Statistical mechanics in the portfolio optimization with Kusuoka’s representation

R Restrepo\textsuperscript{1}, and Y Herrera\textsuperscript{2}

\textsuperscript{1} Google, M"unich, Germany
\textsuperscript{2} Departamento de Matemáticas y Ciencias Naturales, Universidad Autónoma de Bucaramanga, Bucaramanga, Colombia

E-mail: ricrest@gmail.com, yherrera743@unab.edu.co

Abstract. Portfolio optimization under a risk measure consists of finding the efficient curve of the given measure in the plane of expected return vs risk. The portfolios corresponding to points of this curve are portfolios which minimize the risk under the given measure and whose expected total return is greater than a pre-specified level (benchmark). The optimization problems involved are mostly non-convex and therefore cumbersome. In this paper, we provide a method to approximate the efficient frontier under a law-invariant coherent risk measure by relaxing these optimization problems involved to a finite-dimensional linear model. Our method is based on three tools: the Kusuokas’s representation of the risk measure, the properties of the conditional value at risk functions and the finite elements theory. Finally, we mention the strong connection that the former kind of optimization have with problems in statistical mechanics following the classic framework of simulating annealing.

1. Introduction

Portfolio optimization (or selection) deals with one of the most fundamental aspects in economics and finance. It consists in finding optimal portfolios, that is, portfolios with the smallest risk among those with an expected (total) return surpassing a pre-established level or benchmark. Modern portfolio theory (MPT) was powered by papers such as [1] and [2], where Markowitz provided a mathematical frame of this type of problem by means of his well known mean-variance model (MVM). This model produces a collection of portfolios with the smallest variance among the ones surpassing a pre-specified level of expected return, that is, the efficient frontier under variance. In addition, Markowitz highlighted the benefits of the investment diversification since risk can be reduced by creating a diversified portfolio of unrelated assets (also stated in [3] and [4]). The two of the main deficiencies of the MVM are: (1) variance is not appropriate when the losses distribution is heavy-tailed, and (2) covariance matrix is computationally infeasible if the number of asset categories is large. Another important era of risk measurement in MPT was the one initiated with the implementation of value-at-risk (VaR), this is a risk measurement which basically addresses the question of how large is the capital one can lose, with a given probability $1 - \alpha$, at the end of a given period of time.

In 1994, J P Morgan created this measurement to be used in the whole institution [5], and the following year the Basel Committee on Banking Supervision stipulated a capital requirement based upon it [6]. VaR has been associated with an alternative measure which aims at quantifying
the losses that will be held when they exceed the VaR threshold, the conditional value at risk (CVaR), which defined as the weighted average of the losses strictly exceeding the VaR. Many are the advantages of CVaR optimization compared to the VaRs, the most relevant is that it can be condensed into a simple linear programming formula, which makes it computationally amenable [7]. In addition, this measure is law-invariant and coherent.

In [8], Kusuoka showed that law-coherent risk measures with the Fatou property have a representation of the form given in Equation (1).

$$\rho(X) = \sup_{\gamma \in \mathcal{M}} \int_0^1 \text{CVaR}_\alpha(X) d\gamma(\alpha), \text{ for all } X \in \mathcal{L}_\infty,$$  \hspace{1cm} (1)

where $\mathcal{L}_\infty$ is the $L_\infty$-space of $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{M}$ is a convex set of probability measures on $[0, 1)$.

The method presented here for approximating the efficient frontier under a law-invariant coherent risk measure is based on taking advantage of the linearity of the CVaR functions to obtain linear relaxations of the optimization problems whose solutions are efficient portfolios. Afterwards, we obtain a further relaxation, which is finite-dimensional, this is done with the help of the finite elements method.

Finally, the formal optimization’s method presented here has the potential of allowing a dual version in statistical mechanics via the classic characterization of simulating annealing along the lines of S. Kirkpatrick et al. [9] and following some of the lines of I. Varga-Haszonits [10]. The extension of this classic and remarkable connection with statistical physics represents a highly relevant path for further research on this fields.

