

**《判别分析》实验报告（ 3 ）**

院 系 数学与统计学院

专 业 应用统计学

姓名-学号 高芳-201813870059

吴佳宝-201813870051

吴琼-201813870050

指导教师 来 鹏

一家公司正试图对其销售员工的质量做评估，并且正寻找一种考察或一系列测试，以期可以揭示是否有创造良好销售额的潜能。该公司已挑选了50个销售人员的随机样本，还已对每一个人就3项表现作了评估：销售增长，销售利润和新客户销售额。这些测量度已被变为同一尺度，其中100表示“平均”业绩。50个人中的每一个接受4项测试，分别测量创造力、机械推理、抽象推理和数学能力。数据表中列出在p=7个变量上的n=50次观察值。请解决如下问题。

1、假设对标准化变量有正交因子模型，求m=2和m=3个公共因子的主成分解和极大似然解。

2、根据问题1的解，求m=2和m=3的旋转后因子载荷。比较这两组旋转载荷。解释m=2和m=3的因子解。

3、列出m=2和m=3的共性方差、特殊方差和AA'+\Fai的估计值。比较这些结果。此时，你更愿意选择m

等于什么值？为什么？

4、设随机选取一个新的销售人员，得到测试分数为：x'=(110,98,105,15,18,20,35)。利用回归方法，

计算该销售人员的因子得分。

备注：fac4=principal(X,2,rotate="varimax")

fac4$weights 可以输出因子得分的回归系数

（1）

代码：

a<-read.table(file.choose(),header=F)

b<-as.matrix(a)

cor(b) #计算样本数据的相关系数矩阵

factanal(b,factors=2,rotation="none") #极大似然法做因子分析

library(mvstats) #加载mvstats包

fac=factpc(b,2) #主成分法做因子分析

fac$loadings

factanal(b,factors=3,rotation="none")

fab=factpc(b,3)

fab$loadings

运行结果：

m=2时

极大似然法：

0.695 0.669

0.669 0.695

0.795 0.494

0.983 -0.167

0.655 0.312

0.250 0.569

0.558 0.812

主成分法：

0.9731 -0.107976

0.9429 0.028297

0.9448 0.008891

0.6603 0.645814

0.7833 0.284971

0.6488 -0.620657

0.9141 -0.193592

m=3时

极大似然法：

0.901 0.381

0.775 0.600

0.931 0.202

0.733 -0.118 0.666

0.689 0.225 0.169

0.757 -0.132 -0.636

0.762 0.608 -0.110

主成分法：

0.9731 -0.107976 -0.053267

0.9429 0.028297 -0.312228

0.9448 0.008891 0.144083

0.6603 0.645814 0.318588

0.7833 0.284971 0.004179

0.6488 -0.620657 0.426048

0.9141 -0.193592 -0.306273

（2）

代码：

fac1=factpc(b,2,rotation="varimax") #用主成分法采用方差最大化作因子正交旋转

fac1$loadings

fac2=factpc(b,3,rotation="varimax")

fac2$loadings

运行结果：

m=2时

0.78538 0.58455

0.67037 0.66363

0.68498 0.65072

0.04262 0.92265

0.37862 0.74256

0.89781 -0.01155

0.80065 0.48173

m=3时

0.7795 0.38719 0.45154

0.9082 0.35623 0.18882

0.6163 0.54771 0.48334

0.2128 0.95242 0.04694

0.5523 0.60712 0.14534

0.2866 0.06067 0.94969

0.9093 0.18073 0.32773

分析：

m=2时

因子1为：

y1=0.78538x1+0.67037x2+0.68498x3+0.04262x4+0.37862x5+0.89781x6+0.80065x7

第一个公共因子中平均业绩、抽象推理、数学能力有大的正负荷；创造力和机械推理有较小的正负荷。因此可将这一因子称为抽象因子思维因子

因子2为：

y2=0.58455x1+0.66363x2+0.65072x3+0.92265x4+0.74256x5-0.01155x6+0.48173x7

第二个公共因子上，平均业绩、创造力、机械推理有较大的正负荷，抽象推理有较小的负负荷，数学能力有较小的正负荷。因此可将这一因子称为创造力因子

m=3时

因子1为：

y1=0.7795x1+0.9082x2+0.6163x3+0.2128x4+0.5523x5+0.2866x6+0.9093x7

第一公共因子上，平均业绩，数学能力有较大的正负荷，可称为业绩因子

y2=0.38719x1+0.35623x2+0.54771x3+0.95242x4+0.60712x5+0.06067x6+0.18073x7

第二公共因子上，创造力，机械推理有较大的正负荷，可称为创造因子

y3=0.45154x1+0.18882x2+0.48334x3+0.04694x4+0.14534x5+0.94969x6+0.32773x7

第三公共因子上，抽象推理有较大的正负荷。可称为抽象推理能力。

（3）

代码：

p=cor(b)

