Principal Components Analysis

Lester Cajegas

March 10, 2018

Principal Components Analysis

PCA is a technique for reducing the dimension of a n x p data matrix $\mathbf{X}$. The *first principal component* direction of the data is that along which the observations *vary the most*.

$$N(\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(\frac{x-\mu}{\sigma})^2}$$

Normality Test

The normality test is conducted to validate the assumption of normality in the residuals. The Jarque-Bera test (`jb.norm.test`) in `normstest` package provide a good level of indication of the data's normality.

Table 1: Normality Test for the dataset

Method:	Jarque-Bera test for normality
Statistic:	4.48587543735093
p-value:	0.077

Plot

The values presented here reflects the shape of the native distribution of the data.

Table 2: Principal Components Loadings

	PC1	PC2	PC3	PC4
Sepal.Length	0.5038236	-0.4549987	0.7088547	0.1914757
Sepal.Width	-0.3023682	-0.8891442	-0.3311628	-0.0912541
Petal.Length	0.5767881	-0.0337880	-0.2192793	-0.7861873
Petal.Width	0.5674952	-0.0354563	-0.5829003	0.5804474

Table 3: Trial Table for caption

PC1	PC2	PC3	PC4
-2.406639	-0.3969554	0.1939647	0.0047795
-2.223539	0.6901804	0.3500015	0.0488684
-2.581105	0.4275418	0.0188976	0.0499095
-2.450869	0.6860074	-0.0687460	-0.1496465
-2.536853	-0.5082516	0.0293226	-0.0400482
-1.841495	-1.2899381	-0.2527683	0.1638906

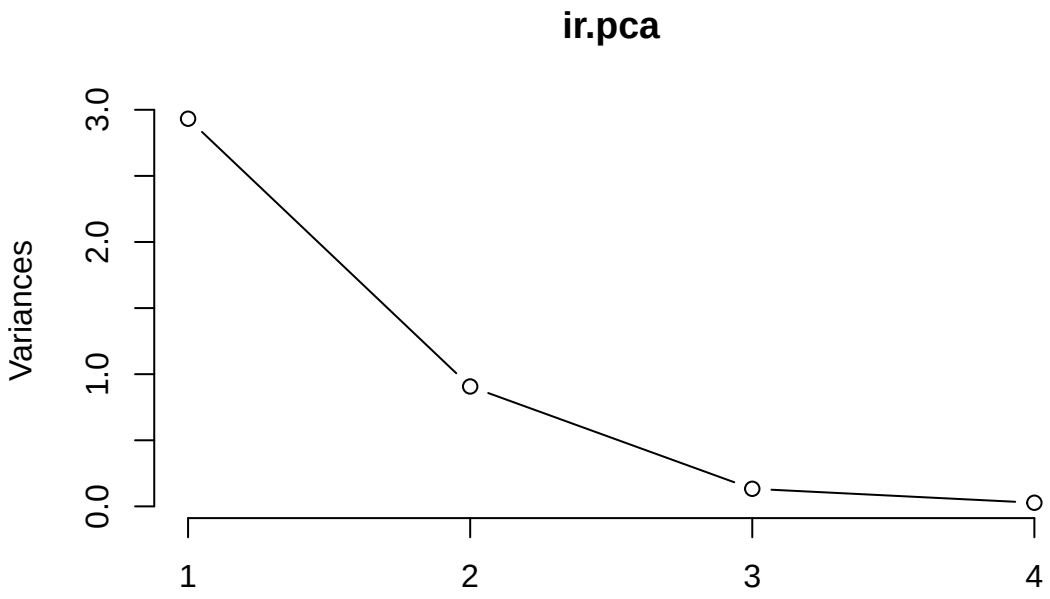

Figure 1: Eigenvalue plot