二值矩阵分解的认知建模方法研究

张 猛[1,2] 付丽华[3] 何婷婷[1] 杨 青[1]
(华中师范大学计算机学院 武汉 430079)[1] (华中师范大学教育信息化协同创新中心 武汉 430079)[2]
(中国地质大学(武汉)数学与物理学院 武汉 430074)[3]

摘 要 根据考试反馈数据,提出新颖的逻辑斯提克二值矩阵分解方法,来预测未来的学生考试成绩并自动对考题进行模式分类,同时设计新的算法对建模中遇到的非凸优化问题进行求解。在模拟数据和真实的美国SAT考试数据上进行的实验发现,新方法不仅可以准确地预测学生的考试表现,而且能够将考题按照知识点进行自动模式分类。实验结果表明,新的方法相比经典方法在结果的可解释性和估计精度方面有明显的提升。

关键词 认知建模,二值矩阵分解,考题分类,学生成绩预测
中图法分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.10.048

Cognitive Modeling Based on Binary Matrix Factorization

ZHANG Meng[1,2] FU Li-hua[3] HE Ting-ting[1] YANG Qing[1]
(School of Computer,Central China Normal University,Wuhan 430079,China)[1]
(Collaborative & Innovative Center for Educational Technology,Central China Normal University,Wuhan 430079,China)[2]
(School of Mathematics and Physics,China University of Geoscience,Wuhan 430074,China)[3]

Abstract A novel logistic binary matrix factorization (LBMF) was proposed to predict the students' performance and to classify the exam items. Besides a new algorithm was designed to tackle the non-convex optimization problem involved in LBMF. The experiments are performed on both simulated data and real data. The results indicate that LBMF can not only predict the students' academic performance but also classify the examination items according to the knowledge points they require. And it can be concluded that LBMF outperforms significantly the out-of-date algorithms in the applications.

Keywords Cognitive modeling,Binary matrix factorization,Item classification,Student performance prediction

1 相关介绍

学生成绩预测和考题分类是认知建模的两个重要应用,也是教育学、心理学和计算机科学的交叉课题。许多数据挖掘和机器学习领域的专家已在此方面做了大量的研究工作[1-8]。

目前相关的工作大致可分为两类,一种是根据学习者在学习系统中的反馈数据建立概率统计模型,其代表是项目反应理论模型(Item Response Theory,IRT)[4]、DINA模型(Deterministic Inputs,Noisy "And" gate)[5]和稀疏因子分析模型(Sparse Factor Analysis,SPARFA)[1]等。其中,DINA模型和SPARFA模型都是在IRT模型中增加题目难度因子和学生能力因子而发展起来的[1,5]。这类方法将成绩预测任务转化为概率模型参数因子估计问题,能够较准确地估计学生的学习状况和学习内容所含知识点的情况,从而判断学生能否正确完成未做过的试题。但是它们一般只给出概率型的输出结果,研究者常常需要根据经验和实际应用环境选取阈值,才能对最终结果进行解释[1-5]。另一种方法是矩阵分解方法,它在考题分类中表现出较好的可解释性,这类方法大致包括:非负矩阵分解方法(Non-negative Matrix Factorization,NMF)[6-7]、逻辑矩阵分解方法[8]等。虽然非负矩阵分解方法相比概率统计模型在可解释性方面具有一定优势,但是目前存在的矩阵分解方法并非完全的二值分解,用户仍然会遇到结果解释的困难[6-7]。逻辑矩阵分解方法依据布尔逻辑对矩阵进行分解,采用迭代方法对矩阵元素逐点运算,计算量较大,不适合大规模运算[8]。

二值矩阵分解(Binary Matrix Factorization,BMF)方法被认为是非负矩阵分解的拓展[9-11],它能够将大型的矩阵分解为低秩的0-1矩阵,并保留原始数据的内在信息,从而达到降维的目的,另外其解具有天然的稀疏性[9]。BMF已经被广

