

基于加速遗传算法的组合证券投资决策

王 硕¹, 唐小我², 曾 勇²

(1. 东南大学经济理管学院, 南京 210096; 2. 电子科技大学管理学院, 成都 610054)

[摘要] 应用加速遗传算法解决组合证券投资决策问题, 可以克服传统遗传算法的缺点: 对搜索空间(优化变量空间)的大小变化适应能力差, 计算量大, 易出现早熟收敛, 控制参数的设置技术无明确准则指导等, 与已有结果相比, 对协方差矩阵无正定性要求, 目标函数可以推广到规模庞大, 提高预测精度等优点。

[关键词] 加速遗传算法; 组合证券; 投资决策

[中图分类号] G304 **[文献标识码]** A **[文章编号]** 1009-1742(2002)09-0059-04

1 引言

为了分散投资风险并取得适当的投资收益, 投资者往往采用组合投资方式, 把一笔资金同时投资于若干不同的证券, 投资者对其投资行为最关心的问题是预期收益率的高低和预期风险的大小。由 H.M. Markowitz 创立的现代组合证券投资理论^[1]是分散投资风险决策的有效途径。组合证券理论已引起研究界的广泛重视, 国内外学者对此进行了深入研究^[2]。笔者在前期工作的基础上, 提出用加速遗传算法解决组合证券投资决策问题, 可以取得较理想的结果。

2 组合证券投资决策模型

沿用文献 [2] 中的记号, 假设投资者选定 n 种证券进行投资, 各单项证券的收益率均值为 R_i , ($i = 1, 2, \dots, n$), E 为 n 阶协方差矩阵, 投资比例系数向量为 $W = (w_1, w_2, \dots, w_n)^T$, 预期投资收益率为 R_0 , 风险为 σ^2 。于是, 最优组合证券投资决策问题可转化为如下二次规划模型^[2]:

$$\min \sigma^2(w_1, w_2, \dots, w_n) = W^T E W, \quad (1.1)$$

$$\text{s.t. } w_1 + w_2 + \dots + w_n = 1, \quad (1.2)$$

$$w_1 R_1 + w_2 R_2 + \dots + w_n R_n = R_0, \quad (1.3)$$

$$w_1, w_2, \dots, w_n \geq 0. \quad (1.4)$$

称 σ^2 为优化准则函数, w_j ($j = 1, 2, \dots, n$) 为优化变量。

3 基于加速遗传算法的组合证券投资决策算法

加速遗传算法 (AGA) 是我国学者金菊良教授在传统遗传算法 (TGA) 基础上提出的一种优化方法^[3], AGA 可以克服 TGA 的缺点: 对搜索空间(优化变量空间)的大小变化适应能力差, 计算量大, 易出现早熟收敛, 控制参数的设置技术无明确准则指导。利用在传统遗传算法运行过程中搜索到的优秀个体, 逐步调整优化变量的搜索区间, 即可形成加速遗传算法^[3]。

加速遗传算法解决组合证券投资问题的步骤为:

Step1: 优化变量初始化变化空间的离散和二进制编码。设编码长度为 e , 把每个优化变量 w_j 的初始化变化区间 $[0, 1]$ 等分成 $2^e - 1$ 个子区间, 则

$$w_j = I_j d_j, (j = 1, 2, \dots, n), \quad (2)$$

[收稿日期] 2002-01-15; 修回日期 2002-05-15

[基金项目] 国家杰出青年科学基金资助项目 (79725002); 信息产业部软科学资助项目 (信产信科 [2000] 8号)

[作者简介] 王 硕 (1964-), 男, 安徽合肥市人, 东南大学副教授, 博士研究生

式中,子区间长度 $d_j = 1/(2^e - 1)$ 是常数,搜索步数 I_j 为 $<2^e$ 的任意十进制非负整数是变数。

经过编码,优化变量的搜索空间离散成 $(2^e)^n$ 个网格点(个体),它对应 n 个优化变量的一种可能取值状态,并用 n 个 e 位二进制数 $\{ja(j, k) | j=1, 2, \dots, n; k=1, 2, \dots, e\}$ 表示:

$$I_j = \sum_{k=1}^e ia(j, k) \cdot 2^{k-1}, (j=1, 2, \dots, n)。(3)$$

通过式(2)和式(3)的编码, n 个优化变量 w_j 的取值状态、网格点、个体、 n 个二进制数 $\{ia(j, k)\}$ 之间可建立一一对应关系。

Step2: 初始父代群体的随机生成。设群体规模大小为 q , 从上述 $(2^e)^n$ 个网格点中均匀随机选取 q 个满足约束条件式(1.2)和式(1.3)点作为初始父代群体。若在网格点中随机选取的点不满足约束条件式(1.2)和式(1.3), 则在网格点中重复产生随机点, 直至产生 q 个满足约束条件式(1.2)和式(1.3)的点, 于是, 生成 q 组(如 $q=300$) 在 $[0, 1]$ 区间上的均匀随机数, 每组 n 个, 即 $\{u(j, i) | j=1, 2, \dots, n; i=1, 2, \dots, q\}$, 经下式转换得到相应的随机搜索步数

$$I_j(i) = \text{int}(u(j, i) \cdot 2^e),$$

$$(j=1, 2, \dots, n; i=1, 2, \dots, q), (4)$$

它们可由式(3)对应二进制数 $\{ia(j, k, i) | j=1, 2, \dots, n; k=1, 2, \dots, e; i=1, 2, \dots, q\}$, 又可由式(2)与 q 组优化变量 $\{w_j(i) | j=1, 2, \dots, n; i=1, 2, \dots, q\}$ 一一对应, 把它们作为初始父代个体。

Step3: 二进制的解码和父代个体适应度评价。把父代个体编码串 $ia(j, k, i)$ 经过式(3)和式(2)解码成优化变量 $w_j(i)$, 把后者代入式(1.1)得相应的优化准则函数 σ_i^2 , 把 $\{\sigma_i^2 | i=1, 2, \dots, q\}$ 从小到大排序, 对应的优化变量 $\{w_j(i)\}$ 和二进制数 $\{ia(j, k, i)\}$ 也跟着排序, 称排序后最前面几个个体为优秀个体。记优秀个体数目为 s 。定义排序后的第 i 个父代个体的适应度函数值为

$$F_i = \frac{1}{(\sigma_i^2)^2 + 0.001}, (i=1, 2, \dots, q)。(5)$$

Step4: 父代个体的概率选择。取比例选择方式, 个体 i 的选择概率为

$$p'_i = \frac{F_i}{\sum_{i=1}^q F_i}, (i=1, 2, \dots, q)。(6)$$

令

$$p_i = \sum_{k=1}^i p'_k, (i=1, 2, \dots, q)。$$

序列 $\{p_i | i=1, 2, \dots, q\}$ 把 $[0, 1]$ 区间分成 q 个子区间, 并与 q 个父代个体一一对应。

生成 q 个随机数 $\{u(k) | k=1, 2, \dots, q\}$ 。若 $u(k) \in [p_{i-1}, p_i]$, 则第 i 个个体被选中, 其二进制数记为 $ia_1(j, k, i)$ 。同理可得另外的 q 个父代个体 $ia_2(j, k, i)$ 。于是从原父代群体中以概率 p'_i 选择第 i 个个体, 共选择两组各 q 个个体。

Step5: 父代个体的杂交。由于杂交概率 p_c 控制杂交算子应用的频率, 在每代新群体中, 有 qp_c 对串进行杂交, p_c 越高, 群体中串的更新越快, TGA 搜索新区域的机会越大^[4], 因此取定杂交率 $p_c = 1.0$ ^[3]。由 step4 得到的两组父代个体随机两两配对, 成为 q 对双亲, 将双亲的二进制数组的任意一段值互换, 得到两组子代个体。

Step6: 子代个体的变异。采用两点变异, 比单点变异更有助于增强群体的多样性^[4]。设变异率为 p_m (如 $p_m = 1.0$ ^[3]), 它是子代个体发生变异的概率。任取 step5 中的一组子代个体, 将它们的二进制数组的任意两值以概率 p_m 进行翻转(原值为 0 的变为 1, 原值为 1 的变为 0)。

