PORTFOLIO OPTIMIZATION WHEN ASSET RETURNS HAVE THE GAUSSIAN MIXTURE DISTRIBUTION

IAN BUCKLEY, GUSTAVO COMEZAÑA, BEN DJERROUD, AND LUIS SECO

Abstract. Portfolios of assets whose returns have the Gaussian mixture distribution are optimized in the static setting to find portfolio weights and efficient frontiers using the probability of outperforming a target return and Hodges' modified Sharpe ratio objective functions. The sensitivities of optimal portfolio weights to the probability of the market being in the distressed regime are shown to give valuable diagnostic information. A two-stage optimization procedure is presented in which the high-dimensional non-linear optimization problem can be decomposed into a related quadratic programming problem, coupled to a lower-dimensional non-linear problem.

Contents

Date: Tuesday, February 18, 2003 at 20:32 and, in revised form, ...

Key words and phrases. Mixture of normals distribution, Gaussian mixture distribution, portfolio optimization, market distress, hedge fund portfolio, Sharpe ratio, efficient frontier, Hodges' modified Sharpe ratio, exponential utility, correlation switching, regime switching, asset allocation, commodity trading advisor, probability of shortfall, distress sensitivities.

This work was completed with the support of .... Scholarship.

1. Introduction

In this article Markowitz mean-variance portfolio theory [12], the foundation for single-period investment theory, is generalized to describe portfolios of assets whose returns are described by the (finite) *Gaussian mixture* (GM) (alternatively *mixture of normals*) distribution. Whilst the assets in the universe could be of the conventional variety, such as equities or bonds, our primary goal is to develop a framework which lends itself to the management of portfolios of hedge funds or even for optimally combining the recommendations from a group of commodity trading advisors (CTAs). That is to say, we seek an approach suitable for finding an optimal fund of funds. Because of the infinite variety of hedge fund and CTA strategies and the speed at which a given fund's composition can be changed, the alternative assets that we describe are not expected to behave like conventional assets such as individual equities, bonds or even long-only funds such as index tracker funds and exchange traded funds (ETFs). Indeed we need to be prepared for their prices to be less predictable, more volatile and to have more exotic distributions. The new approach is ideal for an industrial setting, providing considerable additional flexibility over and above a standard Markowitz approach, with only a modest increase in complexity.

The assumption that asset returns have the multivariate Gaussian distribution is a reasonable first approximation to reality and gives rise to tractable theories. Many theories forming the foundations of mathematical finance adopt this conjecture, including Black-Scholes-Merton option pricing theory, Markowitz portfolio theory, and the CAPM and APT equity pricing models.

However, it is well-known that for assets, both in the conventional sense of equities and bonds, but also in a broader sense, for example in the form of country or sector-based equity or bond indices, and even more so for alternative investments such as hedge funds and CTAs, the situation is more complex. The purpose of the generalization described in this paper is to address two well-known limitations with the assumption that asset returns obey the multivariate Gaussian distribution with constant parameters over time:

- The skewed (asymmetric around the mean) and leptokurtotic (more kurtotic or "fat-tailed" than a Gaussian distribution) nature of marginal probability density functions (pdfs)
- The *asymmetric correlation* (or *correlation breakdown*) phenomenon, which describes the tendency for the correlations between asset returns to be dependent on the prevailing direction of the market. Typically correlations are larger in a bear market than a bull market.

The first point refers to the univariate distributions for returns that are observed if assets are considered one at a time. The second point describes effects that can only be observed when the returns to multiple assets are investigated together. One way to capture the dependence structure of multiple random variables in a risk management setting is by using *copulas* [22], [7], [9], [13], [18]. Copulas describe that part of the shape of the pdf that cannot be described by the marginal distributions. A (finite) Gaussian mixture distribution, as described in this paper, can be used to approximate a general multivariate distribution, as equivalently expressed either as a pdf or decomposed into a series of univariate marginal distributions and a copula.

In fact, strictly the term "asymmetric correlation" describes a proposed explanation for a phenomenon, rather than the underlying, fundamental phenomenon itself. The latter is simply that the iso-probability contours of the multivariate pdf for asset returns are less symmetric than the ellipsoidal contours of the multivariate normal distribution. Apart from those distributions with pdfs with ellipsoidal iso-probability surfaces, such as the Gaussian and multivariate t, for all other multivariate distributions the correlation matrix does not provide an adequate summary of the dependence structure of the constituent risks. Considering multiple regimes within which the ordinary constant parameter Gaussian assumptions prevail successfully gives rise to a model that reflects reality better than a standard Gaussian model, without having to depart too radically from it. However, it is certainly not the only way to construct a model with non-ellipsoidal iso-probability contours. Indeed alternative assumptions to achieve the same ends may render the concept of correlation redundant and statements about its evolution over time, meaningless.

Apart from the distributional assumptions, the second fundamental difference between the novel approach described in this paper (the *GM approach*) and the Markowitz mean-variance approach, is that we adopt a different objective function. This is essential in order to get different optimal portfolio weights. When we refer to the GM approach, the use of an alternative objective will be implicit. A drawback of using a more exotic objective than the variance is that the resultant optimization problem is not a linear-quadratic program (LQP), as it is when variance is the chosen risk measure. However, the nonlinear optimization problem with multiple local extrema that replaces it can be solved reliably and quickly with commonly available routines, at least with a moderate number of assets (i.e. less than ten).

The plan for the paper is as follows. Evidence of covariance regimes over time is given in Section 2, including a summary of existing results from the literature for conventional assets and original results for alternative investments. We are introduced to the GM distribution in Section 3 in which we find definitions of GM distributed random variables, estimation issues, and some identities concerning moments and linear combinations of GM variables. In Section 4 we take the probability of shortfall objective, assume GM distributed asset returns and consider a theory based on these ingredients. Key theoretical and numerical results are described. Section 5 contains our conclusions.

2. Evidence of covariance regimes

During the aftermath of the 1987 market crash, a deficiency in the risk models based on the multivariate normal distribution received increased attention: a simultaneous downward movement in all the markets of the world was a more frequent occurrence than the models predicted when calibrated using asset returns observed during tranquil periods. Diversifying amongst different assets or markets was less effective at reducing risk than many participants had hitherto believed. Important investigations of the contagion phenomenon include [8], [4], [20].

In common with their conventional asset counterparts, alternative assets exhibit the correlation breakdown phenomenon. As evidence we present Figure 1, which shows the correlation matrices between hedge fund returns in tranquil and distressed regimes. During tranquil periods correlations are small, whereas during periods of market distress, the asset returns become highly correlated, with off-diagonal correlation values close to one.

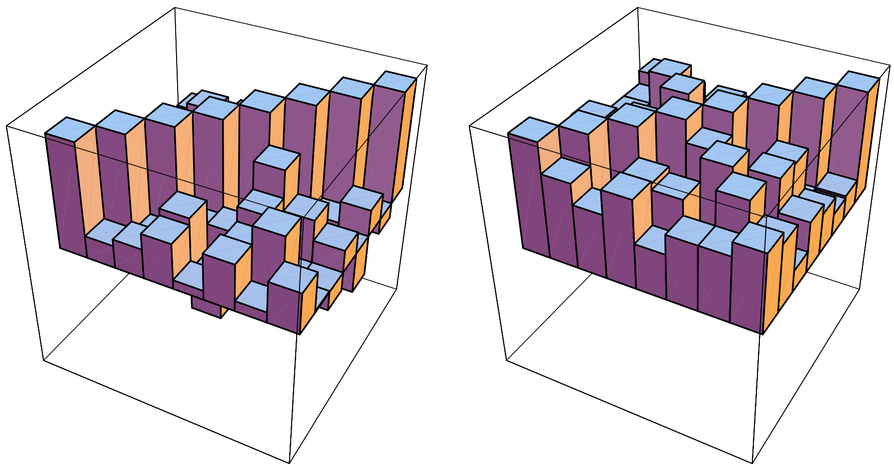

FIGURE 1. Bar charts to show typical values of correlations between the returns of a group of eight hedge funds during tranquil (left) and distressed (right) periods.

