

# 項目反応理論を用いた受験者およびテストの評価

熊本大学大学院自然科学教育部博士前期課程  
機械数理工学専攻数理工学教育プログラム  
210D8551 西郷 虎太郎

# 目次

- ① はじめに
- ② 準備
- ③ 能力値や項目パラメータの推定
  - 最尤推定法
  - ベイズ推定法
    - EAP 法 (Expected a posteriori 法)
- ④ 実際の推定
- ⑤ ロジスティックモデルの多次元化
  - 多次元ロジスティックモデル
  - シミュレーション

# 項目反応理論

## 反応行列 $U$

受験者  $i$  ( $i = 1, 2, 3, \dots, N$ ) の項目  $j$  ( $j = 1, 2, 3, \dots, n$ ) に対する反応を正答を 1、誤答を 0 として  $u_{ij}$  として表す。

- $u_{ij} = 1$  (正解のとき)
- $u_{ij} = 0$  (不正解のとき)

この  $u_{ij}$  を  $(i, j)$  成分にもつ  $N \times n$  行列  $U$  を反応行列  $U$  という。また、反応行列  $U$  の第  $i$  行目を抜き出したベクトル  $\mathbf{u}_i$  を受験者  $i$  の全項目に対する反応として扱う。

## 能力値 $\theta$

受験者  $i$  の能力値を  $\theta_i$  で表す。

## 識別力 $a$

推定したい受験者の能力値  $\theta$  をその項目でどのくらい正確に推定できているのかを表したパラメータである。以下より、項目  $j$  の識別力を  $a_j$  で表す。

## 困難度 $b$

項目  $j$  の難しさを表すパラメータであり、 $b_j$  で表現する。2 に近いほど能力値  $\theta$  が高くないと正答できない難しい項目とされる。

## 偶然率 $c$

能力値  $\theta$  に関係なく偶然正答になる確率を表すパラメータである。一般的には、当て推量と呼ばれることもあるが以下では、偶然率という言葉を用いている。項目  $j$  の偶然率を  $c_j$  で表す。

## 項目特性関数

能力値  $\theta_i$  の受験者  $i$  が項目  $j$  に正解する確率は、受験者の能力値  $\theta_i$  と項目パラメータに依存するため、正解率を  $p_j(\theta_i)$  で表す。 $p_j(\theta_i)$  は  $0 \leq p_j(\theta_i) \leq 1$  を満たす  $\theta$  の増加関数であり、その関数形により項目  $j$  の能力値  $\theta_i$  に対する特徴をモデル化することが必要となる。

## 正規累積モデル

標準正規分布の累積分布関数を利用した  $\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2}) dz$  を利用する。横軸に能力値  $\theta$ 、縦軸に正答率を配した関数  $\Phi(f(\theta))$  を正規累積モデルといい、以下のように定義する。

$$\Phi(f(\theta)) = \int_{-\infty}^{f(\theta)} \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2}) dz \quad (1)$$

## 項目特性関数

$\theta$  の関数  $f(\theta)$  を

$$f(\theta) = a(\theta - b) \quad (2)$$

として、項目  $j$  の項目特性関数を

$$p_j(\theta) = \Phi(a_j(\theta - b_j)) \quad (3)$$

と表すことができる。

## ロジスティックモデル

正規累積モデルでは、式中に積分を含んでいるので実際の推定の際には計算が複雑になってしまう。そこでロジスティック分布を利用した項目特性関数が提案された。提案されたモデルは以下のようなものである。

$$\int_{-\infty}^{f(\theta)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \simeq \frac{1}{1 + \exp(-D \times f(\theta))} \quad (4)$$

この式の右辺をロジスティックモデルという。一般に、 $D = 1.7$  のとき、左辺と右辺の誤差が 0.01 以下になることが知られているので今後は  $D = 1.7$  として扱う。 $f(\theta)$  に使われる式のパラメータの数によって 1 パラメータロジスティックモデルから 4 パラメータロジスティックモデルまで考えられる。

