General Linear Model for fMRI: bases of statistical analyses

Cyril Pernet, PhD
University of Edinburgh

Objectives

- Intuitive understanding of the GLM
- Get an idea how t-tests, ANOVA, regressions, etc .. are instantiation of the GLM
- Learn key concepts: linearity, model, design matrix, contrast, colinearity, orthogonalization

Overview

What is linearity?

Why do we speak of models?

A simple fMRI model

Contrasts

Issues with regressors

What is linearity?

Linearity

- Means created by lines
- In maths it refers to equations or functions that satisfy 2 properties: additivity (also called superposition) and homogeneity of degree 1 (also called scaling)
- Additivity → y = x1 + x2 (output is sum of inputs)
- Scaling → y = βx1 (output is proportional to input)

http://en.wikipedia.org/wiki/Linear

Examples of linearity – non linearity

- X = randn(10,1)
- Linear correlation
- Y1 = 3x + 2
- Pearson r = 1
- Non linear correlation
- Y2 = abs(2x)
- Pearson r = 0.38

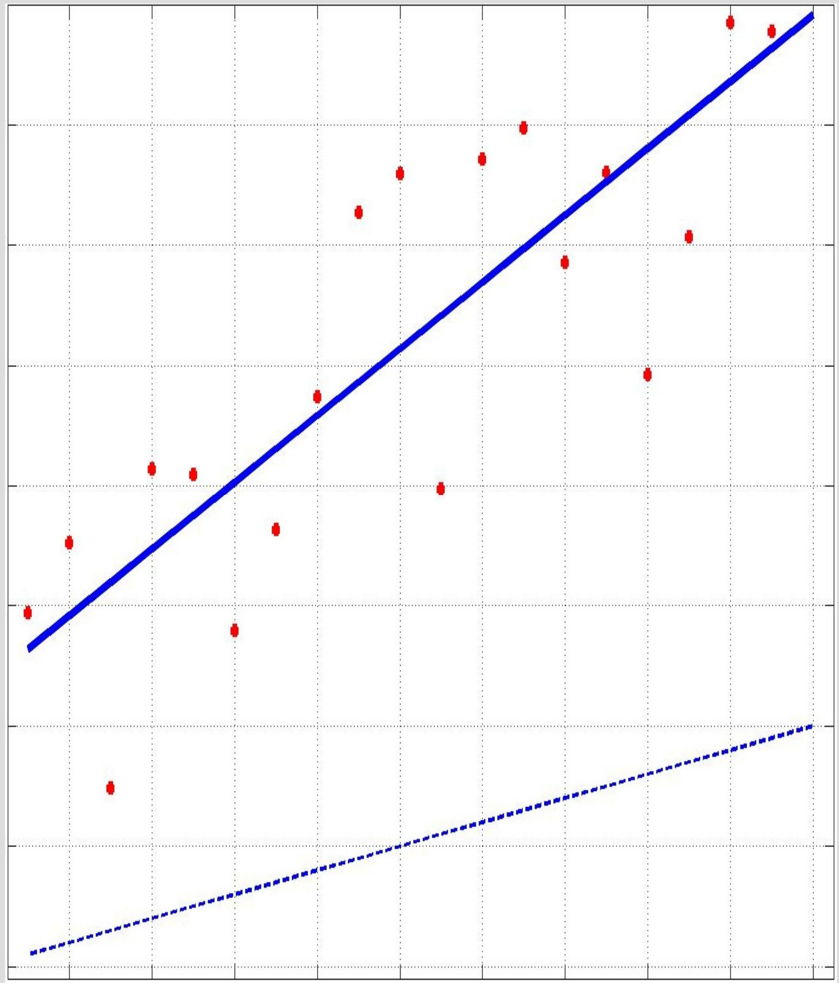
What is a linear model?

What is a linear model?

- An equation or a set of equations that models data and which corresponds geometrically to straight lines, plans, hyperplans and satisfy the properties of additivity and scaling.
- Simple regression: $y = \beta 1x1 + \beta 2 + \varepsilon$
- Multiple regression: $y = \beta 1x1 + \beta 2x2 + \beta 3 + \varepsilon$
- One way ANOVA: $y = \mu + \alpha i + \varepsilon$
- Repeated measure ANOVA: $y = u + Si + \alpha i + \varepsilon$
- …

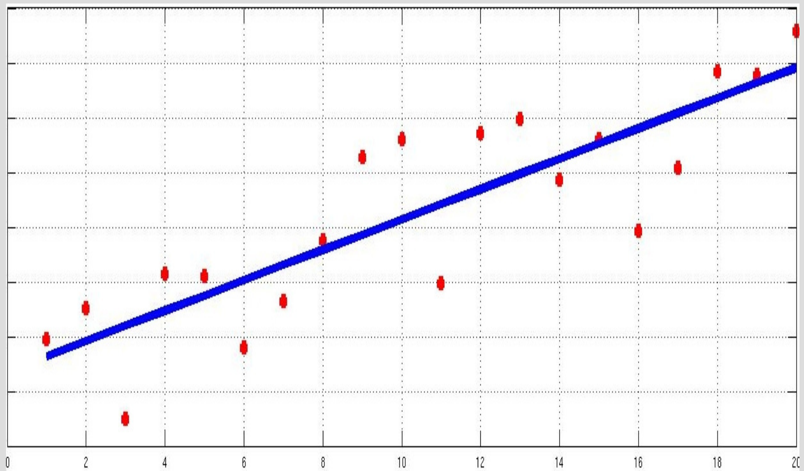
A regression is a linear model

- We have an experimental measure x (e.g. stimulus intensity from 0 to 20)
- We then do the expe and collect data y (e.g. RTs)
- Model: $y = \beta 1 x + \beta 2$
- Do some maths / run a software to find $\beta 1$ and $\beta 2$
  $\hat{y} = 2.7x + 23.6$

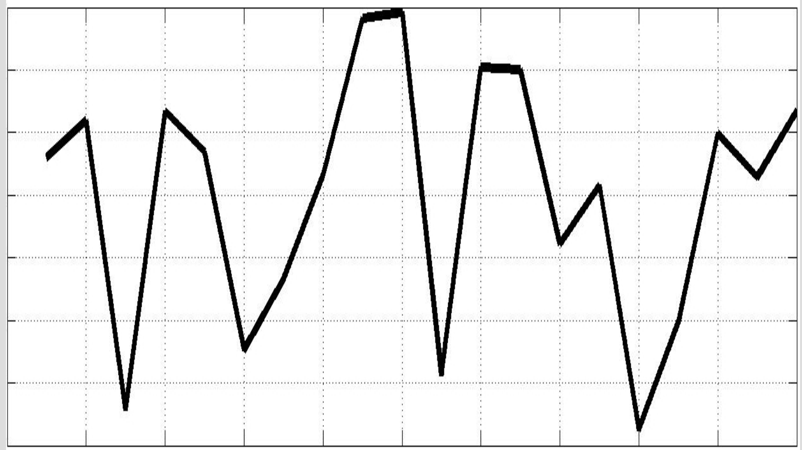


A regression is a linear model

- The error is the distance between the data and the model
- F = (SSeffect / df) / (SSerror / df_error)
- <u>SSeffect</u> = norm(model - mean(model)).^2;
- <u>SSerror</u> = norm(residuals).^2;



Error = Y – XB



Summary

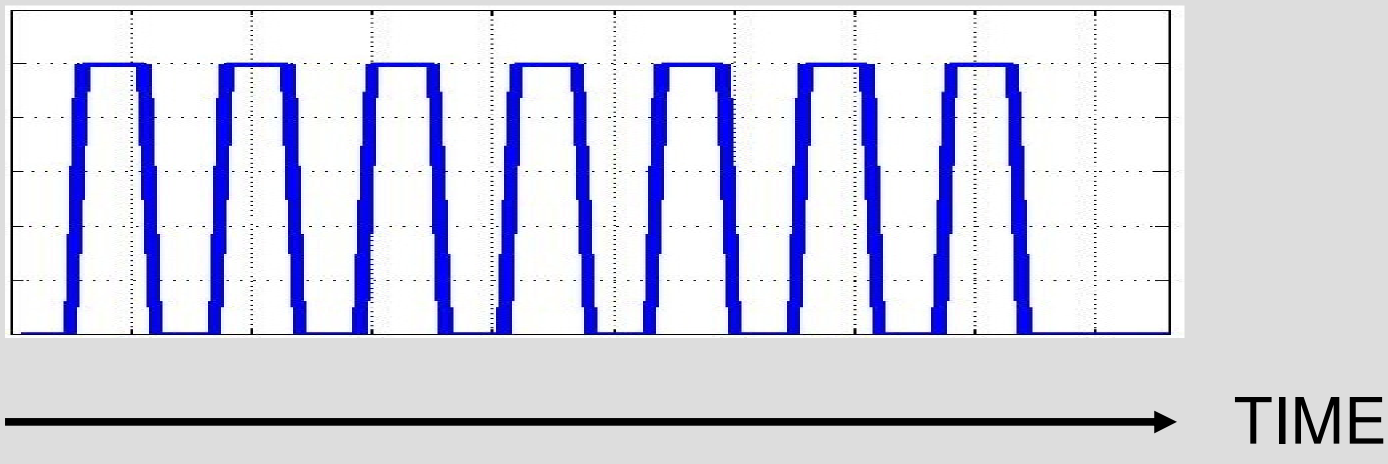
- Linear model: $y = \beta 1 x1 + \beta 2 x2$ (output = additivity and scaling of input)

