

文章编号: 1000-5862(2015)04-0365-06

# 多值 $Q$ 矩阵理论

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:** 罗列了特殊多值  $Q$  矩阵与 0-1 可达阵的相互转换算法 给出多值扩张算法和多值理想反应模式的计算方法. 在多值扩张算法的基础上, 证明了在一定条件下, 多值拟可达阵作为测验  $Q$  矩阵的子矩阵, 可以使得多值的知识状态和理想反应模式一一对应, 从而可指导多级评分认知诊断测验蓝图编制.

**关键词:** 多值  $Q$  矩阵;  $Q$  矩阵理论; 扩张算法; 认知诊断测验蓝图

**中图分类号:** B 841.7; TP 301.6 **文献标志码:** A **DOI:** 10.16357/j.cnki.issn1000-5862.2015.04.07

## 0 引言

K. K. Taksuoka<sup>[1-2]</sup> 建立规则空间模型 (rule space model, RSM), 包括  $Q$  矩阵理论和分类方法. 该  $Q$  矩阵理论用 0-1 布尔矩阵  $Q$  描述项目与属性的关联, 并试图构建知识状态 (KS) 与理想反应模式 (IRP) 的等价对应, 而通过分类方法使得观察反应模式 (ORP) 与理想反应模式对应, 最终获得可观察的反应模式的解释, 即获得隐藏在观察反应模式背后的知识状态, 并期望规则空间模型中的  $Q$ -矩阵理论, 能够起到桥梁作用——将可观察的反应向量对应到不可观察的知识状态<sup>[1]</sup>.

悉知 根据知识状态和测验  $Q$  矩阵 ( $Q_i$ ), 可以计算理想反应模式, 即测验  $Q$  矩阵可以建立知识状态到理想反应模式之间的联系, 但有可能若干个知识状态对应同一个理想反应模式 (这时是一种多对一的对应). 此时, 欲由理想反应模式寻找对应于它的知识状态是困难的. 这表明, 有些测验  $Q$  矩阵建立的知识状态到理想反应模式之间的对应不是一个一一映射. 这 2 个集合之间的一一映射是否有什么优良性质; 若有某种优良性质, 是否存在这样的  $Q_i$ , 能够实现一一映射, 或者说如何构造这样的测验  $Q$  矩阵. 这是笔者感兴趣的 2 个问题.

首先, 考察这种一一映射的优良性. 如果通过认知诊断模型 (CDM), 建立起观察反应模式与理想反

应模式之间的联系 (ORP $\rightarrow$ IRP), 又有办法建立理想反应模式与知识状态之间的联系 (IRP $\rightarrow$ KS), 则就可以建立 ORP $\rightarrow$ IRP $\rightarrow$ KS 这种关联, 从而建立起 ORP 与 KS 的联系, 即达到诊断分类的目的. 当观察反应模式中噪音较小时, 观察反应模式必然和理想反应模式相当接近, 这时由认知诊断模型实现观察反应模式和理想反应模式之间的对应, 如果一个理想反应模式只是对应一个知识状态, 则观察反应模式和知识状态的对应 (映射) 就相当准确. 故建立这种 KS 到 IRP 之间的一一对应有利于准确地进行诊断. 这可以更好地达到 K. K. Taksuoka<sup>[1-2]</sup> 建立  $Q$  矩阵的目的. 所以将这种建立 KS 和 IRP 之间一一映射的  $Q_i$  称为优良的测验  $Q$  矩阵. 当属性之间不可以补偿并且采用 0-1 评分方式, 测验  $Q$  矩阵包含可达阵时, 可以建立 KS 和 IRP 之间一一映射, 丁树良等<sup>[3]</sup> 称包含可达阵的测验  $Q$  矩阵为充分必要  $Q$  矩阵. 其实, 充分必要  $Q$  矩阵的核心是测验  $Q$  矩阵是否能够建立 KS 和 IRP 之间的一一映射.

