

文章编号: 1000-5862(2016)03-0290-05

可达阵功能的不可替代性

丁树良 汪文义 罗 芬 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: 可达阵有2个重要性质: (i) Q 矩阵中的列均可由可达阵的列线性表示; (ii) 在0-1评分、属性之间作用不可相互补偿条件下, 若可达阵(或者其列的置换)是测验 Q 矩阵的子矩阵, 则任何2个不同的属性掌握模式(知识状态)对应的理想反应模式仍然不同. 在 Q 矩阵当中, 是否有其他的 K 阶子矩阵, 具有其中1个或者2个性质? 这对于认知诊断测验蓝图设计和计算机自适应诊断测验(CD-CAT)的选题策略的制定非常重要. 但是, 十分遗憾, 可以证明这2个性质都不可以由其他 Q 矩阵代替. 在一定条件下必要 Q 矩阵才能够表示知识结构, 才能够提高认知诊断测验的构念效度.

关键词: 可达矩阵; 不可替代性; 充分 Q 矩阵; 必要 Q 矩阵

中图分类号: B 841 文献标志码: A DOI: 10.16357/j.cnki.issn1000-5862.2016.03.15

0 引言

本文使用的记号, 如邻接矩阵 A , 可达阵 R , 潜在 Q 矩阵 Q_p , 学生 Q 矩阵 Q_s , 以及测验 Q 矩阵 Q_t , 其定义均与文献[1]相同. 本文假设考察的领域中包含且仅包含 K 个属性.

由扩张算法^[2-4]可知, 所谓潜在 Q 矩阵(Q_p)中由 R 扩张出来的部分(以下简称为扩张部分)的每一列均和可达阵 R 的列不相同, 而 Q_p 和学生 Q 矩阵(Q_s)的列 α 均可由可达阵 R 的列线性表示(组合系数为0, 1). 在本文中, 将 R 中参与表示 α 的列(比如 r_j)称为 α 的合成向量, 从而 α 的任何一个合成向量 $r_j \leq \alpha$, 即 r_j 的每一个分量不超过 α 的相应分量, 显然 r_j 的长度均不大于 α 的长度. 如果 α 是经由 R 扩张出来的列而不是 R 的列, 那么 α 的长度一定大于参与合成 α 的向量的长度, 这里用欧式范数定义向量的长度. 还有研究认为, 在0-1评分和属性之间无补偿作用条件下, 可达阵在认知诊断测验蓝图的设计中, 起着重要作用, 即 R 作为测验 Q 矩阵的子矩阵, 可以使知识状态和理想反应模式一一对应, 更加确切地说, 任何一个具有知识状态 α 的被试, 参加测验蓝图为 R 的认知诊断测验, 在不受任何随机干扰的条件下, 其作答反应模式的转置仍然是 α ^[5-6]. 这

表明, 可达阵在认知诊断测验中起着重要作用.

然而, 是否有且仅有 R 才能够表达 Q_s 中所有的列? 是否在上述条件下, 仅有 R 才能够将知识状态 α 映射为理想反应模式 α 的转置? 即 R 是否可以由其他 K 阶0-1矩阵(或者更加严格地说 Q_p 中 K 阶子矩阵)代替? 这个问题不仅仅在理论研究中很有趣, 而且在众说纷纭的认知诊断测验蓝图设计的讨论中很重要.

认知诊断测验蓝图设计问题, 本质上是测验 Q 矩阵的设计问题. 假设测验 Q 矩阵 Q_t 为 K 行 m 列, Q_t 的每一列代表一类题目的属性向量, 为了行文简洁, 下文有时候将 Q_t 的列称为题目(项目), 将 Q_s 的列称为被试.

通常 Q 矩阵的元素是0-1, 但是最近有人讨论元素为非负整数的多值 Q 矩阵^[7-10]. 如无另外声明, 本文讨论0-1矩阵.

在众多测验中, 诊断测验或许是最复杂的^[11]. 这种复杂性产生的原因, 有观念上的, 也有技术上的: 习惯于编制成就测验的专家对认知诊断测验感到陌生; 认知诊断测验蓝图的设计复杂且和成就测验的双向细目表有较大差异. J. S. Gorin^[11]认为, 好的诊断测验要从被试的作答反应中获得其知识方面的缺陷和不足的信息, 找出他们为什么作出这种反应的原因. 好的认知诊断测验要有“穿透性”(penetration), 即从测验成绩获取包括概念、知

收稿日期: 2016-03-29

基金项目: 国家自然科学基金(31500909, 31360237, 31300876, 31160203, 31100756, 30860084, 11401271), 教育部人文社会科学青年基金(13YJC880060)和江西省教育科学2013年度一般课题(13YB032)资助项目.

