IMPLEMENTATION OF VALID AND STABLE ALGORITHM OF QL1-NMF FOR ANALYZING ENVIRONMENTAL ELF MAGNETIC SIGNALS

Motoaki Mouri[⋆]*, Ichi Takumi*[†] *and Hiroshi Yasukawa*[‡]

[⋆]Faculty of Business Administration, Aichi University
Hiraike-cho, Nakamura-ku, Nagoya, Aichi, Japan
[†]Nagoya Institute of Technology
Gokiso-cho, Showa-ku, Nagoya, Aichi, Japan
[‡]Aichi Prefectural University
Ibaragabasama, Nagakute, Aichi, Japan

ABSTRACT

Previously, we developed two NMF algorithms named QL1-NMF using quasi-L1 norm for analyzing environmental ELF magnetic field measurements. When the data includes many outliers, the QL1-NMF algorithms returned better results than other BSS algorithms using L1 norm. However, the derivative of cost function in our first algorithm was not based on a monotonically increasing function. This problem decreased the validity of algorithm. Our second algorithm had problems about stability though it was based on a monotonically increasing derivative. The method therefore required improvement of validity and stability. In the work described in this paper, we newly introduced the update functions that were based on a monotonically increasing derivative. Computer simulation results and real data's results confirm the new algorithm works more stable than the previous one.

Index Terms— magnetic field, BSS, outlier, L1 norm

1. INTRODUCTION

Observing environmental electromagnetic (EM) signals has many possibilities. One of them is that anomalous radiation has been reported to be a pre-seismic earthquakes [1]. We has been measuring extremely low frequency (ELF) magnetic fields 223 Hz all over Japan since 1985. The amount of data we obtain sometimes becomes large as outliers occur and often includes important information. Then, the ELF measurements are mixtures of signals associated with thunderclouds, human activity, and other phenomena. We need to separate signals for each factor, or extract signals depending on a specific factor. It is therefore important to develop a blind source separation (BSS) method that is appropriate for analyzing the ELF magnetic signals.

This work was supported by The Nitto Foundation and Scientific Research (A) 26249060 from the JSPS

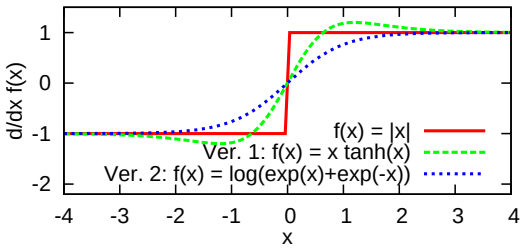

Fig. 1. Derivative of quasi-abs. functions in previous methods

We had concluded that nonnegative matrix factorization (NMF) [2] is suitable for analyzing our data because of its mathematical model [3]. In previous research, we developed two NMF algorithm named QL1-NMF that is based on minimizing the quasi-L1 norm of an error matrix. We found that QL1-NMF worked better than BSS algorithms using L1 norm when the data includes many outliers [4]. However, the differential of cost function in our first algorithm (hereafter call it QL1-NMF1) is not based on a monotonically increasing function. Figure 1 shows the derivative of quasi-absolute functions in previous methods. The red solid lines correspond to absolute function, the green dashed lines correspond to QL1-NMF1 and the blue dotted lines correspond to QL1-NMF2. The distortion of function corresponding to QL1-NMF1 may change the search direction of a solution into an undesirable one. The derivative function corresponding to QL1-NMF2 monotonically increases. However, the algorithm is not stable; the solutions it provides sometimes diverge to infinity. Previously, we controlled stability by small and decreasing step-size parameters. Nevertheless, when the solutions diverged, we restarted algorithm with new initial matrices and smaller step-size parameter.

