



Portfolio construction with Gaussian mixture returns and exponential utility via convex optimization

Eric Luxenberg¹ · Stephen Boyd¹

Received: 26 May 2023 / Revised: 27 June 2023 / Accepted: 28 June 2023 /

Published online: 27 July 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

We consider the problem of choosing an optimal portfolio, assuming the asset returns have a Gaussian mixture distribution, with the objective of maximizing expected exponential utility. In this paper we show that this problem is convex, and readily solved exactly using domain-specific languages for convex optimization, without the need for sampling or scenarios. We then show how the closely related problem of minimizing entropic value at risk can also be formulated as a convex optimization problem.

Keywords Convex optimization · portfolio construction · Gaussian mixture

1 Introduction

1.1 Asset return distributions

There is a long history of researchers observing that the tails of asset returns are not well modeled by a Gaussian distribution, going back to the thesis of Fama (1965), who observed that while somewhat symmetric, the tails of the return distribution were much heavier than those of a Gaussian distribution. Additionally, asset returns are skewed, violating normality (Neuberger 2012). There is also long history of researchers proposing alternative distributions to model asset returns, including elliptical distributions (Bingham and Kiesel 2001), and Gaussian mixtures (GMs) (Ball and Torous 1983; Akgiray and Booth 1987), the focus of this paper.

GMs can in principle approximate any continuous distribution; for asset returns, it has been observed that good approximations can be obtained with just a handful of mixture components (Kon 1984). We can interpret the components of a GM return distribution as market regimes, with a latent variable that represents the active regime, and

✉ Eric Luxenberg
ericlux@stanford.edu

¹ Department of Electrical Engineering, Stanford University, Stanford, USA

a return distribution that is Gaussian, given the regime. Many authors have observed that the correlations among asset returns can change during different market regimes, for example, increased correlations during bear markets (Campbell et al. 2002; Ang and Bekaert 2004, 2015). A GM can model such regime-dependent correlation structures. Another desirable attribute of a GM is that it can model skewness in return distributions, for which many authors have argued real investors exhibit preferences (Arditti 1967; Scott and Horvath 1980).

GM return models arise in a hidden Markov Gaussian model (Rydén et al. 1998; Nystrup et al. 2019) of returns, which models the regimes as Markovian, and the returns as Gaussian, given the regime. In such a model the means and covariances of the Gaussian components corresponding to the regimes are fixed, but the component weights change in each period (Gupta and Dhingra 2012); but in each period, the asset return distribution, conditioned on the past returns, is GM, so the methods in this paper can be applied.

1.2 Mean-variance versus expected utility

Mean-variance portfolio construction In mean-variance portfolio construction, pioneered by Markowitz (1952), portfolio construction is viewed as an optimization problem with two main objectives: the mean or expected return of the portfolio, and the risk, taken to be the variance of the portfolio return. These objectives are combined into a risk-adjusted return using a positive weight parameter, interpreted as setting the level of risk aversion. Mean-variance portfolio construction can be carried out analytically, when there are very simple constraints, or numerically, with realistic constraints, by solving a convex optimization problem such as a quadratic program (Grinold and Kahn 1999; Boyd Vandenberghe 2004; Boyd et al. 2017; Stellato et al. 2020). With current convex optimization methods, mean-variance construction can be done reliably and quickly for up to thousands of assets, and many more when a factor risk model is used. These optimization problems can be solved in well under one second, allowing back-tests (what-if simulations, based on real or simulated data) to be carried out quickly (Boyd et al. 2017; Schaller et al. 2022; MathWorks). In addition, the mean-variance framework is easily adapted to active portfolio management, where exceptional return forecasts are incorporated with the goal of outperforming a benchmark. It is less natural to specify the active portfolio management problem in terms of expected utility.

One obvious criticism of mean-variance portfolio construction is that the (quadratic) objective function penalizes returns that are well above the mean (a desirable event) just as much as returns that are well below the mean (an undesirable event) (Hanoch and Levy 1970). Another is that it only uses the first two moments of the return distribution, and so cannot take into account skewed or fat-tailed distributions. Nevertheless it is very widely used in practice. There is work in analyzing portfolios under higher order Taylor approximations of utilities (Jondeau and Rockinger 2006), but these are not used in practice due to both semantic and computational complexity.

Expected utility portfolio construction In work that predates mean-variance portfolio construction, von Neumann et al. (1944) introduced the notion of utility to model

decision making with uncertain outcomes. A utility function specifies a value indexing an agent's preference for each specific outcome; their theory posits that the agent makes a choice so as to maximize her expected utility.

Portfolio construction by expected utility maximization also frames the problem as an optimization problem. The trader specifies a utility function that is concave and increasing, and the objective (to be maximized) is the expected utility under the return distribution. This formulation avoids the awkward situation in mean-variance portfolio construction where high portfolio returns are considered bad. Expected utility maximization better captures the asymmetry in downside versus upside risks than mean-variance optimization. Since the return distribution is arbitrary, expected utility can directly handle return distributions with skew or fat tails.

Expected utility maximization, like mean-variance optimization, leads to a convex optimization problem, more specifically, a stochastic optimization problem (Shapiro et al. 2021). Almost all expected utility methods for constructing portfolios work with samples of the asset returns. This can be considered an advantage, since it means that such methods can work with any return distribution from which we can sample returns. The disadvantage is that sample-based optimization, while tractable, can be slow compared to mean-variance methods, and scales poorly with problem size.

There are several related portfolio construction methods that rely on return samples and stochastic convex optimization. One is based on conditional value at risk (CVaR) (Rockafellar and Uryasev 2000, 2002). A more recently proposed method uses entropic value at risk (EVaR) (Ahmadi-Javid 2012; Cajas 2021), which we address in Sect. 4. Both of these are coherent measures of risk (Artzner et al. 1999; Rockafellar 2007), and result in convex stochastic optimization problems.

Sample based stochastic optimization methods are used in practice (Grinold 1999), but far less often than methods based on mean-variance optimization, which do not involve samples. This is partly because solving sample based stochastic convex optimization problems is tractable, but far more involved than solving the convex optimization problems that do not involve return samples, e.g., mean-variance optimization or the methods proposed in this paper.

Comparison These two main approaches, mean-variance optimization and expected utility maximization, are not as different as they might seem. Levy and Markowitz (1979) show that maximizing a second order Taylor approximation of a utility function is equivalent to mean-variance optimization. So very roughly speaking, mean-variance optimization is the second order approximation of expected utility optimization.

