

实验一 股票投资组合分析

董玉龙 PB14207015

一、实验内容

选取国内外股票各五支，选取投资组合并计算其回报率，Sharpe Ratio。

二、实验条件

1. 国内股票数据的获取通过 python 库 tushare 的 API 获取，国外股票数据通过 pandas_datareader 库的 API 获取；
2. 数据存储采用 pandas 库提供的 dataframe，获取数据后存入 Excel 表格中进行数据分析；
3. 由于做完国内股票分析之后发现 pandas_datareader 需要 python3 的环境，所以除了获取国外股票用的程序 get_data2.py 在 python3 环境下调试外，获取国内股票用的程序 get_data1.py，股票数据分析用的程序 data_processing.py 均在 python2 环境下调试。

三、程序说明

1、调用方法

- 1) python get_data1.py, python get_data2.py
将在运行的地址创建 xlsx 文件存放获取的股票数据，第一列为时间，后五列为收盘价格；
- 2) python data_processing data_file saving_path
 - A. 使用两个命令行参数，第一个是需要分析的 excel 表格的文件名，若不在当前文件夹应添加路径，第二个是分析得到的图片的存储地址，程序运行完成后将输出一个 output.txt 记录分析的数据结果，四张图分别是收盘价格和时间的关系、直方图、Monte Carlo 对有效边界的估计和最优化计算的近似有效边界、有效边界和市场线
 - B. 对于多支股票的投资组合问题，在程序中通过 `N_of_Assets = data.columns.size` 通过删去时间列的数据表自动得到股票数，因此可以对不同数量的投资组合进行分析

2、函数说明

- 1) return, volatility, Sharpe ratio 都是通过定义进行计算，在第一部分输出了各股票年化收益率的均值，协方差矩阵和相关性矩阵，用于分析各股票的相关性；
- 2) rand_vect(number_of_assets)
生成随机投资权重向量，用于 Monte Carlo 模拟
- 3) data_statistic(weights)
计算输入投资组合的 return, volatility, Sharpe ratio

- 4) 最有投资组合的计算
这里采用两种计算方法，最大化 Sharpe ratio 和最小化方差，实现都是通过 scipy 库提供的最优化 API，方法采用均方误差最小，约束取各投资权重为正数且总权重为 1
- 5) 有效边界
通过最小化投资组合的 volatility，输入 return，计算其对应的最小 volatility，并采用样条插值得到近似的图形
- 6) 市场线
假设无风险收益为 0.02，通过样条插值得到的近似有效边界满足的方程解得市场线的直线方程

四、结果分析

1. 国内股票

1) 国内股票选择['000002', '000063', '002415', '000333', '600036']五支

2) 数据结果

```
===== Info of 5 Assets =====
----- mean -----
000002    0.439057
000063    0.193814
000333    0.348079
002415    0.391141
600036    0.267696
dtype: float64
----- covariance matrix -----
           000002    000063    000333    002415    600036
000002  0.186548  0.068046  0.065538  0.035059  0.056510
000063  0.068046  0.186892  0.084845  0.085284  0.045044
000333  0.065538  0.084845  0.145816  0.077597  0.048852
002415  0.035059  0.085284  0.077597  0.158332  0.026881
600036  0.056510  0.045044  0.048852  0.026881  0.085567
----- correlation coefficient matrix -----
           000002    000063    000333    002415    600036
000002  1.000000  0.365212  0.404832  0.208941  0.434998
000063  0.365212  1.000000  0.515818  0.524591  0.353229
000333  0.404832  0.515818  1.000000  0.537444  0.440152
002415  0.208941  0.524591  0.537444  1.000000  0.244330
600036  0.434998  0.353229  0.440152  0.244330  1.000000
===== optimal portfolio wrt Sharpe Ratio =====
0.298942684227    0.0  0.0634232773536    0.351734843542    0.285899194878
----- optimal return, volatility, Sharpe ratio -----
```

```

0.367441135557    0.272025603567    1.35075937977
===== optimal portfolio wrt Minimizing Variance =====
0.0954920784634    0.0384874755462    0.0627933236033    0.240958763398
0.562268358989
----- optimal return, volatility, Sharpe ratio -----
0.316008633661    0.254535579267    1.2415106547
===== market line wrt tangent point =====
0.02  1.27749122857    0.274272584857

