Professor Lesly ORTIZ-CEREZO, MSc
E-mail: lortizc1401@alumno.ipn.mx
Instituto Politécnico Nacional
Professor Alin CARSTEANU, PhD
E-mail: acarsteanu@ipn.mx
Instituto Politécnico Nacional
Professor Julio B. CLEMPNER, PhD
E-mail: julio@clempner.name
Instituto Politécnico Nacional

OPTIMAL CONSTRAINED PORTFOLIO ANALYSIS FOR INCOMPLETE INFORMATION AND TRANSACTION COSTS

Abstract. *Asset price dynamics can be explained by means of a mathematical model. The problem is that these models simulate real situations, which are not always attainable. The fact that real markets are characterized by inadequate information is one of the challenging issues that mathematical models must solve as they are compelled to make assumptions about the financial market.*

This paper aims to introduce and create a unique Markov model for incomplete information in order to address these deficiencies and solve the portfolio optimization problem while taking pricing and risk minimization management into account. We assume that trading in securities occurs in discrete time increments. Investors can trade in securities and track their short-term pricing. We make the following assumptions: short-selling is not permitted on the market, all assets have equal selling and buying prices, trading costs apply, and investors can buy and sell an infinite number of shares on an arbitrage-free market. We suggest the construction of an observer for partial information to address the portfolio optimization problem. The ideal portfolio and the problem observer are the results of the process. We provide the formulae for recovering the important variables. We can recover the observation kernel from the resulting observer, which is understood to be the manager's intuition during the decision-making process. This is the main result of this paper. To the best of our knowledge, this is the first time that a solution of this kind for portfolio optimization using partial data has been proposed in the literature. Under these assumptions, a Markov model is used in a numerical example that demonstrates the effectiveness and use of the suggested approach for creating workable portfolio models.

Keywords: *Portfolio, observer design, Markov chains, optimization.*

JEL Classification: G11, C61, C69

DOI: 10.24818/18423264/56.4.22.07

1.Introduction

1.1.Brief review

The Markowitz single period mean-variance portfolio [18] is defined as a model able to maximize the terminal wealth while, in the meantime, minimize risk employing the variance as a criterion. The goal is to make possible that an investor looks for the highest return by establishing a tolerable risk level. The model was extended to support continuous-time model based on the geometric Brownian motion represented by differential equations. However, there exists several limitations in the employment of continuous-time models because certain parameters are discrete: actions (controls), interest rates, appreciation, and volatility rates, etc. These parameters are insensible and can produce radical changes in the behavior of the market. Markets are ruled by trends, which are perceived as tendencies of securities to move in a particular direction over time. Investors try to predict securities trends using mathematical models, which identify tendencies when the price reaches support and resistance levels over time.
For solving these problems, we formulate the Markowitz's single period mean-variance portfolio as a system of stochastic difference equations whose variables are represented by a discrete-time Markov chain. We focus on a class of discrete-time Markov mean-variance portfolio model and compute the portfolio policy that minimizes the overall risk given a fixed expected return.

1.2.Related work

There is large body of literature on portfolio optimization for Markov chains with constraints. For a survey in the effects of transaction costs on portfolio optimization see [22] and [11]. There is large body of literature on portfolio optimization for Markov chains with constraints. Sanchez et al. [24] suggested a new mean-variance customer portfolio optimization algorithm for a class of ergodic finite controllable Markov chains. Sanchez et al. [23] improved the work presented in [24] proposing a recurrent reinforcement-learning approach according to an actor-critic architecture subject of a penalty regularized expected utilities [7]. For computing the mean-variance client portfolio with transaction costs in controlled partially observable Markov decision processes, Asiain et al. [2] proposed a reinforcement-learning framework. For continuous-time discrete-state Markov decision processes, Garcia-Galicia et al. [16] took into account a continuous-time mean-variance portfolio with transaction costs comprising temporal penalization. Dominguez and Clempner [14] used the extraproximal technique to solve the multi-period Markowitz's portfolio optimization issue. For the purpose of financial portfolio management, Garcia-Galicia et al. [15] examined the problem of policy optimization in the context of continuous-time reinforcement learning. The portfolio problem aims to redistribute a fund among various financial assets. Risk, return, and transaction cost were the three objectives of the tri-objective portfolio optimization model given by Meghwani and Thakur [19]. They also suggested an algorithm that can handle equality constraints successfully without the need for any constraint management techniques. Options in literature can be found depending on the subject [1, 12, 13, 17, 25, 21, 20].

1.3.Main results

This research provides a market model-based approach to the portfolio optimization issue that allows for the formulation of securities returns and risk mitigation strategies as well as performance monitoring. We assume that the trading of securities (tradable financial assets, including stocks, bonds, and derivatives) occurs in discrete time increments. Investors can trade in securities and track their short-term pricing. The investor makes decisions regarding the makeup of his portfolio based on data from observed securities prices. When there is insufficient market information, we provide a solution. The investor begins with a base amount to purchase securities, and money imports and exports are prohibited. The value of the portfolio will thus only fluctuate as a result of the rate of increase or decrease for each security.

