

红药丸还是蓝药丸？量化交易策略研究的常见误区

卡图拉^{a,1}

^a *cryptosmartbeta.com*

E-mail: katula@cryptosmartbeta.com

ABSTRACT: 量化交易策略研究是金融分析中最富于挑战性的方向之一，因为任何环节都可能出现难以察觉的问题，真正在残酷的市场环境中生存下来并不容易。帕累托原理和马太效应带来的累积优势在资产管理行业一如其它任何行业般存在，甚至更加普适。量化分析通常可能步入哪些误区？资深的投资组合经理会如何处理这些问题？本文介绍交易策略开发过程中可能出现的问题，成因以及可能的解决方案。

¹通讯作者。

Contents

1 引言	1
2 过度拟合	2
3 前视偏差	3
4 选择偏差	3
5 合并价格的可执行度	5
6 日内交易的 K 线可靠度	6
7 交易成本	6
8 结语和头脑风暴	7

1 引言

在著名赛博朋克电影《黑客帝国》中，如果你选择蓝药丸，故事将就此结束，你在床上醒来之时只会记得你之前所知的一切——谬误的现实，虚幻的安全感、幸福感和选择性无视真相带来的欢欣鼓舞。然而，如果你选择红药丸，你会得到知识和自由，并以直面逆境的勇气去承受关于现实世界的残酷真相。



你是否觉得整个市场都在和你作对？为什么总是一买就跌，一卖就涨？庄家为啥老是盯着我这小仓位反向操作？暴涨的怎么全都是别人家的币？

首先，这是个心理学效应，是物种进化的结果——人类天然地厌恶风险（risk averse），你做得其实比你想象的要好，只不过本能地把不好的结果放大了。你并不是唯一一个体验者，所有专业的对冲基金交易员都需要克服类似的心理障碍：市场下跌给人带来的负面情绪和记忆比起市场上涨同样幅度带来的正面情绪要强得多，如果不加以控制，一定会影响你的交易决定。这就是为什么许多散户选择在错误的点位追涨杀跌。你需要坚持使用一套系统的交易方法，并几年如一日地去慢慢改进它。

当然，事情要是仅仅这么简单就好了。在心理学效应之外，交易策略的构造很容易陷入若干误区，即使是老资格的基金经理也需要时时小心。下面让我们谈谈新策略的研究容易出现哪些问题，以及如何妥善处理。本文第2-7章节分别讨论过度拟合，前视偏差，选择偏差，合并价格可靠度，日内价格曲线和交易成本的常见问题，原因和可能的解决方案，并在第8章节给出总结和若干头脑风暴式的开放性问题供感兴趣的读者思考。

本文面向所有背景的读者，讨论的内容也适用于数字货币之外的任意交易策略研究。

2 过度拟合

你找到一只币，采用几种常见的指标，对每个指标在一个很大的参数区间进行优化调节，再汇总得到最优的一套参数。可是当你把它用在虚拟或者实盘交易的时候，信号消失了或者大幅度减弱。

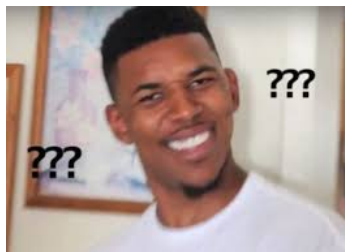


卡图拉如是说： 过于简单的指标，例如技术分析指标通常没有很好的预测能力，当你通过历史回测对它们进行优化的时候，首先已经默认这个指标是有价值的，但这个假定未必是对的，事前需要仔细审视。其次，即使这个指标有一定指导意义，也需要时刻避免过度拟合，因为它会带来比真实情况乐观得多的结果。

解决方案： 按照系统性的标准选择合适的指标进行研究。好的策略通常基于一个十分直观合理的解释，不应该太复杂，但同时也不能过分简单，所有人都能看到的数据（例如 K 线图）里通常不容易找到特别好的策略。采用保守的方法进行数据拟合，或是采用人工智能、机器学习的方法寻找信号。这些进阶的研究方式适合在大数据中寻找并提取微弱的信号，最近时有 IT 界高端人才加入著名高频交易公司或对冲基金。随着 AI 技术的持续进步，金融市场也势必变得越来越高效，阿尔法策略的研究将逐渐出现业余爱好者难以企及的门槛。如果你有心在业余时间往机器学习的方向发展，Kaggle 的题库是一个不错的入门方法。

