第 25 卷 第 1 期　　应 用 数 学 学 报　　Vol. 25 No. 1
2002 年 1 月　　ACTA MATHEMATICAE APPLICATAE SINICA　　Jan., 2002

模糊时间数列的分析与预测: 以台湾地区加权股价指数为例

吴柏林　　林玉钧

(台湾政治大学应用数学系, 台湾)

摘 要 随着我国经济快速成长, 衍生性金融商品的投资分析, 已成为国内财务数学研究热门课题. 以股票市场而言, 人们总希望比别人早一步掌握行情的脉动, 以获取最高的报酬率. 然而, 影响股市加权股价指数波动的因素众多, 要如何进行趋势分析与预测, 是很多学者相当感兴趣与研究的主题. 本文考虑以模糊统计方法, 作模糊时间数列的趋势分析与预测. 期望应用模糊统计分析方法比传统的时间数列分析方法能得到更合理的解释, 且预测结果可以提供决策者更多的信息, 做出正确的决策. 最后以台湾地区加权股票指数为例, 做一实证上的详细探讨.

关键词 模糊时间数列, 预测, 台湾加权股价指数

1 前言

台湾股票市场自 1986 年 10 月突破一千点, 开始快速攀升, 至 1990 年 2 月 10 日达到 12495.34 点的历史性高点, 但于同年 10 月 1 日回跌至 2560 点, 随后于次年 5 月 9 日反弹至 6305 点, 直到近年多在六、七千点附近游走. 在此短时间内暴涨暴跌的特性, 使台湾股票市场被国际公认为是一个充满投机色彩的大赌场, 并使一般正常的投资人视股票市场为畏途. 长此下去, 台湾证券市场前途不容乐观.

证券市场交易的热络, 虽然诱发较多投资机会, 但也伴随而来不可忽视的风险. 面对诡局多变的证券市场, 投资人应如何做理性的抉择是件重要的事. 所以, 培养一般投资人对于股票市场的信心与兴趣可以说是第一要物. 由于资料收集的误差、时间的迟延 (time lag). 使得单一度量的数值, 形式上看似精确, 但它实际上所表达的应该是某一时间点的可能值. 例如, 加权股票指数在收盘时的指数, 并不能完全表示当天的股票指数. 是故, 加权股票指数这个数字的本身就具有不确定性与模糊性. 在此情况下, 加权股票指数的模糊特性, 在经济的实证研究过程中, 应予以慎重考虑.

模糊理论 (Fuzzy theory) 本身具有语言变量 (linguistic variables) 蕴含特性, 这种特性可以减少在处理不确定性问题时可能造成的困扰. 因此, 模糊理论日前已被广泛地应用于各种领域. 例如航空、机械、医学、电力、地质等等, 近来也逐渐应用在社会科学方面. 而 Clymer, Corey 与 Gardner[1] 提出离散事件的机场起落模糊控制方法, Cutsem 与 Gath[2] 曾提出利用模糊分类的程序来侦测离群值及稳健地估计参数, Hathaway 与 Bezdek[3] 则对模糊回归模式进行参数估计与分类. Yoshinari, Pedrycz 与 Hirota[4] 经由模糊分类方法建立模糊模式, Romer 与 Kandel[5] 在统计推论上采用模糊分割的理论及可能性理论, Wu 与 Hung[6] 提出模糊认定法则, 以作为 ARCH 模式族与 Billinear 模式族的决策判定过程, Wu 与 Chen[7] 利用模糊分类法来检定时间数列数据结构转变的

本文 1999 年 8 月 27 日收到. 2000 年 1 月 2 日收到修改稿.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

转折区间. 近年来, 在人文社会科学的测度理论里, 模糊统计与模糊相关性日渐受到重视, 这应是复杂的人文社会现象无法以传统数值模型充分合理解释的一种自然发展结果. 有鉴于此, 本文乃尝试以模糊时间数列来分析台湾地区加权股票指数, 并考虑以模糊关系矩阵, 来认定台湾地区加权股票指数的走势. 这对于了解台湾的经济波动将深具意义.

2 模糊时间数列的建构

2.1 模糊统计分析

我们介绍一个模糊模式的建构方式, 从经济或社会问题的立场来分析人类思考与行为的模式. 此方法称为模糊逻辑模式信念 (the fuzzy logic model of belief). 模糊逻辑的应用主张个人的喜好程度不需要非常清晰或是条理井然. 这种概念与布尔逻辑理念刚好相对力, 例如依照布尔逻辑中 A 和 B 的比较, 其结果只有三种可能: (1) $A > B$; (2) $A < B$; (3) $A = B$. 但是, 人类思维的运作远比布尔逻辑的结果来得复杂, 尤其是在人的思维中具有许多不明确的偏好. 因此, 如何掌握真实状况作全盘描绘的构想, 早已成为许多逻辑学家的努力目标. 对我们人类而言, 凭借模糊模式的呈现方式要比直接指定单一物体的特定值来得适切些, 此一概念也比较合于用来评估物体与物体间的相关特性. 此外, 由于其它特性往往也会有助于某一意见的评定, 因此使用模糊逻辑时, 必须对所谓的 "其它特性" 加以说明, 以便将人们的喜好程度转换成便于计算的效用函数 (utility function).

