Bi-Clustering Algorithm

**Objective**

The goal of this project was to implement a bi-clustering algorithm that partitions data into, well, clusters. For this study, we took the following matrix:

Methodology

*The algorithm was as follows:*

Removing single rows or columns according to…

1. Calculate contribution of every row and every column to *H*and remove the row/column with the corresponding largest  
   contribution

* Contribution of row *m*:
* Contribution of column *i*:

1. Repeat step 1 until *H < δ*

The algorithm calculates an optimal δ-cluster of a matrix of gene expression values. The algorithm terminates if either the remaining matrix consists of only one row or column or if the heterogeneity H is smaller than the predefined threshold δ. The algorithm was applied using δ = 0.1:

In every iteration of the algorithm, I have printed which row or column is being deleted along with the final resulting submatrix.