Avagyan, Vahe; **Alonso, Andrés M.**; **Nogales, Francisco J.**

D-trace estimation of a precision matrix using adaptive lasso penalties. (English)

Zbl 1414.62224
Adv. Data Anal. Classif., ADAC 12, No. 2, 425-447 (2018).

Summary: The accurate estimation of a precision matrix plays a crucial role in the current age of high-dimensional data explosion. To deal with this problem, one of the prominent and commonly used techniques is the ℓ_1 norm (Lasso) penalization for a given loss function. This approach guarantees the sparsity of the precision matrix estimate for properly selected penalty parameters. However, the ℓ_1 norm penalization often fails to control the bias of obtained estimator because of its overestimation behavior. In this paper, we introduce two adaptive extensions of the recently proposed ℓ_1 norm penalized D-trace loss minimization method. They aim at reducing the produced bias in the estimator. Extensive numerical results, using both simulated and real datasets, show the advantage of our proposed estimators.

MSC:

62H30 Classification and discrimination; cluster analysis (statistical aspects)
62J10 Analysis of variance and covariance (ANOVA)
65S05 Graphical methods in numerical analysis

Keywords:

adaptive thresholding; D-trace loss; Gaussian graphical model; gene expression data; high-dimensionality

Software:

glasso

Full Text: DOI

References:

[1] Anderson TW (2003) An introduction to multivariate statistical analysis. Wiley-Interscience, New York

[2] Banerjee, O.; Ghaoui, L.; d'Aspremont, A., Model selection through sparse maximum likelihood estimation for multivariate Gaussian or binary data, J Mach Learn Res, 9, 485-516, (2008) · Zbl 1225.68149

[3] Banerjee, S.; Ghosal, S., Bayesian structure learning in graphical models, J Multivar Anal, 136, 147-162, (2015) · Zbl 1308.62119 · doi:10.1016/j.jmva.2015.01.015

[4] Bickel, PJ; Levina, E., Regularized estimation of large covariance matrices, Ann Stat, 36, 199-227, (2008) · Zbl 1132.62040 · doi:10.1214/009053607000000758

[5] Cai, T.; Liu, W.; Luo, X., A constrained ${\ell _1}$ minimization approach to sparse precision matrix estimation, J Am Stat Assoc, 106, 594-607, (2011) · Zbl 1232.62087 · doi:10.1198/jasa.2011.tm10155

[6] Cai, T.; Yuan, M., Adaptive covariance matrix estimation through block thresholding, Ann Stat, 40, 2014-2042, (2012) · Zbl 1257.62060 · doi:10.1214/12-AOS999

[7] Cui, Y.; Leng, C.; Sun, D., Sparse estimation of high-dimensional correlation matrices, Comput Stat Data Anal, 93, 390-403, (2016) · Zbl 1468.62044 · doi:10.1016/j.csda.2014.10.001

[8] d'Aspremont, A.; Banerjee, O.; Ghaoui, L., First-order methods for sparse covariance selection, SIAM J Matrix Anal Appl, 30, 56-66, (2008) · Zbl 1156.90423 · doi:10.1137/060670985

[9] Dempster, A., Covariance selection, Biometrics, 28, 157-175, (1972)· doi:10.2307/2528966

[10] Deng, X.; Tsui, K., Penalized covariance matrix estimation using a matrix-logarithm transformation, J Comput Graph Stat, 22, 494-512, (2013)· doi:10.1080/10618600.2012.715556

[11] Duchi J, Gould S, Koller D (2008) Projected subgradient methods for learning sparse Gaussians. In: Proceeding of the 24th conference on uncertainty in artificial intelligence, pp 153-160. arXiv:1206.3249

[12] Karoui, N., Operator norm consistent estimation of large-dimensional sparse covariance matrices, Ann Appl Stat, 36, 2717-2756, (2008) · Zbl 1196.62064 · doi:10.1214/07-AOS559