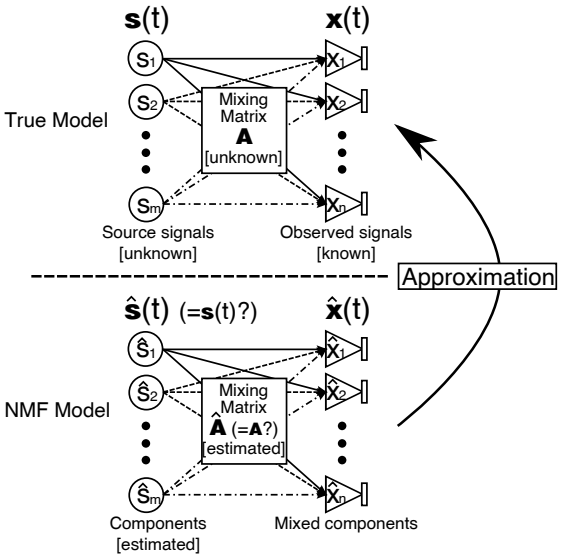

Fig. 2. NMF model

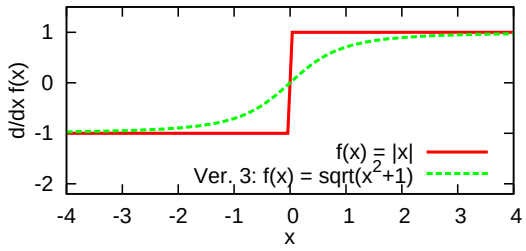

Fig. 3. Derivative of quasi-abs. function in proposed method

2. NONNEGATIVE MATRIX FACTORIZATION BASED ON MINIMIZING QUASI-L1 NORM

The NMF algorithm model approximately factorizes a given nonnegative matrix under nonnegativity constraints. With this model the $n \times T$ matrix $\boldsymbol{X}$, which has only nonnegative values, is approximated by NMF as

$$\boldsymbol{X} \approx \boldsymbol{AS} \quad \boldsymbol{X}, \boldsymbol{A}, \boldsymbol{S} \geq 0 \tag{1}$$

where $\boldsymbol{A}$ is an $n \times r$ mixing matrix and $\boldsymbol{S}$ is an $r \times T$ component matrix. Both $\boldsymbol{A}$ and $\boldsymbol{S}$ have only nonnegative values. The rank of factorization, r, is chosen as $nT > nr + rT$. Eq. (1) can be written column by column as $\boldsymbol{x}(t) \approx \boldsymbol{A}\boldsymbol{s}(t)$, where $\boldsymbol{x}(t)$ and $\boldsymbol{s}(t)$ correspond to the tth columns in $\boldsymbol{X}$ and $\boldsymbol{S}$. This model is an approximation of a linear mixture signal model (Fig. 2).

NMF finds $\boldsymbol{A}$ and $\boldsymbol{S}$ by using iterative updates based on an arbitrary cost function. In order to robustly analyze data including outliers, we newly propose the quasi-absolute function as follows:

$$\mathrm{qabs}_3(x) \equiv \sqrt{x^2 + 1/\alpha} \tag{2}$$

where α is an approximation parameter. The larger α is, the more fit the curve to the absolute function. We set $\alpha = 1$ tentatively. Figure 3 shows the graph of derivative of $\mathrm{qabs}_3(x)$.

Then, the cost function using quasi-L1 norm becomes

$$D(\boldsymbol{X}||\boldsymbol{AS}) \equiv \sum_{i,k} \sqrt{E_{ik}^2 + 1/\alpha} \tag{3}$$

where E_{ik} is the value of the error matrix $\boldsymbol{E} = \boldsymbol{X} - \boldsymbol{AS}$ whose index is (i, k).

The nonnegative update functions based on the cost function Eq. (3) are as follows.

$$A_{ij} \leftarrow (1 - \beta_A)A_{ij} + \frac{\sum_k S_{jk} \frac{X_{ik}}{\sqrt{E_{ik}^2 + 1/\alpha}}}{\sum_k S_{jk} \frac{[\boldsymbol{AS}]_{ik}}{\sqrt{E_{ik}^2 + 1/\alpha}}} \beta_A A_{ij} \tag{4}$$

$$S_{jk} \leftarrow (1 - \beta_S)S_{jk} + \frac{\sum_i A_{ij} \frac{X_{ik}}{\sqrt{E_{ik}^2 + 1/\alpha}}}{\sum_i A_{ij} \frac{[\boldsymbol{AS}]_{ik}}{\sqrt{E_{ik}^2 + 1/\alpha}}} \beta_S S_{jk} \tag{5}$$

