

漫谈CTA及投资观

王生 2014年 8月

CTA概览

CTA简介

- CTA一词起源于美国，全称为商品交易顾问（Commodity Trading Advisor），是侧重于期货行业的投资经理人。这一称呼原来专指美国农产品期货交易顾问，但是随着其他种类期货的增多，尤其20世纪70年代金融期货的大规模发展，使得CTA业务范围扩大。
- 根据投资方向的不同CTA基金可以分为多元化投资基金和专业化投资基金；根据交易策略的不同。分为系统化投资模式和非系统化投资模式。根据投资风格的不同。趋势交易型和非趋势交易型两类。

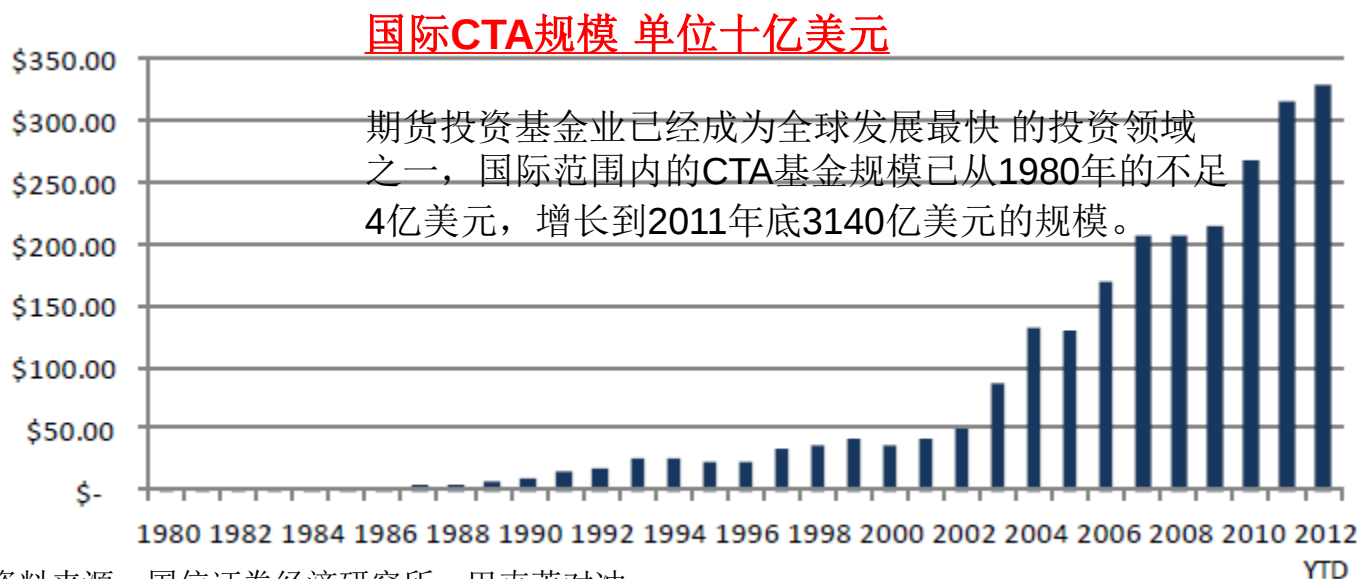
CTA Industry - Assets Under Management

CTA分类

Assets Under Management	1st Qtr 12	4th Qtr 11	3rd Qtr 11
<u>Managed Futures</u>	\$328.3B	\$314.3B	\$320.3B
Sectors			
<u>Agricultural Traders</u>	\$1.15B	\$1.09B	\$1.38B
<u>Currency Traders</u>	\$26.11B	\$27.28B	\$28.09B
<u>Diversified Traders</u>	\$181.87B	\$181.91B	\$186.59B
<u>Financial/Metal Traders</u>	\$80.11B	\$81.64B	\$86.84B
<u>Discretionary Traders</u>	\$26.02B	\$26.98B	\$27.57B
<u>Systematic Traders</u>	\$260.24B	\$259.88B	\$269.33B

