

Readings for Module 5: Dimension Reduction.

- N.S. Holter, M. Mitra, A. Maritan, M. Cieplak, J.R. Banavar, and N.V. Fedoroff (2000). Fundamental patterns underlying gene expression profiles: Simplicity from complexity. *PNAS* 97: 8409-8414.
- O. Alter, P.O. Brown, and D. Botstein (2000). Singular value decomposition for genome-wide expression data processing and modeling. *PNAS* 97: 10101-10106.
- Misra J., Schmitt W., Hsiao Li-Li, Gullans S., Stephanopoulos G., Stephanopoulos G. (2002). Interactive exploration of microarray gene expression patterns in a reduced dimensional space. *Genome Research* 12:1112-1120.
- J. Landgrebe, W. Wurst, G. Welzl (2002). Permutation-validated principal components analysis of microarray data. *Genome Biology* 3(4): research 0019.1-0019.11.
- D Ghosh (2002). Resampling methods for variance estimation of singular value decomposition analyses from microarray experiments. *Functional & Integrative Genomics* 2(3): 92-97
- Wolfram Liebermeister (2002). Linear modes of gene expression determined by independent component analysis. *Bioinformatics* 18:51-60.
- O Alter, PO Brown, D Botstein (2003). Generalized singular value decomposition for comparative analysis of genome-scale expression data sets of two different organisms. *PNAS* 100(6):3351-3356.

Exploration and graphics:

- C. Posse (1995). Tools for two-dimensional exploratory projection pursuit. *JCGS*, 4(2).
- A. Buja, Cook D., Swayne D.F. (1996) Interactive high dimensional data visualization. *JCGS*, 5(1).

Non-linear dimension reduction:

- Hastie T., W. Stuetzle (1989) Principal Curves. *Journal of the American Statistical Association* 84 (406), 502-516.
- J. Tenenbaum, V. De Silva, J.C. Langford (2000). A global geometric framework for non-linear dimensionality reduction. *Science* 290, 2319-2323.
- S.T. Roweis, L.K. Saul (2000). Non-linear dimensionality reduction by locally linear embedding. *Science* 290, 2323-2326.

Classification & Regression – Supervised dimension reduction:

- Dudoit S., Jane Fridlyand, Speed T.P. (2002). Comparison of discrimination methods for the classification of tumors using gene expression data. *Journal of the American Statistical Association* 97(457):77-87.
- Chiaromonte F., Martinelli J.A. (2002). Dimension reduction strategies for analyzing global gene expression data with a response. *Math'l Biosciences* 176(1):123-144.
- Antoniadis A., Lambert-Lacroix S., Leblanc F. (2003). Effective dimension reduction methods for tumor classification using gene expression data. *Bioinformatics* 19:563-570. [[webpage from the authors: www-lmc.imag.fr/SMS/software/microarrays/](http://www-lmc.imag.fr/SMS/software/microarrays/)]
- Bura E., Pfeiffer R.M. (2003). Graphical methods for class prediction using dimension reduction techniques on DNA microarray data. *Bioinformatics* 19:1252-1258.