DOCUMENT RESUME

ED 227 171 TM 830 195

AUTHOR Reckase, Mark D.; McKinley, Robert L.
TITLE The Feasibility of a Multidimensional Latent Trait Model.
INSTITUTION American Coll. Testing Program, Iowa City, Iowa.
SPONS AGENCY Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
PUB DATE Aug 82
CONTRACT NR0014-81-K0817
NOTE 11p.; Paper presented at the Annual Meeting of the American Psychological Association (Washington, DC, August 23-27, 1982).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Correlation; Item Analysis; *Latent Trait Theory; *Mathematical Models; Maximum Likelihood Statistics; Psychometrics; Research Methodology; Scores; *Test Items
IDENTIFIERS *Multidimensional Approach; Multiple Measures Approach; *Rasch Model

ABSTRACT

A class of multidimensional latent trait models is described. The properties of the model parameters, and initial results on the accuracy of a maximum likelihood procedure for estimating the model parameters are discussed. The model presented is a special case of the general model described by Rasch (1961), with close similarities to the models suggested by Bock and Aitkin (1981) and Samejima (1974). The concepts of item difficulty and discrimination were discussed in reference to this model as generalizations of the same concepts used in the unidimensional latent trait models. For this case, difficulty was shown to be defined by a function rather than a single value, and discrimination was shown to be related to the slope of the item characteristic surface at its intersection with the .5-plane. Both the difficulty function and discrimination parameters are most easily interpreted when determined conditional on a particular dimension. The maximum likelihood estimation procedure that was developed for the model was given an initial trial on a set of simulation data that was generated to contain two distinct dimensions. The item parameters estimated from the simulated test data were shown to be very highly related to the true parameters. (Author/PN)

The Feasibility of a Multidimensional Latent Trait Model

Mark D. Reckase and Robert L. McKinley
The American College Testing Program

Because of the complexity of the interaction between a person and a set of test items, there has been a continuing search for multidimensional models of test taking behavior. Although test results are sometimes thought to be an indicant of some ideal continuum, this hypothesized state of nature is never truly achieved. Even conceptually simple tests such as vocabulary and number series are found to have several components if they are studied closely enough (Holzman, Glaser and Pellegrino, 1980). Many tests are not even intended to measure a single trait. For example, achievement tests are typically designed to measure a number of content areas. A single achievement dimension is not usually defined.

The approach taken in the past to deal with the multidimensional nature of test data has been to determine the number and composition of the components in a test and to use that information to form tests measuring a single dimension. Unfortunately, the procedures that have been developed for the purpose of sorting test items into sets measuring a single dimension do not work well with the data yielded by the typical, dichotomously scored, multiple choice test items. Factor analysis is plagued by problems dealing with the selection of a similarity coefficient, the effects of guessing on the solution, and the conceptual problem of using a model derived for continuous data on dichotomous scores (Kim and Mueller, 1978). The use of non-metric multidimensional scaling as an alternative to factor analysis has not been well researched, and the results that have been reported are inconclusive (Reckase, 1981).

An alternate approach to the problem of multidimensionality in test data is to develop a model that is designed to explain the responses to dichotomously scored test items using a number of hypothetical dimensions. In a sense, this has been done by Christoffersson (1975) and Muthén (1978) from a factor analysis perspective, but their approach requires extensive computation. The approach presented in this paper is to describe the multidimensional interaction using a latent trait approach, which will have the advantages of using the sample-free properties of latent trait theory and the availability of useful statistics such as the item and test information functions.

To date, little work has been done with multidimensional latent trait models, despite the fact that several variations have been described in the literature (Bock and Aitkin, 1981; Rasch, 1961; Samejima, 1974; Sympson, 1978; Mulaik, 1972; Whitely, 1980). For the most part, the references to

Paper presented at the annual meeting of the American Psychological Association, Washington, D.C., August 1982. This research was supported by contract number NR0014-81-K0817 from the Personnel and Training Research Programs of the Office of Naval Research.

the models in the literature are only descriptions of the mathematical forms with little information about actual applications. Guidelines for the interpretation of results obtained using these models are nonexistent.

This paper will discuss one class of multidimensional latent trait models that is related to the models proposed by Bock and Aitkin (1981), Rasch (1961), and Samejima (1974). The form of the model will be presented first and information concerning the interpretation of the model parameters will be given. Also, the applicability of the model to test data will be demonstrated using simulated item responses.

