The Pairwise-Method for Parameter Estimation within the ordinal Rasch-Model

Jörg–Henrik Heine[1,2]

[1]Technical University of Munich [2]Centre for International Student Assessment (ZIB)

Workshop

Technische Universität München

ZENTRUM FÜR INTERNATIONALE BILDUNGSVERGLEICHSSTUDIEN

Background

Three widespread standard estimation methods for item parameters (and person parameters) in the Rasch Model

1. **Condtional Maximum Likelihood – CML**:
   - Seperat estimation of item and person parameters
   - Independent from score distribution
   - might be problematic with higher proportion of missing values (e.g. Booklet–Design)

Background

Three widespread standard estimation methods for item parameters (and person parameters) in the Rasch Model

1. **Condtional Maximum Likelihood – CML**:
   - Seperat estimation of item and person parameters
   - Independent from score distribution
   - might be problematic with higher proportion of missing values (e.g. Booklet–Design)
2. **Maximum Marginal Likelihood – MML**:
   - Seperat estimation of item and person parameters
   - No problems when using Booklet–Desigs (missing by design)
   - might be problematic if assumption of normal distribution is violated (e.g. small data sets, selective samples)

Background

Three widespread standard estimation methods for item parameters (and person parameters) in the Rasch Model

1. **Condtional Maximum Likelihood – CML**:
   - Seperat estimation of item and person parameters
   - Independent from score distribution
   - might be problematic with higher proportion of missing values (e.g. Booklet–Design)
2. **Maximum Marginal Likelihood – MML**:
   - Seperat estimation of item and person parameters
   - No problems when using Booklet–Desigs (missing by design)
   - might be problematic if assumption of normal distribution is violated (e.g. small data sets, selective samples)
3. **Joint Maximum Likelihood – JML**:
   - Simultaneous estimation of item and person parameters
   - Person parameter estimation always less precise than item parameter estimation

Background

Three widespread standard estimation methods for item parameters (and person parameters) in the Rasch Model

1. **Condtional Maximum Likelihood – CML**:
    - Seperat estimation of item and person parameters
    - Independent from score distribution
    - might be problematic with higher proportion of missing values (e.g. Booklet–Design)
2. **Maximum Marginal Likelihood – MML**:
    - Seperat estimation of item and person parameters
    - No problems when using Booklet–Desigs (missing by design)
    - might be problematic if assumption of normal distribution is violated (e.g. small data sets, selective samples)
3. **Joint Maximum Likelihood – JML**:
    - Simultaneous estimation of item and person parameters
    - Person parameter estimation always less precise than item parameter estimation

Short in between summary – CML, MML, JML

In general using multistage (polytomous) response formats in combination with small sample size will result in low frequencies in response categories. This in turn is often accompanied with estimation problems.

- „*These conditional methods of parameter estimation for items with more than two ordered categories involve sufficient statistics for the threshold parameters that involve directly the frequencies of responses in the adjacent categories. This is also a feature of other methods of estimation, for example, the marginal maximum likelihood (Bock and Aitkin, 1981) and the joint maximum likelihood (Wright and Douglas, 1977) methods. As a result, and in the case of some dislocation of persons and items which can result in* **some categories** *of* **some items** *having* **low frequencies**, *the* **estimates of corresponding thresholds will be unstable.** "(Andrich & Luo, 2003)

The „Pairwise“ Method as an Alternative

Explicit calculation of item parameters using the Method of „paarwise item comparison “ and then in turn estimation of person parameters using ML based methodology

1. **Central References to Pairwise Method**:
   - Early theoretical Discussion by Georg Rasch with Gustaf Leunbach (vgl. Leunbach, 1961), within the framework of his mathematical derivation of the Rasch Modell (Rasch, 1960, p. 172).
   - Further development for use in item calibration within item banks (Choppin, 1968, 1983)
   - Theoretical considerations about a generalization for polytomous items (Wright & Masters, 1982; Garner & Engelhard Jr, 2002)