A simple fMRI model

http://www.fil.ion.ucl.ac.uk/spm/data/auditory/

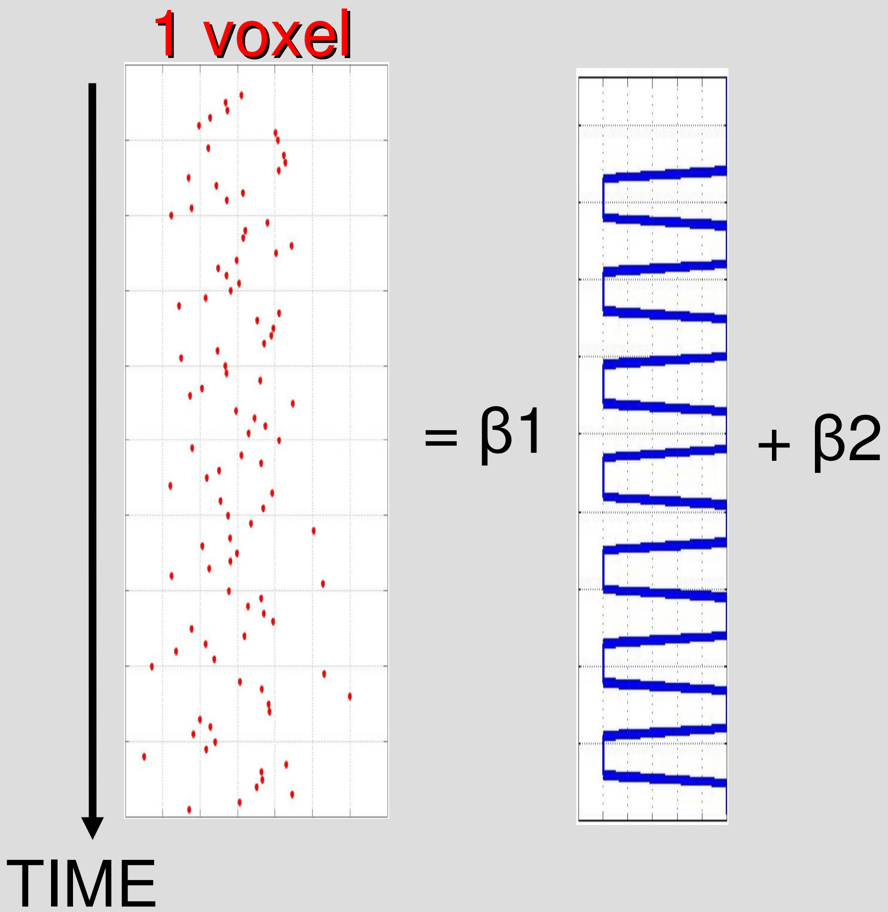
FMRI experiment

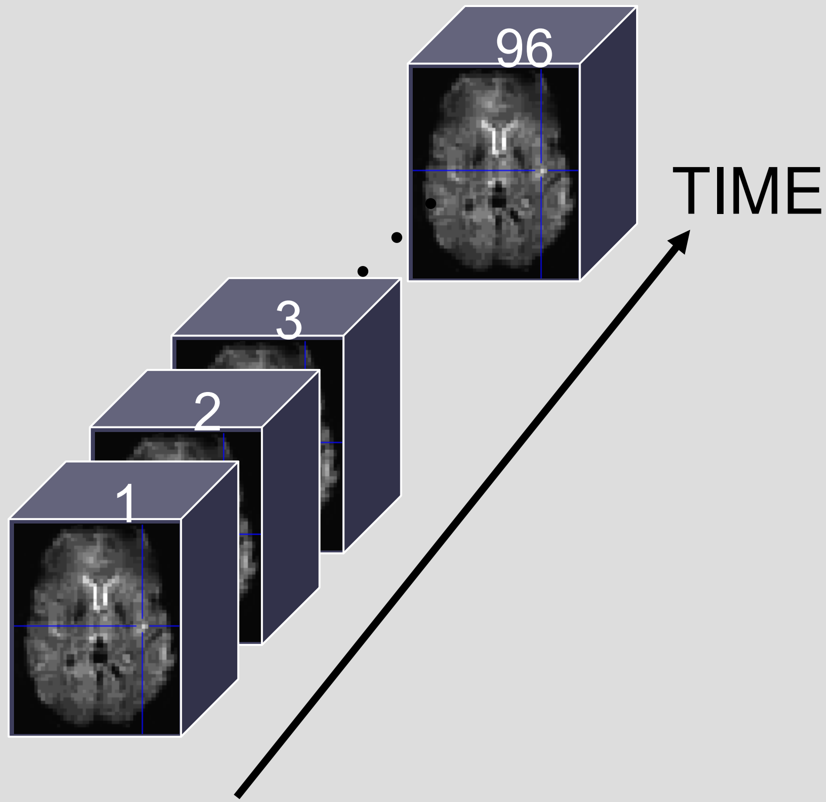
- *SPM data set*: which areas are activated by the presentation of bi-syllabic words presented binaurally (60 per minute)
- Experimental measure x: 7 blocks of 42 sec of stimulation