We note in passing that because finding portfolio weights in a mean-variance setting is tantamount to inverting the covariance matrix, then errors in the optimal portfolio weights will be most sensitive to errors in the covariance matrix when the latter is closest to being singular (having a determinant close to zero). The covariance matrix will be more singular in the distressed regime than in the tranquil regime.

3. Gaussian mixture distribution

The Gaussian mixture distribution is selected from the range of parametric alternatives to the normal distribution for its tractability: calculations using it often closely resemble those using the normal distribution. Whilst there are many univariate parametric probability distributions, (e.g. hyperbolic, t, generalized beta, α-stable) the list for multivariate distributions is shorter (e.g. t, α-stable).

The GM distribution has been used before in the field of finance, mostly in its univariate guise for the estimation of Value at Risk (VaR). [16] develop a model for estimating VaR in which the user is free to choose any probability distribution for the daily changes in each market variable and employ the univariate mixture of normals distribution as an example. In the same field, [28] assumes probability distributions for each of the parameters describing the mixture of normals and uses a Bayesian updating scheme; and [26] uses a quasi-Bayesian maximum likelihood estimation procedure. The current RiskMetrics$^{\text{TM}}$ methodology uses GM with a mixture of two normal distributions. More recently GM models have been used [19] to model futures markets and for portfolio risk management and by [10] for credit risk. [27] develops an efficient analytical Monte Carlo method for generating changes in asset prices using a multivariate mixture of normal distributions with arbitrary covariance matrix. [25] describe computational tools for the calculation of VaR and other more sophisticated risk measures such as shortfall, Max-VaR, conditional VaR and conditional risk measures that aim to take account of the heteroskedastic structure of time series.

This paper describes the static, single period setting only, in which distributions of random variables are sufficient to specify the model. In this setting the key idea is that we build an exotic distribution by mixing simple ones, namely copies of the normal distribution. Our only assumption is that over an interval of time, the returns to the assets in the universe are described by the multivariate GM distribution. It is unnecessary to make further assumptions about the nature of the asset price evolution during this interval. However, we do maintain an interest in the dynamic case, i.e. the multiple period discrete or continuous time setting, because we wish to motivate the use of the GM distribution and we prefer to construct static models that extend naturally to the dynamic case.

There is a growing body of work in which exotic (asset return) stochastic *processes* have been constructed by mixing simpler ones. Processes for asset returns may be constructed from unconditional or conditional distributions. As an example of the latter case, by mixing autoregressive processes such as ARCH and GARCH, processes can be constructed that can account for both the heteroscedastic and leptokurtic nature of financial time series. See [23] and [11].

GM distributions can arise naturally as the level of certain stochastic processes at a point in time, conditional on the level at an earlier time e.g. Markov (regime) switching models and jump processes. Regime switching models describe processes in which parameters of a continuous diffusion process may change discontinuously according to the realized stochastic path through an associated Markov chain. We conclude that GM distributions are better motivated and less contrived than first impressions might suggest. Recent applications of regime switching asset return processes include: in the field of Merton-style option pricing theory, [17], and in portfolio management [5] (CAPM) and [2], [1] (international diversification).

Mixture distributions have the appeal that by adding together a sufficient number of component distributions, any multivariate distribution may be approximated to arbitrary accuracy. With an infinite number of contributions, any distribution can be reconstructed exactly.

As far as estimation is concerned, a disadvantage of using the GM distribution is that the log-likelihood function does not have a global minimum. A resolution to this problem explored in [14] is to use a modified log likelihood function. Because of the use of the GM distribution and other mixture distributions in image processing, clustering, and unsupervised learning a host of estimation techniques have been developed to address this problem [6]. When using the GM distribution to model asset returns, [27] employs the EM algorithm.

3.1. **Definitions and identities.** In Appendix A the definition of the GM distribution and various identities involving it are presented.

3.2. **Visualizing the GM distribution.** In Figure 2, in four contour plots of pdfs, we see how two bivariate Gaussian distributions (top row) can be added to yield a GM distribution (bottom, left). Note the potential for highly non-elliptic iso-probability contours when the GM distribution is used. For comparison a bivariate normal with the same sample means, variances and covariances is included (bottom, right). If the GM distribution given were used to describe the returns to two assets, the lobe pointing down and to the left of the figure would describe the

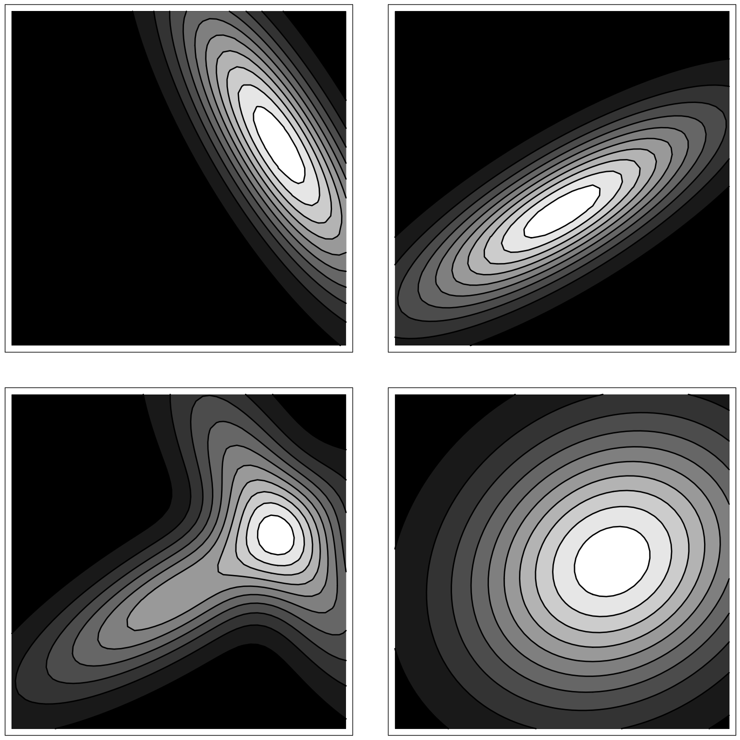

FIGURE 2. Contour plots of probability density functions. The top row contains two bivariate Gaussian distributions - potentially for the tranquil (left) and distressed (right) regimes. The bottom row illustrates the composite Gaussian mixture distribution obtained by mixing the two distributions from the top row (left) and a bivariate normal distribution with the same means and variance-covariance matrix as the composite (right).

propensity of the two assets to decline sharply together (the asymmetric correlation phenomenon). The Gaussian distribution, with its elliptic contours, is clearly unable to capture this feature.

These examples illustrate the two asset case. With three assets the contours for the Gaussian distribution are three-dimensional ellipsoids (rather than ellipses in two dimensions), and for the GM distribution the contours are complicated lobed surfaces embedded in a three-dimensional space.

4. Portfolio optimization

In our framework, the n Gaussian contributions to the GM distribution are associated with n regimes. In all the numerical examples, we take $n = 2$, and call the two regimes the *tranquil* and *distressed* regimes.

4.1. **Probability of shortfall as a risk measure.** When variance is used as the optimization objective, the efficient frontier is independent of the distribution for the asset returns. Therefore, to achieve non-trivial results in the GM setting, we adopt alternative, non-quadratic objective functions. One with the advantage of being intuitive for practitioners is the probability of shortfall below a target return (PoS). When returns have a normal distribution, minimizing the PoS is equivalent to maximizing the out-performance Sharpe ratio (ratio of difference between realized return and target return in the numerator to volatility in the denominator). Therefore adopting the GM approach will not represent a significant departure from current practice for many investors.

Note that the PoS is *not* the same as the VaR: the former is the probability beyond a given point on the distribution (e.g. a target return of 18.2%), the latter is the point on the distribution such that the probability of being beyond that point is equal to a given value (e.g. 1% or 5%).

