量化策略研究及应用

同花顺量化实验室负责人-陈城

目 录

CONTENTS

01 量化投资概述

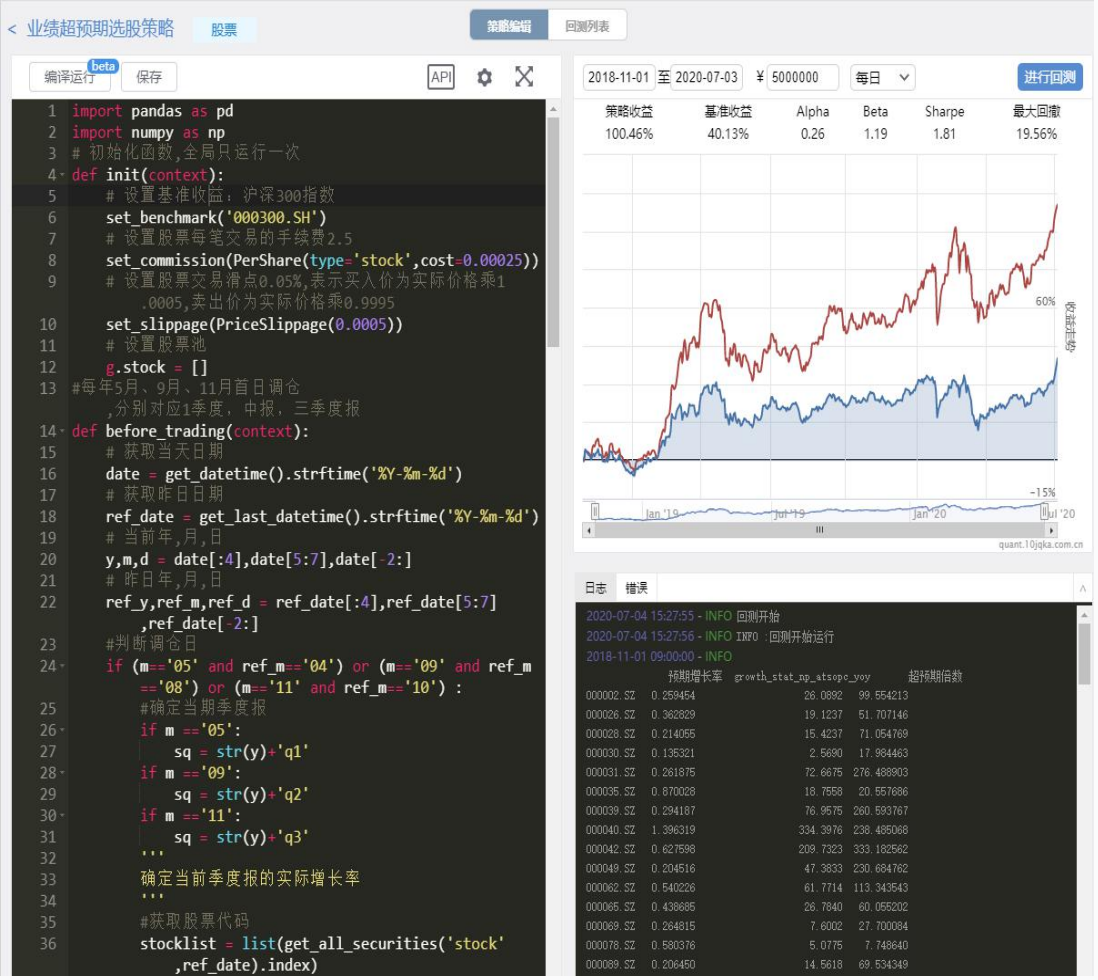
02 股票数据

03 量化选股策略

04 量化技术面择时策略

01 量化投资概述

（1）量化投资的概念



量化策略概念：

通过**数量化方法**作出投资决策，使用计算机技术完成交易指令，力求获得稳定收益的交易方式

ps：数量化方法：定量地对评价对象的各个环节中每一影响因素进行分析，研究各因素之间的数量关系、数量变化和规律

【案例】

投资决策：全历史数据统计，沪深300指数触发分钟涨幅大于等于0.5%事件，次分钟平均涨幅0.05%，交易手续费0.02%，预期盈利0.03%

交易指令：使用计算机技术监控该事件，当事件触发时，买入沪深300ETF，次分钟收盘卖出沪深300ETF。

ps：账户已持仓足量的沪深300ETF份额

（2）量化投资流程

基于完整、可靠、规范的数据，通过**数量化方法**研究收益(alpha)、风控、成本等模型，并使用计算机技术完成交易执行，包括资管管理、风控管理、订单管理等，最终获得投资收益。

（3）量化投资的应用

1 定量分析

运用计算机技术(python、C、R编程语言)对历史数据进行处理、分析、可视化，得出定量的结论。

【案例】

MACD指标是否可作为投资参考?

