

非負値制約行列因子分解による地震予測の試み

鎌田 暉¹, 段 正楠², 増井 隆治³, 久富 望⁴, 曾我部 舞奈⁵, 大関 真之^{6*}

¹奈良県立青翔高等学校, ²京都大学農学部, ³京都大学工学部, ⁴京都大学大学院情報学研究科, ⁵京都大学大学院医学研究科, ⁶東北大学大学院情報科学研究科

要旨

今まで日本で発生した地震データから、地震が発生する特徴を表すパターンを抽出し、地震予測に用いることができないかを考えた。そのため、2004年2月から2016年の5月に日本で発生した地震のデータを「YAHOO! 天気・災害」から取得し、行列として表現し、非負値行列制約因子分解を用いて解析を行った。その結果、地震発生の典型的なパターンと、突発性に伴う困難を示唆する結果を得た。この研究結果は、未来の技術として繋がることを考えている。

序論 1

本研究の動機

地震予測とは、地震が起きる時、場所、大きさの3つの要素を精度よく限定して予測することである。地震予測が可能になれば、地震に関する被害と安全の度合いを伝えることによって、被害者数を減少させる対策が取れるなど、多くの社会的価値がある。

しかし、気象庁をはじめとした様々な努力にも関わらず、一般的な地震予測は未だ不可能である⁽¹⁾。いくつかの地震の予兆と考えられる説も挙げられているが、実用的なものはまだ存在しない。

そこで本研究では、地震の予兆や発生の原因などの情報を用いず、過去の地震データのみに着目し、地震が発生する特徴を表すパターンを抽出することから、地震予測に繋げる可能性について考えることにした。

データから何らかの規則性を抽出する手法として機械学習がある。機械学習においては、データの背後にある規則性を抽出するために、データを表現する関数に潜むパラメータを推定することが基本方針となる。その手法の一つに、非負値制約行列因子分解 (Non-negative matrix factorization, 以下、NMF と略す) がある。

序論 2

非負値行列因子分解

数字を長方形に並べたものを行列といい、構成する値 (成分) が負の値でない行列を非負値行列という。非負値行列因子分解 (NMF) は、非負値行列を2つの非負値行列の積に変形 (因子分解) する技術であり、非負値のパラメータを引き出すことができる。非負値という制約によって、元のデータを構成する各成分を、正の係数をもつ足し算で表すことになり、地震発生回数のような負値に意味のないデータの解析に対し、特に有用である⁽²⁾。

例えば、コンピュータでは画像を非負値の集まりとしてみなすことができる。そこで、Lee & Seung (1999)⁽³⁾ ではたくさんの顔画像を並べて作った非負値行列に対して NMF を適用している。この結果、それら顔画像は、人が目や鼻として解釈できる顔の部分構造を表す行列と、元の顔画像を復元するために部分構造をどのように足し合わせるかというパラメータの行列の、2つの積に分解されることが示されている (図1)。脳は部分を基礎とした認知を行っているという心理学及び生理学的知見を連想させる、興味深い結果である。

解析方法

地震データの取得と整理

本研究では、2004年3月から2016年5月の147か月分の日本国内における26804件の地震データを「YAHOO! 天気・災害」(地震情報)⁽⁴⁾ から取得し、日付、震源地の緯度、経度を利用した。表1は、そのデータの一部である。

次に、日本地図を分割し、各月に起きた地震の回数をマスごとに数え、147か月分の地震発生回数のデータとした。図2は、分割数を30×30マスに設定したときの、各マスにおける147か月分の合計を表している。

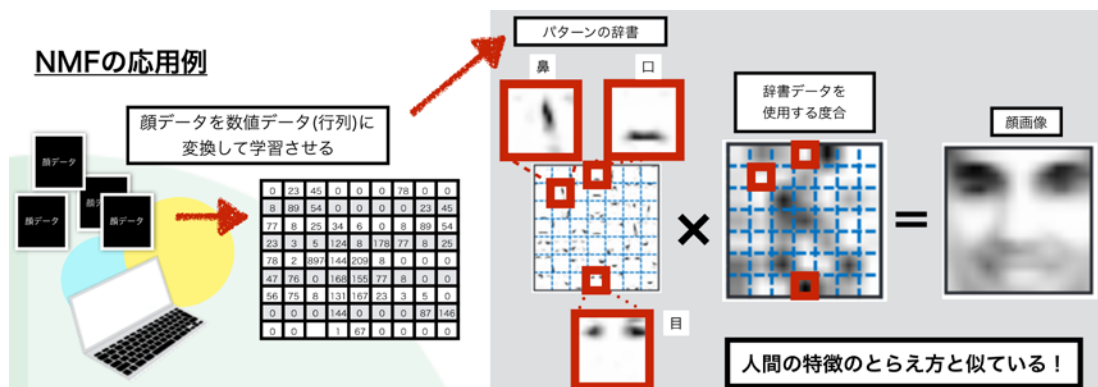


図1. NMFの応用例 (Lee & Seung (1999)⁽³⁾, Figure 1 を利用して編集)。

* 内容に関する連絡先: mohzeki@smapi.is.tohoku.ac.jp

表 1. 地震データの一例.

