The international financial crisis of September 2008 and May 2010 showed the importance of liquidity as an attribute to be considered in portfolio decisions. This study proposes an optimization model based on available public data, using Markov chain and Genetic Algorithms concepts as it considers the classic duality of risk versus return and incorporating liquidity costs. The non-linear models were tested using Genetic Algorithms with twenty five Brazilian stocks from 2007 to 2009. The results suggest that this is an innovative development methodology and useful for developing an efficient and realistic financial portfolio, as it considers many attributes such as risk, return and liquidity.

Keywords: Investment management, Markov chain, Genetic Algorithms
1 Introduction

The fundamentals of the Modern Finance Theory are represented by articles written by Markowitz (1952) and Sharpe (1964). Markowitz broke the paradigms of portfolio selection that considered only the return aspect. His proposed formulation based on the risk-return duality, explains why diversification is an advantage when it comes to portfolio selection and demonstrates that there is an optimal mix of assets in a portfolio that achieves both maximum return with a minimum risk.

Markowitz formulated the variance (or risk) theory of a generic portfolio composed of $n$ assets and showed that it depends on the variances of individual assets and the covariances between pairs of assets involved, as originally published in the following formula:

$$V = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} X_i X_j$$  \hspace{1cm} (1)

Where:

- $X_i$ = asset participation in the portfolio
- $\sigma_{ij}$ = covariance between asset $i$ and asset $j$
- $n$ = number of assets

Sharpe (1964) developed the fundamentals of asset pricing by taking into account the conclusions of Markowitz portfolio risk. Among its conclusions, he emphasizes that there is a linear relationship between the rates of return on assets and their covariance with the market portfolio. This relationship is expressed by beta ($\beta$), a standardized covariance to the market portfolio variance. Therefore, there is a linear relationship between the return on assets and $\beta$ defined by:

$$\bar{R} = R_f + \beta \times (\bar{R}_m - R_f)$$ \hspace{1cm} (2)

Where:

- $\bar{R}$ = asset expected return
- $R_f$ = risk-free rate
- $\beta$ = beta of the asset
- $\bar{R}_m$ = market expected return

According to the Modern Portfolio Theory, the risk of a portfolio can be divided into two components: (i) a factor that affects a large number of assets, each with a higher or lower intensity, called systematic and (ii) a factor that specifically affects a single asset or a small group of assets, called unsystematic or specific (ROSS et al. 1999).

Over time, operational and conceptual problems have been identified in the original formulation of Markowitz. The most important are:

- a) There are computational difficulties related to solving large-scale quadratic programming problems (KONNO AND YAMAZAKI, 1991; YOUNG, 1998; PARRA et al. 2001);
- b) Generally, the portfolios obtained by the original formulation concentrate on few assets, which is against the idea of diversification (JANA et al. 2009);
- c) The absence of transaction costs and liquidity (or illiquidity) can result in inefficient portfolios (ARNOTT AND WAGNER, 1990; AMIHUD AND MENDELSON, 1991);
- d) In large portfolios the model would suggest the purchase of a small fraction of assets, often lower than the minimum traded in the market (KONNO AND YAMAZAKI 1991);
e) The resolution of the quadratic programming model is intractable for entire portfolios with more than 20 assets (KONNO AND YAMAZAKI, 1991);
f) The model assumes there are no difficulties in liquidating the portfolio formed, in other words, the market would absorb any type and amount of assets allocated by optimization (POGUE, 1970).

It is a fact that liquidity or transaction costs, are implicitly incorporated by investors in their investment allocation decisions. In other words, all else being equally constant, investors prefer more liquid than less liquid assets, particularly in the short-term.

Recently there has been an increased interest in studies of financial models with parameters modulated by Markov chains in an attempt to reflect the dynamics of the markets under conditions of financial distress (BAUERLE AND RIEDER, 2004; CAKMAK AND OZEKICI, 2006; COSTA AND ARAUJO, 2008; REBOREDO, 2002).

The job data were collected from daily reports provided by the BMF&Bovespa from January 2007 to September 2009. The selection, comparison and testing of hypotheses applied to the chosen liquidity indicator comprised the years 2007 and 2008. To perform the simulations, an application based on an Excel spreadsheet using the Microsoft Excel ® nonlinear programming solver for the Markowitz model was developed. The proposed model used the search tool Evolver ® with Genetic Algorithms. Portfolios were formed and compared from an arbitrary initial application of $ 5,000, after Brazilian taxes. The estimated price of these shares was based on the average behavior of the Brazilian financial market in the first half of 2009.

