Information Technology and Management Science
doi: 10.1515/itms-2014-0016
2014 / 17

Modeling VIX Index Based on Semi-parametric Markov Models with Frank Copula

Andrejs Matvejevs [1], Jegors Fjodorovs [2], [1, 2] *Riga Technical University*

***Abstract* – The research studies the estimation of a semi-parametric stationary Markov models based on a Frank copula density function. Described techniques allow us to estimate the parameters of the Frank copula, which has a better fit compared to previously selected regression models (estimators of the marginal distribution and the copula parameters are provided). We show how to apply our technique to the financial index VIX – a market mechanism that measures the 30-day forward implied volatility of the underlying index, the S&P500. Moreover, using MatLab we made VIX option index study – found the best copula fit under our condition, estimated nonlinear parameters and showed evaluation steps for copula based semi-parametric models.**

***Keywords* – Diffusion processes, Frank copula, Markov models, semi parametric regressions, VIX index.**

I. Introduction

Copulas have become popular in the finance and insurance community in the past years, where modeling and estimating the dependence structure between several univariate time series are of great interest; see Frees and Valdez (1998) [1] and Embrechts et al. (2002) [2] for review.

A copula function is a multivariate distribution function with standard uniform marginals. By Sklar's (1959) [3] theorem, one can always model any multivariate distribution by modeling its marginal distributions and its copula function separately, where the copula captures all the scale-free dependence in the multivariate distribution [4]. The central result of this theorem, which states that any continuous N-dimensional cumulative distribution function F, evaluated at a point $x = (x_1, \dots, x_n)$ can be represented as

$$F(x) = C(F_1(x_1), \dots, F_n(x_n)),$$

where C is called a copula function and $F_i(x_i)$, $i = 1, \dots, n$ are the marginal distributions [4]. The use of copulas, therefore, splits a complicated problem (finding a multivariate distribution) into two simpler tasks. The first task is to model the univariate marginal distributions and the second task is finding a copula that summarizes the dependence structure between them.

It is also useful to represent copulas as joint distribution functions of standard uniform random variables:

$$U = F(X_1) \text{ and } V = F(X_2)$$
$$C(u,v) = P(U \le u, V \le v)$$

The outcome of uniform random variables falls into the interval [0, 1]; therefore, the domain of a copula must be the N-dimensional unit cube. Similarly, because the mapping represents a probability, the range of the copula must also be the unit interval. It is also easy to determine the value of a copula on the border of its domain. When one argument equals zero, the probability of any joint event must also be zero. Similarly, when all but one of the inputs is equal to one, the joint probability must be equal to the (marginal) probability of the argument that does not equal one. Finally, the function must be increasing in all its arguments [4].

Apart from the standard distribution functions, copulas have associated densities:

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}$$

which permit the bivariate density f(u; v) as the product of the copula density and the density functions of the margins

$$f(u,v) = c(F_1(u), F_2(v)) f_1(u) f_2(v)$$

The expression above indicates how the simple product of two marginal distributions will fail to properly measure the joint distribution of two asset prices unless they are in fact independent and the dependence information captured by the copula density [4], $c(F_1(u), F_2(v))$, is normalized to unity.

II. Copula-Based Semi-parametric Models for Stohastic Simulations

The possibility of identifying nonlinear time series using nonparametric estimates of the conditional mean and conditional variance was studied in many papers (see, for example, [5], and references there). As a rule, analyzing the dependence structure of stationary time series $\{x_t, t \in Z\}$ regressive models defined by invariant marginal distributions and copula functions that capture the temporal dependence of the processes. As indicated in [5], this allows separating out the temporal dependence (such as tail dependence) from the marginal behavior (such as fat tails) of a time series. One more advantage of this type regressive approach is a possibility to apply probabilistic limit theorems for transition from deference equations to continuous time stochastic differential equations ([5], [6]). In our paper, we also study a class of copula-based semi-parametric stationary Markov models in a form of scalar difference equation

