Simulated Annealing vs Genetic Algorithm to Portfolio Selection

Joël N. Kapiamba
Faculty of Sciences
University of Kinshasa
Kinshasa, DR Congo
joskapiambal@yahoo.fr

Berthold Ulungu E.-L.
Faculty of Mechanics
High College of Applied Engineering
Kinshasa, DR Congo
ulungu_lukata@yahoo.fr

Pascal K. Mubenga
Faculty of Sciences
University of Kinshasa
Kinshasa, DR Congo
mubengakpd@gmail.com

Abstract: Portfolio selection problem consists on choice of best titles to constitute a portfolio which maximizes revenue and minimizes risk. We then deal with a bicriteria problem having two conflicting criteria to optimize simultaneously. It is well known that such a combinatorial problem is intractable with exact methods for large dimension problems. Then metaheuristics are useful to find a good approximation of the efficient set.

The purpose of this article is to evaluate the efficiency of two metaheuristic methods, namely the simulated annealing and the genetic algorithm to solve Portfolio selection problem. In order to compare both of them, metaheuristics have been implemented in the same language, Matlab. Statistical estimator and variance analysis allowed us to discriminate numerical experiments results. We observe that differences are significant. In terms of calculation time, simulated annealing appears more efficient than genetic algorithm. Variance analysis shows that both methods are independent one to another, and the result depends on the used method. In conclusion, some suggestions for future research are proposed.

Keywords: Portfolio selection, Simulated Annealing, Genetic Algorithm, Optimization of a Portfolio, Markowitz model, Analysis of Variance (ANOVA).

1. INTRODUCTION

First introduced by Markowitz [1], the mean-variance model for portfolio selection problem was the benchmark formulation in the financial field and was the basis for the development of modern financial theory over the past 60 years [2, 3].

A review of the portfolio management literature listed several articles on the subject since the early works of Markowitz [1]. In [4], the proposed algorithm based on hybridizing Genetic Algorithm and Simulated Annealing is efficient and applicable. The combination of Genetic Algorithm and Simulated Annealing is used to solve the portfolio investment problem, and the strategic restriction is introduced to the mutation process of Genetic Algorithm. It provides a high efficient decision-making method for portfolio investment, and it can also be used in other fields related to optimization. In [5], a modern theory of portfolio management is exposed. In [16], fuzzy number appears for the first time in this field. In [7, 8, 9, 10, 11, 12], authors introduced multi-objective paradigm modelling. A resolution with Goal Programming method is described in [13, 14]. In several other articles [8, 10, 13, 15, 16, 17, 18, 19, 20] various related issues are discussed, it is among other investment diversification, the portfolio management expected return of the capital markets or mutual funds [18, 21].

In his model, Markowitz considered the mathematical expectation as portfolio investment revenue (see [1, 22]) and variance as risk of investment.

In this article, we briefly present the benchmark model, the simulated annealing algorithm used with its neighborhood systems that we have built, inspired by [5] and adapted to our genetic algorithm.
problem. A numerical example will illustrate the quality of solutions, and finally, we conclude with some remarks and perspectives for future research.

2. Modeling

2.1 Classical Mathematical Formulation

Consider a portfolio $P$ for which the expected performance $E(R_p)$ and risk $\sigma_p^2$ are known. The classical management portfolio problem is formulated as follows:

$$
\begin{align*}
\max & \quad (1 - w) \sum_{i=1}^{n} r_i x_i - w \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j \\
\text{such that} & \quad \sum_{i=1}^{n} x_i = 1 \\
& \quad 0 \leq x_i \leq 1; i = 1, 2, \ldots, n
\end{align*}
$$

Where

- $n$ is the headline number of titles;
- $x_i$ is the proportion of capital invested in title $i$;
- $r_i$ is the result on title $i$;
- $r_i = E(R_i)$ is the expected result of title $i$;
- $\sigma_{ij}$ is the covariance of results of title $i$ and $j$;
- $\sigma_{ij} = \text{cov}(R_i, R_j) = E[(R_i - E(R_i))(R_j - E(R_j))]$;
- $w$ is the coefficient of risk aversion characterizing the investor; with $0 < w < 1$ (w \equiv 1 means a high risk aversion).

