Usefulness and estimation of proportionality constraints

The propcnsreg package

Maarten L. Buis

Department of Social Research Methodology
Vrije Universiteit Amsterdam
http://home.fsw.vu.nl/m.buis/
Outline

usefulness
proportionality constraint
a latent variable
scale for a categorical variable

estimation
Outline

usefulness
  proportionality constraint
  a latent variable
  scale for a categorical variable

estimation
Hypothesis:
Effect of father’s and mother’s socioeconomic status on child’s education can change over cohorts,
Hypothesis:
Effect of father’s and mother’s socioeconomic status on child’s education can change over cohorts, but the relative contribution of the father and the mother remains constant.
Hypothesis:
Effect of father’s and mother’s socioeconomic status on child’s education can change over cohorts, but the relative contribution of the father and the mother remains constant.

\[ ed = \beta_0 + \beta_1 coh + (1 + \lambda_1 coh)(\gamma_1 pasei + \gamma_2 masei) \]
Hypothesis:
Effect of father’s and mother’s socioeconomic status on child’s education can change over cohorts, but the relative contribution of the father and the mother remains constant.

\[ ed = \beta_0 + \beta_1 coh + (1 + \lambda_0)(\gamma_1 pasei + \gamma_2 masei) \]
Hypothesis:
Effect of father’s and mother’s socioeconomic status on child’s education can change over cohorts, but the relative contribution of the father and the mother remains constant.

\[
ed = \beta_0 + \beta_1 \text{coh} + (1 + \lambda_1 0)(\gamma_1 \text{pasei} + \gamma_2 \text{masei})
\]

\[
ed = \beta_0 + \beta_1 \text{coh} + \gamma_1 \text{pasei} + \gamma_2 \text{masei}
\]
Hypothesis:
Effect of father’s and mother’s socioeconomic status on child’s education can change over cohorts, but the relative contribution of the father and the mother remains constant.

\[ ed = \beta_0 + \beta_1 \text{coh} + (1 + \lambda_1)(\gamma_1pasei + \gamma_2masei) \]
Hypothesis:
Effect of father’s and mother’s socioeconomic status on child’s education can change over cohorts, but the relative contribution of the father and the mother remains constant.

\[
ed = \beta_0 + \beta_1 coh + (1 + \lambda_1)(\gamma_1 pasei + \gamma_2 masei)\]

\[
ed = \beta_0 + \beta_1 coh + (1 + \lambda_1)\gamma_1 pasei + (1 + \lambda_1)\gamma_2 masei\]
empirical example

- 7 surveys held between 1994 and 2006 in the USA from the General Social Survey (GSS) containing data on 2,500 white male.
empirical example

- 7 surveys held between 1994 and 2006 in the USA from the General Social Survey (GSS) containing data on 2,500 white males.
- Variable *degree*: educational attainment in pseudo years
empirical example

- 7 surveys held between 1994 and 2006 in the USA from the General Social Survey (GSS) containing data on 2,500 white male.
- Variable *degree*: educational attainment in pseudo years
- Variable *byr*: cohort centered in 1940 and measuring time in decades, ranges between 1929 and 1979.
empirical example

- 7 surveys held between 1994 and 2006 in the USA from the General Social Survey (GSS) containing data on 2,500 white male.
- Variable *degree*: educational attainment in pseudo years
- Variable *byr*: cohort centered in 1940 and measuring time in decades, ranges between 1929 and 1979.
- Variables *pasei* and *masei*: Father’s and mother’s occupational status, ranges between 0 and 1.
Usefulness and estimation of proportionality constraints

example output

```
. propcnsreg degree byr, lambda(byr) constrained(masei pasei) lcons
Constraint: [lambda]_cons = 1

                  degree | Coef. Std. Err.  z    P>|z|   [95% Conf. Interval]
-----------------+--------------------------------------------------
unconstrained    |        
        byr |  0.0392033  0.1418648  0.28  0.782  -0.2388465   0.3172531
        _cons |   10.2406   0.2762536  37.07  0.000    9.699157   10.78205
constrained      |        
        masei |   3.363018   0.3688164  9.12  0.000    2.640152   4.085885
        pasei |   3.948723   0.3972388  9.94  0.000    3.170149   4.727296
lambda           |        
        byr |  -0.0323712  0.0378542  -0.86  0.392   -0.1065637   0.0418212
        _cons |     1.000   .0000000   .000  1.000   .0000000   .0000000
ln_sigma         |        
        _cons |    0.837853  0.0141989  59.01  0.000    0.8100234   0.8656826

LR test vs. unconstrained model: chi2(1) =  0.04  Prob > chi2 =  0.849
```
alternative way of looking

