Efficient Bayesian inference of systemic risk interlinkages

Veni **Arakelian**[1] and Apostolos **Chalkis**[2]

[1]Piraeus Bank

[2]UoA, ATHENA, GeomScale org

Naxos, July 16, 2021

Acknowledgement: This paper has received funding from the European Union's Horizon 2020 research and innovation program "FIN-TECH: A Financial supervision and Technology compliance training programme" under the grant agreement No 825215 (Topic: ICT-35-2018, Type of action: CSA). Also, it is based upon work from COST Action 19130, supported by COST (European Cooperation in Science and Technology). **Disclaimer**: This study expresses the views of the authors and it should not be reported as representing the views of the Piraeus Bank.

Presentation structure

Introduction

Presentation of the problem

We discuss Bayesian models to estimate the inverse covariance matrix –also known as the concentration matrix or precision matrix– from Gaussian data.

More formally, the problem is to estimate the concentration matrix when k observations $y_1, \ldots, y_k$ are given such that,

$$y_t \sim N(0, \Sigma), \; y_t \in \mathbb{R}^n, \; t = 1, \ldots, k. \tag{1}$$

Bayesian inference for the covariance matrix

Bayesian estimation of inverse covariance matrix

Let the concentration matrix $=$ ^{1}. Our approach is to use the decomposition of introduced in Barnard et al. (2000), that is,

$$= T \quad C \quad T; \tag{2}$$

where C is a correlation matrix and T is a diagonal matrix. Each entry of matrix T is $T_i = \ _{ii}^{1=2}; \ i = 1; : : : ; n$ so that T_i^2 corresponds to the inverse partial variance of $y_{i;t}$.

Bayesian inference for the covariance matrix

Bayesian estimation of inverse covariance matrix

Moreover, the partial correlation coe□cients ^{ij} are given by,

$$_{ij} = \quad _{ij}=(\quad _{ii} \quad _{jj})^{1=2} = \quad C_{ij}; \qquad (3)$$

which means that C contains the negative of the partial correlation coe□cients.

As in Barnard et al. (2000), we assume that in the prior the elements $_{ii}$ are independently and identically distributed, and are independent of the elements of C.

As they notice in Barnard et al. (2000), the flexibility in dealing with tails of individual components is a key practical advantage of this separation strategy.

Contribution : The geometrical framework to represent correlation matrices and the MCMC sampler we introduce, allow us to ***use several priors for both the*** T_i ***and the correlation matrix*** C ***while our sampler handles all the possible different choices of priors without requiring any modification.***

Priors

Priors for the partial precisions Ω_{ii}.

Set a prior such that T_i are independently and identically distributed and the prior is also uninformative. In Barnard et al. (2000), they consider T_i as a vector $T\ 2\ \mathrm{R}^n$ and they use the prior

$$T \quad N(\ ; \ ); \tag{4}$$

where the matrix $\ 2\ \mathrm{R}^{n \quad n}$ is diagonal, that is, they choose independent log normal distributions for each of the standard deviations.

In Wong et al. (2003), the T_i follows a gamma distribution with parameters and ,

$$p(T_i)\ /\ T_i^2 \quad ^1 exp(\quad T_i^2): \tag{5}$$

In Barnard et al. (2000), they consider two choices for the prior $p(C)$. The first is the jointly uniform prior for the partial correlations $_{ij}$ and the second is the marginally uniform prior for each ^{ij}.

A third popular choice for the prior of is the Jeffreys prior. However, the use of Jeffreys prior in many cases fails to shrink the eigenvalues appropriately and suffers from estimation errors when C is close to a singular matrix.

Geometric representation of correlation matrices

We introduce a new geometric and computational framework to represent the set of correlation matrices C_R with a convex body $K \in R^p$, where p is the number of non-diagonal elements in the upper (or lower) triangular part.

Then, each point in the interior of K represents a correlation matrix. We define the body K as the intersection of the hypercube $H = [-1, 1]^p$ and a spectrahedron $S \in R^p$. The first guarantees that the non-diagonal elements lie in $[-1, 1]$ and the second that the matrix has unit diagonal elements and is positive definite.

