sciendo

Exchangeable models of financial correlations matrices. Bayesian nonparametric models and network derived measures of financial assets

Sorin OPINCARIU
Bucharest University of Economic studies, Bucharest Romania
sopincariu@live.com

Stefan IONESCU
Romanian American University, Bucharest Romania
stefan.ionescu@profesor.rau.ro

Abstract. *De Finetti theorem establishes the conceptual basis of Bayesian inference replacing the independent and identically distributed sampling hypothesis prevalent in frequentist statistics with the much easier to justify in practical settings hypothesis of exchangeability. In this paper we make use of the extension of the concept of exchangeability from sequences to arrays arguing that the invariance to ordering is a much more tenable assumption than independent and identically distributed sampling in the financial modeling problems. Making use of the celebrated Aldous-Hoover representation theorem of exchangeable matrix we construct a Bayesian non-parametric model of the financial returns correlation matrices arguing that a Bayesian approach can mitigate many of the known shortcomings of the usual Pearson correlation coefficient. We posit the correlation matrix to be an exchangeable matrix and construct a Bayesian neural network to estimate the functions from the Aldous-Hoover representation theorem. The correlation matrix model is coupled with a Student-t likelihood (accounting for the heavy tails of financial returns). The model is estimated with a Hamiltonian Monte Carlo sampler. The samples are used to construct an ensemble of networks where each edge is weighted by the size of the correlation between two financial instruments. Various centrality measures are being calculated (betweenness, eigenvector) for each network of the ensemble allowing us to obtain a probabilistic view of each financial instrument's importance. We also construct a minimum spanning tree associated with the mean correlation matrix allowing us to visualize the most important financial instruments from the universe selected.*

Keywords: Exchangeability, Correlation, Bayesian neural networks, Centrality, Networks.

Introduction

Modeling the financial dependence across assets is one of the corner stones of quantitative approach in finance. Failing to account for dependence between financial assets or using poorly specified models with unrealistic assumptions can lead to disastrous consequences as exemplified by the subprime crisis of 2007-2009. The collateralized debt obligations (CDO) were securitized instruments which contained pools of mortgages, and their value was calculated using poorly fit correlation models. The pricing of the CDOs was made by having an overly optimistic assumption that the correlation between mortgages defaults was low. When the housing market started to fall, the number of mortgage defaults was higher than was than predicted by the valuation models leading to the collapse of the CDO market and of the banking organizations that held those instruments in their portfolios.

In this paper we argue for a principled approach in the modeling of the dependence between financial assets taking into account the uncertainty that is inherent to every modeling endeavor. We consider that the Bayesian approach to modeling dependence between financial instruments is well suited to the particularities of financial data.

Bayesian framework lays bare all the modeling assumptions about the data generating process (in the form of likelihood function) about the plausible values of the parameters (in the form of the prior distribution of the parameters) and about the uncertainties in the measurement process (the prior distribution of the data). Having all the assumptions laid out in a probabilistic form, forces the model's stakeholders to always take into consideration the model limits. The output of a Bayesian model is again presented in the form of a probability distribution (posterior) forcing again the stakeholder to consider the uncertainty of the results. This is where Bayesian approach departs the frequentist philosophy: in the Bayesian approach one does not reject or fail to reject a hypothesis, in the Bayesian approach one only updates one's beliefs in one alternative or the other.

In this paper we construct a probabilistic model of pairwise correlations between financial assets returns as general as possible, a model that takes into account the extreme complexity and its associated uncertainty of financial instruments evolution. We placed our approach under the framework of exchangeable sequences of random variables pioneered by de Finetti (De Finetti, 2017) and extended to the case of exchangeable arrays by Aldous and Hoover (Aldous, 1981 and Hoover, 1979). The representation theorem of Aldous and Hoover, discussed below, establishes the necessary and sufficient conditions for an array to be exchangeable. As we consider exchangeability of correlation matrices (invariance to permutations) as a safe modeling assumption, we use the representation theorem to build a Bayesian model of correlation matrices.

The Aldous and Hoover theorem only proves the existence of a representation of an exchangeable random matrix as a function of independently and identically distributed (i.i.d.) sequence of uniform random variables without offering a constructive method of finding that representation. Being universal approximators (Hornik et al, 1989), we use Bayesian models of neural networks to estimate the distribution of the correlation matrix from data. We used a Hamiltonian Monte Carlo sampler to estimate our model from data.