Step7: 进行迭代。由 step7 得到的 q 个子代个体作为新的父代, 算法转入 step3, 进入下一次进化过程, 如此循环两次, 优秀个体将逼近最优点。

Step8: 加速循环。用第一次、第二次进化迭代所产生的优秀个体的变量变化区间作为新的初始变化区间, 算法进入 step1, 重新运行 TGA, 如此加速循环, 优秀个体的变量变化区间将逐步调整和收缩, 与最优点的距离越来越近, 直到最优个体的优化准则函数值达到稳定值, 结束整个算法的运行。此时, 把当前群体中最佳个体指定为 AGA 的结果。

组合证券投资决策加速遗传算法的计算机流程图见图 1。

4 AGA 控制参数的配置

AGA 控制参数包括编码长度 e , 父代个体数目 q , 优秀个体数目 s 和变异率 p_m 。这些参数的不同取值会对 AGA 的寻优性能产生很大影响, 要得到 AGA 运行的最优性能, 必须对这些参数进行

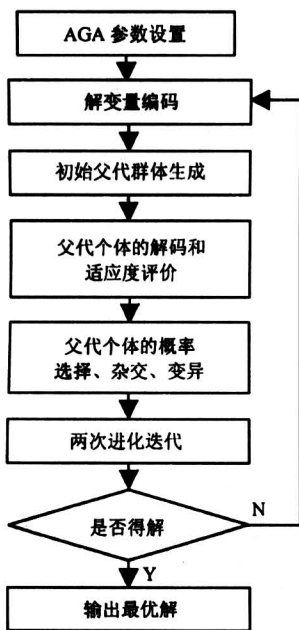


图 1 加速遗传算法计算机流程图

Fig.1 The computer flow chart of AGA model

适当配置。在 AGA 中，控制参数的配置已有明确的指导准则，可以事先设置^[3]。

1) 关于编码长度 在 TGA 中，对实变量优化问题， e 太小则解的精度太低，失去优化的意义； e 越大则优化变量搜索区间的网格分得越细，网格点就越多，搜索效率越低，需相应增加进化迭代次数或父代群体规模以提高搜索到最优解的成功率，从而导致计算量猛增，影响计算效率^[4]。因 AGA 在算法 step8 中能逐步调整优化变量的变化空间，随着 AGA 的运行网格自动细分，它采用的是一种动态编码策略， e 可取定 $10^{[3]}$ 。

2) 关于变异率 AGA 的 p_m 反映个体向其他个体网格点变迁的概率。由于其可以保证算法的收敛性^[3]， p_m 越大，搜索区域越大，优秀个体包围、接近全局最优点的机会也越大，越有利于克服早熟收敛， p_m 可取定 $1.0^{[3]}$ 。

3) 关于父代个体数目 q 和优秀个体数目 s 由 AGA 原理知， q 个父代个体是按选择概率从 $(2^e)^n$ 个优化变量空间网格中的一个随机抽样。大的群体可更好代表优化准则函数在搜索空间上的变化特性，增强基于二进制编码的杂交操作能力，阻止早熟收敛^[3]。然而，当 q 继续增加而 s 仍维持较小值不变时，真正优秀个体产生后代的概率相对减少，从而使优秀个体包围、接近最优点的机会也减

少。在 q 一定时， s 越大，优秀个体包围、接近最优点的机会越大，但收敛速度越慢。

在 TGA 中，当优秀优化变量的变化空间很大时，其编码长度、父代个体数目也必须很大，导致计算量大增，算法控制参数的设置技术也趋复杂化^[4]。而这些在 AGA 中却是不必要的。由于 AGA 利用进化迭代过程中产生的优秀个体所包含的优化准则函数在最优点附近各参数方向的变化特性的信息，来调整整体优化变量变化空间的大小，使 AGA 同时在 n 个参数方向寻优且收敛，只要 s 与 q 配合合理，就可望增强 AGA 对优化问题优化变量变化区间大小变化的适应能力。研究表明， q 和 s 的配置应满足 $s/q > q / (2^e)^{[3]}$ 。文献 [3] 建议在应用 AGA 时，取 $q = 300$ ， $s = 10$ 。

