

## 高频交易及量化投资的策略与误区

### 量化投资一般依赖复杂的模型，而高频交易一般依赖运行高效的代码

广发期货 李尉

#### 1 高频交易公司和量化投资公司的区别

一般来说，高频交易公司和量化投资公司既有联系，又有区别。在美国，人们常说的高频交易公司一般都是自营交易公司，这些公司主要有 Getco、Tower Research、Hudson River Trading、SIG、Virtu Financial、Jump Trading、RGM Advisor、Chopper Trading、Jane Street 等；而常说的量化投资公司一般都是对冲基金，包括 RenTec、DE Shaw、Two Sigma、WorldQuant、AQR、Winton、BlueCrest、Citadel 等。此外，Citadel、Two Sigma 等公司，既有高频交易业务，又有量化投资业务；DE Shaw 等公司，既有量化投资，又有非量化投资——很多公司朝着更综合的方向发展。

从历史上看，很多高频交易公司的创始人都是交易员出身，原来就从事衍生品的做市、套利等业务。一开始这些工作并不需要多高深的知识。随着计算机技术的发展，交易的自动化程度和频率也逐渐提高，这些公司逐渐聘请一些数学、统计、计算机背景较强的人员加入以适应形势的发展。当然，这个过程也出现了一些分化，有的公司还是保留了交易员在公司的主导地位，并且始终未放弃人工交易，最终形成了人机结合的半自动交易；而另外一些公司对新鲜技术的接受程度更高一些，往往采取全自动的交易模式。事实上，也没有证据表明全自动交易的公司就比半自动交易的公司更为优越，到目前为止，也只能说是各有利弊。

人工交易的最大弊端在于手动下单的地方离交易所较远，在行情剧变的时候往往抢不到单。在这一点上，全自动交易的公司可以通过托管机房来最大程度减少信号传输的时间，不过自动化交易往往因为程序过于复杂，加上很多公司人员流动较大，在程序的维护上会出现一些失误，最终程序出错酿成大祸，比如著名的骑士资本。

至于过度拟合无法抵御黑天鹅事件，那是人工交易和自动交易都无法避免的问题。一般来说，Getco、Jane Street、SIG、Virtu Financial 等是半自动交易，Tower Research、Hudson River Trading、Jump Trading 等是全自动交易。

量化投资公司跟高频交易公司则有很大的不同。首先，美国的量化投资公司基本上都是量化背景极强的人创办的，比如说文艺复兴的创始人西蒙斯是数学家出身，DE Shaw 的创始人 David Shaw 是计算机教授出身，AQR 的创始人 Cliff Asness 是金融学家出身，而高频交易公司则更多是传统交易员创办的；其次，量化投资一般依赖于复杂的模型，而高频交易一般依赖于运行高效的代码。

量化投资公司的持仓时间往往达到 1—2 个星期，要预测这么长时间的价格趋势需要处理的信息自然非常庞大，模型也因此更为复杂，对程序的运行速度反而没那么敏感；高频交易处理信息的时间极短（微秒或毫秒级），不可能分析很多的信息，因此模型也趋于简单，竞争优势更多依靠代码运行的效率，很多人甚至直接在硬件上写程序；而最后，量化投资的资金容量可达几百亿美元，而高频交易公司往往只有几千万至几亿美元，但由于高频交易的策略表现远比量化投资稳定，如 Virtu Financial 交易 1238 天只亏 1 天，因此一般都是自营交易，而量化投资基金一般来说都是帮客户投资。

#### 2 量化交易的模型

下面介绍一下量化交易的模型，从简单到复杂：

最简单的以约翰·墨菲的《期货市场技术分析》为代表，最多用到指数、对数等高中层面的数学知识，通俗易懂，更适合主观交易，或者计算机计算并发出交易信号由人手动下单的半自动交易。

