Journal of Classification (2019) 36:566–575
https://doi.org/10.1007/s00357-018-9298-2

A Note on Applying the BCH Method Under Linear Equality and Inequality Constraints

L. Boeschoten[1,2] · M. A. Croon[1] · D. L. Oberski[3]

Published online: 18 December 2018

Abstract
Researchers often wish to relate estimated scores on latent variables to exogenous covariates not previously used in analyses. The BCH method corrects for asymptotic bias in estimates due to these scores' uncertainty and has been shown to be relatively robust. When applying the BCH approach however, two problems arise. First, negative cell proportions can be obtained. Second, the approach cannot deal with situations where marginals need to be fixed to specific values, such as edit restrictions. The BCH approach can handle these problems when placed in a framework of quadratic loss functions and linear equality and inequality constraints. This research note gives the explicit form for equality constraints and demonstrates how solutions for inequality constraints may be obtained using numerical methods.

Keywords Classification · Latent class analysis · Three-step procedure · BCH method

1 Introduction

Researchers in many different disciplines apply latent structure models in which observed variables are treated as indicators of an underlying latent variable that cannot be measured directly. An often used strategy in this context consists of three steps (Vermunt 2010). First, the parameters of the measurement model are estimated, describing the relationship between the latent variable and its indicators. Second, each respondent is assigned a latent score based on his/her scores on the indicators. Finally, the relationships between the latent scores and scores on exogenous variables are assessed.

Croon (2002) showed that for general latent structure models, such a strategy leads to inconsistent estimates of the parameters of the joint distribution of the latent variable and the

✉ L. Boeschoten
l.boeschoten@tilburguniversity.edu

[1] Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands

[2] Statistics Netherlands, The Hague, The Netherlands

[3] Utrecht University, Utrecht, The Netherlands

exogenous variables. Bolck et al. (2004) discussed this problem in the context of latent class analysis where observed variables are categorical. They also derived a correction procedure that produces consistent estimates, known as the BCH correction method. Subsequent simulation studies by Vermunt (2010), Bakk et al. (2013), Bakk and Vermunt (2016), and Nylund-Gibson and Masyn (2016) have demonstrated that this procedure produces unbiased parameter estimates and correct inference for a large range of simulation conditions. When applying the BCH correction method in cases of categorical exogenous variables, two problems can arise. First, negative cell proportion estimates can be obtained (Asparouhov and Muthén 2015). Second, the approach cannot deal with situations where marginals need to be constrained. An example is edit restrictions in official statistics, leading to certain marginals being fixed to zero (De Waal et al. 2012), which is also used in combination with latent class modelling (Boeschoten et al. 2017).

In this research, note the BCH method is extended to solve these two problems. We allow for linear equality and inequality constraints by noting the correction method minimizes a quadratic loss function and give a closed form solution for linear equality restrictions. Next, we demonstrate how solutions for inequality constraints may be obtained using numerical methods. We first discuss the three-step approach to the latent class model and the BCH correction method. We then show how to impose linear restrictions and how to extend this to including non-negativity constraints. At last, the extended BCH method is applied on a dataset from the Political Action Survey. In the Appendix, R code is given to apply the procedure.

2 The Three-Step Approach to the Latent Class Model and the BCH Correction Method

Let us denote a set of observed exogenous variables $\mathbf{Q}$ and an unobserved latent variable X. All variables involved are assumed to be categorical. Let $\mathbf{Q} = (\mathbf{Q}_1, \mathbf{Q}_2, ..., \mathbf{Q}_J)$ be the Cartesian product of J different discrete random variables $\mathbf{Q}_j$. If the variable $\mathbf{Q}_j$ is defined for n_j categories, the distribution of $\mathbf{Q}$ can be specified as a multinomial distribution with $n = \prod_{j=1}^{J} n_j$ categories.

