General Linear Model & Classical Inference

Lyon, SPM-M/EEG course

April 2012

C. Phillips, Cyclotron Research Centre, ULg, Belgium

Overview

- Introduction
  - ERP example
- General Linear Model
  - Definition & design matrix
  - Parameter estimation & interpretation
  - Contrast & inference
  - Correlated regressors
- Conclusion

Overview

- Introduction
  - ERP example
- General Linear Model
  - Definition & design matrix
  - Parameter estimation & interpretation
  - Contrast & inference
  - Correlated regressors
- Conclusion

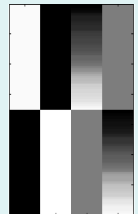
Overview of SPM

Raw EEG/MEG data

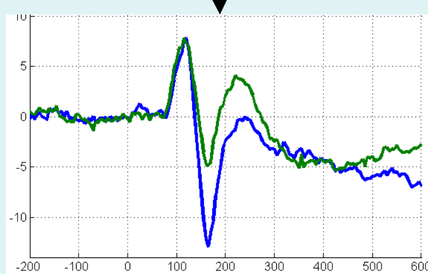
Pre-processing:
- Converting
- Filtering
- Resampling
- Re-referencing
- Epoching
- Artefact rejection
- Time-frequency transformation
- …

Image convertion

Design matrix

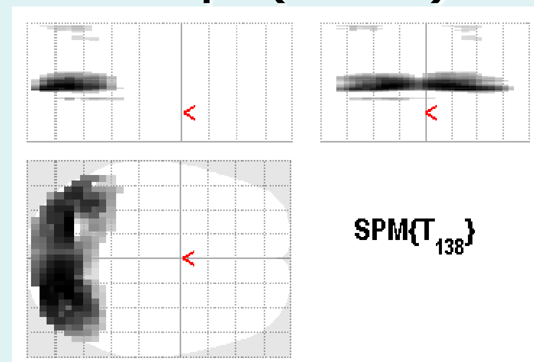


General Linear Model



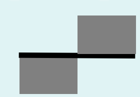
Parameter estimates

Statistical Parametric Map (SPM)



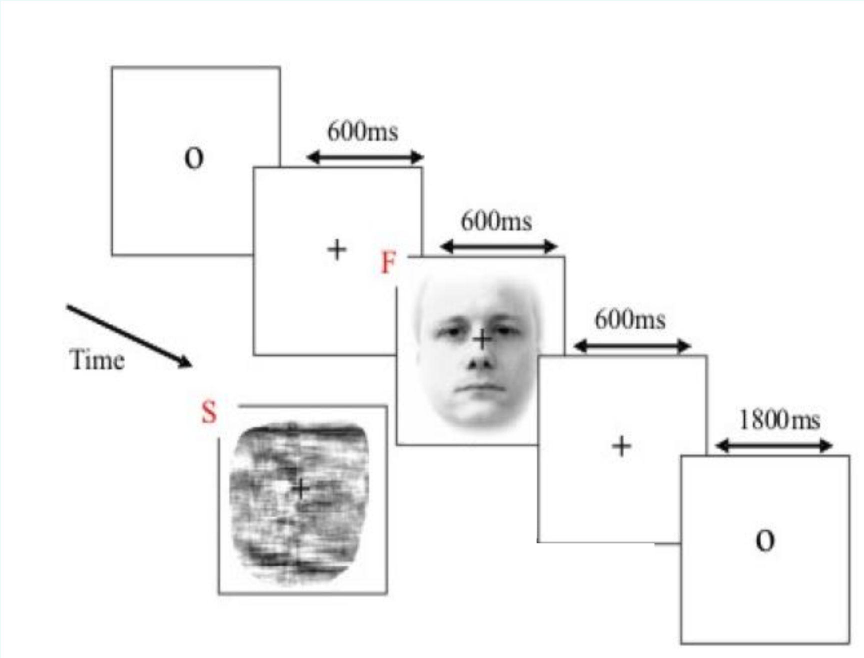
Inference & correction for multiple comparisons

Contrast:
c’ = [-1 1]



ERP example

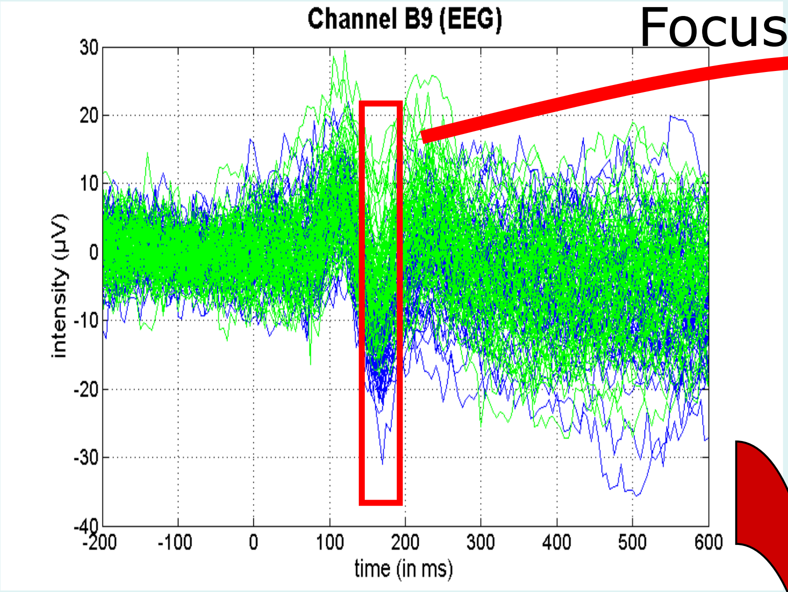
- Random presentation of ‘*faces*’ and ‘*scrambled faces*’
- 70 trials of each type
- 128 EEG channels



