

GenePattern

Principal Components Analysis Documentation

Module name: PCA
Description: Principal Components Analysis
Author: Joshua Gould (Broad Institute) gp-help@broad.mit.edu
The Institute for Genomics Research
Date: October 22, 2003
Release: 1.0

Performs a principal components analysis. This implementation is an adaptation of TIGR's MultiExperiment Viewer PCA module (<http://www.tm4.org/mev/>).

References:

Raychaudhuri S, Stuart JM, and Altman RB. (2000) Principal components analysis to summarize microarray experiments: application to sporulation time series. Pac. Symp. Biocomput. 455-466.

Parameters:

Name	Description
input_filename	The input file (.res, .gct, .odf)
cluster_by	Cluster by rows or columns
output_file	The stub for the output files

Return Value:

1. s matrix The s matrix file, the eigen vector 1/ Principal Component
2. t matrix The t matrix file, the eigen values 1/Principal Component
3. u matrix The u matrix file, covariance metrics for the raw data (ie component scores matrix).
4. Stdout.txt: the "stdout" text output from running the program.

Platform dependencies:

Task type:	Projection
CPU type:	any
OS:	any
Java JVM level:	1.3
Language:	Java