到稿日期:2016-09-12 返修日期:2016-12-04 本文受教育部新世纪人才计划(NCET-13-1011),华中师范大学中央高校基本科研业务费教育科学专项(CCNU16JYKX21),中央高校科研业务费项目(CCNU15A05022)资助。

张 猛(1977-),男,博士,副教授,CCF会员,主要研究方向为机器学习、教育数据挖掘,E-mail:m.zhang@mail.ccnu.edu.cn;**付丽华**(1979-),女,博士,教授,主要研究方向为机器学习与智能信息处理;**何婷婷**(1964-),博士,主要研究方向为计算语言学与数据库;**杨 青**(1965-),硕士,副教授,主要研究方向为人工智能、软件工程和数据库。

泛地应用于生物信息学等实际问题[9-11]。

然而,经典的 BMF 算法不适合直接应用于成绩预测和考题分类问题(后文将说明)。本文提出利用新颖的逻辑斯提克二值矩阵分解(Logistic BMF,LBMF)方法对问题进行建模,并设计新的算法来求解问题,新算法具有分解精度高和可解释性好的优点。在模拟数据和真实的 SAT 考试数据上进行的实验发现,新方法不但可以应用于学生成绩预测问题,而且能够进行考题的自动模式分类。实验结果表明,所提方法相比传统方法在可解释性和预测与分类精度方面都有明显的性能提升。

2 基于非负矩阵分解的方法

2.1 理论

给定学生在学习系统中的反馈数据 0-1 矩阵 $\boldsymbol{V}\in\boldsymbol{B}^{q\times s}$,其中,$\boldsymbol{B}=\{0,1\}$,$q$ 和 s 分别指考题的数目和学生的人数,$\boldsymbol{V}$ 被称为题目-学生矩阵。定义元素 $\boldsymbol{V}_{ij}$ 为第 j 个学生是否能够正确回答第 i 个题目的逻辑值:

$$\boldsymbol{V}_{ij}=\begin{cases}0, & \text{wrong}\\ 1, & \text{correct}\end{cases} \tag{1}$$

矩阵分解方法假设题目-学生矩阵 $\boldsymbol{V}$ 中包含隐含变量-知识点,若学生掌握某个题目包含的所有知识点,则能够正确完成该题目,反之将不能正确完成[6-8]。

基于上述假设,基于矩阵分解的方法将 $\boldsymbol{V}$ 分解为两个低秩的 0-1 矩阵 $\boldsymbol{W}\in\boldsymbol{B}^{q\times c}$ 和 $\boldsymbol{H}\in\boldsymbol{B}^{c\times s}$,其中 c 表示隐含变量-知识点的数目。矩阵 $\boldsymbol{W}$ 表示题目包含知识点的情况,称作题目-知识点矩阵,定义元素 $\boldsymbol{W}_{ij}$ 为第 i 个题目是否包含第 j 个知识点的逻辑值:

$$\boldsymbol{W}_{ij}=\begin{cases}0, & \text{No}\\ 1, & \text{Yes}\end{cases} \tag{2}$$

矩阵 $\boldsymbol{H}$ 表示学生掌握知识点的情况,称作学生-知识点矩阵,定义元素 $\boldsymbol{H}_{ij}$ 为第 j 个学生是否掌握第 i 个知识点:

$$\boldsymbol{H}_{ij}=\begin{cases}0, & \text{No}\\ 1, & \text{Yes}\end{cases} \tag{3}$$

为了便于读者理解,特举例说明。

例 1 问题包含 5 个考生、3 个知识点、4 个测试题目,假设实际矩阵 $\boldsymbol{W}$ 和 $\boldsymbol{H}$ 为:

$$\boldsymbol{W}=\begin{array}{c}\\ \text{题目}\end{array}\overset{\text{知识点}}{\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\\0&0&1\end{pmatrix}}\qquad \boldsymbol{H}=\begin{array}{c}\text{知}\\\text{识}\\\text{点}\end{array}\overset{\text{学生}}{\begin{pmatrix}1&1&1&0&1\\0&0&1&0&1\\0&1&0&1&1\end{pmatrix}} \tag{4}$$

