Fully Flexible Extreme Views[1]

Attilio Meucci
attilio_meucci@symmys.com

David Ardia
da@aeris-capital.com

Simon Keel
sk@aeris-capital.com

To appear: Journal of Risk - 2011
This revision: January 22, 2011
Last version available at `http://ssrn.com/abstract=1542083`

Abstract

We extend the Fully Flexible Views generalization of the Black-Litterman approach to effectively handle extreme views on the tails of a distribution. First, we provide a recursive algorithm to process views on the conditional value at risk, which cannot be handled directly by the original implementation of Fully Flexible Views. Second, we represent both the prior and the posterior distribution on a grid, instead of by means of Monte Carlo scenarios: this way it becomes possible to cover parsimoniously even the far tails of the underlying distribution. Documented code is available for download.

JEL Classification: C1, G11

Keywords: Entropy Pooling, Kullback-Leibler, Black-Litterman, VaR, CVaR, grid-probability pair, Monte Carlo, Gauss-Hermite polynomials, Newton-Raphson, kernel estimator

[1] The opinions expressed in this article are the sole responsibility of the authors and do not necessarily reflect the opinions of their employers Kepos Capital LP, aeris CAPITAL AG, or any of their affiliates.

1 Introduction

The combination of subjective views within a broadly accepted risk model is one of the main challenges in quantitative portfolio management. Indeed, any risk model, be it based on historical scenarios, parametric fits, or Monte Carlo scenarios generated according to a given distribution, is subject to estimation risk and thus it is inherently flawed. Therefore, it is important to provide a framework that allows practitioners to overlay their judgement to any risk model in a statistically sound way.

The mainstream approach to perform this task is the celebrated model by Black and Litterman (1990). In this case the distribution of the risk factors is assumed normal and the views are processed using a Bayesian formalism. For non-normal markets, one can choose an ad-hoc parametrization and modify the Black-Litterman approach accordingly. For instance, this is the route chosen in Giacometti, Bertocchi, Rachev, and Fabozzi (2007).

To allow for fully general market distributions and for views on general features of such distributions, one can use the "Fully Flexible Views" approach in Meucci (2008), FFV in the sequel. FFV combines an arbitrary market model, which is referred to as the "prior" and fully general views or stress-tests on the underlying market. The output is a distribution, referred to as the "posterior", which incorporates all the inputs and which can be used for risk management and portfolio optimization.

In FFV, the posterior is obtained by warping the prior distribution so that the views are fulfilled, in such a way that the least possible amount of spurious structure is imposed. Specifically, the posterior distribution minimizes the entropy relative to the prior, which is the natural measure of discrepancy between two distributions. As a final step, opinion pooling assigns different confidence levels to different views and users, leading to a mixture of distributions.

FFV is advantageous in that it allows for full flexibility in the specification of the views: indeed, not only views on expectations, but also views on volatility, or on value at risk (VaR), among others, are possible.

FFV can be implemented in three ways: analytical, non-parametric, and parametric. The analytical solution, which only applies to normally distributed markets, and the non-parametric approach are discussed in the original paper Meucci (2008). The parametric approach is studied in Meucci, Ardia, and Keel (2011).

In the non-parametric approach to FFV, the prior distribution of the generic risk factor, say a return, is represented in terms of a set of scenarios and their associated probabilities. This representation presents two drawbacks when processing extreme views on the tails.

First, this representation allows for views of many sorts, including VaR, but cannot accommodate views on the conditional value at risk (CVaR). Second, the scenarios in the original FFV article were generated by standard Monte Carlo. Monte Carlo scenarios are inadequate when expressing extreme views on the tails of the distribution, unless the number of scenarios is unrealistically large.

This article addresses these two issues of the non-parametric approach. In

Section 2 we review FFV. In Section 3 we discuss a recursive algorithm to cover views on the CVaR. In Section 4 we replace the Monte Carlo scenarios with a deterministic grid and the respective non-equal probabilities in such a way that extreme views can be processed. In Section 5 we present a case study: we model the distribution of a hedge fund return non-parametrically and we process extreme views on its expectation and its CVaR. Fully documented code is available at MATLAB Central File Exchange.

