

(67) ニューラルネットワークを用いた劣化予測モデルの予測区間推定手法の提案

奥田 知之¹・鈴木 康豊²・神武 直彦³

¹正会員 慶應義塾大学院 システムデザイン・マネジメント研究科
(〒223-8526 横浜市港北区日吉4-1-1協生館)
E-mail:t.okuda@keio.jp

²株式会社パスコ 中央事業部 技術センター インフラマネジメント部
(〒153-0043 東京都目黒区東山1-1-2 東山ビル5階)

³慶應義塾大学院 システムデザイン・マネジメント研究科
(〒223-8526 横浜市港北区日吉4-1-1協生館)

社会資本管理効率化のため、施設情報や点検履歴等の情報をデータベース化した情報基盤を整備し、蓄積した情報を活用する事が求められている。それらの情報から、数理モデルによる社会資本の劣化予測が可能であるが、個々の管理単位に対する精度の高い予測は、観測不可能な変数が相互に影響しており難しい。一方で、社会資本のある区間の状態が将来どの範囲に収まるのか推定する事で、ライフサイクルコストの範囲も推定する事が可能になる。

本研究は数理モデルの中でも、点検履歴を時系列データと捉え、各管理単位区間ごとの予測を生成する事が可能なニューラルネットワークモデルにおいて、ブートストラップ法により信頼区間と予測区間を推定する際に、計算量を大幅に抑えて推定するスナップショット・ブートストラップ法を提案する。

Key Words : *pavement management, pavement condition survey, bootstrap, recurrent neural network, prediction interval, confidence interval*

1. はじめに

近年、社会資本の維持管理を効率化していく事が求められている。インフラ長寿命化基本計画(行動計画)¹⁾において、情報基盤の整備のため一元的に施設情報とその点検履歴を収集するデータベースを整備する事とされており、そのようなデータベースを利活用し、維持管理を効率化する事が求められている。

その様な点検履歴データを数理モデルによりモデル化し、社会資本の管理単位ごとの将来予測を精度高く行う事が可能であれば、点検結果からその都度補修や更新を判断する対処療法的な事後保全ではなく、予防保全によるライフサイクルコスト削減が可能である。しかし、社会資本の劣化は施工、環境、素材等の様々な観測不可能な劣化要因が相互に関係しているため、精度高く予測する事は難しい。一方で、ある区間の劣化が将来どの範囲に収まるか推定する事で、最もライフサイクルコスト削減になる補修時期の範囲や、将来必要な補修コスト範囲の算出等が可能になる。

その様な範囲を求める手法として、本研究では、ニ

ューラルネットワーク(NN: Neural Network)モデルを劣化予測に用いた際の、信頼区間・予測区間の推定手法について提案する。具体的には、ブートストラップ法により生成したサンプルに対し、計算量を抑えて複数のパラメータ推定値を得る事が出来るスナップショットアンサンブル²⁾を応用した提案手法を用いる事で、予測の信頼区間と予測区間を推定する手法を提案する。

2. 劣化予測モデルの信頼区間と予測区間の推定

(1) 信頼区間と予測区間

ある目的変数の観測値と説明変数の組み合わせ $\{d(\mathbf{x}), \mathbf{x}\}$ があるとすると、観測値は次のように生成されると考えられる。

$$d(\mathbf{x}) = y(\mathbf{x}; \mathbf{w}^*) + \varepsilon(\mathbf{x}) \quad (1)$$

ここで、 $y(\mathbf{x}; \mathbf{w}^*)$ は回帰モデルで \mathbf{w}^* はその真のパラメータ、 $\varepsilon(\mathbf{x})$ は誤差を表す。予測モデルのパラメータを観測データにより推定する際には、真のパラメータ \mathbf{w}^* を近似した $\hat{\mathbf{w}}$ を推定すると考える事が出来る。その推