The mathematical formulation of portfolio selection problem given below is due to Markowitz [1]. It has become the reference formulation because it had generated other developments.

2.2 Multi-Objective Formulation

The multi-objective paradigm has emerged over the past thirty years. It is a realistic model and it allows to cohabit several conflicting objectives (see [8, 16, 23, 24, 25]).

In view of the problem $(PS)$, we see that it is a bi-criteria problem, which can be formulated as follow:

$$
\begin{align*}
\max & \quad \sum_{i=1}^{n} R_i x_i \\
\min & \quad \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j \\
\text{such that} & \quad \sum_{i=1}^{n} x_i = 1 \\
& \quad 0 < x_i < 1; i = 1, \ldots, n
\end{align*}
$$

(BCPS)

We can deduce problem $(PS)$ from $(BCPS)$ considering an aggregation of two criteria with a weighted sum considering weights $w_1$ and $w_2$ such that $w_1 + w_2 = 1$. Let $w_2 = w$, then $w_1 = 1 - w$. Therefore, the weight $w$ expresses the importance of criterion risk that corresponds to the aversion coefficient, hence the formulation $(PS)$. 
Simulated Annealing vs Genetic Algorithm to Portfolio Selection

3. **Metaheuristic Resolution Methods**

Simpressive number in literature but well known are: Simulated Annealing, Genetic Algorithm and Tabu Search. In this article, we focus on Simulated Annealing and Genetic Algorithm.

3.1 The simulated annealing algorithm (SA) [10, 23, 26]

The probably best-known trajectory method is Simulated Annealing (SA), introduced in [26]. SA was conceived for combinatorial problems, but can easily be used for continuous problems where the algorithm pseudocode is given below:

3.1.1. Algorithm

Set $R_{\text{max}}$ and $T_0$

Randomly generate current solution $x^0$

For $i=1$ to $R_{\text{max}}$ do

While stopping criteria not met do

generate $x^n \in V(x^0)$ (neighbor to current solution)

compute $\Delta = f(x^n) - f(x^0)$ and generate $u$ (uniform random variable)

if $\Delta < 0$ or $(\frac{\Delta}{T_0} > u)$ then $x^i = x^n$

end while

reduce $T_0$

end for

$x_{\text{opt}} = x^0$

3.1.2. Adaptation for BCPS

We can use simulated annealing with each of the specificity on how to create neighborhoods to solve the problem where $f$ is the criterion to be maximized.

A. Settings and input variables:

- $R$: matrix values history titles $I$;
- $n$: number of portfolio titles;
- $w$: coefficient of risk aversion;
- $x^0$: initial solution;
- $\alpha$: coefficient of cooling;
- $L$: length of the cooling bear;
- $T_0$: initial temperature;
- $N$: maximum number of iterations;
- Parameter of neighbourhood selection

B. Presentation of results:

- $x_{\text{opt}}$ as the optimal solution;
- $f_{\text{opt}}$ as optimal value of $f$:
- Result of the optimal portfolio:
- The risk of the optimal portfolio;
- The execution time.
3.2 How Create Neighborhoods

The subtlety of this step in the simulated annealing method is great. Indeed, the construction of the vicinity of a solution should take into account the nature of the problem and the characteristics of the set of eligible solutions. For our problem, the set of feasible solutions is defined by:

\[ D = \{ x \in \mathbb{R}^n / \sum_{i=1}^{n} x_i = 1, x_i \geq 0, i = 1, \ldots, n \} \]

Take a portfolio \( x^* = (x_1^*, \ldots, x_n^*) \). The objective is to construct a set of interesting portfolio belonging to \( D \), by making on \( x^* \) elementary transformation. In what follows, we propose two procedures for creating neighborhoods:

- The first is based on the definition of a rentability;
- The second is based on the correlation of returns of the portfolio securities.

