

# An Application Study of Factor Analysis in TCM Acupuncture Clinical Data Analysis\*

Jun Wu<sup>1</sup>, Qinfeng Huang<sup>2</sup>, Xiangyang Feng<sup>1</sup>, Houqin Su<sup>1</sup>

<sup>1</sup>College of Computer Science & Technology, Donghua University, Shanghai

<sup>2</sup>Shanghai Research Institute of Acupuncture and Meridian, Shanghai

Email: aqaya@126.com

Received: Nov. 7<sup>th</sup>, 2013; revised: Dec. 1<sup>st</sup>, 2013; accepted: Dec. 6<sup>th</sup>, 2013

Copyright © 2013 Jun Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In accordance of the Creative Commons Attribution License all Copyrights © 2013 are reserved for Hans and the owner of the intellectual property Jun Wu et al. All Copyright © 2013 are guarded by law and by Hans as a guardian.

**Abstract:** TCM acupuncture has a long history in China, but analysis about aspects of related degrees and contributed ratios among acupuncture points used to treat the same disease is seldom reported. To find if there are some correlations in clinical data, factor is applied in analysis of clinical data. Through actual analysis of the treatment cases for the disease of facioplegia recorded in a large number of the acupuncture literature materials, the influence degrees and potential relations in meanings of statistics among the different acupuncture points for treatment effects of the same disease have been exposed. The result shows the correlation in clinical data and arrival at a conclusion that factor analysis can be applied in analysis of clinical data.

**Keywords:** Factor Analysis; TCM (Traditional Chinese Medicine); Acupuncture; Clinical Data

## 因子分析在中医针灸临床数据分析中的应用研究\*

吴俊<sup>1</sup>, 黄琴峰<sup>2</sup>, 冯向阳<sup>1</sup>, 苏厚勤<sup>1</sup>

<sup>1</sup>东华大学计算机科学与技术学院, 上海

<sup>2</sup>上海市针灸经络研究所, 上海

Email: aqaya@126.com

收稿日期: 2013年11月7日; 修回日期: 2013年12月1日; 录用日期: 2013年12月6日

**摘要:** 中医针灸在中国历史悠久, 但针对治疗同种疾病所用穴位之间的关联程度和贡献率方面的分析鲜有报道。针对挖掘临床数据是否存在内在联系这一课题, 提出将因子分析应用在临床数据分析领域的设想, 通过对大量针灸文献资料记载的面神经麻痹疾病的治疗案例的实际分析了不同穴位对同一种疾病治疗效果的影响程度以及它们潜在的意义上的内在联系, 研究结果表明临床数据之间存在一定的相关性, 并得出因子分析可以应用于针灸临床数据分析这一结论。

**关键词:** 因子分析; 中医; 针灸; 临床数据

### 1. 引言

在 20 世纪 70 年代, 以中国向全世界公布针刺麻醉的研究成果为契机, 国际社会掀起了一股渴望了解

\*资助信息: 国家重点基础研究发展计划(973 项目)(No. 2009CB522900)。

针灸学和应用针灸治病的热潮。在针灸国际化的进程中, 世界卫生组织发挥了重要的推动和引导作用, 如在一些国家设立针灸研究培训合作中心、支持并创建世界针灸学联合会、制定《经络穴位名称国际标

准》《针灸临床研究规范》等。如今，在世界各国，已有 140 多个国家和地区开展针灸医疗，从事针灸的人数约 20~30 万人<sup>[1]</sup>。

在人体上，大约有 600 个穴位，每个穴位对于治疗特定的疾病都有可能有一定的疗效，而不同穴位的组合方法也可能影响治疗的效果。对针灸数据进行分类统计，以便找出不同穴位对针灸效果的影响程度和穴位之间针对某种疾病的内在联系，从而为临床医学人员在穴位选择的决策过程中提供量化依据。

因子分析是数据挖掘分析中的一种重要的统计方法，它可以用较少的综合指标分析治疗疾病的穴位，找出其中起主要作用的穴位。这样，为穴位的快速准确选择提供决策依据，保证针灸方案的可靠性。

为充分挖掘针灸临床数据间的联系，本文通过将因子分析应用于针灸临床数据分析中的方法，选取具有代表性的疾病数据进行分析研究，数值分析结果表明，针灸临床数据间存在相当的内在联系，将因子分析应用于这一领域的是可行的。