The Bayesian model we specify, results in a distribution of probability over the space of correlation matrices. We decided that listing the correlation matrices is simply not very useful even for moderately large matrices as it obscures the relationships between various components of the matrices. We express the correlation matrix as a weighted network, and we calculate various measures of nodes (financial instruments) importance. The utility of Bayesian approach becomes visible in this context, as we can derive with minimum effort a probability distribution of various centrality measures based on samples from the posterior distribution of correlation matrices. We argue that due to the concentration of measure in high dimensional spaces (i.e. most of the probability mass is concentrated around the mean), the mean is the most informative summary of the posterior distribution of the correlation matrix. We build a representation of the mean correlation matrix under the form of a minimum spanning tree (MST). We will also show how one can intuitively visualize these probability distributions by making use of boxplots.

In the conclusion section we discuss the importance of the results obtained and trace out paths for future work.

Literature review

There is a vast literature dedicated to modeling of financial correlation matrices. When modeling large random matrices, it was observed that sometimes one observes larger fluctuations from the mean value of the components than it was observed in the corresponding one-dimensional distribution. Following this observation, the econophysics literature (Laloux, 1999; Plerou, 1999; Sengupta, 1999) have proposed using the tools of random matrix theory to filter the empirical correlation matrices.

A different avenue of modeling correlation matrices is through the use of network theory (Mantenga, 1999; Kumar, 2012; Fieodor, 2014; Onnela, 2013; Onnela, 2014; Tumminello, 2010; Keneti, 2013; Namki, 2011). Network theoretic approach sets out to uncover the topological structure of dependence which is deemed to be more resilient to spurious relationships.

The heavy tail property of returns distribution has been widely documented in the econometrics theory (Peiro, 1999; Cont, 2010). These authors argue for a departure from the gaussian distribution in modeling financial returns.

The use of Aldous-Hoover theorem in the context of exchangeable models of random arrays and graphs can be found in (Orbanz et al, 2015). The application of Aldous-Hoover theorem in a machine learning setting can be found in (Loyd et al, 2012) where a gaussian process prior is used to infer the form of the graph generating function. Bayesian nonparametric models have also been proposed in the context of missing link prediction in large networks (Wolfe et al, 2013; Miller et al, 2009).

Exchangeable random arrays. Aldous Hoover representation

Following (Orbanz, 2013), we define an infinite sequence (X_i) of random variables as exchangeable if the joint distribution of the sequence is identical with the joint distribution of any permutations of the sequence:

$$P(X_1 \in A_1 \cdots) = P\left(X_{\sigma(1)} \in A_1 \cdots\right)$$

for every permutation σ.

In other words, an exchangeable sequence is invariant to permutation.

The definition of exchangeability extends this principle of symmetry to random sequences of 2 indices $\left(X_{ij}\right)$. We have the following definition of an exchangeable array:

Definition. A random 2-array $\left(X_{ij}\right)$ is called jointly exchangeable if:

$$\left(X_{ij}\right) =_d \left(X_{\sigma(i)\sigma(j)}\right)$$

The joint exchangeability implies the random matrix is invariant to the simultaneous reordering of columns and rows of the matrix that preserves pairwise relationships.

The Aldous-Hoover representation theorem (Aldous, 1981 and Hoover, 1979) establishes the equivalence between the exchangeable matrices and functions of sequences of i.i.d uniform random variables. The Aldous Hoover theorem states that:

Theorem: A random array $\left(X_{ij}\right)$ is jointly exchangeable if and only if it can be represented:

$$\left(X_{ij}\right) =_d \left(F\left(U_i, U_j, U_{ij}\right)\right)$$

where $F: [0,1]^3 \to X$ and (U_i) and $\left(U_{ij}\right)$ are a sequence respectively an array of i.i.d of uniform random variables independent of F (the equality above is in distribution).

Bayesian model of financial correlations

Pearson correlation coefficient for measuring financial correlations

In this section we lay out the assumptions and the principles of constructing a Bayesian model of financial assets dependence.

Traditionally the correlation between the financial assets returns is measured by Pearson correlation coefficient. The Pearson correlation coefficient for the random variables X and Y is defined as:

$$\rho_{X,Y} = \frac{E\big[\big(X - E(X)\big)\big(Y - E(Y)\big)\big]}{\sigma_X \sigma_Y}$$

where σ is the standard deviation of the two random variables.

