Distance metric learning Vs. Fisher discriminant analysis

Ali Ghodsi* Michael Biggs† Babak Alipanahi ‡

In many fundamental machine learning problems, the Euclidean distances between data points do not represent the desired topology that we are trying to capture. Kernel methods address this problem by mapping the points into new spaces where Euclidean distances may be more useful. An alternative approach is to construct a Mahalanobis distance (quadratic Gaussian metric) over the input space and use it in place of Euclidean distances. This approach can be equivalently interpreted as a linear transformation of the original inputs, followed by Euclidean distance in the projected space. This approach has attracted a lot of recent interest [2, 3, 4, 5].

Some of the proposed algorithms are iterative and computationally expensive. In this paper, We propose a closed-form solution to one algorithm that previously required expensive semidefinite optimization. We provide a new problem setup in which the algorithm performs better or as well as some standard methods, but without the computational complexity. Furthermore, we show a strong relationship between these methods and the Fisher Discriminant Analysis (FDA)[1]. We also extend the approach by kernelizing it, allowing for non-linear transformations of the metric.

References

[1] Ronald A. Fisher. The use of multiple measurements in taxonomic problems. *Annals Eugen.*, 7:179–188, 1936.

[2] Ali Ghodsi, Dana F. Wilkinson, and Finnegan Southey. Improving embeddings by flexible exploitation of side information. In Manuela M. Veloso, editor, *International Joint Conference on Artificial Intelligence*, pages 810–816, Hyderabad, India, 2007.

*Department of Statistics, University of Waterloo, 200 University Ave. W., Waterloo, Ontario, Canada N2L 3G1, aghodsib@uwaterloo.ca.

†Department of Statistics, University of Waterloo, 200 University Ave. W., Waterloo, Ontario, Canada N2L 3G1, mike@doubleplum.net.

‡School of Computer Science, University of Waterloo, 200 University Ave. W., Waterloo, Ontario, Canada N2L 3G1, vavasis@math.uwaterloo.ca.

[3] Amir Globerson and Sam Roweis. Metric learning by collapsing classes. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *Advances in Neural Information Processing Systems 18*, pages 451–458, Cambridge, MA, 2006. MIT Press.

[4] Kilian Weinberger, John Blitzer, and Lawrence Saul. Distance metric learning for large margin nearest neighbor classification. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *Advances in Neural Information Processing Systems 18*, pages 1473–1480, Cambridge, MA, 2006. MIT Press.

[5] Eric P. Xing, Andrew Y. Ng, Michael I. Jordan, and Stuart Russell. Distance metric learning with application to clustering with side-information. In S. Thrun S. Becker and K. Obermayer, editors, *Advances in Neural Information Processing Systems 15*, pages 505–512, Cambridge, MA, 2003. MIT Press.