The „Pairwise“ Method as an Alternative

Explicit calculation of item parameters using the Method of „paarwise item comparison “ and then in turn estimation of person parameters using ML based methodology

1. **Central References to Pairwise Method**:
   - Early theoretical Discussion by Georg Rasch with Gustaf Leunbach (vgl. Leunbach, 1961), within the framework of his mathematical derivation of the Rasch Modell (Rasch, 1960, p. 172).
   - Further development for use in item calibration within item banks (Choppin, 1968, 1983)
   - Theoretical considerations about a generalization for polytomous items (Wright & Masters, 1982; Garner & Engelhard Jr, 2002)
2. **Principle**:
   - Counting of frequencies solving item j given that item i was not solved.
   - Forming a reciprocal frequencies matrix
   - Logarithmize the matrix entries and taking the row means
   - Detailed presentation in the following example for dichtomous items - see also (Heine et al., 2011; Heine & Tarnai, 2015).

The „Pairwise“ Method as an Alternative

Explicit calculation of item parameters using the Method of „paarwise item comparison “ and then in turn estimation of person parameters using ML based methodology

1. **Central References to Pairwise Method**:
   - Early theoretical Discussion by Georg Rasch with Gustaf Leunbach (vgl. Leunbach, 1961), within the framework of his mathematical derivation of the Rasch Modell (Rasch, 1960, p. 172).
   - Further development for use in item calibration within item banks (Choppin, 1968, 1983)
   - Theoretical considerations about a generalization for polytomous items (Wright & Masters, 1982; Garner & Engelhard Jr, 2002)
2. **Principle**:
   - Counting of frequencies solving item j given that item i was not solved.
   - Forming a reciprocal frequencies matrix
   - Logarithmize the matrix entries and taking the row means
   - Detailed presentation in the following example for dichtomous items - see also (Heine et al., 2011; Heine & Tarnai, 2015).

The „Pairwise“ method – simple example – data

As a simple example suppose **M** is a data matrix containing the answers of eight persons on four dichotomous items. This matrix **M** might look like the one depicted below.

	Item1	Item2	Item3	Item4
Person 1	1	1	1	0
Person 2	0	1	0	1
Person 3	1	0	0	1
Person 4	1	0	0	0
Person 5		1	1	
Person 6	1	1		1
Person 7		0	0	
Person 8	1		0	0

Please note that there are some missing data points in this example data which corresponds to a missing proportion of 18.75%.

R–Code:

```
load("simple.RData")
simple
```

The „Pairwise“ method – simple example – step 1

The first step when applying the pairwise comparison algorithm according to Choppin (1968), is to count the conditional response frequencies for each item. That is counting the number of persons who got item i right under the condition of having answered wrong to item j for every item $k (i \neq j)$. This results in a symmetrical pairwise comparison matrix **C** with entries f_{ij} above diag. and f_{ji} below diag. which is depicted below based on the example data on the previous slide.

$$C_{f_{i,j}} = \begin{bmatrix} 0 & 2 & 3 & 3 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 2 & 0 \end{bmatrix}$$

R–Code:

```
library(pairwise)
C <- pair(simple,pot=F,ccf = T,zerocor=F)
C
```

The „Pairwise" method – simple example – additional note

It should be noted that the matrix **C** has an analogy to the Burt matrix **B** (Burt, 1950), which is used in correspondence analysis (e.g. Blasius, 2011, for an overview) computed from the (item category) indicator (super) matrix **Z**.

