**主成分分析**

目录

[一．与主成分分析有关的函数 1](#_Toc431228233)

[二．例题 2](#_Toc431228234)

[三．练习题 11](#_Toc431228235)

主成分分析是通过降维技术将多个变量化为少数几个主成分的一种统计分析方法，这些主成分能够反映原始变量的绝大部分信息，他们通常表示为原始变量的线性组合。

# 一．与主成分分析有关的函数

（1）princomp()函数——主成分分析函数

使用格式为princomp(x, cor = FALSE, scores = TRUE, covmat = NULL,...)

其中：

x为用于主成分分析的数据，以数值矩阵或者数据框的形式给出。

cor是逻辑变量，cor=T表示用样本的相关系数矩阵R作主成分分析，cor=F（缺省值）表示用样本的协方差矩阵作主成分分析。

covmat是协方差阵，如果数据不用x提供，可由协方差阵提供。

另，prcomp()函数的意义和使用方法与princomp()函数相同。

（2）summary()函数——提取主成分信息

使用格式为summary(object, loadings = FALSE, ...)

其中：

object是由princomp()得到的对象；

loadings是逻辑变量，loadings=T表示显示loadings的内容，loadings=F表示不显示loadings的内容。

（3）loadings()函数——显示主成分分析的载荷（主成分对应的各列，即正交矩阵Q）

使用格式为loadings(object)

其中：

object是由princomp()得到的对象。

（4）predict()函数——计算主成分得分（预测主成分值）

使用格式为predict(object, newdata, ...)

其中：

object是由princomp()得到的对象；

newdata是由预测值构成的数据框，当newdata缺省时，预测已有数据的主成分值。

（5）screeplot()函数——主成分碎石图

使用格式为screeplot(object, npcs = min(10, length(x$sdev)), type = c("barplot", "lines") ...)

其中：

object是由princomp()得到的对象；

npcs是画出的主成分的个数；

type是画出的碎石图的类型，"barplot"是直方图，"lines"是折线图。

（6）biplot()函数——信息重叠图

biplot()可画出数据关于主成分的散点图和原坐标在主成分下的方向。

使用格式为

biplot(object, choices = 1:2, scale = 1, pc.biplot = FALSE, ...)

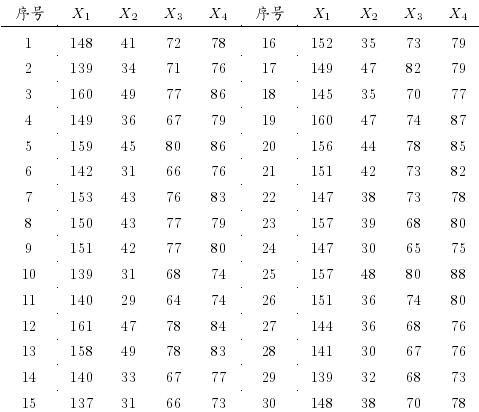
其中，

object是由princomp()得到的对象；

choices是选择的主成分，缺省为第一主成分和第二主成分。

# 二．例题

例1：在某中学随机抽取某年级30名学生，测量其身高（X1）、体重（X2）、胸围（X3）和坐高（X4），数据如下，试对这30名中学生身体四项指标数据做主成分分析。



#### 用数据框形式输入数据

>student<-data.frame(

X1=c(148, 139, 160, 149, 159, 142, 153, 150, 151, 139,

140, 161, 158, 140, 137, 152, 149, 145, 160, 156,

151, 147, 157, 147, 157, 151, 144, 141, 139, 148),

X2=c(41, 34, 49, 36, 45, 31, 43, 43, 42, 31,

29, 47, 49, 33, 31, 35, 47, 35, 47, 44,

42, 38, 39, 30, 48, 36, 36, 30, 32, 38),

X3=c(72, 71, 77, 67, 80, 66, 76, 77, 77, 68,

64, 78, 78, 67, 66, 73, 82, 70, 74, 78,

73, 73, 68, 65, 80, 74, 68, 67, 68, 70),

X4=c(78, 76, 86, 79, 86, 76, 83, 79, 80, 74,

74, 84, 83, 77, 73, 79, 79, 77, 87, 85,

82, 78, 80, 75, 88, 80, 76, 76, 73, 78)

)

> student.pr<-princomp(student, cor=TRUE) #### 作主成分分析，选择使用相关系数矩阵

> summary(student.pr, loadings=TRUE) ###显示结果，方差累积率以及载荷矩阵

Importance of components:

Comp.1 Comp.2

主成分标准差 Standard deviation 1.8817805 0.55980636

方差贡献率 Proportion of Variance 0.8852745 0.07834579

累积方差贡献率 Cumulative Proportion 0.8852745 0.96362029

Comp.3 Comp.4

0.28179594 0.25711844

0.01985224 0.01652747

0.98347253 1.00000000

Loadings:

Comp.1 Comp.2 Comp.3 Comp.4

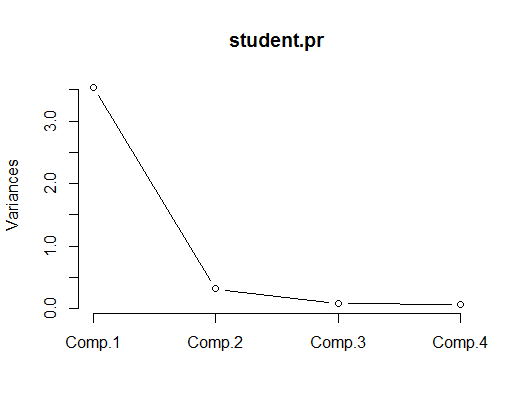
X1 -0.497 0.543 -0.450 0.506

X2 -0.515 -0.210 -0.462 -0.691

X3 -0.481 -0.725 0.175 0.461

X4 -0.507 0.368 0.744 -0.232

> screeplot(student.pr,type="lines") ###碎石图



由累积方差贡献率和碎石图可以看到，前两个主成分的累计方差贡献率达到96%，另外两个属成分可以舍弃，达到降维的目的。

并且可由载荷矩阵得到：





第一主成分对应的符号都相同，其值在0.5左右，它反映了中学生身材魁梧程度：身体高大的学生，他的4个部分的尺寸都比较大，因此，第一主成分的值就较小（因为系数均为负值）；而身材矮小的学生，他的4部分的尺寸都比较小，因此第一主成分的值较大。第二主成分是高度与围度的差，第2主成分值大的学生表明该学生“细高”，而第二主成分值越小的学生表明该学生“矮胖”，因此，称第二主成分为体型因子。