年	月	日	緯度	経度
2004	3	1	34.4	139.2
2004	3	2	40.8	141.3
		⋮		
2016	5	30	24.3	125.2
2016	5	31	32.9	130.8

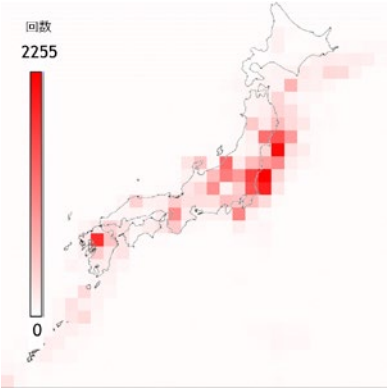


図 2. 147 か月分合計.

NMF による地震データの解析

この 147 か月分の一覧に対して NMF を用いてパターンの抽出を試みた.

まず, 30×30 のデータを図 3 のように列ベクトルに直し, 1 か月目, 2 か月目, 3 か月目...と 147 ヶ月分のデータを並べ, 900×147 の行列 V を作った. そして, NMF に基づき

$$V_{il} \approx (WH)_{il} = \sum_{k=1}^K (W_{ik}H_{kl})$$

という等式を満たすよう, 900 行 K 列の非負値行列 W と, K 行 147 列の非負値行列 H の積に分解することとした (図 4).

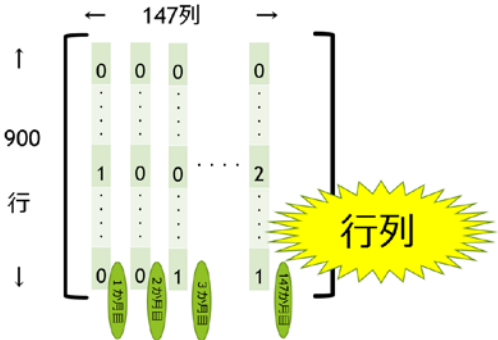


図 3. 行列の仕組み その 1.

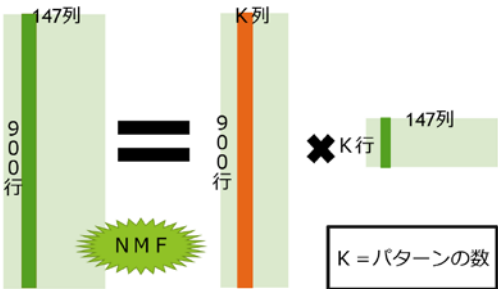


図 4. 行列の仕組み その 2.

ここで, K はパターン数であり, 行列 W の各列は 1 つのパターンを表し, 行列 H の各列は, 各月の出現度を表している. 分解にあたっては, Tikhonov 正則化を行った 2 乗誤差

$$\frac{1}{2} \left\{ \sum_{i,l} (V_{il} - (WH)_{il})^2 + \alpha_W \sum_{i,k} W_{ik}^2 + \alpha_H \sum_{k,l} H_{kl}^2 \right\}$$

を最小にするため Multiplicative Iterative Algorithm [2] を用い, W と H の更新を

$$\begin{aligned} W_{ik} &\leftarrow W_{ik} \left(\sum_l \frac{V_{il}}{(WH)_{il} + \alpha_W W_{ik}} H_{kl} \right)^\omega \\ W_{ik} &\leftarrow \frac{W_{ik}}{\sum_j W_{jk}} \\ H_{kl} &\leftarrow H_{kl} \left(\sum_i W_{ik} \frac{V_{il}}{(WH)_{il} + \alpha_H H_{kl}} \right)^\omega \end{aligned}$$

に従って行った. なお, $\alpha_W = 1$, $\alpha_H = 1$, $\omega = 1.2$ とした.

結果・考察

$K = 5$ の場合の各パターンの出現度を各月毎に表したものが図 5 である. 横軸は地震発生年を表し, 例えば「05」は 2005 年 1 月を表す. 縦軸は各パターンの出現度である. 例えば, 「11」のあたりで急激に出現し, 他と比べて少しずつ大きさが減少したパターンがある. これは 2011 年に発生した東日本大震災に相当するパターンであることが認められる. この他にも, 新潟県中越地震, 熊本地震のような大規模な地震に相当するパターンが認められる. また, 緑の線は 2 回大きく出現するが, これは新潟で起きた 2 つの地震を表しており, 似た地震を同一のパターンとして分類できていることを示唆している.

一方, K が 5 より大きい場合には, 大きな地震が複数のパターンに分けられただけとなった. 一例として, パターン数 $K = 8$ の場合を表しているのが図 6 である.