The following is a summary of the main results:

- Considers the problem of portfolio selection with transaction costs.
- For such problems, the optimal portfolio is computed very rapidly using a proximal algorithm.
- Deals with incomplete information, transaction costs and arbitrage-free market.
- Proposes a new auxiliary variable, which stands for the product of the observer and the (prior) distribution vector.
- Conceptualizes the observer design as the product of the policy and observation kernel which denotes the relationship between the real and the estimated state.
- Interprets managers' intuition in making decisions from the generated observation kernel. This is paper's key finding.
- Provides a powerful resolution to the problem, by combining a financial mathematical model with rising computer power.

1.4.Organization of the paper

The remaining sections of the paper are structured as follows. The single-period portfolio selection issue is discussed in the next section. With regard to partially observable Markov theory and the formulation of the portfolio issue, Section 3 discusses the portfolio model with imperfect information. Section 4 offers a portfolio solution approach for partial data that takes transaction costs and arbitrage-free into account. In Section 5, a numerical example is provided. Section 6 details our findings and closing remarks.

2.Portfolio model for incomplete information

2.1.Partially observable Markov chains

Let S be a finite state space and $Y = \{y_1, \ldots, y_m\}$ the set of observations (pseudostates) [10]. If at the time $t \in [0, T]$ the state $s(t) = s_i$ is possible then with probability $q_{m|i}$ the pseudostate y_m appears. The control goal for partially observable Markov chains has the same structure as in the case of totally observable states, but the sets of feasible strategies are different.

We consider finite controllable partially observable Markov chains with utility u given by $M = \{S, A, P, Y, Q, Q_0, P_0, u\}$. Although we believe that their states s are unobservable, we can nonetheless observe their pseudostates y from a set Y, such that $|Y| = M$, $|S| = N$ and $N = M$. These states appear according to following the rule: if at the time t the Markov chain is in the state s_i then the pseudostate y_m will appear at the same moment with probability $q_{m|i}$ which does not depend on the history. Given $y(t) = y_m$ and $a(t) = a_k$, then $s(t) = s_i$ will have a probability $q_{m|ik} := P(s(t) = s_i | y(t) = y_m, a(t) = a_k)$, $m = \overline{1, M}$, $i = \overline{1, N}$, $k = \overline{1, K}$ that denotes the relationship between the state and the observation when an action $a_k \in A(s_i)$ is chosen at time t. The probabilities of $q_{m|i}$ are recovered from the observation kernel, which is a stochastic kernel on Y, recovered from $Q = [q_{m|ik}]_{m=\overline{1,M}, i=\overline{1,N}, k=\overline{1,K}}$.

In $M = \{S, A, P, Y, Q, Q_0, P_0, u\}$ we have that S is a finite state space; A is a finite control space; $P = [p_{j|ik}]$ is a controlled transition matrix; Y is the observation set, which takes values in a finite space $\{1, \ldots, M\}$, $M \in \mathbb{N}$; $Q = [q_{m|i}]_{m=\overline{1,M}, i=\overline{1,N}}$ denotes the observation kernel is a stochastic kernel on Y such that $\sum_{m=1}^{M} q_{m|i} = 1$ for all $i = \overline{1, N}$, which means that the each state i is observable with probability one; $Q_0 = [q_{m|i}]_{m=\overline{1,M}, i=\overline{1,N}}$ denotes the initial observation kernel, is a stochastic kernel on Y given S; P_0 is the (a prior) initial distribution; $u_{ijmk} : S \times S \times Y \times A \to \mathbb{R}$ is the utility function. The payoff function u_{imk} could depend upon the current state s_i, the estimated state y_m and the action a_k taken.

The partially observable system dynamics at time t is given by

$$(s(0), y(0), a(0), s(1), y(1), a(1), \ldots) \in H := (SYA)^{\infty},$$

where $s(0)$ has a given distribution $P(s(0) = s_i)$ and $\{a_t\}$ is a control sequence in A determined by a control policy. To define a policy, we cannot use the (unobservable) states $s(0), s(1), \ldots$. Then, we introduce the observable histories $h_0 := (y(0)) \in H_0$ and $h_t := (s(0), y(0), a(0), \ldots, y(t-1), a(t-1), y(t)) \in H_t$ for all $t \geq 1$ and $H_t := H_{t-1}AY$ if $t \geq 1$. Then, a *policy* is defined as a sequence $\{\pi_{k|m}(t)\}$. The collection of all policies is denoted by Π.

