

# Notes on Distributionally Robust Optimization for the calculation of Ruin Probabilities

Alex Infanger

March 18, 2018

*All of the following is based on or directly from [1].*

The following assumptions will be made throughout the work.

**Assumption 1.** (A1) The cost function  $c : S \times S \rightarrow \mathbb{R}^+$  is a non-negative lower semicontinuous function satisfying  $c(x, y) = 0$  if and only if  $x = y$ .

**Assumption 2.** (A2) The function  $f : S \rightarrow \mathbb{R}$  whose expectation we are interested in lives in  $L^1(d\mu)$  and is upper semicontinuous.

We will take for granted the “DRO duality” proved in class for the compact  $S$  case (with discussion over the extension). We restate it here.

**Theorem 1.** (*Strong Duality for DRO*) Under the assumptions (A1) and (A2), we have that the primal problem

$$\begin{aligned} \sup \quad & \int f d\nu \\ \text{s.t.} \quad & d_c(\mu, \nu) \leq \delta \end{aligned}$$

is equal to the dual problem,

$$\begin{aligned} \inf \quad & \lambda\delta + \int \phi d\mu \\ \text{s.t.} \quad & \lambda \geq 0 \\ & \phi(x) + \lambda c(x, y) \geq f(y) \text{ for all } x, y \end{aligned}$$

(which upon noting the form of the optimizer can be written as a one dimensional problem)

$$\inf_{\lambda \geq 0} \lambda\delta + \mathbf{E}_\mu \left[ \sup_{y \in S} (f(y) - \lambda c(X, y)) \right].$$

With this strong duality come the complementary slackness conditions,

$$\begin{aligned} \lambda (d_c(\mu, \nu) - \delta) &= 0 \\ \phi(x) &= \sup_y (f(y) - \lambda c(x, y)) \quad \pi^* \text{ a.s.} \end{aligned}$$

Our mission now is to simplify the above when we our function  $f = \mathbf{1}_A$  is an indicator, so that we can calculate ruin probabilities. Let me extend our definition slightly from class,

$$c(x, A) = \begin{cases} \inf\{c(x, y), y \in A\} & x \notin A \\ 0 & \text{otherwise} \end{cases}.$$

We begin by recognizing that

$$\sup_{y \in S} \{\mathbf{1}_A(y) - \lambda c(x, y)\} = (1 - \lambda c(x, A))^+,$$

which is clear casewise. Notice the left hand side is bounded above by one because  $\lambda, c \geq 0$ . Then if  $x \in A$  choose  $y = x$  to achieve that bound. On the other hand if  $x \notin A$  the choice  $y = x$  gets you zero, and the question is whether we can choose  $y \in A$  that is  $c$ -close enough to  $x$  to have  $\mathbf{1}_A(y) - \lambda c(x, y) \geq 0$ . hence we have,

$$\sup_{y \in S} \{\mathbf{1}_A(y) - \lambda c(x, y)\} = \max\{1 - \lambda c(x, A), 0\} = (1 - \lambda c(x, A))^+.$$

Potentially, for the sake of time, skip the build up of theorem 3.

**Theorem 2.** (*Simplified Theorem 3*) Suppose that Assumption (A1) is in force. If  $A$  is a nonempty closed subset of the Polish space  $S$ , and  $\lambda \in [0, \infty)$  attains the infimum in Theorem 1, and the function,

$$h(u) \triangleq \int_{\{x: c(x, A) \leq u\}} c(x, A) d\mu(x)$$

is continuous, we can guarantee

$$\sup \{P(A) : d_c(\mu, P) \leq \delta\} = \mu \left\{ x : c(x, A) \leq \frac{1}{\lambda^*} \right\}.$$

## An Example

We work here with the Cramer-Lundberg model for insurance claims. This is a continuous time stochastic process, where the amount of money in the bank  $R(t)$  satisfies

$$R(t) = u + ct - \sum_{i=1}^{N_t} X_i.$$

Here  $u$  is the initial money in the bank,  $c$  is the premium rate, and the  $X_i$  are claim sizes. They come in at a rate that is a Poisson process,  $N_t$ . You have to assume the  $X_i$  are distributed in some way. Let this distribution have first and second moments  $m_1$  and  $m_2$  respectively. We can assume  $m_1 < \infty$  else the insurance company would have no point insuring such a risk. Just for the insurance company to make even, the amount of money coming in has to be the amount of money being paid out on average, which corresponds to the choice

$$c = \nu m_1$$

where  $\nu$  is the rate at which policies are sold. To make money, the firm adds a safety loading – the so-called risk premium – so that,

$$c = (1 + \eta)\nu m_1.$$

Hence our final model is of the form,

$$R(t) = u + (1 + \eta)\nu m_1 t - \sum_{i=1}^{N_t} X_i.$$

We are interested in the probability of ruin

$$\psi(u, T) \triangleq \mathbf{Pr} \left[ \inf_{t \in [0, T]} R(t) \leq 0 \right].$$