其次, 考察是否存在这样的  $Q_i$  能够实现一一映射. 通常认为  $Q$  矩阵是诊断范围内的属性 (attribute) 和项目的关联矩阵, 若正确回答项目  $j$  必须掌握属性  $i$ , 则  $q_{ij} = 1$ , 否则  $q_{ij} = 0$ . 其实,  $Q$  矩阵也是属性和知识状态的关联矩阵. 此时, 若第  $j$  种知识状态中包含属性  $i$ , 则  $q_{ij} = 1$ , 否则  $q_{ij} = 0$ . 元素仅取 0 或 1 的矩阵称为布尔矩阵. 在一定条件下, 当  $Q$  矩阵为布尔矩阵时, 已经有一些关于优良的测验  $Q$  矩阵构造

收稿日期: 2015-03-25

基金项目: 国家自然科学基金 (30860084, 31160203, 31100756, 31360237), 教育部人文社会科学研究青年基金 (13YJC880060) 和江西省教育厅科技计划 (GJJ13207, GJJ13208, GJJ13209, GJJ13226, GJJ13227) 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.

(设计)的相关研究成果<sup>[3-7]</sup>. 本文重点也是研究这个问题.

$Q_i$  矩阵又称为测验蓝图. 称对应同一个理想反应模式的知识状态的集合为一个知识状态的等价类 (equivalent class), 知识状态等价类也是由  $Q_i$  矩阵决定的. 一个知识状态等价类如果包含多个知识状态, 这时由理想反应模式到知识状态的联系 (IRP→KS) 就难以实现; 反之, 如果一个知识状态等价类只包含一个知识状态, 理想反应模式和知识状态就可以一一对应. 要做到这一点, 就必须对  $Q_i$  矩阵进行设计. K. K. Tatsuoka 给出充分  $Q$  阵和充分题库的概念, 希望提高认知诊断分类准确率, 然而充分  $Q$  矩阵和充分题库的概念并不能保证认知诊断测验的分类准确率高, 即使用充分  $Q$  矩阵的测验蓝图, 并不是提高认知诊断分类准确率的“充分”条件. 由于使用标定正确、性质优良的测验  $Q$  矩阵, 才能够达到比较准确地找到隐藏在观察反应模式后面的知识状态, 从而提高诊断准确率的目的, 所以认为 K. K. Tatsuoka 的  $Q$  矩阵理论应该拓展, 即应该包括优良认知诊断测验蓝图的设计, 甚至  $Q$  矩阵的在线辅助标定、 $Q$  矩阵的修正等内容<sup>[8-9]</sup>.

近年来有一些关于认知诊断评估的文章中引入、使用和研究多值  $Q$  矩阵<sup>[10-12]</sup>, 此时  $Q$  矩阵就不再是布尔矩阵了. 文献[12]叙述了引入多值  $Q$  矩阵的原因有: (i) 因为二值(0-1)  $Q$  矩阵不能充分反映同一个属性在不同题目中难度水平的变化, 而多值  $Q$  矩阵可以; (ii) 如果测验要考察综合的问题解决过程, 属性粒度 (grain size) 不能够太细, 因此要设计多值  $Q$  矩阵以刻画属性的大粒度. 文献[12]将认知诊断分类和  $Q$  矩阵理论进行整合, 开发了多值广义距离判别方法 (GDD-P), 并对多值的  $Q$  矩阵理论展开了讨论, 然而多值的  $Q$  矩阵理论中的相关结果并不是不证自明的, 本文要给出相应的证明. 如上所述, 测验蓝图的设计是拓展的  $Q$  矩阵理论中的重要内容. 测验蓝图的设计不仅仅和属性之间是否补偿、采用 0-1 还是多级评分有关, 而且至少还和  $Q$  矩阵采用什么样的元素(0-1 还是多值) 以及采用什么样的评分规则(或者说和理想反应的得分规则) 等因素有关.