2. Methodology

In this section we provide the mathematical concepts and results fundamental for this work. In addition, we state in mathematical terms the problem of interest. Portfolios, risk measures, portfolio optimization, eff and Kusuoka’s representation.

2.1. Mathematical setting

A financial portfolio (or position) refers to a distribution of the wealth to be invested in a set of investment categories, for example, stocks, securities and bonds. Let $\Omega = (\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $L = L_\Omega$ be the set of classes of all finite-valued random variables modulo the equivalence relation of $\mathbb{P}$-a.e. equality. For $X, Y \in L$, we write $X \preceq Y$ to mean that $X(\omega) \leq Y(\omega)$, for $\omega$ $\mathbb{P}$-a.e in $\Omega$. Let us assume without loss of generality that there is a wealth $W = 1$ to be invested in $d$ investment categories with random vector of (total) returns $\xi = (\xi_1, \xi_2, \ldots, \xi_d)$. Then, a portfolio is determined by a decision vector $w = (w_1, w_2, \ldots, w_d)$, where $0 \leq w_i \leq 1$, for $1 \leq i \leq d$, and such that $w \cdot 1 \leq 1$, with $1 = (1, 1, \ldots, 1) \in \mathbb{R}^d$. We assume further that the return of the portfolio associated with $w$ is $w \cdot \mathbb{E}_\mathbb{P}[\xi] = \sum_{i=1}^d w_i \mathbb{E}_\mathbb{P}[\xi_i]$. Let $\mathcal{AP}$ be the subspace of $L$ consisting of all possible returns which is also closed under addition and taking the positive part of the variable, that is, $(X - c)^+ \in \mathcal{AP}$, for all $X \in \mathcal{AP}$ and $c \in \mathbb{R}$, where $z^+ = \max\{z, 0\}$. In this paper, an element $Y$ of $\mathcal{AP}$ will represent both an admissible portfolio and its total return.

Another intrinsic feature of a portfolio is its risk. Quantifying the risk of a financial portfolio is one of the fundamental aspects of risk management and portfolio optimization and has sparked the interest of researchers to create innovative risk measures. In our mathematical setting the notion of risk measure is defined as follows.

Definition 1. A functional $\rho : \mathcal{AP} \to \mathbb{R}$ is a risk measure (or risk function) if it satisfies the following axioms for all $X, Y \in \mathcal{AP}$:
Definition 2. A risk measure $\rho$ is coherent if the following two conditions hold for all $X, Y \in \mathcal{A}$:

(A1). Translation invariance: $\rho(X + r) = \rho(X) - r$, for $r \in \mathbb{R}$.

(A2). Monotonicity: $\rho(X) \leq \rho(Y)$, whenever $X \leq Y$.

(A3). Sub-additivity: $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

(A4). Positive homogeneity: $\rho(\tau X) = \tau \rho(X)$, for $\tau \geq 0$.

Example. The variance $\sigma^2$, defined by $\sigma^2(X) = E[X^2] - E[X]^2 = X \cdot (\Sigma X)$, where $\Sigma$ is the covariance matrix of $\xi$ (see [11]), is not a risk measure since $\sigma^2(X + r) = \sigma^2(X) \neq \sigma^2(X) - r$, for $r \neq 0$, failing Axiom (A1).

Definition 3 (value at risk). For $X \in \mathcal{A}$, let $F_X$ denote its probability distribution with respect to $\mathbb{P}$, that is, $F_X(t) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq t\}$). For $\alpha \in (0, 1)$, we define $VaR_\alpha(X)$, the value at Risk with confidence level $\alpha$, as the $\alpha$-quantile of $X$, i.e., is given by Equation (2).

$$VaR_\alpha(X) = \inf\{u \in \mathbb{R} : \mathbb{P}\{\omega \in \Omega : X(\omega) > u\} \leq 1 - \alpha\} = \inf\{u \in \mathbb{R} : F_X(u) > \alpha\}. \quad (2)$$

The risk measure $VaR_\alpha$, for $\alpha \in (0, 1)$, is not sub-additive as can be seen in [12]. One coherent alternative to $VaR_\alpha$ is the following.

