

---

# 多元数据处理

## ---因子分析方法

多元数据处理主要包括多元随机变量，协方差分析，趋势面分析，聚类分析，判别分析，主成分分析，因子分析，典型相关分析，回归分析以及各个分析方法的相互结合等等。本文主要针对其中的因子分析方法展开了论述，并举了一个因子分析法在我国房地产市场绩效评价中的应用实例。

### 第一章 因子分析方法概述

#### 1.1 因子分析的涵义

为了更全面和准确的测量和评估对象的特征，在实际的应用中，我们往往尽可能多的选用特征指标进行系统评估，选取的指标越多，就越能全面、客观的反映评价对象的特征。选取众多指标的同时也带来了统计分析的困难：一、不同的指标，不同重要程度需要赋予不同的权重，而靠主观的评价避免不了一些失误与错误。二、收集到的指标之间可能存在较大的相关性，大量收集指标带来了人力、物力和财力的浪费。而因子分析方法则较好的解决了上述问题。

因子分析<sup>[1]</sup>是一种多元统计方法，该方法起源于 20 世纪初 Karl Pearson 和 Charles Spearman 等人关于心理测试的统计分析，它的核心是用最少的相互独立的因子反映原有变量的绝大部分信息。<sup>[2]</sup>通过分析事物内部的因果关系来找出其主要矛盾，找出事物内在的基本规律。

因子分析的基本思想是通过变量的相关系数矩阵内部结构的研究，找出能控制所有变量的少数几个随机变量去描述多个变量之间的相关关系，但是，这少数几个随机变量是不可观测的，通常称为因子。然后根据相关性的大小把变量分组，使得同组内的变量之间相关性较高，使不同组内的变量相关性较低<sup>[3]</sup>。对于所研究的问题就可试图用最少数个数的所谓因子的线性函数与特殊因子之和来描述原来观测的每一变量<sup>[4]</sup>。因子变量的特点：第一，因子变量的数量远小于原指标的数量，对因子变量的分析能够减少分析的工作量；第二，因子变量不是原有变量的简单取舍，而是对原有变量的

重新组构，他们能够反映原有变量的绝大部分信息，不会产生丢失；第三，因子变量之间线性相关性较低；第四，因子变量具有命名解释性<sup>[5]</sup>。因子分析可以消除指标间的信息重叠，抽象出事物的本质属性，不仅可以综合评价，还可以综合分析对其产生影响的主要因素。

## 1.2 因子分析统计模型

设  $p$  个可以观测的指标为  $X_1, X_2, X_3, \dots, X_p$ ， $m$  个不可观测的因子为  $F_1, F_2, F_3, \dots, F_m$ ，则因子分析模型描述如下：<sup>[6][7]</sup>

$$\begin{aligned} X_1 &= a_{11}F_1 + a_{12}F_2 + \dots + a_{1m}F_m + \varepsilon_1 \\ X_2 &= a_{21}F_1 + a_{22}F_2 + \dots + a_{2m}F_m + \varepsilon_2 \\ &\dots\dots\dots \\ X_p &= a_{p1}F_1 + a_{p2}F_2 + \dots + a_{pm}F_m + \varepsilon_p \\ m &< p \end{aligned}$$

其中： $m < p$

$F = (F_1, F_2, \dots, F_m)$  是不可测的向量，我们把  $F$  称为  $X$  的公共因子，其均值向量  $E(F) = 0$ ，协方差矩阵  $Cov(F) = 1$ ，即向量的各分量是相互独立的  $\varepsilon(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)$  是特殊因子，与  $F$  相互独立，且  $E(\varepsilon) = 0$ 。

$A = (a_{ij})$ ， $a_{ij}$  为因子载荷，数学上可以证明，因子载荷  $a_{ij}$  就是第  $i$  指标与第  $j$  因子的相关系数，载荷越大，说明第  $j$  个指标与第  $i$  个因子的关系越密切；反之载荷越小，关系越疏远<sup>[8]</sup>。

## 1.3 因子分析步骤

### (1) 原始数据的标准化

原始数据的标准化包括指标正向化合和无量纲化处理两方面。在多指标的评价中，有些指标数值越大，评价越好；有些指标数值越小，评价越好，这种指标称为逆向指标；还有些指标数值越靠近某个具体数值越好，这种指标称为适度指标。根据不同类型的指标需要将逆向指标、适度指标转化为正向指标，此过程称为指标的正向化。指标正向化过程既可以在无量纲化前处理也可以在无量纲化时处理。逆向指标可以选用公式  $X'_i = (X_{\max} - X_i) / (X_{\max} - X_{\min})$ 。其中， $X_{\max}$ 、 $X_{\min}$  分别为指标的最大与最小值。适度指标方面，叶宗裕<sup>[9]</sup>认为正向化可以采用指标值减去适度值的绝对值的相反数。公式为  $Y_{xy} = -|X_{xy} - M|$ 。其中  $Y_{xy}$  为正向后数据， $X_{xy}$  为原始数据， $M$  为适度值。

