



因子分析

心理データ解析演習

2007/12/12

M1 山添愛



因子分析とは？



因子分析とは？

- 多変量解析の一種

- ある **観測変数** (質問項目に対する回答) が
どのような **潜在変数** (因子) から
影響を受けているかを探るもの

- 例
 - 授業評価についてのアンケート

授業評価についてのアンケート

性別 1. 男 2. 女

所属学部・学科

外国語学部 11. 外国語学科 12. 国際関係学科 経済学部 21. 経済学科 22. 経営情報学科
文学部 31. 比較文化学科 32. 人間関係学科 法学部 41. 法律学科 42. 政策科学科 50. その他

1. 内容は理解しやすかった

全く当てはまらない 1 _____ 2 _____ 3 _____ 4 _____ 5
少し当てはまらない 2 _____ 3 _____ 4 _____ 5
どちらでもない 3 _____ 4 _____ 5
少し当てはまる 4 _____ 5
よく当てはまる 5

2. 内容は面白かった

全く当てはまらない 1 _____ 2 _____ 3 _____ 4 _____ 5
少し当てはまらない 2 _____ 3 _____ 4 _____ 5
どちらでもない 3 _____ 4 _____ 5
少し当てはまる 4 _____ 5
よく当てはまる 5

3. 内容はためになった

全く当てはまらない 1 _____ 2 _____ 3 _____ 4 _____ 5
少し当てはまらない 2 _____ 3 _____ 4 _____ 5
どちらでもない 3 _____ 4 _____ 5
少し当てはまる 4 _____ 5
よく当てはまる 5

4. 授業の進み具合は適切だった

全く当てはまらない 1 _____ 2 _____ 3 _____ 4 _____ 5
少し当てはまらない 2 _____ 3 _____ 4 _____ 5
どちらでもない 3 _____ 4 _____ 5
少し当てはまる 4 _____ 5
よく当てはまる 5

5. 教員の熱意は感じられた

全く当てはまらない 1 _____ 2 _____ 3 _____ 4 _____ 5
少し当てはまらない 2 _____ 3 _____ 4 _____ 5
どちらでもない 3 _____ 4 _____ 5
少し当てはまる 4 _____ 5
よく当てはまる 5

