Professor Edgar VAZQUEZ, MsC
E-mail: enapoleon@esfm.ipn.mx
Instituto Politécnico Nacional
Mexico
Professor Julio B. CLEMPNER, PhD
E-mail: julio@clempner.name
Instituto Politécnico Nacional
Mexico

CUSTOMER PORTFOLIO MODEL DRIVEN BY CONTINUOUS-TIME MARKOV CHAINS: AN l_2 LAGRANGIAN REGULARIZATION METHOD

Abstract. *This paper provides a solution to the customer portfolio for a given fixed desired expected rate of return under constraints based. We restrict the solution to a class of finite, ergodic, controllable continuous-time, finite-state Markov chains. We propose a regularized Lagrange method for the portfolio representation that ensures the strong convexity of the objective function and the existence of a unique solution of the portfolio. The solution is obtained by using the standard Lagrange method introducing the positive parameters θ and δ, and the Lagrange vector-multipliers μ_0 and μ_1 for the equality and inequality constraints, respectively, and forming the Lagrangian. We prove that if the ratio $\frac{\theta_n}{\delta_n}$ tends to zero, then the solution of the original portfolio converges to a unique solution with the minimal weighted norm. We introduce a recurrent procedure based on the projection-gradient method for finding the extremal points of the portfolio. In addition, we prove the convergence of the method. A numerical example validates the effectiveness of the regularized portfolio Lagrange method.*

Keywords: *Mean-variance portfolio selection; Tikhonov regularization; Lagrange; Continuous-time Markov chains; Applications in finance.*

JEL Classification: G11, C61, C69

1.Introduction

1.1.Brief review

The mean-variance model was first proposed by Markowitz [16], its aim is to minimize the risk of the investment, expressed by the variance of the terminal wealth, with a given level of expected return for portfolio construction in a single period. The

DOI: 10.24818/18423264/54.2.20.02

mean-variance approach has become the foundation of modern finance theory and has inspired numerous extensions and applications. A relevant property of the model is that it makes possible for an investor to seek highest expected return after determining the acceptable risk level (measured by the variance of the expected return). A factor that dominates the movement of a stock is the trend of the market. To reflect the market trend, it is necessary to allow the key parameters to respond to the general market movements.

Motivated by the importance of Markov-chain solutions for solving Markowitz's portfolio [7, 10, 13, 14], there have been also continuing efforts in extending portfolio selection to continuous-time Markov models. Zhou and Yin [25] developed the continuous-time version of Markowitz's mean-variance portfolio selection with regime switching and derived the efficient portfolio and efficient frontier explicitly. Yin and Zhou [23] suggested a random regime switching delineated by a finite-state Markov chain, based on a discrete-time Markov modulated portfolio selection model where the connections between discrete-time models and their continuous-time counterpart are revealed. They proved that the process of interest yields a switching diffusion limit using weak convergence methods and devised the portfolio selection strategies for the original problem and demonstrate their asymptotic optimality. BauerleandRieder [1] solved the problem using stochastic control methods utility maximization in a continuous-time-model with switching drift and volatility. Sass and Haussmann [20] consider only a switching drift which yields a setting with partial information where they derive explicit strategies using martingale arguments and Malliavin calculus. Taksar and Zeng [21] look at the discretized model and discuss relation to the continuous-time results. Brodiea et al. [2] presented a portfolio selection as a constrained least-squares regression problem considering a penalty approach. DeMiguel et al. [14,15] considered different portfolio algorithms including both a linear and a quadratic penalty regularization parameter. Putschögland Sass[17] extended the approach of [39] considering in addition convex constraints (as no-short-selling). Fan et al. [11] justified the use of a linear regularization parameter to identify sparse and stable portfolios validating the empirical results given in [15]. Sanchez et al. [19] presented an iterated method for solving a mean-variance customer portfolio optimization problem for Markov chains employing Tikhonov's regularization method to ensure the convergence of the objective-function to a single optimal portfolio solution. On the same lines, Sanchez et al. [18] suggested a recurrent reinforcement learning approach that adjusts the Markov chains policies according to a preprocessing and an actor-critic architecture for computing the mean-variance customer portfolio. It is important to note that in this paper we present a completely different method for computing the mean-variance customer portfolio a new approach for regularization. Clempner and

DOI: 10.24818/18423264/54.2.20.02

Poznyak [5] considered a penalty-regularized expected utility and investigates the applicability of the method for computing the mean-variance Markowitz customer portfolio optimization problem. Zhang et al. [24] developed a method that combines the l_2 and l_p norm penalties as well as the nonnegativity constraints to compute the simple portfolio model. For other applications of regularization methods for portfolio selection see [3, 12].

1.2. Main results

Although several models have been used in a wide variety of situations, they have certain limitations. The fundamental problem is in the fact that the concave functional of the portfolio is not strongly concave and the evaluation of the portfolio optimization problem determines several admissible portfolio solutions. To address these shortcomings, this paper presents the following contributions:

- Suggests a solution based on a continuous-time finite-state customer portfolio.
- Proposes an l_2 regularized Lagrange method for the portfolio representation.
- Provides a poly-linear programming problem formulation of the problem
- Shows that there exists a positive regularization parameter δ for which the Hessian matrix is strictly positive definite.
- Proves that if the ratio $\frac{\theta_n}{\delta_n} \downarrow 0$ then, the solution of the original portfolio converges to a unique solution with the minimal weighted norm.
- Introduces a recurrent procedure based on the projection-gradient method.
- Shows the convergence to a unique portfolio.

1.3.Organization of the paper

The reminder of the paper is organized as follows. The next Section presents the formulation of the problem. Section 3 describes the regularization method for Markowitz's portfolio presenting the Theorems that describes the dependence of the saddle-point of the regularized portfolio Lagrange function on the regularizing parameters δ, θ and analyses its asymptotic behavior. A continuous-time Markov portfolio approach is suggested in Section 4. Section 5 computes the Markov portfolio presenting the convergence of the method. A numerical example is presented in Section 6. Section 7 concludes the paper.

2.Regularized portfolio problem formulation

The goal is to find a unique admissible portfolio $x \in X_{adm}$ (among all the admissible portfolios) whose expected terminal wealth is $E(x)$ so that the risk measured by the variance of the terminal wealth

$$Var(x) := E[x - Ex]^2$$

DOI: 10.24818/18423264/54.2.20.02

is minimized. Finding such a portfolio $x \in X_{adm}$ where

$$f(x) = -E(x) + \frac{\xi}{2} Var(x) \to \min_{x \in X_{adm}}$$

is referred to as the *mean-variance portfolio* selection problem where $\frac{\xi}{2}$ can be any positive number. Specifically, we have the following poly-linear programming problem re-formulation

$$\left.\begin{array}{l} f(x) = \alpha_1 \sum_{j_1=1}^{N} c_{j_1} x_{j_1} + \alpha_2 \sum_{j_1=1}^{N} \sum_{j_2=1}^{N} c_{j_1,j_2} x_{j_1} x_{j_2} \\ +\alpha_3 \sum_{j_1=1}^{N} \sum_{j_2=1}^{N} \sum_{j_3=1}^{N} c_{j_1,j_2,j_3} x_{j_1} x_{j_2} x_{j_3} + \ldots + \\ \alpha_{N-1} \sum_{j_1=1}^{N} \sum_{j_2=1}^{N} \cdot\cdot \sum_{j_{N-1}=1}^{N} c_{j_1,..,j_{N-1}} x_{j_1} \ldots x_{j_{N-1}} + \\ \alpha_N \sum_{j_1=1}^{N} \sum_{j_2=1}^{N} \cdot\cdot \sum_{j_N=1}^{N} c_{j_1,..,j_N} x_{j_1} \cdots x_{j_N} \to \min_{x \in X_{adm}} \\ \alpha_j = \{0;1\} (j = 1, \ldots, N) \text{ are binary variables} \\ X_{adm} := \{x \in R^N : x \geq 0, V_0 x = b_0 \in R^{M_0}, V_1 x \leq b_1 \in R^{M_1}\} \\ \text{is a bounded set} \end{array}\right\} \quad (1)$$