3.2.1. Neighborhood with threshold: \( V_1 \)

Here is our procedure to build a neighborhood of \( x \):

A. As a preliminary point, given the trade-off between risk and return, defined in the criterion to maximize, we will introduce the concepts of the overestimated and underestimated title. To do so, we define a threshold rating, depending on the risk aversion \( w \) of the investor. We take:
   - \( r_M \) the greatest expected return on all titles:
     \[ r_M = \max_{1 \leq i \leq n} r_i \]
   - \( r_m \) the smallest expected return on all titles:
     \[ r_m = \min_{1 \leq i \leq n} r_i \]
   - \( S \) the threshold is defined by:
     \[ S = (1 - w)r_M + wr_m \]

B. Let us put \( I = \{1, \ldots, n\} \) the set of portfolio titles.
   - We define an underestimated title as a title whose return is strictly below the threshold \( S \).
     Let \( S^- \) be the set of underestimated title:
     \[ S^- = \{ i \in I / r_i < S \} \]
   - Similarly we call a title overestimated when its rentability is greater than or equal to the threshold \( S \).
     We note \( S^+ \) the set of overestimated title:
     \[ S^+ = \{ i \in I / S \leq r_i \} \]

N.B: it is trivial \( S^+ \) and \( S^- \) form a partition of \( I \).

C. A neighbourhood \( x' \) of \( x \) is constructed as follow:
   - It is drawn at random two titles \( i \) and \( j \) \((i \neq j \in I)\).
   - If the titles \( i \) and \( j \) are of the same nature, that is to say simultaneously overestimated or underestimated, \( i.e \) \((i \in S^- \text{ et } j \in S^-) \text{ ou } (i \in S^+ \text{ et } j \in S^+) \). Then transferring a portion of the amount to be invested in \( i \) to the amount invested in \( j \) is performed.
     Note \( \delta \) the rate of this transfer, we get \( x' = (x'_1, \ldots, x'_n) \) such that:
Simulated Annealing vs Genetic Algorithm to Portfolio Selection

If the titles \textit{i} and \textit{j} are different in nature
\textit{i.e} (\textit{i} \in S^+ and \textit{j} \in S^-) or \textit{i.e} (\textit{i} \in S^- and \textit{j} \in S^+) so we had an exchange between the amounts invested in titles \textit{i} and \textit{j}.

Thus, we obtain \( x' = (x'_1, ..., x'_n) \) such that
\[
\begin{cases}
  x'_k = x_k & \text{if } k \neq i, j \\
  x'_i = x_i - \delta x_i \\
  x'_j = x_j + \delta x_j
\end{cases}
\]

3.2.2. \textit{Neighborhood with covariance :} \( V_2 \)

The nature of correlations can induce a mode of generation solutions. Indeed, when two tracks are positively correlated, they tend to "grow" in the same manner which encourages some transformations on the current holdings to improve the refinements. Similarly, in the case of securities, negative correlation could yield a priori to "upgrading" current portfolio so that the test is maximized over a set of alternatives.

Note:
- if \( \text{cov}(R_i, R_j) > 0 \) return of titles \textit{i} and \textit{j} are said to be positively correlated;
- if \( \text{cov}(R_i, R_j) < 0 \) return of titles \textit{i} and \textit{j} are said to be negatively correlated;
- if \( \text{cov}(R_i, R_j) = 0 \) we say that there is no correlation between the returns of titles \textit{i} and \textit{j}.

In this dynamic, the process of creating covariance-based neighborhoods of safety measure can be described as follow: \( x = (x_1, ..., x_n) \) is the current portfolio \( x' = (x'_1, ..., x'_n) \) is generated in the portfolio \( V(x) \).

1. Pick randomly \textit{i} and \textit{j} (\textit{i} \neq \textit{j}) from \( I \).
2. If the titles \textit{i} and \textit{j} are positively correlated or uncorrelated and held an exchange between \( x_i \) and \( x_j \) in a way that the portfolio \( x' \) is the best of the two alternatives on the criterion to optimize.
3. If two titles are negatively correlated, \( x' \) it obtained by carrying out the criterion that optimizes the transfer of the set of two alternatives.

