Using ADNI data for Cusp model fitting

Simple way

Auditory Verbal Learning Test fitted with Whole gray matter and covariables.

```
 # Estevez-Gonzalez, A., Kulisevsky, J., Boltes, A., Otermin, P., & Garcia-
Sanchez, C. (2003).
 # Rey verbal learning test is a useful tool for differential diagnosis in
the preclinical phase
 # of Alzheimer's disease: comparison with mild cognitive impairment and
normal aging.
 # International Journal of Geriatric Psychiatry. 18 (11), 1021.
```

```
library("ADNIMERGE")
library(cusp)
library(psych) #for composite scores
# Let's get the data
tmp_np <- merge(adas, neurobat, by=c("RID", "VISCODE") )
mt2fa <- merge(tmp_np, adnimerge, by=c("RID", "VISCODE") )
rm(tmp_np)
# Calculate the subject age at every point
mt2fa$vAGE = mt2fa$AGE + mt2fa$Years
data <- data.frame(mt2fa$WholeBrain, mt2fa$ICV, mt2fa$vAGE, mt2fa$PTGENDER,
mt2fa$PTEDUCAT, mt2fa$AVDEL30MIN, mt2fa$AVDELTOT)
datac <- data[complete.cases(data),]
datac$WB = datac$mt2fa.WholeBrain/datac$mt2fa.ICV
fit_avd <- cusp(y ~ mt2fa.AVDEL30MIN, alpha ~ WB +mt2fa.vAGE +
mt2fa.PTGENDER +mt2fa.PTEDUCAT, beta ~ WB +mt2fa.vAGE + mt2fa.PTGENDER
+mt2fa.PTEDUCAT, datac)
summary(fit_avd)
```

Amazing results

```
Call:
cusp(formula = y ~ mt2fa.AVDEL30MIN, alpha = alpha ~ WB + mt2fa.vAGE +
    mt2fa.PTGENDER + mt2fa.PTEDUCAT, beta = beta ~ WB + mt2fa.vAGE +
    mt2fa.PTGENDER + mt2fa.PTEDUCAT, data = datac)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-2.06955  -0.26210  -0.03226   0.63723   3.40775

Coefficients:
                         Estimate Std. Error  z value Pr(>|z|)
a[(Intercept)]          -5.842305   0.325414  -17.953  < 2e-16 ***
a[WB]                    5.365145   0.296437   18.099  < 2e-16 ***
```

```
a[mt2fa.vAGE]             0.004777   0.002000    2.388   0.0169 *
a[mt2fa.PTGENDERFemale]   0.364218   0.026875   13.552  < 2e-16 ***
a[mt2fa.PTEDUCAT]         0.078235   0.005217   14.995  < 2e-16 ***
b[(Intercept)]            7.001814   0.509934   13.731  < 2e-16 ***
b[WB]                    -5.970685   0.470236  -12.697  < 2e-16 ***
b[mt2fa.vAGE]            -0.026695   0.003309   -8.069 7.11e-16 ***
b[mt2fa.PTGENDERFemale]   0.395185   0.044766    8.828  < 2e-16 ***
b[mt2fa.PTEDUCAT]         0.033672   0.008002    4.208 2.58e-05 ***
w[(Intercept)]           -1.763659   0.012301 -143.381  < 2e-16 ***
w[mt2fa.AVDEL30MIN]       0.257004   0.001838  139.800  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


  Null deviance:   8677.7  on 6754  degrees of freedom
Linear deviance: 110910.4  on 6749  degrees of freedom
Logist deviance:       NA  on   NA  degrees of freedom
 Delay deviance:   3610.5  on 6743  degrees of freedom

             R.Squared     logLik npar      AIC     AICc      BIC
Linear model 0.1557933 -19036.661    6 38085.32 38085.33 38126.23
Cusp model   0.6142339  -7321.477   12 14666.95 14667.00 14748.77
---
Note: R.Squared for cusp model is Cobb's pseudo-R^2. This value
      can become negative.

    Chi-square test of linear vs. cusp model

X-squared = 2.343e+04, df = 6, p-value = 0

Number of optimization iterations: 40
```

Z-scores

Now let's compare the weights of each variable on the model. We need to translate everything to z-scores (or just do another linear transformation that carry every thing to comparable values)

```
datac$zWB = (datac$WB - mean(datac$WB))/sd(datac$WB)
datac$zAge = (datac$mt2fa.vAGE -
mean(datac$mt2fa.vAGE))/sd(datac$mt2fa.vAGE)
datac$zEduc = (datac$mt2fa.PTEDUCAT -
mean(datac$mt2fa.PTEDUCAT))/sd(datac$mt2fa.PTEDUCAT)
datac$zAVD = (datac$mt2fa.AVDEL30MIN -
mean(datac$mt2fa.AVDEL30MIN))/sd(datac$mt2fa.AVDEL30MIN)
fit_avd_z <- cusp(y ~ zAVD, alpha ~ zWB + zAge + mt2fa.PTGENDER + zEduc,
beta ~ zWB +zAge + mt2fa.PTGENDER + zEduc, datac)
summary(fit_avd_z)
```

