

Lab8: 因子分析

1. 内容: 练习因子分析方法的使用
 2. 作业提交: 完成后面的作业, 现场演示给助教并解释结果.
-

1 因子分析的推断

R中可以用来进行因子分析的函数包括

- **factanal** 包含在R的基本安装里. 其使用最大似然方法来求解因子负荷阵. 因子的个数必须指定, 可以使用varimax和proximax 两种因子旋转方法. 输入数据可以为矩阵或者数据框.
- **paf** 包含在rela包里. 使用主轴方法(迭代主因子法)来求解因子负荷阵. 因子的个数通过特征根准则(eigencrit, 默认为1)来自动决定. 自动计算KMO和MSA, 也有因子旋转功能, 输入必须为矩阵.
- **fa** 包含在psych包里. 可以使用fm参数指定估计方法("minres"=最小残差; "ml"=最大似然;"pa"=主轴方法). 因子个数必须通过nfactors指定. 可以使用各种因子旋转方法(rotate=option ("none", "varimax", "quartimax", "bentlerT", "geominT", "oblimin", "simplimax", "bentlerQ", "geominQ", and "cluster"))
- **cortest.bartlett** 包含在psych包里, 用来进行Bartlett's 球性检验.
- **fa.parallel** 包含在psych包里, 使用平行因子分析法返回充分因子个数(Horn,1965). 平行分析可以通过以下几个步骤来完成:首先,生成一组随机数据的矩阵,这些矩阵和真实的数据矩阵要有相同的变量个数和样本个数. 然后,求出这组随机数据矩阵的特征值,并计算这些特征值的平均值. 最后,通过比较真实数据中特征值的碎石图和这组随机矩阵的平均特征值的曲线,我们可以找到两条特征值曲线的交点,根据交点的位置,我们就可以确定要抽取因子的绝对最大数目. 如果真实数据的特征值落在随机矩阵的平均特征值曲线之上,则保留这些因子;反之,则舍弃这些因子.《平行分析在探索性因素分析中的应用》(孔明, 2007)

```
library(foreign)
x <- read.delim("factorexdata05.txt")
# The data x consists of 538 cases with 102 variables.

Ps <- x[,4:43] # Extract variables p1-p40
Ps <- subset(Ps, complete.cases(Ps)) # Omit missings (511 cases remain)
res0<-factanal(Ps,factors=10,scores="reg",rotation="varimax")
```

```

library(rela)
res <- paf(as.matrix(Ps))
summary(res)
# Automatically calculate KMO with MSA, determine the number of factors,
# calculate chi-square of Bartlett's sphericity test, communalities and
# factor loadings. Communalities are 1 minus uniquenesses.
barplot(res$Eigenvalues[,1]) # First column of eigenvalues.
resv <- varimax(res$Factor.Loadings)
# Varimax rotation is possible later.
print(resv)
barplot(sort(colSums(loadings(resv)^2),decreasing=TRUE))
# screeplot using rotated SS loadings.

library(psych)
install.packages("GPArotation")

cortest.bartlett(Ps) # Bartlett's sphericity test.
res2 <- fa.parallel(Ps)
res3 <- fa(Ps, fm="minres", nfactors=8, rotate="varimax")
print(res3)

```

练习 1. 学习上述因子分析的包，使用其帮助文档的例子学习其用法。

上机作业:

练习 2. 三家房地产股票和两家白酒公司股票数据的每日收盘价，使用因子分析进行分析，能得出什么结论？

可以使用R的quantmod包下载股票交易数据, 代码示例如下

```

library(quantmod)
#保利地产#
setSymbolLookup(BLDC=list(name='600048.ss',src='yahoo'))
getSymbols("BLDC",from="2016-1-1", to="2016-11-16")
chartSeries(BLDC)

```

#绿地

```
setSymbolLookup(LD=list(name='600606.ss',src='yahoo'))
```

```
getSymbols("LD",from="2016-1-1", to="2016-11-16")
```

```
chartSeries(LD)
```

#宝安地产

```
setSymbolLookup(BADC=list(name='000040.sz',src='yahoo'))
```

```
getSymbols("BADC",from="2016-1-1", to="2016-11-16")
```

```
chartSeries(BADC)
```

#茅台"

```
setSymbolLookup(GZMT=list(name='600519.ss',src='yahoo'))
```

```
getSymbols("GZMT",from="2016-1-1", to="2016-11-16")
```

```
chartSeries(GZMT)
```

#五粮液#

```
setSymbolLookup(WLY=list(name='000858.sz',src='yahoo'))
```

```
getSymbols("WLY",from="2016-1-1", to="2016-11-16")
```

```
chartSeries(WLY)
```