2 Review of fully flexible views

In this section we follow Meucci (2008); please access this reference for more details. Consider a univariate market X, say the return of a hedge fund. We denote by f the probability density function (pdf) of the "prior", i.e. the estimated distribution of X. Suppose that we express generalized views, i.e. a set of scenarios, directional statements, or stress-tests. We denote the views by $\mathcal{V}$. The posterior is a new distribution $\widetilde{f}$ which is as close as possible to the prior f_X but satisfies the views $\mathcal{V}$. More precisely

$$\widetilde{f} \equiv \underset{g \in \mathcal{V}}{\operatorname{argmin}} \left\{ \mathcal{E}\left(g, f\right) \right\}, \tag{1}$$

where the distance is measured in terms of the relative entropy, or Kullback-Leibler divergence

$$\mathcal{E}(\widetilde{f}, f) \equiv \int \widetilde{f}\left(x\right) \ln \frac{\widetilde{f}\left(x\right)}{f\left(x\right)} dx. \tag{2}$$

To implement FFV we represent the distribution f by a set of scenario-probability pairs

$$f \Leftrightarrow \left\{ \left(x_j, p_j\right) \right\}_{j=1,\ldots,J}. \tag{3}$$

The probabilities p_j are determined in such a way that for generic domains D the following approximation holds true

$$\int_D f\left(x\right) dx \approx \sum_{x_j \in D} p_j. \tag{4}$$

We represent the posterior $\widetilde{f}$ as in (3) in terms of the same scenarios, but with different probabilities $\widetilde{f} \Leftrightarrow \{(x_j, \widetilde{p}_j)\}$, where we dropped the running subscripts for notational ease. The discrete version of the relative entropy (2) then reads

$$\mathcal{E}\left(\widetilde{p}, p\right) \equiv \sum_{j=1}^{J} \widetilde{p}_j \ln \frac{\widetilde{p}_j}{p_j}. \tag{5}$$

Denote by g a generic distribution and by q its discrete representation as in (3). Denote with minor abuse of notation by $q \in \mathcal{V}$ the discrete counterpart of the fully flexible constraint $g \in \mathcal{V}$ in (1). Then from (1) we obtain that the discrete posterior satisfies

$$\widetilde{p} = \underset{q \in \mathcal{V}}{\operatorname{argmin}} \left\{ \mathcal{E}\left(q, p\right) \right\}, \tag{6}$$

If the constraints $q \in \mathcal{V}$ are linear, the minimization (6) becomes a linear program that can be easily computed numerically by switching to the dual.

As it turns out, very general views can be expressed as linear constraints on the discrete posterior, including views on expectations, volatilities, correlations, see Meucci (2008).

3 Views on the conditional value at risk

Views on CVaR do not belong to the above list of flexible views that turn into linear constraints on the discrete posterior.

Consider a view on the CVaR

$$\mathcal{V} : \widetilde{\mathrm{E}} \left\{ X | X \leq \widetilde{qu}_{\gamma} \right\} \equiv \widetilde{cv}_{\gamma}. \tag{7}$$

In this expression $\widetilde{cv}_{\gamma}$ is the target $(1-\gamma)$-CVaR of the view, where $\gamma \in (0,1)$; and $\widetilde{qu}_{\gamma}$ denotes the posterior γ-quantile

$$\int_{-\infty}^{\widetilde{qu}_{\gamma}} \widetilde{f}\left(x\right) dx \equiv \gamma. \tag{8}$$

The constraint (7)-(8) cannot be mapped into a linear constraint for q in (6) and thus we cannot compute the posterior directly by linear programming. The reason for this is that if we have a view on the CVaR, we do not know a-priori what the VaR will be. Put differently, the number of scenarios falling below the VaR-quantile is not known in advance. Obviously we are looking for the VaR-level, or equivalently the number of scenarios below the VaR-level, such that the posterior has minimum-entropy distance to the prior. As we proceed to show, we can compute the posterior as a recursion of linear programs.

