

算法交易研究系列（六）

2011 年 10 月 27 日

主成分分解方法在 VWAP 策略中的应用

在研究证券市场的特点时，对价格或者收益率的重视程度通常要远远高于对成交量的分析。但每一个市场的实际参与者在交易过程中都会面临由流动性变化所带来的冲击成本。这一成本可以通过比较一系列交易指令得到完全执行后的成交价与某一基准的差异来进行度量。一个常用的基准就是成交量加权平均价格(Volume Weighted Average Price, VWAP)。它是一段时期内，所有成交价格经由成交量加权之后的和。基于此，衍生出一系列算法交易策略，以跟踪或近似真实的 VWAP。

- 交易的 VWAP 策略被定义为投资者拆分大额委托单，在约定时间段内分批执行，以期使得最终买入或卖出的成交均价尽量接近该段时间内整个市场成交均价的算法交易策略。作为一个比较标准，VWAP 不仅简单有效，而且有着天然的优良性。与任意未来某个固定时刻的价格相比，VWAP 并不能被人为操控。对一个以跟踪 VWAP 下单的投资者来说，他所要做的就是找到在交易前找到一个策略使得事后的平均成交价格尽可能接近 VWAP，从而达到降低成本、增加收益的目的。
- 通过主成分分解，成交量被分成两个部分——市场成分和特殊成分。市场成分描述了市场上所有股票成交量的共同变化形态，并且抓住了该变化具有周期性这一重要特点。特殊成分则度量了短时间内单个股票成交量反常或意外的波动，常常是因为一些重大事件的信息提早泄露所导致的。市场成分可以由历史同期的均值较好地刻画，而特殊成分则展现出 ARMA 模型的典型特征。
- VWAP 策略理论执行的结果充分说明了主成分分解方法的巨大优势，其预测的每个样本股的 VWAP 策略的执行成本都低于传统方法。不仅如此，主成分分解方法在较短的时间间隔下能获得更小的误差，这一特点对实际交易有着极其重要的意义。而在相同的采样区间下，AR 模型的滞后阶数越大，MAPE 的改善也越多。当采样区间为 2 分钟，且用 AR(3)来拟合特殊成分时，相应于传统方法，主成分分解方法所得的平均 MAPE 降低了 8.921bp，为所有模型的最优结果。此外，该模型所得 MAPE 的标准差为 1.108bp，95%的分位数为 6.874bp，均对传统方法有着很大的改进。
- 样本外检验表明，和传统方法相比，主成分分解方法的 VWAP 动态执行策略在跟踪误差上有着明显的改善。我们用 2011 年 2 月 1 日到 2011 年 8 月 31 日期间内上证 50 的所有股票进行了实证分析，49 个股票的 PC-AR(3)模型的 VWAP 策略执行成本低于传统方法，其中 16 个股票的成本比传统方法减少了 3bp 以上，占到所有样本股的 36%，有效降低了 VWAP 策略的执行成本。
- 传统方法的预测误差越大，主成分分解方法的动态策略获得的改进也越大。在传统方法较为有效的区间 ($MAPE < 10bp$)，用主成分分解加入成交量日内的动态变化并不能显著降低 VWAP 策略的执行成本。相反，一旦传统方法的误差变大时，改善就变得相当明显。这一结论保证了本文方法在任何情况下都是有效的。

特别声明：本报告所涉及模型和策略的原理、假设和计算方法已在专题研究报告中披露，模型使用的数据源均来自于市场公开信息。本报告中的投资建议是数量化模型自动计算的结果，研究员未进行主观判断调整。

相关研究

股票市场均价下单策略 (VWAP-D)	20111020
关于股票配对交易的补充说明	20110729
统计套利之股票配对交易	20110610
套利交易策略综述	20110111
算法交易在国内的应用	20101228

金融工程高级分析师

吴先兴

SAC 执业证书编号：S0850511010032

电话：021-23219449

Email: wuxx@htsec.com

联系人

朱剑涛

电话：021-23219745

Email: zhujt@htsec.com

联系人

冯佳睿

电话：021-23219732

Email: fengjr@htsec.com

目 录

1. VWAP和VWAP策略.....	2
2. 日内成交量的分解与建模.....	2
2.1 成交量预测和VWAP跟踪.....	2
2.2 成交量的日内分解.....	3
2.3 日内成交量建模.....	4
3. VWPA策略的应用.....	6
4. 结论.....	11

在研究证券市场的特点时，对价格或者收益率的重视程度通常要远远高于对成交量的分析。但每一个市场的实际参与者在交易过程中都会面临由流动性变化所带来的冲击成本。这一成本可以通过比较一系列交易指令得到完全执行后的成交价与某一基准的差异来进行度量。一个常用的基准就是成交量加权平均价格(Volume Weighted Average Price, VWAP)。它是一段时期内，所有成交价格经由成交量加权之后的和。基于此，衍生出一系列算法交易策略，以跟踪或近似真实的 VWAP。

1. VWAP和VWAP策略

算法交易策略中常用到的比较基准有 VWAP，TWAP，Market On Open(MOO)，Market On Close(MOC)，Price In Line，但其中最流行的还是成交量加权平均价格(VWAP)。主要原因有以下两个。首先，日 VWAP 的计算非常简便，任何人都可以通过股票每天的交易记录来获取。更重要的是，采用 VWAP 作为比较基准，方便直观且不能被人为操控，有着天然的优良性。简单说来，如果一笔交易的买入价低于 VWAP，那么这就是一笔好交易。反之，就可以认为交易是不成功的。对于卖出交易，只需将这一比较标准相反进行即可。对于一个以跟踪 VWAP 下单的投资者来说，他所要做的就是在交易前找到一个策略使得事后的平均成交价格尽可能接近 VWAP，从而达到降低成本、增加收益的目的。

