SIMULATION OF MARKOWITZ’S PORTFOLIO SELECTION USING CELLULAR AUTOMATA

AGNIESZKA ULFIK
Politechnika Częstochowska

Summary


H. Markowitz has published his first paper about portfolio theory in 1952, considering two criteria: risk and rate of return. After more than half a century, this theory is still not widely applied in its clear form. The problem is numerical complexity.

Proposed solution to this problem is parallel computational environment – cellular automata with use of artificial intelligence.

There are created models of functioning cellular automata with abilities to realize computation to maximize expected income or minimize risk of stock investment. In different models cells are communicating with each other in different ways and exchanging different information. Different models are giving different results with different data. Simulation of these models, used to select optimal portfolio, are compared. Data used in the simulations come from Warsaw Stock Exchange.

In the beginning of this article is presented classical Markowitz’s portfolio analysis with all formulas used during computation. In further part are presented cellular automata with their abilities to realize computation. Later are presented simulations (with all assumptions) and its results. Simulations are done on data from Polish stock market. All simulations are executed in author’s computer program written in Borland C++ Builder. Conclusions are presented at the end of this article.

Key words: portfolio selection, portfolio analysis, portfolio management, cellular automata, simulation, stock market, Warsaw Stock Exchange

1. Introduction

In last few years, Polish stock market has strongly developed. On Warsaw Stock Exchange on 14th April 2010 were quoted 379 stock companies (www.gpw.pl). Stock market in Poland is currently strongly developed and despite the global capital market crisis it is a good place for researches of behaviour of stock prices. Apart from stock, investors on Polish capital market can invest money in bonds, futures contracts and also other financial instruments. The main aim for investors is to estimate correctly future value of securities and then choose those with the biggest expected profit. Problem is that very often those securities are characterized by high level of risk. Investors can reduce this risk by investing in more than one security. One of basic techniques of capital market analysis is portfolio analysis.
Problem of selection effective portfolio is not simple in spite of existence of theoretical solution given by Harry Markowitz in 1952. With choosing portfolio investor has to first choose securities in which he wants to invest, and then establish how the invested capital will divided among securities.

Markowitz’s portfolio analysis gives investors the best portfolio, depending on their expectations. Disadvantage of this theory is its numerical complexity. Even though we can use very fast multi threads computers and other types of parallel machines, we still cannot find the best solution using classical methods in reasonable time of computations. First of all we have to gather many data and with use of this data calculate expected rate of return and standard deviation for every security. We also need to calculate coefficients of correlation between all considered companies. There are \( \frac{n(n-1)}{2} \) coefficients of correlation, so for 50 stocks we have to count 1 225 coefficients of correlation, for 100 stocks – 4 950 coefficients and for 379 stocks quoted on 14.04.2010 on Warsaw Stock Exchange, we have to calculate 71 631 coefficients of correlation.

When we will count all necessary characteristics we have to choose portfolio. Number of possible portfolios is \( 2^n - 1 \), where \( n \) is number of stocks. For 5 stocks we can have 31 possible portfolios, for 10 – already 1 023. For 50 stocks we will have 1 125 899 906 842 623 possible portfolios, for 100 we will have \( 1.26 \times 10^{15} \) portfolios and for 379 stocks this number will reach \( 1.23 \times 10^{114} \). Additionally every portfolio can have different percentage of every stock that it includes. To solve this problem we not only need very fast computers but also special techniques that will allow to find effective portfolio.

Very often the best solution we got as a result of watching nature. The most popular usage of such processes are genetic algorithms and neural networks. Another one is a cellular automaton. Cellular automata were created in fortieth of last century to emulate processes occurring in nature. Soon after this invention, cellular automata occurred very interesting and useful also for engineers. There are still some researches done to find new application for cellular automata. One of them can be problem of selection of portfolio in stock market.

2. Cellular Automata

Cellular automata (CA) are discrete models used mainly in physics, mathematics and computability theory. They were created by Stanisław Ulam and later developed by his colleague also working at Los Alamos National Laboratory – John von Neumann. CA are structures of the same cells putted mostly common into lattice. It is usually one, two or three-dimensional and involves large number of cells (theoretically this number is infinite but reality it is impossible to implement such model). Every cell has defined type and starting value \( C(i,j,t) \). It also has its algorithm called function of conversion or evaluation rule – \( F \). This function defines what will be future value of cell, depending on values of neighbouring cells in current time.