Assume that we have sorted the scenarios in such a way that $x_1 \leq \cdots \leq x_J$ and rearranged the respective prior discrete probabilities $\{p_j\}$ accordingly. For an arbitrary value $s \in \{1, \ldots, J\}$, define the following J-dimensional vector of discrete probabilities

$$p^{(s)} \equiv \underset{q \in \mathcal{C}_s}{\operatorname{argmin}} \left\{ \mathcal{E}\left(q, p\right) \right\}, \tag{9}$$

where the constraints read

$$\mathcal{C}_s : \left\{ \begin{array}{l} x_1 q_1 + \cdots + x_s q_s \equiv \gamma \cdot \widetilde{cv}_{\gamma} \\ q_1 + \cdots + q_s \equiv \gamma. \end{array} \right. \tag{10}$$

These constraints are linear for the J-dimensional vector $q \equiv (q_1, \ldots, q_s, q_{s+1}, \ldots, q_J)'$ and thus (9) can be solved by switching to the dual as in Meucci (2008). The grid-probabilities pair $\{x_j, p_j^{(s)}\}$ represents a minimum-distortion posterior that displays a VaR equal to $\widetilde{qu}_\gamma$ and a CVaR equal to $\widetilde{cv}_\gamma$.

The posterior (6) that satisfies the CVaR view (7)-(8) is represented by the vector of probabilities among all the $p^{(s)}$ defined in (9) which displays the least distortion overall. In formulas, $\widetilde{p} \equiv p^{(\widetilde{s})}$ where

$$\widetilde{s} \equiv \underset{s\in\{1,\ldots,J\}}{\operatorname{argmin}} \mathcal{E}\left(p^{(s)}, p\right). \tag{11}$$

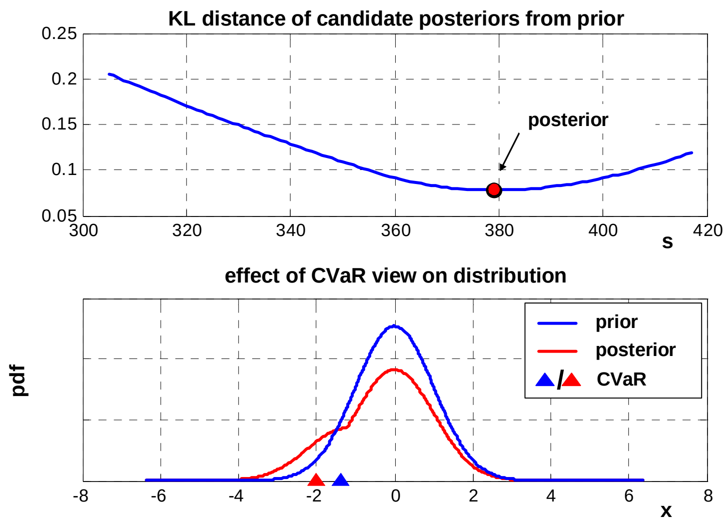

Figure 1: View on CVaR: extensive search of minimum relative-entropy posterior

To illustrate how this recursion works in practice, we consider as prior the standard normal distribution $X \sim \mathrm{N}(0,1)$ with a view on the 80%-CVaR, namely $\widetilde{\mathrm{E}}\{X|X \leq \widetilde{qu}_{0.2}\} \equiv -2$. In the top portion of Figure 1 we display the profile of the Kullback-Leibler distance $\mathcal{E}(p^{(s)}, p)$ from the prior provided by each distribution satisfying (10) as a function of s. The true posterior is represented by the minimum of this profile and is illustrated in the bottom portion of Figure 1.

In practice, the computation of all the terms in (11) is very computationally intensive. However, the relative entropy $\mathcal{E}(p^{(s)}, p)$ is a concave function of s. Therefore we can greatly simplify the process using a discrete counterpart of the Newton-Raphson search method. First we define the empirical derivative of the relative entropy as a function of s

$$D\mathcal{E}\left(p^{(s)}, p\right) \equiv \mathcal{E}\left(p^{(s+1)}, p\right) - \mathcal{E}\left(p^{(s)}, p\right). \tag{12}$$

