APPLICATION OF ASSET LIABILITY MANAGEMENT ON LIABILITY MODEL BASED ON CLUSTER ANALYSIS

Jan Fojtík – Jiří Procházka – Pavel Zimmermann – Markéta Švehláková – Simona Macková

Abstract

Proper and accurate modelling of liabilities and assets according to the current market condition is an important actuarial task which is mandatory by Solvency II or IFRS. The proper liability modelling requires a testing of the sensitivity of results of hundreds-thousands policies on thousands of scenarios which is extremely demanding on computer time. In our previous research, we proved that cluster analysis can be applied to decrease the computational time while preserving the accuracy of the life insurance liability estimates. The goal of this paper is to develop a method reducing computational time, based on the cluster analysis, for the purpose of dynamic asset liability management tasks. These tasks focus on asset and liability mismatch stemming from different sensitivity on different risk factors such as the interest rate, mortality rate or lapse rate. Faster estimates can be used to deliver more stress tests and sensitivity tests, which allow actuaries to understand the risks of asset and liability mismatch in higher detail. In this article, we focus on the sensitivity of values of liabilities and assets on the change in interest rates.

Key words: life insurance liability modelling, asset liability management, cluster analysis

JEL Code: C38, C63, G22

Introduction

New regulation of Solvency II or IFRS have further increased the importance of responsible administration of the market values of assets and liabilities of insurance contracts. The common method used for this administration is asset liability management (ALM) (Fernandez, 2018). The scope of ALM techniques is to manage investment strategies to follow regulatory or competitive goals. The usual goal of ALM modelling is the maximization of investment returns while minimizing the reinvestment risk. This goal can be reached by setting "*optimal*" investment strategy. ALM methods designed for optimizing investment strategies are based on cash flow matching (Xidonas, 2016), duration immunization (Shang, 2016) or analysis of value at risk (Balestreri, 2011).

The important role of ALM takes place especially in life insurance business where most of the products include profit participation (Aas, 2018), therefore the value of the insurer's liability may depend on the market or accounting value of invested assets and assets' returns. Actuarial models of expected liabilities or some risk measures of the liability value are therefore generally based on extensive Monte Carlo simulation techniques. Commonly, the liability modelling is based on cash flow projection of each contract (per-policy modelling), which is time demanding. Therefore, a faster method for liability modeling is required for testing a high number of scenarios. Let's suppose, the calculation of one ALM scenario on the mid-sized portfolio (300 000 model-points) lasts about 10 minutes. The simulation of thousand scenarios then lasts 10 000 minutes which is almost a week. Results derived with such a delay are outdated, especially when testing different investments strategies because they do not reflect current market position. The time aspect of liability modeling is crucial for the actuality of ALM models. An algorithm based on cluster analysis was previously introduced in (Fojtík, 2017) and (Freedman, 2008), as a good approximate method that reduces significantly computational time of the liability models. The primary focus of this paper is to implement these ideas for simultaneous asset and liability modeling, which is essential for dynamic asset liability management. The results will be presented on the simulated distribution of different investment strategies.

1 Liability model

In this part, we introduce differences between traditional liability model based on per-policy cash flow projection and faster liability model based on cluster analysis.

1.1 Liability modelling using per-policy cash flow model

The traditional approach is based on cash flow projection of each contract separately and then the results are cumulated for the whole portfolio. Cash flow model can be represented by the following formula:

$$CF_t = Prem_t - Surr_t - Death_t - Mat_t - Comm_t - Exp_t + Inv_t\,, \qquad (1)$$

where $Prem_t$ stands for expected premium at the beginning of the period t, $Surr_t$ stands for expected value of surrenders, $Death_t$ stands for expected value of deaths outgo, Mat_t stands for expected value of maturities, $Comm_t$ stands for expected value of commissions, Exp_t stands for expected value of expenses at the end of the period t and Inv_t stands for investment income in time t. The investment income is then split to guaranteed investment income (usually with flat interest) and profit share. The profit share is given by a multiple of an excess of assets return over the guaranteed return.