There are two types of neighborhoods in CA: von Neumann’s and Moore’s. Von Neuman’s neighborhood is made up of four cells adjacent vertically and horizontally. Moore’s neighborhood are eight cells surrounding specific cell from right, left, top, bottom and four diagonally. In every iteration current value of every cell is calculated on base of values of adjacent cells from previous iteration, with evaluation rule, formulated as follows:
The best known example of CA is game “Life” created by John Conway. In this game every cell receives starting state: it can be active or inactive. This game simulates real environment where animals can be born or day when they don’t have enough food. In this game there are settled rules of behaviour of every cell. If it is inactive and three other active cells surround it, it comes to life so becomes active. If cell is active and in its surrounding are two or three other active cells, it stays active also. In every different case cell is inactive – it stays inactive if such she was earlier or she days if earlier conditions are not fulfilled. So simple rules leads to astounding solutions. The final effect of CA can be systematize as following: CA achieves stable state, in which nothing does not it change (all cells stay in determinate state); state of CA changes cyclically after some quantity of iteration; CA achieves chaotic state in which it is hard to find any order; in CA we can find stable local configurations with long time of life.

The cellular automata are already since a long time applied in many fields of science. In every issue it is necessary to establish three basic parameters CA as: type of cells creating CA, which means to establish what kind of information they have to contain; starting value of every cell; function of conversion which is algorithm deciding what will be the state of cells in current iteration on the base of values of neighbouring cells in previous iteration.

It has been proved that CA can simulate mechanisms occurring at the stock market and how this simulation can be useful for potential capital market investors using its collective work to choose effective portfolios, although further researches are needed, especially in time of capital markets crisis. Structure of CA simulating stock market using collective work of all cells can be organized in many ways, as it was shown before.
3. Portfolio analysis

According to theory of portfolio analysis we can predict future values of stock basing on historical quotation. From this data is counted rate of return interpreted as expected profit and standard deviation which is measure of dispersion and is interpreted as risk connected with this expected profit. Investments are usually interested in shares with large profit and low risk level. Portfolio analysis studies how these values will change if we will invest in more than one share. This theory also shows how to choose assets during constructing portfolio to get diversified risk which means that risk of portfolio would be lower than risk of shares that are this portfolio’s components.

Harry Markowitz in 1952 had published his first paper about portfolio selection. It has started a real revolution for capital markets and created completely new technique of making investment decisions today, called portfolio analysis. Although this theory was showing how to choose the best stocks to get the highest income with lowest level of risk, it was very hard to apply it in practice because of computational complexity. After few years another scientist W. Sharp has simplified this model and made portfolio analysis able to apply in practice. He has added to this theory two coefficients giving investors a clear hint which stocks are giving better results than a general tendency on considered market. As factor reflecting market tendency he has suggested stock exchange index. On stock market in Poland investors most commonly use Warsaw Stock Exchange Index for this purpose.

4. Rate of return and risk

In classical Markowitz’s model of portfolio selection the most important profiles of assets are rate of return and standard deviation. These two values have to be calculated for all shares taken by investor into consideration. Positive value of rate of return is interpreted as expected profit and negative value – as expected loss. This quantity is count in period of time \( t \) with following formula:

\[
R_t = \frac{P_t - P_{t-1} + D_t}{P_{t-1}}
\]

where:
- \( R_t \) – the rate of return in period \( t \),
- \( P_t \) – the price of securities in period \( t \),
- \( P_{t-1} \) – the price of securities in period \( t - 1 \),
- \( D_t \) – paid-out dividend in period \( t \).

Rate of return is appointed for every period of time \( t \), there for it becomes function of time. The real value of income depends on many factors and investor cannot be sure that he will get calculated profit. This is the reason that we use expression: expected rate of return. It is counted with following formula:

\[
R = \frac{\sum_{t=1}^{N} R_t}{N}
\]
where:
$R$ – expected rate of return from securities,
$R_t$ – the rate of return in period $t$,
$N$ – number of all analyzed rates of return.

Definite in this way level of profit or loss always is accompanying the investment risk. Risk in portfolio analysis is calculated using statistics. Standard deviation is interpreted as amount of risk. Below is shown formula to calculate variation of stock’s rate of return.

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (R_t - R)^2$$

where:
$S^2$ – the variation of rate of return,
$R, R_t$ – as above,
$n$ – number of all analyzed rate of return.
$S$ – standard deviation of rate of return.

5. Map of risk and income

Both values: expected income and risk can be shown on a graph. It is called map of risk and income. This graph makes possible retrieval companies with possibly lowest risk and largest income. Figure 1 shows general schema of map risk and income with division into quarters.