It turns out that it is computationally intractable to solve the above so we instead work with the following Brownian motion approximation,

$$\begin{aligned} R_B(t) &\triangleq u + (1 + \eta)\nu m_1 t - (\nu m_1 t + \sqrt{\nu m_2} B(t)) \\ &= u + \eta \nu m_1 t - \sqrt{\nu m_2} B(t). \end{aligned}$$

Our mission is to robustify this estimate by allowing for some non-trivial movement away from the Brownian motion. For this purpose, we identify the Polish space where the stochastic processes of interest live as the *Skorokhod space*,

$$S = D([0, T], \mathbb{R})$$

the space of real valued right-continuous functions with left limits equipped with the  $J_1$  metric.

## A Quick Digression: The $J_1$ Metric

The  $J_1$  metric is supposed to be like a sup metric that is robust against time shifts that go to zero. The formal definition of the  $J_1$  metric is, if we let  $\Lambda$  be the set of strictly increasing functions  $\lambda : [0, T] \rightarrow [0, T]$  such that  $\lambda \in \Lambda$  implies  $\lambda, \lambda^{-1}$  are continuous up to a set of measure zero, and let  $e$  be the identity map on  $[0, T]$ . Then we define

$$d_{J_1}(x_1, x_2) \triangleq \inf_{\lambda \in \Lambda} \{ \|x_1 \circ \lambda - x_2\|_\infty \vee \|\lambda - e\|_\infty \} \quad (1)$$

where  $\|\cdot\|_\infty$  is the sup norm and  $\vee$  is the max function.

Intuitively, this is suppose to capture the sup norm except you're allowed to perturb the function input a litte bit. So it's a bit weaker than the sup norm. In fact, plugging in  $\lambda = e$  in Equation 1 tells us that everything is closer together with the  $J_1$  metric compared to the sup metric. Take for example when  $T = 1$ ,

$$x_n = (1 + n^{-1})\mathbf{1}_{\{[\frac{1}{2} + \frac{1}{n}, 1]\}}, \quad x = \mathbf{1}_{\{[\frac{1}{2}, 1]\}}.$$

Then  $\|x_n - x\|_\infty \geq 1$  for all  $n$  but  $d_{J_1}(x_n, x) \rightarrow 0$  as  $n \rightarrow \infty$ . To show this notice that,

$$\begin{aligned} d_{J_1}(x_n, x) &= \inf_{\lambda \in \Lambda} \{ \|x_n \circ \lambda - x\|_\infty \vee \|\lambda - e\|_\infty \} \\ &\leq \max \{ \|x_n \circ \lambda_n - x\|_\infty, \|\lambda_n - e\|_\infty \}, \end{aligned}$$

for some feasible choices  $\lambda_n$ . Let us choose

$$\lambda_n = \begin{cases} x & x \in [0, \frac{1}{2}] \\ (1 - \frac{2}{n})x + \frac{2}{n} & x \in [\frac{1}{2}, 1], \end{cases}$$

which is effectively the identity from  $[0, \frac{1}{2}]$  and then jumps to  $\frac{1}{2} + \frac{1}{n}$  and then is the line which connects  $(\frac{1}{2}, \frac{1}{2} + \frac{1}{n})$  to  $(1, 1)$ . Notice now that the difference between  $x_n \circ \lambda_n$  and  $x$  now only sees contribution from the part of the domain  $\frac{1}{2} + \frac{1}{n}$ , so that,

$$\|x_n \circ \lambda_n - x\|_\infty = \frac{1}{n}.$$

On the other hand we have that the identity map and  $\lambda$  differ the most at the point  $x = \frac{1}{2}$  where they differ by  $\frac{1}{n}$ , so that we conclude,

$$\begin{aligned} d_{J_1}(x_1, x_2) &= \inf_{\lambda \in \Lambda} \{ \|x_n \circ \lambda\|_\infty \vee \|\lambda - e\|_\infty \} \\ &\leq \max \{ \|x_n \circ \lambda_n\|_\infty, \|\lambda_n - e\|_\infty \} \\ &\leq \frac{1}{n} \rightarrow 0. \end{aligned}$$

