1.论文题目

2.摘要

3.关键词

# 关键词：python、仓位管理、最优风险资产组合、多因子模型、

1. **摘要**

针对第一问，依据题意，利用题目给出的指数增强策略，根据收益率、斜率、逐个进行单项排名，将排名累计后取综合排名前十支股票，此为强势股。确定选股后建立模型，将粤港澳指数大湾区指数30支成分股的每天的收益率，对应得出2011-2020年的收益率曲线，进而对该方案做出了在收益预测方面，表现良好，累计收益达到36.1918%，但亏损次数偏多，胜率偏低的评价。将我们的收益率曲线与粤港澳指数大湾区指数曲线进行对比，发现从2011年开始连续两年该策略基本跑赢粤港澳指数大湾区指数，在2013年后，策略的收益率无法跟上粤港澳指数大湾区指数走势，说明该策略在小熊市时期表现良好；但在单边上涨大行情中表现不佳，仍有很大改进空间。

针对第二问，题目给定策略交易的时机，因此只需要通过仓位分配来提高收益并且尽量降低风险。因此我们采用经典的仓位分配理论哈里马尔维茨的均值-方差投资组合理论来进行仓位的分配。我们利用最小方差模型和最大夏普率模型分别进行交易，比较两种交易策略在2011年-2020年的收益率、波动率、方差和夏普率。采用SLSQP优化算法，在一定误差范围内和限制条件下，从大量随机数据组合中得出最小方差模型和最高夏普率模型的权重分配，采取该权重分配方案并购入股票。最小方差模型以年收益率12%的微弱优势胜于最大夏普率10.2%。

针对第三问，调仓时间不变，单只股票的投资额为本金10%，我们需要选择自己的策略。通过分析总市值因子、净市率、权益回报率、资产回报率、换手率相对波动率、6日乖离率这六个因子，之后分别分析六个因子的单独收益，然后加权每支股票六个因子的IC值，演算后对其进行排序，选出前十支股票，此即我们的选股策略。

针对第四问，海龟策略来对仓位的控制和对交易时间的把握，通过第三问的多因子策略，综合选择出排名前十五的十五支股票。

**二、问题重述**

指数增强策略采用量化增强模型，追求高于标的指数回报水平的投资，同时力求进行有效的风险控制、降低交易成本、优化投资组合。指数增强策略不会对跟踪标的成份股进行完全复制，而是会对部分看好的股票增加权重，不看好的股票则减少权重，甚至完全去掉。通过对交易成本模型的不断监测，尽可能让交易成本降到最小。综合来看，就是既做到超额收益，又控制主动风险。某证券公司选取三十支大湾区金股(表1)构建了大湾区指数 (399999),指数行情如图2所示，附录一为成份股的行情数据，附录二为大湾区指数行情数据。



图1： 大湾区指数行情数据



图2：大湾区指数行情实况

目前该公司采用的大湾区指数增强策略为：在30支成份股中选取上一周最强势的10支股票，作为下一周投资标的，每周星期一换仓，以当天的收盘价为准考虑收益，每支股票投资额固定为本金的10%，手续费万分之2.5。

1）请计算该投资策略从2011-2020年的收益率曲线，并且参照大湾区指数评价其投资成效。

2）如果调仓时间不变，但单支股票的投资额可以灵活变化，请找出该投资策略最优的收益率曲线。

3）调仓时间不变，单支股票的投资额为本金的10%，根据市场交易数据建模，设计你自己的选股策略。

4）如果调仓时间和单支股票的投资额都可以灵活变化，根据市场交易数据建模，设计你自己的选股策略。要求：前三个交易日完成建仓，仓位不低于50%，单支股票不超过总资产10%。

**三、问题分析**

**第一问：**

根据指数增强模型里对强势股的定义，选择收益率、斜率、为判断依据。从第一周开始以一周为单位，计算出（）并择出10支强势股，作为下一周投资标的。通过python运行得出2011到2020年的收益率曲线并与大湾区指数对比。

