

关于私募产品的最优投资组合模型

摘 要

本文求解最优规划问题，对产品净值数据和同期大盘指数进行分析处理，根据投资者对于风险和收益的要求，确定约束条件，建立相应的优化模型进行求解。

针对问题一：

要求自定规则为 45 只私募产品进行分类，将上证指数与产品净值的变化趋势绘制曲线，再用相关系数进行数据分析，根据相关系数不同区间的特征差异性将 45 只产品分成四类。

针对问题二：

由资金到位时间可知，应从 2016 年 5 月 31 日开始分析各产品净值变化趋势和同期大盘指数，再根据所给条件(年化收益要求、风险偏好以及强制清盘时对收益增长率的限制)，作出约束条件，结合 MPT 模型，建立最优规划模型，用有效前沿理论筛选出符合要求的产品组合。对 2018 年 5 月 31 日后的投资问题，根据历史数据进行了期望回报率以及方差分析进行预测，进而选取最优的产品组合。

针对问题三：

在问题二的约束上增加了新的约束，并对 3-5 号投资者增加了产品数量的约束，改进最优投资组合模型，根据有效前沿理论，可知最佳组合总是分布在有效前沿上，利用 Portfolio 函数设立约束，求解出对应的最佳组合，并在后面利用 checkFeasibility 函数核对投资组合对象的可行性，进行了结果的验证。

本模型本模型建立中运用 MATLAB r2021b，利用 MPT 模型和 Portfolio 函数对投资组合问题进行求解，使建模过程顺利进行，使所建模型更加精简，减少了人为的计算复杂度和计算错误度。

关键词：优化模型、相关系数、MRT 模型、有效前沿、MATLAB

一 问题重述

投资者在面对众多的私募产品时，如何精准挑选符合自己收益期望和风险预期的产品，实现资产增值；高净值投资者追求多策略产品的组合投资，以抵御市场的系统性风险，同时实现财富长期保值增值。根据金融市场中 45 只优质私募产品近三年的净值数据和同期上证指数每日收盘数据，下表 1 是某资产管理公司的意向投资客户及其投资需求。

表 1: 投资需求

意向投资者	资金量(百万元)	投资周期(月)	年化收益要求	风险偏好
1 号投资者	500	24	至少 2 倍同期存款利率	止损线 0.98
2 号投资者	100	24	8%–13%	止损线 0.96, 净值 1.04 以上最大回撤不超过 6%
3 号投资者	50	12	13%–18%	止损线 0.93, 净值 1.03 以上最大回撤不超过 8%
4 号投资者	10	12	18%–23%	止损线 0.9, 净值 1.02 以上最大回撤不超过 10%
5 号投资者	5	12	不低于 20%	止损线 0.85, 最大回撤不超过 15%
6 号投资者	80	12	15% 左右	止损线 0.95, 净值 1.03 以下最大回撤不超过 7%, 净值 1.03 以上最大回撤不超过 10%

本题考察的是利用相关数据，结合数学规划合理设计最优投资组合模型，并基于模型协助公司解决以下产品投资问题：

1. 请根据净值数据与同期大盘指数，自订规则为 45 只私募产品进行分类。
2. 假设 1 号、2 号投资者资金于 2016 年 5 月 31 日到位，请从 45 只私募产品中选取恰当的产品组合，要求：(1) 尽可能满足客户收益与风险要求；(2) 触及止损线时（强制清盘）力求不输于同期大盘涨跌幅。根据投资策略，请在 2018 年 5 月 31 日为 1 号、2 号投资者选取他们第二个投资周期的产品组合。
3. 分别在 2016 年 1 月、2017 年 3 月、2018 年 6 月三个时间点为 3-6 号投资者选取最优投资组合，要求：(1)(2) 同前一问；(3) 其中 3-5 号投资者分别投资不超过 10 只、3 只、2 只产品。

二 问题分析

该问题主要讨论私募产品的最优投资组合，通过对已知净值数据和大盘指数分析得出对其预测结果，并规划出适合投资者的组合。

问题一：

根据净值数据与同期大盘指数表格，对二者时间及数据进行处理将大盘与 45 只私募产品分别进行相关性分析，考虑单个产品，用相关系数反映单个产品与上证指数的联动性。

根据相关性和相关系数量化确定其风险，根据 45 只私募产品走势分析量化确定其收益。然后将风险与收益综合考虑，建立私募产品分类模型，对私募产品进行分类。

问题二：

基于私募产品分类模型，根据投资者资金到位时间，与从所有产品组合构成的收益风险可行区域，选择其最优组合。将收益与标准差关系构建坐标系，运用 MPT 模型，建立最优投资组合模型，在可行区域内选择符合收益要求且风险尽可能小的组合。

由于投资者追求的是风险低并且收益高的资产配置，固定收益率不变，将风险降到最低，也就是要得到最小的横轴值；或者固定风险不变，将收益率提升到最高，也就是要得到最大的竖轴值。

对于在 2018 年 5 月 31 日之后选取的投资组合，需要对产品和大盘进行预测，通过预测的走势。结合每个产品的风险大小，以及多个产品对冲情况，进行投资组合。

问题三：

在问题二条件的基础上，加入对投资产品数量的限制，再根据投资者的要求筛选出符合条件的产品组合。

三 模型假设

1. 投资者对一个资产的风险的认知等同于该资产的收益率变量的标准差。
2. 市场上所有无风险资产的收益率是一样的，叫做无风险利率。
3. 全球无相关政策，影响产品净值以及大盘的因素较小。
4. 市场存在买空和卖空。
5. 不存在通货膨胀，折现率也不变。
6. 所有投资者投资产品只考虑净值与大盘走势，不受其他因素的影响。

四 符号说明

符号表示	文字说明
n	第 n 件产品
σ_i^2	产品 i 的方差
σ_i	产品 i 的标准差
A	大盘数据集
\min_A	A 中最小数据
\max_A	A 中最大数据
a_i	大盘数据
\hat{a}_i	映射后的大盘数据
\bar{a}	大盘数据均值
B	净值数据集
b_i	产品净值
\bar{b}	产品净值均值
$r_{(A,B)}$	A 与 B 的相关系数
r_i	第 i 种资产的收益率
$E[r_i]$	产品 i 收益率的预期
r_f	无风险利率
P	风险资产配置
w_i	每一个资产 i 在 P 中的权重
r_p	P 的收益变量
$E[r_p]$	P 的预期收益
α	比例系数
S_p	夏普比率
μ	资本市场线
$b_{i,j}$	第 i 个投资者选取的第 j 个产品的净值
b_{\max}	在该时刻后产品的最大净值
b_{\min}	在该时刻后产品的最小净值
$x_{i,j}$	产品 j 是否符合第 i 个投资者的要求
N	大盘涨幅