モデル

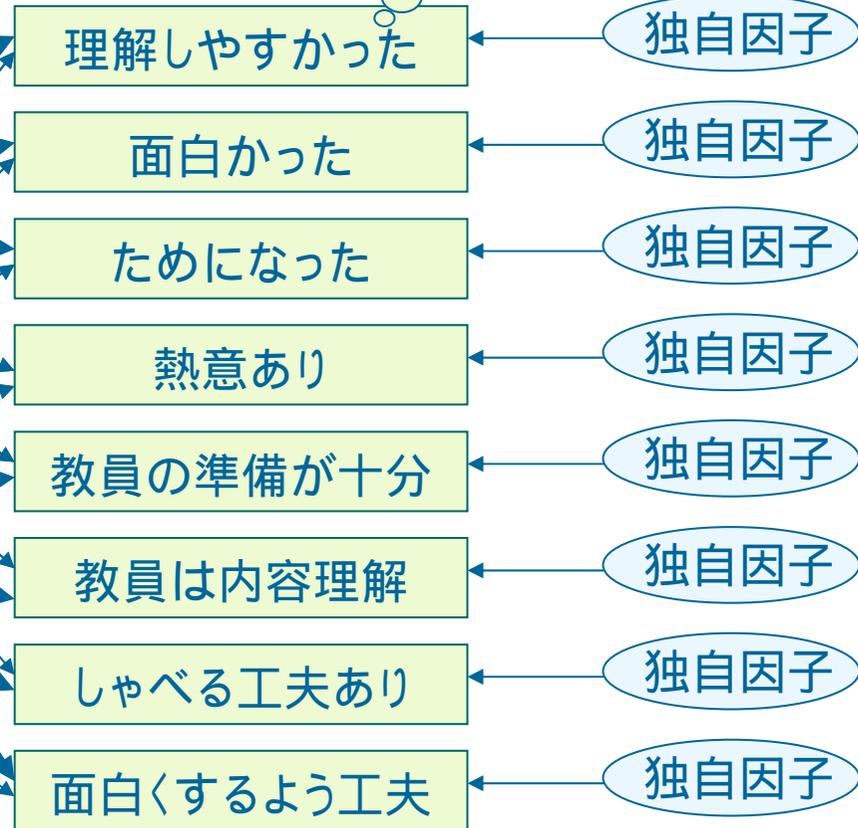
質問項目

授業内容に関する因子

教員の努力に関する因子

⋮

潜在変数

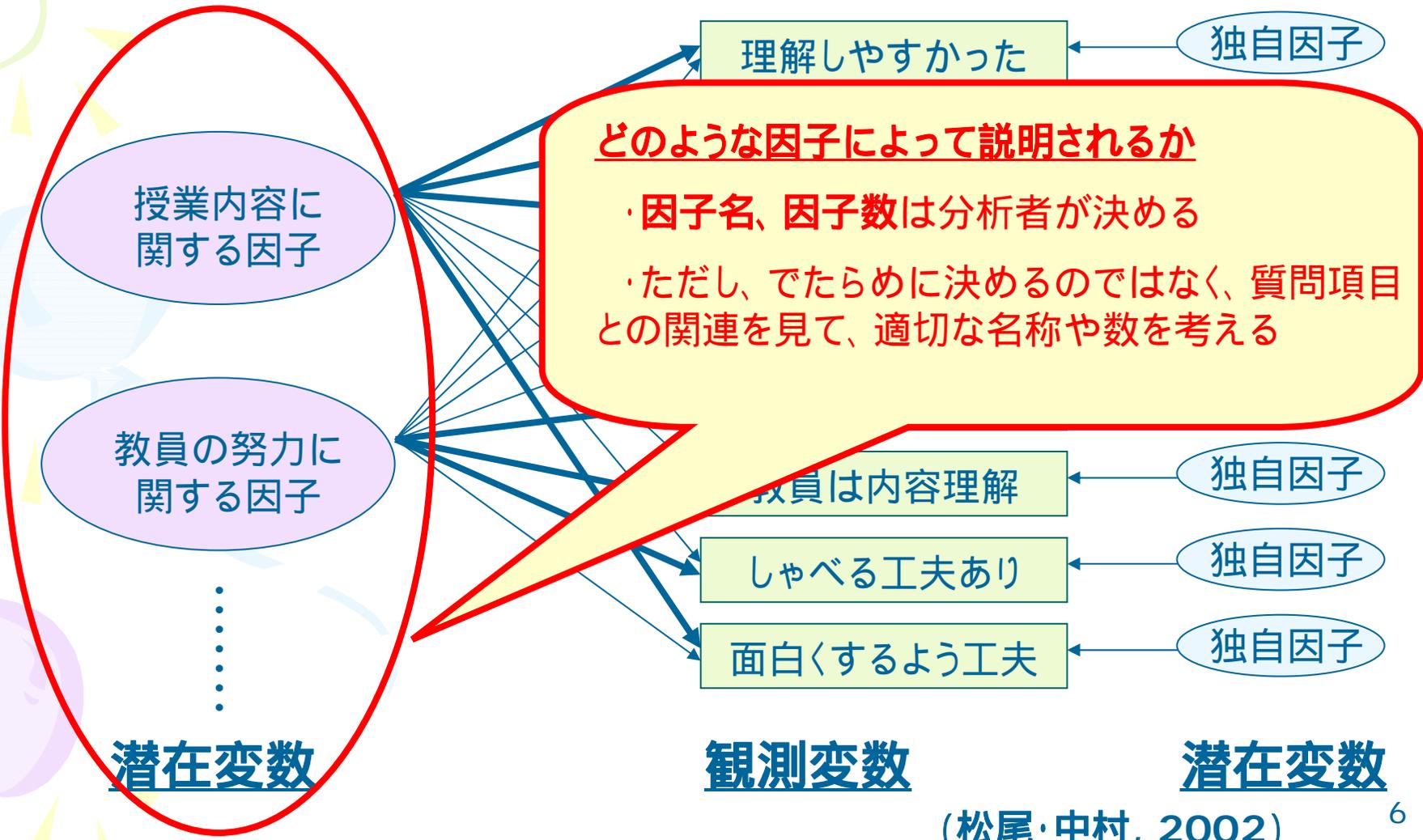


観測変数

潜在変数

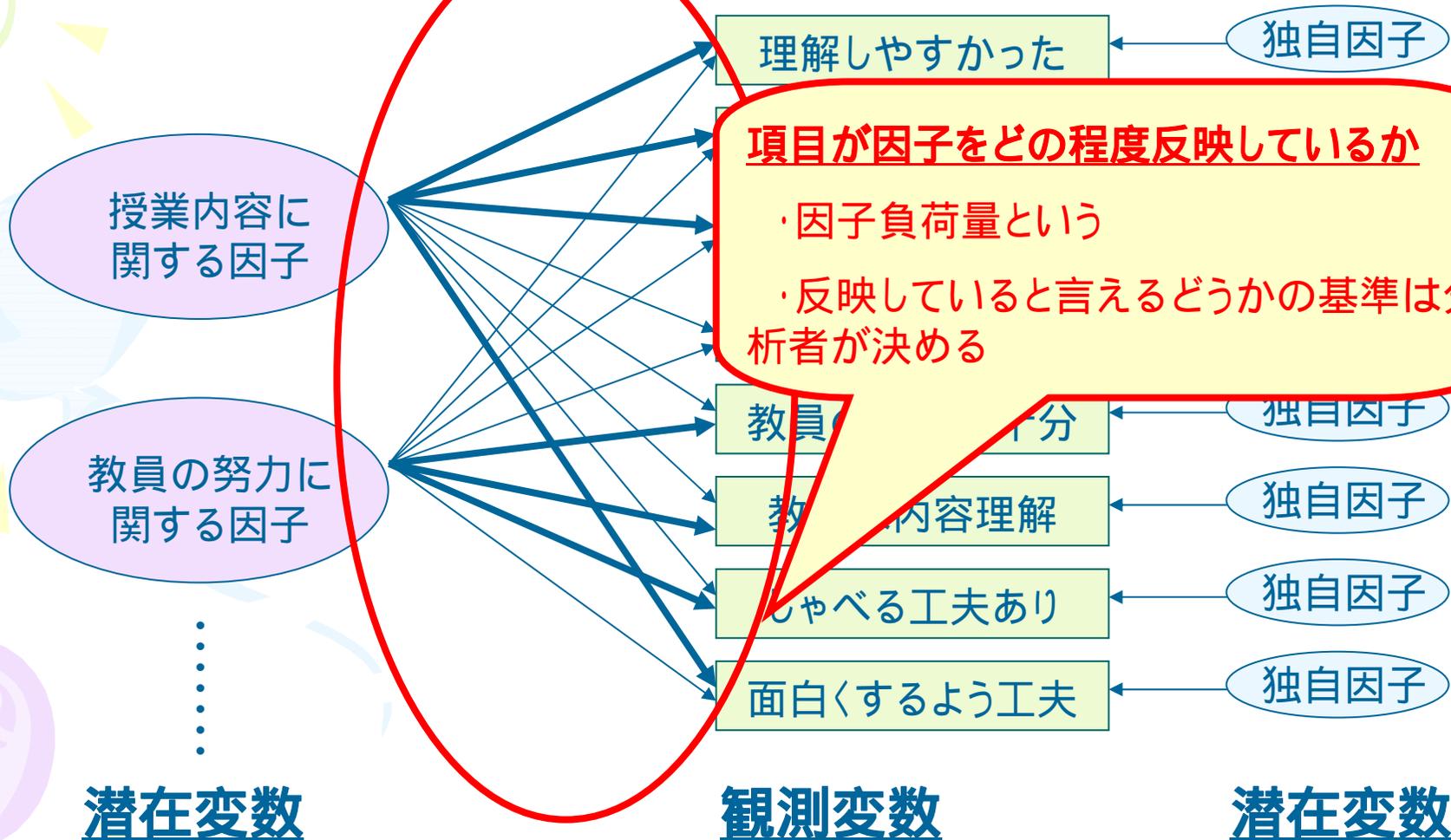
(松尾・中村, 2002)

わかること



(松尾・中村, 2002)

わかること

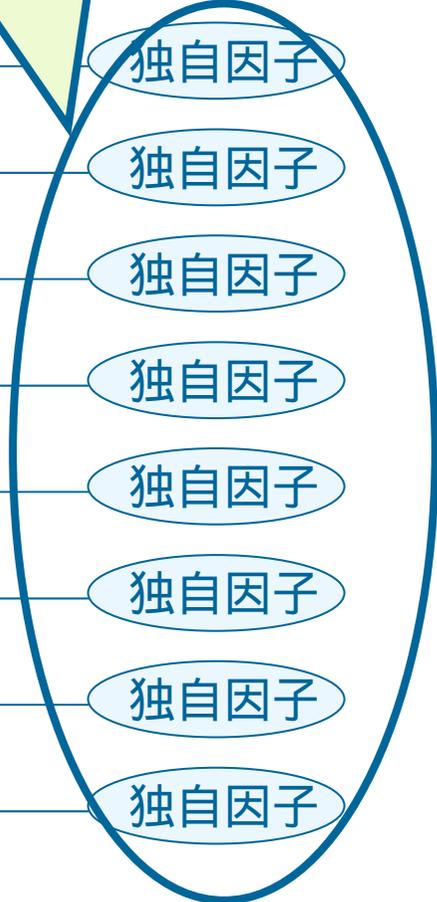
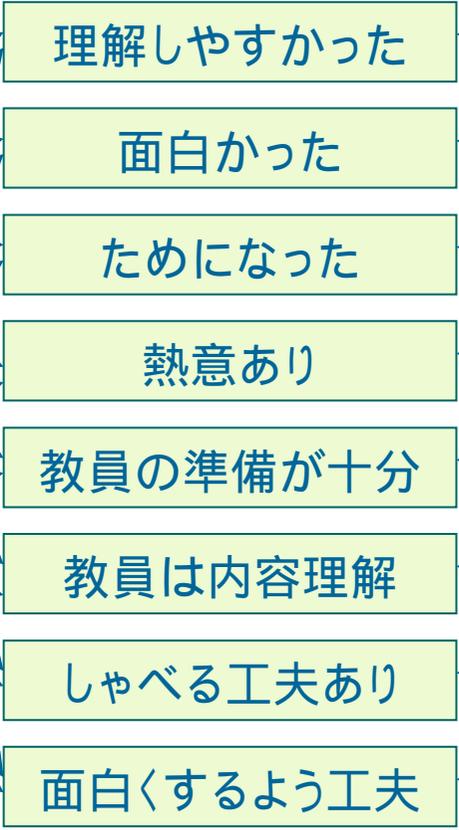
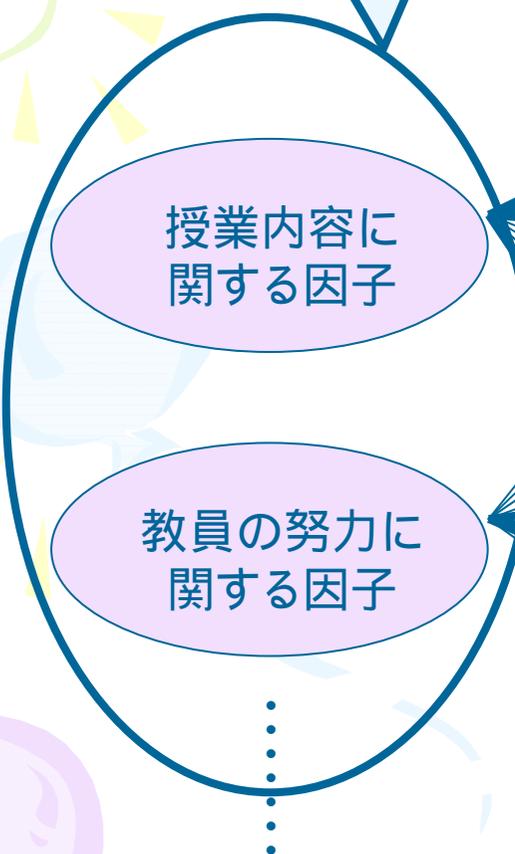


(松尾・中村, 2002)