定精度を表す

$$y(\mathbf{x}; \mathbf{w}^*) - y(\mathbf{x}; \hat{\mathbf{w}}) \quad (2)$$

の分布は信頼区間を表す。また、観測値と予測値間の差 $d(\mathbf{x}) - y(\mathbf{x}; \hat{\mathbf{w}})$ の分布は予測区間であり

$$d(\mathbf{x}) - y(\mathbf{x}; \hat{\mathbf{w}}) = [y(\mathbf{x}; \mathbf{w}^*) - y(\mathbf{x}; \hat{\mathbf{w}})] + \varepsilon(\mathbf{x}) \quad (3)$$

と信頼区間を含む形で表すことが出来る。

予測区間は観測値と予測値間の誤差分布を表すため、予測モデルのパラメータの推定精度を表す信頼区間より広い範囲を示し、観測値がどの範囲に収まるかを推定する事が出来る。

(2) 劣化予測モデルの信頼区間と予測区間の推定事例

舗装劣化予測モデルは、長期的な維持管理計画策定のために、マルコフ連鎖モデルを用いる事例が多い³⁾。この手法は同じ傾向があると見なせる区間をグループ化し、平均的な状態の変化を、離散化した健全度の遷移としてハザード関数で定式化する事で、グループに属する区間の期待寿命を短期のデータから求める事が出来る。マルコフ連鎖モデルのパラメータの推定はマルコフ連鎖モンテカルロ(MCMC: Markov Chain Monte Carlo)法が主に用いられる。それにより MCMC で生成したパラメータのサンプルから、任意のパーセンタイルを求める事で、信頼区間を求める手法が用いられている⁴⁾⁵⁾。

一方、個々の管理単位区間に対する補修や更新等の維持管理を効率化するためには、単位区間に対する予測が必要である。そのような予測モデルとして、NN モデルがある。NN では、誤差を 1 次のテイラー展開で近似し、誤差分布を平均 0 の正規分布に従うと仮定する事で、学習データに対するパラメータの傾きから予測区間を求める手法⁶⁾がある。しかし学習データから予測区間を求めるこの手法は過学習している可能性がある。また、ブートストラップ法により NN モデルの信頼区間を推定した例⁷⁾では、学習データ数が 1000 点程度と少ない事から、計算量が限定的である。しかし、データ数が多くなれば計算量が大幅に増加するため、ブートストラップ法により生成した数百以上のブートストラップサンプルに対してそれぞれ NN モデルを学習し、大量のパラメータのブートストラップ推定値を生成する事は困難になる。

そこで本研究は、ブートストラップ法により NN モデルの信頼区間・予測区間を推定する際の計算量増加を抑える手法を提案する。

3. スナップショット・ブートストラップによる信頼区間・予測区間の推定手法

スナップショットアンサンブル²⁾は、NN モデルの学習過程のパラメータ(スナップショット)を記録しておき、

それらの複数のパラメータによりモデルアンサンブルを行う事で、予測精度を向上する為に提案された手法である。一方、本研究ではブートストラップ法を用いる事により、信頼区間・予測区間の推定に用いる事から、以下スナップショット・ブートストラップと称する。

スナップショット・ブートストラップによる信頼区間・予測区間推定アルゴリズムを次に示す。

- ①ブートストラップを用いない学習データに対して、クロスバリデーション等を用いて収束するまで学習し得た NN モデルのパラメータを $\hat{\mathbf{w}}$ 、予測値を $\hat{\mathbf{Y}}$ とする。
- ②学習データのブートストラップサンプルを生成する。学習データを B 回復元抽出し、ブートストラップサンプルを B 個生成する。
- ③パラメータのスナップショット・ブートストラップ推定値を求める。①で得た $\hat{\mathbf{w}}$ を初期値とし、②で得た各ブートストラップサンプルについてそれぞれ数 epoch 追加で学習し、パラメータのスナップショット・ブートストラップ推定値 $\hat{\mathbf{w}}_b^*$ ($b=1, \dots, B$) を生成する。初期値として既に収束済みの $\hat{\mathbf{w}}$ を用いているため、最初からモデルを学習するより大幅に計算量を削減できる。この点が本研究の最も特徴的な所である。なお、本研究の場合は時系列データを対象としているため、ブロックブートストラップにより、単位区間に対する時系列データ全てを 1 つに纏めてリサンプリングを行った。
- ④信頼区間を求める。信頼区間を求めたいある単位区間 o の入力データを \mathbf{x}^o とする。 $\mathbf{y}(\mathbf{x}^o; \hat{\mathbf{w}}_b^*)$ を全ての b に対して出力した結果を並べたベクトルをブートストラップ予測値 $\hat{\mathbf{Y}}^*$ とする。 $\hat{\mathbf{Y}}^*$ を小さい順にソートし、 $0 < p < 1$ に対して pB 番目に小さい数値を $\hat{\mathbf{Y}}^{*(p)}$ と表すと、 $100(1-\alpha)\%$ 両側信頼区間は