When the returns are Gaussian, and we use an exponential utility, mean-variance and expected utility optimization are not merely close, but exactly the same. It appears that Merton was the first to note this connection (Merton 1969), but his observation does not seem to be mentioned often after that (see also Sect. 3.3).

We remark that the exponential utility has the property of constant absolute risk aversion (CARA). Under constant absolute risk aversion, allocations to risky assets are independent of starting wealth level. As a result, financial economists tend to view the exponential utility as less realistic than utilities with constant relative risk aversion. However, the Nobel Prize winning work of Kahneman and Tversky (1979) shows that utility maximization in general is incompatible with human behavior. So we should not be too concerned with the properties of any one utility function, and choose the

exponential utility because of its desirable mathematical properties when applied to Gaussian returns.

1.3 This paper

In this paper we consider a GM model for asset returns, and maximize expected utility with a generic utility function, the exponential utility (Saha 1993). We refer to this type of portfolio construction as EGM, for exponential utility with Gaussian mixture returns. We show that the EGM portfolio construction problem can be solved exactly as a convex optimization problem, without the need for any samples from the distribution or other approximations. The EGM portfolio construction problem is not only convex, but is easily specified in just a few lines of code in domain-specific languages (DSLs) for convex optimization such as CVXPY (Diamond and Boyd 2016), CVX (Grant and Boyd 2014), or CVXR (Fu et al. 2020). Thus EGM combines the efficiency, reliability, and scalability of mean-variance optimization with the ability of expected utility maximization to handle non-Gaussian returns and the asymmetry in our preferences. When the GM has only one component, our return model is Gaussian, and EGM reduces to mean-variance optimization. Thus we can think of EGM as an extension of mean-variance optimization, or as a special case of expected utility where the problem can be solved exactly, without any return samples.

We also show that EGM is closely related to portfolio construction methods based on the entropic value at risk (EVaR). With GM return model, we show that EVaR portfolio construction problem leads to a convex optimization problem that, like EGM, does not involve sample based stochastic optimization.

1.4 Previous and related work

Portfolio construction with Gaussian mixture returns In Buckley et al. (2008), consider a two-component Gaussian mixture of tranquil and distressed regimes, and analyze several objectives, including Markowitz, Sharpe ratio, exponential utility, and lower partial moments. In Sect. 3.3.6 of their paper they derive the closed form expression for expected exponential utility under Gaussian mixture returns, but do not observe that maximizing exponential utility leads to a convex problem. Studying single period portfolios consisting of a risk free asset and a risky asset, Prigent and Kaffel analyze optimal portfolios under arbitrary utility functions, and show on historical data that GM return models lead to significantly different portfolios than those arising from a Gaussian return model (Hentati-Kaffel and Prigent 2014).

EVaR portfolio construction In recent work Cajas (2021), Cajas develops a disciplined convex (DCP) formulation of EVaR, with return samples, which allows it to be used as either the objective or as a constraint in portfolio optimization problems specified using DSLs for convex optimization such as CVXPY. Since a return distribution that takes on a finite set of values (e.g., the empirical distribution of samples) is a special case of GM, we can consider EGM (with EVaR) as a generalization of Cajas' formulation.

1.5 Outline

We describe the GM return model in Sect. 2, and in Sect. 3 we show that portfolio optimization with exponential utility is a convex optimization problem. In Sect. 4 we show that the closely related function EVaR is also convex, so minimizing it, or adding a limit on it as a constraint, results in a convex optimization problem.

2 Gaussian mixture return model

2.1 Asset return distribution

We let $r \in \mathbf{R}^n$ denote the return of n assets over some specific period. We model r as having a GM distribution with k components,

$$r \sim \mathbf{GM}(\{\mu_i, \Sigma_i, \pi_i\}_{i=1}^k),$$

where $\pi_i \in \mathbf{R}$ are the (positive) component probabilities, $\mu_i \in \mathbf{R}^n$ are the component means, and $\Sigma_i \in \mathbf{R}^{n \times n}$ are the (symmetric positive definite) component covariance matrices.

The GM return distribution includes two interesting special cases. When there is only one component, it reduces to Gaussian, with $r \sim \mathcal{N}(\mu_1, \Sigma_1)$. Another special case arises when $\Sigma_1 = \dots = \Sigma_k = 0$. Here r takes on only the values μ_1, \dots, μ_k , with probabilities π_1, \dots, π_k . We refer to this as a finite values return distribution.

2.2 Portfolio return distribution

Let $w \in \mathbf{R}^n$ denote the weights in an investment portfolio, with $\mathbf{1}^T w = 1$, where $\mathbf{1}$ is the vector with all entries one. For $w_i \geq 0$, w_i is the fraction of the total portfolio value invested in asset i ; for $w_i < 0$, $-w_i$ is the fraction of total portfolio value that is held in a short position in asset i . The portfolio return is $R = w^T r$. This scalar random variable is also GM with component probabilities π_i , and means and variances

$$v_i = w^T \mu_i, \quad \sigma_i^2 = w^T \Sigma_i w, \quad i = 1, \dots, k.$$

We observe that various quantities associated with the portfolio return R can be evaluated analytically, without the need for Monte Carlo or other sampling methods. For example its cumulative distribution function (CDF) is given by

$$\Phi_R(w, a) = \sum_{i=1}^k \pi_i \Phi\left(\frac{a - v_i}{\sigma_i}\right) = \sum_{i=1}^k \pi_i \Phi\left(\frac{a - w^T \mu_i}{(w^T \Sigma_i w)^{1/2}}\right), \tag{1}$$

where Φ is the CDF of a standard Gaussian.

2.3 Moment and cumulant generating functions

Two other quantities we will encounter later are the moment generating function

$$M(w, t) = \mathbf{E} \exp(tR) = \sum_{i=1}^k \pi_i \exp\left(tv_i + \frac{t^2}{2} \sigma_i^2\right) = \sum_{i=1}^k \pi_i \exp\left(t\mu_i^T w + \frac{t^2}{2} w^T \Sigma_i w\right), \quad (2)$$

where we use $\mathbf{E} \exp Z = \exp(\mu + \sigma^2/2)$ for $Z \sim \mathcal{N}(\mu, \sigma^2)$, and the cumulant generating function

$$\begin{aligned} K(w, t) &= \log \mathbf{E} \exp(tR) \\ &= \log \left(\sum_{i=1}^k \pi_i \exp\left(tv_i + \frac{t^2}{2} \sigma_i^2\right) \right) \\ &= \log \left(\sum_{i=1}^k \pi_i \exp\left(t\mu_i^T w + \frac{t^2}{2} w^T \Sigma_i w\right) \right). \end{aligned} \quad (3)$$

We observe for future use the identity

$$K(w, t) = K(tw, 1), \quad (4)$$

i.e., the parameter t simply multiplies the argument w .