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

识表示和认知加工的心理信息,提供个体在各个知识和加工上的水平信息。

认知诊断中常用一个 K 维 0-1 向量表示被试的属性掌握状况,称之为知识状态,如果其第 j 个分量等于 1,表示被试掌握了第 j 个属性,否则这个分量等于 0。一个测验的反应数据是带有随机性的,一般来说,带有随机性的数据的处理比不带随机性的数据的处理更加困难。一个好的认知诊断测验的设计,至少应该在测验项目(题目)质量相当好而且被试反应不受随机干扰(即反应数据是确定性)时,能够准确获得被试知识状态。这是好的认知诊断测验设计应该满足的必要条件。不带随机误差的反应模式称为理想反应模式。用数学语言描述,好的认知诊断测验的设计应该满足的必要条件是,对于任意 2 个不同的知识状态,它们在这个测验之下的理想反应模式也不相等。对于任意 2 个不相等的知识状态 α 与 β ,有 $\alpha \circ Q_i \neq \beta \circ Q_i$,这里 Q_i 为测验 Q 阵 $\alpha \circ Q_i$ 表示知识状态为 α 的被试在测验 Q 矩阵为 Q_i 的认知诊断测验上的理想反应模式。若将 Q_i 按列剖分,且 q_j 为其第 j 个列,则 $\alpha \circ Q_i$ 的第 j 个分量为 $\alpha \circ q_j$,有时候也记为 $(\alpha \circ Q_i)_j$,显然 $\alpha \circ q_i$ 中 α 与 q_j 的位置不可以交换。

在 2007 年以前,有一些关于认知诊断测验编制(或者称为认知诊断测验蓝图的设计)的讨论^[12-15],有的不一定很成功,如 R. Henson 等^[15]虽然给出一种认知诊断测验选题和组卷的方法,但是从他们的模拟结果来看,当属性数目等于 8 时,模式判准率在 0.232 ~ 0.450 之间,这样低的模式判准率,很难说是十分成功的方法;J. P. Leighton 等^[14]的方案(包括 M. J. Gierl^[16]的方案)是将潜在 Q 矩阵 Q_p 作为测验 Q 矩阵,这个方案对于结构紧密^[4]的属性层级结构是合理的,但是对于属性数目 K 比较大(比如 $K > 8$)且结构松散的属性层级结构 Q_p 的列数太多,测验不可能安排这么多题目;因此,如何从 Q_p 中寻找尽可能少的题目以代表 Q_p 是一个重要的问题。

F. Samejima^[12]从属性数目等于 4 的例子入手,K. K. Tatsuoaka^[13]从一般情况入手,均认为单位矩阵作为测验 Q 矩阵的子矩阵,则知识状态和理想反应模式一一对应。但是 K. K. Tatsuoaka^[13]认为通常 Q 矩阵不会具有这样十分简单的形式,并且她给出充分 Q 矩阵的概念,即如果 Q 矩阵的行表示属性 Q 矩阵的行的关于包含关系的逐对比较产生可达矩阵,则称这个 Q 矩阵为这个感兴趣的领域内表达认知模型是充分(sufficient)的^[13]。K. K. Tatsuoaka 认为充分 Q 矩阵是知识结构的核心,可以提升认知诊

断测验的效度^[13]。知识结构是题目的集合和这个题目集合的结构,如果沿着 J. P. Leighton 等^[14]的思路,知识结构就是属性及其层级关系。但是充分 Q 矩阵并没有预期那样发挥巨大的作用,这可以从下面的反例看出:一方面它不一定可以将属性及其层级关系可以对应的题目都表达出来;另一方面充分 Q 矩阵并不比非充分 Q 矩阵的理论构念效度^[17]更高。

例 1 4 个独立属性 $A_i, i = 1, 2, 3, 4$ 学生 Q 矩阵 Q_s 包含 16 个不同的知识状态,给出 2 个测验 Q 矩阵 Q_1 是充分 Q 矩阵 Q_2 不是可达阵,但 Q_2 表达的属性层级结构中 A_1, A_2, A_3 相互独立,而 A_3 是 A_4 的先决属性,其中 Q_1 和 Q_2 为