Let's assume that liability estimate of the whole portfolio calculated by traditional per-policy approach lasts time T_{CF}. The actuaries usually need to calculate a thousand scenarios to obtain full information about insurance portfolio. The total calculation time T of $N_{scenarios}$ calculated by traditional approach is then given by the formula:

$$T = T_{CF} \cdot N_{scenarios} \tag{2}$$

1.2 Liability modelling using cluster analysis

The clustering approach is based on cluster analysis which is a technique used for grouping objects into clusters in such a way that entities from the same cluster are alike and dissimilar to entities in other clusters. All objects belonging to the same cluster can be characterized using a single object and thus the portfolio can be reduced to the number of clusters. This significantly reduces the computational time of all calculations performed on the dataset.

The similarities between model points are measured with respect to a certain set of clustering variables. The best choice of clustering variables is the variables whose values are intended to be reproduced with the compressed model. In this case, the best choice of clustering variables seems to be metrics of economic profit such as present value of profit and loss (PVPL), future cash flow (PVFC), distributive earnings (PVDE), premium (PVP) or values of individual cash flow projections (CF). Since the objective is to model liabilities development rather than nominal values, the clustering variables need to be adjusted to their relative values using

$$R_{i,k} = \frac{100 \cdot \left(X_{i,k} - V_i\right)}{V_i}, \tag{3}$$

where $R_{i,k}$ represents the adjusted value of the k^{th} variable of the i^{th} model point, $X_{i,k}$ is the non-adjusted value and V_i is the reference variable. As a reference variable PVFC or PVPL is usually chosen.

Similarities between model points are measured using Euclidian distance measure. The distance between the i^{th} and the j^{th} model point is defined as

$$d\left(MP_i, MP_j\right) = \sqrt{\sum_{k=1}^{K}\left(X_{k,i} - X_{k,j}\right)^2}, \tag{4}$$

where $X_{k,i}$ is the value of the k^{th} clustering variable.

Once clustering variables and distance measure have been determined, one can perform a clustering algorithm and group the objects into clusters. The algorithm used in this paper is CLARA algorithm (Dmitriev, 2018). This method represents each cluster with one of the model points from the cluster known as a medoid. Objects selected as medoids are model points with a minimum average distance to other members of the relevant cluster. In the case of a large portfolio, the process of finding the medoids may be rather time-consuming which is why CLARA algorithm utilizes the technique of sampling. It selects a random sample from the data set and it only searches medoids in the sample. The whole portfolio is then clustered using the medoids retrieved from the random sample. The process is repeated a pre-specified number of times and in every iteration, a better choice of a medoid is searched in each cluster. After CLARA procedure each cluster is revised and better representant is searched in order to reduce approximation error. The idea behind this second step is that the CLARA procedure selects as medoid model point with the lowest total distance within the cluster, but if there is a very large model point whose metrics of economic profit are much bigger than the others, it is better to select this very large model point even when his sum of distance within the cluster is not the lowest.

Finally, a representative portfolio can be created using the medoids weighted by the sum of reference variable over each model point within the cluster. The representative portfolio can be used for all calculations in place of the original portfolio.

Our previous research has shown that both the computational time and accuracy increase with the increasing number of clusters. Using 500 clusters, one can reduce the computational time of one scenario more than 100 times with the average error of 0.02%[1]. Therefore, it could be concluded that clustering enables one to gain quite accurate results with a significant reduction in the computational time. The total calculation time T of $N_{scenarios}$ calculated by clustering approach is then given by the formula:

$$T = T_{CL} + T_{CF} \cdot \frac{N_{clusters}}{N_{modelpoints}} \cdot N_{scenarios}, \qquad (5)$$

where $N_{modelpoints}$ in the size of the portfolio (number of contracts) and T_{CL} is clustering time.