\[ ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) (\gamma_1 pasei + \gamma_2 masei) \]

latent family sei

Need to identify the latent variable by fixing the origin and scale. If the minimum value of \( pasei \) and \( masei \) is 0 then the origin is fixed to when both variables are minimum.
alternative way of looking

\[
ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) (\gamma_1 pasei + \gamma_2 masei)
\]

- Need to identify the latent variable by fixing the origin and the scale.
alternative way of looking

\[ ed = \beta_0 + \beta_1 \text{coh} + (\lambda_0 + \lambda_1 \text{coh}) (\gamma_1 \text{pasei} + \gamma_2 \text{masei}) \]

- Need to identify the latent variable by fixing the origin and the scale.
- If the minimum value of \text{pasei} and \text{masei} is 0 then the origin is fixed to when both variables are minimum.
alternative way of looking

$$ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) (\gamma_1 pasei + \gamma_2 masei)$$

- Need to identify the latent variable by fixing the origin and the scale.
- If the minimum value of $pasei$ and $masei$ is 0 then the origin is fixed to when both variables are minimum.

$$\text{latent family sei} = \gamma_1 pasei + \gamma_2 masei$$
alternative way of looking

\[ ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) (\gamma_1 pasei + \gamma_2 masei) \]

- Need to identify the latent variable by fixing the origin and the scale.
- If the minimum value of *pasei* and *masei* is 0 then the origin is fixed to when both variables are minimum.

\[ \text{latent family sei} = \gamma_1 0 + \gamma_2 0 = 0 \]
alternative way of looking

\[ ed = \beta_0 + \beta_1 \text{coh} + (\lambda_0 + \lambda_1 \text{coh}) (\gamma_1 \text{pasei} + \gamma_2 \text{masei}) \]

- Need to identify the latent variable by fixing the origin and the scale.
- If the minimum value of \( \text{pasei} \) and \( \text{masei} \) is 0 then the origin is fixed to when both variables are minimum.
- If the maximum value of \( \text{pasei} \) and \( \text{masei} \) is 1, and
alternative way of looking

\[ ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) (\gamma_1 pasei + \gamma_2 masei) \]

- Need to identify the latent variable by fixing the origin and the scale.
- If the minimum value of \( pasei \) and \( masei \) is 0 then the origin is fixed to when both variables are minimum.
- If the maximum value of \( pasei \) and \( masei \) is 1, and their parameters are constrained to sum to 1,
alternative way of looking

\[ ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) (\gamma_1 pasei + \gamma_2 masei) \]

- Need to identify the latent variable by fixing the origin and the scale.
- If the minimum value of \( pasei \) and \( masei \) is 0 then the origin is fixed to when both variables are minimum.
- If the maximum value of \( pasei \) and \( masei \) is 1, and their parameters are constrained to sum to 1, then the unit is fixed to the distance between both variables at minimum and both variables at maximum.
alternative way of looking

\[ ed = \beta_0 + \beta_1 \text{coh} + (\lambda_0 + \lambda_1 \text{coh}) (\gamma_1 \text{pasei} + \gamma_2 \text{masei}) \]

▶ Need to identify the latent variable by fixing the origin and the scale.
▶ If the minimum value of pasei and masei is 0 then the origin is fixed to when both variables are minimum.
▶ If the maximum value of pasei and masei is 1, and their parameters are constrained to sum to 1, then the unit is fixed to the distance between both variables at minimum and both variables at maximum.

latent family sei = \gamma_1 \text{pasei} + \gamma_2 \text{masei}
alternative way of looking

\[ ed = \beta_0 + \beta_1 \text{coh} + (\lambda_0 + \lambda_1 \text{coh}) (\gamma_1 \text{pasei} + \gamma_2 \text{masei}) \]

- Need to identify the latent variable by fixing the origin and the scale.
- If the minimum value of \( \text{pasei} \) and \( \text{masei} \) is 0 then the origin is fixed to when both variables are minimum.
- If the maximum value of \( \text{pasei} \) and \( \text{masei} \) is 1, and their parameters are constrained to sum to 1, then the unit is fixed to the distance between both variables at minimum and both variables at maximum.

