

# 基于非线性规划的最优投资组合

**摘要：**投资者常常希望实现既定风险下的收益最大化和既定收益下的风险最小化，而面对这两个相互冲突和矛盾的投资目标，投资决策应该是同时投资于几种资产，即实行投资的分散化。因此，本文采用 SAS 软件和非线性规划方法选择最优投资组合。

首先，通过在锐思数据库下载的 2007 年至 2017 年股票原始数据，对其进行合并清洗后，得到个股的月持有期收益率数据集 `res_sas.monret`，以及 A 股市场月持有期收益率数据集 `res_sas.monretm`。挑选没有缺失值的 8 只股票，本文挑选的 8 只股票分别为：深物业 A、南玻 A、深科技、飞亚达 A、深证能源、凌云 B 股、锦港 B 股、凯马 B。

接着，建立股票代码数据集 `res_sas.stk8`，再与数据集 `res_sas.monret`、`res_sas.monretm` 进行横向合并，得到数据集 `res_sas.return`。

然后，选择非线性规划结合 SAS 软件进行计算选择最优投资组合。我们任选 8 只股票中的 4 只股票进行两两组合，按不同权重水平来看，如果风险一致，则投资组合 3 收益率最高，判断其为最优投资组合；后求得不同权重水平下，风险不一致，则上述判断无效。重新根据同一权重水平时的收益与风险的关系选取最优投资组合，根据其关系，绘制出各个投资组合的最小方差前沿。从而得到三个有效边界，经过对比后，得到合理投资 4 只股票构成的投资组合比合理投资 2 只股票构成的投资组合要好些，选择更多的合理的股票，能尽量分散投资风险。

最后，为选择出更优的投资组合，我们用 PROC CORR 过程求得两个投资组合（组合 2 和组合 3）与 PROC IML 过程求得的一个投资组合做“优劣”对比，选择三者中最优的投资组合。用 PROC NLP 以及 PROC OPTMODEL 过程优化由 8 只股票构成的投资组合的权重，得到最优投资组合的权重。这说明风险是伴随着收益的，收益越高，相应的风险也会越高。

**关键字：**投资组合，最小方差前沿，有效边界，收益与风险，非线性规划。

# 目录

一、 引言.....	3
二、 问题描述.....	3
(一) 如何选择最优的投资组合.....	3
(二) 最优投资组合的条件.....	4
三、 指标选择.....	4
(一) 数据及数据集的选择.....	4
(二) 股票的选择.....	6
四、 数据描述.....	8
(一) 证券的预期收益.....	8
(1) 单一证券的预期收益.....	8
(2) 证券组合的预期收益.....	9
(二) 证券的风险.....	9
(1) 单一证券的风险.....	9
(2) 证券组合的风险.....	9
(三) 每种证券与其他各种证券之间的协方差或相关系数.....	9
五、 模型建立.....	10
六、 模型求解、检验及评价.....	10
(一) 使用 PROC CORR 过程步 .....	10
(二) 使用 PROC IML 过程.....	19
(三) 优化四只股票构成的投资组合的权重.....	21
(四) 使用 PROC NLP 过程.....	24
(五) 使用 PROC OPTMODEL 过程 .....	25
结论与建议.....	26
参考文献.....	27
附    录.....	28

## 一、引言

马科维茨均值方差模型广泛应用于资金在各种证券资产上的合理分配，在应用时一般分为两个步骤进行。

第一步，估计各个证券的期望收益率、方差以及每一对证券之间的相关系数。

一般的，对期望收益、方差、相关系数的估计可利用历史数据通过统计技术来完成，在市场相对稳定的情况下，这种估计具有较好的精确性，在不稳定的情况下还需要投资者在对未来形势作出判断分析的基础上对这些估计作出改进。

第二步，对于给定的期望收益水平，计算证券组合最小方差。

因为每个人的最佳证券组合是不同的，因此，不可能有一个公共的计算过程为每个投资人计算出最佳证券组合。作为一般性的方法，我们只能追求一个次要的目标，即对于给定的期望收益水平，计算最小方差证券组合。这个目标一旦达到，我们就得到了有效边界，或者得到了有效组合。

这个步骤分为允许卖空和不允许卖空两种情况。当允许卖空时，为求得每一给定期望收益率水平最小方差组合，实际上只要对两个不同的期望收益水平分别计算其方差组合即可，因为此时的方差集可由其上的两个组合的再组合产生。而对于给定的某期望收益率水平计算其最小方差组合可通过数学上的拉格朗日惩罚数来完成，或通过计算机的试错程序来确定。在不允许卖空的情况下，其计算情况会相对更复杂。

无论如何，马科维茨模型在应用时面临的最大困难是计算十分复杂，所以在实际中，马科维茨模型并不单一应用于一般的资产分配问题。而利用数学中的有关约束条件下求得优化问题的解的技术，可以计算出有效边界或有效组合，因此，我们利用数学中的非线性规划模型进行选择最优投资组合。

## 二、问题描述

面对市场上众多的证券，存在着很多的证券组合。按照马科维茨的证券组合选择模式，在分别计算出市场上众多证券的预期收益、方差或标准差、协方差和相关系数后，运用非线性规划方法，可以从中结合若干证券组成许多种可行的组合，再通过对这些组合收益和风险相对关系比较，选出一系列有效组合以供选择，同时满足以下两个条件的一组证券组合，称为有效组合：第一，在各种风险条件下，提供最大的预期收益率；第二，在各种预期收益率水平下，提供最小的风险。显然，这种有效证券组合正是投资者希望得到的最优组合。

### （一）如何选择最优的投资组合

投资者如何在有效边界中选择一个最优的投资组合，取决于投资者对风险的态度。并可由无差异曲线表示出来。无差异曲线的特点：某投资者对同一条无差异曲线上的投资点有相同的偏好；投资者更偏好位于左上方的无差异曲线；无差异曲线的斜率为正。

无差异曲线的形状因人而异，风险偏好投资者的无差异曲线较为平坦，而风险厌恶者的无差异曲线较为陡峭。

(二) 最优投资组合的条件

最优投资组合的选择应符合以下条件：

- ①最优投资组合应位于有效边界上，只有在有效边界上的组合才是有效组合。
- ②最优投资组合又应同时位于投资者的无差异曲线上，而且应位于左上方的无差异曲线上。
- ③由于无差异曲线斜率为正、非满足性和回避风险的特性使无差异曲线呈凹凸型，而有效集一般呈凹型，两者有可能相切并且只有一个切点。

三、指标选择

(一) 数据及数据集的选择

本次课程设计，数据来源于锐思数据库官方网站（<http://www.resnet.cn/>）。一个是个股的月收益率，下载的时间段为 2007 年 1 月 1 日到 2017 年 12 月 31 日，共下载到 7 个 EXCEL 文件。另一个是市场的月收益率，下载的时间段为 2007 年 1 月 1 日到 2017 年 12 月 31 日，共 1 个 EXCEL 文件。

下面详细介绍个股信息数据集的建立。在锐思数据库中查询个股的月收益率如图 1 所示。



图 1 锐思数据库中查询个股月收益率

然后将这 7 个 EXCEL 下载到本地的“D:\Res\_sas\”这个具体的物理路径中，创建永久逻辑库 res\_sas,其连接路径为“D:\Res\_sas”,然后将数据都导入到 SAS 系统的 res\_sas 永久逻辑库中。个股数据导入 SAS 系统结果如图 2 所示。

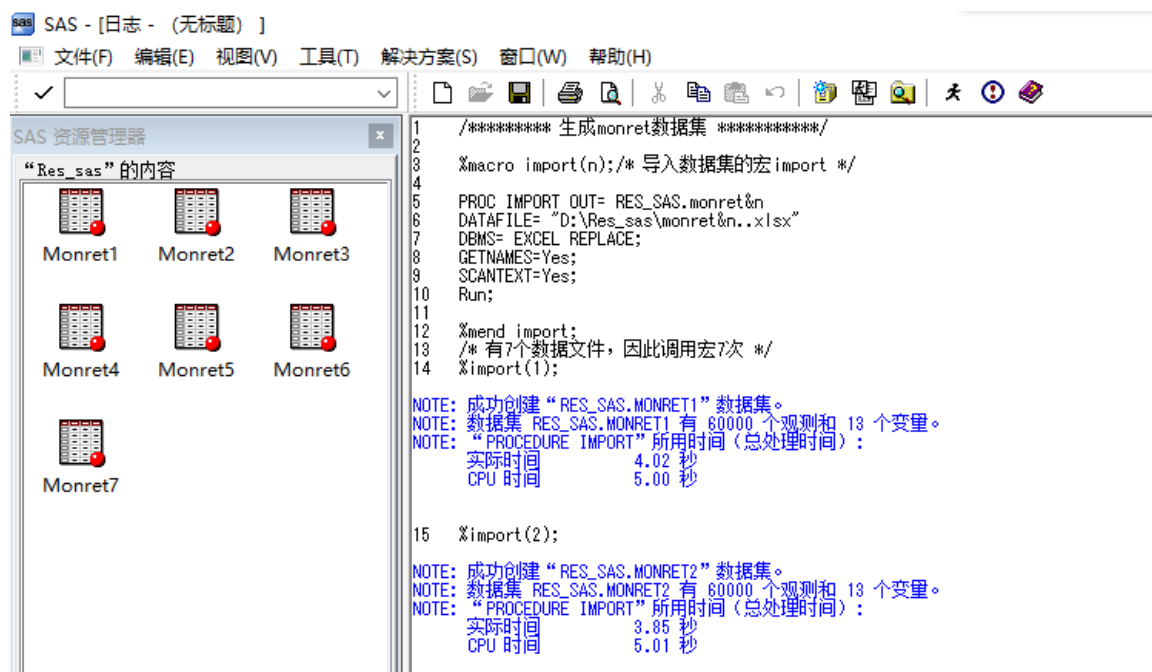


图2 个股数据导入 SAS 系统结果

根据上图2可知,数据集 res\_sas.monret1 和数据集 res\_sas.monret2 都是6万个观测,因此没有遗漏 EXCEL 中的数据。然后在将7个数据集合并到一个数据集中,新建数据集 res\_sas.monret, 将数据集 res\_sas.monret1 到 res\_sas.monret7 中的数据全部追加到 res\_sas.monret 中。得到的数据集 res\_sas.monret 的数据如表1所示。

表1 数据集 res\_sas.monret (部分)

	股票代码	最新股票名称	交易时股票代码	日期	月交易天数	月收益率	月资本收益率	等权平均市场月收益率	流通市值加权平均市场月收益率	总市值加权平
1	000001	平安银行	000001	31JAN2007	20	0.322	0.322	0.2278	0.1976	0.2055
2	000001	平安银行	000001	28FEB2007	15	-0.0042	-0.0042	0.192	0.134	0.13
3	000001	平安银行	000001	30MAR2007	22	-0.0089	-0.0089	0.2107	0.1314	0.1281
4	000001	平安银行	000001	30APR2007	19	0.3745	0.3745	0.3191	0.3087	0.302
5	000001	平安银行	000001	31MAY2007	9	0.1056	0.1056	0.0953	0.1394	0.122
6	000001	平安银行	000001	29JUN2007	7	-0.0408	-0.0408	-0.1468	-0.0971	-0.0907
7	000001	平安银行	000001	31JUL2007	21	0.3185	0.3185	0.2237	0.2075	0.2029
8	000001	平安银行	000001	31AUG2007	23	0.0489	0.0489	0.1317	0.1436	0.1439
9	000001	平安银行	000001	28SEP2007	20	0.0521	0.0521	0.0362	0.0488	0.055
10	000001	平安银行	000001	31OCT2007	18	0.2019	0.2019	-0.0892	-0.0369	-0.045
11	000001	平安银行	000001	30NOV2007	22	-0.2491	-0.2491	-0.0787	-0.1554	-0.159
12	000001	平安银行	000001	28DEC2007	19	0.0698	0.0698	0.2073	0.1742	0.1865
13	000001	平安银行	000001	31JAN2008	22	-0.1373	-0.1373	-0.0679	-0.0847	-0.096
14	000001	平安银行	000001	29FEB2008	16	-0.0045	-0.0045	0.0783	0.0447	0.0488
15	000001	平安银行	000001	31MAR2008	21	-0.1493	-0.1493	-0.1826	-0.1848	-0.2041
16	000001	平安银行	000001	30APR2008	21	0.0504	0.0504	-0.0376	-0.0158	-0.0004
17	000001	平安银行	000001	30MAY2008	20	-0.1482	-0.1482	-0.0361	-0.0681	-0.0586
18	000001	平安银行	000001	30JUN2008	19	-0.2338	-0.2338	-0.2516	-0.2418	-0.2405
19	000001	平安银行	000001	31JUL2008	23	0.076	0.076	0.0936	0.0505	0.048
20	000001	平安银行	000001	29AUG2008	21	-0.0288	-0.0288	-0.2369	-0.2145	-0.2186
21	000001	平安银行	000001	26SEP2008	19	-0.2579	-0.2579	-0.096	-0.0753	-0.0711
22	000001	平安银行	000001	31OCT2008	19	-0.2719	-0.2741	-0.25	-0.253	-0.2538
23	000001	平安银行	000001	28NOV2008	20	0.0741	0.0741	0.1942	0.1677	0.1637
24	000001	平安银行	000001	31DEC2008	23	0.0523	0.0523	0.1058	0.0273	0.0245
25	000001	平安银行	000001	23JAN2009	15	0.2304	0.2304	0.1475	0.1201	0.1047
26	000001	平安银行	000001	27FEB2009	18	0.1856	0.1856	0.0855	0.0783	0.0808
27	000001	平安银行	000001	31MAR2009	22	0.1551	0.1551	0.2207	0.1925	0.1911
28	000001	平安银行	000001	30APR2009	21	0.0238	0.0238	0.076	0.0557	0.0589
29	000001	平安银行	000001	27MAY2009	17	0.0931	0.0931	0.074	0.0647	0.0621
30	000001	平安银行	000001	30JUN2009	16	0.2231	0.2231	0.0884	0.091	0.0919
31	000001	平安银行	000001	31JUL2009	22	0.1998	0.1998	0.1463	0.1651	0.1632

根据上述的表1可知,个股信息数据集已经建立。接下来,开始创建市场信息数据集。首先,在锐思数据库中查询市场月收益率数据,查询结果如图3所示。



图 3 锐思数据库中查询市场月收益率数据

然后，将其下载到本地电脑的“D:\Res\_sas\”这个具体的物理路径中，然后将其导入到 SAS 系统中，导入市场的月收益数据集的结果如表 2 所示。

表 2 市场的月收益数据集（部分）

交易标识	市场标识	日期	月交易天数	等权平均市场月收益率	流通市值加权平均市场月收益率	总市值加权平均市场月收益率	等权平均市场月资本收益率	流通市值加权
1	0	A	31JAN2007	20	0.2376	0.1914	0.0869	0.2376
2	0	A	28FEB2007	15	0.1885	0.1042	0.0531	0.1885
3	0	A	30MAR2007	22	0.1884	0.123	0.1116	0.1882
4	0	A	30APR2007	21	0.3415	0.2974	0.2194	0.3409
5	0	A	31MAY2007	18	0.1009	0.1023	0.0805	0.0998
6	0	A	29JUN2007	21	-0.1555	-0.0862	-0.0723	-0.1571
7	0	A	31JUL2007	22	0.2325	0.1969	0.177	0.2314
8	0	A	31AUG2007	23	0.1303	0.1593	0.1655	0.1301
9	0	A	28SEP2007	20	0.0432	0.0488	0.0616	0.0432
10	0	A	31OCT2007	18	-0.0882	-0.017	0.0496	-0.0882
11	0	A	30NOV2007	22	-0.0701	-0.1449	-0.1551	-0.0701
12	0	A	28DEC2007	20	0.2035	0.1494	0.0955	0.2035
13	0	A	31JAN2008	22	-0.0674	-0.1123	-0.1559	-0.0674
14	0	A	29FEB2008	16	0.0637	0.0351	0.0027	0.0636
15	0	A	31MAR2008	21	-0.1925	-0.1925	-0.2034	-0.1925

