

一、概述

在处理信息时，当两个变量之间有一定相关关系时，可以解释为这两个变量反映此课题的信息有一定的重叠，例如，高校科研状况评价中的立项课题数与项目经费、经费支出等之间会存在较高的相关性；学生综合评价研究中的专业基础课成绩与专业课成绩、获奖学金次数等之间也会存在较高的相关性。而变量之间信息的高度重叠和高度相关会给统计方法的应用带来许多障碍。

为了解决这些问题，最简单和最直接的解决方案是削减变量的个数，但这必然又会导致信息丢失和信息不完整等问题的产生。为此，人们希望探索一种更为有效的解决方法，它既能大大减少参与数据建模的变量个数，同时也不会造成信息的大量丢失。主成分分析正是这样一种能够有效降低变量维数，并已得到广泛应用的分析方法。

主成分分析以最少的信息丢失为前提，将众多的原有变量综合成较少几个综合指标，通常综合指标（主成分）有以下几个特点：

- ★主成分个数远远少于原有变量的个数

原有变量综合成少数几个因子之后，因子将可以替代原有变量参与数据建模，这将大大减少分析过程中的计算工作量。

- ★主成分能够反映原有变量的绝大部分信息

因子并不是原有变量的简单取舍，而是原有变量重组后的结果，因此不会造成原有变量信息的大量丢失，并能够代表原有变量的绝大部分信息。

- ★主成分之间应该互不相关

通过主成分分析得出的新的综合指标（主成分）之间互不相关，因子参与数据建模能够有效地解决变量信息重叠、多重共线性等给分析应用带来的诸多问题。

- ★主成分具有命名解释性

总之，主成分分析法是研究如何以最少的信息丢失将众多原有变量浓缩成少数几个因子，如何使因子具有一定的命名解释性的多元统计分析方法。

二、基本原理

主成分分析是数学上对数据降维的一种方法。其基本思想是设法将原来众多的具有一定相关性的指标 X_1, X_2, \dots, X_p （比如 p 个指标），重新组合成一组较少个数的互不相关的综合指标 F_m 来代替原来指标。那么综合指标应该如何去提取，使其既能最大程度的反映原变量 X_p 所代表的信息，又能保证新指标之间保持相互无关（信息不重叠）。

设 F_1 表示原变量的第一个线性组合所形成的主成分指标，即 $F_1 = a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p$ ，由数学知识可知，每一个主成分所提取的信息量可用其方差来度量，其方差 $\text{Var}(F_1)$ 越大，表示 F_1 包含的信息越多。常常希望第一主成分 F_1 所含的信息量最大，因此在所有的线性组合中选取的 F_1 应该是 X_1, X_2, \dots, X_p 的所有线性组合中方差最大的，故称 F_1 为第一主成分。如果第一主成分不足以代表原来 p 个指标的信息，再考虑选取第二个主成分指标 F_2 ，为有效地反映原信息， F_1 已有的信息就不需要再出现在 F_2 中，即 F_2 与 F_1 要保持独立、不相关，用数学语言表达就是其协方差 $\text{Cov}(F_1, F_2) = 0$ ，所以 F_2 是与 F_1 不

相关的 X_1, X_2, \dots, X_p 的所有线性组合中方差最大的，故称 F_2 为第二主成分，依此类推构造出的 F_1, F_2, \dots, F_m 为原变量指标 X_1, X_2, \dots, X_p 第一、第二、……、第 m 个主成分。

$$\begin{cases} F_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \\ F_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p \\ \dots\dots\dots \\ F_m = a_{m1}X_1 + a_{m2}X_2 + \dots + a_{mp}X_p \end{cases}$$

根据以上分析得知：

(1) F_i 与 F_j 互不相关，即 $\text{Cov}(F_i, F_j) = 0$ ，并有 $\text{Var}(F_i) = a_i' \Sigma a_i$ ，其中 Σ 为 X 的协方差阵

(2) F_1 是 X_1, X_2, \dots, X_p 的一切线性组合（系数满足上述要求）中方差最大的，……，即 F_m 是与 F_1, F_2, \dots, F_{m-1} 都不相关的 X_1, X_2, \dots, X_p 的所有线性组合中方差最大者。

