

# 非負値行列因子分解を用いた道後平野の地下水データマイニング

## Data Mining for the Variation of Groundwater Level in Dogo Plain

### Using Non-negative Matrix Factorization

○泉 智揮\*・川原千奈\*・中村壮志\*\*・小林範之\*  
○Tomoki Izumi, China Kawahara, Masashi Nakamura, Noriyuki Kobayashi

#### 1. はじめに

道後平野は、重信川とその支流により形成された沖積平野である。瀬戸内式気候に属し、年平均降水量が約 1,300mm の少雨地域にあるため、古くから地下水や伏流水が農業用水として盛んに利用されてきた。しかしながら、近年の臨海工業地区の発展や都市化の進展にともない、水需要の増加、雨水浸透領域の減少、河川流量の減少等による地下水位の低下が懸念されている[1]。そこで、泉・小林(2014)は、データマイニングの一手法である非負値行列因子分解 (NMF: Non-negative Matrix Factorization) [3]を用いて、松山市における降水量や地下水位データの特徴抽出を試みている。本報告では、道後平野における地下水位観測井の時系列データに対して NMF を適用し、その特徴を抽出する。

#### 2. NMF

NMF とは、非負値の観測データを行列で表し、その観測行列を再現するような 2 つの非負値の行列に分解することで、観測行列がもつ特徴を抽出するための技術である。いま、 $n$  次元の  $m$  個の非負値観測ベクトルが表す  $n \times m$  の行列  $V$  を非負値の要素から成る  $n \times r$  の行列  $W$  と  $r \times m$  の行列  $H$  の積で近似できるとすると、行列  $V$  と  $WH$  の関係は次式で表される。

$$V \approx WH$$

or

$$V_{i\mu} \approx (WH)_{i\mu} = \sum_{a=1}^r W_{ia} H_{a\mu}$$

ここで、 $r$  は特徴の数を表し、一意に決まらないため次式を満たすように与える。

$$r \leq \frac{nm}{n+m} \quad (2)$$

行列  $W$  と  $H$  は、 $V$  と  $WH$  の二乗誤差で定義される目的関数(3)式を乗法的更新則(4)式により最小化することで求められる。

$$\|V - WH\|^2 = \sum_{i\mu} (V_{i\mu} - (WH)_{i\mu})^2 \quad (3)$$

$$H_{a\mu} \leftarrow H_{a\mu} \frac{(W^T V)_{a\mu}}{(W^T WH)_{a\mu}}, \quad W_{ia} \leftarrow W_{ia} \frac{(VH^T)_{ia}}{(WHH^T)_{ia}} \quad (4)$$

ここで、(3)式は(4)式のもとで非増加であることが示されている[4]。

#### 3. NMF の適用

本報告では、水文水質データベース[5]と国土交通省四国地方整備局松山河川国道事務所から提供を受けた地下水位データのうち、欠測を含まない長期の時系列データが存在する見奈良観測井 (N33° 47' 26" , E132° 52' 36" ) の地下水位データ (1995 年~2012 年) に対して NMF を適用し、見奈良における 1995 年~1999 年の 5 年間と 2008 年~2012 年の 5 年間の特徴を比較し、その変化について検討する。図 1 に見奈良観測井の位置を示す。

#### 4. 結果と考察

1995 年~1999 年の 5 年間と 2008 年~2012 年の 5 年間の観測データに対して、行を月、列を年とする  $12 \times 5$  の観測行列  $V$  を作成し、 $r = 2$  としたときの NMF の適用結果について、それぞれ図 2、図 3 に示す。ここで、行列  $V$  の成分は、水位低下の特徴を捉えるために、任意に定めた基準面から地下水面までの距離 (低下水位) を表す。行列  $W$  は抽出された特徴を表し、低下水位が大きい月ほど値が大きくなる。行列  $H$  は抽出された特徴の適合度を表し、特徴によくあてはまる年ほど値が大きくなる。図 2 より、特徴 1 として 5 月、11 月、

\* 愛媛大学農学部, Faculty of Agriculture, Ehime University

\*\* 愛媛大学大学院農学研究科, Graduate School of Agriculture, Ehime University

キーワード: 道後平野, 地下水, データマイニング

2月, 6月, 9月の順で, 特徴2として12月, 3月, 8月, 1月の順で低下水位が大きいことが分かる. また, 特徴1には1996年が, 特徴2には1999年がよくあてはまることが分かる. 一方で, 図3より, 特徴1として2月, 1月の順で, 特徴2として9月, 6月, 10月の順で低下水位が大きく, 特徴1には2011年が, 特徴2には2009年がよくあてはまることが分かる. 抽出されたこれらの特徴は, 道後平野における地下水位の一般的な年変動の特徴と一致する. しかしながら, 1995年~1999年と2008年~2012年の特徴を比較すると, 2008年~2012年では低下水位が大きくなる月が絞られていることが分かり, 行列Vの低下水位の値が大きくなっていることを考慮すると, 水位低下事象が極端化している可能性が示唆される.

