**LINEAR MODELING OF EEG DATA**

**Validation for limo\_glm1.m**

**Cyril Pernet**

[**cyril.pernet@ed.ac.uk**](mailto:cyril.pernet@ed.ac.uk)

**limo\_glm1.m : Two-samples t-tests, ANOVA, regression, ANCOVA**

All of those models can be performed using limo\_glm1.m. In short, one build a design matrix X which includes dummy variables (1/0) to code groups and continuous variables for covariates. If we have 2 groups only, this is equivalent to a two-samples t-test; if we have more than 2 groups, this is an ANOVA, if we have continuous regressors only, this is equivalent to a regression, and if we have a combination of dummy and continuous variables, this is equivalent to an ANCOVA.

All is needed is data (Y), the design matrix (X), vectors describing factors and levels (nb\_conditions, nb\_interactions, nb\_continuous) and the method use (here ‘OLS’). Note the design matrix is always built in the same order: 1st factors and levels, 2nd interaction terms if any, 3rd covariates and 4th the constant term.

* model = limo\_glm(Y, X, nb\_conditions, nb\_interactions, nb\_continuous, method)

We show here the results obtain with limo\_glm1 vs. Statistica® using ‘random’ data.

***2 independent conditions (ANOVA – two samples t-test)***

The data Y are

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Y | 5 | 6 | 8 | 7 | 9 | 3 | 2 | 1 | 5 | 6 |
| Gp | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 |

*Using limo\_glm1*

directory = pwd;

% create 3D data for limo\_design\_matrix

Y = NaN(1,1,10);

Y(1,1,:) = [5,6,8,7,9,3,2,1,5,6]';

% specify categorical and continuous variables

Cat = [1 1 1 1 1 2 2 2 2 2]';

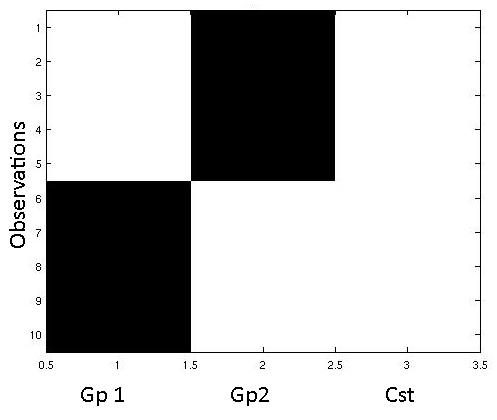
Cont = [];

% create the design matrix

%[X,nb\_conditions,nb\_interactions,nb\_continuous] = limo\_design\_matrix(Y,Cat,Cont,directory,zscoring,full\_factorial,flag)

[X,nb\_conditions,nb\_interactions,nb\_continuous] =limo\_design\_matrix(Y,Cat,Cont,directory,1,0,1);

We obtain the following design matrix



% compute the ANOVA sending 1D data

model = limo\_glm1(squeeze(Y),X,nb\_conditions,nb\_interactions,nb\_continuous,'OLS');

model.conditions.F = 9.5294

model.conditions.df = [1 8]

model.conditions.p = 0.015

Statistica® results are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **SS** | **Df** | **MS** | **F** | **p** |
| **Gp** | 32.4000 | 1 | 32.4000 | 9.52941 | 0.014958 |
| **Error** | 27.2000 | 8 | 3.4000 |  |  |

***3 independent conditions of different size (ANOVA)***

The data Y are

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Y | 5 | 6 | 8 | 7 | 9 | 3 | 2 | 1 | 5 | 6 | 4 | 5 | 8 | 9 | 6 |
| Gp | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 |

*Using limo\_glm1.m*

directory = pwd;

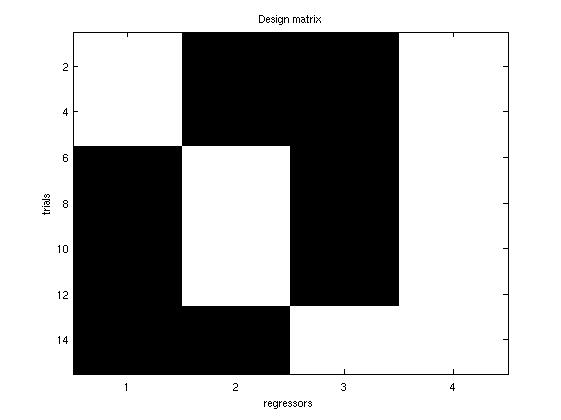
Y=NaN(1,1,15);