3 Portfolio optimization with exponential utility

3.1 Expected exponential utility

Our objective is to choose w to maximize the expected exponential utility $\mathbf{E} U_\gamma(R)$, where

$$U_\gamma(a) = 1 - \exp(-\gamma a),$$

with $\gamma > 0$ the risk aversion parameter. Using (2), we can express this as

$$\mathbf{E} U_\gamma(R) = 1 - \mathbf{E} \exp(-\gamma w^T r) = 1 - M(w, -\gamma).$$

It follows that we can maximize $\mathbf{E} U_\gamma(R)$ by minimizing the moment generating function $M(w, -\gamma)$, or equivalently the cumulant generating function

$$K(w, -\gamma) = \log \left(\sum_{i=1}^k \exp\left(\log \pi_i - \gamma \mu_i^T w + \frac{\gamma^2}{2} w^T \Sigma_i w\right) \right). \quad (5)$$

Convexity The function $K(w, -\gamma)$ is a convex function of w . To see this, we note that for each i , $\log \pi_i - \gamma \mu_i^T w + \frac{\gamma^2}{2} w^T \Sigma_i w$ is a convex quadratic function of w , and therefore convex. The function $K(w, -\gamma)$ is the log-sum-exp function (also called the soft-max function),

$$S(u) = \log \left(\sum_{i=1}^k \exp u_i \right), \quad (6)$$

of these arguments. The log-sum-exp function is convex and increasing in all arguments, so the composition f_γ is convex (Boyd Vandenberghe 2004, §3.1.5).

3.2 EGM portfolio construction

Our portfolio construction optimization problem has the form

$$\begin{aligned} & \text{minimize } K(w, -\gamma) \\ & \text{subject to } \mathbf{1}^T w = 1, \quad w \in \Omega, \end{aligned} \quad (7)$$

where \mathcal{W} is a convex set of portfolio constraints. This is evidently a convex optimization problem. One implication is that we can efficiently solve this problem globally using a variety of methods.

DSL specification The EGM problem (7) is not just convex. It is readily specified in domain-specific languages (DSLs) for convex optimization, since all such systems include the log-sum-exp function, and all such systems can handle the convex function composition rules that we used to establish convexity of $K(w, -\gamma)$ in w . No special methods (or gradient or other derivatives) are needed; the function $K(w, -\gamma)$ can be specified in a DSL by just typing it in as is. As a simple example, CVXPY code for specifying the EGM construction problem (7) and solving it, with a long only portfolio (i.e., $w \geq 0$), is given below. (This code snippet is also available at the repository https://github.com/cvxgrp/exp_util_gm_portfolio_opt.)

```

1  import cvxpy as cvx
2
3  def K(w):
4      u = cvx.vstack([cvx.log(pi[i])
5                     - gamma * mus[i] @ w
6                     + (gamma**2/2) * cvx.quad_form(w, Sigmas[i])
7                     for i in range(len(pi))])
8      return cvx.log_sum_exp(u)
9
10 w = cvx.Variable(n)
11 objective = cvx.Minimize(K(w))
12 constraints = [ w >= 0, cvx.sum(w) == 1 ]
13 egm_prob = cvx.Problem(objective, constraints)
14 egm_prob.solve()
```

15 w.value

Here it is assumed that SPSVERBc1, SPSVERBc2, and SPSVERBc3 are constants corresponding to n , π and γ , and SPSVERBc4 and SPSVERBc5 are lists of the μ_i and Σ_i , respectively. In lines 3–7 the objective $K(w, -\gamma)$ is formed, and in lines 9–12 the EGM optimization problem is formed. The problem is solved in line 13, which populates SPSVERBc6 with optimal weights. In this simple example our portfolio constraint set \mathcal{W} is simple. One of the advantages of using a DSL is that more complex constraints can be added by just appending them to the list of constraints defined in line 11, without requiring modification of the solving algorithm. DSL's allow practitioners to write code which can be developed quickly and is highly extensible. Additionally, the problem being convex means a high precision solution can be obtained quickly and with high reliability.

Soft-max interpretation We can give an interpretation of the objective $K(w, -\gamma)$ in (5) in terms of the soft-max function, which can be thought of as a smooth approximation to the max, since it satisfies

$$\max_i u_i \leq S(u) \leq \max_i u_i + \log k. \tag{8}$$

The objective (5) can be expressed as

$$K(w, -\gamma) = S(u), \quad u_i = \log \pi_i + \gamma \left(-\mu_i^T w + \frac{\gamma}{2} w^T \Sigma_i w \right), \quad i = 1, \dots, k.$$

We recognize $-\mu_i^T w + \frac{\gamma}{2} w^T \Sigma_i w$ as the negative risk-adjusted return of the portfolio under the i th Gaussian component. Thus $K(w, -\gamma)$ is the soft-max of these negative risk-adjusted returns, offset by the terms $\log \pi_i$, and scaled by γ . Roughly speaking, our objective is an approximation of the maximum of the negative risk-adjusted returns under the component distributions.

From (8) we have

$$K(w, -\gamma) \geq \max_{i=1, \dots, k} \left(\log \pi_i - \gamma \mu_i^T w + \frac{\gamma^2}{2} w^T \Sigma_i w \right), \tag{9}$$

$$K(w, -\gamma) \leq \log k + \max_{i=1, \dots, k} \left(\log \pi_i - \gamma \mu_i^T w + \frac{\gamma^2}{2} w^T \Sigma_i w \right). \tag{10}$$

3.3 Special cases

Gaussian returns When $k = 1$ our GM return distribution reduces to Gaussian, and the problem (7) reduces to the standard Markowitz problem Markowitz (1952, 1959)

$$\begin{aligned} &\text{maximize } \mu_1^T w - \frac{\gamma}{2} w^T \Sigma_1 w \\ &\text{subject to } \mathbf{1}^T w = 1, \quad w \in \Omega. \end{aligned}$$

Finite values returns We can model a finitely supported return distribution by setting $\Sigma_i = 0$, in which case r takes on only the values μ_1, \dots, μ_k with probabilities

