GMX: Extended Graphical Model Checks

A Versatile Replacement of the plotGOF() Function of eRm

Rainer W. Alexandrowicz[1]

Abstract

The article introduces the R-package `GMX`, which extends the standard graphical model check of the `eRm` package. It supports the Rasch model, the PCM, and the RSM providing multiple group splits and options for selecting items, split groups, or specific parameters. Along with several graphical features, the package may prove useful for psychometric analyses extending the capabilities of `eRm`. It is freely available at https://osf.io/2ryd8.

Keywords: Rasch model, graphical model check, conditional maximum likelihood, multi-group split, R-package

[1]*Correspondence concerning this article should be addressed to:* rainer.alexandrowicz@aau.at
Rainer W. Alexandrowicz, University of Klagenfurt, Institute of Psychology, Methods Department, Universitaetsstrasse 67–69, 9020 Klagenfurt, Austria;

Introduction

The `eRm` package (Mair, Hatzinger, & Maier, 2020) has gained enormous importance in the IRT world, because it is one of the few packages supporting the conditional maximum likelihood (CML) estimation principle for dichotomous and polytomous item response models (see Mair & Hatzinger, 2007). This estimation principle excels the more frequently used marginal maximum likelihood as it conditions on the sufficient statistics of the incidental parameters thus avoiding any distributional assumptions regarding the person parameters for item parameter estimation (Baker & Kim, 2004; Pawitan, 2001). It comes closest to the ideas formulated by Rasch (cf. Rasch, 1966a, 1966b). As a downside we face estimation problems when response categories are missing in the data set. For a detailed introduction and discussion see e. g., de Ayala (2022), Baker and Kim (2004), Fischer and Molenaar (1995), or van der Linden (2016a, 2016b). The CML approach is applicable to the dichotomous Rasch model (RM; Rasch, 1960), the Partial Credit Model (PCM; Masters, 1982), the Rating Scale Model (RSM; Andrich, 1978), and several of its direct derivatives (like the linear logistic extensions, for example; see Fischer, 1973, 1983, 1995; Kubinger, 2008).

Next to its advantages for parameter estimation, the CML approach further allows for a rigid assessment of model adequacy, most prominently in form of the Conditional Likelihood Ratio Test (cLRT; Andersen, 1973; see also Alexandrowicz & Draxler, 2014). Basically, this test splits the data set into two or more person subgroups and estimates the item parameters for the entire sample and separately for each group. If the model holds, the sum of the conditional log-likelihoods across split samples equals but for random variation the conditional log-likelihood of the entire data. This test is frequently accompanied by a graphical check, which can, in principle, be traced back to Rasch (1960). This check simply draws a scatterplot juxtaposing the item parameter estimates from two split groups. Model-conforming items exhibit similar estimates (if model identification is attained by constraining parameters to sum zero; cf. Andersen, 1973, p. 38; Baker & Kim, 2004, p. 135, or Molenaar, 1995, p. 40). These items will be located in the vicinity of the 45°-line (diagonal) through the diagram, whereas items not conforming to the model assumptions will exhibit larger distances to the diagonal. The advantage of this quick check is that deviating item parameters (i. e., those with markedly unequal estimates across split samples) are easily identifiable (see Kubinger, Litzenberger, & Mrakotsky, 2006, Fig. 1 there, for an example). Additionally, confidence ellipses can add an inferential note to the conclusions drawn from this diagram (ellipses not touching the 45°-line differ significantly between the two subgroups, but alpha-inflation has to be taken into account). Such a diagram is available in the `eRm` package with the `plotGOF()` function, however, only for two split groups (see also Debelak, Strobl, & Zeigenfuse, 2022, p. 62).

One particularity of `plotGOF()` deserves some attention: This function draws what is termed `"betas"` in `eRm`. These are cumulative item easiness parameters (denoted β^* here), because the sophisticated implementation of `eRm` reformulates all supported models as an LPCM (the Linear PCM; see Fischer & Ponocny, 1994). While the case is clear for the dichotomous model, users might be tempted to misinterpret the points' coordinates as the split samples' threshold parameters in the polytomous case. Therefore, the transformation of these β^* into the commonly used thresholds (denoted τ here) shall be illustrated by means of a mini example of a four-categorical item with three threshold parameters and the following transformation matrix:

	β_1^*	β_2^*	β_3^*
τ_1	1	0	0
τ_2	-1	1	0
τ_3	0	-1	1

Applying the transformations, we yield

$$\beta_1^* = \tau_1 \tag{1a}$$

$$\beta_2^* = \beta_1^* + \tau_2 = \tau_1 + \tau_2 \tag{1b}$$

$$\beta_3^* = \beta_2^* + \tau_3 = \tau_1 + \tau_2 + \tau_3 \tag{1c}$$

or, correspondingly:

$$\tau_1 = \beta_1^* \tag{2a}$$

$$\tau_2 = -\beta_1^* + \beta_2^* = \beta_2^* - \beta_1^* \tag{2b}$$

$$\tau_3 = -\beta_2^* + \beta_3^* = \beta_3^* - \beta_2^* \tag{2c}$$

Unfortunately, the `plotGOF()` function of `eRm` only supports plotting the β^* parameters for a two-group split. The GMX package introduced in the next chapter overcomes these limitations in several respects.

The GMX Package

The `GMX` package provides a versatile function `gmx()`, which allows for plotting several variants of the graphical model check: First, it extends to an arbitrary number of split groups (which could be selected, see below), second, it allows for plotting the beta parameters (as does `plotGOF()`; default), the threshold parameters (for PCM/RSM applications), or the person parameters (RM, PCM, and RSM) of the split groups

obtained by partitioning the sample "horizontally", i.e., according to the score r or any other criterion of substantive interest. Third, users may select subsets of items or parameters to plot, and, finally, GMX provides further graphical options currently not available in the `plotGOF()` function.

The package requires `eRm` and uses the output of the `LRtest()` function. By default, it extracts the subgroup parameters and draws the graphical model checks of all pairs of split groups in a diagram matrix. The `type="betas"` option allows for selecting the estimates of $\hat{\beta}$ (in `eRm`-notation, see Equation set 2 above; default), the `type="thresholds"` option draws the threshold estimates $\hat{\tau}$ (according to Equation set 1), and the `type="perspar"` option draws the person parameter estimates $\hat{\theta}_r$ for all scores r. The `extwin=` option (default: `TRUE`) opens a new graphics window for plotting – a feature particularly interesting for users of RStudio, as they gain control over the plot's size and proportions rather than having to rely on the internal graphics pane size and proportion. The diagram size option (`ds=`) takes the size of the subplots (in inches) and the options `title=` and `annot=` allow for fine-tuning plot titles and annotations, respectively. The option `col=` allows for either specifying a color vector to be used for the parameter dots or takes one of the two keywords `"items"` or `"thresholds"`: The former selects automatically the same color for all parameters of each item and the latter uses the same color for all first thresholds/betas across the items, all second thresholds/betas, and so on. (see Listings 2 and 3 and Figures 4 and 5 below). The number of parameters per item may vary, as permissible for the PCM.

The `EQlims=` option chooses equal axis limits for all subplots to foster comparisons across pairs of split groups. It overrides manually set `xlim`/`ylim` and has to be set to `FALSE` if the explicit ranges are to be used. The `glabs="auto"` (default) option extracts the split group names from the output object of `LRtest()`[1] or it takes a group name vector. The `tlabs=` option determines the point labeling with the same options as `plotGOF()`. Additionally, only subsets of items (`isel=`) or parameters (`psel=`) can be shown (with `isel=` taking precedence over `psel=` if both are specified). Note that the `isel=` and `psel=` options refer to the item numbers as available in the result object of `LRtest()`. These may differ from the item numbers of the original data set, because `LRtest()` drops items with missing categories (with a warning). The option `gsel=` selects split groups to be drawn. Additionally, the order, in which the groups are specified, also controls how groups are assigned to the horizontal and the vertical axis.

The `confell=` option selects whether or not to draw confidence ellipses, the color of which can be set with the `confcol=` option (accepting a color vector) and the level

---

[1]Note that due to a bug `LRtest()` requires the split groups to be defined as a `factor` (see https:// r-forge.r-project.org/tracker/?atid=363&group_id=80, bug # 6759).

of an error of the first kind (α) with `conferr=` (note the difference to `plotGOF()`, which expects $1 - \alpha$).