$$t \in Z : X_t = X_{t-1} + \varepsilon f(X_{t-1}, \varepsilon) + \varepsilon g(X_{t-1}, \varepsilon)\xi_t \quad (1)$$

where $\{\xi_t, t \in Z\}$ is i.i.d., N(0; 1), and ε is a small positive parameter, which will be used for diffusion approximation of (1). Regressions (1) are high-usage equations for simulation and parameter estimation of stochastic volatility models ([5]). Unfortunately, defined by (1) Markov chain has incompact phase space that complicates an application of probabilistic limit theorem. Copula approach helps to simplify asymptotic analysis of (1). Due to persistence of small parameter ε after a substitution $U_t = F(X_t)$ in equation (1) for a further diffusion approximation one can write a difference equation in the same form like (1):

$$t \in Z : U_t = U_{t-1} + \varepsilon f(U_{t-1}, \varepsilon) + \varepsilon g(U_{t-1}, \varepsilon)\xi_t \quad (2)$$

But now this equation defines a Markov chain on the compact [0, 1]. This makes it easier to formulate construction for transition probability and further estimators of functions $\hat{f}(u)$ and $\hat{g}(u)$. After diffusion approximation of (2) one can make a inverse substitution and derive a stochastic differential equation as diffusion approximation for (1).

Described algorithm allows evaluating the parameters of copula, which have the best fit to previously selected model. In our copula dependence study we used MatLab, which helps to evaluate copula parameters and choose the best copula class, based on log likelihood estimation, for the selected financial market data. These copula based models are easy to simulate and can be expressed as semi-parametric regression transformation models. Moreover, using this MatLab we made VIX option index simulation – found the best copula fit under our condition and built semi-parametric autoregression.

III. Proposed Estimation Algorithm and Evaluation of Parameters

Let us take into consideration VIX – Market Volatility Index – daily data from 25 October 2009 to 1 October 2014. The VIX is a market mechanism that measures the 30-day forward implied volatility of the underlying index, the S&P 500. Being able to meaningfully interpret movements in the VIX and its reaction to market events can give investors an edge in managing the risk and profitability of their trading book and in designing portfolio strategies using VIX derivatives to capitalize on the dynamic and time-varying correlation of the VIX with its underlying S&P 500 Index. Let us build for this option index semi-parametric copula based model, using AIC and BIC criteria.

The easiest way of parameter estimating of the semi-regressive model for the VIX index would be to hold the algorithm:

- To simulate U_t points, which are R[0,1] (uniform) or transform the existing sample into R[0,1];
- To build a scatter plot for (U_{t-1}, U_t);
- To make several statistical tests to find the suited distribution of data;
- To take into account a scatter plot and distribution of data, try to choose copula from the existing class or build your own copula, if you know marginal distributions;
- To test copula consistency to data (for example AIB and BIC);
- To find regression parameters.

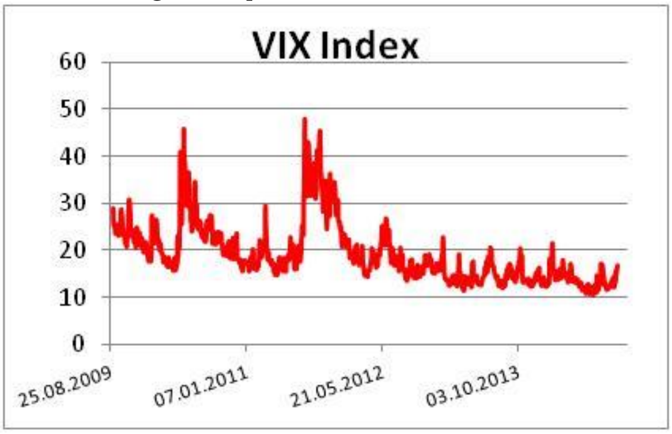

Fig. 1. Historical VIX index levels.