In the GM setting of this paper the probability of shortfall objective is the probability that a univariate mixture of normals random variable, with regime means $\bar{\mu}_i = \boldsymbol{\mu}_i.\boldsymbol{\theta}$ and regime variances $\bar{\sigma}_i^2 = \boldsymbol{\theta}'.\boldsymbol{V}_i.\boldsymbol{\theta}$, exceeds the target return k. From the expression for the univariate mixture of normals CDF Eqn. A.1.2, above, it can be shown that:

Proposition 4.1.1. *The probability that the portfolio return falls short of the target* k *is:*

$$F_k(\boldsymbol{\theta}) = \sum_{i=1}^{n} w_i \Phi\left(\frac{\boldsymbol{\mu}_i.\boldsymbol{\theta} - k}{\sqrt{\boldsymbol{\theta}'.\boldsymbol{V}_i.\boldsymbol{\theta}}}\right) \tag{4.1.1}$$

4.2. **Hodges ratio.** To address the paradoxes inherent in using the Sharpe ratio [24] as a measure for ranking the desirability of payoff distributions, Hodges [15] introduced a measure of risk for performance ranking based on the exponential utility function

$$U(w) = -e^{-\lambda w}, \tag{4.2.1}$$

which is intuitive and is a generalization of the Sharpe ratio, reducing to it for normally distributed returns. We refer to Hodges modified Sharpe ratio as the Hodges ratio (HR).

Because the HR uses a utility function with constant absolute risk aversion, the composition of the optimal portfolio is independent of the coefficient of risk tolerance and so without loss of generality this can be taken to be one. It reduces to the Sharpe ratio for normally distributed returns and yet can be applied to any return distribution. It is compatible with stochastic dominance: an investment opportunity that outperforms another in every state of the world necessarily has a higher HR. This is one of the properties of a coherent risk measure [3].

Madan and McPhail [21] point out an apparent limitation with the HR, namely that it sometimes perversely considers distributions with large negative skew to be desirable investments, because the approach effectively shorts the risky asset. It turns out that this problem is simply resolved by scaling the utility function by the sign of the weight in the risky assets, ξ. In this paper, for all sensible parameter values $\xi > 0$, so this refinement was unnecessary.

4.3. **Lower partial moments.** Partial moments of a random variable are its moments over a sub-interval of the domain of the pdf. In particular, the lower and upper nth-partial moments (LPM, UPM) are the nth order moments for the random variable when it takes nonzero values below and above a given value, respectively.

Necessarily, as a statistic it summarises the properties of the distribution only to one side of the parameter.

The LPM is a good candidate for a risk measure because it only considers those states in which return on an asset is below a pre-specified target rate. To real individuals a downside risk measure such as this corresponds more closely to their concept of risk than the ubiquitous variance risk measure.

We present some useful identities:

$$\int_{-\infty}^{a} (a-x)\ \phi(x)\,dx = a\,\Phi(a) + \phi(a) \tag{4.3.1}$$

$$= \int_{-\infty}^{a} \Phi(x)\,dx \tag{4.3.2}$$

$$\int_{-\infty}^{a} (a-x)^2\ \phi(x)\,dx = \left(1+a^2\right)\ \Phi(a) + a\,\phi(a) \tag{4.3.3}$$

$$\int_{-\infty}^{a} \frac{a-x}{\sigma}\,\phi_{\mu,\sigma}(x)\,dx = \frac{a-\mu}{\sigma}\,\Phi_{\mu,\sigma}(a) + \sigma\,\phi_{\mu,\sigma}(a) \tag{4.3.4}$$

$$\frac{1}{\sigma^2}\int_{-\infty}^{a} (a-x)^2\,\phi_{\mu,\sigma}(x)\,dx = (a-\mu)\ \phi_{\mu,\sigma}(a) + \left(1+\frac{(a-\mu)^2}{\sigma^2}\right)\ \Phi_{\mu,\sigma}(a) \tag{4.3.5}$$

4.4. **Optimization problems for different objectives.** In each case μ_{T} is the target portfolio expected return.

Markowitz:

$$\begin{array}{ll} \min\limits_{\boldsymbol{\theta}} & \boldsymbol{\theta}'.\boldsymbol{V}_{\mathrm{S}}.\boldsymbol{\theta} \\ \text{s.t.} & \boldsymbol{\theta}.\mathbf{1} = 1 \\ & \boldsymbol{\mu}_{\mathrm{S}}.\boldsymbol{\theta} \geq \mu_{\mathrm{T}} \end{array} \tag{4.4.1}$$

where $\boldsymbol{\mu}_{\mathrm{S}} = \sum_{i=1}^{n} w_i \boldsymbol{\mu}_i$ and $\boldsymbol{V}_{\mathrm{S}}$ are the sample mean and sample variance-covariance matrix, respectively. Note that the latter is *not* equal to the weighted sum of the regime variances. Instead it is calculated using the identity in Equation A.2.5 or equivalently A.2.6. Moments of GM distributions are the sums of the component distribution moments. However, for central moments, such as the variance, the expressions are more complicated.

Sharpe Ratio:

$$\begin{array}{ll} \max\limits_{\boldsymbol{\theta}} & \dfrac{\boldsymbol{\mu}_{\mathrm{S}}.\boldsymbol{\theta} - k}{\sqrt{\boldsymbol{V}_{\mathrm{S}}}} \\ \text{s.t.} & \boldsymbol{\theta}.\mathbf{1} = 1 \\ & \boldsymbol{\mu}_{\mathrm{S}}.\boldsymbol{\theta} \geq \mu_{\mathrm{T}} \end{array} \tag{4.4.2}$$

Probability of shortfall:

$$\begin{aligned} \max_{\boldsymbol{\theta}} \quad & \sum_{i=1}^{n} w_i \Phi\left(\frac{\boldsymbol{\mu}_i.\boldsymbol{\theta} - k}{\sqrt{\boldsymbol{\theta}'.\boldsymbol{V}_i.\boldsymbol{\theta}}}\right) \\ \text{s.t.} \quad & \boldsymbol{\theta}.\mathbf{1} = 1 \\ & \sum_{i=1}^{n} w_i \boldsymbol{\mu_i}.\boldsymbol{\theta} \geq \mu_{\mathrm{T}} \end{aligned} \tag{4.4.3}$$

Hodges ratio:

$$\begin{aligned} \max_{\boldsymbol{\theta},\xi} \quad & -\sum_{i=1}^{n} w_i \exp\left(-\left(\lambda\,\xi\,\boldsymbol{\mu}_i.\boldsymbol{\theta} - \frac{1}{2}\lambda^2\,\xi^2\,\boldsymbol{\theta}'.\boldsymbol{V}_i.\boldsymbol{\theta}\right)\right) \\ \text{s.t.} \quad & \boldsymbol{\theta}.\mathbf{1} = 1 \\ & \sum_{i=1}^{n} w_i \boldsymbol{\mu_i}.\boldsymbol{\theta} \geq \mu_{\mathrm{T}} \end{aligned} \tag{4.4.4}$$

where ξ is the weight in the risky assets and λ is the parameter for the exponential utility function. The remainder, $1 - \xi$ is invested in a risk-free bank account, taken to have zero return for simplicity. The above optimization problem is clearly equivalent to a similar one in which $\xi = 1$ and the budget constraint is removed. However, in the numerical examples, both the total weight in risky assets variable, ξ, and the budget constraint, are retained so that the normalised weights of the individual assets within the risky asset portion can be compared with the results from other objective functions. E.g., see 6.

Note that the exponential term in the objective can be derived by observing that if the random variable X has mean μ and variance V then

$$E[U(\xi X)] = -\exp\left(-\left(\lambda\,\xi\,\mu - \frac{1}{2}\lambda^2\,\xi^2\,V\right)\right)$$

where the utility function $U(x) = -e^{-\lambda\,x}$.