五 模型建立

5.1 模型一 私募产品分类模型

5.1.1 数据预处理

首先对产品净值以及大盘指数进行处理 [1]，观察二者之间的关系。因此先以产品 A 为例建立对产品的分析模型，下图为产品 A 净值与大盘波动图。

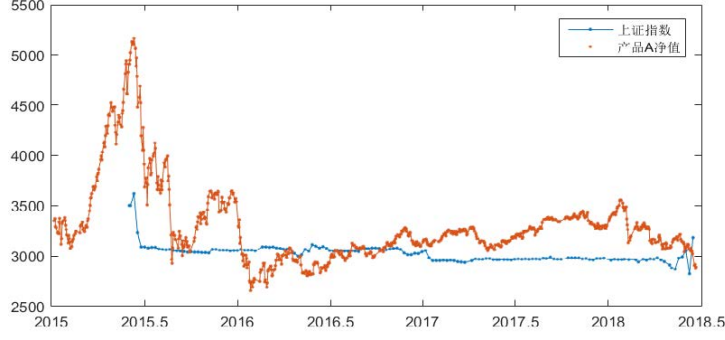


图 1: 产品 A 净值与大盘波动图

由此得出，产品 A 在上证指数浮动时有较大的浮动。观察波动图，通过分析单个产品与整体市场的联动性，就可以衡量单个产品的风险；通过分析单个产品的走势，就可以大致预测单个产品净值走势。

5.1.2 数据规范化

当数据集的值域范围差别较大时，度量的单位也可能影响数据分析的精度。对大盘指数而言，用较小的单位表示数据的值将导致该数据具有较大的权重。为了避免数据分析的结果对度量单位的选择产生依赖性，在进行数据相关性分析之前，应当采用最小-最大规范化方法 [2]，对原始的大盘数据进行线性变换。

设大盘数据集为数据集 $A = \{a_1, a_2, \dots, a_n\}$ ，其最大值和最小值分别为 max_A 和 min_A 对于 A 中任意值 a_i ，可以利用最小-最大规范化方法将其映射成区间 $[min_{new}, max_{new}]$ 之中的 a'_i ：

$$a'_i = \frac{a_i - min_A}{max_A - min_A} (max_{new} - min_{new}) + min_{new} \quad (1)$$

将大盘数据规范化后，原始数据间的联系还能继续保持而不被破坏，同时，各个数据值在进行后续与产品净值数据的相关性计算时将被赋予相等的权重。

5.1.3 数据相关性分析

利用数据标准差度量数据分布的离散程度,对于数据集 $A = \{a_1, a_2, \dots, a_n\}$ ，其标准差为：

$$\sigma_A = \sqrt{\frac{1}{n} \sum_{i=1}^n (a_i - \bar{a})^2} \quad (2)$$

其中 \bar{a} 是 A 中所有数据的均值，利用公式 2 计算 45 个产品和大盘数据的标准差。

以产品 A 为例，设产品 1 的净值数据集为 $B = \{b_1, b_2, \dots, b_n\}$ ，已知大盘数据集 A，A 和 B 的相关系数 $r_{(A,B)} \in [-1, 1]$ 可计算如下：

$$r_{(A,B)} = \frac{\sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b})}{n\sigma_A\sigma_B} \quad (3)$$

若 $r_{(A,B)} > 0$ ，则 A 和 B 之间是正相关的；若 $r_{(A,B)} < 0$ ，则 A 和 B 之间是负相关的；若 $r_{(A,B)} = 0$ ，则表明 A 和 B 是独立的，即两者不存在相关性。

5.1.4 产品分类

$|r_{(A,B)}|$ 的值越大，A 和 B 的相关性越强，以此为基准对产品进行分类。

表 2: 产品分类标准

相关系数绝对值	产品分类
0.0—0.3	微相关产品
0.3—0.5	实相关产品
0.5—0.8	显著相关产品
0.8—1.0	高度相关产品

利用 MATLAB 导入数据对 45 个产品分别计算相关系数，相关系数与股票类别关系如下图：

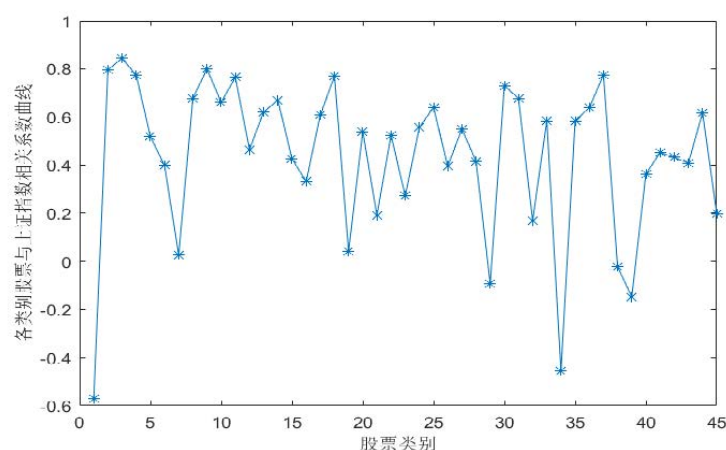


图 2: 相关系数与股票类别关系图

将 45 个产品按照分类标准进行分类：

表 3: 产品相关系数

产品编号	相关系数	产品分类	产品编号	相关系数	产品分类	产品编号	相关系数	产品分类
1	-0.57300	显著相关	16	0.33162	实相关	31	0.67670	显著相关
2	0.79647	显著相关	17	0.60768	显著相关	32	0.16881	微相关
3	0.84391	高度相关	18	0.76964	显著相关	33	0.58442	显著相关
4	0.77427	显著相关	19	0.04214	微相关	34	-0.45620	实相关
5	0.51844	显著相关	20	0.53865	显著相关	35	0.58504	显著相关
6	0.39822	实相关	21	0.19205	微相关	36	0.63875	显著相关
7	0.02516	微相关	22	0.52421	显著相关	37	0.77461	显著相关
8	0.67822	显著相关	23	0.27169	微相关	38	-0.02340	微相关
9	0.79952	显著相关	24	0.55645	显著相关	39	-0.14820	微相关
10	0.66070	显著相关	25	0.64050	显著相关	40	0.36334	实相关
11	0.76486	显著相关	26	0.39506	实相关	41	0.45195	实相关
12	0.46469	实相关	27	0.54971	显著相关	42	0.43204	实相关
13	0.62145	显著相关	28	0.41682	实相关	43	0.40876	实相关
14	0.67008	显著相关	29	-0.09310	微相关	44	0.61709	显著相关
15	0.42679	实相关	30	0.73050	显著相关	45	0.19936	微相关