A sequence of random stochastic matrices $\Pi(n) = \{\pi_{k|m}(t)\}_{k=\overline{1,K},m=\overline{1,M}}$ is said to be *a randomized control policy* such that for any random action $\pi_{k|m}(t)$ at time t

$$\sum_{k=1}^{K} \pi_{k|m}(t) = 1, \qquad \pi_{k|m} \geq 0, \qquad m = 1, \dots, M$$

For any random action $\{\pi_{k|m}(t)\}_{k=\overline{1,K},m=\overline{1,M}}$ the conditional transition probability matrix

$$P(\pi_{k|m}(t)) = [P_{j|i}(\pi_{k|m}(t))]_{i,j=\overline{1,N}}$$

can be defined as follows

$$P_{j|i}(\{\pi_{k|m}(t)\}) = \sum_{k=1}^{K}\sum_{m=1}^{M} P(s(t+1) = s_j | s(t) = s_i, a(t) = a_k) \cdot$$
$$P(s(t) = s_i | y(t) = y_m)\pi_{k|m}(t) = \sum_{k=1}^{K}\sum_{m=1}^{M} p_{j|ik} q_{m|i} \pi_{k|m}(t).$$

2.2.Formulation of the portfolio problem

We are considering a portfolio problem where, at moment $t = 0$, the state $s(0)$ has a distribution P, as well as the state $y(0)$ is recovered from $Q_0(y(0)|s(0))$. If at moment t the state of the system is $s(t)$ and the control $a(t) \in A$ is applied, then each of policy is allowed to randomize, with distribution $\pi_{k|m}(t)$, over the pure action choices $a(t) \in A(s(t))$, $m = \overline{1,M}$ and $k = \overline{1,K}$. These options generate the utility $\mathcal{U}_{ik}$. The system makes an effort to minimize the associated one-step reward. Thereafter, the system makes a transition to the state $s(t) = s_i$ following to the transitions given by $P(\{\pi_{k|m}(t)\}_{k=\overline{1,K},m=\overline{1,M}})$. Next, the unobserved state $y(t)$ is generated by the kernel $Q(y(t)|s(t))$. Following the resulting reward, the system computes the policy $\pi_{k|m}(t+1)$ for the next selection of the control actions.

The *reward* $\mathcal{U}$ of a portfolio is

$$\sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{j=1}^{N} J_{ijmk} p_{j|ik} q_{m|i} \pi_{k|m} P_i = \sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{k=1}^{K} V_{imk} z_{imk} \to \max_{z \in Z_{adm}}$$

such that $V_{imk} = \sum_{j=1}^{N} J_{ijmk} p_{j|ik}$ and $z := [z_{imk}]_{i=\overline{1,N};m=\overline{1,M};k=\overline{1,K}}$ is defined as the *joint policy* given by

$$z_{imk} = \pi_{k|m} q_{m|i} P_i$$

Considering an ergodic Markov chain [7], for any fixed collection of stationary strategies $\pi_{k|m}(t) = \pi_{k|m}$ we have that $P\big(s(t+1) = s_j\big) \to P_j, \quad t \to \infty$. Then, for any stationary policy $\pi_{k|m}$ the distributions $P\big(s_j\big)$ exponentially converge to $P(s_i) = P_i$ such that

$$P_j = \sum_{i=1}^{N}\left(\sum_{m=1}^{M}\sum_{k=1}^{K} p_{j|ik}q_{m|i}\pi_{k|m}\right)P_i$$

Let us denote by p_m the probability to observe the estimated state m , as follows

$$p_m = \sum_{k=1}^{K}\left(\sum_{i=1}^{N}\left(q_{m|i}\pi_{k|m}\right)\right)P_i$$

The reward function $\mathcal{U}$ in the stationary regime is represented by

$$\mathcal{U}(o) := \sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{k=1}^{K} V_{imk}o_{mi}\,\pi_{k|m}$$

where $V_{imk} = \sum_j J_{ijmk}p_{j|ik}$ and the observer is defined as $o_{mi} = q_{i|m}P_i$.

Let $\mathcal{O}_{adm}$ be the set of admissible observers. An observer $o^* = \{o^*_{mi}\}$ is *optimal*, if it satisfies $o^* := argmax_{o\in\mathcal{O}_{adm}}\mathcal{U}(o)$.

The portfolio optimization problem maximizes the mean value $\mathcal{U}(o)$ while minimizing the *variance* $\mathcal{V}ar(o)$ defined as follows

$$\mathcal{V}ar(o) \coloneqq \sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{k=1}^{K}[V_{imk} - \mathcal{U}(o)]^2 z_{imk} \to \min_{o\in\mathcal{O}}$$

Finally, the mean-variance customer model of [18] is defined as follows

$$\Psi(o) := \mathcal{U}(o) - \frac{\xi}{2}\,\mathcal{V}ar(o) = \sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{k=1}^{K} V_{imk}\pi_{k|m}o_{mi} - \frac{\xi}{2}\sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{k=1}^{K}[V_{imk} - \mathcal{U}(o)]^2\, z_{imk} \to \max_{o\in\mathcal{O}}$$

where ξ is *the risk-aversion* parameter.

