

**Fan, Jianqing; Feng, Yang; Wu, Yichao**

**Network exploration via the adaptive LASSO and SCAD penalties.** (English) Zbl 1166.62040  
*Ann. Appl. Stat.* 3, No. 2, 521-541 (2009).

Summary: Graphical models are frequently used to explore networks, such as genetic networks, among a set of variables. This is usually carried out via exploring the sparsity of the precision matrix of the variables under consideration. Penalized likelihood methods are often used in such explorations. Yet, positive-definiteness constraints of precision matrices make the optimization problem challenging.

We introduce nonconcave penalties and the adaptive LASSO penalty to attenuate the bias problem in the network estimation. Through the local linear approximation to the nonconcave penalty functions, the problem of precision matrix estimation is recast as a sequence of penalized likelihood problems with a weighted  $L_1$  penalty and solved using the efficient algorithm of *J. Friedman* et al. [*Biostatistics* 9, 432–441 (2008; [Zbl 1143.62076](#))]. Our estimation schemes are applied to two real datasets. Simulation experiments and asymptotic theory are used to justify our proposed methods.

**MSC:**

62H12 Estimation in multivariate analysis  
65C05 Monte Carlo methods  
65C60 Computational problems in statistics (MSC2010)

Cited in **73** Documents

**Keywords:**

adaptive LASSO; covariance selection; Gaussian concentration graphical model; genetic network; LASSO; precision matrix; SCAD; telephone call center data; breast cancer data

**Software:**

MIM; glasso; HdBCS

**Full Text:** [DOI](#)

**References:**

- [1] Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A. F. and Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: An overview. *Bioinformatics* 16 412-424. · [Zbl 0992.92024](#)
- [2] Breiman, L. (1996). Heuristics of instability and stabilization in model selection. *Ann. Statist.* 24 2350-2383. · [Zbl 0867.62055](#) · [doi:10.1214/aos/1032181158](#)
- [3] d'Aspremont, A., Banerjee, O. and Ghaoui, L. E. (2008). First-order methods for sparse covariance selection. *SIAM J. Matrix Anal. Appl.* 30 56-66. · [Zbl 1156.90423](#) · [doi:10.1137/060670985](#)
- [4] Dempster, A. P. (1972). Covariance selection. *Biometrics* 28 157-175.
- [5] Dobra, A., Hans, C., Jones, B., Nevins, J. R., Yao, G. and West, M. (2004). Sparse graphical models for exploring gene expression data. *J. Multivariate Anal.* 90 196-212. · [Zbl 1047.62104](#) · [doi:10.1016/j.jmva.2004.02.009](#)
- [6] Drton, M. and Perlman, M. (2004). Model selection for Gaussian concentration graphs. *Biometrika* 91 591-602. · [Zbl 1108.62098](#) · [doi:10.1093/biomet/91.3.591](#)
- [7] Edwards, D. M. (2000). *Introduction to Graphical Modelling*. Springer, New York. · [Zbl 0952.62003](#) · [doi:10.1007/978-1-4612-0493-0](#)
- [8] Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R. (2004). Least angle regression (with discussions). *Ann. Statist.* 32 409-499. · [Zbl 1091.62054](#) · [doi:10.1214/009053604000000067](#)
- [9] Fan, J. (1997). Comment on "Wavelets in statistics: A review," by A. Antoniadis. *J. Italian Statist. Soc.* 6 131-138.
- [10] Fan, J. and Fan, Y. (2008). High-dimensional classification using features annealed independence rules. *Ann. Statist.* 36 2605-2637. · [Zbl 1360.62327](#) · [doi:10.1214/07-AOS504](#)
- [11] Fan, J., Feng, Y. and Wu, Y. (2008). Supplement to "Network exploration via the adaptive LASSO and SCAD penalties." DOI: 10.1214/08-AOAS215SUPP.
- [12] Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.* 96 1348-1360. · [Zbl 1073.62547](#) · [doi:10.1198/016214501753382273](#)
- [13] Fan, J. and Peng, H. (2004). Nonconcave penalized likelihood with a diverging number of parameters. *Ann. Statist.* 32 928-961. · [Zbl 1092.62031](#) · [doi:10.1214/009053604000000256](#)