表 4: 产品分类结果

产品类型	产品
微相关产品	7 19 21 23 29 32 38 39 45
实相关产品	6 12 15 16 26 28 34 40 41 42 43
显著相关产品	1 2 4 5 8 9 10 11 13 14 17 18 20 22 24 25 27 30 31 33 35 36 37 44
高度相关产品	3

5.2 模型二 最优投资组合模型

5.2.1 MPT 模型

根据 MPT 模型 [3], 假设投资者对一个资产的风险的认知等同于该资产的收益率变量的标准差。因此, 任何一个资产或者资产组合都可以根据其收益和风险被放在一个二维坐标图上, 该图的横轴是收益率的标准差, 竖轴是收益率的预期值。为求解风险低并且收益高的资产配置, 可以固定收益率不变, 得到最小横轴值和最大竖轴值。

对于某一资产 $i=1,2,\dots,45$, 用 r_i 表示该资产的收益率的随机变量, $E[r_i]$ 表示收益率的预期, σ_i 表示 r_i 的标准差。其中市场上所有 $\sigma_i > 0$ 的资产叫做风险资产, 收益率没有不确定性的资产叫做无风险资产。并且, 假设市场上所有无风险资产的收益率是一样的, 叫做无风险利率 r_f 。

一个风险资产配置 P 是由风险资产 $i=1,2,\dots,n$ 按照某个权重比例组成的, 每一个资产 i 在 P 中的权重是 $w_i \in R$, 满足 $\sum_{i=1}^n w_i = 1$ 。根据单个资产的收益率, 可以计算资产配置 P 的收益变量的一些性质:

$$\text{资产组合收益率的随机变量: } r_p = \sum_{i=1}^n w_i r_i \quad (4)$$

$$\text{预期收益: } E[r_p] = E\left[\sum_{i=1}^n w_i r_i\right] = \sum_{i=1}^n w_i E[r_i] \quad (5)$$

$$\text{方差: } \sigma^2(r_p) = E[r_p - E[r_p]] = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \text{Cov}(r_i, r_j) \quad (6)$$

(1) 有效前沿 [4]

对于任意一个预期收益值 μ , 找到一个由配置权重 w 定义的资产配置 P (净值数据), 要求 P 的预期收益率为 μ 。并且, 在所有可以配置出的预期收益为 μ 的组合中, P 的方差是最小的。这个问题的最优解可以用拉格朗日乘子法找出, 对于每一个值 μ , 求得一个风险资产配置 P 满足 $E[r_p] = \mu$, 并且 σ_p 是最小的。

(2) 无风险资产 [5]

有效前沿左侧的区域是通过风险资产无法配置出的。但是如果把无风险资产加入资产配置, 那么左侧的一些位置是可以获取的。选择有效前沿上的一个资产配置 P , 并选择比例 $\alpha \geq 0$, 将本金的 α 配置于 P , 并将 $1 - \alpha$ 配置于无风险资产。

若 $\alpha \leq 0$, 即 $1 - \alpha \geq 0$, 意思是将 $1 - \alpha$ 倍的本金存入银行或买入债券, 获取那部分的无风险利率; 若 $\alpha \geq 1$, 即 $1 - \alpha \leq 0$, 意思是贷款本金 $\alpha - 1$ 倍的资金, 支付无风险利率, 并用贷款连同本金一并配置于 P 。如此, 以 α 为系数, 使用 P 和无风险资产配置出一个组合, 将它的收益随机变量记为 r_α 。计算得到:

$$E[r_\alpha] = E[\alpha r_p + (1 - \alpha)r_f] = \alpha E[r_p] + (1 - \alpha)r_f \quad (7)$$

$$\sigma_\alpha = \sqrt{\sigma^2(\alpha r_p + (1 - \alpha)r_f)} = \sqrt{\alpha^2 \sigma^2(r_p)} = \alpha \sigma_p \quad (8)$$

使用常用的直线坡度计算方法, 所得的坡度为夏普比率:

$$S_p = \frac{E[r_p] - r_f}{\sigma_p - \sigma_{r_f}} = \frac{E[r_p] - r_f}{\sigma_p} \quad (9)$$

(3) 市场组合 [6]

经计算可以证明，用市场信息计算得来的有效前沿上必定有一个夏普比率最高的点，将其叫做市场组合 M ，其夏普比率为 S_M 。一个配置的夏普比率等同于连接无风险资产和这个配置的直线的坡度，那么市场组合的夏普比率最高，就意味着它连接无风险资产的线坡度最陡。因此，这条线必定与有效前沿构成切线关系。这条线为资本市场线，其公式为：

$$\mu = r_f + S_M \sigma \quad (10)$$

资本市场线的意义在于，固定标准差，那么市场上收益预期最高的投资组合在这条线上；或者，固定预期收益，那么市场上标准差最低的投资组合在这条线上。所以，资本配置线可以直观地理解为理论上的“最佳配置线”。

资本市场线和有效前沿相切在夏普比率最高的资产配置上，我们将这个组合称作 M 。该新组合的预期收益率高于它的几个单一产品的收益，但风险比其中的每一支都低。

5.2.2 利用 MPT 模型求解

$$\text{目标函数} \begin{cases} \min \sigma^2(r_p) = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \text{Cov}(r_i, r_j) \\ \max E(r_p) = \sum_{i=1}^n w_i r_i \end{cases} \quad (11)$$

