

## 证券研究报告—深度报告

### 金融工程

### 数量化投资

## 交易性数据挖掘系列报告

2012年08月06日

### 专题报告

#### 相关研究报告:

《金融工程专题研究: 基于模式聚类的短线选股模型》——2011-8-30

《金融工程专题研究: 模式识别选股模型的优化——支撑线和压力线的组合识别》——2011-12-26

《金融工程专题研究: 基于动态时间弯曲的形态匹配在指数增强中的实证研究》——2012-3-12

《金融工程专题研究: 基于动态时间规整的择时策略》——2012-5-29

《交易性数据挖掘系列报告: 均线回抽的超跌反弹选股策略》——2012-7-30

#### 联系人: 郑亚斌

电话: 021-60933150

E-MAIL: zhengyb@guosen.com.cn

#### 证券分析师: 林晓明

电话: 0755-82136165

E-MAIL: linxiaom@guosen.com.cn

证券投资咨询执业资格证书编码: S0980512020001

#### 证券分析师: 戴军

电话: 0755-82133129

E-MAIL: daijun@guosen.com.cn

证券投资咨询执业资格证书编码: S0980510120064

# 价量联动的形态匹配择时策略

## ●以史为鉴可以知兴替: 历史总是在不断的重演

历史会重演是技术分析的重要假设和基石。市场上的技术分析派通过长期的盘面观察积累市场经验, 并形成适合自己的投资逻辑和策略。正是从这个角度出发, 我们试图探讨量化技术分析的实现, 之前已有一些关于价量模式、支撑线压力线模式以及形态分析的模式的研究报告。大多数情况下, 我们从交易性数据出发, 总结和归纳历史数据中的经验规律, 并验证其在不同市场环境下的效果。我们始终认为, 以史为鉴可以知兴替, 从历史会重演的角度出发, 可以为我们提供更多准确有效的信息。

## ●技术分析之价量关系

技术分析的流派五花八门, 例如各种技术指标的测算、波浪分析、分形理论等。但究其根源, 无外乎对价量的分析, 大部分的技术分析手段皆是由价量数据衍生得到。一般而言, 大多数投资者认为价格反应了趋势, 而成交量是对趋势的确认, 在实际应用中不能将两者互相隔离分析。我们此前的系列报告也是从交易性数据出发, 形成以价格为主, 成交量为辅的研究思路, 从不同角度分析了价量关系。

## ●价量联动的形态匹配择时策略

在上一篇报告《基于动态时间规整的择时策略》中, 在给定指数近期走势的前提下, 我们从历史会重演的角度出发, 对指数的历史走势进行回溯, 挖掘出与近期走势较为相似的历史片段, 进而观察在那些历史片段后的走势, 最终依此给出当前走势的后期预测结果。上篇报告仅从价格的维度对形态进行分析, 虽有一定的效果, 但我们认为加入成交量的信息将有助于提高策略的有效性和稳定性。基于此, 本篇报告利用 Brownian 相关系数将价格和成交量两个维度同时纳入考虑范围, 并检验其在择时策略上的效果。以观察 30 个交易日, 预测 10 个交易日为参数设置, 策略判断涨跌方向的准确率为 57%; 与随机预测结果相比, 策略的胜率为 76%; 预测涨跌幅与真实涨跌幅差异在可接受范围内。

## ●未来研究方向

我们已经将价量联动的形态匹配方法用于择时策略, 并验证了其有效性。未来的研究方向包括以下几个方面: 1. 从行业的交易性数据出发, 结合行业指数走势的价量关系, 对行业的走势进行判断指导行业配置: 超配预期上涨最为强烈的行业, 低配预期下跌最为强烈的行业。2. 从个股的交易性数据出发, 分析个股的价量数据, 从而对个股走势的判断, 最终形成价量联动的选股策略。3. 总结所有的历史数据, 对其进行聚类分析, 最终形成有效的价量联动形态库。总而言之, 我们试图从量化的技术分析出发, 结合计算机强大的计算能力, 为投资者在不同的市场环境和投资方向上提供投资建议和参考, 最大限度地融合计算机数据运算能力及投资者的主观决策能力。

#### 独立性声明:

作者保证报告所采用的数据均来自合规渠道, 分析逻辑基于本人的职业理解, 通过合理判断并得出结论, 力求客观、公正, 结论不受任何第三方的授意、影响, 特此声明。

## 内容目录

前言 .....	4
基于交易性数据的量化技术分析 .....	4
研究框架 .....	5
价量联动的利器：Brownian Correlation 方法 .....	6
Brownian 相关系数范例 .....	6
Brownian 相关系数计算 .....	7
价量联动的形态匹配择时策略 .....	8
择时策略流程实例 .....	9
择时策略效果评价及分析 .....	11
总结和扩展讨论 .....	15
国信证券投资评级 .....	17
分析师承诺 .....	17
风险提示 .....	17
证券投资咨询业务的说明 .....	17

## 图表目录

图 1: 基于交易性数据的量化策略研究框架 .....	4
图 2: 报告研究框架 .....	6
图 3: Pearson 相关系数范例 .....	7
图 4: Brownian 相关系数范例 .....	7
图 5: 上证指数价量联动的形态匹配实例 .....	9
图 6: 预测走势与真实走势对比图 (2009 年 7 月 2 到 2009 年 8 月 12 日) .....	10
图 7: 涨跌方向预测准确率与上证指数走势 .....	12
图 8: 10 日涨跌幅预测误差经验分布图 .....	12
图 9: 带阈值的涨跌方向预测准确率 .....	14
图 10: 上证指数价量联动形态匹配 (截至 2012 年 8 月 2 日) .....	14
图 11: 上证指数后期走势预测 (截至 2012 年 8 月 2 日) .....	15
表 1: 相似历史片段时点 .....	10
表 2: 涨跌方向预测分年正确率统计 .....	13
表 3: 最新走势相似历史片段时点 .....	15

## 前言

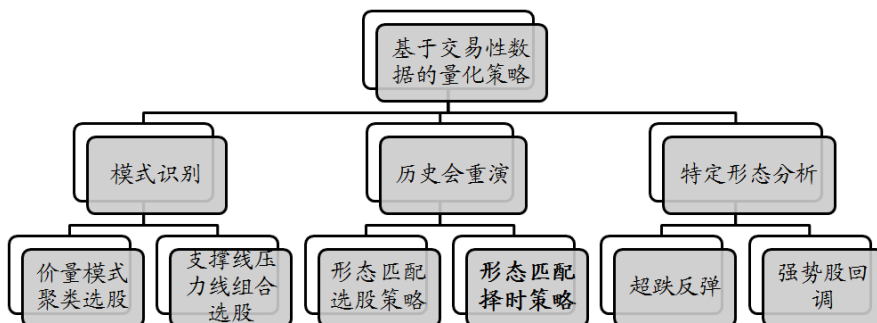
### 基于交易性数据的量化技术分析

从主要应用的数据划分，量化策略大体可以分为从交易性数据出发的策略以及从基本面数据出发的策略。前者偏重于研究隐藏在交易性数据背后的规律，例如：市场上的技术分析派总是通过 K 线的走势图勾勒出指数或个股的运行趋势和通道，并依此进行决策。他们并不把重心放在研究个股基本面或宏观经济方面。后者将研究重心放在个股基本面数据上（包括财务指标、负债情况、净利润等方面），诸如多因子选股模型、GARP 选股模型等。