由此，我们完成了本课程设计的原始数据的下载和导入。那么，接下来就是确定要拿来构建投资组合的股票了。

## （二）股票的选择

选择具体的八只股票，在这八只股票中，抽取两只或者两只以上的股票建立投资组合。选择股票为 000011（深物业 A）、000012（南玻 A）、000021（深科技）、000026（飞亚达 A）、000027（深证能源）、900957（凌云 B 股）、900952（锦港 B 股）、900953（凯马 B）八只股票。建立股票代码数据集 res\_sas.stk8 如表 3 所示。

表 3 股票代码数据集 res\_sas.stk8

	stkcd
1	000011
2	000012
3	000021
4	000026
5	000027
6	900957
7	900953
8	900952

然后将股票代码数据集 res\_sas.stk8 与数据集 res\_sas.monret 进行横向合并。得到新的数据集 res\_sas.return。然后再将处理过后的市场月收益数据 res\_sas.monretm 和数据集

res\_sas.return 横向合并到数据集 res\_sas.return 中，得到收益率数据集 res\_sas.return。具体结果如表 4 所示。

表 4 收益率数据集（部分）

VIEWTABLE: Res_sas.Return						
	股票代码	日期	月收益率	year	month	总市值加权平均市场月收益率
1	000011	31JAN2007	0.2088	2007	1	0.0869
2	000011	28FEB2007	0.0682	2007	2	0.0531
3	000011	30MAR2007	0.1698	2007	3	0.1116
4	000011	30APR2007	0.2938	2007	4	0.2194
5	000011	31MAY2007	0.1993	2007	5	0.0805
6	000011	29JUN2007	-0.0988	2007	6	-0.0723
7	000011	31JUL2007	-0.0274	2007	7	0.177
8	000011	31AUG2007	0.2323	2007	8	0.1655
9	000011	28SEP2007	0.0118	2007	9	0.0616
10	000011	31OCT2007	-0.0938	2007	10	0.0496
11	000011	30NOV2007	-0.1361	2007	11	-0.1551
12	000011	28DEC2007	0.1076	2007	12	0.0955
13	000011	31JAN2008	-0.1722	2008	1	-0.1559
14	000011	29FEB2008	-0.0162	2008	2	0.0027
15	000011	31MAR2008	-0.2008	2008	3	-0.2034
16	000011	30APR2008	-0.1456	2008	4	0.0516
17	000011	30MAY2008	0.1378	2008	5	-0.0667
18	000011	30JUN2008	-0.2472	2008	6	-0.2076

由此，我们选取了构建投资组合的股票，并且建立了收益率数据集 res\_sas.return。下面我们可以抽取这八只股票的月收益率进行数据的初步探索。将这八只股票每只股票的收益率分别作为一个变量，创建新的数据集 res\_sas.return2，里面含有这八只股票的月收益率，观测为 132 个，变量 x1 到变量 x8 分别代表股票 000011（深物业 A）、000012（南玻 A）、000021（深科技）、000026（飞亚达 A）、000027（深证能源）、900957（凌云 B 股）、900952（锦港 B 股）、900953（凯马 B）八只股票在 2007 年 1 月到 2017 年 12 月的月收益率，具体结果如表 5 所示。

表 5 八只股票每只股票的月收益率（部分）

VIEWTABLE: Res_sas.Return2								
	x1	x2	x3	x4	x5	x6	x7	x8
1	0.2088	0.1627	0.3205	0.1564	0.1645	0.3333	0.5836	0.1603
2	0.0682	0.1337	0.1555	0.2614	0.1887	0.4471	0.2112	0.1237
3	0.1698	0.0276	0.0629	0.3691	0.0989	-0.0764	-0.097	0.0239
4	0.2938	0.1322	0.2909	0.2666	0.2391	0.1709	0.0929	0.2204
5	0.1993	-0.1227	-0.0212	0.0702	0.2006	0.4424	0.585	0.2482
6	-0.0988	-0.0784	-0.1289	-0.1089	0.2053	-0.3195	-0.347	-0.2212
7	-0.0274	0.2563	0.1396	0.2413	0.1228	0.3537	0.4396	0.2704
8	0.2323	0.2643	0.1092	-0.003	0.0827	-0.0474	-0.0772	-0.0214
9	0.0118	0.1456	-0.0131	0.1244	0.0385	0.0655	0.0933	0.0656
10	-0.0938	-0.0788	-0.0028	-0.0336	-0.1711	-0.1287	-0.1308	-0.1026
11	-0.1361	-0.156	-0.2022	-0.4251	-0.1296	-0.1137	-0.1122	-0.1029
12	0.1076	0.1415	0.1692	0.2911	0.1926	0.1239	0.1106	0.1232
13	-0.1722	-0.1437	0.0089	-0.1043	-0.2588	-0.2008	-0.2432	-0.0328
14	-0.0162	0.2527	0.0148	0.1367	0.0427	0.0279	-0.0376	0.0117
15	-0.2008	-0.2105	-0.2286	-0.0675	-0.243	-0.1741	-0.0693	-0.2101
16	-0.1456	0.0798	-0.1606	-0.2349	-0.1074	-0.0193	-0.2042	0.0326
17	0.1378	-0.0364	-0.0827	-0.0396	-0.1015	-0.0237	-0.1871	0.1043
18	-0.2472	-0.203	-0.3056	-0.2301	-0.2688	-0.1394	-0.1386	-0.1788
19	0.0282	0.0709	0.1211	0.0424	0.0671	0.0282	0.0274	0.0854
20	-0.1597	-0.2383	-0.292	-0.1975	-0.1111	-0.2489	-0.25	-0.2937
21	-0.2917	-0.1349	-0.0565	-0.1969	0.0947	-0.1337	-0.12	-0.1136
22	-0.2005	-0.4048	-0.2615	-0.282	-0.0337	-0.3404	-0.2525	-0.3
23	0.3729	0.1849	0.1324	0.1976	0.1331	0.1596	0.0811	0.0073
24	-0.0395	0.1533	-0.0119	-0.0122	-0.1076	0.0275	0.075	0.0982
25	0.162	0.3059	0.1643	0.1732	0.0295	0.2054	0.1279	0.0762
26	-0.0199	-0.0541	0.3237	0.1213	0.3142	0.037	0.1959	0.1938
27	0.4944	0.6448	0.1285	0.2586	0.0909	0.1714	0.4138	0.201
28	0.3353	0.0628	0.1681	-0.0361	-0.14	-0.0152	0.378	0.0429
29	0.0068	-0.0348	0.3032	0.0517	0.031	0.1146	0.3473	0.0556
30	-0.009	-0.1319	0.0356	0.0418	0.0489	0.0417	-0.0722	0.115
31	0.2483	0.0862	0.1179	0.1262	0.2563	0.072	0.1717	-0.0021

根据表 5 得到的结果，我们可以对八只股票的每只股票的月收益做数据探索性分析，即数据描述。

## 四、数据描述

对所选的8只A股市场月收益率股票进行排序,计算月收益率,并检查数据完整性,是否存在缺失值,对得到的SAS数据集中\_FREQ\_变量检查,发现在2007年1月至2017年12月,8只股票的2007年1月至2017年12月间各股票月收益数据均完整,不存在缺失值,保证数据处理过程中免去对缺失值的处理,这样保证计算的准确性。

表6 八只股票个股收益及数据完整性检验(部分)

VIEWTABLE: SUMMARY STATISTICS					
	股票代码	_TYPE_	_FREQ_	_STAT_	月收益率
1	000011	0	132	N	132.0000
2	000011	0	132	MIN	-0.3671
3	000011	0	132	MAX	0.4944
4	000011	0	132	MEAN	0.0209
5	000011	0	132	STD	0.1538
6	000012	0	132	N	132.0000
7	000012	0	132	MIN	-0.4048
8	000012	0	132	MAX	0.6448
9	000012	0	132	MEAN	0.0154
10	000012	0	132	STD	0.1511
11	000021	0	132	N	132.0000
12	000021	0	132	MIN	-0.3416
13	000021	0	132	MAX	0.3966
14	000021	0	132	MEAN	0.0156
15	000021	0	132	STD	0.1429
16	000026	0	132	N	132.0000
17	000026	0	132	MIN	-0.4251
18	000026	0	132	MAX	0.3691
19	000026	0	132	MEAN	0.0185
20	000026	0	132	STD	0.1475
21	000027	0	132	N	132.0000
22	000027	0	132	MIN	-0.3405
23	000027	0	132	MAX	0.3142
24	000027	0	132	MEAN	0.0100
25	000027	0	132	STD	0.1105
26	900952	0	132	N	132.0000
27	900952	0	132	MIN	-0.3000
28	900952	0	132	MAX	0.3418
29	900952	0	132	MEAN	0.0085
30	900952	0	132	STD	0.1029
31	900953	0	132	N	132.0000

在马科维茨模型中,为确定一个有效的投资组合,有三个变量是必须的,这三个变量即证券的预期收益、证券的风险和每种证券与其他各种证券之间的协方差或相关系数。

### (一) 证券的预期收益

#### (1) 单一证券的预期收益

表7 八只股票个股预期收益

VIEWTABLE: Res_sas.M_out1a		
	股票代码	月收益率
1	000011	0.0209
2	000012	0.0154
3	000021	0.0156
4	000026	0.0185
5	000027	0.0100
6	900952	0.0085
7	900953	0.0214
8	900957	0.0210



## （2）证券组合的预期收益

证券组合的预期收益率可以用包含各种证券的预期收益率的加权平均数来表示。对此，我们从 8 只股票中选择了四只股票进行两两组合，并按不同权重进行组合计算其预期收益的详细情况，其协方差矩阵详细情况，见模型的详细求解。

## （二）证券的风险

### （1）单一证券的风险

表 8 八只股票个股风险度量（标准差）

VIEWTABLE: Res_sas.M_out1b		
	股票代码	月收益率标准差
1	000011	0.1538
2	000012	0.1511
3	000021	0.1429
4	000026	0.1475
5	000027	0.1105
6	900952	0.1029
7	900953	0.1706
8	900957	0.1646

### （2）证券组合的风险

证券组合的风险度量，是一个比较复杂的问题。这是因为，第一，可供选择的机会增加了，投资者不仅能在多种证券之间选择，而且可以将其资金按不同搭配方式投放在所选证券上。每一种搭配方式，就是一种证券组合，一种可供选择的机会。第二，与预期收益不同，证券组合的风险并不等于组合中单个证券风险的加权平均值，在许多情况下，前者要小于后者。

## （三）每种证券与其他各种证券之间的协方差或相关系数

证券组合的风险不仅取决于构成它的各种证券的风险，而且取决于他们之间相互关联的程度。这种相关的程度可以用统计学上的协方差或相关系数来表示。

如果用 $\sigma_i$ 和 $\sigma_j$ 分别表示证券 i 和证券 j 的标准差， $\sigma_{ij}$ 表示两种证券之间的协方差，

则用 $\rho_{ij} = \sigma_{ij} / \sigma_i \sigma_j$ 表示两种证券之间的相关系数。在本次实验过程中，我们从 8 只股票

中选择了四只股票进行两两组合，其协方差矩阵详细情况，见模型的详细求解。

## 五、模型建立

非线性规划的一般模型为：

$$\left. \begin{array}{l} \min f(X) \\ \text{s. t. } h_i(X) = 0 (i = 0, 1, 2, \dots, m) \\ g_j(X) \geq 0 (j = 0, 1, 2, \dots, m) \end{array} \right\} \mathbf{P}$$

$X \in E^n$ ,  $f(X)$ ,  $h_i(X)$ ,  $g_j(X)$  为  $E^n$  上的实函数。

在马科维茨模型中，股票收益指的是平均收益（也叫算术平均收益）。股票风险指的是股票的标准差，证券投资组合收益是每只股票的加权收益，在不允许卖空（即组合权重在 0 和 1 之间，且和为 1）的条件下，证券投资组合收益为：

$$R_p = x_1 R_1 + x_2 R_2 + \dots + x_N R_N = \sum_{i=1}^N x_i R_i$$

两只股票组合的风险定义为：

$$\sigma_p = (x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + 2x_1 x_2 \sigma_{12})^{1/2}$$

$$\sigma_{12} = \rho_{12} \sigma_1 \sigma_2$$

其中： $\sigma_1^2$  为股票 1 的方差； $\sigma_2^2$  为股票 2 的方差； $\sigma_{12}$  为股票 1 和股票 2 收益间的协方差，也可以用相关系数  $\rho_{12}$  表示为  $\sigma_{12} = \rho_{12} \sigma_1 \sigma_2$ 。

利用马科维茨模型，投资者求出使组合风险最小的证券投资组合权重，但马科维茨模型是非线性的，有时很难求解。因此，我们在非线性规划的基础上利用 PROC NLP、PROC IML、PROC OPTMODEL 三种方法计算多只股票的最优投资选择，用到了最小化的二次函数模型：

$$f(X) = 0.5X^T HX + g^T X + c$$

其中， $X$  为投资组合权重， $H$  为协方差阵， $g$  为线性系数向量， $c$  为常数。

## 六、模型求解、检验及评价

### （一）使用 PROC CORR 过程步

我们首先选取四只股票，从中选取两只股票构建投资组合。

考虑从 000011（深物业 A）、000012（南玻 A）、000021（深科技）、000026（飞亚达 A）四只股票里面选出两只股票构造投资组合。最优投资组合的选取标准为：既定风险下，收益最大的投资组合；既定收益下，风险最小的投资组合。

首先，我们使用宏的方法，从数据集 Res\_sas.return 中选取我们需要的 000011（深物业 A）、000012（南玻 A）、000021（深科技）、000026（飞亚达 A）四只股票的月收益率输出到数据集 Res\_sas.return1 中，即为四只股票的个股月收益率，具体如表 9 所示。

表 9 四只股票的个股月收益率

VIEWTABLE: Res_sas.Return1				
	月收益率	月收益率	月收益率	月收益率
1	0.2088		0.3205	0.1564
2	0.2088	0.1627	0.3205	0.1564
3	0.0682	0.1337	0.1555	0.2614
4	0.0682	0.1337	0.1555	0.2614
5	0.1698	0.0276	0.0629	0.3691
6	0.1698	0.0276	0.0629	0.3691
7	0.2938	0.1322	0.2909	0.2666
8	0.2938	0.1322	0.2909	0.2666
9	0.1993	-0.1227	-0.0212	0.0702
10	0.1993	-0.1227	-0.0212	0.0702
11	-0.0988	-0.0784	-0.1289	-0.1089
12	-0.0988	-0.0784	-0.1289	-0.1089
13	-0.0274	0.2563	0.1396	0.2413