3.Portfolio solution method

3.1.Feasibility of Markowitz's portfolio

The joint policy $z_{imk} = \pi_{k|m}o_{mi} = \pi_{k|m}q_{m|i}P_i$ belongs to Z_{adm} and satisfies the constraints:

1) The matrix $z := [z_{imk}] \in Z_{adm}$ is a stationary policy that belongs to the simplex

$$\mathcal{S}^{NMK} := \left\{ z \in \mathbb{R}^{NMK} \; : \; \text{for } z_{imk} \geq 0 \text{ where } \sum_{m=1}^{M} \sum_{i=1}^{N} \sum_{k=1}^{K} z_{imk} = 1 \right\}$$

$$\sum_{l=1}^{N} \sum_{k=1}^{K} z_{ilk} = P_i > 0$$

2) The variable $z_{imk} \in Z_{adm}$ satisfies the ergodicity constraints i.e.:

$$\sum_{m=1}^{M} \sum_{k=1}^{K} [\delta_{ij} - p_{j|ik}] z_{imk} = 0 \, , j = 1, \ldots, N$$

$$\sum_{k=1}^{K} z_{imk} = q_{m|i} \sum_{l=1}^{M} \sum_{k=1}^{K} z_{ilk} = 0$$

$$\sum_{l=1}^{M} \sum_{k=1}^{K} [\delta_{lm} - q_{m|i}] z_{ilk} = 0 \, , m = 1, \ldots, M$$

Once the model (ergodic Markov decision process) is solved in order to recover the quantities of interest we have that:

$$\pi^*_{kim} = \begin{cases} \dfrac{z^*_{imh}}{\sum_{h=1}^{K} z^*_{imh}} & if \quad \sum_{h=1}^{K} z^*_{imh} > 0 \\ 0 & if \quad \sum_{h=1}^{K} z^*_{imh} = 0 \end{cases}$$

$$P^*_i = \sum_{l=1}^{M} \sum_{k=1}^{K} z^*_{ilk} \; and \; q^*_{m|i} = \frac{\sum_{k=1}^{K} z^*_{imk}}{\sum_{l=1}^{M} \sum_{k=1}^{K} z^*_{ilk}}$$

To close, to get the optimal observer we must use the formula

$$o^*_{mi} = q^*_{i|m} P^*_i$$

The resulting expression $q^*_{m|i}$ is not trivial. The inaccurate knowledge about the system state, which is often gleaned intuitively, is represented by the matrix $[q^*_{m|i}]$.

The policy $\pi^*_{k|m}$, constructed from π^*_{kim}, is given by

$$\pi^*_{k|m} = \frac{1}{N} \sum_{i=1}^{N} \pi^*_{kim}$$

The distribution is given by

$$p_m^* = \sum_{i=1}^{N} \sum_{k=1}^{K} z_{imk}^*$$

Using the Lagrange method, the mean-variance Markowitz portfolio $\Psi(o)$ can be re-written as follows:

$$L(o, \mu_0, \mu_{N+1}, \mu_{N+2,j}) \;\; := \sum_{m=1}^{M} \sum_{i=1}^{N} \sum_{k=1}^{K} V_{imk} \pi_{k|m} o_{mi} - \frac{\xi}{2} \sum_{m=1}^{M} \sum_{i=1}^{N} \sum_{k=1}^{K} \left[V_{imk} - V_{imk} \pi_{k|m} o_{mi}\right]^2 z_{imk} + \sum_{j=1}^{N} \mu_{0,j} \left[\left(\sum_{m=1}^{M} \sum_{k=1}^{K} [\delta_{ij} - p_{j|ik}] z_{imk}\right) - b_{eq,j}\right] + \mu_{N+1} \left(\sum_{m=1}^{M} \sum_{i=1}^{N} \sum_{k=1}^{K} z_{imk} = 1 - b_{eq,N+1}\right) + \sum_{j=1}^{N} \mu_{N+2,j} \left[\left(\sum_{l=1}^{M} \sum_{k=1}^{K} [\delta_{lm} - q_{i|m}] z_{ilk}\right) - b_{eq,N+1+j}\right]$$

3.2.Transaction costs

Let us consider $\alpha = col(z_{imk})$ and let $\mathcal{A}_{adm} = Z_{adm}$. Each α_i reprents current portfolio holding of asset. Each time an investor buys or sells an asset an expense is incurred. We study the case where the transaction cost to move from one asset to a better one is penalized by a cost. In the work of Markowitz, the expenses associated with trading equities, were excluded from his model. The importance of considering the transaction cost in a new portfolio and in revising an existing portfolio are well acknowledged. Transaction cost should be as low as possible. Thus, a portfolio manager must carefully consider trading and its resulting cost.

We will define the behavior of an investor as a sequence $(\alpha_n)_{n \in \mathbb{N}}$ where there are possible changes in the assets, such that $\alpha_j \neq \alpha_i$ or holding the same asset, $\alpha_j = \alpha_i$ in order to maximize the utility and minimize the risk of the portfolio. We also consider problems with transaction costs where the costs are paid on all transactions irrespective of the volume of the transaction. They consist of brokerage commissions and transfer fees. Then, at each step $n \in \mathbb{N}$ the investor chooses to change or to stay in $\alpha_i \in \mathcal{A}$. The function ψ represents the function which determines the decision to change asset α_i . The change from an asset to a different one produces a transaction cost, which can be defined as a function $\varphi : \mathcal{A} \times \mathcal{A} \to \mathbb{R}$ which can interpreted as a distance