**第二问：**

该题为配资仓位管理问题，可运用最优风险资产组合理论求最优解。

**第三问：**

我们组选用经典多因子选股理论基础，通过分析总市值因子（PE）、净市率（BP）、权益回报率（ROE）、资产回报率（ROA）、换手率相对波动率（volatility）、6日乖离率(bias 6)这六个因子，之后分别分析六个因子，其次加权每支股票六个因子的IC值，演算出的后对其进行排序，选出前十支股票。此即为我们的策略。

**第四问：**

根据前三问的结果，（）

**四、名词解释**

**五、模型的建立与求解**

**5.1第一问：**

**5.1.1理论基础与原理**

1、先将收益率从最高开始排序，即 ，排序后为每一支股票附上 1-30 的排序序号。

2、将斜率按从大到小排序。即用最小二乘法公式：设拟合直线的公式为：



其中，拟合直线的斜率为：



计算出斜率后，进行排序。排序后为每一支股票附上 1-30 的排序序号。

1. 将按从大到小排序，若平方大于0.75，表示模型拟合度很好，可解释程度较高；小于0.5，表示模型拟合有问题。

=

计算出后，对进行排序，之后为每一支股票附上 1-30 的排序序号。

4、对三个指标进行等权累加后进行对等权累加后的指标进行综合排序，取综合排序前十只股票，即为每周选出的10支强势股票。

**5.1.2模型的建立**

涉及到为什么选此三个指标来选股，思路大概是想从哪些方面可以体现出强势股票。首当其冲是收益率，通过分析过去数据，考察上周的上涨程度，来分析下一周中的股票强势程度。其次是拟合直线的斜率，斜率可以反映整体股票的走势，斜率越大股票上涨程度越大。最后是，相关系数，值越大，说明我们拟合的直线越能反应该股票的真实走势。故三个指标综合排名前十的股票即为我们选出的下周强势股。

运用附录一的粤港澳大湾区指数30支成分股的数据，使用Python进行策略的编写最后使用Auto-Trader 进行策略的回测 最后得出2011-2020年的收益率曲线，如图3

