認知診断における複雑度指標の開発

尾崎幸謙

データ科学研究系 助教

認知診断とは

$$\frac{1}{2} + \frac{1}{3} = \frac{3}{6} + \frac{2}{6} = \frac{5}{6}$$
約分の知識 分数の足し

属性(Attribute)=問題に正答するために必要となる知識のうちの1つ

算の知識

Q matrix の例(要素は q_{ik})

		-5
1	2	3
1	1	0
0	1	0
1	0	0
0	0	1
1	1	1
	0 1 0	$ \begin{array}{cccc} 0 & 1 \\ 1 & 0 \\ 0 & 0 \\ \end{array} $

受験者1は項目3のみに正答するだろう。 受験者2は項目1,2,3に正答するだろう。 knowledge state ベクトル (α) の例 (要素は α_{ik})

1	2	3
1	0	0
1	1	0
0	1	0
1	1	1
1	0	0
	1 1 1 0 1 1	1 2 1 0 1 1 0 1 1 1 1 0

認知診断の目的は、knowledge stateベクトルαを推定することである。これにより、受験者は「なぜ問題が解けないのか」が分かり、教育者の効率的な指導が可能になる。

DINAモデルとは

$$\eta_{ij} = \prod_{k=1}^{K} \alpha_{ik}^{q_{jk}}$$

項目iで必要とされているAttributeをすべて持っている場合に、受験者jのηij=1になる。1つでも持っていないならば、ηij=0になる。

$$P(Y_{ij} = 1 | \alpha_i) = (1 - s_j)^{\eta_{ij}} g_j^{1 - \eta_{ij}}$$

Q matrixとknowledge stateの例において, 受験者1と受験者2のηは,

$$\begin{array}{lll} \eta 11 = 1^{1*} 0^{1*} 0^{0} = 1 & \eta 21 = 1^{1*} 1^{1*} 0^{0} = 1 \\ \eta 12 = 1^{0*} 0^{1*} 0^{0} = 0 & \eta 22 = 1^{0*} 1^{1*} 0^{0} = 1 \\ \eta 13 = 1^{1*} 0^{0*} 0^{0} = 1 & \eta 23 = 1^{1*} 1^{0*} 0^{0} = 1 \\ \eta 14 = 1^{0*} 0^{0*} 0^{1} = 0 & \eta 24 = 1^{0*} 1^{0*} 0^{1} = 0 \\ \eta 15 = 1^{1*} 0^{1*} 0^{1} = 0 & \eta 25 = 1^{1*} 1^{1*} 0^{1} = 0 \end{array}$$

DINAモデル(αは受験者パラメタ, sとgは項目パラメタ)

αiは受験者iのknowledge stateベクトル

sjは項目jのslipパラメタ。項目jに正解するために必要なattributeをすべて持っているにも関わらず誤答する確率。

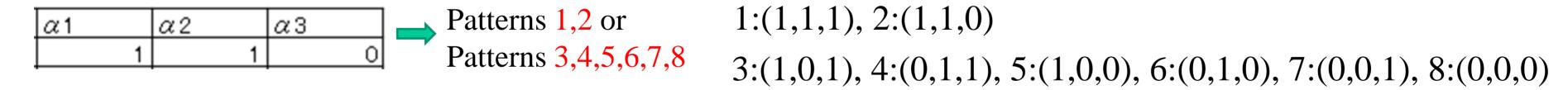
gjは項目jのguessingパラメタ。項目jに正解するために必要なattributeのうちのいくつかを持っていないにも関わらず正答する確率。 パラメタ推定には、周辺最尤推定法やMCMCが使われる。

多肢選択問題のための、多肢選択型DINAモデルもある(de la Torre, 2009; Ozaki, 2011)。

属性パターンの分類問題

各項目によって受検者集団は属性パターンによって構成される集合に分類される。

例えば, 属性数が3でα1とα2を必要とする項目がある場合には, 受検者集団は2つの集合のいずれかに分類される。



αの推定精度は何によって決まるのか?

項目数、属性数、選択肢数などが候補となるが、本研究ではそれらを統合的に捉えた「複雑度指標」を開発し、「複雑度指標」 の影響がより強いことをシミュレーション研究で確認する。

複雑度指標(CP Index)=B/A

A=(属性数の合計がKのとき)分類可能な属性パターンの総数。

B = (属性数の合計がKのとき)分析対象となっている項目によって分類可能な属性パターンの総数。

シミュレーションの概要・結果