Lower partial moment, order $n = 1$:

$$\begin{aligned} \max_{\boldsymbol{\theta}} \quad & \sum_{i=1}^{n} w_i \left(\frac{k - \mu_i}{\sigma_i}\right) \Phi_{\mu_i,\sigma_i}(k) + \sigma_i\,\phi_{\mu_i,\sigma_i}(k) \\ \text{s.t.} \quad & \boldsymbol{\theta}.\mathbf{1} = 1 \\ & \sum_{i=1}^{n} w_i \mu_i \geq \mu_{\mathrm{T}} \end{aligned} \tag{4.4.5}$$

where $\mu_i = \boldsymbol{\mu}_i.\boldsymbol{\theta}$ and $\sigma_i = \sqrt{\boldsymbol{\theta}'.\boldsymbol{V}_i.\boldsymbol{\theta}}$

Lower partial moment, order $n = 2$:

$$\begin{aligned} \max_{\boldsymbol{\theta}} \quad & \sum_{i=1}^{n} w_i\,(k - \mu_i)\,\phi_{\mu_i,\sigma_i}(k) + \left(1 + \frac{(k - \mu_i)^2}{{\sigma_i}^2}\right) \Phi_{\mu_i,\sigma_i}(k) \\ \text{s.t.} \quad & \boldsymbol{\theta}.\mathbf{1} = 1 \\ & \sum_{i=1}^{n} w_i \mu_i \geq \mu_{\mathrm{T}} \end{aligned} \tag{4.4.6}$$

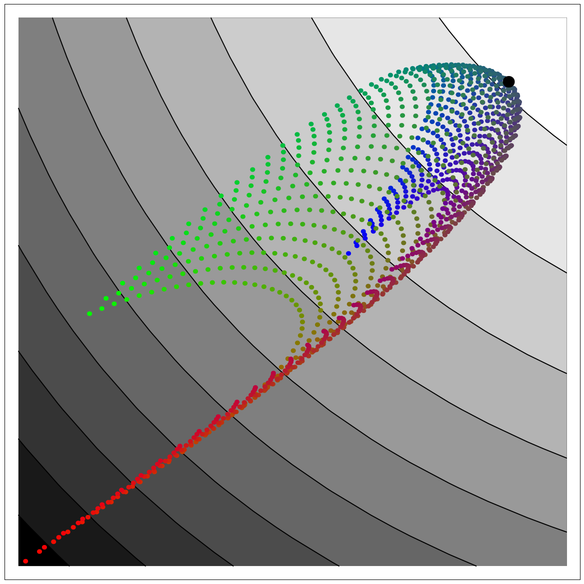

FIGURE 3. Scatter plot of portfolios in the investment opportunity set, overlaid on the contour plot of the probability of shortfall objective function. The axes are the out-performance Sharpe ratios in the tranquil (x) and distressed (y) regimes. The GM approach optimal portfolio is marked in the top righthand corner, on the efficient frontier. Typically, the portfolios that are optimal with respect to the mean-variance objectives in the tranquil and distressed regimes will be sub-optimal with respect to the probability of shortfall objective used in the GM approach. This is a three asset example.

4.5. **Investment opportunity set.** Figure 3 shows the investment opportunity set (IOS) in the 3-asset case, overlaid on the contour plot for the PoS objective function, in tranquil and distressed regime out-performance Sharpe ratio space. The investment opportunity set is the set of all attainable points in a given space that may be reached by constructing portfolios of the assets in a given universe. Typically variance and return are taken as the dimensions of a space to explore, but any statistics may be used. The GM approach optimal portfolio is marked in the top righthand corner, on the efficient frontier.

Compare this with Figure 4, which shows the IOSs for the tranquil and distressed regimes superimposed onto the same plot. The axes are the portfolio means and variances in the two regimes. Clearly the portfolio that maximizes the PoS objective, will be suboptimal with respect to the mean-variance efficient frontier in each of the component regimes.

4.6. **Numerical examples.** We present a three asset example with five objectives:

- Variance
- Probability of shortfall
- Hodges ratio
- Lower partial first moment

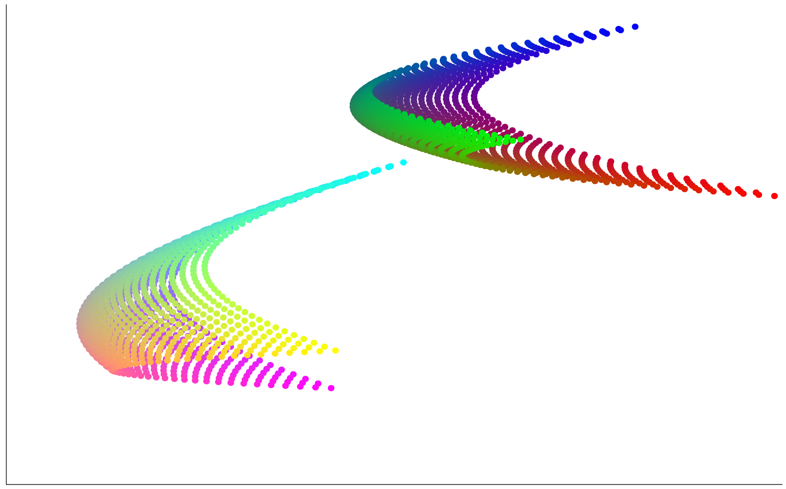

FIGURE 4. Investment opportunity sets for the tranquil and distressed regimes superimposed onto the same plot. The axes are the portfolio mean and variance. Typically the GM approach optimal portfolio will be sub-optimal with respect to both the tranquil and distressed mean-variance objectives. This is a three asset example.

- Lower partial second moment

All results are presented graphically. Parameter values are:

$$\mu_{\mathrm{T}} = \begin{pmatrix} 0.21 \\ 0.29 \\ 0.41 \end{pmatrix} \qquad \mu_{\mathrm{D}} = \begin{pmatrix} 0.21 \\ 0.29 \\ 0.41 \end{pmatrix} \tag{4.6.1}$$

$$\sigma_{\mathrm{T}} = \begin{pmatrix} 0.2 \\ 0.3 \\ 0.4 \end{pmatrix} \qquad \sigma_{\mathrm{D}} = \begin{pmatrix} 0.2 \\ 0.308 \\ 0.412 \end{pmatrix} \tag{4.6.2}$$

$$\rho_{\mathrm{T}} = \begin{pmatrix} 1. & 0. & 0. \\ 0. & 1. & 0. \\ 0. & 0. & 1. \end{pmatrix} \qquad \rho_{\mathrm{D}} = \begin{pmatrix} 1. & 0.974 & 0.971 \\ 0.974 & 1. & 0.984 \\ 0.971 & 0.984 & 1. \end{pmatrix} \tag{4.6.3}$$

Figure 5 shows 2-dimensional projections of the GM pdf as contour plots. The contours indicate iso-probability curves. The pictures are slices in the sense that the suppressed dimension has asset weight zero.

Figure 6 shows the optimal portfolio weights along the frontier. All heights on the plot are optimal objectives for a given target expected return.

Figure 7 shows the objective values for all five objectives based on optimal portfolio weights obtained by using a given, single objective. For this example, all optimal portfolio weights are similar and the five plots correspond closely.

Figure 8 shows the dependence of the optimal portfolio weights on the mixing parameter w. Note that pure tranquil (distressed) regime is on the right (left) of the figure. Therefore assets with positive (negative) slope on this diagram, i.e., those with a tendency to have small (big) positions in the distressed regime, are good for insuring against (speculating on) the market entering a distressed state.

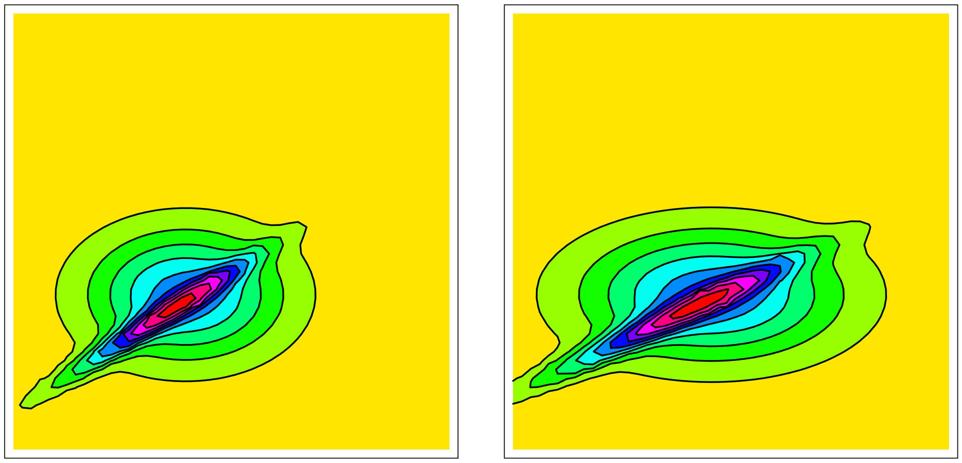

FIGURE 5. Contour plots of 2d projections of 3d asset return probability density function

Figure 9 shows the dependence of the rate of change of asset weight with respect to mixing parameter along the efficient frontier, i.e., for different values of desired expected return.