3.3. Genetic Algorithm (GA)[26]

3.3.1. Algorithm

1: Randomly generate initial population \( P \) of solutions
2: While stopping criteria not met do
3: \quad Select \( P' \subset P \) (mating pool), initialize \( P'' = \emptyset \) (set of children)
4: \quad For \( i=1 \) to \( n \) do
5: \quad \quad randomly select individuals \( x^a \) and \( x^b \) from \( P' \)
6: \quad Apply crossover to \( x^a \) and \( x^b \) produce \( x^{\text{child}} \)
7: randomly mutate produced child $x^{child}$
8: $P'' = P'' \cup x^{child}$
9: end for
10: $P = \text{survive}(P', P'')$
11: end while
12: return best solution

The genetic algorithm that we propose to solve the problem (P) was inspired by Yusen Xia et al. [20]. The mathematical formulation of the problem in [20] is substantially different from that in (P), in particular, we have brought modifications in the genetic algorithm as:

- representation of chromosomes and their mode of regeneration
- mutation procedure

We have also kept the parameters related to the method:

- probability of crossover;
- probability of mutation;
- crossover operator;
- total number of iterations.

3.2.3. Algorithm Description

A. Input Parameters

(a) $W$: coefficient of risk aversion;
(b) $n$: number of portfolio titles
(c) $N$: population size
(d) $a$: parameter of selection of the good people;
(e) $P_c$: probability of crossing;
(f) $P_m$: probability of mutation;
(g) $N_{bit}$: number of iteration of the algorithm

B. Evaluation function

An evaluation function, «eval», is built to assign to each chromosome a selection probability as big as its quality is good, regarding the criterion to maximize.

Let $\{W_1, ..., W_N\}$ population of $N$ chromosome at the current stage. We calculate the «images» of $W_i$ by the objective function of problem (P) and we sort it in a way that the index values decrease as $i$ increase. We obtain the population $\{W_1, ..., W_N\}$ identical to the previous one but verify

$W_1 \geq W_2 \geq ... \geq W_N$

Note:

$W_i \geq W_{i+1}$ means $W_i$ is better than $W_{i+1}$.

Given $a \in ]0,1[$ real parameter, the evaluation function is defined by:

$$\text{eval}(W_i) = a(i - a)^{i-1}$$

The evaluation of the chromosome is based on its rank rather than its value in the objective function. We have set the parameter $a$ such that the probability of selection does not too quickly decrease with the rank of the chromosome.
C. Chromosomes Selection

a) Calculation of the accumulated probability \( q_i \) of each chromosome \( W_i \):

\[
q_0 = 0, \quad q_i = \sum_{j=1}^{i-1} \text{eval}(W_j), \quad i = 1, \ldots, N
\]

b) Generation of a real \( r \) at random in \([0, q_N]\)

N.B: We generate first randomly a number in \([0,1]\) which is then multiplied by \( q_N \).

c) \( i^{th} \) selection of the chromosome \( W_i \), satisfying \( q_{i-1} < r \leq q_i \)

d) Repeating steps b) and c) \( N \) times to get a good selection of \( N \) people.

e) Finally we obtain the population \( \{Z_1, \ldots, Z_N\} \).

D. Crossing

a) Consider \( P_c \in [0,1] \) a parameter: probability of crossover.

b) To determine the parents it is generated randomly, \( N \) times a real \( r \) in \([0,1]\). \( Z_i \) is selected as a parent if the \( i^{th} \) generation of \( r \) we have \( r < P_c \). This produces at most \( N \) individuals, \( Z'_1, \ldots, Z'_N \), which are grouped in pairs: \( (Z'_1, Z'_2), (Z'_3, Z'_4), \ldots, (Z'_{N-1}, Z'_N) \)

c) It generates a random number \( c \in [0,1] \).