Y(1,1,:) = [5,6,8,7,9,3,2,1,5,6,4,5,8,9,6];

Cat = [1 1 1 1 1 2 2 2 2 2 2 2 3 3 3];

Cont = 0;

[X,nb\_conditions,nb\_interactions,nb\_continuous] =limo\_design\_matrix(Y,Cat,Cont,directory,1,0,1);

we obtain the following design matrix

model = limo\_glm1(squeeze(Y),X,nb\_conditions,nb\_interactions,nb\_continuous,'OLS');

model.conditions.F = 8.3598

model.conditions.df = [2 12]

model.conditions.p = 0.0053

Statistica® results are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **SS** | **Df** | **MS** | **F** | **p** |
| **Gp** | 47.5048 | 2 | 3.7524 | 8.3598 | 0.005321 |
| **Error** | 27.2000 | 12 | 2.8413 |  |  |

***Single continuous variable (Simple Regression)***

The data Y are

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Y | 5 | 6 | 8 | 7 | 9 | 3 | 2 | 1 | 5 | 6 | 4 | 5 | 8 | 9 | 6 |
| Cov | 0.1978 | 1.3107 | 0.5688 | -0.5441 | -1.286 | -0.915 | -0.1731 | 0.1978 | 0.5688 | 0.9398 | 1.6817 | -0.915 | 0.9398 | -1.286 | -1.286 |

directory = pwd;

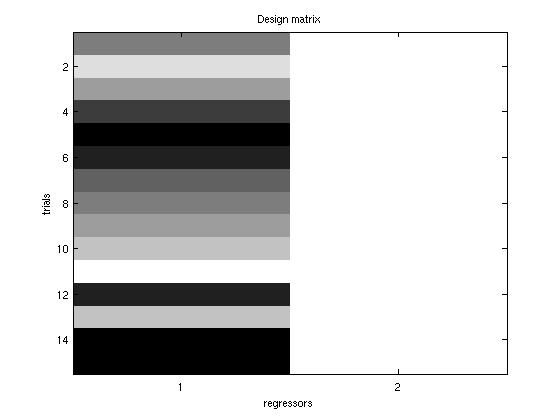
Y=NaN(1,1,15);

Y(1,1,:) = [5,6,8,7,9,3,2,1,5,6,4,5,8,9,6];

Cat = []

Cont = [0.1978 1.3107 0.5688 -0.5441 -1.286 -0.915 -0.1731 0.1978 0.5688 0.9398 1.6817 -0.915 0.9398 -1.286 -1.286]

[X,nb\_conditions,nb\_interactions,nb\_continuous] =limo\_design\_matrix(Y,Cat,Cont,directory,1,0,1);

we obtain the following design matrix

model = limo\_glm1(squeeze(Y),X,nb\_conditions,nb\_interactions,nb\_continuous,'OLS');

model.R2\_univariate = 0.0316

model.F=0.4244 (same as model.continuous.F)

model.df = [1 13] (same as model.continuous.df)

model.p=0.5261 (same as model.continuous.p)

Statistica® results are:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **R2** | | | **Df** | | **Dfe** | **F** | | **p** | |
| **R2** | | 0.031613 | | | 1 | | 13 | 0.42439 | | 0.526105 | |
|  | | **SS** | **Df** | | **MS** | | | **F** | | **p** | |
| **Cov1** | | 2.57 | 1 | | 2.5796 | | | 0.42439 | | 0.526105 | |
| **Error** | | 79.0204 | 13 | | 6.0785 | | |  | |  | |

***Multiple continuous variables (Multiple regression)***

The data Y are

|  |  |  |
| --- | --- | --- |
| Y | Cov1 | Cov2 |
| 5 | 0.1978 | -1.0185 |
| 6 | 1.3107 | -0.3542 |
| 8 | 0.5688 | 0.31 |
| 7 | -0.5441 | -0.3542 |
| 9 | -1.286 | -1.0185 |
| 3 | -0.915 | 0.9742 |
| 2 | -0.1731 | -0.3542 |
| 1 | 0.1978 | 0.31 |
| 5 | 0.5688 | 2.3026 |
| 6 | 0.9398 | 1.6384 |
| 4 | 1.6817 | -0.3542 |
| 5 | -0.915 | -1.0185 |
| 8 | 0.9398 | -1.0185 |
| 9 | -1.286 | -0.3542 |
| 6 | -1.286 | 0.31 |

directory = pwd;