层次高一点的以丹尼斯的《海龟交易法则》为代表，数学上毕竟使用了均值、方差、正态分布等大学低年级数学的内容，策略的测试也更具科学性，而且提出了可靠的资金管理办法，但缺点是依旧没有摆脱传统的、依靠交易规则的排列组合进行交易的思路。不过，如果策略设计得好且行情出大趋势的话还是可以有不错的效果。

更高一级的层次主要体现在交易信号的整合方面，比如运用更现代的统计方法——回归分析、神经网络、支持向量机等对传统的技术指标进行有机整合，并使用更严格的统计方法进行变量的筛选及测试。考虑到金融数据的时间特征，往往需要使用滚动优化来获取样本外的测试结果，这样得出的模型也更为稳健。

不过，一般的程序化交易系统都难以实现这些功能，需要自己用更通用的编程语言来实现。

如果是量化投资，那么除了行情信息，还要收集整理其他基本面的信息，整理出对应的时间序列，并融入到预测模型中。一般来说，成功的模型不在于运用了多高深的数学理论，而在于它整合了多少不同来源的信息。即使是最简单的线性回归，如果各个参数都有很强的预测能力，且相关性很低，那么模型的预测效果也会很好；相反，即使运用复杂的深度学习理论，如果选取的参数毫无意义，最后得出的模型也没有用。目前美国一些公司不仅利用新闻等文本信息建模，甚至用到谷歌卫星拍摄到的港口集装箱的图像来建模，通过对商品集装箱的数目来预测商品价格的走势，取得了很好的预测效果。

建模是一回事，求解模型其实也同等重要。比如说物理学上有很多模型能精确描述现实，但经常由于缺乏高效的科学计算方法而难以求解。量化交易也一样的。参数的计算、筛选、优化、回测等往往伴随着巨大的计算量，如何巧妙求解是一门颇为高深的学问。据西蒙斯透露，著名的文艺复兴公司内部有着明确的分工——计算机程序员从各个来源收集数据，物理学家分析数据建立模型，数学家构建优化算法并求解模型等。

### 3 高频、量化领域常见的误区

#### 量化模型无法战胜黑天鹅事件

事实上，任何投资方法都是依靠历史预测未来，都害怕黑天鹅事件，都会有回撤。量化的好处在于遇到回撤之后，可以迅速把最新的情况纳入模型，及时调整，重新回测、优化、模拟，争取在最短的时间内扭转损失。比如文艺复兴在 2007 年 8 月遭遇历史上罕见的 9% 回撤之后，西蒙斯采取果断措施，重新建模，在致投资者的信中他宣称“我们新的模型已经发现了 3 个很强的交易信号”，结果在接下来的日子很快扳回损失，当年的收益率达到 80%。

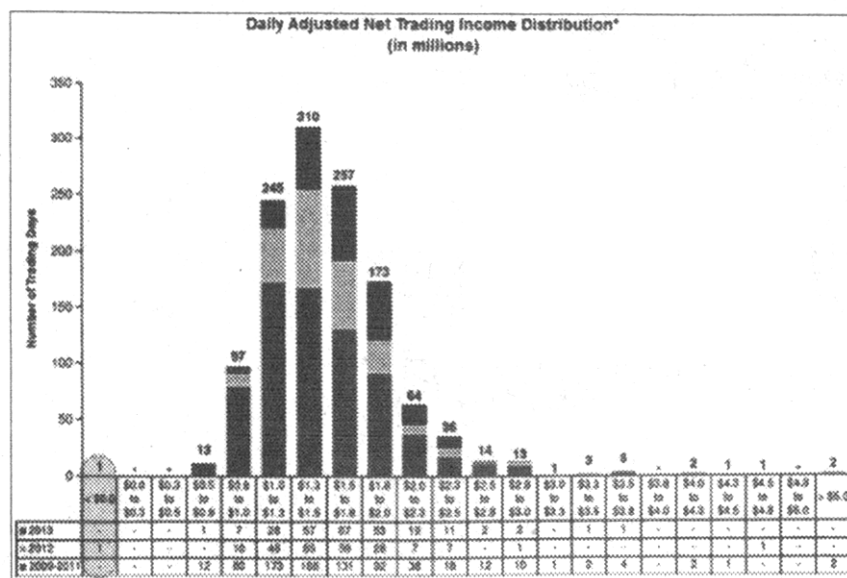
长期资本管理公司（LTCM）就是因为用了量化模型而破产的。事实上，LTCM 是一个多策略基金，它的纯量化交易策略最后在 1998 年还是赚了 1 亿美元，它亏损最多的策略都是交易流动性极差的柜台衍生品，很多甚至是它自己设计来跟投行对赌的产品，遇到黑天鹅事件无法及时清理头寸。这些产品一般只是在定价时候使用量化模型辅助一下，具体的交易执行、产品设计、销售等都跟量化无关，一般认为 LTCM 的破产更多是因为流动性风险，跟模型关系不大。