北上资金流入的股票，后市上涨概率会更大吗?

2 市场监测

运用计算机程序，实时对整个市场进行监控，包括个股价格波动、市场消息、突发事件等。

【案例】

全市场近4000只股票，如何监控刚涨跌停的股票?

如何监控分钟涨幅大于5%的股票?

3 交易执行

运用计算机程序，完成精确的、及时的交易工作。

【案例】

100个股票账户都需要买入50只股票，该如何操作?

某股票账户需要买入平安银行2个亿，如何完成这笔交易?

（4）量化投资的优势

1 **纪律性** 量化投资严格执行投资策略，不会受到投资者情绪的干扰

2 **及时性** 快速跟踪市场变化，不断发现提供超额收益的新模型，寻找新的交易机会

3 **准确性** 客观评价并捕捉交易机会，克服因人的情绪、懒惰而产生行为偏差

4 **分散化** 控制风险的前提下，研究可覆盖所有资产，并构建大概率取胜的组合

5 **系统性**
多层次的量化模型（大类资产配置、选股、择时、统计套利等）
多维度的分析角度（宏观周期、市场结构、估值、成长、盈利质量、市场情绪等）
海量的金融数据（基本面、技术面、高频、舆情等）

（5）量化投资的陷阱

未来函数

时间靠后的数据对靠前的结果产生影响

例如：幸存者偏差

讲故事

事先确定结论，寻找数据验证

例如：片面分析

数据挖掘

强行在海量数据中找出逻辑

例如：最优参数

异常值

剔除异常值时，同时剔除有用信息

例如：确定上下限

对称性

多空收益特征存在对称性

例如：动量因子

02 股票数据

（1）股票数据分类

1 行情数据

市场中所有交易者的交易行为数据集合

2 资金数据

市场中所有交易者的成交数据集合

3 财务数据

上市公司经营过程中公布的财务报表数据集合

4 事件数据（舆情）

上市公司经营过程中公开事件数据集合

5 行业数据

上市公司所属行业分类的数据集合

6 其他数据

其他可分析数据，比如同花顺用户行为数据（App）、投顾持仓账户、基金数据等等

（1）行情数据

行情数据是市场中所有交易者的交易行为数据集合，例如：成交量、成交金额、开高低收价，换手率，委托量，委托价等

运用数学公式及交易理论，将行情数据演化成技术面数据，用于指导选股与交易

技术面：2020年5月后，东方铁塔股价持续创60日新高，均线呈现多头排列，表明买方力量大于卖方，投资者看多股价。

预测：股价或将持续上涨

（2）资金数据

资金数据是市场中所有交易者的成交数据集合，例如：主力流入、主力流出、主力净流入等

运用数学公式，将资金数据演化成资金因子，结合技术面用于指导选股与交易

资金面：2020-08-07，大连重工主力流入率1952.2万元，其大单净比2.8%，在全市场排407名。

预测：当日资金流入，股价短期处于强势上攻阶段。

（3）财务数据

财务数据是上市公司经营过程中公布的财务报表数据集合，例如：总资产、总负责、所有者权益、净利润等

基于财务分析理论，运用数学公式，将财务报表数据演化成基本面数据，用于指导选股与交易

财务面：2020-04-29，东方铁塔业绩披露：2020年一季报净利润4541.86万元，同比去年增长101.11%。

预测：公司净利润或持续增多

（4）事件数据（舆情）

事件数据泛指上市公司经营过程中公开事件数据集合，例如：项目中标、资产重组、新闻舆论、舆论事件等

基于事件驱动理论，运用非结构化数据分析方法，将舆情数据演化成舆情因子，用于指导选股与交易

事件：2020年7月，东方铁塔披露两项中标信息事件，预计利好当年净利润

预测：公司净利润或持续增多

（5）行业数据

行业数据是上市公司所属行业分类的数据集合，例如：食品饮料、银行、非银金融等

基于行业动量、反转、轮动等理论，运用数学公式，**在行情、资金、财务数据的基础上，细化行业分析**，用于指导行业选股与交易

行业：食品饮料、医药生物、家用电器属于非周期行业，适合长期投资

预测：食品饮料行业中，更容易诞生大牛股

