Model Selection in Weighted Stochastic Block models

Selezione del modello per i modelli a blocchi stocastici

Roberto Casarin, Michele Costola, Erdem Yenerdag

Abstract We propose the weighted stochastic block model (WSBM) as generative model for the financial networks and exploit the topological features of its blocks. This model considers both the edge existence and the edge weight of the network and is independent from the methodology implemented on the estimation of the network. In this paper, we discuss three specifications of the model with by analysing the European financial network.

Abstract In questo lavoro, viene proposto il modello a blocchi stocastici pesato (WSBM) come modello generativo per i network finanziari dove vengono esplorate le caratteristiche topologiche dei blocchi. Questo modello considera sia il peso che l'esistenza del legame ed é indipendente dalla metodologia di stima del network. In questo articolo, discutiamo tre specifiche del modello analizzando il network finanziario Europeo.

Key words: financial networks, stochastic block model

1 Introduction

The study of the financial network topologies and connectedness provides useful tools in monitoring financial stability.

Roberto Casarin
Department of Economics, Ca' Foscari University of Venice, Dorsoduro 3246, 30123 Venice (Italy). e-mail: r.casarin@unive.it

Michele Costola
SAFE, House of Finance, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 3, 60323 Frankfurt am Main (Germany). e-mail: costola@safe.uni-frankfurt.de

Erdem Yenerdag
Department of Economics, Washington University in St. Louis, One Brookings Drive Campus Box 1208, Saint Louis, MO 63130 (USA). e-mail: erdemyenerdag@wustl.edu

Nodes exhibit often the propensity to be clustered into groups which are internally densely connected but show fewer connections outside. This feature is known as community structure or network modularity [Newman, 2010]. The aim of this paper is to investigate the network topology with focus on the detection of blocks in a broader sense where nodes represents financial assets. The communities are obtained through the Weighted Stochastic Block Model (WSBM) introduced recently by [Aicher et al., 2014] which allows through the blocks to have a compact characterization of the network's structure. This network generative model considers both the edge existence and the edge weight of the network and represents a generalization of the SBM [Holland et al., 1983] which is independent from the methodology implemented on the estimation of the network. By analysing the European financial institutions, we compare three specifications of the model such as the Stochastic Block Model (SBM), the balanced Weighted Stochastic Block Model (WSBM) and the pure WSBM (pWSBM).

2 Weighted Stochastic Block Models

The likelihood function of the basic Stochastic Block Model (SBM) is:

$$\mathscr{L}(A|z,\theta) = \prod_{ij} \theta_{z_i z_j}^{A_{ij}} (1-\theta_{z_i z_j})^{1-A_{ij}} \tag{1}$$

where A is an adjacency matrix of the network, which contains binary values representing edge existence, $A_{ij} \in \{0,1\}$, z is a vector that contains the group label of each node $z_i \in \{1,\ldots,K\}$, where K is the number of latent groups. For example, if A is an $n \times n$ adjacency matrix then z is a $(1 \times n)$ vector such as $z = (z_1, z_2, \ldots, z_i, \ldots, z_n)$. Therefore, vector z represents the partition of the nodes into K blocks and each pair of groups kk' represents a bundle of edge between the groups. The parameter θ in Equation 1 represents a $(K \times K)$ matrix and its elements represent the edge existence parameters, $\theta_{z_i z_j}$, of each edge bundle. The existence probability of an edge A_{ij} is given by the parameter $\theta_{z_i z_j}$ that depends only the membership of nodes i and j. In Equation 1 A_{ij} is conditionally independent given z and θ. Number of latent groups, K, is a free parameter that must be chosen before the model and it controls model's complexity. The model can be also expressed as an exponential family:

$$\mathscr{L}(A|z,\theta) \propto \exp\left(\sum_{ij} T(A_{ij}) \cdot \eta(\theta_{z_i z_j})\right) \tag{2}$$

where $T(x) = (x,1)$ is the vector-valued function of sufficient statistics of the Bernoulli random variable and $\eta(x) = (\log(x/(1-x)), \log(1-x))$ is the vector-valued function of natural parameters.

