

自動車交通シミュレーションの因子分析と分析結果の考察

内種岳詞¹, 伊藤伸泰^{1,2}

¹ 理化学研究所 計算科学研究機構

² 東京大学大学院工学系研究科

概要

著者らは、神戸市の自動車交通のシミュレーションの実現を目指してきた。都市規模の自動車交通現象では、取り扱う道路ネットワークの規模も自動車の台数の規模も多くなる。それゆえに、シミュレーション結果で観測される多次元データからシミュレーションの妥当性を検証しなくてはならない。そして、次元数が多いことはデータ同化が困難であること、さらには、そのようなデータを収集する困難があることが問題となる。より少ない次元数で自動車交通を説明するために、因子分析を利用した現象の説明が試みられた。そして、数千の自由度を持つシミュレーション結果が、数十の因子とよばれる変数で定式化できることが示された。しかし、因子の数は確率的なシミュレーション実験を重ねることにより精度が変化することが予想される。本稿では、シミュレーションの実験数と因子分析結果の関係を明らかにする。そして、交通現象を説明するモデルの自由度には上限があることを発見した。

Applying Factor Analysis to Results of Vehicle Traffic Simulation and Discussion for Results of the Analysis

Takeshi Uchitane¹, Nobuyasu Ito^{1,2}

¹ Advanced Institute for Computational Science, RIKEN

² Department of Applied Physics, Graduate School of Engineering, The University of Tokyo

Abstract

A goal of this study is to simulate vehicle traffic flow in Kobe-city. To make sure that the simulations are acceptable, it is required to verify the simulation results. When verifying simulation results, we should compare flows on each road in the map and we should get such flow data in the real world. In city scale traffic flow simulations, it is too difficult to compare such flows. It is because that there are a lot of roads on the map and there also are a lot of vehicles on the map. To reduce the comparison costs, we have suggested to apply factor analysis to the result of simulations. From the result of the factor analysis, simulation results which have thousands degree of freedom are formulated only with tens variables called factors. However, it is expected that the number of factors get larger as the number of simulation results become larger. In this paper, we make clear relationship between the number of factors and the number of simulation results. Then, it is found that there may be an upper value of the number of factors.

1 はじめに

自動車交通現象をシミュレーションで再現するためには、道路ネットワークの構造や自動車の発生台数などが与えられなければならない。シミュレーションに利用する道路ネットワークの構造に関しては、十分な精度のデータが入手可能である。Open Street Map[1] に代表されるデジタル地図が広く普及し、無償で利用できることや、より精度の高い地図情報も地図を販売する企業から購入することが可能となっている。一方、自動車の発生台数・時刻、出発地点・目的地点に関しては、入手が困難である。なぜなら、日々の交通活動において変化する可能性があるデータは個人情報と密接に関係しているため、十分な精度のデータを容易に収集・利用することはできない。そのため、入手可能な部分データからさまざまな仮定を置いてデータ同化を実施する必要がある。

自動車交通の地域が街や国家規模となると、データ同化ですら困難となる。たとえば、Asano らは、株式会社ゼンリンの地図と NTT Docomo の人口分布推移データを利用して神戸市の現実の自動車交通をシミュレーションで再現する試みを報告した [2]。神戸市の道路ネットワークに含まれる交差点や道路の数は数千本で、1日あたりの自動車交通も数十万台である。道路交通センサスのデータとシミュレーション結果との比較において、差が十分に小さくならなかった。その原因は、シミュレーション結果が道数と同程度の次元を持ち、多次元の結果を調整が必要とされたためだと考えられる。すなわち、ある道におけるシミュレーション結果を調整しても、他の道の誤差が蓄積され、全体として大きな誤差となった。このような多次元のデータの同化を容易にするために、内種らは、因子分析 [3] により、交通現象シミュレーション結果を次元数の少ないモデルで説明することを提案している [4]。実際に、人工的な Origin-Destination (OD) を仮定した神戸市の自動車交通シミュレーション 2097 次元の結果は、因子分析により 33 変数の線形モデルで説明された。

本稿では、自動車交通シミュレーションの結果に因子分析を適用した分析結果について論じる。内種らは、シミュレーション結果 160 試行に対して因子分析を行ったが、分析に利用するシミュレーションの試行数が増えると、ノイズと有意に区別できる因子の数が増える。前述したように、因子数の増加はデータ同化を困難にするため、因子数と実験数の関

係を明らかにすることは重要であると考える。

2 神戸市自動車交通シミュレーション

本研究では、内種らの行った神戸市自動車シミュレーション [4] の設定を利用する。設定より、シミュレータには、都市規模での動作実績が報告されている SUMO[5] を利用した。シミュレーション結果から通過交通量を測定するためには、測定する道がシミュレーションで定義されていなければならない。そのため、地図データに株式会社ゼンリンの神戸市のデジタル地図を利用した。また、現実の交通を再現することが目的ではなく、因子分析を適用することを目的としたシミュレーションでは、以下に示す人工的な OD の分布が利用された。神戸市都市計画総局による調査結果 [6] によると、神戸市の都心部の自動車交通の特徴は次の 2 点である。

