

传统算法交易策略中的相关参数研究

——算法交易系列研究之二

安宁宁 资深分析师
电话: 0755-23948352
eMail: ann@gf.com.cn
执业编号: S0260512020003

算法交易概述

算法交易是实现对大规模母单进行拆分,并对拆分后的子单进行定时、定量交易的一种程序化交易方式。算法交易通过事先设定好的策略,由投资者编制完成相关的计算机自动化交易程序,并通过连接交易系统接口,利用计算机实现大规模交易的拆分、报价、下单、撤单等一系列动作,并在交易后对已完成的交易进行分析和评估,从反馈中进一步修正算法模型。算法交易产生的根本目的,在于其可以减小市场摩擦,有效降低交易中的冲击成本,从而使得整个交易可以以最优价格完成。目前,国内外使用算法交易的投资者主要是各类机构投资者,包括基金公司、保险公司、养老金、投资银行以及各类资产管理机构。由于国内证券市场起步较晚,算法交易还没有大规模普及,但在海外发达金融市场,算法交易已成为一种成熟的证券交易模式。在本篇报告中,我们主要对几类传统的算法模型(策略)进行研究,包括 TWAP 策略、VWAP 策略,以及 MVWAP 策略,并着重对策略中的一些细节进行思考与分析,特别对实际操作中可能遇到的几类参数进行了实证研究。

传统算法交易策略中的相关参数研究

通过分析算法交易的整个流程,我们认为在传统算法策略中,有三类参数需要进行重点研究。一类是预测成交量(成交量比例)参数,包括成交量历史样本天数和在预测成交量(成交量比例)过程中历史样本(每个交易日)的权重;另一类是订单拆分参数,包括 VWAP (MVWAP) 策略的分段周期和下单频率;最后一类是下单量参数,包括相对于预测成交量的实际下单比例和 MVWAP 策略中下单量的增减比例参数。

预测成交量(成交量比例)参数

通过对不同的市场环境进行回测,我们发现针对国内 A 股市场,历史样本天数取 50 至 70 个交易日较为合适。而历史样本权重则与欧美一些投资研究中所指出的权重随时间间隔增大而递减的情况不同,在预测成交量(成交量比例)分布时,我们建议采用简单算术平均方法即可。

订单拆分参数

通过回测跟踪偏差和成交执行率两个指标,我们发现取 2 分钟至 10 分钟的 VWAP (MVWAP) 分段周期对策略结果的影响不大,因此在实际操作过程中可以在此范围内任意选择;下单频率则是对于 TWAP 和 VWAP (MVWAP) 通用的参数,通过研究,我们推导出了下单频率的一个表达式,其与股票的交易时间、流通股本、流动性、历史成交量分布特征、交易者对下单比例和成交量分布拟合误差的要求等有关。

下单量参数

通过盘口数据测算,我们认为相对于预测成交量的实际下单比例不应该超过 40%,并且更好的办法是针对不同个股测算该参数的具体数值。根据这个参数的结果,可以进一步确定 MVWAP 下单量增减比例这个参数,我们给出了增加比例参数上限的表达式,并针对股票在某个交易日出现价格持续高于或低于市场实时 VWAP 的情况给出了一些定性建议。

目录索引

| | |
|-------------------------|----|
| 一、算法交易概述..... | 3 |
| 二、传统算法交易策略介绍..... | 3 |
| (一) TWAP 策略..... | 3 |
| (二) VWAP 策略..... | 4 |
| (三) MVWAP 策略..... | 5 |
| 三、传统算法交易策略参数的实证研究..... | 5 |
| (一) 预测成交量（成交量比例）参数..... | 5 |
| (二) 订单拆分参数..... | 9 |
| (三) 下单量参数..... | 11 |
| 四、总结..... | 13 |

图表索引

| | |
|---|----|
| 图 1: 各股票预测偏差 \bar{r} 在不同市场环境中的表现..... | 7 |
| 图 2: VWAP 历史样本权重曲线..... | 8 |
| 表 1: 传统算法实证股票样本（截止至 2012.6.11.）..... | 6 |
| 表 2: 算术平均与加权平均预测 \bar{r} 结果比较..... | 8 |
| 表 3: 各种市场环境中不同分段周期投资组合的 MPE 与 VFA | 10 |

一、算法交易概述

在金融市场中对股票、期货等金融产品进行交易，当交易规模很大时，容易带来较大的冲击成本。算法交易是实现大规模订单进行拆分，并对拆分后的子单进行定时、定量交易的一种程序化交易方式。算法交易通过事先设定好的策略，由投资者编制完成相关的计算机自动化交易程序，并通过连接交易系统接口，利用计算机实现大规模交易的拆分、报价、下单、撤单等一系列动作，并在交易后对已完成的交易进行分析和评估，从反馈中进一步修正算法模型。

算法交易产生的根本目的，在于其可以减小市场摩擦，有效降低交易中的冲击成本，从而使得整个交易可以以最优价格完成。同时，算法交易可以在降低人力成本的同时，通过自动化下单的方式，有效提高交易的执行效率。另外，针对大规模交易，算法交易可以通过一些策略有效隐藏交易行为，从而避免竞争对手根据自己的“套路”出牌。

目前，国内外使用算法交易的投资者主要是各类机构投资者，包括基金公司、保险公司、养老金、投资银行以及各类资产管理机构。由于国内证券市场起步较晚，算法交易还没有大规模普及，但在海外发达金融市场，算法交易已成为一种成熟的证券交易模式。