CTA特点及其投资规模

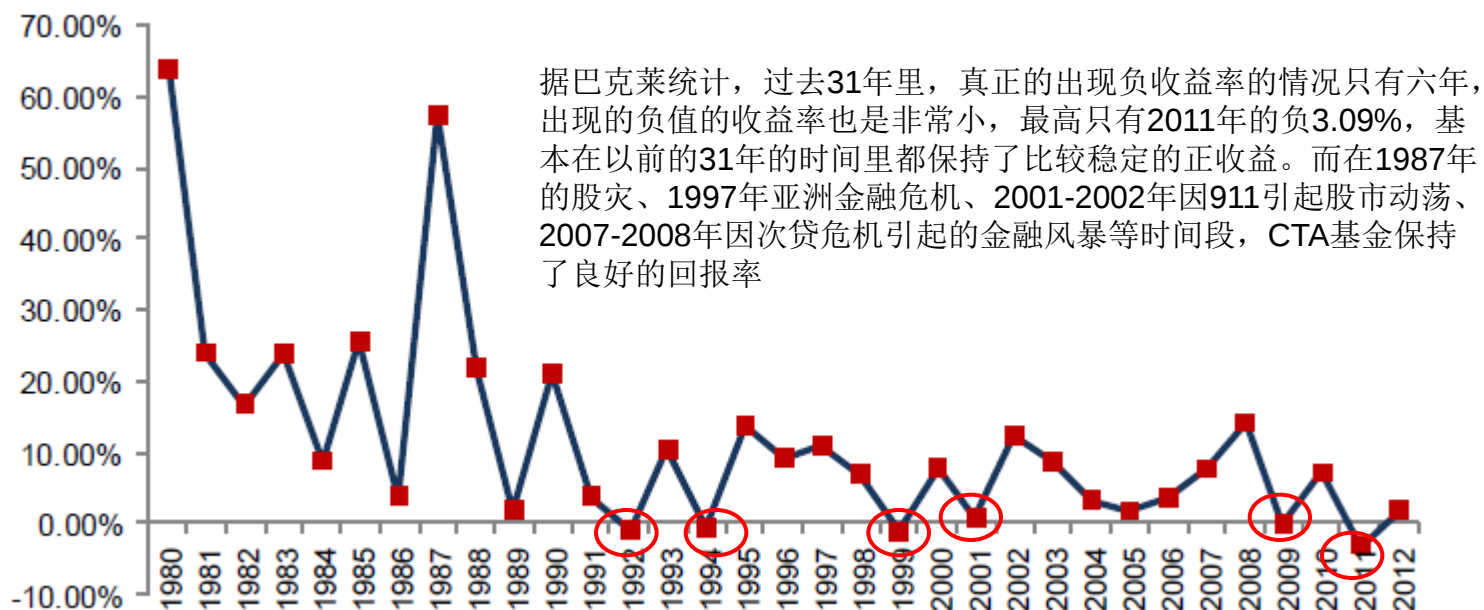
- 从美国市场的经验看，CTA与期货市场上的个人投资者相比具有三大特点：第一，CTA都具有明确的、长期的投资策略，以及具有资金管理和风险控制体系。第二，CTA会持续不断地监控市场，分析市场的价格变化所带来的风险与机会，从而在相对恰当的时机调整风险头寸。第三，CTA有着严格的交易纪律，可以避免投资者心理波动对交易策略效果的影响



资料来源：国信证券经济研究所，巴克莱对冲

CTA重要的市场功能

- CTA由于目标市场具有双向交易机制和高倍杠杆作用，呈现出与股票和债券组合的低相关性，并且具备丰富的策略设计、组合和运用空间，且不易受市场波动影响。尤其当股票市场处于熊市时，在资产组合中加入CTA基金，可以使防御风险，还能提高收益。



资料来源：国信证券经济研究所，巴克莱对冲

CTA策略讨论

P-Quant or Q-Quant?

- Q是指风险中性测度。风险中性的意思主要是说历史数据不能帮助你预测未来的走势，所以决策是没有风险补偿。模型可以给出漂亮的数学性质，而且可以在缺乏数据的情况下得到一些结论。涉及的数学技术主要是随机过程，偏微分方程之类。基于经济学理论、数学理论、交叉学科理论成为主要的策略产生源泉。例如：混沌、分形、工程、生物、生态理论等。整体上 Q重模型和理论，主要投资银行衍生品定价和产品设计。
- P是指真实概率测度。模型依赖的概率分布是从历史数据上估算出来的。方法主要依赖数据，数据量越大估算的效果越好。初期由于计算工具限制，主要是技术指标和技术系统的方法，经典的KDJ、MACD、海龟系统等。后期，伴随统计和计量经济学发展，以及计算机工具和大数据、云及并行结构计算网络的成熟，各类统计计量模型以及复杂算法产生并得以应用。主要有时间序列（ARIMA，GARCH类），Bayesian，MINE方法以及机器学习等。拼设备和速度已经白热化。P重数据和验证，是买方（Hedge Fund类）所需要的技术，他们需要针对大量证券做出筛选和投资决策，业务是数据驱动的。
- PQ本是一家。其实模型和数据的重视是相对的。目前在P领域也会大量借鉴理论模型，Q领域也得益于大数据对模型和理论创新的帮助。