多值广义距离判别方法 (GDD-P) 是一个准确率高的认知诊断方法, 文献[12]给出了一个新的理想反应评分方案, 即

$$S_j(\alpha) = \sum_{k=1}^K q_{kj} I(\alpha_k \geq q_{kj}), \quad (1)$$

其  $\alpha$  为被试知识状态,  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_K)$ ;  $Q = (q_{ij})_{K \times m}$  为多值  $Q$  阵,  $\alpha$  的分量和  $q_{ij}$  均为非负整数, 将  $Q$  按列剖分为  $Q = (Q_1, Q_2, \dots, Q_j, \dots, Q_m)_{K \times m}$ ,  $Q_j = (q_{1j}, q_{2j}, \dots, q_{Kj})^T$ ,  $K$  为考察范围中属性数,  $m$  为测验中的题目数, 不失一般性, 可设  $Q_1, Q_2, \dots, Q_m$  为  $m$  列互不相同的列向量,  $S_j(\alpha)$  是知识状态为  $\alpha$  的被试在测验第  $j$  题上的理想得分,  $j = 1, 2, \dots, m$ .  $I(\alpha_k \geq q_{kj})$  是一个取值 0 或者 1 的示性函数, 它表示当且仅当  $\alpha$  的第  $k$  个分量不小于  $Q_j$  中第  $k$  个分量时, 其值为 1, 这时被试在第  $j$  题上的理想得分才能增加  $q_{kj}$  分, 否则为 0 分; 理想得分是既不猜测也不失误时的得分.

定义 1 称 2 个向量满足关系式  $\alpha \leq \beta$ , 如果  $\beta - \alpha$  的分量均非负.

引理 1 对于评分方式 (1), 有  $S_j(\alpha) \leq \sum_{k=1}^K q_{kj}$ , 当且仅当  $Q_j \leq \alpha$  时,  $S_j(\alpha) = \sum_{k=1}^K q_{kj}$ .

给定理想反应的评分方式, 就可以计算理想反应模式, 由 (1) 设被试知识状态为  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_K)^T$ , 测验  $Q$  矩阵  $Q_i^{(p)}$  的第  $j$  列为  $(q_{1j}, q_{2j}, \dots, q_{Kj})^T$ . 若  $\alpha_t \geq q_{tj}$ , 则  $S_j(\alpha) \leftarrow S_j(\alpha) + q_{tj}$ ,  $t = 1, 2, \dots, K$ ,  $j = 1, 2, \dots, m$ .

如果要计算理想反应模式, 可以对题目循环:

```
for j = 1 to m
  S_j(α) = 0
  for t = 1 to K
    if α_t ≥ q_tj then S_j(α) = S_j(α) + q_tj
  next t
next j
```

文献[12]给出基于 0-1 可达阵 (记为  $R_2$ ) 和各属性的最高水平数  $w_j$  构造一个多值矩阵  $R$  的方法 (本文称之为 Sun Jia'nan 的 DPR 转换方法). DPR 转换方法可以总结为  $R_2$  的  $(j, j)$  元用  $(1, 2, \dots, w_j)$  代替  $R_2$  的第  $j$  列的其他元素  $r_{ij}$  用 1 行  $w_j$  列的向量  $r_{ij}(1, \dots, 1)$  代替,  $j = 1, 2, \dots, K$ , 这样便获得多值矩阵  $R$ . 这里记多值矩阵  $R$  为  $R_p$ , 基于  $R_p$  又导出一个多值  $Q$  阵, 以及基于  $R_p$  和扩张算法获得多值  $Q$  阵的方法. 他们还表示在理想计分方式为 (1) 条件下,  $R_p$  可以使多值理想反应模式和知识状态一一对应. 这是一个重要的  $Q$  矩阵理论方面的命题, 但是他们没有给出相应的证明. 作为一种理论的成果, 总是希望能够获得严格的证明. 本文欲解决以下几个问题:

1) 在多值情况下, 扩张算法应该如何叙述? 因

为文献[12]的相关计算结果有一些疏漏(他们列出的线型和收敛型相应的列有冗余)。

2)  $R_p$  通过扩张算法获得多值  $Q$  阵, 这样获得的多值  $Q$  阵是否包含了所有非零的多值知识状态? 等价地说, 多值扩张算法是否可以表明任何一个非零的多值知识状态, 均可以用  $R_p$  的某些列表达出来?

