risks

MDPI

Commentary

Kalman Filter Learning Algorithms and State Space Representations for Stochastic Claims Reserving

Nataliya Chukhrova † and **Arne Johannssen** *,†

Faculty of Business Administration, University of Hamburg, 20146 Hamburg, Germany; nataliya.chukhrova@uni-hamburg.de

* Correspondence: arne.johannssen@uni-hamburg.de

† The authors contributed equally to this work.

Abstract: In stochastic claims reserving, state space models have been used for almost 40 years to forecast loss reserves and to compute their mean squared error of prediction. Although state space models and the associated Kalman filter learning algorithms are very powerful and flexible tools, comparatively few articles on this topic were published during this period. Most recently, several articles have been published which highlight the benefits of state space models in stochastic claims reserving and may lead to a significant increase in its popularity for applications in actuarial practice. To further emphasize the merits of these papers, this commentary highlights various additional aspects that are useful for practical applications and offer some fruitful directions for future research.

Keywords: adaptive learning; dependence modeling; evolutionary models; insurance; Kalman filter; machine learning; multivariate analysis; quantitative risk management; state space models; time series forecasting

Citation: Chukhrova, Nataliya, and Arne Johannssen. 2021. Kalman Filter Learning Algorithms and State Space Representations for Stochastic Claims Reserving. *Risks* 9: 112. https://doi.org/10.3390/risks9060112

Academic Editors: Mogens Steffensen and Alexandra Dias

Received: 22 April 2021
Accepted: 4 June 2021
Published: 6 June 2021

In non-life insurance, loss reserves often make up a large share of the liability side of the balance sheet, so forecasting these liabilities and quantifying their uncertainty is a key actuarial issue. Thus, numerous loss-reserving techniques have been developed over the last 40 years, but only a few of them are based on time series models. This is particularly astonishing in view of the fact that the claims reserving process is a stochastic process and claims development data correspond to time series data. In this commentary, we deal with time series models, particularly state space models (SSMs, also referred to as evolutionary models) and the associated Kalman filter (KF) learning algorithms. SSMs and the KF are powerful and flexible tools, originating in the early 1960s in the field of aerospace engineering. Since then, they have been applied to a wide range of problems, e.g., in the field of technical process control, speech recognition, robotic motion planning, trajectory optimization, oceanography, agriculture, and stochastic claims reserving.

Recently, some articles have been published that deal with SSMs in stochastic claims reserving (see Avanzi et al. 2020; Costa and Pizzinga 2020; Hendrych and Cipra 2021). The main aspects involved are multivariate SSMs, which allow for the incorporation of claims activity dynamics and modeling of dependencies between correlated lines of business (Avanzi et al. 2020; Hendrych and Cipra 2021), as well as the concept of row-wise stacking the claims development data within run-off triangles to form a univariate time series (Costa and Pizzinga 2020; Hendrych and Cipra 2021). Moreover, the authors address issues such as calendar year reserve prediction, inclusion of tail effects (Costa and Pizzinga 2020), a two-dimensional evolution of factors across accident and calendar years (Avanzi et al. 2020), and a unified modeling of various common approaches (log-normal models, Hoerl curve approaches etc) within the same kind of state space representation (Hendrych and Cipra 2021). Especially in the today's data-driven insurance business, the handling of Big Data and of dependencies across a huge number of multiple business segments (business lines and their various subsets) are crucial characteristics of claims reserving for insurance companies. Thus, the adaptive learning approach underlying SSMs and the KF is a powerful tool for the improvement in handling and forecasting of claims development data.

To further emphasize the merits of these papers, this commentary intends to highlight various additional aspects that are useful for practical applications but were not deeply considered in these papers. Firstly, the papers are primarily focused on the construction of SSMs, and the powerful technique of KF learning algorithms that usually comes along with SSMs is kept almost completely in the background. However, knowledge of this technique makes deeper interpretations and reasonable adjustments possible. Secondly, the papers mostly deal with a specific state space representation, but there is no comparative discussion regarding other common approaches to construct state space representations in stochastic claims reserving. Thus, this commentary intends to highlight the methodology behind KF learning algorithms on the one hand, and to provide a brief discussion on alternatives for state space modeling on the other hand. In addition, some promising directions for future research are given.

