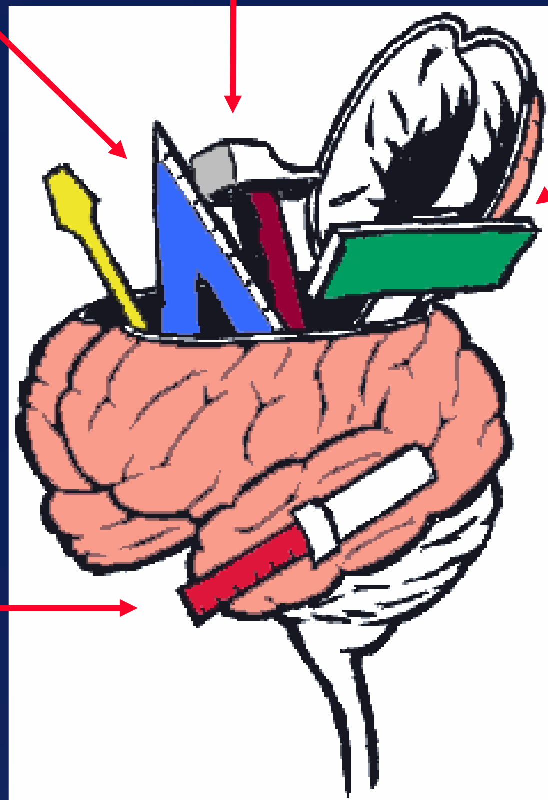
SPM short course at Yale – April 2005
Linear Models and Contrasts

T and F tests :
(orthogonal projections)

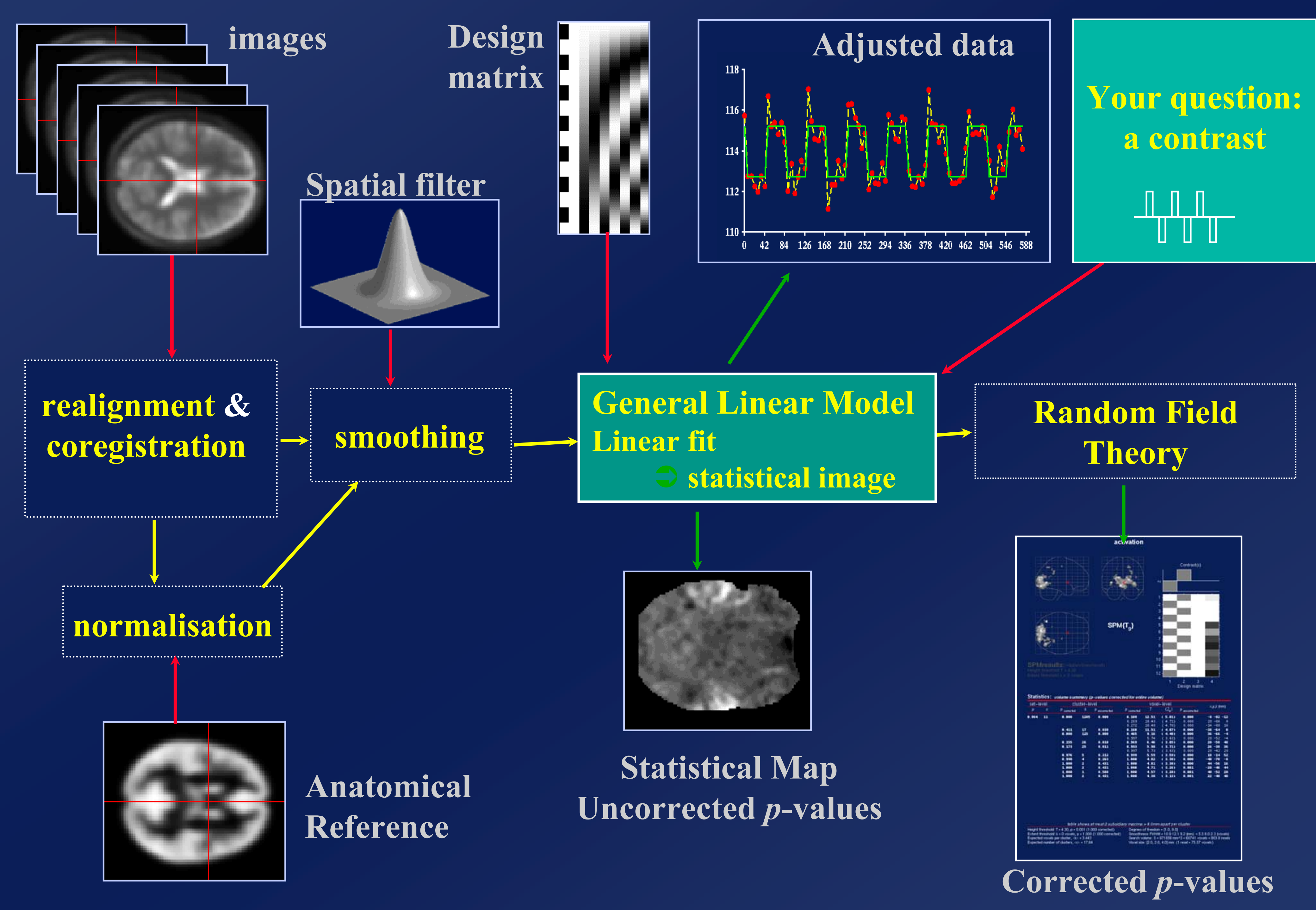
Hammering a Linear Model

The random field theory



Use for Normalisation

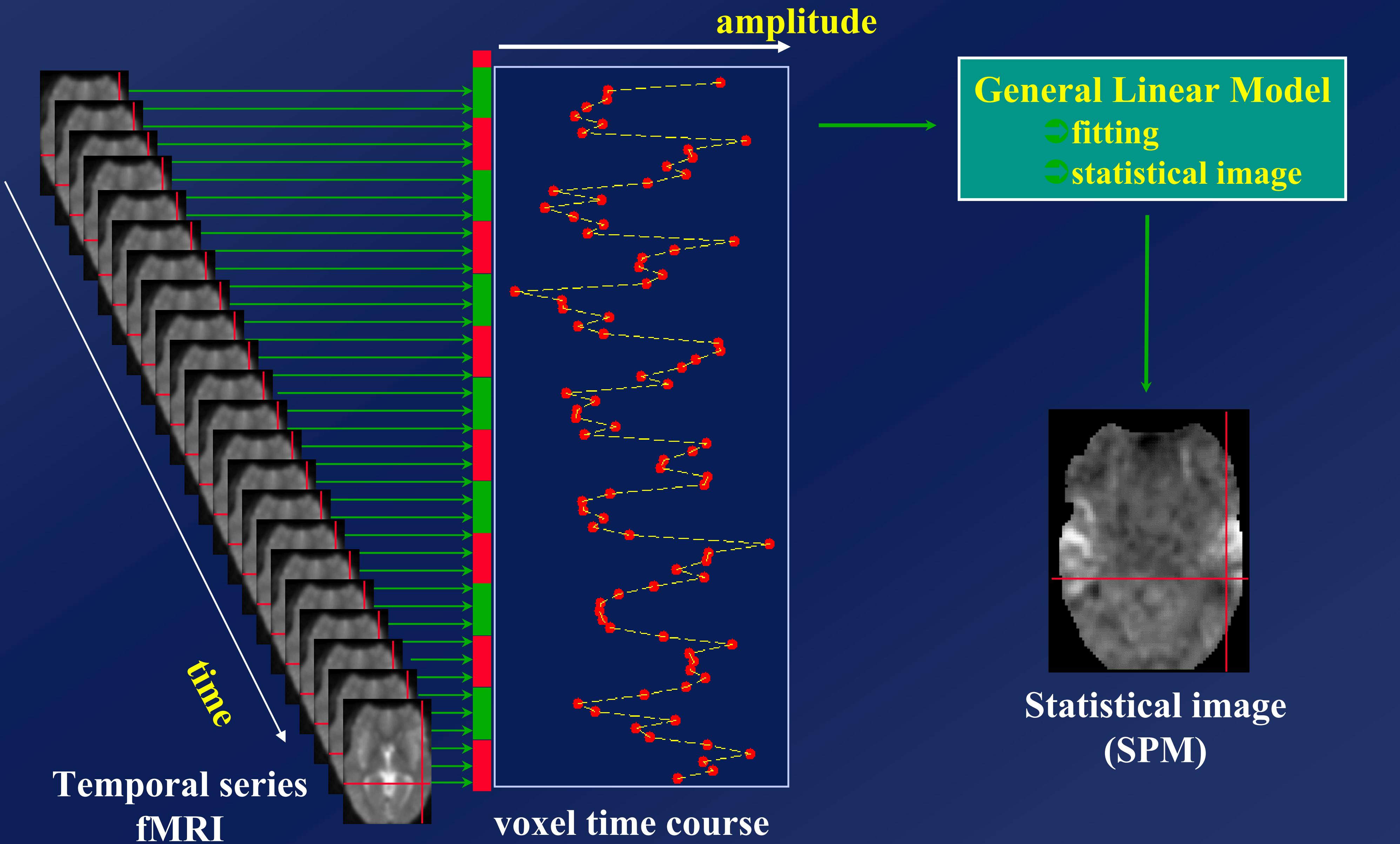
Jean-Baptiste Poline
Orsay SHFJ-CEA
www.madic.org


images
Design matrix
Adjusted data
118
116
114
112
110
0 42 84 126 168 210 252 294 336 378 420 462 504 546 588
Your question: a contrast
Spatial filter
realignment & coregistration
smoothing
General Linear Model
Linear fit
statistical image
Random Field Theory
normalisation
Anatomical Reference
Statistical Map
Uncorrected p-values
Corrected p-values

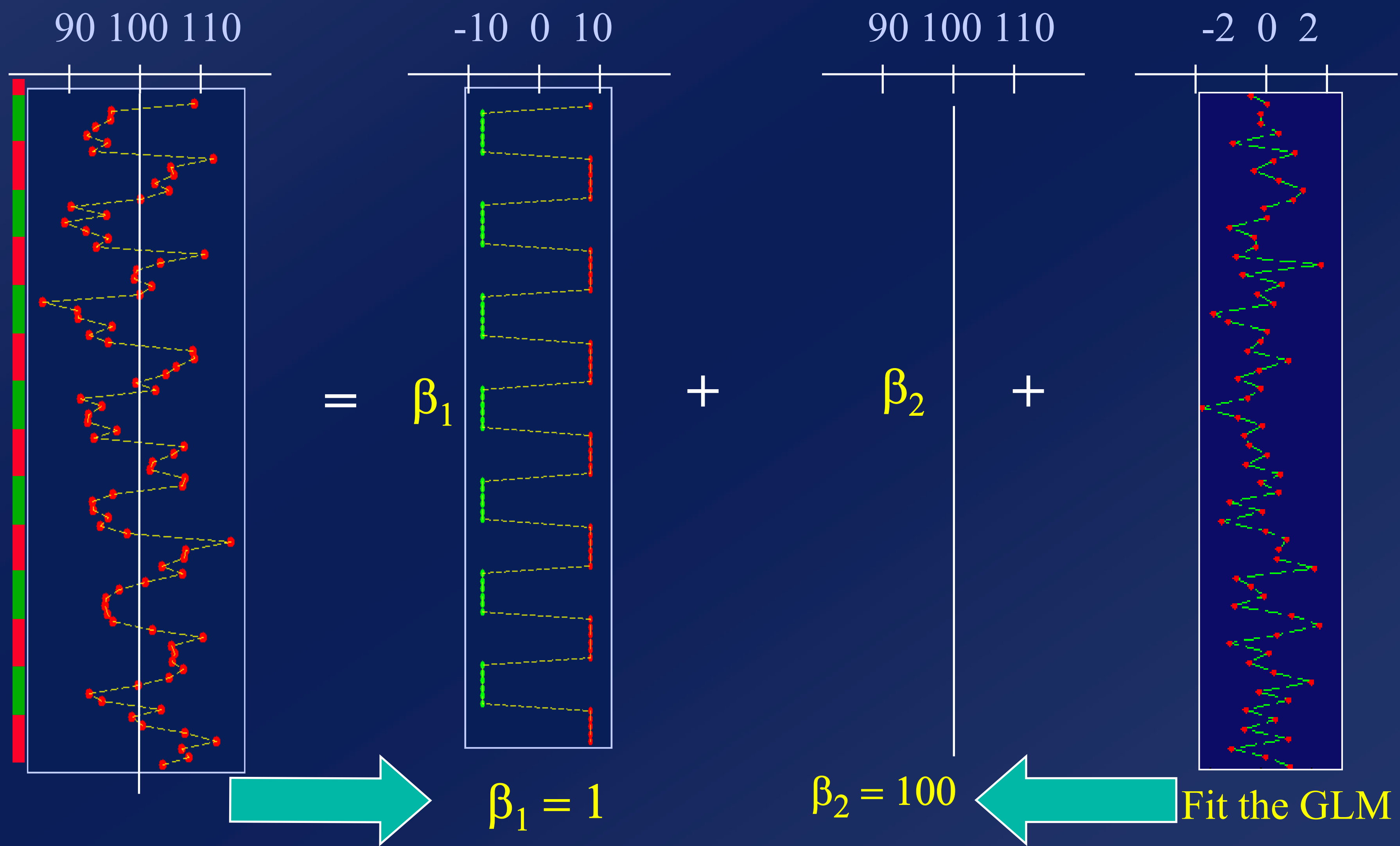
Plan

- *Make sure we know all about the estimation (fitting) part ....*
- *Make sure we understand the testing procedures : t- and F-tests*
- *A bad model ... And a better one*
- *Correlation in our model : do we mind ?*
- *A (nearly) real example*

One voxel = One test (t, F, ...)