## 1 パラメータロジスティックモデル

$$p_j(\theta) = \frac{1}{1 + \exp(-Da(\theta - b_j))} \quad (5)$$

これを1パラメータロジスティックモデルという。

- $a$  は全ての項目に共通している値であるため、項目特性関数は項目同士で交わることはない。
- 困難度が高くなるほど、グラフは右に移動していく。



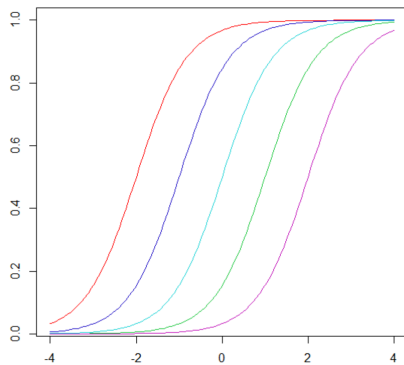


Figure:  $a$  は 1 で固定、 $b$  は  $-2, -1, 0, 1, 2$

## 2 母数ロジスティックモデル

$$p_j(\theta) = \frac{1}{1 + \exp(-Da_j(\theta - b_j))} \quad (6)$$

- 1パラメータロジスティックモデルのときとは違い、項目特性関数は項目同士で交わることもある。
- 下図のように、 $a$ の値が大きくなるほどグラフの傾きが大きくなる。つまり、識別力 $a$ の値が大きくなると、能力値 $\theta$ の値の少しの変化でも、大きく正答確率が変わることになる。このことから、パラメータ $a$ は識別力を表すパラメータを表していることが分かる。

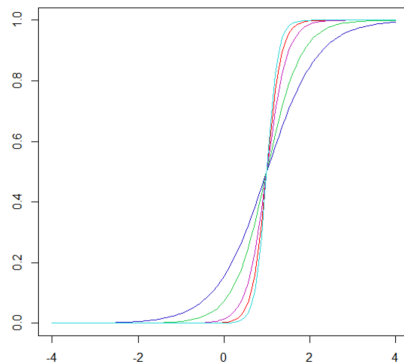


Figure:  $b$  は 1 で固定、 $a$  は  $-4, -2, 0, 2, 4$

### 3 母数ロジスティックモデル

$$p_j(\theta) = c_j + \frac{1 - c_j}{1 + \exp(-Da_j(\theta - b_j))} \quad (7)$$

- 2 母数ロジスティックモデルに加えて、偶然率  $c_j$  が追加されたモデル
- 下方漸近線といわれる
- 項目特性関数の各項目の正答率は  $c_j$  から 1 の間で推移する。
- $c$  の値が大きくなるほど、グラフの下限が上がっていく。

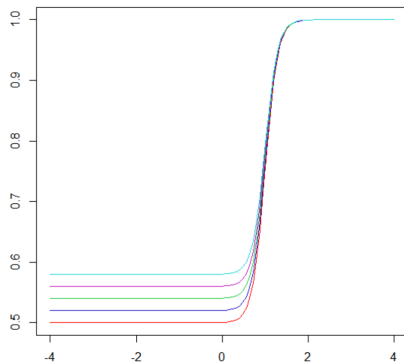


Figure:  $a$  は 4 で固定、 $b$  は 1 で固定、 $c$  は 0.58, 0.56, 0.54, 0.52, 0.5

## 最尤推定法

項目パラメータが既知の場合に、項目パラメータと反応行列から能力値  $\theta$  の推定値を求める方法を紹介する。

受験者  $i$  の能力値  $\theta_i$  は現実においては未知である。一方で、反応行列は収集して実際の値として固定し用いることができる。そこで、反応行列を既知として、能力値  $\theta_i$  を変数として扱うことにする。 $L(\theta_i|\mathbf{u}_i)$  を次のように定義し、

$$L(\theta_i|\mathbf{u}_i) = \prod_{j=1}^n p_j(\theta_i)^{u_{ij}} (1 - p_j(\theta_i))^{1-u_{ij}} \quad (8)$$

この  $L(\theta_i|\mathbf{u}_i)$  を尤度関数という。このとき、尤度が最大になるときの項目パラメータの値を、パラメータの推定値として利用する。このような推定法を最尤推定法という。つまり、この推定法は実際のデータが最も得られやすいようなパラメータの値を決める方法である。

