Operations Research Letters 43 (2015) 46–51

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Incorporating views on marginal distributions in the calibration of risk models[☆]

Santanu Dey, Sandeep Juneja*, Karthyek R.A. Murthy

Tata Institute of Fundamental Research, Mumbai, India

ARTICLE INFO

Article history:
Received 17 January 2013
Received in revised form
7 November 2014
Accepted 7 November 2014
Available online 15 November 2014

Keywords:
Model selection
I-divergence minimization
Portfolio modeling
Options pricing

ABSTRACT

We apply entropy based ideas to portfolio optimization and options pricing. The known abstracted problem corresponds to finding a probability measure that minimizes relative entropy with respect to a specified measure while satisfying moment constraints on functions of underlying assets. We generalize this to also allow constraints on marginal distribution of functions of underlying assets. These are applied to Markowitz portfolio framework to incorporate fatter tails as well as to options pricing to incorporate implied risk neutral densities on liquid assets.

1. Introduction

Entropy based ideas have found a number of popular applications in finance over the last two decades. A key application involves portfolio optimization where we often have a prior probability model and some independent expert views on the assets involved. If such views are of the form of constraints on moments, entropy based methods are used (see, e.g., Meucci [17]) to arrive at a 'posterior' probability measure that is closest in the sense of minimizing relative entropy or I-divergence to the prior probability model while satisfying those moment constraints. Another important application involves calibrating the risk neutral probability measure used for pricing options (see, e.g., Buchen and Kelly [6], Stutzer [21], Avellaneda et al. [2]). Here, entropy based ideas are used to arrive at a probability measure that correctly prices given liquid options (which are expectations of option payoffs) while again being closest to a specified prior probability measure.

As indicated, in the existing literature the conditions imposed on the posterior measure correspond to constraints on the moments of the underlying random variables. However, the constraints that arise in practice may be more general. For instance, in portfolio optimization settings, an expert may have a view that a certain index of stocks has a fat-tailed t-distribution, and is looking for a posterior joint distribution as a model of stock returns that satisfies this requirement while being closest to a prior model, that may, for instance, be based on historical data.

Similarly, a view on the risk neutral density of a certain financial instrument would also be reasonable if it is heavily traded, e.g., futures contract on a market index, and such views on marginal densities can be used to better price less liquid instruments that are correlated with the heavily traded instrument. There is now a sizable literature that focuses on estimating the implied risk neutral density from the observed option prices of an asset that has a highly liquid options market (see [15] for a comprehensive review). In [12], Figlewski notes that the implied risk neutral density of the US market portfolio, as a whole entity, implicitly captures market's expectations, investors' risk preferences and sensitivity to information releases and events. This is usually not possible with just a finite number of constraints on expected values of payoffs from options. So in the options pricing scenario, views on the posterior measure could include, for example, those on the implied risk neutral density of a security price estimated from certain heavily traded options written on that security. See, for example, Avellaneda [1] for a discussion on the need to use all the available econometric information and stylized market facts to accurately calibrate mathematical models.

Motivated by these considerations, in this paper we devise a methodology to arrive at a posterior probability measure when the constraints on this measure are of a general nature that, apart from moment constraints, include specifications of marginal distributions of functions of underlying random variables as well.

[☆] An extended version of the paper with all the proofs can be accessed at http://arxiv.org/abs/1411.0570.

* Corresponding author.
E-mail addresses: dsantanu@tifr.res.in (S. Dey), juneja@tifr.res.in (S. Juneja), kamurthy@tifr.res.in (K.R.A. Murthy).

http://dx.doi.org/10.1016/j.orl.2014.11.004
0167-6377/

Related literature: The evolving literature on updating models for portfolio optimization to include specified views builds upon the pioneering work of Black and Litterman [4]. They consider variants of Markowitz's model where the subjective views of portfolio managers are used as constraints to update models of the market using ideas from Bayesian analysis. Their work focuses on Gaussian framework with views restricted to linear combinations of expectations of returns from different securities. Since then a number of variations and improvements have been suggested (see, e.g., [18–20]). Earlier, Avellaneda et al. [2] used weighted Monte Carlo methodology to calibrate asset pricing models to market data (also see Glasserman and Yu [13]). Buchen and Kelly in [6] and Stutzer in [21] use the entropy approach to calibrate one-period asset pricing models by selecting a pricing measure that correctly prices a set of benchmark instruments while minimizing *I*-divergence from a prior specified model, that may, for instance be estimated from historical data (see the recent survey article [16]).