接下来看预测值：

> predict(student.pr) #### 作预测

Comp.1 Comp.2 Comp.3 Comp.4

[1,] 0.06990950 -0.23813701 -0.35509248 -0.266120139

[2,] 1.59526340 -0.71847399 0.32813232 -0.118056646

[3,] -2.84793151 0.38956679 -0.09731731 -0.279482487

[4,] 0.75996988 0.80604335 -0.04945722 -0.162949298

[5,] -2.73966777 0.01718087 0.36012615 0.358653044

[6,] 2.10583168 0.32284393 0.18600422 -0.036456084

[7,] -1.42105591 -0.06053165 0.21093321 -0.044223092

[8,] -0.82583977 -0.78102576 -0.27557798 0.057288572

[9,] -0.93464402 -0.58469242 -0.08814136 0.181037746

[10,] 2.36463820 -0.36532199 0.08840476 0.045520127

[11,] 2.83741916 0.34875841 0.03310423 -0.031146930

[12,] -2.60851224 0.21278728 -0.33398037 0.210157574

[13,] -2.44253342 -0.16769496 -0.46918095 -0.162987830

[14,] 1.86630669 0.05021384 0.37720280 -0.358821916

[15,] 2.81347421 -0.31790107 -0.03291329 -0.222035112

[16,] 0.06392983 0.20718448 0.04334340 0.703533624

[17,] -1.55561022 -1.70439674 -0.33126406 0.007551879

[18,] 1.07392251 -0.06763418 0.02283648 0.048606680

[19,] -2.52174212 0.97274301 0.12164633 -0.390667991

[20,] -2.14072377 0.02217881 0.37410972 0.129548960

[21,] -0.79624422 0.16307887 0.12781270 -0.294140762

[22,] 0.28708321 -0.35744666 -0.03962116 0.080991989

[23,] -0.25151075 1.25555188 -0.55617325 0.109068939

[24,] 2.05706032 0.78894494 -0.26552109 0.388088643

[25,] -3.08596855 -0.05775318 0.62110421 -0.218939612

[26,] -0.16367555 0.04317932 0.24481850 0.560248997

[27,] 1.37265053 0.02220972 -0.23378320 -0.257399715

[28,] 2.16097778 0.13733233 0.35589739 0.093123683

[29,] 2.40434827 -0.48613137 -0.16154441 -0.007914021

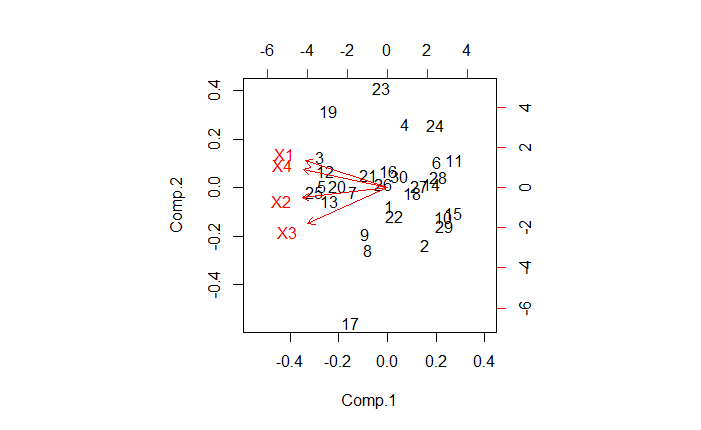
[30,] 0.50287468 0.14734317 -0.20590831 -0.122078819

从第一主成分来看，较小的几个值是25号、3号和5号，因此说明这几个学生身材魁梧；而11号、15号和29号的值较大，说明这几个学生身材瘦小。

从第二主成分来看，较大的几个值是23号、19号和4号，因此说明这几个学生属于“细高”型；而17号、8号和2号的值较小，说明这几个学生身材属于“矮胖”型。

做出第一主成分和第二主成分样本的散点图：

> biplot(student.pr)



从该图中可以容易地看出：哪些学生属于高大魁梧型，如25号，哪些学生属于身材瘦小型，如11号或者15号；哪些属于细高型，如23号，哪些属于矮胖型，如17号。哪些同学属于正常体型，如26号，等等。

> library(mvstats)

> princomp.rank(student.pr,m=2,plot=T)

计算综合得分：###使用mvstats函数包

Comp.1 Comp.2 PC rank

[1,] 0.06990950 -0.23813701 0.04486421 15

[2,] 1.59526340 -0.71847399 1.40714824 22

[3,] -2.84793151 0.38956679 -2.58471127 2

[4,] 0.75996988 0.80604335 0.76371582 19

[5,] -2.73966777 0.01718087 -2.51552608 3

[6,] 2.10583168 0.32284393 1.96086837 25

[7,] -1.42105591 -0.06053165 -1.31044040 9

[8,] -0.82583977 -0.78102576 -0.82219623 11

[9,] -0.93464402 -0.58469242 -0.90619170 10

[10,] 2.36463820 -0.36532199 2.14268263 27

[11,] 2.83741916 0.34875841 2.63508211 30

[12,] -2.60851224 0.21278728 -2.37913045 4

[13,] -2.44253342 -0.16769496 -2.25758089 5

[14,] 1.86630669 0.05021384 1.71865182 23

[15,] 2.81347421 -0.31790107 2.55888215 29

[16,] 0.06392983 0.20718448 0.07557694 16

[17,] -1.55561022 -1.70439674 -1.56770710 8

[18,] 1.07392251 -0.06763418 0.98110985 20

[19,] -2.52174212 0.97274301 -2.23762793 6

[20,] -2.14072377 0.02217881 -1.96487203 7

[21,] -0.79624422 0.16307887 -0.71824781 12

[22,] 0.28708321 -0.35744666 0.23468061 17

[23,] -0.25151075 1.25555188 -0.12898115 14

[24,] 2.05706032 0.78894494 1.95395798 24

[25,] -3.08596855 -0.05775318 -2.83976377 1

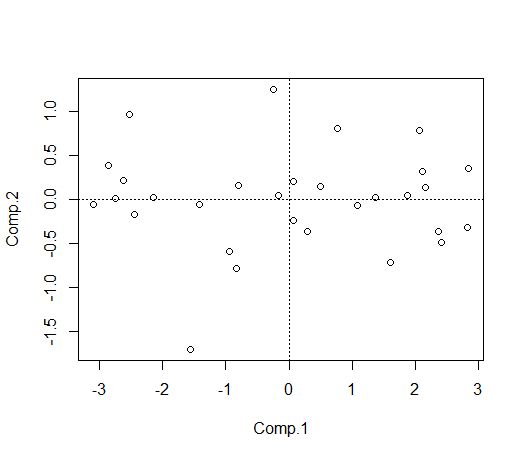
[26,] -0.16367555 0.04317932 -0.14685750 13

[27,] 1.37265053 0.02220972 1.26285484 21

[28,] 2.16097778 0.13733233 1.99644814 26

[29,] 2.40434827 -0.48613137 2.16934189 28

[30,] 0.50287468 0.14734317 0.47396869 18



例2：为了研究我国31个省、市、自治区2001年城镇居民生活消费的分布规律。用主成分分析法对我国31个省、市、自治区2001年城镇居民生活消费水平作分析评价，并根据因子得分和综合得分对各省、市、自治区的人均消费水平进行综合分析。