## 2. 概述

因子分析在 1904 年由心理学家 Charles Spearman 提出，主要用于解决相关测验的得分统计分析。目前，因子分析已被应用到许多科学领域，并且取得了一定的成果。

因子分析可以把大量的检测指标提炼为几个反映综合情况的指标。当待处理问题中含有较多变量时，由于变量较多，变量之间很有可能存在相关性，这样多变量反映的信息就会产生重叠。因子分析方法研究变量的相关系数矩阵或协方差矩阵，从多个变量中找出较少、同时又能尽可能地反映原变量信息综合变量。

经过分析，原始变量可以表示为一个线性组合，前半部分是由变量的几个公共因子构成，后半部分是影响变量的特殊因子<sup>[2]</sup>。

设  $X = (X_1, \dots, X_p)'$  是观测到的随机向量， $F = (F_1, \dots, F_m)'$  ( $m < p$ ) 为公共随机向量。设  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_p)$  与  $F$  不存在相关性，且  $E(\varepsilon) = 0$ ， $D(\varepsilon) = \text{diag}(\sigma_1^2, \dots, \sigma_p^2) = D$  为对角矩阵， $\sigma_i^2$  是特殊因子  $\varepsilon_i$  的方差，仅和变量  $X_i$  有关，即剩余方差，由此可得正交因子模型：

$$\begin{aligned} X_1 &= a_{11}F_1 + a_{12}F_2 + \dots + a_{1m}F_m + \varepsilon_1 \\ &\vdots \\ X_p &= a_{p1}F_1 + a_{p2}F_2 + \dots + a_{pm}F_m + \varepsilon_p \end{aligned}$$

该模型的矩阵式为： $x = AF + \varepsilon$ ，其中： $A = (a_{ij})_{p \times m}$  系数矩阵(因子载荷矩阵)， $a_{ij}$  是变量  $i$  在第因子  $j$  上的载荷，反映了第  $i$  个变量对第  $j$  个因子上的贡献度<sup>[3]</sup>，特殊因子  $\varepsilon_i$  独立于其他变量，之和相应的变量  $X_i$  相关。

应用因子分析解决实际问题：主要需要正确选择因子，并进行合理解释<sup>[4]</sup>。

因此，围绕上述两点，得出以下四个步骤：

- 1) 检验因子分析是否适用原始变量的分析；
- 2) 公共因子的选择；
- 3) 旋转因子载荷矩阵，更好地解释公共因子；
- 4) 计算因子得分，为后续解释和分析提供数据。

图 1 给出从针灸信息系统数据库中获得针灸治疗临床数据样本信息对应因子分析的处理流程。

如图 1 所示，首先将数据从数据库中取出，根据上述分析进行分组和计算，并选择临床使用次数较多，更具有统计研究意义的变量研究。再根据分析结果选择因子提取的方法进行因子分析，如需更好地解释因子的实际意义，可以进一步进行因子旋转。最后，决定是否计算因子得分，根据最终分析结论得到反馈结果。

## 3. 分析应用

### 3.1. 测评样本数据

本文分析的针灸临床治疗样本数据选自自己建《中医针灸信息库》<sup>[5]</sup>中有关疾病“面神经麻痹”(数据库中对疾病编码 0511)的针灸治疗数据集。该数据库收集和整理了 74 种 1705 册中医和中西医结合期刊中 6 万多篇文献登载的中医针灸临床治疗案例数据。在搜集、整理和创建《中医针灸信息库》的过程中曾得到国家重点基础研究发展计划(973 项目, 2009CB522900)的资助。在对选择的数据集进行初步筛选之后，选择其中出现频率最高的 32 个穴位作为变量进行分析，它们是：GB34(阳陵泉)、GB30(环跳)、EX-HN4(鱼腰)、EX-HN5(太阳)、BL57(承山)、PC6(内关)、W21(牵正)、W23(阿是穴)、BL40(委中)、GV26(水沟)、GB20(风池)、BL2(攒竹)、CV24(承浆)、SI18(颧髻)、GV20(百会)、GB14(阳白)、LR3(太冲)、TE5(外关)、LI11(曲池)、