交易的 VWAP 策略被定义为投资者拆分大额委托单，在约定时间段内分批执行，以期使得最终买入或卖出的成交均价尽量接近该段时间内整个市场成交均价的算法交易策略。这一策略在算法交易中得到迅速的发展，完全得益于其操作上的简易性。对于选择 VWAP 策略的投资者在进行大量证券的买卖时，为了减少市场冲击，通常会选择拆分委托单并分批交易的方式进行，通过承担一定的时间风险提高自身的收益。不仅如此，这种交易法则也提高了整个证券市场的透明度和有效性。

为了实现 VWAP 策略，就必须对日内成交量的变化方式进行建模。众所周知，几乎所有股票的日内成交量都呈现出 U 型的周期性特征。而这种周期性波动恰恰成了对日内成交量精确建模的最大桎梏。我们希望找到一个办法把这一周期性给分离出来，作为整个市场的共有特性，剩余部分则作为每个股票独有的特点进行研究。

2. 日内成交量的分解与建模

如前文所述，日内成交量建模所面临的最大问题就是其高度的周期性。通常有两种方法来解决这一问题。一是对任意一个股票，取其历史成交量的均值作为周期性的估计。二是在一固定时刻以所有股票的平均成交量来描述这种周期性特征。这两种方法都只研究了成交量变化特点的一个侧面，有着一定的局限性。因此，在本文中，我们将采用一个基于主成分分析的新方法弥补上述两种技术的缺陷。

通过主成分分解，成交量被分成两个部分——市场成分和特殊成分。市场成分描述了市场上所有股票成交量的共同变化形态，并且抓住了该变化具有周期性这一重要特点。特殊成分则度量了短时间内成交量反常或意外的波动，常常是因为一些重大事件的信息提早泄露所导致的。市场成分可以由历史同期的均值较好地刻画，而特殊成分则展现出 ARMA 模型的典型特征。

在具体介绍这一方法之前，我们先讨论一些与 VWAP 跟踪策略相关的概念。

2.1 成交量预测和VWAP跟踪

因为 VWAP 策略是事前制定的，所以在日内的任意一个时刻，只要知道当日剩余时间内成交量的分布情况，我们就能及时调整交易策略。这样，一个好的 VWAP 策略关键

就在于如何准确地预测日内成交量。为了解释 VWAP 策略的执行以及好的成交量预测的重要性，我们通过下面这个简单的例子加以说明。

假定在一个简化了的金融市场上，证券交易只能在时刻 $i (i=1, \dots, 16)$ 发生，一个投资者想要以 VWAP 交易 100,000 股某股票。如果她事先知道当天成交量 (V_i) 或换手率，

$x_i = \frac{V_i}{N}$ ，其中 N 是该股票的流通市值，在每个交易时点上的分布形态，她就能计算出

各时点上需要成交的比例 $\tilde{x}_i = \frac{x_i}{\sum x_i} = \frac{V_i}{\sum V_i}$ ，然后按照这个比例以价格 P_i

交易她的 100,000 股。其中，每个时点上交易的量 $\underline{V}_i = 100000 \tilde{x}_i$ 。如此，这笔交易的 VWAP 即为，

$$VWAP = \sum_{i=1}^{16} P_i \left(\frac{\underline{V}_i}{\sum_{i=1}^{16} \underline{V}_i} \right) = \sum_{i=1}^{16} P_i \left(\frac{100000 \tilde{x}_i}{100000 \sum_{i=1}^{16} \tilde{x}_i} \right) = \sum_{i=1}^{16} P_i \tilde{x}_i。$$

由上式可以看出，想要跟踪某个股票或者投资组合每天的 VWAP，只需知道当日的成交量形态。但是，事实上在每天收盘前我们都不知道当日的成交量会如何分布，所以就需要在每日开盘前预测换手率序列 x_i 。当然了，预测越精确，最终的成交价格也会越接近理论 VWAP。这个简单的例子强调了寻找一个好的成交量模型的重要性，这也是本文余下部分讨论的重点。

2.2 成交量的日内分解

本文的思想来源于资产管理的理论和实践，由 CAPM 模型可知，任何一个投资组合都能被分解成市场组合和套利组合。同样的思路 and 过程也可以被应用到日内成交量，它有一个市场成分和特殊成分。任意一个股票在某一天的成交量或换手率都可以被看作是平均项与误差项的和。为了得到这两部分，我们需要依赖于主成分分析。

假设 $x_{it}^j = \frac{V_{it}^j}{N_i^j}$ ， $i=1, \dots, I$ ， $t=1, \dots, T$ 为第 j 天第 i 个股票 t 时刻的换手率，为记号

简便，在后文的叙述中，省略上标 j 。根据这个定义，市场的换手率可以被定义为：

$$x_t^I = \frac{\sum_i P_{it} V_{it}}{\sum_k P_{kt} N_k} = \frac{\sum_i P_{it} N_i \frac{V_{it}}{N_i}}{\sum_k P_{kt} N_k} = \sum_i w_{it} x_{it}，$$

其中 P_{it} 是第 i 个股票 t 时刻的成交价格。将序列 x_{it} 中心化之后做主成分分解，可得

$$x_{it} = c_{i,t} + y_{i,t}。$$

由此可见，第 i 个股票 t 时刻的换手率 x_{it} 是市场成分 $c_{i,t}$ 与特殊成分 $y_{i,t}$ 的和。一方面，市场成分描述了股票成交量的长期变化，并且抓住了该变化具有周期性这一重要特