[13] Fan, J.; Feng, J.; Wu, Y., Network exploration via the adaptive Lasso and SCAD penalties, Ann Appl Stat, 3, 521-541, (2009) · Zbl 1166.62040 · doi:10.1214/08-AOAS215

[14] Fan, J.; Li, R., Variable selection via nonconcave penalized likelihood and its oracle properties, J Am Stat Assoc, 96, 1348-1360,

(2001) · Zbl 1073.62547 · doi:10.1198/016214501753382273

[15] Frahm, G.; Memmel, C., Dominating estimator for minimum-variance portfolios, J Econom, 159, 289-302, (2010) · Zbl 1441.62264 · doi:10.1016/j.jeconom.2010.07.007

[16] Friedman, J.; Hastie, T.; Tibshirani, R., Sparse inverse covariance estimation with the graphical Lasso, Biostatistics, 9, 432-441, (2008) · Zbl 1143.62076 · doi:10.1093/biostatistics/kxm045

[17] Goto, S.; Xu, Y., Improving mean variance optimization through sparse hedging restrictions, J Finan Quant Anal, 50, 1415-1441, (2015)· doi:10.1017/S0022109015000526

[18] Haff, LR, Estimation of the inverse covariance matrix: random mixtures of the inverse Wishart matrix and the identity, Ann Stat, 8, 586-597, (1980) · Zbl 0441.62045 · doi:10.1214/aos/1176345010

[19] Hsieh C-J, Dhillon IS, Ravikumar PK, Sustik MA (2011) Sparse inverse covariance matrix estimation using quadratic approximation. In: Advances in neural information processing systems, vol 24, pp 2330-2338

[20] Huang, S.; Li, J.; Sun, L.; Ye, J.; Fleisher, A.; Wu, T.; Chen, K.; Reiman, E., Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation, NeuroImage, 50, 935-949, (2010)· doi:10.1016/j.neuroimage.2009.12.120

[21] Johnstone, IM, On the distribution of the largest eigenvalue in principal component analysis, Ann Stat, 29, 295-327, (2001) · Zbl 1016.62078 · doi:10.1214/aos/1009210544

[22] Jorissen, RN; Lipton, L.; Gibbs, P.; Chapman, M.; Desai, J.; Jones, IT; Yeatman, TJ; East, P.; Tomlinson, IP; Verspaget, HW; Aaltonen, LA; Kruhøffer, M.; Orntoft, TF; Andersen, CL; Sieber, OM, DNA copy-number alterations underlie gene expression differences between microsatellite stable and unstable colorectal cancers, Clin Cancer Res, 14, 8061-8069, (2008)· doi:10.1158/1078-0432.CCR-08-1431

[23] Kourtis, A.; Dotsis, G.; Markellos, N., Parameter uncertainty in portfolio selection: shrinking the inverse covariance matrix, J Bank Finan, 36, 2522-2531, (2012)· doi:10.1016/j.jbankfin.2012.05.005

[24] Kuerer, HM; Newman, LA; Smith, TL; Ames, FC; Hunt, KK; Dhingra, K.; Theriault, RL; Singh, G.; Binkley, SM; Sneige, N.; Buchholz, TA; Ross, MI; McNeese, MD; Buzdar, AU; Hortobagyi, GN; Singletary, SE, Clinical course of breast cancer patients with complete pathologic primary tumor and axillary lymph node response to doxorubicin-based neoadjuvant chemotherapy, J Clin Oncol, 17, 460-469, (1999)· doi:10.1200/JCO.1999.17.2.460

[25] Lam, C.; Fan, J., Sparsistency and rates of convergence in large covariance matrix estimation, Ann Stat, 37, 4254, (2009) · Zbl 1191.62101 · doi:10.1214/09-AOS720

[26] Lauritzen S (1996) Graphical models. Clarendon Press, Oxford · Zbl 0907.62001

