Insurance: Mathematics and Economics 82 (2018) 124–140

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Estimating loss reserves using hierarchical Bayesian Gaussian process regression with input warping

Nathan Lally [a], Brian Hartman [b,*]

[a] Hartford Steam Boiler, United States[1]
[b] Brigham Young University, United States

A R T I C L E I N F O

Article history:
Received September 2017
Received in revised form June 2018
Accepted 21 June 2018

Keywords:
Loss reserves
Claims triangles
Input warping
Gaussian process regression
Spatial statistics

A B S T R A C T

In this paper, we visualize the loss reserve runoff triangle as a spatially-organized data set. We apply Gaussian Process (GP) regression with input warping and several covariance functions to estimate future claims. We then compare our results over a range of product lines, including workers' comp, medical malpractice, and personal auto. Even though the claims development of the lines are very different, the GP method is very flexible and can be applied to each without much customization. We find that our model generally outperforms the classical chain ladder model as well as the recently proposed hierarchical growth curve models of Guszcza (2008) in terms of point-wise predictive accuracy and produces dramatically better estimates of outstanding claims liabilities.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Insurance is one of the few industries where the cost of the product is not known when it is priced for sale. To determine the rate for a policy, the insurer must predict the future claims from that policy. To predict those future claims, insurers use past claims from similar policies. Ideally, the most recently written policies will closely match the future policies and should be prominently included in the model. Unfortunately, the total cost of a policy is not known immediately after policy expiration. Reporting lags, litigation, settlement negotiation, and other adjustments to the ultimate claims can all lengthen the time until the ultimate cost of the policy is known.

Loss reserves represent the insurer's best estimate of their outstanding loss payments. These reserves include both incurred, but not reported (IBNR) losses (losses incurred by the policyholder during the policy period, but not yet reported to the insurer as of the valuation date) and incurred, but not enough reported (IBNER) losses (the insurer knows about these losses, but the predicted ultimate costs as of the valuation date are often smaller than the actual ultimate losses). Properly estimating these ultimate losses is important for future pricing and company valuation.

For comprehensive reviews on the prediction of loss reserves and their associated variability, see Taylor (2012) and Wüthrich and Merz (2008). In recent years, a variety of regression models have been proposed for forecasting loss reserves; supplementing a litany of deterministic and stochastic link ratio based models. Earlier work tends to build on parallels between link ratio methods and certain types of regression (Barnett and Zehnwirth, 2000). Many authors have built upon this framework using mixed models (Antonio and Beirlant, 2008), Bayesian models (de Alba and Nieto-Barajas, 2008; Shi et al., 2012), and unique link functions (Peters et al., 2009). Zhang and Dukic (2013) added copulas to account for the correlation between various lines, while Shi and Hartman (2016) account for those same correlations using a Bayesian hierarchical model. However, a common criticism of link ratio models and their regression-based derivatives is that they tend to be heavily parameterized for a problem with few degrees of freedom (England and Verrall, 2002). To address this shortcoming, interest in research on nonlinear regression approaches deviating from traditional actuarial models has grown. Two broad categories of nonlinear regressions can be considered, parametric models, or those where the nonlinear relationship between covariates and losses takes an explicit functional form, and non-parametric models, where a nonlinear relationship is defined more generally and learned from the data.

The parametric class of nonlinear reserve models is well represented by Stelljes (2006), modeling incremental losses with exponential curves, and later by Guszcza (2008) and Zhang et al.

* Corresponding author.
E-mail address: hartman@stat.byu.edu (B. Hartman).

1 Nathan Lally is an employee of The Hartford Steam Boiler Inspection and Insurance Company (HSB). This article is presented for informational purposes only. Views and opinions expressed in the article do not necessarily reflect those of HSB. HSB makes no warranties or representations as to the accuracy or completeness of the content of this article. Under no circumstances shall HSB or any party involved in creating or delivering this article be liable to you for any loss or damage that results from the use of the information contained in this article.

https://doi.org/10.1016/j.insmatheco.2018.06.008
0167-6687/© 2018 Elsevier B.V. All rights reserved.

(2012) who forecast cumulative losses. The latter two papers allow the intercept to vary by accident year using a hierarchical structure and model the development lag effect with Weibull and lognormal distributions. These two papers differ slightly as Guszcza (2008) uses a maximum likelihood estimation (MLE) while Zhang et al. (2012) use Bayesian estimation and consider serial correlation in the errors. An unaddressed concern with nonlinear parametric regressions is that the choice of functional form can drastically affect ultimate reserve estimates, as observed in Guszcza (2008). In the absence of concrete prior knowledge about the loss generating process, these methods may produce poor results.

Though less popular than their parametric counterparts, nonparametric nonlinear regressions have also been employed in the literature. In a comprehensive review of stochastic reserving methods England and Verrall (2002) introduce generalized additive models (GAM) with cubic regression splines as a method to forecast losses; noting flexibility and the ability to reproduce ad hoc adjustments to deterministic models as distinct advantages to this approach when compared to parametric models. Extending this methodology, Spedicato et al. (2014) use generalized additive models for location, scale and shape (GAMLSS) to model the conditional scale parameter as well as the location parameter for a variety of distributions. However, the authors conclude the methodology produces mixed results due to problems with convergence, large differences in variability when compared to standard models, and poor predictive accuracy when compared to the best linear unbiased estimator (BLUE) chain ladder approach.