Definition 4 (conditional value at risk). For $X \in \mathcal{A}$ and $\alpha \in (0, 1)$, we define $CVaR_\alpha$, the Conditional (or Average) value at risk with confidence level $\alpha$, by Equation (3).

$$CVaR_\alpha(X) = E_\mathbb{P}[X \mid X > VaR_\alpha(X)] = (1 - \alpha)^{-1} \int_\alpha^1 VaR_\beta d\mathbb{P}(\beta). \quad (3)$$

In was shown in [13] that CVaR is a coherent risk measure and that it can also be defined as the solution of the linear optimization problem given in Equation (4).

$$CVaR_\alpha(X) = \inf_{c \in \mathbb{R}} \{c + (1 - \alpha)^{-1} E_\mathbb{P}[(X - c)^+]\}. \quad (4)$$

The function $CVaR_\alpha$ can be equivalently defined in the following way.

- Lemma. For $Y \in \mathcal{A}$ and $\alpha \in (0, 1)$, $CVaR_\alpha(X) = \inf_{c \in \mathbb{R}, \mathbb{Z} \in \mathcal{A}} \{c + (1 - \alpha)^{-1} E_\mathbb{P}[Y] \mathbb{Z} \geq \max(0, X - c)\}.

2.2. Kusuoka’s representation

It was shown by Kusuoka in [8], for $p = \infty$, and by Pflug and W. Römisch in [14], for $1 \leq p < \infty$, that a risk measure $\rho$ is a law-invariant coherent risk measure with the Fatou property if and only if $\rho$ satisfies Equation (5) for some compact convex set $\mathcal{M}$ of probability measures on $[0, 1]$.

$$\rho(X) = \sup_{\gamma \in \mathcal{M}} \int_0^1 CVaR_\alpha(X)d\gamma(\alpha), \text{ for all } X \in \mathcal{L}_p. \quad (5)$$

The right hand side of Equation (5) is known as the Kusuoka’s representation of $\rho$. 

2.3. Efficient frontier under a risk measure
In the MVM, variance $\sigma^2$ is used for quantifying the risk (see [1]). The goal consisted in finding the portfolio that had both an expected return exceeding a benchmark ($\theta$) and the smallest variance among all of the portfolios with this feature. In mathematical terms the can be written as Equation (6).

$$\begin{array}{c}
\text{minimize } \sigma^2(w \cdot \xi) \\
\text{subject to: } w \in \mathbb{R}^d \\
w \geq 0, \ w \cdot 1 = 1 \\
w \cdot E_p[\xi] \geq \theta_k
\end{array}$$

where $\{\theta_k\}_{k=1}^K$ is a collection of numbers such that $0 < \theta_1 < \theta_2 < \cdots < \theta_K = \xi^\text{max}$, with $\xi^\text{max}$ being the maximum historical return of the collected data. If $Y(\theta_k)$ are the solutions of these problems, the points $(Y(\theta_k), \theta_k)$ can be used to approximate the efficient frontier under variance in the variance vs expected return plane.

Therefore, in order to approximate the efficient frontier under a risk measure, we solve the optimization problems in the MVM, but replacing the variance $\sigma^2$ by the risk measure at issue.

3. Results and discussion
In this work, we focus on risk measures where the supremum in its Kusuoka’s representations is actually a maximum, that is, the risk measures $\rho$ of the form of Equation (7).

$$\rho(X) = \int_0^1 CVaR_\alpha(X) \, d\mu(\alpha), \text{ for all } X \in \mathcal{L}_\infty,$$

for some probability measure $\mu$ on $[0,1)$. Let $\xi = (\xi_1, \xi_2, \ldots, \xi_d)$ be the vector of random returns of asset categories $1,2,\ldots,d$. The (total) return of the associated portfolio with the decision vector $w = (w_1, w_2, \ldots, w_d)$ is $w \cdot \xi$. Therefore, the optimization problems we have to consider are of the form of Equation (8).

$$\begin{array}{c}
\text{minimize } \rho(w \cdot \xi) \\
\text{subject to: } w \in \mathbb{R}^d \\
w \geq 0, \ w \cdot 1 \leq 1 \\
w \cdot E_p[\xi] \geq \theta
\end{array}$$

for some $0 < \theta \leq \xi^\text{max}$, where $\xi^\text{max}$ is the maximum of the recorded returns of the $\xi$’s. In this case, we have the equivalent Equation (9).