Our contributions: As mentioned earlier, we focus on examples related to portfolio optimization and options pricing. It is well known that for views expressed as a finite number of moment constraints, the optimal solution to the *I*-divergence minimization can be characterized as a probability measure obtained by suitably exponentially twisting the original measure; this exponentially twisted measure is known in the literature as the Gibbs measure (see, for instance, [9]). We generalize this to allow cases where the expert views may specify marginal probability distribution of functions of random variables involved. We show that such views, in addition to views on moments of functions of underlying random variables can be easily incorporated. In particular, under technical conditions, we characterize the optimal solution with these general constraints, when the objective is *I*-divergence and show the uniqueness of the resulting optimal probability measure.

As an illustration, we apply our results to portfolio modeling in Markowitz framework where the returns from a finite number of assets have a multivariate Gaussian distribution and expert view is that a certain portfolio of returns is fat-tailed. We show that in the resulting probability measure, under mild conditions, all correlated assets are similarly fat-tailed. Hence, this becomes a reasonable way to incorporate realistic tail behavior in a portfolio of assets. Generally speaking, the proposed approach may be useful in better risk management by building conservative tail views in mathematical models. We also apply our results to price an option which is less liquid and written on a security that is correlated with another heavily traded asset whose risk neutral density is inferred from the options market prices. We conduct numerical experiments on practical examples that validate the proposed methodology.

Organization of the paper: We formulate the model selection problem as an optimization problem in Section 2, and derive the posterior probability model as its solution in Section 3. In Section 4, we apply our results to the portfolio problem in the Markowitz framework and develop explicit expressions for the posterior probability measure. There we also show how a view that a portfolio of assets has a 'regularly varying' fat-tailed distribution renders a similar fat-tailed marginal distribution to all assets correlated to this portfolio. Further, we numerically test our proposed algorithms on practical examples. In Section 5, we illustrate the applicability of the proposed framework in options pricing scenario. All the proofs are presented in [10].

2. The model selection problem

Let the random vectors $\mathbf{X} = (X_1, \ldots, X_m)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ denote the risk factors associated with a prior reference risk model which is specified as a joint probability density $f(\mathbf{x}, \mathbf{y})$ over $\mathbf{X}$ and $\mathbf{Y}$. This model, typically arrived using statistical analysis of historical data, is used for risk analysis (such as calculating expected shortfall, and VaR) or for choosing optimal positions in portfolios. However, the market presents itself with additional information, usually in the form of 'views' of experts (or) current market observables. These views can be simple moment constraints as in,

$$\int_{\mathbf{x},\mathbf{y}} h_i(\mathbf{x}, \mathbf{y}) \mathbb{P}(d\mathbf{x}, d\mathbf{y}) = c_i, \quad i = 1, \ldots, k,$$

(or) as detailed as constraints over marginal densities:

$$\int_{\mathbf{y}} \mathbb{P}(d\mathbf{x}, d\mathbf{y}) = g(\mathbf{x}) \quad \text{for all } \mathbf{x},$$

where $c_i,\ i = 1, \ldots, k$ are constants, $g(\cdot)$ is a given marginal density of $\mathbf{X}$ and $\mathbb{P}(\cdot)$ is the unknown probability measure governing the risk factors. Let $\mathcal{P}(f)$ denote the collection of probability density functions which are absolutely continuous with respect to the density $f(\cdot,\cdot)$ (a density $\tilde{f}(\cdot,\cdot)$ is said to be absolutely continuous with respect to $f(\cdot,\cdot)$ if for almost every x and y such that $f(x, y) = 0$, $\tilde{f}(x, y)$ also equals 0). For any probability density $\tilde{f}(\mathbf{x}, \mathbf{y})$, the relative entropy of $\tilde{f}$ with respect to f (also known as I-divergence or Kullback–Leibler divergence) is defined as

