

案例 2 四川省 2007 年各地市经济

. 根据四川省 2007 年各地市经济指标数据，见表求解：

(1) 利用主成分分析对 17 个地市的经济发展进行分析，给出排名；

(2) 此时能否只用第一主成分进行排名？为什么？

地区	工业总产值	资产合计	工业增加值	实收资本	长期负债	业务收入	业务成本	利润
成都	491.70	380.31	158.39	121.54	22.74	439.65	344.44	17.43
绵阳	21.12	30.55	6.40	12.40	3.31	21.17	17.71	2.03
德阳	1.71	2.35	0.57	0.68	0.13	1.48	1.36	-0.03
广元	9.83	9.05	3.13	3.43	0.64	8.76	7.81	0.54
乐山	64.06	77.86	20.63	30.37	5.96	63.57	52.15	4.71
自贡	30.38	46.90	9.19	9.83	17.87	28.24	21.90	3.80
宜宾	31.20	70.07	8.93	18.88	33.05	31.17	26.50	2.84
雅安	79.18	62.09	20.78	24.47	3.51	71.29	59.07	6.78
甘孜	47.81	40.14	17.50	9.52	4.14	45.70	34.73	4.47
阿坝	104.69	78.95	29.61	25.96	5.39	98.08	84.81	3.81
南充	21.07	17.83	6.21	6.22	1.90	20.24	16.46	1.09
广安	214.19	146.78	65.16	41.62	4.39	194.98	171.98	11.05
内江	31.16	27.56	8.80	9.44	1.47	28.83	25.22	1.05
宜宾	12.79	14.16	3.66	4.07	1.57	11.95	10.24	0.73
大足	6.45	5.37	2.39	2.20	0.40	5.97	4.79	0.52
邛崃	39.43	44.60	15.17	15.72	3.27	36.03	27.87	3.48

遂宁	5.02	3.62	1.63	1.42	0.53	4.45	4.04	0.02
----	------	------	------	------	------	------	------	------

解答：

(1)

```
>> clear
>> A=[491.70, 380.31, 158.39, 121.54, 22.74, 439.65, 344.44, 17.43;
21.12, 30.55, 6.40, 12.40, 3.31, 21.17, 17.71, 2.03;
1.71, 2.35, 0.57, 0.68, 0.13, 1.48, 1.36, -0.03;
9.83, 9.05, 3.13, 3.43, 0.64, 8.76, 7.81, 0.54;
64.06, 77.86, 20.63, 30.37, 5.96, 63.57, 52.15, 4.71;
30.38, 46.90, 9.19, 9.83, 17.87, 28.24, 21.90, 3.80;
31.20, 70.07, 8.93, 18.88, 33.05, 31.17, 26.50, 2.84;
79.18, 62.09, 20.78, 24.47, 3.51, 71.29, 59.07, 6.78;
47.81, 40.14, 17.50, 9.52, 4.14, 45.70, 34.73, 4.47;
104.69, 78.95, 29.61, 25.96, 5.39, 98.08, 84.81, 3.81;
21.07, 17.83, 6.21, 6.22, 1.90, 20.24, 16.46, 1.09;
214.19, 146.78, 65.16, 41.62, 4.39, 194.98, 171.98, 11.05;
31.16, 27.56, 8.80, 9.44, 1.47, 28.83, 25.22, 1.05;
12.76, 14.16, 3.66, 4.07, 1.57, 11.95, 10.24, 0.73;
6.45, 5.37, 2.39, 2.20, 0.40, 5.97, 4.79, 0.52;
39.43, 44.60, 15.17, 15.72, 3.27, 36.03, 27.87, 3.48;
5.02, 3.62, 1.63, 1.42, 0.53, 4.45, 4.04, 0.02];
```

得到的相关系数矩阵为：

```
>> R=corrcoef(A)
R =

    1.0000    0.9877    0.9988    0.9820    0.4281    0.9999
    0.9980    0.9510
    0.9877    1.0000    0.9884    0.9947    0.5438    0.9885
    0.9835    0.9485
    0.9988    0.9884    1.0000    0.9824    0.4294    0.9984
    0.9948    0.9462
    0.9820    0.9947    0.9824    1.0000    0.5051    0.9829
    0.9763    0.9391
    0.4281    0.5438    0.4294    0.5051    1.0000    0.4311
    0.4204    0.4557
    0.9999    0.9885    0.9984    0.9829    0.4311    1.0000
    0.9986    0.9530
    0.9980    0.9835    0.9948    0.9763    0.4204    0.9986
    1.0000    0.9569
```

0.9510	0.9485	0.9462	0.9391	0.4557	0.9530
0.9569	1.0000				

计算特征值与特征向量：

```
>> [v,d]=eig(R)
```

v =

-0.3723	0.1179	0.1411	-0.2543	-0.0459	0.5917
-0.5641	0.3041				
-0.3741	-0.0343	0.1606	0.2247	-0.1514	-0.6284
-0.1535	0.5841				
-0.3719	0.1152	0.1957	-0.1954	-0.6909	-0.1351
0.0383	-0.5244				
-0.3713	0.0096	0.2368	0.7875	0.2168	0.2385
0.0303	-0.2845				
-0.1949	-0.9689	-0.0004	-0.1242	0.0119	0.0628
0.0151	-0.0593				
-0.3725	0.1143	0.1222	-0.2302	0.0924	0.2259
0.7946	0.2988				
-0.3716	0.1272	0.0353	-0.3800	0.6591	-0.3521
-0.1557	-0.3428				
-0.3613	0.0596	-0.9185	0.1165	-0.0872	0.0302
0.0022	-0.0096				

d =

7.1135	0	0	0	0	0
0	0				
	0	0.7770	0	0	0
0	0				
	0	0	0.0810	0	0
0	0				
	0	0	0	0.0237	0
0	0				
	0	0	0	0	0.0041
0	0				
	0	0	0	0	0
0	0				
	0	0	0	0	0
0.0000	0				
	0	0	0	0	0
0	0.0001				

各主成分贡献率：

```
>> w=sum(d)/sum(sum(d))
```

w =

```
    0.8892    0.0971    0.0101    0.0030    0.0005    0.0001
0.0000    0.0000
```

计算各个主成分得分：

```
>> F=[A-ones(17,1)*mean(A)]*v(:,8)
```

F =

```
224.3503
-24.0409
-40.0941
-35.9075
  4.7573
-12.6102
 -2.8573
  1.8038
-13.9012
 13.4541
-29.3847
 62.3383
-23.3175
-32.4285
-38.1309
-14.8637
-39.1675
```

```
>> [F1,I1]=sort(F,'descend')
```

F1 按从大到小的顺序给个主成分得分排名：

F1 =

```
224.3503
 62.3383
 13.4541
  4.7573
  1.8038
 -2.8573
-12.6102
-13.9012
-14.8637
```

-23.3175
-24.0409
-29.3847
-32.4285
-35.9075
-38.1309
-39.1675
-40.0941

I1 给出各个名次的序号：

I1 =

1
12
10
5
8
7
6
9
16
13
2
11
14
4
15
17
3

>> [F2, I2]=sort(I1)

F2 =

1
2
3
4
5
6
7
8
9

10
11
12
13
14
15
16
17

I2 给出个城市排名，即所求排名：

I2 =

1
11
17
14
4
7
6
5
8
3
12
2
10
13
15
9
16

(2) 由于第一主成分的贡献率大于 80%，其他各成分贡献率都太小，所以只能用第一主成分进行排名。