$F_1, F_2, \dots, F_m (m \leq p)$ 为构造的新变量指标，即原变量指标的第一、第二、……、第 m 个主成分。

由以上分析可见，主成分分析法的主要任务有两点：

(1) 确定各主成分 $F_i (i=1, 2, \dots, m)$ 关于原变量 $X_j (j=1, 2, \dots, p)$ 的表达式，即系数 $a_{ij} (i=1, 2, \dots, m; j=1, 2, \dots, p)$ 。从数学上可以证明，原变量协方差矩阵的特征根是主成分的方差，所以前 m 个较大特征根就代表前 m 个较大的主成分方差值；原变量协方差矩阵前 m 个较大的特征值 λ_i （这样选取才能保证主成分的方差依次最大）所对应的特征向量就是相应主成分 F_i 表达式的系数 a_i ，为了加以限制，系数 a_i 启用的是 λ_i 对应的单位化的特征向量，即有 $a_i' a_i = 1$ 。

(2) 计算主成分载荷，主成分载荷是反映主成分 F_i 与原变量 X_j 之间的相互关联程度：

$$P(Z_k, x_i) = \sqrt{\lambda_k} a_{ki} (i=1, 2, \dots, p; k=1, 2, \dots, m)$$

三、主成分分析法的计算步骤

主成分分析的具体步骤如下：

(1) 计算协方差矩阵

计算样品数据的协方差矩阵： $\Sigma = (s_{ij})_{p \times p}$ ，其中

$$s_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j) \quad i, j=1, 2, \dots, p$$

(2) 求出 Σ 的特征值 λ_i 及相应的正交化单位特征向量 a_i

Σ 的前 m 个较大的特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m > 0$, 就是前 m 个主成分对应的方差, λ_i 对应的单位特征向量 a_i 就是主成分 F_i 的关于原变量的系数, 则原变量的第 i 个主成分 F_i 为:

$$F_i = a_i' X$$

主成分的方差 (信息) 贡献率用来反映信息量的大小, α_i 为:

$$\alpha_i = \lambda_i / \sum_{i=1}^m \lambda_i$$

(3) 选择主成分

最终要选择几个主成分, 即 F_1, F_2, \dots, F_m 中 m 的确定是通过方差 (信息) 累计贡献率 $G(m)$ 来确定

$$G(m) = \sum_{i=1}^m \lambda_i / \sum_{k=1}^p \lambda_k$$

当累积贡献率大于 85% 时, 就认为能足够反映原来变量的信息了, 对应的 m 就是抽取的前 m 个主成分。

(4) 计算主成分载荷

主成分载荷是反映主成分 F_i 与原变量 X_j 之间的相互关联程度, 原来变量 X_j ($j=1, 2, \dots, p$) 在诸主成分 F_i ($i=1, 2, \dots, m$) 上的荷载 l_{ij} ($i=1, 2, \dots, m; j=1, 2, \dots, p$)。

$$l(Z_i, X_j) = \sqrt{\lambda_i} a_{ij} \quad (i=1, 2, \dots, m; j=1, 2, \dots, p)$$

在 SPSS 软件中主成分分析后的分析结果中, “成分矩阵”反应的就是主成分载荷矩阵。

(5) 计算主成分得分

计算样品在 m 个主成分上的得分:

$$F_i = a_{1i} X_1 + a_{2i} X_2 + \dots + a_{pi} X_p \quad i = 1, 2, \dots, m$$

实际应用时, 指标的量纲往往不同, 所以在主成分计算之前应先消除量纲的影响。消除数据的量纲有很多方法, 常用方法是原始数据标准化, 即做如下数据变换:

$$x_{ij}^* = \frac{x_{ij} - \bar{x}_j}{s_j} \quad i=1, 2, \dots, n; j=1, 2, \dots, p$$

$$\text{其中: } \bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}, \quad s_j^2 = \frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2$$