In the basic latent class model considered by Bolck et al. (2004), a single categorical latent variable X with m categories is introduced. The variable X itself is not directly observed but only indirectly via a set of indicator variables $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, ..., \mathbf{Y}_K)$. Let the joint distribution of the categorical variables $\mathbf{Q}$, X, and $\mathbf{Y}$ be denoted by

$$\mathrm{p}(\mathbf{Q} = \mathbf{q}, X = x, \mathbf{Y} = \mathbf{y}) = \mathrm{p}(\mathbf{q}, x, \mathbf{y}).$$

Then, a possible factorization is

$$\mathrm{p}(\mathbf{q}, x, \mathbf{y}) = \mathrm{p}(\mathbf{q})\mathrm{p}(x|\mathbf{q})\mathrm{p}(\mathbf{y}|x, \mathbf{q}).$$

Since in the basic latent class model $\mathbf{Q}$ is assumed to have no direct effect on $\mathbf{Y}$, the latter result simplifies to

$$\mathrm{p}(\mathbf{q}, x, \mathbf{y}) = \mathrm{p}(\mathbf{q})\mathrm{p}(x|\mathbf{q})\mathrm{p}(\mathbf{y}|x).$$

The three-step approach to the estimation of the parameters of the latent class model starts with the estimation of the parameters of the measurement model represented by the conditional probability distribution $\mathrm{p}(\mathbf{y}|x)$. Once this estimation procedure is completed, individual research units may be assigned to one of the latent classes solely on the basis of their observed scores on $\mathbf{Y}$. This defines the second step of the estimation procedure

and results in an assignment of each individual to a latent class. If the random variable W represents the latent classes individuals are assigned to, and assignment is done using a modal rule where each individual is assigned to the class for which its posterior membership probability is the largest, this can be expressed as

$$\mathrm{p}(w|\mathbf{y}) = \begin{cases} 1 & \text{if } \mathrm{p}(x_1|\mathbf{y}) > \mathrm{p}(x_2|\mathbf{y}) \; \forall \; x_1 \neq x_2, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Different assignment rules than the modal rule will yield a different form for Eq. 2. All subsequent results also apply to other assignment rules, such as proportional or random assignment (Bakk 2015).

Since $\mathbf{Y}$ and $\mathbf{Q}$ are conditionally independent given X, so are W and $\mathbf{Q}$ and the conditional distributions are related by

$$\mathrm{p}(w|\mathbf{q}) = \sum_{x=1}^{\mathbf{X}} \mathrm{p}(w|x)\mathrm{p}(x|\mathbf{q}).$$

In terms of the joint distribution, this becomes

$$\mathrm{p}(\mathbf{q}, w) = \sum_{x=1}^{\mathbf{X}} \mathrm{p}(\mathbf{q}, x)\mathrm{p}(w|x).$$

The latter result can be recast as a matrix equation

$$\mathbf{E} = \mathbf{AD},$$

with the elements of the three matrices defined as $e_{\mathbf{q}w} = \mathrm{p}(\mathbf{q}, w)$, $a_{\mathbf{q}x} = \mathrm{p}(\mathbf{q}, x)$, and $d_{xw} = \mathrm{p}(w|x)$. After completing the first and the second estimation steps, the elements of the matrices $\mathbf{E}$ and $\mathbf{D}$ are known. The joint distribution of $\mathbf{Q}$ and the latent variable X is then given by

$$\mathbf{A} = \mathbf{ED}^{-1}.$$

Here, it is assumed that matrix $\mathbf{D}$ is not singular so that its inverse exists (see Bolck et al. (2004, pp. 13–14) for a discussion on when this assumption may be violated). A consistent estimate of $\mathbf{A}$ is $\hat{\mathbf{E}}\hat{\mathbf{D}}^{-1}$.