数据集：d7.2。

x1：人均食品支出 x2：人均衣着商品支出 x3：人均家庭设备用品及服务支出 x4：人均医疗保健支出 x5：人均交通和通信支出 x6：人均娱乐教育文化服务支出 x7：人均居住支出 x8：人均杂项商品和服务支出。

解：

###读取数据

> X=read.table("clipboard",header=T) ##数据集d7.2

进行主成分分析：

> PCA=princomp(X,cor=T)

#特征值开方（即主成分的方差开方——主成分的标准差）

> PCA

Call:

princomp(x = Z, cor = T)

Standard deviations: 主成分的方差开方

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6

2.2787134 1.1227556 0.8044059 0.6231343 0.4843913 0.3823558

Comp.7 Comp.8

0.2964918 0.2068370

8 variables and 31 observations.

###累积方差贡献率与载荷矩阵

> summary(PCA,loadings=T)

Importance of components:

Comp.1 Comp.2 Comp.3

主成分的方差开方Standard deviation 2.2787134 1.1227556 0.8044059

方差贡献率Proportion of Variance 0.6490668 0.1575725 0.0808836

累积方差贡献率Cumulative Proportion 0.6490668 0.8066394 0.8875230

Comp.4 Comp.5 Comp.6 Comp.7 Comp.8

0.62313432 0.48439131 0.38235576 0.29649185 0.206836981

0.04853705 0.02932937 0.01827449 0.01098843 0.005347692

0.93606002 0.96538939 0.98366388 0.99465231 1.000000000

Loadings:

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8

X1 -0.400 0.301 0.133 0.492 -0.215 0.604 -0.274

X2 -0.141 0.752 0.358 -0.488 -0.183 -0.103

X3 -0.363 -0.492 -0.492 0.321 0.526

X4 -0.342 0.262 -0.535 0.328 -0.521 -0.116 0.367

X5 -0.401 -0.135 0.377 -0.181 0.344 0.110 0.714

X6 -0.410 -0.211 0.286 -0.618 -0.463 0.329

X7 -0.288 -0.576 0.140 -0.427 -0.485 -0.222 -0.310

X8 -0.399 0.107 0.215 0.455 0.322 -0.521 -0.447

Loadings的内容是主成分对应于原始变量X1、X2、X3、X4、X5、X6、X7、X8的系数，即前面介绍的矩阵Q。因此可得到：（这里的\*指标准化后的结果）

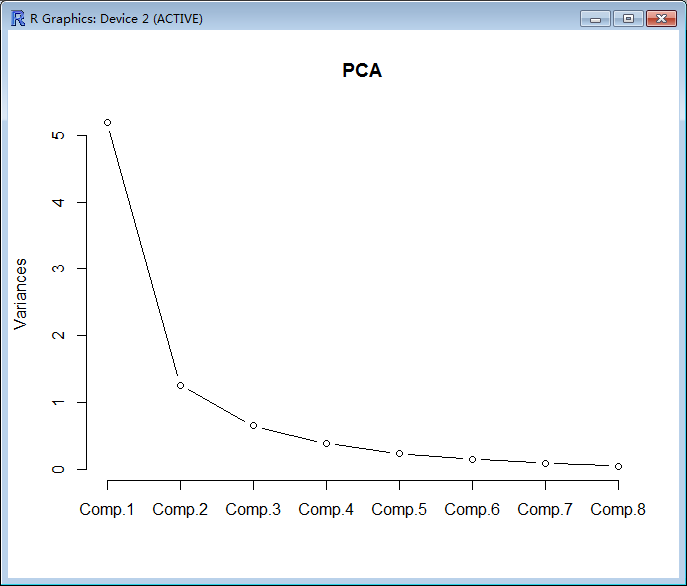


……..