3 前视偏差

你找到市值最高的几十只币，查看它们的历史业绩表现，看起来很有特点，完美！于是开始用它们构造你的交易策略。可是历史回测结果明明不错的信号，在实盘测试里大相径庭。



卡图拉如是说： 你的历史数据是有没有附带来自将来的信息？在任何量化策略研究里，数据的实时性（point-in-time）是非常重要的。这意味着可用的数据必须是在那个时间节点确切可以得到的数据。当你选择当前市值最高的一些主流币做历史回测的时候，实际上暗含了一个前提——你已经知道它们今天是成功的。可是几年前的你并不知道这一切！根据我们在引文 [1] 中的统计，历史上的高市值币种后来大部分都跑不赢比特币。你如果在当时用同样的方法选币，结果将会大不相同。

还有一种隐含的前视偏差和数据提供商的数据回填（backfill）方式有关，例如一个数据公司 2015 年成立，但提供 2011 年起的数据。这时一定要问清楚 2011-2015 年间他们的数据是怎么来的——是实时收集的呢，还是别处买的，或者自己按照某种规则回填的？简单的价格、交易量、市值数据还好，但略微复杂的数据如新闻、市场情绪、币种评级、公司基本面必须实时采集才是真实可靠的，因为它们很容易在事后被有意无意地修改以符合某种预期。回填的数据往往带有某种后知后觉（hindsight），可信度需要视具体情况打一个不小的折扣。

解决方案： 尽可能基于实时的数据和当时能够获取的信息构造交易策略。一切信息、新闻、数据的采集注意提取准确的时间戳（timestamp）。回填数据时尽量采取保守的估值方法，避免带入那个节点不可能获取的任何来自未来的信息。

4 选择偏差

经过一段时间的精心准备和大量试验（trial and error），你找到了 100 只可以系统交易的策略。它们有的基于单个币种，有的针对多个币种，研究方式各不相同，但都有一定道理，并且历史回测表现良好。你把它们组合起来进行实盘测试，但是观察了一段时间发现，业绩与历史表现大不相同，基本上是平的！

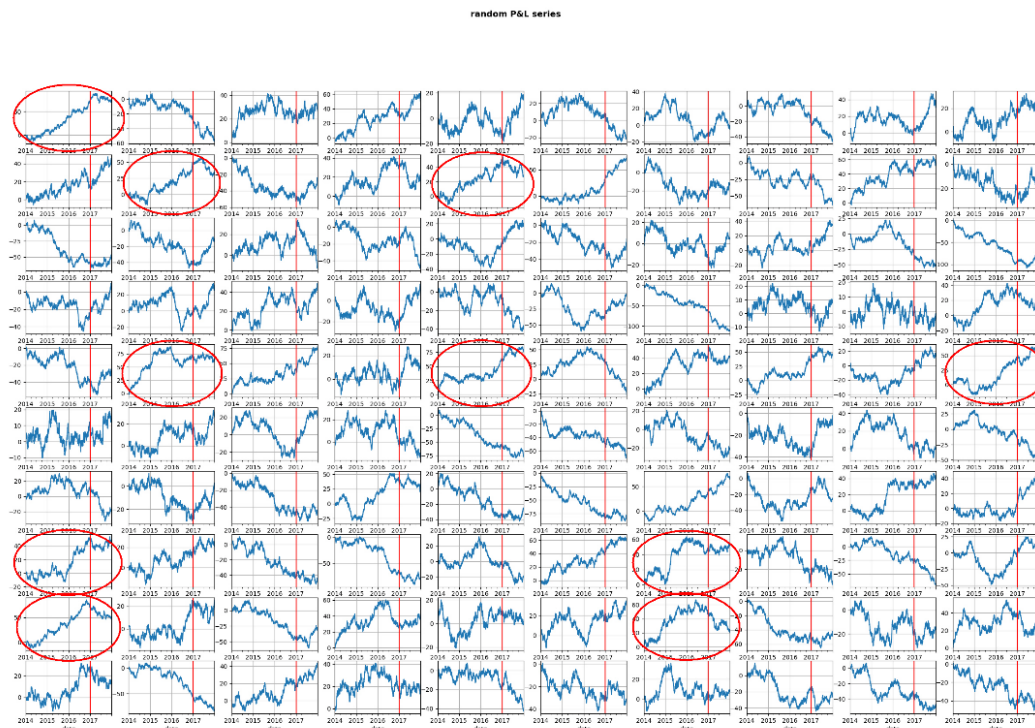


Figure 1. 由标准正态分布产生的 100 个日收益率随机时间序列的累积 P&L。样本内和样本外数据分居竖直红线的左、右边。圈出的图形是样本内总收益最高的 10 个序列。