算法交易的基本流程包括算法模型（策略）研究、交易系统的设计与开发、交易的执行、交易后分析等。在本篇报告中，我们主要对几类传统的算法模型（策略）进行研究，包括 TWAP（Time Weighted Average Price）策略、VWAP（Volume Weighted Average Price）策略，以及 MVWAP（Modified Volume Weighted Average Price）策略，并着重对策略中的一些细节进行思考与分析，特别对实际操作中可能遇到的几类参数进行了实证研究。

二、传统算法交易策略介绍

（一）TWAP策略

TWAP（Time Weighted Average Price），时间加权平均价格算法，是一种最简单的传统算法交易策略。该模型将交易时间进行均匀分割，并在每个分割节点上将拆分的订单进行提交。例如，可以将某个交易日的交易时间平均分为 N 段，TWAP 策略会将该交易日需要执行的订单均匀分配在这 N 个时间段上去执行，从而使得交易均价跟踪 TWAP

$$TWAP = \frac{\sum_{t=1}^N price_t}{N} \quad (1)$$

TWAP 策略设计的目的是在使交易对市场影响最小化的同时提供一个较低的平均成交价格，从而达到减小交易成本的目的。在分时成交量无法准确估计的情况下，该模型可以较好地实现算法交易的基本目的。但是使用 TWAP 过程中的一个问题是，在订单

规模很大的情况下，均匀分配到每个节点上的下单量仍然较为可观，仍有可能对市场造成一定的冲击。另一方面，真实市场的成交量总是在波动变化的，将所有的订单均匀分配到每个节点上显然是不够合理的。因此，算法交易研究人员很快建立了基于成交量变动预测的 VWAP 模型。不过，由于 TWAP 操作和理解起来非常简单，因此其对于流动性较好的市场和订单规模较小的交易仍然适用。

（二）VWAP策略

VWAP (Volume Weighted Average Price)，成交量加权平均价格算法，是目前市场上最为流行的算法交易策略之一，也是很多其它算法交易模型的原型。首先定义 VWAP，它是一段时间内证券价格按成交量加权的平均值

$$VWAP = \frac{\sum_t price_t \times volume_t}{\sum_t volume_t} \quad (2)$$

其中 $price_t$ 和 $volume_t$ 分别是某个时点上证券的成交价格 and 成交量。VWAP 是对一段时间市场上所有交易活动平均价格的衡量。

VWAP 算法交易策略的目的就是尽可能地使订单拆分成成交所得的 $VWAP_{成交}$ 盯住市场的 $VWAP_{市场}$ 。从 VWAP 的定义(2)式来看，若希望能够使得 $VWAP_{成交}$ 跟住 $VWAP_{市场}$ ，则需要将拆分订单按照市场真实的成交量分时比例进行提交，这就需要对市场分时成交量（成交量比例）进行预测。通常来说，VWAP 策略会使用过去 M 个交易日分段成交量的平均值作为预测成交量，这里就要涉及到 M 和平均权数的确定，我们将在后文进行讨论。更为严格地说，假设需要在某段时间买入或卖出一定数量的股票，采用算法交易将这段时间分为 N 部分，并预测每部分时间的成交量比例（占当天所需交易量）为 vp_i ，而市场真实的分段成交比例（占当天市场真实成交量）为 vm_i ，市场在每个时点的成交价格为 P_i ，则可以定义跟踪误差

$$TE = \sum_i P_i \times (vm_i - vx_i) \quad \begin{cases} vx_i = vp_i & \text{if } vm_i > vp_i \\ vx_i = vm_i & \text{if } vm_i \leq vp_i \end{cases} \quad (3)$$

从 TE 的定义(3)式可以看出两点：1、跟踪误差与成交量预测的关系非常紧密，预测结果的好坏直接影响到 VWAP 算法交易的结果。2、当某段时间的 vp_i 超过市场真实 vm_i 时，有可能造成订单无法全部成交，这样就会造成算法交易执行效率的下降，因此，更为常用的是被称为“带反馈的”VWAP 算法交易策略。

所谓带反馈的 VWAP 算法交易策略，是指在 VWAP 策略的基础之上，将每个分段未成交的订单按分摊至后面的时间段中，这样可以有效提高成交执行率。之前所讨论的 TWAP 策略也可以采用该类反馈技术，使流动性相对较弱股票的成交比率大幅提升。

（三）MVWAP策略

MVWAP (Modified Volume Weighted Average Price)，成交量加权平均价格优化算法。在原始 VWAP 的基础之上有很多优化和改进的算法，最为常见的一种策略是根据市场实时价格和 $VWAP_{\text{市场}}$ 的关系，对下单量的大小进行调整与控制，我们统一将这一类算法称为 MVWAP。

当市场实时价格小于此时的 $VWAP_{\text{市场}}$ 时，在原有计划交易量的基础上进行放大，如果能够将放大的部分成交或部分成交，则有助于降低 $VWAP_{\text{成交}}$ ；反之，当市场实时价格大于此时的 $VWAP_{\text{市场}}$ 时，在原有计划交易量的基础上进行缩减，也有助于降低 $VWAP_{\text{成交}}$ ，从而达到控制交易成本的目的。