指标的无量纲化则是通过标准化处理，将不同的指标通过数学变换转化为统一的相对值，消除各个指标不同量纲的影响。常用的无量纲化包括：标准化法、均值法和极差正规化法。本文采用最常见的标准化法进行无量纲化处理，公式处理如下：（ $\bar{X}$  是  $X$  的期望值， $\sigma_x$  是  $X$  的标准差）

$$Z = \frac{X - \bar{X}}{\sigma_x}$$

#### （2）计算相关矩阵 $R$ 的特征值和特征向量

根据特征方程  $|R - \lambda E| = 0$ ，计算相关矩阵的特征值  $\lambda$  及对应的特征向量  $A$ ， $\lambda$  的大小描述了各个因子在解释对象所起的作用的大小。

#### （3）计算因子贡献率及累积贡献率，确定公共因子个数

因子贡献率表示每个因子的变异程度占所有因子变异程度的比率，公式为：

$$C_i = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i}$$

， $C_i$  表示方差贡献率。当累积贡献率达到 85% 以上或者特征根  $\lambda$  不小于 1，即确定了公因子的个数。

#### （4）求解初始因子载荷矩阵

$X=AF$ ，因子载荷矩阵  $A$  并不唯一，软件则是运用不同的参数估计方法求出相应的估计矩阵，参数估计方法主要包括：最小平方法、极大似然法、主成分法、主因子法、多元回归法。

#### （5）因子载荷矩阵的旋转

若因子载荷较为平均，初始的因子载荷矩阵描述的经济含义不太明显，难以判断与各个因子的关系时，就需要进行因子旋转。通过因子旋转，使旋转后公共因子的贡献更加分散，并对主因子进行命名，确定经济含义<sup>[10]</sup>。因子旋转主要有正交旋转法和斜交旋转法。

#### （6）计算样本的综合得分

通过因子载荷矩阵，可以得出因子的因子得分系数矩阵  $B$ 。然后计算出每个因子的得分  $F=BZ$ ，最后以各因子的方差贡献率占因子总方差的贡献率的比重作为权重加权汇总，得到应变综合得分

$$F = \frac{\lambda_1}{\sum_{i=1}^m \lambda_i} F_1 + \frac{\lambda_2}{\sum_{i=1}^m \lambda_i} F_2 + \cdots + \frac{\lambda_m}{\sum_{i=1}^m \lambda_i} F_m$$

## 第二章 我国房地产市场绩效实证研究

运用构建的房地产市场绩效评价指标体系及因子分析方法实证分析 2007、2008、2009 三年我国 31 个省市各房地产市场的绩效水平,利用 SPSS 软件进行因子分析与评价。我国房地产业市场绩效指标体系见表 2-1。

表 2-1 我国房地产业市场绩效指标体系

总体层	状态层	变量层次	单位
房地产业市场绩效	产业规模	X1 企业个数	个
		X2 本年完成投资额	万元
		X3 本年资金来源合计	万元
		X4 本年购置土地面积	m <sup>2</sup>
	商品房开发	X5 房屋施工面积	m <sup>2</sup>
		X6 新开工面积	m <sup>2</sup>
		X7 商品房屋竣工面积	m <sup>2</sup>
	商品房销售	X8 商品房销售额	万元
		X9 商品房销售面积	m <sup>2</sup>
		X10 商品房销售套数	套
	企业经济效益	X11 企业利润总额	万元
		X12 企业经营收入	万元
		X13 企业所有者权益	万元
	社会评价	X14 从业人数	人
		X15 房屋销售价格	元/ m <sup>2</sup>
		X16 商品房空置面积	m <sup>2</sup>

由上面的指标体系可以看出,我国房地产业市场绩效包括三个层次,第一层为为总体层,总体层又称为目标层,它反映了此指标体系的目的。第二层为状态层,本文分五个方面对目标层进行解释。

### 2.1 数据的采集和整理

根据本文构建的房地产绩效评价指标体系,选取了《房地产统计年鉴》中 2007-2009 年我国 31 个省市基础数据(不包括台湾省、香港特别行政区、澳门特别行政区 3 个地区)。

## 2.2 数据的同向化处理

本文中评价房地产业市场绩效的指标中 X16 房屋空置面积这个指标是绩效评价的逆向指标，并不是越大越好，为了与其他指标保持同向性，需要将其转化为正向指标，转化公式为  $X'_i = (X_{\max} - X_i) / (X_{\max} - X_{\min})$ 。其中， $X_{\max}$  为评价指标最大值， $X_{\min}$  为评价指标最小值。

