МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2017
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2017
Proceedings of the Forty-sixth Spring Conference
of the Union of Bulgarian Mathematicians
Borovets, April 9–13, 2017

NOTES ON THE PARAMETER ESTIMATION OF SOME IRT MODELS BY MEANS OF THE EM-ALGORITHM*

Dimiter P. Tsvetkov, Lyubomir Ya. Hristov, Ralitsa L. Angelova-Slavova

In this paper we consider an implementation of the EM-algorithm to the parameter estimation of some IRT models. The scheme is described in major important details. We point out an online author's realization of the above approach for various particular IRT models, including the binary logistic models with one, two and three parameters (1PL, 2PL, 3PL), the nominal response model (NRM), the partial credit model (PCM) and the generalized partial credit model (GPCM).

Introduction. The statistical approach to the social and behavioral sciences measurement requires observation of statistical units – individuals from some normal population. The Item Response Theory (IRT) is one of the common methods to these measurements (see for example [1, 5, 6]). The main purpose is to give a quantitative estimate of a latent scalar individual feature θ usually called "ability". We call a *result indicator* the observable individual outcome v of the measurement. A key step here is the definition of a relationship between the latent trait θ and the result indicator v by means of a certain probabilistic model.

The term IRT is somewhat metaphorical and difficult for translation in Bulgarian – for the time being "theory of a response to the test unit" is preferred (see for example [16]). The usual test is composed of a finite number I of test items (test units). The outcome from a test unit with number i is written by one of the markers $0, 1, \dots, M_i$. The natural numerical order between markers of the most models have an objective sense.

The probabilistic model for test unit i requires a definition of probabilities $P_{im}(\theta)$ (Item Category Response Function) for the result with marker m at individual of ability θ, according to the normalizing condition

$$P_{i0}(\theta) + P_{i1}(\theta) + \cdots + P_{iM_i}(\theta) = 1$$

The probabilities $P_{im}(\theta)$ usually depend on certain parametric set $\xi_i = (\xi_{i1}, \xi_{i2}, \dots)$. We will denote by $\eta = (\xi_1, \xi_2, \dots, \xi_I)$ the set of all the items parameters. Also we will assume a local stochastic independence condition, which means that the outcomes from different items are independent.

The result indicator $v \in \mathbb{V}$ admits convenient form as a result jagged array $v = (v_{im})$, $1 \le i \le I$, $0 \le m \le M_i$, where $v_{im} = 1$, if for the test unit i marker m is pointed,

*2010 Mathematics Subject Classification: 97K80, 62P25, 62P15.
Key words: Item Response Theory, EM-algorithm.

and $v_{im} = 0$ otherwise. Another convenient notation is the result vector $u = (u_i)$, in which the coordinate u_i shows the marker for the test unit i. This study systematically uses the suggestion about the presence of convenient probability distribution for θ with a density $\varphi(\theta)$. The parameters of $\varphi(\theta)$ are adopted as a part of the overall model ones. The distribution of θ is assumed to be normal $N(\mu, \sigma^2)$ (see for example [1, 3, 4, 5, 6]) in accordance with the common idea for a normal population as a set of prototype variants. Therefore, the unobserved values of θ appear as realizations of a random variable $\Theta \sim N(\mu, \sigma^2)$.

From the assumptions it follows that the probability distribution of the result indicator v is determined by

$$f(v|\theta, \eta) = \prod_{i=1}^{I} \prod_{m=0}^{M_i} (P_{im}(\theta, \xi_i))^{v_{im}} \tag{1}$$

defining a sample space $\mathbb{V}$ with a normalizing condition

$$\sum_{v \in \mathbb{V}} f(v|\theta, \eta) = 1.$$

Formula (1) describes a $\mathbb{V}$-valued discrete random variable V with a distribution $f(v|\theta, \eta)$. For both random variables V and Θ we have a joint distribution

$$f(v, \theta|\eta) = f(v|\theta, \eta)\,\varphi(\theta|\eta) = \left(\prod_{i=1}^{I} \prod_{m=0}^{M_i} (P_{im}(\theta, \xi_i))^{v_{im}}\right) \varphi(\theta|\eta). \tag{2}$$

