

Efficient valuation of a large portfolio of variable annuity contracts

Zusammenfassung der Masterarbeit an der Universität Ulm

Stefanie Burkart

Eine variable Annuität (VA, engl. Variable Annuity) ist eine Form der Lebensversicherung, die es Versicherungsnehmern ermöglicht, in den Finanzmarkt zu investieren, und gleichzeitig Mindestleistungen bietet. Bei Vertragsbeginn wird die Prämie typischerweise durch eine Einmalzahlung eingezogen und in mehrere Fonds investiert. Zum Ende der Laufzeit kann der Versicherte entscheiden, ob er diese Investitionen als einmalige Kapitalauszahlung oder als lebenslange Rente erhalten möchte. Ein attraktives Merkmal von VAs ist, dass man eine oder mehrere Garantien kaufen kann, um die Investition im Todesfall und/oder bei einer schlechten Fondsperformance zu schützen. Oftmals werden VAs so abgeschlossen, dass sie ein Einkommen nach Renteneintritt bieten.

Die Herausforderung der Masterarbeit besteht darin, die Versicherungsleistungen dieser Garantien für ein großes Portfolio von VA-Verträgen zu bewerten. Zusätzlich wird ein Risikomaß berechnet, das als Solvenzkapitalanforderung (SCR, engl. Solvency Capital Requirement) bezeichnet wird und den Betrag angibt, den eine Versicherungsgesellschaft zur Deckung unerwarteter Verluste innerhalb des nächsten Jahres vorhalten muss. Innerhalb des in der Arbeit betrachteten Modells kann die SCR-Berechnung auf die Bewertung der Versicherungsleistungen unter Betrachtung verschiedener Finanzmarktentwicklungen reduziert werden.

Da die Garantiezahlungen meist komplex und von der Entwicklung des Kapitalmarkts abhängig sind, gibt es im Allgemeinen keine geschlossene Lösung für das Bewertungsproblem. In der Praxis greifen Versicherungsunternehmen häufig auf Monte-Carlo(MC)-Simulationen zurück. Ein stochastisches Modell wird verwendet, um zukünftige ökonomische Szenarien zu simulieren. Für jedes Szenario werden die zukünftigen Zahlungen eines VA-Vertrag berechnet. Das Mittel der diskontierten Werte über alle Szenarien ergibt die MC-Schätzung des heutigen Zeitwerts der Leistungen eines Vertrages. Die Anwendung

dieser Simulationsmethode auf ein großes Portfolio von VA-Verträgen ist jedoch rechenintensiv, da jeder Vertrag über viele Szenarien für einen langen Zeithorizont projiziert werden muss.

Das Ziel der Arbeit ist es, die Berechnung in Bezug auf die Laufzeit effizienter zu gestalten, während die Genauigkeit der Ergebnisse so weit wie möglich erhalten bleibt.

Grundidee

Die Idee besteht darin, die Berechnung des heutigen Zeitwerts durch Anwendung von Methoden des maschinellen Lernens zu beschleunigen:

1. Zunächst wird eine kleine Teilmenge von VA-Verträgen aus dem Portfolio ausgewählt. Diese Verträge werden als Repräsentanten bezeichnet.
2. Für jeden der Repräsentanten wird eine MC-Simulation durchgeführt, um den heutigen Zeitwert des Vertrages zu ermitteln.
3. Basierend auf den Repräsentanten und deren Werte wird ein Regressionsmodell gefittet.
4. Mithilfe des Modells werden schließlich die Werte für das gesamte Portfolio vorhergesagt.

Dieser Ansatz ist in der Lage, die Rechenzeit erheblich zu reduzieren, da er nur eine geringe Anzahl an zeitintensiven MC-Simulationen benötigt.

Inhalt

Die Arbeit ist folgendermaßen aufgebaut. Nach einer Einführung gebe ich in Kapitel 2 eine kurze Beschreibung der VA-Produkte und ihrer grundlegenden Garantien. Es folgt ein universelles Rahmenwerk zur Bestimmung des heutigen Zeitwerts der Versicherungsleistungen und die traditionelle MC-Simulation wird vorgestellt, um diesen Wert näherungsweise zu bestimmen. Kapitel 3 führt in die Methodik des SCR ein und beschreibt, wie dieses über eine geschachtelte MC-Simulation geschätzt werden kann. In Kapitel 4 erläutere ich drei gängige Methoden des maschinellen Lernens: verschiedene Ausprägungen von Regressionsbäumen, verallgemeinerte lineare Modelle und neuronale Netze. In Kapitel 5 werden die numerischen Ergebnisse der Anwendung

dieser Vorhersagemodelle auf die Bestimmung des heutigen Zeitwerts und des SCRs betrachtet. Kapitel 6 schließt die Arbeit ab.