由 2016-2018 年央行同期存款两年定期年利率均为 2.1%，1 号投资者的年化收益与风险偏好约束条件为：

$$\begin{cases} r_1 \geq 2.1\% \\ b_{1,j} + d_1^- - d_1^+ = 0.98 \\ n \geq N(b_{1,j} = 0.98) \end{cases} \quad (12)$$

同理 2 号投资者的约束条件：

$$\begin{cases} 8\% \leq r_2 \leq 13\% \\ b_{2,j} + d_2^- - d_2^+ = 0.96 \\ b_{i,j} - b_{min} \leq 6\%(b_{i,j} \geq 1.04) \\ n \geq N(b_{2,j} = 0.96) \end{cases} \quad (13)$$

利用 MATLAB 进行规划求解：

(1) 单一产品选择

风险偏好的单一产品：

表 5: 单一资产选择

	2016.5.31	2018.5.31
1 号投资者	4 8 10 15 19 20 22 23 28 33 36 40 44 45	4 8 10 15 19 20 22 23 28 33 36 40 44 45
2 号投资者	15 19 40	15

(2) 组合选择

被淘汰掉的产品权重为 0，在满足风险偏好的单一产品中进行组合选择，重新生成有效前沿以及收益要求区间。如图 3，要求组合资产的选择，即求解投资者买入若干个产品的最佳投资组合，其组合整体满足每个投资者的年化收益与风险偏好。

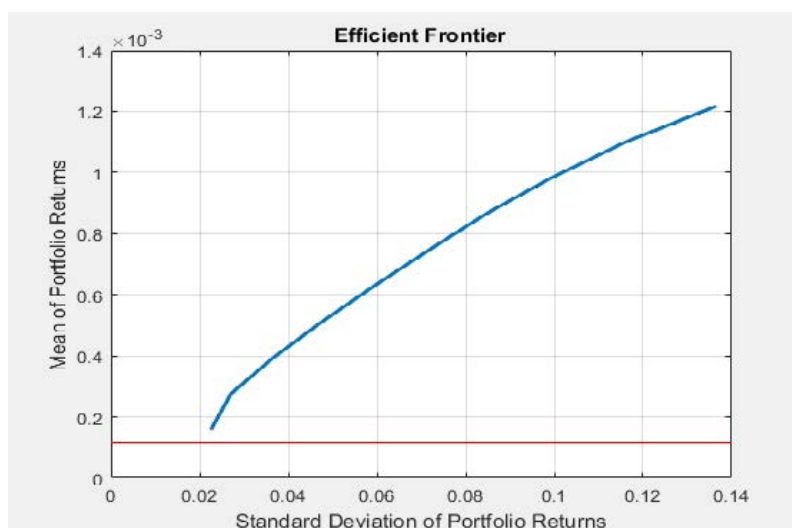


图 3: 满足客户 1 的所有组合的有效前沿

可以看出，满足客户收益与风险要求的投资组合为红线以上蓝线以下的范围。

(3) 组合资产选择

随机生成任意符合收益要求的 5 个组合，任一组合均满足客户的要求。

1 号投资者组合资产选择：

表 6: 组合资产选择

组合	产品	投资资金 (百万元)	风险系数	预测每日盈利 (百万元)
1	4	11.75	0.0226	0.0798
	19	172.65		
	22	62.35		
	23	211.90		
	28	41.35		
2	8	54.65	0.0393	0.2120
	19	145.20		
	23	139.15		
	28	53.35		
	45	107.60		
3	8	143.35	0.0656	0.3442
	19	101.80		
	28	46.80		
	45	208.05		
4	8	169.30	0.0954	0.4764
	45	330.70		
5	45	500.00	0.1366	0.6000

若考虑最大损失，1 号投资者最大损失约为 10 百万元。

2 号投资者组合资产选择：

选择产品 15 进行投资 1 百万元，则预计盈利大约为 37.41153 百万元，最大损失为 4 百万元。

5.3 模型三 含产品数的最优投资组合模型

在 2016 年 1 月、2017 年 3 月、2018 年 6 月三个时间点为 3-6 号投资者选取最优投资组合。利用模型二，在每个投资者年化收益与风险偏好约束条件下，进行规划求解。

5.3.1 目标函数和约束条件

目标函数：

$$\begin{cases} \min \sigma^2(r_p) = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \text{Cov}(r_i, r_j) \\ \max E(r_p) = \sum_{i=1}^n w_i r_i \end{cases} \quad (14)$$

约束条件:

$$\left\{ \begin{array}{l} x_{i,j} = \begin{cases} 1, (\text{产品 } j \text{ 符合第 } i \text{ 个投资者的要求}) \\ 0, (\text{产品 } j \text{ 不符合第 } i \text{ 个投资者的要求}) \end{cases}, (i = 3, 4, 5; j = 1, 2, \dots, 45) \\ \sum_{j=1}^{45} x_{3,j} \leq 10 \\ \sum_{j=1}^{45} x_{4,j} \leq 3 \\ \sum_{j=1}^{45} x_{5,j} \leq 2 \end{array} \right. \quad (15)$$

$$\text{3 号投资者} \left\{ \begin{array}{l} 13\% \leq r_3 \leq 18\% \\ b_{3,j} + d_3^- - d_3^+ = 0.93 \\ b_{i,j} - b_{min} \leq 8\% (b_{i,j} \geq 1.03) \\ n \geq N(b_{3,j} = 0.93) \end{array} \right. \quad (16)$$

$$\text{4 号投资者} \left\{ \begin{array}{l} 18\% \leq r_4 \leq 23\% \\ b_{4,j} + d_4^- - d_4^+ = 0.9 \\ b_{i,j} - b_{min} \leq 10\% (b_{i,j} \geq 1.02) \\ n \geq N(b_{4,j} = 0.9) \end{array} \right. \quad (17)$$

$$\text{5 号投资者} \left\{ \begin{array}{l} r_5 \geq 20\% \\ b_{5,j} + d_5^- - d_5^+ = 0.85 \\ b_{i,j} - b_{min} \leq 15\% \\ n \geq N(b_{5,j} = 0.85) \end{array} \right. \quad (18)$$

$$\text{6 号投资者} \left\{ \begin{array}{l} r_6 + d_7^- - d_7^+ = 15\% \\ b_{6,j} + d_6^- - d_6^+ = 0.95 \\ b_{max} - b_{i,j} \leq 7\% \\ b_{i,j} \leq 1.03 \\ b_{i,j} - b_{min} \leq 10\% (b_{i,j} \geq 1.03) \\ n \geq N(b_{6,j} = 0.95) \end{array} \right. \quad (19)$$