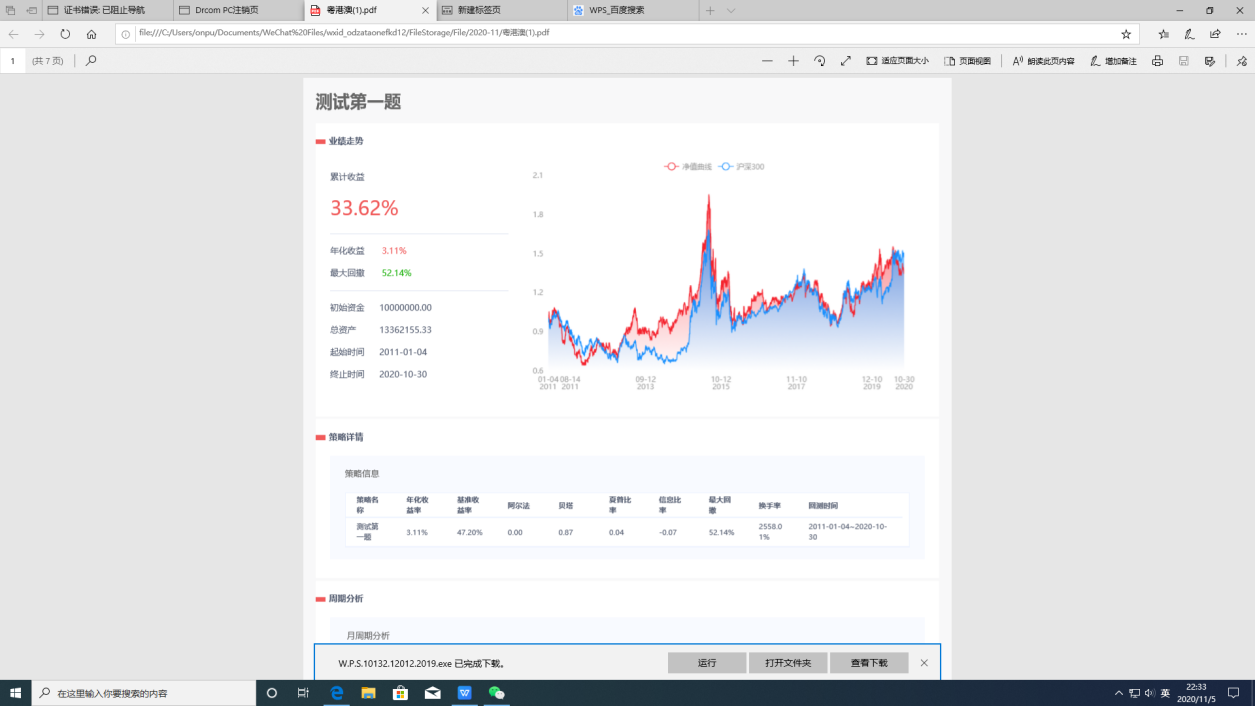


图 3： 第一题交易策略回测图

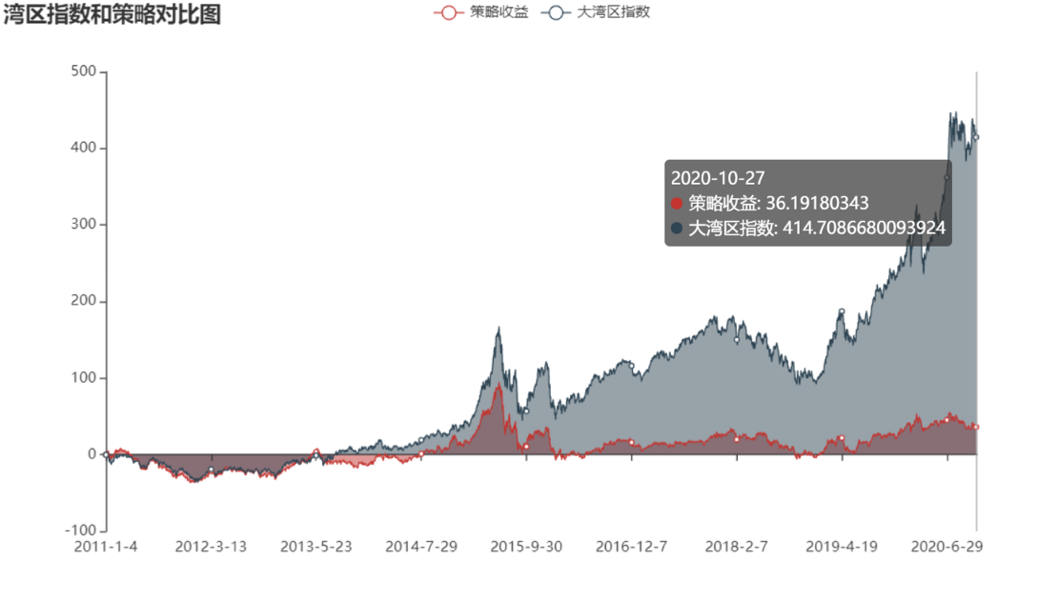


图 4：策略收益曲线与大湾区指数的对比图

现分析投资策略成效，据图2，策略累计收益为36.1918%，据图3，年化收益为3.11%，最大回撤为52.14%。以股票策略而言，我们认为好的量化策略必须具备以下的共性：

1、收益预测的有效性与特异性

有效性很容易理解，量化策略的收益预测模型必须是有很扎实的理论依据与逻辑基础，并且是适用于所要投资的市场的一个模型，例如把美股的动量因子用到A股，会发现A股的反转效应更突出；特异性是指预测模型和市场上其他参与者的模型的相关性要低，最直观的例子就是大家如果都认可并使用的Alpha因子会逐渐变为Beta因子，同质化会影响收益预测的表现。