The previously obtained algebraic solution for matrix $\mathbf{A}$ can also be derived via a rather trivial minimization of a least squares function. Let $\mathbf{E}$ and $\mathbf{D}$ be matrices with known elements. Matrix $\mathbf{E}$ is of order $n \times m$ and $\mathbf{D}$ is an invertible matrix of order $m \times m$. Let $\mathbf{A}$ be an $n \times m$ matrix of unknown elements and consider the following least squares function:

$$\varphi = \frac{1}{2}\mathrm{tr}(\mathbf{AD} - \mathbf{E})'(\mathbf{AD} - \mathbf{E}).$$

Minimizing φ with respect to the unknown matrix $\mathbf{A}$ yields $\mathbf{A} = \mathbf{ED}^{-1}$, for which φ attains the truly minimal value of zero. Note that the factor $1/2$ is introduced to obtain simpler expressions for the first derivatives. Its introduction does not change the solution of the minimization problem.

3 The Correction Procedure Under Linear Equality Constraints

In some applications, simple linear restrictions may be imposed on the elements of matrix $\mathbf{A}$. For instance, some of the probabilities in the joint distribution of $\mathbf{Q}$ and X may be set equal to zero, for example for combinations of $\mathbf{Q}$ and X that cannot occur in practice. After

imposing such zero constraints, all the non-zero cell probabilities should still add to one. The quadratic loss function φ can be minimized under equality constraints on the unknown elements of matrix $\mathbf{A}$ by applying the method of Lagrangian multipliers.

We first rewrite the quadratic loss function φ in the following way using vectorization operations on matrices (see Schott 1997, pp. 261–266). For the vector of residuals $\mathbf{r}$, we obtain

$$\begin{aligned}\mathbf{r} &= \text{vec}(\mathbf{AD} - \mathbf{E}) \\ &= \text{vec}(\mathbf{I}_{n\times n}\mathbf{AD}) - \text{vec}(\mathbf{E}),\end{aligned}$$

where $\mathbf{I}_{n\times n}$ is an $n \times n$ identity matrix. Applying Theorem 7.15 from Schott (1997, p. 263) yields

$$\mathbf{r} = (\mathbf{D}' \otimes \mathbf{I}_{n\times n}) \cdot \text{vec}(\mathbf{A}) - \text{vec}(\mathbf{E}),$$

in which $\otimes$ is the Kronecker product of two matrices (Graham 1982). Defining $\mathbf{P} = \mathbf{D}' \otimes \mathbf{I}_{n\times n}$, $\mathbf{a} = \text{vec}(\mathbf{A})$ and $\mathbf{e} = \text{vec}(\mathbf{E})$, we are able to write

$$\mathbf{r} = \mathbf{Pa} - \mathbf{e},$$

so that the least squares function becomes

$$\begin{aligned}\varphi &= \frac{1}{2}\mathbf{r}'\mathbf{r} \\ &= \frac{1}{2}(\mathbf{a}'\mathbf{P}'\mathbf{Pa} - 2\mathbf{e}'\mathbf{Pa} + \mathbf{e}'\mathbf{e}).\end{aligned}$$

The completely unconstrained solution to the minimization problem is given by

$$\mathbf{a}_0 = (\mathbf{P}'\mathbf{P})^{-1} \cdot \mathbf{P}'\mathbf{e}.$$

Now suppose that the S linear equality constraints can be represented by a matrix equation

$$\mathbf{Ha} = \mathbf{c}.$$

The matrix $\mathbf{H}$ is of order $S \times N$, N being the number of cells in matrices $\mathbf{A}$ and $\mathbf{E}$. We may assume that $\mathbf{H}$ is of rank S; otherwise, the linear equality constraints would not be linearly independent. To minimize the least square function φ under a set of S linear constraints on the elements of $\mathbf{A}$, the Lagrangian is defined as

$$\mathbf{L} = \varphi - \lambda'(\mathbf{Ha} - \mathbf{c}). \tag{2}$$

Setting the first derivatives of $\mathbf{L}$ with respect to $\mathbf{a}$ equal to the zero vector, and solving for $\mathbf{a}$ yields:

$$\mathbf{a} = (\mathbf{P}'\mathbf{P})^{-1}(\mathbf{P}'\mathbf{e} + \mathbf{H}'\boldsymbol{\lambda}),$$

which can be rewritten as:

$$\mathbf{a} = \mathbf{a}_0 + (\mathbf{P}'\mathbf{P})^{-1}\mathbf{H}'\boldsymbol{\lambda}.$$