function where if $\varphi(\alpha_i, \alpha_j) = c(\alpha_i, \alpha_j)$ we have that $c(\alpha_i, \alpha_j) = 0$ if $\alpha_j = \alpha_i$ or $c(\alpha_i, \alpha_j) > 0$ if $\alpha_i < \alpha_j$. In the classical model for transaction costs is that there are none, i.e., $c(\alpha_i, \alpha_j) = 0$. Then, the function $\varphi(\alpha_i, \alpha_j)$ can be reexpressed as $\varphi(\alpha_i, \alpha_j) = \delta(\alpha_i)c(\alpha_i, \alpha_j)$ where $\delta(\alpha_i) \in [0,1]$ is the costs for each transaction proportional to the distance to move from α_i to α_j and $c(\alpha_i, \alpha_j)$ is the one-step cost function.
The case $\psi(x_i) - \psi(x_j) \geq 0$ is the advantage to change from α_i to α_j and $\zeta(\alpha_i)$ is the weight the investor puts on his advantages to change from a given asset to another. The advantages to change from α_i to α_j are given by $\Xi(\alpha_i, \alpha_j) = \zeta(\alpha_i)\big(\psi(\alpha_i) - \psi(\alpha_j)\big)$.
If any number of convex transaction costs and convex constraints are combined, the resulting problem is convex. Linear transaction costs, as well as all the portfolio constraints describe above, are convex. Such problems can be solved with great efficiency, even for problems with a large number of assets and constraints. We have that the general portfolio problem with transaction cost can be defined by

$$\Psi(\alpha_i) = \left\{\alpha_j \in \mathcal{A} \;:\; \zeta(\alpha_i)\big(\psi(\alpha_i) - \psi(\alpha_j)\big) \geq \delta(\alpha_i)c(\alpha_i, \alpha_j)\right\}$$

Then, the acceptance criterion to change or stay in the same process satisfies the condition

$$\zeta(\alpha_n)(\psi(\alpha_n) - \psi(\alpha_{n+1})) \geq \delta(\alpha_n)c(\alpha_n, \alpha_{n+1})$$

After that, considering

$$\alpha_n = \alpha \;,\quad \delta(\alpha)c(\alpha, \alpha^*) = \delta_n\|(\alpha - \alpha^*)\|^2 \quad \text{and} \quad \Xi(\alpha, \alpha^*) := -\zeta_n[\psi(x) - \psi(x^*)]$$

we obtain

$$\alpha^* = \arg\max_{\alpha \in \mathcal{A}}\{-\delta_n\|(\alpha - \alpha^*)\|^2 + \gamma_n(\delta_n(\alpha)[\psi(x) - \psi(x^*)])\}$$

4.Numerical example

4.1.Description of the example

We suppose that a portfolio containing all risky assets and can be calculated by considering the sum of the risk-free rate plus the excess return (market return minus risk-free rate) multiplied by ξ , the sensitivity of the portfolio to market movements

$$\Psi(o) := \mathcal{U}(o) - \frac{\xi}{2}\,\mathcal{V}ar(o) \rightarrow \max_{o \in \mathcal{O}}$$

where ξ is *the risk-aversion* parameter [3].
We assume that investors target the portfolio with the lowest risk over the same one-period horizon and anticipate returns with the same probability distribution. We believe that the markets are in equilibrium and that there has been no inflation or change in interest rates. To make trading more realistic, we take into account transaction costs and the fact that investors can trade an unlimited number of shares on an arbitrage-free

market.