## 2.3 数据的标准化处理

由于不同变量之间存在了不同量纲、不同数量级的情况，需要对原指标数据进行标准化处理，把不同指标数据转化成相同量纲的数据，是得各指标数据具有可比性。

标准化处理方法如下：

$$Z = \frac{X - \bar{X}}{\sigma_x} \quad (\bar{X} \text{ 是 } X \text{ 的期望值, } \sigma_x \text{ 是 } X \text{ 的标准查差})$$

## 2.4 确定是否适合因子分析：KMO 检验和 Bartlett 球形检验

KMO 检验给出抽样充足量的测度，检验变量间的偏相关系数是否过小。Bartlett 球形检验检验相关系数矩阵是否是单位阵，如果是单位阵，则表明不适合采用因子模型。经 SPSS 检验结果如下表 2-2。

根据 Kaiser 给出的是否做因子分析的 KMO 标准为：KMO>0.9，非常适合；0.9>KMO>0.8，适合；0.8>KMO>0.7，一般；0.7>KMO>0.6，不太适合；KMO<0.5 不适合，故 KMO 检验通过。同时，相伴概率为 0.000，小于显著水平 0.05，表明 Bartlett 球形检验通过，所以本文所选的变量适合做因子分析。

表 2-2 KMO 检验结果和 Barlett 球形检验结果

KMO 和 Bartlett 的检验		
取样足够度的		0.834
Kaiser-Meyer-Olkin 度量。		
Bartlett	近似卡方	1115.445
的球形度	df	120
检验	Sig.	0

## 2.5 指标相关性检验

根据本文构建的房地产业市场绩效指标体系，借助于多元分析软件 SPSS，利用我国 2007 年房地产业基础数据对我国房地产业市场绩效做出分析与评价。通过 SPSS 软件进行的相关性分析，得到 2007 年全国各省市房地产市场绩效评价指标的相关系数矩阵，如表 2-3

表 2-3 Correlation Matrix (相关系数矩阵)

	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>
X <sub>1</sub>	1	0.932	0.895	0.758	0.947	0.887	0.915	0.888	0.869	0.891
X <sub>2</sub>	0.93	1	0.962	0.664	0.958	0.877	0.95	0.96	0.848	0.865
X <sub>3</sub>	0.9	0.962	1	0.503	0.888	0.748	0.865	0.984	0.757	0.753
X <sub>4</sub>	0.76	0.664	0.503	1	0.774	0.898	0.778	0.508	0.825	0.886
X <sub>5</sub>	0.95	0.958	0.888	0.774	1	0.933	0.972	0.896	0.923	0.909
X <sub>6</sub>	0.89	0.877	0.748	0.898	0.933	1	0.955	0.763	0.91	0.964
X <sub>7</sub>	0.92	0.95	0.865	0.778	0.972	0.955	1	0.893	0.927	0.944
X <sub>8</sub>	0.89	0.96	0.984	0.508	0.896	0.763	0.893	1	0.766	0.776
X <sub>9</sub>	0.87	0.848	0.757	0.825	0.923	0.91	0.927	0.766	1	0.927
X <sub>10</sub>	0.89	0.865	0.753	0.886	0.909	0.964	0.944	0.776	0.927	1
X <sub>11</sub>	0.72	0.723	0.805	0.276	0.681	0.493	0.652	0.856	0.549	0.566
X <sub>12</sub>	0.86	0.922	0.97	0.434	0.846	0.691	0.835	0.987	0.713	0.724
X <sub>13</sub>	0.68	0.726	0.81	0.246	0.64	0.446	0.639	0.855	0.539	0.529
X <sub>14</sub>	0.96	0.874	0.857	0.744	0.89	0.839	0.866	0.84	0.873	0.888
X <sub>15</sub>	0.52	0.684	0.792	0.033	0.522	0.309	0.511	0.798	0.342	0.332
X <sub>16</sub>	-0.8	-0.829	-0.819	-0.527	-0.766	-0.726	-0.793	-0.809	-0.726	-0.76

续上表

	$X_{11}$	$X_{12}$	$X_{13}$	$X_{14}$	$X_{15}$	$X_{16}$
$X_1$	0.718	0.858	0.677	0.955	0.52	-0.795
$X_2$	0.723	0.922	0.726	0.874	0.684	-0.829
$X_3$	0.805	0.97	0.81	0.857	0.792	-0.819
$X_4$	0.276	0.434	0.246	0.744	0.033	-0.527
$X_5$	0.681	0.846	0.64	0.89	0.522	-0.766
$X_6$	0.493	0.691	0.446	0.839	0.309	-0.726
$X_7$	0.652	0.835	0.639	0.866	0.511	-0.793
$X_8$	0.856	0.987	0.855	0.84	0.798	-0.809
$X_9$	0.549	0.713	0.539	0.873	0.342	-0.726
$X_{10}$	0.566	0.724	0.529	0.888	0.332	-0.76
$X_{10}$	0.566	0.724	0.529	0.888	0.332	-0.76
$X_{11}$	1	0.91	0.918	0.7	0.735	-0.639
$X_{12}$	0.91	1	0.904	0.827	0.826	-0.784
$X_{13}$	0.918	0.904	1	0.682	0.811	-0.587
$X_{14}$	0.7	0.827	0.682	1	0.492	-0.736
$X_{15}$	0.735	0.826	0.811	0.492	1	-0.512
$X_{16}$	-0.639	-0.784	-0.587	-0.736	-0.512	1