从上图可知，尽管 Res\_sas.return1 数据集中的变量标签都是“月收益率”，但是变量名确不是一样的。可以用 proc contents data=Res\_sas.return1;run;这条语句来检查数据集 Res\_sas.return1 变量是否一致。结果如表 10 所示。

表 10 查看数据集 Res\_sas.return1 的属性

按字母排序的变量和属性列表				
#	变量	类型	长度	标签
1	COL000011	数值	8	月收益率
2	COL000012	数值	8	月收益率
3	COL000021	数值	8	月收益率
4	COL000026	数值	8	月收益率

然后使用 PROC CORR 过程步，求出这四只股票的月收益率的协方差矩阵，输出到数据集 Res\_sas.cov\_out1 里。协方差如表 11 所示。

表 11 个股月收益率的协方差矩阵

协方差矩阵，自由度 = 263					
		COL000011	COL000012	COL000021	COL000026
COL000011	月收益率	0.0235723688	0.0116404857	0.0113645773	0.0133120333
COL000012	月收益率	0.0116404857	0.0227481811	0.0114591723	0.0134397519
COL000021	月收益率	0.0113645773	0.0114591723	0.0203563627	0.0127577588
COL000026	月收益率	0.0133120333	0.0134397519	0.0127577588	0.0216827993

数据集 Res\_sas.cov\_out1 中的数据如表 12 所示，变量\_TYPE\_、\_NAME\_分别为观测类型以及该观测属于哪一个股票。从图中，我们可以看出，数据集 Res\_sas.cov\_out1 中不仅包含着协方差矩阵，也包含着每一只股票的月收益率的平均值。

由此，可以尝试算出包含两只股票的投资组合的在各个权重水平下的平均收益率。

表 12 数据集 Res\_sas.cov\_out1 中的数据

VIEWTABLE: Res_sas.Cov_out1 (Pearson 相关矩阵)						
	_TYPE_	_NAME_	月收益率	月收益率	月收益率	月收益率
1	COV	COL000011	0.0235723688	0.0116404857	0.0113645773	0.0133120333
2	COV	COL000012	0.0116404857	0.0227481811	0.0114591723	0.0134397519
3	COV	COL000021	0.0113645773	0.0114591723	0.0203563627	0.0127577588
4	COV	COL000026	0.0133120333	0.0134397519	0.0127577588	0.0216827993
5	MEAN		0.0209075758	0.0154492424	0.0156386364	0.0185151515
6	STD		0.1535329566	0.1508250015	0.1426757256	0.1472508041
7	N		264	264	264	264
8	CORR	COL000011	1	0.5026852597	0.5188019296	0.5888234648
9	CORR	COL000012	0.5026852597	1	0.5325125185	0.6051460984
10	CORR	COL000021	0.5188019296	0.5325125185	1	0.6072487397
11	CORR	COL000026	0.5888234648	0.6051460984	0.6072487397	1

接下来,开始用试错法求解投资组合的平均收益率。X 代表权重,增加变量 X,然后令 X 从 0 到 1,步长是 0.05,将变量名 COL000011 重命名为 r000011,代表股票 000011 (深物业 A) 的收益,其他三只股票也以同样的规则进行重命名。再然后,计算出投资组合的收益,由于是从 4 只股票中选取 2 只股票构成投资组合,共构建了 6 种投资组合,具体组合情况如表 13 所示。

表 13 在 4 只股票选 2 只构成的投资组合方式

组合序号	股票 1	股票 2
1	000011 (深物业 A)	000012 (南玻 A)
2	000011 (深物业 A)	000021 (深科技)
3	000011 (深物业 A)	000026 (飞亚达 A)
4	000012 (南玻 A)	000021 (深科技)
5	000012 (南玻 A)	000026 (飞亚达 A)
6	000021 (深科技)	000026 (飞亚达 A)

这 6 个投资组合的收益如表 14 所示:

表 14 各投资组合在不同权重水平下的平均收益

两种股票投资组合的收益												
Obs	_TYPE_	r000011	r000012	r000021	r000026	x	pfol_m1	pfol_m2	pfol_m3	pfol_m4	pfol_m5	pfol_m6
1	MEAN	0.020908	0.015449	0.015639	0.018515	0.00	0.015449	0.015639	0.018515	0.015639	0.018515	0.018515
2	MEAN	0.020908	0.015449	0.015639	0.018515	0.05	0.015722	0.015902	0.018635	0.015629	0.018362	0.018371
3	MEAN	0.020908	0.015449	0.015639	0.018515	0.10	0.015995	0.016166	0.018754	0.015620	0.018209	0.018227
4	MEAN	0.020908	0.015449	0.015639	0.018515	0.15	0.016268	0.016429	0.018874	0.015610	0.018055	0.018084
5	MEAN	0.020908	0.015449	0.015639	0.018515	0.20	0.016541	0.016692	0.018994	0.015601	0.017902	0.017940
6	MEAN	0.020908	0.015449	0.015639	0.018515	0.25	0.016814	0.016956	0.019113	0.015591	0.017749	0.017796
7	MEAN	0.020908	0.015449	0.015639	0.018515	0.30	0.017087	0.017219	0.019233	0.015582	0.017595	0.017652
8	MEAN	0.020908	0.015449	0.015639	0.018515	0.35	0.017360	0.017483	0.019352	0.015572	0.017442	0.017508
9	MEAN	0.020908	0.015449	0.015639	0.018515	0.40	0.017633	0.017746	0.019472	0.015563	0.017289	0.017365
10	MEAN	0.020908	0.015449	0.015639	0.018515	0.45	0.017905	0.018010	0.019592	0.015553	0.017135	0.017221
11	MEAN	0.020908	0.015449	0.015639	0.018515	0.50	0.018178	0.018273	0.019711	0.015544	0.016982	0.017077
12	MEAN	0.020908	0.015449	0.015639	0.018515	0.55	0.018451	0.018537	0.019831	0.015534	0.016829	0.016933
13	MEAN	0.020908	0.015449	0.015639	0.018515	0.60	0.018724	0.018800	0.019951	0.015525	0.016676	0.016789
14	MEAN	0.020908	0.015449	0.015639	0.018515	0.65	0.018997	0.019063	0.020070	0.015516	0.016522	0.016645
15	MEAN	0.020908	0.015449	0.015639	0.018515	0.70	0.019270	0.019327	0.020190	0.015506	0.016369	0.016502

从上述的表 14 可知, pfol\_m1 到 pfol\_m6 分别列出了各权重水平下 (表 14 只显示了权重为 0 到 0.70 时) 六种投资组合的收益。从每一只股票的平均收益可以看出, 股票 000011 (深物业 A) 的平均月收益率最高, 股票 000012 (南玻 A) 的平均月收益率最低; 从每一个投资组合的平均月收益率可以看出, 投资组合 3 (股票 000011 与股票 000026 组合) 的各个权重水平下的平均收益率最高, 投资组合 4 (股票 000012 与股票 000021 组合) 的各个权重水平下的平均收益率是最低的。因此, 若各种投资组合的风险都相同, 则投资组合 3 为最优投资组合, 因为它是平均收益率最大的两种股票的线性组合。

但是, 从表格中得到的结论并不一定是准确无误的。所以接下来, 要绘制图像来辅助我们的判断。通过 SAS 系统绘制得到如下投资组合对权重的线图, 具体如图 4 所示。

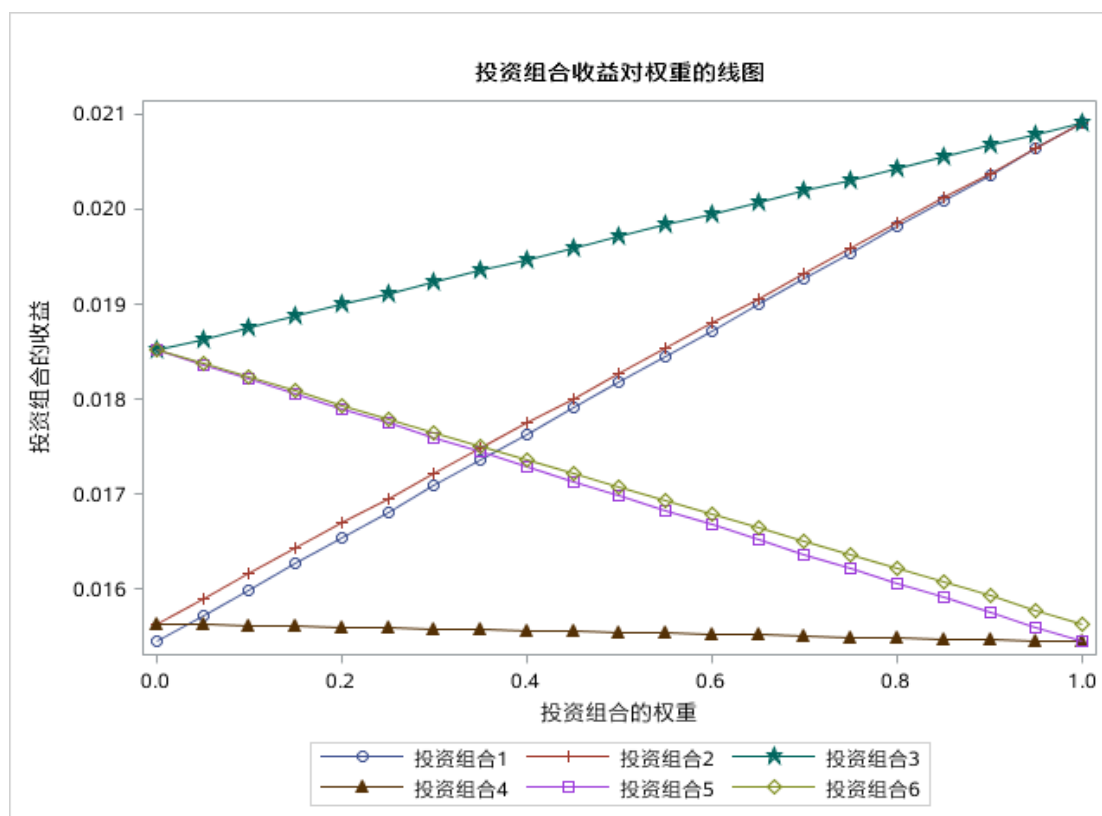


图 4 投资组合收益对权重的线图

从上图 4 我们可以看出, 各权重水平下, 投资组合 3 的平均收益率最高, 投资组合 2、投资组合 1、投资组合 5、投资组合 6 的平均收益率次之, 投资组合 4 的平均收益率最低。因此, 若各种投资组合的风险都相同, 则投资组合 3 为最优投资组合。

但是, 各个权重水平下的各种投资组合的风险相同只是我们的一个假设, 并没有得到实际的验证。如果上述的猜想被验证是正确的, 那么, 投资组合 3 即为最优投资组合, 反之, 我们就不能得出投资组合 3 为最优投资组合的结论。

所以, 在各个权重水平下, 各种投资组合的风险是否都相同呢? 我们将各个投资组合的风险先算出来, 然后再做判断。投资组合的风险可以通过下式计算:

$$\sigma_p = (x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + 2x_1 x_2 \sigma_{12})^{1/2}$$

计算得到的含有两只股票的投资组合的在各个权重水平下的风险大小如表 15 所示。

表 15 在各个权重水平下各个投资组合的风险

含有两只股票投资组合的风险水平							
Obs	x	pfol_r1	pfol_r2	pfol_r3	pfol_r4	pfol_r5	pfol_r6
1	0.00	0.15060	0.14247	0.14703	0.14247	0.14703	0.14703
2	0.05	0.14707	0.13947	0.14433	0.13950	0.14436	0.14412
3	0.10	0.14386	0.13680	0.14190	0.13683	0.14195	0.14144
4	0.15	0.14098	0.13447	0.13976	0.13449	0.13981	0.13900
5	0.20	0.13846	0.13250	0.13793	0.13249	0.13795	0.13682
6	0.25	0.13631	0.13090	0.13641	0.13084	0.13639	0.13490
7	0.30	0.13455	0.12969	0.13521	0.12956	0.13514	0.13328
8	0.35	0.13320	0.12888	0.13436	0.12866	0.13420	0.13194
9	0.40	0.13227	0.12848	0.13384	0.12814	0.13358	0.13090
10	0.45	0.13178	0.12848	0.13368	0.12802	0.13329	0.13018
11	0.50	0.13172	0.12890	0.13385	0.12829	0.13332	0.12976
12	0.55	0.13209	0.12972	0.13438	0.12894	0.13369	0.12967
13	0.60	0.13289	0.13095	0.13524	0.12999	0.13437	0.12989
14	0.65	0.13412	0.13256	0.13645	0.13141	0.13538	0.13043
15	0.70	0.13577	0.13454	0.13797	0.13319	0.13670	0.13128

根据表 15，可以得到含有两只股票的投资组合在各个权重下的风险水平，从列表中可以看出风险大小随着投资组合的权重变化。但是不能靠简单的观察，就判断在不同权重下投资组合间的风险大小的关系。由此通过 SAS 系统做出投资组合对权重的线图，具体如图 5 所示。

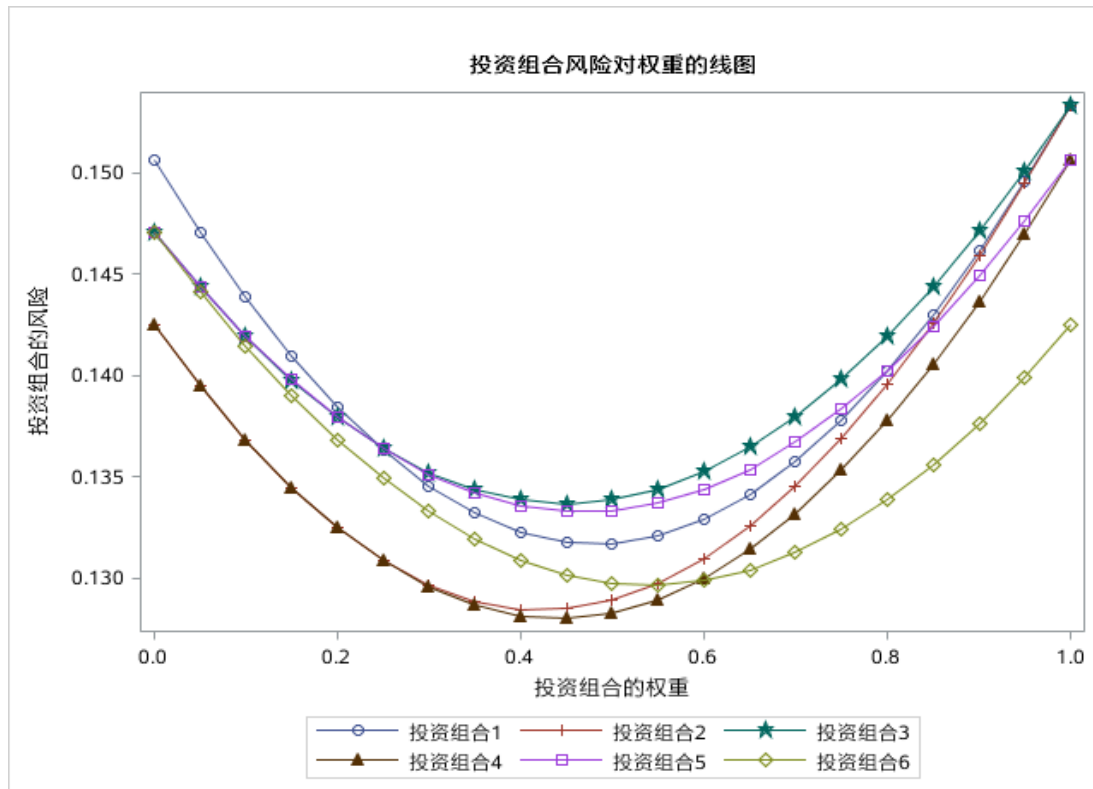


图 5 投资组合风险对权重的线图



从上图 5 可以看出, 投资组合 2、投资组合 4、投资组合 6 在各个权重下具有较小的风险; 投资组合 1、投资组合 5、投资组合 3 在各个权重下具有较大的风险。