Spectrahedron Definition

Spectrahedra are probably the most well studied shapes after polyhedra. A spectrahedron $S \subseteq \mathbb{R}^p$ is the feasible set of a linear matrix inequality. That is, let

$$F(x) = A_0 + x_1 A_1 + \cdots + x_p A_p, \tag{6}$$

where $A_i \in \mathbb{R}^{n \times n}$ are symmetric matrices; then the corresponding spectrahedron is the set,

$$S = \{x \in \mathbb{R}^p \mid F(x) \succeq 0\}, \tag{7}$$

where $\succeq$ denotes positive semidefiniteness.

A three-dimensional example

Let the hypercube $H := [-1, 1]^3$ and the spectrahedron

$$S := \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid F(x) \succeq 0\}$$

where,

$$F(x) = A_0 + x_1 A_1 + x_2 A_2 + x_3 A_3,$$

is the Linear Matrix Inequality (LMI), and,

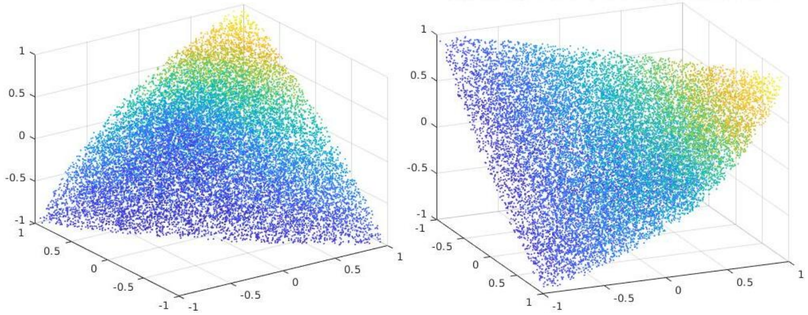
$$A_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, A_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Note that for any point $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ in the interior of $K = S \cap H$, the matrix

$$F(x) = \begin{bmatrix} 1 & x_1 & x_2 \\ x_1 & 1 & x_3 \\ x_2 & x_3 & 1 \end{bmatrix}$$

is a correlation matrix. Thus, the set of 3×3 correlation matrices can be represented by the interior of $K = S \cap H$.

We sample 20000 uniformly distributed points and plot them from two different angles.



Simulation study

Comparing prior choices

Our experiments show that the efficiency of Reflective Hamiltonian Monte Carlo (ReHMC) depends on the number of Gaussian observations k (or sample size) and the ratio between the largest over the minimum eigenvalue $_{max}= _{min}$ of the covariance matrix .

The sample size k determines the estimation error, i.e. the larger the sample size the smaller the estimation error. The ratio $_{max}= _{min}$ determines the mixing rate of the MCMC algorithm.

We perform the following experiment :

1. Generate a covariance matrix $2\ \mathrm{R}^{n\ \ n}$ from the Wishart distribution.
2. Modify the element $_{11}$ such that the ratio of the maximum over the minimum eigenvalue of to be $\frac{\max}{\min} = 5$.
3. We set the sample size $k = 20000$ and generate $y_1; : : : ; y_k\ 2\ \mathrm{R}^n$, i.i.d. vectors from $N(0; \ )$.
4. Sample from the posterior distribution for uniform prior with ReHMC.
5. End sampling when the Potential Scale Reduction Factor (PSFR) (Gelman (1992)) of each marginal (partial correlation) is smaller than $1{:}1$, and then compute the sample average to estimate the covariance matrix $2\ \mathrm{R}^{n\ \ n}$.

Simulation results

n	p	N	PSRF	Ef	t(sec)	$jj\ \ jj_1$	$jj\ \ jj_F$
10	55	2000	1.0724	32.635	3.3	0.0357	0.0343
20	210	4000	1.0577	12.751	10.4	0.0759	0.0716
30	465	6000	1.0659	4.093	25.6	0.1234	0.1206
40	820	8000	1.0959	1.965	50.6	0.1704	0.1695
50	1275	10000	1.0831	0.873	118.6	0.2096	0.2081
60	1830	14000	1.0807	0.380	302.1	0.2496	0.2573
70	2485	18000	1.0895	0.153	645.3	0.2974	0.3045
80	3240	18000	1.0968	0.086	1035	0.3488	0.3509
90	4095	18000	1.0945	0.059	1477	0.3958	0.4063
100	5050	28000	1.0812	0.027	3739	0.4502	0.4604

Motivation (maybe)

Source: ECB and ECB calculations.