模糊逻辑弥补了布尔逻辑的不足, 但是并没有完全取代布尔逻辑, 两个方法基本上是可以互存的. 关键在于每一理论各自描述不同理念. 模糊子集理论的主要特质, 在于对于无限的可能偏好, 能做出无限的解说. 因此, 使用模糊逻辑来测度, 需要使用模糊统计来代替传统统计, 而模糊逻辑测度主要根据模糊集合的基本理论. 有关模糊集合的基本理论, 在此不做详细介绍, 有兴趣的读者可参考 [8,9].

在传统的集合理论中, 一元素要就属于某集合, 要就不属于该集合. 但是在模糊集合中的元素, 其隶属程度可能只有部分是属于该集合. 但是在模糊集合中的元素, 其隶属程度可能只有部分是属于某集合. 例如: "年轻人" 此一名词, 究竟几岁以上才算得是 "不再年轻"? 在一般人们的印象中并没有一个确切的界限来区分 "年轻" 和 "不再年轻". 在模糊集合的定义中, 可以显示出一个 20 岁的人是 " 90% 的年轻", 而某一个 60 岁的人相对的只有 " 30隶属度 (membership) 的全距通常设定在 0 到 1 之间. 因此, 每个语言变项, "年轻", 代表一个可能性的分布. 此一分布的平均值用以表示人类真实评定是 "年轻人" 的数值. 而关于分布的评定, 可能因人而异, 且不需呈现常态.

传统的社会和经济研究, 投入了很多有关人类的互动关系及模式分析. 在一个典型的模式建构中, 经常面对一些不确定的例子, 例如: 每年的学生注册人数, 以年初? 年中或年尾为准? 期间所得数值往往各有不同. 又如新台币对美元的汇率, 是以开盘? 收盘? 或是最高最低价之平均为准? 结果亦有相当的差距. Hendershot, Placek[10] 曾对此一领域的文献做一广泛的回顾. 在社会科学及有关经济领域的研究中, 问题的答案很少是确切的真或伪. 如果我们尝试着去分析人类的信念, 将会发现必须去面对行为的不确定性. 而模糊集合中连续区间值具有能处理亦真亦伪状况的能力, 因此应用区间值的模糊特性所做的分析, 使研究者可以处理有关的不确定, 在实际的运用上, 的确是比较符合实际状况的一种测量工具.

2.2 模糊时间数列分析

在讨论应用模糊时间数列来进行预测时, 首先, 要给定几个模糊时间数列的定义.

定义 2.1 模糊时间数列

令 $\{X(t) \in R,\ t = 1,2,\cdots,n\}$ 为一个时间数列, U 为其论域. 给定 U 的一个次序分割集合 (ordered partition set), $\{P_i,\ i = 1,2,\cdots,r,\ \bigcup_{i=1}^{r} P_i = U\}$ 其相对于语言变量

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

为 $\{L_i,\ i=1,2,\cdots,r\}$. 若在 $\{L_i,\ i=1,2,\cdots,r\}$ 上相对于 $X(t)$ 的模糊集合 $F(t)$ 有隶属度函数为 $\{\mu_1(X(t)),\mu_2(X(t)),\cdots,\mu_r(X(t))\}$, $0\le\mu_i\le 1$, $i=1,2,\cdots,r$, 则我们称 $\{F(t)\}$, $F(t)$ 的集合为布于 $X(t)$ 上的一个模糊时间数列. 并且记为

$$F(t)=\frac{\mu_1(X(t))}{L_1}+\frac{\mu_2(X(t))}{L_2}+\cdots+\frac{\mu_r(X(t))}{L_r},\tag{2.1}$$

其中 $\frac{\mu_i(X(t))}{L_i}$ 表示 $X(t)$ 相对于语言变量 L_i 及其隶属度 $\mu_i(X(t))$ 的对应关系, "+" 表示连结符号, $\mu_i:R\to[0,1]$, 且 $\sum\limits_{i=1}^{r}\mu_i(X(t))=1$, 对所有 $t=1,2,\cdots,n$. 为了方便起见, 以后我们将 $F(t)$ 简写成 $F(t)=(\mu_1,\mu_2,\cdots,\mu_r)$.

在计算模糊隶属度时, 一般根据三角形隶属度函数、梯形隶属度函数、钟形隶属度函数. 本文考虑以三角形隶属度函数, 进行时间数列转换为模糊时间数列的程序, 其方法如下:

令 $\{X_t\in R,\ t=1,2,\cdots,n\}$ 为一个时间数列, $\{P_i,\ i=1,2,\cdots,r,\ \bigcup\limits_{i=1}^{r}P_i=U\}$ 为论域 U 的一个次序分割集合.

设 m_i 为其分割集合的中间值, $i=1,2,\cdots,r$, 且其相对于语言变量为 $\{L_i,\ i=1,2,\cdots,r\}$, 则对任意 t, 若 X_t 介于 m_i 与 m_{i+1} 之间, 则其属于语言变量 L_i 的隶属度为 $\frac{m_{i+1}-X_t}{m_{i+1}-m_i}$, 属于语言变量 L_{i+1} 的隶属度为 $\frac{X_t-m_i}{m_{i+1}-m_i}$, 属于其它之语言变量 $\{L_i,\ i=1,2,\cdots,i-1,i+2,\cdots,r\}$ 的隶属度皆为 0.