```

从表中可以看出，

- A. 各股票之间存在比较强的正相关关系，说明各个产业之间不是完全独立的；
- B. 通过最大化 Sharpe ratio 得到的最佳投资组合为 0.298942684227
0.0 0.0634232773536 0.351734843542 0.285899194878，最大收益为 0.367441135557，波动率为 0.272025603567，Sharpe ratio 为 1.35075937977；
- C. 通过最小化方差得到的最佳投资组合为 0.0954920784634
0.0384874755462 0.0627933236033 0.240958763398
0.562268358989，最大收益为 0.316008633661，波动率为 0.254535579267，Sharpe ratio 为 1.2415106547；
- D. 两种方法得到的投资组合各有优势，从波动率和 Sharpe ratio 来看各有好处；
- E. 切线的切点和最大化 Sharpe ratio 得到的波动率基本一致，说明切点处 Sharpe ratio 最大。

3) 图片结果

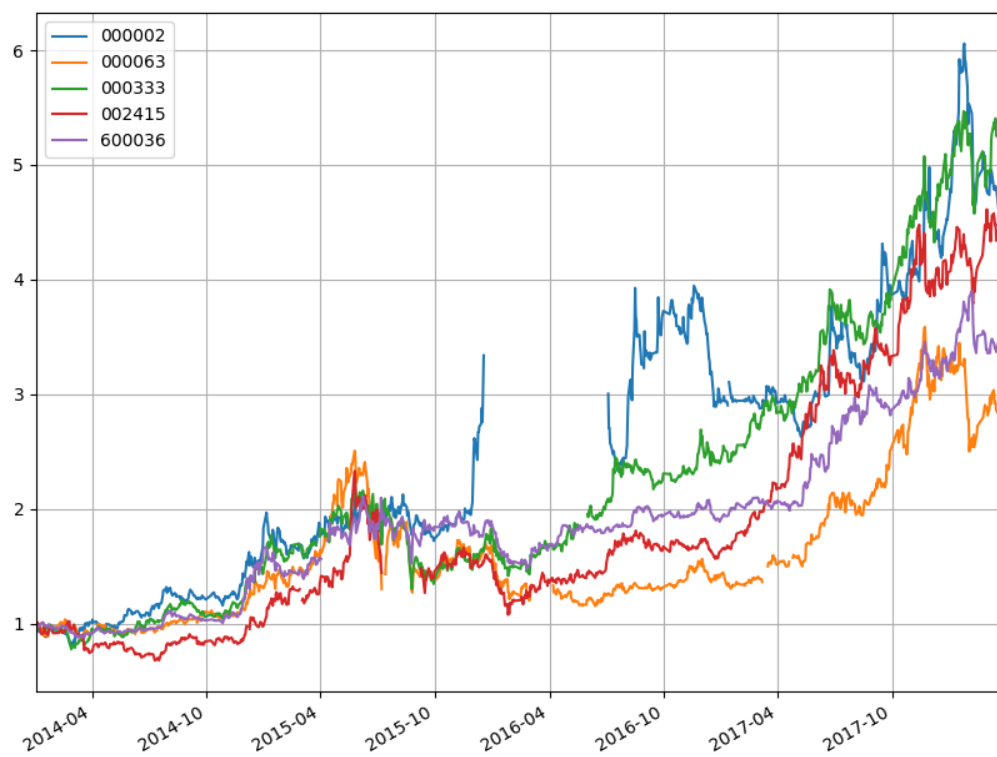


图 1 . 收盘价格随时间变化图

可以看出，收盘价格大体呈现增长趋势

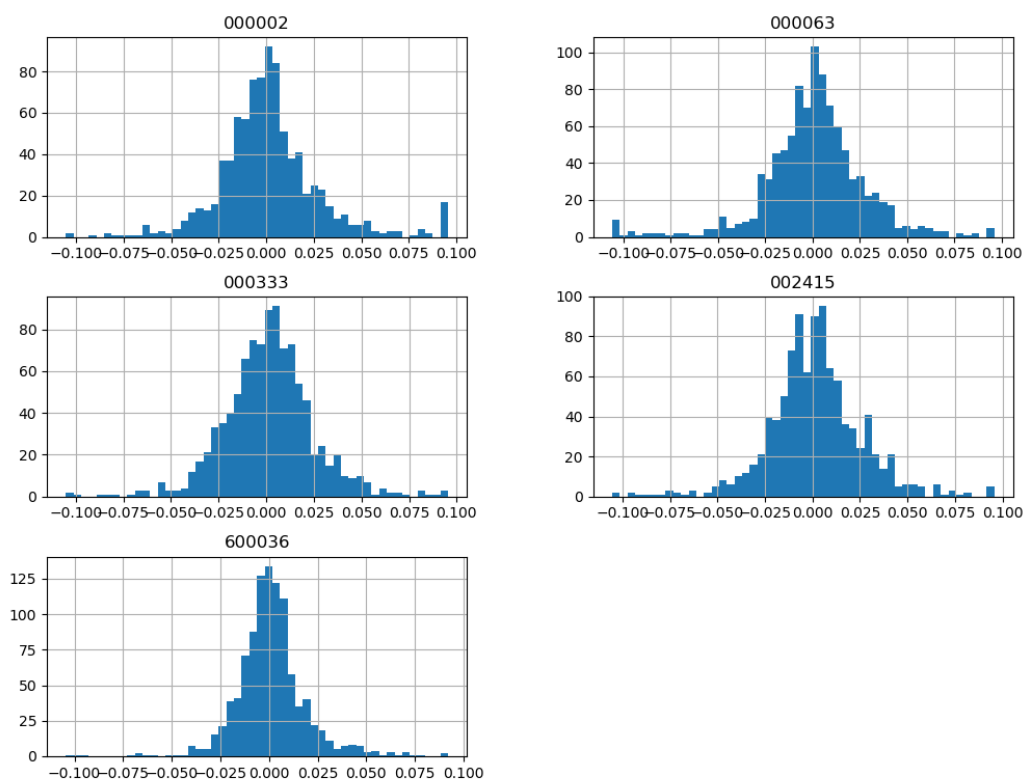


图 2 . 收盘价格直方图
可以看出，600036 的收盘价格峰宽较小，说明波动不太大

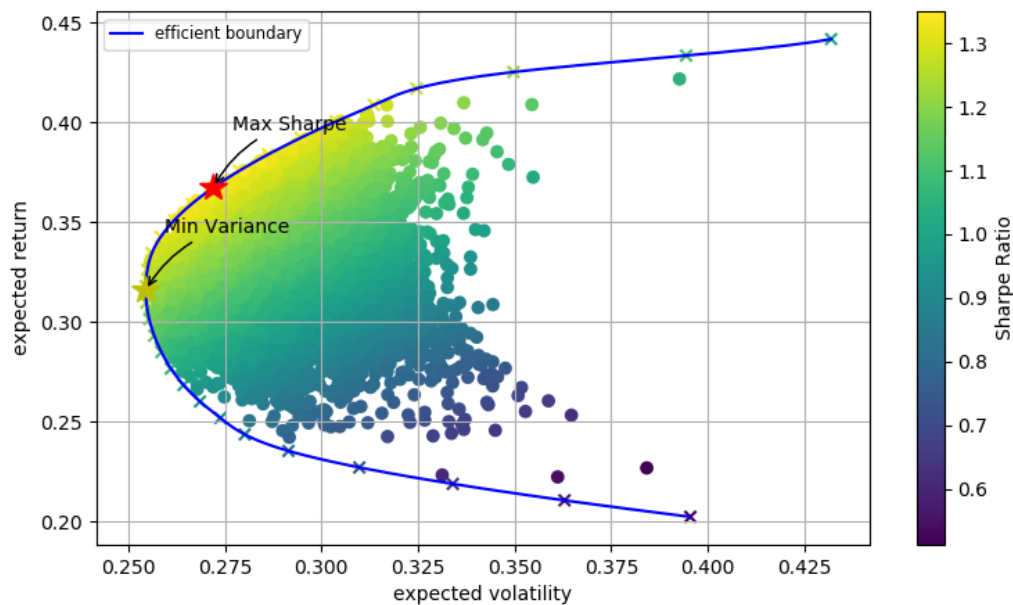


图 3 . Monte Carlo 模拟和最优化计算得到的有效边界
可以看出，模拟结果能较好的给出上边界，但对下边界的估计很不准

