

マトリクス分解法を用いた周期性感染拡大予測に関する研究

担当責任者 廣瀬 英雄 九州工業大学 教授

研究要旨

感染症の拡大予測方法には、SIR モデル、エージェントモデル、統計的モデルなど、使われる目的に応じてそれぞれ実際に使われている。ここでは全く別の手法からの予測を試みる。感染症の拡大には季節性を伴うものも多い。この性質を、年のトレンドと季節（週など）のトレンドとの組み合わせだと考えると一つのマトリクスを構成することができる。当該年のトレンドを早期にとらえることができたなら、過去の季節性のトレンドの情報を利用して当該年の感染拡大を予測できる可能性がある。このことを実現しているモデルが推薦システムに使われているマトリクス分解法である。

そこで、ここではマトリクス分解法によって季節性の感染症の拡大予測の可能性について検討することとした。これまで、このようにマトリクス分解法を季節性の感染症拡大予測に適用した例は見あたらない。

ここでは、まずノロウイルスによって引き起こされた感染性胃腸炎感染者の予測に疑似シリンダーマトリクスにおけるマトリクス分解法を適用し、早期時期での感染規模予測を行った結果について述べた。更に、デング熱感染拡大の予測についても適用の可能性を検討した。予測結果を季節性感染拡大予測に一般的に用いられている時系列解析結果との比較を行った結果、提案法では従来法よりも良い予測結果が得られている。

この結果により、季節性感染症拡大予測にマトリクス分解法を用いることができる可能性を示した。この方法はインフルエンザ感染拡大予測にも適用できる可能性があり、その簡便性から様々なフィールドに適用することが出来ることを示唆した。

A. 研究目的

感染症の拡大予測方法には、SIR モデル、エージェントモデル、統計的モデルなど、使われる目的に応じてそれぞれ実際に使われている。

ここでは全く別の手法からの予測を試みる。感染症の拡大には季節性を伴うものも多い（図1）。この性質を、年のトレンドと季節（週など）のトレンドとの組み合わせだと考えると一つのマトリクスを

構成することができる(図2)。当該年のトレンドを早期にとらえることができたなら、過去の季節性のトレンドの情報を利用して当該年の感染拡大を予測できる可能性がある。このことを実現しているモデルが推薦システムに使われているマトリクス分解法である。

そこで、ここではマトリクス分解法によって季節性の感染症の拡大予測の可能性について検討することとした。これまで、このようにマトリクス分解法を季節性の感染症拡大予測に適用した例は見あたらない。

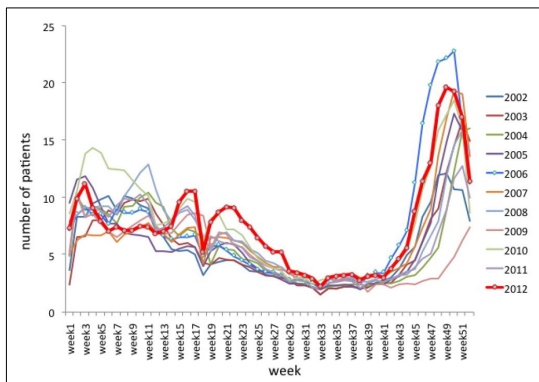


図1. 感染症の拡大例(ノロウイルス)

| | 1週目 | 2週目 | 3週目 | ... | 46週目 | 47週目 | ... | 52週目 |
|------|------|-------|-------|-----|-------|-------------|-----|------|
| 2002 | 4.99 | 6.22 | 6.75 | ... | 8.3 | 9.64 | ... | 7.99 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2009 | 4.98 | 10.23 | 8.24 | ... | 2.68 | 2.86 | ... | 7.39 |
| 2010 | 8.63 | 10.53 | 13.87 | ... | 10.68 | 12.74 | ... | 8.65 |
| 2011 | 7.98 | 8.49 | 9.16 | ... | 4.7 | 5.09 | ... | 9.97 |
| 2012 | 7.33 | 9.87 | 11.21 | ... | 11.39 | データがまだ存在しない | ... | |

図2. 年のトレンドと季節(週など)のトレンドを組み合わせたマトリクス

B. 研究方法

マトリクス分解法は、推薦システムによく用いられている。図3は映画推薦の例である。ユーザが映画を評価し、評価値がマトリクスの対応する要素の値で示されている(通常、1-5の離散評価値)。全ての映画を評価している訳ではないため不完全なマトリクスとなっている。そこ