3)  $R_p$  在测验蓝图中所起的作用和 0-1 可达阵的作用<sup>[5]</sup>是否相当? 即是否可以建立起(多值)知识状态集合和(多值)理想反应模式集合的一一对应关系?

这几个问题应当都属于  $Q$  矩阵理论的范畴, 而  $Q$  矩阵理论是认知诊断评估的重要组成部分, 值得化力气仔细研究。当然这些问题中的重点是 3), 但是 1)、2) 是解决 3) 的基础, 所以下面逐个进行讨论。

## 1 多值扩张算法

先定义布尔并运算, 若  $\alpha, \beta$  均为  $K$  维 0-1 向量,  $\alpha \vee \beta$  的分量是  $\alpha$  和  $\beta$  相应分量的布尔并  $\alpha_k \vee \beta_k = \max(\alpha_k, \beta_k)$ , 这实际上等同于以下 4 个式子:  $0 \vee 0 = 0, 0 \vee 1 = 1, 1 \vee 0 = 1, 1 \vee 1 = 1$ 。这里定义的布尔并运算既适用于 0-1 向量, 又可用于多值向量, 有别于逻辑代数中布尔并运算。

在 0-1 矩阵中, 设  $R$  是可达阵, 基于前面定义的向量布尔并运算, 由  $R$  通过如下扩张算法<sup>[13-15]</sup>可以得到潜在  $Q$  阵  $Q_p$ , 即所有非零的知识状态, 在  $Q_p$  中增加一个零列, 记为  $Q_s$ , 即得到所有知识状态, 称  $Q_s$  为学生  $Q$  阵。

### 1.1 0-1 条件下扩张算法

将 0-1 可达阵  $R$  按列剖分:  $R = (r_1 r_2 \cdots r_K)_{K \times K}$

(i)  $B \leftarrow R, b_j \leftarrow r_j, j = 1, 2, \cdots, K$ ;

(ii)  $j \leftarrow 1$ ;

(iii) 将  $b_j$  和  $b_j$  右边所有列作布尔并, 若产生新的列, 则添加在  $B$  的最右边, 仍记为  $B$ ;

(iv)  $j \leftarrow j + 1$ ;

(v) 若  $j > K$ , 则停止, 否则转 (iii)。

### 1.2 多值条件下扩张算法

设属性  $A_j$  的最高水平为正整数  $w_j, j = 1, 2, \cdots,$

$K$ ; 记  $w = \sum_{j=1}^K w_j$ , 多值条件下扩张算法总结为

(i) 将  $R_p$  按列剖分:  $R_p = (r_1 r_2 \cdots r_w)_{K \times w}$ ;

(ii) 布尔并定义为  $\alpha_k \vee \beta_k = \max(\alpha_k, \beta_k)$ ;

(iii) 将上述扩张算法中  $R$  的列数  $K$  修改为  $w$ , 就可以导出多值扩张算法。

将基于  $R_p$  使用多值扩张算法获得的矩阵称为多值潜在  $Q$  阵(记为  $Q_p^{(p)}$ ), 在  $Q_p^{(p)}$  基础上增加一个零列得到的矩阵称为多值学生  $Q$  阵, 记为  $Q_s^{(p)}$ , 又记多值测验  $Q$  阵为  $Q_t^{(p)}$ , 不致混淆时,  $Q_p^{(p)}$  和  $Q_s^{(p)}, Q_t^{(p)}$  上标均省写, 仍记为  $Q_p, Q_s, Q_t$ 。

可以得到如下的引理。

引理 2 任取  $Q_p^{(p)}$  中 1 列记为  $\alpha$ , 则  $\alpha = \bigvee_{t=1}^h r_{i_t} r_{i_t}$

