

Wikipedia of Trend Following (II)

原创：石川 川总写量化 1周前



作者：石川，北京量信投资管理有限公司创始合伙人，清华大学学士、硕士，麻省理工学院博士。

封面来源：<https://www.pexels.com>

未经授权，严禁转载。

摘要 本文在 Wikipedia of Trend Following 的基础上，继续讨论趋势策略的方面。



引言

近日，《贫穷的本质》这本书突然火了。它的英文原作是 *Poor Economics*，出版于 2011 年，并于当年就斩获了 Financial Times and Goldman Sachs Business Book of the Year Award。中文版《贫穷的本质》于 2013 年 4 月（六年多前）就由中信出版集团带到了国内。然而它却突然在 2019 年 10 月的某一天火了。究其原因，是因为其作者——来自 MIT 的 Abhijit Banerjee 以及 Esther Duflo 教授——连同 Harvard 的 Michael Kremer 教授分享了 2019 年诺贝尔经济学奖。从该书的作者获奖到该书在国内迅速走红毫无疑问是一个 trend following（趋势追踪）的鲜活实例，而这样的例子每时每刻都发生在我们身边。

在金融市场，trend following 更是最具生命力的一种交易策略。之前，公众号转载了我的好友刀疤连的关于趋势追踪的两万字雄文（我将其命名为《Wikipedia of Trend Following》），该文系统阐述了趋势追踪策略的是与非。考虑到我自己之前也写过多篇关于趋势追踪的文章（很多文章写的很早，恐怕只有最早关注公众号的朋友才有印象），因此今天也来就这个话题系统梳理一下；我将本文命名为 Wikipedia of Trend Following (II)。本文的正文部分共包含七小节：

第一节：时序与截面
第二节：确定计算趋势的时间尺度
第三节：哪个趋势指标更好使？
第四节：移动平均的本质
第五节：对趋势追踪的正确预期
第六节：从风险角度看趋势策略
第七节：结语

下面就让我们开启趋势之旅。

01

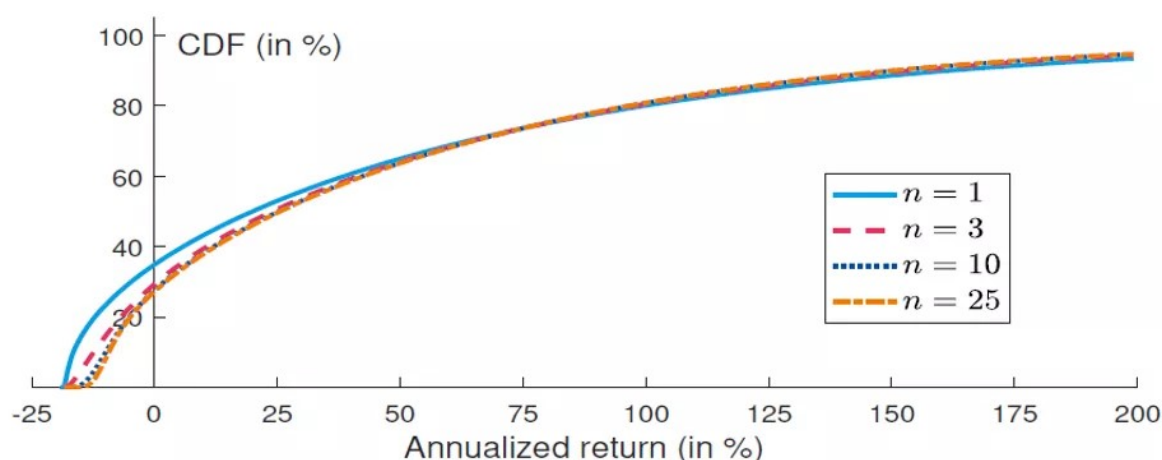
时序与截面

Trend Following 策略指的是投资者根据资产的上涨或者下跌趋势做出相应的多、空交易。具体来说，趋势追踪又可以分为时序动量（time-series momentum）和截面动量（cross-sectional momentum）两类。

在时序动量中，投资者做多之前上涨趋势显著的资产、做空之前下跌趋势显著的资产；而在截面动量中，投资者做多在截面上收益率**相对**较高的资产、做空在截面上收益率**相对**较低的资产。在时序动量中，我们希望资产可以各自延续其在之前的趋势。而在截面动量中，相比于每个资产绝对走势，我们更关注的是它们相对的强弱走势，希望强者恒强、弱者恒弱，因此截面动量又称为 winners-minus-losers 策略。

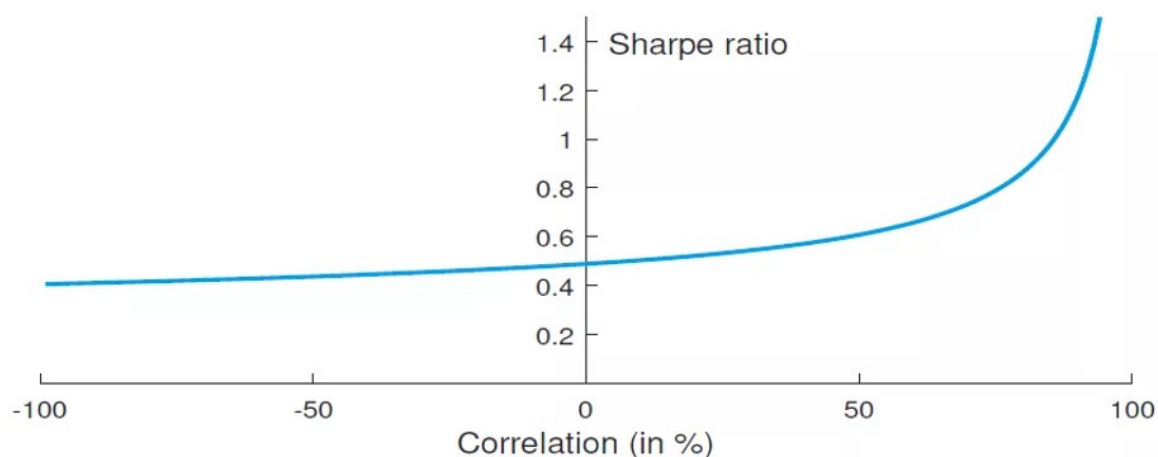
以下从资产相关性的角度比较一下时序和截面动量的差异。

多样化（diversification）是规避投资风险的重要思想，其体现是将资金分配到不同的资产中。对于时序动量策略来说，要想获得多样化收益（diversification gain），就必须考虑相关性低的资产。Roncalli（2017）指出，当资产之间的相关性很高时，时序动量策略的表现不会因为资产个数的增加而显著提升。下图为假设资产相关系数为 80% 的情况下，资产个数和时序动量策略年化收益率的累积分布函数的关系。由于资产之间的相关性很高，当资产个数超过 3 时，多样化带来的额外收益就几乎不发生改变。时序动量喜欢相关性低的资产。



