THE EFFECT OF COPULA ON SCENARIO TREE STRUCTURE

Henrikas Pranevicius, Kristina Sutiene
Department of Business Informatics
Kaunas University of Technology
Studentu str. 56-301, 51424 Kaunas, Lithuania
E-mail: kristina.sutiene@stud.ktu.lt

The multistage stochastic programs can be solved if the underlying stochastic process is approximated by a scenario tree. The correlation matrix is usually used to describe the dependence between the stochastic variables in a multivariate structure. Due to assumptions of using the Pearson's correlation coefficient, the usefulness of such correlations is restricted. In statistics and finance the alternative method – copula-based dependency measure (Embrechts et al. 2002, Aas 2004) – has been used for some time. The main advantage of employing copulas is the possibility to model the nonlinear dependencies between non-elliptically distributed stochastic variables. To our knowledge, the copulas still are not very popular in generation of scenario trees. According to this, we propose to approximate the multivariate stochastic process by a scenario fan with multivariate structure using copulas. Then, the form of scenario tree is constructed out of individual scenarios using a cluster analysis. In the analysis, some copulas with different features are employed. While investigating the copula effect on the scenario tree structure, at least theoretically, we consider the question of whether the copula features are captured in the approximate representation of the underlying stochastic process in the form of scenario tree. Such generation of scenario tree can be useful in cases when it is difficult to construct the adequate scenario tree from the stochastic differential equations or time-series models; and the sampled paths can be obtained by sampling or resampling techniques, using copulas as a dependence measure. The discussed way of scenario tree generation is explained below in more details.

Stochastic programming models. The concept of scenarios is usually employed for the modeling of randomness in stochastic programming models, in which data evolve over time and decisions have to be made independent upon knowing the actual paths that will occur. Such data are usually subject to uncertainty or some kind of risk. For instance, the random variables are the return values of each asset on an investment in portfolio management problems, and the investment decisions must be implemented before the asset performance can be observed. Each scenario can be viewed as one realization of an underlying multivariate stochastic data process. The modeling of randomness employees the set of available past data with the aim of building sub-models for each individual stochastic parameter. These sub-models are then used to generate a set of scenarios that encapsulate the consistent depictions of pathways to possible futures based on assumptions about economic and technological developments. Thus, the factors driving the risky events are approximated by a discrete set of scenarios, or sequence of events.

Scenario generation. A good approximation of the underlying stochastic process may involve a very large number of scenarios and their probabilities. This process is known as scenario generation. There are a lot of scenario generation methods, see for example (Dupačová et al. 2002; Pflug 2001; Yu et al. 2003). They are based on different principles: conditional sampling, moment matching, path based methods, optimal discretization. A better accuracy of uncertainties is described when scenarios are constructed via a simulated data path structure, also named as scenario fan (Hibiki 2003) (see Figure 1).

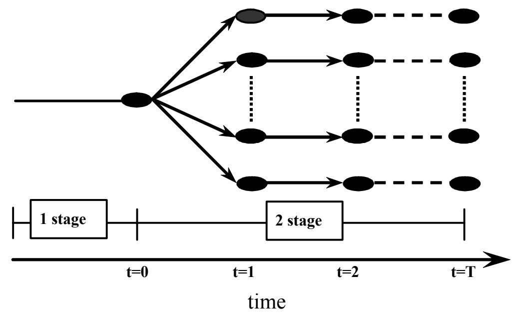

Figure 1. Scenario fan

Scenario tree construction by K-means clustering. According to the complexity of stochastic programming model, the scenario tree structure (see Figure 2) is used as an input to the stochastic programming model. Due to this, in this paper we propose to generate a scenario tree out of individual scenarios fan, using the cluster analysis. An approach similar to our work is introduced in the article (Dupačová et al. 2000), but without a detailed clustering algorithm. According to this, we propose the strategy to generate multistage scenario tree from the set of individual scenarios by bundling scenarios based on the K-means clustering method (Kaufmann and Rousseeuw 1990). This method was modified to cluster the data paths, capturing the inter-stage dependence, in order to model the sequential decisions. In the Figure 3, the idea of bundling the scenarios to the clusters is depicted. The fan of 11 scenarios is schematically illustrated. At time $t = 0$ all these scenarios (which are the same) form the root of the tree. Next, two clusters are