Y=NaN(1,1,15);

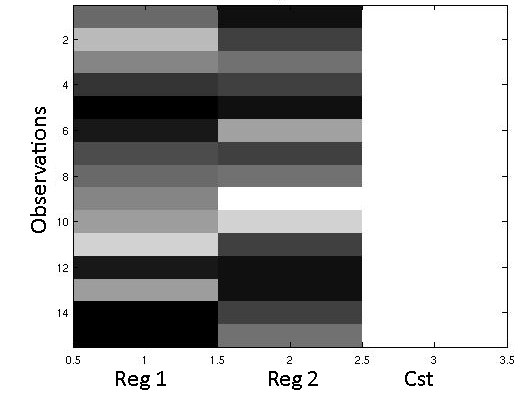
Y(1,1,:) = [5,6,8,7,9,3,2,1,5,6,4,5,8,9,6];

Cat = []

Cont = [0.1978 1.3107 0.5688 -0.5441 -1.286 -0.915 -0.1731 0.1978 0.5688 0.9398 1.6817 -0.915 0.9398 -1.286 -1.286 ; -1.0185 -0.3542 0.31 -0.3542 -1.0185 0.9742 -0.3542 0.31 2.3026 1.6384 -0.3542 -1.0185 -1.0185 -0.3542 0.31]';

[X,nb\_conditions,nb\_interactions,nb\_continuous] =limo\_design\_matrix(Y,Cat,Cont,directory,1,0,1);

we obtain the following design matrix



model = limo\_glm1(squeeze(Y),X,nb\_conditions,nb\_interactions,nb\_continuous,'OLS');

Results are

model.R2\_univariate = 0.0814

model.F=0.5315

model.df = [2 12]

model.p=0.6009

model.continuous.F = [0.2363 0.6501]

model.continuous.p = [ 0.6355 0.4358]

model.continuous.df = [1 12]

Statistica® results are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Multiple** | **Df** | **Dfe** | **F** | **p** |
| **R2** | 0.081383 | 2 | 12 | 0.5315 | 0.60090 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **SS** | **Df** | **MS** | **F** | **p** |
| **Cov1** | 1.4775 | 1 | 1.4775 | 0.23653 | 0.635482 |
| **Cov1** | 4.0612 | 1 | 4.0612 | 0.65014 | 0.435751 |
| **Error** | 74.9592 | 12 | 6.2466 |  |  |

***3 groups (different sample size) and 2 continuous variables (ANCOVA)***

The data Y are

|  |  |  |  |
| --- | --- | --- | --- |
| Y | Gp | Cov1 | Cov2 |
| 5 | 1 | 0.1978 | -1.0185 |
| 6 | 1 | 1.3107 | -0.3542 |
| 8 | 1 | 0.5688 | 0.31 |
| 7 | 1 | -0.5441 | -0.3542 |
| 9 | 1 | -1.286 | -1.0185 |
| 3 | 2 | -0.915 | 0.9742 |
| 2 | 2 | -0.1731 | -0.3542 |
| 1 | 2 | 0.1978 | 0.31 |
| 5 | 2 | 0.5688 | 2.3026 |
| 6 | 2 | 0.9398 | 1.6384 |
| 4 | 2 | 1.6817 | -0.3542 |
| 5 | 2 | -0.915 | -1.0185 |
| 8 | 3 | 0.9398 | -1.0185 |
| 9 | 3 | -1.286 | -0.3542 |
| 6 | 3 | -1.286 | 0.31 |

*Using limo\_glm1.m*

directory = pwd;

Y=NaN(1,1,15);

Y(1,1,:) = [5,6,8,7,9,3,2,1,5,6,4,5,8,9,6];