Notice that this portfolio problem may have non-unique solution and $\det(V_0^{\intercal} V_0) = 0$. The problem (1) is called *feasible* if there is at least one portfolio satisfying all the constraints. An optimal portfolio to the above problem is called an *efficient portfolio* corresponding to $x \in X_{adm}$. Define by $X^* \subseteq X_{adm}$ the set of all solutions of the problem (1). The set of all the efficient points $x \in X^*$ is called the *efficient frontier*.

Following [6], let us consider the Markowitz *Regularized Portfolio Lagrange function (RPLF)*

$$\mathcal{L}_{\theta,\delta}(x, \mu_0, \mu_1) := \theta f(x) + \mu_0^{\intercal}(V_0 x - b_0) + \mu_1^{\intercal}(V_1 x - b_1) + \frac{\delta}{2}(\|x\|^2 - \|\mu_0\|^2 - \|\mu_1\|^2) \quad (2)$$

where the parameters θ, δ are positive and the Lagrange vector-multipliers $\mu_1 \in \mathbb{R}^{M_1}$

DOI: 10.24818/18423264/54.2.20.02

are non-negative and the components of $\mu_0 \in \mathbb{R}^{M_0}$ may have any sign. Obviously, the optimization problem

$$\mathcal{L}_{\theta,\delta}(x,\mu_0,\mu_1) \to \min_{x\in X_{adm}} \max_{\mu_0,\mu_{1\geq 0}} \quad (3)$$

has a unique saddle-point on x since the optimized RPLF (2) is *strongly convex* ([35]) if the parameters θ and $\delta > 0$ provide the condition

$$\frac{\partial^2}{\partial x \partial x^{\intercal}} \mathcal{L}_{\theta,\delta}(x,\mu_0,\mu_1) > 0, \forall x \in X_{adm} \subset \mathbb{R}^N \quad (4)$$

and is strongly concave on the Lagrange multipliers μ_0, μ_1 for any $\delta > 0$. In view of these properties RLF has the unique saddle point $\left(x^*(\delta), \mu_0^*(\theta,\delta), \mu_1^*(\theta,\delta)\right)$ for which the following inequalities hold: for any μ_0, μ_1 with nonnegative components and any $x \in \mathbb{R}^n$

$$\mathcal{L}_{\theta,\delta}\left(x, \mu_0^*(\theta,\delta), \mu_1^*(\theta,\delta)\right) \geq \mathcal{L}_{\theta,\delta}\left(x^*(\delta), \mu_0^*(\theta,\delta), \mu_1^*(\theta,\delta)\right) \geq \mathcal{L}_{\theta,\delta}(x^*(\delta), \mu_0, \mu_1) \quad (5)$$

As for the non-regularized function $\mathcal{L}_{1,0}(x,\mu_0,\mu_1)$, it may have several (not necessarily unique) saddle points $(x^*, \mu_0^*, \mu_1^*) \in X^* \otimes \Lambda^*$.

Given the Markowitz *Regularized Portfolio Lagrange function* represented in Eq. (2). If the parameter θ and the regularizing parameter δ tends to zero by a particular manner ($0 < \theta_n \downarrow 0, \frac{\theta_n}{\delta_n} \downarrow 0$ when $n \to \infty$), then we may expect that $x^*(\theta,\delta)$ and $\mu_0^*(\theta,\delta), \mu_1^*(\theta,\delta)$, which are the solutions of the min-max portfolio optimization problem (3) tend to the set $X^* \otimes \Lambda^*$ of all saddle points of the original portfolio optimization problem (1), that is,

$$\rho\{x^*(\theta,\delta), \mu_0^*(\theta,\delta), \mu_1^*(\theta,\delta); X^* \otimes \Lambda^*\} \underset{\theta,\delta\downarrow 0}{\to} 0 \quad (6)$$

where $\rho\{a; X^* \otimes \Lambda^*\}$ is the Hausdorff distance defined as

$$\rho\{a; X^* \otimes \Lambda^*\} = \min_{z^* \in X^* \otimes \Lambda^*} \|a - z^*\|^2$$

In the next Section we will define exactly how the parameters θ and δ should tend to zero to provide the property (6).

3.Uniqueness of the Markowitz portfolio

The next theorems describe the dependence of the saddle-point $x^*(\theta,\delta)$ and $\mu_0^*(\theta,\delta), \mu_1^*(\theta,\delta)$ of the RPLF on the regularizing parameters δ, θ.

Theorem 3.1 *Let the portfolio be defined as*

$$\mathcal{L}_{\theta,\delta}(x,\mu_0,\mu_1) := f(x) + \mu_0^{\intercal}(V_0 x - b_0) + \mu_1^{\intercal}(V_1 x - b_1) + \frac{\delta}{2}(\|x\|^2 - \|\mu_0\|^2 - \|\mu_1\|^2)$$

Then, there exists a positive parameter δ for which the Hessian matrix $H :=$

DOI: 10.24818/18423264/54.2.20.02

$\frac{\partial^2}{\partial x \partial x^\intercal} \mathcal{L}_{\theta,\delta}(x, \mu_0, \mu_1)$ is strictly positive definite.

Proof. First, let us prove that the Hessian matrix $H := \frac{\partial^2}{\partial x \partial x^\intercal} \mathcal{L}_{\theta,\delta}(x, \mu_0, \mu_1)$ is strictly positive definite for all $x \in \mathbb{R}^N$ and for some positive θ and δ, satisfying a special relation, namely, $H > 0$. We have

$$\frac{\partial^2}{\partial x^2} \mathcal{L}_{\theta,\delta}(x, \mu_0, \mu_1) = \theta \frac{\partial^2}{\partial x^2} f(x) + \delta I_{N\times N} \geq \delta \left(1 + \frac{\theta}{\delta} \lambda^-\right) I_{N\times N} > 0 \forall \delta > \theta |\lambda^-|$$

$$\lambda^- := \min_{x \in X_{adm}} \lambda_{\min} \left(\frac{\partial^2}{\partial x^2} f(x)\right)$$

fulfilling the property $H > 0$ if $\delta > \theta |\lambda^-|$. This means that RLF (2) is strongly convex on x and, hence, has a unique minimal point defined below as x^*.