The crossing applied to the pair \( (Z'_1, Z'_2) \) gives two chromosomes

\[
X = cZ'_1 + (1-c)Z'_2 \quad \text{et} \quad Y = (1-c)Z'_1 + cZ'_2,
\]

replacing \( Z'_1 \) and \( Z'_2 \). Idem for the other selected pairs.

d) At the end of crossing we obtain the population \( \{X_1, \ldots, X_N\} \).

E. Mutation

a) We define a parameter \( P_m \in [0,1] \): mutation probability. For \( i = 1, \ldots, N \) we generate a random number \( \eta_i \in [0,1] \) The chromosome \( X_i \) will mutate if \( \eta_i < P_m \).

b) For each \( X \) selected for mutation, a random vector is generated

\[
d = (d_1, \ldots, d_n) \quad \text{which} \quad d_i \in [0,1] \]

c) The mutation consists to build the chromosome \( Z = X + d \) and make it eligible through the transformation described in the initialization procedure.

Remark

Repeating steps B) to E) \( N_{iter} \) time, each corresponding to an iteration of the genetic algorithm, and we hold the best chromosome on all iterations as the optimal solution of the problem (P).

4. Numerical Application

We give a numerical example to illustrate the resolution of the problem (P) by the two methods presented earlier. For comparison, we also give the solution of the same problem obtained by the optimization software LINGO.

Note: results of LINGO are taken from Yusen Xia et al [27]. Rentability and risk have been recalculated on the basis of portfolios from that software.

We consider indeed, the returns of 6 titles on a history of 8 periods shown in the table below.

<table>
<thead>
<tr>
<th>Titles</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>periods</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-7</td>
<td>0.04</td>
<td>0.14</td>
<td>0.13</td>
<td>0.12</td>
<td>0.18</td>
<td>0.15</td>
</tr>
<tr>
<td>t-6</td>
<td>0.07</td>
<td>0.06</td>
<td>0.13</td>
<td>0.04</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>t-5</td>
<td>0.09</td>
<td>0.08</td>
<td>0.11</td>
<td>0.18</td>
<td>0.22</td>
<td>0.08</td>
</tr>
<tr>
<td>t-4</td>
<td>0.13</td>
<td>0.15</td>
<td>0.15</td>
<td>0.13</td>
<td>0.15</td>
<td>0.06</td>
</tr>
<tr>
<td>t-3</td>
<td>0.14</td>
<td>0.11</td>
<td>0.10</td>
<td>0.19</td>
<td>0.14</td>
<td>0.13</td>
</tr>
<tr>
<td>t-2</td>
<td>0.17</td>
<td>0.13</td>
<td>0.07</td>
<td>0.16</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>t-1</td>
<td>0.21</td>
<td>0.10</td>
<td>0.14</td>
<td>0.14</td>
<td>0.08</td>
<td>0.10</td>
</tr>
<tr>
<td>t</td>
<td>0.24</td>
<td>0.11</td>
<td>0.11</td>
<td>0.11</td>
<td>0.09</td>
<td>0.09</td>
</tr>
</tbody>
</table>
4.1. Simulated Annealing Parameters

For simulated annealing, we implement two types of neighborhood defined in (3.1.1) and use the following parameters:

- The temperature $T_0 = 0.005$
- The coefficient of cooling $\alpha = 0.9$
- The length of the landing $L = 7$

We consider an initial portfolio sufficiently distinct from the optimal portfolio to highlight the performance of the algorithm.

- Initial Portfolio $X_0 = [0 \ 0 \ 0 \ 0 \ 1]$

4.2. Genetic Algorithm Parameters

With regard to the genetic algorithm, we consider:

- a population of 30 individuals
- a probability of crossing $p_c = 0.3$
- a mutation probability $p_m = 0.2$
- an evaluation function with $\alpha = 0.3$

We run the genetic algorithm on 50,000 iterations, delivering results for different values of coefficient of risk aversion.

