

第IV 部門

ベイジアンネットワークを用いた道路巡回イベントの発生構造に関する考察

京都大学工学部 学生会員 ○由 雅涵
 京都大学大学院 正会員 松島 格也
 京都大学大学院 フェロー 小林 潔司

1 はじめに

2017年3月に公表した新しい舗装点検要領によると、路面性状測定車を用いる路面性状調査は原則として行わず、かわりに路面の状態を原則として目視点検により把握することとなった。その場合従前の路面性状調査と同様の精度を持ったデータが十分に確保される保証はない。したがって、路面性状調査データに頼らない新たな舗装マネジメントの方法論の提案が求められている。土木施設のアセットマネジメントでは、ライフサイクルコストを限られた予算のもとで低減化する最適な維持管理戦略を求めること必要である。

本研究では、複数の変数間における因果関係や相関構造を推計するベイジアンネットワークを用いる。ベイジアンネットワークは、複数変数間の依存関係を条件付き確率で表現したネットワークモデルである。グラフ構造は既存の知識により事前に作成されることが多く、このグラフ構造に基づき、確率推論を行う。確率推論では、確率変数間の条件付き確率の積から、全ての確率変数の生起確率を計算することができる。変数間の依存関係を条件付き確率で表現することにより、柔軟なモデルを作成することが可能である。ベイジアンネットワークを用いて舗装の各種の変状間の因果関係を把握すれば、目視点検範囲を絞ることができ、コストを下げる事が可能である。

本研究では、これまでに蓄積されてきた日常巡回記録を用いて、現場の技術者が日常巡回の記録に基づき容易に修繕の優先順位や修繕の計画を決定することができる舗装マネジメントモデルの構築を目指して、ベイジアンネットワークを用いた道路舗装変状の発生構造を分析する。

2 ベイジアンネットワークモデル

2.1 ベイジアンネットワーク

今、 $x = \{x_{1t}, \dots, x_{Nt}\}$ を離散 N 変数集合とし、各変数 x_{it} は r_{it} 個の状態集合の中から一つの値をとるとする。ここで、変数 x_{it} は変状の種別が $i \in N$ で、変状が発見された日付の順序が $t \in T$ のときに、値 k をとる時に $x_{it}=k$ と表している。変状間の因果関係を表現する際に変状間の時間的な関係を明示的に考慮するために、変数 x_{it} には時点を表す添え字 t がついている。ベイジアンネットワークの構造 B_s が与えられ、変数 x_{it} の親ノード集合を $\Pi_{it} \subseteq \{x_{1t}, \dots, x_{Nt}\}$ とするとき、変数 $x = \{x_{1t}, \dots, x_{Nt}\}$ の同時確率分布を以下の式のように表す。

$$P(x_{1t}, \dots, x_{Nt}) = \prod_{i=1}^N P(x_{it} | \Pi_{it}) \quad (1)$$

道路舗装の変状に影響をもたらす因子としては、構造条件、環境条件等の様々なものが考えられる。変状間の因果関係に関するグラフ構造が未知であるため、観測されるデータを用いてそのグラフ構造を推定する。データからグラフ構造を決定することをベイジアンネットワークの統計的学習と呼ぶ。式(1)を全ての道路区間に関して積をとることにより、モデルの対数尤度は式(2)と求められる。ただし、 P_{ijk} を親変数集合が j 番目のパターンをとった時変数 x_{it} が値 k をとる条件付き確率パラメータとする。 P_{ijk} ($i = 1, \dots, n, j = 1, \dots, q_{it}, k = 1, \dots, r_{it}$)、 r_{it} は変数 x_{it} がとり得る値の数、 q_{it} は Π_{it} の数、 N_{ijk} は $\Pi_{it} = j$ のとき $x_{it} = k$ となるデータ数を示している。

$$L = \sum_{i=1}^n \sum_{j=1}^{q_{it}} \sum_{k=1}^{r_{it}} N_{ijk} \log_{10} P_{ijk} \quad (2)$$