从某种意义上来说，两种方法各有优缺点，交易性数据出发的策略通常具有更好的可扩展性。例如：技术分析的思路既可以用于个股的判断，也可以向上推广用于行业涨跌预测以及择时等。而基本面数据出发的策略更多地偏重于研究个股或者指数的各类基本面指标，符合行业研究的思路，具有较好的逻辑解释能力。两种方法并不存在矛盾，在具体投资的过程中，我们可以同时参考两类策略的结论。

本篇报告延续之前基于交易性数据的研究框架，在之前形态匹配择时策略的基础上，加入了成交量的因素，对已有的策略进行了一定的改进。在此之前，我们首先对已有的基于交易性数据的量化策略进行梳理，如图 1 所示。

图 1：基于交易性数据的量化策略研究框架



资料来源：国信证券经济研究所

从图 1 中我们可以看出，目前我们基于交易性数据的量化策略大体分为三个方向：模式识别策略、基于历史会重演假设的策略以及对某种特定形态进行分析的策略。从模式识别来说，我们对成交价格、成交量模式进行聚类分析，并形成价量模式聚类选股策略，参见《金融工程专题研究：基于模式聚类的短线选股模型》。在此基础上，我们对模式识别策略进行优化，针对成交价格的角度，研究了一套自动绘制支撑线压力线的策略，并研究支撑线压力线特征对后期涨跌的判断，参见《金融工程专题研究：模式识别选股模型的优化——支撑线和压力线的组合识别》。

此外，我们也针对某些特定的技术形态进行专门的分析。从技术分析的均线理论出发，我们观察到当股价长期偏离均线时，均线对短期股价形成吸引，会引发股价向上反弹会抽均线。这种交易性机会被验证是非常有效的，是一种尤其适用于熊市及震荡市的策略。参见《交易性数据挖掘系列报告：均线回抽的超跌反弹选股策略》。

从历史会重演的角度，我们试图找到与大盘或个股近期走势较为相近的历史片段，并试图利用相似历史片段之后的走势预测当下可能的后期走势。我们首先研究了形态匹配在选股方面的效果，并将其推广至指数择时。参见《金融工程专题研究：基于动态时间弯曲的形态匹配在指数增强中的实证研究》及《金融工程专题研究：基于动态时间规整的择时策略》。在这篇报告中，我们尝试在之前仅研究价格走势的研究基础上，加入成交量的因素，检验是否可以进一步提高策略的效果。从直观上来说，价量是不可分割的整体，两段走势的相似性不仅要体现在价格上，在成交量的趋势和方向上也应类似。

在价格的技术上同时成交量的变化需要解决两个难点：1. 价格和成交量的变化程度不一致，由于涨跌停板的限制，指数的极端涨跌幅在 $\pm 10\%$ ，而成交量可能出现极度的放大或萎缩的情况。2. 价格和成交量同时刻画了趋势，传统的方法是分别按照价格和成交量计算两段走势的相似性，最终用线性加权的方式给出最终相似度的分值。这种办法的缺点在于需要定量给出价格和成交量的权重，可扩展性较差。在本篇报告中，我们引入了 Brownian 相关系数较好地同时考虑价格和成交量两个维度。Brownian 相关系数是一种可以度量多维随机向量相似性的方法，我们甚至可以在价量的基础上加入开盘价、最高价、最低价、收盘价等维度，更好的捕捉形态的趋势。在对 Brownian 相关系数展开具体介绍之前，我们首先介绍本篇报告的研究框架。

### 研究框架

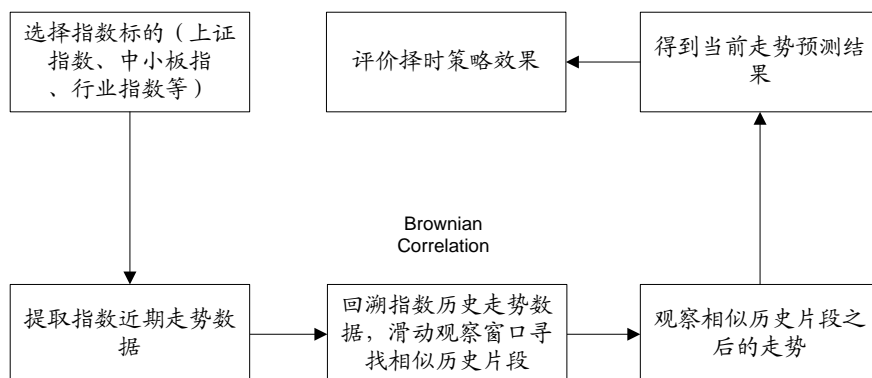
本篇报告的研究框架与之前并无较大区别，均从历史会重演的角度出发。当需要对指数或个股进行后期走势预测时，我们提取标的指数或个股在观察窗口内的走势数据，并对其历史的交易数据进行回溯，找到历史数据中与当前走势较为相近的片段。那么从历史会重演的角度出发，相似的历史走势倾向于产生相似的未来走势，换言之，我们有较大概率会重演历史。从这个角度来说，相似历史片段之后的走势为我们当下的预测提供了非常有意义的借鉴和参考。当然，在实际投资的过程中，需要考虑非技术层面因素的影响，例如宏观市场环境、政策因素等。

从上述的讨论不难看出，策略的核心在于如何定量给出两段走势的相似性。在我们之前的报告中，我们借鉴了语音识别领域的动态时间规整算法，较好地解决了形态之间的伸缩以及时间点不对齐的问题，取得了一定的效果。但不可回避的一点是，我们并没有考虑价格之外因素的影响，诸如成交量、换手率等。技术分析时常把价格当作骨骼，成交量当作血液，两者是不可分割的整体。按照价格上涨下跌及成交量的放大萎缩，我们可以将价量关系大致分为量增价涨、量缩价涨、量增价跌、量缩价跌四种情况。量是价格的先行指标，股价在上涨过程中如若没有成交量的配合，将极大地影响上涨的真实性。成交量很大程度上反应了投资者的心理活动，人气高涨的时候往往表现为买卖活跃，成交量放大，而投资者举棋不定，犹豫不决时，往往伴随着成交量的萎缩。俗话说地量见地价，也再次说明了价量的关系。

本篇报告的研究框架如图 2 所示，在给定标的指数后，数据提取模块得到标的指数在观察窗口内（短期、中期、长期趋势均可）的走势。其次，我们按照滑动窗口对历史数据进行回溯，结合 Brownian 相关系数的方法，从价量联动的角度，定位出与当前走势较为相近的历史片段。第三，根据相似历史片段之后的走势得到当前走势的预测结果，假若相似历史片段大都出现上涨趋势，那么我们在当前时点买入应是有利可图的。最后，我们对择时策略进行全面的评价，以期对后续的策略改进提供参考。



图 2: 报告研究框架



资料来源：国信证券经济研究所

由于我们需要借助其他的手段融价格 and 成交量，如前所述，我们借鉴 Brownian 相关系数的办法同时考虑价格两个维度。在接下来的章节中，我们将具体展开介绍 Brownian 相关系数。