π_1, \dots, π_k . This is because each Gaussian in the mixture is degenerate and is supported on its mean. Taking $\pi_i = \frac{1}{k}$ captures the sample based approach. Then, the problem (7) can be expressed as

$$\begin{aligned} &\text{minimize } \log\left(\sum_{i=1}^k \pi_i \exp(-\gamma \mu_i^T w)\right) \\ &\text{subject to } \mathbf{1}^T w = 1, \quad w \in \Omega. \end{aligned} \tag{11}$$

3.4 Simple example

To illustrate the difference between EGM and mean-variance portfolios, we construct a very simple example for which the two portfolios can be analytically found. We take a finite values distribution with $n = 2$ assets and $k = 2$ components, with $\Sigma_1 = \Sigma_2 = 0$,

$$\mu_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \mu_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Thus $r = (-1, 0)$ with probability π_1 and $r = (1, 0)$ with probability π_2 . The first asset is risky, and the second is riskless, with zero return. We take $\Omega = \mathbf{R}^2$, so the only constraint on the portfolio weight is $w_1 + w_2 = 1$.

Markowitz portfolio The mean and covariance of r are

$$\mu = \begin{bmatrix} 1 - 2\pi_1 \\ 0 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 4\pi_1(1 - \pi_1) & 0 \\ 0 & 0 \end{bmatrix}.$$

The Markowitz optimal portfolio is

$$w_1^M = \frac{1 - 2\pi_1}{4\gamma\pi_1(1 - \pi_1)},$$

with $w_2^M = 1 - w_1^M$.

EGM portfolio The EGM portfolio minimizes

$$\pi_1 \exp(\gamma w_1) + (1 - \pi_1) \exp(-\gamma w_1),$$

so the EGM portfolio is

$$w_1^E = \frac{\log(1/\pi_1 - 1)}{2\gamma},$$

with $w_2^E = 1 - w_1^E$.

Comparison The two portfolios are the same for $\pi_1 = 1/2$, with $w^M = w^E = (0, 1)$. They are not too far from each other for other values of π_1 and γ , but can differ substantially for others. For example with $\pi_1 = 0.05$ and $\gamma = 1$, the Markowitz and EGM portfolios are

$$w^M = (4.74, -3.74), \quad w^E = (1.47, -.47).$$

The value at risk 5% is 4.74 for the Markowitz portfolio compared to 1.47 for the EGM portfolio.

3.5 High and low risk aversion limits

High risk aversion limit Here we consider the case where $\gamma \rightarrow \infty$. Dividing (9) and (10) by γ , we find that

$$\frac{K(w, -\gamma)}{\gamma} = \max_{i=1, \dots, k} \left(-\mu_i^T w + \frac{\gamma}{2} w^T \Sigma_i w \right) + O(1/\gamma).$$

So for large risk aversion parameter γ , the EGM portfolio construction problem (7) is approximately

$$\begin{aligned} &\text{minimize } \max_{i=1, \dots, k} \left(-\mu_i^T w + \frac{\gamma}{2} w^T \Sigma_i w \right) \\ &\text{subject to } \mathbf{1}^T w = 1, \quad w \in \Omega. \end{aligned}$$

Thus in the limit of high risk aversion, the EGM portfolio minimizes the maximum of the risk adjusted returns under each of the components, regardless of the π_i . This is similar to solving a minimax Markowitz problem, where we use as a risk model the maximum risk over a set of covariance matrices (see Boyd et al. 2017, §4.2, p. 30).

Low risk aversion limit Here we consider the case where $\gamma \rightarrow 0$. We start with the well known expansion

$$\frac{1}{\gamma} \log \left(\sum_{i=1}^k \pi_i \exp \gamma z_i \right) = \sum_{i=1}^k \pi_i z_i + \frac{\gamma}{2} \left(\sum_{i=1}^k \pi_i z_i^2 - \left(\sum_{i=1}^k \pi_i z_i \right)^2 \right) + O(\gamma^2),$$

for any z_i . (We recognize the first term on the righthand side as the mean of z , and the second as $\gamma/2$ times the variance of z , when z is a random variable taking values z_1, \dots, z_k with probabilities π_1, \dots, π_k .) Substituting $z_i = -\mu_i^T w + \frac{\gamma}{2} w^T \Sigma_i w$ we obtain

$$\frac{K(w, -\gamma)}{\gamma} = -\mu^T w + \frac{\gamma}{2} w^T \Sigma w + O(\gamma^2),$$

where μ and Σ are the mean and covariance of r ,

$$\begin{aligned} \mu &= \mathbf{E} r = \sum_{i=1}^k \pi_i \mu_i, & \Sigma &= \mathbf{E} r r^T - (\mathbf{E} r)(\mathbf{E} r)^T \\ & & &= \sum_{i=1}^k \pi_i \left(\Sigma_i + (\mu_i - \mu)(\mu_i - \mu)^T \right). \end{aligned}$$

This has a very nice interpretation: in the limit of small risk aversion, EGM reduces to Markowitz, using the mean and covariance of the return.

4 Portfolio optimization with entropic value at risk

4.1 Entropic value at risk

The traditional measure of downside risk is the value at risk (VaR) with probability α , which is the $(1 - \alpha)$ quantile of the negative return $-R$,

$$\text{VaR}_\alpha(R) = -\inf\{x \in \mathbf{R} \mid \mathbf{Prob}(R \leq x) > \alpha\}.$$

(We are typically interested in values such as $\alpha = 0.05$ or $\alpha = 0.01$.) For example if the value at risk of a portfolio with probability 5% is 15%, the probability of a loss exceeding 15% (*i.e.*, $R \leq -0.15$) is 5%. Value at risk is interpretable and widely used, but it is not a coherent risk measure (Rockafellar and Uryasev 2000). For example, VaR is not sub-additive, so the sum of two portfolios can have a higher VaR than the sum of the component VaRs. Several coherent risk measures have been proposed, including the conditional value at risk CVaR $_\alpha$ (Rockafellar and Uryasev 2000) and entropic value at risk EVaR $_\alpha$ (Ahmadi-Javid 2012).

The entropic value at risk EVaR $_\alpha$ is the tightest Chernoff upper bound on VaR $_\alpha$, which can be expressed in terms of the cumulant generating function as

$$\text{EVaR}_\alpha(R) = \inf_{\lambda > 0} \frac{K(w, -\lambda) - \log \alpha}{\lambda} \geq \text{VaR}_\alpha(R)$$

(It is also an upper bound on CVaR $_\alpha$). Cajas (2021) and Shen et al. (2022) describe convex optimization problems involving EVaR with the expectation replaced with its sample approximation.

