

文章编号: 1000-5862(2016)03-0295-04

属性蕴含 Q 矩阵理论的认知诊断模型

武永华¹ 杨淑群^{2*}

(1. 福建江夏学院, 福建 福州 350108; 2. 福建师范大学软件学院, 福建 福州 350007)

摘要: 引入属性蕴含的 Q 矩阵理论扩展了传统 Q 矩阵理论. 属性的先决关系是属性蕴含的特殊形式, 寻找属性蕴含 Q 矩阵理论中合适的认知诊断模型具有重要的理论与应用意义. 就满足属性蕴含的 Q 矩阵, 视其为形式概念分析中的形式背景, 形势背景诱导的概念格亦为认知诊断模型. 研究成果可进一步丰富 Q 矩阵理论与方法.

关键词: 属性蕴含; 概念格; 认知诊断; Q 矩阵

中图分类号: B 841 **文献标志码:** A **DOI:** 10.16357/j.cnki.issn1000-5862.2016.03.16

0 引言

认知诊断(Cognitive Diagnosis, CD)是测量个体特定的知识结构(knowledge structure)和加工技能(processing skills)^[1]. 认知诊断模型在认知诊断中的重要性是不言而喻的, Q 矩阵(项目与属性的关联矩阵)对于大多数认知诊断模型也是不可缺少的, 无论如何评论 Q 矩阵以及 Q 矩阵理论在认知诊断中作用都不过份^[3]. Q 矩阵的研究吸引了众多学者的兴趣, 亦有大量的研究成果. Q 矩阵从提出^[3-4], 到 J. P. Leighton 等^[5]提出先确定属性层级关系, 对被试进行分类诊断的属性层级方法(AHM), 在 AHM 的基础上, 丁树良等^[6]对 Q 矩阵理论作了进一步补充和修正, 修正后的 Q 矩阵理论更合理、科学. 杨淑群等^[2]将 AHM 拓广至更一般的属性蕴含, 并给出基于属性蕴含的 Q 矩阵理论^[7], 但引入属性蕴含的认知诊断模型的研究尚未有报道.

1 国内外研究现状及发展动态分析

迄今为止, 心理测量学者们已开发了近百种认知诊断模型. Q 矩阵对于大多数认知诊断模型不可或缺, 其中规则空间方法(RSM)、属性层级方法(AHM)与 DINA 等模型都要求必须有 Q 矩阵. RSM

通过测验项目得到描述项目与属性关系的 Q 矩阵, 用代数方法把不可观察的属性掌握模式转化为可观察的理想反应模式, 然后进行判别分类, 达到认知诊断的目的. 与 RSM 不同的是, AHM 先确定属性之间的先决关系, 给出满足先决关系的理想反应模式, 再判别分类. 杨淑群等^[8]基于给定的先决关系, 生成满足先决关系的 Q 矩阵, 视此 Q 矩阵为形式背景^[9]. 形式背景诱导的概念格可作为认知诊断模型, 具有良好的诊断能力^[10-12].

2 基本概念

2.1 属性蕴含关系

蕴含关系的定义为: A, B 为属性集, A 蕴含 B 当且仅当每个具有 A 中属性的对象就一定具有 B 中的属性. 蕴含的例子非常多, 数学中的所有定理都表明了蕴含关系. 如大家熟悉“三角形两边的和大于第三边”, A 为“三角形”, B 为“两边的和大于第三边”, A 蕴含 B . 再如, 集合论中的关系, 一般从自反、对称、反对称及传递来考察关系的性质. 众所周知, 具有对称与反对称的关系一定具有传递性, 即对称与反对称蕴含传递. 如属性 A 为同分母分数减法($5/6 - 2/6$), 属性 A_1 为整数减法($5 - 2$), 属性 A_2 为整数的除法. 要掌握属性 A , 首先要掌握属性 A_1 与属性 A_2 ,

收稿日期: 2016-01-17

基金项目: 国家自然科学基金青年科学基金(51503036), 教育部人文社会科学研究一般项目青年基金(10YJCXLX049)和福建省自然科学基金(2013J01119)资助项目.

作者简介: 武永华(1975-), 男, 河北石家庄人, 讲师, 主要从事嵌入式系统开发和大数据挖掘的研究.

通信作者: 杨淑群(1970-), 女, 江西临川人, 教授, 博士, 主要从事人工智能和计算机辅助教学的研究.