点。另一方面，特殊成分则度量了短时间内成交量反常或意外的波动，常常是因为一些重大事件的信息提早泄露所导致的。

2.3 日内成交量建模

上述分解只是理论上的推导，在实际应用中，我们使用如下的模型作为日内第 i 个股票 t 时刻换手率的估计。

$$\hat{x}_{it} = \bar{c}_{i,t} + y_{i,t}$$

其中 $\bar{c}_{i,t}$ 是过去 L 个交易日同一时刻历史换手率的平均。如前文所述， $c_{i,t}^j$ 代表了第 j 天的换手率，应该被写作 $c_{i,t}^j$ ，所以 $\bar{c}_{i,t}$ 为

$$\bar{c}_{i,t} = \frac{1}{L} \sum_{l=1}^L c_{i,t}^{j-l}。$$

采用这一形式是因为在较短的时期内(通常不长于 3 个月)，市场成分应当是较为稳定的。第二项 $y_{i,t}$ 则代表了每个股票的特殊成分，可以用 ARMA 模型较好地拟合。

在进行实证分析前，我们需要对提出的模型和理论做一番验证，保证其适用性和有效性。为此，我们选取了上证 50 的所有股票作为样本股，时间段为 2011 年 2 月 1 日至 2011 年 8 月 31 日，共 143 个交易日。以 15 分钟为长度在交易时间内提取换手率，每天可以得到 16 个样本点。

上述成交量分解方法的第一步便是对所有样本股进行主成分分析，因此我们需要确定选用多少个交易日的数据来完成这一步骤。表 1 给出了选择不同长度的数据做主成分分析后前三个主成分所占的比例。

表 1 上证 50 样本股日内成交量的主成分分解

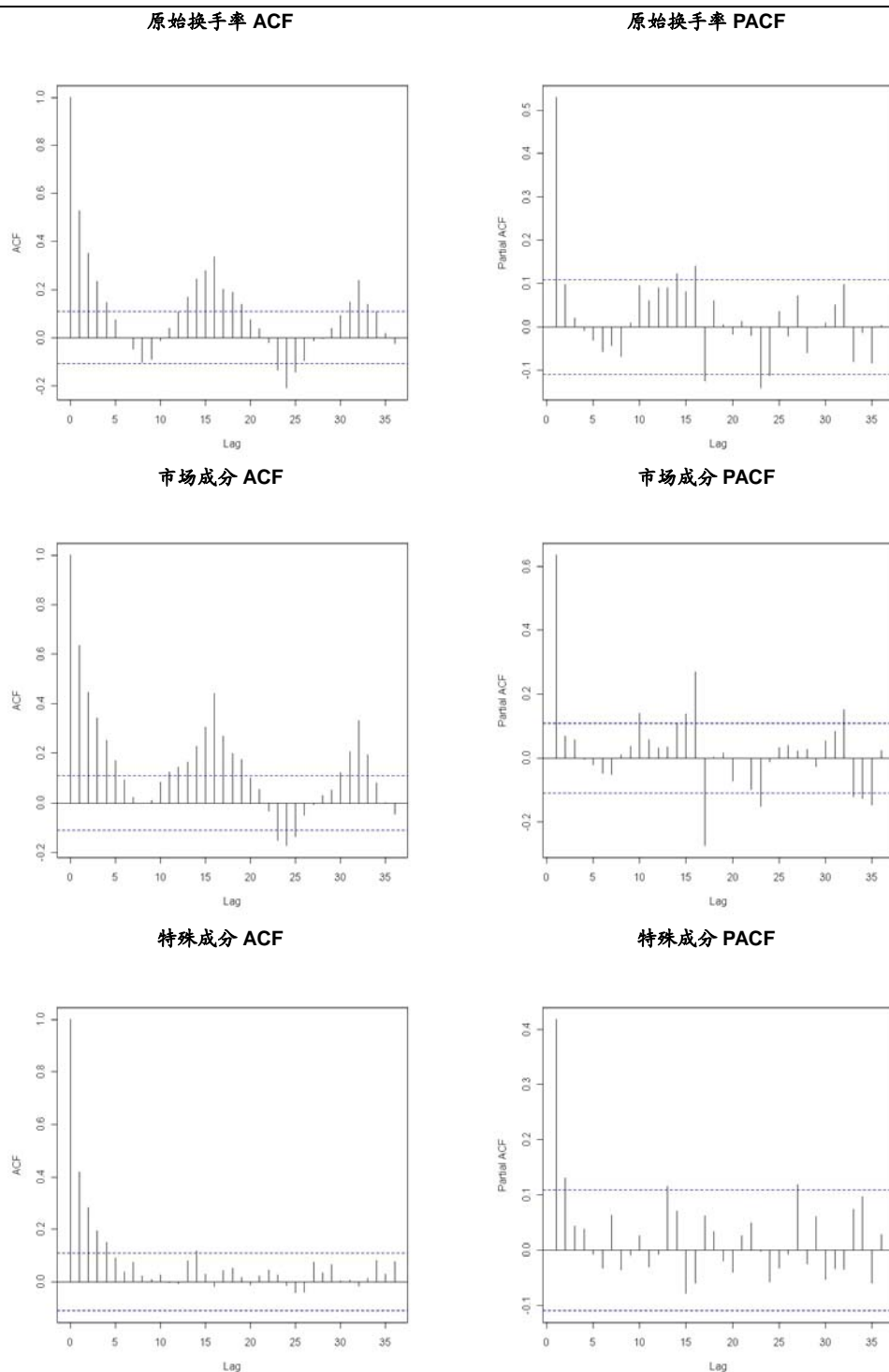
时间区间	特征值排序	特征值	特征值之差	百分比	累计百分比
20 个交易日 (2011 年 2 月 1 日-2011 年 3 月 7 日)	1	15.166	10.204	30.331	30.331
	2	4.962	1.920	9.924	40.255
	3	3.042	0.243	6.084	46.339
40 个交易日 (2011 年 2 月 1 日-2011 年 4 月 6 日)	1	14.174	9.613	28.349	28.349
	2	4.561	1.194	9.122	37.471
	3	3.367	1.018	6.734	44.205
60 个交易日 (2011 年 2 月 1 日-2011 年 5 月 5 日)	1	13.430	9.070	26.861	26.861
	2	4.361	1.447	8.721	35.582
	3	2.914	0.786	5.828	41.410

