Model building

Time-use variables (D modifiable predictors)

Transform the time-use data into ilr coordinates

$$\boldsymbol{z} = V^T \ln(\boldsymbol{x})$$

Multivariate scale ("whiten") *ilr* coordinates $\boldsymbol{z}$ using

$$\boldsymbol{z}^{(s)} = (\boldsymbol{z} - \hat{\boldsymbol{\mu}}_z)^T \hat{\Sigma}_z^{-1/2}$$

Other covariates (other participant predictors)

Re-parameterise categorical variables as contrast indicators/"one-hot-coded" w.r.t. to a reference level

Multivariate scale each covariate group-wise to create block-wise, main-effect design matricies

$$\boldsymbol{w}^{(s)} = (\boldsymbol{w} - \hat{\boldsymbol{\mu}}_w)^T \hat{\Sigma}_w^{-1/2}$$

Create all two way interactions as well as squared polynomial terms (for continuous variables only) for the set of multidimensional scaled ilr-coordinates and other group/block-wise scaled predictor variables. Resulting in the group-wise set of predictor matricies $X^{(g)}$, $g = 1, 2, ..., G$; where G is the total number of variable groups (interaction/polynomial blocks are seperate predictor groups to main effects)

Cognative variables (outcomes)

$$\boldsymbol{y}_1, \ldots, \boldsymbol{y}_5$$

Univariate scale cognative response variables

$$\boldsymbol{y}_1^{(s)}, \ldots, \boldsymbol{y}_5^{(s)}$$

Fit Group Lasso model for each scaled cognitive outcome (response variable) $\boldsymbol{y}_r^{(s)}$, $r = 1, 2, \ldots, 5$:

$$\boldsymbol{\beta}_r^* = \underset{\boldsymbol{\beta}_r}{\operatorname{argmin}} \left|\left| \boldsymbol{y}_r^{(s)} - \sum_{g=1}^{G} X^{(g)} \boldsymbol{\beta}_r^{(g)} \right|\right|_2^2 + \lambda_r \sum_{g=1}^{G} \sqrt{\nu_g} \left|\left| \boldsymbol{\beta}_r^{(g)} \right|\right|_2$$

That is, find the regularised/penalised coefficients $\boldsymbol{\beta}_r^*$ for each response variable block-wise over the scaled main, interaction and polynomial (order 2, continuous only) predictors

Minimise 10-fold cross-validation prediction error over grid of group lasso penaltys to find optimised λ_r^* for each response variable r

User side

Personalised optimisation

Specify response r of interest

Enter demographic info

Calculate the sample (*ilr*-transformed) mean vector and covariance matrix of the k^{th} strata, $\hat{\boldsymbol{\mu}}_k$ and $\hat{\Sigma}_k$, respectively

Person specific strata k can be deduced from the 2^3=8 strata defined by tuples of dichotimised categories (age$_{<75}$, sex, BMI$_{<30}$)

Estimate the surface ("fence") of an ellipsoid (3D generalisation of an ellipse in $D-1 = 3$-dimensional Euclidean *ilr*-space) that captures the 80% highest density of the *ilr*-transformed time-use values specific to the strata

- The fence is calculated assuming a 3-variate Gaussian distribution with points $\boldsymbol{z}^*$ that satisfy:

$$(\boldsymbol{z}^* - \hat{\boldsymbol{\mu}}_k)^T \hat{\Sigma}_k^{-1} (\boldsymbol{z}^* - \hat{\boldsymbol{\mu}}_k) \leq 4.641628$$

Generate exhaustive grid of all possible 5-min spaced time-use compositions in 24-hours and ilr transform

Restrict ilr transformed exhaustive grid to only points $\boldsymbol{z}_{(\#)}$ that lie within strata k's distributional ellipsoid "fence" and then multivariate scale these values as before to create $\boldsymbol{z}_{(\#)}^{(s)}$

Create all scaled main effects, two way interactions as well as squared polynomial terms like before using the personalised user predictors as covariates with the ilr grid points

Make predictions for response r from Lasso model ($\boldsymbol{\beta}_r^*$) over all constrained grid points

Extract top 5% of grid predctions and average associated ilrs to create the optimal ilr coords, $\boldsymbol{z}_{(opt)}^{(s)}$

Reverse transform ilr scaling on the optimal ilr and then inverse ilr tranform back to time-use composition scale, $\boldsymbol{x}_{(opt)}$

Report associated optimal time-use composition $\boldsymbol{x}_{(opt)}$ to the user