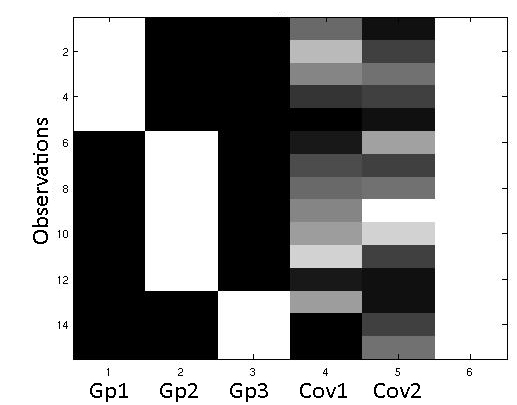
Cat = [1 1 1 1 1 2 2 2 2 2 2 2 3 3 3];

Cont = [0.1978 1.3107 0.5688 -0.5441 -1.286 -0.915 -0.1731 0.1978 0.5688 0.9398 1.6817 -0.915 0.9398 -1.286 -1.286 ; -1.0185 -0.3542 0.31 -0.3542 -1.0185 0.9742 -0.3542 0.31 2.3026 1.6384 -0.3542 -1.0185 -1.0185 -0.3542 0.31]';

[X,nb\_conditions,nb\_interactions,nb\_continuous] =limo\_design\_matrix(Y,Cat,Cont,directory,1,0,1);

model = limo\_glm1(squeeze(Y),X,nb\_conditions,nb\_interactions,nb\_continuous,'OLS');

we obtain the following design matrix



Results are

model.R2\_univariate = 0.5986

model.F = 3.7277

model.p = 0.0416

model.df = [4 10]

model.conditions.F = 6.4416

model.continuous.p=0.0159

model.continuous.df = [2 10]

model.continuous.F = [0.01 74 0.4052]

model.continuous.p = [0.8978 0.5387]

model.continuous.df = [1 10]

Statistica® results are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Multiple** | **Df** | **Dfe** | **F** | **p** |
| **Model** | 0.598562 | 4 | 10 | 3.727608 | 0.041627 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **SS** | **Df** | **MS** | **F** | **p** |
| **Cov1** | 0.0569 | 1 | 0.0569 | 0.0174 | 0.897759 |
| **Cov2** | 1.3272 | 1 | 1.3272 | 0.4052 | 0.538737 |
| **Gp** | 42.2018 | 2 | 21.1009 | 6.4416 | 0.015938 |
| **Error** | 32.7574 | 10 | 3.2757 |  |  |

***3\*3 independent conditions of different sizes and no interaction (2-way ANOVA)***

The data Y are

|  |  |  |
| --- | --- | --- |
| Y | V1 | V2 |
| 5 | 1 | 1 |
| 6 | 1 | 1 |
| 8 | 1 | 2 |
| 7 | 1 | 2 |
| 9 | 1 | 3 |
| 3 | 2 | 1 |
| 2 | 2 | 1 |
| 1 | 2 | 2 |
| 5 | 2 | 2 |
| 6 | 2 | 3 |
| 4 | 2 | 3 |
| 5 | 2 | 3 |
| 8 | 3 | 1 |
| 9 | 3 | 2 |
| 6 | 3 | 3 |

Using LIMO,

directory = pwd;

Y=NaN(1,1,15);

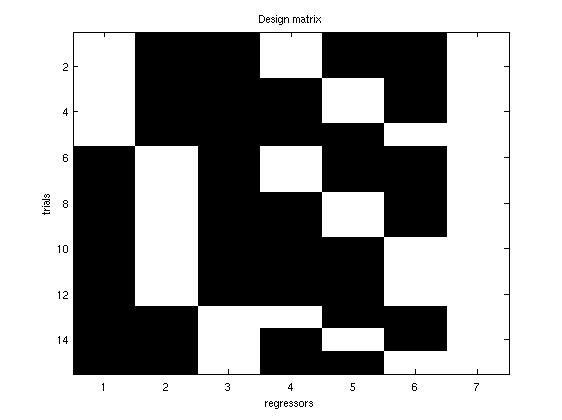
Y(1,1,:) = [5,6,8,7,9,3,2,1,5,6,4,5,8,9,6];

Cat = [1 1 1 1 1 2 2 2 2 2 2 2 3 3 3 ; 1 1 2 2 3 1 1 2 2 3 3 3 1 2 3];

Cont = 0;