$$[\hat{\mathbf{Y}}^{*(\alpha/2)}, \hat{\mathbf{Y}}^{*(1-\alpha/2)}] \quad (4)$$

- ⑤ブートストラップ残差を求める。ブートストラップ残差 \mathbf{R}^* は以下ようになる。

$$\mathbf{R}^* = d(\mathbf{x}^o) - \hat{\mathbf{Y}}^* \quad (5)$$

- ⑥ブートストラップ予測誤差を求める。 \mathbf{R}^* をランダムに並び替えたベクトルを \mathbf{R}' とすると、ブートストラップ予測誤差 \mathbf{E}^* は次式で求められる。

$$\mathbf{E}^* = \mathbf{R}' - \hat{\mathbf{Y}}^* \quad (6)$$

\mathbf{E}^* を小さい順にソートし、 $0 < p < 1$ に対して pB 番目に小さい数値を $\mathbf{E}^{*(p)}$ と表すと、 $100(1-\alpha)\%$ 両側予測区間は次のようになる。

$$[E^{*(\alpha/2)}, E^{*(1-\alpha/2)}] \quad (7)$$

この提案手法の特徴は、NN モデルの信頼区間と予測区間を、予め学習済みのパラメータ $\hat{\mathbf{w}}$ から学習する事で、各ブートストラップサンプルについて最初から学習を行うより計算量を大幅に削減する事が可能な事である。

4. 適用事例

本研究で用いる点検履歴データは、同じ単位区間について、複数回点検を行った履歴がある時系列データと見なす事が出来る。従って、提案手法を時系列データをモデル化できる、再帰的ニューラルネットワーク (RNN : Recurrent Neural Network) モデルを用いた予測時系列生成に適用し、予測区間・信頼区間の推定を行う。

(1) RNNモデル

RNN モデルの概要を図-1 に示す。RNN モデルは NN モデルに時系列方向の接続を追加したモデルである。この時系列方向の接続は、NN モデルの入力層からの層数を l 、時刻を t で表すと、 $t-1$ の l 層の内部状態 $\mathbf{h}_{t-1}^{(l)}$ から、 $\mathbf{h}_t^{(l)}$ への再帰的な接続である。この接続により、 $\mathbf{h}_{t-1}^{(l)}$ を通して過去全ての自身の内部状態の情報を引き継いでいる。それにより時系列データをモデル化し、過去全ての時系列を考慮して現在の予測をすることが出来る。

本研究では、次の時刻の予測値を生成するだけでなく、RNN モデルによる予測時系列データ生成⁸⁾手法により、ある時刻 s 以前 $t \leq s$ のデータを元に、それ以降の予測時系列 $t = s + q (q = 1, \dots, Q)$ を生成した。 \mathbf{x}_t には RNN モデルの時刻 $t-1$ における予測値 $y(\mathbf{x}_{t-1}; \hat{\mathbf{w}})$ に対応する観測値 $d(\mathbf{x}_{t-1})$ が含まれている。予測時系列を生成する際には、図-1 の点線で示した接続のように、 \mathbf{x}_t に含まれる $d(\mathbf{x}_{t-1})$ 、本研究ではわだち掘れ量を $y(\mathbf{x}_{t-1}; \hat{\mathbf{w}})$ 、前時刻の予測わだち掘れ量に置き換える事で、予測時系列を生成した。

