doi:10.16055/j.issn.1672-058X.2019.0001.011

我国各类物价指数关系的因子模型分析*

石 凯[1,2]

(1. 乐山师范学院 数学与信息科学学院,四川 乐山 614000; 2. 西南财经大学 统计学院,成都 611130)

摘 要:针对现实经济从投资、生产到消费各领域物价水平之间存在联动关系的现象,提出运用多元统计中因子模型分析的方法,从我国现行编制的各类物价指数中提炼具有综合性的公共因子,以此反映整体物价的共同波动规律,研究结果显示:各类物价指数之间具有较强的相关性,能从中提炼出有现实意义的潜在公共因子;并且,相比较于传统的单一物价指数,用具有很强综合性的公共因子能更全面反映生产、消费、投资等领域的物价总水平的变动,从而为监测我国实体经济整体通胀水平提供科学的参考。

关键词:物价指数;因子分析;通货膨胀

中图分类号:F224.0 **文献标志码**:A **文章编号**:1672-058X(2019)01-0060-05

1 问题提出

在我国国民经济核算中,基于不同的目的编制有各类物价指数,它们在宏观经济政策调控和企业经营决策中起着非常重要的作用。根据国家统计局定期发布的《中国统计摘要》中第五部分“物价指数”显示,我国物价指数统计主要有居民消费价格指数(CPI)、商品零售价格指数(RPI)、农业生产资料价格指数(AMPI)、农产品生产价格指数(APPI)、工业生产者出厂价格指数(PPI)、工业生产者购进价格指数(IPI)、固定资产投资价格指数(FAI)等 7 类。以上 7 类物价指数分别反映从投资、工农业生产到消费的各个领域物价总水平变动,但现实经济是一个有着相互关联的整体,组成实体经济的各领域之间并非是相互独立的,那么这 7 类物价指数间是否存在共同的波动规律,以及能否综合成一个全面反映我国物价整体水平的因子,这是个值得研究分析的课题。有鉴于此,现搜集整理得 1990—2016 年我国的各类物价指标的年度数据,并对各变量进行初步的相关系数测算和检验,整理形成表 1 和表 2 的数据资料(数据来源于《中国统计摘要 2017》)。

表 1 1990—2016 年我国各类物价指数

Table 1 The price indices of China from 1990 to 2016

年 份	CPI	RPI	AMPI	APPI	PPI	IPI	FAI
1990	103.1	102.1	105.5	97.4	104.1	105.6	108.0
1991	103.4	102.9	102.9	98.0	106.2	109.1	109.5
1992	106.4	105.4	103.7	103.4	106.8	111.0	115.3
1993	114.7	113.2	114.1	113.4	124.0	135.1	126.6
1994	124.1	121.7	121.6	139.9	119.5	118.2	110.4
1995	117.1	114.8	127.4	119.9	114.9	115.3	105.9
1996	108.3	106.1	108.4	104.2	102.9	103.9	104.0
1997	102.8	100.8	99.5	95.5	99.7	101.3	101.7

收稿日期:2018-04-09;**修回日期**:2018-05-12.

* **基金项目**:中央高校博士研究生科研课题(JBK1607K02),四川省教育厅人文社科一般项目(18SB0223).

作者简介:石凯(1976—),男,四川泸州人,副教授,博士,从事应用统计研究.