[X,nb\_conditions,nb\_interactions,nb\_continuous] =limo\_design\_matrix(Y,Cat,Cont,directory,1,0,1);

we obtain the following design matrix



the 1st three column are for the factor 1

the next three columns are for factor 2

model = limo\_glm1(squeeze(Y),X,nb\_conditions,nb\_interactions,nb\_continuous,'OLS');

Results are

model.conditions.F = [10.3755 1.8256]

model.conditions.df = [2 10]

model.conditions.p = [0.0036 0.2109]

Statistica® results are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **SS** | **Df** | **MS** | **F** | **p** |
| **V1** | 51.8251 | 2 | 25.9126 | 10.3755 | 0.003637 |
| **V2** | 9.1204 | 2 | 4.5602 | 1.8259 | 0.210885 |
| **Error** | 24.97 | 10 | 2.4975 |  |  |

***2-way ANCOVA with 3\*3 independent conditions of different sizes and no interaction***

The data Y are

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Y | V1 | V2 | Cov1 | Cov2 |
| 5 | 1 | 1 | 0.1978 | -1.0185 |
| 6 | 1 | 1 | 1.3107 | -0.3542 |
| 8 | 1 | 2 | 0.5688 | 0.31 |
| 7 | 1 | 2 | -0.5441 | -0.3542 |
| 9 | 1 | 3 | -1.286 | -1.0185 |
| 3 | 2 | 1 | -0.915 | 0.9742 |
| 2 | 2 | 1 | -0.1731 | -0.3542 |
| 1 | 2 | 2 | 0.1978 | 0.31 |
| 5 | 2 | 2 | 0.5688 | 2.3026 |
| 6 | 2 | 3 | 0.9398 | 1.6384 |
| 4 | 2 | 3 | 1.6817 | -0.3542 |
| 5 | 2 | 3 | -0.915 | -1.0185 |
| 8 | 3 | 1 | 0.9398 | -1.0185 |
| 9 | 3 | 2 | -1.286 | -0.3542 |
| 6 | 3 | 3 | -1.286 | 0.31 |

Using LIMO,

directory = pwd;

Y=NaN(1,1,15);

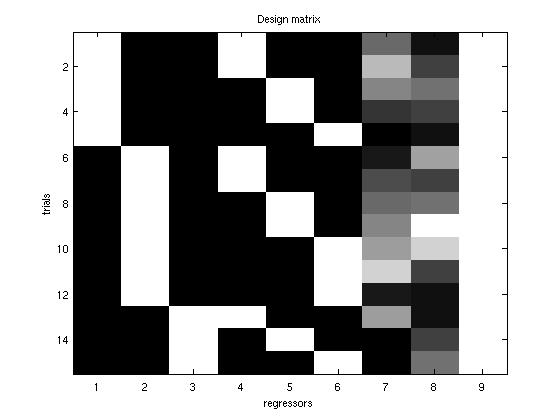
Y(1,1,:) = [5,6,8,7,9,3,2,1,5,6,4,5,8,9,6];

Cat = [1 1 1 1 1 2 2 2 2 2 2 2 3 3 3 ; 1 1 2 2 3 1 1 2 2 3 3 3 1 2 3];

Cont = [0.1978 1.3107 0.5688 -0.5441 -1.286 -0.915 -0.1731 0.1978 0.5688 0.9398 1.6817 -0.915 0.9398 -1.286 -1.286 ; -1.0185 -0.3542 0.31 -0.3542 -1.0185 0.9742 -0.3542 0.31 2.3026 1.6384 -0.3542 -1.0185 -1.0185 -0.3542 0.31]';

[X,nb\_conditions,nb\_interactions,nb\_continuous] =limo\_design\_matrix(Y,Cat,Cont,directory,1,0,1);

we obtain the following design matrix



the 1st three column are for the factor 1

the next three columns are for factor 2

the following columns are for the covariates

model = limo\_glm1(squeeze(Y),X,nb\_conditions,nb\_interactions,nb\_continuous,'OLS');

Results are

model.conditions.F = [7.2751 ; 1.4768]

model.conditions.df = [2 8 ; 2 8]

model.conditions.p = [0.0158 0.2845]

model.continuous.F = [0.0593 ; 0.2401]

model.continuous.p = [0.8138 ; 0.6373]

model.continuous.df = [1 8]

