Derivation of Metric Scales from Ordinal Data with Guttman-Goode's Scaling

Víthor Rosa Franco (vithorfranco@gmail.com)
Postgraduate Program in Psychology, Department of Psychology, São Francisco University, Campinas Campus, Rua Waldemar César da Silveira, Jardim Cura D'Ars, 13045510 - Campinas, SP, BRA

Keywords: Measurement theory; Guttman scale; Goode scaling; Psychometrics

Introduction

Psychometric modeling usually assumes that the observed behavior is caused by a set of metric latent variables. For instance, the Rasch model, one of the most traditional models from the Item Response Theory (Embretson & Reise, 2013), assumes that the probability of getting an answer right (or saying yes, or agreeing to the statement, or simply that $X = 1$) is equal to a logistic transformation of an additive interaction between the respondent's true score θ and the item's difficulty b. Formally, the model is represented as:

$$P(X = 1) = \frac{1}{1 + \exp(\theta - b)}. \quad (1)$$

This type of model is used mainly to estimate and develop interval measures for θ and b. Perline et al. (1979) argued that this is possible because the Rasch model is a stochastic variant of the Additive Conjoint Measurement Theory (Luce & Tukey, 1964). The Additive Conjoint Measurement Theory is a formal theory of continuous quantities which allows for the derivation of interval scales from ordinal data, as long as some empirical relations are observed.

However, some authors have disputed this view that the Rasch model is a stochastic variant of the Additive Conjoint Measurement Theory (e.g., Michell, 2008). More specifically, it has been argued that if the Rasch model is a probabilistic version of the Guttman scale (Guttman, 1944), which allows only for θ and b to be measured in the same ordinal scale, then the Rasch model provides an interval measure only because it is modeling response error. This apparent inconsistency is called the Rasch paradox.

On the other hand, the Rasch paradox has also been disputed (e.g., Borsboom & Zand Scholten, 2008). Regardless of whether the Rasch paradox is real or not, it would be interesting for psychometric researchers if interval, or even ratio, (i.e., metric) scales could be derived from Guttman scales without reliance on response errors. The aim of the present study is to propose a procedure that combines the probabilistic Guttman scaling (Proctor, 1970) with Goode's method (Coombs, 1964) to obtain metric scales from dichotomous psychometric data. We call this procedure the Guttman-Goode's Scaling (GGS).

Guttman-Goode's Scaling

The GGS procedure combines two methods for deriving interval and ratio scales from psychometric data. The first is the probabilistic Guttman scaling (Proctor, 1970). Guttman scales assume that the respondent will answer $X = 1$ if and only if $\theta > b$. Otherwise, the respondent will answer $X = 0$. If this condition is exactly met, the matrix (or Guttman scalogram) of response patterns averaged by sum scores (for an instrument with 5 items) will be equal to the matrix represented in Table 1. It is possible to see that all cells are equal to 0 or 1, representing that all individuals with a specific ordinal θ level answered to the items in the same way (e.g., a correct answer, 1, or an incorrect answer, 0).

Table 1: Perfect Guttman scalogram of response patterns averaged by sum scores for an instrument with 5 items.

θ level	**Item 1**	**Item 2**	**Item 3**	**Item 4**	**Item 5**
0	0	0	0	0	0
1	1	0	0	0	0
2	1	1	0	0	0
3	1	1	1	0	0
4	1	1	1	1	0
5	1	1	1	1	1

However, real data seldomly result in a perfect Guttman scalogram, as represented in Table 2, which was calculated from a toy dataset with actual answers from respondents. Therefore, traditional Guttman scaling cannot be applied to this type of scenarios. The probabilistic Guttman scaling, then, estimates the probability of both the order of the items as well as the ordinal θ level by assuming that only the $v + 1$, where v is the number of items, levels of θ are distinguishable. This differs from the Rasch model, for instance, that allows for more than $v + 1$ values of θ to be estimated.

Table 2: Empirical Guttman scalogram of response patterns averaged by sum scores for an instrument with 5 items.

θ level	**Item 1**	**Item 2**	**Item 3**	**Item 4**	**Item 5**
0	0	.253	.126	.149	.092
1	1	0	.243	.216	.027
2	1	1	0	.285	.228
3	.727	1	1	0	0
4	.760	.880	1	1	0
5	.836	.873	.945	.727	1

After estimating the ordinal θ level and the order of the items, our procedure uses Goode's method to analytically derive an interval scale from an ordered metric scale of respondents and items. The ordered metric scale is a scale derived from the data dependent on empirical relations regarding the distance between a respondent ordinal θ level and two items' orders (i.e., b_1 and b_2), formally stated as:

$$\overline{\theta I_1} > \overline{\theta I_2}. \quad (2)$$

Equation 2 is simply an order relation of order relations (i.e., order relation of distances). Research in measurement theory has shown that ordered metric scales impose constraints on the uniqueness of numerical representations that can be derived from simple ordinal data (Coombs, 1964, p. 359). We propose that, for Guttman scales, the order relation of distances can be found by taking the average of the rows and the complement (i.e., 1 minus) of the average of the columns of the empirical Guttman scalogram. Because we know that the ordinal θ level represented with 0 is the smallest possible value and that the ordinal θ level represented with 5 is the largest possible value, we can use the aforementioned averages to create a dominance matrix, such as the one represented in Table 3. This table is created based on the distance between each point (i.e., a θ level or an item) and the ordinal θ level represented with 0.