Characteristics of the Model

The particular multidimensional latent trait model presented here was selected on the basis of a detailed analysis of the general Rasch model (1961). A full report of that analysis was given elsewhere (Reckase and McKinley, 1982). The equation for the model is given by

$$P(X_{ip} = 1 | \underline{\sigma}_i, \underline{\theta}_p) = \frac{e^{\sum_{j=1}^{n} \sigma_{ij}\theta_{pj} + \sigma_{i(n+1)}}}{1 + e^{\sum_{j=1}^{n} \sigma_{ij}\theta_{pj} + \sigma_{i(n+1)}}} \qquad (1)$$

where x_{ip} is the score on Item i for Person p,
$\underline{\sigma}_i$ is a vector of item parameters for Item i,
and $\underline{\theta}_p$ is a vector of person parameters for Person p.

The relationship between the probability of a correct response and the θ-vector for this model defines an item response surface. An example of the item response surface for the two dimensional case of this model is given in Figure 1 with $\sigma_1 = 1.5$, $\sigma_2 = .5$ and $\sigma_3 = .650$. Note that this surface is monotonically increasing in both θ_1 and θ_2.

Insert Figure 1 about here

In order to interpret the parameters of this model, the traditional concept of item difficulty and discrimination must be extended to the multidimensional case. In unidimensional latent trait models, the difficulty of an item is defined as the point on the θ-scale below the point of inflection of the item characteristic curve (See Figure 2). This value can be determined by setting the second partial derivative of the model with respect to θ to zero and solving for θ. The same procedure can be followed for the multidimensional model, except that the second partial derivative is now taken with respect to the θ-vector. For this particular model, the solution of the second partial derivative yields a function rather than a single value for the difficulty of the item. The function is given by

$$\sum_{j=1}^{n} \sigma_{ij}\theta_{pj} + \sigma_{i(n+1)} = 0 \qquad (2)$$

where the σ- and θ-terms are the elements of the σ- and θ-vectors. For the special case of the model shown in Figure 1, this function is given by the equation

$$1.5\theta_{p1} + .5\theta_{p2} - .650 = 0. \qquad (3)$$

This is the formula for a line in the θ_1, θ_2 plane. The line is shown by the dashed line on the surface shown in Figure 1. Individuals whose θ-vectors locate them on one side of this line have a greater than .5 probability of a correct response, while those located on the other side of the line have less than a .5 probability of a correct response. The difficulty function is defined by the intersection of the item response surface with the plane parallel to the θ_1, θ_2 plane at a probability of .5.

Insert Figure 2 about here

If all of the θ-values but one are set equal to zero, the result is a "conditional" difficulty for the item on that dimension. For the example given above, the conditional difficulty on Dimension 1 is .43 and the conditional difficulty on Dimension 2 is 1.3. Hence, the item requires proportionally more ability on Dimension 2 than Dimension 1 to obtain a high probability of a correct response. Note that all of the parameters of the model enter into the definition of the item difficulty function.

Just as the concept of item difficulty can be generalized from unidimensional latent trait theory to multidimensional latent trait theory, so to can the concept of item discrimination. In unidimensional theory, the discrimination parameter for an item is a function of the slope of the item characteristic curve at the point of inflection (see Figure 2). For the two parameter logistic model, this function is given by

$$a_g = \frac{(\text{slope at } b_g) \times 4}{1.7} \qquad (4)$$

where a_g is the discrimination parameter and b_g is the difficulty parameter. For the multidimensional model, the slope can be determined by solving the first partial derivative with respect to θ for values on the difficulty function. If the slope is measured parallel to a Dimension d the slope is given by the expression $\sigma_{id}/4$ for Item i. Thus, the discrimination parameter for the item on Dimension d can be defined as four times the slope measured parallel to that dimension. For the example given in Figure 1, the slope with respect to Dimension 1 is .375, and with respect to Dimension 2 it is .125. The surface is, therefore, "flatter" with respect to Dimension 1 than Dimension 2. On the basis of this information, the first n terms of the σ-vector can be considered as discrimination parameters.

Estimation of Parameters

The development of a model alone is not sufficient to show that a multidimensional latent trait model is a practical possibility. A procedure must be developed to estimate the parameters of the model with sufficient accuracy to judge whether they yield useful information about the test items and the examinees.