The Pearson correlation coefficient arises thus from a linear regression setting with normally distributed errors and it measures the strength of linear dependence between two random variables.

One of the disadvantages of the Pearson correlation coefficient is that it assumes both linearity of dependence and the normality of the random variables. The vast empirical research on financial returns (see, for example, Cont, 2001) has proven the absence of linear correlations (the efficient markets arbitrage away easy to trade linear relationships) and the non-gaussian character (the heavy tails are driven by asymmetry profit/loss).

The Pearson coefficient not only obscures from the model stakeholder (usually a financial markets professional not very proficient in statistics) the assumptions of linearity and gaussianity of financial returns distribution but it also obscure from the same stakeholder the uncertainty in the estimation of the coefficient.

We consider that the classical significance tests as a tool to ascertain the confidence of the correlation coefficients estimations are inapplicable in the financial setting because the non-stationarity of financial returns invalidates the assumption of i.i.d sampling.

Bayesian model of exchangeable financial correlations

The Bayesian model we propose, takes into account the stylized facts of financial returns: non-gaussian character, absence of linear correlations and non-stationarity. The return of a financial asset is defined as the difference of the logarithm of the price of that asset. The logarithmic return thus defined is an approximation of the percentual change in the price of that asset.

Our model likelihood function (data generating process) is a multivariate Student-t distribution. The multivariate Student-t distribution has the density function:

$$f(x|\nu, \mu\Sigma) = \frac{\Gamma\big((\nu + p)/2\big)}{\Gamma(\nu/2)\nu^{p/2}|\Sigma|^{1/2}\pi^{p/2}} \left[\left(1 + \frac{1}{\upsilon}(x - \mu)^T \Sigma^{-1} (x - \mu)\right)\right]^{-(\upsilon+p)/2}$$

where Σ is the covariance matrix, μ is the mean vector, ν is the degrees of freedom and p the dimensionality of the random vector x.

The degrees of freedom ν controls the heavy tailness of the distribution. In the limit of infinite degrees of freedom the multivariate t distribution converges to the multivariate normal distribution. We will attach a flat (uniform) prior on the degrees of distribution ν quantifying our complete ignorance about them. The uniform distribution is the maximum entropy distribution (maximum ignorance) on the real line.

For the mean vector μ we specify the maximum entropy distribution with a known empirical mean (financial returns have an empirical mean closed to 0) and standard deviation of 1 that is the standard normal distribution $N(0,1)$.

The prior for the covariance matrix $\sum$ will be factorized into a prior for variances (a standard normal distribution restricted to the positive real line $N^+(0,1)$) and a prior for the correlation matrix $Corr$.

The main assumption of this paper is that correlation matrices are *exchangeable.* In our opinion this assumption is sensible and safe. No matter how we change de order of the financial instruments in the matrix, the elements of the matrix refer to the same pairwise dependence of financial assets. Note that the exchangeability assumption is much weaker than the i.i.d assumption, and it only has to do with our method of construction of the correlation matrix itself rather than referring to the objective properties of return distributions.

We will specify a prior for the correlation matrix in a nonparametric fashion making use of the Aldous Hoover theorem. The Aldous Hoover theorem indicates that the problem of finding a prior for an exchangeable matrix can be translated into a problem of finding a function of uniformly distributed sequences of random variables. We will define a Bayesian model of neural network as a way to approximate the function F of the Aldous Hoover theorem.

Bayesian neural networks consist of layers of affine transformations passed through a nonlinearity, on whose parameters a prior distribution has been specified. For more details on Bayesian learning of neural networks we refer the reader to (Neal, 2012). Following Neal we propose the following model of Bayesian neural network to specify the function F of the Aldous Hoover theorem:

$$U_i, U_{ij} \sim_{iid} Uniform([0,1])$$
$$W_i \sim N(0,1)$$
$$act_0 = tanh(W_i U)$$
$$W_o \sim N(0,1)$$
$$Corr = F((U_i, U_j, U_{ij})) = tanh(W_{oi} act_0)$$

where W_i and W_o are the parameters of the network.