必须认识一个大的背景：中国量化投资相当落后，但成熟速度相当快

- 主要原因是：中国金融市场的历史短，功能欠缺。使得市场主体没有必要进行过度的量化交易。在这种历史阶段，量化交易反而成了成本高昂意义不大的事情。相当于一个拖拉机不需要飞机的引擎。
- 但是中国市场的成熟度从两个维度迅速在提高。第一就是制度层面，中央的金融体制改革，已经使得一个大资管行业的格局十分清晰。未来各个证券金融产品可以自由交易，资管行业也完全自由竞争。交易产品线和工具利益丰富，体现在期权推出等方面。第二就是交易主体层面。传统的私募投资机构已经开始转量化投资。同时，多年大陆金融教育和海龟人员也带来了许多量化投资技术，因此大陆的量化投资已经迎来的一个迅速发展期。
- 结合到前文我们划分的大类来看，中国的量化投资还基本处于Pquant中的初级阶段。据笔者了解大多数的投资基金的量化模型还是基于技术和统计分析，甚至计量统计分析用的程度都相当有限。——本质上还是因为市场不成熟的原因。

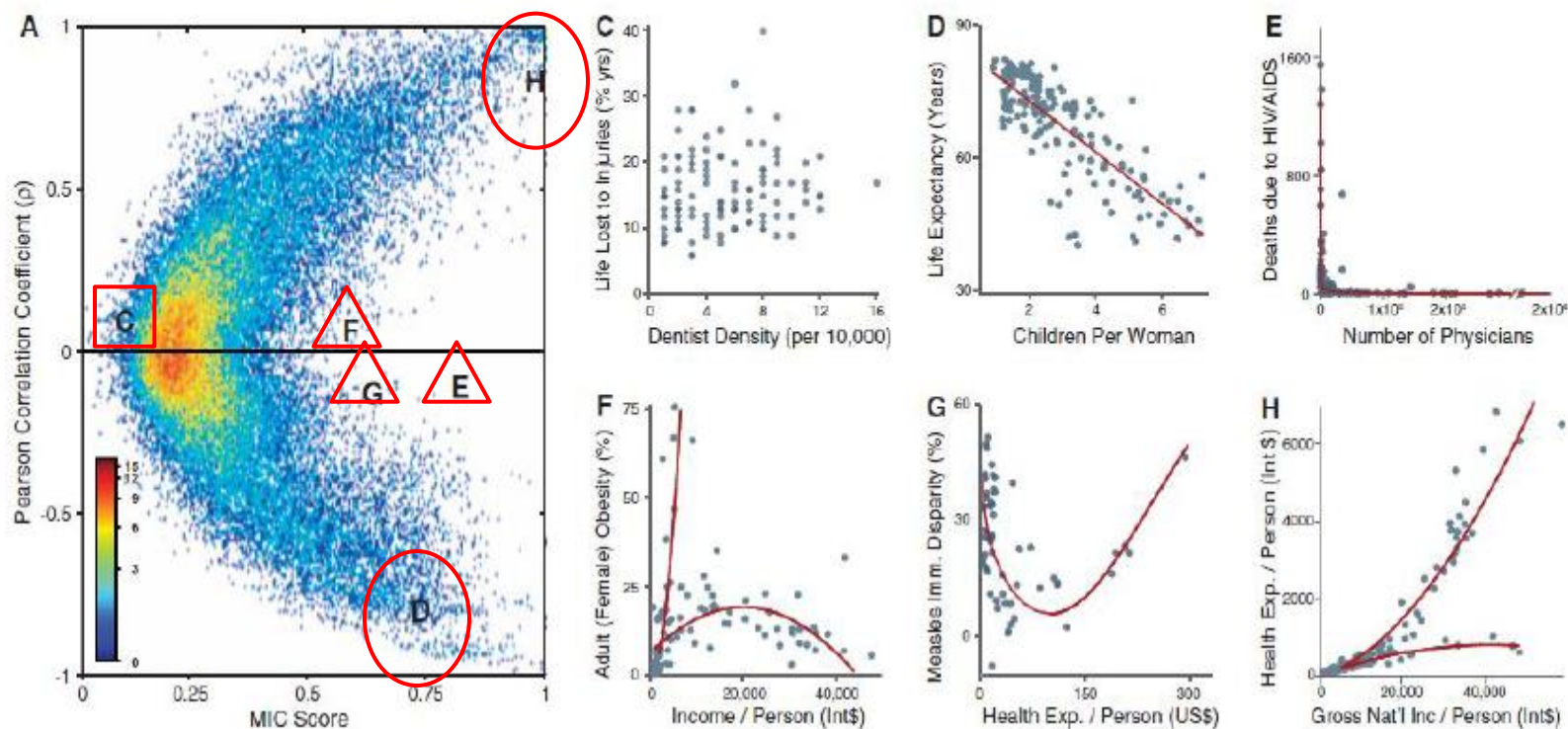
必须认识一个大的背景：巨大的跨市场差异

- 不成熟的市场短期定价失效，短线和高频均可盈利。长来看，短期定价能力迅速上升，长期定价能力回复较慢，市场会转向中线趋势策略为主。这个角度挣钱从短期向长期过渡。
- 交易者机构化使得简单的策略迅速失效，跨学科或具有独特思想的策略中长期才能生存，核心是持续开发能力。整个角度从赚散户的钱过渡到机构博弈。

Exp	恒生指数	CAC - 40		D - Mark		IBM	Dow	
MA	1988~	1983~	1988~	1983~	1988~	1982~	1983~	1988~
Days	1993	1987	1992	1987	1992	1987	1987	1992
5	22.3	40.5	16.6	5.7	1.3	18.0	-47.5	-36.2
15	17.5	33.1	20.7	9.7	4.9	4.0	-22.3	-26.8
25	12.1	32.8	19.9	6.7	7.4	4.3	-22.7	-16.1