で、評価済みの要素の値を用いて未評価な要素の値を予測する。

不完全マトリクスでの未評価要素を埋めるためにいろいろな方法が提案されている。最も有望な方法の一つはマトリクス分解法を用いることである。マトリクス分解法は基本的には特異値分解(SVD)の方法を踏襲している。

マトリクス分解法では、観測値マトリクスAをUとVの2つのマトリクスに分解する。UとVの深さを表す指数は潜在要因数と呼ばれている。UとVの未知要素は最小2乗法により求めることができる。実際には観測値と予測値の間の2乗誤差の合計に正則化項を加えたものを最小化することにより行う。この際に使う方法は確率的勾配法である。確率的勾配法では、観測値と予測値の間の誤差を計算しながら、誤差最小になるように最適値の方向に向かって勾配を下りながらUとVを繰り返して求めている。このとき、勾配方向は確率的に決定されている。

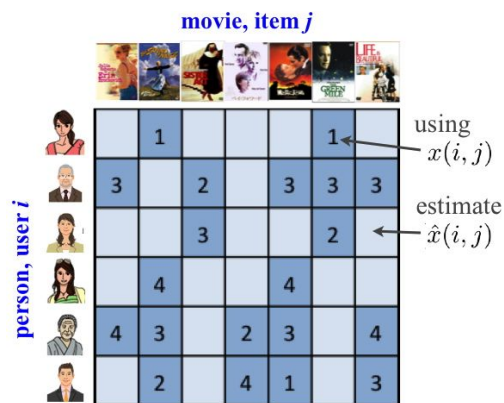


図3. 映画推薦の例

C. 研究結果

C-1 ノロウイルスの場合

図4に2012年ノロウイルス感染症拡大予測結果を示す。2002年から2011年ま

でのデータと2012年の1週目から5週目までのデータを用いて2012年の6週目から52週目までの予測された結果を、2012年の観測データと併せて示す。図から、かなり早期であっても感染拡大が最も広がった2006年のケースに類似した予測ができていることが分かる。

提案したマトリクス分解による感染性胃腸炎の予測が正確であるかどうかを判断するために、他の方法を用いた予測結果と比較した。ここでの比較には季節性時系列解析によく用いられるニューラルネットワークやARIMAを用いた手法によった。比較法としては平均2乗誤差の平方根を用いる。提案法により、2011年の1-5週までのデータをトレーニングデータに使い、6週目以降をテストデータとして予測した結果、マトリクス分解法では季節性が反映された良好な予測結果になっていることが分かる。

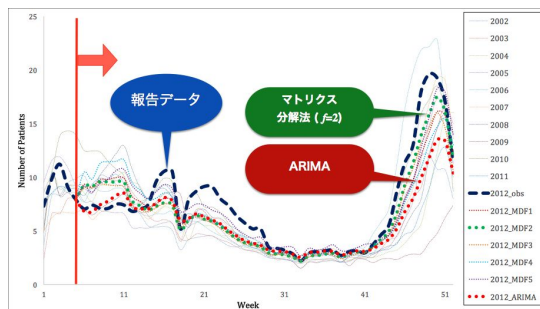


図4. 2012年のノロウイルス感染拡大予測 (2012年5週目までのデータで6週目以降を予測)

C-2 デング熱の場合

2014年、突如として日本に出現したデング熱。季節性を持つが日本には過去のデータが極めて少ない。このような場合、単純に過去のデータをトレーニングデータとして考えることはできないかもしれ

ない。しかし、考えられるあらゆる方法を試みて予兆段階で最終の予測を早期に行うことが極めて重要である。

植物の植生分布に注目したケッペンの気候区分によると、日本と台湾はほぼ同一の温暖湿潤気候と区分されている。そこで、台湾の過去のデング熱感染データを用いることにした(図5)。

マトリクス分解法を用いて予測した結果を図6に示す。観測データの大きさがピークの時点では、観測データよりも少し高めの前測結果になっているが、予防対策の安全性のためには低めに予測されるよりもよいと考えられる。

