Portfolio Construction and Analytics 读书笔记

1　资产管理的介绍

2　随机变量、概率分布和重要的统计概念

3　常见的分布函数介绍

4　统计学模型

5　模型模拟

6　模型优化

本章我们主要讨论的是优化问题——在一系列限制条件下如何使得我们的模型是最优的。

6.1　优化公式

一个优化问题的数学表达式主要由以下三个部分组成：

1.一系列的决策变量（一般由一个$N \times 1$维的向量组成）

2.一个目标函数

3.一系列的限制条件$(g_i(x), h_j(x))$满足：$g_i(x) \leq 0, h_j(x) = 0$

一般来说我们的目标函数总是可以写成：

Maximize：资产预期回报率

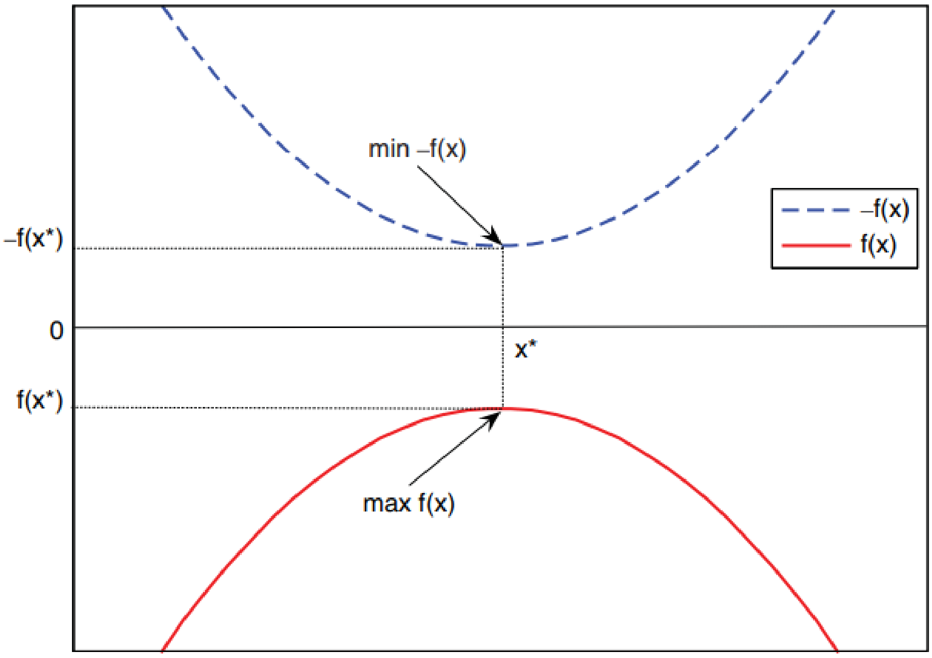
6.1.1最大化和最小化

通常来说，最优化问题可以写成如下的表达式：

$$\begin{aligned} &\text{minimize } f(x) \\ &\text{subject to } g_i(x) \leq 0 \ \ i \in \{1, \ldots, I\} \\ &\qquad\qquad h_j(x) = 0 \ \ j \in \{1, \ldots, J\}. \end{aligned}$$

同时最大化问题和最小化问题也是可以相互转化的：

$$\max_x f(x) = -\min_x -f(x)$$



6-1 最大化f(x)相当于最小化-f(x)

6.1.2局部最优和全局最优

在求解最优解的时候，部分算法（例如梯度下降法）容易陷入局部最优解而无法得到全局最优解。

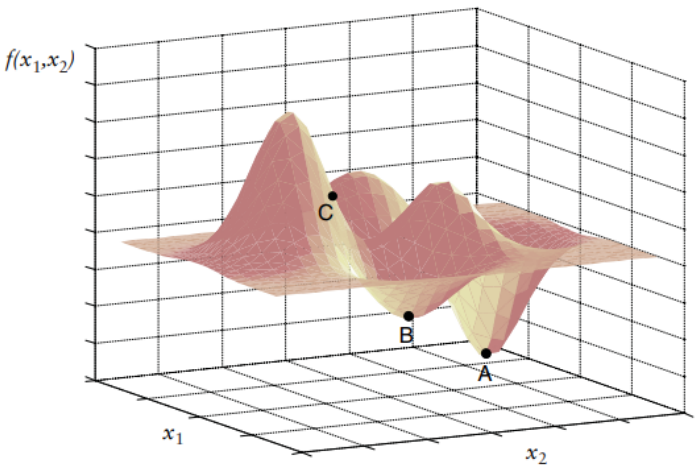


Exhibit 6.2 Global (point A) versus local (point B) minimum for a function of two variables x_1 and x_2.