4.3. Numerical Results

In this subsection we present the results of our simulations with the caption:

Ling : results by the Lingo Optimization software

GA : the result of the genetic algorithm

SA1 : the results of simulated annealing by neighborhood with threshold

SA2 : The result of simulated annealing by neighborhood with covariance

Table 1. Portfolio with coefficient of risk aversion $w = 1$

<table>
<thead>
<tr>
<th>Titles</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Criterion</th>
<th>Return</th>
<th>risk</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LING</td>
<td>0</td>
<td>0.2598</td>
<td>0.4659</td>
<td>0.0782</td>
<td>0</td>
<td>0.2171</td>
<td>-0.0004</td>
<td>0.1128</td>
<td>0.0004</td>
<td></td>
</tr>
<tr>
<td>SA1</td>
<td>0.0676</td>
<td>0.1423</td>
<td>0.5748</td>
<td>0.1930</td>
<td>0.0000</td>
<td>0.0222</td>
<td>-0.0002</td>
<td>0.1202</td>
<td>0.0003</td>
<td>2.61</td>
</tr>
<tr>
<td>SA2</td>
<td>0.0728</td>
<td>0.0280</td>
<td>0.6400</td>
<td>0.2329</td>
<td>0</td>
<td>0.0263</td>
<td>-0.0002</td>
<td>0.1217</td>
<td>0.0003</td>
<td>3.81</td>
</tr>
<tr>
<td>GA</td>
<td>0.0909</td>
<td>0.1425</td>
<td>0.5445</td>
<td>0.1566</td>
<td>0.0215</td>
<td>0.0440</td>
<td>-0.0002</td>
<td>0.1195</td>
<td>0.0003</td>
<td>55.6</td>
</tr>
</tbody>
</table>

Table 2. Portfolio with coefficient of risk aversion $w = 0.8$

<table>
<thead>
<tr>
<th>Titles</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Criterion</th>
<th>Return</th>
<th>risk</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LING</td>
<td>0.2925</td>
<td>0</td>
<td>0</td>
<td>0.3707</td>
<td>0</td>
<td>0</td>
<td>0.0225</td>
<td>0.1345</td>
<td>0.0018</td>
<td></td>
</tr>
<tr>
<td>SA1</td>
<td>0.3289</td>
<td>0</td>
<td>0</td>
<td>0.0346</td>
<td>0.6270</td>
<td>0.0094</td>
<td>0</td>
<td>0.0255</td>
<td>0.1339</td>
<td>0.0019</td>
</tr>
<tr>
<td>SA2</td>
<td>0.3217</td>
<td>0.0030</td>
<td>0.0361</td>
<td>0.6361</td>
<td>0.0030</td>
<td>0</td>
<td>0.0255</td>
<td>0.1339</td>
<td>0.0019</td>
<td>2.31</td>
</tr>
<tr>
<td>GA</td>
<td>0.3711</td>
<td>0.0157</td>
<td>0.0717</td>
<td>0.4896</td>
<td>0.0487</td>
<td>0.0032</td>
<td>0.0254</td>
<td>0.1324</td>
<td>0.0019</td>
<td>56.3</td>
</tr>
</tbody>
</table>

Table 3. Portfolio with coefficient of risk aversion $w = 0.5$

<table>
<thead>
<tr>
<th>Titles</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Criterion</th>
<th>Return</th>
<th>risk</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LING</td>
<td>0.3943</td>
<td>0</td>
<td>0</td>
<td>0.6057</td>
<td>0</td>
<td>0</td>
<td>0.0664</td>
<td>0.1347</td>
<td>0.0019</td>
<td></td>
</tr>
<tr>
<td>SA1</td>
<td>0.5042</td>
<td>0</td>
<td>0</td>
<td>0.4958</td>
<td>0</td>
<td>0</td>
<td>0.0665</td>
<td>0.1350</td>
<td>0.0027</td>
<td>2.67</td>
</tr>
<tr>
<td>SA2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.0658</td>
<td>0.1338</td>
<td>0.0022</td>
<td>3.67</td>
</tr>
<tr>
<td>GA</td>
<td>0.502</td>
<td>0.0093</td>
<td>0.013</td>
<td>0.4499</td>
<td>0.023</td>
<td>0.0028</td>
<td>0.0661</td>
<td>0.1342</td>
<td>0.0026</td>
<td>59</td>
</tr>
</tbody>
</table>
Simulated Annealing vs Genetic Algorithm to Portfolio Selection

5. ANALYSIS OF RESULTS

By observing the means and variances obtained in tables 1 to 5, we find some differences between our medium and to justify them, we use the variance analysis as a statistical estimator.

