

- J. Abernethy, F. Bach, T. Evgeniou, and J.-P. Vert. A new approach to collaborative filtering: Operator estimation with spectral regularization. *Journal of Machine Learning Research*, 10:803–826, 2009.
- Y. Amit, M. Fink, N. Srebro, and S. Ullman. Uncovering shared structures in multiclass classification. In *Proceedings of the 24th international conference on Machine Learning (ICML)*, 2007.
- C. Archambeau and F. Bach. Sparse probabilistic projections. In *Advances in Neural Information Processing Systems 21 (NIPS)*, 2008.
- A. Argyriou, C.A. Micchelli, and M. Pontil. On spectral learning. *Journal of Machine Learning Research*, 2009.
- F. Bach. High-Dimensional Non-Linear Variable Selection through Hierarchical Kernel Learning. Technical Report 0909.0844, arXiv, 2009a.
- F. Bach. Exploring large feature spaces with hierarchical multiple kernel learning. In *Advances in Neural Information Processing Systems*, 2008c.
- F. Bach. Consistency of trace norm minimization. *Journal of Machine Learning Research*, 9:1019–1048, 2008d.
- F. Bach, G. R. G. Lanckriet, and M. I. Jordan. Multiple kernel learning, conic duality, and the SMO

- algorithm. In Proceedings of the International Conference on Machine Learning (ICML), 2004a.
- F. Bach, R. Thibaux, and M. I. Jordan. Computing regularization paths for learning multiple kernels. In Advances in Neural Information Processing Systems 17, 2004b.
- F. Bach, J. Mairal, and J. Ponce. Convex sparse matrix factorizations. Technical Report 0812.1869, ArXiv, 2008.
- R. Baraniuk. Compressive sensing. IEEE Signal Processing Magazine, 24(4):118–121, 2007.
- R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde. Model-based compressive sensing. Technical report, arXiv:0808.3572, 2008.
- L. Bottou and O. Bousquet. The tradeoffs of large scale learning. In Advances in Neural Information Processing Systems (NIPS), volume 20, 2008.
- W. Buntine and S. Perttu. Is multinomial PCA multi-faceted clustering or dimensionality reduction. In International Workshop on Artificial Intelligence and Statistics (AISTATS), 2003.
- E. Candès and M. Wakin. An introduction to compressive sampling. IEEE Signal Processing Magazine, 25(2):21–30, 2008.
- E.J. Candès and Y. Plan. Matrix completion with noise. 2009a.

- S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Review*, 43(1):129–159, 2001.
- A. d’Aspremont, El L. Ghaoui, M. I. Jordan, and G. R. G. Lanckriet. A direct formulation for sparse PCA using semidefinite programming. *SIAM Review*, 49(3):434–48, 2007.
- A. d’Aspremont, F. Bach, and L. El Ghaoui. Optimal solutions for sparse principal component analysis. *Journal of Machine Learning Research*, 9:1269–1294, 2008.
- D.L. Donoho and J. Tanner. Neighborliness of randomly projected simplices in high dimensions. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27):9452, 2005.
- M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736–3745, 2006.
- M. Fazel, H. Hindi, and S.P. Boyd. A rank minimization heuristic with application to minimum order system approximation. In *Proceedings of the American Control Conference*, volume 6, pages 4734–4739, 2001.
- C. Févotte, N. Bertin, and J.-L. Durrieu. Nonnegative matrix factorization with the itakura-saito divergence. with application to music analysis. *Neural Computation*, 21(3), 2009.
- R. Jenatton, G. Obozinski, and F. Bach. Structured sparse principal component analysis. Technical report, arXiv:0909.1440, 2009b.

- S. Lacoste-Julien, F. Sha, and M.I. Jordan. DiscLDA: Discriminative learning for dimensionality reduction and classification. *Advances in Neural Information Processing Systems (NIPS)* 21, 2008.
- G. R. G. Lanckriet, N. Cristianini, L. El Ghaoui, P. Bartlett, and M. I. Jordan. Learning the kernel matrix with semidefinite programming. *Journal of Machine Learning Research*, 5:27–72, 2004b.
- L. Mackey. Deflation methods for sparse PCA. *Advances in Neural Information Processing Systems (NIPS)*, 21, 2009.
- J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Supervised dictionary learning. *Advances in Neural Information Processing Systems (NIPS)*, 21, 2009b.
- J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In *International Conference on Computer Vision (ICCV)*, 2009c.
- R. Mazumder, T. Hastie, and R. Tibshirani. *Spectral Regularization Algorithms for Learning Large Incomplete Matrices*. 2009.
- C.A. Micchelli and M. Pontil. Learning the kernel function via regularization. *Journal of Machine Learning Research*, 6(2):1099, 2006.
- B. Moghaddam, Y. Weiss, and S. Avidan. Spectral bounds for sparse PCA: Exact and greedy algorithms. In *Advances in Neural Information Processing Systems*, volume 18, 2006b.
- G. Obozinski, B. Taskar, and M.I. Jordan. Joint covariate selection and joint subspace selection for multiple classification problems. *Statistics and Computing*, pages 1–22, 2009.

- M. Pontil, A. Argyriou, and T. Evgeniou. Multi-task feature learning. In *Advances in Neural Information Processing Systems*, 2007.
- R. Raina, A. Battle, H. Lee, B. Packer, and A.Y. Ng. Self-taught learning: Transfer learning from unlabeled data. In *Proceedings of the 24th International Conference on Machine Learning (ICML)*, 2007.
- A. Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet. SimpleMKL. *Journal of Machine Learning Research*, 9:2491–2521, 2008.
- B. Recht, W. Xu, and B. Hassibi. Null Space Conditions and Thresholds for Rank Minimization. 2009.
- J. D. M. Rennie and N. Srebro. Fast maximum margin matrix factorization for collaborative prediction. In *Proceedings of the 22nd international conference on Machine Learning (ICML)*, 2005.
- R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. In *Advances in Neural Information Processing Systems*, volume 20, 2008.
- S. Sonnenburg, G. Raetsch, C. Schaefer, and B. Schoelkopf. Large scale multiple kernel learning. *Journal of Machine Learning Research*, 7:1531–1565, 2006.
- N. Srebro, J. D. M. Rennie, and T. S. Jaakkola. Maximum-margin matrix factorization. In *Advances in Neural Information Processing Systems 17*, 2005.
- B. K. Sriperumbudur, D. A. Torres, and G. R. G. Lanckriet. A d.c. programming approach to the sparse generalized eigenvalue problem. Technical Report 0901.1504v2, ArXiv, 2009.
- E. van den Berg, M. Schmidt, M. P. Friedlander, and K. Murphy. Group sparsity via linear-time projection. Technical Report TR-2008-09, Department of Computer Science, University of British Columbia, 2009.
- D.M. Witten, R. Tibshirani, and T. Hastie. A penalized matrix decomposition, with applications to

sparse principal components and canonical correlation analysis. *Biostatistics*, 10(3):515–534, 2009.

H. Zou and R. Li. One-step sparse estimates in nonconcave penalized likelihood models. *Annals of Statistics*, 36(4):1509–1533, 2008.