4.7. **Alternative algorithm for solving the non-linear problem.** The PoS objective gives rise to the non-linear optimization problem (Equation 4.4.6) of dimension equal to the number of assets in the universe, m. This can be solved using standard non-linear program (NLP) optimizers, but such tools are unable to exploit the close resemblance between the non-linear problem and a linear-quadratic program (LQP). An alternative approach is to recognize that the optimal portfolio weights, for a given target return μ_{T}, will be embedded in the $2(n-1)$-dimensional solution hyper-surface of a related LQP, where n is the number of regimes. The latter is that of minimizing a linear sum of the regime variances subject to the linear constraint that a linear sum of the regime expected returns be less than the return target. An NLP is still required to find the optimal solution, but the dimensionality of the problem is reduced from m to $2(n-1)$. Typically the former exceeds the latter. E.g., if there are thirty assets in the problem and two regimes.

We define some notation:

- Linear and quadratic functions of the portfolio weight vector, θ and the Sharpe ratio, for each regime i:

$$L_i(\theta) = \mu_i.\theta \tag{4.7.1}$$

$$Q_i(\theta) = \theta^T V_i.\theta \tag{4.7.2}$$

$$x_i(\theta) = \frac{L_i(\theta) - k}{\sqrt{Q_i(\theta)}} \tag{4.7.3}$$

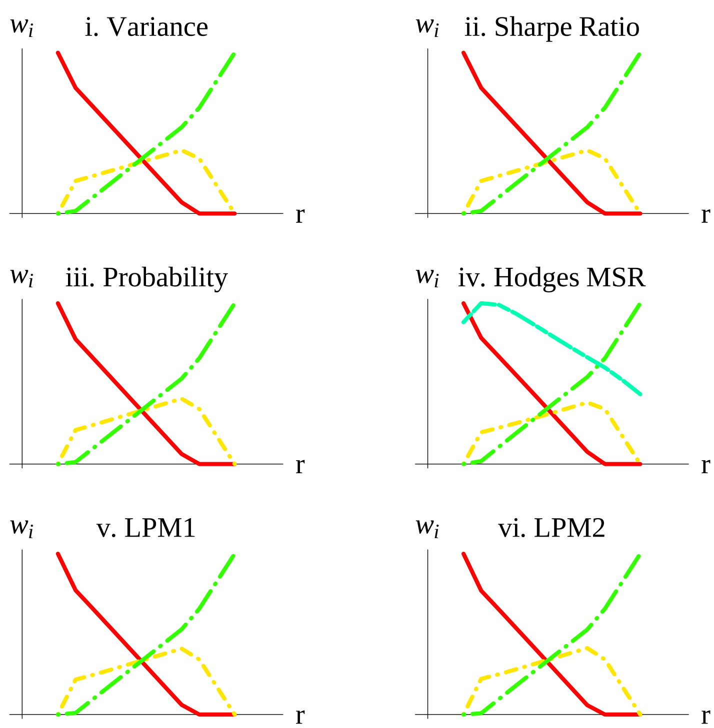

FIGURE 6. Optimal asset weights along the efficient frontier

- Objective functions, one non-linear, one quadratic in θ:

$$f_{\boldsymbol{\alpha}}(\theta) = \sum_{i=1}^{n} \alpha_i \Phi(x_i(\theta)) \tag{4.7.4}$$

$$g_{\boldsymbol{\beta}}(\theta) = \sum_{i=1}^{n} \beta_i Q_i(\theta)) \tag{4.7.5}$$

where $\alpha_i, \beta_i \geq 0 \forall i$.

The non-linear and linear-quadratic programmes are:

Definition 4.7.1. NLP(1)

$$\begin{aligned} \max_{\boldsymbol{\theta}} \quad & f_{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \\ \text{s.t.} \quad & \boldsymbol{\theta}.\mathbf{1} = 1 \\ & \sum_{i=1}^{n} \alpha_i L_i(\boldsymbol{\theta}) \geq p \end{aligned} \tag{4.7.6}$$

where $\boldsymbol{\alpha}.\mathbf{1} = 1$.

Solutions to NLP(1) will be denoted by $\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}$.

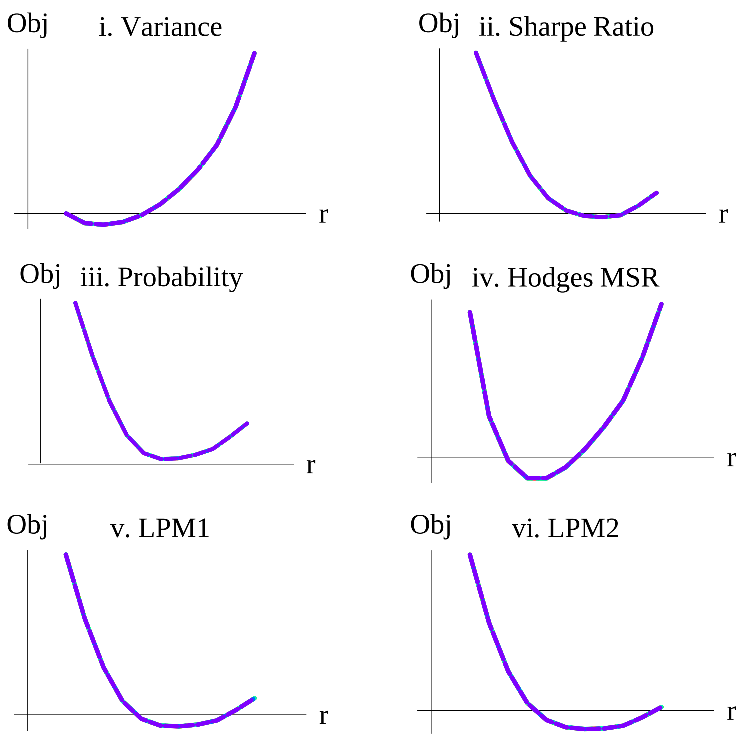

FIGURE 7. Efficient frontiers for different objectives

Definition 4.7.2. LQP(2)

$$\begin{aligned} \min_{\boldsymbol{\theta}} \quad & g_{\boldsymbol{\beta}}(\boldsymbol{\theta}) \\ \text{s.t.} \quad & \boldsymbol{\theta}.\mathbf{1} = 1 \\ & L_i(\boldsymbol{\theta}) \geq q_i \end{aligned} \tag{4.7.7}$$

where $\boldsymbol{\beta}.\mathbf{1} = 1$

Solutions to LQP(2) will be denoted by $\boldsymbol{\theta}^*_{\boldsymbol{\beta},\boldsymbol{q}}$.