For the model presented in Equation 1, an empirical maximum likelihood procedure has been developed for the estimation of the person and item parameters. The item parameter estimation procedure begins with an initial estimate of the item parameters based on a weighted sum of ability parameter estimates. These values are used as the starting point for an iterative Newton-Raphson procedure that modifies the item parameter estimates on each dimension before estimating those on the next dimension. The iterations continue until successive estimates of the parameters do not differ by more than a specified value. A full description of the procedure is given in Reckase and McKinley (1982).

The ability parameter estimates are determined in a similar fashion. Initial parameter estimates are obtained from a weighted sum of the item parameter estimates. The initial values are used as the starting point in a Newton-Raphson procedure for finding the maximum of the likelihood function for the data. This procedure is also described in Reckase and McKinley (1982).

In order to determine the practicality of the estimation procedure that was developed, it was applied to a set of simulated test data that was generated to fit the model. The simulated test data contained responses on 50 items for 1,000 exaimees. The data were generated to model two distinct ability dimensions. The item parameters used to generate data to fit the model are given in the second through fourth columns of Table 1. The first of these three columns gives the $\sigma_{i(n+1)}$ term from the model and the other two columns give the discrimination parameters. The ability parameters used to generate the data were sampled from the bivariate normal distribution with $\rho = 0$, $\underline{\mu} = 0$ and $\underline{\Sigma} = \underline{I}$.

Insert Table 1 about here

The results of the item parameter estimation procedure for the two-dimensional model are presented in the last three columns of Table 1. These estimates have been scaled to have the same mean and standard deviation as the true parameters. The correlations between the parameter estimates and the true values are given in Table 2. From the information presented, it can be seen that the parameter estimates are very highly related to the true values. While this does not conclusively show the value of this estimation procedure, the results do suggest that the procedure is very promising.

Insert Table 2 about here

Summary

The purpose of this paper was to describe a class of multidimensional latent trait models, discuss the properties of the model parameters, and give some initial results on the accuracy of a maximum likelihood procedure for estimating the model parameters. The model presented is a special case of the general model described by Rasch (1961). It also has close similarities to the models suggested by Bock and Aitkin (1981) and Samejima (1974).

The concepts of item difficulty and discrimination were discussed in reference to this model as generalizations of the same concepts used in the unidimensional latent trait models. For this case, difficulty was shown to be defined by a function rather than a single value, and discrimination was shown to be related to the slope of the item characteristic surface at its intersection with the .5-plane. Both the difficulty function and discrimination parameters are most easily interpreted when determined conditional on a particular dimension.

The maximum likelihood estimation procedure that was developed for the model was given an initial trial on a set of simulation data that was generated to contain two distinct dimensions. The item parameters estimated from the simulated test data were shown to be very highly related to the true parameters. Based on these results, and on the interpretive results presented earlier, this model seems very promising as a means of describing the interaction between a person and a test item in a multidimensional latent space.

References

Bock, R.D. and Aitkin, M. Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm. Psychometrika, 1981, 46, 443-459.

Christoffersson, A. Factor analysis of dichotomized variables. Psychometrika, 1975, 40, 5-32.

Holzman, T.G., Glaser, R. and Pellegrino, J.W. Cognitive determinants of series completion: Individual and developmental differences. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Kim, J. and Mueller, C.W. Factor analysis: Statistical methods and practical issues. Beverly Hills, CA: Sage Publications, 1978.

Mulaik, S.A. A mathematical investigation of some multidimensional Rasch models for psychological tests. Paper presented at the meeting of the Psychometric Society, Princeton, NJ, March 1972.

Muthén, B. Contributions to factor analysis of dichotomous variables. Psychometrika, 1978, 43, 551-560.

Rasch, G. On general laws and the meaning of measurement in psychology. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, 1961, 4, 321-334.

Reckase, M.D. The formation of homogeneous item sets when guessing is a factor in item responses (Research Report 81-5). Columbia: University of Missouri, Department of Educational Psychology, August 1981.

Reckase, M.D. and McKinley, R.L. Some latent trait theory in a multidimensional latent space. Paper presented at the invitational conference on item response theory and computerized adaptive testing. Wayzata, MN, July 1982.

Samejima, F. Normal ogive model on the continuous response level in the multidimensional latent space. Psychometrika, 1974, 39, 111-121.