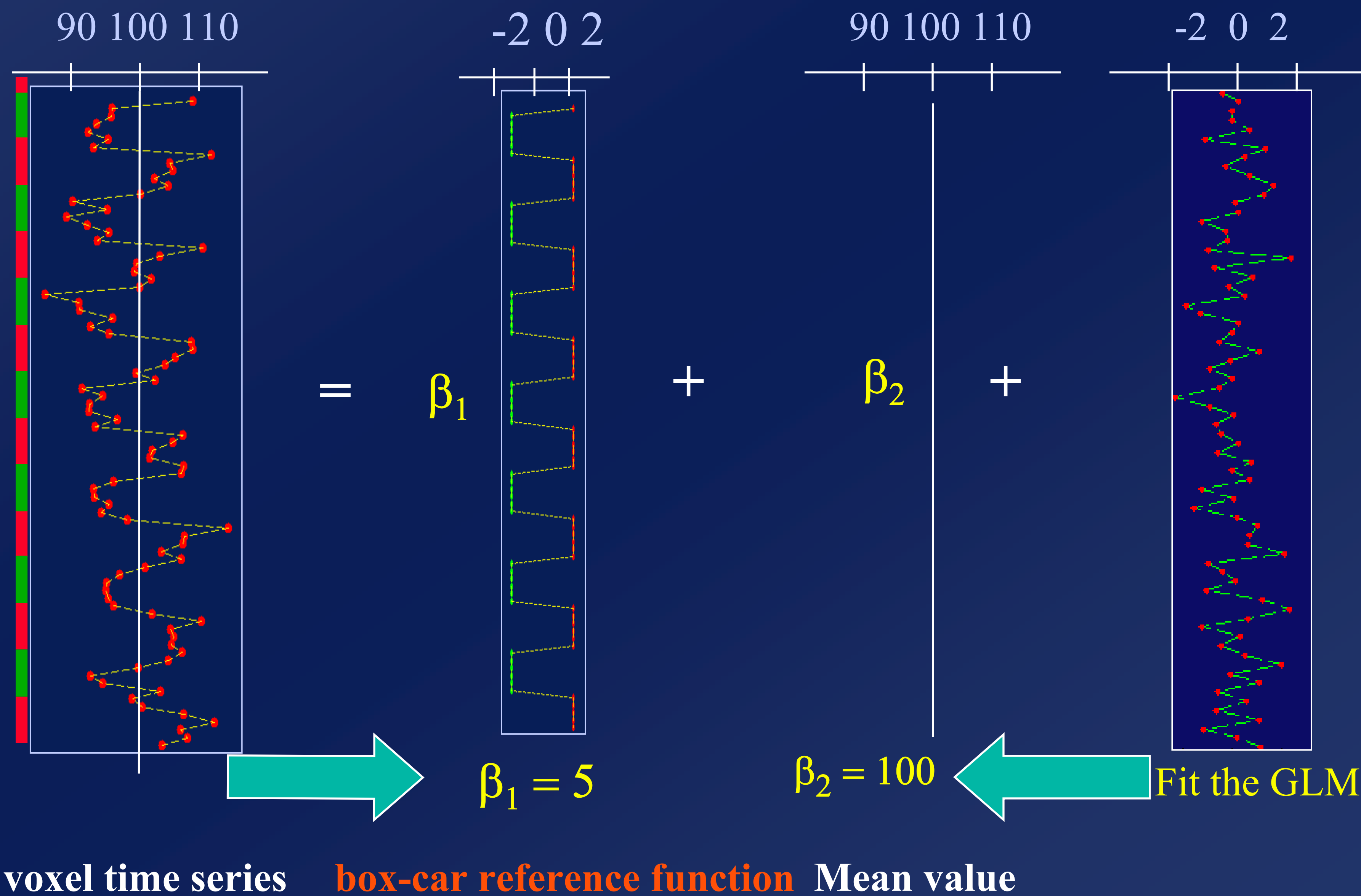


Regression example…

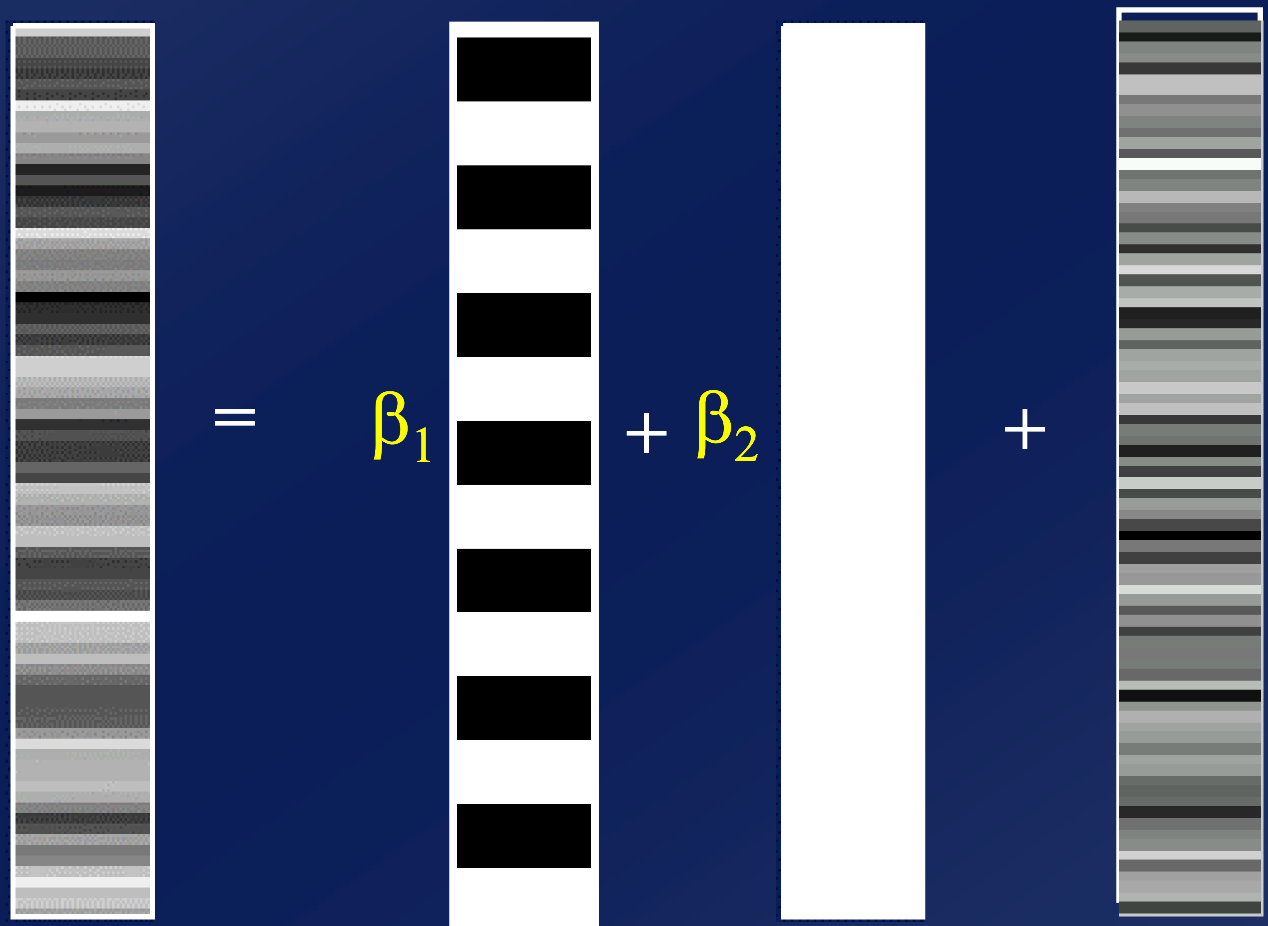


voxel time series **box-car reference function** **Mean value**

Regression example…



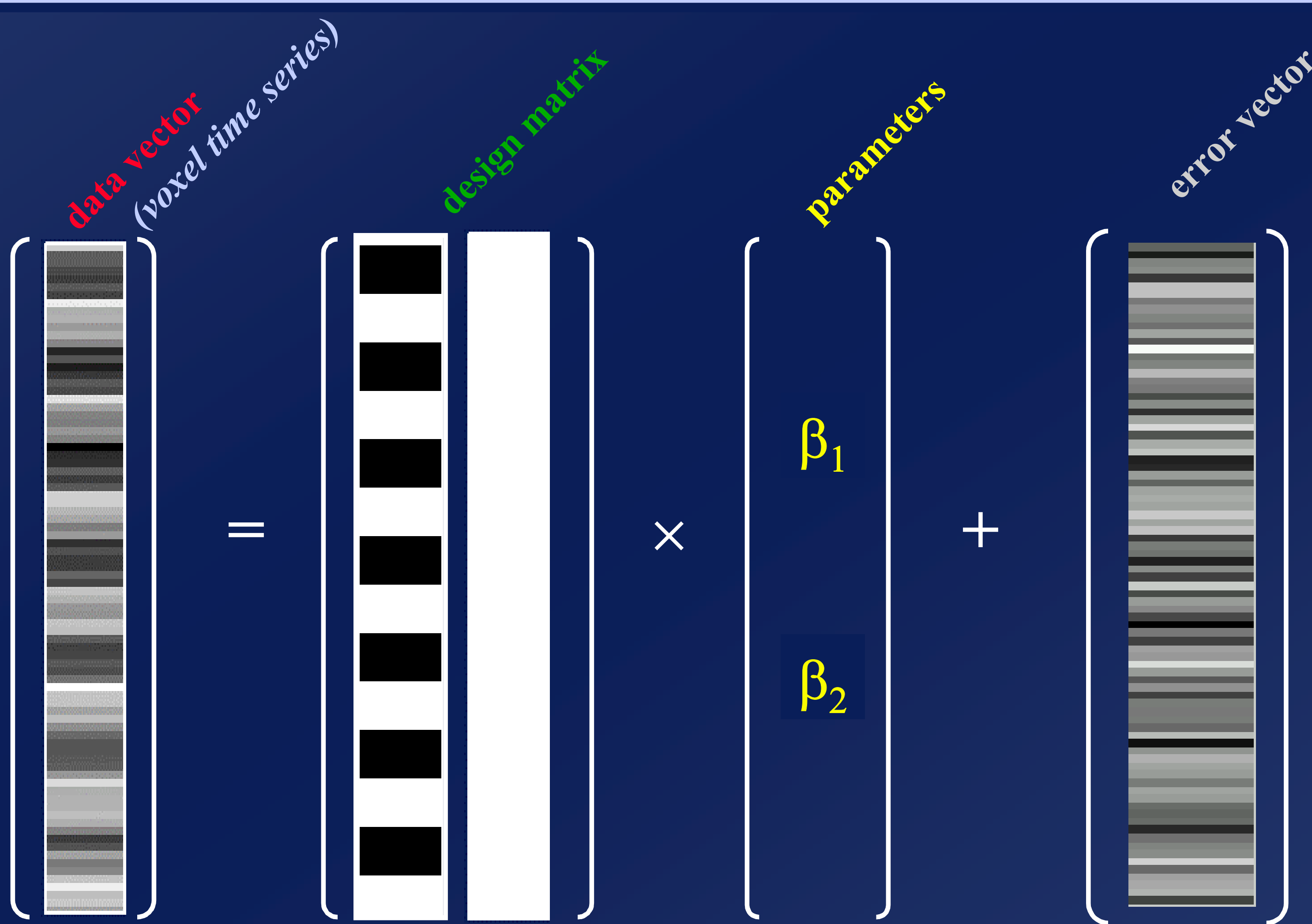
…revisited : matrix form



error

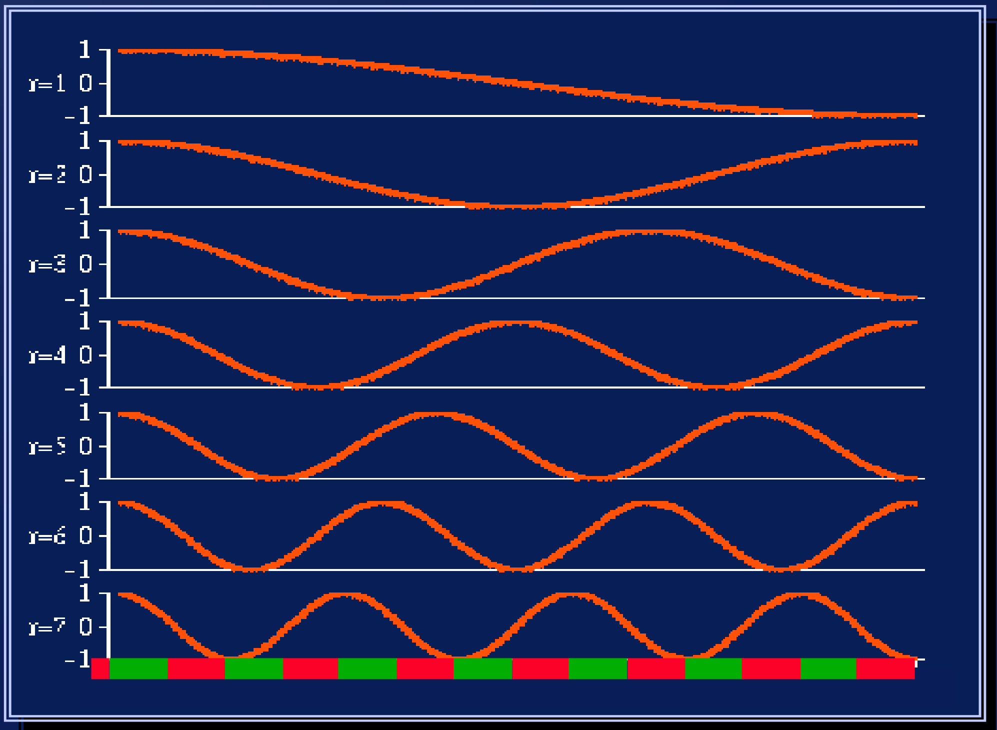
$$Y = \beta_1 \times f(t) + \beta_2 \times 1 + \varepsilon_s$$