5.3.2 求解结果

对于 2016 年 1 月，2017 年 3 月，根据程序产生的结果，45 个产品中单一的产品均不满足 4 个投资者的风险偏好。因此在这里需要将产品进行投资组合，这里暂时不考虑要完全满足投资者的风险偏好，而是在考虑更低风险以及更好收益同时，使得风险满足于投资者的偏好。

2016 年 1 月的投资组合：

3 号投资者在有效前沿生成图上的有效区间为 $[0.13, 0.18]$ ，其最大日化收益为 $0.18/365$ ，利用 MATLAB 计算得到最佳投资组合在有效前沿上的位置如下图所示。

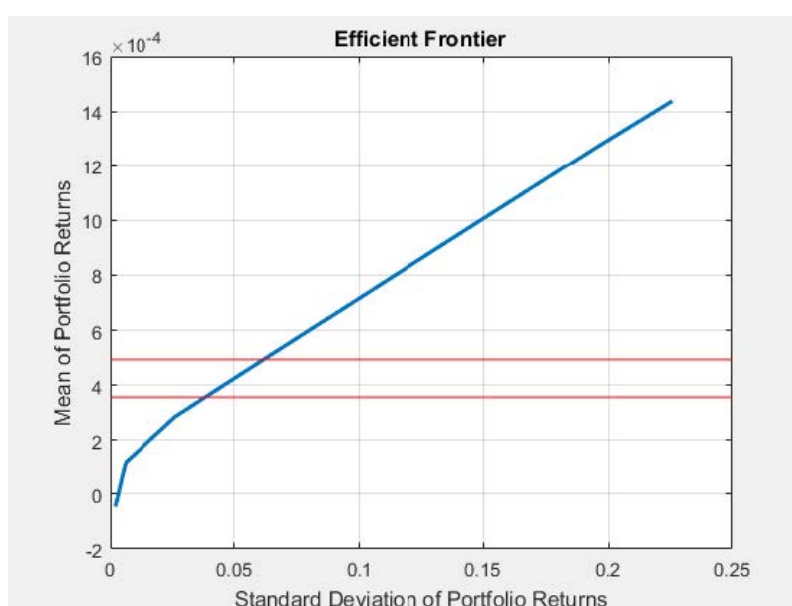


图 4: 3 号投资者在有效前沿生成图上的有效区间

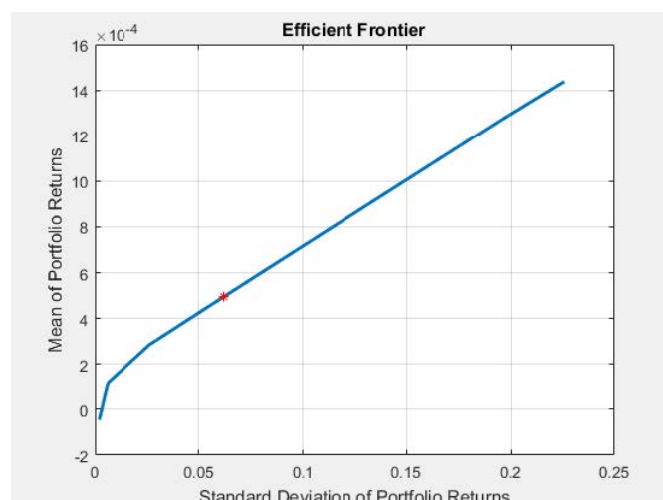


图 5: 3 号投资者最佳投资组合在有效前沿上的位置

由此生成 3-6 号投资者的最佳投资组合如下图所示：

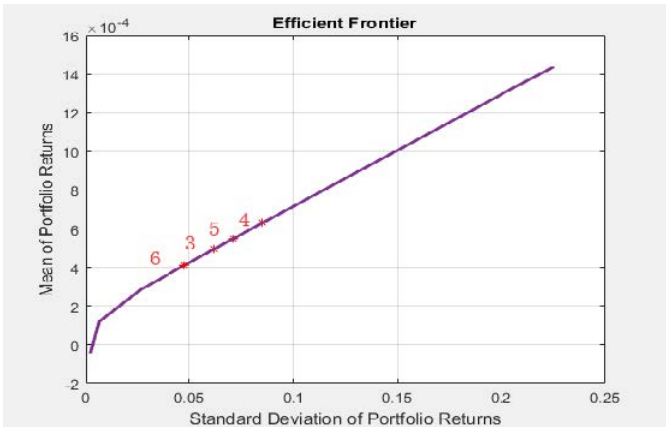


图 6: 3-6 号投资者的最佳投资组合

对应的投资组合分别是：

表 7: 投资者 2016 年 1 月的投资组合表

投资者	产品			风险系数	回报率
	20	28	44		
3 号	0.732671	0.208708	0.058621	0.061660193	4.93E-04
4 号	0.607124	0.314488	0.078387	0.085192408	6.30E-04
5 号	0.682452	0.251020	0.066528	0.071064966	5.48E-04
6 号	0.807999	0.145240	0.046761	0.047589352	4.11E-04

2017 年 3 月的投资组合：

根据程序产生的结果，45 个产品中单一的产品 10、33 号满足客户 3 的风险偏好，14 号产品满足客户 4 的风险偏好要求，没有单一的产品满足客户 5、6 的风险偏好要求。则在满足单一风险偏好下，其余不满足的作不考虑要完全满足投资者的风险偏处理好处理。则我们可以给予满足风险偏好的产品较大的权重 (设为 0.4)，使得组合在风险偏好上可以比较符合。