Statistica® results are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **SS** | **Df** | **MS** | **F** | **p** |
| **V1** | 43.51 | 2 | 21.7565 | 7.2751 | 0.01584 |
| **V2** | 8.83294 | 2 | 4.4165 | 1.4768 | 0.2845 |
| **Cov1** | 0.1773 | 1 | 0.1773 | 0.0593 | 0.8137 |
| **Cov2** | 0.7180 | 1 | 0.7120 | 0.2401 | 0.6372 |
| **Error** | 23.9245 | 8 | 2.9906 |  |  |

***2-way ANOVA with 2\*3 independent conditions of SAME sizes and the interaction***

When the interaction terms are present, interactions and main effects are orthogonal when the number of trials is the same for each level – LIMO doesn't deal with different n and instead, if full factorial is selected during the design matrix specification, the smallest n is used across levels taking randomly trials in each level making the full factorial ANOVA with equivalent n across levels. For the same reason, LIMO does not do ANCOVA with interaction terms.

The data Y are

|  |  |  |
| --- | --- | --- |
| Y | V1 | V2 |
| 5 | 1 | 1 |
| 6 | 1 | 2 |
| 8 | 1 | 3 |
| 7 | 1 | 1 |
| 9 | 1 | 2 |
| 3 | 1 | 3 |
| 2 | 2 | 1 |
| 1 | 2 | 2 |
| 5 | 2 | 3 |
| 6 | 2 | 1 |
| 4 | 2 | 2 |
| 5 | 2 | 3 |

If one runs a model without interaction (as above)

Using LIMO,

directory = pwd;

Y=NaN(1,1,12);

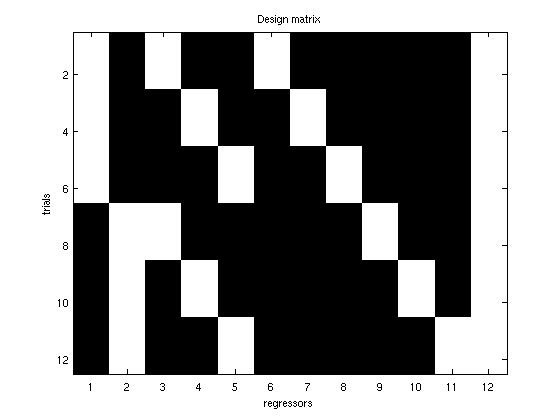
Y(1,1,:) = [5,6,8,7,9,3,2,1,5,6,4,5];

Cat = [1 1 1 1 1 1 2 2 2 2 2 2 ; 1 2 3 1 2 3 1 2 3 1 2 3 ];

Cont = 0;

[X,nb\_conditions,nb\_interactions,nb\_continuous] =limo\_design\_matrix(Y,Cat,Cont,directory,1,1,1); % note the flag for interaction is 1

we obtain the following design matrix



the 1st two column are for the factor 1

the next three columns are for factor 2

the following 6 columns is the interaction

load Yr % here because of the factor structure Y is reorganized

model = limo\_glm1(squeeze(Yr),X,nb\_conditions,nb\_interactions,nb\_continuous,'OLS')

Results are

model.conditions.F = [3.57 ; 0.1]

model.conditions.df = [1 6 ; 2 6]

model.conditions.p = [0.1 0.98]

model.interactions.F = 1

model.interactions.p = 0.4219

model.interactions.df = [2 6]

Statistica® results are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **SS** | **Df** | **MS** | **F** | **p** |
| **V1** | 18.75 | 1 | 18.75 | 3.57143 | 0.095452 |
| **V2** | 0.1667 | 2 | 0.0833 | 0.1587 | 0.984283 |
| **Error** | 42 | 8 | 5.25 |  |  |

If we add the interaction, Statistica® results are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **SS** | **Df** | **MS** | **F** | **p** |
| **V1** | 18.75 | 1 | 18.75 | 3.57143 | 0.107679 |
| **V2** | 0.1667 | 2 | 0.0833 | 0.01587 | 0.9842 |
| **V1\*V2** | 10.5 | 2 | 5.25 | 1 | 0.421875 |
| **Error** | 31.5 | 6 | 5.25 |  |  |

Because the interaction is orthogonal to the main effects and simply added to the model, the SS of each factor are unchanged, and the error term is now the same as before minus the SS interaction.