The probabilities of the principal events $((v \in \mathcal{V})(\theta \in \Delta))$, where $\mathcal{V} \subseteq \mathbb{V}$ and $\Delta \subseteq \mathbb{R}$ is an interval, are

$$\Pr((v \in \mathcal{V})(\theta \in \Delta)) = \sum_{v \in \mathcal{V}} \int_{\Delta} f(v|\theta, \eta)\,\phi(\theta|\eta)\,d\theta.$$

The integrals with respect to θ are always from $-\infty$ to ∞. Hereafter the notation η will stand for all parameters of the model – both for the test units and for the distribution of Θ.

We call the joint distribution of Θ and V, given by (2), *basic IRT probability model* (see for example [17, 18]), because it gives the probability association between the main measurement purpose θ and its observed result indicator v.

Thus, for the marginal distribution of V we get

$$f(v|\eta) = \int f(v, \theta|\eta)\,d\theta = \int \left(\prod_{i=1}^{I} \prod_{m=0}^{M_i} (P_{im}(\theta, \xi_i))^{v_{im}}\right) \phi(\theta|\eta)\,d\theta. \tag{3}$$

Getting the computer generated data about a given probability model is always an important task related to the study of its properties. For the model (2) this is realized by the following iterative scheme.

Algorithm for generating data about the basic IRT model

1	Choose values for η and the sample size J
2	$\theta_{(j)} \sim \varphi(\theta|\eta)$
3	$v_{(j)} \sim f(v|\theta_{(j)}, \eta)$
4	Repeat steps $2 \rightarrow 3$ until the sample size is attained

Here the generated observations $\left(\theta_{(j)}, v_{(j)}\right)$, $1 \leq j \leq J$, are the complete model data, and $\left(v_{(j)}\right)$ represent an i.i.d. sample of result indicators. The algorithm above makes more clear the role of the latent variable Θ.

Let an i.i.d. sample of result indicators $v = \left(v_{(j)}\right)$ be given. Then according to (3) for the likelihood function we have

$$L(v|\eta) = \prod_{j=1}^{J} f\left(v_{(j)}|\eta\right)$$

therefore the log-likelihood function $l(v|\eta) = \ln L(v|\eta)$ has the form

$$l(v|\eta) = \sum_{j=1}^{J} \ln \int \left(\prod_{i=1}^{I} \prod_{m=0}^{M_i} \left(P_{im}(\theta, \xi_i)\right)^{v_{(j)im}} \right) \varphi(\theta|\eta)\, d\theta \tag{4}$$

The presentation (4) allows to find a maximum likelihood estimation (MLE)

$$\hat{\eta} = \arg\max_{\eta} l(v|\eta)$$

based on the known general EM-algorithm scheme.

EM-algorithm. The EM-algorithm (see for example [2, 8, 11]) is a very common method for obtaining MLE under an existence of latent variables. An important benchmark for the present study is the monograph of Baker and Kim [1].

In this section we give a description of the EM-algorithm in a terminology oriented to the basic IRT model (2). Consider some (arbitrary) manifest random variable v and some (arbitrary) latent random variable θ with a joint distribution

$$f(v, \theta|\eta) = f(v|\theta, \eta)\, f(\theta|\eta)$$

obtained from the known density shapes $f(v|\theta, \eta)$ and $f(\theta|\eta)$. The marginal distribution of v is

$$f(v|\eta) = \int f(v, \theta|\eta)\, d\theta = \int f(v|\theta, \eta)\, f(\theta|\eta)\, d\theta$$

and for the posterior distribution $f(\theta|v, \eta)$ we have

$$f(\theta|v, \eta) = \frac{f(v, \theta|\eta)}{f(v|\eta)} = \frac{f(v|\theta, \eta)\, f(\theta|\eta)}{\int f(v|\theta, \eta)\, f(\theta|\eta)\, d\theta}.$$

The MLE

$$\hat{\eta} = \arg\max_{\eta} f(v|\eta)$$

can be found by the following iterative EM-procedure. First choose initial value $\eta^{(0)}$. Then the next two steps are fulfilled until a stop criterion is met.