## EAP 法 (Expected a posteriori 法)

これは、 $\theta_i$  の事後分布の密度関数の  $\theta_i$  に関する平均値を推定値とする推定法である。これは、事後分布の平均が 2 乗損失を最小にするという事実に基づいた推定法である。式にすると、

$$E[\theta_i | \mathbf{u}_i] = \int_{-\infty}^{+\infty} \theta_i \frac{p(\mathbf{u}_i | \theta_i) \pi(\theta_i)}{\int_{-\infty}^{+\infty} p(\mathbf{u}_i | \theta_i) \pi(\theta_i) d\theta_i} d\theta_i = \frac{\int_{-\infty}^{+\infty} \theta_i \pi(\theta_i) p(\mathbf{u}_i | \theta_i) d\theta_i}{\int_{-\infty}^{+\infty} \pi(\theta_i) p(\mathbf{u}_i | \theta_i) d\theta_i} \quad (9)$$

となる。ただし、 $\pi(\theta)$  は  $\theta$  の事前分布の密度関数である。式中の積分は区分別求積法によって求めることができる。

## EAP 法 (Expected a posteriori 法)

能力値  $\theta$  だけでなく、項目パラメータに用いることで推定値を求めることができる。そのためには、項目パラメータの事前分布を仮定しておく必要がある。

**能力値  $\theta$  の事前分布** 能力値  $\theta$  の事前分布には正規分布が仮定されることが多い。正規分布のパラメータは平均  $\mu_\theta$  と分散  $\sigma_\theta^2$  である。

**識別力  $a$  の事前分布** 識別力の推定値は数学的には負の値もとる得るが、識別力的事前分布には正の範囲で定義される対数正規分布が仮定されることが多い。

**困難度  $b$  の事前分布** 困難度的事前分布にも正規分布が仮定されることが多い。

**偶然率  $c$  の事前分布** 偶然率の推定値は確率を表現するパラメータである。このため事前分布はベータ分布が仮定されることが多い。ベータ分布のパラメータは  $\alpha, \beta$  であり、密度関数は

$$\frac{x^{\alpha-1}(1-x)^{\beta-1}\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \quad (10)$$



## 実際の推定

- 30 人の教育部学生に対して教育教養の問題を 8 問出題する。

## 反応行列

集計して反応行列  $U$  は以下ようになった。

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}^t$$

## 実際の推定

2パラメータロジスティックモデルと3パラメータロジスティックモデルを用いて推定する。それぞれのパラメータの推定値は以下のようになった。

	困難度	識別力
問題 1	0.329	0.369
問題 2	0.767	0.843
問題 3	1.091	0.705
問題 4	1.019	0.766
問題 5	0.197	3.083
問題 6	0.958	0.534
問題 7	0.259	0.431
問題 8	-0.18	3.212

	困難度	識別力	偶然率
問題 1	1.414	1.622	0.350
問題 2	0.958	1.674	0.140
問題 3	1.275	1.787	0.137
問題 4	1.274	1.785	0.136
問題 5	0.026	2.970	0.038
問題 6	1.334	1.779	0.200
問題 7	1.414	1.623	0.350
問題 8	-0.487	3.175	0.044