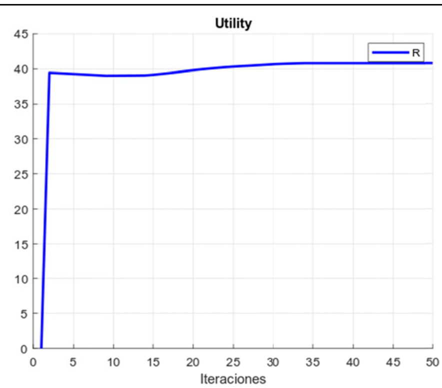

Figure 1. Utility value of the portfolio

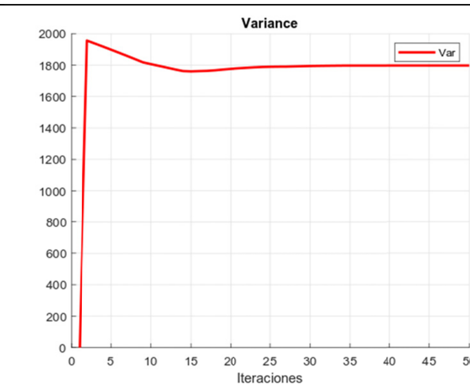

Figure 2. Variance of the portfolio

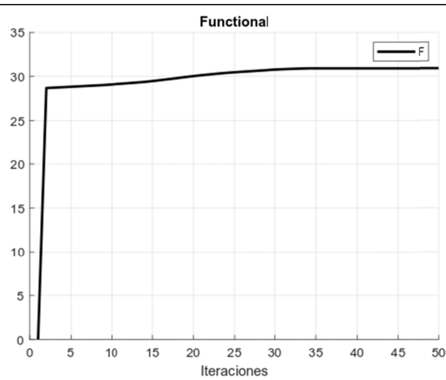

Figure 3. Portfolio's functional value

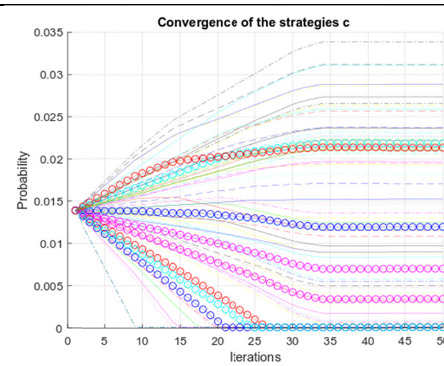

Figure 4. Policies' convergence

4.2.Computing a single period portfolio

Considering $\xi = 1.1 \times 10^{-2}$ and $\gamma = 1.3 \times 10^{-4}$ and the resulting values for the observer design o^*_{mi}, the observability matrix $q^*_{m|i}$ and optimal portfolio $\pi^*_{k|m}$ are given by

$$o_{im}^{*}=\begin{bmatrix}0.0214 & 0.0354 & 0.0238 & 0.0398 & 0.0002 & 0.0477\\ 0.0287 & 0.0353 & 0.0099 & 0.0458 & 0.0283 & 0.0212\\ 0.0584 & 0.0172 & 0.0252 & 0.0270 & 0.0056 & 0.0362\\ 0.0403 & 0.0369 & 0.0010 & 0.0042 & 0.0552 & 0.0290\\ 0.0196 & 0.0315 & 0.0183 & 0.0242 & 0.0338 & 0.0320\\ 0.0269 & 0.0238 & 0.0153 & 0.0473 & 0.0052 & 0.0486\end{bmatrix}\quad \pi_{k|m}^{*}=\begin{bmatrix}0.4158 & 0.5842\\ 0.1927 & 0.8073\\ 0.6100 & 0.3900\\ 0.5792 & 0.4208\\ 0.6434 & 0.3566\\ 0.2862 & 0.7138\end{bmatrix}$$

$$q_{m|i}^{*}=\begin{bmatrix}0.1272 & 0.1695 & 0.3445 & 0.2419 & 0.1229 & 0.1609\\ 0.2103 & 0.2089 & 0.1013 & 0.2218 & 0.1973 & 0.1425\\ 0.1413 & 0.0583 & 0.1485 & 0.0063 & 0.1149 & 0.0915\\ 0.2363 & 0.2709 & 0.1594 & 0.0249 & 0.1521 & 0.2831\\ 0.0012 & 0.1671 & 0.0331 & 0.3312 & 0.2123 & 0.0312\\ 0.2837 & 0.1253 & 0.2133 & 0.1739 & 0.2005 & 0.2909\end{bmatrix}\quad p_{m}^{*}=\begin{bmatrix}0.1952\\ 0.1801\\ 0.0935\\ 0.1883\\ 0.1283\\ 0.2146\end{bmatrix}$$

Figure 1 shows the convergence of the utility, Figure 2 shows the variance and Figure 3 shows the convergence of the functional. Figure 4 shows the convergence of the portfolio strategies.

4.3.Computing the efficient frontier

The main goal of the investor is to gain a given return. A rational investor makes an effort to identify the portfolio with minimal risk which satisfies this goal. For fulfilling this goal, we outline all the possible portfolios of risky assets in a mean-variance diagram, where the points represent the expected returns $\mathcal{U}$ and the risk $\mathcal{V}ar$ (variance) of the portfolios. Furthermore, we call the set of all points $\mathcal{V}ar(o^{*})=\mathcal{U}(o^{*})$ the efficient frontier and has the shape of a hyperbola (Pareto front [6]) see Figure 5.

A portfolio is called mean-variance efficient (or just efficient), if for a given volatility there is no portfolio with a higher return such that $\mathcal{U}(o^{*})\leq\mathcal{U}(o)$ and $\mathcal{V}ar(o^{*})\geq\mathcal{V}ar(o)$. It is the upper boundary of all portfolios in the mean-variance diagram from Figure 5. According to our paradigm, the rational investor is specifically seeking the following set of portfolios: They both increase the expected return while minimizing the risk for a given return.

The portfolio on the efficient frontier with the lowest volatility is called minimum-variance portfolio (red circle in Figure 6). If a risk-free asset exists (zero volatility) then the set of mean-variance efficient portfolios, established by the risk-free and risky assets, is the tangency point on the efficient frontier (blued circle in Figure 6). Considering $\xi=5.5\times10^{-5}$ and $\gamma=1.3\times10^{-3}$ and the resulting values for the