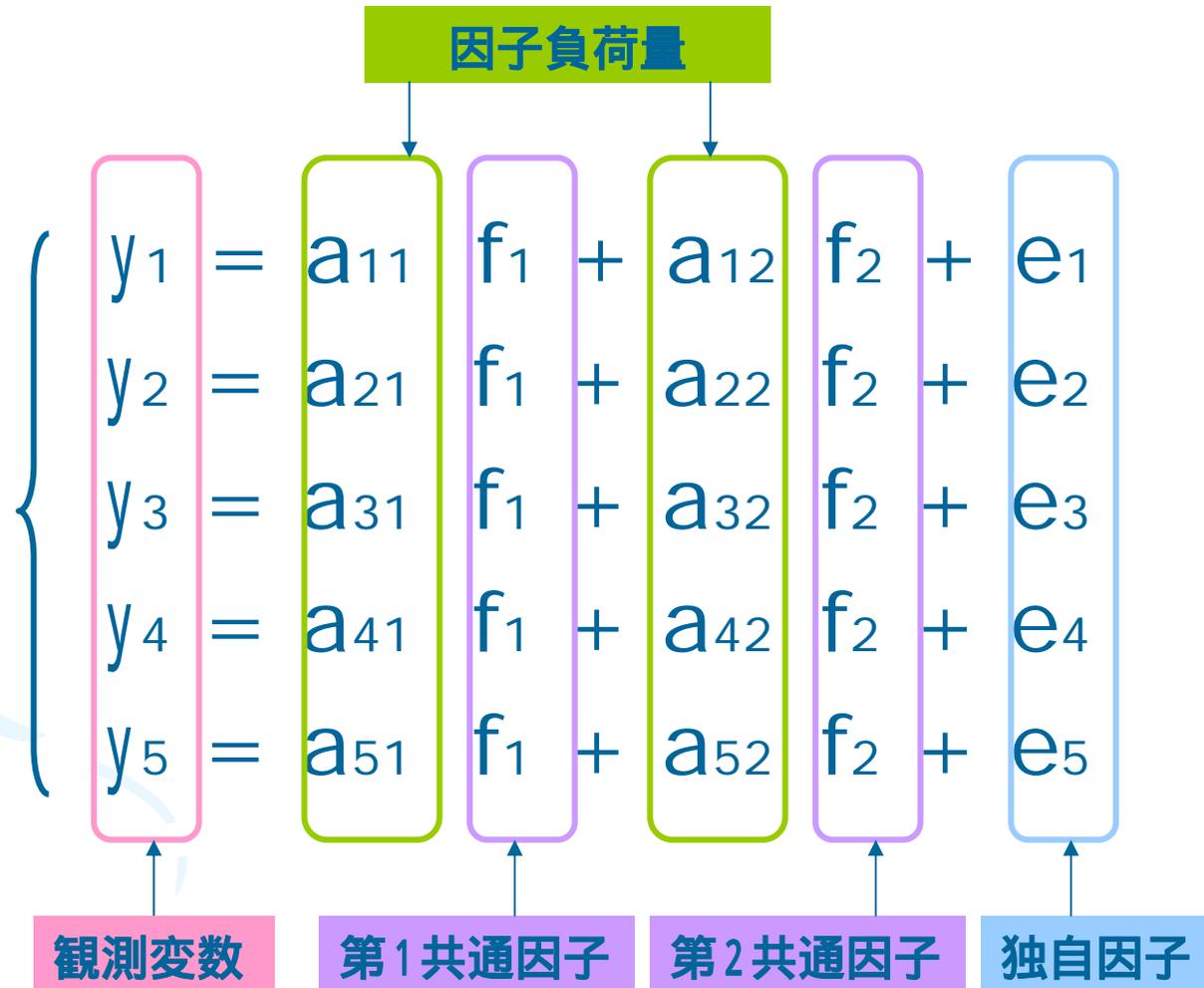
ちなみに

共通因子
複数の観測変数
に関わる因子

独自因子
ある観測変数だけに
関わる因子

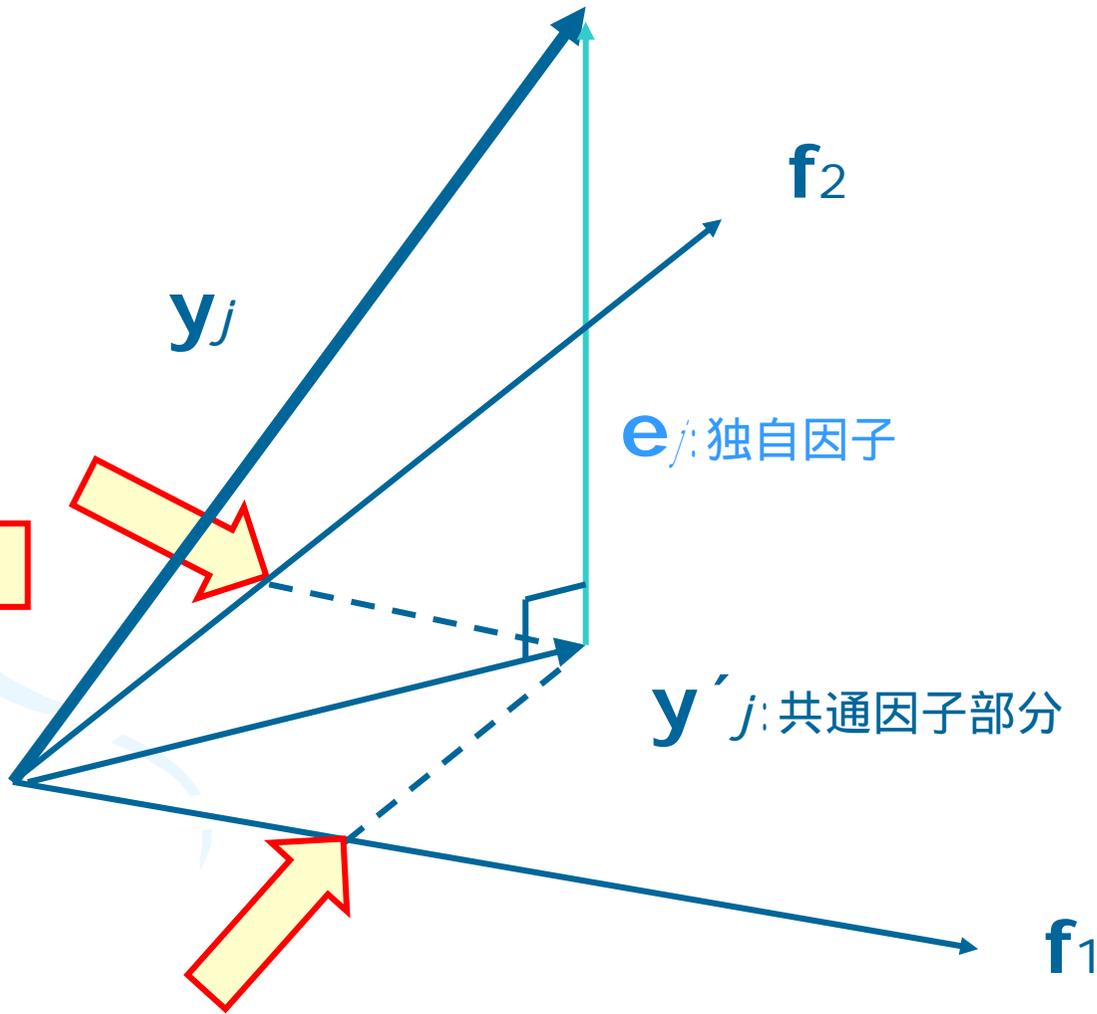


式であらわすと...



ベクトルであらわすと...

因子負荷





大事な用語

- **共通性**

- 観測変数の分散のうち、共通因子で説明できる部分の割合

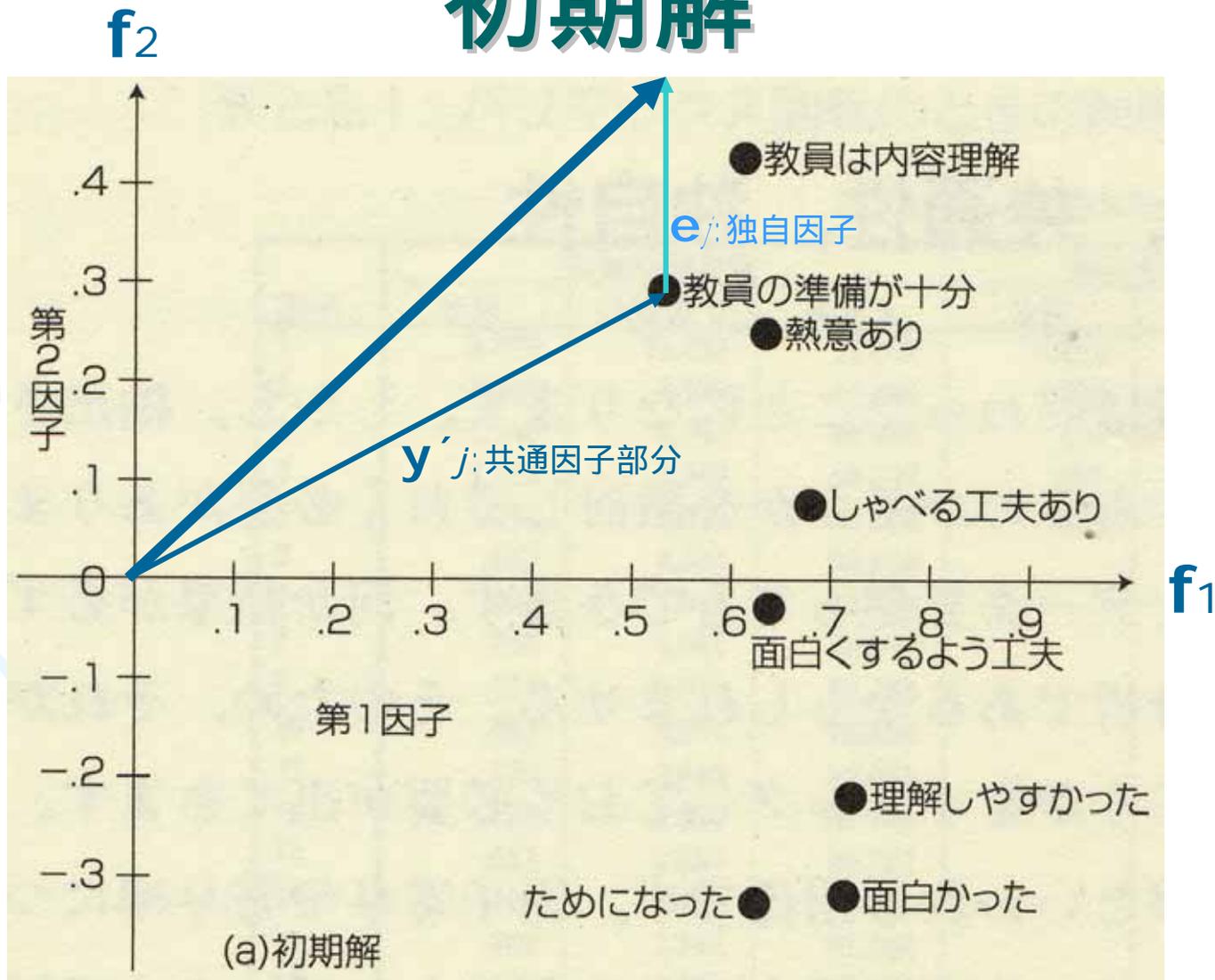
- (共通性) = (共通因子部分の分散) / (観測変数の分散)
- ベクトル y' の長さの2乗

