

情報技術学習支援システムの開発と学習評価

- MT 法による解答要因の推定 -

Development of the Information Technology Learning Supporting System and Learning Evaluation

- Presumption of Answer Factor by the MT Method -

瀬沼 航太郎[†]宮川 裕介[†]泉 隆[†]

Kotaro Senuma

Yusuke Miyakawa

Takashi Izumi

1.はじめに

インターネットの普及により、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、視認性や利便性を考慮しつつ、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^[1]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。そこで、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識別し、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識別手法では、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識別には精度が欠ける点や、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、本報告では特徴空間が明瞭であり詳細な分類が可能な、マハラノビス・タグチ(MT)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た。

2.システム概要

本システムは、基本情報技術者試験を対象としている。システムにログイン(ユーザ認証)後、モード選択によってユーザは必要に応じた学習を行う。各モードの問題形式は試験の午前問題と同様の四肢択一問題である。

3.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、個人により傾向が異なる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。

3.1.MT システム^[2]

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。パターン認識の応用分野には、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。本研究では、特徴量として様々な解答情報を扱うため、全特徴量の相関を考慮して識別できるマハラノビス距離が適しているため MT 法を用いる。

3.2.MT 法

MT(マハラノビス・タグチ)法は目的に対して均質な集団を単位空間(特徴空間)として定義し、未知データの単位空間の中心からの距離をマハラノビス距離(MD)として求める。図 1 のように MD が小さければ単位空間に近いパターン、MD が大きければ単位空間から遠いパターンであると判断することができる。

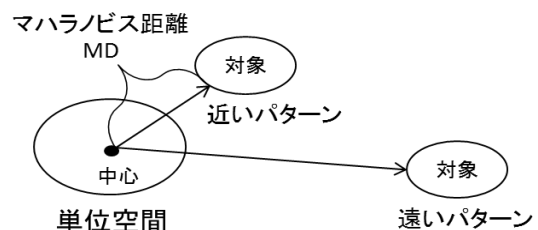


図 1. MT 法のご概念

3.3.MT 法の手順

MT 法の手順と本研究の使用法について述べる。

- ① 目的の明確化
学習者の解答要因を推定することを目的とする。
- ② 基準となる状態の定義
本報告では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”、“不得意問題”、“学習不足”の五つを基準となる状態と定義する。
- ③ 特徴量の決定
本システムで取得できる特徴量は、学習者の正

[†] 日本大学

誤、解答時間、確信度、不正解時の解答、テスト(20問)の正解数、総解答時間、問題の正解率、分野などが挙げられる。本報告では、学習者の解答時間、確信度、正解数、総解答時間、解答時間割合、問題の正解率、分野を用いる。

④ 単位空間の作成

単位空間とは、多数の情報(変数)を系統化、組織化、知識として保有する関数である。MT法においては、相関行列の逆行列が単位空間に相当する。すなわち、②で定義した基準状態ごとの単位空間を作成する。以下に単位空間の作成方法を示す。

単位空間の特徴量の変数 x 、その変数の数 k 、サンプル数 n に対し平均値 $\bar{x}$ と標準偏差 σ を求め、これに対してデータの規準化する。

$$X_{ij} = (x_{ij} - \bar{x}_{ij}) / \sigma_{ij} \quad (1)$$

$$(i=1, 2, \dots, n; j=1, 2, \dots, k)$$

k 種類の変数について、相関行列 R を計算する。相関行列は $k \times k$ の正方行列であり、成分 r_{ij} は変数 i と変数 j の相関係数である。

$$R = \begin{pmatrix} 1 & \cdots & r_{1k} \\ \vdots & \ddots & \vdots \\ r_{k1} & \cdots & 1 \end{pmatrix} \quad (2)$$

ここで、 r_{ij} は以下である。

$$r_{ij} = \left(\sum_{l=1}^n X_{il} X_{jl} \right) / n \quad (3)$$

次に、相関行列 R の逆行列 R^{-1} を計算し、これを行列 A とする。

$$A = R^{-1} = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} \quad (4)$$

⑤ 識別処理

識別処理とは、対象から得た未知データを単位空間に与え、その結果を MD のとして求める処理である。

未知データの変数、

$$y = (y_1, y_2, \dots, y_k) \quad (5)$$

が与えられたとき、この変数の MD ($=D^2$) は、 y を式(1)により規準化する。これを、

$$Y = (Y_1, Y_2, \dots, Y_k) \quad (6)$$

として、以下の式より MD を求める。

$$D^2 = YAY^T / k$$

$$= \sum_{i=1}^k \sum_{j=1}^k a_{ij} Y_i Y_j / k \quad (7)$$

ここで Y^T は行列 Y の転置行列である。

未知のデータをすべての単位空間との MD を求め、MD の値が最小である単位空間がその解答要因とする。

4. 評価実験

本手法を検証するために、学習評価実験を行った。実験概要を表1に示す。

表1. 評価実験概要

実施期間	2013年4月13日～4月21日
対象者数	15名
全データ数	703問
サンプルデータ数	455問
未知データ数	248問
単位空間数	5

予め学習者からアンケートにより取得した解答要因既知のデータ 703 問を用いる。単位空間の作成データに対し解答要因ごとに単位空間をサンプルデータ 455 問から作成し、未知データ 248 問の識別率を調べる。表2に解答要因ごとの識別率の実験結果を示す。ここで、正識別とは、識別結果が正しい解答要因に識別できものであり、誤識別とは、識別結果が異なる解答要因に識別されたものである。

表2. 解答要因ごとの識別率

解答要因	識別数	正識別	誤識別	識別率
	[問]	[問]	[問]	[%]
学習良好	134	134	0	100
無作為解答	42	27	15	64
ケアレスミス	47	23	24	49
不得意	22	12	10	55
学習不足	3	1	2	33
計	248	197	51	79

表2より解答要因が“学習良好”では100%の高識別率を得ることができた。“無作為解答”、“ケアレスミス”、“不得意”では、65～50%程度であった。また、“学習不足”は未知データに含まれるデータ数が極端に少ないため検証することができなかった。全体では79%と良好な識別率を得ることができた。

誤識別として、“無作為解答”を“ケアレスミス”として識別したものが17問と最も多く、“ケアレスミス”を“無作為解答”としたものが8問見られた。これは、“無作為解答”と“ケアレスミス”のMDの値が近く、特徴量の解答傾向が似ていたためと考えられる。

“学習良好”と他の解答要因の識別結果が大きく異なった原因として、単位空間を作成する際に用いたサンプルデータ数455問に対して“学習良好”が246問、他の解答要因では80～50問程度しかなくデータ数不足が考えられる。このため、異常なサンプルデータの解答を抑制することができず、単位空間が不十分であったと考えられる。このことよりサンプルデータを増やす必要がある。

5. まとめ

本報告ではMT法を用いて、学習者の解答要因の推定について検討した。5つのカテゴリに対する識別率は79%であった。特に“学習良好”では100%の高識別結果が得られた。

今後は、MT法の単位空間の作成時に用いるサンプルデータを増加して、単位空間の更新を行う。また単位空間を作成する際の特徴量として、学習者の分野における得意・不得意の程度を特徴量として扱えないか検討を行う。

参考文献

- [1]久津間啓右, 金子勇太, 泉隆:「インターネットを利用した情報技術学習支援システムー学習状況の評価に用いる問題に関する検討ー」, FIT2011, N-018(2011-09)
- [2]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)