$$\begin{array}{c}
\text{minimize } \int_0^1 CVaR_\alpha(w \cdot \xi) \, d\mu(\alpha) \\
\text{subject to: } w \in \mathbb{R}^d \\
w \geq 0, \ w \cdot 1 \leq 1 \\
w \cdot E_p[\xi] \geq \theta
\end{array}$$

The following proposition can be easily proven and it enables us to write a relaxation of Equation (9).
Proposition 1. For $\mu$ as above let $\eta(\mu) := \int_0^1 (1-\alpha)^{-1} d\mu(\alpha)$ and assume that $\eta(\mu) < \infty$. Then, inequality holds (Equation (10)).

$$\rho(w \cdot \xi) \leq \inf_{c \in \mathbb{R}, w \in \mathbb{R}^d, Z \in \mathcal{AP}} \{ c + \eta(\mu)E_P[Z] \}. \quad (10)$$

Proposition 1 allows us to get the Equation (11) whose solution is an approximation to the solution of Equation (9), but with the enormous advantage of being a linear problem.

$$\text{minimize} \{ c + \eta(\mu)E_P[Z] \}$$

subject to: $w \in \mathbb{R}^d$, $c \in \mathbb{R}$, $Z \in \mathcal{AP}$

$$w \geq 0, \ w \cdot 1 \leq 1$$

$$w \cdot E_P[Z] \geq \theta$$

$$Z \geq \max\{0, w \cdot \xi - c\}. \quad (11)$$

It is worth noticing that this problem is infinite-dimensional.

3.1. Finite dimensional relaxation

Now we consider one further relaxation of Equation (11) obtained with the finite element method. For simplicity, we assume that $d = 2$, so we consider only two asset categories with vector of random returns $\xi = (\xi_1, \xi_2)$ in a time interval $I$ and our decision vectors are of the form $w = (w_1, w_2)$, with $w_1, w_2 \geq 0$ and $w_1 + w_2 \leq 1$.

Let $K, L \in \mathbb{N}$, and let us consider two partitions of $\mathbb{R}$, say $I = \{I_1, \ldots, I_K\}$ and $J = \{J_1, \ldots, J_L\}$. Let $Q_{k,l} = \xi^{-1}(I_k \times J_l)$ for $k = 1, 2, \ldots, K$ and $l = 1, 2, \ldots, L$. Denote by $\chi_{k,l}$ the indicator function of the set $Q_{k,l}$, that is, $\chi_{k,l}(x, y) = 1$ when $(x, y) \in Q_{k,l}$ and $\chi_{k,l}(x, y) = 0$, otherwise.

Next, let $\mathcal{H} = \mathcal{H}_{I,J}$ be the positive cone subspace of random variables generated by the indicator functions $\chi_{k,l}$'s, i.e., $\mathcal{H} = \{ \sum_{k=1}^K \sum_{l=1}^L t_{k,l} \chi_{k,l} \mid t_{k,l} \geq 0 \}$. Now, in order to approximate the efficient frontier under the risk measure $\rho$, we restrict $Z$ to the subspace $\mathcal{H}$ to get the finite-dimensional linear optimization Equation (12).

$$\text{minimize} \{ c + \eta(\mu)E_P[Z] \}$$

subject to: $w \in \mathbb{R}^d$, $c \in \mathbb{R}$, $Z \in \mathcal{H}$

$$w \geq 0, \ w \cdot 1 \leq 1$$

$$w \cdot E_P[\xi] \geq \theta$$

$$Z \geq w \cdot \xi - c. \quad (12)$$

Due to the fact that $\mathcal{H} \subset \mathcal{AP}$, Equation (12) is indeed a relaxation of Model (11), but with the tremendous advantage of being a finite-dimensional linear programming problem. Note that $Z = \sum_{j=1}^n t_{k,l} \chi_{k,l} \in \mathcal{H}$ satisfies Equation (13).

