Guojun Gan and Emiliano A. Valdez
Department of Mathematics
University of Connecticut
Storrs, CT, USA

Metamodeling for Variable Annuities
To our students

– Guojun and Emiliano

To my mom Norma and my dad Ponso

– Emiliano
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>xi</td>
</tr>
<tr>
<td>I Preliminaries</td>
<td>1</td>
</tr>
<tr>
<td>1 Computational Problems in Variable Annuities</td>
<td>3</td>
</tr>
<tr>
<td>1.1 Variable Annuities</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Computational Problems Related to Daily Hedging</td>
<td>6</td>
</tr>
<tr>
<td>1.3 Computational Problems Related to Financial Reporting</td>
<td>7</td>
</tr>
<tr>
<td>1.4 Summary</td>
<td>9</td>
</tr>
<tr>
<td>2 Existing Approaches</td>
<td>11</td>
</tr>
<tr>
<td>2.1 Scenario Reduction</td>
<td>11</td>
</tr>
<tr>
<td>2.1.1 Scenario Ranking</td>
<td>11</td>
</tr>
<tr>
<td>2.1.2 Representative Scenarios</td>
<td>13</td>
</tr>
<tr>
<td>2.1.3 Importance Sampling</td>
<td>14</td>
</tr>
<tr>
<td>2.1.4 Curve Fitting</td>
<td>14</td>
</tr>
<tr>
<td>2.1.5 Random Sampling</td>
<td>16</td>
</tr>
<tr>
<td>2.2 Inforce Compression</td>
<td>16</td>
</tr>
<tr>
<td>2.2.1 Cluster Modeling</td>
<td>17</td>
</tr>
<tr>
<td>2.2.2 Replicating Liabilities</td>
<td>17</td>
</tr>
<tr>
<td>2.2.3 Replicated Stratified Sampling</td>
<td>19</td>
</tr>
<tr>
<td>2.3 Summary</td>
<td>20</td>
</tr>
<tr>
<td>3 Metamodelling Approaches</td>
<td>21</td>
</tr>
<tr>
<td>3.1 A General Framework</td>
<td>21</td>
</tr>
<tr>
<td>3.2 Literature Review</td>
<td>23</td>
</tr>
<tr>
<td>3.3 Summary</td>
<td>25</td>
</tr>
<tr>
<td>II Experimental Design Methods</td>
<td>27</td>
</tr>
<tr>
<td>4 Latin Hypercube Sampling</td>
<td>29</td>
</tr>
<tr>
<td>4.1 Description of the Method</td>
<td>29</td>
</tr>
<tr>
<td>4.2 Implementation</td>
<td>32</td>
</tr>
<tr>
<td>4.3 Examples</td>
<td>36</td>
</tr>
<tr>
<td>4.4 Summary</td>
<td>44</td>
</tr>
</tbody>
</table>
5 Conditional Latin Hypercube Sampling 45
  5.1 Description of the Method ......................... 45
  5.2 Implementation ................................... 46
  5.3 Examples ......................................... 48
  5.4 Summary .......................................... 52

6 Hierarchical $k$-Means 53
  6.1 Description of the Method ....................... 53
  6.2 Implementation .................................. 55
  6.3 Examples ......................................... 58
  6.4 Summary .......................................... 64

III Metamodels 67

7 Ordinary Kriging 69
  7.1 Description of the Model ....................... 69
  7.2 Implementation .................................. 73
  7.3 Applications .................................... 77
    7.3.1 Ordinary Kriging with Latin Hypercube Sampling . 80
    7.3.2 Ordinary Kriging with Conditional Latin Hypercube Sampling ........................................... 84
    7.3.3 Ordinary Kriging with Hierarchical $k$-Means .... 88
  7.4 Summary .......................................... 93

8 Universal Kriging 95
  8.1 Description of the Model ....................... 95
  8.2 Implementation .................................. 98
  8.3 Applications .................................... 100
    8.3.1 Universal Kriging with Latin Hypercube Sampling . 101
    8.3.2 Universal Kriging with Conditional Latin Hypercube Sampling ........................................... 105
    8.3.3 Universal Kriging with Hierarchical $k$-Means .... 110
  8.4 Summary .......................................... 114

9 GB2 Regression Model 117
  9.1 Description of the Model ....................... 117
  9.2 Implementation .................................. 119
  9.3 Applications .................................... 123
    9.3.1 GB2 with Latin Hypercube Sampling ............. 124
    9.3.2 GB2 with Conditional Latin Hypercube Sampling ........................................... 125
    9.3.3 GB2 with Hierarchical $k$-Means ............... 127
  9.4 Summary .......................................... 129
Contents