$$Z = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ \mathbf{0} & \mathbf{0} & 0 & 1 & 0 & 1 & \mathbf{0} & \mathbf{0} \\ 0 & 1 & 0 & 1 & \mathbf{0} & \mathbf{0} & 0 & 1 \\ \mathbf{0} & \mathbf{0} & 1 & 0 & 1 & 0 & \mathbf{0} & \mathbf{0} \\ 0 & 1 & \mathbf{0} & \mathbf{0} & 1 & 0 & 1 & 0 \end{bmatrix} \qquad B_{f_{i,j}} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 5 & 2 & 2 & 3 & 1 & 3 & 2 \\ 0 & 2 & 3 & 0 & 3 & 0 & 1 & 1 \\ 1 & 2 & 0 & 4 & 1 & 2 & 1 & 2 \\ 1 & 3 & 3 & 1 & 5 & 0 & 2 & 2 \\ 0 & 1 & 0 & 2 & 0 & 2 & 1 & 0 \\ 0 & 3 & 1 & 1 & 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 2 & 2 & 0 & 0 & 3 \end{bmatrix}$$

Multiplying the transposed indicator matrix **Z** by itself results in the Burt matrix (depicted above on the right side).
That is $B = Z^T Z$.

The „Pairwise“ method – simple example – additional note

Contrary to the Burt matrix **B** (computed based on **Z** from **M**) some columns and rows in the matrix **C** used for itemparameter calculation are skipped.

$$B_{f_{i,j}} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ \mathbf{0} & 5 & \mathbf{2} & 2 & \mathbf{3} & 1 & \mathbf{3} & 2 \\ 0 & 2 & 3 & 0 & 3 & 0 & 1 & 1 \\ \mathbf{1} & 2 & \mathbf{0} & 4 & \mathbf{1} & 2 & \mathbf{1} & 2 \\ 1 & 3 & 3 & 1 & 5 & 0 & 2 & 2 \\ \mathbf{0} & 1 & \mathbf{0} & 2 & \mathbf{0} & 2 & \mathbf{1} & 0 \\ 0 & 3 & 1 & 1 & 2 & 1 & 3 & 0 \\ \mathbf{1} & 2 & \mathbf{1} & 2 & \mathbf{2} & 0 & \mathbf{0} & 3 \end{bmatrix} \quad C_{f_{i,j}} = \begin{bmatrix} 0 & 2 & 3 & 3 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 2 & 0 \end{bmatrix}$$

In reference to the simple example, **C** could be derived out of the full Burt matrix **B** by skipping rows 1, 3, 5 and 7 and columns 2, 4, 6 and 8.

R–Excursus: The Burt matrix on the „simple“ example

```
load("simple.RData"); library(pairwise); simple

### install and load the anacor package
require(anacor)

### this returns indicator (super) matrix Z for simple
Z <- expandFrame(simple,returnFrame = F);Z

### this computes B by multiplying transposed Z by itself
B <- t(Z)%*%Z ;B

### skiping some columns and rows ...
B[-c(1, 3, 5, 7) , -c(2, 4, 6, 8)]

### ... yields in the same result as:
pair(simple,pot=F,ccf = T,zerocor=F)
```

The „Pairwise“ method – simple example – step 2

Naturally, the diagonal of this Matrix **C** is filled with zeros as there is no meaningful comparison of one item with itself. There are also off diagonal zeros which would cause numerical problems in the following steps if not incorporated in an adequate way.

$$C_{f_{i,j}} = \begin{bmatrix} 0 & 2 & 3 & 3 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 2 & 0 \end{bmatrix} \qquad C^3_{f_{i,j}} = \begin{bmatrix} 8 & 15 & 28 & 26 \\ 7 & 6 & 14 & 11 \\ 1 & 2 & 4 & 6 \\ 8 & 8 & 17 & 9 \end{bmatrix}$$

It was shown algebraically by Choppin (1968) that instead of using the direct comparisons in matrix **C**, powers of **C** can be used. Under the condition that all items are proper linked no off diagonal zeros should occur when using the power of 3 from **C**.