The clustering algorithm was programmed in R using R package *Cluster* (Maechler, et al. 2017).

2 Asset model

[1] Computer used in this application is i5-6500 CPU @ 3.20 GHz with 4 GB RAM running Windows10.

In this part, we present basic asset model of the insurance company. Assets commonly used in the insurance business to cover the value of liability are mainly bonds because of the low-risk profile. For the purpose of this paper, the asset portfolio will consist of fixed bonds. In the next sub-chapter, we introduce a general method for bond valuation, calculation of total asset value and return realized from these assets.

2.1 Bond valuation

The general principle of bond valuation is based on the principle of the present value of future expected cash flows. The general formula for bond valuation (Cipra, 2010) has the following form:

$$P = \sum_{t=1}^{T} \frac{CF_t}{(1+r)^t} + \frac{F}{(1+r)^T}, \qquad (6)$$

where r is a spot rate with a tenor in time t. CF_t is a coupon at time t and F is a face value at maturity time ($t = T$). The total value of asset portfolio in time t is given by a sum of present values of all bonds in time t.

The important part of bond valuation is a construction of yield curve. To simulate yield curves one of the most popular models was used, Vasicek model (Vasicek, 1977) given by the formula:

$$dr_t = a(b - r_t)dt + \sigma dW_t. \qquad (7)$$

Vasicek model has four parameters determining the range of modeled yield curve or it's variability. It has a mean-reverting property. The drift parameter $a(b\text{-}r_t)$ represents pulling effect from initial rate r_0 at the time 0 towards to long-term rate b. The speed of the increase is given by parameter a and volatility of the process by parameter σ. The W_t is a Wiener process modeling random market factor.

2.2 Return from asset model

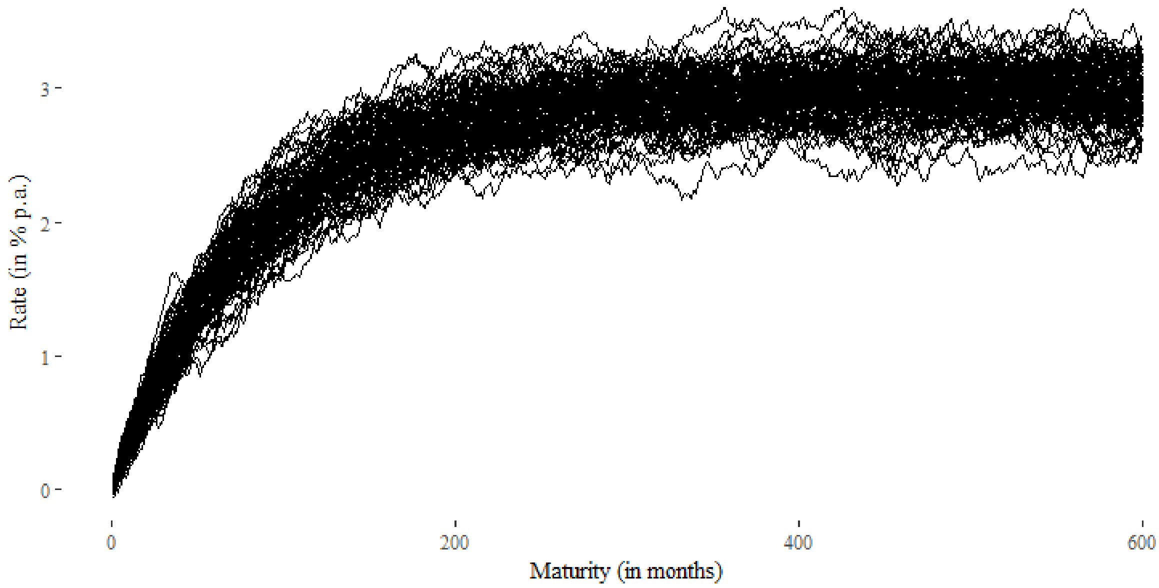
In general, the calculation of assets return depends on several factors such as the type of asset or accounting scheme. In this paper we focus only on fixed bonds, therefore the assets return in time t is given by a sum of all cash flows from asset portfolio (coupons or face value if the bond is at maturity). This asset return will be applied to liability model in profit sharing scheme.

