

References

- Beck, A., and Teboulle, M. 2009a. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sciences* 2(1):183–202.
- Beck, A., and Teboulle, M. 2009b. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* 2(1):183–202.
- Berwin A. Turlach, William N. Venables, S. J. W. 2005. Simultaneous variable selection. In *Technometrics*, 349–363.
- Boyd, S., and Vandenberghe, L. 2004. *Convex Optimization*. Cambridge University Press.
- Chang, X.; Shen, H.; Wang, S.; Liu, J.; and Li, X. 2014. Semi-supervised feature analysis for multimedia annotation by mining label correlation. In *Advances in Knowledge Discovery and Data Mining*, 74–85.
- Chartrand, R., and Staneva, V. 2008. Restricted isometry properties and nonconvex compressive sensing. *Inverse Problems* 24(035020):1–14.
- Ding, C., and Peng, H. 2003. Minimum redundancy feature selection from microarray gene expression data. In *J Bioinform Comput Biol*, 523–529.
- Fan, J., and Li, R. 2001. Variable selection via nonconcave penalized likelihood and its oracle properties. 96(456):1348–1359.
- Hastie, T.; Tibshirani, R.; and Friedman, J. H. 2001. *Elements of Statistical Learning*. Springer.
- Huang, J.; Horowitz, J.; and Ma, S. 2008. Asymptotic properties of bridge estimators in sparse high-dimensional regression models. *Annals of Statistics* 36(2):587–613.
- Kohavi, R., and John, G. H. 1997. Wrappers for feature subset selection. *ARTIFICIAL INTELLIGENCE* 97(1):273–324.
- Kong, D., and Ding, C. H. Q. 2013. Efficient algorithms for selecting features with arbitrary group constraints via group lasso. In *ICDM*, 379–388.
- Kong, D.; Ding, C. H. Q.; Huang, H.; and Zhao, H. 2012. Multi-label relief and f-statistic feature selections for image annotation. In *CVPR*, 2352–2359.
- Kong, D.; Ding, C. H. Q.; and Huang, H. 2011. Robust nonnegative matrix factorization using l21-norm. In *CIKM*, 673–682.
- Kong, D.; Zhang, M.; and Ding, C. H. Q. 2013. Minimal shrinkage for noisy data recovery using Schatten-p norm objective. In *ECML/PKDD (2)*, 177–193.
- Kononenko, I. 1994. Estimating attributes: Analysis and extensions of relief. 171–182. Springer Verlag.
- Langley, P. 1994. Selection of relevant features in machine learning. In *In Proceedings of the AAAI Fall symposium on relevance*, 140–144. AAAI Press.
- Liu, H., and Motoda, H. 1998. *Feature Selection for Knowledge Discovery and Data Mining*. Springer.
- Liu, J.; Ji, S.; and Ye, J. 2009. Multi-task feature learning via efficient l2, 1-norm minimization. In *UAI*, 339–348.
- Liu, H.; Palatucci, M.; and Zhang, J. 2009. Blockwise coordinate descent procedures for the multi-task lasso, with applications to neural semantic basis discovery. In *ICML*.
- Masaeli, M.; Fung, G.; and Dy, J. G. 2010. From transformation-based dimensionality reduction to feature selection. In *ICML*.
- Nesterov, Y. 2004. *Introductory Lectures on Convex Optimization: A Basic Course*. Kluwer Academic Publishers.
- Oliva, A., and Torralba, A. 2001. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision* 42(3):145–175.
- Quattoni, A.; Collins, M.; and Darrell, T. 2008. Transfer learning for image classification with sparse prototype representations. In *CVPR*.
- Schmidt, M.; Murphy, K.; Fung, G.; and Rosales, R. 2008. Structure learning in random fields for heart motion abnormality detection. In *CVPR*.
- Tropp, J. A.; Gilbert, A. C.; Martin, J.; Strauss, J.; Tropp, J. A.; Gilbert, A. C.; and Strauss, M. J. 2006. Algorithms for simultaneous sparse approximation. part ii: Convex relaxation. *Signal Processing* 589–602.
- Tsoumakas, G.; Katakis, I.; and Vlahavas, I. 2010. Mining multi-label data. *Data Mining and Knowledge Discovery Handbook*.
- Turlach, B. A.; Venables, W. N.; and Wright, S. J. 2005. Simultaneous variable selection. In *Technometrics*, 349–363.
- Vogt, J. E., and Roth, V. 2012. A complete analysis of the $l_{1,p}$ group-lasso.
- Yang, Y. 1999. An evaluation of statistical approaches to text categorization. *Journal of Information Retrieval* 1:67–88.
- Yuan, L.; Liu, J.; and Ye, J. 2011. Efficient methods for overlapping group lasso. In *NIPS*, 352–360.
- Zhang, Y. 2011. A probabilistic framework for learning task relationships in multi-task learning. In *Ph.D Thesis, The Hong Kong University of Science and Technology*.