续表(表1)

年　份	CPI	RPI	AMPI	APPI	PPI	IPI	FAI
1998	99.2	97.4	94.5	92.0	95.9	95.8	99.8
1999	98.6	97.0	95.8	87.8	97.6	96.7	99.6
2000	100.4	98.5	99.1	96.4	102.8	105.1	101.1
2001	100.7	99.2	99.1	103.1	98.7	99.8	100.4
2002	99.2	98.7	100.5	99.7	97.8	97.7	100.2
2003	101.2	99.9	101.4	104.4	102.3	104.8	102.2
2004	103.9	102.8	110.6	113.1	106.1	111.4	105.6
2005	101.8	100.8	108.3	101.4	104.9	108.3	101.6
2006	101.5	101.0	101.5	101.2	103.0	106.0	101.5
2007	104.8	103.8	107.7	118.5	103.1	104.4	103.9
2008	105.9	105.9	120.3	114.1	106.9	110.5	108.9
2009	99.3	98.8	97.5	97.6	94.6	92.1	97.6
2010	103.3	103.1	102.9	110.9	105.5	109.6	103.6
2011	105.4	104.9	111.3	116.5	106.0	109.1	106.6
2012	102.6	102.0	105.6	102.7	98.3	98.2	101.1
2013	102.6	101.4	101.4	103.2	98.1	98.0	100.3
2014	102.0	101.0	99.1	99.8	98.1	97.8	100.5
2015	101.4	100.1	100.4	101.7	94.8	93.9	98.2
2016	102.0	100.7	100.1	103.4	98.6	98.0	99.4

表 2　各类物价指数之间的相关系数

Table 2　Correlation coefficient between the price indices

指　数	CPI	RPI	AMPI	APPI	PPI	IPI	FAI
CPI	1						
RPI	0.995**	1					
AMPI	0.843**	0.865**	1				
APPI	0.837**	0.860**	0.824**	1			
PPI	0.868**	0.880**	0.797**	0.720**	1		
IPI	0.746**	0.766**	0.715**	0.620**	0.972**	1	
FAI	0.646**	0.671**	0.557**	0.452*	0.854**	0.900**	1

注:“**”表示在 1% 的水平下显著,“*”表示在 5% 的水平下显著.

将表 1 和表 2 结合起来分析,可以得到大量的信息。首先,从表 1 可见,各类物价指数的变动方向基本一致,即表现出同时上升或者下降。例如,1992—1994 年,各类指数均呈现出较大幅度的涨幅;1998—1999 年,各类物价指数都小于 1,表现出通缩的态势;2007—2008 年,经济过热,各类物价指数纷纷上扬,而 2009 年,受全球金融危机的影响,物价指数整体明显下跌,2011 年又整体上涨。表 2 则是在表 1 数据的基础上计算变量间的相关系数,并对相关系数进行显著性检验。原假设为总体相关系数为 0,所采用的统计量为 $t=r\sqrt{n-2}/\sqrt{1-r^2}$,在原假设成立的条件下,服从自由度为($n-2$)的 t 分布。表 2 显示,各物价指数基本都在 1% 的水平下表现出显著不为 0 的相关关系,农产品生产价格指数(APPI)和固定资产投资价格指数(FAI)也在 5% 的水平下表现出显著不为 0 的相关关系。而且各个变量间的相关系数均为大于 0 的正相关关系,有不少的相关系数数值高于 0.8,表现出高度相关的密切关系。这进一步验证了从表 1 中观察到的结果,即各类物价指数有同升同降的特征。

其次,各类物价指数编制的目的不一样,反映的内容不同,变动幅度势必会存在着差异性。例如,PPI 反映生产环节价格水平,而 CPI 反映消费环节的价格水平,整体价格水平的波动一般首先

出现在生产领域,然后通过产业链向下游产业扩散,最后波及消费品。同时,由于 CPI 不仅包括消费品价格,还包括服务价格,CPI 与 PPI 在统计口径上并非严格的对应关系,因此 CPI 与 PPI 的变化出现差异是必然的。同样,虽然各类物价指数之间彼此联动的现象比较明显,但是由于含义、构成和用途各有不同,因此变动幅度并非总是保持一致,有时会出现走势差异乃至背离(从表 1 中有些年份也能找出此类情况)。

基于上述分析,如果能够在各类不尽相同的物价指数中找出其变动的潜在共同因子,则对我国通胀水平的监测和预警有着重要的参考价值。因为该共同因子综合了生产、消费等各个领域的价格变动,从而更能全面监测我国经济运行的整体态势。

