Extracting Time-evolving Latent Skills from Examination Time Series

Shinichi OEDA
Department of Information and Computer Engineering,
Kisarazu National College of Technology
2–11–1 Kiyomidai, Kisarazu-shi, Chiba, 292–0041, Japan
oeda@j.kisarazu.ac.jp

Kenji YAMANISHI
Department of Mathematical Informatics,
The University of Tokyo
7–3–1 Hongo, Bunkyo-ku, Tokyo, 113–8656, Japan
yamanishi@mist.i.u-tokyo.ac.jp

ABSTRACT

Examination results are used to judge whether an examinee possesses the desired latent skills. In order to grasp the skills, it is important to find which skills a question item contains. The relationship between items and skills may be represented by what we call a Q-matrix. Recent studies have been attempting to extract a Q-matrix with non-negative matrix factorization (NMF) from a set of examinees' test scores. However, they did not consider the time-evloving nature of latent skills. In order to comprehend the learning effects in the educational process, it is significant to study how the distribution of examinees' latent skills changes over time. In this paper, we propose novel methods for extracting both a Q-matrix and time-evolving latent skills from examination time series, simultaneously.

1. INTRODUCTION

The relationship between items and skills may be represented by what we call a *Q-matrix* [5]. Its original idea came from the rule space method (RSM) developed by Tatsuoka et al. [5]. The Q-matrix allows us to determine which skills are necessary to solve each item. However, the process of determining the skills involved in a given item is a boring and heavy task. Recently, there exist several studies on how to extract a Q-matrix from a set of examinees' test scores [1, 2]. These studies applied the *non-negative matrix factorization* (NMF) to the problem of establishing the skills from examinee performance data. They were applied to only static examination results provided at a certain time. However, in order to comprehend the learning effects in the educational process, it is significantly important to study how the distribution of examinees' latent skills changes over time. There have been studied on cognitive modeling from student performance over time [3, 4]. From another aspect, we propose novel methods for extracting time-evolving latent skills. It enables us to extract both a Q-matrix and time-evolving latent skills from examination time series, simultaneously.

2. EXTRACTION OF Q-MATRIX WITH NMF

In the Q-matrix extraction with NMF [1, 2], $\mathbf{R}$ represents an observed examination outcome data for m question items and n examinees. We define $\mathbf{Q}$ as a Q-matrix with m items and k skills, while we define $\mathbf{S}$ as an S-matrix with k skills and the n examinees. We assume that $\mathbf{R}$ is factorized into two matrices $\mathbf{Q}$ and $\mathbf{S}$. If Q-matrix is a conjunctive model, an examination result may be obtained according to the equation: $\neg\mathbf{R} = \mathbf{Q} \circ (\neg\mathbf{S})$ [1]. The operator $\neg$ denotes a Boolean negation, which is defined as a function that maps a value of 0 to 1 and any other value to 0. This equation will yield 0 in $\mathbf{R}$ whenever an examinee is missing one or more skills for a given item, and will yield 1 whenever all the necessary skills are mastered by an examinee.

3. Q-MATRIX EXTRACTION FROM EXAMINATION TIME SERIES

We propose an *online NMF* to extract a stable Q-matrix from examination time series in an online fashion. The key idea of the online NMF is that the initial values of matrix inherits those of the decomposed matrix at the previous data. The online NMF runs sequentially every time an examination result is input. When applying the conventional NMF-based method into such a sequential scenario, it must calculate a Q-matrix of which the initial values are set to random ones ignoring the lastest ones. Meanwhile, the online NMF produces a Q-matix letting the initial values be those obtained at the last time. This enables us to learn a Q-matrix in an online fashion.

In order to make the obtained Q-matrix more stable, we further propose an *online NMF with regularization.* In it we add the regularization term to the squared error function in the objective function to be minimized so that the Q-matrix does not change so much. We introduce here a cost function as follows:

$$\min_{\mathbf{Q}_t,\mathbf{S}_t}\{\|\neg\mathbf{R}_t - \mathbf{Q}_t\neg\mathbf{S}_t\|_F^2 + \lambda(t)(\|\mathbf{Q}_{t-1} - \mathbf{Q}_t\|_F^2)\}, \qquad (1)$$

where $\lambda(t)$ is a monotonous increasing function of time. It is defined as $\lambda(t) = \alpha t/T$, where t is a time in $(1,\ldots,T)$, and α is a constant parameter of the increasing rate. At each time t, we find $\mathbf{Q}_t$ and $\neg\mathbf{S}_t$ according to (1), so that the sum of the factorization error and regularization term is minimized. We can do this through an iterative procedure in which each iteration involves two successive steps corresponding to suc-

