search along with global search emerging from the neighboring experience or the experience of the whole swarm.

PSO has gained widespread appeal amongst researchers and has been shown to offer good performance and efficiency in a variety of application domains such as power and voltage control (Abido [1]), mass-spring system (Brandstatter and Baumgartner [4]), and task assignment (Salman et al. [13]). The comprehensive survey of the PSO algorithms and applications can be found in Kennedy and Eberhart [11].

Particle swarm optimization was first introduced as an optimization method for solving continuous problem. Later, Kennedy and Eberhart [10], proposed a
binary version of PSO (BPSO) to accommodate discrete binary variables and allow it to operate in a binary problem space. A particle moves in a search space restricted to 0 or 1 on each dimension. Since the performance ability of BPSO is not good enough, several improved versions of BPSO have been proposed to modify the velocity vector when binary variables are involved. Recently, Chen et al. [5] proposed a modified version of BPSO, so called Essential Particle Swarm Optimization (EPSO). In this research work, authors dismantle the BPSO algorithm qualitatively, breaking it into its essential components and then reinterpreting it in another ways as new program. After, as pheromone array in Ant Colony Optimization (Alaya et al. [2]), they introduce the queen informant particle in EPSO. They identify this implementation of this idea with the acronym EPSOq that is considered as another improved version of BPSO.

This paper deals with the application of BPSO, EPSO and EPSOq in the field of Combinatorial Optimization (CO) problems, which is a quite rare field tackled by PSO. The constrained problem discussed in this paper is the well-known to be NP-hard CO problem. The 0-1 multidimensional knapsack problem (MKP), which consists in selecting a subset of n given objects (or items) in such a way that the total profit of the selected objects is maximized while a set of knapsack constraints are satisfied. More formally, the MKP01 can be stated as follows:

\[
\text{(MKP)} \begin{cases} 
\text{Max} & \sum_{j=1}^{n} c_j x_j \\
\text{s.s.} & \sum_{j=1}^{n} a_{ij} x_j \leq b_i ; \quad \forall i \in M = \{1, ..., m\} \\
& x_j \in \{0; 1\} \quad \forall j \in N = \{1, ..., n\}
\end{cases}
\]

Equation (1) describes the objective function for the MKP. Each of the m constraints described in condition (2) is called a knapsack constraint, so the MKP is also called the m-dimensional knapsack problem. Let $M= \{1, 2, ..., m\}$ and $N = \{1, 2, ..., n\}$, with $b_i > 0$ for all $i \in M$ and $a_{ij} \geq 0$ for all $i \in M, j \in N$, a well-stated MKP assumes that $c_j > 0$ and $a_{ij} \leq b_i \leq \sum_{j=1}^{n} a_{ij}$ for all $i \in M, j \in N$.

MKP is one of the most intensively studied discrete programming problems, mainly because its simple structure which can be seen as a general model for any kind of binary problems with positive coefficients.

In this paper, we propose an algorithm which applies the BPSO, EPSO and EPSOq approaches for solving the 0-1MKP. This is achieved by using a generator based on the Chi-squared distribution in order to generate particles well-varied in the research space during the initialization phase. This algorithm has also the advantage of using a repair operator based on the pseudo-utility ratios derived from the surrogate duality approach to guarantee generating feasible solutions.

The performance assessment of the BPSO, EPSO and EPSOq is a critical point in this paper. That is why, we experiment these approaches on a very large variety of bigger size MKP01 instances from OR-Library, which are considered
to be rather difficult for optimization approaches. Then, we compare our results with the best known ones in the literature (Angelelli et al. [3], Chu and Beasley [6], Vasquez and Vimont [14]).

2 BPSO Methods for MKP Resolution

Since we have found any literature concerning neither the EPSO nor EPSOq algorithm applied to the 0-1 MKP problems, we select some bigger size MKP instances from OR-Library to assess their performance, and we compare their results not only with that of the BPSO but also with the best known profits in the literature. To make a fair comparison, we set the same ring topology as the neighborhood structure with number of neighbors set to 2 for the three PSO improved versions. The parameters given in (Chen et al. [5]) are regarded as optimal for our algorithm. That is why; we conserve their values even for BPSO. We also use the same initialization generator for the particles position and the same reparation method for the unfeasible solutions. These methods are described as follows:

Initialization method: Monte Carlo methods use the computer together with the generation of random numbers and mathematical models to generate statistical results that be able to simulate and experiment with the behavior of various business, engineering and scientific systems. Monte Carlo simulations usually use the application of random numbers that are uniformly distributed over the interval [0, 1]. These uniformly distributed random numbers are employed in order to obtain stochastic variables from various probability distributions. These stochastic variables can then be useful to approximate the behavior of worthwhile system variables.

Indeed, most computer languages have form of random number generator that generates uniformly distributed random numbers between 0 and 1. Almost these random number generators use modulo-arithmetic so as to generate numbers that appear to be uniformly distributed. As a result, the random number generators are called pseudorandom number generators since they are not really random. They only simulate the behavior of a uniformly distributed random number on the interval [0, 1].

As an example, we apply in the present paper the chi-square goodness of fit test to the random number generator associated with JAVA our computer programming language in order to initialize the particles positions. We try to build a computer program to generate, for instance, 1000 random numbers between 0 and 1. We can then divided the interval 0 to 1 into 10 classes using the intervals (0, .1), (.1, .2), …, (.9, 1.0) and then we sort the 1000 random numbers to precise the number in each class. These values are the observed frequencies designed by the experiment. If the pseudorandom number generator is truly uniform, then the theoretical frequency associated with each class would have a value of 100.

So, for each position bit, a number will be generated thanks to this approach and consequently, it will be rounded either 0 or 1. This method would be very helpful in creating a swarm whose particles are well-scattered in the search
space and even quick not only for the particles positions initialization step but also for the particles velocities start up.

**Reparation method:** Usually, in PSO, a given particle is dynamically attracted by the social and the cognitive components. Therefore, during evolution towards the best global solution, the algorithm can pass through regions of unfeasible solutions.

Indeed, an unfeasible particle now can be changed to a feasible one later. A particle representing an unfeasible solution is allowed to exist in the swarm. PSO usually makes use of penalty function technique in order to reduce the constrained problem to an unconstrained problem by imposing a penalty to the fitness of such particle (Hu and Eberhart [8]).

Instead of this methodology, we incorporate the heuristic repair operator suggested in (Kong and Tian [12]) specially designed for MKP. The repair operator utilizes the notion of the pseudo-utility ratios $u_j$ that is explained in the following equation:

$$ u_j = \frac{e_j}{\sum_{i=1}^{m} w_i a_{ij}} $$  \hspace{1cm} (4)

where $w=(w_1, w_2, ..., w_m)$ is a set of surrogate multipliers (or weights) of some positive real numbers. To obtain reasonably good surrogate weights we can solve the LP relaxation of the original MKP and utilize the values of the dual variables as the weights. Otherwise, $w_i$ is set equal to the shadow price of the $i^{th}$ constraint in the LP relaxation of the MKP.

A brief description behind this reparation method is given as follows: The first phase, which is called DROP phase, changes the value of each bit of the solution from one to zero in increasing order of $u_j$ if feasibility is violated. The second phase, so called ADD phase, reverses the process by changing each bit from zero to one in decreasing order of $u_j$ as long as feasibility is not violated.

Adopting this procedure, we are able to solve the problem of an unfeasible particle existence in the search space.

### 3 Computational Results and Discussions

Combining all the ideas described above, we aggregate the BPSO, EPSO and EPSOq in a unique algorithm applied on a variety large MKP instances with maximum number of cycles of $n$, where $n$ is the objects number. Indeed, we concentrated in MKP instances with $m \in \{5, 10, 30\}$ constraints, $n \in \{100, 250, 500\}$ variables and $\alpha \in \{0.25, 0.5, 0.75\}$ tightness ratios. Experiments were performed in an Intel® Core ™ i5 with 4 Gb of RAM and 2.67 GHz CPU.

The following table reports just the test results of some instances.

<table>
<thead>
<tr>
<th>Instance Name</th>
<th>Best Known Solutions</th>
<th>BPSO Fitness Average</th>
<th>BPSO CPU Time Average</th>
<th>EPSO Fitness Average</th>
<th>EPSO CPU Time Average</th>
<th>EPSOq Fitness Average</th>
<th>EPSOq CPU Time Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1: Comparison of results obtained by BPSO, EPSO, EPSOq and the optimal known solutions

<table>
<thead>
<tr>
<th>Instances</th>
<th>Best-known</th>
<th>BPSO</th>
<th>EPSO</th>
<th>EPSOq</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CPU time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OB5x100.02.1</td>
<td>243561</td>
<td>21774</td>
<td>0.012</td>
<td>24514.3</td>
</tr>
<tr>
<td>OB5x100.03.1</td>
<td>45545</td>
<td>41805</td>
<td>0.15</td>
<td>41320</td>
</tr>
<tr>
<td>OB5x100.03.2</td>
<td>39802</td>
<td>44794</td>
<td>0.87</td>
<td>39494.5</td>
</tr>
<tr>
<td>OB5x100.03.2</td>
<td>61000</td>
<td>57550</td>
<td>0.09</td>
<td>61069</td>
</tr>
<tr>
<td>OB5x100.03.2</td>
<td>106485</td>
<td>106255</td>
<td>0.23</td>
<td>105734</td>
</tr>
<tr>
<td>OB5x100.03.2</td>
<td>155109</td>
<td>154639</td>
<td>0.015</td>
<td>154417.5</td>
</tr>
<tr>
<td>OB5x200.02.1</td>
<td>114181</td>
<td>109606</td>
<td>327.994</td>
<td>112182</td>
</tr>
<tr>
<td>OB5x200.03.1</td>
<td>216742</td>
<td>215622.5</td>
<td>222.85</td>
<td>215091.5</td>
</tr>
<tr>
<td>OB5x200.03.1</td>
<td>303199</td>
<td>298667</td>
<td>122.01</td>
<td>300267</td>
</tr>
</tbody>
</table>

This table and for all the 270 instances of MKP that we tested and which are considered to be rather difficult for optimization approaches, Our experiments show that BPSO, EPSO and EPSOq are able to find good solutions. But, it is obvious that EPSOq outperforms BPSO and even EPSO with better solution quality, and with quick convergence to satisfied solution, as the size of the problem increases.

In fact, incorporating the queen informant in the EPSO doesn’t increase the number of function evaluations because it was added just a new informer that only offers information to the other particles. Moreover, the initialization generator based on the Chi-squared distribution achieves a higher exploration of solutions at the start of algorithm and the repair operator plays a critical role in a higher exploitation near the global optimum solutions at the end of algorithm.

We also compare the EPSOq fitness with the best known solutions in the literature. Indeed, the average performance of EPSOq for the 270 instances considered was 1.028% of the known optimum. This reveals that PSO can perform well for this class of combinatorial problem, even for large instances. That is, EPSOq seems to be efficient in navigating the hyper-surface of the search space and finding good solutions (and, sometimes, the best solution). It is also as competitive as the existing discrete PSO approaches for the MKP thanks to the initialization generator, the repair operator and the queen informant.

4 Conclusions
This paper presents the application of the BPSO, EPSO and EPSOq on the 0-1 multidimensional knapsack problem. The particles positions are initialized using a random number generator based on the Chi-squared distribution in the aim to obtain a swarm well-varied. To tackle
the problem of the movement to regions at very large distances from the main
swarm or even outside the search space, we implement the reparation method
using the notion of the pseudo-utility ratios. So as to evaluate the effectiveness
and viability of these discrete binary PSO versions, we take many large MKP
instances from OR-Library to test its performance. Indeed, its simulation results
are compared with the best known solutions. The EPSOq results were close and
sometimes equal to the optimal solution known, even considering that the
parameters were not optimized. These experiments were not optimal. Little
effort was taken to find the best parameters. In future versions, we will do our
best in these aspects. Eventually, the EPSOq is seemed as a new path for
building a fast and easy discrete PSO by its smart representation.

References
2. I. Alaya, C. Solnon and K. Ghédira. Optimisation par colonies de fourmis pour le
problème du sac à dos multidimensionnel. TSI-Technique et Science Informatiques,
3. E. Angelelli, M. G. Speranza and M. W. P. Savelsbergh. Competitive analysis for
dynamic multiperiod uncapacitated routing problems. Networks, 49, 4, 308–317,
2007.
4. B. Brandstatter and U. Baumgartner. Particle Swarm Optimization—Mass-Spring
5. E. Chen, J. Li and X. Liuc. In search of the essential binary discrete particle swarm.
6. P. Chu and J. Beasley. A genetic algorithm for the multidimensional knapsack
7. R. C. Eberhart and J. Kennedy. A new optimizer using particle swarm theory,
Proceedings of the 6th International Symposium on Micro Machine and Human
8. X. Hu and R. Eberhart. Solving Constrained Nonlinear Optimization Problems with
Particle Swarm Optimization. Proceedings of the 6th World Multiconference on
10. J. Kennedy and R.C. Eberhart. A discrete binary version of the particle swarm
algorithm. IEEE International Conference on Systems, Man, and Cybernetics, 5,
12. M. Kong and P. Tian. Apply the Particle Swarm Optimization to the multidimensional
13. A. Salman, I. Ahmad and S. Al-Madani. Particle Swarm Optimization for Task
14. M. Vasquez and Y. Vimont. Improved results on the 0-1 multidimensional knapsack
Generalised Asymmetric Linnik distributions and process

Kuttykrishnan A.P., Associate Professor, Department of Statistics, Sir Syed College, Taliparamba, Kerala, India-670142.

Abstract

In recent years there has been an increasing interest in developing the theory and applications of geometric stable distributions. The class of geometric stable distributions is a four-parameter family of distributions denoted by $\text{GS}_\alpha(\sigma, \beta, \mu)$ and conveniently described in terms of characteristic function

$$\Phi(t) = \frac{1}{1 + \sigma^\alpha |t|^\alpha}$$

where

$$\varpi_{\alpha, \beta}(t) = \begin{cases} 1 - i\beta \text{sign}(t) \tan(\pi \alpha / 2) & \text{if } \alpha \neq 1 \\ 1 + i\beta \frac{2}{\pi} \text{sign}(t) \log|t| & \text{if } \alpha = 1. \end{cases}$$

The parameter $\alpha \in (0, 2]$ is the index of stability and determines the tail of the distribution. These classes of distributions arise as a limiting distribution of geometric random sums of independent and identically distributed random variables. Since the geometric random sums frequently appear in many applied problems in various areas, the geometric stable distributions have wide variety of applications especially in the field of reliability, biology, economics, financial mathematics etc. When $\beta = 0, \mu = 0$, the geometric stable distribution have the characteristic function $\Phi(t) = \frac{1}{1 + \sigma^\alpha |t|^\alpha}$, and the corresponding distribution is called Linnik distribution and is named after Ju.V. Linnik, who showed that the above function is a bona fide characteristic function of a symmetric distribution for any $0 < \alpha \leq 2$. It may be noted that probability density and distribution functions of the Linnik random variable are not in closed form except for $\alpha = 2$, which corresponds to the Laplace
distribution. The Laplace distribution is symmetric, and there were several asymmetric extensions in generalizing the Laplace distribution.