$$E_P[Z] = \sum_{k=1}^K \sum_{l=1}^L t_{k,l} E_P[Q_{k,l}] = \sum_{k=1}^K \sum_{l=1}^L t_{k,l} P\{ \xi \in I_k \times J_l \}. \quad (13)$$

In addition, if $n_{k,l} = \#\{(\xi_1, \xi_2) \in I_k \times J_l\}$, we approximate $P[Q_{k,l}]$ with $n_{k,l}/N$, where $N = \#\{\xi \in I_k \times J_l\}$. Hence, we get our final model, Equation (14).
\[
\begin{align*}
\text{minimize} & \quad \left\{ c + \eta(\mu) \frac{1}{N} \sum_{k=1}^{K} \sum_{l=1}^{L} t_{k,l} n_{k,l} \right\} \\
\text{subject to:} & \quad w \in \mathbb{R}^2, \ w \geq 0, \ c \in \mathbb{R}, \ t_{k,l} \geq 0 \\
& \quad w_1 + w_2 \leq 1, \ w_1 E_P[\xi_1] + w_2 E_P[\xi_2] \geq \theta \\
& \quad t_{k,l} \geq w_1 \xi_1 + w_2 \xi_2 - c, \text{ when } (\xi_1, \xi_2) \in I_k \times J_l
\end{align*}
\] (14)

It is worth noticing that the fact that the Equation (14) is a relaxation of Equation (8), therefore, its solution is a suitable approximation of the solution of latter.

4. Numerical example
In order to validate our proposed method we approximate the efficient frontier under a law-invariant coherent risk measure with the Fatou property.

To this end, we consider the daily returns of Google and Microsoft shares, which we refer to as Asset 1 and Asset 2, respectively. Let \( \xi_i \) be the daily total return of Asset \( i \), for \( i = 1, 2 \), for the time interval corresponding to the dates from the January 1, 2017 to the April 11, 2017. We find an approximate efficient frontier under the risk measure \( \rho \) defined by Equation (15).

\[
\rho(X) = \int_0^1 CVaR_\alpha(X) d\mu(\alpha),
\] (15)

where \( \mu \) is the uniform probability measure on the interval \([0.9, 0.95]\).

Now, we proceed to make the computations needed for the relaxantion model described in Section 3.1.

- The total number of records are \( N = 69 \).
- Consider one partition of \( \mathbb{R} \) with \( m = 130 \) intervals \( I_k \)'s and another partition with \( n = 120 \) intervals \( J_l \)'s. Hence, \( P[Q_{k,l}] = P\{ (\xi_1, \xi_2) \in I_k \times J_l \} \), which we approximate with \( n_{k,l}/69 \).
- The value of \( \eta(\mu) \), is \( \eta(\mu) = \int_0^1 (1 - \alpha)^{-1} d\mu(\alpha) = \frac{1}{0.95} \ln(\alpha)|_{0.9}^{0.95} \approx 1.08 \).
- The means of the assets are \( E_P[\xi_1] \approx 0.0007 \), and \( E_P[\xi_2] \approx 0.008 \), and the maximum recorded return is \( \xi_2^{\text{max}} = 0.0235 \), then we consider \( \theta < 0.0235 \).

Thus, in order to approximate the efficient frontier of the risk measure \( \rho \) we apply the proposed method to get the two-dimensional linear Equation (16).

\[
\begin{align*}
\text{minimize} & \quad \left\{ c + \frac{1.08}{69} \sum_{k=1}^{K} \sum_{l=1}^{L} n_{k,l} \right\} \\
\text{subject to:} & \quad w \in \mathbb{R}^2, \ w \geq 0, \ c \in \mathbb{R}, \ t_{k,l} \geq 0 \\
& \quad w_1 + w_2 \leq 1, \ 0.007w_1 + 0.008w_2 \geq \theta \\
& \quad t_{k,l} \geq w_1 \xi_1 + w_2 \xi_2 - c, \text{ when } (\xi_1, \xi_2) \in I_k \times J_l
\end{align*}
\] (16)

To solve Equation (16) we use the Python’s scipy module (https://www.scipy.org). Table 1 shows the weights of the efficient portfolios, with their respective expected return and \( \rho \)-risk.