例 2.1 令时间数列 $\{X_t\}=\{0.8,1.7,2.9,4.1,3.5,3.2,4.3,3.6\}$, 表示某股票连续 8 天之涨跌值. 观察其全距我们选择一个次序分割集合, 例如 $U=\{(0,1],(1,2],(2,3],(3,4],(4,5]\}$, 且其语言变量为: 很低 $=L_1\alpha(0,1]$, 低 $=L_2\alpha(1,2]$, 中等 $=L_3\alpha(2,3]$, 高 $=L_4\alpha(3,4]$, 很高 $=L_5\alpha(4,5]$, 其中 "α" 表 "相对于", 其分割集合的中间值为 $m_1=0.5$, $m_2=1.5$, $m_3=2.5$, $m_4=3.5$, $m_5=4.5$, 且其相对于语言变量为 $\{L_i,\ i=1,2,\cdots,5\}$. 因为 X_1 落于 (0,1] 与 (1,2] 之间, 则可得相对于 X_1 的模糊时间数列 $F_1=(0.7,0.3,0,0,0)$

$$\frac{1.5-0.8}{1.5-0.5}=0.7\in L_1,\qquad \frac{0.8-0.5}{1.5-0.5}=0.3\in L_2.$$

故可得相对于 $\{X_t\}$ 的模糊时间数列 $\{F_t\}$ 如下:

	很低	低	中等	高	很高
F_1	0.7	0.3	0	0	0
F_2	0	0.8	0.2	0	0
F_3	0	0	0.6	0.4	0
F_4	0	0	0	0.4	0.6
F_5	0	0	0	1	0
F_6	0	0	0.3	0.7	0
F_7	0	0	0	0.2	0.8
F_8	0	0	0	0.9	0.1

在时间数列分析过程中, 自相关值的计算相当重要. 对模糊时间数列而言, 模糊关系及模糊关系矩阵, 亦是一主要探讨重点. 我们定义模糊关系及模糊关系矩阵如下:

定义 2.2 模糊关系

设 $\{P_1,P_2,\cdots,P_r\}$ 为论域 U 的一个次序分割. 若在 U 上的模糊集 F 和 G 分别为 $F=(\mu_1,\mu_2,\cdots,\mu_r)$, $G=(\nu_1,\nu_2,\cdots,\nu_r)$, 其中 μ_i 和 ν_i 为布于 U 上的隶属度, 则介于 F 和 G 之间的模糊关系记为

$$R=F^t\circ G=[R_{ij}]_{m\times m},\tag{2.2}$$

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

其中 '$\circ$' 表示模糊合成运算, 在此我们使用最大 - 最小 (max-min), '$A\#t$' 表示转置, R_{ij} 表介于 F 和 G 之间关系的隶属度.

定义 2.3 设 $R(t,t-1)$ 是 FAR(1) (fuzzy autoregression of order one) 的模糊关系矩阵. 若对任何时点 t, $R(t,t-1)$ 与 t 无关, 亦即对任何 t,

$$R(t,t-1)=R(t-1,t-2), \tag{2.3}$$

则称 FAR(1) 具有非时变的 (time-invariant) 性质, 否则称为具时变的 (time-variant) 性质.

为了方便分析起见, 这里只探讨一阶的模糊时间数列关系矩阵.

定义 2.4 设 $F(t)$ 为一 FAR(1) 之非时变模糊时间数列. 若对任何时点 t, 恒有 $F(t)=F(t-1)$ 且 $F(t)$ 是由有限个 $f_{ik}(t)$, $k=1,2,\cdots,m$ 所组成, 则

$$\begin{aligned} R(t,t-1)=&f_{i1}(t-1)\otimes f_{j1}(t)Vf_{i2}(t-2)\\ &\otimes f_{j2}(t-1)V\cdots Vf_{im}(t-m)f_{jm}(t-m+1), \end{aligned} \tag{2.4}$$

其中 I,j 为对应的语言程度指针 (index of linguistic hedge), $m>0$, V 是一个取最大的运算.

2.3 模糊数值转换过程及预测

如何将模糊数值转换成语言变量值并进行预测, 是本文研究的重点. 根据模糊逻辑的特性及非线性时间数列的不稳定走势, 我们考虑以 (1) 最大隶属度法 (2) 隶属度加权平均法, 来做模糊数值之转换.

最大隶属度法

设 $w_1,w_2,\cdots,w_n$ 为论域 U 上输出元素的隶属度. 令 $w_0=\max\{w_1,w_2,\cdots,w_n\}$, 我们选择此 w_0, 即输出元素的最大隶属度, 其所在区间的中点为输出值. 当输出的隶属度有两个以上最大隶属度时, 则取各最大隶属度所在区间中点之平均为输出值.

隶属度加权平均法

设 $w_1,w_2,\cdots,w_n$ 为论域 U 上输出元素的隶属度. 我们考虑以标准化的隶属度加权数, 对各组中点做平均为其输出值. 亦即输出值 O

$$O=S_1*M_1+\cdots+S_n*M_n, \tag{2.5}$$

其中 $S_i=w_i/\{\sum w_i\}$, 为标准化后的隶属度, M_i 为指数分区的组中点, $i=1,2,\cdots,n$.