我们从第一主成分对应系数的符号可以看出，x1到x8消费越高，Z1\*的值越小，Z1\*的绝对值越大。从第二主成分来看，正号大小多过负号大小，可认为x1到x8消费越高，Z2\*的值越大。

###碎石图

> screeplot(PCA,type="lines")



按照累计方差贡献率大于80%的原则，选定了两个主成分，其累计方差贡献率为80.7%，本例取m=2。从碎石图上也可以看出m取2比较合适。

计算主成分得分：

> predict(PCA)

或> PCA$scores

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8

北京 -6.0881641 2.09605700 -0.967784530 0.25776803 0.0005352398 0.37264097 -0.259313644 0.101898159

天津 -2.6531538 -0.89692239 -0.891557396 -1.07712010 0.0244270657 -0.27252772 0.270831386 -0.278031583

河北 1.1621365 0.30058573 -0.784504691 0.02193894 0.7896821227 -0.09836765 0.650193865 0.212926202

山西 1.6499715 0.43010054 -0.460958474 0.40646178 0.3690874371 0.07020928 -0.204588360 0.050284508

内蒙古 1.6314462 0.57608823 0.441231853 0.06715704 0.2598856038 -0.12973344 -0.510229607 0.174135101

辽宁 1.2429282 0.75205167 -0.051753055 0.33518942 0.4212833512 0.49894181 0.430659441 0.003292285

吉林 1.6459349 0.25354850 0.112123455 0.21903096 0.4579875644 0.32578858 0.139150622 -0.007332857

黑龙江 1.8162784 0.31636032 -0.254967650 0.25832368 0.5430370763 0.42992066 0.261995255 0.089928521

上海 -5.9388303 -0.16127086 0.413064801 1.23263843 -0.5779564955 0.73833944 0.239540230 0.019023710

江苏 -0.1682782 0.03012142 -0.233423513 -0.26671292 -0.8701181382 -0.65679059 0.021191974 0.137926258

浙江 -4.4178377 0.39587163 -0.969755399 -0.75820971 -0.0758300384 -0.56847898 0.052117591 -0.286139927

安徽 1.8800460 -0.38729538 0.304359563 -0.03960882 -0.8785949113 0.26598394 0.146905761 0.144447324

福建 -0.4665698 -0.90174490 0.729825867 -0.32744781 -0.3734575948 -0.26879668 0.652548503 -0.146259386

江西 2.5741394 -1.49544912 0.305732170 -0.13992335 -0.4700609392 0.05081704 -0.269775123 0.013447409

山东 0.1042408 1.12235349 -0.187443665 -0.81020523 -0.1991289510 0.07042515 -0.353062197 0.196000659

河南 1.8817327 -0.80438760 -0.167097210 -0.56844225 0.9028319618 0.13039124 0.125025151 -0.408239412

湖北 1.1608958 -0.21343763 0.346801704 -0.80620043 -0.0452773092 0.72568824 -0.295122773 -0.070358039

湖南 -0.4165050 -0.44372643 -0.001396215 -0.47340571 0.2432147756 0.40431003 -0.467658737 0.040373645

广东 -4.6096563 -3.09209873 1.517086976 0.33542504 1.0428558256 -0.41620859 -0.203544237 0.394551345

广西 0.2393494 -1.95810917 -0.100413776 -0.46735567 -0.1007335309 0.06618172 -0.285896780 -0.068351777

海南 1.7618874 -1.80161082 -0.120225434 1.35763576 -0.6997838002 -0.10853058 0.017022890 -0.187015549

重庆 -0.4425809 0.03293607 -0.152221620 -0.64286419 -0.4487145420 0.58529200 0.171194246 0.123427881

四川 0.5004198 -0.41276027 -0.203583270 -0.19934662 -0.3867866891 0.11601698 0.065305904 -0.156010593

贵州 1.9329234 0.06752372 -0.003314839 0.10134689 -0.5829378816 -0.21588180 0.190903255 0.396356661

云南 0.1084411 0.11915042 0.467213205 0.85613520 0.4402763200 -0.02640783 0.005888688 -0.326523841

西藏 -0.2021215 2.59081438 3.214746408 -0.43987122 -0.1134174018 -0.36790757 0.166188437 -0.133106158

陕西 0.7689799 -0.20239904 -1.151012313 0.01361699 -0.0680352499 -0.37792016 -0.111388971 0.102067293

甘肃 1.2875337 0.80112651 -0.525399252 0.57331632 -0.0233449641 -0.08444037 -0.091012163 -0.034341153

青海 0.6706571 0.97432899 -0.168785155 1.32837316 -0.0563299105 -0.45636528 -0.471031209 -0.401834404

宁夏 0.7537904 1.01381166 -0.718171966 0.34970205 0.3157298525 -0.70287387 0.172563888 0.326958031

新疆 0.6299651 0.89838206 0.261583420 -0.69734565 0.1596741511 -0.09971596 -0.256603285 -0.023500314

Comp.1 Comp.2 PC rank

> library(mvstats)

> princomp.rank(PCA,m=2,plot=T) #对主成分结果PCA做排名，m=2指选择2个主成分，plot=T指做出第一主成分为横轴，第二主成分为纵轴的散点图，结果中的PC为综合得分