Box car regression: design matrix…

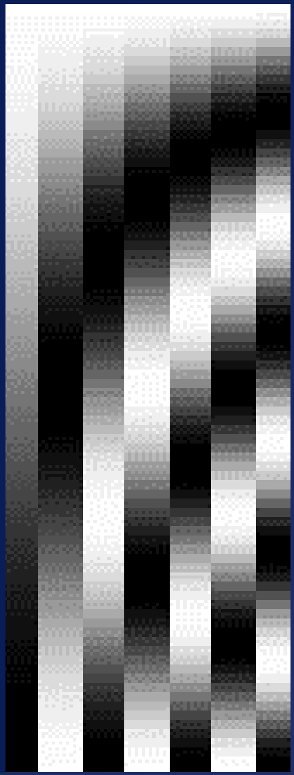


$$\underline{Y} = X \times \underline{\beta} + \underline{\varepsilon}$$

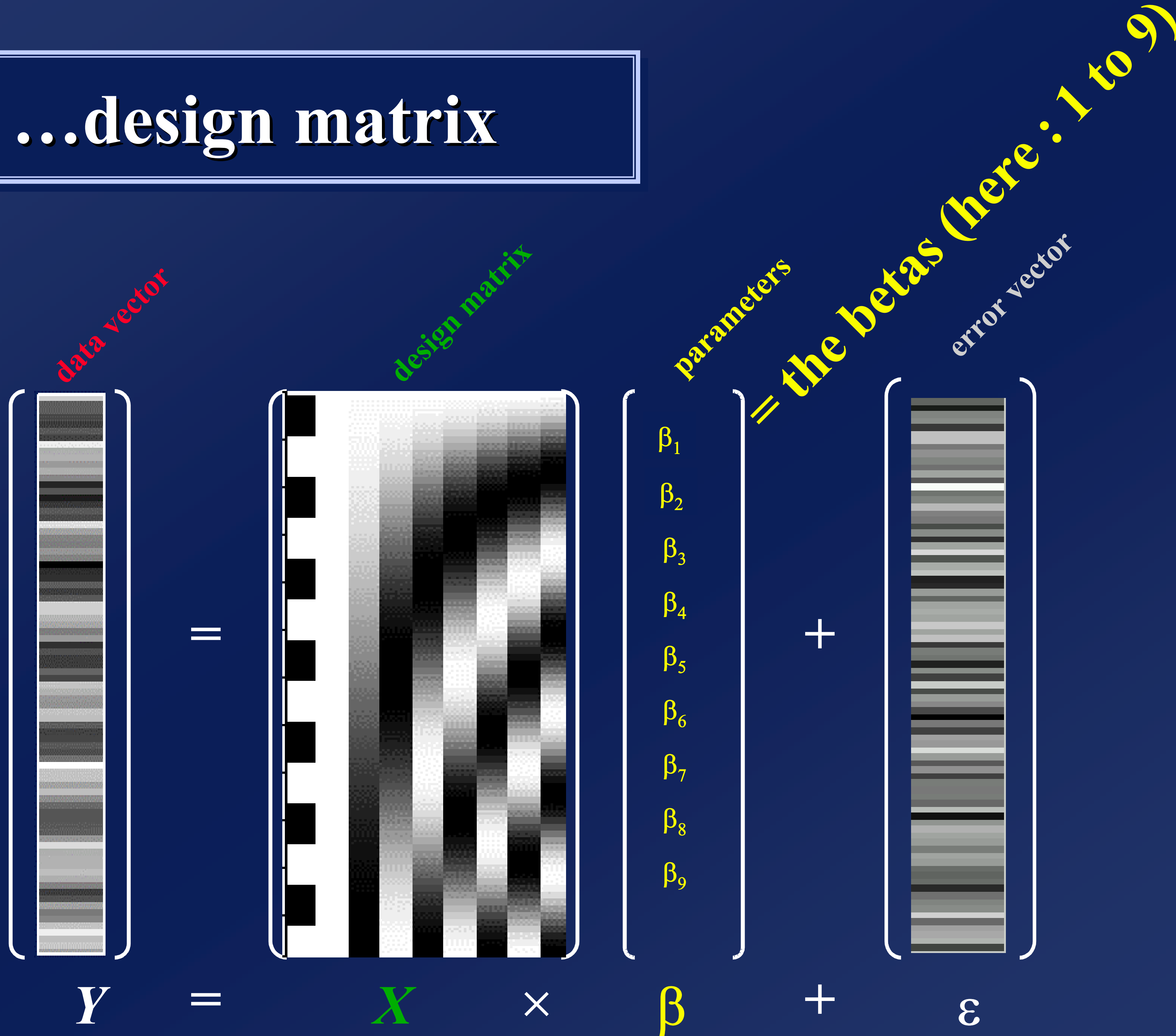
Add more reference functions ...



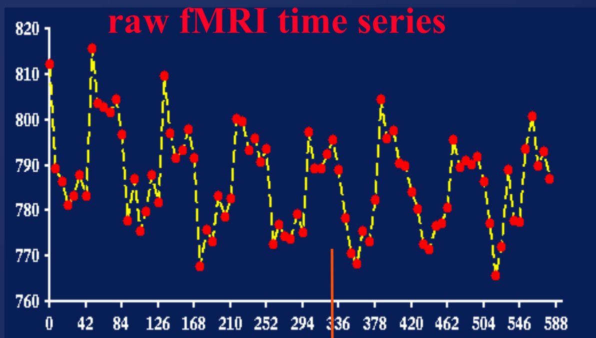
Discrete cosine transform basis functions

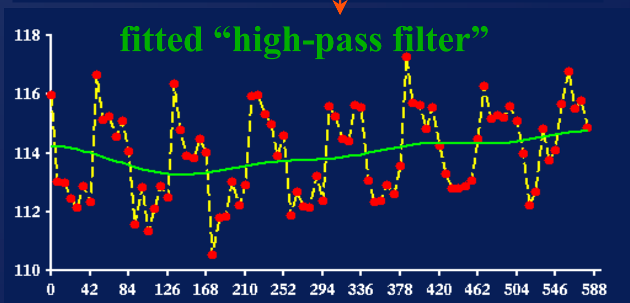


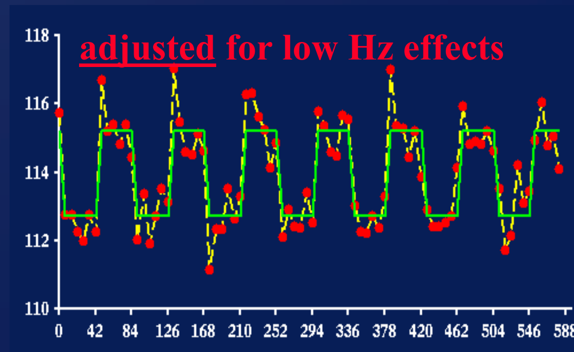
…design matrix



Fitting the model = finding some estimate of the betas = minimising the sum of square of the residuals S^2

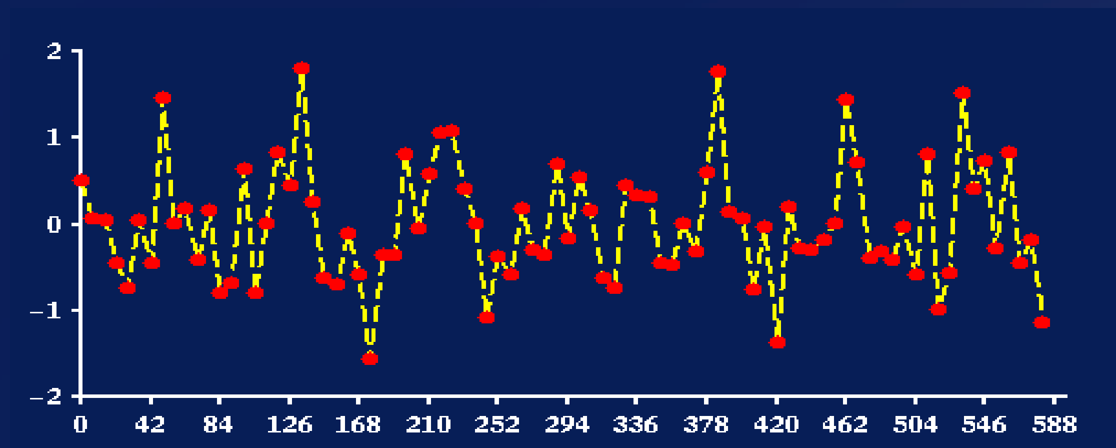






fitted box-car

residuals



$$\frac{\sum \text{the squared values of the residuals}}{\text{number of time points minus the number of estimated betas}} = s^2$$

Summary ...

- *We put in our model regressors (or covariates) that represent how we think the signal is varying (of interest and of no interest alike)*

- *Coefficients (= parameters) are estimated using the Ordinary Least Squares (OLS) or Maximum Likelihood (ML) estimator.*

- *These estimated parameters (the "betas")* ***depend*** *on the scaling of the regressors. But entered with SPM, regressors are normalised and comparable.*

- *The residuals, their sum of squares and the resulting tests (t,F),* ***do not*** *depend on the scaling of the regressors.*

Plan

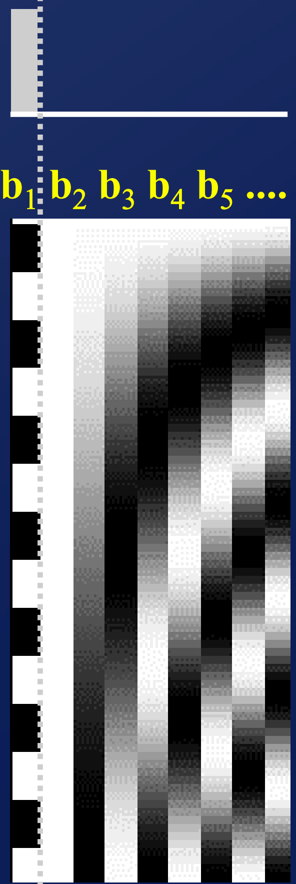
- *Make sure we all know about the estimation (fitting) part ....*
- *Make sure we understand t and F tests*
- *A (nearly) real example*
- *A bad model ... And a better one*
- *Correlation in our model : do we mind ?*

T test - one dimensional contrasts - SPM{t}

A contrast = a linear combination of **parameters:** $c' \times \beta$

$c' = 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0$

$b_1\ b_2\ b_3\ b_4\ b_5\ ....$



box-car amplitude > 0 ?

=

$\beta_1 > 0$?

=>

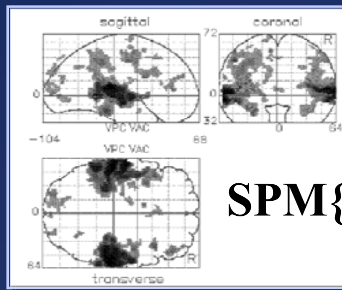
Compute $1 \mathrm{x} b_1 + 0 \mathrm{x} b_2 + 0 \mathrm{x} b_3 + 0 \mathrm{x} b_4 + 0 \mathrm{x} b_5 + \ldots$

and

divide by estimated standard deviation

$$T = \frac{\textit{contrast} \text{ of estimated parameters}}{\sqrt{\text{variance estimate}}}$$

$$T = \frac{c'b}{\sqrt{s^2 c'(X'X)^{+} c}}$$



SPM{t}

How is this computed ? (t-test)

$$\frac{\textit{contrast}\text{ of estimated parameters}}{\sqrt{\text{variance estimate}}}$$

Estimation [Y, X] [b, s]

$Y = X\beta + \varepsilon$ $\qquad$ $\varepsilon \sim \sigma^2 N(0,I)$ *(Y : at one position)*

$b = (X'X)^+ X'Y$ $\qquad$ *(b: estimate of β) -> beta??? images*

$e = Y - Xb$ $\qquad$ *(e: estimate of ε)*

$s^2 = (e'e/(n - p))$ $\qquad$ *(s: estimate of σ, n: time points, p: parameters)*

-> 1 image ResMS

Test [b, s², c] [c'b, t]

$\mathrm{Var}(c'b) = s^2c'(X'X)^+c$ $\qquad$ *(compute for each contrast c)*

$t = c'b / \mathrm{sqrt}(s^2c'(X'X)^+c)$ $\qquad$ *(c'b -> images spm_con???*

compute the t images -> images spm_t???)

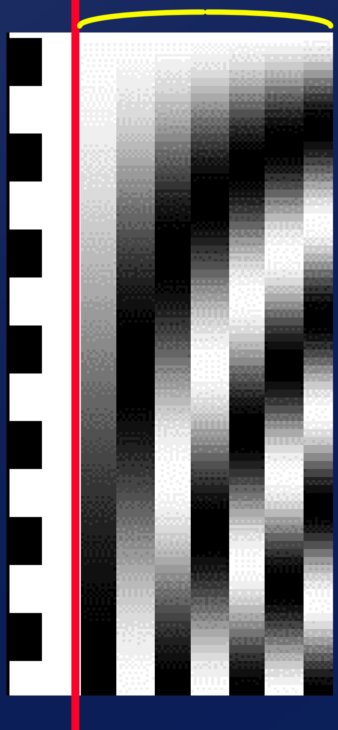
under the null hypothesis $H_0 : t \sim \text{Student-}t(df)$ $\qquad$ $df = n-p$

F-test (SPM{*F*}) : a reduced model or ...

Tests multiple linear hypotheses : Does X1 model anything ?

H$_0$: True (reduced) model is X_0

X_0 X_1



$\longrightarrow$ S^2

This (full) model ?

