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

example

Hypothesis:

Effect of father's and mother's socioeconomic status on child's education can change over cohorts,

example

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.

example

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.

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

example

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.

$$ed = \beta_0 + \beta_1 coh + (1 + \lambda_1 0)(\gamma_1 pasei + \gamma_2 masei)$$

example

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.

$$ed = \beta_0 + \beta_1 coh + (1 + \lambda_1 0)(\gamma_1 pasei + \gamma_2 masei)$$

$$ed = \beta_0 + \beta_1 coh + \gamma_1 pasei + \gamma_2 masei$$

example

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.

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

example

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.

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

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

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 male.
- 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.

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]
-------------+----------------------------------------------------------------
unconstrai~d |
         byr |   .0392033   .1418648     0.28   0.782    -.2388465    .3172531
       _cons |    10.2406   .2762536    37.07   0.000     9.699157    10.78205
-------------+----------------------------------------------------------------
constrained  |
       masei |   3.363018   .3688164     9.12   0.000     2.640152    4.085885
       pasei |   3.948723   .3972388     9.94   0.000     3.170149    4.727296
-------------+----------------------------------------------------------------
lambda       |
         byr |  -.0323712    .037854    -0.86   0.392    -.1065637    .0418212
       _cons |          1          .        .       .            .           .
-------------+----------------------------------------------------------------
ln_sigma     |
       _cons |    .837853    .014199    59.01   0.000     .8100234    .8656826
------------------------------------------------------------------------------
LR test vs. unconstrained model: chi2(1) =      0.04   Prob > chi2 =     0.849
```

alternative way of looking

$$ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) \underbrace{(\gamma_1 pasei + \gamma_2 masei)}_{\text{latent family sei}}$$

alternative way of looking

$$ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) \underbrace{(\gamma_1 pasei + \gamma_2 masei)}_{\text{latent family sei}}$$

- Need to identify the latent variable by fixing the origin and the scale.

alternative way of looking

$$ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) \underbrace{(\gamma_1 pasei + \gamma_2 masei)}_{\text{latent family sei}}$$

- 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.

alternative way of looking

$$ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) \underbrace{(\gamma_1 pasei + \gamma_2 masei)}_{\text{latent family sei}}$$

- 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.
  latent family sei $= \gamma_1 pasei + \gamma_2 masei$

alternative way of looking

$$ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) \underbrace{(\gamma_1 pasei + \gamma_2 masei)}_{\text{latent family sei}}$$

- 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.
  latent family sei $= \gamma_1 0 + \gamma_2 0 = 0$

alternative way of looking

$$ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) \underbrace{(\gamma_1 pasei + \gamma_2 masei)}_{\text{latent family sei}}$$

- 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

alternative way of looking

$$ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) \underbrace{(\gamma_1 pasei + \gamma_2 masei)}_{\text{latent family sei}}$$

- 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

$$ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) \underbrace{(\gamma_1 pasei + \gamma_2 masei)}_{\text{latent family sei}}$$

- 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

$$ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) \underbrace{(\gamma_1 pasei + \gamma_2 masei)}_{\text{latent family sei}}$$

- 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 $= \gamma_1 pasei + \gamma_2 masei$

alternative way of looking

$$ed = \beta_0 + \beta_1 coh + (\lambda_0 + \lambda_1 coh) \underbrace{(\gamma_1 pasei + \gamma_2 masei)}_{\text{latent family sei}}$$

- 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 $= \gamma_1 1 + \gamma_2 1 = 1$

example output