In this paper we introduce and study new classes of distributions, namely Pakes generalized asymmetric Linnik distribution and geometric Pakes generalized asymmetric Linnik distribution. First order autoregressive process with Geometric Pakes generalized asymmetric Linnik distribution as marginal distribution is developed. Higher order extensions are discussed. A bivariate distribution related to geometric Pakes asymmetric Laplace and Linnik distribution is introduced and bivariate time series model corresponding to this distribution is developed.

Key words: Autoregressive process, Geometric infinite divisibility, Geometric exponential distribution, Geometric marginal asymmetric Laplace and Linnik distribution, Geometric Pakes generalized asymmetric Linnik distribution, Geometric Stable distribution.

1. Introduction

In recent years there has been an increasing interest in developing the theory and applications of geometric stable distributions. The class of geometric stable distributions is a four-parameter family of distributions denoted by \( GS_\alpha (\sigma, \beta, \mu) \) and conveniently described in terms of characteristic function

\[
\Phi(t) = \frac{1}{1 + \sigma^\alpha |t|^{\alpha}} e^{\sigma |t|^{\alpha - 1} \omega_{\alpha, \beta}(t) - i\mu t}
\]

where \( \omega_{\alpha, \beta}(t) = \begin{cases} 
1 - i \beta \text{sign}(t) \tan(\pi \alpha / 2) & \text{if } \alpha \neq 1 \\
1 + i \beta \frac{2}{\pi} \text{sign}(t) \log |t| & \text{if } \alpha = 1.
\end{cases} \)

The parameter \( \alpha \in (0, 2] \) is the index of stability and determines the tail of the distribution. These classes of distributions arise as a limiting distribution of geometric random sums of independent and identically distributed random variables. Since the geometric random sums frequently appear in many applied problems in various areas (see Gnedenko and
Korolev (1996)), the geometric stable distributions have wide variety of applications especially in the field of reliability, biology, economics, financial mathematics etc.

When $\beta = 0, \mu = 0$, the geometric stable distribution have the characteristic function

$$\Phi(t) = \frac{1}{1 + \sigma^\alpha |t|^\alpha},$$

and the corresponding distribution is called Linnik distribution and is named after Ju.V.Linnik, who showed that the above function is a bona fide characteristic function of a symmetric distribution for any $0 < \alpha \leq 2$. The probability density function of the Linnik random variable when $\sigma = 1$ has the representation

$$f_\alpha(x) = \frac{\sin \pi \alpha}{\pi} \int_0^\infty \frac{v^{\alpha} \exp(-v|x|)}{1 + v^2 + 2v^\alpha \cos \frac{\pi \alpha}{2}} \, dv$$

for $x > 0$ and for $x < 0$,

$$f_\alpha(x) = f_\alpha(-x).$$

It may be noted that probability density and distribution functions of the Linnik random variable are not in closed form except for $\alpha = 2$, which corresponds to the Laplace distribution. The Laplace distribution is symmetric, and there were several asymmetric extensions in generalizing the Laplace distribution. Kozubowski and Podgórski (2000) studied asymmetric Laplace (AL) distribution with characteristic function

$$\Phi(t) = \frac{1}{1 + \sigma^2 t^2 - i \mu t},$$

$-\infty < \mu < \infty$, $\sigma \geq 0$, and discussed applications of it in the fields of financial mathematics. Kotz et al. (2001) discussed a new class of distributions, namely generalized asymmetric Laplace distributions, with characteristic function

$$\Phi(t) = \left( \frac{1}{1 + \sigma^2 t^2 - i \mu t} \right)^\tau,$$

$-\infty < \mu < \infty$, $\tau \geq 0$. By specifying $\tau = 1$, $\sigma = 0$ and $\mu > 0$, we have an exponential distribution with mean $\mu$ and obtain symmetric Laplace distribution if
\( \tau = 1, \mu = 0 \) and \( \sigma \neq 0 \). When \( \sigma = 0 \), the function reduced to characteristic function of a gamma variable with the scale parameter \( \mu \) and the shape parameter \( \tau \).

In this paper we introduce and study new classes of distributions, namely Pakes generalized asymmetric Linnik distribution and geometric Pakes generalized asymmetric Linnik distribution. In Section 2 we introduce Pakes generalized asymmetric Linnik distribution and obtain the representation of the random variable. In Section 3 we introduce and study geometric Pakes generalized asymmetric Linnik distribution. Time series model equivalent to TEAR (1) model discussed in Lawrance and Lewis (1981) with geometric Pakes generalized asymmetric Linnik distribution as marginal distribution is introduced and studied in Section 4.

2. Pakes generalized asymmetric Linnik distribution

Pakes (1998) generalized the Linnik distribution and introduced a symmetric distribution, namely generalized Linnik distribution with characteristic function

\[
\Phi(t) = \left( \frac{1}{1 + \sigma^{\alpha}|t|^{\alpha}} \right)^{\tau}, \quad \sigma, \tau \geq 0, 0 < \alpha \leq 2.
\]

It may be noted that when \( \alpha = 2 \) this reduces to the characteristic function of generalized Laplace distribution of Mathai (1993). Similar to generalized asymmetric Laplace distribution we can define an asymmetric distribution with characteristic function

\[
\Phi(t) = \left( \frac{1}{1 + \sigma^{\alpha}|t|^{\alpha} - i\mu t} \right)^{\tau}, \quad -\infty < \mu < \infty, \sigma, \tau \geq 0, 0 < \alpha \leq 2.
\]

We shall refer this distribution as the Pakes generalized asymmetric Linnik distribution and denoted by \( \text{PGAL}_{\alpha}(\mu, \sigma, \tau) \).
When $\alpha = 2, \tau = 1$, it reduces to the asymmetric Laplace distribution of Kozubowski and Podgórski (2000).

**Theorem 2.1** A $\text{PGAL}_\alpha (\mu, \sigma, \tau)$ random variable $X$ with characteristic function (2.1) admits the representation $X \sim \mu W + \sigma W^{1/\alpha} Z$, where $Z$ is symmetric stable with characteristic function $\Psi(t) = \exp(-\sigma^\alpha |t|^\alpha)$ and $W$ is a gamma random variable with probability density function $g(w) = \frac{1}{\Gamma(\tau)} w^{\tau-1} e^{-w}, w > 0, \tau > 0$, independent of $Z$.

**Proof:**

Conditioning on $W$, we obtain the characteristic function $\Phi(t)$ of $\mu W + \sigma W^{1/\alpha} Z$ as

$$
\Phi(t) = \mathbb{E}(\mathbb{E}(e^{it(\mu W + \sigma W^{1/\alpha} Z)} / W))
= \int_0^\infty \mathbb{E}(e^{it(\mu W + \sigma W^{1/\alpha} Z)}) g(w) dw
= \frac{1}{\Gamma(\tau)} \int_0^\infty w^{\tau-1} e^{-w(1+\sigma^\alpha |t|^\alpha - i\mu t)} dw
= \left(\frac{1}{1+\sigma^\alpha |t|^\alpha - i\mu t}\right)^\tau.
$$

Hence the theorem.

3. **Geometric Pakes generalized asymmetric Linnik distribution**

Pillai (1990) introduced geometric exponential distribution and studied the properties of the renewal process with geometric exponential waiting time distribution. Jose and Seetha Lekshmi (1999, 2003), studied geometric gamma and geometric Laplace distributions and developed autoregressive time series models. Jayakumar and Ajitha
(2003) introduced geometric Mittag-Leffler distribution and studied its properties including infinite divisibility and attraction to stable laws. Seetha Lekshmi and Jose (2006) introduced and studied geometric Pakes generalized Linnik distribution and developed time series model using this distribution.

Since the distribution with characteristic function (2.1) is infinitely divisible, using the result of Klebanov et al. (1984), we can define a geometrically infinitely divisible distribution with characteristic function \( \Psi(t) \) such that \( \Phi(t) = \exp \left\{ -\frac{1}{\Psi(t)} \right\} \).

The characteristic function (2.1) can be written as

\[
\left( \frac{1}{1 + \sigma^\alpha |t|^\alpha - i \mu t} \right)^\tau = \exp \left\{ -\frac{1}{(1 + \tau \log (1 + \sigma^\alpha |t|^\alpha - i \mu t))^{-1}} \right\}.
\]

Hence \( \Psi(t) = \frac{1}{1 + \tau \log (1 + \sigma^\alpha |t|^\alpha - i \mu t)} \) is a characteristic function of a geometrically infinitely divisible distribution.

A distribution with characteristic function

\[
\Psi(t) = \frac{1}{1 + \tau \log (1 + \sigma^\alpha |t|^\alpha - i \mu t)}, -\infty < \mu < \infty, \sigma, \tau \geq 0, 0 < \alpha \leq 2, \quad (3.1)
\]

is called geometric Pakes generalized asymmetric Linnik (GPGAL) distribution with parameters \( \mu, \sigma, \alpha \) and \( \tau \).

If \( X \) is a random variable with characteristic function (3.1), we represent it as \( X \overset{d}{\sim} \text{GPGAL}_\alpha (\mu, \sigma, \tau) \). It may be noted that when \( \tau = 1 \) in (3.1), the corresponding distribution is the geometric version of asymmetric Linnik distribution and in such case we call it as geometric asymmetric Linnik distribution (GPGAL_\alpha (\mu, \sigma, \tau = 1)).

Now we consider the asymptotic behavior of the GPGAL_\alpha (\mu, \sigma, \tau) distribution.
Theorem 3.1. The GPGAL$_{\alpha}(\mu, \sigma, \tau)$ distribution is the limit distribution of geometric sums of PGAL$_{\alpha}(\mu, \sigma, \tau)$ random variables.

Proof:

Let \( \Phi(t) \) be the characteristic function of PGAL$_{\alpha}(\mu, \sigma, \frac{\tau}{n})$ random variable. Then

\[
\Phi(t) = \left( \frac{1}{1 + \sigma^\alpha |t|^\alpha - i\mu t} \right)^{\frac{\tau}{n}}.
\]

Define \( \Theta(t) = \frac{1}{\Phi(t)} - 1 = (1 + \sigma^\alpha |t|^\alpha - i\mu t)^{\frac{\tau}{n}} - 1. \) Hence, using Lemma 3.2 of Pillai (1990), \( \Phi_n(t) = \frac{1}{1 + p \Theta(t)} \), where \( p > 1 \) is the characteristic function of a geometric sum of random variables. By choosing \( p = n \), we have

\[
\Phi_n(t) = \left(1 + n \left[ (1 + \sigma^\alpha |t|^\alpha - i\mu t)^{\frac{\tau}{n}} - 1 \right] \right)^{-1}.
\]

So \( \Phi_n(t) \) is the characteristic function of a geometric sum of PGAL$_{\alpha}(\mu, \sigma, \frac{\tau}{n})$ random variables.

Consider

\[
\lim_{n \to \infty} \Phi_n(t) = \frac{1}{1 + \lim_{n \to \infty} \left[ (1 + \sigma^\alpha |t|^\alpha - i\mu t)^{\frac{\tau}{n}} - 1 \right]}
\]

\[= \frac{1}{1 + \tau \log(1 + \sigma^\alpha |t|^\alpha - i\mu t)}, \]

which is the characteristic function of GPGAL$_{\alpha}(\mu, \sigma, \tau)$ random variables.

Hence GPGAL$_{\alpha}(\mu, \sigma, \tau)$ distribution is the limit distribution of geometric sum of PGAL$_{\alpha}(\mu, \sigma, \frac{\tau}{n})$ random variables.

Hence the theorem.
Now we prove stability property of $\text{GPGAL}_\alpha(\mu, \sigma, \tau)$ random variables with respect geometric summation.

**Theorem 3.2** Let $\{X_n\}$ be a sequence of independent and identically distributed random variables and let $N_p$ be a geometric random with mean $1/p$. Further, assume that $N_p$ is independent of the $X_i$'s. If $U_N = \sum_{i=1}^{N_p} X_i$ then the random variables $U_{N_p}$ and $X_i$ are identically distributed if $X_i$ follows $\text{GPGAL}_\alpha(\mu, \sigma, \tau)$ distribution.

**Proof:**

Let $\Phi(t)$ and $\Theta(t)$ be the characteristic functions of $X_i$ and $U_{N_p}$ respectively.

Then

$$
\Theta(t) = \frac{p \Phi(t)}{1 - (1-p) \Phi(t)}.
$$

(3.2)

Suppose $X_i \overset{d}{=} \text{GPGAL}_\alpha(\mu, \sigma, \tau)$, then by (3.2) we have

$$
\Theta(t) = \frac{p}{p + \tau \log(1 + \sigma^\alpha |t|^\alpha - i \mu t)}
$$

$$
= \frac{1}{1 + \frac{\tau}{p} \log(1 + \sigma^\alpha |t|^\alpha - i \mu t)}
$$

Hence

$$
U_{N_p} \overset{d}{=} \text{GPGAL}_\alpha(\mu, \sigma, \frac{\tau}{p})
$$

Hence the theorem.

4. **Autoregressive model with GPGAL$_\alpha(\mu, \sigma, \tau)$ marginal distribution**

Time series in which observations are of clearly non-Gaussian nature are very common in many areas. A number of literatures have developed in recent years in modeling time series data with non-Gaussian, and more generally asymmetric marginal distributions. Gaver and Lewis (1980) discussed and studied conventional first order linear autoregressive model $X_n = \rho X_{n-1} + \varepsilon_n$ with exponential marginal distribution.
Subsequently, Lawrance and Lewis (1981), Dewald and Lewis (1985), and Jayakumar and Pillai (1993) developed autoregressive models with different marginal distributions such as Laplace, gamma, Mittag-Leffler distributions. These first order autoregressive models are developed using the self-decomposability property of the corresponding marginal distributions. Now we develop a time series model using $GPGAL(\alpha, \mu, \sigma, \tau)$ marginal distribution on the basis of geometric infinitely divisible property of the distribution. This model is equivalent to the one-parameter TEAR (1) model discussed in Lawrance and Lewis (1981).

**Theorem 4.1**

Let $\{X_n, n \geq 1\}$ be defined as

\[
X_n = \begin{cases} 
\varepsilon_n & \text{w.p. } \theta \\
X_{n-1} + \varepsilon_n & \text{w.p. } 1 - \theta 
\end{cases} \quad (4.1)
\]

where $0 \leq \theta \leq 1$ and $\{\varepsilon_n\}$ is a sequence of independent and identically distributed random variables. A necessary and sufficient condition that $\{X_n\}$ is a stationary process with $GPGAL(\alpha, \mu, \sigma, \tau)$ marginal is that $\{\varepsilon_n\}$ is distributed as $GPGAL(\alpha, \mu, \sigma, \tau)$.