资料来源：海通证券研究所

从表中最后一列的累积百分比可以看出，时间区间越长，前三个主成分能够解释的偏离度也越低。当时间周期为一个月（20 个交易日）时，这三个主成分可以解释 46.40% 的偏离度。而当使用三个月（60 个交易日）的数据进行主成分分析时，这个比例就下降到了 41.41%。作为一种静态的分析技术，主成分分析并不能抓住长期特征，因此在下文的实证分析中，我们将基于一个月的数据来分解换手率。

下面，我们以特变电工（600089）为例按前述的模型分解换手率，考察市场成分是否能抓住换手率的主要特征——周期性，ARMA 模型又是否能较好地刻画特殊成分。图 1 从上至下分别是原始换手率、市场成分和特殊成分的自相关系数（ACF）和偏自相关系数（PACF）。

图 1 特变电工(600089)20 个交易日换手率分解后的自相关和偏自相关系数



资料来源：海通证券研究所

原始换手率的 ACF 反映出成交量的典型特征——周期性，而市场成分恰能很好地捕

请务必阅读正文之后的信息披露和法律声明

捉这一特点，其 ACF 基本呈现出和原始换手率相同的变化形态，这表明对原始换手率进行主成分分解是成功的。最后两幅图是特殊成分的 ACF 和 PACF，自相关系数的快速衰减和偏自相关系数仅在二阶显著表明 AR(2)模型能够很好地刻画该序列。

通过以上的论证与分析，我们展示了主成分分解既能有效地抓住原始换手率的特点，也能体现出每个股票的个性。以此为基础，我们将在下一部分集中讨论该方法在降低 VWAP 策略的执行成本时所体现的优势。

3. VWPA 策略的应用

预测成交量或换手率的传统方法是求过去 L 个交易日的历史平均。因此日内成交量 $x_{i,t}$ 可以被估计为

$$\hat{x}_{i,t} = \frac{1}{L} \sum_{l=1}^L x_{i,t}^{j-l}.$$

毫无疑问，传统方法的最大优势就在于其简单的形式，然而它却忽略了成交量日内的动态变化，而这恰恰对预测质量有着重大的影响。如果本文提出的方法能够把这一动态变化考虑在内，并能有效地降低 VWAP 策略执行时的成本，那对于实际交易将会有极大的价值。

在该报告系列的上一篇中，我们使用了两种 VWAP 策略的执行方法：VWAP 策略静态执行和 VWAP 策略动态执行。这里，我们再介绍一种新的方法——VWAP 策略理论执行。根据主成分分解方法构建的模型，只要获得 t 时段内的真实换手率 $x_{i,t}$ 就能很轻易地预测下一时间段的换手率 $\hat{x}_{i,t+1}$ 。在此基础上便可以确定当日每一个时间区间内需要被

交易的证券数量的占比为 $\frac{\hat{x}_{i,t}}{\sum_{t=1}^T \hat{x}_{i,t}}$ 。但是，理论 VWAP 执行策略在实际应用中是无法

实现的，因为这一执行方法不仅要求投资者能够根据最新的换手率数据实时更新下一期的预测，更需要知晓当日所有的预测值以求出交易数量的占比。显然，后者在收盘前是无法做到的。所以，这一执行方法只能是理论上的。

本文研究的重点是 VWAP 策略动态执行，它是依据最新得到的数据实时更新的一种技术，有着极大的实战价值。以 2011 年 2 月 1 日至 2011 年 8 月 31 日间上证 50 所有股票的成交数据为研究对象，分别以 2 分钟、5 分钟、10 分钟和 15 分钟为采样区间提取换手率和平均成交价，考察不同的时间间隔对 VWAP 策略执行成本的影响。由前文的论证可知，采用 20 天的数据进行主成分分解对偏离度的解释程度最高。所以，在实证分析时，我们以 20 天为一个周期进行模型的估计并向后预测一天的 VWAP，计算和真实 VWAP 的误差。然后向后滚动一天，同时舍去最远那天的数据重新估计模型进而预测后一天的 VWAP。这样，便可得到上证 50 所有样本股 2011 年 3 月 8 日至 8 月 31 日间每日的 VWAP 预测。在实际计算中，市场成分可由历史均值预测，而对特殊部分的建模，我们分别使用了 AR(1)、AR(2)和 AR(3)进行拟合，考察模型的效果是否对滞后阶数敏感。

我们首先从 VWAP 策略理论执行的结果开始分析和比较，平均绝对值百分比误差 (MAPE) 在日层面上被用来度量预测 VWAP 相对于真实值的偏离。假设我们预测了 K 个交易日的 VWAP，以 VWAP_PC_AR 记采用主成分分解方法预测的 VWAP，VWAP_TRUE 记真实的 VWAP，那么主成分分解方法得到的 MAPE 被定义为

$$\text{MAPE} = \frac{1}{K} \sum_{k=1}^K \frac{|\text{VWAP_PC_AR}_k - \text{VWAP_TRUE}_k|}{\text{VWAP_TRUE}_k}.$$

同理可得传统方法下每个股票的 MAPE。表 2 展示的是在不同采样区间和滞后阶数下由主成分分解方法和传统方法所得 50 个样本股 MAPE 的均值、标准误等描述性统计量。

表 2 上证 50 样本股日 VWAP 预测的 MAPE (策略的理论执行, 单位: bp)

