A Hybrid Portfolio Selection Model based on Grey Relational Analysis and Feature Selection: Evidence on Listed Firms in Tehran Stock Exchange

1Milad Lajevardi and 2Farshad Faezy Razi

1Department of Management, Semnan Branch, Islamic Azad University, Semnan, Iran.
2Department of Management, Semnan Branch, Islamic Azad University, Semnan, Iran.

INTRODUCTION

Identifying the ‘‘best’’ portfolio of assets for an individual investor is one of the principal challenges in the world of finance. A portfolio is a collection of investments held by an institution or private individual. Portfolio selection is about choosing which assets and in what proportion will best respect the investor’s preferences for achieving an expected return assuming a minimum risk. Objective: write the main objective for your paper. Results: The purpose of stock portfolio selection is how to allocate the capital to a large number of stocks in order to bring a most profitable return for investors. The main aim of this research is selection optimized portfolio using grey relational analysis (GRA) and feature selection methods. First, this paper used of feature selection method to identify key indicators as criteria for recognition of proper firms. Second, uses of grey relational analysis (GRA) method to ranking of examined firms for selecting a basket of stocks that has the best performance. 50 companies were selected as the sample among listed companies in Tehran Stock Exchange during 2000 to 2013. Conclusion: The Results of this study indicate that a grey relational analysis (GRA) technique is employed to specify an appropriate weighting of the selected stocks such that the portfolio rate of return is maximized. Finally, was found that, both of (GRA) and feature selection methods are fit.

portfolio selection model. Solving the portfolio selection problem based on Markowitz’ model and, simultaneously, considering minimum transaction lots are of practical significance. However, it appears that no methods in the past solving the portfolio selection problem with minimum transaction lots were based on Markowitz’ model. Deng (1982) founded context of gray relations based on the theory of systems. This method consults togetherness among components of one system and reference series (Deng, 1988; Huang et al., 2008). This theory is used to solve ambiguous problems and the problems having disconnected and incomplete data. It provides satisfactory and popular outputs by a little data and with many changes in criteria. Gray theory, like fuzzy theory, is an effective mathematic model to solve indefinite and ambiguous problems. This theory is used in many fields and has been utilized in the field of solving multi criterion decision making problems named gray relational analysis. Gray relational analysis being one of gray relation is used for solving complex relations between factors and variations for solving problems. Theory of Gray systems is an algorithm that analyzes non-logical relations of one system members with a reference member and it includes capability of solving multi criterion decision making problems. In many research problems, such as pattern recognition, it is important to choose a group of set of attributes with more prediction information. That is, if the number of irrelevant or redundant features is reduced drastically, the running time of a learning algorithm is also reduced. Moreover, a more general concept can be yielded. Performing feature selection can lead to many potential benefits, which are facilitating data visualization and data understanding, reducing the measurement and storage requirements, reducing training and utilization times, defying the curse of dimensionality to improve prediction performances, etc. (Guyon, Elisseeff, 2005; Kim, 2006; Mladenic ‘Grobelnik, 2003).

By using of grey relational analysis (GRA) and feature selection methods in this research having been a little focused and ignored in previous studies; we are going to introduce a hybrid model for optimal selection of portfolio. The purpose of stock portfolio selection is how to allocate the capital to a large number of stocks in order to bring a most profitable return for investors. Therefore, the main aim of this research is selection optimized portfolio using grey relational analysis (GRA) and feature selection methods. First, this paper used of feature selection method to identify key indicators as criteria for recognition of proper firms. Second, uses of grey relational analysis (GRA) method to ranking of examined firms for selecting a basket of stocks that has the best performance.

**Literature:**

The main researches that conducted in order to stock selection issue, used of some models such as: Markowitz, DEA, AHP, MCDM, Electra and TOPSIS. Furthermore, there are a few researches that used of gray relational analysis (GRA) for portfolio selection. This section reviews the literature of previous work that used of above models.