\[
\text{latent family sei} = \gamma_1 1 + \gamma_2 1 = 1
\]
Usefulness and estimation of proportionality constraints

A latent variable scale for a categorical variable

Example output

```stata
. propcnsreg degree byr, lambda(byr) constrained(masei pasei) unit(masei pasei)
Constraint: [constrained]masei + [constrained]pasei = 1

|         | Coef.    | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|---------|----------|-----------|-------|------|---------------------|
| unconstrained |         |           |       |      |                     |
| byr     | 0.0392033 | 0.1418647 | 0.28  | 0.782| -0.238464 to 0.3172529 |
| _cons   | 10.2406  | 0.2762534 | 37.07 | 0.000| 9.699158 to 10.78205 |

|         | Coef.    | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|---------|----------|-----------|-------|------|---------------------|
| constrained |       |           |       |      |                     |
| masei   | 0.4599477 | 0.0323745 | 14.21 | 0.000| 0.3964949 to 0.5234005 |
| pasei   | 0.5400523 | 0.0323745 | 16.68 | 0.000| 0.4765995 to 0.6035051 |

|         | Coef.    | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|---------|----------|-----------|-------|------|---------------------|
| lambda  |         |           |       |      |                     |
| byr     | -0.2366899 | 0.2935214 | -0.81 | 0.420| -0.8119814 to 0.3386015 |
| _cons   | 7.311741  | 0.601956  | 12.15 | 0.000| 6.131929 to 8.491553  |

|         | Coef.    | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|---------|----------|-----------|-------|------|---------------------|
| ln_sigma |         |           |       |      |                     |
| _cons   | 0.837853 | 0.014199  | 59.01 | 0.000| 0.8100234 to 0.8656826 |

LR test vs. unconstrained model: chi2(1) = 0.04  Prob > chi2 = 0.849
```
Example

Differences in the effect of education in 5 dummies on occupational status between white and black US men:

- < highschool (reference)
- highschool \((hs)\)
- some college \((sc)\)
- college \((c)\)
- graduate \((g)\)
Example

Differences in the effect of education in 5 dummies on occupational status between white and black US men:

- < highschool (reference)
- highschool (hs)
- some college (sc)
- college (c)
- graduate (g)

\[ ise_i = \beta_0 + (\lambda_0 + \lambda_1 \text{black})(\gamma_1 hs + \gamma_2 sc + \gamma_3 c + \gamma_4 g) \]
Example

- Differences in the effect of education in 5 dummies on occupational status between white and black US men:
  - < highschool (reference)
  - highschool (hs)
  - some college (sc)
  - college (c)
  - graduate (g)

\[
isei = \beta_0 + (\lambda_0 + \lambda_1 \text{black})(\gamma_1 \text{hs} + \gamma_2 \text{sc} + \gamma_3 \text{c} + 1 \text{g})
\]
Example

- Differences in the effect of education in 5 dummies on occupational status between white and black US men:
  - < highschool (reference)
  - highschool (hs)
  - some college (sc)
  - college (c)
  - graduate (g)

\[
isei = \beta_0 + (\lambda_0 + \lambda_1 \text{black})(\gamma_1 hs + \gamma_2 sc + \gamma_3 c + 1 g)
\]

\(\gamma_1, \gamma_2, \text{ and } \gamma_3\) now measure the position of highschool, some college, and college education, relative to less than highschool (0) and graduate (1).
Usefulness and estimation of proportionality constraints

