

1 3. 多変量解析 3 (因子分析)

因子分析

2

探索的因子分析

モデルを作成する

観測データから共通因子を抽出する

確認（認）的因子分析

モデルを検証する

(探索的) 因子分析

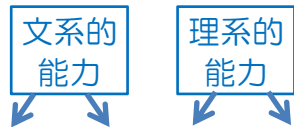
背景にある共通要素 (因子) を見つける

$$x_1 = a_{11}f_1 + a_{12}f_2 + \cdots + a_{1m}f_m + \varepsilon_1$$

$$x_2 = a_{21}f_1 + a_{22}f_2 + \cdots + a_{2m}f_m + \varepsilon_2$$

⋮

$$x_i = a_{i1}f_1 + a_{i2}f_2 + \cdots + a_{im}f_m + \varepsilon_i$$

 x_i : 観測可能な変数 f : 共通因子 (一般因子) ε : 特殊因子 (独自因子) a : 因子負荷量

英語	国語	社会	数学	理科	英語	国語	社会	数学	理科
60	65	76	54	45	85	73	77	85	91
70	68	67	66	56	90	82	84	93	94
43	56	54	23	34	77	63	75	74	84
82	78	78	45	69	69	58	62	76	67
89	89	93	78	82	56	63	45	72	82
67	78	92	56	45	59	73	67	54	62
87	92	70	69	72	79	87	65	83	85
72	58	74	78	68	68	66	74	80	78
67	82	77	67	62	82	72	81	89	93
58	67	84	78	83	72	78	83	95	89

5

```
eigo <- c(60, 70, 43, 82, 89, 67, 87, 72, 67, 58, 85, 90, 77, 69, 56, 59, 79, 68, 82, 72)
kokugo <- c(65, 68, 56, 78, 89, 78, 92, 58, 82, 67, 73, 82, 63, 58, 63, 73, 87, 66, 72, 78)
syakai <- c(76, 67, 54, 78, 93, 92, 70, 74, 77, 84, 77, 84, 75, 62, 45, 67, 65, 74, 81, 83)
sugaku <- c(54, 66, 23, 45, 78, 56, 69, 78, 67, 78, 87, 93, 74, 76, 72, 54, 83, 80, 89, 95)
rika <- c(45, 56, 34, 69, 82, 45, 72, 68, 62, 83, 91, 94, 84, 67, 82, 62, 85, 78, 93, 89)
kyoka <- cbind(eigo, kokugo, syakai, sugaku, rika)
colnames(kyoka) <- c("英語", "国語", "社会", "数学", "理科")
rownames(kyoka) <- 1:20
```

	英語	国語	社会	数学	理科
1	60	65	76	54	45
2	70	68	67	66	56
3	43	56	54	23	34
4	82	78	78	45	69
5	89	89	93	78	82
6	67	78	92	56	45
7	87	92	70	69	72

解の推定方法

6

主因子法

第1因子の因子寄与をもっとも大きくする

最小二乗法

観測相関行列とモデルの再生相関行列のずれを最小化する

最尤法

確率密度により解を推定する

モデルについての検定が可能

軸の回転

7

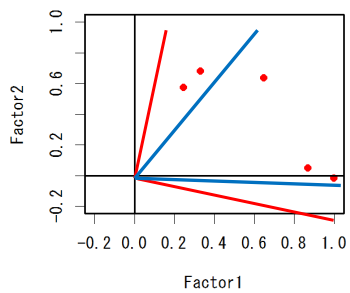
直交回転

バリマックス回転

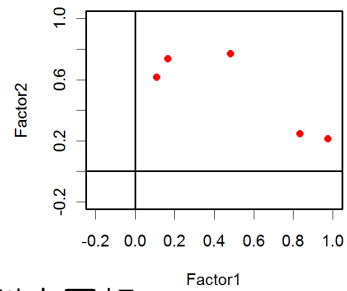
斜交回転

プロマックス回転

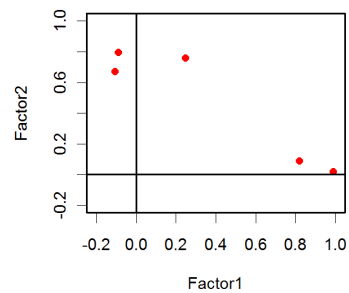
回転なし



直交回転



斜交回転



8

因子分析のためのRの関数

factanal()

fa()

psychパッケージ

factanal2()

青木先生の関数

因子分析 初期解 (回転なし) 最尤法 10

results <- factanal2(kyoka, rotation=c("none"))

H0: 2 factors are sufficient.

Chi sq.	d.f.	P value
1.395507	1	0.2374773

因子負荷量

Factor loadings(rotation:none)

	Factor1	Factor2	Communality
英語	0.6461454	0.63806732	0.8246338
国語	0.3301486	0.68133826	0.5732182
社会	0.2454512	0.57457783	0.3903816
数学	0.8672887	0.05023194	0.7547146
理科	0.9973613	-0.01647854	0.9950000

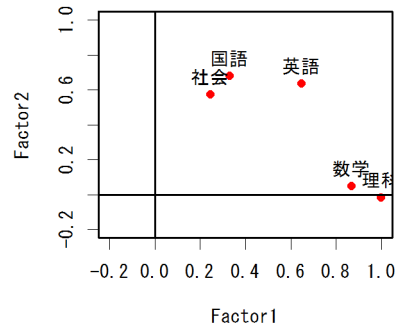
← 共通性

SS. loadings	2.3336676	1.2042	因子負荷量の二乗和 (固有値)
Proportion	46.6733514	24.0857	分散に対する寄与率
Cum. Prop.	46.6733514	70.7590	分散に対する累積寄与率

```

par(cex=1.8, pch=16, lwd=3)
plot(results$loadings, col="red", xlim=c(-0.2, 1),
ylim=c(-0.2, 1))
abline(v=0, h=0)
text(results$loadings, labels=colnames(kyoka),
pos=3)
    
```