[27] Ledoit, O.; Wolf, M., A well-conditioned estimator for large-dimensional covariance matrices, J Multivar Anal, 88, 365-411, (2004) · Zbl 1032.62050 · doi:10.1016/S0047-259X(03)00096-4

[28] Ledoit, O.; Wolf, M., Nonlinear shrinkage estimation of large-dimensional covariance matrices, Ann Stat, 40, 1024-1060, (2012) · Zbl 1274.62371 · doi:10.1214/12-AOS989

[29] Mardia KV, Kent JT, Bibby JM (1979) Multivariate analysis. Academic Press, New York · Zbl 0432.62029

[30] Matthews, BW, Comparison of the predicted and observed secondary structure of T4 phage lysozyme, Biochim Biophys Acta, 405, 442-451, (1975)· doi:10.1016/0005-2795(75)90109-9

[31] Maurya, A., A joint convex penalty for inverse covariance matrix estimation, Comput Stat Data Anal, 75, 15-27, (2014) · Zbl 06983941 · doi:10.1016/j.csda.2014.01.015

[32] McLachlan S (2004) Discriminant analysis and statistical pattern recognition. Wiley, New Jersey · Zbl 1108.62317

[33] Meinshausen, N., Relaxed Lasso, Comput Stat Data Anal, 52, 374-393, (2007) · Zbl 1452.62522 · doi:10.1016/j.csda.2006.12.019

[34] Meinshausen, N.; Bühlmann, P., High-dimensional graphs and variable selection with the Lasso, Ann Stat, 34, 1436-1462, (2006) · Zbl 1113.62082 · doi:10.1214/009053606000000281

[35] Nguyen, TD; Welsch, RE, Outlier detection and robust covariance estimation using mathematical programming, Adv Data Anal Classif, 4, 301-334, (2010) · Zbl 1284.62057 · doi:10.1007/s11634-010-0070-7

[36] Ravikumar, P.; Wainwright, M.; Raskutti, G.; Yu, B., High-dimensional covariance estimation by minimizing $\ell _1$-penalized log-determinant divergence, Electr J Stat, 5, 935-980, (2011) · Zbl 1274.62190 · doi:10.1214/11-EJS631

[37] Rothman, A.; Bickel, P.; Levina, E., Generalized thresholding of large covariance matrices, J Am Stat Assoc, 104, 177-186, (2009) · Zbl 1388.62170 · doi:10.1198/jasa.2009.0101

[38] Rothman, A.; Bickel, P.; Levina, E.; Zhu, J., Sparse permutation invariant covariance estimation, Electr J Stat, 2, 494-515, (2008) · Zbl 1320.62135 · doi:10.1214/08-EJS176

[39] Rothman, AJ, Positive definite estimators of large covariance matrices, Biometrika, 99, 733-740, (2012) · Zbl 1437.62595 · doi:10.1093/biomet/ass025

[40] Ryali, S.; Chen, T.; Supekar, K.; Menon, V., Estimation of functional connectivity in fMRI data using stability selection-based sparse partial correlation with elastic net penalty, NeuroImage, 59, 3852-3861, (2012)· doi:10.1016/j.neuroimage.2011.11.054

[41] Schafer J, Strimmer K (2005) A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. Stat Appl Genet Mol Biol 4(1):Article 32

[42] Scheinberg K, Ma S, Goldfarb D (2010) Sparse inverse covariance selection via alternating linearization methods. In: Advances in neural information processing systems, vol 23, pp 2101-2109

[43] Shi, L.; Reid, LH; Jones, WD; Shippy, R.; Warrington, JA; Baker, SC; Collins, PJ; deLongueville, F.; Kawasaki, ES; Lee, KY; Luo, Y.; Sun, YA; Willey, JC; Setterquist, RA; Fischer, GM; Tong, W.; Dragan, YP; Dix, DJ; Frueh, FW; Goodsaid, FM; Herman, D.; Jensen, RV; Johnson, CD; Lobenhofer, EK; Puri, RK; Scherf, U.; Thierry-Mieg, J.; Wang, C.; Wilson, M.;