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad Q_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

16 个不同的知识状态在 Q_1 上仅仅产生 10 个不同的理想反应模式,而在 Q_2 上却产生 12 个不同的理想反应模式。由此可知 Q_2 的理论构念效度高于 Q_1 的理论构念效度。这表明,充分 Q 矩阵既不能提供 4 个独立属性应该表达的 16 种知识状态,从而无法准确表达知识结构;又不能够提高测验效度。因为认知诊断测验的目的是探查被试的属性掌握情况,所以对所探查的知识状态的准确估计的程度应该是其效度。此例说明充分 Q 矩阵并不比非充分 Q 矩阵具有更高的效度。

其他人的研究结果也没有引起足够的重视。M. J. Gierl^[16]说 2006 年由 Steven M. Downing Thomas M Haladyna 编辑的 Elebaum 出版的《测验开发手册》(Handbook of test development)里面没有关于认知诊断测验开发的章节,也没有诊断评估的词条,尽管编辑对测验开发的概念、实践和创新提供了出色的综述,并且声称他们描述了 21 世纪测验开发的最新成果(the state of the art)。

Chiu Chia-Yi 等^[18]给出包含单位矩阵的设计原则,这对于属性层级关系是独立型结构的情形,是一个合理的解决方案,但是对于其他属性层级结构^[5-6, 14],单位阵并不是 Q_p 的子矩阵,从而这样做不是正确的解决方案。

M. J. Madison 等^[19]认为认知诊断评估的设计是 Q 矩阵设计,并给出 Q 矩阵设计的定义是根据每一个题目测量的特定的属性子集仔细地安排一组测验题目。他们认为 Q 矩阵设计的基本特征包括评估的题目数和测量的属性数,其他特征影响 Q 矩阵的复杂性(complexity),一般来讲 Q 矩阵中非零元增

加则复杂性上升;复杂性因测量每一个属性的题目数、每一个题目测量的属性数和其他属性联合测量的属性数的变化而变化. 他们一方面和 Chiu Chia-Yi 等^[18] 的观点基本一致,认为测验 Q 矩阵包含单位矩阵能够改善分类准确性、可靠性和收敛速度;另一方面他们意识到对于不同的属性层级结构,他们的这种设计有缺陷,于是他们提出一个解决方案,即通过属性“打包”的方式,将若干个粒度比较细的属性组合在一起,形成一个粒度比较大的属性,从而将其他属性层级结构化成独立结构. 笔者认为他们关于 Q 矩阵设计的定义不完整,没有说清楚仔细安排一组测验题目的目的,这样无法考察 Q 矩阵是否设计合格;另外他们提出将非独立层级结构的若干属性“打包”以变成独立层级结构的解决方案是很难行得通的,特别是完全不适用 J. P. Leighton 等所定义的属性层级结构,因为 J. P. Leighton 等给出的层级结构都有一个公共的先决属性,按照“打包”方案 J. P. Leighton 等^[14] 的属性层级结构对应的只是一个属性,而这样又和认知诊断测验欲诊断各个属性的掌握情况的目的相违背.

对于 0-1 评分,在属性之间不存在补偿关系条件下,丁树良等^[5-7] 指出在可达阵 R 作为测验 Q 矩阵时,任何一个属性掌握模式(知识状态)在这个测验上的理想反应模式的转置仍然是这个属性掌握模式,从而建立可达阵 R 可以使知识状态和理想反应模式一一对应的结论. 这个结论包含了前面的一些结果,而且这个结论使得前面那些难以解决的问题^[12-14] 迎刃而解,但是,仔细审视这个结论,笔者认为还是存在应该解决的如下 2 个问题:

(i) 学生 Q 矩阵(Q_s)中任何一列均可以表成可达阵 R 的列的线性组合(组合系数只能是 0 或者 1),问题是潜在 Q 矩阵 Q_p 中是否还存在另外的 K 阶方阵 Q_0 ,使得 Q_s 中任何一列均可表成 Q_0 的列的线性组合?