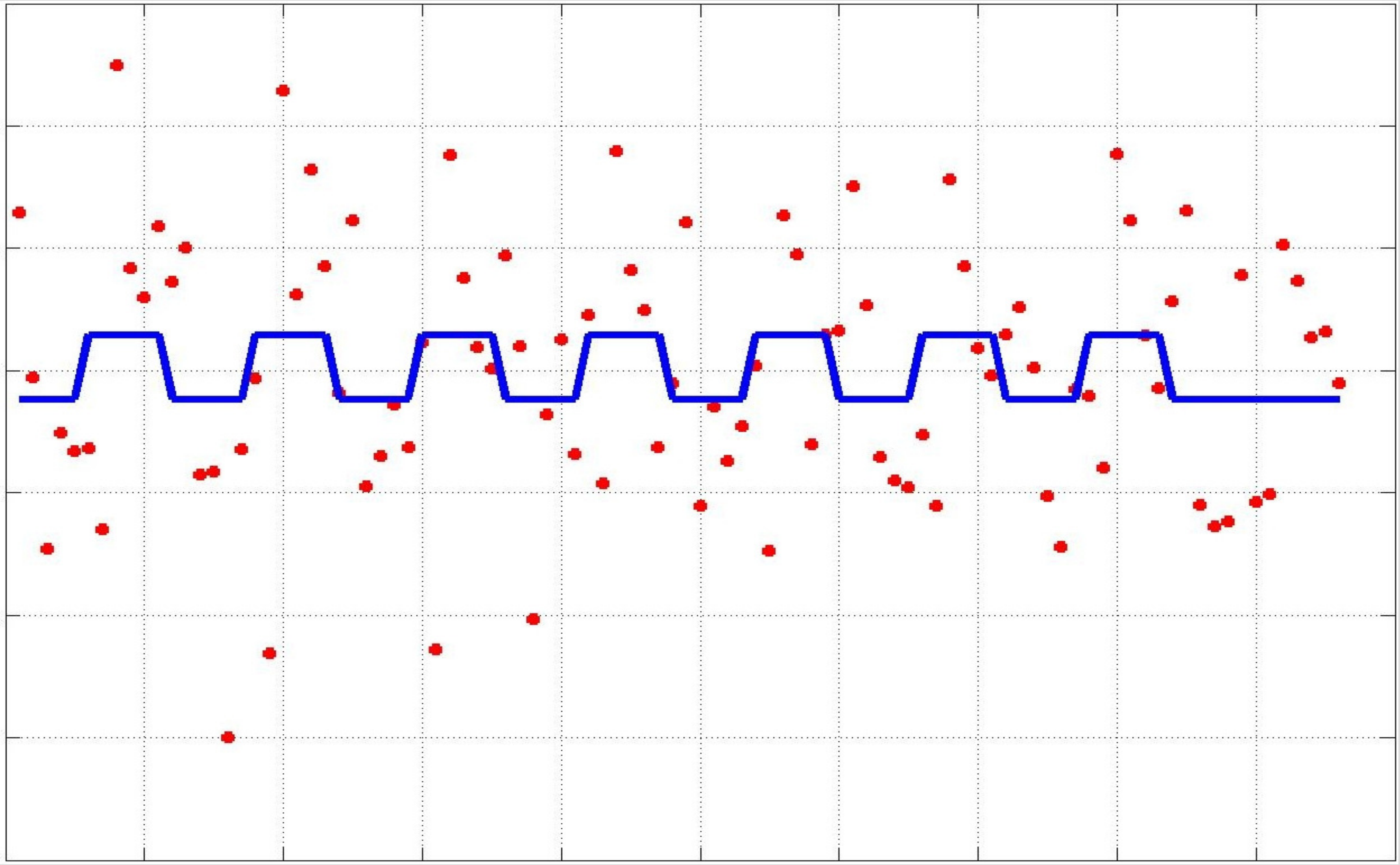
FMRI experiment

- Collect the data : 96 fMRI volumes (RT=7s)
- Model: y = β1x+ β2



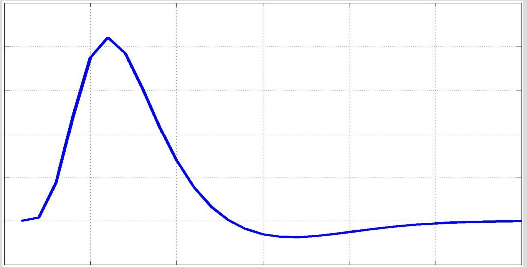


FMRI experiment

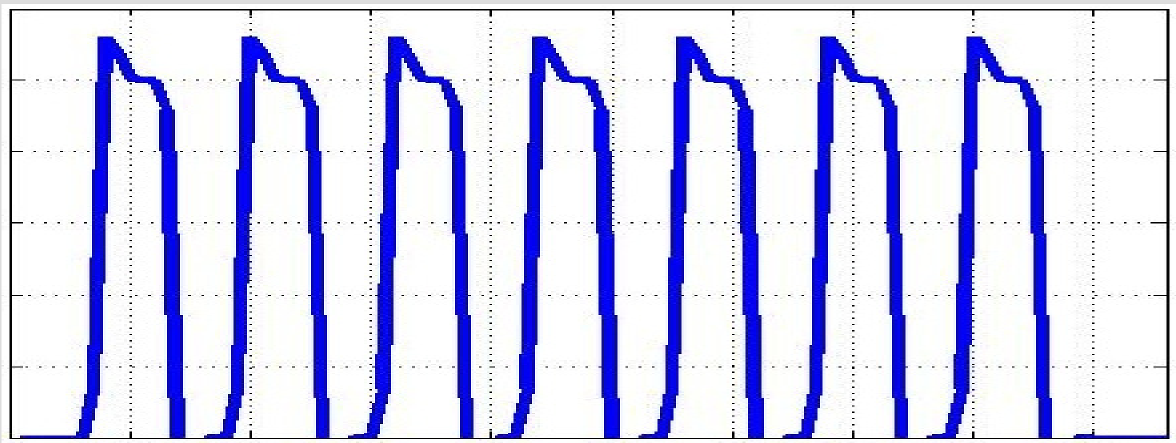


FMRI experiment

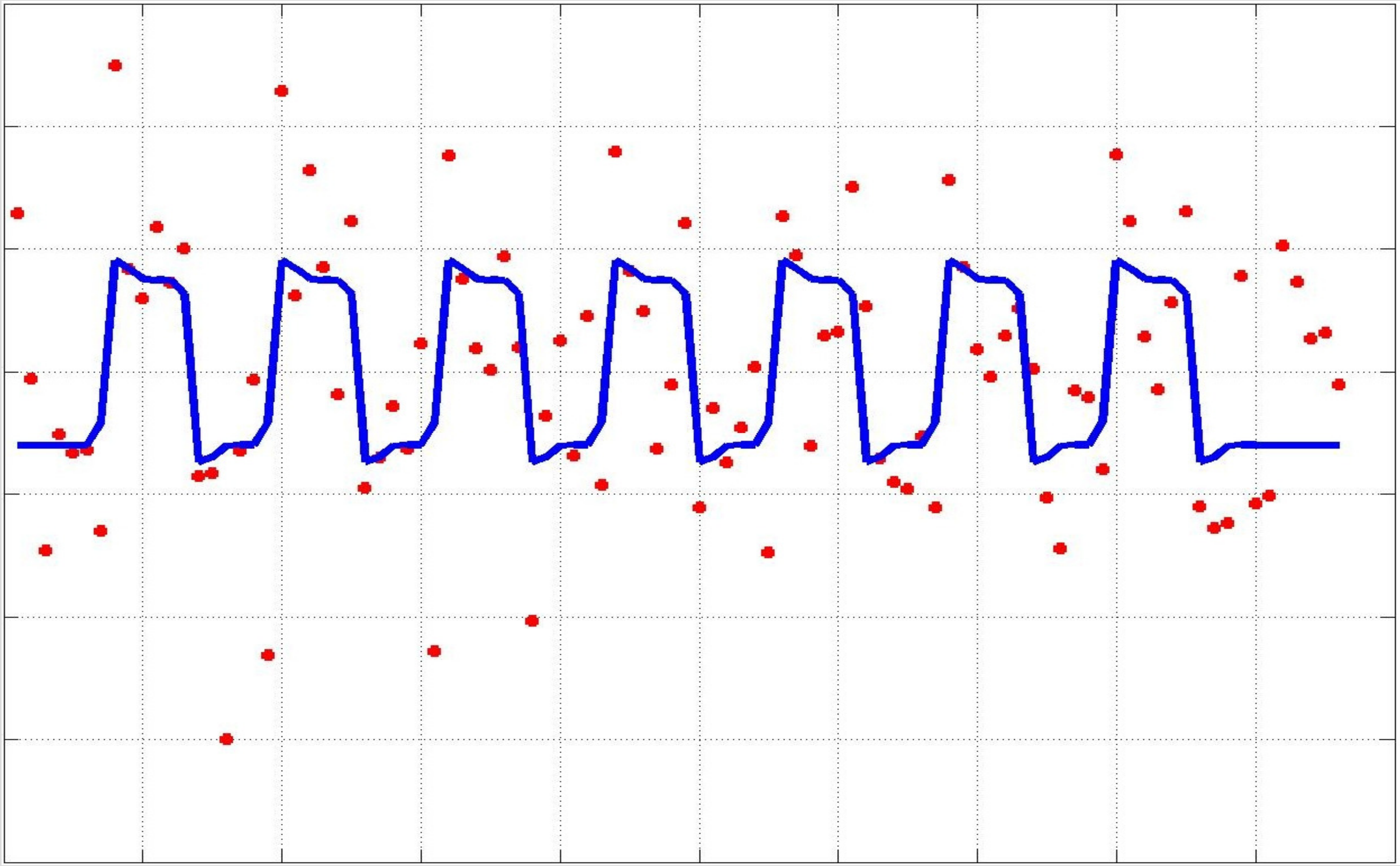
- A better model: we know the shape of the BOLD response



- Convolution by the hrf: x ⊗ hrf



FMRI experiment

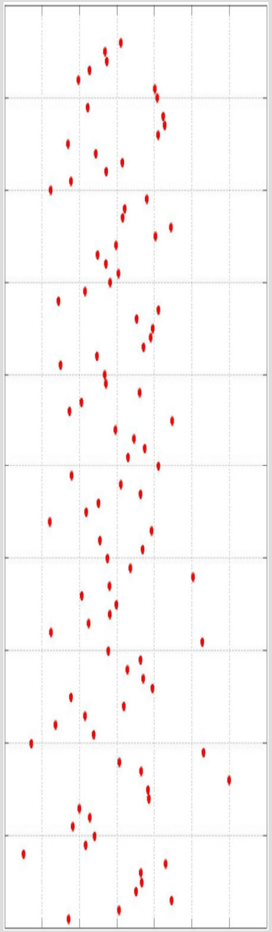


FMRI experiment

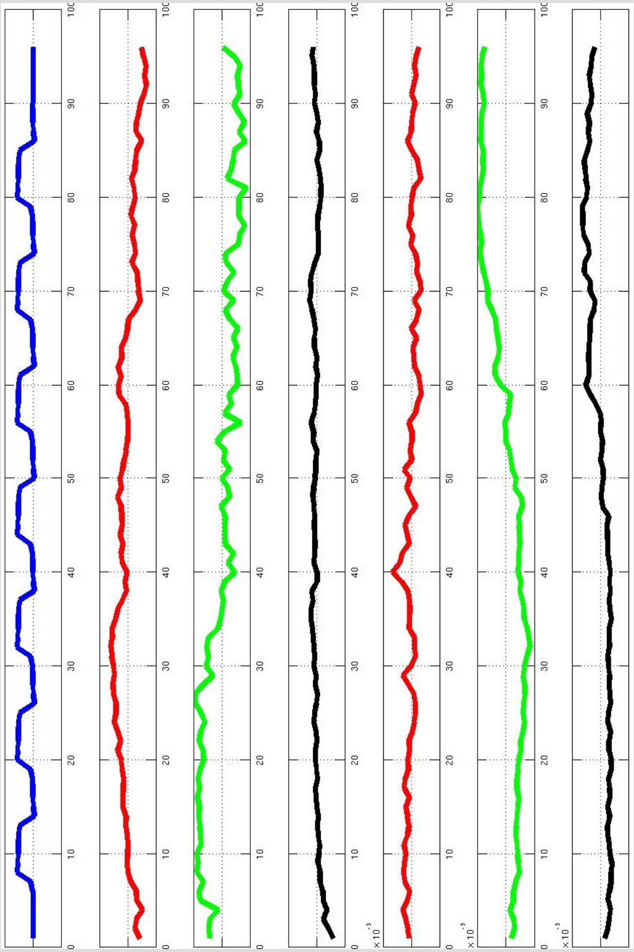
- An even better model: add motion parameters