X_0

$\longrightarrow$ S_0^2

Or this one?

$$F = \frac{\text{additional variance accounted for by tested effects}}{\text{error variance estimate}}$$

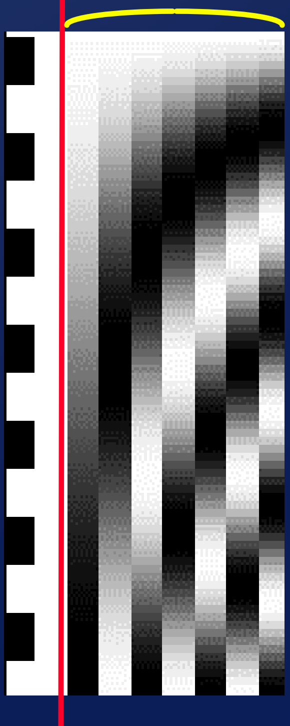
$$F \sim (S_0^2 - S^2) / S^2$$

F-test (SPM{F}) : a reduced model or ... multi-dimensional contrasts ?

tests multiple linear hypotheses. Ex : does DCT set model anything?

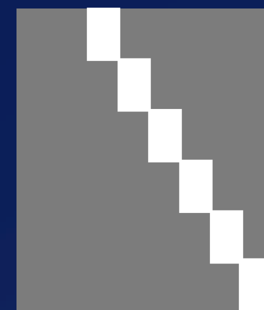
H_0: True model is X_0 $\qquad$ H_0: $\beta_{3\text{-}9} = (0\ 0\ 0\ 0\ \ldots)$ $\qquad$ test H_0 : $c' \times b = 0$?

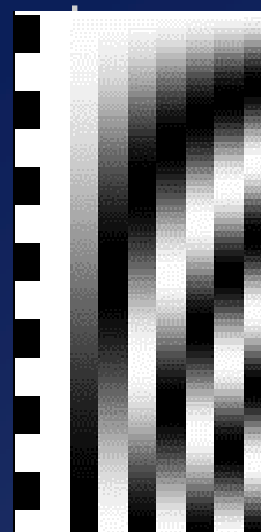
X_0 $\qquad$ X_1 $(\beta_{3\text{-}9})$



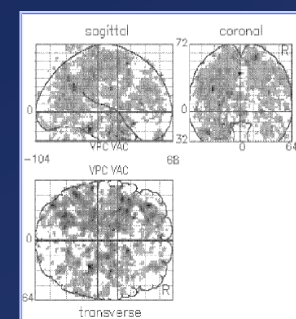
X_0

$$c' = \begin{matrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{matrix}$$





SPM{F}



This model ? $\qquad$ Or this one ?

How is this computed ? (F-test)

$$\frac{\text{additional variance accounted for by tested effects}}{\text{Error variance estimate}}$$

Estimation [Y, X] [b, s]

$$Y = X\beta + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I)$$

$$Y = X_0 \beta_0 + \varepsilon_0 \qquad \varepsilon_0 \sim N(0, \sigma_0^2 I) \qquad X_0 : X \textit{ Reduced}$$

Estimation [Y, X_0] [b_0, s_0]

$$b_0 = (X_0{}'X_0)^+ X_0{}'Y$$

$$e_0 = Y - X_0 b_0 \qquad (e_\Diamond\text{: estimate of } \varepsilon_\Diamond)$$

$$s^2{}_0 = (e_0{}'e_0/(n - p_0)) \qquad (s_\Diamond\text{: estimate of } \sigma_\Diamond,\ n\text{: time, } p_\Diamond\text{: parameters})$$

Test [b, s, c] [ess, F]

$$F \sim \boxed{(s_0 - s) / s^2}$$

-> image spm_ess???

-> image of F : spm_F???

under the null hypothesis : $F \sim F(p - p0, n\text{-}p)$

Plan

- *Make sure we all know about the estimation (fitting) part ....*
- *Make sure we understand t and F tests*
- *A (nearly) real example : testing main effects and interactions*
- *A bad model ... And a better one*
- *Correlation in our model : do we mind ?*

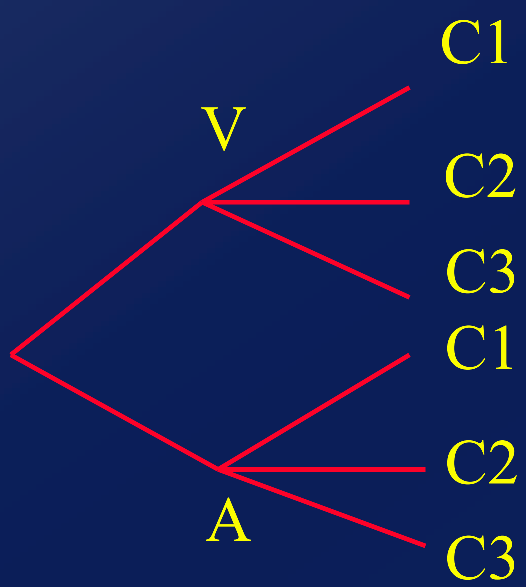
A real example (almost !)

Experimental Design → Design Matrix

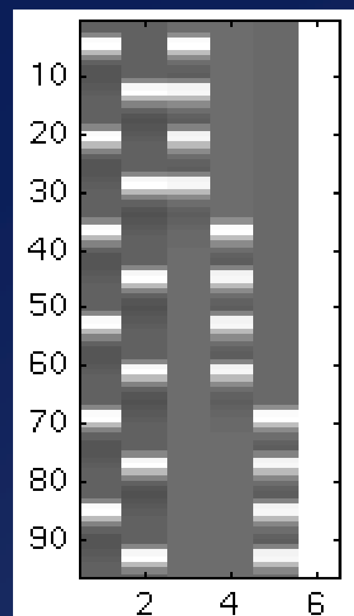
Factorial design with 2 factors : modality and category

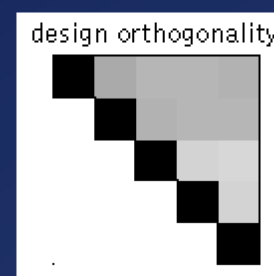
2 levels for modality (eg Visual/Auditory)

3 levels for category (eg 3 categories of words)

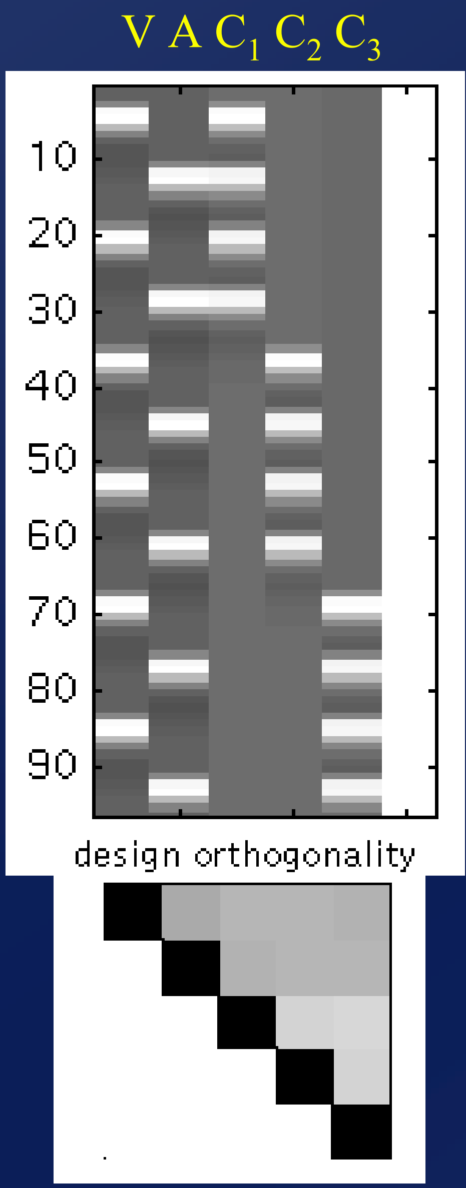


V A C_1 C_2 C_3





Asking ourselves some questions ...



Test C1 > C2 : c = [0 0 1 -1 0 0]

Test V > A : c = [1 -1 0 0 0 0]

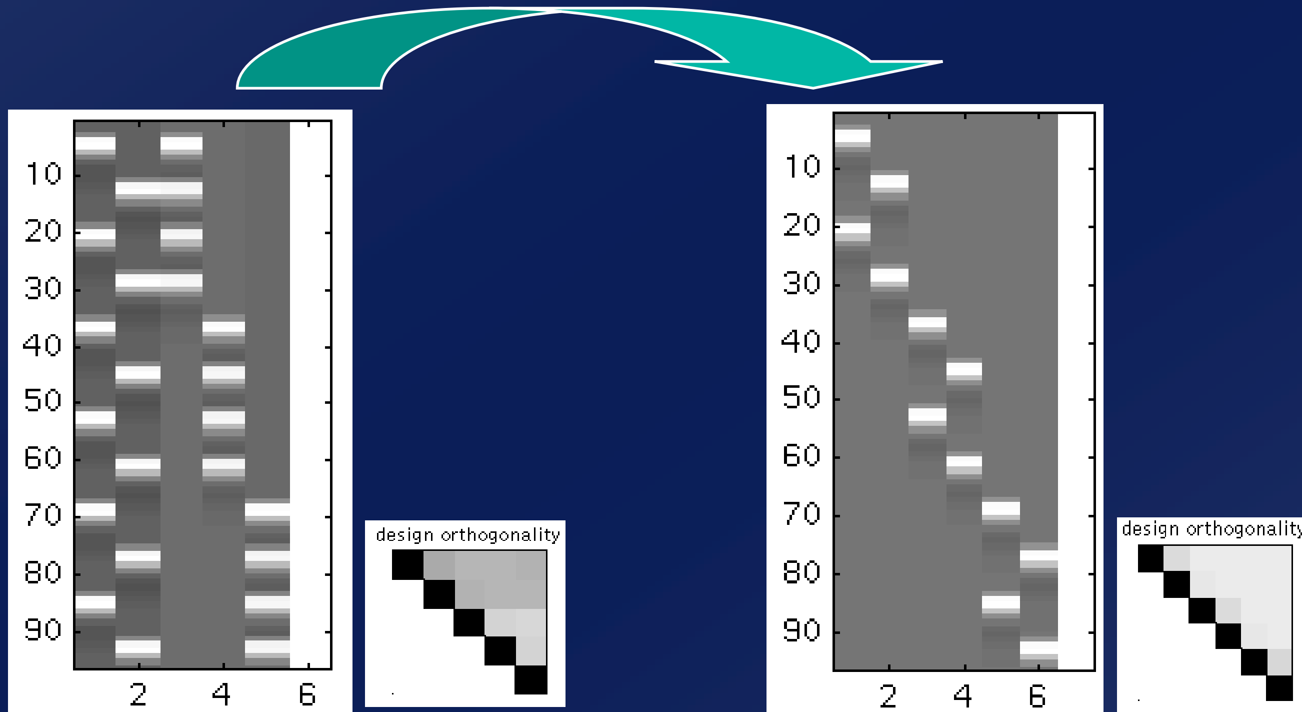
Test C1,C2,C3 ? (F)

$$c = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

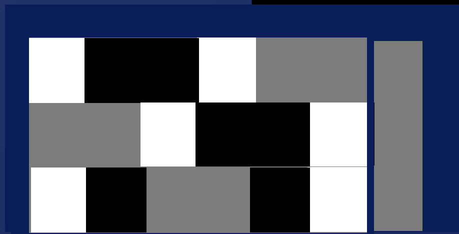
Test the interaction MxC ?

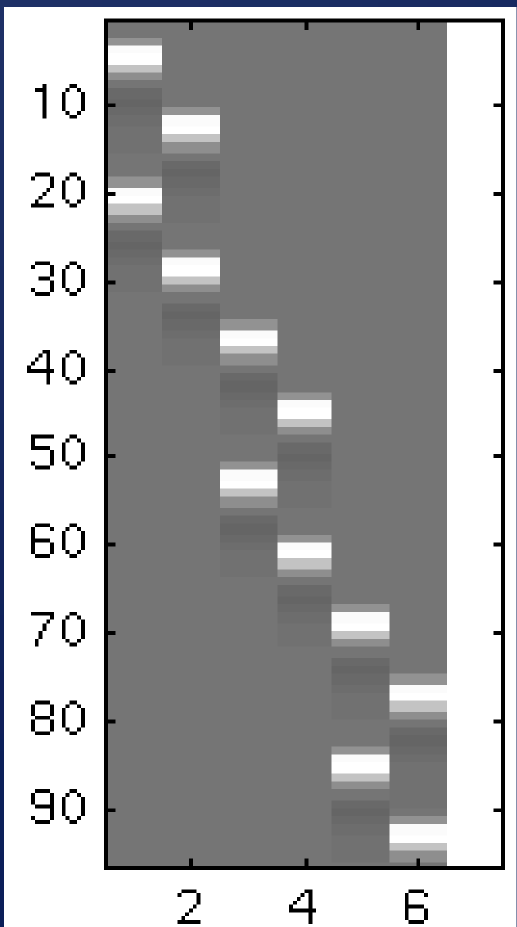
- Design Matrix not orthogonal
- Many contrasts are non estimable
- Interactions MxC are not modelled

Modelling the interactions



Asking ourselves some questions ...