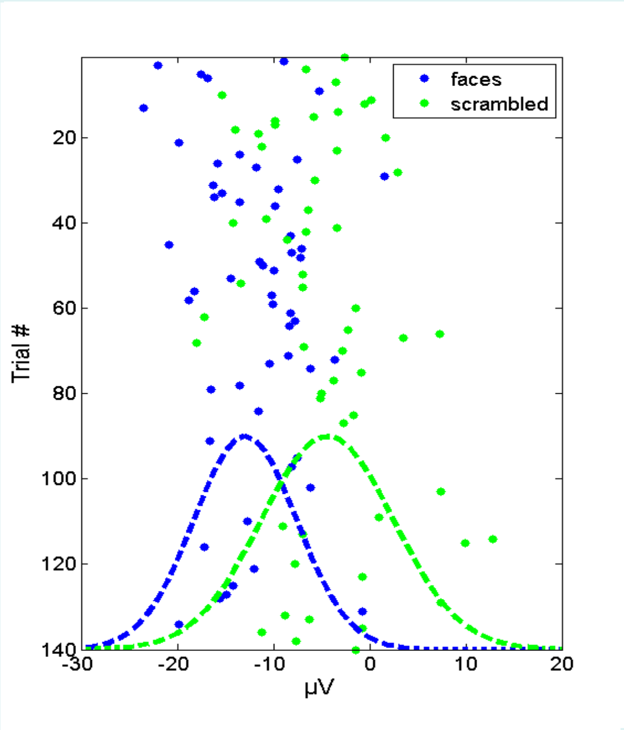
<u>Question:</u>

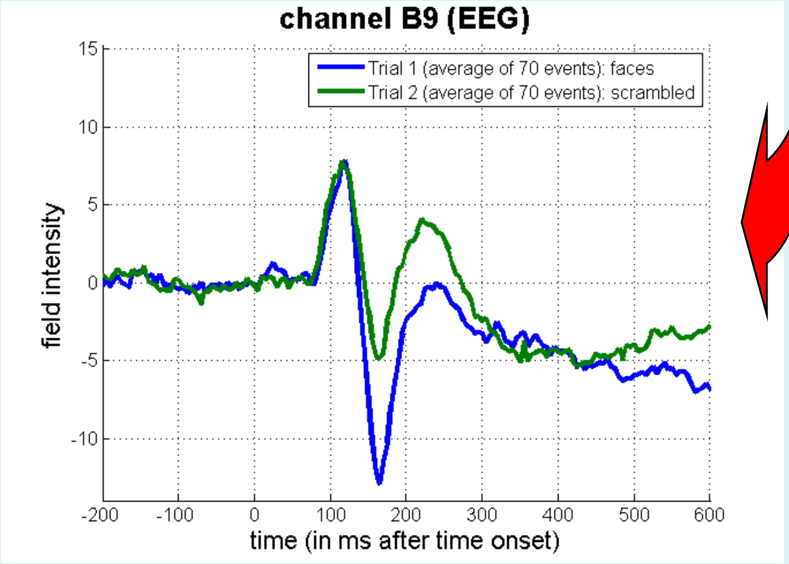
is there a difference between the ERP of ‘*faces*’ and ‘*scrambled faces*’?