卡图拉如是说： 不知道你平均需要尝试几次才能找到一个可用的策略呢？问题出在你的尝试次数上。如果你试了 100 次，最后发现其中只有 10 个能用，就需要十分小心。策略研究是一个残酷的领域，如果研究方式不当，并不是像爱迪生一样大量试错，最后做出一个重大发现就可以逆袭改变一切；恰恰相反，这种情况下你的新发现很可能只是统计学上的一次涨落。让我们看一个直观的例子。按照标准正态分布随机产生 100 个日收益率的时间序列，它们近年来的累积收益（cumulative P&L）曲线如图1。

直觉和简单的统计知识告诉我们，这一百只“策略”的最终回报也符合正态分布。我们把前三年取为测试样本（in-sample），最后一年用作样本外测试（out-of-sample test），分别对应红线左边和右边的区域。现在让我们挑选前三年表现最好的 10 个“策略”（在图1中

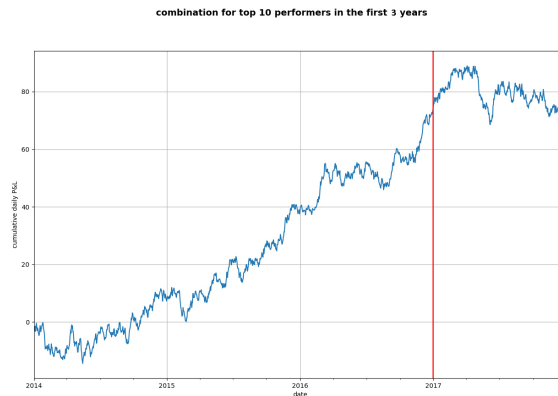


Figure 2. 测试样本中表现最好的 10 个策略等权重叠加之后的累积 P&L。

圈出做为标记)，用相等的权重把它们组合起来。图2是它的业绩表现。我们看到它在前三年表现强劲，但最后一年表现平平。原因是显而易见的，我们自始至终都只使用了一些随机的时间序列，实盘测试出来也只能是大量随机数叠加的结果，而根据大数定理，策略数目越多，实盘结果会越来越趋向于一条水平的直线。从图1也可以很直观地看到，任何一个策略盈利的概率大约是 50%，不论前三年还是最后一年都如此。现在你知道了吗？哪怕是随机产生的时间序列，只要数目足够多，做投资决定时就很容易挑中一些似是而非的策略，但它们其实只是统计学上的涨落，并不是真正可用的。更何况还有许多你觉得有一定道理，但实际上行不通的策略呢？

当你的研究进入到一定阶段，就会意识到真正的策略和随机序列的有效区分其实才是优秀的基金经理和一般人最大的差别。基金的交易小组里最重要的是两类人：投资组合经理（portfolio manager）基于特定的策略决定投资仓位，对自己的 P&L 负责并承担风险，而量化研究员（quantitative researcher）负责研究阿尔法（alpha）或贝塔（beta）模型，自己不承担交易风险。这两类人的比例在典型的基金里通常是 1:3 到 1:10，而区分他们最重要分水岭就是策略选择和投资决策的能力。这是每一个对冲基金都花费巨大的资源虚心对待的工作，很可能比起你之前想象的困难得多。

解决方案： 尽量选择有一定经济或金融学原理、符合价值投资理念的策略，避免过度复杂的研究方法。默认新策略历史回测和虚拟交易的表现可能大不相同，降低心理预期。任何策略都不能马上大规模交易，而需要先虚拟交易（paper trade）一段时间，再慢慢增加到最优仓位。虚拟交易的策略表现和所需的观察时间有关，如果很好可以略微缩短，反之则需要适当延长。

5 合并价格的可执行度

你对 CoinMarketCap 的数据仔细分析，找到了一些有趣的特征，构造了几个策略并实时跟踪了几个月（外样本测试，out-of-sample test），没有问题！但是在交易所进行实盘操作时，拿到的 P&L 差得多。