$$D(\tilde{f}\|f) := \begin{cases} \displaystyle\int \log\left(\frac{\tilde{f}(\mathbf{x}, \mathbf{y})}{f(\mathbf{x}, \mathbf{y})}\right) \tilde{f}(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y}, & \text{if } \tilde{f} \in \mathcal{P}(f) \\ \infty, & \text{otherwise.} \end{cases}$$

Though relative entropy $D(\cdot\|\cdot)$ is not a metric, it has been widely used to discriminate between probability measures in the context of model calibration (see [7,6,21,3,1,2,17,16]). Our objective in this paper is to identify a probability model that has minimum relative entropy with respect to the prior model $f(\cdot,\cdot)$ while agreeing with the views on moments of $\mathbf{Y}$ and marginal distribution of $\mathbf{X}$. Formally, the optimization problem $\mathbf{O_1}$ we attempt to solve is:

$$\min_{\tilde{f} \in \mathcal{P}(f)} \int \log\left(\frac{\tilde{f}(\mathbf{x}, \mathbf{y})}{f(\mathbf{x}, \mathbf{y})}\right) \tilde{f}(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y}$$

subject to:

$$\int_{\mathbf{y}} \tilde{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} = g(\mathbf{x}) \quad \text{for all } \mathbf{x}, \quad \text{and} \tag{1a}$$

$$\int_{\mathbf{x},\mathbf{y}} h_i(\mathbf{x}, \mathbf{y}) \tilde{f}(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} = c_i, \quad i = 1, 2, \ldots, k. \tag{1b}$$

3. Solution to the optimization problem $\mathbf{O_1}$

Some notation is needed to proceed further. For any $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_k) \in \mathbb{R}^k$, let

$$f_{\boldsymbol{\lambda}}(\mathbf{y}|\mathbf{x}) := \frac{\exp\left(\sum_{i=1}^k \lambda_i h_i(\mathbf{x}, \mathbf{y})\right) f(\mathbf{y}|\mathbf{x})}{\int_{\mathbf{y}} \exp\left(\sum_{i=1}^k \lambda_i h_i(\mathbf{x}, \mathbf{y})\right) f(\mathbf{y}|\mathbf{x}) d\mathbf{y}} = \frac{\exp\left(\sum_{i=1}^k \lambda_i h_i(\mathbf{x}, \mathbf{y})\right) f(\mathbf{x}, \mathbf{y})}{\int_{\mathbf{y}} \exp\left(\sum_{i=1}^k \lambda_i h_i(\mathbf{x}, \mathbf{y})\right) f(\mathbf{x}, \mathbf{y}) d\mathbf{y}}$$

whenever the denominator exists. Further, let $f_{\boldsymbol{\lambda}}(\mathbf{x}, \mathbf{y}) := f_{\boldsymbol{\lambda}}(\mathbf{y}|\mathbf{x}) \times g(\mathbf{x})$ denote a joint density function of $(\mathbf{X}, \mathbf{Y})$ and $\mathbb{E}_{\boldsymbol{\lambda}}[\cdot]$ denote the expectation under $f_{\boldsymbol{\lambda}}(\cdot,\cdot)$. Let $m_g(\cdot)$ be the measure corresponding to the probability density $g(\cdot)$ on $\mathbb{R}^m$. For a mathematical claim that depends on $\mathbf{x} \in \mathbb{R}^m$, say $S(\mathbf{x})$, we write $S(\mathbf{x})$ for almost all $\mathbf{x}$, with respect to $g(\mathbf{x})d\mathbf{x}$ to mean that $m_g(\{\mathbf{x} : S(\mathbf{x}) \text{ is false}\}) = 0$.