ERP example: channel B9



Focus on N170





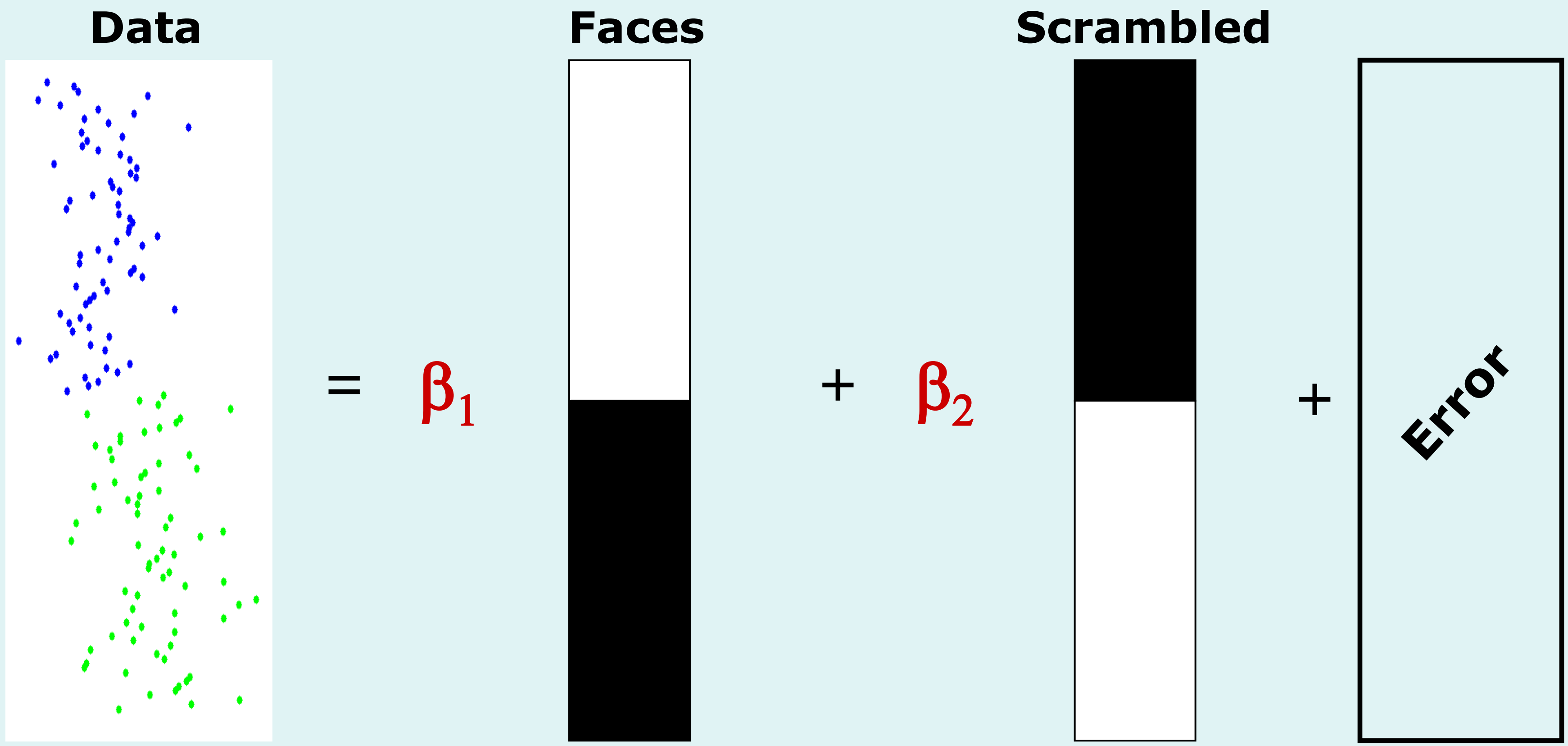
$$t = \frac{\mu_f - \mu_s}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n_f} + \frac{1}{n_s} \right)}}$$

compares size of effect to its error standard deviation

Overview

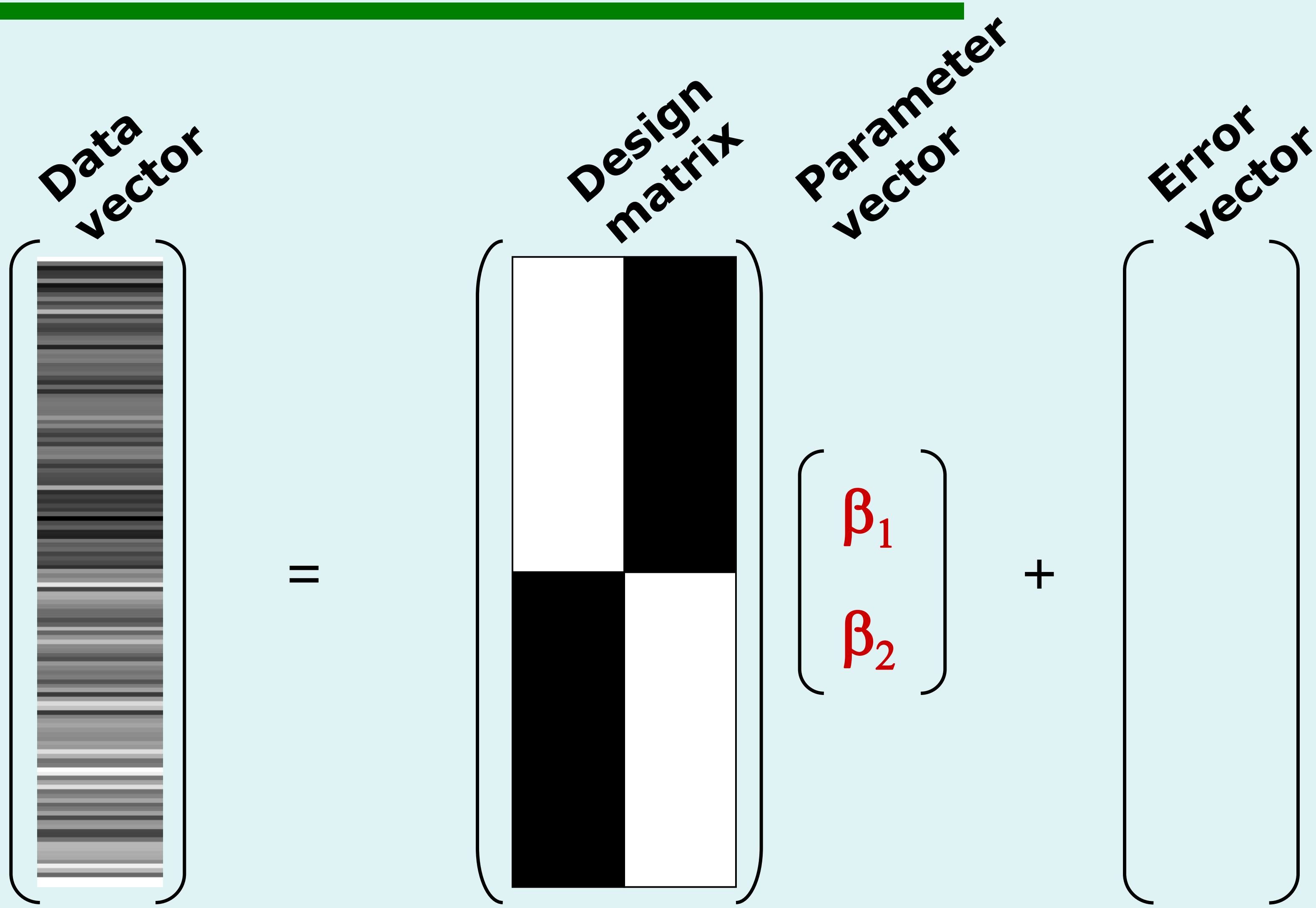
- Introduction
  - ERP example
- General Linear Model
  - Definition & design matrix
  - Parameter estimation & interpretation
  - Contrast & inference
  - Correlated regressors
- Conclusion