***3-way ANOVA with 3\*3\*2 independent conditions of 'different' sizes and the interaction***

As stated above, LIMO doesn't deal with unbalanced designs – we illustrate here how the design matrix specification steps sample the data.

If the data Y are

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Y | V1 | V2 | V3 | condition |
| 5 | 1 | 1 | 1 | A |
| 6 | 1 | 2 | 1 | B |
| 8 | 1 | 3 | 1 | C |
| 7 | 1 | 1 | 1 | A |
| 9 | 1 | 2 | 1 | B |
| 3 | 1 | 3 | 1 | C |
| 2 | 2 | 1 | 1 | D |
| 1 | 2 | 2 | 1 | E |
| 5 | 2 | 3 | 2 | F |
| 6 | 2 | 1 | 2 | G |
| 4 | 2 | 2 | 2 | H |
| 5 | 2 | 3 | 2 | F |
| 3 | 3 | 1 | 2 | I |
| 4 | 3 | 2 | 2 | J |
| 5 | 3 | 3 | 2 | K |
| 9 | 3 | 1 | 2 | I |
| 8 | 3 | 3 | 2 | K |
| 4 | 3 | 3 | 2 | K |

We have 11 conditions out of 3\*3\*2 = 18 possible; in this case it is impossible to create balance and LIMO will 1. return a message telling that it can't balance the data and 2. create a factorial design instead (without interaction)

By contrast, if the data that have all possible outcomes, LIMO\_design\_matrix simply sample randomly to create a 'new' set, however for the data Y =

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Y | V1 | V2 | V3 | condition |
| 5 | 1 | 1 | 1 | A |
| 6 | 1 | 2 | 1 | B |
| 8 | 1 | 3 | 1 | C |
| 7 | 1 | 1 | 2 | D |
| 9 | 1 | 2 | 2 | E |
| 3 | 1 | 3 | 2 | F |
| 2 | 2 | 1 | 1 | G |
| 1 | 2 | 2 | 1 | H |
| 5 | 2 | 3 | 1 | I |
| 6 | 2 | 1 | 2 | J |
| 4 | 2 | 2 | 2 | K |
| 5 | 2 | 3 | 2 | L |
| 3 | 3 | 1 | 1 | M |
| 4 | 3 | 2 | 1 | N |
| 5 | 3 | 3 | 1 | O |
| 9 | 3 | 1 | 2 | P |
| 8 | 3 | 2 | 2 | Q |
| 4 | 3 | 3 | 2 | R |
| 5 | 1 | 1 | 1 | A2 |
| 9 | 1 | 1 | 2 | D2 |
| 6 | 2 | 1 | 1 | G2 |
| 7 | 2 | 1 | 1 | G3 |
| 8 | 3 | 1 | 2 | P2 |

After sampling to create balance we have only 1 observation per condition like e.g.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Y | V1 | V2 | V3 | condition |
| 5 | 1 | 1 | 1 | A |
| 7 | 1 | 1 | 2 | D |
| 6 | 1 | 2 | 1 | B |
| 9 | 1 | 2 | 2 | E |
| 8 | 1 | 3 | 1 | C |
| 3 | 1 | 3 | 2 | F |
| 7 | 2 | 1 | 1 | G3 |
| 6 | 2 | 1 | 2 | J |
| 1 | 2 | 2 | 1 | H |
| 4 | 2 | 2 | 2 | K |
| 5 | 2 | 3 | 1 | I |
| 5 | 2 | 3 | 2 | L |
| 3 | 3 | 1 | 1 | M |
| 9 | 3 | 1 | 2 | P |
| 4 | 3 | 2 | 1 | N |
| 8 | 3 | 2 | 2 | Q |
| 5 | 3 | 3 | 1 | O |
| 4 | 3 | 3 | 2 | R |

in this case it is impossible to get any F or p values since df(error) = 0 and LIMO will 1. return a message telling that the design is over specified and 2. create a factorial design instead (without interaction)

