

情報技術学習支援システムの開発と学習評価

—解答要因の推定—

瀬沼 航太郎[†]宮川 裕介[†]泉 隆[†]日本大学[†]

1.はじめに

インターネットを用いて学習を行う e-Learning システムが教育機関や企業の研修で多く利用されている。e-Learning システムを利用する学習者に対する利点として時間と場所の制約がないことや、自分のペースで学習が可能なが挙げられる。欠点として、IT スキルおよび環境がない学習者は受講が困難なことや、学習の頻度が学習者の学習意欲に依存することが挙げられる。特に学習意欲が低下することは、目標とする試験や達成度に大きな影響を与える。そのため学習意欲を向上するシステムが求められる。

本研究では、先行研究で開発した基本情報技術者試験対策 e-Learning システムを利用する。本システムで獲得できる解答情報をもとに学習評価のフィードバックをすることで学習意欲の向上をはかる。その方法として、各学習者が苦手な問題やケアレスミスしやすい問題などのカテゴリに解答要因を分類して、学習者のオリジナルな問題集を作成し、学習者の復習や試験直前対策の教材として提供することを考える。

解答要因をカテゴリに分類する際、パターン認識における分類手法では、 k 近傍法やニューラルネットワークなどが考えられるが、 k 近傍法は局所的な近似に過ぎず詳細な分類には精度が欠ける点、ニューラルネットワークは教師データから作成した特徴空間がブラックボックスであるため解析が困難である点が問題となる。そこで、特徴空間が明瞭であり詳細な分類が可能な、MT(Mahalanobis Taguchi)システムを用いて学習者の解答要因を推定する手法を検討している。本報告では、MT システムの中でも多くの単位空間を定義することが可能である RT(Recognition Taguchi)法について検討した。

2.学習評価

e-Learning システムにおいて、学習の場所や時間の制約が取り除かれる一方で、強制力が低く学習意欲の低下により学習が継続しないことが考えられる。この原因として、学習者は何を勉強すべきかが不明確であることや学習者自身の学習状況の把握が困難であるためと考えられる。また、基本情報技術者試験に採用されている、合計得点を評価対象とする素点方式では、学習者の弱点や各問題の傾向を把握することができない。それらの問題を解決するため、それぞれの問題の解答に要した時間(以下、解答時間)を加え、学習者の解答要因を、MT システムを用いて推定する手法を検討する。

2.1.MT システム^[1]

MT システムは、ある集団に対する多次元の情報を処理して、集団全体に 1 つの距離の指標を作成し、パターン認識をする手法である。応用分野には、医療診断や機器・設備の監視、売上推移の予測、文字・音声認識など多くの分野で利用されている。

MT システムには主に MT 法、MTA 法、TS 法、T 法の 4 つの手法がある。本報告では T 法の中でも複数の単位空間(均質空間)が存在する場合に適している、RT 法を用いた解答要因の推定を検討した。

2.2.RT 法

RT 法は特徴量が複数個あった場合でも、2 個の変数に情報圧縮をしてパターン認識を行う手法であり、出力値が数値で定義できない場合に適用することができる。また、2 個の変数に圧縮することによって、多くの単位空間を定義することが可能である。

2.2.1.RT 法の手順

本研究における RT 法について述べる。

2.2.2.単位空間と特徴量の決定

本手法では、単位空間を解答要因と考えられる“学習良好”、“ケアレスミス”、“無作為解答”、“不得意”、“学習不足”の 5 つとし、特徴量は正解数、正解率、確信度、解答時間、総解答時間の 5 個とする。

2.2.3.単位空間の作成

単位空間を構成するデータ(以下、メンバー) n 個及びそれぞれの特徴量 k 個(本報告の場合 5 個)について、メンバー番号 $i(=1\sim n)$ 、特徴量番号 $j(=1\sim k)$ 、とする。特徴量ごとの平均 m_i および平均の二乗和 r を求める。

次に単位空間のメンバーごとの 2 変数 Y_{i1} 、 Y_{i2} を次式で求める。 Y_{i1} は感度、 Y_{i2} は特徴量ごとの平均からのばらつきを表す。

$$L_i = \sum_{j=1}^k m_j x_{ij} \quad (i = 1, 2, \dots, n) \quad (1)$$

$$S_{Ti} = \sum_{j=1}^k x_{ij}^2 \quad (i = 1, 2, \dots, n) \quad (2)$$

$$S_{\beta i} = \frac{L_i^2}{r} \quad (i = 1, 2, \dots, n) \quad (3)$$

$$S_{ei} = S_{Ti} - S_{\beta i} \quad (i = 1, 2, \dots, n) \quad (4)$$

$$V_{ei} = \frac{S_{ei}}{k-1} \quad (i = 1, 2, \dots, n) \quad (5)$$

“Development of the Information Technology Learning Supporting System and Learning Evaluation - Presumption of Answer Factor -”

[†]Kotaro Senuma, [†]Yusuke Miyakawa, [†]Takashi Izumi

[†]Nihon University

$$Y_{i1} = \beta_i = \frac{L_i}{r} \quad (i = 1, 2, \dots, n) \quad (6)$$

$$Y_{i2} = \sqrt{V_{ei}} \quad (i = 1, 2, \dots, n) \quad (7)$$

Y_{i1}, Y_{i2} から分散と共分散を求め分散共分散行列 \mathbf{V} を求める.