As a way of illustration, Figure 1 shows the approximate efficient frontier obtained with 30 benchmarks.
Table 1. Some optimal portfolios.

| $w_1$  | $w_2$  | $\approx \rho$ | ER         |
|--------|--------|-----------------|------------|
| 4.69627494e-02 | 5.78332907e-02 | 1.54955357e-09 | 7.83155473e-05 |
| 9.39254988e-02 | 1.15666581e-01 | 3.09910713e-09 | 1.56631095e-04 |
| 1.4088248e-01 | 1.73499872e-01 | 4.64866069e-09 | 2.34946642e-04 |
| 1.87850998e-01 | 2.31333163e-01 | 6.19821426e-09 | 3.13262189e-04 |
| 2.34813747e-01 | 2.89166453e-01 | 7.74776783e-09 | 3.91577737e-04 |

Figure 1. Efficient frontier with 30 benchmarks.

5. Conclusions
The method here provided for approximating the efficient frontier under a law-invariant coherent risk measure constitutes an effective mechanism due to the fact that it is agnostic in the sense that it can be used regardless of the distributions of the returns of the stocks.

In addition, this method represents a convenient computational tool to approximate efficient frontiers since it transforms a infinite-dimensional non-linear problem into a finite-dimensional linear problem, which is a considerable simpler task.

Lastly, the former optimization method has a huge potential for being extended in the context of statistical mechanics via the traditional framing of simulating annealing due to Kirkpatrick and colleagues and further developed by Varga-Haszonits et al. Moreover, as pointed before our approach has certain computational soundness that can serve as a desired feature for enriching the former connection via modern algorithmic tool coming from computational physics.

Acknowledgments
The second author thanks the Universidad Autónoma de Bucaramanga, Bucaramanga, Colombia, for their grant I56074.

References
[1] Markowitz H 1952 Portfolio selection The Journal of Finance 7(1) 77–91
[2] Markowitz H 1959 Portfolio Selection: Efficient Diversification of Investments (New York: Wiley)
[3] Rubinstein M 2002 Markowitz’s "portfolio selection": A fifty-year retrospective Journal of Finance 57(3) 1041–1045
[4] Watkins G P 1922 Knight’s risk, uncertainty and profit The Quarterly Journal of Economics 36(4) 682–690
[5] Dowd K 2013 Measuring Market Risk (Hoboken: Wiley-Blackwell)
[6] Danielsson J, Shin H S and Zigrand J P 2004 The impact of risk regulation on price dynamics Journal of Banking & Finance 28(5) 1069–1087
[7] Rockafellar R T and Uryasev S 2000 Optimization of conditional value-at-risk Journal of Risk 2 21–41
[8] Kusuoka S 2001 On law invariant coherent risk measures Advances in Mathematical Economics (Advances in Mathematical Economics vol 3) (Tokyo: Springer) pp 83–95
[9] Kirkpatrick S, Gelatt C D and Vecchi M P 1983 Optimization by simulated annealing Science 220(4598) 671–680
[10] Varga-Haszonits I, Caccioli F and Kondor I 2016 Replica approach to mean-variance portfolio optimization Journal of Statistical Mechanics: Theory and Experiment 2016(12) 123404:1–24
[11] Chin L, Chendra E and Sukmana A 2018 Analysis of portfolio optimization with lot of stocks amount constraint: case study index LQ45 IOP Conference Series: Materials Science and Engineering 300 012004
[12] Danielsson J, Jorgensen B N, Samorodnitsky G, Sarma M and de Vries C G 2013 Fat tails, var and subadditivity Journal of Econometrics 172(2) 283–291
[13] Pflug G C 2000 Some remarks on the value-at-risk and the conditional value-at-risk Probabilistic constrained optimization (Nonconvex Optimization and Its Applications vol 49) (Netherlands: Kluwer Academic Publishers) pp 272–281
[14] Pflug G C and Römisch W 2007 Modeling, Measuring and Managing Risk (Hackensack: World Scientific Publishing Co. Pte. Ltd.)