2、风险控制模型

好的量化模型必须有应对风险的能力，即需要一个能准确预测组合风险的风险模型。例如我们在实施量化对冲模型的时候，我们会利用优矿的风险模型去控制组合在各个风格上的暴露，控制组合与基准指数的跟踪误差，确保组合能稳定地获取超额收益；可以通过绩效归因去对收益进行分解， 一个好的量化策略应该是收益绝大部分来自于选股而非风格与行业的策略，风控能力是量化策略好坏的核心之一。

3、交易模型

交易模型就是策略在实施的能最小化交易成本的模型，例如你的交易如何应对流动性的问题，如何去尽可能地减少冲击成本，做空如何处理期指的升贴水等等，这是一个策略实际运作无法避免的问题，也是策略实际运作的保证。

具体到本指数增强策略，在收益预测方面，表现良好，累计收益达到36.1918%，但亏损次数偏多，胜率偏低。

**5.2第二问[1]：**

本题目中，我们掌握了一种风险资产，且股票组合、换仓时间等已确立，只通过仓位分配来提高收益和降低风险。可利用Harry M. Markowitz的均值-方差投资组合理论分析10支股票的收益率、波动率、方差和夏普率。

**5.2.1 理论基础与原理**

**（1）Markowitz均值-方差模型**

马科维茨把风险定义为[收益率](https://baike.baidu.com/item/%E6%94%B6%E7%9B%8A%E7%8E%87/4533890)的[波动率](https://baike.baidu.com/item/%E6%B3%A2%E5%8A%A8%E7%8E%87/8350659)，首次将数理统计的方法应用到[投资组合](https://baike.baidu.com/item/%E6%8A%95%E8%B5%84%E7%BB%84%E5%90%88/3520475)选择的研究中。这种方法使收益与风险的多目标优化达到最佳的平衡效果

依据以下几个假设：

1、投资者在考虑每一次投资选择时，其依据是某一持仓时间内的证券收益的[概率分布](https://baike.baidu.com/item/%E6%A6%82%E7%8E%87%E5%88%86%E5%B8%83/828907)。

2、投资者是根据证券的[期望收益率](https://baike.baidu.com/item/%E6%9C%9F%E6%9C%9B%E6%94%B6%E7%9B%8A%E7%8E%87)的方差或标准差估测[证券组合](https://baike.baidu.com/item/%E8%AF%81%E5%88%B8%E7%BB%84%E5%90%88)的风险。