根据数学公式知道, ①任何随机变量对其作标准化变换后, 其协方差与其相关系数是一回事, 即标准化后的变量协方差矩阵就是其相关系数矩阵。②另一方

面,根据协方差的公式可以推得标准化后的协方差就是原变量的相关系数,亦即,**标准化后的变量的协方差矩阵就是原变量的相关系数矩阵**。也就是说,在**标准化前后变量的相关系数矩阵不变化**。

根据以上论述,为消除量纲的影响,将变量标准化后再计算其协方差矩阵,就是直接计算原变量的相关系数矩阵,所以主成分分析的实际常用计算步骤是:
☆计算相关系数矩阵

☆求出相关系数矩阵的特征值 λ_i 及相应的正交化单位特征向量 a_i

☆选择主成分

☆计算主成分得分

总结:原指标相关系数矩阵相应的特征值 λ_i 为主成分方差的贡献,方差的贡献率为 $\alpha_i = \lambda_i / \sum_{i=1}^p \lambda_i$, α_i 越大,说明相应的主成分反映综合信息的能力越强,可根据 λ_i 的大小来提取主成分。每一个主成分的组合系数(原变量在该主成分上的载荷) a_i 就是相应特征值 λ_i 所对应的单位特征向量。

主成分分析法的计算步骤

1、原始指标数据的**标准化**采集 p 维随机向量 $x = (x_1, x_2, \dots, x_p)^T$ n 个样品 $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$, $i=1, 2, \dots, n$,

$n > p$, 构造样本阵, 对样本阵元进行如下标准化变换:

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, i = 1, 2, \dots, n; j = 1, 2, \dots, p$$

$$\bar{x}_j = \frac{\sum_{i=1}^n x_{ij}}{n}, s_j^2 = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n-1}, \text{得标准化阵 } Z.$$

2、对标准化阵 Z 求**相关系数**矩阵

$$R = [r_{ij}]_{p \times p} = \frac{Z^T Z}{n-1}$$

$$r_{ij} = \frac{\sum z_{kj} \cdot z_{kj}}{n-1}, i, j = 1, 2, \dots, p$$

其中,

3、解样本相关矩阵 R 的特征方程 $|R - \lambda I_p| = 0$ 得 p 个特征根,确定主成分

$$\frac{\sum_{j=1}^m \lambda_j}{\sum_{j=1}^p \lambda_j} \geq 0.85$$

按 确定 m 值, 使信息的利用率达 85% 以上, 对每个 $\lambda_j, j=1, 2, \dots, m$, 解方程组 $Rb = \lambda_j b$ 得单位特征向量 b_j^o 。

4、将标准化后的指标变量转换为主成分

$$U_{ij} = z_i^T b_j^o, j = 1, 2, \dots, m$$

U_1 称为第一主成分, U_2 称为第二主成分, ..., U_p 称为第 p 主成分。

5、对 m 个主成分进行综合评价

对 m 个主成分进行加权求和, 即得最终评价值, 权数为每个主成分的方差贡献率。

一、主成分分析基本原理

概念: 主成分分析是把原来多个变量划为少数几个综合指标的一种统计分析方法。从数学角度来看, 这是一种降维处理技术。

思路: 一个研究对象, 往往是多要素的复杂系统。变量太多无疑会增加分析问题的难度和复杂性, 利用原变量之间的相关关系, 用较少的新变量代替原来较多的变量, 并使这些少数变量尽可能多的保留原来较多的变量所反应的信息, 这样问题就简单化了。

原理: 假定有 n 个样本, 每个样本共有 p 个变量, 构成一个 $n \times p$ 阶的数据矩阵,

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

记原变量指标为 x_1, x_2, \dots, x_p ，设它们降维处理后的综合指标，即新变量为 $z_1, z_2, z_3, \dots, z_m$ ($m \leq p$)，则

$$\begin{cases} z_1 = l_{11}x_1 + l_{12}x_2 + \dots + l_{1p}x_p \\ z_2 = l_{21}x_1 + l_{22}x_2 + \dots + l_{2p}x_p \\ \dots\dots\dots \\ z_m = l_{m1}x_1 + l_{m2}x_2 + \dots + l_{mp}x_p \end{cases}$$