由矩阵 $\boldsymbol{H}$ 知,第 4 位学生仅仅掌握第 3 个知识点,而由矩阵 $\boldsymbol{W}$ 知第 2 题和第 4 题仅考察第 3 个知识点,因此第 4 位学生能够正确回答第 2 题和第 4 题,但不能回答第 1 题和第 3 题,故 $\boldsymbol{V}$ 矩阵的第 4 列向量为$[0,1,0,1]^{\mathrm{T}}$。以此类推,可以得到学生对题目的反馈矩阵 $\boldsymbol{V}$:

$$\boldsymbol{V}=\boldsymbol{W}\odot\boldsymbol{H}=\begin{array}{c}\text{题}\\\text{目}\end{array}\overset{\text{学生}}{\begin{pmatrix}0&0&1&0&1\\0&1&0&1&1\\1&1&1&0&1\\0&1&0&1&1\end{pmatrix}} \tag{5}$$

其中,$\odot$指通过矩阵 $\boldsymbol{W}$ 和 $\boldsymbol{H}$ 合成矩阵 $\boldsymbol{V}$ 的运算。虽然较容易根据矩阵 $\boldsymbol{W}$ 和 $\boldsymbol{H}$ 合成矩阵 $\boldsymbol{V}$,但是根据给定的题目-学生矩阵 $\boldsymbol{V}$ 分解出 $\boldsymbol{W}$ 和 $\boldsymbol{H}$ 则是一项非常具有挑战性的工作。

2.2 前人工作

Michel C. Desmarais 的研究团队提出利用 NMF 方法进行考题分类[6-7]。他们证明了在每个问题仅含一个知识点的情况下,下列等式成立:

$$\boldsymbol{V}=\boldsymbol{W}\cdot\boldsymbol{H} \tag{6}$$

其中,$\cdot$ 表示普通矩阵乘法。利用如下经典 NMF 方法可以求解 $\boldsymbol{W}$ 和 $\boldsymbol{H}$[6]:

$$\begin{gathered}\min_{\boldsymbol{W},\boldsymbol{H}}\|\boldsymbol{V}-\boldsymbol{W}\boldsymbol{H}\|_F^2+\lambda\sum_{i,j}(\|\boldsymbol{H}\|_F^2+\|\boldsymbol{W}\|_F^2)\\ \text{subject to } \boldsymbol{H}_{i,j}\geqslant 0,\boldsymbol{W}_{i,j}\geqslant 0,\forall\, i,j\end{gathered} \tag{7}$$

其中,目标函数第一项表示矩阵分解精度;第二项是正则项,用来光滑目标函数;$\|\cdot\|_F$ 是矩阵的 Frobenius 范数,约束条件使得解矩阵元素均为非负数。然而 NMF 的输出并非 0,1 矩阵,需要利用聚类方法或阈值方法对所得结果进行进一步分析[6-7]。此外,当单个题目的知识点数目大于或等于 2 时会出现联合矩阵(Conjunctive matrix)[7],此时式(6)不成立。例 2 进一步说明了联合矩阵产生的问题。

例 2 假设第 i 个考生掌握 3 个知识点的情况如下:

$$\boldsymbol{h}_i=\begin{array}{c}\text{知}\\\text{识}\\\text{点}\end{array}\overset{\text{学生}}{\begin{pmatrix}0\\1\\1\end{pmatrix}}$$

第 j 个试题涉及到的知识点如下:

$$\boldsymbol{w}_j=\begin{array}{c}\text{题}\\\text{目}\end{array}\overset{\text{知识点}}{\begin{pmatrix}0&1&1\end{pmatrix}}$$