3 Results

In results, we present application of different investment strategies. For the purpose of this analysis, we use 100 simulations of the yield curve for each investment strategy to describe the distribution of the assets and liability values. Because the assets portfolio consists only of fixed bonds, each investment strategy will differ in the maturity of the bonds. The studied strategies have maturity 5 years, 10 years, 15 years and 20 years.

The results will be presented on liability portfolio (model-points) taken from our previous research (Fojtík, 2017), where we were able to reduce bigger insurance portfolio of 100 000 policies into smaller "reference" portfolio consisted of 500 policies (model-points). This reduction of portfolio size leads to a significant speedup of liability modeling and preserving the accuracy of the life liability estimates. The assets portfolio consists of fixed bonds, where each year one new bond is purchased. The yield curve for bond valuation is calculated by Vasicek model with following parameters: $a = 0.001$, $b = 0.03$, $r_0 = 0$, $\sigma = 0.00009$. The random market factor W_t is generated by normal distribution with zero mean and unit variance and has independent variance. A hundred simulations of the yield curve can be seen in Figure 1.

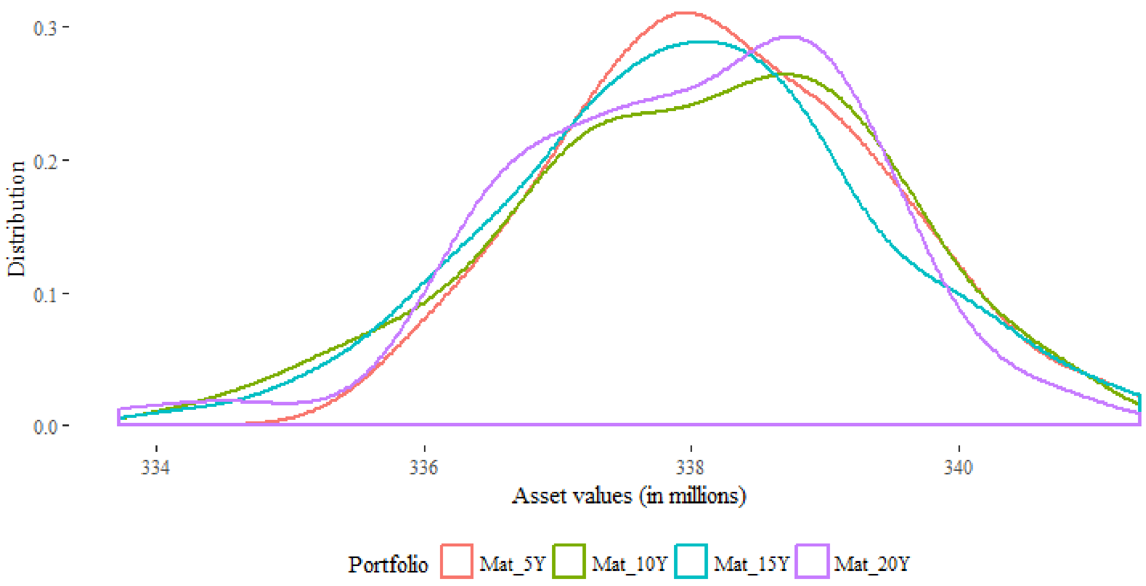
Fig. 1: Simulation of 100 yield curves



Source: Authors work

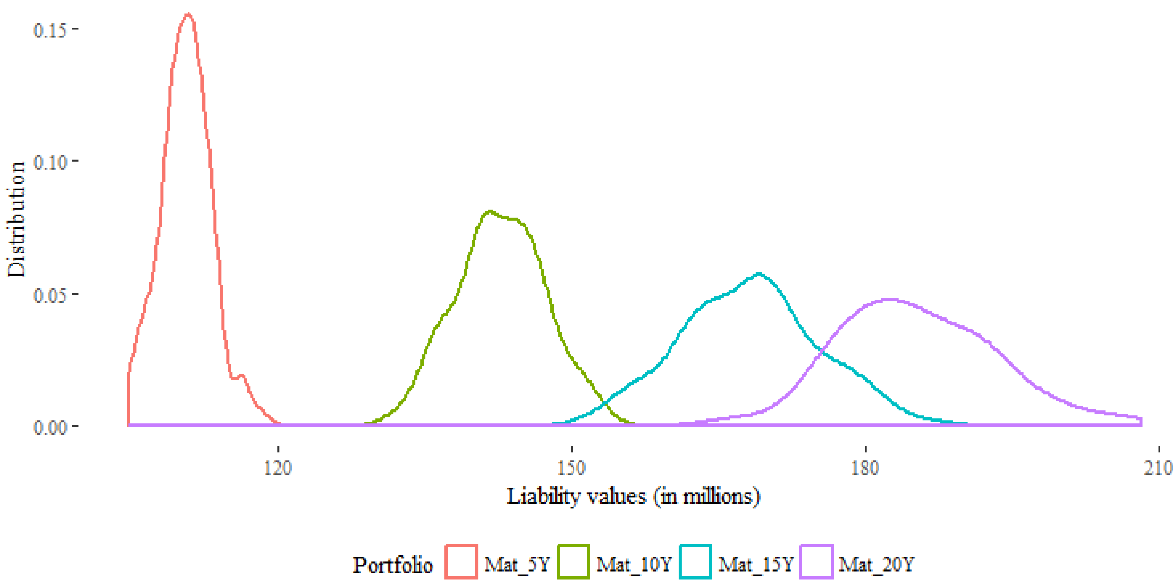
Simulating different yield curves, we obtain the distribution of assets in figure 2 and distribution of liabilities in figure 3.

Fig. 2: Distribution of Assets (values in millions)



Source: Authors work

Fig. 3: Distribution of Liabilities (values in millions)



Source: Authors work