3、投资者的决定仅仅是依据证券的风险和收益。

4、在一定的风险水平上，投资者期望收益最大；相对应的是在一定的收益水平上，投资者希望风险最小。

给出数学形式：

资产组合的均值和方差为：

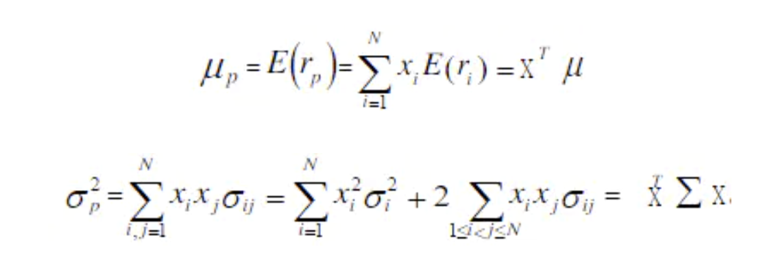


图5： 均值方差计算公式

**（2）优化求最优解：SLSP 方法**

使用Sequential Least Squares Programming 算法方法，给予

边界条件：每个权重在0到1之间；

约束条件：权重之和为1。

可以得到最佳的权重分配，借助这个算法，我们可以去 求得可以获得 最大收益率 和 最小方差的股票组合。

**5.2.2 模型的建立与求解**

**（1）初步统计数据**

首先计算个股的对数收益率 (Log Return)，用当天的收盘价与前一天相比较。计算年化收益率和协方差矩阵。

**（2）随机取得不同投资组合的夏普比率 和 方差值**

每只股票都有一定权重。此资产组合的收益率（均值）为组合中个股收益率（均值）的权重之和。

使用蒙特卡洛模拟生成随机数权重随机分配给10支股票，产生大量投资组合，每一组投资组合由一组随机权重组成。将其收益率和波动率用散点图描述，并计算每个组合的夏普率。定义函数，使输入权重分配时输出该组合的收益率、波动率和夏普比率。

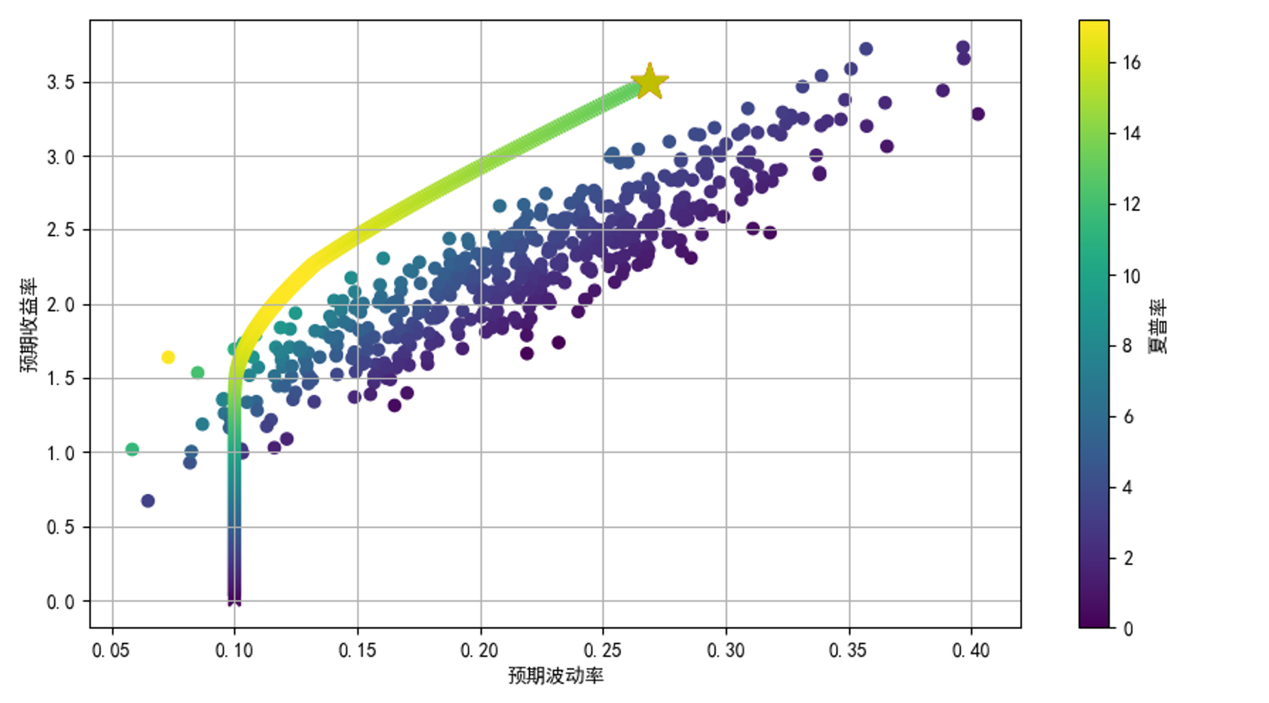


图6：蒙特卡洛模拟数据和有效边界曲线

**（3）最大夏普比率和最小方差优化**

若某投资组合的夏普比率最高，则该组合夏普比率最优，但夏普比率最优的点不一定是所产生的随机投资组合中的一个。所以需要通过优化算法，寻找出一个恰当的权重分配输入来输出最大的夏普比率。