where β_A and β_S are adjustment parameters which behave like a step size ($0 < \beta_A \leq 1$, $0 < \beta_S \leq 1$). Also previous algorithms have the same parameters, and we set $\beta_A = \beta_S = 0.2$, and gradually decrease them in iterations in order to improve the algorithm's stability and convergence. When the results were diverged to infinity, the initial βs are decreased to 90 % at the time of algorithm restart. In each iteration, we standardize matrices as follows:

$$S_{jk} \leftarrow \max_j(\boldsymbol{A}) \cdot S_{jk}, \quad A_{ij} \leftarrow \frac{A_{ij}}{\max_j(\boldsymbol{A})}. \tag{6}$$

Hereafter we call this algorithm QL1-NMF3.

3. COMPUTER SIMULATION

First, we generated four source signals of $\boldsymbol{s}(t)$ like shown in Fig. 4. In the figure, the horizontal axis indicates the sampling index and the vertical axes indicate amplitudes. Figure 4(a) shows a large common signal observed at all observation sites. We assumed it as the background signal in the ELF band. Figure 4 (b) and 4 (c) show common signals observed at several observation sites. We assumed them as electromagnetic waves from thunderclouds and artifacts. Figure 4 (d) shows outliers observed at only one observation site.

Second, we generated a mixing matrix. The values corresponding to background signal $s_1(t)$ are almost uniform, while those in the second and third column vary. In the fourth column, only one value is non-zero. We made 12 observed signals $\boldsymbol{x}(t)$ by mixture signal of $\boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{e}(t)$, where $\boldsymbol{e}(t)$

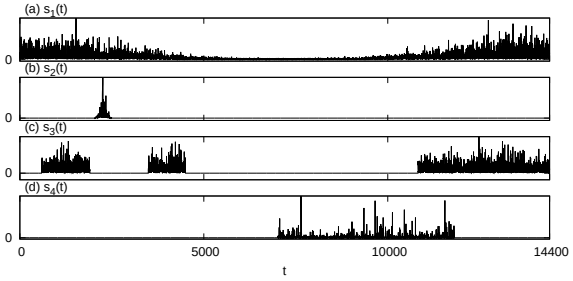

Fig. 4. An example of generated source signals

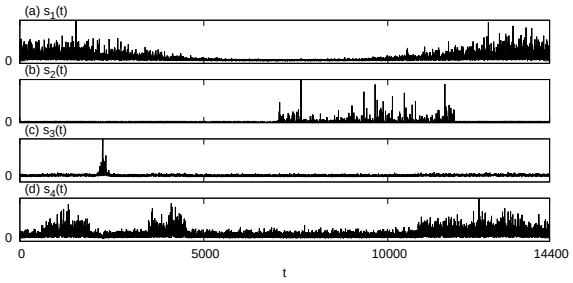

Fig. 5. The components corresponding to Fig. 4 estimated using QL1-NMF3

is the signals distributed absolute Gaussian. We applied NMF to these signals with $r = 4$.

Third, we evaluated the accuracy of the estimated components using a criterion defined as

$$C_j \equiv \frac{100}{n} \sum_{i=1}^{n} \frac{\sum_t \left(A_{ij} s_j(t) - \hat{A}_{ij} \hat{s}_j(t) \right)^2}{\sum_t x_i^2(t)} \qquad (7)$$

where $\hat{A}_{ij}$ is estimated A_{ij}, and $\hat{s}_j(t)$ is estimated $s_j(t)$. Larger values of C_j produce more accurate solutions. We also evaluated the stability of the methods by comparing the number of restarts.

We processed 100 trials using the same procedure. Of course we newly generated $\boldsymbol{s}(t)$ and $\boldsymbol{A}$ by each trial.

The components corresponding to Fig. 4 estimated using QL1-NMF3 are shown in Fig. 5. Though the components are of different orders, they were accurately estimated.