在MVWAP策略中，除了对成交量做出定量预测之外（通常也是按照历史成交量取平均进行预测），同样重要的是对于交易量放大或减小的定量控制。一种简单的办法是在市场实时价格低于或高于 $VWAP_{\text{市场}}$ 时，将下一时段的下单量按固定比例放大或缩小，那么这个比例参数就存在一个最优解的问题。如果考虑得更为复杂和细致，这个比例还可以是一个随价格偏差（市场实时价格与 $VWAP_{\text{市场}}$ 之差）变化的函数。在后文中，我们主要对前一种较为简单的方法（固定成交量缩放比例）进行定量分析。

三、传统算法交易策略参数的实证研究

通过分析算法交易的整个流程，我们认为在传统算法策略中，有三类参数需要进行重点研究。一类是预测成交量（成交量比例）参数，包括成交量历史样本天数和在预测成交量（成交量比例）过程中历史样本（每个交易日）的权重；另一类是订单拆分参数，包括 VWAP (MVWAP) 策略的分段周期和下单频率；最后一类是下单量参数，包括相对于预测成交量的实际下单比例和 MVWAP 策略中下单量的增减比例参数。接下来我们分别对这三类六个参数进行实证研究。

（一）预测成交量（成交量比例）参数

在 VWAP (MVWAP) 策略、以及需要控制下单量的 TWAP 策略中，都需要对拟交易日的分段成交量或成交量比例进行预测。从之前关于 VWAP 策略介绍的分析当中可以看出，成交量预测准确与否直接决定了 VWAP 策略的执行效果。

预测成交量或成交量比例，按照传统的算法交易策略，一般是对过去 N 个交易日的分段成交量求平均，从而得到最新一个交易日的成交量（成交量比例）预测结果。这样的话，就会涉及到两个参数，一个是 N ，另一个是这 N 个交易日成交量求平均时的权重。

1、历史样本天数

长期来看，每日的成交量总是在波动的。在传统算法交易策略中，某个交易日的成交量情况可以通过前 N 个交易日成交量的加权平均来进行预测。为了研究历史样本天数 N 这个参数，我们首先将各交易日成交量的权重设定为等权重，待确定 N 后再对权重进行进一步研究。

如果 N 取得太小，样本短期内的波动可能引起较大的预测偏差；如果 N 取得太大，由于在较长的时间中市场环境已经发生了较大的变化，远离预测日的交易量情况可能对该预测日不具有实际意义，并且还可能给预测带来偏差。因此，我们需要为 N 确定一个合适的范围，从而使得对过去的成交量数据进行平均后，所得到的预测结果最为准确。

这里我们以预测成交量比例为例进行研究。通过扫描不同的 N ，可以得到不同的成交量比例预测序列 $\{V_p\}$ ，通过与该交易日真实成交量比例 $\{V_r\}$ 进行比较，就可以看出预测结果是否准确。具体来讲，我们将每个交易日均分为 48 段，每段 5 分钟。通过某个交易日之前 N 个交易日分段成交量比例的数据，我们可以由算数平均得到一个成交量比例预测序列 $\{V_p\}$ ；而该交易日的真实成交量比例序列为 $\{V_r\}$ 。通过最小二乘法，我们可以得到序列 $\{V_p\}$ 与序列 $\{V_r\}$ 的距离平方和

$$r = \sum_{i=1}^{48} (V_{pi} - V_{ri})^2 \quad (4)$$

这个值越小，说明预测越为准确。一个交易日的 r 并不能说明问题，因此我们在实证计算中，取连续 5 个交易日的平均值 \bar{r} 作为结果，即通过扫描不同的 N ，得到 \bar{r} 的一个序列。

上述操作是指对于一支股票而言的，为了代表市场整体，我们选择了 5 支市场流动性较好、分处不同行业的蓝筹股作为样本（如表 1 所示），并分别选择不同的市场环境进行预测（熊市：2011 年 12 月 9 日-2011 年 12 月 15 日；牛市：2009 年 6 月 25 日至 2009 年 7 月 1 日；震荡市：2010 年 1 月 18 日至 2010 年 1 月 22 日）。通过回测历史数据，我们得到的每支股票 \bar{r} 与 N 的关系，如图 1 所示。

表 1：传统算法实证股票样本（截止至 2012.6.11.）

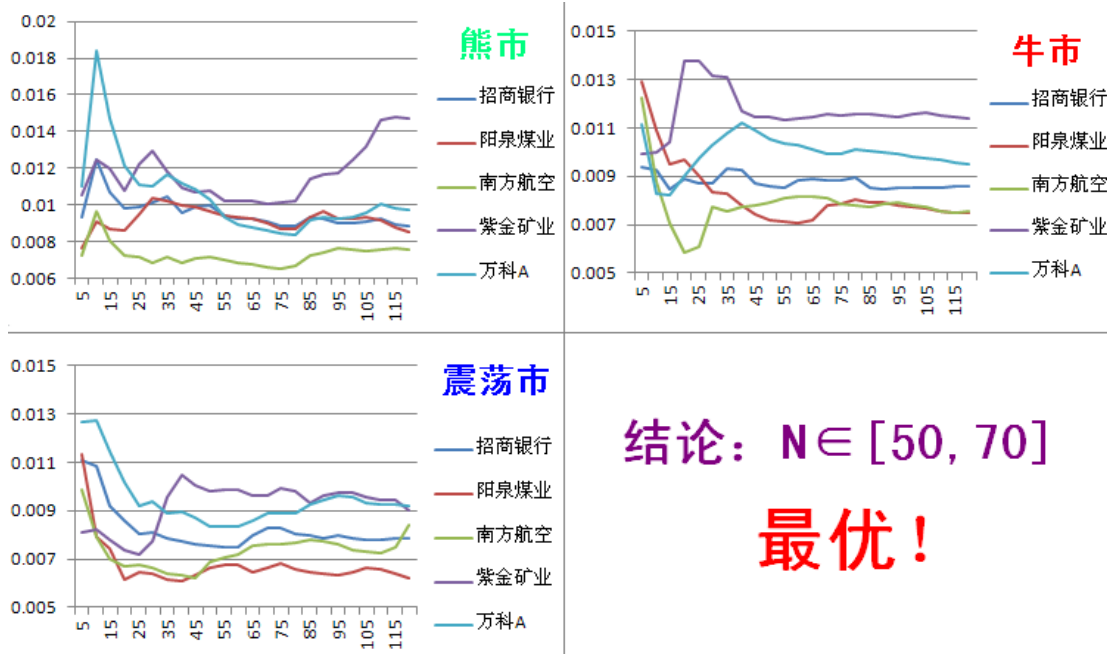
| 股票代码 | 股票名称 | 流通市值（亿元） | 近 3 年平均流通换手率 |
|--------|------|----------|--------------|
| 600036 | 招商银行 | 1363.94 | 0.44 |
| 600348 | 阳泉煤业 | 176.75 | 2.27 |
| 600029 | 南方航空 | 134.67 | 1.41 |
| 601899 | 紫金矿业 | 289.99 | 1.09 |
| 000002 | 万科 A | 739.93 | 0.75 |