试题 j 的知识点数目多于 2,导致 $\boldsymbol{w}_i\boldsymbol{h}_j=2$,使得 $\boldsymbol{W}\cdot\boldsymbol{H}$ 不是 0,1 矩阵,从而 $\boldsymbol{V}\neq\boldsymbol{W}\cdot\boldsymbol{H}$。

为了解决联合矩阵问题,文献[7]提出:

$$\boldsymbol{V}=\overline{\boldsymbol{W}\overline{\boldsymbol{H}}} \tag{8}$$

其中,$\bar{\cdot}$ 是对矩阵逐个元素取反的运算,即将 0 映射为 1,任意非 0 数映射为 0。式(8)表示 $\boldsymbol{V}_{ij}=0$ 当且仅当第 i 个考生未掌握第 j 个试题涉及的所有知识点。文献[7]对式(8)两边取反得:

$$\overline{\boldsymbol{V}}=\boldsymbol{W}\overline{\boldsymbol{H}} \tag{9}$$

此处需要对 $\overline{\boldsymbol{V}}$ 规范化,使其成为 0-1 矩阵[7]。虽然文献[7]能够解决联合矩阵的问题,但是所得分解结果 $\boldsymbol{W}$ 和 $\boldsymbol{H}$ 不是 0,1 矩阵,需要借助聚类或选取阈值的方法对结果 0,1 化。

3 基于二值矩阵分解的方法

3.1 理论

BMF 在生物信息学与集成电路设计等领域已有广泛的应用[9-11]。BMF 作为 NMF 方法的延伸,除了能将高维矩阵降维(BMF 的分解结果一般非常稀疏),还能够给予信息紧凑的表示形式,且容易解释。

然而,目前尚无将 BMF 应用于成绩预测和考题分类方面的研究工作,本文设想利用 BMF 将式(9)中的 $\overline{\boldsymbol{V}}$ 分解为 $\boldsymbol{W}$

和 $\bar{H}$。经典的 BMF 优化问题的目标函数通过含有如下项[9-11]来控制分解结果的精确度:

$$\min_{W,H} \| \bar{V} - W\bar{H} \|_F^2 \tag{10}$$

但这种目标函数并不适用于认知建模。假设估计矩阵 $\hat{V} = W\bar{H}$,显然 $\hat{V}$ 中元素 $\hat{V}_{ij}$ 不一定为 0 或者 1。

基于上述考虑,本文提出了新颖的 LBMF 模型:

$$\begin{aligned} &\min_{W,H} \| \bar{V} - logistic(W\bar{H}) \|_F^2 \\ &\text{s. t. } (\bar{H}_{i,j} - 1)\bar{H}_{i,j} = 0, (W_{ij} - 1)W_{i,j} = 0, \forall i, j \end{aligned} \tag{11}$$

其中 Logistic 函数将 $\hat{V}_{ij}$ 压缩映射到 0 或者 1,定义如下:

$$logisitic(x) = 1 - \frac{1}{1 + \exp(100x - 10)} \tag{12}$$

3.2 算法

与大多数非负矩阵分解问题类似,式(11)并非是一个凸优化问题,但是针对所有变量矩阵 W 和 H,它是凸的,文献[12]称这类优化问题为块多凸(Block Multi-convex)问题。文献[12]提出利用推广的块坐标下降法(Generalized Block-coordinate Descent Method)来求解这类问题,基本思路为:将大的全局非凸优化问题分解为多个较小且容易求解的局部凸优化子问题,然后迭代使用拉格朗日乘子法逐块优化变量,并通过协调子问题的解来得到大的全局问题的解。文献[12]讨论了这类方法的收敛性,证明其具有收敛速度快、计算量小的优点。目前已经有一些学者将其用于求解矩阵分解问题[1,9]。基于此,本文提出如下迭代优化方法。