1 voxel

TIME

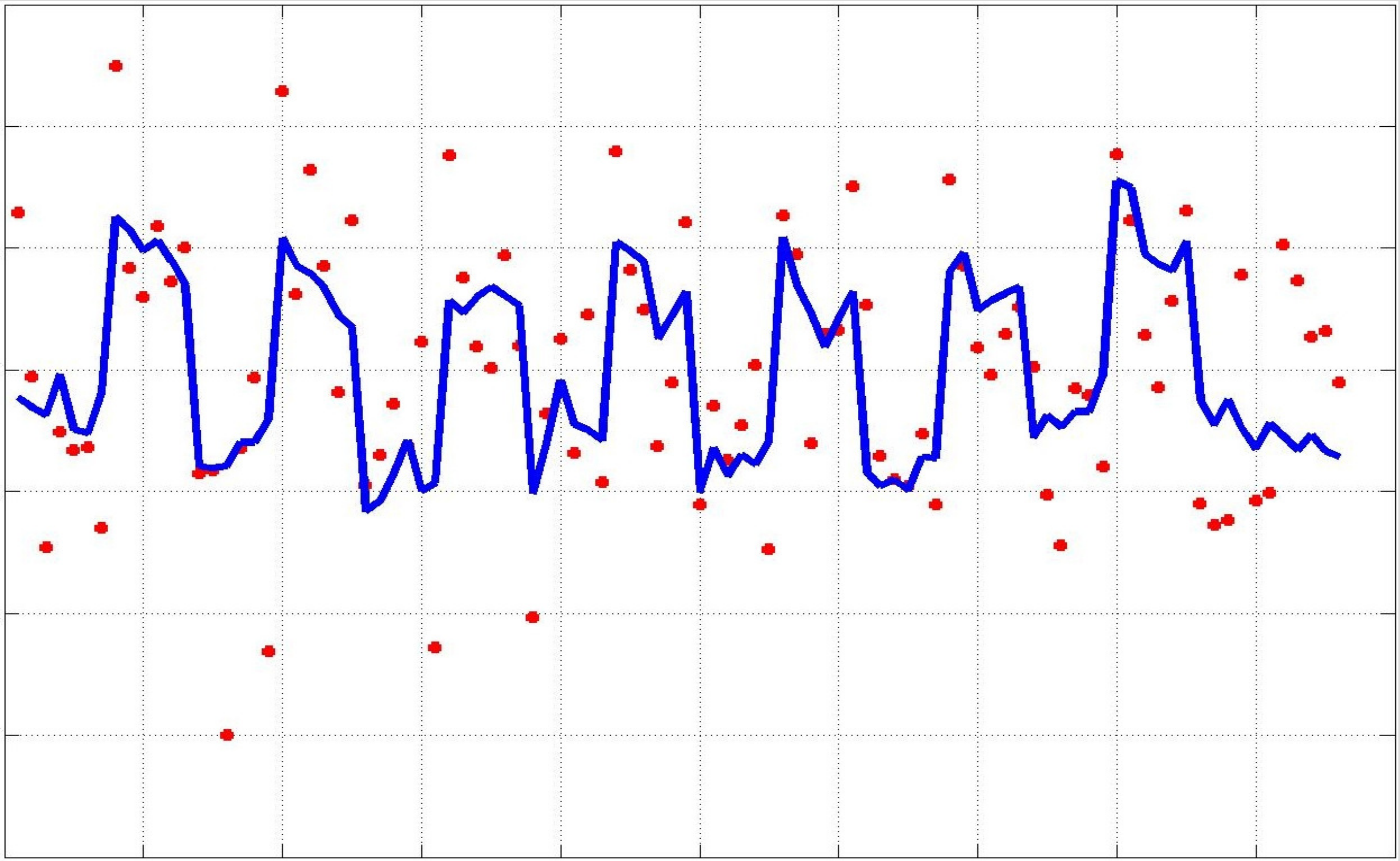


=



$* [\beta 1\ \beta 2\ \beta 3\ \beta 4\ \beta 5\ \beta 6\ \beta 7\ \beta 8] + \beta 8$

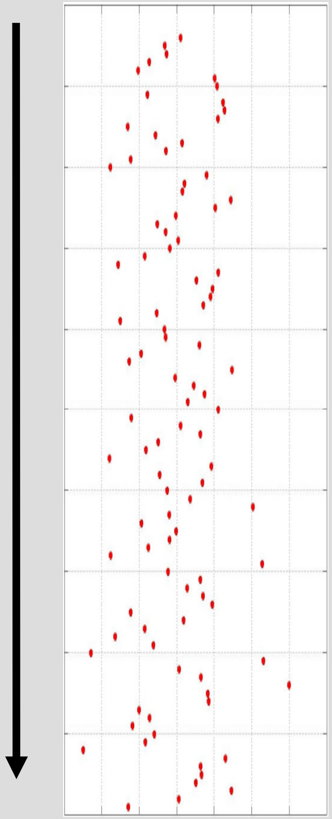
FMRI experiment



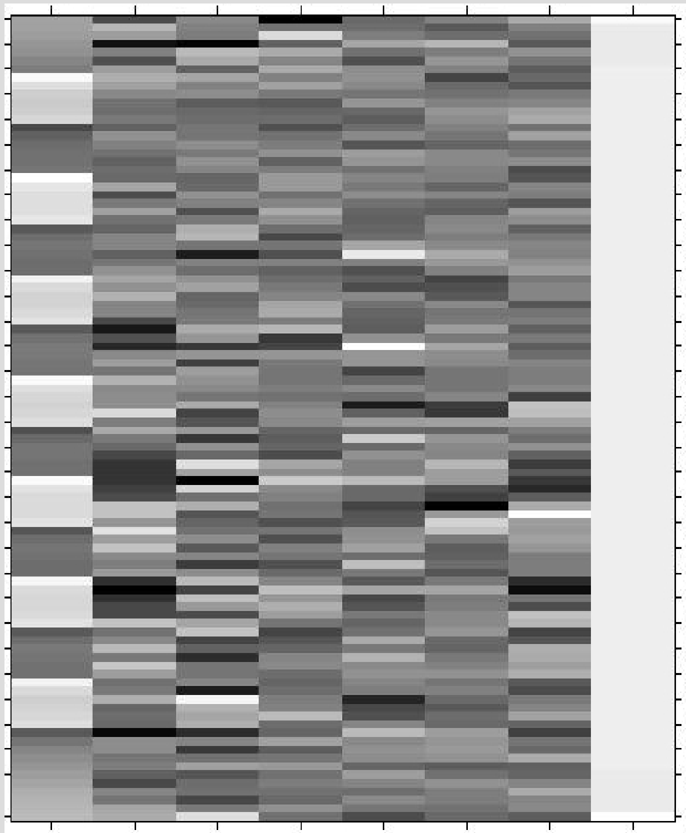
FMRI experiment

- Matrix formulation and SPM colour coding

1 voxel



=

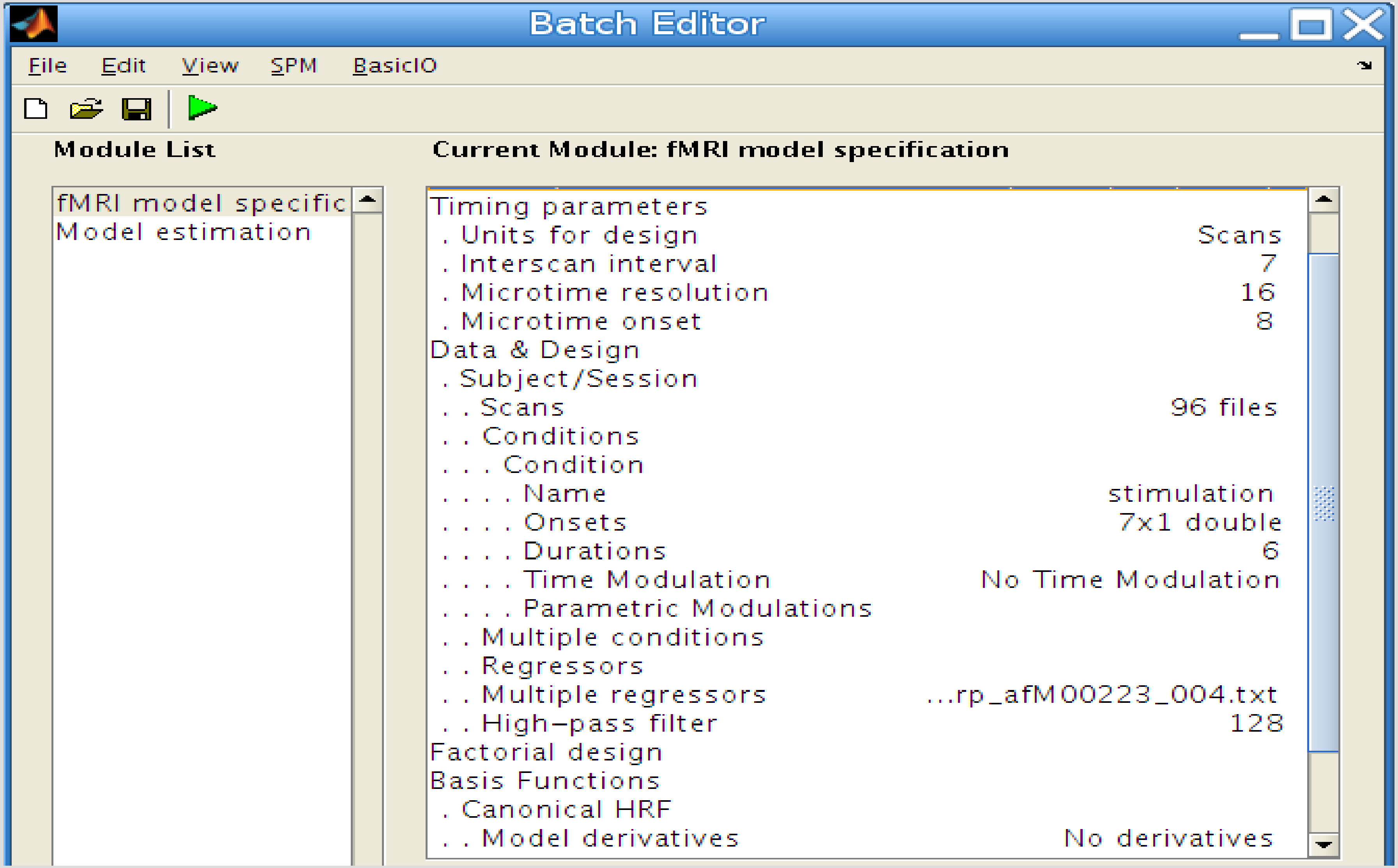


* [β1 β2 β3 β4 β5 β6 β7 β8]

BetaXXXX.hdr
BetaXXXX.img

FMRI data (Y) = Design matrix (X = SPM.mat) * B + E (ResMS)

FMRI experiment



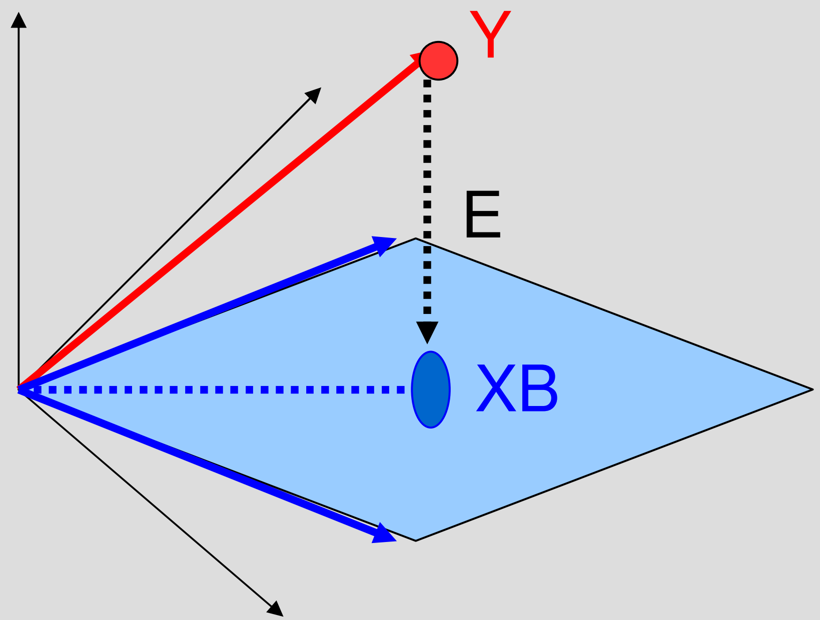
Summary

- Linear model: $y = \beta 1 x1 + \beta 2 x2$ (output = additivity and scaling of input)
- GLM: $Y = XB + E$ (matrix formulation, works for any statistics, express the data Y as a function of the design matrix X)