- 都心に流入する自動車交通の 7 割を通過交通が占めている。(1日に通過する 10 万台の内 7 万台は通過交通である。)
- 小規模な駐車場が散在しており、うろつき自動車交通が発生している。

特徴を満たすために、神戸市を図 1 のように分割し、area 間の自動車の OD 発生数を表 1 のように定めた。area 1 から area 4 および area 4 から area 1 への移動は都心の通過交通で、24 時間あたり 7 万台、6 時間あたりでは 17500 台、東西の 1 方向あたり 8750 台通過し、表 1 に太字で示される。また、area 5 と area 6 の間で移動する自動車は都市交通に含まれない。しかし、area 5 と area 6 に背景流が存在しないことは不自然なので、ここでは、他の area 間の背景流と同じ値を設定した。都心交通に含まれない OD 発生数は、表 1 における数字の下線で示される。

シミュレーション 160 試行から得られた各道の通過台数を図 2 に示す。図 2 では、通過台数を縦軸に、各道の通過台数を箱ひげ図にし、箱ひげ図を平均通過台数の多い順に左から並べた。自動車発生台数は同じでも試行ごとに異なる OD 分布のため、道の利用頻度に偏りが存在する。

因子分析では、シミュレーション結果を多変量として扱い、データの相関行列をデータの次元より少ない数の固有値と固有ベクトルで近似する。因子分析における固有値は、固有ベクトルで表現される軸

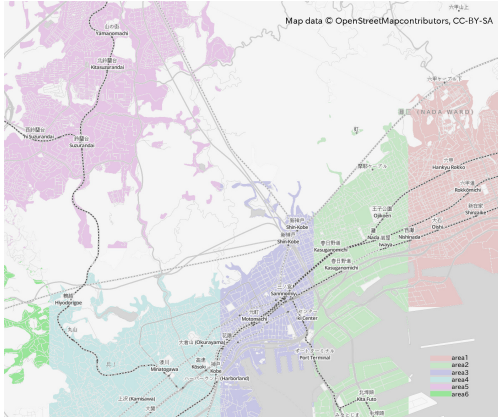


図 1: 神戸市と地域分割 [4]. 地図データ: Open Street Map Contributors[1].

におけるデータの分散に相当する。相関行列は主成分分析のように、データの次元数個の固有値と固有ベクトルに分解できるが、値の小さな固有値に対応する固有ベクトルの軸におけるデータの分散は、因子分析ではノイズとしてみなされる。シミュレーション結果と説明モデルとの誤差であるノイズを最小とするため、最小二乗法を利用した。また、データの次元数に対しデータ点数が小さい場合、理論上では無相関な変数間にもある確率で相関が存在するように推定される。そのため、ノイズと区別できない大きさ固有値の因子を取り除く平行分析 [7] を因子数の特定法として採用した。また、因子負荷量の回転にバリマックス回転を採用した。因子分析をシミュレーション結果に適用した結果、有効な固有値と対応する固有ベクトルで 2097 本の道の交通量の 69% が 33 因子で近似できた。残りの 31% の交通量はノイズとして表現された。

		Departure					
		a1	a2	a3	a4	a5	a6
Arrive	a1	250	250	250	8750	250	250
	a2	250	250	250	250	250	250
	a3	250	250	250	250	250	250
	a4	8750	250	250	250	250	250
	a5	250	250	250	250	<u>250</u>	<u>250</u>
	a6	250	250	250	250	<u>250</u>	<u>250</u>

表 1: 6 時間あたりの自動車発生台数と発生・到着 area[4]. area1 は a1 と略して記載される。

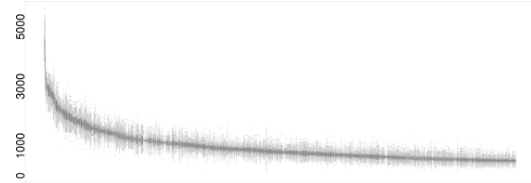


図 2: 平均通過台数が 500 台以上の道 2097 本を降順に並び替え、それぞれの道の自動車通過台数 (縦軸) を箱ひげ図で示した [4].

文献 [4] における得られた因子の考察より、図 3 に寄与率が 1 番と 2 番の因子がモデル化した道路を赤線と青線で示す。なお、それぞれの因子にたいして、正の相関を持つ道が青で、負の相関を持つ道が赤で示される。人工的な OD 分布を利用したにも関わらず、国道や県道の通貨交通を説明する因子が獲得されたように見える。



図 3: 第 1 因子 (左) と第 2 因子 (右) がモデル化した道の分布. 地図データ: Open Street Map Contributors[1].