Test C1 > C2 : c = [1 1 -1 -1 0 0 0]

Test V > A : c = [1 -1 1 -1 1 -1 0]

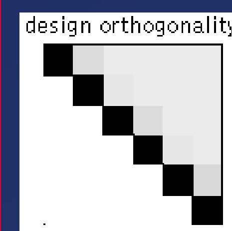
Test the categories :

c =
[1 1 -1 -1 0 0 0]
[0 0 1 1 -1 -1 0]
[1 1 0 0 -1 -1 0]

Test the interaction MxC :

c =
[1 -1 -1 1 0 0 0]
[0 0 1 -1 -1 1 0]
[1 -1 0 0 -1 1 0]

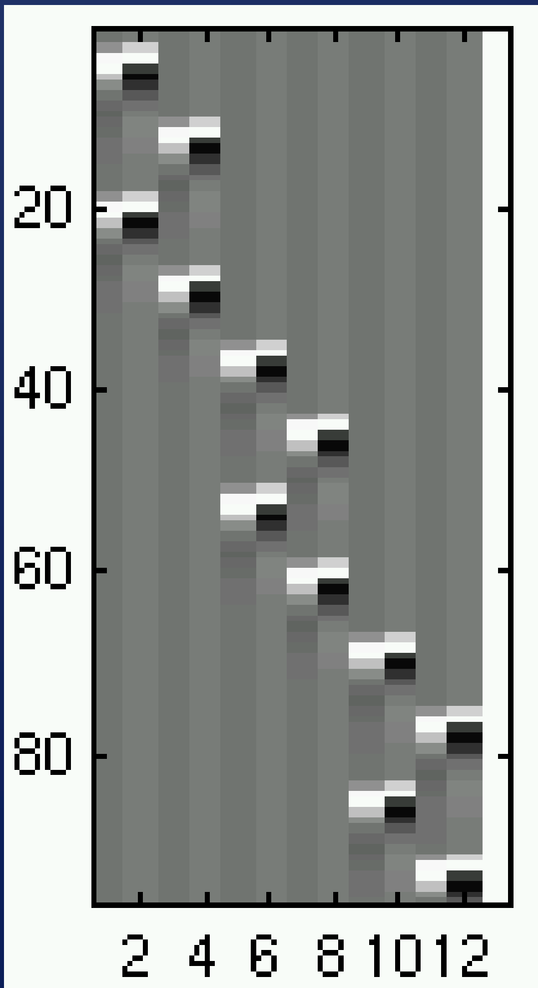
- Design Matrix orthogonal
- All contrasts are estimable
- Interactions MxC modelled
- If no interaction ... ? Model is too “big” !



Asking ourselves some questions ... With a more flexible model

C_1 C_1 C_2 C_2 C_3 C_3

V A V A V A



Test C1 > C2 ?

Test C1 different from C2 ?

from

c = [1 1 -1 -1 0 0 0]

to

c = [1 0 1 0 -1 0 -1 0 0 0 0 0 0]

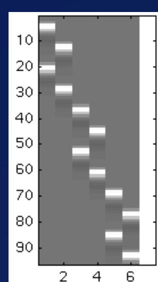
[0 1 0 1 0 -1 0 -1 0 0 0 0 0]

becomes an F test!

Test V > A ?

c = [1 0 -1 0 1 0 -1 0 1 0 -1 0 0]

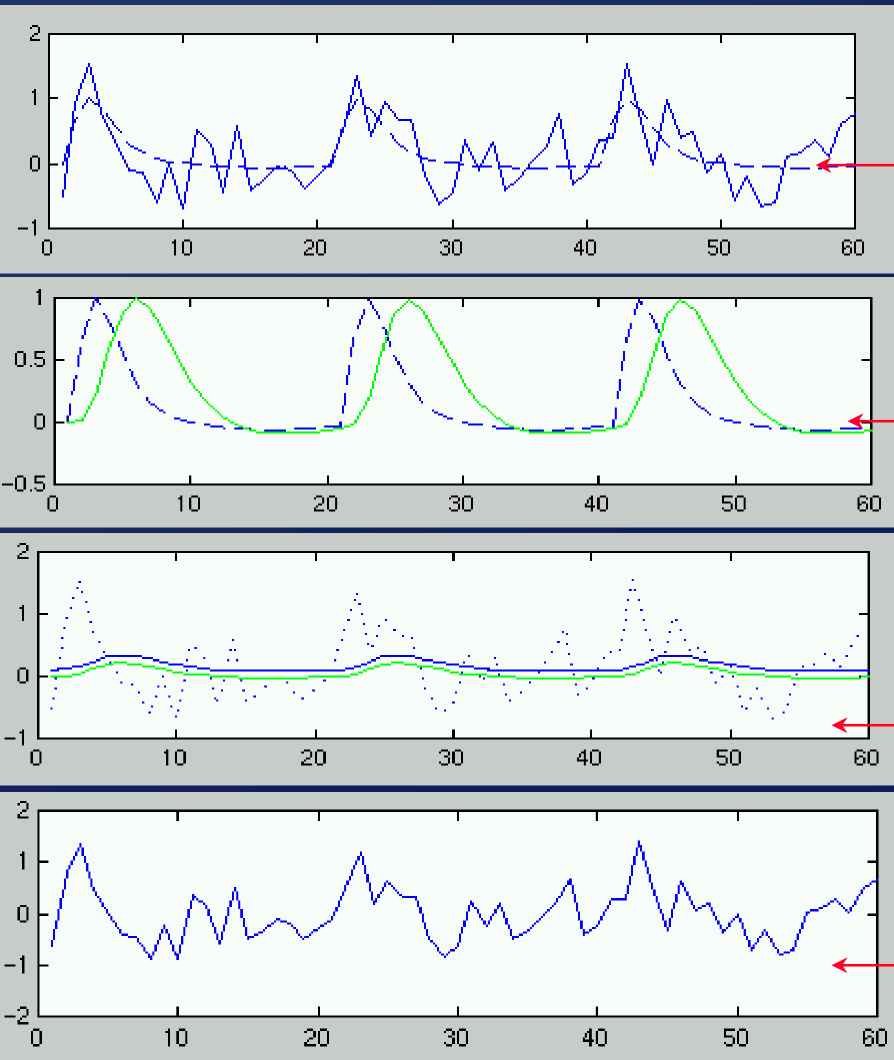
is possible, but is OK only if the regressors coding for the delay are all equal



Plan

- *Make sure we all know about the estimation (fitting) part ....*
- *Make sure we understand t and F tests*
- *A (nearly) real example*
- *A bad model ... And a better one*
- *Correlation in our model : do we mind ?*

A **bad** model ...



True signal and observed signal (---)

Model (green, pic at 6sec)
TRUE signal (blue, pic at 3sec)

Fitting (b1 = 0.2, mean = 0.11)

Residual (still contains some signal)

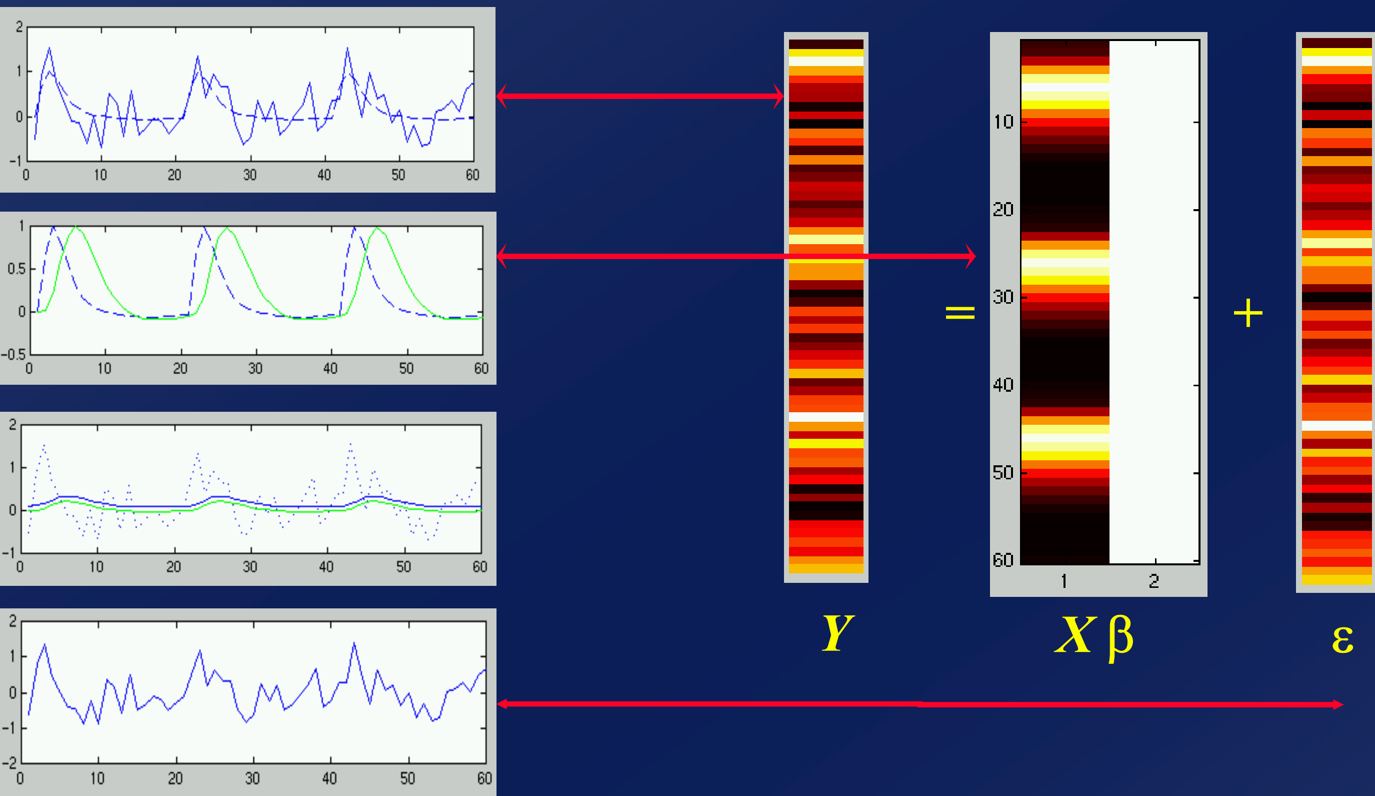
=> Test for the green regressor not significant

A bad model ...

$\beta_1 = 0.22$

$\beta_2 = 0.11$

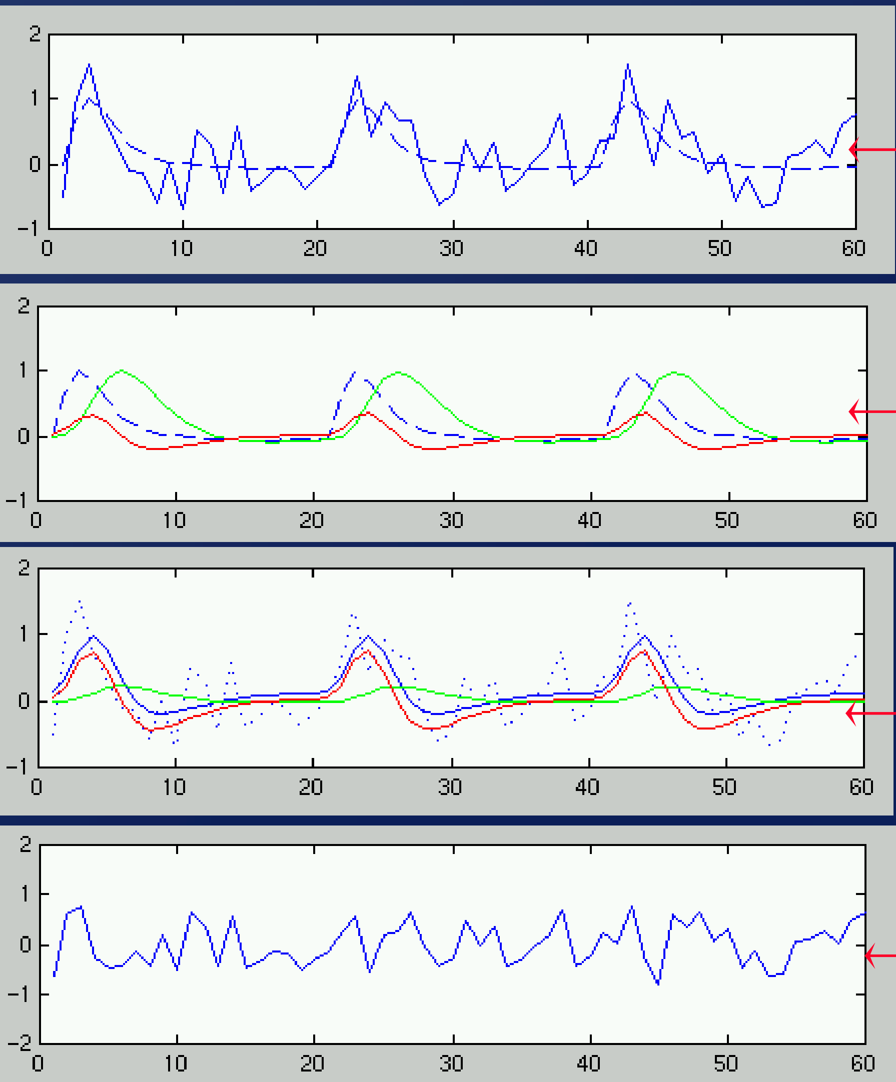
Residual Variance $= 0.3$



$P(Y|\ \beta_1 = 0) \Rightarrow$
p-value = 0.1
(t-test)

$P(Y|\ \beta_1 = 0) \Rightarrow$
p-value = 0.2
(F-test)

A « better » model ...



