

高速道路実測データの機械学習による分析

只木進一

佐賀大学理工学部

概要

日本の高速道路では、およそ 2km 毎にインダクションループ計測器が設置され、流量と速度が計測されている。これらのデータに対して機械学習を行うことで、自然に 3 つの状態への分類が行えることを示す。また、渋滞発生の前兆を捉えられる可能性についても議論する。

Machine Learning Analyses of Observed Highway Traffic Data

Shin-ichi Tadaki

Department of Information Science, Saga University

Abstract

On Japanese highways, induction loop devices are installed approximately every 2 kilometers to monitor traffic flow and speed. By applying machine learning to this data, it is possible to naturally classify the states of the traffic into three categories. Furthermore, the possibility for capturing early signs of congestion is discussed.

1 はじめに

日本の高速道路では、およそ 2km 毎に、インダクションループという計測器が埋設されている。そこでは、通過する車両の数と平均速度を計測している。筆者らは、東名高速道路を中心に、高速道路の車両の流れの全体像を把握するために、当該データの分析を行ってきた [1, 2, 3]。本研究では、機械学習を用いてデータ分析を行い、その特徴を抽出する。

今回用いるデータは、東名高速道路 172.65 キロポスト上り車線において、1996 年に取得されたものである。この地点は、日本坂 PA の西にあたり、2 車線の区間である。下流側に日本坂トンネルがある。データには、走行車線と追越車線の、5 分間の流量と平均速度が記録されている。

2 K 平均法

K 平均法 (K-means method) は、データを自動的にグループ (クラスター) 分けする機械学習の一種である。今回は、流量と速度が 2 車線分あるため、

4 次元のデータをクラスターに分ける。 k 個の中心に対して、データの距離に応じてクラスターへと分割する。中心がクラスターの重心と会うように自動的に調整する。データ点間の距離を用いるため、今回は 4 つの指標のそれぞれの平均値を用いて、規格化している。

なお、異なる k に対して、分割の質を評価し、適切な k を求めることができる。

3 分析結果

3.1 8月のデータ

8月のデータには、はっきりした渋滞が見られる。このデータを K 平均法で分析する。 $k = 3$ とした場合の分析結果を、基本図の形で示したものが図 1 である。

紫の点は、流量が少ない自由流である。平均速度が高いことが、傾きの大きさとして現れている。流量の上限は、概ね 5 分間に 100 台である。青の点は、流量が多い自由流である。また、時系列を見ると、

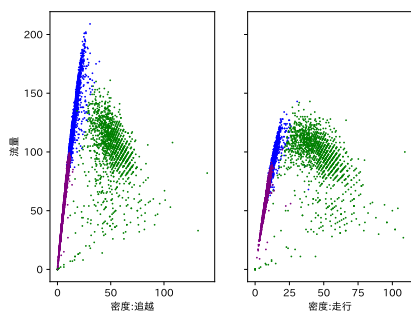


図 1: 8月データの基本図。左は追越車線、右は走行車線。3つのクラスタに分かれている。

時間とともに平均速度が低下し、渋滞へと至る場合が多いことが分かる。

緑の点は、渋滞流である。広い範囲にデータ点が分散しているが、一つのクラスタに分類される。また、追越車線では、青の点と緑の点の間は、きれいに分かれているように見える。一方、走行車線では、二つの領域は連続的になっている。

また、青と緑のデータ点のほとんどでは、追越車線の流量が走行車線のそれを上回る流量逆転が発生している。

3.2 他の月のデータ

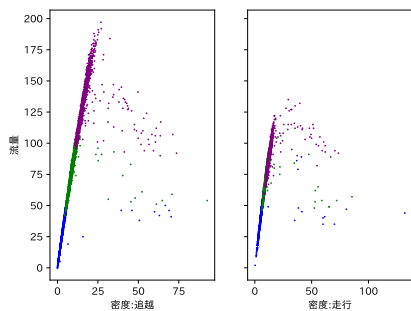


図 2: 2月データの基本図。左は追越車線、右は走行車線。3つのクラスタに分かれている。

当該観測点では、8月には毎日のように渋滞が発生している。一方、他の月ではそれほど渋滞は発生していない。例として2月のデータの分析を行った。 $k = 2$ と $k = 3$ では、分類の質はほぼ同質である。ここでは、 $k = 3$ の場合を示す。図 2 中の紫の点が表すように、流量の多い自由流と渋滞流が一つのクラスタに分類される。

渋滞流に相当するデータ点が少なく、まとまりとして認識されていないと推測される。

4 予測

機械学習によるクラスタリングは、単にデータを分類するだけでなく、訓練されたモデルを使って、未知のデータをグループ分けすることができることが重要である。

そこで、8月のデータで訓練したモデルが、2月のデータを分類の様子を確認する。図 3 に示すように、少ないデータ点である緑の点が、渋滞として分類することが出来た。

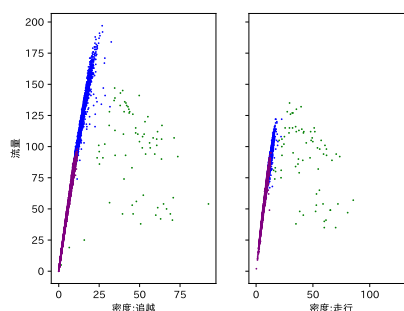


図 3: 8月のモデルを2月データに適用した基本図。左は追越車線、右は走行車線。3つのクラスタに分かれている。

5 議論

最も単純はクラスタリング手法である K 平均法を用いて、高速道路での実測データの分類を実施した。今回は、時系列としての分類を行っていないが、高流量の自由流として分類したデータと速度低下がほぼ一致している。このことから、渋滞発生の前兆も捉えられていると考えることができるのではないかと。

参考文献

- [1] 西成活佑, 林幹久 「東名高速道路における交通量資料集 I」, 交通流数理研究会 (1999).
- [2] S. Tadaki, K. Nishinari, M. Kikuchi, Y. Sugiyama, and S. Yukawa, J. Phys. Soc. Japan, **71** (2002) 2326.
- [3] S. Tadaki, K. Nishinari, M. Kikuchi, Y. Sugiyama, and S. Yukawa, Physica A **315** (2002) 156.