3 因子分析における試行数と因子数の関係

シミュレーション実験を繰り返し行うことにより、平行分析において、ランダムサンプリングするデータ点数が増えるため、ランダムサンプルデータの相関行列の全ての固有値の大きさは理論上 1 に近づく。よって、シミュレーション試行数が増えると、ノイズと区別できる因子の数が増加する。一方、シミュレーション結果の多変量データの相関行列の固有値は、変数間になにかしらの関係性を含んでいるため、固有値の大きさはなにある分布に収束する。これらのことより、因子数には上限が存在する。シミュレーション試行数と因子数の関係を調べるため、試行数を 160, 500, 1000, 1500, 2000, 2500, 3000, 3200 の

8通り定め、因子分析を行った。図4に実験から得られた試行数と因子数の関係を示す。

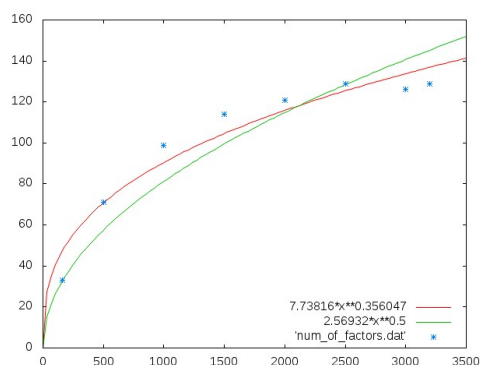


図4: 試行数(横軸)にたいする因子数(縦軸)の増加傾向が点で示される。赤と緑の曲線はデータ点を関数に当てはめた結果である。

図4より因子数は実験数が増えるに連れて増加したことが分かる。しかし、実験数が2500以上の場合、因子数の増加は確認できなかった。これらのことより、2097本の道における交通量は、128変数で説明できると推定される。

実験数と因子数の関係をモデル化するために、因子に関係する固有値の大きさは分散がカイ自乗分布に従っていると仮定し、べき関数に因子数と実験数を当てはめた。図2に、乗数を変数として当てはめた場合と平方根を仮定した場合との関数を示す。データ点と、当てはめた関数を比較すると、シミュレーション試行数が2500より少ない範囲では因子数の方がモデルより多く、シミュレーション試行数が2500より多い範囲では、因子数の方がモデルより少なく見積もられている。このことより、因子数の増加傾向をべき関数で近似することは困難であり、よりよいモデルの提案が必要である。

4 まとめ

神戸市の自動車交通シミュレーション結果に因子分析を適用することにより、道を通る自動車台数を128因子で近似できた。シミュレーション試行数を増やし発見された因子数128は、神戸市に存在する道の数より十分少なく、データ同化を行うときに評価する道数を大きく削減できる可能性がある。しかし、シミュレーション試行数と因子数との関係については、明確なモデルを得ることができなかった。

本稿の目的は因子分析の適用結果を議論すること

であり、人工のOD分布を利用したシミュレーション結果は現実の神戸市の交通現象とは異なる可能性がある。しかし、シミュレーション結果の因子分析結果から、神戸市における主要な道と考えられる国道や県道が因子として発見できたことから、シミュレーション結果が神戸市の自動車交通現象の本質を捉えている可能性はある。

謝辞

本研究は、JST CRESTの支援を受けたものである。本研究での、シミュレーション3200試行の実行・管理にシミュレーション実行・管理ソフトOACIS[8]を利用した。ここにOACISの開発者らに謝意を表す。地図画像の作成に、理化学研究所計算科学研究機構 稲岡創の助力を得た。ここに謝意を表す。

参考文献

- [1] OpenStreetMap, <http://www.openstreetmap.org>
- [2] Yuta Asano, Nobuyasu Ito, Hajime Inaoka, Tetsuo Imai, Takeshi Uchitane, “Traffic Simulation of Kobe-city”, proceedings of SMSEC, Springer, pp. 255–264, (2015).
- [3] 浅野 長一郎, “因子分析法通論”, 共立出版, (1972).
- [4] 内種岳詞, 伊藤信康, “因子分析による都市規模自動車交通シミュレーション結果の解釈”, 計測自動制御学会論文集, (投稿中).
- [5] Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Laura Bieker. “Recent Development and Applications of SUMO - Simulation of Urban Mobility”, *International Journal On Advances in Systems and Measurements*, 5 pp. 128–138, (2012).
- [6] 神戸市都市計画総局, “「都市の概況と人の移動実態」および「想定される課題」”, (2012).
- [7] Horn, John, “A rationale and test for the number of factors in factor analysis”, *Psychometrika*, vol. 30, pp. 179–185, 1965
- [8] Y. Murase, T. Uchitane, and N. Ito, “A tool for parameter-space explorations”, *Physics Procedia*, 57, p73-76 (2014)