Contrasts

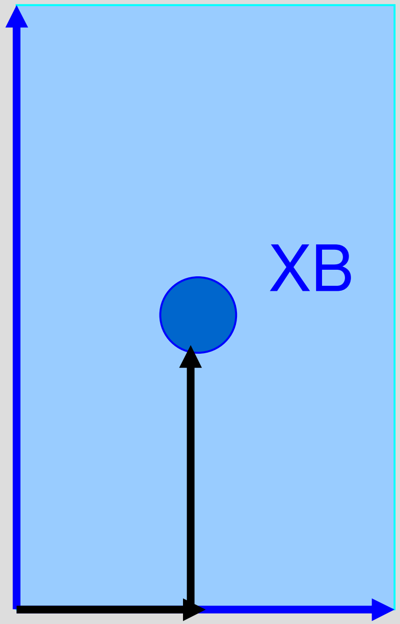
Model = R^2

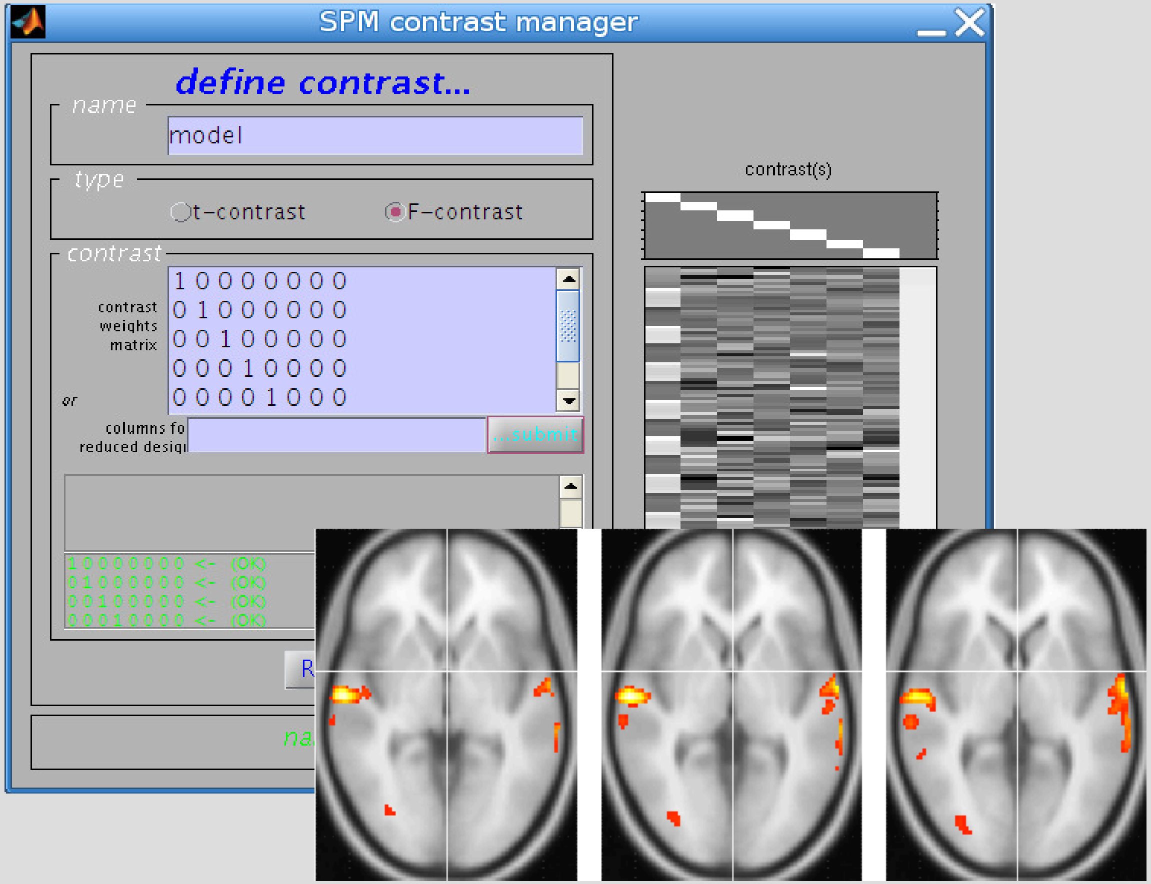
- Geometrical perspective
- Y = 3 observations X = 2 regressors
- Y = XB+E



Model = R^2

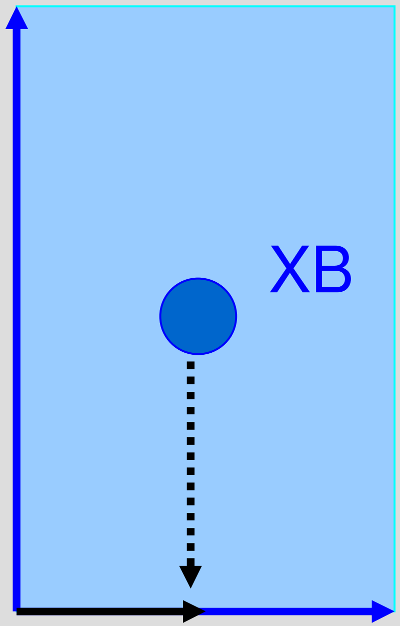
- Where does the model fit the data?
- F test for all regressors: $y = 1/2x1 + 1/2x2 + \varepsilon$

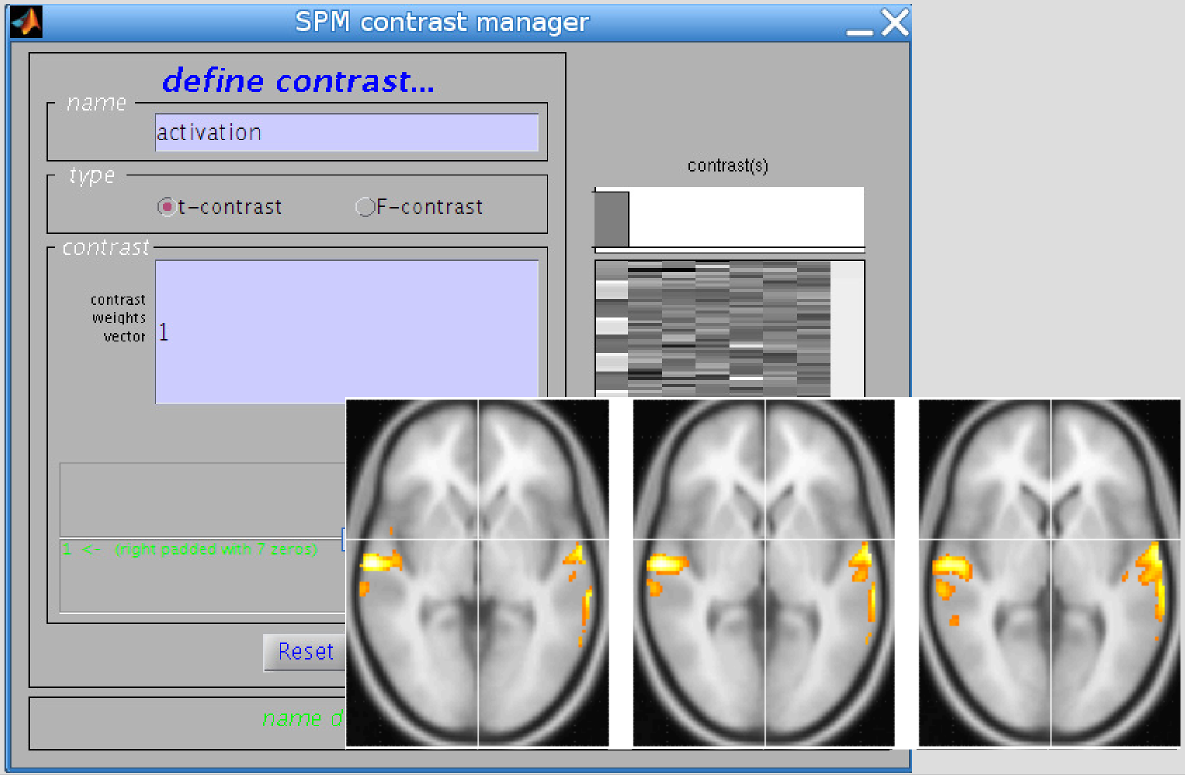




Contrast = effect to test

- Where does the regressor for activation only explain the data (given the model)
- y = 1/2x1+ε   (the orientation of x1 and value of β1 are fixed by the model)





Summary

- Linear model: $y = \beta 1 x1 + \beta 2 x2$ (output = additivity and scaling of input)
- GLM: $Y = XB+E$ (matrix formulation, works for any statistics, express the data Y as a function of the design matrix X)
- Contrasts: F or t test for the effect of 1 or several regressors given the design matrix

