

研究論文

Group fused lasso による料率区分の自動セグメンテーション

野村 俊一 *

2017 年 1 月 24 日投稿

2017 年 10 月 24 日受理

概要

本稿では、スパース回帰手法である group fused lasso を利用して、リスクファクター内の料率区分を同等のリスク水準をもつグループへと自動的にクラスタリングする新たな料率算定手法を提案する。タリフ分析では、信頼できる最良推計を得るために、多数のクラスをもつ料率ファクターを少数のカテゴリーへと再グループ化することが多い。しかしながら、料率区分の分割方法の組合せは膨大になることが多く、そのため可能な全ての分割方法を検討することは計算量的に困難である。このような状況下で、fused lasso と呼ばれる L1 正則化法は、推定の過程の中でリスク水準に有意差のない隣接区分を自動的に統合してくれるため非常に適している。ここでは特に、料率ファクター間の交互作用がある場合や、クレーム頻度と規模を同時にモデル化する場合の料率算定のための一般化線形モデルに対するグループ正則化項を導入する。正則化項付き対数尤度からのモデルパラメータ推定には交互方向乗数法を利用し、また、正則化パラメータは交差確認法により選択する。解析例として、提案手法を日本の自動車損害賠償責任保険のクレームデータへと適用し、都道府県を同水準のクレーム頻度または規模をもつクラスターへとグループ化を行う。

キーワード：タリフ分析，一般化線形モデル，Group fused lasso，交互方向乗数法，自動車損害賠償責任保険

1 はじめに

保険業界では古くから、保険契約ごとの保険料を決めるのにタリフ (tariff) と呼ばれる保険料率表が利用されてきた。タリフでは料率ファクターと呼ばれる保険リスクに関わる属性情報について、各ファクター内で属する料率区分から該当契約の保険料が参照できるようになっている。

タリフに定める保険料を過去のクレームデータから算定するための方法論は、現代の推測統計学が発展を遂げる以前から考案されており、その代表的なものとして Bailey(1963) による minimum bias 法がある。その後、Jung(1968) はクレーム頻度についてポアソン分布を仮定して最尤法により推定する手法を提案し、minimum bias 法とともに保険料率算定の実務的手法として利用されてきた。その後、90 年代に各国で保険の自由化が進むと、より柔軟なタリフ分析手法が求められるようになり、Nelder and Wedderburn(1972) による一般化線形モデルが利用されるようになった。Ohlsson and Johansson(2010) などの損害保険に特化した一

* 統計数理研究所 E-mail: nomura@ism.ac.jp