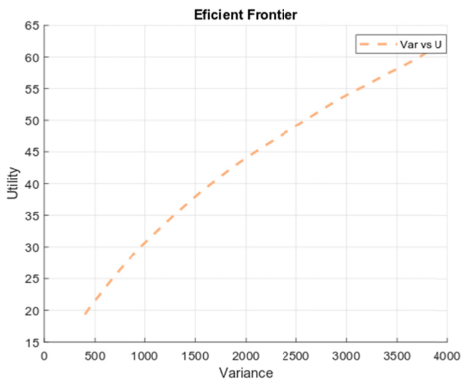

Figure 5. Efficient frontier

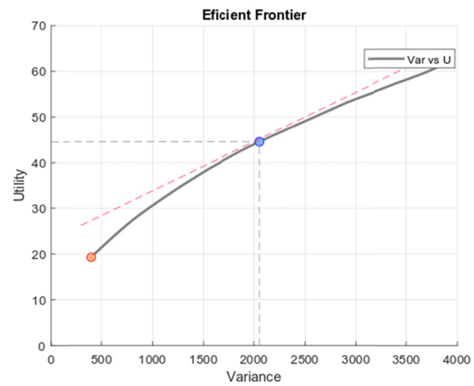

Figure 6. Tangency point

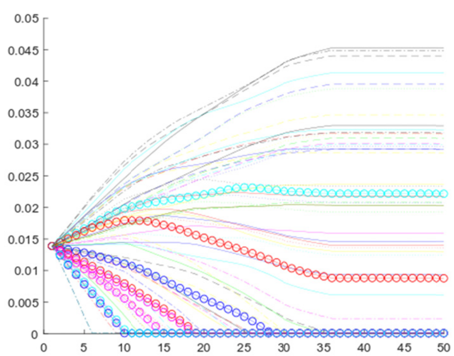

Figure 7. Strategies' Convergence

observer design o^*_{mi} , the observability matrix $q^*_{m|i}$ and optimal portfolio $\pi^*_{k|m}$ are given by

$$o^*_{im} = \begin{bmatrix} 0.0002 & 0.0378 & 0.0731 & 0.0002 & 0.0002 & 0.0436 \\ 0.0838 & 0.0002 & 0.0002 & 0.0002 & 0.0002 & 0.0746 \\ 0.1128 & 0.0483 & 0.0002 & 0.0002 & 0.0002 & 0.0002 \\ 0.0002 & 0.0002 & 0.0969 & 0.0002 & 0.0002 & 0.0754 \\ 0.1901 & 0.0002 & 0.0002 & 0.0002 & 0.0002 & 0.0002 \\ 0.0719 & 0.0002 & 0.0869 & 0.0002 & 0.0002 & 0.0002 \end{bmatrix} \quad \pi^*_{k|m} = \begin{bmatrix} 0.6664 & 0.3336 \\ 0.3341 & 0.6659 \\ 0.7494 & 0.2506 \\ 0.5000 & 0.5000 \\ 0.5000 & 0.5000 \\ 0.5833 & 0.4167 \end{bmatrix}$$

$$q_{m|i}^{*} = \begin{bmatrix} 0.0013 & 0.5264 & 0.6966 & 0.0012 & 0.9948 & 0.4505 \\ 0.2436 & 0.0013 & 0.2984 & 0.0012 & 0.0010 & 0.0013 \\ 0.4716 & 0.0013 & 0.0012 & 0.5599 & 0.0010 & 0.5445 \\ 0.0013 & 0.0013 & 0.0012 & 0.0012 & 0.0010 & 0.0013 \\ 0.0013 & 0.0013 & 0.0012 & 0.0012 & 0.0010 & 0.0013 \\ 0.2810 & 0.4686 & 0.0012 & 0.4355 & 0.0010 & 0.0013 \end{bmatrix} \quad p_{m}^{*} = \begin{bmatrix} 0.4590 \\ 0.0869 \\ 0.2576 \\ 0.0012 \\ 0.0012 \\ 0.1941 \end{bmatrix}$$

Figure 7 shows the convergence of the portfolio strategies.

5.Conclusion and future work

Financial market research has grown in importance as a result of the use of advanced mathematical techniques in decision-making. The significance of using the right modeling policy to address the portfolio optimization problem has expanded due to the rapidly expanding variety of financial assets.

A Markovian model of the financial market was presented. A brief explanation of the model was followed by a solution to the problem of incomplete information. We take into account a one-period model in which trading in securities occurs in discrete time increments. Securities are traded by investors, who track their short-term pricing. We assumed that the market does not permit short selling, that all assets have equal selling and purchasing prices, that there are expenses associated with trading, and that investors can transact in an infinite number of shares on an arbitrage-free market. The use of partially observed Markov chains has been used to resolve all of these problems.

Within the parameters of the study, we introduced a unique method for partially observed Markov chains based on observer design. As far as we are aware, this is the first piece of work that demonstrates how to create an observer for missing data. The product of the observer and distribution vector was added as a new variable. To get the necessary variables, we developed the equations. The resultant $q_{m|i}$, is a non-trivial solution to the portfolio optimization issue and is regarded as the manager′s intuition during the decision-making process. In many decision-making processes related to strategic management, intuitive rationality is seen as being crucial.