出处：Roncalli (2017)

与时序动量不同，截面动量策略喜欢相关性高的资产。同样是来自 Roncalli (2017) 的研究表明，截面动量策略的夏普率随着资产相关性的增加而提高（下图）。关于时序动量和截面动量的更多讨论，感兴趣的朋友可参考《动量策略的是与非》。



出处：Roncalli (2017)

一般当人们提到 trend following 的时候，指的往往是时序动量策略，即多、空交易信号根据资产自己上涨或下跌的趋势而构建。本文接下来的讨论也将围绕时序动量展开。



确定计算趋势的时间尺度

想要通过 trend following 策略赚钱，最重要的一点是什么？很多文献告诉我们，趋势追踪中的重中之重包括 trend measure（即选择最优的指标来计算趋势）以及 position sizing（仓位管理）。毫无疑问，这两点对于趋势策略至关重要，但在我看来，**更重要的要数选择计算趋势的时间尺度。**

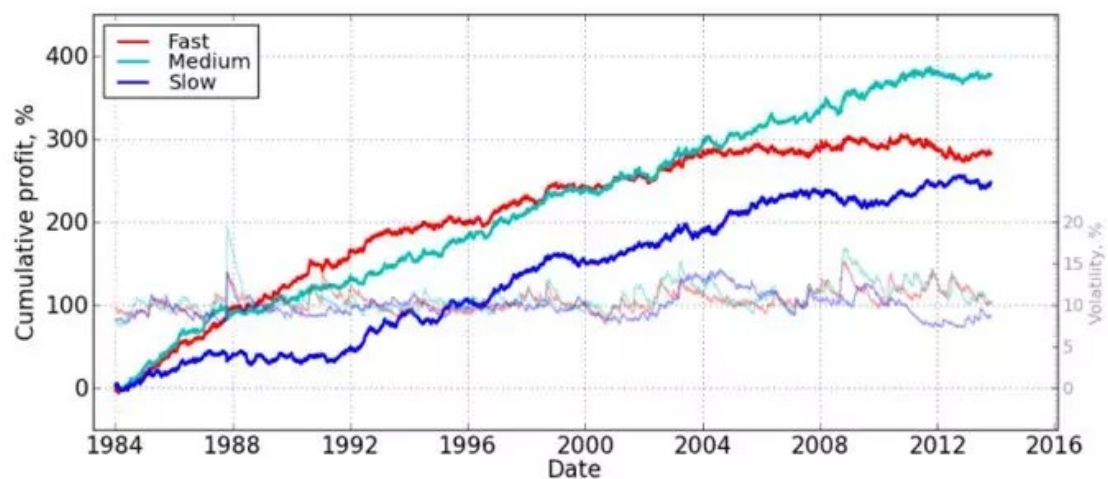
举个例子，Winton Capital（元盛——每当提及趋势追踪时，我脑海中想到的第一个名字）在 2013 年的一篇 research note（Winton 2013）中以下列来自四大类的 20 个资产进行了趋势策略时间尺度的研究。

Fixed Income	Stock Indices	Currencies	Commodities
US 10-yr T-note	S&P 500	Euro	Gold
Eurodollar	Nikkei 225 (3)	Japanese yen	Live cattle
Bunds (1)	Hang Seng (4)	British pound	Copper (7, 8)
JGB (2)	ASX 200 (5)	Swiss franc	WTI crude oil
Bonds	FTSE 100 (6)	Canadian dollar	Corn

These are now all traded on the CME except: (1) Eurex, (2) TSE, (3) OSE, (4) HKFE, (5) SFE, (6) LIFFE, (7) LME. (8) We synthesise prices for a Copper futures contract with a fixed expiry date, rather than a fixed time to expiry as found on the LME.

出处：Winton (2013)

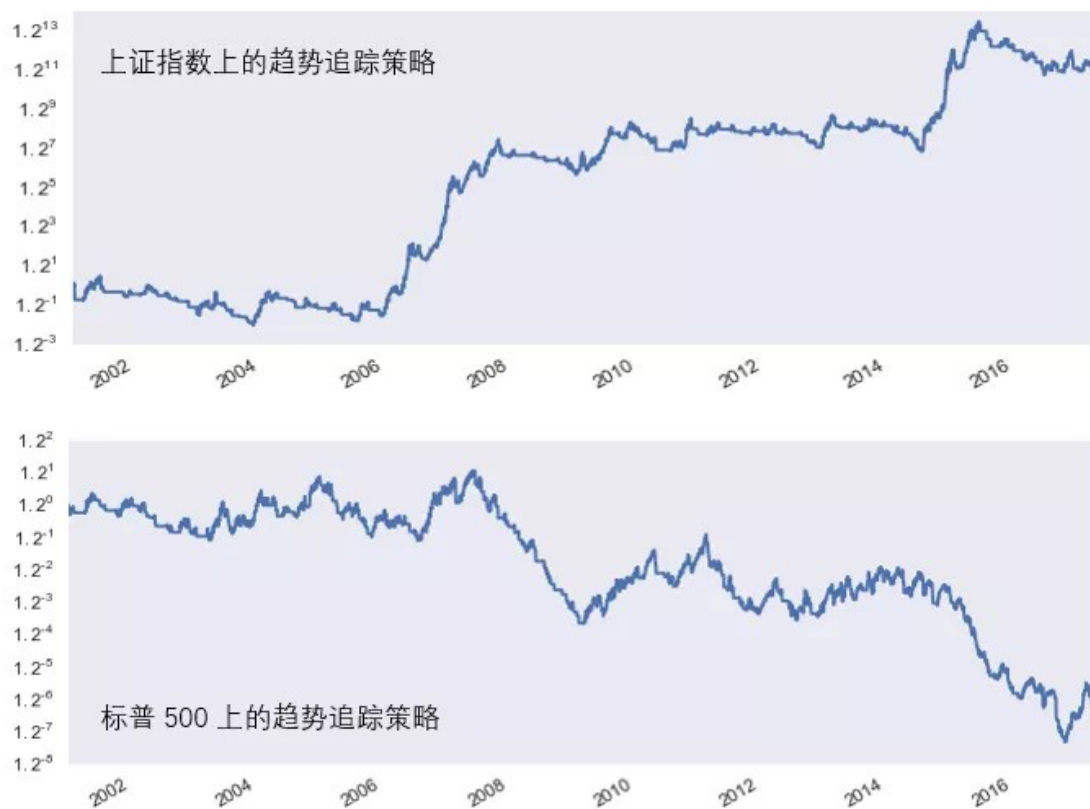
Winton (2013) 依据趋势追踪策略换手率的高低定义了 fast、medium 以及 slow 三个级别的策略（换手率越高意味着时间尺度越短），并交易上述 20 种资产。风险调整后，这三个级别策略的表现如下图所示。从中不难看出，fast 级别（换手率高）自 2004 年开始就基本失效了，而 medium 和 slow 级别策略在 2004 年之后依然有效；自 2012 年之后，medium 级别策略较其之前的表现也有所减弱，而 slow 级别的表现相对稳健，虽然其绝对收益不如 fast 和 medium 级别。