而 A_1 与 A_2 之间没有先决关系, 因此属性集 $\{A_1, A_2\}$ 蕴含属性 A .

K. K. Tatsuoka 对先决关系给出过数学化定义, 即将 Q 矩阵行向量进行两两比较, 采用集合的包含关系以产生属性之间的先决关系. 属性的先决关系为蕴含关系的特殊情况^[6].

2.2 形式概念分析

形式概念分析中, 数据是以二元关系的形式背景来表示. 本文沿用标准的 FCA 记法^[13]. 下面给出一个形式背景, 如表 1 所示.

表 1 一个形式背景

编号	a	b	c	d
1	1	0	0	1
2	1	1	0	0
3	1	0	1	1
4	1	0	1	1

给定一个形式背景 (G, M, I) , 定义 2 个单目运算:
 $\forall A \in P(G) \quad A^* = \{m \in M \mid gIm, \forall g \in A\} \in P(M)$,
 $\forall B \in P(M) \quad B^* = \{g \in G \mid gIm, \forall m \in B\} \in P(G)$.
 可检验公式 2 个运算的复合分别是偏序集 $(P(G), \subseteq)$ 和 $(P(M), \subseteq)$ 的闭包算子.

设 (G, M, I) 为一个形式背景, 有序对 (A, B) , $A \in P(G)$, $B \in P(M)$, P 为幂集. 若 $(A, B) \in P(G) \times P(M)$ 满足 $A^* = B$ 且 $B^* = A$, 则称 (A, B) 为形式概念. 形式背景的概念可定义序关系 $(A, B) \leq (C, D) \Leftrightarrow A \subseteq C$ (或 $B \subseteq D$), 可诱导出 $(P(G) \times P(M), \leq)$ 为一个格, 称为概念格.

3 基于形式概念分析的认知诊断模型

3.1 属性先决关系下的认知诊断模型

图 1 为满足属性先决关系的一个属性层级结构, 属性 A_1 是其他 3 个属性的先决属性, 属性 A_3 是 A_4 的先决属性, A_2 与 A_4 不是任何属性的先决属性.

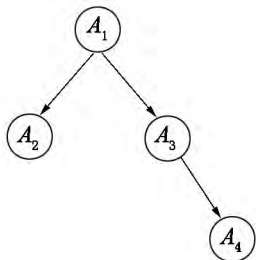


图 1 一个满足属性先决关系的属性层次结构

根据文献[8]中的算法, 满足图 1 的 Q 矩阵为

$$Q_1 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Q_1 转置后为形式背景, 从文献[12]可得此形式背景诱导的概念格如图 2 所示.

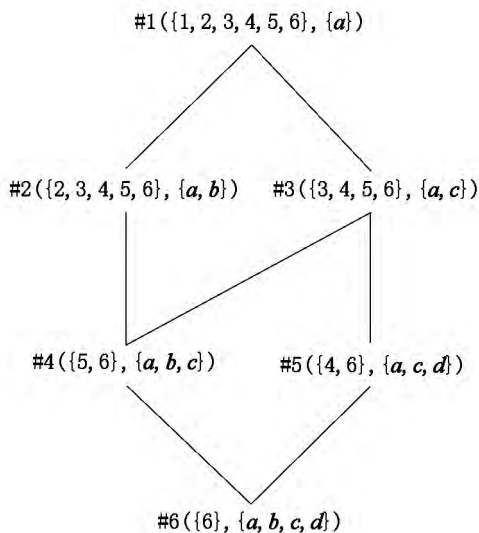


图 2 Q_1 诱导的概念格

每个概念有二元信息, 一是项目集, 二是属性集. 项目集的共同属性构成了属性集, 具有属性集的项目就是项目集. 概念格不仅表达了概念之间的序关系, 还表达了概念之间的包含关系. 比如, 概念 #2 包含了项目集 $\{2, 5, 6\}$, 属性集 $\{a, b\}$, 表示项目 2、项目 5 与项目 6 都具有属性 a 与 b , 被试测试了这 3 个项目意味着诊断其是否掌握了属性 a 与 b . 概念格自上而下, 概念中的属性集的属性数目逐层递增 1, 项目难度依次增强, 是理想的认知诊断模型. Monte Carlo 模拟实验表明, 满足先决关系的 Q 矩阵诱导的概念格作为认知诊断模型, 被试的知识状态较好的被诊断, 能力估计精度也较高^[11].

