第 29 卷 第 1 期　　数学的实践与认识　　Vol. 29 No. 1
1999 年 1 月　　MATHEMATICS IN PRACTICE AND THEORY　　Jan., 1999

一类投资组合问题的建模与分析

陈叔平　　谭永基

(浙江大学，杭州 310027)　(复旦大学，上海 200433)

摘 要　本文介绍了 1998 年全国大学生数学建模竞赛 A 题的特点、建模与求解过程，并对参赛队的答案作出评述.

在今年的全国大学生数学建模竞赛中，陈叔平提供的赛题 (即 A 题) 被竞赛全国组委会采用.

该题是一类关于投资组合的决策问题. 构思这道题目的主要动机是它有广泛的背景和一定的难度，同时也让参赛学生有较大的发挥余地. 在人类的经济活动中，收益和风险往往是相伴而来的，不同的经营 (或投资) 项目有不同的收益和风险. 此外，每一项目都需要经营 (或投资) 成本，而可用的资金都是有限的. 因此相当大的一类决策问题可归结为在资金总额约束下，对一批成本 — 收益 — 风险各不相同的项目进行权衡并作出经营 (或投资) 抉择. 我们可以列出许多这样的问题. 下面就是一个例子.

某出版社有 n 种不同的书籍可供选择出版. 确定出版一种图书时，要支付固定数额的稿酬和制版费. 当印数达到某一数量时，包括印刷发行费在内的整个成本可视为印数的线性函数. 此外每种图书都有各自的利润率. 但当第 i 种图书印数为 y_i 时，将可能有 $g_i(y_i)$ 的量销不出去，这部分图书只能作为废纸回收处理，相当于损失. 如果 $g_i(y_i) = q_i y_i$ 其中 q_i 是常值，那么这一问题就与 A 题所给出的情形基本相同.

我们选择证券投资组合为具体背景是因为它是目前社会上的一个热点，接近大众生活，容易引起兴趣. 但与此同时也带来了一些弊病. 下面我们具体介绍这道赛题的几个特点，同时分析一下命题中的一些缺陷.

1. 交易费问题. 一项商业活动的成本一般由固定成本和可变成本两部分组成. 研制开发费、基建费等与产量无关的基本费用属固定成本. 原材料、能耗等与产量或经营规模相关的属可变成本. 赛题中给出的这类成本函数是有代表性和有趣的，开发生产一种新药或新软件如此，乘坐出租车也如此. 从数学上看，这类成本函数非凸也非连续. 因此放到目标函数和约束条件中求解优化问题有一定的困难，至少没有完全现成的方法可直接套用.

2. 风险问题. 风险是人们都能意会、十分熟悉但又没有统一标准的一个量. 为了明确起见，我们选择 $\max\limits_{1\le i\le N} q_i x_i$ 作为投资 x_i 于 $S_i, 1 \le i \le n$, 的总体风险. 它反映了风险损失一般不会同时发生的现实，也与一些文献 (如 [1]) 中所采用的一致. 本拟给出这个具体表达式，后考虑到检验参赛学生的建模能力而改成文字叙述. 由于问题的背景是证券投资，一些参赛学生从各种不同的参考资料中查阅了金融风险的其它公式 (如 Markowitz 的定义，见 [2])，从而引起了一定程度的混乱. 对金融风险的不同度量公式进行比较，在完成本题基本要求基础上才予讨论.

3. 多目标优化问题. 人们的经济行为之所以呈现多样性，主要原因是兼顾收益和风险时有不同的考虑，且各人可支配的资金也不同. 追求大的收益和小的风险构成一个两目标优化问题. 传统上，这类问题大多用某种方式化为单目标问题来求解. 例如，将一个指标作为约束来优化另一指标或将两指标加权合成一个指标. 这一过程依赖于决策者对收益和风险的理解与偏好，因而往往是不唯一的. 对固定的一个单目标优化问题求解不是主要目的，因为所得的结果对本题的具体背景并无太大的实际意义. 本题的决策包含两个层次：首先是决定买哪些股票，然后是决定各买多少. 对在不同的偏好考虑下算得的结果进行分析、比较则更有意义，也有更大的讨论和发挥余地.