Theorem 3.2 *Given the Portfolio Lagrange function represented in Eq. (2). Let us assume that*

1. the bounded set X^* of all solutions of the original portfolio optimization problem (1) is not empty and the Slater's condition holds, that is, there exists a point $x \in X_{adm}$ such that

$$V_1 x < b_1 \tag{7}$$

2. The parameters θ and δ are time-varying, i.e.,

$$\theta = \theta_n, \delta = \delta_n (n = 0,1,2, \dots .)$$

such that

$$0 < \theta_n \downarrow 0, \frac{\theta_n}{\delta_n} \downarrow 0 \text{ when} n \to \infty \tag{8}$$

Then

$$\left.\begin{array}{l} x_n^* := x^*(\theta_n, \delta_n) \underset{n\to\infty}{\to} x^{**} \\ \mu_0^*(\theta_n, \delta_n) \underset{n\to\infty}{\to} \mu_0^{**} \\ \mu_1^*(\theta_n, \delta_n) \underset{n\to\infty}{\to} \mu_1^{**} \end{array}\right\} \tag{9}$$

where $x^{**} \in X^*$ and $(\mu_0^{**}, \mu_1^{**}) \in \Lambda^*$, define the solution of the original problem (1) with a unique minimal norm, i.e.,

$$\left.\begin{array}{l} \|x^{**}\|^2 + \|\mu_0^{**}\|^2 + \|\mu_1^{**}\|^2 \leq \|x^*\|^2 + \|\mu_0^*\|^2 + \|\mu_1^*\|^2 \\ \\ \text{for all} x^* \in X^*, (\mu_0^*, \mu_1^*) \in \Lambda^* \end{array}\right\} \tag{10}$$

Proof.

In view of the properties

DOI: 10.24818/18423264/54.2.20.02

$$\left(\nabla f(x),(y-x)\right) \leq f(y)-f(x)$$
$$\left(\nabla f(x),(x-y)\right) \geq f(x)-f(y)$$

valid for any convex function $f(x)$ and any $x, y,$ for RPLF at any admissible points x, μ_0, μ_1 and $x_n^* = x^*(\theta_n, \delta_n),\ \mu_{0,n}^* = \mu_0^*(\theta_n, \delta_n), \mu_{1,n}^* = \mu_1^*(\theta_n, \delta_n)$ we have

$$\begin{gathered}\left(x-x_n^*, \frac{\partial}{\partial x}\mathcal{L}_{\theta_n,\delta_n}(x,\mu_0,\mu_1)\right)-\left(\mu_0-\mu_{0,n}^*, \frac{\partial}{\partial \mu_0}\mathcal{L}_{\theta_n,\delta_n}(x,\mu_0,\mu_1)\right)- \\ \left(\mu_1-\mu_{1,n}^*, \frac{\partial}{\partial \mu_1}\mathcal{L}_{\theta_n,\delta_n}(x,\mu_0,\mu_1)\right)=\mathcal{L}_{\theta_n,\delta_n}\left(x,\mu_{0,n}^*,\mu_{1,n}^*\right)- \\ \mathcal{L}_{\theta,\delta}(x_n^*,\mu_0,\mu_1)+\frac{\delta_n}{2}+\left(\|x-x_n^*\|^2+\left\|\mu_0-\mu_{0,n}^*\right\|^2+\left\|\mu_1-\mu_{1,n}^*\right\|^2\right)\end{gathered} \tag{11}$$

which by the saddle-point condition (5) implies

$$\begin{gathered}\theta_n(x-x_n^*)^{\intercal}\frac{\partial}{\partial x}f(x)+(x-x_n^*)^{\intercal}[V_0^{\intercal}\mu_0+V_1^{\intercal}\mu_1+\delta_n x]+ \\ \left(\mu_0-\mu_{0,n}^*\right)^{\intercal}(\delta_n-V_0x+b_0)+\left(\mu_1-\mu_{1,n}^*\right)^{\intercal}(\delta_n-V_1x+b_1)\geq \\ \frac{\delta_n}{2}\left(\|x-x_n^*\|^2+\left\|\mu_0-\mu_{0,n}^*,\right\|^2+\left\|\mu_1-\mu_{1,n}^*\right\|^2\right)\end{gathered} \tag{12}$$

Selecting in (12) $x := x^* \in X^*$ (x^* is one of admissible solutions such that $V_0 x^* = b_0$ and $V_1 x^* \leq b_1$) and $\mu_0 = \mu_0^*, \mu_1 = \mu_1^*$ in view of the complementary slackness conditions

$$(\mu_1^*)_i (V_1 x^* - b_1)_i = \left(\mu_{1,n}^*\right)_i (V_1 x_n^* - b_1)_i = 0$$

we obtain

$$\begin{gathered}\theta_n(x^*-x_n^*)^{\intercal}\frac{\partial}{\partial x}f(x^*)+(x^*-x_n^*)^{\intercal}[V_0^{\intercal}\mu_0^*+V_1^{\intercal}\mu_1^*+\delta_n x^*]+ \\ \left(\mu_0^*-\mu_{0,n}^*\right)^{\intercal}(\delta_n\mu_0^*-V_0x^*+b_0)+\left(\mu_1^*-\mu_{1,n}^*\right)^{\intercal}(\delta_n\mu_1^*-V_1x^*+b_1)\geq \\ \frac{\delta_n}{2}\left(\|x^*-x_n^*\|^2+\left\|\mu_0^*-\mu_{0,n}^*\right\|^2+\left\|\mu_1^*-\mu_{1,n}^*\right\|^2\right)\geq 0\end{gathered}$$

Simplifying the last inequality, we have

$$\begin{gathered}\theta_n(x^*-x_n^*)^{\intercal}\frac{\partial}{\partial x}f(x^*)+\delta_n(x^*-x_n^*)^{\intercal}x^*+\delta_n\left(\mu_0^*-\mu_{0,n}^*\right)^{\intercal}\mu_0^*+ \\ \left(\mu_1^*-\mu_{1,n}^*\right)^{\intercal}\delta_n\mu_1^*\geq 0\end{gathered}$$

Dividing both sides of this inequality by δ_n and taking $\frac{\theta_n}{\delta_n} \underset{n\to\infty}{\to} 0$ we get

DOI: 10.24818/18423264/54.2.20.02

$$0 \leq \limsup_{n\to\infty}\left[(x^* - x_n^*)^\intercal x^* + \left(\mu_0^* - \mu_{0,n}^*\right)^\intercal \mu_0^* + \left(\mu_1^* - \mu_{1,n}^*\right)^\intercal \mu_1^*\right] \quad (13)$$

This means that there necessarily exist subsequences δ_k and $\theta_k (k \to \infty)$ on which there exist the limits

$$x_k^* = x^*(\theta_k, \delta_k) \to \tilde{x}^*, \mu_{0,k}^* = \mu_0^*(\theta_k, \delta_k) \to \tilde{\mu}_0^*$$

$$\mu_{1,k}^* = \mu_1^*(\theta_k, \delta_k) \to \tilde{\mu}_1^* \text{as} k \to \infty$$

Suppose that there exist two limit points for two different convergent subsequences, i.e., there exist the limits

$$x_{k'}^* = x^*(\theta_{k'}, \delta_{k'}) \to \bar{x}^*, \mu_{0,k'}^* = \mu_0^*(\theta_{k'}, \delta_{k'}) \to \bar{\mu}_0^*$$

$$\mu_{1,k'}^* = \mu_1^*(\theta_{k'}, \delta_{k'}) \to \bar{\mu}_1^* \text{as} k \to \infty$$