Returns are not compounded and volatility is determined by a 100-day exponentially-weighted estimate. The Sharpe ratios are 0.87, 1.12 and 0.81 for the fast, medium and slow strategies, respectively, for the 30-year period. The standard error on the Sharpe ratio is approximately 0.2 for 30 years of data. The realised volatility is around 10.5% for all three strategies.

出处：Winton (2013)

再来看两个离我们更近的例子。下图中，我用**同样的参数**分别以上证指数和标普 500 指数为资产，构建了仅做多的趋势追踪策略。结果显示，适合上证指数的趋势参数在标普 500 指数上完全不好使。



出处：川总写量化

上面的强烈对比说明：**趋势策略好使的前提是计算趋势的时间尺度和资产的收益率特征匹配。**

前文《什么样的收益率特性适合趋势追踪策略》使用如下的数学模型对上述观点进行了详细的论述，这里简单回顾下。考虑到资产收益率序列通常包含低频均值、非常微弱的自相关性以及白噪声三个部分构成，利用如下步骤人工合成对数收益率：

$$\begin{aligned}\text{Step 1: } b_0 &\sim \mathcal{N}(0, \sigma) \\ \text{Step 2: } e_t &\sim \mathcal{N}(0, \sigma) \\ \text{Step 3: } b_t &= \rho \times b_{t-1} + e_t, \quad t = 1, 2, \dots \\ \text{Step 4: } r_t &= b_t - \mu_b + \mu, \quad t = 0, 1, 2, \dots\end{aligned}$$

上述过程中， e_t 是白噪声序列，它满足均值为 0，标准差为 σ 的正态分布。 r_t 是人工对数收益率序列。为了使它满足给定的长期均值 μ 和自相关性 ρ ，构造过程中使用了辅助序列 b_t 。首先利用迭代构造出满足自相关性 ρ 的序列 b_t 。由于在构造 b_t 时可能会引入非零样本均值 μ_b ，因此在利用 b_t 构造 r_t 时，将 μ_b 去除并把给定的均值 μ 加回来。这样最终生成的对数收益率序列 r_t 同时满足的低频均值 μ 和自相关性 ρ 。

利用这个数学模型，以 μ 、 ρ 、 σ 作为输入，《什么样的收益率特性适合趋势追踪策略》一文进行了大量的仿真分析，研究了趋势策略的效果与参数组 μ 、 ρ 、 σ 的关系并得到如下结论：

- 1. 长期均值 μ 主宰自相关性 ρ —— 比起缥缈的自相关性，低频分量 μ 才是真正的发展趋势（想想股市的大牛市）；**
- 2. 收益率均值 μ 和收益率标准差 σ 的数量级之差影响趋势效果 —— 这意味着合理选择计算趋势的时间尺度至关重要。**

通常来说，对于日频数据，收益率标准差比均值高两个数量级，在这个级别上趋势策略难有作为。如果把时间拉长，在更低的频率（即更长的时间尺度）上计算收益率，可以有效提高均值并降低其标准差，当均值和标准差相差一个数量级时，趋势策略或大有可为。这就是为什么同一组趋势参数在 A 股好使（A 股大牛市的时候日收益率均值太高）却在美股上不好使（例如，在过去 3000 个交易日中，SPX 的日频收益率标准差是其均值的近 40 倍，而其月频收益率的标准差仅是其均值的 7 倍）。

以上就是关于时间尺度的讨论。一旦确定了计算趋势的时间尺度，下一步就是选择合适的指标了，比如是使用价格均线还是收益率均值来衡量趋势强弱？是使用简单的移动平均还是对近期数据权重更高的指数平均？这都属于 trend measure 要回答的问题。而关于构建趋势策略的另一方面 position sizing，刀疤连在[之前的文章](#)中有过详细说明，本文不再赘述。