True signal + observed signal

Model (green and red)
and true signal (blue ---)
Red regressor : temporal derivative of the green regressor

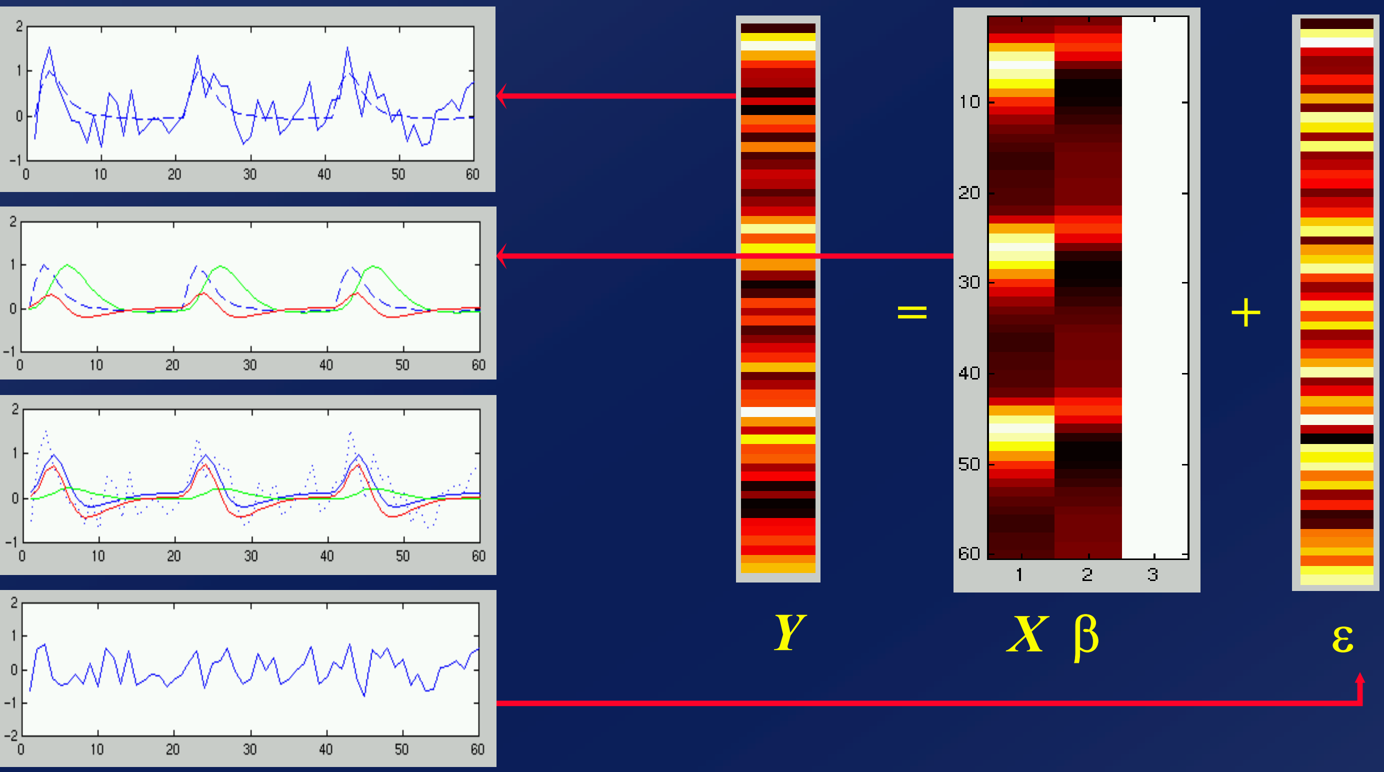
Global fit (blue)
and partial fit (green & red)
Adjusted and fitted signal

Residual (a smaller variance)

=> t-test of the green regressor significant
=> F-test very significant
=> t-test of the red regressor very significant

A better model ...

$\beta_1 = 0.22$
$\beta_2 = 2.15$
$\beta_3 = 0.11$

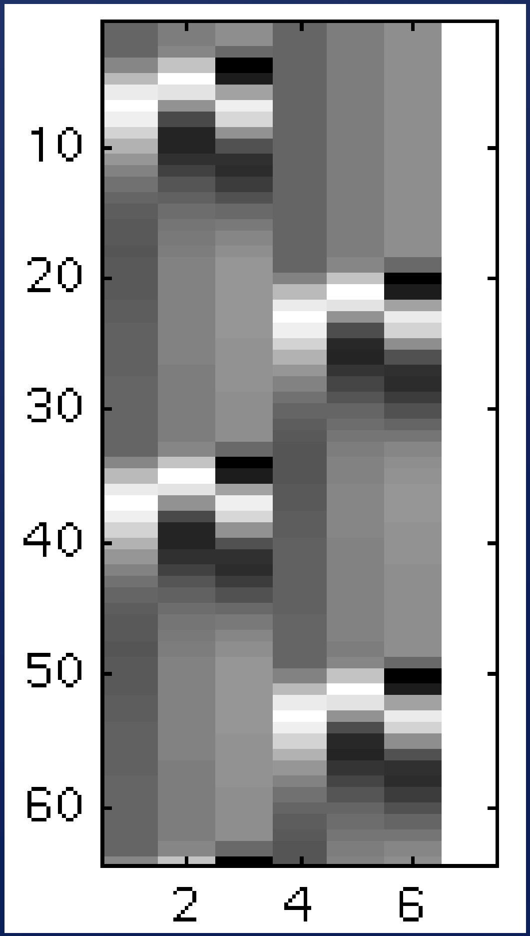


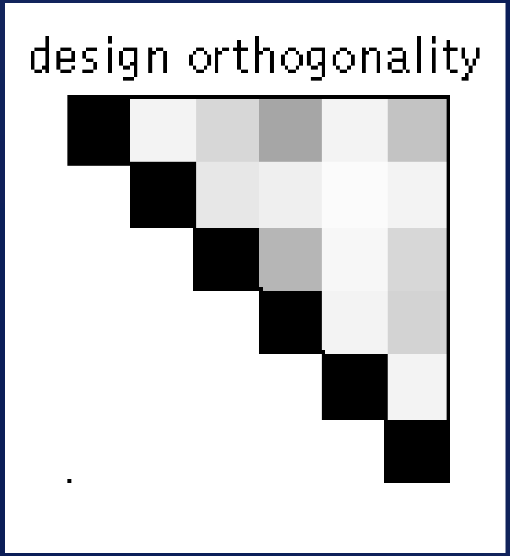
Residual Var = 0.2

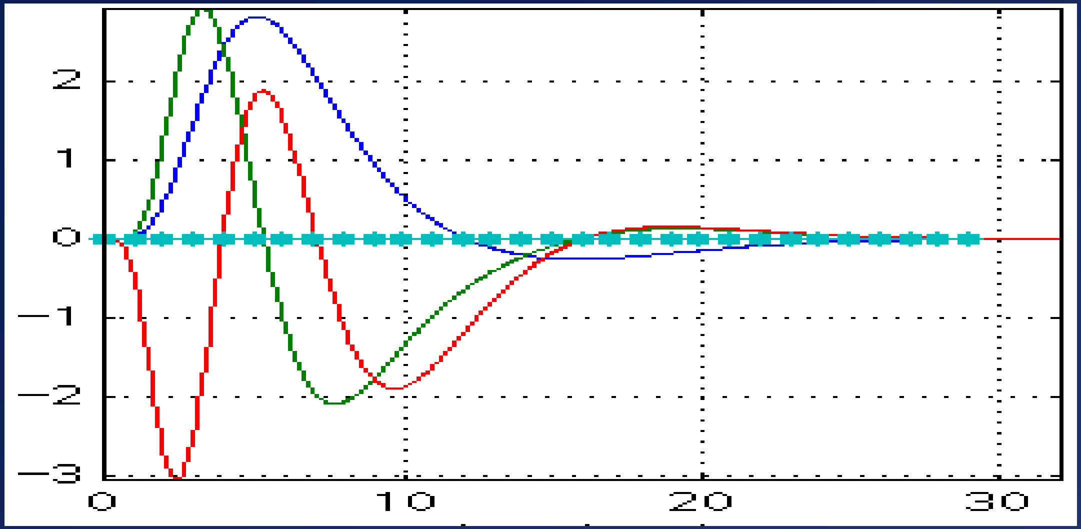
$P(Y|\ \beta_1 = 0)$
p-value = 0.07
(t-test)

$P(Y|\ \beta_1 = 0, \beta_2 = 0)$
p-value = 0.000001
(F-test)

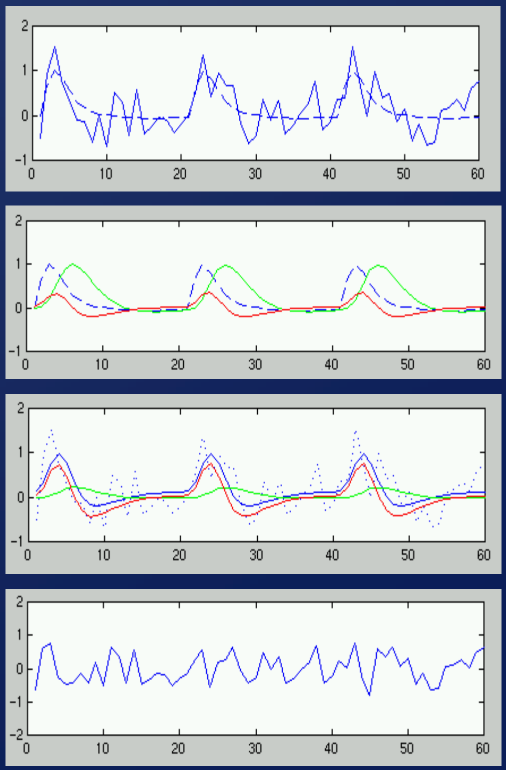
Flexible models : Gamma Basis







Summary ... (2)

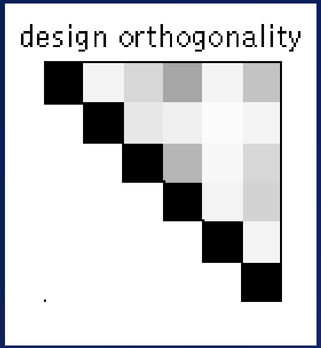


- *The residuals should be looked at ...!*
- *Test flexible models if there is little a priori information*
- *In general, use the F-tests to look for an overall effect, then look at the response shape*

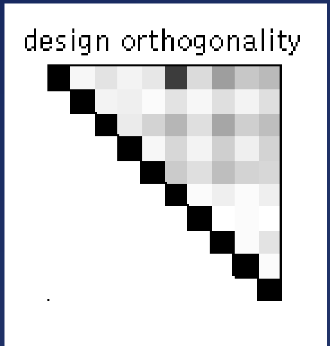
- *Interpreting the test on a single parameter (one regressor) can be difficult: cf the delay or magnitude situation*
- *BRING ALL PARAMETERS AT THE 2nd LEVEL*

Plan

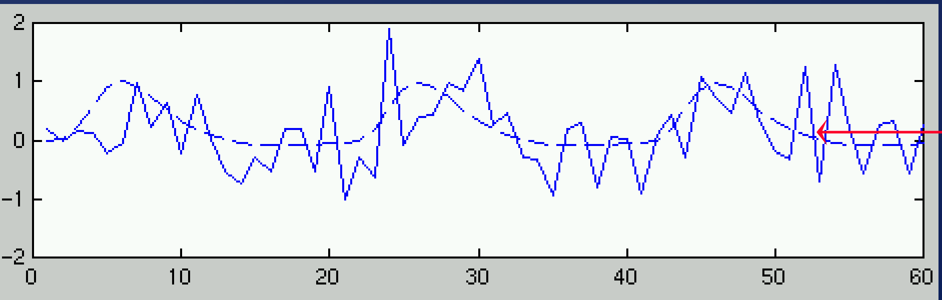
- *Make sure we all know about the estimation (fitting) part ....*
- *Make sure we understand t and F tests*
- *A (nearly) real example*
- *A bad model ... And a better one*
- *Correlation in our model : do we mind ?*



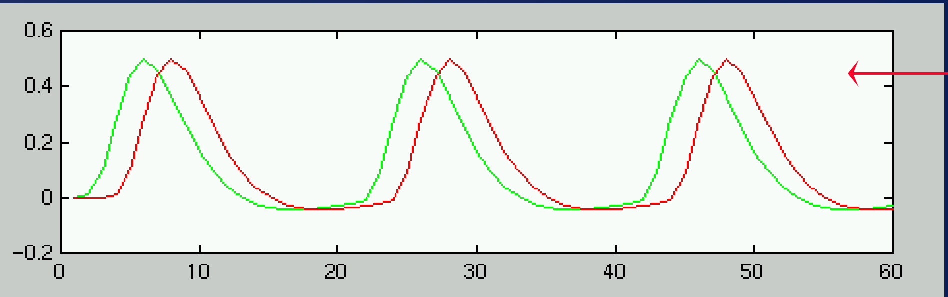
?