（6）其他数据

其他数据

- 同花顺（东财、雪球、大智慧等）用户行为
- 分析师预测、公私募基金、券商研报观点
- 期货期权衍生品、可转债、指数等品种数据
- 龙虎榜，陆港通……

03 量化选股策略

量化选股分类

01 技术面选股

02 基本面选股

03 资金面选股

04 行业选股

05 组合优化

量化技术面选股策略

量化技术面选股策略，**主要是选定初始股票池，再采用单个或多个技术指标筛选出若干只股票进行数量化组合交易。**

通常情况下，技术面选股有效性时长较短，需要时刻关注市场变化，包括政策变化，投资者结构变化，事件等因素，根据市场变化及时调整。

技术指标选股在业内通常采用相对选股法。比如，均线技术指标的相对法选强势股，**首先计算初始股票池内股票的均线与现价，再计算出现价偏离度，然后按偏离度排序，接着挑选首或尾部的若干只股票，最后组合交易。**相对法选股的优势，主要是从单个股票角度衍生到群体股票的角度，从概率上取得优势，另外，该方法可以确保选出股票，避免大涨行情出现空仓的情况。

另外，**业内一般选择创业板、中证500、科创板等指数成份股作为初始股票池**，因为这些指数内的股票散户参与度比较高，更符合技术指标的特征。

量化MACD选股(正向)-创业板指数成份股

MACD选股：从创业板指数成份股中，采用MACD(12,26,9)的MACD相对值(现价的偏离度)，**选取前30只股票**作为股票组合，每周更新持股。

效果简评：**与整体市场环境保持一致，但无法获得超额收益率。**

量化MACD选股(反向)-创业板指数成份股

MACD选股：从创业板指数成份股中，采用MACD(12,26,9)的MACD相对值(现价的偏离度)，**选取后30只股票**作为股票组合，每周更新持股。

效果简评：**与整体市场环境保持一致，获得20%+超额收益率，比正向选股的效果更优。**

技术选股策略分析

1 正反向法

按技术指标排序后，通过构建正反向两个组别，准确了解MACD指标反向组收益更好，表明MACD值偏离度越小的股票，后续收益率比偏离度大的股票。可能原因是MACD相对值偏离度较大时，股价分歧越大，影响整体收益率。

2 超额收益率

通常情况下，选股策略长期处于满仓运行状态，我们采用超额收益率及衍生指标评价选股策略，超额收益率越高，说明选股有效性越强。

3 市场一致性

选股策略的收益率与市场整体收益率的相关性较高，在市场行情上涨时，选股策略获得更多的收益率。当市场行情下跌时，选股策略损失更少的收益率，长期坚持，获取较大的超额收益率。

量化基本面选股策略

量化基本面选股策略，**主要是选定初始股票池，再采用多个财务指标筛选出若干只股票进行数量化组合交易。**

一般而言，基本面股票组合有效性较长，短期内可能持续亏损，但长期往往能取得较好的收益。财务指标选股既可以采用相对选股法，也可以采用绝对值法。另外，**业内一般选择上证50指数成份股作为价值股票池**，因为这指数内的股票往往价值较高，也可以选择中证500等指数成份股作为成长股票池，因为这些指数内的股票往往成长性更好。

根据不同财务指标性质，构建的股票组合也会呈现一定特征。如果采用净利润增长率、营收增长率等财务指标，构建的投资组合更偏向成长性，业内俗称“成长组合”。如果采用股息率、市盈率等财务指标，构建的投资组合更偏向价值性，业内俗称“价值组合”。

价值组合-沪深300指数成份股

价值组合选股：从沪深300指数成份股中，按股息率指标从大到小排序，**选取前30只股票**作为股票组合，每月更新持股。

效果简评：策略Beta小于1，策略非常稳定保守，符合价值派。**长期而言，组合具备盈利能力，但是盈利能力一般。**

成长组合-沪深300指数成份股

成长组合选股：从沪深300指数成份股中，筛选出**销售毛利率大于50**的股票，并按股东权益收益率ROE指标从大到小排序，**选取前30只股票**作为股票组合，每月更新持股。

