Quantum Clustering of Microarray Data in a Truncated SVD Space

Axel, I. and Horn, D.
School of Physics and Astronomy, Tel Aviv University

We describe the application of a novel clustering method to microarray expression data. Its first stage involves compression of dimensions that can be achieved by applying SVD to the gene-sample matrix in microarray problems.

Thus the data (samples or genes) can be represented by vectors in a truncated space of low dimensionality, 4 and 5 in the examples studied here. We find it preferable to project all vectors onto the unit sphere before applying our clustering algorithm, the quantum clustering method. Although the method is not hierarchical, it can be modified to allow hierarchy in terms of its free scale parameter.

We test our method on three data sets and obtain promising results. On cancer cell data we obtain a dendrogram that reflects correct groupings of cells. In an AML/ALL data set we obtain very good clustering of samples into four classes of the data. Finally, in clustering of genes in yeast cell cycle data we obtain four groups in a problem that is estimated to contain five families.