**Proof**

Let $\Phi_{X_n}(t)$ be the characteristic function of $\{X_n\}$. Then from (4.1), we get

\[
\Phi_{X_n}(t) = \theta \Phi_{\varepsilon_n}(t) + (1 - \theta) \Phi_{X_{n-1}}(t) \Phi_{\varepsilon_n}(t). \quad (4.2)
\]

Assuming stationarity, we have

\[
\Phi_X(t) = \theta \Phi_{\varepsilon}(t) + (1 - \theta) \Phi_X(t) \Phi_{\varepsilon}(t).
\]

Hence

\[
\Phi_{\varepsilon}(t) = \frac{\Phi_X(t)}{\theta + (1 - \theta) \Phi_X(t)}. \quad (4.3)
\]
Suppose $X_n \sim \text{GPGAL}_\alpha (\mu, \sigma, \tau)$ then $\Phi_{X_n}(t) = \frac{1}{1 + \tau \log(1 + \sigma^\alpha |t|^\alpha - i \mu t)}$.

Substituting this in (4.3) and simplifying, we get

$$\Phi_{\varepsilon_n}(t) = \frac{1}{1 + \theta \tau \log(1 + \sigma^\alpha |t|^\alpha - i \mu t)}.$$  

Hence $\varepsilon_n \sim \text{GPGAL}_\alpha (\mu, \sigma, 0 \tau)$.  

Conversely, if $\{\varepsilon_n\}$ is a sequence of independent and identically distributed $\text{GPGAL}_\alpha (\mu, \sigma, 0 \tau)$ random variables and $X_0 \sim \text{GPGAL}_\alpha (\mu, \sigma, \tau)$. 

Then from (4.2), when $n=1$, we have $\Phi_{X_1}(t) = \frac{1}{1 + \tau \log(1 + \sigma^\alpha |t|^\alpha - i \mu t)}$.

If $X_{n-1} \sim \text{GPGAL}_\alpha (\mu, \sigma, \tau)$ then we get $X_n \sim \text{GPGAL}_\alpha (\mu, \sigma, \tau)$.

Thus, using inductive argument $\{X_n\}$ is a stationary process with $\text{GPGAL}_\alpha (\mu, \sigma, \tau)$ marginal distribution.

Hence the theorem.

We call the process defined by (4.1) with $X_0 \sim \text{GPGAL}_\alpha (\mu, \sigma, \tau)$ and $\{\varepsilon_n\}$ is a sequence of independent and identically distributed $\text{GPGAL}_\alpha (\mu, \sigma, 0 \tau)$ random variables, as the first order autoregressive process with $\text{GPGAL}_\alpha (\mu, \sigma, \tau)$ marginal distribution.

From the definition (4.1) of the model it is easily verified that

$$\Phi_{X_n}(t) = \theta \Phi_{\varepsilon_n}(t) \frac{1 - (1 - \theta)^n \Phi_{\varepsilon_n}(t)}{1 - (1 - \theta) \Phi_{\varepsilon_n}(t) + (1 - \theta)^n \Phi_{X_0}(t) \Phi_{\varepsilon_n}(t)}.$$  

When $n \to \infty$, $\Phi_{X_n}(t) = \theta \Phi_{\varepsilon_n}(t) \frac{1}{1 - (1 - \theta) \Phi_{\varepsilon_n}(t)}$. 

570
Let $X_0$ is distributed arbitrarily and $\{e_n\}$ is a sequence of independent and identically distributed GPGAL $\alpha$ $(\mu, \sigma, \theta \tau)$ random variables then the characteristic function

$$
\Phi_{X_n}(t) = \frac{1}{1 + \tau \log(1 + \sigma^\alpha |t| - i \mu t)}.
$$

Hence if $X_0$ is distributed arbitrarily, then also the autoregressive process is asymptotically Markovian with GPGAL $\alpha$ $(\mu, \sigma, \tau)$ marginal distribution.

Now from the joint characteristic function of $(X_n, X_{n+1})$ of the process it is easily verified that

$$
\Phi_{(X_n, X_{n+1})}(t_1, t_2) \neq \Phi_{(X_n, X_{n+1})}(t_2, t_1).
$$

Hence the AR (1) process given by (4.1) with GPGAL $\alpha$ $(\mu, \sigma, \tau)$ marginal distribution is not time reversible.

5. Conclusion

Most of the data sets in the areas of financial mathematics, reliability, environmental studies etc, often do not follow the normal law but with asymmetric and heavy tailed character. A large number of research journals discussed applications of Laplace and asymmetric Laplace distributions in different fields where data exhibits asymmetric and heavy tailed character. In communication theory, frequently encountered impulsive noise possesses heavy tails, and so Laplace noise has been suggested as a best model. Also it is established that the Laplace distribution is considered as a model for the distribution of speech waves and the distribution is commonly encountered in image and speech compression. Empirical analysis of some important time series data, especially in the field of financial mathematics, environmental studies etc. shows that asymmetric and heavy tailed distributions, related to Laplace distribution, are more suitable for modeling the data. The applications of Laplace distribution in modeling sizes of sand particle, diamonds etc are also well established in many research articles (for more details see Kotz...
et al. (2001)). In this paper we examined the distributions related to Laplace and asymmetric Laplace distributions and can be used as an appropriate model in the areas where Laplace and Linnik distributions do not provide a better fit. Much will be learned in future by further analysis related to the applications of this new class of distributions across a range of contexts.

References


One dimensional embedding for nonnegative data visualization

Lazhar Labiod and Mohamed Nadif

Department of Computer sciences, LIPADE, University of Paris 5, 45 rue des Saints Pères 75006 Paris, France
(E-mail: l.labiod@yahoo.fr)
(E-mail: mohamed.nadif@parisdescartes.fr)

Abstract. Data visualization has attracted more attention this last decade as a powerful tool for a better understanding of data. In this paper we propose a new theoretical framework for data visualization, this framework is based on Rank-one SVD looking up an appropriate leading left and right singular vectors of a rows or/and columns normalized $m$ by $n$ adjacency matrix $A$. This involved computing of a truncated rank one singular value decomposition of a suitable normalized data matrix, constructing an one dimensional embedding for both rows and columns data. The visualization of $A$ consists in a simple permutation of rows and columns data according the sorted first left and right singular vectors, which involves an optimal data reorganization revealing homogeneous blocks. Finally, we link our approach to spectral co-clustering and show its usefulness in the context of co-clustering.

Keywords: Data visualization, Stochastic data, Power method, SVD, Co-clustering.

1 Introduction

Data visualization has attracted more attention this last decade as a powerful tool for a better understanding of the data. Prominent authors in the discipline of information visualization [1] have identified that the data mining community gives minimal attention to information visualization, but believe that there are hopeful signs that the narrow bridge between data mining and information visualization will be expanded in the coming years. Bertin [7] has described the visualization procedure as simplifying without destroying and was convinced that simplification was ‘no more than regrouping similar things’. Spath [3] considered such matrix permutation approaches to have a great advantage in contrast to the cluster algorithms, because no information of any kind is lost, and because the number of clusters does not have to be presumed, it is easily and naturally visible. Murtagh [4], Arabie and Hubert [6,5] have referred to similar advantages calling such an approach a non-destructive data analysis, emphasizing the essential property that no transformation or reduction of the data itself takes place.

In certain problems it may be useful to perform co-clustering, where both objects and features are assigned to groups simultaneously. One approach to the co-clustering problem is to view it as the task of partitioning a weighted bipartite graph. Dhillon [8] proposed a spectral approach to approximate the optimal normalised cut of a bipartite graph, which was applied for document clustering. This involved computing a truncated singular value decomposition
(SVD) of a suitably normalised term-document matrix, constructing an embedding of both terms and documents, and applying k-means to this embedding to produce a simultaneous k-way partitioning of both documents and terms. Finally, data visualization is obtained by reorganizing rows and columns data according to the co-clustering result. Despite the advantages of co-clustering, all methods require the knowledge of the number of blocks, in this paper we will not tackle co-clustering but we will see how an appropriate visualization of data involves a reorganization into homogeneous blocks, and we will show the usefulness of our approach in the context of co-clustering. We propose a new theoretical framework, specifically we develop an efficient iterative procedure to find a one dimensional embedding of both rows and columns data. This involved an optimal simultaneous rows and columns data reordering. We show that the solution is given by the leading left and right singular vectors of data matrix.

The rest of paper is organized as follows. Section 2 introduces the problem formulation and the aims. Section 3 is devoted to the proposed algorithm for data visualization. In Section 5, we discuss the relationship between our approach and spectral co-clustering. Section 5 presents numerical experiments on real and simulated data. Finally, the conclusion summarizes the advantages of our contribution.

2 Problem formulation

An interesting connection between data matrices and graph theory can be established. Let \( A \) be a \( m \times n \) data matrix, it can be viewed as a weighted bipartite graph \( G = (V, E) \), where \( V \) is the set of vertices and \( E \) is the set of edges, it is said to be bipartite if its vertices can be partitioned into two sets \( I \) and \( J \) such that every edge in \( E \) has exactly one end in \( I \) and the other in \( J: V = I \cup J \). The data matrix \( A \) can be viewed as a weighted bipartite graph where each node \( i \) in \( I \) corresponds to a row and each node \( j \) in \( J \) corresponds to a column. The edge between \( i \) and \( j \) has weight \( a_{ij} \), denoting the element of the matrix in the intersection between row \( i \) and columns \( j \), for convenience of discussion we also call the vertices in \( I \) as the documents (rows) while vertices in \( J \) as words (columns). The adjacency matrix of a bipartite graph is:

\[
B = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix},
\]

from which we define a stochastic data matrix as follow,

\[
S = D^{-1}B = \begin{bmatrix} 0 & D_c^{-1}A \\ D_r^{-1}A^T & 0 \end{bmatrix} \text{ where } D = \begin{bmatrix} D_r & 0 \\ 0 & D_c \end{bmatrix},
\]

where \( D_r \) and \( D_c \), are diagonal matrices such that \( D_r = \text{diag}(A1) \) and \( D_c = \text{diag}(A^T1) \). Let us focus our attention on the first leading eigenvectors of \( S \), since \( S \) is nonnegative and stochastic, The Perron-Frobenius theorem tell us that the first vector is also nonnegative and constant. The power method is
the well known technique used to compute the leading eigenvector of $S$. The power method consists in the following iterative process:

$$\pi^{(t)} = S^{(t)}\pi^{(0)} \text{ and } \pi^{(t)} = \frac{\pi^{(t)}}{||\pi^{(t)}||} \quad (3)$$

By Perron-Frobenius theorem [9] all eigenvalues are real and belong to $[-1, 1]$. Since $S$ is stochastic, it is known that for every right eigenvector there is a corresponding left eigenvector that corresponds to the same eigenvalue $\lambda_1 = 1$; the greatest eigenvalue called the Perron root. The right eigenvector corresponding to the uniform distribution $(\frac{1}{m+n}, ..., \frac{1}{m+n})^T$. The corresponding left eigenvector $\pi = 1$ represents the constant left eigenvector of $S$ so that $\pi^T1 = m + n$. In the matrix notation we have $\pi = S\pi$ and $S1 = 1$.

At first sight, this process might seem uninteresting since it eventually leads to a vector with all rows and columns coincide for any starting vector. However our practical experience shows that, first the vector $\pi$ very quickly collapses into rows and columns blocks and these blocks move towards each other relatively slowly. If we stop the power method iteration at this point, the algorithm would have a potential application for data reordering. The structure of $\pi^{(t)}$ during short-run stabilization makes the discovery of data ordering straightforward. The key is to look for values of $\pi^{(t)}$ that are approximately equal and reordering data accordingly.

3 Rank one SVD algorithm for data visualization

3.1 SVD algorithm

Singular value decomposition (SVD) is a widely-used multivariate data analysis technique [11] which has many potential applications in machine learning, document clustering, pattern recognition, and signal processing. Successful applications of SVD is due to its ability to learn parts-based representation through a matrix decomposition. Given a data matrix $A$ of size $m \times n$, SVD algorithms seek to find factors $U$, $\Lambda$ and $V$ such that

$$A = U_{m \times g}A_{g \times g}V_{n \times g}^T$$

where $U^TU = I_g$ and $V^TV = I_g$. \quad (4)

The steps of Jordan’s SVD are algorithm are described in [11].

3.2 Rank one SVD algorithm

Now, let us consider $\pi = \begin{bmatrix} u \\ v \end{bmatrix}$, where $u \in \mathbb{R}_+^m$ and $v \in \mathbb{R}_+^n$. The upper part of $\pi$, i.e $u$ is for documents weight and the lower part $v$ is for words weights. Exploiting now the diagonal structure of $S$, then we can write

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 & D^{-1}A \\ D^{-1}A^T & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \Leftrightarrow \begin{cases} u = D^{-1}Av \quad (a) \\ v = D^{-1}A^Tu \quad (b) \end{cases} \quad (5)$$
For numerical computation of the leading singular vectors, we use rank one SVD algorithm noted R1SVD, which is a variation of the power method adapted to rectangular data matrix. This iterative process starts with arbitrary vector $u^0$ and repeatedly performs the updates of $v$ and $u$ by alternating between formulas (a) and (b) given in equation 5 until convergence.

Algorithm 1: R1SVD

Input: data $A \in \mathbb{R}^{m \times n}$, $D_r$ and $D_c$

Output: $u$, $\lambda$, $v$

Initialize: $\tilde{u} = D_r^{-1}A1$, $u = \frac{\tilde{u}}{||\tilde{u}||}$

repeat
\[
\begin{align*}
\tilde{v}^{(t+1)} & = D_c^{-1}A^T u^{(t)} \\
v^{(t+1)} & = \frac{\tilde{v}^{(t+1)}}{||\tilde{v}^{(t+1)}||} \\
\tilde{u}^{(t+1)} & = D_r^{-1}A v^{(t)} \\
u^{(t+1)} & = \frac{\tilde{u}^{(t+1)}}{||\tilde{u}^{(t+1)}||} \\
\gamma^{(t+1)} & = ||u^{(t+1)} - u^{(t)}|| + ||v^{(t+1)} - v^{(t)}||
\end{align*}
\]

until stabilization of $u$, $v$, $|\gamma^{(t+1)} - \gamma^{(t)}| \leq \text{threshold}$

Documents and words weight are collected by $u$, $v$ respectively, the corresponding component value of $u$ and $v$ give document and word weights, respectively. We can sort word and document in decreasing (or increasing) order of theirs weights and reorganize data matrix accordingly to reveal the homogeneous blocks structure of $A$. We have developed a mutually reinforcing optimization procedure to exploit duality between documents set and words set, if a word is shared by many documents associated with a block, then word has a high weight associated with the block. On the other hand, if a document is shared by many words associated with a block, then the document has a high weight associated with the same block.

3.3 Relationship with spectral co-clustering

R1SVD is related to spectral co-clustering in that it finds a low dimensional embedding of data, and k-means or another clustering techniques is used to produce the final co-clustering. But as a result, in this paper it is not necessary to find any singular vector (as most co-clustering methods do), in order to find a low dimensional embedding for co-clustering, the embedding just needs to be a good linear combination of left singular vectors and of the right singular vectors respectively. In this respect R1SVD is very different approach from.