- **独自性**

- 観測変数の分散のうち、共通因子で説明できない部分の割合

- (共通性) + (独自性) = 1

初期解

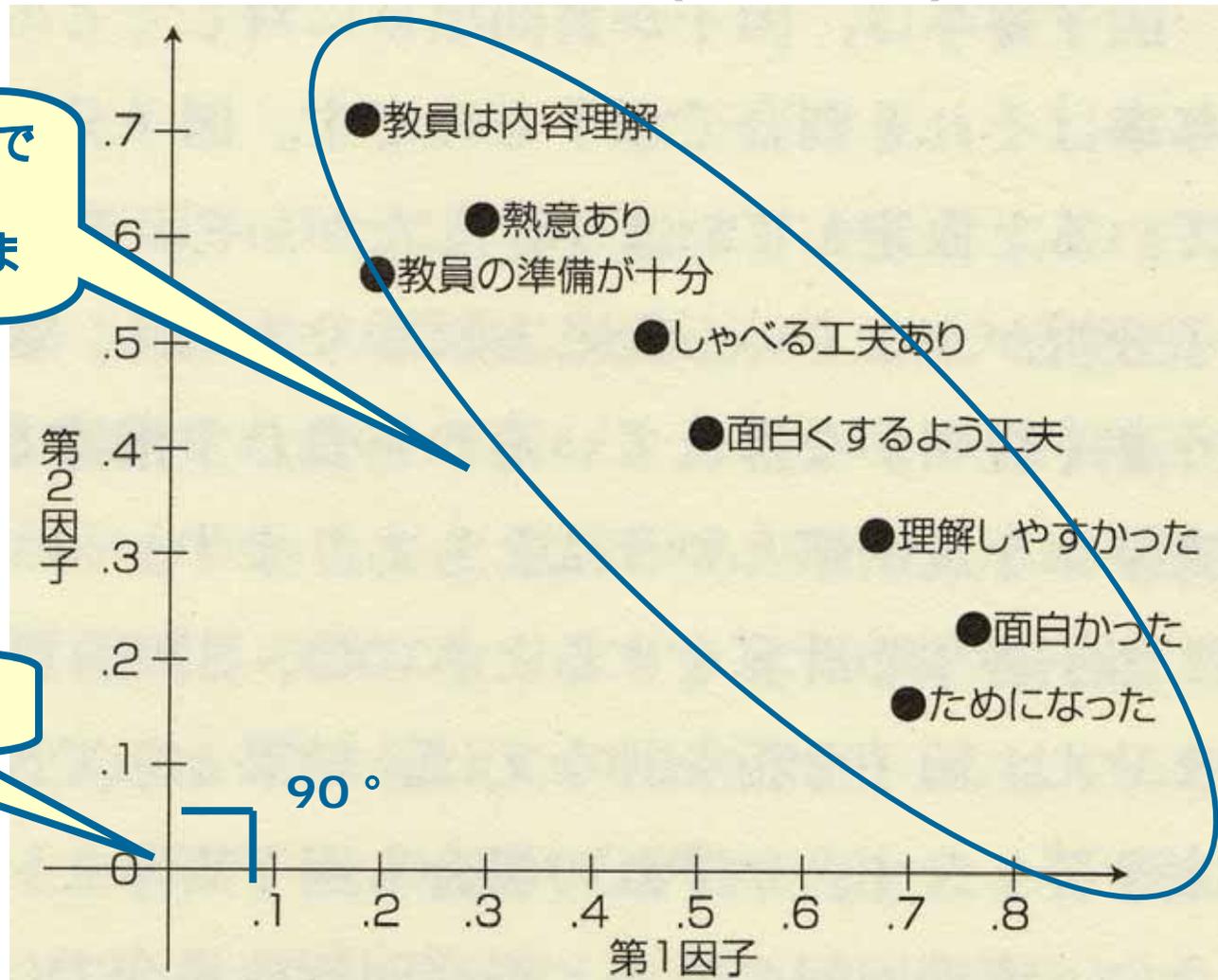


(松尾・中村, 2002)

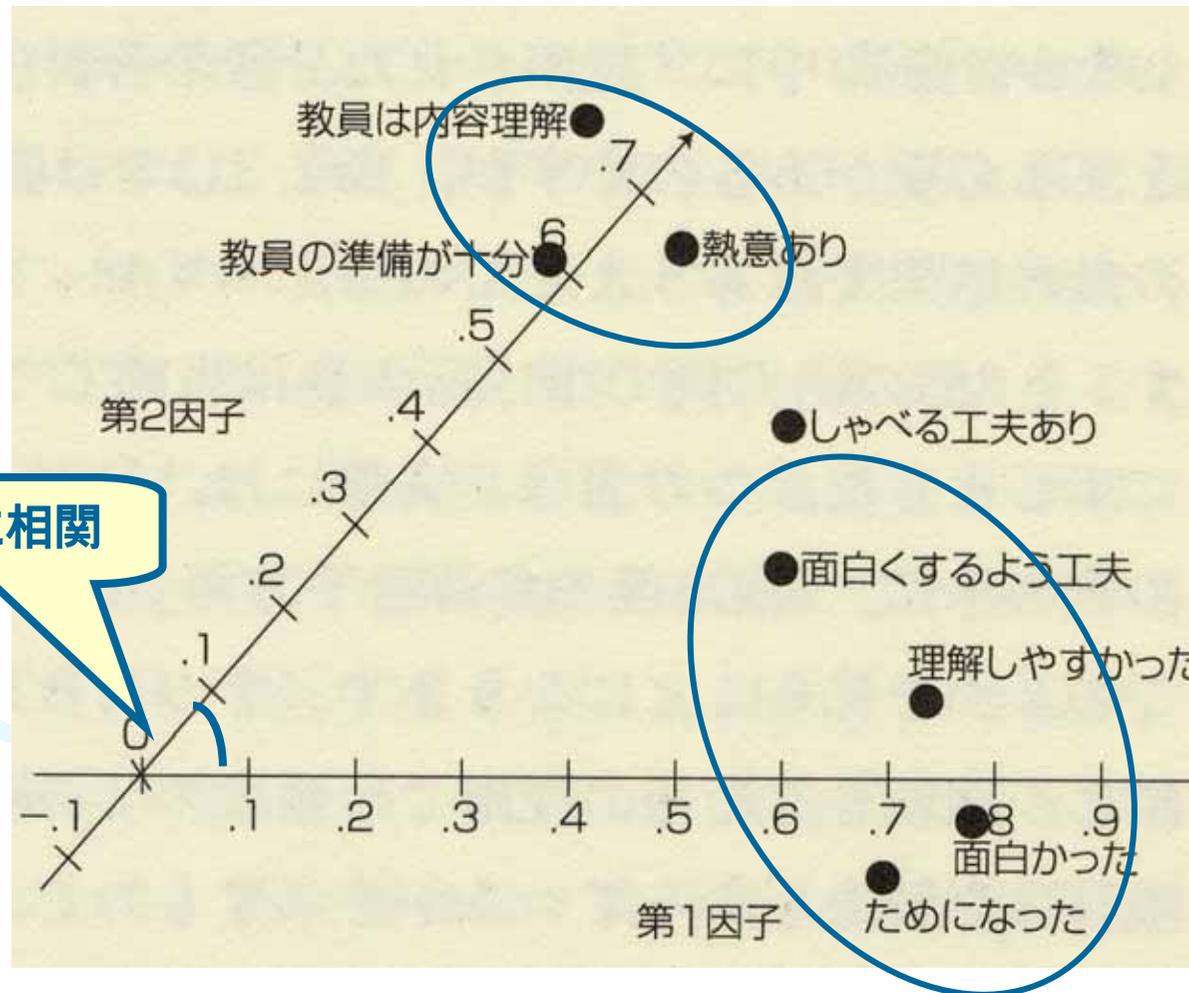
バリマックス回転(直交)

軸を回転させただけで
プロットの点の
位置関係はそのまま

因子間は無相関



プロマックス回転(斜交)



(松尾・中村, 2002)

まとめ：因子分析の流れ

START

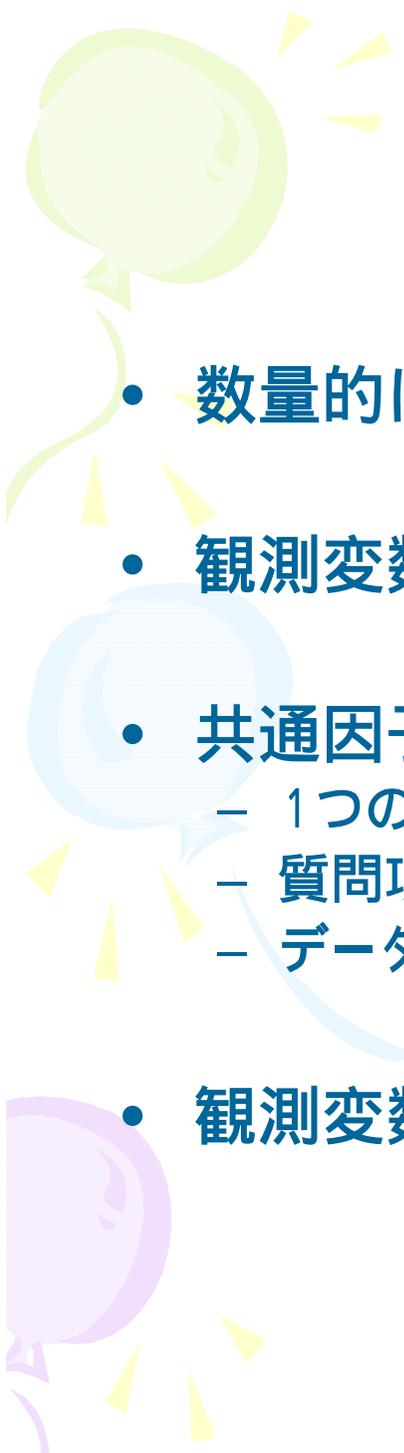
共通性の推定

共通因子と独自因子に分解

共通因子を複数の因子に分解

回転

因子負荷量の算出



補足

- 数量的に表わされているデータ(間隔尺度か比例尺度)
- 観測変数間の相関関係は直線的でなければならない
- 共通因子の個数を分析前に仮定しておく必要がある
 - 1つの因子には最低でも3～4つの項目が関連していることが必要
 - 質問項目は想定する因子の数の3～4倍
 - データ数(回答者数)は、項目数の5～10倍
- 観測変数の選定が大事(仮説をある程度立ててから)

The background features three large, stylized swirls in purple, green, and blue. Interspersed among these swirls are several yellow starburst shapes, each consisting of five triangular points radiating from a central point. The overall aesthetic is bright and celebratory.