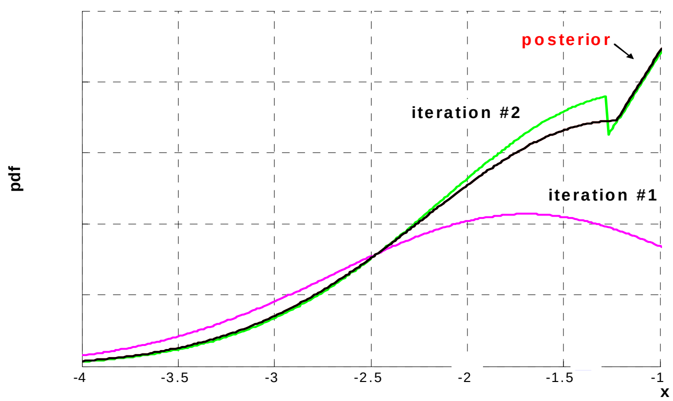

Figure 2: View on CVaR: discrete Newton-Raphson recursion

Next, we initialize a value $\overline{s} \in \{1, \ldots, J\}$ in terms of the prior probabilities $p \equiv (p_1, \ldots, p_s, p_{s+1}, \ldots, p_J)'$ as the threshold that satisfies

$$p_1 + \cdots + p_{\overline{s}} \equiv \gamma. \tag{13}$$

Then we apply the Newton-Raphson method to the empirical derivative, updating and iterating the following two steps until convergence is reached

$$\overline{s} \equiv \mathrm{int}\left(\underline{s} - \frac{D\mathcal{E}\left(p^{(\underline{s})}, p\right)}{D^2\mathcal{E}\left(p^{(\underline{s})}, p\right)}\right) \tag{14}$$

$$\underline{s} \equiv \overline{s}, \tag{15}$$

where $\mathrm{int}(z)$ denotes the integer closest to z and D^2 applies twice the operator D defined in (12). Numerical studies show that convergence is typically reached within a few steps. For instance, we can reproduce the extensive search of Figure 1 with only three recursions, see Figure 2.

4 Grid representation of a distribution

Plain Monte Carlo is the most direct approach to represent a distribution in terms of the scenarios-probabilities pair as specified by (3): the scenarios $\{x_j\}$ are generated randomly according to the desired distribution and the probabilities are set as $p_j \equiv 1/J$. However, this approach concentrates the scenarios on a limited portion of the domain of the distribution. When the views are on the far tails, as is the case with VaR or CVaR relative to confidence levels larger than 95%, plain Monte Carlo fails to offer enough scenarios to cover these views.

Alternative approaches are available to represent the distribution of the market, such as importance sampling or stratified sampling, see Glasserman (2004).

However, these techniques focus on sub-domains of the distribution, and thus, in order to apply such methods, we must forego the full flexibility on the specification of the views.

A different approach, which preserves the generality of the views specification, consists in selecting the scenarios $\{x_j\}$ deterministically as a grid and then use (4) to associate with each of them the suitable probability,

$$p_j \equiv \int_{I_j} f(x)\, dx, \tag{16}$$

where the generic interval I_j contains the j-th point of the grid

$$I_j \equiv \left[x_j - \frac{x_j - x_{j-1}}{2}, x_j + \frac{x_{j+1} - x_j}{2} \right). \tag{17}$$

Once the grid is defined, the entropy optimization (6) can be applied to replace (16) with the new posterior probabilities $\{\widetilde{p}_j\}$ that reflect the views.

The choice of the grid $\{x_j\}$ determines the success of the approach. In particular, it is important that the grid spans a domain large enough to include extreme events that would not be covered by simple Monte Carlo simulations. Also, the grid should be sufficiently sparse and homogeneously spaced. A natural choice for $\{x_j\}$ is the equally spaced grid. More precisely, consider the lower extreme $\underline{x}$ and the upper extreme $\overline{x}$ of the distribution of X, defined as

$$\int_{-\infty}^{\underline{x}} f(x)\, dx \equiv \epsilon \equiv \int_{\overline{x}}^{+\infty} f(x)\, dx, \tag{18}$$

where ϵ is a very small tail probability, such as $\epsilon \approx 10^{-9}$. Then the grid $\{x_j\}$ is determined by the points

$$x_j \equiv \underline{x} + \frac{2j-1}{2}\Delta, \quad j = 1, \ldots, J, \tag{19}$$

where $\Delta \equiv (\overline{x} - \underline{x})/J$ is the constant width of all the intervals (17).

To compute the probabilities (16) that correspond to the grid $\{x_j\}$ we perform numerical integration, see also the case study in Section 5.