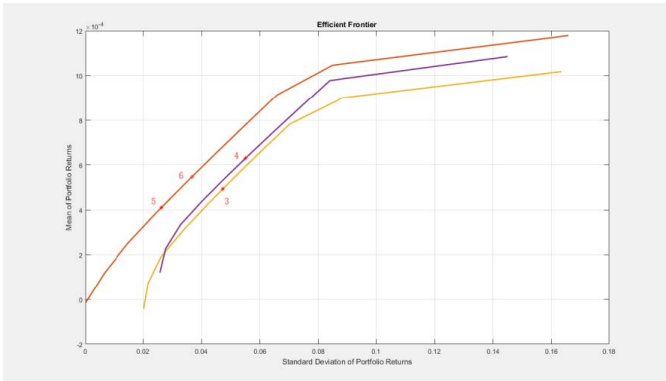


图 7: 不同投资者不同时间资产选择

表 8: 投资者 2017 年 3 月的投资组合表

投资者	产品	权重	风险系数	回报率
3 号	6	0.311265	0.047223260	4.93E-04
	12	0.250502		
	33	0.400000		
	38	0.038233		
4 号	6	0.333265	0.055047243	6.30E-04
	12	0.266735		
	14	0.400000		
5 号	5	0.020703	0.026252947	4.11E-04
	6	0.307454		
	12	0.330291		
	15	0.269489		
	38	0.082062		
6 号	5	0.036817	0.036724349	5.48E-04
	6	0.305945		
	12	0.466830		
	15	0.190369		

2018 年 6 月的投资组合：

由于产品市场的不可控性，本模型并不能有效计算出未来的投资组合。

六 模型评价

本模型用的是 MPT 模型，在可行区域内选择符合收益要求且风险尽可能小的组合，一个资产组合的预期收益是它的成分资产的线性加权平均，但组合的波动性和风险却更复杂，是由成分资产两两之间的相关性决定的。在 MPT 模型的分析中，通过合理地配置资产，可以在保证高收益的同时也降低风险。理论上，最优的风险组合就是夏普比率最高的市场组合 [7]。

本模型是规划私募产品最优投资组合问题，通过对已知净值数据和大盘指数分析得出对其预测结果，规划出适合投资者的组合。本模型可用于对股票、基金走势的预测，对投资者进行方向性分析，方便了各种不同应用对象的建模与分析，进而可以对一些复杂问题做出合理的判断决策，做出有效的预测和估计。

当然，本模型太过于理想，模型一没有考虑上证指数本身的波动性，只考虑了产品净值与上证指数波动性的关系，因此本模型距离现实生活中的应用还有一些差距。

参考文献

- [1] 李荣. 关于数据挖掘技术在经济统计中的应用 [J]. 中国商贸.2015(2):173- 175.
- [2] 李双琦. 上证股票市场与基金市场相关性实证研究 [J]. 商.2015(22):169- 171.
- [3] 爱德华. 技术指标分析. 股市趋势技术分析 [M].1997,(3):45-48.
- [4] 庾建设, 云天铨, 郭志明. 计算证券理论 [M]. 北京: 科学出版社,2008:176- 181.
- [5] 邓成成. 我国开放式基金收益率与上证指数的相关性分析—以嘉实增长开放式基金为例 [J]. 时代金融,2012(20):35-36.
- [6] 刘素凤. 我国基金市场与股票市场关系的实证分析 [J]. 经济师,2007(10):97 98.
- [7] 马永开, 唐小我. 不允许卖空的证券组合选择模型研究 [J]. 预测.1999,18(2) :49-52

附录

以下代码均用 MATLAB 编写。

6.1 模型一 q1.m

```
close all
clear
clc
[NUM,TXT]=xlsread('附件1.xlsx',1);
TXT=TXT(3:end,:);
x2=(2015.9178:0.01:2018.3014)';
YY=[];
for i=1:45
    x1=[];
    y1=[];
    if i<26
        t=TXT(:,2*(i-1)+1);
        shuju=TXT(:,2*i);
        for j=1:length(t)
            riqi=t{j};
            if ~isempty(riqi)
                nian=str2num(riqi(1:4));
                yue=str2num(riqi(6:7));
                ri=str2num(riqi(end-1:end));
                x1(j,1)=nian+data2days(nian,yue,ri)/365;
                y1(j,1)=str2num(shuju{j});
            end
        end
        y2=pchip(x1,y1,x2);
        YY(:,i)=y2;
    elseif i==26
        t=TXT(:,2*(i-1)+1);
        shuju=NUM(:,2*i-8);
        t{1}='20150603';
        for j=1:length(t)
            riqi=t{j};
            if ~isempty(riqi) && ~isempty(shuju(j))
                nian=str2num(riqi(1:4));
                yue=str2num(riqi(5:6));
                ri=str2num(riqi(end-1:end));
                x1(j,1)=nian+data2days(nian,yue,ri)/365;
                y1(j,1)=shuju(j);
            end
        end
    end
end
```

```

y2=pchip(x1,y1,x2);
YY(:,i)=y2;
else
    t=TXT(:,2*(i-1)+1);
    shuju=NUM(:,2*i-8);
    for j=1:length(t)
        riqi=t{j};
        if ~isempty(riqi)
            nian=str2num(riqi(1:4));
            yue=str2num(riqi(5:6));
            ri=str2num(riqi(end-1:end));
            x1(j,1)= nian+data2days(nian,yue,ri)/365;
            y1(j,1)=shuju(j);
        end
    end
    y2=pchip(x1,y1,x2);
    YY(:,i)=y2;
end
end
[ NUM,TXT]=xlsread('附件2处理_日期.xlsx',1);
t1=NUM(:,1);
y1=NUM(:,5);
Y=pchip(t1,y1,x2);
for i=1:45
    c=corrcoef(Y,YY(:,i));
    C(i,1)=c(1,2);
end
figure
plot(C,'*-'); xlabel('股票类别')
ylabel('与上证指数相关系数')
ylabel('各类别股票与上证指数相关系数曲线')

```

6.2 模型二 q2.m

```

clear all;
close all;
jingzhi='基金净值';
dapanzhishu='大盘指数';
all1=xlsread(jingzhi);
all2=xlsread(dapanzhishu);
begintime=516; %2016.5.31
endtime=1216; %2018.5.31
format long
for i=1:45