数据来源：wind 资讯

从图 1 的回测结果来看，无论是在牛市、熊市，还是震荡的市场环境下， N 的范围设定在 50 至 70 之间，即取过去 50 至 70 个交易日成交量（成交量比例）的平均值，

总体上可以使得预测成交量（成交量比例）与市场真实成交量（成交量比例）的偏差最小。

图 1：各股票预测偏差 \bar{r} 在不同市场环境中的表现



数据来源：广发证券发展研究中心

2、历史样本权重

在上述对历史样本天数的研究中，我们采用了加权平均的算法。但是在实际情况中，由于市场环境以及市场微观结构的变化，算数平均所带来的结果有可能不够精确。

从经验上来看，相隔越近的交易日，其市场形态应该越相似，这也是波动率聚集和趋势交易的技术分析思想之一。那么成交量分布是否也具有这样的性质呢？国外有部分学术研究表明，在欧美一些市场中使用传统算法交易策略进行成交量预测时，通过图 2 中随时间逐步递减的权重，即对越近的交易日样本赋予越大的权重，可以对成交量（成交量比例）的分布做出更为准确的判断。

假设按照之前的研究结果，取 $N=60$ ，即采用 60 个交易日的历史成交量数据进行加权平均预测。按照越近交易日的成交量在加权平均时获得越大的权重，我们对 60 个交易日的权重 w_i 作以下设置

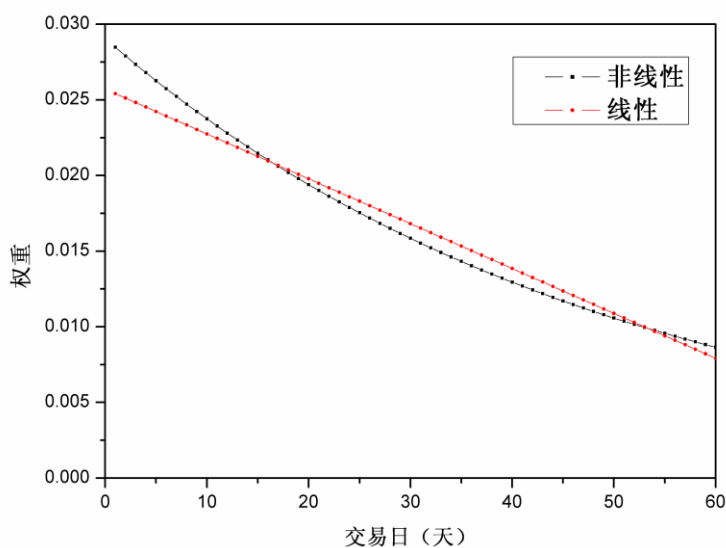
$$w_t = 0.97 w_{t+1} \quad (5)$$

并且令其满足约束条件

$$\sum_{i=1}^{60} w_i = 1 \quad (6)$$

即 t 日的权重是 $t+1$ 日权重的 97%。这样，权重随交易日天数的变化曲线如图 2 中黑线所示。在国外一些文献中还有如图 2 中红线所示的线性权重缩减，两者思想类似，我们不做过多讨论，这里只研究黑线所代表的非线性递减方式。