In spectral co-clustering the embedding is formed by the bottom left and right eigenvectors of a normalized data matrix [8]. In R1SVD, embedding is defined as weighted linear combination of singular vectors, then $u$ is a defined as linear combination of all left singular vectors of $A_r = D_r^{-1}A$ and $v$ as weighted linear combination of all right singular vectors of $A_c = AD_r^{-1}$. The left and right embedding turn out to be very interesting for data reordering and co-clustering.
From the start, the first largest left of $A_r$ and right singular vectors of $A_c$ are not very interesting since they move towards the uniform distribution via a long run times. However the intermediate $u$ and $v$ obtained by R1SVD after a short run time are very interesting. The experimental observation suggests that an effective reordering might run R1SVD for a small number of iterations.

Let us define the data matrix $M$ as follow

$$
M = \begin{pmatrix}
0 & A_r \\
A_c^T & 0
\end{pmatrix}
$$

Assuming that $M$ is diagonalizable i.e, that there exists a non singular matrix $Q$ of eigenvectors such that

$$
Q^{-1}M = \text{diag}(\lambda_1, \lambda, ..., \lambda_{n+m}).
$$

Furthermore, assuming that the eigenvalues are ordered $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq .... \geq |\lambda_{n+m}|$.

and expanding the initial approximation $\pi^{(0)} = \begin{bmatrix} u^{(0)} \\ v^{(0)} \end{bmatrix}$ in terms of the eigenvectors of $M$

$$
\pi^{(0)} = c_1 q_1 + c_2 q_2 + .... + c_{n+m} q_{n+m}
$$

with $q_k = \begin{bmatrix} y_k \\ z_k \end{bmatrix}$

where the upper part $y_k$ is for the rows of $A_r$ and the lower part $z_k$ is for the columns of $A_c$, $c_k \neq 0$ is assumed. We have

$$
\pi^{(t)} = M^{(t)} \pi^{(0)} = \lambda_1^{(t)} c_1 q_1 + \sum_{k=2}^{n+m} c_k (\lambda_k^{(t)}) q_k
$$

for $k = 2, ..., n + m$, we have $|\lambda_k| < |\lambda_1|$, so the second term tends to zero, and the power method converges to the eigenvector $q_1 = \begin{bmatrix} y_1 \\ z_1 \end{bmatrix}$ corresponding to the dominant eigenvalue $\lambda_1$. The rate of convergence is determined by the ratio $|\lambda_2|$, if this is close to one the convergence is very slow.

In equation (9), we expand the left eigenvector $\pi^{(t)}$ as a linear combination of the eigenvectors of $M^{(t)}$. It is easy to see that $\pi^{(t)} = \begin{bmatrix} u^{(t)} \\ v^{(t)} \end{bmatrix}$ is the leading eigenvector of $M^{(t)}$. Mathematically, in equation (10) we show that the leading eigenvector of $M^{(t)}$ is closely related to the first singular vectors of $\hat{A}^{(t)}$, $\pi_r^{(t)}$ and $\pi_c^{(t)}$ are the left and right singular vector of $\hat{A}^{(t)}$ respectively.

$$
\begin{bmatrix}
\begin{bmatrix} u^{(t)} \\ v^{(t)} \end{bmatrix} \\
\begin{bmatrix} u^{(0)} \\ v^{(0)} \end{bmatrix}
\end{bmatrix} = \begin{bmatrix}
\begin{bmatrix} A_r^{(t)} u^{(0)} \\ A_c^{T(t)} v^{(0)} \end{bmatrix}
\end{bmatrix}
$$

Instead of constructing $M$ (like the most spectral co-clustering methods do) which is bigger and sparser than the data matrix $A$, we provide a way to visualize and co-cluster data, not using $M$ but directly from $A$. 

579
Now, by exploiting the block structure of $M$ in equation (9), we show in equation (11); $u(t)$ as a linear combination of the left singular vectors of $A_r$ and $v(t)$ as a linear combination of the right singular vectors of $A_c$.

$$
\begin{bmatrix}
u(t) \\ v(t) \end{bmatrix} =
\begin{bmatrix}
\lambda(t)c_1y_1 + \sum_{i=2}^n c_i(\frac{\lambda_i}{\lambda_1})y_i \\
\lambda_1(c_1z_1 + \sum_{j=n+1}^m c_j(\frac{\lambda_j}{\lambda_1})z_j)
\end{bmatrix}
$$

(11)

4 Experiments

We now provide experimental results to illustrate the behavior of the R1SVD algorithm. We argue that R1SVD allows us to capture the trends of objects over a subset of attributes and then reorganizes data matrix into homogeneous blocks. We apply our algorithm on different real-world and word-document simulated data sets (using a Bernoulli latent block model [10]). Different patterns are considered to show the ability of the R1SVD algorithm to rediscover the hidden blocks in data without fixing any parameters on the rows and columns ordering.

Table 1. 16 Townships Data.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Agriculture Coop</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Rail station</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>One Room School</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Veterinary</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>No Doctor</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>No Water Supply</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Police Station</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Land Reallocation</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2. Reorganization of 16 Townships data after co-clustering.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Agriculture Coop</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rail station</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>One Room School</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Veterinary</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No Doctor</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No Water Supply</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

We now try to visualize the homogeneous block structure that might be discovered by our algorithm. Table 1 shows the original characteristics-townships data matrix and table 2 the reordered matrix obtained by arranging rows and columns based on the sorted $u$ and $v$. The table 2 reveals the hidden sparsity structure of both characteristics and townships clusters. The three diagonal blocks in this table correspond to the three clusters. It clearly appears that we can characterize each cluster of townships by a cluster of characteristics: \{H, K\} by \{High School, Railway Station, Police Station\}, \{B, C, D, G, L, O\} by \{Agriculture Coop, Veterinary, Land Reallocation\} and \{M, N, J, I, A, A, P, F, E\} by \{One Room School, No Doctor, No Water Supply\}.

Figure 1 shows in order from the left to right, the original characteristics-townships data matrix, the reordered data matrix, the reordered rows similarity matrix $S_r = AA^T$, the reordered columns similarity matrix $S_c = A^TA$ and the plots of sorted singular vectors $u$ and $v$. 

580
The three diagonal blocks in figures 2 and 3 correspond to the three clusters. The rough block diagonal structure indicates the cluster structure relation between documents and words. Hence by exploiting the duality of the data and features, incorporating the features information in data reordering at each stage, our algorithm yields better reordering solution that one dimensional clustering approaches, especially for high dimensional sparse datasets.

The R1SVD framework seems to have the potential to address the question of the number of clusters underlying the data, it detects the suitable number of blocks by analyzing the evolution of the first left $u$ of $A_r$ and right singular vectors $v$ of $A_c$. A performance study has been conducted to evaluate our method. In this subsection, we try to answer the question; is this reordering meaningful? In order to be able to answer this question we use confusion matrices to measure the clustering performance of the co-clustering result provided by our method. The co-clustering task is to recover groups of rows and columns. After the learning stage, the clusters indicators are given by the vectors $u$ and $v$. It can be seen that our method reconstructs efficiently all co-clusters for balanced and unbalanced data sets used in our experiments.

From table 3, we observe that data reordering provided by R1SVD can be useful in a co-clustering context. It is very interesting to underly the fact that R1SVD visualization does not destroy data and, unlike most co-clustering algorithms, it does not require the number of blocks.
Table 3. Confusion Matrix evaluation on rows and columns data.

<table>
<thead>
<tr>
<th>data1 (rows)</th>
<th>data2 (rows)</th>
<th>data1 (columns)</th>
<th>data2 (columns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 205 0</td>
<td>0 0 795</td>
<td>0 0 40</td>
<td>155 0 0</td>
</tr>
<tr>
<td>1614 0 5</td>
<td>626 0 0</td>
<td>397 0 0</td>
<td>0 133 0</td>
</tr>
<tr>
<td>0 0 176</td>
<td>0 579 0</td>
<td>0 63 0</td>
<td>0 0 212</td>
</tr>
<tr>
<td>626 0 0</td>
<td>397 0 0</td>
<td>0 63 0</td>
<td>0 0 212</td>
</tr>
</tbody>
</table>

5 Conclusion

In this paper we have presented an iterative matrix-vector multiplication procedure called rank-one SVD for data visualization. The procedure is to apply iteratively an appropriate stochastic adjacency data matrix associated to a bipartite graph, and then compute the first leading left singular vector associated to the eigenvalue $\lambda_1$ of this matrix. Stopping the algorithm after a few iterations involves a visualization of data matrix into homogeneous blocks. This approach appears therefore very interesting in co-clustering context.

References

2. B. Shneiderman, Inventing discovery tools: combining information visualization with data mining, Inf Vis 1 (2002), 5-12.
Approach for Vehicle Routing Problem with Resource Constraints

A. LAMAMRI(1), A. NAGIH(2), H. AIT HADDADENE(2)
(1) University Khemis-Miliana, Algeria,
(2) Laboratory LITA, University Paul Verlaine – Metz, France.
(3) Laboratory LITA, Lorraine University Ile-Saulcy-Metz, France.
E-mail: plamamri@yahoo.fr

Abstract

The vehicle routing problem with resource constraints is a generalization of the classical vehicle routing problem. Given the large size of the problems encountered in practice, these models are solved by an approach based on column generation that can handle implicitly all feasible solutions and a master problem determining the best solution. In this paper, we present the development of a new approach to improve the acceleration of the method of column generation for solving the problem of construction vehicle routing. It is projected in each arc, the resources a vector of size smaller by using a Lagrangean relaxation algorithm to determine the coefficients of the projection arc, combined with an algorithm for re-optimization, then generates a subset of solutions to the master problem. Numerical tests on problems from instances of random vehicle routing giving competitive results.

Keywords: Vehicle routing problem, path problems in graphs, decomposition methods, column generation.

1 Introduction

The vehicle routing problem with resource constraints (VRPRC) is given by a set of customers \( N \) and a set of \( K \) vehicles available in a repository. This problem is to find a set of minimum cost route, departing and returning to a single repository, where each customer is visited by one vehicle to satisfy some demand. Each customer must be served during a given time window. A vehicle arriving in advance at a customer waits until the start date of service without additional cost, time windows in this case are called 'hard'. Some models penalize the hold with an extra cost, these models are called 'soft', but much research is devoted to the time windows 'hard'. A vehicle arriving late at a customer is not allowed to perform his service. A feasible
tour is a series of visits (conducted by the same vehicle) respecting the time windows, which begins and ends at the same depot.

The \( (VRPRC) \) is defined on the network \( G^k = (X^k, A^k), X = N \cup \{d^k, s^k\} \), where the depot is represented by the two nodes \( s^k \) and \( d^k \), and \( N = \{1, \ldots, n\} \) is the set of vertices representing the customer, and \( A \) the set of arcs that interconnect the customers and the depot. An arc \( (i, j) \in A \) means the possibility of linking the service customers \( i \) and \( j \). To write the formulation of this problem, we introduce the following notations:

- \( c_{ij} \) is the cost of the arc \( (i, j) \in A \).
- \( t_{ij} \) The arc duration \( (i, j) \in A \).
- \([a_i, b_i]\) The time window during which the customer service \( i \in N \) must start.
- \( d_i \) Customer demand \( i \in N \).
- \( Q \) capacity of each vehicle.

Assignment of customers to vehicles is called feasible if:

- The combined demand of customers visited by a vehicle does not exceed its capacity.
- Time constraints are met by each vehicle.
- Each customer is visited by one vehicle.
- Every vehicle that leaves the depot back to depot after completing his route.

The problem is to find a feasible assignment of vehicles to tour the minimum cost.

2 Formulation

The \( (VRPRC) \) can then be formally described as the following multi-commodity network flow model with time window and resource constraints:
\[
\min \sum_{(k=1)}^{K} \sum_{(i,j) \in A} c_{ij} x_{ij}^k \\
\text{a.t} \sum_{(k=1)}^{K} \sum_{(i,j) \in A} x_{ij}^k = 1 \quad \text{for } i \in N = \{1, \ldots, n\} \\
\sum_{(k,j) \in A} d_j x_{ij}^k \leq Q \quad \text{for } k \in K \\
\sum_{(l,j) \in A} x_{lj}^k = 1 \quad \text{for } k \in K \\
\sum_{(l,i) \in A} x_{il}^k = 1 \quad \text{for } k \in K \\
\sum_{(i,j) \in A} x_{ij}^0 = \sum_{(i,j) \in A} x_{ij}^k \quad \text{for } j \in N \\
x_{ij}^k(T_{ij}^{(k,q)} + t_{ij}^{(k,q)} - T_{ij}^{(k,q)}) \leq 0 \quad \text{for } (i, j) \in A, k \in K, q \in Q \\
a_{ij}^{(k,q)} \leq T_{ij}^{(k,q)} \leq b_{ij}^{(k,q)} \quad \text{for } i \in N, k \in K, q \in Q \\
x_{ij}^q \in \{0, 1\}, T_{ij}^{(k,q)} \geq 0 \quad \text{for } (i, j) \in A, k \in K, q \in Q
\]

Binary variables \(x_{ij}^k\) indicate if the tour takes the arc \((i, j) \in A\), while the variable \(T_{ij}^{(k,q)}\) indicates the cumulative consumption of each resource \(q\) at each node \(i\).

The objective function (1) minimizes the total travel cost. The constraints (2) ensure that each customer is visited exactly once, and (2') state that a vehicle can only be loaded up to its capacity. Next, equations (3–5) indicate that each vehicle must leave the depot \(a\); after a vehicle arrives at a customer it has to leave for another destination, and finally, all vehicles must arrive at the depot \(d\). The inequalities (6) establish the relationship between the vehicle departure time from a customer and its immediate successor. Finally, constraints (7) affirm that the time windows are observed, and (8) are the integrality constraints. Note that an unused vehicle is modeled by driving the "empty" route \((a, d)\), and the constraints (9) provide the cumulative consumption of resource \(q\) at node \(j\), since we have:

\[
T_{ij}^{(k,q)} = \max(\alpha_{ij}^{(k,q)}, T_{ij}^{(k,q)} + t_{ij}^{(k,q)})
\]

Note that the constraints (3–7) are local constraints valid only for the network \(G\). Only the partitioning constraints (2) are global constraints linking the \(K\) sub-networks. The relaxation of these binding constraints and the decomposition of the initial problem by sub-network will be an interesting option for a resolution. Finally, note that resource constraints (8–7) make the problem \((VRPFC) NP\)-hard. Even the problem of realizability is associated \(NP\)-complete [5].
3 Solving approaches

And as the number of coefficients to adjust will be more important for the approach of projection arcs, finding the optimal multiplier \( w_{ij}^* \) require several iterations of DPA-L [2], this method can be expensive. To quickly obtain good heuristic solutions (feasible), our approach applied once DPA-L and then apply DPA-LND [2], using multipliers to find \( w_{ij} \) generate feasible columns and negative marginal cost. Specifically, we first choose a sequence of steps \( (p_k) \) as the standard \( (\sum p_k) \) is divergent and \( \lim_{k \to \infty} p_k = 0 \), i.e. conditions ensuring convergence of the algorithm sub gradient. It applies primarily DP-L using multipliers \( w_{ij}^{(k-1)} \) of the previous iteration, we find the sub gradients \( Sg_{ij}^{(k)} \) corresponding \( arc (i,j) \) then we calculates the new Lagrange multipliers \( w_{ij}^{(k)} \). This heuristic is certainly based on the fact that when \( k \) is large, the vector \( C_k \) reduced costs on the arcs of the network do not change much from one iteration to another of the algorithm for column generation. Thus for \( k \) large, we can expect to see \( w_{ij}^{(k)} \) converge to an optimal value.

