THE THEORETICAL DETECT INDEX OF DIMENSIONALITY AND ITS APPLICATION TO APPROXIMATE SIMPLE STRUCTURE

JINMING ZHANG
EDUCATIONAL TESTING SERVICE

WILLIAM STOUT
DEPARTMENT OF STATISTICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

In this paper, a theoretical index of dimensionality, called the theoretical DETECT index, is proposed to provide a theoretical foundation for the DETECT procedure. The purpose of DETECT is to assess certain aspects of the latent dimensional structure of a test, important to practitioner and researcher alike. Under reasonable modeling restrictions referred to as "approximate simple structure", the theoretical DETECT index is proven to be maximized at the correct dimensionality-based partition of a test, where the number of item clusters in this partition corresponds to the number of substantively separate dimensions present in the test and by "correct" is meant that each cluster in this partition contains only items that correspond to the same separate dimension. It is argued that the separation into item clusters achieved by DETECT is appropriate from the applied perspective of desiring a partition into clusters that are interpretable as substantively distinct between clusters and substantively homogeneous within cluster. Moreover, the maximum DETECT index value is a measure of the amount of multidimensionality present. The estimation of the theoretical DETECT index is discussed and a genetic algorithm is developed to effectively execute DETECT. The study of DETECT is facilitated by the recasting of two factor analytic concepts in a multidimensional item response theory setting: a dimensionally homogeneous item cluster and an approximate simple structure test.

Key words: item response theory, dimensionality, multidimensionality, generalized compensatory model, simple structure, approximate simple structure, genetic algorithm, dimensionally homogeneous item cluster, latent space, test space.

1. Introduction

Classically, educational or psychological tests have been idealized as having exactly one latent trait. Often a more realistic viewpoint is that a test has an intentionally concentrated measurement focus, which can be modeled by a unidimensional composite of latent traits called the "test composite", whereas the test itself is multidimensional for a given examinee population. The test composite is defined to be a particular linear combination of the test's complete latent trait variables. This test composite, formalized in section 2 in a manner consistent with our non-parametric emphasis, should be thought of intuitively as being aligned in the linear direction in which the number correct test score best measures. Typically a test is designed to have its items dispersed away from the test composite because of construct-relevant contextual dimensions, like "science" paragraph in a reading comprehension test, or because items are designed to measure only certain components of the complete latent variables, like "geometry" in a mathematics test.

Portions of this article were presented at the annual meeting of the National Council on Measurement in Education, New York, April 1996.

The authors would like to thank Ting Lu, Ming-mei Wang, and four anonymous reviewers for their comments and suggestions on an earlier version of this manuscript. The research of the first author was partially supported by an ETS/GREB Psychometric Fellowship, and by Educational Testing Service Research Allocation Projects 794-33 and 883-05. The research of the second author was partially supported by NSF grants DMS 94-04327 and DMS 97-04474.

Requests for reprints should be sent to Jinming Zhang, Educational Testing Service, MS 02-T, Rosedale Road, Princeton NJ 08541. E-mail: jnzhang@ets.org

0033-3123/1999-2/1996-0504-A $00.75/0
© 1999 The Psychometric Society
Here an item’s direction is intuitively viewed as the direction in which the item best measures, often in the sense of being most discriminating in that direction. The test composite should also be viewed intuitively as a sort of average or centroid of all the item directions. Geometrically, each item is appropriately visualized as both rotated away from and rotated “around” the test composite. Rotation away is interpreted as an item measuring a dimension in addition to the test composite, while rotation around the test composite separates such dimensions from each other.

A certain pattern of separated clusters of items about the test composite should typically result from the categorical nature of many test specifications. For example a “mathematics” test might consist of three test specification-mandated clusters of items: algebra, geometry, and trigonometry. Suppose the test is well fit by a three dimensional latent space with coordinate system \( (\Theta_{AI}, \Theta_{Ge}, \Theta_{Tr}) \) and “mathematics” test composite \( \Theta_T = (\Theta_{AI} + \Theta_{Ge} + \Theta_{Tr})/3 \). Consider the algebra cluster. A “pure” algebra item is the one with direction in the \( \Theta_{AI} \) axis direction. Certainly the “algebra” cluster should contain all the pure algebra items of the test. But, consider a predominantly algebraic item that also measures to a lesser extent geometry and trigonometry. We would still want to consider this an algebra item. Thus, substantively speaking, it should be assigned to the algebra cluster. Depending on the size of the geometry and trigonometry influences, the item’s rotation away from \( \Theta_T \) will be only a fraction of the total \( \Theta_{AI} \) rotation away from \( \Theta_T \) and hence the item will not be aligned with \( \Theta_{AI} \). Furthermore, the item’s angular rotation around the \( \Theta_T \) axis should be relatively close to the \( \Theta_{AI} \) axis rotation (exactly equal to the \( \Theta_{AI} \) rotation if a symmetry of influence between geometry and trigonometry exists). Thus based on these substantive considerations, we expect the “algebra” cluster to consist of items rotated an arbitrary amount away from \( \Theta_T \) roughly towards the \( \Theta_{AI} \) axis. Also, we would expect the three clusters of this test likely to be sufficiently separated from each other, thereby creating a clumping into three geometrically separated and dimensionally homogeneous clusters. Hence, informally, a cluster being “dimensionally homogeneous” means that although its items are allowed to be rotated away from the test composite \( \Theta_T \) by arbitrary angles, all its item rotations around the test composite should be restricted to a somewhat narrow sector, with all item rotation angles close to one of the axis rotation angles, say, the \( \Theta_{AI} \) axis rotation angle. Substantively, this is equivalent to an item being more influenced by algebra than either geometry or trigonometry. The main point of this paragraph is that our intentionally geometrically relaxed notion of item directional measurement similarity that characterizes membership in a dimensionally homogeneous cluster provides an appropriate clustering principle for partitioning a test into substantively meaningful clusters.

Informally, if a test consists of dimensionally homogeneous clusters that are sufficiently separated in their varying rotations around the test composite, the test will be said to display “approximate simple structure”. The unique partition into clusters when approximate simple structure holds is often referred to as the correct partition or the dimensionality-based partition, in the sense that when a statistical procedure finds this partition, it has found the “correct” partition into substantively meaningful and dimensionally separate clusters. Approximate simple structure of a test and dimensional homogeneity of a cluster will be defined rigorously in section 3. The key point to note from the formal quantitative definition of approximate simple structure is that rather than being heavily restrictive, approximate simple structure is very consistent with the geometric clustering we expect from the categorical nature of typical test specifications. In particular, the formal definition allows surprisingly wide angular variation among items that are members of a “dimensionally homogeneous” cluster, thereby making approximate simple structure a widely applicable and realistic model-based concept.

Given a test with test composite \( \Theta_T \), it is very important to identify the number of substantively meaningful and distinct latent dimensions, to estimate the amount of test multidimensionality, and to correctly assign items to the resulting unique dimensionally homogeneous clusters when approximate simple structure exists. Growing out of the \( \hat{r} \) index (Junker & Stout, 1994), Kim (1994) proposed a data-driven index of dimensionality called the DETECT index intended