北京 -6.0881641 2.09605700 -4.4894219 2

天津 -2.6531538 -0.89692239 -2.3100837 5

河北 1.1621365 0.30058573 0.9938373 20

山西 1.6499715 0.43010054 1.4116764 26

内蒙古 1.6314462 0.57608823 1.4252879 27

辽宁 1.2429282 0.75205167 1.1470382 22

吉林 1.6459349 0.25354850 1.3739399 25

黑龙江 1.8162784 0.31636032 1.5232777 29

上海 -5.9388303 -0.16127086 -4.8102161 1

江苏 -0.1682782 0.03012142 -0.1295219 10

浙江 -4.4178377 0.39587163 -3.4775063 4

安徽 1.8800460 -0.38729538 1.4371334 28

福建 -0.4665698 -0.90174490 -0.5515789 6

江西 2.5741394 -1.49544912 1.7791678 31

山东 0.1042408 1.12235349 0.3031235 12

河南 1.8817327 -0.80438760 1.3570140 24

湖北 1.1608958 -0.21343763 0.8924274 19

湖南 -0.4165050 -0.44372643 -0.4218225 7

广东 -4.6096563 -3.09209873 -4.3132099 3

广西 0.2393494 -1.95810917 -0.1899120 9

海南 1.7618874 -1.80161082 1.0657778 21

重庆 -0.4425809 0.03293607 -0.3496913 8

四川 0.5004198 -0.41276027 0.3220351 13

贵州 1.9329234 0.06752372 1.5685279 30

云南 0.1084411 0.11915042 0.1105331 11

西藏 -0.2021215 2.59081438 0.3434630 14

陕西 0.7689799 -0.20239904 0.5792264 15

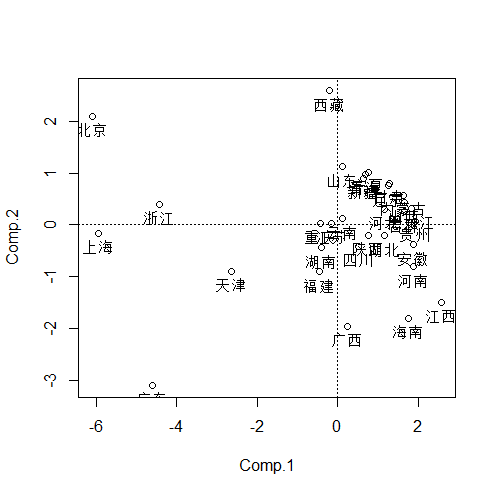
甘肃 1.2875337 0.80112651 1.1925167 23

青海 0.6706571 0.97432899 0.7299777 17

宁夏 0.7537904 1.01381166 0.8045841 18

新疆 0.6299651 0.89838206 0.6823989 16

以第一主成分为横轴，第二主成分为纵轴，绘制各省、市、自治区的成分图。



在日常必需消费主成分C1上得分最高的前五个省、市、自治区依次是上海、北京、广东、浙江和天津，且上海、北京和广东绝对值明显高于其他省、市、自治区。这就是说，以食品和交通通信等为主的日常必需消费而言，北京、广东和上海的消费水平远远高于其他省、市、自治区；而江西和贵州在这方面的消费相对较低。西藏、北京和山东在主成分C2上的得分较高，可见这些地区人们用于衣着和住房方面的消费支出不小，西藏排到全国最前，主要是从人均来说，西藏在这方面占有优势。对衣着因子而言，西藏、北京的得分最高，得分较低的是广东、广西和海南。这说明衣着因子受气候影响最大，北部、西北部省、市、自治区的人们为了御寒，因此在这方面的支出较多。其次影响衣着因子的就是各地人们的衣着习惯了，例如天津和广东，他们的经济都比较发达。但排名却较后，根据资料可知，天津虽和北京一样同为直辖市，且与北京相邻，但由于衣着习惯的原因，北京人是非常注重衣着的，而天津人就没有北京人那么注重着装，因而他的衣着因子得分较低。同样的道理，同为经济发达地区的广东和上海相比，上海人的穿着就比广东人要讲究的多，广东人平时的穿着很随意，因而该省人们用于衣着方面的人均消费支出相对较少也就不足为奇了。就综合得分来看，上海、北京、广东、浙江、天津这五个省、市的得分最高，江西、贵州、黑龙江得分位于全国之末，故可知上海、北京、广东、浙江、天津这五个省、市的综合人均消费水平居于全国水平前列，江西、贵州、黑龙江省的综合人均消费水平居于全国水平之末。

做信息重叠图：

>biplot(PCA)



图中红色箭头线和旁边标注的原始变量，表示原始变量在以第一主成分和第二主成分为横纵坐标的坐标系中的方向。

# 三．练习题

用主成分分析法对我国31个省、市、自治区2006年城镇居民生活消费水平作分析评价.