Data modeling



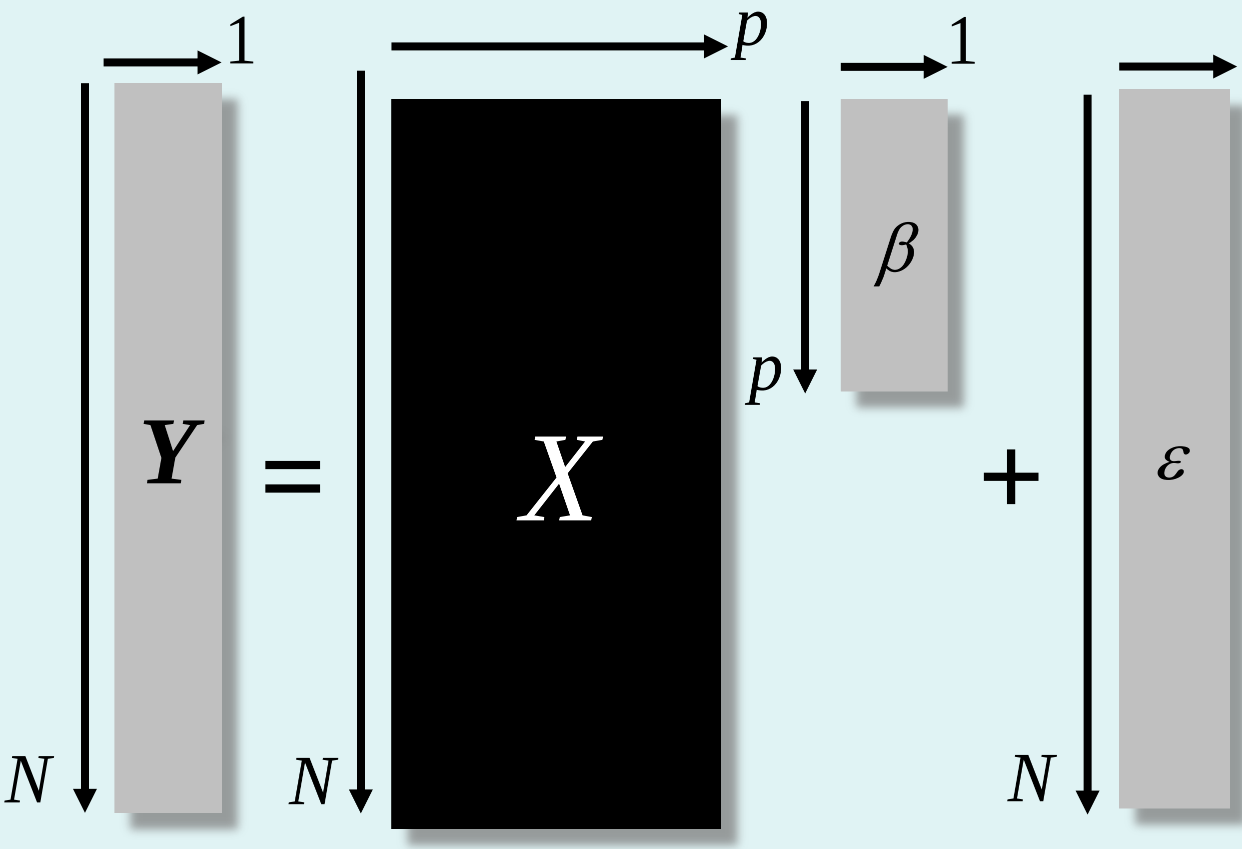
$$Y = \beta_1 \bullet X_1 + \beta_2 \bullet X_2 + \varepsilon$$