效果简评：整体市场环境较好时，容易获得超额收益率，**但当行情较差时，组合会出现较大回撤。长期而言，组合盈利能力较强。**

基本面选股策略分析

1 价值与成长

基本面选股中最多的两个流派是价值和成长，价值派策略更为稳定，成长派策略更激进，两者并没有绝对优势，需要根据市场风格来调整。当前市场更属于成长风格，尤其是当行情较好时。

2 长期收益率

无论是价值还是成长选股，基本面选股方法往往更容易获得长期收益率，这个与技术面选股有着本质的差异。而基本面选股取得长期收益率的根本动力是挑选出盈利持续增长的上市公司(股票)。

3 市场风格

随着监管政策，科技发展以及投资者结构变化等因素，市场的风格并不是一尘不变的。2017年和2018年，价值组合更容易取得较好的收益率，而2019年和2020年，成长组合更容易取得较好的收益率。当然，投资者可以构建成长&价值组合。

量化资金面选股策略

量化资金面选股策略，**主要是选定初始股票池，再采用多个资金指标筛选出若干只股票进行数量化组合交易。**

一般而言，资金面股票组合有效性较短，且长期持有不一定能取得较好的收益。资金指标选股与技术指标选股类似，业内通常采用相对选股法。另外，**业内一般选择中证500、创业板指数成份股作为股票池**，因为这指数内的股票散户参与度较高，符合资金指标的特性。**根据指标方向，可以构建资金流入或资金流出的股票组合。**需要注意的是，资金流入的股票不代表未来涨幅一定更好。

根据同花顺免费行情终端资金展示，主要资金指标有大单净量、DDE、资金净流入率等衍生指标。当然，在研究时可以选取大中小单的流入流出等基础指标，设计新型的衍生资金指标。

资金流入组合-创业板指数成份股

资金流入选股：从创业板指数成份股中，**按资金净流入率指标从大到小排序，选取前30只股票**作为股票组合，每周更新持股。

效果简评：**市场整体行情较好时，构建资金流入组合，可以容易获得超额收益；**

但是面临下跌和震荡行情时，容易持续亏损。

资金流出组合-创业板指数成份股

资金流入选股：从创业板指数成份股中，**按资金净流入率指标从小到大排序，选取前30只股票**作为股票组合，每周更新持股。

效果简评：资金流出组合与流入组合相比，两者效果接近，说明该指标选股效果差异性小，并非是选股的好指标。

资金面选股策略分析

1 同质化与差异性

从资金流入和流出两个组合的效果差异分析，两个组合的收益高度相关，说明该指标对组合收益的影响很小，无法影响下一期的收益率。

2 指标失效

同一个指标不同分层下组合的效果差异接近时，业内俗称“指标失效”，投资者应该优化指标算法或者更换指标。

3 指标优化

一般而言，指标优化有2个方向。一个是纵向深度优化，将指标细化，从月级别细化到周级别，周级别优化到日级别，甚至优化到分钟高频级别，重构指标。另一个是横向优化，将该指标与其他指标相互结合，优劣互补。

量化行业选股策略

量化行业选股策略，**主要是选定单个或多个行业，再采用基本面选股法筛选出若干只股票进行数量化组合交易。**

通常情况下，业内主要做单行业选股或者某类行业的选股，因为行业间的差异性较大，很难用同一套逻辑去筛选行业内的优质股票，比如市盈率指标选股，如果不区分行业，那么银行股的市盈率普遍较低，因此市盈率选股结果都是银行股，比如负债率，如果不区分行业，那么银行股的负债率普遍很高，因此负债率选股结果很难出现银行股。以上都是个行业特性决定的。另外，也可以做某类行业，比如煤炭和钢铁行业，两者接近度很高，但是不能组合通信和煤炭，这两者差异就很大。

行业选股中，主要采用基本面选股，因为技术面、资金面等选股法是不区分行业的，这些方法中所有股票都是一致的，而基本面选股法中，每个行业都可以区别对待，其背后的上市公司运作模式大不相同。