03

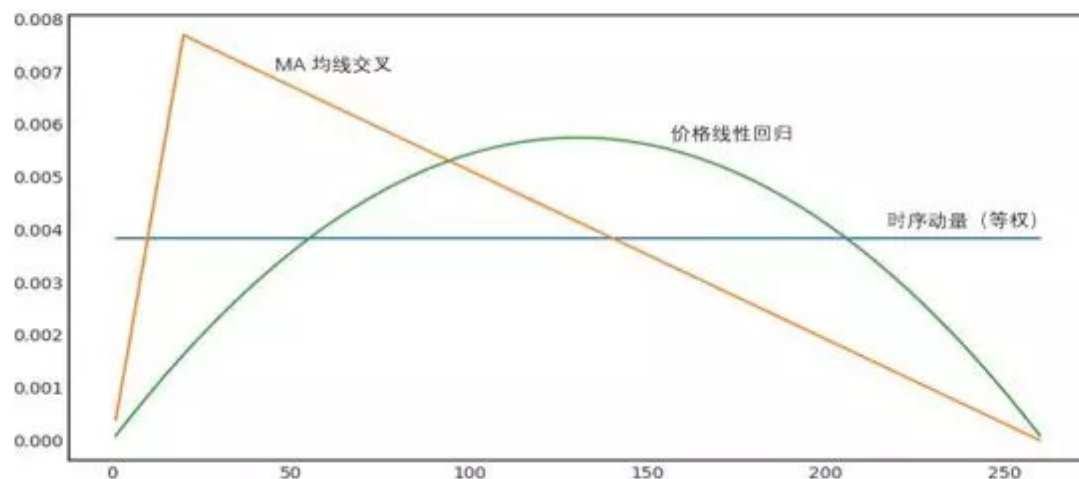
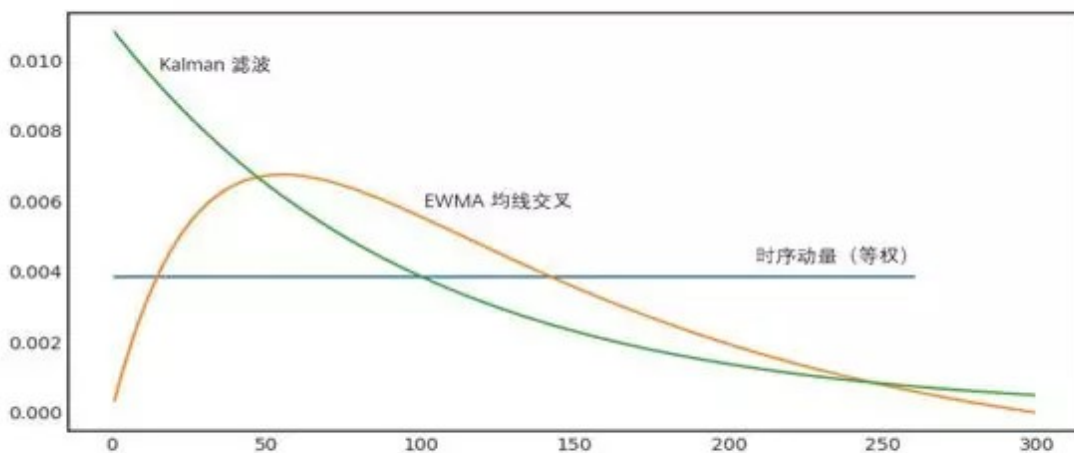
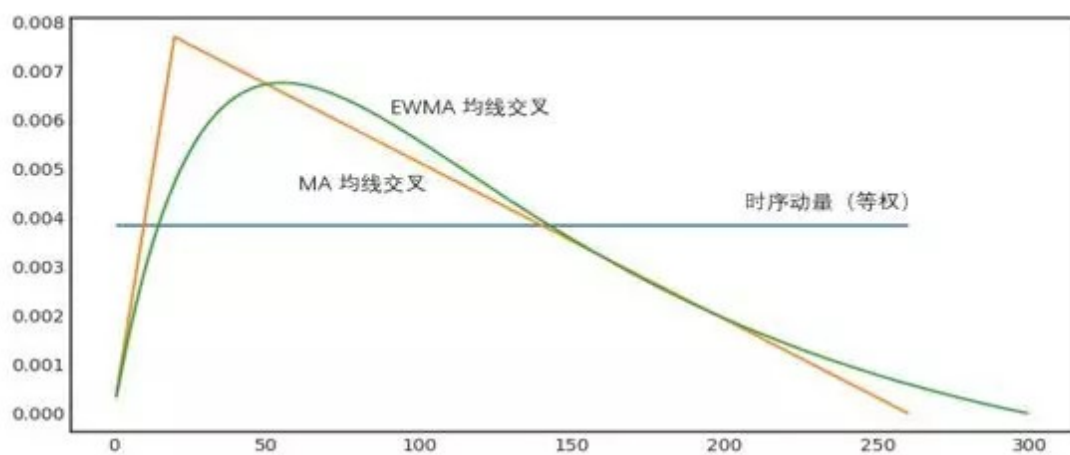
哪个趋势指标更好使？

关于不同 trend measures 的比较，Levine and Pedersen (2016) 有过精彩的讨论，《[有没有哪个趋势指标更好使？](#)》一文也做过解读。结论是 —— 先泼盆冷水吧 —— 并没有哪种 measure 吊打其他的。

在技术层面，各种捕捉趋势的方法也是层出不穷：基于收益率的时序动量、来自技术分析的均线交叉或通道突破、各种 state space 模型比如 Kalman 滤波、基于价格的线性回归、甚至是频域分析。当某种方法的回测结果不是那么给力的时候，人们的第一反应总是寻找下一个更复杂的方法。然而究其核心，上述方法都是基于价格或者收益率的时间序列计算出来的。

Levine and Pedersen (2016) 通过使用 trend signature plots (趋势签名图) 这个概念比较了不同的 trend measures 方法，认为不同的方法并没有本质区别。

以时序收益率均值，技术分析中的长、短均线交叉，Kalman 滤波，以及价格对时间的时序回归这些常见的 trend measures 为例，它们的 trend signature plots 如下图所示。在给定的计算趋势的时间尺度下，这些方法的 trend signature plots 显然没有它们的名字听上去差异那么大。**它们虽然形态有些差异，但其本质仍然都是收益率的某种加权平均。**



出处：川总写量化

除了从数学上证明不同 trend measures 并无实质差异外，Levine and Pedersen (2016) 也通过一个实例进行了说明。它们采用 Moskowitz, Ooi, and Pedersen (2012) 一文中的 58 种资产为交易对象，在不同时间尺度上分别使用两种不同的 trend measures 来构建策略、考察趋势策略的效果。得到的结论和本小节以及上一节完全一致：

1. 计算趋势的时间尺度至关重要；

2. 在给定的时间尺度下，具体使用哪种 trend measure 差异并不大。

既然不同 trend measures 的差异很小，接下来我们就以最简单的移动平均为例来探究一下趋势信号背后的本质。

04

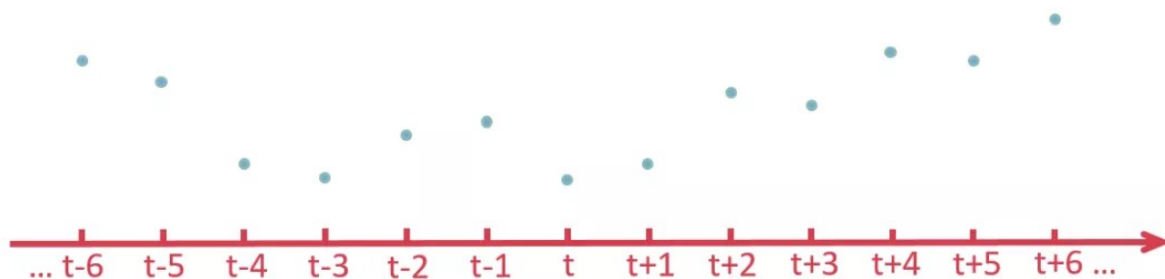
移动平均的本质

在人们的认知中，趋势追踪是一个右侧交易策略，它使用某个 trend measure 来描述过去一段时间上涨或下跌趋势的强弱，并假设该趋势会持续。而无论采取何种 trend measure，我们都无法提前预判趋势何时会结束，因此趋势策略往往在趋势发生逆转时录得较大的亏损，这也正是趋势策略的风险收益特征。在良好的风控下，这并不是太大的缺点，本文的第六节将从风险的角度进一步讨论这一点。

在本小节，我们试图回答的问题是，趋势策略为什么注定是一个右侧交易策略，为什么无法预知趋势什么时候会结束，意味着**趋势信号有明显的滞后性**。这就要从趋势信号（以移动平均为例）的本质说起。