使用最小优化算法 SLSP 方法，将最大化夏普比率问题转变为最小化负的夏普比率问题。利用Scipy.optimize库中的minimize函数，以及边界条件及约束条件，得出结果，即最大夏普比率的投资组合的权重分配。



图7： 最大夏普比率组合回测图

**5.2. 模型的分析**

。。。。。。。

**5.3第三问：**

**5.3.1理论基础与原理**

本题目中谈到调仓时间不变，单只股票的投资额为本金10%，需要选择自己的策略。我们组选用经典多因子选股理论基础，先通过分析规模类：总市值因子（PE）；估值类：净市率（BP）；每股指标类：权益回报率（ROE）、资产回报率（ROA）；流动性类：换手率相对波动率（volatility）和动量类：6日乖离率(bias 6)这六个因子，之后分别分析六个因子，其次加权每支股票六个因子的IC值，演算出的后对其进行排序，选出前十支股票。此即为我们的策略。

**（1）经典的多因子选股理论基础**

**多因子模型**是关于资产定价的模型。与资本资产定价模型和单指数模型不同，多因子模型认为股票价格并不仅仅取决于股票的风险，还取决于其他一些因素。我们组从众多因子中选出六个，下面将详细分析每个因子及其IC值，对其进行排序。

这里改IC，IR的解释 说明 以及 IC加权的算法

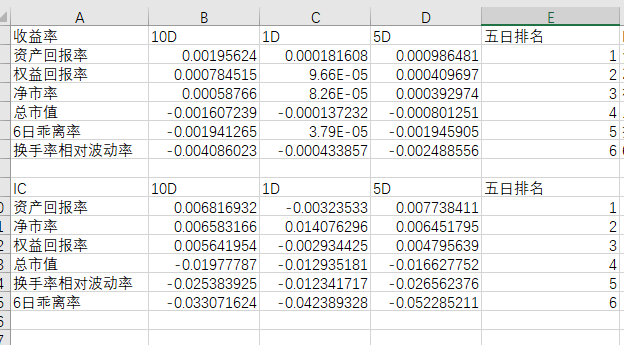


图8: 六因子的回测IC 和收益率排名图

我们将分析出的IC值进行排序，得到排名前10支的股票，进行策略的演算，得到如下收益图。



图9： 多因子策略回测图

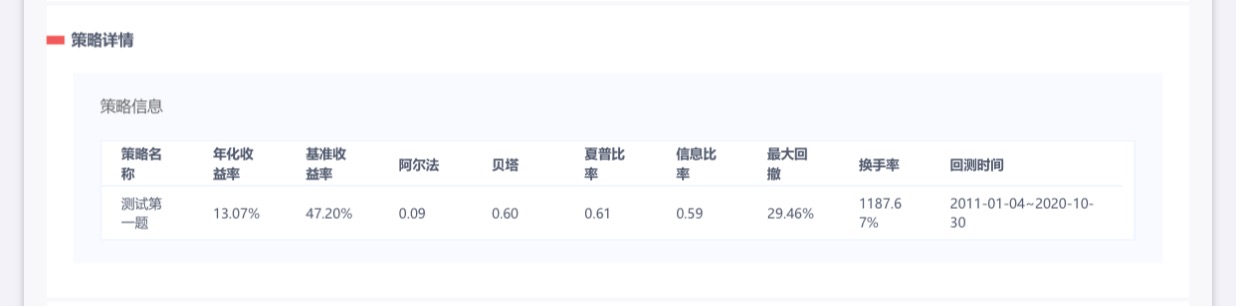


图10： 多因子策略回测数据图

**5.4第四问：**

**5.4.1理论基础与原理**