Design matrix



$$Y = X \bullet \beta + \varepsilon$$

General Linear Model



N: # trials

p: # regressors

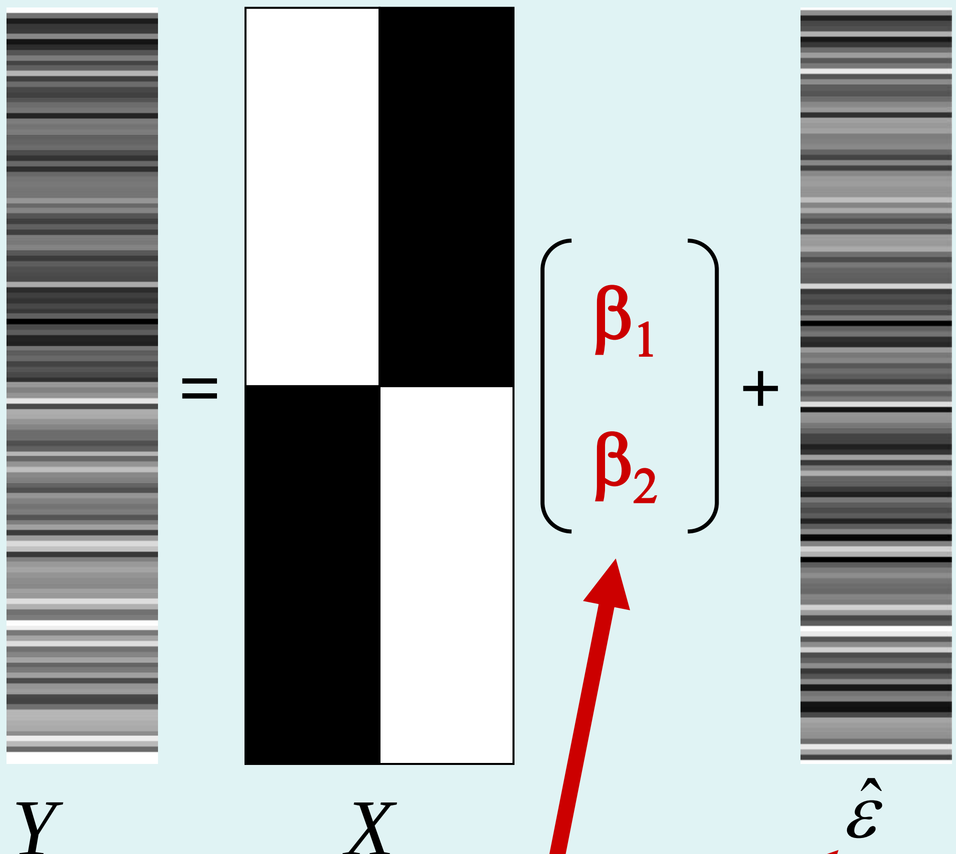
$$\boxed{Y = X\beta + \varepsilon}$$

GLM defined by $\begin{cases} \text{design matrix } X \\ \text{error distribution } \varepsilon \sim N(0, \sigma^2 I) \end{cases}$

General Linear Model

- The *design matrix* embodies all available knowledge about experimentally controlled factors and potential confounds.
- Applied to all channels & time points
- Mass-univariate parametric analysis
  - one sample *t*-test
  - two sample *t*-test
  - paired *t*-test
  - Analysis of Variance (ANOVA)
  - factorial designs
  - correlation
  - linear regression
  - multiple regression

Parameter estimation



Estimate parameters

such that $\sum_{i=1}^{N} \hat{\varepsilon}_i^2$ minimal

$$Y = X\beta + \varepsilon$$

Residuals: $\hat{\varepsilon} = Y - X\hat{\beta}$

Assume iid. error:

$$\varepsilon \sim N(0, \sigma^2 I)$$

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

Ordinary Least Squares
parameter estimate

Hypothesis Testing

The Null Hypothesis H_0

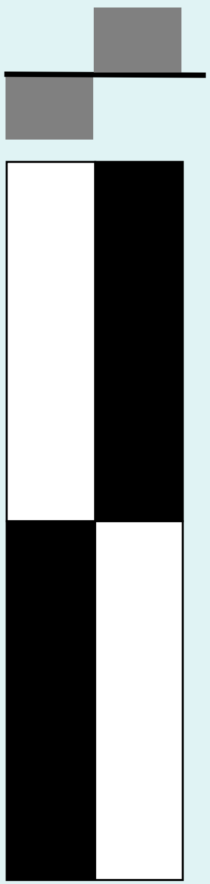
Typically what we want to disprove (i.e. *no* effect).

$\Rightarrow$ Alternative Hypothesis H_A = outcome of interest.

Contrast & *t*-test

Contrast **:** specifies linear combination of parameter vector: $c'\beta$

$c' = -1 \quad +1$



ERP: faces < scrambled ?

=

$\hat{\beta}_1 < \hat{\beta}_2$? ($\hat{\beta}_i$: **estimation of** β_i)

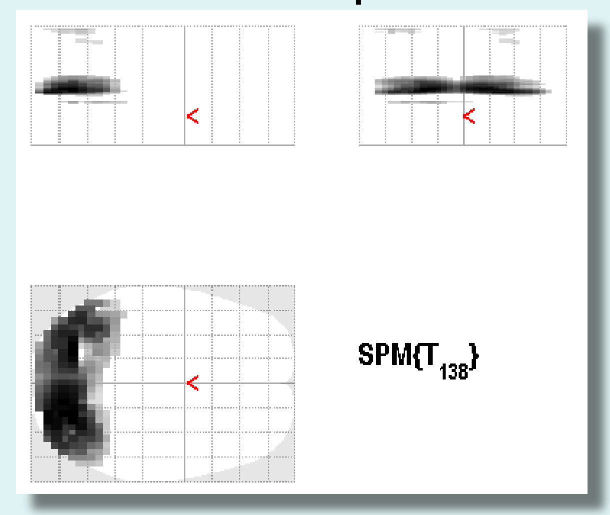
=

$-1 \times \hat{\beta}_1 + 1 \times \hat{\beta}_2 > 0$?