Then, on these subsequences we have

$$0 \leq (x^* - \tilde{x}^*)^\intercal x^* + (\mu_0^* - \tilde{\mu}_0^*)^\intercal \mu_0^* + (\mu_1^* - \tilde{\mu}_1^*)^\intercal \mu_1^*$$

$$0 \leq (x^* - \bar{x}^*)^\intercal x^* + (\mu_0^* - \bar{\mu}_0^*)^\intercal \mu_0^* + (\mu_1^* - \bar{\mu}_1^*)^\intercal \mu_1^*$$

From this inequalities it follows that $(\tilde{x}^*, \tilde{\mu}_0^*, \tilde{\mu}_1^*)$ and $(\bar{x}^*, \bar{\mu}_0^*, \bar{\mu}_1^*)$ correspond to the minimum point of the function

$$s(x^*, \mu_0^*, \mu_1^*) := \frac{1}{2}(\|x^*\|^2 + \|\mu_0^*\|^2 + \|\mu_1^*\|^2)$$

defined on $X^* \otimes \Lambda^*$ for all possible saddle-points of the non-regularized Lagrange function. But the function $s(x^*, \mu_0^*, \mu_1^*)$ is strictly convex, and, hence, its minimum is unique that gives $\tilde{x}^* = \bar{x}^*, \tilde{\mu}_0^* = \bar{\mu}_0^*, \tilde{\mu}_0^* = \bar{\mu}_0^*$. The theorem is proved.

The following property also takes place.

Lemma 3.3 *Under the assumptions of the Theorem 3.2 there exist positive constants C_μ and C_δ such that*

$$\|x_n^* - x_m^*\| + \left\|\mu_{0,n}^* - \mu_{0,m}^*\right\| + \left\|\mu_{1,n}^* - \mu_{1,m}^*\right\| \leq C_\theta|\theta_n - \theta_m| + C_\delta|\delta_n - \delta_m| \quad (14)$$

Proof. It follows also from the necessary and sufficient conditions (11) for the points $x_n^* = x^*(\theta_n, \delta_n)$, $\mu_{0,n}^* = \mu_0^*(\theta_n, \delta_n), \mu_{1,n}^* = \mu_1^*(\theta_n, \delta_n)$ to be the extremal points of the function $\mathcal{L}_{\theta_n,\delta_n}(x, \mu_0, \mu_1)$.

Corollary 1 *Given the portfolio Lagrange function represented in Eq. (2) we have that:*

1. $\mathcal{L}_{\theta,\delta}(x, \mu_0, \mu_1)$ is strictly convex on $x_n^* = x^*(\theta_n, \delta_n)$, and strictly concave on $\mu_0^* = \mu_0^*(\theta_n, \delta_n)$, and $\mu_1^* = \mu_1^*(\theta_n, \delta_n)$
2. The Markowitz portfolio admits a unique solution.

DOI: 10.24818/18423264/54.2.20.02

4.Continuous-time Markov portfolio

4.1.Continuous-time Markov process

In this section we introduce the (continuous-time, discrete-state) Markov chains we are interested in [4,22].

Let $\{X(t), t \geq 0\}$ a stochastic process that satisfies the *Markov property* if, letting $\mathcal{F}_{X(\tau)}$ denote all the information pertaining to the history of X up to time τ, and $\tau \leq t P(X(t)|\mathcal{F}_{X(\tau)}) = P(X(t)|X(\tau))$ we say that the process is time homogeneous if given $\tau, t' \leq t,\ t = t' + \tau$. In other words, this property means that any distribution in the future depends only on the value $X(\tau)$ and is independent on the past values.

Throughout the remainder

$$CTMDP = (S, A, \{A(s)\}_{s\in S}, Q, u) \quad (15)$$

stands for a *continuous-time Markov decision process* (CTMDP), where the state-space S is a *finite* set $\{s_1, \ldots, s_N\}$, for some $N \in \mathbb{N}$ indexed by $i = \overline{1, N}$, and the finite action set $A = \{a_1, \ldots, a_M\}$ is the action (or control) space for some $M \in \mathbb{N}$ indexed by $k = \overline{1, M}$.

For each $s \in S$, $A(s) \subset A$ is the nonempty set of admissible actions at s. Whereas, the set $\mathbb{K} := \{(s, a): s \in S, a \in A(s)\}$ is the class of admissible pairs, which is considered as a subspace of $S \times A$.

The matrix $Q = \left[q_{j|ik}\right]_{i,j=\overline{1,N},k=\overline{1,M}}$ denotes the transition rates which satisfy that $q_{j|ik} \geq 0$ for all $s \in S$ and $j \neq i$. The transition rates $q_{j|ik}$ are conservative, i.e., $\sum_{j=1}^{N} q_{j|ik} = 0$ and stable, which means that $q_i^* := \sup_{a\in A(i)} q_i(a) < \infty \forall i \in S$ where $q_i := -q_{i,i} \geq 0$ for all $s \in S$.

Finally, $u \in \mathcal{B}(\mathbb{K})$ is the (measurable) one-stage utility function.

We denote the probability transition matrix by $\Pi(t) = [\pi_{(t',i,\tau,j,k)}]_{i,j=\overline{1,N},k=\overline{1,M}}$, $\tau \geq t'$ such that, $\pi_{(t',i,\tau,j,k)} = \pi_{(0,i,t,j,k)}, t = \tau - t' \forall s_i, s_j \in S$ and where $\sum_{j=1}^{N} \pi_{j|ik} = 1,\ \forall s_i \in S$.

The Kolmogorov forward equations, can be written as the matrix differential equation: $\Pi'(t) = \Pi(t)Q; \Pi(0) = I \Pi(t) \in \mathbb{R}^{N\times N}$, $I \in \mathbb{R}^{N\times N}$ is the identity matrix. This system can be solved by $\Pi(t) = \Pi(0)e^{Qt} = e^{Qt} := \sum_{n=0}^{\infty} \frac{t^n Q^n}{n!}$ and at the stationary state, the probability transition matrix is defined as

$$\Pi^* = \lim_{t\to\infty} \Pi(t) \quad (16)$$

We also point out that given a state-space S, the infinitesimal generator Q completely determines the CTMC as it contains all the local information pertaining to the transitions rates λ_{ij}. Thus, it is sufficient to characterize a chain by simply providing a state-space, S, and the generator Q.

DOI: 10.24818/18423264/54.2.20.02

Definition 4.1. *The vector $p \in \mathbb{R}^N$ is called stationary distribution vector if*

$$\Pi^{\top *} p = p \quad (17)$$

where $\sum_{i=1}^{N} p_i = 1$ and $p_i = P(X(t) = s_i)$ this vector can be seen as the long run proportion of time that the process is in state $s_i \in S$.

Theorem 4.2. *Let $X(t)$ be an irreducible and recurrent CTMC then the following statements are equivalent: $Q^{\top} p = 0$ and $\Pi^{\top *} p = p$.*

Proof. See [22].

A *strategy* is then defined as a sequence $d = \{d(t), t \geq 0\}$ of stochastic kernels $d(t)$ such that: a) for each time $t \geq 0 d_{k|i}(t)$ is a probability measure on A such that $d_{A(i)|i}(t) = 1$ and, b) for every $E \in \mathcal{B}(A) d_{E|i}(t)$ is a Borel measurable function in $t \geq 0$. We denoted by D the family of all strategies. From now on, we will consider only stationary strategies $d_{k|i}(t) = d_{k|i}$.