系数 l_{ij} 的确定原则：

- ① z_i 与 z_j ($i \neq j; i, j=1, 2, \dots, m$) 相互无关；
- ② z_1 是 x_1, x_2, \dots, x_p 的一切线性组合中方差最大者， z_2 是与 z_1 不相关的 x_1, x_2, \dots, x_p 的所有线性组合中方差最大者； z_m 是与 z_1, z_2, \dots, z_{m-1} 都不相关的 x_1, x_2, \dots, x_p 的所有线性组合中方差最大者。

新变量指标 z_1, z_2, \dots, z_m 分别称为原变量指标 x_1, x_2, \dots, x_p 的第 1, 第 2, \dots , 第 m 主成分。

从以上的分析可以看出，主成分分析的实质就是确定原来变量 x_j ($j=1, 2, \dots, p$) 在诸主成分 z_i ($i=1, 2, \dots, m$) 上的荷载 l_{ij} ($i=1, 2, \dots, m; j=1, 2, \dots, p$)。

从数学上可以证明，它们分别是相关矩阵 m 个较大的特征值所对应的特征向量。

二、主成分分析的计算步骤

1、计算相关系数矩阵

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1p} \\ r_{21} & r_{22} & \dots & r_{2p} \\ \vdots & \vdots & & \vdots \\ r_{p1} & r_{p2} & \dots & r_{pp} \end{bmatrix}$$

r_{ij} ($i, j=1, 2, \dots, p$) 为原变量 x_i 与 x_j 的相关系数， $r_{ij}=r_{ji}$ ，其计算公式为

$$r_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}}$$

2、计算特征值与特征向量

解特征方程 $|\mathbf{A}-\lambda\mathbf{I}|=0$ ，常用雅可比法（Jacobi）求出特征值，并使其按大小顺序排列 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ ；

分别求出对应于特征值 λ_i 的特征向量 $\mathbf{e}_i (i=1,2,\dots,p)$ ，要求 $\|\mathbf{e}_i\|=1$ ，即 $\sum_{j=1}^p e_{ij}^2 = 1$ 其中 e_{ij} 表示向量 \mathbf{e}_i 的第 j 个分量。

3、计算主成分贡献率及累计贡献率

$$\begin{aligned} \text{贡献率: } & \frac{\lambda_i}{\sum_{k=1}^p \lambda_k} \quad (i=1,2,\dots,p) \\ \text{累计贡献率: } & \frac{\sum_{k=1}^i \lambda_k}{\sum_{k=1}^p \lambda_k} \quad (i=1,2,\dots,p) \end{aligned}$$

一般取累计贡献率达 85%–95% 的特征值， $\lambda_1, \lambda_2, \dots, \lambda_m$ 所对应的第 1、第 2、 \dots 、第 m ($m \leq p$) 个主成分。

4、计算主成分载荷

$$l_{ij} = p(z_i, x_j) = \sqrt{\lambda_i} e_{ij} (i, j=1,2,\dots,p)$$

5、各主成分得分

$$Z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1m} \\ z_{21} & z_{22} & \cdots & z_{2m} \\ \vdots & \vdots & & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nm} \end{bmatrix}$$