From figure 3, we can see that the longer maturity of bonds results in the higher value of liabilities. If the assets have longer maturity, then we can expect more coupon payment which increases the the assets return. This higher asset return then causes higher profit sharing which increase the liability value. Figure 3 presents that the higher returns also increase the volatility of liabilities due to the variability of investment income. The bond portfolio with 5 years maturity reaches lower asset returns therefore many contracts do not have any profit share and their investment income is flat. The bond portfolio with higher maturities reaches higher assets return therefore the more contracts have profit share and their investment income is more volatile. The summary of liability distribution is in table 1.

Tab. 1: Summary of distribution

	mean	median	sd. dev	var. coef
Mat_5Y	110.6067	110.5286	2.6061	0.0235
Mat_10Y	142.7322	142.7783	4.5441	0.0318
Mat_15Y	168.1504	168.3498	6.9759	0.0414
Mat_20Y	185.2868	184.5312	7.9985	0.0431

Source: Authors work

Conclusion

A method to reduce computation time based on cluster analysis in life insurance seems to be a good approximate approach for liability modelling. In previous research, we have built a clustering approach for the purposes of the estimate of expected life insurance liability value and tested the precision of approximation on different scenarios (parallel change in interest rates or mortality and lapse shocks). In this article, we applied this approach for the asset liability management model and studied whether this faster liability model can be used also for the purpose of the ALM (e,g, applying different investment scenarios). We made four different investment strategies and observed that the results behave as we expected. In our demo portfolio, the higher assets return results in an increase of the liability through the profit-sharing scheme. The final calculation time of one ALM scenario is significantly faster from 2.2 hours by traditional per-policy approach to 1.2 minutes by clustering approach with the average error 0.019 %. Such a time reduction can now allow actuaries to test more investment scenarios and obtain better information about the risk profile of life insurance portfolio.