2 因子分析模型与求解

因子分析是通过研究众多变量之间的内部依赖关系来探析数据的基本结构,它的基本思想是将一些有相关性的变量由少数几个潜在的不相关的共同因子表现出来,而不被共同因子包含的部分(即每个变量自己的个性特征)则由特殊因子表达出来。据此,建立模型如下:

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_7 \end{bmatrix} = \begin{bmatrix} \alpha_{11}\alpha_{12}\cdots\alpha_{1m} \\ \alpha_{21}\alpha_{22}\cdots\alpha_{2m} \\ \vdots \quad \vdots \qquad \vdots \\ \alpha_{71}\alpha_{72}\cdots\alpha_{7m} \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_7 \end{bmatrix}$$

或者简写为矩阵形式:$\boldsymbol{X}=\boldsymbol{AF}+\boldsymbol{\varepsilon}$。其中,$X_1,X_2,\cdots,X_7$ 为各类物价指数,$F_1,F_2,\cdots,F_m$ 为公共因子,$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_7$ 是特殊因子,是不能被前 m 个公共因子包含的部分,$\alpha_{ij}(i=1,2,\cdots,7;j=1,2,\cdots,m)$ 为因子载荷。对变量进行标准化处理(即减去均值除以标准差)后满足以下 3 个条件:$\boldsymbol{F},\boldsymbol{\varepsilon}$ 不相关;$F_1,F_2,\cdots,F_m$ 互不相关,方差均为 1;$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_7$ 互不相关。

因子分析模型的关键是求解因子载荷矩阵 $\boldsymbol{A}$。估计 $\boldsymbol{A}$ 的一般方法有主成分法、主因子法、极大似然估计法等,现采用最常用的主成分方法。

对 $\boldsymbol{X}=\boldsymbol{AF}+\boldsymbol{\varepsilon}$ 两边求方差,则有:

$\mathrm{Var}(\boldsymbol{X})=\boldsymbol{A}\mathrm{Var}(\boldsymbol{F})\boldsymbol{A}'+\mathrm{Var}(\boldsymbol{\varepsilon})$

又因为 $\mathrm{Var}(\boldsymbol{F})=\boldsymbol{I}$,$\mathrm{Var}(\boldsymbol{X})=\sum_X$,可写为

$\sum_X=\boldsymbol{AA}'+\mathrm{Var}(\boldsymbol{\varepsilon})$

其中,$\mathrm{Var}(\boldsymbol{\varepsilon})=\mathrm{diag}(\sigma_1^2,\sigma_2^2,\cdots,\sigma_7^2)$,若 $\mathrm{Var}(\boldsymbol{\varepsilon})$ 主对角线上的元素数值越小,则说明公共因子共享的成分越多。因子分析的目的在于希望找到的公共因子能占有尽量大的共享成分,因此假定特殊因子的方差为 0,即 $\sum_X=\boldsymbol{AA}'$。又因为 $\sum_X$ 为实对称矩阵,则总能通过正交变化为

$$\sum_X=\boldsymbol{U}\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_7 \end{bmatrix}\boldsymbol{U}'$$

$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_7$ 为 $\sum_X$ 的特征根,$\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_7$ 为对应的标准化向量。所以也可写成:

$$\sum_X=\lambda_1\boldsymbol{u}_1\boldsymbol{u}'_1+\lambda_2\boldsymbol{u}_2\boldsymbol{u}'_2+\cdots+\lambda_7\boldsymbol{u}_7\boldsymbol{u}'_7=$$

$$[\sqrt{\lambda_1}\boldsymbol{u}_1,\sqrt{\lambda_2}\boldsymbol{u}_2,\cdots,\sqrt{\lambda_7}\boldsymbol{u}_7]\begin{bmatrix} \sqrt{\lambda_1}\boldsymbol{u}'_1 \\ \sqrt{\lambda_2}\boldsymbol{u}'_2 \\ \vdots \\ \sqrt{\lambda_7}\boldsymbol{u}'_7 \end{bmatrix}=$$

$$\boldsymbol{U}\begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_7} \end{bmatrix}\begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_7} \end{bmatrix}\boldsymbol{U}'$$

