1. 对原始数据进行标准化处理.

Data standardization

假设进行主成分分析的指标变量有$m$个，分别为$x\_1$，$x\_2$,$\cdot$ ,$x\_ m$,共有$n$个评价对象，第$i$个评价对象的第$j$个指标的取值为$a\_{ij}$。将各指标的值a\_{ij}转换成标准化指标值，有$\widetilde{a}\_{ij}$

$\widetilde{a}\_{ij}=\displaystyle\frac{a\_{ij}-\mu\_j}{s\_j},i=1,2,\cdots,n;j=1,2,\cdots,m$

其中\[\mu\_{j}=\displaystyle\frac{1}{n}\sum\limits\_{i=1}^{n}a\_{ij},s\_j=\sqrt{\displaystyle\frac{1}{n-1}\sum\limits\_{i=1}^{n}\left(a\_{ij}-\mu\_j\right)^2},j=1,2,\cdots,m\]，即$\mu\_j$,$s\_j$为第$j$个指标的样本均值和样本标准差。对应的，称

$$\widetilde{x}\_j=\displaystyle\frac{x\_j-\mu\_j}{s\_j},j=1,2,\cdots,m$$

为标准化指标变量。

1. 计算相关系数矩阵$R$，相关系数矩阵$\it R=\left(r\_{ij})\_{m\times n}\right$，有

$$r\_{ij}=\displaystyle\frac{\sum\limits\_{k=1}^{n}\widetilde a\_{ki}\cdot\widetilde a\_{kj}}{n-1},I,j=1,2,\cdots,m$$

其中：$r\_{ii}=1$,$r\_{ij}=r\_{ji}$,$r\_{ij}$是第个$i$指标和第$j$个指标的相关系数

1. 计算特征值和特征向量。计算相关系数矩阵$R$的特征值$$\lambda\_1 \geqslant \lambda\_2\geqslant\cdots\geqslant\lambda\_m\geqslant 0$$j及对应的特征向量$u\_1,u\_2,\cdots,u\_m$,其中$u\_j=[u\_{1j},u\_{2j},\cdots,u\_{mj}]^T$，由特征向量组成$m$个新的指标变量:

$$\begin{cases}

y\_1=u\_{11}\widetilde x\_1+u\_{21}\widetilde x\_2+\cdots+u\_{m1}\widetilde x\_m\\

y\_2=u\_{12}\widetilde x\_1+u\_{22}\widetilde x\_2+\cdots+u\_{m2}\widetilde x\_m\\

\qquad\qquad\quad\quad\quad\vdots\\

y\_m=u\_{1m}\widetilde x\_1+u\_{2m}\widetilde x\_2+\cdots+u\_{mm}\widetilde x\_m

\end{cases}$$