The results are of course the same but the coefficients must be meaningful now,

```
Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
a[(Intercept)]          -0.708321   0.023240 -30.478  < 2e-16 ***
a[zWB]                   0.290456   0.016048  18.099  < 2e-16 ***
a[zAge]                  0.034254   0.014342   2.388   0.0169 *
a[mt2fa.PTGENDERFemale]  0.364218   0.026875  13.552  < 2e-16 ***
a[zEduc]                 0.223024   0.014873  14.995  < 2e-16 ***
b[(Intercept)]           1.603759   0.046739  34.313  < 2e-16 ***
b[zWB]                  -0.323238   0.025457 -12.697  < 2e-16 ***
b[zAge]                 -0.191434   0.023726  -8.069 7.11e-16 ***
b[mt2fa.PTGENDERFemale]  0.395185   0.044766   8.828  < 2e-16 ***
b[zEduc]                 0.095988   0.022812   4.208 2.58e-05 ***
w[(Intercept)]          -0.688350   0.009603 -71.682  < 2e-16 ***
w[zAVD]                  1.133500   0.008108 139.800  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Composite scores

First I'm going to try another NP test (Recognition)

```
fit_avr <- cusp(y ~ zAVR, alpha ~ zWB + zAge + mt2fa.PTGENDER + zEduc, beta
~ zWB +zAge + mt2fa.PTGENDER + zEduc, datac)
```

and this is not so good but still an improvement is done

```
> summary(fit_avr)
Call:
cusp(formula = y ~ zAVR, alpha = alpha ~ zWB + zAge + mt2fa.PTGENDER +
    zEduc, beta = beta ~ zWB + zAge + mt2fa.PTGENDER + zEduc,
    data = datac)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.3337  -0.8146  -0.1929   0.2446   2.5763

Coefficients:
                        Estimate Std. Error z value Pr(>|z|)
a[(Intercept)]           0.74119    0.02784  26.625  < 2e-16 ***
a[zWB]                   0.40520    0.01943  20.854  < 2e-16 ***
a[zAge]                  0.08352    0.01681   4.967 6.79e-07 ***
a[mt2fa.PTGENDERFemale] -0.09332    0.03128  -2.983  0.00285 **
a[zEduc]                 0.10035    0.01495   6.713 1.91e-11 ***
b[(Intercept)]           1.09138    0.05476  19.932  < 2e-16 ***
b[zWB]                   0.02940    0.02921   1.007  0.31416
b[zAge]                  0.11858    0.02729   4.345 1.39e-05 ***
```

```
b[mt2fa.PTGENDERFemale]  0.53540    0.04991  10.726  < 2e-16 ***
b[zEduc]                 0.11832    0.02474   4.782 1.74e-06 ***
w[(Intercept)]           0.71871    0.01153  62.322  < 2e-16 ***
w[zAVR]                  0.99105    0.00889 111.482  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


  Null deviance: 6633.7  on 6754  degrees of freedom
Linear deviance: 5952.1  on 6749  degrees of freedom
Logist deviance:     NA  on   NA  degrees of freedom
 Delay deviance: 5049.6  on 6743  degrees of freedom

             R.Squared    logLik npar      AIC     AICc      BIC
Linear model 0.1187225 -9157.572    6 18327.14 18327.16 18368.05
Cusp model   0.3758839 -7966.518   12 15957.04 15957.08 16038.85
---
Note: R.Squared for cusp model is Cobb's pseudo-R^2. This value
      can become negative.

    Chi-square test of linear vs. cusp model

X-squared = 2382, df = 6, p-value = 0

Number of optimization iterations: 38
```

Now, let's try a composite score

```
gfam <- data.frame(datac$zAVD, datac$zAVR)
famod <- fa(gfam, scores="regression")
datac$cs <- famod$scores
fit_cs <- cusp(y ~ cs, alpha ~ zWB + zAge + mt2fa.PTGENDER + zEduc, beta ~
zWB +zAge + mt2fa.PTGENDER + zEduc, datac)
```

And we get very bad fit result

```
> summary(fit_cs)

Call:
cusp(formula = y ~ cs, alpha = alpha ~ zWB + zAge + mt2fa.PTGENDER +
    zEduc, beta = beta ~ zWB + zAge + mt2fa.PTGENDER + zEduc,
    data = datac)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.9864  -0.5145   0.0386   0.5796   2.8034

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
```

```
a[(Intercept)]          -0.166529   0.015351 -10.848  < 2e-16 ***
a[zWB]                   0.508523   0.009612  52.905  < 2e-16 ***
a[zAge]                  0.124850   0.017119   7.293 3.03e-13 ***
a[mt2fa.PTGENDERFemale]  0.251292   0.020108  12.497  < 2e-16 ***
a[zEduc]                 0.230918         NA      NA       NA
b[(Intercept)]          -0.224522   0.022933  -9.791  < 2e-16 ***
b[zWB]                  -0.231656   0.010449 -22.171  < 2e-16 ***
b[zAge]                 -0.128716   0.012640 -10.183  < 2e-16 ***
b[mt2fa.PTGENDERFemale]  0.845789   0.017774  47.587  < 2e-16 ***
b[zEduc]                 0.204127   0.007479  27.293  < 2e-16 ***
w[(Intercept)]          -0.035730   0.011210  -3.187  0.00144 **
w[cs]                    1.012113   0.006090 166.187  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


  Null deviance: 5487.3  on 6754  degrees of freedom
Linear deviance: 4490.5  on 6749  degrees of freedom
Logist deviance:     NA  on   NA  degrees of freedom
 Delay deviance: 5324.3  on 6743  degrees of freedom

              R.Squared    logLik npar      AIC     AICc      BIC
Linear model 0.16171127 -8205.808    6 16423.62 16423.63 16464.52
Cusp model   0.03112304 -8567.331   12 17158.66 17158.71 17240.48
---
Note: R.Squared for cusp model is Cobb's pseudo-R^2. This value
      can become negative.

    Chi-square test of linear vs. cusp model

X-squared = 723, df = 6, p-value = 0

Number of optimization iterations: 106
```

That is, the composite score is not related through a cusp model to the independent variable analyzed here