

文章编号: 1000-5862(2018)02-0130-04

Q 矩阵标定的一种简便方法

丁树良, 罗 芬, 汪文义, 熊建华

(江西师范大学计算机信息工程学院, 江西 南昌 330022)

摘要: 基于认知诊断的 Q 矩阵理论, 由潜在 Q 矩阵向量的累赘表达式的分解, 给出一种简便的 Q 矩阵标定方法, 并且给出这种方法的证明.

关键词: 认知诊断; Q 矩阵标定; 累赘表达式; Q 矩阵理论

中图分类号: B 841.7 文献标志码: A DOI: 10.16357/j.cnki.issn1000-5862.2018.02.02

0 引言

Q 矩阵在认知诊断中的重要性是不言而喻的. 要得到准确的认知诊断分类结果, 影响因素较多, 比如准确确定认知模型, 即属性数目和属性层级结构, 选择和数据拟合的并且判准率高的认知诊断模型, 设计优良的测验蓝图, 精心打造试题和试卷等, 还有一点不容忽视, 即 Q 矩阵必须标定准确. Q 矩阵标定实质上是标定题目属性向量, 题目属性向量标定不准确, 比如对于 2 值 (0-1) Q 矩阵而言, 可以分成 3 种情况: 1) “过度标定”: 题目没有涉及的属性但是却认为涉及, 即本来为 0 的元素标注为 1, 这样可能认为刚好可以答对这道题目的被试(其知识状态和题目属性向量刚好相等)是猜测正确, 因此这种题目猜测参数高; 2) “不足标定”: 本来涉及的属性却认为没有涉及, 即 Q 矩阵的元素本来应该为 1 却标定为 0, 这样从题目属性向量来看, 应该答对的题目却答不对, 显示这个题目的失误率很高; 3) 同一个题目的属性向量既有“过度”标定、又有“不足”标定, 很可能这一类题目猜测和失误参数均比较高. 题目猜测或者失误参数高, 则题目质量不太好; 题目质量不太好, 就难以准确评定被试的知识状态.

因此 Q 矩阵的准确标定, 是认知诊断评估是否成功的一个关键因素. 然而准确标定 Q 矩阵, 却不是轻而易举的工作, 它是一个比较复杂的过程.

若按照 J. P. Leighton 等^[1] 推荐的方案, 先认知诊断测验设计, 再命题, 这个过程有许多环节: 1) 必须确定认知诊断的范围, 比如准备诊断某一门课程的一个章节的某一个具体的内容, 理清这个诊断内容的所有属性及其层级关系, 包括考虑属性的粒度的大小; 2) 计算这个属性及其层级关系对应的邻接矩阵、可达阵, 基于可达阵, 使用扩张算法计算潜在 Q 矩阵^[2-3], 潜在 Q 矩阵也就是简化 Q 矩阵^[4-5]; 3) 根据认知诊断测验设计的原理, 选择潜在 Q 矩阵的一个子矩阵作为测验 Q 矩阵, 再根据测验 Q 矩阵命题, 命题以后要对每一个题目标注题目属性向量. 问题: 编制好认知诊断测验的题目后, 题目的属性向量是不是标注正确? 这个环节往往争论最多.

若采用 K. K. Tatsuo^[4-5] 的方案, 可以使用以往为别的目的(而不是以认知诊断为目的)进行测试的数据进行认知诊断分析(这个方案称为“翻新”), 则 Q 矩阵标定的工作或许更加困难. 因为这时候属性及其层级关系的认定更加困难, 题目属性向量不像 J. P. Leighton 等^[1] 的方案那样存在初始的“值”, 从而如何标定的争论更多, 并且由于这时试卷缺乏认知诊断测验设计, 对其提供数据诊断分析结果的准确性如何, 也难以预料. 因此, 根据实测数据进行数据挖掘以获得题目属性向量的做法可能更加困难.

对于 K. K. Tatsuo^[6] 的测验数据和测验 Q 矩阵, 往后有多位学者对其进行研究, 发现 Q 矩阵的标注存在一些问题^[7-8]. 这个例子说明 Q 矩阵正确

收稿日期: 2017-08-06

基金项目: 国家自然科学基金(31500909, 31360237, 31300876), 教育部人文社会科学研究青年基金(BYJC880060)和江西省教育科学 2013 年一般课题(BYB032)资助项目.

作者简介: 丁树良(1949-), 男, 江西樟树人, 教授, 主要从事计算机辅助教学及教育和心理测量方面的研究. E-mail: ding06026@163.com

标定是较困难的.