Minimum EVaR portfolio To minimize EVaR $_\alpha(R)$, we solve the optimization problem

$$\begin{aligned} &\text{minimize } \frac{K(w, -\lambda) - \log \alpha}{\lambda} \\ &\text{subject to } \mathbf{1}^T w = 1, \quad w \in \Omega, \quad \lambda > 0, \end{aligned} \tag{12}$$

with variables $w \in \mathbf{R}^n$ and $\lambda \in \mathbf{R}$.

There is a close connection between this problem and the exponential utility maximization problem (7). Suppose that w^* and λ^* are optimal for (12). Then w^* is also optimal for the exponential utility problem (7), with risk aversion parameter λ^* . Thus we can think of minimizing EVaR $_\alpha(R)$ as simply choosing a value of the risk aversion parameter in EGM so as to minimize the tightest Chernoff upper bound on CVaR $_\alpha$. We will refer to portfolio construction using (12) also as EGM, since any such portfolio is optimal for EGM with some value of risk aversion, and also, conveniently, entropic and exponential both start with E.

4.2 Convex formulation

The problem (12) is not a convex optimization problem since the objective is not jointly convex in w and λ . But a change of variable can give us an equivalent convex

problem. Instead of using the variable λ , we use the new variable $\delta = 1/\lambda$, and the problem (12) becomes

$$\begin{aligned} & \text{minimize } \delta K(w/\delta, -1) - \delta \log \alpha \\ & \text{subject to } \mathbf{1}^T w = 1, \quad w \in \Omega, \quad \delta > 0, \end{aligned} \quad (13)$$

with variables $w \in \mathbf{R}^n$ and $\delta \in \mathbf{R}$. (We use the identity (4) above.) This objective is jointly convex in the variables w and δ , since it is the perspective function of $K(w, -1)$, which is convex in w (Boyd Vandenberghe 2004, §3.2.6), so (13) is a convex optimization problem, which is readily solved. (The constraint $\delta > 0$ is actually not needed, since the perspective function is defined to be $+\infty$ if $\delta \leq 0$.)

Unfortunately, current DSLs for convex optimization do not automate the creation of the perspective of a function, so the problem (13) cannot simply be typed in; we must form the the perspective function by hand, as outlined below in “Appendix A”.

There are also simple methods that can be used to solve it, with a modest loss in efficiency, that are immediately compatible with DSLs. One method is alternating optimization, where we alternate between fixing δ and optimizing over w (easy with current DSLs), and fixing w and optimizing over δ (minimization of a scalar convex function, which can be done by many simple methods). To start we can replace K with the lower bound (9) (or the upper bound (10)), to obtain the approximate problem

$$\begin{aligned} & \text{minimize } \max_i \left(\delta \log(\pi_i/\alpha) - \mu_i^T w + \frac{w^T \Sigma_i w}{2\delta} \right) \\ & \text{subject to } \mathbf{1}^T w = 1, \quad w \in \Omega, \quad \delta > 0. \end{aligned} \quad (14)$$

This problem is convex, and also immediately representable in DSLs using the quadratic-over-linear function for the last term in the objective. (Here too the constraint $\delta > 0$ is redundant, since the quadratic-over-linear function is defined as $+\infty$ if the denominator is not positive.)

4.3 Special cases

Gaussian returns When $k = 1$, our GM return distribution is Gaussian and we have

$$\delta K(w/\delta, -1) - \delta \log \alpha = -\delta \log \alpha - \mu_1^T w + \frac{w^T \Sigma_1 w}{2\delta}, \quad (15)$$

with variables w and γ . This objective is readily minimized using DSLs, using the quadratic-over-linear function for the last term.

The value of δ that minimizes this, with fixed w , is

$$\delta = \left(\frac{w^T \Sigma_1 w}{-2 \log \alpha} \right)^{1/2}.$$

Thus we see that for Gaussian returns, the portfolio that minimizes EVaR_α is in fact Markowitz, with the specific choice of risk aversion parameter

$$\gamma = \left(\frac{-2 \log \alpha}{w^T \Sigma_1 w} \right)^{1/2}.$$

(This depends on w , so to find it we must solve the convex problem with objective (15).) We see that as α decreases, the associated risk aversion increases, which makes sense.

Substituting the optimal value of δ into the objective (15), we find that the objective is

$$-\mu_1^T w + (-2 \log \alpha)^{1/2} \left(w^T \Sigma_1 w \right)^{1/2},$$

plus a constant. Thus we maximize a risk adjusted return, using the standard deviation instead of the traditional variance as risk, and the very specific risk aversion constant $(-2 \log \alpha)^{1/2}$.

Finite values returns When $\Sigma_i = 0$, so r takes on only the values μ_1, \dots, μ_k , the problem (13) can be expressed as

$$\begin{aligned} &\text{minimize } \delta \log \left(\sum_{i=1}^k (\pi_i / \alpha) \exp(-\mu_i^T w / \delta) \right) \\ &\text{subject to } \mathbf{1}^T w = 1, \quad w \in \Omega. \end{aligned} \tag{16}$$

5 Conclusions

In this paper we have shown that two specific portfolio construction problems, maximizing expected exponential utility and minimizing entropic value at risk, with a Gaussian mixture return model, can be formulated as convex optimization problems, and exactly solved with no need for return samples or Monte Carlo approximations. The resulting problems are not much harder to solve than a mean-variance problem, but have the advantage of directly handling return distributions with substantial skews or large tails.

Our focus in this paper is on the formulation of these portfolio construction problems as tractable convex optimization problems that do not need return samples. In a future paper we will report on practical portfolio construction using these methods.

Data availability Data sharing is not applicable to this article as no datasets were generated or analysed in the current study.

A Graph form representation of EVaR

A.1 Graph form representation

In this section we show how to express the objective of (13) in graph form (Grant and Boyd 2008), which is the basic representation of a function in DSLs for convex optimization, that rely on disciplined convex programming (DCP) (Grant et al. 2006). In a recent paper Cajas gave a graph form description of EVaR, for the special case when $\Sigma_i = 0$, *i.e.*, for a finite values return model (Cajas 2021). Thus we are extending his formulation from a finite values return model to a GM return model.