Using Matlab program, we built scatter plots for VIX index transformed into uniform distribution (R[0,1]) and non-transformed data.

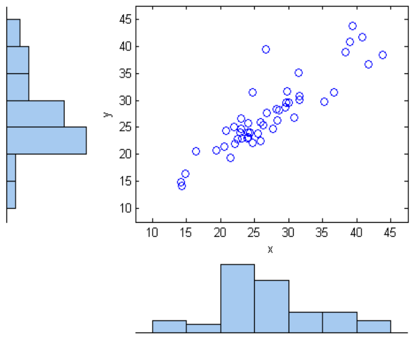

Fig. 2. Scatter plot for non-transformed VIX index data.

An important issue faced by an applied researcher interested in using the class of semi-parametric copula-based time series models is the choice of an appropriate parametric copula [7]. In different papers, Chen et al. (1998) [8] propose two simple tests for the correct specification of a parametric copula in the context of modeling the contemporaneous dependence between several univariate time series and of the innovations of univariate GARCH models used to filter each univariate time series.

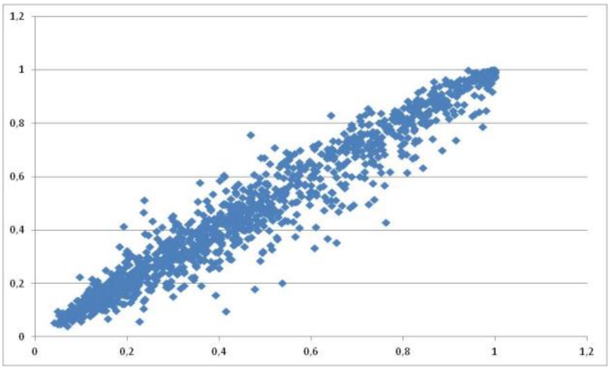

Fig. 3. Scatter plot for transformed into R[0,1] VIX index data.

Chen and Fan (2004b) [9] establish pseudo-likelihood ratio tests for selection of parametric copula models for multivariate i. i. d. observations under copula misspecification [4]. But our suggestion is simpler – we can choose the best copula fit using AIC and BIC criteria or using χ^2 test for data distribution. We take for different copula comparisons AIC and BIC.

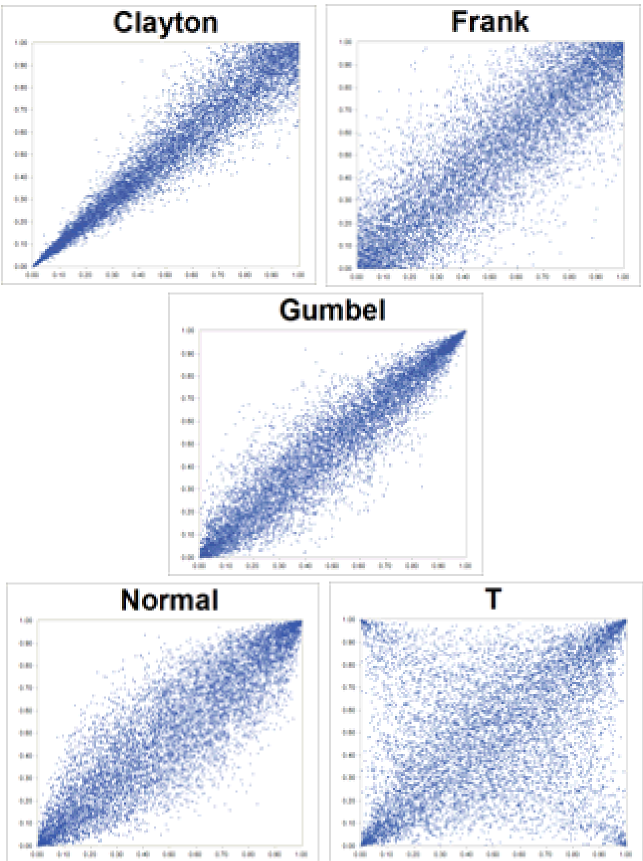

Fig. 4. Most common types of copula in finance (theoretical illustrations).