We are now in position to specify the fully hierarchical Bayesian model of financial returns as:

$$Corr = F(U_i, U_j, U_{ij})$$
$$\sigma_i \sim N^+(0,1)$$
$$\Sigma = diag(\sigma) \cdot Corr \cdot diag(\sigma)$$
$$\mu \sim N(0,1)$$
$$\nu \sim Uniform([1,35])$$
$$return \sim Student - t(\mu, \Sigma, \nu)$$

The above model makes little assumptions about the plausible values of the parameters. In specifying the hierarchical model; we made use as much as possible of maximum entropy distributions so that we do not bias the results with our prior beliefs.

Note that even though our primary subject of interest is the correlation matrix, we model the full multivariate distribution of the financial returns.

Network representation of the Bayesian correlation matrix

The inference of the hierarchical model specified in the previous section consists of estimating the posterior distribution of all the parameters of the models. Of all the parameters we focus our attention on $Corr$ as a measure of pairwise financial returns dependence.

The posterior probability distribution of the correlation matrix is given as a sequence of samples obtained through the application of a Hamiltonian Monte Carlo samples. Having a large number of n by n matrices is not a very illuminating result of the analysis. Even for a single estimation of the correlation matrix, for a medium sized financial instruments universe visualizing the correlation matrix will not reveal any meaningful relationships.:

To create a more informative view of the posterior distribution we create an ensemble of networks out of an ensemble of correlation matrices.

A network is defined as a tuple (V, E) of vertices and edges between those vertices. The canonical representation of a network is through its adjacency matrix A in which $A_{ij} = 1$ if there is an edge between nodes i and j and 0 otherwise. Using the correlation matrix to bootstrap a weighted network has a vast literature dedicated to it (see for example Kumar, 2012; Fiedor, 2014; Onnela, 2013; Onnela, 2014; Tumminello, 2010; Keneti, 2013; Namki, 2011). To each non-zero entry in the correlation matrix we associate a link between the corresponding two nodes (financial instruments) with an associated weight W_{ij} equal to the correlation coefficient between nodes i and j.

Measures of centrality of the correlation matrix

Having converted the correlation matrix to a network representation will not advance in itself the analysis. The resulted network is densely connected (every node is connected to every other node) which does not allow visualize important dependences between nodes.

An apparent solution would be to prune de edges below a certain threshold. However, this solution seems arbitrary. Another solution is to define measures of node importance taking into account the whole set of links and their associated weights. Of the measures of centrality presented in the literature we decided to focus on: *eigenvector* and *betweenness* centrality (for a more detailed discussion of the measures of centrality we refer the reader to Newman, 2010).

The simplest measure of a node importance is the degree centrality measuring the number of link that a node has. However, for networks elicited from the correlation matrices this measure is superfluous as every node is connected to every other node. Another possibility is to define the centrality of a node as proportional to the weighted sum of centralities of the other nodes the node is connected to:

$$c_j = \lambda \sum_i W_i c_i$$

Hence the centrality c is an eigenvector of the weight (correlation) matrix W. The eigenvector centrality measures not only the number of connections a node has but also the importance of the nodes the node is connected to. In the context of financial markets eigenvector centrality measures the influence a financial instrument has in the market.

The betweenness centrality measures the numbers of shortest paths a given node is part of. The betweenness centrality measures a node importance as an ”Influence broker”. In the contexts of financial assets, a high betweenness centrality translates into he node being a bridge between different sectors of the market.

A last network theoretic construct that we discuss here is the minimum spanning tree. The use of minimum spanning trees in financial correlation graphs was pioneered by Onnela, et al. 2013. A minimum spanning tree (MST) is a connected acyclic graph with exactly n-1 edges, of minimum cost measured by the edges' weights. An MST has the minimum number of edges such that the graph is still connected. Nodes with a large number of links in an MST are important in transmission of information in the network. The more connected a node is in an MST, the more that node can project influence on the other nodes of the network.

Experimental results.

We fitted the hierarchical Bayesian model presented in this paper to a number of 16 currency pairs over the course of 2 years (January 2017 December 2018).

The inference was performed using a variant of the Hamiltonian Monte Carlo sampler called No U Turn sampler (see Hoffman, 2014). The reason we choose the Monte Carlo simulation over variational inference methods has to do with dependence introduced in the matrix $Corr$ by the sequences U, making it impossible to use mean field approximation required by the variational models.