E-step. For a known $\eta^{(t)}$ find the posterior distributions

$$f\left(\theta|v, \eta^{(t)}\right) = \frac{f\left(v, \theta|\eta^{(t)}\right)}{f\left(v|\eta^{(t)}\right)}.$$

M-step. Update the η value by the rule

$$\eta^{(t+1)} = \arg\max_{\eta} Q\left(\eta|\eta^{(t)}\right),$$

where the EM-quantity $Q\left(\eta|\eta^{(t)}\right)$ is defined as

$$Q\left(\eta|\eta^{(t)}\right)=\int f\left(\theta|v,\eta^{(t)}\right)\ln f\left(v,\theta|\eta\right)d\theta. \tag{5}$$

Given probability densities f and g, the Kullback-Leibler divergence (see [7]) from g to f is defined by the formula

$$D_{KL}\left(f|g\right)=\int f\left(x\right)\ln\frac{f\left(x\right)}{g\left(x\right)}dx\quad\left(0\cdot\ln 0=0\right)$$

and the following inequality holds

$$D_{KL}\left(f|g\right)\geq 0, \tag{6}$$

where the equality is reached only in the case $f=g$.

Proposition 1 (see [2])**.** *The log-likelihood function* $\ln f\left(v|\eta\right)$ *does not decrease at any combined step of the EM-algorithm.*

Proof. It is not difficult to see that

$$\begin{aligned}&\ln f\left(v|\eta^{(t+1)}\right)-\ln f\left(v|\eta^{(t)}\right)\\&\qquad=Q\left(\eta^{(t+1)}|\eta^{(t)}\right)-Q\left(\eta^{(t)}|\eta^{(t)}\right)+D_{KL}\left(f\left(\theta|v,\eta^{(t)}\right)|f\left(\theta|v,\eta^{(t+1)}\right)\right).\end{aligned}$$

Now according to (6)

$$D_{KL}\left(f\left(\theta|v,\eta^{(t)}\right)|f\left(\theta|v,\eta^{(t+1)}\right)\right)\geq 0$$

and $Q\left(\eta^{(t+1)}|\eta^{(t)}\right)-Q\left(\eta^{(t)}|\eta^{(t)}\right)\geq 0$ by the choice of $\eta^{(t+1)}$, therefore

$$\ln f\left(v|\eta^{(t+1)}\right)-\ln f\left(v|\eta^{(t)}\right)\geq 0. \tag{7}$$

The usual stop condition for the EM-algorithm is a sufficiently small change in the $\ln f\left(v|\eta\right)$ value. □

The announced experience in numerous applications of the EM-algorithm (see for example [8, 11]) indicates favorable properties from computational perspective such as stability and convergence. Typically the algorithm converges to the global maximum but also there are examples where the EM-algorithm converges to a local maximum or a saddle point (see for example [8]).

The presence of prior distributions $f\left(\eta|\tau\right)$ for the parameters, leads to minor changes in the general EM-scheme, which is another advantage of the EM-algorithm. In this case consider the joint distribution

$$f\left(v,\theta,\eta|\tau\right)=f\left(v|\theta,\eta\right)f\left(\theta|\eta\right)f\left(\eta|\tau\right),$$

where τ stands for the hyper-parameters set. For the marginal distribution of (v,η) we find

$$f\left(v,\eta|\tau\right)=\int f\left(v|\theta,\eta\right)f\left(\theta|\eta\right)f\left(\eta|\tau\right)d\theta=\left(\int f\left(v|\theta,\eta\right)f\left(\theta|\eta\right)d\theta\right)f\left(\eta|\tau\right)$$