Let's now try a 'good' unbalanced set (at least 2 observations per conditions)

|  |  |  |  |
| --- | --- | --- | --- |
| Y | V1 | V2 | V3 |
| 5 | 1 | 1 | 1 |
| 6 | 1 | 1 | 2 |
| 8 | 1 | 1 | 1 |
| 7 | 1 | 1 | 2 |
| 9 | 1 | 2 | 1 |
| 3 | 1 | 2 | 2 |
| 2 | 1 | 2 | 1 |
| 1 | 1 | 2 | 2 |
| 5 | 2 | 1 | 1 |
| 6 | 2 | 1 | 2 |
| 4 | 2 | 1 | 1 |
| 5 | 2 | 1 | 2 |
| 3 | 2 | 2 | 1 |
| 4 | 2 | 2 | 2 |
| 5 | 2 | 2 | 1 |
| 9 | 2 | 2 | 2 |
| 8 | 3 | 1 | 1 |
| 4 | 3 | 1 | 2 |
| 5 | 3 | 1 | 1 |
| 9 | 3 | 1 | 2 |
| 6 | 3 | 2 | 1 |
| 7 | 3 | 2 | 2 |
| 8 | 3 | 2 | 1 |
| 9 | 3 | 2 | 2 |
| 7 | 1 | 1 | 1 |
| 6 | 1 | 1 | 1 |
| 9 | 3 | 2 | 1 |
| 8 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 |

Using LIMO,

directory = pwd;

Y=NaN(1,1,29);

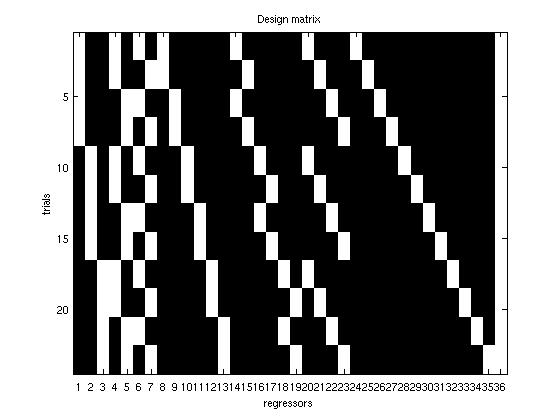
Y(1,1,:) = [5,6,8,7,9,3,2,1,5,6,4,5,3,4,5,9,8,4,5,9,6,7,8,9,7,6,9,8,2];

Cat = [1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 1 1 3 2 2; 1 1 1 1 2 2 2 2 1 1 1 1 2 2 2 2 1 1 1 1 2 2 2 2 1 1 2 2 2; 1 1 2 2 1 1 2 2 1 1 2 2 1 1 2 2 1 1 2 2 1 1 2 2 1 1 1 2 2];

Cont = 0;

[X,nb\_conditions,nb\_interactions,nb\_continuous] =limo\_design\_matrix(Y,Cat,Cont,directory,1,1,1);

we obtain the following design matrix



the 1st three column are for the factor 1, then factor 2,

then factor 3, followed by interactions 1\*2, 1\*3, 2\*3

and the 3 way interaction 1\*2\*3 – resampling operates

on this higher interaction to have at least 2 trials per 'unit'

load Yr % here because of the factor structure Y is reorganized

model = limo\_glm1(squeeze(Yr),X,nb\_conditions,nb\_interactions,nb\_continuous,'OLS')

model.conditions

F = [2.16 ; 0.69 ; 0.17]

p = [0.15 ' 0.42 ; 0.68]

df = [2 12; 1 12; 1 12]

model.interactions

F = [2.68 1.70 0 3.35]

p = [0.1 ; 0.22 ; 1 ; 0.06]

df = [2 12; 2 12; 1 12; 2 12]

If one runs a model, Statistica® results are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **SS** | **Df** | **MS** | **F** | **p** |
| **V1** | 16.58 | 2 | 8.29 | 2.16 | 0.157 |
| **V2** | 2.66 | 1 | 2.66 | 0.69 | 0.420 |
| **V3** | 0.66 | 1 | 0.66 | 0.17 | 0.684 |
| **V1\*V2** | 20.58 | 2 | 10.29 | 2.68 | 0.108 |
| **V1\*V3** | 13.08 | 2 | 6.54 | 1.70 | 0.222 |
| **V2\*V3** | 0 | 1 | 0 | 0 | 1 |
| **V1\*V2\*V3** | 25.75 | 2 | 12.87 | 3.35 | 0.069 |
| **Error** | 46 | 12 | 3.88 |  |  |

Of course those values might be different if you try since data have been resampled