```

```

    time1(:,i)=all1(:,5*i-1); %四十五个产品的时间数据
    data1(:,i)=all1(:,5*i); %四十五个产品的净值数据
end
time2=all2(:,4);
shoupan=all2(:,8);
fangcha=[];
shouyilv=zeros(743,45)*NaN;
duiying=zeros(743,45)*NaN;
%数据插值 补充缺失的数据
time=beginTime:endTime;
shoupanall=interp1(time2,shoupan,time,'spline');
for i=1:45
    findend=find(isnan(data1(:,i)));
    if isempty(findend)
        findend=size(data1(:,i));
    end
    datatem=interp1(time1(1:findend(1)-1,i),data1(1:findend(1)-1,i),time,'spline');
    data(:,i)=datatem';
end
%数据标准化 计算方差作为风险系数 计算收益率
normalized_data=normalize(data);
shouyilv(1,:)=0;
for i=1:45
    fangcha(i)=nanvar(normalized_data(:,i));
    biaoZhuncha(i)=sqrt(fangcha(i));
    p=find(isnan(data(:,i)));
    if isempty(p)
        p=size(data(:,i));
    end
    for j=2:p(1)-1
        shouyilv(j,i)=(data(j,i)-data(j-1,i))/data(j-1,i);
    end
    pinjunshouyilv(i)=nanmean(shouyilv(:,i)); %期望收益
end
normalized_shoupan=normalize(shoupanall);
dapanfangcha=var(normalized_shoupan);
dapanshouyilv(1,:)=0;
for j=2:size(shoupanall)
    dapanshouyilv(j,:)=(shoupanall(j)-shoupanall(j-1))/shoupanall(j-1);
end
%计算协方差矩阵
xiefangcha=cov(data);
%模型求解
%投资人 1 和 2 的投资组合

```

```

%%基于已知数据的单一资产的选择（利用已知数据）
% nianshouyilv=(data(end,:)-data(1,:))./data(1,:)/2; %两年收益率/2=年收益率
% findzhisun1=[];
% findzhisun2=[];
% for i=1:45
%     if isempty(find(data(:,i)<=0.98))==0
%         findzhisun1=[findzhisun1 i];
%     end
%     if isempty(find(data(:,i)<=0.96))==0
%         findzhisun2=[findzhisun2 i];
%     continue
%     else if isempty(find(data(:,i)>=1.04))==0
%         minjingzhi=1.04/1.06;
%         if isempty(find(data(:,i)<=minjingzhi))==0
%             findzhisun2=[findzhisun2 i];
%         end
%     end
% end
% end
% findshouyi1=find(nianshouyilv<=0.042);
% findshouyi2=find(nianshouyilv<=0.08 | nianshouyilv>=0.13)
% taotai1=union(findshouyi1, findzhisun1);
% taotai2=union(findshouyi2, findzhisun2);
% disp('客户 1 的单一资产的选择:');
% choose1=setdiff(1:45, taotai1)
% disp('客户 2 的单一资产的选择:');
% choose2=setdiff(1:45, taotai2)
%基于历史数据对未来投资的选择（利用整体平均值和方差）
pinshouyilv=pinjunshouyilv.*365; %平均年收益率
findzhisun1=[];
findzhisun2=[];
for i=1:45
    if isempty(find(data(:,i)<=0.98))==0
        findzhisun1=[findzhisun1 i];
    end
    if isempty(find(data(:,i)<=0.96))==0
        findzhisun2=[findzhisun2 i];
    elseif isempty(find(data(:,i)>=1.04))==0
        minjingzhi=1.04/1.06;
        if isempty(find(data(:,i)<=minjingzhi))==0
            findzhisun2=[findzhisun2 i];
        end
    end
end
end
end

```

```

findshouyi1=find(pjnshouyilv<=0.042);
findshouyi2=find(pjnshouyilv<=0.08 | pjnshouyilv>=0.13);
taotai1=union(findshouyi1,findzhisun1);
taotai2=union(findshouyi2,findzhisun2);
disp('客户1的单一资产的选择:');choose1=setdiff(1:45,taotai1)
disp('客户2的单一资产的选择:');choose2=setdiff(1:45,taotai2);
%多种组合选择
%所有投资组合的有效前沿figure(1)
p = Portfolio('assetmean',pinjunshouyilv,'assetcovar',xiefangcha,'lowerbudget',1,'
    upperbudget',1,'lowerbound',0); plotFrontier(p)
%满足客户1的区间以及有效前沿
beginshouyi=0.042/365;
endshouyi=1;
W(choose1)=1;x=0:0.001:0.14;
y=beginshouyi*ones(1,size(x,2)); figure(2)
p = Portfolio('assetmean',pinjunshouyilv,'assetcovar',xiefangcha,'lowerbudget',1,'
    upperbudget',1,'lowerbound',0);
p = setEquality(p,W,1); %被淘汰掉的产品将不给予组合机会 剩余产品权重和为1
bestwgt=estimateFrontier(p,5); %生成组合
[brsk,bret]=estimatePortMoments(p,bestwgt) %计算风险brsk与回报率bret
xlswrite('baocun.xls',[bestwgt;brsk;bret]);
plotFrontier(p)
hold on
plot(x,y,'r-');

```

6.3 模型三 q3.m

```

clear all;
close all;
jingzhi='基金净值';
dapanzhishu='大盘指数';
all1=xlsread(jingzhi);
all2=xlsread(dapanzhishu);
 begintime=790; %2016.1.1-365 2017.3.1-790
 endtime=1155; %2017.1.1-731 2018.3.1-1155
format long
for i=1:45
    time1(:,i)=all1(:,5*i-1); %四十五个产品的时间数据
    data1(:,i)=all1(:,5*i); %四十五个产品的净值数据
end
time2=all2(:,4);
shoupan=all2(:,8);
fangcha=[];

```

```

shouyilv=zeros(743,45)*NaN;
duiying=zeros(743,45)*NaN;
%数据插值 补充缺失的数据
time=beginTime:endTime;
shoupanall=interp1(time2,shoupan,time,'spline');
for i=1:45
    findend=find(isnan(data1(:,i)));
    if isempty(findend)
        findend=size(data1(:,i));
    end
    datatem=interp1(time1(1:findend(1)-1,i),data1(1:findend(1)-1,i),time,'spline');
    data(:,i)=datatem';
end
%数据标准化 计算方差作为风险系数 计算收益率
normalized_data=normalize(data);
shouyilv(1,:)=0;
for i=1:45
    fangcha(i)=nanvar(normalized_data(:,i));
    biaoZhuncha(i)=sqrt(fangcha(i));
    p=find(isnan(data(:,i)));
    if isempty(p)
        p=size(data(:,i));
    end
    for j=2:p(1)-1
        shouyilv(j,i)=(data(j,i)-data(j-1,i))/data(j-1,i);
    end
    pinjunshouyilv(i)=nanmean(shouyilv(:,i)); %期望收益
end
normalized_shoupan=normalize(shoupanall);
dapanfangcha=var(normalized_shoupan);
dapanshouyilv(1,:)=0;
for j=2:size(shoupanall)
    dapanshouyilv(j,:)=(shoupanall(j)-shoupanall(j-1))/shoupanall(j-1);
end
%计算协方差矩阵
xiefangcha=cov(data);
%模型求解
%%投资人 3~6 的投资组合 2016.1 2017.3
%3~6
%基于已知数据的单一资产的选择（利用已知数据）
nianshouyilv=(data(end,:)-data(1,:))./data(1,:)/2; %两年收益率/2=年收益率
findzhisun3=[];
findzhisun4=[];
findzhisun5=[];

