因子分析

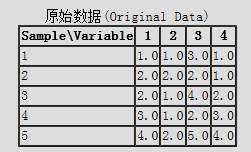
1. 资源

老师网上教案：<http://210.34.136.253:8488/MultiAnalysis/Chapter10.htm>

因子分析程序（第5个）：<http://210.34.136.253:8488/contribute.asp>

1. 因子分析步骤

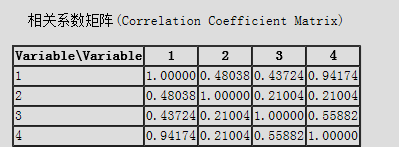
原始数据**【因子分析.xls】**：

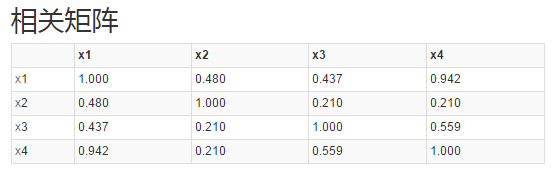


①计算相关矩阵R**（与主成分相同）**

【注】调pearson相关系数的方法，或者你的主成分中的相关系数的方法。

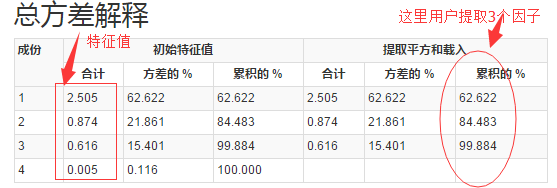
规格化后的数据矩阵X得到的相关系数矩阵为：



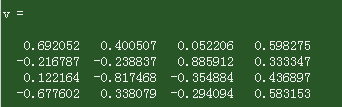


②计算该相关矩阵R的特征值和对应的特征向量**（与主成分相同）**

特征值为：



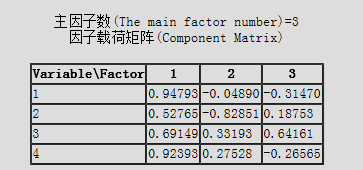
特征向量为：（我们把这张特征向量表也打印出来）



【注】与主成分的特征值和特征向量的计算方法是相同的。前台会让用户选择是用基于特征值还是用因子的固定数量来提取因子的个数。

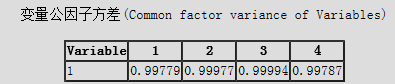
③根据提取出来的个数，得到因子的载荷（即成分）矩阵。**（与主成分相同）**

用对应的特征向量乘以根号特征值得到。





④得到公因子方差表**（与主成分相同）**





提取的公因子方差是上面得到的成分矩阵的每一行的元素的平方累加。

例如：0.947^2+(-0.049)^2+(-0.314)^2=0.9978……

⑤方差极大正交旋转矩阵**（拷贝学长代码）**

拷贝学长的**public** **void** Varimax\_orthogonal\_rotation()这个方法。

其中有5个变量名字说明：

aa[][]——是我们计算的出来的成分矩阵。你可以对应调整成你的变量名。

m——是成分矩阵的行数，也即变量的个数。你可以对应调整下你的变量。

p——是成分矩阵的列数，也即提取了多少个成分的数目。你可以对应调整下你的变量名。

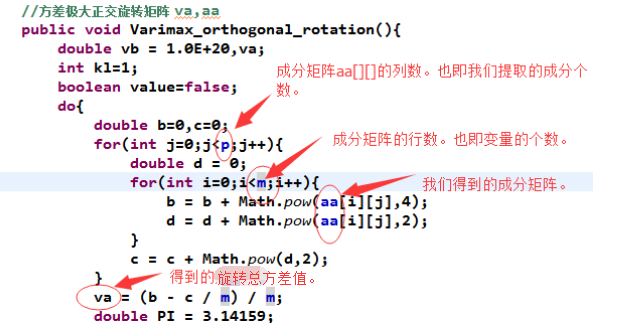
其他直接套用。

Parameters.Variance——方差旋转的精度值，我会让前台传这个值进来，你默认是0.001。

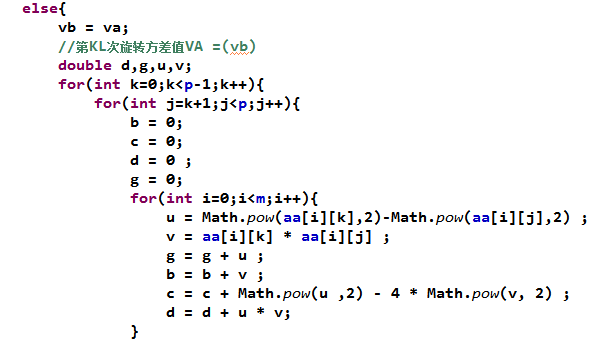
Va——旋转总方差值，我们想要得到最大的va，然后得到sin和cos的旋转角度值。