![Figure 1. Map of risk and income with division into quarters](image)

Source: Own study.

In the center of figure 1 we can see point that shows characteristics of index. In Warsaw Stock Exchange investors most commonly choose index WIG or WIG20 as an image of an average
situation on stock market. First quarter of map includes securities with risk and rate of return higher than those characteristics calculated for index. Securities in this quarter are interesting for aggressive investors, willing to take extra risk but also expecting additional income. Second quarter is part of map with securities characterized by lower risk than risk of index and rate of return higher than rate of return calculated for index. Those securities are the most interesting for all investors, because they most probably will give extra profit without extra risk. Third quarter contains securities with risk lower than risk calculated for index, but at the same time rate of return also lower for rate of return calculated for index. This quarter contains securities good for investors that are very careful and don’t like to take additional risk. Fourth quarter is part of map of risk and income containing securities worst for all investors, because there are securities with risk higher than risk calculated for index and rate of return lower than rate of return calculated for index.

Figure 2 represents map of risk and income for all securities quoted on Warsaw Stock Exchange counted with one year history in period from 1st April 2009 till 1st April 2010. Figure 3 represents map of risk and income for all securities quoted on Warsaw Stock Exchange counted with two years history in period from 1st April 2008 till 1st April 2010. On both figures there is also marked index WIG, as a central point of those graphs. As it was shown by many authors, too long history can include information that don’t have any influence on current situation. Because of that there is no point of calculating rate of return and standard deviation on very long quotation history, especially of late crisis on financial markets.

Figure 2. Map of risk and income for companies on Warsaw Stock Exchange and WIG with one year history, form 01.04.2009 to 01.04.2010.

Source: Own study.
On both maps of risk and income, most of stock companies have achieved standard deviation much higher than standard deviation of WIG. In one year history there has been only 1.06% of stock companies with dispersion of its quotation lower than WIG, and in two years of history there has been only 1.28% of stock companies with dispersion of its quotation lower than WIG, but at the same time dispersion of WIG was higher for two years history, because of fluctuations resulting from the crisis in global financial markets.

At the same time 53.19% of stock companies had expected rate of return lower than index WIG in one year history and 50.21% of stock companies had expected rate of return lower than index WIG in two years history, but at the same time expected rate of return of WIG was below zero for two years history, because of fluctuations resulting from the crisis in global financial markets.

Table 1 shows what percentage of securities were in adequate quarters on maps of risk and income with one and two years history of quotation.
Table 1. Percentage of securities in quarters on maps of risk and income with one and two years history

<table>
<thead>
<tr>
<th>Quarter of map risk and income</th>
<th>1 year history</th>
<th>2 years history</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>53.2%</td>
<td>49.4%</td>
</tr>
<tr>
<td>II</td>
<td>0%</td>
<td>0.8%</td>
</tr>
<tr>
<td>III</td>
<td>1.1%</td>
<td>0.4%</td>
</tr>
<tr>
<td>IV</td>
<td>45.7%</td>
<td>49.4%</td>
</tr>
</tbody>
</table>

Source: Own study.

Second quarter of map of risk and income which is the most interesting for investors is empty or almost empty. It means that investors don’t have easy task, to chose securities to invest.

6. Two element portfolios

Expected rate of return and standard deviation are two basic characteristics of individual securities. When we consider more than one security there is one more important value. It is the coefficient of correlation between two securities. It defines connection of rates of return of two stocks. The coefficient of correlation can be calculated by following formula:

$$\rho_{12} = \frac{\sum_{i=1}^{n} (R_{1i} - R_1) \cdot (R_{2i} - R_2)}{S_1 \cdot S_2}$$

where:

- $\rho_{12}$ – coefficient of correlation of rates of return of stock 1 and 2,
- $R_{1i}$ – rate of return in period $i$,
- $R_1$ – expected rate of return from first stock,
- $S_1$ – standard deviation of rate of return first stock,
- $R_{2i}, R_2, S_2$ – the same for second stock,
- $n$ – number of all analyzed rate of return.

Investors are seeking the possibility of investing capital in stocks with high rate of return. However these securities very often characterize with high level of risk. For investor will be interesting such investments, in which with growth of rate of return, the risk will go down. The portfolio analysis gives us such possibility.