また、本研究で用いる点検時系列データのように定常的ではない誤差が大きい事が想定される場合、 q が大きくなるほど予測精度が低下すると考えられる。従って、 $\hat{\mathbf{Y}}^*$ と \mathbf{R}^* 、 \mathbf{E}^* を q ごとに求める事で、 q ごとの信頼区間・予測区間を作成した。

(2) 適用データ

a) データの概要

本研究で用いたデータは川崎市の路面性状調査により取得された 1989 年～2015 年の点検履歴である。入力データ(説明変数)の概要を表-1 に示した。比例尺度のデータは正規化、名義尺度のデータはダミー変数化した。

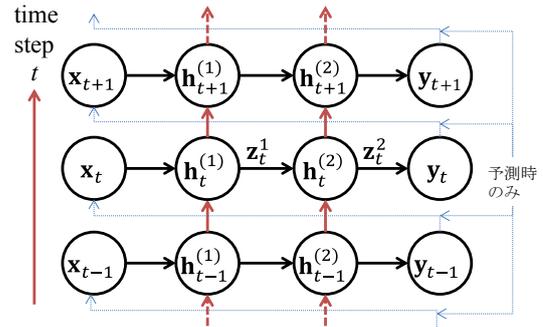


図-1 RNNモデル概念図

表-1 RNNモデルで用いた入力データ

説明変数	前処理
補修後の経過年数	正規化
大型車交通量	
大型車累積交通量	
車道幅員	
予測対象区間の平均わだち掘れ量	
点検時間間隔	
舗装厚	
CBR(路床の支持力比)	ダミー変数化
交差点の有無	
構造物(橋又はトンネル)	
最新の補修工法	

データの管理単位区間は 50m で、施設情報と点検結果が単位区間ごとに記録されている。また、殆どの単位区間で 3 年の点検時間間隔で点検履歴データが存在する。

b) データクレンジング

外れ値除去のため、 t から $t+1$ の間に補修記録がある区間と無い区間を別にし、それぞれで年あたりのわだち掘れ変化量の上下限 5 パーセンタイル外のデータを除去した。

c) クロスバリデーションと評価のためのデータ分割

クロスバリデーションのため、ランダムに全単位区間の 20% を選択して学習時には用いない試験データとし、残りを学習データとした。それにより試験データの単位区間数は 984 区間、学習データは 3,976 区間であり、単位区間の平均時系列数はおおよそ 5 回分である。

また、アスファルト舗装の標準的な設計寿命が 10 年である事、殆どの単位区間において 3 年間隔で点検データが存在する事から、点検 3 回分、9 年後までの予測を評価する事とした。そのため、試験データにおいて最新の点検結果である 2015 年から 9 年前の 2006 年以前のデータを元に、2006 年以降を予測し、3 年間隔で 9 年後までの信頼区間と予測区間を求めた。