(1)初始化:令 $L=0$,λ,W 和 H 中的元素初始值由中心为 0、方差为 1 的正态分布随机产生。

(2)将 W 和 H 规范化为 0-1 矩阵,记为 W_0 和 H_0。

(3)在第 L 步,固定 $W = W_L$,求解下列拉格朗日优化问题。

$$\min_{H} \| \bar{V} - logisitic(W\bar{H}) \|_F^2 + \frac{\lambda}{2} \| \bar{H}^2 - \bar{H} \|_F^2$$

其中,$\bar{H}^2$ 指矩阵 $\bar{H}$ 按元素取平方数,得到的解为 H_{L+1},令 $H = H_{L+1}$,求解下列优化问题:

$$\min_{W} \| \bar{V} - logisitic(W\bar{H}) \|_F^2 + \frac{\lambda}{2} \| W^2 - W \|_F^2$$

得到的解为 W_{L+1},令 $W = W_{L+1}$。

(4)若 $(H_{ij}^2 - H_{i,j})^2 + (W_{ij}^2 - W_{i,j})^2 \leqslant \varepsilon, \forall i, j$,则令 $H = \delta(H - 0.5)$,$W = \delta(W - 0.5)$,其中:

$$\delta(x) = \begin{cases} 0, & x \leqslant 0 \\ 1, & x > 0 \end{cases}$$

否则 $\lambda = 10\lambda$,$L = L + 1$,并返回步骤(3)。

(5)输出 W 和 H。

4 实验

实验在模拟数据和真实的 SAT 考试数据上进行,将本文提出的新算法 LBMF 与经典的认知建模方法进行比较,经典的认知建模方法包括 IRT,SPARFA 和 NMF,其中 LBMF,IRT 和 SPARFA 采用 Matlab 来实现,NMF 算法采用 R 语言来实现。

4.1 模拟数据

本实验旨在说明 LBMF 估计学生知识结构和挖掘题目包含知识点的能力。仿真数据矩阵 V,W 和 H 的产生方法请参见文献[1]。仿真实验数据包含 3 个知识点,50 个题目,100 名测试学生。

本实验的任务为根据考试结果推测各考题包含的知识点及各学生掌握知识点的情况,即根据学生的作答矩阵 V,分解得到 W 和 H 的近似矩阵 $\hat{W}$ 和 $\hat{H}$,并利 $\hat{W}$ 和 $\hat{H}$ 推断 V 的近似矩阵 $\hat{V}$,最后利用矩阵估计错误率来判断算法性能。

图 1 给出了 LBMF 的实验结果,用黑白图像表示原始矩阵、近似矩阵和误差矩阵的绝对值,其中黑色块表示 0,白色块表示 1。图 1 中的第一列为各个矩阵的估计误差绝对值矩阵,第二列为各个原始矩阵,第三列为相应的估计矩阵。经过 30 次独立实验得 LBMF 方法估计正确率的平均值如下:$error\_V = 0.014$,$error\_W = 0.003$,$error\_H = 0.097$。将文献[1]的 SPARFA 应用于相同的训练和测试数据集中以便比较,得到的平均测试误差分别为 $error\_V = 0.019$,$error\_W = 0.124$,$error\_H = 0.163$。实验结果表明,新提出的 LBMF 算法能够估计学习内容包含的知识点及挖掘学生掌握知识点的情况。

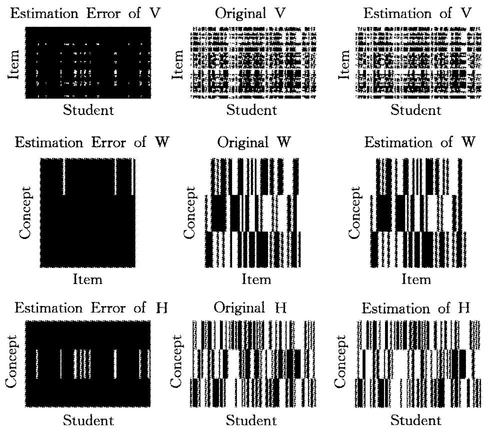
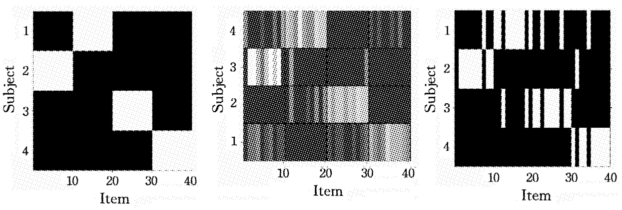


