

行動的な傾向を判定するための移動経路パターンの分類法

The classification of moving route pattern to judge active tendency

瀬藤 諒

Ryo Seto

岡山大学 大学院自然科学研究科

Graduate school of Natural Science and Technology, Okayama University

1. はじめに

近年、位置情報を用いたライフログが注目されている、これにより、地図上に行動記録を記録するといったアプリ¹⁾や滞在地において LDA (Latent Dirichlet Allocation) を用いて行動分析を行い、各被験者の平日と休日の行動パターンの違いを抽出する研究²⁾も行われている。さらに、これらの研究のように位置情報から人の行動分析を行うことで企業のマーケティングにも活用されている。このように、位置情報から得られるデータを用いて、行動の特徴を抽出することやその人がどのような行動をする傾向であるのかについて行動分析することは今後ますます重要になると考えられる。その中でも、人の移動した経路のデータに対して、移動経路パターンの抽出や分類という研究は少ない。

そこで本研究では、行動診断と題し、その人がどの程度行動的であるのかについて考え、それを判定するための移動経路パターンの分類法についての検証を行う。

2. 本研究の概要

本研究では、行動診断システムというものを考える。これは、位置情報から得られるデータを用いて、人がどの程度行動的であるのかを様々な行動の特徴から判定するシステムである。行動の特徴としては、行動範囲や滞在箇所、滞在時間、滞在頻度などが考えられる。その中で本稿は、行動範囲に着目する。行動範囲とは、人の移動範囲のことであり、日々の移動データを比較することで、その人の移動パターンを分類できると考える。また、他の人の移動パターン数との比較を行うことで、自分がより行動的であるかもわかるはずである。たとえば、移動パターンを分類し、そのパターン数が多ければ、いろいろな場所へ行っていると捉えることができ、行動的であるといえる。

3. 主成分分析を用いた移動経路パターンの分類法

本稿では、日々の移動データに対して、主成分分析を用いることで、個人がどの程度行動的であるのかを判定するための移動経路パターンの分類を行う。主成分分析とは、多次元データのもつ情報をできるだけ損なわずに低次元に情報を縮約する方法である。よって、まず主成分分析により、固有値 (主成分) を求め、さらに固有値 (主成分) における寄与率累積を求める。固有値 (主成分) と累積寄与率を分析することで、どの程度の情報量で、どの程度次元を縮約することができるのかを推定できる。つまり、ある累積寄与率の数値に対しての主成分の数を求めることで、その数が多ければ移動経路のパターンが多いとみなせ、行動範囲が広く、行動的であるといえる。

4. 実験に用いる移動データの作り方

まず、30秒間隔で取得した GPS データを用いて、個人の移動データを作る。このとき、移動データは隙間が開いてしまうため、直線補間アルゴリズムに従い、Geohash³⁾を用いてデータの補間を行う。Geohash とは世界中を隙間のない長方形ポリゴンの hash で敷き詰め、GPS データである緯度と経度を hash コードと呼ばれる文字列により量子化する方

法である。これにより、移動経路を表す移動データとなる。本研究では、他人との比較評価を行うにあたり全く異なる地域を生活圏としている人同士を比較しても意味がないため、評価対象範囲⁴⁾を設定する。この評価対象範囲に該当する Geohash に対して、左上から順に一つのセルとして考え、配列 (2304 次元=48 次元×48 次元) を作る。そして、日々の移動データと照らし合わせて、1度でも通ったことのある場所 (セル) を 1 とし、1度でも通ったことのない場所 (セル) を 0 とし、配列に格納する。これにより、評価対象範囲における 1, 0 で表される日々の移動データとなる。また、この 1, 0 で表される日々の移動データに対して、行が 1 日ごとの移動データとしたもの (日数)、列が評価対象範囲を示す 2304 次元となる移動データ行列を作成する。

5. 実験と結果

4章で作成した移動データ行列に対して、主成分分析を用いることで、個人がどの程度行動的であるのかを判定するための移動経路パターンの分類を行う。

本稿では、被験者 5 名 (A, B, C, D, E) の 2012 年 11 月の 1ヶ月分のデータに対して、主成分分析を用いて移動経路パターンの分類を行った。また、情報量の割合を示す累積寄与率が 70% を超えるときを主成分の数として定めた。

結果は、表 1 のとおりである。各被験者の主成分の数を比較すると、被験者 B と E は最も少なく、被験者 D は最も多くなっていることが分かる。これより、被験者 D がより行動的であるといえる。

表 1 累積寄与率 70% のときの各被験者の主成分の数

被験者	A	B	C	D	E
主成分の数	7	4	8	10	4

6. まとめ

本稿では、被験者 5 名の日々の移動データを用いて、各被験者がどの程度行動的であるのかを判定するために移動パターンの分類を主成分分析により検証した。

今後は、主成分分析により低次元に縮約できた結果から、移動経路の復元を行う。これにより、その人のよく通る移動経路を推定することができると思われる。

7. 参考文献

- 1) <http://latlonglab.yahoo.co.jp/service/bokumichi.html>, 僕の来た道.
- 2) Katayoun, F., and Daniel, G., "What did you do today? Discovering daily routines from large-scale mobile data", Proceeding of the 16th ACM international conference on Multimedia, pp.849-852, 2008.
- 3) <http://geohash.org/>, geohash.org
- 4) 瀬藤諒他, "他人との行動ログ比較による個人の行動特徴分析", 電子情報通信学会 2014 総合大会 講演論文集, D-9-4, Mar. 2014.