An improvement on the equally spaced grid for compact-domain, unimodal distributions draws inspiration from the literature on Gaussian, and in particular Gauss-Hermite, quadratures. First, we introduce the Hermite polynomial of order J

$$H_J(x) \equiv (-1)^J e^{-x^2/2} \frac{d^J e^{-x^2/2}}{dx^J}. \tag{20}$$

As we see in Figure 3, the roots are symmetrical around zero and the growth is sub-linear. Then we construct a suitable grid by shifting and stretching the roots to cover the the extremes $\underline{x}$ and $\overline{x}$ defined in (18).

We test our methodology under the normal assumption with an extreme view on the expectation so that we can benchmark the results with the analytical

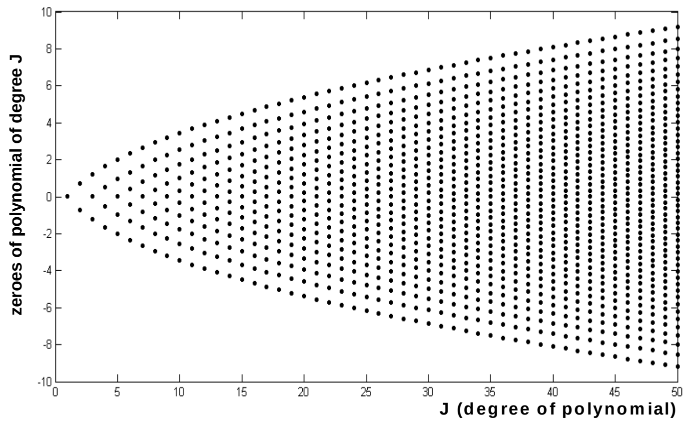

Figure 3: Zeroes of Hermite polynomials as function of polynomial degree

formula in Meucci (2008). In particular, we consider a standard normal prior $X \sim \mathrm{N}(0,1)$ and we overlay the extreme view $\widetilde{\mathrm{E}}\{X\} \equiv -3$, which turns into the linear constraint $\sum_{j=1}^{J} q_j x_j \equiv -3$ in (6). As we appreciate in Figure 4, a parsimonious grid of $J = 10^3$ points performs considerably better than plain Monte Carlo with $J = 10^5$ scenarios.

We emphasize that the grid representation is by no means less flexible than Monte Carlo scenarios. As a matter of fact, it is immediate to generate good quality Monte Carlo scenarios from the posterior distribution as computed on the grid. To do so, first we compute the posterior cumulative distribution function at the grid points

$$\widetilde{F}(x_j) \equiv \textstyle\sum_{s=1}^{j} \widetilde{p}_s. \tag{21}$$

Then we use uniform simulations to obtain the desired posterior scenarios by linear interpolation of (21), without having to invert the cdf, see Meucci (2006).

5 Case study: analysis of hedge fund returns

Here we consider a real world example for modeling and expressing extreme views on the distribution of the return of a hedge fund, for which we collected a time series of $T \equiv 83$ monthly returns, from January 2003 to November 2009.

We denote by $x_1, \ldots, x_T$ the observed time series of the returns realizations and we estimate the prior non-parametrically as the exponentially weighted, kernel-smoothed empirical function. More precisely

$$f(x) \equiv \sum_{t=1}^{T} w_t k_t(x), \tag{22}$$

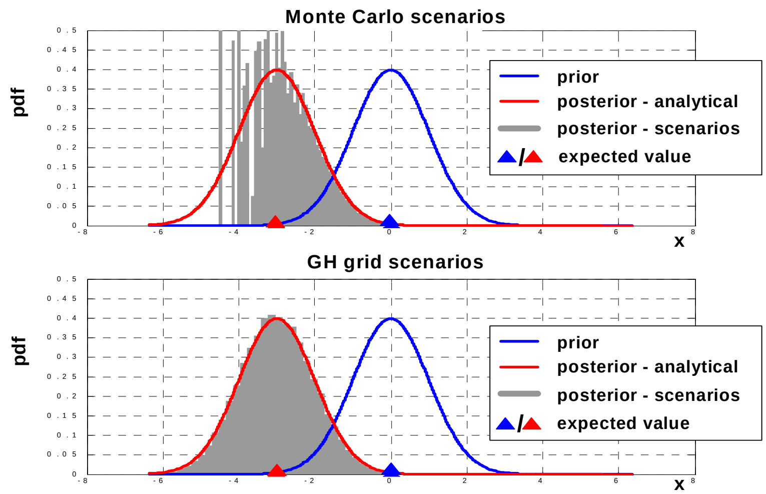

Figure 4: Entropy Pooling on simple Monte Carlo scenarios versus Gauss-Hermite grid