Ergebnisse

In der Arbeit wurde die Leistungsfähigkeit der drei maschinellen Lernverfahren auf einem synthetischen Datensatz hinsichtlich der Schätzung des heutigen Zeitwerts und des SCRs empirisch untersucht.

Alle Methoden des maschinellen Lernens waren in der Lage die Laufzeit drastisch zu reduzieren. Bei der Bestimmung des heutigen Zeitwerts des gesamten Portfolios waren Verfahren mindestens zehnmals schneller als die MC-Simulation. Die Berechnung des SCR per geschachtelter MC-Simulation war aufgrund des hohen Rechenaufwands nicht einmal durchführbar. Die Laufzeit wird auf circa 394 Stunden geschätzt. Dagegen konnten die maschinellen Lernverfahren das SCR innerhalb von maximal 6 Minuten bestimmen, was nur 0,023% der ursprünglichen Zeit entspricht.

Anhand verschiedener Metriken wurde die Vorhersagequalität und Stabilität der einzelnen Methoden untersucht. Die Unterschiede werden im Folgenden dargestellt.

Baumbasierte Methoden werden häufig für komplexe, nicht-lineare Zusammenhänge angewandt und gelten als nicht-parametrische Modelle. Aus diesem Grund sind sie nicht in der Lage Trends zu erkennen und können keine extrapolierten Werte außerhalb des Bereichs der Repräsentantenwerte vorhersagen. Das führt in unserem Anwendungsfall meistens zu einer Unterschätzung der Barwerte. Insgesamt ist die Modellqualität also sehr abhängig davon, wie gut die Barwerte der Repräsentanten die Werte des gesamten Portfolios aufspannen. Diese Eigenschaft kann jedoch zum Zeitpunkt der Auswahl nicht abgeschätzt werden, da die Barwerte aller Verträge noch unbekannt sind.

Die verallgemeinerten linearen Modelle (GLMs, engl. Generalized Linear Models) sind parametrische Modelle, die eine transformierte lineare Beziehung annehmen. Dies ermöglicht einerseits eine direkte Interpretation der Zusammenhänge, macht das Modell aber gleichzeitig sehr unflexibel. Für die Bestimmung des heutigen Barwerts erzielt das

GLM gute und stabile Ergebnisse. Es ist insbesondere zu empfehlen, wenn eine geringe Anzahl an Repräsentanten vorgegeben ist. Für die Bestimmung des SCRs scheint das Modell allerdings Schwierigkeiten zu haben, die ökonomische Situation in seiner vorgegebenen Struktur sinnvoll mit einzubeziehen.

Das neuronale Netz scheint in diesem Kontext eine sehr vielversprechende Technik zu sein. Mit einer ausreichenden Anzahl an Repräsentanten liefert es innerhalb der Arbeit die besten Ergebnisse für die Bestimmung des heutigen Zeitwertes und des SCRs. Aufgrund der Vielzahl an Parametern ist das Modell flexibler als ein GLM und kann andererseits auch extrapolieren. Ein „Over-Fitting“ des Modells wird durch Regularisation verhindert. Eine Schwachstelle des Netzes stellt die zufällige Initialisierung der Parameter dar. Abhängig davon kann der Trainingsprozess in verschiedenen lokalen Maxima enden und zu unterschiedlichen Vorhersagen führen. Die Verwendung von vielen Repräsentanten stabilisiert diesen Effekt bereits. Zusätzlich könnten in Zukunft Ensemble-Techniken eingesetzt werden.

Fazit

Zusammenfassend haben die numerischen Ergebnisse gezeigt, dass das neuronale Netz in Bezug auf die in diesem Zusammenhang betrachteten Methoden am besten abschneidet. Mit einer ausreichenden Anzahl von Repräsentanten ist es in der Lage, das Verhalten der Werte sehr genau vorherzusagen.

Dieses Ergebnis ist sehr zufriedenstellend; insbesondere, da MC-Werte auch nur Näherungen an die tatsächlichen Barwerte darstellen und die vorgestellte Methodik gleichzeitig eine immense Zeitersparnis ermöglicht.

Innerhalb der Arbeit wurden viele Annahmen und Vereinfachungen getroffen, um einen Bewertungsrahmen zu schaffen und daher sind die Ergebnisse nur in diesem Kontext gültig. Die Projektion auf realistische Verhältnisse und reale Datensätze muss sehr sorgfältig durchgeführt werden.