3.2 属性蕴含的认知诊断模型

属性先决关系是属性蕴含关系的特例. 显然, 当属性蕴含为先决关系时, 上述概念格是良好的认知诊断模型. 但是, 属性关系满足一般的属性蕴含关系时, 概念格是否依然是理想的认知诊断模型是本文讨论的重点.

文献[7]中介绍了满足属性蕴含的 Q 矩阵, 给定属性蕴含集 $\{\{a, b\} \rightarrow c, c \rightarrow d\}$, 其对应的充分 Q 矩阵为

$$Q_2 = \begin{pmatrix} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ a & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ b & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ c & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ d & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix},$$

Q_2 转置后可得形式背景如表 2 所示.

表 2 Q_2 转置所得形式背景

编号	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
1	0	0	0	1
2	0	0	1	1
3	0	1	0	0
4	0	1	0	1
5	0	1	1	1
6	1	0	0	0
7	1	0	0	1
8	1	0	1	1

按照概念格的生成算法,其诱导的概念格如图 3 所示.

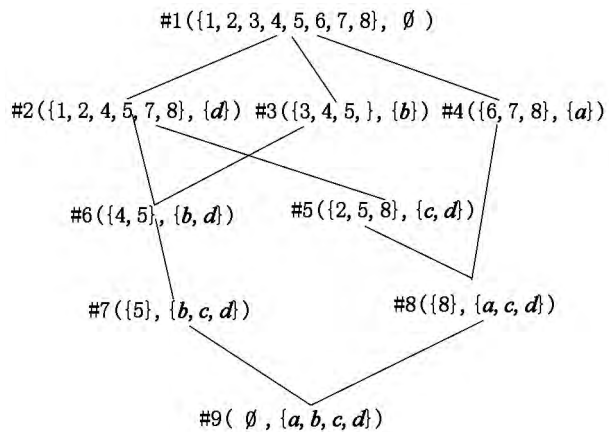


图 3 表 2 诱导的概念格

从图 3 来看,与基于属性先决关系生成的概念格相比,概念包含的信息相同,概念格自上而下,概念的属性数目递增,项目难度依次增强,也可作为认知诊断模型.所不同的是,格中上下层之间属性集的数目不一定相差 1,比如概念 #4 的属性集有一个属性 *a*,概念 #8 有 3 个属性 *a*、*c* 与 *d*.诊断中,被试掌握了属性 *a*,如果要诊断属性 *c*,只能同时诊断属性 *c* 与 *d*,原因是受蕴含关系的约束.

既然属性蕴含诱导的概念格可作为认知诊断模型,其诊断方法的思想、算法与步骤可参考文献 [11].基于属性蕴含的题库可囊括属性先决关系的题库,只要效率与精度不低于属性先决关系的诊断就是进步.以后的论文会专门阐述这个问题.

4 结论

属性先决关系是属性蕴含的特例,满足属性蕴含关系的充分 Q 矩阵视为形式背景,诱导的概念格可作为认知诊断模型.较之基于属性先决关系的概念格研究理论,本文所不同的在于充分 Q 矩阵与其诱导的概念格特性,对认知诊断的效率与精度无影响,因此基于属性蕴含的概念格依然是理想的认知诊断模型.

诊断中涉及的属性多,如果充分 Q 矩阵与概念格的生成算法效率不能容忍,在自适应的考试中难以应用.因此进一步的工作是,研究属性蕴含关系的性质,得出高效的充分 Q 矩阵与其诱导的概念格生成算法.