Issues with regressors

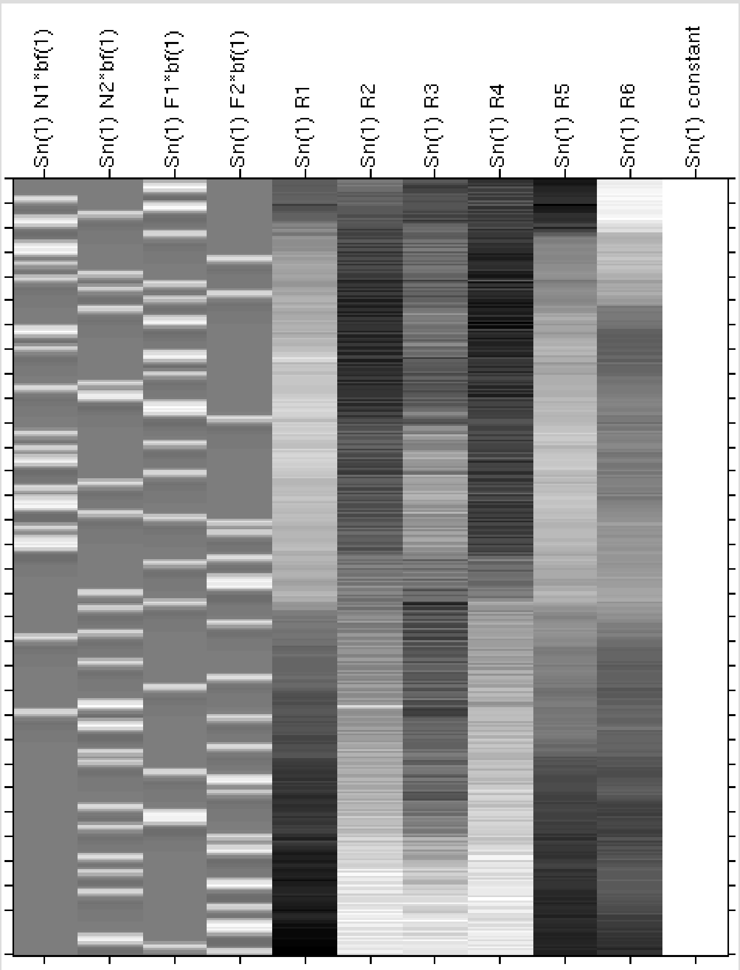
More contrasts

- New experiment: (Famous vs. Nonfamous) x (1st vs 2nd presentation) of faces against baseline of chequerboard
- 2 presentations of 26 Famous and 26 Nonfamous Greyscale photographs, for 0.5s, randomly intermixed, for fame judgment task (one of two right finger key presses).

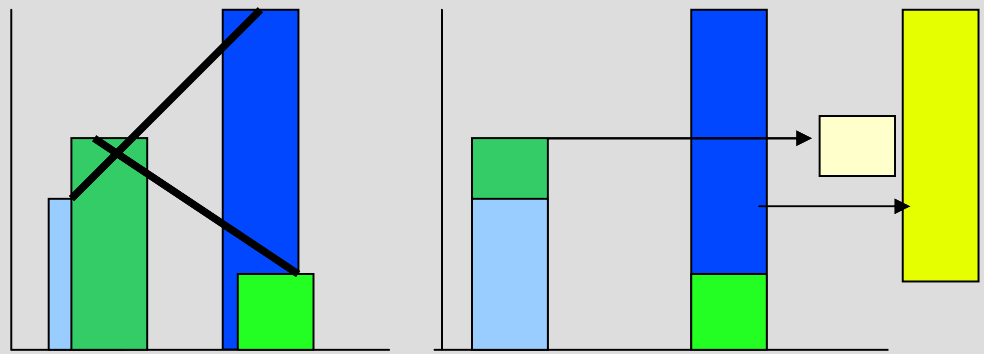
http://www.fil.ion.ucl.ac.uk/spm/data/face_rep/face_rep_SPM5.htm

More contrasts

- SPM design matrix

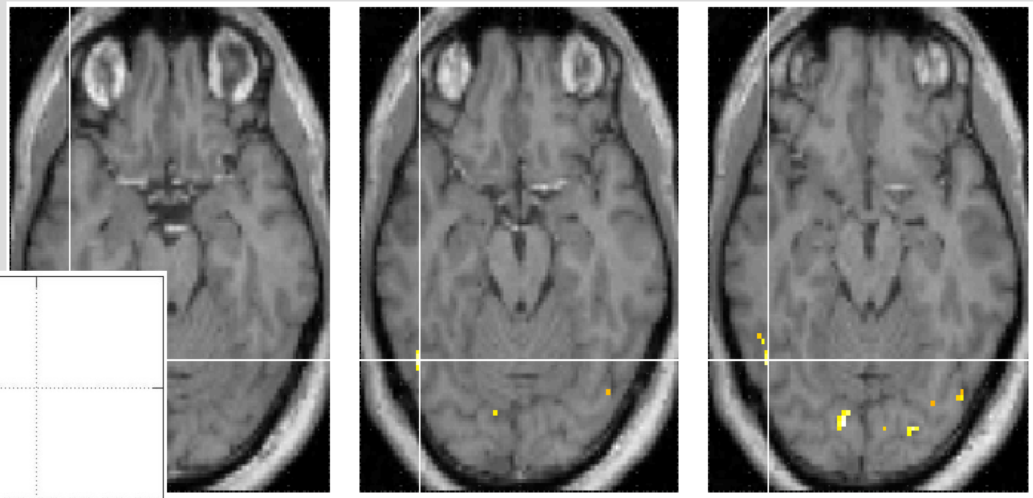


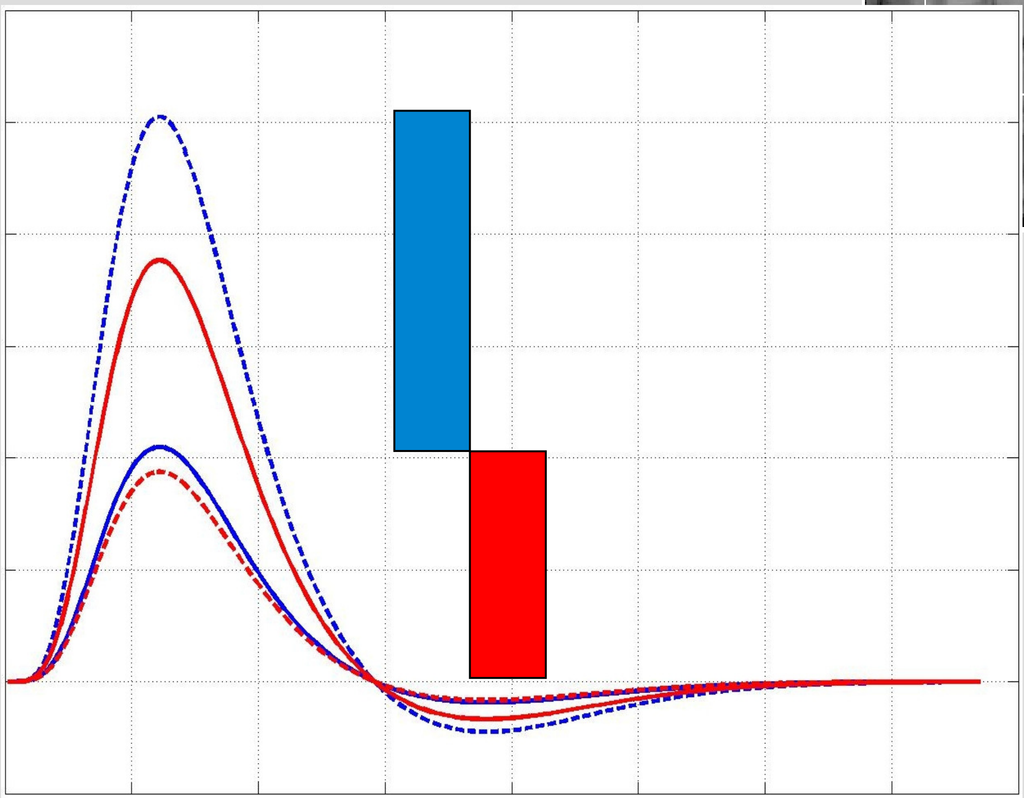
- <u>Questions:</u>
- Main effects
- Fame: [1 1 -1 -1 0 0 ...]
- Rep: [1 -1 1 -1 0 0 ...]
- Interaction [1 -1 -1 1 0 0 ..]



(1 -1) – (1 -1)