4. 计算和数据问题. 由于交易成本函数的非凸, 非连续, 该题在一般情况下的求解是有一定难度的. 例如, 一维装箱问题就可化为固定风险水平条件下极大化收益这类问题的特殊情况 (见附录). 吴雄伟的论文[3] 曾对加权和模型的算法作过巧妙但较复杂的分析. 当资金总额不足以购买这 N 种证券 "起步值" 之和, 即 $M \le \sum_{i=1}^{n} u_i$ 时, 情况变得有些复杂. 因些我们加了 M "相当大" 这一条件. 当 N 和 $\{(\tilde{r}_i, q_i, p_i, u_i); 1 \le i \le N\}$ 的值由数据具体给出时, 就可构造各种数值算法. 由于命题较匆忙且宥于我国股市目前的实际情况, 赛题中给出的数据不够理想, 使得把交易成本当作线性函数处理对计算结果的影响不大. 这就造成了解题和评卷时的又一些混乱.

一、参考建模与求解过程

1. 模型的建立.

设购买 S_i 的金额为 x_i, 所需的交易费 $c(x_i)$ 为

$$c_i(x_i) = \begin{cases} 0, & x_i = 0, \\ p_i u_i, & 0 < x_i < u_i, \quad i = 1, 2, \cdots n. \\ p_i x_i, & x_i \ge u_i, \end{cases}$$

设存银行的金额为 x_0, 显然 $c_0(x_0) = 0$.

对 S_i 投资的净收益为

$$R_i(x_i) = (1 + r_i)x_i - (x_i + c_i(x_i)) = r_i x_i - c_i(x_i),$$

投资组合 $x = (x_0, x_1, \cdots, x_n)$ 的净收益为

$$R(x) = \sum_{i=0}^{n} R_i(x_i)$$

由题意, 投资的风险

$$Q(x) = \max_{1 \le i \le n} q_i x_i.$$

投资所需资金为

$$F(x) = \sum_{i=0}^{n} (x_i + c_i(x_i)).$$

因此, 问题的数学模型是一个双目标优化:

$$\min \left\{ \begin{pmatrix} Q(x) \\ -R(x) \end{pmatrix} \middle| F(x) = M, \quad x \ge 0 \right\}$$

2. 模型的约化.

上述双目标优化模型可用多种方式记为单目标优化问题, 主要有以下三种:

模型 a 固定风险水平, 优化收益:

$$\begin{aligned} &\max\ R(x), \\ &\text{s.t. } Q(x) \le k, \\ &F(x) = M \qquad x > 0. \end{aligned}$$

模型 b 固定盈利水平，极少化风险：

$$\begin{aligned}&\min\ Q(x),\\&\text{s.t. } R(x)\geq k,\\&F(x)=M,\qquad x\geq 0\end{aligned}$$

模型 c 确定投资者对风险 — 收益的相对偏好参数 $\rho>0$，求解

$$\begin{aligned}&\min\ \rho Q(x)-(1-\rho)R(x),\\&\text{s.t. } F(x)=M, x\geq 0.\end{aligned}$$

在上述三个问题中，选择 k,h 的不同水平和 ρ 的不同值进行求解就可揭示投资和风险之间相互依存规律，再根据投资者对风险的承受能力，确定投资方案.

3. 化简与求解.

因为 M 相当大，S_i 若被选中，其投资额 x_i 一般都超过 u_i，投资费用可简化为

$$C_i(x_i)=P_ix_i.$$

在进行计算时，可设 $M=1$，此时 $(1+p_i)x_i$ 可视作投资 S_i 的比例.