高回报组合-食品饮料

高回报组合：从食品饮料行业成份股中，筛选市盈率在0-40区间的股票，并**按投入资本回报率指标从大到小排序，选取前5只股票**作为股票组合，每月更新持股。

效果简评：**由于选择非周期食品饮料行业，并挑出高资本回报率的股票组合，因此整体策略收益率非常可观。**

捡便宜组合-房产金融

捡便宜组合：从房地产和金融服务成份股中，选出破净股票，并**按股东权益收益率（ROE）指标从大到小排序，选取前5只股票**作为股票组合，每月更新持股。

效果简评：**2015年市场过热时，全市场无破净股票，策略保持空仓规避风险。该策略具备长期盈利能力，单盈利能力一般。**

行业选股策略分析

行业选股策略研究过程中，需要抓住行业特征。这些特征需要从行业内公司的经营模式出发，采用业务模式对标财务指标的方法，从中找到有效的财务指标，这些财务指标可以快速衡量该行业内股票的经营处境

除了上述传统方法以外，**业内普遍采用遍历法，将所有的财务指标对某个行业进行测试，寻找到有效的财务指标，**其本质就是数量化投资。该方法的优势是门槛相对较低，效率更高。

组合优化器概述

组合优化器，**主要用于确定股票组合内各个股票的占比权重，业内目前使用并不频繁。**

通常情况下，我们选出若干只股票后，采用等权或总市值加权的方式确定股票组合内股票的权重。**实际上，等权配置并不一定是最优配置，组合优化器主要用于配置权重。其工作原理是根据各股票的历史表现，并确定股票组合的优化方向，最终计算最优权重。**