5 实例研究

例^[2]设三种证券的协方差矩阵为

$$E = \begin{bmatrix} 140 & -120 & -150 \\ -120 & 240 & 100 \\ -150 & 100 & 300 \end{bmatrix}.$$

收益率均值分别为 $R_1 = 16\%$ ， $R_2 = 18\%$ ， $R_3 = 20\%$ ，预期收益率 $R_0 = 19\%$ ，试确定非负最优投资比例系数和相应的投资风险。

在文献 [2] 中，得到如下结果： $w_1 = 0.233$ ， $w_2 = 0.034$ ， $w_3 = 0.733$ ， $\sigma_{\min}^2 = 120.91$ 。

用加速遗传算法 (AGA) 求解该问题，计算过程见表 1。于是，有 $w_1 = 0.23496530$ ， $w_2 = 0.03228606$ ， $w_3 = 0.73274860$ ， $\sigma_{\min}^2 = 120.31520000$ 。它优于文献 [2] 的结果。

表 1 用 AGA 计算组合证券投资模型的权重和最优值

Table 1 The weighs and optimal values of utilizing AGA to calculate the portfolios investment model

加速次数	优秀个体各参数的变化值			最优化 准则函数值
	w_1	w_2	w_3	
1	0.237 536 70	0.027 167 63	0.735 295 70	120.322 900 00
5	0.234 986 30	0.032 244 19	0.732 769 50	120.315 300 00
10	0.234 982 00	0.032 252 79	0.732 765 20	120.315 300 00
15	0.234 965 30	0.032 286 06	0.732 748 60	120.315 200 00
AGA 结果	0.234 965 30	0.032 286 06	0.732 748 60	120.315 200 00

6 结语

用加速遗传算法 (AGA) 求解组合证券投资

决策问题有如下特点:

1) 算法的控制参数已有明确设置, 比传统遗传算法循环次数少, 计算量小, 可以避免早熟现象发生, 易于全局收敛。

2) 协方差矩阵无正定性要求, 更接近实际问题的需要。

3) 易于处理有非卖空(即投资比例系数为非负)条件要求的组合证券投资决策问题。

4) 算法可以推广到目标函数规模庞大甚至无明确解析表达式的情形。

致谢 笔者写作得到金菊良教授的热情帮助,

谨此致谢

参考文献

- [1] Markowitz H M. Foundations of Portfolio [J]. J Finance, 1991, (2): 469~477
- [2] 唐小我. 经济预测与决策的新方法及其应用研究 [M]. 成都: 电子科技大学出版社, 1997. 58~83, 127~128
- [3] 金菊良, 丁晶. 遗传算法及其在水科学中的应用 [M]. 成都: 四川大学出版社, 2000. 42~57
- [4] 刘勇, 康立山. 非数值并行算法——遗传算法 [M]. 北京: 科学出版社, 1997. 75~121

Research on Decision Approach of Portfolios Investment Based on Accelerating Genetic Algorithm

Wang Shuo¹, Tang Xiaowo², Zeng Yong²

(1. Economic and Management School of Southeast University, Nanjing 210096, China;

2. Management School of University of Electronic Science and Technology of China,
Chengdu 610054, China)

[Abstract] It can win through traditional genetic algorithm's shortcomings by applying accelerating genetic algorithm in solving combination forecasting problems. These shortcomings include poor adapt ability in search space (i.e. optimizing variable space), large measure quantity, premature convergence, no definitude instruct rule for setting technique of control parameter, etc. The new approach does not need canonicity in forecasting error information matrix, the objective function scale may extend widely, and the forecasting precision is high.

[Key words] accelerating genetic algorithm; portfolio; investment decision

汪旭光当选为第八届国际岩石爆破破碎学术会议组委会主席

[本刊讯] 2002年8月12—15日在北京召开了第七届国际岩石爆破破碎学术会议。由来自5大洲的11名知名专家所组成的国际爆破破岩委员会,也于同期在北京先后两次举行全体委员会会议,一致投票选举我国炸药及爆破专家汪旭光院士为第八届国际岩石爆破破碎学术会议组委会主席,亚洲爆破学会筹委会主席;同时决定第八届国际爆破破碎学术会议将于2005年8月在智利首都圣地亚哥市召开。