Working Examples

The examples use the `pcmdat2` data set from the `eRm` package. Listing 1 illustrates the basic call and the three diagram types:

Listing 1: Example calls using the example data set of `eRm`.

```
1    library(eRm)
2    data(pcmdat2)
3
4    mod = PCM(pcmdat2)
5    grp = rep(1:3, each=100)
6    tst = LRtest(mod, splitcr=grp)
7
8    library(GMX)
9    gmx(tst)                          # betas (= default)
10   gmx(tst,type="t")                 # thresholds
11   gmx(tst,type="p",EQlims=FALSE,    # person parameters
12       xlim=c(-3,3),ylim=c(-3,3))
```

Note that `GMX` does not require the `eRm` package, it was used in Listing 1 to obtain the result object of the LR-test. Figures 1–3 show the diagrams obtained with these commands.
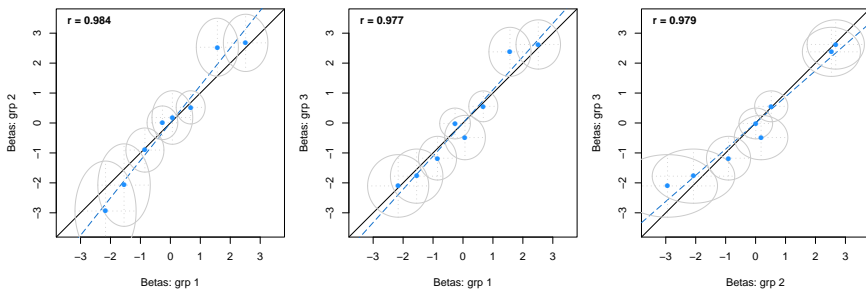

Figure 1: Output from `gmx(tst)`. Notes: The solid diagonal line is the 45°-line and the blue dashed line is the regression line of the vertical axis on the horizontal axis; In the top left corner, we find the correlation coefficient of the two sets of estimates.

Figure 1 shows the typical graphical model checks as will be obained with `plotGOF()` of `eRm` (yet with different colors). By default, it draws the (negative) `"beta"` estimates, i. e., the cumulative thresholds $-\hat{\beta}^*$, as does `plotGOF()`.
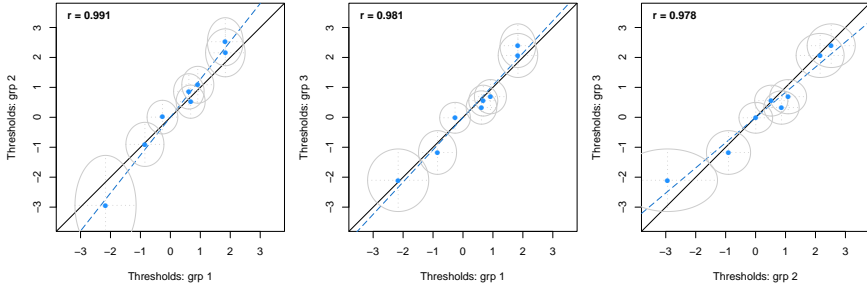
Figure 2: Output from `gmx(tst,type="t")`.

Figure 2 shows the decumulated threshold estimates, i. e., the values obtained after applying Equation set 1. These are the commonly referred parameters. Note that the plot (and the conclusions drawn thereof) may change drastically when switching between β and τ.
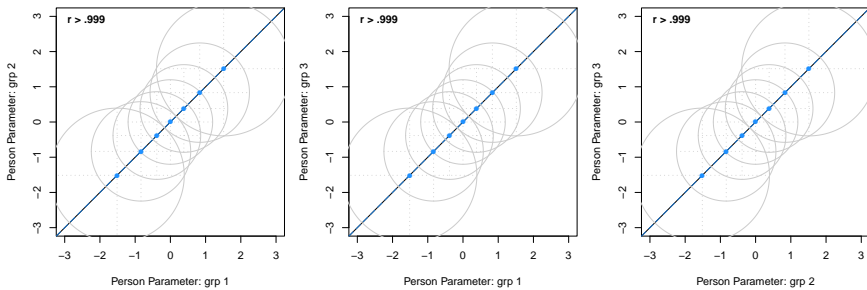
Figure 3: Output from `gmx(tst,type="p"`, `xlim=c(-3,3)`, `ylim=c(-3,3)`, `EQlims=FALSE)`. The axis limits were manually adapted to make this diagram comparable to Figures 1 and 2.

