

文章编号 1000-5013(2000)01-0107-04

模糊推理方法在股市预测中的应用

胡 一 朗

(华侨大学管理信息科学系, 泉州 362011)

摘要 以日成交金额和15日相对强弱指标的绝对量和增量共4个变量作为影响股指走势因素,应用模糊推理方法对沪市中短期上升行情进行预测,结果较为满意. 模糊推理方法对每一种影响因子(x_i)与输出因子(y),统计其在每一分级中占的次数,取两者次数的比值即为分级模糊关系值,从而得到模糊关系矩阵 $R(x_i, y)$. 再取 $R(x_i, y)$ 每行最大值相加,可得到4个数值. 以每个数值除以4个数值之和,即为每个影响因子权重 w_i . 最后,建立预测模型 $y = w \circ R^*$. R^* 是由所预测行情的4个变量决定的矩阵.

关键词 模糊推理方法, 股市预测, 成交金额, 相对强弱指标

分类号 O 159 : F 830. 91

文献标识码 A

成交量是股价(或股指)上升和下跌的原动力,成交量的大小反映了市场人气的聚散,反映了市场上的一切信息. 强弱指标是基于供需平衡原理产生的,用以分析市场买方和卖方的强弱程度,从而预测股价或股指未来变化趋势^[1]. 成交资金量分析是重要的分析方法之一^[2]. 本文以成交金额和强弱指标(RSI)作为影响股指走势因素,应用模糊推理方法预测沪市股指中短期上升走势,结果较为满意. 在推理合成运算中,选用的算子为(\cdot , \oplus).

1 原始数据与等级划分

首先,我们对上升行情作如下规定:一次上升行情是以涨幅大于5%计,并且在该次行情上升过程中的回调不大于5%. 按此规定取样,1996年1月份至1999年1月份共有26次上升行情(表1). 其次,根据成交量与股价指数的关系^[3],欲使股价指数上升,成交量必须放大.

通过对上升行情的分析可以发现,涨幅(y)大小不仅与成交金额绝对量(x_1 指的是该次行情顶部区域的日最大成交金额)有关,而且与成交金额增量(x_3 指的是成交金额绝对量与该次底部区域的日最小成交金额之差)有关. 如表1中,序号11和序号13的涨幅分别为14.6%, 8.0%,前者绝对量为65.6亿元小于后者的87.2亿元,但前者增量为49.2亿元大于后者的17.6亿元. 因而,本文对成交金额分别考虑其绝对量和增量. 类似地,对15日相对强弱指标(RSI(15))也考虑其绝对量(x_2)和增量(x_4). 最后,股指涨幅及4个指标数据取自《上海证券交易所每日证券行情》或通过该行情表计算而得(表1). 通过实例的反复验证,采用表2的等级划分,分级值见表3.

表1 原始数据

序号	起止时间	$y / (\%$	$x_1 / \text{亿元}$	x_2	$x_3 / \text{亿元}$	x_4
1	1996.01.19 ~ 1996.03.05	18.2	26.9	74.2	24.9	57.5
2	1996.03.29 ~ 1996.04.29	34.4	76.5	85.5	73.7	39.0
3	1996.05.03 ~ 1996.05.20	12.4	45.8	68.4	31.1	11.3
4	1996.05.24 ~ 1996.07.24	41.9	81.6	71.1	74.8	24.6
5	1996.08.01 ~ 1996.08.12	9.8	42.2	63.8	13.9	11.9
6	1996.08.27 ~ 1996.09.05	9.4	60.8	48.5	40.0	13.9
7	1996.09.13 ~ 1996.10.29	38.0	111.3	79.0	95.1	43.5
8	1996.11.06 ~ 1996.11.09	16.8	108.0	66.1	80.7	18.7
9	1996.11.22 ~ 1996.12.11	37.3	198.6	79.6	149.3	30.6
10	1996.12.25 ~ 1996.12.30	9.7	49.4	44.1	19.6	7.8
11	1997.01.07 ~ 1997.02.18	14.6	65.6	58.3	49.2	19.3
12	1997.02.20 ~ 1997.05.12	73.5	167.3	83.3	117.6	43.5
13	1997.05.16 ~ 1997.05.21	8.0	87.2	52.2	17.6	6.5
14	1997.05.26 ~ 1997.06.05	12.9	131.7	53.9	76.6	14.1
15	1997.06.13 ~ 1997.06.23	14.5	91.3	52.4	56.5	13.7
16	1997.07.08 ~ 1997.07.23	14.8	65.7	49.0	30.7	18.4
17	1997.07.29 ~ 1997.08.08	7.4	60.1	51.9	31.2	12.9
18	1997.08.18 ~ 1997.09.11	13.3	63.4	64.7	40.2	27.2
19	1997.09.23 ~ 1997.10.28	21.1	70.0	64.8	58.0	37.8
20	1997.11.15 ~ 1998.01.08	12.2	90.1	72.2	72.2	35.1
21	1998.01.13 ~ 1998.02.09	13.7	50.2	63.2	21.2	18.8
22	1998.03.23 ~ 1998.06.04	20.8	117.2	73.5	92.6	34.1
23	1998.08.18 ~ 1998.09.17	21.2	100.7	59.5	72.0	42.8
24	1998.09.25 ~ 1998.10.16	6.0	63.9	60.6	26.7	23.4
25	1998.10.22 ~ 1998.11.17	9.1	95.1	69.3	52.1	22.8
26	1999.01.05 ~ 1999.01.11	6.3	36.4	48.7	22.6	21.2