为  $R_p$  中的列, 即  $\alpha$  可以表示为  $R_p$  中列的布尔并。

证 由扩张算法<sup>[13-15]</sup>即得。

定义 2 称引理 2 中  $R_p$  的列  $r_{i_t} (t = 1, 2, \cdots, h)$  为  $\alpha$  的构成向量。

须注意  $\alpha$  的构成向量集合一般不唯一。

### 1.3 多值非零知识状态集合

由  $K$  行  $w$  列的多值矩阵  $R_p$  通过膨胀算法<sup>[11]</sup>获得  $w$  行  $w$  列的 0-1 可达阵, 及由 0-1 可达阵通过扩张算法获得相应的 0-1 潜在  $Q$  阵, 并且在已知各属性的最高水平数的条件下, 由潜在  $Q$  阵通过压缩算法<sup>[11]</sup>可获得多值潜在  $Q$  阵, 比较这个潜在  $Q$  阵和基于  $R_p$  通过扩张算法获得的多值  $Q$  阵, 便知使用多值扩张算法, 的确可以由  $R_p$  获得多值潜在  $Q$  阵, 即获得所有可能的多值非零知识状态。这表明, 多值知识状态的确可以由  $R_p$  的列通过上述的布尔并表示。

注意到  $R_p$  通过膨胀算法可以获得 0-1 可达阵(记之为  $M$ ), 而  $M$  通过压缩算法也可以获得  $R_p$ , 并且  $R_p$  通过扩张算法可以获得所有的多值知识状态, 这和 0-1 可达阵的作用相当, 所以可以称  $R_p$  为拟可达阵(quasi-reachability matrix)。

定义 3 在 0-1 可达阵的基础上, 通过文献[12]的 DPR 转换方法获得的  $R_p$  称为拟可达阵。

至于扩张算法的正确性的证明, 可参见文献[11, 16]。

## 2 拟可达阵 $R_p$ 在认知诊断测验蓝图设计中重要作用

### 2.1 $R_p$ 的重要作用

众所周知, 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设计中有重要作用, 即可达阵  $R$  作为测验  $Q$  阵  $Q_t$  的子矩阵, 必可

使知识状态和理想反应模式一一对应<sup>[3,5]</sup>,但在多级评分条件下,如果属性之间无补偿作用,在每多掌握一个项目所测属性则理想得分增加1分的记分方式下,可达阵仍可以使理想反应模式和知识状态一一对应,并且完备 $Q$ 阵可以使理想反应模式与知识状态一一对应而包含列数最少<sup>[6-7]</sup>.对于文献[12]的评分方式,他们认为拟可达阵仍可以使理想反应模式与知识状态一一对应.然而他们对此没有给出数学证明.以下给出数学证明.

设 $R_p = (r_1, r_2, \dots, r_w)$ 为 $K \times w$ 拟可达阵 $\alpha$ 为多值 $K$ 维知识状态 $\alpha$ 在以 $R_p$ 为测验蓝图时,记分方式为(1)时理想反应模式记为 $\alpha \circ R_p$ ,这里 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_K)^T$ ,  $r_j = (r_{1j}, r_{2j}, \dots, r_{Kj})^T$ , 其中 $j=1, 2, \dots, m$ .

定理1 设采用评分方式为(1),对 $R_p$ 中2个不同的列 $r_i$ 与 $r_j$ ,  $r_i \neq r_j$ ,有 $r_i \circ R_p \neq r_j \circ R_p$ ,即 $R_p$ 为测验 $Q$ 阵,对 $R_p$ 中任2个不同列对应的知识状态,它们对应的理想反应模式不同.

证 不妨记 $r_i = \alpha$ ,  $r_j = \beta$ ,  $\alpha \neq \beta$ . (i) 如果 $\alpha$ 与 $\beta$ 可以比较,即或者 $\alpha \leq \beta$ ,或者 $\beta \leq \alpha$ ,为明确计,不妨设 $\alpha \leq \beta$ ,由于 $\alpha \neq \beta$ ,故必有 $\beta$ 某个分量 $\beta_k$ 大于 $\alpha$ 的对应分量 $\alpha_k$ ,由记分方式(1)知 $\alpha$ 不可能在 $r_j (= \beta)$ 上获得满分,而 $\beta$ 在 $r_j$ 上获得满分,故 $\alpha \circ R_p \neq \beta \circ R_p$ ; (ii) 若 $\alpha$ 与 $\beta$ 不可以比较,则 $\alpha (= r_i)$ 在 $r_j (= \beta)$ 上不可能得满分,但 $\beta (= r_j)$ 在 $r_j$ 上得满分,  $\beta$ 在 $r_i$ 上不可能得满分,但 $\alpha$ 可以,由此可知 $\alpha \circ R_p \neq \beta \circ R_p$ . 定理1证毕.

