**Principal Components Analysis, briefly**

To find the first principal component of a matrix $A$ one should find such linear combinations

$$Z^T = CA^T$$

of a row-vectors of $A$ that the column-vectors $\mathbf{z}_1, ..., \mathbf{z}_n$ of $Z$ have the maximal variance,

$$\max \sum_{i=1}^{n} D\mathbf{z}_1$$

with the conditions $||C||_2^2 = 1$ and $CC^T = I_n$.

As Rau, C.R. shows, the column vectors of the matrix $C$ is a eigenvectors of a matrix $\Sigma$. Elements of this matrix are covariance coefficients of row-vectors of $A$. The interpretation of the principal components is the next. Assume one should replace a $p$-dimensional random value with $k < p$ linear function with condition that he does not want to loss to much information. How to choose these $k$ linear functions? Any choice depends on the quality of the $p$ original random value reconstruction. One of the random variable reconstruction method is to make the best predictor on a $k$ linear functions basis. In this case the predictor effectiveness can be estimated with the residual variance $\sigma_i^2$, $i \in \{1, ..k\}$. The overall effectiveness measure is $\sum \sigma_i^2$. The best choice of linear functions for which this sum is minimal is the choice of the first $k$ principal components.

A principal components computation procedure is the following. Make the covariation matrix $\Sigma = [\sigma_{jk}]$:

$$\sigma_{jk} = \frac{1}{n-1} \sum_{i=1}^{n} (a_{ij} - \bar{a}_{\cdot j})(a_{ik} - \bar{a}_{\cdot k}), \quad j, k = 1, ..., n.$$

Here the variables $a_{ij}, a_{ik}$ are the items of the matrix $A$. Notice that the variable $\bar{a}_{\cdot j}$ (and $\bar{a}_{\cdot k}$ as well) in the classic algorithm is $\frac{1}{n} \sum_{\chi=1}^{n} a_{\chi j}$ while in the case of the indices computation the variable $\bar{a}_{\chi j} = 0$ so that all the objects data $a_{ij}$ are positive. Find the eigenvector $\mathbf{c}_1 = [c_1, ..., c_n]^T$ of the matrix $\Sigma$ that corresponds to its maximal eigenvalue. For each object compute its index

$$q_i = c_{11}(a_{i1} - \bar{a}_{\cdot 1}) +, ..., + c_{n1}(a_{in} - \bar{a}_{\cdot n}).$$

From the matrix $A$ singular value decomposition $A = U\Lambda V^T$ one can deduce that

$$A^T A V^T = V^T \Lambda^2.$$

It means the matrices $V^T$ and $\Lambda^2$ are eigensystem of $A^T A = \Sigma$. So, assume the first principal component $\mathbf{c}_1$ as the first row vector of $V$. The indices vector $\mathbf{q} = [q_1, ..., q_m]^T$ is

$$\mathbf{q} = A\mathbf{c}_1.$$

FOR FURTHER READING

Rao, C. R. Linear statistical inference and its applications. Wiley, 1965.
Jackson, J.E., A User's Guide to Principal Components, Wiley, 1988.
Jolliffe, I.T. Principal Component Analysis, 2nd ed., Springer, 2002.
Krzanowski, W.J., Principles of Multivariate Analysis, Oxford University Press, 1988.
Seber, G.A.F., Multivariate Observations, Wiley, 1984.