Mantoux, Clément; **Durrleman, Stanley**; **Allassonnière, Stéphanie**
Asymptotic analysis of a matrix latent decomposition model. (English) Zbl 07533061
ESAIM, Probab. Stat. 26, 208-242 (2022)

Summary: Matrix data sets arise in network analysis for medical applications, where each network belongs to a subject and represents a measurable phenotype. These large dimensional data are often modeled using lower-dimensional latent variables, which explain most of the observed variability and can be used for predictive purposes. In this paper, we provide asymptotic convergence guarantees for the estimation of a hierarchical statistical model for matrix data sets. It captures the variability of matrices by modeling a truncation of their eigendecomposition. We show that this model is identifiable, and that consistent Maximum A Posteriori (MAP) estimation can be performed to estimate the distribution of eigenvalues and eigenvectors. The MAP estimator is shown to be asymptotically normal for a restricted version of the model.

MSC:

62R30 Statistics on manifolds
62F12 Asymptotic properties of parametric estimators
62H11 Directional data; spatial statistics

Keywords:

hierarchical model; matrix data sets; low rank; Stiefel manifold; identifiability; strong consistency; asymptotic normality

Software:

rstiefel

Full Text: DOI

References:

[1] C. Aicher, A.Z. Jacobs and A. Clauset, Learning Latent Block Structure in Weighted Networks. J. Complex Netw. 3 (2015) 221-248. · Zbl 1397.68151 · doi:10.1093/comnet/cnu026

[2] M. Ali and J. Gao, Classification of matrix-variate fisher—bingham distribution via maximum likelihood estimation using manifold valued data. Neurocomputing 295 (2018) 72-85.· doi:10.1016/j.neucom.2018.01.048

[3] S. Allassonniere, Y. Amit and A. Trouvé, Toward a coherent statistical framework for dense deformable template estimation. J. Royal Stat. Soc. B 69 (2007) 3-29.· doi:10.1111/j.1467-9868.2007.00574.x

[4] E.S. Allman, C. Matias and J.A. Rhodes, Identifiability of parameters in latent structure models with many observed variables. Ann. Stat. 37 (2009) 3099-3132. · Zbl 1191.62003 · doi:10.1214/09-AOS689

[5] T.W. Anderson and Y. Amemiya, The asymptotic normal distribution of estimators in factor analysis under general conditions. Ann. Stat. 16 (1988) 759-771. · Zbl 0646.62051 · doi:10.1214/aos/1176350834

[6] O.E. Barndorff-Nielsen, Identifiability of mixtures of exponential families. J. Math. Anal. Appl. 12 (1965) 115-121. · Zbl 0138.12105 · doi:10.1016/0022-247X(65)90059-4

[7] O.E. Barndorff-Nielsen, Information and exponential families. In: Statistical theory, Wiley series in probability and mathematical statistics. Wiley, Chichester, New York (1978).

[8] P.J. Bickel, Y. Ritov and T. Rydén, Asymptotic normality of the maximum-likelihood estimator for general hidden Markov models. Ann. Stat. 26 (1998) 1614-1635. · Zbl 0932.62097 · doi:10.1214/aos/1024691255

[9] S. Bonhomme and J.-M. Robin, Consistent noisy independent component analysis. J. Econ. 149 (2009) 12-25. · Zbl 1429.62215 · doi:10.1016/j.jeconom.2008.12.019

[10] J. Chen, G. Han, H. Cai, J. Ma, M. Kim, P. Laurienti and G. Wu, Estimating common harmonic waves of brain networks on Stiefel manifold, in A.L. Martel, P. Abolmaesumi, D. Stoyanov, D. Mateus, M.A. Zuluaga, S.K. Zhou, D. Racoceanu and L. Joskowicz (editors), Medical Image Computing and Computer Assisted Intervention — MICCAI 2020, Lecture Notes in Computer Science, Springer International Publishing, Cham (2020) 367-367.

[11] J. Chevallier, V. Debavelaere and S. Allassonniére, A coherent framework for learning spatiotemporal piecewise-geodesic trajectories from longitudinal manifold-valued data. SIAM J. Imag. Sci. 14 (2021) 349-388. · Zbl 1474.62463 · doi:10.1137/20M1328026

[12] Y. Chikuse, Concentrated matrix Langevin distributions. J. Multivar. Anal. 85 (2003) 375-394. · Zbl 1016.62065 · doi:10.1016/S0047-259X(02)00065-9

[13] Y. Chikuse, Statistics on Special Manifolds, Lecture Notes in Statistics, Springer-Verlag, New York (2003). · Zbl 1026.62051 · doi:10.1007/978-0-387-21540-2

[14] Y. Chikuse, State space models on special manifolds. J. Multivar. Anal. 97 (2006) 1284-1294. · Zbl 1099.62098 · doi:10.1016/j.jmva.2006.03.002

[15] R. Douc, Non Singularity of the Asymptotic Fisher Information Matrix in Hidden Markov Models. arXiv:math/0511631 (2005).