Beshkooh and Afshari (2012), used of AHP and GRA theory in Tehran stock exchange to portfolio selection and encountered some indicators such as: Stock Prices, Operating income ratio, Company size, Current Ratio, Inventory turnover, weighted average index. Tasi and Hsio (2012) investigated Genetic algorithm in Taiwan stock exchange to portfolio selection and used of some variables i.e. financial variables, Macroeconomic variables, Stock price information, Technical indicators.

Lin and liu (2006), surveyed Taiwanese investment funds Data on the basis of Markowitz, Fuzzy multi-objective decision making models and encountered some indicators such as: Risk, Monthly returns, Return covariance between assets per year. Wang (2003) investigated Fuzzy rough set model in Taiwan stock exchange to portfolio selection and used of some variables such as: Stock Prices, Trends Stock, Mining Rules, Details Buy / Sell. Huang (2012) used of Genetic algorithms, Support vector regression in Taiwan stock exchange and applied of some indicator i.e. Profitability Growth, Share prices, Liquidity. Bulgarca (2012) investigated TOPSIS model in Istanbul stock exchange to portfolio selection and used of some variables i.e. Current Ratio, Total debt ratio, Debt, Current assets, Turnover,

Turnover of fixed assets, Profit (loss), net. Yi and et al. (2010) surveys 8 firms of chines steel industry and implies of AHP and GRA models. Furthermore, they used of several indicators such as: Profitability, Profitability to shareholders, Asset Management. Chen and hirasawa (2010) used of GRA and Genetic Network Programming in Tokyo stock exchange to portfolio selection and encountered some indicators such as: Return, Budget and Profitability. See summary of the literature in table 1.

On the basis above literature review, this paper used of a hybrid model to portfolio selection with using of GRA and feature selection methods. Therefore, current study has a new point of view to portfolio selection problem. In this paper used some indicators such as: company size, ROE, ROA, P/BV, stock price, Current Ratio, Liquidity ratio, Operating Profit, Net profit to sales.

**Methods:**

**Grey relational analysis (GRA):**

Grey system theory, which was presented by Deng in 1982, is a new mathematic theory that was born by idea of grey set theory. It is one of efficient methods that are used to solve uncertainty and vogue problems
under discrete data and incomplete information. In grey system theory, according to information degree, if the system information is fully known, the system is called a white system, while the system information is unknown; it is called a black system. We give concept a grey system as Figure 1.

Table 1: Summary of the literature

<table>
<thead>
<tr>
<th>#</th>
<th>Author(s)</th>
<th>Indicator(s)</th>
<th>Method(s)</th>
<th>Nature of decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Beshkooh and Afshari</td>
<td>Stock Prices, Operating income ratio, Company size, Current Ratio, Inventory turnover, weighted average index</td>
<td>AHP, GRA</td>
<td>Non Fuzzy</td>
</tr>
<tr>
<td>2</td>
<td>Tasi and Hsio</td>
<td>Financial variables, Macroeconomic variables, Stock price information, Technical indicators</td>
<td>Genetic algorithm</td>
<td>Non Fuzzy</td>
</tr>
<tr>
<td>3</td>
<td>Lin and liu</td>
<td>Risk, Monthly returns, Return covariance between assets per year</td>
<td>Markowitz, Fuzzy multi-objective decision making</td>
<td>Fuzzy</td>
</tr>
<tr>
<td>5</td>
<td>Huang (2012)</td>
<td>Profitability, Growth, Share prices, Liquidity</td>
<td>Genetic algorithms, Support vector regression</td>
<td>Non Fuzzy</td>
</tr>
<tr>
<td>6</td>
<td>Bulgurca (2012)</td>
<td>Current Ratio, Total debt ratio, Debt, Current assets, Turnover, Turnover of fixed assets, Profit (loss), net</td>
<td>TOPSIS</td>
<td>Non Fuzzy</td>
</tr>
<tr>
<td>7</td>
<td>Yi and et al. (2010)</td>
<td>Profitability, Profitability to shareholders, Asset Management</td>
<td>AHP, GRA</td>
<td>Non Fuzzy</td>
</tr>
<tr>
<td>8</td>
<td>Chen and hirasawa (2010)</td>
<td>Return, Budget, Profitability</td>
<td>GRA, Genetic Network Programming</td>
<td>Non Fuzzy</td>
</tr>
</tbody>
</table>

Fig. 1: the concept of a grey system.