5 参考文献

- [1] 漆书青,戴海崎,丁树良.现代教育与心理测量学原理[M].北京:高等教育出版社,2002.
- [2] 杨淑群.认知评估中的属性关系[J].江西师范大学学报:自然科学版,2015,39(2):132-137.
- [3] Tatsuo K K. Boolean algebra applied to determination of universal set of knowledge states [R]. Princeton: Education Testing Service, 1991.
- [4] Tatsuo K K. Architecture of knowledge structure and cognitive diagnosis: a statistical pattern recognition and classification approach [C]. Hillsdale: Erlbaum, 1995.
- [5] Leighton J P, Cderl M J, Hunka S M. The attribute hierarchy method for cognitive assessment: a variation on Tatsuo-ka's rule-space approach [J]. Journal of Educational Measurement, 2004, 41(3): 205-237.
- [6] 丁树良,祝玉芳,林海菁,等.矩阵理论的修正[J].心理学报,2009,41(2):101-112.
- [7] 杨淑群.基于属性蕴含的 Q 矩阵理论[J].江西师范大学学报:自然科学版,2015,39(6):619-622.
- [8] 杨淑群,蔡声镇,丁树良,等.求解简化 Q 矩阵的扩张算法[J].兰州大学学报:自然科学版,2008,44(3):87-96.
- [9] Wille R. Restructuring lattice theory: an approach based on hierarchies of concepts [M]. Dordrecht-Boston: Reidel, 1982: 445-470.
- [10] Yang Shuqun, Ding Shuliang, Cai Shenzheng, et al. An algorithm of constructing concept lattices for CAT with cognitive diagnosis [J]. The Journal of Knowledge Based Systems, 2008, 21(8): 852-855.
- [11] Yang Shuqun, Ding Shuliang, Ding Qiulin. The theory about CD-CAT based on FCA and its application [J]. In-

- ternational Journal of Distance Education Technologies , 2009 4: 61-78.
- [12] Yang Shuqun ,Ding Shuliang. A FCA-based cognitive diagnosis model for CAT [EB/OL]. [2015-11-19]. 10. 4018/978-1-60960-539-1. ch010.
- [13] Ganter B ,Wille R. Formal concept analysis: mathematical foundations [M]. Berlin Heidelberg: Springer-Verlag , 1999.

The Cognitive Diagnosis Model of Q Matrix Theory with Attribute Implication

WU Yonghua¹ ,YANG Shuqun^{2*}

(1. Fujian Jiangxia University ,Fuzhou Fujian 350108 ,China 2. Faculty of Software ,Fujian Normal University ,Fuzhou Fujian 350007 ,China)

Abstract: The Q matrix theory introduced with attribute implication has expanded the traditional Q matrix theory. The prerequisite relationship of attributes is a distinctive form of attribute implication and an appropriate cognitive diagnosis model for Q matrix theory can realize profound significance both theoretically and pragmatically. The Q matrices that meet the conditions of the formal concept analysis are considered as the formal context of formal cognitive analysis and the concept lattices induced by formal concept are regarded as cognitive analysis models. The results of this study can broaden the prospects of Q matrix theory and its methods.

Key words: attribute implication; concept lattices; cognitive diagnosis; Q matrix

(责任编辑: 冉小晓)

(上接第 294 页)

- [19] Madison M J ,Bradshaw L P. The effects of Q -matrix design on classification accuracy in the log-linear cognitive diagnostic diagnosis model [J]. Educational and Psychological Measurement 2015 ,75(3) : 491-511.
- [20] 罗欢 ,丁树良 ,汪文义 ,等. 属性不等权重的多级评分属性层级方法 [J]. 心理学报 2010 42(4) : 528-538.
- [21] 李瑜. 多选题认知诊断测验编制及多策略的多选题认知诊断模型的开发 [D]. 南昌: 江西师范大学 2014.
- [22] 李瑜 ,丁树良 ,唐小娟. 多项选择题认知诊断潜能的最大化 [J]. 心理科学进展 2014 22(5) : 866-880.
- [23] 唐小娟. 粗糙集理论在认知诊断中的应用 [D]. 南昌: 江西师范大学 2013.

The Irreplaceability of a Reachability Matrix

DING Shuliang ,WANG Wenyi ,LUO Fen ,XIONG Jianhua

(College of Computer Information Engineering ,Jiangxi Normal University ,Nanchang Jiangxi 330022 ,China)

Abstract: A reachability matrix R has two important properties: one is that any column in Q matrix can be expressed by a linear combination of the columns of R with the combination coefficients being 1 or 0 ,the other is that under the conditions of 0-1 scoring rubric and the noncompensatory among the attributes ,if R (or a permutation of its columns) is a sub-matrix of the test Q matrix ,then the ideal response patterns corresponding to any two different knowledge states are different. It is proved that these properties of R are irreplaceability ,i. e. ,any other Q matrix does not have one of these properties. A counterexample is provided to explain that the concept of necessary Q matrix instead of the concept of sufficient Q matrix can promote the construct validity.

Key words: reachability matrix; irreplaceability; sufficient Q matrix; necessary Q matrix

(责任编辑: 冉小晓)