Wolber, PK, The microarray quality control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models, Nat Biotechnol, 28, 827-838, (2010)· doi:10.1038/nbt.1665

[44] Stifanelli, PF; Creanza, TM; Anglani, R.; Liuzzi, VC; Mukherjee, S.; Schena, FP; Ancona, N., A comparative study of covariance selection models for the inference of gene regulatory networks, J Biomed Inf, 46, 894-904, (2013)· doi:10.1016/j.jbi.2013.07.002

[45] Tibshirani, R., Regression shrinkage and selection via the Lasso, J R Stat Soc, 58, 267-288, (1996) · Zbl 0850.62538

[46] Touloumis, A., Nonparametric Stein-type shrnikage covariance matrix estimators in high-dimensional settings, Comput Stat Data Anal, 83, 251-261, (2015) · Zbl 06984137 · doi:10.1016/j.csda.2014.10.018

[47] van de Geer S, Buhlmann P, Zhou S (2010) The adaptive and the thresholded Lasso for potentially misspecified models. arXiv preprint arXiv:1001.5176

[48] Wang, Y.; Daniels, MJ, Computationally efficient banding of large covariance matrices for ordered data and connections to banding the inverse Cholesky factor, J Multivar Anal, 130, 21-26, (2014) · Zbl 1292.62082 · doi:10.1016/j.jmva.2014.04.026

[49] Warton, DI, Penalized normal likelihood and ridge regularization of correlation and covariance matrices, J Am Stat Assoc, 103, 340-349, (2008) · Zbl 1471.62362 · doi:10.1198/016214508000000021

[50] Whittaker J (1990) Graphical models in applied multivariate statistics. Wiley, Chichester · Zbl 0732.62056

[51] Witten, DM; Friedman, JH; Simon, N., New insights and faster computations for the graphical Lasso, J Comput Graph Stat, 20, 892-900, (2011)· doi:10.1198/jcgs.2011.11051a

[52] Xue, L.; Ma, S.; Zou, H., Positive-definite $\ell _1$-penalized estimation of large covariance matrices, J Am Stat Assoc, 107, 1480-1491, (2012) · Zbl 1258.62063 · doi:10.1080/01621459.2012.725386

[53] Yin, J.; Li, J., Adjusting for high-dimensional covariates in sparse precision matrix estimation by $\ell _1$-penalization, J Multivar Anal, 116, 365-381, (2013) · Zbl 1277.62146 · doi:10.1016/j.jmva.2013.01.005

[54] Yuan, M., High dimensional inverse covariance matrix estimation via linear programming, J Mach Learn Res, 11, 2261-2286, (2010) · Zbl 1242.62043

[55] Yuan, M.; Lin, Y., Model selection and estimation in the Gaussian graphical model, Biometrika, 94, 19-35, (2007) · Zbl 1142.62408 · doi:10.1093/biomet/asm018

[56] Zerenner, T.; Friederichs, P.; Lehnertz, K.; Hense, A., A Gaussian graphical model approach to climate networks. Chaos: an interdisciplinary, J Nonlinear Sci, 24, 023103, (2014) · Zbl 1345.86008

[57] Zhang C-H, Huang J (2008) The sparsity and bias of the Lasso selection in high-dimensional linear regression. Ann Stat 36(4):1567-1594 · Zbl 1142.62044

[58] Zhang, T.; Zou, H., Sparse precision matrix estimation via Lasso penalized D-trace loss, Biometrika, 88, 1-18, (2014) · Zbl 1285.62063

[59] Zou, H., The adaptive Lasso and its oracle properties, J Am Stat Assoc, 101, 1418-1429, (2006) · Zbl 1171.62326 · doi:10.1198/016214506000000735