=

test H_0 : $c' \hat{\beta} > 0$?

SPM-t over time & space



$$T = \frac{\textit{contrast of estimated parameters}}{\sqrt{\text{variance estimate}}} \;\rightarrow\; T = \frac{c'\hat{\beta}}{\sqrt{s^2 c'(X'X)^+ c}}$$

Hypothesis Testing

The Null Hypothesis H_0

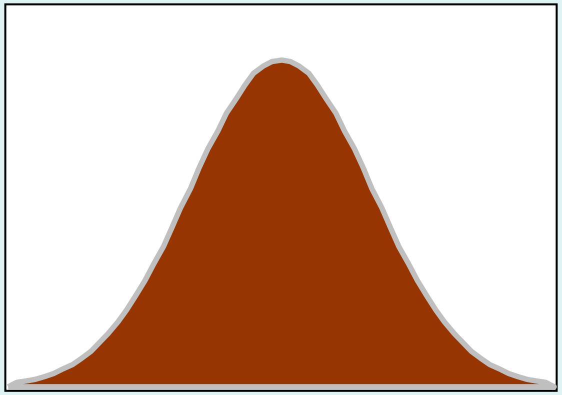
Typically what we want to disprove (i.e. *no* effect).

$\Rightarrow$ Alternative Hypothesis H_A = outcome of interest.

The Test Statistic T

- summarises evidence about H_0.
- (typically) small in magnitude when H_0 is true and large when false.

$\Rightarrow$ know the distribution of T under the null hypothesis.



Null Distribution of T

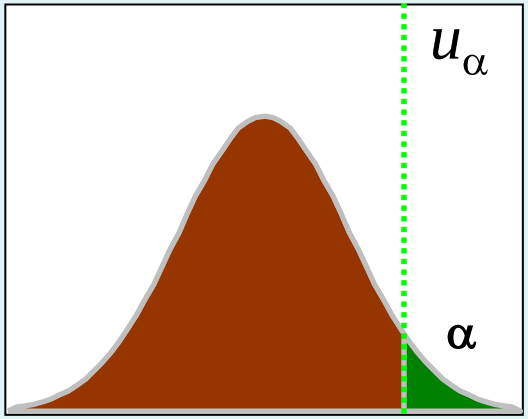
Hypothesis Testing

Significance level α:

Acceptable *false positive rate α*.

⇒ threshold u_α, controls the false positive rate

$$\alpha = p(T > u_\alpha \mid H_0)$$



Null Distribution of T

Observation of test statistic t, a realisation of T

⇒ Conclusion about the hypothesis:

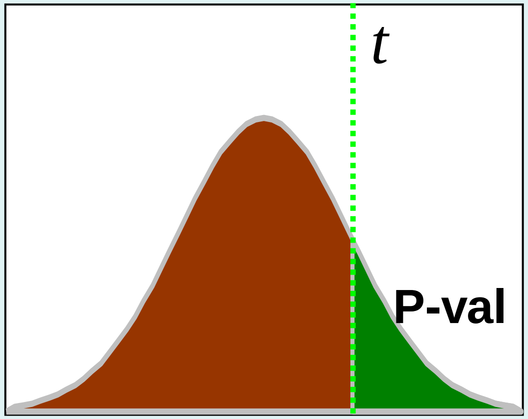
reject H_0 in favour of H_a if $t > u_\alpha$

⇒ *P*-value:

summarises evidence against H_0.

= chance of observing value more extreme than t under H_0.

$$\boxed{p(T > t \mid H_0)}$$



Null Distribution of T

Contrast & *T*-test, a few remarks

- Contrasts = simple linear combinations of the betas
- *T*-test = signal-to-noise measure (ratio of estimate to standard deviation of estimate).
- *T*-statistic, NO dependency on scaling of the regressors or contrast
- Unilateral test:

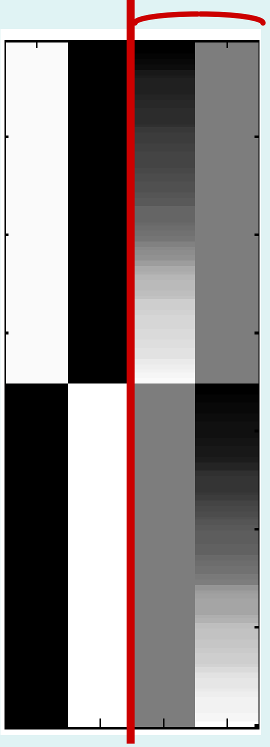
$$H_0: c^T \beta = 0 \qquad \text{vs.} \qquad H_A: c^T \beta > 0$$

Extra-sum-of-squares & *F*-test

Model comparison: *Full vs. Reduced model?*

Null Hypothesis $\mathbf{H_0}$: True model is X_0 (reduced model)

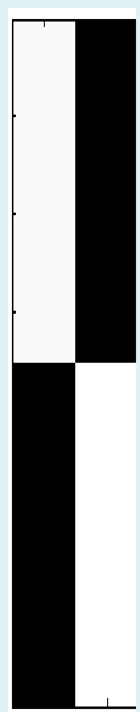
$\boldsymbol{X_0}$ $\boldsymbol{X_1}$



Full model ?