and the posterior distribution of θ has the form

$$f\left(\theta|v,\eta,\tau\right)=\frac{f\left(v|\theta,\eta\right)f\left(\theta|\eta\right)f\left(\eta|\tau\right)}{\left(\int f\left(v|\theta,\eta\right)f\left(\theta|\eta\right)d\theta\right)f\left(\eta|\tau\right)}=\frac{f\left(v|\theta,\eta\right)f\left(\theta|\eta\right)}{\int f\left(v|\theta,\eta\right)f\left(\theta|\eta\right)d\theta},$$

which shows that

$$f(\theta|v,\eta,\tau) = f(\theta|v,\eta).$$

Therefore the E-step of the EM-algorithm remains intact, while in the M-step the EM-quantity $Q\left(\eta|\eta^{(t)}\right)$ becomes

$$\begin{aligned} Q\left(\eta|\eta^{(t)}\right) &= \int f\left(\theta|v,\eta^{(t)}\right) \ln f(v,\theta,\eta|\tau)\, d\theta \\ &= \int f\left(\theta|v,\eta^{(t)}\right) \ln f(v,\theta|\eta)\, d\theta + \int f\left(\theta|v,\eta^{(t)}\right) \ln f(\eta|\tau)\, d\theta \\ &= \int f\left(\theta|v,\eta^{(t)}\right) \ln f(v,\theta|\eta)\, d\theta + \ln f(\eta|\tau). \end{aligned}$$

Here the change consists of the addend $\ln f(\eta|\tau)$ to the right-hand side of (5). Thus we prove

Proposition 2 (see for example [17]). *The presence of a prior distribution for the parameters set η (or for a part of them) modifies the EM-algorithm only with the addend $\ln f(\eta|\tau)$ to the usual EM-quantity* (5).

Applying the above approach to the basic IRT model, we have to set $v = \left(v_{(j)}\right)$ and $\theta = \left(\theta_{(j)}\right)$, where $\left(v_{(j)}\right)$ is an i.i.d. sample of result indicators. In this case it is not difficult to see that the EM-quantity $Q\left(\eta|\eta^{(t)}\right)$ has the form (the observations additivity of the EM-quantity $Q\left(\eta|\eta^{(t)}\right)$)

$$Q\left(\eta|\eta^{(t)}\right) = \sum_{j=1}^{J} \int f\left(t|v_{(j)},\eta^{(t)}\right) \ln f\left(v_{(j)},t|\eta\right) dt + \left[\ln f(\eta|\tau)\right], \tag{8}$$

where the addend $[\ln f(\eta|\tau)]$ exists only in the case of an existence of some prior parameter distribution.

Now we can present the scheme of the EM-algorithm for the MLE of the parameters η given an i.i.d. result indicators sample $\left(v_{(j)}\right)$ from the basic model (2).

EM-algorithm for the basic IRT model

1	Choose initial value $\eta^{(0)}$
2	E-step. Given $\eta^{(t)}$, find the distributions $f\left(\theta|v_{(j)},\eta^{(t)}\right) = \dfrac{\left(\prod_{i=1}^{I}\prod_{m=0}^{M_i}\left(P_{im}\left(\theta,\xi_i^{(t)}\right)\right)^{v_{(j)im}}\right)\phi\left(\theta|\eta^{(t)}\right)}{\int\left(\prod_{i=1}^{I}\prod_{m=0}^{M_i}\left(P_{im}\left(\theta,\xi_i^{(t)}\right)\right)^{v_{(j)im}}\right)\phi\left(\theta|\eta^{(t)}\right)d\theta}$
3	M-step. Update $\eta^{(t)} \to \eta^{(t+1)}$ according to the rule $\eta^{(t+1)} = \arg\max_{\eta}\left(\sum_{j=1}^{J}\int f\left(\theta|v_{(j)},\eta^{(t)}\right)\ln f\left(v_{(j)},\theta|\eta\right)d\theta + [\ln f(\eta|\tau)]\right)$ where $f\left(v_{(j)},\theta|\eta\right) = \left(\prod_{i=1}^{I}\prod_{m=0}^{M_i}\left(P_{im}(\theta,\xi_i)\right)^{v_{(j)im}}\right)\phi(\theta|\eta)$
4	Repeat steps 2 $\to$ 3 until a stop condition is met.