For each of the samples we calculated the eigenvector centrality. In Figure 1 we present the boxplot of the distribution of probability of centralities for each asset. The horizontal line in the boxplot chart represents the mean of the distribution while the box represents the interquartile interval (Q1-Q3). The mean eigenvector centralities are presented in Table 1. We notice the high eigen vector centrality of the EUR currency and the surprisingly small centrality of GBP and JPY.

Table 1. Mean eigenvector centrality

DKK	**EUR**	**CZK**	**HUF**	**RON**	**PLN**	**SEK**	**NOK**	**CHF**
0.335	0.334	0.325	0.324	0.319	0.315	0.259	0.250	0.238
GBP	**AUD**	**NZD**	**ZAR**	**CAD**	**JPY**	**TRY**		
0.193	0.19	0.176	0.148	0.144	0.129	0.099		

Source: Authors' own research.

The betweenness centrality is plotted in Figure 2 for the 5 assets that are non-zero. We notice that high values of betweenness centrality of the heavily traded emerging market currencies (ZAR South African rand, TRY Turkish lira) and of the carry currencies (JPY Japanese yen, CHF Swiss franc, GBP Great Britain pound). We conclude that these pairs act as bridges between developed and emerging markets.

The minimum spanning tree of the mean correlation matrix is presented in Figure 3. Following Ledoux (Ledoux, 2013) we argue that in high dimensional spaces (16 dimensions in our case) the probability mass of the distribution is concentrated around the mean. The concentration of measure phenomenon makes the mean relevant for the whole probability distribution. If we analyze the Figure 1 we will notice the concentration of measure phenomenon (very narrow boxes).