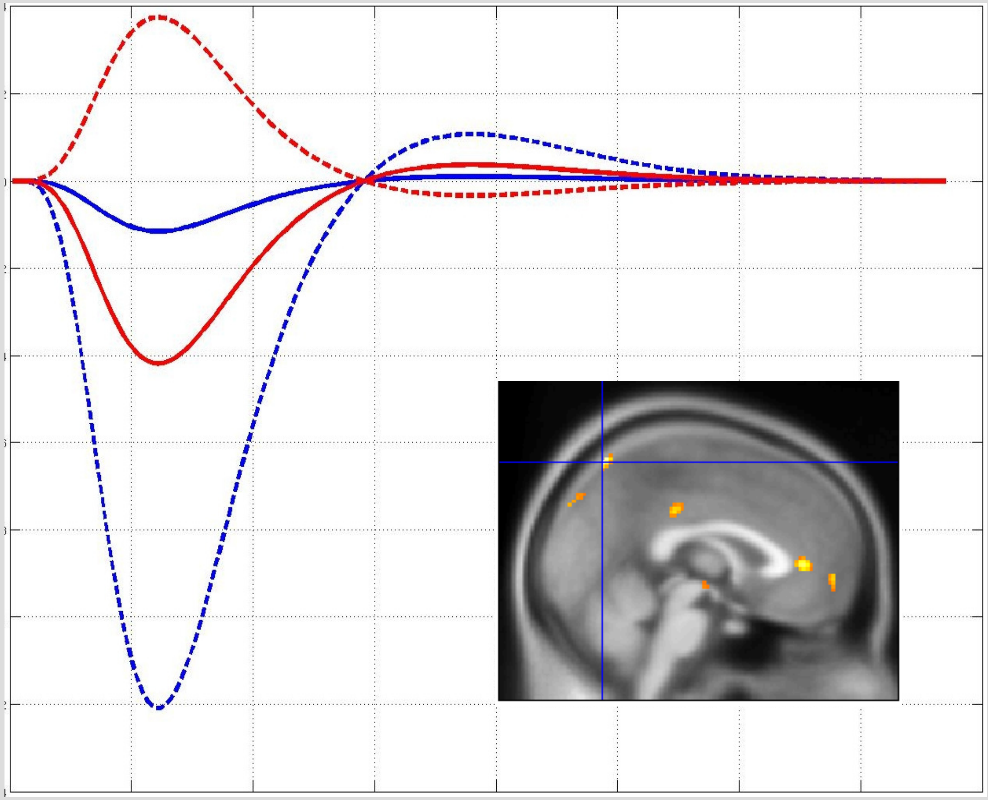
More contrasts

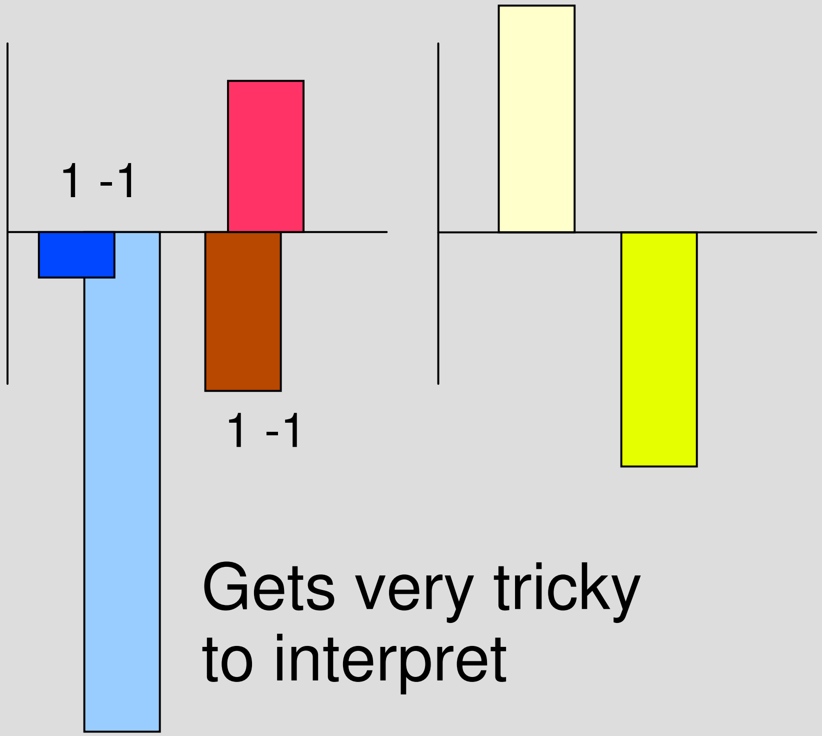




More contrasts

- Search in areas where all regressors are positive or all negative otherwise ...



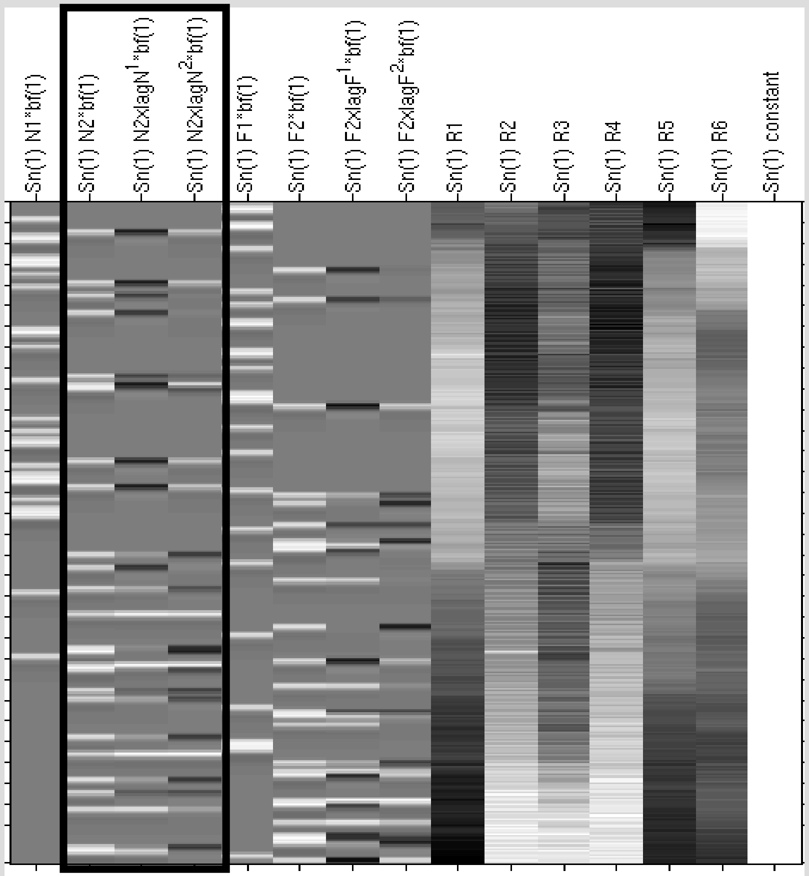


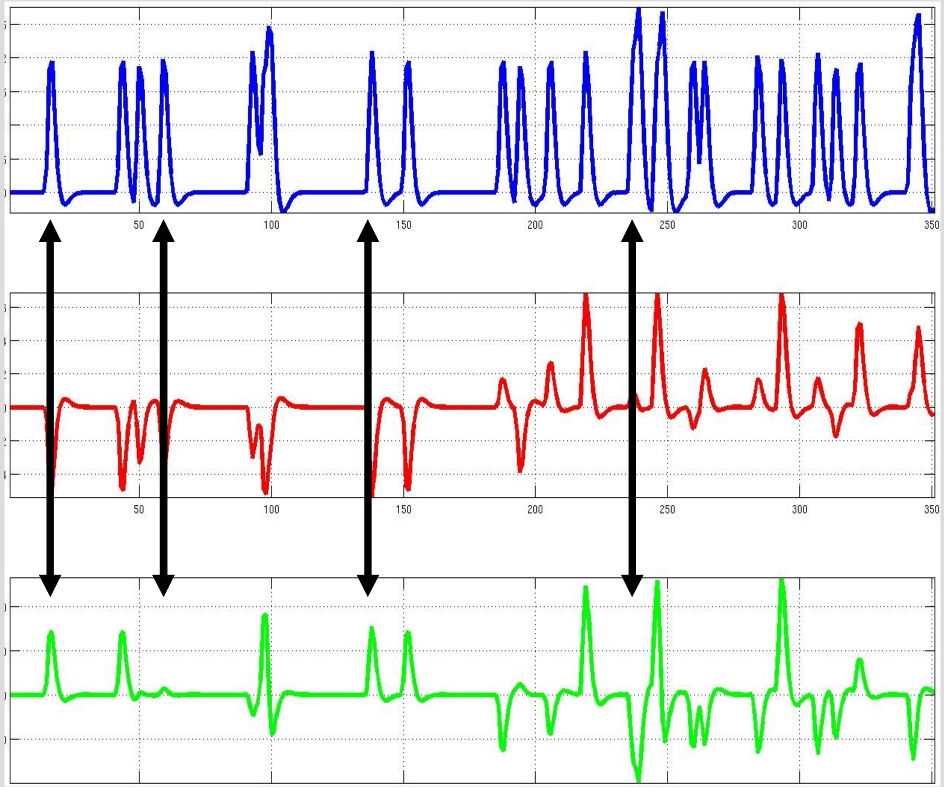
Gets very tricky to interpret

Always plot the results

More Regressors

- Same design as before but added a 'parametric' regressor – here the lag between presentations

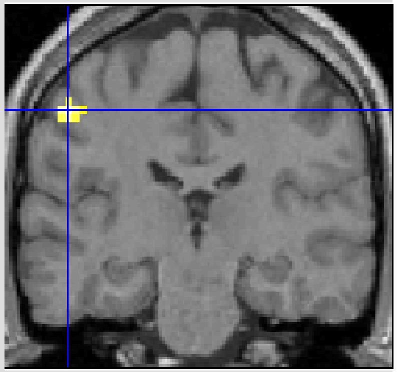


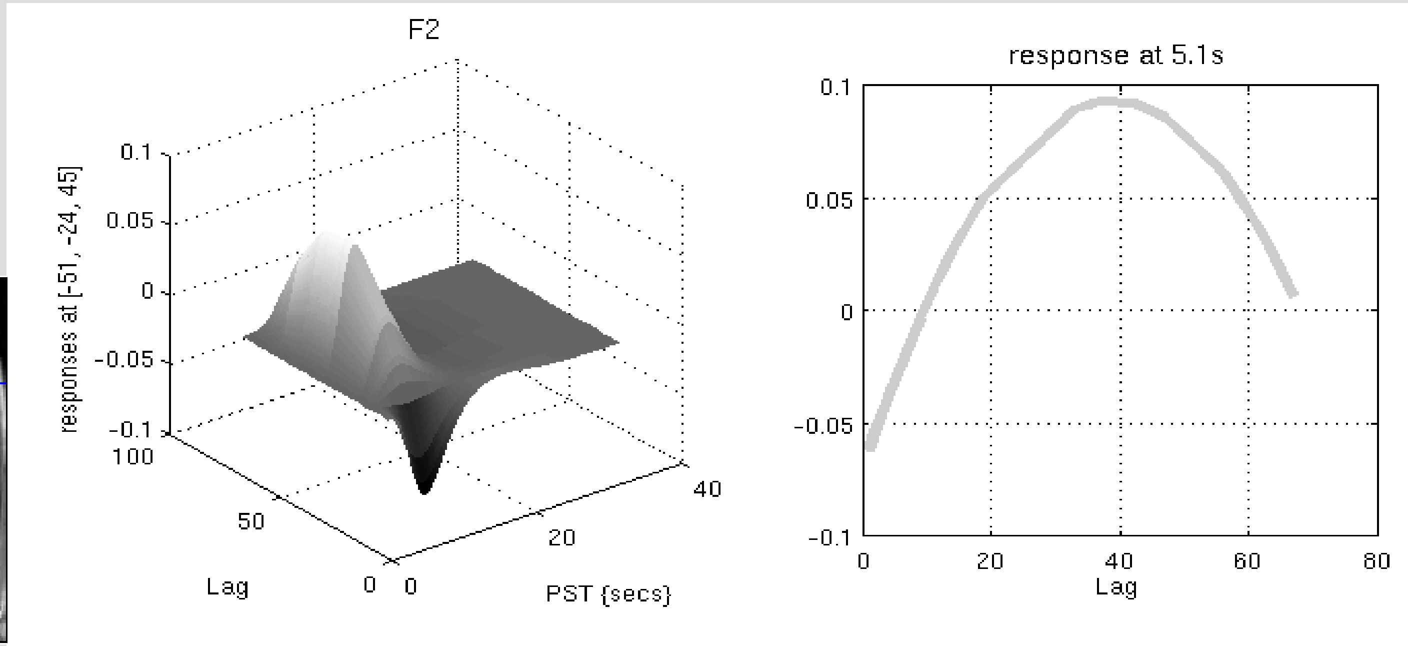


3 regressors all at the same time (diff. values)