因此 Q 矩阵标定的研究引起很多学者的关注. de la Torre 提出 δ 方法^[7], 涂冬波等^[9] 提出 γ 方法, 汪文义等^[10] 提出新颖的交差方法, 喻晓锋等^[11] 在 Liu Jingchen 等^[12] 的基础上提出联合估计方法, 该方法计算量大, 完全是数据驱动方法, 陈平等^[13] 也对 Q 矩阵标定进行了探索. 这些探索很多和认知诊断模型有关. 即在给定认知诊断模型的前提下, 考察 Q 矩阵标定的方法. 以上方法都是借助被试知识状态的估计值来校正项目的属性向量, 显然被试知识状态估计精度会影响 Q 矩阵标定方法的表现^[10]. 特别在测验长度较短时, 知识状态估计精度难以保证, 这个问题尤为突出.

Q 矩阵中非零列的累赘表达式和简洁表达式^[14] 的概念能否用于 Q 矩阵进行标定, 并且避免上述问题, 这是本文欲探讨的问题; 特别地, 若能够使用 Q 矩阵理论对二值 Q 矩阵进行标定, 能否将其解决多值 Q 矩阵^[15-17] 的标定问题? 而据笔者所知, 多值 Q 矩阵的标注的研究目前尚属阙如.

本文首先介绍一种基于 Q 矩阵理论基础上的 2 值 Q 矩阵标定的新方法及其理论, 再指出这种方法也可用于多值 Q 矩阵的标定, 然后是讨论.

1 关于 Q 矩阵标定新方法的理论

1.1 Q 矩阵中非零列的不同表达方式

丁树良等^[14] 从非零知识状态 x 均可表示为可达阵的列的布尔并, 但是表示方法不唯一入手, 研究了 x 的 2 种表示, 即累赘表达式和简洁表达式. 所谓累赘表达式是先定义一个集合: 令 $S_x = \{r | (r \text{ 是 } R \text{ 的列}, R \text{ 是可达阵}) \text{ 并且 } (r \leq x)\}$. 这里 $w \leq y$ 表示 $y - w$ 的每一个分量均为非负, 称 $w \leq y$; 若 $w \leq y$, 并且 y 中至少有一个元素大于 w 中相应的元素, 则称 $w < y$.

定义 1 S_x 中所有向量的布尔并称为 x 的“累赘”(redundant) 表达式, 并且 S_x 中每一个列均称为 x 的构成向量. 而 S_x 就是 x 构成向量的集合.

假设 w 是 S_x 中的任意一个分量, 将 w 和 x 的其他构成向量分别进行比较, 若 w 小于某一个构成向量, 则从 S_x 中删除 w , 并且将所得到的集合记为 S'_x . 将这个过程称之为比较删除步骤. 将 S_x 中每一个向量都作上述的比较删除步骤, 最后所得到的集合仍然记为 S'_x . S'_x 中的元素的布尔并称为 x 的简洁表达式.

Q 矩阵中非零列不仅可以表示非零知识状态, 还可以表示题目属性向量, 所以知识状态的 2 种不同表示也可以说是题目属性向量的 2 种不同表示, 即累赘表达式和简洁表达式. 显然, 有如下的引理.

引理 1 Q 矩阵中任何一列的累赘表达式是唯一确定的, 从而简洁表达式也是唯一确定的.

事实上, 假设 x 是潜在 Q 矩阵的列, 那么 x 的构成向量集合是唯一的, 因此对应的累赘表达式唯一. 而根据简洁表达式的定义, 简洁表达式也唯一. 由引理 1 可知, 若获得潜在 Q 矩阵列的累赘表达式构成向量的集合, 这一列就应该视为已知. 引理 1 表明潜在 Q 矩阵的任何 1 列的累赘表达式是唯一的, 所以如果约定使用累赘表达式进行讨论, 那么可达阵的列就相当于“基底(base)”, 由扩张算法^[2-3] 得知 Q 矩阵中任何一个非零向量都可以由这个“基底”唯一表示. 当然如果约定仅仅使用简洁表达式, 也可以得到相仿结论.