数据集：e8.3

> X=read.table("clipboard",header=T)

> cor(X)

X1 X2 X3 X4 X5 X6 X7 X8

X1 1.0000000 0.6423319 0.8842360 0.9281409 0.8755226 0.8364431 0.7513519 0.9296232

X2 0.6423319 1.0000000 0.8055496 0.7920670 0.8629616 0.8252951 0.8825457 0.6857909

X3 0.8842360 0.8055496 1.0000000 0.9339795 0.8875355 0.8331069 0.8308087 0.8359878

X4 0.9281409 0.7920670 0.9339795 1.0000000 0.8873376 0.8941652 0.8404040 0.8651084

X5 0.8755226 0.8629616 0.8875355 0.8873376 1.0000000 0.9311442 0.9125810 0.8913567

X6 0.8364431 0.8252951 0.8331069 0.8941652 0.9311442 1.0000000 0.9197204 0.8529065

X7 0.7513519 0.8825457 0.8308087 0.8404040 0.9125810 0.9197204 1.0000000 0.7714116

X8 0.9296232 0.6857909 0.8359878 0.8651084 0.8913567 0.8529065 0.7714116 1.0000000

> PCA=princomp(X,cor=T)

> PCA

Call:

princomp(x = Z, cor = T)

Standard deviations:

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5

2.6371051 0.7171103 0.4601558 0.3562669 0.2822771

Comp.6 Comp.7 Comp.8

0.2286716 0.2120411 0.1258082

8 variables and 31 observations.

> PCA$loadings

Loadings:

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8

X1 -0.349 -0.514 0.190 0.489 0.577

X2 -0.330 0.607 0.250 0.457 0.388 0.150 0.273

X3 -0.357 0.628 -0.426 -0.527

X4 -0.363 -0.164 0.348 -0.375 0.447 -0.171 0.214 -0.552

X5 -0.369 -0.217 0.264 -0.238 0.683 0.132 -0.449

X6 -0.361 0.120 -0.428 -0.429 0.348 0.201 -0.518 0.241

X7 -0.351 0.393 -0.238 -0.339 -0.538 -0.421 0.290

X8 -0.348 -0.396 -0.366 0.527 -0.496 -0.212 -0.136

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7

SS loadings 1.000 1.000 1.000 1.000 1.000 1.000 1.000

Proportion Var 0.125 0.125 0.125 0.125 0.125 0.125 0.125 Cumulative Var 0.125 0.250 0.375 0.500 0.625 0.750 0.875

Comp.8

1.000

0.125

1.000

> summary(PCA)

Importance of components:

Comp.1 Comp.2 Comp.3

Standard deviation 2.6371051 0.71711027 0.46015579

Proportion of Variance 0.8692904 0.06428089 0.02646792

Cumulative Proportion 0.8692904 0.93357129 0.96003921

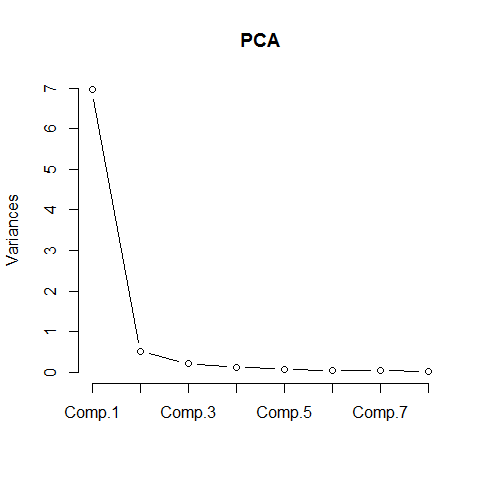
Comp.4 Comp.5 Comp.6 Comp.7 Comp.8

0.35626689 0.282277145 0.228671557 0.212041117 0.125808190

0.01586576 0.009960048 0.006536335 0.005620179 0.001978463

0.97590497 0.985865023 0.992401358 0.998021537 1.000000000

> screeplot(PCA,type="lines")