(ii) 任给学生 Q 矩阵中一个知识状态 α ,有 $\alpha \circ R = \alpha$ ^[5],如果有 Q_1 是 Q_p 的 K 阶子矩阵,并且有 $\alpha \circ Q_1 = \alpha$,问题是 Q_1 是否必须等于 R ?

第 1 个问题是讨论 Q_p 中是否存在一个子矩阵可以代表 Q_p ,这个问题的解决可以作为解决 J. P. Leighton 等^[14] 方案困境的备选策略.

第 2 个问题的解决,可以使得认知诊断测验的设计和题库建设方案灵活多样,而且对于构建有认知诊断功能的计算机化自适应测验(CD-CAT)的选题策略有帮助.

上述 2 个问题实际上是可达阵 R 的可替代性问题.

1 第 1 个问题的解

J. P. Leighton 等^[14] 和 M. J. Gierl^[16] 甚至 J. S. Gorin^[11] 都建议使用潜在 Q 矩阵 Q_p 作为测验 Q 矩阵,如前所述这个方案的缺陷是:有时候 Q_p 的列比较多,测验不可能那么长. 因为 Q_p 中任何一列都可以用可达阵 R 的列表达,即 R 是代表,所以前面建议用包含可达阵 R 的测验 Q 矩阵(称之为充分必要 Q 矩阵)代替,如果频频使用可达阵 R ,这容易使 R 对应的题目(即题目的属性向量和可达阵的列对应)过度曝光,危害测验的安全性. 所以产生第 1 个问题,存在其他的 K 阶子矩阵,可以将 Q_p 的列都表达出来?这时候当然只要考虑除可达阵 R 之外,还有扩张出来的部分属性层级结构. 记扩张出来的部分为 Q_e .

定理 1 设 Q_0 是 Q_p 的一个 K 阶子矩阵,则 Q_s 中任何一列可以由 Q_0 的列线性表示当且仅当 $Q_0 = R$.

证 由于可达阵可以表示成为上三角矩阵,要将 Q_s 中所有列表示出来,势必要将可达阵的列表示出来,所以不妨假设 Q_0 是上三角矩阵.

“ \Leftarrow ”充分性. 由扩张算法^[2-4] 立即可得.

“ \Rightarrow ”必要性. 只须证明其逆否命题,即若 Q_0 不等于 R ,则 Q_s 中存在一列 q ,使得 q 不可由 Q_0 的列线性表示.

首先注意到 R 是 Q_s 的子矩阵,而且 R 可以排列成为对角元素全为 1 的上三角 0-1 矩阵,故 R 的任意一列均不可能由 R 的其他列表示.

若 Q_0 不等于 R ,则 Q_0 至少缺少 R 中某一列,如 r_j ,注意 r_j 的第 j 个分量 $r_{jj} = 1$,而对于所有 $t > j$ $r_{tj} = 0$,对凡是第 j 个分量等于 1 并且 $t > j$ $q_{tj} = 0$ 的列 q (如果它存在),注意到 q 是经过 R 扩张出来的列,从而 q 的长度大于 r_j ,所以 r_j 不可能由 Q_0 的列表示. 如果 Q_0 中根本不存在第 j 个分量等于 1,并且所有 $t > j$ 第 t 个分量均等于 0 的向量,当然 Q_0 的列不可能表出 r_j .

2 第 2 个问题的解

定理 2 在 0-1 评分并且属性之间不存在补偿作用条件下,设 Q_1 是 Q_p 的 K 阶子矩阵且 α 是 Q_s 中的任意 1 个列向量,即是一个属性掌握模式(知识状态),则 $\alpha \circ Q_1 = \alpha \Leftrightarrow Q_1 = R$.

证 “ \Leftarrow ”. 设 $\alpha \circ Q_1 = \beta^T \alpha = (\alpha_1 \alpha_2 \cdots \alpha_K)^T \beta = (\beta_1 \beta_2 \cdots \beta_K)^T$,要证对所有 $i = 1, 2, \cdots$,

$K \alpha_i = 1 \Leftrightarrow \beta_i = 1$ 且 $\alpha_i = 0 \Leftrightarrow \beta_i = 0$.

由于 α 和 β 都是 0-1 向量, 于是 $\alpha_i = 1 \Leftrightarrow \beta_i = 1$ 的逆否命题是 $\alpha_i = 0 \Leftrightarrow \beta_i = 0$, 所以只需证明 $\alpha_i = 1 \Leftrightarrow \beta_i = 1$.