图 1　LBMF 在模拟数据中的效果图

4.2 考题自动模式分类

本实验利用 SAT 数据来考察新方法 LBMF 进行考题自动模式识别的能力,即由 V 提取出 W。SAT 数据可由文献[13]公开下载,其中包含 4 个科目的考试情况:数学、生物、世界史和法语。每个科目包含 10 个题目,本实验随机选取 100 名学生的考试结果,得到学生-考题矩阵。图 2(a)-图 2(c)分别给出了原始的 W 矩阵及 NMF 与 LBMF 的估计结果。与上个实验类似,原始的 W 矩阵和 LBMF 的估计结果矩阵用黑白两色表示,黑色块为 0,白色块为 1。由于 NMF 方法的分解结果为非 0-1 矩阵,因此用灰度图表示,较亮的色块表示趋近于 1 的值,反之趋近于 0。因此当使用 NMF 方法时,用户需要自行确定阈值来判定色块的值表示 1 或者 0。若将阈值设为0.5,则 NMF 方法的准确率为 62.5%,而新的 LBMF 方法的准确率为 77.5%。本实验说明与传统的 NMF 相比,LBMF 不但使用方便,不需要自行设置阈值,而且在挖掘题目包含知识点的问题中能够获得较准确的逼近效果。



(a)真实 W 矩阵　(b)NMF 的估计结果　(c)LBMF 的估计结果

图 2　考题自动模式分类效果图

4.3　预测未观测到的学生反馈

本实验利用模拟数据将 LBMF 与经典的 IRT 模型[4]、SPARFA 模型[1]相比较,以考察它们预测未观测到的学生反馈的能力。观测不到学生的反馈数据的原因有多种,如智能学习系统漏洞产生数据缺失、学生未做某些考题等[14]。实验采用文献[1]中的方法随机产生 5 组数据,数据的具体描述如表 1 所列。使用两个指标来检验预测效果:估计精度(Correctness,COR)和接受者操作曲线下面积(Area Under the receiver operation Characteristic curve,AUC),其中定义估计精度 COR 为正确估计样本和总体待预测样本数目之比。在每个数据集中随机抽取 20%的学生做 20%的试题的数据作为待预测数据。

表 1　随机产生数据集描述

数据集	1	2	3	4	5
学生人数	90	40	50	1700	1600
题目数	200	150	100	60	60

表 2 列出了各种方法在 30 次随机采样后的平均效果。在 5 个数据集合的 10 个评价指标中,LBMF 6 次排名第一,4 次排名第二,这说明新提出的 LBMF 与目前的流行方法相比在预测精度方面具有一定的竞争力。

表 2　对数据集进行 30 次随机采样时各方法的平均实验效果

		IRT	SPARFA	LBMF
1	COR	0.753	0.791	**0.812**
	AUC	0.796	0.805	**0.838**
2	COR	0.693	0.786	**0.791**
	AUC	0.735	**0.780**	0.765
3	COR	0.783	**0.843**	0.830
	AUC	0.864	**0.891**	0.871
4	COR	0.737	0.649	**0.862**
	AUC	**0.746**	0.693	0.721
5	COR	0.709	0.723	**0.744**
	AUC	0.682	0.710	**0.714**