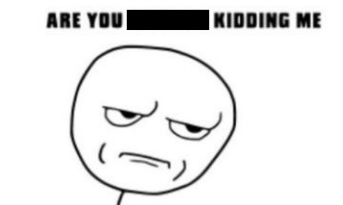


卡图拉如是说： 如果你的策略对价格较为敏感，通常是因为交易频繁导致换手率 (turnover) 高。这本身并没有问题，然而有没有想过历史回测的数据可能不大可靠呢？CoinMarketCap 和许多数据网站提供的是合并价格 (consolidated price)，也就是几个大交易所按照某种权重平均之后的价格，但落实到具体的交易所，未必能拿到这个价格，尤其如果你的交易量比较大的话。

解决方案： 尽量减少交易次数，并使用交易所提供的真实价格数据做历史回测。当然，这样的价格一般不够长，因此有时候不得不用合并价格去测试，但实盘交易之前至少可以用找得到的交易所数据确认一下近年的策略业绩表现。通常这比你的回测结果要差一些，日内交易尤其如此。

6 日内交易的 K 线可靠度

作为大数据分析达人的你发现币市价格波动比股市剧烈，特征信号强烈，于是基于 5 分钟线发展了一套机器学习策略，历史业绩强劲，但实盘交易差别很大。



卡图拉如是说： 越短的 K 线通常越不可靠，哪怕不同交易所都可能大不相同。另外，你知道“回”字有四种写法么？5 分钟线的定义也可以有 5 种，分别从第 1、2、3、4、5 分钟开始，你的日内策略如果只对其中一种定义有效，另外 4 种给出结果差不少，那就可能存在隐性的过度拟合。要知道实盘交易是需要时间也存在一定误差的，不同交易所的盘口深度 (order book depth) 不一，回测用的价格和成交价格很可能大不相同。

解决方案： 对于日内交易策略，用不同的方式构造 K 线并测试策略表现。如果不能在最坏的情况下盈利，用起来就需要格外小心。

7 交易成本

你的历史回测业绩是直线上升的！激动之余，你用小资金做了几天测试，结果业绩反过来了，直线下降？



卡图拉如是说： 历史回测可以选择考虑或者不考虑交易成本，但对于换手率一定要心知肚明。一个月换一次仓和每天换仓的交易成本差别巨大，大多数时候一个阿尔法看似存在，但扣除交易费用之后就什么都没有了。而且加密货币市场目前总体上交易量不算大，流动性整体不足，单单用买价和卖价 (bid and ask) 估计交易费用是不够的，还需要根据你的资金量计算盘口深度，或者按照每个币种的平均交易量去估计你的策略能够容纳的资金 (capacity)。

解决方案： 减少交易次数，降低交易金额，尽可能准确地估计交易费用。

8 结语和头脑风暴

本文讨论了量化策略研究中的常见误区，可能的成因和解决方案。当然，这只是一个轻量级的综述性分析，主要为非买方量化金融背景的读者准备。

看了这些例子，相信你对量化交易策略研究有了进一步的了解。下面让我们通过一些开放性问题来做一个有趣的头脑风暴。这些问题有可能出现在量化分析师和交易员的面试中，并且不限于数字货币领域，如果有一天你参加面试时被问起，可千万不要感到意外！

1. 你觉得现在币圈形形色色的百倍币预测可能步入哪些误区？如果你同时预测几十只币会怎么样？即使没有事后修改，那些事先保存的预测截图可信吗？
2. 如果你经常做预测，并且有很多人参考你的预测做交易，他们会相信你吗？在不作弊的前提下，你能想到哪些几乎百发百中的预测方法？小提示：如果预测的时候只给进场点，不给出场点会怎么样呢？
3. 为什么可审计或验证的历史交易记录 (track record) 那么重要？为什么专业的对冲基金都会对投资人强调“未来业绩和过往表现可能完全无关”？为什么他们只接受符合资质的投资人 (accredited investor) 参与？
4. 对冲基金需要发行去中心化的通证吗？有哪些优点和缺点？你能想到哪些应用场景？通证怎样定价？会比基金提供的赎回价格高吗？为什么？如果在币市交易，应该向投资人收取和返还法币还是加密货币？基金收益 (P&L) 的结算应该用法币还是加密货币？共同基金和对冲基金的结算方式会有区别吗？
5. 你会怎样挑选合适的基金经理去投资呢？换个角度，如果你是 CEO，会怎么挑选优秀的投资经理和研究员为你工作？

References

- [1] 卡图拉, *cryptosmartbeta.com* 创始人, <https://cryptosmartbeta.com/fallen-angels/>, 2018.