Part B (IRT part)

In previous part, we made strong assumption: we assumed that the guessing is a part of the ability of each candidate and all questions have equivalent discriminations, which means that one parameter is enough to describe a certain question. In other words, though each individual question might have secondary factors, these features are assumed to be mutually independent and collectively orthogonal, which is also called one-parameter model (1PL). Similar to the Naïve Bayes model, we know that too strong assumptions might not match the real situation and over simplified fitted models may not have that good performance in prediction accuracy.

Based on the issues mentioned above, a more general model with more parameters should be fitted. Before that, we can take a look back at the ICC (item characteristic curve), which is the base of IRT that describes the properties of a certain item. We can tell that these features that have influence on the shape of ICC include:

* question difficulty :the point at which candidates with the ability level or even higher level that scores the top fifty percentile on this question
* discrimination power : the ability of this question that can separate all candidates accurately, represents how well it differentiates good and bad performed candidates
* pseudo-guessing parameter : the probability that low-skilled people get answer correctly
* upper asymptote : the probability that skilled students can answer correctly, which is rarely used and usually close to value 1

Based on this, we rewrite the general statistic model as following:

Let denotes the ( entry of a 2-dimential matrix with size . if student answers question correctly and equals to 0 otherwise.

: represents parameters of question , including ()

: discrimination power of question

: question difficulty of question

: pseudo-guessing parameter of question

Interpretation of parameters:

: discrimination power, if for question and question , , and then question discriminates more than question

: student’s ability necessary for the probability of solving the question equals ; the higher value of is the more difficult level the question is

: the probability of random correct answer

To estimate question parameters, marginal maximum likelihood method is applied. Before conducting MLE, two hypotheses are set up:

1. every student’s answers are independent from each other
2. different questions are solved independently by students (with different ability)

Assume there are number of different response patterns among all candidates, saying where . The probability density function of a certain response pattern is decided both by the property of question and the ability . Further assume the ability variable follows a known distribution with parameter , saying. In addition, represents the number of occurrences of response pattern among all the candidates. Thus, we have:

is a vector where

each has three parameters , to simplify the expression, we rewrite it as

Apply independence assumption,

: conditional distribution of given

Question parameters can be estimated individually based on the hypothesis.

Then, the ability parameter is estimated using MLE too. We have:

Thus by applying Newton-Raphson algorithm, candidates’ abilities are easily estimated.

References:

Frank B. Baker, Seock-Ho Kim (2004), ‘Item Response Theory: Parameter Estimation Techniques’.

Bock, R. D. & Aitkin, M. (1981), ‘Marginal maximun likelihood estimation of item parameters: An application of a EM algorithm’.

Fisher, G. H. (1995), ‘Some neglected problems in IRT’.

Link for resources:

<https://github.com/pluralsight/irt_parameter_estimation> (check the doc and code zlc\_mle.py which is modified and can ignore the abc\_mle and baker\_mle. The zlc\_mle.py code include all three different IRT models, the one used in part A is 1PL, and the mle estimate mentioned above is 3PL)

<https://github.com/aimir/irt/blob/master/irt/irt.py> (another sample code to estimate theta – student skills parameter)

<https://www.emis.de/journals/RCE/V27/V27_1_27Cepeda.pdf> (formula shown above and symbols are consistent with this article)