## 价格联动的利器：Brownian Correlation 方法

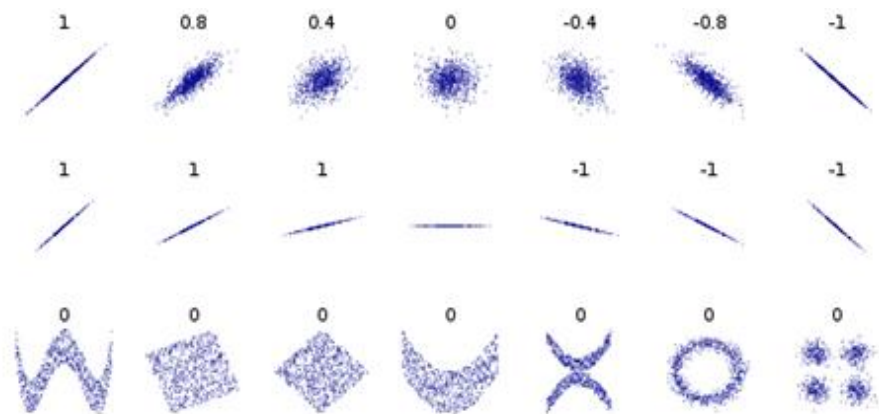
在之前的报告《基于动态时间规整的择时策略》中，我们基于某个指数（上证指数、中小板指数、创业板指数、行业指数等）的近期走势，根据技术分析中历史总是会重演的假设，我们回溯其历史走势，挖掘与当前走势非常类似的历史片段，根据那些相似的历史片段之后的走势预测当前走势之后的发展。在上述报告中，挖掘与当前走势相近的历史片段，只用到指数本身的数据，亦即只用到价格信息，然而，两个指数走势片段相近，不应该只是价格走势相近，而应该是各方面走势都相近。在本报告中，我们不仅考虑价格，还将成交量考虑进来，挖掘和近期指数价格和成交量都相近的历史片段。

### Brownian 相关系数范例

度量两个 1 维随机向量的相似性，有许多不同的方法，可以简单的使用欧式距离来度量，也可以使用我们之前报告中提到的动态规整距离来度量，还可以使用通常的 Pearson 相关系数。然而，度量两个 2 维随机向量的相似性，这些方法都不再适用，Szekely、Rizzo、Bakirov 在 2007 年提出的 Brownian Correlation 方法，提供了一个可行的度量任意维等长随机向量相似性的方法。与 Pearson 相关系数不同，Brownian 相关系数取值范围为  $[0,1]$ ，并且当 Brownian 相关系数为 0，意味着两个随机向量是独立的，越接近 0，说明独立性越强，越接近 1，说明相关性越强，当然，这种相关有可能是正相关，也有可能是负相关。因此，在应用中，我们还需要利用 Pearson 相关系数来辅助判断是正相关还是负相关。

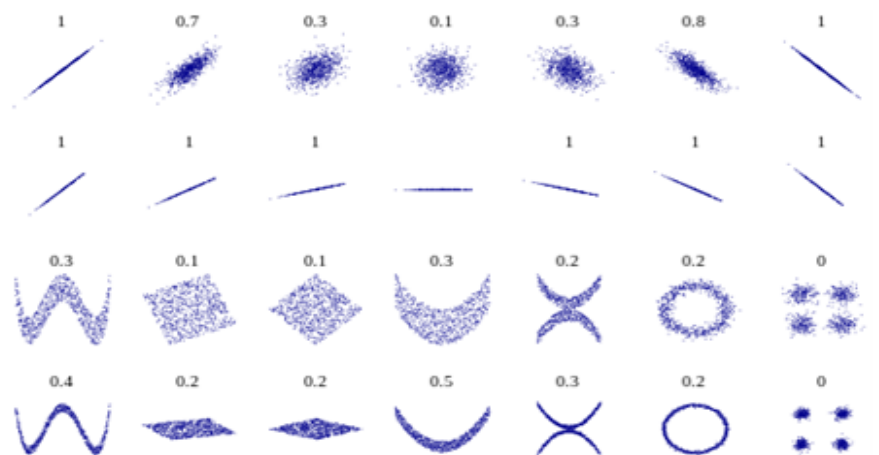
图 3 是一些  $(x,y)$  点集和它们对应的 Pearson 相关系数，图 4 是同样的  $(x,y)$  点集和它们对应的 Brownian 相关系数，认真比较这两个图，可以看出通常的 Pearson 相关系数和 Brownian 相关系数的一些不同。对于 Pearson 相关系数来说，取值范围为  $[-1,1]$ ，从图 3 第一行可以看出，它不仅反映了噪声，还反映了线性相关关系的方向，但从图 3 最后一行可以看出，它对非线性关系的许多方面不能很好的刻画出来，Pearson 相关系数为 0，只能说明不相关，不能说明独立。相对的，从图 4 第一行可以看出，Brownian 相关系数不能够反映出线性关系的方向，但从图 4 第三行可以看出，它能够比较好的反映非线性关系的许多方面。

图 3: Pearson 相关系数范例



资料来源：维基百科，国信证券经济研究所整理

图 4: Brownian 相关系数范例



资料来源：维基百科，国信证券经济研究所整理

## Brownian 相关系数计算

Brownian 相关系数计算方法如下：

给定一对随机向量 $(X, Y)$ 的  $n$  个样本对 $(X_k, Y_k), k = 1, 2, \dots, n$ ，首先，计算所有的两两距离

$$a_{k,l} := \|X_k - X_l\|, \quad b_{k,l} := \|Y_k - Y_l\|, \quad \text{其中 } k, l = 1, 2, \dots, n$$

亦即计算两个 $n \times n$ 的距离矩阵 $(a_{k,l})$ 和 $(b_{k,l})$ ，这里距离取欧式距离即可；然后，计算全部中心化距离 $A_{k,l} := a_{k,l} - \bar{a}_{k.} - \bar{a}_{.l} + \bar{a}_{..}$ 和 $B_{k,l} := b_{k,l} - \bar{b}_{k.} - \bar{b}_{.l} + \bar{b}_{..}$ ，这里 $\bar{a}_{k.}$ 和 $\bar{a}_{.l}$ 分别是 $(a_{k,l})$ 和 $(b_{k,l})$ 的第  $k$  行均值， $\bar{a}_{.l}$ 和 $\bar{b}_{.l}$ 分别是 $(a_{k,l})$  和 $(b_{k,l})$ 的第  $l$  列均值， $\bar{a}_{..}$ 和 $\bar{b}_{..}$ 分别是 $(a_{k,l})$  和 $(b_{k,l})$ 的全矩阵均值。则 $(X, Y)$ 的样本 Brownian 协方差的平方定义为所有乘积和 $A_{k,l}B_{k,l}$ 的算数平均，即：

$$dCov_n^2(X, Y) := \frac{1}{n^2} \sum_{k,l} A_{k,l} B_{k,l}$$

样本 **Brownian** 相关系数定义如下：

$$dCor_n(X, Y) = \frac{dCov_n(X, Y)}{\sqrt{dCov_n(X, X) dCov_n(Y, Y)}}$$

在本篇报告的 **Brownian Correlation** 方法择时模型中， $X$  和  $Y$  都是既包含价格信息也包含成交量信息的 2 维向量， $X$  的第一个维度是窗口区间经中心化的价格，第二个维度是窗口区间经中心化的成交量， $Y$  的第一个维度是与  $X$  等长的历史匹配区间经中心化的价格，第二个维度是历史匹配区间经中心化的成交量。 $dCor_n(X, Y)$  即为  $X$  和  $Y$  的样本 **Brownian** 相关系数。

## 价量联动的形态匹配择时策略

在前面我们已经具体介绍过从历史会重演的角度出发，研究择时策略的大体思路。在需要对指数进行预测时，我们单从交易性数据出发，试图通过回溯历史走势数据，找到在历史上曾经与当前走势较为相近的片段（这个相近可以从各种维度考虑，包括价格、成交量、换手率等信息）。以史为鉴可以知兴替，我们可以借鉴这些历史片段之后的信息，指导我们进行当下的投资。