例 2.2 设论域 U 上输出元素的隶属度为 $\{0.2,0.3,0.6,0.1,0.4\}$, 其组区域为 $\{(-2,-1),(-1,0),(0,1),(1,2),(2,3)\}$, 组中点值 m_i 分别为 $\{-1.5,-0.5,0.5,1.5,2.5\}$, 则根据 (1) 最大隶属度法, 来做模糊数值转换. 因为 $0.6=\max\{0.2,0.3,0.6,0.1,0.4\}$ 而所在的组区域 (0,1) 之组中点值为 0.5. 故输出值为 0.5, 若根据 (2) 隶属度加权平均法, 来做模糊数值之转换, 得到 $\{0.2/1.6,0.3/1.6,0.6/1.6,0.1/1.6,0.4/1.6\}$, 则输出值为 $1.1875=-1.5*0.2/1.6-0.5*0.3/1.6+0.5*0.6/1.6+1.5*0.1/1.6+2.5*0.4/1.6$.

2.4 模糊时间数列模式建构步骤与流程

根据上一节的定义, 将进一步说明模糊时间数列建立流程与方法, 基本上, 模糊时间数列模式需要依据下列六个步骤, 如图 2.1.

步骤 1. 架构论域的模糊集合.

步骤 2. 资料模糊化.

步骤 3. 计算模糊相关矩阵.

步骤 4. 建构合适的模糊时间数列模式.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

步骤 5. 运算后的模糊时间数列模式来做预测.

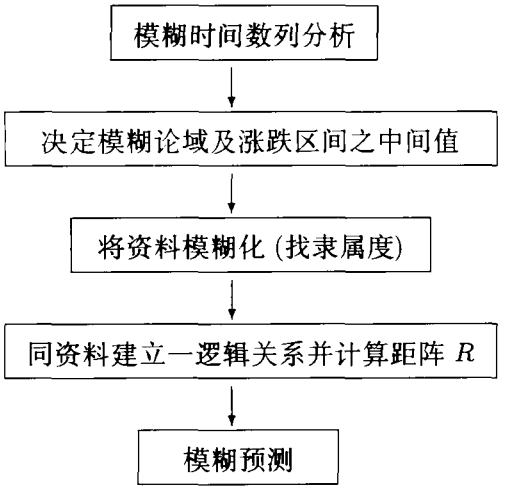

图 2.1 模糊时间数列建立的流程图

3 台湾股票加权指数之模糊时间数列

在社会科学的测度理念上，模糊统计和模糊相关性的使用可说是一种数字模式的推广. 单一数值性质的资料，主要限制是具有过度解释的潜在危机. 而使用比较模糊且朴实的语言资料，则可以避免这些无法预料的危险. 事实上，利用语言模式作为预测，可以发现每一期皆增加了结果的模糊性. 从一般的生活中这种模糊化似乎也是很正常的现象，但是从另一角度来看，如果数值处理的观念没有改变，预测方法没有突破，往往却阻碍了长期预测的可能性. 以模糊时间数列来预测台湾的股票加权指数，在建构模糊时间数列时，必需注意要让历史资料达到稳定. 亦即，当资料不是趋于稳定时，利用时间数列分析的方法，我们需将原始资料差分，使资料达到稳定的情况，再继续我们的步骤.

3.1 资料分析

本资料来自日盛证券投资公司，其原始资料来自台湾证券交易所. 共收集自 1999 年 4 月 1 日至 1999 年 4 月 30 日的每日加权股价开盘指数，共 23 笔资料. 其走势图绘于图 3.1.

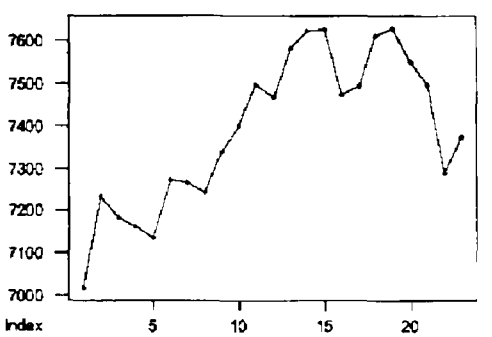

图 3.1 台湾加权股价日指数走势图 (1/4/1999~30/4/1999)

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

根据时间数列分析的方法, 其趋势为不稳定, 故我们先将原始资料做一次差分, 再观察其是否满足稳定型要求. 如图 3.2

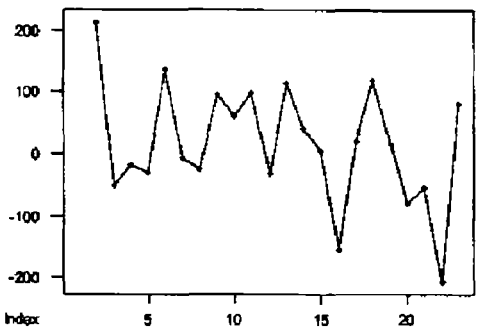

图 3.2 差分后的台湾加权股价日指数走势图 (1/4/1999~30/4/1999)

