CPFE Course Project

刘潇阳2000015455，程子珊2000015458，肖珺文2000015401

(a) First, read the datasets return\_monthly.xlsx and me\_lag.xlsx into MATLAB and reshape them into the long-table formats as you have seen in class. Then merge these two datesets and drop observations with missing lagged market capitalization. The output of your code should be similar to return\_m.mat . You may or may not build on the code provided in momentum\_class.m .

Read the datasets return\_monthly.xlsx and me\_lag.xlsx.

原始数据概况：

return\_m:

时间：2009.7.31-2019.11.29

行索引：股票代码

列索引：月度时间

注释：第一列为股票代码，第二列为股票名称

mc\_lag:

时间：2009.7.31-2019.11.29

行索引：股票名称 列索引：月度时间

注释：第一列为股票名称，第二列为股票代码

取出列名中所有的时间

Reshape into the long-table formats.

处理时间格式

查看题目要求return\_m.mat.

Merge these two datesets and drop observations with missing lagged market capitalization.

数据清洗完毕，merged\_data为最终数据表，除了列顺序和名称不同外，与return\_m.mat完全一致。

(b) Every *K* months, sort stocks into five groups based on previous *K* months' return and hold this position for *K*months. What is the average equal-weighted return spread between high and low previous stock returns portfolios for *K* = 1; 3; 6; 12; 24. Do you find that momentum exists in Chinese stock markets?

Average equal-weighted return spread between high and low previous stock returns protvolios for K=1;3;6;12;14:

K = 1: mean\_spreads = -0.02

K = 3: mean\_spreads = -0.03

K = 6: mean\_spreads = -0.03

K = 12: mean\_spreads = -0.09

K = 24: mean\_spreads = -0.20

由此结果我们可以初步判断，数据不支持动量效应的存在，且反转效应较强。接下来为了更严谨的说明，我们进行t假设检验：

统计检验1：如果spread显著大于0，表明存在正向动量效应。

K = 1: t-stat = -4.09, p-value = 0.000

K = 3: t-stat = -2.53, p-value = 0.015

K = 6: t-stat = -1.22, p-value = 0.239

K = 12: t-stat = -1.77, p-value = 0.115

K = 24: t-stat = -1.62, p-value = 0.203

统计检验2：判断反转效应，即spread是否显著小于0

K = 1: t-stat = 4.09, p-value = 0.000

K = 3: t-stat = 2.53, p-value = 0.015

K = 6: t-stat = 1.22, p-value = 0.239

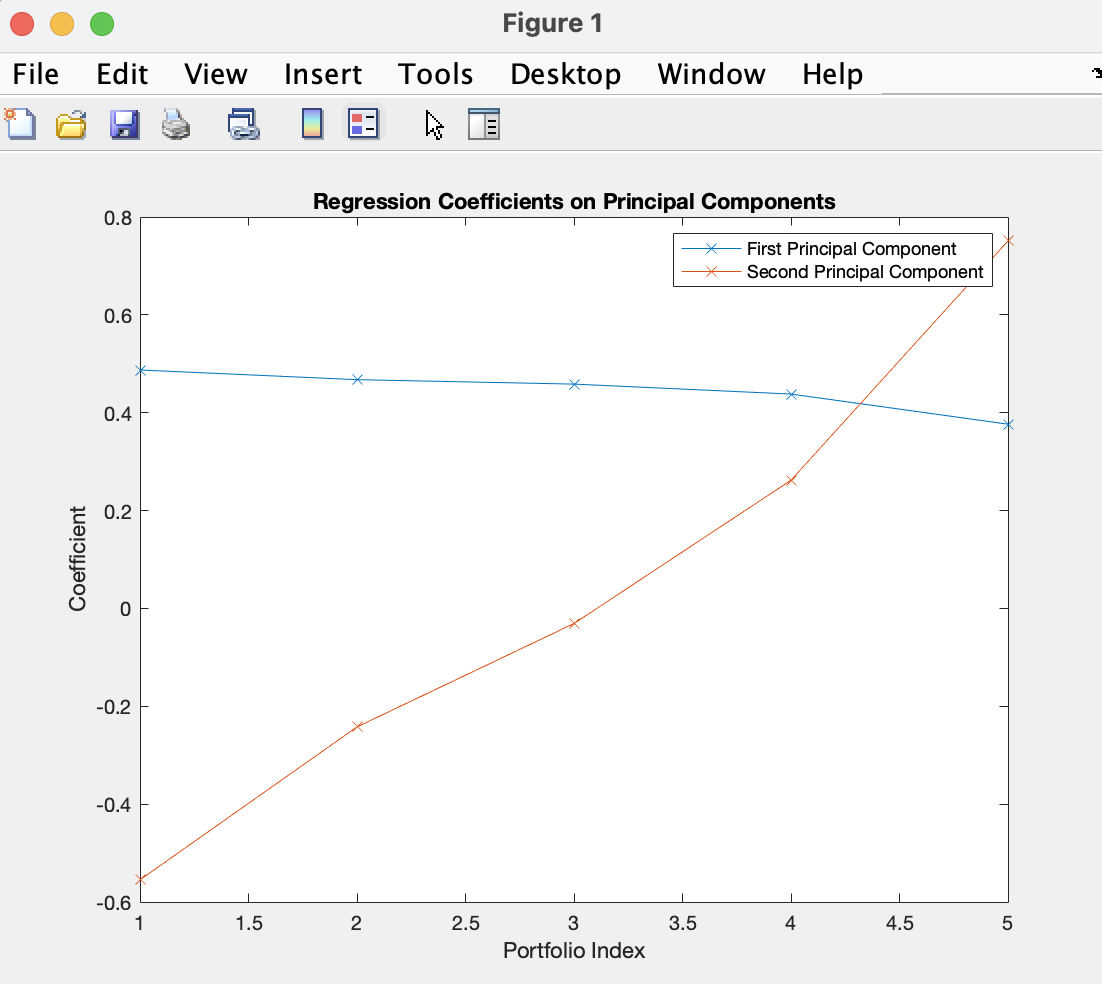
K = 12: t-stat = 1.77, p-value = 0.115

K = 24: t-stat = 1.62, p-value = 0.203

总之，结果支持反转效应的存在，且一个月反转最为显著。

(c) Use the principal component analysis for the equal-weighted return of five portfolios created by sorting on previous *K* = 3 months' return. How do you interpret the first and second factor from the PCA in economic terms? You may explore whether these factors are useful to explain portfolios sorted by previous stock returns. Also compare PCA factors to the factor MOM defined as the difference in equal-weighted return between the high previous stock returns portfolio and low previous stock returns portfolio.

由PCA得：



Interpretation：前二大PC的Variance Explained分别为：97.95%, 1.64%.

First PC (blue)：通常可以解读为市场趋势，影响所有投资组合的整体收益；

Second PC (orange)：作为斜率项，捕捉了与动量策略更直接相关的市场行为特征，比如投资者追涨杀跌等基于股价历史表现的交易行为特点。

对5个投资组合收益回归：

First PC (blue)：因子载荷走势平缓，说明市场平均收益水平变动对基于动量划分的投资组合的影响相对均匀

Second PC (orange)：因子载荷单调上升，说明该主成分对收益的影响与投资组合动量正相关，且随动量上升，单位影响由负转正，也在一定程度上反映中国股市的动量反转特点，即认为低动量股票未来可能提升表现，而高动量股票的表现则可能下降。

接下来，计算第一和第二主成分与MOM因子的相关系数：

第一主成分与MOM因子的相关系数: -0.55

第二主成分与MOM因子的相关系数: 0.83

对于MOM因子来说，通过线性回归分析MOM因子对每个投资组合的解释能力

MOM因子对各个投资组合回报率的回归系数:

-1.6160

-1.3798

-1.2468

-1.0469

-0.6160

所有投资组合的MOM因子回归系数均为负值，这表明在过去3个月表现良好的投资组合在未来表现较差，这与传统的正向动量效应相反。此现象在金融市场上被称为“反转效应”，即过去表现好的股票在未来一段时间内表现不佳，而过去表现差的股票则相反。

对于PCA因子来说，进行线性回归，对第一和第二主成分进行回归

投资组合的R^2值:

0.9957

0.9973

0.9953

0.9954

0.9954

回归系数的p-value:

1.0e-43 \*

0.1908 0.1908

0.0000 0.0000

0.7568 0.7568

0.6340 0.6340

0.6515 0.6515

R²值非常接近1，表明PCA因子在统计上能非常好地解释投资组合收益的变化，即模型能够解释大部分收益的波动。

p-value结果表明，对于第二个投资组合，第一和第二主成分在统计上是显著的（p-value接近0），而其他投资组合的PCA因子在统计上不显著（p-value较高）。这可能意味着虽然PCA因子捕捉到了投资组合收益的主要变动方向，但这些变动并非总是与市场中的已知风险因子相关。

比较MOM因子和PCA因子：

输出整体模型的R^2值和MOM因子的p-value

投资组合1的MOM因子回归系数: -1.106727

投资组合2的MOM因子回归系数: 1.264941

投资组合3的MOM因子回归系数: 0.722353

投资组合4的MOM因子回归系数: -0.783120

投资组合5的MOM因子回归系数: -0.106727

整体模型的R^2值: 0.995391

MOM因子的p-value: 0.000000

当MOM因子与PCA因子一起使用时，R²值仍然很高，这表明结合模型同样能很好地解释投资组合收益的变化。对于不同的投资组合，MOM因子的系数有正有负，这可能表明不同投资组合对动量效应的敏感性不同。MOM因子的p-value为0，表明MOM因子对投资组合收益的解释在统计上是显著的，即MOM因子是影响投资组合表现的一个重要因素。这表明MOM因子在投资组合收益的解释中起到了关键作用，可能作为一个独立的预测因子，提供了超出PCA因子的信息。