6.1.3多目标优化

多目标规划是数学规划的一个分支。研究多于一个的目标函数在给定区域上的最优化。又称多目标最优化。通常记为MOP(multi-objective programming)。求解多目标线

性规划的基本思想是将多目标转化为单目标，常见的方法有理想点法、线性加权法、最大最小法、目标规划法、模糊数学解法等。

6.2　重要的优化问题

6.2.1凸优化

所谓的凸优化问题是指我们的目标函数和约束条件都是凸函数：

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0 \; i = 1, \ldots, I \\ & A\mathbf{x} \leq \mathbf{b} \end{aligned}$$

其中$f(x)$和$g_i(x)$都是凸函数。

6.2.2线性规划

线性规划（Linear programming,简称LP）是运筹学中研究较早、发展较快、应用广泛、方法较成熟的一个重要分支，它是辅助人们进行科学管理的一种数学方法。研究线性约束条件下线性目标函数的极值问题的数学理论和方法。线性规划问题的标准形式：

$$\begin{aligned} \min_{x} \quad & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

6.2.3二次规划

二次规划的一般形式可以表示为：

$$\begin{aligned} \min_{x} \quad & \tfrac{1}{2}\mathbf{x}^{\mathrm{T}}Q\mathbf{x} + \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} \leq \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

其中:

x是一个N维决策变量

Q是一个$N \times N$维矩阵

c是一个N维向量

A是一个$J \times N$维矩阵

b是一个J维矩阵

6.2.4二阶锥规划

一个二阶锥规划问题(SOCP)是指具有如下形式的图规划问题：

$$\begin{aligned} \min_x \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & \|A_i x + d_i\|_2 \le c_i^T x + e_i, \quad i = 1, \dots, I \end{aligned}$$

其中:

c是一个N维变量

A是一个$J \times N$的矩阵

b是一个J维向量

C_i是一个$I_i \times N$维矩阵

d_i是一个I_i维矩阵

e_i是标量

6.2.5整数规划

整数规划是指规划中的变量（全部或部分）限制为整数，若在线性模型中，变量限制为整数，则称为整数线性规划。

整数规划又分为：

1、纯整数规划：所有决策变量均要求为整数的整数规划

2、混合整数规划：部分决策变量均要求为整数的整数规划

3、纯0－1整数规划：所有决策变量均要求为0－1的整数规划

4、混合0－1规划：部分决策变量均要求为0－1的整数规划

6.3 一个简单的优化例子:资产分配

现在有一个资产管理者计划投资1千万美元于下面四种基金：

EXHIBIT 6.4 Data for the portfolio manager's problem.

Fund Type	Growth	Index	Bond	Money Market
Fund #	1	2	3	4
Expected return	20.69%	5.87%	10.52%	2.43%
Risk level	4	2	2	1
Max investment	40%	40%	40%	40%

他对自己的投资分配有如下的限定：投资于任何一种基金的比例不超过40%、投资于基金1和基金3的金额之和不超过总投资的60%、平均的投资风险水平不能超过2。

我们假定$x=(x_1,x_2,x_3,x_4)$表示投资于四种基金的金额。那么我们的目标函数可以写成：

$$f(x)=\mu^T x=(20.69\%)x_1+(5.87\%)x_2+(10.52\%)x_3+(2.43\%)x_4$$

我们的约束条件可以写成：

1.总的投资金额应该等于1千万美元：

$$x_1+x_2+x_3+x_4=10,000,000$$

2.投资于基金1和基金3的金额之和不超过总投资的60%：

$$x_1+x_3\le 6,000,000$$

3.平均的投资风险水平不能超过2:

$$\frac{4x_1+2x_2+2x_3+x_4}{x_1+x_2+x_3+x_4}\le 2$$

由于$x_1+x_2+x_3+x_4=10,000,000$,所以我们化简得到：

$$4x_1+2x_2+2x_3+x_4\le 20,000,000$$

4.投资于各个基金的比例不超过40%:

$$x_1\le 4,000,000, x_2\le 4,000,000, x_3\le 4,000,000, x_4\le 4,000,000$$

5.当然，投资金额是非负的：

$$x_1\ge 0, x_2\ge 0, x_3\ge 0, x_4\ge 0$$

最后我们的问题可以写成如下的形式：

$$\max_{x_1,x_2,x_3,x_4}\begin{bmatrix}0.2069 & 0.0587 & 0.1052 & 0.0243\end{bmatrix}\cdot\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}$$

$$\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix}\cdot\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=10,000,000$$

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 4 & 2 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \leq \begin{bmatrix} 6,000,000 \\ 20,000,000 \\ 4,000,000 \\ 4,000,000 \\ 4,000,000 \\ 4,000,000 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

6.4 优化算法

求解线性规划问题的基本方法是单纯形法。单纯形算法利用多面体的顶点构造一个可能的解，然后沿着多面体的边走到目标函数值更高的另一个顶点，直至到达最优解为止。虽然这个算法在实际上很有效率，在小心处理可能出现的“循环”的情况下，可以保证找到最优解，但它的最坏情况可以很坏：可以构筑一个线性规划问题，单纯形算法需要问题大小的指数倍的运行时间才能将之解出。事实上，有一段时期内人们曾不能确定线性规划问题是NP完全问题还是可以在多项式时间里解出的问题。

第一个在最坏情况具有多项式时间复杂度的线性规划算法在1979年由前苏联数学家Leonid Khachiyan提出。这个算法建基于非线性规划中Naum Shor发明的椭球法（ellip-soid method），该法又是Arkadi Nemirovski（2003年冯诺伊曼运筹学理论奖得主）和D. Yudin的凸集最优化椭球法的一般化。

理论上，“椭球法”在最恶劣的情况下所需要的计算量要比“单形法”增长的缓慢，有希望用之解决超大型线性规划问题。但在实际应用上，Khachiyan的算法令人失望：一般来说，单纯形算法比它更有效率。它的重要性在于鼓励了对内点算法的研究。内点算法是针对单形法的“边界趋近”观念而改采“内部逼近”的路线，相对于只沿着可行域的边沿进行移动的单纯形算法，内点算法能够在可行域内移动。

1984年，贝尔实验室印度裔数学家卡马卡（Narendra Karmarkar）提出了投影尺度法（又名Karmarkar’s algorithm）。这是第一个在理论上和实际上都表现良好的算法：它的最坏情况仅为多项式时间，且在实际问题中它比单纯形算法有显著的效率提升。自此之后，很多内点算法被提出来并进行分析。一个常见的内点算法为Mehrotra predictor-corrector method。尽管在理论上对它所知甚少，在实际应用中它却表现出色。

6.5 优化软件

matlab、python、c++、java 理论上都可以用于求解。

6.6 一个求解的例子

6.6.1Excel求解

这里我采用的是matlab求解。代码如下：

```
f=[-0.2069 -0.0587 -0.1052 -0.0243];
Aineq=[1 0 1 0;4 2 2 1;1 0 0 0 ;0 1 0 0 ;0 0 1 0 ;0  0 0
    1];
bineq=[6 20 4 4 4 4];
Aeq=[1,1,1,1];
beq=10;
lb=[0 0 0 0 ];
%% Start with the default options
options = optimoptions('linprog');
%% Modify options setting
options = optimoptions(options,'Display', 'off');
options = optimoptions(options,'Algorithm', 'interior-
    point');
[x,fval,exitflag,output,lambda] = ...
linprog(f,Aineq,bineq,Aeq,beq,lb,[],[],options);
```

6.6.2求解结果

运行上述代码得到求解的结果：

$$x = \begin{bmatrix} 2,000,000 \\ 0 \\ 4,000,000 \\ 4,000,000 \end{bmatrix}$$

$$fmax = 931,800$$

所以我们投资于四种基金的金额分别为：2,000,000美元，0美元，4,000,000美元，4,000,000美元。一年后我们的预期收益为931,800美元。