10 Rank Order Kriging 131
10.1 Description of the Model ............................... 131
10.2 Implementation ........................................ 133
10.3 Applications .......................................... 134
10.3.1 Rank Order Kriging with Latin Hypercube Sampling ........ 136
10.3.2 Rank Order Kriging with Conditional Latin Hypercube Sampling ......................... 141
10.3.3 Rank Order Kriging with Hierarchical k-Means .......... 145
10.4 Summary .................................................. 150

11 Linear Model with Interactions 151
11.1 Description of the Model ............................... 151
11.2 Implementation ........................................ 153
11.3 Applications .......................................... 154
11.3.1 Linear Model with Latin Hypercube Sampling .......... 156
11.3.2 Linear Model with Conditional Latin Hypercube Sampling ............................... 158
11.3.3 Linear Model with Hierarchical k-Means ........ 160
11.4 Summary .................................................. 163

12 Tree-Based Models 165
12.1 Description of the Model ............................... 165
12.2 Implementation ........................................ 167
12.3 Applications .......................................... 168
12.3.1 Regression Trees with Latin Hypercube Sampling ........ 169
12.3.2 Regression Trees with Conditional Latin Hypercube Sampling ............................... 173
12.3.3 Regression Trees with Hierarchical k-Means ........ 177
12.4 Summary .................................................. 182

A Synthetic Datasets 183
A.1 The Synthetic Inforce ................................... 183
A.2 The Greeks .............................................. 183

Bibliography 189

Index 195
Preface

Variable annuities are life insurance products that offer various types of financial guarantees. Insurance companies that have a large block of variable annuity business face many challenges. For example, guarantees embedded in variable annuity policies demand sophisticated models for pricing, financial reporting, and risk management. In practice, insurance companies rely heavily on Monte Carlo simulation to calculate the fair market values of the guarantees because the guarantees are complicated and no closed-form valuation formulas are available. One drawback of Monte Carlo simulation is that it is extremely time-consuming and often prohibitive to value a large portfolio of variable annuity contracts because every contract needs to be projected over many scenarios for a long time horizon.

This monograph is devoted to metamodeling approaches, which have been proposed recently in the academic literature to address the computational problems associated with the valuation of large variable annuity portfolios. A typical metamodeling approach involves the following four steps:

1. select a small number of representative variable annuity contracts (i.e., experimental design),
2. use Monte Carlo simulation to calculate the fair market values (or other quantities of interest) of the representative contracts,
3. build a metamodel (i.e., a predictive model) based on the representative contracts and their fair market values, and
4. use the metamodel to estimate the fair market value for every contract in the portfolio.

Using metamodeling approaches can significantly reduce the runtime of valuing a large portfolio of variable annuity contracts for the following reasons: first, building a metamodel only requires using the Monte Carlo simulation model to value a small number of representative contracts; second, the metamodel is usually much simpler and more computation efficient than Monte Carlo simulation.

This book is primarily written for undergraduate students, who study actuarial science, statistics, risk management, and financial mathematics. It is equally useful for practitioners, who work in insurance companies, consulting firms, and banks. The book is also a source of reference for researchers and graduate students with scholarly interest in computational issues related to
variable annuities and other similar insurance products. The methods presented in the book are described in detail and implemented in R, which is a popular language and environment for statistical computing and graphics. Using the R code and datasets that accompany this book, readers can easily replicate the numerical results presented in the book. In addition, readers can modify the R code included in this book for their own use.

This book is divided into three parts. The first part, which consists of Chapters 1, 2, and 3, introduces the computational problems associated with variable annuity valuation, reviews existing approaches, and presents the metamodeling approach in detail. The second part includes Chapters 4, 5, and 6. This part introduces some experimental design methods, which are used to select representative variable annuity contracts. In particular, we describe Latin hypercube sampling, conditional Latin hypercube sampling, and hierarchical $k$-means for selecting representative policies. The third part includes Chapters 7 to 12 and introduces six metamodels: ordinary kriging, universal kriging, GB2 regression, rank order kriging, linear model with interactions, and tree-based models. These metamodels are predictive models that have been studied in the academic literature for speeding up variable annuity valuation. The dependencies of the chapters are shown in Figure 1. We implement all the experimental design methods and metamodels using a synthetic dataset, which is described in the appendix of this book.

This book has grown from several research projects on variable annuity valuation undertaken by the authors at the University of Connecticut. First,
Preface

we would like to thank the financial support provided through a CAE (Centers of Actuarial Excellence) grant (i.e., *Applying Data Mining Techniques in Actuarial Science*) from the Society of Actuaries. Second, we would like to acknowledge two CKER (Committee on Knowledge Extension Research) Individual Grants (i.e., *Regression Modeling for the Valuation of Large Variable Annuity Portfolios* and *Valuation of Large Variable Annuity Portfolios with Rank Order Kriging*).

Guojun Gan and Emiliano A. Valdez
Storrs, Connecticut, USA
June 1, 2019