The set of solutions for all possible parameter values are denoted for the NLP and LQP respectively as:

$$\begin{aligned} \Theta_{\text{NLP}} &= \{\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}\}_{\boldsymbol{\alpha}\in\mathbb{R}^n,\boldsymbol{\alpha}.\mathbf{1}=1,p\in\mathbb{R}} \\ \Theta_{\text{LQP}} &= \{\boldsymbol{\theta}^*_{\boldsymbol{\beta},\boldsymbol{q}}\}_{\boldsymbol{\beta}\in\mathbb{R}^n,\boldsymbol{\beta}.\mathbf{1}=1,\boldsymbol{q}\in\mathbb{R}^n,\boldsymbol{q}.\mathbf{1}=1} \end{aligned} \tag{4.7.8}$$

Proposition 4.7.3. *All solutions to NLP(1) are solutions of LQP(2). That is to say, there always exist parameter values of $\boldsymbol{\beta}, \boldsymbol{q}$ in LQP(2), such that the solution $\boldsymbol{\theta}^*_{\boldsymbol{\beta},\boldsymbol{q}}$ corresponds to the solution $\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}$ to NLP(1), for all parameter values $\boldsymbol{\alpha}, p$.*

$$\Theta_{\text{NLP}} \subset \Theta_{\text{LQP}} \tag{4.7.9}$$

Proof. The proof is by contradiction. The strategy is to consider a solution of NLP(1) that will necessarily satisfy the constraints of LQP(2), and hypothesize the existence of another portfolio, which also satisfies the constraints of LQP(2), but

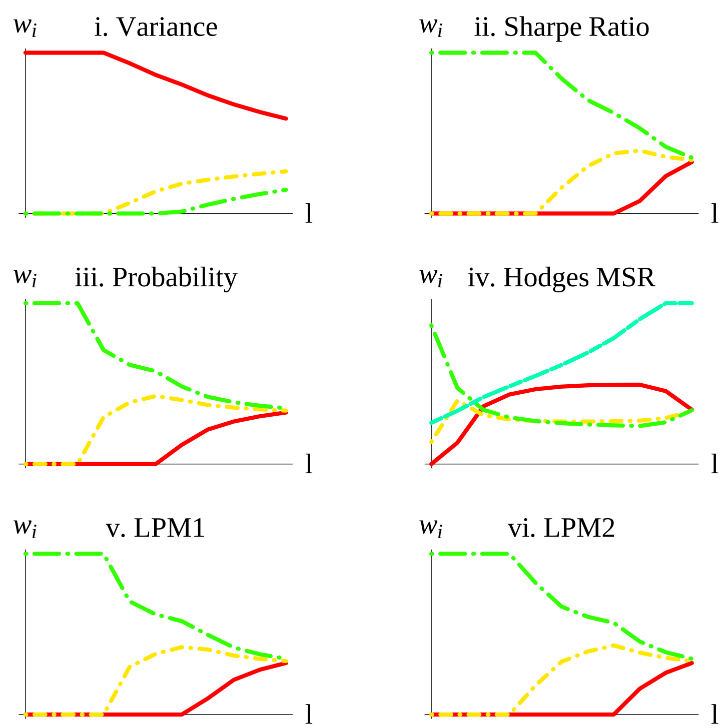

FIGURE 8. Dependence of optimal weights on the mixing parameter

which gives a smaller objective for LQP(2) (i.e., is more optimal) and show that this leads to a contradiction.

Given values for the parameters for NLP(1), choose specific values for the LQP(2) parameters:

$$q_i = L_i(\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}) \tag{4.7.10}$$

$$\beta_i = -\frac{1}{N}\frac{\partial f_{\boldsymbol{\alpha}}}{\partial Q_i}\Big|_{\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}} \tag{4.7.11}$$

$$= -\frac{1}{N}\frac{\partial f_{\boldsymbol{\alpha}}}{\partial x_i}\frac{\partial x_i}{\partial Q_i}\Big|_{\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}} \tag{4.7.12}$$

$$= \frac{1}{2N}\alpha_i \phi(x_i)\frac{x_i}{Q_i}\Big|_{\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}} \tag{4.7.13}$$

where N is a normalisation factor to ensure that $\boldsymbol{\beta}.\mathbf{1} = 1$. For the LQP objective to be convex, we require that $\beta_i \geq 0$, and this holds because $f_{\boldsymbol{\alpha}}(x)$ is increasing in x, whilst $\frac{1}{\sqrt{Q}}$ is decreasing in Q.

Due to the choice of the chosen LQP parameters, $\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}$ satisfies the constraints of the LQP problem constraints. What is less obvious is whether $\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}$ is the optimal solution to the problem.

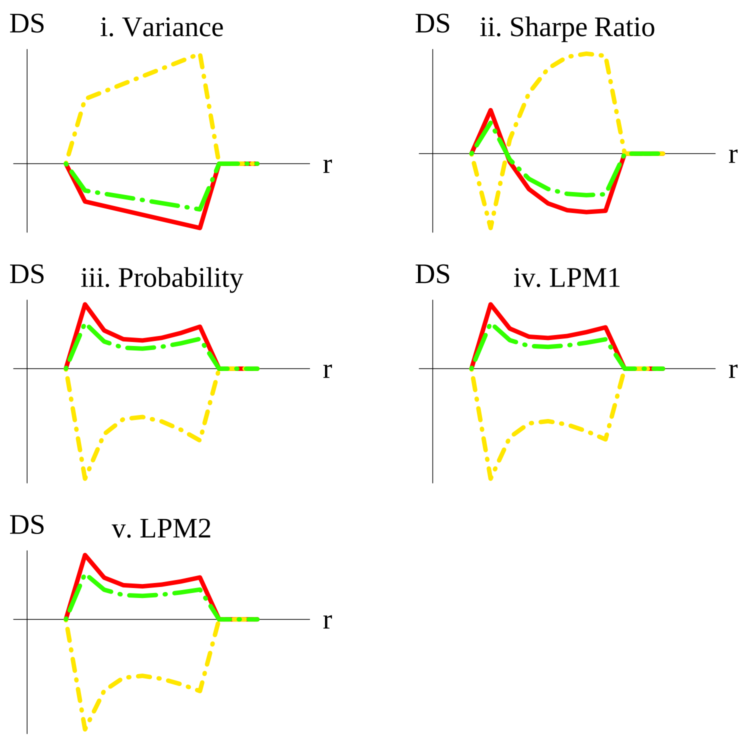

FIGURE 9. Distress sensitivities along the efficient frontier

If $\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}$ is not the solution to LQP(2), then there exists a $\boldsymbol{\theta}^\dagger = \boldsymbol{\theta}^*_{\boldsymbol{\alpha},p} + \boldsymbol{\epsilon}$ that gives a smaller value for the LQP (constrained) objective, $g_{\boldsymbol{\beta}}(\boldsymbol{\theta}^\dagger)$, with parameter values $\boldsymbol{\beta}(\boldsymbol{\alpha}, p)$, (objective parameters) and $\boldsymbol{q}(\boldsymbol{\alpha}, p)$ (constraint parameters), than $\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}$.

$$g_{\boldsymbol{\beta}}(\boldsymbol{\theta}^\dagger) < g_{\boldsymbol{\beta}}(\boldsymbol{\theta}_{\boldsymbol{\alpha},p}) \tag{4.7.14}$$

Taylor's expansion on NLP(1):

$$\begin{aligned} f_{\boldsymbol{\alpha}}(\boldsymbol{\theta}^\dagger) &\approx f_{\boldsymbol{\alpha}}(\boldsymbol{\theta}_{\boldsymbol{\alpha},p}) + \sum_{i,j=1}^{n,m} \left(\frac{\partial f_{\boldsymbol{\alpha}}}{\partial L_i} \frac{\partial L_i}{\partial \theta_j} + \frac{\partial f_{\boldsymbol{\alpha}}}{\partial Q_i} \frac{\partial Q_i}{\partial \theta_j} \right) \Big|_{\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}} \epsilon_j \\ &= f_{\boldsymbol{\alpha}}(\boldsymbol{\theta}_{\boldsymbol{\alpha},p}) + \sum_{i,j=1}^{n,m} \frac{\partial f_{\boldsymbol{\alpha}}}{\partial Q_i} \frac{\partial Q_i}{\partial \theta_j} \Big|_{\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}} \epsilon_j \end{aligned} \tag{4.7.15}$$

The term involving L_i vanishes because $\boldsymbol{\theta}^\dagger$ and $\boldsymbol{\theta}_{\boldsymbol{\alpha},p}$ both satisfy the linear constraints of LQP(2) by assumption and design, respectively. Therefore,

$$\sum_{j=1}^{m} \frac{\partial L_i}{\partial \theta_j} \Big|_{\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}} \epsilon_j = 0. \tag{4.7.16}$$