Portfolio is set of stock, which we have or we want to buy. The rate of return of portfolio is the sum of rates of return individual values multiplied by their parts in investment.

$$R_p = x_A \cdot R_A + x_B \cdot R_B$$

where:

- $R_p$ – the rate of return from two component portfolio,
- $x_A$ – the part of share $A$ in portfolio,
- $R_A$ – the rate of return of stock $A$,
- $x_B, R_B$ – the same for stock $B$.  


where $0 \leq x_A \leq 1, \ 0 \leq x_B \leq 1, \ x_A + x_B = 1$.

Calculations of risk for two elements investment portfolio are more complicated. Variation of two component portfolio is defined as:

$$S_p^2 = x_A^2 \cdot S_A^2 + x_B^2 \cdot S_B^2 + 2 \cdot x_A \cdot x_B \cdot S_A \cdot S_B \cdot r_{AB}$$

where:
- $S_p^2$ – variation of investment portfolio,
- $x_A, x_B, R_A, R_B$ – as before,
- $r_{AB}$ – the coefficient of correlation between the rate of return of stock $A$ and $B$.
- $S_p$ – standard deviation of two component portfolio.

7. Many elements portfolios.

For more than two elements of portfolio, formulas for are as follow:

$$R_p = \sum_{i=1}^{N} x_i \cdot R_i$$

$$S_p^2 = \sum_{i=1}^{N} x_i^2 \cdot S_i^2 + 2 \cdot \sum_{i=1}^{N} \sum_{j=1}^{N} x_i \cdot x_j \cdot S_i \cdot S_j \cdot r_{ij}$$

Formulas above show computational complexity of problem of selection effective portfolios. Portfolio is effective when rate of return higher than any other portfolio with the same risk, or level of risk is lowest from portfolios with the same rate of return.

8. Effective portfolios

As it is written in literature, portfolio is effective when rate of return is higher than for any other portfolio with the same risk or level of risk is lowest from portfolios with the same rate of return.

Problem of selection of effective portfolio is not simple in spite of existence of theoretical solution. With choosing portfolio investor has to first choose securities in which he wants to invest, and then establish how the invested capital will divided among securities.

Sharpe measure of portfolio effectiveness value is also called the Sharpe efficiency of the portfolio. It is quotient of risk premium achieved by the portfolio and risk of the total portfolio which has been taken. It can be calculated with the following formula:

$$SH_J = \frac{E(R_J) - R_F}{E(S_J)}$$

where:
- $SH_J$ – Sharpe measure of portfolio effectiveness value,
- $E(R_J)$ – the value of the expected rate of return on portfolio $J$,
- $R_F$ – risk-free rate of return,
- $S_J$ – standard deviation of portfolio returns $J$.

The higher is the value of this measure, the better the quality of portfolio management. Portfolios lying on the line of greatest slope in the total system risk are preferred additional income.
9. Simulations

Between 1st April 2009 and 1st April 2010 on Warsaw Stock Exchange there were 470 stock companies that had enough long history, to be able to calculate their characteristics that gave reasonable information for investors. For simulation were chosen all quotation with at least 70 quotation history in examined time.

To find effective portfolio it was necessary to establish work of CA and their three basic parameters: type of cells, starting value and function of passage from one iteration to the other. In suggested solution all cells of cellular automata work collectively exchanging information between each other.

All cells of CA have one aim: choose the effective portfolio so they have to include information about composition of this portfolio and during iterations improve it. To be able to choose portfolio, they also have to have access to all input data. Every cell can be interpreted as artificial investor choosing the efficient portfolio. This artificial investor is working on his own choosing the best result but at the same time he exchanges all information with other artificial investors which are in CA arranged in grid.

At the beginning of simulation all cells of CA get randomly chosen portfolio. Sample randomly chosen portfolios used for simulations are shown on figure 4 on map risk and income. In the center of this graph there are characteristics of index WIG.

![Figure 4. Map of risk and income of randomly chosen portfolios used for simulations and index WIG](image)

Source: Own study.

Cellular automata after getting randomly chosen portfolios, calculates basic characteristics of portfolios: rate of return and standard deviation. Then every cell communicates with its neighbours with Moore’s or von Neumann’s. If in another cell there is portfolio with better characteristics (with given criteria) then cell gets this portfolio from its neighbour. If portfolios in adjoining cells
have worst characteristics (comparing with given criteria), then cell stays with its original portfolio. In next step all cells work individually by looking for new portfolio with given criteria. They get this portfolio as completely new, randomly chosen one. If new portfolio is better it stays as current one if not, then it is passed over. This is repeated until all cells in CA get the same effective portfolio.