11



```

バリマックス回転 (直交回転)
factanal2(kyoka, rotation=c("varimax"))
    
```

12

H0: 2 factors are sufficient. モデルの適合度検定
 Chi sq. d.f. P value (pは大きいほどよい)
 1.395507 1 0.2374773

バリマックス回転後の因子負荷量 共通性 = 1-独自性

	Factor1	Factor2	Communality
英語	0.4826418	0.7692144	0.8246338
国語	0.1651416	0.7388831	0.5732182
社会	0.1071776	0.6155477	0.3903816
数学	0.8326691	0.2477399	0.7547146
理科	0.9745718	0.2126286	0.9950000
SS. loadings	1.9148300	1.6231238	3.5379538
Proportion	38.2965991	32.4624763	70.7590754
Cum. Prop.	38.2965991	70.7590754	NA

プロマックス回転 (斜交回転)

13

factanal2(kyoka)

H0: 2 factors are sufficient.

Chi sq.	d.f.	P value
1.395507	1	0.2374773

プロマックス回転後の因子負荷量

Factor loadings (rotation: promax)

	Factor1	Factor2	Communality
英語	0.24649429	0.75992183	0.8246338
国語	-0.08954746	0.79767067	0.5732182
社会	-0.10783511	0.67141882	0.3903816
数学	0.81981229	0.09110478	0.7547146
理科	0.98779235	0.01922607	0.9950000

factanal2(kyoka, scores="Bartlett")

14

因子得点

Factor scores (Bartlett)

	Factor1	Factor2
1	-1.56788146	-0.68447667
2	-0.92682428	-0.19989255
3	-2.19827871	-2.38284265
4	-0.21697106	0.85396196
5	0.54266994	1.85666801
6	-1.59235514	0.44640475
7	-0.03492474	1.51619720
8	-0.21638092	-0.36809545
9	-0.59219442	0.16849307
10	0.65475839	-0.94620121
11	1.09890578	0.75347486
12	1.26201307	1.45803747
13	0.69879842	-0.06307521

15

```
(results <- factanal2(kyoka, scores="Bartlett"))
```

Call:

```
factanal(x = dat, factors = factors, scores = scores,
rotation = rotation)
```

Uniquenesses: 独自性 (1 - 共通性)

```
英語 国語 社会 数学 理科
0.175 0.427 0.610 0.245 0.005
```

Loadings: 因子負荷量

```
          Factor1 Factor2
英語  0.246    0.760
国語           0.798
社会 -0.108    0.671
数学  0.820
理科  0.988
```

0.1以下の時は非表示

16

	Factor1	Factor2	
SS loadings	1.728	1.673	因子負荷量の2乗和
Proportion Var	0.346	0.335	寄与率
Cumulative Var	0.346	0.680	累積寄与率

因子間の相関係数

Factor Correlations:

```
          Factor1 Factor2
Factor1  1.000 -0.498
Factor2 -0.498  1.000
```

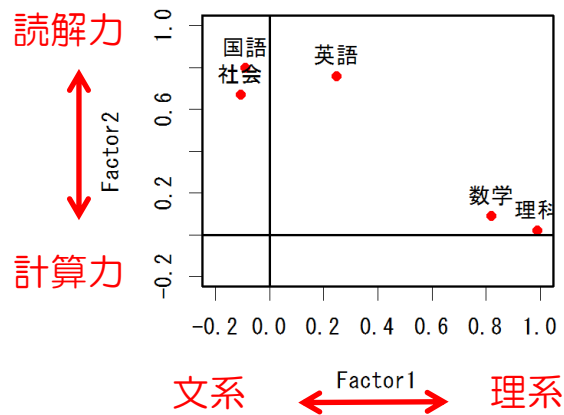
モデルの適合度検定

Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 1.4 on 1 degree of freedom.
The p-value is 0.237

因子負荷量の散布図

17

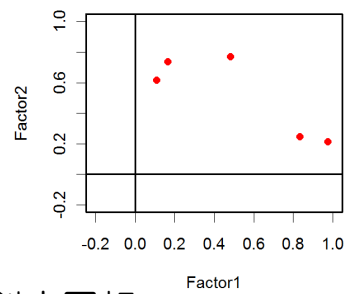
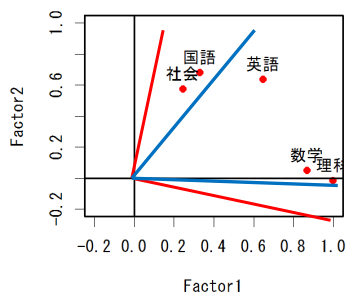
```
plot(results$loadings, col="red", xlim=c(-1.2,1) ,ylim=c(-0.2,1))
abline(v=0, h=0)
text(results$loadings, labels=colnames(kyoka), pos=3)
```



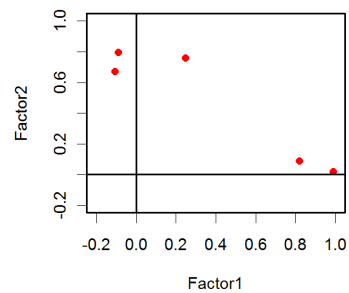
回転なし

直交回転

18



斜交回転



因子得点の散布図

19

```
plot(results$scores, col="red")  
abline(v=0, h=0)  
text(results$scores, labels=rownames(kyoka),  
pos=3)
```

