

検証的因子分析

青木繁伸

2020年3月17日

1 目的

2 使用法

```
import sys
sys.path.append("statlib")
from multi import cfa
cfa(r, n, loc, init_val=0.5, verbose=True)
```

2.1 引数

r	相関係数行列
n	サンプルサイズ
loc	因子負荷量または設定情報（使用例を参照のこと）
init_val	推定パラメータの初期値
verbose	必要最小限のプリント出力をする

2.2 戻り値の名前

"loadings"	因子負荷量
"faccor"	因子間相関係数行列
"chisq"	あてはまりの良さを示す χ^2 値
"df"	自由度
"pvalue"	p 値
"GFI"	GFI
"AGFI"	AGFI
"SRMR"	SRMR
"RMSEA"	RMSEA

3 使用例

```
r = [[1, 0.412, 0.521, 0.358, 0.334, 0.346],
      [0.412, 1, 0.495, 0.499, 0.293, 0.248],
      [0.521, 0.495, 1, 0.525, 0.364, 0.323],
      [0.358, 0.499, 0.525, 1, 0.607, 0.517],
      [0.334, 0.293, 0.364, 0.607, 1, 0.506],
      [0.346, 0.248, 0.323, 0.517, 0.506, 1]]
n = 100
```

3.1 各変数がどれか一つの因子に含まれる場合

それぞれの変数が、どの因子に含まれるかをベクトルで表す。

この指定法では、ある変数が複数の因子に含まれるという状況は指定できない。そのような場合には次項の指定法による。

下の例では、1～3番目の変数は第2因子、4～6番目の変数は第1因子に含まれることを示している。

```
loc = [2, 2, 2, 1, 1, 1]

import sys
sys.path.append("statlib")
from multi import cfa

a = cfa(r, n, loc)
```

Loadings

```
[[0.        0.6419522 ]
 [0.        0.65232029]
 [0.        0.78165933]
 [0.86290139 0.        ]
 [0.70953311 0.        ]
 [0.62220815 0.        ]]
```

Factor correlation

```
[[1.        0.7402033]
 [0.7402033 1.        ]]
```

chisq = 10.081, df = 8, p value = 0.25940

GFI = 0.96672, AGFI = 0.91265, SRMR = 0.035631, RMSEA = 0.051256

3.2 一般的な場合

各変数（行）がどの因子（列）に含まれるかを二次元配列で表す。含まれる場合は1をセットする。

この指定法では、ある変数が複数の因子に含まれるという状況も指定できる。

下の例では、1～3番目の変数は第2因子、5、6番目の変数は第1因子、4番目の変数は第1、第2因子に共に含まれることを示している。

```
loc = [[0,1],
        [0,1],
        [0,1],
        [1,1],
        [1,0],
        [1,0]]

a = cfa(r, n, loc)
```

Loadings

```
[[0.        0.62951454]
 [0.        0.66158537]
 [0.        0.78273515]
 [0.60757149 0.28688868]
 [0.76758737 0.        ]
 [0.65922144 0.        ]]
```

Factor correlation

```
[[1.        0.63072538]
 [0.63072538 1.        ]]
```

chisq = 7.3879, df = 7, p value = 0.38964

GFI = 0.97652, AGFI = 0.92956, SRMR = 0.029429, RMSEA = 0.023659

もちろん、前項に示したモデルを本項のやり方で指定しても同じ答えになる。

```
loc = [[0,1],
        [0,1],
        [0,1],
        [1,0],
        [1,0],
        [1,0]]

a = cfa(r, n, loc)
```

Loadings

```
[[0.        0.6419522 ]
 [0.        0.65232029]
 [0.        0.78165933]
 [0.86290139 0.        ]
 [0.70953311 0.        ]
 [0.62220815 0.        ]]
```

Factor correlation

```
[[1.        0.7402033]
 [0.7402033 1.        ]]
```

[0.7402033 1.]]

chisq = 10.081, df = 8, p value = 0.25940

GFI = 0.96672, AGFI = 0.91265, SRMR = 0.035631, RMSEA = 0.051256