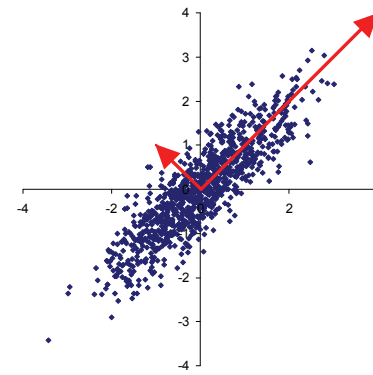


Principal Component Analysis

F. Edward Boas
2001-08-30

Principal Component Analysis



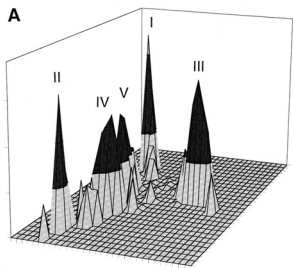
Data points
Principal components

Given a set of data points in a high-dimensional space, PCA:

- Identifies the major independent directions of variation
- Allows you to reduce the dimensionality of the data set

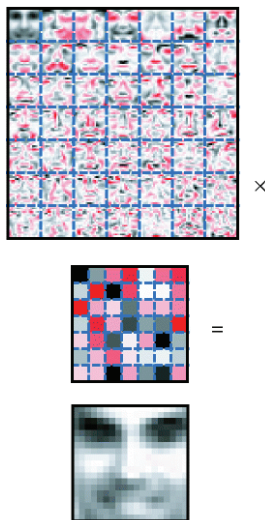
Examples of PCA

Protein classification



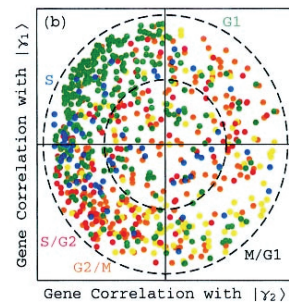
Science (1996) 273:595

Face recognition



Nature (1999) 401: 788

Cell-cycle genes



PNAS (2000) 97: 10101