The main steps of our approach are summarized below:

- **Master problem**
- **Sub-problem** - calculate the Lagrange multipliers \( w_{ij}^{(k)} \) (projection arc).
  - calculate the solution \( ZL(w_{ij}^{(k)}) \), used in DPA-L.
  - calculate feasible solutions \( ZLND(w_{ij}^{(k)}) \), used in DPA-LND.

Generated the solutions negative reduced cost.

4 Numerical results

This section presents the preliminary evaluation of our approach to the problem of construction of vehicle routing with a single resource.

Solomon's 100-customer Euclidean (VRPTW) instances are used to test our algorithm. In these instances, the travel time and the Euclidean distance between two customer locations are the same and this value is truncated to two decimal places. There are six different classes of instances depending on the geographic location of the customers (\( R \): random; \( C \): clustered; \( RC \): mixed) and width of the scheduling horizon (1: short horizon; 2: long horizon). In this work, instances of type 1 are discarded due to the short horizon that does not allow a significant number of routes to be sequenced to form a workday. Results are thus reported for \( R2 \), \( C2 \) and \( RC2 \). Due to the limitations of our exact approach, the computational study focuses on instances obtained by taking only the first 25 customers from each original instance.
Solomon's (VRPTW) test instances are modified to fit our problem. In particular, a value \( t_{\text{max}} \) to limit route duration is needed. This value was first set to 100 in the case of \( R2 \) and \( RC2 \), and 200 in the case of \( C2 \). The value is larger for \( C2 \) because the service or dwell time at each customer is 90, as opposed to 10 for \( R2 \) and \( RC2 \). Finally, a gain of 1 is associated with each customer and weighted by an arbitrarily large constant to maximize the first number of served customers, and then minimize the total distance.

The results for the instances with reduced time windows are shown in Table1. In the table, a particular instance is identified by its class and its index followed by a dot and the number of customers considered. For example, \( RC202.25 \) is the second instance of class \( RC2 \), where only the first 25 customers are considered. In these tables, column Problem is the identifier of the problem instance, \( Itr_{\text{GC}} \) is the total number of iteration of (PM) solved by Simplex, \( \text{Col} \) is the total number of columns generated during the branch-and-price algorithm, \( T(\text{ssp}) \) is the computation time in seconds and \( \text{Obj} \) is the total distance.

<table>
<thead>
<tr>
<th>Problems</th>
<th>( Itr_{\text{GC}} )</th>
<th>( \text{Col} )</th>
<th>( T(\text{ssp}) )</th>
<th>( \text{Obj} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC201.25</td>
<td>123</td>
<td>609</td>
<td>0.9</td>
<td>957.9</td>
</tr>
<tr>
<td>RC202.25</td>
<td>110</td>
<td>1132</td>
<td>221.0</td>
<td>961.6</td>
</tr>
<tr>
<td>RC203.25</td>
<td>713</td>
<td>2589</td>
<td>2566.2</td>
<td>751.3</td>
</tr>
<tr>
<td>RC205.25</td>
<td>218</td>
<td>944</td>
<td>5.4</td>
<td>974.9</td>
</tr>
<tr>
<td>RC206.25</td>
<td>444</td>
<td>1703</td>
<td>4.6</td>
<td>977.1</td>
</tr>
<tr>
<td>RC207.25</td>
<td>3119</td>
<td>13989</td>
<td>418.4</td>
<td>819.6</td>
</tr>
<tr>
<td>R201.25</td>
<td>218</td>
<td>577</td>
<td>1.0</td>
<td>772.8</td>
</tr>
<tr>
<td>R202.25</td>
<td>108</td>
<td>1030</td>
<td>127.0</td>
<td>694.0</td>
</tr>
<tr>
<td>R205.25</td>
<td>1326</td>
<td>4930</td>
<td>60.1</td>
<td>761.2</td>
</tr>
<tr>
<td>R210.25</td>
<td>71</td>
<td>918</td>
<td>121.4</td>
<td>704.6</td>
</tr>
<tr>
<td>R211.25</td>
<td>57</td>
<td>1150</td>
<td>42.9</td>
<td>623.7</td>
</tr>
<tr>
<td>C201.25</td>
<td>329</td>
<td>3448</td>
<td>5.1</td>
<td>679.5</td>
</tr>
<tr>
<td>C202.25</td>
<td>4023</td>
<td>13860</td>
<td>782.8</td>
<td>677.3</td>
</tr>
</tbody>
</table>

Table1

The comparison between the different methods and our approach has revealed that it has provided good results. These are best when certain conditions are met:

The initialization of the algorithm and the choice of Lagrange multipliers and the displacement step.
Conclusions

In this paper, we proposed an algorithm for vehicle routing problem with resource constraints (VRPRC) which is an extension of the classical vehicle routing problem (VRP) to take into account the more practical problem; we have mainly developed approaches to column generation and decomposition master problem and sub-problem. The difficulty of solving the sub-problem is directly related to the number of resources; we particularly studied the techniques of reduction of space resources, and this notion of reduction is a key element of the effectiveness of the overall resolution of problem.

References


SGR modeling of fake ordinal data with correlational structures

Luigi Lombardi\(^1\), Massimiliano Pastore\(^2\),
Massimo Nucci\(^3\), and Andrea Bobbio\(^4\)

\(^1\) Dipartimento di Psicologia e Scienze Cognitive, Università di Trento,
Corso Bettini, 31, I-38068 Rovereto (TN), Italy
(E-mail: luigi.lombardi@unitn.it)

\(^2\) Dipartimento di Psicologia dello Sviluppo e della Socializzazione,
Università di Padova, via Venezia, 8, I-35131 Padova, Italy
(E-mail: massimiliano.pastore@unipd.it)

\(^3\) Dipartimento di Psicologia Generale, Università di Padova,
via Venezia, 8, I-35131 Padova, Italy
(E-mail: massimo.nucci@unipd.it)

\(^4\) Dipartimento di Filosofia, Sociologia, Pedagogia e Psicologia Applicata,
Università di Padova, via Venezia, 8, I-35131 Padova, Italy
(E-mail: andrea.bobbio@unipd.it)

Abstract. In many psychological inventories (i.e., personnel selection surveys and diagnostic tests) the collected samples often include fraudulent records. This confronts the researcher with the crucial problem of biases yielded by the usage of standard statistical models. In this paper we generalize a recent probabilistic perturbation procedure, called SGR - Sample Generation by Replacements - (Lombardi & Pastore\(^4\)), to simulate fake data with correlational structures. To mimic these more complex faking data we proposed a novel extension of the SGR conditional replacement distribution which is based on a discrete version of the truncated multivariate normal distribution. We also applied the new procedure to real behavioral data on the role of perceived affective self-efficacy in social contexts.

Keywords: Sample Generation by Replacement, Fake-good data, Truncated multivariate normal distribution.

Many self-report measures of attitudes, beliefs, personality, and pathology include items that can be easily manipulated by respondents. For example, an individual may deliberately attempt to manipulate or distort responses to personality inventories and attitude tests to create positive impressions (e.g., Paulhus\(^7\); Zickar & Robie\(^8\)). In other circumstances, some individuals may tend to malinger responses on a symptom checklist to simulate grossly exaggerated physical or psychological symptoms in order to reach specific goals such as, for example, obtaining financial compensation, avoiding being charged with a crime, avoiding military duty, or obtaining drugs (e.g., Hall & Hall\(^2\)).

Sample Generation by Replacement (SGR) is a recent data simulation procedure to artificially generate samples of fake ordinal data (Lombardi & Pastore\(^4\)). SGR is based on a two-stage sampling algorithm characterized by two distinct generative models: the model representing the process that generates the data prior to any fake perturbation (data generation process) and the
model representing the faking process to perturb the data (data replacement process). This approach has been recently applied to evaluate the impact of hypothetical faking good manipulations in ordinal data on the performances of a set of widely known and commonly used stand-alone SEM-based fit indices (Lombardi & Pastore[4]). SGR has also been used to study the sensitivity of reliability indices to fake perturbations in dichotomous and ordered data generated by factorial models (Pastore & Lombardi[6]).

Up to now the SGR approach has been limited to the modeling of conditionally independent fake data. In this contribution, we propose a novel generalization of the conditional replacement distribution that accounts for possible correlated structures in the simulated fake data.

1 Main features of the SGR approach

1.1 Data generation process

We think of the original data as being represented by an $n \times m$ matrix $D$, that is to say, $n$ observations (e.g., participants) each containing $m$ elements (e.g., participant’s responses). We assume that entry $d_{ij}$ of $D$ ($i = 1, \ldots, n; j = 1, \ldots, m$) takes values on a small ordinal range $1, 2, \ldots, Q$. In particular, let $d_i$ be the $(1 \times m)$ array of $D$ denoting the simulated pattern of responses of participant $i$. The response pattern $d_i$ is a multidimensional ordinal random variable with probability distribution $p(d_i | \theta)$, where $\theta$ indicates the vector of parameters of the generative probabilistic model of the data. Moreover, we assume that the simulated response patterns are independent and identically distributed (i.i.d.) observations.

1.2 Data replacement process

The basic principle of the SGR approach is to generate a new $n \times m$ ordinal data matrix $F$, called the fake data matrix of $D$, by manipulating each element $d_{ij}$ in $D$ according to a replacement probability distribution. Let $f_i$ be the $(1 \times m)$ array of $F$ denoting the hypothetical pattern of fake responses of participant $i$. The fake response pattern $f_i$ is a multidimensional ordinal random variable with conditional replacement probability distribution $p(f_i | d_i, \theta_F)$ where $\theta_F$ indicates the vector of parameters of the probabilistic faking model. In general, $\theta_F$ represents hypothetical a priori knowledge about the distribution of faking (e.g., the chance of observing a fake observation in the data) or empirically based knowledge about the process of faking (e.g., the direction of faking - fake good vs. fake bad -). In the SGR framework the replacement distribution $p(f_i | d_i, \theta_F)$ is restricted to satisfy a conditional independence assumption (see Lombardi & Pastore[4]; Pastore & Lombardi[6]). More precisely, in the replacement distribution each fake response $f_{ij}$ only depends on the corresponding data observation $d_{ij}$ and the model parameter $\theta_F$. Therefore, because the patterns of fake responses are also i.i.d. observations, the
simulated data array \((D,F)\) is drawn from the joint probability distribution

\[
p(D,F|\theta, \theta_F) = \prod_{i=1}^{n} p(d_i|\theta)p(f_i|d_i, \theta_F) \tag{1}
\]

\[
= \prod_{i=1}^{n} p(d_i|\theta) \prod_{j=1}^{m} p(f_{ij}|d_{ij}, \theta_F) \tag{2}
\]

By repeatedly sampling data from the two generative models we can simulate the so called fake data sample (FDS). We can then study the distribution of some relevant statistics computed on this FDS.

### 1.3 The problem of the independence assumption

We recall that the SGR simulation procedure generates fake perturbations that are restricted to satisfy the conditional independence assumption. Unfortunately, this restriction clearly limits the range of empirical faking processes that can be mimicked by the SGR simulation procedure. In particular, because the replacement distribution acts as a perturbation process for the original data, the resulting fake data will always show correlations that are (on average) weaker than the ones observed for the original data, thus showing a sort of residual correlation effect. However, it is known that some empirical contexts may require different model assumptions about the faking process that cannot be captured by this simple framework. For example, different modulations of graded faking such as slight faking and extreme faking (e.g., Zickar & Robie[8]) are not consistent with the simple independence hypothesis.

To fill this gap, in this contribution we propose a novel conditional replacement distribution that allows to modulate different levels of correlational patterns in the simulated fake data. Because our new proposal is based on a discrete version of the truncated multivariate normal distribution, in what follows we present some relevant properties of this important distribution function before introducing the new perturbation model that does not hinge on the independence assumption.

### 2 A SGR framework for correlated fake-good data

The SGR approach offers an elegant way to simulate faking good scenarios. In general, faking good can be conceptualized as an individual’s deliberate attempt to manipulate or distort responses to create a positive impression (Paulhus[7]; Zickar & Robie[8]). Notice that, the faking good (as well as the faking bad) scenario always entails a conditional replacement model in which the conditioning is a function of response polarity. This model represents a perturbation context in which responses are exclusively subject to positive feigning:

\[ f_{ij} \geq d_{ij}; \quad i = 1, \ldots, n; j = 1, \ldots, m \]
2.1 The truncated multivariate replacement distribution

As a kernel for the conditional replacement distribution we consider the truncated multivariate normal distribution $TN(\mu, \Sigma, a, b)$ (e.g., Horrace[3]). This distribution can be expressed as

$$f(x|\mu, \Sigma, a, b) = \frac{\exp \left\{ -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right\}}{\int_a^b \exp \left\{ -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right\} dx} \quad (3)$$

for $a \leq x \leq b$ and 0 otherwise. The $(1 \times m)$ vectors $a$ and $b$ are the lower and upper truncation points ($a_j < b_j$; $j = 1, \ldots, m$) for the multivariate normal distribution with $m$ dimensions. Finally, $\mu$ and $\Sigma$ are the location parameter vector and the covariance matrix of the (not truncated) multivariate normal distribution.

Now, let $f_i = (k_1, \ldots, k_m)$ and $d_i = (h_1, \ldots, h_m)$ be the replaced values and the original values for the $i^{th}$ simulated observation, respectively. According to the Underlying Variable Approach (UVA; Muthén[5]) we can set

$$p(f_i|d_i, \theta_F) = \left\{ \begin{array}{ll} \int_{\alpha_{k_j - 1}}^{\alpha_k} \cdots \int_{\alpha_{h_j - 1}}^{\alpha_h} f(x|0, \Sigma, a^i, b^i)dx, & \forall j : 1 \leq h_j \leq k_j \leq Q, \\ 0, & \exists j : k_j < h_j \end{array} \right.$$ 

In the replacement distribution $\mathbf{0}$ is the $1 \times m$ array of zeros representing the location parameter (that is to say $\mu = \mathbf{0}$), whereas the pair $(\alpha_{k_j - 1}, \alpha_{k_j})$ are the thresholds corresponding to the discrete value $k_j$ ($j = 1, \ldots, m$). Following the UVA framework we assume that there exists a continuous data matrix $F^*$ underlying the fake ordinal data matrix $F$. The connection between the ordinal variable $f_{ij}$ and the underlying variable $f^*_{ij}$ in $F^*$ is given by

$$f_{ij} = k_j \iff \alpha_{k_j - 1} < f^*_{ij} < \alpha_{k_j}; \quad i = 1, \ldots, n; j = 1, \ldots, m.$$ 

Recall that to represent an ordinal item with $Q$ categories we need $Q + 1$ thresholds. Finally, the bounds $a^i = (a^i_1, \ldots, a^i_m)$ and $b^i = (b^i_1, \ldots, b^i_m)$ are set to

$$a^i_j = \alpha_{h_j - 1} \quad b^i_j = +\infty, \quad j = 1, \ldots, m$$

where we recall that $(\alpha_{h_j - 1}, \alpha_{h_j})$ is the pair of thresholds corresponding to the value $h_j$ for the original variable $d_{ij}$ in $d_i$. Figure 1 shows an example for a one dimensional case. In sum, the parameter array for the faking model is given by

$$\theta_F = (\alpha, \Sigma).$$

with $\alpha$ being the $m \times (Q - 1)$ threshold matrix.