Table: 2パラメータロジスティックモデル Table: 3パラメータロジスティックモデル

## 考察

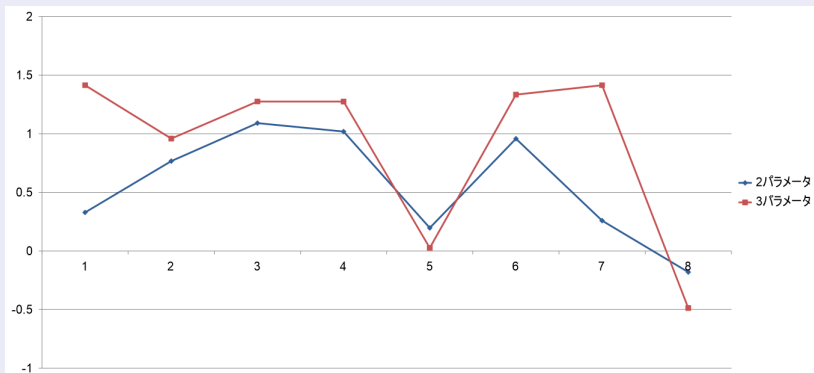


Figure: 困難度の推定値

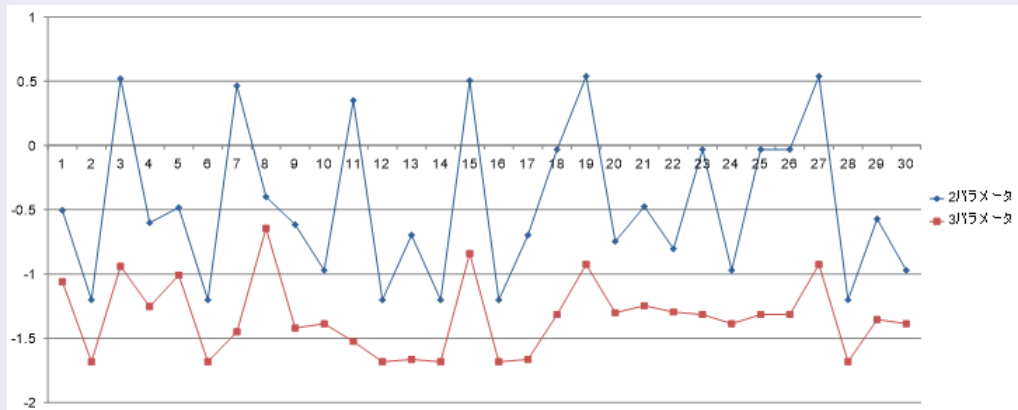


Figure: 能力値  $\theta$  の推定値

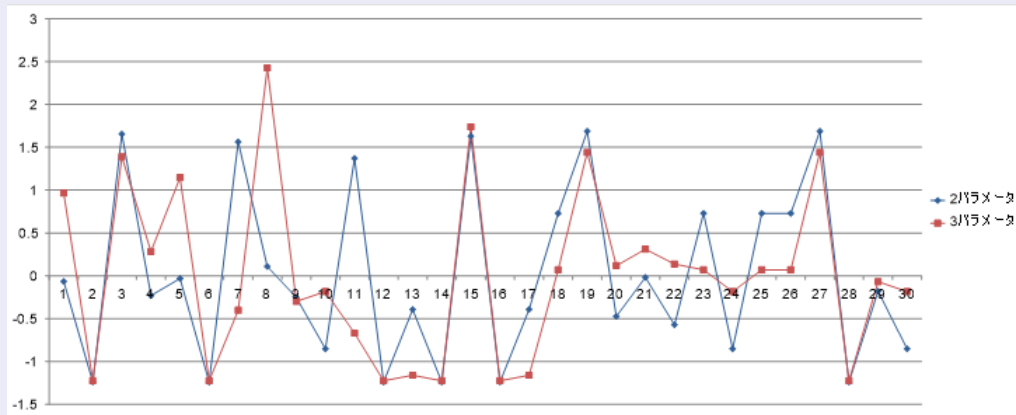


Figure: 能力値  $\theta$  の推定値 (標準化 ver)

## 最尤推定法との比較

最尤推定法も用いた推定値との比較もしていく。まずは、最尤推定法を用いた推定値とベイズ推定法の値(2パラメータ)を表にまとめたものと能力値 $\theta$ を標準化しグラフに表したものを以下に示す。

	困難度	識別力
問題 1	0.329	0.369
問題 2	0.767	0.843
問題 3	1.091	0.705
問題 4	1.019	0.766
問題 5	0.197	3.083
問題 6	0.958	0.534
問題 7	0.259	0.431
問題 8	-0.18	3.212

Table: ベイズ推定法

	困難度	識別力
問題 1	-6.87	0.640
問題 2	-2.87	0.841
問題 3	-0.87	3.999
問題 4	0.461	2.705
問題 5	1.461	1.091
問題 6	2.261	0.671
問題 7	2.928	0.266
問題 8	3.499	0.194