As for achieving the KF recursions, let us consider a general multivariate SSM describing the dynamics of $\{\mathbf{Y}_t\}_{t\in\mathbb{N}}$ and $\{\mathbf{X}_t\}_{t\in\mathbb{N}}$,

$$\begin{aligned} \mathbf{Y}_t &= \mathbf{G}_t\mathbf{X}_t + \mathbf{W}_t \qquad &\text{(observation equation)},\\ \mathbf{X}_{t+1} &= \mathbf{F}_t\mathbf{X}_t + \mathbf{V}_t \qquad &\text{(state equation)}, \end{aligned}$$

where $\{\mathbf{W}_t\}_{t\in\mathbb{N}}$ (measurement noise), $\{\mathbf{V}_t\}_{t\in\mathbb{N}}$ (process noise) with $\mathbb{E}[\mathbf{W}_t] = \mathbf{0}$, $\mathbb{E}[\mathbf{V}_t] = \mathbf{0}$,

$$\mathbb{E}\left[\mathbf{W}_s\mathbf{W}_t^T\right] = \begin{cases} \mathbf{R}_t & \text{if } s = t \\ \mathbf{O} & \text{otherwise} \end{cases} \quad \text{and} \quad \mathbb{E}\left[\mathbf{V}_s\mathbf{V}_t^T\right] = \begin{cases} \mathbf{Q}_t & \text{if } s = t \\ \mathbf{O} & \text{otherwise} \end{cases},$$

and $\mathbb{E}\left[\mathbf{V}_s\mathbf{W}_t^T\right] = \mathbf{O}$ for all $s, t \in \mathbb{N}$. The KF recursions with initial conditions $\widehat{\mathbf{X}}_{1|0}$ and $\mathbf{P}_{1|0}$ are then given by

$$\begin{aligned} \widehat{\mathbf{X}}_{t+1|t} &= \mathbf{F}_t\widehat{\mathbf{X}}_{t|t-1} + \mathbf{F}_t\mathbf{K}_t\left(\mathbf{Y}_t - \mathbf{G}_t\widehat{\mathbf{X}}_{t|t-1}\right) \qquad &\text{(one-step prediction)}\\ \widehat{\mathbf{X}}_{t+h|t} &= (\mathbf{F}_{t+h-1}\mathbf{F}_{t+h-2}\cdots\mathbf{F}_{t+1})\widehat{\mathbf{X}}_{t+1|t} \qquad &\text{(h-step prediction)}\\ \widehat{\mathbf{X}}_{t|t} &= \widehat{\mathbf{X}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{Y}_t - \mathbf{G}_t\widehat{\mathbf{X}}_{t|t-1}\right) \qquad &\text{(filtering)}\\ \widehat{\mathbf{X}}_{t|s} &= \widehat{\mathbf{X}}_{t|s-1} + \mathbf{K}_s^S\left(\mathbf{Y}_s - \mathbf{G}_s\widehat{\mathbf{X}}_{s|s-1}\right) \qquad &\text{(fixed-point smoothing)} \end{aligned}$$

and

$$\begin{aligned} \mathbf{P}_{t+1|t} &= \mathbf{F}_t\mathbf{P}_{t|t-1}\mathbf{F}_t^T + \mathbf{Q}_t - \mathbf{F}_t\mathbf{K}_t\mathbf{Y}_t^T\mathbf{F}_t^T\\ \mathbf{P}_{t+h|t} &= \mathbf{F}_{t+h-1}\mathbf{P}_{t+h-1|t}\mathbf{F}_{t+h-1}^T + \mathbf{Q}_{t+h-1}\\ \mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} - \mathbf{K}_t\mathbf{Y}_t^T\\ \mathbf{P}_{t|t-1}^{s+1|s} &= \mathbf{P}_{t|t-1}^{s|s-1}\left(\mathbf{F}_s - \mathbf{Y}_s\boldsymbol{\Delta}_s^{-1}\mathbf{G}_s\right)^T\\ \mathbf{P}_{t|s} &= \mathbf{P}_{t|s-1} - \mathbf{K}_s^S\left(\mathbf{Y}_s^S\right)^T \end{aligned}$$