从数学上说，移动平均的本质是一种低通滤波。它的目的是过滤掉时间序列中的高频扰动，保留有用的低频趋势。如何从时间序列中抽取出真正的低频趋势呢？下面将从数学上解释移动平均的本质，并揭示趋势信号滞后性的成因。

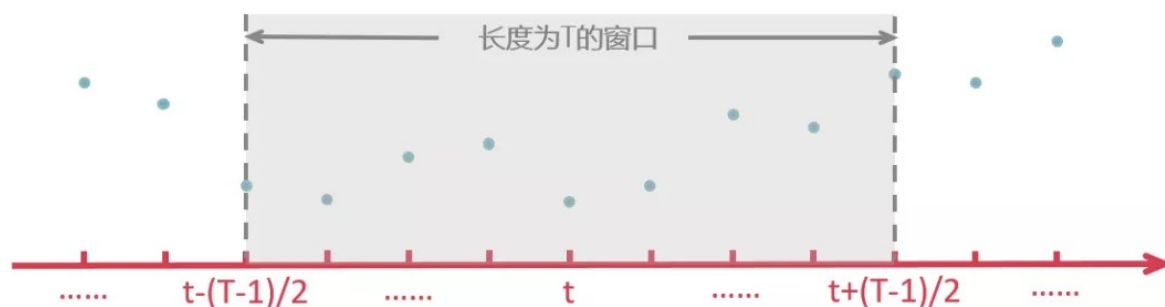
假设我们有一个时间序列 $y = \{..., y_{(t-2)}, y_{(t-1)}, y_t, y_{(t+1)}, y_{(t+2)}, ...\}$ ，如下图所示。另外，假设我们有一个作用在时域 t 上的过滤函数 F (F 的具体形式由选择的移动平均算法决定)。



在理论上，在任意 t 时刻的低频滤波（用 x_t 表示）在数学上可以表示为该时间序列 y 和过滤函数 F 在整个时域上的卷积，即

$$x_t = \sum_{i=-\infty}^{+\infty} F_i y_{t-i}$$

其中， F_i 为过滤函数 F 在时刻 i 的取值（可以理解为 y_{t-1} 的权重；比如简单平均就是等权）。在实际中通常使用长度为 T 的窗口来代替上面这个无穷级数，这意味着过滤函数 F 只在窗口长度 T 内有效、在窗口之外为 0，如下图所示：

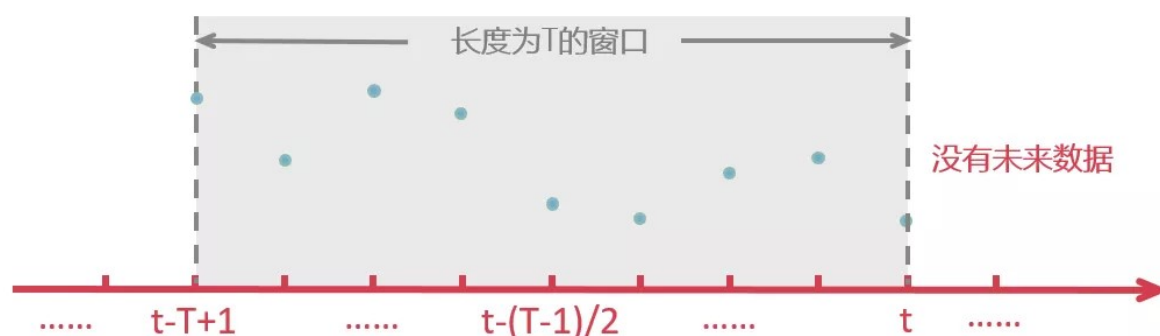


加入长度为 T 的窗函数后，在时刻 t 的低频滤波变为该时间序列 y 和过滤函数 F 在这个窗口内的卷积：

$$x_t = \sum_{i=t-(T-1)/2}^{t+(T-1)/2} F_i y_{t-i}$$

上述公式最大的问题是：在计算 t 时刻的低频分量时，利用到了未来的数据。换句话说，理论上的低通滤波（或者移动平滑）必须要用事后数据，其假设所有数据都发生后再在全球上计算所有时点的低频分量。但这在实时数据中是不可能的，因为在任何当前时刻 t ，我们都没有未

来数据可以利用。由于现实中无法使用 $t-(T-1)/2$ 到 $t+(T-1)/2$ 之间的数据，只能退而求其次将整个窗口在时间轴上向左平移 $(T-1)/2$ 个单位、使用 $t-(T-1)$ 到 t 之间的数据来计算 x_t ，如下图所示：



如此处理后，对于实时数据，在当前时刻 t 的低频滤波变为该时间序列 y 和过滤函数 F 在 $t-T+1$ 到 t 之间的卷积：

$$x_t = \sum_{i=t-T+1}^t F_i y_{t-i}$$

以 simple moving average 为例，在窗口 T 内，过滤函数在每个时点的取值都是 $1/T$ （等权）。利用上述公式计算得到的实际上是 $t-(T-1)/2$ 时刻、而非 t 时刻的低频趋势，而我们却只能把它当作 t 时刻的低频趋势使用，如此就产生了 $(T-1)/2$ 的滞后。