定理2 若 $\alpha$ 与 $\beta$ 为多值非负 $K$ 维知识状态,且 $\alpha \neq \beta$ ,则 $\alpha \circ R_p \neq \beta \circ R_p$ .

证 为简化记号而使版面清楚起见,由引理2,不妨设 $\alpha = \bigvee_{t=1}^h r_t$ , 而 $\beta = \bigvee_{h=1}^p s_h$ , 且 $r_t$  ( $t=1, 2, \dots, h$ )和 $s_k$  ( $k=1, 2, \dots, p$ )分别为 $\alpha$ 和 $\beta$ 的构成向量,即均为 $R_p$ 中列向量,由于 $\alpha \neq \beta$ ,若 $\alpha \leq \beta$ ,但 $\alpha \neq \beta$ ,则与定理1的证明类似,可知 $\alpha \circ R_p \neq \beta \circ R_p$ . 若 $\alpha$ 与 $\beta$ 不可以比较,则可断言存在 $R_p$ 中某一列,不妨记之为 $u$ ,有 $u \leq \alpha$ ,而 $u$ 与 $\beta$ 不可以比较或 $u \leq \beta$ 而 $u$ 与 $\alpha$ 不可比较,即 $u$ 为 $\alpha$ 的构成向量而不是 $\beta$ 的构成向量(或 $u$ 为 $\beta$ 的构成向量而不是 $\alpha$ 的构成向量).实际上,若 $\alpha$ 的构成向量均为 $\beta$ 的构成向量,则必有 $\alpha \leq \beta$ ,矛盾.同理,若 $\beta$ 的构成向量为 $\alpha$ 的构成向量也会导出矛盾.为明确计,设有 $u \leq \alpha$ 但 $u \leq \beta$ 不成立,则 $\alpha$ 在 $u$ 上得满分但 $\beta$ 在 $u$ 上不可能得满分,故 $\alpha \circ R_p \neq \beta \circ R_p$ .

定理3 设 $Q_i^{(p)}$ 以拟可达阵 $R_p$ 为其子矩阵,则如果 $\alpha \neq \beta$ 有 $\alpha \circ Q_i^{(p)} \neq \beta \circ Q_i^{(p)}$ .

证 由引理2和定理2以及向量相等的定义,知结论成立.

注意到 $\alpha \neq 0$ ,有 $\alpha \circ R_p \neq 0$ ,故任意非零知识状态 $\alpha$ 与零知识状态在 $R_p$ 上的理想反应必定能够区分.

当 $R$ 和 $Q$ 均为0-1矩阵,(1)式化成掌握题目中一个属性理想评分增加一分的情形,上述结论同样成立,并可获得更强的结论<sup>[6-7]</sup>.

定义4(多值充分必要 $Q$ 矩阵) 如果采用(1)的方式对理想反应评分,称多值测验 $Q$ 矩阵为多值充分必要 $Q$ 矩阵,如果拟可达阵 $R_p$ 作为测验 $Q$ 矩阵的子矩阵,则多值充分必要 $Q$ 矩阵也可以简称为多值充要 $Q$ 阵,不至混淆时,还可以简称为充要 $Q$ 阵.