#### 高频交易损害投资者利益

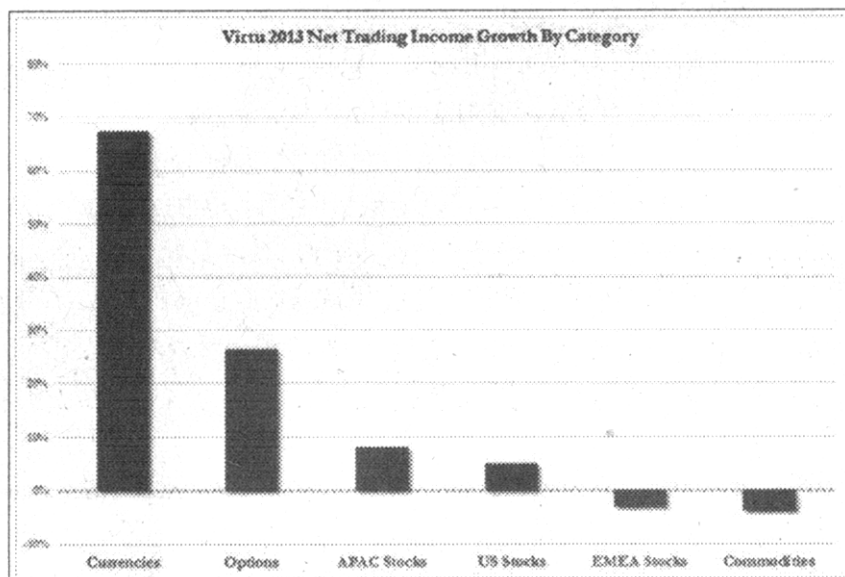
像《Flash Boys》等书籍的观点其实都很有争议的，只不过作者文笔极佳，叙事手法极富煽动性，所以才吸引了众多的眼球。除了媒体之外，应该说美国目前要求禁止高频交易最为强烈的，基本上都是当年的传统交易员。正因为新兴的、依靠先进技术的高频交易公司把他们打败了，他们心有不甘，所以才组织更多的力量来进行反击。由于这些人都是市场老手，所以对这个市场还是非常熟悉的，提的观点也有可取的地方。

在国内，现在期权准备上市，股票也很可能开放 T+0。对这两块“肥肉”，国外高频交易商早就垂涎已久。如果说在期货高频领域，我们还能依靠在程序化交易上的丰富经验与国外抗衡一下，

那么在期权和股票高频领域，我们的实践经验为零，跟国外的差距更大。对此，笔者认为，我们一方面不能妄自菲薄，觉得外资太厉害就干脆不做了；另一方面也不能急于求成，妄想一年半载就要取得很大的成绩。凡事都要本着谦虚谨慎的态度，国外很多高手来到国内都是先研究一年才能稳定盈利，国内的人基础较薄，研究周期要长一些，比如第一年做准备工作开发系统，第二年逐渐打平手续费，之后开始盈利，或许更为合理。策略研究要慢工出细活，急于求成，频繁改变研究方向，最终很可能一事无成。



图为 Virtu Financcail 每日调整后净交易利润统计图(2009年1月1日至2013年12月31日)



图为 Virtu Financial 各领域盈亏状况