

# 北京大学第二十二届“江泽涵杯”数学建模竞赛题目

请先阅读竞赛规则：

<http://portal.math.pku.edu.cn/htdocs/showarticle.php?id=18977>

## B题题目：沪深300指数交易策略及收益率

### 赛题背景

沪深300指数作为中国证券市场的重要指标，其变化直接影响着投资者的决策和市场情绪。该指数由沪深两市市值大、流动性好的300只A股股票组成，自2005年4月发布以来，已成为衡量市场表现的关键工具。沪深300指数的预测对于投资者理解市场趋势、制定投资策略具有重要意义。本赛题旨在通过数学建模和数据分析，提高对沪深300指数未来走势的预测准确性，并在此基础上构建有效的交易策略。

本大题共分2个小题目，请依次作答。

### 题目一：交易频率限制下的沪深300指数交易策略构建

本题要求在已知指数未来收益率的情况下，限制交易频率，最大化策略收益。

赛题说明：

- 交易规则设置：交易开始时仓位为0，每天的仓位上限是100单位，下限是0。每天最多买入或卖出单位是10；
- 优化目标：在满足交易限制的前提下，实现整个回测周期内最大化的收益；
- 在策略设计中明确考虑交易限制，并解其对策略性能的影响；
- 交易执行的时间点为每个交易日的收盘时间点，交易价格为收盘价；
- 回测时间窗：2020年1月2日至2023年12月29日。

提交结果要求：

- 参赛者需提供策略在历史数据上的回测结果，包括累计收益、最大回撤、夏普比率等评估指标（指标计算方法见附录）；
- 提交策略的代码实现，以便评审能够复现优化结果。

## 题目二：基于沪深300成分股预测信号的交易策略设计与组合优化

参赛者将获得一组沪深300指数的成分股及其对应的预测信号。这些信号是基于各种金融模型或算法生成，旨在预测每只股票未来的表现。参赛者的任务是设计一个交易策略，利用这些预测信号来构建一个投资组合，并优化该组合以实现在回测期间超额收益率的最大夏普比率。

赛题说明：

- 交易信号为利用t日及之前的信息计算出的对未来收益率的预测的排序。投资组合的构建和调整应基于提供的成分股预测信号；
- 交易策略仅允许每天收盘时以当日收盘价买入或卖出任意数量的股票；
- 投资组合中的股票持有期不受限制，可以是短期或长期；
- 假设不存在交易成本，即所有交易均无需支付手续费或税费；
- 组合优化的目标是在回测期间超额收益率的最大化夏普比率，即在考虑风险的基础上最大化投资组合的超额回报；
- 回测时间窗口：2021年1月4日至2023年12月29日。

超额收益率的定义：

$$ER_p = R_p - R_{HS300}$$

其中ER代表超额收益率。

提交结果要求：

- 提交的交易策略应包括策略的详细描述、预测信号的使用方式、以及投资组合构建和调整的具体规则；
  - 提供回测期间投资组合的表现分析；
  - 提交投资组合优化过程的代码实现，以便评审能够复现优化结果。
-

## 附录

### 评估指标计算

#### 累计收益 (Cumulative Return):

累计收益是指在一段时间内投资组合收益的总和，计算公式为：

$$\text{Cumulative Return} = \prod_{t=1}^N (1 + r_t) - 1$$

其中，

$r_t$  是第t期的收益率

#### 最大回撤 (Maximum Drawdown):

最大回撤是指投资组合在某一段时间内从峰值下跌到谷底的最大跌幅，计算公式为：

$$\text{Max Drawdown} = \frac{\text{Peak Value} - \text{Trough Value}}{\text{Peak Value}}$$

其中，Peak Value是峰值时的投资组合价值，Trough Value是随后的谷底价值。

#### 夏普比率 (Sharpe Ratio):

夏普比率是衡量风险调整后收益的指标，计算公式为：

$$\text{Sharpe Ratio} = \frac{R_p}{\sigma_p}$$

其中，

$R_p$  是投资组合的平均收益率；

$\sigma_p$  是投资组合的标准差（风险）；

注：在本比赛中假设无风险利率为0；

## 数据说明

hs300.csv: 包含 2007-2023 年 hs300 指数的加量数据，涨跌幅；

hs300\_constituent\_stocks.csv: 包含 2007-2023 年 hs300 指数成分股的价格数据，涨跌幅和权重；

hs300\_constituent\_stocks\_predict.csv: 包含 2021-2023 每个交易日对 hs300 指数成分股预测值；