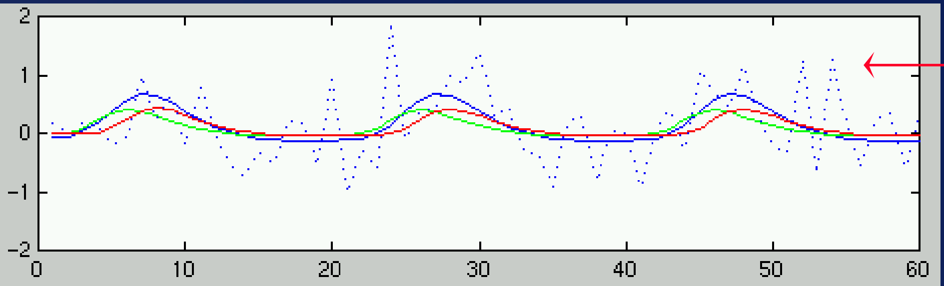
Correlation between regressors



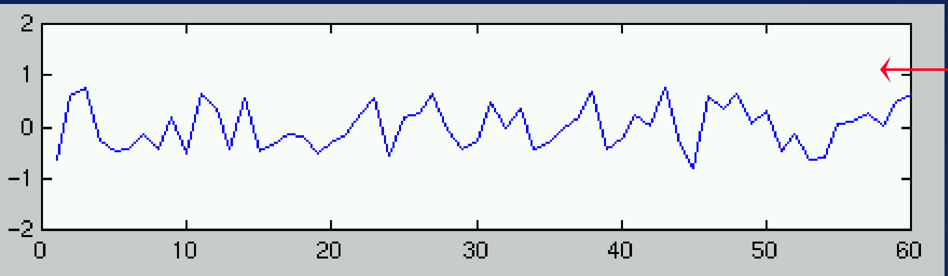
True signal



Model (green and red)



Fit (blue : global fit)



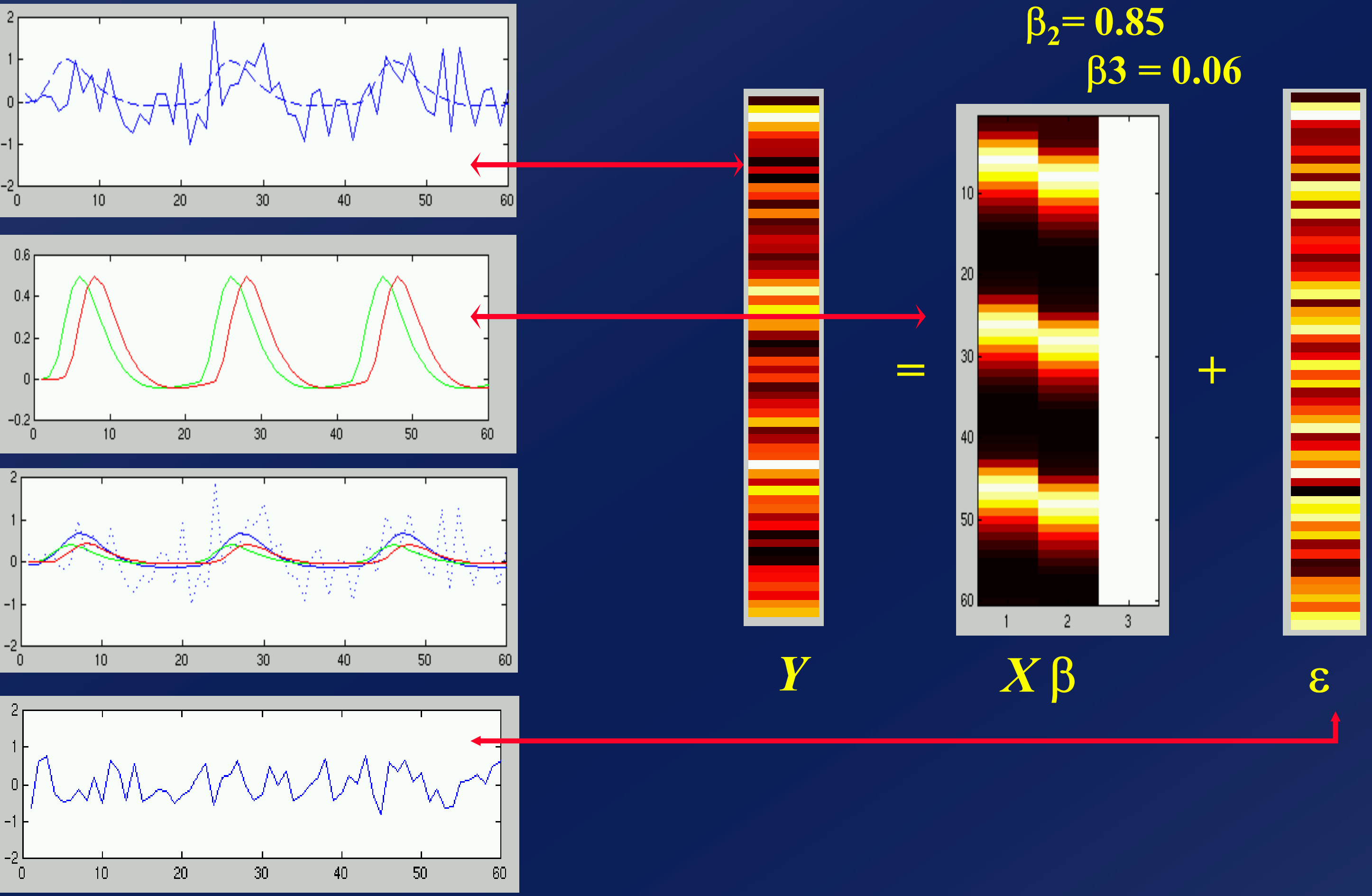
Residual

Correlation between regressors

$\beta_1 = 0.79$

$\beta_2 = 0.85$

$\beta 3 = 0.06$



Residual var. = 0.3

$P(Y|\ \beta_1 = 0)$

p-value = 0.08

(t-test)

$P(Y|\ \beta_2 = 0)$

p-value = 0.07

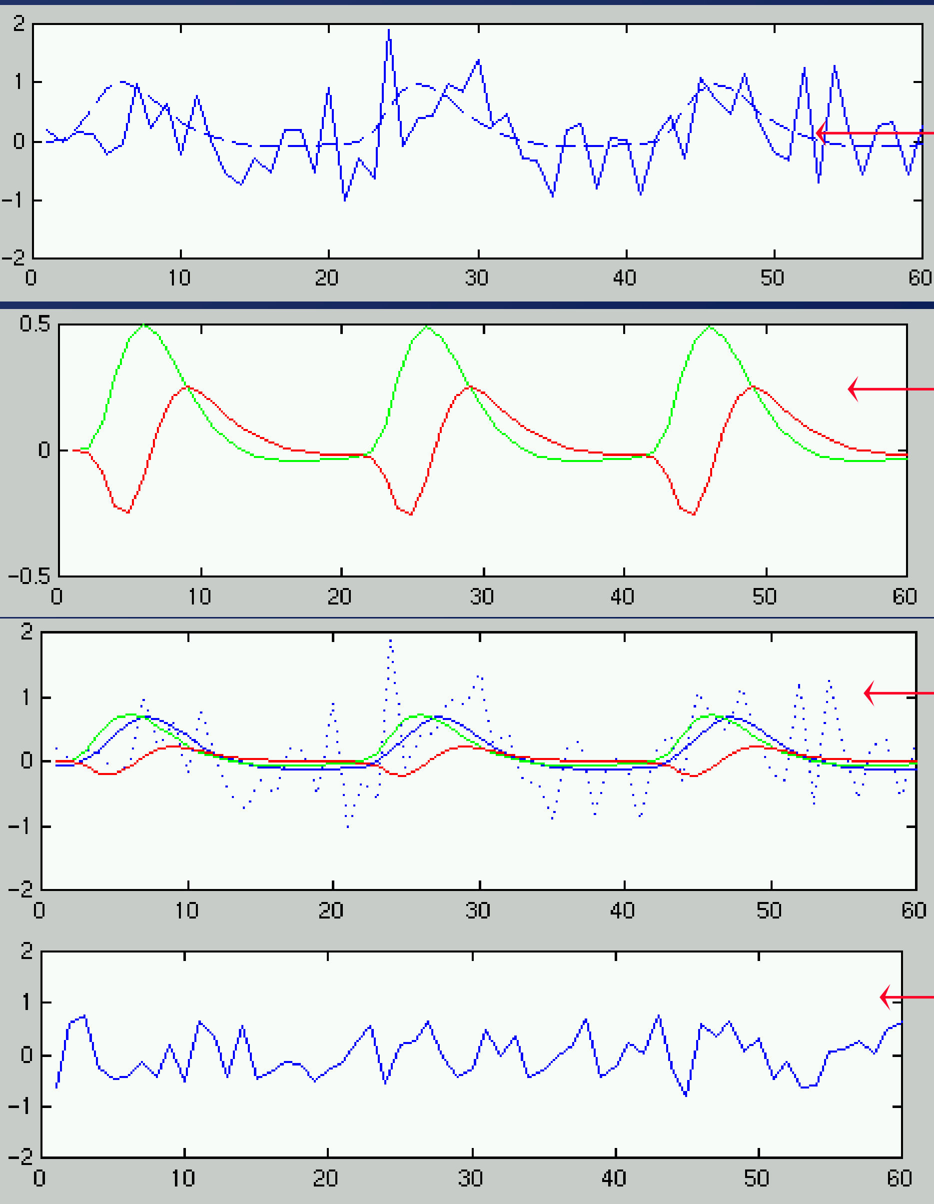
(t-test)

$P(Y|\ \beta_1 = 0, \beta_2 = 0)$

p-value = 0.002

(F-test)

Correlation between regressors - 2



true signal

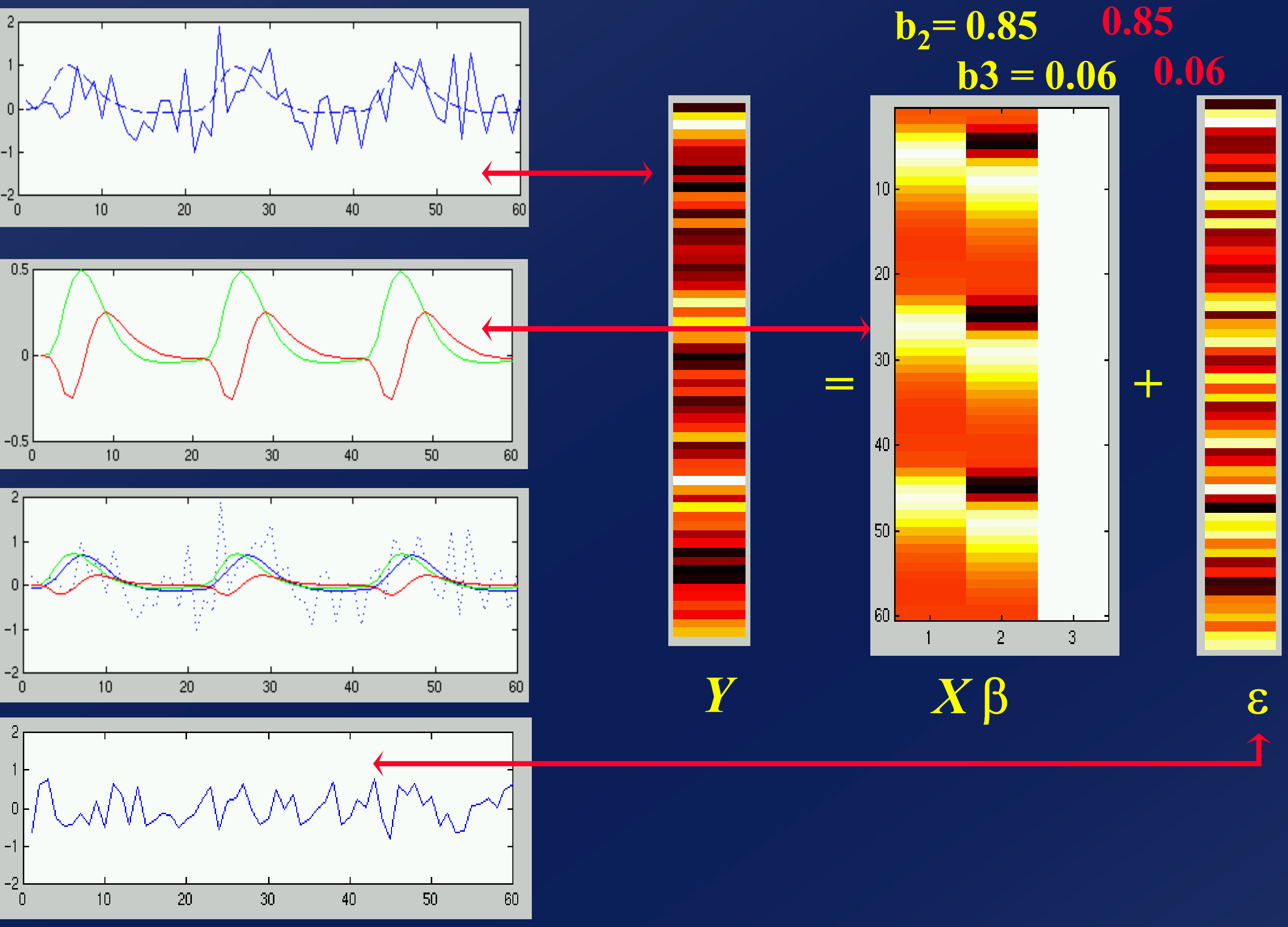
Model (green and red)
red regressor has been
orthogonalised with respect to the green one
⇔ remove everything that correlates with
the green regressor

Fit

Residual

Correlation between regressors -2

$b_1 = 1.47$ 0.79***
$b_2 = 0.85$ 0.85
b3 = 0.06 0.06



Residual var. = 0.3

$P(Y|\ \beta_1 = 0)$
p-value = 0.0003
(t-test)

$P(Y|\ \beta_2 = 0)$
p-value = 0.07
(t-test)

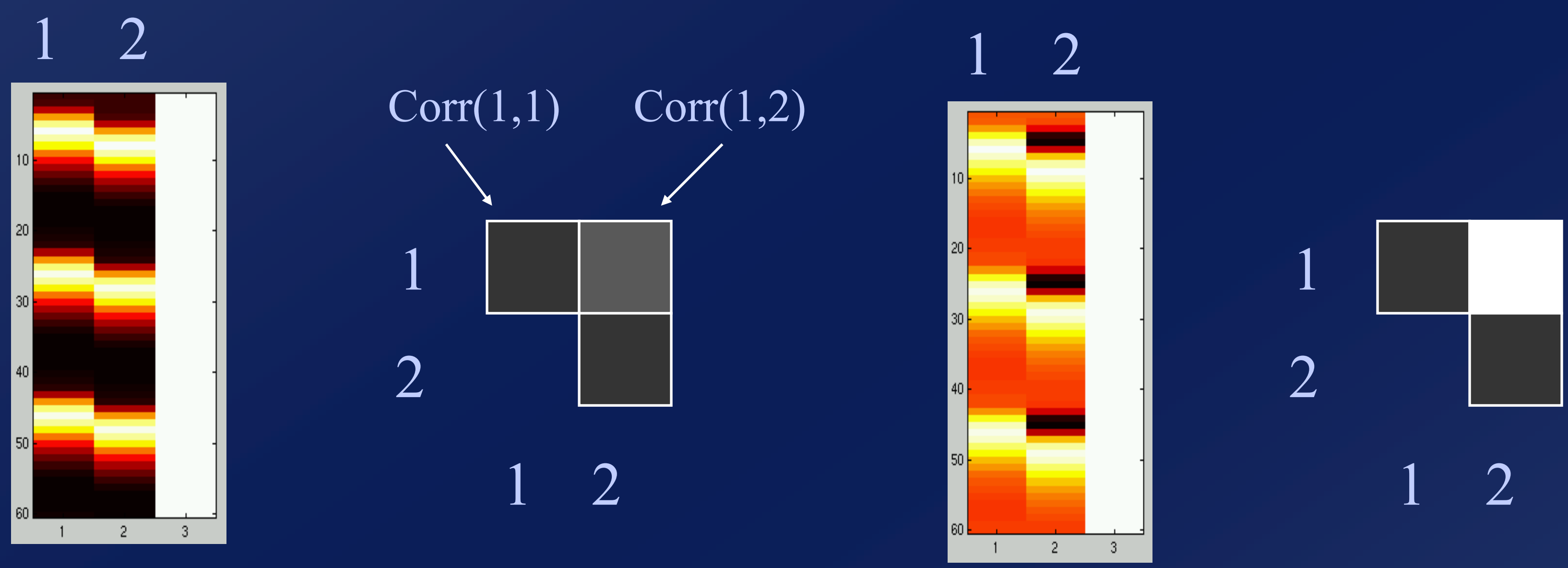
$P(Y|\ \beta_1 = 0, \beta_2 = 0)$
p-value = 0.002
(F-test)