由差分后的资料可以看出, 每日涨幅的最大值为 214, 每周跌幅的最小值为 -207. 通常, 论域要包括最大值与最小值, 所以选择了集合 $(-210, 220)$. 接下来的步骤是将集合模糊化. 首先, 将 $(-210, 220)$ 均分成 k 个区间 (本文我们令 $k = 5$) 亦即, $I_1 = (-210, -124)$, 其中间值为 -167, $I_2 = (-124, -38)$, 其中间值为 -81, $I_3 = (-38, 48)$, 其中间值为 5, $I_4 = (48, 134)$, 其中间值为 91, $I_5 = (134, 220)$, 其中间值为 177. 然后, 在 $(-210, 220)$ 之中定义 5 个模糊集. 令 F_1 ="大跌", F_2 ="跌", F_3 ="平盘", F_4 =q2"涨", F_5 ="大涨", 其中每一个模糊集的元素为 I_1, I_2, I_3, I_4 及 I_5 与其相对应的隶属度所组成. 这些集合表示如下:

$$
\begin{aligned}
F_1 &= \{1/I_1, 0.5/I_2, 0/I_3, 0/I_4, 0/I_5\},\\
F_2 &= \{0.5/I_1, 1/I_2, 0.5/I_3, 0/I_4, 0/I_5\},\\
F_3 &= \{0/I_1, 0.5/I_2, 1/I_3, 0.5/I_4, 0/I_5\}\\
F_4 &= \{0/I_1, 0/I_2, 0.5/I_3, 1/I_4, 0.5/I_5\}\\
F_5 &= \{0/I_1, 0/I_2, 0/I_3, 0.5/I_4, 1/I_5\},
\end{aligned}
$$

其中, 隶属度的给定是由 Song and Chissom[11] 所提的原则. 为了方便起见, 集合中的元素可使用相对应的隶属度来表示.

$$
\begin{aligned}
&F_1 = (1, 0.5, 0, 0, 0)^t, \quad F_2 = (0.5, 1, 0.5, 0, 0)^t, \quad F_3 = (0, 0.5, 1, 0.5, 0)^t,\\
&F_4 = (0, 0, 0.5, 1, 0.5)^t, \quad F_5 = (0, 0, 0, 0.5, 1)^t
\end{aligned}
$$

经由时间数列分析方法所得的数据应都为稳定的.

3.2 模糊模式建构

台湾加权股票指数的资料模糊化后, 即在每个集合 F_i $(i = 1, 2, \cdots, 5)$ 中, 再将每日增加的加权股票指数找对应的隶属度, 如表 3.1.

由表 3.1 可以得到 1999 年 4 月的每日台湾加权股票指数涨跌的情况. 假设每日的加权股票指数涨跌的最大隶属度在 F_i $(i = 1, 2, \cdots, 5)$, 则认为这日的台湾加权股票指数涨跌为 $F_{date,i}$. 以 1999 年 4 月 1 日为例, 最大的隶属度在 F_5, 则可以认为 1999 年 4 月 1 日的台湾加权股票指数涨跌为 $F_{0401,5}$, 或称为数量是属于 "大涨" 的一日. 透过许多过

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

去的模糊资料，理论上可以找出模糊关系矩阵. 因此，可以就模糊集合中探讨连续两日台湾加权股票指数涨跌的模糊关系. 假定某个日 (t) 的台湾加权股票指数涨跌为 $F_{t,m}$，而其下一日 $(t+1)$ 的台湾加权股票指数涨跌为 $F_{t+1,n}$，期间的模糊关系以 $F_{t,m} \to F_{t+1,n}$ 表示之. 因此，可以获得台湾加权股票指数涨跌的模糊关系如下 (在此我们将时间的下标省略)：

$F_5 \to F_5$、$F_5 \to F_2$、$F_2 \to F_3$、$F_3 \to F_3$、$F_3 \to F_5$、$F_5 \to F_3$、$F_3 \to F_4$、$F_4 \to F_4$、$F_4 \to F_3$、$F_3 \to F_1$、$F_1 \to F_2$、$F_3 \to F_2$、$F_2 \to F_2$、$F_1 \to F_5$、$F_1 \to F_4$.

表 3.1 台湾加权股票日指数涨跌值之隶属度函数

日	指数涨跌值	F_1	F_2	F_3	F_4	F_{t5}
1999-04-01	137	0.00	0.00	0.00	0.47	0.53
1999-04-02	214	0.00	0.00	0.00	0.00	1.00
1999-04-03	−50	0.00	0.64	0.36	0.00	0.00
1999-04-06	−18	0.00	0.27	0.73	0.00	0.00
1999-04-07	−28	0.00	0.38	0.62	0.00	0.00
1999-04-08	138	0.00	0.00	0.00	0.45	0.55
1999-04-09	−8	0.00	0.15	0.85	0.00	0.00
1999-04-12	−23	0.00	0.33	0.67	0.00	0.00
1999-04-13	95	0.00	0.00	0.00	0.95	0.05
1999-04-14	61	0.00	0.00	0.35	0.65	0.00
1999-04-15	100	0.00	0.00	0.00	0.90	0.10
1999-04-16	−31	0.00	0.42	0.58	0.00	0.00
1999-04-17	115	0.00	0.00	0.00	0.72	0.28
1999-04-19	42	0.00	0.00	0.57	0.43	0.00
1999-04-20	5	0.00	0.00	1.00	0.00	0.00
1999-04-21	−154	0.85	0.15	0.00	0.00	0.00
1999-04-22	20	0.00	0.00	0.83	0.17	0.00
1999-04-23	118	0.00	0.00	0.00	0.69	0.31
1999-04-26	16	0.00	0.00	0.87	0.13	0.00
1999-04-27	−79	0.00	0.98	0.02	0.00	0.00
1999-04-28	−54	0.00	0.69	0.31	0.00	0.00
1999-04-29	−207	1.00	0.00	0.00	0.00	0.00
1999-04-30	82	0.00	0.00	0.10	0.90	0.00