5.1. Applicability Condition of Test

Considering the size of the sample, the general form of the variance analysis is based on the Fisher’s test, and the normality of the distributions of the independent samples.

We will check two hypotheses:

- The null hypothesis $H_0$ corresponds to the case where the distributions follow the same normal distribution, and in other words, these differences are due to chance, the result is the same despite the method used and there is no difference between the averages found.

- The alternative hypothesis $H_1$ means there exists at least a distribution whose average deviates other means, that is to say, the methods are independent, which means that each method gives a result independently to the others and each result depends on the method used.

The fundamental equation of the variance analysis

$$
\sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2 = \sum_{i=1}^{p} n_i (\bar{y}_i - \bar{y})^2 + \sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2
$$

$$
SS_{total} = \sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2;
$$

$$
SS_{treatments} = \sum_{i=1}^{p} n_i (\bar{y}_i - \bar{y})^2;
$$

$$
SS_{error} = \sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2.
$$

Where

$SS_{total}$ = sum of the total deviation or total variation;

$SS_{treatments}$ = sum of the differences between titles;

$SS_{error}$ = sum of deviations with respect to each title.
Joël N. Kapiamba et al.

Sub $H_0$ : the method used does not influence the results; $\mu_1 = \mu_2 = \mu_3 = \mu_4$

By con $H_1$ : the method used influences the results; $\exists\{i, j\} \in \{1,2,3,4\} i \neq j / \mu_i \neq \mu_j$

$$F_{obs} = \frac{SS_{Treatments}}{MS_{Error}} = \frac{p - 1}{N - p}$$ follows the Fisher-Snedecor law

$F_{obs}$ Compared to $F_{Seal}$ read from the table of the law of Fisher -Snedecor for a fixed error risk $\alpha$ and $(p - 1, N - p)$ degrees of freedom.

- If $F_{obs} > F_{Seal}$ the hypothesis $H_0$ is rejected at the risk of error $\alpha$ .
- If $F_{obs} \leq F_{Seal}$ the hypothesis $H_0$ is accepted at the risk of error $\alpha$ .

Let $\alpha = 0.01$, $p = 6$, $N = 24$ and degrees of freedom $(5,18)$

<table>
<thead>
<tr>
<th>$w$ = risk aversion</th>
<th>$F_{obs}$</th>
<th>$F_{Seal}$</th>
<th>Accepted hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32.8347</td>
<td>4.24788</td>
<td>$H_1$</td>
</tr>
<tr>
<td>0.8</td>
<td>143.106</td>
<td>4.24788</td>
<td>$H_1$</td>
</tr>
<tr>
<td>0.5</td>
<td>14.5279</td>
<td>4.24788</td>
<td>$H_1$</td>
</tr>
<tr>
<td>0.2</td>
<td>30.0838</td>
<td>4.24788</td>
<td>$H_1$</td>
</tr>
<tr>
<td>0.01</td>
<td>17.1048</td>
<td>4.24788</td>
<td>$H_1$</td>
</tr>
</tbody>
</table>

Compared to the table above, we can conclude that our results are significant at 99% and our methods are independent for a coefficient of aversion to any hazard as well the results are convergent.

6. REMARKS AND PERSPECTIVES

This article emerging from the junction between operations research, statistics, and finance has inevitably caused some difficulties in the comprehension and mathematical modeling of certain financial parameters; it opens the way to possible research topics

We found that the procedure with neighborhood threshold (V1) gives a better result than that with covariance (V2) to the computation time, the nature of the correlation is very important in choosing the optimal portfolio for the Markowitz model, we consider it important to maintain the general idea of the procedure neighborhood V2, even see how to improve it in the future.

