What is a q-matrix?

Q-matrix is a matrix describing relations of questions and concepts required for their understanding. It is a domain-independent model of knowledge represented by a binary matrix showing the relationship between test items and latent or underlying attributes, or concepts.

How does a q-matrix look like?

Matrix elements

Q-matrix is a M*x*N matrix, where M equals the number of questions in an assessment, and N equals the total number of concepts required for understanding all questions. The matrix element A[i,j] equals 1 if the i-th concept is required for correctly answering j-th question and 0 if the i-th concept is NOT required for correctly answering j-th question. Alternatively, matrix values can be not just {0,1}, but real numbers from the interval [0,1], describing the probability that a student who knows i-th concept will correctly answer j-th question.

Matrix example

An example of a q-matrix is shown in the following image.

	Questions				
	1	2	3	4	5
Concept 1	0	0	1	1	1
Concept 2	1	1	1	1	0

Understanding matrix values and implications

From the matrix, one can read that knowledge of concept 1 is required in order to answer correctly questions 3, 4 and 5. One can also read that questions 1 and 2 test only the knowledge of concept 2.

Furthermore, one could also say that the *ideal response* of a student taking the test formed of those 5 questions who knows only concept 1 should be "00001". This is so since he does not know concept 2 which is required for questions 1 and 2 (therefore the leading "00"). Yet this concept is also required for correct ansqering of questions 3 and 4 so he can not answer those questions neither (therefore the following "00"). Finally, 5th question requires only knowledge about the concept 1 so the student can answer this question correctly (therefore the ending "1", forming all together "00001").

What can I do with a q-matrix?

Q-matrix can be used for understanding students' performance. Due to various knowledge and assessment characteristics, students' responses rarely match *ideal responses* generated from the matrix. Still, by assigning the closest *ideal response* to a student's response vector, it can be assumed which concepts the student does, and which he does not know. This information can be used in order to direct him in further learning.

How do I create a q-matrx?

The goal of q-matrix construction is to extract underlying, or latent, variables, which account for students' differential performance on questions. A q-matrix can be created in two ways:

- by having the domain experts analyze assessment items (questions) and define the concept (often referred to as attribute) corresponding to each item and thereby construct the q-matrix. In this case all the concepts in the q-matrix are exactly defined and labeled as elements of the domain. Q_matrix and student's knowledge is therefore easily interpretable. This method relies on a simple hill-climbing algorithm that creates a matrix representing relationships between concepts and questions directly. The algorithm varies c, the number of concepts, and the values in the q-matrix, minimizing the total error for all students for a given set of n questions. To avoid of local minima, each hill-climbing search is seeded with different random Q-matrices and the best of these is kept. Problem: “*a q-matrix is a much more abstract measure of the relationships of questions to concepts. We might assume that the questions designed to test students are a more accurate reflection of the teaching objectives than an abstract construct which relates questions to underlying concepts.*”
- by automatically extracting the matrix directly from obtaines students' performance on the test items. in this case, the number of concepts is determined by the algorithm (and is usually much smaller than the number of assessment items) and their labels are automatically assigned (and do not reffer to exact domain concepts). Experts can examine the resulting q-matrix to ensure that the extracted relationships seem to be valid, and then use that q-matrix to guide the generation of new problems.

Factor analysis

Factor analysis can be considered as an alternative to q-matrices. Concepts are in that case automatically determined using covariance matrix. Number of concepts should be smaller than number of questions. Still, this methos has proven to be less fault tollerant. Still, when forming a correlation matrix, we lose individual student data in favor of calculating average relationships between questions. The q-matrix method is optimized to assign each student the most appropriate knowledge state, using all available response data for each student.

Discussion

As Sellers found in her research, the results obtained through q-matrix analysis seem to describe relationships among variables in interpretable ways. Factor analysis and principal components analysis, on the other hand, do not readily offer interpretable results.

Later researchers found that, although the q-matrix model was a good way to compare student data to a concept model, expert-constructed q-matrices did not correspond to student data any better than random q-matrices did.

The findings in Brewer's previous research, which found that the factor analysis method performed poorly in comparison with the q-matrix method when fewer observations were available.

Literature

Shuqun, Yang, Cai Shenzheng, Yao Zhiqiang, and Ding Shuliang. “The Concept Lattices of Q-Matrices.” In First International Workshop on Education Technology and Computer Science, 2009. ETCS ’09, 1:413 –417, 2009. doi:10.1109/ETCS.2009.101.

Barnes, Bitzer. Fault Tolerant Teaching and Automated Knowledge Assessment.

Barnes, Tiffany Michelle. “The Q-Matrix Method of Fault-Tolerant Teaching in Knowledge Assessment and Data Mining.” North Carolina State University, 2003.

Barnes, M. T. The Q-Matrix Method: Mining Student Response Data for Knowledge.