Solving for the unknown Lagrangian multipliers by taking the derivative of the Lagrangian (Eq. 2), and setting it to zero, or equivalently by imposing linear constraints $\mathbf{Ha} - \mathbf{c} = 0$ yields:

$$\boldsymbol{\lambda} = [\mathbf{H}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{H}']^{-1}(\mathbf{c} - \mathbf{Ha}_0).$$

So that the final solution for $\mathbf{a}$ is:

$$\mathbf{a} = \mathbf{a}_0 + (\mathbf{P}'\mathbf{P})^{-1}\mathbf{H}'[\mathbf{H}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{H}']^{-1}(\mathbf{c} - \mathbf{Ha}_0).$$

Note that the vector $\mathbf{c} - \mathbf{Ha}_0$ represents the deviations of the unconstrained solution from the linear equality constraints. Again a consistent estimate of $\mathbf{a}$ can be obtained by replacing $\mathbf{P}$ and $\mathbf{a}_0$ with their sample estimates.

4 The Correction Procedure Under Linear Equality and Inequality Constraints

A second issue with the BCH procedure is that in finite samples the consistent estimate $\hat{\mathbf{A}}$ hat may contain negative values. This issue is similar to the occurrence of Heywood cases in factor analysis (Heywood 1931). Such negative values in the probability table estimate $\hat{\mathbf{A}}$ may prevent subsequent analyses. We suggest preventing such inadmissible solutions by imposing inequality constraints. The resulting minimization problem is a quadratic program that can be solved by an iterative method.

Such a numerical iterative method for an equality and inequality constrained minimization of a quadratic function has been described by Goldfarb and Idnani (1983). Their numerical algorithm solves the quadratic programming problem of the form

$$\min\left(\frac{1}{2}\mathbf{b}'\mathbf{D}_{\text{mat}}\mathbf{b} - \mathbf{d}'_{\text{vec}}\mathbf{b}\right),$$

subject to the constraints

$$\mathbf{H}'\mathbf{b} \geq \mathbf{b}_0,$$

with respect to the n unknown parameters in vector $\mathbf{b}$. The matrix $\mathbf{D}_{\text{mat}}$ is a given $n \times n$ symmetric positive definite matrix whereas $\mathbf{d}_{\text{vec}}$ is a given $n \times 1$ vector.

To apply the Goldfarb-Idnani optimization procedure in the present context, the following definitions have to be implemented. First, to include non-negativity constraints, we make use of Theorem 7.6 from Schott (1997, p. 254) to obtain

$$\begin{aligned}\mathbf{D}_{\text{mat}} &= \mathbf{P}'\mathbf{P} \\ &= (\mathbf{D}\mathbf{D}') \otimes \mathbf{I}_{n\times n}.\end{aligned}$$

and

$$\mathbf{d}_{\text{vec}} = \mathbf{P}'\mathbf{e}$$

Since it is assumed that matrix $\mathbf{D}$ is of full rank, the matrix $\mathbf{P}'\mathbf{P}$ is positive-definite. This ensures that the quadratic loss function φ is strictly convex. Moreover, the type of equality and inequality constraints considered here (the sum of the elements in matrix $\mathbf{A}$ is equal to 1, where all elements ≥ 0 and some are fixed to 0), define a convex region in the parameter space.

