第八章 因子分析

一、盐泉水化学分析资料的因子分析

1、盐泉水化学特征系数的数据

| **Obs** | **x1** | **x2** | **x3** | **x4** | **x5** | **x6** | **x7** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **1** | 11.835 | 0.480 | 14.360 | 25.210 | 25.21 | 0.810 | 0.98 |
| **2** | 45.596 | 0.526 | 13.850 | 20.040 | 26.01 | 0.910 | 0.96 |
| **3** | 3.525 | 0.086 | 24.400 | 49.300 | 11.30 | 6.820 | 0.85 |
| **4** | 3.681 | 0.370 | 13.570 | 25.120 | 26.00 | 0.820 | 1.01 |
| **5** | 42.287 | 0.386 | 14.500 | 25.900 | 23.32 | 2.180 | 0.93 |
| **6** | 17.956 | 1.280 | 9.750 | 17.050 | 37.20 | 0.464 | 0.98 |
| **7** | 7.370 | 0.506 | 13.600 | 34.280 | 10.69 | 8.800 | 0.56 |
| **8** | 4.223 | 0.340 | 3.800 | 7.100 | 88.20 | 1.110 | 0.97 |
| **9** | 6.442 | 0.190 | 4.700 | 9.100 | 73.20 | 0.740 | 1.03 |
| **10** | 16.234 | 0.390 | 3.100 | 5.400 | 121.50 | 0.420 | 1.00 |
| **11** | 10.585 | 0.420 | 2.400 | 4.700 | 135.60 | 0.870 | 0.98 |
| **12** | 23.535 | 0.230 | 2.600 | 4.600 | 151.80 | 0.310 | 1.02 |
| **13** | 5.398 | 0.120 | 2.800 | 6.200 | 111.20 | 1.140 | 1.07 |
| **14** | 283.149 | 0.148 | 1.763 | 2.986 | 215.86 | 0.140 | 0.98 |
| **15** | 316.604 | 0.317 | 1.453 | 2.432 | 263.41 | 0.249 | 0.98 |
| **16** | 307.310 | 0.173 | 1.627 | 2.729 | 235.70 | 0.214 | 0.99 |
| **17** | 322.515 | 0.312 | 1.382 | 2.320 | 282.21 | 0.024 | 0.99 |
| **18** | 254.580 | 0.297 | 0.899 | 1.476 | 410.30 | 0.239 | 0.93 |
| **19** | 304.920 | 0.283 | 0.789 | 1.357 | 438.36 | 0.193 | 1.01 |
| **20** | 202.446 | 0.042 | 0.741 | 1.266 | 309.77 | 0.290 | 0.99 |

SAS程序

data chapter8;

input x1 x2 x3 x4 x5 x6 x7;

cards;

11.835 0.480 14.360 25.210 25.21 0.810 0.98

45.596 0.526 13.850 20.040 26.01 0.910 0.96

3.525 0.086 24.400 49.300 11.30 6.820 0.85

3.681 0.370 13.570 25.120 26.00 0.820 1.01

42.287 0.386 14.500 25.900 23.32 2.180 0.93

17.956 1.280 9.750 17.050 37.20 0.464 0.98

7.370 0.506 13.600 34.280 10.69 8.800 0.56

4.223 0.340 3.800 7.100 88.20 1.110 0.97

6.442 0.190 4.700 9.100 73.20 0.740 1.03

16.234 0.390 3.100 5.400 121.50 0.420 1.00

10.585 0.420 2.400 4.700 135.60 0.870 0.98

23.535 0.230 2.600 4.600 151.80 0.310 1.02

5.398 0.120 2.800 6.200 111.20 1.140 1.07

283.149 0.148 1.763 2.986 215.86 0.140 0.98

316.604 0.317 1.453 2.432 263.41 0.249 0.98

307.310 0.173 1.627 2.729 235.70 0.214 0.99

322.515 0.312 1.382 2.320 282.21 0.024 0.99

254.580 0.297 0.899 1.476 410.30 0.239 0.93

304.92 0.283 0.789 1.357 438.36 0.193 1.01

202.446 0.042 0.741 1.266 309.77 0.290 0.99

;

proc print;

run;

proc factor rotate=varimax;

run;

3、结果分析

由于没有确定因子分析方法，所以系统默认是主成分分析法，可选用的有主成分分析法，最大似然分析法，主因子分析法；

Rotate是选择因子旋转的方法，quartimax是四次方最大法，应为选用最大方差旋转法时显示Conditional variance of x1 is zero.

| **相关矩阵的特征值: 总计 = 7 平均值 = 1** | | | | |
| --- | --- | --- | --- | --- |
|  | **特征值** | **差分** | **比例** | **累积** |
| **1** | 4.18700647 | 2.81656907 | 0.5981 | 0.5981 |
| **2** | 1.37043740 | 0.52791091 | 0.1958 | 0.7939 |
| **3** | 0.84252650 | 0.40020840 | 0.1204 | 0.9143 |
| **4** | 0.44231809 | 0.32739248 | 0.0632 | 0.9775 |
| **5** | 0.11492561 | 0.07517167 | 0.0164 | 0.9939 |
| **6** | 0.03975394 | 0.03672195 | 0.0057 | 0.9996 |
| **7** | 0.00303199 |  | 0.0004 | 1.0000 |

由上表可以看出，当选择三个因子时，因子累积达到0.7939，快超过了0.80。

即选用两个因子即可。

| **因子模式** | | | | |
| --- | --- | --- | --- | --- |
|  | **Factor1** | | **Factor2** | |
| **x1** | -0.72157 | | 0.50038 | |
| **x2** | 0.31612 | | -0.54053 | |
| **x3** | 0.91449 | | -0.03224 | |
| **x4** | 0.94956 | | 0.09351 | |
| **x5** | -0.84158 | | 0.42556 | |
| **x6** | 0.82479 | | 0.52113 | |
| **x7** | -0.66327 | | -0.60450 | |
|  |  | |  | |
| **每个因子已解释方差** | | | |
| **Factor1** | | **Factor2** | |
| 4.1870065 | | 1.3704374 | |

这个表给出了因子符合矩阵

| **最终的公因子方差估计: 总计 = 5.557444** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **x1** | **x2** | **x3** | **x4** | **x5** | **x6** | **x7** |
| 0.77104313 | 0.39210572 | 0.83732584 | 0.91041224 | 0.88935728 | 0.95185369 | 0.80534597 |

方差最大的正交旋转

|  | **1** | **2** |
| --- | --- | --- |
| **1** | 0.76862 | -0.63971 |
| **2** | 0.63971 | 0.76862 |

| **旋转因子模式** | | |
| --- | --- | --- |
|  | **Factor1** | **Factor2** |
| **x1** | -0.23451 | 0.84620 |
| **x2** | -0.10281 | -0.61769 |
| **x3** | 0.68226 | -0.60979 |
| **x4** | 0.78967 | -0.53557 |
| **x5** | -0.37461 | 0.86546 |
| **x6** | 0.96732 | -0.12708 |
| **x7** | -0.89650 | -0.04033 |

| **每个因子已解释方差** | |
| --- | --- |
| **Factor1** | **Factor2** |
| 3.0343866 | 2.5230573 |

| **最终的公因子方差估计: 总计 = 5.557444** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **x1** | **x2** | **x3** | **x4** | **x5** | **x6** | **x7** |
| 0.77104313 | 0.39210572 | 0.83732584 | 0.91041224 | 0.88935728 | 0.95185369 | 0.80534597 |

由此可以看出，经过正交旋转后，对factor1起作用的主要是x4和x6，对factor起作用的主要是x1和x5。