Figure 3 shows two interesting phenomena: First, we find much larger confidence ellipses for the person parameters compared to the item parameters. These are necessarily larger, as we dispose of much less information to obtain the person parameters (the number of items is usually much smaller than the number of observations). Second, the person

parameter estimates exhibit no visible differences across subgroups, which is also almost necessarily the case yet less obvious (for an explanation see Alexandrowicz, 2016).

Next, two graphical features shall be illustrated. The `col=` vector takes a set of colors used to highlight the points. Or, it takes one of the keywords `"items"` or `"thresholds"` to color the points using the currently defined color palette. The former can be used to mark items (Listing 2):

Listing 2: Distinguishing items with `gmx()`.

```
1    palette(c("darkorange2","deepskyblue1","green3","steelblue"))
2    gmx(tst,col="items",type="t",cex=2,tlab="nam")
```

Figure 4 shows all thresholds belonging to one item with the same color (i. e., all thresholds of item 1 in `darkorange2`, all thresholds of item 2 in `deepskyblue1`, and so on):
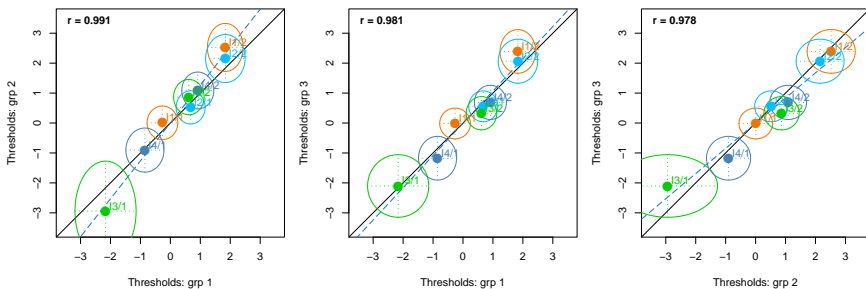

Figure 4: Highlighting items with the `col="items"` option.

Similarly, all threshold estimates $\hat{\tau}_1/\hat{\tau}_2/\ldots$ (or $\hat{\beta}_1/\hat{\beta}_2/\ldots$) could be highlighted with such a color vector (Listing 3).

Listing 3: Distinguishing thresholds with `gmx()`.

```
1    palette(c("steelblue1","orangered2"))
2    gmx(tst,col="thresholds",type="t",cex=2,tlab="nam")
```

Listing 3 yields Figure 5 with the $\hat{\tau}_1$ of all items colored in `steelblue1` and the $\hat{\tau}_2$ of all items colored in `orangered2`:
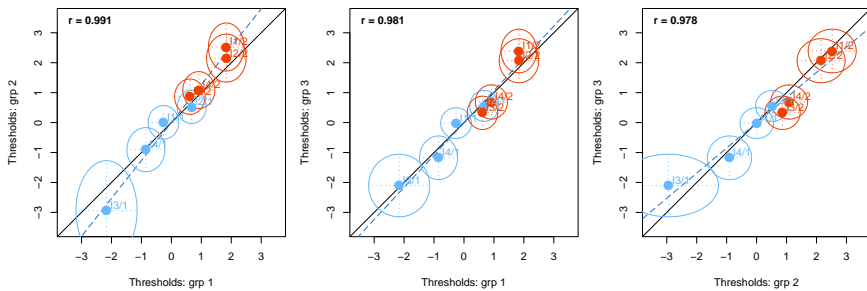
Figure 5: Color coding thresholds with the `col="thresholds"` option.

Figure 5 allows for visualizing effects appearing for specific categories (which might prove useful in cases of threshold disordering; see Andrich, 2013).

Discussion

This article introduced the GMX package, which provides for extended graphical model checks of the RM, the PCM, or the RSM based on the CML estimation method. It extends the original functionality of the `plotGOF()` function of the `eRm` package by supporting a multi-group comparison and providing various further graphical options. For example, we can inspect only a subset of items (with the `isel=` option) exhibiting differences in certain split groups and this way gain insight about their malfunctioning. Furthermore, we can scrutinize the response format by highlighting the thresholds across items (with the `col="thresholds"` option). The various possibilities of the GMX package allow for detailed analyses of items and categories thus fostering in-depth understanding of possible model violations. The package is freely available at https://osf.io/2ryd8.