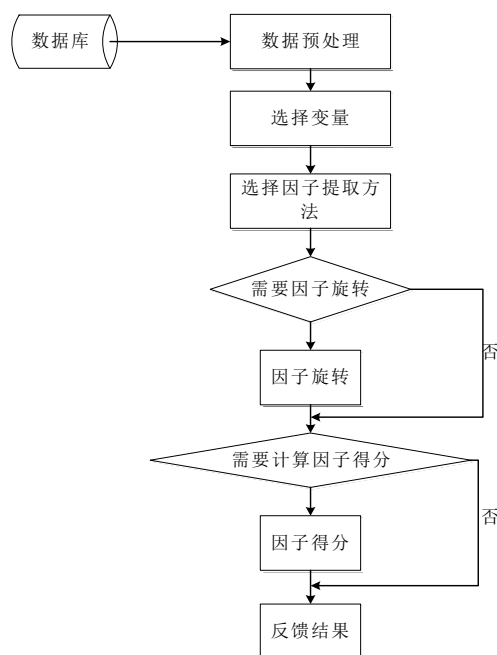


Figure 1. Factor analysis flow diagram of clinical data  
图 1. 临床数据因子分析处理流程示意图

TE23(丝竹空)、BL60(昆仑)、LI20(迎香)、BL23(肾俞)、LI4(合谷)、TE17(翳风)、GB39(悬钟)、ST2(四百)、SP6(三阴交)、ST4(地仓)、ST36(足三里)、ST6(颊车)、ST7(下关)(下文叙述使用穴位代码)。由于原样本中穴位变量取值为离散值(使用标记为 0, 未使用标记为 1), 且穴位过多, 不易于筛选掉极少用到的穴位(这些穴位仅在极少数案例中出现, 在统计意义上无研究意义), 因而将大量数据划分为 60 组, 这样, 每组样本的变量取值为连续值, 适用因子分析方法, 且易于过滤掉极少使用的穴位。以组为单位作为样本研究。通过过滤和计算各组内穴位使用情况的累计值获得的穴位使用次数, 如表 1 所列。

### 3.2. 相关系数矩阵

样本数据的相关系数矩阵是因子分析的基础。根据相关系数计算公式:

$$r_{XY} = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2}}$$

对应表 1 所列的 32 个评估变量的相关系数如表 2 所列。

由表 2 可得所有变量值两两之间的相关程度, 如

Table 1. Sixty groups acupuncture treatment data of Facial Paralysis

表 1. “面神经麻痹” 60 组样本的针灸治疗穴位数据

穴位 样本 组号	GB34	GB30	EXHN4	...	ST36	ST6	ST7
1	18	19	7	...	24	25	23
2	21	18	18	...	28	43	27
3	17	19	9	...	38	33	23
...	...	...	...	...	...	...	...
59	17	9	19	...	45	43	35
60	18	7	9	...	27	31	17

Table 2. Correlation matrix

表 2. 相关系数矩阵

	GB34	GB30	...	ST36	ST6	ST7
GB34	1.000	0.742	...	0.577	0.328	-0.644
GB30	0.742	1.000	...	0.741	0.071	-0.523
EX-HN4	-0.371	-0.444	...	-0.236	-0.209	0.418
...	...	...	...	...	...	...
ST6	-0.591	-0.618	...	0.335	1.000	0.847
ST7	-0.541	-0.552	...	0.286	0.847	1.000

GB34 与 GB30、ST6 与 ST7 等高度相关, 因此在统计意义上表 1 中的指标数据具有一定的相关性。

### 3.3. Bartlett 和 KMO 检验值

KMO 检验统计量是用于比较变量间偏相关系数和简单相关系数的指标, KMO 统计量的检验值反映因子分析对原始变量的分析效果, 越接近 1, 效果越好。计算公式为:

$$KMO = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} a_{ij}^2}$$

其中  $r_{ij}$  表示简单相关系数,  $a_{ij-1,2,3,\dots,k}^2$  为偏相关系数。显然, 当  $a_{ij-1,2,3,\dots,k}^2 \approx 0$  时,  $KMO \approx 1$ ; 当  $a_{ij-1,2,3,\dots,k}^2 \approx 1$  时, KMO 的取值介于 0 和 1 之间<sup>[6]</sup>, KMO 的度量标准如表 3 所列。

Bartlett 检验以多元正态总体协方差矩阵检验理论为基础。后者主要检验: 1) 总体协方差矩阵的元素是否均为已知协方差矩阵元素的  $\sigma^2$  倍, 2) 总体协方差矩阵  $\Sigma$  与已知矩阵  $\Sigma_0$  的相等性, 3) 多总体协方差矩