## 2.2 如何从多值 $Q$ 阵挖掘属性层级

对于布尔矩阵情形的 $Q$ 矩阵理论,有比较丰富的成果,比如当 $Q$ 矩阵的行对应属性,列对应题目时,可以通过可达矩阵行的逐对比较获得属性之间的层级关系<sup>[17]</sup>,而对于多值 $Q$ 矩阵,拟可达阵的行的逐对比较能够获得属性之间的层级关系吗?其实拟可达阵是特殊的多值 $Q$ 矩阵,相同属性的不同水平可以认为是一种层级关系(注意,这是水平之间的层级关系),如果将同一属性的所有水平对应的列收缩为一列(实际上施行文献[12]的从0-1可达阵导出拟可达阵的DPR算法的逆算法),并且收缩结果中的非零元素用1标注,对这样导出的0-1矩阵进行矩阵行的逐对比较,就可以获得属性层级关系.如果对于多值 $Q$ 矩阵,可以使用如0-1 $Q$ 矩阵的“缩减”算法<sup>[18]</sup>,将由拟可达阵通过布尔并运算生成的列删除(缩减),然后对剩余下来的多值 $Q$ 矩阵实施上述的收缩算法,就可获得属性层级关系.

## 3 结论与讨论

文献[10,12]引入多值 $Q$ 阵,这是对取值仅仅为0-1的 $Q$ 阵的发展.对同一个题目,采用多级评分可以比采用0-1评分提供更丰富的诊断信息,但除了评分方式外,认知诊断测验蓝图的设计也是一个至关重要的问题,对此迄今为止没有太多的成果.上面讨论了理想反应评分方式(1)下,拟可达阵可以使理想反应模式和知识状态一一对应的问题.这种

评分方式表明,同一属性的不同水平不可以补偿,准确地讲相同属性的低水平不能补偿高水平,但是根据评分方式(1)可知,在同一个题目上不同属性的得分是累加的,所以不同属性之间可以相互部分补偿。

丁树良等<sup>[6-7]</sup>在可达阵和 $Q$ 阵为0-1矩阵条件下,即(1)中 $q_{kj}$ 和 $\alpha_k$ 仅仅取0-1值时,讨论了用较少题目确可以获得对知识状态较准确测量的测验蓝图设计问题,给出了完备 $Q$ 阵,基本完备 $Q$ 阵的概念,除了独立型之外,其他层级关系对应的基本完备 $Q$ 阵均是可达阵的真子矩阵,即其列数少于可达阵的列数。而对于(1)中 $q_{kj}$ 和 $\alpha_k$ 取非负整数条件下,相对应的完备 $Q$ 阵和基本完备 $Q$ 阵如何构造值得讨论。一方面,课堂评估是一种形成性评估,而课堂评估往往不可能使用太多的题目,所以能够用较少的题目达到较准确的诊断的目的是很有实用价值的。另一方面,在文献[12]的评分方式下,下例表明,这时基本完备 $Q$ 阵的构造和 $q_{kj}$ 及 $\alpha_k$ 均为0或1的情况不一定相同。

例1 2个属性的线型结构,先决属性 $A_1$ 最高水平为3, $A_2$ 最高水平为2。

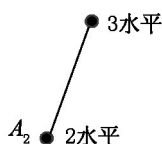


图1 2个属性多水平线型结构

对应的0-1可达阵、拟可达阵和多值潜在 $Q$ 阵

分别为:  $R_2 = \begin{pmatrix} 11 \\ 01 \end{pmatrix}$ ,  $R_p = \begin{pmatrix} 12311 \\ 00012 \end{pmatrix}$ ,  $Q = \begin{pmatrix} 123112233 \\ 000121212 \end{pmatrix}$ .  $Q$ 中的列给出所有非零知识状态。采

用理想反应评分方式(1),由于本例中 $Q$ 矩阵中元素不只是0或者1,则仅仅取 $R_p$ 的某些列而不是所有的列,不能够使理想反应模式与知识状态一一对应;但是如果 $Q$ 依照“膨胀算法”导出相应的0-1 $Q$ 矩阵,并且掌握题目中1个属性则理想得分增加1分的评分方式,则按照文献[6-7],只需要可达阵中2列就可以使知识状态和理想反应模式一一对应。

每多掌握一个项目所测属性则理想得分增加1分的计分方式(简称为一属性一分方式)<sup>[1]</sup>与文献[12]提出的计分方式在布尔矩阵情况下是等价的,但是有人认为,对于多值 $Q$ 矩阵情形却不等价,而且认为,“一属性一分”的计分方式包含了文献[12]的计分方式,并且比文献[12]的计分方式更加细致,即提供的信息应该更丰富。这导致2种评分方式