Simulation were done in application written in Borland C++ Builder by author. In those simulations were used two-dimensional CA with 10,000 cells (100 rows and 100 columns in two dimensional grid). The aim of this simulation was achieving stable state by cellular automata, in which all cells have the same portfolio. This simulations were repeated several times with necessary number of iteration. In cells portfolios changed when new randomly chosen portfolio was better comparing to given criteria then present one.

In simulation as main aim was assumption to achieve portfolio with maximum income and risk as less as it is possible. On figure 5 we can see results of 30 simulations on map risk and income, together with index WIG. All results were unique and often did not appear in next execution of simulation even though input data were the same.

![Figure 5. Map of risk and income of portfolios chosen by cellular automata in simulations and index WIG](source: Own study)

Because aim was to have maximum income in all portfolios cellular automata has chosen, have rate of return higher than rate of return calculated for index WIG. At the same time those portfolios have lowest risk comparing to all randomly chosen portfolios.
10. Conclusions

Aim of this work was to present methods of simulation of Markowitz’s portfolio selection using cellular automata.

On 1st March 2006 on Warsaw Stock Exchange there were quotation of 248 joint-stock companies. We have chosen 10 companies to simplified simulation of portfolio selection process. Criterions of choosing portfolios were: high value of expected rate of return and low value of standard deviation. Companies and portfolio with such characteristics are most interesting for investors that want to have high level of income from its portfolio and they don’t want to take height level of risk. All characteristics were counted basing on one year history of quotation from 1st April 2009 to 1st April 2010. Longer history is very commonly used but if we consider world financial market crisis, longer history would now only disturb proper view of characteristics and current situation.

Conducted simulations confirmed that cellular automata can select portfolio on stock market, basing on classical Markowitz’s model of portfolio analysis using collective work of cells. All portfolios obtained in simulation were different in every simulation.

Bibliography

SYMULACJA DOBORU SKŁADNIKÓW DO PORTFELA INWESTYCJNEGO MARKOWITZA Z UŻYCIEM AUTOMATÓW KOMÓRKOWYCH

Streszczenie

Artykuł przedstawia symulację procesu doboru elementów do portfela inwestycyjnego z zastosowaniem teorii klasycznego portfela Markowitza, używając do tego równoległego środowiska obliczeniowego – automatu komórkowego.

H. Markowitz opublikował pierwszą pracę na temat teorii portfelowej w 1952 roku, w której to rozpatrywał dwa kryteria: ryzyko i stopę zwołto. Po ponad pół wieku, ta teoria nadal nie jest szeroko stosowana w czystej postaci. Problem stanowi złożoność obliczeniowa.

Zaproponowanym rozwiązaniem tego problemu jest równoległe środowisko obliczeniowe – automat komórkowy z użyciem sztucznej inteligencji.

Stworzone zostały modele funkcjonowania automatu komórkowego ze zdolnością do wykonywania obliczeń prowadzących do maksymalizacji oczekiwanej stopy zwołto lub minimalizacji ryzyka inwestycji giełdowej. W różnych modelach komórki komunikują się ze sobą w różny sposób, wymieniając tym samym informacje. Różne modele dają różne rezultaty z różnymi danymi wejściowymi. Porównano symulacje przeprowadzone zgodnie z poszczególnymi modelami. Dane użyte do symulacji to dane pochodzące z Giełdy Papierów Wartościowych w Warszawie.

Na początku artykułu jest przedstawiona klasyczna teoria analizy portfela Markowitza wraz ze wszystkimi wzorami używanymi w trakcie obliczeń. W dalszej części przedstawiono zagadnienie automat komórkowego wraz ze zdolnością do przeprowadzenia obliczeń, później zaprezentowano symulacje (ze wszystkimi założeniami) oraz ich rezultaty. Symulacje są przeprowadzone na danych pochodzących z polskiej giełdy. Wszystkie obliczenia są przeprowadzone przez autorski program komputerowy napisany w środowisku Borland C++ Builder. Wyniki zaprezentowano na końcu artykułu.

Słowa kluczowe: wybór portfela, analiza portfelowa, zarządzanie portfelem inwestycyjnym, automat komórkowy, symulacje, giełda, Giełda Papierów Wartościowych w Warszawie

Agnieszka Ulfik
Instytut Ekonometrii i Informatyki
Wydział Zarządzania
Politechnika Częstochowska
e-mail: ulfik@zim.pcz.pl