The graph form of a function $f : \mathbf{R}^n \rightarrow \mathbf{R}$ expresses the epigraph of f as the inverse image of a cone under an affine mapping. (For practical use, the cone must be a Cartesian products of cones supported by the solver.) The graph form of f is

$$\text{epi } f = \{(x, t) \mid f(x) \leq t\} = \{(x, t) \mid \exists z \ Fx + Gz + td + e \in C\}, \quad (17)$$

where $F \in \mathbf{R}^{p \times n}$, $G \in \mathbf{R}^{p \times m}$, $d \in \mathbf{R}^p$, and $e \in \mathbf{R}^p$ are the coefficients, and $C \subseteq \mathbf{R}^p$ is a cone, typically a Cartesian product of simple, standard cones, such as the nonnegative cone, second-order cone, and exponential cone. Such a representation allows f to be used in any DSL based on DCP, in the objective or constraint functions.

Specifically we work out a graph form for the perspective

$$P(w, \delta) = \delta K(w/\delta),$$

where $K = S(g_1, \dots, g_k)$, with

$$g_i(w) = \log(\pi_i) - \mu_i^T w + \frac{1}{2} w^T \Sigma_i w, \quad i = 1, \dots, k,$$

and S is the soft-max or log-sum-exp function (6).

A.2 Graph form calculus

We view P as a composition of four operations: an affine pre-composition, then an affine post-composition, then composition, and finally, the perspective. We show here generic methods for carrying out these operations using graph form representations. The first three operations, affine pre-composition, affine post-composition, and composition, are known (and indeed, used in all DSLs for convex optimization); we give them here for completeness. The last one, the perspective transform, is not well known, but is mentioned in Moehle and Boyd (2015).

Affine pre-composition Suppose f has graph form

$$\text{epi } f = \{(x, t) \mid \exists z \ Fx + Gz + td + e \in C\},$$

and g is the affine pre-composition $g(x) = f(Ax + b)$. Then g has graph form

$$\text{epi } g = \{(x, t) \mid \exists z \, FAx + Gz + td + (Fb + e) \in C\}. \tag{18}$$

Affine post-composition Suppose f has graph form

$$\text{epi } f = \{(x, t) \mid \exists z \, Fx + Gz + td + e \in C\},$$

and $h(x) = af(x) + b$, where $a \in \mathbf{R}_+$ and $b \in \mathbf{R}$. Then h has graph form

$$\text{epi } h = \{(x, t) \mid \exists z \, Fx + Gz + t(d/a) + (e - (b/a)d)\}. \tag{19}$$

Composition Suppose that g_i are convex functions with graph forms

$$\text{epi } g_i = \{(w, t_i) \mid \exists z_i \, F_i w + G_i z_i + t_i d_i + e_i \in C_i\}, \quad i = 1, \dots, k,$$

and S is a convex function with graph form

$$\text{epi } S = \{(u, t) \mid \exists z_0 \, F_0 u + G_0 z_0 + td_0 + e_0 \in C_0\}.$$

We assume that S is increasing in each of its arguments, so the composition $K = S(g_1, \dots, g_k)$ is convex. Then K has graph form

$$\text{epi } K = \left\{ (w, t) \mid \begin{array}{l} \exists z_0, t_i, z_i \, F_0(t_1, \dots, t_k) + G_0 z_0 + td_0 + e_0 \in C_0, \\ F_i w + G_i z_i + t_i d_i + e_i \in C_i, \quad i = 1, \dots, k \end{array} \right\}. \tag{20}$$

(We can stack the affine functions of (w, t) , and use the product cone $C_0 \times \dots \times C_k$ as the cone in the representation of K .)

Perspective Here we show how to construct a graph form of the perspective of a function given in graph form. The perspective of $f : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$ is the function $p : \mathbf{R}^{n+1} \rightarrow \mathbf{R} \cup \{\infty\}$ defined by

$$p(x, s) = \begin{cases} sf(x/s) & s > 0 \\ 0 & s = 0, \, x = 0 \\ \infty & \text{otherwise.} \end{cases}$$

(See Boyd Vandenberghe 2004, §3.2.6 or Urruty and Lemaréchal 1993, §IV.2.2.) Then p has graph form given by

$$\text{epi } p = \{(x, s, t) \mid sf(x/s) \leq t\} = \{(x, s, t) \mid f(x/s) \leq t/s\}.$$

Substituting this expression into the graph form of f given in (17), we have

$$\text{epi } p = \{(x, s, t) \mid \exists z \, F(x/s) + Gz + d(t/s) + e \in C\}.$$

Since $s > 0$ and C is a cone, we have

$$F(x/s) + Gz + d(t/s) + e \in C \iff Fx + G(sz) + dt + se \in C.$$

Thus, introducing a new affine description

$$\tilde{F} = [F \ e], \quad \tilde{e} = 0,$$

and with \tilde{z} a new auxilliary variable, p has graph form

$$\text{epi } p = \{(x, s, t) \mid \exists \tilde{z} \tilde{F}(x, s) + G\tilde{z} + td + \tilde{e} \in C\}. \tag{21}$$

Graph form of log-sum-exp

$$\begin{aligned} S(x) \leq t &\iff \log \left(\sum_{i=1}^k \exp(x_i) \right) \leq t \\ &\iff \log \left(\sum_{i=1}^k \exp(x_i - t) \right) \leq 0 \\ &\iff \sum_{i=1}^k \exp(x_i - t) \leq 1 \\ &\iff \sum_{i=1}^k u_i \leq 1, \quad (x_i - t, 1, u_i) \in C_{\text{exp}}, \quad i = 1, \dots, k, \end{aligned}$$

where

$$C_{\text{exp}} = \{(a, b, c) \mid e^{a/b} \leq c/b, b > 0\} \cup \{(a, 0, c) \mid a \leq 0, c \geq 0\}$$

is the exponential cone Glineur (2000); Chandrasekaran and Shah (2017), which is supported by several solvers. So S has graph form given by

$$\text{epi } S = \{(x, t) \mid \exists u \ F^{\text{LSE}}x + G^{\text{LSE}}u + td^{\text{LSE}} + e^{\text{LSE}} \in C^{\text{LSE}}\}, \tag{22}$$

with

$$F^{\text{LSE}} = \begin{bmatrix} 0 \\ e_1^T \\ 0 \\ 0 \\ \vdots \\ e_k^T \\ 0 \\ 0 \end{bmatrix}, \quad G^{\text{LSE}} = \begin{bmatrix} \mathbf{1}^T \\ 0 \\ 0 \\ e_1^T \\ \vdots \\ 0 \\ 0 \\ e_k^T \end{bmatrix}, \quad d^{\text{LSE}} = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ \vdots \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad e^{\text{LSE}} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

$$F^{\text{LSE}} \in \mathbf{R}^{(3k+1) \times k}, \quad G^{\text{LSE}} \in \mathbf{R}^{(3k+1) \times k}, \quad d^{\text{LSE}} \in \mathbf{R}^{3k+1}, \quad e^{\text{LSE}} \in \mathbf{R}^{3k+1},$$

and

$$C^{\text{LSE}} = \mathbf{R}_- \times C_{\text{exp}} \times \cdots \times C_{\text{exp}}.$$

The horizontal dividers denote separate blocks. After the first row, blocks of size 3 are repeated k times.