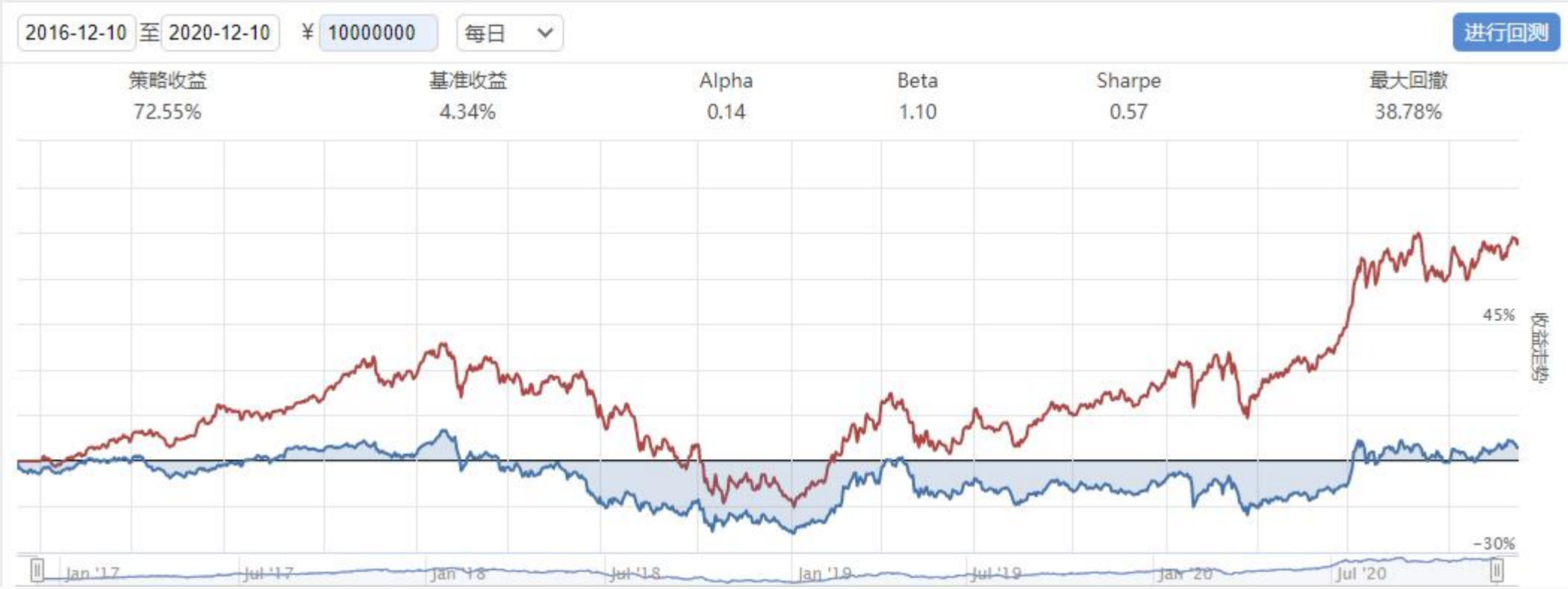
业内常见的优化方式由最大信息比率、最大夏普比率、效用最大化。

最大信息比率：指的是组合超额风险带来的超额收益最优；

最大夏普比率：指的是组合承受单位风险所获得的超额报酬最优；

效用最大化：指的是组合一定风险厌恶下收益最优。

初始组合



初始组合：从沪深300指数成份股中，**按投入股东权益收益率（ROE）指标从大到小排序，选取前60只股票**作为股票组合，每月更新持股。

夏普率最大化

基于初始组合，采用组合优化器，优化方法为夏普率最大化。

效果简评：**经过组合优化后，整体策略的效果下滑较为明显，并没有增强收益和控制风险。**

效用最大化

基于初始组合，采用组合优化器，优化方法为效用最大化。

效果简评：**经过组合优化后，整体策略的效果依然下滑较为明显，并没有增强收益和控制风险。但比夏普率最大化的结果稍微好一些。**

最大信息比率

基于初始组合，采用组合优化器，优化方法为最大信息比率。

效果简评：**经过组合优化后，整体策略的效果依然下滑明显，并没有增强收益和控制风险。**

优化分析

1 历史与未来

目前，所有的组合优化器都是基于历史数据，推导出最优权重，试图在未来行情中增强组合收益，该方法整体上是符合量化投资逻辑，但是效果并不是特别理想，主要原因是历史数据的信息含量太少，且采用很早之前的马科维茨投资组合理论。

2 等权与加权

股票组合权重层面，业内主要还是使用等权或者总市值加权的方式为主，目前还无法定论等权/总市值加权是最佳方法，但普遍使用。

04 量化技术面择时策略

量化技术面择时策略

量化技术面择时策略，**主要采用单个或多个技术指标，对某个具体标的进行数量化交易。**

通常情况下，业内选择对指数进行择时，包括上证指数、沪深300、创业板等，并交易指数对应的ETF品种。因为交易个股面临系统风险以外，还有非系统风险，包括经营风险、暴雷事件等，所以个股择时的难度系数更高，需要考虑的因素更多。

另外，各技术指标间的择时效果差异很多，且应用行情不同，有些技术指标主要在行情上涨时使用，有些是行情震荡，还有一些是行情下跌，因此需要综合考虑后再使用。

除此以外，**单个指标并不一定比多个指标的效果差，使用多个技术指标时，需要仔细分析各个指标的优劣势后再使用。**

量化MA择时-沪深300指数

MA择时：采用20日和60日双均线，对沪深300指数进行日级择时，20日均线大于60日均线满仓，否则空仓。

效果简评：**策略在指数上涨阶段大幅盈利，但在震荡行情中持续亏损。**

量化MACD指标择时-沪深300指数

MACD择时：采用MACD(12,26,9)指标，对沪深300指数进行日级择时，MACD大于0时满仓，否则空仓。

效果简评：**策略在指数上涨阶段大幅盈利，但无法躲避下跌行情，出现较大回撤，在震荡行情中亏损较小。**

量化双指标择时-沪深300指数

MACD+MA择时：采用MACD(12,26,9)和MA(20,60)指标，对沪深300指数进行日级择时，双指标看多则满仓，否则空仓。

效果简评：**策略在指数上涨阶段盈利，且可以躲避下跌行情，在震荡行情中亏损非常少。**

技术择时策略分析

1 MA优劣势

均线作为最传统的技术指标，其优势是在大涨大跌行情中获利并规避大跌。但是遇到震荡行情时，往往反复打脸，出现持续亏损的情况，因此需要寻找辅助指标，剔除掉震荡行情中的伪上涨信号。

2 MACD优劣势

MACD指标起源于均线指标，相对于均线指标，MACD指标还能反应价格涨跌动量，也就说只有股价上涨且出现动量增强，MACD指标才会提示开仓。但是当行情下跌时，只要下跌动量减弱，MACD指标也会开仓，这是劣势。

3 指标组合

多指标组合秉承优劣互补原则，上涨行情时，MACD和MA指标可同时看多，当行情涨过头时，动量下降，MACD提示空仓，提前获利了解。行情下跌时，均线可以完美规避风险。行情震荡由于动量不足，MACD指标可以弥补MA指标的劣势，寻找有效上涨信号

谢谢聆听

同花顺量化实验室负责人-陈城