对固定风险的情形，问题可化为求如下线性规划问题：

$$\begin{aligned}&\max\ \sum_{i=0}^{n}(r_i-p_i)x_i,\\&\text{s.t. } q_ix_i\leq k,\qquad i=1,2,\cdots,n,\\&\sum_{i=0}^{n}(1+p_i)x_i=M.\end{aligned}$$

对于有相对偏好参数 ρ 的优化问题，引入变量 x_{n+1}，可化为如下线性规划：

$$\begin{aligned}&\min\ L(x)=\rho x_{n+1}-(1-\rho)\sum_{i=0}^{n}(r_i-p_i)x_i,\\&\text{s.t. } \sum_{i=0}^{n}(1+p_i)x_i=M,\qquad x\geq 0,\\&q_ix_i\leq x_{n+1},\qquad i=1,2,\cdots,n.\end{aligned}$$

这两个问题易于用 MATHEMATICA, MATLAB, LINDO 等数学软件来求解，对固定收益优化风险的问题也可作类似处理.

二、几点评述

多数参赛队能够正确地写出净收益的表达式. 有些参赛队用投资 S_i 的全部费用（包括投资与手续费）$y_i=x_i+c_i(x_i)$ 作基本变量. 当 M 较大时 $c_i(x_i)=p_ix_i$，从而 $y_i=x_i+p_ix_i=(1+p_i)x_i$，因此收益表达式为

$$R=\sum\frac{(r_i-p_i)y_i}{1+p_i}.$$

固定风险，优化收益问题就成为

$$\max \sum \frac{(r_i - p_i)}{1 + p_i} y_i,$$
$$\text{s.t. } \max \frac{q_i y_i}{1 + p_i} \le k,$$
$$\sum_{i=0}^{n} y_i = M, \quad y_i \ge 0, \qquad i = 1, 2, \cdots, n.$$

此问题既可以用线性规划求解，但也可以证明，将 $\frac{r_i - p_i}{1+p_i}$ 按由大至小次序排队，投资时尽可能将资金投向 $\frac{r_i - p_i}{1+p_i}$ 最大的投资项目即能获得最优解. 这样，无需求解线性规划也能获得最优解.

本问题建模中主要的缺陷表现在二个方面. 其一，计算投资总支出时遗漏交易费用，约束条件误为

$$\sum_{i=0}^{n} x_i = M.$$

另一个缺陷是未用题目所要求的 $\max\limits_{1 \le i \le n} q_i x_i$ 作为风险的度量而是采用 Markowitz 定义的风险.

在数值求解时，有些参赛队用随机模拟或枚举的方法求最优解，但计算量较大，精度也略差一些. 有些参赛队直接解非线性规划，但没有注意到交易费不是一个光滑函数，因此目标函数和约束条件均不是可微的，而直接调用一些需用导数的非线性规划程序，这样做是不可取的.

各参赛队的答卷中总的来说对模型的验证、评价和敏感性分析等方面似嫌不足，有待提高.

附 录

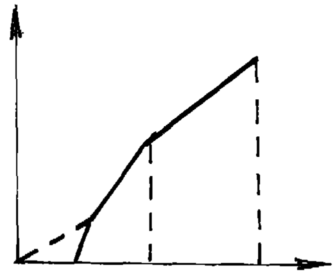
定理一 设模型 a 的最优解 (固定风险水平，优化效益) 为 $x^* = (x_0^*, x_1^*, \cdots, x_n^*)$，则至多有一个 $i \ge 1$，使得 $x_i^* \in (t_i, u_i) \cup (u_i, s_i)$. (其中符号的意义见定理二的证明)

定理二 模型 a 是 NP 难问题.

证明 构造一个一维装箱问题，已知是 NP 难问题，下面说明它可化为模型 a.