REFERENCES

[1] **M. Akian, A. Sulem, and M. Taksar(2001),*Dynamic Optimization of Long-Term Growth Rate for a Portfolio with Transaction Costs and Logarithmic Utility*.** *Mathematical Finance*, 11:152–188;

[2] **E.Asiain, J. B. Clempner and A. S. Poznyak(2018),*A Reinforcement Learning Approach for Solving the Mean Variance Customer Portfolio in Partially Observable Models*.** *International Journal on Artificial Intelligence Tools,* 27(08), 1850034;

[3] **J. B. Clempner(2016),** ***Necessary and Sufficient Karush-Kuhn-Tucker Conditions for Multiobjective Markov Chains Optimality***. *Automatica*, 71:135–142;
[4] **J. B. Clempner and A. S. Poznyak(2014),*Simple Computing of the Customer Lifetime Value: A Fixed Local-Optimal Policy Approach***. *Journal of Systems Science and Systems Engineering,* 23(4):439–459;
[5] **J. B. Clempner and A. S. Poznyak(2018)**,***Computing Multiobjective Markov Chains Handled by the Extraproximal Method***. *Annals of Operations Research*, 271:469–486;
[6] **J. B. Clempner and A. S. Poznyak(2018)**,***Constructing the Pareto Front for Multiobjective Markov Chains Handling a Strong Pareto Policy Approach***. *Journal of Computational and Applied Mathematics*, 37:567–591;
[7] **J. B. Clempner and A. S. Poznyak(2018),*Sparse Mean-Variance Customer Markowitz Portfolio Selection for Markov Chains: A Tikhonov's Regularization Penalty Approach***. *Optimization and Engineering*, **19**, 383–417;
[8] **J. B. Clempner and A. S. Poznyak (2018),*A Tikhonov Regularization Parameter Approach for Solving Lagrange Constrained Optimization Problems***. *Engineering Optimization*, 50(11):1996–2012;
[9] **J. B. Clempner and A. S. Poznyak(2018),*A Tikhonov Regularized Penalty Function Approach for Solving Polylinear Programming Problems.*** *Journal of Computational and Applied Mathematics*,328:267–286;
[10] **J. B. Clempner and A. S. Poznyak(2018),*Observer and Control Design in Partially Observable Finite Markov Chains***. *Automatica*,110:108587;
[11] **G. Constantinides(1979),*Multiperiod Consumption and Investment Behavior with Convex Transaction Costs***. *Management Science*, 25:1127–1137;
[12] **J. Cvitanic and I. Karatzas(1996),** ***Hedging and Portfolio Optimization under Transaction Costs: A Martingale Approach***. *Mathematical Finance*, 6:133–166;
[13] **M. Davis and A. Norman(1990),** ***Portfolio Selection with Transaction Costs.*** *Mathematics of Operations Research*, 15:676–713;
[14] **F. Domínguez and J. B. Clempner(2019),** ***Multiperiod Mean-variance Customer Constrained Portfolio Optimization for Finite Discrete-time Markov Chains***. *Economic Computation and Economic Cybernetics Studies and Research; ASE Publishing,* 1:39–56;
[15] **M. Garcia-Galicia, A. A. Carsteanu and J. B. Clempner(2019),** ***Continuous-time Learning Method for Customer Portfolio with Time Penalization***. *Expert Systems with Applications*, 129:27–36;
[16] **M. Garcia-Galicia, A. A. Carsteanu and J. B. Clempner (2019),** ***Continuous-time Mean Variance Portfolio with Transaction Costs: A Proximal Approach Involving Time Penalization***. *International Journal of General Systems*, 48:91–111;

[17] **H. Liu(2005),** ***Optimal Consumption and Investment eith Transaction Costs and Multiple Risky Assets***. *Jornal of Finance*, 59:289–338;
[18] **H. Markowitz(1952),*Portfolio Selection*.** *Journal of Finance*, 7:77–91;
[19**] S. Meghwani and M. Thakur(2018),** ***Multi-objective Heuristic Algorithms for Practical Portfolio Optimization and Rebalancing with Transaction Cost***. *Applied Soft Computing*, 67:865–894;
[20**] L. L. Ortiz-Cerezo, A. A. Carsteanu, J. B. Clempner(2022),*Sharpe-Ratio Portfolio in Controllable Markov Chains: Analytic and Algorithmic Approach for Second Order Cone Programming***. *Mathematics,* 10(18):3221;
[21] **P. Nowak and M. Romaniuk(2018),** ***Valuing Catastrophe Bonds Involving Correlation and CIR Interest Rate Model***. *Journal of Computational and Applied Mathematics*, 37:365–394;
[22**] P. Samuelson(1969),** ***Lifetime Portfolio Selection by Dynamic Stochastic Programming***. *The Review of Economics and Statistics*, 51:239–246;
[23] **E. M. Sánchez, J. B. Clempner and A. S. Poznyak(2015),** ***A Priori-Knowledge/Actor-Critic Reinforcement Learning Architecture for Computing the Mean'variance Customer Portfolio: The Case of Bank Marketing Campaigns***. *Engineering Applications of Artificial Intelligence*, 46, Part A:82–92;
[24**] E. M. Sánchez, J. B. Clempner and A. S. Poznyak(2015),*Solving the Mean-Variance Customer Portfolio in Markov Chains Using Iterated Quadratic/Lagrange Programming: A Credit-Card Customer-Credit Limits Approach.*** *Expert Systems with Applications*, 42(12):5315–5327;
[25] **S. S. Zhu, D. Li and S. Y. Wang,(2004),*Risk Control over Bankruptcy in Dynamic Portfolio Selection: A Generalized Mean-Variance Formulation***. *IEEE Transactions on Automatic Control*, 49:447–457.