图 2:、VWAP 历史样本权重曲线



数据来源：广发证券发展研究中心

此时，我们对表1中5支具有代表性个股的 \bar{r} 再次进行回测。所得结果如表2所示。

表 2: 算术平均与加权平均预测 \bar{r} 结果比较

| 熊市 | 算术平均 \bar{r} | 加权平均 \bar{r} | 算术平均 \bar{r} -加权平均 \bar{r} |
|------|----------------|----------------|--------------------------------|
| 招商银行 | 0.00926 | 0.009012 | 0.000248 |
| 阳泉煤业 | 0.009316 | 0.009165 | 0.000151 |
| 南方航空 | 0.006887 | 0.006811 | 7.58E-05 |
| 紫金矿业 | 0.010203 | 0.0098 | 0.000402 |
| 万科 A | 0.008947 | 0.007998 | 0.000949 |
| 牛市 | 算术平均 | 加权平均 | 差值 |
| 招商银行 | 0.008818 | 0.008916 | -9.8E-05 |
| 阳泉煤业 | 0.007086 | 0.006663 | 0.000423 |
| 南方航空 | 0.008133 | 0.00859 | -0.00046 |
| 紫金矿业 | 0.011413 | 0.011134 | 0.000278 |
| 万科 A | 0.010283 | 0.010358 | -7.5E-05 |
| 震荡市 | 算术平均 | 加权平均 | 差值 |
| 招商银行 | 0.007484 | 0.008682 | -0.0012 |
| 阳泉煤业 | 0.006784 | 0.007841 | -0.00106 |
| 南方航空 | 0.007203 | 0.008804 | -0.0016 |
| 紫金矿业 | 0.00987 | 0.010588 | -0.00072 |
| 万科 A | 0.008354 | 0.007549 | 0.000806 |

数据来源：广发证券发展研究中心

可以看出，加权平均预测成交量比例分布所得到的 \bar{r} 除了在熊市中普遍缩减（即加权平均所预测的成交量分布与真实市场成交量分布更为接近）以外，在其他两种市场环境并不比简单算术平均所得到的效果更好。并且在很多情况下，投资者很难判断出自己是处于什么样的市场环境中。因此，我们认为将历史样本的权重设置为等权重即可，即在预测成交量（成交量比例）分布时建议采用简单算术平均方法。

（二）订单拆分参数

在对成交量（成交量比例）进行预测之后，就会进入实际的拆分下单环节。在这个环节中，拆分的一些细节可能会直接影响到算法交易的实际效果。对于几类传统算法交易策略而言，我们认为有两个较为重要的参数需要确定。一个是在 VWAP (MVWAP) 策略中的分段周期；另一个是下单频率，对于 TWAP 来说就是整个区间的划分频率（次数），而对于 VWAP (MVWAP) 来说则是在一个分段周期内进行平均划分的频率（次数）。

1、VWAP (MVWAP) 分段周期

对于 VWAP (MVWAP) 策略来说，需要按照分段的预测成交量在市场上进行下单。如果这个分段周期取得太长，对 VWAP 的跟踪误差会加大；如果分段周期取得太细，由于市场存在噪声，成交量的预测偏差会加大，从而同样会造成价格跟踪偏差增大。可以看出，在订单划分时需要取定合适的分段周期使得价格的跟踪偏差最小。

对此我们进行了实证研究。对于算法交易执行的绩效而言，主要关注两个指标。

第一个是成交价格对于 $VWAP_{\text{市场}}$ 的跟踪偏差 MPE

$$MPE = \frac{VWAP_{\text{成交}} - VWAP_{\text{市场}}}{VWAP_{\text{市场}}} \quad (7)$$

该指标越小越好；第二个是成交执行率，即实际成交量占预计成交量的比例，该指标越大（越接近 100%）越好。对于成交执行率而言，由于市场微观结构总是随实际成交发生变化的，并且结合一些算法交易用户的反馈，我们认为在限价委托的方式下仅仅使用历史数据很难对成交执行率做出较为准确的估算。同时，按照（3）式，我们认为成交量预测的准确程度可以在较大程度上反映出实际成交执行率的情况。因此，我们建立成交量预测准确度指标 VFA (Volume Forecasting Accuracy)，其定义为

$$VFA = \sum_i VP_i / \sum_i VPP_i = \sum_i VP_i \quad (8)$$

其中 VPP_i 是计划要完成的成交量比例，其总和为 1；而成交比例 VP_i 定义为

$$VP_i = \begin{cases} VPP_i & VPP_i < VPR_i \\ VPR_i & VPP_i \geq VPR_i \end{cases} \quad (9)$$

其中 VPR_i 是市场实际的成交比例。从（8）式来看， VFA 越大，说明成交量（成交量比例）的预测越准确，从而成交执行率越高。

按照之前表 1 中所提供的股票样本，同时按照 $N=60$ 以及不同的市场环境（如前，熊市：2011 年 12 月 9 日-2011 年 12 月 15 日；牛市：2009 年 6 月 25 日至 2009 年 7 月 1 日；震荡市：2010 年 1 月 18 日至 2010 年 1 月 22 日），取分段周期从 2 分钟至 10 分钟不等进行测试，并在最

后计算5支成分股 MPE 和 VFA 的均值（相当于计算一个投资组合的 MPE 和 VFA ），我们得到表3所示的结果。

表 3：各种市场环境中不同分段周期投资组合的 MPE 与 VFA

| 熊市 | MPE | VFA | 牛市 | MPE | VFA | 震荡市 | MPE | VFA |
|--------|------------|-------|--------|------------|-------|--------|------------|-------|
| 2min | -0.0000898 | 63.8% | 2min | -0.0005015 | 71.9% | 2min | -0.0004947 | 73.8% |
| 3 min | -0.0000870 | 63.8% | 3 min | -0.0004961 | 71.9% | 3 min | -0.0004970 | 73.8% |
| 4 min | -0.0000878 | 63.8% | 4 min | -0.0004953 | 72.0% | 4 min | -0.0005018 | 73.8% |
| 5 min | -0.0000916 | 63.8% | 5 min | -0.0004976 | 72.0% | 5 min | -0.0004971 | 73.8% |
| 6 min | -0.0000908 | 63.8% | 6 min | -0.0004926 | 72.0% | 6 min | -0.0005031 | 73.8% |
| 8 min | -0.0000942 | 63.8% | 8 min | -0.0004936 | 72.0% | 8 min | -0.0005011 | 73.8% |
| 10 min | -0.0000934 | 63.8% | 10 min | -0.0004887 | 72.0% | 10 min | -0.0005037 | 73.8% |