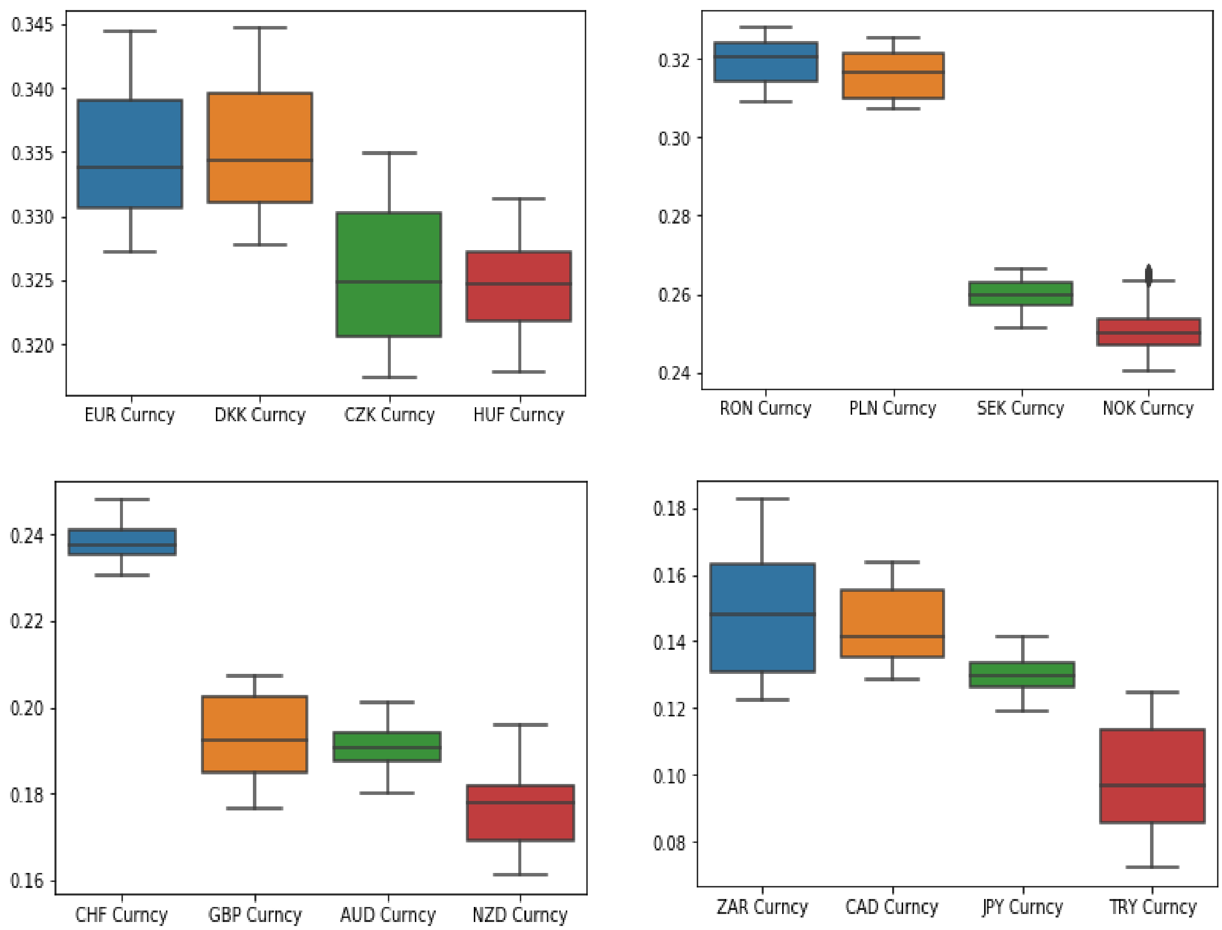

Figure 1. Eigenvector centrality. Boxplot
Source: Authors research

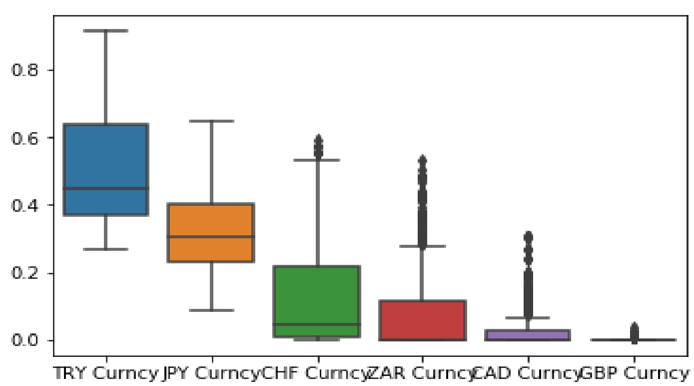

Figure 2. Betweenness centrality. Boxplot
Source: Authors research

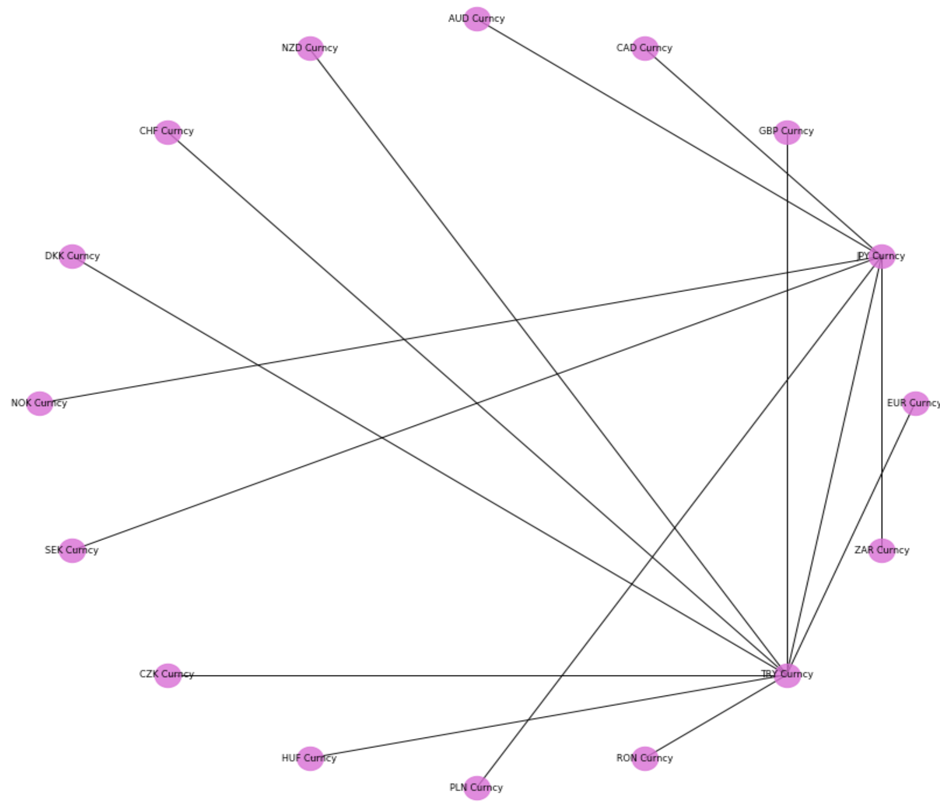

Figure 3. Minimum spanning tree
Source: Authors research

Conclusion

In this paper we presented a general Bayesian approach in modeling and measuring pairwise multivariate correlation between assets. We showed how by using the Aldous-Hoover theorem and nonparametric priors on function spaces, one can specify very flexible and general Bayesian models of correlation matrices.