R–Code:

```
C3 <- C %*% C %*% C
C3
```

The „Pairwise“ method – simple example – step 3

By further applying the pairwise comparison algorithm a so called positive reciprocal matrix **D** with entries $^{f_{ij}}/_{f_{ji}}$ must be formed.

$$C^3_{f_{i,j}} = \begin{bmatrix} 8 & 15 & 28 & 26 \\ 7 & 6 & 14 & 11 \\ 1 & 2 & 4 & 6 \\ 8 & 8 & 17 & 9 \end{bmatrix} \qquad D = \begin{bmatrix} ^8/_8 & ^7/_{15} & ^1/_{28} & ^8/_{26} \\ ^{15}/_7 & ^6/_6 & ^2/_{14} & ^8/_{11} \\ ^{28}/_1 & ^{14}/_2 & ^4/_4 & ^{17}/_6 \\ ^{26}/_8 & ^{11}/_8 & ^6/_{17} & ^9/_9 \end{bmatrix}$$

R–Code:

```
D <- t(C3) / C3
D
```

The „Pairwise“ method – simple example – step 4

The final difficulty parameters for the four items are obtained by taking the natural logarithm of the entries in **D** and computing the row means.

R–Code:

```
Dlog <- log(D)
Dlog
rowMeans(Dlog)
```

For a comparison you can just use the `pairwise` function `pair()` to compute the Item parameters for the simple example data.

R–Code:

```
summary(pair(simple,m=2))
```

The „Pairwise“ method – polytomous items

When analyzing a data matrix comprising polytomous items with more than two answer categories the procedure of pairwise comparison follows in general the same underlying principle, that is:

- „dummy code“ each item in the data matrix forming a 0,1 coded item answer category data matrix. In fact this is the indicator (super) matrix $\mathbf{Z}$ mentioned above.
- This matrix in turn is analyzed in the manner described by means of the simple example above.

The „Pairwise“ method – polytomous items

- Form the symmetrical pairwise comparison matrix **C**. In this case $C_{f_i f_{j_c}}$ with entries $f_{i_c j_c}$ represents the number of respondents who answered to item *i* in category *c* and to item *j* in category *c-1*.
- As there is no meaningful comparison of the answer category frequencies within one item with itself, there are symmetrical square areas with size $(m-1)*(m-1)$ with *m* beeing the number of answer categories containing zero entries along the diagonal of the resulting pairwise comparison matrix **C**
- Forming the reciprocal matrix **D**
- Calculate the final (category) difficulty parameters by taking the natural logarithm of the entries in **D** and computing the row means. In case of polytomous items these parameters represent the "Thurstonian Thresholds".

The „Pairwise“ Method Summary

1. **Pros**:

The „Pairwise“ Method Summary

1. **Pros**:
   - Separate estimation of item and person parameters

The „Pairwise“ Method Summary

1. **Pros**:
   - Separate estimation of item and person parameters
   - *„the method is simpler to implement than the CML“* (Andrich & Luo, 2003)

The „Pairwise“ Method Summary

1. **Pros**:
   - Separate estimation of item and person parameters
   - „*the method is simpler to implement than the CML*“ (Andrich & Luo, 2003)
   - Consistency proven for dichotomous items (Zwinderman, 1995)

The „Pairwise“ Method Summary

1. **Pros**:
   - Separate estimation of item and person parameters
   - „*the method is simpler to implement than the CML*“ (Andrich & Luo, 2003)
   - Consistency proven for dichotomous items (Zwinderman, 1995)
   - And finaly: . . .
     „*A special advantage of PAIR is that it can be used to analyze the incomplete data matrices which result when some items are not taken by some persons in the calibration sample.*“
     (Wright & Masters, 1982, S. 69)

The R–Package {pairwise}

1. **Version History**:
   - First version (0.1.3) on CRAN scince 11. Februar 2013
   - Current Version is version 0.3.1 (Heine, 2015) (since 3. Dezember 2015 on CRAN)
   - Forthcoming preliminary versions are available under: `www.joerg-henrik-heine.de`
   - Current preliminary version is 0.3.2

