

Q-Matrix Development

NCME 2009
Workshop

Introduction

- We will define the Q-matrix
- Then we will discuss method of developing your own Q-matrix
- Talk about possible problems of the Q-matrix to avoid

The Q-matrix

- To begin, let's assume that I am interested in measuring math ability
- In addition, let's assume that I have four well defined and validated attributes
 - Addition
 - Subtraction
 - Multiplication
 - Division

The Q-matrix

- Possible items may be:
 - $2+3-1$
 - $4/2$
 - $(4 \times 2) + 3$
- Not all items measure all attributes.
- A Q-matrix is used to indicate which attributes are required for each item.

The Q-matrix

- An example of a Q-matrix using our math test.

	Add	Sub	Mult	Div
2+3-1	1	1	0	0
4/2	0	0	0	1
(4 x 2)+3	1	0	1	0

Examinees

- In this case, examinees are characterized by profiles that specify which attributes have been mastered.
- This is done in the same way as the Q-matrix.
- For example:

	Add	Sub	Mult	Div
Examinee A	1	1	0	0

Examinees

Q-matrix

	Add	Sub	Mult	Div
2+3-1	1	1	0	0
4/2	0	0	0	1
(4 x 2)+3	1	0	1	0

By knowing what attributes are required by each item and what attributes have been mastered by an individual, we can determine which items will likely be answered correctly by each examinee.

Examinee Mastery

	Add	Sub	Mult	Div
Examinee 1	1	1	0	0
Examinee 2	0	1	0	1
Examinee 3	1	0	1	0
Examinee 4	1	1	1	0

Prob Ans #1

Prob Ans #2

Prob Ans #3

Prob Ans #1 & #3

Diagnostic Classification Models

- The models define the chances of a correct response based on the examinee's profile.
 - Most assume that the Q-matrix defines the ONLY set of attribute measured by the item
- The models range in their complexity.
- Each model determines the basic set of assumptions that relate a persons profile to how they are expected to respond.

The Q-Matrix

- Notice that by specifying the Q-matrix we have defined the attributes of interest
- If this is done carelessly, it is possible that the attributes are not well defined and as a result your parameters will be meaningless
- In many ways, development of the Q-matrix is one of the most important steps of DCMs

Q-matrix Construction

- It should be noted that if great care has been placed on the development of your assessment, Q-matrix construction should be trivial
- By “great care” we mean that:
 - Attributes have been well defined
 - Attributes have been validated
 - Items were developed based on these attributes

Q-matrix

- In many ways, Q-matrix can be seen as a second level of attribute validation
- That is, items are developed to measure specific attributes
- These attributes should be clear
- That said, the following describes the set of methods that could be used to develop the Q-matrix

The Q-matrix Construction

- Basic methods of Q-matrix construction:
 - Simple inspection of the items.
 - Multiple Rater Methods.
 - Iterative procedures based on item parameters.

Simple Inspection

- In using simple inspection, we are evaluating the item and determining what attributes are required to answer each item.
- All one must do is determine which of the attributes are required to correctly answer each item.
- To do this, we recommend working through each question and making note of which attributes were used.

Examples

- A basic math test designed to measure (addition, subtraction, multiplication, and division).
 - $2+3-1$
- A questionnaire designed to measure the 10 criteria used to define a pathological gambler.
 - *For example, “I find it difficult to stop gambling.”*

Examples

- Other examples may include tests that have been designed to measure specific parts of speech or verbal ability.
- The important thing is that the tests were created to measure multiple attributes

Multiple-Raters

- A more likely situation is where a set of experts/researchers are working on the same project.
- In that case, each of the researchers may follow the same procedures as previously outlined.
 - Determine the attributes.
 - Specify required attributes for each item.
 - Refine Q-matrix.

Multiple-Raters

- However, it is unlikely that they will all provide the same answer.
- Therefore, as a second possibility, we consider the procedures of Q-matrix development for multiple raters.

Determine Skills

- To begin, we recommend ensuring that all experts (or a sub-committee) agree with the definitions of your attributes.
 - This procedure is the same as the validation before, only now they must agree on the set of attributes
 - Given that the basic set of attributes have been agreed upon, a thorough definition for each should be refined.
 - These definitions should be given to all experts.