具体的说，在这篇报告中我们利用 **Brownian Correlation** 方法进行相似区间挖掘的思路制定择时策略，针对某个标的指数或个股，我们首先提取其在观察窗口内的近期价格走势和成交量数据（根据观察频率的不同，可以选择 15 分钟线、30 分钟线、日 K 线等），在这篇报告中，我们经验性的选取 30 的日收盘价走势和日成交量数据。

其次，我们根据 **Brownian Correlation** 方法对标的指数的历史数据进行挖掘，我们按照滑动窗口逐一计算历史上相同长度的区间与窗口区间的 **Brownian** 相关系数和价格序列的 **Pearson** 相关系数，我们将历史区间按相应的 **Brownian** 相关系数从高到低排序，挑出 **Brownian** 相关系数最高且价格序列的 **Pearson** 相关系数为正的 7 个区间，作为和当前待预测的窗口区间指数走势和成交量数据最接近的历史区间。这些历史区间将作为我们对窗口区间未来走势进行预测的重要依据。最后，根据历史片段后期的走势来预测当前窗口区间后期走势。我们简单的使用历史相似片段涨跌幅的平均值作为当前窗口区间往后涨跌的预测。

特别的，为了降低择时策略发出信号的频率，同时提高其预测的准确率。我们可以设置高低两个阈值，当历史相似片段涨跌幅平均值大于高阈值时，我们判断当前窗口区间往后会涨，当历史相似片段涨跌幅平均值小于低阈值时，我们判断当前窗口区间往后会跌，而历史相似片段涨跌幅平均值在高低阈值之间的，我们则不发表预测。假若策略有效，那么加入阈值条件之后，准确率等指标理应有提升。

理论上，如果我们这种用过去相似片段的涨跌幅预测窗口区间未来涨跌的思路是对的，那么用这种方法判断出涨跌的正确率应该高于随机预测的正确率，而且设了高低阈值之后，预测的正确率应该会上升，实证的结果也验证了我们的猜想。在评价的模块中，我们引入随机模型评价的方式，更加客观准确地评价择时策略的真实效果，具体的讨论将在后续展开。另一方面，我们可以调整形态匹配择时策略的参数，譬如观察时间窗口，拟合方案，阈值设定等方面。从提供参考的角度来说，策略可以给出相似的历史片段，供投资者借鉴。



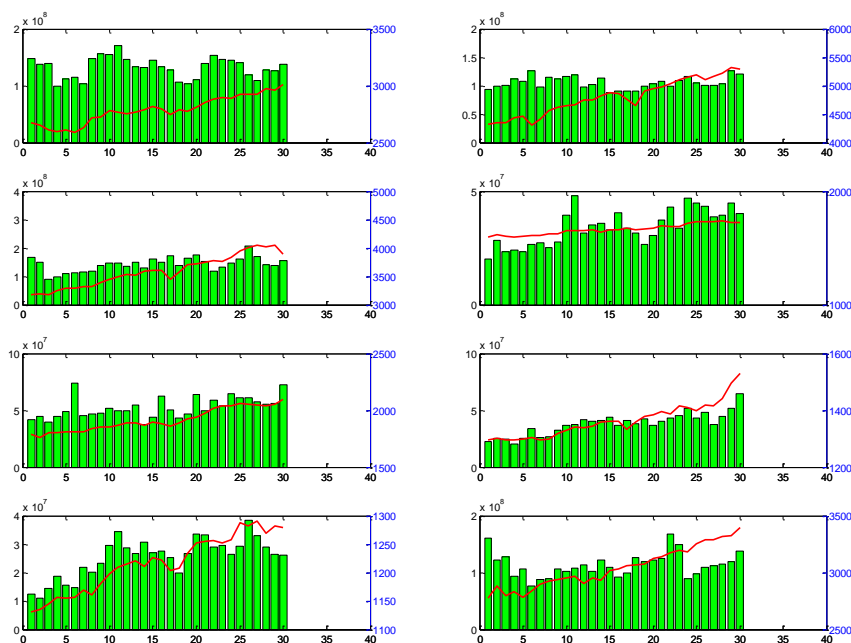
## 择时策略流程实例

为了更好的描述上述基于 Brownian Correlation 方法的择时策略过程，我们以上证指数为例子描述具体的匹配过程和步骤。我们选取总的历史时间段为 2005 年 1 月 4 日到 2012 年 3 月 19 日，将观察窗口设为 30 个交易日，提取其 2009 年 5 月 19 到 2009 年 7 月 1 日的日 K 线收盘价数据和成交量数据。为了避免指数在不同时间区间内绝对数值相差较大的问题，我们将观察时间序列指数的中位数作为标准化因子，进行适当的标准化处理；由于成交量的变化极大，我们先对成交量取对数运算后，再除以标准化因子进行标准化。数据的预处理对于相似度匹配有一定程度的影响，后续我们将对数据预处理的方法进行深入的分析。

其次，我们回溯总的历史时间段（去除观察区间）的指数走势和成交量数据，根据 Brownian Correlation 方法，挑出价格序列的 Pearson 相关系数为正并且 Brownian 相关系数最高的 7 个匹配区间。图 5 展示了观察窗口区间和挖掘的 7 个最类似区间的指数走势和成交量数据图，其中左上角为观察窗口区间（2009 年 5 月 19 到 2009 年 7 月 1 日）。

需要特别说明的是，从技术分析的角度出发，我们有理由认为形态的出现与时间及观察样本均无关。因此，在数据分析的过程中，我们将观察时间窗口数据从原始走势图中剔除，在观察窗口前后的历史数据中挖掘相似形态。通过这种方式可以得到更多的相似片段，提高策略效果。

图 5：上证指数价量联动的形态匹配实例



资料来源：天软科技，国信证券经济研究所整理

相应地，我们也列出了图 5 中 7 个最为相似的历史片段所处的时点信息，以及计算得到的价格、成交量两个维度的 Brownian 相关系数。需要特别指出的是，单从价格维度计算得到的 Pearson 相关系数也保持在较大的数值，再次从价格的维度验证了相似性。

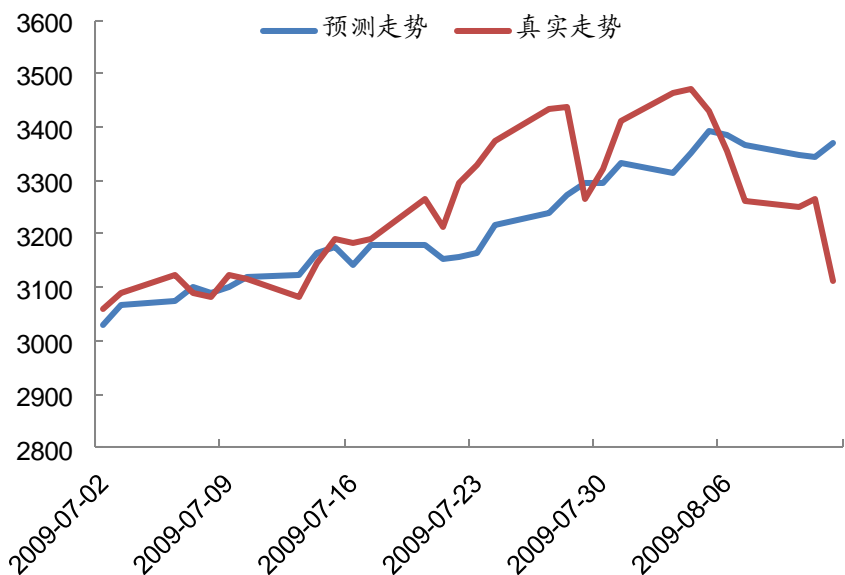
表 1: 相似历史片段时点

开始时点	结束时点	Brownian 相关系数	Pearson 相关系数
2007/7/25	2007/9/4	0.9647	0.9654
2007/3/28	2007/5/15	0.9607	0.9467
2006/8/15	2006/9/25	0.9545	0.9492
2006/10/20	2006/11/30	0.9516	0.9474
2006/3/22	2006/5/9	0.9516	0.9283
2005/12/21	2006/2/13	0.9494	0.9406
2007/2/27	2007/4/9	0.9476	0.9521