通过以上相关系数矩阵的分析可以看出,各个房地产市场绩效指标之间有较大的相关性,如果单纯以一个指标来评价市场绩效指标就会存在不够准确甚至重迭。为了消除指标间的重迭,简化计算,可以采用因子分析的方法进行降维处理,把原来比较复杂的相关矩阵内部找出几个综合指标,使综合指标为原来变量的线性组合,利用相对较少的因子研究市场绩效。

## 2.6 共同度分析

根据变量共同度的统计意义,它刻画了全部公共因子对于原始变量的总方差所作的贡献,它说明了全部公共因子反映出原变量信息的百分比<sup>[11]</sup>。如下表 2-4 所示的

变量共同度可知，除了 X13 企业所有者权益、X14 从业人数、X15 房屋销售价格、X16 商品房空置面积、的共同度为 0.880、0.878、0.877、0.705，其余变量的共同度都在 90%以上，因此这四个公共因子对各变量的解释能力是比较强的。采用因子分析房地产市场绩效的效果是比较好的。

表 2-4 共同性公因子方差

	初始	提取
X <sub>1</sub>	1	0.937
X <sub>2</sub>	1	0.958
X <sub>3</sub>	1	0.957
X <sub>4</sub>	1	0.921
X <sub>5</sub>	1	0.957
X <sub>6</sub>	1	0.967
X <sub>7</sub>	1	0.958
X <sub>8</sub>	1	0.986
X <sub>9</sub>	1	0.907
X <sub>10</sub>	1	0.957
X <sub>11</sub>	1	0.847
X <sub>12</sub>	1	0.993
X <sub>13</sub>	1	0.88
X <sub>14</sub>	1	0.878
X <sub>15</sub>	1	0.877
X <sub>16</sub>	1	0.705

## 2.7 公共因子分析

因子载荷是公共因子与指标变量之间的相关系数，载荷越大，说明公共因子与指标变量之间的关系越密切。在确定公共因子个数时，先选择与原变量数目相等的因子个数，其因子计算结果见下表 2-5。取初始特征矩阵大于 1 的因子为公共因子。经过总方差分解，可以明显看出有二个因子旋转后特征值大于 1，它们的方差贡献率分别为 50.574%、41.199%，累计贡献率为 91.772%。当累积贡献率达到 85%以上，因此完



全可以采用这二个因子概况原始数据对全国 31 个省市的房地产市场绩效做出评价是合适的。

表 2-5 总方差分解

成份	初始特征值			提取平方和载入			旋转平方和载入		
	合计	方差 的 %	累积 %	合计	方差 的 %	累积 %	合计	方差的 %	累积 %
1	12.549	78.432	78.432	12.549	78.432	78.432	8.092	50.574	50.574
2	2.134	13.341	91.772	2.134	13.341	91.772	6.592	41.199	91.772

从未转轴的因素矩阵（见下表 2-6）可以看出，结果并不非常令人满意，有 2 个因素被抽取，所以本文采用方差最大化正交旋转方法对因子进行了旋转，得到了因子载荷矩阵，进而更清楚地观察样本。从旋转后的因素矩阵（见下表 2-7）可以看出：

（1）X4 本年购置土地面积、X6 新开工面积、X10 商品房销售套数、X9 商品房销售面积、X7 商品房屋竣工面积、X5 房屋施工面积、X1 企业个数、X14 从业人数、X2 本年完成投资额、X16 商品房空置面积为第一主因子，他们的载荷值分别为：0.959、0.948、0.926、0.894、0.848、0.835、0.802、0.781、0.728、0.630；这些指标都是从一个方面反映关于房地产开发销售方面的情况，故可以命名此公共因子 F1 为：房地产市场开发销售。

表 2-6 未旋转因素矩阵表

	1	2
X <sub>2</sub>	0.979	0
X <sub>5</sub>	0.965	-0.162
X <sub>8</sub>	0.962	0.246
X <sub>1</sub>	0.961	-0.115
X <sub>7</sub>	0.961	-0.186
X <sub>3</sub>	0.951	0.231
X <sub>12</sub>	0.936	0.343
X <sub>14</sub>	0.929	-0.12
X <sub>10</sub>	0.906	-0.368
X <sub>9</sub>	0.891	-0.336
X <sub>6</sub>	0.888	-0.421
X <sub>16</sub>	-0.839	0
X <sub>11</sub>	0.79	0.471
X <sub>13</sub>	0.774	0.531
X <sub>4</sub>	0.705	-0.651
X <sub>15</sub>	0.645	0.678