[16] R. Douc, E. Moulines, J. Olsson and R. van Handel, Consistency of the maximum likelihood estimator for general hidden Markov models. Ann. Stat. 39 (2011) 474-513. · Zbl 1209.62194

[17] R. Douc, F. Roueff and T. Sim, Necessary and sufficient conditions for the identifiability of observation-driven models. J. Time Ser. Anal. 42 (2021) 140-160. · Zbl 07364926 · doi:10.1111/jtsa.12559

[18] N.S. D'Souza, M.B. Nebel, N. Wymbs, S. Mostofsky and A. Venkataraman, A generative-discriminative basis learning framework to predict clinical severity from resting state functional MRI data, in A.F. Frangi, J.A. Schnabel, C. Davatzikos, C. Alberola-Léopez and G. Fichtinger (editors), Medical Image Computing and Computer Assisted Intervention — MICCAI 2018. Springer International Publishing, Cham (2018), vol. 11072, 163-163.· doi:10.1007/978-3-030-00931-1$_1$9

[19] N.S. D'Souza, M.B. Nebel, N. Wymbs, S. Mostofsky and A. Venkataraman, Integrating neural networks and dictionary learning for multidimensional clinical characterizations from functional connectomics data, in D. Shen, T. Liu, T.M. Peters, L.H. Staib, C. Essert, S. Zhou, P.-T. Yap and A. Khan (editors), Medical Image Computing and Computer Assisted Intervention — MICCAI 2019. Springer International Publishing, Cham (2019), vol. 11766, 709-709.· doi:10.1007/978-3-030-32248-9$_7$9

[20] L.L. Duan, G. Michailidis and M. Ding, Spiked Laplacian Graphs: Bayesian Community Detection in Heterogeneous Networks. arXiv:1910.02471

[21] A. Edelman, T.A. Arias and S.T. Smith, The geometry of algorithms with orthogonality constraints. SIAM J. Matrix Anal. Appl. 20 (1998) 303-353. · Zbl 0928.65050 · doi:10.1137/S0895479895290954

[22] K. Fan, On a theorem of weyl concerning eigenvalues of linear transformations I. Proc. Natl. Acad. Sci. 35 (1949) 652-655.· doi:10.1073/pnas.35.11.652

[23] P.J. Forrester, Log-gases and random matrices (LMS-34). Vol. 34 of London Mathematical Society Monographs. Princeton University Press (2010). · Zbl 1217.82003

[24] C. Fraikin, K. Hüper and P.V. Dooren, Optimization over the Stiefel Manifold, in vol. 7 of PAMM: Proceedings in Applied Mathematics and Mechanics. Wiley Online Library (2007) 1062205-1062205.· doi:10.1002/pamm.200700861

[25] Y. Gu and G. Xu, Identifiability of Hierarchical Latent Attribute Models. arXiv:1906.07869

[26] P.D. Hoff, Simulation of the matrix Bingham—von Mises—Fisher distribution, with applications to multivariate and relational data. J. Comput. Graph. Stat. 18 (2009) 438-456.· doi:10.1198/jcgs.2009.07177

[27] H. Holzmann, A. Munk and B. Stratmann, Identifiability of finite mixtures - with applications to circular distributions. Sankhya 66 (2004) 440-449. · Zbl 1192.62144

[28] S. Janson, Graphons, Cut Norm and Distance, Couplings and Rearrangements, Vol. 4 of New York Journal of Mathematics. NYJM Monographs, State University of New York, University at Albany, Albany, NY 4 (2013) 76. · Zbl 1268.05001

[29] M. Jauch, P.D. Hoff and D.B. Dunson, Random Orthogonal Matrices and the Cayley Transform. Bernoulli 26 (2020) 1560-1586. · Zbl 1466.60010 · doi:10.3150/19-BEJ1176

[30] P.E. Jupp and K.V. Mardia, Maximum Likelihood Estimators for the Matrix Von Mises-Fisher and Bingham Distributions. Ann. Stat. 7 (1979) 599-606. · Zbl 0406.62012

[31] J.T. Kent, Identifiability of Finite Mixtures for Directional Data, Ann. Stat. 11 (1983). · Zbl 0515.62018

[32] C.G. Khatri and K.V. Mardia, The von Mises—Fisher Matrix Distribution in Orientation Statistics. J.R. Stat Soc. Ser. B (Methodological) 39 (1977) 95-106. · Zbl 0356.62044

[33] A. Khetan and M. Mj, Cheeger Inequalities for Graph Limits, arXiv:1807.02225

[34] T.N. Kipf and M. Welling, Semi-Supervised Classification with Graph Convolutional Networks, in ICLR 2017 (2017).