Table 3: Dominance matrix of θ levels and item orders.

	I4	**θ_1**	**I3**	**θ_2**	**I1**	**θ_3**	**I2**	**θ_4**	**I5**	**θ_5**
I4	0	1	1	1	1	1	1	1	1	1
θ_1	0	0	1	1	1	1	1	1	1	1
I3	0	0	0	1	1	1	1	1	1	1
θ_2	0	0	0	0	0	1	1	1	1	1
I1	0	0	0	1	0	1	1	1	1	1
θ_3	0	0	0	0	0	0	1	1	1	1
I2	0	0	0	0	0	0	0	1	1	1
θ_4	0	0	0	0	0	0	0	0	1	1
I5	0	0	0	0	0	0	0	0	0	1
θ_5	0	0	0	0	0	0	0	0	0	0

Note. I is an acronym for "Item".

For the next step of Goode's method, one must choose the value for the smallest distance (represented as Δ_0) and then analytically derive the next distances, Δ_j, for each distance j. We adapt Goode's original equation to the current scenario and propose the following equation for calculating Δ_j:

$$\Delta_j = \Delta_0(CS(j) + 1) + CS(j), \quad (3)$$

where $CS(\cdot)$ is the sum of the column representing the distance j. For instance, CS(I4) is equal to 0 and $CS(\theta_5)$ is equal to 9. Finally, the last step involves attributing values for each point. Arbitrarily, the smallest point θ_0 may be set to 0. The other points can simply be attributed their Δ_j values, as these were calculated based on the points distance in relation to θ_0.

After analytically deriving all the scale values, which are measured in an interval level, one may wish to estimate how well this numeric approximation represents the data. One way of doing this is using a logistic or hyperbolic tangent function on the linearly transformed scale values and compare the results with the empirical Guttman scalogram. For the logistic function, we propose:

$$AM(\theta, I) = \frac{1}{1 + \exp\,(\psi_\theta - \psi_I)}, \quad (4)$$

where ψ_θ and ψ_I are, respectively, the normalized interval level measure for the ordinal θ level and for the item. For the hyperbolic tangent function, we propose:

$$DM(\theta, I) = 2\frac{1}{1 + \exp\,[-2(\varphi_\theta \varphi_I)]} - 1, \quad (5)$$

where $\varphi_l = \exp\,(\psi_l)$, as the exponential transform of the normalized interval scale values results in a ratio scale (Fishburn, 1974). Applying this procedure to the data that generated Table 2 results in a RMSE of .015 to the logistic function approach and in a RMSE of .184 to the hyperbolic tangent function approach. This result suggests that an interval representation is better than a ratio representation of the points.

Final Considerations

The GGS procedure can be used with any data following a direct response design (such as attitude or performance psychometric scales). The main advantage of the GGS procedure is that, different from Item Response Theory models, the scales are derived from ordered metric information in the data and, therefore, should be less reliant on response error. However, it should be noted that this is an initial implementation of the GGS procedure and limitations abound. For instance, Table 3 presents an intransitivity for θ_2 which is not dealt with. We also do not estimate the distances between intermediary points such as $\overline{\theta_2 I_1} > \overline{\theta_1 I_2}$, which hides an implicit assumption that a unidimensional representation is the most appropriate (Coombs, 1964). Future studies should deal with these limitations to provide more robust metric scales for psychometric data.

References

Borsboom, D., & Zand Scholten, A. (2008). The Rasch model and conjoint measurement theory from the perspective of psychometrics. *Theory & Psychology, 18*, 111–117.

Coombs, C.H. (1964) *A theory of data*. Wiley.

Embretson, S. E., & Reise, S. P. (2013). *Item response theory for psychologists*. Lawrence Erlbaum.

Fishburn, P. C. (1974). von Neumann-Morgenstera Utility Functions on Two Attributes. *Operations Research*, *22*(1), 35-45.

Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review*, *9*, 139–150.

Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, *1*(1), 1-27.

Michell, J. (2008). Conjoint measurement and the Rasch paradox: A response to Kyngdon. *Theory & Psychology*, *18*(1), 119-124.

Perline, R., Wright, B. D., & Wainer, H. (1979). The Rasch model as additive conjoint measurement. *Applied Psychological Measurement*, *3*(2), 237-255.

Proctor, C. H. (1970). A probabilistic formulation and statistical analysis for Guttman scaling. *Psychometrika*, *35*, 73-78.