For future work we set up to relax the exchangeability assumption allowing one to elaborate more general dependent Bayesian model to account for temporal evolution of correlation matrices.

References

Aldous, D. J. (1981). Representations for partially exchangeable arrays of random variables. *Journal of Multivariate Analysis*, *11*(4), 581-598.
Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues.
De Finetti, B. (2017). *Theory of probability: A critical introductory treatment* (Vol. 6). John Wiley & Sons.Aldag, R.J., & Stearns, T.M. (1991). *Management* (2nd ed.). Cincinnati, OH: South-Western Publishing.
Fiedor, P. (2014). Networks in financial markets based on the mutual information rate. *Physical Review E*, *89*(5), 052801.
Hoover, D. N. (1979). Relations on probability spaces and arrays of random variables. *Preprint, Institute for Advanced Study, Princeton, NJ*, *2*.'

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15*(1), 1593-1623.
Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural networks*, *2*(5), 359-366.
Kenett, D. Y., Huang, X., Vodenska, I., Havlin, S., & Stanley, H. E. (2015). Partial correlation analysis: Applications for financial markets. *Quantitative Finance*, *15*(4), 569-578.
Kumar, S., & Deo, N. (2012). Correlation and network analysis of global financial indices. *Physical Review E*, *86*(2), 026101.
Laloux, L., Cizeau, P., Bouchaud, J. P., & Potters, M. (1999). Noise dressing of financial correlation matrices. *Physical review letters*, *83*(7), 1467.
Ledoux, M. (2001). *The concentration of measure phenomenon* (No. 89). American Mathematical Soc..
Lloyd, J., Orbanz, P., Ghahramani, Z., & Roy, D. M. (2012). Random function priors for exchangeable arrays with applications to graphs and relational data. In *Advances in Neural Information Processing Systems* (pp. 998-1006).
Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, *11*(1), 193-197.
Miller, K., Jordan, M. I., & Griffiths, T. L. (2009). Nonparametric latent feature models for link prediction. In *Advances in neural information processing systems* (pp. 1276-1284).
Namaki, A., Shirazi, A. H., Raei, R., & Jafari, G. R. (2011). Network analysis of a financial market based on genuine correlation and threshold method. *Physica A: Statistical Mechanics and its Applications*, *390*(21-22), 3835-3841.
Neal, R. M. (2012). *Bayesian learning for neural networks* (Vol. 118). Springer Science & Business Media.
Newman, M. (2010). *Networks: an introduction*. Oxford university press.
Onnela, J. P., Kaski, K., & Kertész, J. (2004). Clustering and information in correlation based financial networks. *The European Physical Journal B*, *38*(2), 353-362.
Onnela, Jukka-Pekka, Anirban Chakraborti, Kimmo Kaski, Janos Kertesz, and Antti Kanto. "Asset trees and asset graphs in financial markets." *Physica Scripta* 2003, no. T106 (2003): 48.
Orbanz, P., & Roy, D. M. (2015). Bayesian models of graphs, arrays and other exchangeable random structures. *IEEE transactions on pattern analysis and machine intelligence*, *37*(2), 437-461.
Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., & Stanley, H. E. (1999). Universal and nonuniversal properties of cross correlations in financial time series. *Physical review letters*, *83*(7), 1471.
Peiro, A. (1999). Skewness in financial returns. *Journal of Banking & Finance*, *23*(6), 847-862.
Sengupta, A. M., & Mitra, P. P. (1999). Distributions of singular values for some random matrices. *Physical Review E*, *60*(3), 3389.
Tumminello, M., Lillo, F., & Mantegna, R. N. (2010). Correlation, hierarchies, and networks in financial markets. *Journal of Economic Behavior & Organization*, *75*(1), 40-58.
Wolfe, Patrick J., and Sofia C. Olhede. "Nonparametric graphon estimation." *arXiv preprint arXiv:1309.5936* (2013)