Graph form of quadratic To derive a graph form for the function $f(x) = x^T x$ with $x \in \mathbf{R}^n$, we first observe that

$$x^T x \leq t \iff \left\| \begin{bmatrix} x \\ \frac{t-1}{2} \end{bmatrix} \right\|_2 \leq \frac{t+1}{2}.$$

Therefore,

$$\text{epi } f = \{(x, t) \mid F^{\text{quad}}x + td^{\text{quad}} + e^{\text{quad}} \in C_{\text{SOCP}}\}, \tag{23}$$

with $F^{\text{quad}} \in \mathbf{R}^{(n+2) \times n}$, $d^{\text{quad}} \in \mathbf{R}^{n+2}$, $e^{\text{quad}} \in \mathbf{R}^{n+2}$ defined by

$$F^{\text{quad}} = \begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix}, \quad d^{\text{quad}} = \begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix}, \quad e^{\text{quad}} = \begin{bmatrix} 0 \\ -1/2 \\ 1/2 \end{bmatrix},$$

and where $C_{\text{SOCP}} = \{(x, t) \mid \|x\|_2 \leq t\}$ is the second order cone (Boyd Vandenberghe 2004, §4.4.2; Nesterov and Nemirovskii 1994).

A.3 Graph form of EVaR

Using the calculus outlined above, we can now develop a graph form of P , where $P(w, \delta) = \delta K(w/\delta)$. First, we use affine pre-composition to write

$$g_i(w) = \log(\pi_i) - \mu_i^T w + \frac{1}{2} w^T \Sigma_i w \leq t$$

as

$$f(A_i w + b_i) - \frac{1}{2} \mu_i^T \Sigma_i^{-1} \mu_i + \log(\pi_i) \leq t,$$

where

$$f(w) = w^T w, \quad A_i = \frac{1}{\sqrt{2}} \Sigma_i^{1/2}, \quad b_i = -\frac{\sqrt{2}}{2} \Sigma_i^{-1/2} \mu_i.$$

Thus, using our affine pre-composition expression (18) together with our graph form of the quadratic (23) and affine post-composition (19), we have

$$\text{epi } g_i = \left\{ (w, t_i) \mid \exists z_i (F^{\text{quad}} A_i)w + t_i d^{\text{quad}} + e_i^{\text{quad}} \in C_{\text{SOCP}} \right\},$$

with

$$e_i^{\text{quad}} = F^{\text{quad}} b_i + e^{\text{quad}} + \left(\frac{1}{2} \mu_i^T \Sigma_i^{-1} \mu_i - \log(\pi_i) \right) d^{\text{quad}}.$$

Then, using the graph form of log-sum-exp given in (22) and the composition rule give in (20), we can write the composition in graph form as

$$\begin{aligned} \text{epi } K &= \left\{ (w, t) \mid \exists z_0, t_1, \dots, t_k \begin{array}{l} F^{\text{LSE}}(t_1, \dots, t_k) + G^{\text{LSE}} z_0 + t d^{\text{LSE}} + e^{\text{LSE}} \in C^{\text{LSE}}, \\ F^{\text{quad}} A_i w + t_i d^{\text{quad}} + e_i^{\text{quad}} \in C_{\text{SOCP}}, \quad i = 1, \dots, k \end{array} \right\} \\ &= \{(w, t) \mid \exists z F^K w + G^K z + t d^K + e^K \in C^K\}, \end{aligned}$$

with

$F^K \in \mathbf{R}^{(3k+1+k(k+2)) \times k}$, $G^K \in \mathbf{R}^{(3k+1+k(n+2)) \times 2k}$, $d^K \in \mathbf{R}^{3k+1+k(n+2)}$, $e^K \in \mathbf{R}^{3k+1+k(n+2)}$ defined by

$$\begin{aligned} F^K &= \begin{bmatrix} 0 \\ F^{\text{quad}} A_1 \\ \vdots \\ F^{\text{quad}} A_k \end{bmatrix}, & G^K &= \begin{bmatrix} G^{\text{LSE}} & F^{\text{LSE}} \\ 0 & d^{\text{quad}} e_1^T \\ \vdots & \vdots \\ 0 & d^{\text{quad}} e_k^T \end{bmatrix}, \\ d^K &= \begin{bmatrix} d^{\text{LSE}} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, & e^K &= \begin{bmatrix} e^{\text{LSE}} \\ e_1^{\text{quad}} \\ \vdots \\ e_k^{\text{quad}} \end{bmatrix}, \end{aligned}$$

and

$$C^K = C^{\text{LSE}} \times C^{\text{SOCP}} \times \dots \times C^{\text{SOCP}}.$$

Finally, using the perspective rule given in (21), the perspective of K has graph form given by

$$\{(w, \delta, t) \mid \delta K(w/\delta) \leq t\} = \{(w, \delta, t) \mid \exists z \tilde{F}^K(w, \delta) + G^K z + t d^K \in C^K\},$$

with $\tilde{F}^K = [F^K \ e^K]$.

CVXPY specification CVXPY code for EVaR portfolio optimization using the graph form of $\delta K(w/\delta)$ is available at the repository https://github.com/cvxgrp/exp_util_gm_portfolio_opt.