2-7 转轴因素矩阵

	1	2
X <sub>4</sub>	0.959	0
X <sub>6</sub>	0.948	0.263
X <sub>10</sub>	0.926	0.315
X <sub>9</sub>	0.894	0.329
X <sub>7</sub>	0.848	0.488
X <sub>5</sub>	0.835	0.509
X <sub>1</sub>	0.802	0.541
X <sub>14</sub>	0.781	0.517
X <sub>2</sub>	0.728	0.655
X <sub>16</sub>	-0.63	-0.555
X <sub>15</sub>	0	0.935
X <sub>13</sub>	0.238	0.908
X <sub>11</sub>	0.289	0.874
X <sub>12</sub>	0.483	0.872
X <sub>8</sub>	0.566	0.816
X <sub>3</sub>	0.568	0.796

（2）X15 房屋销售价格、X13 企业所有者权益、X11 企业利润总额、X12 企业

经营收入、X8 商品房销售额、X3 本年资金来源合计为第二因子，他们的载荷值分别为：0.935、0.908、0.874、0.872、0.816、0.796；这些指标都是在一定程度上能够反映房地产企业的经济效益，故可以命名此公共因子 F2 为：房地产企业综合效益水平。

## 2.8 计算因子得分计算分析及结果

表 2-8 为因子得分系数矩阵，根据因子得分系数和原始变量的值可以计算出每个观测值的各因子的分数，并可以据此对观测值进行下一步的分析。旋转后的因子得分表达式与计算结果如下：

$$F_1 = 0.093X_1 + 0.053X_2 - 0.013X_3 + 0.242X_4 + 0.108X_5 + 0.183X_6 + 0.115X_7 - 0.018X_8 + 0.157X_9 + 0.167X_{10} - 0.097X_{11} - 0.049X_{12} - 0.116X_{13} + 0.093X_{14} - 0.169X_{15} - 0.048X_{16}$$

$$F_2 = 0.009X_1 + 0.058X_2 + 0.131X_3 - 0.194X_4 - 0.007X_5 - 0.103X_6 - 0.016X_7 + 0.137X_8 - 0.073X_9 - 0.083X_{10} + 0.208X_{11} + 0.170X_{12} + 0.228X_{13} + 0.006X_{14} + 0.274X_{15} - 0.047X_{16}$$

表 2-8 2007 年各因子得分系数矩阵

	成份	
	1	2
X <sub>1</sub>	0.093	0.009
X <sub>2</sub>	0.053	0.058
X <sub>3</sub>	-0.013	0.131
X <sub>4</sub>	0.242	-0.194
X <sub>5</sub>	0.108	-0.007
X <sub>6</sub>	0.183	-0.103
X <sub>7</sub>	0.115	-0.016
X <sub>8</sub>	-0.018	0.137
X <sub>9</sub>	0.157	-0.073
X <sub>10</sub>	0.167	-0.083
X <sub>11</sub>	-0.097	0.208
X <sub>12</sub>	-0.049	0.17
X <sub>13</sub>	-0.116	0.228
X <sub>14</sub>	0.093	0.006
X <sub>15</sub>	-0.169	0.274
X <sub>16</sub>	-0.048	-0.047

由估计出的因子的得分，可以描述我国各省市绩效水平，利用因子得分可以从不同的角度对我国各省市房地产市场绩效水平进行比较分析。为了对我国各省市房地产业市场绩效进行评价，现利用各省市因子得分表计算综合得分，各省市房地产市场绩效的获取是基于总方差分解表中旋转后各因子的方差贡献率及计算所得的上市公司各因得分所得，其具体计算公式为：

