Robust Methods in Portfolio Theory
by Lucia Petrusova
Referee Report

This thesis focuses on the recent developments in the robust methods in portfolio theory. The arising portfolio optimization problems are studied in the relationship with the lower partial moments, the value at risk, or the conditional value at risk that are used as risk measures. A more recent method based on the worst case conditional value at risk is proposed.

The theoretical part of the thesis represents mostly a survey of the existing literature. The thesis heavily builds on the paper of Zhu and Fukushima: Worst Case Conditional Value at Risk with Application to Robust Portfolio Management, Operations Research 57: 1155–1168 to the point that the major results such as theorems or propositions that appear in the thesis are verbatim copies of their counterparts that appear in the above mentioned paper. The results are properly referenced, so it is not of a concern of academic ethics, but at the same time there is no attempt of the author of the thesis to contribute to the existing portfolio optimization theory. So the value of this part is that the author has learned these techniques.

The main contribution of the thesis is in numerical applications. Three examples are given:

1. CVaR approach with 10 most expensive stocks on NYSE,
2. WCVaR under box uncertainty on the same data,
3. WCVaR under mixture distribution on 5 financial stocks.

The conclusions are in line with the prior expectation, more specifically that the requirement on higher expected return leads to more risky portfolio and that the proposed robust approach performs better with increasing uncertainty.

In general, the financial data are dominated by noise and thus any estimation of drift parameters is a statistical trickery. For instance, if we pick a random stock from the data set such as SEB from Table 4.1, the 95% confidence interval for the daily drift is approximately

$$\mu \pm 2\sigma = 0.000559 \pm 2\sqrt{0.000526} = 0.000559 \pm 2 \cdot 0.02293 = [-0.04531, 0.04642],$$

so in particular one cannot be even sure about the sign of the drift, not to mention to get any idea about the specific value. A daily move of more than $\pm 4\%$ seems to be perfectly fine for that stock. The signal to noise ratio $\frac{\mu}{\sigma}$ is only 0.0243, which is statistically indistinguishable from zero. The situation becomes slightly better if we look at the longer term drift, in this case a 5 year drift. The drift scales linearly with the number of observations, the standard deviation scales as a square root, so the 5 year confidence interval is

$$\mu n \pm 2\sigma\sqrt{n} = 0.000559 \cdot 1257 \pm 2\sqrt{0.000526 \cdot 1257} = 0.70266 \pm 2 \cdot 0.81313 = [-0.9235, 2.3289].$$

Obviously a zero drift is statistically not far from the estimated mean value of 0.70266, so one cannot be sure about the sign of the drift on such a long observational horizon as well. Other stocks in the market suffer from the same problem. On a longer horizon, one can start to detect statistically significant drift from zero, but the estimated drift tends to copy the general GDP growth. This is not a problem of the author of this thesis, but one should take all the theoretical results and the subsequent conclusions with some reservations as the parameters that enter these models are not obtainable from the data with the required precision.

A minor comment regarding the data set used in Section 4.1 – the most expensive stocks do not represent a natural pick as there is no deep logic about how many stocks a given company has, which determines the value of a single stock. A small company with a small number of stocks can easily have a more expensive stock than the largest companies on the market if they have a large number of stocks. Thus the pick of the author is somewhat puzzling, a more natural choice would be to take the companies with the largest market capitalization. This is just a comment that I find it somewhat odd.

In conclusion, the author has summarized the existing literature and applied it to specific numerical examples. This is perhaps what is expected from this kind of thesis and it can be recommended as successfully defendable. However, there is no effort to build up on the theoretical model and thus the work lacks innovation that would make it more outstanding in comparison to other theses. I would rank it as an average work at best. The thesis is written in a solid English, so that can count as a plus.

August 31, 2016

Jan Vecer
Department of Probability and Mathematical Statistics
MFF UK
Sokolovska 83
18675 Praha 8