As a stop criterion one can use a sufficiently small change in the current numerical value of the log-likelihood function

$$\sum_{j=1}^{J} \ln \int \left(\prod_{i=1}^{I} \prod_{m=0}^{M_i} \left(P_{im} \left(\theta, \xi_i^{(t)} \right) \right)^{v_{(j)im}} \right) \varphi \left(\theta | \eta^{(t)} \right) d\theta + \left[\ln f \left(\eta^{(t)} | \tau \right) \right].$$

It is recommended to apply the minimal reasonable tolerance of order 10^{-15} under the 64-bit floating point arithmetic calculations.

The EM-algorithm is the common approach to find MLE for the IRT parametric models. In some earlier works the EM-algorithm was applied under the name MMLE (Marginal Maximum Likelihood Estimation) (see for example [1]). Using of prior distributions is necessary for some models to obtain realistic estimation values (see for example [1]). Especially in the case of 3PL model, the use of a prior distribution for the pseudo-guessing parameter is inevitable.

Optimization. Obviously the M-step of the EM-algorithm requires solving of an optimization problem with many variables (see for example [13, 14, 15]). In this case, however, exists a specific facility because the practical optimization is carried out separately for each test unit, which reduces the dimensionality of the problem to a trivial level and makes the use of the Newton-Raphson method practically almost seamlessly. Here one can apply also various quasi-Newton methods (see for example [9, 13]), such as the method of conjugate gradients and the BFGS method. Here it is possible to apply even some atypical optimization approaches as genetic algorithms (see for example [10]).

The Newton-Raphson scheme for finding critical points of the function $\phi(x)$ requires a choice of an appropriate initial value $x^{(0)}$, after which to iterate

$$x^{(t+1)} = x^{(t)} - H^{-1} \left(x^{(t)} \right) \nabla \phi \left(x^{(t)} \right),$$

where $H(x) = \left(\partial^2 \phi(x) / \partial x_i \partial x_j \right)$ is the corresponding Hesse matrix. The method is super convergent which leads to a small number of iterations (usually less than 16). However the success depends heavily on the initial value choice. Fortunately, in the problem considered, this stage does not appear to be an obstacle.

Numerical integration. In some of the expressions above there are integrals with a weight function $\varphi(\cdot|\sigma)$ ($N\left(0, \sigma^2\right)$ density). The present paper utilizes the traditional Gaussian quadrature integration rule (see for example [14, 15])

$$\int g(\theta) \varphi(\theta|\sigma) d\theta \approx \sum_{s=1}^{S} w_s g(\sigma\theta_s),$$

where the weights (w_s) and nodes (θ_s) are calculated for the case $\sigma = 1$. Using this formula can be regarded as a replacement of the continuous normal distribution for θ with a discrete one, supported on the nodes (θ_s) with probabilities (w_s), which can be assumed as a specific advantage of the Gaussian quadrature.

Using too many nodes does not lead to an essential improvement, because such a precision does not apply appropriately to the 64-bit arithmetic due to the fact that the nodes are gathered in the tails of the density $\varphi(\cdot|\sigma)$.

Ability estimation. The most important objective of the IRT models is to measure the abilities (see for example [1, 3, 4, 5, 6, 12]), which technically lies in the data estimation of θ in the context of the presented model (2). Suppose we have a parameter estimation $\hat{\eta}$. Then the conditional distribution of θ has the form

$$f\left(\theta|v,\hat{\eta}\right)=\frac{\left(\prod_{i=1}^{I}\prod_{m=0}^{M_i}\left(P_{im}\left(\theta,\hat{\xi}\right)\right)^{v_{im}}\right)\varphi(\theta|\hat{\eta})}{\int\left(\prod_{i=1}\prod_{m=0}^{M_i}\left(P_{im}\left(\theta,\hat{\eta}_i\right)\right)\right)\varphi(\theta|\hat{\eta})d\theta}, \tag{9}$$