In the further research, the ALM model should be extended by reinvestments to study also impact on the value of assets. Because the clustering approach of liability estimation allows actuaries to test more investment strategies the next steps should also take place in dynamic ALM modelling to develop optimization procedures as cash flow matching, duration matching and partial duration matching.

Acknowledgment

The support of grant scheme Methods for fast estimation of life insurance liabilities with respect to different investment strategies IG410017 is gladly acknowledged.

References

Aas, K., & Neef, L. R., & Williams, L. (2018). Interest rate model comparisons for participating products under Solvency II. *Scandinavian Actuarial Journal*,(3), 203-224.

Balestreri, A., & Kent, J., & Morgan, E. (2011). Dynamic Asset Liability Management A Method for Optimising Investment Strategy. *European Actuarial Journal*,*1*(S1), 29-46. doi:10.1007/s13385-011-0011-7

Cipra, T. (2010). Financial and Insurance Formulas. doi:10.1007/978-3-7908-2593-0

Dmitriev, I. N. (2018). FAST ALGORITHM OF CLUSTER ANALYSIS k-MEDOIDS. *Prikladnaya Diskretnaya Matematika*,(39), 116-127. doi:10.17223/20710410/39/11

Fernandez, J., L., & Ferreiro-Ferreiro, A., M., & Garcia-Rodriguez, J., A. (2018). GPU Parallel Implementation for Asset-liability Management in Insurance Companies. *Journal of Computational Science*, (24), 232-254.

Fojtík, J., & Procházka, J., & Zimmermann, P., & Macková, S., & Švehláková, M. (2017). Alternative Approach for Fast Estimation of Life Insurance Liabilities. *Applications of Mathematics and Statistics in Economics*. doi:10.15611/amse.2017.20.11

Freedman, A., & Reynold, C., W. (2008). Cluster analysis: A spatial approach to actuarial modeling. http://www.milliman.com/uploadedFiles/insight/research/life-rr/cluster-analysis-a-spatial-rr08-01-08.pdf.

Maechler, M., & Rousseeuw, P., & Struyf, A., & Hubert, M., & Hornik, K. (2017). cluster: Cluster Analysis Basics and Extensions. *R package version 2.0.6*.

Shang, D., & Kuzmenko, V., & Uryasev, S. (2016). Cash flow matching with risks controlled by buffered probability of exceedance and conditional value-at-risk. *Annals of Operations Research*,*260*(1-2), 501-514. doi:10.1007/s10479-016-2354-6

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2), 177-188. doi:10.1016/0304-405X(77)90016-2

Xidonas, P., & Hassapis, C., & Bouzianis, G., & Staikouras, C. (2016). An Integrated Matching-Immunization Model for Bond Portfolio Optimization. *Computational Economics*,*51*(3), 595-605. doi:10.1007/s10614-016-9626-8

Contact

Jan Fojtík
University of Economics, Prague
Czech Republic W. Churchill Sq. 1938/4
130 67 Prague 3 - Žižkov
xfojj00@vse.cz

Jiří Procházka
University of Economics, Prague
Czech Republic W. Churchill Sq. 1938/4
130 67 Prague 3 - Žižkov
xproj16@vse.cz

Pavel Zimmermann
University of Economics, Prague
Czech Republic W. Churchill Sq. 1938/4
130 67 Prague 3 - Žižkov
zimmerp@vse.cz

Markéta Švehláková
University of Economics, Prague
Czech Republic W. Churchill Sq. 1938/4
130 67 Prague 3 - Žižkov
xsvem35@vse.cz

Simona Macková
University of Economics, Prague
Czech Republic W. Churchill Sq. 1938/4
130 67 Prague 3 - Žižkov
macs03@vse.cz