Non-Asymptotic Properties of Sparse Approximate Factor Models

Benjamin Poignard, Osaka University, Graduate School of Engineering Science
Yoshikazu Terada, Osaka University, Graduate School of Engineering Science - RIKEN AIP

The factor modelling approach consists in summarizing the information through a small number of variables called factors. The quantity of interest is the variance covariance matrix Σ which is decomposed as

$$\Sigma = \Lambda\Lambda^T + \Psi,$$

where Ψ is the $p \times p$ variance covariance matrix of the idiosyncratic variables and Λ is the $p \times m$ matrix of the loading factors. Although the factor model approach aims at reducing the statistical complexity, the total number of parameters remains significant for large p-dimensional random vectors. As a consequence, the total number of parameters to estimate in the idiosyncratic variance covariance matrix is $O(p^2)$. To foster sparsity, Bai and Liao (2016) proposed a Lasso/Scad penalization procedure based on a Gaussian quasi maximum likelihood approach for estimating a sparse idiosyncratic variance covariance matrix. They derive asymptotic bounds and consistency results for the penalized estimate $\hat{\Psi}$.

In our study, we provide a finite-sample theoretical analysis of sparse idiosyncratic variance covariance matrices Ψ together with the conditions to satisfy the support recovery property using the theoretical framework developed by Loh and Wainwright (2017). To do so, we propose a two-step estimation, where the statistical criterion is given by

$$\begin{cases} (\tilde{\Lambda}, \tilde{\Psi}) & = & \underset{\Lambda,\Psi}{\arg\min} \{\mathbb{G}_n(\Lambda,\Psi)\}, \text{ with } \mathbb{G}_n(\Lambda,\Psi) = \frac{1}{2n}\big(\log(|\Sigma(\Lambda,\Psi)|) + tr(\hat{S}\Sigma(\Lambda,\Psi)^{-1})\big), \\ \hat{\Psi} & = & \underset{\Psi:\Psi\succeq 0, g(vec(\Psi))\leq R}{\arg\min} \{\mathbb{F}_n(\Psi;\tilde{\Lambda},\tilde{\Psi}) + \boldsymbol{p}(\gamma_n, vec(\Psi))\}, \text{ where} \\ \mathbb{F}_n(\Psi;\tilde{\Lambda},\tilde{\Psi}) & = & vec(\Psi-\tilde{\Psi})'\{\nabla^2_{vec(\Psi)vec(\Psi)'}\mathbb{G}_n(\tilde{\Lambda},\tilde{\Psi})\}vec(\Psi-\tilde{\Psi}), \end{cases}$$

where $\hat{S}$ is the sample variance covariance estimate and $vec(.)$ is the vectorization operator that stacks the columns of any matrix on top of one another into a vector. In a first step, both the loading factor matrix and idiosyncratic matrix are obtained based on a Gaussian maximum likelihood estimator. Conditionally on these first step estimates, the regularized idiosyncratic matrix corresponds to the solution of a minimum distance criterion plus the regularizer. This sparse estimator is thus based on a least square type loss function weighted by the Hessian matrix of the first step Gaussian likelihood function. The regularization procedure is performed by the regularizer $\boldsymbol{p}(\gamma_n, .) : \mathbb{R}^{p^2} \to \mathbb{R}$, where γ_n is the regularization parameter, which depends on the sample size, and enforce a particular type of sparse structure in the solution $\hat{\Psi}$ and $R > 0$ is a supplementary regularization parameter. Due to the potential non-convexity of this penalty, we include the side condition $g(vec(\Psi)) \geq \|vec(\Psi)\|_1$ with $g : \mathbb{R}^{p^2} \to \mathbb{R}$ a convex function, and R to ensure the existence of local/global optima.

Under the restricted strong convexity of the unpenalized loss function $\mathbb{F}_n(.;\tilde{\Lambda},\tilde{\Psi})$ and regularity conditions on the regularizer, we prove non-asymptotic error bounds on the regularized idiosyncratic matrix estimator. Moreover, based on the primal-dual witness method, we establish variable selection consistency, including the case when the regularizer is non-convex. These theoretical results are supported by empirical studies.

References

Bai, J. and Liao, K. (2016). *Efficient estimation of approximate factor models via penalized maximum likelihood.* Journal of Econometrics, Vol. 191, 1-18.

Loh, P.L. and Wainwright, M.J. (2017). *Support recovery without incoherence: a case for non-convex regularization.* The Annals of Statistics, Vol. 45, No. 6, 2455-2482.