Similarly Taylor's expansion on LQP(2):

$$
\begin{aligned}
g_{\boldsymbol{\beta}}(\boldsymbol{\theta}^{\dagger}) &\approx g_{\boldsymbol{\beta}}(\boldsymbol{\theta}_{\boldsymbol{\alpha},p}) + \sum_{i,j=1}^{n,m} \frac{\partial g_{\boldsymbol{\beta}}}{\partial Q_i} \frac{\partial Q_i}{\partial \theta_j}\Big|_{\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}} \epsilon_j \\
&= g_{\boldsymbol{\beta}}(\boldsymbol{\theta}_{\boldsymbol{\alpha},p}) + \sum_{i,j=1}^{n,m} \beta_i \frac{\partial Q_i}{\partial \theta_j}\Big|_{\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}} \epsilon_j \\
&= g_{\boldsymbol{\beta}}(\boldsymbol{\theta}_{\boldsymbol{\alpha},p}) - \frac{1}{N} \sum_{i,j=1}^{n,m} \frac{\partial f_{\boldsymbol{\alpha}}}{\partial Q_i} \frac{\partial Q_i}{\partial \theta_j}\Big|_{\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}} \epsilon_j
\end{aligned} \tag{4.7.17}
$$

suggesting that $\sum_{i,j=1}^{n,m} \frac{\partial f_{\boldsymbol{\alpha}}}{\partial Q_i} \frac{\partial Q_i}{\partial \theta_j}\Big|_{\boldsymbol{\theta}^*_{\boldsymbol{\alpha},p}} \epsilon_j > 0$ and therefore that

$$
f_{\boldsymbol{\alpha}}(\boldsymbol{\theta}^{\dagger}) > f_{\boldsymbol{\alpha}}(\boldsymbol{\theta}_{\boldsymbol{\alpha},p}) \tag{4.7.18}
$$

which contradicts the optimality of $\boldsymbol{\theta}_{\boldsymbol{\alpha},p}$ to NLP(1), which was our initial assumption.

□

5. Conclusions

We are certain that the GM approach, namely the assumption of a multivariate finite Gaussian mixture distribution for asset returns used in conjunction with a probability of shortfall objective, will be useful for a whole range of portfolio management applications. It is only slightly harder to implement than standard Markowitz, with many features in common (e.g. we retain the use of covariance matrices). However, it is more flexible because of its ability to handle non-elliptic return distributions. Because it is intuitive, the technique is unlikely to face resistance from practitioners already familiar with mean-variance approaches. With two regimes, the objective function does not possess more than two maxima, so our numerical examples have been robust and quick to solve.

We have compared the GM and mean-variance approaches. The obvious questions to ask are whether the two approach give different optimal weights from one another, and if so, whether holding the GM weights will give improved performance using measures preferred by practitioners. The response to both questions is in the affirmative. The optimal weights are significantly different between the approaches. The optimal weights for the GM approach will by definition better serve an investor seeking to minimize the probability of shortfall in an environment with multiple regimes. Because the GM distribution is a better model for reality than the Gaussian distribution, we believe that the GM approach will do a better job for managing portfolios in the real world.

Acknowledgment. IB would like to thank Mr. Gerry Salkin and Prof. Nicos Christofides of the CQF, Imperial College for funding.

Appendix A. Gaussian mixture distribution

A.1. Definitions.

Definition A.1.1. A (scalar) random variable Z has the univariate GM distribution if its probability density function $f_Z(z)$ is of the form

$$f_Z(z) = \sum_{i=1}^{n} w_i f_{X_i}(z) = \sum_{i=1}^{n} w_i \phi(\frac{z-\mu_i}{\sigma_i}) \tag{A.1.1}$$

where the random variables X_i are normally distributed with normal probability density functions $\phi_{X_i}(x) = \phi(\frac{z-\mu_i}{\sigma_i})$, $\phi(z)$ is the standard normal probability density function and the weights w_i sum to one. The random variables X_i have means μ_i and variances σ_i^2. The finite sum is over the desired number of normal components to combine, n.

Remark A.1.2. The cumulative distribution function is trivially:

$$F_Z(z) = \sum_{i=1}^{n} w_i \Phi(\frac{z-\mu_i}{\sigma_i}) \tag{A.1.2}$$

where $\Phi(z)$ is the standard normal cumulative density function. We will made use of this observation in Section 4.1 when we define the PoS objective.

Similarly,

Definition A.1.3. A vector random variable $\boldsymbol{Z}$ has the multivariate GM distribution if its probability density function $f_{\boldsymbol{Z}}(\boldsymbol{z})$ is of the form

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \sum_{i=1}^{n} w_i f_{\boldsymbol{X}^{(i)}}(\boldsymbol{z}) = \sum_{i=1}^{n} w_i \phi_{\boldsymbol{\mu}^{(i)},\boldsymbol{V}^{(i)}}(\boldsymbol{z}) \tag{A.1.3}$$

where the (vector) random variables $\boldsymbol{X}^{(i)}$ are multivariate normally distributed with probability density functions $\phi_{\boldsymbol{\mu}^{(i)},\boldsymbol{V}^{(i)}}(\boldsymbol{z})$, and the weights w_i sum to one. The vector random variable $\boldsymbol{X}^{(i)}$ has mean $\boldsymbol{\mu}^{(i)}$ and variance-covariance matrix $\boldsymbol{V}^{(i)}$. E.g. if we take the ath and bth components of $\boldsymbol{X}^{(i)}$, their covariance is the element (a, b) of $\boldsymbol{V}^{(i)}$; i.e. $\mathrm{Cov}(X_a^{(i)}, X_b^{(i)}) = V_{ab}^{(i)}$. Also $\mathbb{E}[X_a^{(i)}] = \mu_a^{(i)}$.

Remark A.1.4. Note that by definition $\mathrm{Cov}(X_a^{(i)}, X_b^{(j)}) = 0$ for $i \neq j$.

Remark A.1.5. In the numerical experiments described later, the mixture distribution contains two normal components, describing asset returns under *tranquil* and *distressed* conditions. We shall refer to the weights w_i as *regime weights*.

A.2. Moments.

The mean of a random variable with the mixture distribution is simply expressed as a linear combination of the means of the component normal distributions.

Proposition A.2.1. *The expectation of a function f of a random variable with the GM distribution can be expressed in terms of the expectations of functions of the component normally distributed variables:*

$$\mathbb{E}[f(\boldsymbol{Z})] = \sum_{i=1}^{n} w_i \mathbb{E}[f(\boldsymbol{X}^{(i)})] \tag{A.2.1}$$

Remark A.2.2. In particular,

$$\mathbb{E}[\boldsymbol{Z}] = \sum_{i=1}^{n} w_i \boldsymbol{\mu}^{(i)} \tag{A.2.2}$$

N.B. This is a potential source of confusion given that it is *not* true in general that $\boldsymbol{Z} = \sum_{i=1}^{n} w_i \boldsymbol{X}^{(i)}$.

The variance depends not only on the variances of the components, but also on the differences between the means of the components.

