量化辅助系统可行性报告

何彦涛

摘要:

本项目旨在通过量化分析的技术手段对中国股票二级市场进行择股、择时方面的研究，以期能达到帮助分析师进行决策以及辅助印证分析师判断的目的。计划使用R语言进行开发，全系统分为两大类，分别为多因子选股模型以及量化择时模型。成果展现形式现暂时规划为以软件自主生成图、文、数据相结合的研究报告，日后可能依据开发过程更改。

1. 设计目标

本系统的设计目标主要是希望能辅助分析师对股票市场进行研判。具体地说，量化选股模型既可以做到为分析师进行大量股票初步筛查以减小工作量，又可以在经过充分调教以后对分析师的判断进行佐证。择时模型可以同时监视整体市场的方向变动以及某些标的的变动，对于买卖时机的选择、行情反转的监控可以起到提示作用。

第一期的目标在于搭建起基本的模型框架，为之后进一步进行模型的优化、调整提供基础。

1. 基础设计
   1. 选用工具

考虑使用R语言以及万得数据库进行开发。考虑到R语言较为强大的统计分析能力以及高速稳定的处理效率，同时考虑到Wind数据库对于R的接口设计成熟易用，这样的搭配可以基本保证系统的功能得到实现。

* 1. 系统组成

现暂规划为两个部分：择股模型以及择时模型。

其中择股模型初步计划采用多因子模型框架；择时模型初步考虑为三个子系统，包括短期择时，中长期择时以及市场异常监控。

1. 详细设计
   1. 多因子模型
      1. 设计理念

多因子模型由于采用了大量基本面数据，因而计划更新周期为一季度一次，即模型应在每个季度导入新数据进行分析，其所得结论也应该在一个季度或更长时间内有效。基于其更新周期，该模型的核心能力在于找出被低估的或者被市场长期追捧的股票类型。其不应当追求短时利润、进行盲目的追涨杀跌，而应当在更长的时间周期内寻求长期表现稳定的标的。

* + 1. 开发方案

第一步：多因子模型初步因子选定为规模类、量价指标类、估值水平类、盈利能力类、成长能力类。

第二步：通过历史数据回测判定有效的因子种类

第三步：进行主成分分析，合并类似因子达到降维目的

第四步：进行回归分析以计算各因子权重数据

第五步：依据得到的回归方程，投入使用进行预测，并根据预测结果与实际的偏差进行参数优化

* + 1. 后续改进方案

在模型基础搭建完成后，进行以下改进：

* 将模型的参数优化改为自适应形式，即将近期的数据赋予更高的权重，以确保模型适用于最新的市场
* 按照不同的行情走势阶段（牛、熊、振荡）以及不同的行业分别建立回归方程，以更好地适应各个市场环境
* 请教研究部资深同事们对于行业和个股分析的见解，有针对性地对一些重要的因子权重进行加强，或者引入新的因子
* 进行较多实验以明确模型优缺点，摸索其有利的使用环境和合适的使用方法
* 通过请教研究员与阅读书籍，寻找其它可能的定价因子，进一步扩充模型。现计划阅读的书目为格雷厄姆的《证券分析》
* 反复优化参数，反复优化参数，反复优化参数
  1. 择时模型

择时模型现暂时考虑使用三个子系统：中长期择时，短线择时以及异常监控。

* + 1. 中长线择时
       1. 设计理念

本模型旨在判断趋势市的出现与消失，判断持续期高于六十天的中长期趋势。计划采用技术分析中的三重指数平滑线以及MACD线或其他周期较长的移动均线。目标调仓周期一个月。

* + - 1. 开发方案

第一步：通过万得接口读取历史行情数据，制作技术指标

第二步：判读技术指标是否产生买入或卖出以及反转等信号

第三步：通过历史数据回测，每月产生新的持仓标的，计算策略效益

第四步：调整模型灵敏度，综合考量交易成本，优化参数

* + - 1. 后续改进方案
* 由于为较长期的择时模型，有必要增加对于经济形势，市场热度以及市场信心的判断功能，具体方法尚在开发之中。
* 通过阅读技术分析书籍，寻找更多更好的技术分析指标，现在考虑阅读的书目为约翰墨菲的《期货市场技术分析》
* 考虑引入神经网络，小波分析和支持向量机等机器学习的方法
* 反复优化参数，反复优化参数，反复优化参数
  + 1. 短线择时
       1. 设计理念

本模型旨在判断具体买入卖出时机，目标调仓周期两到三天。计划采用布林线，资金流向，以及短期均线等周期较短的技术指标。

* + - 1. 开发方案

第一步：通过万得接口读取历史行情数据，制作技术指标

第二步：判读技术指标是否产生买入或卖出以及反转等信号

第三步：通过历史数据回测，每日产生新的持仓标的，计算策略效益

第四步：调整模型灵敏度，综合考量交易成本，优化参数

* + - 1. 后续改进方案
* 由于为较短期的择时模型，有必要增加对于热点趋势，资金流向乃至政策面和主题类的判读能力。具体操作方法有待进一步学习。
* 通过阅读技术分析书籍，寻找更多更好的技术分析指标，现在考虑阅读的书目为约翰墨菲的《期货市场技术分析》
* 反复优化参数，反复优化参数，反复优化参数
  + 1. 异常监控
       1. 设计理念

本模型旨在判断市场异常动向，判断市场大拐点的出现。本模型具体的设计思路尚待明确，现在可以确定的是，兴登堡凶兆的监测将列入此模型。

1. 进度安排

多因子模型已经于八月十三日开始开发，现处于读取因子数据与计算因子大小的开发阶段。希望于一周内基本完成模型开发，进入参数修正阶段。择时模型之前已经开发过MATLAB为基础的类似模型，而且相对而言工作量小于因子模型。希望在一周内完成基础模型开发，但其参数优化需要大量测试，时间难以预计。总的来说，希望在一个月到一个半月内完成第一阶段初步搭建模型框架的工作。

何彦涛