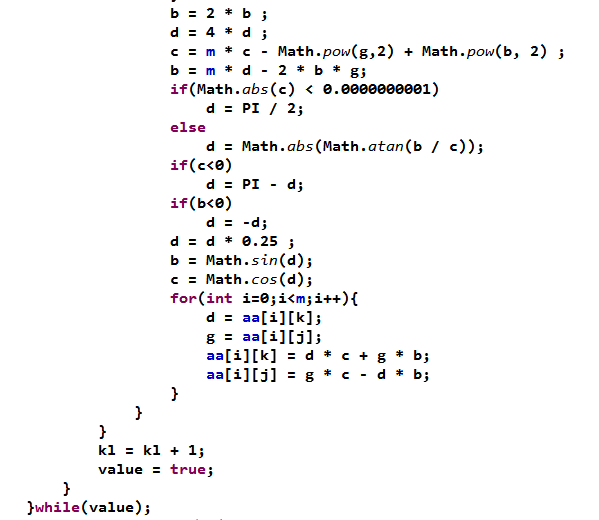
其他的照抄。

**经过对成分矩阵aa[][]的不断循环，最后得到最大方差旋转值，学长这里最后输出的极大方差旋转矩阵为aa[][]（是经过变换后的aa）。我们要的就是这个极大方差旋转矩阵aa[][]。**









⑥正交因子得分

1. 数据正规化处理得到数据矩阵X
2. 按照

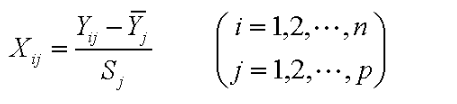
【注】其中X是第（1）步数据正规化后得到的数据矩阵。A是第⑤步得到的方差极大正交矩阵aa[][]，表示A的转置矩阵。R是第①步的相关系数矩阵，表示R的逆矩阵。

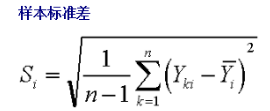
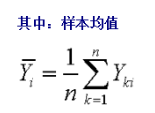
**F就是我们要得到的正交因子得分矩阵。**

------------------------------------------------------------------------------------------------------

（1）需要先做进行数据的正规化处理：

按照如下进行正规化处理：即原始数据矩阵元素减去每列对应的平均值，减去标准差。





【注】有可能S会等于零。你要做好当S等于0的时候，给S一个很小的值，如0.0000001之类的。老师的代码在正规化处理的时候没有把这个判断写进去，所以，运行对应分析.xls的时候会有问题，因为对应分析.xls里面有一列都是1，它的标准差为0。

【注】学长的代码中，标准差是除以n，我们按除以n-1计算。（这样我们得到的因子得分矩阵会和老师的程序运行的结果不一样，原因就在这里。）

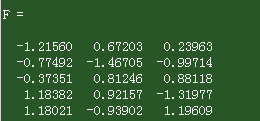
上述数据经过正规化后的结果是:（标准差除以n-1的）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 样本号\变量名 | X1 | X2 | X3 | X4 |
| 1 | -1.228 | -0.730 | -0.153 | -0.920 |
| 2 | -0.351 | 1.095 | -0.920 | -0.920 |
| 3 | -0.351 | -0.730 | 0.613 | -0.153 |
| 4 | 0.526 | -0.730 | -0.920 | 0.613 |
| 5 | 1.143 | 1.095 | 1.380 | 1.380 |

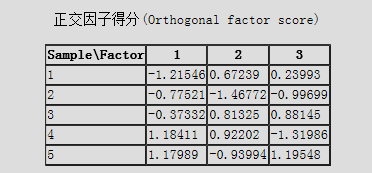
顺便把老师的代码中的正规化后的结果列出来，你用来测试（标准差除以n的）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 样本号\变量名 | X1 | X2 | X3 | X4 |
| 1 | -1.373 | -0.816 | -0.172 | -1.029 |
| 2 | -0.392 | 1.224 | -1.029 | -1.029 |
| 3 | -0.392 | -0.816 | 0.686 | -0.172 |
| 4 | 0.588 | -0.816 | -1.029 | 0.686 |
| 5 | 1.569 | 1.224 | 1.544 | 1.544 |

1. 按照公式，得到矩阵F为：



老师的结果：

（一致）