2.2 Some relevant properties of the new replacement distribution

It is important to point out two interesting properties of the novel replacement distribution for correlated fake data.
Fig. 1. Example of a truncated normal distribution (one dimensional) with \( d = 2 \) and \( f = 3 \). The mean \( \mu \) of the distribution is 0, the bounds \( a \) and \( b \) are \( \alpha_1 \) and \( +\infty \), respectively. The conditional probability \( p(3|2, \theta_F) \) is the shaded area between \( \alpha_2 \) and \( \alpha_3 \).

The first property is related to the probability of replacement \( \pi \). According to the truncated multivariate replacement distribution the value \( 1 - \pi_i \) of the conditional probability of non-replacement

\[
f_i = d_i = (h_1, \ldots, h_m)
\]

is equivalent to

\[
1 - \pi_i = \int_{\alpha_{h_1}}^{\alpha_{h_1-1}} \cdots \int_{\alpha_{h_m}}^{\alpha_{h_m-1}} f(x|0, \Sigma, a^i, b^i)dx
\]

and consequently, the probability of replacement \( \pi_i \) is

\[
\pi_i = 1 - \left( \int_{\alpha_{h_1}}^{\alpha_{h_1-1}} \cdots \int_{\alpha_{h_m}}^{\alpha_{h_m-1}} f(x|0, \Sigma, a^i, b^i)dx \right)
\]

Note that in the original model of replacement (Lombardi & Pastore[6]) the probability of replacement \( \pi \) was an explicit parameter in the replacement distribution, whereas in this new proposal \( \pi \) is an implicit parameter that can be derived by taking the average across the \( n \) distincts \( \pi_i \):

\[
\bar{\pi} = \frac{1}{n} \sum_{i=1}^{n} \pi_i.
\]

The second property is related to the correlational structure of the simulated fake data set \( F \). Unlike the standard model of replacement, in the new configuration we can directly represent correlations between the replaced values for the \( m \) ordinal variables. In particular, the correlation matrix \( R_F \) of \( F \) can be modulated by the covariance matrix \( \Sigma \) in the replacement model. Note, however, that \( \Sigma \) is the covariance matrix of the original (not truncated) multivariate normal distribution. In general, the computation of the first and
second moments is not trivial for the truncated case, since they are obviously not the same as \( \mu \) and \( \Sigma \) from the parametrization of \( TN(0, \Sigma, a, b) \). In particular, it can be seen that truncation can significantly reduce the variance and change the covariance between variables. In the next section we will show some examples of how the resulting correlation matrix \( R_f \) can be affected from the interaction between different modulations of faking (represented by different configurations of threshold values \( \alpha \)) and the structure of the covariance matrix \( \Sigma \).

3 Applicative example

The new replacement distribution is illustrated using data from a questionnaire about the role of perceived affective self-efficacy in personality and social psychology (Bandura, Caprara, Barbaranelli, Gerbino, and Pastorelli[1]). Participants were 463 undergraduate students (389 females) at the University of Padua (Italy). Ages ranged from 18 to 48, with a mean of 20.64 and a standard deviation of 2.71. Data consisted of the participants’ responses to three of the 12 items of the AEP/A scale (Caprara, 2001) scored on a 4-point agree-disagree scale (value 1 denotes that a participant totally disagrees with the statement, whereas value 4 means total agreement with the statement). However, the 463 participants were divided into two groups. The first group \((n_1 = 231)\) received a neutral set of instructions, whereas the second group \((n_2 = 232)\) received ad lib faking instructions. The resulting responses were collected into two empirical data matrices: \( D_e \) for the neutral group and \( F_e \) for the ad lib faking group. As expected the observed responses for the ad lib faking group were affected by fake good observations. More precisely, the participants deliberately manipulated their responses using larger values of the scale to create better impressions. This hypothesis was partially supported by the moderate ceiling effects observed in the data (see Fig. 2). Because no additional items on social desirability was available in the AEP/A inventory, we decided to perform an SGR analysis on the basis of two hypothetical scenarios: slight faking and uniform faking.

3.1 Comparing faking models

An SGR analysis was used to evaluate the mimicking ability of four different faking models with respect to the empirical fake set \( F_e \). In particular, because \( D_e \) and \( F_e \) contained responses collected from two different groups of individuals, we decided to evaluate the simulated fake samples against the empirical marginal means of the three items as well as against their empirical correlations in \( F_e \). To that end, we defined four perturbation models derived by the combination of two factors with two levels each. The first factor defined two structures for the covariance matrix in the truncated replacement distribution: a) the identity matrix \( I \) (denoting an independent model) b) the empirical correlation matrix of \( R_e \) computed on the observed matrix \( F_e \) (representing the correlational structure in \( F_e \)). The second factor represented
two different options for the thresholds in the replacement distribution: a) $\alpha_1 = -0.674, \alpha_2 = 0.0, \alpha_3 = 0.674$ denoting a uniform support fake-good distribution (Lombardi and Pastore[4]) b) $\alpha_1 = -1.591, \alpha_2 = 1.596, \alpha_3 = 3.332$ denoting a slight fake-good distribution (e.g., Zickar & Robie[8]). All the other components (parameters’ values) were set to identical values in all the four faking models. In particular, the original data set $D$ in the conditional replacement distribution was set equal to the empirical data set $D_e$. The latter means that in this application we did not simulate the data $D$, instead we directly used the observed data $D_e$ in the replacement distribution equation. The four faking models were then used to simulated new FDSs.

The results of the SGR analysis are shown in Figures 2,3. Figure 2 represents the simulated marginal means of the fake-good data as a function of the two factors. The results showed that the slight faking model with mild dependency was preferred to the other three models. Similarly, figure 3 shows the simulated correlations of the fake-good data as a function of the two factors. Like for the marginal means, also for the correlations the slight faking model with mild dependency showed a better performance as compared to the other three models. In other words, the ad lib faking instructions seem to be more consistent with a mild positive impression effect that boosts the correlations between the items.

![Fig. 2. Boxplots for the simulated marginal means of the fake-good data for the four models. The solid line denotes the observed pattern for the marginal means in $D_e$. The dashed line indicates the observed pattern for the marginal means in $F_e$. UF = uniform support fake-good distribution, SL = slight fake-good distribution; S0 = independent faking model, SF = faking model with a correlational structure. $a_1, a_2$, and $a_3$ denote the three selected items of the AEP/A scale. The data represented in each boxplot were derived from 500 FDSs.](image-url)
Fig. 3. Boxplots for the simulated correlations of the fake-good data for the four models. The solid line denotes the observed correlational pattern in $D_e$. The dashed line indicates the observed correlational pattern in $F_e$. UF = uniform support fake-good distribution, SL = slight fake-good distribution; S0 = independent faking model, Sf = faking model with a correlational structure. The variable $rf_{jj'}$ denotes the correlation between item $a_j$ and item $a_j'$ of the AEP/A scale. The data shown in each boxplot were derived from 500 FDSs.

References

Adaptive Monitoring Rules based on
Scan Statistics for Compound Patterns with
an Application to Quality Control
in Blood Banking

Wendy Lou¹, James C. Fu², and Qilong Yi³

¹ Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, M5T 3M7, Canada
(E-mail: wendy.lou@utoronto.ca)
² Department of Statistics, University of Manitoba, Winnipeg, Manitoba, R3T 2N2, Canada
(E-mail: fu@ad.umanitoba.ca)
³ National Epidemiology and Surveillance, Canadian Blood Services, Ottawa, Ontario, K1G 4J5, Canada
(E-mail: qi-long.yi@blood.ca)

Abstract. Scan Statistics have been used widely in various biomedical applications, such as genomic sequence analysis, and also in public health sciences, including healthcare surveillance. This study is motivated by a practical application involving quality control of allogeneic blood components, where a set of fixed sampling criteria is used to test all blood production sites regardless of the number of blood products collected at individual sites. It is commonly found in real data that the variation of production ranges 5 to 15 fold across blood centers over the monitoring duration, and hence, for example, testing a fixed percentage of blood units will likely result in sub-optimal usable blood production.

We propose a scan-like approach using compound patterns to provide a flexible adaptive strategy for determining the quantity of samples to be tested. The approach will allow varying the sampling criteria based on sequential monitoring, to minimize the number of testing samples and therefore optimize the usable total production of blood units. The proposed approach utilizes the concept of multiple waiting times, in the sense of runs and patterns, and the decision rules to modify the testing criteria can be considered under a Markov-dependent process. An example using real data from the quality testing of blood products will be presented to illustrate the methodology.

Keywords: Runs and Patterns, Sampling Inspection, Markov Chain.

1 Introduction

Blood product components are tested on a multitude of criteria to ensure patient safety; a sample of test criteria from the Canadian Blood Services is shown in Table 1. Typically, a fixed fraction of each product is tested regardless of the quality history of the particular production process, as per guidelines from national health authorities in Canada and the US, for example. There is concern among health professionals that, on the one hand, inadequate quality control
Table 1. Example of monthly quality control testing requirements.

<table>
<thead>
<tr>
<th>Component</th>
<th>Test</th>
<th>Criteria</th>
<th>Sample Age</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Blood Cells, with SAGM additive</td>
<td>Hemoglobin</td>
<td>≥40g/unit in 90% of units tested and ≥35g/unit in 100% units tested</td>
<td>Within 24 hours after product expiry</td>
<td>1% of production, minimum of 10 per month</td>
</tr>
<tr>
<td></td>
<td>Sterility</td>
<td>No growth</td>
<td>Expired</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Volume</td>
<td>±10% labeled volume</td>
<td>At any time</td>
<td></td>
</tr>
<tr>
<td>Red Blood Cells, LR, Deglycerolated</td>
<td>Hemoglobin</td>
<td>≥35g/unit in 95% units tested</td>
<td>Within 24 hours after expiry</td>
<td>1% of production, minimum of 1 per month</td>
</tr>
<tr>
<td></td>
<td>%RBC Recovery</td>
<td>≥80%</td>
<td>Within 24 hours after expiry</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sterility</td>
<td>No growth</td>
<td>Expired</td>
<td></td>
</tr>
</tbody>
</table>

could cause harm, and that, on the other, excessive testing is wasting precious resources. A need exists for more advanced sampling plans that are tailored to the particular manufacturing processes, and that are sufficiently easy to implement by processing sites of various sizes.

In this article, we are reporting some of our ongoing research on the use of scan statistics in multi-level sampling plans, based on the finite Markov chain imbedding technique. In particular, we focus on two levels of inspection, normal and tightened, to allow for reduced sampling compared to fixed sampling schemes currently in use.

2 Adaptive monitoring rules

Let \( \{X_i\}_{i=1}^n \) be a sequence of Bernoulli trials with probabilities \( P(X = 1) = p \) and \( P(X = 0) = 1 - p \), where \( p \) is the probability of lot nonconformance. Let

\[
S_n = \sum_{i=1}^{n} X_i
\]

(1)

be the sum of the sequence. The scan statistic \( S_n(r) \) of window size \( r \) for the sequence \( \{X_i\}_{i=1}^n \) is defined as [4,5]

\[
S_n(r) = \max_{r \leq t \leq n} S(r, t),
\]

(2)

where \( S(r, t) \), for \( t = r, \ldots, n \), are partial sums given by

\[
S_n(r, t) = \sum_{k=t-r+1}^{t} X_k.
\]

Recently, Lachenbruch et al. [6] suggested the use of the scan statistic for quality control in blood product manufacturing. For example, given \( n, r \) and \( p \),
after finding \( k \) such that \( P(S_n(r) \geq k) = 0.05 \), we have a test for \( p \) and for clustering in the sequence \( \{X_i\}_{i=1}^{n} \). If too many \( k \) lots are nonconforming within any window, then the process is deemed nonconforming and has potentially failed. This procedure is highly dependent on the predetermined sample size \( n \) and on the window of size \( r \) sliding continuously until the \( n \)-th observation. Currently in Canada, one percent of blood products are typically slated for quality testing and subsequently discarded. Many automated production processes, however, are very stable after validation, and for such processes it may be desirable to reduce the number of samples in order to lower the various costs associated with quality testing. Sampling plans that switch between multiple levels of inspection depending on the quality history of the process have been developed and used in various manufacturing industries (e.g. Military Standard 105E, Lou and Fu [7]). In this article, we propose a two-level sampling plan for blood product manufacturing, based on either the sum \( S_n = \sum_{i=1}^{n} X_i \) or the scan statistic \( S_n(r) \).

### 2.1 Two-stage switching control rule based on the sum \( S_n \)

(i) Normal Inspection

- Step 1. Select a small or moderate sample size \( m_1 \) (e.g. \( m_1 = 40, 50 \), or some fraction of \( n \)).
- Step 2. Find the integers \( k_1 \) and \( k_2 \) such that

\[
P(S_{m_1} \geq k_i | p) = \alpha_i, \ i = 1, 2,
\]

for given small significance levels \( \alpha_1 < \alpha_2 \) (e.g. \( \alpha_1 = 0.01 \) and \( \alpha_2 = 0.05 \)). This step can be carried out efficiently using the finite Markov chain imbedding technique.
- Step 3. After \( S_{m_1}^{ob} \) is observed, it is compared with \( k_1 \) and \( k_2 \):
  (a) if \( S_{m_1}^{ob} \geq k_1 \), the process is deemed nonconforming (out-of-control);
  (b) if \( S_{m_1}^{ob} < k_2 \), sampling continues under normal inspection;
  (c) if \( k_2 \leq S_{m_1}^{ob} < k_1 \), sampling switches to tightened inspection.

(ii) Tightened Inspection

For \( k_2 \leq S_{m_1}^{ob} < k_1 \), we define a sample-size switching function

\[
1 < \rho(k_1, k_2, S_{m_1}^{ob}),
\]

and set

\[
m_2 = [1 + \rho]m_1. \tag{3}
\]

The tightened inspections should be conducted with a larger sample size, and with a larger \( \alpha_3 \) \( (\alpha_3 > \alpha_2) \), so that the test will have better power. This means finding \( k_3 \) such that

\[
P(S_{m_2} \geq k_3 | p) = \alpha_3, \tag{4}
\]

and the power \( \beta(k_1, m_2, p + \delta) \) of the test under the larger rate of nonconformance \( p + \delta \), where \( \delta \) is a small quantity greater than zero, is given by

\[
\beta(k_1, m_2, p + \delta) = P(S_{m_2} \geq k_3 | p + \delta). \tag{5}
\]
Although the normal approximation could be used to evaluate the probabilities associated with equations (4) and (5), it often lacks accuracy. Here we provide the following simple theorem (for proof, see Fu and Lou[2]) to compute these probabilities.