35	9.8	30.0	19.7	9.3	6.2	6.9	-18.9	-14.9
45	8.7	29.1	10.8	8.5	9.3	8.6	-12.9	-14.4
55	11.7	28.5	13.5	9.3	6.3	11.7	-0.7	-8.7
65	11.1	27.7	18.3	11.7	5.8	11.1	0.8	-6.6
75	18.0	27.8	13.5	11.2	5.8	18.0	-0.7	-11.6
Avg	10.4	31.2	16.6	9.0	6.2	10.3	-15.6	-16.9

案例分享

- 为什么一定是线性？或某种非线性模型？
- 过分依赖于分析者的经验
- 对于非连续的离散数据难以处理



案例分享

- 《Science》上的文章《Detecting Novel Associations in Large Data Sets》
 - 方法概要：用网格判断数据的集中程度，集中程度意味着是否有关联关系
 - 方法具有一般性，即无论数据是怎样分布的，不限于特定的关联函数类型，此判断方法都是有效
-
- MIC: the Maximal Information Coefficient
 - MINE: Maximal Information-based Nonparametric Exploration

案例分享

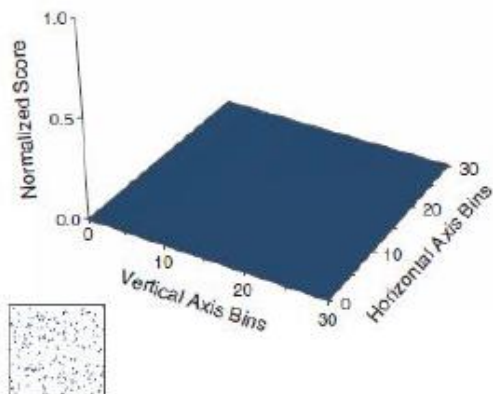
- 坐标平面被划分为 (x,y) 网格 G （未必等宽），其中 $xy < n^{0.6}$
- 在 G 上可以诱导出“自然概率密度函数” $p(x,y)$ ，任何一个方格（box）内的概率密度函数值为这个方格所包含的样本点数量占全体样本点的比例
- 计算网格划分 G 下的 **mutual information值** I_G

$$I(X; Y) = \int_Y \int_X p(x, y) \log \left(\frac{p(x, y)}{p(x) p(y)} \right) dx dy.$$