资料来源：天软科技，国信证券经济研究所整理

挖掘出 7 个最相近的匹配区间后，我们需要根据挖掘出的历史区间的后期走势来定量预测窗口区间的未来走势。因为窗口区间的长度是 30 个交易日，我们最多也只能预测 30 个交易日的涨跌。为了预测窗口区间未来  $n(1 \leq n \leq 30)$  天的涨跌，我们计算匹配区间对应天数的涨跌幅，简单取所有匹配区间涨跌幅的均值，得到最终的预测结果。在评价时，若涨跌幅均值大于 0，则预测窗口区间未来  $n$  天会涨，如果涨跌幅均值小于 0，则预测窗口区间未来  $n$  天会跌，这是最简单的关于涨跌方向的判断。我们也可以设定阈值，即预测涨跌幅大于某个阈值时发出看涨信号，预测涨跌幅小于某个阈值时发出看跌信号。另一方面，假若预测涨跌幅与真实涨跌幅的差距越小，则表明预测效果越好。同时，我们借鉴了上一篇报告的随机评价方式，更全面地评价了策略。

图 6: 预测走势与真实走势对比图（2009 年 7 月 2 到 2009 年 8 月 12 日）



资料来源：天软科技，国信证券经济研究所整理

针对图 5 中的示例，我们给出了在 2009 年 7 月 1 日之后的真实走势与预测走势的对比，如图 6 所示。图中我们不难观察到，基于价量联动的预测方法可以较好地给出后期走势的预报，取得了不错的效果。

对于每一个时点下的择时预测，我们重复上述策略过程，得到每个时点下对于标的指数的后期预测结果。接下来，我们可以利用标的指数的真实走势评价预测结果的效果。在下一节中我们将对价量联动的形态匹配择时策略进行检验，并给出一些定量和定性的分析结果。

## 择时策略效果评价及分析

择时策略最主要关心是对于不同周期下涨跌方向的判断，假设我们采用抛硬币的方式做判断，那么应取得差不多 50% 的准确率。假若择时策略从统计意义上可以取得大于 50% 的准确率，那么认为其有效。此外，在判断涨跌方向的基础上，我们关心其预测的幅度，即涨跌的绝对值。理想情况下，如果预测涨跌幅与真实涨跌幅越相近，则择时效果越好。如无特殊说明，后续择时策略的起点为 2005 年 1 月 4 日，终点为 2012 年 3 月 19 日，共计 1750 个交易日。具体地，我们从几个方面来评价择时策略的效果：

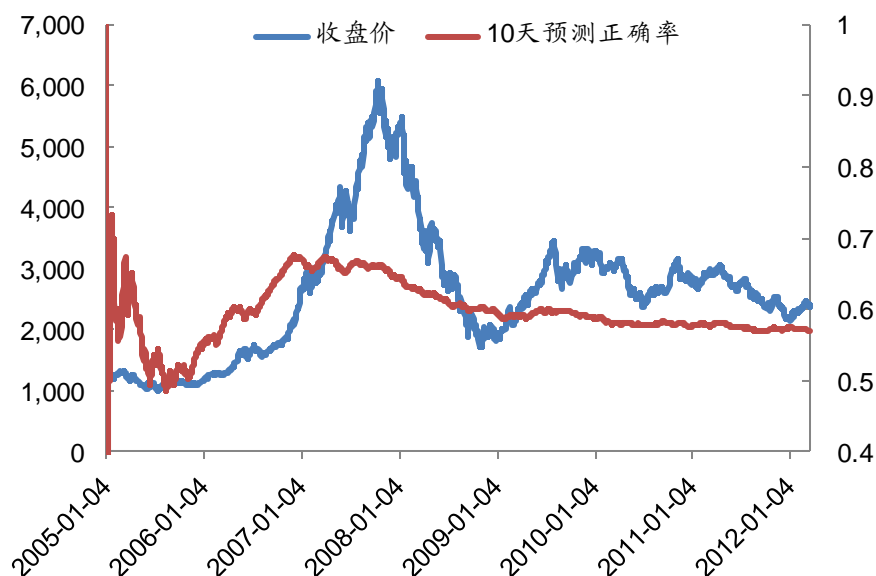
1. 涨跌方向判断的准确率：即看涨看跌方向的判断。假若策略给出  $n$  天后涨幅大于 0，则预测上涨，反之预测下跌。这种评价方式的缺点在于忽略了预测值的相对大小，即预测涨幅 10% 和预测涨幅 0.1% 被同样视作上涨。
2. 预测涨跌幅与真实对比：对于第一种评价方式的改进，我们不仅关心预测的方向，同时也考虑预测涨跌幅与真实涨跌幅的差异。例如，策略预测涨幅 0.1% 而指数实际下跌 0.1%，在第一种评价方式看来，这属于判断错误。但从相差的绝对数值来看，则认为预测效果较好。
3. 随机评价胜率：在我们之前的报告中，提出过一种基于随机模型绩效评估的方法，可以更科学合理地评价效果。具体地说，在某个择时时点，我们利用价量联动的形态匹配择时策略给出后期标的指数的走势预测，并计算预测结果与真实结果的相似性。另一方面，我们也可以采用随机的方式预测标的指数的后期走势，得到每一次随机预测的效果以及其与真实结果的相似性。进而我们可以计算择时策略预测结果的相似性在随机结果中的百分位排名，假若择时策略有效，那么其应该可以战胜大部分的随机结果。我们将 50% 作为界定胜率的界限，即：择时策略可以战胜一半以上的随机结果则认为有效。
4. 择时策略分年效果：由于市场一直处于变化的过程中，我们期待策略在各种市场环境下的效果相近，尽量可以做到市场中性。但一般而言，这是比较困难的，从后面的结果我们可以看到，在 08 年单边下跌的市场中，择时策略表现较差。究其原因，这是因为 08 年市场单边下跌的走势无法从历史数据中找到相似的片段，导致策略失效。
5. 加入阈值修正涨跌方向判断：直观上来说，如果择时策略预测后期上涨幅度为 0.1%，那么这个时候我们做多市场的把握并不大。相反，如果策略预测上涨幅度为 10%，看涨的信心将会更充分。换言之，我们会对涨跌方向的判断加入阈值的处理，当预测上涨幅度大于某个阈值时，我们才发出看涨信号。这样会在一定程度上减少信号的触发频率，但是会提高策略的准确率。

