

サンプル選択バイアスを考慮した社会基盤施設の統計的劣化予測モデル

大阪大学大学院工学研究科 学生会員 ○二宮陽平
大阪大学大学院工学研究科 正会員 貝戸清之

1. はじめに

近年、点検データに基づく社会基盤施設の統計的劣化予測手法が多数開発されている。施設の健全性は、健全度と呼ばれる離散的な数値で記録されている場合が少なくないため、健全度を劣化指標とした劣化予測モデルが多数開発されている。中でも、マルコフ劣化ハザードモデルは、施設の健全度に関する時系列データさえ存在すれば、施設の劣化を統計的に予測できる高い汎用性を持ち、実際の施設のマネジメントにも活用されている¹⁾。

一方で、劣化予測に用いる点検データは対象とする施設全体からランダムに獲得されていることが前提である。実際の施設の中には、補修・補強を経験しているものが少なくなく、獲得されている点検データが健全な施設から獲得されたものに偏っている可能性がある。このような点検データのサンプル選択メカニズムを無視して単純な劣化予測を行うことにより、劣化予測結果にサンプル選択バイアスが生じる可能性がある²⁾。

以上の問題意識のもと、本研究では、サンプル選択バイアスを補正する社会基盤施設の劣化予測手法を提案する。まず、点検データに基づいた社会基盤施設の劣化予測モデルを開発する。その上で、社会基盤施設に対して補修や補強が適用されるか否かを、ヘックマンのプロビット選択モデルを拡張してモデル化することにより、点検データの選択バイアスを明示的に考慮する。

2. モデルの定式化

(1) 施設の劣化モデル

対象とするデータが観測されるかどうかランダムではなく、特定の対象のデータのみが観測されるメカニズムは一般に、サンプル選択メカニズム (sample selection mechanism) と呼ばれる。また、サンプル選択メカニズムを無視して単純な推定を行うことで生じる

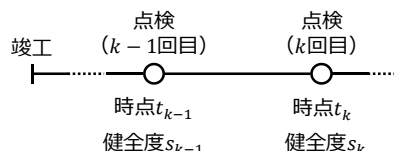


図-1 想定される点検データ



図-2 補修によりサンプル欠損がある点検データ

バイアスは一般に、サンプル選択バイアス (sample selection bias) と呼ばれる。図-1 に本研究で想定するサンプル選択バイアスのない点検データを示す。マルコフ劣化ハザードモデルを用いると、任意の経過時間 z で健全度が i から j に推移する確率は、

$$\begin{aligned} \text{Prob}[s_{k-1} = i, s_k = j, t_k - t_{k-1} = z] \\ = \sum_{k=1}^i \prod_{m=1}^{k-1} \frac{\theta_m}{\theta_m - \theta_k} \prod_{m=k}^{j-1} \frac{\theta_m}{\theta_{m+1} - \theta_k} \exp(-\theta_k z) \quad (1) \\ \theta_1 = \exp(\beta_1), \theta_2 = \exp(\beta_2), \theta_3 = \exp(\beta_3) \end{aligned}$$

と表せる¹⁾。ただし、 $\beta_1, \beta_2, \beta_3$ は未知パラメータである。

(2) サンプル選択バイアスを考慮した施設の劣化モデル

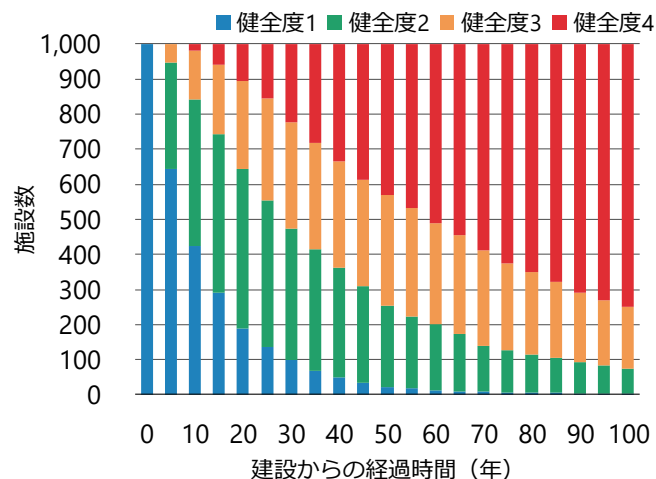
補修選択指標 $u_d(t_k)$ を導入する。 $u_d(t_k) > 0$ の場合、補修により点検データが欠損すると考える。図-2 のような点検データが得られる尤度関数は、

$$\begin{aligned} \ell'_d(\beta | \mathcal{E}_d) = \\ \prod_{k=1}^{K'} \left\{ \text{Prob}[s_{k-1} = i, s_k = j, t_k - t_{k-1} = z | u_d(t_k) \leq 0] \right. \\ \left. \cdot \text{Prob}[u_d(t_k) \leq 0] \right\} \quad (2) \\ \cdot \prod_{k=K'+1}^K \text{Prob}[u_d(t_k) > 0] \end{aligned}$$