Sympson, J.B. A model for testing with multidimensional items. In D.J. Weiss (Ed.), Proceedings of the 1977 Computerized Adaptive Testing Conference. University of Minnesota, 1978.

Whitely, S.E. Multicomponent latent trait models for ability tests. Psychometrika, 1980, 45, 479-494.

Table 1
True and Estimated Item Parameters

Item	True Parameters			Maximum Likelihood Estimates		
	1	2	3	1	2	3
1	-0.65	1.50	0.50	-0.48	1.27	0.36
2	-1.40	0.50	1.25	-1.56	0.45	1.29
3	-0.20	1.35	0.15	-0.06	1.69	0.25
4	0.40	1.60	0.55	0.59	1.58	0.63
5	0.00	0.50	1.15	0.10	0.51	1.04
6	-1.30	0.35	1.05	-1.68	0.42	1.25
7	0.05	1.45	0.35	0.34	1.63	0.35
8	0.19	0.25	1.40	0.25	0.20	1.23
9	-0.17	0.85	0.85	-0.07	0.89	0.67
10	0.14	1.75	0.45	0.34	1.45	0.57
11	0.37	0.60	0.80	0.29	0.61	0.88
12	0.87	1.65	0.65	0.70	1.38	0.46
13	-0.93	0.35	1.35	-1.00	0.24	1.65
14	1.85	0.65	1.65	1.59	0.54	1.29
15	0.06	0.65	0.65	0.00	0.80	0.52
16	-0.41	0.45	1.45	-0.38	0.23	1.60
17	-1.54	0.75	1.25	-1.55	0.71	1.14
18	0.34	1.55	0.25	0.39	1.46	0.29
19	-0.15	0.65	1.35	0.04	0.65	1.28
20	1.48	1.25	0.45	1.42	1.32	0.22
21	-1.45	1.65	0.45	-1.73	1.80	0.48
22	0.75	0.45	1.35	0.67	0.40	1.35
23	-0.75	0.35	1.55	-0.63	0.21	1.72
24	1.10	1.10	0.30	0.93	1.11	0.35
25	-0.55	1.20	0.15	-0.38	1.37	0.24
26	0.50	0.50	1.00	0.28	0.49	1.12
27	-0.15	1.45	0.45	-0.02	1.32	0.42
28	0.65	0.70	0.70	0.56	0.68	0.86
29	-1.00	1.00	0.30	-0.95	1.14	0.38
30	1.00	0.30	1.00	0.98	0.27	1.14
31	-0.25	0.95	0.25	-0.04	1.11	0.24
32	-0.70	0.15	1.50	-0.68	0.03	1.51
33	0.85	1.15	0.45	0.98	1.29	0.30
34	0.05	0.10	0.95	0.03	0.20	1.09
35	-0.95	1.35	0.50	-0.80	1.28	0.52
36	-1.50	0.20	1.20	-1.38	0.47	1.03
37	1.80	1.55	0.55	2.06	1.51	0.44
38	-2.00	0.15	1.15	-2.06	0.30	1.09
39	-0.90	1.40	0.35	-0.70	1.18	0.40
40	1.00	1.00	1.00	1.04	0.87	0.79
41	0.15	1.25	0.70	0.28	1.17	0.75
42	-1.50	0.25	0.95	-1.60	0.49	1.14
43	-1.25	0.35	1.45	-1.13	0.26	1.30
44	1.25	1.30	0.25	0.95	1.26	0.26
45	-2.00	1.15	0.15	-2.01	1.30	0.13
46	1.75	0.50	0.50	1.65	0.68	0.64
47	0.65	0.65	1.30	0.71	0.49	1.27
48	-0.25	1.00	0.45	-0.17	1.03	0.53
49	0.35	0.55	1.15	0.24	0.48	1.04
50	0.00	0.95	0.15	-0.07	1.11	0.26

Table 2

Intercorrelation Matrix for True and Estimated Item Parameters

Parameter	True Parameters			Estimates		
	σ_{T1}	σ_{T2}	σ_{T3}	$\hat{\sigma}_{C1}$	$\hat{\sigma}_{C2}$	$\hat{\sigma}_{C3}$
σ_{T1}	1.00	0.21	-.12	0.99	0.15	.18
σ_{T2}	0.21	1.00	-.75	0.15	0.96	-.79
σ_{T3}	-.12	-.75	1.00	-.18	-.79	-.96
$\hat{\sigma}_{C1}$	0.99	0.15	-.18	1.00	0.20	-.22
$\hat{\sigma}_{C2}$	0.15	0.96	-.79	0.20	1.00	-.88
$\hat{\sigma}_{C3}$	-.18	-.79	-.96	-.22	-.88	1.00