Table 1 shows the averaged values of C_j for 100 trials. The method ISRA [5] is a basic NMF algorithm using L2 norm. The methods PRMF [6] and VSMF [7] are NMF algorithms using L1 norm. The phrases $\beta \leq 0.2$ mean that the method sets $\beta_A = \beta_S = 0.2$ and decreases them in iterations. The phrases $\beta = 1.0$ mean that the method sets $\beta_A = \beta_S = 1.0$ and does not decreases them in iterations. Note that the βs are decreased when the algorithm restart. From this table, the criteria for the QL1-NMF3 are small and almost the same with the criteria for the QL1-NMF1.

Table 1. Averages of accuracy criteria for 100 trials

Method	C_1	C_2	C_3	C_4
ISRA	9.90	8.59	13.57	0.27
PRMF	0.74	4.74	1.98	3.51
VSMF	8.84	10.58	8.10	0.10
QL1-NMF1 ($\beta \leq 0.2$)	0.35	6.35	0.97	0.06
QL1-NMF1 ($\beta = 1.0$)	0.32	5.43	2.74	0.04
QL1-NMF2 ($\beta \leq 0.2$)	0.33	4.54	0.18	0.05
QL1-NMF2 ($\beta = 1.0$)	201.44	71.19	139.65	6.06
QL1-NMF3 ($\beta \leq 0.2$)	0.35	6.35	0.97	0.06
QL1-NMF3 ($\beta = 1.0$)	0.28	4.52	0.27	0.04

Table 2. Restart statistics for 100 trials

Method	Trials with restarts	Average number of restarts
QL1-NMF1 ($\beta \leq 0.2$)	0	0.00
QL1-NMF1 ($\beta = 1.0$)	100	6.95
QL1-NMF2 ($\beta \leq 0.2$)	2	0.05
QL1-NMF2 ($\beta = 1.0$)	100	8.17
QL1-NMF3 ($\beta \leq 0.2$)	0	0.00
QL1-NMF3 ($\beta = 1.0$)	0	0.00

We maintained restart statistics to compare the stability of the methods. The results are shown in Table 2. In the case of using QL1-NMF1 ($\beta = 1.0$), restart were operated in all trials though the case of using the QL1-NMF1 ($\beta \leq 0.2$) did not need restart. This means the solutions were diverged to infinity when βs were large when using previous method. In the case of using QL1-NMF2, the stability were low. However, in the case of using QL1-NMF3, there were no restart. We can conclude the stability of QL1-NMF3 was quite improved from previous methods.

4. APPLYING TO ENVIRONMENTAL MAGNETIC MEASUREMENTS

Figure 6 shows our observed ELF signals (6 of 28 sites) on March 17 in 2005. Each vertical axis indicates the EM energy $[(\mathrm{pT})^2/\mathrm{Hz}]$, and each horizontal axis indicates the time course [hour]. They have common changes that are the effect of background signal. An anomalous signal was observed at Unzen (e) before the 2005 Fukuoka Earthquake (M 7.0, on March 20). This site is 112 [km] from the epicenter.

Figure 7 shows source signals estimated by QL1-NMF3 ($\beta = 1.0$). Each vertical axis indicates the amplitude of signals, and each horizontal axis indicates the time course [hour]. The signal $s_1(t)$ is similar to common change. The signal