A system with partial information known and partial information unknown is grey system. The theory includes five major parts, which include grey prediction, GRA, grey decision, grey programming, and grey control (Tzeng and Huang, 2011).

The GRA is an important approach of grey system theory in the application of estimating alternatives through decision attributes. In GRA, the data that contain same features are regarded as a sequence. As a tool of quantitative and qualitative analyses, the GRA can be used to measure the relation between two sequences by calculation their correlative degrees, which is called grey relational grade (GRG). The GRG is expressed by a scalar between 0 and 1. Up to now, the method used in resolving MADM issues (Li et al., 2009). The main idea of GRA is first transforming the performance of all alternatives into a comparability sequence. This step is called grey relational generating. According to these sequences, a reference sequence (ideal target sequence) is defined. Then, grey relational coefficient between all comparability sequences and the reference sequence is calculated. In the final step, based on these grey relational coefficients, the GRG between the reference sequence and every comparability sequences is calculated. If a comparability sequence translated from an alternative has the highest GRG between the reference sequence and that alternative will be the best choice (Kuo et al., 2008). The procedures of GRA are shown in Figure 2. The detail of the proposed GRA procedure is presented below:
Step 1 Determination of comparability sequences and reference sequence. Let $X_1~X_m$ be comparability sequences. Then comparability matrix $D$ is formulated as follows:

$$D = \begin{bmatrix} X_{11} & \cdots & X_{1m} \\ \vdots & \ddots & \vdots \\ X_{m1} & \cdots & X_{mm} \end{bmatrix}$$

Let $y_1 = (y_1(1), y_1(2), \ldots, y_1(n))^T$ be reference sequence. The GRGs between the comparability sequences and reference sequence denote the relation degrees between the Technology characteristics and Alternatives.

Step 2 Data processing.
Grey data processing must be performed before GRGs can be calculated. A series of various units must be transformed to be dimensionless. This step is called 'grey relational generating'. The data can be treated with the following three situations. If the value of the original sequence is larger-the-better, it can be normalized by

$$X_i'(k) = \frac{x_i(k) - x_i^{\text{min}}(k)}{x_i^{\text{max}}(k) - x_i^{\text{min}}(k)}$$

If the value of the original sequence is smaller-the-better, it can be normalized by

$$X_i'(k) = \frac{x_i^{\text{max}}(k) - x_i(k)}{x_i^{\text{max}}(k) - x_i^{\text{min}}(k)}$$

But, if there is 'a specific target value', the original sequence is normalized using

$$X_i'(k) = 1 - \frac{x_i - x_i(k)}{x_t}$$

Where $x_t$ is the specific target value. After normalization, $D$ becomes matrix $D'$,

$$D = \begin{bmatrix} X_{11}' & \cdots & X_{1m}' \\ \vdots & \ddots & \vdots \\ X_{m1}' & \cdots & X_{mm}' \end{bmatrix}$$

Similarly, $Y_1$ can be normalized as $Y_1'$,

$$Y_1 = (y_1'(1), y_1'(2), \ldots, y_1'(n))^T.$$

Step 3 Calculations of grey relational coefficients. The grey relational coefficient is defined as follows (Dhas and Satheesh, 2012):

$$\gamma(y_i'(k), x_i'(k)) = \frac{\delta_{i_{\text{min}}} + \varepsilon \delta_{i_{\text{max}}}}{\delta_i'(k) + \varepsilon \delta_{i_{\text{max}}}},$$

where $\delta_{i_{\text{min}}}$ is the deviation sequence of reference sequence $y_1'(k)$ and comparability sequence $x_i'(k)$, $\delta_{i_{\text{max}}} = y_1'(k) - x_i'(k)$, $\delta_{i_{\text{max}}} = \max \delta_{i_{\text{min}}}$, $\delta_{i_{\text{min}}} = \min \min \delta_{i_{\text{min}}}$, $\varepsilon$ is the distinguishing coefficient, $\varepsilon \in [0, 1]$.