図 5, 6 に共通して見られることは, 典型的なパターンについては突如現れた後, しばらく尾を引いているという事実である.

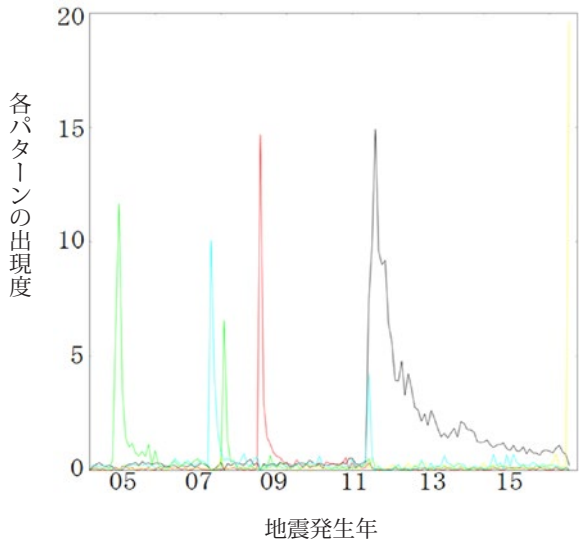


図 5. パターン数, $K = 5$.

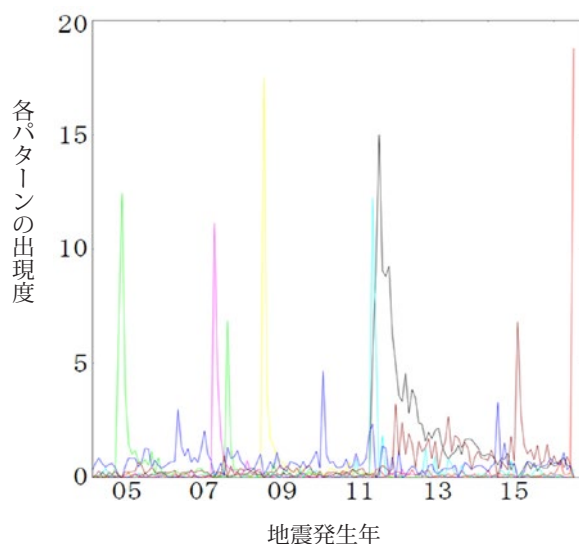


図6. パターン数: $K = 8$.

これは、地震の突発性と余震がしばらく持続的に続く特徴を表している。特に、各パターンが大きく現れる前の出現度の増加はあっても大変短く、そこにある規則性を見出すことは難しい。地震予測という観点で見ると、この突発性の特徴から困難であることを認めざるを得ない。

今後の展望

地震予測を困難にしている突発性の問題はありますが、様々な地域で頻発する日本における地震をいくつかの少ないパターンで分けることができた。本研究の手法は膨大な量のデータを処理できるため、長期間のデータを使用して、どのようなパターンを抽出できるか観察したい。また、日本だけでなく、世界の地震情報を使用することにより、世界中の地震をパターンに分けることも試みてみたいと考えている。

参考文献

1. 気象庁. 地震予知について, <<http://www.jma.go.jp/jma/kishou/known/faq/faq24.html>> [accessed 12 Dec. 2016]. (2016).
2. Cichocki, A. *et al.* Nonnegative Matrix and Tensor Factorizations. Wiley, Hoboken. (2009).
3. Lee, D. D. & H. S. Seung. Learning of the parts of objects by non-negative matrix factorization. *Nature* 401:788–791. (1999).
4. YAHOO! JAPAN, 天気・災害, 地震情報 <<http://typhoon.yahoo.co.jp/weather/jp/earthquake/>> [accessed 12 Dec. 2016]. (2016).

Earthquake Prediction Using Non-negative Matrix Factorization

AKIRA KAMADA¹, ZHENGNAN DUAN², RYUJI MASUI³,
NOZOMU KUTOMI⁴, MAINA SOGABE⁵ & MASAYUKI
OHZEKI^{6*}

¹Nara Prefectural Seisho High School, ²Faculty of Agriculture, Kyoto University, ³Faculty of Engineering, Kyoto University, ⁴Graduate School of Informatics, Kyoto University, ⁵Graduate School of Medicine, Kyoto University, ⁶Graduate School of Information Sciences, Tohoku University

Abstract

Using data collected from past earthquakes in Japan, we extracted patterns representing earthquake occurrence characteristics and explored the possibility of using these for earthquake prediction. We obtained data from “Yahoo! Weather and Disaster” concerning earthquakes that occurred from 2004 February to 2016 May, and expressed them in the form of a matrix. We analyzed the resulting matrix using non-negative matrix factorization. Our results suggest some typical patterns of earthquake occurrence as well as difficulties accompanying sudden outbreaks. We expect that this method will lead to the development of future technologies.

*Correspondence Researcher: mohzeki@smapi.is.tohoku.ac.jp