所以, 各个投资组合的风险是随着权重的变化而变化。因此, 风险不变的假设不成立, 即投资组合 3 为最优投资组合的假设不成立。

上述全是对各个权重水平下, 对各投资组合的收益与风险分别进行讨论。综合上述叙述内容, 可知, 收益随着权重的变化而变化、风险随着权重的变化而变化, 但是同一个权重下, 收益与风险之间的关系并不知道。

所以, 下面研究同一权重下, 收益与风险之间的关系, 从而绘制出投资组合的可行集的边缘弧线, 即最小方差前沿。具体如图 6 所示。

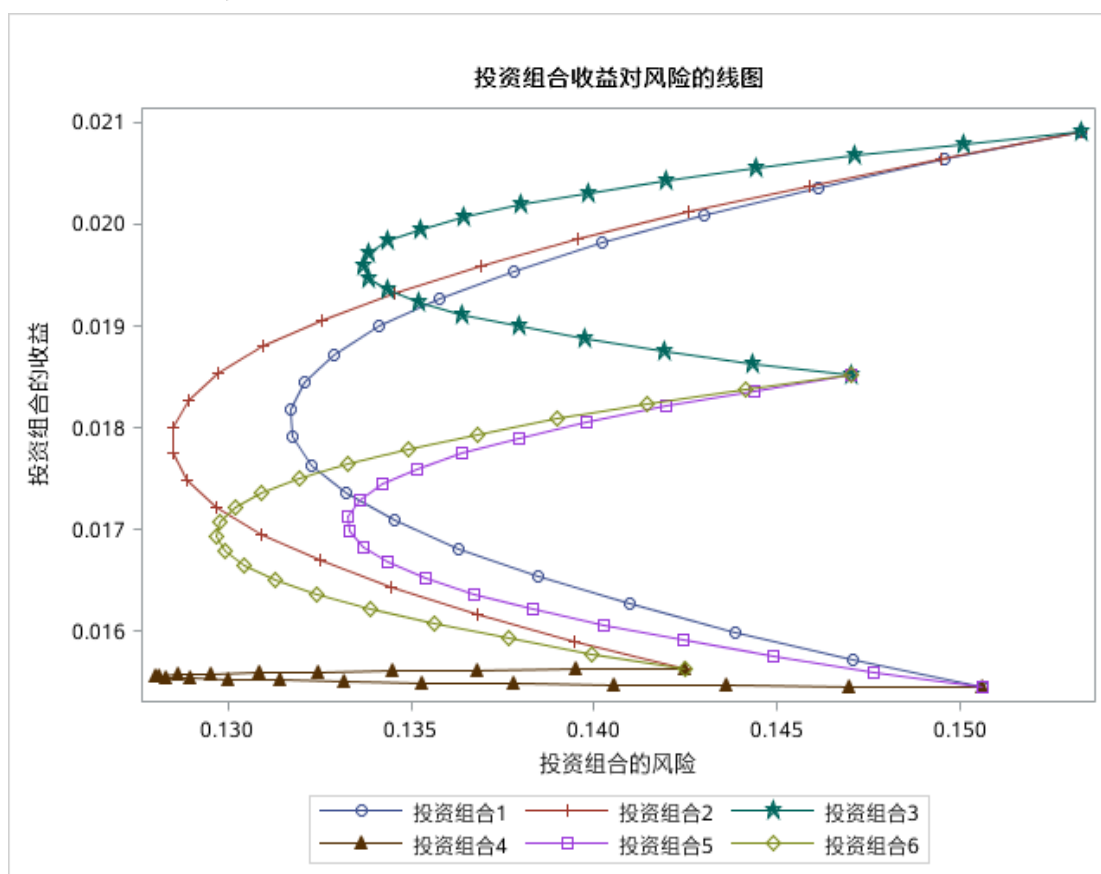


图 6 投资组合的最小方差前沿

根据上图 6, 可以看出投资组合 3 具有高风险高收益, 投资组合 4 具有低风险低收益, 符合生活实际, 因此该马科维茨模型通过检验。

但总体来说, 投资组合 1、投资组合 2、投资组合 3 的可行集的边缘弧线 (最小方差前沿) 位于投资组合收益对风险的线图的左上角, 是 6 个投资组合中, 最接近最优投资组合的投资组合。接下来, 抽取这 3 个投资组合做相关理论的分析。

通过 SAS 系统再一次绘制出投资组合 1、投资组合 2、投资组合 3 的可行集的边缘弧线 (最小方差前沿), 具体如下图 7 所示。

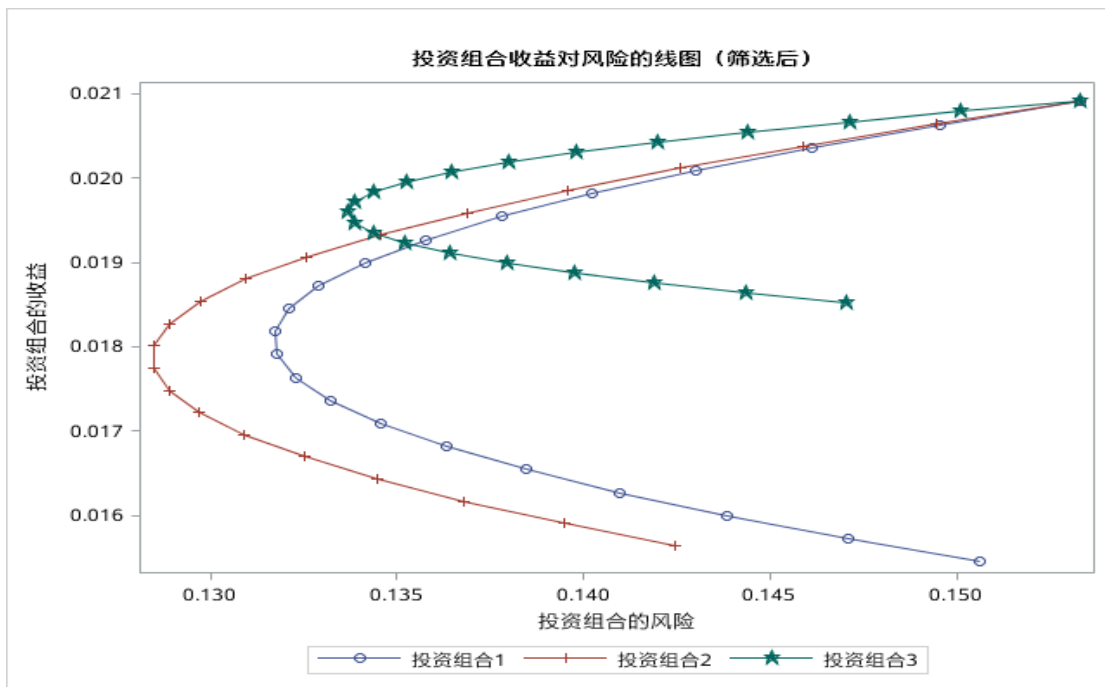


图 7 较优投资组合的最小方差前沿

经过筛选后，投资组合由原先的 6 个变为 3 个。在图 7 中的投资组合 1，在既定风险下，收益比投资组合 2 要小；在既定的收益下，风险比投资组合 2 要大。所以，根据均值方差理论，可以判断出投资组合 1 比投资组合 2 要差些。但是投资组合 2 和投资组合 3 之间不好比较，因为一个是低风险低收益，一个是高风险高收益。是否为最优的投资组合，是需要根据投资人的偏好来决定的。

紧接着，绘制出 3 个投资组合在各个权重下的风险函数的图像。具体如图 8 所示。

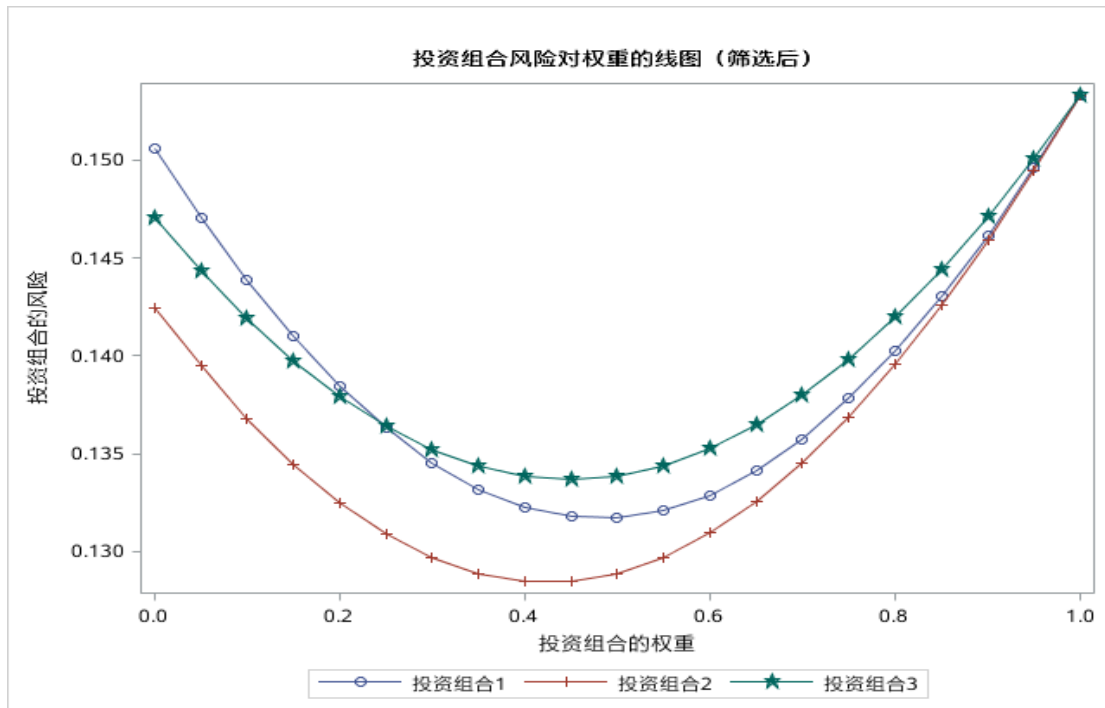


图 8 较优投资组合在各个权重下的风险



从上图 8 可知, 在各个权重水平下, 投资组合 3 的风险较高, 投资组合 1 的风险次之, 投资组合 2 的风险最小。

然后绘制出 3 个投资组合在各个权重下的收益函数图像, 具体如下图 9 所示。

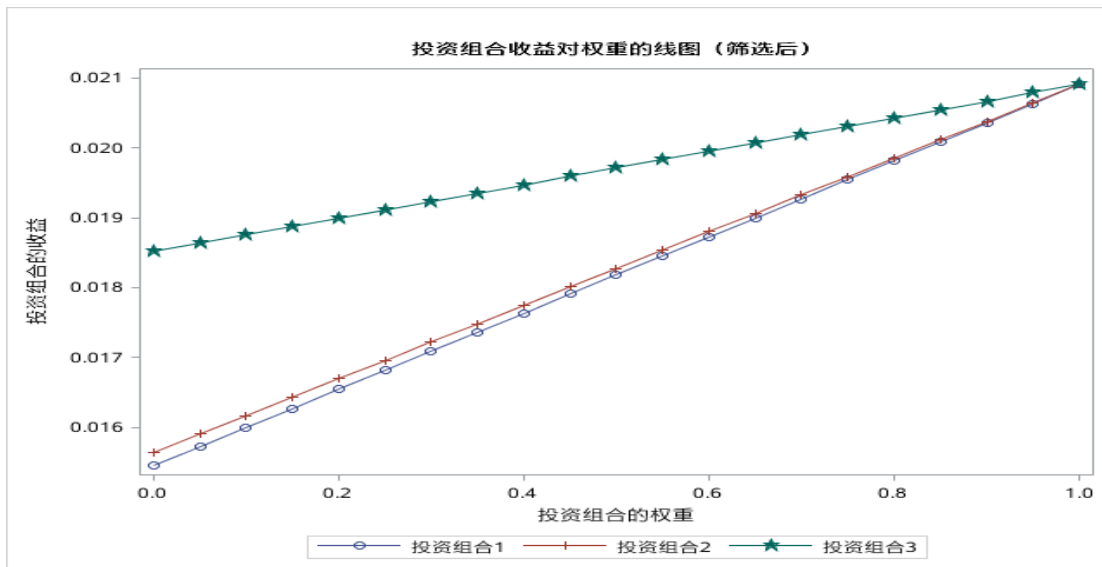


图 9 较优投资组合在各个权重下的收益

从上述的图 9 可以得出, 在各个权重下, 投资组合 3 具有最高的收益, 投资组合 2 次之, 投资组合 1 的收益最小。

综上图 7 到图 9 的结果, 可以得知: 从筛选后的 3 个投资组合的各个权重下的收益、各个权重下的风险、以及既定收益下的风险、既定风险下的收益的比较之后。可以得出, 投资组合 1 不是最优的投资组合。

下面, 开始在投资组合 2、投资组合 3 的可行集的边缘弧线 (最小方差前沿) 中搜寻最优投资组合的有效边界。具体的如下图 10 所示:

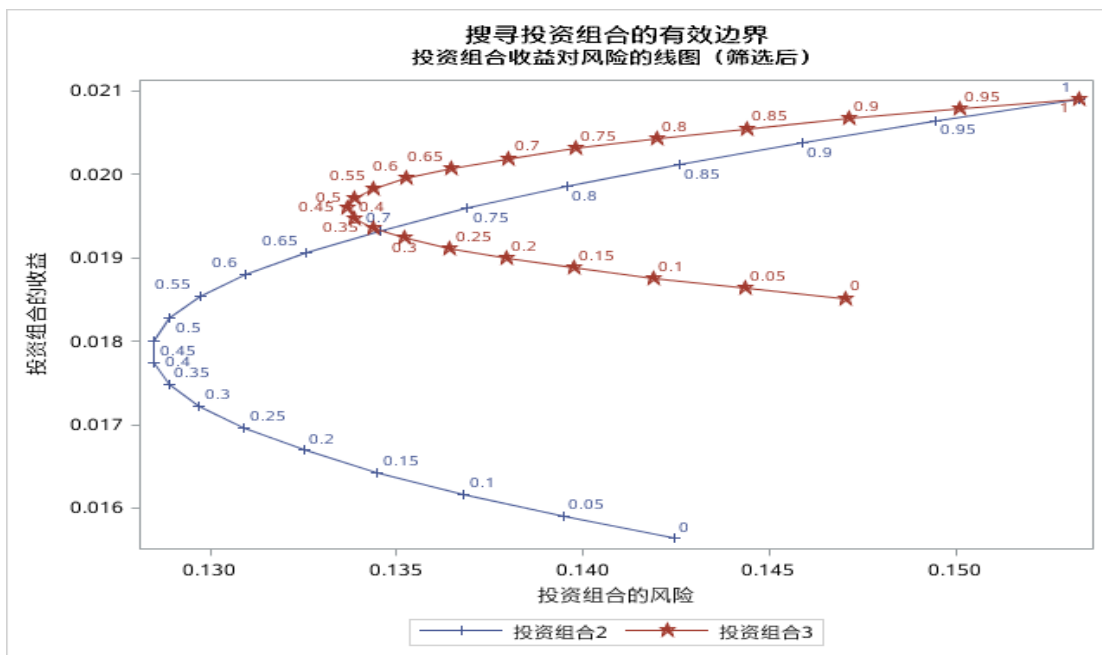


图 10 绘制较优投资组合的最小方差前沿

在图 10 中，可行集边缘弧线（最小方差前沿）标注的数字是各个权重水平。可以较为直观的看到，当权重水平介于 0.4-0.45 间，两个投资组合的风险都是最小的。我们可以根据投资组合的最小方差前沿绘制出投资组合的有效边界，进而得到最优投资组合的有效边界。