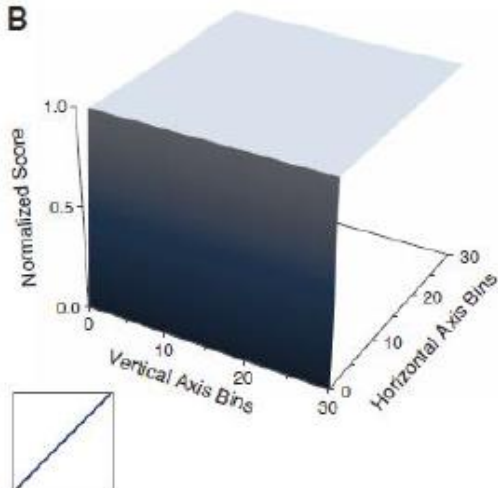
- 构造**特征矩阵** $\{m_{xy}\}$ ，矩阵的元素
 $m_{xy} = \max\{I_G\} / \log \min\{x, y\}$ 。max取遍
所有可能的 (x,y) 网格 G
- $MIC = \max\{m_{xy}\}$ 。Max取遍所有可能的
 (x,y) 对

案例分享

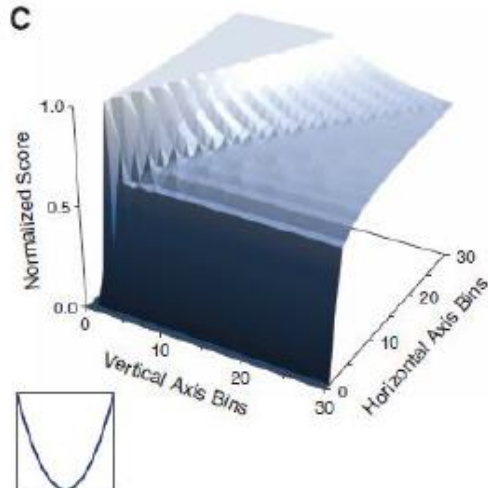
A



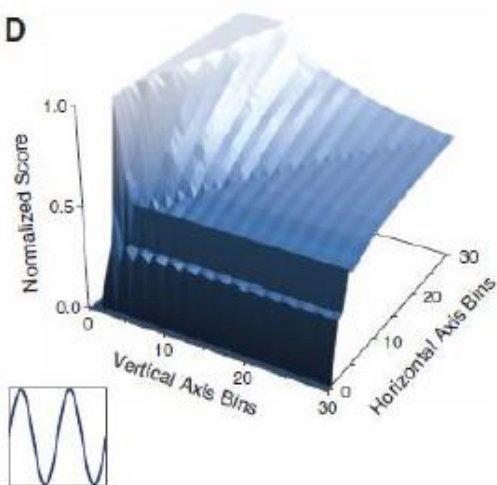
B



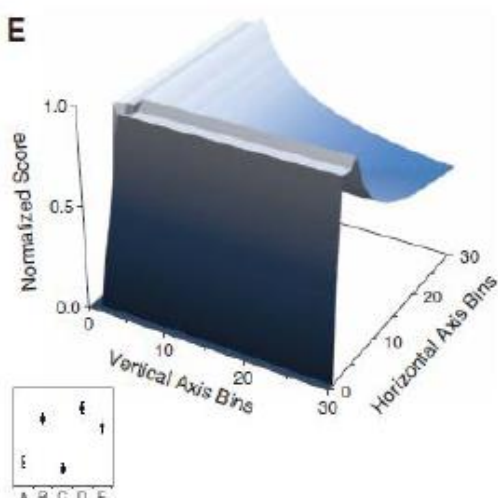
C



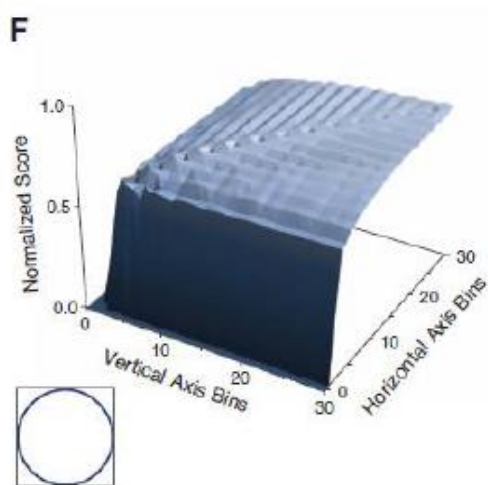
D



E



F



结束，谢谢！

- 上海通联期货有限公司
- 陆家嘴西路99号7楼
- 通联期货研发部 联系电话68865935 传真68866958