三、主成分分析法在 SPSS 中的操作

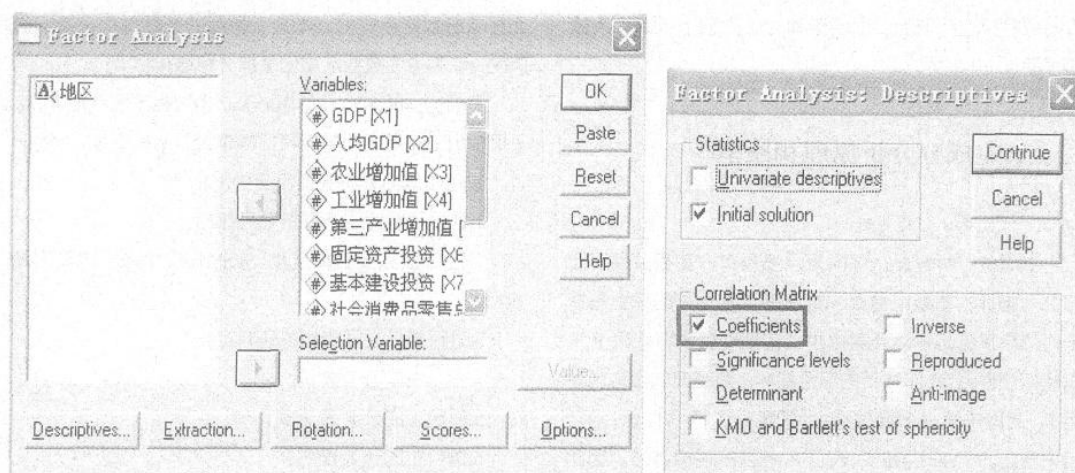
1、指标数据选取、收集与录入（表 1）

表 1 沿海 10 个省市经济数据

地区	GDP	人均 GDP	农 业 增加值	工 业 增加值	第三产业 增加值	固定资 产投资	基本建 设投资	社会消费品 零售总额	海关出 口总额	地方财 政收入
辽宁	5458.2	13000	14883.3	1376.2	2258.4	1315.9	529.0	2258.4	123.7	399.7
山东	10550.0	11643	1390.0	3502.5	3851.0	2288.7	1070.7	3181.9	211.1	610.2
河北	6076.6	9047	950.2	1406.7	2092.6	1161.6	597.1	1968.3	45.9	302.3
天津	2022.6	22068	83.9	822.8	960.0	703.7	361.9	941.4	115.7	171.8
江苏	10636.0	14397	1122.6	3536.3	3967.2	2320.0	1141.3	3215.8	384.7	643.7
上海	5408.8	40627	86.2	2196.2	2755.8	1970.2	779.3	2035.2	320.5	709.0
浙江	7670.0	16570	680.0	2356.5	3065.0	2296.6	1180.6	2877.5	294.2	566.9
福建	4682.0	13510	663.0	1047.1	1859.0	964.5	397.9	1663.3	173.7	272.9
广东	11770.0	15030	1023.9	4224.6	4793.6	3022.9	1275.5	5013.6	1843.7	1202.0
广西	2437.2	5062	591.4	367.0	995.7	542.2	352.7	1025.5	15.1	186.7

2、Analyze →Data Reduction →Factor Analysis, 弹出Factor Analysis 对话框:

表 2 Factor Analyze 对话框与 Descriptives 子对话框



3、把指标数据选入 Variables 框, Descriptives: Correlation Matrix 框组中选中 Coefficients, 然后点击 Continue, 返回 Factor Analysis 对话框, 单击 OK。

注意: SPSS 在调用Factor Analyze 过程进行分析时, SPSS 会自动对原始数据进行标准化处理, 所以在得到计算结果后的变量都是指经过标准化处理后的变量, 但SPSS 并不直接给出标准化后的数据, 如需要得到标准化数据, 则需调用Descriptives 过程进行计算。

表 3 相关系数矩阵

Correlation Matrix										
	GDP	人均 GDP	农 业 增加值	工 业 增加值	第三 产业 增加值	固定 资产 投资	基本 建设 投资	社会消 费品零 售总额	海关 出口 总额	地方 财政 收入
GDP	1.000	-0.094	-0.052	0.967	0.979	0.923	0.922	0.941	0.637	0.826
人均 GDP	-0.094	1.000	-0.171	0.113	0.074	0.214	0.093	-0.043	0.081	0.273
农业增加值	-0.052	-0.171	1.000	-0.132	-0.050	-0.098	-0.176	0.013	-0.125	-0.086
工业增加值	0.967	0.113	-0.132	1.000	0.985	0.963	0.939	0.935	0.705	0.898
第三产业增加值	0.979	0.074	-0.050	0.985	1.000	0.973	0.940	0.962	0.714	0.913
固定资产投资	0.923	0.214	-0.098	0.963	0.973	1.000	0.971	0.937	0.717	0.934
基本建设投资	0.922	0.093	-0.176	0.939	0.940	0.971	1.000	0.897	0.624	0.848
社会消费品零售总额	0.941	-0.043	0.013	0.935	0.962	0.937	0.897	1.000	0.836	0.929
海关出口总额	0.637	0.081	-0.125	0.705	0.714	0.717	0.624	0.836	1.000	0.882
地方财政收入	0.826	0.273	-0.086	0.898	0.913	0.934	0.848	0.929	0.882	1.000