To represent the constraints on the cell probabilities we now define matrix $\mathbf{H}$ in such a way that the first row of $\mathbf{H}$ has all its elements equal to 1. This row represents a constraint on the sum of all cell probabilities. We represent this row vector as matrix $\mathbf{H}_0$. Let $J = \{1, 2, 3, ..., N\}$ be an index set corresponding to the column numbers of matrix $\mathbf{H}$. This index set can be partitioned in two non-overlapping subsets J_1 and J_2:

- Subset J_1 contains the indices of the elements of vector $\mathbf{a}$ which are set exactly equal to zero: for those indices j we require $\mathbf{a}_j = 0$;
- Subset J_2 contains the indices of the elements of vector $\mathbf{a}$ which are required to be non-negative: for those indices j we require $\mathbf{a}_j \geq 0$.

Now let $\boldsymbol{I}_n$ be an $N \times N$ identity matrix and permute the rows of this matrix so that the upper part contains the rows corresponding with the index numbers in J_1, and the lower part

of the permuted identity matrix contains the rows corresponding with the index numbers in J_2. Referring to the two parts of the permuted identity matrix as $\mathbf{H}_1$ and $\mathbf{H}_2$, respectively, the matrix $\mathbf{H}$ is obtained by

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \\ \mathbf{H}_2 \end{pmatrix},$$

where $\mathbf{H}$ is used to obtain the final solution for $\mathbf{a}$. Note that in cases where we are not interested in applying equality constraints, but we are interested in applying the inequality constraints we simply omit $\mathbf{H_1}$. Vector $\mathbf{b_0}$ is of length $N + 1$, with its first element equal to 1 and all the remaining elements equal to 0.

With this procedure, we are able to find a solution for $\mathbf{A}$ (the joint distribution of latent variable X and exogenous covariates $\mathbf{Q}$) where the sum of the elements is equal to 1, where no negative elements are created, and where impossible combinations of scores can be set to have a probability of zero. Having defined $\mathbf{b}$, $\mathbf{D}_{\text{mat}}$ and $\mathbf{H}$, the solution can be obtained using standard software for quadratic programming, such as the R package `quadprog` (Turlach and Weingessel 2013).

5 Application

As an illustration, the extended BCH method is applied on a dataset from the Political Action Survey (Barnes et al. 1979; Jennings and Van Deth 1990). The dataset consists of five dichotomous indicators on political involvement and tolerance ("System Responsiveness"; "Ideological Level"; "Repression Potential"; "Protest Approval"; "Conventional Participation") and three nominal covariates ("Sex"; "Level Of Education"; "Age"). This dataset has previously been used in Hagenaars (1993) and Vermunt and Magidson (2000) and in the Latent GOLD user's manual (Vermunt and Magidson 2005). The dataset as well as the syntax used in this illustration can be found in Latent GOLD version 5.1 under "syntax examples" → LCA → restrictions → equalities → Model C.

In the first step, a four class restricted model is applied to distinguish between four latent classes on involvement and tolerance. In this model, response probabilities are restricted to be equal for the items "System Responsiveness" and "Conventional Participation," and the response probability for the variable "Ideological Level" is fixed to 0 by specifying a logit of 100.

In the second step, cases are assigned to a latent class by using modal assignment, resulting in the imputed latent variable W. In the third step, the relationship between the imputed latent variable "Involvement And Tolerance" (W) and exogenous covariate "Age" (Q) is investigated. The $\mathbf{E}$-matrix containing the joint probabilities of these variables is:

$$\mathbf{E} = \begin{array}{c} \\ Q_{16\text{-}34} \\ Q_{35\text{-}57} \\ Q_{58\text{-}91} \end{array} \begin{array}{c} \begin{array}{cccc} W_1 & W_2 & W_3 & W_4 \end{array} \\ \begin{pmatrix} 0.05795848 & 0.15743945 & 0.01643599 & 0.09256055 \\ 0.08477509 & 0.17560554 & 0.05276817 & 0.03979239 \\ 0.12802768 & 0.10034602 & 0.06920415 & 0.02508651 \end{pmatrix} \end{array}.$$