又因为 $\sum_X=\boldsymbol{AA}'$,似乎 $\boldsymbol{A}$ 已经求解出来。但是,因子分析的目的是要用少数几个重要的公共因子来解释变量,所以还应去掉特征根小的,它们起到的解释作用较少。若认为前 m 个特征根较大,则:

$$\sum_X=\boldsymbol{AA}'\approx\lambda_1\boldsymbol{u}_1\boldsymbol{u}'_1+\lambda_2\boldsymbol{u}_2\boldsymbol{u}'_2+\cdots+\lambda_m\boldsymbol{u}_m\boldsymbol{u}'_m=$$

$$[\sqrt{\lambda_1}\boldsymbol{u}_1,\sqrt{\lambda_2}\boldsymbol{u}_2,\cdots,\sqrt{\lambda_m}\boldsymbol{u}_m]\begin{bmatrix} \sqrt{\lambda_1}\boldsymbol{u}'_1 \\ \sqrt{\lambda_2}\boldsymbol{u}'_2 \\ \vdots \\ \sqrt{\lambda_m}\boldsymbol{u}'_m \end{bmatrix}$$

以上即是用主成分法求解因子载荷矩阵的过程,$\boldsymbol{A}=[\sqrt{\lambda_1}\boldsymbol{u}_1,\sqrt{\lambda_2}\boldsymbol{u}_2,\cdots,\sqrt{\lambda_m}\boldsymbol{u}_m]$。现在代入 7 个物价指数变量的数据,标准化处理后可用表 2 的相关系数矩阵进行计算特征根,得 $\lambda_1=5.705$,$\lambda_2=0.831$,$\lambda_3=0.201$,$\lambda_4=0.159$,$\lambda_5=0.097$,$\lambda_6=0.005$,$\lambda_7=0.003$。根据一般筛选原则,选择特征根大于 1,所以选择 $\lambda_1=5.705$,而且可见 λ_1 明显高于其他特征根,符合要寻求的最大解释的目的。λ_1 所对应的标准化的特征向量为

$\boldsymbol{u}_1=(0.395,0.402,0.372,0.353,0.404,0.379,0.335)$

求解因子载荷矩阵:

$\boldsymbol{A}=\sqrt{\lambda_1}\boldsymbol{u}_1=\sqrt{5.705}\boldsymbol{u}_1=$

$(0.944,0.959,0.889,0.843,0.965,0.905,0.801)$

即(变量首先进行了标准化处理，即减去均值除以标准差，用 x_i^* 表示)：

$$x_1^* = 0.944F, x_2^* = 0.959F, x_3^* = 0.889F,$$

$$x_4^* = 0.843F, x_5^* = 0.965F, x_6^* = 0.905F, x_7^* = 0.801F$$

F 即为综合物价指数因子，对 $\boldsymbol{X}$ 的贡献为5.705。在所有的特征根中，该因子的方差贡献率达到了81.50%，说明了原始变量间(即各类物价指数之间)有非常强的相关性，因此可以从中提取出具有很强解释意义的公共因子 F。

3　因子得分的计算

上述因子分析中，从我国的 7 个物价指数中综合出一个非常有意义的公共因子 F。但要将公共因子用于实际，来判断我国各个年份总体物价的涨跌程度，还需要将公共因子进行测度量化，从而为判断宏观经济整体的通胀情况提供重要的参考。将 F 进行量化，则要计算因子得分，即把公共因子反过来表达为原变量的线性组合：$F=\beta_1 x_1+\beta_2 x_2+L+\beta_7 x_7$。

