UNIVERSITÄT BIELEFELD

Fakultät für Wirtschaftswissenschaften

IMW

Institut für Mathematische Wirtschaftsforschung Universität Bielefeld

BiGSEM Bielefeld Graduate School of Economics and Management

A parametric approach to the estimation of convex risk functionals based on Wasserstein distance

Max Nendel & Alessandro Sgarabottolo, Bielefeld University

Abstract

We explore a static setting for the assessment of risk in the context of mathematical finance and actuarial science that takes into account model uncertainty in the distribution of a possibly infinite-dimensional risk factor. We allow for perturbations around a baseline model, measured via Wasserstein distance, and we investigate to which extent this form of probabilistic imprecision can be parametrized. The aim is to come up with a convex risk functional that incorporates a safety margin with respect to nonparametric uncertainty and still can be approximated through parametrized models. The particular form of the parametrization allows us to develop a numerical method, based on neural networks, which gives both the value of the risk functional and the optimal perturbation of the reference measure.

1. Introduction

Given a random variable Y on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in a separable Hilbert space H, one is usually interested in expressions of the form

$$\mathbb{E}_{\mathbb{P}}[f(Y)] = \int_H f(y)\,\mu(\mathrm{d}y),$$

where $f\colon H \to \mathbb{R}$ is a continuous loss (in insurance and actuarial science) or payoff function (in mathematical finance) and $\mu = \mathbb{P} \circ Y^{-1}$ is the distribution of Y.

Lack of data or a reliable calibration procedure usually leads to **uncertainty in the distribution of the risk factor**. We measure this uncertainty via Wasserstein distance and study the functional

$$\mathcal{I}(h)f := \sup_{\nu \in \mathcal{P}_p} \left(\int_H f(z)\,\nu(\mathrm{d}z) - \varphi_h(\mathcal{W}_p(\mu, \nu)) \right),$$

where

- $\mathcal{W}_p(\mu, \nu)$ is the Wasserstein distance of order p between two probability measures μ and ν;
- $\mathcal{P}_p$ is the space of probability measures over H with finite p-th moment;
- $\varphi\colon [0, \infty) \to [0, \infty]$ is an increasing penalization function and, for $h > 0$, $\varphi_h(u) = h\varphi(u/h)$.

This formulation allows for a general **nonparametric uncertainty**, considering a very large set of probability measures, where the measure μ represents a *preferred model*. Moreover, the functional gives structure to the *ambiguity set*, penalizing models that are *far* from the reference one, and the scaling in the penalization function allows to control the uncertainty through the parameter h. Via standard representation theorems, this functional represents a **convex risk measure**.

What has to be tackled:

- computation of **Wasserstein distance**;
- the problem corresponds to an **infinite dimensional optimization**.

In the literature on distributionally robust stochastic optimization similar problems are tackled via duality methods ([3]). In [1], the authors focus on the coherent case and look for a linear expansion of the functional around the central point.

2. First order approximation

We aim for a **first order approximation** of the functional $\mathcal{I}(h)$ for $h > 0$, identifying, for asymptotically small levels of uncertainty, **relevant probability measures** for the optimization problem.

2.1. Relevant models

For $\theta\colon H \to H$, we define a *shifted version* of μ, $\mu_\theta \in \mathcal{P}_p$ via

$$\int_H f(z)\,\mu_\theta(\mathrm{d}z) := \int_H f(y + \theta(y))\,\mu(\mathrm{d}y).$$

and, for $\Theta \subseteq L_p(\mu; H)$,

$$\mathcal{I}_\Theta(h)f := \sup_{\theta \in \Theta} \left(\int_H f(y)\,\mu_\theta(\mathrm{d}y) - \varphi_h(\|\theta\|_{L_p(\mu;H)}) \right).$$

For $p > 1$ and for every function f satisfying suitable **growth conditions** ([4]), we show that, if Θ is dense in $L_p(\mu; H)$, then

$$\lim_{h \downarrow 0} \frac{\mathcal{I}(h)f - \mathcal{I}_\Theta(h)f}{h} = 0.$$

3. Neural network approximation

When $H = \mathbb{R}^d$, using universal approximation theorems (cf. [2]), we can compute the quantity $\mathcal{I}_\Theta(h)f$ numerically by modeling the parameter set Θ with **neural networks**.