注:黑体表示最优结果

结束语　本文通过对二值矩阵分解方法在认知结构建模中的应用研究,提出了新的认知结构模型,即 Logistic BMF 模型。该模型借鉴非负矩阵分解和二值矩阵分解算法的主体思路,保证了分解之后的 $\boldsymbol{W}$ 矩阵和 $\boldsymbol{H}$ 矩阵都为 0,1 矩阵。该模型不仅具有较高的认知诊断准确率,而且不需要具备专家知识,避免了因人为主观因素而带来的偏差。此外,该模型还具有简单且易于理解的特点,其分解方法的过程性较强,分解结果具有良好的可解释性。未来研究将在以下 3 个方面展开:

1)解的稳定性问题。针对 LBMF 算法对于某些实际数据会产生不能快速收敛到最优解的问题,文中采用多次随机重启实验的方法,保留较优结果作为最终解。虽然类似的算法已经在生物数据和高光谱图像上取得成功[9,12],但是如何从理论角度探索 LBMF 在学生反馈数据上的收敛性仍然值得进一步研究。

2)应用领域拓展问题。将研究其他因素(例如样本量、测验长度、属性个数及分布、题目难度因子、学生个人能力因子等)对 LBMF 模型准确率的影响。

3)从表 2 可以看出,当题目数和学生人数相差不大时,LBMF 的效果不太理想。该问题是普遍存在还是仅仅存在于本实验所采用的模拟数据中还需要大量实际数据来进行进一步验证。

参 考 文 献

[1] LAN A,WATERS A,STUDER C,et al. Sparse Factor Analysis for Learning and Content Analytics[J]. Journal of Machine Learning Research,2014,15(6):1959-2008.

[2] WU R,LIU Q,LIU Y,et al. Cognitive Modelling for Predicting Examinee Performance[C]// Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI 2015). 2015:1017-1024.

[3] LAN A,GOLDSTEIN T,BARANIUK R. Dearlbreaker:A Nonlinear Latent Variable Model for Educational Data[C]// Proceedings of the 33rd International Conference on Machine Learning(ICML 2016). 2016:266-275.

[4] RASCH G. Probabilistic Models for Some Intelligence and Attainment Tests [M]. MESA Press,1993.

[5] TORRE J. The generalized DINA model framework[J]. Psychometrika,2011,76(2):179-199.

[6] DESMARAIS M. Mapping Question Items to Skill with Non-negative Matrix Factorization [J]. ACM KDD-Explorations,2011,13(2):30-36.

[7] DESMARAIS M,MEHESHTI B,NACEUR R. Item to Skill Mapping:Deriving a Conjunctive Q-matrix from Data[C]// Proceedings of the 11th International Conference Intelligent Tutoring Systems. 2012:454-463.

[8] SUN Y,YE S,INOUE S,et al. Alternating Recursive Method of Q-matrix Learning [C]//Proceedings of the 7th Educational Data Mining. 2014:14-20.

[9] ZHANG Z,LI T,DING C. Binary Matrix Factorization for Analyzing Gene Expression Data[J]. Data Mining and Knowledge Discovery,2010,20(1):28-52.

[10] LI T. A General Model for Clustering Binary Data [C]// Proceedings of the 11th International Conference on Knowledge Discovery and Data Mining. 2005:188-197.

[11] SHEN B,JI S,YE J. Mining Discrete Patterns via Binary Matrix Factorization [C]//Proceedings of 15th International Conference on Knowledge Discovery and Data Mining. 2009:757-966.

[12] XU Y,YIN W. A Block Coordinate Descent Method for Regularized Multi-convex Optimization with Applications to Nonnegative Tensor Factorization and Completion[J]. SIAM Journal of Imaging Sciences,2013,6(3):1758-1789.

[13] WINTERS T. Datasets for Topic Clustering [DB/OL]. http://alumni.cs.ucr.edu/~titus.

[14] HUANG J M. Application of Bayesian Network to predicting Students' Achievement[J]. Computer Science,2012,39(11A):280-282. (in Chinese)
黄建明. 贝叶斯网络在学生成绩预测中的应用[J]. 计算机科学,2012,39(11A):280-282.