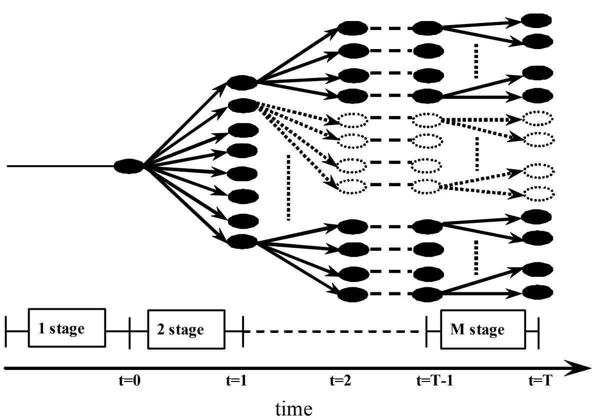

Figure 2. Multistage scenario tree

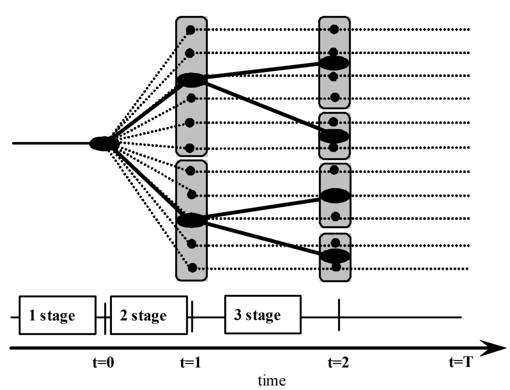

Figure 3: Illustration of 3-stage Tree Construction

formed by the first iteration of proposed clustering algorithm. It results that we have six and five scenarios in each cluster. The centers of each cluster are computed, which represent the one-level nodes at time $t = 1$. Two black points denote the nodes corresponding to the conditional decisions. The formed clusters are then divided into sub-clusters in the next time period $t = 2$. We have four, two, three and two paths in each cluster, representing two-level nodes, since the centers are calculated. These nodes are denoted by four black points in the scenario tree. Such strategy of bundling scenarios to the clusters continues till the end of time horizon is reached. Joining the black points by line, the scenario tree structure is obtained.

The numerical example. In the paper (Pranevicius and Sutiene 2006) the effect of using two copulas – Gaussian and Student's t_2 – as the dependence measure between real interest rate and inflation rate was investigated on the scenario fan of asset returns. The authors employed the multivariate Hibbert, Mowbray and Turnbull (HMT) stochastic asset model for long-term financial planning (Hibbert et al. 2001), where the nominal interest rates are modeled as a function of both real interest rates and inflation rates. In this paper the proposed scenario tree generation based on HMT model will be demonstrated. The effect of these two copulas – Gaussian and Student's t_2 – will be investigated on the scenario tree structure of nominal interest rate.

References

Embrechts, P.; McNeil, A. and Straumann, D. 2002. „Correlation and Dependency in Risk Management: Properties and Pitfalls". In *Risk management: value at risk and beyond,* M.A.H. Dempster (Eds.), 176–224.

Aas, K. 2004. Modeling the Dependence Structure of Financial Assets: a Survey of Four Copulas, *Norwegian Computing Center: Applied Research and Development*, 1-19.

Hibiki, N. 2003. "A Hybrid Simulation / Tree Stochastic Optimization Model for Dynamic Asset Allocation". In *Asset and Liability Management Tools*, B. Scherer, (Eds.)*,* 279–304.

Dupačová, J.; Hurt, J. and Štěpán, J. 2002. *Applied Optimization 75: Stochastic Modeling in Economics and Finance.* Kluwer Academic Publishers.

Pflug, G. 2001. "Scenario Tree Generation for Multiperiod Financial Optimization by Optimal Discretization." *Mathematical Programming B89,* 251–271.

Yu, L.Y.; Ji, X.D. and Wang, S.Y. 2003. "Stochastic Programming Models in Financial Optimization: Survey." *AMO – Advanced Modeling and Optimization 5(1),* 1–26.

Dupačová, J.; Consigli, G. and Wallace, S.W. 2000. "Scenarios for Multistage Stochastic Programs." *Annals of* Operations *Research 100,* 25–53.

Kaufmann, L. and Rousseeuw, P.J. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis.* John Wiley & Sons.

Pranevicius, H. and Sutiene, K. 2006. "Simulation of Dependence between Asset Returns in Insurance". In *Proceedings of the 5th International Conference on Operational Research: Simulation and Optimization in Business and Industry.* Tallinn, Estonia, 23–28.

Hibbert, J.; Mowbray, P. and Turnbull, C. 2001. "A stochastic Asset Model & Calibration for Long-Term Financial Planning Purposes". Technical Report. Barrie&Hibbert Limited.