将上述模糊时间数列关系 R 记为

$R_1 = F_5 \times F_5^t$、$R_2 = F_5 \times F_2^t$、$R_3 = F_2 \times F_3^t$、$R_4 = F_3 \times F_3^t$、$R_5 = F_3 \times F_5^t$、$R_6 = F_5 \times F_3^t$、$R_7 = F_3 \times F_4^t$、$R_8 = F_4 \times F_4^t$、$R_9 = F_4 \times F_3^t$、$R_{10} = F_3 \times F_1^t$、$R_{11} = F_1 \times F_3^t$、$R_{12} = F_3 \times F_2^t$、$R_{13} = F_4 \times F_1^t$、$R_{14} = F_2 \times F_1^t$、$R_{15} = F_1 \times F_4^t$.

其中 $\times$ 为取最小运算，t 为转置. 然后，

$$R(t, t-1) = R = \bigvee_{i=1}^{15} R_i. \tag{3.1}$$

其中，$\vee$ 为取最大运算. 经过计算之后，可得一阶模糊关系矩阵 R 为

$$R = \begin{bmatrix} 0.5 & 0.5 & 1 & 1 & 0.5 \\ 1 & 1 & 1 & 0.5 & 0.5 \\ 1 & 1 & 1 & 1 & 1 \\ 0.5 & 0.5 & 1 & 1 & 0.5 \\ 0.5 & 1 & 1 & 0.5 & 1 \end{bmatrix}, \tag{3.2}$$

所以，预测模式为

$$F(t) = F(t-1)R. \tag{3.3}$$

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

其中，$F(t-1), F(t)$ 分别是表示模糊集合中，第 $t-1$ 及 t 周的台湾加权股票指数涨跌数. 最后，计算出模式输出的隶属度，并将隶属度予以标准化. 处理的结果如表 3.2.

表 3.2 台湾加权股票指数涨跌数之输出隶属度与标准化隶属度与预测值

日	模式输出的隶属度	标准化的隶属度	预测值
1999-04-02	0.50, 0.77, 1.00, 0.74, 0.77	0.13, 0.20, 0.27, 0.20,0.20	16
1999-04-03	0.50, 1.00, 1.00, 0.50, 0.50	0.14, 0.29, 0.29, 0.14, 0.14	−7
1999-04-06	1.00, 1.00, 1.00, 0.68, 0.68	0.23, 0.23, 0.23, 0.16, 0.16	−14
1999-04-07	1.00, 1.00, 1.00, 0.87, 0.87	0.21, 0.21, 0.21, 0.18, 0.18	−2
1999-04-08	1.00, 1.00, 1.00, 0.81, 0.81	0.22, 0.22, 0.22, 0.18, 0.18	−6
1999-04-09	0.50, 0.78, 1.00, 0.73, 0.78	0.13, 0.21, 0.26, 0.19, 0.21	16
1999-04-12	1.00, 1.00, 1.00, 0.93, 0.93	0.21, 0.21, 0.21, 0.19, 0.19	1
1999-04-13	1.00, 1.00, 1.00, 0.84, 0.84	0.21, 0.21, 0.21, 0.18, 0.18	−4
1999-04-14	0.50, 0.53, 1.00, 0.98, 0.53	0.14, 0.15, 0.28, 0.28, 0.15	17
1999-04-15	0.68, 0.68, 1.00, 1.00, 0.68	0.17, 0.17, 0.25, 0.25, 0.17	12
1999-04-16	0.50, 0.55, 1.00, 0.95, 0.55	0.14, 0.15, 0.28, 0.27, 0.15	17
1999-04-17	1.00, 1.00, 1.00, 0.79, 0.79	0.22, 0.22, 0.22, 0.17, 0.17	−7
1999-04-19	0.50, 0.64, 1.00, 0.86, 0.64	0.14, 0.18, 0.27, 0.24, 0.18	17
1999-04-20	0.79, 0.79, 1.00, 1.00, 0.79	0.18, 0.18, 0.23, 0.23, 0.18	9
1999-04-21	1.00, 1.00, 1.00, 1.00, 1.00	0.20, 0.20, 0.20, 0.20, 0.20	5
1999-04-22	0.58, 0.58, 1.00, 0.93, 0.50	0.16, 0.16, 0.28, 0.26, 0.14	10
1999-04-23	0.92, 0.92, 1.00, 1.00, 0.92	0.19, 0.19, 0.21, 0.21, 0.19	7
1999-04-26	0.50, 0.66, 1.00, 0.85, 0.66	0.14, 0.18, 0.27, 0.23, 0.18	17
1999-04-27	0.94, 0.94, 1.00, 1.00, 0.94	0.19, 0.19, 0.21, 0.21, 0.19	6
1999-04-28	1.00, 1.00, 1.00, 0.51, 0.51	0.25, 0.25, 0.25, 0.13, 0.13	−26
1999-04-29	1.00, 1.00, 1.00, 0.66, 0.66	0.23, 0.23, 0.23, 0.15, 0.15	−16
1999-04-30	0.50, 0.50, 1.00, 1.00, 0.50	0.14, 0.14, 0.29, 0.29, 0.14	17