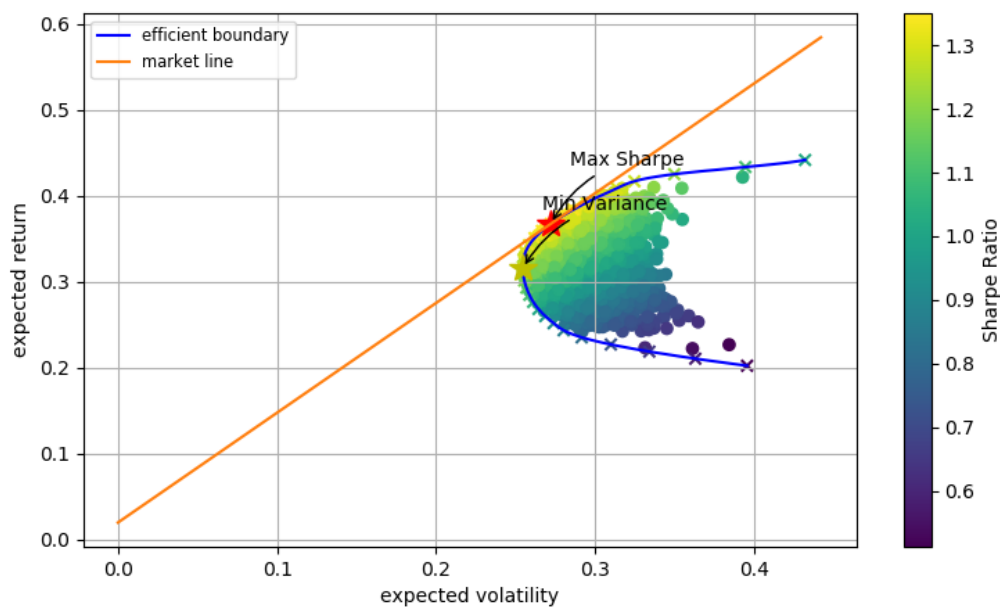


图 4 . 有效边界和市场线
可以看出，市场线与有效边界的切点是最大化 Sharpe ratio 对应的点

2. 国外股票

1) 国外股票选取['AAPL', 'GOOGL', 'KO', 'GE', 'IBM']

2) 数据结果

```
===== Info of 5 Assets =====
----- mean -----
AAPL      0.212649
GE        -0.013873
GOOGL     0.145785
IBM       0.021761
KO        0.053908
dtype: float64
----- covariance matrix -----
          AAPL      GE      GOOGL      IBM      KO
AAPL  0.064263  0.019437  0.026299  0.017782  0.010474
GE    0.019437  0.049217  0.021056  0.021122  0.014297
GOOGL 0.026299  0.021056  0.057532  0.018427  0.011979
IBM   0.017782  0.021122  0.018427  0.036468  0.011745
KO    0.010474  0.014297  0.011979  0.011745  0.021406
----- correlation coefficient matrix -----
          AAPL      GE      GOOGL      IBM      KO
AAPL  1.000000  0.345608  0.432511  0.367316  0.282383
GE    0.345608  1.000000  0.395697  0.498571  0.440470
GOOGL 0.432511  0.395697  1.000000  0.402296  0.341353
IBM   0.367316  0.498571  0.402296  1.000000  0.420371
KO    0.282383  0.440470  0.341353  0.420371  1.000000
===== optimal portfolio wrt Sharpe Ratio =====
0.618711687127    0.0  0.264000485543    6.03358508988e-17  0.117287827331
----- optimal return, volatility, Sharpe ratio -----
0.176378592934    0.199393625481    0.884574883015
===== optimal portfolio wrt Minimizing Variance =====
0.0822366061979    0.0346375559061    0.0687376914307    0.192736235985
0.62165191048
----- optimal return, volatility, Sharpe ratio -----
0.0647340978689    0.133265862766    0.485751538504
===== market line wrt tangent point =====
0.02  0.790459692406    0.225415787557
```