下面开始绘制投资组合的有效边界，具体如下图 11 到图 13 所示。

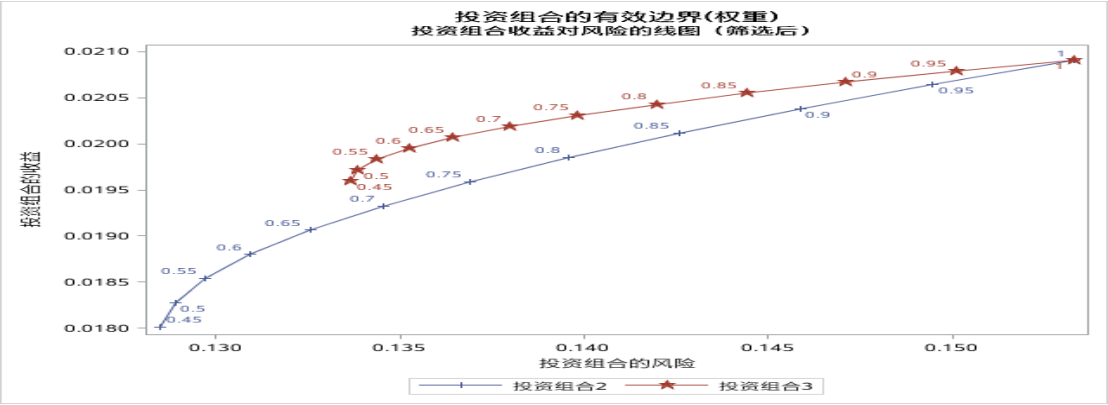


图 11 投资组合的有效边界（权重）

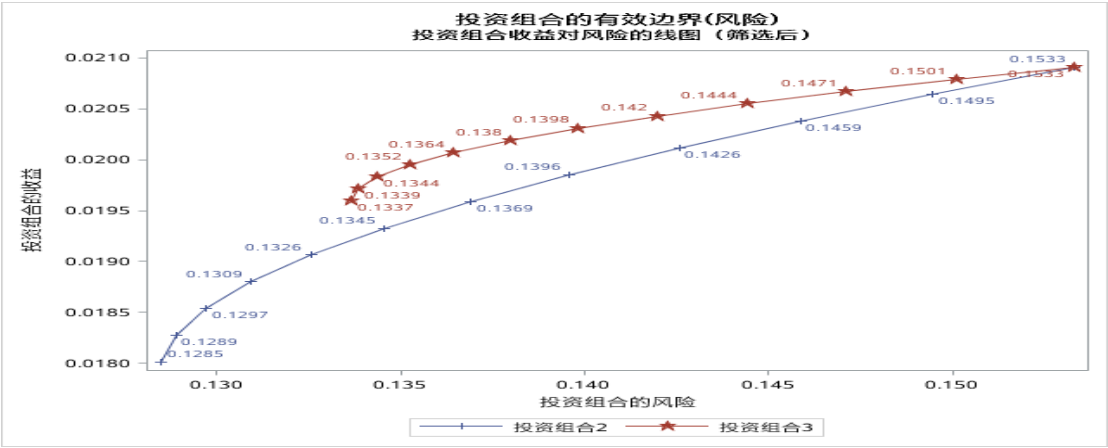


图 12 投资组合的有效边界（风险）

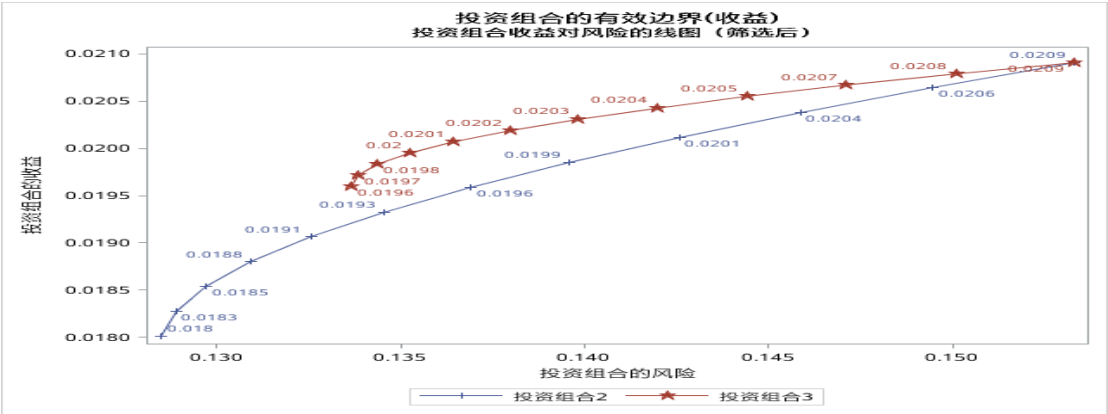


图 13 投资组合的有效边界（收益）

从上述的图 11 到图 13 可以看到最优投资组合的有效前沿。但是，投资人的偏好是未知的，在只讨论两只股票建立的投资组合中，组合 2 更受保险型投资者的喜爱，投资组合 3 更受激进型投资者的喜爱。但是不管最后的最优投资组合是哪一个，都只是在两

只股票构建起来的投资组合的范围内。如果超出了两只股票构建的投资组合，那么该最优投资组合的选取的分析就是片面的。也就是说，当构建投资组合的股票数越多，会有更好的最小方差前沿在图 10 中的左上边。

所以，下面由 PROC IML 过程分析当构建投资组合的股票数为 4 的时，有效边界的形状。

## （二）使用 PROC IML 过程

通过 SAS 系统的 IML 过程，绘制出构建投资组合股票数为 4 的时候的最优投资组合的最小方差前沿，可以根据最小方差前沿而绘制出最优投资组合的有效边界。具体的最小方差前沿图形如下图 14、图 15 所示。

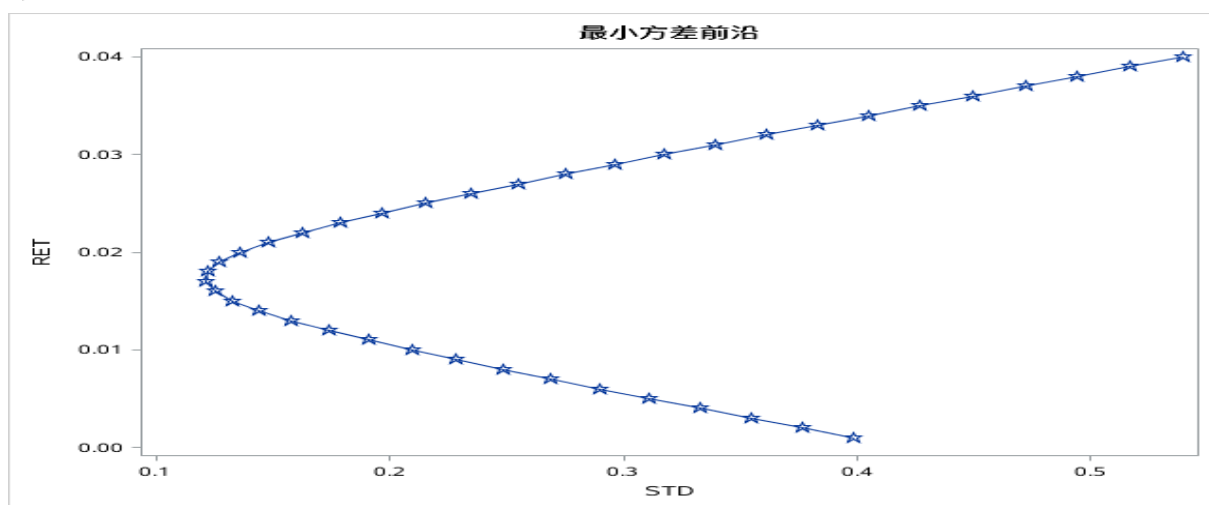


图 14 最优投资组合的最小方差前沿

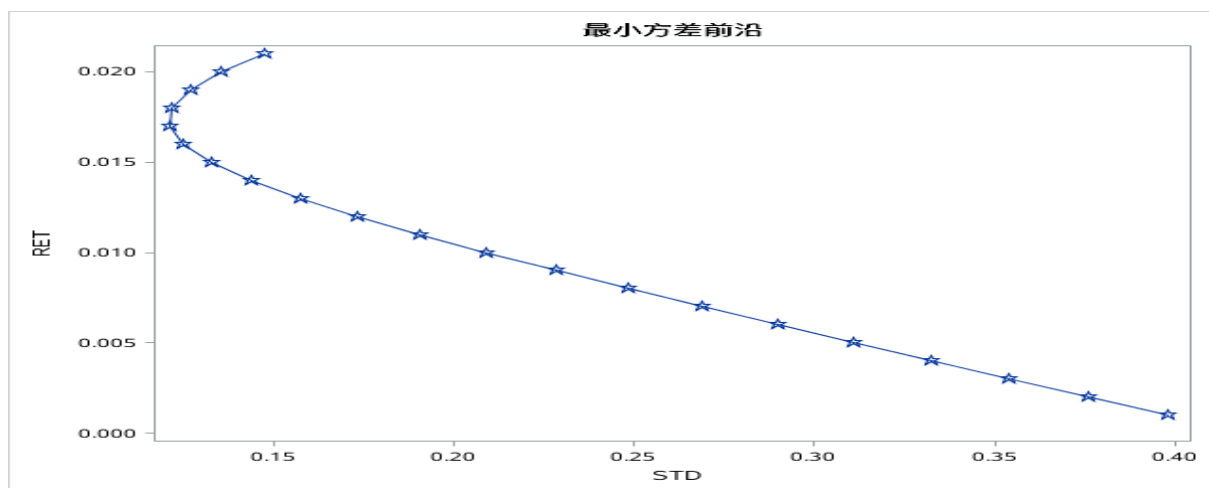


图 15 最优投资组合的最小方差前沿（放大）

根据上述图 14，可得四只股票构建的投资组合的最小方差前沿。但是根据图 14 的结果，并不能判断，图 14 中的这个投资组合的最小方差前沿在图 11 中的两个投资组合的最小方差前沿的左上方。因此，需要进一步将这个投资组合的有效边界绘制出来，再进行对比分析。IML 过程绘制的最优投资组合的有效边界的具体结果如图 16 所示。

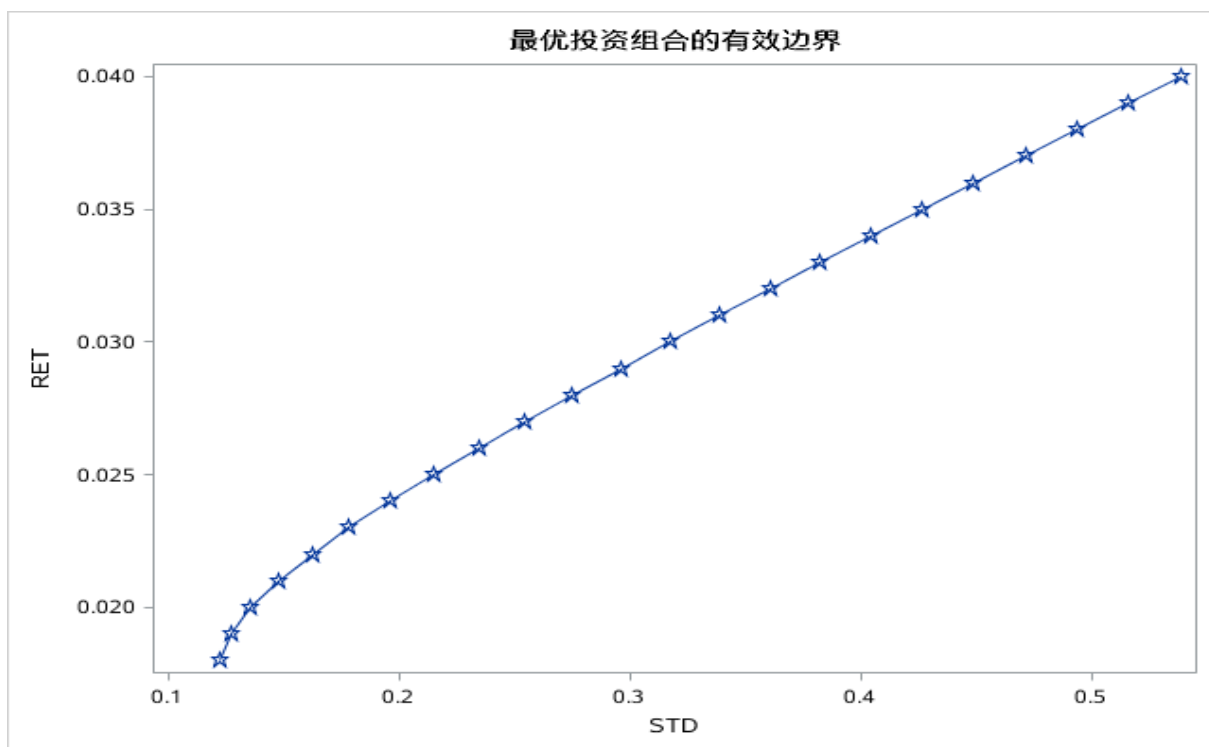


图 16 最优投资组合的有效边界

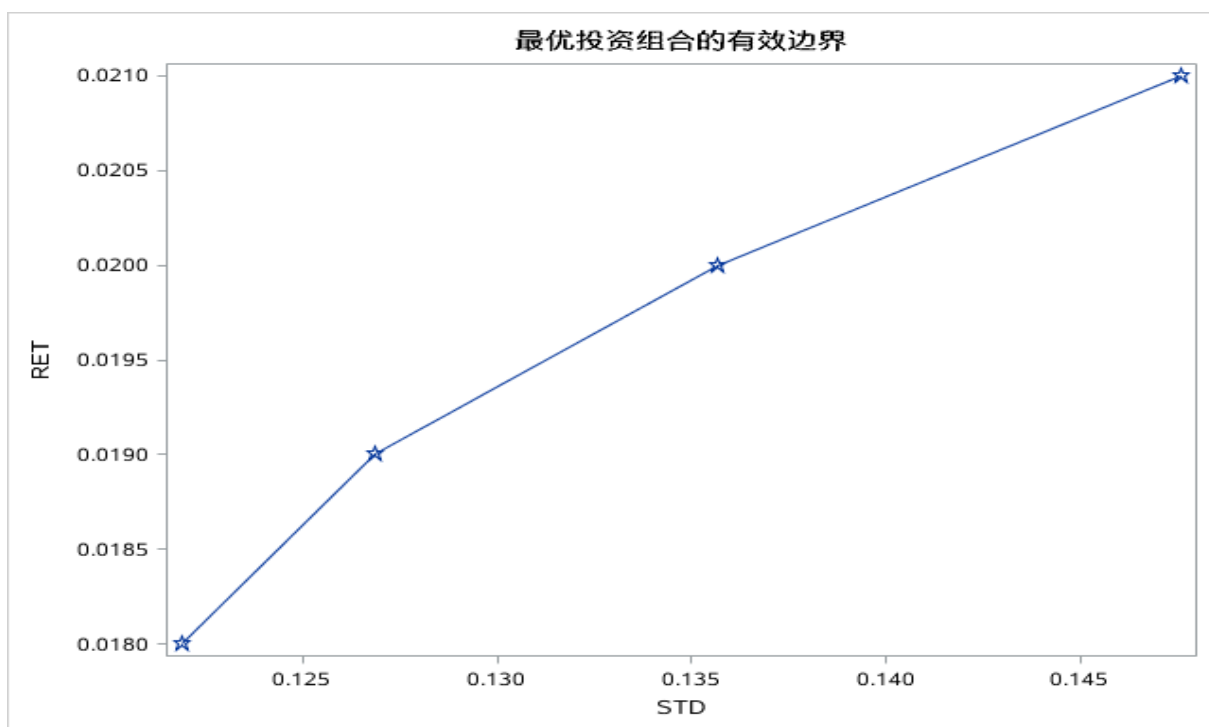


图 17 最优投资组合的有效边界（放大）

根据上述图 17 的结果，我们还是不能判断这个四只股票构建的一个投资组合和上述两只股票构建的两个投资组合，三个投资组合中，哪个是最优的投资组合。接下来，将进一步的分析，寻找出这三个较优投资组合中的最优投资组合。