表 4 方差分解主成分提取分析表

Total Variance Explained						
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7.220	72.205	72.205	7.220	72.205	72.205
2	1.235	12.346	84.551	1.235	12.346	84.551
3	0.877	8.769	93.319			
4	0.547	5.466	98.786			
5	0.085	0.854	99.640			
6	0.021	0.211	99.850			
7	0.012	0.119	99.970			
8	0.002	0.018	99.988			
9	0.001	0.012	100.000			
10	0.000	0.000	100.000			

Extraction Method: Principal Component Analysis.

表 5 初始因子载荷矩阵

	Component	
	1	2
GDP	0.949	0.195
人均 GDP	0.112	-0.824
农业增加值	-0.109	0.677
工业增加值	0.978	-0.005
第三产业增加值	0.986	0.070
固定资产投资	0.983	-0.068
基本建设投资	0.947	-0.024
社会消费品零售总额	0.977	0.176
海关出口总额	0.800	-0.051
地方财政收入	0.954	-0.128

Extraction Method: Principal Component Analysis.
a. 2 components extracted.

从表3 可知GDP 与工业增加值, 第三产业增加值、固定资产投资、基本建设投资、社会消费品零售总额、地方财政收入这几个指标存在着极其显著的关系, 与海关出口总额存在着显著关系。可见许多变量之间直接的相关性比较强, 证明他们存在信息上的重叠。

主成分个数提取原则为主成分对应的特征值大于1的前m个主成分。特征值在某种程度上可以被看成是表示主成分影响力度大小的指标, 如果特征值小于1, 说明该主成分的解释力度还不如直接引入一个原变量的平均解释力度大, 因此一般可以用特征值大于1作为纳入标准。通过表4(方差分解主成分提取分析) 可知, 提取2个主成分, 即m=2, 从表5(初始因子载荷矩阵) 可知GDP、工业增加值、第三产业增加值、固定资产投资、基本建设投资、社会消费品零售总额、海关出口总额、地方财政收入在第一主成分上有较高载荷, 说明第一主成分基本反

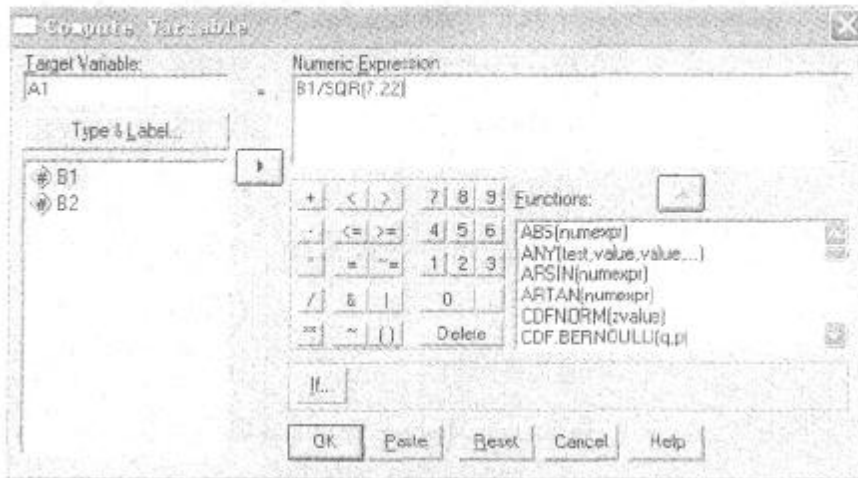
映了这些指标的信息；人均GDP 和农业增加值指标在第二主成分上有较高载荷，说明第二主成分基本反映了人均GDP 和农业增加值两个指标的信息。所以提取两个主成分是可以基本反映全部指标的信息，所以决定用两个新变量来代替原来的十个变量。但这两个新变量的表达还不能从输出窗口中直接得到，因为“Component Matrix”是指初始因子载荷矩阵，每一个载荷量表示主成分与对应变量的相关系数。