$$\text{综合绩效} = (50.574\% \times F_1 + 41.199\% \times F_2) / 91.772\%。$$

详细情况见表 2-9。

表 2-9 2007 年 31 个省（市）各因子得分及排名

地区	房地产市场开发销售		房地产企业综合效益	
	因子得分	排名	因子得分	排名
北京	-0.541	21	2.6789	2
天津	-0.6929	24	0.3734	5
河北	-0.093	15	-0.4336	19
山西	-0.6313	22	-0.4278	18
内蒙古	-0.0726	14	-0.6782	30
辽宁	1.175	5	-0.2956	12
吉林	-0.4363	18	-0.4775	22
黑龙江	-0.2948	17	-0.3253	13
上海	-0.8737	26	3.4528	1
江苏	2.5437	1	0.3358	6
浙江	0.9249	6	1.024	4
安徽	0.4774	8	-0.5424	24
福建	-0.0034	12	0.3237	7
江西	-0.0084	13	-0.5737	26
山东	2.0432	2	-0.5865	27
河南	0.4345	10	-0.6015	28
湖北	0.1252	11	-0.2822	11
湖南	0.4673	9	-0.6931	31
广东	1.6465	3	2.0979	3
广西	-0.1966	16	-0.4466	21
海南	-1.3167	30	-0.0612	9
重庆	0.7485	7	-0.5106	23
四川	1.5469	4	-0.6275	29
贵州	-0.5118	19	-0.5564	25
云南	-0.5284	20	-0.0188	8
西藏	-1.3458	31	-0.3402	14
陕西	-0.6617	23	-0.2693	10
甘肃	-0.9342	27	-0.3912	17
青海	-1.2624	29	-0.347	15
宁夏	-1.0306	28	-0.356	16
新疆	-0.6979	25	-0.4444	20

为了更直观的观察，本文也给出了样本城市因子的柱状图，如下图 2-1，由表 2-9 可以看出 2007 年北京、上海在房地产市场开发因子 F1 得分仅-0.541、-0.8737，在 31 个省市自治区排名分别为 21、26；而在房地产企业综合绩效 F2 得分为 2.6789、3.4528，排名为第二、第一。江苏、广东在因子 F1 得分分别为：2.5437、1.6465，排名第一、第三；在因子 F2 得分为：0.3358、2.0979，排名第六、第三。江苏、广东在因子 F1、F2 得分均靠前。

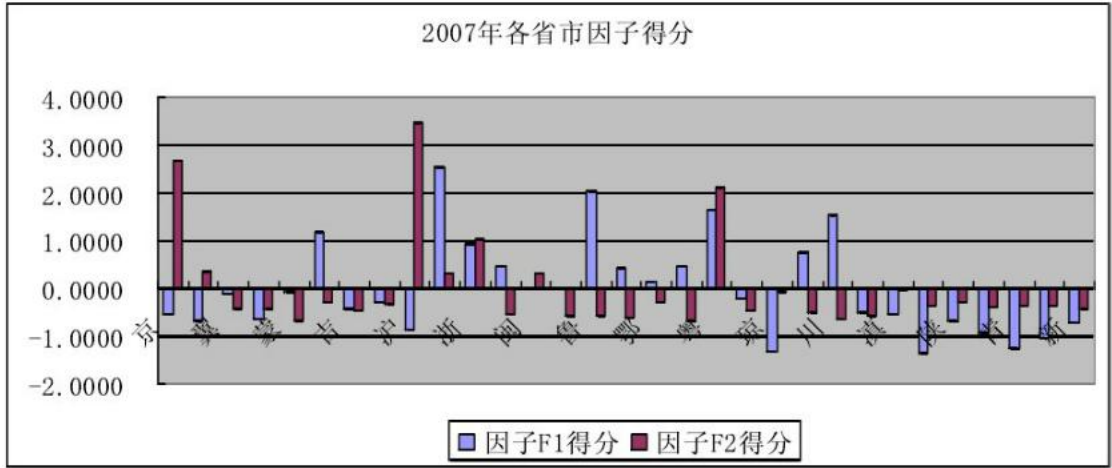


图 2-1 2007 年各省市因子得分

使用同样的方法我们也给出了 2008 年及 2009 年各省市因子得分及排名情况。如表 2-10 和 2-11 所示。

通过表 2-9、2-10、2-11 的分析可知，广东、江苏、浙江三省份在因子 F1 房地产市场开发销售、F2 房地产企业综合效益得分均较高，发展相对平衡，总的绩效水平高。而其他各省市都有某一个因子或两个得分相对较低，即所谓的“短板”，发展不平衡影响了其总的绩效水平。

由表 2-12 可以看出，2007-2009 三年内北京综合绩效得分排名从第五下降到第八，上海从第三下降到第七，市场绩效没有得到提高，房地产发展存在一定问题；江苏从 2007 年的第二到 2009 年的第一，山东从 2007 年的第六到 2009 年第三，绩效持续提高。西藏、陕西、甘肃、青海、宁夏、新疆等省市排名一直靠后，绩效水平没有得到提高。