Figure 1

Example of an Item Response Surface for the Two-Dimensional Case

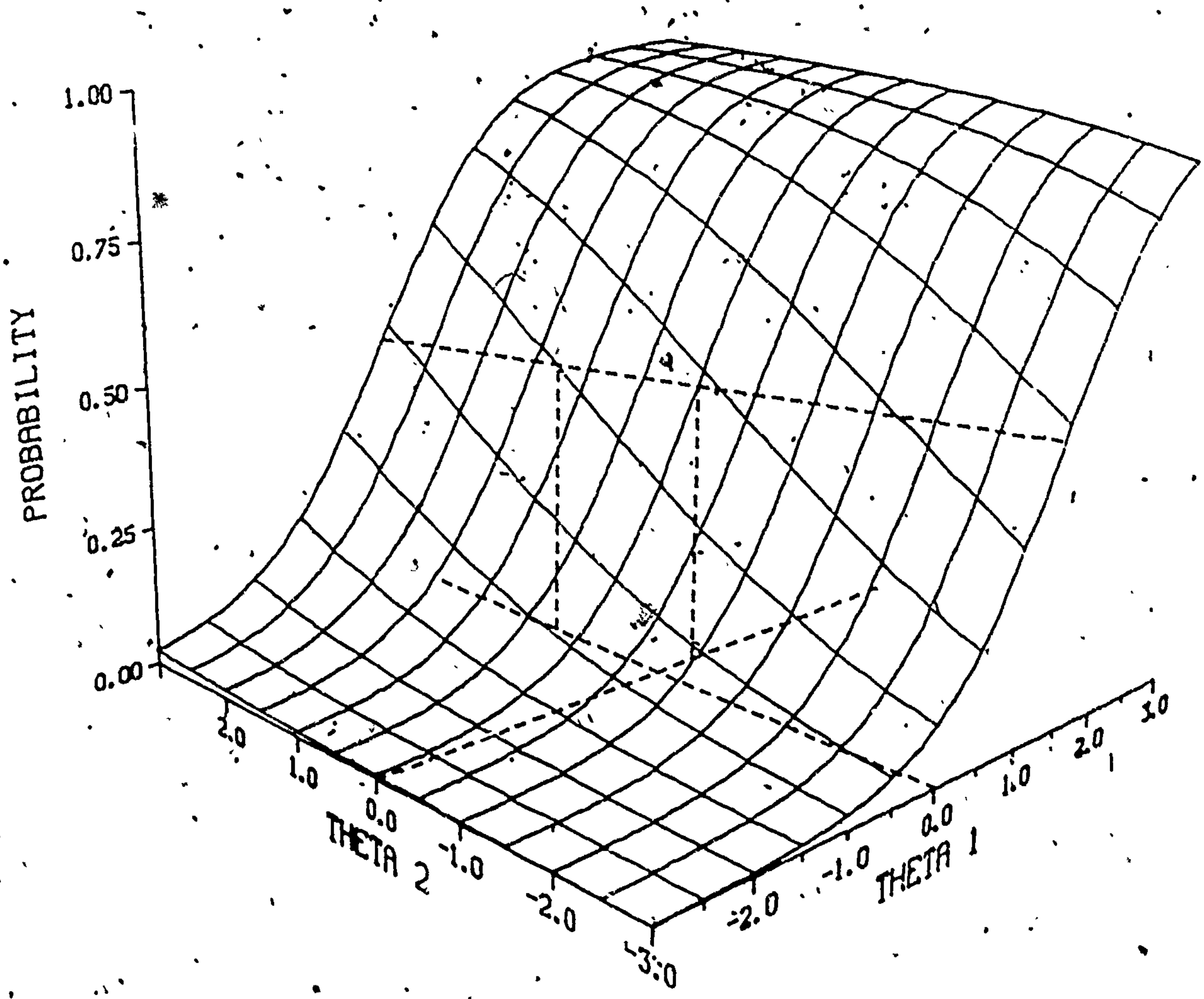

Figure 2

Slope and Point of Inflection
for the Unidimensional Two-Parameter Logistic Model