See « explore design »

Design orthogonality : « explore design »

Black = completely correlated

White = completely orthogonal



Beware: when there are more than 2 regressors (C1,C2,C3,...), you may think that there is little correlation (light grey) between them, but C1 + C2 + C3 may be correlated with C4 + C5

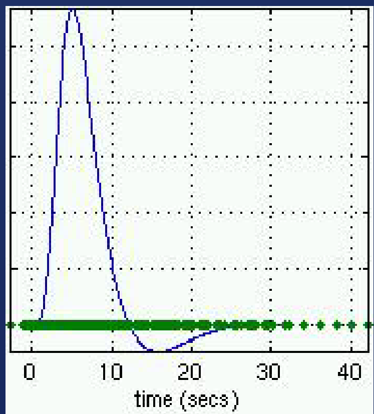
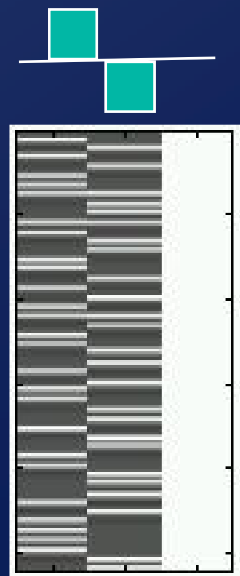
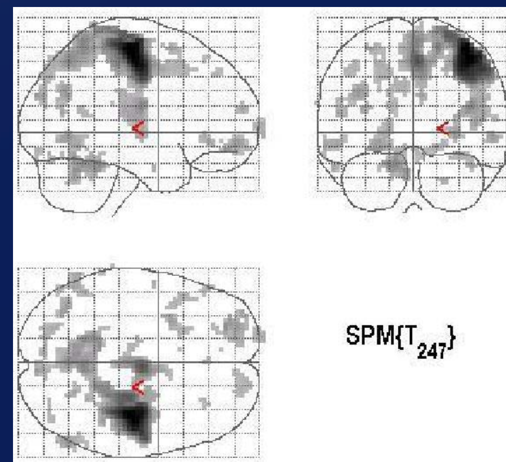
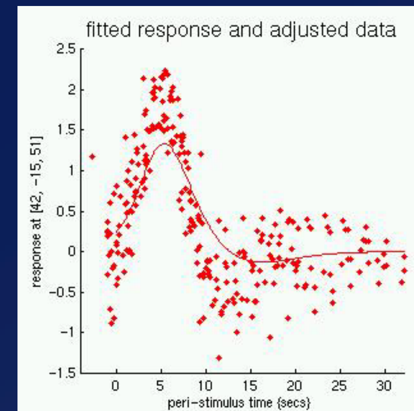
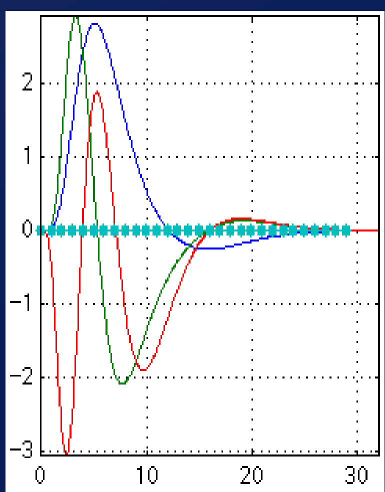
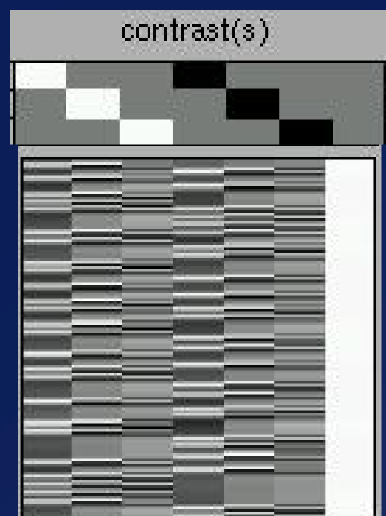
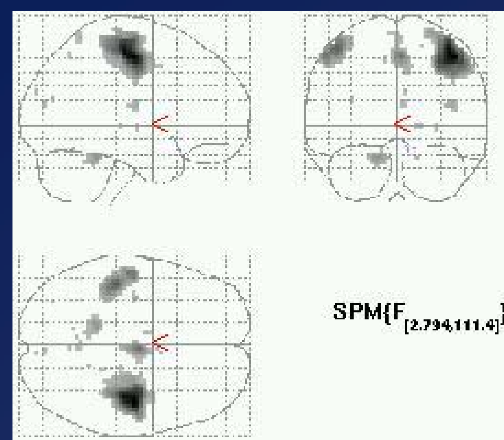
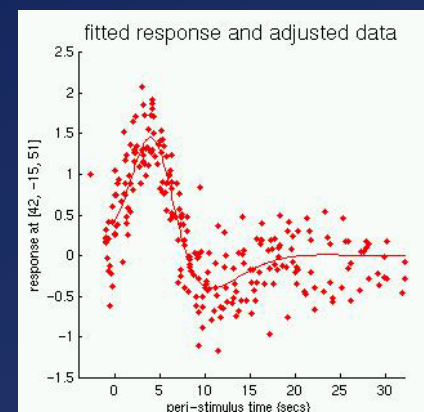
Summary ... (3)

- *We implicitly test for an <u>additional</u> effect only, be careful if there is correlation*

- *Orthogonalisation = decorrelation*
- This is not generally needed
- Parameters and test on the non modified regressor change

- *It is always simpler to have orthogonal regressors and therefore designs !*

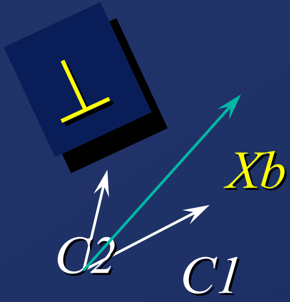
- *In case of correlation, use F-tests to see the overall significance. There is generally no way to decide to which regressor the « common » part should be attributed to*

Convolution model	Design and contrast	SPM(t) or SPM(F)	Fitted and adjusted data
			
			

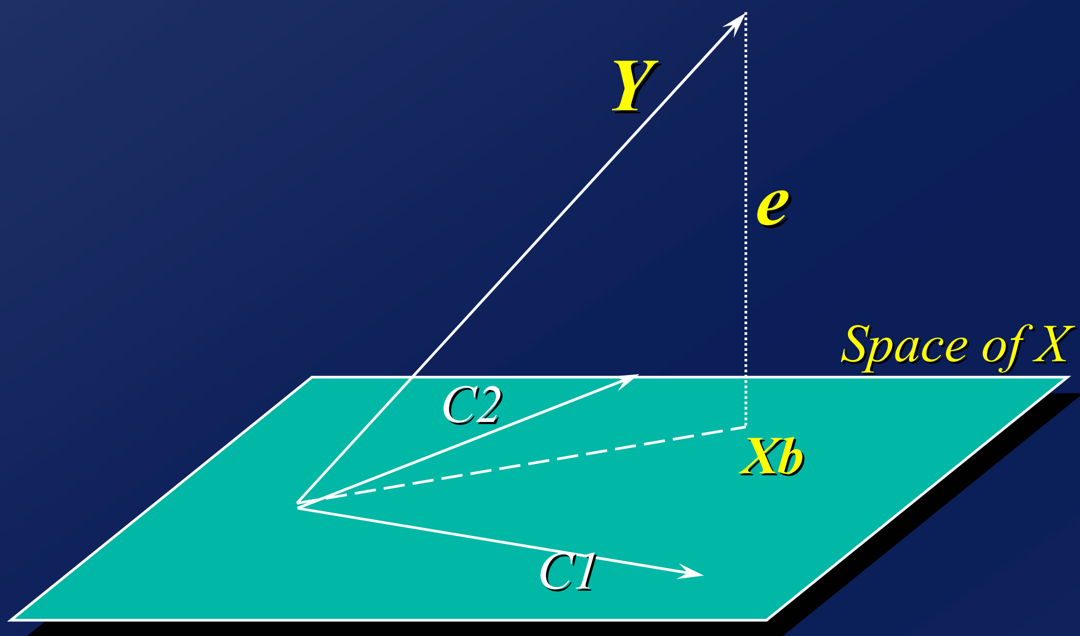
Conclusion : check your models

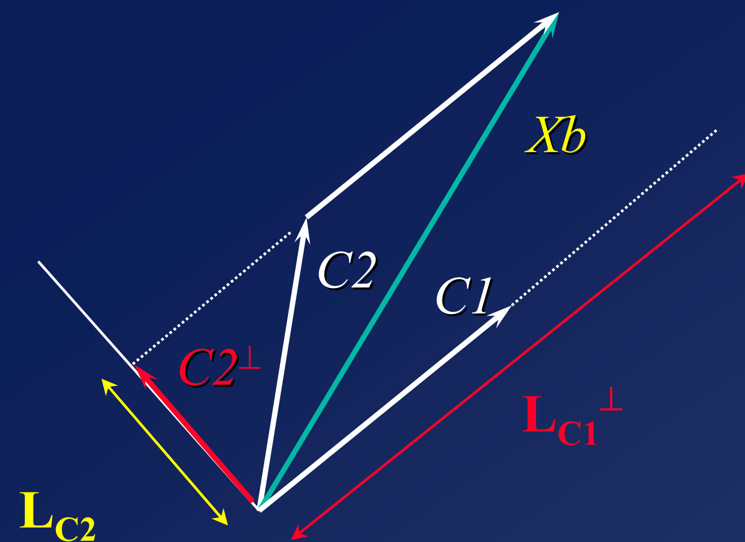
- *Check your residuals/model*
  - multivariate toolbox
- *Check your HRF form*
  - HRF toolbox
- *Check group homogeneity*
  - Distance toolbox

www.madic.org !



Implicit or explicit (⊥) decorrelation (or orthogonalisation)





This generalises when testing several regressors (F tests)

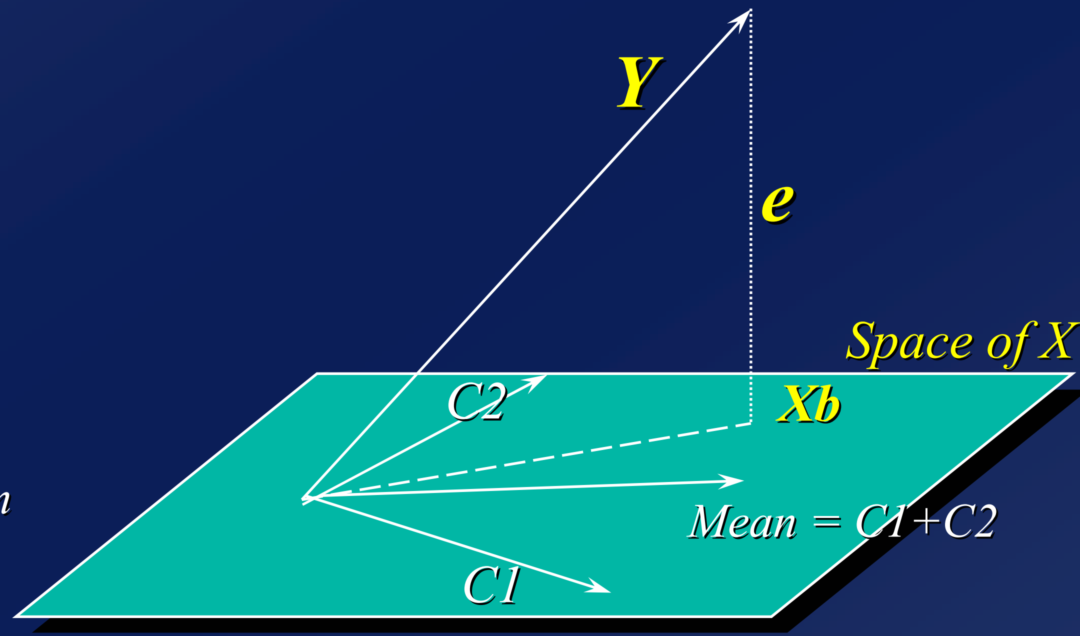
cf Andrade et al., NeuroImage, 1999

L_{C2} : test of C2 in the implicit ⊥ model

$L_{C1}{}^{\perp}$: test of C1 in the explicit ⊥ model

"completely" correlated ...

$$Y = Xb + e; \quad X = \begin{array}{ccc} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ \textit{Cond 1} & \textit{Cond 2} & \textit{Mean} \end{array}$$



Parameters are not unique in general ! Some contrasts have no meaning: NON ESTIMABLE

c = [1 0 0] is **not** estimable (no specific information in the first regressor);

c = [1 -1 0] is estimable;