预测方法	优于传统方法的比例 (%)	均值	标准误	最小值	最大值	95%分位数
采样区间=2 分钟						
PC-AR(1)	100	6.454	1.451	3.016	10.024	8.764
PC-AR(2)	100	5.555	1.229	2.702	8.536	7.598
PC-AR(3)	100	5.072	1.108	2.561	7.603	6.874
传统方法		13.976	4.290	6.992	23.503	22.289
采样区间=5 分钟						
PC-AR(1)	100	6.320	1.549	3.081	10.092	9.193
PC-AR(2)	100	5.764	1.434	2.917	9.703	8.521
PC-AR(3)	100	5.523	1.383	2.784	9.455	8.124
传统方法		13.591	4.282	6.856	22.840	21.838
采样区间=10 分钟						
PC-AR(1)	100	6.746	1.773	3.355	11.727	10.007
PC-AR(2)	100	6.485	1.776	3.131	11.827	9.680
PC-AR(3)	100	6.423	1.775	3.100	11.871	9.627
传统方法		13.112	4.112	6.616	21.875	21.035
采样区间=15 分钟						
PC-AR(1)	100	7.272	2.027	3.486	12.407	11.003
PC-AR(2)	100	7.175	2.041	3.462	12.594	10.885
PC-AR(3)	100	7.194	2.041	3.412	12.687	10.870
传统方法		12.699	3.978	6.499	20.952	20.209

资料来源: 海通证券研究所

上表的结果充分说明了主成分分解方法的巨大优势, 其预测的每个样本股的 VWAP 策略的执行成本都低于传统方法 (第 2 列)。不仅如此, 主成分分解方法在较短的时间间隔下能获得更小的误差, 这一特点对实际交易有着极其重要的意义。而在相同的采样区间下, AR 模型的滞后阶数越大, MAPE 的改善也越多。当采样区间为 2 分钟, 且用 AR(3) 来拟合特殊成分时, 相应于传统方法, 主成分分解方法所得的平均 MAPE 降低了 8.921bp, 为所有模型的最优结果。此外, 该模型所得 MAPE 的标准差为 1.108bp, 95% 的分位数为 6.874bp, 均对传统方法有着很大的改进。

然而, 上述巨大的改进只是 VWAP 策略理论执行的结果, 在实际交易中是无法达到的, 因此 VWAP 策略动态执行的效果才是我们真正关心的。从表 2 的结果来看, PC-AR(3) 在 VWAP 策略理论执行中表现最好, 因此我们在研究动态执行时都以该模型为基础, 并在不同的采样时间长度上和传统方法一一对比, 检验主成分分解方法的效果。

表 3 上证 50 样本股 PC-AR(3)模型日 VWAP 预测的 MAPE (策略的动态执行, 单位: bp)

预测方法	优于传统方法的比例 (%)	均值	标准误	最小值	最大值	95%分位数
采样区间=2 分钟						
PC-AR(3)	98	11.241	3.241	6.370	20.582	16.715
传统方法		13.976	4.290	6.992	23.503	22.289
采样区间=5 分钟						
PC-AR(3)	100	10.833	3.225	6.096	19.752	16.703
传统方法		13.591	4.282	6.856	22.840	21.838
采样区间=10 分钟						
PC-AR(3)	100	10.553	3.131	5.665	18.780	16.328
传统方法		13.112	4.112	6.616	21.875	21.035
采样区间=15 分钟						
PC-AR(3)	100	10.543	3.241	5.515	18.390	17.238
传统方法		12.699	3.978	6.499	20.952	20.209

资料来源: 海通证券研究所

表 3 的结果显示, 当我们使用动态执行时, 主成分分解方法依然对传统方法有着压倒性的优势 (第 2 列) 这也肯定了本文模型在实际交易中的有效性。然而, 在确定哪一种时间区间的长度更适合进行主成分分解并进行 VWAP 策略动态执行时, 却并不如理论情形那么一目了然。从所有股票的平均 MAPE 这一重要变量的数值来看, 较长的时间区间确实能获得更低的 MAPE。这种单调性似乎提示我们应该选择更长的采样区间来分解主成分并预测 VWAP。但是, 如果放在实际交易层面考虑, 那么较短的时间间隔应该是更优的选择。因为对于那些流动性较好的公司, 一旦时间较长, 如何在这段时间内分配需要交易的证券数量同样是个问题, 而且想以这么长时间内的平均成交价完成交易也是不太现实的。而且, 从 MAPE 的均值来看, 采样区间为 2 分钟的结果并没有比 15 分钟的差太多。所以, 我们认为用 2 分钟的换手率数据来进行主成分分解和 VWAP 策略动态执行是最优的选择。表 5 比较了每个股票用传统方法, PC-AR(3)模型的理论 and 动态执行这三者得到的 MAPE。

表 4 上证 50 每个样本股三种方法的日 VWAP 预测的 MAPE 比较 (采样区间=2 分钟, 单位: bp)

证券名称	传统方法	理论 PC-AR(3)	动态 PC-AR(3)	两种方法的差		
				理论 PC-AR(3)	动态 PC-AR(3)	理论 PC-AR(3)
				传统方法	传统方法	动态 PC-AR(3)
浦发银行	12.693	5.348	9.957	-7.344	-2.736	-4.609
华夏银行	16.330	5.749	12.971	-10.581	-3.359	-7.222
民生银行	13.404	4.442	10.624	-8.962	-2.780	-6.182
宝钢股份	10.136	4.105	8.995	-6.031	-1.141	-4.890
中国石化	9.813	5.024	7.639	-4.789	-2.174	-2.615
中信证券	15.537	4.513	13.444	-11.024	-2.093	-8.931
三一重工	16.463	4.731	12.871	-11.731	-3.592	-8.139
招商银行	11.651	4.902	8.679	-6.750	-2.972	-3.777
保利地产	18.206	6.020	12.396	-12.186	-5.810	-6.376
中国联通	11.747	5.262	10.421	-6.485	-1.326	-5.159