where w_t are exponentially smoothed weights $w_t \propto e^{-\lambda(T-t)}$ normalized to sum to one, and k_t is a Gaussian smoothing kernel of bandwidth h for the t-th observation

$$k_t\left(x\right) \equiv \frac{1}{\sqrt{2\pi h^2}} \exp\left(-\frac{\left(x - x_t\right)^2}{2h^2}\right). \tag{23}$$

We set the exponential decay half-life as half the observation period $\lambda \equiv 2\ln\left(2\right)/T$, and we set the bandwidth as in Silverman (1986)

$$h \equiv \min\left\{\widehat{\sigma}, \frac{\widehat{ir}}{1.34}\right\} T^{-\frac{1}{5}}, \tag{24}$$

where $\widehat{\sigma}$ is the sample standard deviation and $\widehat{ir}$ is the sample interquartile range of the observations.

We represent the pdf (22) on a Gauss-Hermite grid of $J \equiv 10^3$ points. The pdf can be easily integrated analytically to yield the probabilities associated with the grid as in (16). The prior distribution for the monthly fund return features an expected value of 52 basis points and a 95%-confidence CVaR of -5.61%.

We consider an analyst with a bearish outlook for the fund: both the expected value and the CVaR should be 100 basis points less than prescribed by the prior estimate. We process the view on the expected value by applying FFV on the grid as discussed in Section 4 and we process the view on the CVaR through the recursive application of FFV as discussed in Section 3.

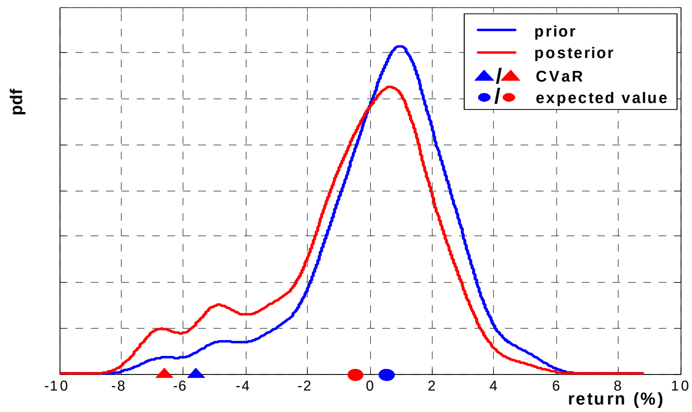

Figure 5: Hedge fund return distribution: non-parametric prior, posterior after views on body and tails

In Figure 5 we display the results. As expected, compared with the prior, in the posterior the negative outcomes are overweighted. However, the posterior distribution is not obtained by simply shifting the prior to the left: all the probability masses are rearranged to reflect the views and the domain of the distribution remains correctly unaltered.

References

Black, F., and R. Litterman, 1990, Asset allocation: combining investor views with market equilibrium, *Goldman Sachs Fixed Income Research*.

Giacometti, M., I. Bertocchi, T.S. Rachev, and F. Fabozzi, 2007, Stable distributions in the Black-Litterman approach to asset allocation, *Quantitative Finance* 7, 423–433.

Glasserman, P., 2004, *Monte Carlo Methods in Financial Engineering* (Springer).

Meucci, A., 2006, Beyond Black-Litterman in practice: A five-step recipe to input views on non-normal markets, *Risk* 19, 114–119 Extended version available at http://ssrn.com/abstract=872577.

——— , 2008, Fully flexible views: Theory and practice, *Risk* 21, 97–102 Extended version available at http://ssrn.com/abstract=1213325.

——— , D. Ardia, and S. Keel, 2011, Fully flexible views with parametric multivariate distributions, *Working Paper* Available at http://ssrn.com/abstract=1742559.

Silverman, B. W., 1986, *Density Estimation for Statistics and Data Analysis* (Chapman and Hall).