usefulness
estimation
a latent variable
scale for a categorical variable

example output

```
. propcnsreg sei black, lambda(black) constrained(hs sc c g) unit(g)
Constraint: [constrained]g = 1

|            | Coef.  | Std. Err. | z      | P>|z|  | [95% Conf. Interval] |
|------------|--------|-----------|--------|------|---------------------|
|            |        |           |        |      |                     |
| unconstrained |      |           |        |      |                     |
| black      | -.042371 | .009563  | -4.43 | 0.000 | -.0611141 -.0236279 |
| _cons      | .3638307  | .0076114  | 47.80 | 0.000 | .3489126 .3787488  |
|            |        |           |        |      |                     |
| constrained |      |           |        |      |                     |
| hs         | .2226429  | .016662  | 13.36 | 0.000 | .1899861 .2552997  |
| sc         | .4411229  | .0206904  | 21.32 | 0.000 | .4005705 .4816753  |
| c          | .7185653  | .01676   | 42.87 | 0.000 | .6857163 .7514144  |
| g          | 1       |          | .      | .    | .                   |
|            |        |           |        |      |                     |
| lambda     |      |           |        |      |                     |
| black      | .0458751  | .0227816  | 2.01  | 0.044 | .0012239 .0905263  |
| _cons      | .38541  | .0099432  | 38.76 | 0.000 | .3659217 .4048983  |
|            |        |           |        |      |                     |
| ln_sigma   |      |           |        |      |                     |
| _cons      | -1.859163 | .0090043 | -206.48 | 0.000 | -1.876811 -1.841515 |
|            |        |           |        |      |                     |
| LR test vs. unconstrained model: chi2(3) = 5.42  Prob > chi2 = 0.144
```

Maarten L. Buis
 Scaling of education

Maarten L. Buis
Usefulness and estimation of proportionality constraints
Scaling of education

usefulness estimation
proportionality constraint
a latent variable
scale for a categorical variable

Maarten L. Buis
Usefulness and estimation of proportionality constraints
EM-algorithm for starting values

\[ y = \beta_0 + \beta_1 x_1 + (\lambda_0 + \lambda_1 x_1)(\gamma_1 z_1 + \gamma_2 z_2) \]
EM-algorithm for starting values

\[ y = \beta_0 + \beta_1 x_1 + (\lambda_0 + \lambda_1 x_1)(\gamma_1 z_1 + \gamma_2 z_2) \]

1. Given current estimates/starting value for \( \gamma \), create a new variable containing the latent variable.
EM-algorithm for starting values

\[ y = \beta_0 + \beta_1 x_1 + (\lambda_0 + \lambda_1 x_1)(\gamma_1 z_1 + \gamma_2 z_2) \]

1. Given current estimates/starting value for \( \gamma \), create a new variable containing the latent variable. This simplifies the problem to: \( y = \beta_0 + \beta_1 x_1 + \lambda_0 \text{latent} + \lambda_1 x_1 \text{latent} \)
EM-algorithm for starting values

\[ y = \beta_0 + \beta_1 x_1 + (\lambda_0 + \lambda_1 x_1)(\gamma_1 z_1 + \gamma_2 z_2) \]

1. Given current estimates/starting value for \( \gamma \), create a new variable containing the latent variable. This simplifies the problem to: \( y = \beta_0 + \beta_1 x_1 + \lambda_0 \text{latent} + \lambda_1 x_1 \text{latent} \)

2. Estimate \( \beta \) and \( \lambda \) using `regress`

3. Given current estimates of \( \lambda \), create a new variable containing the effect of the latent variable. This simplifies the problem to: \( y = \beta_0 + \beta_1 x_1 + \gamma_1 \text{effect} z_1 + \gamma_2 \text{effect} z_2 \)

4. Estimate \( \beta \) and \( \gamma \) using `cnsreg` imposing the constraint specified in the unit option.

5. Repeat steps 1-4 till convergence.
EM-algorithm for starting values

\[ y = \beta_0 + \beta_1 x_1 + (\lambda_0 + \lambda_1 x_1)(\gamma_1 z_1 + \gamma_2 z_2) \]

1. Given current estimates/starting value for \( \gamma \), create a new variable containing the latent variable. This simplifies the problem to: \( y = \beta_0 + \beta_1 x_1 + \lambda_0 \text{latent} + \lambda_1 x_1 \text{latent} \)

2. Estimate \( \beta \) and \( \lambda \) using `regress`

3. Given current estimates of \( \lambda \), create a new variable containing the effect of the latent variable.
EM-algorithm for starting values

\[ y = \beta_0 + \beta_1 x_1 + (\lambda_0 + \lambda_1 x_1)(\gamma_1 z_1 + \gamma_2 z_2) \]

1. Given current estimates/starting value for \( \gamma \), create a new variable containing the latent variable. This simplifies the problem to: \( y = \beta_0 + \beta_1 x_1 + \lambda_0 \text{latent} + \lambda_1 x_1 \text{latent} \)