For the first copula choosing step, it is reasonable to compare graphical parametric copulas with VIX data scatter plot (Fig. 2), (Fig. 3). As we can see, the most suitable copulas for our data are Gumbel, Frank and Normal. For this sample of copulas, is useful to calculate AIC and BIC criteria.

TABLE I

AIC AND BIC CRITERIA FOR VIX INDEX DATA

Copula	AIC	BIC
Gumbel copula	-109.3	-105.7
Frank copula	-255.1	-247.1
Normal copula	-248.9	-242.2

Taking into account AIC and BIC criteria, we should choose the Frank copula for further model estimation. Let us see how to derive semi-parametric regression parameters using Frank copula representation:

$$C(u_{t+1},u_t) = -a^{-1}\ln\left(1+\frac{g_{u_{t+1}}g_{u_t}}{g_1}\right) \qquad (3)$$

$$g_{ut} = e^{-au_t}-1$$

Inserting expression (3) into conditional expectation, we get our parameters:

$$E(U_{t+1}\,|\,U_t=u) = \int_0^1 u_{t+1}\, dF_{u_{t+1}|u_t}(u) = \int_0^1 u_{t+1}\, p(u_{t+1}\,|\,u_t)\, du_{t+1} =$$

$$= \int_0^1 u_{t+1}\frac{\partial C(u_{t+1},u_t)}{\partial u_{t+1}\partial u_t}du_{t+1} = \int_0^1 u_{t+1}\, c(u_{t+1},u_t)\, du_{t+1} =$$

$$= -a\int_0^1 u_{t+1}\frac{g_1\left(1+g_{u_t+u_{t+1}}\right)}{\left(g_{u_t}g_{u_{t+1}}+g_1\right)^2}du_{t+1} =$$

$$= \frac{e^{au_t}\left(\left(e^a-1\right)\left(\ln\left(e^{au_t}\left(e^a-1\right)\right)-\ln\left(e^a\left(e^a-1\right)\right)\right)+a\left(e^a-e^{au_t}\right)\right)}{a\left(e^{au_t}-1\right)\left(e^a-e^{au_t}\right)}$$

(4)

$$\mathrm{g}((U_{t+1}\,|\,U_t=u) = E((U_{t+1}-f(U_t))^2/U_t=u) =$$

$$= \int_0^1 (U_{t+1}-f(U_t))^2 c(U_{t+1},U_t)\,dU_{t+1} =$$

$$= \int_0^1 \frac{1}{t}\left(\alpha\ln(t) - \frac{\alpha\ln(t)(\exp(\alpha)-1)}{(1-\frac{1}{t})(\exp(\alpha)-t)} + \frac{t}{\exp(\alpha)-t}\right)^2 x$$

$$x\frac{(1-\exp(-\alpha))\exp(-\alpha u_t)\frac{1}{t}}{(1-\exp(-\alpha)-(1-\frac{1}{t})(1-\exp(-\alpha u_t)))^2}dt$$

(5)

where $t = \exp(U_{t+1})$

It is impossible to solve analytically (4) and (5) expressions. But numerically it is doable, for example, in the Matlab or Mathematica. For the Frank copula, we can use an inverse function with the aim to return to our base equation (1). Besides integral (5) diverges at 0 point, as a result it is impossible to find solution to (5).