```
. propcnsreg degree byr, lambda(byr) constrained(masei pasei) unit(masei pasei)
Constraint: [constrained]masei + [constrained]pasei = 1
------------------------------------------------------------------------------
      degree |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
unconstrai~d |
         byr |   .0392033   .1418647     0.28   0.782    -.2388464    .3172529
       _cons |    10.2406   .2762534    37.07   0.000     9.699158    10.78205
-------------+----------------------------------------------------------------
constrained  |
       masei |   .4599477   .0323745    14.21   0.000     .3964949    .5234005
       pasei |   .5400523   .0323745    16.68   0.000     .4765995    .6035051
-------------+----------------------------------------------------------------
lambda       |
         byr |  -.2366899   .2935214    -0.81   0.420    -.8119814    .3386015
       _cons |   7.311741    .601956    12.15   0.000     6.131929    8.491553
-------------+----------------------------------------------------------------
ln_sigma     |
       _cons |    .837853    .014199    59.01   0.000     .8100234    .8656826
------------------------------------------------------------------------------
LR test vs. unconstrained model: chi2(1) =      0.04   Prob > chi2 =     0.849
```

scale for a categorical variable

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*)

scale for a categorical variable

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*)

$$isei = \beta_0 + (\lambda_0 + \lambda_1 black)(\gamma_1 hs + \gamma_2 sc + \gamma_3 c + \gamma_4 g)$$

scale for a categorical variable

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*)

$$isei = \beta_0 + (\lambda_0 + \lambda_1 black)(\gamma_1 hs + \gamma_2 sc + \gamma_3 c + 1 g)$$

scale for a categorical variable

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*)

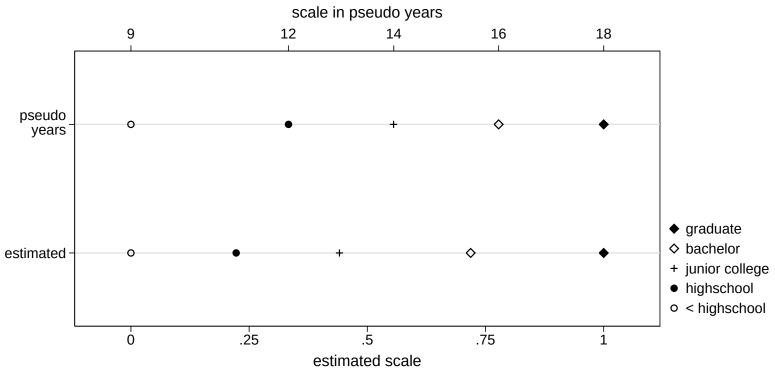
$$isei = \beta_0 + (\lambda_0 + \lambda_1 black)(\gamma_1 hs + \gamma_2 sc + \gamma_3 c + 1g)$$

γ_1, γ_2, and γ_3 now measure the position of highschool, some college, and college education, relative to less than highschool (0) and graduate (1).

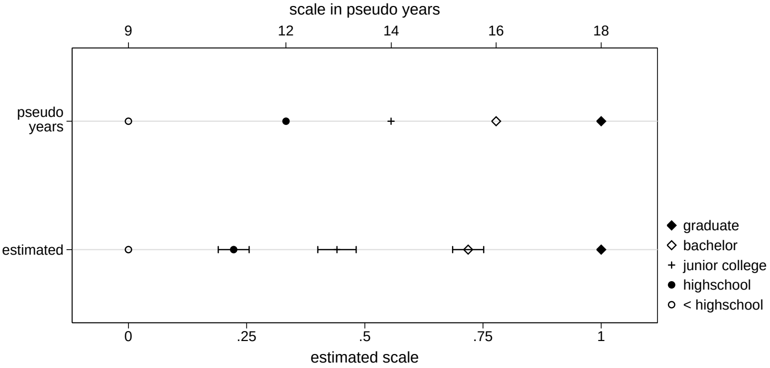
example output

```
. propcnsreg sei black, lambda(black) constrained(hs sc c g) unit(g)
Constraint: [constrained]g = 1
------------------------------------------------------------------------------
         sei |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
unconstrai~d |
       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
```

Scaling of education



Scaling of education



Outline

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

estimation

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 γ, 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 γ, create a new variable containing the latent variable. This simplifies the problem to: $y = \beta_0 + \beta_1 x_1 + \lambda_0 latent + \lambda_1 x_1 latent$

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 γ, create a new variable containing the latent variable. This simplifies the problem to: $y = \beta_0 + \beta_1 x_1 + \lambda_0 latent + \lambda_1 x_1 latent$
2. Estimate β and λ using `regress`

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 γ, create a new variable containing the latent variable. This simplifies the problem to: $y = \beta_0 + \beta_1 x_1 + \lambda_0 latent + \lambda_1 x_1 latent$
2. Estimate β and λ using `regress`
3. Given current estimates of λ, 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 γ, create a new variable containing the latent variable. This simplifies the problem to: $y = \beta_0 + \beta_1 x_1 + \lambda_0 latent + \lambda_1 x_1 latent$
2. Estimate β and λ using `regress`
3. Given current estimates of λ, create a new variable containing the effect of the latent variable. This simplifies the problem to: $y = \beta_0 + \beta_1 x_1 + \gamma_1 effectz_1 + \gamma_2 effectz_2)$

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 γ, create a new variable containing the latent variable. This simplifies the problem to: $y = \beta_0 + \beta_1 x_1 + \lambda_0 latent + \lambda_1 x_1 latent$
2. Estimate β and λ using `regress`
3. Given current estimates of λ, create a new variable containing the effect of the latent variable. This simplifies the problem to: $y = \beta_0 + \beta_1 x_1 + \gamma_1 effectz_1 + \gamma_2 effectz_2)$
4. Estimate β and γ using `cnsreg` imposing the constraint specified in the `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 γ, create a new variable containing the latent variable. This simplifies the problem to: $y = \beta_0 + \beta_1 x_1 + \lambda_0 latent + \lambda_1 x_1 latent$
2. Estimate β and λ using `regress`
3. Given current estimates of λ, create a new variable containing the effect of the latent variable. This simplifies the problem to: $y = \beta_0 + \beta_1 x_1 + \gamma_1 effectz_1 + \gamma_2 effectz_2)$
4. Estimate β and γ using `cnsreg` imposing the constraint specified in the `unit` option.
5. Repeat steps 1-4 till convergence.

speed and standard errors

speed and standard errors

- To speed up convergence every 5^{th} iteration will consist of two `ml` iterations for the complete model.

speed and standard errors

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

example iteration log

```
improving starting values
------------------------------------------------------
iteration  unconstrained  constrained   full model
           part only      part only
------------------------------------------------------
1           2712.7047      2716.1367
2           2716.4376      2716.5608
3           2716.6246      2716.6572
4            2716.674      2716.6825
------------------------------------------------------
5          two iterations from full model
                                         2716.6914
------------------------------------------------------
6           2716.6914      2716.6914
------------------------------------------------------


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 it's 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 it's 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 it's 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.