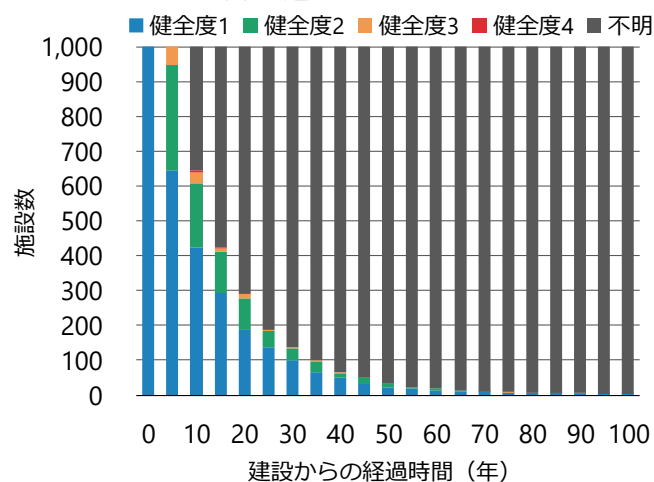
となる。補修選択指標を導入したことにより、サンプル選択バイアスを解消できる。複数の施設 $1, \dots, D$ それぞれに関して獲得されている点検データ $\mathcal{E}_1, \dots, \mathcal{E}_D$ 全体に

キーワード：点検データ、サンプル選択バイアス、劣化予測、アセットマネジメント

連絡先：〒565-0871 大阪府吹田市山田丘 2-1 AR 棟 05 TEL:06-6879-7622



(a) 通常サンプル



(b) 欠損サンプル

図-3 シミュレーションにより発生させた点検データ
関する尤度関数は、

$$\mathcal{L}(\boldsymbol{\beta}|\mathcal{E}) = \prod_{d=1}^D \ell'_d(\boldsymbol{\beta}|\mathcal{E}_d) \quad (3)$$

と定式化できる。

3. 実証分析

シミュレーションにより発生させた点検データに、本研究で開発したモデルを適用する。いま、1,000の施設を管理していると考え、すべての施設は供用から100年が経過しており、各施設の健全性を評価する4段階の離散的な健全度が5年毎の周期で記録されているとする。施設の劣化過程は、マルコフ劣化ハザードモデルに従うとし、それぞれの点検で記録される健全度をモンテカルロ・シミュレーションにより発生させた。

図-1(a)には、発生させた点検データを示している。これを通常サンプルと呼ぶ。一方で、欠損のある点検データを作成するために、健全度2以上が記録されればその次の点検までに補修が実施されるとし、補修後には

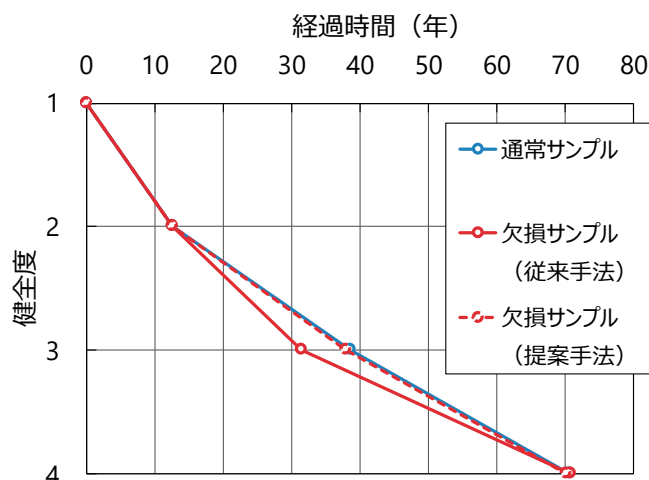


図-4 劣化予測結果

健全度が観測されないと考え、データを除外し、健全度が不明になるとする。この点検データを欠損サンプルと呼び、図-1(b)に示す。

それぞれのサンプルに対して、MCMC (マルコフ連鎖モンテカルロ) 法によるパラメータのベイズ推定 (事後確率分布の推定) を実施した。10,000回のパラメータのサンプリングを実行し、通常サンプルおよび欠損サンプルそれぞれの未知パラメータの推定結果を比較した。

図-2には、推計結果を示している。健全度2→3段階では、寿命の過小評価となり、健全度3→4段階では、寿命の過大評価となった。ただし、健全度4に達するまでの期待寿命が一致しているのはあくまで偶然であると考えられる。両者の結果を比較すると、期待寿命を精度良く推定できているといえる。

4. おわりに

本研究では、点検データの選択バイアスを補正して、劣化予測を行う方法論を示した。提案モデルにより、バイアスを解消した推定がある程度可能となることを示した。

今後の課題として、推計結果の差異を統計的検定によって客観的に評価すること、実際の社会基盤施設の点検データに適用して分析を行うことが挙げられる。

参考文献

- 1) 津田尚胤, 貝戸清之, 青木一也, 小林潔司: 橋梁劣化予測のためのマルコフ推移確率の推定, 土木学会論文集, No.801/I-73, pp.68-82, 2005.
- 2) 星野崇宏: 調査観察データの統計科学—因果推論・選択バイアス・データ融合, 岩波書店, 2009.