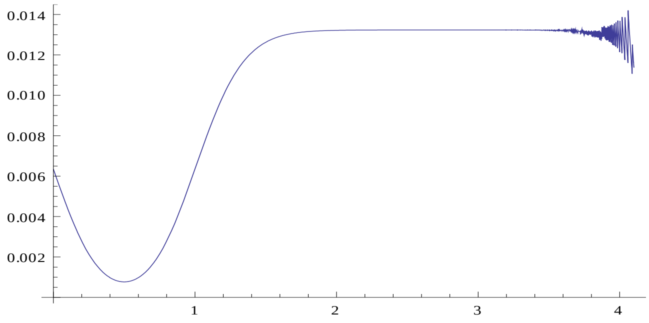

Fig. 5. Second moment of U_t in an interval [0.001; 4] with a parameter α = 8.68.

However, using Mathematica software it is possible to draw values of equation (5). As it is seen in Fig. 4, equation (5) in the interval (0;1) is nonlinear. This second moment adds a nonlinear effect to regression (1). This second moment starts to diverge from point 1, but visual divergence is seen from point 2.7.

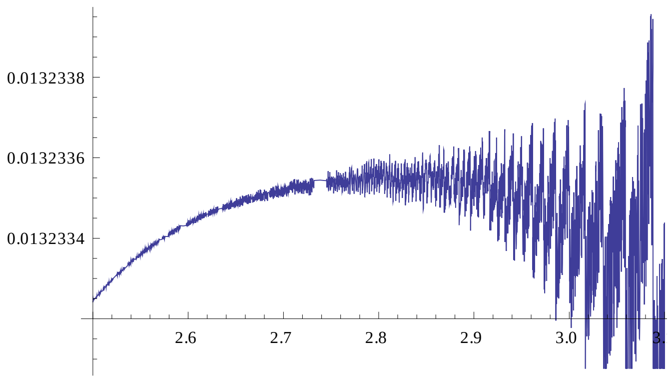

Fig. 6. Second moment of U_t in an interval [0.001; 4] with a parameter α = 8.68.

The final Markov copula based semi-parametric regression model in a form of (2) is equal to the sum of (4) and (6):

$$t \in Z : U_t = U_{t-1} + f(U_{t-1}, \varepsilon) + g(U_{t-1}, \varepsilon)\xi_t =$$

$$= U_{t-1} +$$

$$+ \frac{e^{au_t}\left(\left(e^a - 1\right)\left(\ln\left(e^{au_t}\left(e^a - 1\right)\right) - \ln\left(e^a\left(e^a - 1\right)\right)\right) + a\left(e^a - e^{au_t}\right)\right)}{a\left(e^{au_t} - 1\right)\left(e^a - e^{au_t}\right)} +$$

$$+ \int_0^1 \frac{1}{t}\left(\alpha \ln(t) - \frac{\alpha \ln(t)(\exp(\alpha) - 1)}{(1 - \frac{1}{t})(\exp(\alpha) - t)} + \frac{t}{\exp(\alpha) - t}\right)^2 x$$

$$x \frac{(1 - \exp(-\alpha))\exp(-\alpha u_t)\frac{1}{t}}{(1 - \exp(-\alpha) - (1 - \frac{1}{t})(1 - \exp(-\alpha u_t)))^2} dt\xi_t$$

where $\{\xi_t, t \in Z\}$ is i. i. d., N(0; 1).

But if we deal with copulas, we should not skip some facts. For example, it is not easy to say which parametric copula best fits a given dataset, since some copulas may fit better near the center and other near the tails and many copulas do not have moments that are directly related to the Pearson correlation; it is difficult to compare financial models based on correlation. Of course, if we want to use this model in practice, it is crucial to compare different class models which could be suitable for these data. This can give an applied added value for this method.

IV. Conclusions and Further Work

The algorithm for copula simulation and semi-parametric regression coefficient finding using a Markov chain has been presented. For Option VIX index data using MatLab the best fitted copula model, which is the Frank copula, has been found. According to this copula, principals have been shown and parameters of the semi-parametric regression model coefficients have been evaluated in a copula space. The future research may be devoted to developing numerical algorithms for simulation of trajectories of various random processes, evaluating the characteristic of semi-parametric models based on copula, and finding continuous stochastic models using limit theorems [12].