用表5(主成分载荷矩阵)中的数据除以主成分相对应的特征值开平方根便得到两个主成分中每个指标所对应的系数。将初始因子载荷矩阵中的两列数据输入(可用复制粘贴的方法)到数据编辑窗口(为变量B1、B2)，然后利用“Transform→Compute Variable”，在Compute Variable对话框中输入“A1=B1/SQR(7.22)”[注：第二主成分SQR后的括号中填1.235，即可得到特征向量 A_1 (见表6)。同理，可得到特征向量 A_2 。将得到的特征向量与标准化后的数据相乘，然后就可以得出主成分表达式[注：因本例只是为了说明如何在SPSS 进行主成分分析，故在此不对提取的主成分进行命名，有兴趣的读者可自行命名。

$$F_1 = 0.353ZX_1 + 0.042ZX_2 - 0.041ZX_3 + 0.364ZX_4 + 0.367ZX_5 + 0.366ZX_6 + 0.352ZX_7 + 0.364ZX_8 + 0.298ZX_9 + 0.355ZX_{10}$$

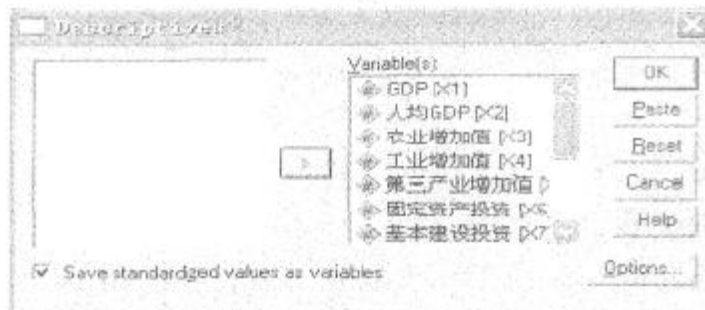
$$F_2 = 0.175ZX_1 - 0.741ZX_2 + 0.609ZX_3 - 0.004ZX_4 + 0.063ZX_5 - 0.061ZX_6 - 0.022ZX_7 + 0.158ZX_8 - 0.046ZX_9 - 0.115ZX_{10}$$

表 6 Compute Variable 对话框



标准化：通过Analyze→Descriptive Statistics→Descriptives 对话框来实现：弹出Descriptives 对话框后，把 $X_1 \sim X_{10}$ 选入Variables 框，在Save standardized values as variables 前的方框打上钩，点击“OK”，经标准化的数据会自动填入数据窗口中，并以Z开头命名。

表 7 Descriptives 对话框



以每个主成分所对应的特征值占所提取主成分总的特征值之和的比例作为权重计算主成分综合模型，即用第一主成分F1 中每个指标所对应的系数乘上第一主成分F1 所对应的贡献率再除以所提取两个主成分的两个贡献率之和，然后加上第二主成分F2 中每个指标所对应的系数乘上第二主成分F2 所对应的贡献率再除以所提取两个主成分的两个贡献率之和，即可得到综合得分模型：

$$F=0.327ZX_1-0.072ZX_2+0.054ZX_3+0.310ZX_4+0.323ZX_5+0.304ZX_6+0.297ZX_7+0.334ZX_8+0.248ZX_9+0.286ZX_{10}$$

根据主成分综合模型即可计算综合主成分值，并对其按综合主成分值进行

排序，即可对各地区进行综合评价比较，结果见表8。

表 8 综合主成分值

城市	第一主成分 F_1	排名	第二主成分 F_2	排名	综合主成分 F	排名
广东	5.23	1	0.11	6	4.48	1
江苏	2.25	2	0.23	5	1.96	2
山东	1.96	3	0.50	2	1.75	3
浙江	1.16	4	- 0.19	8	0.96	4
上海	0.30	5	- 2.36	10	- 0.09	5
辽宁	- 1.24	6	1.96	1	- 0.78	6
河北	- 1.35	7	0.41	4	- 1.10	7
福建	- 1.97	8	- 0.07	7	- 1.70	8
天津	- 3.04	9	- 1.01	9	- 2.74	9
广西	- 3.29	10	0.41	3	- 2.75	10

具体检验还需进一步探讨与学习