由预测模式为 $F(t)=F(t-1)R$ 的输出结果为一模糊集合. 若模糊集合的结果能满足需求，则预测模式的建构工作便告完成了. 但是，在很多情况下所要求的结果往往是数值资料. 所以往往也需要将模糊集合转换成数值资料，以便于分析.

(1) 由 2.3 节的最大隶属度法，来做模糊数值之转换. 由表 (3.2) 得到论域 U 上输出元素的隶属度及组区域 $\{(-210,-124),(-124,-38),(-38,48),(48,134),(134,220)\}$，组中点值分别为 $\{-167,-81,5,91,177\}$，经计算后得表 3.3

(2) 由 2.3 节的隶属度加权平均法，来做模糊数值之转换. 即输出值 $=S_1*(-167)+S_2*(-81)+S_3*5+S_4*91+S_5*177$，其中 S_i 为标准化后的隶属度，$i=1,2,\cdots,5$.

根据上述原则所获得的预测输出值如表 3.3.

表 3.3 由最大隶属度法做模糊数值之转换

	组区域	组中点值	输出值
$\text{Max}\{0.50,0.77,1.00,0.74,0.77\}=1$	$(-38,48)$	5	5
$\text{Max}\{0.50,1.00,1.00,0.50,0.50\}=1$	$(-124,-38)$ $(-38,48)$	−81,5	−38
$\text{Max}\{1.00,1.00,1.00,0.68,0.68\}=1$	$(-210,-124)$ $(-124,-38)$ $(-38,48)$	−167,−81,5	−81
⋮	⋮	⋮	⋮
$\text{Max}\{1.00,1.00,1.00,0.66,0.66\}=1$	$(-210,-124)$ $(-124,-38)$ $(-38,48)$	−167,−81,5	−81
$\text{Max}\{0.50,0.50,1.00,1.00,0.50\}=1$	$(-38,48)$ $(48,134)$	5,91	48

为了比较预测结果，我们以准确率为标准. 即 $P=\sum_{i=1}^{N} I_{RL_t}(FL_t)/N$，其中 FL_t 表示预测语言变量值，RL_t 表示实际语言变量值，N 为样本长度，I 为一指针函数，即

$$I_t(x)=\begin{cases}1, & \text{if } x=t,\\ 0, & \text{if } x\neq t.\end{cases}$$

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

表 3.4 台湾加权股票指数之预测值比较

日	真实语言 (值)	ARIMA	最大隶属度法	隶属度加权平均法
1999-04-02	大涨 (7233)	大涨 (7484)	平盘 (7024)	平盘 (7035)
1999-04-03	跌 (7182)	大涨 (7511)	跌 (7195)	平盘 (7226)
1999-04-06	平盘 (7164)	大涨 (7439)	跌 (7101)	平盘 (7168)
1999-04-07	平盘 (7136)	大涨 (7374)	跌 (7083)	平盘 (7162)
1999-04-08	大涨 (7273)	大涨 (7412)	跌 (7055)	平盘 (7130)
1999-04-09	平盘 (7266)	大涨 (7527)	平盘 (7278)	平盘 (7289)
1999-04-12	平盘 (7242)	大涨 (7602)	跌 (7185)	平盘 (7267)
1999-04-13	涨 (7338)	大涨 (7566)	跌 (7161)	平盘 (7238)
1999-04-14	涨 (7399)	大涨 (7483)	平盘 (7343)	平盘 (7355)
1999-04-15	涨 (7498)	涨 (7471)	涨 (7447)	平盘 (7411)
1999-04-16	平盘 (7467)	涨 (7563)	平盘 (7503)	平盘 (7515)
1999-04-17	涨 (7582)	大涨 (7669)	跌 (7386)	平盘 (7460)
1999-04-19	平盘 (7623)	涨 (7683)	平盘 (7587)	平盘 (7599)
1999-04-20	平盘 (7628)	平盘 (7607)	平盘 (7671)	平盘 (7632)
1999-04-21	大跌 (7474)	跌 (7552)	平盘 (7633)	平盘 (7633)
1999-04-22	平盘 (7495)	涨 (7602)	平盘 (7479)	平盘 (7484)
1999-04-23	涨 (7613)	大涨 (7717)	涨 (7543)	平盘 (7502)
1999-04-26	平盘 (7629)	大涨 (7780)	平盘 (7618)	平盘 (7630)
1999-04-27	跌 (7550)	涨 (7735)	平盘 (7677)	平盘 (7635)
1999-04-28	跌 (7497)	涨 (7656)	跌 (7469)	平盘 (7524)
1999-04-29	大跌 (7290)	大涨 (7656)	跌 (7416)	平盘 (7481)
1999-04-30	涨 (7371)	大涨 (7755)	涨 (7338)	平盘 (7407)
准确率		0.18	0.5	0.41