```

```

findzhisun6=[];
for i=1:45
    %客户 3
    if isempty(find(data(:,i)<=0.93))==0
        findzhisun3=[findzhisun3 i];
    else if isempty(find(data(:,i)>=1.03))==0
        minjingzhi=1.03/1.08;
        if isempty(find(data(:,i)<=minjingzhi))==0
            findzhisun3=[findzhisun3 i];
        end
    end
end
%客户 4
if isempty(find(data(:,i)<=0.9))==0
    findzhisun4=[findzhisun4 i];
else if isempty(find(data(:,i)>=1.02))==0
    minjingzhi=1.04/1.10;
    if isempty(find(data(:,i)<=minjingzhi))==0
        findzhisun4=[findzhisun4 i];
    end
end
%客户 5
if isempty(find(data(:,i)<=0.85))==0
    findzhisun5=[findzhisun5 i];
else
    minjingzhi=max(data(:,i))/1.15;
    if isempty(find(data(:,i)<=minjingzhi))==0
        findzhisun5=[findzhisun5 i];
    end
end
%客户 6
if isempty(find(data(:,i)<=0.95))==0
    findzhisun6=[findzhisun6 i];
else
    maxjingzhi=1.03*1.10;
    minjingzhi=1.03/1.07;
    if isempty(find(data(:,i)<=minjingzhi))==0|isempty(find(data(:,i)>=maxjingzhi))==0
        findzhisun6=[findzhisun6 i];
    end
end
end
findshouyi3=find(nianshouyilv<=0.13 | nianshouyilv>=0.18);

```

```

findshouyi4=find(nianshouyilv<=0.18 | nianshouyilv>=0.23);
findshouyi5=find(nianshouyilv<0.20);
findshouyi6=find(nianshouyilv>=0.175 | nianshouyilv <=0.125);
taotai3=union(findshouyi3, findzhisun3);
taotai4=union(findshouyi4, findzhisun4);
taotai5=union(findshouyi5, findzhisun5);
taotai6=union(findshouyi6, findzhisun6);
disp('客户3的单一资产的选择:');choose3=setdiff(1:45, taotai3);
disp('客户4的单一资产的选择:');choose4=setdiff(1:45, taotai4);
disp('客户5的单一资产的选择:');choose5=setdiff(1:45, taotai5);
disp('客户6的单一资产的选择:');choose6=setdiff(1:45, taotai6);
%多种组合选择
%所有投资组合的有效前沿figure(1)
p = Portfolio('assetmean', pinjunshouyilv, 'assetcovar', xiefangcha, 'lowerbudget', 1, 'upperbudget', 1, 'lowerbound', 0);
plotFrontier(p)
%满足客户3的区间以及有效前沿
beginshouyi=0.13/365; endshouyi=0.18/365;
W=ones(1,45); %每个产品都可以加权
x=0:0.001:0.25;
y1=beginshouyi*ones(1,size(x,2));
y2=endshouyi*ones(1,size(x,2)); figure(2)
p = Portfolio('assetmean', pinjunshouyilv, 'assetcovar', xiefangcha, 'lowerbudget', 1, 'upperbudget', 1, 'lowerbound', 0);
% W=ones(1,45);
% p = setEquality(p,W,1); %所有产品加权和为1
%5号投资者
p = setDefaultConstraints(p); %所有产品加权和为1
pwgt = estimateFrontierByReturn(p,0.20/365) %生成收益最大的组合
[brsk, bret]=estimatePortMoments(p,pwgt) %计算风险 brsk 与回报率
bret plotFrontier(p)
hold on plot(brsk, bret, '*r'); hold on
%6号投资者
pwgt = estimateFrontierByReturn(p,0.15/365) %生成收益最大的组合
[brsk, bret]=estimatePortMoments(p,pwgt) %计算风险 brsk 与回报率
bret plotFrontier(p)
hold on plot(brsk, bret, '*r'); hold on
%3号投资者
W=zeros(1,45);
W([10 33])=1;
p = setGroups(p,W,0.4,1); %产品10, 33加权和大于40%
pwgt = estimateFrontierByReturn(p,0.18/365) %生成收益最大的组合
[brsk, bret]=estimatePortMoments(p,pwgt) %计算风险 brsk 与回报率 bret
%xlswrite('baocun.xls',[bestwgt;brsk;bret]); %保存数据

```



```

plotFrontier (p)
hold on plot(brsk, bret, '*r'); hold on
%4号投资者
W=zeros(1,45);
W([14])=1;
p = setGroups(p,W,0.4,1); %产品 14 权重大于 40%
pwgt = estimateFrontierByReturn (p,0.23/365) %生成收益最大的组合
[brsk, bret]=estimatePortMoments(p,pwgt) %计算风险 brsk 与回报率
bret plotFrontier (p)
hold on plot(brsk, bret, '*r'); hold on

```

6.4 函数一 data2days.m

```

function t=data2days(y,m,d)
    if ((rem(y,400)==0)||((rem(y,100)==0&&rem(y,4)==0))&&(m>2))
        a=1;
    else
        a=0;
    end
    month_day=[0 31 59 90 120 151 181 212 243 273 304 334];
    b=month_day(m);
    t=a+b+d;
end

```

6.5 函数二 normalize.m

```

function [normalized_data] = normalize(source_data)
[m, n] = size(source_data);
for i = 1:n
    for j=1:m
        ma = max(source_data(:, i)); % Matlab 中变量名不宜和函数名相同, 所以不用
        % max、min、mean 等变量名
        mi = min(source_data(:, i));
        normalized_data(:, i) = (source_data(:, i)-mi) / (ma-mi);
    end
end
end

```