记 $Q_1 = (q_1 \ q_2 \ \cdots \ q_K) = (r_1 \ r_2 \ \cdots \ r_K)$ $\beta_i = 1 \Leftrightarrow \alpha \circ r_i = \alpha \circ q_i = 1 \Leftrightarrow r_i = q_i \leq \alpha \Leftrightarrow r_i$ 参与 α 的复合, 即 r_i 是 α 的一个合成向量, 且由于 $r_{ii} = 1 \Rightarrow \alpha_i = 1$.

由 $\alpha_i = 1$ 知 r_i 是 α 的一个合成向量 $\Rightarrow r_i \leq \alpha$, 因此 $\beta_i = 1$.

由于 $\beta_i = 0 \Leftrightarrow \alpha_i = 0$ 是 $\beta_i = 1 \Leftrightarrow \alpha_i = 1$ 的逆否命题, 故由上述证明知对一切的 $\alpha \in Q_s$ $\alpha \circ R = \alpha$.

“ \Rightarrow ”. 反之, 如果 K 阶方阵 Q_1 对任意 $\alpha \in Q_s$, 有 $\alpha \circ Q_1 = \alpha^T$, 要证 $Q_1 = R$, 令 $\alpha \circ Q_1 = \beta^T = (\beta_1, \beta_2, \cdots, \beta_K)$ $\alpha^T = (\alpha_1, \alpha_2, \cdots, \alpha_K)$.

以下总是假设 Q_1 是梯形矩阵且如果 Q_1 不等于 R :

(i) 如果 Q_1 可以排成对角元均为 1 的上三角阵 $Q_1 = (q_1 \ q_2 \ \cdots \ q_K)$ 则 Q_1 存在 1 列, 比如 $q_j \ q_j$ 不等于 r_j , 于是由 $q_{jj} = r_{jj} = 1$ 而 q_j 不等于 r_j , 可知 r_j 是 q_j 的合成向量, 于是 $r_j^T r_j < q_j^T q_j$, 从而 $r_j \circ q_j = 0$, 即 $r_j \circ Q_1$ 的第 j 个分量 $\beta_j = (r_j \circ Q_1)_j = 0 \neq r_{jj} = 1$. 所以当 Q_1 为对角元均为 1 的上三角阵时, 不可能对所有的 Q_s 中的列 α 使 $\alpha \circ Q_1 = \alpha^T$; 即 $r_j \circ Q_1$ 的第 j 个分量 $\beta_j = (r_j \circ Q_1)_j = 0 \neq r_{jj} = 1$.

(ii) Q_1 不可以化成对角元等于 1 的对角矩阵, 比如存在第 $j+1$ 列 q_{j+1} , 使得 $q_{jj+1} = 1$, 但对于任意 $t > j$ $q_{tj+1} = 0$ 这时,

(a) 因为可以对 Q_1 的列进行置换, 不妨设 $q_j = q_{j+1}$ 则考察 q_j 在测验 Q 矩阵上的理想反应模式, 知 $\beta_{j+1} = (q_j \circ Q_1)_{j+1} = q_j \circ q_{j+1} = 1 > q_{j+1j+1} = 0$. 由此可知存在 $\alpha \in Q_s$ 使得 $\alpha \circ Q_1 \neq \alpha^T$.

由此可见, 如果 Q_1 与 R 相比, 仅仅是第 $j+1$ 列不同, 且 $q_{j+1} = r_j$ (也就是 Q_1 是将 R 中某 1 列用其他 $K-1$ 列中的 1 列代替, 其他 $K-1$ 列保持不变, 即 Q_1 中有 2 列相同) 也不能够满足任意 $\alpha \in Q_s$ 使 $\alpha \circ Q_1 = \alpha^T$.

(b) 若 $q_j \neq q_{j+1}$ 则由于 $q_{jj+1} = 1$, 但 $\forall t > j$, $q_{tj+1} = 0$ 故或者 $q_{j+1} = r_j$ 或者 $(r_j \leq q_{j+1})$ 且 $(r_j \neq q_{j+1})$, 于是当 $q_{j+1} = r_j$ 时, 考察 r_j 在测验 Q 矩阵 Q_1 上的理想反应模式, 有 $r_j \circ q_{j+1} = 1 = \beta_{j+1} > r_{j+1} = 0$, 故 $r_j \circ Q_1 \neq r_j^T$;

当 $(r_j \leq q_{j+1})$ 且 $(r_j \neq q_{j+1})$ 时, 考察 q_{j+1} 在测验 Q 矩阵 Q_1 上的理想反应模式, 知

$\beta_{j+1} = (q_{j+1} \circ Q_1)_{j+1} = 1 > q_{j+1j+1} = 0$, 则可知 $\exists \alpha \in Q_s$ 使得 $\alpha \circ Q_1 \neq \alpha^T$. 定理证毕.

