The Eigen Brothers: Vector and Value

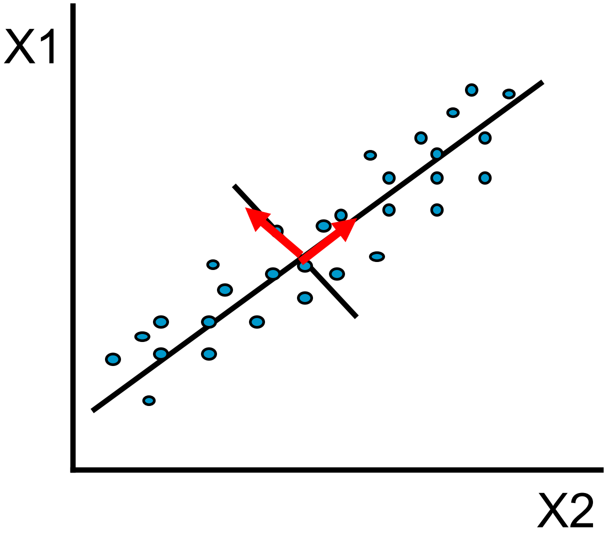
KOGOD SCHOOL of BUSINESS
AMERICAN UNIVERSITY • WASHINGTON, DC

Factor Extraction

- The goal of factor analysis is to reduce the variable space to a few more general variables by **extracting** factors that **explain** a **high %** of the **variance** in the data
- The factors are found by one of the many **"factor extraction"** methods involving axis **rotation** (e.g., principal components, maximum likelihood, principal axis, etc.)
- **Principal Components** is a very popular **factor extraction** (see PCA regression lecture), which is what we use in this class
- It is customary to first **group variables** based on **business intuition** and then do **factor analysis** on the data to **confirm** that the business grouping of variables into factors matches the statistical grouping, which is why this method is referred to as **"Confirmatory Factor Analysis"** (**CFA**)
- To understand factor extraction with PCA we need to briefly define two key concepts: **eigenvectors** and **eigenvalues**

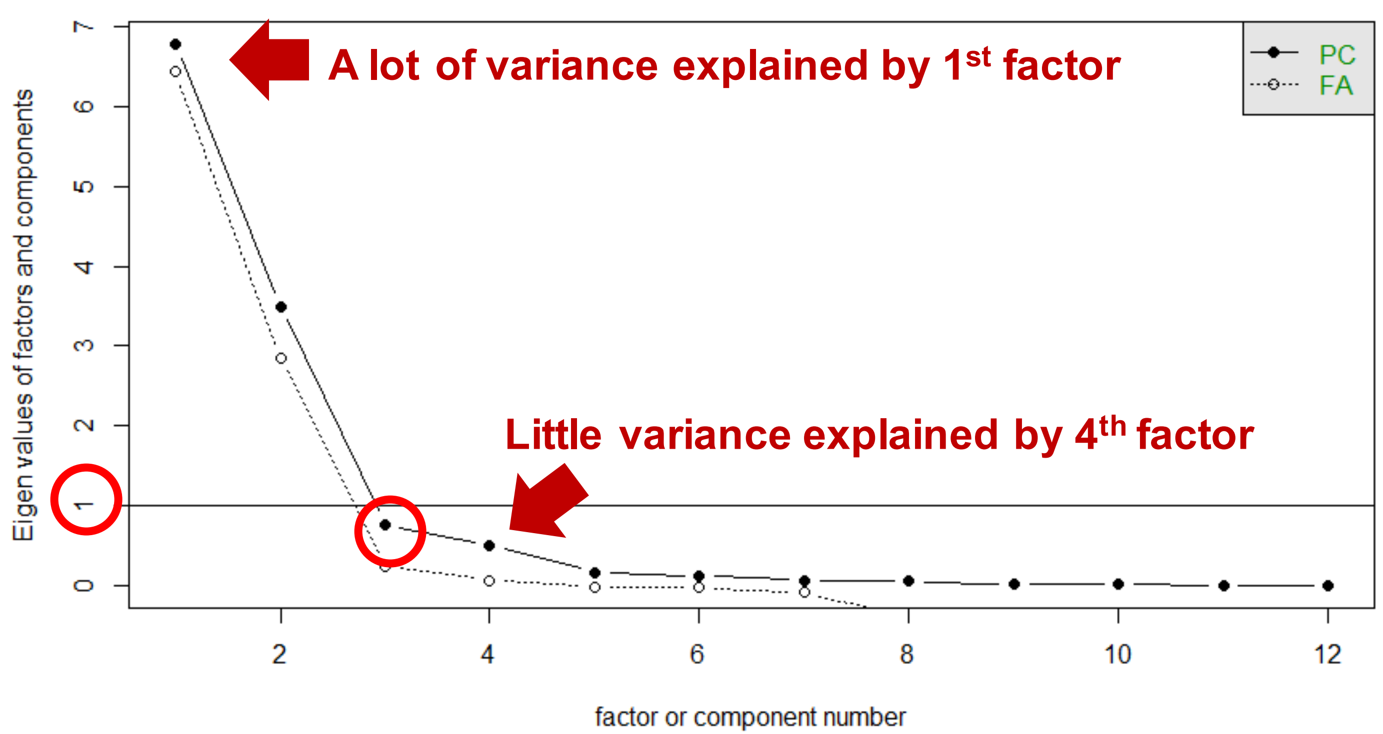
Eigenvectors and Eigenvalues

- **Eigenvectors** and **eigenvalues** are essential to understanding factor analysis using **PCA**.
- **Eigenvectors** are simply **vectors** of **length = 1 aligned** with the **principal components**. They provide the direction of the PC's
- **Eigenvalues** measure the **variance** of the data along each of the eigenvectors, but the variance is standardized so the **average** eigenvalue is **1** (like all standardized variables) – why this is important will become apparent shortly
- Think of **Eigenvalues** as the **size** of each principal component relative to each other



Scree Plot

- **Great tool** to identify the optimal **number of factors** to extract
- It plots the **eigenvalues** against the **number of factors**
- The optimal number of factors to use is around **Eigenvalue = 1** or around the **"elbow"** → where the additional variance explained by one more factor begins to bottom out
- The example below supports a **3-Factor solution**



R Tips

`scree(){psych}` → Principal components function in the `{psych}` package, which works well for Factor Analysis

`x.dataFrame=data.frame(dataName$x1,`
`dataName.x2, etc.)` → Most of the work in Factor Analysis is working with and rotating the covariance or correlation matrix, so the first order of business is to create a data frame with the predictor variables to factor analyze

`cor.x=cor(x.dataFrame)` → We then generate the corresponding correlation object

`scree(cor.x, factors=TRUE)` → Scree Plot

`eigen(){base}` → Function to extract eigenvalues and eigenvectors

`eigen.vect.x=eigen(cor.x)` → Generate eigenvectors

`eigen.val.x=eigen(cor.x, only.values = TRUE)` → And eigenvalues

KOGOD SCHOOL

of

BUSINESS