which admits an interpretation as its posterior distribution given $\varphi\left(\theta|\hat{\eta}\right)$ as a prior one. Our experience shows that $f\left(\theta|v,\hat{\eta}\right)$ appears to be unimodal in the typical case. Formula (9) reveals various estimation options for θ with respect to the result indicator v. The most mentioned in the literature method is the EAP-estimation (Expected A Posteriori)

$$\begin{aligned}\bar{\theta}_{EAP}&=E[f(\theta|v,\hat{\eta})]=\int\theta f(\theta|v,\hat{\eta})d\theta\\&=\frac{\int\theta\left(\prod_{i=1}^{I}\prod_{m=0}^{M_i}\left(P_{im}\left(\theta,\hat{\xi}_i\right)\right)^{v_{im}}\right)\varphi(\theta|\hat{\eta})d\theta}{\int\left(\prod_{i=1}^{I}\prod_{m=0}^{M_i}\left(P_{im}\left(\theta,\hat{\xi}_i\right)\right)^{v_{im}}\right)\varphi(\theta|\hat{\eta})d\theta}\end{aligned} \tag{10}$$

under which one can find the variance

$$\begin{aligned}\sigma^2_{EAP}&=E[(\theta-\bar{\theta}_{EAP})^2]=\int(\theta-\bar{\theta}_{EAP})^2 f(\theta|v,\hat{\eta})d\theta\\&=\frac{\int(\theta-\bar{\theta}_{EAP})^2\left(\prod_{i=1}^{I}\prod_{m=0}^{M_i}\left(P_{im}\left(\theta,\hat{\xi}_i\right)\right)^{v_{im}}\right)\varphi(\theta|\hat{\eta})d\theta}{\int\left(\prod_{i=1}^{I}\prod_{m=0}^{M_i}\left(P_{im}\left(\theta,\hat{\xi}_i\right)\right)^{v_{im}}\right)\varphi(\theta|\hat{\eta})d\theta}.\end{aligned} \tag{11}$$

Practical implementation of (10) and (11) requires only numerical integration without iterations. The other widely used option is the MAP-estimation (Maximum A Posteriori)

$$\hat{\theta}_{MAP}=\arg\max_{\theta}f(\theta|v,\hat{\eta})=\arg\max_{\theta}\left(\left(\prod_{i=1}^{I}\prod_{m=0}^{M_i}\left(P_{im}\left(\theta,\hat{\xi}_i\right)\right)^{v_{im}}\right)\varphi(\theta|\hat{\eta})\right),$$

or simply the ordinary MLE

$$\hat{\theta}_{ML}=\arg\max_{\theta}f(\theta|v,\hat{\eta})=\arg\max_{\theta}\left(\prod_{i=1}^{I}\prod_{m=0}^{M_i}\left(P_{im}\left(\theta,\hat{\xi}_i\right)\right)^{v_{im}}\right)$$

for which one can find the asymptotic variance by means of the Fischer information

$$\mathcal{J}\left(\theta\right)=-E\left[\frac{d^2}{d\theta^2}\ln f\left(v|\theta,\hat{\eta}\right)\right].$$

Remember that $\hat{\theta}_{ML}\sim N(\hat{\theta}_{ML},1/\mathcal{J}(\hat{\theta}_{ML}))$ asymptotically.

Any central tendency measure can also be used as a specific estimation.

Conclusions and online implementation. The review of various models of the modern test theory shows that the most important ones can be dealt with in a relatively simple formal mechanism – the basic IRT model (2). The measured characteristic θ is considered as a latent variable, which in turn directs the estimation of the parameters to be done by means of the common EM-algorithm. Thus one can relatively easily obtain a maximum likelihood estimation for the model parameters. Using prior distributions needs only minor modifications in the original scheme.

