On the study of premium renewal problem in non-life insurance based on two families of customer renewal probability through reinforcement learning

Christian Ngnie Fokoua

Pr Fono Louis Aimé

African Institute for Mathematical Sciences and University of Douala

May 18, 2021

Overview

1. Motivation
   - Problem and Context
   - Previous work
2. Reinforcement Learning
   - General aspect reinforcement learning
   - value functions and Bellman equation
3. Results / Contribution
   - Main results
   - Discussions

IAA AAI ASTIN Non-Life Insurance

Overview

1. Motivation
   - Problem and Context
   - Previous work
2. Reinforcement Learning
   - General aspect reinforcement learning
   - value functions and Bellman equation
3. Results / Contribution
   - Main results
   - Discussions

Problem Statement

Context

Each end of year and based on the recorded claims history, Non-Life Insurance companies examine premiums of their products and strategies to get more tailored ones to the local economic environment. This helps them

Problem Statement

Context

Each end of year and based on the recorded claims history, Non-Life Insurance companies examine premiums of their products and strategies to get more tailored ones to the local economic environment. This helps them

- to be more adapted to customer's needs.

Problem Statement

Context

Each end of year and based on the recorded claims history, Non-Life Insurance companies examine premiums of their products and strategies to get more tailored ones to the local economic environment. This helps them

- to be more adapted to customer's needs.
- to satisfy solvency requirements

Problem Statement

Context

Each end of year and based on the recorded claims history, Non-Life Insurance companies examine premiums of their products and strategies to get more tailored ones to the local economic environment. This helps them

- to be more adapted to customer's needs.
- to satisfy solvency requirements
- to increase their revenues.

Optimization Problem

Optimization Problem

Find $(r_i)_{i\in\mathcal{V}}$, $r_i \in \mathbb{R}$ maximizing:

$$\begin{cases} R(\mathcal{V}) = \sum_{i\in\mathcal{V}} \Phi_i(r_i) * P_i * r_i, \\ T(\mathcal{V}) = \frac{1}{N}\sum_{i\in\mathcal{V}} \rho_i(r_i). \end{cases}$$

Overview

1. Motivation
   - Problem and Context
   - Previous work
2. Reinforcement Learning
   - General aspect reinforcement learning
   - value functions and Bellman equation
3. Results / Contribution
   - Main results
   - Discussions

Literature overview

- Shapiro et al. (2003), Fuzzy clustering and non-linear integer programming for optimal ratio estimation.
- Dutang (2012), logistic regression for lapse rate estimation and deduces customer sensitivity as derivative of lapse rate.
- Dutang et al. (2013) , competition of market, multinomial model for lapse rate estimation, Nash and stackelberg equilibrium as concept solution for optimal ratio estimation.

Literature overview

- Shapiro et al. (2003), Fuzzy clustering and non-linear integer programming for optimal ratio estimation.
- Dutang (2012), logistic regression for lapse rate estimation and deduces customer sensitivity as derivative of lapse rate.
- Dutang et al. (2013) , competition of market, multinomial model for lapse rate estimation, Nash and stackelberg equilibrium as concept solution for optimal ratio estimation.
- Elena et al. (2019), dynamical approach through Q-learning model of reinforcement learning.

Overview

Reinforcement learning

definition

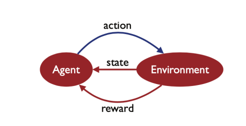
Dynamic study of an environment by an active learner for achieving certain goal.

Reinforcement learning

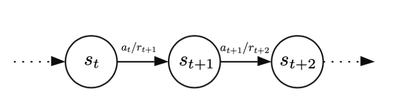
definition

Dynamic study of an environment by an active learner for achieving certain goal.