[35] A. Kume, S.P. Preston and A.T.A. Wood, Saddlepoint Approximations for the Normalizing Constant of Fisher—Bingham Distributions on Products of Spheres and Stiefel Manifolds. Biometrika 100 (2013) 971-984. · Zbl 1452.62998 · doi:10.1093/biomet/ast021

[36] P. Latouche and S. Robin, Variational Bayes Model Averaging for Graphon Functions and Motif Frequencies Inference in W-graph Models, Stat. Comput. 26 (2016) 1173-1173. · Zbl 1349.62081 · doi:10.1007/s11222-015-9607-0

[37] M. Lavielle and L. Aarons, What Do We Mean by Identifiability in Mixed Effects Models?. J. Pharmacokinet. Pharmacodyn. 43 (2016) 111-122.· doi:10.1007/s10928-015-9459-4

[38] E.L. Lehmann and G. Casella, Theory of Point Estimation, Springer Texts in Statistics, 2nd edn., Springer, New York (2003). · Zbl 0916.62017

[39] X. Li, N.C. Dvornek, Y. Zhou, J. Zhuang, P. Ventola and J.S. Duncan, Graph Neural Network for Interpreting Task-fMRI Biomarkers, in D. Shen, T. Liu, T.M. Peters, L.H. Staib, C. Essert, S. Zhou, P.-T. Yap and A. Khan (editors), Medical Image Computing and Computer Assisted Intervention — MICCAI 2019, Lecture Notes in Computer Science, Springer International Publishing, Cham (2019) 485-485.

[40] X. Liang, L. Wang, L.-H. Zhang and R.-C. Li, On Generalizing Trace Minimization. arXiv:2104.00257

[41] L. Lin, V. Rao and D. Dunson, Bayesian Nonparametric Inference on the Stiefel Manifold. Stat. Sin. 27 (2017) 535-553. · Zbl 1362.62071

[42] L. Lovász, Large Networks and Graph Limits. Colloquium Publications, vol. 60, American Mathematical Society, Providence, Rhode Island (2012). · Zbl 1292.05001 · doi:10.1090/coll/060

[43] C. Mantoux, B. Couvy-Duchesne, F. Cacciamani, S. Epelbaum, S. Durrleman and S. Allassonnière, Understanding the Variability in Graph Data Sets through Statistical Modeling on the Stiefel Manifold, Entropy 23 (2021) 490.· doi:10.3390/e23040490

[44] S.S. Mukherjee and S. Chakrabarti, Graphon Estimation from Partially Observed Network Data. arXiv:1906.00494

[45] S.C. Olhede and P.J. Wolfe, Network Histograms and Universality of Blockmodel Approximation. Proc. Natl. Acad. Sci. 111 (2014) 14722-14727.· doi:10.1073/pnas.1400374111

[46] S. Pal, S. Sengupta, R. Mitra and A. Banerjee, Conjugate Priors and Posterior Inference for the Matrix Langevin Distribution on the Stiefel Manifold. Bayesian Anal. 15 (2020) 871-908. · Zbl 1459.62238

[47] T.P. Peixoto, Bayesian Stochastic Blockmodeling, in P. Doreian, V. Batagelj and A. Ferligoj (editors), Advances in Network Clustering and Blockmodeling, Wiley Series in Computational and Quantitative Social Science, 289-332, Wiley (2020) .

[48] Z. Ren, T. Sun, C.-H. Zhang and H.H. Zhou, Asymptotic Normality and Optimalities in Estimation of Large Gaussian Graphical Models. Ann. Stat. 43 (2015). · Zbl 1328.62342

[49] A.A. Shabalin and A.B. Nobel, Reconstruction of a Low-Rank Matrix in the Presence of Gaussian Noise. J. Multivar. Anal. 118 (2013) 67-76. · Zbl 1280.15022 · doi:10.1016/j.jmva.2013.03.005

[50] B. Sischka and G. Kauermann, EM-based Smooth Graphon Estimation Using MCMC and Spline-Based Approaches. Soc. Netw. 68 (2022) 279-295.· doi:10.1016/j.socnet.2021.08.007

[51] E. Tabrizi, E.B. Samani and M. Ganjali, A Note on the Identifiability of Latent Variable Models for Mixed Longitudinal Data. Stat. Probab. Lett. 167 (2020) 108882. · Zbl 1455.62102 · doi:10.1016/j.spl.2020.108882

[52] H. Teicher, Identifiability of Finite Mixtures. Ann. Math. Stat. 34 (1963) 1265-1269. · Zbl 0137.12704 · doi:10.1214/aoms/1177703862

[53] T. Traynor, Change of Variables for Hausdorff Measure (from the Beginning). Universitá degli Studi di Trieste. Dipartimento di Scienze Matematiche 26 suppl. (1994) 327-327. · Zbl 0876.28010

[54] A.W. van der Vaart, Asymptotic Statistics, Cambridge Series in Statistical and Probabilistic Mathematics, 1st edn., Cambridge Univ. Press, Cambridge (1998).

[55] J. Xu, Rates of Convergence of Spectral Methods for Graphon Estimation, in International Conference on Machine Learning (2018) 5433-5433. · Zbl 07123442

[56] S.J. Yakowitz and J.D. Spragins, On the Identifiability of Finite Mixtures. Ann. Math. Stat. 39 (1968) 209-214. · Zbl 0155.25703 · doi:10.1214/aoms/1177698520