1.2 定理及其证明

假设 x 为潜在 Q 矩阵的 1 列, 但是 x 未知, x 作为一个题目的属性向量, 要对 x 的元素进行标注. 设 b 是任意一个满足下述条件的知识状态: b 是潜在 Q 矩阵的 1 列. 在理想反应情况下, $x \leq b$ 则具有知识状态 b 的被试在题目 x 上正确反应, 即 $x \leq b$, 记 $S_b(x) = \{x | r \in R \wedge (r \leq b) \wedge (x \leq b)\}$, $S_b(x)$ 表示对 x 的理想反应正确的知识状态的累赘表达式的构成向量的集合, 则有

定理 1 $\bigcap_{b \in Q_p} S_b(x) = \{r | (r \in R) \wedge (r \leq x)\}$, 并且 $x = \bigvee_{r \in \{r | r \in R \wedge (r \leq x)\}} r$.

证 若 $x \leq b$, 则 $S_x(x) \subseteq S_b(x)$, 所以 $S_x(x) \subseteq \bigcap_{b \in Q_p} S_b(x)$. 注意 $x \in Q_p$, 并且由偏序关系的自反性, 有 $x \leq x$, 从而 $\bigcap_{b \in Q_p} S_b(x) \subseteq S_x(x)$. 由于 $(x \leq x)$ 总是为真, 所以 $S_x(x) = \{r | (r \in R) \wedge (r \leq x) \wedge (x \leq x)\}$. 根据 x 的累赘表达式的定义以及引理 1, 即累赘表达式唯一, 故 $x = \bigvee_{r \in S_x(x)} r$.

定理 1 的条件是严格的, 因为给定测验 Q 矩阵, 它在测验 Q 矩阵上所有反应均为理想反应模式; 而且要求潜在 Q 矩阵对应的所有知识状态都要对测验 Q 矩阵反应, 缺一不可; 然后挑选那些在 x 上正确反应的被试的知识状态的累赘表达式构成向量的集合作交集. 这个交集对应 x 的累赘表达式的构成向量. 悉知, 一般观察反应模式不一定和理想反应模式重合.

2 新的题目属性标定方法

假设 x 是潜在 Q 矩阵中的 1 列, 其元素未知. 潜在 Q 矩阵包含 n 列, 可达阵 R 是 K 阶方阵 $R = (r_1 r_2 \cdots r_K)$. 为了标注潜在 Q 矩阵中未知列, 不妨采取一列一列标注的方式进行. 令 $Q_i = (R | x)$ 为测验 Q 矩阵 x 为未知向量. 在理想反应条件下, 进行标注的步骤如下:

(i) $j = 1$;

(ii) 对于潜在 Q 矩阵的列(即非零知识状态) q_j , 若 q_j 在 x 上的理想反应为正确, 则收集 q_j 在 R 的所有列上的反应; 构造集合 $S_{q_j}(x) = \{r | (r \in R) \wedge (r \leq q_j) \wedge (x \leq q_j)\}$;

(iii) $j = j + 1$;

(iv) 若 $j < n + 1$, 转 (i), 否则令 $S_x(x) = \{r | (r \in R) \wedge (r \leq x) \wedge (x \leq x)\}$, 则 $x = \bigvee_{r \in S_x(x)} r$.

例 1 假设独立属性层级结构, 属性数目 $K = 3$, 又设 $x = (1 \ 0 \ 1)^T$ 未知; 此时, 可达阵是 3 阶单位阵, 潜在 Q 矩阵包含 7 列, $Q_p = (q_1 q_2 \cdots q_7)$, 而测验 Q 矩阵是除单位阵之外, 增加未知向量 x , x 取自潜在 Q 矩阵 $Q_i = (R | x)$,

(i) Q_p 中所有列对 x 理想反应: 除 $q_5 = (1 \ 0 \ 1)^T$, $q_7 = (111)^T$ 对 x 正确反应之外, 其他均错误;

(ii) q_5 对应的累赘表达式分量的集合 $S_1 = \{r_1 r_3\}$, q_7 对应的累赘表达式分量的集合为 $S_2 = \{r_1 r_2 r_3\}$, 则 x 对应的累赘表达式的分量的集合是 S_1 与 S_2 的交集, 它等于 S_1 , 由于累赘表达式的唯一性, 所以 $x = (1 \ 0 \ 1)^T$.