数据来源：广发证券发展研究中心

可以看出，不同的分段周期对 VFA 的影响很小，也就是说如果仅仅采用限价委托方式，分段周期的调整对交易执行率的影响较小；同样， MPE 随分段周期虽然显示出了一定的波动，但是换算至人民币单位后波动仍不足一分钱，因此我们认为 VWAP（MVWAP）分段周期取在 2 分钟至 10 分钟这一范围都可以，只要接下来的下单频率取得合适，就不会对执行效果造成太大影响。

2、下单频率

对于 TWAP 而言，将一个交易日的的时间进行平均划分，并为每个划分周期的子单赋予相同的下单量，这个划分周期就已经反映了下单频率；而对于 VWAP（MVWAP）而言，按照预测成交量取定分段周期后，往往还需要在其内部再进行成交量平均划分（相当于在 VWAP 策略中嵌入 TWAP 策略），从而可以达到隐藏交易行为、增加成交执行率、减小冲击成本的目的，而这样二次划分所得到的周期所对应的才是真实的下单频率。

在欧美等金融发达地区，由于市场由做市商制度保证流动性，很多算法交易都采用市价委托（Market Price Order）方式，并将下单频率设置得很高（约几个 seconds 量级），每次只提交很少量订单。

虽然理论上来说，母单拆分得越细，交易的连续性越可以得到保证，但是在国内市场，由于股票交易还是传统的市场化制度，交易频率太高可能会要求某些流动性不好的股票母单数量巨大。举个例子，某支流动性相对较弱的股票拥有 $c = 5$ 个亿的股本，平均每个交易日的换手率 $r = 0.2\%$ ，也就是说平均每个交易日的成交量为 $cr = 1$ 万手。如果按照该成交量均值买入或卖出一定比例的股票，例如 $p = 10\%$ ，则该交易日的母单数为 $crp = 1000$ 手。这意味着要将 1000 手的订单均摊（TWAP）或按一定比例分摊（VWAP 或 MVWAP）到 $t = 4$ 个小时的交易时间中。这样一来，如果是 TWAP，至少要在约每 14 秒提交一手订单；如果是 VWAP（MVWAP），提交订单数量需要按照预测成交量形成一个分布，也就是说每次的下单量都有可能大于 1 手，这样的话如果要在 4 个小时消化掉 1000 手母单，就需要更低的下单频率才能实现预测成交分布的有效追踪。

具体定量分析，假设 VWAP（MVWAP）的订单提交量序列（根据预测成交量得到）为 $\{V_i\}$ ，对其进行从小到大排序，得到新的有序序列 $\{V_i'\}$ ，假设 V_i' 变化得足够连续，

则最小子单提交量 $V_{\min}' = V_1'$ 与次子单提交量 V_k' 会有一个差值，这个差值最小为 1 手。如果希望对成交量分布的拟合控制在一定误差范围内，并且保证 V_1' 和 V_k' 都是正整数（单位为手），则需要设定可以容忍的误差范围 e ，使得 V_1' 必须满足 $V_1' \cdot e \cong 1$ 。例如当可以容忍 $e = 10\%$ 的成交量分布跟踪误差时， $V_{\min}' = V_1' = 10$ 手，则当 V_k' 与 V_1' 具有一手的差值时，可以满足 $V_k' = (1 + e)V_1' = 11$ 手。

从经验上来看，经过对过去大量成交量样本求均值后的预测成交量在一个交易日内往往呈现出微笑状的 U 型分布，即在上午开盘和下午收盘时的成交量相对较大，而中间时段的成交量相对较小。我们这里进行粗略估算：假设某个交易日子单提交量的均值是最小的子单提交量 V_1' 的 k 倍（例如， $k = 1.5$ ），则根据上面的分析，母单提交量为 crp ，平均子单提交量为 $k \times \frac{1}{e}$ ，则拆单数应该是两者之商，即

$$M = \frac{crpe}{k} \quad (10)$$

如果将这 M 份子单平均分摊到 t 秒中（例如，如果分摊到整个交易日，一般情况下 $t = 4 \times 60 \times 60$ 秒 = 14400 秒），则下单的最短周期为

$$T = \frac{t}{M} = \frac{kt}{crpe} \quad (11)$$

根据最短周期 T 的计算公式（11）可以看出，在 VWAP（MVWAP 策略中），下单频率和股票本身的股本、流动性、历史成交量分布特征有关，也与交易者对下单比例和成交量分布拟合误差的要求有密切联系。另一方面，由于国内市场算法交易目前多采用限价委托方式进行订单提交，因此如果下单频率过高，也会引起撤单频率高、交易执行效率低的问题。因此，我们认为在实施算法交易时，可以按照（11）式，根据个股的具体特征对下单频率进行设定和调整。

（三）下单量参数

在对成交量分布进行预测、并对下单周期进行划分之后，面对的就是订单的提交与执行。此时提交的下单量是一个重要参数——提交的下单量太大，会对市场造成较大的冲击，使得交易成本增加；而提交的下单量太小，会使得交易节奏降低，影响交易执行的效率。因此下单量对于整个算法交易流程是一个值得着重研究的参数。