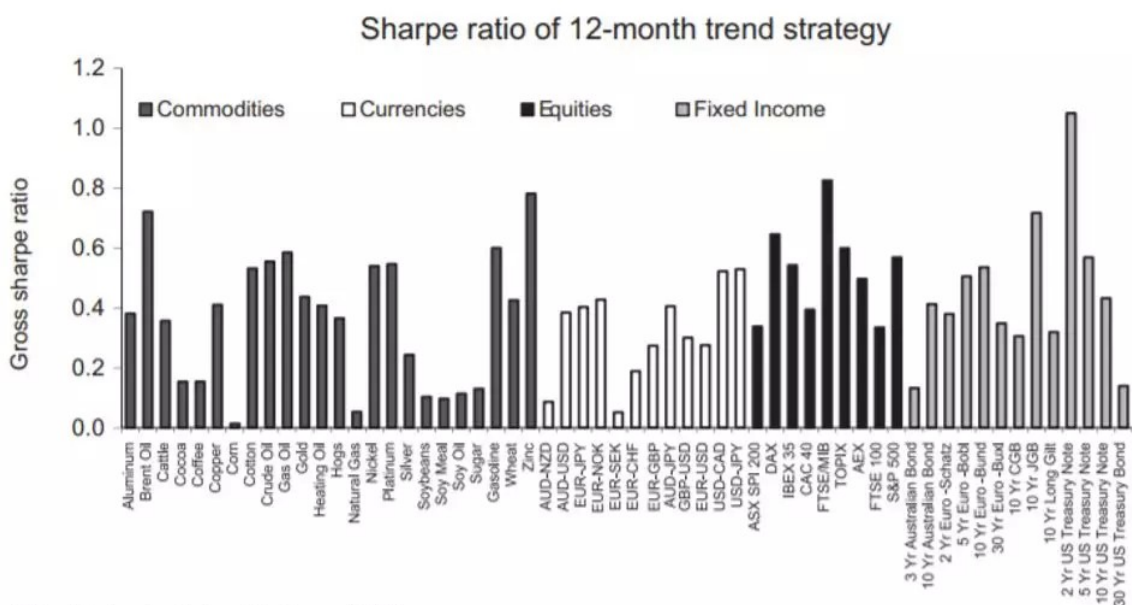
没有未来数据便是趋势信号滞后的根本原因。

从上述公式可知，计算均线的时间窗口 T 越大，滞后 $(T-1)/2$ 也就越多（比如我们看股票数据里面 MA20、MA30、MA50 等日均线，随着均线窗口 T 增大，得到的移动平滑曲线也越滞后）。正是由于滞后性和计算均线计算均线窗口的关系，人们本能上更加抵触使用长时间窗口来计算趋势。同时，人们也更加偏好使用指数加权或其他复杂的加权方法，以期更快捕捉趋势变化的原因（业界为此开发出了很多不同的移动平均算法，见《移动平均：你知道的与你不知道的》），但没有未来数据造成的滞后性使得趋势追踪注定成为右侧交易策略。

对趋势追踪表现的正确预期

要想成功的使用趋势追踪策略，除了对其各个方面有深入的理解外，还要对它的表现有合理的预期，这就是本小节的话题。

在关于趋势策略（时序动量）的研究中，最重要的一篇文章当属 Moskowitz, Ooi, and Pedersen (2012) 这篇发表于顶刊 Journal of Financial Economics 的题为 Time series momentum 的文章。下图正来自这篇文章。它展示了对不同大类的资产使用 12-month trend measure 进行趋势追踪策略的年化夏普率。虽然趋势追踪策略在所有资产均获得正的夏普率，但大部分年化夏普率都在 0.4 以下。以当前市场上资金的主流（yet 非理性）看法来看，这样的表现毫无吸引力，但这个范畴内的夏普率恰恰就是趋势策略在单一资产上的代表性表现。

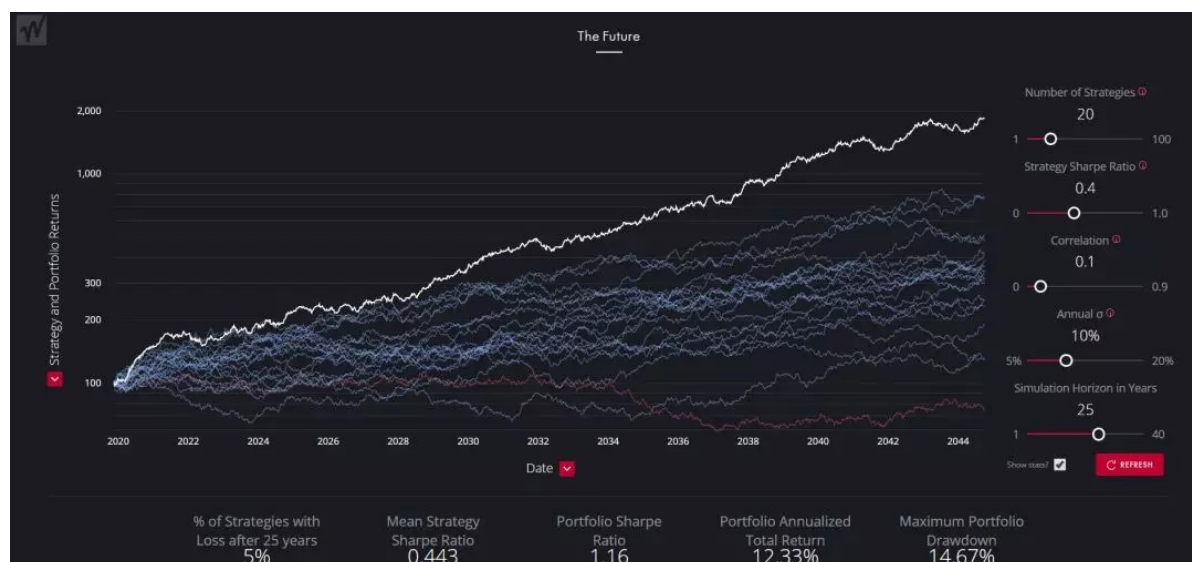


出处：Moskowitz, Ooi, and Pedersen (2012)

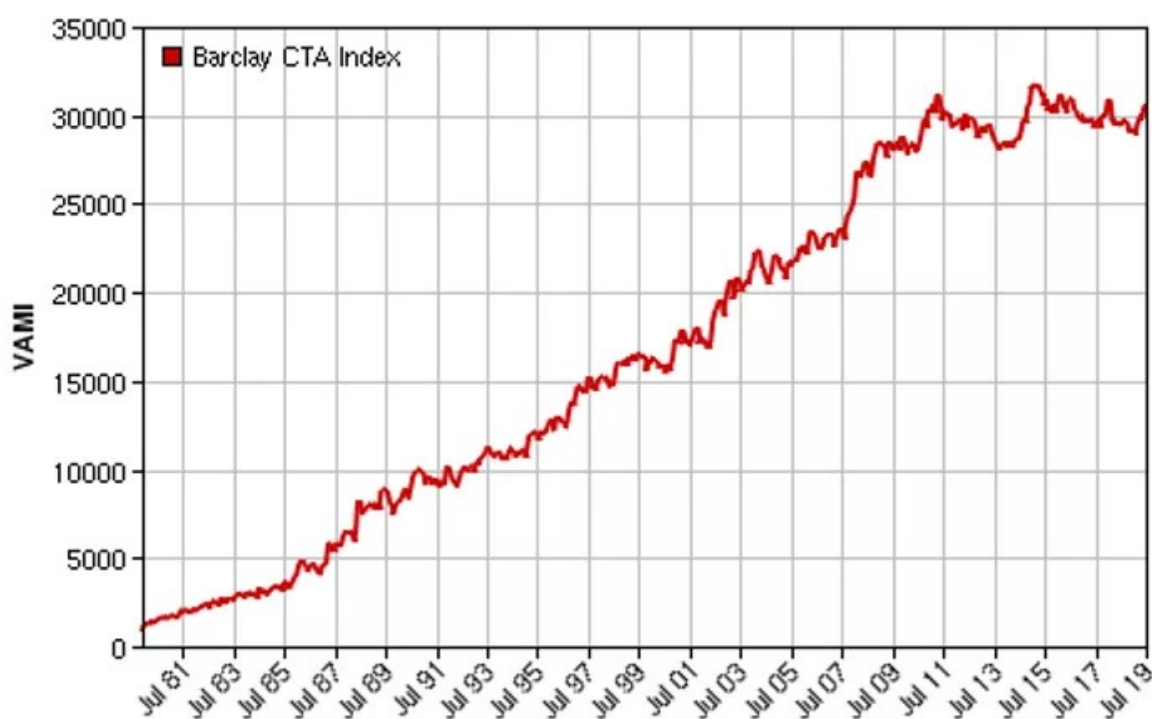
虽然趋势追踪在单一策略上提高空间有限，但好处是不同类资产的相关性较低，因此我们可以把它们趋势策略放在一起，既实现了分散化，又能提升风险收益特征。这便给了人们无限的想象空间。下面我们就来模拟一下。