References

[1] E. W. Frees, E. A, Valde, “Understanding relationships using copulas”. *North American Actuarial Journal* 2, 1998, pp. 1–25. http://dx.doi.org/10.1080/10920277.1998.10595667

[2] P. Embrechts, A. McNeil, D. Straumann, “Correlation and dependence properties in risk management: properties and pitfalls,” in: Dempster, M. (Ed.), *Risk Management: Value at Risk and Beyond*. Cambridge University Press, Cambridge, pp. 176–223, 2002. http://dx.doi.org/10.1017/CBO9780511615337.008

[3] A. Sklar, “Fonctions de repartition an dimensionset leurs marges,” *Publ. Inst. Statis. Univ*. Paris 8,229–231, 1959.

[4] W. Darsow, B. Nguyen, E. Olsen, “Copulas and Markov processes,” *Illinois Journal of Mathematics* 36, 600–642, 1992..

[5] X. Chen, Y. Fan. “Estimation of copula-based semiparametric time series models,” *Journal of Econometrics*, 2006. http://dx.doi.org/10.1016/j.jeconom.2005.03.004

[6] D. B. Nelson, “ARCH models as difusion approximations,” *Journal of Econometrics*, vol. 45, no. 1–2, pp. 7–38, 1990.

[7] Y. Ait-Sahalia, R. Kimmel, “Maximum likelihood estimation of stochastic volatility models,” *Journal of Financial Economics*, 2007. http://dx.doi.org/10.1016/j.jfineco.2005.10.006

[8] X. Chen, L. P. Hansen, M. Carrasco, “Nonlinearity and temporal dependence,” Working Paper, University of Chicago, 1998.

[9] X. Chen, Y. Fan, “Pseudo-likelihood ratio tests for model selection in semiparametric multivariate copula models,” *Canadian Journal of Statistics*, 2004.

[10] H. R. Stoll, R. E. Whaley, “The Dynamics of Stock Index and Stock Index Futures Returns,” *Journal of Financial and Quantitative Analysis*, no. 25, 1990, pp. 441–468. http://dx.doi.org/10.2307/2331010

[11] H. Joe, *Multivariate Models and Dependence Concepts.* Chapman & Hall/CRC, 1997. http://dx.doi.org/10.1201/b13150

[12] J. Carkovs, "On Diffusion Approximation of Discrete Markov Dynamical Systems. In Computational Geometry," *Proceedings of World Academy of Science, Engineering and Technology*, vol. 30 pp. 1–6, 2008.

Jegors Fjodorovs is a Doctoral student. He received the Master Degree in Mathematics from the University of Latvia. He received the Qualification of Mathematician–Statistician from the same university. Since 2011 he has been studying at the doctoral study program at Riga Technical University, the Chair of Probability Theory and Mathematical Statistics. The main fields of research are stochastic differential equations, Markov processes and copulas with application in finance mathematics and econometrics.

Presently he is a Lecturer with Riga Technical University in the course "Random Processes" and a Risk Engineer with JSC Swedbank in the Wealth Management Division. He has also worked for seven years at the State Treasury of the Republic of Latvia, where he was able to manage market risks of government assets/liabilities.

E-mail: Jegors.Fjodorovs@rtu.lv.

Andrejs Matvejevs received the Doctoral Degree in 1989. He graduated from the Faculty of Automation and Computing Technique, Riga Technical University, and became an Associate Professor with Riga Technical University in 2000 and a Full Professor in 2005. He has made the most significant contribution in the field of actuarial mathematics. Andrejs Matvejevs is a Doctor of Technical Sciences in Information Systems. Until 2009 he was a Chief Actuary at the insurance company "BALVA". For more than 25 years he has taught at Riga Technical University and Riga International College of Business Administration, Latvia. His previous research was devoted to solving dynamical systems with random perturbation. His current professional research interests include applications of Markov chains to actuarial technologies: mathematics of finance and security portfolio. He is the author of about 40 scientific publications, two study books and numerous conference papers.