由于 $\boldsymbol{X}=\boldsymbol{AF}$ 中，$\boldsymbol{A}$ 并非方阵，所以不能由求逆阵的方法来求解(即 $F=\boldsymbol{A}^{-1}\boldsymbol{X}$)，只能通过估计。运用巴特莱特因子得分(也称为加权最小二乘)方法来进行估计：

若 $\boldsymbol{X}=\boldsymbol{AF}+\boldsymbol{\varepsilon}$，假定因子载荷矩阵 $\boldsymbol{A}$ 和特殊因子方差已知，$\mathrm{Var}(\varepsilon_i)=\sigma_i^2(i=1,2,\cdots,7)$，令 $\mathrm{Var}(\boldsymbol{\varepsilon})=\boldsymbol{D}$，将 $\boldsymbol{\varepsilon}$ 看作误差项，用加权最小二乘的方法令其最小，从而估计公因子 F 的值。即：

$$\sum_{i=1}^{7}\frac{\boldsymbol{\varepsilon}_i^2}{\boldsymbol{\sigma}_i^2}=\boldsymbol{\varepsilon}'\boldsymbol{D}^{-1}\boldsymbol{\varepsilon}=(\boldsymbol{X}-\boldsymbol{A}F)'\boldsymbol{D}^{-1}(\boldsymbol{X}-\boldsymbol{A}F)\overset{\mathrm{def}}{=\!=}\pi(F)$$

求 F 的估计值 $\hat{F}$，使 $\pi(F)$ 最小。将 $\pi(F)$ 对 F 求一阶导并令其等于 0 即可得 $\hat{F}$。

$$\hat{F}=(\boldsymbol{A}'\boldsymbol{D}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{D}^{-1}\boldsymbol{X}$$

可得：

$$\hat{F}=0.166x_1^*+0.168x_2^*+0.157x_3^*+0.148x_4^*+0.170x_5^*+0.158x_6^*+0.140x_7^*$$

代入各个年份的数据计算相应的因子得分：

表 3　各类物价指数的综合因子得分

Table 3　The factor score of all kinds of price indices

年份	F	年份	F	年份	F	年份	F	年份	F
1990	−0.054	1996	0.217	2002	−0.801	2008	0.837	2014	−0.662
1991	0.083	1997	−0.570	2003	−0.346	2009	−1.118	2015	−0.855
1992	0.516	1998	−1.154	2004	0.405	2010	0.125	2016	−0.611
1993	2.427	1999	−1.164	2005	−0.098	2011	0.552		
1994	2.694	2000	−0.575	2006	−0.325	2012	−0.422		
1995	1.854	2001	−0.660	2007	0.246	2013	−0.542		

4　结论与建议

通过上述因子模型分析，并结合表 1 的实际数据以及表 3 计算的综合因子得分，可见当 $F>0$ 时，说明各类物价指数综合来看基本都在历年平均水平之上，整体经济中显现出通胀趋势，人们已经明显感觉到物价水平的上涨。F 数值越大，通胀程度越明显，整体经济出现的问题越严重，需要采用越紧缩的政策进行治理。反之，则呈现通缩趋势，明显低于平均水 平时，就应采取宽松的政策进行治理。

我国 1990—2016 年间，F 明显大于 0 的年份有：1992—1996 年，2004 年，2007—2008 年，2011 年。结合这些年份的经济背景来看，均表现出普遍物价上涨，经济过热的情况。可见所构建的综合因子能有效对宏观经济的通胀现象进行检测，当 F 呈现出大于 0 的趋势时，说明整体经济中已经显现出全面物价上涨的状态。相比较下，若根据单一指数对通胀进行诊断，有时各个指数涨跌程度不同，甚至涨跌方向也不一样，从而不能得到准确的判断。运用因子分析方法，计算综合因子得分，能综合生产、消费、投资等各个环节以及能综合各次产业中物价水平的变动，因而能更全面反映国民经济的整体运行态势，更有利于对通胀水平进行全面有效的监测和预警，为宏观决策提供更科学的依据。