More Regressors

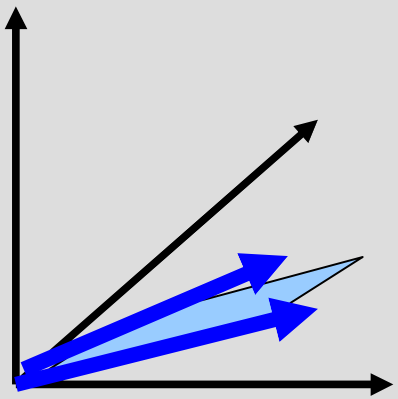
- The parametric regressors express the amplitude of signal as a function of the lag, i.e. the signal amplitude changes from trial to trial





More Regressors: collinearity

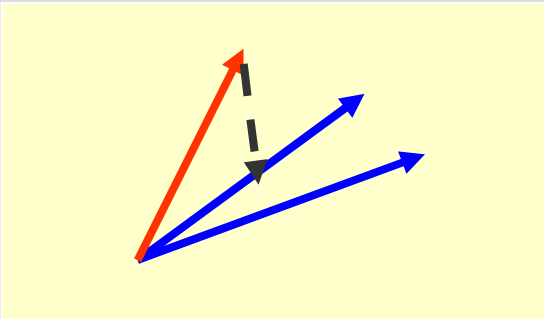
- SPM *orthogonalizes* the parametric regressors making the regressors *non collinear.*
- Three or more points are said to be collinear if they lie on a single straight line.
- Regressors are collinear if they are perfectly correlated (note corr of 2 vectors = cosθ)

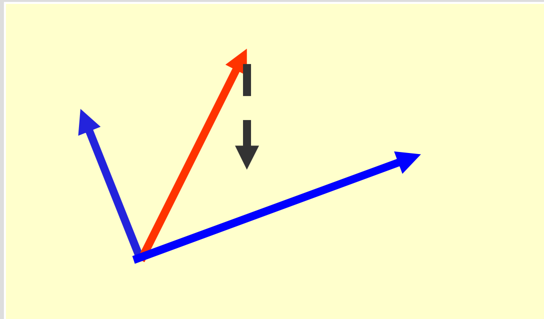


- Can make solution impossible
- Often make the model ok but individual regression values unstable
- Classical height and weight regression pblm

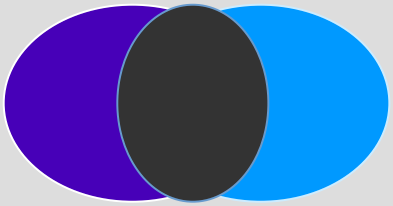
http://en.wikipedia.org/wiki/Multicollinearity
http://mathworld.wolfram.com/Collinear.html

More Regressors: orthogonalization

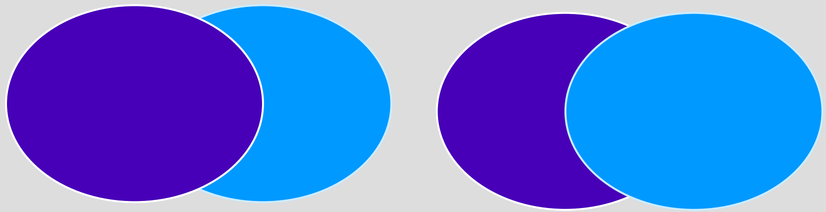




Lot of variance shared – because we look for the unique part of variance, the shared part goes into the error



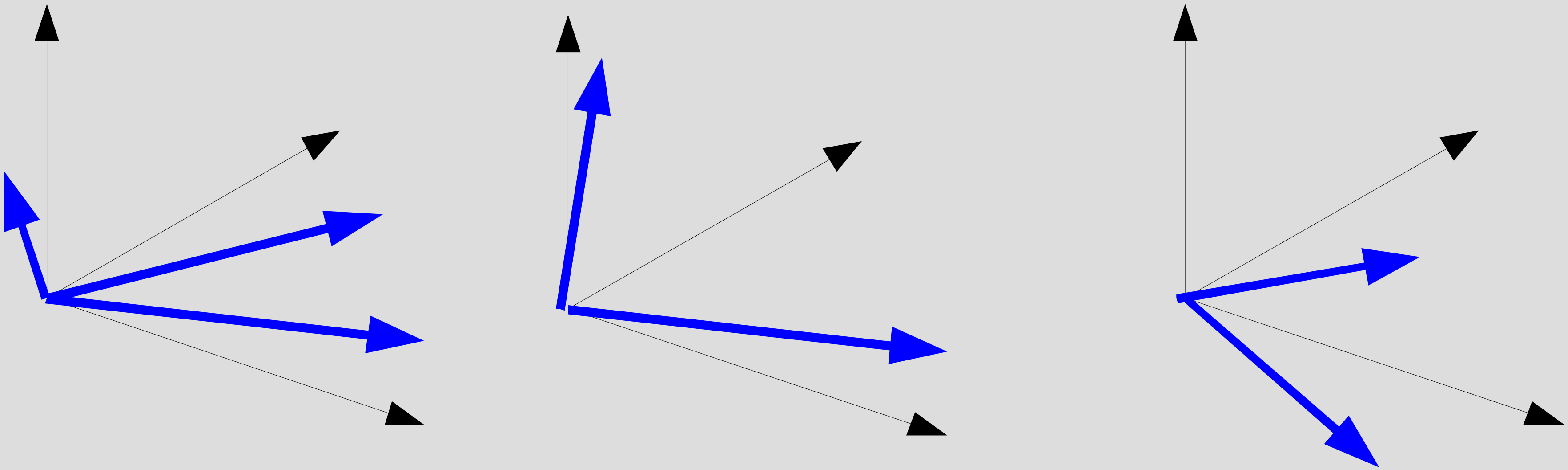
Orthogonalization (θ = 90°) removes shared variance BUT order matters !



http://en.wikipedia.org/wiki/Gram-Schmidt_process

More regressors

- Linearly independent (X2 ≠ aX1), orthogonal (X1'Y2 = 0) and uncorrelated (X1-mean(X1))'(X2-mean(X2))=0) variables



More regressors

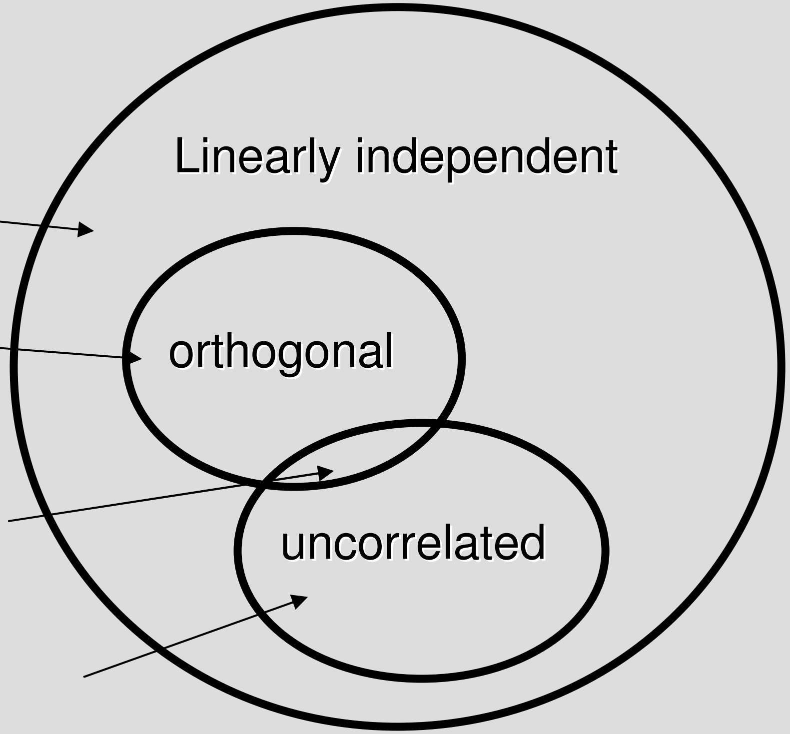
- Linearly independent (X2 ≠ aX1), orthogonal (X1’Y2 = 0) and uncorrelated (X1-mean(X1))’(X2-mean(X2))=0) variables

[1 1 2 3] and [2 3 4 5]
Independent, correlated, not orthogonal

[1 -5 3 -1] and [5 1 1 3]
Independent, correlated and orthogonal

[-1 -1 1 1] and [1 -1 1 -1]
Independent, uncorrelated and orthogonal

[0 0 1 1] and [1 0 1 0]
Independent, uncorrelated, not orthogonal



http://www.jstor.org/pss/2683250

Summary

- Linear model: $y = \beta 1 x1 + \beta 2 x2$ (output = additivity and scaling of input)
- GLM: Y = XB+E (matrix formulation, works for any statistics, express the data Y as a function of the design matrix X)
- Contrasts: F or t test for the effect of 1 or several regressors given the design matrix
- More regressor is better as it captures more of the signal but it may bring instability if regressors are collinear (and cost df) – SPM orthogonalizes parametric regressors