**Theorem 1.** Given an integer \( m \), \( 0 < p < 1 \) and \( 0 \leq k \leq m \), the statistic \( S_m \) is finite Markov chain imbeddable in the sense that there exists a homogeneous Markov chain \( \{ Y_t \}_{t=1}^m \) defined on the state space \( \Omega = \{ 0, 1, \cdots, k-1, \alpha \} \), such that

\[
P(S_m < k | p) = P(Y_t \in \Omega \setminus \alpha, \forall t = 1, \cdots, m) = \xi_0 N^{m1'},
\]

where \( \alpha \) is an absorbing state, \( \xi_0 \) is the initial distribution, \( 1 = (1, 1, \cdots, 1) \), and \( N \) is the fundamental matrix of the transition matrix \( M \), which has the form

\[
M = \begin{pmatrix}
0 & 1 - p & p & 0 & \cdots & 0 & 0 \\
1 & 0 & 1 - p & p & 0 & \cdots & 0 \\
2 & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
k - 1 & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\alpha & 0 & 0 & \cdots & 0 & 1 & p \\
0 & 0 & \cdots & 0 & 1 - p & p & 0 \\
0 & 0 & \cdots & 0 & 1 & 0 & 0 \\
\end{pmatrix}
\]

(7)

For example, to compute the probability of the type II error \((1 - \beta)\), one needs only to replace \( p \) by \((p + \delta)\) and the computation remains the same.

Given \( p \), \( \delta \), and predetermined \( \alpha \) and power \( \beta \) under nonconformance rate \( p + \delta \), equation (6) provides for finding, via numerical study, optimal \( m_1 \) and \( m_2 \) (also \( k_i \)'s), and a switching rule with power higher or equal to the predetermined power \( \beta \).

### 2.2 Adaptive scan statistics control rule

If the rate of nonconformance and the clustering of nonconforming lots both are relevant to the quality assessment, then the scan statistic \( S_n(r) \) could be employed in a control rule in the same way as \( S_n \) was used in the previous section. It is well known the scan statistic \( S_n(r) \) is associated with compound patterns and is finite Markov chain imbeddable in the following sense (see Fu[1] and Lou and Fu[7]):

**Theorem 2.** Given the sequence \( \{ X_i \}_{i=1}^n \), window size \( r \), and \( s \) \((s \leq r)\), then (i) the event

\[ \{ S_n(r) < s \} \leftrightarrow \{ \text{no compound pattern } A_{r,s} \text{ occurred in the sequence } \{ X_i \}_{i=1}^n \} \]

where the compound pattern \( A_{r,s} \) is the union of \( l \) distinct simple patterns:

\[ A_{r,s} = \bigcup_{A_i \in \mathcal{F}(r,s)} A_i, \]

where

\[ \mathcal{F}(r,s) = \{ A_1 = 11 \cdots 1, \ldots, A_t = 100 \cdots 01 \}_{r}, \]

and

\[ l = \sum_{v=0}^{r-s} \binom{s - 2 + v}{v}, \]

600
(ii) there exists a Markov chain \( \{Y_t\}_{t=1}^\infty \) induced by the compound pattern \( \Lambda_{r,s} \) with transition matrix

\[
M = \begin{pmatrix}
N_{r,s} & C_{r,s} \\
0 & 1
\end{pmatrix}
\]

such that

\[
P(S_n(r) < s) = \xi_0 N^m_{r,s} 1'.
\]

For detailed proofs of the above results, see Fu [1] and Fu et al. [3]. Note also that equation (9) plays the same role as equation (6).

(i) Normal Inspection

Given an initial sample size \( m_1 \), observations \( \{X_i\}_{i=1}^{m_1} \) and \( \alpha_1 < \alpha_2 \), find \( s_i \) such that

\[
P(S_m(r) \geq s_i | p) = \alpha_i, \quad i = 1, 2,
\]

analogous to the previous section. If the observed \( S_{m_1}^{\text{ob}}(r) \) is greater than \( s_1 \), we declare the process is nonconforming. If \( S_{m_1}^{\text{ob}}(r) \) is less than \( s_2 \), then we continue the normal inspections by taking \( m_2 = m_1 \) observations again. If \( S_{m_2}^{\text{ob}} \geq s_1 \), we declare the process is nonconforming; otherwise we declare the process is conforming. If the observed \( S_{m_1}^{\text{ob}}(r) \) is between \( s_2 \) and \( s_1 \) \((s_2 \leq S_{m_1}^{\text{ob}}(r) < s_1)\), we switch to a tightened inspection rule by taking \( m_2 \) observations with larger \( \alpha_2 \):

\[
m_2 = [1 + \rho(s_1, s_2, S_{m_1}^{\text{ob}}(r))] m_1,
\]

where \( \rho(s_1, s_2, S_{m_1}^{\text{ob}}(r)) \geq 0 \), a switching coefficient.

(ii) Tightened Inspection

Under tightened inspection, the sample size \( m_2 \) is at least twice larger than \( m_1 \), and \( \alpha_2 \) is also larger than (or equal to) \( \alpha_1 \). First, we find \( s_2 \) such that

\[
P(S_{m_2}(r) \geq s_2 | p) = \alpha_2.
\]

Then we declare the process is nonconforming if \( S_{m_2}^{\text{ob}} \geq s_2 \); otherwise we say the process is conforming.

The most difficult part is to properly arrange \( \alpha_1, \alpha_2, m_1 \) and \( \rho(s_1, s_2, S_{m_1}^{\text{ob}}) \) so that the above two-stage switching control rule has approximately the same type I error as the procedure with fixed sample size \( n \), and also that

\[
\beta(m_1, m_2, p, \delta) = P\text{(two-stage switching control rule}|p + \delta) \\
\geq \text{pre-determined power (e.g. 95%).}
\]

3 Remarks

Numerical results for these two-stage switching control rules and also for fixed sample-size control rules will be given and compared for various parameters during the presentation at the ASMDA2013. A real application to quality testing of allogeneic blood products will also be discussed.
References

Robust extrapolation of functionals of stochastic sequences with stationary increments

M. Luz and M. Moklyachuk

Department of Probability Theory, Statistics, and Actuarial Mathematics, National Taras Shevchenko University of Kyiv, Kyiv 01601, Ukraine
(E-mail: mmp@univ.kiev.ua)

Abstract. The problem of optimal estimation of the functional $A \xi = \sum_{k=0}^{\infty} a(k)\xi(k)$ depending on the unknown values of a stochastic sequence $\xi(k)$ with stationary increments of order $n$ from observations of the sequence for the moments of time $k = -1, -2, \ldots$ is considered. Formulas for calculating mean square error and spectral characteristic of optimal linear estimation of the functional are proposed in the case where spectral density is exactly known. When the set admissible spectral densities is given, formulas that determine the least favorable spectral densities are proposed.

Keywords: Sequence with stationary increments, minimax-robust estimation, mean square error, least favorable spectral density, minimax-robust spectral characteristic.

1 Introduction

Stochastic processes with $n$th stationary increments were being considered by A.M. Yaglom (see [15]). He described a spectral representation of these processes and a canonical representation of the spectral density, solved an extrapolation problem. An extrapolation problem of the process with stationary increments was defined and solved in term of increments of the process. Stochastic processes with $n$th stationary increments were investigated by M.S. Pinsker [11], A.M. Yaglom and M.S. Pinsker [10] as well.

Traditional methods of solution of the linear extrapolation, interpolation and filtering problems for stationary stochastic processes were proposed by A. N. Kolmogorov [5], N. Wiener [13], A. M. Yaglom [14]. When the complete information on the spectral density is not available but the set of admissible spectral densities is given, minimax estimation methods are applied. These methods consist of finding the estimation which minimizes the value of mean square error for each density in the set. The minimax approach to extrapolation problem for stationary processes was proposed by Ulf Grenander [1].

J. Franke [2], J. Franke and H. V. Poor [3] investigated the minimax extrapolation and filtering problems for stationary sequences with the help of convex optimization methods. In the papers by M. P. Moklyachuk [6] -[9] the minimax approach to extrapolation, interpolation and filtering problems was applied to the functionals that depend on the unknown values of stationary processes and sequences.

In this article we are interested in finding the estimations of the unknown values of the functional $A \xi = \sum_{k=0}^{\infty} (k)\xi(k)$ that depends on the unknown values of a stochastic sequence $\xi(k)$ with stationary increments of order $n$ from
observations of the sequence $\xi(k)$ for $k = -1, -2, \ldots$. The extrapolation problem for the sequences with stationary increments is solved in case where the spectral density is known. In case where the spectral density is not known but the set of admissible spectral density is given, the least favorable spectral density for the optimal extrapolation of the functional $A\xi$ was found.

2 Stationary increments. Spectral representation

Definition 1. For a given stochastic sequence $\{\xi(m), m \in Z\}$ a function

$$\xi^{(n)}(m, \mu) = (1 - B_\mu)^n \xi(m) = \sum_{l=0}^{n} (-1)^l C^n_l \xi(m - l\mu),$$  \hspace{1cm} (1)

where $B_\mu$ is a backward operator with step $\mu \in Z$, $B_\mu \xi(m) = \xi(m - \mu)$, is called stochastic $n$-th increment with step $\mu \in Z$.

Definition 2. Stochastic increment $\xi^{(n)}(m, \mu)$ of some stochastic sequence $\{\xi(m), m \in Z\}$ is called stationary increment in wide sense if mathematical expectations

$$E\xi^{(n)}(m_0, \mu) = c^{(n)}(\mu),$$

$$E\xi^{(n)}(m_0 + m, \mu_1)\xi^{(n)}(m_0, \mu_2) = D^{(n)}(m, \mu_1, \mu_2)$$

exist for all $m_0, \mu, m, \mu_1, \mu_2 \in Z$ and do not depend on $m_0$. The function $c^{(n)}(\mu)$ is called the mean value and the function $D^{(n)}(m, \mu_1, \mu_2)$ is called the structural function of the stationary $n$th increment. Stochastic sequence $\{\xi(m), m \in Z\}$ which defines a stationary $n$th increment $\xi^{(n)}(m, \mu)$ by formula (1) is called the sequence with stationary $n$th increments.

Theorem 1. The functions $c^{(n)}(\mu)$ and $D^{(n)}(m, \mu_1, \mu_2)$ can be presented in the forms

$$c^{(n)}(\mu) = c\mu^n,$$

$$D^{(n)}(m, \mu_1, \mu_2) = \int_{-\pi}^{\pi} e^{i\lambda m} (1 - e^{-i\mu_1 \lambda})^n (1 - e^{-i\mu_2 \lambda})^n \frac{1}{\lambda^{2n}} dF(\lambda),$$ \hspace{1cm} (2)

where $c$ is a constant, function $F(\lambda)$ is a left-continuous nondecreasing bounded function, $F(-\pi) = 0$. That is more the constant $c$ and the function $F(\lambda)$ are uniquely defined by the stochastic increment $\xi^{(n)}(m, \mu)$.

Using the representation (2) of the structural function of the stationary $n$th increment $\xi^{(n)}(m, \mu)$ and the Karunen’s theorem [4], we get the following representation of the stationary $n$th increment $\xi^{(n)}(m, \mu)$:

$$\xi^{(n)}(m, \mu) = \int_{-\pi}^{\pi} e^{im\lambda} (1 - e^{-i\mu \lambda})^n \frac{1}{(i\lambda)^n} dZ(\lambda),$$ \hspace{1cm} (3)

where $Z(\lambda)$ is an orthogonal random measure on $[-\pi, \pi]$ related with the structural function $F(\lambda)$.
Let the stationary $n$th increment $\xi^{(n)}(m, \mu)$ admit the one sided moving average representation

$$\xi^{(n)}(m, \mu) = \sum_{k=0}^{\infty} \varphi^{(n)}(k, \mu) \varepsilon(m - k) \tag{4}$$

with some renovating sequence $\{\varepsilon_m : m \in \mathbb{Z}\}$ [15] and some complex sequence $\{\varphi^{(n)}(k, \mu) : m \geq 0\}$, $\sum_{k=0}^{\infty} |\varphi^{(n)}(k, \mu)|^2 < \infty$. Then the spectral function $F(\lambda)$ of the increment $\xi^{(n)}(m, \mu)$ has a spectral density $f(\lambda)$ admitting a canonical factorization

$$f(\lambda) = |\Phi(e^{-i\lambda})|^2, \tag{5}$$

where the function $\Phi(z) = \sum_{k=0}^{\infty} \varphi(k) z^k$ has a convergence radius $r > 1$ and does not have any zero in $|z| \leq 1$. Let us define $\Phi_{\mu}(z) = \sum_{k=0}^{\infty} \varphi_{\mu}(k) z^k$ where $\varphi_{\mu}(k) = \varphi^{(n)}(k, \mu)$ are the coefficients from the representation (4). Then the following equality takes place

$$|\Phi_{\mu}(e^{-i\lambda})|^2 = \frac{|1 - e^{-i\lambda \mu}|^{2n}}{\lambda^{2n}} f(\lambda). \tag{6}$$

We use the representation (4) to find the optimal mean square estimation of unknown values of the sequence with stationary increment.

3 Extrapolation of the linear functional $A\xi$

Let a stochastic sequence $\{\xi(m), m \in \mathbb{Z}\}$ determine a stationary $n$th increment $\xi^{(n)}(m, \mu)$ with absolutely continuous spectral function $F(\lambda)$ which has spectral density $f(\lambda)$. Let the increment $\xi^{(n)}(m, \mu)$ admit the one sided moving average representation (4) and the spectral density $f(\lambda)$ admit the canonical factorization (5). From now we make the assumption that observation of the sequence $\xi(m)$ for $m < 0$ are known. Consider a problem of finding optimal in mean square sense linear estimation of the functional $A\xi = \sum_{k=0}^{\infty} a(k) \xi(k)$ depending on unknown values $\xi(m), m \geq 0$.

Let $\{d_{\mu}(k) : k \geq 0\}$ be the coefficients from a relation $\sum_{k=0}^{\infty} d_{\mu}(k)x^k = \left(\sum_{j=0}^{\infty} x^{\mu j}\right)^{-1}$. The functional $A\xi$ can be presented as $A\xi = B\xi - V\xi$, where

$$B\xi = \sum_{k=0}^{\infty} b_{\mu}(k) \xi^{(n)}(k, \mu), \quad V\xi = \sum_{k=-\mu n}^{-1} v(k) \xi(k),$$

$$v_{\mu}(k) = \sum_{l=-\mu n}^{n} (-1)^l C^l_{\mu} b_{\mu}(l\mu + k), \quad k = -1, -2, \ldots, -\mu n, \tag{7}$$

$$b_{\mu}(k) = \sum_{m=k}^{\infty} a(m)d_{\mu}(m-k) = (D^\mu a)_k, \quad k \geq 0, \tag{8}$$
and $[x]$ denotes the least integer number among numbers greater or equal to $x$. $D^\mu$ is a linear operator defined by the elements $D^\mu_{k,j} = d_\mu(j - k)$ if $0 \leq k \leq j$, and $D^\mu_{k,j} = 0$ if $j < k$; vector $a = (a(0), a(1), a(2), \ldots)$.