具体来讲，传统算法交易策略中，下单量参数有两类：一类是对于 TWAP 或 VWAP 而言，在根据历史成交量进行预测后，总会根据预测成交量的一定比例进行下单；另一类是在 MVWAP 中，当市场的实时价格低于（高于）实时 $VWAP_{\text{市场}}$ 时，会对拟下单量进行一定比例的调整。接下来我们就分别对这两类下单量参数进行研究。

1、相对于预测成交量的实际下单比例

根据之前所述，通过历史上一定长度（如 $N=60$ 个交易日）的样本，可以获得拟交

易日的成交量分布。但是在进行真实订单提交的时候，并不能直接按照这个预测的成交量进行下单。通常的做法是在预测成交量的基础之上乘以一个比例系数，得到下单量。这个下单比例系数要取得合适——既不能太大，造成较大的市场冲击；又不能太小，影响交易执行效率。

由于通过历史数据进行回测仅具有一定的参考意义，特别是对于交易策略而言，交易者在市场中每进行一笔交易，都会改变市场的微观结构，使得市场在之后的演化与不真实进行该笔交易的情况完全不同。因此，我们这里采用高频盘口数据对下单比例系数进行估计，也仅仅是一种粗略的计算模型。

以 VWAP 买入股票为例，考虑到国内市场一般以限价委托进行订单提交，假设以最近成交价进行下单，则会出现两种可能：一种是被动挂单，一种是主动买盘，从长期、统计的角度来看，两者情况出现的概率各为 0.5。如果是主动买盘，为了避免对市场造成较大冲击，最好的办法是所提交的订单数不要超过当时卖一盘口上的数量。

按照 1 分钟下单频率（5 分钟 VWAP 分段周期），我们计算了沪深 300 成分股 30 个交易日（截止至 2012 年 6 月 11 日，剔除停牌部分）卖一盘口数量（1 分钟频率）占预测成交量（ $N=60$ ）的比例均值 \bar{R} 与标准差 $\sigma_{\bar{R}}$ ，计算中当卖一盘口数量大于占预测成交量时，为不造成市场冲击，我们仅按卖一价下单，则该比例记为 1。经过对 30 个交易日 300 支样本股票的统计，我们得到结果

$$\begin{cases} \bar{R} = 0.419 \\ \sigma_{\bar{R}} = 0.228 \end{cases} \quad (12)$$

也就是说，从统计角度上来看，沪深 300 成分股盘口实时的卖一委托单约占 1 分钟的预测成交量的 40%。换句话说，如果按照 1 分钟的下单频率，对于子单的提交量，建议不超过该分钟预测成交量的 40%，这样才能够实现对 VWAP 的有效跟踪。

当然，这个数值只是一个投资组合的统计的结果，对于不同股票，这个下单比例应该有所不同，建议单独进行测算。我们这里仅仅是从市场微观结构的角度出发提供一种思路。另一方面，按照不超过预测成交量 40% 的比例下单，只是为了盯紧市场均价并减小冲击成本，并不能避免由于买盘量增加而带来的股价提升。在实际交易中，这种由于自身参与而改变市场微观结构所带来的隐形成本也是投资者应该关注的一个问题。

2、MVWAP 下单量增减比例

采用 MVWAP 策略的目的之一是为了有效战胜市场，获得更优的成交均价（对于买入交易实现成交均价低于 $VWAP_{\text{市场}}$ ，对于卖出交易实现成交均价高于 $VWAP_{\text{市场}}$ ）。但是该策略也同样会带来两个问题：第一问题是，对于增加下单量来说，可能带来一定的冲击成本，削弱算法交易本来的作用和意义；第二个问题是，如果股票在某个交易日出现价格持续高于或低于实时 $VWAP_{\text{市场}}$ 的情况，按照 MVWAP 策略，有可能会出现订单无法以较高交易执行率完成的情况，也有可能出现订单提前完成、跟踪误差放大的情况。

对于第一个问题，可以通过调整 MVWAP 的下单量增减比例来控制冲击成本。之前已经估算过，对于沪深 300 指数来说，应按照不超过预测成交量 40% 的订单量进行下单。如果相对于预测成交量的实际下单比例为 ϕ ，在这个比例的基础上再增加一定的

下单量，但要保证总比例不超过 40%，即 $\phi(1+x) \leq 0.4$ 。此时 MVWAP 下单量增加比例的上限为

$$x = \frac{0.4}{\phi} - 1, \quad \phi \in (0, 0.4] \quad (13)$$

在此范围内，冲击成本可以得到有效控制。如前所述，这里的参数 0.4 是通过沪深 300 指数成分股统计出的一个整体参数，具体到个股，建议按照之前盘口数据的方法对参数进行测算。

对于第二个问题，已经不仅仅是参数研究所讨论的范畴，更是一种对于风险和相对收益进行权衡的策略问题。这里我们仅提出几点定性建议：

1、对于交易执行率没有严格要求的订单（例如一张母单要在若干个交易日完成），推荐使用 MVWAP 算法。

2、趋势行情使用 MVWAP 的效果会打折扣，MVWAP 在震荡市中表现相对会更好。

3、一般情况下不建议开盘使用 MVWAP 策略；在尾盘时适当使用 MVWAP 会提升算法交易的绩效。例如在买入交易中，当尾盘（如 14:30 后）市场价格小于 $VWAP_{\text{市场}}$ 时，或在卖出交易中，当尾盘市场价格大于 $VWAP_{\text{市场}}$ 时，可以将 VWAP 转换为 MVWAP，适当增加下单数量。