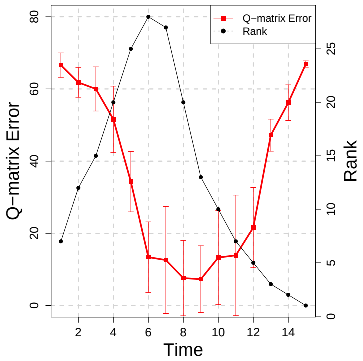

Figure 1: The conventional NMF

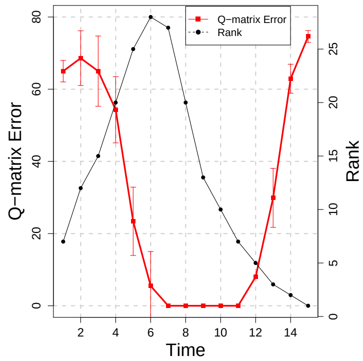

Figure 2: The online NMF

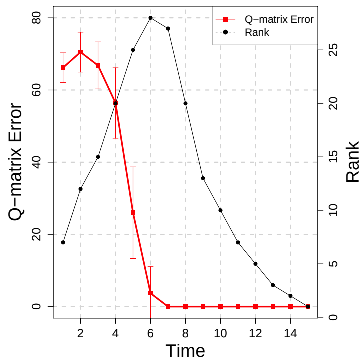

Figure 3: The online NMF with Regularization

cessive optimizations with respect to $\mathbf{Q}_t$ and $\neg\mathbf{S}_t$. First, we set $\mathbf{Q}_t$ by inheriting $\mathbf{Q}_{t-1}$, and choose $\neg\mathbf{S}_t$ by giving random non-negative values. Then, in the first phase, we minimize the cost function (1) with respect to $\neg\mathbf{S}_t$, keeping $\mathbf{Q}_t$ fixed. In the second phase, we minimize the same cost function with respect to $\mathbf{Q}_t$, keeping $\neg\mathbf{S}_t$ fixed. This two-stage optimization is then repeated until convergence.

4. EXPERIMENTAL RESULTS

In order to verify the effectiveness of our methods, we made a synthetic examination time series. We generated a time-varying S-matrix and a fixed Q-matrix to obtain $\neg\mathbf{R}_t$ according to the equation $\neg\mathbf{R}_t = \mathbf{Q} \circ (\neg\mathbf{S}_t)$. A conjunctive Q-matrix consisted of 31 items and 6 skills. We designed a time series of $\neg\mathbf{S}_t$ as a process of acquiring skills, on the basis of the item response theory (IRT) [6].

As a measure of the performance for Q-matrix extraction, we introduce a *Q-matrix error* e_t between $\mathbf{Q}$ and an extracted matrix $\widehat{\mathbf{Q}}_t$ as follows:

$$e_t = \|\widehat{\mathbf{Q}}_t - \mathbf{Q}\|_F^2. \tag{2}$$

Figures 1 and 2 show the experimental results obtained using the coventional NMF and the online NMF. Note that the factorized solutions obtained using the NMF may not be unique due to the randomness of the initial matrices. Hence we calculated the mean of Q-matrix errors from 10-fold simulations and indicated error bars indicating the standard deviations. The Q-matrix errors both in Figures 1 and 2 were large at the initial and the final stages, while the ranks of matrices were small at the same stages. We calculated the rank of each matrix $\neg\mathbf{R}_t$ by means of QR decomposition. As a result, there was a correlation between the Q-matrix error and the rank of matrix. Note that in the conventional NMF, the Q-matrix error did not become zero at any stage, and the standard deviations were uniformly large. In the online NMF, the Q-matrix errors became zero at the middle stage of $t = 7, \cdots, 11$. However, the Q-matrix error gradually increased after $t = 12$. Figure 3 shows the result of the online NMF with regularization. It overcame the problem as above. That is, the Q-matrix error became zero after $t = 12$.

5. CONCLUSIONS

In this paper, we have introduced the online NMF with regularization for the purpose of extracting a Q-matrix and a time-evolving S-matrix from time series of examination results. We have designed it in order to extract a stable Q-matrix in an online fashion. We have employed a synthetic data set to demonstrate that it performs more accurately and in a more stable way than the conventional NMF in the extraction of the Q-matrix.

6. ACKNOWLEDGMENTS

We would like to thank Ryota TOMIOKA and Hisashi KASHIMA for their valuable comments and discussions. This research has been conducted at the University of Tokyo. This work was partially supported by MEXT KAKENHI 23240019, Aihara Project, the FIRST program from JSPS, initiated by CSTP, and NTT Corporation.

7. REFERENCES

[1] B. Beheshti, M. C. Desmarais, and R. Naceur. Methods to find the number of latent skills. *Proceedings 5th International Conference on of Educational Data Mining, EDM2012*, pages 81–86, 2012.

[2] M. C. Desmarais, B. Beheshti, and R. Naceur. Item to skills mapping: Deriving a conjunctive q-matrix from data. *Proceedings of the 11th International Conference Intelligent Tutoring Systems, ITS2012*, pages 454–463, 2012.

[3] J. P. Gonzalez-Brenes and J. Mostow. Dynamic cognitive tracing: Towards unified discovery of student and cognitive models. *Proceedings 5th International Conference on of Educational Data Mining, EDM2012*, pages 49–56, 2012.

[4] C. Studer, B. Junker, and H. Chan. Incorporating learning into the cognitive assessment framework. In *SREE Fall 2012 Conference on Effective Partnerships: Linking Practice and Research*, 2012.

[5] K. Tatsuoka. Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*, 20:345–354, 1983.

[6] W. J. van der Linden and R. K. Hambleton. Handbook of modern item response theory. *Springer*, 1996.