**SPSSで
やってみよう**

社会医療に関するアンケート

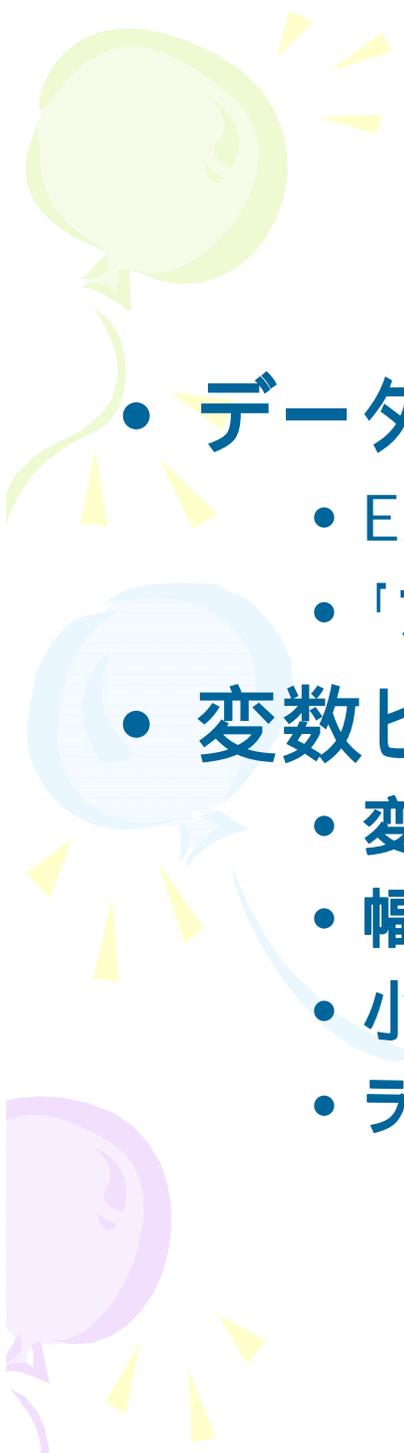
*無題2 [データセット1] - SPSS データ エディタ

ファイル(F) 編集(E) 表示(V) データ(D) 変換(T) 分析(A) グラフ(G) ユーティリティ(U) ウィンドウ(W) ヘルプ(H)

1: ストレス 3

	ストレス	健康行動	健康習慣	社会支援	社会役割	健康度	生活環境	医療機関
1	3	0	5	4	8	3	2	3
2	3	0	1	2	5	3	2	2
3	3	1	5	8	7	3	3	3
4	3	2	7	7	6	3	2	3
5	2	1	5	8	4	2	2	4
6	7	1	2	2	6	4	5	2
7	4	1	3	3	5	3	3	3
8	1	3	6	8	8	2	3	2
9	5	4	5	6	6	3	3	3
10	3	1	5	3	6	3	3	3
11	5	1	4	7	5	5	3	3
12	6	1	2	7	6	3	4	3
13	4	0	0	2	7	3	3	3
14	5	0	0	0	5	3	3	2
15	7	2	3	4	8	4	4	3
16	3	0	1	8	5	3	3	3
17	0	1	3	8	7	3	3	3
18	4	0	5	3	5	3	3	2
19	5	1	7	3	7	4	4	3
20	3	1	5	0	5	3	3	3
21	3	1	6	8	6	3	2	3
22	1	1	3	3	4	1	3	3
23	5	0	0	8	8	5	4	5

(石村, 1998)



準備

- データビューに入力
 - Excelに入力したデータを直接開ける
 - 「ファイル」「開く」「データ」ファイルを選択
- 変数ビューの設定
 - 変数の名前: 今回はExcelにすでに書いてある
 - 幅(桁数): デフォルトは「11」
 - 小数の桁数: デフォルトは「0」
 - ラベル: 分析結果の表に出力される
実際の質問項目をここに入れる

因子分析の手順

変数の選択

因子抽出

因子軸の回転

因子得点

記述統計

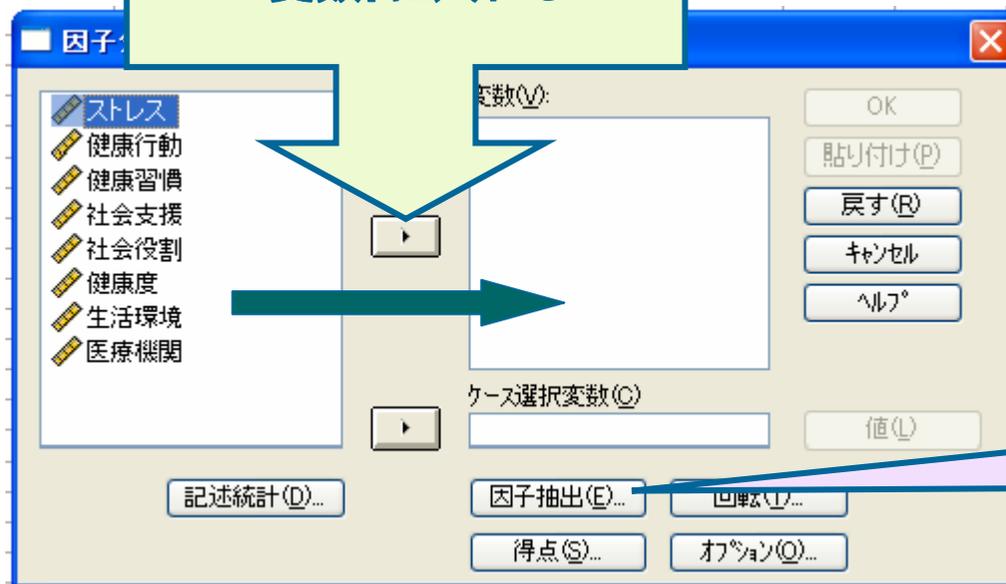
出力チェック

- ・抽出法は適切か
- ・因子数の決定
- ・因子の解釈
- ・項目の取舍選択

変数の選択

- 「分析」 「データの分解」 「因子分析」

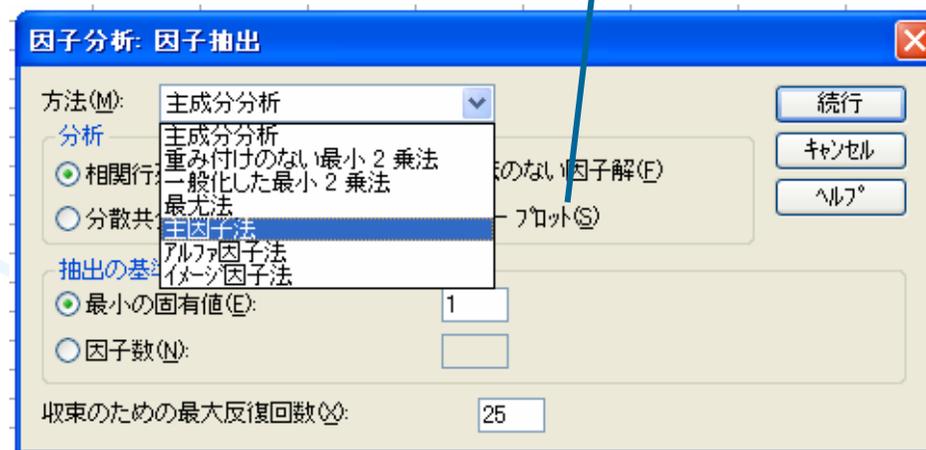
分析対象の項目を選択し
ここをクリックして
「変数」に入れる



次は
ここを
クリック

因子抽出

- 「方法」: 「主因子法」を選択
- 「スクリープロット」にチェックを入れる
- 「続行」



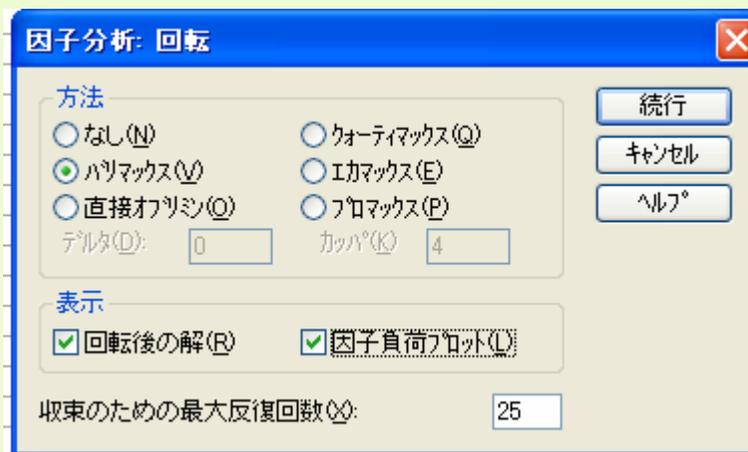


参考：因子抽出法

- **主成分分析**
 - 因子分析のときには使わない
- **主因子法**
 - 第1因子から順に因子寄与が最大となるように因子を抽出
- **重み付けのない最小二乗法**
 - 元のデータと因子分析のモデルから算出される共分散行列の間の差を最小にするように行う
- **一般化された最小二乗法**
 - 上記の最小二乗法に重み付けをし、尺度の単位の影響されないように行う。適合度の検定が可能。
- **最尤法**
 - 因子分析に関わるパラメータから尤度といわれる指標を算出し、これを最大にするように行う。適合度の検定が可能。