The basic IRT model presented above is implemented online by means of the .net platform and the C# language at the domain `http://arsmath.org/`. This implementation until now deals only with items of equal length ($M_i = \text{const}$ for all items). Therein the user can place his/her own data and find the parameters MLE also EAP and MAP estimation for the ability θ. More detailed description can be found in [17].

REFERENCES

[1] F. Baker, S.-H. Kim. Item Response Theory: Parameter Estimation Techniques. New York, Marcel Dekker, 2004.

[2] A. Dempster, N. M. Laird, D. B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B (Methodological)*, **39**, No 1. (1977), 1–38.

[3] R. Hambleton, H. Swaminathan. Item response theory: Principles and applications. Boston, Kluwer-Nijhoff Pub., 1985.

[4] W. Linden, R. Hambleton. Handbook of Modern Item Response Theory. New York, Springer, 1997.

[5] F. Lord. Applications of Item Response Theory to Practical Testing Problems, Hillsdale, N. J, L. Erlbaum Associates, 1980.

[6] F. Lord, M. Novick. Statistical Theories of Mental Test Scores. Reading Mass, Addison-Wesley, 1968.

[7] S. Kullback. Information theory and statistics. New York, Wiley, 1959.

[8] G. McLachlan, T. Krishnan. The EM-Algorithm and Extensions. Hoboken, N. J, Wiley, 2008.

[9] J. Nocedal, S. J. Wright. Numerical optimization. New York, Springer, 1999.

[10] S. Stefanov, D. Tsvetkov. Genetic Algorithm Based Optimization in the Two-Parametric Logistic Model. Varna, CompSysTech'2005, 2005.

[11] M. Watanabe, K. Yamaguchi. The EM-Algorithm and Related Statistical Models. New York, Marcel Dekker, 2004.

[12] B. Wright, M. Stone. Best Test Design. Chicago, MESA Press, 1979.

[13] К. Бабенко. Основы численного анализа. Москва, Наука, 2002.

[14] Б. Боянов. Лекции по числени методи. София, ДАРБА, 1998.

[15] Б. Сендов, В. Попов. Числени методи. Част първа. София, Наука и изкуство, 1976.

[16] Е. Стоименова. Измерителни качества на тестовете. София, НБУ, 2000.

[17] Д. Цветков. Вероятностни модели в теория на тестовете. Велико Търново, Университетсто издателство „Св. св. Кирил и Методий 2015.

[18] Д. Цветков, Л. Христов. Основен вероятностен модел в теория на тестовете. Велико Търново, Майски четения на СУБ, 2009.

Dimiter Petkov Tsvetkov
e-mail: dimiter.petkov.tsvetkov@gmail.com
Lyubomir Yanakiev Hristov
e-mail: lyubomir.hristov@gmail.com
Department of Mathematical Analysis and Applications
Faculty of Mathematics and Informatics
St Cyril and St Methodius University of Veliko Tarnovo
5000 Veliko Tarnovo, Bulgaria

Ralitsa Lyubomirova Angelova-Slavova
e-mail: r.angelova.slavova@abv.bg
Department of Communication Systems and Technologies
Vasil Levski National Military University at Veliko Tarnovo
5000, Veliko Tarnovo, Bulgaria

БЕЛЕЖКИ ВЪРХУ ОЦЕНКАТА НА ПАРАМЕТРИТЕ ЗА НЯКОИ IRT МОДЕЛИ ЧРЕЗ EM-АЛГОРИТЪМА

Димитър П. Цветков, Любомир Я. Христов, Ралица Л. Ангелова-Славова

В тази статия се разглежда имплементация на EM-алгоритъма за оценка на параметрите на някои IRT модели. Схемата е описана в главните детайли. Посочена е връзка към авторска онлайн реализация на използвания метод относно различни конкретни IRT модели, включително бинарните модели с един, два и три параметъра (1PL, 2PL, 3PL), номиналния модел на Бок (NRM), модела на частичния кредит (PCM) и обобщения модел на частичния кредит (GPCM).