4、如果在使用 MVWAP 的过程中发现价格持续高于或低于实时 $VWAP_{\text{市场}}$ ，并对算法交易造成不利影响时，应该及时对算法交易策略做出调整。

四、总结

本篇报告首先介绍了几类传统算法交易策略，包括 TWAP、VWAP 以及 MVWAP。在此基础上，针对这几类策略在实际操作细节中可能遇到的一些参数，我们进行了实证研究。这些参数可以分为三类，分别是预测成交量（成交量比例）参数、订单拆分参数和下单量参数，其中每类各包括两个具体参数。

预测成交量（成交量比例）参数包括历史样本天数和历史样本权重。历史样本天数取得太小，样本短期内的波动可能引起较大的成交量（成交量比例）预测偏差；如果取得太大，由于在较长的时间中市场环境已经发生了较大的变化，远离预测日的交易量情况可能对该预测日不具有重要意义，并且还可能给预测带来偏差。通过对不同的市场环境进行回测，我们发现针对国内 A 股市场，历史样本天数取 50 至 70 个交易日较为合适。而历史样本权重则与欧美一些投资研究中所指出的权重随时间间隔增大而递减的情况不同，在预测成交量（成交量比例）分布时，我们建议采用简单算术平均方法即可。

订单拆分参数包括 VWAP（MVWAP）分段周期和下单频率。通过回测跟踪偏差和成交执行率两个指标，我们发现取 2 分钟至 10 分钟的 VWAP（MVWAP）分段周期对策略结果的影响不大，因此在实际操作过程中可以在此范围内任意选择；下单频率则是对于 TWAP 和 VWAP（MVWAP）通用的参数，通过研究，我们推导出了下单频率的一个表达式，其与股票的交易时间、流通股本、流动性、历史成交量分布特征、交易者对下单比例

和成交量分布拟合误差的要求等有关。

下单量参数包括相对于预测成交量的实际下单比例和 MVWAP 下单量增减比例。其中通过盘口数据测算，我们认为相对于预测成交量的实际下单比例不应该超过 40%，并且更好的办法是针对不同个股测算该参数的具体数值。根据这个参数的结果，可以进一步确定 MVWAP 下单量增减比例这个参数，我们给出了增加比例参数上限的表达式，并针对股票在某个交易日出现价格持续高于或低于市场实时 VWAP 的情况给出了一些定性建议。

广发金融工程研究小组

罗军，首席分析师，华南理工大学理学硕士，2010 年进入广发证券发展研究中心。

俞文冰，首席分析师，CFA，上海财经大学统计学硕士，2012 年进入广发证券发展研究中心。

叶涛，资深分析师，CFA，上海交通大学管理科学与工程硕士，2012 年进入广发证券发展研究中心。

安宁宁，资深分析师，暨南大学数量经济学硕士，2011 年进入广发证券发展研究中心。

胡海涛，分析师，华南理工大学理学硕士，2010 年进入广发证券发展研究中心。

夏满阳，分析师，上海交通大学金融工程硕士，2012 年进入广发证券发展研究中心。

汪鑫，分析师，中国科学技术大学金融工程硕士，2012 年进入广发证券发展研究中心。

李明，分析师，伦敦城市大学卡斯商学院计量金融硕士，2010 年进入广发证券发展研究中心。

蓝昭钦，分析师，中山大学理学硕士，2010 年进入广发证券发展研究中心。

史庆盛，研究助理，华南理工大学金融工程硕士，2011 年进入广发证券发展研究中心。

张超，研究助理，中山大学理学硕士，2012 年进入广发证券发展研究中心。

相关研究报告

积小流以成江海——关于算法交易的一个综述：算法交易系列研究之一 罗军 2012-06-11

| | 广州市 | 深圳市 | 北京市 | 上海市 |
|------|----------------------------|----------------------------|-----------------------------|--------------------------------|
| 地址 | 广州市天河北路 183 号 大都会广场 5 楼 | 深圳市福田区民田路 178 号华融大厦 9 楼 | 北京市西城区月坛北街 2 号 月坛大厦 18 层 | 上海市浦东南路 528 号 上海证券大厦北塔 17 楼 |
| 邮政编码 | 510075 | 518026 | 100045 | 200120 |
| 客服邮箱 | gfyf@gf.com.cn | | | |
| 服务热线 | 020-87555888-8612 | | | |

免责声明

广发证券股份有限公司具备证券投资咨询业务资格。本报告只发送给广发证券重点客户，不对外公开发布。

本报告所载资料的来源及观点的出处皆被广发证券股份有限公司认为可靠，但广发证券不对其准确性或完整性做出任何保证。报告内容仅供参考，报告中的信息或所表达观点不构成所涉证券买卖的出价或询价。广发证券不对因使用本报告的内容而引致的损失承担任何责任，除非法律法规有明确规定。客户不应以本报告取代其独立判断或仅根据本报告做出决策。

广发证券可发出其它与本报告所载信息不一致及有不同结论的报告。本报告反映研究人员的不同观点、见解及分析方法，并不代表广发证券或其附属机构的立场。报告所载资料、意见及推测仅反映研究人员于发出本报告当日的判断，可随时更改且不予通告。

本报告旨在发送给广发证券的特定客户及其它专业人士。未经广发证券事先书面许可，任何机构或个人不得以任何形式翻版、复制、刊登、转载和引用，否则由此造成的一切不良后果及法律责任由私自翻版、复制、刊登、转载和引用者承担。