Comp.1 Comp.2 Comp.3 Comp.4

> PCA$scores

北京 -6.0830445 -1.82216912 -0.67496049 -0.328269412

天津 -0.7626158 -0.92180915 0.35896706 0.409760359

河北 0.8400787 -0.33623204 0.37213733 -0.013039644

山西 1.0856463 -0.81756405 -0.08678151 0.185225157

内蒙古 0.1849889 -0.53112701 -0.73080992 -0.009811459

辽宁 -0.5057344 -0.60482393 -0.30858778 0.458609758

吉林 0.1306953 -0.36629827 -0.74417992 0.210809416

黑龙江 0.5096477 -0.78752713 0.08342849 0.128919613

上海 -9.5090573 1.03381724 0.89920329 -0.393628226

江苏 -2.1631855 0.55165429 -0.41171564 0.113426253

浙江 -5.8711926 -0.21737131 0.23102153 -0.011840045

安徽 1.0840292 -0.01326624 -0.04118083 -0.285760525

福建 -1.0972392 0.75807196 -0.06546941 0.790390153

江西 0.8804707 0.40876564 -0.18364049 -0.140966911

山东 -0.4561218 -0.23968219 -0.02431928 -0.183497237

河南 1.3558516 -0.15001878 0.37196076 0.148479842

湖北 0.3934507 0.60394573 -0.33565232 0.173585701

湖南 0.5471425 1.29770184 -0.60193711 -0.694499999

广东 -1.3992189 1.62023655 -0.42002219 0.891058183

广西 1.5246952 0.64302764 0.25418969 -0.358525927

海南 1.5925446 0.95340974 -0.57446275 0.076448865

重庆 1.7371109 0.13914743 0.11131299 -0.536151254

四川 1.3877436 0.19373085 0.17383886 -0.265062934

贵州 2.7760936 0.19569035 0.19756046 -0.130913730

云南 1.8910326 0.32904625 0.36089886 -0.340734184

西藏 2.0287134 0.05756430 1.36391722 0.312841834

陕西 1.3235337 -0.38818803 -0.27085397 -0.303128315

甘肃 2.2498297 -0.07082232 -0.11228359 -0.430423407

青海 1.4346852 -0.47993606 0.30189596 0.198632809

宁夏 1.2060704 -0.19034498 0.18311126 0.203377371

新疆 1.6833556 -0.84862921 0.32341344 0.124687894

Comp.5 Comp.6 Comp.7 Comp.8

北京 0.23691596 -0.059592362 -0.3938359466 -0.061749757

天津 -0.42884223 0.493225433 -0.0765644536 0.007838461

河北 -0.09768740 0.208635580 0.3392038083 -0.237745355

山西 0.73258577 0.049462853 0.2334382877 0.181854324

内蒙古 -0.02130721 0.287731227 0.0755949612 -0.029951276

辽宁 -0.11391883 -0.286336838 0.0839962964 0.137743399

吉林 -0.02359604 -0.416953791 0.0007592711 -0.013975838

黑龙江 -0.54796492 -0.239426850 0.3187143250 0.155325341

上海 0.07252429 -0.293215622 0.0947034021 -0.080717324

江苏 0.41687955 0.342306984 0.1679370468 0.022717644

浙江 -0.31045585 0.205048926 0.1440747076 0.233778877

安徽 0.13654426 0.055599410 0.0916624410 -0.065641386

福建 0.30896988 0.300301988 -0.1439182557 -0.104806858

江西 0.06145558 0.152357766 -0.1107619180 0.238352430

山东 0.16065956 0.177013462 0.2269779362 -0.171988981

河南 0.03971211 -0.062280131 0.1760012791 -0.136377119

湖北 0.27191788 -0.397582630 -0.0174681844 -0.045322417

湖南 -0.32561402 -0.112453079 0.0717250230 0.001918867

广东 -0.36450218 0.034545932 -0.1660818707 0.017332117

广西 -0.20535838 0.349041437 -0.0624260866 -0.084910398

海南 0.13071614 -0.118034569 0.0455160512 0.009240823

重庆 0.12025473 0.041549352 -0.5291999485 0.021982750

四川 0.06821391 0.006600059 -0.4187764328 0.134094675

贵州 0.09855003 0.025062995 0.1121374130 0.041390876

云南 -0.29064562 -0.022591718 0.0885927535 0.202148173

西藏 0.53051158 -0.212070735 -0.0657547623 0.091367377

陕西 -0.02555145 -0.163783410 0.2312324870 -0.102275367

甘肃 0.07997399 0.145935233 0.0752852901 -0.024872930

青海 -0.30170869 -0.046989570 -0.4096867530 -0.189702065

宁夏 -0.19378509 -0.346671788 -0.0456423802 -0.204119674

新疆 -0.21544732 -0.096435546 -0.1374357879 0.057070612