1. **Version History**:
   - First version (0.1.3) on CRAN scince 11. Februar 2013
   - Current Version is version 0.3.1 (Heine, 2015) (since 3. Dezember 2015 on CRAN)
   - Forthcoming preliminary versions are available under: `www.joerg-henrik-heine.de`
   - Current preliminary version is 0.3.2
2. **Functional range**:
   - Itemparameter Calculations for dichotomous and polytomous items – even mixed response formats in one dataset.
   - Standerrors (on item parameters) – implemented by means of bootstrap or rather jack–knive procedure
   - WLE person parameter estimation following Warm (1989)
   - Graphical model test (unsing various splitting–criteria)
   - Person- and item-fit statistics following Wright und Masters (1982)
   - Point-biseriale correlation for item categories (to scale theta)
   - Supports Booklet-Designs
   - S3–methods for result objects: `summary()` and `plot()`

The R–Package {pairwise}

1. **Version History**:
   - First version (0.1.3) on CRAN scince 11. Februar 2013
   - Current Version is version 0.3.1 (Heine, 2015) (since 3. Dezember 2015 on CRAN)
   - Forthcoming preliminary versions are available under: `www.joerg-henrik-heine.de`
   - Current preliminary version is 0.3.2
2. **Functional range**:
   - Itemparameter Calculations for dichotomous and polytomous items – even mixed response formats in one dataset.
   - Standerrors (on item parameters) – implemented by means of bootstrap or rather jack–knive procedure
   - WLE person parameter estimation following Warm (1989)
   - Graphical model test (unsing various splitting–criteria)
   - Person- and item-fit statistics following Wright und Masters (1982)
   - Point-biseriale correlation for item categories (to scale theta)
   - Supports Booklet-Designs
   - S3–methods for result objects: `summary()` and `plot()`

Thank you for listening

- And now . . .

Thank you for listening

- And now ...
- Hands on 'pairwise' ...

References

Andrich, D. & Luo, G. (2003). Conditional pairwise estimation in the rasch model for ordered response categories using principal components. *Journal of Applied Measurement*, *4* (3), 205–221.

Blasius, J. (2011). Correspondence analysis. In M. Lovric (Hrsg.), *International encyclopedia of statistical science.* Berlin Heidelberg: Spektrum Akademischer Verlag.

Burt, C. (1950). The factorial analysis of qualitative data. *British Journal of Statistical Psychology*, *3* (3), 166–185.

Choppin, B. (1968). Item bank using sample-free calibration. *Nature*, *219* (5156), 870–872.

Choppin, B. (1983). *A fully Conditional Estimation Procedure for Rasch Model Parameters* (CSE Report Nr. 196).

Garner, M. & Engelhard Jr, G. (2002). An eigenvector method for estimating item parameters of the dichotomous and polytomous rasch models. *Journal of Applied Measurement*, *1* (3), 107–128.

Heine, J.-H. (2015). *pairwise: Rasch model parameters by pairwise algorithm (v 0.3.1).* Zugriff auf `http://cran.r-project.org/web/packages/pairwise/index.html` (R package version 0.3.1)

Heine, J.-H. & Tarnai, C. (2015). Pairwise Rasch model item parameter recovery under sparse data conditions. *Psychological Test and Assessment Modeling*, *57* (1), 3–36. Zugriff auf `http://www.psychologie-aktuell.com/fileadmin/download/ptam/1-2015_20150327/01_Heine.pdf`

Heine, J.-H., Tarnai, C. & Hartmann, F. G. (2011). Eine Methode zur Parameterbestimmung im Rasch-Modell bei fehlenden Werten. Bamberg. (21.09.2011 - 23.09. 2011)

Leunbach, G. (1961). On Quantitative Models for Qualitative Data. *Acta Sociologica*, *5* (3), 144–156. Zugriff am 2014-04-30 auf `http://www.jstor.org/stable/4193536`

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* (Nr. 1). Copenhagen: Danmarks pædagogiske Institut.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, *54* (3), 427–450.

Wright, B. D. & Masters, G. N. (1982). *Rating scale analysis.* Chicago: MESA Press.

Zwinderman, A. H. (1995). Pairwise parameter estimation in rasch models. *Applied Psychological Measurement*, *19* (4), 369–375.