This is a basic and classical SBM for unweighted networks since, as they are defined, the functions (T,η) produces binary edge values. With a different and ap-

propriate functions of (T, η), a specific form of Weighted Stochastic Block Model (WSBM) can be established by weights that are drawn from an exponential family distribution over the domain of T. In this case, each $\theta_{z_i z_j}$ denotes the parameters governing the weight distribution of the edge bundle $(z_i z_j)$.
However, the models SBM and WSBM that presented above, produce complete graphs. In order to model sparse networks by SBM and WSBM, [Aicher et al., 2014] assumes $A_{ij} = 0$ as a directed edge from node i to j is existed with zero weight. In this case, to denote absence of an edge from node i to j is $A_{ij} = \text{NaN}$. By this, sparse networks can be modelled with two types of information, edge existence and edge weight values, in together by a simple tuning parameter with the following way:

$$\log \mathscr{L}(A|z,\theta) = \alpha \left(\sum_{ij \in E} T_e(A_{ij}) \cdot \eta_e(\theta^{(e)}_{z_i z_j}) \right) + (1-\alpha) \left(\sum_{ij \in W} T_w(A_{ij}) \cdot \eta_e(\theta^{(w)}_{z_i z_j}) \right) \quad (3)$$

where the pair (T_e, η_e) denotes the family of edge existence distribution and the pair (T_w, η_w) denotes the family of edge-weight distribution, $\alpha \in [0,1]$ is a simple tuning parameter that combines their contributions in the likelihood function. E is the set of observed interactions (including non-edges) and W is the set of weighted edges with $W \subset E$.
It is possible to see in Equation 3 that if $\alpha = 1$ then the model reduces to SBM, to Equation 1, and if $\alpha = 0$ the model ignores edge existence information then we call such models as pure WSBM (pWSBM). When $0 < \alpha < 1$, the likelihood function combines both information set, i.e. it is called balanced WSBM if $\alpha = 0.5$.[1]

3 An application to the European financial market

We provide an application of the model and analyse the European financial institutions using the closing price series from Worldscope lists downloaded in Datastream at a daily frequency from 29^{th} May 1997 to 27^{th} May 1998. The network is estimated through Granger causality tests [Billio et al., 2012]. In SBM ($\alpha = 1$), Figure 1, the most systematically important community is Block 3 where the members have the highest out degree. However, there are not particular edge weight characteristics in SBM ($\alpha = 1$) communities. In balanced WSBM ($\alpha = 0.5$) the most important community is Block 2.

Acknowledgement

The author acknowledges financial support from the Marie Skłodowska-Curie Actions, European Union, Seventh Framework Program HORIZON 2020 under REA

[1] See [Aicher et al., 2014] for more detailed information.

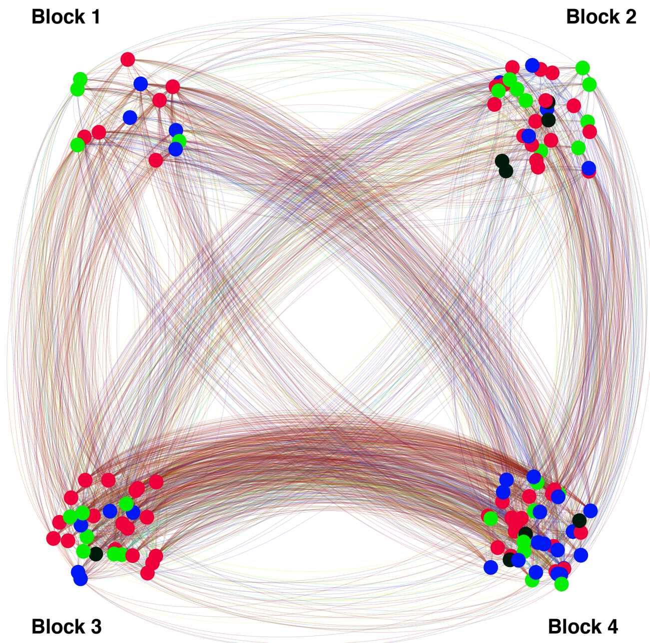

Fig. 1: European Financial Institution Network. Red nodes are banks, blue nodes are insurance companies, green nodes are companies in Financial Service sector and black nodes are Real Estate Companies.

grant agreement n.707070. He also gratefully acknowledges research support from the Research Center SAFE, funded by the State of Hessen initiative for research LOEWE.

References

[Aicher et al., 2014] Aicher, C., Jacobs, A. Z., and Clauset, A. (2014). Learning latent block structure in weighted networks. *Journal of Complex Networks*, 3(2):221–248.

[Billio et al., 2012] Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, 104(3):535–559.

[Holland et al., 1983] Holland, P. W., Laskey, K. B., and Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Social Networks*, 5(2):109–137.

[Newman, 2010] Newman, M. (2010). *Networks: an introduction*. Oxford University Press.