将三个投资组合的有效边界全部绘制带到一张图上，具体如下图 18 所示：

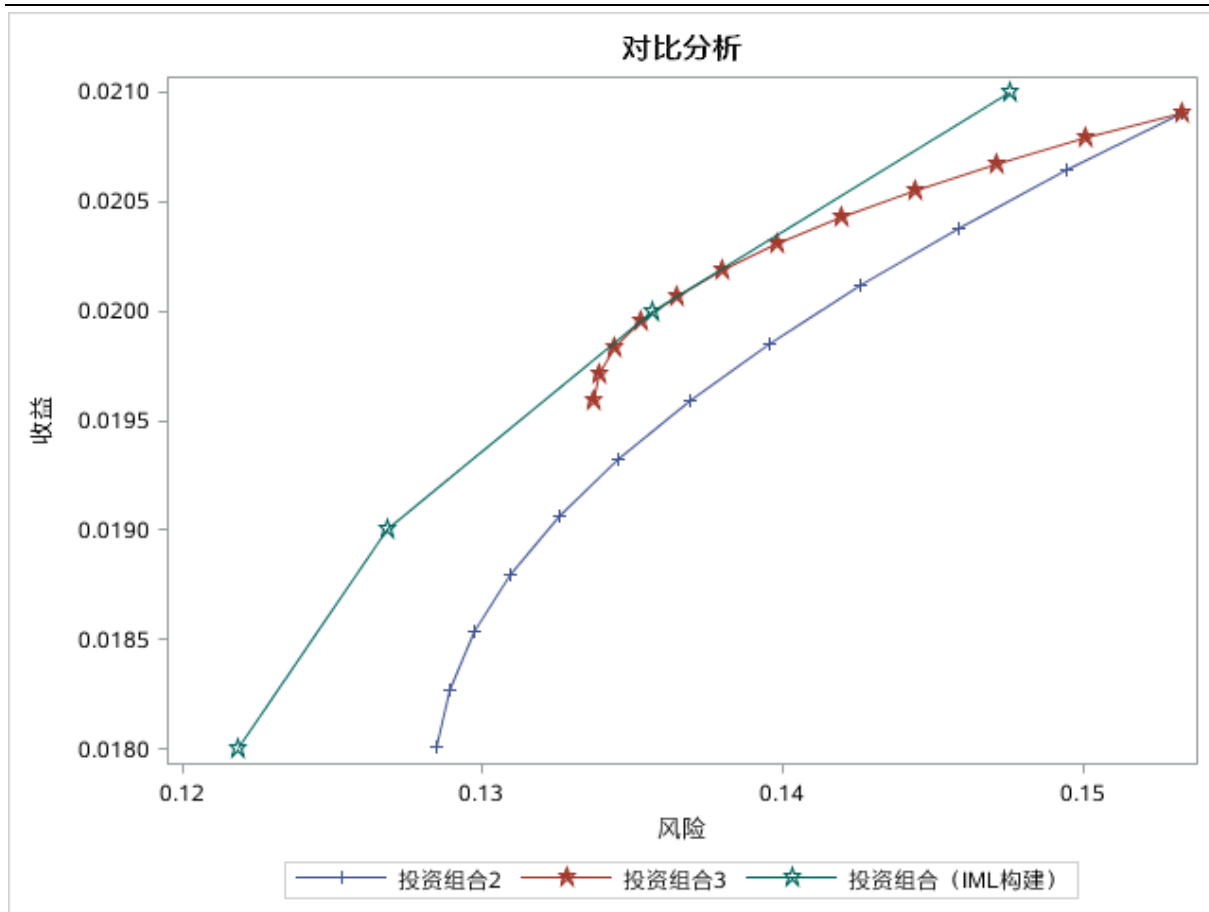


图 18 三个投资组合的有效边界的对比分析图

根据上述的图 18 可得，三个投资组合，其中由四只股票构建的最优投资组合的有效边界，在由两只股票构建的两个投资组合的左上方。这说明，将一定量的资产合理投资两只股票构建的最优投资组合，没有合理投资四只构建的最优投资组合好。

但是，在四只股票构建的投资组合里，如何确定每一只股票的投资权重呢？下面，将优化四只股票构建的投资组合的权重。

### （三）优化四只股票构成的投资组合的权重

首先，优化四只股票构成的投资组合的权重是一个非线性规划问题。上述，我们求出来的收益率的协方差矩阵  $H$  如下：

$$H = \begin{pmatrix} 0.0235085216 & 0.0116089568 & 0.0113337956 & 0.0132759769 \\ 0.0116257509 & 0.0227193859 & 0.0114446670 & 0.0134227395 \\ 0.0113501917 & 0.0114446670 & 0.0203305951 & 0.0127416097 \\ 0.0132951827 & 0.0134227395 & 0.0127416097 & 0.0216553527 \end{pmatrix}$$

令  $X = (x_1, x_2, x_3, x_4)$ 。然后，我们要求的非线性规划的目标函数和约束条件如下：

$$\min f = XHX^T$$

$$st. \begin{cases} 0 \leq x_i \leq 1, i=1,2,3,4 \\ x_1 + x_2 + x_3 + x_4 = 1 \\ 0.0209075758x_1 + 0.0154492424x_2 + 0.0156386364x_3 + 0.0185151515x_4 \geq \text{预期收益} \end{cases}$$

规定期望收益率为 0.008 到 0.020，步长为 0.002。然后通过编制宏，进行循环操作。下面，当预期收益为 0.008 时，PROC NLP 过程和 PROC OPTMODEL 过程得到的结果如下图 19 和图 20 所示。

PROC NLP: Nonlinear Minimization			
Optimization Results			
Parameter Estimates			
N	Parameter	Estimate	Gradient Objective Function
1	x1	0.235783	0.014643
2	x2	0.248349	0.014643
3	x3	0.336197	0.014643
4	x4	0.179671	0.014643

Value of Objective Function = 0.0073214165

图 19 PROC NLP 过程求得的结果

根据上述的图 19，可以知道：股票代码 000011 的权重（x1，下同）为 0.235783，股票代码 000012 的权重（x2，下同）为 0.248349，股票代码 000021 的权重（x3，下同）为 0.336197，股票代码 000026 的权重（x4，下同）为 0.179671。然后，NLP 过程求得的目标函数的值是投资组合方差的一半，此次目标函数的值为 0.0073214165，那么投资组合的方差为 0.014642833，再取平方根就可得到投资组合的风险。

Solver	QP
Algorithm	Interior Point
Objective Function	f
Solution Status	Optimal
Objective Value	0.0146428331
Primal Infeasibility	0
Dual Infeasibility	0
Bound Infeasibility	0
Duality Gap	1.367751E-17
Complementarity	0
Iterations	4
Presolve Time	0.00
Solution Time	0.01

[1]	x
1	0.23578
2	0.24835
3	0.33620
4	0.17967

图 20 PROC OPTMODEL 过程求得的结果

根据上述的图 20, 可以求得  $x_1$  为 0.23578,  $x_2$  为 0.24835,  $x_3$  为 0.33620,  $x_4$  为 0.17967。然后 PROC OPTMODEL 过程求得的目标函数是投资组合的方差, 此次为 0.0146428331。

综合上述的 PROC NLP 过程和 PROC OPTMODEL 过程, 得出两种过程优化投资组合的权重, 具体如下表 16:

表 16 NLP 过程优化四只股票构成的投资组合的权重

收益	000011	000012	000021	000026	目标函数	方差
0.008	0.235783	0.248349	0.336197	0.179671	0.007321417	0.014642833
0.010	0.235783	0.248349	0.336197	0.179671	0.007321417	0.014642833
0.012	0.235783	0.248349	0.336197	0.179671	0.007321417	0.014642833
0.014	0.235783	0.248349	0.336197	0.179671	0.007321417	0.014642833
0.016	0.235783	0.248349	0.336197	0.179671	0.007321417	0.014642833
0.018	0.332056	0.180030	0.263379	0.224535	0.007434791	0.0148695824
0.020	0.635734	0	0.012549	0.351717	0.009218055	0.0184361098

根据表 16 可得, 表格的第一列是预期收益率, 第二列是股票代码为 000011 的股票的权重, 第三列是股票代码为 000012 的股票的权重, 第三列是股票代码为 000021 的股票的权重, 第四列是股票代码为 000021 的股票的权重, 第五列目标函数是投资组合的方差的一半, 第六列是投资组合的方差。

OPTMODEL 过程得到的结果如下表 17 所示。

表 17 OPTMODEL 过程优化四只股票构成的投资组合的权重

收益	000011	000012	000021	000026	目标函数
0.008	0.23578	0.24835	0.33620	0.17967	0.014642833
0.010	0.23578	0.24835	0.33619	0.17967	0.014642833
0.012	0.23578	0.24835	0.33619	0.17967	0.014642833
0.014	0.23578	0.24835	0.33619	0.17967	0.014642833
0.016	0.23578	0.24835	0.33619	0.17967	0.014642833
0.018	0.33206	0.18003	0.26338	0.22454	0.014869583
0.020	0.6357724186	0.0000011434	0.0125795349	0.3516469031	0.018436112

根据表 17 可得, 表格的第一列是预期收益率, 第二列是股票代码为 000011 的股票的权重, 第三列是股票代码为 000012 的股票的权重, 第三列是股票代码为 000021 的股票的权重, 第四列是股票代码为 000021 的股票的权重, 第五列的目标函数即为投资组合的方差。

综合上述表 16 和表 17 对比来看，PROC NLP 过程和 PROC OPTMODEL 过程计算所得的结果差不多。同时，两个表都反应了：随着收益率的上升，风险也上升了。这说明风险是伴随着收益的。

#### （四）使用 PROC NLP 过程

紧接着，我们开始进行八只股票构建的投资组合权重的优化。得到八只股票的收益率协方差矩阵为 A 如下：

$$A = \begin{pmatrix} 0.0236623396 & 0.0116849150 & 0.0114079535 & 0.0133628426 & 0.0086300872 & 0.0117391590 & 0.0139428238 & 0.0087103809 \\ 0.0116849150 & 0.0228350062 & 0.0115029096 & 0.0134910486 & 0.0073542574 & 0.0138939055 & 0.0130127117 & 0.0093163350 \\ 0.0114079535 & 0.0115029096 & 0.0204340587 & 0.0128064525 & 0.0085550014 & 0.0122646062 & 0.0127601782 & 0.0084790341 \\ 0.0133628426 & 0.0134910486 & 0.0128064525 & 0.0217655581 & 0.0097825423 & 0.0153428905 & 0.0143835881 & 0.0099116323 \\ 0.0086300872 & 0.0073542574 & 0.0085550014 & 0.0097825423 & 0.0122101665 & 0.0090092111 & 0.0096812371 & 0.0063703016 \\ 0.0117391590 & 0.0138939055 & 0.0122646062 & 0.0153428905 & 0.0090092111 & 0.0270952158 & 0.0229878926 & 0.0129685768 \\ 0.0139428238 & 0.0130127117 & 0.0127601782 & 0.0143835881 & 0.0096812371 & 0.0229878926 & 0.0291195228 & 0.0137789440 \\ 0.0087103809 & 0.0093163350 & 0.0084790341 & 0.0099116323 & 0.0063703016 & 0.0129685768 & 0.0137789440 & 0.0105924184 \end{pmatrix}$$

令  $X = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8)$ 。然后要求的非线性规划的目标函数和约束条件如下：

$$\begin{aligned} \min f &= XAX^T \\ \text{st.} \left\{ \begin{aligned} 0 &\leq x_i \leq 1, i=1, 2, \dots, 8 \\ \sum_{i=1}^8 x_i &= 1 \\ 0.0209075758x_1 &+ 0.0154492424x_2 + 0.0156386364x_3 + 0.0185151515x_4 + \\ &0.0099515152x_5 + 0.0209613636x_6 + 0.0214431818x_7 + 0.0085136364x_8 \geq \text{预期收益} \end{aligned} \right. \end{aligned}$$

通过 SAS 系统的 NLP 过程，可以得到八只股票的最优投资组合的权重，具体如下表 18 所示。

表 18 NLP 过程优化八只股票构成的投资组合的权重

收益	000011	000012	000021	000026	000027	900957	900953	900952	目标函数
0.008	0.001562	0.017894	0.020065	0	0.406378	0	0	0.554101	0.004404233
0.010	0.036934	0.030474	0.036050	0	0.389738	0	0	0.506804	0.004420546
0.012	0.151848	0.071344	0.087985	0	0.335678	0	0	0.353146	0.004698742
0.014	0.206349	0.076490	0.104306	0.026156	0.304926	0.076733	0	0.205040	0.005238356
0.016	0.256288	0.081099	0.118789	0.055168	0.275252	0.146477	0.010075	0.056853	0.005947576
0.018	0.320200	0.061811	0.122054	0.099252	0.157632	0.188147	0.050905	0	0.006858583
0.020	0.406283	0.012956	0.102946	0.148870	0	0.237573	0.091371	0	0.00814115



根据表 3 可得，表格的第一列是预期收益率，第二列是股票代码为 000011 的股票的权重，第三列是股票代码为 000012 的股票的权重，第三列是股票代码为 000021 的股票的权重，第四列是股票代码为 000021 的股票的权重，第五列是股票代码为 000027 的股票的权重，第六列是股票代码为 900957 的股票的权重，第七列是股票代码为 900953 的股票的权重，第八列是股票代码为 900952 的股票的权重，第九列是非线性规划目标函数的值，即为投资组合的方差的一半。由上述的表 6.4 可知，风险是伴随着收益的，收益越高，相应的风险也会越高。

