Optimization of PCA parameters:

|  |
| --- |
| Cangelosi, Richard, and Alain Goriely. "Component retention in principal component analysis with application to cDNA microarray data." *Biology direct* 2.1 (2007): 1. |
| **Kaiser method**  Retain components with eigenvalues greater than 1.  **Scree test**  The ideal pattern in a scree plot is a steep curve, followed by a bend and then a flat or horizontal line. Retain those components or factors in the steep curve before the first point that starts the flat line trend. You might have difficulty interpreting a scree plot. Use your knowledge of the data and the results from the other methods of selecting components or factors to help decide the number of important components or factors.  **Percentage of variation explained**  Retain components that cumulatively explain a certain percentage of variation. The acceptable level of explained variance depends on how you use Principal Components. For descriptive purposes, you might only need 80% of the variance explained. However, if you are doing other analyses on these data, you might want to have at least 90% of the variance explained.  <http://papers.nips.cc/paper/1853-automatic-choice-of-dimensionality-for-pca.pdf>  Automatic Choice of Dimensionality for PCA |  |