在模拟中，我使用了元盛开发的 The Future 工具 (<https://tools.winton.com/thefuture/>)，它以单一策略的夏普率、相关性等参数作为输入，非常方便的可视化出这些单一策略组合在一起之后的效果。假设我们有 20 个单一趋势追踪策略，每个策略的夏普率为 0.4，策略之间的相关性为 0.1，则可以得到以下的仿真结果：

portfolio 的夏普率提升到了 1.16，显著高于单一策略的 0.4。当然，这仅仅是一次仿真；复合策略夏普率的分布可以通过数学推导、在合理的假设下得到，或通过更多的仿真得到。



上面这个小实验表明众多品种的趋势策略放在一起后的综合效果确实能够显著提升风险收益特征。然而我们也必须认识到，多品种的趋势追踪策略仍然可能在短到中期（1 - 3年）的范围内持续亏损。下图是大名鼎鼎的 Barclay CTA Index（考虑到趋势追踪占 CTA 中很高的比重，因此以该 Index 的表现代表趋势策略在历史上的表现），它在 2011 年 7 月之后的表现就有些惨不忍睹，从 2015 年开始几乎在持续的回撤。



出处：BarclayHedge

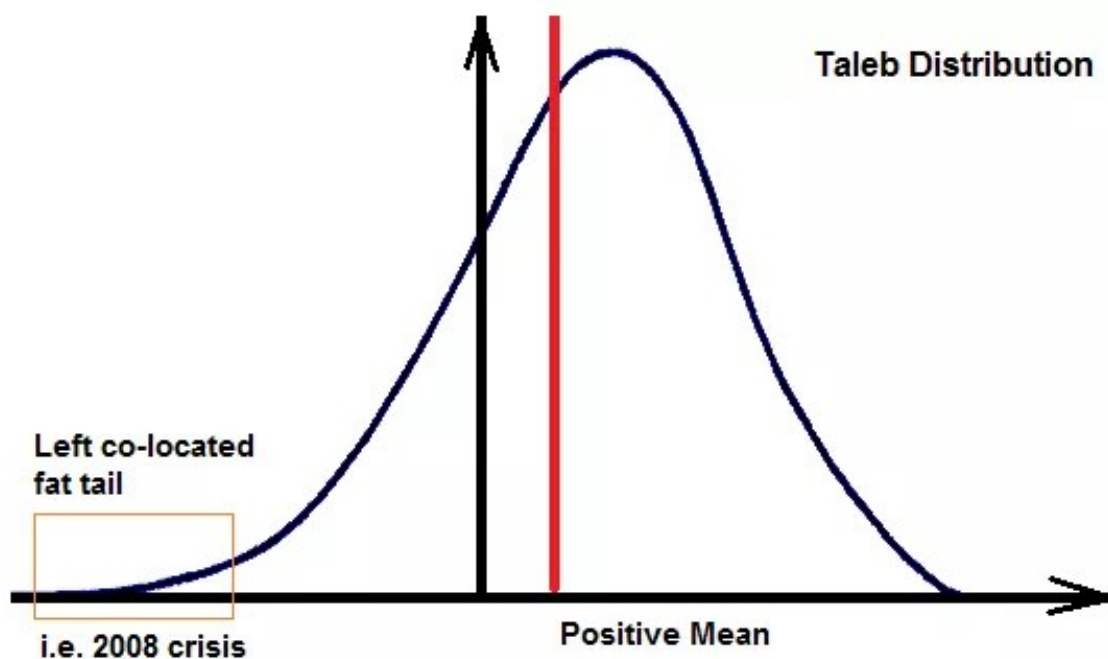
趋势追踪策略并不是圣杯；根据市场所处的状态，它有好使的时候，更有持续亏损的时候。长期来看，趋势追踪策略可以获得很好的收益、值得配置；**但真正能配置以及驾驭它的人，一定是对短到中期内它的表现能有多差有非常清醒的认知。**