For each action the matrix $Q(a_k) := [q_{j|ik}]$, $a_k \in A$ denotes the transition rates matrix for the action a_k such that

$$q_{ji}(a_k)] := [q_{j|ik}] = \begin{cases} -\sum_{i \neq j} \lambda_{ij}(a_k), & if \quad i = j \\ \lambda_{ij}(a_k), & if \quad i \neq j \end{cases}$$

while, for each strategy d the associated transition rate matrix is defined as:

$$Q(d) := [q_{ji}(d)] = \sum_{k=1}^{M} q_{j|ik} d_{k|i} \quad (18)$$

such that on a stationary state distribution for all $d_{k|i}$ and $t \geq 0$ from Eq. (16) we have that, $\Pi^*(d) = \lim_{t \to \infty} e^{Q(d)t}$, where $\Pi^*(d)$ is a stationary transition controlled matrix.

$$\Pi^*(d) := [\pi_{ji}(d)] = \sum_{k=1}^{M} \pi_{j|ik} d_{k|i}$$

Pondering the long-run expected average reward over the states at steady state, the following linear functional $U(d)$ under the fixed strategy $d_{k|i}(t) = d_{k|i}$ can be defined as follows

$$U(d) := \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{M} u_{ik} \pi_{j|ik} d_{k|i} p_i \quad (19)$$

In an ergodic chain, we consider a stationary distribution and by the property in Eq. (17), the value function in Eq. (19), under a given optimal policy is:

$$U(d) = \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{M} u_{ik} \pi_{j|ik} p_i d_{k|i} \longrightarrow \max_{d} \quad (20)$$

4.2. Feasibility of Markowitz's portfolio

The model presented in Eq. (20) is nonlinear. Then, if we introduce a new decision variable $c_{i|k}$, called the *joint strategy variable*, then the problem presented above can be reformulated as a linear programming as follows

$$U(d) = \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{M} u_{ik} \pi_{j|ik} p_i d_{k|i} = \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{M} u_{ik} \pi_{j|ik} c_{i|k}$$

where $c_{i|k} = p_i d_{k|i}$. The variable c_{ik} belongs to the set of matrices C_{adm} and it is restricted by the following constraints:

DOI: 10.24818/18423264/54.2.20.02

1. Each vector from the matrix $c := [c_{i|k}]$ represents a stationary mixed strategy that belongs to the simplex

$$S^{N\times M} := \left\{c \in \mathbb{R}^{N\times M} : \text{for}\, c_{i|k} \geq 0 \ \ \text{where} \sum_{i=1}^{N}\sum_{k=1}^{M} c_{i|k} = 1\right\} \quad (21)$$

2. The variable c_{ik} satisfies the ergodicity constraints, i.e.:

$$p_j(d) = \sum_{i=1}^{N}\sum_{k=1}^{M} \pi_{j|ik} p_i d_{k|i}$$

that in terms of $c_{i|k}$ takes the form:

$$g(c) = \sum_{i=1}^{N}\sum_{k=1}^{M} \pi_{j|ik} c_{i|k} - \sum_{k=1}^{M} c_{j|k} = 0 \quad (22)$$

3. From Eq. (18) we obtain:

$$\sum_{i=1}^{N} q_{j|i}(d) p_i(q) = \sum_{i=1}^{N}\sum_{k=1}^{M} q_{j|ik} d_{k|i} p_i = 0$$

this expression in terms of $c_{i|k}$ takes the form:

$$h(c) = \sum_{i=1}^{N}\sum_{k=1}^{M} q_{j|ik} c_{i|k} = 0 \quad (23)$$

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

$$p_i(d) = \sum_{k=1}^{M} c_{i|k} \qquad d_{i|k} = \frac{c_{i|k}}{\sum_{k=1}^{M} c_{i|k}}$$

Then, in terms of c-variables the reward function U becomes:

$$U(c) = \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{M} u_{ik}\pi_{j|ik} c_{i|k}$$

4.3.Markov portfolio formulation

One may formally state Markowitz's decision model for mean-variance customer portfolio as follows. We consider diversification with respect to the number of customers chosen in the portfolio problem. Then, $\alpha_{i|k}$ is the *number of customers* at state i applying action k, $0 \leq \alpha_{(i|k)}(n)$. The *utility* U of a customer portfolio is calculated as the sum of the weighted net presenting the value u as follows

$$U(\alpha, c) = \sum_{i=1}^{N}\sum_{k=1}^{M} W_{i|k}\alpha_{i|k} c_{i|k} \to \max_{\alpha\in A_{adm} c\in C_{adm}}$$

where

$$W_{i|k} := \sum_{j=1}^{N} u_{j|ik}\pi_{j|ik} \text{and} c_{i|k} := d_{k|i} p_i.$$

The customer portfolio optimization problem attempts to maximize the *mean value* ($U(\alpha, c)$) generated by all the customers while minimizing the *variance* ($\mathrm{Var}(U(\alpha, c))$)

$$\mathrm{Var}(\alpha, c) := \sum_{i=1}^{N}\sum_{k=1}^{M} \left[\alpha_{i|k} W_{i|k} - U(\alpha, c)\right]^2 c_{i|k} = \sum_{i=1}^{N}\sum_{k=1}^{M} \alpha_{i|k}^2 W_{i|k}^2 c_{i|k} - U^2(\alpha, c) \ \ \to \min_{\alpha\in A_{adm} c\in C_{adm}}$$

For practical purposes, the resulting *customer portfolio optimization* problem includes a

DOI: 10.24818/18423264/54.2.20.02

model-user's tolerance for risk, and it is represented by the following expression:

$$\Phi(\alpha, c) := U(\alpha, c) - \frac{\xi}{2}\mathrm{Var}(\alpha, c) \to \max_{\substack{\alpha\in A_{adm} \\ c\in C_{adm}}} \tag{25}$$

where

$$A_{adm} = \left\{\alpha = [\alpha_{i|k}] : \sum_{i=1}^{N} \sum_{k=1}^{M} \alpha_{i|k} \le \alpha^{+}, \alpha_{i|k} \in [\varepsilon, \alpha^{+}], \varepsilon > 0\right\} \tag{26}$$

and $\frac{\xi}{2}$ is a *monetary factor* that reflects the price per unit of risk. Because, the purpose is to obtain a higher mean value return $(U(\alpha, c))$ also the corresponding risk level $\mathrm{Var}(U(\alpha, c)$in creases. Here the goal is to find the values of α and c that maximize the objective function in Eq. (25) subject to the following constrains:

$$\sum_{i=1}^{N} \sum_{k=1}^{M} \alpha_{i|k} c_{i|k} \eta_{i|k} \le b_{ineq} \tag{27}$$

where $\eta_{i|k}$ are the resources destined for carrying out in state i a promotion k and the admissible sets are as in Eq. (21), Eq. (22) and Eq. (23). The following optimization properties are the key to find efficient portfolios: a) the Markowitz model in Eq. (25) is a quadratic optimization problem (quadratic objective function and linear constraints in Eq. (27), Eqs. (21)-(22)-(23) and Eq. (26)), b) the feasibility set C_{adm} is convex since it is the intersection of hyperplanes, c) the factor $\frac{1}{2}$ of *the risk-aversion* parameter ξ is chosen for notational convenience and, d) the parameter b_{ineq} is endogenously given (the budget is chosen by the decision maker in the respective model). The mean-variance Markowitz portfolio given in Eq. (2) can be re-written:

$$\begin{aligned}
&\mathcal{L}_{\theta_n,\delta_n}(c, \alpha, \mu_0, \mu_{N+1}, \mu_{N+2}, \mu_1) := \\
&\theta\Big[\sum_{i=1}^{N} \sum_{k=1}^{M} W_{i|k}\alpha_{i|k}c_{i|k} + \frac{\xi}{2}\sum_{i=1}^{N} \sum_{k=1}^{M} W_{i|k}\alpha_{i|k}c_{i|k} \cdot \\
&\sum_{\hat{\imath}=1}^{N} \sum_{\hat{k}=1}^{M} W_{\hat{\imath}|\hat{k}}\alpha_{i|\hat{k}}c_{\hat{\imath}|\hat{k}} - \frac{\xi}{2}\sum_{i=1}^{N} \sum_{k=1}^{M} W_{i|k}^{2}\alpha_{i|k}^{2}c_{i|k}\Big] + \\
&\sum_{j=1}^{N} \mu_{0,j}\Big[\Big(\sum_{i=1}^{N} \sum_{k=1}^{M} \pi_{j|ik}c_{i|k} - \sum_{k=1}^{M} c_{j|k}\Big) - b_{eq,j}\Big] - \\
&\mu_{N+1}\Big(\sum_{i=1}^{N} \sum_{k=1}^{M} c_{i|k} - b_{eq,N+1}\Big) - \\
&\sum_{j=1}^{N} \mu_{N+2,j}\Big[\Big(\sum_{i=1}^{N} \sum_{k=1}^{M} q_{j|ik}c_{i|k}\Big) - b_{eq,N+1+j}\Big] \\
&+\mu_1\Big(\sum_{i=1}^{N} \sum_{k=1}^{M} \alpha_{i|k}c_{i|k}\eta_{i|k} - b_{ineq}\Big) - \\
&\frac{\delta}{2}(\|c\|^2 + \|\alpha\|^2 - \|\mu_0\|^2 - \|\mu_1\|^2 - \mu_{N+1}^2 - \mu_{N+2}^2)
\end{aligned} \tag{28$$

5. Computing the Markov portfolio

Let us denote

$$\mathcal{L}_{\theta_n,\delta_n} = -\mathcal{L}_{\theta_n,\delta_n}\big(c_n, \alpha_n, \mu_{0,n}, \mu_{N+1,n}, \mu_{N+2,n}, \mu_{1,n}\big)$$

DOI: 10.24818/18423264/54.2.20.02

Then, we have that

$$
\begin{aligned}
c_{n+1} &= \left[c_n - \gamma_{c,n} \frac{\partial}{\partial c} \mathcal{L}_{\theta_n,\delta_n}\right]_+ \\
\alpha_{n+1} &= \left[\alpha_n - \gamma_{\alpha,n} \frac{\partial}{\partial \alpha} \mathcal{L}_{\theta_n,\delta_n}\right]_+ \\
\mu_{0,n+1} &= \mu_{0,n} + \gamma_{\mu_{0,n}} \frac{\partial}{\partial \mu_0} \mathcal{L}_{\theta_n,\delta_n} \\
\mu_{N+1,n+1} &= \mu_{N+1,n} + \gamma_{\mu_{N+1,n}} \frac{\partial}{\partial \mu_{N+1}} \mathcal{L}_{\theta_n,\delta_n} \\
\mu_{N+2,n+1} &= \mu_{N+2,n} + \gamma_{\mu_{N+1,n}} \frac{\partial}{\partial \mu_{N+2}} \mathcal{L}_{\theta_n,\delta_n} \\
\mu_{1,n+1} &= \left[\mu_{1,n} + \gamma_{\mu_{1,n}} \frac{\partial}{\partial \mu_1} \mathcal{L}_{\theta_n,\delta_n}\right]_+
\end{aligned}
\tag{29}
$$

where the operator $[\cdot]_+$ acts from $\mathbb{R}^n$ into $\mathbb{R}^n$ as follows:

$$
[z]_+ = ([z_1]_+, \ldots, [z_n]_+), [z_i]_+ := \begin{cases} z_i & \text{if} \quad z_i \geq 0 \\ 0 & \text{if} \quad z_i < 0 \end{cases} \tag{30}
$$

Theorem 5.1 (on the convergence of the method) *If for some positive sequences* $\{\varepsilon_n\}$

$$
\sum_{n=0}^{\infty} \gamma_n \delta_n = \infty, \frac{|\theta_n - \theta_{n-1}|^2 + |\delta_n - \delta_{n-1}|^2}{\varepsilon_n \delta_n} \underset{n\to\infty}{\to} 0
$$

$$
\frac{\varepsilon_n}{\gamma_n \delta_n} \text{ and } \frac{\gamma_n}{\delta_n} \text{are small enough} \tag{31}
$$

then

$$
\begin{aligned}
G_n &:= \|c - c_n^*\|^2 + \|\alpha - \alpha_n^*\|^2 + \left\|\mu_0 - \mu_{0,n}^*\right\|^2 + \\
&\left\|\mu_{N+1} - \mu_{N+1,n}^*\right\|^2 + \left\|\mu_{N+2} - \mu_{N+2,n}^*\right\|^2 + \left\|\mu_1 - \mu_{1,n}^*\right\|^2 \underset{n\to\infty}{\to} 0
\end{aligned}
\tag{32}
$$

Proof. Let us denote

$$
\mathcal{L}_{\theta_n,\delta_n} = \mathcal{L}_{\theta_n,\delta_n}\left(c_n, \alpha_n, \mu_{0,n}, \mu_{N+1,n}, \mu_{N+2,n}, \mu_{1,n}\right)
$$

In view of Eq. (29) it follows

$$
\begin{aligned}
G_{n+1} \leq &\left\|c_n - \gamma_n \frac{\partial}{\partial c} \mathcal{L}_{\theta_n,\delta_n} - c_{n+1}^*\right\|^2 + \left\|\alpha_n - \gamma_{\alpha,n} \frac{\partial}{\partial \alpha} \mathcal{L}_{\theta_n,\delta_n} - \alpha_{n+1}^*\right\|^2 + \\
&\left\|\mu_{0,n} + \gamma_n \frac{\partial}{\partial \mu_0} \mathcal{L}_{\theta_n,\delta_n} - \mu_{0,n+1}^*\right\|^2 + \left\|\mu_{N+1,n} + \gamma_{\mu_{N+1,n}} \frac{\partial}{\partial \mu_{N+1}} \mathcal{L}_{\theta_n,\delta_n} - \mu_{N+1,n}\right\|^2 + \\
&\left\|\mu_{N+2,n} + \gamma_{\mu_{N+1,n}} \frac{\partial}{\partial \mu_{N+2}} \mathcal{L}_{\theta_n,\delta_n} - \mu_{N+2,n}\right\|^2 + \left\|\mu_{1,n} + \gamma_n \frac{\partial}{\partial \mu_1} \mathcal{L}_{\theta_n,\delta_n} - \mu_{1,n+1}^*\right\|^2
\end{aligned}
$$

For strongly convex (concave) functions, the following inequalities hold.