The **D**-matrix describing the relationship between the imputed latent variable "involvement and tolerance" (W) and the latent variable "involvement and tolerance" (X) is also obtained:

$$\mathbf{D} = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \end{array} \begin{array}{c} \begin{array}{cccc} W_1 & W_2 & W_3 & W_4 \end{array} \\ \begin{pmatrix} 0.67389148 & 0.1570985 & 0.02678610 & 0.1422239 \\ 0.01898361 & 0.7891416 & 0.05879905 & 0.1330757 \\ 0.17186997 & 0.2725275 & 0.54176422 & 0.0138383 \\ 0.12184782 & 0.3220914 & 0.01975761 & 0.5363031 \end{pmatrix} \end{array}.$$

The BCH method can now be applied by estimating $\mathbf{ED}^{-1}$, resulting in the **A** matrix:

$$\mathbf{A}_{\text{unconstraint}} = \begin{array}{c} \\ Q_{16\text{-}34} \\ Q_{35\text{-}57} \\ Q_{58\text{-}91} \end{array} \begin{array}{c} \begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \end{array} \\ \begin{pmatrix} 0.0577223 & 0.13465976 & 0.008359502 & 0.123652898 \\ 0.1018944 & 0.17635045 & 0.073167182 & 0.001529175 \\ 0.1618782 & 0.06157076 & 0.113576159 & -0.014360760 \end{pmatrix} \end{array}.$$

As can be seen, this result is inadmissable since the cell $Q_{58\text{-}91} \times X_4$ contains a negative value. Therefore, it will not be possible to estimate posterior membership probabilities and to do subsequent analyses here.

When the extended BCH method is applied, the following constrained **A** matrix is obtained:

$$\mathbf{A}_{\text{constraint}} = \begin{array}{c} \\ Q_{16\text{-}34} \\ Q_{35\text{-}57} \\ Q_{58\text{-}91} \end{array} \begin{array}{c} \begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \end{array} \\ \begin{pmatrix} 0.05741718 & 0.13472999 & 0.007627791 & 0.1229631559 \\ 0.10158926 & 0.17642067 & 0.072435471 & 0.0008394325 \\ 0.15689781 & 0.05436459 & 0.114714655 & 0.0000000000 \end{pmatrix} \end{array}.$$

The cell $Q_{58\text{-}91} \times X_4$ does not contain a negative value anymore, so this matrix can now be used to estimate posterior membership probabilities and to do subsequent analyses.

Since there are no combinations of scores between "Involvement And Tolerance" and "Age" that are not possible in practice, it is not needed to fix any marginals to zero.

6 Conclusion

We have modified the BCH method to include linear equality and inequality constraints solving the problem of negative solutions and allowing for restrictions on arbitrary cell margins. With these adjustments, analysts interested in relating covariates to assignments on latent class variables will now be able to, for example, impose edit restrictions, further analyse solutions that were previously inadmissible, and analyse datasets involving more complex marginal restrictions. The application demonstrates that when a negative value is obtained using the regular BCH method, this can be solved by using the extended BCH method. In the Appendix, R code is given to apply the extended BCH method, and an addition to the example is given that demonstates how margins can be fixed to zero using the extended BCH method.

Appendix

This Appendix consists of two sections. In Appendix 1, R code is given to apply the extended BCH method as described in the research note. In Appendix 2, it is illustrated how the code can be used and how margins can be fixed to zero.

Appendix 1

The iterative method for an equality and inequality constrained minimization of a quadratic function described by Goldfarb and Idnani (1983) has been implemented in the R package `quadprog` available in the repository CRAN (Turlach and Weingessel 2013).

The minimization procedure is implemented in the function `solve.QP` which is called as

`solve.QP(Dmat,dvec,Amat,bvec,meq)`.