## Back to the Process of Robustification

We intend to invoke Theorem 2 so that we can show that our robust ruin calculation simplifies to a standard Brownian motion ruin calculation with a different initial money-in-the-bank  $u$ . The set of interest is the ruin set,

$$A_u \triangleq \left\{ x \in S : \sup_{t \in [0, T]} (\sqrt{\nu m_2} x(t) - \nu m_1 t) \geq u \right\}.$$

In order to invoke Theorem 2, we must show that the above is closed. Before that we will want the following simple proposition,

**Proposition 1.** I may be mistaken, but I do not believe the following is true:

$$\sup_{t \in [0, T]} y(t) \leq \sup_{t \in [0, T]} x(t) + d_{J_1}(x, y).$$

*Proof.* Let  $T = 1$  and  $y = \mathbf{1}_{\{[0, \frac{1}{2}]\}}$  and  $x = 0$ . Then  $\sup_{t \in [0, 1]} y = 1$  and  $\sup_{t \in [0, 1]} x = 0$  but for the choice  $\lambda = \frac{1}{2}t + \frac{1}{2}$  we have  $\|\lambda - e\|_\infty = \frac{1}{2}$  and  $y \circ \lambda = x$  so that  $d_{J_1}(x, y) \leq \frac{1}{2}$ . We conclude that,

$$\sup_{t \in [0, T]} y(t) = 1 > 0 + \frac{1}{2} = \sup_{t \in [0, T]} x(t) + d_{J_1}(x, y).$$

□

In the Appendix, we consider first the simpler set

$$A_u = \left\{ x \in D([0, T], \mathbb{R}) : \sup_{t \in [0, 1]} x(t) \geq u \right\},$$

and attempt to prove  $A_u$  is closed. I claim however that the example in Proposition 1 lends us a way to show that  $A_u$  does not contain all of its limit points. See the below proposition.

**Proposition 2.**  $A_u$  is not closed.

*Proof.* Consider the choice

$$y_n = \mathbf{1}_{\{[0, \frac{1}{n}]\}}$$

where we notice  $y_n \in A_{\frac{1}{2}}$  for all  $n$ . On the other hand, we have that for the choice  $\lambda_n(t) = \frac{1}{n} + (1 - \frac{1}{n})t$  that  $y_n \circ \lambda_n = 0$  for all  $n$  and  $\|\lambda_n - e\|_\infty = \frac{1}{n} \rightarrow 0$  as  $n \rightarrow \infty$ . We conclude that  $d_{J_1}(y_n, 0) \rightarrow 0$ . But  $0 \notin A_{\frac{1}{2}}$  and therefore  $A_{\frac{1}{2}}$  does not contain one of its limit points. The above proof can be modified by a scaling factor to prove for any  $u$ , that  $A_u$  is not closed. □

## Numerical Example

In the numerical example from Jose Blanchet and Karthyek Murthy, they use the Pareto distribution for the claim sizes, with  $\alpha = 2.2$ .

## Questions and things still to solve

1. I am worried that I might have implemented that Khosnevisan embedding incorrectly for the following reason: I have to simulate more Brownian motion time steps than Poisson timesteps. This stems from the fact that

$$\sigma(t) \triangleq \inf\{s : A(s) = m_1 t\}.$$

But  $A(t) = N(t) + S(t)$  is effectively the sup of a Brownian motion. Because of this it goes like  $\sqrt{t} = o(t)$  such that the above equality (or in the computer it is a  $\geq$ ) exists only for  $t$  less than that simulated.

2. How would I implement the large deviations simulation?
3. The uncountable inf of lower semi-continuous is not necessarily lower semicontinuous. How do I show that for our case,

$$h(u) = \int_{\{x: c(x, A) \leq u\}} c(x, A) d\mu(x)$$

is continuous? This is a stochastic integral in our case.

4. I seemed to have disproven the fact that  $A_u$  is closed; where is my mistake or is this a typo?

## References

- [1] Jose Blanchet, Karthyek R.A. Murthy. “Quantifying Distributional Model Risk via Optimal Transport.” <https://arxiv.org/abs/1604.01446>.