---

[2]See https://r-forge.r-project.org/tracker/?atid=363&group_id=80, bug # 6759. Note that another bug concerning the person parameter estimation may occur under rare circumstances in connection with small samples and missing response categories (see bug # 6770).

References

Alexandrowicz, R. W. (2016). The Rasch-Model From an Individual's Perspective: The Item Rank Plot and the Compensation Test. *Journal of Person-Oriented Research*, *2(1–2)*, 87–101. doi: 10.17505/jpor.2016.09

Alexandrowicz, R. W., & Draxler, C. (2014). The Andersen Likelihood Ratio Test: Sample Size Requirements and Bootstrap Algorithms. *Journal of Statistical Distributions and Applications*, *3*. doi: 10.1186/s40488-016-0039-y

Andersen, E. B. (1973). A goodness of fit test for the Rasch model. *Psychometrika*, *38*, 123–140.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*, 561–573.

Andrich, D. (2013). An Expanded Derivation of the Threshold Structure of the Polytomous Rasch Model That Dispels Any "Threshold Disorder Controversy". *Educational and Psychological Measurement*, *73*, 78–124. doi: 10.1177/0013164412450877

Baker, F. B., & Kim, S.-H. (2004). *Item Response Theory. Parameter Estimation Techniques.* NY: Marcel Dekker.

de Ayala, R. J. (2022). *The Theory and Practice of Item Response Theory* (2nd ed.). The Guilford Press.

Debelak, R., Strobl, C., & Zeigenfuse, M. D. (2022). *An Introduction to the Rasch Model with Examples in R*. Chapman & Hall.

Fischer, G. H. (1973). The Linear Logistic Test Model as an Instrument in Educational Research. *Acta Psychologica*, *37*, 359–374.

Fischer, G. H. (1983). Logistic Latent Trait Models with Linear Constraints. *Psychometrika*, *48*, 3–26.

Fischer, G. H. (1995). The Linear Logistic Test Model. In G. H. Fischer & I. W. Molenaar (Eds.), *Rasch Models. Foundations, Recent Developments, and Applications* (pp. 131–155). Springer.

Fischer, G. H., & Molenaar, I. W. (Eds.). (1995). Springer.

Fischer, G. H., & Ponocny, I. (1994). An extension of the partial credit model with an application to the measurement of change. *Psychometrika*, *59*, 177–192.

Kubinger, K. D. (2008). On the revival of the Rasch model-based LLTM: From constructing tests using item generating rules to measuring item administration effects. *Psychology Science Quarterly*, *50*, 311–327.

Kubinger, K. D., Litzenberger, M., & Mrakotsky, C. (2006). Practised intelligence testing based on a modern test conceptualization and its reference to the common intelligence theories. *Learning and Individual Differences*, *16*, 175–193. doi: 10.1016/j.lindif.2005.08.001

Mair, P., & Hatzinger, R. (2007). CML based estimation of extended Rasch models with the eRm package in R. *Psychology Science*, *49*, 26–43.

Mair, P., Hatzinger, R., & Maier, M. J. (2020). eRm: Extended Rasch Modeling [Computer software manual]. Retrieved from https://cran.r-project.org/package=eRm (1.0-2)

Masters, G. N. (1982). A Rasch Model for Partial Credit Scoring. *Psychometrika*, *47*, 149–174.

Molenaar, I. W. (1995). Estimation of Item Parameters. In G. H. Fischer & I. W. Molenaar (Eds.), *Rasch Models. Foundations, Recent Developments, and Applications* (pp. 39–51). Springer.

Pawitan, Y. (2001). *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford: Clarendon.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Danmarks Pædagogiske Institut.

Rasch, G. (1966a). An Individualistic Approach to Item Analysis. In P. F. Lazarsfeld & N. W. Henry (Eds.), *Readings in Mathematical Social Science* (pp. 89–107). Cambridge, MA: The M.I.T. Press.

Rasch, G. (1966b). *An informal report on the present state of a theory of objectivity in comparisons.* Retrieved 2022-08-03, from https://www.rasch.org/

memo1966.pdf

van der Linden, W. J. (Ed.). (2016a). *Handbook of Item Response Theory. Volume One: Models*. Taylor & Francis/CRC.

van der Linden, W. J. (Ed.). (2016b). *Handbook of Item Response Theory. Volume Two: Statistical Tools*. Taylor & Francis/CRC.