### （五）使用 PROC OPTMODEL 过程

使用 SAS 的 OPTMODEL 过程优化八只股票构成的投资组合的权重如下表 19，具体的代码见附录。

表 19 OPTMODEL 过程优化八只股票构成的投资组合的权重

收益	000011	000012	000021	000026	000027	900957	900953	900952	方差
0.008	0.003721	0.01751	0.019650	0.000001	0.405752	0.000000	0.000000	0.553364	0.008808
0.010	0.036941	0.03050	0.036068	0.000010	0.389721	0.000004	0.000003	0.506751	0.008841
0.012	0.151171	0.07093	0.087571	0.000649	0.335823	0.000601	0.000002	0.353253	0.009397
0.014	0.206198	0.076388	0.104197	0.026684	0.30478	0.076459	0.000132	0.205161	0.010476
0.016	0.255974	0.081174	0.118713	0.055183	0.275364	0.145544	0.011251	0.056797	0.011895
0.018	0.320144	0.061829	0.122027	0.099270	0.157597	0.187947	0.051146	0.000040	0.013717
0.020	0.406291	0.012966	0.102918	0.148859	0.000011	0.237577	0.091378	0.000000	0.016282

根据表 19 可得，表格的第一列是预期收益率，第二列是股票代码为 000011 的股票的权重，第三列是股票代码为 000012 的股票的权重，第三列是股票代码为 000021 的股票的权重，第四列是股票代码为 000021 的股票的权重，第五列是股票代码为 000027 的股票的权重，第六列是股票代码为 900957 的股票的权重，第七列是股票代码为 900953 的股票的权重，第八列是股票代码为 900952 的股票的权重，第九列是非线性规划目标函数的值，即为投资组合的方差。由上述的表 6.5 可知，风险是伴随着收益的，收益越高，相应的风险也会越高。

## 结论与建议

根据实验结果我们得到几个结论：当构建投资组合的股票数越多，会有更好的最小方差前沿；风险是伴随着收益的，收益越高，相应的风险也会越高，建议投资者合理多选择股票，分散风险。

实验的不足就是我们选择进行投资组合的股票样本太少，具有偶然性，面对实际的证券市场随机因素特别多，我们还需要考虑更多因素才能准确帮助投资者选择最优投资组合。

将所学的数学模型应用到金融方面，解决投资组合中计算复杂的问题，特别是能够利用我们所学的数学和软件编程结合有效帮助投资者尽可能规避投资风险，选择最优投资组合，这是一种计量的新方式方法。不仅只是在金融领域，在其他任何需要数据分析的领域中，也需要软件分析来实现，这是一种大的趋势。

## 参考文献

- [1] 肖枝洪. SAS 软件与金融数据库[M]. 北京:科学出版社, 2017.
- [2] 肖枝洪. 时间序列分析与 SAS 应用[M]. 武汉:武汉大学出版社, 2012.
- [3] 朱世武. SAS 编程技术教程[M]. 北京:清华大学出版社, 2013.
- [4] 朱世武. 金融计算与建模:理论、算法与 SAS 程序[M]. 北京:清华大学出版社, 2007.
- [5] 谷鸿秋. SAS 编程演义[M]. 北京:清华大学出版社, 2017.
- [6] 胡金焱. 证券投资学[M]. 北京:高等教育出版社, 2017.
- [7] 王燕. 应用时间序列分析[M]. 北京:中国人民大学出版社, 2016.
- [8] 夏坤庄. 深入解析 SAS:数据处理、分析优化与商业应用[M]. 北京:机械工业出版社, 2015.
- [9] SAS Institute. SAS 9.4 sql procedure user's guide, third edition [M]. SAS Institute, Incorporated, 2015.
- [10] SAS Institute. SAS 9.4 macro language: Reference, second edition [M]. SAS Institute, Incorporated, 2014.
- [11] Institute SAS 9.4 graph template language reference, third edition [M]. SAS Institute, 2014.

## 附 录

```
/****** 生成monret数据集 *****/

%macro import(n);/* 导入数据集的宏import */

PROC IMPORT OUT= RES_SAS.monret&n
DATAFILE= "D:\Res_sas\monret&n..xlsx"
DBMS= EXCEL REPLACE;
GETNAMES=Yes;
SCANTEXT=Yes;
Run;

%mend import;
/* 有7个数据文件，因此调用宏7次 */
%import(1);
%import(2);
%import(3);
%import(4);
%import(5);
%import(6);
%import(7);
run;

%macro he(m);/* 将所有个股的收益率追加到数据集res_sas.monret中，循环调用宏7次 */
proc append data=res_sas.monret&m base=res_sas.monret;
run;
%mend he;
%he(1);%he(2);%he(3);%he(4);%he(5);%he(6);%he(7);
run;

/* 给数据集res_sas.monret中的变量加标签，其中Monret为月收益率，Mretmc为总市值加权平均
市场月收益率 */
data res_sas.monret;
set res_sas.monret;
Label Stkcd="股票代码";
Label Lstknm="最新股票名称";
Label Stkcdotrd="交易时股票代码";
Label Date="日期";
Label Montrds="月交易天数";
Label Monret="月收益率";
Label Monaret="月资本收益率";
Label Mreteq="等权平均市场月收益率";
Label Mrettmv="流通市值加权平均市场月收益率";
```

```

Label Mretmc="总市值加权平均市场月收益率";
Label Mareteq="等权平均市场月资本收益率";
Label Marettmv="流通市值加权平均市场月资本收益率";
Label Maretmc="总市值加权平均市场月资本收益率";
run;

/***** 生成monretm数据集 *****/
PROC IMPORT OUT= res_sas.monretm1
DATAFILE= "D:\Res_sas\monretm.xlsx"
DBMS= EXCEL REPLACE;
GETNAMES=Yes;
SCANTEXT=Yes;
Run;

/* 给数据集res_sas.monretm1的变量加上标签 */
data Res_sas.monretm1;
set Res_sas.monretm1;
Label Exchflg="交易所标识";
Label Mktflg="市场标识";
Label Date="日期";
Label Montrds="月交易天数";
Label Mreteq="等权平均市场月收益率";
Label Mrettmv="流通市值加权平均市场月收益率";
Label Mretmc="总市值加权平均市场月收益率";
Label Mareteq="等权平均市场月资本收益率";
Label Marettmv="流通市值加权平均市场月资本收益率";
Label Maretmc="总市值加权平均市场月资本收益率";
set Res_sas.monretm1;
run;

/* 选出2007-2017年A股市场月持有期收益率 */
data Res_sas.monretm; /* 数据集res_sas.monretm为市场月持有期收益率数据集 */
set Res_sas.monretm1;
year=year(date); /* 提取出年份 */
month=month(date); /* 提取出月份 */
If 2007<=year<=2017 and Exchflg='0' and Mktflg='A';
/* Exchflg为交易所标识, Mktflg为市场标识 */
/* Exchflg='0'表示所有交易所, Mktflg='A'表示A股市场 */
keep date year month Mretmc;
run;

data Res_sas.Stk8; /* 创建要选定的股票代码数据集 */
input stkcd $6.;

```

```

cards;
000011
000012
000021
000026
000027
900957
900953
900952
;
run;

proc sort data=Res_sas.stk8; /* 对数据集res_sas.stk8排序 */
by stkcd;
run;

proc sort data=Res_sas.monret; /* 对数据集res_sas.monret排序 */
by stkcd;
run;

/* 股票代码数据集和个股月收益数据集合并，得到所选八只股票的月持有期收益率 */
data Res_sas.return;
merge Res_sas.monret Res_sas.stk8 (in=samp); /* 数据集选项in=samp创建自动变量samp */
/*
by stkcd; /* 按照股票代码合并 */
if samp; /* 保留samp自动变量存在的观测行 */
year=year(date);
month=month(date);
if 2007<=year<=2017;
keep stkcd date year month monret;
run;

proc sort data=Res_sas.return; /* 对res_sas.return数据集按照年、月进行排序 */
by year month;
run;

proc sort data=Res_sas.monretm; /* 对res_sas.monretm数据集按照年、月进行排序 */
by year month;
run;

data Res_sas.return; /* 将数据集res_sas.return和res_sas.monretm两个数据集按照年、
月进行横向合并，最终输出到数据集res_sas.return中 */
merge Res_sas.return Res_sas.monretm;
by year month;
run;

proc sort data=Res_sas.return; /* 对数据集res_sas.return进行排序 */

```

```

by stkcd year month;
run;

/* 计算单只股票期望收益 */
proc means data=Res_sas.return noprint;
by stkcd;
var monret;
output out=Res_sas.m_out;
run;

data Res_sas.m_out1a;
set Res_sas.m_out;
where _STAT_='MEAN';
keep stkcd monret;
run;

/* 计算单只股票风险度量 */
data Res_sas.m_out1b;
set Res_sas.m_out;
where _STAT_='STD';
keep stkcd monret;
rename monret=std;
label monret='月收益率标准差';
run;

/*=====*/
/*
/*===== 用非线性规划选择投资组合1 =====*/
/*
/*=====*/
/* CORR求解有效边界 */
/* ===用宏的方式选择股票=== */
%macro select(n);

proc sql;
create table res_sas.return&n as select Monret as COL&n from res_sas.return where
stkcd= "&n";
run;
quit;

%mend select;

%select(000011);
%select(000012);

```

```

%select(000021);
%select(000026);
/*
%select(000027);
%select(900957);
%select(900953);
%select(900952);
*/
run;

data res_sas.return1;
merge res_sas.return000011 res_sas.return000012 res_sas.return000021
res_sas.return000026
/* res_sas.return000027 res_sas.return900957
   res_sas.return900953 res_sas.return900952 */ ;
run;

proc corr data=res_sas.return1 cov outp=res_sas.cov_out1 nosimple;
var COL000011 COL000012 COL000021 COL000026 /* COL000027 COL900957 COL900953
COL900952 */;
title "Markowitz Model";
run;
quit;

/* 计算两只股票投资组合的平均收益率 */

/* 数据集cov_out1中, 增加变量x, 表示权重 */
data res_sas.cov_out2(drop=_name_);
set res_sas.cov_out1;
if _type_ ne 'MEAN' then delete;
do x=0 to 1 by .05;
output;
end;
rename COL000011=r000011 COL000012=r000012 COL000021=r000021
COL000026=r000026 ;
label x='投资组合的权重';
run;

/* 计算投资组合的收益 */
data res_sas.mean1;
set res_sas.cov_out2;
pfol_m1=x*r000011 + (1-x)*r000012 ;
pfol_m2=x*r000011 + (1-x)*r000021 ;
pfol_m3=x*r000011 + (1-x)*r000026 ;
pfol_m4=x*r000012 + (1-x)*r000021 ;

```



```

pfol_m5=x*r000012 + (1-x)*r000026 ;
pfol_m6=x*r000021 + (1-x)*r000026 ;
label pfol_m1='投资组合的收益' pfol_m2='投资组合的收益';
run;

proc print data=res_sas.mean1;
title 'Markowitz Model';
title2 '两种股票投资组合的收益';
run;

proc sgplot data=res_sas.mean1;
series x=x y=pfol_m1 / legendlabel='投资组合1' markers
markerattrs=(symbol=circle);
/* 用圆圈来表示投资组合1, 下同 */
series x=x y=pfol_m2 / legendlabel='投资组合2' markers
markerattrs=(symbol=plus);
/* 用加号来表示投资组合2, 下同 */
series x=x y=pfol_m3 / legendlabel='投资组合3' markers
markerattrs=(symbol=starfilled);
/* 用填满的星星表示投资组合3, 下同 */
series x=x y=pfol_m4 / legendlabel='投资组合4' markers
markerattrs=(symbol=trianglefilled);
/* 用填满的三角形表示投资组合4, 下同 */
series x=x y=pfol_m5 / legendlabel='投资组合5' markers
markerattrs=(symbol=square);
/* 用方形表示投资组合5, 下同 */
series x=x y=pfol_m6 / legendlabel='投资组合6' markers
markerattrs=(symbol=diamond);
/* 用菱形表示投资组合6, 下同 */
title1 ' ';
title2 '投资组合收益对权重的线图';
run;

/* 计算投资组合的风险 */
data res_sas.risk1;
set res_sas.mean1;
pfol_v1=x**2*0.0235028630 + (1-x)**2*0.0226811056 + 2*x*(1-x)*0.0116061625 ;
pfol_v2=x**2*0.0235028630 + (1-x)**2*0.0202963397 + 2*x*(1-x)*0.0113310676 ;
pfol_v3=x**2*0.0235028630 + (1-x)**2*0.0216188652 + 2*x*(1-x)*0.0132727814 ;
pfol_v4=x**2*0.0226811056 + (1-x)**2*0.0202963397 + 2*x*(1-x)*0.0114253837 ;
pfol_v5=x**2*0.0226811056 + (1-x)**2*0.0216188652 + 2*x*(1-x)*0.0134001233 ;
pfol_v6=x**2*0.0202963397 + (1-x)**2*0.0216188652 + 2*x*(1-x)*0.0127201411 ;
array pfol_v(6) pfol_v1-pfol_v6;
array pfol_r(6) pfol_r1-pfol_r6;
do i=1 to 6;

```

```

pfol_r(i)=sqrt(pfol_v(i));
end;
label pfol_r1='投资组合的风险' pfol_r2='投资组合的风险';
run;

/* 打印含有两只股票投资组合的风险水平 */
proc print data=res_sas.risk1;
var x pfol_r1-pfol_r6 ;
title1 ' ';
title2 '含有两只股票投资组合的风险水平';
run;

/* 投资组合风险对权重的散点图 */
proc sgplot data=res_sas.risk1;
series x=x y=pfol_r1 / legendlabel='投资组合1' markers
markerattrs=(symbol=circle);
series x=x y=pfol_r2 / legendlabel='投资组合2' markers
markerattrs=(symbol=plus);
series x=x y=pfol_r3 / legendlabel='投资组合3' markers
markerattrs=(symbol=starfilled);
series x=x y=pfol_r4 / legendlabel='投资组合4' markers
markerattrs=(symbol=trianglefilled);
series x=x y=pfol_r5 / legendlabel='投资组合5' markers
markerattrs=(symbol=square);
series x=x y=pfol_r6 / legendlabel='投资组合6' markers
markerattrs=(symbol=diamond);
title1 ' ';
title2 '投资组合风险对权重的线图';
run;

/* 投资组合收益对风险散点图 */
proc sgplot data=res_sas.risk1;
series x=pfol_r1 y=pfol_m1 / legendlabel='投资组合1' markers
markerattrs=(symbol=circle);
series x=pfol_r2 y=pfol_m2 / legendlabel='投资组合2' markers
markerattrs=(symbol=plus);
series x=pfol_r3 y=pfol_m3 / legendlabel='投资组合3' markers
markerattrs=(symbol=starfilled);
series x=pfol_r4 y=pfol_m4 / legendlabel='投资组合4' markers
markerattrs=(symbol=trianglefilled);
series x=pfol_r5 y=pfol_m5 / legendlabel='投资组合5' markers
markerattrs=(symbol=square);
series x=pfol_r6 y=pfol_m6 / legendlabel='投资组合6' markers
markerattrs=(symbol=diamond);
title1 ' ';