3 多值 Q 矩阵的标注

多值 Q 矩阵中属性的水平可以是非负整数^[15-16]. 给出 K 个属性, 假设第 j 个属性的水平数为 w_j , 丁树良等^[17] 讨论在 K 阶 0-1 可达阵 R 的基础上, 结合属性的水平数, 使用“膨胀算法”, 获得一个 w 阶 ($w = w_1 + \cdots + w_K$) 的 0-1 方阵 S , 然后在 S 的行的方向上使用“压缩算法”, 将 S 压缩成为 K 行 w 列的矩阵, 称之为“拟可达阵”, 记之为 R_p . 将扩张算法中计算 2 个元素布尔并的方法, 修改为求 2 个元素的最大值, 即对扩张算法进行修改, 在 R_p 的基础上, 使用修改过的扩张算法, 就可以获得所有的多值知识状态, 即获得多值潜在 Q 矩阵.

将 1.2 节中的可达阵 R 改为拟可达阵 R_p , 假设 x, y 是多值潜在 Q 矩阵的列, x 是测验 Q 矩阵的未知列, 若 $x \leq y$, 则 y 在 x 上的理想反应正确.

例 2 3 个属性 A_1, A_2, A_3 , A_1 是根节点, A_2, A_3 是叶节点(如图 1 所示), 它们的最高水平数分别为 3、2、2,

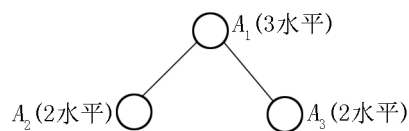


图 1 3 属性多水平的例子

$$R_2 = \begin{pmatrix} 111 \\ 010 \\ 001 \end{pmatrix} \rightarrow R_p = \begin{pmatrix} 1231111 \\ 0001200 \\ 0000012 \end{pmatrix} = (r_1 r_2 r_3 r_4 r_5 r_6 r_7),$$

R_p 是 3×7 矩阵.

$$Q_p^{(p)} = \begin{pmatrix} 1231111 & j=2 & j=3 & j=4 & j=5 & j=6 & j=7 \\ 2222 & 2222 & 3333 & 112233 & 112223331 & & \\ 0001200 & 1200 & 1200 & 111111 & 222222222 & & \\ 0000012 & 0012 & 0012 & 121212 & 120120122 & & \end{pmatrix},$$

Q_p 是 3×30 矩阵.

取 $x = (2 \ 1 \ 2)^T$, $Q_i = (R_x | x)$, 在 x 上理想反应正确的 $Q_p^{(p)}$ 中的列有 $y_1 = (2 \ 1 \ 2)^T$, $y_2 = (3 \ 1 \ 2)^T$, $y_3 = (2 \ 2 \ 2)^T$, $y_4 = (3 \ 2 \ 2)^T$, 且 $S_{y_1} = \{r_1 r_2 r_4 r_6 r_7\}$, $S_{y_2} = \{r_1 r_2 r_3 r_4 r_6 r_7\}$, $S_{y_3} = \{r_1 r_2 r_4 r_5 r_6 r_7\}$, $S_{y_4} = \{r_1 r_2 r_3 r_4 r_5 r_6 r_7\}$ 且 $\bigcap_{h=1}^4 S_{y_h} = S_{y_1}$, 所以 $x = r_1 \vee r_2 \vee r_4 \vee r_6 \vee r_7 = (2 \ 1 \ 2)^T$.

4 小结与讨论

本文应用 Q 矩阵理论中潜在 Q 矩阵的列的累赘表达式, 证明在理想反应模式下, 可以标定二值和多值 Q 矩阵中未知元素. 对于多值 Q 矩阵的标定, 以前未见报道. 对于“翻新”数据, 本文新方法不一定可以使用, 因为“翻新”数据无法保证测验 Q 矩阵中包含可达阵, 也就是说, 此时的测验 Q 矩阵不一定是必要 Q 矩阵, 而可达阵是使用新方法的前提.

如上所述, 定理 1 的条件很苛严, 面对一般的观察反应模式, 如何能够应用这个定理对 Q 矩阵进行标注? 是否使用修改以后层级相合性指标 (MH-CI)^[18] 对观察反应模式进行筛选? 试验将 MHCI 等于 0.95 作为分界点, 只有大于 0.95 的观察反应模式才能够入选, 但是效果不太稳定. 究其原因, 可能和题目的猜测、失误参数, 即题目质量息息相关. 当然上述测验 Q 矩阵是由已知的可达阵和未知的潜在 Q 矩阵的 1 列构成, 命题专家可以好好打造可达阵的列对应的题目, 使它们的质量比较好, 但是这个

未知列对应的题目的质量很难保证,因此使用这种方法标注很难成功.因此,寻找一种解决方案,能够使定理1从理想状态扩展到实际应用状态,是十分有意义的.