(3) 信頼区間及び予測区間の推定結果

試験データに属する国道 409 号線の 2006 年以降補修

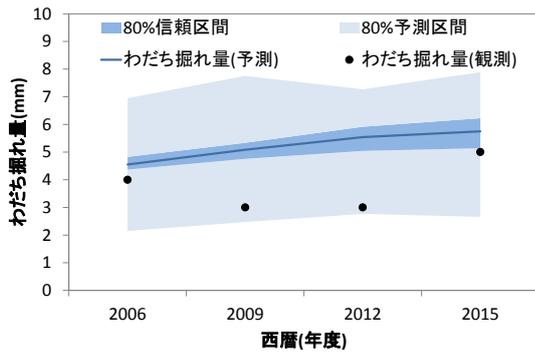


図-2 国道 409 号線(8.650kp)の 9 年予測

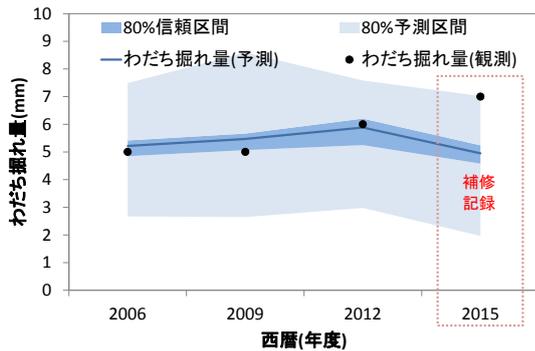


図-3 国道 409 号線(3.600kp)の 9 年予測

の記録が無い単位区間(8.650kp~8.700kp)と、補修の記録がある単位区間(3.600kp~3.650kp)について、提案手法によりパラメータのスナップショット・ブートストラップ推定値を 1000 個生成した結果から 80%信頼区間, 80%予測区間を求め、観測値と比較した結果をそれぞれ図-2, 図-3 に示す。補修記録の無い図-2 は、2006 年以前の点検履歴 6 時刻分のデータを用いて、3, 6, 9 年後を予測している。ただし 2006 年時点はそれより前の 5 時刻分のデータから予測した結果である。年数が経過するにつれて徐々にわだち掘れ量が悪化していく様子が予測されている。それにつれて、 q ごとに信頼区間を作成しているため、信頼区間が拡大することが解る。また通常、何らかの補修が無ければわだち掘れ量が改善する事は考えにくい。観測値の値が予測区間内ではあるものの増減しており、測定誤差のばらつきにより観測値が増減したと考えられる。

一方、2015 年に補修記録がある図-3 において、2015 年の予測は、入力データの補修後の経過年数が 0 年であるため、補修によりわだち掘れ量が回復する状況を予測している。また、補修があったという情報により信頼区間が縮小していることが特徴的である。しかし、川崎市の路面性状調査履歴では、補修記録は年度単位のみで記録されている事から、点検年度と補修年度が同じ場合、補修が点検の前後どちらに行われたかを判断する事が出来ない。この場合、2015 年度の観測値のわだち掘れ量

は回復していない事から、点検後に補修された可能性があり、この点における観測値と予測値が乖離する要因となっている。

5. まとめ

本研究では、NN モデルについてブートストラップ法により信頼区間・予測区間を推定する際の計算量を大幅に抑える事が可能なスナップショット・ブートストラップ手法を提案した。提案手法を RNN モデルによる川崎市のわだち掘れ量点検履歴予測に適用し、信頼区間・予測区間の推定が可能であることを確認した。また、ブートストラップ残差を予測時系列数 q ごとに求める事で、 q ごとの予測精度変化を推定出来る事が確認できた。

謝辞：本研究を遂行するにあたり、川崎市の貴重な路面性状調査履歴データを使用させて頂いた。ここに記して謝意を表す。

参考文献

- 1) 国土交通省：インフラ長寿命化基本計画（行動計画）, p. 55, 2014.
- 2) Huang, G., Li, Y., Pleiss, G., Hopcroft, J. E., and Weinberger, K. Q. : Snapshot ensembles : train 1, get m for free, *ICLR 2017*, pp. 1-14, 2017.
- 3) Abaza, K. A. : Back-calculation of transition probabilities for Markovian-based pavement performance prediction models, *International Journal of Pavement Engineering*, Vol. 17, No. 3, pp. 1-12, 2014.
- 4) 貝戸清之, 小林潔司：マルコフ劣化ハザードモデルのベイズ推定, 土木学会論文集 A, Vol. 63, No. 2, pp. 336-355, 2007.
- 5) 全邦釘, 瀬戸大輔, 大賀水田生：確率分布表現されたマルコフ連鎖モデルによる劣化予測区間の導出とベイズ理論による個体劣化特性の反映, 土木学会論文集 A2 (応用力学), Vol. 68, No. 2, pp. 771-781, 2012.
- 6) Chrysolouris, G., Lee, M., and Ramsey, a. : Confidence interval prediction for neural network models., *IEEE transactions on neural networks*, Vol. 7, No. 1, pp. 229-232, 1996.
- 7) Goh, W. Y., Lim, C. P., Peh, K. K., and Subari, K. : Application of a recurrent neural network to prediction of drug dissolution profiles., *Neural Computing & Applications*, Vol. 10, No. 4, pp. 311-317, 2002.
- 8) Graves, A. : Generating Sequences with Recurrent Neural Networks, *arXiv*, pp. 1-43, 2013.