```

```
title2 '投资组合收益对风险的线图';
run;

/* 通过观察,得出投资组合1、投资组合2、投资组合3在既定风险下有较高的收益 */
proc sgplot data=res_sas.risk1;
series x=pfol_r1 y=pfol_m1 / legendlabel='投资组合1' markers
markerattrs=(symbol=circle);
series x=pfol_r2 y=pfol_m2 / legendlabel='投资组合2' markers
markerattrs=(symbol=plus);
series x=pfol_r3 y=pfol_m3 / legendlabel='投资组合3' markers
markerattrs=(symbol=starfilled);
title1 ' ';
title2 '投资组合收益对风险的线图(筛选后)';
run;

proc sgplot data=res_sas.risk1;
series x=x y=pfol_r1 / legendlabel='投资组合1' markers
markerattrs=(symbol=circle);
series x=x y=pfol_r2 / legendlabel='投资组合2' markers
markerattrs=(symbol=plus);
series x=x y=pfol_r3 / legendlabel='投资组合3' markers
markerattrs=(symbol=starfilled);
title1 ' ';
title2 '投资组合风险对权重的线图(筛选后)';
run;

proc sgplot data=res_sas.mean1;
series x=x y=pfol_m1 / legendlabel='投资组合1' markers
markerattrs=(symbol=circle);
series x=x y=pfol_m2 / legendlabel='投资组合2' markers
markerattrs=(symbol=plus);
series x=x y=pfol_m3 / legendlabel='投资组合3' markers
markerattrs=(symbol=starfilled);
title1 ' ';
title2 '投资组合收益对权重的线图(筛选后)';
run;

/* 有效边界的准备工作 */
proc sgplot data=res_sas.risk1;
series x=pfol_r2 y=pfol_m2 / datalabel = x legendlabel='投资组合2' markers
markerattrs=(symbol=plus);
series x=pfol_r3 y=pfol_m3 / datalabel = x legendlabel='投资组合3' markers
markerattrs=(symbol=starfilled);
title1 '搜寻投资组合的有效边界';
title2 '投资组合收益对风险的线图(筛选后)';
```

```

run;

/* 绘制有效边界 */
data res_sas.risk2;
set res_sas.risk1;
if x<0.4 then delete;
run;

/* 显示权重 */
proc sgplot data=res_sas.risk2;
series x=pfol_r2 y=pfol_m2 / datalabel = x legendlabel='投资组合2' markers
markerattrs=(symbol=plus);
series x=pfol_r3 y=pfol_m3 / datalabel = x legendlabel='投资组合3' markers
markerattrs=(symbol=starfilled);
title1 '投资组合的有效边界(权重)';
title2 '投资组合收益对风险的线图(筛选后)';
run;

/* 显示风险 */
proc sgplot data=res_sas.risk2;
series x=pfol_r2 y=pfol_m2 / datalabel = pfol_r2 legendlabel='投资组合2' markers
markerattrs=(symbol=plus);
series x=pfol_r3 y=pfol_m3 / datalabel = pfol_r3 legendlabel='投资组合3' markers
markerattrs=(symbol=starfilled);
title1 '投资组合的有效边界(风险)';
title2 '投资组合收益对风险的线图(筛选后)';
run;

/* 显示收益 */
proc sgplot data=res_sas.risk2;
series x=pfol_r2 y=pfol_m2 / datalabel = pfol_m2 legendlabel='投资组合2' markers
markerattrs=(symbol=plus);
series x=pfol_r3 y=pfol_m3 / datalabel = pfol_m3 legendlabel='投资组合3' markers
markerattrs=(symbol=starfilled);
title1 '投资组合的有效边界(收益)';
title2 '投资组合收益对风险的线图(筛选后)';
run;

title ' ';

/*=====*/
/*
/*===== 用非线性规划选择投资组合2 =====*/
/*
/*=====*/

/* IML求解有效边界 */
/* 求解得到协方差数据集convariance和平均值数据集mean */

```

```

data convvariance mean;
set res_sas.cov_out1;
if _type_ = 'COV' then output convvariance;
if _type_ = 'MEAN' then output mean;
drop _type_ _name_;
run;

%macro shicuo;
%do i=1 %to 40;
/* 下面开始用IML过程步求解最优投资组合 */
proc iml;
/* 把数据集convvariance里面所有变量都放入矩阵H中 */
use convvariance var _all_;
read all into H;
/* 把平均值数据集里面所有变量读入到收益率矩阵R中 */
use mean var _all_;
read all into R;
show H R;
return=&i*0.001; /* 自行设定的投资组合报酬率 */
Rp=return||1;
show Rp;
/* 计算投资组合权重 */
r1=(R`)||repeat(1,nrow(H),1);
show r1;
/* 权重矩阵 */
W=inv(H)*r1* inv( (r1)` * inv(H) * r1 ) * Rp`;
show W;
/* 风险矩阵 */
V=(return||sqrt(W`*H*W));
show V;
/* 将风险矩阵的两列分别命名为RET（收益）和STD（标准差/风险） */
varnames={RET STD};
create fengxian from V[colname=varnames];
/* 将数据集的两列分别命名为RET和STD */
append from V;
run;
quit;
/* 将数据集fengxian中的观测强行导入ret_std数据集 */
proc append data=fengxian base=ret_std force;
run;
quit;
%end;
%mend shicuo;

proc datasets noprint;

```

```

delete ret_std;
run;
quit;

%shicuo;

/* 绘制可行集边缘曲线 */
proc sgplot data=ret_std;
series x=STD y=RET / legendlabel='最优投资组合' markers
markerattrs=(symbol=star);
title'最小方差前沿';
run;

/* 绘制有效边界 */
data ret_std1;
set ret_std;
if RET>=0.018;
run;

proc sgplot data=ret_std1;
series x=STD y=RET / legendlabel='最优投资组合' markers
markerattrs=(symbol=star);
title'最优投资组合的有效边界';
run;

title ' ';

/*=====*/
/*
/*=====将40改为21 =====*/
/*
/*=====*/

%macro shicuo;
%do i=1 %to 21;
/* 下面开始用IML过程步求解最优投资组合 */
proc iml;
/* 把数据集convariance里面所有变量都放入矩阵H中 */
use convariance var _all_;
read all into H;
/* 把平均值数据集里面所有变量读入到收益率矩阵R中 */
use mean var _all_;
read all into R;
show H R;
return=&i*0.001; /* 自行设定的投资组合报酬率 */

```

```
Rp=return||1;
show Rp;
/* 计算投资组合权重 */
r1=(R`)||repeat(1,nrow(H),1);
show r1;
/* 权重矩阵 */
W=inv(H)*r1* inv( (r1)` * inv(H) * r1 ) * Rp`;
show W;
/* 风险矩阵 */
V=(return||sqrt(W`*H*W));
show V;
/* 将风险矩阵的两列分别命名为RET（收益）和STD（标准差/风险） */
varnames={RET STD};
create fengxian from V[colname=varnames];
/* 将数据集的两列分别命名为RET和STD */
append from V;
run;
quit;
/* 将数据集fengxian中的观测强行导入ret_std数据集 */
proc append data=fengxian base=ret_std force;
run;
quit;
%end;
%mend shicuo;

proc datasets noprint;
delete ret_std;
run;
quit;

%shicuo;

/* 绘制可行集边缘曲线 */
proc sgplot data=ret_std;
series x=STD y=RET / legendlabel='最优投资组合' markers
markerattrs=(symbol=star);
title'最小方差前沿';
run;

/* 绘制有效边界 */
data ret_std1;
set ret_std;
if RET>=0.018;
run;
```

```

proc sgplot data=ret_std1;
series x=STD y=RET / legendlabel='最优投资组合' markers
markerattrs=(symbol=star);
title'最优投资组合的有效边界';
run;

/*****对比分析*****/
data risk2;
set res_sas.risk2;
keep pfol_r2 pfol_r3 pfol_m2 pfol_m3;
label pfol_r2=' ' pfol_r3=' ' pfol_m2=' ' pfol_m3=' ';
run;
data all;
merge risk2 ret_std1 ;
label pfol_r2='风险' pfol_r3='风险' STD='风险' pfol_m2='收益' pfol_m3='收益' RET='
收益';
run;

proc sgplot data=all;
series x=pfol_r2 y=pfol_m2 / legendlabel='投资组合2' markers
markerattrs=(symbol=plus);
series x=pfol_r3 y=pfol_m3 / legendlabel='投资组合3' markers
markerattrs=(symbol=starfilled);
series x=STD y=RET / legendlabel='投资组合（IML构建）' markers
markerattrs=(symbol=star);
title'对比分析';
run;

title ' ';

/*=====*/
/*
/*===== 用非线性规划选择投资组合3 =====*/
/*
/*=====*/

%macro opt(m);
/* NLP优化四只股票组成的投资组合的权重 */
proc nlp;
minquad H,g,0; /* 确定所求为最小风险 */
/* 输入协方差矩阵 */
array H[4,4]
0.0235425303 0.0116257509 0.0113501917 0.0132951827
0.0116257509 0.0227193859 0.0114446670 0.0134227395
0.0113501917 0.0114446670 0.0203305951 0.0127416097
0.0132951827 0.0134227395 0.0127416097 0.0216553527 ;

```



```

array g[4] 0 0 0 0 ;
parms x1-x4; /* 规定权重为x1到x4 */
bounds 0<=x1-x4<= 1; /* 对权重的取值范围进行限制 */
/* 所有权重的和为1 */
lincon 1=x1+x2+x3+x4,
&m <= 0.0209075758*x1 + 0.0154492424*x2 + 0.0156386364*x3 + 0.0185151515*x4 ;
run;
quit;
/* OPTMODEL优化四只股票组成的投资组合的权重 */
proc optmodel;
var x{1..4} >=0 <=1;
num coeff{1..4,1..4}=[
0.0235425303 0.0116257509 0.0113501917 0.0132951827
0.0116257509 0.0227193859 0.0114446670 0.0134227395
0.0113501917 0.0114446670 0.0203305951 0.0127416097
0.0132951827 0.0134227395 0.0127416097 0.0216553527
];
num r{1..4}=[ 0.0209075758 0.0154492424 0.0156386364 0.0185151515 ];
minimize f = sum{i in 1..4 , j in 1..4}coeff[i,j]*x[i]*x[j];
con sum{i in 1..4}x[i]=1;
con sum{i in 1..4}r[i]*x[i]>=&m;
solve with qp;
print x;
run;
quit;
%mend opt;
%opt(0.008);
%opt(0.010);
%opt(0.012);
%opt(0.014);
%opt(0.016);
%opt(0.018);
%opt(0.020);

run;

title ' ';

/*=====*/
/*
/*===== 用非线性规划选择投资组合4 =====*/
/*
/*=====*/
/* NLP过程步优化八只股票构成的投资组合的权重 */

```

```

/* ===用宏的方式选择股票=== */
%macro select(n,m);

proc sql;
create table res_sas.return&n as select Monret as x&m from res_sas.return where
stkcd= "&n";
run;
quit;

%mend select;

%select(000011,1);
%select(000012,2);
%select(000021,3);
%select(000026,4);
%select(000027,5);
%select(900957,6);
%select(900953,7);
%select(900952,8);
run;

data res_sas.return2;
merge res_sas.return000011 res_sas.return000012 res_sas.return000021
res_sas.return000026
      res_sas.return000027 res_sas.return900957
      res_sas.return900953 res_sas.return900952 ;
run;

proc corr data=res_sas.return2 cov outp=res_sas.cov_out2 nosimple;
var x1-x8 ;
title "Markowitz Model";
run;
quit;

/* 使用PROC NLP过程步优化权重 */
%macro yue(n);
proc nlp;
minquad H,g,0; /* 确定所求为最小风险 */
/* 输入协方差矩阵 */
array H[8,8]
0.0236623396 0.0116849150 0.0114079535 0.0133628426 0.0086300872 0.0117391590
0.0139428238 0.0087103809
0.0116849150 0.0228350062 0.0115029096 0.0134910486 0.0073542574 0.0138939055
0.0130127117 0.0093163350
0.0114079535 0.0115029096 0.0204340587 0.0128064525 0.0085550014 0.0122646062

```

```

0.0127601782 0.0084790341
0.0133628426 0.0134910486 0.0128064525 0.0217655581 0.0097825423 0.0153428905
0.0143835881 0.0099116323
0.0086300872 0.0073542574 0.0085550014 0.0097825423 0.0122101665 0.0090092111
0.0096812371 0.0063703016
0.0117391590 0.0138939055 0.0122646062 0.0153428905 0.0090092111 0.0270952158
0.0229878926 0.0129685768
0.0139428238 0.0130127117 0.0127601782 0.0143835881 0.0096812371 0.0229878926
0.0291195228 0.0137789440
0.0087103809 0.0093163350 0.0084790341 0.0099116323 0.0063703016 0.0129685768
0.0137789440 0.0105924184 ;
array g[8] 0 0 0 0 0 0 0 0 ;
parms x1-x8; /* 规定权重为x1到x8 */
bounds 0<=x1-x8 <= 1; /* 对权重的取值范围进行限制 */
/* 所有权重的和为1 */
lincon 1=x1+x2+x3+x4+x5+x6+x7+x8,
&n <= 0.0209075758*x1 + 0.0154492424*x2 + 0.0156386364*x3 + 0.0185151515*x4 +
0.0099515152*x5 + 0.0209613636 *x6 + 0.0214431818 *x7
+ 0.0085136364*x8 ;
run;
%mend yue;
%yue(0.008);
%yue(0.010);
%yue(0.012);
%yue(0.014);
%yue(0.016);
%yue(0.018);
%yue(0.020);

run;

title ' ';

/*=====*/
/*
/*===== 用非线性规划选择投资组合5 =====*/
/*
/*=====*/
/* OPTMODEL过程步优化八只股票构成的投资组合的权重 */
%macro yueshu(m);

proc optmodel;
var x{1..8} >=0 <=1;
num coeff{1..8,1..8}=[
0.0236623396 0.0116849150 0.0114079535 0.0133628426 0.0086300872 0.0117391590

```

```

0.0139428238 0.0087103809
0.0116849150 0.0228350062 0.0115029096 0.0134910486 0.0073542574 0.0138939055
0.0130127117 0.0093163350
0.0114079535 0.0115029096 0.0204340587 0.0128064525 0.0085550014 0.0122646062
0.0127601782 0.0084790341
0.0133628426 0.0134910486 0.0128064525 0.0217655581 0.0097825423 0.0153428905
0.0143835881 0.0099116323
0.0086300872 0.0073542574 0.0085550014 0.0097825423 0.0122101665 0.0090092111
0.0096812371 0.0063703016
0.0117391590 0.0138939055 0.0122646062 0.0153428905 0.0090092111 0.0270952158
0.0229878926 0.0129685768
0.0139428238 0.0130127117 0.0127601782 0.0143835881 0.0096812371 0.0229878926
0.0291195228 0.0137789440
0.0087103809 0.0093163350 0.0084790341 0.0099116323 0.0063703016 0.0129685768
0.0137789440 0.0105924184
];
num r{1..8}=[0.0209075758 0.0154492424 0.0156386364 0.0185151515 0.0099515152
0.0209613636 0.0214431818 0.0085136364];
minimize f = sum{i in 1..8 , j in 1..8}coeff[i,j]*x[i]*x[j];
con sum{i in 1..8}x[i]=1;
con sum{i in 1..8}r[i]*x[i]>=&m;
solve with qp;
print x;
run;
quit;

%mend yueshu;
%yueshu(0.008);
%yueshu(0.010);
%yueshu(0.012);
%yueshu(0.014);
%yueshu(0.016);
%yueshu(0.018);
%yueshu(0.020);

run;
quit;

```