It would be desirable to consider the improvement of the estimators used in finance in order to reduce bias in the criteria for the expected return and portfolio risk are usually estimated by the expected value and variance. This requires that the probability distributions of these criteria be normal when Markowitz is applied; but, it happens that the random variables obtained are not always Gaussian.

Investment strategies in the era of globalization are such that so restrictive approach to the problem of the portfolio may be inadequate for investors in terms of capital gain.

As the portfolio problem is basically a portfolio of multi-criteria problem and our experiments with different values of the coefficient of aversion have shown that a portfolio can be effectively managed by simulated annealing, we hope that it is possible to address resolution to the using other methods such as MOSA (Multiple Objective Simulated Annealing), the precise cooperative approaches, approximated cooperative approaches and exact method to two phases.

The rejection of the null hypothesis or $H_0$ does not allow us to know what are the averages differ significantly. For this, the contrast method or Scheffe method associated with the variance analysis allow to answer this question.

In this work, we wrote three programs in Matlab with respect to different methods exposed and we hope that the concept of the algorithm complexity can be addressed in order to compare the three programs objectively.
Simulated Annealing vs Genetic Algorithm to Portfolio Selection

We experienced our methods in portfolio management with a sample of six and eight titles, but hopefully in the future it would be possible to see if there is convergence for a sample of 30 titles vis-à-vis of any one portfolio.

To be more realistic, it would be interesting for better modeling to take into account tax (VAT, income tax property values, etc.)

Note that our simulation is a daily fact and for significant results, various sets of parameters must be considered in order to strengthen the robustness of the proposed solution.

7. CONCLUSION

Our goal in this paper was to see how the portfolio could be managed with Simulated Annealing method and compare the result with the one of genetic algorithm one. The portfolio problem is basically a multi-criteria one, and our experiments with different values of aversion coefficient have shown that a portfolio can be managed efficiently when Simulated Annealing is used. Note that from the variance analysis, we observe that both different used methods are independent and the obtained results are related to the respective used methods.

We conclude that the above mentioned result is interesting in the mathematics and finance field. This is the proof that common metaheuristics which are used to approximate solution to hard combinatorial optimization problem, are often devoted to find optimum-close solution or potential efficient solution set. An adapted tool in solving problem that any investor may face when confronted to risky financial assets: to maximize his profit and minimize the risk.

Way for further research in this field is double: first hand to use metrics and indicators for verifying of obtained solutions quality and second and to hybridize both methods to solve portfolio selection problem.

REFERENCES


AUTHORS’ BIOGRAPHY

Joël Kapiamba Ntumba is PhD student at university of Kinshasa. His main fields are Operations Research, Decision-making Analysis, Multi-Criteria Decision-making Aid, Statistics, Multi-objective Combinatorial Optimization, Computer programming, Portfolio Management Theory, and Mathematical Programming.

Berthold Ulungu Ekunda Lukata is Professor at Institut Supérieur des Techniques Appliquées (High College of Applied Engineering), Kinshasa (Democratic Republic of the Congo). He worked at MATHRO laboratory (University of Mons) for many years and has taught at many colleges and universities in DR Congo, Algeria, Burkina Faso, and Belgium. In collaboration with Jacques Teghem, he was the first to formalize use of metaheuristics (simulated annealing, taboo search, genetic algorithm) for solving Multi-Objective Combinatorial Optimization problems. He is a fellow (executive secretary) of two Applied Mathematics societies: ROCARO and RAMAD.

Pascal Mubenga Kampotu is Professor at University of Kinshasa where he has taught Operations Research, General topology, and Functional analysis for forty years. He is a PhD from Western Michigan University (Kalamazoo, Michigan, USA, 1972). His works are essentially related to General Topology, Functional Analysis, optimization, Mathematical programming, and Graphs Theory.