- [14] Friedman, J., Hastie, T. and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9 432-441. · [Zbl 1143.62076](#) · [doi:10.1093/biostatistics/kxm045](#)
- [15] Hess, R. K., Anderson, K., Symmans, W. F., Valero, V., Ibrahim, N., Mejia, J. A., Booser, D., Theriault, R. L., Buzdar, A. U., Dempsey, P. J., Rouzier, R., Sneige, N., Ross, J. S., Vidaurre, T., Gomez, H. L., Hortobagyi, G. N. and Pusztai, L. (2006). Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *Journal of Clinical Oncology* 24 4236-4244.
- [16] Huang, J., Liu, N., Pourahmadi, M. and Liu, L. (2006). Covariance matrix selection and estimation via penalised normal likelihood. *Biometrika* 93 85-98. · [Zbl 1152.62346](#) · [doi:10.1093/biomet/93.1.85](#)
- [17] Hunter, D. R. and Li, R. (2005). Variable selection using mm algorithm. *Ann. Statist.* 33 1617-1642. · [Zbl 1078.62028](#) · [doi:10.1214/009053605000000200](#)
- [18] Kuerer, H. M., Newman, L. A., Smith, T. L. et al. (1999). 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.
- [19] Lam, C. and Fan, J. (2008). Sparsistency and rates of convergence in large covariance matrices estimation. Manuscript. · [Zbl 1191.62101](#)
- [20] Levina, E., Zhu, J. and Rothman, A. J. (2008). Sparse estimation of large covariance matrices via a nested LASSO penalty. *Ann. Appl. Statist.* 2 245-263. · [Zbl 1137.62338](#) · [doi:10.1214/07-AOAS139](#)
- [21] Li, H. and Gui, J. (2006). Gradient directed regularization for sparse Gaussian concentration graphs, with applications to inference of genetic networks. *Biostatistics* 7 302-317. · [Zbl 1169.62378](#) · [doi:10.1093/biostatistics/kxj008](#)
- [22] Lin, S. P. and Perlman, M. D. (1985). A Monte Carlo comparison of four estimators of a covariance matrix. *Multivariate Anal.* 6 411-429. · [Zbl 0593.62051](#)
- [23] Mardia, K. V., Kent, J. T. and Bibby, J. M. (1979). *Multivariate Analysis*. Academic Press, New York. · [Zbl 0432.62029](#)
- [24] Meinshausen, N. and Bühlmann, P. (2006). High-dimensional graphs with the lasso. *Ann. Statist.* 34 1436-1462. · [Zbl 1113.62082](#) · [doi:10.1214/009053606000000281](#)
- [25] Rothman, A. J., Bickel, P. J., Levina, E. and Zhu, J. (2008). Sparse permutation invariant covariance estimation. *Electron. J. Statist.* 2 494-515. · [Zbl 1320.62135](#) · [doi:10.1214/08-EJS176](#)
- [26] Schäfer, J. and Strimmer, K. (2005). An empirical Bayes approach to inferring large-scale gene association networks. *Bioinformatics* 21 754-764.
- [27] Shen, H. and Huang, J. (2005). Analysis of call centre arrival data using singular value decomposition. *Appl. Stoch. Models Bus. Ind.* 21 251-263. · [Zbl 1089.62155](#) · [doi:10.1002/asmb.598](#)
- [28] Tibshirani, R. J. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B* 58 267-288. · [Zbl 0850.62538](#)
- [29] Vandenberghe, L., Boyd, S. and Wu, S.-P. (1998). Determinant maximization with linear matrix inequality constraints. *SIAM J. Matrix Anal. Appl.* 19 499-533. · [Zbl 0959.90039](#) · [doi:10.1137/S0895479896303430](#)
- [30] Wong, F., Carter, C. K. and Kohn, R. (2003). Efficient estimation of covariance selection models. *Biometrika* 90 809-830. · [Zbl 1436.62346](#) · [doi:10.1093/biomet/90.4.809](#)
- [31] Yuan, M. and Lin, Y. (2007). Model election and estimation in the Gaussian graphical model. *Biometrika* 94 19-35. · [Zbl 1142.62408](#) · [doi:10.1093/biomet/asm018](#)
- [32] Zou, H. (2006). The adaptive lasso and its oracle properties. *J. Amer. Statist. Assoc.* 101 1418-1429. · [Zbl 1171.62326](#) · [doi:10.1198/016214506000000735](#)
- [33] Zou, H. and Li, R. (2008). One-step sparse estimates in nonconcave penalized likelihood models (with discussion). *Ann. Statist.* 36 1509-1566. · [Zbl 1282.62112](#) · [doi:10.1214/009053607000000802](#)

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.