由表 3.3 可知，模糊时间数列模式的预测结果，较单变量 ARIMA 的预测结果来得准确，可见此一方法是相当有效的预测工具. 但在模糊时间数列模式的数值转换中，又以最大隶属度法比隶属度加权平均法来得准确，因此，一般学者常用的隶属度加权平均法在台湾加权股票指数上未必适用，因为影响股票指数的因素众多，如汇率、景气指针、国民所得，等等，故有时也可考虑用最大隶属度法来做模糊数值的转换，应用于其它的模糊时间数列的数值转换上. 有了合理的预测模式，才可依据预测的结果，拟定适当的投资策略，否则在无头绪的情况下，将造成投资者无所适从的窘境.

4 结论

在社会科学的测度理念上，模糊统计和模糊相关性的使用可说是一种数字模式的推广. 单一数值性质的资料，主要限制是具有过度解释的潜在危机. 而使用比较模糊且朴实的语言资料，则可以避免这些无法预料的危险. 事实上，利用语言模式作为预测，可以发现每一期皆增加了结果的模糊性. 从一般的生活中这种模糊化似乎也是很正常的现象，但是从另一角度来看，如果数值处理的观念没有改变，预测方法没有突破，往往却阻碍了长期预测的可能性.

本文尝试利用模糊时间数列方法，针对台湾地区加权股票日指数，建构预测模式. 以 1999 年 4 月 1 日至 1999 年 4 月 30 日的日指数为历史资料，建立适当的预测模型. 我们以准确度来衡量模糊时间数列模式，与单变量 ARIMA 模式的预测能力，并且清楚地显示出模糊时间数列模式相对于 ARIMA 模式，有较佳的预测能力. 这对于投资者而言，不外乎提供了一种新的预测方法，又可以获得较好的预测结果，使得投资者在正确的信息下，做出更明确的判断.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

参 考 文 献

1 Clymer J P Corey, Gardner J. Discrete Event Fuzzy Airport Control. *IEEE Transactions on Systems, Man, and Cybernetics*, 1992, 22(2): 343–351

2 Cutsem B V, Gath I. Detection of Outliers and Robust Estimation Using Fuzzy Clustering. *Computational Statistics and Data Analysis*, 1993, 15: 47–61

3 Hathaway R J, Bezdek J C. Switching Regression Models and Fuzzy Clustering. *IEEE Transactions of Fuzzy Systems*, 1993, 1: 195–204

4 Yoshinari Y, Pedrycz W, Hirota K. Construction of Fuzzy Models through Clustering Techniques. *Fuzzy Sets and Systems*, 1993, 54: 157–165

5 Romer C, Kandel A, Backer E. Fuzzy Partitions of the Sample Space and Fuzzy Parameter Hypotheses. *IEEE Transactions on Systems, Man and Cybernetics*, 1995, 25(9): 1314–1321

6 Wu B, Hung S. A Fuzzy Identification Procedure for Nonlinear Time Series: with Example on ARCH and Bilinear Models. *Fuzzy Set and System*, 1999, 108: 275–287

7 Wu B, Chen M. Use Fuzzy Statistical Methods in Change Periods Detection. *Applied Mathematics and Computation*, 1999, 99: 241–254

8 Klir G J, Folger T A. Fuzzy Set, Uncertainty,and Information. Englewood Cliffs, NJ: Prentice Hall, 1988

9 Zimmermann H J. Fuzzy Set Theroy and Its Applications. Boston: Kluwer Academi, 1991

10 Hendershot G, Placek P (eds.) Predicting Fertility. Lexington, MA: Health and Co., 1981

11 Song Q, Chissom B S. Fuzzy time Series and its Models. *Fuzzy Sets and Systems*, 1993, 54: 269–277

12 Song Q, Chissom B S. Forecasting Enrollments with Fuzzy Time Series Part Ⅱ. *Fuzzy Sets and Systems*, 1994, 62: 1–8

FUZZY TIME SERIES ANALYSIS AND FORECASTING: WITH AN EXAMPLE TO TAIWAN WEIGHTED STOCK INDEX

WU BERLIN, LIN YU-CHUN

(*Department of Mathematical Sciences, National Chengchi University, Taiwan*)

Abstract In recent years, along with the advance of technology and industry, the innovation and improvement of forecasting techniques have caught more and more attention. Especially, in the issues of financial engineering, economic development, population policy, and management planning and control, forecasting provides indispensable information in decision-making process. Regarding stock market as an example, the essence of closing price is uncertain and indistinct. Therefore, if we merely consider closing price of yesterday to build our forecasting model, not only will we misestimate the future trend, but also we will suffer unnecessary loss. Based on this reason, in this research we propose an integrated procedure for fuzzy time series modeling and forecasting through fuzzy relation equations. We apply this technique to construct a fuzzy time series model for Taiwan Weighted Stock Index and forecast future trend while comparing the forecasting performance by average forecasting accuracy. We strongly believe that this model will be profound of meaning in forecasting future trend of financial market.

Key words Fuzzy time series, forecasting, Taiwan weighted stock index

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net