To improve upon existing approaches to loss reserve forecasting, we propose a hierarchical Bayesian Gaussian process (GP) regression with input warping similar to Snoek et al. (2014). GP regression is a flexible nonparametric statistical/machine learning method which provides a robust and smooth fit to a wide variety of data types, structures, and distributions. We reserve a detailed description of GP regression for Section 2.

Recently, Lopes et al. (2012) proposed using hybrid chain ladder/kernel machine (both support vector machines and GP regression) models for incurred but not reported claim reserve estimation. This paper is the first to introduce GP regression to reserving literature but not as a stand-alone methodology. GP regression was only used to adjust residuals from the chain ladder model with the hopes of obtaining more accurate predictions and the authors struggle to find an expression for a IBNR variance estimator.[2]

In reserve modeling, GP regression with input warping offers several advantages over popular methods used in industry and contemporary literature,

- Because it is a nonparametric method, the relationship between accident years, development lag, and losses is learned from the data rather than being pre-specified as in parametric models. Nor is any post-hoc adjustment (as used with deterministic methods) necessary. Defining losses as a smooth function of accident period and development lag is more consistent with reality than the random intercept model proposed in Guszcza (2008) and Zhang et al. (2012) and has the added benefit of enabling interpolation/extrapolation along both time dimensions.[3]
- Through its covariance function, GP regression naturally models the dependence structure between losses across both the accident period and development lag dimensions. Methods and concepts from spatial/geo-statistics can be borrowed to visualize and make sense of this dependence structure.
- GP regression is parsimonious, only requiring the estimation of a few hyperparameters to learn potentially complex relationships present in the data. For example, it is possible to fit GP models to incremental loss data.[4] without relying on external models or residual analysis as in Stelljes (2006) For relatively simple covariance functions, GP hyperparameters are easily to interpret and enable substantial posterior inference.
- GP regression models can be implemented efficiently and easily though standard software, we use Stan (Team, 2017) in this application, using Hamiltonian Monte Carlo (HMC) with automatic tuning provided by the No-U-Turn (NUTS) sampler (Hoffman and Gelman, 2014). This paradigm affords a great deal of flexibility, allowing the practitioner to easily adjust the details of the model to take an objective stance or to incorporate prior assumptions based on previous experience. Parameter and process variability can be measured exactly and directly rather than through asymptotic methods or bootstrapping. Finally, given reasonable hyperprior elicitation, posterior computation using HMC is more forgiving than classical methods for fitting non-linear regressions. We experienced no problems with convergence when fitting GP regressions to reserve data.
- Input warping through hierarchical Bayes automates feature engineering and incorporates major non-stationary effects (Snoek et al., 2014).

In Section 2 we provide a brief overview of GP regression and covariance functions.[5] In Section 3 we present the intuition behind our approach; viewing the problem as a geostatistician. Section 4 details our proposed model. Section 5 applies our method to several sets of paid loss data from the NAIC Schedule P and compares predictive accuracy with the popular chain ladder and hierarchical growth curve models.[6] We conclude in Section 6 by suggesting potential modifications extensions to our GP reserve models for future research. We include an appendix with Stan code to implement a GP regression with input warping. Stan is freely available allowing practitioners to take our code and implement these models on their own data.

2. Gaussian process regression

Given a training matrix $X \in \mathbb{R}^{n,p}$ and an associated target vector $y \in \mathbb{R}^n$, a GP regression can be applied to learn an unknown function $f(x)$ (where $x \in \mathbb{R}^p$ is any row vector in X) which models the target observations y. In most applications it is assumed the observations deviate from $f(x)$ according to some noise parameter but this need not be the case if the process is truly noise-free (ex: output from a deterministic computer simulation model). Before moving on we formalize the definitions of stochastic processes and GPs for reference.

Definition 1 (*Stochastic Process*)**.** Defining Ω as a sample space, $\mathcal{F}$ a set of events, and $\mathcal{P}$ a function assigning probabilities to events, a *stochastic process* is a sequence of random variables defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and ordered with respect to a time index t, taking values in an index set $\mathcal{S}$ (the state space).

[2] If the chain ladder estimate is to be considered fixed/deterministic then Williams and Rasmussen (2006) chapter 2 section 2.7 provide solutions for the predictive mean and variance under the GP model in Lopes et al. (2012).

[3] Accident period may only be extrapolated to whole unit values (ex. year 1 to year 2) since losses will begin at 0 and accumulate for each period. Along the development lag dimension, interpolation/extrapolation can be performed at any continuous value.

[4] This paper focuses on modeling cumulative losses, but the methodology extends naturally to incremental losses.

[5] For a far more thorough overview of GP regression and covariance functions we recommend Williams and Rasmussen (2006).

[6] It was our intention to include the additive models of Spedicato et al. (2014) for comparison to GP regression however, as the authors warned, we experienced convergence issues and generally poor results.

Download English Version:

https://daneshyari.com/en/article/7354564

Download Persian Version:

https://daneshyari.com/article/7354564

Daneshyari.com