表2 等级划分

等级	1	2	3	4	等级	1	2	3	4
y	(5, 10]	(10, 16]	(16, 23]	(23, +)	x_3	(0, 30]	(30, 65]	(65, 95]	(95, +]
x_1	(0, 65]	(65, 95]	(95, 120]	(120, +)	x_4	(0, 14]	(14, 28]	(28, 38]	(38, 100]
x_2	(0, 53]	(53, 65]	(65, 78]	(78, 100]					

表3 分级值

序号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
y	3	4	2	4	1	1	4	3	4	1	2	4	1	2	2	2	1	2	3	2	2	3	3	1	1	1
x_1	1	2	1	2	1	1	3	3	4	1	2	4	2	4	2	2	1	1	2	2	1	3	3	1	3	1
x_2	3	4	3	3	2	1	4	3	4	1	2	4	1	2	1	1	1	2	2	3	2	3	2	2	3	1
x_3	1	3	2	3	1	2	4	3	4	1	2	4	1	3	2	2	2	2	2	3	1	3	3	1	2	1
x_4	4	4	1	2	1	1	4	2	3	1	2	4	1	2	1	2	1	2	3	3	2	3	4	2	2	2

2 数学模型及其预测

2.1 数学模型

本文的模糊推理采用如下方法^[6]: 把各种影响因素分级后, 由统计求得权重及模糊关系. 对于每一种影响因子(x_i) 与输出因子(y), 统计其在每一分级中占的次数. 取两者次数的比值即为分级模糊关系值, 从而得到模糊关系矩阵 $R(x_i, y)$. 再取 $R(x_i, y)$ 每行最大值相加, 可得到4个数值. 以每个数值除以4个数值之和, 即为每个影响因子权重 w_i . 最后建立预测模糊 $y = w \circ R^*$. R^* 是由所预测行情的4个变量决定的矩阵. 在推理合成运算中, 选用的算子“ \circ ”为(\cdot, \oplus). 其中 \oplus 为有界和, $a \oplus b \triangleq 1 - (a + b)$.

以表3中的序号1~25为建模样本, 应用上述方法可得到模糊关系矩阵 $R(x_i, y)$, 因子权重 w 和预测模型 y .

$$R(x^1, y) = \begin{bmatrix} 0.556 & 0.333 & 0.111 & 0 \\ 0.125 & 0.500 & 0.125 & 0.250 \\ 0.200 & 0 & 0.600 & 0.200 \\ 0 & 0.333 & 0 & 0.667 \end{bmatrix}, R(x^2, y) = \begin{bmatrix} 0.667 & 0.333 & 0 & 0 \\ 0.250 & 0.500 & 0.250 & 0 \\ 0.143 & 0.286 & 0.428 & 0.143 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$
$$R(x^3, y) = \begin{bmatrix} 0.667 & 0.167 & 0.166 & 0 \\ 0.333 & 0.556 & 0.111 & 0 \\ 0 & 0.286 & 0.428 & 0.286 \\ 0 & 0 & 0 & 1 \end{bmatrix}, R(x^4, y) = \begin{bmatrix} 0.714 & 0.286 & 0 & 0 \\ 0.222 & 0.556 & 0.111 & 0.111 \\ 0 & 0.250 & 0.500 & 0.250 \\ 0 & 0 & 0.400 & 0.600 \end{bmatrix},$$
$$y = w \circ R^* = (0.234, 0.261, 0.267, 0.238) \circ R^*.$$