另外,由于定理中要求使用质量好的可达阵(每1列不仅标定准确而且对应的题目猜测和失误参数都很小),而每一次标定新题属性向量都要使用可达阵作为“基准”,所以可达阵会频繁使用,这在实际使用中必定会面临可达阵对应的题目频频曝光问题.解决的方案是命制足够多的可达阵的列备用.至于潜在 Q 矩阵中列的累赘表达式的其他应用以及简洁表达式的应用,值得进一步研究.

5 参考文献

- [1] Leighton J P, Gierl M J, Hunka S M. The attribute hierarchy method for cognitive assessment: a variation on Tatsuoka's rule-space approach [J]. Journal of Educational Measurement 2004 41(3): 205-237.
- [2] Ding Shuliang, Luo Fen, Cai Yan, et al. Complement to Tatsuoka's Q matrix theory [C]//Shigemasu K, Okada A, Imaizumi T, et al. New Trends in Psychometrics. Tokyo: Universal Academy Press 2008: 417-423.
- [3] 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): 87-91 96.
- [4] Tatsuoka K K. Architecture of knowledge structure and cognitive diagnosis: a statistical pattern recognition and classification approach [C]//Nichols P D, Chipman S F, Brennan R L. Cognitively diagnostic assessment. Hillsdale, NJ: Erlbaum, 1995: 327-361.
- [5] Tatsuoka K K. Cognitive assessment: an introduction to the rule space method [M]. New York: Taylor & Francis Group 2009.
- [6] Tatsuoka K K. Toward an integration of item-response theory and cognitive error diagnosis [C]//Frederiksen N, Glaser R, Lesgold A, et al. Diagnostic monitoring of skill and knowledge acquisition. Hillsdale, NJ: Erlbaum, 1990: 453-488.
- [7] de la Torre J. An empirically based method of Q -matrix validation for the DINA model: development and applications [J]. Journal of Educational Measurement 2008 45(4): 343-362.
- [8] De Carlo L T. Recognizing uncertainty in the Q -matrix via a Bayesian extension of the DINA model [J]. Applied Psychological Measurement 2012 36(6): 447-468.
- [9] 涂冬波, 蔡艳, 戴海琦. 基于 DINA 模型的 Q 矩阵修正方法[J]. 心理学报 2012 44(4): 558-568.
- [10] 汪文义, 丁树良, 游晓峰. 计算机化自适应诊断测验中原始题的属性标定[J]. 心理学报 2011 43(8): 964-976.
- [11] 喻晓峰, 罗照盛, 秦春影 等. 基于作答数据的模型参数和 Q 矩阵联合估计[J]. 心理学报 2015 47(3): 273-282.
- [12] Liu Jingchen, Xu Gongjun, Ying Zhiliang. Data-driven learning of Q -matrix [J]. Applied Psychological Measurement 2012 36(7): 548-564.
- [13] 陈平, 辛涛. 认知诊断计算机化自适应测验中在线标定方法的开发[J]. 心理学报 2011 43(6): 710-724.
- [14] 丁树良, 罗芬, 汪文义 等. 知识状态的不同表达及其应用[J]. 江西师范大学学报: 自然科学版 2017 41(3): 296-301.
- [15] Chen Jinsong, de la Torre J. A general cognitive diagnosis model for expert-defined polytomous attributes [J]. Applied Psychological Measurement 2013 37(6): 419-437.
- [16] Sun Jianan, Xin Tao, Zhang Shumei, et al. A polytomous extension of the generalized distance discriminating method [J]. Applied Psychological Measurement 2013 37(7): 503-521.
- [17] 丁树良, 汪文义, 罗芬 等. 可达阵功能的不可替代性[J]. 江西师范大学学报: 自然科学版 2016 40(3): 290-294 298.
- [18] 丁树良, 毛萌萌, 汪文义 等. 教育认知诊断测验与认知模型一致性的评估[J]. 心理学报 2012 44(11): 1535-1546.

A Simple Method to Specify Q Matrix

DING Shuliang, LUO Fen, WANG Wenyi, XIONG Jianhua

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

Abstract: Based on the Q matrix theory, using the decomposing of the redundant expression of a column of the potential Q matrix, it can offered a simple method to calibrate attribute vector, and the method is proved mathematically.

Key words: cognitive diagnosis; specification of Q matrix; redundant expression; Q matrix theory

(责任编辑: 冉小晓)