Proposition A.2.3. *The variance of a random variable with the (univariate) GM distribution can be expressed in terms of the expectations and variances of the component normally distributed variables:*

$$\begin{aligned} \mathrm{Var}[Z_a] &= \sum_{i=1}^{n} w_i \mathrm{Var}[X_a^{(i)}] + \sum_{i,j<i}^{n,n} w_i w_j (\mathbb{E}[X_a^{(i)}] - \mathbb{E}[X_a^{(j)}])^2 \\ &= \sum_{i=1}^{n} w_i (\sigma_a^{(i)})^2 + \sum_{i,j<i}^{n,n} w_i w_j (\mu_a^{(i)} - \mu_a^{(j)})^2 \end{aligned} \tag{A.2.3}$$

Remark A.2.4. If we permit the variance and expectation operators to thread over the components of the vector arguments $f(\boldsymbol{X})_a := f(X_a)$, this can be written in alternative vector form as

$$\begin{aligned} \mathrm{Var}[\boldsymbol{Z}] &= \sum_{i=1}^{n} w_i \mathrm{Var}[\boldsymbol{X}^{(i)}] + \sum_{i,j<i}^{n,n} w_i w_j (\mathbb{E}[\boldsymbol{X}^{(i)}] - \mathbb{E}[\boldsymbol{X}^{(j)}])^2 \\ &= \sum_{i=1}^{n} w_i {\boldsymbol{\sigma}^{(i)}}^2 + \sum_{i,j<i}^{n,n} w_i w_j (\boldsymbol{\mu}^{(i)} - \boldsymbol{\mu}^{(j)})^2 \end{aligned} \tag{A.2.4}$$

The variance result, above is a special case of the following result for the covariance:

Proposition A.2.5. *The covariance between two elements of a vector random variable with the (multivariate) GM distribution can be expressed in terms of the expectations of functions of the component normally distributed variables:*

$$\begin{aligned} \mathrm{Cov}[Z_a, Z_b] = \sum_{i=1}^{n} & w_i \mathrm{Cov}[X_a^{(i)}, X_b^{(i)}] + \\ & \sum_{i,j<i}^{n,n} w_i w_j (\mathbb{E}[X_a^{(i)}] - \mathbb{E}[X_a^{(j)}])(\mathbb{E}[X_b^{(i)}] - \mathbb{E}[X_b^{(j)}]) \\ = \sum_{i=1}^{n} & w_i V_{ab}^{(i)} + \sum_{i,j<i}^{n,n} w_i w_j (\mu_a^{(i)} - \mu_a^{(j)})(\mu_b^{(i)} - \mu_b^{(j)}) \end{aligned} \tag{A.2.5}$$

Remark A.2.6. In matrix notation, where $W_{ab} := \mathrm{Cov}[Z_a, Z_b]$,

$$\boldsymbol{W} = \sum_{i=1}^{n} w_i \, \boldsymbol{V}^{(i)} + \sum_{i,j<i}^{n,n} w_i w_j (\boldsymbol{\mu}^{(i)} - \boldsymbol{\mu}^{(j)}).(\boldsymbol{\mu}^{(i)} - \boldsymbol{\mu}^{(j)})' \tag{A.2.6}$$

where $'$ indicates matrix transpose.

A.3. **Linear combinations of random variables with the GM distribution.**

Proposition A.3.1. *Linear combinations of random variables with the (multivariate) GM distribution will themselves have a (univariate) mixture of normals distribution. In particular, the (scalar) random variable* $Y = \sum_{a=1}^{m} \theta_a Z_a$ *where the* m*-vector random variable* $\boldsymbol{Z}$ *has the multivariate GM distribution and* $\boldsymbol{\theta}$ *is an* m*-vector of real coefficients, has probability density function*

$$f_Y(y) = \sum_{i=1}^{n} w_i \phi_{\bar{\mu}_i, \bar{\sigma}_i^2}(y) \tag{A.3.1}$$

where $\bar{\mu}_i = \boldsymbol{\mu}_i.\boldsymbol{\theta}$ *and* $\bar{\sigma}_i^2 = \boldsymbol{\theta}'.\boldsymbol{V}_i.\boldsymbol{\theta}$.

Similar identities may be found in [27].

References

1. Andrew Ang and Geert Bekaert, *How do regimes affect asset allocation?*, Online at http://www.gsb.columbia.edu/faculty/aang/papers/inquire.pdf, April 2002.
2. ———, *International asset allocation with regime shifts*, On-line, April 2002.
3. F. Delbaen J. Eber Artzner, P. and D. Heath, *Thinking coherently*, RISK **10** (1997), no. 11, 68–71.
4. K. Koedijk Campbell, R. and P. Kofman, *Increased correlation in bear markets*, Financial Analyst Journal (2002), 87–94.
5. Lorenzo Capiello and Tom A. Fearnley, *International capm with regime switching garch parameters*, On-line at http://www.fame.ch:8080/research/papers/CapielloFearnley.pdf, July 2000.
6. David Dowe, *David dowe's clustering, mixture modelling and unsupervised learning page*, On-line at http://www.csse.monash.edu.au/ dld/mixture.modelling.page.html, 1999.
7. Straumann D Embrechts P, McNeil AJ, *Correlation: pitfalls and alternatives*, RISK (1999), 69–71.
8. C. R. Harvey Erb, C. B. and T. E. Viskanta, *Forecasting international equity correlations*, Financial Analysts Journal (1994), 32–45.
9. E. Frees and E. Valdez, *Understanding relationships using copula*, North American Actuarial Journal (1998), no. 2, 1–25.
10. McNeil A.J. Nyfeler M. Frey, R., *Copulas and credit models*, RISK (2001), 111–114.
11. A. Graflund and B. Nilsson, *Dynamic portfolio selection: The relevance of switching regimes and investment horizon*, http://swopec.hhs.se/lunewp/abs/lunewp2002_008.htm, March 2002.
12. Markowitz H., *Portfolio selection*, Journal of Finance **7** (1952), 77–71.
13. C.N. Haas, *On modeling correlated random variables in risk assessment*, Risk Analysis **19** (1999), 1205–1214.
14. James Hamilton, *A quasi-bayesian approach to estimating parameters for mixtures of normal distributions*, Journal of Business and Economic Statistics **9** (1991), no. 1, 2739.
15. S Hodges, *A generalization of the sharpe ratio and its application to valuation bounds and risk measures*, University of Warwick Financial Options Research Centre, FORC Pre-Print 98/88, April 1998.
16. John Hull and Alan White, *Value at risk when daily changes in market variables are not normally distributed*, Journal of Derivatives **5** (1998), no. 3, 9–19.
17. Martin Sola John Driffill, Turalay Kenc, *Merton-style option pricing under regime switching*, On-line at http://www.ms.ic.ac.uk/turalay/DriffillKencSola.pdf, January 2002.
18. S.A. Klugman and R. Parsa, *Fitting bivariate loss distribution with copulas*, Insurance : Mathematics and Economics **24** (1999), 139–148.
19. Chiraz Labidi and Thierry An, *Revisiting the finite mixture of gaussian distributions with applications to futures markets*, Computing in Economics and Finance (2000), no. 67.
20. F. Longin and B. Solnik, *Correlation structure of international equity markets during extremely volatile periods*, Journal of Finance **56** (2001), no. 2, 649–676.

21. Dilip B. Madan and Gavin S. McPhail, *Investing in skews*, http://www.rhsmith.umd.edu/Finance/dmadan/hmr.pdf, June 2000.
22. Y. Malevergne and D. Sornette, *Testing the gaussian copula hypothesis for financial assets dependences*, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=291140, November 2001.
23. D. Pelletier, *Regime switching for dynamic correlations*, http://www.crde.umontreal.ca/crde-cirano/pelletier.pdf, March 2002.
24. W.F. Sharpe, *The sharpe ratio*, The Journal of Portfolio Management **Fall** (1994).
25. A. Suárez and S. Carrillo, *Computational tools for the analysis of market risk*, Computing in Economics and Finance (2000), no. 144.
26. Subu Venkataraman, *Value at risk for a mixture of normal distributions: The use of quasi-bayesian estimation techniques*, Economic Perspectives (1997).
27. Jin Wang, *Modeling and generating daily changes in market variables using a multivariate mixture of normal distributions*, On-line at http://www.valdosta.edu/ jwang/paper/MixNormal.pdf, January 2000.
28. Peter Zangari, *An improved methodology for measuring var*, RiskMetrics Monitor **2Q** (1996).

E-mail address: i.buckley@imperial.ac.uk

Centre for Quantitative Finance,, Imperial College,, London,, SW7 2BX,, UK

E-mail address: gustavo@sigmanalysis.com

E-mail address: ben_d@achilles.math.utoronto.ca

Sigmanalysis, Toronto, Canada

E-mail address: seco@math.toronto.edu

RiskLab,, University of Toronto,, Toronto,, Canada