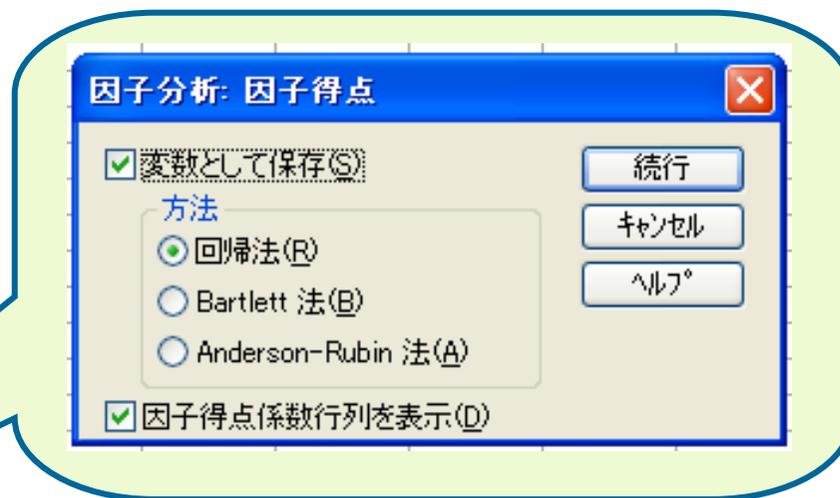
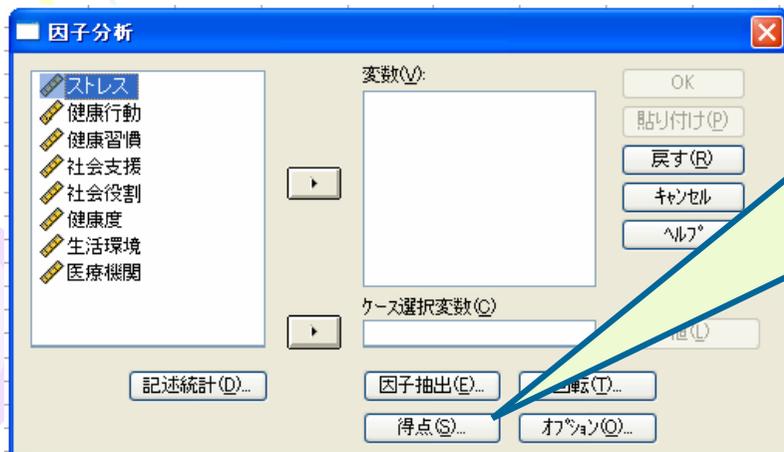
因子軸の回転

- 「バリマックス」にチェック
- 「因子負荷プロット」
- 「続行」



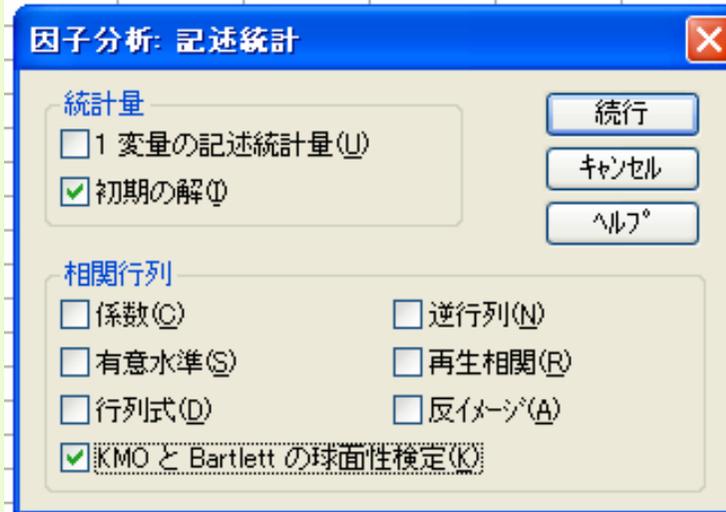
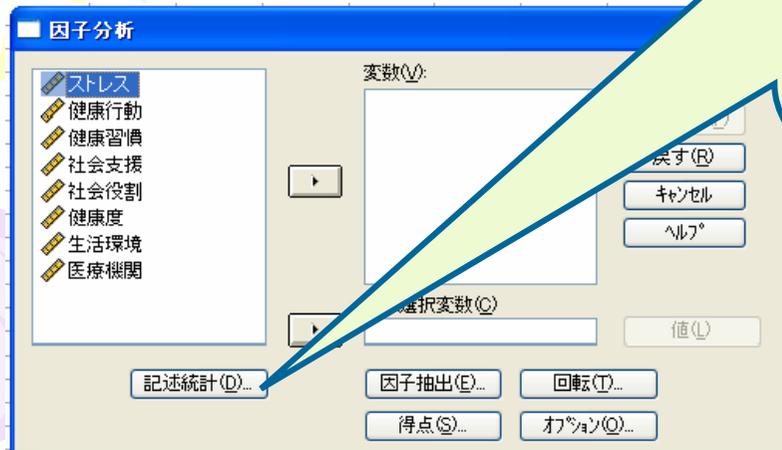
因子得点

- 「変数として保存」
- 「回帰法」
- 「因子得点係数行列を表示」
- 「続行」



記述統計

- 「KMOとBartlettの球面性検定」
- 「続行」



出力を見るポイント

妥当性

KMO および Bartlett の検定

Kaiser-Meyer-Olkin の標本妥当性の測度		.637
Bartlett の球面性検定	近似カ2乗	222.694
	自由度	28
	有意確率	.000

0.5以上であれば
問題がない

0.05より小さければ
観測変数の間に
関連がある

・0に近い項目は取り除くことも
・1より大きいと、データ数が少ない、抽出法が合っていないなどの問題がある可能性がある

共通性

	初期	因子抽出後
ストレス	.214	.332
健康行動	.067	.151
健康習慣	.107	.190
社会支援	.094	.147
社会役割	.124	.214
健康度	.228	.559
生活環境	.147	.662
医療機関	.111	.126

因子抽出法: 主因子法

出力を見るポイント

因子数の決定

固有値: 因子負荷量の2乗和

説明された分散の合計

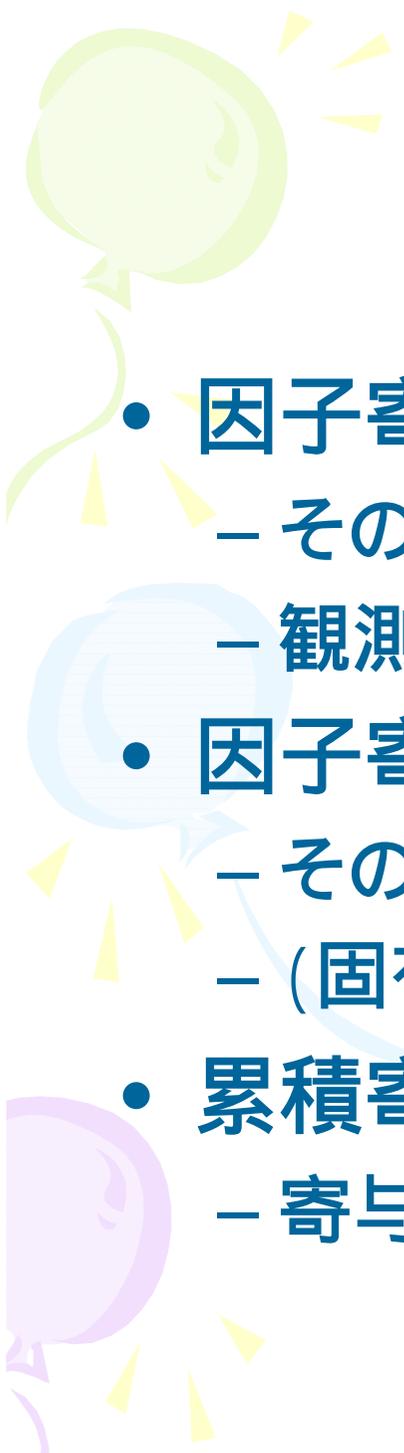
因子	初期の固有値			抽出後の負荷量平方和			回転後の負荷量平方和		
	合計	分散の %	累積 %	合計	分散の %	累積 %	合計	分散の %	累積 %
1	2.051	25.640	25.640	1.417	17.708	17.708	.966	12.069	12.069
2	1.161	14.512	40.152	.615	7.686	25.394	.801	10.016	22.085
3	1.062	13.279	53.431	.351	4.383	29.777	.615	7.692	29.777
4	.974	12.170	65.601						
5	.846	10.579	76.181						
6	.729	9.112	85.293						
7	.645	8.068	93.361						
8	.531	6.639	100.000						

因子抽出法: 主因子法

因子寄与

因子寄与率

累積寄与率



参考：因子寄与

- 因子寄与

- その因子がデータをどの程度説明しているか
- 観測変数の数に影響される

- 因子寄与率(説明率)

