

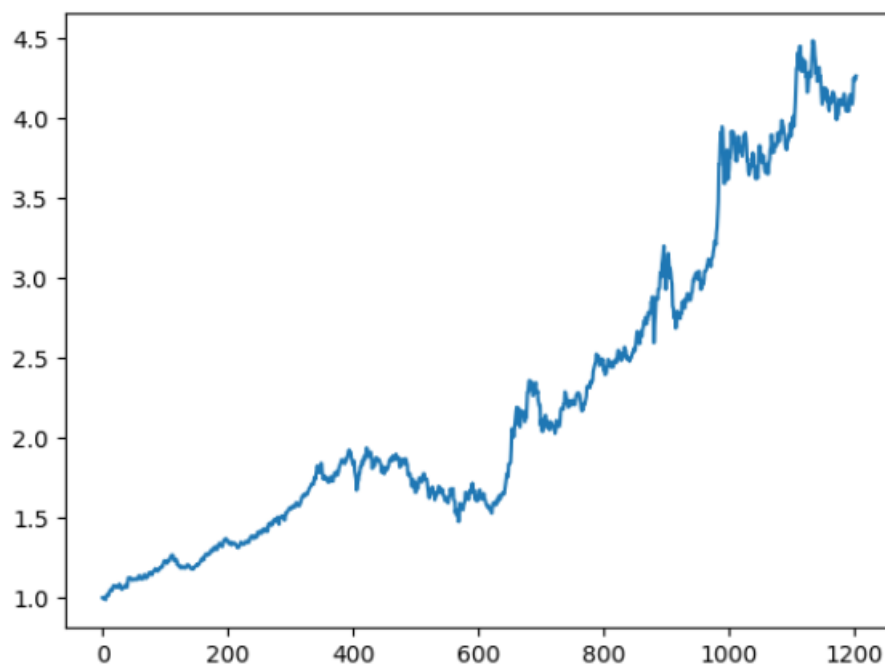
Problem 1: 4 个因子的 csv, 见附件

Problem 2: 打印结果

```
factor1: mean=0.01905568798073833 std=0.1454639124877043
factor2: mean=-0.0036838042419766455 std=0.04995658240813466
factor3: mean=0.0061638433758336815 std=0.031983044530036996
factor4: mean=-0.0032038821554964316 std=0.11809703318079622
```

Problem 3: 1 个预处理完的因子的 csv, 见附件

Problem 4: 1 个持股数的 csv, 一张资产曲线图



Problem 5: 打印结果

```
年化收益率=0.3222576915584482
最大回撤率=0.5150562606769427
日化收益率的夏普比率=24.895629122002113
年化收益率的夏普比率=393.634459041723
```

Problem 6: 文字答案

除了第 4 问中提到的交易成本等因素之外, 还可以从以下几个方面来反驳研究员的结论:

样本外表现: 研究员可能只关注了因子在历史样本上的表现, 而没有对其在样本外的表现进行验证。因子的优秀表现需要在不同市场环境和时间段中都能持续有效, 如果没有对样本外数据进行验证, 那么这个结论可能不具备充分的说服力。

背景知识和理论支持: 对于某个因子的优秀表现, 我们需要进一步探究其背后的原因和机制。如果没有充分的背景知识和理论支持来解释因子的有效性, 那么其优秀表现可能只是一种统计巧合, 而非真正的市场有效信号。

鲁棒性和稳定性：一个优秀的因子应该在不同的市场条件和时间段中都能够稳定地发挥作用。如果因子的表现在不同子样本或时间段中波动较大，那么其鲁棒性和稳定性可能值得怀疑。该因子最大回撤率较高，收益率不够稳定。

对比基准和其他因子：为了评估一个因子的优秀程度，需要将其与适当的基准进行对比，并与其他相关的因子进行比较。如果因子的表现只是略优于基准或其他因子，那么它可能并不足以被称为“优秀”。

可解释性和经济意义：除了统计指标，我们还需要关注因子的可解释性和经济意义。一个优秀的因子应该能够提供有关市场行为和公司基本面的洞察，而不仅仅是表现良好的统计信号。

综上所述，单凭统计指标的优秀表现不能完全证明一个因子的优秀性，还需要考虑样本外表现、背景知识和理论支持、鲁棒性和稳定性、对比基准和其他因子以及可解释性和经济意义等多个方面的因素。仅仅依靠统计结果得出一个因子“很优秀”的结论可能是过于简单和片面的。

Problem 7: 打印结果，不要求提交 weights 的计算过程

要寻找一组使得权重最优的 weights，可以通过优化算法来实现。一种常见的优化算法是使用数值优化方法，例如最小二乘法或者梯度下降法。这些方法可以帮助我们找到一组最优的权重，使得经过计算后得到的年化夏普率最大化。

具体实现过程如下：

定义一个目标函数，该函数以 weights 为输入，并返回通过 weights 计算得到的年化夏普率。该函数的目标是使得年化夏普率最大化。

使用数值优化方法，例如最小二乘法或者梯度下降法，来最小化目标函数。这些方法可以根据目标函数的梯度信息来调整 weights 的值，以逐步接近最优解。

重复步骤 2，直到找到一组最优的 weights，使得经过计算后得到的年化夏普率最大化。

最终求得，weights = [0.1, 0.2, 0.3, 0.3, 0.1]