特变电工	11.989	4.319	9.283	-7.671	-2.706	-4.964
上海汽车	15.870	4.837	14.976	-11.032	-0.893	-10.139
包钢稀土	23.638	6.801	20.582	-16.836	-3.056	-13.780
兖州煤业	19.783	6.871	16.715	-12.912	-3.068	-9.844
阳泉煤业	22.157	6.239	18.810	-15.918	-3.347	-12.571
江西铜业	15.830	5.098	12.024	-10.731	-3.805	-6.926
金地集团	17.271	5.163	10.861	-12.108	-6.410	-5.698
中金黄金	12.089	2.922	11.155	-9.167	-0.934	-8.233
贵州茅台	13.366	4.279	7.876	-9.087	-5.490	-3.597
山东黄金	12.823	3.285	10.445	-9.538	-2.378	-7.160
海螺水泥	17.723	5.070	12.855	-12.653	-4.868	-7.784
海通证券	12.649	4.811	9.846	-7.838	-2.803	-5.036
长江电力	6.740	2.561	6.471	-4.179	-0.269	-3.910
大秦铁路	9.958	3.844	7.031	-6.114	-2.927	-3.187
中国神华	17.229	5.320	12.491	-11.909	-4.738	-7.171
中国国航	17.315	5.539	13.439	-11.776	-3.875	-7.900
海南橡胶	23.204	6.206	14.566	-16.998	-8.637	-8.360
兴业银行	13.101	4.059	10.382	-9.042	-2.719	-6.323
西部矿业	17.033	5.860	14.216	-11.174	-2.817	-8.357
北京银行	10.537	4.355	9.463	-6.182	-1.074	-5.108
农业银行	10.871	6.039	8.210	-4.832	-2.661	-2.171
中国平安	10.606	4.335	9.713	-6.271	-0.893	-5.378
交通银行	8.296	5.134	7.278	-3.161	-1.018	-2.144
中国中铁	7.347	3.527	6.370	-3.820	-0.977	-2.843
工商银行	10.076	5.237	9.149	-4.839	-0.927	-3.912
中国铝业	19.705	5.545	15.317	-14.161	-4.388	-9.772
中国太保	15.232	6.873	12.138	-8.359	-3.094	-5.264
中国人寿	7.677	4.075	7.760	-3.602	0.083	-3.685
中国建筑	12.508	6.150	10.636	-6.359	-1.872	-4.487
华泰证券	15.326	6.583	12.689	-8.744	-2.637	-6.107
潞安环能	20.064	7.603	16.293	-12.462	-3.771	-8.691
中国南车	15.171	3.910	10.200	-11.260	-4.971	-6.289
光大银行	11.325	5.002	10.014	-6.323	-1.311	-5.012
中国石油	8.860	4.272	7.621	-4.588	-1.239	-3.349
中煤能源	14.330	5.866	11.510	-8.464	-2.820	-5.644
紫金矿业	13.442	3.838	11.812	-9.605	-1.630	-7.974
中国远洋	7.265	3.956	6.659	-3.309	-0.607	-2.702
建设银行	9.242	5.241	7.520	-4.001	-1.722	-2.278
金钼股份	18.167	7.236	16.529	-10.931	-1.637	-9.293
中国重工	17.863	5.664	13.129	-12.200	-4.735	-7.465

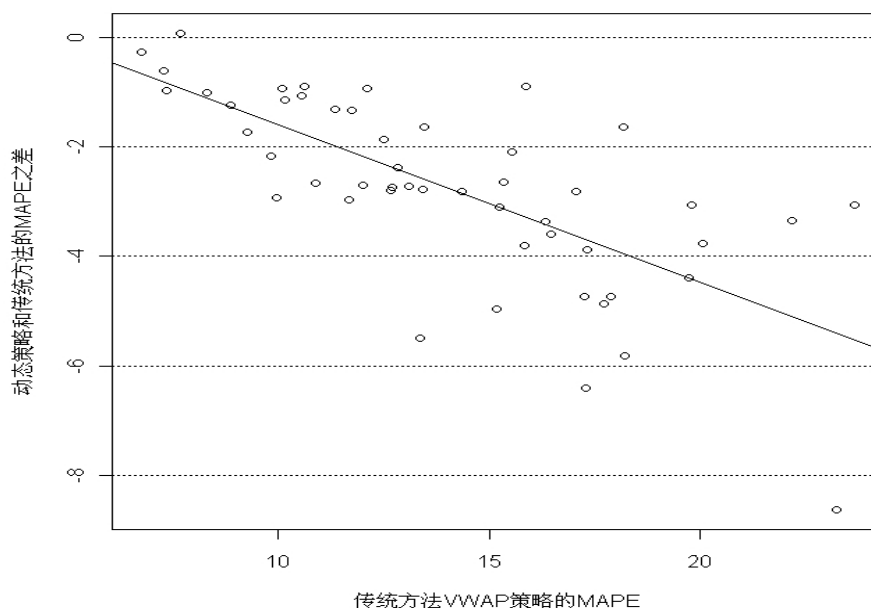
资料来源：海通证券研究所

我们先从分析理论执行的结果开始。如上表所示，PC-AR(3)模型的效果是显著的，不仅 50 个样本股都超越了对应传统方法的结果，而且 MAPE 降低幅度最小的也有 3.2bp，40% 的股票（20 个）都降低了至少 10bp。这些优良的结果让我们对主成分方法充满了信心，但在实务中还是需要考察策略动态执行的效果，因为理论方法毕竟只是一个无法实现的最优情况。