- その因子がデータ全体を説明している割合
- $(\text{固有値}) / (\text{変数の数}) \times 100$

- 累積寄与率

- 寄与率を第1因子から順番に加算したもの

出力を見るポイント

因子数の決定

説明された分散の合計

因子	初期の固有値			抽出後の負荷量平方和			回転後の負荷量平方和		
	合計	分散の %	累積 %	合計	分散の %	累積 %	合計	分散の %	累積 %
1	<u>2.051</u>	25.640	25.640	1.417	17.708	17.708	.966	12.069	12.069
2	<u>1.161</u>	14.512	40.152	.615	7.686	25.394	.801	10.016	22.085
3	<u>1.062</u>	13.279	53.431	.351	4.383	29.777	.615	7.692	<u>29.777</u>
4	.974	12.170	65.601						
5	.846	10.579	76.181						
6	.729	9.112	85.293						
7	.645	8.068	93.361						
8	.531	6.639	100.000						

3つの因子でデータ全体の
29.777%を説明できている

因子抽出法: 主因子法

8個の因子のうち
固有値が1より
大きい因子だけを
取り上げる

因子分析: 因子抽出

方法(M): 主成分分析

分析

- 相関行列
- 分散共分散行列

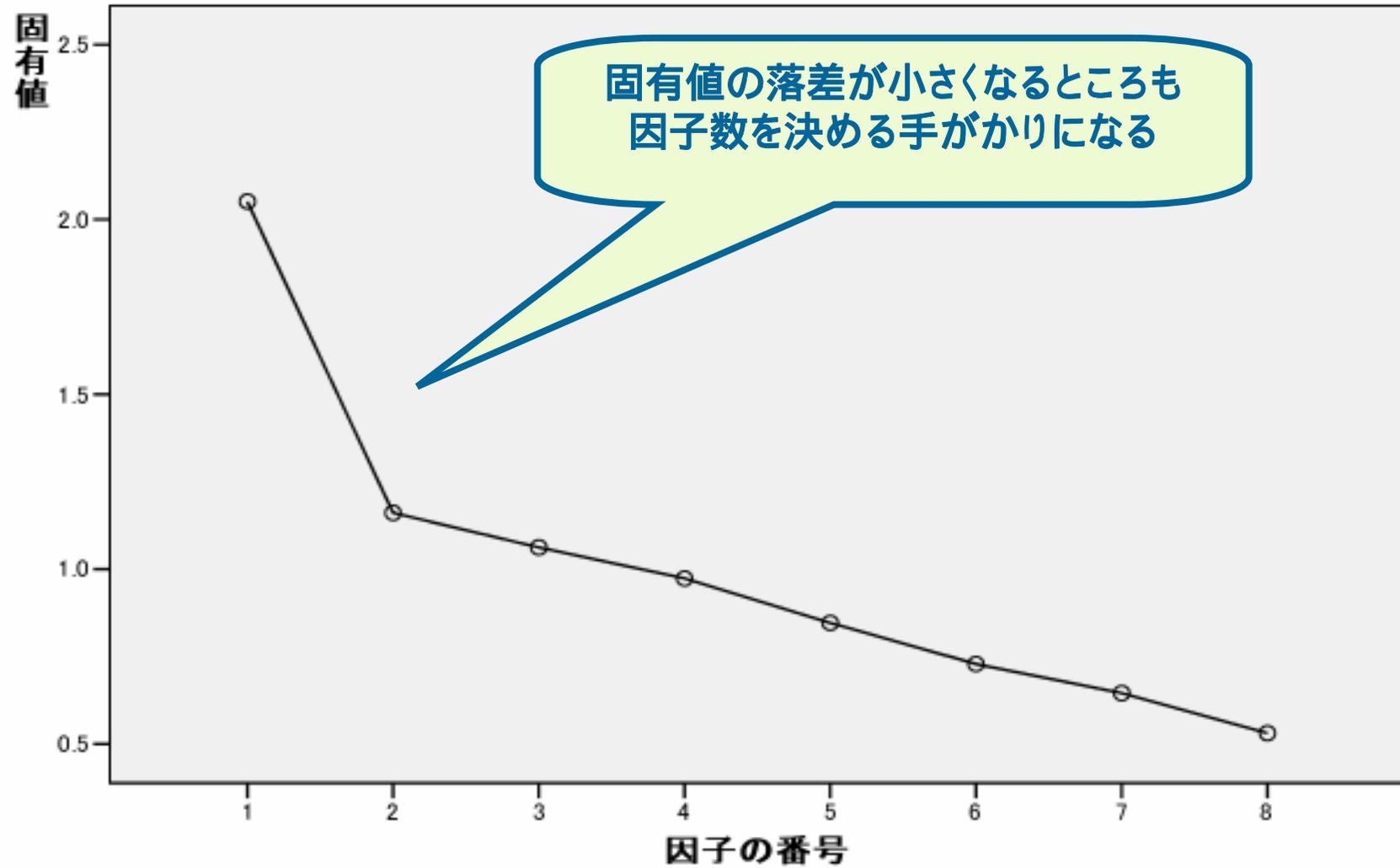
抽出の基準

- 最小の固有値(E): 1
- 因子数(N):

収束のための最大反復回数(O): 25

続行
キャンセル
ヘルプ

因子のスクリープロット



出力を見るポイント

因子の解釈

因子負荷量

因子行列^a

	因子		
	1	2	3
ストレス	.546	-.143	.113
健康行動	-.230	.047	.310
健康習慣	-.378	.026	.215
社会支援	-.325	.177	.101
社会役割	-.326	.248	.215
健康度	.635	-.208	.335
生活環境	.496	.644	.041
医療機関	.245	.202	-.158

因子抽出法: 主因子法

- a. 3 個の因子の抽出が試みられました。25 回以上の反復が必要です。(収束基準 =.010)。抽出が終了しました。

回転後の因子行列^a

	因子		
	1	2	3
ストレス	.519	.130	-.213
健康行動	-.007	-.103	.375
健康習慣	-.165	-.183	.360
社会支援	-.243	-.015	.297
社会役割	-.203	.035	.415
健康度	.735	.094	-.099
生活環境	.168	.796	.032
医療機関	.020	.311	-.170

因子抽出法: 主因子法

回転法: Kaiser の正規化を伴うバリマックス法

- a. 6 回の反復で回転が収束しました。

回転

因子変換行列

因子	1	2	3
1	.740	.486	-.466
2	-.346	.868	.357
3	.577	-.103	.810

因子抽出法: 主因子法

回転法: Kaiser の正規化を伴うバリマックス法

出力を見るポイント

因子の解釈

因子行列^a

	因子		
	1	2	3
ストレス	.546	-.143	.113
健康行動	-.230	.047	.310
健康習慣	-.378	.026	.215
社会支援	-.325	.177	.101
社会役割	-.326	.248	.215
健康度	.635	-.208	.335
生活環境	.496	.644	.041
医療機関	.245	.202	-.158

因子抽出法: 主因子法

- a. 3 個の因子の抽出が試みられました。25 回以上の反復が必要です。(収束基準 = .010)。抽出が終了しました。

回転後の因子行列^a

	因子		
	1	2	3
ストレス	.519	.130	-.213
健康行動	-.007	-.103	.375
健康習慣	-.165	-.183	.360
社会支援	-.243	-.015	.297
社会役割	-.203	.035	.415
健康度	.735	.094	-.099
生活環境	.168	.796	.032
医療機関	.020	.311	-.170