注意, 可达阵 R 经过列的置换以后记为 R_1 , 则 R_1 作为测验 Q 阵的子矩阵也可以使不同的知识状态对应的理想反应模式不同, 但是任给一个知识状态 α $\alpha \circ R_1 \neq \alpha^T$, 因为 $\alpha \circ R_1$ 只能对应 α^T 的一个置换.

3 小结与讨论

对于 0-1 评分和属性之间不存在补偿的条件下, 对已有的定量认知诊断测验蓝图的编制原则 (方案) 和它们的缺陷进行归纳和总结: 第 1 类是在测验 Q 矩阵中植入单位矩阵而不注意属性不同的层级结构, 须知在其他属性层级结构之下, 某些单位矩阵的列是不可能纳入测验 Q 矩阵的; 第 2 类是将潜在 Q 矩阵作为测验 Q 矩阵, 而有的属性层级结构测验 Q 矩阵的列数太多; 第 3 类是将可达阵植入测验 Q 矩阵, 虽然避免了上述 2 类缺陷, 但是没有考虑可达阵是否可以由其他矩阵代替的问题. 本文针对第 3 类认知诊断测验蓝图设计方案的不足, 讨论可达阵的替代性问题, 证明了可达阵的 2 种不可替代性: 第 1 个不可替代性是对所有元素为 0-1 的 Q 矩阵, 当且仅当可达阵的列才可以将潜在 Q 矩阵的列线性表示; 第 2 个不可替代性是对 0-1 评分且属性之间不可补偿条件下, 当且仅当可达阵 R 作为测验 Q 矩阵的子矩阵才能够将知识状态和理想反应模式一一对应. 这 2 个不可替代性进一步反映了可达阵的重要作用. 第 2 个不可替代性表明在 0-1 评分且属性之间的作用不可以相互补偿条件下, 将可达阵作为测验 Q 矩阵的子矩阵可以使知识状态和理想反应模式一一对应. 这样设计的测验 Q 矩阵应该称为必要 Q 矩阵 (necessary Q matrix). 这个名字比以前说的充分必要 Q 矩阵更加准确, 更加简洁.

由于理想反应模式不带随机误差, 而观察反应模式往往带有随机误差, 所以定理 2 只是在一定条件下, 好的认知诊断测验编制的必要条件. 它在实际应用中能够起多大作用, 还应该结合实际数据进行分析. 实际应用时认知诊断测验的题目质量肯定非常重要, 题目质量越差, 观察反应模式偏离理想反应模式越远, 定理 2 对认知诊断测验编制的指导作用越小, 也就是说在这种情况下, 纵使可达阵作为测验 Q 矩阵的子矩阵, 对于提高认知诊断测验的模式判断率的作用有限; 另外认知诊断测验蓝图如何编制, 至少必须考虑测验的长度和是否利用选择项中的诊断信息, 显然题目越多, 测验误差的控制越好. 特别是独立型属性结构, 可达阵对应的题目的确比

较简单,当 K 比较大时,挑选什么样的题目属性向量,很有讲究,罗欢等^[20]提出过一个方案,是否有更加合适的方案,如何平衡属性数目和题目数,值得讨论;至于利用选择项的诊断信息的认知诊断测验蓝图的编制问题,李瑜等^[21-22]提出多项选择题认知诊断测验的编制原理和方法,认为既要有干扰项和正答项的属性向量相似,又要有干扰项和正答项的属性向量区分。

定理 2 的充分性部分于 2010 年就已经提出^[6],只不过那时候的证明比较繁琐,需要几条引理才能够导出。本文的证明比较简洁易懂。

定理 2 的成立是有条件的,如果违背属性之间作用不可补偿的条件,即属性之间可以补偿或者部分属性之间作用可以补偿的情况,或者 0-1 评分的条件变成多级评分条件,甚至变成多值 Q 矩阵的条件,认知诊断测验蓝图的设计原理的讨论是一个非常重要而又具有挑战性的论题。当然多值 Q 矩阵的条件下,也有相应的扩张算法^[7],定理 1 的讨论可能比较容易。