References

- Ahmadi-Javid A (2012) Entropic value-at-risk: a new coherent risk measure. *J Optim Theory Appl* 155(3):1105–1123
- Akgiray V, Booth G (1987) Compound distribution models of stock returns: an empirical comparison. *J Financ Res* 10(3):269–280
- Ang A, Bekaert G (2004) How regimes affect asset allocation. *Financ Anal J* 60(2):86–99
- Ang A, Bekaert G (2015) International Asset Allocation With Regime Shifts. *Rev Financ Stud* 15(4):1137–1187
- Arditti F (1967) Risk and the required return on equity. *J Financ* 22(1):19–36
- Artzner P, Delbaen F, Eber J-M, Heath D (1999) Coherent measures of risk. *Math Finance* 9(3):203–228
- Ball C, Torous W (1983) A simplified jump process for common stock returns. *J Financ Quant Anal* 18(1):53–65
- Bingham N, Kiesel R (2001) Semi-parametric modelling in finance: theoretical foundations. *Quant Finance* 1:1–10
- Boyd S, Busseti E, Diamond S, Kahn R, Koh K, Nystrup P, Speth J (2017) Multi-period trading via convex optimization. arXiv preprint [arXiv:1705.00109](https://arxiv.org/abs/1705.00109)
- Boyd S, Vandenberghe L (2004) *Convex Optimization*. Cambridge University Press, Cambridge
- Buckley I, Saunders D, Seco L (2008) Portfolio optimization when asset returns have the Gaussian mixture distribution. *Eur J Oper Res* 185(3):1434–1461
- Cajas D (2021) Entropic portfolio optimization: a disciplined convex programming framework. *SSRN Electronic Journal*, 01
- Campbell R, Koedijk K, Kofman P (2002) Increased correlation in bear markets. *Financ Anal J* 58(1):87–94
- Chandrasekaran V, Shah P (2017) Relative entropy optimization and its applications. *Math Program* 161(1):1–32
- Diamond S, Boyd S (2016) CVXPY: a python-embedded modeling language for convex optimization. *J Mach Learn Res* 17(1):2909–2913
- Fama E (1965) The behavior of stock-market prices. *J Bus* 38(1):34–105
- Fu A, Narasimhan B, Boyd S (2020) CVXR: an R package for disciplined convex optimization. *J Stat Softw* 94(14):1–34
- Glineur F (2000) An extended conic formulation for geometric optimization. *Found Comput Decis Sci* 25(3):161–174
- Grant M, Boyd S (2008) Graph implementations for nonsmooth convex programs. In: *Recent advances in learning and control*, Springer, pp 95–110
- Grant M, Boyd S (2014) CVX: Matlab software for disciplined convex programming, version 2.1
- Grant M, Boyd S, Ye Y (2006) Disciplined convex programming. In: *Global optimization*, Springer, pp 155–210
- Grinold R (1999) Mean-variance and scenario-based approaches to portfolio selection. *J Portf Manag* 25(2):10–22
- Grinold R, Kahn R (1999) *Active portfolio management: a quantitative approach for producing superior returns and selecting superior returns and controlling risk*. McGraw-Hill Library of Investment and Finance, Mcgraw-hill
- Gupta A, Dhingra B (2012) Stock market prediction using hidden Markov models. In: *2012 students conference on engineering and systems*, IEEE, pp 1–4
- Hanoch G, Levy H (1970) Efficient portfolio selection with quadratic and cubic utility. *J Bus* 43(2):181–189
- Hentati-Kaffel R, Prigent J-L (2014) Portfolio Optimization within Mixture of Distributions. Working Papers 01066105, HAL, Sept
- Jondeau E, Rockinger M (2006) Optimal portfolio allocation under higher moments. *Eur Financ Manag* 12(1):29–55
- Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47(2):263–291
- Kon S (1984) Models of stock returns—a comparison. *J Financ* 39(1):147–165
- Levy H, Markowitz H (1979) Approximating expected utility by a function of mean and variance. *Am Econ Rev* 69(3):308–317
- Markowitz H (1952) Portfolio selection. *J Finance* 7:77–91
- Markowitz H (1959) *Portfolio selection: efficient diversification of investments*. Yale University Press, London

- MathWorks. Backtest investment strategies. <https://www.mathworks.com/help/finance/backtest-investment-strategies.html>, accessed 2022-04-25
- Merton R (1969) Lifetime portfolio selection under uncertainty: the continuous-time case. *Rev Econ Stat* 51:247–257
- Moehle N, Boyd S (2015) A perspective-based convex relaxation for switched-affine optimal control. *Syst Control Lett* 86:34–40
- Nesterov Y, Nemirovskii A (1994) Interior-point polynomial algorithms in convex programming. SIAM, New Delhi
- Neuberger A (2012) Realized Skewness. *Rev Financ Stud* 25(11):3423–34509
- Nystrup P, Boyd S, Lindstrom E, Madsen H (2019) Multi-period portfolio selection with drawdown control. *Ann Oper Res* 282:245–271
- Rockafellar R (2007) Coherent approaches to risk in optimization under uncertainty. In: *OR tools and applications: glimpses of future technologies*, pp 38–61. *Inform*s
- Rockafellar R, Uryasev S (2000) Optimization of conditional value-at-risk. *J Risk* 04:21–41
- Rockafellar R, Uryasev S (2002) Conditional value-at-risk for general loss distributions. *J Bank Finance* 26(7):1443–1471
- Rydén T, Teräsvirta T, Åsbrink S (1998) Stylized facts of daily return series and the hidden Markov model. *J Appl Econom* 13(3):217–244
- Saha A (1993) Expo-power utility: a ‘flexible’ form for absolute and relative risk aversion. *Am J Agric Econ* 75(4):905–913
- Schaller M, Banjac G, Diamond S, Agrawal A, Stellato B, Boyd S (2022) Embedded code generation with CVXPY. *IEEE Control Syst Lett* 6(2022):2653–2658
- Scott R, Horvath P (1980) On the direction of preference for moments of higher order than the variance. *J Financ* 35(4):915–919
- Shapiro A, Dentcheva D, Ruszczyński A (2021) *Lectures on stochastic programming: modeling and theory*. SIAM, New Delhi
- Shen X, Ali A, Boyd S (2022) Minimizing oracle-structured composite functions. *Optim Eng* pp 1–35
- Stellato B, Banjac G, Goulart P, Bemporad A, Boyd S (2020) OSQP: an operator splitting solver for quadratic programs. *Math Program Comput* 12(4):637–672
- Urruty J-B, Lemaréchal C (1993) *Convex analysis and minimization algorithms*. Springer-Verlag, Berlin
- von Neumann J, Morgenstern O, Rubinstein A (1944) *Theory of games and economic behavior* (60th Anniversary Commemorative Edition). Princeton University Press, Princeton

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.