Reinforcement learning Process



Reinforcement learning Process



Overview

Reinforcement learning

$$V_\pi(s) = E_\pi(\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t = s) \tag{2.1}$$

$$Q_\pi(s,a) = E_\pi(\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t = s, A_t = a) \tag{2.2}$$

Reinforcement learning

$$V_\pi(s) = E_\pi\left(\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \middle| S_t = s\right) \tag{2.1}$$

$$Q_\pi(s,a) = E_\pi\left(\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \middle| S_t = s, A_t = a\right) \tag{2.2}$$

Theorem

Any finite MDP admits an optimal deterministic policy.

Reinforcement learning

Bellman equation

$$\pi_*(s) = argmax_{a' \in \mathcal{A}} Q_*(s, a) \tag{2.3}$$

$$Q_*(s, a) = E(R_{t+1} | S_t = s, A_t = a) + \gamma \sum_{s' \in \mathcal{S}} P^a_{ss'} max_{a' \in \mathcal{A}} Q_*(s', a') \tag{2.4}$$

Stochastic approximation

Q-learning

$$q(S_t, A_t) \longleftarrow q(S_t, A_t) + \alpha(R_{t+1} + \gamma \max_{a'} q(S_{t+1}, a') - q(S_t, A_t)) \tag{2.5}$$

ASTIN Non-Life Insurance

Stochastic approximation

Q-learning

$$q(S_t, A_t) \longleftarrow q(S_t, A_t) + \alpha(R_{t+1} + \gamma \max_{a'} q(S_{t+1}, a') - q(S_t, A_t)) \tag{2.5}$$

Theorem

Consider a finite MDP. Assume that for all $s \in S$ and $a \in A$, $\sum_{t=0}^{+\infty} \alpha_t(s,a) = +\infty$ and $\sum_{t=0}^{+\infty} \alpha_t^2(s,a) < +\infty$ with $\alpha_t(s,a) \in [0,1]$. Then, the Q-learning algorithm converges to the optimal value q*.

Stochastic approximation

Q-learning

$$q(S_t, A_t) \longleftarrow q(S_t, A_t) + \alpha(R_{t+1} + \gamma \max_{a'} q(S_{t+1}, a') - q(S_t, A_t)) \quad (2.5)$$

Theorem

Consider a finite MDP. Assume that for all $s \in S$ and $a \in A$, $\sum_{t=0}^{+\infty} \alpha_t(s, a) = +\infty$ and $\sum_{t=0}^{+\infty} \alpha_t^2(s, a) < +\infty$ with $\alpha_t(s, a) \in [0, 1]$. Then, the Q-learning algorithm converges to the optimal value q*.

SARSA model

$$q(S_t, A_t) \longleftarrow q(S_t, A_t) + \alpha(R_{t+1} + \gamma q(S_{t+1}, A_{t+1}) - q(S_t, A_t))$$

Forward and backward approach

SARSA forward

$$
\begin{cases}
G_t^{(n)} = R_{t+1} + \gamma R_{t+2} + ... + \gamma^n q(S_{t+n}, A_{t+n}), \\
G_t^{\lambda} = (1-\lambda) \sum_{n=1}^{\infty} \lambda^{n-1} G_t^{(n)}, \\
q(S_t, A_t) \longleftarrow q(S_t, A_t) + \alpha(G_t^{\lambda} - q(S_t, A_t)).
\end{cases}
$$

Forward and backward approach

SARSA forward

$$\begin{cases} G_t^{(n)} = R_{t+1} + \gamma R_{t+2} + \dots + \gamma^n q(S_{t+n}, A_{t+n}), \\ G_t^{\lambda} = (1-\lambda)\sum_{n=1}^{\infty} \lambda^{n-1} G_t^{(n)}, \\ q(S_t, A_t) \longleftarrow q(S_t, A_t) + \alpha(G_t^{\lambda} - q(S_t, A_t)). \end{cases}$$