Step 4 Determination of GRG. After the grey relational coefficients are got, the mean of the coefficients is often used as the GRG.

$$\gamma(Y_1, X_i) = \frac{1}{n} \sum_{k=1}^{n} \gamma(y_i'(k), x_i'(K)).$$
Normalization is performed to get the final GRG.

\[ \gamma(Y_1, X_i) = \frac{\gamma(Y_1, X_i)}{\sum_{i=1}^{n} \gamma(Y_1, X_i)} \]  

Then \( \gamma(Y_1, X_i) \) is the relation degree between the alternatives and attributes (Li et al., 2008).

**Fig. 2:** Gray relational analysis.

Fig 3 shows the output of GRA method. And table 1 shows ranks and degree of confidence.

**Feature selection:**

Feature selection is one of the two important factors contributing to the performance of a prediction model for any classification problem. The objectives of feature selection are three-fold: (a) better performance, (b) faster and more cost-effective models, and (c). deeper insight into the underlying processes. (Guyon & Elisseeff, 2003). In many research problems, such as pattern recognition, it is important to choose a group of set of attributions with more prediction information. That is, if the number of irrelevant or redundant features is reduced drastically, the running time of a learning algorithm is also reduced. Moreover, a more general concept can be yielded. Performing feature selection can lead to many potential benefits, which are facilitating data visualization and data understanding, reducing the measurement and storage requirements, reducing training and utilization times, defying the curse of dimensionality to improve prediction performances, etc. [Guyon, Elisseeff, 2005] [Kim, 2006] [Mladenic’, Grobelnik, 2003]. see feature selection process in fig 3 that run with clementine software.

**Fig. 3:** Feature selection process by clementine software.

**Fig. 3:** Grey relational analysis to ranking of stocks.

<table>
<thead>
<tr>
<th>Stock</th>
<th>A5</th>
<th>A10</th>
<th>A16</th>
<th>A19</th>
<th>A4</th>
<th>A9</th>
<th>A2</th>
<th>A11</th>
<th>A1</th>
<th>GRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>A8</td>
<td>0.4246</td>
<td>0.4246</td>
<td>0.4276</td>
<td>0.4293</td>
<td>0.4306</td>
<td>0.4316</td>
<td>0.4397</td>
<td>0.4411</td>
<td>0.4707</td>
<td>0.9898</td>
</tr>
<tr>
<td>A15</td>
<td>0.4130</td>
<td>0.4136</td>
<td>0.4148</td>
<td>0.4163</td>
<td>0.4165</td>
<td>0.4166</td>
<td>0.4184</td>
<td>0.4209</td>
<td>0.4218</td>
<td>0.4226</td>
</tr>
<tr>
<td>Rank</td>
<td>20</td>
<td>19</td>
<td>18</td>
<td>17</td>
<td>16</td>
<td>15</td>
<td>14</td>
<td>13</td>
<td>12</td>
<td>11</td>
</tr>
</tbody>
</table>
Conclusion:

The main aim of this research was selection optimized portfolio using grey relational analysis (GRA) and feature selection methods. First, this paper used of feature selection method to identify key indicators as criteria for recognition of proper firms. Second, uses of grey relational analysis (GRA) method to ranking of examined firms for selecting a basket of stocks that has the best performance. The Results of this study indicate that a grey relational analysis (GRA) technique is employed to specify an appropriate weighting of the selected stocks such that the portfolio rate of return is maximized. Finally, was found that, both of (GRA) and feature selection methods are fit. Performing feature selection can lead to many potential benefits, which are facilitating data visualization and data understanding, reducing the measurement and storage requirements, reducing training and utilization times, defying the curse of dimensionality to improve prediction performances, etc. in this paper used of this method to reduce indicators and lead to bounded group of indicators such as: company size, ROE, ROA, P/BV, stock price, Current Ratio.

REFERENCES