表 2-10 2008 年 31 个省（市）各因子得分及排名

地区	房地产市场开发销售因子		房地产企业综合效益	
	因子得分	排名	因子得分	排名
北京	-1.2078	29	2.9111	2
天津	-0.8173	24	0.3854	6
河北	0.3335	11	-0.5948	27
山西	-0.5715	21	-0.5376	20
内蒙古	0.2699	12	-0.6283	29
辽宁	1.4896	4	-0.2015	11
吉林	-0.2545	18	-0.5218	19
黑龙江	-0.3949	19	-0.3632	14
上海	-1.0129	26	3.0716	1
江苏	2.5965	1	1.1075	4
浙江	0.5769	9	1.2212	4
安徽	0.6802	7	-0.5844	26
福建	-0.2407	17	0.2226	7
江西	-0.1408	15	-0.5494	22
山东	2.1169	2	-0.3857	15
河南	0.96	5	-0.7032	31
湖北	0.0191	13	-0.1741	10
湖南	0.5786	9	-0.6505	30
广东	1.5236	3	1.891	3
广西	-0.0898	14	-0.5402	21
海南	-1.1361	28	-0.0262	8
重庆	0.3444	10	-0.2851	12
四川	0.7786	6	-0.1564	9
贵州	-0.6073	23	-0.5567	23
云南	-0.2285	16	-0.5967	27
西藏	-1.3723	31	-0.3401	13
陕西	-0.4633	20	-0.3924	16
甘肃	-0.8235	25	-0.5717	24
青海	-1.213	30	-0.4549	18
宁夏	-1.0667	27	-0.4157	17
新疆	-0.6268	23	-0.5798	25

表 2-11 2009 年 31 个省（市）各因子得分及排名

地区	房地产市场开发销售因子		房地产企业综合效益	
	因子得分	排名	因子得分	排名
北京	-0.9623	26	2.6616	2
天津	-0.8373	24	0.3071	7
河北	0.537	10	-0.6389	28
山西	-0.3881	20	-0.6591	29
内蒙古	-0.0156	13	-0.517	23
辽宁	1.5225	3	-0.2796	11
吉林	-0.4307	21	-0.4621	19
黑龙江	-0.3534	18	-0.3846	16
上海	-1.3397	31	3.2511	1
江苏	2.3768	1	1.2599	5
浙江	0.6413	8	1.3077	4
安徽	0.5503	9	-0.4039	16
福建	-0.2849	17	0.3387	6
江西	-0.2653	16	-0.5001	22
山东	2.0222	2	-0.2818	12
河南	1.0857	5	-0.7157	30
湖北	0.2943	12	-0.315	13
湖南	0.6853	7	-0.7674	31
广东	1.4719	4	1.8048	3
广西	-0.244	15	-0.378	14
海南	-1.2381	29	-0.079	8
重庆	0.4469	11	-0.2782	10
四川	1.0623	6	-0.2227	9
贵州	-0.6369	22	-0.4661	20
云南	-0.195	14	-0.5823	27
西藏	-1.3237	30	-0.4956	21
陕西	-0.3788	19	-0.454	17
甘肃	-0.9276	25	-0.5143	23
青海	-1.1804	28	-0.5311	25
宁夏	-1.0138	27	-0.4521	17
新疆	-0.6806	23	-0.5522	26

表 2-12 2007-2009 年 31 个省（市）综合得分及排名

地区	2007 年综合得分	排名	2008 年综合得分	排名	2009 年综合得分	排名
北京	0.9045	5	0.4251	7	0.5179	8
天津	-0.2142	15	-0.3405	19	-0.3698	21
河北	-0.2459	16	-0.0345	13	0.0567	13
山西	-0.54	25	-0.5581	24	-0.4988	24
内蒙古	-0.3444	21	-0.0862	16	-0.2204	16
辽宁	0.5148	8	0.8192	5	0.7864	5
吉林	-0.4548	22	-0.3604	20	-0.4435	23
黑龙江	-0.3085	20	-0.3823	22	-0.3661	20
上海	1.0686	3	0.6064	6	0.5354	7
江苏	1.5525	2	2.0062	1	1.9206	1
浙江	0.9694	4	0.8324	4	0.9135	4
安徽	0.0196	11	0.1788	10	0.1605	10
福建	0.1434	10	-0.057	14	-0.0302	15
江西	-0.2622	17	-0.3028	18	-0.3612	19
山东	0.8627	6	1.1248	3	1.0811	3
河南	-0.0306	12	0.3006	9	0.3499	9
湖北	-0.0577	14	-0.0575	15	0.0454	14
湖南	-0.0536	13	0.0913	12	0.0919	12
广东	1.8492	1	1.6693	2	1.6079	2
广西	-0.3088	19	-0.2683	17	-0.2987	17
海南	-0.7531	29	-0.6961	27	-0.7647	28
重庆	0.1833	9	0.0949	11	0.1507	11
四川	0.5707	7	0.4079	8	0.5374	6
贵州	-0.5318	24	-0.5873	25	-0.5672	25
云南	-0.2996	18	-0.3745	21	-0.3532	18
西藏	-0.8944	31	-0.9631	31	-0.9855	31
陕西	-0.4855	23	-0.4352	23	-0.4095	22
甘肃	-0.6904	27	-0.7237	28	-0.7588	27
青海	-0.8515	30	-0.9125	30	-0.9152	30
宁夏	-0.7277	28	-0.8086	29	-0.7843	29
新疆	-0.5841	26	-0.6082	26	-0.6281	26