where $\mathbf{K}_t = \mathbf{Y}_t\boldsymbol{\Delta}_t^{-1}$, $\mathbf{Y}_t = \mathbf{P}_{t|t-1}\mathbf{G}_t^T$, $\boldsymbol{\Delta}_t = \mathbf{G}_t\mathbf{P}_{t|t-1}\mathbf{G}_t^T + \mathbf{R}_t$, $\boldsymbol{\Psi}_t = \mathbf{P}_{t|t}\mathbf{F}_{t+1}\mathbf{P}_{t+1|t}^{-1}$, $\mathbf{K}_s^S = \mathbf{Y}_s^S\boldsymbol{\Delta}_s^{-1}$, $\mathbf{Y}_s^S = \mathbf{P}_{t|t-1}^{s|s-1}\mathbf{G}_s^T$ for $h \geq 2$, $s \geq t$ (see Johannssen 2016). While the KF prediction recursions are required to forecast future observations, the KF filtering and smoothing recursions are useful to identify outliers and to interpolate gaps in the data (e.g., resulting from a merger). Note that the KF can also be seen as an evolutionary credibility model (see Wüthrich and Merz 2008).

An explicit calculation of the optimal predictor $\widehat{\mathbf{X}}_{t|s}$ using observations is, in general, impossible, unless the noise processes $\{\mathbf{W}_t\}_{t\in\mathbb{N}}$, $\{\mathbf{V}_t\}_{t\in\mathbb{N}}$ are Gaussian or the optimal predictor is a linear function of the observations under the restriction of a quadratic loss function. Due to the Hilbert projection theorem, $\widehat{\mathbf{X}}_{t|s}$ is the best approximation of $\mathbf{X}_t$ by a vector $\widetilde{\mathbf{X}}_{t|s} \in \mathcal{Y}_s$, i.e.,

$$\mathbb{E}\left[||\mathbf{X}_t - \widehat{\mathbf{X}}_{t|s}||^2\right] = \min_{\widetilde{\mathbf{X}}_{t|s} \in \mathcal{Y}_s} \mathbb{E}\left[||\mathbf{X}_t - \widetilde{\mathbf{X}}_{t|s}||^2\right],$$

where $\mathcal{Y}_s$ is a vector subspace spanned by $\mathbf{Y}_1, \ldots, \mathbf{Y}_s$ and $\widehat{\mathbf{X}}_{t|s}$ is the orthogonal projection of $\mathbf{X}_t$ onto $\mathcal{Y}_s$. Thus, the KF is generally distribution-free and provides the best linear predictors in the sense of minimizing the mean squared error. Provided that the second moment of $\mathbf{X}_t$ exists, orthogonal projection is actually identical with conditional expectation, i.e., $\widehat{\mathbf{X}}_{t|s} = \mathbb{E}[\mathbf{X}_t|\mathbf{Y}_1, \ldots, \mathbf{Y}_s]$. That is, $\widehat{\mathbf{X}}_{t|s}$ is the best (not necessarily linear) predictor of $\mathbf{X}_t$ using $\mathbf{Y}_1, \ldots, \mathbf{Y}_s$. Since this is also achieved when using the Gaussian assumption, which, in addition, considerably simplifies parameter estimation, the Gaussian assumption is common in the literature on SSMs and KF in stochastic claims reserving. However, the Gaussian assumption is not a mandatory requirement when dealing with SSMs and KF, much less justified in any case. For instance, Hoerl curve approaches are generally distribution-free.