→ **RSS** $\sum \hat{\varepsilon}^2_{full}$

$\boldsymbol{X_0}$



Or reduced model?

→ $\mathbf{RSS_0}$ $\sum \hat{\varepsilon}^2_{reduced}$

Test statistic: ratio of explained and unexplained variability (error)

$$F \propto \frac{RSS_0 - RSS}{RSS}$$

$$F \propto \frac{ESS}{RSS} \sim F_{\nu_1,\nu_2}$$

$\nu_1 = \mathrm{rank}(X) - \mathrm{rank}(X_0)$
$\nu_2 = N - \mathrm{rank}(X)$

F-test & multidimensional contrasts

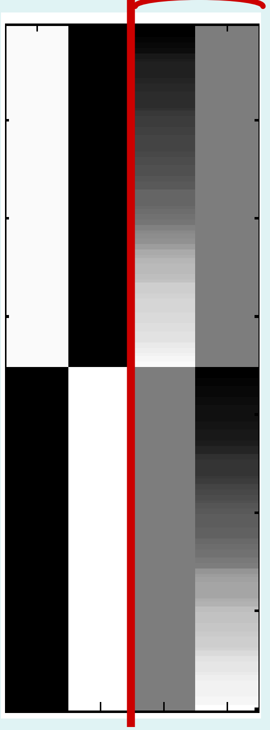
Tests multiple linear hypotheses:

H$_0$: True model is X_0

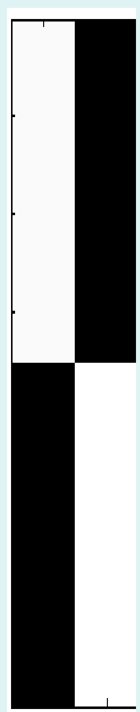
H$_0$: $\beta_3 = \beta_4 = 0$

test H$_0$: $c^T\beta = 0$?

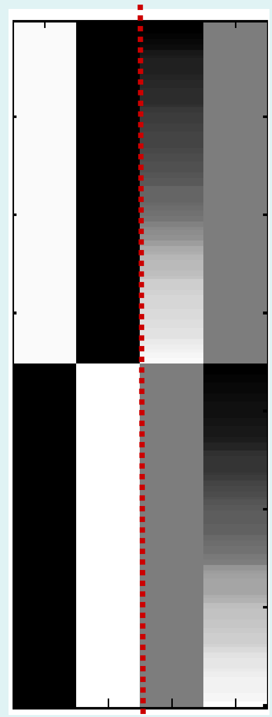
X_0 | X_1 $(\beta_{3\text{-}4})$

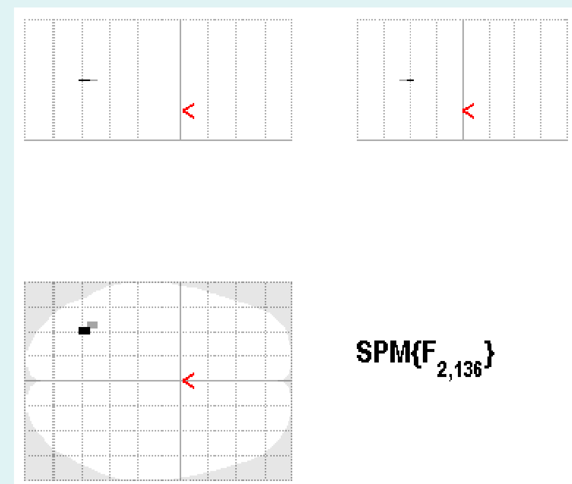


X_0



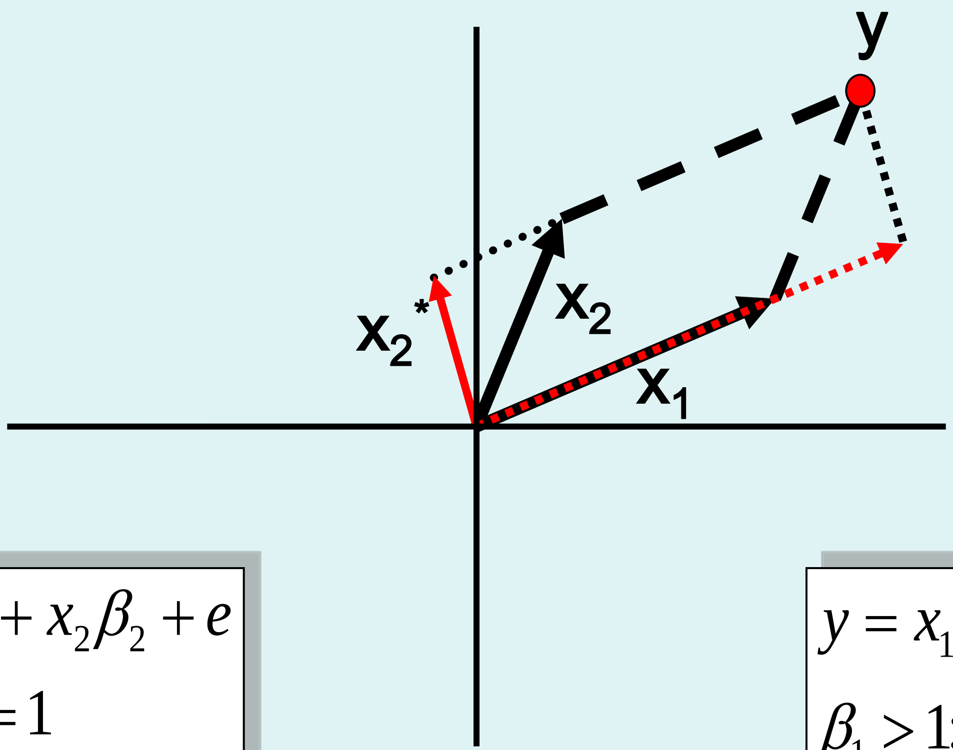
$$c^T = \begin{matrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{matrix}$$