The concept of liquidity can be found in many ways. In accounting, liquidity is associated with the ease or speed which an asset can be turned into cash. In economic terms, an asset is considered liquid if its value is both easily negotiable and experiences little volatility over time.

In financial terms, liquidity can be defined as the ease which an asset can be exchanged within a short period of time (trading) without causing significant changes in its price (transaction cost). It is a systemic phenomenon that depends on the interaction between economic agents, where one wants to buy the asset (tangible or intangible) from another. Often, the information required for calculating the cost is not available, especially intraday trading data. Because of this, there is a variety of indicators available for measuring liquidity.

The Brazilian stock market releases and offers the public daily information on the bid-ask spread for each traded stock and releases the necessary data to evaluate the negotiability index (IN). The IN has several advantages: (i) available data; (ii) reliability because the data come from the BMF&Bovespa; (iii) measuring the intensity of trading action is consistent with the purposes of this study and (iv) by assessing the quantity and total financial volume traded in a period of time, IN avoids price distortion in the analysis of a long time series, for example, cases of split and bonus. However, IN does not assess the exposure frequency of the stock during the analyzed period.

Due to the facts presented, IN was adopted as liquidity measure for measuring and classifying analyzed assets and a new liquidity ratio ($IN_p$) was created whose calculation method is shown below. The $IN_p$ will be used for evaluating the negotiation probability.
\[ IN_p = \frac{\text{stock trading days}}{\text{trading days}} \times \frac{IN_{\text{max}}}{IN_{\text{max}}} \] (3)

The work is structured as follows. Section 2 describes a measure of liquidity for the portfolio based on Markov chain concepts. Section 3 shows the formulation of Markowitz model of portfolio optimization. Section 4 presents the proposed model. In section 5 the tests applied are characterized and the results presented. Section 6 shows the conclusions of the study.

2 Portfolio Liquidity

All portfolios are formed to be sold one day. Initially, consider two portfolios, A and B, with two assets in each one. The probabilities of trading these assets are, respectively, \( P_1 \) and \( P_2 \). These probabilities can be obtained from past observations that associate the quality of these assets to the amounts traded, or by a subjective estimate originating from the intuition of experts. The random variable \( X_i \) represents the possibility of trading of portfolio at time \( i \) with \( i = 1,2,3, ..., n \). Assuming that \( X_i \) are independent events and the probabilities remain the same throughout the \( n \) attempts in the portfolio, then, from the viewpoint of Markov stochastic processes there are two possible states for the portfolio: (i) \( S_1 \), the portfolio is full (complete), or (ii) \( S_2 \), the portfolio is empty (sold or traded). Figure 1 below shows the state diagram of this situation.

![State diagram of the negotiation portfolio comprising two assets](image)

Figure 1 – State diagram of the negotiation portfolio comprising two assets

What is the probability of negotiating the portfolio after \( n \) attempts?

Markov chains represent the simplest mathematical models of random phenomena that involve time. They are applied in a wide range of applications and are considered the most important example of stochastic processes. Indeed, the whole field of mathematical study of stochastic processes can be considered to be a form of generalization of the theory of Markov chains (Norris, 2006).

The stability condition of the Markov chain requires that the transition probabilities are for \( n = 1,2, ..., \) and all known possible sequences of states \( s_1, s_2, s_3, ..., s_{n+1} \) with \( X_1, X_2, ..., X_{n+1} \) are given by:

\[ P(X_{n+1} = s_{n+1} / X_1 = s_1, X_2 = s_2, ..., X_n = s_n) = P(X_{n+1} = s_{n+1} / X_n = s_n) \] (4)

The respective transition matrix of the examined case of a portfolio with two assets would be:

\[
\begin{pmatrix}
P_{11} & P_{12} \\
P_{21} & P_{22}
\end{pmatrix}
= \begin{pmatrix}
1 - P_1P_2 & P_1P_2 \\
P_1P_2 & 1 - P_1P_2
\end{pmatrix}
\sum_{j=1}^{k} p_{ij} = 1 \quad i = 1, 2, ..., p_{ij} \geq 0
\]
Similarly, the transition matrix of a portfolio of $n$ assets would be:

$$
\begin{pmatrix}
1 - \prod_{i=1}^{n} p_i & \prod_{i=1}^{n} p_i \\
\prod_{i=1}^{n} p_i & 1 - \prod_{i=1}^{n} p_i
\end{pmatrix}
\quad \quad \sum_{j=1}^{k} p_{ij} = 1 \quad i = 1, 2, \ldots
$$

(5)

$p_{ij} \geq 0$

Assuming $P$ as the initial probability matrix, the method for calculating the general matrix of transition probabilities in $n$ steps is:

$$
P_{(n)} = P_{(n-1)} P
$$

As an example, the main idea of the proposed liquidity portfolio measure, in accordance with the buyer’s perspective, is to form a portfolio which presents the most favorable trading outcomes measured by the probability of being traded, after an estimated number of sequential attempts, according to the concepts of Markov stochastic processes.