因子抽出法: 主因子法

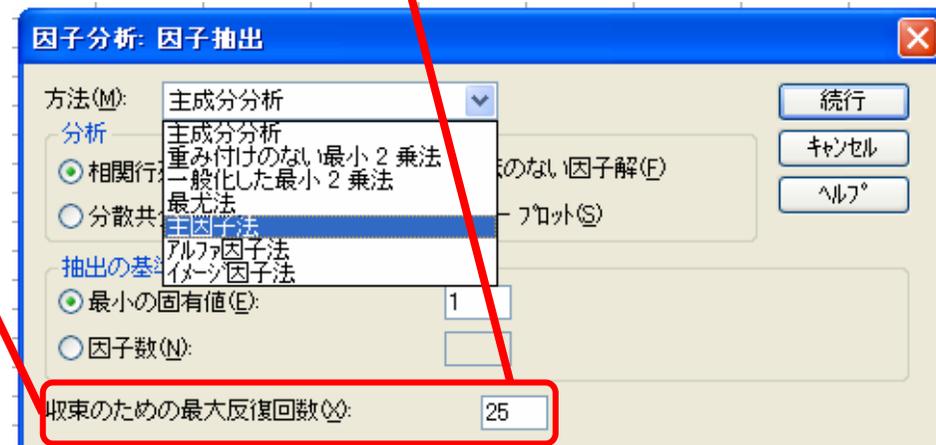
回転法: Kaiser の正規化を伴うバリマックス法

- a. 6 回の反復で回転が収束しました。

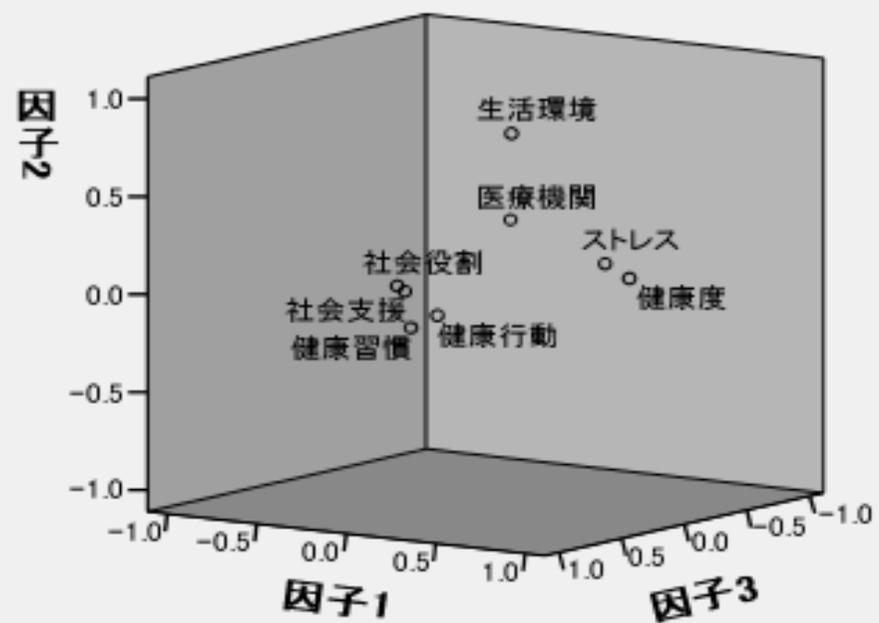
基準に達するまで反復計算

共通性

因子負荷量



回転後の因子空間の因子プロット



チェック

- 因子抽出法
 - 共通性が1を超えていないか
 - 収束のための最大反復回数
 - データ数は十分か
- 因子数の決定
 - 最小の固有値
 - スクリーンプロット
- 因子の解釈
 - 因子負荷量
 - 因子間の相関係数(斜交解の場合)
 - 回転法は適切か
- 項目の取捨選択
 - 共通性や因子負荷が低い項目
 - 複数の因子に関して因子負荷が高い項目

参考：因子得点

$$\text{因子得点} = \text{因子得点係数} \times \text{回答値}$$

因子得点係数行列

	因子		
	1	2	3
ストレス	.252	.012	-.096
健康行動	.059	-.032	.267
健康習慣	-.004	-.057	.243
社会支援	-.070	.030	.176
社会役割	-.035	.017	.286
健康度	.612	-.077	.069
生活環境	.008	.771	.126
医療機関	-.034	.106	-.111

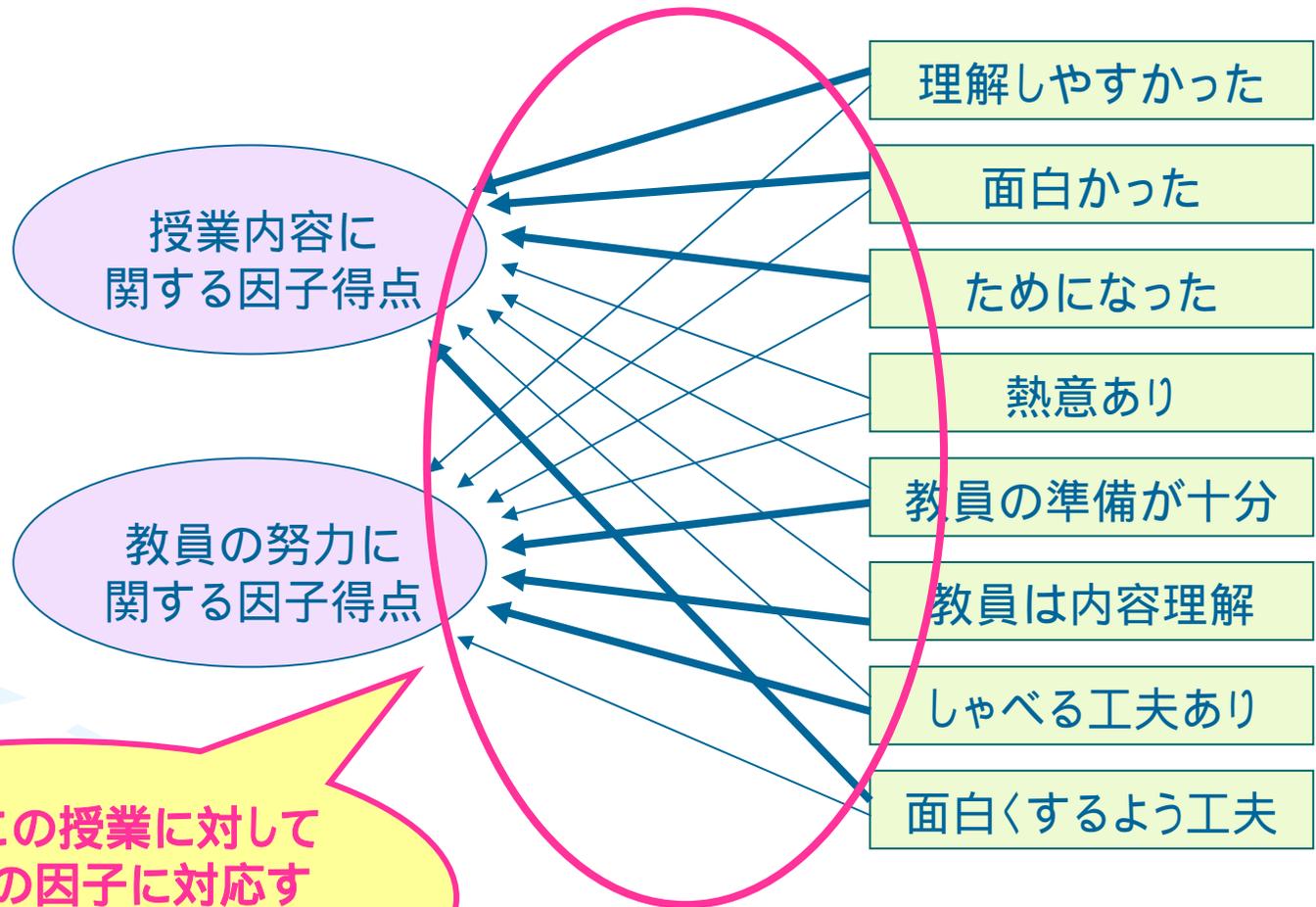
因子抽出法: 主因子法
回転法: Kaiser の正規化を伴うバリマックス法
因子得点の計算方法: 回帰法

因子得点共分散行列

因子	1	2	3
1	.606	.081	-.123
2	.081	.655	-.005
3	-.123	-.005	.395

因子抽出法: 主因子法
回転法: Kaiser の正規化を伴うバリマックス法
因子得点の計算方法: 回帰法

授業評価アンケートの例で言うと...



ある人がこの授業に対してそれぞれの因子に対応する項目について何点をつけたかがわかる

= 因子得点係数 × 回答値
(松尾・中村, 2002)



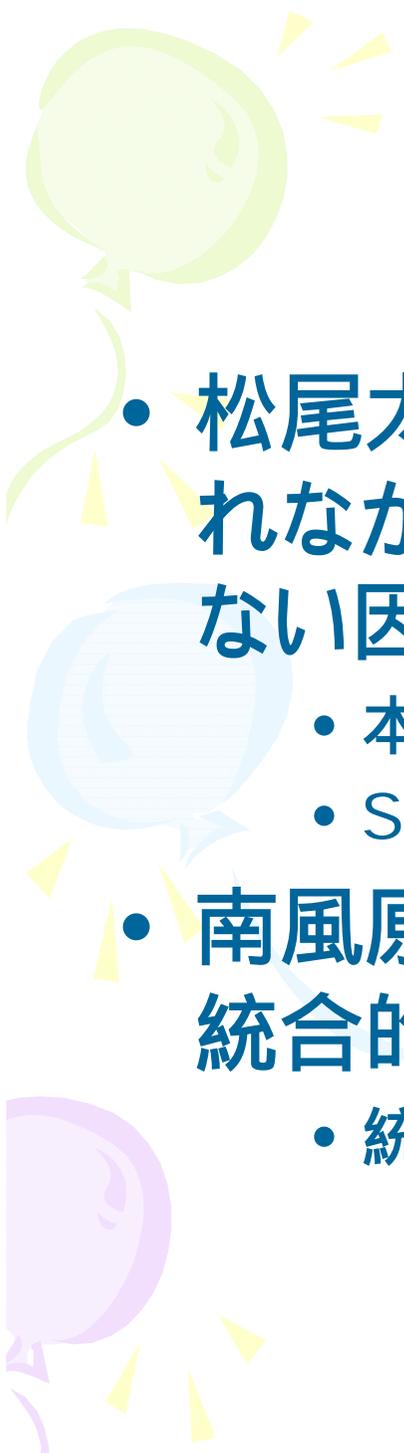
- 石村貞夫 1998 SPSSによる多変量データ解析の手順 東京図書

- さまざまな多変量解析の手順が掲載されている
- 分析自体についての説明はない



- 高橋信 2005 マンガでわかる統計学 因子分析編 オーム社

- アンケート作成の注意点がわかる
- 図表が多く、とっつきやすい
- 後半は、行列の知識がないと理解するのは難しい

- 
- 松尾太加志・中村知靖 2002 誰も教えてくれなかった因子分析 - 数式が絶対に出てこない因子分析入門 - 北大路書房

- 本当に数式はでてこない
- SPSSの手順や出力の読み取り方も掲載されている

- 南風原朝和 2002 心理統計学の基礎 - 統合的理解のために 有斐閣

- 統計のベーシックなテキスト