E-mail: Andrejs.Matvejevs@rtu.lv.

Andrejs Matvejevs, Jegors Fjodorovs. VIX indeksa modelēšana, izmantojot neparametriskos Markova modeļus ar Franka kopulu

Šajā rakstā tiek aprakstīts neparametriskās Markova modeļa novērtēšanas algoritms, izmantojot Franka kopulas blīvumu. Uz kopulām bāzētas neparametriskās regresijas atšķiras ar to, ka pētnieks varētu sadalīt dažāda veida efektus (risku avotus) un katru modelēt atsevišķi (neparametriskais marginālais sadalījums un parametriskā kopulas funkcija), bet kopula savieno visu no mēroga neatkarīgo laika sakarību. Rakstā tika izmantots finanšu indekss VIX – mēra 30 dienas uz priekšu netiešu svārstīgumu, bāzes indeksa S & P 500. Šis indekss aprēķināms no akciju opciju cenām. Aprakstītā pieeja ļauj mums novērtēt parametrus Franka kopulai, kura pēc noteiktajiem statistiskajiem kritērijiem ir labākais variants VIX indeksa datu vēsturiskās atkarības aprakstīšanai. Turpmāk, balstoties uz Franka kopulas blīvuma funkciju, tika aprakstīts neparametriskas Markova regresijas koeficientu atrašanas mehānisms. Šī modeļa parametru noteikšana ir sarežģīta – analītiski nav atrisinājuma (parametriskais integrāls diverģē 0 punktā). Tādējādi tika izmantotas MatLab un Mathematica skaitliskās metodes, kas ļauj pārliecināties par pareizo metodoloģisko pieeju – no ilustrētajiem grafikiem var redzēt, ka otrais moments pievieno nelinearitāti aprakstītajam vienādojumam. Rezultātā, izmantojot aprakstīto metodoloģiju, var imitēt VIX indeksa dinamiku dažādiem laika intervāliem un iegūtos novērtējumus izmantot finanšu riska vadībai (riska ierobežošanas operācijas, izmantojot ar opcijām) vai pieņemot spekulatīvas tirdzniecības pozīcijas ar opcijām.

Андрей Матвеев, Егор Фёдоров. Моделирование VIX индекса посредством непараметрических Марковских моделей с копулой Франка

Данная статья описывает алгоритм оценки непараметрической Марковской модели с помощью плотности копулы Франка. Копульные непараметрические регрессии отличаются тем, что исследователь может разделить различные виды (источники) риска, каждый смоделировать отдельно (непараметрические маргинальные распределения и параметрическая копульная функция) и соединить копулой, свободной от маштаба временной зависимостью. В статье был использован финансовый индекс VIX, измеряющий 30-дневную будущую внутреннюю волатильность на основе индекса акций S & P 500. Этот индекс рассчитывается, исходя из цен опционов. Описанный подход позволяет оценить параметры копулы Франка, правильность выбора которой устанавливается с помощью статистических критериев и является лучшим для данных индекса VIX. То есть эта копула лучше остальных копул описывает историческую зависимость. Далее, на основе функции плотности Франка копулы, был описан механизм оценки коэффициентов непараметрической Марковской регрессии. Такая оценка параметров трудоёмка – нет аналитического решения (параметрический интеграл расходится в точке 0). Таким образом, вычисление параметров происходит с использованием численных методов в пакетах Matlab и Mathematica. Проверить правильность подхода позволяют графические иллюстрации, где можно видеть, что второй момент, добавленный к уравнению, является нелинейным. В результате, используя описанную методологию, можно имитировать индекс VIX в разные промежутки времени и полученные результаты использовать в управлении финансовыми рисками (операции хеджирования через опционы) или принятии спекулятивных торговых позиций с опционами.