Development of the Q-matrix

- Each expert is now asked to create the Q-matrix.
- Here we have two possible options:
 - Use 0/1 for the Q-matrix.
 - Rate each skill based on his or her impression of its relevance to each item (e.g. on a scale of 1 to 5).

Development of the Q-matrix

- When they have finished, they should consider possible refinement of the Q-matrix.
- The experts ratings are collected and aggregated.
- Next, we consider how this information is used.

Multiple Rater Results

- Use the results to determine the most likely Q-matrix.
- Use an iterative procedure asking raters for justifications if they deviate from the most common conclusions.
- Use rater scores to determine probabilities each skill is required for each item.

Multiple Rater Summary

- In general, Multiple Raters is no different than a single rater, only now more information is obtained.
- This allows for more options of how one determines the final Q-matrix to be used.
- Summary of raters conclusions can range from very simple (e.g., the most common Q) to more complicated statistical procedures in aggregating the ratings.

Refinement based on Item Parameters

- Finally, we get to the last of the basic methods for Q-matrix construction.
- Even if a lot of care has been placed in determining an initial Q-matrix, it is possible that the Q-matrix is incorrect.
 - Think in terms of a confirmatory factor analysis.

Additional Indicators of Q-matrix Misspecification

- Slow convergence/lack of convergence if using an MCMC.
- If many of the class probabilities are very low.
 - In many models this can be detected using attribute associations.
- Poorly fit test score distribution.

Refinement based on Item Parameters

- In any event, these are simply indicators of possible problems.
 - There are other reasons for item to perform poorly.
- Given these results one should:
 - Revisit any “trouble” items .
 - Consider if the entries of the Q-matrix should be changed.
 - Look for theoretically supported reasons.

Possible Problems to Avoid

- After the Q-matrix has been developed there are certain considerations that must be made.
 - Have I tried to measure too many attributes?
 - Are there attributes that are very similar?
 - Are some attributes required by most or all items?
 - Have I specified too many attributes on a single item?

Too Many Attributes?

	Skill 1	Skill 2	...	Skill 20
1				
.				
30				

- Must consider reducing the number of attributes.
- You will not have enough information to estimate all of these Attributes.
- Attributes are too finely defined.

Similar Attributes?

	Skill 1	Skill 2	Skill 3
1	1	1	0
2	1	1	1
3	1	1	0
4	0	0	1
...
20	1	1	0

- In this example Skill 1 and Skill 2 are measured by most of the same items.

- It will be difficult to determine whether items are being missed because of lacking Skill 1, Skill 2, or both (“blocking”).

- Consider combining the two skills or selecting one of the two skills for each item.

Attributes Required by Many Items?

- In this case, a single attribute is measured by every item.
- The item alone will determine whether you will have a high or low score.
- Also, if you lack this attribute it may be difficult to determine mastery of the other attribute.
- Consider breaking the attribute into two attribute (difficult level of attribute and easy level of attribute).

	Skill 1	Skill 2	Skill 3
1	1	0	0
2	1	0	1
3	1	1	0
4	1	1	1
20	1	0	1

Too Many Attributes for an Item

- In some cases, it may be tempting to specify several (more than 4 or 5) attributes for an item.
- This can begin to cause problems if it is frequent.
- Re-evaluate your skills.
 - Are they too fine grade?
 - Can the meaning of each attribute be broadened so that fewer defined attributes are required on each item?

Summary

- Generally speaking, whether you have a set of experts or it is only you.
 - You should determine the attribute.
 - Determine which items require which attributes.
 - Consider possible refinements of the Q-matrix.
 - Fit a preliminary model and evaluate item parameters.
 - Consider refinements and fit model (repeat).

Conclusions

- There is no substitution for a well defined theory and well defined attributes.
- Given these attributes multiple raters can provide their opinion of a possible Q-matrix and refine it to a Q-matrix that will be used in a basic analysis.
- In many cases, simple inspection of the results from the estimation algorithm may provide additional insight as to a reasonable Q-matrix.