2.2 预测实例

由表3知, 序号1的 x_1, x_2, x_3, x_4 分别为1级, 3级, 1级, 4级. 由此可得

$$y = (0.234, 0.261, 0.267, 0.238) \begin{bmatrix} 0.556 & 0.333 & 0.111 & 0 \\ 0.143 & 0.286 & 0.428 & 0.143 \\ 0.667 & 0.167 & 0.166 & 0 \\ 0 & 0 & 0.400 & 0.600 \end{bmatrix} =$$
$$(0.644, 0.622, 0.776, 0.381).$$

按最大隶属原则, 预测为3级与实际吻合. 类似地, 对序号2~26进行预测, 预测结果列于表4, 并与实际结果比较.

表4 预测结果与实际结果比较

序号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
预测结果	3	4	2	2	1	1	4	3	4	1	2	4	2	2	2	2	1	2	2	2	1	3	3	1	1	1
评定	√	√	√	×	√	√	√	√	√	√	√	√	×	√	√	√	√	√	×	√	×	√	√	√	√	√

表4中序号6、序号17的预测结果为1级或2级. 在两个计算隶属度相等的情况下, 则可根据模糊关系矩阵 R^* 中各级隶属度总的情况来评判. 本文序号26预测准确, 序号1~25拟合率为84%. 对于变幻莫测的股指走势而言, 这样的结果是较为满意的. 故本文方法可以作为股市投资者的一种参考工具.

3 结束语

应该指出的是,由于成交金额、RSI 是动态变化的,上升行情的日成交金额、RSI 能达到的最大值是事先未知的.这就需要投资者一方面应用其它方法或凭经验对其(主要是日成交金额)进行研判;另一方面,基于量先行于价的经验法则,在股指涨幅大于5%后,利用本文预测模型进行跟踪预测.随着股市规模的不断扩大,日最大成交金额和日最小成交金额将会有所变化,因而对等级的划分,投资者要根据市场实际情况适时修改.对于后续行情(如本文序号26起)的实际数值可以加入原始数据列,重新建立预测模型,以期完善并提高预测精度.

对股市下跌行情以及个股股价走势,也可类似地建立预测模型.限于篇幅,本文略去.

参 考 文 献

- 1 夏时柏. 上海股市实用操作技巧[M]. 北京: 经济日报出版社, 1995. 73, 107~108, 110
- 2 袁尔豪. 股票投资技术新概念——数据分析方法[M]. 广州: 华南理工大学出版社, 1996. 32~33
- 3 陈鸣钊, 张志烈, 樊宝康. 模糊数学及其实用[M]. 南京: 河海大学出版社, 1993. 176~178

Application of Fuzzy Inference to the Prediction of Stock Price

Hu Yilang

(Dept. of Manag. Info. Sci., Huaqiao Univ., 362011, Quanzhou)

Abstract Four variables, absolute amount and increment of daily turnover and relative strength of 15 days, are taken as factors influencing the trend of share exchange. The method of fuzzy inference is applied to predicting medium term and short term bull market of shanghai share exchange, with fairly satisfactory result. The method is adopted in proper steps. Firstly, various influencing factors are graded, and their weights and fuzzy relationship are solved statistically. Secondly, the orders of every influencing factors in every grade (x_i) and the output factor in it (y) are counted up, the ratio of their orders is just the graded fuzzy relationship from which the fuzzy relational matrix $R(x_1, y)$ can be obtained. Thirdly, to add up the maximum values in every line of $R(x_1, y)$, four numerical values can be obtained; and to divide every numerical value by the summation of four ones, the weight of every influencing factor can be obtained. Finally, to set up a prediction model $y = w^0 R^*$, where R^* is the matrix decided by four variables for predicting market.

Key words fuzzy inference, stock price prediction, turnover, relative strength index