2. Estimate \( \beta \) and \( \lambda \) using \textit{regress}

3. Given current estimates of \( \lambda \), create a new variable containing the effect of the latent variable. This simplifies the problem to: \( y = \beta_0 + \beta_1 x_1 + \gamma_1 \text{effect}z_1 + \gamma_2 \text{effect}z_2 \)
EM-algorithm for starting values

\[ y = \beta_0 + \beta_1 x_1 + (\lambda_0 + \lambda_1 x_1)(\gamma_1 z_1 + \gamma_2 z_2) \]

1. Given current estimates/starting value for \( \gamma \), create a new variable containing the latent variable. This simplifies the problem to: \( y = \beta_0 + \beta_1 x_1 + \lambda_0 \text{latent} + \lambda_1 x_1 \text{latent} \)

2. Estimate \( \beta \) and \( \lambda \) using \texttt{regress}

3. Given current estimates of \( \lambda \), create a new variable containing the effect of the latent variable. This simplifies the problem to: \( y = \beta_0 + \beta_1 x_1 + \gamma_1 \text{effectz}_1 + \gamma_2 \text{effectz}_2 \)

4. Estimate \( \beta \) and \( \gamma \) using \texttt{cnsreg} imposing the constraint specified in the \texttt{unit} option.
EM-algorithm for starting values

\[ y = \beta_0 + \beta_1 x_1 + (\lambda_0 + \lambda_1 x_1)(\gamma_1 z_1 + \gamma_2 z_2) \]

1. Given current estimates/starting value for \( \gamma \), create a new variable containing the latent variable. This simplifies the problem to: \( y = \beta_0 + \beta_1 x_1 + \lambda_0 \text{latent} + \lambda_1 x_1 \text{latent} \)

2. Estimate \( \beta \) and \( \lambda \) using `regress`

3. Given current estimates of \( \lambda \), create a new variable containing the effect of the latent variable. This simplifies the problem to: \( y = \beta_0 + \beta_1 x_1 + \gamma_1 \text{effect} z_1 + \gamma_2 \text{effect} z_2 \)

4. Estimate \( \beta \) and \( \gamma \) using `cnsreg` imposing the constraint specified in the `unit` option.

5. Repeat steps 1-4 till convergence.
speed and standard errors
To speed up convergence every 5\textsuperscript{th} iteration will consist of two \texttt{ml} iterations for the complete model.
speed and standard errors

- To speed up convergence every 5\textsuperscript{th} iteration will consist of two \texttt{ml} iterations for the complete model.
- Once the EM has converged, these estimates are fed into \texttt{ml} for the complete model to get the variance covariance matrix.
### Example Iteration Log

#### Improving Starting Values

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Unconstrained Part Only</th>
<th>Constrained Part Only</th>
<th>Full Model Part Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2712.7047</td>
<td>2716.1367</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2716.4376</td>
<td>2716.5608</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2716.6246</td>
<td>2716.6572</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2716.674</td>
<td>2716.6825</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>two iterations from full model</td>
<td>2716.6914</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2716.6914</td>
<td>2716.6914</td>
<td></td>
</tr>
</tbody>
</table>

#### Estimating Full Model

Iteration 0: log likelihood = 2716.6899
Iteration 1: log likelihood = 2716.6914
Iteration 2: log likelihood = 2716.6914
Conclusion

- A proportionality constraint means that the effects of a group of variables changes, but that the relative differences in the sizes of the effect remain constant.
Conclusion

▶ A proportionality constraint means that the effects of a group of variables changes, but that the relative differences in the sizes of the effect remain constant.

▶ This can be of interest in its own right, e.g. the effect on child’s education of father’s and mother’s status change, but the relative contribution of each parent can remain constant.
Conclusion

▶ A proportionality constraint means that the effects of a group of variables changes, but that the relative differences in the sizes of the effect remain constant.

▶ This can be of interest in its own right, e.g. the effect on child’s education of father’s and mother’s status change, but the relative contribution of each parent can remain constant.

▶ It can also be interpreted in terms of a latent variable, e.g. father’s and mother’s status both measure family status.
Conclusion

▶ A proportionality constraint means that the effects of a group of variables changes, but that the relative differences in the sizes of the effect remain constant.

▶ This can be of interest in its own right, e.g. the effect on child’s education of father’s and mother’s status change, but the relative contribution of each parent can remain constant.

▶ It can also be interpreted in terms of a latent variable, e.g. father’s and mother’s status both measure family status.

▶ Standard ml can have a hard time converging, so starting values are created using a EM algorithm.