We will suppose that the following conditions hold true

$$\sum_{k=0}^{\infty} |b_\mu(k)| = \sum_{k=0}^{\infty} |(D^\mu a)_k| < \infty, \sum_{k=0}^{\infty} (k+1)|b_\mu(k)|^2 = \sum_{k=0}^{\infty} (k+1)(|D^\mu a)_k|^2 < \infty.$$  

(9)

Under these conditions the functional $B\xi$ has the second moment and the operator $B^\mu$ defined below is compact.

Let $A\xi$ denote the mean square optimal linear estimation of the functional $A\xi$ from observations of the sequence $\xi(m)$ for $m < 0$, $B\xi$ denote the mean square optimal linear estimation of the functional $B\xi$ from observations of the stochastic increment $\xi^{(1)}(m, \mu)$, $\mu > 0$, for $m < 0$. Let $\Delta(f, \tilde{A}) := E|A\xi - A\xi|^2$ and $\Delta(f, \tilde{B}) := E|B\xi - \tilde{B}\xi|^2$ denote the mean square errors of the estimations $A\xi$ and $\tilde{B}\xi$. Since the values $\xi(m)$ for $m = -1, -2, \ldots, -\mu$ are known, we have the equality $A\xi = B\xi - 2\xi$. Therefore the following equalities take place:

$$\Delta(f, \tilde{A}) = E|A\xi - \tilde{A}\xi|^2 = E|A\xi + V\xi - B\xi|^2 = E|B\xi - \tilde{B}\xi|^2 = \Delta(f, \tilde{B}).$$

Let $L^0_2(f)$ be a subspace of the space $L_2(f)$ generated by the functions $\{e^{i\lambda k} : k \leq 1\}$. The optimal estimation $\tilde{B}\xi$ can be written as

$$\tilde{B}\xi = \int_{-\pi}^{\pi} h_\mu(\lambda)(1 - e^{-i\lambda})^n \frac{1}{(i\lambda)^n} dZ(\lambda),$$

(10)

where $h_\mu(\lambda) \in L^0_2(f)$ is a spectral characteristic which provides the minimum value of the mean square error $\Delta(f, \tilde{B})$. Using the projection method of the theory of Hilbert spaces, we obtained he following formula for calculation he optimal spectral characteristic

$$h_\mu(\lambda) = B^\mu(e^{i\lambda}) - r_\mu(e^{i\lambda})\varphi^{-1}_\mu(e^{-i\lambda}),$$

(11)

$$r_\mu(e^{i\lambda}) = \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} b_\mu(m+j)\varphi_\mu(m)e^{i\lambda j} = \sum_{j=0}^{\infty} (B^\mu \varphi_\mu)_j e^{i\lambda j},$$

where $B^\mu$ is a linear symmetric operator defined by the matrix with elements $B^\mu_{k,j} = b_\mu(k + j)$, $k, j \geq 0$. $\varphi_\mu = (\varphi_\mu(0), \varphi_\mu(1), \varphi_\mu(2), \ldots); \varphi_\mu(k)$, $k \geq 0$, are the coefficients from the moving average representation (4).

Further the estimation of the functional $A\xi$ can be presented in the form

$$\tilde{A}\xi = -\sum_{k=-\mu}^{-1} v_{\mu,i}(k)\xi(k) + \int_{-\pi}^{\pi} h_\mu^{(n)}(\lambda)(1 - e^{-i\lambda})^n \frac{1}{(i\lambda)^n} dZ(\lambda),$$

(12)

where the coefficients $v_{\mu,i}$ for $i = -1, -2, \ldots, -\mu$ are defined in (7). Using he relationship (8) between the coefficients $a(k)$ and $b_\mu(k)$, we obtain the equality $(B^\mu \varphi_\mu)_k = (D^\mu A\varphi_\mu)_k$, where the linear operator $A$ is defined by the coefficients
If conditions (9) hold true, the optimal linear estimation \( \hat{A}_\xi \) of the functional \( A_\xi \) of unobserved values \( \xi(m), m \geq 0 \), from observations of the sequence \( \xi(m), m = -1, -2, \ldots \), can be calculated by formula (12). Spectral characteristic \( h_\mu^{(a)}(\lambda) \) of the optimal linear estimation \( \hat{A}_\xi \) can be calculated by formula (13). The value of mean square error \( \Delta(f, \hat{A}) \) is calculated by formula (15).

\[
\Delta(f, \hat{A}) = E|A\xi - \hat{A}_\xi|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |r_\mu^{(a)}(e^{i\lambda})|^2 d\lambda = ||D^\mu A_{\varphi_\mu}||^2.
\]

**Theorem 2.** Let the stochastic sequence \( \{\xi(m), m \in Z\} \) generate the stationary random increments \( \xi^{(a)}(m, \mu), \mu > 0 \), with absolutely continuous structural function \( F(\lambda) \) and spectral density \( f(\lambda) \) admitting the canonical factorization (5). If conditions (9) hold true, the optimal linear estimation \( \hat{A}_\xi \) of the functional \( A_\xi \) of unobserved values \( \xi(m), m \geq 0 \), from observations of the sequence \( \xi(m), m = -1, -2, \ldots \), can be calculated by formula (12). Spectral characteristic \( h_\mu^{(a)}(\lambda) \) of the optimal linear estimation \( \hat{A}_\xi \) can be calculated by formula (13). The value of mean square error \( \Delta(f, \hat{A}) \) is calculated by formula (15).

### 4 Minimax-robust extrapolation

If the spectral density \( f(\lambda) \) of the sequence \( \xi(m) \) with stationary \( n \)th increments is known, the value of mean square error \( \Delta(h_\mu^{(a)}(f); f) := \Delta(f, \hat{A}) \) and the spectral characteristic \( h_\mu^{(a)} \) of the optimal linear estimation \( \hat{A}_\xi \) of the functional \( A_\xi \) can be calculated by formulas (13) and (15). In the case where the spectral density is not known, but the only set \( \mathcal{D} \) of possible spectral densities is given, the minimax (robust) approach to estimation of functionals of the unknown values of random sequence with stationary increments is reasonable. In other words we find an estimation that minimizes the mean square error for all spectral densities from the given class \( \mathcal{D} \) simultaneously.

**Definition 3.** For a given class of spectral densities \( \mathcal{D} \) spectral density \( f_0(\lambda) \in \mathcal{D} \) is called least favorable in \( \mathcal{D} \) for the optimal linear estimate of the functional \( A_\xi \) if the following relation holds true

\[
\Delta(f_0) = \Delta(h_\mu^{(a)}(f_0); f_0) = \max_{f \in \mathcal{D}} \Delta(h_\mu^{(a)}(f); f).
\]

**Definition 4.** For a given class of spectral densities \( \mathcal{D} \) spectral characteristic \( h_0^{(a)}(\lambda) \) of the optimal linear estimate of the functional \( A_\xi \) is called minimax-robust if there are satisfied conditions

\[
h_0^{(a)}(\lambda) \in H_\mathcal{D} = \bigcap_{f \in \mathcal{D}} L_2^{0,-}(f), \quad \min_{h \in H_\mathcal{D}} \max_{f \in \mathcal{D}} \Delta(h; f) = \sup_{f \in \mathcal{D}} \Delta(h_0^{(a)}; f).
\]
Lemma 1. Spectral density $f^0$ which admits factorization (5) is the least favorable in $D$ for the optimal linear estimate of the functional $A\xi$ if

$$f^0(\lambda) = \left| \sum_{k=0}^{\infty} \varphi^0(k)e^{-i\lambda k} \right|^2,$$

where $\varphi^0 = \{\varphi^0(k) : k = 0, 1, 2, \ldots\}$ is a solution to the extremum problem

$$||D^\mu A\varphi^\mu||^2 \rightarrow \max, \quad f(\lambda) = \left| \sum_{k=0}^{\infty} \varphi(k)e^{-i\lambda k} \right|^2 \in D.$$

If $h^\mu_{(a)}(f_0) \in H_D$, the minimax-robust spectral characteristic can be calculated by the formula $h^0_\mu = h^\mu_{(a)}(f_0)$. The functions $h^0$ and $f_0$ form a saddle point of the function $\Delta(h; f)$ on the set $H_D \times D$. The saddle point inequalities

$$\Delta(h; f_0) \geq \Delta(h^0; f_0) \geq \Delta(h^0; f) \quad \forall f \in D, \forall h \in H_D$$

hold when $h^0 = h^\mu_{(a)}(f_0)$ and $h^\mu_{(a)}(f_0) \in H_D$ where $f_0$ is a solution to the conditional extremum problem

$$\tilde{\Delta}(f) = -\Delta(h^\mu_{(a)}(f_0); f) \rightarrow \inf, \quad f \in D,$$

where $r^\mu$ is defined by the formula (14) when $f(\lambda) = f_0(\lambda)$. The previous problem is equivalent to the unconditional extremum problem

$$\Delta_D(f) = \tilde{\Delta}(f) + \delta(f|D) \rightarrow \inf.$$

A solution $f_0$ to this unconditional extremum problem is characterized by the condition $0 \in \partial\Delta_D(f_0)$ [12].

5 Least favorable spectral densities in the class $D_0$

Consider the following set of spectral densities

$$D_0 = \left\{ f(\lambda) : \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda)d\lambda \leq P_0 \right\}.$$

Using the condition $0 \in \partial\Delta_D(f_0)$, we get an equation $|r^\mu_{(a)}(e^{i\lambda})|^2(f^0(\lambda))^{-1} = \psi(\lambda) + c^{-2}$, where $\psi(\lambda) \leq 0$ and $\psi(\lambda) = 0$ when $f^0 > 0$. Therefore the least favorable density in the class $D_0$ for optimal linear estimate of the functional $A\xi$ can be presented as

$$f^0(\lambda) = \left| c \sum_{k=0}^{\infty} (D^\mu A\varphi^\mu)ke^{i\lambda k} \right|^2.$$
Consider the following set of spectral densities

6 Least favorable spectral densities in the class moving average representation. If \( \nu(16) \) is the least favorable in the class of the equation (20) that satisfy condition with stationary 

\[
D^\mu A\varphi_\mu = \alpha \varphi, \quad \alpha \in R. 
\]

For each solution of it such that \( \|\varphi\|^2 = P_0 \) the following equality holds true

\[
f^0(\lambda) = \left| \sum_{k=0}^{\infty} \varphi(k)e^{i\lambda k} \right|^2 = \left| \sum_{k=0}^{\infty} (D^\mu A\varphi_\mu)_k e^{i\lambda k} \right|^2 .
\]

Let \( \nu_0 P_0 \) be the maximum value of \( \|D^\mu A\varphi_\mu\|^2 \) on the set of solutions of the equation (20) that satisfy condition \( \|\varphi\|^2 = P_0 \) and define canonical factorization of the spectral density \( f(\lambda) \). Let \( \nu_0 P_0 \) be the maximum value of \( \|D^\mu A\varphi_\mu\|^2 \) on the set of such \( \varphi \) that satisfy condition \( \|\varphi\|^2 = P_0 \) and define canonical factorization of the spectral density \( f^0(\lambda) \). Consequently, the following statements holds true.

**Theorem 3.** If there exists a sequence \( \varphi^0 = \{ \varphi^0(m) : m \geq 0 \} \) that satisfies conditions \( \|\varphi^0\|^2 = P_0 \) and \( \nu_0 P_0 = \nu_0^2 P_0 = \|D^\mu A\varphi_\mu\|^2 \), the spectral density (16) is the least favorable in the class \( D_0 \) for optimal extrapolation of the functional \( A \xi \) of unknown values \( \xi(k), k = 0, 1, 2, \ldots \), of the stochastic sequence with stationary nth increments. The increment \( \xi^{(n)}(m, \mu) \) admits the one side moving average representation. If \( \nu_0 < \nu_0^2 \), the density (19) which admits the canonical factorization (5) is the least favorable in the class \( D_0 \). Sequence \( c\varphi_\mu = \{ c\varphi_\mu(k) : k \geq 0 \} \) is defined by conditions (17) and \( \int_{-\pi}^{\pi} f(\lambda)d\lambda = 2\pi P_0 \).

6 Least favorable spectral densities in the class \( D_M \)

Consider the following set of spectral densities

\[
D_M = \left\{ f(\lambda) \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) \cos(m\lambda)d\lambda = \rho_m, m = 0, 1, 2, \ldots, M \right. \right\},
\]

where \( \rho_0 = P_0 \) and \( \{ \rho_m, m = 0, 1, 2, \ldots, M \} \) form a strictly positive sequence. From condition \( 0 \in \partial D(f_0) \) we get the equation \( |\varphi^0_\mu(e^{i\lambda})|^2 (f^0(\lambda))^{-1} = \psi(\lambda) + c \sum_{m=1}^{M} \psi_m \cos m\lambda \). Thus the list favorable density in the class \( D_M \) for the optimal linear estimate of the functional \( A \xi \) can be presented as

\[
f^0(\lambda) = \left| \frac{\varphi_0 \sum_{k=0}^{\infty} (D^\mu A\varphi_\mu)_k e^{i\lambda k}}{\sum_{k=1}^{M} c_m e^{-i\lambda k}} \right|^2 .
\]

where the parameters \( c_m, m = 0, 1, 2, \ldots, M, \varphi_\mu = (\varphi_\mu(0), \varphi_\mu(1), \varphi_\mu(2), \ldots) \) can be calculated using conditions (17) and \( \int_{-\pi}^{\pi} f(\lambda) \cos(m\lambda)d\lambda = 2\pi \rho_m, m = 0, 1, 2, \ldots, M, \) factorization (5).

Let \( \nu_M P_0 \) be the maximum value of \( \|D^\mu A\varphi_\mu\|^2 \) on the set of solutions of the equation (20) that satisfy condition \( \|\varphi\|^2 = P_0 \) and define canonical
factorization of the spectral density \( f(\lambda) \). Let \( \nu_M^+P_0 \) be the maximum value of \( ||D^m A\varphi_m||^2 \) on the set of such \( \varphi \) that satisfy condition \( ||\varphi||^2 = P_0 \) and define canonical factorization of the spectral density \( f^0(\lambda) \in \mathcal{D}_M \) defined by (21).

**Theorem 4.** If there exists a sequence \( \varphi^0 = \{\varphi^0(m) : m \geq 0\} \) that satisfies conditions \( ||\varphi^0||^2 = P_0 \) and \( \nu_0P_0 = \nu_M^+P_0 = ||D^m A\varphi^0_m||^2 \), the spectral density (16) is the least favorable in the class \( \mathcal{D}_M \) for optimal extrapolation of the functional \( A\xi \) of unknown values \( \xi(k), k = 0, 1, 2, \ldots \) of the stochastic sequence with stationary \( n \)th increments. If \( \nu_M < \nu_M^+ \), the density (21), which admits the canonical factorization (5), is the least favorable in the class \( \mathcal{D}_M \). Sequence \( \varphi_m = \{\varphi_m(k) : k \geq 0\} \) and unknown parameters \( c_m, m = 0, 1, 2, \ldots, M \), are defined by conditions (17) and \( \int_{-\pi}^{\pi} f(\lambda) \cos(m\lambda)d\lambda = 2\pi \rho_m, m = 0, 1, 2, \ldots, M. \)

**References**