2.2 モデル選択に関する情報量基準

ある現象を説明するグラフ構造が複数存在するとき、その中から情報量基準をもとに最適なモデルを選択する。評価基準としてよく用いられる情報量基準にはAIC, MDL(= Minimum Description Length)がある。AICは「より適合し、しかし短く」の思想から定められた指標であるが、データサイズが大きくなればなるほど複雑なモデルが有利になってしまう恐れを含んでいる。そこでデータサイズも一種のコストと見なした指標がMDLである。指標値MDLはモデルの尤度を L 、パラメータ数を k 、データサイズ n を用いて以下の式のように表される。

$$MDL = -\log L + \frac{k}{2} \log n \quad (3)$$

AIC同様に指標値が小さいほど優れたモデルであると言える。AICと比較すると、データサイズが大きくなると記述の複雑なモデルは評価されなくなる特徴を持つ。今回の研究においては、過度な学習を防ぐために評価基準としてMDL(BICスコア)を採用した。

2.3 グラフ構造推定アルゴリズム

上記のスコアを最小化するために、greedy hill-climbing algorithmを用いて逐次探索を行う。具体的には、現在のグラフ構造のスコアと、その近いグラフ構造のスコアと比較し、高いスコアのグラフを保存することにより、逐次最適なグラフ構造に近づけていく。近いグラフ構造とは、現在のグラフ構造から、ある変数(ノード)に対して、(a)新しい親変数をつかえる。(b)いま繋がっている親変数をつ取り除く。(c)いま繋がっている子変数とのリンクを反転させる。(d)何もしない。の四つの操作のいずれか一つを行ったグラフ構造のことである。これらのグラフ構造のうち、情報量基準に基づき、最適な操作を選択する。

3 適用事例

3.1 利用データの概要

前章で提案したベイジアンネットワーク手法を用いて、路面性状の変化に関する事象間の因果関係の分析を行う。本研究では、国土交通省近畿地方整備

局大阪国道事務所が管理する国道全線(バイパスを除く)に対して2015年9月から2016年3月の間に行った道路巡回において得られた舗装の変状を対象に分析を行った。巡回記録には、巡回日時、転記、路線名、巡回ルート、距離標、対象、状況といった項目が記載されている。本研究ではそのうち、道路区間を表す「距離標」、記録された事象が該当するものを表す「対象」、各対象における変状の種類を表す「状況」に関するデータを用いて分析を行う。「対象」に含まれる全128種類のカテゴリーの内、舗装と強く関係すると考えられる「アスファルト舗装」「コンクリート舗装」「舗装」「排水性AS舗装」「透水性AS舗装」を分析の対象として抽出する。抽出した対象の変状を表す「状況」としては、「剥離」、「クラック」、「くぼみ」などを合せて合計10種類をとりあげる。当該対象における「状況」の一覧は表1のようにとりまとめられる。

表 1: 各路線の変状の種類と個数

	クラック	剥離	段差	ポットホール	その他
1号	24	17	5	11	3
2号	32	9	1	2	2
25号	125	23	4	23	19
26号	15	27	7	36	15
43号	28	52	7	34	1
163号	15	3	0	2	1
171号	7	18	3	7	8
合計	246	149	27	115	49
	くぼみ	盛り上り	轍ぼれ	波状凹凸	陥没
1号	5	5	3	2	1
2号	3	3	0	0	0
25号	14	1	15	2	0
26号	14	7	4	1	0
43号	5	1	2	0	1
163号	11	0	4	0	1
171号	1	0	0	2	0
合計	50	17	28	7	3

3.2 分析結果

変状に関する確率変数 x_{it} ($1 \leq i \leq 10, t = 0, 1, 2, 3$)を定義する。道路巡回は二日に一回の頻度で実施されているが、変状をとりまとめる道路区間の距離を100mとしたところ、変状が発見された度数は各区分あたり最大で4となったので、 t の値を0から3までに設定した。また、 t の値が大きいほど変状が比較的に新しく発見されることを意味する。モデルの対数尤度を式(2)から出して構造のMDLスコア式(3)に基づいて最適なグラフ構造 B_s^* を決定する。紙面の都合上、推計したベイジアンネットワークの結果については、講演時に発表する。