至于属性“打包”的问题,虽然不能够将其他属性层级结构化成独立结构,但是却可能可以解决属性过多、认知诊断模型难于处理的问题^[23]。

以上讨论均假设 Q 矩阵的元素标注正确,而这是十分理想的情况,实际情况可能与此有相当大的出入。如何修正 Q 矩阵的元素,是另外的重要议题,但囿于本文的主旨,在此不做讨论。

北京师范大学张淑梅教授指出,可达阵的列进行置换后,在 0-1 评分及属性之间不可相互补偿条件下,仍然可以使得不同的知识状态对应不同的理想反应模式,弥补了本文的不足。在此表示感谢。

4 参考文献

- [1] 丁树良,汪文义,罗芬. 认知诊断中 Q 矩阵和 Q 矩阵理论 [J]. 江西师范大学学报: 自然科学版, 2012, 36(5): 441-445.
- [2] Ding Shuliang, Luo Fen, Cai Yan, et al. Complement to Tatsuoka's Q matrix theory [A]. Shigemasa K, Okada A, Imaizumi T, et al. New Trends in Psychometrics [C]. Tokyo: Universal Academy, 2008: 417-423.
- [3] 丁树良,祝玉芳,林海菁,等. Tatsuoka Q 矩阵理论的修正 [J]. 心理学报, 2009, 41(2): 175-181.
- [4] 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.
- [5] 丁树良,汪文义,杨淑群. 认知诊断测验蓝图的设计 [J]. 心理科学, 2011, 34(2): 258-265.
- [6] 丁树良,杨淑群,汪文义. 可达矩阵在认知诊断测验编制中的重要作用 [J]. 江西师范大学学报: 自然科学版, 2010, 34(5): 490-494.
- [7] 丁树良,罗芬,汪文义,等. 0-1 和多值可达矩阵的性质及应用 [J]. 江西师范大学学报: 自然科学版, 2015, 39(1): 64-68.
- [8] 蔡艳,涂冬波. 属性多级化的认知诊断模型拓展及其 Q 矩阵设计 [J]. 心理学报, 2015, 47(10): 1300-1308.
- [9] Chen Jinsong, Jimmy de la Torre. A general cognitive diagnosis model for expert-defined polytomous attributes [J]. Applied Psychological Measurement, 2013, 37(6): 419-437.
- [10] Sun Jia'nan, Xin Tao, Zhang Shumei, et al. A polytomous extension of the generalized distance discriminating method [J]. Applied Psychological Measurement, 2013, 37(7): 503-521.
- [11] Gorin J S. Test construction and diagnostic testing [A]. Leighton J P, Gierl M J. Cognitive diagnostic assessment for education: theory and applications [C]. Cambridge: Cambridge University Press, 2007.
- [12] Samejima F. A cognitive diagnosis method using latent trait models: competency space approach and its relationship with DiBello and Stout's unified cognitive-psychometric diagnosis model [A]. Nichols P D, Chipman S F, Brennan R L. Cognitively diagnostic assessment [C]. Hillsdale, NJ: Erlbaum, 1995: 391-410.
- [13] Tatsuoka K K. Architecture of knowledge structure and cognitive diagnosis: a statistical pattern recognition and classification approach [A]. Nichols P D, Chipman S F, Brennan R L. Cognitively diagnostic assessment [C]. Hillsdale, NJ: Erlbaum, 1995: 327-361.
- [14] 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.
- [15] Henson R, Douglas J. Test construction for cognitive diagnosis [J]. Applied Psychological Measurement, 2005, 29(4): 262-277.
- [16] Gierl M J. Making diagnostic inferences about cognitive attributes using the rule-space model and attribute hierarchy method [J]. Journal of Educational Measurement, 2007, 44(4): 325-340.
- [17] 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, 43(11): 1535-1546.
- [18] Chiu Chia-Yi, Douglas J A, Li Xiaodong. Cluster analysis for cognitive diagnosis: theory and applications [J]. Psychometrika, 2009, 74(4): 633-665.

(下转第 298 页)

- 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)