3.1. A toy model for an earthquake

We consider a one period model for an earthquake where the risk factor is the location of the epicenter. The loss function is modeled in a way that damages decrease as the epicenter moves away from a point of interest, e.g. a city center.

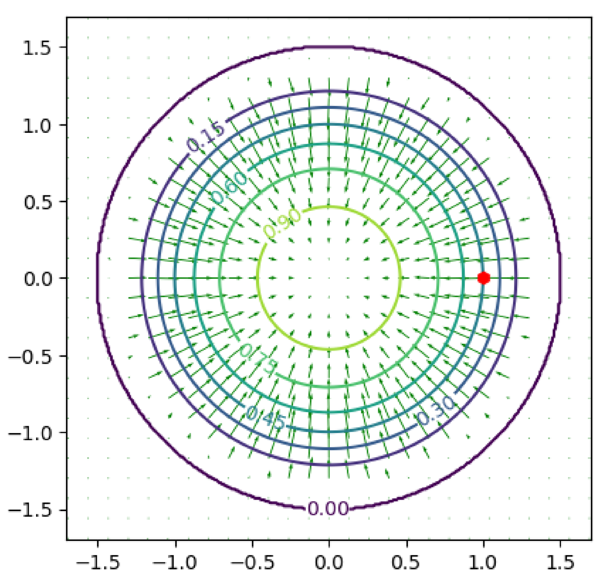

Figure 1: Contour plot of the loss function and vector field θ giving the worst case measure. The reference measure is a gaussian random variable with mean located in the red dot.

Using the neural network approach, we obtain the **worst case loss** and the **worst case measure**, which is given in terms of the optimal shift.

4. Extensions

We can deal with **additional constraints** on the set of models that we consider. In particular, we find approximation results when the measures in the ambiguity set satisfy a **mean** or a **martingale constraint** with respect to the reference model.

4.1. Bounds for option pricing

As an application of the martingale constraint, we compute **model free bounds for option pricing** in an arbitrage-free market.

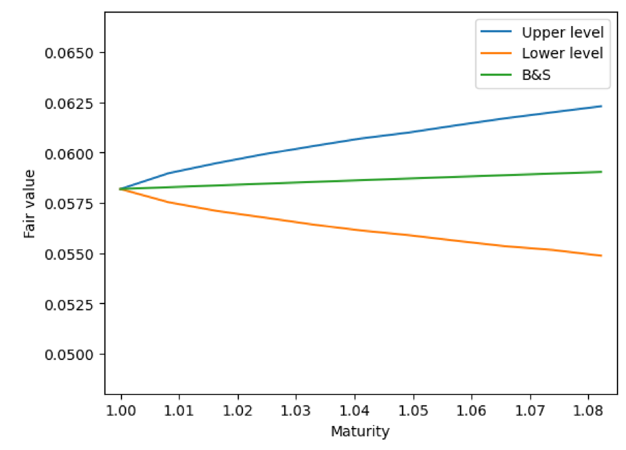

Figure 2: Option: bull spread option with strike long 1 and strike short 1.2. Reference model: Black-Scholes with $S_0 = 1$, $r = 0$, and $\sigma = 20\%$.

References

[1] D. Bartl, S. Drapeau, J. Obłój, and J. Wiesel. Sensitivity analysis of Wasserstein distributionally robust optimization problems. *Proc. R. Soc. A.*, 477(2256):Paper No. 20210176, 19, 2021.

[2] K. Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4(2):251–257, 1991.

[3] P. Mohajerin Esfahani and D. Kuhn. Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. *Math. Program.*, 171(1-2, Ser. A):115–166, 2018.

[4] M. Nendel and A. Sgarabottolo. A parametric approach to the estimation of convex risk functionals based on wasserstein distance. *Preprint*, arXiv.2210.14340, 2022.