There are some interesting interpretations of the system matrices and set screws within the KF recursions that could be revealing in actuarial practice. The KF recursions all have the same basic structure: the optimal predictor is a linear combination of the preceding prediction and the innovation (i.e., the difference between current and predicted observation) weighted by the Kalman gain $\mathbf{K}$. The Kalman gain is a weighting matrix that determines the weight of the innovation that is to be incorporated into the current prediction of the state vector, i.e., it quantifies the relative importance of the most recent observation. The Kalman gain is made up of two matrices, the cross-covariance matrix $\mathbf{Y}$ between the state to be predicted and the innovation as well as the inverse of the error covariance matrix $\mathbf{\Delta}$. To depict this clearly, $\mathbf{K}$ can be interpreted as the "ratio" of two covariance terms. When both matrices are real numbers, one could also state: The higher the covariance between the state to be predicted and the innovation and/or the lower the variance of the innovation, the higher the trust in the new observation and, therefore, the higher the Kalman gain. Thus, the Kalman gain is also an important set screw for systematic weighting of observations, e.g., to assign higher weights to more recent observations, lower weights to outliers or zero-weights to missing observations (as in the h-step prediction recursion).

Alongside the row-wise stacking approach (see Atherino et al. 2010; Costa and Pizzinga 2020; Hendrych and Cipra 2021) there are some other approaches to construct state space models in stochastic claims reserving. The most prevalent approach is a calendar-year-based modeling where claims development data of different calendar years are stacked into separate observation vectors. This approach can be found in De Jong and Zehnwirth (1983), Verrall (1989, 1994), Taylor et al. (2003), De Jong (2006), and Li (2006). Following Chukhrova and Johannssen (2017), the main reasons for the popularity of the calendar-year-based approach are (1) natural modeling of the claims data, (2) appropriate embedding of calendar year effects (e.g., inflation factor or changes in legislation, see also Avanzi et al. 2018; Wüthrich and Merz 2013), and (3) higher weighting of more recent observations. In addition to the calendar-year-based approach, there are accident-year- and development-year-based modeling approaches appearing only in Taylor et al. (2003) and De Jong and Zehnwirth (1983), respectively. However, both approaches have no significant advantages compared to the calendar year approach. In contrast to the row-wise stacking approach, the above approaches have the drawback that the dimensions of the system vectors and matrices are time-variant due to the varying numbers of observations in different calendar years. This can complicate parameter estimation, practical handling, and the simultaneous involvement of multiple run-off triangles considerably.

Finally, we would like to point out some promising directions for future research in stochastic claims reserving:

- In addition to the hitherto considered full reserve risk, one could investigate the one-year reserve risk by quantifying the claims development result (see, e.g., Merz and Wüthrich 2008, 2012) via SSMs and KF;
- Instead of assuming linear systems, one could consider non-linear systems, where the extended KF is applicable (see, e.g., Julier and Uhlmann 2004);
- It is also feasible to conduct micro-level claims reserving (see, e.g., De Felice and Moriconi 2019; Duval and Pigeon 2019) by means of SSMs and KF;
- One could consult an outlier-robust KF (see, e.g., Agamennoni et al. 2011) or an interval KF for interval-linear systems (see, e.g., Chen et al. 1997).

Beyond the field of stochastic claims reserving, insurance companies may benefit from using SSMs and KF in any area where time series data are available, such as investments, expenditures, demand for insurance, and other economic processes that evolve over time. In particular, due to the high flexibility of SSMs, they can be applied for modeling stationary or non-stationary and univariate or multivariate time series, as well as in cases with missing data, interventions, structural changes or other irregularities in the data. Moreover, SSMs are able to detect the temporal dynamics of a system accurately, especially when compared to other time series models.

Author Contributions: Conceptualization, N.C. and A.J.; methodology, N.C. and A.J.; formal analysis, N.C. and A.J.; investigation, N.C. and A.J.; writing—original draft preparation, A.J.; writing—review and editing, N.C.; project administration, A.J. Both authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank both anonymous reviewers for their valuable feedback and suggestions, which were helpful in further improving the commentary.

Conflicts of Interest: The authors declare no conflict of interest.

References

Agamennoni, Gabriel, Juan I. Nieto, and Eduardo M. Nebot. 2011. An outlier-robust Kalman filter. Paper presented at IEEE International Conference on Robotics and Automation (ICRA 2011), Shanghai, China, May 9–13.