Using the notion of Markov stochastic processes, the probability of trading a portfolio of $n$ assets after two attempts starting from an empty position is given by:

$$
P_{12}^2 = \left(1 - \prod_{i=1}^{n} p_i \right) \left( \prod_{i=1}^{n} p_i \right) = \left(1 - \prod_{i=1}^{n} p_i \right)^2 + \left( \prod_{i=1}^{n} p_i \right)^2
$$

Based on the concepts of a finite Markov chain after $n$ trials, there is the possibility of the convergence of the probability matrix to state of equilibrium (since at least one $p_i < 1$). The probabilities of this state are obtained by solving a linear system of equations. For the case analyzed of a matrix with two states, the probabilities $\pi_1$ and $\pi_2$ of liquidating the portfolio at the state of equilibrium would be obtained by solving the following system of equations:

$$
\begin{pmatrix}
1 - \prod_{i=1}^{n} p_i & \prod_{i=1}^{n} p_i \\
\prod_{i=1}^{n} p_i & 1 - \prod_{i=1}^{n} p_i
\end{pmatrix}
\begin{pmatrix}
\pi_1 \\
\pi_2
\end{pmatrix}
= 1 - \prod_{i=1}^{n} p_i + \prod_{i=1}^{n} p_i
$$

(6)

$\pi_1 + \pi_2 = 1$

Assuming $\prod_{i=1}^{n} p_i \neq 0$, then, $\pi_1 = \pi_2 = 0.5$

3 The Markowitz Model

The approach developed by Markowitz (1952) assumes that the expected returns of the examined assets are known and so the allocation of available capital is possible. He suggests the use of past observations as an alternative to projecting expected returns.
$R_i$ is a random variable representing the rate of return per period of asset $i$ with $i = 1, 2, 3, \ldots$, and $X_i$ is the amount of capital to be invested in asset $i$. The expected return of the investment for the analyzed period is given by:

$$r(x_1, x_2, \ldots, x_n) = E \left[ \sum_{i=1}^{n} R_i x_i \right] = \sum_{i=1}^{n} E[R_i] x_i$$  \hspace{1cm} (7)$$

The duality return-risk is characterized by the expectation the investor has of obtaining maximum return for minimum risk. The risk measure used was the standard deviation of returns in a given period:

$$\sigma(x_1, x_2, \ldots, x_n) = \sqrt{ E \left[ \left( \sum_{i=1}^{n} R_i x_i - E \left[ \sum_{i=1}^{n} R_i x_i \right] \right)^2 \right] }$$  \hspace{1cm} (8)$$

One interpretation of the Markowitz model as a quadratic programming problem is given by Konno and Yamazaki (1991):

$$\min Z = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j$$  \hspace{1cm} (9)$$

subject to:

$$\sum_{i=1}^{n} r_i x_i \geq \rho M_0$$  \hspace{1cm} (10)$$

$$\sum_{i=1}^{n} x_i = M_0$$  \hspace{1cm} (11)$$

$$0 \leq x_i \leq u_i, \hspace{0.5cm} i = 1, 2, \ldots, n$$

Where $M_0$ is the total available capital for investment, $\rho$ is the minimum rate of return desired by the investor, $\mu_i$ is the maximum amount of money that can be invested in asset $i$, $R_i = E[R_i]$ and $\sigma_{ij} = E[(R_i - r_i)(R_j - r_j)]$.

4 The Proposed Model

The proposed approaches adopt the same assumptions as Markowitz model, plus the liquidity condition, based on Markov chains. The nonlinear models proposed aim to form a portfolio that simultaneously, minimize risk and maximize liquidity, after $k$ sequential attempts of trading, exceeding a minimum rate of return and deducting the operating costs of trading ($\alpha$). The risk of the proposed optimization model ($P_1$) is based on the covariance matrix of Markowitz.

The objective function was developed using the concept of goal programming (Hillier and Lieberman, 2005). The risk goal ($R_g$) used was a small value, but close to zero (e.g: 0.1). A natural candidate for the liquidity goal is the probability at the state of equilibrium explained in equation 6 ($\pi e = 0.5$).