从表中可以看出，

- A. 各股票之间存在比较强的正相关关系，说明各个产业之间不是完全独立的；
- B. 通过最大化 Sharpe ratio 得到的最佳投资组合为 0.618711687127
0.0 0.264000485543 6.03358508988e-17 0.117287827331，最

大收益为 0.176378592934，波动率为 0.272025603567，Sharpe ratio 为 1.35075937977；

- C. 通过最小化方差得到的最佳投资组合为 0.0822366061979
0.0346375559061 0.0687376914307 0.192736235985
0.62165191048，最大收益为 0.0647340978689，波动率为
0.133265862766，Sharpe ratio 为 0.485751538504；
- D. 两种方法得到的投资组合各有优势，从波动率和 Sharpe ratio 来看各有好处；
- E. 相比于国内股票，国外股票的总体收益较低，但波动率有所降低，回报低但风险较小，从 Sharpe ratio 来看，国外股票相对国内股票小，说明投资国内股票性价比比较高；
- F. 切线的切点和最大化 Sharpe ratio 得到的波动率相差不大，说明在误差允许范围内可以认为切点处 Sharpe ratio 最大。

3) 图片结果

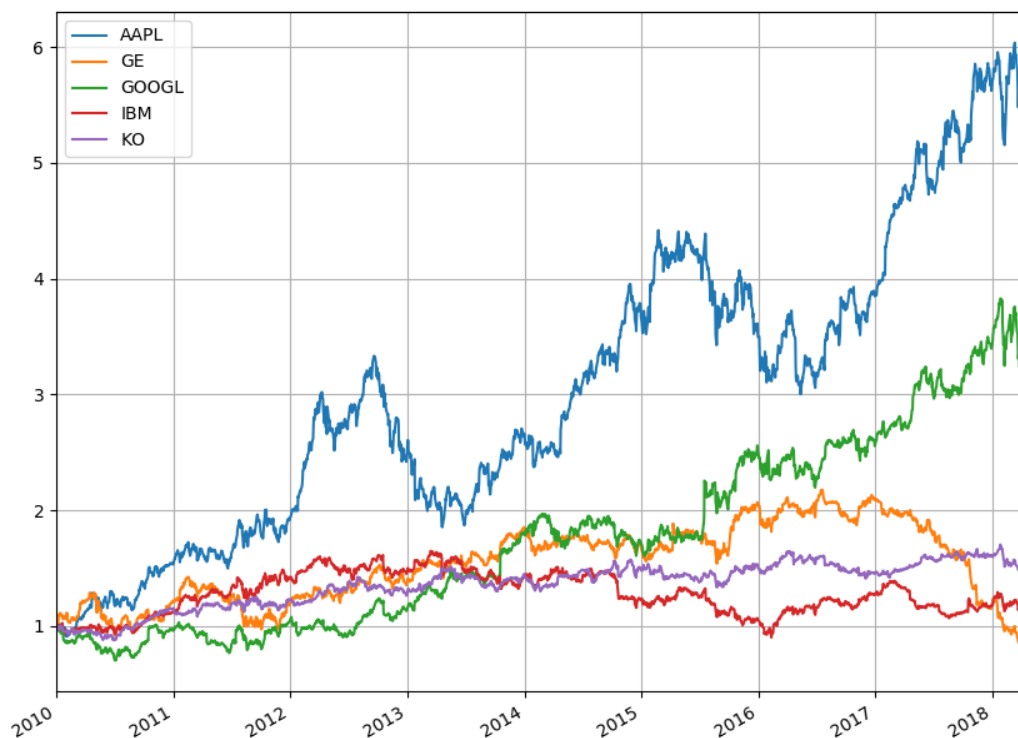


图 5 . 收盘价格随时间变化图

可以看出，除了 AAPL 和 GOOGL 的收盘价格大体呈现增长趋势，其他三支股票基本保持不变

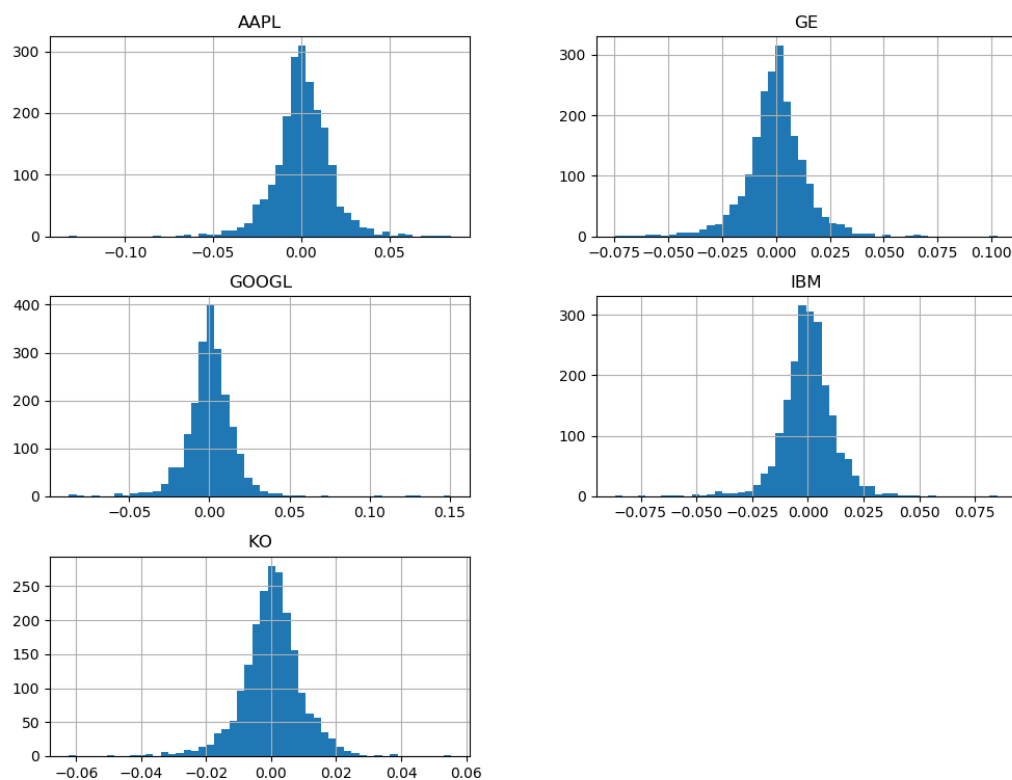