下测验 $Q$ 矩阵设计的基本条件有所不同这样一个重要问题。根据计分方式提供的信息的多少,可以得出如下结论:

- 1) 若采用“一属性一分”的计分方式,设计的充要 $Q$ 阵或完备 $Q$ 阵的所要求的列数应该会更少;
- 2) 若按文献[12]的计分方式(1),要达到知识状态集合和理想反应模式集合之间的一一对应,所要求的列数应该更多。

如上所述,尽管都是多级评分,由于采用的评分标准不同,就可能导致不同的测验构造。当然,这里所说的认知诊断测验的构造,是指某种优良性准则下的构造,而不是拼凑一份测验。

文章中定理2的证明方法通过一点修改,可以用于证明下面的命题:属性之间不可以补偿并且采用0-1评分方式的认知诊断测验蓝图如果包含可达阵作为子矩阵,可以使期望反应模式和知识状态一一对应<sup>[3-5]</sup>。由此可以得到文献[5]的基本结论,但是这个证明和文献[5]的证明相比,简单得多。

## 4 参考文献

- [1] Tatsuo K K. Architecture of knowledge structures and cognitive diagnosis: a statistical pattern classification approach [C]. Nichols P D, Chipman S F, Brennan R L. Cognitively diagnostic assessments. Erlbaum: Hillsdale, 1995: 327-359.
- [2] Tatsuo K K. Cognitive assessment: an introduction to the rule space method [M]. New York: Taylor & Francis Group, 2009.
- [3] 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.
- [4] Chiu Chia-yi, Jeffrey A Douglas, Li Xiaodong. Cluster analysis for cognitive diagnosis: theory and applications [J]. Psychometrika, 2009, 74(4): 633-665.
- [5] 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): 490-495.
- [6] 丁树良,罗芬,汪文义. 多级评分认知诊断测验蓝图的设计—独立型和收敛型结构[J]. 江西师范大学学报: 自然科学版, 2014, 38(3): 265-269.
- [7] 丁树良,汪文义,罗芬. 多级评分认知诊断测验蓝图的设计—根树型结构[J]. 江西师范大学学报: 自然科学版, 2014, 38(2): 111-118.
- [8] 丁树良,罗芬,汪文义.  $Q$ 矩阵理论的扩展[J]. 心理学探新, 2012, 32(5): 410-422.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] Ding Shuliang ,Luo Fen ,Cai Yan ,et al. Complement to Tatsuoka's  $Q$  matrix theory [C]. Shigemasuk ,Okada A , Imaizumi T ,et al. New Trends in Psychometrics. Tokyo : Universal Academy Press 2008: 417-423.
- [14] 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.
- [15] 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.
- [16] 杨淑群,丁树良. 有效对象的判定理论与方法 [J]. 江西师范大学学报: 自然科学版 2011 35(1): 1-4
- [17] 丁树良,罗芬. 由偏序关系的可达阵导出 Hasse 图的有效算法—兼谈其在认知诊断中的作用 [J]. 江西师范大学学报: 自然科学版 2013 37(5): 441-444.
- [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.

## The Polytomous $Q$ - Matrix Theory

DING Shuliang ,WANG Wenyi ,LUO Fen ,XIONG Jianhua

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

**Abstract:** An algorithm to make the translation from the polytomous quasi-reachability matrix ( $R_p$ ) to a dichotomous reachability matrix ( $R_2$ ) has been given. An expansion algorithm of  $R_p$  and a method to compute polytomous ideal response patterns (PIRP) are also provided under the polytomous  $Q$ -matrix. Given certain item scoring rules proposed by Sun Jia'nan ,et al ,the statement that the  $R_p$  matrix can be used as the submatrix of the polytomous  $Q$ -matrix to guarantee bijective mapping from the set of PIRP to the set of the polytomous knowledge states has been proved. This statement has not been proved in the study of Sun Jia'nan.

**Key words:** polytomous  $Q$ -matrix;  $Q$ -matrix theory; expansion algorithm; test blueprint of diagnostic testing

( 责任编辑: 冉小晓)