下面我们将具体地对上述的评价方法展开论述。

### 涨跌方向判断准确率

首先，我们针对最简单的情况对策略进行分析，即判断涨跌方向的准确率。具体地说，我们观察上证指数 30 个交易日的走势，预测其 10 个交易日后的涨跌情况，如图 7 所示。从图 7 中我们不难看出，关于涨跌方向判断的准确率大致收敛在 57% 附近（除去策略起始点 2005 年 1 月 4 日附近的样本）。同时，我们也列出了上证指数对应的收盘价，从图中我们观察到：在 2006 年初的大牛市起涨时，择时策略关于涨跌方向判断的准确率稳步上升，较好地抓住了牛市的起点。到 2007 年 10 月附近，指数开始单边下跌，策略的准确率也开始下降，主要的原因是从历史数据中较难发现相似的持续下跌的历史片段。

图 7：涨跌方向预测准确率与上证指数走势

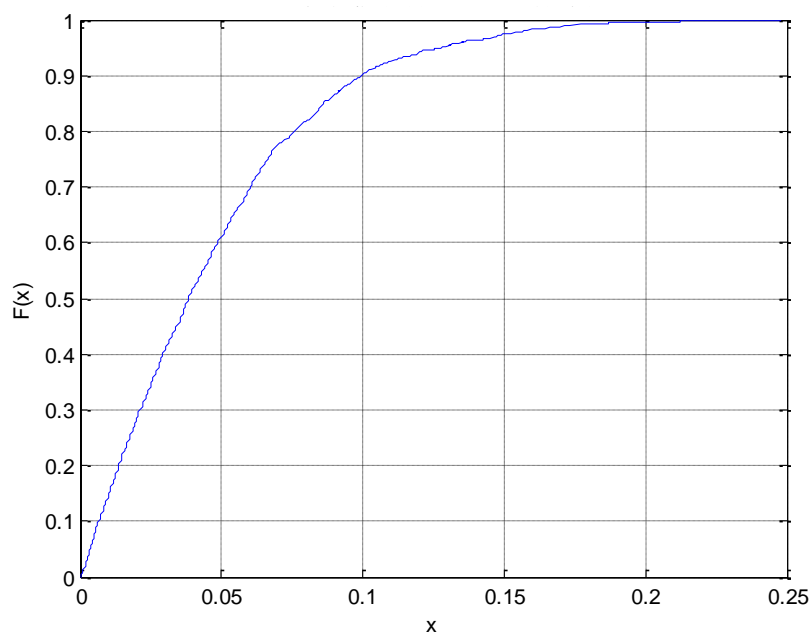


资料来源：天软科技，国信证券经济研究所整理

### 预测涨跌幅与真实对比

如前所述，我们同时也考虑预测涨跌幅与真实涨跌幅的绝对值差异，差异值越小，证明预测效果越好。我们观察上证指数 30 个交易日的走势，计算 10 日涨跌幅预测值与真实结果的差异。1750 个交易日涨跌幅预测误差的经验分布如图 8 所示。我们可以看到，50%（875 个交易日）的预测误差在 3.83% 以内，60%（1050 个交易日）的预测误差在 4.86% 以内，预测误差在可接受范围内。

图 8：10 日涨跌幅预测误差经验分布图



资料来源：天软科技，国信证券经济研究所整理

### 随机评价胜率

我们在之前的报告中引入了随机模型评价的方式，这种评价方式可以更科学合理地评价策略的效果，避免市场因素的影响。具体地说，对于每一个时点，按照价量联动的形态匹配择时策略我们给出指数后期走势的预测，并计算其与真实走势的相似性。相似性越大，则说明预测结果越好。另一方面，在每一个时点下我们也可以随机地预测后期走势。如果策略有效，那么其理应战胜大部分的随机预测结果。以 50% 作为界定胜率的界限，计算择时策略在 1750 个交易日内战胜随机结果的次数。实证结果表明，共有 1330 个交易日内择时策略战胜随机预测结果，即胜率为 76%。

### 择时策略分年统计效果

在图 7 中我们分析了择时策略 10 天涨跌方向判断的准确率，为了更全面分析策略在不同市场环境下的效果，我们在表 2 中列出了策略在不同年份，分别在 10 天、20 天、30 天时涨跌方向预测的正确率。从表 2 中的结果中我们不难看出，择时策略在 2006 年的表现最优，在不同时间长度下的准确率均在 75% 左右。在 2008 年表现最差，这与图 7 中的结论是相符的。市场在 2008 年出现了单边的下跌，从历史会重演的角度来说，站在当时的时点，我们并不能从历史的角度找到可以借鉴的参考。总体而言，择时策略在 10 天、20 天、30 天时，关于涨跌方向判断的准确率在 57% 左右。

表 2：涨跌方向预测分年正确率统计

年份	10 天涨跌方向预测准确率	20 天涨跌方向预测准确率	30 天涨跌方向预测准确率
2005	54.96%	60.74%	58.68%
2006	79.25%	74.69%	76.76%
2007	59.50%	60.74%	62.40%
2008	45.12%	54.47%	60.16%
2009	55.33%	53.28%	47.54%
2010	52.07%	50.00%	50.00%
2011	56.56%	57.38%	51.23%
总体	56.91%	57.49%	56.97%

资料来源：天软科技，国信证券经济研究所整理

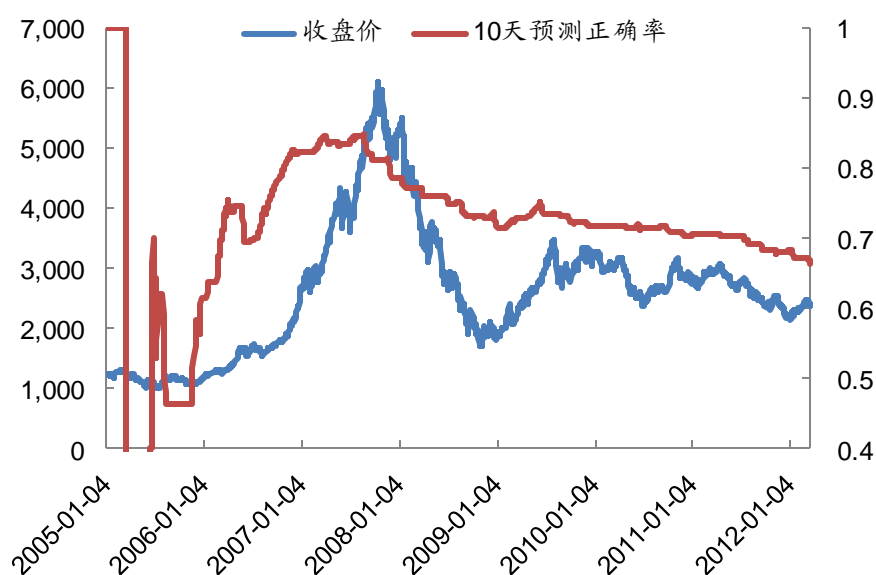
### 加入阈值判断涨跌方向

在具体预测操作时，我们也需要考虑预测涨跌幅度的判断。例如，择时策略预测指数 10 日内涨幅为 10% 和 0.1%，从方向上说都是上涨，但从幅度的角度来说，显然 10% 的涨幅显得更有把握。因此，我们也在判断方向涨跌方向时加入阈值的考虑，即认为涨幅大于某个高阈值时，策略才发出看涨信号；当认为涨幅小于某个低阈值时，策略才发出看跌信号；在高低阈值间的涨幅不做判断。不难看出，带阈值的方法其实是对信号的触发加入了更为严格的条件，可以看作是对图 7 中数据的抽样。择时策略并不对每个时点都发表意见，仅当出现较为明显的上涨或下跌趋势时才给出方向的判断。通过参数优化，在预测 10 天涨跌幅度时，高阈值设置为 2.71%，低阈值设置为 -6.38%。设置了阈值之后，策略关于涨跌方向判断的准确率如图 9 所示。