SARSA backward

$$\begin{cases} e_0(s,a) = 0, \forall s \in \mathcal{S}, \forall a \in \mathcal{A}, \\ e_t(s,a) \longleftarrow \lambda\gamma e_{t-1}(s,a) + 1_{S_t=s, A_t=a}, \\ \delta_t = R_{t+1} + \gamma q(S_{t+1}, A_{t+1}) - q(S_t, A_t), \\ q(S_t, A_t) \longleftarrow q(S_t, A_t) + \alpha\delta_t e_t(S_t, A_t) \end{cases}$$

Overview

Portfolio segmentation

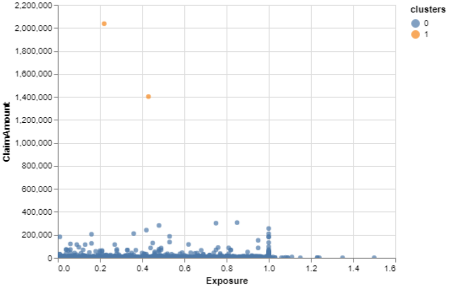


Figure 1: Clusters visualization

Renewal probability

$$p_\mu = \begin{cases} 1, \text{if, } r\pi \leq \pi + \mu\pi \\ 0, \text{else.} \end{cases} \quad r \in [1,2], \text{and} \mu \in [0,1]$$

Results

(a) Evolution of cumulative rewards using Logistic model

(b) Evolution of cumulative rewards using intuitive probability model

Figure 2: SARSA results

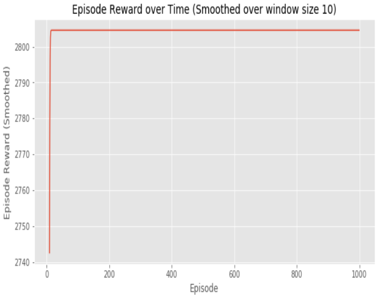
Reinsurance

Renewal probability with reinsurance

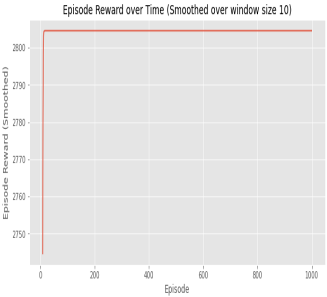
Let C_a the claims amount of a customer i. By considering an excess of loss contract a Xs b, the new model for estimating renewal probability is defined for customer i by:

$$\rho_\mu = \begin{cases} 1, \text{if}, r\pi \leq \pi + \mu\pi \text{ or } C_a > b, \\ 0, \text{else}. \end{cases}$$

Results with reinsurance considerations



(a) Evolution of cumulative rewards using intuitive probability model



(b) Evolution of cumulative rewards considering reinsurancel

Figure 3: SARSA results with reinsurance considerations

Overview

1. Motivation
   - Problem and Context
   - Previous work
2. Reinforcement Learning
   - General aspect reinforcement learning
   - value functions and Bellman equation
3. Results / Contribution
   - Main results
   - Discussions

Summary

- Optimization's curse in Q-learning convergence
- Rapid convergence of Backward SARSA got from new probability
- Estimation of sensitivity and new definition of strategy concept
- Contribution of reinsurance in renewal optimal strategies

Conclusion and discussion

In conclusion our model, improved by reinsurance considerations, found for each customer the best renewal ratio solving the trade-off between revenue and retention.

References

Arnold Shapiro, Lakshmi Chand Jain.
Intelligent and Other Computational Techniques in Insurance.
World Scientific Publishing, 2003.

Christophe Dutang.
The customer, the insurer and the market.
Bulletin Français d'Actuariat. 2(1):50–74, 2012.

Christophe Dutang, Hansjoerg Albrecher, Stephane Loisel.
Competition among non-life insurers under solvency constraints: A game-theoretic approach.
European Journal of Operational Research, Elsevier. 3(1):702–711, 2013.

Elena Krasheninnikova, J. Garcia, R. Maestre.
Reinforcement learning for pricing strategy optimization in the insurance industry.

Thank you for your attention!