The model incorporates real practices of the financial market such as fees, taxes and dividend payments there by making them more realistic. Besides these features the following assumptions are made:
a) The planning horizon of the investor is the short term;
b) The planning horizon consists of a single continuous period;
c) The investor is risk averse, so, the higher the risk the higher the expected return;
d) Variable and fixed operating costs were considered.

Model P1

\[
\min Z = \left[ \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\sigma_{ij} x_{ij} - R_e}{R_e} \right] + \frac{p^{(k)} - \pi_e}{\pi_e} \quad \text{(minimizes risk and maximizes liquidity)} \tag{12}
\]

subject to:

\[
\sum_{i=1}^{n} r x_i (1 - \alpha)(1 - \lambda) + D_i - C_i \geq \rho M \quad \text{(profitability is greater than benchmark)} \tag{13}
\]

\[
\sum_{i=1}^{n} x_i - C_i = M \quad \text{(allocated value is equal to initial investment less cost)} \tag{14}
\]

\[
x_i \geq u_i \quad \text{(invested value per stock is greater than last negotiated value)} \tag{15}
\]

\[
\sum_{j=1}^{k} p_{ij} = 1 \quad i = 1, 2, \ldots, k \quad \text{(the sum of probabilities is equal to 1)} \tag{16}
\]

\[
p_{ij} \geq 0 \quad \text{(positive probability)}
\]

\[
x_i \geq 0 \quad i = 1, 2, \ldots, n
\]

\[
y_i \in \{0, 1\}
\]

Metaheuristics are powerful search engines inspired by models of human life or nature. They can achieve good solutions in a short computational time for problems that have no exact mathematical solution. Metaheuristics are more complex simulations that have the ability to incorporate patterns of human behavior during the simulation process, such as adaptation and learning, allowing for the selection of superior solutions. For this reason, some metaheuristics are considered to be artificial intelligence (eg: genetic algorithms). Examples of metaheuristics: (i) Genetic Algorithm (GA), (ii) Ant System, (iii) Tabu Search, (iv) Simulated Annealing (SA) and (v) Hybrids.

Financial decisions in the short term, such as the portfolio, are inserted in the context of optimization. There is a clear predominance of the use of genetic algorithms in relation to other metaheuristics. The following list shows the classification of some metaheuristic applications in financial decisions.

b) Rules of marketing actions (ALLEN AND KARJALAINEN, 1999).
c) Financial difficulty prediction and insolvency risk (MCKEE AND LENSBERG, 2002; VARETO, 1998).
d) Investment Recommendations (LI AND TSANG, 1999, 2000)

According to reasons presented, genetic algorithms were chosen as search engine to select the best combination of stocks for the portfolio by the proposed model. In GA, the term chromosome typically refers to a candidate solution. Functionally, the genetic algorithm uses the following operators:
Reproduction
The initial solution is formed by a sequence of bits that represent the characteristics of the product. The selection operator selects a subset of $m$ chromosomes of size $M$ of the population that can reproduce, on average, better adapted chromosomes produce more offspring than the less well adapted. Generally, the size of the chromosome is maintained in successive generations.
b) Crossover
The operator of the crossover exchange parts of chromosomes positions specifically chosen for the formation of new offspring.
c) Mutation
The mutation operator changes the values of some attributes at random.

5 Results
Liquidity was estimated by weighting the frequency $F$ obtained by dividing the trading days of each action in period ($F_a$) by the number of trading days in the period ($F_p$) ($F = F_a / F_p$). This frequency was weighted by its respective average $IN$ divided by the maximum number recorded between the studied securities ($F x IN_{avg} / IN_{max}$). The weighted ratio was grouped into quartiles arbitrarily assuming, the average probability of 1, 2, 3 and 4, respectively, 1.00, 0.75, 0.50 and 0.05. This was the possible liquidity estimate that could be obtained from the public data available. The list of shares participating in the simulations with their respective quartiles and trading probabilities are presented in Table 1 below.

The Brazilian financial market defines the validity of a buy or sell order by the number of days and not by the number of attempts. Twenty attempts were adopted as an intermediate value between the minimum and maximum used by the market.