$$V_{11} = \frac{1}{n} \sum_{i=1}^n (Y_{i1} - \bar{Y}_1)^2 \quad (i = 1, 2, \dots, n) \quad (8)$$

$$V_{12} = \frac{1}{n} \sum_{i=1}^n (Y_{i1} - \bar{Y}_1)(Y_{i2} - \bar{Y}_2) \quad (i = 1, 2, \dots, n) \quad (9)$$

$$V_{22} = \frac{1}{n} \sum_{i=1}^n (Y_{i2} - \bar{Y}_2)^2 \quad (i = 1, 2, \dots, n) \quad (10)$$

$$\mathbf{V} = \begin{pmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} \end{pmatrix} \quad (11)$$

そして、分散共分散行列の余因子行列 \mathbf{A} は次のように表される。これを単位空間とする。

$$\mathbf{A} = \begin{pmatrix} V_{11} & -V_{12} \\ -V_{12} & V_{22} \end{pmatrix} \quad (12)$$

この余因子行列 \mathbf{A} を用いて単位空間のメンバーに対してマハラノビスの距離 D を次式で求める。

$$D_i^2 = \frac{1}{2} [V_{22}(Y_{i1} - \bar{Y}_1)^2 - 2V_{12}(Y_{i1} - \bar{Y}_1)(Y_{i2} - \bar{Y}_2) + V_{11}(Y_{i2} - \bar{Y}_2)^2] \quad (13)$$

$$D_i = \sqrt{D_i^2} \quad (14)$$

ここで \bar{Y}_1, \bar{Y}_2 は Y_{i1}, Y_{i2} の平均である

この単位空間における D_i のばらつきの尺度として、 D_i のゼロ点からの分散 (D^2 の平均) を σ^2 ($\sigma > 0$) とおき、その標準偏差を σ とする。

2.3.4.単位空間のメンバーによるばらつきの評価

単位空間のばらつきの評価を行うために、他の単位空間のメンバーを用いて 2 変数 Y_{i1}, Y_{i2} を算出する。ここで (8), (9)式と同様に Y_{i1}, Y_{i2} を求め、(13)式の D_i^2 を求める。ただし計算の際、 $m_k, r, V_{11}, V_{12}, V_{22}$ の値は単空間作成時の値を用いる。求めた D_i^2 の平均 \bar{D} から分類力が評価できる。

2.3.5.未知データの分類

2.3.4 で求めた D と同様に未知データに対して算出する。ここで、しきい値を単位空間における D_i のばらつきの尺度の 2 倍、すなわち 2σ 以内であれば単位空間内の正分類とし、超えていれば単位空間外であり、誤分類とする。

3.評価実験

RT法による解答要因の推定について検討するために、単位空間によるばらつきの評価と分類実験を行った。実験概要を表 1 に示す。データは学習者が問題解答後のアンケートに基づく、解答要因が既知ものを用いた。

表 1. 評価実験概要

実施期間	2013 年 4 月 13 日～8 月 27 日
対象者数	25 名
全データ数	1469 問
サンプルデータ数	949 問
未知データ数	520 問
単位空間数	5

3.1.単位空間によるばらつきの評価結果

図 1 に横軸 Y_{i1} 、縦軸 Y_{i2} とした単位空間と信号データの散布図を示す。

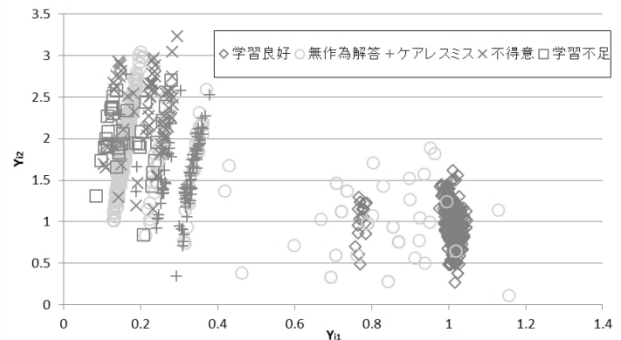


図 1. 単位空間と信号データの散布図

図 1 において“学習良好”は二つに分かれ、それらはまとまりがあることから、“学習良好”は、その他の単位空間と区別できることがわかる。また“学習良好”以外の単位空間は重なりが多く分類が困難であることがわかる。しかし、“無作為解答”と“ケアレスミス”はある程度まとまりがあり、分類可能と考える。

3.2.分類実験結果

分類結果を表 2 に示す。

表 2. 分類結果

	分類数	正分類数	分類率[%]
学習良好	300	267	89
無作為解答	100	67	67
ケアレスミス	70	48	69
不得意	30	4	13
学習不足	20	11	55

全体の正分類率は 76[%]となった。特に“学習良好”は 89[%]と高分類率が得られた。誤分類の主な原因としては、“学習良好”の単位空間は確信度が強く、確信度の値が小さいことによる。なお、“不得意”は 13[%]、“学習不足”は 55[%]と低い値であった。すなわち、本手法では“不得意”、“学習不足”の分類は困難であると考えられる。

4.まとめ・今後の課題

本報告では RT 法による解答要因の推定について検討した。単位空間と信号データのばらつきによる評価では“学習良好”は分類可能であり、分類実験全体の正分類率は 76[%]であった。

今後は、“学習良好”、“無作為解答”、“ケアレスミス”の分類率の更なる向上を図る。

参考文献

[1]立林和夫, 手島昌一, 長谷川良子:「入門 MT システム」, 日科技連出版社(2008-12)