从策略动态执行的结果来看，49 个股票的 PC-AR(3)模型的 VWAP 策略执行成本低于传统方法，其中 16 个股票的成本比传统方法减少了 3bp 以上，占到所有样本股的 32%。唯一一个执行成本高于传统方法的是中国人寿(0.083 bp)，基本可以认为等同于传统方法。可见动态执行能够有效地降低成本，但离改善的上限（理论执行）还有一定的距离。造成这一损失的原因有两个，一是在开盘前我们无法获知当天交易的所有信息；二是即使有新的信息表明我们在当天已完成的交易中犯了错误，也无法擦除或修改它们。我们能做的只是在获得新的有关成交量的信息后及时调整当天剩余时间的策略。那些已经犯下的错误而导致的损失可以通过计算理论和动态执行这两者 MAPE 的差来衡量。上表中最后一列显示，这一差值相当不稳定，资源类股票动态执行的损失最大，包钢稀土从理论执行的 6.801bp 上升至 20.582bp，阳泉煤业从 6.239bp 升至 18.810bp。而在大市值的银行股上，这两种执行方式的表现较为接近，农业银行、建设银行和交通银行动态执行的损失都没有超过 2.5bp。

最后，我们讨论一个有趣的发现。直观上看，传统方法预测的 MAPE 越大，主成分分解方法的改进幅度也越大。如果真是这样，就表明本文的方法能够有效地改善那些由于使用传统方法而造成的巨大误差。下面，我们以传统方法的 MAPE 为自变量，动态执行和传统方法的差为因变量做回归分析，考察是否真的存在这样一个显著的关系。图 2 是对应的散点图和回归直线。可以看出，传统方法的 MAPE 越大，动态执行获得的改进也越大。当然，在传统方法较为有效的区间（MAPE<10bp），用主成分分解加入成交量日内的动态变化并不能显著降低 VWAP 策略的执行成本。相反，一旦传统方法的误差变大时，改善就变得相当明显。这一结论保证了本文方法在任何情况下都是有效的。

图 2 传统方法的 MAPE 与动态策略改善程度的关系



资料来源：海通证券研究所

4. 结论

本文中，我们提出了一个全新的方法对日内成交量的动态变化进行建模，从而有效降低了 VWAP 策略的执行成本。基于主成分分析，成交量或换手率被分解成两个部分：一个反映了因整个市场波动而导致的成交量变化，可以用静态的历史均值来估计；另一个则描述了每个股票特有的成交量形态，我们用 ARMA 模型来拟合。

采用这一方法使我们获得了一个预测成交量较为精确的统计模型，不仅解决了成交量呈周期性波动这一建模中的难点，而且成功将预测结果应用到实际交易中，有效降低了 VWAP 策略的执行成本。

当然，本文的模型依然存在改进的空间，我们将在后续的报告予以呈现。

信息披露

分析师声明

吴先兴：金融工程

本人具有中国证券业协会授予的证券投资咨询执业资格，以勤勉的职业态度，独立、客观地出具本报告。本报告所采用的数据和信息均来自市场公开信息，本人不保证该等信息的准确性或完整性。分析逻辑基于作者的职业理解，清晰准确地反映了作者的研究观点，结论不受任何第三方的授意或影响，特此声明。

法律声明

本报告仅供海通证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议。在任何情况下，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。

本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。

市场有风险，投资需谨慎。本报告所载的信息、材料及结论只提供特定客户作参考，不构成投资建议，也没有考虑到个别客户特殊的投资目标、财务状况或需要。客户应考虑本报告中的任何意见或建议是否符合其特定状况。在法律许可的情况下，海通证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

本报告仅向特定客户传送，未经海通证券研究所书面授权，本研究报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。如欲引用或转载本文内容，务必联络海通证券研究所并获得许可，并需注明出处为海通证券研究所，且不得对本文进行有悖原意的引用和删改。

根据中国证监会核发的经营证券业务许可，海通证券股份有限公司的经营经营范围包括证券投资咨询业务。

海通证券股份有限公司研究所

汪异明 所长
(021) 63411619
wangym@htsec.com

高道德 副所长
(021) 63411586
gaodd@htsec.com

路颖 副所长
(021) 23219403
luying@htsec.com

江孔亮 所长助理
(021) 23219422
kljiang@htsec.com

宏观经济研究团队

李明亮 (021) 23219434
汪辉 (021) 23219432
刘铁军 (021) 23219394

lml@htsec.com
wanghui@htsec.com
liutj@htsec.com

策略研究团队

陈瑞明 (021) 23219197
吴一萍 (021) 23219387
荀玉根 (021) 23219658

chenrm@htsec.com
wuyiping@htsec.com
xyg6052@htsec.com

基金研究团队

娄静 (021) 23219450
单开佳 (021) 23219448
倪韵婷 (021) 23219419
罗震 (021) 23219326
唐洋运 (021) 23219004

loujing@htsec.com
shankj@htsec.com
niyt@htsec.com
luozh@htsec.com
tangyy@htsec.com