Its arguments are:

- `Dmat`: the matrix $\mathbf{D}$ appearing in the quadratic function: $(\mathbf{DD}') \otimes \mathbf{I}_{n\times n}$;
- `dvec`: the vector $\mathbf{d}$ appearing in the quadratic function: $\mathbf{e}'\mathbf{P}$;
- `Amat`: The transpose of $\mathbf{H}$ ($\mathbf{H}'$) defining the linear constraints on the parameters $\mathbf{b}$;
- `bvec`: A vector of length $N+1$, with its first elements equal to 1 and the remaining N elements all equal to 0, these are the constants $\mathbf{b}_0$ in the constraints.
- `meq`: 1+ the number of elements in J_1

The minimization procedure can be applied using the following function:

```
qpsolve <- function(e,d,iequal){
  nr    <- nrow(e)
  nc    <- ncol(e)
  ncel  <- nr*nc
  evec  <- as.vector(e)
  id    <- diag(nr)
  p     <- kronecker(t(d),id)
  dmat  <- kronecker(d %*% t(d),id)
  dvec  <- as.vector(evec %*% p)
  im    <- diag(ncel)
  i1    <- iequal
  i2    <- setdiff(1:ncel,i1)
  index <- c(i1,i2)
  im2   <- im[index,]
  at    <- rbind(rep(1,ncel),im2)
  amat  <- t(at)
  bvec  <- c(1,rep(0,ncel))
  meq   <- 1 + length(iequal)
  res   <- solve.QP(dmat,dvec,amat,bvec,meq)
  return(res)
}
```

The function is used by defining the $\mathbf{E}$-matrix, the $\mathbf{D}$-matrix and the inequality constraints:

`res <- qpsolve(E,D,iequal)`.

Appendix 2

In Section 5, the extended BCH method is applied on a dataset from the Political Action Survey. There are no combinations of scores between the latent variable and the exogenous covariate that are not possible in practice, so therefore it is not needed to fix any marginals to zero. However, in this appendix, a margin of the **A**-matrix is fixed for illustrative purposes.

As can be seen in Appendix 1, the `qsolve()` function can be used by defining the **E**-matrix, the **D**-matrix and the inequality constraints. In the application section, the **E**-matrix, the **D**-matrix are defined, and since there are no inequality constraints, these are omitted for the function by specifying

```
iequal <- c()
```

By using the function `qpsolve(E,D,iequal)`, both the unconstrained and the constrained solutions for the **A**-matrix are given. The output is saved under the name `res`: `res <- qpsolve(E,D,iequal)`. The unconstrained solution can be requested by:

```
res$unconstrained.solution
```

and the constrained solution can be requested by:

```
res$solution
```

For illustration purposes, the cell $Q_{16\text{-}34} \times X_3$ of the **A**-matrix is fixed to zero. When vectorizing the **A**-matrix, this cell is the seventh element, so this needs to be specified:

```
iequal <- c(7)
```

It can now be seen that the constrained solution is not only without negative values, also the cell $Q_{16\text{-}34} \times X_3$ is fixed to zero:

$$\mathbf{A}_{\text{constraint}} = \begin{array}{c} \\ Q_{16\text{-}34} \\ Q_{35\text{-}57} \\ Q_{58\text{-}91} \end{array} \begin{array}{c} \begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \end{array} \\ \begin{pmatrix} 0.06007299 & 0.13800030 & 0.0000000 & 0.12215613 \\ 0.10183738 & 0.17636356 & 0.0730305 & 0.00140033 \\ 0.15732017 & 0.05457865 & 0.1152400 & 0.00000000 \end{pmatrix} \end{array}.$$

References

Asparouhov, T., & Muthén, B. (2015). Auxiliary variables in mixture modelling: using the BCH method in Mplus to estimate a distal outcome model and an arbitrary secondary model. *Mplus Web Notes 21*. Version was retrieved April 26th, 2017 from https://www.statmodel.com/examples/webnotes/webnote21.pdf.

Bakk, Z. (2015). *Contributions to bias adjusted stepwise latent class modeling* (Doctoral thesis, Tilburg University, Tilburg, The Netherlands). Retrieved from https://pure.uvt.nl/portal/files/8521154/Bakk_Contributions_16_10_2015.pdf.