Atherino, Rodrigo, Adrian Pizzinga, and Cristiano Fernandes. 2010. A row-wise Stacking of the Runoff Triangle: State Space Alternatives for IBNR Reserve Prediction. *ASTIN Bulletin* 40: 917–46.

Avanzi, Benjamin, Gregory C. Taylor, and Bernard Wong. 2018. Common Shock Models for Claim Arrays. *ASTIN Bulletin* 48: 1109–36. [CrossRef]

Avanzi, Benjamin, Gregory C. Taylor, Phuong A. Vu, and Bernard Wong. 2020. A multivariate evolutionary generalised linear model framework with adaptive estimation for claims reserving. *Insurance: Mathematics and Economics* 93: 50–71.

Chen, Guanrong, Jianrong Wang, and Leang-san Shieh. 1997. Interval Kalman filtering. *IEEE Transactions on Aerospace and Electronic Systems* 33: 250–59. [CrossRef]

Chukhrova, Nataliya, and Arne Johannssen. 2017. State Space Models and the Kalman-Filter in Stochastic Claims Reserving: Forecasting, Filtering and Smoothing. *Risks* 5: 30. [CrossRef]

Costa, Leonardo, and Adrian Pizzinga. 2020. State-space models for predicting IBNR reserve in row-wise ordered runoff triangles: Calendar year IBNR reserves & tail effects. *Journal of Forecasting* 39: 438–48.

De Felice, Massimo, and Franco Moriconi. 2019. Claim Watching and Individual Claims Reserving Using Classification and Regression Trees. *Risks* 7: 102. [CrossRef]

De Jong, Piet. 2006. Forecasting Runoff Triangles. *North American Actuarial Journal* 10: 28–38. [CrossRef]

De Jong, Piet, and Ben Zehnwirth. 1983. Claims Reserving, State-Space Models and the Kalman Filter. *Journal of the Institute of Actuaries* 110: 157–81. [CrossRef]

Duval, Francis, and Mathieu Pigeon. 2019. Individual Loss Reserving Using a Gradient Boosting-Based Approach. *Risks* 7: 79.

Hendrych, Radek, and Tomas Cipra. 2021. Applying State Space Models to Stochastic Claims Reserving. *ASTIN Bulletin* 51: 267–301. [CrossRef]

Johannssen, Arne. 2016. *Stochastische Schadenreservierung unter Verwendung von Zustandsraummodellen und des Kalman-Filters*. Hamburg: Dr. Kovac.

Julier, Simon J., and Jeffrey K. Uhlmann. 2004. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE* 92: 401–22. [CrossRef]

Li, Jackie. 2006. Comparison of Stochastic Reserving Methods. *Australian Actuarial Journal* 12: 489–569.

Merz, Michael, and Mario V. Wüthrich. 2008. Modelling the Claims Development Result for Solvency Purposes. *Casualty Actuarial Society E-Forum* 2008: 542–68.

Merz, Michael, and Mario V. Wüthrich. 2012. Full and 1-year runoff risk in the credibility-based additive loss reserving method. *Applied Stochastic Models in Business and Industry* 28: 362–80. [CrossRef]

Taylor, Gregory C., Grainne McGuire, and Alan Greenfield. 2003. *Loss Reserving: Past, Present and Future*. Research Paper No. 109. Australia: University of Melbourne.

Verrall, Richard J. 1989. A State Space Representation of the Chain Ladder Linear Model. *Journal of the Institute of Actuaries* 116: 589–610. [CrossRef]

Verrall, Richard J. 1994. A Method for Modelling varying Run-Off Evolutions in Claims Reserving. *ASTIN Bulletin* 24: 325–32. [CrossRef]

Wüthrich, Mario V., and Michael Merz. 2008. *Stochastic Claims Reserving Methods in Insurance*. West Sussex: John Wiley & Sons.

Wüthrich, Mario V., and Michael Merz. 2013. *Financial Modeling, Actuarial Valuation and Solvency in Insurance*. Berlin and Heidelberg: Springer.