**Table 3. Standard of KMO**  
**表 3. KMO 度量标准**

KMO 值	分析的适用性
0.90~1.00	非常好
0.80~0.89	好
0.70~0.79	一般
0.60~0.69	差
0.50~0.59	很差
<0.5	不适用

阵的相等性检验。其中，对第一点检测的原假设是  $H_0: \Sigma = \sigma^2 \Sigma_0$ ,  $H_1: \Sigma \neq \sigma^2 \Sigma_0$  ( $\Sigma$ : 总体协方差矩阵),  $\Sigma_0 > 0$  已知,  $\sigma^2$  未知, 此时检验的似然比统计量为:

$$\lambda = \frac{|\sum_0^{-1} A|^{n/2}}{[tr(\sum_0^{-1} A)/p]^{pn/2}}, A = \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})'$$

$$\text{或等价于 } W = (\lambda)^{2/n} = \frac{p^p |\sum_0^{-1} A|}{[tr(\sum_0^{-1} A)]^p}$$

应用中, 当  $n$  很大时, 可利用下列近似公式进行  $\chi^2$  检验:

$$[(n-1) - (2p^2 + p + 2)] \ln W = \chi_{(p+2)(p-1)/2}^2$$

基于以上检验, 令  $\sum_0 = I_p$  即检验相关矩阵是单位矩阵 ( $\rho = I$ )。此时原假设  $H_0: \Sigma = \sigma^2 I$ ,  $H_1: \Sigma \neq \sigma^2 I$ , 这时, 检验统计量为:

$$\Lambda = \frac{|S|^{n/2}}{[tr(S)/p]^{np/2}} = \left[ \frac{\prod_{i=1}^p \hat{\lambda}_i}{\left(\frac{1}{p} \sum_{i=1}^p \hat{\lambda}_i\right)^p} \right]^{n/2} = \left[ \frac{\text{几何均值 } \hat{\lambda}_i}{\text{算数均值 } \hat{\lambda}_i} \right]^{np/2}$$

其中,  $S$  是样本数据的协方差矩阵,  $S = A/(n-1)$ ,  $p$  是变量的个数,  $n$  是样本的容量,  $\lambda_i$  则是  $S$  的特征值。

若假设成立, 样本容量够大, Bartlett 认为  $-2[1 - (p^2 + p + 2)/6pn] \ln \Lambda$  近似服从  $\chi_{(p+2)(p-1)/2}^2$ , 因此, 这个样本的  $\alpha$  临界值为  $\chi_{(p+2)(p-1)/2}^2(\alpha)$ 。

实际检验中, 可根据检验统计量公式计算出概率值  $P$ , 若小于 0.05, 拒绝原假设, 即因子分析适用原

始数据的分析, 反之则不适用。因为若  $\Sigma = \sigma^2 I$ , 则常密度轮廓线为球面, 因此称该检验为 Bartlett 球形检验。

另外, 还采用如下近似公式计算 Bartlett 球形检验统计量:

$$\chi^2 = \frac{(11+2p)-6n}{6} \ln |R|$$

此处自由度是  $p(p-1)/2$ ,  $p$  为变量数,  $n$  是样本容量,  $|R|$  是相关系数矩阵。在应用中, 有时为配合 Bartlett 球形检验, 还需计算  $\ln|R|$  的值, 并根据其值大小判断因子分析的适用性。若因子分析使用协方差矩阵, 则  $\ln|R|$  就取该矩阵的行列式的自然对数值<sup>[7]</sup>。

如表 2 所示相关系数和相应偏相关系数统计的 Bartlett 球形和 KMO 检验值<sup>[8]</sup>。KMO 检验值接近 1 (为 0.838), 由此推断各变量之间有相关性, 适用因子分析; Bartlett's 球形检验值较大 (为 496), 也佐证了这一推断。

### 3.4. 选取因子

当确定原始数据适用因子分析后, 需要计算相关系数矩阵的特征值, 由于该矩阵是实对称矩阵, 因而采用雅克比方法计算其特征值, 计算步骤归纳如下:

令  $A = A_0$ , 依次构造矩阵序列  $A_k$ , 使  $A_k = R_k^T(p, q) A_{k-1} R_k(p, q)$  ( $k = 1, 2, \dots$ ), 其中  $R(p, q)$  是在  $(p, q)$  平面上转过角度  $\theta$  的一个旋转。其中:

$$\begin{aligned} R_k(i, i) &= 1, i \neq p, q & R_k(i, i) &= \cos \theta, i = p, q \\ R_k(i, j) &= -R_k(j, i) = \sin \theta, i = p, j = q \\ R(i, j) &= 0, i, j = \text{其他}。 \end{aligned}$$

选择平面和确定旋转角度  $\theta$  的方法比较简单: 寻找  $A_s$  主对角线上方元素确定最大模的项  $a_{pq}^{(s)}$ , 选择  $\theta$ , 使  $a_{pq}^{(s+1)} = 0$ , 由于平面旋转  $R(p, q)$  进行相似变换后仅影响第  $p, q$  行和列的元素。  $P, q, \theta$  的确定原则如下:

记  $A_{k-1} = (a_{ij}^{(k-1)})$ ,  $A_k = (a_{ij}^{(k)})$ , 则:

$$\begin{aligned} a_{ij}^{(k)} &= a_{ij}^{(k-1)} (i, j \neq p, q), \\ a_{pi}^{(k)} &= a_{ip}^{(k)} = a_{ip}^{(k-1)} \cos \theta - a_{iq}^{(k-1)} \sin \theta (i \neq p, q), \\ a_{ip}^{(k)} &= a_{ip}^{(k-1)} \sin \theta - a_{iq}^{(k-1)} \cos \theta (i \neq p, q), \end{aligned}$$

$$\begin{aligned}
 a_{pp}^{(k)} &= a_{pp}^{(k-1)} \cos^2 \theta - 2a_{pp}^{(k-1)} \cos \theta \sin \theta + a_{pp}^{(k-1)} \sin^2 \theta, \\
 a_{pp}^{(k)} &= a_{pp}^{(k-1)} \sin^2 \theta - 2a_{pp}^{(k-1)} \cos \theta \sin \theta + a_{pp}^{(k-1)} \cos^2 \theta, \\
 a_{pp}^{(k)} &= a_{qp}^{(k)} = \left( a_{pp}^{(k-1)} - a_{qq}^{(k-1)} \right) \cos \theta \sin \theta \\
 &\quad + a_{pp}^{(k-1)} \left( \cos^2 \theta - \sin^2 \theta \right) \\
 &= \frac{1}{2} \left( a_{pp}^{(k-1)} - a_{qq}^{(k-1)} \right) \sin 2\theta + a_{qq}^{(k-1)} \cos 2\theta
 \end{aligned}$$

为使  $a_{pq}^{(k)} = 0$  (即消去  $(p,q)$  位的元素), 必须选择这样的旋转角:

1)  $\operatorname{tg} 2\theta = \frac{-2a_{qp}^{(k-1)}}{a_{pp}^{(k-1)} - a_{qq}^{(k-1)}}$ ; 2)  $\theta$  满足:

$-\frac{\pi}{4} \leq \theta \leq \frac{\pi}{4}$ ; 此外, 若  $a_{pp}^{(k-1)} - a_{qq}^{(k-1)} = 0$  那么选择

$\theta = a_{qp}^{(k-1)} a_{pq}^{(k-1)} = \frac{\pi}{4}$ 。注意,  $a_{pq}^{(k-1)} > 0$  时,  $\theta = -\frac{\pi}{4}$ ,

$a_{pq}^{(k-1)} < 0$ 。这样确定的旋转角所对应的旋转矩阵就能使  $A_k = R_k^T A_{k-1} R_k$  中的  $a_{pq}^{(k)} = 0$ , 然而即使  $a_{pj}^{(k-1)} = 0$  或  $a_{qj}^{(k-1)} = 0$ , 由于旋转变换后它们通常不为 0, 所以旋转变换消除实对称矩阵的非对角元素法的过程不是有

限的, 而是需要无限迭代的, 但只需进行到所有非对角元素小于预先设定的精度即可。由于经每一次正交相似变换后矩阵的非对角元素平方和  $E(A) = \sum_{i \neq j} a_{ij}^2$

都会减少, 因而雅克比方法是收敛的。

通过特征值得到贡献率和累积贡献率, 从而开源确定因子的个数。针对表 2 所示 32 个变量指标值的相关系数矩阵, 经计算后的贡献率、特征值以及累计贡献率如表 4 所列。