Bakk, Z., Tekle, F.B., Vermunt, J.K. (2013). Estimating the association between latent class membership and external variables using bias-adjusted three-step approaches. *Sociological Methodology*, *43*, 272–311. https://doi.org/10.1177/0081175012470644.

Bakk, Z., & Vermunt, J.K. (2016). Robustness of stepwise latent class modelling with continuous distal outcomes. *Structural Equation Modelling: A Multidisciplinary Journal*, *23*, 20–31. https://doi.org/10.1080/10705511.2014.955104.

Barnes, B.H., Kaase, M., Allerback, K.R., Farah, B., Heunks, F., Inglehart, R., Jennings, M.K., Klingemann, A.M., Rosenmayr, L. (1979). *Political Action, Mass participation in five Western Democracies*. Beverly Hills: Sage Publications. ISBN-10: 0803909578; ISBN-13: 978-0803909571.

Bolck, A., Croon, M., Hagenaars, J. (2004). Estimating latent structure models with categorical variables: one-step versus three-step estimators. *Political Analysis*, *12*, 3–27. https://doi.org/10.1093/pan/mph001.

Boeschoten, L., Oberski, D., de Waal, T. (2017). Estimating classification errors under edit restrictions in composite survey-register data using multiple imputation latent class modelling (MILC). *Journal of Official Statistics*, *33*(4), 921–962. https://doi.org/10.1515/jos-2017-0044.

Croon, M. (2002). Using predicted latent scores in general latent structure models. In I. Marcoulides, & G.A. Moustaki (Eds.), *Latent variable and latent structures models* (pp. 195–224). Mahwah: Lawrence Erlbaum.

De Waal, T., Pannekoek, J., Scholtus, S. (2012). The editing of statistical data: methods and techniques for the efficient detection and correction of errors and missing values. *Wiley Interdisciplinary Reviews: Computational Statistics*, *4*, 204–210. https://doi.org/10.1002/wics.1194.

Goldfarb, D., & Idnani, A. (1983). A numerically stable dual method for solving strictly convex quadratic programs. *Mathematical Programming*, *27*, 1–33. https://doi.org/10.1007/BF02591962.

Graham, A. (1982). *Kronecker products and matrix calculus: with applications*. New York: Wiley.

Hagenaars, J.A. (1993). *Loglinear models with latent variables*. Newbury Park: Sage.

Heywood, H.B. (1931). On finite sequences of real numbers. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *134*, 486–501. https://doi.org/10.1098/rspa.1931.0209.

Jennings, M.K., & Van Deth, J.W. (1990). *Continuities in political action: a longitudinal study of political orientations in three western democracies* (Vol. 5). Walter de Gruyter GmbH Co KG.

Nylund-Gibson, K., & Masyn, K.E. (2016). Covariates and mixture modeling: results of a simulation study exploring the impact of misspecified effects on class enumeration. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*, 782–797. https://doi.org/10.1080/10705511.2016.1221313.

Schott, J.R. (1997). *Matrix analysis for statistics*. New York: Wiley.

Turlach, B.A., & Weingessel, A. (2013). Quadprog: functions to solve quadratic programming problems. *R package version 1.5–5*. Version was retrieved May 1st, 2017 from https://cran.r-project.org/web/packages/quadprog/quadprog.pdf.

Vermunt, J.K. (2010). Latent class modeling with covariates: two improved three-step approaches. *Political Analysis*, *18*, 450–469. https://doi.org/10.1093/pan/mpq025.

Vermunt, J.K., & Magidson, J. (2000). Graphical displays for latent class cluster and latent class factor models. In W. Jansen, & J.G. Bethlehem (Eds.), *Proceedings in Computational Statistics 2000. Statistics Netherlands. ISSN 0253-018X* (pp. 121–122).

Vermunt, J.K., & Magidson, J. (2005). *Latent GOLD 4.0 User's Guide*. Belmont: Statistical Innovations Inc. https://www.statisticalinnovations.com/wpcontent/uploads/LGusersguide.pdf.