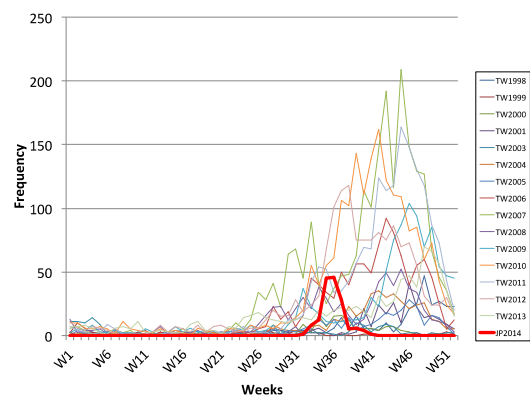


図5. 台湾の過去のデング熱感染データと2014年日本のデング熱データ

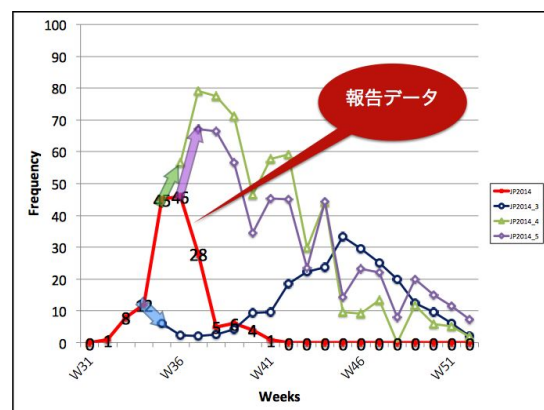


図6. 台湾の過去のデング熱感染データと2014年日本のデング熱データから予測され

た感染拡大

D . 考察

デング熱のデータはまだ日本には少ないため、今回のような手法では予測の精度はよくないと考えられる。この精度を向上させるには、流行のトレンドの時期の微調整を統計的に合理的に行うなどの工夫が必要になると思われる。今後はこのような工夫も加えていきたい。

E . 結論

マトリクス分解法は推薦システムに用いられる最も強力な解法の 1 つである。マトリクス分解法を用いることにより、ユーザの評価値を使って未評価な部分はどう評価されるかを予測することができる。この方法は、映画推薦、音楽推薦、書籍推薦などいろいろな推薦システムに適用されているが、ここでは、まずノロウイルスによって引き起こされた感染性胃腸炎感染者の予測に疑似シリンダーマトリクスにおけるマトリクス分解法を適用し、早期時期での感染規模予測を行った結果について述べた。更に、デング熱感染拡大の予測についても適用の可能性を検討した。

予測結果を季節性感染拡大予測に一般的に用いられている時系列解析結果との比較を行った結果、提案法では従来法よりも良い予測結果が得られている。

この結果により、季節性感染症拡大予測にマトリクス分解法を用いることができる可能性を示した。この方法はインフルエンザ感染拡大予測にも適用できる可能性があり、その簡便性から様々なフィールドに適用することが出来ることを示唆した。

F . 健康危険情報

なし

G . 研究発表

1 . 論文発表

なし（本分担研究は初年度である）

2 . 学会発表

H. Hirose, M. Tokunaga, T. Sakumura, J. Sulaiman, and H. Darwis, "Matrix Approach for the Seasonal Infectious Disease Spread Prediction," 6th Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling (APARM 2014), pp.137-144, August 21-23, 2014, Sapporo, Japan (2014)

H. Hirose, "A seasonal infectious disease spread prediction method by using the singular-value decomposition," The First BMIRC International Symposium on Frontiers in Computational Systems Biology and Bioengineering, February 28 - March 1, 2013, Fukuoka, Japan.

H. Hirose, "Some Methods to Predict Risks Earlier," Ishigaki International Conference on Modern Statistics Theories, Practices, and Education in the 21st Century, page 7, November 9 - 10, 2013, Okinawa, Japan. Invited

H. Hirose, T. Nakazono, M. Tokunaga,

T. Sakumura, S.M. Sumi, J. Sulaiman,
“ Seasonal Infectious Disease Spread
Prediction Using Matrix Decomposition
Method,” the 4th International
Conference on Intelligent Systems,
Modelling and Simulation (ISMS 2013),
pp.121-126, January 29-31, 2013,
Bangkok, Thailand.

H . 知的財産権の出願・登録状況 (予定を含む)

1 . 特許取得

なし

2 . 実用新案登録

なし

3 . その他

なし