## 2.9 房地产市场绩效综合评价实证结果分析

### 2.9.1 房地产市场绩效的综合得分分析

通过因子分析法得到 31 个省（市）房地产市场绩效的 2 项因子得分及综合得分，按照综合得分从高到低排名如表 2-12 所示。综合得分越高，表明其市场绩效水平越好。综合得分为零或者负值，并不代表其绩效水平为负，这里仅是将各省市的平均绩效水平作为零点。根据综合得分可以给我国各省市绩效水平层次分类，分为四类，见表



2-13。

我国房地产市场绩效跟经济水平有很大关系，沿海东部省市绩效水平靠前，经济发展落后的西部地区绩效水平相对较差。

表 2-13 我国各省市绩效水平层次分类

层次水平	省（市）	数量
绩效水平高	江苏、广东	2
绩效水平较高	山东、上海、浙江、北京、辽宁、四川、重庆、河南安徽	9
绩效水平较差	湖南、湖北、河北、内蒙古、福建、江西、广西、天津、吉林、黑龙江、云南、陕西	12
绩效水平差	山西、贵州、甘肃、新疆、宁夏、海南、青海、西藏	8

## 2.9.2 各省市房地产绩效的主因子得分分析

房地产市场开发销售因子 F1 反映了一个省市在房地产市场的开发、投入、销售等方面的水平。方差贡献率最大。2007 年为 50.574%、2008 年为 55.895%、2009 年为 54.716%。

根据 2007-2009 三年来因子 F1 得分可以把我国 31 个省市分为四个层次。如表 2-14 所示

表 2-14 我国各省市因子 F1 层次分类

层次水平	省（市）	数量
因子 F <sub>1</sub> 水平高	江苏、山东	2
因子 F <sub>1</sub> 水平较高	广东、浙江、辽宁、四川、重庆、河南、安徽、湖南、湖北	9
因子 F <sub>1</sub> 水平较低	河北、内蒙古、福建、江西、广西、吉林、黑龙江	7
因子 F <sub>1</sub> 水平低	北京、天津、上海、云南、陕西、山西、贵州、甘肃、新疆 宁夏、海南、青海、西藏	13

根据 2007-2009 三年来因子 F2 得分可以把我国 31 个省市分为三个层次。如表 2-15 所示。

因子 F2 大部分城市没有达到平均水平，都处于相对落后的水平，这跟当地的房地产企业发展水平有很大的关系。北京、上海为国际化大都市，房地产企业众多，房地产综合效益水平较高，此两个直辖市房地产业综合效益要比其他地区有优势，房地



产企业 经营收入、销售利润更多、房地产企业发展更快。而中部及西部落后城市，经济发展较为缓慢，房地产企业较少，企业综合效益水平低。导致得分相对较低。

表 2-15 我国各省市因子 F2 层次分类

层次水平	省（市）	数量
因子 F <sub>2</sub> 水平高	北京、上海	2
因子 F <sub>2</sub> 水平较高	广东、浙江、江苏、天津、福建	5
因子 F <sub>1</sub> 水平较低	海南、辽宁、湖北、四川、重庆、山东、河南、安徽、湖南、河北、内蒙古、江西、广西、吉林、黑龙江、陕西、云南、山西、贵州、甘肃、新疆、宁夏、青海、西藏	24

## 参考文献

[1] 晓群, 统计学. 现代统计分析方法与应用[M]. 中国人民大学出版社, 1998.

[2] 张丽. 基于因子分析法的基金综合绩效评价研究 [D]. 中南大学, 2007.

[3] 李斌, 蒋涛, 吴俊芳, 等. 房地产行业上市公司绩效评价的实证研究[J]. 特区经济, 2007 (3): 122-123.

[4] 陆璇. 实用多元统计分析[J]. 2001.

[5] 薛微. 统计分析与 SPSS 的应用[J]. 北京:中国人民大学出版社, 2001: 234-256.

[6] Anderson T W. An Introduction To Multivariate Stsisical Analysis[J]. 1954.

[7] 张斌, 府亚军. 我国房地产上市公司经营业绩实证研究[J]. 北京机械工业学院学报, 2006, 20(3): 62-66.

[8] Harva M, Kabán A. Variational learning for rectified factor analysis[J]. Signal Processing, 2007, 87(3): 509-527.

[9] 余有贤. 我国地方政府绩效评估指标体系构建与实际测评[D]. 浙江大学, 2005.

[10] 王家远, 袁红平. 基于因子分析法的建筑业综合评价[J]. 深圳大学学报: 理工版, 2007, 24(4): 373-378.

[11] 周刚, 孙尧, 许远明. 我国房地产业市场结构探析[J]. 重庆建筑大学学报 (社科版), 2001, 3.