DOI: 10.24818/18423264/54.2.20.02

$$(c_n - c_n^*)^\intercal \frac{\partial}{\partial c}\mathcal{L}_{\theta_n,\delta_n} \geq \delta_n \left(1 + \frac{\theta_n}{\delta_n}\lambda_c^-\right) \|c - c_n^*\|^2$$
$$\delta_n(1-\epsilon)\|c - c_n^*\|^2, \left|\frac{\theta_n}{\delta_n}\lambda_c^-\right| \leq \epsilon$$

$$(\alpha_n - \alpha_n^*)^\intercal \frac{\partial}{\partial \alpha}\mathcal{L}_{\theta_n,\delta_n} \geq \delta_n \left(1 + \frac{\theta_n}{\delta_n}\lambda_\alpha^-\right) \|\alpha - \alpha_n^*\|^2$$
$$\delta_n(1-\epsilon)\|\alpha - \alpha_n^*\|^2, \left|\frac{\theta_n}{\delta_n}\lambda_\alpha^-\right| \leq \epsilon$$

$$\left(\mu_{0,n} - \mu_{0,n}^*\right)^\intercal \frac{\partial}{\partial \mu_0}\mathcal{L}_{\theta_n,\delta_n} \leq -\delta_n \left\|\mu_{0,n} - \mu_{0,n}^*\right\|^2$$
$$\left(\mu_{N+1,n} - \mu_{N+1,n}^*\right)^\intercal \frac{\partial}{\partial \mu_{N+1}}\mathcal{L}_{\theta_n,\delta_n} \leq -\delta_n \left\|\mu_{N+1,n} - \mu_{N+1,n}^*\right\|^2$$

$$\left(\mu_{N+2,n} - \mu_{N+2,n}^*\right)^\intercal \frac{\partial}{\partial \mu_{N+2}}\mathcal{L}_{\theta_n,\delta_n} \leq -\delta_n \left\|\mu_{N+2,n} - \mu_{N+2,n}^*\right\|^2$$

$$\left(\mu_{1,n} - \mu_{1,n}^*\right)^\intercal \frac{\partial}{\partial \mu_1}\mathcal{L}_{\theta_n,\delta_n} \leq -\delta_n \left\|\mu_{1,n} - \mu_{1,n}^*\right\|^2$$

By the Λ-inequality $2(a,b) \leq (a,\Lambda a) + (b,\Lambda^{-1}b)$ valid for any vectors a, b and any matrix $\Lambda > 0$, we get for $\Lambda = \varepsilon_n I$ the following estimate

$$2(c_n - c_n^*)^\intercal(c_n^* - c_{n+1}^*) + 2(\alpha_n - \alpha_n^*)^\intercal(\alpha_n^* - \alpha_{n+1}^*) + 2\left(\mu_0 - \mu_{0,n}^*\right)^\intercal\left(\mu_{0,n}^* - \mu_{0,n+1}^*\right) +$$
$$2\left(\mu_{N+1} - \mu_{N+1,n}^*\right)^\intercal\left(\mu_{N+1,n}^* - \mu_{N+1,n+1}^*\right) + 2\left(\mu_{N+2} - \mu_{N+2,n}^*\right)^\intercal\left(\mu_{N+2,n}^* - \mu_{N+2,n+1}^*\right) +$$
$$2\left(\mu_1 - \mu_{1,n}^*\right)^\intercal\left(\mu_{1,n}^* - \mu_{1,n+1}^*\right) \leq \varepsilon_n G_n + 2\varepsilon_n^{-1}\left(C_\theta^2|\theta_n - \theta_m|^2 + C_\delta^2|\delta_n - \delta_m|^2\right)$$

If a nonnegative $\{u_n\}$ sequence satisfies the recurrent inequality

$$u_{n+1} \leq u_n(1-\zeta_n) + \beta_n$$
$$0 < \zeta_n \leq 1, \sum_{n=0}^{\infty} \zeta_n = \infty, \frac{\beta_n}{\zeta_n} \underset{n\to\infty}{\to} p$$

then $u_n \underset{n\to\infty}{\to} p$. Defining

$$\zeta_n := 2\gamma_n\delta_n(1-\epsilon)\left(1 - \frac{1}{1-\epsilon}\frac{\varepsilon_n}{\delta_n} - \frac{1}{2(1-\epsilon)}\frac{\varepsilon_n}{\gamma_n\delta_n} - \frac{(1+L)}{2(1-\epsilon)}\frac{\gamma_n}{\delta_n}\right)$$
$$\beta_n := C\gamma_n\varepsilon_n^{-1}(|\theta_n - \theta_{n+1}|^2 + |\delta_n - \delta_{n+1}|^2)$$

and applying the conditions in Eq. (31) of this theorem for $p = 0$ we obtain the desired result. The theorem is proven.

To satisfy the convergence conditions in Eq. (31) let us select the parameters $\gamma_n, \delta_n, \varepsilon_n$ of the algorithm in Eq. (29) as small positive constants and

DOI: 10.24818/18423264/54.2.20.02

$$\theta_n = \frac{\theta_0}{1+\ln(n+1)}, \theta_0 > 0 \tag{33}$$

The condition $\frac{\theta_n}{\delta_n} \downarrow 0$ also holds. Notice that in this case the rate of convergence G_n to zero will have the same order as the term

$$|\theta_n - \theta_{n-1}|^2 \sim \frac{\theta_0^2}{n^2(\ln n)^4}(1 + o(1))$$

6.Numerical example

Let us now illustrate the practical implications of the theoretical issues discussed above. We consider the problem of maximizing the portfolio given by $f(x) = E(x) - \frac{\xi}{2} Var(x) \to \max_{c \in C_{adm}}$. We assume that the number of states is $N = 6$ and the number of actions is $M = 3$. Fixing $\theta_0 = 1 \times 10^{-9}$, $\delta_0 = 1 \times 10^{-8}$, $\gamma_{c0} = 2.6 \times 10^{-9}$, $\gamma_{\alpha 0} = 0.99$, $\gamma_{u0} = [0.01; \ldots; 0.01]$, $\gamma_{\mu_{N+1,0}} = 0.005$, $\gamma_{\mu_{N+1,0}} = [0.01; \ldots; 0.01]$, $\gamma_{\mu_{1,n}} = 1 \times 10^{-5}$, $b_{ineq} = 100$, $\xi = 0.05$ and

$$\eta_{i|k} = \begin{bmatrix} 3.0240 & 1.4599 & 1.8620 \\ 1.9963 & 2.1583 & 0.9906 \\ 2.6344 & 7.7400 & 2.4484 \\ 2.0840 & 4.9203 & 1.6975 \\ 3.2843 & 0.8358 & 4.7582 \\ 3.1399 & 0.5311 & 4.6017 \end{bmatrix}$$

The resulting optimal portfolio (see Figure 1 and Figure 2) generated by the recurrent method is given by

$$\alpha^*_{i|k} = \begin{bmatrix} 100.0020 & 100.0025 & 100.0026 \\ 100.0026 & 100.0027 & 100.0029 \\ 82.9756 & 88.1662 & 93.1698 \\ 100.0022 & 100.0026 & 100.0024 \\ 100.0028 & 100.0029 & 100.0025 \\ 88.6045 & 92.4799 & 86.0521 \end{bmatrix} \quad d^*_{i|k} = \begin{bmatrix} 0.3256 & 0.3300 & 0.3444 \\ 0.4108 & 0.2786 & 0.3106 \\ 0.3248 & 0.3231 & 0.3521 \\ 0.4059 & 0.2680 & 0.3261 \\ 0.3183 & 0.3419 & 0.3398 \\ 0.4151 & 0.2355 & 0.3494 \end{bmatrix}$$

The satisfaction of the constraints $(A_{eq} - b_{eq})$ is shown in Figure 3.