按照选取因子的规则, 通常只需考虑选取大于 1 的特征值作为因子, 并结合特征值对应的累计贡献率决定被选取因子的个数。通常被选取因子的累计贡献率需达到 80%, 表 5 列出了适合作为因子的 7 个特征值(序号 1 到 7), 其累计贡献率达到了 80.324%, 依次称为因子 1、因子 2, ..., 因子 7, 它们反映了表 1 中所有测评样本数据信息量的 80.324%。

### 3.5. Promax 旋转

为了更好地解释因子的实际意义, 对因子载荷矩

**Table 4. Eigenvalues calculated contribution rate and contribution rate computed from correlation matrix**  
**表 4. 由表 2 相关系数矩阵计算获得的特征值与贡献率**

特征值序	特征值	贡献率%	累积贡献率%	特征值序	特征值	贡献率%	累积贡献率%
1	15.663	48.948	48.948	17	0.248	0.775	95.144
2	3.451	10.785	59.733	18	0.222	0.694	95.838
3	2.393	7.479	67.212	19	0.216	0.674	96.512
4	1.382	4.318	71.53	20	0.189	0.591	97.103
5	1.039	3.247	74.776	21	0.164	0.513	97.616
6	0.936	2.925	77.701	22	0.153	0.478	98.094
7	0.839	2.623	80.324	23	0.122	0.381	98.475
8	0.804	2.511	82.836	24	0.108	0.336	98.811
9	0.636	1.988	84.823	25	0.088	0.275	99.086
10	0.618	1.931	86.754	26	0.074	0.23	99.316
11	0.537	1.679	88.433	27	0.063	0.197	99.513
12	0.461	1.441	89.874	28	0.044	0.138	99.651
13	0.407	1.273	91.147	29	0.038	0.12	99.771
14	0.372	1.164	92.31	30	0.035	0.111	99.881
15	0.355	1.109	93.419	31	0.025	0.079	99.96
16	0.304	0.951	94.37	32	0.013	0.04	100

**Table 5. Rotation component matrix of selected seven factors**  
**表 5. 所选取 7 个因子的旋转成分矩阵**

穴位	因子						
	1	2	3	4	5	6	7
GB34	-0.406	-0.74	0.285	0.099	-0.074	-0.166	0.118
GB30	-0.347	-0.859	-0.079	-0.004	-0.125	-0.009	-0.051
EX-HN4	0.792	0.14	-0.108	0.131	0.078	-0.158	0.193
EX-HN5	0.813	0.216	-0.046	-0.059	0.109	0.145	0.06
BL57	-0.234	-0.772	-0.273	-0.159	0.073	-0.062	0.05
PC6	-0.391	-0.046	0.407	0.416	0.325	-0.117	-0.121
W21	0.651	0.41	-0.082	-0.176	0.134	0.209	-0.204
W23	-0.142	-0.092	-0.019	0.051	-0.9	-0.054	-0.031
BL40	-0.268	-0.857	-0.018	-0.005	-0.06	-0.151	-0.178
GV26	0.717	0.143	-0.019	-0.098	0.137	0.452	-0.059
GB20	0.562	0.245	0.453	0.014	0.133	0.176	0.354
BL2	0.815	0.21	0.067	0.179	0.158	0.151	-0.128
CV24	0.721	0.172	0.044	0.006	0.082	0.514	0.032
SI18	0.708	0.427	0.166	0.224	0.107	-0.038	-0.093
GV20	-0.131	0.104	0.185	0.873	-0.063	0.101	0.089
GB14	0.871	0.323	0.024	-0.075	0.097	0.068	-0.009
LR3	0.395	0.39	0.284	0.16	0.175	0.143	0.589
TE5	0.147	0.023	0.872	0.011	-0.104	-0.119	-0.13
LI11	-0.119	-0.092	0.768	0.155	-0.016	0.148	0.1
TE23	0.626	0.119	-0.108	0.353	0.42	-0.076	0.139
BL60	-0.33	-0.812	-0.094	-0.049	0.048	-0.074	-0.11
LI20	0.878	0.254	-0.018	-0.057	0	0.154	-0.064
BL23	-0.352	-0.57	0.069	0.081	-0.107	-0.023	0.542
LI4	0.767	0.395	0.351	0.026	0.018	0.149	0.173
TE17	0.812	0.344	0.118	-0.13	0.057	0.123	-0.065
GB39	-0.41	-0.756	0.022	-0.088	-0.105	0.185	-0.003
ST2	0.852	0.281	-0.008	-0.051	-0.151	-0.051	0.048
SP6	0.035	0.327	0.674	0.072	0.188	0.152	0.312
ST4	0.894	0.336	0.015	-0.121	0.076	0.02	0.041
ST36	0.322	0.104	0.356	0.332	-0.002	0.566	0.166
ST6	0.892	0.323	0.069	-0.132	0.026	0.054	0.047
ST7	0.809	0.317	0.034	-0.113	-0.184	0.021	0.155