### 5. まとめ

道後平野に位置する見奈良観測井の地下水位データに対してNMFを適用し, その特徴を抽出した. その結果の一つとして, 近年, 偏った地下水位の低下事象が発生している可能性が示唆された.

### 謝辞

本研究で使用したデータの一部は, 国土交通省四国地方整備局松山河川国道事務所から提供を受けた. ここに記して謝意を表す.

### 引用文献

- [1] 国土交通省四国地方整備局 (2008): 重信川水系河川整備計画, 133p.
- [2] 泉智揮・小林範之 (2014): 非負値行列因子分解を用いた観測データの特徴抽出, 第69回農業農村工学会中国四国支部講演会講演要旨集, pp.61-63.
- [3] Lee, D.D. and Seung, H.S. (1999): Learning the parts of objects by non-negative matrix factorization, Nature, 401, pp.788-791.
- [5] Lee, D.D. and Seung, H.S. (2000): Algorithms for non-negative matrix factorization, In NIPS, 13, pp.556-562.
- [4] 国土交通省 水文水質データベース: <http://www1.river.go.jp/> (2012.12.26)



図1: 観測井の位置

行列V						行列W		
	1995	1996	1997	1998	1999		$r_1$	$r_2$
1月	1.92	1.84	1.39	1.27	2.12	1月	0.7583	0.8266
2月	2.32	2.14	1.63	1.67	1.64	2月	0.9474	0.7470
3月	2.42	1.57	1.41	1.51	1.94	3月	0.6779	1.0319
4月	1.57	1.53	1.38	0.78	1.03	4月	0.7782	0.2819
5月	0.88	2.11	1.48	1.13	0.89	5月	0.9790	0.0048
6月	1.57	1.60	1.89	1.25	0.82	6月	0.9343	0.2256
7月	0.78	0.61	0.79	1.62	1.11	7月	0.2246	0.7671
8月	1.84	1.62	0.95	1.97	1.20	8月	0.5968	0.8577
9月	2.01	2.06	1.48	1.96	1.41	9月	0.8790	0.7340
10月	1.74	1.61	1.25	0.66	1.95	10月	0.6965	0.6177
11月	1.76	1.92	1.94	1.60	1.41	11月	0.9784	0.4993
12月	2.20	1.44	1.59	2.21	1.95	12月	0.6311	1.2038

  

行列H					
	1995	1996	1997	1998	1999
$r_1$	1.28	1.95	1.63	0.87	0.83
$r_2$	1.23	0.30	0.30	1.23	1.27

図2: 見奈良の1995年~1999年の地下水位データに対するNMFの適用結果 ( $r=2$ )

行列V						行列W		
	2008	2009	2010	2011	2012		$r_1$	$r_2$
1月	2.08	1.80	2.26	2.22	1.99	1月	0.8578	0.4710
2月	1.86	1.51	1.40	2.23	1.49	2月	0.8791	0.1587
3月	1.37	1.36	0.96	1.50	1.07	3月	0.5762	0.2102
4月	1.21	1.62	0.93	1.86	1.16	4月	0.6398	0.2040
5月	1.50	2.03	1.30	1.20	2.12	5月	0.3727	0.7577
6月	1.02	2.27	1.07	0.37	1.65	6月	0.0033	0.9691
7月	1.72	0.75	0.55	0.57	0.59	7月	0.4132	0.1172
8月	2.29	0.89	1.87	1.28	1.57	8月	0.6929	0.3206
9月	1.77	2.17	2.35	0.55	1.73	9月	0.1796	1.0769
10月	1.13	1.75	1.86	1.03	1.95	10月	0.2714	0.8210
11月	1.69	1.55	2.19	1.25	2.00	11月	0.4703	0.7181
12月	1.63	2.09	1.99	1.49	1.75	12月	0.4949	0.7223

  

行列H					
	2008	2009	2010	2011	2012
$r_1$	2.09	1.18	1.46	2.33	1.38
$r_2$	1.02	1.92	1.62	0.31	1.69

図3: 見奈良の2008年~2012年の地下水位データに対するNMFの適用結果 ( $r=2$ )