As suggested by Markowitz (1952), for demonstration purposes, the average performance of the 1st half of 2009 was used to estimate the profitability of each stock

<table>
<thead>
<tr>
<th>Rankin</th>
<th>Stock</th>
<th>Probability</th>
<th>Quartile</th>
<th>Rankin</th>
<th>Stock</th>
<th>Probability</th>
<th>Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ambev</td>
<td>0.75</td>
<td>3</td>
<td>14</td>
<td>MMX</td>
<td>0.50</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Bradesco</td>
<td>0.75</td>
<td>3</td>
<td>15</td>
<td>Celese</td>
<td>0.50</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Banco do Brasil</td>
<td>1.00</td>
<td>4</td>
<td>16</td>
<td>OHL</td>
<td>0.50</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>CESP</td>
<td>0.75</td>
<td>3</td>
<td>17</td>
<td>P. Seguro</td>
<td>0.50</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>Gerdau</td>
<td>1.00</td>
<td>4</td>
<td>18</td>
<td>Random</td>
<td>0.50</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>Net</td>
<td>1.00</td>
<td>4</td>
<td>19</td>
<td>Copasa</td>
<td>0.75</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>Petrobras</td>
<td>1.00</td>
<td>4</td>
<td>20</td>
<td>Marco Polo</td>
<td>0.75</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>CSN</td>
<td>1.00</td>
<td>4</td>
<td>21</td>
<td>Klabin</td>
<td>1.00</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>Telemar</td>
<td>0.50</td>
<td>2</td>
<td>22</td>
<td>Caf Brasil</td>
<td>0.05</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Usiminas</td>
<td>1.00</td>
<td>4</td>
<td>23</td>
<td>Sergen</td>
<td>0.05</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>Vale</td>
<td>1.00</td>
<td>4</td>
<td>24</td>
<td>Hercules</td>
<td>0.05</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>Comgás</td>
<td>0.50</td>
<td>2</td>
<td>25</td>
<td>Marisol</td>
<td>0.05</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>TAM</td>
<td>0.75</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Once the portfolio is classified in terms of attributes and levels, an initial population of size $M$ is randomly generated. For purposes of this research, a convergence was found with the following configuration parameters of genetic algorithm: (i) $M = 52$, (ii) a uniform rate of crossover equal to 50%, (iii) a mutation rate of 10%, (iv) stopping criterion after 75,000 iterations and (v) the Evolver ® internal method recipe. The best results and comparisons between the models are shown in Table 2 below.

As expected, the Markowitz model created portfolio with lower volatility relative to the Ibovespa market portfolio according to Sharpe. The proposed model, in turn, have created high volatile portfolio. However, the difference in quality of the selected assets is significant. The Markowitz model allocated values to shares 22 and 23 that have low liquidity while the proposed model $P_1$, as expected, avoided selecting these assets.

The share of less liquid stocks (22 and 23) in portfolio of Markowitz reduced the potential returns of the portfolio because of lower trading frequency of trading and the updating of their prices at auctions. The portfolio formed by the $P_1$ model, incorporated the most liquid stocks and the portfolio that was formed is more realistic in terms of potential trading shares and include well know Brazilian companies such as Bradesco, Usiminas and Marco Polo.

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Markowitz</th>
<th>$P_1$</th>
<th>Ibovespa</th>
</tr>
</thead>
<tbody>
<tr>
<td># stocks</td>
<td>6</td>
<td>6</td>
<td>65</td>
</tr>
<tr>
<td>Stocks</td>
<td>6 – 8 – 18</td>
<td>2 - 3 - 4 - 16 19- 20</td>
<td></td>
</tr>
<tr>
<td>Estimated Profitability</td>
<td>4,26%</td>
<td>4,63%</td>
<td>4,20%</td>
</tr>
<tr>
<td>Observer Profitability</td>
<td>4,75%</td>
<td>5,69%</td>
<td>4,90%</td>
</tr>
<tr>
<td>Risk ($\beta$)</td>
<td>0,2601</td>
<td>0,6937</td>
<td>1,0000</td>
</tr>
<tr>
<td>Liquidity (Markov)</td>
<td>0,0184</td>
<td>0,4998</td>
<td>1,0000</td>
</tr>
</tbody>
</table>

3 65 stocks on average according to BMF&Bovespa
4 from June to Dec/2009
5 according to market portfolio definition

6 FINAL CONSIDERATIONS

The multi-criteria optimization model generated in this study incorporated a measure of liquidity based on the probability of trading the shares included in a stochastic process of Markov chains. This includes two important aspects: (i) an approach to the dynamism of a market that trades shares in several attempts, and (ii) the introduction of liquidity and transaction costs in the decisions. The work reinforced the conclusion obtained in other studies that the absence of transaction costs can generate inefficient or unrealistic portfolios.

This study considered two possible states of negotiation (negotiated or not negotiated). However, other states of the partial liquidation of the portfolio could be simulated, but this
would require a large computational effort for implementing. This opens the way to new lines of research on the subject.

REFERENCES