阵(因子载荷矩阵的值为上文正交因子模型中的系数  $a_{ij}$ , 称为第  $i$  个变量在第  $j$  个因子上的载荷, 反映了第

$i$  个变量在第  $j$  个因子上的重要性)进行 Promax 旋转。Promax 旋转使得各因子载荷发生两极分化, 使得各因

子间不再相互独立，而彼此相关，从而更清晰地反映原变量对因子的潜在关联性。对表 4 选取的 7 个因子的载荷系数矩阵经 Promax 旋转后的矩阵系数如表 5 所列。

从表 4 可以更加清晰地看出，因子 1 在 EX-HN4(鱼腰)、EX-HN5(太阳)、W21(牵正)、GV26(水沟)等穴位上有较高的载荷，由表 3 知道主因子 1 占整个疗效贡献率的 48.948%，是治疗“面神经麻痹”的重要因素。以此类推，其它因子的重要程度逐步降低，这一点通过分析结果的碎石图可以直观地看出，碎石图如图 2 所示。

#### 4. 结语

本文把因子分析的理论和方法用于中医针灸治疗的穴位分析。在穴位统计贡献率的意义上给出治疗“面神经麻痹”的 32 个主要穴位间的潜在联系。通过分析，清晰地揭示了各穴位对“面神经麻痹”治疗的影响程度及其潜在的统计意义上的联系，能够为临床针灸治疗“面神经麻痹”疾病的穴位选择提供量化的辅助决策依据。

本文研究和应用的方法同样适用于其它不同疾病针灸临床治疗数据的分析。通过对已有针灸临床治疗数据的预处理，使用因子分析进行分析可以完整地揭示不同穴位针对同种疾病针灸治疗的影响程度及其统计意义上的内在联系，能够为针灸医学提供极有

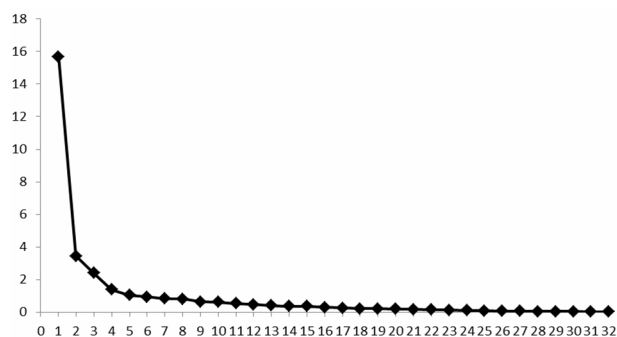


Figure 2. Scree plot of all factors  
图 2. 所有因子的碎石图

意义的分析依据。

#### 参考文献 (References)

- [1] 林海明 (2009) 因子分析模型的改进与应用. *数理统计与管理*, **6**, 998-1012.
- [2] 苏厚勤, 吴磊, 冯娟, 成长生 (2008) 因子分析在 AFC 设备故障事件分析中的应用. *东华大学学报*, **5**, 619-623.
- [3] 高惠璇 (2005) 应用多元统计分析. 北京大学出版社, 北京.
- [4] 刘菊香, 沈霄风 (2006) 用 SPSS 统计软件对学生综合成绩的因子分析. *统计教育*, **1**, 53-56.
- [5] 黄琴峰, 苏厚勤 (2012) 《上海针灸杂志》创建中医针灸信息库之路. *编辑学报*, **增刊 2**, 66.
- [6] 傅德印 (2007) 因子分析统计检验体系的探讨. *统计研究*, **6**, 86-90.
- [7] 傅德印 (2007) 主成分分析中的统计检验问题[J]. *统计教育*, **9**, 4-7.
- [8] BRAND (2013) Exploratory factor analysis. <http://www.stat.sc.edu/habing/courses/530EFA.pdf>