d=fac1$common

e=1-d

A=fac1$loadings

g=A%\*%t(A)+diag(e)

p-g

d=fac2$common

e=1-d

A=fac2$loadings

g=A%\*%t(A)+diag(e)

p-g运行结果：

m=2时

> d共性方差

[1] 0.9585 0.8898 0.8926 0.8531 0.6948 0.8062 0.8731

> e特殊方差

[1] 0.04148 0.11019 0.10737 0.14691 0.30525 0.19380 0.12689

> g AA’+e的估计值

[,1] [,2] [,3] [,4] [,5] [,6] [,7]

[1,] 1.0000 0.9144 0.9183 0.5728 0.7314 0.6984 0.9104

[2,] 0.9144 1.0000 0.8910 0.6409 0.7466 0.5942 0.8564

[3,] 0.9183 0.8910 1.0000 0.6296 0.7425 0.6075 0.8619

[4,] 0.5728 0.6409 0.6296 1.0000 0.7013 0.0276 0.4786

[5,] 0.7314 0.7466 0.7425 0.7013 1.0000 0.3314 0.6609

[6,] 0.6984 0.5942 0.6075 0.0276 0.3314 1.0000 0.7133

[7,] 0.9104 0.8564 0.8619 0.4786 0.6609 0.7133 1.0000

V1 V2 V3 V4 V5 V6 V7

V1 0.0000000 0.0116521 -0.034345 -0.0007655 -0.0233511 -0.02397 0.016897

V2 0.0116521 0.0000000 -0.048506 -0.0993605 -0.0006954 -0.12881 0.087867

V3 -0.0343452 -0.0485065 0.000000 0.0707866 -0.1050755 0.03362 -0.009335

V4 -0.0007655 -0.0993605 0.070787 0.0000000 -0.1105217 0.11930 -0.065951

V5 -0.0233511 -0.0006954 -0.105076 -0.1105217 0.0000000 0.05460 -0.086307

V6 -0.0239651 -0.1288125 0.033624 0.1193030 0.0545977 0.00000 -0.146897

V7 0.0168969 0.0878673 -0.009335 -0.0659510 -0.0863067 -0.14690 0.000000

m=3时

> d

[1] 0.9614 0.9873 0.9134 0.9546 0.6948 0.9877 0.9669

> e

[1] 0.03864 0.01271 0.08661 0.04541 0.30523 0.01229 0.03309

> g

[,1] [,2] [,3] [,4] [,5] [,6] [,7]

[1,] 1.0000 0.9311 0.9107 0.5558 0.7312 0.6757 0.9267

[2,] 0.9311 1.0000 0.8460 0.5414 0.7453 0.4612 0.9521

[3,] 0.9107 0.8460 1.0000 0.6755 0.7431 0.6689 0.8178

[4,] 0.5558 0.5414 0.6755 1.0000 0.7026 0.1633 0.3810

[5,] 0.7312 0.7453 0.7431 0.7026 1.0000 0.3331 0.6596

[6,] 0.6757 0.4612 0.6689 0.1633 0.3331 1.0000 0.5828

[7,] 0.9267 0.9521 0.8178 0.3810 0.6596 0.5828 1.0000

V1 V2 V3 V4 V5 V6

V1 0.0000000 -0.0049794 -0.02667 0.0162049 -0.0231285 -0.001271

V2 -0.0049794 0.0000000 -0.00352 0.0001118 0.0006093 0.004212

V3 -0.0266703 -0.0035197 0.00000 0.0248834 -0.1056776 -0.027762

V4 0.0162049 0.0001118 0.02488 0.0000000 -0.1118530 -0.016431

V5 -0.0231285 0.0006093 -0.10568 -0.1118530 0.0000000 0.052817

V6 -0.0012706 0.0042116 -0.02776 -0.0164308 0.0528175 0.000000

V7 0.0005825 -0.0077598 0.03479 0.0316241 -0.0850269 -0.016410

V7

V1 0.0005825

V2 -0.0077598

V3 0.0347937

V4 0.0316241

V5 -0.0850269

V6 -0.0164103

V7 0.0000000

两者共同度接近，且累计方差贡献率都超过了85%，综上选择m=2

（4）

代码：

library(psych)

x=matrix(c(110,98,105,15,18,20,35),1,7)

fac4=principal(data,2,rotate="varimax")

sum(x\*(fac4$weights[,1]))

sum(x\*(fac4$weights[,2]))

运行结果：

> sum(x\*(fac4$weights[,1]))

[1] 62.57

> sum(x\*(fac4$weights[,2]))

[1] 39.96