Full or reduced model?

Correlated and orthogonal regressors



$$y = x_1\beta_1 + x_2\beta_2 + e$$
$$\beta_1 = \beta_2 = 1$$

$$y = x_1\beta_1 + x_2^*\beta_2^* + e$$
$$\beta_1 > 1; \beta_2^* = 1$$

Correlated regressors
$\Rightarrow$ explained variance shared between regressors

x_2 orthogonalized w.r.t. x_1
$\Rightarrow$ only the parameter estimate for x_1 changes, not that for x_2!

Inference & correlated regressors

- implicitly test for an *additional* effect only
  - be careful if there is correlation
  - orthogonalisation = decorrelation (not generally needed)

    $\Rightarrow$ parameters and test on the non modified regressor change
- always simpler to have orthogonal regressors and therefore designs.
- use F-tests in case of correlation, to see the overall significance. There is generally no way to decide to which regressor the « common » part should be attributed to.
- original regressors may not matter: it's the contrast you are testing which should be as decorrelated as possible from the rest of the design matrix

Overview

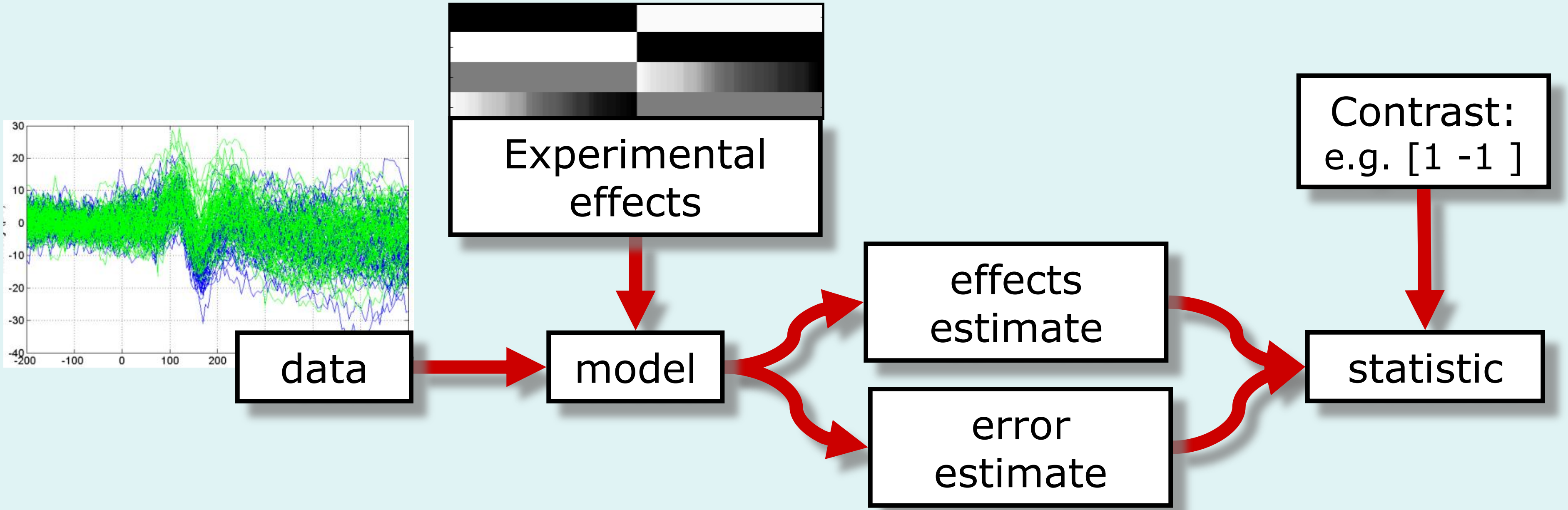
- Introduction
  - ERP example
- General Linear Model
  - Definition & design matrix
  - Parameter estimation & interpretation
  - Contrast & inference
  - Correlated regressors
- Conclusion

Modelling?

Why? Make *inferences* about effects of interest

How?
1. Decompose data into *effects* and *error*
2. Form *statistic* using estimates of effects (of interest) and error

Model? Use any available *knowledge*



Thank you for your attention!

Any question?

Thanks to Klaas, Guillaume, Rik, Will, Stefan, Andrew & Karl
for the borrowed slides!