Table: 最尤推定法

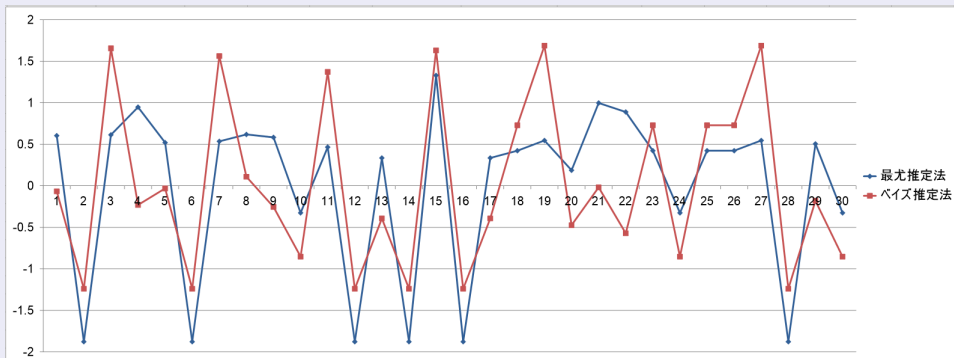


Figure: 能力値  $\theta$  の推定値 (標準化 ver)

## 多次元ロジスティックモデル

$$p_j(\theta) = \frac{1}{1 + \exp(-\mathbf{a}_j^t \boldsymbol{\theta}_i + b_j)} \quad (14)$$

ここで  $\mathbf{a}$  は項目ごとの識別力を表すベクトルである。また、 $\boldsymbol{\theta}$  も受験者ごとの多次元の能力値を並べたベクトルである。つまり式中の  $\mathbf{a}_j^t \boldsymbol{\theta}_i + b_j$  は次のように表現することができる。

$$\mathbf{a}_j^t \boldsymbol{\theta}_i + b_j = \sum_{l=1}^L a_{jl} \theta_{il} + b_j \quad (15)$$

さらに、能力が2次元なものであるとし  $\boldsymbol{\theta} = (\theta_1, \theta_2)^t$  のように項目が受験者の違う2つ能力値を測定していると仮定する。



## 能力値や項目パラメータの推定

- 識別力が高いものを選ぶべきなのか
- 困難度が高いものを選ぶべきなのか
- 過去問題は全体の中でどのくらいの割合で入れるべきなのか

## 計算の流れ

100 人の受験者に 100 問の問題を解いてもらう。そのうち、一定の割合で項目パラメータが既知である問題を入れる。今回は過去問題が 25 問、50 問、75 問の 3 パターンで計算していく。計算の流れとしては、

- 1 過去問題と反応行列から新しい問題の項目パラメータを推定
- 2 項目パラメータが推定された新しい問題を使って受験者の能力値を推定

の 2 段階に分ける。ここで以下の 2 点に注意されたい。

- 過去問題のパラメータは既知扱いであるので、推定の対象外である。
- 受験者の能力値の推定には過去問題は使用せず、新しい問題だけを使用する。

## パラメータの設定

- $\mathbf{a}_1 = (0.5, 0.53061, 0.56122, \dots, 2, 0.5, 0.53061, 0.56122, \dots, 2)$

$$\mathbf{a}_2 = (0.5, 0.53061, 0.56122, \dots, 1, 0.5, 0.53061, 0.56122, \dots, 1)$$

- $\mathbf{b} = (-1, -0.9797, -0.9595, \dots, 1)$

- $\boldsymbol{\theta}_1 = (-2, -1.9595, -1.9191, \dots, 2)$

$$\boldsymbol{\theta}_2 = (-1, -1, -1, \dots, -1, 1, 1, \dots, 1)$$

## 結果

以下に示すのは、能力値  $\theta_1$ 、 $\theta_2$  の推定値と設定値との誤差の平均を、過去問題の割合毎にまとめたものである。

	過去問題 0 問	過去問題 25 問	過去問題 50 問	過去問題 75 問
$\theta_1$	0.397705342	0.408443905	0.416983791	0.520276119
$\theta_2$	0.488909176	0.459919565	0.5123526	0.51364787