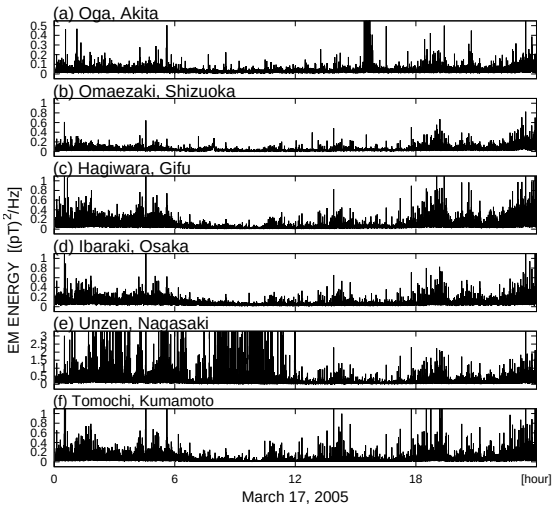

Fig. 6. ELF observed signal

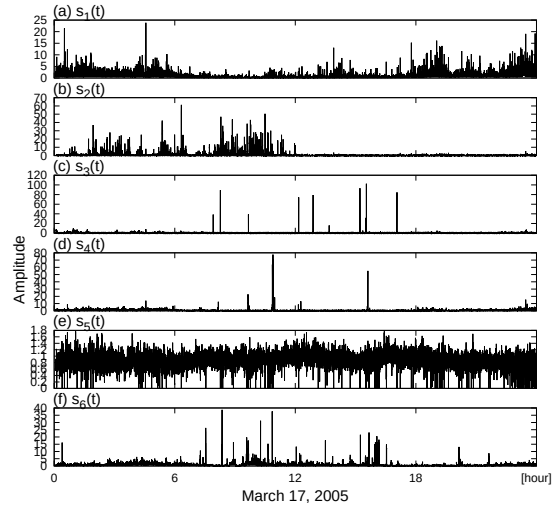

Fig. 7. Source signals estimated by QL1-NMF3

$s_2(t)$ is similar to anomalous signal observed at Unzen. There is a possibility that $s_2(t)$ is earthquake-related source signal.

For numerical comparison, we calculate GICs [3]. The smaller the GIC is, the better the background signal esitimation is. The calculated GICs and number of restarts are shown in Fig. 3. The GICs of QL1-NMF3 were a little worse than QL1-NMF1, but stabilities were quite better.

5. CONCLUSION

This paper detailed how we improved QL1-NMF algorithm and we named QL1-NMF3. The cost function of algorithm was based on a monotonically increasing derivative unlike in the case of QL1-NMF1. About the estimation accuracy of results of QL1-NMF3, they were the same level with QL1-NMF1 in computer simulations, and they were a little worse when applying to ELF observed signals. The stability of QL1-NMF3 was quite improved from previous method.

Subjects for future work will include researching characteristics of QL1-NMF3, e.g. optimizing parameter α. Additionaly, we will apply to other kinds of signals.

Table 3. GICs and number of restarts

Method	GIC	Restarts
ISRA	0.1656	-
PRMF	0.1000	-
VSMF	0.1032	-
QL1-NMF1 ($\beta \leq 0.2$)	0.0939	0
QL1-NMF1 ($\beta = 1.0$)	0.0901	5
QL1-NMF2 ($\beta \leq 0.2$)	0.0938	0
QL1-NMF2 ($\beta = 1.0$)	0.1021	8
QL1-NMF3 ($\beta \leq 0.2$)	0.0996	0
QL1-NMF3 ($\beta = 1.0$)	0.0944	0

6. REFERENCES

[1] M. B. Gokhberg, V. A. Morgunov, T. Yoshino and I. Tomizawa, "Experimental measurements of EM emissions possibly related to earthquakes in Japan," *J. Geophys. Res.*, vol. 87, pp. 7824-7829, 1982.

[2] A. Cichocki, R. Zdunek, A.-H. Phan and S. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, Chichester: Wiley, 2009.

[3] M. Mouri et al., "Global Signal Elimination from Environmental Electromagnetic Signals by Nonnegative Matrix Factorization," *Journal of Signal Processing*, Vol. 14, No. 6, pp. 415-425, 2010.

[4] M. Mouri et al., "Usefulness of Quasi-L1 Norm-Based Nonnegative Matrix Factorization Algorithm to Estimate Background Signal Using Environmental Electromagnetic Field Measurements at ELF Band," *IEEJ Trans. on Fund. Mat.*, Vol. 136, No. 5, pp. 241-251, 2015.

[5] D. D. Lee and H. S. Seung, "Learning the parts of objects by nonnegative matrix factorization, *Nature*, vol. 401, pp. 788-791, 1999.

[6] N. Wang, T. Yao, J. Wang, D. Yeung, "A Probabilistic Approach to Robust Matrix. Factorization," *Computer Vision - ECCV 2012*, LNCS, Vol. 7578, pp. 126-139, 2012.

[7] Y. Li and A. Ngom, "The non-negative matrix factorization toolbox for biological data mining," *Source Code for Biology and Medicine 2013*, pp. 8-10, 2013.