从图 9 中我们可以观察到，经过阈值的设置后，择时策略的准确率有进一步的提高，从之前的 57% 提升至 66%，触发信号的次数从 1750 次减少至 417 次。同时，我们再次观察到，择时策略的准确率在 2006 年牛市起点时稳步上升，直到指数见顶后，在 2008 年有所回落。在后续的时间段内，基本稳定在接近 70% 左右，2009 年的小牛市有所提升。2010 年和 2011 年的市场环境中表现基本稳定。图 9 的结果再次验证了择时策略的有效性。



图 9: 带阈值的涨跌方向预测准确率

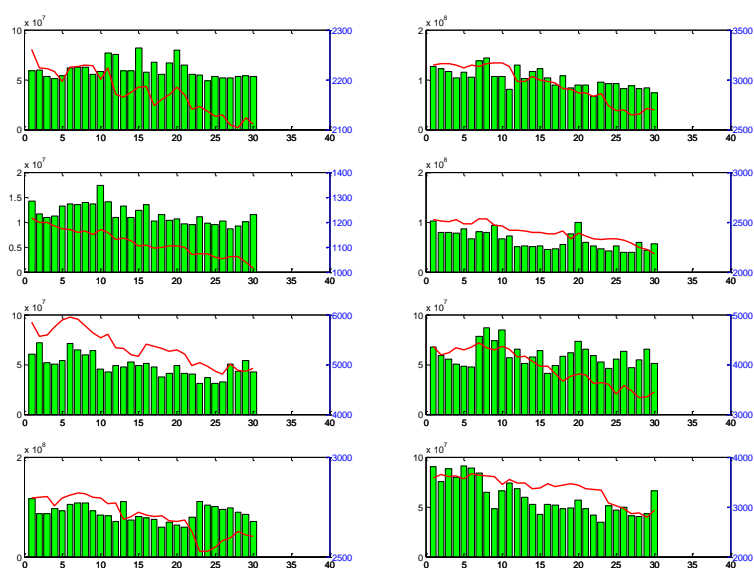


资料来源：天软科技，国信证券经济研究所整理

### 我们现在在哪里

截至最新交易日（2012年8月2日），我们提取上证指数最新30个交易日的价格和成交量数据，并对其历史数据进行回溯。得到最为相近的7个历史片段如图10所示（左上角为观察窗口数据，其余为历史片段数据）。这7个历史片段的时点总结在表3中。最后，我们也根据相似的历史片段对于后期走势给出了预测，如图11所示。从预测结果看来，指数存在反弹的可能性，真实的走势有待观察和分析。

图 10: 上证指数价量联动形态匹配（截至2012年8月2日）



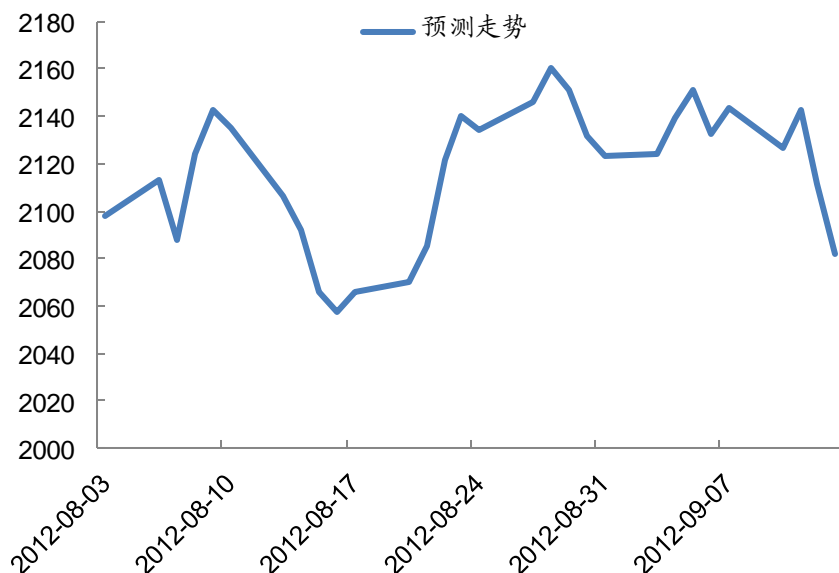
资料来源：天软科技，国信证券经济研究所整理

表 3: 最新走势相似历史片段时点

开始时点	结束时点	Brownian 相关系数	Pearson 相关系数
2010/4/1	2010/5/14	0.9430	0.9424
2005/4/15	2005/6/2	0.9343	0.9335
2011/11/4	2011/12/15	0.9297	0.9181
2007/10/24	2007/12/4	0.9248	0.9108
2008/2/22	2008/4/3	0.9208	0.9273
2011/7/7	2011/8/17	0.9200	0.8769
2008/5/7	2008/6/18	0.9128	0.8939

资料来源：天软科技，国信证券经济研究所整理

图 11: 上证指数后期走势预测（截至 2012 年 8 月 2 日）



资料来源：天软科技，国信证券经济研究所整理

## 总结和扩展讨论

在本篇报告中，我们从技术分析中历史会重演的角度出发，提出了一种价量联动的形态匹配择时策略。在给定标的指数观察期内走势的前提下，我们利用 Brownian 相关系数的方法对指数的历史数据进行回溯，同时考虑价格和成交量两个方面，挖掘出历史上与观察期间内较为相似的形态，并利用历史上的相似形态后期的走势对目前观察期后期走势进行预测。

实证结果表明，同时考虑价格和成交量可以进一步提高择时策略的效果，以观察 30 个交易日，预测 10 个交易日为参数设置，策略判断涨跌方向的准确率为 57%；与随机预测结果相比，策略的胜率为 76%；预测涨跌幅与真实涨跌幅差异在可接受范围内。在加入了阈值的设置后，策略的准确率得到进一步提高，在各种评价方式下均取得了不错的效果。

目前，我们的工作量化技术分析的研究框架下，试图通过对交易性数据的挖掘，总结股票或指数在运行过程中形成的各种模式。后续我们将沿着这个思路继续研究，未来的工作主要包含以下几个方面：

1. 对于择时策略进行更为全面的扩展。从利用的数据出发，我们可以很方便地将目前的择时策略扩展至选股和行业配置策略。只要将指数的数据替换成个股或者行业指数的数据即可。历史会重演不仅表现在指数上，也同时在个股上得到体现，后续我们将检验择时策略的扩展至选股等方面的效果。
2. 加入成交量后的形态库整理。在我们之前的报告中就提到过关于形态库的构建，但更多讨论的是价格方面的形态。从本篇报告后，我们将更多的加入成交量的影响，最大限度地挖掘价量联动背后的宝藏，并形成独特的投资策略。

## 国信证券投资评级

类别	级别	定义
股票 投资评级	推荐	预计 6 个月内，股价表现优于市场指数 20%以上
	谨慎推荐	预计 6 个月内，股价表现优于市场指数 10%-20%之间
	中性	预计 6 个月内，股价表现介于市场指数 $\pm 10\%$ 之间
	回避	预计 6 个月内，股价表现弱于市场指数 10%以上
行业 投资评级	推荐	预计 6 个月内，行业指数表现优于市场指数 10%以上
	谨慎推荐	预计 6 个月内，行业指数表现优于市场指数 5%-10%之间
	中性	预计 6 个月内，行业指数表现介于市场指数 $\pm 5\%$ 之间
	回避	预计 6 个月内，行业指数表现弱于市场指数 5%以上