Empirical Results

We apply our methodology to cross border bank total claims at the two core countries of the Euro zone, France and Germany and at GIIPS, namely Greece, Ireland, Italy, Portugal and Spain.

The prior distributions are the (jointly) uniform, the inverse-Wishart (marginally uniform prior on the partial correlations), and the Jeffrey prior.

We find that under different prior specifications the biggest correlations are found between Germany and Ireland, enhancing the implications of the banking crisis in the country.

	France	Italy	Germany	Greece	Ireland	Spain	Portugal
France	1	0.3481	0.5226	0.1416	-0.1804	0.0975	0.1124
Italy	0.3481	1	0.3736	-0.0086	-0.3928	0.5222	-0.3239
Germany	0.5226	0.3736	1	-0.1345	0.8331	-0.5701	0.5859
Greece	0.1416	-0.0086	-0.1345	1	0.0839	0.0836	0.0288
Ireland	-0.1804	-0.3928	0.8331	0.0839	1	0.4053	-0.5792
Spain	0.0975	0.5222	-0.5701	0.0836	0.4053	1	0.7623
Portugal	0.1124	-0.3239	0.5859	0.0288	-0.5792	0.7623	1

	France	Italy	Germany	Greece	Ireland	Spain	Portugal
France	1	0.3531	0.4626	0.0979	0.0728	0.0835	0.2717
Italy	0.3531	1	0.2170	0.0244	-0.0974	0.2750	0.1016
Germany	0.4626	0.2170	1	-0.1024	0.4913	-0.1284	0.2364
Greece	0.0979	0.0244	-0.1024	1	-0.0108	0.2488	0.0614
Ireland	0.0728	-0.0974	0.4913	-0.0108	1	-0.0894	-0.1657
Spain	0.0835	0.2750	-0.1284	0.2488	-0.0894	1	0.4183
Portugal	0.2717	0.1016	0.2364	0.0614	-0.1657	0.4183	1

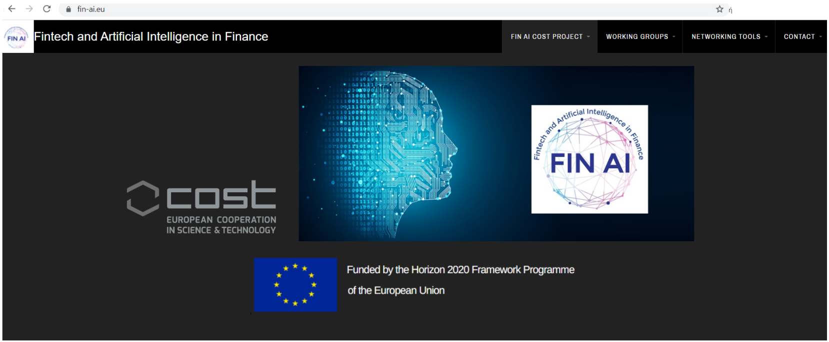
	France	Italy	Germany	Greece	Ireland	Spain	Portugal
France	1	0.3786	0.5185	0.0893	-0.0149	0.0734	0.2382
Italy	0.3786	1	0.2193	-0.0170	-0.1516	0.3516	-0.0090
Germany	0.5185	0.2193	1	-0.1101	0.6064	-0.2466	0.3106
Greece	0.0893	-0.0170	-0.1101	1	0.0397	0.2218	0.0390
Ireland	-0.0149	-0.1516	0.6064	0.0397	1	-0.0231	-0.2467
Spain	0.0734	0.3516	-0.2466	0.2218	-0.0231	1	0.5290
Portugal	0.2382	-0.0090	0.3106	0.0390	-0.2467	0.5290	1

Conclusions - Future work

We find that under different prior specifications the biggest correlations are found between Germany and Ireland, enhancing the implications of the banking crisis in the country.

Thanks for attending! Insert time variation...

Maybe interesting



References

Barnard, J., R. McCulloch, and X. Meng (2000, October). Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage, Statistica Sinica, 10(4), 1281–1311.

Gelman, A. and D. B. Rubin (1992). Inference from iterative simulation using multiple sequences, Statistical Science, 7(4), 457–472.

Wong, F., C. K. Carter, and R. Kohn (2003). E□cient estimation of covariance selection models, Biometrika, 90(4), 809–830. 90(4), 809–830.