DOI: 10.24818/18423264/54.2.20.02

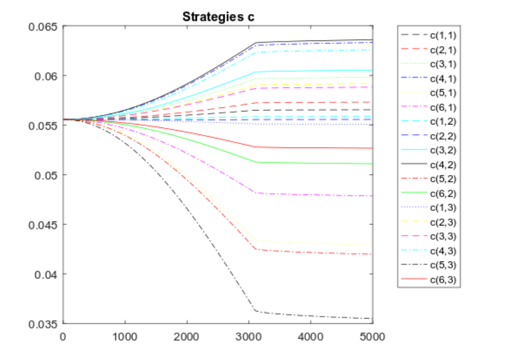

Figure 1: Portfolio - strategies c.

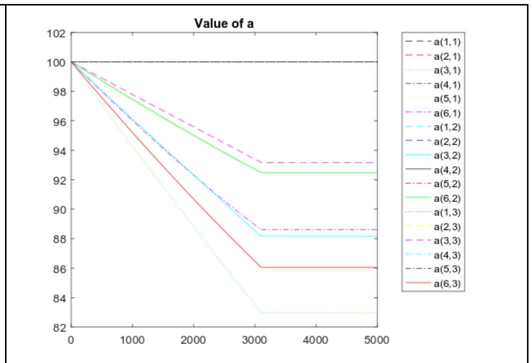

Figure 2: Customers α.

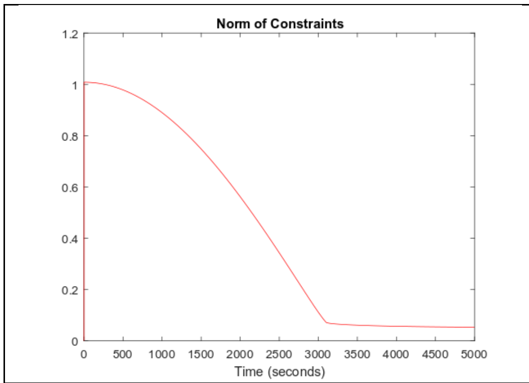

Figure 3: The norm of $(A_{eq} - b_{eq})$ constraints.

6.Conclusion and future work

This paper proposes a customer portfolio optimal selection, restricted to continuous-time Markov chains. The method employed a Lagrange optimization approach for solving the regularization problem that ensured the strong convexity of the objective function and the existence of a unique solution of the portfolio. The regularization approach simplifies the computation of the unique solution of the original portfolio optimization problem. We proved that the solution of the original portfolio converges to a unique solution with the minimal weighted norm. We introduced a recurrent procedure based on the projection-gradient method and proved the convergence of the method. A numerical example has been offered, to validate the effectiveness of our approach.

DOI: 10.24818/18423264/54.2.20.02

REFERENCES

[1] **N. Bauerleand U. Rieder (2004*), Portfolio Optimization with Markov-Modulated Stock Prices and Interest Rates***. *IEEE Transactions on Automatic Control*, 49,442–447;

[2] **J. Brodiea, I. Daubechiesa, C. De Molc, D. Giannoned, and I. Loris (2009),*Sparse and Stable Markowitz Portfolios***. In *Proceedings of the National Academy of Sciences of the United States of America*, 106, 12267–12272;

[3] **M. Carrasco and N. Noumon(2011)**, ***Optimal Portfolio Selection Using Regularization***. Working Paper University of Montreal. Available at http://www.unc.edu/maguilar/metrics/carrasco.pdf;

[4] **L. Carrillo, J. Escobar, J. B. Clempner, and A. S. Poznyak (2016)**,***Solving Optimization Problems in Chemical Reactions Using Continuous-time Markov Chains***. Journal of Mathematical Chemistry, 54,1233–1254;

[5] **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(2),383–417;

[6] **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;

[7] **J. B. Clempner (2016), *Necessary and Sufficient Karush-Kuhn-Tucker Conditions for Multiobjective Markov Chains Optimality***. *Automatica*, 71:135–142, 2016;

[8] **V. DeMiguel, J. Garlappi, J. Nogales, and R. Uppal (2009)**, ***A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms***. Management Science, 55(5):798–812;

[9] **V. DeMiguel, J. Garlappi, and R. Uppal (2009)**, ***Optimal versus Naive Diversification: How Inefficient is the 1/n Portfolio Strategy?***The Review of Financial Studies, 22(5):1915–1953;

[10] **F. Dominguez 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(53):39–56;

[11] **J. Fan, J. Zhang and K. Yu (2012)**, ***Vast Portfolio Selection with Gross Exposure Constraints***. *Journal of the American Statistical Association*, 107(498):592–606;

DOI: 10.24818/18423264/54.2.20.02

[12] **B. Fastrich, S. Paterlini, and P. Winker (2014)**, ***Cardinality versus q-Norm Constraints for Index Tracking***. *Quantitative Finance*, 14:2019–2032;
[13] **M. Garcia-Galicia, 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(2):91–111;
[14] **M. Garcia-Galicia, A. Carsteanu, and J. B. Clempner (2019)**,***Continous-time Reinforcement Learning Approach for Portfolio Management with Time Penalization***. *Expert Systems with Applications*, 129:27–36;
[15] **R. Jagannathan and T. Ma(2003)**, ***Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps***. *Journal of Finance*, 58(4):1651–1683;
[16] **H. Markowitz(1952),*Portfolio Selection*.** *Journal of Finance*, 7:77–91;
[17] **W. Putschögl and J. Sass (2011)**,***Optimal Investment under Dynamic Risk Constraints and Partial Information***. *Quantitative Finance*, 11:1547–1561;
[18] **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;
[19**] 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;
[20**] J. Sass and U. Haussmann (2004)**, ***Optimizing the Terminal Wealth under Partial Information: The Drift Process as a Continuous Time Markov Chain***. *Finance and Stochastics*, 8:553–577;
[21] **M. Taksar and X. Zeng (2007)**,***Optimal Terminal Wealth under Partial Information***. *SIAM Journal on Control and Optimization*, 46:1461–1482;
[22**] Yin and Q. Zhang (1998)**, ***Continuous-Time Markov Chains and Applications: A Singular Perturbations Approach***. *Springer-Verlag,* New York;
[23**] *Markowitz's Mean-variance Portfolio Selection with Regime Switching: From Discrete-time Models to their Continuous-time Limits***. *IEEE Transaction in Automatic Control*, 49(3):349–360;
[24**] C. Zhang, J. Wang, and N. Xiu (2019)**, ***Robust and Sparse Portfolio Model for Index Tracking***. *Journal of Industrial and Management Optimization*, 15(3):1001-1015;
[25] **X. Y. Zhou and G. Yin(2003)**,***Markowitz's Mean-variance Portfolio Selection with Regime Switching: A Continuous-time Model***. *Siam. Journal of Control Optimization*, 42:1466–1482.

DOI: 10.24818/18423264/54.2.20.02