## 分析师承诺

作者保证报告所采用的数据均来自合规渠道，分析逻辑基于本人的职业理解，通过合理判断并得出结论，力求客观、公正，结论不受任何第三方的授意、影响，特此声明。

## 风险提示

本报告版权归国信证券股份有限公司（以下简称“我公司”）所有，仅供我公司客户使用。未经书面许可任何机构和个人不得以任何形式使用、复制或传播。任何有关本报告的摘要或节选都不代表本报告正式完整的观点，一切须以我公司向客户发布的本报告完整版本为准。本报告基于已公开的资料或信息撰写，但我公司不保证该资料及信息的完整性、准确性。本报告所载的信息、资料、建议及推测仅反映我公司于本报告公开发布当日的判断，在不同时期，我公司可能撰写并发布与本报告所载资料、建议及推测不一致的报告。我公司或关联机构可能会持有本报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行业务服务。我公司不保证本报告所含信息及资料处于最新状态；我公司将随时补充、更新和修订有关信息及资料，但不保证及时公开发布。

## 证券投资咨询业务的说明

证券投资咨询业务是指取得监管部门颁发的相关资格的机构及其咨询人员为证券投资者或客户提供证券投资的相关信息、分析、预测或建议，并直接或间接收取服务费用的活动。

证券研究报告是证券投资咨询业务的一种基本形式，指证券公司、证券投资咨询机构对证券及证券相关产品的价值、市场走势或者相关影响因素进行分析，形成证券估值、投资评级等投资分析意见，制作证券研究报告，并向客户发布的行为。



### 国信证券经济研究所团队成员

<b>宏观</b>			<b>固定收益</b>			<b>策略</b>		
周炳林	0755-82130638		侯慧梯	021-60875161		黄学军	021-60933142	
崔 嵘	021-60933159		赵 婧	021-60875168		林丽梅	021-60933157	
张 娜	0755-82133259					技术分析		
李智能						闫 莉	010-88005316	
<b>交通运输</b>			<b>机械</b>			<b>商业贸易</b>		
郑 武	0755-82130422		郑 武	0755-82130422		孙菲菲	0755-82130722	
陈建生	0755-82133766		陈 玲	021-60875162		常 伟	0755-82131528	
岳 鑫	0755-82130432		杨 森	0755-82133343				
糜怀清	021-60933167		后立尧	010-88005327				
<b>汽车及零配件</b>			<b>钢铁及新材料</b>			<b>房地产</b>		
左 涛	021-60933164		郑 东	010-66025270		区瑞明	0755-82130678	
			陈 健	010-88005308		黄道立	0755-82133397	
						刘 宏	0755-22940109	
<b>基础化工及石化</b>			<b>医药</b>			<b>计算机及电子元器件</b>		
刘旭明	010-66025272		贺平鸽	0755-82133396		段迎晨	0755-82130761	
张栋梁	0755-82130532		丁 丹	0755-82139908		高耀华	010-88005321	
吴琳琳	0755-82130833-1867		杜佐远	0755-82130473		刘 翔	021-60875160	
罗 洋	0755-82150633		胡博新	0755-82133263		欧阳仕华	0755-82151833	
朱振坤	010-88005317		刘 勍	0755-82133400				
<b>传媒</b>			<b>有色金属</b>			<b>电力及公共事业</b>		
陈财茂	010-88005322		彭 波	0755-82133909		谢达成	021-60933161	
刘 明	010-88005319		龙 飞	0755-82133920				
			徐张红	0755-22940289				
<b>非银行金融</b>			<b>轻工</b>			<b>建筑工程及建材</b>		
邵子钦	0755-82130468		李世新	0755-82130565		邱 波	0755-82133390	
田 良	0755-82130470		邵 达	0755-82130706		刘 萍	0755-82130678	
童成墩	0755-82130513					马 彦	010-88005304	
<b>家电及通信</b>			<b>电力设备与新能源</b>			<b>食品饮料</b>		
王念春	0755-82130407		杨敬梅	021-60933160		黄 茂	0755-82138922	
			张 弢	010-88005311				
<b>旅游</b>			<b>农业</b>			<b>金融工程</b>		
曾 光	0755-82150809		杨天明	021-60875165		戴 军	0755-82133129	
钟 潇	0755-82132098					林晓明	0755-82136165	
						秦国文	0755-82133528	
						张璐楠	0755-82130833-1379	
						周 琦	0755-82133568	
						郑亚斌	021-60933150	
						陈志岗	0755-82136165	
						马璞清	0755-22940643	
						吴子昱	0755-22940607	
<b>基金评价与研究</b>								
杨 涛	0755-82133339							
康 亢	010-66026337							
李 腾	010-88005310							
刘 洋	0755-82150566							
潘小果	0755-82130843							
蔡乐祥	0755-82130833-1368							
钱 晶	0755-82130833-1367							





### 国信证券机构销售团队

华北区（机构销售一部）			华东区（机构销售二部）			华南区（机构销售三部）		
王立法	010-66026352 13910524551 wanglf@guosen.com.cn		盛建平	021-60875169 15821778133 shengjp@guosen.com.cn		魏宁	0755-82133492 13823515980 weining@guosen.com.cn	
王晓健	010-66026342 13701099132 wangxj@guosen.com.cn		马小丹	021-60875172 13801832154 maxd@guosen.com.cn		邵燕芳	0755-82133148 13480668226 shaoyf@guosen.com.cn	
焦戡	010-66026343 13601094018 jiaojian@guosen.com.cn		郑毅	021-60875171 13795229060 zhengyi@guosen.com.cn		段莉娟	0755-82130509 18675575010 duanlj@guosen.com.cn	
李文英	010-88005334 13910793700 liwying@guosen.com.cn		黄胜蓝	021-60875166 13761873797 huangsl@guosen.com.cn		郑灿	0755-82133043 13421837630 zhengcan@guosen.com.cn	
原玮	010-88005332 15910551936 yuanyi@guosen.com.cn		孔华强	021-60875170 13681669123 konghq@guosen.com.cn		王昊文	0755-82130818 18925287888 wanghaow@guosen.com.cn	
赵海英	010-66025249 13810917275 zhaohy@guosen.com.cn		叶琳菲	021-60875178 13817758288 yelf@guosen.com.cn		甘墨	0755-82133456 15013851021 ganmo@guosen.com.cn	
甄艺	010-66020272 18611847166		崔鸿杰	021-60933166 13817738250 cuihj@guosen.com.cn		徐冉	0755-82130655 13923458266 xuran1@guosen.com.cn	
杨柳	18601241651 yangliu@guosen.com.cn		李佩	021-60875173 13651693363 lipei@guosen.com.cn		颜小燕	0755-82133147 13590436977 yanxy@guosen.com.cn	
			刘塑	021-60875177 13817906789 liusu@guosen.com.cn		林莉	0755-82133197 13824397011 linli2@guosen.com.cn	
			汤静文	021-60875164 13636399097 tangjingwen@guosen.com.cn		赵晓曦	0755-82134356 15999667170 zhaoxxi@guosen.com.cn	
			梁轶聪	021-60873149 18601679992 liangyc@guosen.com.cn				