06

从风险的角度看趋势策略

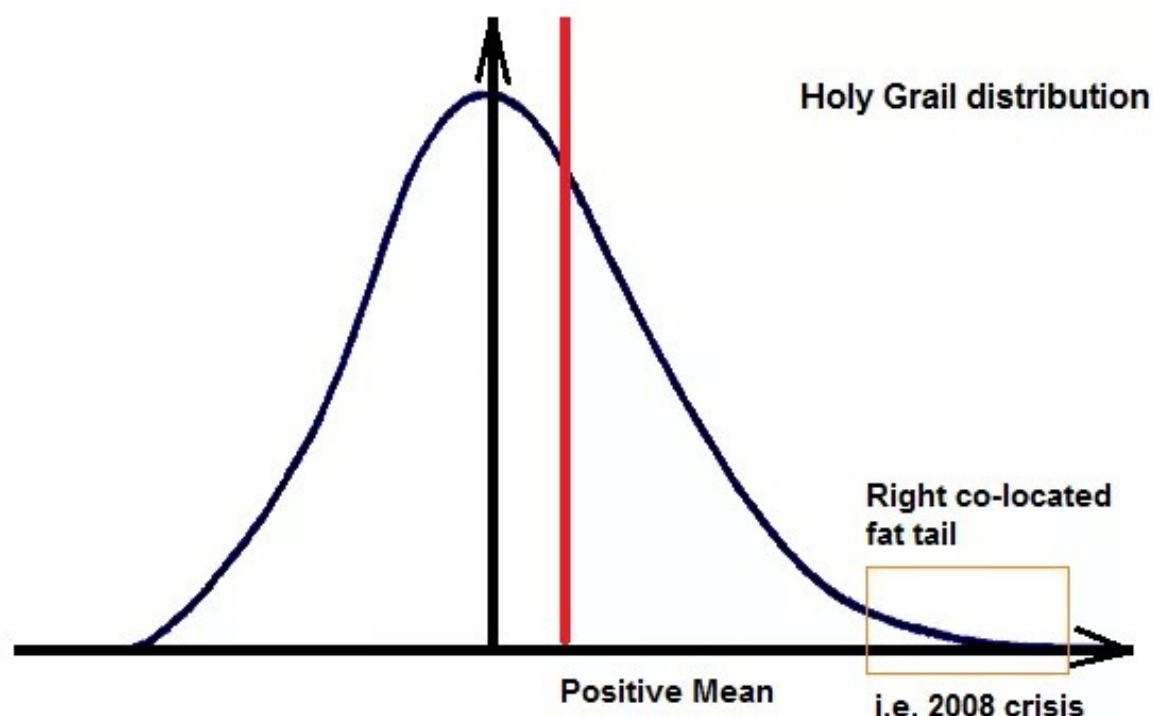
最后来从风险的角度看趋势追踪。

对于大部分常见的资产或者投资策略，它们的收益分布都有如下的形状 —— 均值为正，但是左侧存在着极端事件造成的巨大亏损，从而造成分布的负偏。这个分布通常被称为 Taleb Distribution，以 Nassim Nicholas Taleb 命名；他以《黑天鹅》一书闻名于世（黑天鹅代表了左侧的极端事件）。



而趋势追踪策略的收益分布（假设合理使用杠杆和进行风控）却和上面不同，趋势策略的收益率分布往往是 positive skewed 的。从趋势策略的 PnL 曲线来看，它在更多的时候是连续经历小幅回撤，然后换来一大波（大趋势）的上涨，从而将之前的众多小幅回撤抹平，使得该策

略的长期收益为正。这个“亏损有限、收益无限”的特征造就了趋势策略收益分布的正偏。这个分布又被称作圣杯分布（Holy Grail distribution）。趋势追踪策略虽然不是圣杯，但它的收益分布满足圣杯分布。



正是因为这种独特的特性，使得趋势策略能在其他资产发生左侧尾部风险的时候脱颖而出，这也就是它为什么享有 crisis alpha 的美名。关于 crisis alpha 的特性，[刀疤连](#)的文章中有过详细的描述并针对国内市场进行了实证研究。

为什么不同策略的收益率会表现出不同的分布特征？这可以从两类不同的风险来解释。任何投资策略想要赚钱，都需要承担一定的风险。而策略承担什么类型的风险就决定了它的收益分布具备何种的特征。风险可以被分为 **convergent risks** 和 **divergent risks** 两大类，因此策略也可以被视为 **convergent risk taking** 和 **divergent risk taking** 两大类。

1. **Convergent risks** 是那些被人们理解、可通过建模来测量（尽管不一定准确）的风险；
2. **Divergent risks** 是哪些未知的风险、无法预测的风险、不能很好的被度量的风险。

当构建 convergent risk taking 策略的时候，人们根据金融学或者经济学原理对资产的走势有一个先验信仰，并以此为判断做出投资决策。另一方面，在构建 divergent risk taking 策略时，人们假设资产未来的走势是未知的、不使用任何先验信仰来辅助判断资产会涨还是会跌，而是根据资产价格自身的发展做出相应的判断。

举例来说，股票市场中的价值投资正是 convergent risk taking 的代表。这类策略通过深度行业研究挖掘上市公司的内在价值，当价格低于价值时则买入，耐心的等待其价格向价值回归、无视短期的波动甚至是持续下跌，这是这类策略愿意且主动承担的风险。而常见于 CTA 中的趋势追踪策略则是 divergent risk taking 的代表。这类策略对资产的内在价值不做任何判断，而是严格依据价格的走势进行交易：价格上涨了做多、价格下跌时做空；当盈利时会逐渐加仓，当亏损时会及时清仓。正因如此，趋势追踪策略属于 divergent risk taking 策略，因而其收益满足圣杯分布的特性，这也正是趋势追踪策略能够长盛不衰的原因。

有必要指出的是，圣杯分布是非常美好的，而获得美好的东西注定是十分困难的。趋势追踪策略最困难的地方在于在趋势中拿的住单子，让利润奔跑（只有这样才能实现收益率分布的右尾）。然而，由于认知偏差，人们往往对确定性高的低收益过度偏爱，因而总是倾向早早平掉盈利的单子。唯有克服人性的错误，才有可能真正享受到趋势策略带来的圣杯分布。

07

结语

来回顾一下本文对趋势追踪策略做了哪些思考。

从技术层面来看，一个成功的趋势策略，最重要的是计算趋势的时间尺度，策略的参数应适应资产的收益率特征。当确定了时间尺度后，具体使用哪个 trend measure 并无本质区别，把更多的精力放在 position sizing 来控制风险可能是更好的选择。作为右侧交易策略，使用历史数据计算得到的趋势信号注定有一定的滞后性，这是难以避免的，使用诸如指数平滑来给最新的数据更高的权重是业界对于减少滞后性的一种尝试。

从心理层面来看，想要成功使用趋势追踪策略就必须要对它的风险收益特征有清醒的认识。单资产上的趋势策略的夏普率通常不高，但是把不同低相关资产的趋势策略组合在一起能显著提升 portfolio 的夏普率。即便如此，趋势策略在短到中期内（例如 1 年）也非常有可能出现持

续的回撤。驾驭趋势策略不在于理解它长期来看的正期望，而在于欣然接受它短期内潜在的很差的表现。

从风险角度来看，趋势策略属于 divergent risk taking 策略。它虽然不是圣杯，但它的收益率分布满足圣杯分布，这使得它具备 crisis alpha 等优点，极具配置价值。不过值得一提的是，有研究表明，过度使用杠杆对这类策略的收益和风险的放大是非线性的。因此，科学、有效的风控是使用趋势策略的重要一环。

以上。

好了，我要接着去读最近刚入手的《贫穷的本质》了。各位再见。

参考文献

Levine, A. and L. H. Pedersen (2016). Which trend is your friend? *Financial Analysts Journal*, Vol. 72(3), 51 – 66.

Moskowitz, T., Y. H. Ooi, and L. H. Pedersen (2012). Time Series Momentum. *Journal of Financial Economics*, Vol. 104(2), 228 – 250.

Roncalli, T. (2017). Keep up the momentum. SSRN: <https://ssrn.com/abstract=3083921>.

Winton (2013). Historical Performance of Trend Following. Winton Capital Research Note.

免责声明：文章内容不可视为投资意见。市场有风险，入市需谨慎。

分享量化理论和实证心得



免责声明：

文章内容不可视为投资意见。资本市场有风险，入市投资需谨慎。

文章已于2019-11-08修改