联系人

高远 (021) 23219669
李宁 (021) 23219431

gaoy@htsec.com
lin@htsec.com

联系人

王旭 (021) 23219396
汤慧 (021) 23219733
李珂 (021) 23219396

wx5937@htsec.com
tangh@htsec.com
lk6604@htsec.com

联系人

陈瑶 (021) 23219645
伍彦妮 (021) 23219774
桑柳玉 (021) 23219686
曾逸名 (021) 23219773

chenyao@htsec.com
wyn6254@htsec.com
sly6635@htsec.com
zym6586@htsec.com

金融工程研究团队

吴先兴 (021) 23219449
丁鲁明 (021) 23219068
郑雅斌 (021) 23219395

wuxx@htsec.com
dinglm@htsec.com
zhengyb@htsec.com

固定收益研究团队

姜金香 (021) 23219445
徐莹莹 (021) 23219885

jiangjx@htsec.com
xyy7285@htsec.com

政策研究团队

陈久红 (021) 23219393
陈峥嵘 (021) 23219433

chenjiuhong@htsec.com
zrchen@htsec.com

联系人

冯佳睿 (021) 23219732
朱剑涛 (021) 23219745
张欣慰 (021) 23219370

fengjr@htsec.com
zhujt@htsec.com
zxw6607@htsec.com

联系人

武亮 (021) 23219883
黄轩 (021) 23219886

wl7222@htsec.com
hx7252@htsec.com

计算机及通信行业

陈美凤 (021) 23219409
联系人
袁兵兵 (021) 23219770
蒋科 (021) 23219474

chenmf@htsec.com
ybb6053@htsec.com
jiangk@htsec.com

煤炭行业

朱洪波 (021) 23219438
刘惠莹 (021) 23219441

zbb6065@htsec.com
liuhy@htsec.com

批发和零售贸易行业

路颖 (021) 23219403
潘鹤 (021) 23219423
汪立亭 (021) 23219399
联系人
李宏科 (021) 23219671

luying@htsec.com
panh@htsec.com
wanglt@htsec.com
lhk6064@htsec.com

建筑工程行业

江孔亮 (021) 23219422
联系人
赵健 (021) 23219472
张显宁 (021) 23219813

kljiang@htsec.com
zhaoj@htsec.com
zxn6700@htsec.com

石油化工行业

邓勇 (021) 23219404
联系人
王晓林 (021) 23219812

dengyong@htsec.com
wxl6666@htsec.com

机械行业

龙华 (021) 23219411
联系人
何继红 (021) 23219674
熊哲颖 (021) 23219407

longh@htsec.com
hejh@htsec.com
xzy5559@htsec.com

农林牧渔行业

丁频 (021) 23219405
联系人
夏木 (021) 23219748

dingpin@htsec.com
xiam@htsec.com

纺织服装行业

卢媛媛 (021) 23219610

lly5991@htsec.com

非银行金融行业

董乐 (021) 23219374
联系人
黄媚 (021) 23219638
汤婧 (021) 23219809

dl5573@htsec.com
hm6139@htsec.com
tj6639@htsec.com

电子元器件行业

邱春城 (021) 23219413
联系人
张孝达 (021) 23219697

qiucc@htsec.com
zhangxd@htsec.com

互联网及传媒行业

联系人
白洋 (021) 23219646
薛婷婷 (021) 23219775

baiyang@htsec.com
xtt6218@htsec.com

交通运输行业

钮宇鸣 (021) 23219420
钱列飞 (021) 23219104
联系人
虞楠 (021) 23219382
李晨 (021) 23219817

ymniu@htsec.com
qianlf@htsec.com
yun@htsec.com
lc6668@htsec.com

汽车行业

赵晨曦 (021) 23219473
冯梓钦 (021) 23219402

zhaocx@htsec.com
fengzq@htsec.com

食品饮料行业

赵勇 (0755) 82775282
齐莹 (021) 23219166

zhaoyong@htsec.com
qiy@htsec.com

钢铁行业

刘彦奇 (021) 23219391

liuyq@htsec.com

医药行业

联系人
刘宇 (021) 23219608
刘杰 (021) 23219269
冯皓琪 (021) 23219709

liuy4986@htsec.com
liuj5068@htsec.com
fhq5945@htsec.com

有色金属行业

联系人
刘博 (021) 23219401

liub5226@htsec.com

基础化工行业

曹小飞 (021) 23219267
联系人
易团辉 (021) 23219737
张瑞 (021) 23219634

caoxf@htsec.com
yith@htsec.com
zr6056@htsec.com

家电行业 陈子仪 (021) 23219244 联系人 孔维娜 (021) 23219223	chenzy@htsec.com kongwn@htsec.com	建筑建材行业 联系人 赵健 (021) 23219472	zhaoj@htsec.com	电力设备及新能源行业 张浩 (021) 23219383 牛品 (021) 23219390 联系人 房青 (021) 23219692 徐柏乔 (021) 23219171	zhangh@htsec.com np6307@htsec.com fangq@htsec.com xbq6583@htsec.com
公用事业 陆凤鸣 (021) 23219415 联系人 汤砚卿 (021) 23219768	lufm@htsec.com tyq6066@htsec.com	银行业 联系人 刘瑞 (021) 23219635	lr6185@htsec.com	社会服务业 林周勇 (021) 23219389 联系人 胡宇飞 (021) 23219810	lzy6050@htsec.com hyf6699@htsec.com
房地产业 涂力磊 (021) 23219747 谢盐 (021) 23219436 联系人 贾亚童 (021) 23219421	tll5535@htsec.com xiey@htsec.com jiayt@htsec.com	造纸轻工行业 徐琳 (021) 23219767 联系人 马浩博 (021) 23219822	xl6048@htsec.com mhb6614@htsec.com		

海通证券股份有限公司机构业务部

陈苏勤 总经理
(021) 63609993
chensq@htsec.com

贺振华 总经理助理
(021) 23219381
hzh@htsec.com

深广地区销售团队 蔡铁清 (0755) 82775962 刘晶晶 (0755) 83255933 辜丽娟 (0755) 83253022 高艳娟 (0755) 83254133	ctq5979@htsec.com liujj4900@htsec.com gulj@htsec.com gyj6435@htsec.com	上海地区销售团队 高溱 (021) 23219386 李唯佳 (021) 23219384 胡雪梅 (021) 23219385 黄毓 (021) 23219410 朱健 (021) 23219592 肖文宇 (021) 23219442 王丛丛 (021) 23219454 朱昂 (021) 23219397	gaoqin@htsec.com jiwj@htsec.com huxm@htsec.com huangyu@htsec.com zhuj@htsec.com xiaowu@htsec.com wcc6132@htsec.com za6533@htsec.com	北京地区销售团队 孙俊 (010) 58067988 张广宇 (010) 58067931 王秦豫 (010) 58067930 隋巍 (010) 58067944	sunj@htsec.com zgy5863@htsec.com wqy6308@htsec.com sw7437@htsec.com
--	---	--	--	--	--

海通证券股份有限公司研究所

地址: 上海市黄浦区广东路 689 号海通证券大厦 13 楼

电话: (021) 23219000

传真: (021) 23219392

网址: www.htsec.com