图 6 . 收盘价格直方图

可以看出，各支股票的分布都比较一致，AAPL 和 GOOGL 的峰宽较大

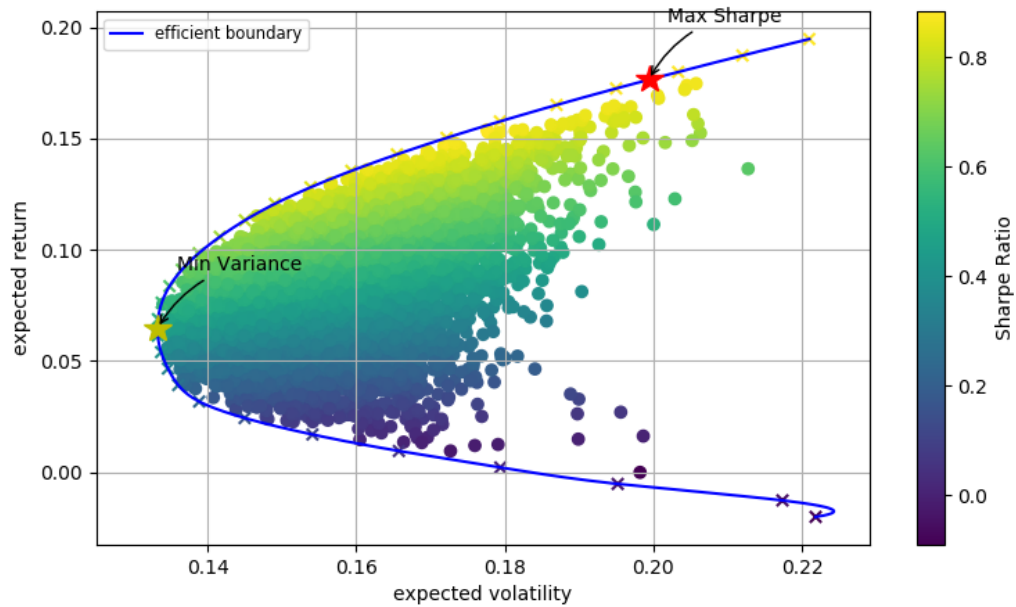


图 7 . Monte Carlo 模拟和最优化计算得到的有效边界
可以看出，模拟结果能较好的给出有效边界，而国内股票的下边界估计不准，相比于国内股票，国外股票通过两种方法得到的投资组合差别较大

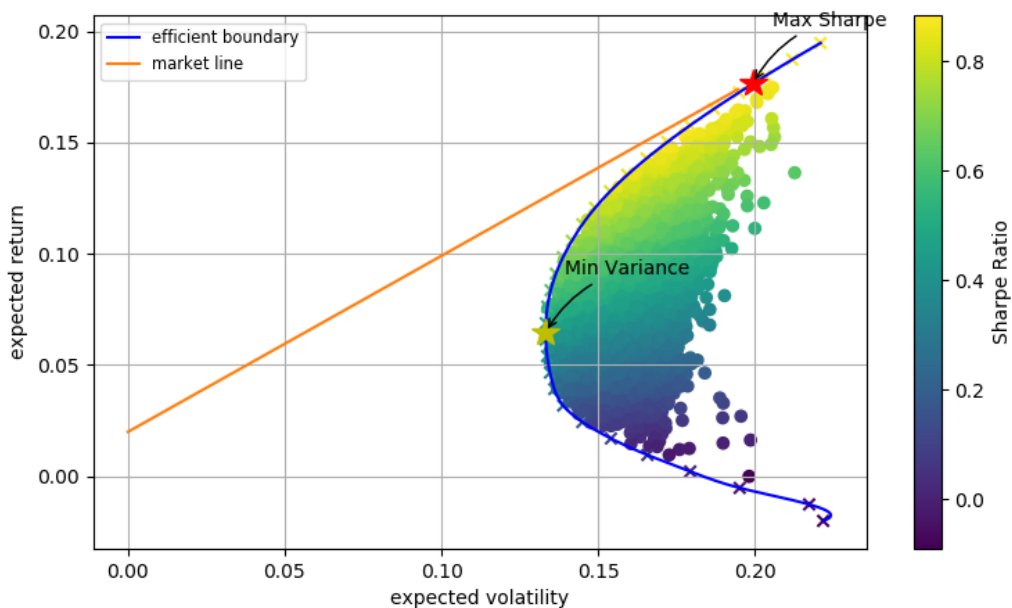


图 8 . 有效边界和市场线
可以看出，市场线与有效边界的切点是最大化 Sharpe ratio 对应的点

五、 总结

1. 各行业股票之间存在一定的相关性；
2. 投资国内股票相对于投资国外股票有较高的性价比；
3. 国外股票用 Monte Carlo 方法对有效边界给出的估计比国内股票准确。