问题甲：设 M 是一个正整数，$a_1, a_2, \cdots, a_n$ 是 n 个小于 M 的正整数，$A = \{1, 2, \cdots, n\}$，求指标集 $B \subset A$，使得

$$\sum_{i \in B} a_i = \max_{B' \subset A} \left\{ \sum_{i \in B'} a_i < M \right\}.$$



问题乙：我们要讨论的投资问题.
问题甲是难问题，下面证明若问题乙存在多项式时间算法，则问题甲也存在多式时间算法. 构造一个乙问题如下：
适当选取 $p_i, q_i, r_i, 0 \le i \le n$ 等参数，可以使得 $r_0 = \frac{1}{10^n M}$，
$\bar{r}_i = 1$，$u_i = a_i, t_i = a_i - \frac{1}{10^{2n}}, s_i = a_i + \frac{1}{10^n}, 1 \le i \le n$.
其中 q_i, p_i, r_i 意义同上文.
$\bar{r}_i = r_i - u_i, s_i = \frac{k}{q_i}$. 特别地，$i = 0$ 表示银行. 如图.

设 $x^* = (x_0^*, x_1^*, \cdots, x_n^*)$ 是乙的最优解. 易知 $x_i^* \in [t_i, s_i] \cup \{0\}, 1 \le i \le n$.

令 $B = \{i | x_i^* \ne 0, i \in A\}$，则 B 即是甲问题的解. [首先可以证明 $x_i^* \notin [t_i, u_i)$，因为若 $x_i^* \in [t_i, u_i)$，则由定理一必有 $x_j^* = u_i$ 或 $x_j^* = s_i (j \ne i, j \in B)$，这时 $\sum\limits_{i \in B} x_i^*$ 不是整数，矛盾.] 所以 $x_i^* \in [u_i, s_i]$.

因对 $i \in B$ 有 $x_i^* = a_i + \tau(i)\frac{1}{10^n}, \tau(i) \in [-1, 1]$. 则

$$R(x^*) = \sum_{i\in B} R_i(x_i^*) + R_0(x_0^*) = \sum_{i\in B} \bar{r}_i x_i^* + r_0 x_0^*$$
$$= \sum_{i\in B}\left(a_i + \tau(i)\frac{1}{10^n}\right) + r_0 x_0^* = \sum_{i\in B} a_i + \left(\sum_{i\in B}\tau(i)\frac{1}{10^n} + r_0 x_0^*\right).$$

下面断言：$\{a_i | i \in B\}$ 使甲最优. 否则记甲的最优解为 $\{a_j | j \in \overline{B} \subset A\}$ 则 $\sum_{i\in\overline{B}} a_i - \sum_{i\in B} a_i \geq 1$. 因为

$$\sum_{i\in\overline{B}} a_i - R(x^*) \geq \sum_{i\in\overline{B}} a_i - \sum_{i\in B} a_i - \left(\sum_{i\in B}\tau(i)\frac{1}{10^n} + r_0 x_0^*\right)$$
$$\geq 1 - \frac{n+1}{10^n} > 0, \qquad (n \geq 1).$$

构造 $\bar{x}$ 如下： $\bar{x}_i = a_i (i \in \overline{B}), \bar{x}_0 = b - \sum_{i\in\overline{B}} a_i, \bar{x}_i = 0 (i \notin \overline{B}, i \neq 0)$. 则 $R(\bar{x}) \geq \sum_{i\in\overline{B}} a_i > R(x^*)$, 与 x^* 是乙的最优解矛盾.

所以乙有多项式时间算法，甲必有.　　证毕

参 考 文 献

[1] Cai, X., Teo, K.L., Yang, X.Q. and Zhou, X.Y., Portfolio Optimization Under Minimax Risk Measure, 1998, to appear in Management Sci.

[2] Elton, E.J. and Gruber, M.J., Modern Portfolio Theory and Investment Analysis, John Wiley & Sons, New York 1995.

[3] 吴雄伟，一类带交易成本的证券组合投资 -- 模型及算法，浙江大学博士学位论文， 1998.

Modeling and Analysis of a Kind of Investment Combination Problems

CHEN SHU-PING

(Zhejiang University, Hangzhou 310027)

TAN YONG-JI

(Fudan University, Shanghai 200433)