

A Toolbox for Bicluster Analysis in R

Sebastian Kaiser and Friedrich Leisch

Department of Statistics, Ludwig-Maximilians-Universität München,
Ludwigstrasse 33, 80539 München, Germany,
firstname.lastname@stat.uni-muenchen.de

Abstract. Over the last decade, bicluster methods have become more and more popular in different fields of two way data analysis and a wide variety of algorithms and analysis methods have been published. In this paper we introduce the R package `biclust`, which contains a collection of bicluster algorithms, preprocessing methods for two way data, and validation and visualization techniques for bicluster results. For the first time, such a package is provided on a platform like R, where data analysts can easily add new bicluster algorithms and adapt them to their special needs.

Keywords: Biclustering, Two-Way-Clustering, Software, R

References

- BARKOW, S., BLEULER, S., PRELIC, A., ZIMMERMANN, P., and ZITZLER, E. (2006): Bicat: a biclustering analysis toolbox. *Bioinformatics*, 22,1282–1283.
- CHENG, Y. and CHURCH, G. M. (2000): Biclustering of expression data. In: *Proceedings of the Eighth International Conference on Intelligent Systems for Molecular Biology*, 1,93–103.
- KLUGER, Y., BASRI, R., CHANG, J. T., and GERSTEIN, M. (2003): Spectral biclustering of microarray data: Coclustering genes and conditions. *Genome Research*, 13,703–716.
- MADEIRA, S. C. and OLIVEIRA, A. L. (2004): Biclustering algorithms for biological data analysis: A survey. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 1(1),24–45.
- VAN MECHELEN, I. and SCHEPERS, J. (2006): A unifying model for biclustering. In: *Compstat 2006 - Proceedings in Computational Statistics*, 81–88.
- MURALI, T. and KASIF, S. (2003): Extracting conserved gene expression motifs from gene expression. In: *Pacific Symposium on Biocomputing*, 8,77–88.
- PRELIC, A., BLEULER, S., ZIMMERMANN, P., WIL, A., BÜHLMANN, P., GRUISSEM, W., HENNING, L., THIELE, L., and ZITZLER, E. (2006): A systematic comparison and evaluation of biclustering methods for gene expression data. *Bioinformatics*, 22(9),1122–1129.
- SANTAMARIA, R., THERON, R., and QUINTALES, L. (2007): A framework to analyze biclustering results on microarray experiments. In: *8th International Conference on Intelligent Data Engineering and Automated Learning (IDEAL'07)*, Springer, Berlin, 770–779.
- TURNER, H., BAILEY, T., and KRZANOWSKI, W. (2005): Improved biclustering of microarray data demonstrated through systematic performance tests. *Computational Statistics and Data Analysis*, 48,235–254.