参考文献(References):

[1] JOHNSON R A, WICHERN D W. Applied Multivariate Statistical Analysis (6th Edition) [M]. New York: Pearson Education, 2007

[2] ANGA, PIAZZESI M. No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables[J]. Journal of Monetary Economics, 2003, 50(4): 745—787

[3] 高惠璇. 应用多元统计分析[M]. 北京:北京大学出版社,2005
GAO H X. Application of Multivariate Statistical Analysis [M]. Beijing: Peking University Press, 2005

[4] 何晓群. 多元统计分析(第四版)[M]. 北京:中国人民大学出版社,2015
HE X Q. Multivariate Statistical Analysis (4th Edition) [M]. Beijing: People's University Publication House, 2005

[5] 林海明. 因子分析精确模型的基本思想与方法[J]. 统计与信息论坛,2006(5):23—25
LIN H M. The Precise Model of Factor Analysis and Its Thought[J]. Statistics & Information Forum, 2006(5): 23—25

[6] 林海明. 因子分析模型的改进与应用[J]. 数理统计与管理,2009 (6):998—1012
LIN H M. The Improvement and Application of Factor Analysis Model [J]. Application of Statistics and Management, 2009(6): 998—1012

[7] 林海明. 因子分析应用中一些常见问题的解析[J]. 统计与决策,2012(15):65—69
LIN H M. Analysis of Some Common Problems in the Application of Factor Analysis [J]. Statistics and Decision, 2012(15): 65—69

[8] 林海明,王翊. 因子分析模型 L 及其解是更好的[J]. 统计研究,2007(8):77—83
LIN H M, WANG Y. Factor Analysis Model L and Improved Solution of Factor Analysis [J]. Statistical Research, 2007(8): 77—83

[9] 汤丹,赵昕东. 核心通货膨胀的度量方法及其应用[J]. 宏观经济研究,2011(7):28—34
TANG D, ZHAO X D. The Measurement Method of Core Inflation and Its Application [J]. Studies on Macroeconomics, 2011(7): 28—34

[10] 杨超. 通货膨胀的测度方法及实证[J]. 统计与决策,2014(8):85—87
YANG C. Measure Method and Empirical Research of Inflation[J]. Statistics and Decision, 2014(8): 85—87

[11] 何光辉. 中国 CPI 与 PPI 的结构与动态作用机制研究[J]. 经济科学,2009(4):15—30
HE G H. Research on the Structure and Dynamic Mechanism of CPI and PPI in China [J]. Economic Science, 2009(4): 15—30

Factor Analysis Model on the Relationship of Price Indexes in China

SHI Kai[1,2]

(1. College of Mathematics and Information Science, Leshan Normal College, Sichuan Leshan 614000, China;
2. School of Statistics, Southwest University of Finance and Economics, Chengdu 611130, China)

Abstract: In view of the relationship between the price level of the real economy from investment, production to consumption, this paper used the method of factor analysis model to extract comprehensive public factors from all kinds of price indices in China, and to reflect the overall price level through this factor. The results show that there is a strong correlation between the various price indexes, so the common factor is effective. At the same time, compared with the single price index, the common factor can more effectively and comprehensively reflect the changes of price in the fields of production, consumption and investment. By this common factor, we can provide more scientific reference for the level of inflation.

Key words: price index; factor analysis; inflation

责任编辑:罗姗姗

引用本文/Cite this paper:
石凯. 我国各类物价指数关系的因子模型分析 [J]. 重庆工商大学学报(自然科学版),2019,36(1):60—64
SHI K. Factor Analysis Model on the Relationship of Price Indexes in China[J]. Journal of Chongqing Technology and Business University(Natural Science Edition), 2019, 36(1): 60—64